



ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΟΙΚΟΝΟΜΙΑΣ & ΔΗΜΟΣΙΑΣ ΔΙΟΙΚΗΣΗΣ
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ & ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΕΦΗΡΜΟΣΜΕΝΑ ΟΙΚΟΝΟΜΙΚΑ ΚΑΙ ΠΕΡΙΦΕΡΕΙΑΚΗ ΑΝΑΠΤΥΞΗ»
ΚΑΤΕΥΘΥΝΣΗ: ΕΦΗΡΜΟΣΜΕΝΑ ΟΙΚΟΝΟΜΙΚΑ ΚΑΙ ΔΙΟΙΚΗΣΗ

ΤΙΤΛΟΣ ΕΡΓΑΣΙΑΣ

**ΕΠΕΚΤΑΣΕΙΣ ΤΟΥ ΜΟΝΤΕΛΟΥ ΤΗΣ ΛΟΓΙΣΤΙΚΗΣ
ΠΑΛΙΝΔΡΟΜΗΣΗΣ**

Επιβλέπων Καθηγητής:
Φοιτητής:
Ακαδημαϊκό Έτος:
Α.Μ.:

Clive Richardson
Καραγεωργίου Ευάγγελος
2020-2021
0819M006

Περιεχόμενα

Περιεχόμενα.....	2
Εισαγωγή.....	4
Περίληψη.....	4
Abstract	5
Κεφάλαιο 1. Λογιστική Παλινδρόμηση.....	6
1.1 Σύντομη Ιστορική αναφορά	6
1.2 Γενικά.....	6
1.3 Πολλαπλή Διωνυμική Παλινδρόμηση	12
1.4 Πολυωνυμική Παλινδρόμηση	26
Κριτήριο ελέγχου της ισότητας των κλίσεων	28
Συντελεστές παλινδρόμησης.....	29
Λόγος των πιθανοτήτων έκβασης.....	29
Κεφάλαιο 2. Το Διαδίκτυο και το Online Dating	33
2.1. Εισαγωγή	33
2.2. Ιστορία και ανάπτυξη του Διαδικτύου.....	33
2.3. Παραπλανητική αυτο-παρουσίαση στο Διαδίκτυο.....	35
2.4. Ηλεκτρονική απάτη	37
2.5. Διαφορές φύλου στην επιλογή συντρόφων	39
2.6. Διαμόρφωση διαδικτυακών σχέσεων.....	40
2.7. Πρόσωπο με πρόσωπο έναντι Διαδικτύου	41
2.8. Μετακίνηση Διαδικτυακών σχέσεων σε σχέσεις εκτός σύνδεσης	44
Κεφάλαιο 3. Μεθοδολογία έρευνας.....	46
3.1. Ερευνητικός σκοπός-ερωτήματα	46
3.2. Σχεδιασμός έρευνας.....	47
3.3. Δείγμα έρευνας	47
3.4. Μεταβλητές έρευνας	47
3.5. Μέθοδοι Στατιστικής ανάλυσης.....	48
3.6. Ηθικά Ζητήματα	49
Κεφάλαιο 4. Αποτελέσματα	51
4.1. Περιγραφική Στατιστική	51
4.2. Επαγωγική Στατιστική	56
4.2.1. 1 ^ο ερευνητικό ερώτημα:.....	56

4.2.2. 2 ^ο ερευνητικό ερώτημα:.....	62
4.2.3. 3 ^ο ερευνητικό ερώτημα:.....	67
4.2.4. 4 ^ο ερευνητικό ερώτημα:.....	71
Κεφάλαιο 5. Συμπεράσματα	80
Βιβλιογραφικές αναφορές	83
Ελληνική	83
Ξενόγλωσση.....	83

Εισαγωγή

Στη σημερινή εποχή είναι μια μεγάλη ανάγκη να κατατάσσουμε παρατηρήσεις σε γνωστές ομάδες – πληθυσμούς καθώς επίσης και να κάνουμε προβλέψεις. Υπάρχουν πολλές μέθοδοι που κάνουν ή έχουν ως σκοπό να κατατάσσουν παρατηρήσεις. Το υπόδειγμα της λογιστικής ή της λογαριθμικής παλινδρόμησης είναι η σημαντικότερη στατιστική μέθοδος που χρησιμοποιείται κυρίως στην αιτιογνωστική, αλλά και στην διαγνωστική και στην προγνωστική έρευνα. Ο λόγος που είναι η σημαντικότερη στατιστική μέθοδος είναι ότι σε πάρα πολλές έρευνες δεν αρκούν μόνο ποσοτικές μεταβλητές, αλλά πρέπει να χρησιμοποιούμε και κάποιες ποιοτικές (κατηγορικές) μεταβλητές, των οποίων οι τιμές τους δεν είναι αριθμοί. Η λογιστική παλινδρόμηση είναι μια τεχνική σχεδιασμένη για την πραγματοποίηση ανάλυσης δεδομένων που αφορούν την μελέτη και την πρόβλεψη τιμών κάποιας κατηγορικής εξαρτημένης μεταβλητής και χρησιμοποιεί ποσοτικές και ποιοτικές ανεξάρτητες μεταβλητές.

Περίληψη

Στην παρούσα εργασία παρουσιάζεται η έννοια της λογιστικής παλινδρόμησης, μιας μεθόδου που χρησιμοποιείται ευρέως σε στατιστικές αναλύσεις. Η συγκεκριμένη μελέτη επικεντρώνεται, στην έννοια της λογιστικής παλινδρόμησης, καθώς και στις επεκτάσεις χρησιμοποίησης του μοντέλου αυτής. Στην περίπτωση αυτή φανερώνεται μια επιπλέον πολυπλοκότητα, και αυτό που εξετάζουμε είναι κατά πόσο και με ποιον τρόπο μπορούμε να χρησιμοποιήσουμε το μοντέλο της λογιστικής παλινδρόμησης και πέραν του συνηθισμένου τρόπου χρησιμοποίησης αυτού, έτσι ώστε να εξάγουμε ασφαλή συμπεράσματα.

Παρουσιάζονται και αναλύονται οι μελέτες περιπτώσεων από μια έρευνα σχετικά με τον online dating και την χρήση του διαδικτύου από τους ανθρώπους στην εποχή μας και στις γνωριμίες μεταξύ ανθρώπων μέσω διαδικτύου συγκριτικά με την κατ' ιδίαν γνωριμία. Για την ανάλυση των δεδομένων χρησιμοποιείται το πρόγραμμα SPSS.

Abstract

This paper presents the concept of logistic regression, a method that is widely used in statistical analysis. This study focuses on the concept of logistic regression, as well as the extensions of the use of this model. In this case, an additional complexity is revealed, and what we are examining is to what extent and in what way we can use the model of logistic regression and beyond the usual way, so that can draw reliable conclusions.

Case studies are presented and analyzed from a study on online dating and the use of the internet by people in our time and on online dating compared to personal dating. The SPSS program is used for data analysis.

Κεφάλαιο 1. Λογιστική Παλινδρόμηση

1.1 Σύντομη Ιστορική αναφορά

Η Λογιστική παλινδρόμηση αναπτύχθηκε από τον στατιστικό David Cox το 1958 (Cramer 2002). Η Λογιστική παλινδρόμηση χρησιμοποιείται σε διάφορους τομείς, συμπεριλαμβανομένης της μηχανικής μάθησης, των περισσότερων ιατρικών πεδίων, των στατιστικών και των κοινωνικών επιστημών. Χρησιμοποιείται επίσης σε εφαρμογές μάρκετινγκ όπως η πρόβλεψη της τάσης του πελάτη να αγοράσει ένα προϊόν ή να σταματήσει μια συνδρομή κ.λπ. Οι ρίζες της εξαπλώθηκαν πολύ πίσω στις αρχές του 19ου αιώνα, η επιβίωση του όρου logistics και η ευρεία εφαρμογή της εξίσωσης της Λογιστικής Παλινδρόμησης καθορίστηκαν αποφασιστικά από τις προσωπικές ιστορίες και τις ατομικές ενέργειες μερικών μελετητών. Η εξίσωση της Λογιστικής Παλινδρόμησης ανακαλύφθηκε εξ'αρχής το 1920 από τον Pearl και Reed και χρησιμοποιήθηκε σε μια μελέτη σχετικά με την ανάπτυξη του πληθυσμού, στις Ηνωμένες Πολιτείες της Αμερικής. Ωστόσο, ο όρος Λογιστική δεν είχε ξαναχρησιμοποιηθεί πριν από τον Βέλγο μαθηματικό, τον Verhulst, όπου και την όρισε, λίγο μετά από το άρθρο των Pearl και Reed το 1920.

1.2 Γενικά

Η λογιστική παλινδρόμηση (Logistic regression) αποτελεί στην ουσία ένα μοντέλο ταξινόμησης των τιμών μιας μεταβλητής απόκρισης Y με βάση τη θεωρία των πιθανοτήτων. Στο μοντέλο αυτό όπου η μεταβλητή Y συνήθως έχει δυαδικό χαρακτήρα (λαμβάνει δύο τιμές) στοχεύεται η πρόβλεψη της έκβασης αυτής από ένα πλήθος προβλεπτικών μεταβλητών που μπορεί να είναι ονομαστικές, τακτικές ή ποσοτικές.

Η σημαντικότερη διαφοροποίηση μεταξύ λογιστικής και γραμμικής παλινδρόμησης βασίζεται στη φύση της επιλεγμένης μεταβλητής απόκρισης, η οποία στην μεν πρώτη μπορεί να είναι κατηγορική, (τακτική ή ονομαστική), στη δε δεύτερη αποκλειστικά ποσοτική. Κατά τη λογιστική παλινδρόμηση η εκτίμηση των παραμέτρων γίνεται με τη μέθοδο του λόγου πιθανοφάνειας (μέθοδος συνήθως εφαρμοζόμενη στα γενικευμένα γραμμικά υποδείγματα), δηλαδή επιλέγονται οι πιο πιθανοφανείς τιμές

των παραμέτρων, προκειμένου να οδηγήσουν στα παρατηρούμενα αποτελέσματα. Ως επακόλουθο, η πρώτη παραδέχεται την ύπαρξη ομοιογένειας (ομοσκεδαστικότητας) στα υπολείμματα των αποκρίσεων ενώ στη δεύτερη αναπτύσσεται πάντα ετεροσκεδαστικότητα σε κάθε προβλεπόμενη τιμή εξαιτίας του μεταβαλλόμενου ποσοστού διακύμανσης που αναλογεί σε αυτήν.

Διακρίνονται τρεις τύποι λογιστικής παλινδρόμησης ανάλογα με την ιδιαίτερη φύση της εξαρτημένης κατηγορικής μεταβλητής η οποία μπορεί να είναι:

1. **Δίτιμη** ή **δυναδική** ή **διχοτομική** (binary) ή **διμερής** εξαρτημένη μεταβλητή. Συνίσταται από δύο κατηγορίες, όπως π.χ. είναι οι εκβάσεις επιτυχία/αποτυχία, ΝΑΙ/ΟΧΙ, γεγονός απών/παρόν.
2. **Τακτική** (ordinal) μεταβλητή. Η εξαρτημένη μεταβλητή συνίσταται από τρεις ή περισσότερες κατηγορίες μεταξύ των οποίων ισχύει η έννοια της ανισότητας, όπως π.χ. σε μια ερώτηση της κλίμακας διαφωνώ καθόλου, λίγο, μέτρια, αρκετά, πολύ, στην κατάταξη ενός στρώματος υλικού ως λεπτού, μεσαίου, παχέος.
3. **Ονομαστική** (Nominal) ή **πολυωνυμική** (polynomial) ή **πολυχοτομική** (polytomus) ή **κατηγορική αδιαβάθμητη** (non-ordered categorical) ή **πολυμερής** μεταβλητή απόκρισης. Περιέχει τρεις ή περισσότερες κατηγορίες χωρίς κάποια φυσική διαβάθμιση, όπως π.χ. ο χαρακτηρισμός ενός τροφίμου ως τραγανού, μαλακού, εύθρυπτου ή του χρώματος αντικειμένων ως ερυθρού, πράσινου, κίτρινου κτλ.

Η λογιστική παλινδρόμηση επινοήθηκε ως εναλλακτική επιλογή της γραμμικής διακριτικής ανάλυσης για την ταξινόμηση των στοιχείων (ονομαστικών ή τακτικών) της εξαρτημένης, με ευρεία απήχηση σε πολλά διαφορετικά επιστημονικά πεδία και κυρίως στην ιατρική και τις κοινωνικές επιστήμες. Χαρακτηριστικά, χρησιμοποιείται στην πρόβλεψη της:

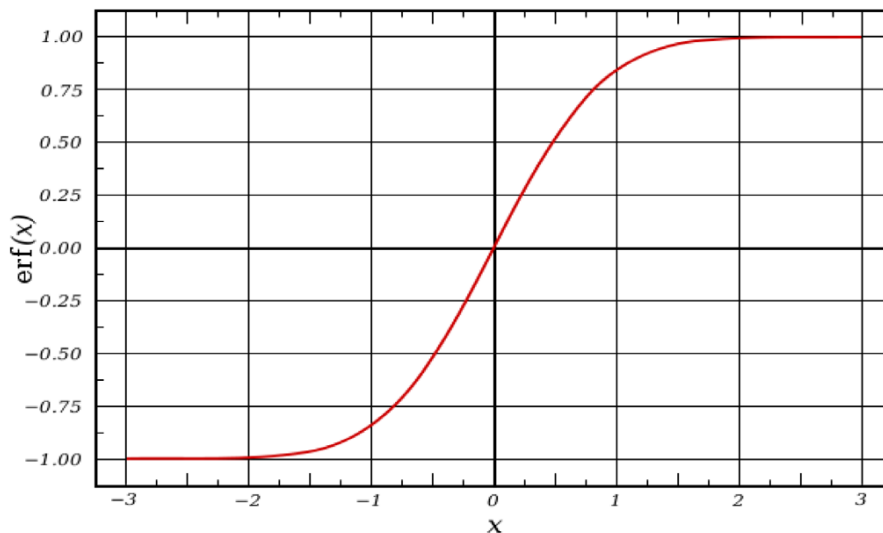
- εμφάνισης ή μη μιας νόσου (π.χ. διαβήτη) από ένα σύνολο διαφορετικών χαρακτηριστικών του πάσχοντος ατόμου (ηλικία, φύλο, αιματολογικά, ηλεκτροκαρδιογράφημα κτλ.)
- επιλογής ενός πολιτικού κόμματος με βάση την καταγραφή των δημογραφικών στοιχείων των πολιτών, όπως είναι η ηλικία, φύλο, φυλή, τόπος διαμονής, εισόδημα, προηγούμενη ψηφοφορία

- πιθανότητας αποτυχίας μιας διεργασίας παραγωγής προϊόντος σε ένα εργοστάσιο τροφίμων
- πρόβλεψη της πρόθεσης αγοράς ενός αγαθού από έναν καταναλωτή (έρευνα αγοράς)
- πιθανότητας αθέτησης από δανειολήπτη της αποπληρωμής του δανείου του.

Λεπτομερής περιγραφή των μεθόδων της λογιστικής παλινδρόμησης παρέχεται από τα συγγράμματα των Cox & Snell (1989), των Hosmer & Lemeshow (2000), των Long & Freese (2014) και συνδυαστικά με τη χρήση των πινάκων ενδεχομένων από τους Everitt (1992) και Agresti (1996). Η κατανόηση των όρων και μαθηματικών τύπων που συνοδεύουν τη μελέτη της λογιστικής παλινδρόμησης αποτελεί κυριολεκτικά πρόκληση για τον απλό επιστήμονα. Ως εκ τούτου, στο κεφάλαιο αυτό γίνεται αναφορά μόνο των αναγκαίων τεχνικών εφαρμογής της λογιστικής παλινδρόμησης με φειδωλή χρήση των μαθηματικών τύπων και με ιδιαίτερο βάρος στην απόδοση της ερμηνείας των ερευνητικών αποτελεσμάτων σε θέματα κοινωνικών επιστημών.

Ανάπτυξη του μοντέλου

Στη γλώσσα της στατιστικής, η λογιστική παλινδρόμηση χρησιμοποιείται για την πρόβλεψη της πιθανότητας εμφάνισης ενός γεγονότος προσαρμόζοντας τα δεδομένα της μελέτης στην εξίσωση της λογιστικής καμπύλης όπως αυτή παρουσιάζεται στο σχήμα 1.1.



Σχήμα 1.1 Τυπική ανάπτυξη σιγμοειδούς καμπύλης.

Η καμπύλη αυτή έχει σιγμοειδή μορφή και χαρακτηρίζεται από ένα στάδιο εκθετικής ανάπτυξης στο οποίο ο ρυθμός αύξησης επιβραδύνεται βαθμιαία και περατώνεται στο ασυμπτωτικό στάδιο κορεσμού της ανάπτυξης (η ευθεία βαίνει τελικά παράλληλα στον άξονα X).

Η δυαδική λογιστική παλινδρόμηση αποτελεί μια διωνυμική εξίσωση στην οποία η μεταβλητή απόκρισης Y είναι το τυχαίο αποτέλεσμα εμφάνισης μιας από δύο δυνητικές εκβάσεις του τύπου επιτυχία ή αποτυχία όπως π.χ. είναι το αποτέλεσμα της ρίψης ενός νομίσματος δύο διαφορετικών όψεων (κορώνα-γράμματα), η ρίψη ενός ζαριού όπου το αποτέλεσμα εμφάνισης του αριθμού 6 θεωρείται επιτυχία και των λοιπών αριθμών αποτυχία, η θετική ψήφος εκλογής ενός πολιτικού εκπροσώπου κτλ.

Η δίτιμη λογιστική παλινδρόμηση έχει τη μορφή:

$$f(z) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$$

όπου z είναι η μεταβλητή εισόδου και f(z) το αποτέλεσμα αυτής. Στα πλεονεκτήματα της εξίσωσης συγκαταλέγεται και το γεγονός ότι η μεταβλητή εισόδου λαμβάνει θετικές και αρνητικές τιμές ενώ το αποτέλεσμα αυτής f(z) περιορίζεται σε εύρος τιμών μεταξύ 0 και 1. Αναλυτικότερα, η μεταβλητή z εκπροσωπεί τη δράση μιας ομάδας ανεξάρτητων μεταβλητών ενώ η f(z) προσδιορίζει την πιθανότητα ενός συγκεκριμένου αποτελέσματος λόγω της δράσης της ομάδας αυτής. Η μεταβλητή z (λογιστική) εκφράζει επίσης το μέτρο της ολικής συνεισφοράς όλων των συμμετεχουσών ανεξάρτητων μεταβλητών στο μοντέλο και ορίζεται ως

$$z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

όπου β_0 είναι το ύψος της κλίσης της γραμμής παλινδρόμησης και ισούται με την τιμή z όταν οι τιμές όλων των ανεξάρτητων μεταβλητών ισούνται με 0, ενώ β_i είναι οι συντελεστές παλινδρόμησης καθένας των οποίων εκφράζει το μέγεθος συνεισφοράς της αντίστοιχης μεταβλητής. Θετική τιμή του συντελεστή δηλώνει ότι η

επεξηγηματική μεταβλητή αυξάνει την πιθανότητα της επιτυχημένης έκβασης (να συμβεί δηλαδή το γεγονός), αρνητική τιμή σημαίνει ότι η μεταβλητή μειώνει την πιθανότητα αυτής της έκβασης. Υψηλή τιμή του συντελεστή σημαίνει ότι η ανεξάρτητη μεταβλητή επηρεάζει πολύ ισχυρά την πιθανότητα να συμβεί το γεγονός ή μη, ενώ χαμηλή τιμή δηλώνει μικρή επίδραση της ανεξάρτητης μεταβλητής στην πιθανότητα εμφάνισης της ανάλογης έκβασης, ανάλογα πάντα και με την κλίμακα μέτρησης της ανεξάρτητης μεταβλητής.

Συνοψίζοντας, η λογιστική παλινδρόμηση χρησιμεύει στην περιγραφή της σχέσης που αναπτύσσεται μεταξύ μιας ή περισσότερων ανεξάρτητων μεταβλητών (π.χ. ηλικία, φύλο, τοξική συγκέντρωση ουσίας) και μιας δυαδικής μεταβλητής απόκρισης εκφρασμένης ως πιθανότητα δυνάμενη να πάρει μία από δύο τιμές, όπως π.χ. θετική (1) αρνητική (0), παρόν ενδεχόμενο (1) απόν ενδεχόμενο (0), επιζών (1) θανάτων (0), αρεστός (1) δυσάρεστος (0).

Η φύση των ανεξάρτητων μεταβλητών εισόδου στην εξίσωση της πολλαπλής λογιστικής παλινδρόμησης μπορεί να είναι ποσοτική, τακτική ή ονομαστική (αδιαβάθμητη κατηγορική). Για παράδειγμα, η πιθανότητα ένα άτομο να υποστεί καρδιακό επεισόδιο σε συγκεκριμένο χρονικό διάστημα (εξαρτημένη μεταβλητή) μπορεί να προβλεφθεί από ένα πλήθος ανεξάρτητων μεταβλητών όπως είναι η ηλικία, το φύλο, ο δείκτης μάζας σώματος, η φυσική αγωγή, το ιστορικό του ασθενούς κτλ. Η ηλικία ενδέχεται να ενταχθεί στη εξίσωση είτε ως ποσοτική (με την πραγματική τιμή της) είτε ως τακτική: 15-25 ετών, 25-40, 40-60, >60. Η φυσική αγωγή ως ονομαστική μεταβλητή (άθληση ή μη), ο δείκτης μάζας σώματος (Body Mass Index - BMI) ως ποσοτική ή τακτική (<25, 25-30, >30), και το ιστορικό ως ονομαστική (ύπαρξη προδιάθεσης ή μη). Επιστήμες όπως η ιατρική, οι κοινωνικές επιστήμες και το marketing καταφεύγουν συχνά στην εφαρμογή της δυαδικής λογιστικής παλινδρόμησης.

Οι πιθανότητες που συγκλίνουν υπέρ της εμφάνισης ενός γεγονότος ή πρόθεσης εκφράζονται ως λόγος ζεύγους ακέραιων τιμών (odds) όπου ο αριθμητής προσδιορίζει την πιθανότητα που έχει το προσδοκώμενο γεγονός να συμβεί και ο παρονομαστής την πιθανότητα να μη συμβεί. Έτσι, αν p είναι η πιθανότητα να εμφανιστεί το γεγονός και $1-p$ η πιθανότητα να μη συμβεί τότε ο λόγος των

πιθανοτήτων θα είναι $p/(1-p)$. Για παράδειγμα, η πιθανότητα να ανασυρθεί μια κάρτα σπαθί από μια τράπουλα 52 φύλλων είναι 25% δηλαδή μία στις τέσσερις ή αριθμητικά $13/52=1:4$ ή και $1/4$, άρα το odd είναι $p/(1-p) = (1/4)/(3/4) = 1/3$. Με ανάλογο τρόπο εκφράζεται και η πιθανότητα μη εμφάνισης μιας κάρτας σπαθί η οποία ισούται με 3:1, αντιστρέφοντας απλώς τους όρους του κλάσματος, $(1-p)/p$.

Η παραπάνω σχέση (logit) κάλλιστα μπορεί να ενσωματωθεί στο μοντέλο της παλινδρόμησης σε λογαριθμική μορφή ως εξής,

$$\text{logit}(p) = \log_e \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Οι συντελεστές της παλινδρόμησης υπολογίζονται με τη βοήθεια της εκτίμησης της **μέγιστης πιθανοφάνειας (Maximum Likelihood Estimate – MLE)**, ως οι τιμές που μεγιστοποιούν

$$L = \prod_{i=1}^n f(x_i, \theta)$$

ή προτιμότερο με τη λογαριθμική έκδοση αυτής,

$$L = \sum_{i=1}^n \log_e f(x_i, \theta)$$

όπου θ είναι μια παράμετρος της μεταβλητής η οποία μπορεί να μεταβάλλεται ελεύθερα. Η προβλεπόμενη τιμή για κάθε παρατήρηση θα ισούται με

$$\hat{l} = \frac{1}{n} \log_e L$$

Η συνάρτηση της πιθανοφάνειας έκβασης ενός γεγονότος (likelihood) δείχνει πόσο κατάλληλα ένα παρατηρούμενο δείγμα περιγράφεται από κάποιες τιμές παραμέτρων π.χ. μέσος όρος, τυπική απόκλιση. Άρα, η μεγιστοποίηση της συνάρτησης της πιθανότητας έκβασης καθορίζει τις παραμέτρους εκείνες που είναι οι πλέον ικανές να

παράγουν τα παρατηρούμενα στοιχεία. Από άποψη στατιστικής βαρύτητας, η MLE προτείνεται για εφαρμογές σε μεγάλα δείγματα καθόσον είναι ευέλικτη, προσαρμόζεται εύκολα στην παραγωγή πολλών διαφορετικού τύπου μοντέλων, το χειρισμό διαφορετικής φύσης στοιχείων και περιέχει ακριβέστερες μετρήσεις.

Η αξιοπιστία των αποτελεσμάτων της λογιστικής παλινδρόμησης επηρεάζεται κατά πολύ από το δειγματοληπτικό μέγεθος της έρευνας. Ένας χρυσός κανόνας υπαγορεύει την αντιστοιχία του αριθμού των επιθυμητών εκβάσεων προς τον αριθμό των ανεξάρτητων μεταβλητών να προσδιορίζεται από τη σχέση 10:1. Εάν υπάρχουν ονομαστικές ανεξάρτητες μεταβλητές, όπως, για παράδειγμα, διχοτομικές, ο παραπάνω κανόνας θα ισχύει για το μέγεθος των παρατηρήσεων της ολιγοπληθέστερης κατηγορίας.

Για δίτιμες εξαρτημένες μεταβλητές, η άριστη άμεση πρόβλεψη της συμμετοχής μιας μεταβλητής ως μέλους σε ομάδα με τη μέθοδο της διακριτικής ανάλυσης επιβάλλει την ύπαρξη της πολυμεταβλητής κανονικότητας των ανεξάρτητων μεταβλητών αφενός και την ισότητα των διακυμάνσεων-συνδιακυμάνσεων (ομοιογένεια) στις δύο ομάδες, προϋποθέσεις όχι απαραίτητες κατά τη δίτιμη λογιστική παλινδρόμηση. Συμπερασματικά, για την εφαρμογή του υποδείγματος της δίτιμης λογιστικής παλινδρόμησης είναι αναγκαίες πολύ λιγότερες προϋποθέσεις από αυτές που απαιτεί η διακριτική ανάλυση. Μάλιστα, ακόμη κι αν ικανοποιούνται όλες οι προϋποθέσεις για την εφαρμογή της διακριτικής ανάλυσης, η δίτιμη λογιστική παλινδρόμηση λειτουργεί εξαιρετικά καλά, με την απαραίτητη, βέβαια, προϋπόθεση ότι το μέγεθος των παρατηρήσεων είναι τουλάχιστον 10-20 ανά ανεξάρτητη μεταβλητή.

1.3 Πολλαπλή Διωνυμική Παλινδρόμηση

Αποτελεί μέρος κατηγορικών στατιστικών μοντέλων γνωστών ως Γενικευμένα Γραμμικά μοντέλα (McCullagh & Nelder, 1989), τα οποία περιλαμβάνουν τη γνωστή κλασική παλινδρόμηση, την ανάλυση διακύμανσης και συνδιακύμανσης και τη λογαριθμογραμμική παλινδρόμηση και άλλα.

Η μέθοδος αυτή επιτρέπει την πρόβλεψη των τιμών εξαρτημένης διμερούς μεταβλητής μορφής από ένα πλήθος ανεξάρτητων μεταβλητών, οι οποίες μπορεί να

είναι ποσοτικές, διχοτομικές ή πολυμερείς ή και συνδυασμοί αυτών. Αντίποδας της λογιστικής παλινδρόμησης είναι η διακριτική ανάλυση κατηγοριών με τη διαφορά ότι στη δεύτερη συμμετέχουν μόνο ποσοτικές ανεξάρτητες μεταβλητές.

Η μεταβλητή απόκρισης στην αλγεβρική της έκδοση λαμβάνει την τιμή 1 με πιθανότητα επιτυχίας p και την τιμή 0 με πιθανότητα αποτυχίας $1-p$ και καλείται δυαδική (Binary) ή δυωνυμική (Binomial) ή μεταβλητή του Bernoulli.

Λόγω της φύσης των συμμετεχουσών μεταβλητών, απουσιάζουν οι προϋποθέσεις της ομαλής κατανομής των τιμών και της ομοιογένειας των διακυμάνσεών τους, η δε έλλειψη της γραμμικότητας μεταξύ της Y και των ανεξάρτητων μεταβλητών βελτιώνεται με τη χρήση της λογαριθμικής εξίσωσης ως εξής,

$$p = \frac{e^z}{1 + e^z}$$

όπου

$$z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Ακολούθως χρησιμοποιείται η σχέση

$$\log_e \left(\frac{p}{1-p} \right) = z$$

ή λογαριθμώντας προκύπτει

$$\left(\frac{p}{1-p} \right) = e^z$$

Σημαντική πληροφόρηση για τις ιδιότητες των δυνωμικών μοντέλων περιγράφεται από τους Collett (2003) και Cox & Snell (1989).

