



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Πολιτικών Μηχανικών

Τομέας Μεταφορών & Συγκοινωνιακής Υποδομής

**Προτυποποίηση της Επιρροής της Χρήσης Κινητού Τηλεφώνου
στη Συμπεριφορά του Οδηγού
Αξιοποιώντας Λεπτομερή Δεδομένα
από Αισθητήρες Έξυπνων Κινητών Τηλεφώνων**

Διπλωματική Εργασία



Αναστασία Αργυροπούλου

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2017

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον κ. Γ. Γιαννή, Καθηγητή της Σχολής Πολιτικών Μηχανικών ΕΜΠ για την εμπιστοσύνη που μου έδειξε με την ανάθεση της παρούσας Διπλωματικής Εργασίας, για την υποστήριξη του καθώς και για την εκπληκτική συνεργασία που είχαμε σε όλα τα στάδια της. Επιπλέον, οφείλω να τον ευχαριστήσω για τη σωστή καθοδήγηση αλλά και για τις γενικότερες γνώσεις που μου μετέδωσε.

Παράλληλα, θα ήθελα να ευχαριστήσω τα υπόλοιπα μέλη της εξεταστικής επιτροπής: κ. Ι. Γκόλια, Πρύτανη του Εθνικού Μετσόβιου Πολυτεχνείου και Καθηγητή της Σχολής Πολιτικών Μηχανικών ΕΜΠ και κ. Ε. Βλαχογιάννη, Επίκουρο Καθηγήτρια της Σχολής Πολιτικών Μηχανικών ΕΜΠ, για τις εποικοδομητικές παρατηρήσεις τους αλλά κυρίως για τις γνώσεις που μου προσέφεραν καθ' όλη τη διάρκεια της φοίτησής μου στη σχολή.

Οφείλω ένα μεγάλο ευχαριστώ στον κ. Δημήτρη Τσελέντη, Υποψήφιο Διδάκτορα ΕΜΠ, για τις παρατηρήσεις και τις συμβουλές του κατά τη διάρκεια εκπόνησης της παρούσας Διπλωματικής Εργασίας. Χωρίς τη βοήθειά του δε θα τα είχα καταφέρει ως εδώ. Εξίσου ευχαριστώ και την κ. Ελεονώρα Παπαδημητρίου, Δρ. Συνεργάτη- Ερευνήτρια ΕΜΠ, για τη βοήθεια και τις υποδείξεις της σε σημαντικά ζητήματα της εργασίας.

Επίσης, ιδιαίτερες ευχαριστίες ανήκουν και στην εταιρεία [OSeven](#), η οποία παραχώρησε τα πολύ χρήσιμα στοιχεία μέτρησης της συμπεριφοράς οδηγού που χρησιμοποιήθηκαν στην παρούσα Διπλωματική Εργασία.

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου για την αγάπη και την ηθική, αλλά και υλική, υποστήριξή τους σε όλη τη διάρκεια των σπουδών μου. Τέλος, ευχαριστώ όλους τους φίλους μου για τις πολύ όμορφες αλλά και δύσκολες στιγμές που περάσαμε μαζί όλα αυτά τα χρόνια.

Αθήνα, Ιούλιος 2017

Αναστασία Αργυροπούλου

Περίληψη

Προτυποποίηση της επιρροής της Χρήσης Κινητού Τηλεφώνου στη Συμπεριφορά του Οδηγού Αξιοποιώντας Λεπτομερή Δεδομένα από Αισθητήρες Έξυπνων Κινητών Τηλεφώνων

Αναστασία Αργυροπούλου

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής ΕΜΠ

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η διερεύνηση και η προτυποποίηση της επιρροής της χρήσης κινητού τηλεφώνου στη συμπεριφορά του οδηγού αξιοποιώντας λεπτομερή δεδομένα από αισθητήρες έξυπνων κινητών τηλεφώνων. Για το σκοπό αυτό αναλύονται στοιχεία που συλλέχθηκαν από 100 οδηγούς οι οποίοι συμμετείχαν σε πείραμα οδήγησης σε πραγματικές συνθήκες για χρονικό διάστημα τεσσάρων μηνών. Η ανάλυση πραγματοποιήθηκε με τη χρήση των στατιστικών μεθόδων της μεικτής γραμμικής και δυαδικής λογιστικής παλινδρόμησης. Μέσω των μοντέλων γραμμικής παλινδρόμησης εξετάστηκε κατά πόσο χαρακτηριστικά οδήγησης που καταγράφηκαν από τους αισθητήρες έξυπνων κινητών τηλεφώνων επηρεάζουν και επομένως μπορούν να προβλέψουν τα απότομα συμβάντα κατά την οδήγηση, ενώ τα μοντέλα λογιστικής παλινδρόμησης αξιοποιήθηκαν για την πρόβλεψη της χρήσης ή μη του κινητού τηλεφώνου κατά την οδήγηση μέσα από την καταγραφή μεγεθών σχετικών με την οδήγηση. Από την εφαρμογή των μοντέλων προέκυψε ότι οι παράγοντες που επηρεάζουν τα απότομα συμβάντα είναι πέντε, με κυριότερο τη μέση ταχύτητα οδήγησης, ενώ οι αντίστοιχοι παράγοντες που επηρεάζουν την πιθανότητα χρήσης του κινητού τηλεφώνου στην οδήγηση είναι έξι, με κυριότερο εκείνον της μέσης στροφικής ταχύτητας.

Λέξεις-Κλειδιά: Απόσπαση προσοχής, κινητό τηλέφωνο, έξυπνα κινητά τηλέφωνα, απότομα συμβάντα, οδική ασφάλεια, μεικτή γραμμική παλινδρόμηση, μεικτή δυαδική λογιστική παλινδρόμηση

Abstract

Modelling mobile phone use impact on driver behaviour through the exploitation of data from smartphone sensors

Anastasia Argyropoulou

Supervisor: George Yannis, Professor NTUA

The aim of this Diploma Thesis is to examine and model the impact of mobile phone use on driver behaviour through the exploitation of data from smartphone sensors. To achieve this objective, data collected from 100 drivers who participated at a naturalistic driving experiment for four months are analyzed. The analysis was carried out with the use of statistical methods of mixed-effects linear and binary logistic regression. Through the linear regression models it was examined whether driving characteristics recorded by smartphone sensors affect and can therefore predict harsh driving events, whereas the logistic regression models were used to predict the use or no use of mobile phone while driving through the recording of driving related measures. The application of the models revealed that the factors affecting the harsh events are five, with the average driving speed being the main one, while the factors affecting the possibility of using the mobile phone while driving are six, with the average angular speed being the main one.

Keywords: distraction, mobile phone, smartphones, harsh events, road safety, mixed-effects linear regression, mixed-effects binary logistic regression

Εκτενής Περίληψη

Αντικείμενο της παρούσας Διπλωματικής Εργασίας αποτέλεσε η διερεύνηση και η **προτυποποίηση της επιρροής της χρήσης κινητού τηλεφώνου στη συμπεριφορά του οδηγού αξιοποιώντας λεπτομερή δεδομένα από αισθητήρες έξυπνων κινητών τηλεφώνων.**

Συγκεκριμένα, εξετάστηκαν **δύο επιμέρους στόχοι**. Πρώτον, κατά πόσο χαρακτηριστικά οδήγησης που καταγράφηκαν από τους αισθητήρες έξυπνων κινητών τηλεφώνων επηρεάζουν τα απότομα συμβάντα κατά την οδήγηση και δεύτερον, η δυνατότητα πρόβλεψης της χρήσης ή μη του κινητού τηλεφώνου κατά την οδήγηση μέσα από την καταγραφή σχετικών μεγεθών.

Μετά τον καθορισμό του επιδιωκόμενου στόχου, ξεκίνησε η **βιβλιογραφική αναζήτηση** ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας σε χώρες της Ευρωπαϊκής Ένωσης, αλλά και παγκοσμίως, καθώς και των διαθέσιμων στοιχείων που ήταν απαραίτητα για τη συγκεκριμένη διερεύνηση.

Αφού μελετήθηκαν τα στοιχεία που συγκεντρώθηκαν από τη βιβλιογραφική ανασκόπηση, σειρά είχε η ανάλυση της βάσης δεδομένων που συλλέχθηκαν από 100 οδηγούς σε **πείραμα σε πραγματικές συνθήκες οδήγησης** μέσω των αισθητήρων έξυπνων κινητών τηλεφώνων. Κατά τη στατιστική επεξεργασία αποφασίστηκε η εξαγωγή μαθηματικών μοντέλων τόσο για το σύνολο των οδικών δικτύων (αστικό περιβάλλον, υπεραστικό περιβάλλον, αυτοκινητόδρομοι) όσο και ξεχωριστά για κάθε τύπο οδού.

Για την εκπλήρωση του πρώτου επιμέρους στόχου αναπτύχθηκαν μοντέλα **γραμμικής παλινδρόμησης** καθώς η εξαρτημένη μεταβλητή (αριθμός συμβάντων ανά απόσταση) είναι συνεχής μεταβλητή. Για την εκπλήρωση του δεύτερου επιμέρους στόχου αναπτύχθηκαν μοντέλα **δυσιαδικής λογιστικής παλινδρόμησης** καθώς η εξαρτημένη μεταβλητή παίρνει δύο μόνο τιμές (0 ή 1) ανάλογα αν ο οδηγός χρησιμοποιεί ή όχι το κινητό τηλέφωνο. Σημειώνεται ότι καθότι υπάρχουν πολλές μετρήσεις-διαδρομές που αντιστοιχούν σε κάθε οδηγό εφαρμόζονται μοντέλα μεικτών επιδράσεων (**mixed-effects models**) ή αλλιώς μεικτά μοντέλα ώστε να αποφευχθούν σφάλματα λόγω της επαναληπτικότητας των δεδομένων.

Από την παραπάνω επεξεργασία προέκυψαν συγκεντρωτικά τα ακόλουθα μοντέλα που συνοψίζονται στον πίνακα που ακολουθεί

Πίνακας: Συγκεντρωτικός πίνακας στατιστικών μοντέλων

[illegible]

Τα γενικά συμπεράσματα που προέκυψαν από την εκπόνηση της παρούσας Διπλωματικής Εργασίας συνοψίζονται ως εξής:

- Στις έρευνες που έχουν διεξαχθεί έως τώρα παγκοσμίως αναφορικά με τη μελέτη της συμπεριφοράς των οδηγών, η συλλογή των δεδομένων γινόταν είτε μέσω του συστήματος διαγνωστικών του οχήματος (OBD) είτε μέσω άλλων πολύπλοκων συνδυασμών οργάνων καταγραφής που συνδέονταν με το όχημα (naturalistic driving, κλπ.) ή με προσομοιωτές οδήγησης και ερωτηματολόγια. Στην παρούσα Διπλωματική Εργασία εξετάζεται **για πρώτη φορά** η παρακολούθηση της συμπεριφοράς των οδηγών με **δεδομένα που συλλέγονται από αισθητήρες έξυπνων κινητών τηλεφώνων** που αφορούν στον οδηγό και όχι στο όχημα. Επιπλέον, επιχειρείται η διερεύνηση της επιθετικής οδήγησης (απότομα συμβάντα) και της απόσπασης προσοχής με επιπλέον μεταβλητές πέραν εκείνων που έχουν χρησιμοποιηθεί στην πλειονότητα παρόμοιων ερευνών.
- Με την ανάλυση διαπιστώθηκε ότι **είναι δυνατή η πρόβλεψη της χρήσης ή μη κινητού τηλεφώνου** με αρκετά υψηλή ακρίβεια κατά την οδήγηση εάν είναι γνωστά ορισμένα μεγέθη, όπως η μέση ταχύτητα οδήγησης, η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας, η μέση στροφική ταχύτητα, η διάρκεια χρήσης του κινητού τηλεφώνου, ο αριθμός απότομων συμβάντων και η περίοδος οδήγησης σε σχέση με την πρωινή και την απογευματινή αιχμή. Πιο συγκεκριμένα:
 - Η πρόβλεψη της χρήσης ή μη του κινητού τηλεφώνου πραγματοποιείται με **μεγαλύτερη ακρίβεια** κατά σειρά πρώτα για την οδήγηση σε **αυτοκινητόδρομο** και στη συνέχεια για υπεραστικό δίκτυο, αστικό δίκτυο και τέλος για το σύνολο των οδικών δικτύων.
 - Το μέγεθος που επηρεάζει περισσότερο σε σχέση με τους υπόλοιπους παράγοντες τη δυνατότητα πρόβλεψης της χρήσης ή μη του κινητού τηλεφώνου συνολικά για όλους τους τύπους οδού είναι **η μέση στροφική ταχύτητα**.
 - **Όσο μειώνεται η μέση ταχύτητα οδήγησης τόσο μεγαλύτερη γίνεται η πιθανότητα ο οδηγός να μιλάει στο κινητό τηλέφωνο.** Το ίδιο παρατηρείται για τη μέση στροφική ταχύτητα και τα απότομα συμβάντα τόσο για όλους τους τύπους οδού όσο και για αστικό και υπεραστικό περιβάλλον. Αυτό ενδεχομένως εξηγείται από το γεγονός ότι ο οδηγός προβαίνει σε **αντισταθμιστική συμπεριφορά** (risk compensation) και έτσι οδηγεί πιο προσεκτικά, με χαμηλότερες ταχύτητες και λιγότερο νευρικά προκειμένου να μειώσει τον κίνδυνο ατυχήματος.
 - Στον **αυτοκινητόδρομο**, παρατηρείται ότι η αύξηση των απότομων συμβάντων οδηγεί σε αύξηση της πιθανότητας ο οδηγός να μιλάει στο κινητό τηλέφωνο, ενδεχομένως διότι η χρήση του κινητού τηλεφώνου που αποσπά την προσοχή του οδηγού είναι μια βασική αιτία που μπορεί να οδηγήσει σε μεγαλύτερες διακυμάνσεις στη συμπεριφορά του οδηγού και περισσότερα απότομα συμβάντα σε ένα τέτοιο περιβάλλον.
 - **Όσο μεγαλώνει η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας συνολικά σε όλους τους τύπους οδού τόσο μειώνεται η πιθανότητα ο**

οδηγός να κάνει χρήση του κινητού τηλεφώνου. Αυτή η διαπίστωση ενδεχομένως εξηγείται καθώς ο οδηγός δε συνηθίζει να μιλάει στο κινητό τηλέφωνο σε υψηλές ταχύτητες εφόσον απαιτείται μεγαλύτερη συγκέντρωση στο έργο της οδήγησης.

- **Ο αριθμός των απότομων συμβάντων** ανά απόσταση οδήγησης εξαρτάται από την περίοδο οδήγησης σε σχέση με την πρωινή και την απογευματινή αιχμή, τη μέση ταχύτητα οδήγησης, τη διάρκεια χρήσης του κινητού τηλεφώνου και επομένως την απόσπαση προσοχής και τη μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας. Πιο συγκεκριμένα:
 - Το μέγεθος που επηρεάζει περισσότερο σε σχέση με τους υπόλοιπους παράγοντες την πραγματοποίηση απότομων συμβάντων είναι η **μέση ταχύτητα οδήγησης**.
 - Όσο αυξάνεται η **διάρκεια χρήσης του κινητού τηλεφώνου τόσο μειώνονται τα απότομα συμβάντα** (ανά απόσταση) ίσως για λόγους αντισταθμιστικής συμπεριφοράς του οδηγού.
 - Όσο αυξάνεται η **μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας συνολικά για όλους τους τύπους οδού τόσο αυξάνονται τα απότομα συμβάντα** ενδεχομένως επειδή ο οδηγός σε περίπτωση απρόσμενου γεγονότος (π.χ. στροφή, εμπόδιο) δεν έχει το χρόνο να επαναφέρει ομαλά την οδηγική του κατάσταση οπότε προβαίνει σε απότομες κινήσεις. Το ίδιο ισχύει και για την οδήγηση σε αστικό περιβάλλον.
 - Όσο αυξάνεται η **μέση ταχύτητα οδήγησης τόσο μειώνονται τα απότομα συμβάντα ανά διανυθείσα απόσταση**, γεγονός που ίσως σχετίζεται με την οδήγηση σε αυτοκινητόδρομο και στο υπεραστικό δίκτυο στον οποίο οι ταχύτητες είναι εν γένει μεγαλύτερες ενώ τα απότομα συμβάντα λιγότερα καθώς δεν υπάρχουν συχνές διασταυρώσεις, σηματοδότες κλπ.
 - Τέλος, όσο **μεγαλύτερη η απόσταση οδήγησης εντός αστικού, υπεραστικού δικτύου και αυτοκινητοδρόμου τόσο χαμηλότερος είναι ο αντίστοιχος αριθμός απότομων συμβάντων** ανά διανυθείσα απόσταση αφού ενδεχομένως ο οδηγός εξοικειώνεται με τις συνθήκες του περιβάλλοντος.
- Τα στατιστικά μοντέλα που αναπτύχθηκαν είναι κατάλληλα για τις συγκεκριμένες αναλύσεις και την επίτευξη των στόχων της Διπλωματικής Εργασίας αυτής. Πιο συγκεκριμένα, επιλέχθηκε η **γραμμική παλινδρόμηση** για την εξαρτημένη μεταβλητή "απότομα συμβάντα ανά απόσταση" επειδή είναι συνεχής, ενώ για την μεταβλητή που αναφέρεται στη χρήση ή μη του κινητού τηλεφώνου και η οποία είναι κατηγορική (0 ή 1) εφαρμόστηκε η **δυσιαδική λογιστική παλινδρόμηση**. Και στις δύο κατηγορίες μοντέλων, επειδή υπήρχαν επαναλαμβανόμενες μετρήσεις-διαδρομές για κάθε οδηγό, αναπτύχθηκαν μοντέλα **μεικτών επιδράσεων** (mixed-effects models).
- Τα σαφή και ερμηνεύσιμα αποτελέσματα, τα οποία μάλιστα συμφωνούν με τη διεθνή βιβλιογραφία καταδεικνύουν την **καταλληλότητα της μεθόδου** ανάλυσης

για την πρόβλεψη τόσο της πιθανότητας χρήσης του κινητού τηλεφώνου όσο και των απότομων συμβάντων κατά την οδήγηση. Η αποτελεσματικότητα της μεθόδου επιτρέπει την εφαρμογή της σε έρευνες με παρόμοιο αντικείμενο έπειτα από κατάλληλη προσαρμογή των συνθηκών. Βέβαια, οι μεταβλητές θα πρέπει να επιλέγονται ανάλογα με την εξεταζόμενη περίπτωση.

