

Ρέθυμνο 4-7-23

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΠΟΛΙΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΕΡΓΑΣΤΗΡΙΟ ΚΟΙΝΩΝΙΚΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΠΟΛΙΤΙΚΗΣ ΕΡΕΥΝΑΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

της Ζαχαρένιας Γκολέμη

ΘΕΜΑ : «Η Πρόβλεψη του Εκλογικού Αποτελέσματος στο πλαίσιο της διαδικασίας Multinomial Logistic Regression».

Επταμελής Εξεταστική Επιτροπή

1. Δαφέρμος Βασίλης, α. Καθηγητής Πανεπιστημίου Κρήτης, Επιβλέπων.
2. Παπαδάκης Νικόλαος, Καθηγητής Τμήματος Πολιτικής Επιστήμης Πανεπιστημίου Κρήτης, Μέλος Τριμελούς Επιτροπής.
3. Αλεξόπουλος Νεκτάριος, Καθηγητής Τμήματος Πολιτικής Επιστήμης Πανεπιστημίου Κρήτης, Μέλος Τριμελούς Επιτροπής.
4. Χατζηπαντελής Θεόδωρος, Καθηγητής Τμήματος Πολιτικής Επιστήμης Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης.
5. Αναστασιάδου Σοφία, Καθηγήτρια Τμήματος Προσχολικής Αγωγής Πανεπιστημίου Δυτικής Μακεδονίας.
6. Γαγάνη Βαλάντη, Καθηγητής Τμήματος Οικονομικών Πανεπιστημίου Κρήτης
7. Τσαγκαράκης Κώστας, Καθηγητής Σχολής Μηχανικών Παραγωγής και Διοίκησης Πολυτεχνείου Κρήτης.



## **Ευχαριστίες**

Ευχαριστώ βαθιά τον Επόπτη της Διατριβής Καθηγητή της Κοινωνικής Στατιστικής, κ. Βασίλη Δαφέρμο, για όλα όσα μου έμαθε, που μας έμαθε, όλα αυτά τα χρόνια στο καμίνι της εμπειρικής γνώσης και της ερευνητικής Μεθοδολογίας, στο εργαστήριο που ο ίδιος ίδρυσε και ονομάζεται Εργαστήριο Κοινωνικής Στατιστικής και Πολιτικής Έρευνας, του Τμήματος Πολιτικής Επιστήμης του Πανεπιστημίου Κρήτης. Το τελευταίο όχι άδικα, απέσπασε θεσμικούς και ερευνητικούς επαίνους κατά την εξωτερική Πιστοποίηση του Τμήματος Πολιτικής Επιστήμης το Δεκέμβριο του 2020, οι οποίοι αναφέρθηκαν με έμφαση στο Social Statistics Lab. Αντελήφθην το βαθύτερο νόημα του όρου **ΔΕΙΓΜΑΤΟΛΗΠΤΙΚΗ ΚΟΥΛΤΟΥΡΑ**. Κατάλαβα πόσο απαραίτητη είναι η απόκτηση της δεξιότητας να συλλέγεις με επιστημονικό τρόπο τα Δεδομένα της Έρευνας. Την ακεραιότητα και την ιερότητα αυτών. Αποκόμισα μεγάλο όγκο γνώσεων γύρω από τα θέματα της κλασικής Στατιστικής Ανάλυσης, των σύγχρονων Μεθόδων Στατιστικής Πρόβλεψης, των μορφών Δειγματοληψίας, των τεχνικών οικοδόμησης των θεωρητικών κατασκευών (Ερωτηματολογίων Έρευνας), των μεθόδων Προσομοίωσης (Simulation) και των Πακέτων Στατιστικού Λογισμικού (SPSS, LISREL, STATA, AMOS, EQS). Δεν θα ξεχάσω παράλληλα την ενθάρρυνση που καθημερινά μας έδινε ο εν λόγω Καθηγητής να συνεχίσουμε το όντως δύσκολο έργο μας, σε μια περιοχή με βαθιές επιστημονικές έννοιες, στην εκπόνηση των Διδακτορικών μας Διατριβών. Αισθάνομαι έτοιμη να βρω δουλειά στην αγορά εργασίας.

Ευχαριστώ επίσης θερμά, τον Καθηγητή του Τμήματος Πολιτικής Επιστήμης κ. Νικόλαο Παπαδάκη, για όσα ιδιαίτερα χρήσιμα μας έμαθε γύρω από το αντικείμενο των ποιοτικών μεθόδων. Χωρίς αυτόν δεν θα ήξερα τι είναι κρίσιμη πληροφορία. Δεν θα ξεχάσω επίσης τα λόγια του Καθηγητή Βασίλη Δαφέρμου που θεραπεύει ποσοτικές Μεθόδους: “Ευγνωμονώ τους Επιστήμονες των ποιοτικών μεθόδων. Από αυτούς εγώ παίρνω την αναγκαία θεωρία για να κατασκευάσω τα Μοντέλα των εμπειρικών μου εφαρμογών. Τα Ερωτηματολόγια της Έρευνας...! Διότι, Επιστήμη χωρίς Θεωρία, δεν υπάρχει, όπως είπε ο Umberto Eco...”.

Ευχαριστώ επίσης θερμά, τον Καθηγητή του Τμήματος Πολιτικής Επιστήμης κ.  
Νεκτάριο Αλεξόπουλο που εμπλέκει ποιοτικές και ποσοτικές Μεθόδους προκειμένου  
να μεγιστοποιήσει την αποτελεσματικότητα των επιστημονικών του ευρημάτων.  
Εμαθα και από εκείνον, πολλά.

## **Περίληψη**

Στην παρούσα διατριβή πραγματοποιήθηκε σε δείγμα χιλίων ατόμων Στρωματοποιημένη, τυχαία Δειγματοληψία τηλεφωνικών συνεντεύξεων από τον τηλεφωνικό κατάλογο του ΟΤΕ στις 13 περιφέρειες της χώρας, στο Εργαστήριο Κοινωνικής Στατιστικής και Πολιτικής Έρευνας του Πανεπιστημίου Κρήτης. Δικαίωμα συμμετοχής στις έρευνες είχαν όλοι οι πολίτες με δικαίωμα ψήφου.

Αρχικά πραγματοποιείται μια συνοπτική παρουσίαση της εκλογικής συμπεριφοράς και των μοντέλων αυτής και ακολουθεί η στατιστική ανάλυση στην οποία έχουν αναπτυχθεί εννιά μοντέλα με σκοπό τον εντοπισμό των βέλτιστων μοντέλων καθώς και η σύνθεση και παρουσίαση των συμπερασμάτων των ερευνών.

Στο μοντέλο 1 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητη τη μεταβλητή SEX (φύλο) και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με το φύλο και στη συνέχεια εφαρμόζοντας τους Δείκτες προσαρμογής του μοντέλου συμπεραίνουμε ότι δεν θα μπορούσε να γίνει λόγος για ένα καλό μοντέλο.

Στο μοντέλο 2 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, SEX (φύλο), CAT\_AGE (ηλικιακές κατηγορίες) και την αλληλεπίδραση μεταξύ αυτών και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με το φύλο, τις ηλικιακές κατηγορίες και την αλληλεπίδραση μεταξύ αυτών. Με βάση το ανωτέρω μοντέλο, καταλήγουμε στο συμπέρασμα ανυπαρξίας αλληλεπίδρασης, μεταξύ των ανεξάρτητων μεταβλητών του μοντέλου μας.

Στο μοντέλο 3 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, SEX (φύλο) και CAT\_AGE (ηλικιακές κατηγορίες) και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με το φύλο και τις ηλικιακές κατηγορίες

και στη συνέχεια εφαρμόζοντας τους Δείκτες προσαρμογής του μοντέλου, διαπιστώνουμε ότι η προσαρμογή του συγκεκριμένου μοντέλου είναι καλή.

Στο μοντέλο 4 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), CAT\_AGE (ηλικιακές κατηγορίες) και την αλληλεπίδραση μεταξύ αυτών και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με την εκπαίδευση, τις ηλικιακές κατηγορίες και την αλληλεπίδραση μεταξύ αυτών. Με βάση το ανωτέρω μοντέλο, καταλήγουμε στο συμπέρασμα ανυπαρξίας αλληλεπίδρασης, μεταξύ των ανεξάρτητων μεταβλητών του μοντέλου μας.

Στο μοντέλο 5 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), SEX (φύλο) και την αλληλεπίδραση μεταξύ αυτών και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με την εκπαίδευση, το φύλο, τις ηλικιακές κατηγορίες και την αλληλεπίδραση μεταξύ αυτών. Εφαρμόζοντας στη συνέχεια τους Δείκτες προσαρμογής του μοντέλου, καταλήγουμε στο συμπέρασμα ότι δεν μπορεί να γίνει αποδεκτό το μοντέλο μας αφενός γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής αφετέρου επειδή δεν μπορεί εύκολα να ερμηνευθεί.

Στο μοντέλο 6 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), SEX (φύλο), CAT\_AGE (ηλικιακές κατηγορίες) και την αλληλεπίδραση μεταξύ αυτών και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με την εκπαίδευση, με το φύλο, τις ηλικιακές κατηγορίες και την αλληλεπίδραση μεταξύ αυτών. Εφαρμόζοντας στη συνέχεια τους Δείκτες προσαρμογής του μοντέλου, καταλήγουμε στο συμπέρασμα ότι δεν μπορεί να γίνει αποδεκτό το μοντέλο μας αφενός γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής αφετέρου επειδή δεν μπορεί εύκολα να ερμηνευθεί.

Στο μοντέλο 7 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), SEX (φύλο), CAT\_AGE (ηλικιακές κατηγορίες) και τις ερωτήσεις Q2, Q11, Q12, Q13, Q14, Q15 και Q16 και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με τις ανεξάρτητες μεταβλητές. Εφαρμόζοντας στη συνέχεια τους Δείκτες προσαρμογής του μοντέλου, καταλήγουμε στο συμπέρασμα ότι μπορεί να γίνει αποδεκτό.

Στο μοντέλο 8 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), SEX (φύλο), AGE (Ηλικία) και τις ερωτήσεις Q2, Q11, Q12, Q13, Q14, Q15 και Q16 και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με τις ανεξάρτητες μεταβλητές. Εφαρμόζοντας στη συνέχεια τους Δείκτες προσαρμογής του μοντέλου,, καταλήγουμε στο συμπέρασμα ότι μπορεί να γίνει αποδεκτό το μοντέλο, δηλαδή είναι ερμηνεύσιμο.

Στο μοντέλο 9 ορίζουμε ως εξαρτημένη τη μεταβλητή PARTY (επιλογή κόμματος), ως ανεξάρτητες μεταβλητές, EDUC (Εκπαίδευση), SEX (Φύλο), AGE (Ηλικία), την αλληλεπίδραση μεταξύ των μεταβλητών Φύλο - Εκπαίδευση και τις ερωτήσεις Q2, Q11, Q12, Q13, Q14, Q15 και Q16 και ως κατηγορία αναφοράς, δηλαδή ως συγκριτικό μέτρο, το κόμμα SYRIZA, με πρόθεσή μας τη σύγκριση του με όλα τα κόμματα σχετικά με τις ανεξάρτητες μεταβλητές. Εφαρμόζοντας στη συνέχεια τους Δείκτες προσαρμογής του μοντέλου, καταλήγουμε στο συμπέρασμα ότι μπορεί να γίνει αποδεκτό το μοντέλο δηλαδή είναι ερμηνεύσιμο.

Συμπεραίνουμε, λοιπόν, ότι τα εκτιμόμενα ποσοστά που προέκυψαν από την στατιστική ανάλυση είναι πολύ κοντά στα ποσοστά των Κομμάτων που έλαβαν μέρος στις εκλογές της 25ης Ιανουαρίου του 2015.Επίσης, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ, σημαντικό ρόλο παίζουν οι μεταβλητές οι ιδεολογικοί λόγοι, η σωτηρία της πατρίδας, οι ανειλημμένες υποχρεώσεις, η στάση των πολιτών απέναντι στο μνημόνιο και στο σύστημα και τέλος ο Αρχηγός .



## **Abstract**

In this thesis, a Stratified Random Sampling of telephone interviews from the telephone directory of OTE in the 13 regions of the country was carried out on a sample of one thousand people, at the Laboratory of Social Statistics and Political Research of the University of Crete. All citizens with the right to vote had the right to participate in the surveys. Firstly, a brief presentation of electoral behavior and its models is carried out, and we continue with the statistical analysis, in which nine models have been developed with the aim of identifying the best models, and the synthesis and presentation of the research conclusions.

In the model 1, we define as dependent the variable PARTY (party choice), as independent variable, SEX (gender) and as reference category, i.e. as comparative measure, the SYRIZA party, with the intention to compare it with all parties regarding the gender and then by applying the Model Fit Indices we conclude that we couldn't consider it as a good model.

In model 2 we define as dependent, the variable PARTY (party choice), as independent variables, SEX (gender), CAT\_AGE (age categories) and the interaction between them and as reference category, i.e. as a comparative measure, the SYRIZA party, with our intention to compare it with all parties regarding gender, age groups and the interaction between them. Based on the above model, we conclude that there is no interaction between the independent variables of our model.

In model 3 we define as dependent the variable PARTY (party choice), as independent variables SEX (gender) and CAT\_AGE (age categories) and as reference category, i.e. as comparative measure, the SYRIZA party, with our intention to compare it with all the parties regarding gender and age categories and then applying the Model Fit Indices, we find that the fit of the particular model is good.

In model 4 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), CAT\_AGE (age categories) and the interaction between

them and as a reference category, i.e., as a comparative measure, the SYRIZA party, with our intention to compare it with all parties regarding education, age groups and the interaction between them. Based on the above model, we conclude that there is no interaction between the independent variables of our model.

In model 5 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), SEX (gender) and the interaction between them and as a reference category, i.e., as a comparative measure, the SYRIZA party, with our intention to compare it with all parties regarding education, gender, age groups and the interaction between them. Applying then the Model Fit Indices, we conclude that our model cannot be accepted on the one hand because we have no guarantee of a good fit and on the other hand because it cannot be easily interpreted.

In model 6 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), SEX (gender), CAT\_AGE (age categories) and the interaction between them and as a reference category, i.e. as a comparative measure, the party SYRIZA, with our intention to compare it with all parties regarding education, gender, age groups and the interaction between them. Applying then the Model Fit Indices, we conclude that our model cannot be accepted on the one hand because we have no guarantee of a good fit and on the other hand because it cannot be easily interpreted.

In model 7 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), SEX (sex), CAT\_AGE (age categories) and questions Q2, Q11, Q12, Q13, Q14, Q15 and Q16 and as a reference category, i.e. as a comparative measure, the SYRIZA party, with our intention to compare it with all parties regarding the independent variables. Applying then the Model Fit Indices, we conclude that it can be accepted.

In model 8 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), SEX (Sex), AGE (Age) and questions Q2, Q11, Q12, Q13, Q14, Q15 and Q16 and as reference category, i.e. as a comparative measure, the SYRIZA

party, with our intention to compare it with all parties regarding the independent variables. Applying then the Model Fit Indices, we conclude that the model can be accepted, i.e. it is interpretable.

In model 9 we define as dependent the variable PARTY (party choice), as independent variables, EDUC (Education), SEX (Gender), AGE (Age), the interaction between variables Gender - Education and questions Q2, Q11, Q12, Q13, Q14, Q15 and Q16 and as a reference category, i.e. as a comparative measure, the SYRIZA party, with our intention to compare it with all parties regarding the independent variables. Applying then the model fit indices, we come to the conclusion that the model can be accepted, i.e., it is interpretable.

We conclude that the estimated by the statistical analysis percentages are very close to the percentages of the Parties that took part in the elections of the 25th January 2015. Also, in the contest ND - SYRIZA, the variables, the ideological reasons, the salvation of the country, the undertaken obligations, the behavior of the citizens concerning the memorandum and the system and finally the Chief.



## **Περιεχόμενα**

Περιεχόμενα .....	ix
Κατάλογος Πινάκων .....	xiii
Κατάλογος Διαγραμμάτων.....	xvii
Κεφάλαιο 1: Σκοπός, Στόχοι και Μεθοδολογία της Έρευνας .....	1
1.1. Ο Σκοπός της Έρευνας.....	1
1.2. Οι Στόχοι της Έρευνας.....	1
1.3. Μεθοδολογία .....	1
1.3.1. Η Ταυτότητα της Έρευνας.....	1
1.3.2. Συμβολισμός, Κωδικοποίηση και Επεξήγηση των μεταβλητών που υπεισέρχονται στην Έρευνά μας.....	2
1.4. Η οικοδόμηση των MNML μοντέλων. Επειδή η αλήθεια δεν είναι μία, και επομένως θα οικοδομήσουμε αρκετά μοντέλα, (για την ακρίβεια 9 προβλεπτικά Μοντέλα) για να την προσεγγίσουμε (MODELLING).....	4
Κεφάλαιο 2: Εκλογική Συμπεριφορά.....	9
2.1. Εισαγωγή .....	9
2.2. Θεωρητικό πλαίσιο.....	11
2.3. Παράγοντες καταλυτικοί στην επιλογή ενός συγκεκριμένου κομματικού φορέα από τους ψηφοφόρους.....	13
2.4. Μοντέλα Ανάλυσης Εκλογικής Συμπεριφοράς .....	16
2.4.1. Εκλογική γεωγραφία .....	17
2.4.2. Ποσοτική οικολογία και η ανάλυση συνθηκών.....	18
2.4.3. Τα Διπολικά Σχήματα Ανάλυσης Της Εκλογικής Συμπεριφοράς.....	18
2.4.4. Η Έννοια Της Κομματικής Ταύτισης Της Σχολής Του MICHIGAN.....	19
2.4.5. Το Μοντέλο Της Ορθολογικής Επιλογής Και η Θεματική Ψήφος.....	20
2.5. Διαιρετικές Τομές-Εκλογική Συμπεριφορά.....	23

Κεφάλαιο 3: Δυαδικά Λογαρίθμογραμμικά Παλίνδρομα Μοντέλα (Binomial Logistic Regression, BLM), ο προάγγελος των Πολυωνυμικών Λογαρίθμογραμμικών Παλίνδρομων Μοντέλων.....	25
3.1. Προκαταρτικά.....	25
3.2. Η κατασκευή του Μοντέλου.....	26
3.3. Η Ερμηνεία του μοντέλου πρόβλεψης.....	28
3.4. Γενικευμένες εκφράσεις για το BLR .....	32
Κεφάλαιο 4: Το Πολυωνυμικό Λογαρίθμογραμμικό Μοντέλο (Multinomial Logit Model, MNLM).....	33
4.1. Σύντομη Ιστορική αναδρομή .....	33
4.2. Η θεμελίωση του πολυωνυμικού μοντέλου (Multinomial Logit Model, MNLM) .....	34
4.3. Η προσαρμογή του μοντέλου MNLM στα δεδομένα.....	37
4.4. Το φαινόμενο της υπερδιασκορπισης (overdispersion) .....	39
4.5. Τα Πρότυπα Συνδιασποράς (Covariate Patterns).....	41
4.6. Οι παραδοχές για το MNLM μοντέλο.....	42
4.7. Η έννοια της αλληλεπίδρασης στο πλαίσιο των MNLM μοντέλων.....	46
4.8. Οι στατιστικοί δείκτες AIC και BIC.....	46
Κεφάλαιο 5: Η Στατιστική Ανάλυση των Δεδομένων της Έρευνας. Ανάπτυξη των Στατιστικών Μοντέλων Πρόβλεψης.....	47
5.1. Συμβολισμός, Κωδικοποίηση και Επεξήγηση των μεταβλητών που υπεισέρχονται στην έρευνα μας.....	47
5.2. Το αρχείο (Data Set) της Έρευνας. Η συσχέτιση της μεταβλητής έκβασης PARTY με τη μεταβλητή CAT_AGE των ηλικιακών κατηγοριών.....	48
5.3. Η κατασκευή του απλούστερου μοντέλου MODEL 1.....	49
5.3.1. Τα αποτελέσματα της εκτέλεσης του MONTELOY 1 .....	54
5.3.2. Δείκτες προσαρμογής για το MODEL1 .....	58
5.3.3. Μέτρα υευδο- $\chi^2$ (pseudo-R-square Measures) .....	59
5.3.4. Ο Πίνακας Ταξινόμησης (Classification Table) για το MODEL1 .....	60

5.3.5. Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 1 .....	62
5.3.6. Η προσαρμογή (goodness-of-fit) του Model 1 .....	63
5.3.7. Τελικά συμπεράσματα για το Model 1 .....	64
5.4. Η κατασκευή του Model 2.....	64
5.5. Η κατασκευή του Model 3.....	70
5.6. Η κατασκευή του Model 4.....	79
5.7. Η κατασκευή του Model 5.....	83
5.8. Η κατασκευή του Model 6.....	85
5.8.1. Multinomial Logistic Regression Analysis: Το βαθύτερο νόημα της Αλληλεπίδρασης (Interaction) .....	89
5.9. Η κατασκευή του Model 7.....	92
5.9.1. Το φαινόμενο της υπερδιασκόρπισης (overdispersion).....	100
5.9.2. Τι είναι τα covariate patterns, ποια είναι η σχέση τους με τους βαθμούς ελευθερίας, τα ανεξάρτητα Logits και το συνολικό αριθμό των παραμέτρων που θέλουμε να εκτιμήσουμε, συμπεριλαμβανομένου και του σταθερού όρου.....	102
5.9.3. Τα Likelihood Ratio Tests .....	103
5.9.4. Η Προβλεπτική Δύναμη (Prediction Power) του MONTEΛΟΥ 7. Ο Πίνακας Ταξινόμησης (Classification Table). .....	104
5.9.5. Η Ερμηνεία των παραμέτρων για το Model 7.....	106
5.9.6. Παρατηρούμενες και Προβλεπόμενες Συχνότητες για το Model 7.....	114
5.9.7. Η δομική περιγραφή του Μηχανισμού Πρόβλεψης .....	115
5.9.8. Είναι το Model 7 το βέλτιστο μοντέλο; Which is the Best Model? .....	116
5.10. Η κατασκευή του Model 8.....	119
5.10.1. Αποτελέσματα για το Model 8. Προσαρμογή, Σπουδαιότητα των Μεταβλητών. Likelihood Ratio tests. Ερμηνεία του Μοντέλου 8. ....	123
5.10.2. Περιθώριες Συχνότητες (Marginal Effects) για το Model 8.....	130
5.10.3. Η Προβλεπτική Δύναμη (Prediction Power) για το Model 8. Ο Πίνακας Ταξινόμησης (Classification Table). .....	131
5.10.4. Είναι το Model 8 το βέλτιστο μοντέλο; Which is the Best Model? .....	132

5.11. Η κατασκευή του Model 9.....	133
5.11.1. Αποτελέσματα για το Model 9. Περιθώριες Συχνότητες - Πρόβλεψη. Προσαρμογή, Σπουδαιότητα των Μεταβλητών. Likelihood Ratio tests. Ερμηνεία του Model 9.....	136
5.11.2. Παρατηρήσεις για το Model 9.....	145
Κεφάλαιο 6: Συμπεράσματα .....	149
Βιβλιογραφία .....	167

## **Κατάλογος Πινάκων**

Πίνακας 1.1. Δημογραφικά Στοιχεία.....	2
Πίνακας 1.2. Λόγοι Επιλογής Κόμματος στις Εκλογές.....	2
Πίνακας 1.3. Σημαντικότερος Λόγος Επιλογής Κόμματος στις Εκλογές. ....	3
Πίνακας 1.4. Επιλογή Κόμματος στις Εκλογές.....	3
Πίνακας 1.5. Μοντέλο 1.....	6
Πίνακας 1.6. Μοντέλο 2.....	6
Πίνακας 1.7. Μοντέλο 3.....	7
Πίνακας 1.8. Μοντέλο 4.....	7
Πίνακας 1.9. Μοντέλο 5.....	7
Πίνακας 1.10. Μοντέλο 6.....	7
Πίνακας 1.11. Μοντέλο 7.....	7
Πίνακας 1.12. Μοντέλο 8.....	8
Πίνακας 1.13. Μοντέλο 9.....	8
Πίνακας 5.1. Ερωτηματολόγιο της Έρευνας. ....	47
Πίνακας 5.2. Συσχέτιση της PARTY με τη CAT_AGE.....	48
Πίνακας 5.3. Έλεγχος Συσχέτισης της PARTY με τη CAT_AGE.....	48
Πίνακας 5.4. Συσχέτιση της PARTY με τη SEX.....	54
Πίνακας 5.5. Εκτιμήσεις των παραμέτρων του Model 1.....	55
Πίνακας 5.6. Σύγκριση Model 1 με Null μοντέλο.....	58
Πίνακας 5.7. Ψεύδο $X^2$ του Model 1.....	60
Πίνακας 5.8. Πίνακας Ταξινόμησης (Classification Table) για το Model 1. ....	60
Πίνακας 5.9. Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 1, ανά φύλο.....	62
Πίνακας 5.10. Προσαρμογή (goodness-of-fit) του Model 1.....	63
Πίνακας 5.11. Σύγκριση Model 2 με Null μοντέλο.....	69
Πίνακας 5.12. Model 3 - Εκτίμηση Αποτελεσμάτων Εθνικών Εκλογών Ιανουαρίου 2015.....	71

Πίνακας 5.13. Σύγκριση Model 3 με Null μοντέλο.....	71
Πίνακας 5.14. Προσαρμογή (goodness-of-fit) του Model 3.....	72
Πίνακας 5.15. Σύγκριση των Reduced Model 3 με Null μοντέλο. ....	72
Πίνακας 5.16. Ερμηνεία των παραμέτρων του Model 3. ....	75
Πίνακας 5.17. Πίνακας Ταξινόμησης (Classification Table) για το Model 3.....	76
Πίνακας 5.18. Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 3.....	78
Πίνακας 5.19. Σύγκριση Model 4 με Null μοντέλο.....	81
Πίνακας 5.20. Προσαρμογή (goodness-of-fit) του Model 4.....	82
Πίνακας 5.21. Σύγκριση των Reduced Model 4 με Null μοντέλο.....	82
Πίνακας 5.22. Σύγκριση Model 5 με Null μοντέλο.....	84
Πίνακας 5.23. Προσαρμογή (goodness-of-fit) του Model 5.....	84
Πίνακας 5.24. Σύγκριση των Reduced Model 5 με Null μοντέλο.....	84
Πίνακας 5.25. Σύγκριση Model 6 με Null μοντέλο.....	87
Πίνακας 5.26. Συσχέτιση μεταβλητών EDUC και SEX.....	89
Πίνακας 5.29. Περιθώριων Συχνοτήτων, Model 7. ....	98
Πίνακας 5.30. Σύγκριση Model 7 με Null μοντέλο.....	98
Πίνακας 5.31. Προσαρμογή (goodness-of-fit) του Model 7.....	98
Πίνακας 5.32. Σύγκριση των Reduced Model 7 με Null μοντέλο.....	103
Πίνακας 5.33. Πίνακας Ταξινόμησης (Classification Table) για το Model 7.....	104
Πίνακας 5.34. Εκτιμήσεις των παραμέτρων του Model 7, για τη Ν.Δ.....	111
Πίνακας 5.35. Εκτιμήσεις των παραμέτρων του Model 7, για το ΠΑ.ΣΟ.Κ.....	112
Πίνακας 5.36. Εκτιμήσεις των παραμέτρων του Model 7, για το Κ.Κ.Ε.....	112
Πίνακας 5.37. Εκτιμήσεις των παραμέτρων του Model 7, για το Χ.Α.....	113
Πίνακας 5.38. Εκτιμήσεις των παραμέτρων του Model 7, για το ΑΝ.Ε.Λ.....	113
Πίνακας 5.39. Παρατηρούμενες και Προβλεπόμενες Συχνότητες, Model 7.....	114
Πίνακας 5.40. Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 7.....	118

Πίνακας 5.41. Πίνακας $\chi^2$ , Model 8.....	123
Πίνακας 5.32. Σύγκριση των Reduced Model 8 με Null μοντέλο.....	124
Πίνακας 5.43. Εκτιμήσεις των παραμέτρων του Model 8.....	125
Πίνακας 5.44. Περιθώριων Συχνοτήτων, Model 8. ....	130
Πίνακας 5.45. Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 8. ....	131
Πίνακας 5.46. Πίνακας Ταξινόμησης (Classification Table) για το Model 8. ....	131
Πίνακας 5.47. Περιθώριων Συχνοτήτων, Model 9. ....	137
Πίνακας 5.48. Πίνακας $\chi^2$ , Model 9.....	137
Πίνακας 5.49. Σύγκριση των Reduced Model 9 με Null μοντέλο.....	139
Πίνακας 5.50. Εκτιμήσεις των παραμέτρων του Model 9.....	140
Πίνακας 5.51. Πίνακας Ταξινόμησης (Classification Table) για το Model 9. ....	144
Πίνακας 5.52. Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 9. ....	145



## **Κατάλογος Εικόνων**

Εικόνα 5.1. Βήμα 2ο για την εκτέλεση Multinomial Logistic Regression.....	50
Εικόνα 5.2. Βήμα 3ο για την εκτέλεση Multinomial Logistic Regression.....	50
Εικόνα 5.3. Βήμα 4ο για την εκτέλεση Multinomial Logistic Regression.....	51
Εικόνα 5.4. Βήμα 5ο για την εκτέλεση Multinomial Logistic Regression.....	51
Εικόνα 5.5. Βήμα 6ο για την εκτέλεση Multinomial Logistic Regression.....	52
Εικόνα 5.6. Βήμα 7ο για την εκτέλεση Multinomial Logistic Regression.....	53
Εικόνα 5.7. Βήμα 7ο για την εκτέλεση Multinomial Logistic Regression.....	53
Εικόνα 5.8. Προσδιορισμός Μεταβλητών Model 2.....	65
Εικόνα 5.9. Προσδιορισμός Model 2 .....	66
Εικόνα 5.10. Επιλογές μενού Statistics, Model 2.. .....	67
Εικόνα 5.11. Επιλογές μενού Criteria, Model 2. ....	67
Εικόνα 5.12. Επιλογές μενού Options, Model 2.....	68
Εικόνα 5.13. Προσδιορισμός Model 3.....	70
Εικόνα 5.14. Προσδιορισμός Μεταβλητών Model 4.....	80
Εικόνα 5.15. Προσδιορισμός Model 4.....	80
Εικόνα 5.16. Επιλογές μενού Criteria, Model 4. ....	80
Εικόνα 5.17. Επιλογές μενού Options, Model 4.....	81
Εικόνα 5.18. Προσδιορισμός Μεταβλητών Model 5.....	83
Εικόνα 5.19. Επιλογές μενού Options, Model 5.....	84
Εικόνα 5.20. Προσδιορισμός Μεταβλητών Model 6.....	86
Εικόνα 5.21. Προσδιορισμός Model 6.....	86
Εικόνα 5.22. Επιλογές μενού Options, Model 6.....	87
Εικόνα 5.23. Προσδιορισμός Μεταβλητών Model 7.....	92
Εικόνα 5.24. Επιλογή Κατηγορίας Αναφοράς Model 7. ....	93
Εικόνα 5.25. Προσδιορισμός Model 7.....	93
Εικόνα 5.26. Επιλογές μενού Statistics Model 7.....	94
Εικόνα 5.27. Επιλογές μενού Criteria, Model 7. ....	95

Εικόνα 5.28. Επιλογές μενού Options, Model 7.....	95
Εικόνα 5.29. Προειδοποιητικό μήνυμα, Model 7.....	96
Εικόνα 5.30. Προσδιορισμός Μεταβλητών Model 8.....	119
Εικόνα 5.31. Επιλογή Κατηγορίας Αναφοράς Model 8. ....	120
Εικόνα 5.32. Προσδιορισμός Model 8. ....	121
Εικόνα 5.33. Επιλογές μενού Statistics, Model 8.. .....	121
Εικόνα 5.34. Επιλογές μενού Criteria, Model 8. ....	122
Εικόνα 5.35. Επιλογές μενού Options, Model 8.....	122
Εικόνα 5.36. Προσδιορισμός Μεταβλητών Model 9.....	134
Εικόνα 5.37. Προσδιορισμός Model 9. ....	135
Εικόνα 5.38. Επιλογές μενού Statistics, Model 9.. .....	135

## **Κεφάλαιο 1: Σκοπός, Στόχοι και Μεθοδολογία της Έρευνας**

### **1.1. Ο Σκοπός της Έρευνας**

Βασικός σκοπός της Έρευνας είναι η πρόβλεψη του εκλογικού αποτελέσματος, μέσω της αναζήτησης του Βέλτιστου Προβλεπτικού Μοντέλου (Best Predictive Model), από τα διαθέσιμα.

### **1.2. Οι Στόχοι της Έρευνας**

Στόχοι της Έρευνας είναι να προσδιοριστούν οι επιδράσεις των παρακάτω παραγόντων στη λειτουργία των προβλεπτικών μοντέλων:

- i. Οικονομία (ανεργία, οικονομική ανέχεια, ακρίβεια κ.τ.λ.)
- ii. Η σωτηρία της πατρίδας
- iii. Οι προσωπικές υποχρεώσεις
- iv. Η στάση του πολίτη απέναντι στα μνημόνια
- v. Η στάση των πολίτη απέναντι στο ‘Σύστημα’
- vi. Η επιλογή του Αρχηγού του Κόμματος
- vii. Η επιλογή των υποψηφίων του Κόμματος
- viii. Η γενική κοσμοθεωρία/ιδεολογία του Κόμματος
- ix. Το Πρόγραμμα και οι Θέσεις του Κόμματος
- x. Το μορφωτικό επίπεδο του ψηφοφόρου
- xi. Το φύλο του ψηφοφόρου
- xii. Η ηλικία του ψηφοφόρου

### **1.3. Μεθοδολογία**

#### **1.3.1. Η Ταυτότητα της Έρευνας**

**Πληθυσμός της Έρευνας:** Όλοι οι Έλληνες, ηλικίας από 17 ετών και άνω.

**Μέθοδος Δειγματοληψίας:** Στρωματοποιημένη (με στρώμα την Περιφέρεια), τυχαία Δειγματοληψία τηλεφωνικών συνεντεύξεων (από σταθερό ή κινητό, αδιακρίτως, με τυχαία προσπέλαση) από τον τηλεφωνικό κατάλογο του ΟΤΕ, στις 13 περιφέρειες της χώρας.

**Δειγματολήπτες:** Τη Δειγματοληψία διεξήγαγαν 40 φοιτητές του Τμήματος Πολιτικής Επιστήμης του Πανεπιστημίου Κρήτης, στο Εργαστήριο Κοινωνικής Στατιστικής και Πολιτικής Έρευνας του Π.Κ.

**Μέγεθος Δείγματος:** 1917 άτομα.

### 1.3.2. Συμβολισμός, Κωδικοποίηση και Επεξήγηση των μεταβλητών που υπεισέρχονται στην Έρευνά μας

Στους Πίνακες 1.1 έως 1.4 παραθέτουμε τον Συμβολισμό, την Κωδικοποίηση και την Επεξήγηση των μεταβλητών που υπεισέρχονται στην Έρευνά μας.

#### Πίνακας 1.1. Δημογραφικά Στοιχεία.

<b>FYLO:</b>
1 = Άνδρας, 2 = Γυναίκα
<b>CAT AGE<sup>1</sup>:</b>
1 = 18-34 ετών, 2 = 35-54 ετών, 3 = 55 και άνω

#### Πίνακας 1.2. Λόγοι Επιλογής Κόμματος στις Εκλογές.

Το πολιτικό Κόμμα που θα ψηφίσετε σε αυτές τις εκλογές, όποιο κι αν είναι αυτό, το ψηφίζετε για:

Q11: Ιδεολογικούς λόγους;	1 = NAI	2 = OXI
Q12: Οικονομικούς λόγους (ανεργία / οικονομική ανέχεια κτλ);	1 = NAI	2 = OXI
Q13: Να σωθεί η Πατρίδα;	1 = NAI	2 = OXI
Q14: Λόγους προσωπικών υποχρεώσεων;	1 = NAI	2 = OXI
Q15: Γιατί είναι αντιμνημονιακό Κόμμα;	1 = NAI	2 = OXI
Q16: Από αντίδραση στο Σύστημα;	1 = NAI	2 = OXI
Q17: Για άλλο λόγο ποιό; .....	1 = NAI	2 = OXI

<sup>1</sup> Οικοδομήσαμε, σκόπιμα, την κατηγορική μεταβλητή των ηλικιακών ομάδων, CAT\_AGE, και μάλιστα με 3 κατηγορίες: 1=17-34 ετών, 2=35-54 ετών και 3=55 και άνω, καθώς ήταν στόχος μας να εξετάσουμε τη διαφορετική, όπως επιτάσσει η θεωρία μας συμπεριφορά αυτών των Ηλικιακών Ομάδων. Πράγματι, όπως έδειξε η μετέπειτα ανάλυση, η μεταβλητή CAT\_AGE αναδείχθηκε στατιστικώς σημαντική, και άμεσο αποτέλεσμα αυτού το γεγονότος υπήρξε η διαφορετική συμπεριφορά των διάφορων ηλικιακών κατηγοριών.

### **Πίνακας 1.3. Σημαντικότερος Λόγος Επιλογής Κόμματος στις Εκλογές.**

**Q2.** Τι είναι ποιο σημαντικό στην απόφασή σας για το Κόμμα αυτό;  
(μια απάντηση παρακαλώ):

1 = Ο αρχηγός του κόμματος;

2 = Οι υποψήφιοι;

3 = Η γενική κοσμοθεωρία/ιδεολογία του κόμματος;

4 = Το πρόγραμμα και οι θέσεις;

5 = Άλλο (ποιο;... .... )

### **Πίνακας 1.4. Επιλογή Κόμματος στις Εκλογές.**

**PARTY:** Ποιο Κόμμα είναι πιο πιθανό να ψηφίσετε την επόμενη Κυριακή;

1 = ΝΔ

2 = ΣΥΡΙΖΑ

3 = ΠΑΣΟΚ

4 = ΚΙΔΗΣΟ του Γ.Παπανδρέου

5 = ΠΟΤΑΜΙ

6 = KKE

7 = ΧΡΥΣΗ ΑΥΓΗ

8 = ΑΝΕΞ. ΕΛΛΗΝΕΣ

9 = ΠΡΑΣΙΝΟΙ-ΔΗΜΑΡ

10 = KKE (μ-λ) και Μ-Λ KKE

11 = ΤΕΛΕΙΑ του Απ. Γκλέτσου

12 = ΑΝΤΑΡΣΙΑ

13 = ΛΑΟΣ

14 = ΆΛΛΟ ποιο; .....

15 = ΑΔΙΕΥΚΡΙΝΙΣΤΗ ΨΗΦΟΣ = ΛΕΥΚΟ/ΑΚΥΡΟ/ΚΑΝΕΝΑ/ΔΕΝ ΘΑ ΠΑΩ ΝΑ  
ΨΗΦΙΣΩ

### **Σημείωση**

Κατά τη Στατιστική Ανάλυση, που θα δούμε παρακάτω (βλ. Κεφ.5), τα Κόμματα με κωδικούς 4 = ΚΙΔΗΣΟ του Γ. Παπανδρέου, 9 = ΠΡΑΣΙΝΟΙ-ΔΗΜΑΡ, 10 = KKE (μ-λ) και Μ-Λ KKE, 11 = ΤΕΛΕΙΑ του Απ. Γκλέτσου, 12 = ΑΝΤΑΡΣΙΑ και 13 = ΛΑΟΣ, ενσωματώθηκαν στο κωδικό 14 = ΆΛΛΟ.

**1.4. Η οικοδόμηση των MNML μοντέλων. Επειδή η αλήθεια<sup>2</sup> δεν είναι μία, και επομένως θα οικοδομήσουμε αρκετά<sup>3</sup> μοντέλα, (για την ακρίβεια 9 προβλεπτικά Μοντέλα) για να την προσεγγίσουμε (MODELLING).**

Η παράγραφος αυτή αφιερώνεται σε μια κορυφαία διαδικασία που φέρει το όνομα MODELLING και αφορά την οικοδόμηση μιας αλληλουχίας Μοντέλων Πρόβλεψης με απότερο σκοπό το Βέλτιστο (The Best Model) αυτών. Βασικά χαρακτηριστικά αυτού του Μοντέλου είναι:

- I. Η Καλή Προσαρμογή στα δεδομένα (Goodness- of- it).
- II. Η ισχυρή Προβλεπτική Δύναμη (Prediction Power).
- III. Η φειδωλότητα. Η οικοδόμηση του μοντέλου θα πρέπει να γίνει με τον ελάχιστο, αλλά απαραίτητο αριθμό, μεταβλητών.
- IV. Η Ερμηνευσιμότητα.
- V. Η Φροντίδα για αποκλεισμό και πιθανότατα αποβολή από το μοντέλο, μεταβλητών που παρουσιάζουν υψηλή συσχέτιση (Collinearity).

---

<sup>2</sup> Η Αλήθεια, όπως έλεγε ο Gaston Bachelard, είναι πολεμική. Θεάσεις, μορφές, όψεις της αλήθειας υπάρχουν, και κανένας δεν είναι ο απόλυτος κάτοχός της. Γι' αυτό κι εμείς, στο πλαίσιο αυτού του κεφαλαίου, πολλά μοντέλα θα κτίσουμε, και πολλούς αγώνες θα κάνουμε, για να τη φτάσουμε, αν τη φτάσουμε.... Υπάρχει και η άποψη του Νομπελίστα Feynman για τη μη κατοχή της αλήθειας από κανέναν: «Δεν υπάρχει πλέον καμιά αυθεντία που να αποφασίζει ποια ιδέα είναι καλή. Δεν είμαστε πιά αναγκασμένοι να απευθυνόμαστε σε αυθεντίες για να μάθουμε κατά πόσο μια ιδέα είναι αληθινή ή όχι. Μπορούμε να διαβάσουμε το έργο της αυθεντίας και να δούμε εκεί τι προτείνει. Τη σχετική πρόταση μπορούμε να την υποβάλουμε σε έλεγχο και να διαπιστώσουμε αν είναι αληθινή ή όχι. Κι αν δεν είναι αληθινή τόσο το χειρότερο. Έτσι, οι 'αυθεντίες', χάνουν κάτι από το κύρος τους....».

<sup>3</sup> Πόσα μοντέλα θα μπορούσαν να οικοδομήσουν με τα δεδομένα που διαθέτουμε, ποιες στρατηγικές θα μπορούσαμε να ακολουθήσουμε για να προσεγγίσουμε το άριστο (the best) μοντέλο, τι προβλέπει η Θεωρία ανάπτυξης των Ιεραρχικών (Hierarchical) Μοντέλων, ποια μοντέλα έχουν το χαρακτηριστικό της ερμηνευσιμότητας, τι ακριβώς περιλαμβάνει ένα τέλειο/πλήρες μοντέλο (saturated model), όλα αυτά είναι άκρως ενδιαφέροντα ερευνητικά ζητήματα. Ωστόσο, ξεφεύγουν από τα στενά όρια της παρούσας Ερευνας. Η διεθνής βιβλιογραφία πάνω στα θέματα αυτά είναι ιδιαίτερα εκτενής [Burnham, Anderson, (2002), Lahiri, P. (2002), Menard, S. (2010), Collet, D. (2003), Vittinghoff, E., Glidden, D. Shboski, S., McCulloch, C. (2005), Vittinghoff, E., Glidden, D. Shboski, S., McCulloch, C. (2012), SPSS ADVANCED MODELS 10.0 (1999), Kuhn, M., Johnson, K. (2013), Agresti, A. (2013), Agresti, A. (2015), Agresti, A. (2018), Wood, S. (1917), Powers, D., Xie, Y. (1991)].

VI. Η Αποφυγή μη συμπεριληψης στο Μοντέλο, μεταβλητών κεφαλαιώδους σημασίας για τη συγκεκριμένη Έρευνα. Θα θέλαμε να γνωρίζουμε αν στις Εθνικές εκλογές του 2015, οι άνδρες ψήφισαν διαφορετικά από τις γυναίκες. Επίσης θέλουμε να πληροφορηθούμε αν είναι διαφορετική η εκλογική συμπεριφορά των διάφορων ηλικιακών ομάδων. Επομένως, η μεταβλητή του φύλου, όπως και η μεταβλητή των ηλικιακών κατηγοριών δεν θα μπορούσαν να απουσιάζουν από τα τελικά μοντέλα ως κύριες επιδράσεις<sup>4</sup> (main effects).

VII. Η χαμηλή συνθετότητα. Εάν εμείς συμπεριλάβουμε στο μοντέλο, αποκλειστικά και μόνον τις κύριες επιδράσεις (main effects), και ταυτόχρονα αποκλείσουμε κάθε αλληλεπίδραση (interaction) μεταξύ των μεταβλητών, μπορεί να εκτεθούμε στον κίνδυνο να παραβλέψουμε ιδιαίτερα σημαντικές επιδράσεις και συνακόλουθα το μοντέλο μας μπορεί να μην έχει καλή προσαρμογή. Από τη άλλη πάλι, αν εμείς συμπεριλάβουμε στο μοντέλο μας κάθε πιθανή κύρια επίδραση, μα και κάθε πιθανή αλληλεπίδραση, οπωσδήποτε θα έχουμε ένα υψηλής συνθετότητας μοντέλο (complicated model), το οποίο θα έχει μέγα αριθμό παραμέτρων οι οποίες θα είναι δύσκολο να εκτιμηθούν, αλλά και να ερμηνευθούν. Υπέρμετρος αριθμός κενών κελιών (cells frequencies) στον Πίνακα Διασταύρωσης (Crostabulation Table). Ανάμεσα σε αυτά τα δύο άκρα κάποτε, γίνεται ιδιαίτερα δύσκολη η ισορροπία<sup>5</sup>. Προφανώς θα πρέπει να αναζητήσουμε τη χρυσή τομή, ενθυμούμενοι την παροιμιώδη φράση του Albert

<sup>4</sup> Θα μπορούσε κάποιος να διερωτηθεί: Μα είναι δυνατόν να συμβεί αυτό; Να μην υπάρχουν στο μοντέλο οι μεταβλητές του φύλου και των ηλικιακών κατηγοριών ως κύριες επιδράσεις (main effects), ενώ σαφώς έχουν προσδιοριστεί στο μοντέλο ως τέτοιες; Ναι, είναι η απάντηση. Εάν μια οποιαδήποτε μεταβλητή έχει εμπλακεί εντός του μοντέλου, σε οποιασδήποτε τάξεως interaction effect, τότε δεν είναι δυνατόν να ελεγχθεί η κύρια επίδρασή της. Συνακόλουθα, στην περίπτωση αυτή η Ερμηνευσιμότητα του μοντέλου, έχει δεχθεί καίρια πλήγματα.

<sup>5</sup> Υπάρχει και μια άλλη στρατηγική, ένας γνωστός ερευνητικός δρόμος για να φτάσουμε στο βέλτιστο μοντέλο, και ενθαρρύνεται από τη διεθνή SPSS Βιβλιογραφία (SPSS Advanced Statistics, 7.5, 1997, SPSS Advanced Models, 9.0, 1999, Norusis). Είναι η εκτέλεση μιας σειράς «φωλιασμένων μοντέλων» («nested models»), επαναληπτική διαδικασία που έχει στον πυρήνα της τη μέθοδο Backward και που θα μας βοηθήσει να εντοπίσουμε/ αναγνωρίσουμε το κατάλληλο ύψος συνθετότητας του μοντέλου μας.

Einstein : “Τα πράγματα θα πρέπει να γίνουν τόσο απλά, όσο είναι δυνατόν.  
Αλλά όχι απλούστερα”.

VIII. Το μέγεθος του Δείγματος. Τα Μοντέλα πολυωνυμικής, λογαριθμογραμμικής προσαρμογής (Multinomial Logit Models, MNML), επειδή ακριβώς χρησιμοποιούν εκτιμητές μεγίστης πιθανοφάνειας (ML estimators), όπως αναφέρει ο Long (1997), υπόκεινται στους παρακάτω περιορισμούς αναφορικά με το μέγεθος του Δείγματος: (α) Ένα δείγμα τάξης μεγέθους μικρότερο του 100 είναι παρακινδυνευμένο, (β) Ένα δείγμα τάξης μεγέθους μεγαλύτερου του 500 φαίνεται να είναι επαρκές, και (γ) Τουλάχιστον 10 παρατηρήσεις ανά παράγοντα.

Στην πράξη εμείς οικοδομήσαμε και στη συνέχεια αναπτύξαμε, 9 πιθανά (possible) λογαριθμογραμμικά μοντέλα, έτσι ώστε να συμπεριλάβουμε ελέγχους για αλληλεπιδράσεις 2ας και 3ης τάξης μεταξύ των ανεξάρτητων μεταβλητών, να ελέγχουμε τη συμπεριφορά των covariates, δηλ. των ανεξάρτητων πλην όμως συνεχούς φύσης μεταβλητών, όπως για παράδειγμα της ηλικίας (AGE), και τέλος να εξεταστεί αν και κατά πόσον αλλάζει η Ερμηνευσιμότητα ενός μοντέλου, αν μέσα σε αυτή υπάρχει στατιστικώς σημαντική αλληλεπίδραση (interaction). Με βάση αυτούς τους στόχους, τα μοντέλα που θα αναπτύξουμε είναι τα εξής:

#### Πίνακας 1.5. Μοντέλο 1

MODEL 1.

Dependent Var: **PARTY**

Independent Vars: **SEX**

#### Πίνακας 1.6. Μοντέλο 2

MODEL 2.

Dependent Var: **PARTY**

Independent Vars(3): **SEX, CAT\_AGE, SEX \* CAT\_AGE**

### **Πίνακας 1.7. Μοντέλο 3**

MODEL 3.
Dependent Var: <b>PARTY</b>
Independent Vars(2): <b>SEX, CAT_AGE</b>

### **Πίνακας 1.8. Μοντέλο 4**

MODEL 4.
Dependent Var: <b>PARTY</b>
Independent Vars (3): <b>EDUC , CAT_AGE , EDUC*CAT_AGE</b>

### **Πίνακας 1.9. Μοντέλο 5**

MODEL 5.
Dependent Var: <b>PARTY</b>
Independent Variables (2): <b>SEX, EDUC, SEX* EDUC</b>

### **Πίνακας 1.10. Μοντέλο 6**

MODEL 6.
Dependent Var: <b>PARTY</b>
Independent Vars(4): <b>SEX, EDUC, CAT_AGE , SEX*EDUC*CAT_AGE</b>

### **Πίνακας 1.11. Μοντέλο 7**

MODEL 7.
Dependent Var: <b>PARTY</b>
Independent Vars (10): <b>SEX, CAT_AGE, EDUC, Q2, Q11, Q12, Q13, Q14, Q15 και Q16.</b>

Παρατήρηση για τα μοντέλα 2, 3, 4, 6 και 7:

Σε αυτά τα μοντέλα δεν υπεισέρχεται η συνεχή μεταβλητή της Ηλικίας (AGE). Για θεωρητικούς/μεθοδολογικούς λόγους, τη διασπάσαμε<sup>6</sup> σε 3 επίπεδα, όπως είδαμε

<sup>6</sup> Η διάσπαση/κατηγοριοποίηση της συνεχούς μεταβλητής AGE σε 3 επίπεδα, έγινε με γνωστή διαδικασία recode into different variable (βλ. βιβλίο ‘Δαφέρμος Β. (2011). Κοινωνική Στατιστική και Μεθοδολογία Έρευνας με το SPSS’. Εκδόσεις ΖΗΤΗ, Θεσσαλονίκη).

παραπάνω. Για να είμαστε πιο σαφείς, θεωρήσαμε, και αποδεικνύεται παρακάτω, ότι αυτές οι 3 ηλικιακές κατηγορίες έχουν διαφορετική εκλογική συμπεριφορά. Πράγματι, θα δούμε ότι εάν εκτελεστεί Crosstabulation μεταξύ των ονομαστικών μεταβλητών PARTY και CAT\_AGE, το  $\chi^2$  που θα προκύψει, θα αναδείξει στατιστικώς σημαντική σχέση (βλ. Πίνακες 4.2 και 4.3).

Παρόλα αυτά που προαναφέρθηκαν για συνεχή μεταβλητή της Ηλικίας (AGE), εμείς θα οικοδομήσουμε ένα 8<sup>o</sup> μοντέλο (MODEL 8), στο οποίο θα ξεδιπλωθεί με τρόπο καθαρό η συμπεριφορά της μεταβλητής της ηλικίας, ως συνεχούς μεταβλητής. Η δομή αυτού του μοντέλου θα είναι η εξής:

**Πίνακας 1.12.** Μοντέλο 8

MODEL 8.
Dependent Var: <b>PARTY</b>
Independent Factors Vars (10): <b>SEX, EDUC, Q2, Q11, Q12, Q13, Q14, Q15 και Q16</b>
Independent Covariates Vars: <b>AGE</b>

Τέλος, χάριν πληρότητας της έρευνας, και όπως ξαναείπαμε παραπάνω, θα οικοδομήσουμε ένα τελευταίο λογαριθμογραμμικό μοντέλο (MODEL 9), για να εξεταστεί αν και κατά πόσον αλλάζει η Ερμηνευσιμότητα ενός μοντέλου, όταν μέσα σε αυτό υπάρχει στατιστικώς σημαντική αλληλεπίδραση (Sig.ificant interaction). Η δομή του τελευταίου αυτού μοντέλου είναι η εξής:

**Πίνακας 1.13.** Μοντέλο 9

MODEL 9.
Dependent Var: <b>PARTY</b>
Independent Factors Vars (10): <b>SEX, EDUC, Q2, Q11, Q12, Q13, Q14, Q15 και Q16</b>
Independent Covariates Vars: <b>AGE</b>

## **Κεφάλαιο 2: Εκλογική Συμπεριφορά**

### **2.1. Εισαγωγή**

Στα πλαίσια ενός Δημοκρατικού πολιτεύματος, η εξουσία πηγάζει από το λαό, ασκείται από το λαό και υπηρετεί τα συμφέροντα του, όπως ακριβώς ορίζει και το άρθρο 1 του Συντάγματος. Το κύριο χαρακτηριστικό της Δημοκρατίας έγκειται στο γεγονός, ότι η λήψη των αποφάσεων πραγματοποιείται από τους πολίτες μέσω της ψηφοφορίας, δηλαδή μέσω της εκλογικής διαδικασίας. Οι εκλογείς, αφενός προσέρχονται στην κάλπη ισότιμα, προκειμένου να αναδείξουν και να νομιμοποιήσουν τους φορείς της πολιτικής εξουσίας και αφετέρου συμμετέχουν στα πολιτικά δρώμενα, με σκοπό να εκφράσουν τις ιδεολογικές και πολιτικές τους πεποιθήσεις στη διάρκεια της συγκεκριμένης εκλογικής αναμέτρησης. Οι εκλογές χαρακτηρίζονται ως η θεμέλια λίθος του δημοκρατικού καθεστώτος, καθώς οι πολίτες καλούνται να αξιολογήσουν, να εκτιμήσουν και να αποφασίσουν όχι μόνο ποιος κομματικός φορέας είναι κατάλληλος να αναλάβει τα ηνία της διακυβέρνησης της χώρας, αλλά και να εκλέξουν αντιπροσώπους, οι οποίοι θα συντάξουν το νομοθετικό σώμα, το οποίο θα εκπροσωπεί το εκλογικό σώμα στα πλαίσια μιας έμμεσης κοινοβουλευτικής Δημοκρατίας.

Για να εξετάσουμε όμως την εκλογική συμπεριφορά, δηλαδή ποια είναι τα κριτήρια με τα οποία επιλέγουμε να δώσουμε τη ψήφο μας σε ένα κόμμα, θα πρέπει αρχικά να ορίσουμε την έννοια της πολιτικής συμπεριφοράς, δηλαδή τον τρόπο με τον οποίο το άτομο εκφράζει, δραστηριοποιείται και συμμετέχει ως πολίτης σε συλλογικές διαδικασίες, σε σωματεία, συλλόγους, εθελοντικές οργανώσεις, λέσχες, κόμματα και εκλογικές διαδικασίες. Κατά το πλαίσιο μιας συγκεκριμένης κοινωνίας τα άτομα γίνονται δέκτες, αλλά και φορείς της πολιτικής κοινωνικοποίησης, δηλαδή συμμετέχουν σε μια συνεχή διαδικασία διαμόρφωσης και αλληλεπίδρασης πολιτικών αξιών, στάσεων, αναπαραστάσεων και συμπεριφορών που μεταβιβάζονται σ' αυτά μέσω των θεσμών κοινωνικοποίησης. Τα άτομα επεξεργάζονται την κουλτούρα αυτή ανάλογα με τα ατομικά και τα κοινωνικά χαρακτηριστικά τους και διαμορφώνουν μια

συγκεκριμένη πολιτική στάση και κατ' επέκταση μια πολιτική συμπεριφορά, δηλαδή τη συμμετοχή ή όχι στα πολιτικά πράγματα. Η διαμόρφωση της πολιτικής και εκλογικής συμπεριφοράς εξαρτάται από τους Κοινωνικούς παράγοντες δηλαδή τα δημοκρατικά και πολυκομματικά συστήματα διακυβέρνησης που ενθαρρύνουν τη συμμετοχή των πολιτών στις εκλογικές διαδικασίες καθώς επίσης και από τους Ατομικούς παράγοντες δηλαδή το φύλο, η ηλικία, το μορφωτικό επίπεδο, το επάγγελμα, η εργασία, η κοινωνικό-οικονομική θέση, οι φιλοσοφικές και θρησκευτικές πεποιθήσεις, τα οποία επιδρούν στην εκλογική και πολιτική συμπεριφορά του ατόμου, ανάλογα με τα βιώματα, τις εμπειρίες και την πολιτική του κοινωνικοποίηση. Ένας άλλος παράγοντας που καθορίζει και επηρεάζει την εκλογική συμμετοχή είναι ο βαθμός ένταξης και ταύτισης με την κοινωνική ομάδα από την οποία προέρχεται ο κάθε ψηφοφόρος.

Στις εκλογές λοιπόν η ψήφος δεν προκύπτει μόνο από την πολυπλοκότητα των κοινωνικών παραγόντων, αλλά είναι αποτέλεσμα και της ατομικής ψυχολογίας και των πολιτικών αντιλήψεων από τις οποίες διακατέχεται ο κάθε ψηφοφόρος. Οι πολιτικές, κοινωνιολογικές, οικονομικές συγκυρίες μεταβάλλονται από περιοχή σε περιοχή και αυτό έχει σαν αποτέλεσμα οι εκλογείς να μην ψηφίζουν με βάση την ταξική τους προέλευση και την κοινωνική τους θέση. Υπάρχουν όμως και οι σταθερά πολιτικοποιημένοι ψηφοφόροι οι οποίοι διαμορφώνουν χρονικά έγκαιρα την ψήφο τους που αναπαράγεται πιστά από την μία εκλογική αναμέτρηση στην άλλη και επιδεικνύουν ενδιαφέρον για την προεκλογική εκστρατεία. Από την άλλη μεριά όμως υπάρχουν και οι αναποφάσιστοι εκλογείς, οι οποίοι είναι δέκτες διάφορων επιρροών και πληροφοριών, όπου βρίσκονται σε καταστάσεις ιδεολογικής σύγχυσης και πιέσεων από διάφορες κατευθύνσεις.

Η σταθερή προσήλωση του ψηφοφόρου σε ένα πολιτικό κόμμα, που περιλαμβάνει την αποδοχή της ιδεολογίας, των αξιών του κόμματος, του πολιτικού του προγράμματος, του ιστορικού του φορτίου και των προσώπων που το συναπαρτίζουν αυξάνεται ανάλογα με τον βαθμό ενδιαφέροντος που εκφράζει για την πολιτική. Από την άλλη, το ποσοστό των μη κομματικά ταυτισμένων, των λεγόμενων «ανεξάρτητων», δηλαδή

εκείνων που ασχολούνται ή ενημερώνονται λιγότερο παρατηρείται ότι σταδιακά αυξάνεται.

Παράλληλα, γίνεται πλέον αποδεκτό ότι η ψήφος δεν αποτελεί μία προσωπική και ορθολογική επιλογή του κάθε εκλογέα αλλά θα μπορούσε κανείς να ισχυριστεί ότι συγκροτείται και διαμορφώνεται από την ύπαρξη ενός κοινωνικού ντετερμινισμού: «Ψηφίζω πολιτικά, ό,τι είμαι κοινωνικά».

## 2.2. Θεωρητικό πλαίσιο

Η μελέτη της εκλογικής συμπεριφοράς των ψηφοφόρων μας δίνει τη δυνατότητα να μελετήσουμε τους παράγοντες που συμβάλλουν στον καθορισμό της λήψη της πολιτικής απόφασης από ένα μεμονωμένο ψηφοφόρο στα πλαίσια μιας εκλογικής αναμέτρησης. Αφού ανατρέξουμε στις απαρχές και στη γέννηση των εμπειρικών αναλύσεων και των ερμηνευτικών μοντέλων της εκλογικής συμπεριφοράς θα καταλήξουμε να συμφωνήσουμε με την άποψη της Mayer (2005), η οποία ισχυρίζεται ότι η μελέτη της σύγχρονης εκλογικής συμπεριφοράς είναι κατά κανόνα μια πολυσύνθετη διαδικασία, η οποία απαιτεί ουσιαστικά μια σύνθεση όλων των ιστορικών θεωρητικών τάσεων της εκλογικής κοινωνιολογίας, λαμβάνοντας βεβαίως υπόψη τις κοινωνικές και πολιτικές ιδιαιτερότητες και παραδόσεις, καθώς και την ιστορική συγκυρία κάθε χώρας. Η ιστορία της εκλογικής συμπεριφοράς δείχνει ότι τα διαφορετικά θεωρητικά - επιστημονικά ρεύματα εμφανίζονταν και αναπτύσσονταν κάθε φορά που οι προηγούμενες κοινωνικές, ιδεολογικές, ταξικές και πολιτισμικές συνθήκες μεταλλάσσονταν και τα «παλαιά» εργαλεία δεν ήταν ικανά να εξηγήσουν από μόνα τους τη διαδικασία της πολιτικής επιλογής.

Μέσα από την μελέτη της εκλογικής συμπεριφοράς παρατηρείται η τάση, οι διάφορες κοινωνικές ομάδες όπως οι αγρότες και οι αστοί να παρουσιάζουν πολιτική ομοιογένεια, το οποίο οφείλεται στο γεγονός ότι οι σχέσεις των ατόμων, που ανήκουν στην ίδια κοινωνική οικογένεια έχουν την τάση να διαμορφώνουν κοινές αξίες και ιδέες, οι οποίες με τη σειρά τους καθορίζουν πως θα ενεργήσει πολιτικά η συγκεκριμένη

κοινωνική ομάδα. Εφόσον, δηλαδή, οι πολίτες είναι ενταγμένοι σε κοινωνικές οικογένειες, μας δίνεται η δυνατότητα να ερμηνεύσουμε την εκλογική συμπεριφορά του καθενός με τη χρήση τριών σημαντικών παραγόντων, συγκεκριμένα με βάση την κοινωνικό - επαγγελματική θέση των πολιτών, την ηλικία, την θρησκευτική ταυτότητα, καθώς και τον τόπο κατοικίας, οι οποίοι για πολλούς ερευνητές αποτελούν τους κύριους διαμορφωτές της ψήφου.

Ενας άλλος παράγοντας που καθορίζει την εκλογική συμπεριφορά είναι η προεκλογική εκστρατεία. Για πολλούς όμως ερευνητές η επιρροή της καθίσταται ιδιαιτέρως περιορισμένη, το οποίο οφείλεται στο γεγονός ότι η πλειονότητα των ψηφοφόρων, πριν ακόμα από την έναρξη της προεκλογικής περιόδου, είναι κατασταλαγμένοι και έχουν αποφασίσει ήδη την εκλογική τους επιλογή και παραμένουν πιστή σε αυτή ακόμα και όταν προσέρχεται στην κάλπη για να ασκήσει το εκλογικό της δικαίωμα. Από την άλλη μεριά υπάρχουν και οι πολιτικοποιημένοι ψηφοφόροι οι οποίοι διαμορφώνουν χρονικά έγκαιρα την ψήφο τους που αναπαράγεται πιστά από τη μία εκλογική αναμέτρηση στην άλλη, και επιδεικνύουν ενδιαφέρον για την προεκλογική εκστρατεία. Σε αντίθεση με τους αναποφάσιστους ψηφοφόρους των οποίων η ψήφος επηρεάζεται, η τελευταία αποκρυσταλλώνεται από δύο άλλους παράγοντες, την «opinion leaders» και τα ΜΜΕ.