Χαρακτηριστικά της εξίσωσης διωνυμικής παλινδρόμησης

Τα λογαριθμικά μοντέλα λόγου (συμπληρωματικών) πιθανοτήτων ή λογιστικά μοντέλα (logit models) χρησιμοποιούνται όταν η εξαρτημένη μεταβλητή είναι δυαδική. Η λογιστική παλινδρόμηση είναι μη γραμμικής μορφής γιατί αναγκάζει τις προβλεπόμενες τιμές να κυμαίνονται μεταξύ 0 και 1.

Τα λογιστικά μοντέλα κυμαίνονται εκτιμών την πιθανότητα της εξαρτημένης μεταβλητής να λαμβάνει την τιμή 1 ($Y=1$), δηλαδή την πιθανότητα ότι κάποιο γεγονός συμβαίνει. Τα μοντέλα αυτά υπακούουν στη συνθήκη πιθανότητας εμφάνισης Pr,

$$Pr (Y = 1 | X_1, X_2, \dots, X_k) = \frac{1}{1 + e^{-z}}$$

Τόσο τα logit όσο και τα probit μοντέλα παρέχουν παρόμοια αποτελέσματα, διαφέρουν μόνο ως προς την κατανομή των στοιχείων. Τα logit ακολουθούν την αθροιστική τυπική λογιστική κατανομή (F) και τα probit την αθροιστική κανονική κατανομή (Φ).

Η λέξη logit προέρχεται συγκοπτόμενη από τη φράση logarithmic unit (λογαριθμομονάδα) σε ακολουθία με την πρωταρχικά αποδοθείσα ονοματολογία της λέξης probit (probability unit - πιθανομονάδα) και έχει την έννοια της νεπέρειας

$$\text{logit}(p) = \log_e \left(\frac{p}{1-p} \right) = \log_e p - \log_e (1-p)$$

(φυσικής) λογαρίθμησης ενός αριθμού p (\log_e ή \ln) που εκπροσωπεί πιθανότητα (αναλογία) και άρα τιμές μεταξύ 0 και 1:

Αν p δηλώνει κάποια πιθανότητα τότε η σχέση $p/(1-p)$ αντιστοιχεί στην επιτυχημένη πιθανότητα έκβασης (odds) και κατ'αντιστοιχία ο λογάριθμος της p στον λογάριθμο της επιτυχημένης πιθανότητας. Με τον ίδιο συλλογισμό, η διαφορά μεταξύ των λογαρίθμων δυο ευνοϊκών πιθανοτήτων p_1 και p_2 αποτελεί και το λογάριθμο του λόγου των ευνοϊκών πιθανοτήτων R σύμφωνα με την ακολουθία των εξισώσεων:

$$\log_e R = \log_e \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \log_e \frac{p_1}{1-p_1} - \log_e \frac{p_2}{(1-p_2)} = \text{logit}(p_1) - \text{logit}(p_2)$$

Πιθανότητα (επιτυχημένης) έκβασης

Ονομάζεται και προβλεπόμενη πιθανότητα p . Αν οι αποκρίσεις αποτιμώνται ως 0 (αποτυχία) και 1 (επιτυχία), τότε p_j είναι η πιθανότητα όπου η i ανεξάρτητη μεταβλητή (ποσοτική ή κατηγορική) δίνει απόκριση 1,

$$p_j = \frac{e^z}{1 + e^z}$$

όπου

$$z = \beta_0 + \beta_1 X_{j1} + \beta_2 X_{j2} + \dots + \beta_k X_{jk}$$

όπου β_0 είναι σταθερή παράμετρος, β_i συντελεστές και X_{ji} η i προβλεπτική μεταβλητή με $j=2$ κατηγορίες.

Λογαριθμική πιθανότητα έκβασης

Εφαρμόζεται για την άριστη εκτίμηση των συντελεστών της παλινδρόμησης και επίσης για τη σύγκριση δύο μοντέλων που διαφέρουν ως προς το σύνολο των ανεξάρτητων μεταβλητών σε καθένα από αυτά

$$L_{(\beta)} = \sum_j [y_j \log_e p_j + (m_j - y_j) \log_e (1 - p_j)]$$

όπου p_j = πιθανότητα επιτυχημένης έκβασης, y_j = απόκριση, m_j = αριθμός προσπαθειών ή παρατηρήσεων σχετιζόμενων με την j ανεξάρτητη μεταβλητή. Αν τα στοιχεία περιέχουν μια παρατήρηση ανά ανεξάρτητη μεταβλητή τότε $m_j=1$.

Συντελεστές β_i

Ο εκτιμώμενος συντελεστής για κάθε ανεξάρτητη μεταβλητή εκφράζει τη μεταβολή του λογαρίθμου του λόγου $p/(1-p)$ για κάθε μονάδα μεταβολής της αντίστοιχης ανεξάρτητης μεταβλητής, ενώ οι λοιπές παραμένουν σταθερές (το αποτέλεσμα τους είναι υπό έλεγχο). Για να βρεθεί η τιμή του β η οποία μεγιστοποιεί τη λογαριθμική πιθανότητα έκβασης $L_{(\beta)}$, η τελευταία διαφορίζεται ως προς β_0 και β_i έτσι ώστε να προκύπτει,

$$\sum_j (y_j - m_j p_j) = 0$$

$$\sum_j X_{ji} (y_j - m_j p_j) = 0$$

Οι εξισώσεις αυτές επειδή δεν είναι γραμμικές, χρησιμοποιούνται αριθμητικές, επαναληπτικές μεθόδους για την επίλυσή τους.

Λόγος των πιθανοτήτων έκβασης (odds ratio)

Ο λόγος των ευνοϊκών πιθανοτήτων έκβασης (**OR**) δυο μεταβλητών με δυο κατηγορίες εκάστη μπορεί να εκφραστεί με την κατάρτιση του πίνακα ενδεχομένων των δυο μεταβλητών στον οποίο ο λόγος OR εκτιμά τη σχέση που αναπτύσσεται μεταξύ ενός αιτίου (π.χ. έκθεση σε ακτινοβολία) και σε μια έκβαση αποτελέσματος (επιβίωση):

		Έκβαση αποτελέσματος	
		+	-
Αίτιο πρόκλησης	+	a	b
	-	c	d

Ειδικότερα, ο λόγος OR εκτιμά την πιθανότητα εκείνη που προκαλεί ένα γεγονός όταν αυτό συμβεί (ευνοϊκή έκβαση) προς την πιθανότητα να μην συμβεί. Υπό την έννοια αυτή, ο λόγος OR χρησιμοποιείται συχνά στην ιατρική επιστήμη, όπως μελετώντας για παράδειγμα την πρόκληση πυρετού (ευνοϊκή έκβαση) ή μη λόγω της εμφάνισης μιας ασθένειας.

Ο λόγος OR προκύπτει από τη σχέση:

$$OR = \frac{a \cdot d}{b \cdot c}$$

στην οποία τα λατινικά γράμματα αφορούν συχνότητες εμφάνισης ατόμων σε ένα επαρκές δείγμα νοσούντων με τις ακόλουθες διαπιστώσεις:

a= πρόκληση αιτίου και αποτελέσματος (+ +), δηλαδή εμπύρετα άτομα με παρούσα τη νόσο

b= πρόκληση αιτίου άνευ αποτελέσματος (+ -), δηλαδή άτομα φυσιολογικά παρά την εμφάνιση της νόσου

c= απουσία αιτίου αλλά πρόκληση αποτελέσματος (- +), εμπύρετα άτομα χωρίς να νοσούν

d= απουσία αιτίου και απουσία αποτελέσματος (- -), άτομα χωρίς νόσο και πυρετό

Η τιμή του λόγου OR ερμηνεύεται με βάση τις ακόλουθες τρεις συνθήκες, εφόσον προηγουμένως έχει διαπιστωθεί στατιστική σημαντικότητα στο αποτέλεσμα ($p < 0,05$):
Όταν $OR=1$ δεν παρατηρείται πρόκληση αιτίου (νόσος) στην έκβαση ενός αποτελέσματος (πυρετός)

Όταν $OR > 1$ η πρόκληση του αιτίου σχετίζεται με υψηλή ή πιθανότητα να συμβεί ευνοϊκό αποτέλεσμα

Όταν $OR < 1$ η πρόκληση του αιτίου σχετίζεται με χαμηλή πιθανότητα εμφάνισης της ευνοϊκής έκβασης

Η τιμή του λόγου OR συνοδεύεται πάντα και από την εκτίμηση του τυπικού σφάλματος, SE:

$$SE = \sqrt{(1/a) + (1/b) + (1/c) + (1/d)}$$

καθώς επίσης και από τα 95% όρια εμπιστοσύνης μέσα στα οποία ο λόγος OR κυμαίνεται στατιστικώς σημαντικά:

$$e^{\log_e(OR) \pm 1,96 \cdot SE}$$

Η εκθετική σχέση χρησιμοποιείται για την ερμηνεία του συντελεστή β : ο λόγος των πιθανοτήτων επιτυχημένης έκβασης αυξάνεται αναλογικά με την τιμή e^{β_1} για κάθε μονάδα αύξησης της X, στην ουσία δηλαδή ο λόγος ισούται με e^{β_1} . Για παράδειγμα, αν $\beta_1=0,85$ τότε $\theta=e^{0,85}=2,34$ που σημαίνει ότι αναμένεται αύξηση κατά 134% ($2,34-1=1,34$) του λόγου επιτυχημένης έκβασης για κάθε αύξηση της X κατά 1 μονάδα.

Τυπικό σφάλμα των συντελεστών

Υπολογίζεται το ασυμπτωτικό τυπικό σφάλμα το οποίο όσο μικρότερη τιμή παρέχει τόσο ακριβέστερη θεωρείται η εκτίμηση. Η στατιστική σημαντικότητα των συντελεστών των ανεξάρτητων μεταβλητών ελέγχεται με δύο κριτήρια:

α) Το κριτήριο του Wald,

$$z = \frac{\beta_i}{SE}$$

Η τιμή z συγκρίνεται με την τιμή 1,96 ή υψούμενη στο τετράγωνο με τη θεωρητική τιμή χ^2 (3,841) για το επίπεδο στατιστικής σημαντικότητας 0.05.

Τιμές του z μεγαλύτερες από 1,96 δείχνουν στατιστική σημαντικότητα της μεταβλητής. Τα 95% όρια εμπιστοσύνης κάθε συντελεστή β_i εξάγονται ως $\beta_i \pm Z_{0,05/2} \cdot SE$ και τα αντίστοιχα όρια εμπιστοσύνης του λόγου επιτυχημένης έκβασης υπολογίζονται αντιλογαριθμίζοντας το ανώτερο και κατώτερο της παραπάνω σχέσης. Εντός του εύρους των ορίων εμπιστοσύνης, ο λόγος των πιθανοτήτων αντιπροσωπεύεται πλήρως και ισοδύναμα από οποιαδήποτε τιμή.

Το κριτήριο του Wald προκαλεί διεύρυνση του τυπικού σφάλματος όταν οι συγκρινόμενοι συντελεστές έχουν υψηλή τιμή, μία ιδιότητα καθόλου επιθυμητή, διότι οδηγεί σε πολύ μικρή τιμή του στατιστικού Wald και στην λανθασμένα αποδοχή της σημαντικότητας του εξεταζόμενου συντελεστή (Hauck and Donner 1977). Στις περιπτώσεις αυτές, είναι προτιμότερη η απόπειρα ανάπτυξης κάποιου υποδείγματος με και χωρίς τη συγκεκριμένη - με υψηλό συντελεστή - μεταβλητή και ο έλεγχος της υπόθεσης να στηρίζεται στη μεταβολή -2LL, δηλαδή του λογάριθμου πιθανοφάνειας όπως περιγράφεται παρακάτω.

β) Το κριτήριο του λόγου ή λογάριθμου πιθανοφάνειας -2LL (Likelihood ratio statistic), το οποίο ελέγχει ένα μικρότερο μοντέλο S με s συντελεστές και πιθανοφάνεια L_s προς ένα μεγαλύτερο μοντέλο L με l συντελεστές (συνήθως ένα παραπάνω) και πιθανοφάνεια L_l και με τον περιορισμό ότι οι παράμετροι s αποτελούν μέρος του συνόλου των παραμέτρων l :

$$-2 \log_e \left(\frac{L_s}{L_l} \right) = -2 [\log_e(L_s) - \log_e(L_l)] = -2(L_s - L_l)$$

Η τιμή του κριτηρίου συγκρίνεται με τη θεωρητική τιμή χ^2

Μέθοδοι επιλογής, προσαρμογής και αξιολόγησης του μοντέλου

Ένα διωνυμικό υπόδειγμα (μοντέλο) παλινδρόμησης για να θεωρείται αποδεκτό οφείλει να υπακούει σε ορισμένα κριτήρια τα οποία σχετίζονται με τον κατάλληλο αριθμό επιλογής των ανεξάρτητων μεταβλητών από ένα υποψήφιο πλήθος αυτών που έχουν καταμετρηθεί. Ελέγχεται επίσης, το ποσοστό ακρίβειας του επιλεγμένου μοντέλου, ο βαθμός καταλληλότητας του μοντέλου στα στοιχεία που το αναπαράγουν και η ποιότητα σύνδεσης των στοιχείων μεταξύ τους.

Μέθοδοι της άριστης επιλογής των ανεξάρτητων μεταβλητών

Σκοπός της λογιστικής παλινδρόμησης είναι να προβλέψει σωστά την καταγραφή της έκβασης των ατομικών παρατηρήσεων υπολογίζοντας το πλέον φειδωλό μοντέλο, δηλαδή εκείνο που περιλαμβάνει μόνο τις στατιστικά σημαντικές ανεξάρτητες μεταβλητές. Αυτό επιτυγχάνεται με τη χρήση της **βηματικής παλινδρόμησης** η οποία συνίσταται από την προοδευτική ένταξη ή προοδευτική απαλοιφή των μεταβλητών στο μοντέλο. Κυρίως συνιστάται η χρησιμοποίηση τεχνικής κατά την οποία όλες οι μεταβλητές εντάσσονται στο αρχικό μοντέλο και σταδιακά απομακρύνονται εκείνες που δεν υπακούουν στις προϋποθέσεις ενός στατιστικού κριτηρίου (t-απόρριψης ή F-απόρριψης) και παραμένουν μόνο οι στατιστικά σημαντικές.

Η λογιστική παλινδρόμηση χρησιμοποιείται κυρίως για δυο λόγους:

- α) Εξασφαλίζει ορθή πρόβλεψη των μελών καθεμιάς από τις δύο κατηγορίες, λαμβάνοντας υπόψη τη δράση πολλών ανεξάρτητων μεταβλητών. Ουσιαστικά η λογιστική παλινδρόμηση υπολογίζει την πιθανότητα επιτυχίας μιας ενέργειας p προς την πιθανότητα αποτυχίας q , ή με μετατροπή, το πηλίκο (λόγο) της ευνοϊκής πιθανότητας (επιτυχημένης έκβασης): $p/q = p/(1-p)$.

β) Παρέχει σημαντική πληροφόρηση της σχέσης και της ισχύος που αναπτύσσεται μεταξύ των ανεξάρτητων μεταβλητών, η οποία διαπιστώνεται από το μέγεθος των τιμών των συντελεστών της παλινδρόμησης. Για παράδειγμα, το κάπνισμα 10 σιγαρέτων την ημέρα (μεταβλητή X_1) προκαλεί μεγαλύτερο κίνδυνο στην ανάπτυξη καρκίνου απ' ό,τι η εργασία σε ασβεστορυχείο (μεταβλητή X_2).

Κριτήριο ελέγχου της ισότητας των κλίσεων

Το **κριτήριο G**, που είναι επίσης λογαριθμικό πηλίκιο πιθανοφάνειας (Log-likelihood ratio), ελέγχει τη διαφορά στη λογαριθμική έκβαση μεταξύ ενός μοντέλου που περιέχει μόνο τους σταθερούς συντελεστές (όρους) β_0 και εκείνου που περιέχει τους όρους του προσαρμοσμένου μοντέλου και εξετάζει αν όλοι οι συντελεστές των ανεξάρτητων μεταβλητών ισούνται με 0. Ο έλεγχος G θα πρέπει να δίνει στατιστικά σημαντικό αποτέλεσμα ($p < 0,05$) ώστε να ισχύει ότι ένας τουλάχιστον συντελεστής διαφέρει από το 0 και άρα υπάρχει ενδεχόμενο να υφίσταται κάποιο λογιστικό μοντέλο.

Ακρίβεια του επιλεγμένου μοντέλου

Ένας εναλλακτικός τρόπος ελέγχου του μοντέλου που παράγεται με την εφαρμογή της βηματικής παλινδρόμησης (πέραν των κριτηρίων F και R^2) είναι και η εκτίμηση του μοντέλου με μια ομάδα στοιχείων που δεν συμπερίλφθηκαν στο μοντέλο. Συνήθως, η διαδικασία περιλαμβάνει την εκτίμηση ενός μοντέλου με το 30-50% των παρατηρήσεων του δείγματος και χρήση του υπόλοιπου 70-50% για τον έλεγχο της ακρίβειας του μοντέλου. Η ακρίβεια μετρείται με την ταυτοποίηση των παρατηρήσεων εκείνων που ταξινομήθηκαν ορθά στο χρησιμοποιούμενο δείγμα. Υπάρχουν 4 δυνατές περιπτώσεις ταξινόμησης για κάθε παρατήρηση. Μια πρόβλεψη 0 όταν το δείγμα που χρησιμοποιείται δίνει 0 (ορθό), μια πρόβλεψη 0 όταν το δείγμα παρέχει ένδειξη 1 (σφάλμα), μια πρόβλεψη 1 όταν το δείγμα δείχνει 0 (σφάλμα) και μια πρόβλεψη 1 όταν το δείγμα προβλέπει 1 (ορθό). Το ποσοστό των ορθά ταξινομημένων παρατηρήσεων αναφέρεται ως ακρίβεια του παραγόμενου μοντέλου.

Μετρήσεις καλής προσαρμογής του μοντέλου

Οποιοδήποτε μοντέλο με k ανεξάρτητες μεταβλητές που επιλέγεται πριν από την τελική αποδοχή του θα πρέπει να ελεγχθεί για την ποιότητα της αξιοπιστίας του με την εφαρμογή ορισμένων κριτηρίων:

1. Ο έλεγχος χ^2 του Pearson βασίζεται στην εκτίμηση των υπολειμμάτων και περιγράφει το μοντέλο προσαρμόζοντάς το στα μετρηθέντα στοιχεία. Το κριτήριο μειονεκτεί όταν ο αριθμός των κατηγορικών τιμών της ανεξάρτητης μεταβλητής προσεγγίζει τον αριθμό των παρατηρήσεων, προτιμάται όμως όταν υπάρχουν επαναληπτικές παρατηρήσεις για καθεμία κατηγορία της μεταβλητής.

$$r_j = \frac{y_j - m_j \hat{p}_j}{\sqrt{m_j \hat{p}_j (1 - \hat{p}_j)}}$$

Το κριτήριο υπολογίζεται ως,

y_j = ο αριθμός των επιτυχημένων εκβάσεων για την j μεταβλητή

m_j = ο αριθμός των προσπαθειών ή επαναληπτικών μετρήσεων για την j μεταβλητή εκτιμώμενη

(προσαρμοσμένη) πιθανότητα για την j μεταβλητή.

r_j = τυποποιημένο υπόλειμμα του Pearson

Υψηλές τιμές χ^2 , όταν αντιστοιχούν σε ακριβή πιθανότητα σφάλματος p , μικρότερη της θεωρητικής τιμής 0,05, δείχνουν ότι το μοντέλο δεν περιγράφει επαρκώς τα στοιχεία.

2. Το κριτήριο απόκλισης D (Deviance) των παρατηρήσεων δείχνει, επίσης, πόσο καλά προσαρμόζεται το επιλεγμένο μοντέλο στα στοιχεία της μελέτης και ουσιαστικά μετράει την ασυμφωνία που υφίσταται μεταξύ του κρινόμενου μοντέλου και του κορεσμένου μοντέλου (με όλους τους όρους αλληλεπίδρασης ενσωματωμένους). Το κριτήριο υπόκειται στους ίδιους περιορισμούς οι οποίοι αναφέρθηκαν στον προηγούμενο έλεγχο καλής προσαρμογής. Το κριτήριο υπολογίζεται ως,

$$D = \sum_j d_j^2$$

$$d_j = \pm 2 \left[y_j \log_e \left(\frac{y_j}{m_j \hat{p}_j} \right) + (m_j - y_j) \log_e \frac{(m_j - y_j)}{m_j (1 - \hat{p}_j)} \right]$$

d_j = η υπολειμματική απόκλιση και το πρόσημο ± εξαρτάται από το αποτέλεσμα της πράξης

$$y_j - m_j \hat{p}_j$$

Όταν $y_j=0$ τότε η υπολειμματική απόκλιση ισούται με

$$d_j = [2m_j |\log_e(1 - \hat{p}_j)|]^{1/2}$$

Όταν $y_j=m_j$ τότε η εξίσωση τροποποιείται σε

$$d_j = [2m_j |\log_e \hat{p}_j|]^{1/2}$$

Υψηλές τιμές D όταν αντιστοιχούν σε ακριβή πιθανότητα σφάλματος $p < 0,05$ δείχνουν ότι το μοντέλο δεν περιγράφει επαρκώς τα στοιχεία.

3. Το κριτήριο πληροφόρησης του Bayes το οποίο χρησιμοποιείται ως ακριβής προσέγγιση ενός θεωρητικά μεγάλου μεγέθους δείγματος, έχοντας διαθέσιμο μόνο ένα μικρό δείγμα μέχρι και 40 παρατηρήσεων,

$$BIC_k = D(M_k) - df_k \log_e N$$

όπου $D(M_k)$ είναι ο συντελεστής απόκλισης του μοντέλου M_k και df_k οι βαθμοί ελευθερίας των αποκλίσεων. Όσον αυξάνεται η αρνητική τιμή BIC_k τόσο καλύτερη προσαρμογή παρουσιάζει το μοντέλο. Ως ένας καλός οδηγός λήψης ορθής κρίσης προτείνεται ο ακόλουθος:

Όταν η εκάστοτε τιμή $|BIC_1 - BIC_2|$ είναι 0-2 τότε το αποτέλεσμα της προσαρμογής κρίνεται επισφαλές. Τιμές 2-6 κρίνονται ικανοποιητικές, τιμές 6-10 παρέχουν ισχυρή μαρτυρία προσαρμογής του μοντέλου ενώ τιμές >10 εξαιρετικά ισχυρή.

4. Ο συντελεστής τύπου R^2 του McFadden γνωστός και ως δείκτης του λόγου πιθανοφανειών (Likelihood-ratio index) ο οποίος συγκρίνει ένα μοντέλο με k εισηγμένες ανεξάρτητες μεταβλητές με το μοντέλο εκείνο στο οποίο

$$R_{MF}^2 = 1 - \frac{\log_e L_M - k}{L_0}$$

απουσιάζουν οι συγκεκριμένες μεταβλητές,

όπου k ο αριθμός των μεταβλητών στο μοντέλο, L_0 η εκτίμηση πιθανοφάνειας στο μοντέλο χωρίς την ένταξη αυτών των μεταβλητών και L_M η αντίστοιχη στο πλήρες μοντέλο με όλες τις μεταβλητές εισηγμένες. Υψηλές τιμές του δείκτη R_{MF}^2 δηλώνουν ένδειξη καλής προσαρμογής του μοντέλου.

5. Ο έλεγχος των Hosmer-Lemeshow αφορά τον έλεγχο της ποσοστιαίας κατανομής των παρατηρήσεων σε ομάδες, με βάση τις προβλεπόμενες πιθανότητες. Αποτελεί εκδοχή του στατιστικού κριτηρίου χ^2 ενός πίνακα $2 \times g$ παρατηρούμενων και αναμενόμενων συχνοτήτων, όπου g είναι ο αριθμός των ομάδων και με βαθμούς ελευθερίας $g-2$ και θεωρείται κριτήριο σημαντικής βαρύτητας. Το κριτήριο εκτιμάται ως,

$$G_{HL}^2 = \sum_{k=1}^g \frac{(O_k - n'_k \bar{p}_k)^2}{n'_k \bar{p}_k (1 - \bar{p}_k)}$$

n'_k = ο αριθμός των ανεξάρτητων μεταβλητών στην k ομάδα

O_k = ο αριθμός των αποκρίσεων μεταξύ των n'_k μεταβλητών

\bar{p}_k = μέση τιμή πιθανότητας σε κάθε ομάδα.

Συνήθως χρησιμοποιούνται 10 ομάδες ίδιας κλάσης μεγέθους των πιθανοτήτων (0-0,1) αφού προηγουμένως καταταθούν οι πιθανότητες με αύξουσα τάξη. Το κριτήριο G_{HL}^2 συγκρίνεται με τη θεωρητική τιμή $\chi_8^2=15,507$ και όταν κρίνεται μη στατιστικά σημαντικό τότε αποφαινόμεστε ότι οι δύο αντιστοιχίες συχνοτήτων

σε κάθε κλάση δεν διαφέρουν στατιστικώς σημαντικά και άρα ισχύει η καλή προσαρμογή των στοιχείων στο μοντέλο. Οπτικά τουλάχιστον, ένα μοντέλο κρίνεται επαρκώς προσαρμοσμένο όταν οι τιμές των παρατηρούμενων και εκτιμώμενων συχνοτήτων πλησιάζουν αρκετά μεταξύ τους.

Διαγνωστικά κριτήρια της διωνυμικής παλινδρόμησης

Η εγκυρότητα του επιλεγμένου μοντέλου ελέγχεται με τη γραφική απεικόνιση των υπολειμμάτων που υπολογίζονται με διάφορους τρόπους και με τη διενέργεια ελέγχων καλής προσαρμογής. Έτσι, ανιχνεύονται οι ανεξάρτητες μεταβλητές αν εμφανίζουν έλλειψη καλής προσαρμογής των στοιχείων τους, αν επηρεάζουν σημαντικά τους συντελεστές παλινδρόμησης και ακόμα αν περιέχουν στοιχεία που δρουν διαστρεβλωτικά στη συμπεριφορά των υπολειμμάτων (εμφάνιση ακραίων τιμών). Υψηλές τιμές των κριτηρίων αυτών ειδοποιούν για πιθανή έλλειψη καλής προσαρμογής των στοιχείων στο μοντέλο.

1. Κριτήρια που ελέγχουν την καλή προσαρμογή των ανεξάρτητων μεταβλητών:

- α) Τα υπολείμματα Pearson (r_j) μετρούν την απόσταση μεταξύ της παρατηρούμενης και προσαρμοσμένης τιμής
- β) Τα τυποποιημένα υπολείμματα του Pearson μετρούν ό,τι και τα προηγούμενα, προέρχονται όμως ως πηλίκο των στοιχείων δια της τυπικής απόκλισης των τιμών των μεταβλητών.
- γ) Τα υπολείμματα των αποκλίσεων (d_j) με παρόμοια ελεγκτική συμπεριφορά όπως και τα προηγούμενα.
- δ) Τα υπολείμματα $\Delta\chi^2$ τα οποία προκύπτουν ως μεταβολές του κριτηρίου χ^2 του Pearson, όταν εκάστοτε αφαιρείται μια παρατήρηση από το σύνολο των ανεξάρτητων μεταβλητών και ελέγχεται η επίδρασή της στα υπόλοιπα στοιχεία.
- ε) Τα υπολείμματα ΔD τα οποία προκύπτουν ως μεταβολές του κριτηρίου απόκλισης D όταν αφαιρείται εκάστοτε μια παρατήρηση από το πλήθος των ανεξάρτητων μεταβλητών και ελέγχεται η επίδραση της αφαιρούμενης επί των υπόλοιπων στοιχείων.

2. Κριτήρια που ελέγχουν το μέγεθος της επίδρασης των ανεξάρτητων μεταβλητών στη διαμόρφωση των τιμών των συντελεστών της παλινδρόμησης:

α) Το κριτήριο $\Delta\beta$ το οποίο μετρά μεταβολές στους συντελεστές, όταν μια παρατήρηση απομακρύνεται από το μοντέλο και βασίζεται υπολογιστικά στα υπολείμματα Pearson.

β) Το τυποποιημένο κριτήριο $\Delta\beta$ το οποίο συμπεριφέρεται όπως και το προηγούμενο, αλλά προέρχεται υπολογιστικά από τα τυποποιημένα υπολείμματα Pearson.

3. Οι τιμές μόχλευσης – επιρροής (leverages) h_i οι οποίες ελέγχουν αν και πόσο ασυνήθιστες προβλέπουσες τιμές εμφανίζονται στο μοντέλο των ανεξάρτητων μεταβλητών. Η έννοια των συντελεστών επιρροής αναπτύχθηκε στο κεφάλαιο της πολλαπλής γραμμικής παλινδρόμησης και έχει ανάλογη ερμηνεία, με τη διαφορά ότι εδώ οι τιμές μόχλευσης εξαρτώνται τόσο από τις τιμές της εξαρτημένης μεταβλητής όσο και από εκείνες των ανεξάρτητων. Κυμαίνονται μεταξύ 0 και 1 και υψηλές τιμές μπορεί να σημαίνουν την ύπαρξη παρατηρήσεων με σημαντική επίδραση στην αξιοπιστία του μοντέλου. Συνήθως οι συντελεστές συγκρίνονται με την τιμή αναφοράς $3p/n$ (p =αριθμός των ανεξάρτητων όρων στο μοντέλο +1 και n =αριθμός παρατηρήσεων) και αν υπάρχουν κάποιοι μεγαλύτεροι αυτής τότε θεωρούνται ύποπτες τιμές.