Περιεχόμενα

1. Εισαγωγή	1
1.1 Γενική Ανασκόπηση	1
1.1.1 Οδικά Ατυχήματα	1
1.1.2 Τα οδικά ατυχήματα στην Ευρώπη	1
1.1.3 Τα οδικά ατυχήματα στην Ελλάδα	2
1.1.4 Τα αίτια των οδικών ατυχημάτων	4
1.2 Στόχος της Διπλωματικής Εργασίας	6
1.3 Μεθοδολογία	7
1.4 Δομή	8
2. Βιβλιογραφική Ανασκόπηση	11
2.1 Εισαγωγή	11
2.2 Συναφείς έρευνες	11
2.2.1 The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model (2015)	11
2.2.2 Driving performance while using a mobile phone: A simulation study of Greek professional (2016)	12
2.2.3 Distracted driving and mobile phone use: Overview of impacts and countermeasures (2014)	13
2.2.4 Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments (2006)	14
2.2.5 Effects of naturalistic cell phone conversations on driving performance (2004)	15
2.2.6 Using mobile telephones: cognitive workload and attention resource allocation (2004)	16
Σκοπός	16
2.2.7 Impact of mobile phone use on car-following behaviour of young drivers (2015)	17
2.2.8 Driving and telephoning: Relative accident risk when using hand-held and hands-free mobile phones (2010)	17
2.2.9 Aggressive driving: an observational study of driver, vehicle, and situational variables (2004)	18
2.2.10 Taking a look behind the wheel: An investigation into the personality predictors of aggressive driving (2012)	19
2.3 Συναφείς μεθοδολογίες	20
2.3.1 Modeling the behavior of novice young drivers during the first year after licensure (2010)	20
2.3.2 Driving behavior and traffic safety: an acceleration-based safety evaluation procedure for smartphones (2014)	21

2.3.3 Development of a driving data recorder (1994).....	21
2.3.4 In-vehicle data recorders for monitoring and feedback on drivers' behavior (2008)	22
2.3.5 Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones (2011)	23
2.3.6 Innovative motor insurance schemes: A review of current practices and emerging challenges (2017)	24
2.4 Σύνοψη	24
3. Θεωρητικό Υπόβαθρο	27
3.1 Εισαγωγή.....	27
3.2 Βασικές έννοιες Στατιστικής	27
3.3 Βασικές Κατανομές	30
3.3.1 Κανονική Κατανομή	30
3.3.2 Κατανομή Poisson.....	30
3.3.3 Αρνητική Διωνυμική Κατανομή	31
3.4 Μαθηματικά Πρότυπα	31
3.4.1 Γραμμική Παλινδρόμηση	31
3.4.2 Λογιστική Ανάλυση Παλινδρόμησης	33
3.4.3 Μοντέλα πολλαπλών επιπέδων ή μεικτών επιδράσεων ή Μεικτά μοντέλα.....	35
3.5 Στατιστική Αξιολόγηση και Κριτήρια Αποδοχής Μοντέλου	36
4. Συλλογή και Επεξεργασία Στοιχείων.....	41
4.1 Εισαγωγή.....	41
4.2 Συλλογή Στοιχείων	41
4.2.1 Μεθοδολογία καταγραφής, συλλογής, επεξεργασίας και οπτικοποίησης δεδομένων	42
4.3 Επεξεργασία Στοιχείων	51
4.3.1 Προκαταρκτική Ανάλυση	51
4.3.2 Εισαγωγή βάσης δεδομένων στο ειδικό λογισμικό στατιστικής ανάλυσης, IBM SPSS 21.0	59
5. Εφαρμογή μεθοδολογίας- Αποτελέσματα	61
5.1 Εισαγωγή.....	61
5.2 Έλεγχος Συσχέτισης	61
5.3 Γραμμική παλινδρόμηση	63
5.3.1 Διαδικασία Γραμμικής Παλινδρόμησης	63
5.3.2 Δοκιμές Γραμμικής Παλινδρόμησης	65
5.4 Λογιστική παλινδρόμηση	65
5.4.1 Διαδικασία Λογιστικής παλινδρόμησης	65

5.4.2 Δοκιμές Δυναμικής Λογιστικής Παλινδρόμησης	67
5.5 Μοντέλα Μεικτών Επιδράσεων ή Μεικτά Μοντέλα (Mixed Models)	68
5.6 Mixed Linear Model	68
5.6.1 Διαδικασία Mixed Linear Model	68
5.6.2 Αποτελέσματα Mixed Linear Regression	71
5.7 Mixed Binary Logistic Model	83
5.7.1 Διαδικασία Mixed Binary Logistic Model	83
5.7.2 Αποτελέσματα Mixed Binary Logistic Regression	86
5.7.2.1 Χρήση ή μη κινητού τηλεφώνου σε όλους τους τύπους οδού.....	86
5.8 Ανάλυση ευαισθησίας	105
5.8.1 Mixed Γραμμική Παλινδρόμηση για το σύνολο των απότομων συμβάντων από όλους τους τύπους οδού	105
5.8.2 Mixed Δυναμική Λογιστική Παλινδρόμηση για την πιθανότητα χρήσης του κινητού τηλεφώνου κατά την οδήγηση.....	107
6. Συμπεράσματα	111
6.1 Σύνοψη Αποτελεσμάτων	111
6.3 Προτάσεις για βελτίωση της Οδικής Ασφάλειας.....	115
6.4 Προτάσεις για περαιτέρω έρευνα	115
7. Βιβλιογραφία	117
8. Παράρτημα.....	121

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 5.1.: Ελαστικότητες και σχετική επιρροή μεταβλητών για το γραμμικό μοντέλο	75
Πίνακας 5.2.: Ελαστικότητες και σχετική επιρροή μεταβλητών για το μοντέλο δυαδικής λογιστικής παλινδρόμησης	92
Πίνακας 6.1.: Συγκεντρωτικός πίνακας στατιστικών μοντέλων	112

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 3.1.: Επεξήγηση πιθανών τιμών του Συντελεστή Συσχέτισης.....	30
Εικόνα 3.2.: Παράδειγμα ευθείας ελαχίστων τετραγώνων.....	33
Εικόνα 5.3.: Σιγμοειδής καμπύλη λογιστικής παλινδρόμησης.....	34
Εικόνα 4.1.: Διάγραμμα διαχείρισης δεδομένων.....	42
Εικόνα 4.2.: Σύστημα ροής δεδομένων OSeven.....	43
Εικόνα 4.3.: Yaw, Pitch, Roll.....	45
Εικόνα 4.4.: Δείκτες κινδύνου οδήγησης.....	46
Εικόνα 4.5.: Εφαρμογή στο κινητό και διαδικτυακή πύλη.....	48
Εικόνα 4.6.: Τρόπος Εισαγωγής δεδομένων στο SPSS.....	59
Εικόνα 4.7.: Πεδίο Δεδομένων SPSS (data view).....	60
Εικόνα 4.8.: Πεδίο μεταβλητών SPSS (Variable View).....	60
Εικόνα 5.1.: Επιλογή εντολής για τον έλεγχο συσχέτισης των μεταβλητών.....	62
Εικόνα 5.2.: Επιλογή παραμέτρων για τον έλεγχο συσχέτισης των μεταβλητών.....	62
Εικόνα 5.3.: Παραδείγματα συσχέτισης μεταβλητών.....	63
Εικόνα 5.4.: Επιλογή εντολής γραμμικής παλινδρόμησης.....	64
Εικόνα 5.5.: Επιλογή ανεξάρτητων και εξαρτημένων μεταβλητών για τη γραμμική παλινδρόμηση.....	64
Εικόνα 5.6.: Σημείωση SPSS για την αφαίρεση της σταθεράς.....	65
Εικόνα 5.7.: Επιλογή εντολής δυαδικής λογιστικής παλινδρόμησης.....	66
Εικόνα 5.8.: Επιλογή ανεξάρτητων και εξαρτημένων μεταβλητών για τη δυαδική λογιστική παλινδρόμηση.....	66
Εικόνα 5.9.: Τρόπος αλλαγής της τιμής cut-off στο SPSS.....	67
Εικόνα 5.10.: Επιλογή εντολής μεικτής γραμμικής παλινδρόμησης.....	68
Εικόνα 5.11.: Επιλογή υποκειμένου επαναλαμβανόμενων μετρήσεων.....	69
Εικόνα 5.12.: Επιλογή εξαρτημένων και ανεξάρτητων μεταβλητών για τη μεικτή γραμμική παλινδρόμηση.....	69
Εικόνα 5.13.: Καθορισμός μεταβλητών με fixed effect.....	70
Εικόνα 5.14.: Καθορισμός random effect.....	70
Εικόνα 5.15.: Επιλογή στατιστικών ελέγχων.....	71
Εικόνα 5.16.: Χαρακτηριστικά του μοντέλου μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού.....	71
Εικόνα 5.17.: Κριτήρια ποιότητας μοντέλου μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού.....	72
Εικόνα 5.18.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού.....	72
Εικόνα 5.19.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού.....	73
Εικόνα 5.20.: Χαρακτηριστικά του μοντέλου της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον.....	76
Εικόνα 5.21.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον.....	76
Εικόνα 5.22.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον.....	77
Εικόνα 5.23.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον.....	77

Εικόνα 5.24.: Χαρακτηριστικά του μοντέλου της γραμμικής παλινδρόμησης μεικτών επιδράσεων για υπεραστικό περιβάλλον	78
Εικόνα 5.25.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον.....	79
Εικόνα 5.26.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον	79
Εικόνα 5.27.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον	79
Εικόνα 5.28.: Χαρακτηριστικά του μοντέλου της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο	81
Εικόνα 5.29.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο	81
Εικόνα 5.30.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο.....	82
Εικόνα 5.31.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο.....	82
Εικόνα 5.32.: Επιλογή εντολής generalized mixed linear μοντέλου.....	83
Εικόνα 5.33.: Επιλογή υποκειμένου επαναλαμβανόμενων μετρήσεων	84
Εικόνα 5.34.: Καθορισμός εξαρτημένης μεταβλητής και τύπου κατανομής της στη μεικτή δυαδική λογιστική παλινδρόμηση.....	85
Εικόνα 5.35.: Επιλογή μεταβλητών που έχουν fixed effects στη μεικτή δυαδική λογιστική παλινδρόμηση.....	85
Εικόνα 5.36. Καθορισμός πεδίου random effect στη μεικτή δυαδική λογιστική παλινδρόμηση.....	86
Εικόνα 5.37.:Χαρακτηριστικά του μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού	87
Εικόνα 5.38.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού.....	88
Εικόνα 5.39.: Επεξηγηματικός πίνακας για τον πίνακα ταξινόμησης.....	88
Εικόνα 5.40.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού	89
Εικόνα 5.41.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού	89
Εικόνα 5.42.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού	90
Εικόνα 5.43.: Χαρακτηριστικά του μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον	93
Εικόνα 5.44.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον.....	94
Εικόνα 5.45.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον	94
Εικόνα 5.46.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον	95
Εικόνα 5.47.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον	95
Εικόνα 5.48.: Χαρακτηριστικά μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον	97

Εικόνα 5.49.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον	98
Εικόνα 5.50.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον	98
Εικόνα 5.51.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον	99
Εικόνα 5.52.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον	99
Εικόνα 5.53.: Χαρακτηριστικά μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο.....	101
Εικόνα 5.54.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο	102
Εικόνα 5.55.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο	102
Εικόνα 5.56.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο.....	103
Εικόνα 5.57.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο.....	103

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1.1.: Θάνατοι σε τροχαία ατυχήματα 2001-2020 στην Ευρώπη	2
Διάγραμμα 1.2. Θάνατοι ανά εκατομμύριο κατοίκων για τις 28 χώρες-μέλη της Ευρωπαϊκής Ένωσης	2
Διάγραμμα 1.3.: Αριθμός ατυχημάτων από Ιανουάριο 2010 έως Ιανουάριο 2017	3
Διάγραμμα 1.4.: Αριθμός ατυχημάτων στην Ελλάδα από το 1991 έως το 2016	3
Διάγραμμα 1.5. Είδη θανατηφόρων τροχαίων ατυχημάτων στην Ελλάδα το 2016.....	4
Διάγραμμα 1.6.: Αιτίες θανατηφόρων ατυχημάτων στην Ελλάδα το 2016	5
Διάγραμμα 1.7.: Χρήση ηλεκτρονικών συσκευών κατά την οδήγηση	6
Διάγραμμα 1.8.: Διάγραμμα ροής σταδίων εκπόνησης της Διπλωματικής Εργασίας .	8
Διάγραμμα 4.1.: Καθαρή διάρκεια οδήγησης ανά τύπο οδού	52
Διάγραμμα 4.2.: Ποσοστό χρήσης κινητού τηλεφώνου σε συνάρτηση με τα συνολικά απότομα γεγονότα (events) προς τη συνολική απόσταση οδήγησης	52
Διάγραμμα 4.3.: Απότομα γεγονότα εντός και εκτός ωρών αιχμής	53
Διάγραμμα 4.4.: Καθαρή διάρκεια οδήγησης σε συνάρτηση με το ποσοστό χρήσης κινητού τηλεφώνου	53
Διάγραμμα 4.5.: Απότομα γεγονότα (events) ανά χρήστη (ενδεικτικά για κάποιους χρήστες).....	54
Διάγραμμα 4.6.: Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε συνάρτηση με το ποσοστό χρήσης του κινητού τηλεφώνου	54
Διάγραμμα 4.7: Καθαρή διάρκεια οδήγησης σε συνάρτηση με το ποσοστό της διάρκειας πάνω από το όριο ταχύτητας	55
Διάγραμμα 4.8.: Μέση ταχύτητα σε συνάρτηση με το ποσοστό του χρόνου πάνω από το όριο ταχύτητας	55
Διάγραμμα 4.9.: Απότομα γεγονότα ανά συνολική απόσταση σε συνάρτηση με τη συνολική απόσταση οδήγησης	56
Διάγραμμα 4.10.: Ποσοστό χρήσης κινητού τηλεφώνου ανά οδηγό.....	56
Διάγραμμα 4.11.: Μέση ταχύτητα σε συνάρτηση με τα συνολικά απότομα (harsh) γεγονότα προς τη συνολική απόσταση οδήγησης	57
Διάγραμμα 4.12.: Ποσοστό χρήσης κινητού τηλεφώνου ανά τύπο οδού	57
Διάγραμμα 4.13.: Ποσοστό χρόνου υπέρβασης του ορίου ταχύτητας ανά τύπο οδού	58
Διάγραμμα 5.1.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για τις ώρες εντός πρωινής και απογευματινής αιχμής	105
Διάγραμμα 5.2.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης ποσοστιαίας υπέρβασης του ορίου ταχύτητας	106
Διάγραμμα 5.3.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για δυο πιθανές τιμές της διάρκειας χρήσης του κινητού τηλεφώνου	106
Διάγραμμα 5.4.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για τις ώρες εντός πρωινής και απογευματινής αιχμής	107
Διάγραμμα 5.5.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές του αριθμού απότομων συμβάντων	108