Η κομματική ταύτιση θεωρείται ένας από τους παράγοντες, ο οποίος σύμφωνα με την αμερικάνικη σχολή καθορίζει τις πολιτικές στάσεις των εκλογέων. Παρατηρείται, μάλιστα, σε μεγάλο βαθμό το φαινόμενο οι εκλογείς να ταυτίζονται με το κόμμα τους και να επικροτούν τις αποφάσεις που λαμβάνονται στα άδυτα της πολιτικής τους ομάδας. Συγκεκριμένα, στους πολιτικοποιημένους εκλογείς η κομματική ταύτιση οδηγεί στην στήριξη των υποψηφίων στον τρόπο που ασκούν την πολιτική τους, καθώς επίσης φτάνουν στο σημείο να ταυτίζουν τις πολιτικές τους απόψεις με εκείνες των υποψηφίων. Όμως, η επίδραση που ασκεί η κομματική ταύτιση στην ομάδα των ψηφοφόρων που δεν ασχολούνται με το πολιτικό πεδίο είναι διαφορετική. Οι συγκεκριμένοι εκλογείς μη έχοντας αναπτύξει επαρκώς τις πολιτικές τους αντιλήψεις

και γνωρίζοντας μονάχα σε ποιο κομματικό στρατόπεδο ανήκουν, η κομματική ταύτιση θα αποτελέσει απλώς το έναυσμα για την πολιτική τους κινητοποίηση, ακόμα και αν οι ίδιοι δεν παρακολουθούν τα δρώμενα της προεκλογικής εκστρατείας προκειμένου να σχηματίσουν τον ιδεολογικό τους προσανατολισμό.

Τέλος, αξίζει να αναφερθούν τα τρία μοντέλα εκλογών όπως διαμορφώθηκαν από τους ερευνητές του αμερικάνικου Πανεπιστημίου. Η πρώτη κατηγορία είναι οι «εκλογές διατήρησης» οι οποίες καθορίζονται κυρίως από το ποσοστό ενίσχυσης της κομματικής ταύτισης στους κόλπους των ψηφοφόρων επειδή σε αυτές τις αναμετρήσεις απουσιάζει ο «χαρισματικός ηγέτης», ο οποίος θα κινητοποιήσει τους εκλογείς και το ισχυρό πολιτικό φρόνιμα το οποίο έχει σαν αποτέλεσμα την εμφάνιση του φαινομένου να προσέρχονται στην κάλπη οι ψηφοφόροι διακατεχόμενοι από το αίσθημα της πίστης και της αφοσίωσης στο κομματικό τους στρατόπεδο. Η δεύτερη είναι οι «εκλογές παρεκτροπής», όπου οι εκλογείς αναιρούν και αποστρέφουν την κομματική τους ταύτιση, υπέρ του αντίπαλου υποψηφίου δηλαδή, η τελική εκλογική επιλογή έρχεται σε αντίθεση με την κομματική ταύτιση τους. Τέλος, οι «εκλογές αποστοίχησης», οι οποίες οδηγούν στην πλήρη αλλαγή και διαφοροποίηση των κομματικών ταυτίσεων, δηλαδή σε μια πολιτική και ιδεολογική σύγχυση, η οποία είναι αποτέλεσμα μιας σοβαρής πολιτικής, κοινωνικής ή οικονομικής συγκυρίας.

Μελετώντας, λοιπόν, τον περίπλοκο θεσμό των εκλογών και προσπαθώντας να αποκωδικοποιήσουμε και να αναλύσουμε την εκλογική συμπεριφορά των ψηφοφόρων καταλήγουμε στο συμπέρασμα ότι τόσο η ανάλυση των εκφραζόμενων προτιμήσεων τους, όσο και ο βαθμός κινητοποίησής τους, αποτελούν και συνιστούν τους κοινωνικούς μηχανισμούς της πολιτικής συμμετοχής (Francois κ.ά, 2008).

### **2.3. Παράγοντες καταλυτικοί στην επιλογή ενός συγκεκριμένου κομματικού φορέα από τους ψηφοφόρους.**

Αρκετοί ερευνητές υποστηρίζουν ότι η ψήφος είναι αποτέλεσμα της ορθολογικής επιλογής, σε αντίθεση με άλλους που είναι προασπιστές της άποψης ότι η εκλογική

επιλογή είναι αποτέλεσμα της ύπαρξης του κοινωνικού ντετερμινισμού. Δηλαδή, ορισμένοι επισημαίνουν την αντίληψη ότι η ψήφος αποτελεί μία προσωπική επιλογή του εκάστοτε ψηφοφόρου, η οποία διέπετε από ορθολογικότητα, ενώ από την άλλη μεριά υπάρχουν οι προασπιστές της άποψης ότι η ψήφος είναι συνυφασμένη με τα κοινωνιολογικά και δημογραφικά χαρακτηριστικά του ψηφοφόρου. Τα κοινωνικά χαρακτηριστικά που σχετίζονται άμεσα με την εκλογική συμπεριφορά είναι οι λεγόμενες «μεταβλητές» της ψήφου, οι οποίες έχουν αξιοποιηθεί σε ποικίλες έρευνες που έλαβαν χώρα σε ένα συγκεκριμένο χρονικό, τοπικό και εκλογικό πλαίσιο. Παρά το γεγονός ότι το αποτέλεσμα της κάλπης δεν παραμένει πάντα σταθερό αλλά μεταβάλλεται από την μια εκλογική αναμέτρηση μέχρι την επόμενη καθώς οι συγκυρίες διαρκώς μεταβάλλονται, ωστόσο, αξίζει να επισημάνουμε τη σημασία και την επιρροή που ασκούν τα κοινωνιολογικά χαρακτηριστικά στη διαμόρφωση της ψήφου.

Η διερεύνηση των κοινωνικών και δημογραφικών παραγόντων της εκλογικής προτίμησης μας δίνει την δυνατότητα να κατανοήσουμε σε ποιο βαθμό οι αντικειμενικές μεταβλητές της ψήφου, όπως το φύλο, η ηλικία, ο τόπος κατοικίας, ο επαγγελματικός προσανατολισμός και η θρησκεία έχουν την τάση να επηρεάζουν και να κατευθύνουν τις πολιτικές και ιδεολογικές επιλογές των εκλογέων. Επίσης, η ένταξη των πολιτών σε μια κοινωνικό - επαγγελματική κατηγορία συνιστά ένα σημαντικό παράγοντα, που συμβάλλει στη διαμόρφωση της εκλογικής τους συμπεριφοράς και συσχετίζεται τόσο με την εργασία, όσο και με τη συνοχή και το βαθμό ταύτισης των εκλογέων με τις αξίες και τους κανόνες που διέπουν την κοινωνική τους ομάδα.

Σύμφωνα με το κοινωνιολογικό μοντέλο, οι δύο βασικοί προσδιοριστικοί παράγοντες της εκλογικής συμπεριφοράς που διαδραματίζουν σημαντικό ρόλο στην εκλογική ανάλυση είναι το θρήσκευμα και η κοινωνική τάξη, οι οποίες έχουν χαρακτηριστεί ως «βαριές μεταβλητές». Έχουν, λοιπόν, παρουσιαστεί δύο διαφορετικά «πολιτιστικά μοντέλα» ψήφου. Το πρώτο χαρακτηρίζεται από τους «βαθιά θρησκευόμενους» οι οποίοι κατά βάση ανήκουν στις μεσαίες τάξεις είτε των πόλεων είτε της υπαίθρου, ενώ

το δεύτερο ορίζεται από εκείνους «χωρίς θρήσκευμα» οι οποίοι συνήθως υπάγονται στην εργατική κοινωνική τάξη.

Ενα άλλο μοντέλο που αποσκοπεί στην ερμηνεία της εκλογικής συμπεριφοράς είναι το «καταναλωτικό μοντέλο» σύμφωνα με το οποίο η απόφαση της ψήφου μοιάζει πολύ με την απόφαση της αγοράς ενός προϊόντος. Όπως ο καταναλωτής έτσι και ο ψηφοφόρος τοποθετείται απέναντι σε ό, τι του προσφέρεται όπως προγράμματα, υποψήφιοι, κ.λπ. Είναι, ωστόσο, επηρεασμένος από τις παλαιές του αγοραστικές συνήθειες, «προηγούμενες ψήφοι», από την προτίμηση σε ορισμένες μάρκες, «κομματική ταύτιση», από την επιρροή των ομάδων αναφοράς (περιοχή κατοικίας, επαγγελματικός χώρος, κ.λπ.). Ωστόσο, σε κάθε εκλογή του δίνεται η ευκαιρία για μια «αλλαγή συνηθειών», διότι τα οφέλη που προσφέρονται δεν είναι ποτέ ίδια.

Παρουσιάζεται, όμως, το φαινόμενο ψηφοφόροι που ανήκουν στους κόλπους της ίδιας πολιτικής και κοινωνικής «οικογένειας» να παρουσιάζουν διαφορετικές εκλογικές προτιμήσεις στην επεξήγηση του οποίου συμβάλει η ύπαρξη δευτερευόντων κοινωνιολογικών παραγόντων. Συγκεκριμένα, έχει παρατηρηθεί ότι η επαγγελματική κατάρτιση σχετίζεται άμεσα με τον τόπο κατοικίας και διαδραματίζει σημαντικό ρόλο στον καθορισμό της εκλογικής προτίμησης των ψηφοφόρων στα πλαίσια μιας κοινωνικής ομάδας. Το χωροταξικό περιβάλλον παρουσιάζει σημαντικό ρόλο σε σχέση με την τάση των εκλογέων να διαφοροποιούνται ιδεολογικά με την τάξη στην οποία είναι κοινωνικά ενταγμένοι. Για παράδειγμα, ο εργάτης οποίος εργάζεται σε μια αστική βιομηχανική πόλη έχει διαφορετική συμπεριφορά σε σχέση με έναν εργάτη, η έδρα του οποίου βρίσκεται σε μία επαρχιακή πόλη, ακόμα και αν απασχολούνται με τον ίδιο τομέα. Πιο αναλυτικά, ο τόπος κατοικίας αλλά και το κοινωνικό περιβάλλον το οποίο αποκτά κάποιος και με το οποίο συναναστρέφεται, καθορίζει σε σημαντικό βαθμό τις διαφορετικές συμπεριφορές που συναντούμε στα πλαίσια της ίδιας κοινωνικής ομάδας. Οι πολιτικές, οικονομικές, κοινωνιολογικές συγκυρίες μεταβάλλονται από περιοχή σε περιοχή και αυτό έχει σαν αποτέλεσμα να παρατηρείται το φαινόμενο οι εκλογείς να μην ψηφίζουν με βάση την ταξική και την κοινωνική τους προέλευση.

Συνοπτικά, η αναγκαιότητα της συγκεκριμένης πρότασης περιλαμβάνει την μελέτη της εκλογικής συμπεριφοράς στον ελληνικό πολιτικό χώρο, τον τρόπο διαμόρφωσης της εκλογικής συμπεριφοράς των ψηφοφόρων, καθώς επίσης ποιοι από τους ήδη υπάρχοντες κοινωνικούς και όχι μόνο παράγοντες διαδραματίζουν σημαντικό ρόλο στην διαμόρφωση του καθορισμού της εκλογικής συμπεριφοράς. Κάνοντας, λοιπόν, μία ιστορική αναδρομή ξεκινώντας από την γέννηση του Δημοκρατικού πολιτεύματος μέχρι σήμερα, με λύπη συμπεραίνουμε ότι το ενδιαφέρον των πολιτών για ενασχόληση με τα κοινά έχει συρρικνωθεί αισθητά. Μπορεί ο θεσμός της ψηφοφορίας να έχαιρε μεγίστου σεβασμού από τους πολίτες της Αρχαίας Αθήνας, αλλά αυτό πλέον δεν ισχύει, καθώς το ενδιαφέρον και η συμμετοχή έχει δώσει τη θέση της στην αδιαφορία, στην απαξίωση και στην αποχή από την εκλογική διαδικασία και κατ' επέκταση του εκλογικού μας συστήματος.

#### **2.4. Μοντέλα Ανάλυσης Εκλογικής Συμπεριφοράς**

Στη σύγχρονη Ελλάδα η εκλογική συμπεριφορά έχει σχεδόν ταυτιστεί με τις πολιτικές έρευνες της κοινής γνώμης και τις δημοσκοπήσεις στερώντας από την εκλογική ανάλυση το θεωρητικό υπόβαθρο που είναι απαραίτητο για τη σωστή κατανόηση και παρατήρηση των συμπεριφορών του δείγματος (Mayer, 2005).

Πριν επιλέξουμε ή αναπτύξουμε ένα μοντέλο ανάλυσης της εκλογικής συμπεριφοράς, θα πρέπει αρχικά να έχουμε κατανοήσει το θεωρητικό πλαίσιο που σχετίζεται με την κοινωνιολογία της εκλογικής συμπεριφοράς, ώστε να μπορέσουμε να στηριχτούμε σε μια θεωρητική μέθοδο όσον αφορά τη διεξαγωγή των συμπερασμάτων μας.

Δύο κύριοι τύποι ανάλυσης έχουν αναπτυχθεί στην μελέτη της εκλογικής συμπεριφοράς. Ο πρώτος είναι η ανάλυση χωρικών δεδομένων (γεωγραφικών, ιστορικών, πολιτισμικών, κοινωνικών) και ο δεύτερος είναι τα ατομικά δεδομένα (δημογραφικά χαρακτηριστικά, ιδεολογία, πεποιθήσεις).

Οι κύριοι μέθοδοι ανάλυσης της εκλογικής συμπεριφοράς που έγκεινται στο θεωρητικό πεδίο της πολιτικής επιστήμης είναι οι εξής:

- I. Εκλογική γεωγραφία
- II. Ποσοτική οικολογία και η ανάλυση συνθηκών
- III. Κοινωνικός ντετερμινισμός του Lazarsfeld
- IV. Κομματική ταύτιση της Σχολής του Μίσιγκαν
- V. Ορθολογική επιλογή.

#### 2.4.1. Εκλογική γεωγραφία

Η εκλογική γεωγραφία είναι η μελέτη και ανάλυση των γεωγραφικών μοντέλων, δηλαδή αναλύεται η εκλογική συμπεριφορά με βάση τα κοινωνικά και πολιτισμικά χαρακτηριστικά που παρουσιάζουν οι ψηφοφόροι των επιμέρους γεωγραφικών ενοτήτων. Η έννοια της Εκλογικής γεωγραφίας εισήχθη το 1913 με το έργο του André Siegfried, «Tableau Politique de la France de l’Ouest sous la IIIeme République» (Πολιτικός Χάρτης της δυτικής Γαλλίας κατά την Γ' Δημοκρατία).

Η συγκεκριμένη μέθοδος βασίζεται στη χωρική συσχέτιση των γεωγραφικά οργανωμένων κοινωνικών δομών και της κατανομής των ψήφων (Mayer, 2005). Έχει ως βασικό χαρακτηριστικό την αποτύπωση των εκλογικών αποτελεσμάτων ή των κοινωνικοπολιτικών χαρακτηριστικών σε χάρτες. Είναι ένα μοντέλο πολυπαραγοντικής ανάλυσης, του οποίου οι μεταβλητές είναι η μορφολογία του εδάφους, ο τύπος κατοικίας, η μορφή ιδιοκτησίας, η επιρροή της θρησκείας και οι κοινωνικοοικονομικές και πολιτισμικές δομές μιας γεωγραφικά προσδιορισμένης ενότητας. Από το 1970 στη μέθοδο προστέθηκαν και άλλοι παράγοντες, συγκεκριμένα η δημιουργία μικρότερων και διάσπαρτων γεωγραφικών ενοτήτων (γειτονιές) και η επιρροή των M.M.E. στη διαμόρφωση της κοινής γνώμης.

Στην Ελλάδα υλοποιήθηκε η παραπάνω μέθοδος στη στατιστική ανάλυση από την εταιρεία ερευνών Public Issue το 2004 και το 2007, με σκοπό την εκτίμηση των εκλογικών αποτελεσμάτων αλλά χωρίς ιδιαίτερη επιτυχία αφού τα γεωγραφικά

χαρακτηριστικά των περιοχών αποτελούσαν παράγοντα εκτίμησης της ψήφου σε ορισμένες μόνο περιπτώσεις. (Μαυρής & Συμεωνίδης, 2005).

#### **2.4.2. Ποσοτική οικολογία και η ανάλυση συνθηκών**

Η ποσοτική οικολογία είναι ο συνδυασμός της στατιστικής ανάλυσης δεδομένων και εκλογικής γεωγραφίας. Δηλαδή, η ανάλυση της εκλογικής συμπεριφοράς βασίζεται στον εντοπισμό μεταβλητών που την προσδιορίζουν ανά περιοχή ή στην κατανομή των περιοχών ανάλογα με τη στάση των ψηφοφόρων, ή τέλος στην κατασκευή μιας κλιμακωτής ομαδοποίησης των ψηφοφόρων με βάση την επιλογή τους.

Αντίστοιχα, η ανάλυση συνθηκών είναι συνδυασμός μεταβλητών, των δημογραφικών και γεωγραφικών δεδομένων χρησιμοποιώντας με τη μέθοδο της δειγματοληψίας και δεδομένων που απεικονίζουν κοινωνικά χαρακτηριστικά των ψηφοφόρων (Mayer, 2005).

#### **2.4.3. Τα Διπολικά Σχήματα Ανάλυσης Της Εκλογικής Συμπεριφοράς**

Στις αρχές της δεκαετίας του 1960 η έρευνα των Guy Michelat και Jean Pierre Thomas ανέπτυξε την έννοια της παραταξιακής ταύτισης, ορίζοντας ως κύριους πόλους διάκρισης τον άξονα Δεξιά / Αριστερά. Από το 1966 και έπειτα προστέθηκε συστηματικά στην πολιτική έρευνα η χρήση της επταβάθμιας κλίμακας Δεξιά / Αριστερά όπου με 1 ορίζεται το αριστερό άκρο και με 7 το δεξιό άκρο του άξονα.

Η παραταξιακή ταύτιση, όπως και η κομματική στηρίζονται στα κοινωνικά και πολιτιστικά χαρακτηριστικά των πολιτών, η οποία εντείνεται με την αύξηση του πολιτικού ενδιαφέροντος και της πολιτικοποίησης αντίστοιχα. Στην εκλογική συμπεριφορά η πρόθεση της ψήφου ερμηνεύοταν με όρους παραταξιακής ταύτισης σε πολυκομματικές χώρες όπως η Γερμανία και η Γαλλία και σε δικομματικά πολιτικά συστήματα όπως αυτά της Βρετανίας και των Η.Π.Α με μεγαλύτερη επιτυχία.

Οι πολιτικοί ερευνητές ανέδειξαν και άλλες μορφές κοινωνικών διπόλων μέσα από τις οποίες μπορούσαν να ερμηνεύσουν την εκλογική συμπεριφορά, όπως εκείνον της

θρησκευτικότητας και τη κοινωνικής τάξης. Σύμφωνα με έρευνες που έχουν πραγματοποιηθεί διαπιστώθηκε ότι συγκριμένα κοινωνικά και θρησκευτικά χαρακτηριστικά ανήκαν στη Δεξιά πλευρά ή στην Αριστερή του πολιτικού ιδεολογικού άξονα. Έτσι, ένας πολίτης πιο θρησκευόμενος επιλέγει να ψηφίσει δεξιά ή κεντροδεξιά κόμματα, σε αντίθεση με έναν πολίτη που ασχολείται ελάχιστα ή καθόλου με θρησκευτικά ζητήματα εντάσσεται πολιτικά στο χώρο της Αριστεράς. Αντίστοιχα, πολίτες που ανήκουν στις εργατικές τάξεις τείνουν να ψηφίζουν αριστερά ή κεντροαριστερά κόμματα, ενώ τα ανώτερα κοινωνικά στρώματα εντάσσονται στο Δεξιό τμήμα του άξονα.

Κατέληξαν, λοιπόν, στο συμπέρασμα ότι μελετώντας τα κοινωνικά χαρακτηριστικά των πολιτών με απλές ερωτήσεις που αφορούν συνήθειες και αυτοχαρακτηρισμό των πολιτών όπως για παράδειγμα η αυτοτοποθέτηση σε κοινωνική τάξη, τα ατομικά χαρακτηριστικά, το εκπαιδευτικό επίπεδο, η θέση στο επάγγελμα, το εισόδημα, οι θρησκευτικές συνήθειες είναι δυνατόν να υλοποιηθούν εκτιμήσεις σε σχέση με την πολιτική τοποθέτηση του ατόμου και την εκλογική του συμπεριφορά (Mayer, 2005).

#### **2.4.4. Η Έννοια Της Κομματικής Ταύτισης Της Σχολής Του MICHIGAN**

Η σχολή του Michigan εξελίσσει τη θεωρία της Σχολής του Columbia, εισάγοντας την ψυχολογία στα κριτήρια που συνδέουν τα χαρακτηριστικά των ψηφοφόρων με την εκλογική τους συμπεριφορά. Σύμφωνα με τη Σχολή του Michigan η άποψη ότι το κοινωνικό υπόβαθρο του ψηφοφόρου επηρεάζει την επιλογή ψήφου δεν είναι αρκετή για να επεξηγήσει τις επιμέρους διασυνδέσεις. Είναι πολύ σημαντικό να ερευνάται και ο τρόπος με τον οποίο δημιουργούνται αυτές οι διασυνδέσεις, το οποίο απαντά η Σχολή του Michigan μέσα από την ανάπτυξη της έννοιας της κομματικής ταύτισης.

Κομματική ταύτιση είναι η σταθερή πίστη ενός ψηφοφόρου προς έναν κομματικό/πολιτικό σχηματισμό. Δηλαδή, ο ψηφοφόρος αποδέχεται την ιδεολογία, τις αξίες, τις προγραμματικές θέσεις και τα πρόσωπα του σχηματισμού που υποστηρίζει. Όμως, η κομματική ταύτιση μπορεί να επεξηγήσει διάφορες μορφές εκλογικής

συμπεριφοράς. Για παράδειγμα, η θεωρία του Lazarsfeld για τις κοινωνικές ομάδες και τα συνδικάτα τα οποία συνήθως συνδέονται με ένα συγκεκριμένο κόμμα. Δηλαδή, μέλη μειονοτικών ομάδων και συνδικάτων που στηρίζουν συγκεκριμένα κόμματα στρέφονται προς τα κόμματα αυτά.

Επίσης ο βαθμός κομματικής ταύτισης είναι ανάλογος του βαθμού ενδιαφέροντος και ενασχόλησης με τα κοινά. Δηλαδή οι πολίτες που ενδιαφέρονται για τα πολιτικά ζητήματα παρουσιάζουν αυξημένο συντελεστή κομματικής ταύτισης, ενώ οι ψηφοφόροι που δεν εκδηλώνουν έντονο ενδιαφέρον εμφανίζουν χαμηλό δείκτη κομματικής ταύτισης. Άρα, η κομματική ταύτιση δεν μπορεί να αποτελέσει μοναδικό κριτήριο ανάλυσης της εκλογικής συμπεριφοράς, αλλά θα πρέπει να λαμβάνεται υπόψη συμπληρωματικά με άλλες μεταβλητές που αφορούν και άλλα κοινωνικά χαρακτηριστικά των ψηφοφόρων, τις συνθήκες διεξαγωγής των εκλογών και τα άτομα που συμμετέχουν σε αυτές.

Για το λόγο αυτό, λοιπόν, η Σχολή του Μίσιγκαν διακρίνει τις εκλογικές αναμετρήσεις σε τρεις ομάδες:

- I. στις εκλογές διατήρησης, που χαρακτηρίζονται σε μεγάλο βαθμό από την κομματική ταύτιση του εκλογικού σώματος
- II. στις εκλογές παρεκτροπής, όπου υπάρχει αντίστοιχη συμμετοχή των μεταβλητών της κομματικής ταύτισης και της πολιτικής επιλογής
- III. στις εκλογές αποστοίχισης, κατά τις οποίες η κομματική ταύτιση μεταφράζεται ως η μετακίνηση των ψηφοφόρων στα νέα κόμματα που δημιουργούνται έπειτα από την ανατροπή των έως τότε υφιστάμενων κομματικών σχηματισμών (Mayer, 2005).

#### 2.4.5. Το Μοντέλο Της Ορθολογικής Επιλογής Και η Θεματική Ψήφος

Με βάση το μοντέλο της ορθολογικής επιλογής συμπεραίνουμε ότι η εκλογική συμπεριφορά δεν περιορίζεται σε πεποιθήσεις και κοινωνικές ή ιδεολογικές καταβολές. Η έννοια της θεματικής ψήφου έπαιξε πολύ σημαντικό ρόλο στην ανάπτυξη του νέου

μοντέλου ανάλυσης της εκλογικής συμπεριφοράς, συγκεκριμένα η θεματική ψήφος ισχυρίζεται ότι οι ψηφοφόροι επιλέγουν τι θα ψηφίσουν με βάση τα θέματα που κυριαρχούν στην πολιτική σκηνή κατά την περίοδο που διεξάγονται οι εκλογές. Το κριτήριο εντοπίζεται στην τοποθέτηση των ψηφοφόρων ως προς τα θέματα αυτά και η συσχέτισή του με τις αντίστοιχες τοποθετήσεις των κομμάτων ή των υποψηφίων που συμμετέχουν στην εκλογική διαδικασία.

Υπάρχουν, όμως, συγκεκριμένες προϋποθέσεις που ευνοούν την ανάπτυξη αυτού του μοντέλου ανάλυσης, συγκεκριμένα, η σωστή ενημέρωση της κοινής γνώμης για τα πολιτικά δρώμενα και το ενδιαφέρον των ψηφοφόρων για αυτά, η γνώση από την πλευρά των ψηφοφόρων όσον αφορά τις τοποθετήσεις των υποψηφίων κομμάτων ή προσώπων πάνω στα ζητήματα πολιτικού ενδιαφέροντος και τέλος η συνειδητή ταύτιση της τελικής τους επιλογής με έναν υποψήφιο σχηματισμό ή πρόσωπο που μοιράζεται παρόμοιες θέσεις πάνω στα θέματα που απασχολούν την πολιτική ζωή κατά την εκλογική περίοδο.

Η ορθολογική επιλογή αναπτύχθηκε εξαιτίας της κρίση που εμφανίστηκε στην έννοια της κομματικής ταύτισης εξαιτίας της αύξησης των ψηφοφόρων που δήλωναν αναποφάσιστοι και ανεξάρτητοι ως προς τις κομματικές τους επιλογές. Ο βαθμός κομματικής ταύτισης, αλλά και η ταύτιση με άλλες έννοιες όπως αυτή του θρησκεύματος και της τοπικότητας εμφάνισαν μεγάλη πτώση κατά τη δεκαετία του 1970 στις αμερικανικές εκλογές.

Παράλληλα, αναπτύχθηκε η θεωρία του *Homo Economicus*, με βάση την οποία ο ψηφοφόρος σκέφτεται και πράττει στηριζόμενος στη σχέση κόστος – όφελος. Στην πολιτική έρευνα αυτή η νέα θεωρία αναδεικνύει τα κόμματα σε «επιχειρήσεις», στις οποίες απευθύνονται οι ψηφοφόροι ως «πελάτες» με σκοπό να ψηφίσουν αυτόν που θα τους προσφέρει το μεγαλύτερο όφελος, με το μικρότερο από την πλευρά τους κόστος.

Το φαινόμενο των πελατειακών σχέσεων εμφανίστηκε στην Ελλάδα κατά τη διάρκεια των νεότερων χρόνων της πολιτικής ιστορίας της. Ένα μεγάλο μέρος των Ελλήνων ψηφοφόρων θέτει ως κριτήριο επιλογής κομμάτων και προσώπων την εξυπηρέτηση

προσωπικών συμφερόντων. Στη συγκεκριμένη περίπτωση, η ψήφος έχει ανταλλακτική αξία, δηλαδή οι ψηφοφόροι απολαμβάνουν προνόμια και διευκολύνσεις σε διάφορους τομείς που τους ενδιαφέρουν ως αντάλλαγμα την ψήφο τους το οποίο έχει ως αποτέλεσμα σταδιακά να χτίζεται ένα διαρκώς ανακυκλούμενο σύστημα που παρέχει πολιτική στήριξη σε πρόσωπα τα οποία εξυπηρετούν τους ψηφοφόρους τους, οι οποίοι με τη σειρά τους εξασφαλίζουν στα πρόσωπα αυτά σταθερή, μακρά και ισχυρή πολιτική θέση.

Μια άλλη μορφή ανάλυσης που εφαρμόζεται χρησιμοποιώντας ως εργαλεία το κοινωνιολογικό μοντέλο και την ορθολογική επιλογή είναι το καταναλωτικό μοντέλο, δηλαδή ο ψηφοφόρος αντιμετωπίζεται ως καταναλωτής και το κόμματα ως προϊόν. Αυτή η μέθοδος, χρησιμοποιείται κυρίως στις εκτιμήσεις ψήφου που βασίζονται στην καταγραφή της πρόθεσης ψήφου σε σύγκριση με την προηγούμενη. Στην συγκεκριμένη περίπτωση, ο καταναλωτής – ψηφοφόρος καλείται να «αγοράσει» ένα από τα προσφερόμενα προϊόντα – κόμματα. Και πάλι ως *Homo Economicus* θα υπολογίσει το «βαθμό χρησιμότητας» του κάθε κόμματος με βάση τις ανάγκες του ίδιου ή του κοινωνικού συνόλου στο οποίο ανήκει.

Η σύγκλιση που παρατηρείται τα τελευταία χρόνια στην Ελλάδα των παραδοσιακών κομμάτων που εξέφραζαν τη διαιρετική τομή Δεξιάς/Αριστεράς προς το κεντρώο χώρο είχε ως αποτέλεσμα την αύξηση των ποσοστών των μη κομματικοποιημένων ψηφοφόρων και των αναποφάσιστων με συνεπακόλουθο η μέθοδος της ορθολογικής επιλογής να αρχίσει να χρησιμοποιείται περισσότερο ως κριτήριο στην ανάλυση της εκλογικής συμπεριφοράς.

Ωστόσο, λόγω του γεγονότος ότι οι προϋποθέσεις για πλήρη εφαρμογή της έννοιας της θεματικής ψήφου δεν καλύπτονται ακόμη στην περίπτωση της Ελλάδας, η έρευνα λαμβάνει υπ' όψιν της και κριτήρια όπως αυτό της κομματικής προτίμησης αλλά με λιγότερο βαθμό συμμετοχής. Με αυτόν τον τρόπο εισάγεται η έννοια της ανάλυσης του εκλογικού ανταγωνισμού υπό το σχήμα «Κόμματα – Θέματα – Πρόσωπα» και τις

συσχετίσεις που δημιουργούνται ανάμεσα στους τρεις αυτούς παράγοντες διαμόρφωσης της εκλογικής συμπεριφοράς (Mayer, 2005).

## 2.5. Διαιρετικές Τομές-Εκλογική Συμπεριφορά

Οι Lipset και Rokkan (1967) αρθρογράφησαν το κεφάλαιο «Cleavage Structures, Party Systems, and Voter Alignments, Cross – National Perspectives» όπου εκτίμησαν ότι τα κομματικά συστήματα δεκαέξι χωρών, χαρακτηρίζονται από τέσσερις διαιρετικές τομές. Αρχικά είναι η ταξική τομή, δηλαδή, εργάτες – εργοδότες, η θρησκευτική τομή, δηλαδή η ρήξη μεταξύ κράτους-εκκλησίας, ακόμα η πολιτισμική τομή, δηλαδή η ρήξη μεταξύ του Κέντρου του νεόδμητου πολιτικού συστήματος και μιας υποτελούς Περιφέρειας και τέλος η αγροτική τομή, δηλαδή η σύγκρουση Πόλης – Υπαίθρου. Αξιοσημείωτο είναι το γεγονός ότι οι δύο πρώτες διαιρετικές τομές, προέκυψαν από τις εθνικές επαναστάσεις σύστασης των σύγχρονων εθνών κρατών και οι άλλες δύο από τη βιομηχανική επανάσταση και τη μετάβαση από τη φεουδαρχία στον καπιταλισμό. Επίσης, πρέπει να αναφερθεί το γεγονός ότι δεν εμφανίζονται και οι τέσσερις αυτές διαιρετικές τομές σε κάθε χώρα, αλλά ούτε οι όροι και οι συνθήκες ανάδυσης της κάθε τομής είναι κοινοί. Η «κοινή» σε όλες τις χώρες τομή είναι η ταξική, δηλαδή η σύγκρουση μεταξύ εργατών και εργοδοτών. Σ' αντίθεση με τις υπόλοιπες τομές που σημειώνουν τις ιδιαιτερότητες και τις διαφορές. Όπως, για παράδειγμα, η τομή Κέντρο – Περιφέρεια είναι διαφορετική στην ιβηρική χερσόνησο, σε σχέση με τα αποσχιστικά κινήματα Βάσκων και Καταλανών από το καστεγιανικό κέντρο της Μαδρίτης. Αντίστοιχα, η αγροτική τομή στη Σκανδιναβία με την ανάδυση σημαντικών εκλογικά αγροτικών κομμάτων διαφέρει σημαντικά σε σχέση με άλλες περιπτώσεις χαμηλής έντασης σύγκρουσης μεταξύ πόλης και υπαίθρου. Τέλος, θα πρέπει να επισημανθεί το γεγονός ότι θα πρέπει να παρουσιάζεται μια συνέπεια ως προς την παρακολούθηση του περιεχομένου των διαιρέσεων στον χρόνο. Χαρακτηριστικό παράδειγμα, αποτελούν οι αλλαγές στη θρησκευτική διαιρετική τομή. Συγκεκριμένα η πρώιμη εννοιολόγησή της αναφέρεται στη διαμάχη προνομίων ανάμεσα στην Εκκλησία, την Καθολική συνήθως, και το Κράτος αναφορικά με το ποιος αποφασίζει και διαχειρίζεται θέματα όπως η

παιδεία και η εκπαίδευση, οι σχέσεις ηθικών θρησκευτικών προσταγών και νόμων του κράτους. Με τα χρόνια όμως το περιεχόμενο ανασυνθέτεται, συγκεκριμένα παρατηρείται ότι η σύγκρουση πλέον μετατίθεται στις εθιμικές θρησκευτικές πρακτικές, όπως για παράδειγμα η ισλαμική μαντίλα, οι εορτασμοί και οι αργίες, αξιακές αναζητήσεις που συνδέονται με την ίδια τη θρησκευτικότητα και την πίστη όπως για παράδειγμα τα κινήματα κατά των αμβλώσεων (Πιερίδης, 2021).

## **Κεφάλαιο 3: Δυαδικά Λογαριθμογραμμικά Παλίνδρομα Μοντέλα (Binomial Logistic Regression, BLM), ο προάγγελος των Πολυωνυμικών Λογαριθμογραμμικών Παλίνδρομων Μοντέλων.**

### **3.1. Προκαταρτικά**

Αν  $p$  είναι η πιθανότητα ενός γεγονότος ή πιθανότητα επιτυχίας, τότε  $1-p$  είναι η πιθανότητα του μη γεγονότος ή αλλιώς πιθανότητα αποτυχίας.

Στη συνέχεια, μπορούμε να σχηματίσουμε το λόγο  $p/(1-p)$  στο οποίο δίνουμε το όνομα **odds**. Συμβολικά έχουμε:

$$\text{odds} = \frac{\text{probability\_of\_event}}{\text{probability\_of\_nonevent}} = \frac{p}{1-p} \quad (3.1)$$

Αν λάβουμε το νεπέριο λογάριθμο αυτού του λόγου λαμβάνουμε μια μαθηματική ποσότητα που έχει το όνομα **Logit**. Συμβολικά έχουμε:

$$\text{Logit}(p) = \ln\left(\frac{p}{1-p}\right) \quad (3.2)$$

Ας υποθέσουμε, ότι σε μια έρευνα έχουμε κωδικοποιήσει με 1 τους άνδρες και με μηδέν τις γυναίκες, δηλ. είχαμε τη μεταβλητή GENDER με 1= male και 0 = female, τότε θα είχαμε για τη σχετική πιθανότητα πραγματοποίησης του γεγονότος στους άνδρες:

$$\text{odds}_\alpha = \frac{p_\alpha}{1-p_\alpha} \quad (3.3)$$

Αντίστοιχα για τη σχετική πιθανότητα πραγματοποίησης του γεγονότος στις γυναίκες θα έχουμε:

$$\text{odds}_\gamma = \frac{p_\gamma}{1-p_\gamma} \quad (3.4)$$

Κι επομένως για το λόγο των σχετικών πιθανοτήτων, odds ratio ανδρών / γυναικών θα έχουμε:

$$\text{odds\_ratio} = \text{O.R.} = \frac{\frac{p_\alpha}{1-p_\alpha}}{\frac{p_\gamma}{1-p_\gamma}} \quad (3.5)$$

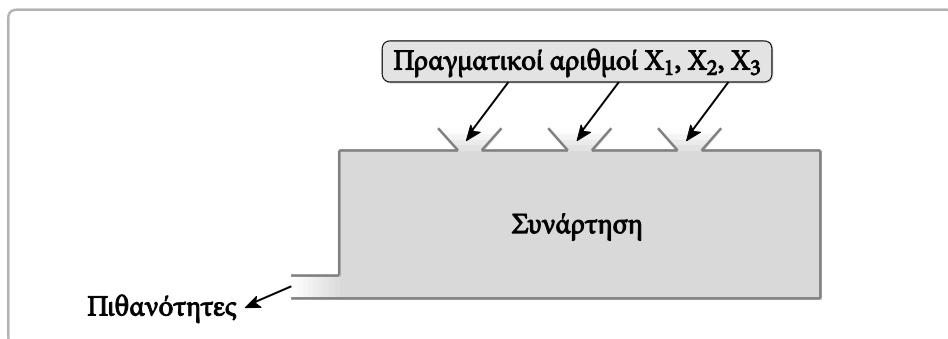
Κεφαλαιώδους σημασίας είναι οι εξής εκδοχές για την τιμή αυτού του λόγου, O.R.:

- I. Να μας προκύψει μεγαλύτερος της μονάδας.
- II. Να μας προκύψει μικρότερος της μονάδας.
- III. Να μας προκύψει ίσος με τη μονάδα.

### 3.2. Η κατασκευή του Μοντέλου

Είναι ανάγκη να κτίσουμε, μια μηχανή, μια συνάρτηση, ένα device, όπως ονομάζεται στη διεθνή ορολογία, που να έχει τη δυνατότητα, από τη μια μεριά να του βάζουμε πραγματικούς αριθμούς, και από την άλλη να μας δίνει πιθανότητες (Διάγραμμα 3.1).

Διάγραμμα 3.1. Διαγραμματική Απεικόνιση Λειτουργίας Συνάρτησης Κατανομής Πιθανότητας.



Ας ξεκινήσουμε λοιπόν από τον πιο γνωστό μηχανισμό μας, από τη γραμμική συνάρτηση παλινδρόμησης, στην πιο απλή της μορφή:

$$y = b_0 + b_1 X_1 \quad (3.6)$$

Από την (3.6) δεν μπορούμε να περιμένουμε και πολλά πράγματα. Πραγματικούς αριθμούς παίρνει, πραγματικούς αριθμούς δίνει. Οπότε, αν θέσουμε όπου γ το p, τότε λαμβάνουμε:

$$p = b_0 + b_1 X_1 \quad (3.7)$$

Αν θεωρήσουμε ότι το πρώτο μέρος αυτής της σχέσης (3.7), λαμβάνει τιμές πιθανότητας, δηλ. τιμές από το διάστημα [0,1], τότε παρατηρούμε ότι το δεύτερο που είναι η ποσότητα  $b_0 + b_1 X$  προφανώς μπορεί να είναι οποιοσδήποτε πραγματικός αριθμός. Επομένως, δεν έχουμε τον κατάλληλο μηχανισμό.

Αν αντικαταστήσουμε τώρα την πιθανότητα  $p$  του γεγονότος με την σχετική πιθανότητα του γεγονότος  $\frac{p}{1-p}$  τότε θα λάβουμε τη συνάρτηση:

$$\frac{p}{1-p} = b_0 + b_1 X_1 \quad (3.8)$$

Σε αυτή τη σχέση το πρώτο μέλος μπορεί να παίρνει τιμές θετικές, αλλά όχι και μικρότερες του 1, οπότε, ακόμη δεν έχουμε το κατάλληλο μοντέλο.

Ας πάρουμε το φυσικό λογάριθμο  $\ln$  της σχετικής πιθανότητας  $\frac{p}{1-p}$  και ας τον βάλουμε στη θέση του  $y$ . Τότε λαμβάνουμε μια καλή σχέση:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1 \quad (3.9)$$

Και λέμε καλή σχέση διότι και το πρώτο μέλος αυτής, αλλά και το δεύτερο είναι πραγματικοί αριθμοί.

Ωστόσο, εμείς, κυνηγάμε την πιθανότητα  $p$  η οποία είναι μέσα στο φυσικό λογάριθμο και μάλιστα σε δυο σημεία, στον αριθμητή, και στον παρονομαστή. Αν όμως λάβουμε τους αντιλογάριθμους των μελών της (3.9), θα πάρουμε τη σχέση:

$$e^{\ln\left(\frac{p}{1-p}\right)} = e^{(b_0 + b_1 X_1)} \quad (3.10)$$

Εύκολα από την (3.10) λαμβάνουμε την παρακάτω σχέση:

$$\frac{p}{1-p} = e^{(b_0 + b_1 X_1)} \quad (3.11)$$

Στη σχέση (3.11) θέτουμε  $z = b_0 + b_1 X_1$ .

Οπότε αυτή γίνεται:

$$\frac{p}{1-p} = e^z \quad (3.12)$$

Επιλύουμε την (3.12) ως προς  $p$  οπότε έχουμε διαδοχικά:

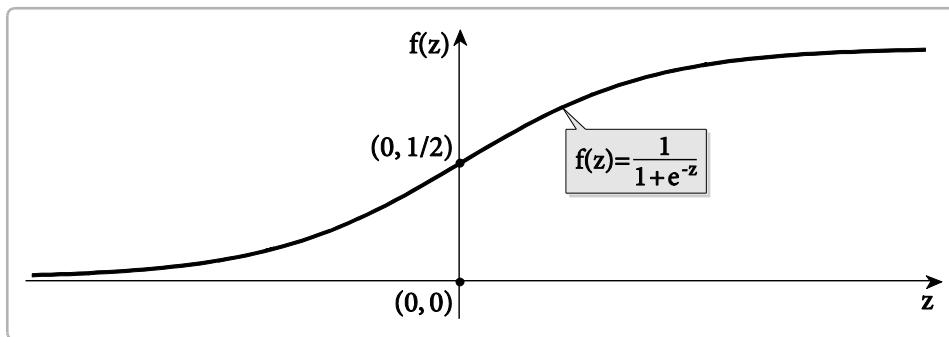
$$\begin{aligned}\frac{p}{1-p} = e^z &\Rightarrow p = e^z(1-p) \Rightarrow p = e^z - pe^z \Rightarrow \\ p + pe^z &= e^z \Rightarrow p(1+e^z) = e^z \Rightarrow p = \frac{e^z}{1+e^z} \Rightarrow \\ p &= \frac{1}{1+e^{-z}}\end{aligned}$$

Δηλαδή η συνάρτηση εκτίμησης της πιθανότητας γεγονότος είναι η  $f(z)$  με:

$$p(\text{event}) = f(z) = \frac{1}{1+e^{-z}} \quad (3.13)$$

Η γραφική παράσταση της συναρτήσεως (3.13) είναι μια σιγμοειδής συνάρτηση (*S-shaped*), την οποία βλέπουμε στο Διάγραμμα 3.2.

**Διάγραμμα 3.2.** Γραφική παράσταση συνάρτησης  $f(z)$ .



### Σημείωση

Έχουμε κατά νου ότι η μεταβλητή  $z$  είναι μια πραγματική μεταβλητή την οποία διαθέτουμε στα δεδομένα μας. Η μεταβλητή  $z$  δεν υπάρχει. Είναι μια αφανής, αδήλωτη μεταβλητή, για την οποία μας μιλάει η σχέση  $z = b_0 + b_1 X_1$  την οποία είδαμε παραπάνω.

### 3.3. Η Ερμηνεία του μοντέλου πρόβλεψης

Ανακαλούμε στη μνήμη την ερμηνεία των συντελεστών στην εξίσωση γραμμικής πολλαπλής παλινδρόμησης [Βλ. Κεφ. 15 στο βιβλίο Δαφέρμος, Β. (2011). Κοινωνική Στατιστική και Μεθοδολογία Έρευνας με το SPSS. Εκδόσεις Ζητη, Θεσσαλονίκη].

Το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης περιγράφεται από τη σχέση:

$$y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k \quad (3.14)$$

όπου κ είναι ο αριθμός των ανεξάρτητων μεταβλητών.

Κάθε ένας από τους παλινδρομικούς συντελεστές:

$$b_1, b_2, b_3, \dots, b_k$$

εκφράζει τη μεταβολή της εξαρτημένης μεταβλητής γ για μια μονάδα αύξησης στη μεταβλητή  $X_i$ , όταν όλες οι υπόλοιπες ανεξάρτητες μεταβλητές παραμένουν σταθερές, ή για να το πούμε πιο αυστηρά, όταν οι γραμμικοί συνδυασμοί όλων των υπόλοιπων ανεξάρτητων μεταβλητών δεν επιδρούν και άρα δεν λαμβάνονται υπόψη.

Πως όμως να προβούμε σε μια ανάλογη ερμηνεία των συντελεστών στη λογαριθμική παλινδρόμηση, όταν δηλ. έχουμε ένα μοντέλο που υπακούει στην εξίσωση (3.13) και στο οποίο την ανεξάρτητη μεταβλητή  $z$  δεν την ξέρουμε; Η απάντηση σε αυτό το ερώτημα είναι να πιάσουμε το μοντέλο μας από το κατάλληλο σημείο. Να βρούμε δηλ. μια μορφή του η οποία να είναι γραμμικής μορφής, με τους ίδιους ανεξάρτητους παράγοντες.

Ας εστιάσουμε την προσοχή μας στην εξίσωση (3.9). Η εξίσωση (3.9) ολοφάνερα είναι εξίσωση του λογαριθμικού μας μοντέλου και μας βιολεύει ιδιαίτερα στην ερμηνεία του καθώς το δεύτερο μέλος της είναι γραμμικής μορφής, και άρα ερμηνεύσιμο. Ας γράψουμε όμως την εξίσωση αυτή για την περίπτωση που έχουμε πολλές ανεξάρτητες μεταβλητές. Θα έχουμε:

$$\ln\left(\frac{P}{1-P}\right) = b_0 + b_1 X_1 + \dots + b_k X_k \quad (3.15)$$

Αφού το β' μέλος της σχέσης (3.15) είναι γραμμικό, μπορούμε να πούμε ότι Κάθε ένας από τους παλινδρομικούς συντελεστές

$$b_1, b_2, b_3, \dots, b_k$$

εκφράζει τη μεταβολή της εξαρτημένης μεταβλητής  $\ln\left(\frac{P}{1-P}\right)$ , που δεν είναι τίποτε άλλο από το φυσικό λογάριθμο της πιθανότητας γεγονότος, για μια μονάδα αύξησης της

αντίστοιχης ανεξάρτητης μεταβλητής  $X_i$ , και όταν οι υπόλοιπες ανεξάρτητες μεταβλητές είναι σταθερές.

Τη σχέση (3.15) ωστόσο, θα πρέπει να τη φέρουμε σε μια πιο ‘καλυτερότερη’, ‘ερμηνεύσιμη’ ας πούμε μορφή, με αντιλογαρίθμιση των μελών της, όπως κάναμε και παραπάνω. Έτσι η (3.15) γίνεται:

$$e^{\ln\left(\frac{p}{1-p}\right)} = e^{b_0+b_1X_1+\dots+b_kX_k} = e^{b_0} \cdot e^{b_1X_1} \cdots e^{b_kX_k}$$

η οποία καταλήγει στη σπουδαία σχέση:

$$\frac{p}{1-p} = e^{b_0} \cdot e^{b_1X_1} \cdots e^{b_kX_k} \quad (3.16)$$

Ας αυξήσουμε τώρα την ανεξάρτητη μεταβλητή  $X_i$  κατά μία μονάδα, δηλ. κατά 1. Τότε έχουμε:

$$e^{b_0} \cdot e^{b_1X_1} \cdots e^{b_i(X_i+1)} \cdot e^{b_kX_k} = e^{b_i} \cdot [e^{b_0} \cdot e^{b_1X_1} \cdots e^{b_kX_k}] = e^{b_i} \cdot \left[ \frac{p}{1-p} \right] \quad (3.17)$$

Η τελευταία αυτή σχέση (3.17), μας λέει ότι, η μεταβολή κατά μία μονάδα της ανεξάρτητης μεταβλητής  $X_i$  είναι πολλαπλάσιο της σχετικής πιθανότητας  $\frac{p}{1-p}$ .

Διακρίνομε τρεις περιπτώσεις:

I. Αν το  $b_i$  είναι θετικό, τότε το  $e^{b_i}$  είναι αριθμός μεγαλύτερος της μονάδας, που σημαίνει ότι περισσότερες από μια φορές είναι η μεταβολή της σχετικής πιθανότητας του επίμαχου γεγονότος, δηλ. αυξάνει.

II. Αν το  $b_i$  είναι αρνητικό, τότε το  $e^{b_i}$  είναι αριθμός μικρότερος της μονάδας, που σημαίνει ότι λιγότερες από μια φορές είναι η μεταβολή της σχετικής πιθανότητας του επίμαχου γεγονότος, δηλ. μειώνεται.

III. Αν το  $b_i$  είναι μηδέν, τότε το  $e^{b_i}$  είναι αριθμός ίσος με τη μονάδα, δηλ. δεν υπάρχει μεταβολή για τη σχετική πιθανότητα του επίμαχου γεγονότος. Είναι αμετάβλητη.

Τώρα αν στο μοντέλο μας έχουμε το φύλο (GENDER), σαν ανεξάρτητη μεταβλητή, και έχουμε κωδικοποιήσει με 1 τους άνδρες και με 0 τις γυναίκες, θα ήταν ενδιαφέρον να δούμε πόσο μεγάλη ή μικρότερη είναι η πιθανότητα πραγματοποίησης του επίμαχου γεγονότος στους άνδρες, αναφορικά με τις γυναίκες.

Ανακαλούμε στη μνήμη τις σχέσεις (3.3) και (3.4) οι οποίες αφορούν τις σχετικές πιθανότητες πραγματοποίησης του επίμαχου γεγονότος στους άνδρες και στις γυναίκες αντίστοιχα. Να εξετάσουμε αυτές οι σχετικές πιθανότητες με πόσο ισούται εκάστη. Ας ξαναγυρίσουμε σε ένα απλό λογαριθμικό μοντέλο, με μια μόνο ανεξάρτητη μεταβλητή, την GENDER. Χωρίς βλάβη της γενικότητας, θα έχουμε:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 \cdot \text{GENDER} \quad (3.18)$$

Η διαφορετικά, με τη γνωστή πλέον αντιλογαρίθμηση:

$$\frac{p}{1-p} = e^{b_0 + b_1 \cdot \text{GENDER}} \quad (3.19)$$

Για τη σχετική πιθανότητα των ανδρών, αφού γι' αυτούς GENDER=1 η σχέση (3.19) δίνει:

$$\frac{p_a}{1-p_a} = e^{b_0 + b_1 \cdot 1} = e^{b_0 + b_1} \quad (3.20)$$

Ενώ για σχετική πιθανότητα των γυναικών αφού γι' αυτές GENDER=0 η σχέση (1.46) δίνει:

$$\frac{p_y}{1-p_y} = e^{b_0 + b_1 \cdot 0} = e^{b_0} \quad (3.21)$$

Αν διαιρέσουμε κατά μέλη τις (1.47) και (1.48) λαμβάνουμε:

$$\frac{\frac{p_a}{1-p_a}}{\frac{p_y}{1-p_y}} = \frac{e^{b_0 + b_1}}{e^{b_0}} = e^{b_0 + b_1 - b_0} = e^{b_1} \quad (3.22)$$

Η τελευταία σχέση (3.22) μας πληροφορεί ότι αν σε ένα λογαριθμικό μοντέλο έχουμε μια ανεξάρτητη διχοτομική μεταβλητή με παλινδρομικό συντελεστή  $b_i$  τότε ο αντιλογάριθμος του συντελεστή της, που συμβολίζεται με:

$$e^{b_i}$$

είναι ο λόγος των σχετικών πιθανοτήτων του επίμαχου γεγονότος γι' αυτές τις δύο κατηγορίες της διχοτομικής μεταβλητής, όταν οι υπόλοιπες ανεξάρτητες μεταβλητές υποθέσουμε ότι είναι σταθερές. Στην SPSS ορολογία συμβολικά έχουμε:

$$e^{b_i} = \exp(b_i) = \exp(B) \quad (3.23)$$

Συνοψίζοντας όλα τα παραπάνω, μπορούμε να καταλήξουμε σε γενικευμένες εκφράσεις γύρω από το μοντέλο δυαδικής λογαριθμογραφικής παλινδρόμησης.

### 3.4. Γενικευμένες εκφράσεις για το BLR

Το μοντέλο BLR μπορεί να γενικευθεί άμεσα στην περίπτωση που εμείς έχουμε διάφορες προβλέπουσες μεταβλητές. Η πιθανότητα πρόβλεψης  $\pi$ , μπορεί τότε να μοντελοποιηθεί ως:

$$\pi = \Pr(Y = 1 | X_1 = x_1, \dots, X_p = x_p) = \frac{e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p}} \quad (3.24)$$

Η εξίσωση (3.24) ονομάζεται logistic regression function, η οποία είναι μη γραμμική, ως προς τις παραμέτρους  $b_0, b_1, \dots, b_p$ , μα μπορεί να γραφεί ως τέτοια, με τη βοήθεια του λεγόμενου logit μετασχηματισμού ( Chatterjee, Hardi, 2012). Ετσι, για την πιθανότητα  $\pi$  του γεγονότος ενδιαφέροντος, ο λόγος  $\pi/(1-\pi)$  ονομάζεται, όπως ξαναείπαμε odds ratio for event :

$$1 - \pi = \Pr(Y = 0 | X_1 = x_1, \dots, X_p = x_p) = \frac{1}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p}} \quad (3.25)$$

Λαμβάνοντας, στην (3.25) τους νεπέριους λογαριθμους αμφοτέρων των μελών έχουμε:

$$g(x_1, \dots, x_p) = \ln\left(\frac{\pi}{1 - \pi}\right) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p \quad (3.26)$$

## **Κεφάλαιο 4: Το Πολυωνυμικό Λογαριθμογραμμικό Μοντέλο (Multinomial Logit Model, MNLM)**

### **4.1. Σύντομη Ιστορική αναδρομή**

Όταν μια μεταβλητή είναι τύπου nominal, τότε οι κατηγορίες αυτής δεν είναι δυνατόν να διαταχθούν. Ωστόσο, σε κάθε περιοχή των Κοινωνικών συναντάμε ονομαστικές εκβάσεις (nominal outcomes). Οι Schmidt και Strauss (1975), εξέτασαν την επαγγελματική επιτυχία στο πλαίσιο μιας πρώιμης εφαρμογής της MNML διαδικασίας. Οι Meng και Miller (1995), εξέτασαν διαφορές λόγω φύλου στα επαγγέλματα στην Κίνα. Οι Arum και Shavit (1995), μελέτησαν επιδράσεις της τριτοβάθμιας Επαγγελματικής Εκπαίδευσης στο πεδίο της επαγγελματικής επιτυχίας. Παραδείγματα επίσης συναντάμε και σε άλλα πεδία έρευνας. Οι Hoffman και Duncan (1988), συνέκριναν υπό συνθήκη λογιστικά (conditional logit) και πολυωνυμικά λογιστικά μοντέλα (multinomial logit models), στο πλαίσιο μιας μελέτης για τη συζυγική κατάσταση και την ευημερία. Επίσης οι Spector και Mazzeo (1980), εξέτασαν τις επιδράσεις ενός προγράμματος πειραματικής διδασκαλίας στην απόδοση μια τάξης μαθητών όπως αυτή αναδείχθηκε από τους βαθμούς. Άλλα παραδείγματα περιλαμβάνουν λόγους παραμονής των νέων στην πατρική εστία (Goldscheider & DaVanzo, 1989), το οργανωμένο πλαίσιο της επιστημονικής παραγωγικότητας (Long & McGinnis, 1981), και τέλος την επιλογή της γλώσσας μέσα σε μια πολυγλωσσική κοινωνία (Stevens, 1992).

Μια ειδική περίπτωση του MNLM παρουσιάστηκε από τον Gurland και τους συνεργάτες του (Gurland e al., 1960). Το μοντέλο MNLM με τη μορφή που ευρύτατα χρησιμοποιήθηκε στις Κοινωνικές Επιστήμες εισήχθη από τον Theil (1969, 1970). McFadden (1973), του οποίου η παραγωγή σχετίστηκε με τη δουλειά των πειραματικών ψυχολόγων, όπως για παράδειγμα ο Luce (1959). Οι Aitchison και Silvey (1957), αλλά και οι Atchison Bennett (1970), ήταν οι πρώτοι που παρουσίασαν το πολυωνυμικό λογαριθμογραμμικό μοντέλο (multinomial probit model), του οποίου οι απαρχές μπορούν να ανιχνευθούν στην εργασία του Thurstone (1927). Το μοντέλο παρείχθη

από τις παραδοχές της λογικής θεωρίας επιλογής του McFadden (1973). Οι Nerlove και Press (1973) δημοσίευσαν μια σύντομη μονογραφία η οποία εξελίχθηκε σε μια σπουδαία συμβολή με τη συμπερίληψη προγραμμάτων Fortran, στην εκτίμηση του πολυωνυμικού λογαριθμογραμμικού μοντέλου.

#### 4.2. Η θεμελίωση του πολυωνυμικού μοντέλου (Multinomial Logit Model, MNLM)

Στο προηγούμενο Κεφάλαιο 4, είχαμε αναπτύξει λογαριθμογραμμικά μοντέλα όπου η εξαρτημένη μεταβλητή  $Y$  περιλάμβανε 2 μόνον κατηγορίες (Binomial Logistic Regression). Τα μοντέλα αυτά μας επέτρεπαν αναλύσεις διχοτομικών κατηγορικών δεδομένων, για παράδειγμα σε Πίνακες Συνάφειας 2x2. Ωστόσο, στη στατιστική ανάλυση, υπάρχει η ανάγκη, να αναλυθούν πίνακες μεγαλύτερων διαστάσεων της μορφής  $I \times J$ . Αυτό πάει παραπέρα την παλινδρομή ανάλυση. Έτσι, γεννήθηκε ως γενίκευση της διωνυμικής, η πολυτομική ή πολυωνυμική λογαριθμογραμμική παλινδρόμηση (Polytomous or Multinomial Logistic Regression), η οποία έχει τον ίδιο στατιστικό/επιστημονικό σκοπό: Να αναπτύξει μοντέλα στα οποία η μεταβλητή έκβασης (outcome variable), η μεταβλητή στόχος (target variable), να σχετίζεται κατά τέτοιο τρόπο με τις προβλέπουσες μεταβλητές, ώστε η πιθανότητα πρόβλεψης (probability), να πέφτει, να κατανέμεται σε κάθε επίπεδο της μεταβλητής απόκρισης. Τα τελευταία αυτά μοντέλα, που καλύπτονται από το διεθνή όρο logits, παράγονται με τον παρακάτω τρόπο.

Ας υποθέσουμε ότι τα επίπεδα της μεταβλητής έκβασης είναι  $J$ . Έστω  $p_j$ , με  $j=1,\dots,J$  ότι είναι η πιθανότητα που πέφτει στο  $j$  επίπεδο της μεταβλητής έκβασης. Ο σκοπός εδώ είναι να κατασκευαστεί ένα μοντέλο για το οποίο η  $p_j$  να συνδεθεί συναρτησιακά, με ένα σύνολο προβλεπουσών μεταβλητών  $X_1, X_2, \dots, X_k$ , ώστε να ισχύει:

$$\sum_j p_j = 1 \quad (4.1)$$

Το μοντέλο, και πάλι βασίζεται σε logits, αλλά η δυσκολία είναι ότι εδώ δεν έχουμε μια απλή κατηγορία κατηγορία ‘επιτυχία’ για να τη χρησιμοποιήσουμε για να

κατασκευάσουμε το Logit. Η λύση είναι (Simonoff, 2003), να κατασκευάσουμε όλα τα Logits που σχετίζονται με ένα από τα επίπεδα της outcome μεταβλητής, το οποίο ονομάζεται βασική κατηγορία αναφοράς (baseline category).

Εάν μια από τις κατηγορίες της outcome μεταβλητής μπορεί να θεωρηθεί ως διαφορετική για κάποιο λόγο από τις άλλες, ή είναι η 'άριστη', τότε είναι φυσικό να την επιλέξουμε ως βασική. Στη δική μας έρευνα που ακολουθεί (βλ. επόμενο Κεφ.5), η μεταβλητή έκβασης του πολιτικού Κόμματος (PARTY), επιλέξαμε ως βασική κατηγορία εκείνη με κωδικό 2=ΣΥΡΙΖΑ, διότι αυτή η κατηγορία αντιστοιχούσε στον ΣΥΡΙΖΑ ο οποίος ήταν, στις εκλογές του Ιανουαρίου του 2015, η ισχυρότερη πολιτική δύναμη, με πολύ υψηλά, απλησίαστα από τις υπόλοιπες πολιτικές δυνάμεις, ποσοστά. Τέλος, αν δεν έχουμε κάποια κατηγορία πολύ διαφορετική από τις άλλες, τότε διαλέγουμε αυθαίρετα μία ως βασική.

Έστω τώρα ότι η βασική μας κατηγορία, η κατηγορία αναφοράς (reference category), όπως λέγεται αλλιώς, είναι η J. Το MNLM μοντέλο ορίζεται τότε από τη σχέση:

$$\ln\left(\frac{p_j}{p_J}\right) = b_{0j} + b_{1j}X_1 + \dots + b_{kj}X_k \quad (4.2)$$

Οπου  $j=1,\dots,J-1$

Και  $\kappa$  = ο αριθμός των προβλεπτουσών μεταβλητών (predictors).

Το μοντέλο που προσδιορίζεται από τη σχέση (4.2), ενσωματώνει  $J-1$  ξεχωριστές εξισώσεις, κάθε μια από τις οποίες βασίζεται σε ένα σύνολο παραμέτρων B. Προφανώς για την κατηγορία αναφοράς J ισχύει:

$$b_{0j}=B_{1j}=\dots=B_{kj}=0 \quad (4.3)$$

Από τη σχέση (4.2) και έχοντας υπόψη την ιδιότητα των λογαρίθμων που περιγράφεται από τη σχέση:

$$\ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b) \quad (4.4)$$

Κι ακόμη, διαλέγοντας την I κατηγορία ως κατηγορία αναφοράς,

Λαμβάνουμε:

$$\begin{aligned}
 \ln\left(\frac{p_j}{p_I}\right) &= \ln\left(\frac{p_J}{p_I}\right) = \ln\left(\frac{p_j}{p_J}\right) - \ln\left(\frac{p_I}{p_J}\right) = \\
 &\quad \frac{p_j}{p_J} \\
 &= (b_{0J} + b_{1j}X_1 + \dots + b_{kj}X_k) - (b_{0I} + b_{1I}X_1 + \dots + b_{kI}X_k) = \\
 &= (b_{0J} - b_{0I}) + (b_{1j} - b_{1I})X_1 + \dots + (b_{kj} - b_{kI})X_k
 \end{aligned} \tag{4.5}$$

Η σχέση (4.5), είναι αποκαλυπτική: Οι Logit συντελεστές για το επίπεδο j σχετικά με την κατηγορία αναφοράς I, είναι η διαφορά ανάμεσα στους συντελεστές για το επίπεδο j σχετικά με την κατηγορία αναφοράς J, και τους συντελεστές για το επίπεδο I σχετικά με την κατηγορία αναφοράς J. Να σημειώσουμε ότι δεν υπάρχει τίποτε, μέσα σε αυτή την παραγοντική διαδικασία που να απαιτεί η κατηγορία I να είναι η βασική κατηγορία, αφού η εξίσωση (4.5) εφαρμόζεται και ισχύει, για οποιοδήποτε ζεύγος κατηγοριών.