Συνήθως επιλέγονται τα κριτήρια $\Delta\chi^2$ και ΔD και ελέγχονται γραφικά (ως Y) θέτοντας στον άξονα X τον αύξοντα αριθμό των παρατηρήσεων ή τις εκτιμώμενες τιμές των πιθανοτήτων. Ακραίες τιμές αντιστοιχούν στα σημεία που εμφανίζονται απομονωμένα των υπολοίπων στις άνω γωνίες των δύο γραφημάτων και, από ποσοτικής πλευράς, όσες τιμές είναι μεγαλύτερες της θεωρητικής 3,841. Με παρόμοιο γραφικό τρόπο ελέγχονται και οι συντελεστές μόχλευσης.

1.4 Πολυωνυμική Παλινδρόμηση

Επιλέγεται στις περιπτώσεις όπου η εξαρτημένη μεταβλητή περιέχει περισσότερες από δύο διαβάθμητες κατηγορίες (π.χ. περιοχές μιας επικράτειας, τυριά ποικίλης προέλευσης, φυτικά είδη ενός υγρότοπου, υποστρώματα ιζήματος κτλ.).

Η τεχνική, γνωστή και ως πολυμερής ή πολυωνυμική παλινδρόμηση, παραδέχεται ότι κάθε ανεξάρτητη μεταβλητή έχει μία μόνο τιμή για κάθε παρατήρηση και ότι η εξαρτημένη δεν μπορεί να προβλεφτεί άριστα από μία μόνο ανεξάρτητη μεταβλητή για καθεμία παρατήρηση. Οι ανεξάρτητες μεταβλητές θα πρέπει να μη συσχετίζονται ισχυρά μεταξύ τους, να υπάρχει δηλαδή έλλειψη πολυσυγγραμμικότητας. Στα μοντέλα όπου η εξαρτημένη μεταβλητή έχει τη μορφή θεμάτων πολλαπλής επιλογής (multiple choice items), με τη δυνατότητα επιλογής περισσότερων της μιας κατηγοριών τη φορά, η τεχνική προϋποθέτει ότι ισχύει η ανεξαρτησία των λοιπών μη σχετικών εναλλακτικών επιλογών.

Τα μοντέλα της πολυωνυμικής λογιστικής παλινδρόμησης εκπροσωπούνται από μία εξαρτημένη μεταβλητή με περισσότερες από δυο κατηγορίες, αποτελούν δηλαδή προέκταση της διχοτομικής παλινδρόμησης και περιγράφονται από την εξίσωση,

$$\log_e \left(\frac{p_{ij}}{p_{il}} \right) = x_i \beta_j$$

όπου $j=2,3,\dots,J$ και $i=1,2,\dots,N$ και η πιθανότητα $p_{ij}=P(Y=j|x)$ προκύπτει ως

$$p_{ij} = \frac{e^{x_i \beta_j}}{\sum_{j=1}^J e^{x_i \beta_j}}$$

όπου \mathbf{x}_i είναι το διάνυσμα των ανεξάρτητων μεταβλητών. Οι άγνωστες παράμετροι β_j υπολογίζονται με την εκτίμηση της μέγιστης πιθανοφάνειας (MLE).

Στην πολυμερή λογιστική παλινδρόμηση, μία από τις κατηγορίες της εξαρτημένης μεταβλητής επιλέγεται ως βασική ή προς σύγκριση ή αλλιώς κατηγορία αναφοράς (baseline category). Χωριστές αναλογίες πιθανοτήτων εκτιμώνται για όλες τις ανεξάρτητες μεταβλητές για καθεμία κατηγορία της εξαρτημένης, εκτός από τη βασική η οποία αποκλείεται από περαιτέρω ανάλυση. Ο εκθετικός συντελεστής β εκφράζει τη μεταβολή της επιτυχημένης πιθανότητας καθεμίας κατηγορίας συγκρινόμενη με τη βασική, όταν κάθε ανεξάρτητη μεταβλητή μεταβάλλεται κατά μια μονάδα.

Έτσι, για τις κατηγορίες $Y=0$, $Y=j$ και $Y=J$ (όπου J είναι η τελευταία κατηγορία) θα προκύψουν οι εξισώσεις,

$$P_i(Y=0) = \frac{1}{1 + \sum_{j=1}^J e^{(x_i\beta_j)}}$$

$$P_i(Y=J) = \frac{e^{(x_i\beta_J)}}{1 + \sum_{j=1}^J e^{(x_i\beta_j)}}$$

Τα πολυωνυμικά λογιστικά μοντέλα θεωρούν ότι οι πολυμερείς μετρήσεις σε κάθε διαφορετικό συνδυασμό των ανεξάρτητων μεταβλητών είναι ανεξάρτητες.

Ο μέγιστος αριθμός των ανεξάρτητων μεταβλητών θα πρέπει να είναι σημαντικά μικρότερος κατά την πολυωνυμική παλινδρόμηση συγκρινόμενος με αυτόν στη γραμμική, αφού η κατηγοριοποίηση της εξαρτημένης μεταβλητής στην πολυωνυμική παλινδρόμηση σημαίνει περιορισμό σε παροχή πλ ηροφοριών.

Για την εφαρμογή της πολυωνυμικής λογιστικής παλινδρόμησης, υποστηρίζεται η ανάγκη ύπαρξης μεγάλου μεγέθους δείγματος παρατηρήσεων (τουλάχιστον 50 παρατηρήσεις ανά ανεξάρτητη μεταβλητή) για την ακρίβεια του ελέγχου των υποθέσεων, ιδιαίτερα όταν η εξαρτημένη μεταβλητή έχει πολλές ομάδες. Αναλυτική

αναφορά στις ιδιότητες των ονομαστικών μεταβλητών δίνεται από το σύγγραμμα του Reynolds (1984).

Χαρακτηριστικά της μεθόδου

Πιθανότητα επιτυχημένης έκβασης

Αν θεωρήσουμε μια μεταβλητή απόκλισης με τρεις κατηγορίες 1, 2 και 3 και κατηγορία αναφοράς την 3 τότε οι δεσμευμένες πιθανότητες προκύπτουν ως

$$P(y=1|X) = \frac{e^{x'\beta_1}}{1 + e^{x'\beta_1} + e^{x'\beta_2}}$$

$$P(y=2|X) = \frac{e^{x'\beta_2}}{1 + e^{x'\beta_1} + e^{x'\beta_2}}$$

$$P(y=3|X) = \frac{1}{1 + e^{x'\beta_1} + e^{x'\beta_2}}$$

όπου $P_k(X) = p(y=k|X)$ για $k=1,2,3$. Κάθε πιθανότητα είναι αποτέλεσμα του ανύσματος $2(p+1)$ παραμέτρων και $\beta' = (\beta'_1, \beta'_2)$.

Το μοντέλο με τις τρεις παραπάνω κατηγορίες απόκρισης (κατηγορία αναφοράς η με αριθμό 3) περιγράφεται από την εξίσωση της λογαριθμικής πιθανότητας έκβασης:

$$L(\beta) = \sum_i^n y_{1i} g_1(X_i) - \log_e(1 + e^{g_1(X_i)} + e^{g_2(X_i)})$$

Κριτήριο ελέγχου της ισότητας των κλίσεων

Το κριτήριο G εξετάζει αν όλοι οι συντελεστές των ανεξάρτητων μεταβλητών ισούνται με 0. Όταν το κριτήριο G είναι στατιστικά σημαντικό ($p < 0,05$) τότε δηλώνει

ότι ένας τουλάχιστον συντελεστής διαφέρει από το 0 και άρα μπορεί να προσαρμοστεί κάποιο λογιστικό μοντέλο.

Συντελεστές παλινδρόμησης

Με παρούσες k κατηγορίες της εξαρτημένης μεταβλητής εκτιμώνται $k-1$ παράμετροι για κάθε ανεξάρτητη μεταβλητή. Οι δράσεις διαφέρουν ανάλογα με την κατηγορία απόκρισης συγκρινόμενες με την κατηγορία αναφοράς. Κάθε λογαριθμικός λόγος πιθανοτήτων παρέχει τις εκτιμώμενες (προσαρμοσμένες) διαφορές μιας κατηγορίας απόκρισης ως προς την κατηγορία αναφοράς. Οι παράμετροι (συντελεστές) των $k-1$ εξισώσεων αφορούν λογαριθμικούς λόγους πιθανοτήτων χρησιμοποιώντας όλα τα λοιπά ζεύγη των κατηγοριών απόκρισης με τη μέθοδο της εκτίμησης της μέγιστης πιθανότητας.

$$z = \frac{\beta_{k_i}}{SE}$$

Η σημαντικότητα των συντελεστών υπολογίζεται με βάση το κριτήριο Wald ως, και τα 95% όρια εμπιστοσύνης από την εξίσωση $\beta_{kj} \pm Z_{0,05/2} \cdot SE$.

Εναλλακτικά, και με περισσότερη αξιοπιστία, χρησιμοποιείται ο λογάριθμος του λόγου πιθανοφάνειας $-2LL$ ο οποίος ελέγχει με τη βοήθεια του κριτηρίου χ^2 την επίδραση καθεμίας ανεξάρτητης μεταβλητής στο τελικό μοντέλο και βασίζεται στη μεταβολή που επέρχεται στην τιμή $-2LL$ με την απομάκρυνση μιας εκάστοτε μεταβλητής.

Λόγος των πιθανοτήτων έκβασης

Χρησιμοποιείται για την εκτίμηση της σχέσης μεταξύ εξαρτημένης και ανεξάρτητης μεταβλητής και παράγει πάντα θετική τιμή. Λόγος ίσος με 1 λαμβάνεται ως λόγος αναφοράς δηλαδή η τιμή $\theta=1$ εκφράζει απουσία κάποιας σχέσης μεταξύ των δυο μεταβλητών X και Y . Αν $\theta > 1$ τότε οι πιθανότητες επιτυχημένης έκβασης είναι υψηλότερες να συμβούν για τη συγκεκριμένη ανεξάρτητη μεταβλητή. Για παράδειγμα, αν ένα μοντέλο περιλαμβάνει k κατηγορίες της εξαρτημένης μεταβλητής

και μία ανεξάρτητη, τότε ο λόγος των πιθανοτήτων έκβασης υποδεικνύει την επιτυχημένη πιθανότητα μιας κατηγορίας συγκρινόμενη με την κατηγορία αναφοράς.

Η παρακάτω εξίσωση περιγράφει το λόγο των πιθανοτήτων μιας εξαρτημένης κατηγορικής μεταβλητής με τρεις κατηγορίες και μιας ανεξάρτητης κατηγορικής με δύο κατηγορίες, α και β :

$$\theta = \frac{P(y=k | X = \alpha) / P(y=3 | X = \alpha)}{P(y=k | X = \beta) / P(y=3 | X = \beta)}$$

Έλεγχοι καλής προσαρμογής του μοντέλου

Στην πολωνιμική παλινδρόμηση υπολογίζονται, όπως και στις προηγούμενες τεχνικές, τα κριτήρια χ^2 του Pearson και της απόκλισης D των παρατηρήσεων (παρατηρούμενες-εκτιμώμενες) και θεωρούνται αξιόπιστες μόνον όταν υπάρχουν πολλές παρατηρήσεις για κάθε συνδυασμό των ανεξάρτητων μεταβλητών. Υψηλές τιμές χ^2 δηλώνουν μοντέλο μη επαρκώς προσαρμοσμένο στα στοιχεία της ανάλυσης.

Οι συντελεστές πολλαπλού προσδιορισμού τύπου R^2 , ή αλλιώς ψευδο-συντελεστές R^2 , αποτελούν διαφορετική προσέγγιση εκτίμησης της καλής προσαρμογής του μοντέλου παρέχοντας ερμηνεία παρόμοια με εκείνη της γραμμικής παλινδρόμησης, δηλαδή εκφράζουν το ποσοστό της διακύμανσης που επεξηγείται από τις ανεξάρτητες μεταβλητές. Εξ αυτών, ο συντελεστής R^2 των Aldrich and Nelson χρησιμοποιείται σε διμερείς και πολυμερείς λογιστικές παλινδρομήσεις αδυνατώντας πάντως να προσεγγίσει το μέγιστο της τιμής R^2 , δηλαδή τη μονάδα. Δημοφιλέστερος θεωρείται ο

$$R^2_{(C \wedge S)} = 1 - \left(\frac{L_0}{L_M} \right)^{2/n}$$

συντελεστής R^2 του McFadden. Ο έλεγχος αυτός στηρίζεται στην εκτίμηση του λογάριθμου της πιθανοφάνειας όπως επίσης και ο ψευδο-συντελεστής των Cox & Snell:

όπου n το πλήθος των παρατηρήσεων και οι παράμετροι L_M και L_0 αφορούν τις πιθανοφάνειες με παρούσες τις μεταβλητές και χωρίς αυτές αντίστοιχα. Ο δείκτης αυτός αδυνατεί επίσης να προσεγγίσει τη μονάδα και το πρόβλημα αυτό παρακάμπτεται με τη χρήση της τροποποιημένης εξίσωσης των Nagelkerke/Cragg & Uhle, ο οποίος επιτρέπει τη διακύμανση του δείκτη στο εύρος 0-1:

$$R^2 = \frac{1 - \left(\frac{L_0}{L_M}\right)^{2/n}}{1 - (L_M)^{2/n}}$$

Επισημαίνεται ότι οι ψευδοσυντελεστές προσδιορισμού R^2 θεωρούνται εσφαλμένως έλεγχοι της καλής προσαρμογής των μοντέλων αφού στην πραγματικά επιζητούν να εκτιμήσουν την ένταση σχέσης που αναπτύσσεται μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών. Ένεκα αυτής της διάστασης μερικές φορές παρατηρείται αναντιστοιχία στην εκτίμηση των αποτελεσμάτων.

Διαγνωστικά κριτήρια της πολυωνυμικής παλινδρόμησης

Η καλή προσαρμογή του μοντέλου υποβοηθείται επίσης και με την εξέταση των τυποποιημένων υπολειμμάτων t_j του Pearson της απόκλισης D , συνηθέστερα όμως με τη διαγραμματική μεταβολή των κριτηρίων $\Delta\chi^2$ και ΔD και τον αύξοντα αριθμό των παρατηρήσεων ή τις εκτιμώμενες τιμές των πιθανοτήτων, όπως περιγράφηκαν στη διωνυμική παλινδρόμηση (ενότητα 1.3).

Ταξινόμηση παρατηρήσεων

Ας υποθέσουμε ότι αξιολογήθηκε επαρκώς ένα μοντέλο πολυωνυμικής παλινδρόμησης με τρεις κατηγορίες παρούσες στην εξαρτημένη μεταβλητή ως αποτέλεσμα της δράσης δύο ανεξάρτητων μεταβλητών. Τότε, είναι εφικτή, με βάση τους συντελεστές παλινδρόμησης, η εκτίμηση της πιθανότητας που έχει ένα άτομο (παρατήρηση) να ανήκει σε κάποια από τις τρεις κατηγορίες της εξαρτημένης, εφαρμόζοντας την εξίσωση (και αντιλογαριθμώντας τις τιμές j):

$$P_1(Y=1) = \frac{e^{g_1}}{\sum_{j=1}^J e^{g_j}}$$

$$P_2(Y=2) = \frac{e^{g_2}}{\sum_{j=1}^J e^{g_j}}$$

όπου P είναι η εκτιμώμενη πιθανότητα για τις κατηγορίες 1 και 2 (η τρίτη τίθεται ως κατηγορία αναφοράς).

Η κατηγορία που εμφανίζει την υψηλότερη πιθανότητα θεωρείται να αντιπροσωπεύει το συγκεκριμένο άτομο.

Με τον τρόπο αυτό υπολογίζονται οι εκτιμώμενες πιθανότητες για όλα τα άτομα να ανήκουν σε μία από τις τρεις κατηγορίες. Γεννιέται όμως το ερώτημα πόσο ορθά έχει πραγματοποιηθεί η ταξινόμηση αυτή αν ληφθεί υπόψη ο βαθμός ταύτισης του αριθμού των προβλεπόμενων παρατηρήσεων με τις εμπειρικές στις κατηγορίες της εξαρτημένης μεταβλητής. Για το λόγο αυτό, καταρτίζεται ένας διασταυρωτός πίνακας δυο κατευθύνσεων ο οποίος περιέχει στα κελιά τις συχνότητες ορθής αντιστοίχισης ανά κατηγορία των προβλεπόμενων (στήλες) και παρατηρούμενων (σειρές). Το ποσοστό των ορθώς προβλέψιμων ανά κατηγορία υπολογίζεται επίσης είτε ατομικά σε κάθε κελί είτε και συνολικά ως αθροίσματα σειρών και στηλών. Όσο μεγαλύτερο ποσοστό ταυτίζεται τόσο πιο αξιόπιστο αναμένεται το μοντέλο πολυωνυμικής παλινδρόμησης.

Η ταξινόμηση των παρατηρήσεων αποτελεί μέτρο εκτίμησης της καλής προσαρμογής του προτεινόμενου μοντέλου, η αξιοπιστία του όμως ενδέχεται να υποβαθμίζεται στην παρουσία άνισου αριθμού παρατηρήσεων μεταξύ των κατηγοριών και συνεπώς θα πρέπει να συνεκτιμάται με τα υπόλοιπα κριτήρια αξιολόγησης της προσαρμογής του μοντέλου.

Κεφάλαιο 2. Το Διαδίκτυο και το Online Dating

2.1. Εισαγωγή

Το κεφάλαιο ξεκινά με την παρουσίαση της ιστορίας και ανάπτυξης του Διαδικτύου και κατόπιν εστιάζει στην έννοια της παραπλανητικής αυτο-παρουσίασης στο Διαδίκτυο. Εξετάζονται τα χαρακτηριστικά της ηλεκτρονικής απάτης, διαφορές που παρουσιάζουν τα δύο φύλα κατά την επιλογή διαδικτυακών συντρόφων και πώς διαμορφώνονται οι διαδικτυακές σχέσεις. Συγκρίνονται οι πραγματικές σχέσεις με τις διαδικτυακές και, τέλος, περιγράφεται η διαδικασία μετάβασης των διαδικτυακών σχέσεων σε σχέσεις εκτός σύνδεσης.

2.2. Ιστορία και ανάπτυξη του Διαδικτύου

Ο πρώτος επαναπρογραμματιζόμενος ψηφιακός υπολογιστής έφερε το όνομα ENIAC και σχεδιάστηκε την περίοδο του Β΄ Παγκοσμίου πολέμου στις ΗΠΑ. Αυτοί οι υπολογιστές, παρόλο που είχαν τη δυνατότητα να ολοκληρώνουν ένα πλήρες εύρος υπολογιστικών προβλημάτων, μπορούσαν να επεξεργάζονται μόνο μία εφαρμογή κάθε φορά, ενώ το μεγάλο μέγεθός τους προϋπέθετε την τοποθέτησή τους σε ειδικά δωμάτια με ψύξη. Οι προγραμματιστές εργάζονταν σε ξεχωριστούς χώρους και στη συνέχεια μετέφεραν τις πληροφορίες στον υπολογιστή μέσω μιας χρονοβόρας διαδικασίας. Το 1957 κατασκευάζεται η πρώτη απομακρυσμένη σύνδεση που επιτρέπει στους προγραμματιστές να συνδέονται απευθείας με τους υπολογιστές από απόσταση (Cohen-Almagor, 2011).

Κατά τη διάρκεια του ψυχρού πολέμου –συγκεκριμένα το 1958–, οι ΗΠΑ ίδρυσαν την υπηρεσία προηγμένης έρευνας για την άμυνα (DARPA), η οποία κατασκεύασε ένα δίκτυο υπολογιστών με κεντρικό υπολογιστή και συνδεδεμένους με αυτόν περιφερικούς υπολογιστές. Η εξέλιξη αυτού του δικτύου το 1966 ονομάστηκε ARPANET και έδινε τη δυνατότητα σε δύο απομακρυσμένους υπολογιστές να συνδεθούν μεταξύ τους, να ανταλλάξουν δεδομένα και να κάνουν χρήση του ηλεκτρονικού ταχυδρομείου. Την ίδια εποχή δημιουργήθηκαν και ακόμα τρία δίκτυα, τα οποία μαζί με το ARPANET αποτέλεσαν τα θεμέλια του σύγχρονου διαδικτύου. Αυτά τα δίκτυα ήταν το στρατιωτικό δίκτυο που κατασκεύασε η RAND Corporation στις ΗΠΑ, το εμπορικό δίκτυο που σχεδίασε το εθνικό εργαστήριο φυσικής στην

Αγγλία και το επιστημονικό δίκτυο CYCLADES στη Γαλλία (Campbell-Kelly & Garcia-Swartz, 2013)

Το 1973 εμφανίζεται το πρωτόκολλο Ιντερνέτ ή TCP/IP το οποίο επιτρέπει τη σύνδεση πολλών δικτύων με ανοιχτή αρχιτεκτονική, έτσι ώστε, σε περίπτωση που δημιουργηθεί κάποια βλάβη σε ένα δίκτυο, όλα τα υπόλοιπα να διατηρούν ομαλά τη σύνδεση και την επικοινωνία τους. Αρχικά, το ARPANET περιείχε το MILNET που χρησιμοποιούνταν από το στρατό και το ARPANET που απευθυνόταν στους ερευνητές και στους ακαδημαϊκούς. Από το 1986 έως το 1987 ο αριθμός των χρηστών αυξήθηκε από τους 2.000 στους 30.000, κάτι που οδήγησε στον διαχωρισμό του ARPANET σε έξι τομείς. Ο τομέας edu αφορούσε την εκπαίδευση, ο gov την κυβέρνηση, ο mil το στρατό, ο com το εμπόριο, ο org τους διάφορους οργανισμούς και ο net τους πόρους των δικτύων. Συγχρόνως, το 1986 το Εθνικό Ίδρυμα Επιστημών των ΗΠΑ (NSF) πραγματοποίησε τη διαδικτυακή σύνδεση και επικοινωνία διαφορετικών πανεπιστημίων. Με τη σύνδεση του ARPANET με το NSFNet έκανε την εμφάνιση του ο όρος διαδίκτυο ή ιντερνέτ. Το ιντερνέτ αρχικά περιέγραφε οποιοδήποτε δίκτυο χρησιμοποιούσε το πρωτόκολλο TCP/IP. Με τη δημιουργία του παγκόσμιου ιστού το 1989 στο CERN, η δημοτικότητα του ιντερνέτ αυξήθηκε ραγδαία, άρχισαν να αναπτύσσονται νέες εφαρμογές και το διαδίκτυο να αποκτά κυρίαρχο ρόλο σε όλο και περισσότερες ανθρώπινες δραστηριότητες (Leiner et al., 2009).

Πλέον χρήστες από όλη την υφήλιο μπορούν να επικοινωνούν μεταξύ τους και να ανταλλάσσουν πληροφορίες και δεδομένα μέσα από τον εικονικό κόσμο του διαδικτύου. Αρχικά, η επικοινωνία πραγματοποιούνταν μέσω chat rooms, forums και online dating sites. Το 2000 δημιουργήθηκαν και οι λεγόμενοι ιστότοποι κοινωνικής δικτύωσης (SNS), όπου ο κάθε χρήστης δημιουργεί το δικό του προσωπικό προφίλ. Μέσα από τα κοινωνικά δίκτυα, κοινότητες ανθρώπων με κοινά χαρακτηριστικά ή ενδιαφέροντα συνδέονται, ανταλλάσσουν πληροφορίες και αναπτύσσουν διαδικτυακές σχέσεις, κάποιες από τις οποίες μπορεί να έχουν την τύχη να μεταφερθούν και στον πραγματικό κόσμο. Δημοφιλή κοινωνικά δίκτυα αποτελούν το facebook, το twitter, το youtube κ.ά., τα οποία μέχρι σήμερα απαριθμούν πάνω από 1 δισεκατομμύριο χρήστες (Shin, 2010).

2.3. Παραπλανητική αυτο-παρουσίαση στο Διαδίκτυο

Μια από τις πιο σημαντικές ιδιαιτερότητες της διαδικτυακής επικοινωνίας είναι η δυνατότητα που παρέχεται στον χρήστη να παραμείνει ανώνυμος κατά τη δράση του στο διαδίκτυο. Η ανωνυμία αναφέρεται στον βαθμό που η ταυτότητα ενός χρήστη αποκαλύπτεται και γνωστοποιείται στους υπόλοιπους χρήστες. Έτσι το κάθε άτομο με την είσοδό του στο διαδίκτυο καλείται να επιλέξει αν η ταυτότητα του θα παραμείνει κρυφή, αν θα αποκαλύψει επιλεκτικά κάποια προσωπικά στοιχεία και θα αποκρύψει άλλα ή ακόμα και αν θα αποκαλύψει τον πραγματικό του εαυτό στους υπόλοιπους χρήστες (Rains, 2007).

Σε αντίθεση λοιπόν με την άμεση επικοινωνία όπου τα μη λεκτικά μέσα χρησιμοποιούνται εξίσου με τη γλώσσα, στο διαδίκτυο όπου η αλληλεπίδραση βασίζεται κυρίως στην ανταλλαγή γραπτού κειμένου, δεν υπάρχει τρόπος να αποκαλυφθούν στοιχεία της ταυτότητας του χρήστη αν ο ίδιος επιθυμεί να τα αποκρύψει. Έτσι κατά την επικοινωνία στο διαδίκτυο συχνά υπάρχουν κάποια κενά που κάνουν δύσκολη τη δημιουργία μιας ξεκάθαρης εικόνας για την ταυτότητα και την προσωπικότητα του διαδικτυακού συνομιλητή. Για να καλυφθεί αυτό το κενό, ο δέκτης συμπληρώνει με τη βοήθεια της φαντασίας του τα κομμάτια της ταυτότητας που του λείπουν, προσδίδοντας έτσι στον συνομιλητή του χαρακτηριστικά που μπορεί να μην ανταποκρίνονται στην πραγματικότητα (Jiang, Bazarova & Hancock, 2011). Με αυτόν τον τρόπο οι δύο διαδικτυακοί σύντροφοι προσθέτουν με τη βοήθεια της φαντασίας τους τα επιθυμητά χαρακτηριστικά στο συνομιλητή και δημιουργούν την ψευδαίσθηση ότι γνωρίζουν πλέον περισσότερα για τον διαδικτυακό τους σύντροφο. Αυτό έχει ως αποτέλεσμα τη ανάπτυξη μεγαλύτερου βαθμού οικειότητας ανάμεσα στους δύο χρήστες (Hancock & Dunham, 2001).

Κάτω από την ασφάλεια της διαδικτυακής ανωνυμίας οι χρήστες δεν αισθάνονται πλέον τον κίνδυνο της έκθεσης και εμφανίζονται πιο πρόθυμοι να μοιραστούν προσωπικές πληροφορίες, ακόμα και να αποκαλύψουν μυστικά που δεν θα τολμούσαν να ειπωθούν στον πραγματικό κόσμο (online disinhibition effect). Ακόμα υπάρχει η δυνατότητα συμμετοχής σε διαδικτυακές ομάδες οι οποίες παρά την ανωνυμία των επιμέρους μελών τους, ως σύνολο έχουν ξεκάθαρη ταυτότητα με την οποία μπορούν εύκολα να ταυτιστούν όλα τα ανώνυμα μέλη της. Τέλος, μέσω της ανωνυμίας είναι πιο εύκολη η συμμετοχή σε διαδικτυακές ομάδες των οποίων η

δράση θεωρείται παράνομη ή κατακριτέα από το κοινωνικό σύνολο (Douglas & McGarty, 2002).

Σύμφωνα λοιπόν με όσα εκτέθηκαν παραπάνω, αν και το διαδίκτυο προσφέρει τη δυνατότητα απόκρυψης της προσωπικής ταυτότητας, σήμερα υπάρχει μια αυξανόμενη τάση των χρηστών να δημιουργούν ατομικά προφίλ στους ιστότοπους κοινωνικής δικτύωσης, μέσα από τα οποία φανερώνουν διάφορα στοιχεία της ατομικότητάς τους με πρόθεση να τα γνωστοποιήσουν στους υπόλοιπους χρήστες. Έτσι, οι κάτοχοι αυτών των προφίλ τα χρησιμοποιούν για να αναρτούν προσωπικά στοιχεία, φωτογραφίες και βίντεο, να μοιράζονται προσωπικές εμπειρίες, σκέψεις και συναισθήματα και να αλληλεπιδρούν μέσω αυτών με τους υπόλοιπους χρήστες. Επιπλέον, το κάθε άτομο έχει το δικαίωμα να επιλέξει ποιοι και πόσοι χρήστες θα έχουν πρόσβαση σε κάθε προσωπική ανάρτηση που ανεβάζει στο διαδίκτυο (Amichai- Hamburger, 2013).

Καθώς λοιπόν ο χρήστης έχει την δυνατότητα να κατασκευάσει ένα ή περισσότερα ατομικά προφίλ στα διάφορα κοινωνικά δίκτυα χωρίς να υπάρχει κανένας έλεγχος για την πιστότητα των δεδομένων που αναρτά, έγκειται πλέον στην πρόθεση του κάθε χρήστη αν το προσωπικό του προφίλ θα ταυτίζεται ή θα διαφέρει από την ατομικότητά του στην πραγματική ζωή. Αν επομένως το επιθυμεί, ο χρήστης μπορεί να παραποιήσει χαρακτηριστικά της εξωτερικής του εμφάνισης ή της προσωπικότητάς του και να παραπλανήσει με αυτόν τον τρόπο τους δέκτες (Amichai- Hamburger, 2013).