Διάγραμμα 5.6.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης στροφικής ταχύτητας	108
Διάγραμμα 5.7.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης ποσοστιαίας υπέρβασης του ορίου ταχύτητας	109

1. Εισαγωγή

1.1 Γενική Ανασκόπηση

1.1.1 Οδικά Ατυχήματα

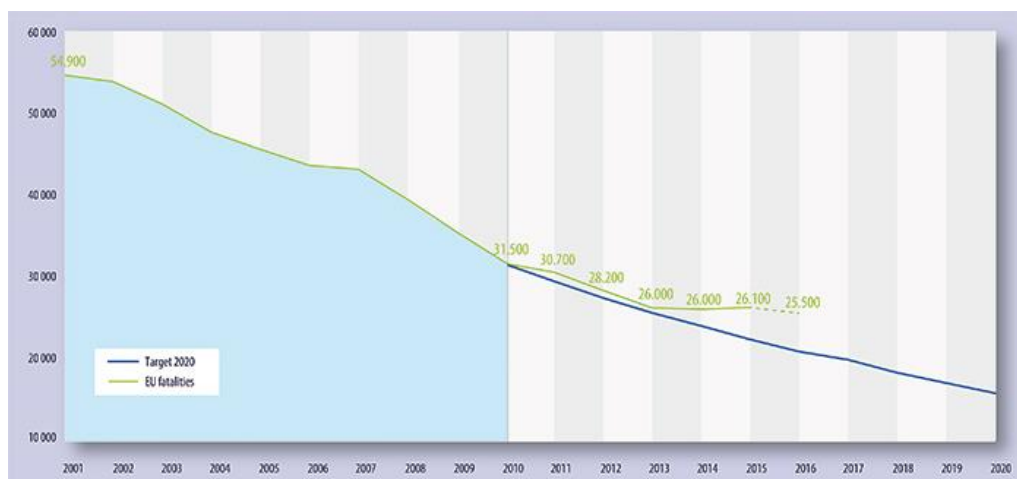
Οι οδικές μεταφορές είναι συνυφασμένες με την ύπαρξη της ανθρώπινης ζωής και αποτελούν δείκτη της εξέλιξης των κοινωνιών, καθώς συμβάλλουν στην αναβάθμιση του βιοτικού επιπέδου. Στη σύγχρονη εποχή, η περαιτέρω πρόοδος στον τομέα των μεταφορών, προσώπων και αγαθών, αποτελεί αντικείμενο της επιστήμης του συγκοινωνιολόγου μηχανικού και έχει στόχο την παροχή ασφαλών, γρήγορων, οικονομικών και άνετων μετακινήσεων. Αν και έχουν γίνει αρκετές **προσπάθειες για τη μείωση των οδικών ατυχημάτων** εκτιμάται διεθνώς ότι χάνονται 1,25 εκατομμύρια ζωές ετησίως (WHO, 2015), εκ των οποίων περίπου οι μισοί είναι «ευάλωτοι χρήστες του οδικού δικτύου»: πεζοί, ποδηλάτες και μοτοσικλετιστές. Δυστυχώς τα οδικά ατυχήματα έχουν καταστεί η κύρια αιτία θανάτου σε νεαρά άτομα από 15 έως 29 ετών (United Nations, 2015) και κατά συνέπεια οι απώλειες αυτές είναι περισσότερο οδυνηρές από οποιαδήποτε άλλη αιτία καθώς τα θύματα είναι κατά κανόνα άτομα υγιή και δραστήρια.

1.1.2 Τα οδικά ατυχήματα στην Ευρώπη

Οι στατιστικές του 2015 για την οδική ασφάλεια, που δημοσίευσε η Ευρωπαϊκή Επιτροπή (2016), επιβεβαιώνουν ότι οι δρόμοι της Ευρώπης παραμένουν οι ασφαλέστεροι στον κόσμο παρά την πρόσφατη επιβράδυνση της μείωσης των θανάτων από τροχαία ατυχήματα. Το 2016 σημειώθηκαν 25.500 θανατηφόρα ατυχήματα στους δρόμους της Ευρωπαϊκής Ένωσης, δηλαδή **600 λιγότεροι θάνατοι** σε σχέση με το 2015 και **6.000 λιγότεροι θάνατοι** σε σχέση με το 2010 (nrso, 2017). Επιπλέον, η Επιτροπή εκτιμά ότι τραυματίστηκαν σοβαρά στους δρόμους της ΕΕ 135.000 άτομα, ενώ το κοινωνικό κόστος (αποκατάσταση, υγειονομική περίθαλψη, υλικές ζημιές, κ.λπ.) των θανάτων και τραυματισμών από τροχαία ατυχήματα εκτιμάται σε τουλάχιστον 100 δισ. ευρώ.

Σύμφωνα με στατιστικές για το 2016 που δημοσιεύτηκαν από την Ευρωπαϊκή Επιτροπή οι περισσότερες ευρωπαϊκές χώρες σημείωσαν μείωση στον αριθμό των θανάτων από οδικά ατυχήματα που κυμάνθηκε από 3% έως 36% μεταξύ του 2010 και του 2015. Προκειμένου να επιτευχθεί ο **στρατηγικός στόχος της ΕΕ για μείωση** στο μισό του αριθμού **των θανάτων** από τροχαία ατυχήματα από το 2010 έως το **2020**, απαιτούνται πρόσθετες ενέργειες. Οι σημαντικές εξελίξεις όσον αφορά την καινοτομία και την τεχνολογία ενισχύουν τις δυνατότητες βελτίωσης της οδικής ασφάλειας στο μέλλον, ιδίως στον τομέα της αυτοματοποίησης των οχημάτων και της συνδεσιμότητας. Συγκεκριμένα, η Επιτροπή σκοπεύει να προωθήσει την ανάπτυξη των συνεργατικών ευφυών συστημάτων μεταφορών —μιας αμφίδρομης επικοινωνίας μεταξύ οχημάτων και οδικών υποδομών. Στο διάγραμμα που ακολουθεί παρουσιάζεται η μείωση των θανατηφόρων τροχαίων ατυχημάτων στην Ευρώπη από το 2001 μέχρι την επιδιωκόμενη τιμή κατά το 2020.

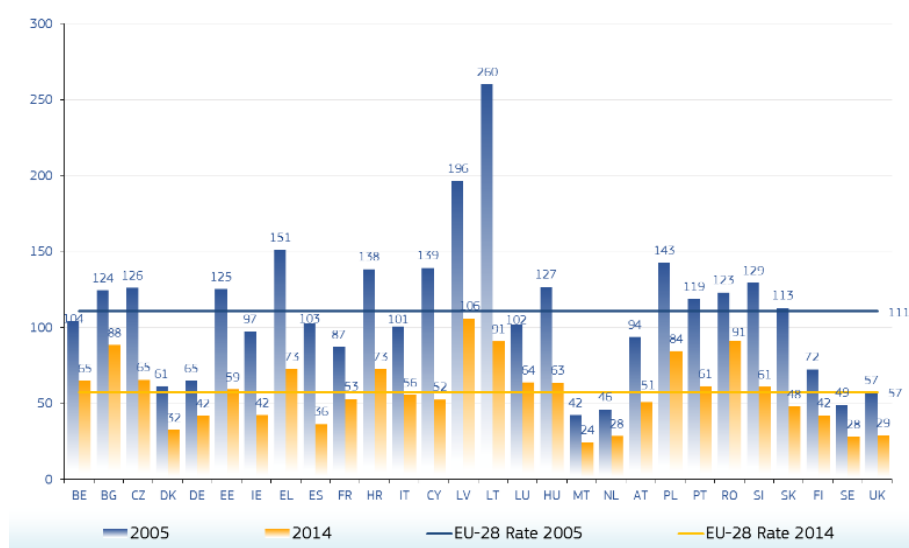
1. Εισαγωγή



Διάγραμμα 1.1.: Θάνατοι σε τροχαία ατυχήματα 2001-2020 στην Ευρώπη

Πηγή: Ευρωπαϊκή Επιτροπή, 2017

Στο διάγραμμα 1.2 απεικονίζονται οι θάνατοι ανά εκατομμύριο κατοίκων για τις 28 χώρες-μέλη της Ευρωπαϊκής Ένωσης κατά το 2005 και το 2014. Η Μάλτα κατέχει τη θέση με τους λιγότερους θανάτους και η Λιθουανία τη θέση με τους περισσότερους θανάτους.



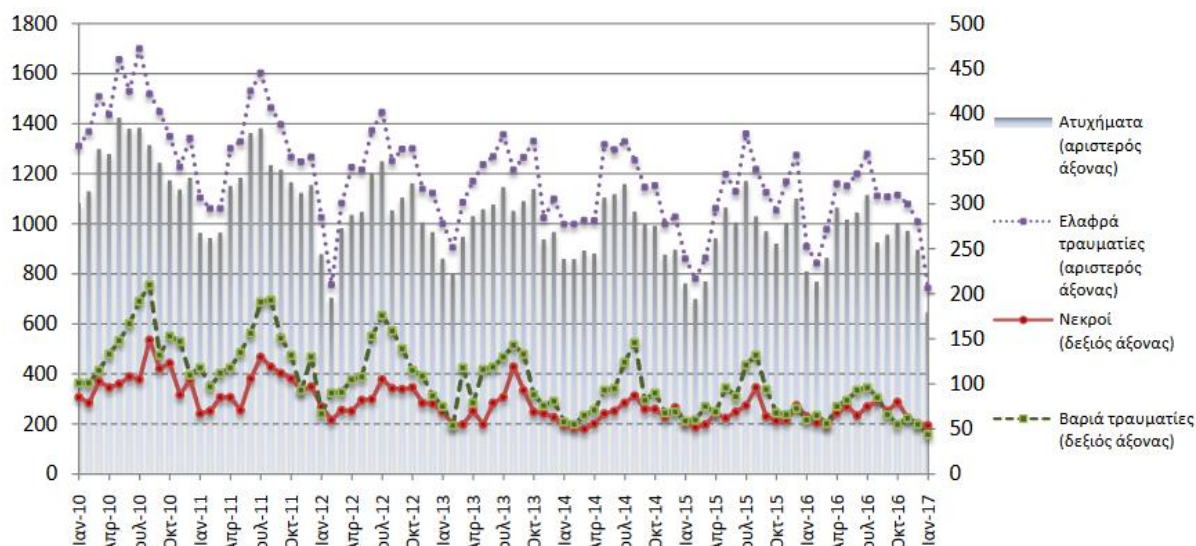
Διάγραμμα 1.2. Θάνατοι ανά εκατομμύριο κατοίκων για τις 28 χώρες-μέλη της Ευρωπαϊκής Ένωσης

Πηγή: Ευρωπαϊκή Επιτροπή

1.1.3 Τα οδικά ατυχήματα στην Ελλάδα

Ειδικότερα για την Ελλάδα, σύμφωνα με έκθεση της Ε. Επιτροπής, η οδική ασφάλεια έχει βελτιωθεί σημαντικά από το 2008 και μετά. Ωστόσο, ο αριθμός των θανατηφόρων τροχαίων παραμένει σημαντικά υψηλότερος από τον ευρωπαϊκό μέσο όρο: 74 ανά ένα εκατομμύριο κατοίκων, έναντι 52 στην ΕΕ, γεγονός που **κατατάσσει την Ελλάδα 4η** ανάμεσα στις χώρες με τους περισσότερους θανάτους σε οδικά ατυχήματα για το

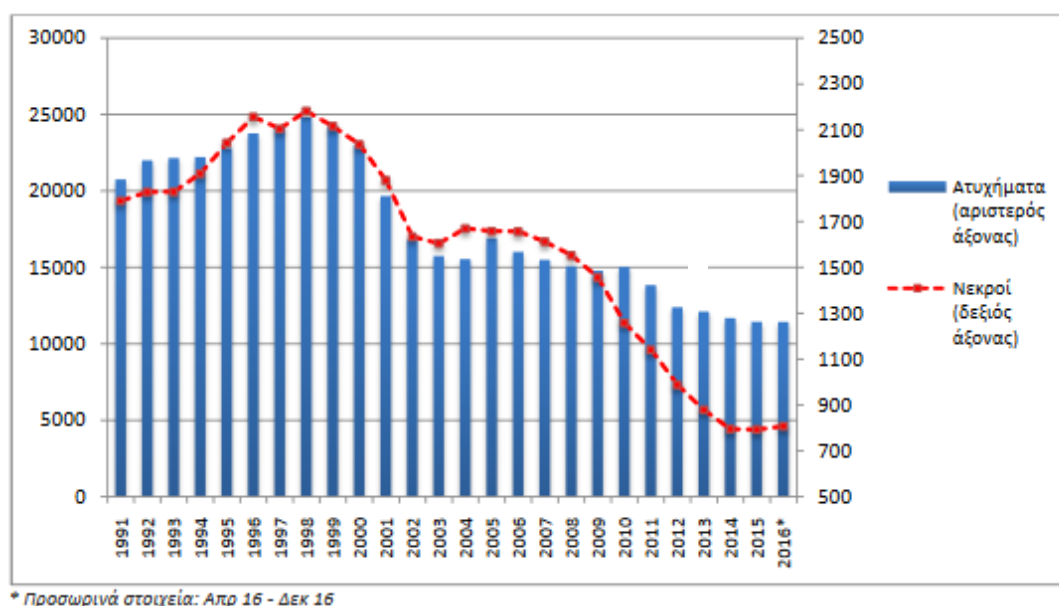
2016. Στο διάγραμμα 1.3 απεικονίζεται ο αριθμός των ατυχημάτων (νεκροί και τραυματίες) όπως εξελίχθηκε από τον Ιανουάριο του 2010 έως τον Ιανουάριο του 2017, στο οποίο διαπιστώνεται μείωση των θανατηφόρων ατυχημάτων με το πέρασμα του χρόνου.



Διάγραμμα 1.3.: Αριθμός ατυχημάτων από Ιανουάριο 2010 έως Ιανουάριο 2017

Πηγή: ΕΛΣΤΑΤ

Ακόμα πιο μακροσκοπικά, η χώρα μας συνολικά παρουσιάζει σημαντική **πτωτική τάση στα ατυχήματα** από το 1991 έως το 2016, όπως παρουσιάζεται και στο διάγραμμα 1.4 παρακάτω οδεύοντας προς τον στόχο της Ευρωπαϊκής Ένωσης για ουσιαστική μείωση των ατυχημάτων έως το 2020.

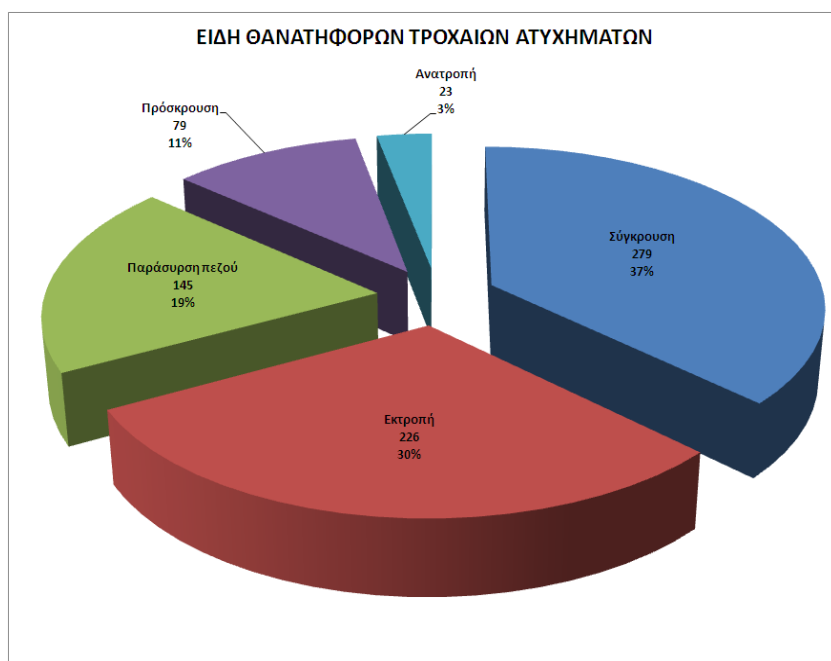


Διάγραμμα 1.4.: Αριθμός ατυχημάτων στην Ελλάδα από το 1991 έως το 2016

Πηγή: ΕΛΣΤΑΤ

1. Εισαγωγή

Στα παρακάτω διαγράμματα πίτας απεικονίζονται τα κυριότερα είδη και αιτίες θανατηφόρων τροχαίων ατυχημάτων στην Ελλάδα για το έτος 2016. Όπως διαπιστώνουμε οι συγκρούσεις και οι εκτροπές είναι οι κυριότεροι τύποι ατυχημάτων μεταξύ οχημάτων που οδηγούν σε θάνατο.