Το παραπάνω αποτέλεσμα μπορεί να είναι χρήσιμο εάν τα odds για ένα συγκεκριμένο ζεύγος επιπέδων είναι το ζεύγος ενδιαφέροντος. Επειδή ακριβώς η επιλογή της βασικής κατηγορία είναι αυθαίρετη, οι εκτιμώμενοι συντελεστές (και ίσως τα τυπικά σφάλματα αυτών) να μπορούν να λαμβάνονται εύκολα χρησιμοποιώντας οποιοδήποτε λογισμικό, αναφορικά με τα MNLM μοντέλα, αλλάζοντας τη βασική κατηγορία με ένα από τα δύο επίπεδα, με την εγγύηση ότι οι εκτιμώμενες πιθανότητες που πέφτουν σε κάθε ένα από τα επίπεδα δεν θα αλλάξουν.

Το μοντέλο που περιγράφει η σχέση (4.2), υπαινίσσεται έναν απλό συναρτησιακό τύπο γι' αυτές τις πιθανότητες. Ο τύπος αυτός είναι το γνωστό σχήμα S-shape στις λογαριθμογραμμικές σχέσεις:

$$p_j = \frac{\exp(b_{0j} + b_{1j}X_1 + \dots + b_{kj}X_k)}{\sum_{l=1}^J \exp(b_{0l} + b_{1l}X_1 + \dots + b_{kl}X_k)} \tag{4.6}$$

Οι εκτιμήσεις για τις παραμέτρους  $[b_1, \dots, b_{J-1}]$  λαμβάνονται με τη χρήση της μεθόδου μέγιστης πιθανοφάνειας, όπου ο λογάριθμος πιθανότητας είναι:

$$L = \sum_{j=1}^J \sum_{y_i=j} \ln p_{j(i)} \quad (4.7)$$

Στη σχέση (4.7) το δεύτερο άθροισμα διατρέχει όλες τις i παρατηρήσεις με επίπεδο απόκρισης j και  $p_{j(i)}$  είναι η πιθανότητα που δίνεται από την (4.6), υποκαθιστώντας τις προβλέπουσες τιμές για την i παρατήρηση.

### 4.3. Η προσαρμογή του μοντέλου MNLM στα δεδομένα

#### 4.3.1. Μέτρα ψευδο- $R^2$ (pseudo-R-square Measures)

Όταν διαπραγματεύμαστε<sup>7</sup> μοντέλα Γραμμικής Παλινδρόμησης, η  $R^2$  στατιστική αναπαρίστανε το ποσοστό της μεταβλητότητας στην εξαρτημένη μεταβλητή που μπορεί να εξηγηθεί από το σύνολο των εξαρτημένων μεταβλητών. Στο πλαίσιο της Λογαριθμογραμμικής Ανάλυσης, δεν έχουμε ένα ανάλογα καλό μέτρο. Δεν έχουμε δηλ. ένα μέτρο, εύκολα ερμηνεύσιμο, που να μιλήσει για την ένταση της σχέσης ανάμεσα στην ανεξάρτητη και τις εξαρτημένες μεταβλητές.

Ωστόσο, εμείς εδώ, και για λόγους πληρότητας, θα αναφέρουμε τρία μέτρα που υποκαθιστούν το  $R^2$  και βέβαια με το χαρακτηρισμό που τα χαρακτηρίζει διεθνώς.

Cox and Snell  $R^2$  για το οποίο ισχύει:

$$R_{CS}^2 = 1 - \left( \frac{L(B^{(0)})}{L(\hat{B})} \right)^{\frac{2}{n}} \quad (4.8)$$

Nagelkerke's  $R^2$  για το οποίο ισχύει:

$$R_N^2 = \frac{R_{CS}^2}{1 - L(B^{(0)})^{\frac{2}{n}}} \quad (4.9)$$

---

<sup>7</sup> Βλ. βιβλίο: Δαφέρμος, Β. (2011). «Κοινωνική Στατιστική και Μεθοδολογία Έρευνας με το SPSS», Εκδόσεις ZHTH, Θεσσαλονίκη.

McFadden's R<sup>2</sup> για το οποίο ισχύει:

$$R_M^2 = 1 - \left( \frac{L(\hat{B})}{L(B^{(0)})} \right) \quad (4.10)$$

Όπου,

$L(\hat{B})$  είναι η συνάρτηση Ln-likelihood για το μοντέλο με τις εκτιμούμενες παραμέτρους,

$L(B^{(0)})$  είναι ο πυρήνας (kernel) του νεπέρειου λογαρίθμου της πιθανότητας του μοντέλου που περιλαμβάνει μόνο το intercept, και τέλος, η ο αριθμός των cases.

#### 4.3.2. Κριτήρια Καλής Προσαρμογής (Goodness-of-fit Tests)

Όταν κτίζουμε ένα μοντέλο, πάντα, μας ενδιαφέρει πόσο πολύ αυτό προσαρμόζεται στα δεδομένα. Ο έλεγχος γίνεται με τη βοήθεια των στατιστικών κριτηρίων Pearson X<sup>2</sup> και Deviance. Συνήθως, υπολογίζουμε το Pearson X<sup>2</sup> από την παρακάτω σχέση:

$$X_{\text{Pearson}}^2 = \sum_{\text{AllCells}} \frac{(\text{observed\_counts} - \text{expected\_counts})^2}{\text{expected\_counts}} \quad (4.11)$$

Μεγάλες τιμές της στατιστικής Pearson X<sup>2</sup>, φανερώνουν ότι το μοντέλο δεν έχει καλή προσαρμογή, δεν έχει καλό fit. Ισοδύναμα εκφραζόμενοι, θα λέγαμε ότι, χαμηλές τιμές στο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας σημαίνουν όχι καλή προσαρμογή. Φυσικά, όσο μεγαλύτερο είναι το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, τόσο καλύτερη είναι η προσαρμογή του μοντέλου μας στα δεδομένα. Πεμπτουσία αυτής της στατιστικής είναι η εκτίμηση της αντίφασης/διαφοράς, ανάμεσα σε παρατηρούμενες και αναμενόμενες συχνότητες, μέσα σε ένα πολυδιάστατο πίνακα διασταύρωσης (Crosstabulation).

Ενας άλλος τρόπος να μετρήσουμε την Goodness-of-fit είναι η απόκλιση  $\chi^2$ , αυτή που ονομάζεται με το διεθνή όρο Deviance. Πεμπτουσία αυτής της στατιστικής είναι η μεταβολή της ποσότητας  $-2LL = -2\ln\text{-likelihood}$  ανάμεσα στο μοντέλο μας και στο κορεσμένο (saturated) μοντέλο, δηλ. εκείνο που διαθέτει όλες τις κύριες επιδράσεις (main effects) και αλληλεπιδράσεις (interactions). Έτσι, αν το μοντέλο μας ταιριάζει καλά στα δεδομένα, τότε η διαφορά ανάμεσα στους λογαρίθμους των πιθανοτήτων θα πρέπει να είναι μικρή.

Συμβολικά για την Deviance και για πολυωνυμικό (multinomial) μοντέλο έχουμε:

$$\text{Deviance} = G^2 = 2 \sum \text{Observed} \cdot \ln\left(\frac{\text{Observed}}{\text{Expected}}\right) \quad (4.12)$$

Στη σχέση (4.12) το άθροισμα αφορά όλα τα κελιά του πίνακα διασταύρωσης τα οποία δεν είναι δομικά μηδέν και δεν έχουν μηδενικές fitted τιμές.

### Σημαντικές παρατηρήσεις

- I. Τόσον η  $\chi^2$  όσον και η  $G^2$  στατιστική, ασυμπτωτικά ακολουθούν  $\chi^2$  -κατανομή.
- II. Ενα κορεσμένο (saturated) μοντέλο, πάντα παράγει τέλεια προσαρμογή, τέλειο Fit.
- III. Στις αναλύσεις κοινωνικών δεδομένων άλλες φορές χρησιμοποιείται η στατιστική  $X^2$ , και άλλες η στατιστική  $G^2$ .
- IV. Εάν διαθέτουμε μεγάλα δειγματικά μεγέθη οι δύο στατιστικές θα πρέπει να οδηγούν στα ίδια συμπεράσματα, αναφορικά με την καλή προσαρμογή.

### 4.4. Το φαινόμενο της υπερδιασκορπισης (overdispersion)

Η υπερδιασκόρπιση (overdispersion), είναι μια ασθένεια που παρουσιάζεται στα Logistic μοντέλα, κι όχι μόνον, και οφείλεται στο γεγονός ότι η παρατηρούμενη/εμπειρική διασπορά (empirical variation) είναι μεγαλύτερη από την προβλεπόμενη/ονομαστική (nominal variation), του μοντέλου. Αυτό πιθανότατα οφείλεται σε 4 λόγους:

- I. Ο πρώτος είναι η συσχέτιση που παρουσιάζουν οι παρατηρήσεις μεταξύ τους. Ας θυμηθούμε, από την Περιγραφική κιόλας Στατιστική, την παραδοχή της ανεξαρτησίας των παρατηρήσεων (independence). Όταν αυτή παραβιάζεται προφανώς και έχουμε correlation. Οι Cases ‘κολλάνε’ μεταξύ τους.
- II. Ο δεύτερος λόγος είναι η μεγάλη μεταβλητότητα (variability) στις πιθανότητες επιτυχίας (Halekoh, U., & Hojsgaard, S., 2007).
- III. Ο τρίτος είναι η παρουσία μιας ή περισσοτέρων ακραίων παρατηρήσεων, (multivariate outliers).
- IV. Και ο τέταρτος το κακοκτισμένο (miss-building) μοντέλο.

Μιλήσαμε ωστόσο για ασθένεια στα νεφρά του Logistic Model. Εννοούμε με αυτό ότι το overdispersion παρεισφέρει στις δομές του μοντέλου προσπαθώντας να περιορίσει τα Standard errors μέσα σε πιο στενά διαστήματα εμπιστοσύνης. Από την άλλη δεν λησμονούμε ότι, για να υπολογίσουμε τη σπουδαιότητα (Importance) ενός συντελεστή, διαιρούμε την τιμή του, δια του Τυπικού σφάλματος (Standard error). Και βέβαια, όσο μικρότερο είναι το Τυπικό σφάλμα, τόσο η ανεξάρτητη μεταβλητή (predictor) βγαίνει σπουδαιότερη, χωρίς αυτή να είναι στην πραγματικότητα. Τόσο μικρότερο δηλ. βγαίνει το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, που την κάνει με τη σειρά του να φαίνεται σπουδαία. Σε τελευταία ανάλυση το μικρόβιο overdispersion μικράίνει τα τυπικά σφάλματα υποχρεώνοντάς τα να χωρέσουν, σε μικρότερα διαστήματα εμπιστοσύνης. Αυτό το τελευταίο είναι σκέτη καταστροφή. Διότι, στενότερα διαστήματα εμπιστοσύνης θα τροφοδοτήσουν την υπέρ-εμπιστοσύνη (overconfidence). Τα μειωμένα διαστήματα, με τη σειρά τους, θα μας μιλήσουν για μεγάλη επιδραστικότητα των predictors πάνω στη μεταβλητή έκβασης (outcome variable). Και τότε, προφανώς, μπαίνουμε σε μεγαλύτερο κίνδυνο να διαπράξουμε σφάλμα τύπου I.

Ανάγκη λοιπόν, πάσα, να μετρήσουμε την υπερ-διασκόρπιση. Το SPSS εμπέσως πλην σαφώς μετράει, με το συντελεστή  $\Phi$  την υπερ-διασκόρπιση. Είναι ακριβώς ο λόγος της  $X^2$  - goodness στατιστικής δια των βαθμών ελευθερίας αυτού. Οπότε, θα έχουμε:

$$\text{overdispersion\_parameter\_}\Phi = \frac{\chi^2_{\text{goodness\_of\_it}}}{\text{df}} \quad (4.13)$$

Υπάρχει overdispersion αν το  $\Phi$  ξεπερνά το 2. Αντίθετα, αν ο λόγος αυτός είναι μεγαλύτερος του 1 και μικρότερος του 2, δεν ανησυχούμε ιδιαίτερα για υπερδιασκόρπιση.

Αφού η υπερδιασκόρπιση είναι ασθένεια στις αναλύσεις κοινωνικών δεδομένων, προφανώς θα πρέπει να αναζητηθούν μέθοδοι ίασης. Οι McCullagh Nelder (1989), ότι είναι δυνατόν να εκτιμηθούν σταθερές (constants), μέσα από τα Δεδομένα της Έρευνας, τις οποίες εμείς μπορούμε να χρησιμοποιήσουμε στη συνέχεια για να 'διορθώσουμε' τη μήτρα διασποράς- συνδιασποράς των παραμέτρων εκτίμησης.

#### 4.5. Τα Πρότυπα Συνδιασποράς (Covariate Patterns)

Τα πρότυπα συνδιασποράς είναι ένα ιδιαίτερα σπουδαίο ζήτημα στην Στατιστική Ανάλυση των Λογαριθμογραμμικών μοντέλων.

Οι στατιστικές καλής προσαρμογής βασίζονται στη σύγκριση παρατηρούμενων και αναμενόμενων τιμών μέσα σε κάθε κελί του Πίνακα Διασταύρωσης (Crossstabulation Table). Ο αριθμός των κελιών είναι το γινόμενο του αριθμού των επιπέδων (levels) για τη μεταβλητή ομαδοποίησης (grouping variable), επί των αριθμό όλων των παραγόντων (factors), και συνδιασπορών (covariates). Πρότυπο Συνδιασποράς (Covariate Pattern) ονομάζεται κάθε διακριτός συνδυασμός τιμών των ανεξάρτητων μεταβλητών (predictors).

Τα στατιστικά τεστ δεν μπορούν να εφαρμοστούν παρά μόνον εάν υπάρχουν πολλαπλές παρατηρούμενες cases για κάθε ένα πρότυπο συνδιασποράς, για κάθε ένα pattern. Εάν διαθέτουμε πολλές cases που έχουν ένα και μόνο covariate pattern, κάτι που πολύ συχνά συμβαίνει όταν δεν έχουμε κατηγορικές ανεξάρτητες, αλλά συνεχείς μεταβλητές, τότε τα στατιστικά τεστ καλής προσαρμογής, δεν μπορούν να εφαρμοστούν, διότι αυτά τα τεστ δεν ακολουθούν  $\chi^2$  κατανομή. Αυτό οφείλεται στο

γεγονός ότι οι αναμενόμενες τιμές στα κελιά είναι τότε πολύ μικρές (Hosmer and Lemeshow, 2000).

Χάριν πληρότητας παραθέσουμε μια σχέση η οποία συνδέει τα covariate patterns με τους με τους βαθμούς ελευθερίας, τα ανεξάρτητα Logits και το συνολικό αριθμό των παραμέτρων που θέλουμε να εκτιμήσουμε, συμπεριλαμβανομένου και του σταθερού όρου.

Οι βαθμοί ελευθερίας σε ένα συγκεκριμένο μοντέλο πρόβλεψης είναι αριθμός ίσος με το γινόμενο των Covariate patterns, επί τον αριθμό των ανεξάρτητων Logits, μείον τον αριθμό των παραμέτρων που έχουμε να εκτιμήσουμε συμπεριλαμβανομένου του σταθερού όρου (intercept). Συμβολικά:

$$\text{DF} = (\text{NUMBER\_OF\_OBSERVED\_COVARIATE\_PATTERNS})^* \\ (\text{NUMBER\_OF\_INDEPENDENT\_LOGITS}) - \text{NUMBER\_ESTIMATED\_PARAMETERS}$$

#### 4.6. Οι παραδοχές για το MNLM μοντέλο

Για να μπορέσει ένα πολυωνυμικό μοντέλο να παραγάγει φερέγγυα αποτελέσματα και ασφαλείς προβλέψεις, θα πρέπει να συμμορφώνεται με τις εξής παραδοχές:

- I. Τυχαίο Δείγμα. Όχι με την αυστηρή μαθηματική/στατιστική έννοια του όρου. Σε κάποιες στροφές του ερευνητικού δρόμου, η Δειγματοληψία, μπορεί να είναι και επιλεκτική. Σημασία έχουν οι όροι κάτω από τους οποίους γίνεται η έρευνα. Θα πρέπει να είναι συμβατοί με αυτήν.
- II. Ανεξαρτησία των odds ratio. Για οποιεσδήποτε δύο κατηγορίες της εξαρτημένης μεταβλητής (respondent variable), τα odds ratio να είναι ανεξάρτητα, που σημαίνει μη γραμμικά εξαρτώμενα από εκείνα όλων των άλλων απαντητικών κατηγοριών. Για παράδειγμα, αν ένα νέο Κόμμα εισάγεται στο πολιτικό φάσμα, αυτή η παραδοχή σημαίνει ότι τα ποσοστά όλων των άλλων Κομμάτων επηρεάζονται εξ ίσου ανάλογα. Η παραδοχή συνδυάζεται με

τον αριθμό των covariate patterns<sup>8</sup>. [Βλ. επίσημες οδηγίες στο manual του IBM SPSS 28 και πιο συγκεκριμένα στο pdf αρχείο με τίτλο IBM SPSS regression.pdf].

- III. Δεδομένου ενός covariate pattern<sup>9</sup>, οι απαντήσεις θεωρούνται/υποτίθενται ότι είναι ανεξάρτητες πολυωνυμικές μεταβλητές.
- IV. Φειδωλός αριθμός συνεχών μεταβλητών στο μοντέλο. Σε περίπτωση ύπαρξης μεγάλου αριθμού συνεχών μεταβλητών ίσως παρουσιαστεί πρόβλημα αδυναμίας σύγκλισης του μοντέλου. Ίσως δηλ. να μην έχουμε λύση.
- V. Φειδωλός αριθμός κατηγορικών μεταβλητών στο μοντέλο. Σε περίπτωση ύπαρξης μεγάλου αριθμού κατηγορικών μεταβλητών ίσως παρουσιαστεί πρόβλημα με τον αριθμό των κενών κελιών, των κελιών δηλ. στον πίνακα διασταύρωσης που έχουν μηδενικές συχνότητες. Τελικές αρνητικές συνέπειες: αδυναμία σύγκλισης του μοντέλου και κακή προσαρμογή.
- VI. Φειδωλός αριθμός κατηγοριών (levels) στις κατηγορικές μεταβλητές του μοντέλου. Όμοια και εδώ. Σε περίπτωση ύπαρξης μεγάλου αριθμού levels, ίσως παρουσιαστεί πρόβλημα υπερβολικού αριθμού κενών κελιών, κελιών δηλ. στον πίνακα διασταύρωσης που έχουν μηδενικές συχνότητες (zero frequencies). Τελικές συνέπειες: αδυναμία σύγκλισης του μοντέλου και κακή προσαρμογή.
- VII. Καλή Προσαρμογή του μοντέλου στα δεδομένα (Goodness-of-fit). Ιδιαίτερα σημαντική και βασικότατη παραδοχή. Μοντέλα χωρίς καλή προσαρμογή δεν μπορούν να παραγάγουν φερέγγυα συμπεράσματα. Δεν μπορούν να κάνουν

---

<sup>8</sup> Θα λέμε covariate pattern κάθε διακριτό συνδυασμό τιμών των ανεξάρτητων/ερμηνευτικών μεταβλητών (predictors variables). Αν τώρα υποθέσουμε ότι έχουμε τις περισσότερες cases να διαθέτουν εντός του δείγματος μοναδικά covariate patterns, κάτι που βλέπουμε συχνά όταν οι ανεξάρτητες μεταβλητές δεν είναι κατηγορικές, τότε τα τεστ καλής προσαρμογής δεν μπορούν να εφαρμοστούν. Και δεν μπορούν να εφαρμοστούν διότι απλά δεν ακολουθούν τη χι τετράγωνο κατανομή. Και δεν ακολουθούν τη χι τετράγωνο κατανομή, επειδή ακριβώς οι αναμενόμενες συχνότητες στα κελιά είναι μικρές ( Hosmer and Lemeshow, 2000).

<sup>9</sup> Οι στατιστικές καλής προσαρμογής, λέει η Norusis (2010), θα πρέπει να χρησιμοποιούνται μόνον όταν υπάρχουν πολλαπλάσιες cases σε κάθε covariate pattern, συμπληρώνοντας/ενισχύοντας τα όσα λέει το manual του IBM SPSS 28, στο οποίο προαναφερθήκαμε.

σωστή στατιστική πρόβλεψη. Οι δείκτες Pearson  $\chi^2$  και DEVIANCE, στο πλαίσιο του SPSS ελέγχουν την προσαρμογή του μοντέλου στα δεδομένα (fit), αν βέβαια οι δείκτες αυτοί, ή ακριβέστερα αυτά τα τεστ ακολουθούν την  $\chi^2$  κατανομή. Όσο μεγαλύτερο είναι το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας<sup>10</sup>, τόσο η προσαρμογή του μοντέλου στα δεδομένα είναι καλύτερη. Επίσης να σημειώσουμε εδώ ότι στις περιπτώσεις που έχουμε μεγάλο δείγμα, τα δύο τεστ θα πρέπει να έχουν παρόμοια αποτελέσματα (Norusis, 2010).

- VIII. Γραμμικότητα<sup>11</sup> (Linearity). Η σχέση ανάμεσα στα Logits που σχηματίζει η εξαρτημένη με τις ανεξάρτητες μεταβλητές θα πρέπει να είναι γραμμική. Προφανώς ο έλεγχος γραμμικότητας δεν έχει νόημα στις περιπτώσεις που έχουμε διχοτομικές μεταβλητές ή κατηγορικές μεταβλητές οι οποίες έχουν μετασχηματισθεί πριν την ανάλυση.
- IX. Ισχυρή προβλεπτική δύναμη<sup>12</sup> του μοντέλου (Prediction Power). Αυτό σημαίνει ότι το μοντέλο μας θα πρέπει να είναι σε θέση να ταξινομεί σωστά. Με άλλα λόγια θα πρέπει, να είναι σε θέση να εκχωρεί μεγάλες πιθανότητες σε παρατηρούμενες τιμές (likelihoods).
- X. Απουσία συγγραμμικότητας (Absence of Collinearity), αν βέβαια έχουμε συνεχείς μεταβλητές στο μοντέλο.

---

<sup>10</sup> Η ποσότητα -2Log-Likelihood που αντιστοιχεί μοντέλο μας συγκρινόμενη με την ποσότητα που αντιστοιχεί στο κορεσμένο μοντέλο (saturated model), δεν θα πρέπει να έχουν μεγάλη διαφορά για να γίνει λόγος για καλή προσαρμογή. Να διευκρινίσουμε άμεσα ότι κορεσμένο θεωρείται το μοντέλο που περιλαμβάνει όλες τις κύριες επιδράσεις (main effects), και όλες τις αλληλεπιδράσεις (interactions effects).

<sup>11</sup> Οι κορυφαίοι ερευνητές Long and Freese (2014, page 192), αποδεικνύουν ότι ο έλεγχος της παραδοχής της Γραμμικότητας δεν έχει πρακτική σπουδαιότητα (practical importance).

<sup>12</sup> Να θυμίσουμε εδώ ότι στις περιπτώσεις των Multinomial Logistic Regression Μοντέλων δεν είναι δυνατή η ακριβής εκτίμηση των αναμενόμενων τιμών (expected values). Αυτό, στις περιπτώσεις που έχουμε ανισομεγέθη groups, έχει σαν συνέπεια οι cases να πάνε να ταξινομηθούν στα μεγαλύτερους μεγέθους groups, χωρίς να επηρεάζονται από το γεγονός ότι το μοντέλο έχει καλό fit. Με άλλα λόγια, ο πίνακας που ελέγχει την προβλεπτική/ταξινομική ικανότητα, ο λεγόμενος Classification Table, να μην έχει σπουδαίες πληροφορίες ούτε για την προβλεπτική δύναμη του μοντέλου, ούτε για την προσαρμογή του στα δεδομένα (Norusis, 2010).

- XI. Απουσία Υπερδιασκόρπισης (Absence of overdispersion). Η εμπειρική διασπορά δεν θα πρέπει να ξεπερνά την προβλεπόμενη από το μοντέλο. Αν συμβαίνει κάτι τέτοιο, το μοντέλο μας, δεν θα έχει καλή προσαρμογή.
- XII. Απουσία Τέλειας Πρόβλεψης (Absence Perfect Prediction). Αν, για παράδειγμα, όλες οι γυναίκες ψηφίζουν ένα συγκεκριμένο Κόμμα τότε το μοντέλο μας δεν μπορεί να κάνει σωστή πρόβλεψη. Είναι ανήμπορο. Με απλά λόγια, δεν θα πρέπει ένα οποιοδήποτε επίπεδο, μιας οποιασδήποτε ανεξάρτητης μεταβλητής, να σχετίζεται με μόνο ένα επίπεδο της εξαρτημένης μεταβλητής.
- XIII. Απουσία πολυδιάστατων ακραίων τιμών (absence multivariate outliers). Ο Tarling (2009), επισημαίνει ότι η ανίχνευση και ο εντοπισμός multivariate outliers είναι ενίστε, δύσκολος<sup>13</sup>.
- XIV. Μέγεθος Δείγματος. Τα MNLM μοντέλα χρησιμοποιούν διαδικασία μεγίστης πιθανοφάνειας (ML). Έτσι, οι Long & Freese, 2014, σελ. 85 ισχυρίζονται: (α) δείγμα με λιγότερες από 100 cases θεωρείται παρακινδυνευμένο, ενώ δείγμα μεγαλύτερο του 500 θεωρείται επαρκές. (β) Ανάγκη πάσα 10 τουλάχιστον cases να αντιστοιχούν σε κάθε παράμετρο. Διαφορετικά, το μοντέλο δεν συγκλίνει, δεν οδηγεί σε λύση. (γ) Εάν τα δεδομένα μας είναι ill-conditioned, δηλ. οι ανεξάρτητες μεταβλητές παρουσιάζουν την ασθένεια της πολυσυγγραμμικότητας (multi-collinearity), θα χρειαστούμε μεγαλύτερο δείγμα. (δ) Αν τέλος, η εξαρτημένη μεταβλητή παρουσιάζει μικρή διασκόρπιση (variation), για παράδειγμα όλες σχεδόν οι εκβάσεις (outcomes) είναι 1, τότε επίσης θα χρειαστούμε μεγαλύτερο δείγμα.

---

<sup>13</sup> Ας μας επιτραπεί να παρατηρήσουμε ότι αν διαθέτεις στατιστικά εργαλεία λογισμικού, όπως το STATA 14 ή νεότερο, αυτή η δουλειά είναι μάλλον εύκολη. Επίσης, με τα ίδια εργαλεία εύκολη δουλειά είναι και η διάγνωση της Γραμμικότητας (Linearity). Η ερευνητική δουλειά με τέτοιου ειδους εργαλεία γίνεται και ευχάριστη. Δεν κουράζεσαι και δεν στεναχωριέσαι.

#### **4.7. Η έννοια της αλληλεπίδρασης στο πλαίσιο των MNLM μοντέλων**

Η έννοια αυτή αναπτύσσεται διεξοδικά, μετά παραδείγματος, στην παράγραφο 5.11.

#### **4.8. Οι στατιστικοί δείκτες AIC και BIC**

Οι στατιστικοί δείκτες AIC και BIC αναπτύσσονται με σαφήνεια επίσης αναλυτικά σε αρκετά σημεία του Κεφαλαίου 5.

## Κεφάλαιο 5: Η Στατιστική Ανάλυση των Δεδομένων της Έρευνας.

### Ανάπτυξη των Στατιστικών Μοντέλων Πρόβλεψης

**5.1. Συμβολισμός, Κωδικοποίηση και Επεξήγηση των μεταβλητών που υπεισέρχονται στην έρευνα μας.**

Στον Πίνακα 5.1 παρατίθεται το ερωτηματολόγιο της Έρευνας

**Πίνακας 5.1. Ερωτηματολόγιο της Έρευνας.**

<b>Q1.</b> Το πολιτικό Κόμμα που θα ψηφίσετε σε αυτές τις εκλογές, όποιο κι αν είναι αυτό, το ψηφίζετε για:		
Q11: Ιδεολογικούς λόγους;	1 = NAI	2 = OXI
Q12: Οικονομικούς λόγους (ανεργία/οικονομική ανέχεια κτλ);	1 = NAI	2 = OXI
Q13: Να σωθεί η Πατρίδα;	1 = NAI	2 = OXI
Q14: Λόγους προσωπικών υποχρεώσεων;	1 = NAI	2 = OXI
Q15: Γιατί είναι αντιμημονιακό Κόμμα;	1 = NAI	2 = OXI
Q16: Από αντίδραση στο Σύστημα;	1 = NAI	2 = OXI
Q17: Για άλλο λόγο ποιό; .....	1 = NAI	2 = OXI
<b>Q2.</b> Τι είναι ποιο σημαντικό στην απόφασή σας για το Κόμμα αυτό; (μια απάντηση παρακαλώ):		
1 = Ο αριθμός του κόμματος;		
2 = Οι υποψήφιοι;		
3 = Η γενική κοσμοθεωρία/ιδεολογία του κόμματος;		
4 = Το πρόγραμμα και οι θέσεις;		
5 = Άλλο (ποιο;... ..)		
<b>PARTY:</b> Ποιο Κόμμα είναι πιο πιθανό να ψηφίσετε την επόμενη Κυριακή;		
1 = ΝΔ		
2 = ΣΥΡΙΖΑ		
3 = ΠΛΑΣΟΚ		
4 = ΚΙΔΗΣΟ του Γ.Παπανδρέου		
5 = ΠΟΤΑΜΙ		
6 = ΚΚΕ		
7 = ΧΡΥΣΗ ΑΥΓΗ		
8 = ΑΝΕΞ. ΕΛΛΗΝΕΣ		
9 = ΠΡΑΣΙΝΟΙ-ΔΗΜΑΡ		
10 = ΚΚΕ (μ-λ) και Μ-Λ ΚΚΕ		
11 = ΤΕΛΕΙΑ του Απ. Γιλέτσου		
12 = ΑΝΤΑΡΣΙΑ		
13 = ΛΑΟΣ		
14 = ΆΛΛΟ ποιο; .....		
15 = ΑΔΙΕΥΚΡΙΝΙΣΤΗ ΨΗΦΟΣ = ΛΕΥΚΟ/ΑΚΥΡΟ/ΚΑΝΕΝΑ/ΔΕΝ ΘΑ ΨΗΦΙΣΩ		
<b>FYLO:</b>		
1 = Άνδρας, 2 = Γυναίκα		
<b>CAT_AGE<sup>14</sup>:</b>		
1 = 18-34 ετών, 2 = 35-54 ετών, 3 = 55 και άνω		
<b>EDUC = Μορφωτικό επίπεδο:</b>		
1 = Στοιχειώδης Εκπαίδευση, 2 = Μέση, 3 = Ανώτερη/Ανώτατη		

14

<sup>14</sup> Οικοδομήσαμε, σκόπιμα, την κατηγορική μεταβλητή των ηλικιακών ομάδων, CAT\_AGE, και μάλιστα με 3 κατηγορίες: 1=17-34 ετών, 2=35-54 ετών και 3=55 και άνω καθώς ήταν στόχος μας να εξετάσουμε τη διαφορετική, όπως εικάζει η θεωρία μας, συμπεριφορά αυτών των Ομάδων. Πράγματι, όπως έδειξε η μετέπειτα ανάλυση, η μεταβλητή CAT\_AGE αναδείχθηκε στατιστικώς σημαντική σε όλα τα προβλεπτικά μοντέλα που συμπεριλήφθηκε, και άμεσο αποτέλεσμα αυτού το γεγονότος υπήρξε η διαφορετική συμπεριφορά των διάφορων ηλικιακών κατηγοριών.

**5.2. Το αρχείο (Data Set) της Έρευνας. Η συσχέτιση της μεταβλητής έκβασης PARTY με τη μεταβλητή CAT\_AGE των ηλικιακών κατηγοριών.**

Οι Πίνακες 5.2 και 5.3 ενέχουν σημαντικές στατιστικές πληροφορίες. Δείχνουν με σαφή τρόπο που υπάρχει ισχυρή συσχέτιση ανάμεσα στη μεταβλητή έκβασης PARTY και στην προβλέπουσα μεταβλητή των ηλικιακών κατηγοριών CAT\_AGE που εμείς κατασκευάσαμε. Όπως λέει ο Abu-Bader (2010), αν σχέση μεταξύ κάποιας προβλέπουσας μεταβλητής (predictor variable), και της μεταβλητής έκβασης είναι στατιστικώς σημαντική, τότε αυτό αποτελεί οιωνό άριστο, στη λειτουργία που προβλεπτικού μοντέλου. Πράγματι, στον Πίνακα 5.3 και τα 3 στατιστικά κριτήρια που χρησιμοποιούνται για να ελέγχουν την σχέση των παραπάνω μεταβλητών, συνηγορούν υπέρ της άποψης ότι υπάρχει ισχυρή συσχέτιση ανάμεσά τους:

- I. Pearson Chi-Square=58,59, df=14, p<0,001
- II. Likelihood Ratio= 57,524, df=14, p<0,001
- III. Linear-by-Linear Association= 8,614, df=1, p=0,003 < 0,05

**Πίνακας 5.2. Συσχέτιση της PARTY με τη CAT\_AGE.**

PARTY * CAT_AGE Crosstabulation					
Count		CAT_AGE			Total
		18-34	35-54	55 plus	
PARTY	ND	51	112	154	317
	SYRIZA	80	200	138	418
	PASOK	4	15	35	54
	POTAMI	20	29	21	70
	KKE	21	24	18	63
	XRYSH AYGH	15	31	15	61
	ANEL	12	27	15	54
	ALLO	31	47	46	124
	Total	234	485	442	1161

**Πίνακας 5.3. Ελεγχος Συσχέτισης της PARTY με τη CAT\_AGE.**

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	58,590 <sup>a</sup>	14	<.001
Likelihood Ratio	57,524	14	<.001
Linear-by-Linear Association	8,614	1	,003
N of Valid Cases	1161		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.88.

Επιτόπιες Παρατηρήσεις:

- I. O Tarling (2009), θεωρεί ότι για τη συνεχή μεταβλητή της ηλικίας, κι όχι μόνον, θα πρέπει να γίνονται επίμονες διαιρέσεις (Segmentation), και αναγωγές/μετατροπές (transformation), σε ονομαστική κλίμακα, μέχρις ότου μια συνεχής μεταβλητή μετατραπεί σε κατηγορική ανεξάρτητη μεταβλητή, που εμφανίζει αλληλεπίδραση με την εξαρτημένη. Αυτό και κάναμε, με τη βοήθεια βέβαια, όπως πάντα, της Θεωρίας.
- II. Όπως θα φανεί παρακάτω, κατά την εκτέλεση των μοντέλων 1, 3, 4, 5, 6 και 7 με τη βοήθεια του SPSS, όλες οι μεταβλητές τέθηκαν στο πλαίσιο με τίτλο Factors, διότι όλες οι ανεξάρτητες μεταβλητές αυτού ήταν κατηγορικού τύπου (βλ. Εικ. 5.1). Στο άλλο πλαίσιο με τίτλο Covariates, τίθενται πάντα οι συνεχείς ανεξάρτητες μεταβλητές του μοντέλου.
- III. Η ανεξάρτητη μεταβλητή Q17 που αφορά μια άλλη αιτία που ο ψηφοφόρος ψηφίζει το Κόμμα που ψηφίζει, δεν περιλαμβάνεται στις ανεξάρτητες μεταβλητές του MODEL 7, διότι απεδείχθη μη σημαντική<sup>15</sup>.

### 5.3. Η κατασκευή του απλούστερου μοντέλου MODEL 1.

Ανακαλούμε στη μνήμη το MONTELO 1, από την παράγραφο 1.4 του Κεφ.1 της ΜΕΘΟΔΟΛΟΓΙΑΣ

MODEL 1

---

---

Dependent Var: PARTY

---

---

Independent Vars: SEX

---

---

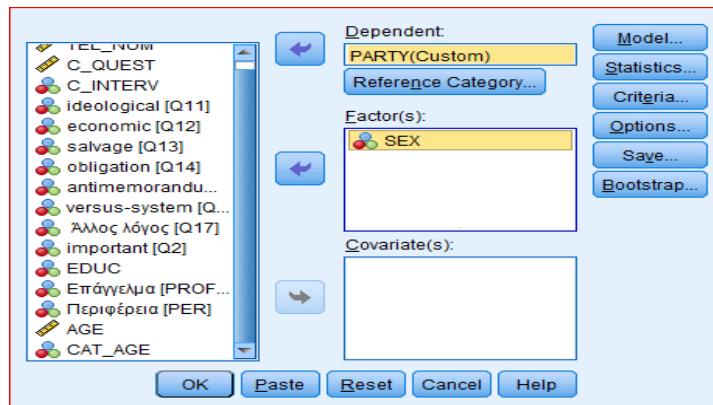
<sup>15</sup> Για να ακριβολογούμε την είχαμε εισαγάγει μαζί με τις υπόλοιπες ανεξάρτητες μεταβλητές στο MODEL 7, που είναι ένα από τα βασικά μας μοντέλα. Τρέξαμε στη συνέχεια τη διαδικασία MLR, και είδαμε ότι δεν είχε τίποτε να προσφέρει στο μοντέλο. Ήταν στατιστικώς ασήμαντη, δεν είχε προσφορά (importance).

**BHMA ΠΡΩΤΟ:** Αφού ανοίξουμε με το γνωστό τρόπο τη βάση των δεδομένων μας ‘GOLEMH.SAV’, απλά ακολουθούμε τη διαδρομή:

Analyze → Regression → Multinomial Logistic (κλικ),

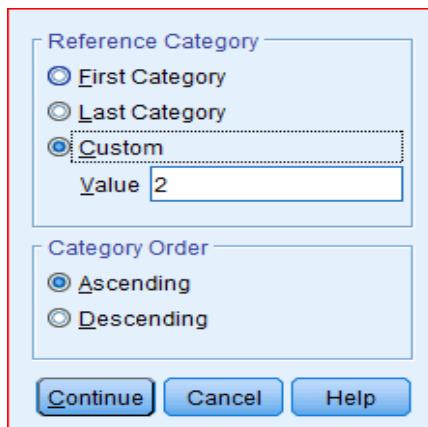
οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.1.

**Εικόνα 5.1.** Βήμα 2o για την εκτέλεση Multinomial Logistic Regression.



**BHMA ΔΕΥΤΕΡΟ:** Στο παράθυρο της Εικόνας 5.1 αναλαμβάνουμε τις εξής δράσεις: Πρώτον, εισάγουμε στο πλαίσιο Dependent την εξαρτημένη μεταβλητή PARTY. Δεύτερον, εισάγουμε στο πλαίσιο Factor(s), τη μοναδική μας ανεξάρτητη μεταβλητή SEX. Τρίτον, κάνουμε κλικ στο πλαίσιο Reference Category και ορίζουμε ως κατηγορία αναφοράς, δηλ. ως συγκριτικό μέτρο, το κόμμα SYRIZA, το οποίο έχει κωδικό 2. Με άλλα λόγια, πρόθεσή μας ήταν, να δηλώσουμε στο SPSS την επιθυμία μας να προβεί σε συγκρίσεις όλων των κομμάτων, με εκείνο του ΣΥΡΙΖΑ. Καθώς όμως κάνουμε κλικ, στο πλαίσιο Reference Category, αναδύεται το παράθυρο της Εικόνας 5.2.

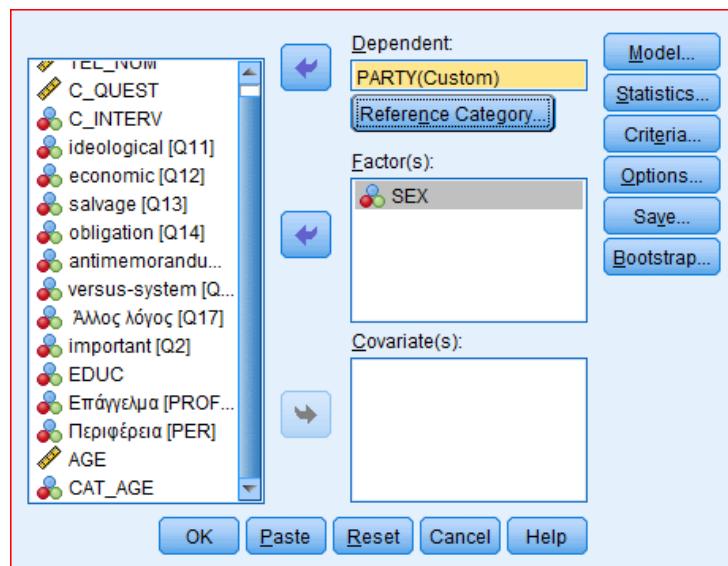
**Εικόνα 5.2.** Βήμα 3o για την εκτέλεση Multinomial Logistic Regression.



**BHMA ΤΡΙΤΟ:** Στο παράθυρο της Εικόνας 5.2 επιλέγουμε με κλικ Custom και θέτουμε 2 στο πλαίσιο Value που είναι ο κωδικός για τη ΝΔ.

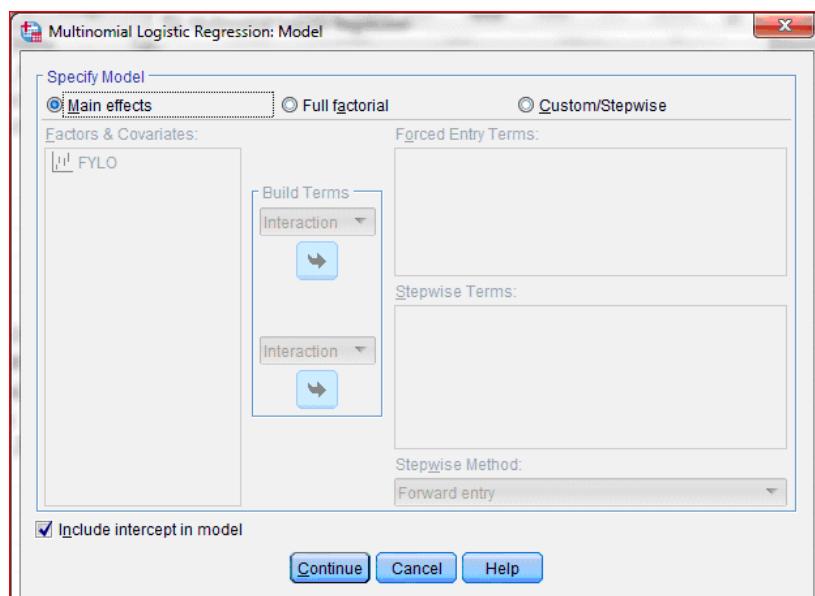
Στη συνέχεια πατάμε το κουμπί Continue και επιστρέφουμε στο αρχικό μας παράθυρο, εκείνο της Εικόνας 5.1, που στο μεταξύ έχει αλλάξει (βλ. Εικόνα 5.3).

**Εικόνα 5.3.** Βήμα 4ο για την εκτέλεση Multinomial Logistic Regression.



**BHMA ΤΕΤΑΡΤΟ:** Ενώ βρισκόμαστε στο παράθυρο της Εικόνας 5.3, πατάμε το κουμπί Model, οπότε μας εμφανίζεται το παράθυρο της Εικόνας 5.4.

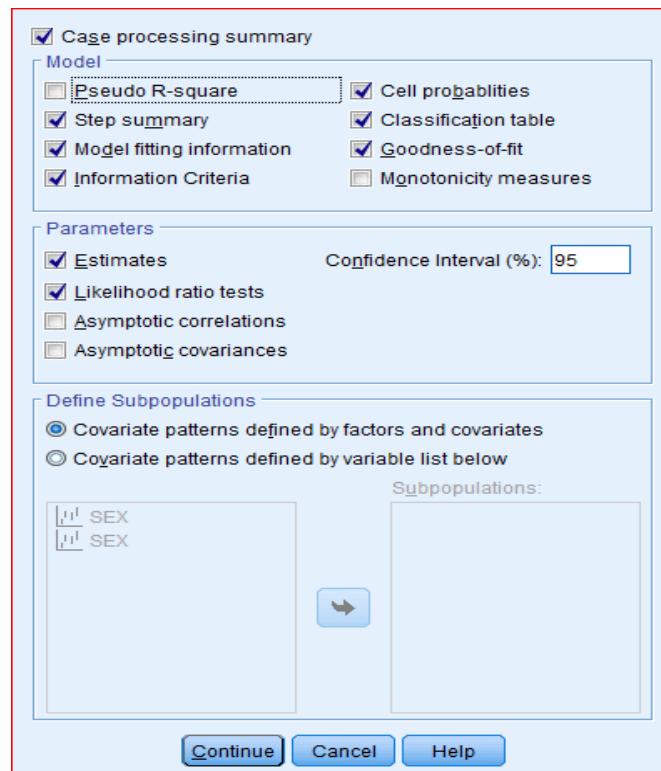
**Εικόνα 5.4.** Βήμα 5ο για την εκτέλεση Multinomial Logistic Regression.



**BHMA ΠΕΜΠΤΟ:** Στο παράθυρο της Εικόνας 5.4 επιλέγουμε με κλικ κύριες επιδράσεις (Main effects). Έτσι κι αλλιώς δεν έχουμε άλλη<sup>16</sup> ανεξάρτητη μεταβλητή πέρα από το SEX.

Με κλικ στο Continue ξαναγυρίζουμε στο βασικό μας παράθυρο, εκείνο της Εικόνας 5.3. Εκεί με κλικ στο κουμπί Statistics, λαμβάνουμε το παράθυρο της Εικόνας 5.5.

**Εικόνα 5.5.** Βήμα 6ο για την εκτέλεση Multinomial Logistic Regression.

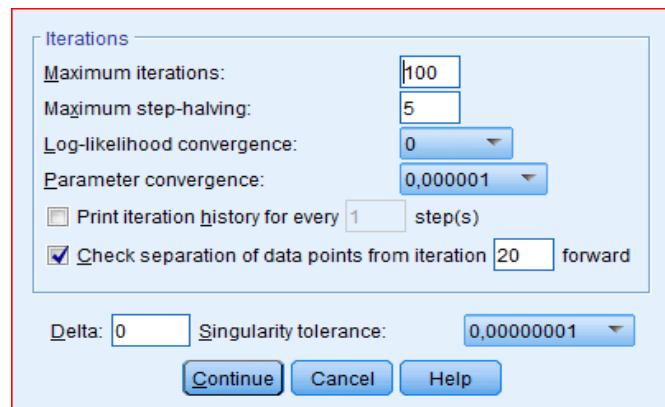


**BHMA ΕΚΤΟ:** Στο παράθυρο της Εικόνας 5.5 ζητάμε πληροφορίες για την προσαρμογή του μοντέλου μας, επιλέγοντας με κλικ τα αντίστοιχα κουμπιά. Με κλικ στο Continue επιστρέφουμε και πάλι στο περιβάλλον της Εικόνας 5.3.

**BHMA ΕΒΔΟΜΟ:** Στο περιβάλλον της Εικόνας 5.3 έχουμε να κάνουμε μια νέα επίσκεψη, με κλικ πάντα, στο κουμπί Criteria, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.6.

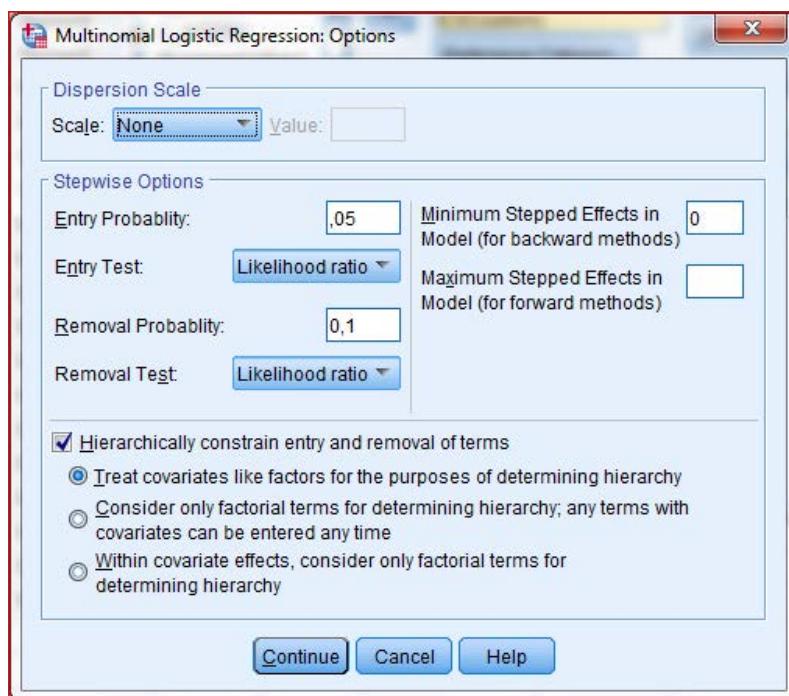
<sup>16</sup> Επομένως, δεν έχει νόημα να ζητήσουμε, να εξεταστεί η αλληλεπίδρασή της (Interaction);

**Εικόνα 5.6.** Βήμα 7ο για την εκτέλεση Multinomial Logistic Regression.



Στο παράθυρο της Εικόνας 5.6 πατάμε το κουμπί Continue, γυρίζουμε στο βασικό μας παράθυρο της Εικόνας 5.3 και πάλι, όπου και πραγματοποιούμε την τελευταία μας επίσκεψη πατώντας κλικ στο κουμπί Options, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.7.

**Εικόνα 5.7.** Βήμα 7ο για την εκτέλεση Multinomial Logistic Regression.



Κλικ και πάλι στο Continue και σαν γυρίσουμε στο βασικό παράθυρο της Εικόνας 5.3, όλα τελειώνουν. Πατάμε OK και άμεσα λαμβάνουμε πολλά και κρίσιμα αποτελέσματα.

### 5.3.1. Τα αποτελέσματα της εκτέλεσης του MONTELOU 1

Υπάρχει ένα πλήθος στοιχείων στα αποτελέσματα και θα σταθούμε σε ένα-ένα διεξοδικά. Στην πορεία θα απαντηθεί και το βασικό ερώτημα της ανάλυσης, το οποίο είναι: Μα ένα μοντέλο που έχει μια μόνο ανεξάρτητη μεταβλητή, είναι δυνατόν να έχει καλή προσαρμογή (fit), και τελικά να μπορεί να δώσει χρήσιμες πληροφορίες (informative); Θα το δούμε αυτό, άμεσα, στα παρακάτω.

Κι ακόμη, πριν σχολιάσουμε τα αποτελέσματα για το MODEL1, ας διασταυρώσουμε<sup>17</sup> με τη διαδικασία Crosstab τις μεταβλητές του φύλου (SEX) και της πρόθεσης ψήφου (PARTY), όπως φαίνεται στον Πίνακα 5.4. Αυτό το κάνουμε για να έχουμε μια αδρομερή εικόνα για τη σχέση των δύο μεταβλητών.

**Πίνακας 5.4.** Συσχέτιση της PARTY με τη SEX.

PARTY * SEX Crosstabulation				
		Count		
		SEX		
		MALE	FEMALE	Total
PARTY	ND	159	159	318
	SYRIZA	204	216	420
	PASOK	35	19	54
	POTAMI	33	37	70
	KKE	30	32	62
	XRYSH AYGH	44	18	62
	ANEL	28	26	54
	ALLO	59	68	127
	Total	592	575	1167

Ας δούμε τώρα και τον Πίνακα 5.5 ο οποίος κάνει λόγο για τις εκτιμήσεις των παραμέτρων του MODEL 1, έτσι ώστε να γίνει εφικτή η ανατομία της διαδικασίας MLR.

<sup>17</sup> Βλ. Βιβλίο : «Δαφέρμος, Β. (2011). Η Κοινωνική Στατιστική και Μεθοδολογία Έρευνας με το SPSS, Εκδόσεις ZHTH, Θεσσαλονίκη», όπου αναπτύσσεται η διαδικασία  $\chi^2$ , στο πλαίσιο της ευρύτερης διαδικασίας Crosstabulation.

**Πίνακας 5.5. Εκτιμήσεις των παραμέτρων του Model 1.**

PARTY <sup>a</sup>	Parameter Estimates						95% Confidence Interval for Exp(B)	
	B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound
ND	Intercept	-,306	,104	8,597	1	,003		
	[SEX=1]	,057	,149	,148	1	,701	1,059	,791 1,417
	[SEX=2]	0 <sup>b</sup>			0			
PASOK	Intercept	-2,431	,239	103,193	1	<,001		
	[SEX=1]	,668	,301	4,919	1	,027	1,950	1,081 3,520
	[SEX=2]	0 <sup>b</sup>			0			
POTAMI	Intercept	-1,764	,178	98,335	1	<,001		
	[SEX=1]	-,057	,259	,049	1	,825	,944	,569 1,568
	[SEX=2]	0 <sup>b</sup>			0			
KKE	Intercept	-1,910	,189	101,627	1	<,001		
	[SEX=1]	-,007	,272	,001	1	,978	,993	,582 1,692
	[SEX=2]	0 <sup>b</sup>			0			
XRYSH AYGH	Intercept	-2,485	,245	102,596	1	<,001		
	[SEX=1]	,951	,296	10,298	1	,001	2,588	1,448 4,626
	[SEX=2]	0 <sup>b</sup>			0			
ANEL	Intercept	-2,117	,208	104,023	1	<,001		
	[SEX=1]	,131	,289	,206	1	,650	1,140	,647 2,010
	[SEX=2]	0 <sup>b</sup>			0			
ALLO	Intercept	-1,156	,139	69,086	1	<,001		
	[SEX=1]	-,085	,203	,175	1	,676	,919	,617 1,367
	[SEX=2]	0 <sup>b</sup>			0			

a. The reference category is: SYRIZA.  
b. This parameter is set to zero because it is redundant.

Εστιάζουμε τώρα την προσοχή μας στο πρώτο διάζωμα του Πίνακα 5.5, το οποίο αφορά τη σύγκριση ND/SΥΡΙΖΑ, αλλά και τη σύγκριση ΑΝΔΡΑ/ΓΥΝΑΙΚΑΣ. Θα εξετάσουμε άμεσα αν το φύλο παρεμβαίνει σε αυτήν την διαμάχη. Παρατηρούμε ότι το φύλο δεν παρεμβαίνει σε αυτή τη σχέση, διότι το επίπεδο στατιστικής σημαντικότητας είναι πολύ ασήμαντο ( $Sig. = 0,701 = 70,1\% > 5\%$ ). Αυτό παραπέρα σημαίνει ότι με ίδιες πιθανότητες άνδρες και γυναίκες κατευθύνονται στην κάλπη της ΝΔ και του ΣΥΡΙΖΑ.

Ας πάμε τώρα στο δεύτερο διάζωμα του ίδιου Πίνακα 5.5. Το επίπεδο στατ. Σημαντικότητας είναι  $Sig. = 0,027 < 0,05$  κι επομένως υπάρχει παρέμβαση του φύλου στη σχέση PASOK/SΥΡΙΖΑ. Επίσης παρατηρούμε ότι, το πρόσημο του συντελεστή  $B = 0,668$  είναι θετικό κάτι το οποίο άμεσα ερμηνεύεται ως εξής: Οι άνδρες συγκρινόμενοι με τις γυναίκες ψηφοφόρους είναι πολύ πιο πιθανό να ψηφίζουν ΠΑΣΟΚ παρά ΣΥΡΙΖΑ. Ας το κάνουμε ωστόσο αυτό πιο συγκεκριμένο, ερμηνεύοντας την τιμή του Odds Ratio που είναι ίση με  $\text{exp}(B) = \text{exp}(0,668) = 1,950$ . Αυτό επακριβώς σημαίνει:

Συγκρινόμενοι με τις γυναίκες, οι άνδρες, εμφανίζουν 1,95 φορές μεγαλύτερη πιθανότητα να ψηφίζουν ΠΑΣΟΚ παρά ΣΥΡΙΖΑ.

Ας δούμε τώρα το λόγο της πιθανότητας μιας γυναίκας συγκρινόμενη με ένα άνδρα να ψηφίζει ΠΑΣΟΚ/ΣΥΡΙΖΑ. Απλά θα λάβουμε τον αντίστροφο (reciprocal) του αριθμού 1,95, ο οποίος είναι ο  $1/1,95 = 0,513$  και σημαίνει ακριβώς το εξής: Μια γυναίκα συγκρινόμενη με ένα άνδρα εμφανίζει μειωμένη πιθανότητα κατά 48,7% ( $100 - 51,3 = 48,7\%$ ) να ψηφίζει ΠΑΣΟΚ παρά ΣΥΡΙΖΑ.

Ας μείνουμε λίγο σε αυτό, το δεύτερο διάζωμα, του Πίνακα 5.5. Παρατηρούμε ότι το intercept = -2,431 και από τη θεωρία ξέρουμε ότι το logit αυτό είναι των γυναικών οι οποίες εδώ συγκροτούν την κατηγορία αναφοράς. Ωστόσο, ανατρέχουμε στον Πίνακα 5.3 για να δούμε αυτός ο αριθμός, από πού έλκει την καταγωγή του. Πράγματι, έχουμε:

$$\ln \left[ \frac{P(PASOK)}{P(SYRIZA)} \right] = \ln \frac{19}{216} = \ln 19 - \ln 216 = -2,431 \quad (5.1)$$

Καθώς από τον Πίνακα 5.3 φαίνεται 19 γυναίκες να ψηφίζουν ΠΑΣΟΚ, και 216 ΣΥΡΙΖΑ.

Τελικά, για τις γυναίκες, που είναι η reference category, η εκτιμούμενη τιμή του κάθε logit ακριβώς είναι το intercept, επειδή ο συντελεστής (coefficient) για την predictor μεταβλητή είναι μηδέν. Ο φυσικός λογάριθμος του λόγου της πιθανότητας μιας γυναίκας να ψηφίζει ΠΑΣΟΚ, προς την πιθανότητα μιας γυναίκας να ψηφίζει ΣΥΡΙΖΑ είναι -2,431.

Ας εστιάσουμε τώρα την προσοχή μας στο διάζωμα του Κόμματος ΠΟΤΑΜΙ στον ίδιο Πίνακα 5.5. Εκεί παρατηρούμε ότι το επίπεδο στατ. σημαντικότητας  $Sig. = 0,825 > 0,05$  πράγμα που σημαίνει ότι το φύλο δεν παρεμβαίνει στη σχέση [ΠΟΤΑΜΙ/ΣΥΡΙΖΑ]. Με άλλα λόγια, άνδρες και γυναίκες ομοιόμορφα μοιράζουν ψήφους σε ΠΟΤΑΜΙ/ΣΥΡΙΖΑ. Αυτή την ερμηνεία, ας την αναδείξουμε περισσότερο από την τιμή του λόγου των Odds η οποία είναι 0,944 δηλ. κοντά στο 1. Αν λάβουμε τον αντίστροφο του 0,944 θα έχουμε:

$$\frac{1}{0,944} = 1,059322 \quad (5.2)$$

Ο αριθμός 1,0593 που βγαίνει, είναι πάλι κοντά στο 1 και αφορά τον αντίστροφο λόγο [ΓΥΝΑΙΚΩΝ/ΑΝΔΡΩΝ]. Πάλι επομένως, οδηγούμαστε στο ίδιο συμπέρασμα: Δεν είναι πιο πιθανό το αντίθετο. Οι γυναίκες συγκρινόμενες με τους άνδρες, δεν είναι πιο πιθανό να πριμοδοτήσουν περισσότερο το ΠΟΤΑΜΙ παρά το ΣΥΡΙΖΑ.

Το επόμενο διάζωμα αφορά το ΚΚΕ. Εδώ παρατηρούμε ότι το φύλο δεν παρεμβαίνει στη σχέση ΚΚΕ/ΣΥΡΙΖΑ, καθώς το παρατηρούμενο επίπεδο σ.σ. είναι κατά πολύ μεγαλύτερο του 5%. (Sig. = 0,978), οπότε οδηγούμαστε στο συμπέρασμα ότι, συγκρινόμενοι με τις γυναίκες, οι άνδρες, δεν εμφανίζουν μεγαλύτερη πιθανότητα να ψηφίζουν περισσότερο το ΚΚΕ παρά το ΣΥΡΙΖΑ. Ας το πούμε αλλιώς: Άνδρες και γυναίκες δεν εμφανίζονται σε διαφορετικό βαθμό διαθέσιμοι να ψηφίσουν ΚΚΕ παρά ΣΥΡΙΖΑ.

Το αμέσως επόμενο διάζωμα, κάνει λόγο για τη σχέση του φύλου με τη σχέση ΧΑ/ΣΥΡΙΖΑ. Εδώ βλέπουμε το φύλο να παρεμβαίνει με στατιστικά σημαντικό τρόπο στη σχέση ΧΑ/ΣΥΡΙΖΑ (Sig. = 0,001 = 0,01% < 5%). Αυτό σημαίνει ότι άνδρες και γυναίκες διαφορετικά ψηφίζουν τα κόμματα ΧΑ και ΣΥΡΙΖΑ. Για να είμαστε πιο ακριβείς οι άνδρες συγκρινόμενοι με τις γυναίκες ψηφίζουν περισσότερο τη ΧΑ παρά το ΣΥΡΙΖΑ, αφού ο συντελεστής για το φύλο είναι με θετικό πρόσημο ( $B = 0,981$ ). Σε ποιο βαθμό ωστόσο γίνεται αυτό; Θα μας το πει ο αντιλογάριθμος του 0,981 για τον οποίο έχουμε  $\exp(0.951) = 2.58$ . Με απλά λόγια, οι άνδρες<sup>18</sup>, συγκρινόμενοι με τις

---

<sup>18</sup> Να σημειώσουμε στο σημείο αυτό, το εξής σημαντικό: Είναι πολύ δύσκολο, έως ακατόρθωτο, να εκτιμηθεί το ακριβές ποσοστό που λαμβάνει η ΧΑ στις εκλογικές αναμετρήσεις. Είναι αδύνατο να ξέρουμε πόσοι θα ψηφίσουν ΧΑ. Διότι απλά, οι πολίτες ντρέπονται ή φοβούνται να σου πουν ότι θα ψηφίσουν ΧΑ. ‘Παιδί μου, εγώ θα ψηφίσω ΧΑ, αλλά μην το γράψεις...’, ήταν η φράση ενός ηλικιωμένου ψηφοφόρου. Στις επανειλημμένες μετρήσεις μας, ιδιαίτερα στην περιοχή της Κρήτης, η Ερευνητική μας Ομάδα οδηγήθηκε στο συμπέρασμα ότι για να βρούμε το ποσοστό που λαμβάνει στον πληθυσμό της Κρήτης η ΧΑ, θα πρέπει να προσθέσουμε το ποσοστό της στο Δείγμα, μιάμιση ή δυο μονάδες, τουλάχιστον...

γυναίκες, εμφανίζονται με υπερδιπλάσιες πιθανότητες (για την ακρίβεια 2,58 φορές) να ψηφίζουν τη ΧΑ παρά το ΣΥΡΙΖΑ.

Το προτελευταίο διάζωμα αφορά τον τρόπο παρέμβασης του φύλου στη σχέση ΑΝΕΛ/ΣΥΡΙΖΑ. Είναι μια στατιστικώς ασήμαντη σχέση ( $Sig. = 0,65 > 0,05$ ), κι επομένως δεν τίθεται ζήτημα διαφορετικής συμπεριφοράς των δύο φύλων απέναντι στο πολιτικό ζεύγος ΑΝΕΛ/ΣΥΡΙΖΑ.