Το κάθε άτομο έχει πλέον την δυνατότητα να ελέγξει την φυσική του εμφάνισή στο διαδίκτυο. Με αυτόν τον τρόπο στο σύγχρονο κόσμο όπου η εξωτερική εμφάνιση θεωρείται από τα σημαντικότερα κριτήρια με βάση τα οποία χαρακτηρίζονται οι άνθρωποι, το διαδίκτυο είναι το μέρος που ο καθένας μπορεί να απαλλαγεί από τα εξωτερικά χαρακτηριστικά που δεν είναι κολακευτικά κατά τη γνώμη του και να αποκτήσει την επιθυμητή εξωτερική ομορφιά. Σύμφωνα με τον Cialdini (1984, όπ. αναφ. στο Amichai-Hamburger, 2013), οι ελκυστικοί και εντυπωσιακοί άνθρωποι απολαμβάνουν διάφορα κοινωνικά πλεονεκτήματα, για παράδειγμα γίνονται πιο εύκολα αρεστοί από τους υπόλοιπους, προθυμοποιούνται περισσότερα άτομα να τους βοηθήσουν, ενώ γενικότερα θεωρείται ότι έχουν καλύτερες επικοινωνιακές δεξιότητες.

Η πρώτη εντύπωση τόσο στην πραγματική, πόσο μάλλον στη διαδικτυακή ζωή που το κάθε άτομο καλείται να επιλέξει ανάμεσα σε εκατομμύρια πιθανά διαδικτυακά ραντεβού, μπορεί να είναι καθοριστική για τη μελλοντική πορεία της κάθε σχέσης. Εφόσον λοιπόν η εξωτερική εμφάνιση διαδραματίζει τόσο σημαντικό ρόλο στο πως θα διαμορφωθεί αυτή η πρώτη εντύπωση, οι χρήστες καταβάλουν μεγάλη προσπάθεια για να δημιουργήσουν ένα προφίλ που δημιουργεί μια θετική πρώτη εντύπωση. Ο όρος παραπλανητική αυτό-παρουσίαση στο διαδίκτυο (deceptive self-presentation) περιγράφει τη δημιουργία ενός παραποιημένου προφίλ που στόχο έχει να παρουσιάσει μια βελτιωμένη εικόνα του χρήστη στο διαδικτυακό κόσμο (Guadagno, Okdie & Kruse, 2011).

Μελέτες που διερεύνησαν τη συμπεριφορά των χρηστών που βρίσκονται σε αναζήτηση διαδικτυακού ερωτικού συντρόφου επιβεβαιώνουν πως ένα μέρος των χρηστών παρουσίαζε μη ρεαλιστική, παραποιημένη εκδοχή του εαυτού του (Byrum & Lenton, 2001). Ο Hitsch και οι συνεργάτες του υποστήριξαν πως οι άντρες παραποιούσαν συχνότερα στοιχεία που ανέβαζαν το κοινωνικό τους στάτους, ενώ οι γυναίκες προσπαθούσαν να φαίνονται πιο ελκυστικές και γοητευτικές (Hitsch, Hortacsu & Ariely, 2009). Άλλη έρευνα παρατηρεί πως το αντρικό φύλο ήταν πιο πιθανό να πει ψέματα σχετικά με το ύψος του, ενώ το γυναικείο φύλο σχετικά με το βάρος του (Toma, Hancock & Ellison, 2008). Γενικότερα φαίνεται πως άντρες και γυναίκες συχνά προσπαθούν να βελτιώσουν την εκδοχή της φυσικής τους εικόνας που προβάλλουν στο διαδίκτυο. Έτσι για παράδειγμα τα υπέρβαρα άτομα συχνά προτιμούν να εμφανίσουν φωτογραφίες που φαίνονται πιο αδύνατοι. Ακόμα οι χρήστες επιδιώκουν να δώσουν την εντύπωση πως ζουν μια ελκυστική, γοητευτική ζωή, γεμάτη μονάχα χαρούμενες στιγμές (Dunn & Guadagno, 2012).

2.4. Ηλεκτρονική απάτη

Χάρη στη δυνατότητα ανωνυμίας που προσφέρει το διαδικτυακό περιβάλλον, καθίσταται πολύ πιο εύκολο για τους χρήστες -σε αντίθεση με την πραγματική ζωή- η δημιουργία προφίλ με παραποιημένα κάποια χαρακτηριστικά, που καταφέρνουν να παραπλανήσουν τους υπόλοιπους χρήστες. Μέσα στο διαδίκτυο κάθε άτομο έχει την ευκαιρία να κατασκευάσει μία ή περισσότερες ψηφιακές αναπαραστάσεις (avatars) του εαυτού του. Με τον όρο avatar θα μπορούσε να περιγραφεί οποιαδήποτε ανθρώπινη αναπαράσταση, όπως για παράδειγμα τα ονόματα, τα διαδικτυακά προφίλ

ή ακόμα και οι κούκλες. Στη σημερινή εποχή συχνότερα χρησιμοποιείται για να ορίσει μια ψηφιακή αναπαράσταση σε ένα εικονικό περιβάλλον (Bailenson et al., 2008).

Η παραπλανητική αυτο-παρουσίαση, όπως αναφέρθηκε παραπάνω, έγκειται στην παραποίηση και βελτίωση κάποιων ατομικών χαρακτηριστικών. Στην περίπτωση όμως που οι χρήστες αλλοιώσουν σε τέτοιο βαθμό τα χαρακτηριστικά τους, ώστε να επιτύχουν τη δημιουργία ενός εντελώς διαφορετικού ψηφιακού εαυτού, γίνεται λόγος για πλήρη αντικατάσταση της ταυτότητας του χρήστη (identity replacement). Αυτή η νέα ταυτότητα που δεν διαθέτει κανένα κοινό γνώρισμα με το άτομο που τη δημιούργησε μπορεί να αναπαριστά κάποιο άλλο άτομο που υπάρχει στον πραγματικό ή στον ψηφιακό κόσμο είτε να είναι μια εντελώς πλασματική ταυτότητα (Segovia & Bailenson, 2013).

Υπάρχει μεγάλη διαφορά ανάμεσα στην ανωνυμία και στην αντικατάσταση ταυτότητας που λαμβάνει χώρα στο διαδίκτυο. Όταν πρόκειται για έναν ανώνυμο χρήστη, η μη αναγνώριση επιτυγχάνεται με την απουσία πληροφοριών όπως το όνομα, η ηλικία ή άλλα χαρακτηριστικά αναγνώρισης. Όμως κατά την αντικατάσταση ταυτότητας ο χρήστης αποκρύπτει σκόπιμα την ύπαρξη της πραγματικής (δεύτερης) ταυτότητάς του και δημιουργεί μια νέα ταυτότητα, αρκετά πειστική ώστε να εξαπατήσει τους υπόλοιπους χρήστες. Επιπλέον, η επιτυχημένη αντικατάσταση ταυτότητας προϋποθέτει από τον χρήστη να υιοθετήσει ενδεχομένως μια νέα εξωτερική εμφάνιση, νέα μη λεκτική συμπεριφορά, σίγουρα ένα νέο όνομα, έναν νέο τρόπο ομιλίας και άλλες αλλαγές προκειμένου να αναπαριστά με ρεαλιστικό τρόπο την ψηφιακή προσωπικότητα που έχει δημιουργήσει (Riva, 2002).

Οι λόγοι που ένα άτομο θα επιλέξει να δημιουργήσει νέες ψηφιακές ταυτότητες στο διαδίκτυο ποικίλουν, αν και οι προθέσεις αυτών των χρηστών σίγουρα δεν είναι πάντα τόσο αθώες. Όσον αφορά τα διαδικτυακά ραντεβού, η δημιουργία μιας σχέσης με ένα ψεύτικο διαδικτυακό προφίλ θα μπορούσε να οδηγήσει σε μια σχέση θύτη και θύματος, στην οποία το άτομο-θύμα αυτο-αποκαλύπτεται και αναπτύσσει συναισθηματικούς δεσμούς με ένα ψηφιακό avatar που δεν υπάρχει στην πραγματική ζωή. Σίγουρα πάντως ο εικονικός κόσμος του διαδικτύου ευνοεί την κατασκευή και την αποτελεσματική απόκρυψη μιας τόσο βαριάς εξαπάτησης. Τέλος, ένα ακόμα πιο κακόβουλο είδος avatar δεν αναπαριστά ένα απλά ανύπαρκτο

πρόσωπο, αλλά παίρνει τη μορφή ενός ατόμου που υπάρχει στον πραγματικό κόσμο. Με τον όρο χειραγώγηση ταυτότητας (identity manipulation) ορίζεται η δημιουργία μιας ταυτότητας που μιμείται κάποιο άλλο άτομο. Κατά τη δράση λοιπόν αυτής της ψηφιακής προσωπικότητας στο διαδίκτυο, οι υπόλοιποι χρήστες θα εξαπατώνται και θα αποδίδουν αυτές τις ενέργειες στο πρόσωπο που μιμείται η αναπαράσταση (Segovia & Bailenson, 2013).

2.5. Διαφορές φύλου στην επιλογή συντρόφων

Μέσα από τον εικονικό κόσμο του διαδικτύου, χρησιμοποιώντας τους ιστότοπους γνωριμιών και τα κοινωνικά δίκτυα, άντρες και γυναίκες βρίσκονται σε αναζήτηση πιθανών ερωτικών συντρόφων. Αυτοί οι διαδικτυακοί σύντροφοι είτε θα καταφέρουν να συναντηθούν και στον πραγματικό κόσμο σε μια άμεση πρόσωπο με πρόσωπο επαφή είτε θα παραμείνουν αιώνια στη σφαίρα του διαδικτύου. Φαίνεται πως τόσο οι άντρες όσο και οι γυναίκες αναζητούν στον ιστό συντρόφους που είναι εύστροφοι, κοινωνικοί, ευγενικοί και έμπιστοι (Botwin, Bus & Shackelfold, 1997, όπ. αναφ. στο Guadagno et al., 2011).

Ωστόσο, τα σύγχρονα ερευνητικά δεδομένα αποκαλύπτουν διαφορές σε συνάρτηση με το φύλο που προκύπτουν κατά την επιλογή και την προσέγγιση των υποψήφιων διαδικτυακών συντρόφων. Όσον αφορά το γυναικείο φύλο, σε έρευνα που διερευνά την προσέγγιση συντρόφων μέσω των ιστότοπων γνωριμιών, φάνηκε πως οι γυναίκες ήταν πιο επιλεκτικές και τελικά προσπαθούσαν να προσελκύσουν μικρότερο αριθμό ερωτικών συντρόφων. Έτσι, το ποσοστό των ανδρών (57%) που δεν είχε λάβει κανένα ερωτικό μήνυμα ήταν διπλάσιο σε σύγκριση με το αντίστοιχο ποσοστό των γυναικών (23%). Επιπλέον, σχετικά με τους παράγοντες με βάση τους οποίους οι χρήστες επιλέγουν τους υποψήφιους συντρόφους, σημαντικό κριτήριο για τις γυναίκες ήταν η οικονομική κατάσταση του πιθανού συντρόφου καθώς το πιο έντονο ενδιαφέρον εκδηλώθηκε για τους πιο πλούσιους άντρες (Hitsch, Hortacsu & Ariely, 2009). Από την άλλη, για τον αντρικό πληθυσμό πιο σημαντικός παράγοντας επιλογής φαίνεται να είναι η εξωτερική εμφάνιση, καθώς σύμφωνα με ερευνητικά δεδομένα οι πιο όμορφες γυναίκες λαμβάνουν τα περισσότερα ερωτικά μηνύματα. Ακόμα, παρατηρήθηκε πως το μικρότερο ερωτικό ενδιαφέρον υπάρχει προς τους κοντούς άντρες και τις υπέρβαρες γυναίκες (Toma, Hancock & Ellison, 2008).

Συγχρόνως, έγιναν προσπάθειες προσδιορισμού των ιδιαίτερων χαρακτηριστικών της προσωπικότητας που δείχνουν να προσελκύουν περισσότερους συντρόφους μέσα στους ιστότοπους γνωριμιών. Τα ερευνητικά δεδομένα υποστηρίζουν πως οι γυναίκες προτιμούσαν να προσεγγίζουν άντρες οι οποίοι ήταν υποστηρικτικοί και βοηθητικοί και ταυτόχρονα διέθεταν κυρίαρχα αρρενωπά χαρακτηριστικά. Γενικότερα το γυναικείο φύλο αφιερώνει περισσότερο χρόνο και ήταν πιο αποτελεσματικό στην αναγνώριση των διαφορετικών στοιχείων της προσωπικότητας των πιθανών συντρόφων και στην επιλογή με βάση αυτά (Jensen-Campbell, Graziano & West, 1996).

Φαίνεται πάντως πως οι χρήστες αντιλαμβάνονται ποια είναι τα χαρακτηριστικά που θα τους κάνουν πιο δημοφιλείς και ελκυστικούς στα κοινωνικά δίκτυα και προσπαθούν να διαμορφώσουν τα προσωπικά τους προφίλ επιδεικνύοντας τα συγκεκριμένα χαρακτηριστικά, προκειμένου να προσελκύσουν περισσότερους ερωτικούς συντρόφους. Στη προσπάθειά τους λοιπόν να δημιουργήσουν ένα ελκυστικό προφίλ, οι άντρες φαίνονται πιο επιρρεπείς στην παραποίηση χαρακτηριστικών που αφορούν την προσωπικότητα ή την εξωτερική τους εμφάνιση προκειμένου να γίνουν πιο αρεστοί στο αντίθετο φύλο. Επιπλέον, σύμφωνα με τα ερευνητικά αποτελέσματα, στις περιπτώσεις που δεν υπήρχε πιθανότητα συνάντησης πρόσωπο με πρόσωπο ή επικοινωνία με βίντεο ήταν ακόμα πιο έντονα τα φαινόμενα παραπλανητικής αυτο-παρουσίασης από τη μεριά των αντρών (Guadagno, Okdie & Kruse, 2011).

2.6. Διαμόρφωση διαδικτυακών σχέσεων

Για να διερευνήσουμε τον τρόπο με τον οποίο διαμορφώνονται οι διαδικτυακές σχέσεις θα πρέπει πρώτα να κατανοήσουμε τον τρόπο με τον οποίο γίνεται η επιλογή του συντρόφου καθώς και τα μέσα τα οποία διαθέτει το διαδίκτυο για την εξέλιξη των σχέσεων αυτών. Μέσα από τους ιστότοπους γνωριμιών οι χρήστες απολαμβάνουν μια σειρά παροχών. Πρώτον, με τη δημιουργία του προσωπικού του προφίλ ο κάθε χρήστης μπορεί να έχει πρόσβαση στα προσωπικά προφίλ των πιθανών ερωτικών συντρόφων. Έτσι έχει πρόσβαση σε φωτογραφίες και προσωπικές πληροφορίες και αναρτήσεις των χρηστών με βάση τις οποίες καλείται να τους αξιολογήσει και στη συνέχεια να επιλέξει αυτούς με τους οποίους θα επιδιώξει διαδικτυακή επαφή. Στη συνέχεια αν υπάρξει η αντίστοιχη ανταπόκριση θα έχει τη δυνατότητα να

αλληλοεπιδράσει με τον διαδικτυακό του σύντροφο με τα μέσα που παρέχει ο κάθε ιστότοπος (Finkel et al., 2012).

Όμως ο αναρίθμητος αριθμός διαφορετικών χρηστών που απαρτίζει τους ιστότοπους γνωριμιών θα καθιστούσε ιδιαίτερα χρονοβόρα την αναζήτηση και την εύρεση των επιθυμητών διαδικτυακών συντρόφων. Για αυτό τον λόγο οι ιστότοποι γνωριμιών με τη χρήση ενός μαθηματικού αλγορίθμου επιτυγχάνουν το λεγόμενο 'ταίριασμα', προβλέπουν δηλαδή με ποια προφίλ έχει περισσότερες πιθανότητες να ταιριάζει ο κάθε χρήστης. Τα συγκεκριμένα προφίλ θα εμφανιστούν πρώτα κατά την αναζήτηση, προκειμένου να εξοικονομήσουν πολύτιμο χρόνο στον χρήστη. Τελικά, η επιλογή του συντρόφου θα γίνει έπειτα από την αναζήτηση πληροφοριών και τη σύγκριση μεταξύ των διαφορετικών προφίλ που έχουν επιλεγεί ξεχωριστά για τον κάθε χρήστη (Amichai- Hamburger, 2013).

Αφού λοιπόν ολοκληρωθεί η επιλογή του πιθανού ερωτικού συντρόφου και υπάρχει η αντίστοιχη ανταπόκριση, οι δύο χρήστες θα έχουν τη δυνατότητα να γνωριστούν καλύτερα μέσα από τις διάφορες μορφές της κοινωνικής δικτύωσης. Το διαδίκτυο παρέχει τη δυνατότητα τόσο για σύγχρονη όσο και για ασύγχρονη επικοινωνία. Κατά τη σύγχρονη επικοινωνία, οι δύο χρήστες βρίσκονται ταυτόχρονα συνδεδεμένοι στον ιστό και επικοινωνούν μέσω μηνυμάτων και βιντεοκλήσεων σε πραγματικό χρόνο. Αντιθέτως, η ασύγχρονη επικοινωνία επιτυγχάνεται με την ανταλλαγή μηνυμάτων κατά την οποία γίνεται η αποστολή και σε δεύτερο χρόνο, ο οποίος καθορίζεται από τον παραλήπτη, συμβαίνει η παραλαβή και η απάντηση στο μήνυμα (Donn & Sherman, 2002).

2.7. Πρόσωπο με πρόσωπο έναντι Διαδικτύου

Καθώς όλο και περισσότεροι χρήστες αρχίζουν να κάνουν χρήση του διαδικτύου με σκοπό τη σύναψη νέων και τη διατήρηση των ήδη υπάρχοντων προσωπικών τους σχέσεων, δημιουργούνται ερευνητικά ερωτήματα σχετικά με την ποιότητα της επικοινωνίας μέσω του ίντερνετ σε σύγκριση με την άμεση επικοινωνία. Ακόμα τίθεται το ερώτημα αν η κοινωνικοποίηση μέσω του διαδικτύου έχει αρνητική επίδραση στη διαμόρφωση σχέσεων στην πραγματική ζωή (Snow, 2007). Υπάρχουν θεωρίες και ερευνητικά δεδομένα που αντικρούονται μεταξύ τους, με τη μία πλευρά να υποστηρίζει την έκπτωση της επικοινωνίας μέσω του διαδικτύου και την άλλη το αντίθετο.

Σύμφωνα με τη θεωρία της κοινωνικής παρουσίας (Social Presence Theory) και τη θεωρία του κοινωνικού πλαισίου (Social Context Cues Theory) το διαδίκτυο παρέχει τη δυνατότητα για μια φτωχότερη και πιο απρόσωπη επικοινωνία, η οποία συγχρόνως δεσμεύει πολύτιμο χρόνο από τον χρήστη που θα μπορούσε να αξιοποιηθεί για τη σύναψη άμεσων επαφών. Κατά την πρώτη λοιπόν θεωρία, στη διαδικτυακή επικοινωνία είναι πολύ πιο περιορισμένα τα μη-λεκτικά και παραγλωσσικά σημεία, γεγονός που μειώνει το αίσθημα της αυτο-αντίληψης και ευνοεί την εξατομίκευση, καθιστώντας έτσι την επικοινωνία πιο απρόσωπη. Συγχρόνως, η ανωνυμία και η δυνατότητα παραπλάνησης μέσω του διαδικτύου αποτελεί ακόμα μία πρόκληση, καθώς οι χρήστες καλούνται να ξεπεράσουν τις αμφιβολίες τους και να εμπιστευτούν στα τυφλά τον συνομιλητή, ενώ αναφέρουν ότι όσο λιγότερες είναι οι προσωπικές εμπειρίες που μοιράζονται με τους χρήστες τόσο φτωχότερη είναι η ποιότητα της επικοινωνίας (Sproull & Kiesler, 1986).

Η δεύτερη θεωρία (Social Context Cues Theory) υποστηρίζει πως η ποιότητα της επικοινωνίας βρίσκεται σε άμεση συνάρτηση με τον πλούτο που διαθέτουν τα μέσα της επικοινωνίας. Έτσι, η ποιότητα εξαρτάται από τη δυνατότητα άμεσης ανταπόκρισης, την ικανότητα του μέσου να μεταδίδει διαφορετικές και ποικίλες ενδείξεις, τη δυνατότητα χρήσης της φυσικής γλώσσας και την προσωπική εστίαση που είναι δυνατό να επιτευχθεί. Με βάση τα παραπάνω κριτήρια η επικοινωνία πρόσωπο με πρόσωπο αποτελεί την πιο πλούσια και σύνθετη μορφή επικοινωνίας, ενώ η σχέση που αναπτύσσεται μεταξύ δύο υποκειμένων μέσω διαδικτύου κρίνεται φτωχότερη ακόμη και από την τηλεφωνική επικοινωνία.

Όσον αφορά την ερωτική σχέση, η ποιότητα και η σημασία της σχετίζεται άμεσα με την οικειότητα που θα αναπτυχθεί μεταξύ των δύο ατόμων. Σύμφωνα με τον Hook και τους συνεργάτες του, η οικειότητα εξαρτάται από τέσσερις επιμέρους μεταβλητές, οι οποίες είναι η αυτο-αποκάλυψη, η εμπιστοσύνη, η προσωπική αξιοπιστία, η αγάπη και η στοργή (Hook et al., 2003). Σε έρευνα που διερευνήθηκε η οικειότητα που αναπτύσσεται στις διαδικτυακές σε σύγκριση με τις εκτός σύνδεσης σχέσεις, φάνηκε πως μεγαλύτερη οικειότητα ανέπτυσαν τα άτομα που είχαν περισσότερη μη λεκτική επικοινωνία. Όπως αναφέρθηκε όμως και παραπάνω, στη διαδικτυακή επικοινωνία είναι περιορισμένα τα μη λεκτικά μέσα επικοινωνίας σε σχέση με την πρόσωπο με πρόσωπο επικοινωνία (Snow, 2007).

Σε αντίθεση υποστηρίζεται πως η ανωνυμία που προσφέρει το διαδίκτυο ευνοεί τη δημιουργία δεσμών, καθώς τα άτομα νιώθουν λιγότερο εκτεθειμένα σε σύγκριση με την άμεση επικοινωνία και είναι πιο πρόθυμα να μοιραστούν προσωπικές πληροφορίες. Μέσα από την ασφάλεια που παρέχει το αχανές δίκτυο δεν είναι λίγες οι περιπτώσεις που οι χρήστες επιλέγουν να φανερώσουν σκέψεις και πληροφορίες που αδυνατούν να εκφράσουν στις πραγματικές τους σχέσεις ίσως λόγω ντροπής και ανασφάλειας. Για αυτά τα άτομα το διαδίκτυο ενσαρκώνει το μέρος όπου μπορούν να είναι ειλικρινή, χωρίς να δέχονται την κριτική ή την απαξίωση του κοινωνικού περιγύρου και έτσι πολύ συχνά σε αυτό το περιβάλλον αναπτύσσουν πιο προσωπικούς και απελευθερωτικούς δεσμούς (Suler, 2004).

Ένας ακόμα παράγοντας που ευνοεί τη δημιουργία διαδικτυακών σχέσεων είναι πως σε αντίθεση με την επικοινωνία πρόσωπο με πρόσωπο, στο διαδίκτυο ο κάθε χρήστης έχει όσο χρόνο επιθυμεί για να σκεφτεί και να απαντήσει με στρατηγική στους διαδικτυακούς ερωτικούς του συντρόφους. Έτσι κατά αυτή την ασύγχρονη επικοινωνία μπορεί να δημιουργείται η ψευδαίσθηση πως ο συνομιλητής συμπάσχει, κατανοεί και απαντά πάντα με τον επιθυμητό τρόπο στον ερωτικό σύντροφο. Αυτή η εξιδανίκευση μπορεί να μην ανταποκρίνεται στην πραγματικότητα και ιδιαίτερα σε περίπτωση άμεσης συνάντησης να οδηγήσει στα αντίθετα αποτελέσματα (Walther, Slovacek & Tidwell, 2001).

Τέλος, ο εικονικός κόσμος του διαδικτύου, αναπαριστά ένα κόσμο της διασκέδασης όπου ο καθένας μπορεί να ενσαρκώσει οποιαδήποτε ανθρώπινη φιγούρα επιθυμεί και να αλληλοεπιδράσει με τους δισεκατομμύρια χρήστες που βρίσκονται συνδεδεμένοι στον ιστό ανταλλάσσοντας μαζί τους πληροφορίες, την πιστότητα των οποίων δεν θα μπορούσε να επαληθεύσει ποτέ. Έτσι, ο καθένας θα μπορούσε να πιστέψει σε οτιδήποτε ήθελε, να δημιουργήσει μια νέα βελτιωμένη έκδοση του εαυτού του, να ταξιδέψει σε μέρη που δεν θα επισκεφθεί ποτέ και να συναναστραφεί με κόσμο που δεν τον ξέρει και είναι στην επιλογή του αν θα τον γνωρίσει. Θα μπορούσαμε να περιγράψουμε το internet ως έναν φανταστικό εικονικό κόσμο φτιαγμένο για τα ενήλικα παιδιά, ιδιαίτερα ελκυστικό και σε θέση να αντικαταστήσει όλο και περισσότερες πτυχές της πραγματικής ζωής (Gibbs, Ellison & Heino, 2006).

2.8. Μετακίνηση Διαδικτυακών σχέσεων σε σχέσεις εκτός σύνδεσης

Τα κοινωνικά δίκτυα και οι ιστότοποι γνωριμιών ενδείκνυνται για την αναζήτηση νέων ερωτικών συντρόφων. Με τη δημιουργία του προσωπικού του προφίλ ο κάθε χρήστης μπορεί να προβάλει τις πληροφορίες που επιλέγει και να προσελκύσει τους ανάλογους πιθανούς μελλοντικούς συντρόφους. Με την ανταλλαγή ηλεκτρονικών μηνυμάτων, φωτογραφιών και βίντεο οι διαδικτυακοί σύντροφοι έχουν τη δυνατότητα να φλερτάρουν, να γνωρίσουν ο ένας τον άλλο έως και να αναπτύξουν στενότερους δεσμούς από τον ασφαλή πάντα κόσμο του διαδικτύου. Μέσω των διαδικτυακών ραντεβού μπορεί να επιδιώκεται μια περιστασιακή εικονική επαφή ή ακόμα και η αναζήτηση ενός συντρόφου ζωής. Το αν μια διαδικτυακή σχέση θα καταφέρει να μεταφερθεί στον πραγματικό κόσμο εξαρτάται από πολλούς παράγοντες που είναι δύσκολο να αποσαφηνιστούν. Από την προσπάθεια των επιστημόνων να διερευνήσουν ποιες διαδικτυακές σχέσεις είναι πιθανότερο να εξελιχθούν σε σχέσεις εκτός σύνδεσης στον πραγματικό κόσμο προέκυψαν αξιοσημείωτα ευρήματα (Boase & Wellman, 2004).

Ένας καθοριστικός παράγοντας για την έκβαση κάθε διαδικτυακής σχέσης είναι οι αρχικές προθέσεις των διαδικτυακών εραστών. Πρόσφατη μελέτη υποστηρίζει πως οι χρήστες που αποσκοπούν στην εύρεση διαδικτυακών συντρόφων οι οποίοι δυνητικά θα μεταφερθούν στην πραγματική ζωή, είναι πιο πρόθυμοι να μοιραστούν προσωπικές πληροφορίες, να αποκαλύψουν κρυφά κομμάτια της προσωπικότητάς τους, ενώ επιδιώκουν την ειλικρίνεια και τη συναισθηματική τριβή προκειμένου να δημιουργήσουν βαθύτερους και ισχυρότερους δεσμούς (Gibbs, Ellison & Heino, 2006).

Ο χρήστης που ελπίζει σε μια μελλοντική συνάντηση στον πραγματικό κόσμο, δεν προσπαθεί να παραπλανήσει τον διαδικτυακό του σύντροφο, παραποιώντας τα χαρακτηριστικά του προκειμένου να εντυπωσιάσει τον δέκτη. Αντιθέτως, θέλουν να παρουσιάσουν μια ρεαλιστική εκδοχή του εαυτού τους, αποκαλύπτοντας και τα σκοτεινά σημεία της πραγματικότητάς τους, έτσι ώστε να μην απογοητεύσουν τον πιθανό σύντροφο στην επικείμενη συνάντηση πρόσωπο με πρόσωπο. Αυτή η υπερβολική ειλικρίνεια και η αποκάλυψη και των αρνητικών χαρακτηριστικών ίσως να έχει τα αντίθετα από τα επιθυμητά αποτελέσματα. Φαίνεται πως η

αποκάλυψη τόσο των θετικών όσο και των αρνητικών χαρακτηριστικών μπορεί να έχει αρνητικό αντίκτυπο στην έκβαση της διαδικτυακής σχέσης (Walther, 1996).