Διάγραμμα 1.5. Είδη θανατηφόρων τροχαίων ατυχημάτων στην Ελλάδα το 2016

Πηγή: Ελληνική Αστυνομία

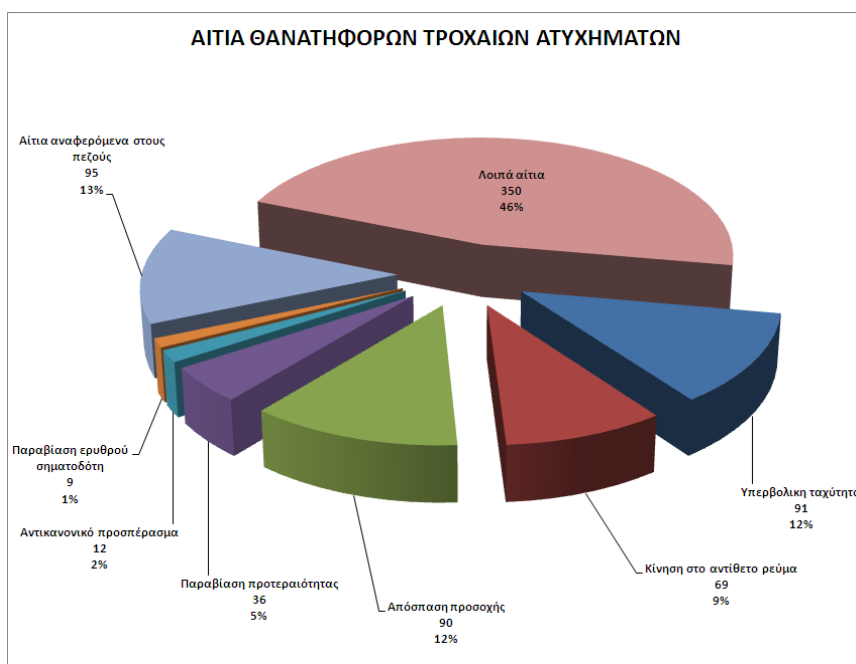
1.1.4 Τα αίτια των οδικών ατυχημάτων

Τρεις είναι οι σημαντικότεροι παράγοντες που επιδρούν στην οδική ασφάλεια, κατά σειρά αυξανόμενης σπουδαιότητας:

- **το όχημα** (τεχνική διαμόρφωση και εξοπλισμός, ενεργητική και παθητική ασφάλεια),
- **η οδός με το περιβάλλον της** (γεωμετρικά χαρακτηριστικά, κατασκευαστική διαμόρφωση, επίπεδο συντήρησης, εξοπλισμός, τοπικές κυκλοφοριακές ρυθμίσεις, χαρακτηριστικά της κυκλοφορίας, καιρικές συνθήκες), και
- **ο χρήστης της οδού** (εμπειρία, ψυχικές και κοινωνικές ιδιαιτερότητες καθώς επίσης πρότυπα συμπεριφοράς)

Σύμφωνα όμως με μελέτες (Sabey and Taylor, 1980) που πραγματοποιήθηκαν στη Μ. Βρετανία ο **ανθρώπινος παράγοντας** ευθύνεται στην συντριπτική πλειοψηφία των οδικών ατυχημάτων. Πιο συγκεκριμένα στο 65% των περιπτώσεων υπεύθυνος υπήρξε αποκλειστικά ο άνθρωπος, ενώ στο 95% τα οδικά ατυχήματα οφείλονταν σ' αυτόν σε συνδυασμό με άλλους παράγοντες. Κατά συνέπεια ο οδηγός δεν είναι μόνο ο παραλήπτης των πληροφοριών και εκείνος που παίρνει τις αποφάσεις, αλλά ο κύριος

παράγοντας στα ατυχήματα, καθώς τα περισσότερα από αυτά αποδίδονται στα ανθρώπινα λάθη (West et al. 1993), τα οποία οφείλονται κατά βάση στην αδύναμη συνειδητοποίηση των κινδύνων του οδηγού. Στο διάγραμμα πίτας 1.5 που ακολουθεί παρουσιάζονται πιο αναλυτικά οι κατηγορίες των παραγόντων που οδηγούν σε θανατηφόρα ατυχήματα στην Ελλάδα και οφείλονται σε ανθρώπινη υπαιτιότητα, με τα ποσοστά τους για το 2016. Όπως διαπιστώνουμε, η **απόσπαση της προσοχής** αποτελεί βασικό παράγοντα που οδηγεί σε θανατηφόρα ατυχήματα (12%) και επομένως κρίνεται απαραίτητη η ανάλυση της σχέσης ανάμεσα στα χαρακτηριστικά της οδικής συμπεριφοράς και των επιδόσεων του οδηγού και στην απόσπαση προσοχής.



Διάγραμμα 1.6.: Αιτίες θανατηφόρων ατυχημάτων στην Ελλάδα το 2016

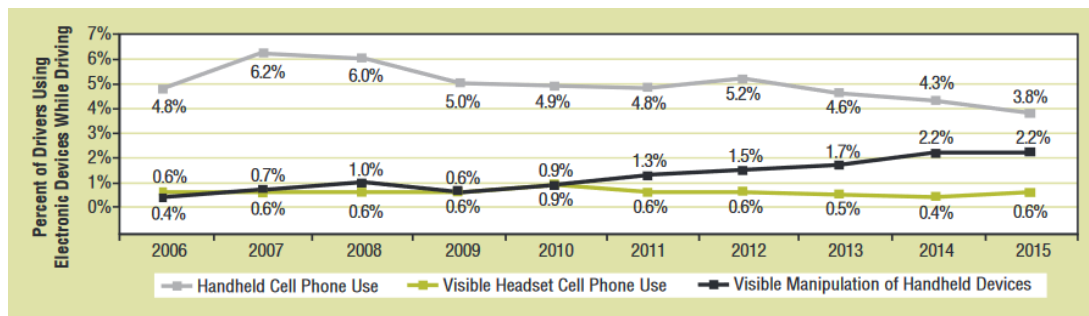
Πηγή: Ελληνική Αστυνομία

Απόσπαση της προσοχής κατά την οδήγηση

Ο οργανισμός NHTSA των Η.Π.Α. (National Highway Traffic Safety Administration) έχει εκτιμήσει ότι περίπου το 25-30% των οδικών ατυχημάτων οφείλεται σε έλλειψη συγκέντρωσης του οδηγού ή σε **απόσπαση της προσοχής** του πριν από το ατύχημα. Η απόσπαση της προσοχής του οδηγού λόγω της χρήσης κινητού τηλεφώνου φαίνεται να αποτελεί έναν από τους σοβαρότερους παράγοντες οδικών ατυχημάτων, αφού ο οδηγός καλείται να ανταπεξέλθει ταυτόχρονα σε παραπάνω από μία ενέργειες. Ενδεικτικά αναφέρεται ότι μόνο το 2015, 3.477 άνθρωποι σκοτώθηκαν και 391.000 τραυματίστηκαν σε αυτοκινητικά ατυχήματα που περιλάμβαναν οδηγούς με απόσπαση προσοχής, ενώ έχει υπολογιστεί ότι κατά τη διάρκεια της ημέρας περίπου 660.000 οδηγοί χρησιμοποιούν το κινητό τους τηλέφωνο κατά την οδήγηση (NHTSA, 2017). Στο παρακάτω διάγραμμα απεικονίζεται η χρήση ηλεκτρονικών συσκευών κατά

1. Εισαγωγή

την οδήγηση από το 2006 έως το 2015 στο οποίο διαπιστώνεται μια αύξηση της τάξεως του 1.8% για την ορατή χρήση συσκευών χειρός με το πέρασμα του χρόνου.



Διάγραμμα 1.7.: Χρήση ηλεκτρονικών συσκευών κατά την οδήγηση

Πηγή: NHTSA

Το πρόβλημα σταματά να είναι μεμονωμένο αφού η **απόσπαση της προσοχής** μπορεί να θεωρηθεί ως ένα τυπικό μέρος της καθημερινής οδήγησης, ενώ, ταυτόχρονα, οι οδηγοί φαίνεται να υποεκτιμούν τις επιπτώσεις αυτού του τρόπου οδήγησης. Η διείσδυση των διαφόρων νέων τεχνολογιών στο εσωτερικό του οχήματος, καθώς και η αναμενόμενη αύξηση της χρήσης των συσκευών αυτών τα επόμενα χρόνια, καθιστά αναγκαία την περαιτέρω διερεύνηση της επίδρασής τους στην προσοχή του οδηγού, στη ροή της κυκλοφορίας και την οδική ασφάλεια.

Πολλές δευτερεύουσες εργασίες ή δραστηριότητες, ιδιαίτερα εκείνες που προκύπτουν από τη χρήση των φορητών ηλεκτρονικών συσκευών, είναι επιζήμιες για την ασφάλεια των οδηγών. Η ανάλυση προήλθε χρησιμοποιώντας το πρόγραμμα SHRP 2 NDS (Strategic Highway Research Program 2 Naturalistic Driving Study) που έγινε στις Ηνωμένες Πολιτείες της Αμερικής. Δημιουργήθηκε μια βάση δεδομένων με στοιχεία από πειράματα οδήγησης σε πραγματικές συνθήκες (naturalistic driving), η οποία περιλαμβάνει τις συνεχείς επί τόπου παρατηρήσεις που έγιναν μέσω πολλαπλών βίντεο από προσαρμοσμένες κάμερες και αισθητήρες. Περισσότεροι από 3.500 οδηγοί κατά την διάρκεια τριών ετών συνέτυχαν στην έρευνα για να αντληθούν όλα τα σχετικά στοιχεία (Virginia Tech Ινστιτούτο Μεταφορών, 2015).

1.2 Στόχος της Διπλωματικής Εργασίας

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η διερεύνηση και η **προτυποποίηση της επιρροής της χρήσης κινητού τηλεφώνου στη συμπεριφορά του οδηγού αξιοποιώντας λεπτομερή δεδομένα από αισθητήρες έξυπνων κινητών τηλεφώνων.**

Πιο συγκεκριμένα, η παρούσα Διπλωματική Εργασία έχει δύο επιμέρους στόχους. Καταρχάς, επιδιώκεται να εξεταστεί κατά πόσο χαρακτηριστικά οδήγησης που καταγράφηκαν από τους αισθητήρες έξυπνων κινητών τηλεφώνων όπως η ταχύτητα, η πέδηση, η επιτάχυνση, η περίοδος οδήγησης κατά τη διάρκεια της ημέρας, καθώς

και η απόσπαση προσοχής μέσω της χρήσης κινητού τηλεφώνου επηρεάζουν και επομένως μπορούν **να προβλέψουν τα απότομα συμβάντα** κατά την οδήγηση.

Ως δεύτερος επιμέρους στόχος, εξετάζεται η δυνατότητα **πρόβλεψης της χρήσης ή μη του κινητού τηλεφώνου κατά την οδήγηση** μέσα από την καταγραφή μεγεθών που σχετίζονται με την οδήγηση.

Για τον σκοπό αυτό αναλύθηκε ένας μεγάλος αριθμός δεδομένων που συλλέχθηκαν από τους αισθητήρες έξυπνων κινητών τηλεφώνων 100 οδηγών και στη συνέχεια **αναπτύχθηκαν μαθηματικά μοντέλα** τα οποία θα επεξηγούν τα παραπάνω ζητήματα και θα προβλέπουν αφενός τη σχέση μεταξύ χαρακτηριστικών οδήγησης και απότομων συμβάντων και αφετέρου τη χρήση ή μη του κινητού τηλεφώνου μέσα από την παρατήρηση συγκεκριμένων μεγεθών.

Τα αποτελέσματα της ανάλυσης και τα συμπεράσματα της εργασίας αναμένεται να αποφέρουν **πολλαπλά και σημαντικά οφέλη στην κοινωνία**, εφόσον θα είναι δυνατό να αξιοποιηθούν τόσο ατομικά από τους οδηγούς όσο και από τις βιομηχανίες αυτοκινήτων και ασφαλειών αλλά και από τις Αρχές για τη βελτίωση της οδικής ασφάλειας με επεμβάσεις στη συμπεριφορά του οδηγού, στην τεχνολογία των οχημάτων και στη διαχείριση της οδικής ασφάλειας γενικότερα.

1.3 Μεθοδολογία

Στο υποκεφάλαιο αυτό περιγράφεται συνοπτικά η μέθοδος που ακολουθήθηκε για την επίτευξη του στόχου της Διπλωματικής Εργασίας.

Αρχικά μετά την οριστικοποίηση του επιδιωκόμενου στόχου πραγματοποιήθηκε **βιβλιογραφική ανασκόπηση** τόσο σε ελληνική όσο και σε διεθνή βιβλιογραφία. Στη φάση αυτή πραγματοποιήθηκε αναζήτηση παρεμφερών ερευνών, επιστημονικών άρθρων καθώς και γενικών πληροφοριών σχετικά με το εξεταζόμενο αντικείμενο που θα μπορούσαν να φανούν χρήσιμες για τη συγκεκριμένη έρευνα. Μέσω των ερευνών αυτών καταβλήθηκε προσπάθεια να αποκτηθεί μια σχετική εμπειρία στην επεξεργασία τέτοιων θεμάτων, καθώς επίσης και να αποφασιστεί η μέθοδος με βάση την οποία θα πραγματοποιηθεί η επεξεργασία των στοιχείων και να επιτευχθεί ο επιδιωκόμενος στόχος.

Αφού μελετήθηκαν τα στοιχεία που συγκεντρώθηκαν από τη βιβλιογραφική ανασκόπηση, σειρά είχε η εύρεση του τρόπου συλλογής στοιχείων. Εφαρμόστηκε ένα καινοτόμο σύστημα συλλογής δεδομένων, μέσα από την εξατομικευμένη καταγραφή της συμπεριφοράς του οδηγού σε πραγματικό χρόνο χρησιμοποιώντας αισθητήρες έξυπνων κινητών τηλεφώνων. **Εκατό (100) οδηγοί συμμετείχαν στο πείραμα οδήγησης σε πραγματικές συνθήκες** (smartphone naturalistic driving experiment) σε χρονικό διάστημα 4 μηνών και δημιουργήθηκε μια μεγάλη βάση δεδομένων με χιλιάδες διαδρομές. Αξιοποιήθηκε η ειδική πλατφόρμα της εταιρείας της συγκεκριμένης εφαρμογής σε έξυπνα κινητά τηλέφωνα (OSeven) για την αποτελεσματική μετάβαση από τη συλλογή στοιχείων στη διαδικασία ανάλυσης των δεδομένων.

Στη συνέχεια, ακολούθησε η επιλογή της μεθόδου στατιστικής επεξεργασίας των στοιχείων και η εισαγωγή της βάσης δεδομένων, που είχε προετοιμαστεί στο Microsoft Excel, στο ειδικό λογισμικό στατιστικής ανάλυσης (IBM SPSS 21.0). Επόμενο στάδιο

1. Εισαγωγή

αποτέλεσε η **ανάπτυξη μαθηματικών μοντέλων γραμμικής και δυαδικής λογιστικής παλινδρόμησης** που περιέγραφαν σε αποδεκτό επίπεδο εμπιστοσύνης την επιρροή των ανεξάρτητων μεταβλητών στην εκάστοτε εξαρτημένη. Στα τελικά μοντέλα γραμμικής παλινδρόμησης λήφθηκε υπόψη και η επαναληπτικότητα στα δεδομένα οδήγησης του κάθε οδηγού για περισσότερες της μίας διαδρομές προκειμένου να αποφευχθούν σφάλματα στο μοντέλο, αναπτύσσοντας μοντέλα ανάμικτων επιδράσεων (mixed-effects models).

Μετά την **αξιολόγηση και ερμηνεία των αποτελεσμάτων**, εξήχθησαν τα αντίστοιχα συμπεράσματα για τον βαθμό και τον τύπο της επιρροής των εκάστοτε ανεξάρτητων μεταβλητών στην εξαρτημένη. Έτσι, προέκυψαν σημαντικές πληροφορίες για το υπό εξέταση πρόβλημα καθώς και διατυπώθηκαν αξιόλογες προτάσεις για περαιτέρω έρευνα.

Στη συνέχεια παρουσιάζονται σχηματικά τα στάδια της μεθοδολογίας που ακολουθήθηκαν για τις ανάγκες της παρούσας Διπλωματικής Εργασίας.



Διάγραμμα 1.8.: Διάγραμμα ροής σταδίων εκπόνησης της Διπλωματικής Εργασίας

1.4 Δομή

Το **πρώτο κεφάλαιο** αποτελεί την εισαγωγή και περιλαμβάνει μια γενική ανασκόπηση όπου παρουσιάζεται το πλαίσιο της Διπλωματικής Εργασίας και αναφέρονται τα δεδομένα όπως έχουν προκύψει ως σήμερα, όσον αφορά την οδική ασφάλεια. Παρατίθενται τα βασικά στατιστικά στοιχεία για την Ελλάδα και τέλος παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, περιγράφοντας συνοπτικά τα βασικά στάδια της εκτέλεσης της παρούσας εργασίας.