### 5.3.2. Δείκτες προσαρμογής για το MODEL1

Ο Πίνακας 5.6 ουσιαστικά εξετάζει εάν και κατά πόσον το MODEL1 έχει λόγο ύπαρξης. Συγκρίνει δύο μοντέλα από τα οποία το ένα είναι το δικό μας, το MODEL1, που δεν έχει άλλη ανεξάρτητη μεταβλητή, που δεν έχει άλλο predictor, πέρα από τη μεταβλητή του φύλου SEX, με εκείνο που δεν περιλαμβάνει καμιά μεταβλητή, αλλά μόνον το σταθερό όρο (intercept).

Η διαφορά ανάμεσα στα δύο αυτά μοντέλα προσδιορίζεται από τη μεταβολή της ποσότητας -2Ln- Likelihood η οποία ακολουθεί τη  $\chi^2$  κατανομή. Υπολογίζουμε δηλ. την τιμή της ποσότητας -2Ln- Likelihood για ένα μοντέλο που διαθέτει το effect του Φύλου, και για ένα μοντέλο που το μόνο που διαθέτει είναι το intercept.

**Πίνακας 5.6.** Σύγκριση Model 1 με Null μοντέλο.

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	103,927	174,798	75,927 <sup>a</sup>	,000	0	.
SEX	107,089	142,524	93,089	17,161	7	,016

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Ετσι, ο Πίνακας 5.6 πρακτικά ελέγχει την εξής μηδενική υπόθεση:

$H_0$  : Όλες οι παράμετροι του μοντέλου, στο ζήτημα της επίδρασης του Φύλου είναι μηδέν.

Ωστόσο, εδώ παρατηρούμε ότι το επίπεδο στατιστικής σημαντικότητας είναι πολύ μικρότερο του 0,05 (Sig. = 0,016 < 0,05). Επομένως, η  $H_0$  απορρίπτεται κι επομένως, ο συντελεστής του Φύλου δεν είναι μηδέν. Επομένως, ένα μοντέλο με μοναδική παράμετρο το φύλο, έχει λόγο ύπαρξης, κάτι έχει να μας πει.

### 5.3.3. Μέτρα ψευδο- $X^2$ (pseudo-R-square Measures)

Όταν κάποιος διαπραγματεύεται<sup>19</sup> μοντέλα Γραμμικής Παλινδρόμησης, η  $R^2$  στατιστική αναπαρίστανε το ποσοστό της μεταβλητότητας στην εξαρτημένη μεταβλητή που μπορεί να εξηγηθεί από το σύνολο των εξαρτημένων μεταβλητών. Στο πλαίσιο της Λογαριθμογραμμικής Ανάλυσης, δεν έχουμε ένα ανάλογα καλό μέτρο. Δεν έχουμε δηλ. ένα μέτρο, εύκολα ερμηνεύσιμο, που να μιλήσει για την ένταση της σχέσης ανάμεσα στην ανεξάρτητη και τις εξαρτημένες μεταβλητές.

Ωστόσο, εμείς εδώ, και για λόγους πληρότητας, θα αναφέρουμε τρία μέτρα που υποκαθιστούν το  $R^2$  και βέβαια με το χαρακτηρισμό που τα χαρακτηρίζει διεθνώς.

I. Cox and Snell  $R^2$  για το οποίο ισχύει:

$$R_{CS}^2 = 1 - \left( \frac{L(B^{(0)})}{L(\hat{B})} \right)^{\frac{2}{n}} \quad (5.3)$$

II. Nagelkerke's  $R^2$  για το οποίο ισχύει:

$$R_N^2 = \frac{R_{CS}^2}{1 - L(B^{(0)})^{\frac{2}{n}}} \quad (5.4)$$

III. McFadden's  $R^2$  για το οποίο ισχύει:

$$R_M^2 = 1 - \left( \frac{L(\hat{B})}{L(B^{(0)})} \right) \quad (5.5)$$

<sup>19</sup> Βλ. βιβλίο Δαφέρμος, Β. (2011). «Κοινωνική Στατιστική και Μεθοδολογία Έρευνας με το SPSS», Εκδόσεις ZHTH, Θεσσαλονίκη.

Όπου,

$L(\hat{B})$ , είναι η συνάρτηση Ln-likelihood για το μοντέλο με τις εκτιμώμενες παραμέτρους,

$L(B^{(0)})$ , είναι ο πυρήνας (kernel) του νεπέρειου λογαρίθμου της πιθανότητας του μοντέλου που περιλαμβάνει μόνο το intercept, και τέλος, n ο αριθμός των cases.

Στην περίπτωση του MODEL1 που πραγματευόμαστε ο Πίνακας 5.7 είναι αυτός που κάνει λόγο για τις 3 ως άνω ψευδό-στατιστικές.

**Πίνακας 5.7.** Ψεύδο  $X^2$  του Model 1.

Pseudo R-Square	
Cox and Snell	,015
Nagelkerke	,015
McFadden	,004

Στον Πίνακα 5.7 παρατηρούμε ότι οι τιμές των 3 ψευδό-χι-τετραγώνων είναι μικρές, και δεν αξιζουν περαιτέρω σχολιασμού.

### 5.3.4. Ο Πίνακας Ταξινόμησης (Classification Table) για το MODEL1

Ο Πίνακας 5.8 είναι ο Πίνακας Ταξινόμησης για το μοντέλο μας, το MODEL1, αλλά είναι αλήθεια ότι δεν χαρακτηρίζεται από στατιστική δύναμη<sup>20</sup> και ορμή. Είναι ο Πίνακας ανίχνευσης/αναγνώρισης των Κομμάτων.

**Πίνακας 5.8.** Πίνακας Ταξινόμησης (Classification Table) για το Model 1.

Observed	ND	Classification							Percent Correct
		SYRIZA	PASOK	POTAMI	KKE	XRYSH AYGH	ANEL	ALLO	
ND	0	318	0	0	0	0	0	0	0,0%
SYRIZA	0	420	0	0	0	0	0	0	100,0%
PASOK	0	54	0	0	0	0	0	0	0,0%
POTAMI	0	70	0	0	0	0	0	0	0,0%
KKE	0	62	0	0	0	0	0	0	0,0%
XRYSH AYGH	0	62	0	0	0	0	0	0	0,0%
ANEL	0	54	0	0	0	0	0	0	0,0%
ALLO	0	127	0	0	0	0	0	0	0,0%
Overall Percentage	0,0%	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	36,0%

<sup>20</sup> Εδώ είναι αδύνατο να μην ανακαλέσουμε στη μνήμη τα λόγια της Marija Norusis (2010): “Ένα μοντέλο μπορεί να είναι σωστό (correct), και φορτωμένο πληροφορίες (informative), αλλά στην ικανότητα ταξινόμησης μπορεί να είναι φτωχό.

Ενας τέτοιος Πίνακας ταξινόμησης, θα λέμε ότι έχει προβλεπτική δύναμη (predictive power), αν καταφέρνει να εκχωρήσει υψηλή προβλεπόμενη πιθανότητα (predicted probability) σε παρατηρούμενες περιπτώσεις (observed cases). Με άλλα λόγια, να μπορεί να ταξινομήσει σωστά (correct classification). Ο Πίνακας ταξινόμησης (classification Table) στέλνει μια case στο group εκείνο για το οποίο αυτή έχει την υψηλότερη προβλεπόμενη πιθανότητα. Στη συνέχεια μπορούμε να συγκρίνουμε τα δυο groups, παρατηρούμενο (observed) και προβλεπόμενο (predicted).

Ωστόσο, αν τα groups είναι ανισομεγέθη, τότε οι cases είναι πιο πιθανό να πάνε να ταξινομηθούν στα μεγαλύτερου μεγέθους groups... Αβίαστα οδηγούμαστε στο συμπέρασμα ότι η προβλεπτική δύναμη ενός πολυτομικού μοντέλου (Power), δεν είναι εύκολο να μετρηθεί (Hosmer, Lemeshow, 2000). Παράλληλα εγείρονται δυο ερωτήματα:

- I. Τι σχέση έχει ο Classification Table με την καλή προσαρμογή ;
- II. Και ακόμη, ένα μοντέλο που χαρακτηρίζεται σωστό (correct) και informative (πληροφοριακό), έχει οπωσδήποτε και καλή προσαρμογή;

Απαντήσεις στα δυο αυτά ερωτήματα δίνει η Norusis (2010): Ο Classification Table ίσως να κουβαλάει επάνω του σπουδαίες και ενδιαφέρουσες πληροφορίες. Άλλα ταυτόχρονα μπορεί να μην έχει καλή προσαρμογή στα δεδομένα. Επομένως, από μόνος του ο Classification Table λίγες πληροφορίες μπορεί να μας δώσει για την προσαρμογή του μοντέλου.

Τελικά, τι ισχύει; Ο Πίνακας 5.8 τι δείχνει; Η απάντηση, με βάση τα παραπάνω, είναι απλή. Πρόκειται για ένα Πίνακα που δείχνει Μοντέλο χαμηλής προβλεπτικής ικανότητας. Στον πυθμένα του, κάτω και δεξιά βλέπουμε τον αριθμό των σωστά ταξινομημένων περιπτώσεων. Και είναι πολύ λίγες... Μόλις 36% . Και το μόνο Κόμμα που καταφέρνει να εντοπίσει και μάλιστα 100% είναι εκείνο με τη μεγαλύτερη συχνότητα εμφάνισης. Εκείνο που πλειοψηφεί. Το Κόμμα του ΣΥΡΙΖΑ.

### 5.3.5. Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 1

Ο Πίνακας 5.9 για το πρώτο μας μοντέλο, το MODEL1, με μία, και μόνη, ανεξάρτητη μεταβλητή –ας μην το ξεχνάμε αυτό– εμμέσως πλην σαφώς, φανερώνει ότι το μοντέλο μας δεν μπορεί να προβλέψει. Ωστόσο, ας δούμε τι έχει να μας πει για τον υποπληθυσμό των ΑΝΔΡΩΝ, αλλά και εκείνο των ΓΥΝΑΙΚΩΝ.

**Πίνακας 5.9.** Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 1,  
ανά φύλο.

Observed and Predicted Frequencies						
SEX	PARTY	Observed	Frequency		Percentage	
			Predicted	Pearson Residual	Observed	Predicted
MALE	ND	159	159,000	,000	26,9%	26,9%
	SYRIZA	204	204,000	,000	34,5%	34,5%
	PASOK	35	35,000	,000	5,9%	5,9%
	POTAMI	33	33,000	,000	5,6%	5,6%
	KKE	30	30,000	,000	5,1%	5,1%
	XRYSH AYGH	44	44,000	,000	7,4%	7,4%
	ANEL	28	28,000	,000	4,7%	4,7%
	ALLO	59	59,000	,000	10,0%	10,0%
FEMALE	ND	159	159,000	,000	27,7%	27,7%
	SYRIZA	216	216,000	,000	37,6%	37,6%
	PASOK	19	19,000	,000	3,3%	3,3%
	POTAMI	37	37,000	,000	6,4%	6,4%
	KKE	32	32,000	,000	5,6%	5,6%
	XRYSH AYGH	18	18,000	,000	3,1%	3,1%
	ANEL	26	26,000	,000	4,5%	4,5%
	ALLO	68	68,000	,000	11,8%	11,8%

The percentages are based on total observed frequencies in each subpopulation.

Οι άνδρες φαίνεται να ετοιμάζονται να επιλέξουν/ψηφίσουν:

- I. Τη ΝΔ σε ποσοστό 26,9%.
- II. Το ΣΥΡΙΖΑ σε ποσοστό 34,5%.
- III. Το ΠΑΣΟΚ σε ποσοστό 5,9%.
- IV. Το ΠΟΤΑΜΙ σε ποσοστό 5,6%.
- V. Το KKE σε ποσοστό 5,1 %.
- VI. Τη ΧΑ σε ποσοστό 7,4 %
- VII. Τους ΑΝΕΛ σε ποσοστό 4,7%.
- VIII. Κάτι ΑΛΛΟ σε ποσοστό 10 %.

Οι γυναίκες, από την άλλη, φαίνεται να ετοιμάζονται να επιλέξουν/ψηφίσουν:

- I. Τη ΝΔ σε ποσοστό 27,7 %.
- II. Το ΣΥΡΙΖΑ σε ποσοστό 37,6 %.
- III. Το ΠΑΣΟΚ σε ποσοστό 3,3%.
- IV. Το ΠΟΤΑΜΙ σε ποσοστό 6,4%.
- V. Το ΚΚΕ σε ποσοστό 5,6%.
- VI. Τη ΧΑ σε ποσοστό 3,1%
- VII. Τους ΑΝΕΛ σε ποσοστό 4,5%.
- VIII. Κάτι ΑΛΛΟ σε ποσοστό 11,8%.

#### 5.3.6. Η προσαρμογή (goodness-of-fit) του Model 1

Ο Πίνακας 5.10 κάνει λόγο για την προσαρμογή του μοντέλου μας, του **Model 1**, στα δεδομένα. Στα δεδομένα που διαθέτουμε. Για να έχουμε καλή προσαρμογή για ένα οποιοδήποτε μοντέλο, καλό fit, όπως λέγεται, θα πρέπει οι δείκτες Pearson και Deviance να είναι καλοί. Και είναι καλοί, όταν το αντίστοιχο, παρατηρούμενο επίπεδο στατιστικής σημαντικότητας (significance level), για κάθε έναν από αυτούς, είναι μεγαλύτερο του 5% .

**Πίνακας 5.10.** Προσαρμογή (goodness-of-fit) του Model 1.

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	,000	0	.
Deviance	,000	0	.

Μα όμως για το **Model 1**, το SPSS δηλώνει αδυναμία να τυπώσει το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας.

#### Επιτόπιο Συμπέρασμα για το ΜΟΝΤΕΛΟ 1:

Είναι σαφές ότι αφού δεν έχουμε καμία πληροφορία για την προσαρμογή του συγκεκριμένου μοντέλου στα δεδομένα, δεν θα μπορούσε να γίνει λόγος για ένα καλό μοντέλο, πολύ δε περισσότερο, δεν θα μπορούσε να είναι, ένα άριστο μοντέλο.

### **5.3.7. Τελικά συμπεράσματα για το Model 1**

- I. Όλα όσα προαναφέρθηκαν, συμβαίνουν μόνον αν υποθέσουμε ότι, ο παράγων του Φύλου παρεμβαίνει, με τους άλλους παράγοντες να παραμένουν αμέτοχοι.
- II. Ωστόσο, αφού με μια μόνο ανεξάρτητη μεταβλητή μπορέσαμε να καταλήξουμε σε προβλεπόμενες πιθανότητες (probabilities) που δεν είναι καθόλου μακριά από τα πραγματικά ποσοστά που έλαβαν τα Κόμματα στις εκλογές του Ιανουαρίου, τότε με μεγάλη πιθανότητα μπορούμε να πούμε ότι βρισκόμαστε σε καλό δρόμο για τον εντοπισμό του τελικού, σωστού, αληθινού, βέλτιστου μοντέλου (best model).
- III. Βέβαια, αυτό καθόλου δεν σημαίνει ότι, όσο πιο πολλές ανεξάρτητες μεταβλητές εισαγάγουμε στο μοντέλο, θα ανέβει συνακόλουθα και η προβλεπτική του δύναμη...! Καλό μοντέλο σημαίνει, συγκεκριμένο, όσο το δυνατόν μικρό αριθμό μεταβλητών, με προσφορά, και καλή προσαρμογή. Αυτό ψάχνουμε.
- IV. Μια τελευταία παρατήρηση που εμπεριέχεται στην όλη αναζητητική του βέλτιστου μοντέλου είναι η εξής: Το τελικό μας μοντέλο (final model), κοντά στα άλλα δεν μπορεί παρά να είναι ουσιωδώς ερμηνεύσιμο Norusis (2010), και ακριβές (accurate) Bλ. SPSS ADVANCED MODELS 10.0, 1999 by SPSS INC.

### **5.4. Η κατασκευή του Model 2.**

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του MONTELOU 2.

MODEL 2.

=====  
Dependent Var: PARTY

=====  
Independent Vars(3): SEX, CAT\_AGE, SEX \* CAT\_AGE

Όπως είχαμε επισημάνει στην παρ. 5.4, το Model 2 περιλαμβάνει δυο ανεξάρτητες μεταβλητές και ακόμη την αλληλεπίδραση μεταξύ αυτών. Επιβάλλεται να ξέρουμε αν υπάρχει μεταξύ τους αλληλεπίδραση διότι αν εντοπίσουμε κάτι τέτοιο η Ερμηνεία του μοντέλου είναι πολύ διαφορετική. Αυτό εύκολα μπορεί να διαπιστωθεί, όπως θα δούμε αμέσως παρακάτω.

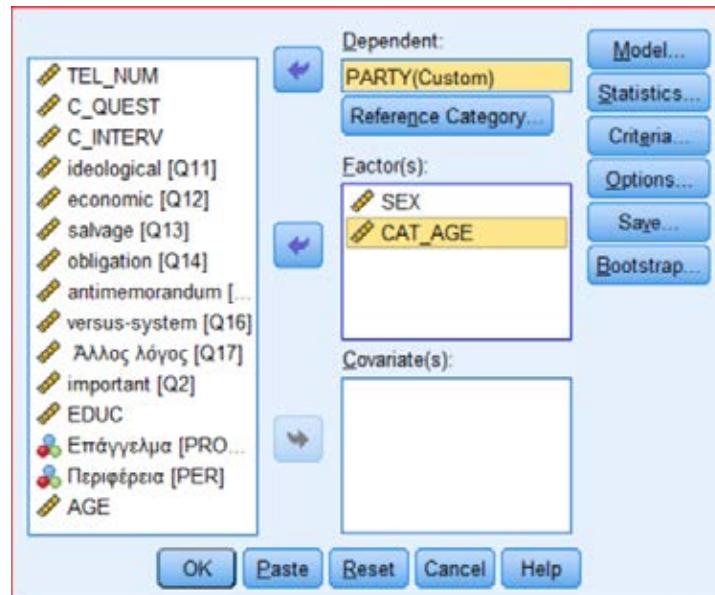
Ας τρέξουμε λοιπόν το Model 2 κατά τα γνωστά.

Με ανοικτή τη βάση των δεδομένων μας ‘GOLEMH.SAV’, απλά ακολουθούμε τη διαδρομή:

Analyze → Regression → Multinomial Logistic (κλικ),

οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.8.

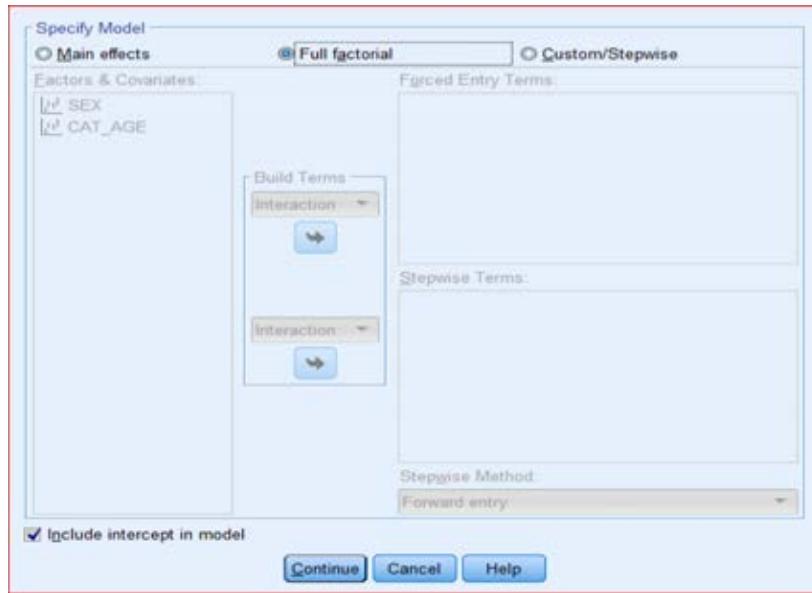
**Εικόνα 5.8.** Προσδιορισμός Μεταβλητών Model 2.



Στην Εικόνα 5.8 βλέπουμε το μοντέλο που οικοδομήσαμε με κατηγορία αναφοράς και πάλι το ΣΥΡΙΖΑ = 2, στην εξαρτημένη μας μεταβλητή PARTY (επιλογή Κόμματος), και με ανεξάρτητες μεταβλητές το φύλο SEX και τις ηλικιακές κατηγορίες CAT\_AGE.

Στη συνέχεια, πατάμε το κουμπί Model οπότε μας εμφανίζεται το νέο παράθυρο της Εικόνας 5.9.

**Εικόνα 5.9.** Προσδιορισμός Model 2.



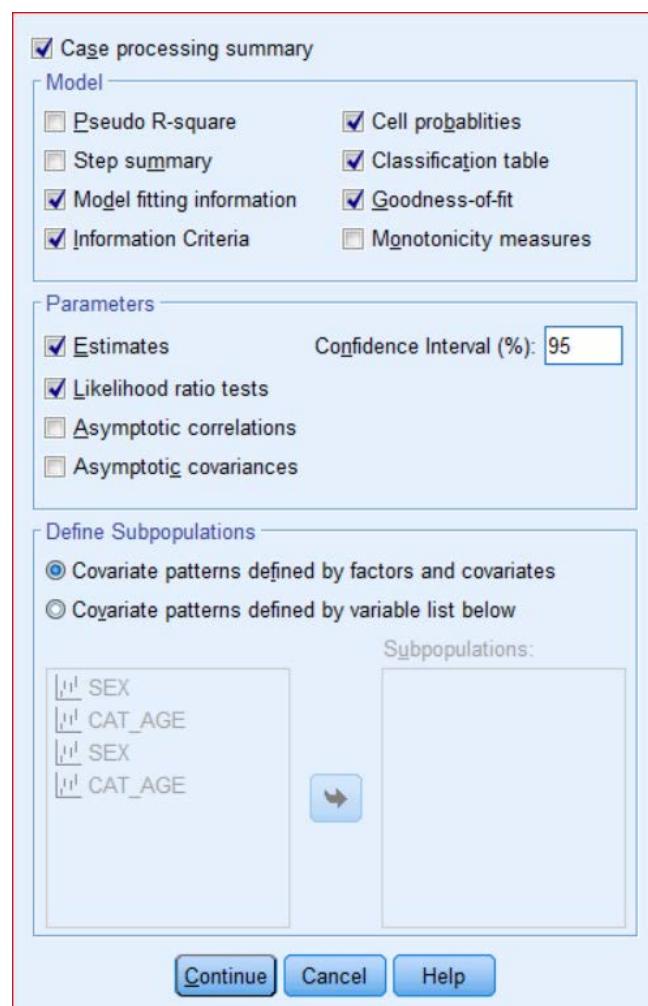
Στην Εικόνα 5.9 επιλέγουμε Full Factorial, και όχι main effects, επειδή ακριβώς μας ενδιαφέρει να ανιχνεύσουμε την ύπαρξη ή μη interaction, μεταξύ των ανεξάρτητων μεταβλητών.

Στη συνέχεια πατάμε το κουμπί continue και τσεκάρουμε την ένδειξη include intercept in model, αν αυτό δεν έχει ήδη γίνει από το SPSS. Έτσι επιστρέφουμε στο βασικό μας παράθυρο, εκείνο της Εικόνας 5.8.

Εκεί πατάμε το κουμπί Statistics, οπότε μας εμφανίζεται το παράθυρο της Εικόνας 5.10, όπου μας δίνεται η ευκαιρία να ζητήσουμε από το SPSS μια σειρά πληροφορίες που σχετίζονται με την προσαρμογή του μοντέλου μας στα δεδομένα, κι όχι μόνον.

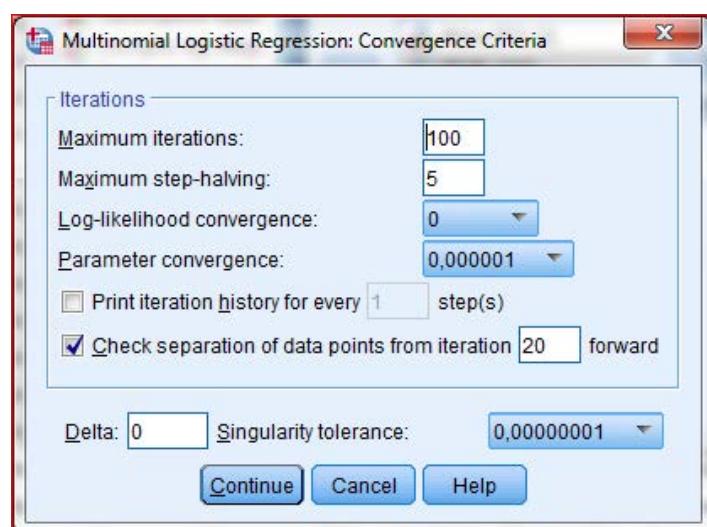
Στην Εικόνα 5.10 είναι σαφές ότι ζητάμε από το SPSS να μας δώσει σε επίπεδο εμπιστοσύνης 95%, ικανότητα προσαρμογής του μοντέλου στα δεδομένα (Goodness-of-fit), ικανότητα ανίχνευσης των Κομμάτων (Classification Table), εκτίμηση της σπουδαιότητας παραμέτρων εκτίμησης (Parameter Estimates) και το σπουδαιότερο πρόβλεψη για την πρόθεση ψήφου, σε κάθε περίπτωση (Observed and Predicted Frequencies).

**Εικόνα 5.10.** Επιλογές μενού Statistics, Model 2.



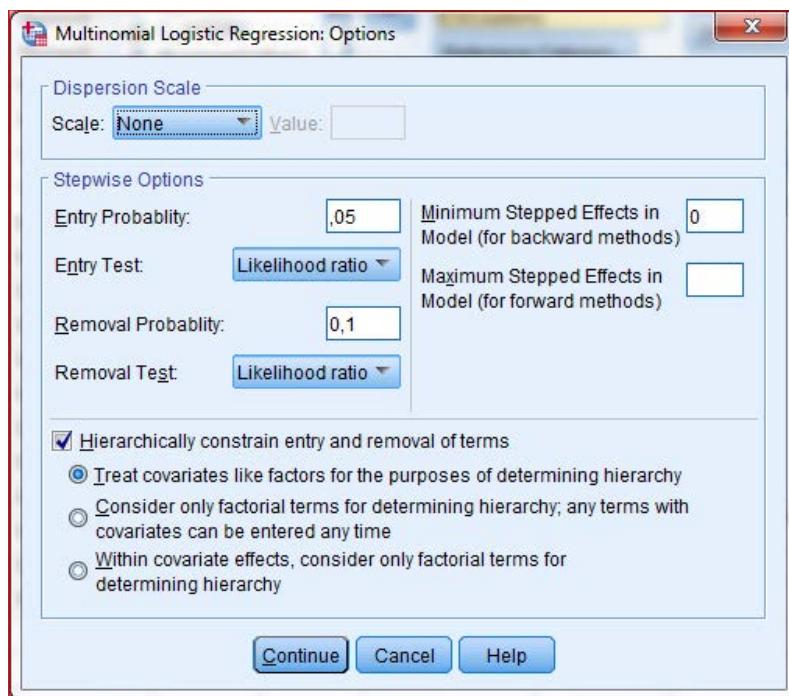
Στη συνέχεια, πατώντας το κουμπί continue επιστρέφουμε στο πρωταρχικό/βασικό μας μενού και πατάμε το κουμπί criteria, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.11.

**Εικόνα 5.11.** Επιλογές μενού Criteria, Model 2.



Στην Εικόνα 5.11 παρατηρούμε τα κριτήρια - προγραμματιστικά και στατιστικά- που από μόνο του (by default), το SPSS έθεσε, και τα οποία προσδιορίζουν τις παραμέτρους σύγκλισης του μοντέλου μας (Model 2).

**Εικόνα 5.12.** Επιλογές μενού Options, Model 2.



Στην Εικόνα 5.12 και πάλι το πρόγραμμα από μόνο του καθορίζει την πιθανότητα εισόδου μιας μεταβλητής στο μοντέλο (5%), και αντίστοιχα την πιθανότητα εξόδου από αυτό (10%). Ο ίδιος Πίνακας 5.12 καθορίζει τους ιεραρχικούς περιορισμούς εισόδου και αφαίρεσης των όρων στη λειτουργία του μοντέλου. Το πρόγραμμα, όπως παραπάνω, από μόνο του δηλ., έχει επιλέξει να θεωρεί τις συνεχείς μεταβλητές (covariates), σαν παράγοντες (factors), προκειμένου να καθορίσει την ιεραρχία. Βέβαια, εμείς εδώ, στο Model 2, δεν έχουμε συνεχείς μεταβλητές, κι επομένως δεν προβληματίζόμαστε πάνω σε αυτό.

Στο σημείο αυτό τρέχουμε το μοντέλο που μόλις οικοδομήσαμε και διαπιστώνουμε την ανυπαρξία αλληλεπίδρασης.

Πράγματι ο Πίνακας 5.11, στην τελευταία του γραμμή, τελευταία στήλη, μας λέει ότι το αντιστοιχούμενο στην αλληλεπίδραση, παρατηρούμενο επίπεδο στατ. σημαντικότητας είναι ασήμαντο ( $Sig. = 0,076 = 7,6\% > 5\%$ ).

Πίνακας 5.11. Σύγκριση Model 2 με Null μοντέλο.

Effect	Likelihood Ratio Tests			
	Model Fitting Criteria -2 Log Likelihood of Reduced Model	Likelihood Ratio Tests		
		Chi-Square	df	Sig.
Intercept	175,678 <sup>a</sup>	,000	0	.
SEX	175,678 <sup>a</sup>	,000	0	.
CAT_AGE	175,678 <sup>a</sup>	,000	0	.
SEX * CAT_AGE	197,830	22,152	14	,076

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

### Συμπέρασμα

Δεν υπάρχει αλληλεπίδραση ανάμεσα στους predictors SEX και CAT\_AGE.

### 1η Σημαντική επισήμανση για τα επόμενα

Το Model 2 έδειξε ότι *interaction δεύτερης τάξης*, μεταξύ φύλου και ηλικιακών κατηγοριών, δεν υπάρχει. Η μεταβλητή SEX\*CAT\_AGE αποδείχθηκε στατιστικώς ασήμαντη δηλ. ο πληθυσμιακός παλινδρομικός συντελεστής (regression coefficient) είναι μηδέν.

### 2η Σημαντική επισήμανση για τα επόμενα

Το Model 2 είναι ένα μοντέλο χωρίς Ερμηνευσιμότητα. Διότι ακόμη κι αν είχε τυπωμένα, παρατηρούμενα επίπεδα για τις μεταβλητές του Φύλου (SEX) και των Ηλικιακών Κατηγοριών (CAT\_AGE) στην τελευταία στήλη του Πίνακα 5.11, δεν θα μπορούσαμε να κάνουμε λόγο για τις αυτόνομες/ ξεχωριστές επιδράσεις αυτών των μεταβλητών (main effects), αφού αυτές συμμετέχουν μαζί σε μια νέα μεταβλητή, τη μεταβλητής της επίδρασης SEX\*CAT\_AGE.

## 5.5. Η κατασκευή του Model 3.

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του ΜΟΝΤΕΛΟΥ 3.

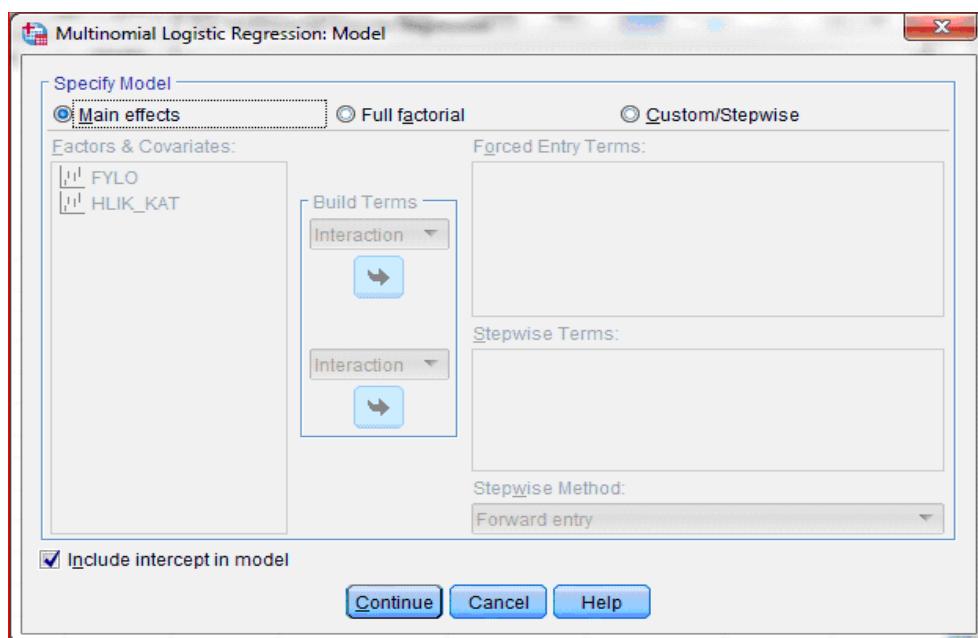
MODEL 3.  
=====

Dependent Var: PARTY  
=====

Independent Vars(2): SEX, CAT\_AGE  
=====

Μετά το ξεκαθάρισμα ανυπαρξίας αλληλεπίδρασης, μεταξύ των ανεξάρτητων μεταβλητών του μοντέλου μας (βλ. Model 2), είμαστε πάντοτε σε θέση να ξέρουμε, αν θα επιλέξουμε μεταξύ κύριων επιδράσεων (main effects), ή πλήρους μοντέλου (Full Model). Προφανώς και επιλέγουμε main effects για το Model 3, καθώς δεν υπάρχει αλληλεπίδραση μεταξύ των Predictors, SEX και CAT\_AGE.

**Εικόνα 5.13.** Προσδιορισμός Model 3.



Στο σημείο αυτό, τρέχουμε το Model 3, επαναλαμβάνουμε όλα όσα κάναμε στην παράγραφο 5.4, με μόνη τη διαφορά ότι τώρα επιλέγουμε για το μοντέλο μας Main effects (βλ. Εικόνα 5.13).

**Πίνακας 5.12.** Model 3 - Εκτίμηση Αποτελεσμάτων Εθνικών Εκλογών Ιανουαρίου 2015.

Case Processing Summary			
		N	Marginal Percentage
PARTY	ND	317	27,3%
	SYRIZA	418	36,0%
	PASOK	54	4,7%
	POTAMI	70	6,0%
	KKE	62	5,3%
	XRYSH AYGH	61	5,3%
	ANEL	54	4,7%
	ALLO	124	10,7%
SEX	MALE	591	50,9%
	FEMALE	569	49,1%
CAT_AGE	18-34	234	20,2%
	35-54	484	41,7%
	55 plus	442	38,1%
Valid		1160	100,0%
Missing		757	
Total		1917	
Subpopulation		6	

Ο πρώτος πίνακας που μας προσφέρει το SPSS όταν εκτελεί Ονομαστική Παλινδρόμηση (Nominal Regression<sup>21</sup>), είναι ο Πίνακας 5.12, οποίος μας θυμίζει πόσο καλά έκαναν τη Δειγματοληψία τα παιδιά του Τμήματος Πολιτικής Επιστήμης του Π.Κ., στο Εργαστήριο Κοινωνικής Στατιστικής και Πολιτικής Έρευνας.

Ο Πίνακας 5.12 δείχνει αρκετά σπουδαία πράγματα, όπως φάνηκε από τα πραγματικά αποτελέσματα των Εθνικών Εκλογών του Ιανουαρίου 2015 (βλ. διαδίκτυο, Υπουργείο Εσωτερικών. Επίσης βλ. Πίνακα 5.36, παρακάτω).

**Πίνακας 5.13.** Σύγκριση Model 3 με Null μοντέλο.

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	285,311	320,704	271,311			
Final	253,830	395,403	197,830	73,481	21	<.001

<sup>21</sup> Ναι, έτσι αποκαλεί το SPSS 28, το Multinomial Logistic Regression.

Ο Πίνακας 5.13 μας πληροφορεί ότι το Model 3, που το ονομάζει Final Model, καθώς είναι το μοντέλο που συγκροτούν οι δυο ανεξάρτητες μεταβλητές του Φύλου και των Ηλικιακών Κατηγοριών, έτσι όπως τις έχουμε εμείς, σαν Ερευνητές, προσδιορίσει. Προφανώς και δεν τίθεται για εκείνο κανένα πρόβλημα ύπαρξης, αφού συγκρινόμενο με το μοντέλο που περιλαμβάνει μόνο το intercept, φαίνεται να έχει με αυτό στατιστικώς σημαντική διαφορά (Sig.<0,001).

**Πίνακας 5.14.** Προσαρμογή (goodness-of-fit) του Model 3.

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	20,765	14	,108
Deviance	22,152	14	,076

Ο Πίνακας 5.14 με τους δείκτες του, τόσο με εκείνον του Pearson (Sig. = 0,108 > 0,05), όσο και με εκείνον της Απόκλισης (Deviance, Sig. 0,076 >0,05), κάνει λόγο για καλή προσαρμογή. Υπενθυμίζεται ότι κάθε Πίνακας με αυτόν τον τίτλο Goodness-of-Fit ελέγχει πάντα την παρακάτω μηδενική υπόθεση  $H_0$ :

$H_0$  : Η προσαρμογή του μοντέλου στα δεδομένα είναι καλή.

Προφανώς εδώ, είτε λάβουμε υπόψη μας το κριτήριο Pearson, είτε το κριτήριο Deviance, το συμπέρασμα είναι το ίδιο. Η μηδενική υπόθεση ισχύει και στις δύο περιπτώσεις. Επομένως, κατά τη μαρτυρία και των δυο στατιστικών κριτηρίων, η προσαρμογή του συγκεκριμένου μοντέλου είναι καλή.

**Πίνακας 5.15.** Σύγκριση των Reduced Model 3 με Null μοντέλο.

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	253,830	395,403	197,830 <sup>a</sup>	,000	0	.
SEX	255,578	361,757	213,578	15,748	7	,028
CAT_AGE	283,461	354,248	255,461	57,631	14	<,001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Ο Πίνακας 5.15 είναι ένας επίσης σημαντικός πίνακας, αφού σχετίζεται με τα επίσης σημαντικά εργαλεία ελέγχου της προσαρμογής, τα λεγόμενα Likelihood Ratio Tests. Αυτά τα τεστ παρακολουθούν τις αυξομειώσεις της ποσότητας -2LL από το αναγμένο (reduced<sup>22</sup>) μέχρι το τελικό (final) μοντέλο. Αν αυτή η διαφορά μεταξύ των δύο μοντέλων είναι στατιστικώς σημαντική, το final model, δηλ. το μοντέλο με τις δύο ανεξάρτητες μεταβλητές, του φύλου και των ηλικιακών κατηγοριών, έχει καλή προσαρμογή.

Ταυτόχρονα με την προσαρμογή, ο Πίνακας 5.15 κάνει λόγο και για τη σπουδαιότητα των παραγόντων του Φύλου και των Ηλικιακών κατηγοριών, αυτό εμμέσως πλην σαφώς. Δηλ. από τη μελέτη του συνάγεται το κρίσιμο συμπέρασμα στην υπόθεση: Αν φύγει μια μεταβλητή από το μοντέλο, θα υπάρξει σημαντική αλλαγή σε αυτό; Για μεν τον παράγοντα του Φύλου(SEX), έχουμε  $Sig. = 0,028 = 2,8\% < 5\%$ , για δε τον παράγοντα των ηλικιακών κατηγοριών (CAT\_AGE), έχουμε  $Sig. = 0,0005 = 0,05\% < 5\%$ . Δηλ. και οι δύο παράγοντες αυτοί έχουν σημαντικές επιδράσεις στο μοντέλο.

Ηρθε, ωστόσο, η ώρα να προσδιορίσουμε με μαθηματικό τρόπο τους δείκτες AIC και BIC.

Για τον AIC (Akaike Information Criterion) ισχύει η σχέση:

$$AIC = -2 \ln(Likelihood) + (2 \times number\_of\_parameters) \quad (5.6)$$

Για τον BIC (Bayesian Information Criterion) ισχύει η σχέση:

$$BIC = -2 \ln(Likelihood) + \ln(n) \times df \quad (5.7)$$

Όπου,

n = το συνολικό μέγεθος του δείγματος, και

df = είναι οι βαθμοί ελευθερίας για το μοντέλο.

Οι Schwarz (1978) και Raftery (1986), είναι οι εισηγητές της στατιστικής BIC.

---

<sup>22</sup> Reduced model, είναι το μοντέλο που σχηματίζεται από το final model, αν αφαιρέσουμε από αυτό ένα effect.

Επιστρέφουμε στον Πίνακα 5.15 για να πούμε ότι κάθε φορά που προσθέτουμε μια μεταβλητή στο μοντέλο μας, η ποσότητα -2LL μειώνεται. Αυτό όμως, όπως έχουμε τονίσει επανειλημμένα, καθόλου δεν σημαίνει ότι τα μοντέλα με περισσότερες μεταβλητές είναι καλύτερα από τα μοντέλα με λιγότερες, επειδή ακριβώς, καθώς προστίθενται οι μεταβλητές, κεφαλαιοποιούνται συσχετίσεις που είναι παρούσες σε ένα συγκεκριμένο δείγμα. Έτσι, ένα μοντέλο που εκτιμήθηκε από ένα δείγμα, συνήθως ταιριάζει καλύτερα στο δείγμα, παρά στον πληθυσμό από τον οποίο αυτό το δείγμα προέρχεται. Αυτός ακριβώς είναι ο λόγος που εμείς θέλουμε ένα μοντέλο να ταιριάζει καλά στα δεδομένα, με τη χρήση ενός μικρού αριθμού μεταβλητών. Με βάση αυτή τη συλλογιστική ξεκίνησαν να εφαρμόζονται τα κριτήρια AIC και BIC. Και είναι τέλος αλήθεια ότι, το BIC κριτήριο, σε αντίθεση με το AIC, ευνοεί μοντέλα με μικρότερο αριθμό μεταβλητών.

Ο Πίνακας 5.16 των παραμέτρων εκτίμησης είναι, όπως ξαναείπαμε, ένας πίνακας με σημαντικές πληροφορίες. Να θυμίσουμε και πάλι ότι ο πίνακας αυτός δομήθηκε με κατηγορία αναφοράς το SYRIZA, που αρχικά είχε κωδικοποιηθεί στο αρχείο δεδομένων (data set) και στη μεταβλητή του ΚΟΜΜΑΤΟΣ (PARTY) με κωδικό 2 (2=SYRIZA).

Εστιάζουμε την προσοχή μας στο πρώτο διάζωμα του Πίνακα 5.16, που κάνει λόγο για την αντίθεση ΝΔ/ΣΥΡΙΖΑ. Το Φύλο σε αυτή την αντίθεση δεν παρεμβαίνει ( $Sig. = 0,849 = 84,9\% > 5\%$ ), όπως είδαμε και στο προηγούμενο μοντέλο [MODEL 2].

Αντίθετα, οι ηλικιακές κατηγορίες παρεμβαίνουν σημαντικά στην αντίθεση ΝΔ/ΣΥΡΙΖΑ και μάλιστα, οι ηλικίες των 18-34 ετών συγκρινόμενες με τις ηλικίες άνω των 55, παρουσιάζουν στατιστικώς σημαντική διαφορά ( $Sig. = 0,009 < 0,05$ ). Ακόμη στο ίδιο διάζωμα του ίδιου Πίνακα 5.16 παρατηρούμε ότι και οι ηλικίες των 35-54 ετών συγκρινόμενες με τις ηλικίες των 55 και άνω παρουσιάζουν στατιστικώς σημαντική διαφορά ( $Sig. = 0,0005 < 0,05$ ).

**Πίνακας 5.16.** Ερμηνεία των παραμέτρων του Model 3.

Parameter Estimates								
PARTY <sup>a</sup>		B	Std. Error	Wald	df	Sig.	95% Confidence Interval for Exp (B)	
							Lower Bound	Upper Bound
ND	Intercept	,095	,141	,455	1	,500		
	[SEX=1]	,029	,150	,036	1	,849	,1,029	,766 1,382
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	-,559	,214	6,821	1	,009	,572	,376 ,870
	[CAT_AGE=2]	-,688	,166	17,089	1	<,001	,502	,363 ,696
	[CAT_AGE=3]	0 <sup>b</sup>			0			
PASOK	Intercept	-1,733	,272	40,525	1	<,001		
	[SEX=1]	,613	,304	4,065	1	,044	,1,846	,1,017 3,349
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	-1,611	,547	8,685	1	,003	,200	,068 ,583
	[CAT_AGE=2]	-1,191	,329	13,131	1	<,001	,304	,160 ,579
	[CAT_AGE=3]	0 <sup>b</sup>			0			
POTAMI	Intercept	-1,847	,268	47,649	1	<,001		
	[SEX=1]	-,070	,259	,074	1	,786	,932	,561 1,549
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	,495	,343	2,087	1	,149	1,640	,838 3,211
	[CAT_AGE=2]	-,052	,307	,028	1	,867	,950	,520 1,735
	[CAT_AGE=3]	0 <sup>b</sup>			0			
KKE	Intercept	-2,025	,287	49,889	1	<,001		
	[SEX=1]	-,024	,273	,008	1	,930	,976	,571 1,668
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	,699	,351	3,972	1	,046	2,011	,1,012 3,999
	[CAT_AGE=2]	-,127	,334	,145	1	,704	,881	,458 1,694
	[CAT_AGE=3]	0 <sup>b</sup>			0			
XRYSH AYGH	Intercept	-2,801	,345	65,911	1	<,001		
	[SEX=1]	,930	,298	9,758	1	,002	2,535	,1,414 4,543
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	,564	,393	2,057	1	,151	1,757	,813 3,795
	[CAT_AGE=2]	,395	,335	1,392	1	,238	1,485	,770 2,862
	[CAT_AGE=3]	0 <sup>b</sup>			0			
ANEL	Intercept	-2,287	,314	53,039	1	<,001		
	[SEX=1]	,129	,290	,198	1	,657	1,137	,645 2,007
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	,325	,412	,621	1	,431	1,384	,617 3,103
	[CAT_AGE=2]	,223	,341	,427	1	,514	1,249	,641 2,437
	[CAT_AGE=3]	0 <sup>b</sup>			0			
ALLO	Intercept	-1,067	,199	28,736	1	<,001		
	[SEX=1]	-,062	,205	,093	1	,761	,939	,628 1,405
	[SEX=2]	0 <sup>b</sup>			0			
	[CAT_AGE=1]	,149	,272	,302	1	,583	1,161	,682 1,977
	[CAT_AGE=2]	-,352	,235	2,244	1	,134	,703	,443 1,115
	[CAT_AGE=3]	0 <sup>b</sup>			0			

a. The reference category is: SYRIZA.  
b. This parameter is set to zero because it is redundant.

Ετσι, θα λέγαμε ακριβολογώντας, αν δεν υπάρχει η επίδραση άλλων παραγόντων, ότι:

Οι ηλικίες 18-34 ετών, συγκρινόμενες με τις ηλικίες των 55 και άνω, εμφανίζουν μειωμένες πιθανότητες, να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ κατά 43%. Αυτό είναι το νοηματικό περιεχόμενο της σχέσης  $[1-\exp(-0,559)=1-0,57=0,43]$ , σε συνδυασμό βέβαια

με το αρνητικό πρόσημο του συντελεστή -0,559. Προφανώς αυτά που μόλις είπαμε για τη σχέση της ηλικιακής κατηγορίας 18-34 και 55 και άνω μπορούμε να τα πούμε κι αλλιώς. Πώς; Μα αναποδογυρίζοντας το λόγο. Να λάβουμε δηλ. τον αντίστροφο (reciprocal) λόγο, να εξετάσουμε δηλ. τη σχέση της ηλικιακής κατηγορίας των 55 και άνω, σε σύγκριση με την ηλικιακή κατηγορία των 18-34. Πράγματι ο αντίστροφος του 0,57 είναι ο  $1/0,57=1,75$  περίπου. Το συμπέρασμα επομένως είναι ότι η ηλικιακή κατηγορία των 55 και άνω, συγκρινόμενη με την ηλικιακή κατηγορία των 18-34 εμφανίζεται να έχει 1,75 μεγαλύτερες πιθανότητες να ψηφίζει ΝΔ, παρά ΣΥΡΙΖΑ.

Τα ίδια και σε μεγαλύτερο βαθμό ισχύουν για τις ηλικίες των 35-54. Δηλ. ηλικίες 35-54 ετών, συγκρινόμενες με τις ηλικίες των 55 και άνω, εμφανίζουν 49,7 % μειωμένες πιθανότητες, να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ. Αυτό είναι το νοηματικό περιεχόμενο της σχέσης  $[1-\exp(-0,688)= 1-0,50258023= 0,49741977=49,7]$ , σε συνδυασμό βέβαια με το αρνητικό πρόσημο του συντελεστή -0,668. Εργαζόμενοι όπως παραπάνω, καταλήγουμε στο συμπέρασμα ότι, η ηλικιακή κατηγορία των 55 και άνω, συγκρινόμενη με την ηλικιακή κατηγορία των 35-54, εμφανίζεται να έχει σχεδόν διπλάσια πιθανότητα να ψηφίζει ΝΔ, παρά ΣΥΡΙΖΑ. Αυτό είναι το νοηματικό περιεχόμενο της σχέσης  $[1/\exp(-0,688)= 1/0,49741977=1,9897321=2$  φορές περίπου].

**Πίνακας 5.17. Πίνακας Ταξινόμησης (Classification Table) για το Model 3.**

Observed	Classification									Percent Correct
	Predicted									
Observed	ND	SYRIZA	PASOK	POTAMI	KKE	XRYSH AYGH	ANEL	ALLO		Percent Correct
ND	154	163	0	0	0	0	0	0		48,6%
SYRIZA	138	280	0	0	0	0	0	0		67,0%
PASOK	35	19	0	0	0	0	0	0		0,0%
POTAMI	21	49	0	0	0	0	0	0		0,0%
KKE	18	44	0	0	0	0	0	0		0,0%
XRYSH AYGH	15	46	0	0	0	0	0	0		0,0%
ANEL	15	39	0	0	0	0	0	0		0,0%
ALLO	46	78	0	0	0	0	0	0		0,0%
Overall Percentage	38,1%	61,9%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%		37,4%

Όπως ο Πίνακας ταξινόμησης 5.17 μας πληροφορεί, το μοντέλο μας κάνει μια απεγνωσμένη απόπειρα να ανιχνεύσει τα Κόμματα στο διάβα του. Φαίνεται να μπορεί να καταφέρνει κάποια πράγματα για τα δυο μεγάλα Κόμματα εξουσίας, [48,6 % για τη

ΝΔ, και 67% για το ΣΥΡΙΖΑ, αλλά τίποτε για τα μικρότερα Κόμματα. Η δε μέση ανιχνευτική του δύναμη (Overall Percentage), ας το πούμε έτσι, δεν ξεπερνά το 37,4%. Να πούμε λοιπόν ότι έχουμε ένα χλωμό μοντελάκι; Να αγνοήσουμε τα όσα ισχυρίζεται; Να το πετάξουμε τελείως, στον καιάδα; Όχι βέβαια. Τίποτε από όλα αυτά. Διότι, όλα όσα προηγήθηκαν (Προσαρμογή, συντελεστές κτλ.) δεν συνηγορούν σε κάτι τέτοιο.

Ο τελευταίος Πίνακας 5.18 του Model 3 είναι ένας πίνακας μεγάλης σπουδαιότητας. Για το λόγο ότι ενέχει το στοιχείο της πρόβλεψης. Δεν είναι στατικός.

**ΕΡΩΤΗΜΑ:** Ας υποθέσουμε, ότι επιλέγουμε στην τύχη ένα άτομο το οποίο είναι άνδρας και ηλικίας μεταξύ 18 και 34 ετών. Ποια η πιθανότητα να ψηφίζει ΝΔ;

**ΑΠΑΝΤΗΣΗ:** Από τον Πίνακα 4.18 παρατηρούμε ότι η προβλεπόμενη πιθανότητα γι' αυτό το άτομο που είναι άνδρας, και που η ηλικία του είναι μεταξύ 18 και 34, να ψηφίζει ΝΔ, είναι 21,4%. Διότι,

$$\text{Predicted\_Percentage} = \frac{\text{Predicted}}{\text{total}} = \frac{25,502}{119} = 0,21430252 \approx 21,4\% \quad (5.8)$$

Η Ερμηνεία του Πίνακα 5.18 καθίσταται πλέον εύκολη:

Ας ερμηνεύσουμε, μια και είναι ατέλειωτος σε πληροφορίες, σκόρπια ορισμένους συνδυασμούς στο εσωτερικό του:

- I. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΝΔ, είναι 21,4%.
- II. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΣΥΡΙΖΑ, είναι 33,1%.
- III. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΠΑΣΟΚ, είναι 2,2%.
- IV. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΠΟΤΑΜΙ, είναι 8 %.
- V. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΚΚΕ, είναι 8,6 %.
- VI. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΧΑ, είναι 9%.
- VII. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει ΑΝΕΛ είναι 5,3%.
- VIII. Η πιθανότητα ένας άνδρας μεταξύ 18 και 34 να ψηφίζει κάτι ΆΛΛΟ, είναι 12,4%.
- IX. Η πιθανότητα μια γυναίκα μεταξύ 18 και 34 να ψηφίζει ΝΔ είναι 22,2%.
- X. Η πιθανότητα μια γυναίκα μεταξύ 18 και 34 να ψηφίζει ΣΥΡΙΖΑ είναι 35,3%.
- XI. Η πιθανότητα μια γυναίκα άνω των 35 και 54 να ψηφίζει ΝΔ είναι 23,7%.
- XII. Η πιθανότητα μια γυναίκα άνω των 35 και 54 να ψηφίζει ΣΥΡΙΖΑ είναι 42,9%.

**Πίνακας 5.18.** Παρατηρούμενες και Αναμενόμενες Συχνότητες για το Model 3.

**Observed and Predicted Frequencies**

CAT_AGE	SEX	PARTY	Observed	Predicted	Frequency	Percentage	
						Pearson Residual	Observed
18-34	MALE	ND	29	25,502	,781	24,4%	21,4%
		SYRIZA	36	39,432	-,668	30,3%	33,1%
		PASOK	4	2,568	,903	3,4%	2,2%
		POTAMI	6	9,506	-1,185	5,0%	8,0%
		KKE	11	10,225	,254	9,2%	8,6%
		XRYSH AYGH	10	10,669	-,215	8,4%	9,0%
		ANEL	7	6,301	,286	5,9%	5,3%
		ALLO	16	14,796	,334	13,4%	12,4%
	FEMALE	ND	22	25,498	-,785	19,1%	22,2%
		SYRIZA	44	40,568	,670	38,3%	35,3%
		PASOK	0	1,432	-1,204	0,0%	1,2%
		POTAMI	14	10,494	1,135	12,2%	9,1%
		KKE	10	10,775	-,248	8,7%	9,4%
		XRYSH AYGH	5	4,331	,328	4,3%	3,8%
		ANEL	5	5,699	-,300	4,3%	5,0%
		ALLO	15	16,204	-,323	13,0%	14,1%
35-54	MALE	ND	55	53,222	,277	23,3%	22,6%
		SYRIZA	98	93,617	,583	41,5%	39,7%
		PASOK	5	9,284	-1,434	2,1%	3,9%
		POTAMI	12	13,067	-,304	5,1%	5,5%
		KKE	10	10,629	-,197	4,2%	4,5%
		XRYSH AYGH	22	21,404	,135	9,3%	9,1%
		ANEL	15	13,506	,419	6,4%	5,7%
		ALLO	19	21,271	-,516	8,1%	9,0%
	FEMALE	ND	57	58,778	-,265	23,0%	23,7%
		SYRIZA	102	106,383	-,562	41,1%	42,9%
		PASOK	10	5,716	1,813	4,0%	2,3%
		POTAMI	17	15,933	,276	6,9%	6,4%
		KKE	13	12,371	,183	5,2%	5,0%
		XRYSH AYGH	9	9,596	-,196	3,6%	3,9%
		ANEL	12	13,494	-,418	4,8%	5,4%
		ALLO	28	25,729	,473	11,3%	10,4%
55 plus	MALE	ND	75	80,276	-,725	31,8%	34,0%
		SYRIZA	70	70,951	-,135	29,7%	30,1%
		PASOK	26	23,148	,624	11,0%	9,8%
		POTAMI	15	10,427	1,449	6,4%	4,4%
		KKE	9	9,147	-,049	3,8%	3,9%
		XRYSH AYGH	11	10,926	,023	4,7%	4,6%
		ANEL	6	8,193	-,780	2,5%	3,5%
		ALLO	24	22,933	,235	10,2%	9,7%
	FEMALE	ND	79	73,724	,767	38,3%	35,8%
		SYRIZA	68	67,049	,141	33,0%	32,5%
		PASOK	9	11,852	-,853	4,4%	5,8%
		POTAMI	6	10,573	-1,444	2,9%	5,1%
		KKE	9	8,853	,050	4,4%	4,3%
		XRYSH AYGH	4	4,074	-,037	1,9%	2,0%
		ANEL	9	6,807	,855	4,4%	3,3%
		ALLO	22	23,067	-,236	10,7%	11,2%

The percentages are based on total observed frequencies in each subpopulation.

## **Συζήτηση**

Αυτές οι δύο τελευταίες προβλέψεις, ασφαλώς και περιέχουν ένα σαφές μήνυμα. Υπάρχει πολύ μεγάλη διαφορά ανάμεσα σε αυτές και σε όλες τις υπόλοιπες προβλεπτικές προτάσεις. Οι μεσαίες ηλικίες 35-54 εμφανίζουν μια έντονη επιλεκτική συμπεριφορά, μια έντονη ροπή, μια ασυνήθιστη τάση προς το κόμμα του ΣΥΡΙΖΑ. Μα, στην περίπτωση των γυναικών αυτή η τάση είναι εμφανέστατα ενισχυμένη. Είναι 19,2 ποσοστιαίες μονάδες. Τι άραγε<sup>23</sup> να συμβαίνει; Το μοντέλο μίλησε. Και μπόρεσε να μιλήσει, αφού έχει καλή προσαρμογή, όπως είδαμε παραπάνω. Μόνο που το μήνυμά του αυτό, δεν είναι μόνο στατιστικό... Είναι και πολιτικό, κοινωνιολογικό, ίσως και είναι ζήτημα μαζικής ψυχολογίας. Εμείς βλέπουμε καθαρά ότι το νόημα, ιδιαίτερα της τελευταίας προβλεπτικής πρότασης, που αφορά τις γυναίκες της συγκεκριμένης ηλικιακής ζώνης, χρήζει περαιτέρω κοινωνιολογικής ανάλυσης.

### **5.6. Η κατασκευή του Model 4.**

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του MONTELOΥ 4.

MODEL 4.  
=====

Dependent Var: PARTY  
=====

Independent Vars(3): EDUC , CAT\_AGE , EDUC\*CAT\_AGE  
=====

Από το μενού Analyze, ακολουθούμε τη διαδρομή:

Analyze → Regression → Multinomial Logistic (κλικ),

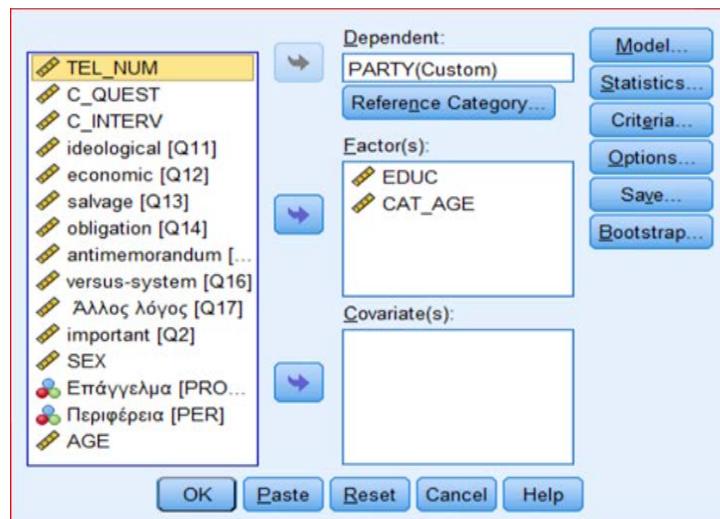
οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.14.

Στη συνέχεια τα παράθυρα των Εικόνων 5.15 - 5.17.

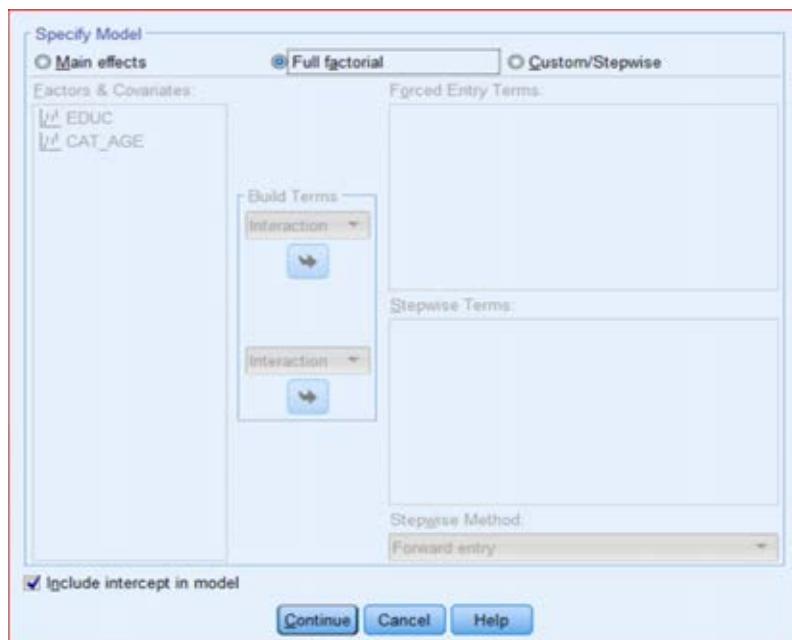
---

<sup>23</sup> Η έρευνά μας κατέγραψε τα λόγια μιας κυρίας από αυτή την ηλικιακή κατηγορία: των 35-54: 'Εγώ παιδί μου όταν θωρά αυτό το κοπέλι, θαρρώ πως θωρά το γιό μου... φρέσκο πρόσωπο που σούρχεται να το φιλήσει'. Και συνέχισε: 'Τέθοιοι αθρώποι θα πρέπει να κάτσουνε στο τιμόνι της χώρας...!'.

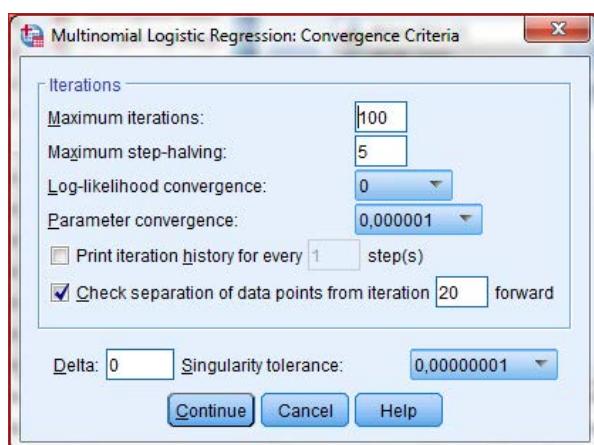
**Εικόνα 5.14.** Προσδιορισμός Μεταβλητών Model 4.