Αντιθέτως, η επιλεκτική αποκάλυψη και προώθηση των θετικών χαρακτηριστικών που στόχο έχουν να εντυπωσιάσουν τον πιθανό σύντροφο φαίνεται να έχουν θετική επίδραση στη μελλοντική εξέλιξη της σχέσης. Έτσι, οι χρήστες που έχουν την ικανότητα να εντοπίζουν άτομα με κοινά γνωρίσματα, να προωθούν την προσωπικότητά τους, να χρησιμοποιούν τη γοητεία τους και τις επικοινωνιακές τους δεξιότητες για να προβάλλουν τα θετικά στοιχεία του χαρακτήρα τους θα σχηματίσουν πιο ισχυρές διαδικτυακές σχέσεις (Gleason, Green & Mc Kenna, 2002).

Τέλος, υπάρχει μια κατηγορία χρηστών που φαίνεται πως μπορεί να αναπτύξει πιο εύκολα ερωτικές σχέσεις μέσω του διαδικτύου σε σύγκριση με την πραγματική ζωή. Για τα πιο εσωστρεφή, μοναχικά και ντροπαλά άτομα που η πρόσωπο με πρόσωπο συνάντηση προκαλεί άγχος και δυσκολεύει την επικοινωνία, ο ασφαλής κόσμος του διαδικτύου προσφέρει ένα λιγότερο στρεσογόνο περιβάλλον και ευνοεί τη σύναψη ισχυρότερων δεσμών. Χαρακτηριστικά αυτά τα άτομα είναι πολύ πιο πρόθυμα να εκτεθούν και να αποκαλύψουν περισσότερα προσωπικά στοιχεία μέσα από το διαδίκτυο, με αποτέλεσμα οι διαδικτυακές τους σχέσεις να αναπτύσσονται με ταχύτερους ρυθμούς σε σχέση με τις σχέσεις εκτός σύνδεσης. Έτσι, είναι πιο εύκολο για αυτή την ομάδα χρηστών να αναπτύξει στενότερες επαφές μέσω του διαδικτύου και να προσπαθήσει να μεταφέρει μελλοντικά αυτές τις σχέσεις και στον πραγματικό κόσμο (Gleason, Green & Mc Kenna, 2002).

Κεφάλαιο 3. Μεθοδολογία έρευνας

3.1. Ερευνητικός σκοπός-ερωτήματα

Σκοπός της παρούσας έρευνας είναι να μελετήσει κατά πόσο μπορούμε να χρησιμοποιήσουμε την δίτιμη λογιστική παλινδρόμηση με διαφορετικούς τρόπους και να την συγκρίνουμε με την πολυωνυμική λογιστική παλινδρόμηση έτσι ώστε να εξακριβώσουμε αρχικά αν τα αποτελέσματα των δυο αυτών μεθόδων είναι κοινά και μας δίνουν τις ιδιές πληροφορίες για να εξάγουμε σωστά συμπεράσματα και αν κάποια απο τις δυο μεθόδους μας παρέχει καλύτερα αποτελέσματα, άρα και κατ' επέκταση και συμπεράσματα. Για τον λόγο αυτό θα χρησιμοποιήσουμε μια έρευνα που μελετά την συμπεριφορά των νέων στο Διαδίκτυο, όσον αφορά την πραγματοποίηση επαφών με άτομα που δεν έχουν γνωρίσει στην πραγματική ζωή αλλά μόνο μέσω του Διαδικτύου. Μελετάται η συχνότητα πραγματοποίησης τέτοιων επαφών αλλά και οι ομάδες των νέων που είναι πιο πιθανό να πραγματοποιήσουν αυτές τις επαφές. Σε δεύτερο επίπεδο, μελετάται η πραγματοποίηση συναντήσεων με άτομα που δεν έχουν γνωρίσει στην πραγματική ζωή αλλά μόνο μέσω του Διαδικτύου καθώς και οι ομάδες των νέων που είναι πιο πιθανό να πραγματοποιήσουν μία τέτοια συνάντηση. Επιπλέον, διερευνάται, αν αυτή η απόφαση διαφοροποιείται στις ομάδες των νέων που ήδη έχουν επαφή στο Διαδίκτυο με άτομα που δεν γνωρίζουν στην πραγματική ζωή. Τέλος, συνδυάζοντας τις αποφάσεις επαφή-συνάντηση, μελετώνται οι ομάδες των νέων οι οποίες απλώς θα μείνουν στην επαφή μέσω του Διαδικτύου με άτομα που δεν γνωρίζουν στην πραγματική ζωή, αυτές που θα πραγματοποιήσουν συνάντηση και αυτές που δεν θα πραγματοποιήσουν ούτε την επαφή. Με βάση τα ανωτέρω, διατυπώνονται παρακάτω τα ερευνητικά ερωτήματα:

- 1.** Ποιοι παράγοντες επηρεάζουν την επαφή στο Διαδίκτυο με κάποιο άτομο που δεν υπάρχει γνωριμία στο παρελθόν στην πραγματική ζωή;
- 2.** Ποιοι παράγοντες επηρεάζουν την συνάντηση με άτομο με το οποίο υπήρξε γνωριμία για πρώτη φορά στο Διαδίκτυο;
- 3.** Όσον αφορά τα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή, ποιοι παράγοντες επηρεάζουν την συνάντηση;
- 4.** Ποιοι παράγοντες επηρεάζουν την γνωριμία μέσω Διαδικτύου μέχρι την συνάντηση ;

3.2. Σχεδιασμός έρευνας

Πραγματοποιήθηκε μία ποσοτική έρευνα, πρωτογενής, μη πειραματική, με χρήση ερωτηματολογίου το οποίο περιλάμβανε ερωτήσεις κυρίως κλειστού τύπου και διχοτομικές ερωτήσεις. Ο συγκεκριμένος σχεδιασμός επιλέχτηκε καθώς οι έννοιες της απόφασης πραγματοποίησης επαφής μέσω Διαδικτύου ή συνάντησης είναι εύκολα μετρήσιμες με χρήση διχοτομικών ερωτήσεων (Ναι-Όχι), συνεπώς η ποσοτική έρευνα κρίνεται κατάλληλη καθώς σκοπός του ερευνητή είναι να μετρήσει την συχνότητα της απόφασης και όχι να διερευνήσει τις αιτίες (Σταλίκας & Κυριάκος, 2019). Επιπλέον, σύμφωνα με τα ερευνητικά ερωτήματα είναι η απαραίτητη η διερεύνηση της εξάρτησης των μεταβλητών, γεγονός που επιτυγχάνεται με χρήση της ποσοτικής μεθόδου στην οποία χρησιμοποιούνται στατιστικές τεχνικές σε αριθμητικά δεδομένα (Ρούσσος, & Τσαούσης, 2011). Τέλος, η έρευνα είναι μη πειραματική, καθώς ο ερευνητής μελετάει τις δημογραφικές ομάδες που διαφέρουν ως προς την απόφαση δημιουργίας επαφής μέσω του Διαδικτύου ή συνάντησης, χωρίς να τον ενδιαφέρει η επίδραση εξωτερικών παραγόντων (Salkind, 2010).

3.3. Δείγμα έρευνας

Πληθυσμός της έρευνας θεωρείται ο μαθητικός πληθυσμός στην Ελλάδα ηλικίας 14-18 ετών. Αναφορικά με το δείγμα, στην έρευνα συμμετείχαν 1625 άτομα, αγόρια και κορίτσια με ισόποση σχεδόν κατανομή, με την πλειοψηφία να είναι ηλικίας 14-16 ετών και άτομα που γνωρίζουν αρκετά καλά το Διαδίκτυο, χωρίς ιδιαίτερη εθιστική συμπεριφορά και χωρίς υπερβολική χρήση του SNS. Οι νέοι της έρευνας ανέφεραν πως ήταν κατά μέσο όρο 11 ετών όταν ξεκίνησαν να χρησιμοποιούν το Διαδίκτυο. Το εκπαιδευτικό επίπεδο των γονέων στις περισσότερες περιπτώσεις χαρακτηρίστηκε υψηλό.

3.4. Μεταβλητές έρευνας

Οι μεταβλητές της έρευνας είναι 9 και είναι

- 1)Φύλο (Γυναίκα, Άντρας)
- 2)Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο; (ελεύθερη επιλογή απάντησης)
- 3)Γνωρίζω πολλά πράγματα για το Διαδίκτυο (Όχι, Ναι)

- 4)Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή; (Όχι, Ναι)
- 5) Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο; (Όχι, Ναι)
- 6)Ηλικία (14-15,9, 16-17,9)
- 7)Εκπαιδευτικό επίπεδο γονέων (Χαμηλό/Μεσαίο, Υψηλό)
- 8)Εθιστική συμπεριφορά στο Διαδίκτυο (Δεν υπάρχουν ενδείξεις, Ήπιες ενδείξεις, Σε κίνδυνο/IAB)
- 9)Βαριά καθημερινή χρήση του SNS (Δεν το χρησιμοποιώ/ <2 ώρες, >= 2 ώρες/ημέρα)

3.5. Μέθοδοι Στατιστικής ανάλυσης

Η στατιστική ανάλυση των δεδομένων πραγματοποιήθηκε στο στατιστικό πρόγραμμα IBM SPSS 24 και στο Microsoft Office Excel 2016. Οι κατηγορικές μεταβλητές παρουσιάστηκαν με συχνότητες, ποσοστά και ραβδογράμματα ενώ οι αναλογικές με μέση τιμή, τυπική απόκλιση, ελάχιστη και μέγιστη τιμή και ιστογράμματα.

Ο έλεγχος υποθέσεων πραγματοποιήθηκε σε στάθμη σημαντικότητας 5%. Στα πρώτα 3 ερευνητικά ερωτήματα χρησιμοποιήθηκε λογιστική παλινδρόμηση καθώς οι εξαρτημένες μεταβλητές ήταν διχοτομικές. Συγκεκριμένα, στο 1^ο ερευνητικό ερώτημα, εξαρτημένη μεταβλητή είναι η «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;» ενώ στο 2^ο και στο 3^ο η «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;». Η διαφορά μεταξύ 2^{ου} και 3^{ου} ερευνητικού ερωτήματος είναι ότι στο 2^ο ερευνητικό ερώτημα χρησιμοποιήθηκε το σύνολο του δείγματος ενώ στο 3^ο μόνο όσα άτομα είχαν απαντήσει ναι στην εξαρτημένη μεταβλητή του 1^{ου} ερευνητικού ερωτήματος «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;». Στο 4^ο ερευνητικό ερώτημα, χρησιμοποιήθηκε πολυωνυμική λογιστική παλινδρόμηση, με εξαρτημένη μεταβλητή την «Γνωριμία μέσω Διαδικτύου» η οποία δημιουργήθηκε συνδυαστικά από τις εξαρτημένες μεταβλητές των προηγούμενων ερωτημάτων «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;» και «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο

Διαδίκτυο;» με 3 πιθανές επιλογές απάντησης (για αυτόν τον λόγο χρησιμοποιήθηκε και πολυωνυμική λογιστική παλινδρόμηση)

- 1) Δεν είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή
- 2) Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή αλλά δεν τον συνάντησα
- 3) Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή και τον συνάντησα.

Ανεξάρτητες μεταβλητές σε όλα τα ερευνητικά ερωτήματα ήταν 1) Φύλο, 2) Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;, 3) Γνωρίζω πολλά πράγματα για το Διαδίκτυο, 4) Ηλικία, 5) Εκπαιδευτικό επίπεδο γονέων, 6) Εθιστική συμπεριφορά στο Διαδίκτυο και 7) Βαριά καθημερινή χρήση του SNS.

Σε όλα τα μοντέλα η αρχική υπόθεση του ελέγχου είναι ότι οι ανεξάρτητες μεταβλητές δεν έχουν στατιστικά σημαντική επίδραση στην εξαρτημένη και η εναλλακτική ότι έχουν. Ο βαθμός προσαρμογής του μοντέλου καθορίστηκε από το Cox & Snell R Square, Nagelkerke R Square και Hosmer-Lemeshow. Επιπλέον, η αρχική υπόθεση για κάθε ανεξάρτητη μεταβλητή είναι ότι δεν έχει στατιστικά σημαντική επίδραση στην εξαρτημένη και η εναλλακτική ότι έχει. Τέλος, παρουσιάστηκαν τα ποσοστά σωστής πρόβλεψης των μοντέλων, δηλαδή αν οι παρατηρήσεις ταξινομούνται σωστά μέσω του μαθηματικού μοντέλου στις κατηγορίες της εξαρτημένης μεταβλητής, συγκρινόμενες πάντα με τις πραγματικές τιμές. (Menard, 2010).

3.6. Ηθικά Ζητήματα

Ο ερευνητής τήρησε τα απαραίτητα ηθικά ζητήματα τα οποία σχετίζονται με την ψυχολογία των νεαρών ατόμων που συμμετείχαν στην έρευνα αλλά και με την φύση της διεξαγωγής της έρευνας (Yip et al., 2016). Συγκεκριμένα:

- ✓ Η διεξαγωγή της έρευνας εγκρίθηκε από τον Ιδρυματικό φορέα του ερευνητή, ενώ καθηγητής του Ιδρυματικού φορέα επέβλεψε την διεξαγωγή της
- ✓ Οι νέοι συμμετέχοντες ηλικίας 14-18 ετών ενημερώθηκαν για τον σκοπό της έρευνας, ότι συμμετέχουν εθελοντικά, με την δική τους συγκατάθεση και ότι έχουν το δικαίωμα να αποχωρήσουν από την έρευνα ή να διαγράψουν την συμμετοχή τους

- ✓ Οι νέοι συμμετέχοντες ενημερώθηκαν ότι οι απαντήσεις τους θα χρησιμοποιηθούν μόνο για ερευνητικούς λόγους.

Κεφάλαιο 4. Αποτελέσματα

4.1. Περιγραφική Στατιστική

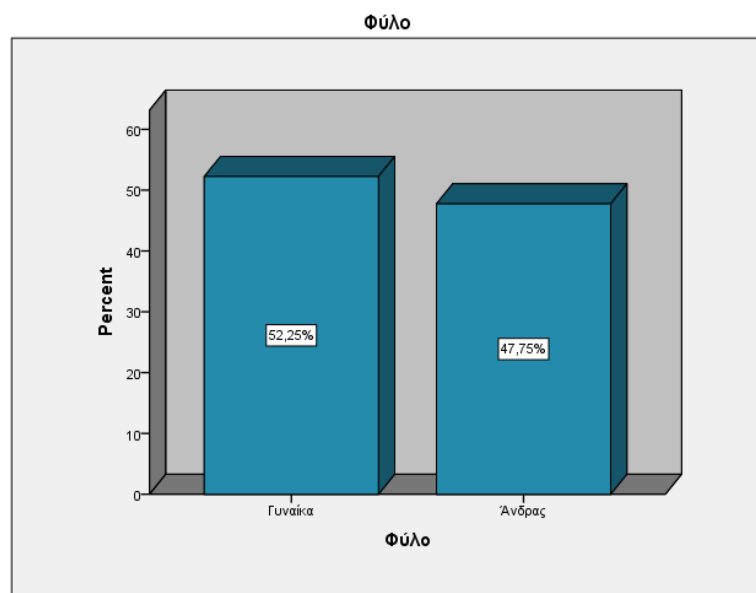
Στον Πίνακα 1 (και τα Γραφήματα 1-8) παρουσιάζονται τα αποτελέσματα σχετικά με τις ποιοτικές μεταβλητές της έρευνας. Όσον αφορά το φύλο, το 52,2% (N=849) αποτελείται από γυναίκες, ενώ το 47,8% (N=776) από άντρες. Το 66,1% (N=1074) ανέφερε πως γνωρίζει πολλά πράγματα για το Διαδίκτυο, ενώ το 33,9% (N=551) ότι δεν γνωρίζει. Ακόμη, το 51,0% (N=828) δήλωσε πως δεν είχε επαφή στο διαδίκτυο, με κάποιον που δεν έχει γνωρίσει στο παρελθόν στην πραγματική ζωή, ενώ το 49,0% (N=797) είχε. Επίσης, το 81,2% (N=1320) δεν έχει ποτέ συναντήσει κάποιον, με τον οποίο γνωρίστηκε για πρώτη φορά στο διαδίκτυο, ενώ το 18,8% (N=305) απάντησε θετικά. Σχετικά με τη ηλικία, το 70,6% (N=1147) είναι 14-15,9 ετών και το 29,4% (N=478) 16-17,9 ετών. Όσον αφορά το εκπαιδευτικό επίπεδο γονέων, το 57,6% (N=936) δήλωσε πως είναι υψηλό, ενώ το 42,4% (N=689) χαμηλό/μεσαίο. Αναφορικά με την εθιστική συμπεριφορά στο διαδίκτυο, το 59,1% (N=961) δήλωσε πως δεν υπάρχουν ενδείξεις, το 29,3% (N=476) ήπιες ενδείξεις και το 11,6% (N=188) βρίσκεται σε κίνδυνο/ IAB. Τέλος, σχετικά με την καθημερινή χρήση του SNS, το 64,1% (N=1042) απάντησε πως δεν το χρησιμοποιεί, ή το χρησιμοποιεί λιγότερο από 2 ώρες, ενώ το 35,9% (N=583) το χρησιμοποιεί 2 ή περισσότερες ώρες την ημέρα.

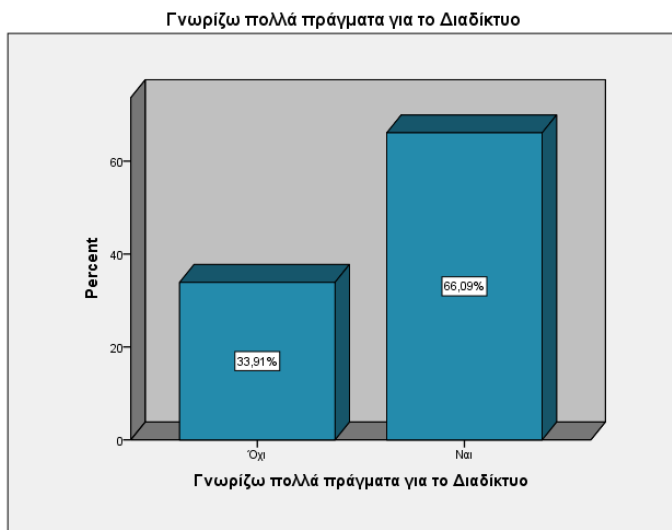
Πίνακας 1: Ποιοτικές μεταβλητές έρευνας

Στοιχείο	Κατηγορία	N	f%
Φύλο	Γυναίκα	849	52,2
	Άντρας	776	47,8
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	Όχι	551	33,9
	Ναι	1074	66,1
Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή	Όχι	828	51,0
	Ναι	797	49,0
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;	Όχι	1320	81,2
	Ναι	305	18,8
Ηλικία	14-15,9	1147	70,6
	16-17,9	478	29,4
Εκπαιδευτικό επίπεδο γονέων	Χαμηλό/Μεσαίο	689	42,4
	Υψηλό	936	57,6
Εθιστική συμπεριφορά στο Διαδίκτυο	Δεν υπάρχουν ενδείξεις	961	59,1
	Ήπιες ενδείξεις	476	29,3
	Σε κίνδυνο/ IAB	188	11,6
Βαριά καθημερινή χρήση του SNS	Δεν το χρησιμοποιώ/ <2 ώρες	1042	64,1
	>= 2 ώρες/ημέρα	583	35,9

N: Συχνότητα

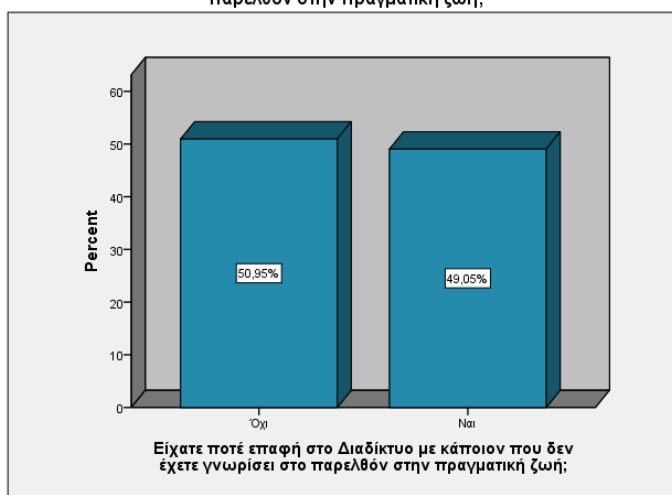
f%: Σχετική συχνότητα %

**Γράφημα 1: Φύλο**



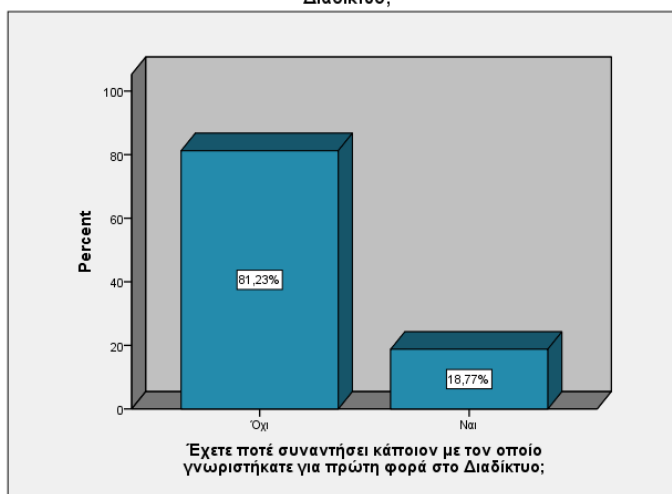
Γράφημα 2: Γνωρίζω πολλά πράγματα για το Διαδίκτυο

Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;

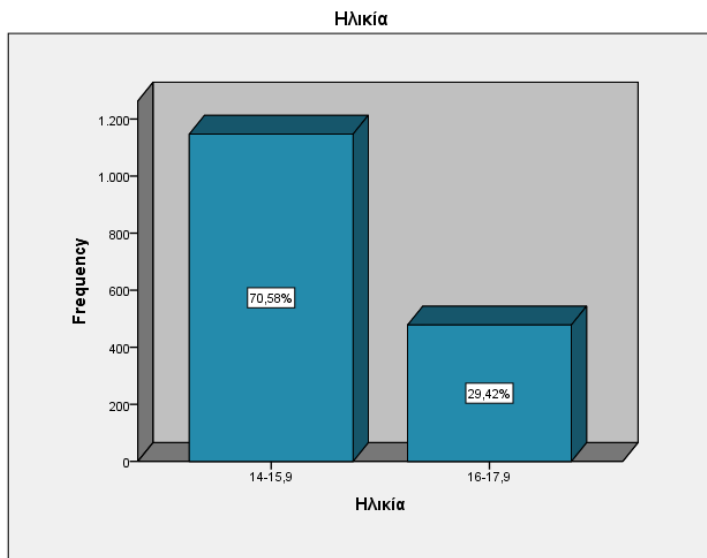


Γράφημα 3: Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;

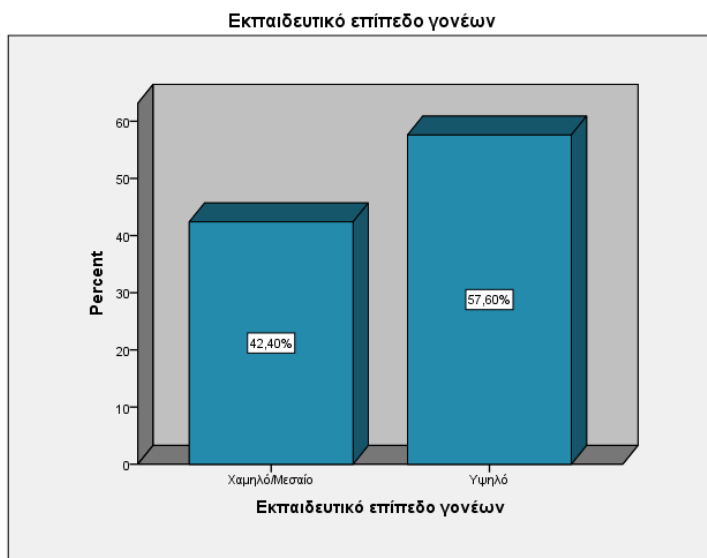
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;



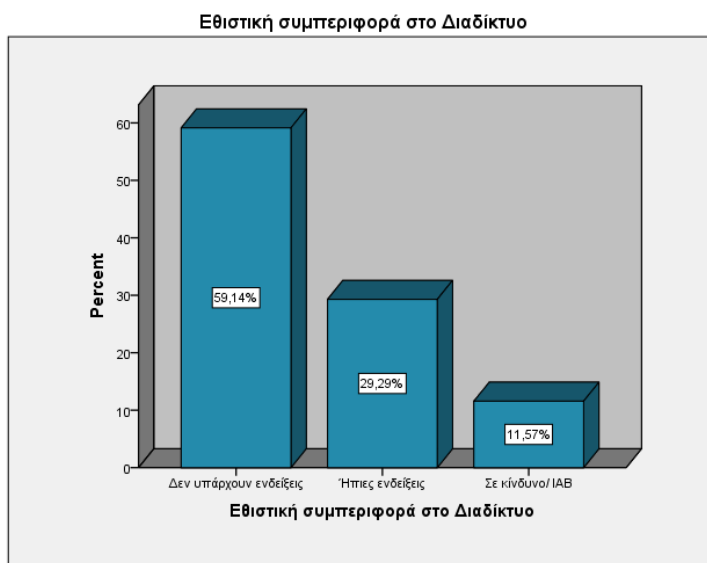
Γράφημα 4: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;



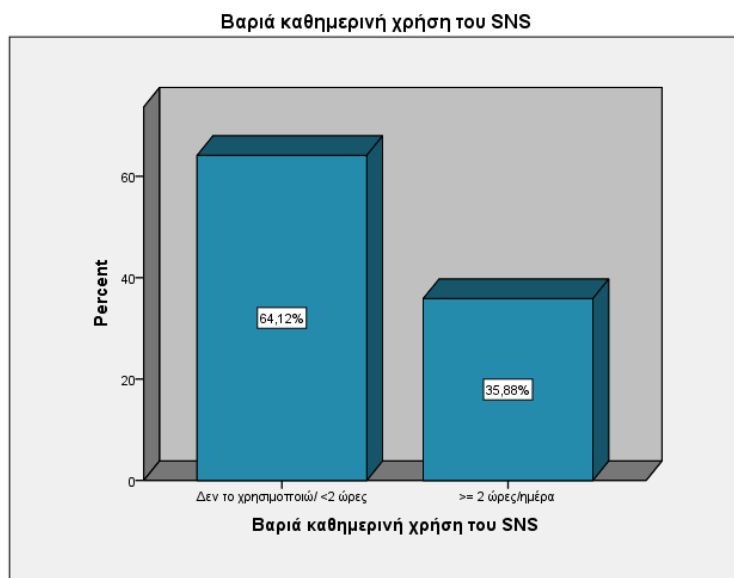
Γράφημα 5: Ηλικία



Γράφημα 6: Εκπαιδευτικό επίπεδο γονέων



Γράφημα 7: Εθιστική συμπεριφορά στο Διαδίκτυο

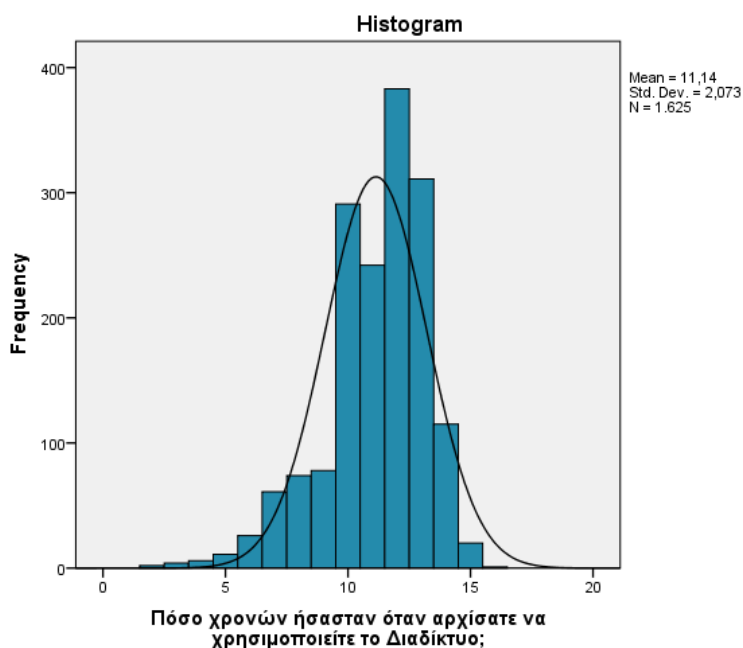


Γράφημα 8: Βαριά καθημερινή χρήση του SNS

Παρακάτω παρατίθενται τα αποτελέσματα σχετικά με την ηλικία έναρξης της χρήσης του διαδικτύου από τους ερωτηθέντες. Από τον Πίνακα 2 (και το Γράφημα 9) προκύπτει ότι οι ερωτηθέντες ήταν κατά μέσο όρο 11 ετών όταν ξεκίνησαν να χρησιμοποιούν το Διαδίκτυο (Μ.Ο.= 11,1±2,07).

Πίνακας 2: Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;

Στατιστικό	Τιμή
Μ.Ο.	11,1
Τ.Α.	2,07
Ελάχιστη	2
Μέγιστη	16



Γράφημα 9: Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;

4.2. Επαγωγική Στατιστική

4.2.1. 1^ο ερευνητικό ερώτημα:

Ποιοι παράγοντες επηρεάζουν την επαφή στο Διαδίκτυο με κάποιο άτομο που δεν υπάρχει γνωριμία στο παρελθόν στην πραγματική ζωή;

Ο Πίνακας 3 παρουσιάζει τα αποτελέσματα της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;» και ανεξάρτητες μεταβλητές το «Φύλο», την «Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;», την «Γνωρίζω πολλά πράγματα για το Διαδίκτυο», την «Ηλικία», το «Εκπαιδευτικό επίπεδο γονέων», την «Εθιστική συμπεριφορά στο Διαδίκτυο» και την «Βαριά καθημερινή χρήση του SNS».