Στο **δεύτερο κεφάλαιο** παρουσιάζονται τα αποτελέσματα της βιβλιογραφικής ανασκόπησης, τα οποία προέκυψαν από την αναζήτηση και την καταγραφή ερευνών

με αντικείμενο την απόσπαση προσοχής καθώς επίσης και τις μεθόδους συλλογής στοιχείων από πειράματα σε πραγματικές συνθήκες οδήγησης. Εξετάζονται εργασίες από Ελλάδα και εξωτερικό, έρευνες που έχουν δημοσιευθεί σε επιστημονικά περιοδικά, βιβλία, άρθρα, πρακτικά συνεδρίων. Το κεφάλαιο κλείνει με τη σύνοψη και κριτική αξιολόγηση των αποτελεσμάτων των ερευνών, προκειμένου να διαπιστωθεί ποιες από αυτές μπορούν να συμβάλλουν ουσιαστικά στην παρούσα εργασία.

Στο **τρίτο κεφάλαιο** παρουσιάζεται το θεωρητικό υπόβαθρο στο οποίο στηρίχθηκε η ανάλυση των στοιχείων της οδηγικής συμπεριφοράς. Αρχικά, επεξηγούνται οι λόγοι επιλογής των δύο στατιστικών αναλύσεων που χρησιμοποιήθηκαν και περιγράφονται βασικές μαθηματικές και στατιστικές έννοιες για την καλύτερη κατανόηση και επεξήγηση των επιλεχθέντων μαθηματικών μοντέλων, δηλαδή της γραμμικής παλινδρόμησης και της δυαδικής λογιστικής παλινδρόμησης. Τέλος το κεφάλαιο ολοκληρώνεται με μία σύντομη αναφορά στα βήματα που ακολουθούνται, για την επεξεργασία των δεδομένων στο ειδικό λογισμικό στατιστικής ανάλυσης.

Το **τέταρτο κεφάλαιο** περιγράφει τη διαδικασία συλλογής και επεξεργασίας των στοιχείων. Αρχικά, περιγράφεται η μέθοδος συλλογής των δεδομένων μέσα από το πείραμα σε πραγματικές συνθήκες οδήγησης που πραγματοποιήθηκε. Στη συνέχεια παρατίθενται συγκεντρωτικοί πίνακες στους οποίους περιγράφονται τα χαρακτηριστικά του δείγματος, ακολουθούμενοι από τον απαραίτητο σχολιασμό. Στο τέλος, αναλύεται ο τρόπος εισαγωγής των δεδομένων στο ειδικό λογισμικό στατιστικής ανάλυσης.

Στο **πέμπτο κεφάλαιο** περιλαμβάνεται η αναλυτική περιγραφή της μεθοδολογίας που εφαρμόστηκε ως την εξαγωγή των τελικών αποτελεσμάτων. Αρχικά περιγράφονται τα βήματα που ακολουθήθηκαν για την εφαρμογή της μεθοδολογίας και παρουσιάζεται η διαδικασία ανάπτυξης του μαθηματικού μοντέλου. Παρουσιάζονται τα δεδομένα εισόδου κι εξόδου, με ιδιαίτερη έμφαση στους στατιστικούς ελέγχους αξιοπιστίας των αποτελεσμάτων, ενώ τα τελικά αποτελέσματα συνοδεύονται από τις αντίστοιχες μαθηματικές σχέσεις και την περιγραφή τα για την διευκόλυνση της κατανόησης τους.

Στο **6ο κεφάλαιο** διατυπώνονται τα κυριότερα συμπεράσματα, τα οποία προέκυψαν από την ανάλυση των στοιχείων της έρευνας. Στην αρχή του κεφαλαίου πραγματοποιείται μία σύνοψη των κυριότερων σημείων της παρούσας εργασίας, ενώ στη συνέχεια διατυπώνονται τα συμπεράσματα, τα οποία προέκυψαν. Στο τέλος του κεφαλαίου, διατυπώνονται προτάσεις, τόσο σχετικά με την αξιοποίηση των αποτελεσμάτων της παρούσας εργασίας, όσο και σχετικά με τη βελτίωση της οδικής ασφάλειας και περαιτέρω έρευνες, σχετικές με το αντικείμενο της εργασίας αυτής.

Στο **7ο κεφάλαιο** παρατίθεται ο κατάλογος των βιβλιογραφικών αναφορών. Το κεφάλαιο αυτό περιλαμβάνει αναφορές, που αφορούν τόσο σε έρευνες που παρουσιάστηκαν στα κεφάλαια της εισαγωγής και της βιβλιογραφικής ανασκόπησης, όσο και σε στατιστικές έννοιες και μεθόδους, που αναλύθηκαν στο θεωρητικό υπόβαθρο.

2. Βιβλιογραφική Ανασκόπηση

2.1 Εισαγωγή

Το κεφάλαιο αυτό αφορά τη βιβλιογραφική ανασκόπηση και περιλαμβάνει έρευνες στον τομέα της οδικής ασφάλειας, το αντικείμενο και η μεθοδολογία των οποίων παρουσιάζει συνάφεια με την παρούσα Διπλωματική Εργασία. Συγκεκριμένα, παρουσιάζονται έρευνες που αναφέρονται στην απόσπαση της προσοχής λόγω της χρήσης του κινητού τηλεφώνου και στην επιθετική οδήγηση καθώς και σε μεθοδολογίες συλλογής στοιχείων παρόμοιες με αυτές που εφαρμόστηκαν για τη συγκεκριμένη Εργασία. Για κάθε επιστημονική εργασία παρατίθεται σύνοψή της που περιλαμβάνει το πλαίσιο της έρευνας, τη μεθοδολογία και τα βασικά αποτελέσματά της με έμφαση σε αυτά που σχετίζονται με την Διπλωματική Εργασία. Μέσω της ανασκόπησης των μεθοδολογιών των ερευνών αυτών, επιχειρήθηκε ο προσδιορισμός μιας κατάλληλης μεθόδου για την αντιμετώπιση του αντικειμένου της Διπλωματικής Εργασίας.

2.2 Συναφείς έρευνες

Το υποκεφάλαιο αυτό παρουσιάζει τα βασικά σημεία ερευνών οι οποίες σχετίζονται με την επιρροή του κινητού στην οδική συμπεριφορά και οδική ασφάλεια καθώς και με την επιθετική οδήγηση. Οι έρευνες αυτές βασίζονται είτε σε προσωπικές εκτιμήσεις των ίδιων των οδηγών μέσω ερωτηματολογίων, είτε στην παρατήρηση της συμπεριφοράς των συμμετεχόντων σε συνθήκες οδήγησης. Σκοπός είναι η σύγκριση των αποτελεσμάτων της παρούσας εργασίας με άλλες έρευνες με παρεμφερές περιεχόμενο.

2.2.1 The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model (2015)

Σκοπός

Το φρενάρισμα είναι ένα κρίσιμο έργο της οδήγησης με άμεσο κίνδυνο σύγκρουσης, ενώ η υπερβολική και η ακατάλληλη πέδηση μπορούν να οδηγήσουν σε συγκρούσεις. **Ο στόχος αυτής της μελέτης είναι να συγκριθεί το προφίλ πέδησης των νέων οδηγών που η προσοχή τους έχει αποσπαστεί** από συνομιλίες στο κινητό τηλέφωνο σε σχέση με την πέδηση χωρίς απόσπαση προσοχής. Ειδικότερα, η συμπεριφορά πέδησης των οδηγών για έναν πεζό που εισέρχεται σε μια διάβαση εξετάσθηκε με τη χρήση του σύνθετου προσομοιωτή οδήγησης Cars-Q.

Μεθοδολογία

Τριάντα-δύο οδηγοί με άδεια οδήγησαν τον **προσομοιωτή σε τρεις συνθήκες τηλεφώνου**: βασική γραμμή (δεν υπάρχει τηλεφωνική συνομιλία), hands-free, και με κανονική χρήση τηλεφώνου. Επιπρόσθετα με την οδήγηση του προσομοιωτή, κάθε συμμετέχων συμπλήρωσε **ερωτηματολόγια** που σχετίζονται με τα δημογραφικά στοιχεία του οδηγού, το ιστορικό οδήγησης, τη χρήση των κινητών τηλεφώνων κατά την οδήγηση, και γενικά την ιστορία χρήσης κινητού τηλεφώνου στην οδήγηση. Οι οδηγοί ήταν 18-26 ετών και χωρίστηκαν ομοιόμορφα κατά φύλο. Ένα **γραμμικό**

μεικτό μοντέλο ανάλυσης των προφίλ πεδήσεως κατά μήκος του οδοστρώματος πριν από την διάβαση πεζών αποκάλυψε συγκριτικά αυξημένες επιβραδύνσεις για τους οδηγούς με αποσπασμένη προσοχή, ιδιαίτερα κατά τη διάρκεια της αρχικής επιβράδυνσης στα 20 kph. Ο αρχικός χρόνος επιβράδυνσης στα 20 kph των οδηγών μοντελοποιήθηκε χρησιμοποιώντας ένα παραμετρικό μοντέλο (AFT) με κατανομή Weibull με σύμπλεγμα ετερογένειας για τα επαναλαμβανόμενα μέτρα σχεδιασμού του πειράματος.

Αποτελέσματα

Οι παράγοντες που βρέθηκε ότι **επηρεάζουν** σημαντικά το έργο πέδησης περιλάμβαναν δυναμικές μεταβλητές του οχήματος όπως η **αρχική ταχύτητα και η μέγιστη επιβράδυνση, η κατάσταση του τηλεφώνου και ειδικές μεταβλητές για τον οδηγό όπως ο τύπος άδειας, το ιστορικό συμμετοχής σε συντριβή, και η αυτοαναφερόμενη εμπειρία** του οδηγού από τη χρήση κινητού τηλεφώνου κατά την οδήγηση. Συμπερασματικά, οι οδηγοί με αποσπασμένη προσοχή εμφανίζονται κατά μέσο όρο να **μειώνουν την ταχύτητα του οχήματός τους πιο γρήγορα και πιο απότομα από ό, τι οι οδηγοί χωρίς απόσπαση προσοχής**, γεγονός που αποκαλύπτει ίσως την **αντιστάθμιση των κινδύνων** που προσπαθούν να επιτύχουν με αυτόν τον τρόπο. Το φρενάρισμα φαίνεται να είναι πιο επιθετικό για τους αποσπασμένους οδηγούς με προσωρινή άδεια σε σύγκριση με τους οδηγούς με ανοικτές άδειες. **Απότομη ή υπερβολική πέδηση από αποσπασμένους οδηγούς** θα μπορούσε να δημιουργήσει σημαντικές ανησυχίες για την ασφάλεια σε οχήματα που ακολουθούν σε ένα ρεύμα κυκλοφορίας.

2.2.2 Driving performance while using a mobile phone: A simulation study of Greek professional (2016)

Σκοπός

Η παρούσα μελέτη έχει ως **στόχο την αξιολόγηση της ικανότητας οδήγησης των επαγγελματιών οδηγών, ενώ χρησιμοποιούν κινητό τηλέφωνο.**

Μεθοδολογία

Ένα δείγμα από 50 άνδρες επαγγελματίες οδηγούς συμμετείχαν στη μελέτη. Το δείγμα λήφθηκε εύκολα από τα **μητρώα των επαγγελματιών οδηγών** και τις κύριες πιάτσες ταξί. Τα **κριτήρια ένταξης** ήταν: (α) ηλικία άνω των 18 ετών (β) κατοχή άδειας οδήγησης, (γ) επαρκείς ικανότητες ανάγνωσης, γραφής και επικοινωνίας, (δ) ενημερωμένη συγκατάθεση πριν από τη συμμετοχή. Εργαστηριακές δοκιμές διεξήχθησαν χρησιμοποιώντας τον **προσομοιωτή VS500M** που κατασκευάζεται από την Virage Simulation Inc. Η παρέμβαση περιλάμβανε διάφορες εργασίες / αναθέσεις: (α) οδήγηση χωρίς τη χρήση κινητού τηλεφώνου (χρόνος ελέγχου), (β) οδήγηση έχοντας μια συνομιλία στο κινητό τηλέφωνο, (Εργασία_1), (γ) οδήγηση διαβάζοντας δυνατά μηνύματα κειμένου (Εργασία_2) και, (δ) οδήγηση συντάσσοντας κείμενο (Εργασία_3). Οι διαφορές στην συμπεριφορά οδήγησης μεταξύ του χρόνου ελέγχου και του χρόνου με αναθέσεις, εξετάστηκαν. Οι συμμετέχοντες αξιολογήθηκαν για τα εξής μέτρα που προέκυψαν: (α) μεταβολή της πορείας του οχήματος ανά

δευτερόλεπτο, (β) απόσταση μεταξύ δυο οχημάτων (προπορευόμενου και ακολουθούμενου) ανά δευτερόλεπτο, (γ) μεταβολή της πλευρικής θέσης στη λωρίδα ανά δευτερόλεπτο, και (δ) άθροισμα των τετραγώνων επιτάχυνσης ανά δευτερόλεπτο.

Αποτελέσματα

Η ανάλυση έδειξε ότι η **"μεταβολή της πορείας του οχήματος ανά δευτερόλεπτο"** επηρεάστηκε σημαντικά από την ανάγνωση μηνυμάτων [$t(50) = -5.443$; $p\text{-value} < 0.0001$] και από τη σύνταξη μηνυμάτων [$t(50) = -5.442$; $p\text{-value} < 0.0001$] και συγκεκριμένα αυξήθηκε. Σημαντική επίδραση παρατηρήθηκε στη μεταβολή της **"απόστασης μεταξύ δυο οχημάτων (προπορευόμενου και ακολουθούμενου) ανά δευτερόλεπτο"** η οποία μειώθηκε και για τις τρεις αναθέσεις κινητού τηλεφώνου ενώ **"η μεταβολή της πλευρικής θέσης στη λωρίδα ανά δευτερόλεπτο"** αυξήθηκε για τη σύνταξη μηνύματος και μειώθηκε για τις άλλες δύο αναθέσεις. Τέλος, **το "άθροισμα των τετραγώνων επιτάχυνσης ανά δευτερόλεπτο"** μειώθηκε σημαντικά κατά τη διάρκεια συνομιλίας στο τηλέφωνο [$t(50) = 2.713$; $p\text{-value} = 0.009$] καθώς και κατά τη διάρκεια "σύνταξης μηνύματος" [$t(50) = 3.428$; $p\text{-value} = 0.001$] σε σχέση με τον χρόνο ελέγχου. **Η μελέτη είναι από τις λίγες υπάρχουσες πειραματικές μελέτες** σε μια χώρα, **την Ελλάδα**, με έναν από τους υψηλότερους αριθμούς θανατηφόρων τροχαίων ατυχημάτων στην Ευρώπη, αλλά με περιορισμένα στοιχεία για τη συμπεριφορά της οδικής κυκλοφορίας. Η μελέτη αυτή θα μπορούσε να καθοδηγήσει το σχεδιασμό των μελετών προσομοίωσης μεγάλης κλίμακας με στόχο να διερευνήσει τις επιπτώσεις των κινητών τηλεφώνων στην οδηγική συμπεριφορά.

2.2.3 Distracted driving and mobile phone use: Overview of impacts and countermeasures (2014)

Σκοπός

Η απόσπαση της προσοχής αποτελεί σημαντικό παράγοντα αυξημένου κινδύνου τροχαίων ατυχημάτων σε ολόκληρο τον κόσμο. Ενώ οι ανθρωπίνι παράγοντες συνολικά είναι οι βασικές αιτίες στο 65-95% των οδικών ατυχημάτων, πρόσφατη έρευνα έδειξε ότι το 30% των οδηγών που συμμετείχαν σε κάποιο οδικό ατύχημα ανέφεραν κάποια πηγή απόσπασης πριν συμβεί το ατύχημα. Επιπλέον, η διείσδυση των διαφόρων νέων τεχνολογιών στο εσωτερικό του οχήματος, καθώς και η αναμενόμενη αύξηση της χρήσης τέτοιων συσκευών μέσα στα επόμενα χρόνια, καθιστά την περαιτέρω διερεύνηση της επιρροής τους στην προσοχή των οδηγών, στη ροή της κυκλοφορίας και την οδική ασφάλεια, πολύ σημαντική. **Ο σκοπός αυτής της μελέτης είναι να παρέχει μια ολοκληρωμένη εικόνα των επιπτώσεων της απόσπασης της προσοχής του οδηγού** όσον αφορά τη χρήση του κινητού τηλεφώνου για την οδική ασφάλεια, καθώς και να προτείνει συγκεκριμένα μέτρα.