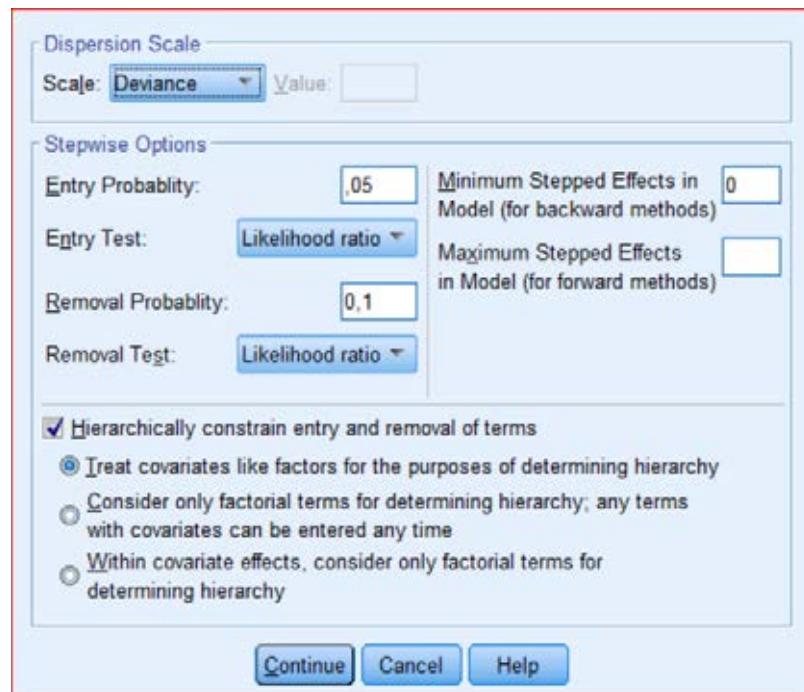
**Εικόνα 5.15.** Προσδιορισμός Model 4.



**Εικόνα 5.16.** Επιλογές μενού Criteria, Model 4.



**Εικόνα 5.17.** Επιλογές μενού Options, Model 4.



Στο παράθυρο της Εικόνας 5.17 κάνουμε κλικ στο κουμπί continue και έτσι επιστρέφουμε στο παράθυρο της Εικόνας 5.14, οπότε με κλικ στο κουμπί OK απογειώνουμε το αεροπλάνο της Multinomial Logistic Regression διαδικασίας.

**Πίνακας 5.19.** Σύγκριση Model 4 με Null μοντέλο.

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	339,617	375,010	325,617			
Final	323,941	607,087	211,941	113,675	49	<.001

Τα αποτελέσματα τα βλέπουμε στους Πίνακες 5.19, 5.20 και 5.21.

Ο Πίνακας 5.19 μας πληροφορεί ότι το Model 4 δικαιολογεί την ύπαρξή του. Έχουν σχέση οι ανεξάρτητες μεταβλητές του μορφωτικού επιπέδου EDUC και των ηλικιακών κατηγοριών CAT\_AGE, δηλ. σχετίζονται με την εξαρτημένη μεταβλητή του Πολιτικού Κόμματος PARTY.

**Πίνακας 5.20.** Προσαρμογή (goodness-of-fit) του Model 4.

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	,000	0	.
Deviance	,000	0	.

Ο Πίνακας 5.20 μας πληροφορεί ότι το μοντέλο μας είναι άγνωστης προσαρμογής.

**Πίνακας 5.21.** Σύγκριση των Reduced Model 4 με Null μοντέλο.

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	323,941	607,087	211,941 <sup>a</sup>	,000	0	.
EDUC	323,941	607,087	211,941 <sup>a</sup>	,000	0	.
CAT_AGE	323,941	607,087	211,941 <sup>a</sup>	,000	0	.
EDUC * CAT_AGE	294,080	471,046	224,080	12,139	21	,936

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Τέλος, ο Πίνακας 5.21 μας πληροφορεί ότι ΔΕΝ υπάρχει αλληλεπίδραση μεταξύ μορφωτικού επιπέδου και ηλικιακών κατηγοριών. Το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι υψηλό, πολύ μεγαλύτερο οπωσδήποτε του 5 % (Sig.=0,936 = 93,6 % > 5%).

### Συμπέρασμα

Δεν υπάρχει αλληλεπίδραση ανάμεσα στους predictors EDUC και CAT\_AGE.

### 1η Σημαντική επισήμανση για τα επόμενα

Το Model 4 έδειξε ότι, interaction δεύτερης τάξης, μεταξύ μορφωτικού επιπέδου και ηλικιακών κατηγοριών, δεν υπάρχει. Η μεταβλητή αλληλεπίδρασης EDUC\*CAT\_AGE αποδείχθηκε στατιστικώς ασήμαντη δηλ. ο πληθυσμιακός παλινδρομικός συντελεστής (regression coefficient) είναι μηδέν.

### 2η Σημαντική επισήμανση για τα επόμενα

Το Model 4 είναι ένα μοντέλο χωρίς ερμηνευσιμότητα. Διότι δεν έχει τυπωμένα, παρατηρούμενα επίπεδα για τις μεταβλητές του μορφωτικού επιπέδου (EDUC) και των ηλικιακών κατηγοριών (CAT\_AGE) στην τελευταία στήλη του Πίνακα 5.21.

## 5.7. Η κατασκευή του Model 5.

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του MONTEΛΟΥ 5.

MODEL 5.

=====

Dependent Var: PARTY

=====

Independent Vars(2): SEX, EDUC, SEX\* EDUC

=====

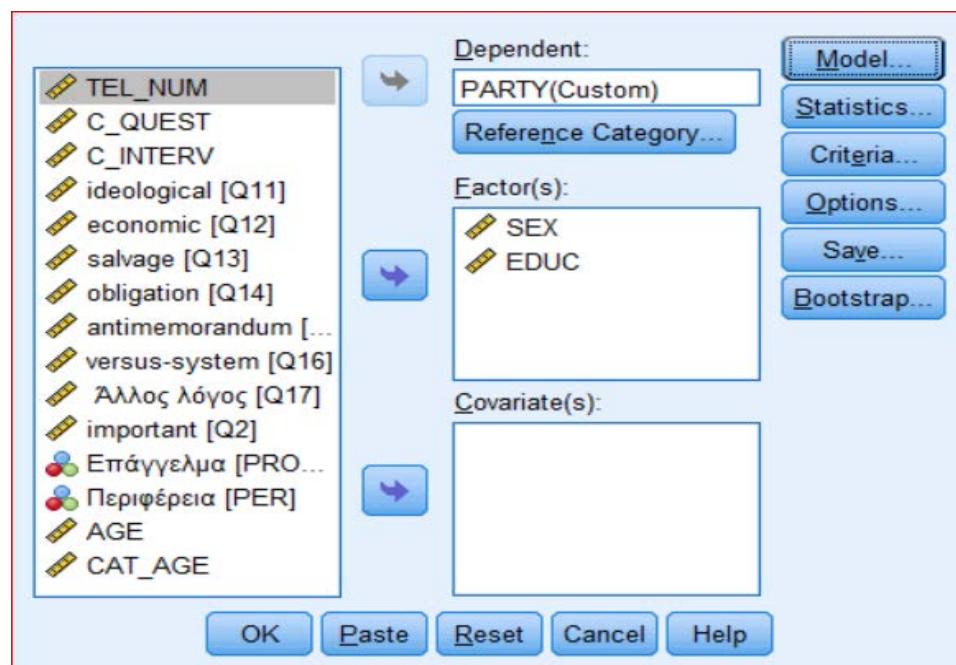
Από το μενού Analyze, ακολουθούμε τη διαδρομή:

Analyze → Regression → Multinomial Logistic (κλικ),

οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.18.

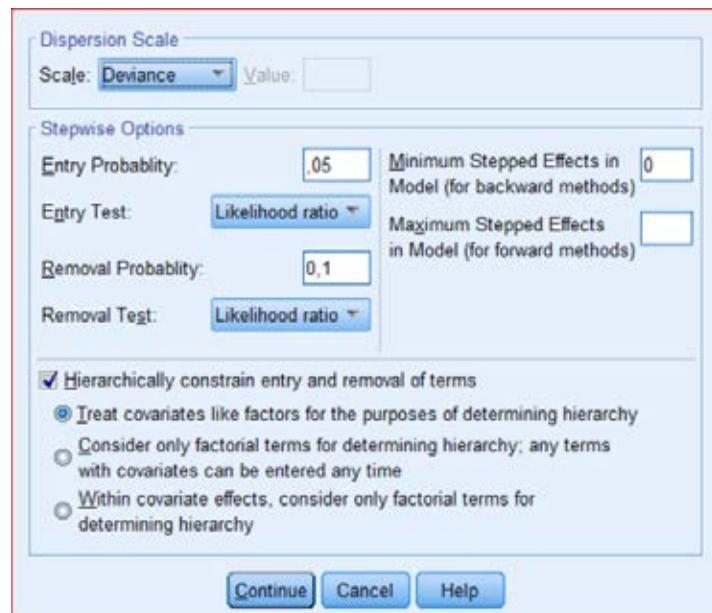
Στη συνέχεια λαμβάνουμε το παράθυρο της Εικόνας 5.19.

**Εικόνα 5.18.** Προσδιορισμός Μεταβλητών Model 5.



Ευρισκόμενοι στο παράθυρο της Εικόνας 5.19, πατάμε το κουμπί Continue και στη συνέχεια το κουμπί OK. Τα αποτελέσματα τα βλέπουμε στους Πίνακες 5.22, 5.23 και 5.24.

**Εικόνα 5.19.** Επιλογές μενού Options, Model 5.



**Πίνακας 5.22.** Σύγκριση Model 5 με Null μοντέλο.

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	279,833	315,257	265,833			
Final	253,279	465,819	169,279	96,555	35	<,001

**Πίνακας 5.23.** Προσαρμογή (goodness-of-fit) του Model 5.

Goodness-of-Fit		
	Chi-Square	df
Pearson	,000	0
Deviance	,000	0

**Πίνακας 5.24.** Σύγκριση των Reduced Model 5 με Null μοντέλο.

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Likelihood Ratio Tests		
				Chi-Square	df	Sig.
Intercept	253,279	465,819	169,279 <sup>a</sup>	,000	0	.
SEX	253,279	465,819	169,279 <sup>a</sup>	,000	0	.
EDUC	253,279	465,819	169,279 <sup>a</sup>	,000	0	.
SEX*EDUC	255,886	397,580	199,886	30,608	14	,006

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Ο Πίνακας 5.22 μας πληροφορεί ότι το μοντέλο μας (final) έχει λόγο ύπαρξης, καθότι έχει στατιστικώς σημαντική διαφορά με το μοντέλο του σταθερού όρου (intercept only), και ότι οι ανεξάρτητες μεταβλητές του παρουσιάζουν συσχέτιση με την εξαρτημένη μεταβλητή PARTY του μοντέλου μας (Model 5).

Ο Πίνακας 5.23 μας πληροφορεί ότι το μοντέλο μας είναι άγνωστης προσαρμογής.

Και τέλος, ο Πίνακας 5.24 μας πληροφορεί ότι οι μεταβλητές του φύλου και της μόρφωσης παρουσιάζουν αλληλεπίδραση, ή αλλιώς η μεταβλητή της αλληλεπίδρασης SEX\*EDUC έχει μη μηδενικό συντελεστή (Sig.= 0,006 = 0,6% <5%). Η τελευταία αυτή η πληροφορία είναι κρίσιμη και θα πρέπει άμεσα να αξιολογηθεί. Πράγματι, η συμπεριφορά των δυο μεταβλητών του φύλου και του μορφωτικού επιπέδου συσκοτίζεται. Αυτό παραπέρα σημαίνει ότι ούτε για τη μία, ούτε για την άλλη, μπορεί να δοθεί στατιστική ερμηνεία, όταν η άλλη παραμένει σταθερά.

### **Επιτόπιο συμπέρασμα για το Model 5**

Το Model 5, δεν μπορεί να γίνει αποδεκτό όχι μόνο γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής, (βλ. Πίνακα 5.23), -αν και αυτό θα έφτανε-, αλλά και γιατί η κεφαλαιώδους σημασίας για την έρευνά μας μεταβλητή του φύλου, στο πλαίσιο του, δεν μπορεί εύκολα να ερμηνευθεί.

### **5.8. Η κατασκευή του Model 6.**

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του MONTELOΥ 6.

MODEL 6.  
=====

Dependent Var: PARTY  
=====

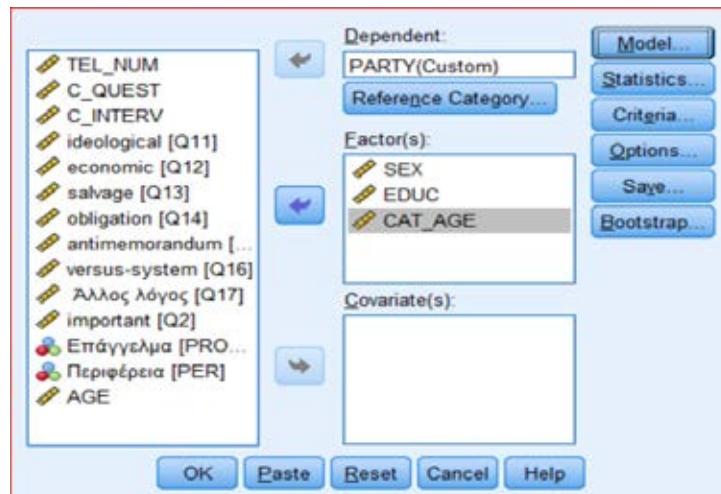
Independent Vars(4): SEX, EDUC, CAT\_AGE , SEX\*EDUC\*CAT\_AGE  
=====

Από το μενού Analyze, ακολουθούμε τη διαδρομή:

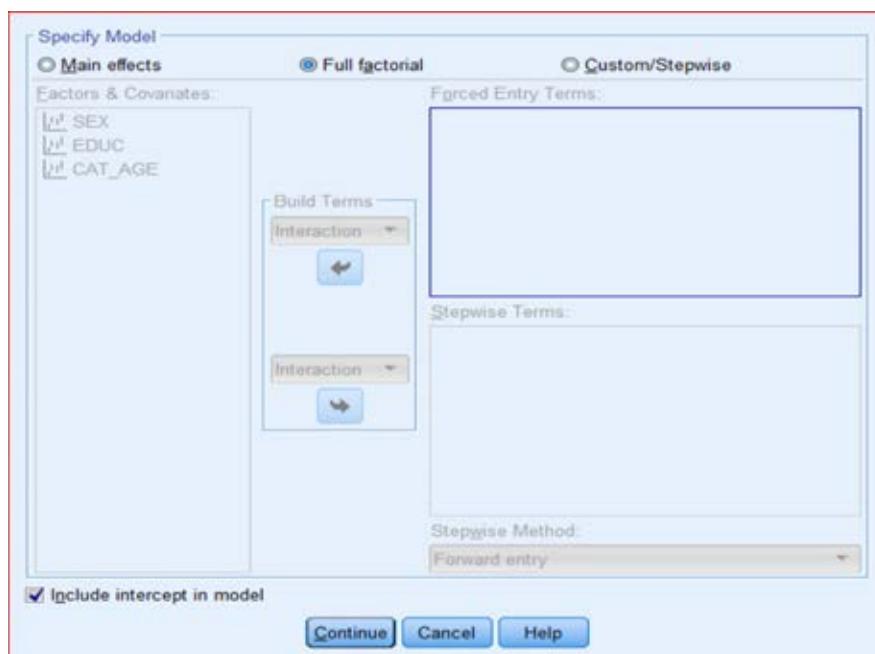
Analyze → Regression →Multinomial Logistic (κλικ),

οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.20. και στη συνέχεια τα παράθυρα των Εικόνων 5.21 και 5.22.

**Εικόνα 5.20.** Προσδιορισμός Μεταβλητών Model 6.



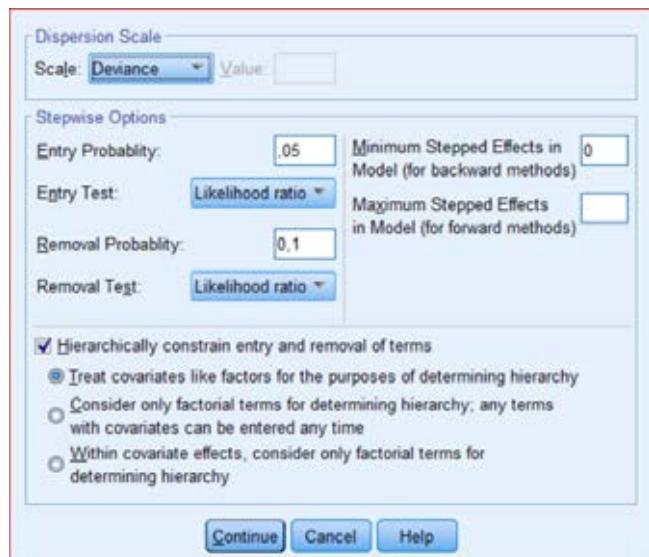
**Εικόνα 5.21.** Προσδιορισμός Model 6.



Στην Εικόνα 5.21 δηλώνουμε Full factorial και στη συνέχεια πατάμε continue... πατάμε το κουμπί Options οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.22. Ρυθμίζουμε στη συνέχεια τα κουμπιά Statistics, Criteria, κατά τα γνωστά. Τα αποτελέσματα τα βλέπουμε στους Πίνακες 5.25, 5.26 και 5.27.

Ευρισκόμενοι στο παράθυρο

**Εικόνα 5.22.** Επιλογές μενού Options, Model 6.



**Πίνακας 5.25.** Σύγκριση Model 6 με Null μοντέλο.

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	550,716	586,103	536,716			
Final	556,315	1122,510	332,315	204,401	105	<,001

**Πίνακας 5.26.** Προσαρμογή (goodness-of-fit) του Model 6.

Goodness-of-Fit		
	Chi-Square	df
Pearson	,000	0
Deviance	,000	0

**Πίνακας 5.27.** Σύγκριση των Reduced Model 6 με Null μοντέλο.

Effect	Likelihood Ratio Tests			Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.			
Intercept	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
SEX	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
EDUC	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
CAT_AGE	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
SEX * EDUC	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
SEX * CAT_AGE	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
EDUC * CAT_AGE	556,315	1122,510	332,315 <sup>a</sup>	,000	0	,			
SEX * EDUC * CAT_AGE	533,105	993,138	351,105	18,790	21	,599			

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model.  
The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Ο Πίνακας 5.25 μας πληροφορεί ότι το μοντέλο μας (final) έχει λόγο ύπαρξης, καθότι έχει στατιστικώς σημαντική διαφορά με το μοντέλο του σταθερού όρου (intercept only), και ότι οι ανεξάρτητες μεταβλητές του παρουσιάζουν συσχέτιση με την εξαρτημένη μεταβλητή PARTY του μοντέλου μας (Model 6).

Ο Πίνακας 5.26 μας λέει ότι το μοντέλο μας είναι άγνωστης προσαρμογής.

Η τελευταία γραμμή του Πίνακα 5.27 μας μεταφέρει την κρίσιμη πληροφορία ότι οι μεταβλητές του φύλου, μόρφωσης και των ηλικιακών κατηγοριών, δεν παρουσιάζουν αλληλεπίδραση, ή αλλιώς η μεταβλητή της αλληλεπίδρασης SEX \* EDUC \* CAT\_AGE έχει μηδενικό συντελεστή, αφού το αντίστοιχο επίπεδο στατιστικής σημαντικότητας είναι ασήμαντο (Sig. = 0,599 = 59,9% > 5%).

### **Επιτόπιο συμπέρασμα**

Το Model 6, δεν μπορεί να γίνει αποδεκτό όχι μόνο γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής, (βλ. Πίνακα 5.26), -αν και αυτό θα έφτανε-, αλλά και γιατί δεν μπορεί εύκολα να ερμηνευθεί.

### **Κρίσιμη Παρατήρηση**

Στον Πίνακα 5.24 είδαμε ότι αποκλείστηκε το interaction 3ης τάξης, SEX \* EDUC \* CAT\_AGE, διότι ο συντελεστής παλινδρόμησης γι' αυτήν την αλληλεπίδραση είναι μηδενικός. Και είναι μηδενικός διότι το αντίστοιχο επίπεδο στατιστικής σημαντικότητας είναι πολύ μεγαλύτερο του 5 %. Αυτό παραπέρα μας παρέχει μια πολύ σημαντική πληροφορία: Αφού δεν έχουμε interaction 3ης τάξης, δεν θα έχουμε ούτε interaction 4ης τάξης, 5ης τάξης, 6ης τάξης κτλ. Γενικά δεν θα έχουμε interaction ανώτερης τάξης. Αυτό πρακτικά σημαίνει ότι καθόλου δεν χρειάζεται να ελέγχουμε άλλα ιεραρχικά μοντέλα. Επομένως ο δρόμος για το ΑΡΙΣΤΟ ΜΟΝΤΕΛΟ (THE BEST MODEL), απλοποιήθηκε, έγινε πιο βατός, λιγότερο σύνθετος, έγινε απλούστερη η διαδρομή. Έχουμε να κάνουμε πλέον στο εξής με αλληλεπιδράσεις δευτέρας τάξεως και μόνον.

### 5.8.1. Multinomial Logistic Regression Analysis: Το βαθύτερο νόημα της Αλληλεπίδρασης (Interaction)

Στην παράγραφο αυτή, καθώς διαπιστώθηκε η ύπαρξη αλληλεπίδρασης ανάμεσα στις μεταβλητές του φύλου SEX και του μορφωτικού επιπέδου EDUC, μας δίνεται η ευκαιρία να διεισδύσουμε βαθύτερα στην έννοια της Αλληλεπίδρασης (Interaction). Το εγχείρημα αυτό αποκτά ιδιαίτερο ενδιαφέρον αφού πρόκειται για κατηγορικές μεταβλητές, οι οποίες συμμετέχουν ως ανεξάρτητες μεταβλητές σε Logistic Models.

Εκτελούμε άμεσα διαδικασία Crosstabs, και το αποτέλεσμα το βλέπουμε στον Πίνακα 5.28.

**Πίνακας 5.28.** Συσχέτιση μεταβλητών EDUC και SEX.

EDUC * SEX Crosstabulation				
		SEX		Count
		MALE	FEMALE	
EDUC	ELEMENTARY	106	176	282
	MIDDLE	320	379	699
	HIGHER	421	506	927
	Total	847	1061	1908

Λαμβάνουμε τους φυσικούς λογάριθμους όλων των κελιών του Πίνακα 5.28, και στη συνέχεια οικοδομούμε τις αντίστοιχες εξισώσεις.

Ετσι, για το κελί (1,1) του Πίνακα 5.28 έχουμε:

$\ln(106)=4,66$  κι επομένως η αντίστοιχη εξισωση έχει ως εξής:

$$4,66 = \ln(106) = \mu + \hat{\lambda}^{\text{MALE}} + \hat{\lambda}^{\text{ELEMENTARY}} + \hat{\lambda}^{\text{MALE*ELEMENTARY}} \quad (5.9)$$

Επίσης, επειδή  $\ln(176)=5,17$ , η αντίστοιχη εξισωση έχει ως εξής:

$$5,17 = \ln(176) = \mu + \hat{\lambda}^{\text{FEMALE}} + \hat{\lambda}^{\text{ELEMENTARY}} + \hat{\lambda}^{\text{FEMALE*ELEMENTARY}} \quad (5.10)$$

Επίσης, επειδή  $\ln(320)=5,77$ , η αντίστοιχη εξισωση έχει ως εξής:

$$5,77 = \ln(320) = \mu + \hat{\lambda}^{\text{MALE}} + \hat{\lambda}^{\text{MIDDLE}} + \hat{\lambda}^{\text{MALE*MIDDLE}} \quad (5.11)$$

Επίσης, επειδή  $\ln(379)=5,94$ , η αντίστοιχη εξισωση έχει ως εξής:

$$5,94 = \ln(379) = \mu + \hat{\lambda}^{\text{FEMALE}} + \hat{\lambda}^{\text{MIDDLE}} + \hat{\lambda}^{\text{FEMALE*MIDDLE}} \quad (5.12)$$

Επίσης, επειδή  $\ln(421)=6,00$ , η αντίστοιχη εξίσωση έχει ως εξής:

$$6,00 = \ln(421) = \mu + \hat{\lambda}^{\text{MALE}} + \hat{\lambda}^{\text{HIGHER}} + \hat{\lambda}^{\text{MALE*HIGHER}} \quad (5.13)$$

Επίσης, επειδή  $\ln(506)=6,22$ , η αντίστοιχη εξίσωση έχει ως εξής:

$$6,22 = \ln(506) = \mu + \hat{\lambda}^{\text{FEMALE}} + \hat{\lambda}^{\text{HIGHER}} + \hat{\lambda}^{\text{FEMALE*HIGHER}} \quad (5.14)$$

Με βάση τις σχέσεις (5.9), (5.10), (5.11), (5.12), (5.13) και (5.14) ο Μεγάλος Μέσος (Grand Mean=μ) διαμορφώνεται ως εξής:

$$\begin{aligned} \mu &= \text{intercept} = \text{Grand\_Mean} = \\ &= \frac{\text{sum\_of\_logs\_of\_cell\_counts\_in\_entire\_table}}{\text{number\_of\_logs}} = \\ &= \frac{\ln(106) + \ln(176) + \ln(320) + \ln(379) + \ln(421) + \ln(506)}{6} = \frac{33,8}{6} = 5,63 \end{aligned} \quad (5.15)$$

Υπολογίζουμε τώρα τον περιθώριο μέσο όρο marginal mean για τη στήλη των ανδρών MALE:

$$\text{Marginal\_mean\_MALE} = \frac{\ln(106) + \ln(320) + \ln(421)}{3} = 5,5 \quad (5.16)$$

Υποθέτοντας στη συνέχεια, ότι έχουμε κορεσμένο μοντέλο (Saturated<sup>24</sup> Model), μπορούμε να υπολογίσουμε τις κύριες επιδράσεις (main effects), ως εξής: Βρίσκουμε τη διαφορά ανάμεσα στον Μεγάλο Μέσο (Grand Mean), και στον περιθώριο μέσο (marginal mean) μιας στήλης ή μιας γραμμής (Πίνακας 5.28).

Επιλέγουμε εδώ τη στήλη των ανδρών MALE. Θα έχουμε:

$$\begin{aligned} \hat{\lambda}^{\text{MALE}} &= \text{Difference}(\text{Grand\_mean}, \text{Marginal_mean_of_MALED}) = \\ &= 5,5 - 5,63 = -0,13 \end{aligned} \quad (5.17)$$

Υπολογίζουμε τώρα τον περιθώριο μέσο όρο Marginal mean για τη γραμμή ELEMENTARY.

$$\text{Marginal_mean_ELEMENTARY} = \frac{\ln(106) + \ln(176)}{2} = 4,91 \quad (5.18)$$

---

<sup>24</sup> Saturated Model λέμε αυτό που περιλαμβάνει όλες τις δυνατές κύριες επιδράσεις (main effects), και αλληλεπιδράσεις (interactions effects).

Βρίσκουμε, όπως και παραπάνω τη διαφορά ανάμεσα στον Μεγάλο Μέσο (Grand Mean), και στον περιθώριο μέσο (Marginal mean) της γραμμής ELEMENTARY. (βλ. Πίνακα 5.28). Θα έχουμε:

$$\hat{\lambda}^{\text{ELEMENTARY}} = 4,91 - 5,63 = -0,72 \quad (5.19)$$

Εστιάζουμε τώρα την προσοχή μας στην εξίσωση που γράψαμε παραπάνω, την (5.9).

Παρατηρούμε εκεί ότι έχουμε υπολογίσει:

- I. Το φυσικό λογάριθμο του 106
- II. Τον μεγάλο μέσο Grand Mean = intercept
- III. Την κύρια επίδραση για το επίπεδο MALE
- IV. Την κύρια επίδραση για το επίπεδο ELEMENTARY

και ο μόνος άγνωστος που μας απομένει, είναι η ποσότητα:  $\hat{\lambda}^{\text{MALE*ELEMENTARY}}$

η οποία είναι και το μεγάλο ζητούμενο, επειδή ακριβώς αναπαριστά το effect της αλληλεπίδρασης.

Οπότε αντικαθιστώντας στην (5.9) τις τιμές των παραπάνω θα έχουμε:

$$4,66 = 5,63 - 0,13 - 0,72 + \hat{\lambda}^{\text{MALE*ELEMENTARY}} \quad (5.20)$$

Αν λύσουμε την (5.20) ως προς  $\hat{\lambda}^{\text{MALE*ELEMENTARY}}$  θα έχουμε:

$$\hat{\lambda}^{\text{MALE*ELEMENTARY}} = 4,66 - 5,63 + 0,13 + 0,72 = -0,12$$

Τελικά για το effect της αλληλεπίδρασης έχουμε:

$$\hat{\lambda}^{\text{MALE*ELEMENTARY}} = -0,12 \quad (5.21)$$

### Ερμηνεία

Το με αρνητικό πρόσημο interaction -0,12 που βρήκαμε, σημαίνει ότι λιγότεροι Άνδρες από όσους περιμέναμε να έχουμε βρίσκονται στην κατώτατη εκπαιδευτική βαθμίδα της Στοιχειώδους Εκπαιδεύσεως. Σημαίνει, αυτό το 12% ότι, ακριβώς 12% είναι η διαφορά ανάμεσα στα αθροίσματα των effects των μεταβλητών, όταν αυτά τα αθροίσματα λαμβάνονται ατομικά/ξεχωριστά και όλα μαζί/συλλογικά.

## 5.9. Η κατασκευή του Model 7.

Κρίνουμε σκόπιμο, να υπενθυμίσουμε τη δομή του ΜΟΝΤΕΛΟΥ 7.

MODEL 7.

=====

Dependent Var: PARTY

=====

Independent Vars(10): SEX, CAT\_AGE, EDUC, Q2, Q11, Q12, Q13, Q14, Q15 και Q16.

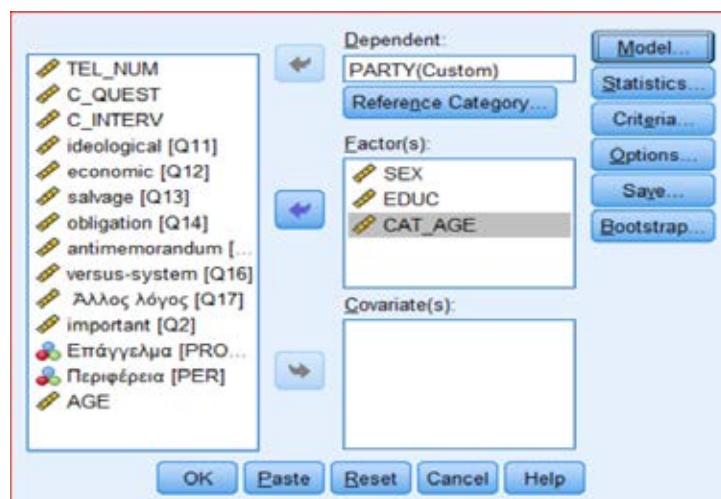
=====

Από το μενού Analyze, ακολουθούμε τη διαδρομή:

Analyze → Regression → Multinomial Logistic (κλικ),

οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.23.

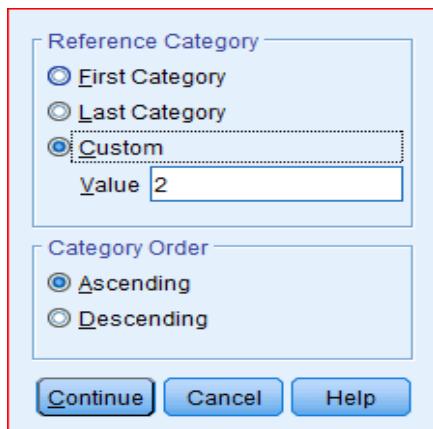
**Εικόνα 5.23.** Προσδιορισμός Μεταβλητών Model 7.



Ευρισκόμενοι στο παράθυρο της Εικόνας 5.23 κάνουμε εισαγωγή στο πλαίσιο Factors των μεταβλητών Q11, Q12, Q13, Q14, Q15, Q16, Q2, CAT\_AGE, EDUC, SEX και επίσης κάνουμε εισαγωγή της εξαρτημένης μας μεταβλητής PARTY στο πλαίσιο Dependent.

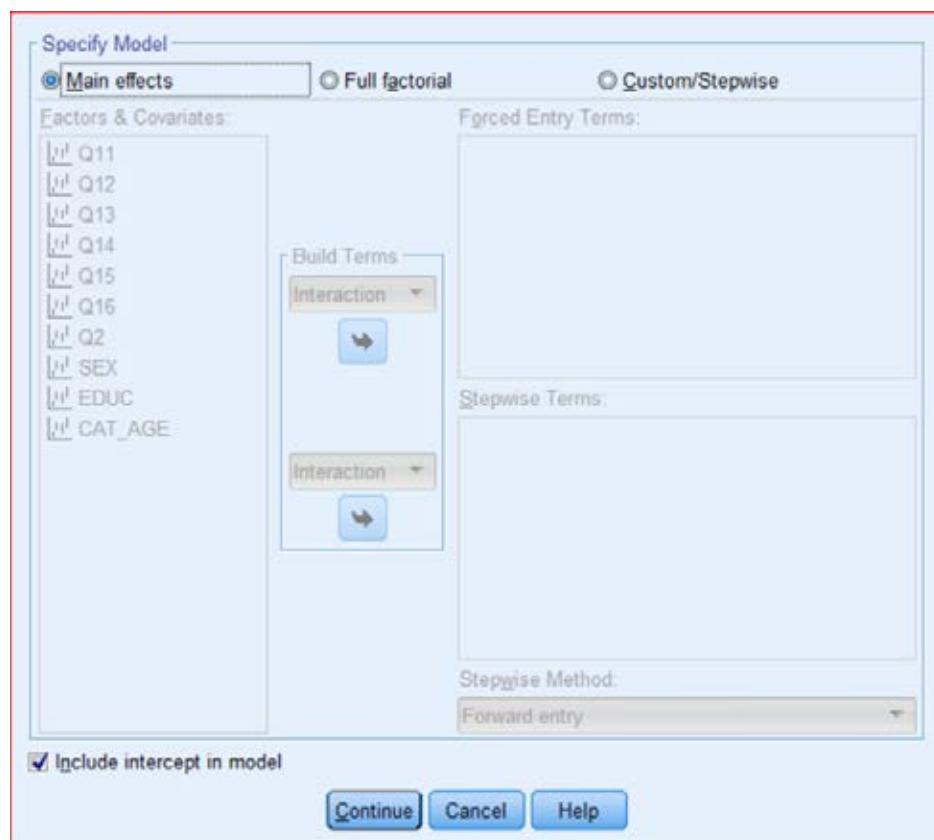
Στη συνέχεια, πατάμε το κουμπί Reference Category, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.24.

**Εικόνα 5.24.** Επιλογή Κατηγορίας Αναφοράς Model 7.



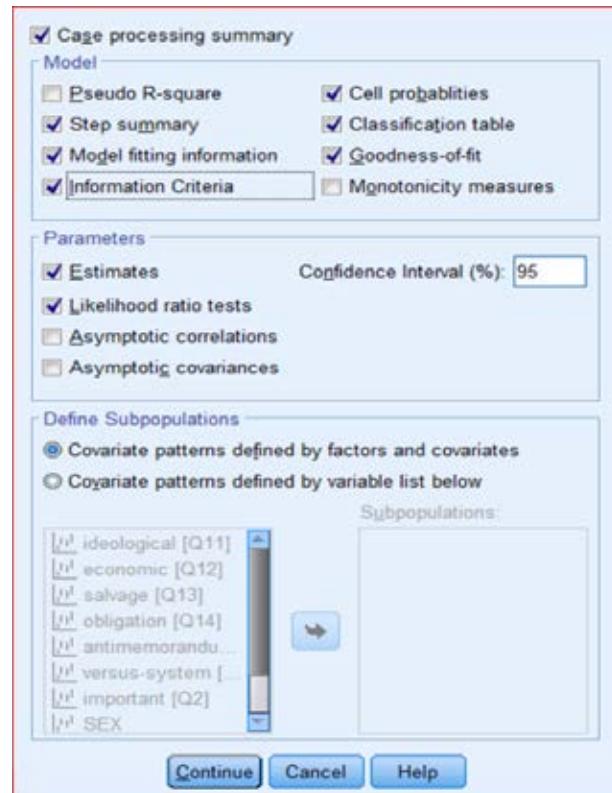
Ευρισκόμενοι στο παράθυρο της Εικόνας 5.24 τσεκάρουμε την ένδειξη Custom, και ακολούθως θέτουμε τον αριθμό 2 μέσα στο ορθογώνιο που έχει στα αριστερά του τη λέξη Value, οπότε σαν κατηγορία αναφοράς (baseline) θα έχουμε τον ΣΥΡΙΖΑ ο οποίος έχει κωδικοποιηθεί έτσι. Συμβολικά: 2=ΣΥΡΙΖΑ. Κατόπιν, με ένα κλικ στο κουμπί Continue ξαναγυρίζουμε στο βασικό μας μενού (βλ. Εικόνα 5.23), και πατάμε το κουμπί Model. Άμεσα θα λάβουμε το παράθυρο της Εικόνας 5.25.

**Εικόνα 5.25.** Προσδιορισμός Model 7.



Ευρισκόμενοι στο περιβάλλον της Εικόνας 5.25 επιλέγουμε main effects διότι επιθυμούμε να ελέγξει το SPSS τις κύριες επιδράσεις των 10 ανεξάρτητων κατηγορικών μεταβλητών, και μόνον αυτές. Με ένα κλικ επίσης στο πλαίσιο continue γυρίζουμε στο βασικό μας menue και στη συνέχεια, πατάμε το κουμπί Statistics , οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.26.

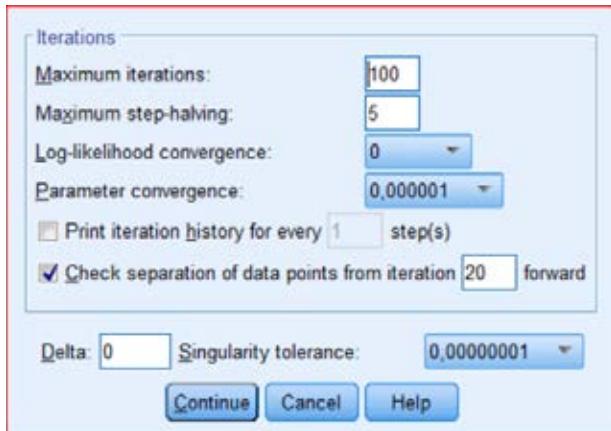
**Εικόνα 5.26.** Επιλογές μενού Statistics Model 7.



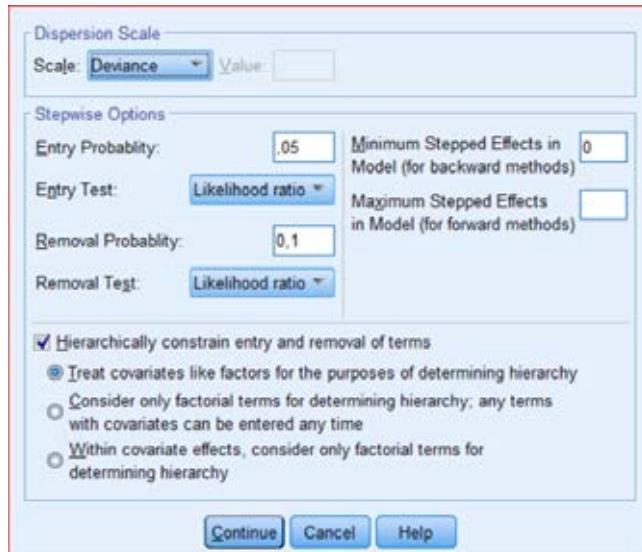
Ευρισκόμενοι στο περιβάλλον της Εικόνας 5.26 τσεκάρουμε κατάλληλα, ώστε το SPSS να ελέγξει τα παρακάτω:

- I. Την προσαρμογή του μοντέλου μας (Goodness-of-fit).
- II. Την αξιολόγηση/αιτιολόγηση ύπαρξης του μοντέλου μας και τη διερεύνηση της σχέσης μεταξύ εξαρτημένης και ανεξάρτητων μεταβλητών (Model fitting information).
- III. Υπολογισμός δεικτών AIC και BIC (Information Criteria).
- IV. Την ικανότητα Πρόβλεψης του μοντέλου μας (Classification table).
- V. Εκτιμήσεις πολινδρομικών συντελεστών (Estimates).
- VI. Πιθανότητες Πρόβλεψης (Cell Probabilities).

**Εικόνα 5.27.** Επιλογές μενού Criteria, Model 7.



**Εικόνα 5.28.** Επιλογές μενού Options, Model 7.



Στη συνέχεια, με κλικ στο Continue επιστρέφουμε στο βασικό μας μενού. Εκεί κάνουμε κλικ στο κουμπί Criteria, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.27.

Ευρισκόμενοι στο παράθυρο της Εικόνας 5.27 λαμβάνουμε μέριμνα έτσι ώστε να παραμείνουν αναλλοίωτες οι ρυθμίσεις για τις παραμέτρους Delta και Singularity tolerance.

Στη συνέχεια, με κλικ στο Continue επιστρέφουμε στο βασικό μας μενού. Εκεί κάνουμε κλικ στο κουμπί Criteria, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.28.

Ευρισκόμενοι στο παράθυρο της Εικόνας 5.28 παραδίνουμε τον έλεγχο Entry Test και removal Test στο κριτήριο Likelihood ratio test επειδή ακριβώς πιστεύουμε αυτό έχει

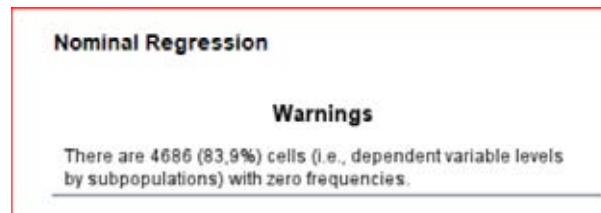
συγκριτικά καλύτερη απόδοση από το Wald test, αναφορικά με τις επιδράσεις (effects). (Norusis, 2010).

Επιστρέφουμε στο βασικό μας μενού κάνοντας κλικ στο Continue και στη συνέχεια με κλικ στο OK, τρέχουμε το MONTELO 7.

### ΑΠΟΤΕΛΕΣΜΑΤΑ για το MODEL 7

Προτού παραθέσουμε και σχολιάσουμε ένα προς ένα τους Πίνακες που παράγει η διαδικασία Multinomial Logistic Regression θα προβούμε σε σύντομο σχολιασμό του προειδοποιητικού μηνύματος που εκείνη παράγει στην αρχή:

**Εικόνα 5.29.** Προειδοποιητικό μήνυμα, Model 7.



Υπάρχουν μας λέει το μήνυμα αυτό 4686 κελιά, ποσοστό της τάξης του 83,9 % , στην ανάλυσή μας, τα οποία είναι με μηδενικές συχνότητες, στον πολυδιάστατο πίνακα διασταύρωσης μεταξύ εξαρτημένης και ανεξάρτητων μεταβλητών. Στο πλαίσιο ωστόσο, της πολυωνυμικής λογαριθμογραμμικής παλινδρόμησης εμπλέκονται πολλές ανεξάρτητες μεταβλητές, και ακόμη πολλές κατηγορίες (επίπεδα) σε αυτές και στην εξαρτημένη, γίνονται ατέλειωτοι συνδυασμοί, ατέλειωτες διασταυρώσεις επιπέδων, και είναι φυσικό, όσο καλή Δειγματοληψία και να κάνουμε- και αποδεικνύεται ότι κάναμε- να υπάρχουν κενά κελιά, με ελλείπουσες τιμές, δηλ. με μηδενικό αριθμό παρατηρήσεων, ενώ δεν θα έπρεπε . Καθόλου δεν αποθαρρυνόμαστε. Με τέτοιους πίνακες, υπερβολικά μεγάλων διαστάσεων, και με δεδομένο ότι είναι αδύνατο να συλλέξεις κοινωνικά δεδομένα για να πληρωθεί αυτός ο τεράστιος, ενίστε μέγας αριθμών κελιών, είναι φυσιολογικό να υπάρχουν ανάλογα προειδοποιητικά μηνύματα.

Εστιάζουμε τώρα την προσοχή μας στον Πίνακα 5.29 είναι (Nominal Regression<sup>25</sup>), είναι ένας πίνακας πρόβλεψης, οποίος έχει στην τελευταία του στήλη τις περιθώριες εκατοστιαίες συχνότητες (marginal percentage) που αντιστοιχούν στα εκτιμούμενα ποσοστά των Κομμάτων που έλαβαν μέρος στις εκλογές της 25<sup>η</sup> Ιανουαρίου του 2015. Ετσι, το συγκεκριμένο μοντέλο (Model 7), προβαίνει στις παραπάνω εκτιμήσεις:

- I. Για τη ΝΔ ποσοστό 27,6%. (Στην πραγματικότητα έλαβε 27,81%).
- II. Για τον ΣΥΡΙΖΑ ποσοστό 36,7%. (Στην πραγματικότητα έλαβε 36,34%).
- III. Για το ΠΑΣΟΚ ποσοστό 4,4%. (Στην πραγματικότητα έλαβε 4,68%).
- IV. Για το ΠΟΤΑΜΙ ποσοστό 6%. (Στην πραγματικότητα έλαβε 6,05%).
- V. Για το ΚΚΕ ποσοστό 5,3%. (Στην πραγματικότητα έλαβε 5,47%).
- VI. Για τη ΧΡΥΣΗ ΑΥΓΗ ποσοστό 5,1%. (Στην πραγματικότητα<sup>26</sup> έλαβε 6,28%).
- VII. Για τους ΑΝΕΛ ποσοστό 4,4%. (Στην πραγματικότητα έλαβε 4,75%).
- VIII. Για τα υπόλοιπα άλλα Κόμματα μαζί ποσοστό 10,4%. (Στην πραγματικότητα αυτά τα Κόμματα έλαβαν: ΚΙΔΗΣΟ 2,47%, ΕΚ 1,79%, ΤΕΛΕΙΑ 1,77%, ΛΑΟΣ 1,03%, ΑΝΤΑΡΣΥΑ 0,64%, ΚΚΕ (μ-λ) 0,48%, ΜΛ ΚΚΕ 0,13%).

Ο Πίνακας 5.29 είναι ένας πίνακας περιθώριων συχνοτήτων (marginal effects), που δείχνει την κατά το Model 7, πρόβλεψη ανά πολιτικό Κόμμα.

Ο Πίνακας 5.30 είναι ένας χρήσιμος πίνακας με κουβαλάει στο εσωτερικό του μια βασική διαπίστωση, γύρω από το εάν έχει νόημα να ασχοληθούμε με το συγκεκριμένο μοντέλο. (Model 7). Για να εκφραστούμε με μεγαλύτερη ακρίβεια συγκρίνει το μοντέλο μας, το Model 7, το οποίο ονομάζει final, με το μοντέλο που έχει μόνο το σταθερό όρο (concept), και αποφαίνεται, αν αυτό είναι σημαντικά καλύτερο από εκείνο. Με άλλα λόγια, ελέγχει την παρακάτω μηδενική υπόθεση :

H<sub>0</sub>: Οι συντελεστές του μοντέλου μας ΔΕΝ απέχουν και πολύ από το μηδέν.

<sup>25</sup> Ναι, έτσι αποκαλεί το SPSS 28, το Multinomial Logistic Regression.

<sup>26</sup> Με τη ΧΡΥΣΗ ΑΥΓΗ ποτέ δεν ξέρεις. Όπως έχουμε ξαναβρεί σε αρκετές έρευνες του Εργαστηρίου Κοινωνικής Στατιστικής και Πολιτικής Έρευνας του Π.Κ., αρκετοί είναι αυτοί που ψηφίζουν ΧΑ, αλλά δεν το λένε...

**Πίνακας 5.29.** Περιθώριων Συχνοτήτων, Model 7.

Case Processing Summary			
	N	Marginal Percentage	
PARTY	ND	306	27,6%
	SYRIZA	407	36,7%
	PASOK	49	4,4%
	POTAMI	67	6,0%
	KKE	59	5,3%
	XRYSH AYGH	57	5,1%
	ANEL	50	4,5%
ideological	ALLO	115	10,4%
	YES	607	54,7%
economic	NO	503	45,3%
	YES	713	64,2%
salvage	NO	397	35,8%
	YES	823	74,1%
obligation	NO	287	25,9%
	YES	140	12,6%
antimemorandum	NO	970	87,4%
	YES	515	46,4%
versus-system	NO	595	53,6%
	YES	486	43,8%
important	NO	624	56,2%
	CHIEF	123	11,1%
	CANDIDATES	42	3,8%
	THEORY	252	22,7%
	PROGRAM	609	54,9%
SEX	OTHER	84	7,6%
	MALE	566	51,0%
	FEMALE	544	49,0%
EDUC	ELEMENTARY	140	12,6%
	MIDDLE	407	36,7%
	HIGHER	563	50,7%
CAT_AGE	18-34	228	20,5%
	35-54	476	42,9%
	55 plus	406	36,6%
Valid		1110	100,0%
Missing		807	
Total		1917	
Subpopulation		698 <sup>a</sup>	
a. The dependent variable has only one value observed in 555 (79,5%) subpopulations.			

**Πίνακας 5.30.** Σύγκριση Model 7 με Null μοντέλο.

Model Fitting Information						
Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	3338,350	3373,435	3324,350			
Final	2881,641	3442,998	2657,641	666,709	105	<,001

**Πίνακας 5.31.** Προσαρμογή (goodness-of-fit) του Model 7.

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	5544,030	4774	<,001
Deviance	2334,228	4774	1,000

Επειδή όμως το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι μικρότερο από 1 %, αδηγούμαστε στο συμπέρασμα ότι υπάρχουν συντελεστές του μοντέλου μας, που δεν είναι μηδέν. Άρα αυτό έχει λόγο ύπαρξης.

Ο Πίνακας 5.31 ελέγχει την κυρίαρχη παραδοχή της Multinomial Logistic Regression διαδικασίας. Την ύπαρξη ή μη καλής προσαρμογής (Goodness-of-fit). Μας ενδιαφέρει πάντα, όταν κτίζουμε ένα μοντέλο, πόσο πολύ αυτό προσαρμόζεται στα δεδομένα. Ο έλεγχος γίνεται με τη βοήθεια των στατιστικών κριτηρίων Pearson  $\chi^2$  και Deviance.

Μεγάλες τιμές της στατιστικής Pearson  $\chi^2$ , φανερώνουν ότι το μοντέλο δεν έχει καλή προσαρμογή, δεν έχει καλό fit. Ισοδύναμα εκφραζόμενοι, θα λέγαμε ότι, χαμηλές τιμές στο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας σημαίνουν όχι καλή προσαρμογή. Φυσικά, όσο μεγαλύτερο είναι το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, τόσο καλύτερη είναι η προσαρμογή του μοντέλου μας στα δεδομένα. Πεμπτουσία αυτής της στατιστικής είναι η εκτίμηση της αντίφασης/διαφοράς, ανάμεσα σε παρατηρούμενες και αναμενόμενες συχνότητες, μέσα σε ένα πολυδιάστατο πίνακα διασταύρωσης (Crosstabulation). Συνήθης τρόπος υπολογισμού η σχέση:

$$\chi^2_{\text{Pearson}} = \sum_{\text{AllCells}} \frac{(\text{observed\_counts} - \text{expected\_counts})^2}{\text{expected\_counts}} \quad (5.22)$$

Στον Πίνακα 5.31 προφανώς και δεν είναι υψηλό το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας. Ωστόσο, λόγος για όχι καλή προσαρμογή του μοντέλου στα δεδομένα, δεν μπορεί να γίνει, διότι ο λόγος  $\chi^2/\text{df}$  είναι μικρότερος από το 2.

Συμβολικά έχουμε:

$$\frac{\chi^2}{\text{DF}} = \frac{5544,030}{4774} \simeq 1,16 < 2 \quad (5.23)$$

Ενας άλλος τρόπος να μετρήσουμε την Goodness-of-fit είναι η απόκλιση  $\chi^2$ , αυτή που ονομάζεται με το διεθνή όρο Deviance. Πεμπτουσία αυτής της στατιστικής είναι η μεταβολή της ποσότητας  $-2\text{LL} = -2\ln\text{-likelihood}$  ανάμεσα στο μοντέλο μας και στο κορεσμένο (saturated) μοντέλο, δηλ. εκείνο που διαθέτει όλες τις κύριες επιδράσεις (main effects) και αλληλεπιδράσεις (interactions). Έτσι, αν το μοντέλο μας ταιριάζει

καλά στα δεδομένα, τότε η διαφορά ανάμεσα στους λογαρίθμους των πιθανοτήτων θα πρέπει να είναι μικρή.

Συμβολικά για την Deviance και για πολυωνυμικό (multinomial) μοντέλο έχουμε:

$$\text{Deviance} = G^2 = 2 \sum \text{Observed} \cdot \ln\left(\frac{\text{Observed}}{\text{Expected}}\right) \quad (5.24)$$

Στη σχέση (5.24) το άθροισμα αφορά όλα τα κελιά του πίνακα διασταύρωσης τα οποία δεν είναι δομικά μηδέν και δεν έχουν μηδενικές fitted τιμές.

Στον Πίνακα 5.31 και για τη στατιστική Deviance έχουμε υψηλότατο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, που αγγίζει το 100%, γεγονός που μας καθησυχάζει ότι η προσαρμογή του μοντέλου μας στα δεδομένα είναι καλή.

### Σημαντικές παρατηρήσεις

- I. Τόσον η  $X^2$  όσον και η  $G^2$  στατιστική, ασυμπτωτικά ακολουθούν  $X^2$ -κατανομή.
- II. Ένα κορεσμένο (saturated) μοντέλο πάντα παράγει τέλεια προσαρμογή, τέλειο Fit.
- III. Στις αναλύσεις κοινωνικών δεδομένων άλλες φορές χρησιμοποιείται η στατιστική  $X^2$ , και άλλες η στατιστική  $G^2$ .
- IV. Εάν διαθέτουμε μεγάλα δειγματικά μεγέθη οι δύο στατιστικές θα πρέπει να οδηγούν στα ίδια συμπεράσματα, αναφορικά με την καλή προσαρμογή.

#### 5.9.1. Το φαινόμενο της υπερδιασκόρπισης (overdispersion).

Ηρθε η ώρα να μιλήσουμε για την έννοια της υπερδιασκόρπισης (overdispersion). Είναι μια ασθένεια που παρουσιάζεται στα Logistic μοντέλα, κι όχι μόνον, και οφείλεται στο γεγονός ότι η παρατηρούμενη/εμπειρική διασπορά (empirical variation) είναι μεγαλύτερη από την προβλεπόμενη/ ονομαστική (nominal variation), του μοντέλου. Αυτό πιθανότατα οφείλεται σε 4 λόγους:

- I. Ο πρώτος είναι η συσχέτιση που παρουσιάζουν οι παρατηρήσεις μεταξύ τους. Ας θυμηθούμε, από την Περιγραφική κιόλας Στατιστική, την παραδοχή της

ανεξαρτησίας των παρατηρήσεων (independence). Όταν αυτή παραβιάζεται προφανώς και έχουμε correlation. Οι Cases ‘κολλάνε’ μεταξύ τους.

- II. Ο δεύτερος λόγος είναι η μεγάλη μεταβλητότητα (variability) στις πιθανότητες επιτυχίας (Halekoh, U., & Hojsgaard, S., 2007).
- III. Ο τρίτος είναι η παρουσία μιας ή περισσοτέρων ακραίων παρατηρήσεων, (multivariate outliers).
- IV. Ο τέταρτος το κακοκτισμένο (miss- building) μοντέλο.

Μιλήσαμε ωστόσο για ασθένεια στα νεφρά του Logistic Model. Εννοούμε με αυτό ότι το overdispersion παρεισφρέει στις δομές του μοντέλου προσπαθώντας να περιορίσει τα Standard errors μέσα σε πιο στενά διαστήματα εμπιστοσύνης. Από την άλλη δεν λησμονούμε ότι, για να βγάλουμε τη σπουδαιότητα (Importance) ενός συντελεστή, διαιρούμε την τιμή του, δια του Τυπικού σφάλματος (Standard error). Και βέβαια, όσο μικρότερο είναι το Τυπικό σφάλμα, τόσο η ανεξάρτητη μεταβλητή (predictor) βγαίνει σπουδαιότερη, χωρίς αυτή να είναι στην πραγματικότητα. Τόσο μικρότερο δηλ. βγαίνει το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, που την κάνει με τη σειρά του να φαίνεται σπουδαία. Σε τελευταία ανάλυση το μικρόβιο overdispersion κονταίνει τα τυπικά σφάλματα υποχρεώνοντάς τα να χωρέσουν, σε μικρότερα διαστήματα εμπιστοσύνης. Αυτό το τελευταίο είναι σκέτη καταστροφή. Διότι, στενότερα διαστήματα εμπιστοσύνης θα τροφοδοτήσουν την υπέρ-εμπιστοσύνη (overconfidence). Τα μειωμένα διαστήματα, με τη σειρά τους, θα μας μιλήσουν για μεγάλη επιδραστικότητα των predictors πάνω στη μεταβλητή έκβασης (outcome variable). Και τότε, προφανώς, μπαίνουμε σε μεγαλύτερο κίνδυνο να διαπράξουμε σφάλμα τύπου I.

Ανάγκη λοιπόν, πάσα, να μετρήσουμε την υπερ-διασκόρπιση. Το SPSS εμμέσως πλην σαφώς μετράει, με το συντελεστή  $\Phi$  την υπερ-διασκόρπιση. Είναι ακριβώς ο λόγος της  $\chi^2$  - goodness στατιστικής δια των βαθμών ελευθερίας αυτού. Οπότε, θα έχουμε:

$$\text{overdispersion\_parameter\_}\Phi = \frac{\chi^2\_goodness\_of\_it}{df} \quad (5.25)$$

Υπάρχει overdispersion αν το Φ ξεπερνά το 2. Αντίθετα, αν ο λόγος αυτός είναι μεγαλύτερος του 1 και μικρότερος του 2, δεν ανησυχούμε ιδιαίτερα για υπερδιασκόρπιση.

Εστιάζουμε ξανά την προσοχή μας στον Πίνακα 5.31. Τελικά έχουμε υπερδιασκόρπιση  
Ναι ή Όχι; Και σε ποιο βαθμό;

Η απάντηση είναι απλή: Ο λόγος

$$\frac{X^2}{DF} = \frac{5544,030}{4774} = 1,1612966 \simeq 1,16$$

Είναι ελαφρά μεγαλύτερος από το 1 και ταυτόχρονα μικρότερος από το 2.

Το γεγονός τούτο δεν αποτελεί επαρκεί ερευνητική μαρτυρία για να τεκμηριώσουμε overdispersion. Επομένως δεν μπορεί να γίνει λόγος για προβληματικό δείκτη υπερδιασκόρπισης (Field, 2018).

Ολοκληρώνοντας τη συζήτηση αναφορικά με την υπερδιασκόρπιση θα λέγαμε ότι αν θέλουμε να έχουμε σωστές εκτιμήσεις τυπικών σφαλμάτων, και σπουδαιότητας θέλουμε σωστούς κι όχι υπερτιμημένους ή υποτιμημένους παλινδρομικούς συντελεστές, και τέλος αν θέλουμε καλή προσαρμογή του μοντέλου μας στα δεδομένα και σωστή πρόβλεψη, θα πρέπει, η παραδοχή της υπερδιασκόρπισης να ελεγχθεί.

**5.9.2. Τι είναι τα covariate patterns, ποια είναι η σχέση τους με τους βαθμούς ελευθερίας, τα ανεξάρτητα Logits και το συνολικό αριθμό των παραμέτρων που θέλουμε να εκτιμήσουμε, συμπεριλαμβανομένου και του σταθερού όρου.**

### Ορισμός 1

Θα λέμε covariate pattern κάθε διακριτό συνδυασμό τιμών των ανεξάρτητων μεταβλητών (predictors).

### Ορισμός 2

Οι βαθμοί ελευθερίας είναι ίσοι με το γινόμενο των Covariate patterns, επί τον αριθμό των ανεξάρτητων Logits, μείον τον αριθμό των παραμέτρων που έχουμε να εκτιμήσουμε συμπεριλαμβανομένου του σταθερού όρου (intercept). Συμβολικά:

$$DF = (\text{NUMBER\_OF\_OBSERVED\_COVARIATE\_PATTERNS}) - (\text{NUMBER\_OF\_INDEPENDENT\_LOGITS}) - \text{NUMBER\_ESTIMATED\_PARAMETERS} \quad (5.26)$$

### 5.9.3. Τα Likelihood Ratio Tests

Τα Likelihood Ratio Tests συμβάλλουν αποφασιστικά στην αποσαφήνιση/ξεκαθάρισμα των effects, και έχουν μεγαλύτερη απόδοση από τα Wald Tests, αναφορικά με αυτό το ζήτημα. Με απλούστερα λόγια εκφραζόμενοι για να διευκρινίσουν τη συμβολή μιας μεταβλητής συγκρίνουν δυο μοντέλα: Αυτό που την έχει, (μαζί ίσως με άλλες), με αυτό που δεν την έχει. Αν η μεταβλητή είναι στατιστικώς σημαντική, και αυτό θα το δούμε στη δεξιά στήλη του Πίνακα 5.32, το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας θα είναι κάτω του 5%. Έτσι βλέπουμε στον Πίνακα 5.32, όλες οι μεταβλητές να είναι σημαντικές, με μία εξαίρεση ίσως, εκείνη του φύλου (SEX), στην οποία βλέπουμε να αντιστοιχεί ένα οριακό παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, της τάξεως του 5,4%.

**Πίνακας 5.32.** Σύγκριση των Reduced Model 7 με Null μοντέλο.

Effect	Likelihood Ratio Tests			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	2881,641	3442,998	2657,641 <sup>a</sup>	,000	0	.
SEX	2881,490	3407,762	2671,490	13,849	7	,054
EDUC	2894,216	3385,403	2698,216	40,575	14	<,001
CAT_AGE	2884,736	3375,924	2688,736	31,095	14	,005
ideological	2893,924	3420,196	2683,924	26,284	7	<,001
economic	2890,381	3416,653	2680,381	22,740	7	,002
salvage	2906,381	3432,653	2696,381	38,740	7	<,001
obligation	2882,532	3408,804	2672,532	14,892	7	,037
antimemorandum	3069,752	3596,024	2859,752	202,111	7	<,001
versus-system	2920,853	3447,126	2710,853	53,213	7	<,001
important	2882,944	3303,961	2714,944	57,303	28	<,001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Ωστόσο, ο Πίνακας 5.32 κομίζει και μια άλλη, ίσως χρήσιμη πληροφορία, αναφορικά με τη Γραμμικότητα (Linearity). Εμμέσως πλην σαφώς, μας δείχνει τη σχέση που έχει η κάθε μια μεταβλητή, ξεχωριστά, με τα Logits της μεταβλητής έκβασης (outcome variable). Εδώ, και με βάση τη λογική των Tabachnick, Fidell (2019), μπορούμε να πούμε ότι η παραδοχή της Γραμμικότητας (Linearity), για το Model 7, δεν ικανοποιείται καθώς στη δεξιά στήλη του Πίνακα 5.32, συχνά πυκνά διαφαίνεται σοβαρή παραβίαση αυτής. (Το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, με εξαίρεση τη μεταβλητή του Φύλου (SEX), είναι για όλες τις μεταβλητές, μικρότερο του 5%). Να διευκρινίσουμε ότι αυτό το τελευταίο, η Γραμμικότητα ως παραδοχή, πάλι κατά τη λογική των Tabachnick, και Fidell (2019), ικανοποιείται όταν για τη μεγάλη πλειοψηφία των ανεξάρτητων μεταβλητών (predictors), το παρατηρούμενο επίπεδο στατ. Σημαντικότητας, είναι μεγαλύτερο του 5%.

Δεν θα μπορούσαμε, να αγνοήσουμε την άποψη των Long και Freese (2014), για τη Γραμμικότητα. Θεωρούν ότι πρόκειται για ένα ζήτημα ήσσονος σημασίας, καθώς η πρακτική του αξία (practical importance) είναι μικρή.

#### **5.9.4. Η Προβλεπτική Δύναμη (Prediction Power) του ΜΟΝΤΕΛΟΥ 7. Ο Πίνακας Ταξινόμησης (Classification Table).**

Ο Πίνακας 5.33 είναι ένας τετραγωνικός Πίνακας Ταξινόμησης (Classification Table), διαστάσεων 8x8, ο οποίος μας πληροφορεί για την ικανότητα του μοντέλου να εκχωρεί υψηλές προβλεπόμενες πιθανότητες (predicted probabilities) σε παρατηρούμενες συχνότητες.