Προκύπτει στατιστικά σημαντική επίδραση των ανεξάρτητων μεταβλητών στην εξαρτημένη ($X^2(8) = 165,189, p < 0,001, Nagelkerke R^2 = 0,129$). Στατιστικά σημαντική θεωρήθηκε η επίδραση των μεταβλητών «Φύλο» ($B = 0,251, Wald = 5,338, p = 0,021$), «Γνωρίζω πολλά πράγματα για το Διαδίκτυο» ($B = 0,256, Wald = 5,071, p = 0,024$), «Εθιστική συμπεριφορά στο Διαδίκτυο(0)» ($B = -1,018, Wald = 56,385, p < 0,001$), «Εθιστική συμπεριφορά στο Διαδίκτυο(1)» ($B = -0,260, Wald = 31,232, p < 0,001$), και «Βαριά καθημερινή χρήση του SNS» ($B = 0,646, Wald = 31,326, p < 0,001$).

Πίνακας 3: Αποτελέσματα λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;»

Εξαρτημένη μεταβλητή	X ²	df	p-value	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square		
Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;	165,189	8	<0,001	2086,948	0,097	0,129		
Hosmer and Lemeshow Test	9,158	8	0,329					
Ανεξάρτητες μεταβλητές	B	S.E.	Wald	df	p-value	Exp (B)	95% C.I.for EXP(B)0	
							Lower	Upper
Φύλο	0,251	0,108	5,338	1	0,021	1,285	1,039	1,589
Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	-0,041	0,027	2,260	1	0,133	0,960	0,911	1,012
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	0,256	0,114	5,071	1	0,024	1,292	1,034	1,615
Ηλικία	0,214	0,115	3,454	1	0,063	1,239	0,988	1,554
Εκπαιδευτικό επίπεδο γονέων	0,141	0,108	1,700	1	0,192	1,151	0,932	1,422
Εθιστική συμπεριφορά στο Διαδίκτυο(0)	-1,018		56,385	2	<0,001			
Εθιστική συμπεριφορά στο Διαδίκτυο(1)	-0,260	0,182	31,232	1	<0,001	0,361	0,253	0,516
Εθιστική συμπεριφορά στο Διαδίκτυο(2)		0,190	1,866	1	0,172	0,771	0,531	1,120
Βαριά καθημερινή χρήση του SNS	0,639	0,116	30,501	1	<0,001	1,895	1,510	2,377
Σταθερά	0,219	0,398	0,304	1	0,582	1,245		

Ο Πίνακας 4 (Γραφήματα 10-13) παρουσιάζει τον τρόπο εξάρτησης της μεταβλητής «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;» με τις μεταβλητές «Φύλο», «Γνωρίζω πολλά πράγματα για το Διαδίκτυο» «Εθιστική συμπεριφορά στο Διαδίκτυο» και «Βαριά καθημερινή χρήση του SNS».

Προκύπτει ότι όσον αφορά το φύλο, μελετώντας τους άντρες το ποσοστό των ατόμων που έχουν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν

στην πραγματική ζωή (53,9%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (49%).

Επίσης, όσον αφορά την γνώση του διαδικτύου, παρατηρείται ότι στα άτομα που ισχυρίζονται ότι γνωρίζουν πολλά πράγματα για το διαδίκτυο, το ποσοστό των ατόμων που έχουν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή (52,8%) είναι υψηλότερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (49%).

Σχετικά με την εθιστική συμπεριφορά στο διαδίκτυο, προκύπτει ότι στα άτομα που έχουν ήπιες ενδείξεις ή ενδείξεις κινδύνου για εθιστική συμπεριφορά, το ποσοστό των ατόμων που έχουν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή (61,6% και 70,2% αντίστοιχα) είναι υψηλότερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (49%).

Αναφορικά με την βαριά καθημερινή χρήση του SNS, προκύπτει ότι στα άτομα που το χρησιμοποιούν 2 ώρες ή περισσότερο καθημερινά, το ποσοστό των ατόμων που έχουν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή (64%) είναι υψηλότερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (49%).

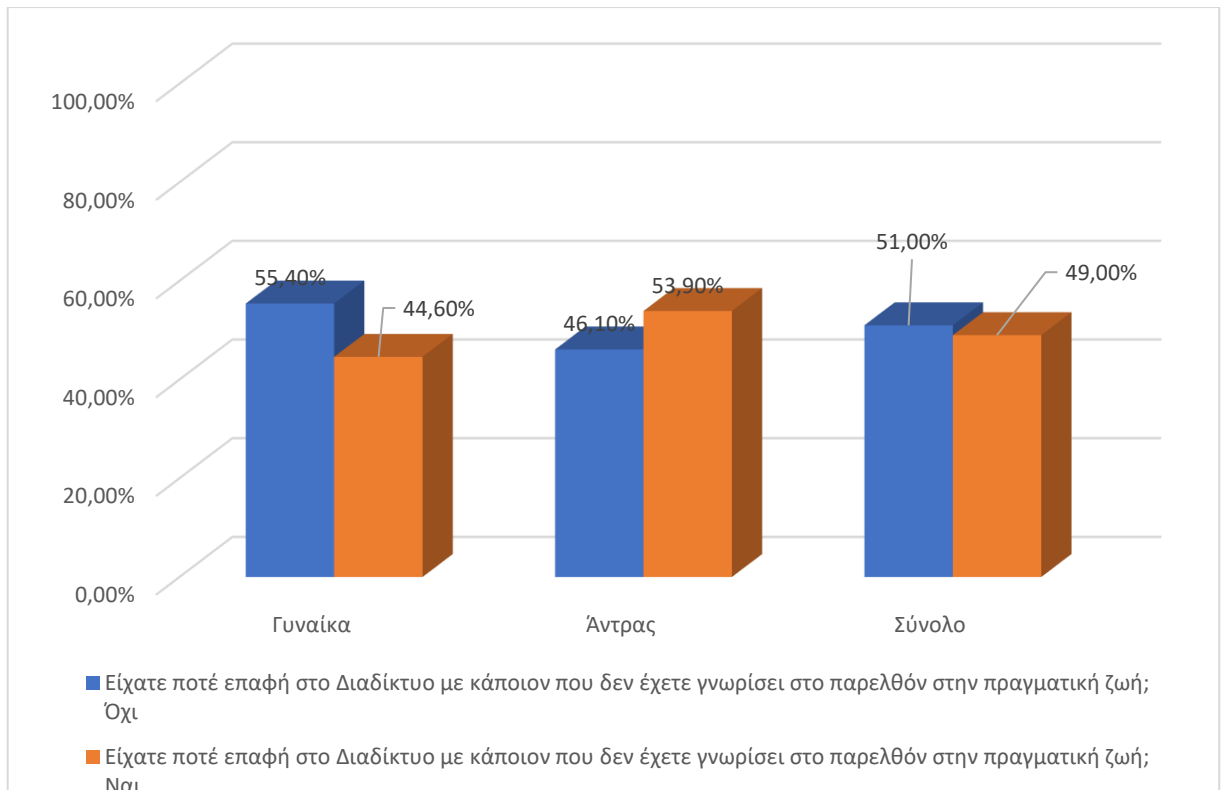
Τέλος, στον Πίνακα 5 παρουσιάζονται τα ποσοστά της σωστής πρόβλεψης του μοντέλου. Προκύπτει ότι το μοντέλο έχει σωστή πρόβλεψη στο 63,80% των περιπτώσεων, με το ποσοστό να είναι υψηλότερο στα άτομα που δεν είχαν ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή (69,60%) και χαμηλότερο για τα άτομα που είχαν επαφή (57,80%).

Πίνακας 4: Στατιστικά σημαντικές εξαρτήσεις για την μεταβλητή «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;»

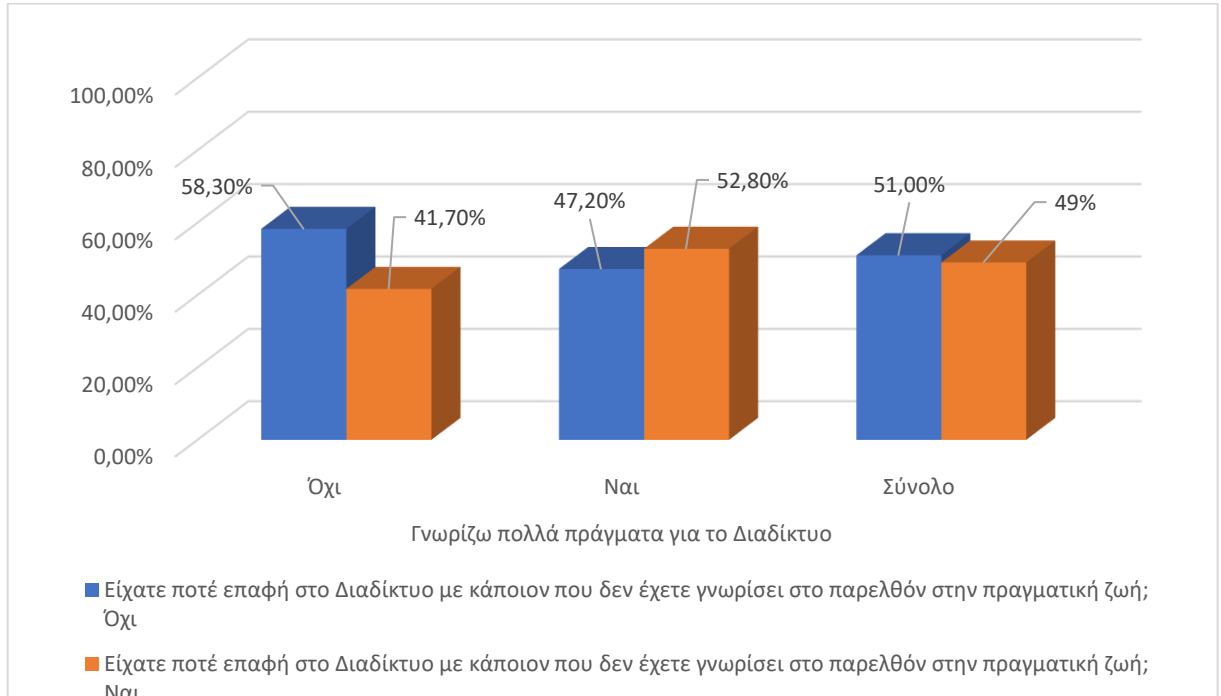
			Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;	
			Όχι	Ναι
Φύλο	Γυναίκα	N	470	379
		%	55,4%	44,6%
	Ανδρας	N	358	418
		%	46,1%	53,9%
	Σύνολο	N	828	797
		%	51,0%	49,0%
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	Όχι	N	321	230
		%	58,3%	41,7%
	Ναι	N	507	567
		%	47,2%	52,8%
	Σύνολο	N	828	797
		%	51,0%	49%
Εθιστική συμπεριφορά στο Διαδίκτυο	Δεν υπάρχουν ενδείξεις	N	589	372
		%	61,3%	38,7%
	Ήπιες ενδείξεις	N	183	293
		%	38,4%	61,6%
	Σε κίνδυνο/ IAB	N	56	132
		%	29,8%	70,2%
Σύνολο	N	828	797	
	%	51,0%	49,0%	
Βαριά καθημερινή χρήση του SNS	Δεν το χρησιμοποιώ/ <2 ώρες	N	618	424
		%	59,3%	40,7%
	>= 2 ώρες/ημέρα	N	210	373
		%	36,0%	64,0%
	Σύνολο	N	828	797
		%	51,0%	49,0%

Πίνακας 5: Αποτελέσματα πρόβλεψης μοντέλου λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;»

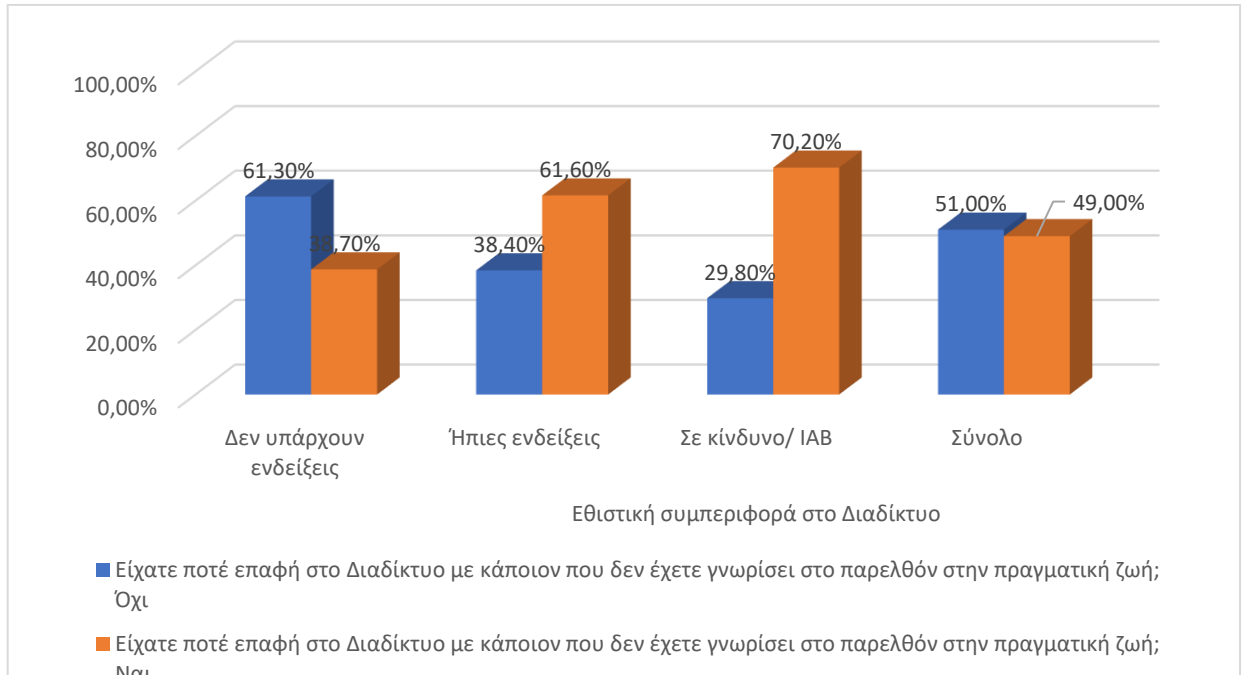
Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή;			
	Σωστά	Λάθος	%
Όχι	576	252	69,60
Ναι	461	336	57,80
Συνολικό	1037	588	63,80



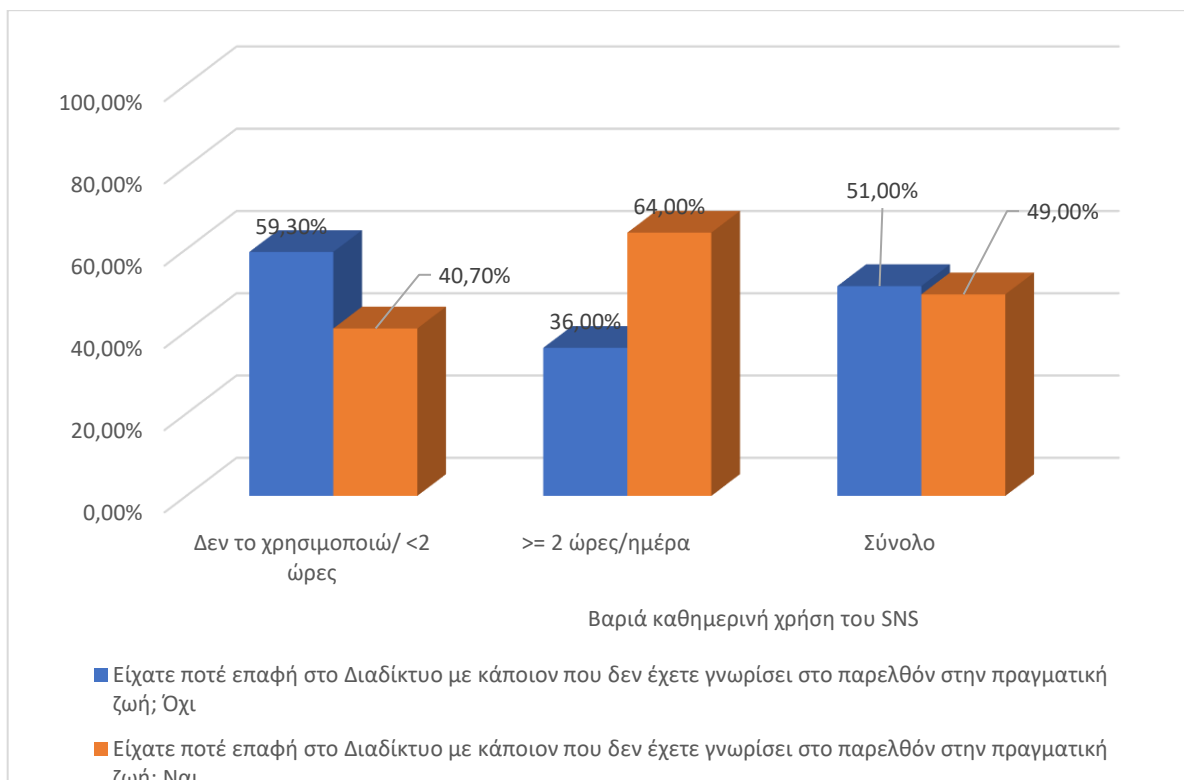
Γράφημα 10: Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή; * Φύλο



Γράφημα 11: Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή; * Γνωρίζω πολλά πράγματα για το Διαδίκτυο



Γράφημα 12: Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή; * Εθιστική συμπεριφορά στο Διαδίκτυο



Γράφημα 13: Είχατε ποτέ επαφή στο Διαδίκτυο με κάποιον που δεν έχετε γνωρίσει στο παρελθόν στην πραγματική ζωή; * Βαριά καθημερινή χρήση του SNS

4.2.2. 2^ο ερευνητικό ερώτημα:

Ποιοι παράγοντες επηρεάζουν την συνάντηση με άτομο με το οποίο υπήρξε γνωριμία για πρώτη φορά στο Διαδίκτυο;

Ο Πίνακας 6 παρουσιάζει τα αποτελέσματα της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» και ανεξάρτητες μεταβλητές το «Φύλο», την «Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;», την «Γνωρίζω πολλά πράγματα για το Διαδίκτυο», την «Ηλικία», το «Εκπαιδευτικό επίπεδο γονέων», την «Εθιστική συμπεριφορά στο Διαδίκτυο» και την «Βαριά καθημερινή χρήση του SNS».

Προκύπτει στατιστικά σημαντική επίδραση των ανεξάρτητων μεταβλητών στην εξαρτημένη ($X^2(7)=150,858, p<0,001, Nagelkerke R^2=0,143$). Στατιστικά σημαντική θεωρήθηκε η επίδραση των μεταβλητών «Φύλο» ($B=0,512, Wald=13,205, p<0,001$), «Ηλικία» ($B=0,379, Wald=7,042, p=0,008$), «Εθιστική συμπεριφορά στο Διαδίκτυο» ($B=0,650, Wald=48,029, p<0,001$) και «Βαριά καθημερινή χρήση του SNS» ($B=0,769, Wald=29,448, p<0,001$).

Πίνακας 6: Αποτελέσματα λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;»

Εξαρτημένη μεταβλητή	X^2	df	p-value	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;	150,858	7	<0,001	1418,435	0,089	0,143
Ανεξάρτητες μεταβλητές	B	S.E.	Wald	df	p-value	Exp (B)
Φύλο	0,512	0,141	13,205	1	<0,001	1,668
Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	-0,035	0,033	1,160	1	0,281	0,966
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	0,095	0,152	0,391	1	0,532	1,099
Ηλικία	0,379	0,143	7,042	1	0,008	1,460
Εκπαιδευτικό επίπεδο γονέων	-0,029	0,137	0,044	1	0,835	0,972
Εθιστική συμπεριφορά στο Διαδίκτυο	0,650	0,094	48,029	1	<0,001	1,916
Βαριά καθημερινή χρήση του SNS	0,769	0,142	29,448	1	<0,001	2,158
Σταθερά	-2,653	0,471	31,770	1	<0,001	0,070

Ο Πίνακας 7 (Γραφήματα 14-17) παρουσιάζει τον τρόπο εξάρτησης της μεταβλητής «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» με τις μεταβλητές «Φύλο», «Ηλικία» «Εθιστική συμπεριφορά στο Διαδίκτυο» και «Βαριά καθημερινή χρήση του SNS».

Προκύπτει ότι όσον αφορά το φύλο, μελετώντας τους άντρες, το ποσοστό των ατόμων που έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (23,3%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (18,8%).

Επίσης, όσον αφορά την ηλικία, μελετώντας τα άτομα ηλικίας 16-17,9 ετών, το ποσοστό των ατόμων που έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (22,8%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (18,8%).

Σχετικά με την εθιστική συμπεριφορά στο διαδίκτυο, προκύπτει ότι στα άτομα που έχουν ήπιες ενδείξεις ή ενδείξεις κινδύνου για εθιστική συμπεριφορά, το ποσοστό των ατόμων που έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (23,7% και 41,5%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (18,8%).

Αναφορικά με την βαριά καθημερινή χρήση του SNS, προκύπτει ότι στα άτομα που το χρησιμοποιούν 2 ώρες ή περισσότερο καθημερινά, το ποσοστό των ατόμων που έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (30,2%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (18,8%).

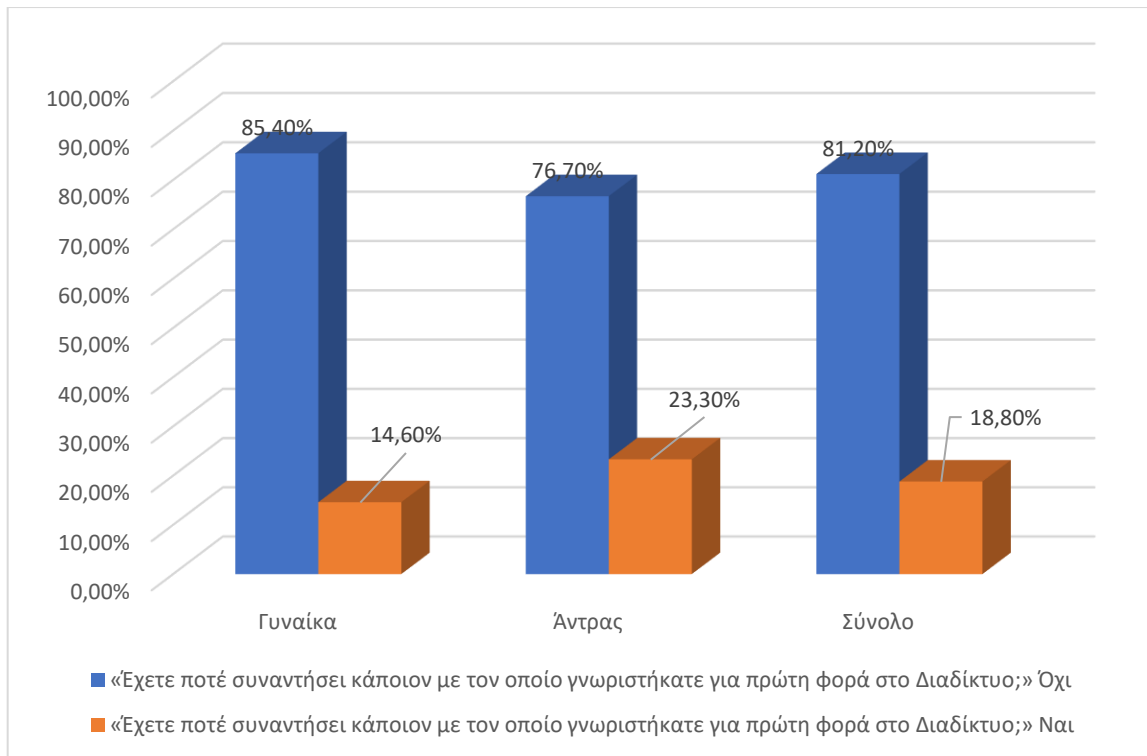
Τέλος, στον Πίνακα 8 παρουσιάζονται τα ποσοστά της σωστής πρόβλεψης του μοντέλου. Προκύπτει ότι το μοντέλο έχει σωστή πρόβλεψη στο 81,8% των περιπτώσεων, με το ποσοστό να είναι υψηλότερο στα άτομα που δεν είχαν ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκανε για πρώτη φορά στο Διαδίκτυο (98,7%) και χαμηλότερο για τα άτομα που είχαν συναντήσει (8,9%).

Πίνακας 7: Στατιστικά σημαντικές εξαρτήσεις για την μεταβλητή «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;»

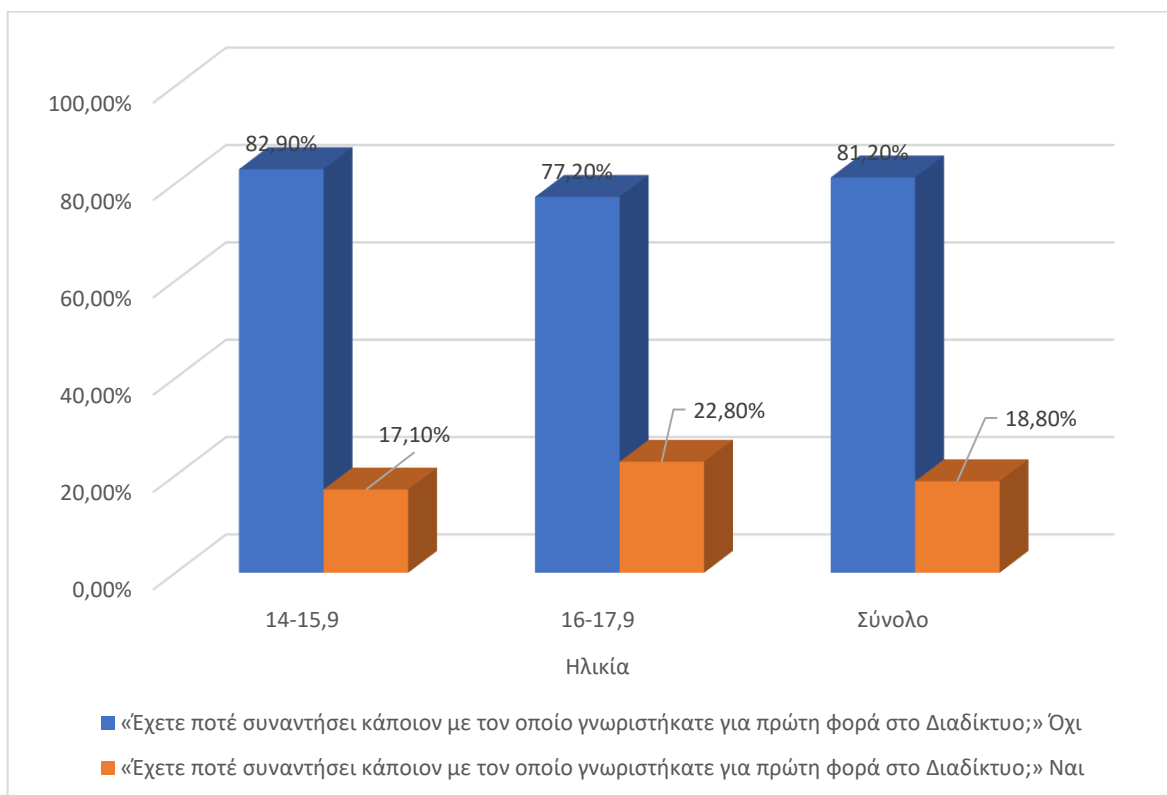
			«Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;»	
			Όχι	Ναι
Φύλο	Γυναίκα	N	725	124
		%	85,4%	14,6%
	Ανδρας	N	595	181
		%	76,7%	23,3%
	Σύνολο		N	1320
		%	81,2%	18,8%
Ηλικία	14-15,9	N	951	196
		%	82,9%	17,1%
	16-17,9	N	369	109
		%	77,2%	22,8%
	Σύνολο		N	1320
		%	81,2%	18,8%
Εθιστική συμπεριφορά στο Διαδίκτυο	Δεν υπάρχουν ενδείξεις	N	847	114
		%	88,1%	11,9%
	Ήπιες ενδείξεις	N	363	113
		%	76,3%	23,7%
	Σε κίνδυνο/ IAB	N	110	78
%		58,5%	41,5%	
Σύνολο		N	1320	305
		%	81,2%	18,8%
Βαριά καθημερινή χρήση του SNS	Δεν το χρησιμοποιώ/ <2 ώρες	N	913	129
		%	87,6%	12,4%
	>= 2 ώρες/ημέρα	N	407	176
		%	69,8%	30,2%
	Σύνολο		N	1320
		%	81,2%	18,8%

Πίνακας 8: Αποτελέσματα πρόβλεψης μοντέλου λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;»