Μεθοδολογία

Για αυτό το σκοπό διεξήχθη **επανεξέταση των υφιστάμενων μελετών** για την επίδραση των κινητών τηλεφώνων κατά την οδήγηση, που συνοδεύτηκε από παρουσίαση πρόσφατων ερευνητικών αποτελεσμάτων πειραμάτων του Εθνικού Μετσόβιου Πολυτεχνείου και από πρόταση αντίστοιχων στρατηγικών ενάντια στην

απόσπαση της οδηγικής προσοχής. Οι παράγοντες απόσπασης της οδηγικής προσοχής μπορούν να κατηγοριοποιηθούν σε αυτούς που συμβαίνουν έξω από το όχημα (εξωτερικούς) και αυτούς που συμβαίνουν μέσα στο όχημα (εντός οχήματος), μερικοί από τους οποίους είναι οι τεχνολογικές συσκευές επικοινωνίας. Πιο συγκεκριμένα, αναφορικά με τη χρήση του κινητού τηλεφώνου (συνομιλία στο κινητό, σύνταξη μηνυμάτων, πλοήγηση κλπ.), μια σειρά μελετών έδειξε ότι **η χρήση κινητού τηλεφώνου έχει αρνητικές συνέπειες στην συμπεριφορά του οδηγού** και στην πιθανότητα εμπλοκής σε ατύχημα. Ωστόσο, η ποσοτικοποίηση αυτών των συνεπειών διαφέρει σημαντικά ανάλογα με το είδος του χρήστη, του οδικού περιβάλλοντος, της επιφάνειας της συσκευής κλπ. Επιπλέον, αποτελέσματα από σχετικές έρευνες στην Ελλάδα (παρατηρήσεις εντός οχήματος, παρατηρήσεις στο πεδίο και πειράματα προσομοιωτή) επιβεβαιώνουν ότι η χρήση κινητού τηλεφώνου οδηγεί σε **αυξημένο κίνδυνο ατυχήματος**, ειδικά όταν συμβαίνουν αναπάντεχα συμβάντα.

Αποτελέσματα

Τα αποτελέσματα της έρευνας δείχνουν ότι η χρήση κινητού τηλεφώνου ίσως είναι ο πιο σημαντικός παράγοντας στην εντός οχήματος απόσπαση προσοχής για τους οδηγούς. Οι οδηγοί που χρησιμοποιούν το κινητό τηλέφωνο κατά τη διάρκεια της οδήγησης εμφανίζουν έως τέσσερις (4) φορές υψηλότερο κίνδυνο ατυχήματος. Επιπλέον, πολλές έρευνες έχουν διαπιστώσει ότι **η ομιλία με ανοιχτή ακρόαση δεν είναι πιο ασφαλής από την κανονική χρήση χειρός**, ενώ άλλες έρευνες έδειξαν ότι οι οδηγοί **επιηρεάζονται ιδιαίτερα όταν συντάσσουν γραπτό μήνυμα** (και κάπως λιγότερο όταν λαμβάνουν).

Αντιμετώπιση

Μέτρα για την αντιμετώπιση της απόσπασης της προσοχής των οδηγών μπορούν να περιλαμβάνουν εστιασμένη **ενίσχυση των κανόνων κυκλοφορίας, καμπάνιες** για την αφύπνιση των οδηγών και **εκπαίδευση** και άσκηση των οδηγών, ιδιαίτερα για τις ομάδες υψηλού κινδύνου (π.χ. νέοι οδηγοί, συχνοί παραβάτες κλπ.). Τεχνολογικές βελτιώσεις με σκοπό έναν πιο εργονομικό σχεδιασμό των εντός του οχήματος συσκευών αναπτύσσονται ταχέως. Ωστόσο, τα σχετικά αποτελέσματα στη βελτίωση της ασφάλειας θα πρέπει να επιβεβαιωθούν. Μελλοντική έρευνα θα πρέπει να επικεντρωθεί στη χρήση του κινητού τηλεφώνου, στην ανάλυση της ξεχωριστής επίδρασης διαφόρων παραγόντων απόσπασης προσοχής, καθώς και στη συνδυασμένη επιρροή όλων των παραγόντων απόσπασης, και στη σχέση μεταξύ της απόσπασης προσοχής, της αντίστοιχης συμπεριφοράς του οδηγού και του κινδύνου ατυχήματος.

2.2.4 Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments (2006)

Σκοπός

Η έρευνα επικεντρώθηκε στις **επιπτώσεις της συνομιλίας στο κινητό τηλέφωνο σε προσομοιωμένη οδήγηση** σε διαφορετικά περιβάλλοντα κυκλοφορίας.

Μεθοδολογία

Σαράντα οχτώ (48) συμμετέχοντες οδήγησαν μια απόσταση 70 km σε μια διαδρομή που διέσχιζε **διαφορετικά περιβάλλοντα**. Τα περιβάλλοντα ήταν: αγροτικό περιβάλλον με όριο ταχύτητας 90 km/h, αγροτικό περιβάλλον με όριο ταχύτητας 70 km/h, αστικό περιβάλλον χαμηλής πολυπλοκότητας, αστικό περιβάλλον μέσης πολυπλοκότητας και αστικό περιβάλλον υψηλής πολυπλοκότητας. Στο πείραμα λήφθηκε υπόψη η **λειτουργία του τηλεφώνου** (ανοιχτή ακρόαση ή κανονική χρήση χειρός) ως ένας μεταξύ των υποκειμένων παράγοντας καθώς και η χρήση τηλεφώνου (ναι / όχι) σε συνδυασμό με το περιβάλλον κυκλοφορίας, ως δυο εντός των υποκειμένων παράγοντες.

Αποτελέσματα

Η απόδοση που απεικονιζόταν σε έναν **ανιχνευτή περιφερειακών εργασιών (PDT)** κατά τη διάρκεια της οδήγησης - το PDT είναι ένα μέτρο του φόρτου πνευματικής ενασχόλησης- χειροτέρευσε λόγω της τηλεφωνικής συνομιλίας σε όλα τα περιβάλλοντα. Η απόδοση PDT ήταν, ωστόσο, αξιοσημείωτα χαμηλή στο σύνθετο αστικό περιβάλλον, ακόμα και όταν οι συμμετέχοντες δεν χρησιμοποιούσαν κινητό τηλέφωνο. **Η ταχύτητα οδήγησης μειώθηκε λόγω τηλεφωνικής συνομιλίας** σε όλα τα περιβάλλοντα για τη χρήση κινητού τηλεφώνου χωρίς ανοιχτή ακρόαση, αλλά μόνο σε δύο περιβάλλοντα για χρήση ανοιχτής ακρόασης- στο αγροτικό περιβάλλον για όριο ταχύτητας 90 km/h και στο σύνθετο αστικό περιβάλλον. Οι επιπτώσεις στην ταχύτητα θα μπορούσαν να ερμηνευθούν ως μια **προσπάθεια αντιστάθμισης του αυξημένου φόρτου πνευματικής ενασχόλησης**.

2.2.5 Effects of naturalistic cell phone conversations on driving performance (2004)

Σκοπός

Ο αριθμός των οδηγών αυτοκινήτων που ομιλούν στο κινητό τηλέφωνο αυξάνεται, όμως η επίπτωση αυτής της συμπεριφοράς στην οδήγηση είναι ασαφής. Επίσης ασαφής είναι και η σχέση ανάμεσα στο επίπεδο δυσκολίας μιας τηλεφωνικής συνομιλίας και στην επακόλουθη διάσπαση προσοχής. **Σκοπός** της παρούσας έρευνας είναι να εξετάσει τις **επιπτώσεις της συνομιλίας στο κινητό τηλέφωνο, σε πραγματικές συνθήκες οδήγησης, στη συμπεριφορά του οδηγού**.

Μεθοδολογία

Η έρευνα χρησιμοποίησε έναν **προσομοιωτή οδήγησης** για να καθορίσει το αποτέλεσμα που έχουν εύκολες και δύσκολες τηλεφωνικές συνομιλίες στο κινητό πάνω στην οδηγική συμπεριφορά.

Αποτελέσματα

Η χρήση κινητού τηλεφώνου οδήγησε τους συμμετέχοντες να έχουν υψηλότερη διακύμανση στη θέση του πεντάλ του γκαζιού, να οδηγούν πιο αργά με διακύμανση

2. Βιβλιογραφική Ανασκόπηση

στην ταχύτητα και να αναφέρουν υψηλότερο επίπεδο φόρτου ενασχόλησης ανεξαρτήτως από το επίπεδο δυσκολίας της συζήτησης. Συμπερασματικά, οι οδηγοί μπορούν να αντιμετωπίσουν το επιπρόσθετο άγχος λόγω τηλεφωνικών συνομιλιών υπομένοντας **υψηλότερους φόρτους ενασχόλησης** ή θέτοντας μειωμένους στόχους επίδοσης.

Επίπτωση στη Βιομηχανία

Καθώς ένας αυξανόμενος αριθμός ανθρώπων ομιλούν στο κινητό τηλέφωνο κατά την οδήγηση, τα ατυχήματα που προκαλούνται από αποσπασμένους οδηγούς που χρησιμοποιούν κινητό τηλέφωνο προκαλούν αναστατώσεις στον **επιχειρηματικό τομέα**, όπως επίσης και τραυματισμούς, αναπηρίες και μόνιμη απώλεια προσωπικού.

2.2.6 Using mobile telephones: cognitive workload and attention resource allocation (2004)

Σκοπός

Η απόσπαση της προσοχής έχει αναγνωριστεί ως μια από τις κυριότερες αιτίες συμβάντων οδικής κυκλοφορίας και τα κινητά τηλέφωνα είναι συσκευές (μεταξύ πολλών άλλων ηλεκτρονικών συσκευών) που μπορούν να αποσπάσουν τον οδηγό μέσα από αλλαγές στο φόρτο ενασχόλησης. **Σκοπός της έρευνας είναι να ανιχνεύσει τα πεδία στα οποία ο οδηγός αφιερώνει την προσοχή του ενώ χρησιμοποιεί το κινητό τηλέφωνο.**

Μεθοδολογία

Σαράντα συμμετέχοντες ολοκλήρωσαν σε επίπεδο χειρισμού οχήματος και επεξεργασίας πληροφοριών μια διαδρομή αυτοκινητόδρομου χαμηλής πολυπλοκότητας. Ένας **ανιχνευτής περιφερειακών εργασιών (PDT)** υιοθετήθηκε για να μετρήσει τον φόρτο πνευματικής ενασχόλησης. Συγκρίθηκαν οι επιπτώσεις διαφόρων τύπων συνομιλίας (απλή, σύνθετη) και τύπων τηλεφωνικής λειτουργίας (ανοιχτή ακρόαση, χειρός) σε βασικές συνθήκες.

Αποτελέσματα

Οι χρόνοι αντίδρασης των συμμετεχόντων αυξήθηκαν σημαντικά όταν συνομιλούσαν αλλά δεν υπήρξε κάποιο όφελος από τη χρήση ανοιχτής ακρόασης αντί για κανονική χρήση χειρός του κινητού σε αγροτικά περιβάλλοντα ή αυτοκινητόδρομους. Αναφορικά με τα κινητά τηλέφωνα, το περιεχόμενο της συζήτησης ήταν πολύ πιο σημαντικό για την οδήγηση και την απόσπαση προσοχής παρά ο τύπος της τηλεφωνικής λειτουργίας κατά την οδήγηση σε αυτοκινητόδρομο ή παρόμοιο τύπο οδού. Όσο πιο δύσκολη και πολύπλοκη η συνομιλία, τόσο σημαντικότερη η πιθανή αρνητική επίπτωση στην απόσπαση προσοχής του οδηγού.

2.2.7 Impact of mobile phone use on car-following behaviour of young drivers (2015)

Σκοπός

Η ταυτόχρονη εκτέλεση πολλαπλών εργασιών, όπως η ταυτόχρονη χρήση του κινητού τηλεφώνου και ο χειρισμός ενός αυτοκινήτου, αποτελεί σημαντικό παράγοντα απόσπασης που χειροτερεύει την απόδοση στην οδήγηση και εξελίσσεται σε κύριο παράγοντα αυτοκινητικών ατυχημάτων. Η παρούσα έρευνα **εξετάζει την επίδραση τηλεφωνικών συνομιλιών στο χειρισμό των αυτοκινήτων στην οδήγηση.**

Μεθοδολογία

Χρησιμοποιήθηκε **προσομοιωτής CARS-Q Advanced Driving** για να εξεταστεί μια ομάδα αυστραλιανών οδηγών ηλικίας 18-26 σε μια εργασία οδήγησης σε τρεις τυχαίες τηλεφωνικές καταστάσεις: βασική (καμία τηλεφωνική συνομιλία), ανοιχτή ακρόαση και κανονική λειτουργία χρήσης κινητού. Εφαρμόστηκε **επαναληπτική παλινδρόμηση (ANOVA)** για να εξεταστεί η επίδραση της απόσπασης λόγω κινητού τηλεφώνου σε επιλεγμένες μεταβλητές σχετικές με το αυτοκίνητο, όπως η ταχύτητα οδήγησης, οι αποστάσεις και οι χρόνοι διαχωρισμού. Για να εξετάσουν περαιτέρω τη συμπεριφορά των οχημάτων των αποσπασμένων οδηγών, οι χρόνοι διαχωρισμού προτυποποιήθηκαν χρησιμοποιώντας εξίσωση γενικευμένης εκτίμησης (GEE).

Αποτελέσματα

Συνολικά, οι οδηγοί έτειναν να επιλέγουν **χαμηλότερες ταχύτητες οδήγησης, μεγαλύτερες αποστάσεις μεταξύ των οχημάτων και μεγαλύτερους χρόνους διαχωρισμού** όταν είχαν είτε ανοιχτής ακρόασης είτε κανονικής χρήσης τηλεφωνικές συνομιλίες, υποδεικνύοντας συμπεριφορά αντιστάθμισης του πιθανού κινδύνου. Επιπλέον, οι τηλεφωνικές συζητήσεις κατά την οδήγηση επηρέασαν τη συμπεριφορά οδήγησης έτσι ώστε αυξήθηκε η μεταβλητότητα της ταχύτητας, των αποστάσεων μεταξύ των οχημάτων και της επιτάχυνσης και επιβράδυνσης. Έπειτα από έλεγχο διαφόρων εξωγενών παραγόντων, το μοντέλο GEE προβλέπει αύξηση 0.33 δευτερολέπτων στους χρόνους διαχωρισμού όταν ο οδηγός έχει συνομιλία ανοιχτής ακρόασης (hands-free) και αύξηση 0.75 δευτερολέπτων σε συνομιλία κανονικής χρήσης τηλεφώνου (handheld). Τα ευρήματα θα βελτιώσουν τη συνολική κατανόηση της απόσπασης στην οδηγική συμπεριφορά, και ιδιαίτερα το χειρισμό του οχήματος (car-following behavior) που είναι περισσότερο κρίσιμος στον καθορισμό οπίσθιων συγκρούσεων.

2.2.8 Driving and telephoning: Relative accident risk when using hand-held and hands-free mobile phones (2010)

Σκοπός

Η πειραματική έρευνα δείχνει ότι η χρήση κινητού τηλεφώνου κατά την οδήγηση χειροτερεύει την οδήγηση και έχει θεωρηθεί σε μεγάλο βαθμό ότι αυτή η χειροτέρευση στην οδήγηση είναι αποτέλεσμα γνωστικής παρά φυσικής απόσπασης. Αυτή η

αντίληψη υποστηρίζεται εν μέρει από εμπειρικά δεδομένα που δείχνουν ότι η χρήση κινητών τηλεφώνων με ανοιχτή ακρόαση συνδέεται με χειρότερη οδήγηση εξίσου όπως και η κανονική χρήση χειρός του κινητού. Στην παρούσα έρευνα **εξετάζεται ο κίνδυνος ατυχήματος κατά την ανοιχτή ακρόαση ή την κανονική λειτουργία χειρός κινητού.**

Μεθοδολογία

Για το σκοπό αυτό **εξετάστηκε ένα δείγμα** 4307 οδηγών που συμμετείχαν σε ατυχήματα το 2007. Επιπρόσθετα, δεδομένα από παρόμοια έρευνα του 1997 (N=5007) χρησιμοποιήθηκαν με σκοπό να ληφθούν περισσότερες παρατηρήσεις. Ο σχετικός κίνδυνος εκτιμήθηκε χρησιμοποιώντας οιονεί επαγόμενη έκθεση σε ατυχήματα με περισσότερα του ενός οχήματα.