**Πίνακας 5.33. Πίνακας Ταξινόμησης (Classification Table) για το Model 7.**

Observed	ND	Classification								Percent Correct
		SYRIZA	PASOK	POTAMI	KKE	XRYSH AYGH	ANEL	ALLO	Predicted	
ND	244	57	0	1	0	2	0	2	79,7%	
SYRIZA	72	322	0	4	1	5	0	3	79,1%	
PASOK	34	15	0	0	0	0	0	0	0,0%	
POTAMI	31	31	0	3	0	0	0	2	4,5%	
KKE	12	47	0	0	0	0	0	0	0,0%	
XRYSH AYGH	10	35	0	1	0	11	0	0	19,3%	
ANEL	7	40	0	0	0	3	0	0	0,0%	
ALLO	49	59	0	1	2	2	0	2	1,7%	
Overall Percentage	41,4%	54,6%	0,0%	0,9%	0,3%	2,1%	0,0%	0,8%	52,4%	

Για παράδειγμα, για τη Νέα Δημοκρατία ο αριθμός 244 σημαίνει ότι, 244 ψηφοφόροι ψήφισαν τη Νέα Δημοκρατία, και τους είχε προβλέψει/εντοπίσει/ανιχνεύσει επίσης, το Model 7.

Αντίστοιχα, για το ΣΥΡΙΖΑ ο αριθμός 322 σημαίνει ότι, 322 ψηφοφόροι ψήφισαν το ΣΥΡΙΖΑ, και τους είχε προβλέψει/εντοπίσει/ανιχνεύσει επίσης, το Model 7.

Ας σημειώσουμε τόσο ο αριθμός 244, όσο και ο αριθμός 322 βρίσκονται στην πρωτεύοντα διαγώνιο του τετραγωνικού πίνακα, κι επομένως το αντίστοιχο νούμερο για το ΠΑΣΟΚ είναι το μηδέν. Τι σημαίνει όπως μηδέν. Ακριβώς ότι κανένας ψηφοφόρος δεν εντοπίστηκε/ ανιχνεύτηκε από το Model 7 να κατευθύνεται στην κάλπη του ΠΑΣΟΚ.

Στην τελευταία δεξιά στήλη του Πίνακα 5.33 βλέπουμε, για κάθε πολιτικό Κόμμα, την ανιχνευτική δύναμη του Model 7. Έτσι έχουμε:

- I. Για τη Νέα Δημοκρατία η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 79,7%.
- II. Για τον ΣΥΡΙΖΑ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 79,1%.
- III. Για το ΠΑΣΟΚ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 0%.
- IV. Για το ΠΟΤΑΜΙ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 4,5%.
- V. Για το ΚΚΕ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 0%.
- VI. Για τη ΧΡΥΣΗ ΑΥΓΗ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 19,3%.
- VII. Για το ΠΑΣΟΚ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 0%.
- VIII. Για τους ΑΝΕΞΑΡΤΗΤΟΥΣ ΕΛΛΗΝΕΣ η ικανότητα ανίχνευσης του Model 7, είναι της τάξεως του 0%.

## **Σημαντικές παρατηρήσεις**

- I. Η πρώτη παρατήρηση είναι ότι για τα δυο μεγαλύτερα Κόμματα, εκείνο της Νέας Δημοκρατίας και του ΣΥΡΙΖΑ, η προβλεπτική δύναμη του Model 7 κυμαίνεται σε αρκετά υψηλά επίπεδα (79,7% και 79,1% αντίστοιχα), ενώ Αυτό είναι απολύτως φυσιολογικό να συμβαίνει διότι, απλά το μοντέλο πρόβλεψης εκχωρεί στα μεγαλύτερα πληθυσμιακώς groups, υψηλότερες probabilities (Norusis, 2010). Με άλλα λόγια, όταν έχουμε ανισομεγέθη groups είναι πιο πιθανό οι cases να ταξινομηθούν σε μεγαλύτερο βαθμό στα μεγαλύτερα groups.
- II. Οι Hosmer, Lemenshow (2000), που δούλεψαν αρκετά πάνω σε αυτό το θέμα, πάνω στον Classification Table, βρήκαν ότι ένας τέτοιος πίνακας μπορεί να είναι informative, μπορεί δηλ. να εμπεριέχει πολλές και χρήσιμες πληροφορίες, αλλά πολύ απέχει από το να μπορεί να δώσει χρήσιμες πληροφορίες για την προσαρμογή του μοντέλου στα δεδομένα, κι ακόμη, να είναι ικανός να ταξινομήσει σωστά.
- III. Αναφορικά με τη γενική ικανότητα του Model 7 να προβλέπει, και με βάση όσα είπαμε παραπάνω και αποτελούν δεδομένο για την οριακή του ικανότητα να προβλέψει σωστά, στην τελευταία του στήλη (percent correct) παρατηρούμε ένα γενικό ποσοστό της τάξεως του 52,4%.
- IV. Επίσης, με βάση όσα είπαμε παραπάνω για το θέμα της εκχώρησης υψηλών probabilities σε παρατηρούμενες cases στα μεγάλα groups και μικρότερων στα χαμηλού μεγέθους, είναι εύλογο για τα μικρότερα Κόμματα, όπως για παράδειγμα το ΠΟΤΑΜΙ, ο Πίνακας 5.33 να εμφανίζει μικρότερη ικανότητα ανίχνευσης (4,5%).

### **5.9.5. Η Ερμηνεία των παραμέτρων για το Model 7.**

Ο Πίνακας 5.34 είναι ένας πίνακας που αναφέρεται στην αναμέτρηση της ΝΔ με το ΣΥΡΙΖΑ, στο ιδεολογικό, οικονομικό, κ.τ.λ πεδίο της έρευνας. Πιο συγκεκριμένα παρατηρούμε τα εξής:

- I. Οι ιδεολογικοί λόγοι είναι μια στατιστικώς σημαντική μεταβλητή του Model 7 (Sig. = 0,046 = 4,6% < 5%). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +0,396$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για ιδεολογικούς λόγους, συγκρινόμενοι με εκείνους που δεν ψήφισαν για ιδεολογικούς λόγους, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε που είναι η βάση των νεπέριων λογαρίθμων, και ο οποίος έχει τιμή στην περίπτωση αυτή ίση με 1,486 μονάδες και είναι προφανώς μεγαλύτερος της μονάδας,  $[\exp(0,396) = 1,486 > 1]$ . Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες /μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει ιδεολογικά, η πιθανότητα να ψηφίσει ΝΔ είναι 1,486 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.
- II. Οι οικονομικοί λόγοι δεν είναι μια στατιστικώς σημαντική μεταβλητή του Model 7 (Sig. = 0,132 = 13,2% > 5%). Δεν παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ. Αυτό παραπέρα σημαίνει ότι όσοι ψήφισαν για οικονομικούς λόγους, συγκρινόμενοι με εκείνους που δεν ψήφισαν για οικονομικούς λόγους, εξ ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ. Αυτό μας το λέει κι ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή 0,736 μονάδες, τιμή προφανώς μικρότερη του 1  $[\exp(-0,306) = 0,736 < 1]$ . Επιτόπιο συμπέρασμα: Ο ΣΥΡΙΖΑ δεν ανέβηκε στην εξουσία ένεκα οικονομικής δυσπραγίας. Σε άλλους λόγους, όπως θα δούμε παρακάτω, οφείλεται η άνοδος του ΣΥΡΙΖΑ στην εξουσία.
- III. Η σωτηρία της πατρίδας είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7 (Sig. < 0,001 = 0,01%). Παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +1,057$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για τη σωτηρία της πατρίδας, συγκρινόμενοι με εκείνους που δεν ψήφισαν για τη σωτηρία της πατρίδας, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή ίση με 2,876 μονάδες και είναι προφανώς είναι μεγαλύτερος της μονάδας,

$[\exp(1,057) = 2,876 > 1]$ . Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες/μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει για τη σωτηρία της πατρίδας, η πιθανότητα να ψηφίσει ΝΔ είναι 2,876 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.

IV. Οι ανειλημμένες υποχρεώσεις είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7 ( $\text{Sig.} = 0,007 = 0,07\%$ ). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +0,758$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για λόγους ανειλημμένων υποχρεώσεων, συγκρινόμενοι με εκείνους που δεν ψήφισαν για λόγους ανειλημμένων υποχρεώσεων, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή ίση με 2,134 μονάδες, και ο οποίος είναι προφανώς είναι μεγαλύτερος της μονάδας,  $[\exp(0,758) = 2,134 > 1]$ . Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες/μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει για τη σωτηρία της πατρίδας, η πιθανότητα να ψηφίσει ΝΔ είναι 2,134 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.

V. Τα πράγματα αλλάζουν στην περίπτωση της μεταβλητής του αντιμνημονιακού θέματος. Η στάση των πολιτών απέναντι στο μνημόνιο είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7 ( $\text{Sig.} < 0,001 = 0,1\%$ ). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει αρνητικό πρόσημο ( $B = -2,613$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν αντιμνημονιακά, συγκρινόμενοι με εκείνους που δεν ψήφισαν αντιμνημονιακά, παρουσιάζουν μειωμένες πιθανότητες να κατευθύνονται προς την κάλπη της ΝΔ. Κατά πόσο; Αυτό θα μας το πει η αφαίρεση  $1 - \exp(-2,613) = 1 - 0,073 = 0,927 = 92,7\%$ . Για κάποιον επομένως, που ψηφίζει αντιμνημονιακά, η πιθανότητα να ψηφίσει ΝΔ είναι μειωμένη κατά 92,7% μικρότερη παρά να ψηφίσει ΣΥΡΙΖΑ.

Επιτόπιο συμπέρασμα: Οι ψηφοφόροι του ΣΥΡΙΖΑ, στην συντριπτική τους πλειοψηφία προέρχονται από το αντιμνημονιακό στρατόπεδο. Ας το πούμε αλλιώς: Το αντιμνημονιακό στρατόπεδο φαίνεται να υπήρξε η κύρια δεξαμενή από την οποία άντλησε ψήφους ο ΣΥΡΙΖΑ.

VI. Στοιχεία όμοια με το αντιμνημονιακό ζήτημα, βλέπουμε και στο αντισυστημικό ζήτημα. Η στάση των πολιτών απέναντι στο σύστημα είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7 (Sig. < 0,001 = 0,1%). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει αρνητικό πρόσημο ( $B = -1,364$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν αντισυστημικά, συγκρινόμενοι με εκείνους που δεν ψήφισαν αντισυστημικά, παρουσιάζουν μειωμένες πιθανότητες να κατευθύνονται προς την κάλπη της ΝΔ. Κατά πόσο; Αυτό θα μας το πει η αφαίρεση  $1-\exp(-1,364) = 1-0,256 = 0,744 = 74,4\%$ . Για κάποιον επομένως, που ψηφίζει αντισυστημικά, η πιθανότητα να ψηφίσει ΝΔ είναι μειωμένη κατά 74,4% μικρότερη παρά να ψηφίσει ΣΥΡΙΖΑ.

Επιτόπιο συμπέρασμα: Οι ψηφοφόροι του ΣΥΡΙΖΑ, στην μεγάλη τους πλειοψηφία προέρχονται από το αντισυστημικό στρατόπεδο. Ας το πούμε αλλιώς: Το αντισυστημικό στρατόπεδο φαίνεται να υπήρξε επίσης μια μεγάλη δεξαμενή από την οποία άντλησε ψήφους ο ΣΥΡΙΖΑ.

VII. Για το θέμα της σπουδαιότητας<sup>27</sup> ο Πίνακας 5.34 μας παρέχει 4 συμπεράσματα:

Πρώτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη του Αρχηγού, συγκρινόμενοι με εκείνους που το ψήφισαν για «Άλλο» λόγο, εμφανίζουν αυξημένες πιθανότητες, για την ακρίβεια 1,477 φορές να κατευθύνονται προς την κάλπη της ΝΔ παρά προς την κάλπη του ΣΥΡΙΖΑ [Sig. = 0,005 = 0,5% < 5% και επίσης  $\exp(1,27) = 1,477$ ].

<sup>27</sup> Υπενθυμίζεται εδώ η σχετική Ερώτηση Q2 του Ερωτηματολογίου “Έρευνας: Ποιος ήταν ο σημαντικότερος λόγος στην επιλογή της ψήφου σας; 1= ο αρχηγός του Κόμματος; 2=οι υποψήφιοι; 3=η γενική κοσμοθεωρία/ιδεολογία του Κόμματος; 4=το Πρόγραμμα και οι Θέσεις; 5=Άλλος λόγος ποιος ....;

Δεύτερο: Εκείνοι που ψήφισαν το Κόμμα για χάρη των Υποψηφίων, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig. = 0,077 = 7,7% > 5% και επίσης  $\exp(1,056) = 2,876$ ].

Τρίτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη της Κοσμοθεωρίας του Κόμματος, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig. = 0,205 = 20,5% > 5% και επίσης  $\exp(0,528) = 1,695$ ].

Τέταρτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη του Προγράμματος και των Θέσεων του Κόμματος, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig.= 0,676 = 67,6 % > 5 % και επίσης  $\exp(0,154) = 1,167$ ].

VIII. Το Φύλο δεν είναι μια στατιστικώς σημαντική μεταβλητή για το Model 7 (Sig. = 0,292 = 29,2% > 5%). Δεν παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ. Αυτό παραπέρα σημαίνει ότι οι άνδρες, συγκρινόμενοι με τις γυναίκες, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ. Επιτόπιο συμπέρασμα: Δεν υπάρχει επαρκής ερευνητική μαρτυρία (evidence) ότι τα δυο φύλλα είχαν διαφορετική στάση απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ.

IX. Για το Μορφωτικό επίπεδο ο Πίνακας 5.34 μας παρέχει 2 συμπεράσματα:

Πρώτο: Οι απόφοιτοι της Στοιχειώδους Εκπαίδευσεως συγκρινόμενοι με τους αποφοίτους του Πανεπιστημίου δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ (Sig. = 0,098 = 9,8% > 5%).

Δεύτερο: Οι απόφοιτοι της Μέσης Εκπαίδευσεως συγκρινόμενοι με τους αποφοίτους του Πανεπιστημίου δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ (Sig.= 0,262 = 26,2% > 5%).

X. Για τις Ηλικιακές Κατηγορίες ο Πίνακας 5.34 μας παρέχει 2 συμπεράσματα:

Πρώτο: Η ηλικιακή κατηγορία των 17-34 συγκρινόμενη με εκείνη των 55 και άνω δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ (Sig.= 0,149 = 14,9% > 5%).

Δεύτερο: Η ηλικιακή κατηγορία των 35-54 συγκρινόμενη με εκείνη των 55 και άνω εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ (Sig.= 0,042 = 4,2% < 5%). Ποιο συγκεκριμένα, η ηλικιακή κατηγορία των 35-54 εμφανίζει μειωμένες πιθανότητες κατά 35,4% (1-0,646 = 0,354 = 35,4%), να κατευθύνεται προς την κάλπη της ΝΔ, συγκριτικά με την κάλπη του ΣΥΡΙΖΑ.

Σημαντική Σημείωση: Αν θέσουμε ως κατηγορία αναφοράς την ηλικιακή κατηγορία των 35-54, θα εντοπίσουμε ενδεχόμενα και άλλες διαφορές στην εκλογική συμπεριφορά των υπόλοιπων ηλικιακών κατηγοριών.

**Πίνακας 5.34.** Εκτιμήσεις των παραμέτρων του Model 7, για τη Ν.Δ.

PARTY*	Parameter Estimates						95% Confidence Interval for Exp(B)	
	B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound
ND	Intercept	,059	,444	,017	1	,895		
	[ideological=1]	,396	,199	3,974	1	,046	1,486	1,007
	[ideological=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[economic=1]	-,306	,203	2,271	1	,132	,736	,495
	[economic=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[salvage=1]	1,057	,234	20,380	1	<,001	2,876	1,818
	[salvage=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[obligation=1]	,758	,279	7,355	1	,007	2,134	1,234
	[obligation=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[antimemorandum=1]	-,2613	,228	131,625	1	<,001	,073	,047
	[antimemorandum=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[versus-system=1]	-,1364	,217	39,318	1	<,001	,256	,167
	[versus-system=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[important=1]	1,270	,449	8,000	1	,005	3,562	1,477
	[important=2]	1,056	,597	3,130	1	,077	2,876	,892
	[important=3]	,528	,416	1,610	1	,205	1,695	,750
	[important=4]	,154	,369	,175	1	,676	1,167	,566
	[important=5]	0 <sup>b</sup>	-	-	0	-	-	-
	[SEX=1]	-,197	,187	1,108	1	,292	,821	,569
	[SEX=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[EDUC=1]	,509	,308	2,733	1	,098	1,664	,910
	[EDUC=2]	,234	,209	1,256	1	,262	1,264	,839
	[EDUC=3]	0 <sup>b</sup>	-	-	0	-	-	-
	[CAT_AOE=1]	-,401	,278	2,085	1	,149	,670	,388
	[CAT_AOE=2]	-,437	,214	4,148	1	,042	,646	,424
	[CAT_AOE=3]	0 <sup>b</sup>	-	-	0	-	-	-

Η Ερμηνεία των Πινάκων 5.35-5.38, γίνεται βάση όσων αναφέρθηκαν στον Πίνακα 5.34.

**Πίνακας 5.35.** Εκτιμήσεις των παραμέτρων του Model 7, για το ΠΑ.ΣΟ.Κ.

PASOK	Intercept	-1,293	,726	3,173	1	,075		
	[ideological=1]	,719	,375	3,685	1	,055	2,052	,985 4,276
	[ideological=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[economic=1]	-,905	,339	7,116	1	,008	,405	,208 ,787
	[economic=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[salvage=1]	,636	,399	2,538	1	,111	1,889	,864 4,133
	[salvage=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[obligation=1]	,835	,482	2,999	1	,083	2,304	,896 5,928
	[obligation=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[antimemorandum=1]	-1,929	,392	24,239	1	<.001	,145	,067 ,313
	[antimemorandum=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[versus-system=1]	-,711	,376	3,591	1	,058	,491	,235 1,026
	[versus-system=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[important=1]	,932	,713	1,706	1	,192	2,539	,627 10,278
	[important=2]	,092	1,234	,006	1	,941	1,096	,098 12,325
	[important=3]	,939	,648	2,099	1	,147	2,557	,718 9,103
	[important=4]	-,323	,621	,271	1	,603	,724	,214 2,445
	[important=5]	0 <sup>b</sup>	.	.	0	.	.	.
	[SEX=1]	,376	,336	1,254	1	,263	1,457	,754 2,816
	[SEX=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[EDUC=1]	,119	,503	,056	1	,813	1,126	,420 3,021
	[EDUC=2]	,183	,362	,255	1	,614	1,200	,591 2,440
	[EDUC=3]	0 <sup>b</sup>	.	.	0	.	.	.
	[CAT_AGE=1]	-1,568	,589	7,085	1	,008	,209	,066 ,661
	[CAT_AGE=2]	-1,063	,372	8,157	1	,004	,346	,167 ,716
	[CAT_AGE=3]	0 <sup>b</sup>	.	.	0	.	.	.

**Πίνακας 5.36.** Εκτιμήσεις των παραμέτρων του Model 7, για το Κ.Κ.Ε.

KKE	Intercept	-2,376	,776	9,369	1	,002		
	[ideological=1]	,885	,341	6,729	1	,009	2,422	,1241 4,726
	[ideological=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[economic=1]	-,419	,313	1,794	1	,180	,658	,356 1,214
	[economic=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[salvage=1]	-,469	,307	2,335	1	,127	,626	,343 1,142
	[salvage=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[obligation=1]	,477	,460	1,075	1	,300	1,612	,654 3,974
	[obligation=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[antimemorandum=1]	-,106	,322	,109	1	,741	,899	,479 1,689
	[antimemorandum=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[versus-system=1]	-,072	,309	,054	1	,815	,930	,507 1,706
	[versus-system=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[important=1]	-,229	,963	,056	1	,812	,796	,120 5,254
	[important=2]	,007	1,234	,000	1	,995	1,007	,090 11,311
	[important=3]	1,185	,671	3,120	1	,077	3,270	,878 12,175
	[important=4]	,224	,649	,119	1	,730	1,251	,351 4,466
	[important=5]	0 <sup>b</sup>	.	.	0	.	.	.
	[SEX=1]	-,019	,290	,004	1	,949	,982	,556 1,731
	[SEX=2]	0 <sup>b</sup>	.	.	0	.	.	.
	[EDUC=1]	,144	,530	,074	1	,785	1,155	,409 3,267
	[EDUC=2]	-,109	,339	,104	1	,748	,897	,462 1,741
	[EDUC=3]	0 <sup>b</sup>	.	.	0	.	.	.
	[CAT_AGE=1]	,516	,411	1,575	1	,209	1,676	,748 3,754
	[CAT_AGE=2]	-,076	,364	,044	1	,834	,927	,454 1,890
	[CAT_AGE=3]	0 <sup>b</sup>	.	.	0	.	.	.

**Πίνακας 5.37.** Εκτιμήσεις των παραμέτρων του Model 7, για το Χ.Α.

KKE	Intercept	-2,376	,776	9,369	1	,002		
	[ideological=1]	,885	,341	6,729	1	,009	2,422	1,241
	[ideological=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[economic=1]	-,419	,313	1,794	1	,180	,658	,356
	[economic=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[salvage=1]	-,469	,307	2,335	1	,127	,626	,343
	[salvage=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[obligation=1]	,477	,460	1,075	1	,300	1,612	,654
	[obligation=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[antimemorandum=1]	-,106	,322	,109	1	,741	,899	,479
	[antimemorandum=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[versus-system=1]	-,072	,309	,054	1	,815	,930	,507
	[versus-system=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[important=1]	-,229	,963	,056	1	,812	,796	,120
	[important=2]	,007	1,234	,000	1	,995	1,007	,090
	[important=3]	1,185	,671	3,120	1	,077	3,270	,878
	[important=4]	,224	,649	,119	1	,730	1,251	,351
	[important=5]	0 <sup>b</sup>	-	-	0	-	-	-
	[SEX=1]	-,019	,290	,004	1	,949	,982	,556
	[SEX=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[EDUC=1]	,144	,530	,074	1	,785	1,155	,409
	[EDUC=2]	-,109	,339	,104	1	,748	,897	,462
	[EDUC=3]	0 <sup>b</sup>	-	-	0	-	-	-
	[CAT_AGE=1]	,516	,411	1,575	1	,209	1,676	,748
	[CAT_AGE=2]	-,076	,364	,044	1	,834	,927	,454
	[CAT_AGE=3]	0 <sup>b</sup>	-	-	0	-	-	-

**Πίνακας 5.38.** Εκτιμήσεις των παραμέτρων του Model 7, για το ΑΝ.ΕΛ.

ANEL	Intercept	-2,551	,712	12,841	1	<.001		
	[ideological=1]	-,326	,325	1,004	1	,316	,722	,382
	[ideological=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[economic=1]	-,391	,338	1,337	1	,248	,677	,349
	[economic=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[salvage=1]	,249	,359	,481	1	,488	1,283	,635
	[salvage=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[obligation=1]	-,091	,493	,034	1	,853	,913	,347
	[obligation=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[antimemorandum=1]	,188	,365	,265	1	,607	1,207	,590
	[antimemorandum=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[versus-system=1]	,153	,336	,208	1	,648	1,166	,603
	[versus-system=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[important=1]	,055	,786	,005	1	,944	1,057	,226
	[important=2]	1,399	,791	3,123	1	,077	4,050	,858
	[important=3]	,662	,580	1,305	1	,253	1,939	,623
	[important=4]	-,199	,532	,139	1	,709	,820	,289
	[important=5]	0 <sup>b</sup>	-	-	0	-	-	-
	[SEX=1]	,147	,307	,228	1	,633	1,158	,634
	[SEX=2]	0 <sup>b</sup>	-	-	0	-	-	-
	[EDUC=1]	,214	,531	,162	1	,687	1,238	,438
	[EDUC=2]	-,026	,343	,006	1	,939	,974	,497
	[EDUC=3]	0 <sup>b</sup>	-	-	0	-	-	-
	[CAT_AGE=1]	,350	,481	,529	1	,467	1,419	,553
	[CAT_AGE=2]	,400	,380	1,112	1	,292	1,492	,709
	[CAT_AGE=3]	0 <sup>b</sup>	-	-	0	-	-	-

### 5.9.6. Παρατηρούμενες και Προβλεπόμενες Συχνότητες για το Model 7.

Ο Πίνακας 5.39 που αφορά τις παρατηρούμενες και προβλεπόμενες συχνότητες (Observed and Predicted Frequencies) είναι, και δεν υπερβάλλουμε, τεραστίων διαστάσεων και είναι αδύνατο να χωρέσει σε χαρτί A4 ή έστω A3. Θα τον δείτε ωστόσο να ξεδιπλώνεται στις οθόνες σας, αν στο output αρχείο του SPSS κινηθείτε με υπομονή δεξιά, αριστερά, πάνω, ή κάτω. Θα λάβετε καταπληκτικές πληροφορίες, για όποιον συνδυασμό ανεξάρτητων μεταβλητών σας ενδιαφέρει. Εμείς, αδυνατούμε να περπατήσουμε αυτόν τον πίνακα για να μαζέψουμε την κρίσιμη πληροφορία, όπως θάλεγε και ο Καθηγητής του Π.Κ. Νίκος Παπαδάκης, που θεραπεύει Ποιοτικές<sup>28</sup> Μεθόδους. Καταγράφουμε μόνον την πρώτη κρίσιμη πληροφορία που έχει να μας δώσει:

#### ΚΡΙΣΙΜΗ ΠΛΗΡΟΦΟΡΙΑ

Αν ένα άτομο είναι 18-34 ετών, είναι απόφοιτος Μέσης εκπαίδευσεως, ανήκει στο ανδρικό φύλο, θεωρεί ότι το πιο σημαντικό πράγμα σε ένα Κόμμα είναι ο Αρχηγός, είναι διακείμενο εχθρικά προς το ‘Σύστημα’, εμφορείται από αντιμνημονιακές απόψεις, ψηφίζει αυτό που ψηφίζει από προσωπική υποχρέωση, ψηφίζει για τη σωτηρία της πατρίδας, επικαλείται λόγους οικονομικής δυσπραγίας στην επιλογή της ψήφου, δεν ψηφίζει για ιδεολογικούς λόγους, τότε κατευθύνεται προς τη ΝΔ με πιθανότητα 10,3%, προς το ΣΥΡΙΖΑ με 46%, προς το ΠΑΣΟΚ με 1,5%, προς το ΠΟΤΑΜΙ με 2,4%, προς το ΚΚΕ με 2,8%, προς τη ΧΑ με 16,3%, προς τους ΑΝΕΛ με 6,8% και τέλος προς διάφορα άλλα Κόμματα με πιθανότητα 13,4%.

**Πίνακας 5.39.** Παρατηρούμενες και Προβλεπόμενες Συχνότητες, Model 7.

CAT_AGE	EDUC	SEX	Important	versus-system	antimemorandum	obligation	salage	economic	ideological	PARTY	Frequency		Percentage		
											Observed	Predicted	Pearson Residual	Observed	Predicted
18-34	MIDDLE	MALE	CHEF	YES	YES	YES	YES	YES	NO	ND	1	.193	2.948	100,0%	10,3%
										SYRIZA	0	.694	-3.930	0,0%	46,4%
										PASOK	0	.015	-1.123	0,0%	1,5%

<sup>28</sup> Εμείς, με Ποσοτικές Μεθόδους, ως γνωστό, ασχολούμαστε. Ωστόσο, σχεδόν πάντα από τις Ποιοτικές Μεθόδους αντλούμε την αναγκαία για την ανάλυσή μας Θεωρία. Διότι, καλά ξέρουμε όλοι –Θετικοί και Θεωρητικοί Αναλυτές– ότι δεν υπάρχει Επιστήμη χωρίς Θεωρία (Umberto Eco).

### 5.9.7. Η δομική περιγραφή του Μηχανισμού Πρόβλεψης

Εστιάζουμε εκ νέου την προσοχή μας στο Model 2, ένα μοντέλο που έχει ως ανεξάρτητες μεταβλητές τις μεταβλητές του φύλου και των ηλικιακών κατηγοριών. Ακόμη θέτουμε σαν κατηγορία αναφοράς και πάλι το ΣΥΡΙΖΑ τον οποίο είχαμε ως συγκριτικό μέτρο για να μιλήσουμε για τα υπόλοιπα Κόμματα της Ελληνικής Βουλής, έτσι όπως αυτά προέκυψαν από τις εκλογές της 25-1-15.

---

**Ας υποθέσουμε ότι λαμβάνουμε στην τύχη έναν άνδρα, ηλικίας μεταξύ 17 και 34, ανήκει δηλ. στην ηλικιακή κατηγορία 1=17-34 ετών. Ποια η πιθανότητα να πάει να ψηφίσει :**

- |            |                |
|------------|----------------|
| • ND=?     | • KKE=?        |
| • ΣΥΡΙΖΑ=? | • ΧΡΥΣΗ ΑΥΓΗ=? |
| • ΠΑΣΟΚ=?  | • ΑΝΕΛ=?       |
| • ΠΟΤΑΜΙ=? | • ΆΛΛΟ=?       |
- 

Ο μηχανισμός που θα περιγράψουμε παρακάτω θα καταφέρει να επιμερίσει Πιθανότητες, σε κάθε ένα από τα προηγούμενα Κόμματα, οι οποίες συνολικά αθροίζουν ποσοστό 100%.

Ο μαθηματικός τύπος ο οποίος υπολογίζει την πιθανότητα P ενός Κόμματος i είναι:

$$P(KOMMA_i) = \frac{\exp(g_i)}{\sum_{k=1}^j \exp(g_k)} \quad (5.26)$$

Οπου, προφανώς,  $i=1,2,3,4,5,6,7,8$  και  $j=8$ .

Οπότε, με βάση τα στοιχεία του Πίνακα 5.39 θα έχουμε :

- I. ND =  $g1 = 0,095 + 0,029 - 0,559 = -0,435 \Rightarrow e^{-0,435} = 0,647.$
- II. SYRIZA =  $g2 = 0$  αφού αυτός αποτελεί την κατηγορία αναφοράς  $e^0 = 1.$
- III. PASOK =  $g3 = -1,733 + 0,613 - 1,611 = -2,731 \Rightarrow e^{-2,731} = 0,065.$
- IV. POTAMI =  $g4 = -1,847 - 0,070 + 0,495 = -1,422 \Rightarrow e^{-1,422} = 0,241.$
- V. KKE =  $g5 = -2,025 - 0,024 + 0,699 = -1,35 \Rightarrow e^{-1,35} = 0,259.$
- VI. XA =  $g6 = -2,801 + 0,930 + 0,564 = -1,307 \Rightarrow e^{-1,307} = 0,27.$
- VII. ANEL =  $g7 = -2,287 + 0,129 + 0,325 = -1,833 \Rightarrow e^{-1,833} = 0,16.$
- VIII. ALLO =  $g8 = -1,067 - 0,062 + 0,149 = -0,98 \Rightarrow e^{-0,98} = 0,375.$

Τελικά, με βάση τον τύπο (5.26) θα έχουμε:

$$P(ND) = \frac{0,647}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,2144 = 21,44\% \quad (5.26)$$

$$P(SYRIZA) = \frac{1}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,3314 = 33,14\% \quad (5.27)$$

$$P(PASOK) = \frac{0,065}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,021 = 2,1\% \quad (5.28)$$

$$P(POTAMI) = \frac{0,241}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,079 = 7,9\% \quad (5.29)$$

$$P(KKE) = \frac{0,259}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,085 = 8,5\% \quad (5.30)$$

$$P(XA) = \frac{0,27}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,089 = 8,9\% \quad (5.31)$$

$$P(ANEL) = \frac{0,16}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,053 = 5,3\% \quad (5.32)$$

$$P(ALLO) = \frac{0,375}{0,647 + 1 + 0,065 + 0,241 + 0,259 + 0,27 + 0,16 + 0,375} = 0,124 = 12,4\% \quad (5.33)$$

### 5.9.8. Είναι το Model 7 το βέλτιστο μοντέλο; Which is the Best Model?

Για το Model 7 παρατηρούμε τα εξής:

- I. Είναι το μοντέλο που μας φέρνει πολύ κοντά στην πραγματικότητα. Με άλλα λόγια, είναι το μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. Πίνακα 5.36). Ειρήσθω εν παρόδω, και όπως πολλές φορές έχουμε επισημάνει, η ΑΠΟΚΛΙΣΗ για τη ΧΡΥΣΗ ΑΥΤΗ είναι η μεγαλύτερη αναφορικά με τα άλλα πολιτικά Κόμματα (-1,18%), διότι

αρκετοί είναι εκείνοι που ψηφίζουν ΧΡΥΣΗ ΑΥΓΗ, αλλά για διάφορους<sup>29</sup> λόγους δεν το λένε.

- II. Αναμφίβολα το Model 7 διαθέτει το σοβαρότερο ίσως χαρακτηριστικό στο πεδίο της κοινωνικής έρευνας που ονομάζεται Ερμηνευσιμότητα.
- III. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα.(βλ. Πίνακας 5.31).
- IV. Για το Model 7 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερδιασκόρπιση (over-dispersion), όπως με ενάργεια δείξαμε στην παράγραφο 5.9.1.
- V. Το Model 7 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης<sup>30</sup> (Long And Freese, 2014).
- VI. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές και ότι η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. Πίνακα 5.32).
- VII. Η ικανότητα ταξινόμησης του Model 7, κινείται σε ικανοποιητικά/αποδεκτά επίπεδα. (βλ. παράγραφο 5.9.4 Classification table). Συνολική Προβλεπτική Δύναμη (Overall Prediction Power) = 58,4%.
- VIII. Το Model 7 δεν περιλαμβάνει στις δομές του καμία συνεχή μεταβλητή, ενώ αποδείχθηκε ότι η συνεχής μεταβλητή της ηλικίας (AGE), καθίσταται χρήσιμη, και έχει συνεισφορά στο μοντέλο, μόνον όταν αυτή μετατραπεί σε κατηγορική. (βλ. σχετική υποσημείωση και σχόλια στην παράγραφο 5.3 του παρόντος. Κι ακόμη, αυτό θα φανεί εναργέστερα στην εξέταση του Μοντέλου 8, το οποίο περιλαμβάνει στη δομή τη μεταβλητή της Ηλικίας, ως συνεχή [covariate]).

---

<sup>29</sup> Ενας από αυτούς τους λόγους είναι η ντροπή: «Εγώ παιδάκι μου θα ψηφίσω ΧΡΥΣΗ ΑΥΓΗ, αλλά μην το γράψεις...!! Έχω πολεμήσει το Ναζισμό ... καταλαβαίνεις ... Το κάνω επειδή θέλω να τιμωρήσω του Πολιτικούς...!!». Στο Εργαστήριο Κοινωνικής Στατιστικής και Πολιτικής Έρευνας του Πανεπιστημίου Κρήτης έγιναν κατά καιρούς 4 έρευνες για τη ΧΡΥΣΗ ΑΥΓΗ. Σε όλες οι Φοιτητές/Δειγματολήπτες μας είπαν ότι ακούστηκαν αυτές οι φράσεις ενοχής...

<sup>30</sup> Ενας απλός τρόπος να το δείξουμε αυτό είναι η διασταύρωση (crosstabulation ), ανάμεσα στην εξαρτημένη μεταβλητή PARTY με κάθε μια από τις ανεξάρτητες μεταβλητές, ξεχωριστά.

IX. Κατά την κατασκευή του Model 7 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο χαμηλής συνθετότητας, χωρίς δηλ. την ύπαρξη αλληλεπιδράσεων (interaction 2ης και 3ης τάξεως, οι οποίες προφανώς, δυσκολεύουν, ενίστε, σοβαρά την Ερμηνεία. Από την άλλη οι αποκλεισμοί των αλληλεπιδράσεων, καθόλου δεν έβλαψαν την προσαρμογή του μοντέλου στα δεδομένα όπως φάνηκε στον έλεγχο καλής προσαρμογής (βλ. Πίνακα 5.31).

#### **ΤΕΛΙΚΟ ΣΥΜΠΕΡΑΣΜΑ για το MODEL 7.**

Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 7 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Στην ερώτηση αν είναι το βέλτιστο ή αλλιώς ‘Το άριστο’, (‘The Best’) μοντέλο, η απάντηση είναι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει η αυτοκρατορία των Γραμμικών Σχέσεων. Αναμφίβολα, η προσπάθεια ανίχνευσης του ΒΕΛΤΙΣΤΟΥ ΜΟΝΤΕΛΟΥ θα ήταν μια σπουδαία ΑΣΚΗΣΗ για τους Σπουδαστές και τους νέους Ερευνητές.

**Πίνακας 5.40.** Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 7.

ΠΟΛΙΤΙΚΟ ΚΟΜΜΑ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΚΛΟΓΩΝ 25-1-15	Η ΔΙΚΗ ΜΑΣ ΠΡΟΒΛΕΨΗ (SPSS MODEL 7)	ΑΠΟΚΛΙΣΗ
ΝΕΑ ΔΗΜΟΚΡΑΤΙΑ	27,81 %	27,6 %	-0,21 %
ΣΥΡΙΖΑ	36,4 %	36,7 %	+0,3 %
ΠΑΣΟΚ	4,68 %	4,4 %	-0,28 %
ΠΟΤΑΜΙ	6,05 %	6,0 %	-0,05 %
ΚΚΕ	5,47 %	5,3 %	-0,17 %
ΧΡΥΣΗ ΑΥΓΗ	6,28 %	5,1 %	-1,18 %
ΑΝΕΞ. ΕΛΛΗΝΕΣ	4,75 %	4,5 %	-0,25 %
ΥΠΟΛΟΙΠΑ ΚΟΜΜΑΤΑ	13,3 %	10,4%	-2,9 %

## 5.10. Η κατασκευή του Model 8.

Στην παράγραφο αυτή θα εξετάσουμε τη συμπεριφορά ενός νέου μοντέλου (Model 8), στο οποίο συμμετέχουν οι μεταβλητές του μοντέλου 7 με μια μόνη αλλαγή: Η μεταβλητή της ηλικίας επιστρέφει στην αρχική φύση της, γίνεται δηλ. και πάλι συνεχής. Επομένως το μοντέλο που θα εξετάσουμε είναι το Model 8.

MODEL 8.

=====

Dependent Var: PARTY

=====

Independent Vars(9): SEX, EDUC, Q2, Q11, Q12, Q13, Q14, Q15 και Q16.  
Independent Covariate Vars: AGE.

=====

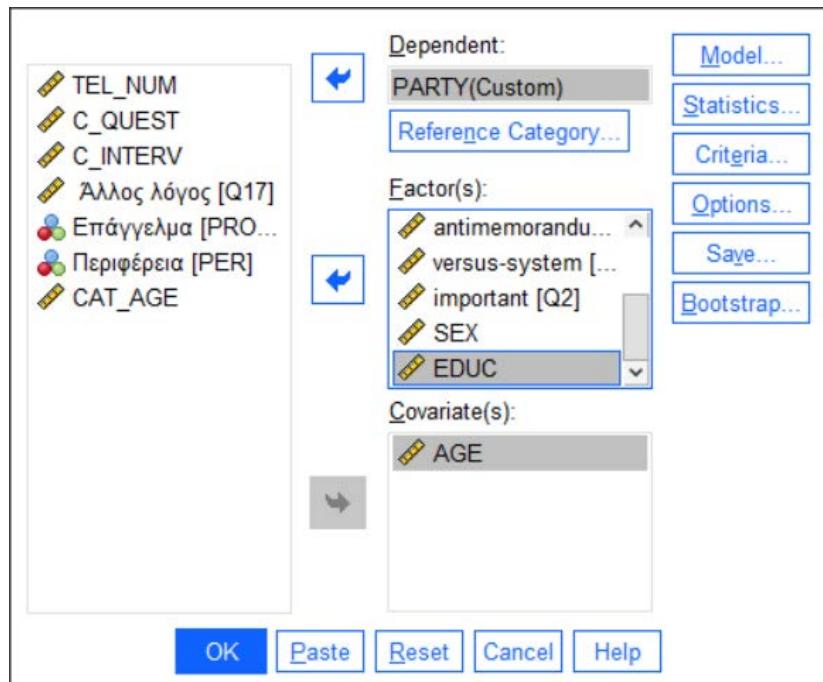
Η εκτέλεση του μοντέλου 8 έχει ως εξής:

Με ανοικτό το data set GOLEMH.SAV, ακολουθούμε τη διαδρομή,

Analyze → Regression → Multinomial Logistic (κλικ)

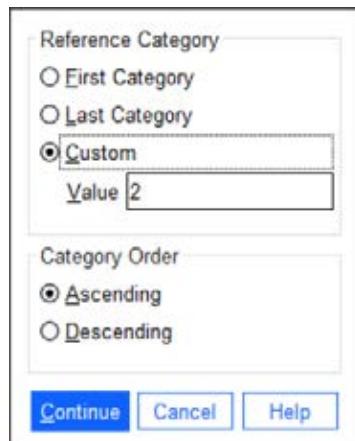
οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.30.

Εικόνα 5.30. Προσδιορισμός Μεταβλητών Model 8.



Ευρισκόμενοι στο περιβάλλον της Εικόνας 5.30 εισάγουμε την εξαρτημένη μεταβλητή PARTY στο πλαίσιο Dependent και κατόπιν πατάμε το κουμπί Reference Category, οπότε μεταφερόμαστε στο περιβάλλον της Εικόνας 5.31.

**Εικόνα 5.31.** Επιλογή Κατηγορίας Αναφοράς Model 8.



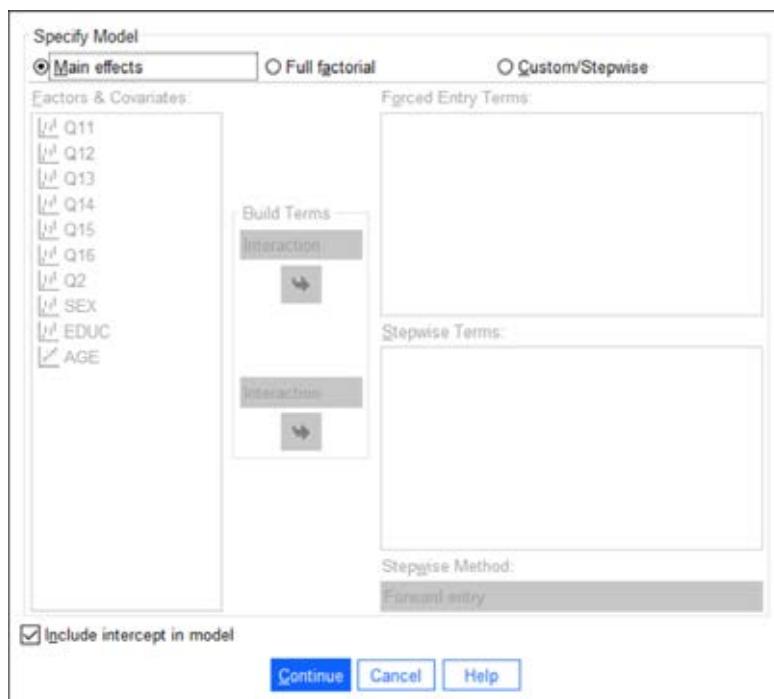
Στο περιβάλλον της Εικόνας 5.31 κάνουμε κλικ στο κουμπί Custom και θέτουμε στο πλαίσιο Value την τιμή 2. Έτσι ορίζουμε σαν κατηγορία αναφοράς το Κόμμα του ΣΥΡΙΖΑ, το οποίο υπενθυμίζεται είχε κωδικό 2.

Στη συνέχεια κάνουμε κλικ στο κουμπί continue οπότε ξαναγυρίζουμε πίσω στο περιβάλλον της Εικόνας 5.30. Εκεί, με το γνωστό τρόπο εισάγουμε στο πλαίσιο Factors τις κατηγορικές μεταβλητές Q11, Q12, Q13, Q14, Q15, Q16, Q2, SEX, EDUC ενώ στο πλαίσιο Covariate θέτουμε τη συνεχή μεταβλητή AGE.

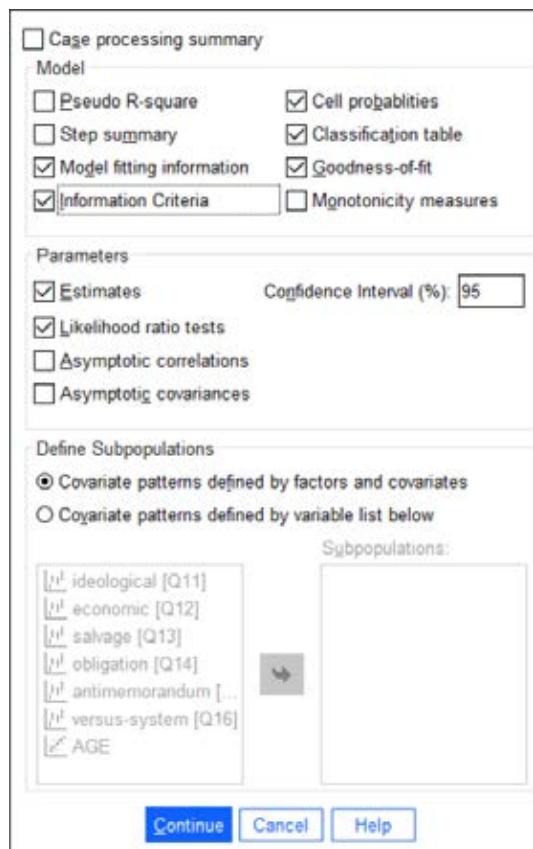
Στη συνέχεια πατάμε το κουμπί Model που βρίσκεται πάνω και δεξιά, οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.32 . Κάνουμε 1 κλικ στο πλαίσιο main effects για να λάβουμε κύριες επιδράσεις, κι άλλο 1 στο include intercept in model για να συμπεριληφθεί ο σταθερός όρος στο μοντέλο.

Με ένα κλικ στο κουμπί continue ξαναγυρίζουμε στη βασικό μας μενού (Εικόνα 5.30). Εκεί με ένα κλικ στο κουμπί Statistics μπαίνουμε στο περιβάλλον Εικόνας 5.33.

**Εικόνα 5.32.** Προσδιορισμός Model 8.

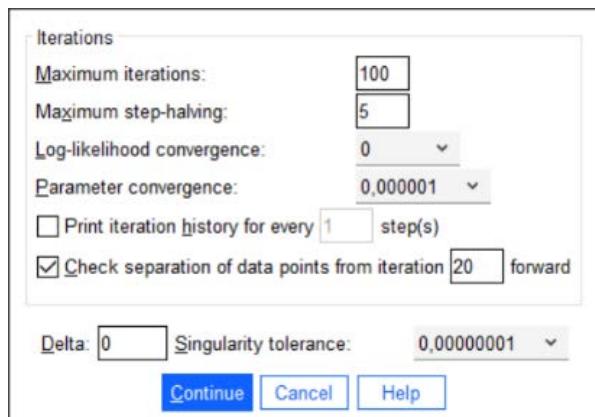


**Εικόνα 5.33.** Επιλογές μενού Statistics, Model 8.

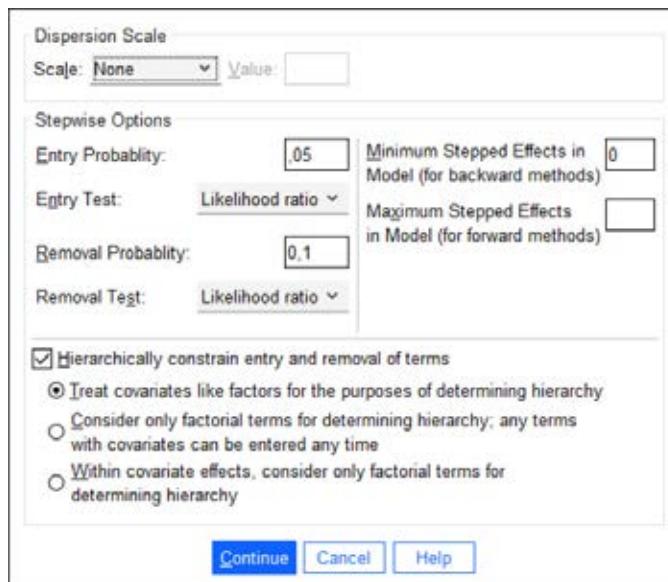


Στο περιβάλλον της Εικόνας 5.33 επιλέγουμε με κλικ τις απαραίτητες ετικέτες για να λάβουμε πληροφορίες αναφορικά με την προσαρμογή του μοντέλου στα δεδομένα (Goodness-of-fit), τις πιθανότητες πρόβλεψης (Cell Probabilities), την ικανότητα του μοντέλου να ταξινομήσει σωστά (Classification Table) και άλλες ιδιαίτερα χρήσιμες πληροφορίες για το μοντέλο μας (information Criteria).

**Εικόνα 5.34.** Επιλογές μενού Criteria, Model 8.



**Εικόνα 5.35.** Επιλογές μενού Options, Model 8.



Στη συνέχεια πάλι με τη βοήθεια του κουμπιού Continue, επιστρέφουμε στο βασικό μας μενού (Εικόνα 5.30). Εκεί με 1 κλικ στο κουμπί Criteria εισερχόμεθα στο περιβάλλον της Εικόνας 5.34. Εκεί αφήνουμε ανέγγιχτες όλες τις τεχνικές ρυθμίσεις που το SPSS έχει προκαθορίσει. Με τη χρήση και πάλι του Continue γυρνάμε και πάλι στο βασικό μας μενού, εκείνο της Εικόνας 5.30. Πραγματοποιούμε στην τελευταία μας επίσκεψη.

Με ένα κλικ στο κουμπί Options μεταφέρόμαστε στο περιβάλλον της Εικόνας 5.35. Όμοια, όπως και παραπάνω, δεν προβαίνουμε σε καμία μεταβολή των παραμέτρων που έχει προκαθορίσει το SPSS. Με τη χρήση και πάλι του Continue γυρνάμε και πάλι στο βασικό μας μενού, εκείνο της Εικόνας 5.30.

Τέλος, με ένα κλικ στο κουμπί OK, τρέχουμε το Model 8.

#### **5.10.1. Αποτελέσματα για το Model 8. Προσαρμογή, Σπουδαιότητα των Μεταβλητών. Likelihood Ratio tests. Ερμηνεία του Μοντέλου 8.**

Προφανώς, η πρώτη πληροφορία την οποία αναζητούμε για το Model 8 είναι προφανώς εκείνη που αφορά την προσαρμογή του στα δεδομένα. Ο Πίνακας 5.41 μας δίνει διπλή πληροφόρηση:

- I. Το κριτήριο του Pearson πληροφορεί ότι ο λόγος  $\chi^2/df$ , δηλ. ο λόγος  $7846,884/6853$  ισούται με 1,15 το οποίο προφανώς είναι μικρότερο από 2, κι επομένως πρόβλημα υπερδιασκόρπισης (overdispersion) δεν υπάρχει. Κατ' επέκταση, σοβαρή απειλή για την καλή προσαρμογή του μοντέλου μας (Model 8), δεν υπάρχει (Field, 2017).
- II. Το κριτήριο Deviance μας πληροφορεί ότι έχουμε μια εξαιρετική/άριστη προσαρμογή του μοντέλου μας (Model 8) στα δεδομένα ( $p=100\%$ ).

**Πίνακας 5.41.** Πίνακας  $\chi^2$ , Model 8.

	Chi-Square	df	Sig.
Pearson	7846,884	6853	<,001
Deviance	2992,367	6853	1,000

Ο Πίνακας 5.42 μας προσφέρει αρκετές και πολύτιμες πληροφορίες για τις ανεξάρτητες μεταβλητές/παραμέτρους του μοντέλου μας (Model 8). Γενικά ο Πίνακας αυτός μας πληροφορεί για το χαρακτήρα των ανηγμένων μοντέλων (reduced Models), δηλ. των μοντέλων τα οποία προκύπτουν από την αφαίρεση μιας (εδώ ακριβώς μιας) ή περισσοτέρων μεταβλητών/παραμέτρων του μοντέλου μας. Εξετάζονται επομένως, σε αυτό τον πίνακα οι συνέπειες της αφαίρεσης μιας μεταβλητής από το αρχικό μας

μοντέλο. Για παράδειγμα, η δεύτερη γραμμή η οποία αφορά τη συνεχή μεταβλητή της ηλικίας (AGE) μας πληροφορεί για την τύχη της παρακάτω μηδενικής υπόθεσης:

$H_0$  : Το μοντέλο που προκύπτει από την αφαίρεση της μεταβλητής AGE, δηλ. το μοντέλο που περιλαμβάνει τις μεταβλητές των ιδεολογικών λόγων (Q11), οικονομικών λόγων (Q12), σωτηρίας της πατρίδας (Q13), προσωπικών υποχρεώσεων (Q14), αντιμνημονιακής στάσης (Q15), αντισυστημικής στάσης (Q16), λόγων επιλογής Κόμματος (Q2), Φύλου (SEX ), και τέλος μορφωτικού επιπέδου (EDUC), παραμένει σημαντικό.

**Πίνακας 5.42.** Σύγκριση των Reduced Model 8 με Null μοντέλο.

Likelihood Ratio Tests				
Effect	Model Fitting Criteria -2 Log Likelihood of Reduced Model	Likelihood Ratio Tests		
		Chi-Square	df	Sig.
Intercept	3087,689*	,000	0	.
AGE	3116,089	28,400	7	<,001
ideological	3113,788	26,099	7	<,001
economic	3109,700	22,010	7	,003
salvage	3127,305	39,615	7	<,001
obligation	3101,822	14,133	7	,049
antimemorandum	3290,997	203,308	7	<,001
versus-system	3141,067	53,378	7	<,001
important	3146,050	58,361	28	<,001
SEX	3102,479	14,790	7	,039
EDUC	3129,674	41,985	14	<,001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Στον Πίνακα 5.32 όλες οι μηδενικές υποθέσεις αυτού του τύπου απορρίπτονται καθώς το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι παντού μικρότερο του 5%. Τελικό συμπέρασμα όλα τα ανηγμένα μοντέλα που προκύπτουν από την αφαίρεση μιας συγκεκριμένης μεταβλητής παραμένουν στατιστικώς σημαντικά.

Τα επίσημα εγχειρίδια της έκδοσης 28 του SPSS, κάνουν λόγο για παράγοντα κινδύνου (risk factor), στην περίπτωση που το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι μικρότερο του 10%.

Υπάρχει ωστόσο, και μια άλλη (ισοδύναμη) ανάγνωση του Πίνακα 5.32. Είναι αυτή που προτείνουν οι Long Freese (2014). Σύμφωνα με αυτήν η δεύτερη γραμμή του Πίνακα 5.32 που αναφέρεται στην παράμετρο της ηλικίας ελέγχει την εξής μηδενική υπόθεση:

$H_0$  : Όλοι οι συντελεστές που σχετίζονται με την ηλικία ΔΕΝ απέχουν και πολύ από το μηδέν, ή αλλιώς είναι μηδέν.

Η τελευταία μηδενική υπόθεση απορρίπτεται ( $p < 0,001$ ), κι επομένως οδηγούμαστε στο συμπέρασμα ότι ΟΛΟΙ οι συντελεστές που σχετίζονται με την ηλικία ΔΕΝ μπορεί να είναι μηδέν.

**Πίνακας 5.43.** Εκτιμήσεις των παραμέτρων του Model 8

PARTY*	Parameter Estimates						95% Confidence Interval for Exp(B)	
	B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound
ND	Intercept	-,796	,512	2,418	1	,120		
	AGE	,012	,006	3,731	1	,053	1,012	1,000 1,025
	[ideological=1]	,391	,198	3,885	1	,049	1,479	1,002 2,181
	[ideological=2]	0 <sup>b</sup>			0			
	[economic=1]	-,305	,203	2,252	1	,133	,737	,495 1,098
	[economic=2]	0 <sup>b</sup>			0			
	[salvage=1]	1,069	,234	20,882	1	<,001	2,913	1,841 4,607
	[salvage=2]	0 <sup>b</sup>			0			
	[obligation=1]	,726	,279	6,798	1	,009	2,067	1,198 3,569
	[obligation=2]	0 <sup>b</sup>			0			
	[antimemorandum=1]	-2,616	,227	132,560	1	<,001	,073	,047 ,114
	[antimemorandum=2]	0 <sup>b</sup>			0			
	[versus-system=1]	-1,360	,217	39,182	1	<,001	,257	,168 ,393
	[versus-system=2]	0 <sup>b</sup>			0			
	[important=1]	1,282	,450	8,133	1	,004	3,604	1,493 8,700
	[important=2]	1,097	,600	3,342	1	,068	2,994	,924 9,702
	[important=3]	,559	,417	1,799	1	,180	1,749	,773 3,957
	[important=4]	,162	,370	,192	1	,661	1,176	,569 2,429
	[important=5]	0 <sup>b</sup>			0			
	[SEX=1]	-,209	,187	1,241	1	,265	,812	,562 1,172
	[SEX=2]	0 <sup>b</sup>			0			
	[EDUC=1]	,470	,313	2,249	1	,134	1,599	,866 2,955
	[EDUC=2]	,199	,208	,914	1	,339	1,220	,811 1,835
	[EDUC=3]	0 <sup>b</sup>			0			

## **ΕΡΜΗΝΕΙΑ του Πίνακα 5.43**

**Στον Πίνακα 5.39 , παρατηρούμε:**

- I. Ο παράγοντας της ηλικίας, μόλις και μετά βίας, δεν παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ (Sig. = 0,053), σε επίπεδο βέβαια 5%. Με απλά λόγια, με βάση το Model 8 οι ψηφοφόροι ανεξαρτήτως ηλικίας, εξίσου κατευθύνονται προς τη ΝΔ και τον ΣΥΡΙΖΑ.
- II. Ο παράγοντας των ιδεολογικών λόγων, μόλις και μετά βίας, παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ (Sig. = 0,049), σε επίπεδο βέβαια 5%. Με απλά λόγια, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν ιδεολογικά συγκρινόμενοι με εκείνους που δεν ψηφίζουν ιδεολογικά, κατευθύνονται με αυξημένες (βλ. πρόσημο συντελεστή  $b=+0,012$ ) πιθανότητες, της τάξης του 47,9 % [βλ.  $\exp(0,012) = 1,479$ ] προς την κάλπη της ΝΔ παρά του ΣΥΡΙΖΑ. Εννοείται και εδώ ότι οι άλλοι παράγοντες του μοντέλου παραμένουν σταθεροί.
- III. Ο παράγοντας των οικονομικών λόγων, δεν παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ (Sig. = 0,133), σε επίπεδο βέβαια 5 %. Με απλά λόγια, με βάση το Model 8 οι ψηφοφόροι που ψήφισαν αυτό που ψήφισαν για οικονομικούς λόγους, συγκρινόμενοι με εκείνους που δεν ψήφισαν για οικονομικούς λόγους κατευθύνονται με ίσες πιθανότητες προς τη ΝΔ και τον ΣΥΡΙΖΑ.
- IV. Ο παράγοντας της σωτηρίας της πατρίδας παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $p < 0,001$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν για τη σωτηρία της πατρίδας συγκρινόμενοι με εκείνους που δεν ψηφίζουν για τη σωτηρία της πατρίδας, έχουν σχεδόν τριπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [βλ.  $\exp(+1,069) = 2,913$ ]. Εννοείται και εδώ ότι οι υπόλοιποι παράγοντες παραμένουν σταθεροί.
- V. Ο παράγοντας των προσωπικών υποχρεώσεων παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $p < 0,001$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν για τη σωτηρία της πατρίδας συγκρινόμενοι με

εκείνους που δεν ψηφίζουν για προσωπικούς λόγους, έχουν διπλάσιες και πλέον πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [βλ.  $\exp(+0,726) = 2,069$ ]. Εννοείται και εδώ ότι οι υπόλοιποι παράγοντες παραμένουν σταθεροί.

VI. Ο παράγοντας της στάσης απέναντι στα μνημόνια παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $p < 0,001$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν αντιμνημονιακά συγκρινόμενοι με εκείνους που δεν ψηφίζουν αντιμνημονιακά, έχουν κατά 92,7% μειωμένη πιθανότητα να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [βλ.  $1-\exp(-2,616)=1-0,073=0,927=92,7\%$ ]. Εννοείται και εδώ ότι οι υπόλοιποι παράγοντες παραμένουν σταθεροί.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ο ΣΥΡΙΖΑ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των αντιμνημονιακών ψηφοφόρων.

VII. Ο παράγοντας της στάσης απέναντι στο ‘σύστημα’ παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $p < 0,001$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν αντισυστημικά συγκρινόμενοι με εκείνους που δεν ψηφίζουν αντισυστημικά, έχουν κατά 74,3% μειωμένη πιθανότητα να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ο λόγος αντισυστημικών/μη αντισυστημικών ψηφοφόρων είναι μικρότερος<sup>31</sup> της μονάδας,  $[\exp(-1,360) = 0,257 < 1]$ . Εννοείται και εδώ ότι οι υπόλοιποι παράγοντες παραμένουν σταθεροί.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 το κόμμα του ΣΥΡΙΖΑ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των αντισυστημικών ψηφοφόρων.

---

<sup>31</sup> Γενικά ισχύει ο εξής κανόνας : Όταν ο λόγος  $\exp(b)$  είναι μικρότερος της μονάδας τότε γίνεται λόγος για μειωμένες πιθανότητες, ενώ όταν ο λόγος  $\exp(b)$  είναι μεγαλύτερος της μονάδας τότε γίνεται λόγος για αυξημένες πιθανότητες. Σε κάθε περίπτωση θα πρέπει, το επίπεδο στατιστικής σημαντικότητας, να είναι χαμηλό. (Long and Freese, 2014).

VIII. Ο παράγων σπουδαιότητας<sup>32</sup> Αρχηγός παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $p < 0,004$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν για τον Αρχηγό, συγκρινόμενοι με εκείνους που ψηφίζουν για ‘άλλο’ λόγο, εμφανίζουν 3,6 φορές μεγαλύτερη πιθανότητα να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 το κόμμα της ΝΔ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των ψηφοφόρων με πίστη στον Αρχηγό του Κόμματος.

IX. Ο παράγων σπουδαιότητας Υποψήφιοι δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig. = 0,068 = 6,8\% > 5\%$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν για τους Υποψήφιους, συγκρινόμενοι με εκείνους που ψηφίζουν για ‘άλλο’ λόγο, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε το κόμμα της ΝΔ ούτε το κόμμα του ΣΥΡΙΖΑ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των ψηφοφόρων με πίστη προς Υποψηφίους.

X. Ο παράγων σπουδαιότητας Ιδεολογία/Κοσμοθεωρία του Κόμματος δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig.=0,18>0,05$ ). Επομένως, με βάση το MODEL 8, οι ψηφοφόροι που ψηφίζουν για την Ιδεολογία του κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για ‘άλλο’ λόγο, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε το κόμμα της ΝΔ ούτε το κόμμα του ΣΥΡΙΖΑ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των ψηφοφόρων με πίστη προς την Ιδεολογία του Κόμματος.

---

<sup>32</sup> Θυμίζουμε εδώ ότι η μεταβλητή της σπουδαιότητας, δηλ. η μεταβλητή που μας αποκαλύπτει το σημαντικότερο λόγο που ψήφισαν το Κόμμα που ψήφισαν, έχει 5 επίπεδα: 1=για χάρη του Αρχηγού, 2=για χάρη των Υποψηφίων, 3= για χάρη της Ιδεολογίας /Κοσμοθεωρίας του Κόμματος, 4= για χάρη του Προγράμματος, για Άλλο’ λόγο, που δεν συμπεριλαμβάνεται στους παραπάνω 4 λόγους.

XI. Ο παράγων σπουδαιότητας Πρόγραμμα του Κόμματος δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig. = 0,661 > 0,05$ ). Επομένως, με βάση το Model 8, οι ψηφοφόροι που ψηφίζουν για το Πρόγραμμα του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για ‘άλλο’ λόγο, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε το κόμμα της ΝΔ ούτε το κόμμα του ΣΥΡΙΖΑ έλαβε τη μερίδα του λέοντος από το στρατόπεδο των ψηφοφόρων με πίστη προς το Πρόγραμμα του Κόμματος.