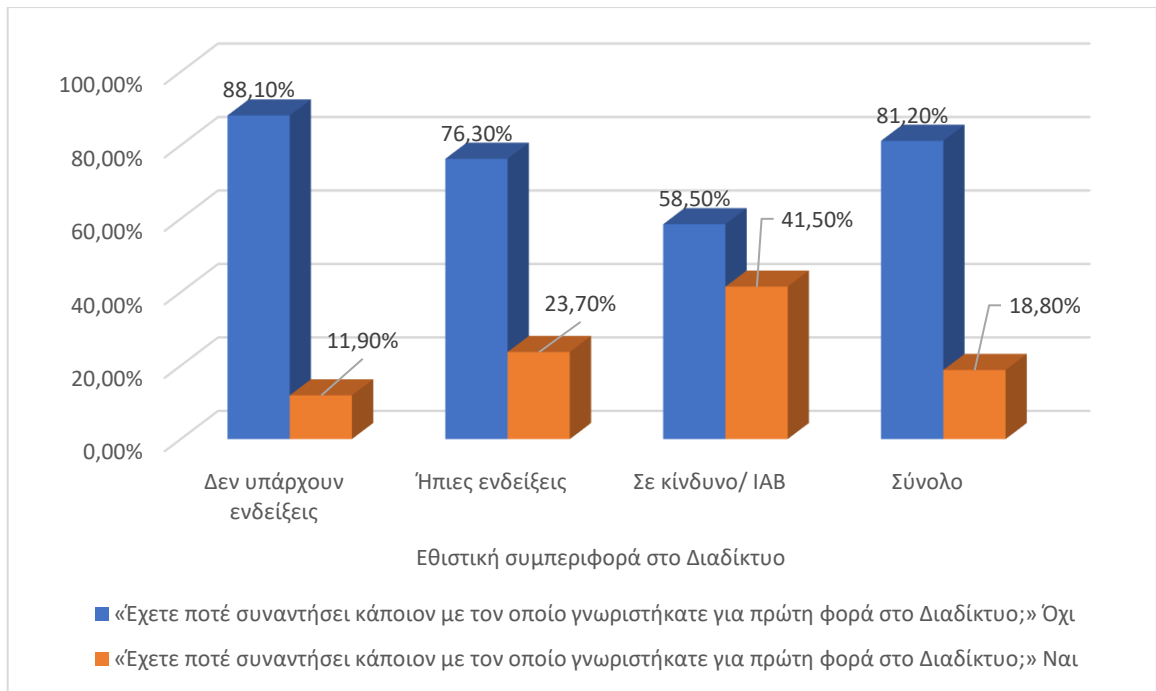
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;	Σωστά	Λάθος	%
Όχι	1303	17	98,7
Ναι	27	278	8,9
Συνολικό	1330	295	81,8



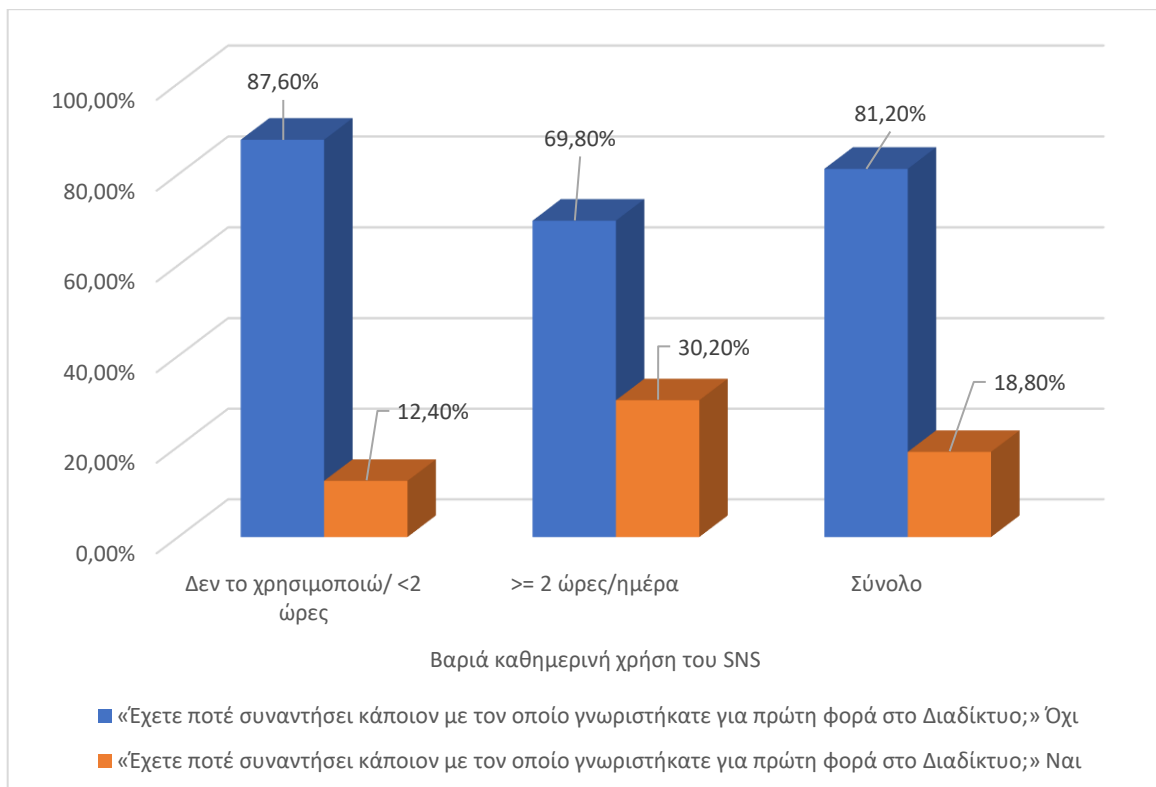
Γράφημα 14: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Φύλο



Γράφημα 15: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Ηλικία



Γράφημα 16: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Εθιστική συμπεριφορά στο Διαδίκτυο



Γράφημα 17: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Βαριά καθημερινή χρήση του SNS

4.2.3. 3^ο ερευνητικό ερώτημα:

Όσον αφορά τα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή, ποιοι παράγοντες επηρεάζουν την συνάντηση;

Ο Πίνακας 9 παρουσιάζει τα αποτελέσματα της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» και ανεξάρτητες μεταβλητές το «Φύλο», την «Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;», την «Γνωρίζω πολλά πράγματα για το Διαδίκτυο», την «Ηλικία», το «Εκπαιδευτικό επίπεδο γονέων», την «Εθιστική συμπεριφορά στο Διαδίκτυο» και την «Βαριά καθημερινή χρήση του SNS», όσον αφορά τα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή

Προκύπτει στατιστικά σημαντική επίδραση των ανεξάρτητων μεταβλητών στην εξαρτημένη ($X^2(7)=56,711$, $p<0,001$, Nagelkerke $R^2=0,093$). Στατιστικά σημαντική θεωρήθηκε η επίδραση των μεταβλητών «Φύλο» ($B=0,462$, $Wald=8,243$, $p=0,004$), «Εθιστική συμπεριφορά στο Διαδίκτυο» ($B=0,447$, $Wald=17,165$, $p<0,001$) και «Βαριά καθημερινή χρήση του SNS» ($B=0,515$, $Wald=10,324$, $p=0,001$).

Πίνακας 9: Αποτελέσματα λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)

Εξαρτημένη μεταβλητή	X^2	df	p-value	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;	56,711	7	<0,001	1003,878	0,069	0,093
Ανεξάρτητες μεταβλητές	B	S.E.	Wald	df	p-value	Exp (B)
Φύλο	0,462	0,161	8,243	1	0,004	1,587
Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	-0,019	0,037	0,264	1	0,607	0,981
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	-0,124	0,174	0,508	1	0,476	0,884
Ηλικία	0,314	0,163	3,695	1	0,055	1,369
Εκπαιδευτικό επίπεδο γονέων	-0,168	0,155	1,173	1	0,279	0,846
Εθιστική συμπεριφορά στο Διαδίκτυο	0,447	0,108	17,165	1	<0,001	1,564
Βαριά καθημερινή χρήση του SNS	0,515	0,160	10,324	1	0,001	1,674
Σταθερά	-1,325	0,533	6,183	1	0,013	0,266

Ο Πίνακας 10 (Γραφήματα 18-20) παρουσιάζει τον τρόπο εξάρτησης της μεταβλητής «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» με τις μεταβλητές «Φύλο», «Εθιστική συμπεριφορά στο Διαδίκτυο» και «Βαριά καθημερινή χρήση του SNS», για τα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή.

Προκύπτει ότι όσον αφορά το φύλο, μελετώντας τους άντρες, το ποσοστό των ατόμων που έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (43,3%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (38,3%).

Σχετικά με την εθιστική συμπεριφορά στο διαδίκτυο, προκύπτει ότι στα άτομα που έχουν ενδείξεις κινδύνου για εθιστική συμπεριφορά, το ποσοστό των ατόμων έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (59,1%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (38,3%).

Αναφορικά με την βαριά καθημερινή χρήση του SNS, προκύπτει ότι στα άτομα που το χρησιμοποιούν 2 ώρες ή περισσότερο καθημερινά, το ποσοστό των ατόμων έχουν συναντήσει άτομο με το οποίο γνωρίστηκαν για πρώτη φορά στο Διαδίκτυο (47,2%) είναι μεγαλύτερο σε σύγκριση με το ποσοστό στο συνολικό δείγμα (38,3%).

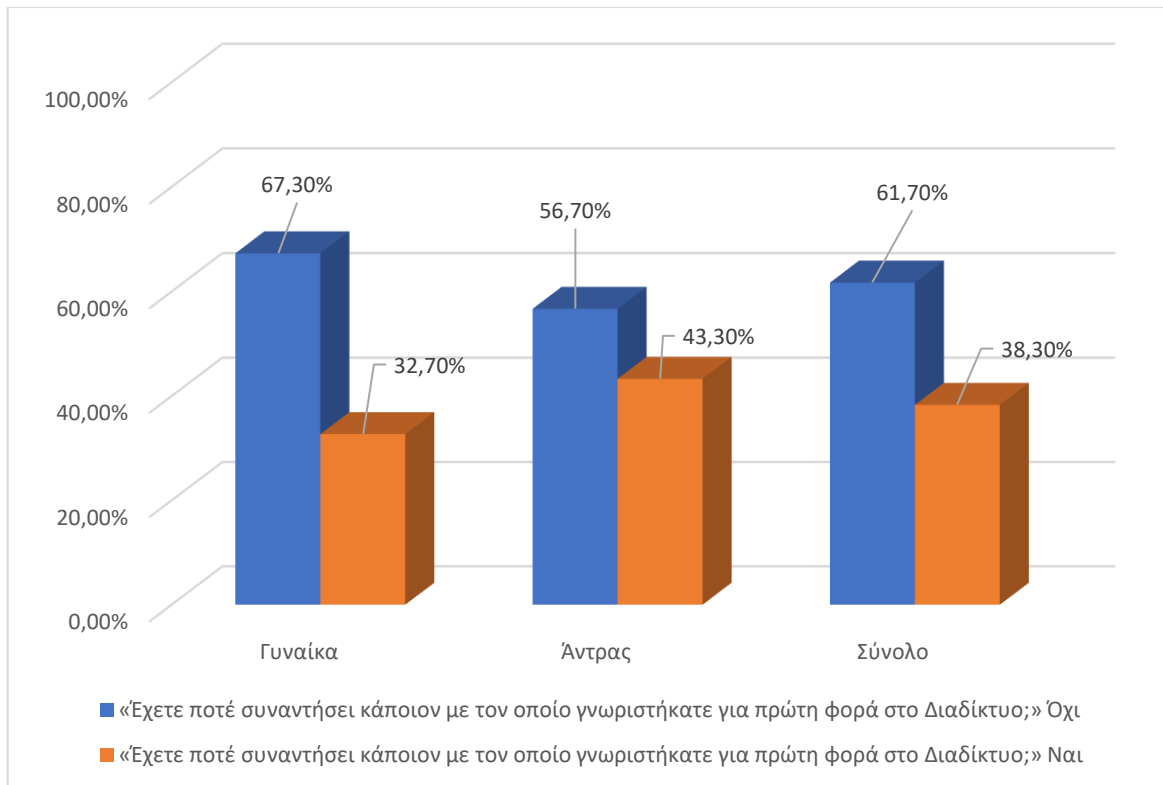
Τέλος, στον Πίνακα 11 παρουσιάζονται τα ποσοστά της σωστής πρόβλεψης του μοντέλου. Προκύπτει ότι το μοντέλο έχει σωστή πρόβλεψη στο 64,6% των περιπτώσεων, με το ποσοστό να είναι υψηλότερο στα άτομα που δεν είχαν ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκανε για πρώτη φορά στο Διαδίκτυο (86,6%) και χαμηλότερο για τα άτομα που είχαν συναντήσει (29,2%).

Πίνακας 10: Στατιστικά σημαντικές εξαρτήσεις για την μεταβλητή «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)

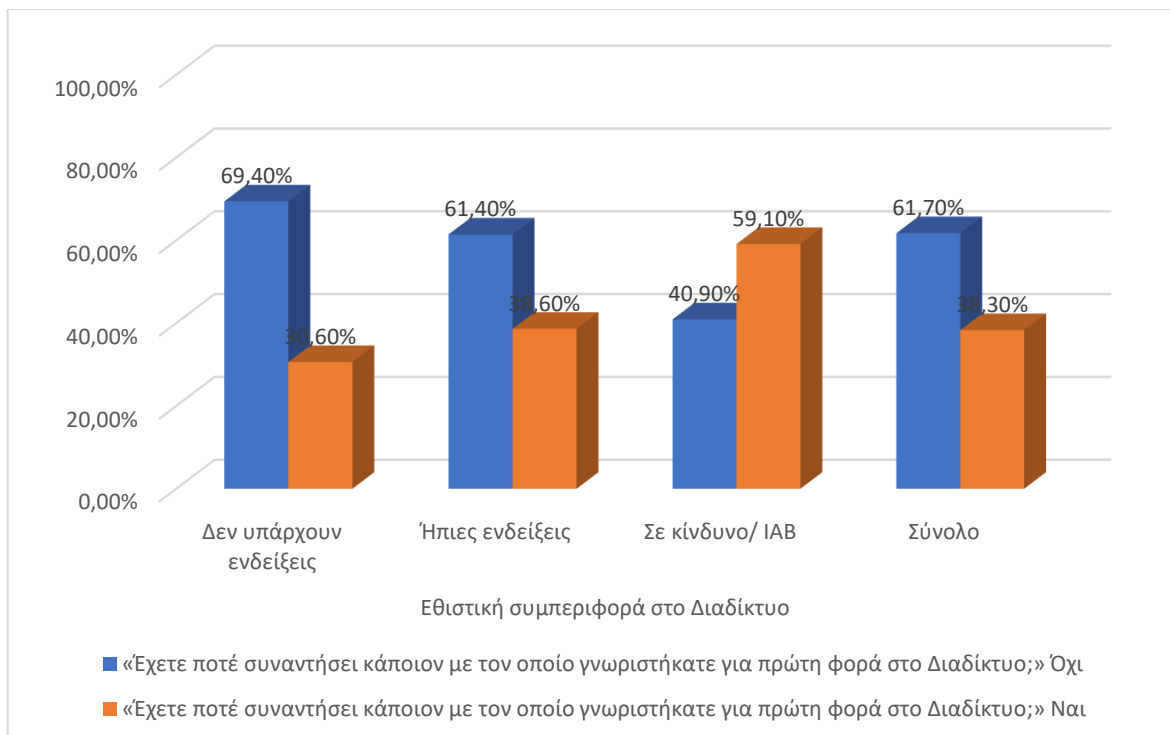
			«Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;»	
			Όχι	Ναι
Φύλο	Γυναίκα	N	255	124
		%	67,3%	32,7%
	Άνδρας	N	237	181
		%	56,7%	43,3%
	Σύνολο		N	492
		%	61,7%	38,3%
Εθιστική συμπεριφορά στο Διαδίκτυο	Δεν υπάρχουν ενδείξεις	N	258	114
		%	69,4%	30,6%
	Ήπιες ενδείξεις	N	180	113
		%	61,4%	38,6%
	Σε κίνδυνο/ IAB	N	54	78
%		40,9%	59,1%	
Σύνολο		N	492	305
		%	61,7%	38,3%
Βαριά καθημερινή χρήση του SNS	Δεν το χρησιμοποιώ/ <2 ώρες	N	295	129
		%	69,6%	30,4%
	>= 2 ώρες/ημέρα	N	197	176
		%	52,8%	47,2%
	Σύνολο		N	492
		%	61,7%	38,3%

Πίνακας 11: Αποτελέσματα πρόβλεψης μοντέλου λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;» (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)

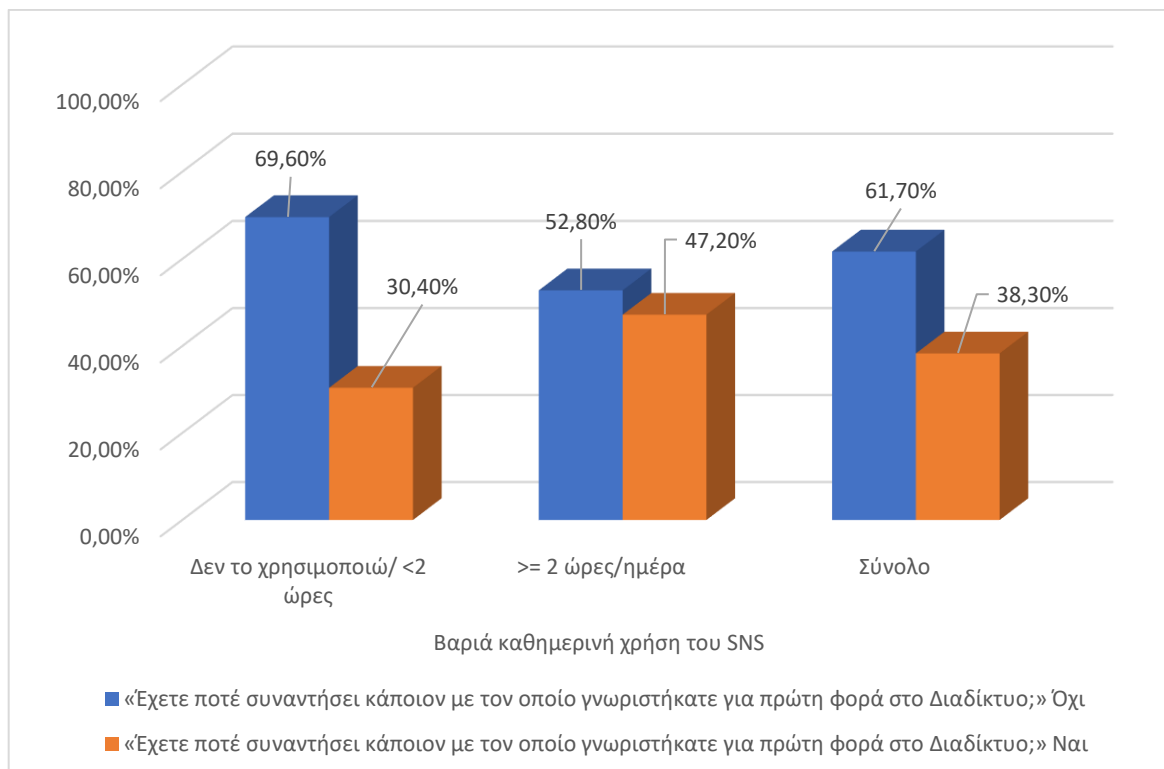
Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;		Σωστά	Λάθος	%
Όχι		426	66	86,6
Ναι		89	216	29,2
Συνολικό		515	282	64,6



Γράφημα 18: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Φύλο (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)



Γράφημα 19: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Εθιστική συμπεριφορά στο Διαδίκτυο (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)



Γράφημα 20: Έχετε ποτέ συναντήσει κάποιον με τον οποίο γνωριστήκατε για πρώτη φορά στο Διαδίκτυο;* Βαριά καθημερινή χρήση του SNS (άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή)

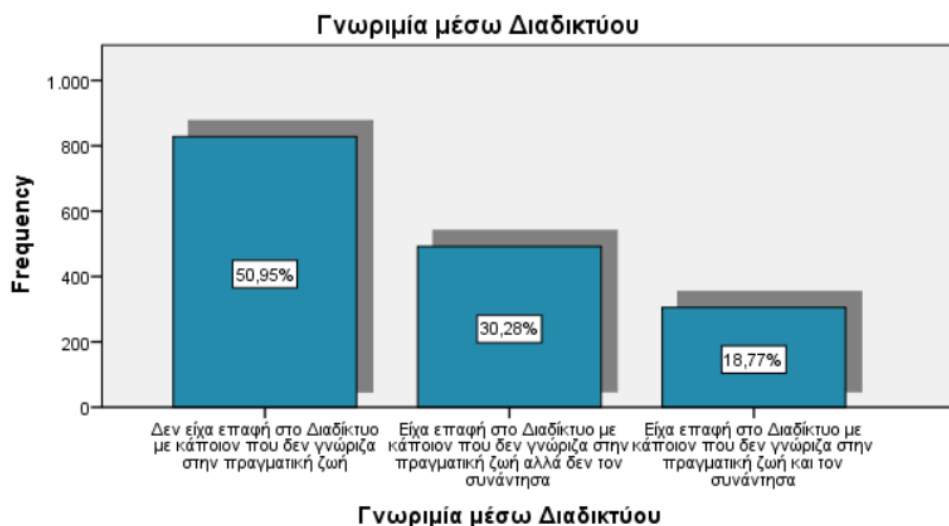
4.2.4. 4^ο ερευνητικό ερώτημα:

Ποιοι παράγοντες επηρεάζουν την γνωριμία μέσω Διαδικτύου μέχρι την συνάντηση ;

Ο Πίνακας 12 (Γράφημα 21) παρουσιάζει τα αποτελέσματα της μεταβλητής που συνδυάζει τις μεταβλητές που αναφέρονται στην γνωριμία μέσω Διαδικτύου και της 1^{ης} συνάντησης. Προκύπτει ότι το 51% (N=828) δεν είχε επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζε στην πραγματική ζωή, το 30,3% (N=492), είχε επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζε στην πραγματική ζωή αλλά δεν τον συνάντησε και το 18,8% (N=305) είχε επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζε στην πραγματική ζωή και τον συνάντησε.

Πίνακας 12: Αποτελέσματα για την γνωριμία μέσω Διαδικτύου

Κατηγορίες	N	%
Δεν είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή	828	51,0
Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή αλλά δεν τον συνάντησα	492	30,3
Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή και τον συνάντησα	305	18,8



Γράφημα 21: Γνωριμία μέσω διαδικτύου

Ο Πίνακας 13 παρουσιάζει τα αποτελέσματα της πολυωνυμικής λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Γνωριμία μέσω Διαδικτύου» και ανεξάρτητες μεταβλητές το «Φύλο», την «Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;», την «Γνωρίζω πολλά πράγματα για το Διαδίκτυο», την «Ηλικία», το «Εκπαιδευτικό επίπεδο γονέων», την «Εθιστική συμπεριφορά στο Διαδίκτυο» και την «Βαριά καθημερινή χρήση του SNS». Προκύπτει στατιστικά σημαντική επίδραση των ανεξάρτητων μεταβλητών στην εξαρτημένη ($X^2(16)=224,086$, $p<0,001$, Nagelkerke $R^2=0,148$). Στατιστικά σημαντική θεωρήθηκε η επίδραση των μεταβλητών «Φύλο» ($X^2(2)=14,078$, $p=0,001$), «Ηλικία» ($X^2(2)=7,485$, $p=0,024$), «Εθιστική συμπεριφορά στο Διαδίκτυο» ($X^2(4)=77,263$, $p<0,001$) και «Βαριά καθημερινή χρήση του SNS» ($X^2(2)=41,294$, $p<0,001$).

Πίνακας 13: Αποτελέσματα πολυωνυμικής λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Γνωριμία μέσω Διαδικτύου»

Εξαρτημένη μεταβλητή	X^2	df	p-value	Nagelkerke R Square
Γνωριμία μέσω Διαδικτύου	224,086	16	<0,001	0,148
Ανεξάρτητες μεταβλητές				
Φύλο	14,078	2	0,001	
Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	2,326	2	0,313	
Γνωρίζω πολλά πράγματα για το Διαδίκτυο	5,299	2	0,071	
Ηλικία	7,485	2	0,024	
Εκπαιδευτικό επίπεδο γονέων	2,483	2	0,289	
Εθιστική συμπεριφορά στο Διαδίκτυο	77,263	4	<0,001	
Βαριά καθημερινή χρήση του SNS	41,294	2	<0,001	
Σταθερά	14,224	2	0,001	

Πίνακας 14: Parameter Estimate πολυωνυμικής λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Γνωριμία μέσω Διαδικτύου»

Meet via internet ^a	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Lower Bound
2 Intercept	,785	,386	4,140	1	,042			
Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	-,033	,030	1,222	1	,269	,967	,912	1,026
Φύλλο=0	-,093	,121	,588	1	,443	,911	,719	1,156
Φύλλο=1	0 ^b	.	.	0
Γνωρίζω πολλά πράγματα για το Διαδίκτυο=0	-,283	,128	4,902	1	,027	,754	,587	,968
Γνωρίζω πολλά πράγματα για το Διαδίκτυο =1	0 ^b	.	.	0
Ηλικία=1	-,096	,130	,543	1	,461	,909	,704	1,172
Ηλικία=2	0 ^b	.	.	0
Εκπαιδευτικό επίπεδο γονέων=0	-,188	,121	2,436	1	,119	,828	,654	1,049
Εκπαιδευτικό επίπεδο γονέων=1	0 ^b	.	.	0
Εθιστική συμπεριφορά στο Διαδίκτυο =0	-,590	,214	7,637	1	,006	,554	,365	,842
Εθιστική συμπεριφορά στο Διαδίκτυο =1	,107	,221	,232	1	,630	1,112	,721	1,716
Εθιστική συμπεριφορά στο Διαδίκτυο =2	0 ^b	.	.	0
Βαριά καθημερινή χρήση SNS=0	-,451	,130	11,951	1	,001	,637	,494	,823
Βαριά καθημερινή χρήση SNS=1	0 ^b	.	.	0

3	Intercept	1,833	,430	18,179	1	,000			
	Πόσο χρονών ήσασταν όταν αρχίσατε να χρησιμοποιείτε το Διαδίκτυο;	-,050	,035	1,960	1	,162	,952	,888	1,020
	Φύλλο=0	-,555	,150	13,626	1	,000	,574	,427	,771
	Φύλλο=1	0 ^b	.	.	0
	Γνωρίζω πολλά πράγματα για το Διαδίκτυο=0	-,206	,160	1,652	1	,199	,814	,595	1,114
	Γνωρίζω πολλά πράγματα για το Διαδίκτυο =1	0 ^b	.	.	0
	Ηλικία=1	-,421	,154	7,507	1	,006	,656	,486	,887
	Ηλικία=2	0 ^b	.	.	0
	Εκπαιδευτικό επίπεδο γονέων=0	-,055	,147	,139	1	,709	,947	,710	1,262
	Εκπαιδευτικό επίπεδο γονέων=1	0 ^b	.	.	0
	Εθιστική συμπεριφορά στο Διαδίκτυο =0	-1,559	,217	51,658	1	,000	,210	,137	,322
	Εθιστική συμπεριφορά στο Διαδίκτυο =1	-,666	,221	9,111	1	,003	,514	,333	,792
	Εθιστική συμπεριφορά στο Διαδίκτυο =2	0 ^b	.	.	0
	Βαριά καθημερινή χρήση SNS=0	-,967	,153	39,806	1	,000	,380	,282	,514
	Βαριά καθημερινή χρήση SNS=1	0 ^b	.	.	0

a. The reference category is: 1.

b. This parameter is set to zero because it is redundant.

$$\text{Log}(p_{i2}/p_{i1}) = 0,785 - 0,093x_1 - 0,096x_2 - 0,59x_3 - 0,451x_4$$

$$\text{Log}(p_{i3}/p_{i1}) = 1,183 - 0,555x_1 - 0,421x_2 - 1,559x_3 - 0,967x_4$$

Ο Πίνακας 15 (Γραφήματα 21-24) παρουσιάζει τον τρόπο εξάρτησης της μεταβλητής «Γνωριμία μέσω Διαδικτύου» με τις μεταβλητές «Φύλο», «Ηλικία» «Εθιστική συμπεριφορά στο Διαδίκτυο» και «Βαριά καθημερινή χρήση του SNS».

Προκύπτει ότι όσον αφορά το φύλο, μελετώντας τους άντρες, το ποσοστό των ατόμων που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και τον συνάντησαν (23,3%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (18,8%). Αντίθετα, στις γυναίκες, το ποσοστό των ατόμων που δεν είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή (55,4%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (51%).

Σχετικά με την ηλικία, μελετώντας τους συμμετέχοντες ηλικίας 16-17,9 ετών, το ποσοστό των ατόμων που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και τον συνάντησαν (22,8%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (18,8%).

Αναφορικά με την εθιστική συμπεριφορά στο διαδίκτυο, προκύπτει ότι στους συμμετέχοντες που δεν έχουν ενδείξεις κινδύνου για εθιστική συμπεριφορά, το ποσοστό των ατόμων που δεν είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή (61,3%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (51%). Επιπλέον, στους συμμετέχοντες που έχουν ήπιες ενδείξεις κινδύνου για εθιστική συμπεριφορά το ποσοστό των ατόμων που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και δεν τον συνάντησαν (37,8%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (30,3%). Τέλος, στους συμμετέχοντες με έντονη ένδειξη κινδύνου για εθιστική συμπεριφορά, το ποσοστό των ατόμων που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και τον συνάντησαν (41,5%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (18,8%).