Αποτελέσματα

Τα αποτελέσματα από τις δύο έρευνες έδειξαν σημαντική αύξηση του κινδύνου ατυχήματος για την ανοιχτή ακρόαση και για την κανονική χρήση χειρός και την ανοιχτή ακρόαση μαζί. Μη σημαντική τάση παρατηρήθηκε για αυξημένο κίνδυνο λόγω χρήσης ανοιχτής ακρόασης. Ωστόσο, χωριστές αναλύσεις από δεδομένα του 2007 σε στατιστικά σημαντικές εκτιμήσεις σχετικού κινδύνου για κανένα από τα δυο είδη χρήσης του τηλεφώνου. Οι χρήστες χειρός τηλεφώνου είχαν περισσότερη τάση να αποδίδουν τα ατυχήματα στα κινητά τηλέφωνα σε σχέση με τους χρήστες της ανοιχτής ακρόασης. Η έρευνα για τη χρήση κινητού τηλεφώνου κατά την οδήγηση **έδειξε αυξημένο κίνδυνο ατυχήματος όταν γίνεται χρήση κινητού τηλεφώνου** ακριβώς πριν το ατύχημα. Επιπλέον, **δεν παρατηρήθηκε διαφορά στον κίνδυνο ανάμεσα στην κανονική χρήση χειρός και την ανοιχτή ακρόαση.** Τέλος, οι κανονικοί χρήστες χειρός τηλεφώνου που συμμετείχαν κατά λάθος στο ατύχημα τείνουν να αποδίδουν το ατύχημα στη χρήση του κινητού τηλεφώνου, σε σχέση με τους χρήστες ανοιχτής ακρόασης που συμμετείχαν κατά λάθος στο ατύχημα.

2.2.9 Aggressive driving: an observational study of driver, vehicle, and situational variables (2004)

Σκοπός

Σκοπός της έρευνας ήταν η παρατήρηση των συμπεριφορών που περιλαμβάνονται συνήθως στη «λίστα της επιθετικής οδήγησης», και περιλάμβαναν **κορνάρισμα, διάσχιση μιας ή περισσότερων λωρίδων μπροστά από άλλα οχήματα και οδήγηση στη λωρίδα έκτακτης ανάγκης.**

Μεθοδολογία

Πάνω από 2000 επιθετικές συμπεριφορές οδήγησης παρατηρήθηκαν σε συνολικά 72 ώρες και σε έξι διαφορετικές θέσεις. Επιπλέον, ένα **δείγμα** 7200 οδηγών εξετάστηκε επίσης στους ίδιους χρόνους και περιοχές. Οι σχετικοί κίνδυνοι (RRs) και οι λόγοι πιθανότητας (ODs) υπολογίστηκαν για να δείξουν τη σχετική πιθανότητα ότι διαφορετικοί οδηγοί θα διαπράξουν επιθετικές συμπεριφορές υπό διαφορετικές συνθήκες.

Αποτελέσματα

Ο ρυθμός των επιθετικών δράσεων που παρατηρήθηκαν σε αυτή τη μελέτη μειώθηκε κατά σειρά από τη συχνότερη συμπεριφορά της διάσχισης κατά μήκος μιας μονής λωρίδας, στο κορνάρισμα και τέλος στις λιγότερο συχνές συμπεριφορές της διάσχισης κατά μήκος πολλαπλών λωρίδων και της οδήγησης στη ΛΕΑ. Αναφορικά με τα ποσοστά στον οδηγικό πληθυσμό, οι **άνδρες ήταν πιο πιθανό από τις γυναίκες να διαπράττουν επιθετικές ενέργειες** και οι διαφορές αυξήθηκαν καθώς η σοβαρότητα της δράσης αυξανόταν. Οι οδηγοί που ήταν 45 ετών ή **μεγαλύτεροι ήταν λιγότερο πιθανό να οδηγήσουν επιθετικά από τους νεότερους**. Η παρουσία επιβατών **συνδέθηκε με ελαφρά αλλά σταθερή μείωση της επιθετικής οδήγησης** όλων των τύπων και ειδικά του κορναρίσματος. Υπήρξε **έντονη γραμμική συσχέτιση** μεταξύ κυκλοφοριακής συμφόρησης και συχνότητας επιθετικών συμπεριφορών, αλλά οφειλόταν στον αριθμό των οδηγών στο δρόμο. Ωστόσο, **όταν η αξία του χρόνου ήταν υψηλή** (όπως στις ώρες αιχμής), **η πιθανότητα επιθετικής οδήγησης**, μετά την προσαρμογή στον αριθμό των οδηγών στο δρόμο, **ήταν υψηλότερη** από ό, τι όταν η αξία του χρόνου ήταν χαμηλή (κατά τη διάρκεια της εβδομάδας εκτός αιχμής ή το σαββατοκύριακο). Τα αποτελέσματα αφήνουν υπαινιγμούς για τροποποιήσεις της συμπεριφοράς των οδηγών και για τον περιβαλλοντικό σχεδιασμό.

2.2.10 Taking a look behind the wheel: An investigation into the personality predictors of aggressive driving (2012)

Σκοπός

Η παρούσα μελέτη αξιολόγησε ένα θεωρητικό μοντέλο της σχέσης ανάμεσα σε έξι πτυχές της **προσωπικότητας του οδηγού** (δηλ. τον θυμό στην οδήγηση και τους 5 Μεγάλους Παράγοντες προσωπικότητας), **την επιθετική οδήγηση** και δύο αποτελέσματα της επιθετικής οδήγησης: **τα ατυχήματα με οχήματα με κινητήρα και τη «βιαιότητα» στις μετακινήσεις**.

Μεθοδολογία

Δεδομένα από 308 οδηγούς που λήφθηκαν από δύο γραφεία αδειοδότησης οχημάτων αναλύθηκαν χρησιμοποιώντας **προτυποποίηση δομικών εξισώσεων**. Όπως αναμενόταν, η επιθετική οδήγηση προέβλεπε συγκρούσεις και «βιαιότητα» στις μετακινήσεις. Με βάση τις συσχετίσεις μηδενικής τάξης, η συναισθηματική σταθερότητα, η συγκαταβατικότητα και η ευσυνειδησία σχετίζονταν με την επιθετική οδήγηση με τον αναμενόμενο τρόπο. Ωστόσο, η εικόνα άλλαξε όταν εξετάστηκε η **κοινή επίδραση όλων των μεταβλητών** μέσω προτυποποίησης δομικών εξισώσεων.

Αποτελέσματα

Τελικώς, ορίστηκε ένα μοντέλο στο οποίο η προσωπικότητα του οδηγού προέβλεπε την επιθετική οδήγηση, η οποία με τη σειρά της προέβλεπε τις συγκρούσεις και τις «βίαιες» μετακινήσεις. **Οι οδηγοί με υψηλό βαθμό οργής στην οδήγηση και χαμηλό βαθμό συγκαταβατικότητας αναφέρθηκε ότι οδήγησαν πιο επιθετικά**. Υπαινιγμοί για τους επαγγελματίες της οδικής ασφάλειας και για τους ερευνητές είναι υπό συζήτηση.

2.3 Συναφείς μεθοδολογίες

Στο υποκεφάλαιο αυτό παρουσιάζονται συναφείς μεθοδολογίες που έχουν χρησιμοποιηθεί σε άλλες έρευνες όσον αφορά στην οδική ασφάλεια.

2.3.1 Modeling the behavior of novice young drivers during the first year after licensure (2010)

Σκοπός

Οι αρχάριοι οδηγοί συχνά εμφανίζονται ευάλωτοι μπροστά στον κίνδυνο ατυχήματος, κάτι το οποίο μεταφράζεται σε μεγάλη συμμετοχή τους στον κατάλογο των οδικών ατυχημάτων. Σκοπός της συγκεκριμένης έρευνας είναι η **ανάλυση της συμπεριφοράς των νέων οδηγών με τη χρήση ενός συστήματος καταγραφής δεδομένων, το οποίο βρίσκεται εντός του οχήματος.**

Μεθοδολογία

Για τις ανάγκες της εργασίας εξετάστηκαν 62 οδηγοί (36 άνδρες – 26 γυναίκες) ηλικίας 17 χρονών κατά την περίοδο των **πρώτων 12 μηνών μετά την απόκτηση της άδειας οδήγησης**. Για τους 3 πρώτους μήνες, ήταν απαραίτητο να υπάρχει στο αυτοκίνητο ένας γονιός κατά τη διάρκεια της οδήγησης, ενώ μετά το πέρας αυτού του χρονικού διαστήματος, μπορούσαν να οδηγούν μόνοι. Το **σύστημα IVDR** (In-Vehicle Data Recorder) κατέγραφε δεδομένα όπως η ταχύτητα και η επιτάχυνση του οχήματος ώστε να μπορεί να διακρίνει πότε ο οδηγός πραγματοποιεί επικίνδυνους ελιγμούς, οι οποίοι κατηγοριοποιούνταν ανάλογα με το είδος τους (αλλαγή λωρίδας, απότομο φρενάρισμα ή επιτάχυνση κτλ.) και τη σοβαρότητά τους (χαμηλή, μέση, υψηλή). Όλες αυτές οι πληροφορίες μεταδίδονταν μέσω ασύρματων δικτύων σε έναν διακομιστή, ο οποίος διατηρούσε αρχείο για κάθε ένα από τα εξεταζόμενα οχήματα, δημιουργώντας έτσι το προφίλ του κάθε αρχάριου οδηγού. Οι γονείς και οι νεαροί οδηγοί ήταν σε θέση να ελέγξουν τα δεδομένα αυτά μέσω μιας εφαρμογής που τους επέτρεπε να έχουν πρόσβαση στο προφίλ τους.

Αποτελέσματα

Όπως ήταν αναμενόμενο, τα αποτελέσματα επιβεβαίωσαν ότι **οι άνδρες οδηγοί είναι πιο επιρρεπείς στον κίνδυνο** και ότι όταν οδηγούν χωρίς επίβλεψη, οι αρχάριοι οδηγοί γίνονται πιο «**επιθετικοί**». Όσον αφορά την οδήγηση με παρόντες τους γονείς, ο κίνδυνος ατυχήματος αποτελεί συνάρτηση και της οδηγικής συμπεριφοράς του ενήλικα. Τέλος, σημαντικό ρόλο στην μείωση της επικινδυνότητας των νέων οδηγών παίζει και ο έλεγχος των δεδομένων που προκύπτουν από το IVDR. **Οι νεαροί οι οποίοι ελέγχονται ηλεκτρονικά από τους γονείς τους** καθ' όλη τη διάρκεια της έρευνας **γίνονται πιο προσεκτικοί** όταν οδηγούν μόνοι τους, ενώ αντίθετα όσοι δεν έχουν την συνεχή επίβλεψη των γονέων μέσω της ηλεκτρονικής εφαρμογής, δεν δείχνουν διάθεση να βελτιώσουν την τάση που έχουν για επικίνδυνη οδήγηση.

2.3.2 Driving behavior and traffic safety: an acceleration-based safety evaluation procedure for smartphones (2014)

Σκοπός

Αυτή η εργασία πραγματεύεται την **ανάπτυξη μιας πρωτότυπης εφαρμογής για ηλεκτρονικές, κινητές συσκευές η οποία θα εκτιμάει τον βαθμό ασφάλειας των οδηγών** που βρίσκονται εν κινήσει μετρώντας την διαμήκη και πλευρική τους επιτάχυνση, μέσω χρήσης των συστημάτων GPS . Όπου απαιτείτο, η εφαρμογή θα εμφανίζει μια προειδοποίηση στους οδηγούς ώστε να συμμορφώσουν την συμπεριφορά τους στον δρόμο.

Μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε ήταν η εξής: Τοποθετώντας σε έναν άξονα x, y τις επιταχύνσεις του οχήματος (πλευρική και διαμήκης αντίστοιχα), εκτιμήθηκε η συμπεριφορά του οδηγού στο τιμόνι (επιθετικός ή μη). Χρησιμοποιήθηκε, επίσης, ο κύκλος τριβής του οχήματος, ο οποίος είναι συνάρτηση των χαρακτηριστικών των ελαστικών του αυτοκινήτου και των χαρακτηριστικών του οδοστρώματος. Λαμβάνοντας, επιπλέον, υπόψη την εμπειρία του οδηγού και τον τύπο του αυτοκινήτου δημιουργήθηκε το **“Διάγραμμα Οδηγικής Συμπεριφοράς” (Driving Style Diagram – DSD)**, το οποίο αποτελεί μία συσχέτιση όλων των παραμέτρων που αναφέρθηκαν παραπάνω. Το πείραμα πραγματοποιήθηκε με το ίδιο όχημα και το ίδιο smartphone και πήραν μέρος 5 οδηγοί με διαφορετικά οδηγικά χαρακτηριστικά. Η συμπεριφορά κάθε οδηγού εκτιμήθηκε με βάση το ποσοστό υπέρβασης των **ορίων επιτάχυνσης του DSD**.

Αποτελέσματα

Η έρευνα είχε ως αποτέλεσμα ότι η καλύτερη τιμή για να διαχωρίσεις τους επιθετικούς με τους ασφαλείς οδηγούς μπορεί να οριστεί στο 9% για το ποσοστό υπέρβασης των ορίων επιτάχυνσης του DSD.

2.3.3 Development of a driving data recorder (1994)

Σκοπός

Η έρευνα αυτή αποτελεί μία από τις πρώτες προσπάθειες ανάπτυξης ενός οργάνου καταγραφής δεδομένων από τον εγκέφαλο του αυτοκινήτου. Ο σκοπός της είναι να παρουσιάσει μία **γενική επισκόπηση του οργάνου αυτού (DDR)** και να παραθέσει τρόπους χρησιμοποίησής του.

Μεθοδολογία

Βασικά στοιχεία:

- Το DDR είναι αρκετά **μικρό** και ελαφρύ ώστε να μην δημιουργεί πρόβλημα στο όχημα.
- Είναι, επίσης, σχεδιασμένο ώστε **να μην καταναλώνει** μεγάλη ποσότητα **μπαταρίας** από το αυτοκίνητο.

2. Βιβλιογραφική Ανασκόπηση

- Μπορεί να χρησιμοποιηθεί για **9.000 ώρες** οδήγησης ή 100.000 χιλιόμετρα.
- Τα δεδομένα, τα οποία αποθηκεύονται αρχικά σε μία **κάρτα μνήμης**, μπορούν να αναλυθούν μέσω προσωπικού Η/Υ.

Υπάρχουν δύο τρόποι με τους οποίους μπορούν να εγγραφούν τα δεδομένα οδήγησης:

- a) Σύστημα που βασίζεται στο χρόνο, όπου τα δεδομένα τοποθετούνται σε διαδοχική σειρά
- b) Σύστημα που βασίζεται στη συχνότητα, όπου η τοποθέτηση των δεδομένων γίνεται με βάση είτε τη συχνότητα χρήσης είτε τη συχνότητα εμφάνισης.

Ανάλογα με τις εντολές εγγραφής, το DDR μπορεί να χρησιμοποιηθεί για να συλλέγει στοιχεία όπως ο συνολικός χρόνος οδήγησης, η συνολική απόσταση που διένυσε ο οδηγός, πόσες φορές άναψε ή έσβησε η μηχανή του αυτοκινήτου κ.τ.λ.

Αποτελέσματα

Εν κατακλείδι, το DDR σχεδιάστηκε για να καταγράψει τις συνθήκες οδήγησης ενός αυτοκινήτου σε κανονικές συνθήκες οδοστρώματος και να αναλύσει τη συμπεριφορά του οδηγού στο δρόμο. Επομένως, ήταν δυνατό να συσσωρευτούν **ποσοτικά στοιχεία οδήγησης** τα οποία θα **καταδεικνύουν το προφίλ του οδηγού**.

2.3.4 In-vehicle data recorders for monitoring and feedback on drivers' behavior (2008)

Σκοπός

Το αντικείμενο της εργασίας αυτής είναι η **διερεύνηση της συμπεριφοράς των οδηγών μέσω ενός συστήματος καταγραφής δεδομένων, ονόματι Drive Diagnostics**, το οποίο είναι προσαρμοσμένο στο εσωτερικό του αυτοκινήτου, έχει διαστάσεις 11x6x3 εκατοστά και φορτίζεται από την μπαταρία του οχήματος. Σκοπός της συγκεκριμένης εργασίας είναι να εκτιμήσει τη συμπεριφορά των οδηγών, καταμετρώντας τον τρόπο οδήγησής τους μέσω συγκεκριμένων παραμέτρων.