XII. Ο παράγων ΦΥΛΟ=SEX δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig. = 0,265 = 26,5\% > 5\%$ ). Επομένως, με βάση το Model 8, οι άνδρες ψηφοφόροι, συγκρινόμενοι με τις γυναίκες ψηφοφόρους, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε οι άνδρες, ούτε οι γυναίκες ψηφοφόροι, πριμοδότησαν ιδιαίτερα το ένα ή το άλλο από τα δυο κόμματα, ΝΔ και ΣΥΡΙΖΑ.

XIII. Ο παράγων ΣΤΟΙΧΕΙΩΔΗΣ ΕΚΠΑΙΔΕΥΣΗ δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig. = 0,134 = 13,4\% > 5\%$ ). Επομένως, με βάση το Model 8, ψηφοφόροι της στοιχειώδους εκπαίδευσης, συγκρινόμενοι με εκείνους της ανώτερης/ανώτατης εκπαίδευσης, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε οι απόφοιτοι της στοιχειώδους εκπαίδευσης, ούτε εκείνοι της ανώτερης/ανώτατης εκπαίδευσης, πριμοδότησαν ιδιαίτερα, το ένα ή το άλλο από τα δυο κόμματα, ΝΔ και ΣΥΡΙΖΑ.

XIV. Ο παράγων ΜΕΣΗ ΕΚΠΑΙΔΕΥΣΗ δεν φαίνεται να παρεμβαίνει στην αναμέτρηση της ΝΔ με τον ΣΥΡΙΖΑ ( $Sig. = 0,339 = 33,9\% > 5\%$ ). Επομένως, με βάση το Model 8, ψηφοφόροι της μέσης εκπαίδευσης, συγκρινόμενοι με εκείνους της ανώτερης/ανώτατης εκπαίδευσης, εμφανίζουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

Τοπικό Συμπέρασμα: Κατά το Model 8, στην εκλογική αναμέτρηση της 25-1-15 ούτε οι απόφοιτοι της μέσης εκπαίδευσης, ούτε εκείνοι της ανώτερης/ ανώτατης εκπαίδευσης, πριμοδότησαν ιδιαίτερα, το ένα ή το άλλο από τα δυο κόμματα, ΝΔ και ΣΥΡΙΖΑ.

#### 5.10.2. Περιθώριες Συχνότητες (Marginal Effects) για το Model 8.

Στον παρακάτω Πίνακα 5.44 θα δούμε την αναλυτική Πρόβλεψη, ανά πολιτικό Κόμμα, του Μοντέλου 8.

Πίνακας 5.44. Περιθώριων Συχνοτήτων, Model 8.

		N	Marginal Percentage
PARTY	ND	309	27,5%
	SYRIZA	411	36,6%
	PASOK	49	4,4%
	POTAMI	67	6,0%
	KKE	61	5,4%
	XRYSH	58	5,2%
	AYGH		
ideological	ANEL	51	4,5%
	ALLO	117	10,4%
economic	YES	612	54,5%
	NO	511	45,5%
salvage	YES	719	64,0%
	NO	404	36,0%
obligation	YES	831	74,0%
	NO	292	26,0%
antimemorandum	YES	141	12,6%
	NO	982	87,4%
m	YES	518	46,1%
	NO	605	53,9%
versus-system	YES	490	43,6%
	NO	633	56,4%
Valid		1123	100,0%
Missing		794	
Total		1917	
Subpopulation		744 <sup>a</sup>	

a. The dependent variable has only one value observed in 585 (78,6%) subpopulations.

Τέλος, στον Πίνακα 5.45 παρατηρούμε μια αντιπαραβολή των στοιχείων που προσφέρει το Model 8, με τα πραγματικά στοιχεία των εκλογών της 25-1-15.

**Πίνακας 5.45.** Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 8.

ΠΟΛΙΤΙΚΟ ΚΟΜΜΑ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΚΛΟΓΩΝ 25-1-15	Η ΔΙΚΗ ΜΑΣ ΠΡΟΒΛΕΨΗ (SPSS MODEL 8)	ΑΠΟΚΛΙΣΗ
ΝΕΑ ΔΗΜΟΚΡΑΤΙΑ	27,81 %	27,50 %	-0,31 %
ΣΥΡΙΖΑ	36,4 %	36,6 %	0,2 %
ΠΑΣΟΚ	4,68 %	4,40 %	-0,28 %
ΠΟΤΑΜΙ	6,05 %	6 %	-0,05 %
ΚΚΕ	5,47 %	5,40 %	-0,07 %
ΧΡΥΣΗ ΑΥΓΗ	6,28 %	5,2 %	-1,08 %
ΑΝΕΞ. ΕΛΛΗΝΕΣ	4,75 %	4,50 %	-0,25 %
ΥΠΟΛΟΙΠΑ ΚΟΜΜΑΤΑ	13,3 %	10,4 %	-2,9 %

#### 5.10.3. Η Προβλεπτική Δύναμη (Prediction Power) για το Model 8. Ο Πίνακας Ταξινόμησης (Classification Table).

Ο παρακάτω Πίνακας 5.46 μας πληροφορεί για την ικανότητα του Μοντέλου 8 να ταξινομήσει σωστά. Αυτή η ικανότητα ανέρχεται συνολικά στο 52,5 %. Με την ευκαιρία να σημειώσουμε ότι περίπου στα ίδια επίπεδα, ελαφρά μειωμένη, υπήρξε η προβλεπτική δύναμη του Μοντέλου 7.

**Πίνακας 5.46.** Πίνακας Ταξινόμησης (Classification Table) για το Model 8.

Observed	Classification										Percent Correct
	Predicted										
ND	SYRIZA	PASOK	POTAMI	KKE	XRYSH AYGH	ANEL	ALLO				
ND	249	58	0	0	0	1	0	0	0	80,8%	
SYRIZA	75	331	0	1	0	4	0	0	0	80,5%	
PASOK	32	16	1	0	0	0	0	0	0	2,0%	
POTAMI	34	32	0	1	0	0	0	0	0	1,5%	
KKE	12	48	0	0	0	0	0	0	0	0,0%	
XRYSH AYGH	11	39	0	1	0	7	0	0	0	12,1%	
ANEL	7	41	0	0	0	3	0	0	0	0,0%	
ALLO	49	66	0	1	0	1	0	0	0	0,0%	
Overall Percentage	41,8%	56,3%	0,1%	0,4%	0,0%	1,4%	0,0%	0,0%	0,0%	52,5%	

#### **5.10.4. Είναι το Model 8 το βέλτιστο μοντέλο; Which is the Best Model?**

Για το Model 8 παρατηρούμε τα εξής:

- I. Είναι και αυτό ένα μοντέλο που μας φέρνει κοντά στην πραγματικότητα. Με άλλα λόγια και αυτό είναι ένα μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. Πίνακα 5.44 Περιθώριες Συχνότητες). Εξαίρεση και πάλι η ΧΡΥΣΗ ΑΥΓΗ, της οποίας η απόκλιση είναι η μεγαλύτερη αναφορικά με τα υπόλοιπα Κόμματα (-1,08%), και για λόγους που εξηγήσαμε παραπάνω.
- II. Αναμφίβολα το Model 8 διαθέτει και εκείνο, όπως και το προηγούμενο (Model 7) το χαρακτηριστικό της Ερμηνευσιμότητας. (βλ. Πίνακα 5.43).
- III. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα. (βλ. Πίνακα 5.32).
- IV. Για το Model 8 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερδιασκόρπιση (over-dispersion).
- V. Το Model 8 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης (Long And Freese, 2014). Δηλ. δεν υπάρχει επίπεδο σε καμία από τις ανεξάρτητες μεταβλητές που να σχετίζεται με μόνο ένα επίπεδο της εξαρτημένης.
- VI. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές και ότι η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. Πίνακα 5.32).
- VII. Η ικανότητα ταξινόμησης του Model 8, κινείται σε ικανοποιητικά/αποδεκτά επίπεδα (Prediction Power = 52,5%), περίπου ίδια με του Model 7 που είναι 52,4%. Το Model 8 περιλαμβάνει στις δομές του μία συνεχή μεταβλητή, της ηλικίας (AGE), που φαίνεται να είναι στατιστικώς ασήμαντη και χωρίς συνεισφορά στο μοντέλο (βλ. στον Πίνακα 5.33 Sig. = 0,053 = 5,3% > 5%).

VIII. Κατά την κατασκευή του Model 8 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο χαμηλής συνθετότητας, χωρίς δηλ. την ύπαρξη αλληλεπιδράσεων (interaction 2ης και 3ης τάξεως, οι οποίες προφανώς, δυσκολεύουν, ενίστε, σοβαρά την Ερμηνεία. Από την άλλη οι αποκλεισμοί των αλληλεπιδράσεων, καθόλου δεν έβλαψαν την προσαρμογή του μοντέλου στα δεδομένα όπως φάνηκε στον έλεγχο καλής προσαρμογής (βλ. Πίνακα 5.41).

#### **ΤΕΛΙΚΟ ΣΥΜΠΕΡΑΣΜΑ για το MODEL 8.**

Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 8 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Προφανώς η Ερμηνεία κατά τόπους αλλάζει. Για παράδειγμα, είδαμε ότι η Ηλικία όταν εισάγεται στο μοντέλο ως κατηγορική μεταβλητή, είναι στατιστικώς σημαντική, όταν εισάγεται ως συνεχής, δεν είναι! Στην επίμονη ερώτηση αν είναι το βέλτιστο ή αλλιώς ‘Το άριστο’, (‘The Best’) μοντέλο , η απάντηση είναι και πάλι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει η αυτοκρατορία των Γραμμικών Σχέσεων. Αναμφίβολα, η προσπάθεια ανίχνευσης του ΒΕΛΤΙΣΤΟΥ ΜΟΝΤΕΛΟΥ θα ήταν μια σπουδαία ΑΣΚΗΣΗ για τους Σπουδαστές και τους νέους Ερευνητές.

#### **5.11. Η κατασκευή του Model 9.**

Στην παράγραφο αυτή παραθέτουμε ένα-ένα τα διαδοχικά βήματα που οδηγούν στην οικοδόμηση του Μοντέλου 9, το οποίο προφανώς είναι μεγαλύτερης συνθετότητας, αφού διαθέτει αλληλεπίδραση 2<sup>ης</sup> τάξης. (SEX\*EDU). Το Μοντέλο 9 έχει την παρακάτω δομή:

MODEL 9.  
=====

Dependent Var: PARTY  
=====

Independent Vars(10): SEX, EDUC, Q2, CAT\_AGE, Q11, Q12, Q13, Q14, Q15, Q16,  
SEX\*EDUC  
=====

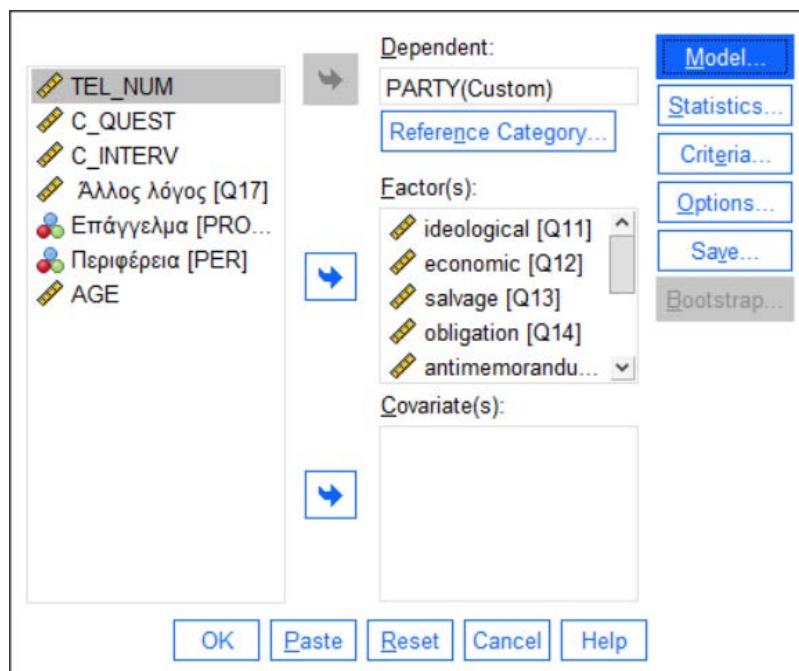
Η εκτέλεση του μοντέλου 9 έχει ως εξής:

Με ανοικτό το data set GOLEMH.SAV, ακολουθούμε τη διαδρομή,

Analyze → Regression → Multinomial Logistic (κλικ)

οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.36.

**Εικόνα 5.36.** Προσδιορισμός Μεταβλητών Model 9.



Στο παράθυρο της Εικόνας 5.36, θέτουμε τη μεταβλητή πολιτικού κόμματος PARTY στο πλαίσιο Dependent και ορίζουμε, κατά τα γνωστά, με τη βοήθεια του υπο-μενού Reference Category, και πάλι ως κόμμα αναφοράς τον ΣΥΡΙΖΑ, που όπως ξέρουμε έχει κωδικό 2.

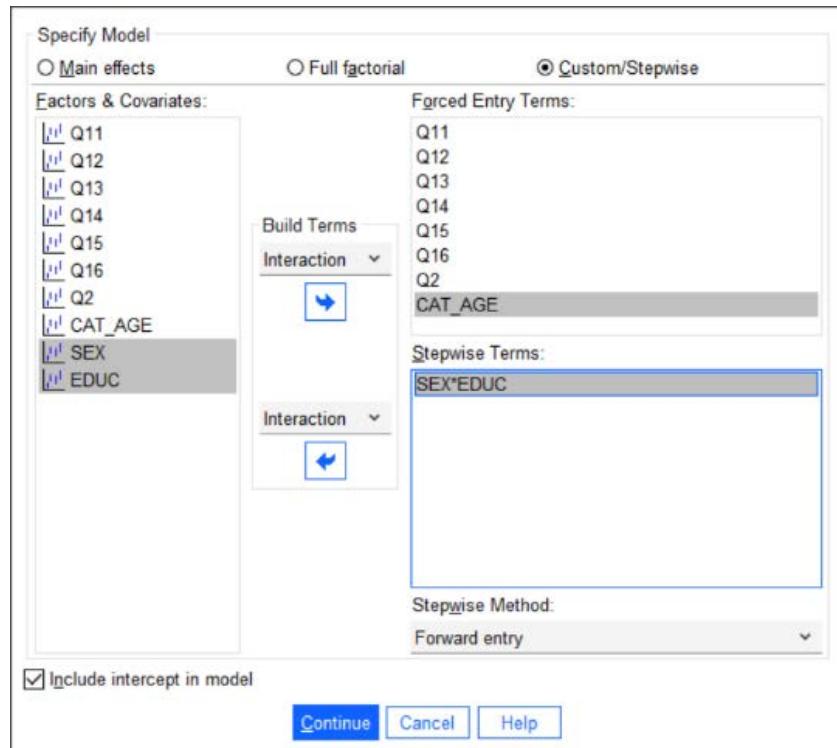
Στη συνέχεια, μεταφέρουμε, με το γνωστό τρόπο, από τα αριστερά, μία-μία, όλες<sup>33</sup> τις ανεξάρτητες μεταβλητές του μοντέλου μας, Q11, Q12, Q13, Q14, Q15, Q16, Q2, CAT\_AGE, SEX και EDUC.

Στη συνέχεια κάνουμε κλικ στο κουμπί continue , οπότε επιστρέφουμε στο βασικό μας μενού (βλ. Εικόνα 5.36).

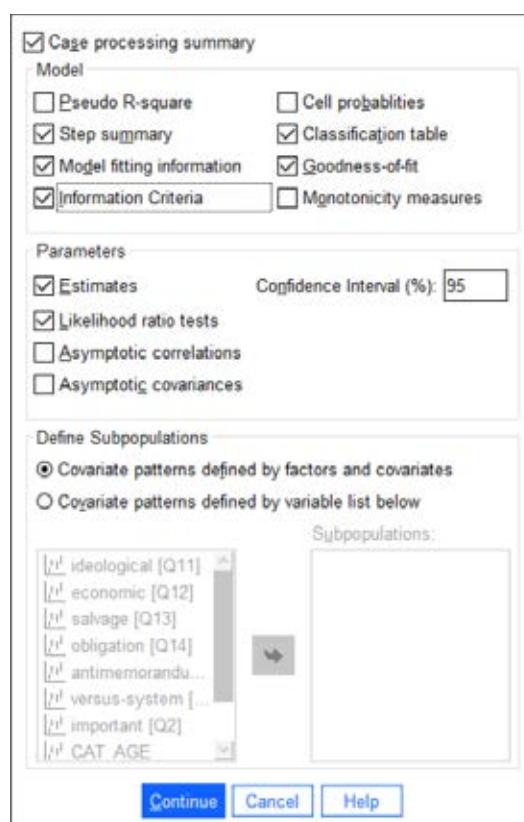
<sup>33</sup> Ο προσεκτικός αναγνώστης θα έχει παρατηρήσει ότι σκόπιμα εισήχθησαν τελευταίες οι μεταβλητές SEX και EDUC, διότι με αυτές θα αναπτυχθεί ευκολότερα το interaction SEX \* EDUC που προαναφέρθηκε.

Στη συνέχεια, κάνουμε κλικ στο κουμπί Model, οπότε μας εμφανίζεται το παράθυρο της Εικόνας 5.37.

**Εικόνα 5.37.** Προσδιορισμός Model 9.



**Εικόνα 5.38.** Επιλογές μενού Statistics, Model 9.



Στο παράθυρο της Εικόνας 5.37 εναπτύσσουμε 3 δράσεις:

- I. Κάνουμε κλικ στο κουμπί Custom/stepwise.
- II. Επιλέγουμε, μία- μία, τις μεταβλητές Q11, Q12, Q13, Q14, Q15, Q16, Q2 και CAT\_AGE, τις μεταφέρουμε μία-μία με το βελάκι που βρίσκεται κάτω από την ένδειξη Build Terms interaction, στο πλαίσιο Forced Entry Terms.
- III. Επιλέγουμε, με τη μέθοδο Shift-Click, μαζί, τις μεταβλητές SEX και EDUC, πατάμε πάλι το βελάκι κάτω από το δεύτερο interaction, και βλέπουμε αυτές να μεταφέρονται στο αποκάτω παράθυρο, κάτω και δεξιά, αλλά έχοντας ανάμεσά τους ένα αστέρι, πράγμα που σημαίνει ότι ζητείται τιμή στατιστικώς σημαντική ή όχι, για τη μεταξύ τους αλληλεπίδραση (βλ. Εικόνα 5.37).

Στη συνέχεια κάνουμε κλικ στο κουμπί Statistics οπότε λαμβάνουμε το παράθυρο της Εικόνας 5.38, όπου επιλέγουμε τις γνωστές παραμέτρους (εμφαίνονται).

Στο παράθυρο της Εικόνας 5.38 πατάμε το κουμπί Continue, επιστρέφουμε στο βασικό μενού (βλ. Εικόνα 5.36). Εκεί, με ένα κλικ στο κουμπί OK εκτελούμε Nominal Logistic Regression διαδικασία, με πολύ σοβαρές πληροφορίες για τα ερευνητικά μας ενδιαφέροντα!

#### **5.11.1. Αποτελέσματα για το Model 9. Περιθώριες Συχνότητες - Πρόβλεψη. Προσαρμογή, Σπουδαιότητα των Μεταβλητών. Likelihood Ratio tests. Ερμηνεία του Model 9.**

Ο Πίνακας 5.47 που βασίζεται στο Μοντέλο 9 ουσιαστικά είναι ένας πίνακας προβλέψεων. Επειδή ακριβώς οι παρατιθέμενες, από το SPSS, στα δεξιά αυτού, περιθώριες συχνότητες (Marginal Percentages) αυτό ακριβώς σημαίνουν. Έτσι:

- I. Η ΝΔ αναμένεται, με βάση αυτό το μοντέλο, να λάβει 27,6% των ψήφων.
- II. Ο ΣΥΡΙΖΑ αναμένεται, με βάση αυτό το μοντέλο, να λάβει 36,7% των ψήφων.
- III. Το ΠΑΣΟΚ αναμένεται, με βάση αυτό το μοντέλο να λάβει 4,4% των ψήφων.
- IV. Το ΠΟΤΑΜΙ αναμένεται, με βάση αυτό το μοντέλο να λάβει 6% των ψήφων
- V. Το ΚΚΕ αναμένεται, με βάση αυτό το μοντέλο, να λάβει 5,3% των ψήφων.

- VI. Η ΧΡΥΣΗ ΑΥΤΗ αναμένεται, με βάση αυτό το μοντέλο, να λάβει 5,1% των ψήφων.
- VII. Οι ΑΝΕΞΑΡΤΗΤΟΙ ΕΛΛΗΝΕΣ αναμένεται, με βάση αυτό το μοντέλο, να λάβουν 4,5% των ψήφων.
- VIII. Τα υπόλοιπα Κόμματα (ALLO) αναμένεται να λάβουν 10,4% των ψήφων.

**Πίνακας 5.47. Περιθώριων Συχνοτήτων, Model 9.**

PARTY	ND	N	Marginal Percentage
		306	27,6%
SYRIZA	407	36,7%	
PASOK	49	4,4%	
POTAMI	67	6,0%	
KKE	59	5,3%	
XRYSH	57	5,1%	
AYGH			
ANEL	50	4,5%	
ALLO	115	10,4%	

Ο Πίνακας 5.48, όπως έχουμε πολλές φορές επισημάνει, ελέγχει την προσαρμογή του Μοντέλου. Εδώ του Μοντέλου 9. Στα εμπειρικά δεδομένα. Δεν φαίνεται να υπάρχει κανένα πρόβλημα για την προσαρμογή του, ούτε επίσης μπορεί να τεθεί πρόβλημα υπερδιασκόρπισης (overdispersion).

**Πίνακας 5.48. Πίνακας  $\chi^2$ , Model 9.**

	Chi-Square	df	Sig.
Pearson	5364,417	4760	<.001
Deviance	2306,241	4760	1,000

Ο Πίνακας 5.49 επίσης μας είναι γνωστός. Ελέγχει ταυτόχρονα, όπως έχουμε επισημάνει ξανά δυο ισοδύναμες μηδενικές υποθέσεις. Η πρώτη έχει να κάνει με τη μηδενικότητα των συντελεστών παλινδρόμησης και η δεύτερη με τη σημαντικότητα του μοντέλου που απομένει αν αφαιρεθεί από αυτό μια συγκεκριμένη μεταβλητή. Η ερμηνεία του, γραμμή-γραμμή, έχει ως εξής:

- I. Αν αφαιρεθεί από το αρχικό μας μοντέλο, ο σταθερός όρος (intercept), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο.

- II. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή των ιδεολογικών λόγων (Q11=ideological), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} < 0,001$  οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή των ιδεολογικών λόγων, δεν μπορεί να είναι όλοι μηδέν.
- III. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή των οικονομικών λόγων (Q12=economic), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} = 0,002$  οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή των οικονομικών λόγων, δεν μπορεί να είναι όλοι μηδέν.
- IV. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή της σωτηρίας της πατρίδας (Q13=salvage), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} < 0,001$  οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή της σωτηρίας της πατρίδας, δεν μπορεί να είναι όλοι μηδέν.
- V. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή των προσωπικών υποχρεώσεων (Q14=obligation), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} = 0,043 < 0,05$ , οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή των προσωπικών υποχρεώσεων, δεν μπορεί να είναι όλοι μηδέν.
- VI. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή της στάσης απέναντι στα μνημόνια (Q15=antimemorandum), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} < 0,001$  οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή της στάσης απέναντι στα μνημόνια, δεν μπορεί να είναι όλοι μηδέν.
- VII. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή της στάσης απέναντι στο ‘σύστημα’ (Q16=versus-system), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $\text{Sig.} < 0,001$  οδηγούμαστε στο συμπέρασμα

ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή της στάσης απέναντι στο ‘σύστημα’, δεν μπορεί να είναι όλοι μηδέν.

- VIII. Αν αφαιρεθεί από το αρχικό μας μοντέλο, η μεταβλητή των ηλικιακών κατηγοριών (CAT\_AGE), τότε το μοντέλο που απομένει είναι ένα σημαντικό μοντέλο. Επίσης, επειδή  $Sig. = 0,05$  οδηγούμαστε στο συμπέρασμα ότι όσοι συντελεστές σχετίζονται με τη μεταβλητή των ηλικιακών κατηγοριών, μπορεί να είναι όλοι μηδέν. Σημείωση: αυτό σε επίπεδο στατιστικής σημαντικότητας 5%. Άλλα αν λάβουμε επίπεδο στατιστικής σημαντικότητας 10% το συμπέρασμα είναι ακριβώς το αντίθετο.
- IX. Η αλληλεπίδραση (interaction) μεταξύ Φύλου και Μόρφωσης είναι στατιστικώς σημαντική ( $Sig. < 0,001$ ). Αυτό σημαίνει ότι έχουμε μια νέα<sup>34</sup> μεταβλητή, την SEX\*EDUC, η οποία αν αφαιρεθεί από το μοντέλο, αφήνει πίσω της ένα επίσης στατιστικώς σημαντικό μοντέλο. Κι ακόμη οι συντελεστές οι οποίοι σχετίζονται με αυτήν τη νέα μεταβλητή δεν μπορεί να είναι όλοι μηδέν.

**Πίνακας 5.49.** Σύγκριση των Reduced Model 9 με Null μοντέλο.

Effect	Likelihood Ratio Tests			Likelihood Ratio Tests		
	Model Fitting Criteria			Chi-Square	df	Sig.
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model			
Intercept	2881,653	3513,180	2629,653 <sup>a</sup>	,000	0	.
ideological	2893,205	3489,647	2655,205	25,552	7	<.001
economic	2889,845	3486,287	2651,845	22,192	7	,002
salvage	2907,132	3503,574	2669,132	39,479	7	<.001
obligation	2882,120	3478,561	2644,120	14,466	7	,043
antimemorandum	3065,811	3662,252	2827,811	198,157	7	<.001
versus-system	2921,136	3517,577	2683,136	53,482	7	<.001
important	2884,128	3375,315	2688,128	58,474	28	<.001
CAT_AGE	2885,215	3446,572	2661,215	31,561	14	,005
SEX * EDUC	2896,440	3352,543	2714,440	84,787	35	<.001

<sup>34</sup> Θα έχετε γίνει αντιληπτό ότι στον Πίνακα 5.45 δεν υπάρχουν αυτόνομες οι μεταβλητές του φύλου (SEX) και του Μορφωτικού επιπέδου (EDUC). Με άλλα λόγια, δεν υπάρχουν κύριες επιδράσεις (main effects). Υπάρχουν όμως interaction effects. Και φυσικά, η Ερμηνεία, όπως θα δούμε στον παρακάτω Πίνακα 5.46, αλλάζει...

**Πίνακας 5.50.** Εκτιμήσεις των παραμέτρων του Model 9.

Parameter Estimates						
PARTY*		B	Std. Error	Wald	df	Sig.
ND	Intercept	-.331	,465	,507	1	,476
	[ideological=1]	,363	,200	3,294	1	,070
	[ideological=2]	0 <sup>b</sup>	.	.	0	.
	[economic=1]	-,327	,205	2,557	1	,110
	[economic=2]	0 <sup>b</sup>	.	.	0	.
	[salvage=1]	1,110	,237	21,937	1	<.001
	[salvage=2]	0 <sup>b</sup>	.	.	0	.
	[obligation=1]	,730	,282	6,716	1	,010
	[obligation=2]	0 <sup>b</sup>	.	.	0	.
	[antimemorandum=1]	-2,601	,230	128,124	1	<.001
	[antimemorandum=2]	0 <sup>b</sup>	.	.	0	.
	[versus-system=1]	-1,395	,220	40,184	1	<.001
	[versus-system=2]	0 <sup>b</sup>	.	.	0	.
	[important=1]	1,381	,452	9,327	1	,002
	[important=2]	1,140	,601	3,600	1	,058
	[important=3]	,643	,420	2,349	1	,125
	[important=4]	,244	,372	,430	1	,512
	[important=5]	0 <sup>b</sup>	.	.	0	.
	[CAT_AGE=1]	-,324	,282	1,324	1	,250
	[CAT_AGE=2]	-,421	,217	3,754	1	,053
	[CAT_AGE=3]	0 <sup>b</sup>	.	.	0	.
	[SEX=1] * [EDUC=1]	,842	,429	3,858	1	,050
	[SEX=1] * [EDUC=2]	-,122	,298	,168	1	,682
	[SEX=1] * [EDUC=3]	,318	,267	1,419	1	,233
	[SEX=2] * [EDUC=1]	,604	,429	1,983	1	,159
	[SEX=2] * [EDUC=2]	,954	,303	9,934	1	,002
	[SEX=2] * [EDUC=3]	0 <sup>b</sup>	.	.	0	.

#### Ερμηνεία του Πίνακα 5.46

- I. Με βάση το Model 9 ο παράγων των ιδεολογικών λόγων δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,07 = 7\% > 5\%$ ). Παραπέρα αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν το Κόμμα που ψηφίζουν για ιδεολογικούς λόγους, έχουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- II. Με βάση το Model 9 ο παράγων των οικονομικών λόγων δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,11 = 11\% > 5\%$ ). Παραπέρα αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για οικονομικούς λόγους, έχουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- III. Αντίθετα, με βάση και πάλι το Model 9, ο παράγων της σωτηρίας της πατρίδας, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. < 0,001$ ), και

μάλιστα περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για λόγους σωτηρίας της, έχουν τριπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(1,11) = 3,036$ ].

- IV. Ακόμη, με βάση και πάλι το Model 9, ο παράγων των προσωπικών υποχρεώσεων, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. < 0,01), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για λόγους προσωπικών υποχρεώσεων, συγκρινόμενοι με εκείνους που δεν ψηφίζουν για λόγους προσωπικών υποχρεώσεων, έχουν διπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(0,73) = 2,076$ ].
- V. Επίσης, με βάση και πάλι το Model 9, ο παράγων της στάσης απέναντι στα μνημόνια, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig.<0,001), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν αντιμνημονιακά, συγκρινόμενοι με εκείνους που δεν ψηφίζουν αντιμνημονιακά, έχουν κατά 92,6% μειωμένες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $1-\exp(-2,601) = 1-0,074 = 0,926 = 92,6\%$ ].
- VI. Επίσης, με βάση και πάλι το Model 9, ο παράγων της στάσης απέναντι στο ‘σύστημα’, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig.<0,001), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για αντισυστημικούς λόγους, έχουν κατά 75,2% μειωμένες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $1-\exp(-1,395) = 1-0,248= 0,752 = 75,2\%$ ].
- VII. Με βάση και πάλι το Model 9, ο παράγων του πρώτου επιπέδου σπουδαιότητας<sup>35</sup> (importance=1), φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig.<0,002), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για τον Αρχηγό, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν σχεδόν τετραπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(1,381) = 3,979$ ].

<sup>35</sup> Να θυμηθούμε ότι είναι ακριβώς εκείνοι που ψήφισαν το Κόμμα που ψήφισαν, το ψήφισαν χάριν του Αρχηγού

- VIII. Με βάση και πάλι το Model 9, ο παράγων του δεύτερου επιπέδου σπουδαιότητας (importance=2), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,058 = 5,8% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για τους Υποψήφιους του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(1,14) = 3,126$ ].
- IX. Με βάση και πάλι το Model 9, ο παράγων του τρίτου επιπέδου σπουδαιότητας (importance=3), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,125 = 12,5% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για την κοσμοθεωρία του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,643) = 1,902$ ].
- X. Με βάση και πάλι το Model 9, ο παράγων του τέταρτου επιπέδου σπουδαιότητας (importance=4), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,512 = 51,2% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για το Πρόγραμμα του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,244) = 1,276$ ].
- XI. Με βάση και πάλι το Model 9, ο παράγων του πρώτου επιπέδου ηλικιακών κατηγοριών (CAT\_AGE = 1 = 17-34 ετών), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,25 = 25% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που είναι από 17-34 ετών, συγκρινόμενοι με εκείνους που είναι 55 και άνω, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(-0,324) = 0,723$ ].
- XII. Με βάση και πάλι το Model 9, ο παράγων του δεύτερου επιπέδου ηλικιακών κατηγοριών (CAT\_AGE = 2 = 35-54 ετών), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,053 = 5,3% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που είναι από 17-24 ετών, συγκρινόμενοι με

εκείνους που είναι 55 και άνω, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(-0,421) = 0,656$ ].

- XIII. Με βάση και πάλι το Model 9, η υποομάδα των ανδρών στοιχειώδους εκπαιδεύσεως [SEX=1]\*[EDUC=1], μόλις που δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,050 = 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι στοιχειώδους εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,842) = 2,321$ ].
- XIV. Με βάση και πάλι το Model 9, η υποομάδα των ανδρών δευτεροβάθμιας εκπαιδεύσεως [SEX=1]\*[EDUC=2], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,682 = 68,2 > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- XV. Με βάση και πάλι το Model 9, η υποομάδα των ανδρών πανεπιστημιακής εκπαιδεύσεως [SEX=1]\*[EDUC=3], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,233 = 23,3%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι πανεπιστημιακής εκπαίδευσης, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- XVI. Με βάση και πάλι το Model 9, η υποομάδα των γυναικών πρωτοβάθμιας δευτεροβάθμιας εκπαιδεύσεως [SEX=2]\*[EDUC=1], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,159 = 15,9% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- XVII. Με βάση τέλος το Model 9, η υποομάδα των γυναικών δευτεροβάθμιας εκπαιδεύσεως [SEX=2]\*[EDUC=2], φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. < 0,002), και μάλιστα, περαιτέρω αυτό σημαίνει

ότι, οι γυναίκες ψηφοφόροι δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενες με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν 2,596 φορές μεγαλύτερες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(0,954) = 2,596$ ].

**Πίνακας 5.51.** Πίνακας Ταξινόμησης (Classification Table) για το Model 9.

Observed	ND	Classification					XRYSH	AYGH	ANEL	ALLO	Percent Correct
		SYRIZA	PASOK	POTAMI	KKE						
ND	248	53	0	2	0		2	0	1		81,0%
SYRIZA	75	324	0	1	1		4	0	2		79,6%
PASOK	33	16	0	0	0		0	0	0		0,0%
POTAMI	30	32	0	0	0		0	0	5		0,0%
KKE	15	43	0	0	1		0	0	0		1,7%
XRYSH AYGH	13	33	0	1	0		10	0	0		17,5%
ANEL	7	39	0	0	0		3	0	1		0,0%
ALLO	48	58	0	0	3		3	0	3		2,6%
Overall Percentage	42,3%	53,9%	0,0%	0,4%	0,5%		2,0%	0,0%	1,1%		52,8%

Ο Πίνακας 5.51 προφανώς κάνει λόγο για την ικανότητα του Model 9 να ταξινομεί σωστά:

- I. Τους οπαδούς της ΝΔ σε ποσοστό 81%.
- II. Τους οπαδούς του ΣΥΡΙΖΑ σε ποσοστό 79,6%.
- III. Τους οπαδούς του ΠΑΣΟΚ σε ποσοστό 0 %.
- IV. Τους οπαδούς του ΠΟΤΑΜΙΟΥ σε ποσοστό 0%.
- V. Τους οπαδούς KKE σε ποσοστό 1,7 %.
- VI. Τους οπαδούς της ΧΡΥΣΗΣ ΑΥΓΗΣ σε ποσοστό 17,5 %.
- VII. Τους οπαδούς των ΑΝΕΞ. ΕΛΛΗΝΩΝ σε ποσοστό 0 %.

Ετσι η Καθολική Ικανότητα Πρόβλεψης (Overall Predictive Power), για το Model 9 κυμαίνεται γύρω στο 52,8 %.

Είδαμε επίσης ότι για το Model 7 η Power κυμαίνεται γύρω στο 52,4% και για το Model 8 γύρω στο 52,5%.

Στο σημείο αυτό παραθέτουμε και αντιπαραβάλουμε τα Αποτελέσματα των Εθνικών Εκλογών της 25-1-15 και τη δική μας πρόβλεψη, με βάση το Model 9.

**Πίνακας 5.52.** Αποτελέσματα Εθνικών Εκλογών 25-1-15 (Στοιχεία ΕΛΣΤΑΤ), συγκριτικά με την πρόβλεψη του Model 9.

ΠΟΛΙΤΙΚΟ ΚΟΜΜΑ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΚΛΟΓΩΝ 25-1-15	Η ΔΙΚΗ ΜΑΣ ΠΡΟΒΛΕΨΗ (SPSS MODEL 8)	ΑΠΟΚΛΙΣΗ
ΝΕΑ ΔΗΜΟΚΡΑΤΙΑ	27,81 %	27,50 %	-0,31 %
ΣΥΡΙΖΑ	36,4 %	36,6 %	0,2 %
ΠΑΣΟΚ	4,68 %	4,40 %	-0,28 %
ΠΟΤΑΜΙ	6,05 %	6 %	-0,05 %
ΚΚΕ	5,47 %	5,40 %	-0,07 %
ΧΡΥΣΗ ΑΥΓΗ	6,28 %	5,2 %	-1,08 %
ΑΝΕΞ. ΕΛΛΗΝΕΣ	4,75 %	4,50 %	-0,25 %
ΥΠΟΛΟΙΠΑ ΚΟΜΜΑΤΑ	13,3 %	10,4 %	-2,9 %

### 5.11.2. Παρατηρήσεις για το Model 9.

Παρατηρούμε τα εξής:

- I. Προφανώς είναι και αυτό ένα μοντέλο που μας φέρνει κοντά στην πραγματικότητα. Είναι ένα μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Ωστόσο, το Model 9 μας πάει μακρύτερα. Εχει τη δυνατότητα να περάσει σε λεπτότερες διαμερίσεις. Για παράδειγμα, μπορεί να διερευνήσει τη συμπεριφορά των κοινωνικών υποομάδων. η υποομάδα των γυναικών δευτεροβάθμιας εκπαίδευσεως [SEX=2]\*[EDUC=2], φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. < 0,002), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι γυναίκες ψηφοφόροι δευτεροβάθμιας εκπαίδευσεως, συγκρινόμενες με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν 2,596 φορές μεγαλύτερες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(0,954) = 2,596$ ].

- II. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. Πίνακα 5.47 Περιθώριες Συχνότητες). Εξαίρεση και πάλι η ΧΡΥΣΗ ΑΥΓΗ, της οποίας η απόκλιση είναι η μεγαλύτερη αναφορικά με τα υπόλοιπα Κόμματα (-1,08 %), και για λόγους που εξηγήσαμε παραπάνω.
- III. Αναμφίβολα το Model 9 διαθέτει και εκείνο, όπως και τα προηγούμενα (Model 7 και Model 8) το χαρακτηριστικό της Ερμηνευσιμότητας. (Βλ. Πίνακα 5.50).
- IV. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα.(βλ. Πίνακα 5.48).
- V. Για το Model 9 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερδιασκόρπιση (over-dispersion), όπως δείχνει ο ίδιος Πίνακας 5.48.
- VI. Το Model 9 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης (Long And Freese, 2014). Δηλ. δεν υπάρχει επίπεδο σε καμία από τις ανεξάρτητες μεταβλητές που να σχετίζεται με μόνο ένα επίπεδο της εξαρτημένης μεταβλητής PARTY του πολιτικού Κόμματος.
- VII. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές, και δεν υπάρχει περίπτωση που να είναι όλοι οι σημαντικοί συντελεστές μηδέν, αναφορικά μια κάποια σημαντική μεταβλητή. Κι ακόμη η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. Πίνακα 5.49).
- VIII. Η ικανότητα ταξινόμησης του Model 9, κινείται σε ικανοποιητικά/αποδεκτά επίπεδα (Prediction Power = 52,8%), κάτι παραπάνω του Model 7 που είναι 52,4%, και του Model 8 (Prediction Power = 52,5%). Το Model 9 δεν περιλαμβάνει στις δομές του καμία συνεχή μεταβλητή.
- IX. Κατά την κατασκευή του Model 9 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο συνθετότητας, δηλ. με ύπαρξη αλληλεπίδρασης (interaction 2ης τάξεως, η οποία ωστόσο, δεν δυσκόλεψε την Ερμηνεία. Αντίθετα, την ενίσχυσε. (βλ. Πίνακα 5.50).

## **ΤΕΛΙΚΟ ΣΥΜΠΕΡΑΣΜΑ για το MODEL 9.**

Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 9 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Προφανώς, η Ερμηνεία αλλάζει εξ αιτίας του ότι υπάρχει αλληλεπίδραση μεταξύ των μεταβλητών του μορφωτικού επιπέδου EDUC, και του φύλου SEX. Έτσι όμως, έγινε δυνατή η εξέταση της συμπεριφοράς των κοινωνικών υποομάδων που περιλήφθηκαν σε αυτό. Και πάλι έχουμε μπροστά μας την επίμονη ερώτηση<sup>36</sup> αν είναι το βέλτιστο ή αλλιώς ‘Το άριστο’, (‘The Best’) μοντέλο. Η απάντηση είναι και πάλι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει, όπως ξαναείπαμε, η αυτοκρατορία των Γραμμικών Σχέσεων.

---

<sup>36</sup> «Οι Ερωτήσεις είναι πάντα οι ίδιες, οι Απαντήσεις αλλάζουν», έλεγε ο Αλβέρτος Αϊνστάιν...



## **Κεφάλαιο 6: Συμπεράσματα**

Η ανάπτυξη των 9 Προβλεπτικών Μοντέλων που προηγήθηκε συσσώρευσε πολλές και σημαντικές πληροφορίες γύρω από την ΠΡΟΒΛΕΨΗ ΤΟΥ ΕΚΛΟΓΙΚΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ:

- I. Η παρούσα Έρευνα, παρήγαγε μεγάλο όγκο πληροφοριών και έδειξε με σαφήνεια ότι, είναι δυνατή η πρόβλεψη του εκλογικού αποτελέσματος των εκλογών της 25-1-15, με τουλάχιστον 3 Στατιστικά Μοντέλα Πρόβλεψης (Model 7, Model 8, Model 9), τα οποία χαρακτηρίζονται από σχετικά απλή δομή, με 10 περίπου προβλέπουσες μεταβλητές (predictors), υψηλή εγγύτητα των προβλεπόμενων ποσοστών των συμμετεχόντων Κομμάτων με τα ποσοστά που έλαβαν στην πραγματικότητα, στη συγκεκριμένη εκλογική αναμέτρηση, και εντυπωσιακά μικρό στατιστικό σφάλμα, όπως δείχνουν οι Πίνακες 5.40, 5.45, και 5.52.
- II. Οι προβλέπουσες μεταβλητές αφορούν Ιδεολογικούς λόγους, Οικονομικούς λόγους (ανεργία, ακρίβεια, οικονομική ανέχεια κ.τ.λ), λόγους Σωτηρίας της Πατρίδας, λόγους Προσωπικών Υποχρεώσεων, λόγους Αντιμνημονιακής Συμπεριφοράς, λόγους Αντισυστημικής Συμπεριφοράς, λόγους επιδοκιμασίας Αρχηγού του Κόμματος, λόγοι επιδοκιμασίας Προγράμματος του Κόμματος, λόγους προτίμησης Κοσμοθεωρίας του Κόμματος, λόγους διαφορετικής εκλογικής συμπεριφοράς των δύο Φύλων, λόγοι διαφορετικής συμπεριφορές των διαφόρων Μορφωτικών Επιπέδων, και υπήρξαν σημαντικοί παράγοντες με συνεισφορά στα διάφορα Προβλεπτικά Μοντέλα.
- III. Η ανάλυση του Μοντέλου 1 (Model 1), έδειξε με σαφήνεια ότι δεν είναι δυνατή η πρόβλεψη του εκλογικού αποτελέσματος, με μια και μόνη προβλέψιμη μεταβλητή, εκείνη του Φύλου (SEX), για λόγους που παρατίθενται (βλ. παρ. 5.3). Ωστόσο, το μοντέλο προσφέρει χρήσιμες πληροφορίες στην Ανάλυση, κι επομένως μπορεί να χαρακτηριστεί πληροφοριακό (Informative), όπως θα έλεγε η Norusis (2010).

- IV. Το Μοντέλου 2 (Model 2), ενώ είναι ένα μοντέλο χωρίς Ερμηνευσιμότητα, (βλ. παρ. 5.4) μεταφέρει μια ιδιαίτερα σημαντική πληροφορία στην παραπέρα στατιστική Ανάλυση: Αναδεικνύει την ανυπαρξία αλληλεπίδρασης δεύτερης τάξης (second-class interaction), ανάμεσα στις μεταβλητές του Φύλου (SEX) και των Ηλικιακών Κατηγοριών (CAT\_AGE).
- V. Το Μοντέλου 3 (Model 3), είναι ένα καλό μοντέλο με πολλές χρήσιμες πληροφορίες. (βλ. παρ. 5.5). Καταφέρνει να προβλέψει, με τη συμμετοχή δύο μόνον προβλεπουσών μεταβλητών, Φύλο (SEX) και ηλικιακές κατηγορίες (CAT\_AGE), κι ακόμη, μας προσφέρει την δυνατότητα να εισαχθούμε βαθύτερα στο νόημα, στο περιεχόμενο και στη χρήση των εργαλείων ελέγχου της προσαρμογής, στα Likelihood Ratio Tests. Επίσης, αναδεικνύει τη μεγάλη σημασία των στατιστικών δεικτών AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion). Είναι ένα μοντέλο με καλή προσαρμογή (fit), κι ακόμη διαθέτει το χαρακτηριστικό της Ερμηνευσιμότητας. Έχει ωστόσο μια σοβαρή αδυναμία: Χαμηλή Προβλεπτική Δύναμη (Predictive Power), όπως τεκμηριώνεται στην παρ. 5.5.
- VI. Το Μοντέλου 4 (Model 4), κομίζει μια σημαντική πληροφορία, ιδιαίτερα χρήσιμη στην οικοδόμηση προβλεπτικών μοντέλων μεγαλύτερης συνθετότητας. Συγκεκριμένα έδειξε με σαφήνεια ότι δεν υπάρχει αλληλεπίδραση ανάμεσα στους predictors EDUC και CAT\_AGE, δηλ. έδειξε ότι, interaction δεύτερης τάξης, μεταξύ μορφωτικού επιπέδου και ηλικιακών κατηγοριών, δεν υπάρχει. Η μεταβλητή αλληλεπίδρασης EDUC\*CAT\_AGE αποδείχθηκε στατιστικώς ασήμαντη δηλ. ο πληθυσμιακός παλινδρομικός συντελεστής (regression coefficient) είναι μηδέν (βλ. παρ. 5.6).
- VII. Το Μοντέλου 5 (Model 5), δεν μπορεί να γίνει αποδεκτό όχι μόνο γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής, (βλ. παρ. 5.7), -αν και αυτό θα έφτανε-, αλλά και γιατί η κεφαλαιώδους σημασίας για την Έρευνά μας μεταβλητή του φύλου, στο πλαίσιό του, δεν μπορεί εύκολα να ερμηνευθεί.

VIII. Το Μοντέλου 6 (Model 6), δεν μπορεί να γίνει αποδεκτό ως προβλεπτικό μοντέλο, όχι μόνο γιατί δεν έχουμε κανένα εχέγγυο καλής προσαρμογής, αλλά και γιατί δεν μπορεί εύκολα να ερμηνευθεί. (βλ. παρ 5.8). Ωστόσο κομίζει μια σημαντικότατη πληροφορία. Αποκλείει interaction 3ης τάξης, ανάμεσα στις μεταβλητές του Φύλου, των Μορφωτικών επιπέδων και των Ηλικιακών Κατηγοριών. Ο συντελεστής παλινδρόμησης γι' αυτήν την αλληλεπίδραση (SEX \* EDUC \* CAT\_AGE) είναι μηδενικός. Και είναι μηδενικός διότι το αντίστοιχο επίπεδο στατιστικής σημαντικότητας είναι πολύ μεγαλύτερο του 5%. Αυτό πάει παραπέρα την ανάλυσή μας. Αφού δεν έχουμε interaction 3ης τάξης, δεν θα έχουμε ούτε interaction 4ης τάξης, 5ης τάξης, 6ης τάξης κτλ. Γενικά, δεν θα έχουμε interaction ανώτερης τάξης. Αυτό πρακτικά σημαίνει ότι καθόλου δεν χρειάζεται να ελέγξουμε άλλα ιεραρχικά μοντέλα. Επομένως ο δρόμος για το ΑΡΙΣΤΟ ΜΟΝΤΕΛΟ (THE BEST MODEL), απλοποιήθηκε, έγινε πιο βατός, λιγότερο σύνθετος, έγινε απλούστερη η διαδρομή. Έχουμε να κάνουμε πλέον στο εξής με αλληλεπιδράσεις δευτέρας τάξεως και μόνον.

IX. Για το Μοντέλου 7 (Model 7) παρατηρούμε τα εξής:

- i. Πέρα από το ότι ικανοποιεί τις παραδοχές που αναφέρθηκαν στο Κεφάλαιο 4, είναι το μοντέλο που μας φέρνει πολύ κοντά στην πραγματικότητα. Με άλλα λόγια, είναι το μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. Πίνακα 5.52). Ειρήσθω εν παρόδω, και όπως πολλές φορές έχουμε επισημάνει, η ΑΠΟΚΛΙΣΗ για τη ΧΡΥΣΗ ΑΥΤΗ είναι η μεγαλύτερη αναφορικά με τα άλλα πολιτικά Κόμματα (-1,18%), διότι αρκετοί είναι εκείνοι που ψηφίζουν ΧΡΥΣΗ ΑΥΤΗ, αλλά για διάφορους λόγους δεν το λένε.
- ii. Αναμφίβολα το Model 7 διαθέτει το σοβαρότερο ίσως χαρακτηριστικό στο πεδίο της κοινωνικής έρευνας που ονομάζεται Ερμηνευσιμότητα. (Βλ. παράγραφο 5.9.5).

- iii. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα.(βλ. παρ. 5.9).
- iv. Για το Model 7 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερ-διασκόρπιση (over-dispersion), όπως με ενάργεια δείξαμε στην παράγραφο 5.9.1.
- v. Το Model 7 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης (Long And Freese, 2014).
- vi. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές και ότι η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. παρ 5.9.3).
- vii. Η ικανότητα ταξινόμησης του Model 7, κινείται σε ικανοποιητικά/ αποδεκτά επίπεδα. (βλ. παράγραφο 5.9.4 Classification table). Συνολική Προβλεπτική Δύναμη (Overall Prediction Power) = 58,4% .
- viii. Το Model 7 δεν περιλαμβάνει στις δομές του καμία συνεχή μεταβλητή, ενώ αποδείχθηκε ότι η συνεχής μεταβλητή της ηλικίας (AGE), καθίσταται χρήσιμη, και έχει συνεισφορά στο μοντέλο, μόνον όταν αυτή μετατραπεί σε κατηγορική. Κι ακόμη, αυτό θα φανεί εναργέστερα στην εξέταση του Μοντέλου 8, το οποίο περιλαμβάνει στη δομή τη μεταβλητή της Ηλικίας, ως συνεχή (covariate).
- ix. Κατά την κατασκευή του Model 7 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο χαμηλής συνθετότητας, χωρίς δηλ. την ύπαρξη αλληλεπιδράσεων (interaction 2ης και 3ης τάξεως, οι οποίες προφανώς, δυσκολεύουν, ενίστε σοβαρά την Ερμηνεία. Από την άλλη οι αποκλεισμοί των αλληλεπιδράσεων, καθόλου δεν έβλαψαν την προσαρμογή του μοντέλου στα δεδομένα όπως φάνηκε στον έλεγχο καλής προσαρμογής (βλ. παρ. 5.9.5). Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 7 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Στην ερώτηση αν είναι το βέλτιστο ή αλλιώς ‘Το άριστο’, (‘The Best’) μοντέλο , η απάντηση είναι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε

κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει η αυτοκρατορία των Γραμμικών Σχέσεων. Αναμφίβολα, η προσπάθεια ανίχνευσης του ΒΕΛΤΙΣΤΟΥ ΜΟΝΤΕΛΟΥ θα ήταν μια σπουδαία ΑΣΚΗΣΗ για τους Σπουδαστές της Κοινωνικής Στατιστικής και τους νέους Ερευνητές του πεδίου.

- x. Ολοκληρώνοντας τη συζήτηση για το Model 7 παραθέτουμε αναπόφευκτα την Ερμηνεία των παραμέτρων αυτού. (βλ. παρ. 5.9.5):
- xii. Οι ιδεολογικοί λόγοι είναι μια στατιστικώς σημαντική μεταβλητή του Model 7, ( $Sig. = 0,046 = 4,6\% <5\%$ ). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +0,396$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για ιδεολογικούς λόγους, συγκρινόμενοι με εκείνους που δεν ψήφισαν για ιδεολογικούς λόγους, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε που είναι η βάση των νεπέριων λογαρίθμων, και ο οποίος έχει τιμή στην περίπτωση αυτή ίση με 1,486 μονάδες και είναι προφανώς μεγαλύτερος της μονάδας, [ $\exp(0,396) = 1,486 > 1$ ]. Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες/μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει ιδεολογικά, η πιθανότητα να ψηφίσει ΝΔ είναι 1,486 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.
- xiii. Οι οικονομικοί λόγοι δεν είναι μια στατιστικώς σημαντική μεταβλητή του Model 7, ( $Sig. = 0,132 = 13,2\% > 5\%$ ). Δεν παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ. Αυτό παραπέρα σημαίνει ότι όσοι ψήφισαν για οικονομικούς λόγους, συγκρινόμενοι με εκείνους που δεν ψήφισαν για οικονομικούς λόγους, εξ ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ. Αυτό μας το λέει κι ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή 0,736 μονάδες, τιμή προφανώς μικρότερη του 1 [ $\exp(-0,306) = 0,736 < 1$ ].

Επιτόπιο συμπέρασμα: Ο ΣΥΡΙΖΑ δεν ανέβηκε στην εξουσία ύποπτη οικονομικής δυσπραγίας. Σε άλλους λόγους, όπως θα δούμε παρακάτω, οφείλεται η άνοδος του ΣΥΡΙΖΑ στην εξουσία.

- xiii. Η σωτηρία της πατρίδας είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7, ( $\text{Sig. } < 0,001 = 0,01\%$ ). Παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +1,057$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για τη σωτηρία της πατρίδας, συγκρινόμενοι με εκείνους που δεν ψήφισαν για τη σωτηρία της πατρίδας, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή ίση με 2,876 μονάδες και είναι προφανώς είναι μεγαλύτερος της μονάδας, [ $\exp(1,057) = 2,876 > 1$ ]. Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες/μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει για τη σωτηρία της πατρίδας, η πιθανότητα να ψηφίσει ΝΔ είναι 2,876 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.
- xiv. Οι ανειλημμένες υποχρεώσεις είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7, ( $\text{Sig. } = 0,007 = 0,07\%$ ). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει θετικό πρόσημο ( $B = +0,758$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν για λόγους ανειλημμένων υποχρεώσεων, συγκρινόμενοι με εκείνους που δεν ψήφισαν για λόγους ανειλημμένων υποχρεώσεων, περισσότερο κατευθύνονται προς την κάλπη της ΝΔ, παρά του ΣΥΡΙΖΑ. Κατά πόσο; Αυτό θα μας το πει ο εκθέτης του ε, οποίος έχει τιμή στην περίπτωση αυτή ίση με 2,134 μονάδες, και ο οποίος είναι προφανώς είναι μεγαλύτερος της μονάδας, [ $\exp(0,758) = 2,134 > 1$ ]. Μα ο εκθέτης αυτός αντιπροσωπεύει το λόγο των δυο πιθανοτήτων ψηφίσαντες/μη ψηφίσαντες, για το θέμα αυτό. Για κάποιον επομένως, που ψηφίζει για τη σωτηρία της πατρίδας, η

πιθανότητα να ψηφίσει ΝΔ είναι 2,134 φορές μεγαλύτερη παρά να ψηφίσει ΣΥΡΙΖΑ.

- xv. Τα πράγματα αλλάζουν στην περίπτωση της μεταβλητής του αντιμνημονιακού θέματος. Η στάση των πολιτών απέναντι στο μνημόνιο είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7, (Sig.<0,001=0,01%). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει αρνητικό πρόσημο ( $B=-2,613$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν αντιμνημονιακά, συγκρινόμενοι με εκείνους που δεν ψήφισαν αντιμνημονιακά, παρουσιάζουν μειωμένες πιθανότητες να κατευθύνονται προς την κάλπη της ΝΔ. Κατά πόσο; Αυτό θα μας το πει η αφαίρεση  $1-\exp(-2,613)=1-0,073=0,927=92,7\%$ . Για κάποιον επομένως, που ψηφίζει αντιμνημονιακά, η πιθανότητα να ψηφίσει ΝΔ είναι μειωμένη κατά 92,7 % μικρότερη παρά να ψηφίσει ΣΥΡΙΖΑ. Επιτόπιο συμπέρασμα: Οι ψηφοφόροι του ΣΥΡΙΖΑ, στην συντριπτική τους πλειοψηφία προέρχονται από το αντιμνημονιακό στρατόπεδο. Ας το πούμε αλλιώς: Το αντιμνημονιακό στρατόπεδο φαίνεται να υπήρξε η κύρια δεξαμενή από την οποία άντλησε ψήφους ο ΣΥΡΙΖΑ.
- xvi. Στοιχεία όμοια με το αντιμνημονιακό ζήτημα, βλέπουμε και στο αντιουστημικό ζήτημα. Η στάση των πολιτών απέναντι στο σύστημα είναι επίσης μια στατιστικώς σημαντική μεταβλητή του Model 7, (Sig.<0,001=0,01%). Παρεμβαίνει επομένως στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ με τον αντίστοιχο παλινδρομικό συντελεστή να έχει αρνητικό πρόσημο ( $B=-1,364$ ), πράγμα που σημαίνει ότι όσοι ψήφισαν αντιουστημικά, συγκρινόμενοι με εκείνους που δεν ψήφισαν αντιουστημικά, παρουσιάζουν μειωμένες πιθανότητες να κατευθύνονται προς την κάλπη της ΝΔ. Κατά πόσο; Αυτό θα μας το πει η αφαίρεση  $1-\exp(-1,364)=1-0,256=0,744=74,4\%$ . Για κάποιον επομένως, που ψηφίζει αντιουστημικά, η πιθανότητα να ψηφίσει ΝΔ είναι μειωμένη κατά 74,4% μικρότερη παρά να ψηφίσει ΣΥΡΙΖΑ. Επιτόπιο συμπέρασμα: Οι ψηφοφόροι του ΣΥΡΙΖΑ, στην μεγάλη

τους πλειοψηφία προέρχονται από το αντισυστημικό στρατόπεδο. Ας το πούμε αλλιώς: Το αντισυστημικό στρατόπεδο φαίνεται να υπήρξε επίσης μια μεγάλη δεξαμενή από την οποία άντλησε ψήφους ο ΣΥΡΙΖΑ.

xvii. Για το θέμα της σπουδαιότητας<sup>37</sup> ο Πίνακας 5.34 μας παρέχει 4 συμπεράσματα:

Πρώτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη του Αρχηγού, συγκρινόμενοι με εκείνους που το ψήφισαν για «Άλλο» λόγο, εμφανίζουν αυξημένες πιθανότητες, για την ακρίβεια 1,477 φορές να κατευθύνονται προς την κάλπη της ΝΔ παρά προς την κάλπη του ΣΥΡΙΖΑ [Sig.=0,005=0,05 % < 5 % και επίσης  $\exp(1,27) = 1,477$ ].

Δεύτερο: Εκείνοι που ψήφισαν το Κόμμα για χάρη των Υποψηφίων, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig.=0,077=7,7%>5 % και επίσης  $\exp(1,056) = 2,876$ ].

Τρίτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη της Κοσμοθεωρίας του Κόμματος, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig.=0,205=20,5 % >5 % και επίσης  $\exp(0,528) = 1,695$ ].

Τέταρτο: Εκείνοι που ψήφισαν το Κόμμα για χάρη του Προγράμματος και των Θέσεων του Κόμματος, συγκρινόμενοι με εκείνους που ψήφισαν για «Άλλο» λόγο, εξ' ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ [Sig. = 0,676 = 20,5% > 5% και επίσης  $\exp(0,154) = 1,167$ ].

xviii. Το Φύλο δεν είναι μια στατιστικώς σημαντική μεταβλητή για το Model 7 (Sig.=0,292=29,2%>5%). Δεν παρεμβαίνει επομένως, στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ. Αυτό παραπέρα σημαίνει ότι οι άνδρες, συγκρινόμενοι με τις

<sup>37</sup> Υπενθυμίζεται εδώ η σχετική Ερώτηση Q2 του Ερωτηματολογίου “Έρευνας: Ποιος ήταν ο σημαντικότερος λόγος στην επιλογή της ψήφου σας; 1= ο αρχηγός του Κόμματος; 2=οι υποψήφιοι; 3=η γενική κοσμοθεωρία/ ιδεολογία του Κόμματος; 4=το Πρόγραμμα και οι Θέσεις; 5=Άλλος λόγος ποιος .... ;

γυναίκες, εξ ίσου κατευθύνονται προς την κάλπη της ΝΔ και του ΣΥΡΙΖΑ.

Επιτόπιο συμπέρασμα: Δεν υπάρχει επαρκής ερευνητική μαρτυρία (evidence) ότι τα δυο φύλλα είχαν διαφορετική στάση απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ.

xix. Για το Μορφωτικό επίπεδο ο Πίνακας 5.34 μας παρέχει 2 συμπεράσματα:

Πρώτο: Οι απόφοιτοι της Στοιχειώδους Εκπαίδευσεως συγκρινόμενοι με τους αποφοίτους του Πανεπιστημίου δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ ( $Sig.=0,098=9,8\% > 5\%$ ).

Δεύτερο: Οι απόφοιτοι της Μέσης Εκπαίδευσεως συγκρινόμενοι με τους αποφοίτους του Πανεπιστημίου δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ ( $Sig. = 0,262 = 26,2\% > 5\%$ ).

xx. Για τις Ηλικιακές Κατηγορίες ο Πίνακας 5.34 μας παρέχει 2 συμπεράσματα:

Πρώτο: Η ηλικιακή κατηγορία των 17-34 συγκρινόμενη με εκείνη των 55 και άνω δεν εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ ( $Sig. = 0,149 = 14,9\% > 5\%$ ).

Δεύτερο: Η ηλικιακή κατηγορία των 35-54 συγκρινόμενη με εκείνη των 55 και άνω εμφανίζουν στατιστικώς σημαντική διαφορά, αναφορικά με την εκλογική τους συμπεριφορά απέναντι στα δυο μεγάλα Κόμματα ΝΔ και ΣΥΡΙΖΑ ( $Sig. = 0,042 = 4,2\% < 5\%$ ). Ποιο συγκεκριμένα, η ηλικιακή κατηγορία των 35-54 εμφανίζει μειωμένες πιθανότητες κατά 35,4% ( $1-0,646 = 0,354 = 35,4\%$ ), να κατευθύνεται προς την κάλπη της ΝΔ, συγκριτικά με την κάλπη του ΣΥΡΙΖΑ.

Σημαντική Σημείωση: Αν θέσουμε ως κατηγορία αναφοράς την ηλικιακή κατηγορία των 35-54, θα εντοπίσουμε ενδεχόμενα και άλλες διαφορές στην εκλογική συμπεριφορά των υπόλοιπων ηλικιακών κατηγοριών.