Αναφορικά με την βαριά καθημερινή χρήση του SNS, προκύπτει ότι στους συμμετέχοντες, που το χρησιμοποιούν λιγότερο από 2 ώρες, το ποσοστό των ατόμων που δεν είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή (59,3%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (51%). Αντίθετα, στους συμμετέχοντες που χρησιμοποιούν το SNS περισσότερο από

2 ώρες, το ποσοστό των ατόμων που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και τον συνάντησαν (30,2%) είναι μεγαλύτερο σε σύγκριση με το συνολικό ποσοστό στο δείγμα (18,8%).

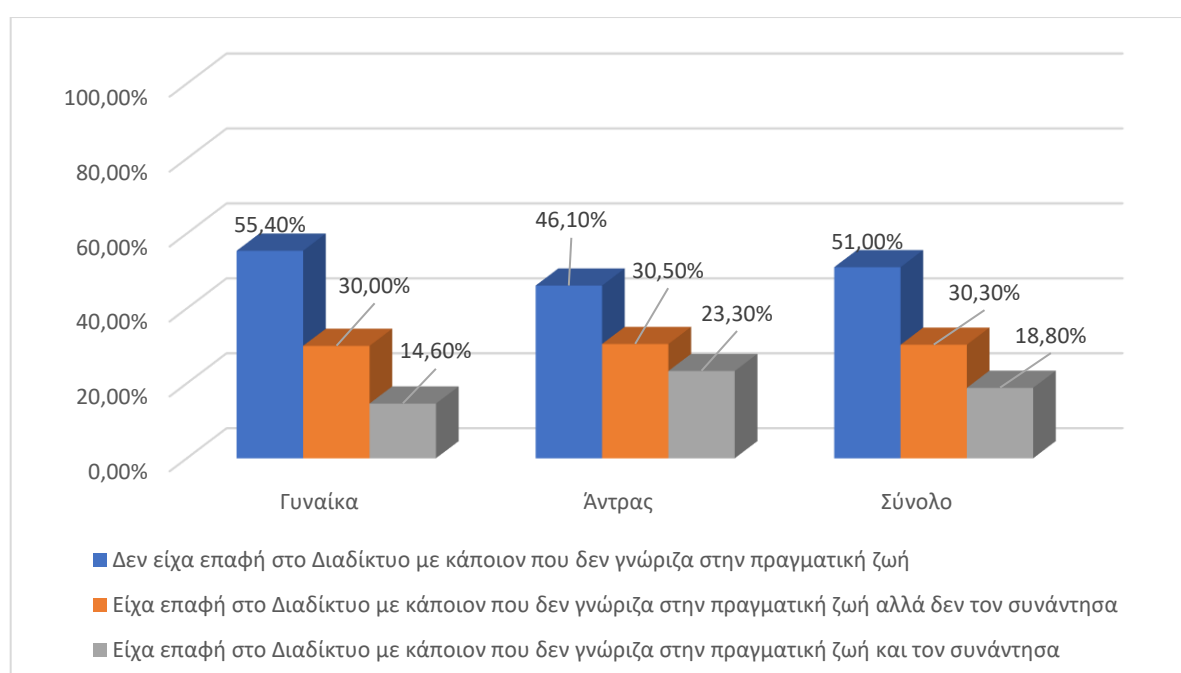
Τέλος, στον Πίνακα 16 παρουσιάζονται τα ποσοστά της σωστής πρόβλεψης του μοντέλου. Προκύπτει ότι το μοντέλο έχει σωστή πρόβλεψη στο 53,9% των περιπτώσεων, με το ποσοστό να είναι υψηλότερο στα άτομα που δεν είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή (85,3%) και χαμηλότερο στα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή και δεν τον συνάντησαν (18,5%) ή τον συνάντησαν (26,2%)

Πίνακας 15: Στατιστικά σημαντικές εξαρτήσεις για την μεταβλητή «Γνωριμία μέσω Διαδικτύου»

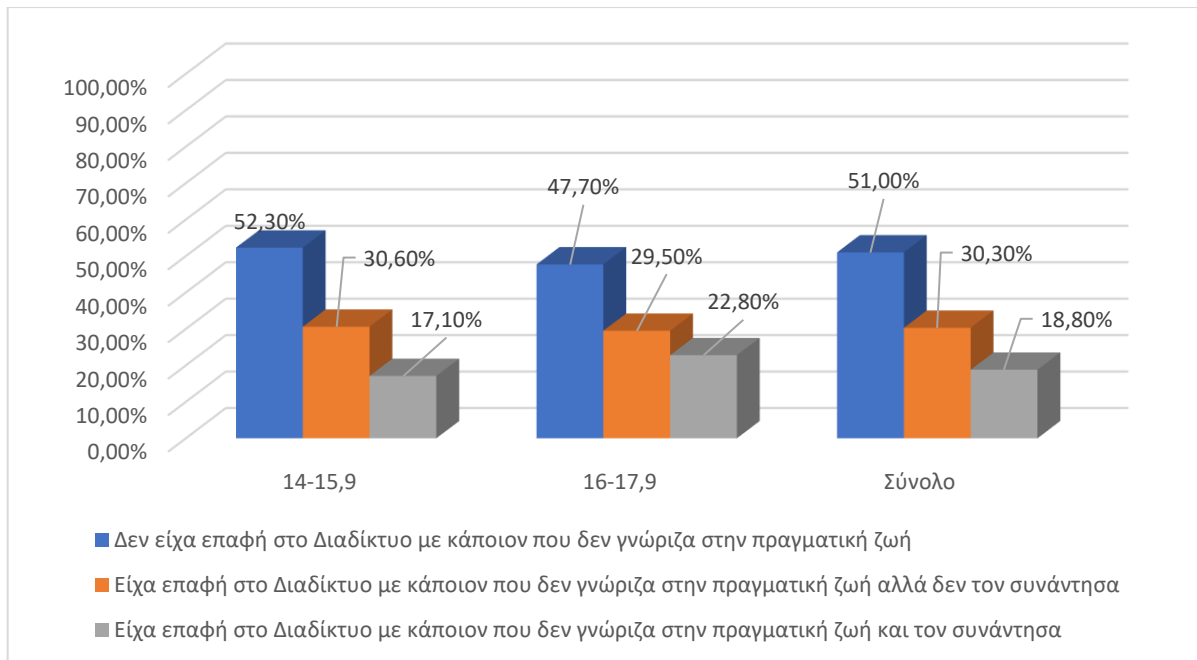
			Γνωριμία μέσω Διαδικτύου		
			Δεν είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή	Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή αλλά δεν τον συνάντησα	Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή και τον συνάντησα
Φύλο	Γυναίκα	N	470	255	124
		%	55,4%	30,0%	14,6%
	Άνδρας	N	358	237	181
		%	46,1%	30,5%	23,3%
	Σύνολο	N	828	492	305
		%	51,0%	30,3%	18,8%
Ηλικία	14-15,9	N	600	351	196
		%	52,3%	30,6%	17,1%
	16-17,9	N	228	141	109
		%	47,7%	29,5%	22,8%
	Σύνολο	N	828	492	305
		%	51,0%	30,3%	18,8%
Εθιστική συμπεριφορά στο Διαδίκτυο	Δεν υπάρχουν ενδείξεις	N	589	258	114
		%	61,3%	26,8%	11,9%
	Ήπιες ενδείξεις	N	183	180	113
		%	38,4%	37,8%	23,7%
	Σε κίνδυνο/ IAB	N	56	54	78
		%	29,8%	28,7%	41,5%
Σύνολο	N	828	492	305	
	%	51,0%	30,3%	18,8%	
Βαριά καθημερινή ή χρήση του SNS	Δεν το χρησιμοποιώ <2 ώρες	N	618	295	129
		%	59,3%	28,3%	12,4%
	>= 2 ώρες/ημέρα	N	210	197	176
		%	36,0%	33,8%	30,2%
	Σύνολο	N	828	492	305

Πίνακας 16: Αποτελέσματα πρόβλεψης μοντέλου πολυωνμικής λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή την «Γνωριμία μέσω Διαδικτύου»

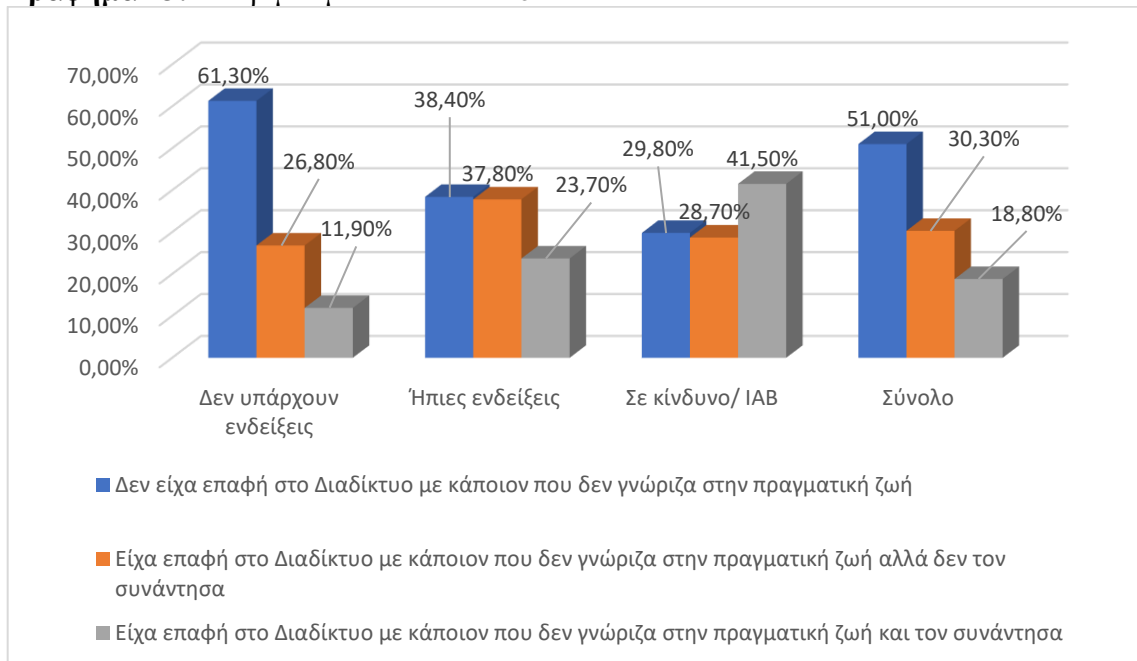
«Γνωριμία μέσω Διαδικτύου»	Σωστά	Λάθος	%
Δεν είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή	706	122	85,3
Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή αλλά δεν τον συνάντησα	91	401	18,5
Είχα επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζα στην πραγματική ζωή και τον συνάντησα	80	225	26,2
Συνολικό	877	748	53,9



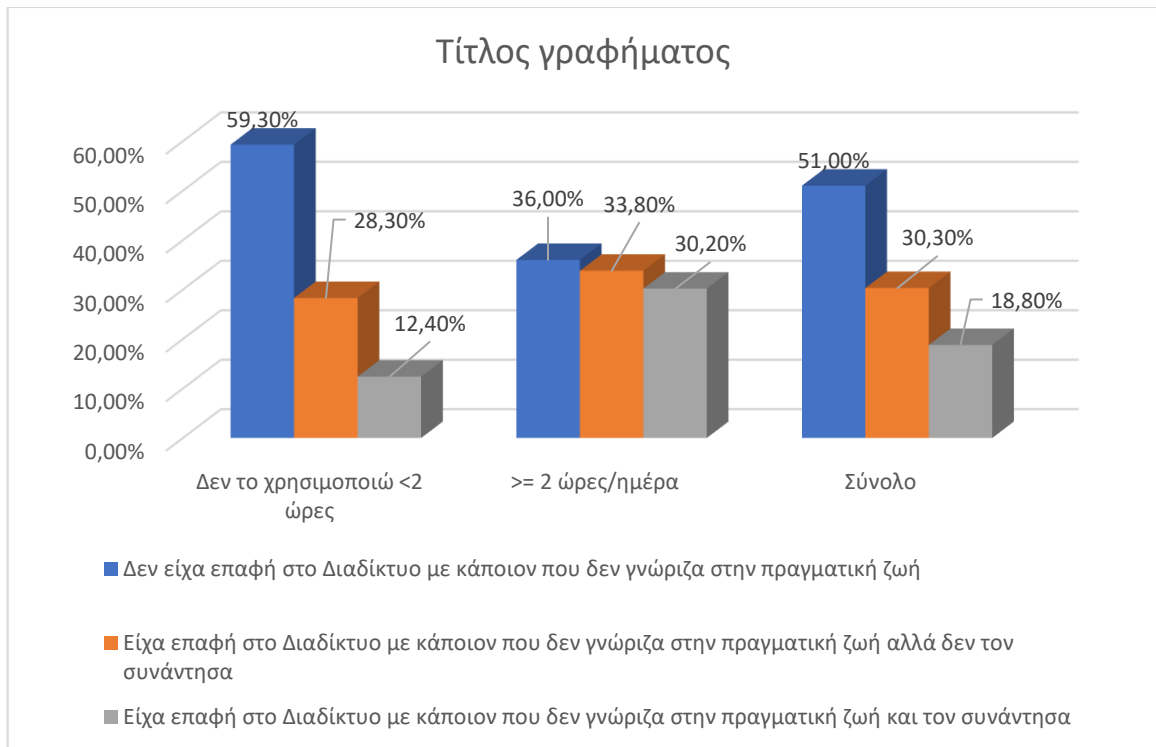
Γράφημα 22: Γνωριμία μέσω Διαδικτύου*Φύλο



Γράφημα 23: Γνωριμία μέσω Διαδικτύου*Ηλικία



Γράφημα 24: Γνωριμία μέσω Διαδικτύου* Εθιστική συμπεριφορά στο Διαδίκτυο



Γράφημα 25: Γνωριμία μέσω Διαδικτύου * Βαριά καθημερινή χρήση του SNS

Κεφάλαιο 5. Συμπεράσματα

Αναφορικά με την έρευνα που χρησιμοποιήσαμε για να ελέγξουμε τις μεθόδους που χρησιμοποιήσαμε συμμετείχαν 1625 άτομα, σχεδόν ισόποσα κατανομημένα ως προς το φύλο, με την πλειοψηφία να είναι ηλικίας 14-16 ετών και άτομα που γνωρίζουν αρκετά καλά το Διαδίκτυο, χωρίς ιδιαίτερη εθιστική συμπεριφορά και χωρίς υπερβολική χρήση του SNS. Οι ερωτηθέντες δήλωσαν πως ήταν κατά μέσο όρο 11 ετών όταν ξεκίνησαν να χρησιμοποιούν το Διαδίκτυο. Οι μισοί συμμετέχοντες ανέφεραν πως είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στην πραγματική ζωή, ενώ η μειοψηφία πραγματοποίησε συνάντηση με κάποιο άτομο το οποίο το γνώρισε για πρώτη φορά στο Διαδίκτυο. Το εκπαιδευτικό επίπεδο των γονέων στις περισσότερες περιπτώσεις χαρακτηρίστηκε υψηλό.

Στο 1^ο ερευνητικό ερώτημα μελετήθηκαν ποιοι παράγοντες επηρεάζουν την επαφή στο Διαδίκτυο με κάποιο άτομο που δεν υπάρχει γνωριμία στο παρελθόν στην πραγματική ζωή. Τα αποτελέσματα έδειξαν πως οι άντρες, άτομα με καλή γνώση του Διαδικτύου, με μέτρια ή υψηλή εθιστική συμπεριφορά το Διαδίκτυο και βαριά καθημερινή χρήση του SNS άνω των 2 ωρών, είναι πιο πρόθυμοι να κάνουν επαφή στο Διαδίκτυο με κάποιο άτομο που δεν υπάρχει γνωριμία στο παρελθόν στην πραγματική ζωή.

Στο 2^ο ερευνητικό ερώτημα μελετήθηκε ποιοι παράγοντες επηρεάζουν την συνάντηση με άτομο με το οποίο υπήρξε γνωριμία για πρώτη φορά στο Διαδίκτυο. Τα αποτελέσματα έδειξαν πως οι άντρες, άτομα ηλικίας 16-18 ετών, με μέτρια ή υψηλή εθιστική συμπεριφορά το Διαδίκτυο και βαριά καθημερινή χρήση του SNS άνω των 2 ετών, είναι πιο πρόθυμοι να πραγματοποιήσουν συνάντηση με άτομο με το οποίο υπήρξε γνωριμία για πρώτη φορά στο Διαδίκτυο.

Στο 3^ο ερευνητικό ερώτημα, μελετήθηκε ποιοι παράγοντες επηρεάζουν την συνάντηση, για τα άτομα που είχαν επαφή στο Διαδίκτυο με κάποιον που δεν έχουν γνωρίσει στο παρελθόν στην πραγματική ζωή. Τα αποτελέσματα έδειξαν πως οι άντρες, άτομα με υψηλή εθιστική συμπεριφορά το Διαδίκτυο και βαριά καθημερινή χρήση του SNS άνω των 2 ετών, είναι πιο πρόθυμοι να πραγματοποιήσουν συνάντηση με άτομο με το οποίο υπήρξε γνωριμία για πρώτη φορά στο Διαδίκτυο.

Στο 4^ο ερευνητικό ερώτημα μελετήθηκαν ποιοι παράγοντες επηρεάζουν την γνωριμία μέσω Διαδικτύου μέχρι την συνάντηση. Οι μισοί συμμετέχοντες πραγματοποίησαν επαφή στο Διαδίκτυο με κάποιον που δεν γνώριζαν στην πραγματική ζωή, ωστόσο μόνο η μικρή πλειοψηφία πραγματοποίησε συνάντηση. Τα αποτελέσματα έδειξαν ότι οι άντρες, τα άτομα ηλικίας 16-18 ετών, άτομα με υψηλή εθιστική συμπεριφορά, θα πραγματοποιούσαν ευκολότερα συνάντηση με κάποιο άτομο που το γνώρισαν στο Διαδίκτυο, ενώ αντίθετα οι γυναίκες και τα άτομα με μικρή εθιστική συμπεριφορά στο Διαδίκτυο και μικρή χρήση του SNS με μεγαλύτερη ευκολία δεν θα πραγματοποιούσαν ούτε επαφή στο Διαδίκτυο με κάποιον που δεν γνωρίζουν στην πραγματική ζωή. Τα άτομα με ήπια εθιστική συμπεριφορά στο Διαδίκτυο θα πραγματοποιούσαν ευκολότερα επαφή στο Διαδίκτυο με άτομο το οποίο δεν γνωρίζουν στην πραγματική ζωή.

Από τα παραπάνω αποτελέσματα καθώς επίσης και από τα στατιστικά συμπεράσματα αναφορικά με την χρήση της λογιστικής παλινδρόμησης σε ένα πιο διευρυμένο πλαίσιο χρησιμοποίησης αυτής, σε περιπτώσεις όπου για παράδειγμα όταν θέλουμε να εξετάσουμε μεταβλητές που παίρνουν παραπάνω από δύο τιμές, σε σύγκριση με το μοντέλο της πολυωνυμικής παλινδρόμησης το οποίο είναι το κατ'εξοχήν μοντέλο που χρησιμοποιείται σε αυτές τις περιπτώσεις, θέλοντας να διαπιστώσουμε την αποτελεσματικότητα των δυο μοντέλων, προκύπτουν κάποια πολύ ενδιαφέροντα συμπεράσματα.

Αρχικά παρατηρούμε ότι και με τις δυο μεθόδους καταλήγουμε στις ίδιες ανεξάρτητες μεταβλητές οι οποίες είναι στατιστικά σημαντικές για την ερευνά μας. Σε αυτό το κομμάτι παρατηρούμε ότι δεν υπάρχει κάποια ουσιαστική διαφορά ανάμεσα στα δυο μοντέλα. Η ουσιαστική διαφορά ανάμεσα στα δυο μοντέλα παρατηρείται στα ποσοστά των αποτελεσμάτων σωστής πρόβλεψης των δυο μοντέλων. Έτσι, τα ποσοστά της δίτιμης λογιστικής παλινδρόμησης παρατηρούμε ότι με τον τρόπο που αυτή έχει χρησιμοποιηθεί κυμαίνονται σε αρκετά υψηλά επίπεδα. Στον αντίποδα, παρατηρούμε ότι το ποσοστό στην πολυωνυμική παλινδρόμηση κυμαίνεται κοντά στο 50%, για την ακρίβεια 53,6%. Πράγμα που σημαίνει ότι μια στις δυο προβλέψεις θα είναι λάθος.

Επομένως, καταλήγουμε στο συμπέρασμα ότι το μοντέλο της δίτιμης λογιστικής παλινδρόμησης μπορεί να επεκταθεί και να χρησιμοποιηθεί σε περιπτώσεις

μεταβλητών που έχουν παραπάνω από δυο τιμές, με συνδυαστικό τρόπο, έτσι ώστε να εξάγουμε πιο έγκυρα αποτελέσματα στις έρευνές μας.

Βιβλιογραφικές αναφορές

Ελληνική

- Γαλάνης, Π. (2012). Εγκυρότητα και αξιοπιστία των ερωτηματολογίων στις επιδημιολογικές μελέτες. *Εφαρμοσμένη Ιατρική Έρευνα*.
- Ρούσσος, Π. Λ., & Τσαούσης, Γ. (2011). *Στατιστική στις Επιστήμες της Συμπεριφοράς με τη Χρήση του SPSS*, 1η Έκδοση. Αθήνα: Τόπος
- Σταλίκας, Α. & Κυριάζος, Θ. (2019). *Μεθοδολογίας έρευνας και στατιστική*. Εκδόσεις: Τόπος.

Ξενόγλωσση

- Amichai- Hamburger, Y. (2013). *The social net: Understanding of our online behavior*. Oxford scholarship online. Doi: 10.1093/acprof:oso/9780199639540.001.0001
- Bailenson, J. N., Yee, N., Blascovich, J., & Guadagno, R. E. (2008). Transformed social interaction in mediated interpersonal communication. Στο Konijn, S. Utz, M. Tanis, & S. B. Barnes. *Mediated Interpersonal Communication* (p. 77–99). New York, NY: Routledge.
- Boase, J. & Wellman, B. (2004). Personal Relationships: On and Off the Internet. *The Cambridge Handbook of Personal Relationships*, edited by Vangelisti, A. L. & Perlman, D., Cambridge, 709-723.
- Brym, R. J., & Lenton, R. L. (2001). Love online: A report on digital dating in Canada.
- Campbell-Kelly, M. & Garcia-Swartz, D.D. (2013). The History of the Internet: The Missing Narratives. *Journal of information technology*, 28(1), 18-33. <https://doi.org/10.1057/jit.2013.4>
- Cohen-Almagor R. (2011). Internet History. *International Journal of Techno ethics*, 2(2), 45-64. Doi: 10.4018/978-1-4666-2931-8.ch002.
- Daft, R. & Lengel, R. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554–571. <https://doi.org/10.1287/mnsc.32.5.554>
- Donn, J.E., & Sherman, R.C. (2002). Attitudes and Practices Regarding the Formation of Romantic Relationships on the Internet. *Cyber Psychology & Behavior*, 5(2):107-23. Doi: 10.1089/109493102753770499
- Douglas, K. M. & McGarty, C. (2002). Internet identifiability and beyond: A model of the effects of identifiability on communicative behavior. *Group Dynamics: Theory, Research, and Practice*, 6(1), 17–26. <https://doi.org/10.1037/1089-2699.6.1.17>
- Dunn, R. A., & Guadagno, R. E. (2012). My avatar and me? Gender and personality predictors of avatar-self-discrepancy. *Computers in Human Behavior*, 28, 97 – 106. <https://doi.org/10.1016/j.chb.2011.08.015>

- Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., Sprecher, S.(2012). Online Dating: A Critical Analysis From the Perspective of Psychological Science. *Psychological Science in the Public Interest*, 13(1), 3-66. <https://doi.org/10.1177/1529100612436522>
- Gibbs, J. L., Ellison N. B., & Heino, R. D. (2006). Self-Presentation in Online Personals. The Role of Anticipated Future Interaction, Self-Disclosure, and Perceived Success in Internet Dating. *Communication Research (COMMUN RES)*. 33(2), 152-177. Doi: 10.1177/0093650205285368
- Gleason, M. E. J., Green, A. S. , & McKenna, K. Y. A. (2002). Relationship Formation on the Internet: What's the Big Attraction? *Journal of social issues*, 58(1), 9-31. <https://doi.org/10.1111/1540-4560.00246>
- Guadagno, R. E., Okdie, B. M., & Kruse, S. A. (2011). Dating deception: Gender, online dating, and exaggerated self-presentation. *Computers in Human Behavior* 28:642-647. Doi: 10.1016/j.chb.2011.11.010
- Hancock, J. T. & Dunham, P. J. (2001). Impression formation in computer-mediated communication revisited: An analysis of the breadth and intensity of impressions. *Communication Research*, 28, 325-347. <https://doi.org/10.1177/009365001028003004>
- Hitsch, G. J., Hortacsu, A., Ariely, D. (2009). *What makes you click: An empirical analysis of online dating*. Unpublished manuscript, University of Chicago. http://home.uchicago.edu/~hortacsu/online_dating_feb2005.pdf
- Hook, M., Gerstein, L., Detterich, L., & Gridley, B. (2003). How close Are we? Measuring intimacy and examining gender differences. *Journal of Counseling & Development*, 81 (4), 462-472. <https://doi.org/10.1002/j.1556-6678.2003.tb00273.x>
- Jensen-Campbell, L. A., Graziano, W. G., & West, S. G. (1996). Dominance, prosocial orientation, and female preferences: Do nice guys really finish last? *Journal of Personality and Social Psychology*, 68, 427-440. <https://doi.org/10.1037/0022-3514.68.3.427>
- Jiang, L. C., Bazarova, N., & Hancock, J. T. (2011). The disclosure-intimacy link in computer-mediated communication: An attributional extension of the hyper personal model. *Human Communication Research*, 37, 58-77. <https://doi.org/10.1111/j.1468-2958.2010.01393.x>
- Leiner, B.M., Kahn, R.E., Postel, J., Cerf, V.G., Kleinrock, L., Roberts, L.G. & Wolff, S. (2009). A Brief History of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22-31. <https://doi.org/10.1145/1629607.1629613>
- Menard, S. (2010). *Logistic Regression: From Introductory to Advanced Concepts and Applications*. Thousand Oaks, CA: Sage.
- Rains, S. A. (2007). The impact of anonymity on perceptions of source credibility and influence in computer-mediated group communication; a test of two competing hypotheses. *Communication Research*, 34(1), 100-125. <https://doi.org/10.1177/0093650206296084>

- Riva, G. (2002). The socio cognitive psychology of computer-mediated communication: The present and future of technology-based interactions. *Cyber Psychology & Behavior*, 5 (6), 581-598. <https://doi.org/10.1089/109493102321018222>
- Salkind, N. J. (2010). *Encyclopedia of research design*. Thousand Oaks, CA: Sage Publications.
- Segovia, K.Y & Bailenson, J.N (2013). Identity Manipulation - What Happens When Identity Presentation is Not Truthful. ΣτοAmichai- Hamburger, Y. *The social net: Understanding of our online behavior*. Oxford scholarship online (pp. 45-61). Doi: 10.1093/acprof: oso/9780199639540.001.0001
- Shin, D.H. (2010). Analysis of online social networks: a cross-national study. *Online Information Review*, 34 (3), 473-495. <https://doi.org/10.1108/14684521011054080>
- Snow, E. (2007). Intimacy and Face-to-Face versus Computer Interaction. *Undergraduate Review*, 3, 37-50. http://vc.bridgew.edu/undergrad_rev/vol3/iss1/9
- Sproull, L. & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, 32, 1492–1512. <https://doi.org/10.1287/mnsc.32.11.1492>
- Suler, J. (2004). The online disinhibition effect. *Cyber Psychology and Behavior*, 7, 321–326. <https://doi.org/10.1089/1094931041291295>
- Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin*, 34, 1023–1036. https://journals.sagepub.com/doi/pdf/10.1177/0146167208318067?casa_token=ZRFie2gHO4wAAAAA:kRWEpxNVOcICkb3QyckbJCHRbgetJS4UAZvk1qpfYdyHPJB0Gj0Q7ojaJ3c7pkz6a_vmiWJ69Lmb4w
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, 3-44. <https://doi.org/10.1177/009365096023001001>
- Walther, J. B., Slovacek, C., & Tidwell, L. (2001). Is a picture worth a thousand words? Photographic images in long-term and short-term computer-mediated communication. *Communication Research*, 28, 105. <https://doi.org/10.1177/009365001028001004>
- Whitty, M. T. (2008). Liberating or debilitating? An examination of romantic relationships, sexual relationships, and friendships on the net. *Computers in Human Behavior*, 24(5), 1837–1850. <https://doi.org/10.1016/j.chb.2008.02.009>
- Whitty, M. T. (2008). Revealing the ‘real’ me, searching for the ‘actual’ you: presentations of self on an internet dating site. *Computers in Human Behavior*, 24 (4),1707-1723. <http://wrap.warwick.ac.uk/83756>
- Yip, C., Han, N. R., & Sng, B. L. (2016). Legal and ethical issues in research. *Indian Journal of Anaesthesia*, 60(9), 684-688.