X. Για το Model 8 παρατηρούμε τα εξής:

- i. Είναι και αυτό ένα μοντέλο που μας φέρνει κοντά στην πραγματικότητα. Με άλλα λόγια και αυτό είναι ένα μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. παρ 5.10.2 Περιθώριες Συχνότητες). Εξαίρεση και πάλι η ΧΡΥΣΗ ΑΥΓΗ, της οποίας η απόκλιση είναι η μεγαλύτερη αναφορικά με τα υπόλοιπα Κόμματα (-1,08%), και για λόγους που εξηγήσαμε παραπάνω.
- ii. Αναμφίβολα το Model 8 διαθέτει και εκείνο, όπως και το προηγούμενο (Model 7) το χαρακτηριστικό της Ερμηνευσιμότητας. (Βλ. παρ 5.10.1).
- iii. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα.(βλ. παρ 5.10.1).
- iv. Για το Model 8 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερ-διασκόρπιση (over-dispersion), όπως με ενάργεια δείξαμε στην παράγραφο 5.10.1.
- v. Το Model 8 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης (Long And Freese, 2014). Δηλ. δεν υπάρχει επίπεδο σε καμία από τις ανεξάρτητες μεταβλητές που να σχετίζεται με μόνο ένα επίπεδο της εξαρτημένης.
- vi. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές και ότι η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. παρ 5.10.1).
- vii. Η ικανότητα ταξινόμησης του Model 8, κινείται σε ικανοποιητικά/ αποδεκτά επίπεδα (Prediction Power = 52,5%), περίπου ίδια με του Model 7 που είναι 52,4%. Το Model 8 περιλαμβάνει στις δομές του μία συνεχή μεταβλητή, της ηλικίας (AGE), που φαίνεται να είναι στατιστικώς ασήμαντη και χωρίς συνεισφορά στο μοντέλο (βλ. στον Πίνακα 5.33 sign.=0,053=5,3% > 5%).

viii. Κατά την κατασκευή του Model 8 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο χαμηλής συνθετότητας, χωρίς δηλ. την ύπαρξη αλληλεπιδράσεων (interaction 2ης και 3ης τάξεως, οι οποίες προφανώς, δυσκολεύουν, ενίστε, σοβαρά την Ερμηνεία. Από την άλλη οι αποκλεισμοί των αλληλεπιδράσεων, καθόλου δεν έβλαψαν την προσαρμογή του μοντέλου στα δεδομένα όπως φάνηκε στον έλεγχο καλής προσαρμογής.

ix. ΤΕΛΙΚΟ ΣΥΜΠΕΡΑΣΜΑ για το MODEL 8.

Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 8 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Προφανώς η Ερμηνεία κατά τόπους αλλάζει. Για παράδειγμα, είδαμε ότι η Ηλικία όταν εισάγεται στο μοντέλο ως κατηγορική μεταβλητή, είναι στατιστικώς σημαντική, όταν εισάγεται ως συνεχής, δεν είναι! Στην επίμονη ερώτηση αν είναι το βέλτιστο ή αλλιώς ‘Το άριστο’, (‘The Best’) μοντέλο, η απάντηση είναι και πάλι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει η αυτοκρατορία των Γραμμικών Σχέσεων. Αναμφίβολα, η προσπάθεια ανίχνευσης του ΒΕΛΤΙΣΤΟΥ ΜΟΝΤΕΛΟΥ θα ήταν μια σπουδαία ΑΣΚΗΣΗ για τους Σπουδαστές και τους νέους Ερευνητές.

XI. Για το Model 9 παρατηρούμε τα εξής:

i. Προφανώς είναι και αυτό ένα μοντέλο που μας φέρνει κοντά στην πραγματικότητα. Είναι ένα μοντέλο που καταφέρνει να αναπαραγάγει σε υψηλό βαθμό την πολιτική πραγματικότητα. Ωστόσο, το MODEL9 μας πάει μακρύτερα. Έχει τη δυνατότητα να περάσει σε λεπτότερες διαμερίσεις. Για παράδειγμα, μπορεί να διερευνήσει τη συμπεριφορά των κοινωνικών υποομάδων. η υποομάδα των γυναικών δευτεροβάθμιας εκπαίδευσεως [SEX=2]\*[EDUC=2], φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. < 0,002), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι γυναίκες ψηφοφόροι δευτεροβάθμιας εκπαίδευσεως, συγκρινόμενες με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν 2,596 φορές

μεγαλύτερες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ  
[ $\exp(0,954)=2,596$ ].

- ii. Πράγματι, με ακρίβεια μικρότερη των 5 δεκάτων καταγράφει τις αποκλίσεις από το πραγματικό ποσοστό που έλαβαν τα πολιτικά Κόμματα στις εκλογές της 25-1-15. (βλ. Πίνακα 5.52 Περιθώριες Συχνότητες). Εξαίρεση και πάλι η ΧΡΥΣΗ ΑΥΓΗ, της οποίας η απόκλιση είναι η μεγαλύτερη αναφορικά με τα υπόλοιπα Κόμματα (-1,08%), και για λόγους που εξηγήσαμε παραπάνω.
- iii. Αναμφίβολα το Model 9 διαθέτει και εκείνο, όπως και τα προηγούμενα (Model 7 και Model 8) το χαρακτηριστικό της Ερμηνευσιμότητας. (Βλ. παρ 5.11.1).
- iv. Είναι ένα μοντέλο με καλή προσαρμογή (Goodness-of-fit) στα εμπειρικά δεδομένα.(βλ. παρ 5.11.1).
- v. Για το Model 9 δεν υπάρχει επαρκής ερευνητική μαρτυρία ότι παρουσιάζει υπερ-διασκόρπιση (over-dispersion).
- vi. Το Model 9 κατά την εφαρμογή του δεν συναντά το φαινόμενο της τέλειας πρόβλεψης (Long And Freese, 2014). Δηλ. δεν υπάρχει επίπεδο σε καμία από τις ανεξάρτητες μεταβλητές που να σχετίζεται με μόνο ένα επίπεδο της εξαρτημένης μεταβλητής PARTY του πολιτικού Κόμματος.
- vii. Τα Likelihood Ratio Tests τεκμηριώνουν με σαφή τρόπο ότι όλες οι ανεξάρτητες είναι σημαντικές, και δεν υπάρχει περίπτωση που να είναι όλοι οι σημαντικοί συντελεστές μηδέν, αναφορικά μια κάποια σημαντική μεταβλητή. Κι ακόμη η απουσία μιας αφήνει πίσω της ένα σημαντικό μοντέλο. (βλ. παρ. 5.11.1).
- viii. Η ικανότητα ταξινόμησης του Model 9, κινείται σε ικανοποιητικά/ αποδεκτά επίπεδα (Prediction Power = 52,8%), κάτι παραπάνω του Model 7 που είναι 52,4%, και του Model 8 (Prediction Power = 52,5%). Το Model 9 δεν περιλαμβάνει στις δομές του καμία συνεχή μεταβλητή.
- ix. Κατά την κατασκευή του Model 9 ελήφθησαν πρόνοιες για να έχουμε ένα μοντέλο συνθετότητας, δηλ. με ύπαρξη αλληλεπίδρασης (interaction 2ης

τάξεως, η οποία ωστόσο, δεν δυσκόλεψε την Ερμηνεία. Αντίθετα, την ενίσχυσε. (βλ. παρ 5.11.1).

x. ΤΕΛΙΚΟ ΣΥΜΠΕΡΑΣΜΑ για το MODEL 9.

Με βάση όλα τα παραπάνω, χωρίς ιδιαίτερη δυσκολία, το Model 9 μπορεί, προφανώς, να χαρακτηρισθεί ως ένα πολύ καλό προβλεπτικό μοντέλο. Προφανώς, η Ερμηνεία αλλάζει εξ' αιτίας του ότι υπάρχει αλληλεπίδραση μεταξύ των μεταβλητών του μορφωτικού επιπέδου EDUC, και του φύλου SEX. Έτσι όμως, έγινε δυνατή η εξέταση της συμπεριφοράς των κοινωνικών υποομάδων που περιλήφθηκαν σε αυτό. Και πάλι έχουμε μπροστά μας την επίμονη ερώτηση αν είναι το βέλτιστο ή αλλιώς 'Το άριστο', ('The Best') μοντέλο. Η απάντηση είναι και πάλι μάλλον απλή: Ναι, μέχρι που να βρεθεί κάποιο καλύτερό του...!! Σε κάθε περίπτωση για την υπόθεση αυτή θα αποφασίσει, όπως ξαναείπαμε, η αυτοκρατορία των Γραμμικών Σχέσεων.

xii. Χάριν πληρότητας της Έρευνας, παραθέτουμε την οπωσδήποτε ενδιαφέρουσα Ερμηνεία του Model 9:

- ✓ Με βάση το Model 9 ο παράγων των ιδεολογικών λόγων δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,07 = 7\% > 5\%$ ). Παραπέρα αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν το Κόμμα που ψηφίζουν για ιδεολογικούς λόγους, έχουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- ✓ Με βάση το Model 9 ο παράγων των οικονομικών λόγων δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,11 = 11\% > 5\%$ ). Παραπέρα αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για οικονομικούς λόγους, έχουν ίσες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- ✓ Αντίθετα, με βάση και πάλι το Model 9, ο παράγων της σωτηρίας της πατρίδας, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. < 0,001$ ), και μάλιστα περαιτέρω αυτό σημαίνει ότι οι

ψηφοφόροι που ψηφίζουν για λόγους σωτηρίας της , έχουν τριπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(1,11) = 3,036$ ].

- ✓ Ακόμη, με βάση και πάλι το Model 9, ο παράγων των προσωπικών υποχρεώσεων, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $\text{Sig.} < 0,01$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για λόγους προσωπικών υποχρεώσεων, συγκρινόμενοι με εκείνους που δεν ψηφίζουν για λόγους προσωπικών υποχρεώσεων, έχουν διπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(0,73) = 2,076$ ].
- ✓ Επίσης, με βάση και πάλι το Model 9, ο παράγων της στάσης απέναντι στα μνημόνια, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $\text{Sig.} < 0,001$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν αντιμνημονιακά, συγκρινόμενοι με εκείνους που δεν ψηφίζουν αντιμνημονιακά, έχουν κατά 92,6% μειωμένες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $1-\exp(-2,601)=1-0,074=0,926=92,6\%$ ].
- ✓ Επίσης, με βάση και πάλι το Model 9, ο παράγων της στάσης απέναντι στο ‘σύστημα’, φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $\text{Sig.} < 0,001$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για αντισυστημικούς λόγους, έχουν κατά 75,2% μειωμένες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $1-\exp(-1,395)=1-0,248=0,752=75,2\%$ ].
- ✓ Με βάση και πάλι το Model 9, ο παράγων του πρώτου επιπέδου σπουδαιότητας (importance=1), φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $\text{Sig.} < 0,002$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι οι ψηφοφόροι που ψηφίζουν για τον Αρχηγό, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν σχεδόν τετραπλάσιες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(1,381) = 3,979$ ].

- ✓ Με βάση και πάλι το Model 9, ο παράγων του δεύτερου επιπέδου σπουδαιότητας (importance=2), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,058 = 5,8% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για τους Υποψήφιους του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(1,14) = 3,126$ ].
- ✓ Με βάση και πάλι το Model 9, ο παράγων του τρίτου επιπέδου σπουδαιότητας (importance=3), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,125 = 12,5% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για την κοσμοθεωρία του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,643) = 1,902$ ].
- ✓ Με βάση και πάλι το Model 9, ο παράγων του τέταρτου επιπέδου σπουδαιότητας (importance=4), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,512 = 51,2% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που ψηφίζουν για το Πρόγραμμα του Κόμματος, συγκρινόμενοι με εκείνους που ψηφίζουν για κάποιο ‘άλλο λόγο’, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,244) = 1,276$ ].
- ✓ Με βάση και πάλι το Model 9, ο παράγων του πρώτου επιπέδου ηλικιακών κατηγοριών (CAT\_AGE = 1 = 17-34 ετών), δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. = 0,25 = 25% > 5%), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που είναι από 17-34 ετών, συγκρινόμενοι με εκείνους που είναι 55 και άνω, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(-0,324) = 0,723$ ].
- ✓ Με βάση και πάλι το Model 9, ο παράγων του δεύτερου επιπέδου ηλικιακών κατηγοριών (CAT\_AGE = 2 = 35-54 ετών), δεν φαίνεται να

παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,053 = 5,3\% > 5\%$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι ψηφοφόροι που είναι από 17-24 ετών, συγκρινόμενοι με εκείνους που είναι 55 και άνω, έχουν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(-0,421) = 0,656$ ].

- ✓ Με βάση και πάλι το Model 9, η υποομάδα των ανδρών στοιχειώδους εκπαιδεύσεως [SEX=1]\*[EDUC=1], μόλις που δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,050 = 5\%$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι στοιχειώδους εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ [ $\exp(0,842) = 2,321$ ].
- ✓ Με βάση και πάλι το Model 9, η υποομάδα των ανδρών δευτεροβάθμιας εκπαιδεύσεως [SEX=1]\*[EDUC=2], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,682 = 68,2 > 5\%$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- ✓ Με βάση και πάλι το Model 9, η υποομάδα των ανδρών πανεπιστημιακής εκπαιδεύσεως [SEX=1]\*[EDUC=3], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,233 = 23,3\% > 5\%$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι πανεπιστημιακής εκπαίδευσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.
- ✓ Με βάση και πάλι το Model 9, η υποομάδα των γυναικών πρωτοβάθμιας δευτεροβάθμιας εκπαιδεύσεως [SEX=2]\*[EDUC=1], δεν φαίνεται να παρεμβαίνει στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ ( $Sig. = 0,159 = 15,9\% > 5\%$ ), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι άνδρες ψηφοφόροι

δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενοι με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν σχεδόν ίδιες πιθανότητες να ψηφίζουν ΝΔ και ΣΥΡΙΖΑ.

- ✓ Με βάση τέλος το Model 9, η υποομάδα των γυναικών δευτεροβάθμιας εκπαιδεύσεως [SEX=2]\*[EDUC=2], φαίνεται να παρεμβαίνει ισχυρά στην αναμέτρηση ΝΔ-ΣΥΡΙΖΑ (Sig. < 0,002), και μάλιστα, περαιτέρω αυτό σημαίνει ότι, οι γυναίκες ψηφοφόροι δευτεροβάθμιας εκπαιδεύσεως, συγκρινόμενες με τις γυναίκες ψηφοφόρους πανεπιστημιακής εκπαίδευσης, έχουν 2,596 φορές μεγαλύτερες πιθανότητες να ψηφίζουν ΝΔ παρά ΣΥΡΙΖΑ [ $\exp(0,954) = 2,596$ ].

- XII. Θα κλείσουμε το κείμενο των Συμπερασμάτων με την επαναφορά στη συζήτηση του στρατηγικού μας στόχου. Το μέγα ζήτημα του Βέλτιστου του Άριστου Μοντέλου. Η συζήτηση για το Βέλτιστο Μοντέλο, μπορεί να τραβήξει σε μάκρος και δικαιολογημένα. Στην πορεία αυτής της Έρευνας διδαχθήκαμε, ανάμεσα στα άλλα ότι, η αναζήτηση του αληθινού, πραγματικού, βέλτιστου μοντέλου είναι κάποτε μια ιδιαίτερα βασανιστική, πλην ιδιαιτέρως ενδιαφέρουσα, διαδρομή. Μοιάζει με το δρόμο για την Ιθάκη. Γεμάτη περιπέτειες, γεμάτη γνώσεις. Αν για παράδειγμα, αφαιρέσουμε μια μεταβλητή από το μοντέλο μας, όλα μπορεί να αλλάξουν. Πάμε σε άλλο μοντέλο. Πολύ ή λίγο όμοιο, με το αρχικό. Η απουσία μιας μεταβλητής οπωσδήποτε αναδεικνύει τη σπουδαιότητά της, ή μη. Άλλα μόνο για μια στιγμή. Όσο κρατάει η αφαίρεση μιας και μόνον, της αφαιρούμενης, μεταβλητής. Αν αφαιρεθούν δυο μαζί και πάμε σε ένα νέο υπο-μοντέλο, με δυο λιγότερες μεταβλητές, πάλι όλα μπορεί να αλλάξουν. Τίποτε δεν είναι δεδομένο. Με την απουσία της μια μεταβλητή, μπορεί να κάνει σπουδαία μια άλλη μεταβλητή που δεν ήταν...!!

- XIII. Ωστόσο, και για να μη μακρηγορούμε, αφαιρώντας ή προσθέτοντας μεταβλητές, δεν είναι βέβαιο ότι θα καταλήξουμε σε καλύτερο μοντέλο. Ούτε, επίσης αυτό σημαίνει, ότι οπωσδήποτε σημειώσαμε μια κάποια πρόοδο. Κάλλιστα, μέσα από μια τέτοια διαδικασία προσθαφαιρέσεων, μπορεί να έχουμε

χειροτέρευση. Να κινούμαστε δηλ. προς την αντίθετη κατεύθυνση. Προς μοντέλο, με όχι καλή προσαρμογή. Ο Simonoff (2003), μας το λέει καθαρά: «Μερικές φορές αυτό το «καλυτερότερο», αυτό που η έρευνα χαρακτηρίζει best model, απλά ... δεν υπάρχει!». Ερευνητή, ήδη θα το κατάλαβες ότι Ιθάκες πιά δεν υπάρχουν... Ωστόσο, Ερευνητή, διόλου μην απογοητεύεσαι .. Η Ιθάκη, (η επίμονη αναζήτηση του Άριστου Μοντέλου), σου έδωσε το ωραίο (το ερευνητικό) το ταξίδι... Χωρίς αυτήν δεν θα μπορούσες...

- XIV. Όλα τα παραπάνω σημαίνουν ότι καθόλου δεν είναι μάταιο να μπούμε στο δρόμο της αναζήτησης του καλύτερου όλων (the best) μοντέλου. Διότι, μέσα από μια τέτοια διαδικασία, είναι βέβαιο, ότι ο Κοινωνικός Ερευνητής, θα εξέλθει σοφότερος...
- XV. Στο σημείο αυτό, αισθανόμαστε την ανάγκη, να προσδιορίσουμε ορθολογικά και επιστημολογικά, το «καλό» μοντέλο. Θα λέγαμε απλά ότι, καλό μοντέλο, είναι αυτό που επιτυγχάνει, στο μέτρο του δυνατού, αναπαραγωγή της πραγματικότητας. Πιο συγκεκριμένα, αυτό που επιτυγχάνει την εξισορρόπηση, τη χρυσή τομή, όπως θα λέγαμε με αρχαιοελληνικούς όρους, ανάμεσα στην προσαρμογή (fit), και στη φειδωλότητα (parsimony). Το τελευταίο, η φειδωλότητα, περιγράφει την ικανότητα του μοντέλου να συγκροτηθεί με τον ελάχιστο δυνατό αριθμό παραγόντων/μεταβλητών. Αυτοί οι παράγοντες ωστόσο, θα πρέπει να είναι σε θέση να περιγράψουν σωστά τις σχέσεις που ενυπάρχουν στα εμπειρικά δεδομένα. Παραπέρα, η διαδικασία περιγραφής που γίνεται μέσω των εκτιμητών (estimators), θα πρέπει να οδηγεί στην πρόβλεψη μελλοντικών γεγονότων (future events), βλέπε Simonoff (2003).

## Βιβλιογραφία

### Αγγλική Βιβλιογραφία

1. *Aberson, C.* (2010). Applied Power Analysis for the Behavioral Sciences. ROUTLEGDE.
2. *Abu-Bader, Soleman Hassan* (2010). Advanced and Multivariate Statistical Methods for the Social science research with a Complete SPSS Guide. LYCEUM BOOKS, INC., Chicago.
3. Acock Alan C. (2014). A Gentle Introduction to Stata, Fourth Edition.
4. Acock, A. (2012). A Gentle Introduction to Stata, Revised Third Edition. Publisher: STATA PRESS.
5. Agresti, A. (2007). An Introduction to Categorical Data Analysis, Second edition, WILEY.
6. Agresti, A. (2010). Analysis of Ordinal Categorical Data. Second edition. WILEY.
7. Agresti, A. (2018). Statistical Methods for the Social Sciences. PEARSON.
8. Aguinis, H. (2004). Regression Analysis for Categorical Moderators. THE GUILFORD PRESS.
9. Aguinis, H. (2004). Regression Analysis for Categorical Moderators (Methodology In The Social Sciences). THE GUILFORD PRESS.
10. Aitchison, J., & Silvey, S.D.(1957). The generalization of probit analysis to the case of multiple responses. BIOMETRIKA, 44, 131-140.
11. Aitchison, J., & Bennett, J. (1970). Polytomous quantal response by maximum indicant. BIOMETRIKA, 57, 253-262.
12. Aitkin, M., Francis, B., Hind, J., and Darnell, R. (2009). Statistical Modelling in R. OXFORD UNIVERSITY PRESS.

13. Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317-332.
14. Allison, P.D. (1978). Measures if inequality. AMERICAN SOCIOLOGICAL REVIEW 43: 865-880.
15. Allison, P.D. (2012b). Logistic Regression Using SAS: Theory and Application. 2<sup>nd</sup> ed. Cary, NC: SAS Institute.
16. Altman, D. (1991). Practical Statistics for Medical Research. Chapman and Hall/CRC.
17. Anderson, J. A. (1984). Regression and ordered categorical variables (with discussion). JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B 46: 1-30.
18. Anderson, J.C, & Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness of fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49,155-173.
19. Arbuckle, J. (2010). *IBM Spss Amos 19 User's Guide*. AMOS DEVELOPMENT CORPORATION.
20. Arbuckle, J. (2011). *IBM Spss Amos 20 User's Guide*. AMOS DEVELOPMENT CORPORATION.
21. Arbuckle, J. and Wothke, W. (1999). *IBM Spss Amos 4.0 User's Guide*. Chicago: SmallWaters Corporation.
22. Arum, R., & Shavit, Y. (1995). Secondary vocational education and the transition from school to work. *SOCIOLOGY OF EDUCATION*, 68, 187-304.
23. Atkinson, L. (1988). The measurement-statistics controversy: Factor analysis and subinterval data. *BULLETIN OF THE PSYCHOMETRIC SOCIETY*, 26, 361-364.
24. Aven, Terje (2011). Quantitative Risk Assessment. The Scientific Platform. CAMBRIDGE UNIVERSITY PRESS.

25. Babakus, E., Ferguson, C.E., Jr., Joreskog, K.G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *JOURNAL OF MARKETING RESEARCH*, 24, 22-228.
26. Bartholomew, D.J., Steele, F., Moustaki, I., Galbraith, J. (2002). *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Chapman and Hall/CRC.
27. Bartholomew, D.J., Steele, F., Moustaki, I., Galbraith, J. (2007). *Ανάλυση Πολυμεταβλητών Δεδομένων για Κοινωνικές Επιστήμες*, Εκδόσεις Επίκεντρο, Θεσσαλονίκη.
28. Bausell, R.B., & Li Yu-Fang (2002). *Power Analysis for Experimental Research. A practical Guide for the Biological, Medical and Social Sciences*. CAMBRIDGE.
29. Bearden, W.O., Sharma, S., and Teel, J.E. (1982). Sample size effects on chi-square and other statistics used in evaluating causal models. *Journal of Marketing Research*, 19, 425-430.
30. Bentler, P.M. (1988). Comparative fit indexes in structural equation models. *Psychological Bulletin*, 107, 238-246.
31. Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
32. Bentler, P.M. (1995). *EQS: Structural Equations Program Manual*. Encino, CA: Multivariate Software, Inc.
33. Bentler, P.M. (2005). *EQS 6 Structural equations program manual*. Encino, CA: MULTIVARIATE SOFTWARE.
34. Bentler, P.M. and Raykov, T. (2000). On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology*, 85, 125-131.

35. Bentler, P.M., and Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis covariance structures. *Psychological Bulletin*, 88,588-606.
36. Bentler, P.M., and Bonett, D.G. (1987). This week's citation classic. CURRENT CONTENTS, 19, 16.
37. Bentler, P.M., and Chou, C.P. (1987). Practical issues in structural equation modeling. *Sociological Methods & Research*, 16,78-117.
38. Bentler, P.M., and Freeman, E.H. (1983). Tests for stability in linear structural equation systems. *Psychometrika*, 48: 143-145.
39. Bentler.P.M. & Yuan, K.H.(1999). Structural Equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, 34, 181-197.
40. Bernstein, S. and Bernstein, R. (1998). Elements of Statistics II: Inferential Statistics. SCHAUM'S OUTLINES.
41. Bluncy, N.J. (2011). Introduction to Structural Education Modelling Using SPSS AND AMOS. SAGE.
42. Bollen, K.A (1989). Structural equations with latent variables. New York: Wiley.
43. Bollen, K.A., and Stine, R.A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods and Research*, 21, : 205-229.
44. Bollen, K.A., Barb, K.. H. (1981). Pearson's r and coarsely categorized measures. *AMERICAN SOCIOLOGICAL REVIEW*, 46, 232-239.
45. Bollen, K.A., Long, J.S. (1993). Introduction. In K.A. Bollen & Long (Eds.), Testing structural equation models. Beverly Hills, CA:SAGE.
46. Boomsma, A. (1983). On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality. Unpublished doctoral dissertation, University of Groningen, The Netherlands.
47. Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R. (2011). Introduction to Meta-Analysis. WILEY.

48. Bozdogan, H. (1987). Model selection and Akaike's information criteria (AIC). The general theory and its analytical extensions. *Psychometrika*, 52, 345-370.
49. Braun, W.J., Murdoch, D.J. (2007). A First Course in Statistical Programming with R. CAMBRIDGE.
50. Breckler, S.J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, 107, 260-273.
51. Breen, R., Karlson, K., and Holm, A. (2011). A reinterpretation of coefficients from logit, probit, and other non-linear probability models: Consequences for comparative sociological research.  
[\(http://papers.ssrn.com/sol3/papaers.cfm?abstract\\_id=1857431\)](http://papers.ssrn.com/sol3/papaers.cfm?abstract_id=1857431)
52. Brouwers E.P., Van Baar A.L. & Pop V.J. (2001) Does the Edinburgh Postnatal Depression Scale measure anxiety? *Journal of Psychosomatic Research* 51, 659-663.
53. Brown, H. , Prescott, R. (2006). *Applied Mixed Models in Medicine*. WILEY.
54. Brown, T. (2006). Confirmatory Factor Analysis for Applied Research. THE GUILFORD PRESS.
55. Browne, M.W. (1968). A comparison of factor analysis techniques. *Psychometrika*, 33, 267-334.
56. Browne, M.W. (1982). Covariance structures. In D.M. Hawkins (ed.), Topics in applied multivariate analysis (pp. 72-141). Cambridge: Cambridge, University Press.
57. Browne, M.W. (1984a). Asymptotically distribution- free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.

58. Browne, M.W., and Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen and J.S. Long (eds). *Testing Structural Models*. Newbury Park, CA: SAGE Publications.
59. Bryman, A. & Cramer, D. (2005). *Quantitative Data Analysis with SPSS. A Guide for Social Scientists*. Routledge.
60. Bryman, A. (2008). *Social Research Methods*. OXFORD.
61. Burnham, P., Lutz, K.G., Grant, W., and Layton-Henry, Z. (2008). *Research Methods in Politics*. PALGRAVE MACMILLAN.
62. Byrne, B. (2001). *Structural Equation Modeling With AMOS, EQS, and LISREL: Comparative Approaches to Testing for the Factorial Validity of a Measuring Instrument*. *International Journal of Testing*, 1(1), pp. 55-86.
63. Byrne, B. (2008). *Structural Equation Modeling With EQS. Basic Concepts, Applications, and Programming*. Second Edition. Routledge.
64. Byrne, B. (2010). *Structural Equation Modelling with AMOS*. Second edition, PSYCHOLOGY PRESS.
65. Byrne, B. M. (1994a). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Educational Research Journal* , 31, 645-673.
66. Cameron A.C., Trivedi P.K.(2005). *Microeconometrics: Methods and Application*. New York: Cambridge University Press.
67. Cameron A.C., Trivedi P.K.(2013). *Regression Analysis of Count Data*.
68. Carlson, M. & Mulaik, S.A.(1993). Trait ratings from descriptions of behaviour as mediated by components of meaning. *Multivariate Behavioral Research*, 28, 111-159.
69. Cattell, R. B., & Jaspers, J.A. (1967). A general palsmode for factor analytic exercises and research. *Multivariate Behavior Research Monographs*.

70. Chatterjee, S., and A. S. Hadi. (2012). *Regression Analysis by Example*. 5th ed. New York: Hoboken, NJ.
71. Chow, Shein-Chung, and Shao, Jun (2008). Sample size Calculations in clinical Research. *Second edition, Chapman & Hall/CRC*.
72. Coakes, S. J., Steed, L.G. (1999). SPSS without anguish. Wiley & Sons.
73. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements*, 20, 37-46.
74. Cohen, L., Manion, L., & Morrison, K. (2001). *Research Methods in Education*. ROUTLEDGE.
75. Comrey, A.L. (1962). The minimum residual method of factor analysis. *Psychological Reports*, 11, 15-18.
76. Comrey, A.L., & Lee, H.B. (1992). A first course in factor analysis. (2<sup>nd</sup>. Ed.). Hillsdale, NJ: Erlbaum.
77. Cox, D.R. (1972). Regression models and life tables (with discussion) *JOURNAL FOR THE ROYAL STATISTICAL SOCIETY*, 34: 187-220.
78. Cramer, J.S. (1986). *Econometric Applications of Maximum Likelihood Methods*. Cambridge: Cambridge University Press.
79. Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
80. Davidson, R., & MacKinnon, J.G. (1993) Estimation and inference in econometrics. New York: Oxford BIOMETRIKA, 57, 253-262.
81. Day, R.A. , Gastel, B. (2006). How to write and Publish a Scientific Paper. Sixth Edition, CAMBRIDGE.
82. De Leeuw, E.D., Hox, J.J. (2008). *International Handbook of Survey Methodology*. PSYCHOLOGY PRESS.

83. De Vaus, D. A. (1991). Surveys in Social Research, third edition, London: George Allen and Unwin.
84. DeCarlo, L.T. (1997). On the meaning and use of kurtosis. PSYCHOLOGICAL METHODS, 2, 292-307.
85. Denscombe, M. (2010). The Good Research Guide, for small-scale social research projects. McGraw Hill, fourth edition.
86. Diebold, F. (2007). Elements of Forecasting. Fourth edition. THOMSON SPOUTH- WESTERN.
87. DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. STRUCTURAL EQUATION MODELING, 9, 327-346.
88. Doble, J., and J. Greene (1999), "Attitudes toward crime and punishment in Vermont: Public opinion about an experiment with restorative justice", 1999 (computer file). Englewood Cliffs, N.J.: Doble Research Associates, Inc. (producer) 2000. Ann Arbor, Mich.: Inter-university Consortium for Political and Social Research (distributor), 2001.
89. Doornik, J. A., and H. Hansen. (2008). An omnibus test for univariate and multivariate normality, Oxford Bulletin of Economics and Statistics 70: 927-939.
90. Dorofeev, S., Grand, P., (2006). Statistics for Real-Life Sample Surveys. Non-Simple-Random Samples and Weighted Data. CAMBRIDGE UNIVERSITY PRESS.
91. du Toit, S., Du Toit, M.(2001), 'Interactive LISREL: User's Guide', SSI.
92. Duncan, O.D., Haller, A.O., and Portes, A. (1968). Peer influences on aspirations: A reinterpretation. AMERICAN GOURNAL OF SOCIOLOGY 74: 119-137.
93. Dunn, G. (1989). Design and analysis of reliability studies: statistical evaluation of measurement errors. Edward Arnold, London.

94. Dunteman, G. (1989). Principal Component Analysis. SAGE University paper series/ Number 07-069. On quantitative applications in the social sciences, London:SAGE.
95. Eliason, S.R. (1993). Maximum Likelihood Estimation: Logic and Practice. Newbury Park, CA: SAGE.
96. Eric Vittinghoff, David V. Glidden, Stephen C. Shiboski, Charles E. McCulloch (2012). Regression Methods in Biostatistics. Linear, Logistic, Survival, and Repeated. (Statistics for Biology and Health). Springer.
97. Everitt, B. (1999). Making sense of statistics in Psychology. OXFORD UNIVERSITY PRESS.
98. Everitt, B. (2010). Multivariable Modelling and Multivariate Analysis for the Behavioral Sciences. Chapman & Hall/ CRC PRESS.
99. Everitt, B.S (2009). Multivariable Modeling and Multivariate Analysis for the Behavioral Sciences. Chapman & Hall/CRC.
100. Everitt, B.S., Landu, S., Leese, M., Stahl, D. (2011). *Cluster Analysis*. WILEY.
101. Felson, R.B., and Bohrnstedt G.W (1979). "Are the good beautiful or the beautiful good ?" The relationship between children's perceptions of ability and perceptions physical attractiveness. SOCIAL PSYCHOLOGY QUARTERLY, 42: 386-392.
102. Field, A. (2009). DISCOVERING STATISTICS USING SPSS. SAGE.
103. Field, A. (2018). DISCOVERING STATISTICS USING IBM SPSS. SAGE.
104. Field, A., Hole, G. (2003). How to Design and Report Experiments. SAGE.
105. Finch, J.F., West, S.G., & MacKinnon, D.P. (1997). Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models. STRUCTURAL EQUATION MODELING, 4, 87-107.

106. Finney, D.J. (1971). Probit analysis. (3rd). Cambridge : Cambridge University Press.
107. Fleiss, J.L.(1981). Statistical methods for rates and proportions. 2<sup>nd</sup> ed. New York: John Wiley and Sons.
108. Fornell, C. And Y. Yi (1992). *Assumptions of the Two-Step Approach to Latent Variable Modeling*. SOCIOLOGICAL METHODS AND RESEARCH 20 (February): 291-320.
109. Forthofer, R.N., Lee, E.S., Hernandez, M. (2007). Biostatistics. Second edition ELSEVIER Inc.
110. Fox, J. (1980). Effect analysis in structural equation models. SOCIOLOGICAL METHODS AND RESEARCH 9: 3-28.
111. Frank, E., and Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems). Elsevier Inc.
112. Freedman, D. (2010). Statistical Models and Causal Inference, CAMBRIDGE UNIVERSITY PRESS.
113. French, J.V. (1951). The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs, 5.
114. Gary T. (2009). How to do Your Research Project. SAGE.
115. Gerbing, D.W., & Anderson, J.C. (1992). Monte Carlo evaluations of goodness of fit indices for structural equation models. Sociological Methods & Research, 21, 132-160.
116. Goldsceider, F.K., & DaVanzo, J. (1989). Pathways to independent living in early adulthood: Marriage, semiautonomy, and premarital residential independence. DEMOGRAPHY, 26, 597-614.
117. Gorsuch, R.L. (1983). Factor Analysis. Hillsdale, NJ: Lawrence Erlbaum.

118. Gould, William Pitblado, Jeffrey, Poi Brian (2010). Maximum Likelihood Estimation with Stata, Fourth Edition. Stata Press
119. Green, S.B., Akey, T.M., Fleming, K.K., Hershberger, S.L. & Marquis, J.G (1997). Effect of the number of scale points on chi-square fit indices in confirmatory factor analysis. *STRUCTURAL EQUATION MODELING*, 4, 108-120.
120. Guadagnoli, E., and Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103, 265-275.
121. Guilford, J.P. (1956). The structure of intellect. *Psychological Bulletin*, 53, 267-293.
122. Gurland, J.Lee, I., Dahm, P.A. (1960). Polychotomous quanttal response in biological assay. *Biometrics*, 16, 382-398.
123. Gustavii, B. (2008). How to write and Illustrate a Scientific Paper. Second Edition, CAMBRIDGE.
124. Hague, R., και Harrop, M. (2011), “Συγκριτική Πολιτική και Διακυβέρνηση”, Εισαγωγή και Επιμέλεια Γ. Κωνσταντινίδης, Πρόλογος Γ. Παγουλάτος, Εκδόσεις Κριτική, Αθήνα.
125. Haigh, J. (2009). Taking Chances. OXFORD UNIVERSITY PRESS.
126. Hair, J.F., Blank, W.C., Babin, B.J., Anderson, R. E. (2010). Multivariate Data Analysis. PEARSON.
127. Hakstian, A.R., Rogers, W.D., & Cattell, R.B. (1982). The behavior of numbers factors rules with simulated data. *Multivariate Behavioral Research*, 17, 193-219.
128. Halekoh, U., & Hojsgaard, S. (2007). Overdispersion. Retrieved March 18, 2007 from:  
<http://gbi.agrsci.dk/statistics/courses/phd207/material/day7/overdispersion-handout.pdf>.

129. Hamilton, L. C. 2013. *Statistics with Stata: Updated for Version 12*. 8th ed. Boston: Brooks/Cole.
130. Hanley, J. A., and B. J. McNeil. 1982. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *RADIOLOGY*, 143:1, 29-36.
131. Harman, H.H. (1967). Modern Factor Analysis. 2<sup>nd</sup> ed. Chicago: University of Chicago Press.
132. Harman, H.H. (1976). Modern Factor Analysis. 3rd ed. Chicago: University of Chicago Press.
133. Harman, H.H. and Jones, W.H. (1966). Factor analysis by minimizing residuals (Minres). *Psychometrika*, 31, 351-368.
134. Haynes, A. (2010). Writing successful Academic Books. CAMBRIDGE.
135. Heck, R.H., Thomas, S.L., Tabata, L.N., (2010). Multilevel and Longitudinal Modeling with IBM SPSS.
136. Henze, N., and B. Zirkler. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics, Theory and Methods* 19: 3595-3617.
137. Henze, N., and B. Zirkler. 1990. A class of invariant consistent tests for multivariate normality. *Communications in Statistics, Theory and Methods* 19: 3595-3617.
138. Hilbe, J.M. (2011). Negative Binomial Regression, second edition, CAMBRIDGE UNIVERSITY PRESS.
139. Hilbe, Joseph (2009). M-Logistic regression models-CRC Press
140. Hills, M. and De Stavola B.L. (2012). A Short Introduction to Stata for Biostatistics (Updated to Stata 12). Publisher: Timberlake Consultants.
141. Ho, R. (2006). Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS. Chapman & Hall/ CRC PRESS.

142. Hoelter, J.W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *SOCIOLOGICAL METHODS & RESEARCH*, 11, 325-344.
143. Hoffman, S.D., & Duncan, G.J. (1988). Multinomial and conditional logit discrete-choice models in demography. *DEMOGRAPHY*, 25, 415-427.
144. Hosmer, D.W., and Lemeshow, S. (1999). Applied survival analysis. New York: John Wiley and Sons.
145. Hosmer, D.W., and Lemeshow, S. (2000). Applied logistic regression. 2<sup>nd</sup> ed. New York: John Wiley and Sons.
146. Hosmer, D.W., and Lemeshow, S., and Sturdivant (2013). Applied logistic regression. 3<sup>rd</sup> ed., Wiley.
147. Howell, D. (1997). Statistical methods for Psychology. Fourth Edition, Duxbury Press.
148. Howell, D. (2007). Statistical methods for Psychology. International Student Edition, THOMSON-WADSWORTH.
149. Hox, J. (2010). Multilevel Analysis. ROUTLEGDE.
150. Hoyle, R.H., and Panter, A.T. (1995). Writing about structural equation models. In R.H. Hoyle (ed.), *Structural Equation Modeling: Concepts, issues and applications* (pp. 158-176). Thousand Oaks, CA: SAGE.
151. Hu L-T., and Bentler, P.M. (1995). Evaluating model fit. In R.H. Hole (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: SAGE.
152. Hu L-T., and Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *PSYCHOLOGICAL METHODS*, 3, 424-453.

153. Hu, L-T., and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
154. J.Scott LongInc (1997),Regression Models for Categorical and Limited Dependent Variables- Sage Publications,Advanced Quantitative Techniques in the Social Sciences 7).
155. Jaccard J., Jacoby, J. (2010). Theory Construction and Model-Building Skills.
156. Jaccard, J.J. (2001). Interaction Effects in Logistic Regression (Quantitative Applications in the Social Sciences). SAGE.
157. James, L.R., Mulaik, S.A., & Brett, J.M. (1982). Causal analysis: Assumptions, models, and data. Beverly Hills, CA:SAGE.
158. Jolliffe, I.T. (1986). Principal Component Analysis. New York, Springer –Verlag.
159. Jomeen J. & Martin C.R. (2005) Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. *Journal of Reproductive and Infant Psychology* 23, 143–154.
160. Joreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis, *Psychometrika*, 34, 183-202.
161. Joreskog, K. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P.R. Krishnaiah (ed). Applications of Statistics. Amsterdam: North-Holland Publishing Co., 265-287.
162. Joreskog, K., Sorbom, D. (1989a), 'LISREL 7: A guide to the program and Applications', Second Edition. Chicago SPSS Publications.
163. Joreskog, K., Sorbom, D. (1989b), 'LISREL 7: User's Reference Guide', SSI.
164. Joreskog, K., Sorbom, D. (1996-2001), 'LISREL 8: User's Reference Guide', SSI.
165. Joreskog, K., Sorbom, D. (1996-2002), 'PRELIS 2: User's Reference Guide', SSI.

166. Joreskog, K., Sorbom, D., du Toit S., Du Toit, M. (2000-2001), 'LISREL 8: New Statistical Features, SSI.
167. Joreskog, K.G. (1979). Basic ideas of factor and component analysis. In K.G. Joreskog & D. Sorbom: Advances in factor analysis and structural equation models. Cambridge, Mass.: Abt Books, 5-20.
168. Joreskog,K., Sorbom, D. (1993), 'LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language', SSI .
169. Judd, C., MacClelland, G., Ryan, C. (2009). Data Analysis: A model Comparison Approach. Second Edition.
170. Kachigan, S.K. (1986). Statistical Analysis: A interdisciplinary introduction to univariate and multivariate methods. New York: Radius.
171. Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational And Psychological Measurement*, 20, 141-151.
172. Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36.
173. Katz, M.H. (2010). Evaluating Clinical and Public Health Interventions. A practical guide to study design and statistics. CAMBRIDGE UNIVERSITY PRESS.
174. Kazmier, L. (2004). Business Statistics, Fourth edition. SCHAUM'S OUTLINES, McGraw, Hill.
175. Kelloway, E. K. (1995). Structural Equation Modeling in perspective. *Journal of Organizational Behavior*, 16, 215-224.
176. Kelloway, E. K. (1998). Using LISREL for Structural Equation Modeling. SAGE.
177. Kenneth P. Burnham, David R. Anderson (2002), Model Selection and Multimodel Inference\_ A Practical Information-Theoretic Approach (2002, Springer-Verlag New York).
178. Kenny, D.A. (1979). Correlation and causality. New York, Willey.

179. Kenny, D.A., & Judd, C.M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201-210.
180. Kim, J.O., Mueller, C.W. (1978). Factor analysis. Statistical methods and practical issues. University paper series/ Number 07-014. On quantitative applications in the social sciences, Newbury Park CA:SAGE.
181. Kluegel, J.R., Singleton, R., Starnes, Jr., and C.E. (1977). Subjective class identification: A multiple indicator approach. AMERICAN SOCIOLOGICAL REVIEW, 42: 599-611.
182. Kohler, U., and F. Kreuter. 2012. *Data Analysis Using Stata*. 3rd ed. College Station, TX: Stata Press.
183. Kremelberg, D. (2011). Practical Statistics. SAGE.
184. Kyle C. Longest (2012). Using Stata for Quantitative Analysis-Sage.
185. Landis, J.R. and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174. [14.3.1].
186. Lawley, D.N., and Maxwell, A.E. (1963). Factor Analysis as a Statistical Method. London: Butterworth.
187. Le, C.T. (2010). Applied Categorical Data Analysis, and Translational Research. Second edition, WILEY.
188. Lee, E. T. (1992). *Statistical Methods for survival data Analysis*. New York: John Wiley and Sons.
189. Lee, E. T., Wang, J. (2003). *Statistical Methods for survival data Analysis*. New York: John Wiley and Sons.
190. Leyland, A.H., Goldstein, H. (2001). Multilevel Modelling of Health Statistics. WILEY.
191. Linn, R.L.(1968). A Monte Carlo approach to the number of factor problem. *Psychometrika*, 33, 37-71.

192. Lipschutz, S., Lipson, M.L. (1992). Discrete Mathematics. SCHAUM'S. McGraw, Hill.
193. Lipschutz, S., Lipson, M.L. (2007). Discrete Mathematics. SCHAUM'S, third edition, McGraw, Hill.
194. Lipschutz, S., Schiller, J. (1998). Introduction to Probability and Statistics. SCHAUM'S OUTLINES.
195. Lisa Daniels, Nicholas Minot - Introduction to Statistics and Data Analysis Using Stata\_ From Research Design to Final Report-SAGE Publishing (2019).
196. Loehlin, J.C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis. 2<sup>nd</sup> edition.* Erlbaum.
197. Long, J. Scott, Freese Jeremy (2006). Regression Models for Categorical Dependent Variables Using Stata. STATA Press.
198. Long, J.\_Scott (1997). Regression\_Models\_for\_Categorical and Limited Dependent Variables.
199. Long, J.\_Scott (1997). Regression\_Models\_for\_Categorical and Limited Dependent Variables. SAGE.
200. Long, J.S, Freese, J. (2014). Regression Models for Categorical Dependent Variables Using Stata. Third Edition, STATA PRESS.
201. Longest, K. (2012). Using Stata for Quantitative Analysis. Publisher: SAGE.
202. Lord, F.M. and Novick, M.R. (1968). Statistical theories of mental test scores. Addison-Wesley, Reading, MA.
203. Lundgren, E., Rastad, J., Thrufjell, E., Akerstrom, G., Ljunghall. (1997). 'Population-based screening for primary hyperparathyroidism with serum calcium and parathyroid hormone values in menopausal women, SYRGERY 121 (3): 287-294.

204. MacCallum, R.C. (1995). Model specification: Procedures, strategies, and related issues. In R.H. Hoyle (ed.), Structural equation modeling: Concepts, issues, and applications (pp.16-36). Thousand Oaks, C.A: SAGE.
205. MacCallum, R.C., and Austin, J.T. (2000). Application of structural equation modeling in psychological research. ANNUAL REVIEW OF PSYCHOLOGY, 51, 201-226.
206. MacCallum, R.C., Browne, M.W., and Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling, PSYCHOLOGICAL METHODS, 1, 130-149.
207. MacCallum, R.C., K.F. Widaman, S. Zhang, and S.Hong (1999). Sample size in factor analysis. Phychological Methods, 4(1): 84-99.
208. MacCullagh, P., and Nelder, J.A. (1989). Generalized Linear Models, 2<sup>nd</sup> ed. London: Chapman and Hall.
209. MacDonald, R.P. (1989). An index of goodness-of-fit based on noncentrality. JOURNAL OF CLASSIFICATION, 6, 97-103.
210. MacFadden, D. (1973). Conditional Logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers of econometrics (pp. 105-142). New York: Academic Press.
211. MacKenzie, S.B., Podsakoff, P.M. and Jarvis, C.B. (2005). *The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions*. JOURNAL OF APPLIED PSYCHOLOGY 90 (July): 710-30.
212. Maindonald, J., Braun, J. (2010). Data Analysis and Graphics Using R. CAMBRIDGE UNIVERSITY PRESS.

213. Malley, J., Malley, K., and Pajevic, S. (2001). Statistical Learning for Biomedical Data (Practical Guides to Biostatistics and Epidemiology), CAMBRIDGE UNIVERSITY PRESS.
214. Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications *Biometrika* 57: 519-530.
215. Mardia, K.V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *SANKHYA*, B36, 115-128.
216. Marsh, C. (1982). The Survey Method: the contribution of surveys to sociological explanation. London: George Allen and Unwin.
217. Marsh, H.W. and Hocevar, D., (1985). 'Application of confirmatory factor analysis to the study of self-concept: First –and higher order factor models and their invariance across groups', *Psychological Bulletin*, 97, 562-582.
218. Marsh, H.W., Balla, J.R., & MacDonald, R.P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: the effect of sample size. *Psychological Bulletin*, 88, 245-258.
219. Mathews, J.R., Mathews, R.W. (2010). Successful Scientific writing. Third edition, CAMBRIDGE.
220. Mayer, N. (2005), "Εκλογική Συμπεριφορά, ιστορικές σχολές και μοντέλα ανάλυσης", Μετάφραση-επιμέλεια-βιβλιογραφία Χριστόφορος Βερναρδάκης, Εκδόσεις Σαββάλας, Αθήνα.
221. McKelvey, R. and Zavoina, W. (1975). A Statistical model for the analysis of ordinal level dependent variables. *JOURNAL OF MATHEMATICAL SOCIOLOGY* 4: 103-120.
222. Medsker, G.J. Williams, L.J., & Holahan, P.J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20, 439-464.

223. Menard, S. (2010). Logistic Regression. SAGE.
224. Meng, X., & Miller, P. (1995). Occupational segregation and its impact on gender wage in China's rural industrial sector. OXFORD ECONOMIC PAPERS, 47, 136-155.
225. Meyer, R., Krueger, D. (2001). A minitab Guide to Statistics. Second Edition. PRENTICE HALL.
226. Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105, 156-166.
227. Michael\_N.\_Mitchell (2004). A Visual Guide to Stata Graphics. STATA Press.
228. Miller, P. W., and Volker, P.A. (1985). On the determination of occupational attainment and mobility. JOURNAL OF HUMAN RESOURCES, 20: 197-213.
229. Mills, T. C., Markellos, Raphael N (2008). The Econometric Modelling of Financial Time Series-Cambridge University Press.
230. Mitchell, M. (2012). A Visual Guide to Stata Graphics, Third Edition. Publisher: STATA PRESS.
231. Mitchell, M. (2012). Interpreting and Visualizing Regression Models Using Stata. Publisher: STATA PRESS.
232. Morrison, D.F. (1976). Multivariate statistical methods (2d ed.). New York: McGraw-Hill.
233. Moser, C., and Kalton, G. (1993). Survey Methods in Social Investigation. ALDERSHOT: DARTMOUTH.
234. Mulaik, S.A. (1972). The Foundation of Factor Analysis. New Work: McGraw-Hill.
235. Mulaik, S.A., James, L.R., Van Altine, J., Bennett, N., Lind, S., and Stillwell, C.D. (1989). An evaluation of goodness of fit indices for structural equation models. *Psychological Bulletin*, 105, 430-445.

236. Muthen, B., Kaplan, D (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. BRITISH JOURNAL OF MATHEMATICAL AND STATISTICAL PSYCHOLOGY, 38, 171-189.
237. Nerlove, M., & Press, S.J. (1973). Univariate and Multivariate log-linear and Logistic Models. SANTA MONICA, CA: RAND.
238. Norman, G., and Streiner, D. (2008). *Biostatistics. The bare Essentials*. Third Edition. BC DECKER.
239. Norusis, M. J. (2006). SPSS 15.0 Statistical procedures companion. Prentice Hall, Inc.
240. Norusis, M.J. (2007), "SPSS 15.0 Advanced Statistical Procedures Companion", Prentice Hall, Chicago.
241. Norusis, Marija (1993). SPSS for Windows: Base Systems User's Guide, *release 6.0*. SPSS Inc.
242. Norusis, Marija (1999). SPSS Base 10.0 Application Guide. SPSS Inc.
243. Norusis, Marija (2000). SPSS 10.0 Guide to Data Analysis. Prentice Hall.
244. Norusis, Marija (2002). SPSS 11.0 Guide to Data Analysis. Prentice Hall.
245. Norusis, Marija (2003). SPSS 12.0 Statistical Procedures Companion. Prentice Hall.
246. Norusis, Marija (2009). PASW Statistics 18 Statistical Procedures Companion. Prentice Hall.
247. Norusis, Marija (2010). PASW Statistics 18, Advanced Statistical Procedures Companion. Prentice Hall.
248. Norusis, Marija (2010). PASW Statistics 18, Guide to Data Analysis. Prentice Hall.

249. NORUSIS, M.J. (2004), "SPSS 13, Advanced Statistical Procedures Companion, Upper Saddle River: Prentice Hall".
250. NORUSIS, M.J. (2005), "SPSS 14.0 Statistical Procedures Companion, Upper Saddle River: Prentice Hall".
251. Pallant J. (2010). SPSS Survival Manual, 4<sup>th</sup> edition. McGraw, Hill.
252. Pavlopoulos, D., Muffels, R. & Vermunt, J.K. (2010). "Wage mobility in Europe. A comparative analysis using restricted multinomial logit regression", Quality and Quantity, 44 (1), p.115-129.
253. Pearson, R.W. (2010). Statistical Persuasion, SAGE.
254. Pop, V.J., Komproe, I.H., van Son, M.J., 1992. Characteristic of the Edinburgh Post Natal Depression Scale in the Netherlands. *J. Affect. Disord.* 26, 105–110.
255. Rabe- Hesketh, S. Skrondal, Anders. (2012). Multilevel and Longitudinal Modeling Using STATA. Vol. I.
256. Rabe- Hesketh, S. Skrondal, Anders. (2012). Multilevel and Longitudinal Modeling Using STATA. Vol. II.
257. Rabe-Hesketh, S. and Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata, third edition. Volume I and Volume II. Publisher: STATA PRESS.
258. Raftery, A.E. (1993). Bayesian model selection in structural equation models. In K.A. Bollen & J. Long (Eds.). Testing structural equation models (pp. 163-180). Newbury Park, C.: SAGE.
259. Raftery, A.E. (1995). Bayesian model selection in social research. In Vol. 25 of Sociological Methodology, ed. P.V. Marsden, 111-163. Oxford: Blackwell.
260. Rich, B., Schmidt, P.A. (2004). Elementary Algebra. SCHAUML'S , third edition.
261. Rogers, Alan (1998). '*H Εκπαίδευση των Ενηλίκων*', Ελληνική μετάφραση, Εκδόσεις Μεταίχμιο, 1998.

262. Ross L.E., Gilbert Evans S.E., Sellers E.M., *et al.* (2003) Measurement issues in postpartum depression part 1: anxiety as a feature of postpartum depression. *Archives of Women's Mental Health* **6**, 51–57.
263. Rossi, R.J. (2010). Applied Biostatistics for the Health Sciences, WILEY.
264. Rothwell, N. (2004). Who wants to be a scientist? CAMBRIDGE.
265. Rummel, R. J. (1970). Applied factor analysis. Evanston, IL: Northwestern University Press.
266. Ruppert, D. Wand, M.P. Carroll, J. R. (2009). Semiparametric Regression. CAMBRIDGE.
267. Sapnas, K. G., and Zeller, R.A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement*, **10** (2), 135-153.
268. Satorra A. & Bentler P.M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. In American Statistical Association 1988 proceedings of the business and economics section (pp. 308-313). Alexandria, VA: AMERICAN STATISTICAL ASSOCIATION.
269. Satorra, A. & Bentler, P.M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C.C. Clogg (Eds.), Latent variables analysis: Applications for developmental research (pp. 399-419). Thousand Oaks, CA: SAGE.
270. Schene, A., Wijngaarden, B., Koeter, M. (1998). "Family Caregiving in Schizophrenia: Domains, Distress". *SCHIZOPHRENIA BULLETIN*, **24** (4): 609-618.
271. Schlesselman James (1982). Case-Control Studies: Design, Conduct, Analysis (Monographs in Epidemiology and Biostatistics). New York and London: Oxford University Press.

272. Schmidt, P., & Strauss, R.P. (1975). The prediction of occupation using multiple logit models. *INTERNATIONAL ECONOMIS REVIEW*, 16, 471-486.
273. Sharma, S. (1996). Applied Multivariate Techniques. New York: John Wiley & Sons.
274. Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86,420-428.
275. Sides John, Vavreck Lynn (2013). The Gamble: Choice and Chance in the 2012 Presidential Election. Princeton University Press.
276. Sorbom, D., & Joreskog, K. (1982). The use of structural equation models in evaluation research. In Fornell (ed.), A Second Generation of Multivariate Analysis: Vol 2. Measurement and Evaluation (pp. 381-418). New York: Praiger.
277. SPSS (1997). Advanced Statistics, 7.5, SPSS inc
278. SPSS (1999). Advanced Models, 9.0, SPSS inc
279. SPSS Base 10.0 (1999). Application Guide, SPSS inc.
280. Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
281. Sternberg, R.J. (2007). Guide to Publishing in Psychology Journals. CAMBRIDGE.
282. Stevens, J. (2002). Applied Multivariate Statistics for the Social Sciences. 4rd ed.
283. Stine, R.A. (1989). An introduction to bootstrap methods: Examples and ideas. *Sociological Methods and Research*, 18, : 243-291.
284. Streibig, J.C., Ritz, C. (2008). Nonlinear Regression with R. Springer.
285. Tabachnick, B.G. and Fidell, L. S. (1996). Using Multivariate Statistics. 3rd ed. NY: Addison-Wesley-Longman.

286. Tabachnick, B.G. and Fidell, L. S. (2007). Using Multivariate Statistics. fifth ed. Pearson International Edition., Allyn and Bacon.
287. Tanaka, J.S. (1993). Multifaceted conceptions of fit. In K.A. Bollen and Long (eds), Testing Structural Models. Newbury Park, CA: SAGE Publications.
288. Tanaka, J.S., & Huba, G.J. (1984). Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology*, 46,621-635.
289. Tanaka, J.S., & Huba, G.J. (1989). A general coefficient of determination for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 42, 233-239.
290. Tarling. R. (2009). Statistical Modelling for Social Researchers. ROUTLEGDE.
291. Teddlie, C., Tashakkori, A. (2009). Foundations of Mixed Methods Research. SAGE.
292. Thurstone, L.L. (1938). Primary mental abilities. *Psychometric Monographs*, 1.
293. Train, Kenneth E. (2009). Discrete Choice Methods with Simulation. Cambridge.
294. Tucker, L.R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Phychometrika*, 38, 1-10.
295. Upton, G., Cook, I. (2008). Oxford Dictionary of Statistics. OXFORD UNIVERSITY PRESS.
296. Van der Ark, Croon, M.A., Sijtsma, K. (2005). New Developments in Categorical Data Analysis for the Social and Behavioral Sciences. LEA.
297. Vittinghoff, E., Glidden, D., Shiboski, S., and McCulloch, C. (2012). Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models, Second Edition. SPRINGER.
298. Vogt, P.W., Johnson, B.R. (2011). Dictionary of statistics & Methodology. Fourth Edition, SAGE.

299. Warner, R. (2008). Applied Statistics. From Bivariate through Multivariate Techniques. SAGE.
300. Warner, R. (2008). Applied Statistics. SAGE.
301. Weakliem, D. (2016). Hypothesis Testing and Model Selection in The Social Sciences. Guilford Press.
302. Wert, C.E., Rock, D.A., Linn, R.L., & Joreskog, K.G. (1977). Validating psychometric assumptions within and between populations. *Educational and Psychological Measurements*, 37, 863-871.
303. West, S.G., Finch, J.F. & Curran,P.J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R.H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.
304. Wheaton, B., Muthen, B., Alwin, D.F., and Summers, G.F. (1977). Assessing reliability and stability in panel models. In *Sociological Methodology 1977*, ed. D.R. Heise, 84-136. San Francisco: Jossey-Bass.
305. Wildenmann, R. (1998), "Η εκλογική έρευνα. Συμπεριφορά του εκλογικού σώματος και ανάλυση εκλογών", Πρόλογος- μετάφραση-σχόλια Β. Γεωργιάδου, Εκδόσεις Παπαζήση, Αθήνα.
306. Williams, (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *SOCIOLOGICAL METHODS and RESEARCH* 37: 531-559.
307. Williams, L.J. & Hollahan, P.J. (1994). Parsimony- based fit indices for multiple-indicator models: Do they work? *Structural equation modeling: Issues and techniques* (pp. 7-56. Mahwah, NJ: Erlbaum.

308. Williams, L.J. (1995). Covariance structure modeling in organizational research: Problems with the method vs. applications of the method. *Journal of Organizational Behavior*, 16, 225-233.
309. Winship, C. and Mare, R. (1984). Regression models with ordinal variables. *AMERICAN SOCIOLOGICAL REVIEW* 49: 512-525.
310. Winstanley, C. (2010). Writing a Dissertation for DUMMIES. WILEY.
311. Witten, I.H., Frank, Hall, M.A. (2011). Data Mining. Third edition. MORGAN KAUFMANN.
312. Wooldridge, J. (2010). Econometric Analysis of Cross Section and Panel Data, 2<sup>nd</sup> ed. Cambridge, MA: MIT PRESS.
313. Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5, 161- 215.
314. Yuan, K.H., & Bentler.P.M. (1998). Normal theorybased test statistics in structural equation modeling. *British Journal of Mathematical and Statistical Psychology*, 51, 289-309.
315. Zeller, R.A. (2005). How few cases is enough to do a credible factor analysis? A Monte Carlo simulation. Manuscript submitted for publication.
316. Zhang, Z. (2016). Model Building Strategy for Logistic Regression: Purposeful Selection. *Annals of Translational Medicine*, 2016; 4(16):111.
317. Zhu, W. (1997). Making bootstrap statistical inference: A tutorial. *Research Quarterly for Exercise and Sport*, 68, 44-55.
318. Zolman, J. (1993). *Biostatistics. Experimental Design and Statistical Inference*. Oxford University Press.
319. Zweig, M. H., and G. Campbell. 1993. Receiver Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine. *CLINICAL CHEMISTRY*, 39:4, 561-577.

## Ελληνική Βιβλιογραφία

1. Βερναρδάκης Χριστόφορος (2011), “Πολιτικά κόμματα, εκλογές και κομματικό σύστημα: Οι μετασχηματισμοί της πολιτικής αντιπροσώπευσης 1990-2010”, Εκδόσεις Σάκκουλα Α.Ε.
2. Δαφέρμος Βασίλης (2013), “Παραγοντική ανάλυση”, Εκδόσεις Ζήτη, Θεσσαλονίκη.
3. Δαφέρμος Βασίλης (2011), “Κοινωνική στατιστική και μεθοδολογία έρευνας με το SPSS”, Εκδόσεις Ζήτη, Θεσσαλονίκη.
4. Μακκομπς, M., και Εϊνσίντελ, E., και Ονίβερ, N. (1996), “Τα μέσα μαζικής ενημέρωσης και η διαμόρφωση της κοινής γνώμης”, Επιμέλεια σειράς Στέλιος Παπαθανασόπουλος, Εκδόσεις Καστανιώτη, Αθήνα. Ρεγγίνα Κασιμάτη, Στράτος Γεωργούλας, Μαρία Παπαϊωαννου, Ιωάννης Πρανταλος “Κοινωνιολογία, Γ' Γενικού Λυκείου”, Βιβλίο του μαθητή.
5. Μαυρής Γ. & Συμεωνίδης Γ. (2005), «Βουλευτικές εκλογές 2004: Πρόβλεψη του εκλογικού αποτελέσματος και επίδραση της προεκλογικής περιόδου», στον τόμο: V-PRC, Η Κοινή Γνώμη στην Ελλάδα 2004: Εκλογές, Κόμματα, Ομάδες Συμφερόντων, Χώρος και Κοινωνία. Αθήνα, Εκδόσεις Σαββάλας.
6. Μενδρινού Μαρία (2014), “Η ψήφος των ευρωεκλογών του 2014 ως αναδρομική οικονομική ψήφος”, στο “Η Ευρώπη φοβάται την Ευρώπη; Μια αποτίμηση του αποτελέσματος των ευρωεκλογών του 2014”, Ινστιτούτο Δημοκρατίας Κωνσταντίνος Καραμανλής, Αθήνα.
7. Νικολακόπουλος, H. (2001), Η Καχεκτική Δημοκρατία: Κόμματα και Εκλογές 1946-1967. Αθήνα, Εκδόσεις Πατάκης.
8. Νικολακόπουλος H. (2003), «Η ανάπτυξη των πολιτικών δημοσκοπήσεων στην Ελλάδα», στο [www.sedea.gr](http://www.sedea.gr).

**9.** Πιερίδης Κωστής(2021), «ΔΙΑΙΡΕΤΙΚΕΣ ΤΟΜΕΣ ΚΑΙ ΕΚΛΟΓΙΚΗ ΣΥΜΠΕΡΙΦΟΡΑ ΣΤΗ ΣΥΓΧΡΟΝΗ ΕΛΛΑΣΑ (2004-2018) »,Εκδόσεις Επίκεντρο.

**10.** Ρεγγίνα Κασιμάτη, Στράτος Γεωργούλας, Μαρία Παπαϊωαννου, Ιωάννης Πρανταλος “Κοινωνιολογία, Γ' Γενικού Λυκείου”, Βιβλίο του μαθητή.