Μεθοδολογία

Η μέθοδος της έρευνας ήταν η εξής: Το σύστημα συλλέγει στοιχεία όπως η επιτάχυνση του οχήματος (κατά x, y), η ταχύτητά του, η ακριβής θέση του μέσω συντεταγμένων με τη βοήθεια του GPS, η κατανάλωση καυσίμων, ο συνολικός χρόνος ταξιδιού κτλ. Σε αντίθεση με άλλα παρόμοια συστήματα, η **μετάδοση των πληροφοριών μέσω του Drive Diagnostics γίνεται συνεχώς και σε πραγματικό χρόνο**, παρόλο που το όχημα βρίσκεται εν κινήσει. Προκειμένου να μετρηθούν συγκεκριμένες παράμετροι, όπως η συχνότητα αλλαγής λωρίδων, τα απότομα φρεναρίσματα και επιταχύνσεις και η υπερβολική ταχύτητα, αυτές οι πληροφορίες μεταφέρονται μέσω ασύρματων δικτύων σε έναν διακομιστή, ο οποίος τις αποθηκεύει ώστε για κάθε οδηγό να υπάρχει ένας συγκεκριμένος φάκελος με τα χαρακτηριστικά του (αριθμός οχήματος, μέση κατανάλωση καυσίμων, συμμετοχή σε οδικά ατυχήματα κτλ.). Τέλος, οι οδηγοί

κατατάσσονται σε τρεις κατηγορίες (προσεκτικοί, κανονικοί και επιθετικοί), ανάλογα με το προφίλ που προκύπτει από τα δεδομένα.

Αποτελέσματα

Τα αποτελέσματα έδειξαν ότι η αρχική έκθεση των συμμετεχόντων στο πείραμα είχε σημαντική επίδραση στη βελτίωση της οδηγικής τους συμπεριφοράς και της οδικής ασφάλειας. Επιπλέον, **η πρόσβαση που είχαν στα δεδομένα του συστήματος τους βοήθησε να καταλάβουν τα σφάλματα** που διαπράττουν στο δρόμο. Παρόλα αυτά, η έρευνα έδειξε ότι αν δεν υπάρξει περαιτέρω ενδιαφέρον για την καταγραφή της οδηγικής συμπεριφοράς των ανθρώπων, μετά από ένα μικρό χρονικό διάστημα δεν παρατηρείται καμία πρόοδος. Στην εν λόγω εργασία, η επίδραση του Drive Diagnostics και στους τριάντα-τρεις οδηγούς που χρησιμοποιήθηκε είχε εξαφανιστεί με το πέρασ πέντε μηνών.

2.3.5 Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones (2011)

Σκοπός

Σκοπός της έρευνας αυτής είναι η **ανάπτυξη μιας εφαρμογής για κινητά τηλεφώνά** με λογισμικό Android η οποία παρακολουθεί το όχημα μέσω του συστήματος OBD-II (On Board Diagnostics) και ανιχνεύει οδικά ατυχήματα.

Το σύστημα OBD αναπτύχθηκε στις Ηνωμένες Πολιτείες Αμερικής και είχε ως στόχο την ανίχνευση μηχανικών προβλημάτων στο όχημα, τα οποία προκαλούσαν υψηλά επίπεδα εκπομπής ρύπων, πάνω από τα αποδεκτά όρια. Για να το επιτύχει αυτό, το σύστημα είναι συνδεδεμένο με την μηχανή του αυτοκινήτου και **όταν ανιχνεύσει ένα πρόβλημα, το αποθηκεύει** στη μνήμη του ώστε να επιδιορθωθεί αργότερα από τους ειδικούς.

Μεθοδολογία

Η μέθοδος της εργασίας αυτής στηρίχθηκε κατά ένα σκέλος στο συγκεκριμένο σύστημα. Το κινητό τηλέφωνο συνδέεται μέσω Bluetooth με το μηχανήμα **OBD-II** και λαμβάνει πληροφορίες σχετικά με την κατάσταση του οχήματος. Επιπλέον, μέσω **GPS**, γνωρίζει την ταχύτητα και την ακριβή τοποθεσία του. Για την ανίχνευση του οδικού ατυχήματος, 2 παράμετροι έχουν ληφθεί υπόψιν. Η μία αφορά την ενεργοποίηση του συστήματος των αερόσακων και η άλλη την επιτάχυνση του αυτοκινήτου, όπου πολύ υψηλές τιμές του G λαμβάνουν χώρα μόνον όταν σοβαρά ατυχήματα συμβαίνουν, καθώς ούτε τα μονοθέσια της Formula 1 μπορούν να πιάσουν τέτοιες τιμές. Επομένως, **μόλις ανιχνευθεί ότι ένα ατύχημα έχει γίνει, αμέσως ειδοποιείται** είτε με γραπτό μήνυμα είτε με email ένα άτομο της επιλογής του οδηγού, το οποίο έχει αποφασιστεί από την αρχή της έρευνας, και γίνεται κλήση στην άμεση βοήθεια και την αστυνομία. Οι συντεταγμένες του οχήματος που λαμβάνονται από το GPS είναι απαραίτητες για την όσο το δυνατόν πιο γρήγορη άφιξη του ασθενοφόρου.

Αποτελέσματα

Τα αποτελέσματα της έρευνας ήταν άκρως **ενθαρρυντικά όσον αφορά την οδική ασφάλεια** και τον χρόνο διακομιδής των εμπλεκομένων σε ατύχημα στο κοντινότερο νοσοκομείο. Ο χρόνος ειδοποίησης από την ώρα που συνέβη το ατύχημα υπολογίστηκε στα 3 δευτερόλεπτα, με τον μέγιστο χρόνο να φτάνει τα 6 δευτερόλεπτα. Η διαφορά έγκειται στην απόδοση σύνδεσης του κινητού τηλεφώνου στο διαδίκτυο.

2.3.6 Innovative motor insurance schemes: A review of current practices and emerging challenges (2017)

Στόχος

Ο στόχος της παρούσας εργασίας είναι να εφαρμόσει μια από τις πιο δημοφιλείς μεθοδολογίες που σχετίζονται με την **ασφάλιση των αυτοκινήτων**, την **Usage-based motor insurance (UBI)**. Συστήματα UBI, όπως Pay-as-you-drive (PAYD) και Pay-how-you-drive (PHYD), είναι νέες καινοτόμες ιδέες που έχουν αρχίσει πρόσφατα να πραγματοποιούνται και να αναλύονται. Η βασική ιδέα είναι ότι αντί για μια σταθερή τιμή, οι οδηγοί πρέπει να πληρώσουν ένα ασφάλιστρο με βάση τα ταξίδια τους και τη συμπεριφορά οδήγησης. Παρά το γεγονός ότι έχει εφαρμοστεί μόνο για μερικά χρόνια, φαίνεται να είναι μια πολλά υποσχόμενη πρακτική με ένα σημαντικό δυνητικό αντίκτυπο στην ασφάλεια της κυκλοφορίας, καθώς και για την άμβλυνση της κυκλοφοριακής συμφόρησης και τη μείωση των εκπομπών ρύπανσης.

Μεθοδολογία

Για το σκοπό αυτό, η **υπάρχουσα βιβλιογραφία** σχετικά με τα συστήματα UBI επανεξετάζει τυχόν κενά που μπορεί να υπάρχουν. Από τα ευρήματα προκύπτει ότι υπάρχει μια πολλαπλότητα και πολυμορφία των διαφόρων ερευνητικών μελετών στη σύγχρονη βιβλιογραφία. Εξετάζεται η συσχέτιση μεταξύ PAYD (βάση της ταξιδιωτικής συμπεριφοράς του οδηγού και της συχνότητας αυτής) και PHYD (βάση της οδηγικής συμπεριφοράς) και του κινδύνου σύγκρουσης των δύο αυτών συστημάτων. Επιπλέον, υπάρχουν ενδείξεις ότι η εφαρμογή UBI θα εξαλείψει το φαινόμενο των cross-subsidies, πράγμα που σημαίνει λιγότερο κόστος ασφάλισης για λιγότερο επικίνδυνους και εκτεθειμένους οδηγούς. Θα προβλέπει επίσης ένα ισχυρό κίνητρο για τους οδηγούς που βελτιώσουν την οδική συμπεριφορά τους, η οποία θα είχε ως αποτέλεσμα τη μείωση της έκθεσης στον κίνδυνο τόσο συνολικά όσο και ατομικά.

Στο τέλος αυτής της έρευνας αναφέρονται οι τρέχουσες και οι αναδυόμενες **προκλήσεις** σε αυτό το πεδίο έρευνας.

2.4 Σύνοψη

Οι εργασίες, οι οποίες εξετάστηκαν παραπάνω, καλύπτουν ευρύ φάσμα αναλύσεων, σχετικών με τη χρήση κινητού τηλεφώνου κατά την οδήγηση και κατά συνέπεια με την απόσπαση της προσοχής του οδηγού. Από τη **σύνθεση των βασικών τους σημείων**, προκύπτουν οι εξής παρατηρήσεις:

- Η χρήση του **κινητού τηλεφώνου** στην οδήγηση οδηγεί σε μεγαλύτερους χρόνους αντίδρασης, χαμηλότερες ταχύτητες, μεγαλύτερες αποστάσεις και χρόνους διαχωρισμού μεταξύ των οχημάτων.
- Ακόμα, οδηγεί σε μεγαλύτερη διακύμανση στη χρήση του πεντάλ του **γκαζιού** και σε πιο απότομα φρεναρίσματα.
- Τα παραπάνω δύο φαινόμενα πιθανόν εξηγούνται από την προσπάθεια του οδηγού για **αντισταθμιστική συμπεριφορά**.
- Ο **φόρτος ενασχόλησης** αυξάνεται λόγω της χρήσης του κινητού τηλεφώνου.
- Η οδήγηση με **ανοιχτή ακρόαση** δεν δείχνει να είναι πιο ασφαλής σε σχέση με την κανονική χρήση χειρός του κινητού τηλεφώνου, ενώ αυτό που φαίνεται ότι επηρεάζει τελικώς περισσότερο την απόσπαση προσοχής του οδηγού ήταν το περιεχόμενο και η δυσκολία της συζήτησης
- **Οι άνδρες** ήταν πιο πιθανό από τις **γυναίκες** να διαπράττουν επιθετικές ενέργειες.
- Οι οδηγοί που ήταν 45 ετών ή **μεγαλύτεροι** ήταν **λιγότερο πιθανό** να οδηγήσουν επιθετικά από τους **νεότερους**.
- Η **παρουσία επιβατών** συνδέθηκε με μείωση της επιθετικής οδήγησης.
- Όταν η **αξία του χρόνου** είναι υψηλή (όπως στις **ώρες αιχμής**), η πιθανότητα επιθετικής οδήγησης είναι υψηλότερη.
- Οδηγοί με πιο **επιθετικό και λιγότερο συγκαταβατικό χαρακτήρα** τείνουν να οδηγούν πιο νευρικά.
- Τα συστήματα καταγραφής της συμπεριφοράς των οδηγών φαίνεται να **βελτιώνουν τα επίπεδα οδικής ασφάλειας**. Ωστόσο, χρειάζεται η παρακολούθηση της συμπεριφοράς να συνεχιστεί σε βάθος χρόνου.
- Τα περισσότερα συστήματα καταγραφής συνδέονται με τον **"εγκέφαλο"** του οχήματος (π.χ. OBD, DDR).
- Η καταγραφή διαφόρων στοιχείων οδήγησης μπορεί να χρησιμοποιηθεί για την σκιαγράφηση του **προφίλ του οδηγού** (νέος, επιθετικός κλπ.), την αυτοβελτίωσή του ή ακόμα και την εφαρμογή ασφαλιστρων με βάση την οδική του συμπεριφορά.
- Η μεθοδολογία που ακολουθήθηκε στις περισσότερες έρευνες περιλάμβανε τη χρήση **προσομοιωτή**, προσομοιωτή σε συνδυασμό με **ερωτηματολόγια** ή οδήγηση σε **πραγματικές συνθήκες**.
- Η κυριότερη μέθοδος ανάλυσης των δεδομένων που εφαρμόστηκε είναι η **παλινδρόμηση μεικτών επιδράσεων** (ANOVA).
- Συχνά υιοθετήθηκε η χρήση **ανιχνευτή περιφερειακών εργασιών** (PDT) για την εκτίμηση του φόρτου πνευματικής ενασχόλησης.

3. Θεωρητικό Υπόβαθρο

3.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο στο οποίο βασίστηκε η ανάλυση του αντικειμένου της Διπλωματικής Εργασίας καθώς και βασικές στατιστικές μέθοδοι ανάλυσης. Για τον σκοπό της παρούσας εργασίας εφαρμόστηκε η γραμμική παλινδρόμηση (linear regression) και η δυαδική λογιστική παλινδρόμηση (binary logistic regression) για την εξαγωγή των δύο βασικών μοντέλων της ανάλυσης. Ο κύριος λόγος στον οποίο βασίστηκε η επιλογή αυτών των μεθόδων είναι η «φύση» της εξαρτημένης μεταβλητής. Συγκεκριμένα, στο μοντέλο που αναζητείται η σχέση μεταξύ των απότομων συμπεριφορών οδήγησης με διάφορα οδηγικά μεγέθη εφαρμόζεται η γραμμική παλινδρόμηση εφόσον η εξαρτημένη μεταβλητή (απότομα γεγονότα ανά απόσταση) είναι συνεχής μεταβλητή. Στο μοντέλο που εξετάζεται η κατάσταση χρήσης ή μη κινητού τηλεφώνου κατά την οδήγηση η εξαρτημένη μεταβλητή είναι διακριτή και παίρνει δύο τιμές (0 ή 1). Στη συνέχεια του κεφαλαίου αναλύεται η γραμμική και η δυαδική λογιστική παλινδρόμηση, τα κριτήρια αποδοχής ενός μαθηματικού προτύπου καθώς και ο τρόπος που επιτεύχθηκαν τα μαθηματικά μοντέλα πρόβλεψης μέσω λειτουργιών του ειδικού στατιστικού λογισμικού. Μεγάλο τμήμα των θεωρητικών στοιχείων αυτών έχει αναπτυχθεί και σε προηγούμενες διπλωματικές εργασίες.

3.2 Βασικές έννοιες Στατιστικής

Ο όρος **πληθυσμός** (population) αναφέρεται στο σύνολο των παρατηρήσεων του χαρακτηριστικού που ενδιαφέρει τη στατιστική έρευνα. Πρόκειται για ένα σύνολο στοιχείων που είναι τελείως καθορισμένα. Ένας πληθυσμός μπορεί να είναι πραγματικός ή θεωρητικός.

Ο όρος **δείγμα** (sample) αναφέρεται σε ένα υποσύνολο του πληθυσμού. Οι περισσότερες στατιστικές έρευνες στηρίζονται σε δείγματα, αφού οι ιδιότητες του πληθυσμού είναι συνήθως αδύνατο να καταγραφούν. Όλα τα στοιχεία που ανήκουν στο δείγμα ανήκουν και στον πληθυσμό χωρίς να ισχύει το αντίστροφο. Τα συμπεράσματα που θα προκύψουν από τη μελέτη του δείγματος θα ισχύουν με ικανοποιητική ακρίβεια για ολόκληρο τον πληθυσμό μόνο εάν το δείγμα είναι αντιπροσωπευτικό του πληθυσμού.

Με τον όρο **μεταβλητές** (variables) εννοούνται τα χαρακτηριστικά που ενδιαφέρουν να μετρηθούν και να καταγραφούν σε ένα σύνολο ατόμων. Οι μεταβλητές διακρίνονται στις παρακάτω κατηγορίες:

- **Ποιοτικές μεταβλητές** (qualitative variables): Είναι οι μεταβλητές των οποίων οι δυνατές τιμές είναι κατηγορίες διαφορετικές μεταξύ τους. Η χρήση αριθμών για την παράσταση των τιμών μίας τέτοιας μεταβλητής είναι καθαρά συμβολική και δεν έχει την έννοια της μέτρησης.
- **Ποσοτικές μεταβλητές** (quantitative variables): Είναι οι μεταβλητές με τιμές αριθμούς, που όμως έχουν τη σημασία της μέτρησης. Οι ποσοτικές μεταβλητές διακρίνονται με τη σειρά τους σε δύο μεγάλες κατηγορίες τις διακριτές και τις συνεχείς. Σε μία διακριτή μεταβλητή η μικρότερη μη μηδενική διαφορά που μπορούν να έχουν δύο τιμές είναι σταθερή ποσότητα. Αντίθετα σε μια συνεχή μεταβλητή δύο τιμές

3. Θεωρητικό Υπόβαθρο

μπορούν να διαφέρουν κατά οποιαδήποτε μικρή ποσότητα. Στην πράξη, συνεχής θεωρείται μια μεταβλητή όταν μπορεί να πάρει όλες τις τιμές σε ένα διάστημα, διαφορετικά θεωρείται διακριτή. Οι ποσοτικές μεταβλητές διακρίνονται με τη σειρά τους σε δυο μεγάλες κατηγορίες, τις **διακριτές** (ή ασυνεχείς) και τις **συνεχείς**.

1. Σε μια **διακριτή** μεταβλητή η μικρότερη μη μηδενική διαφορά που μπορούν να έχουν δύο τιμές της είναι σταθερή ποσότητα. Ένα τέτοιο παράδειγμα είναι ο αριθμός των μελών μιας οικογένειας.
2. Αντίθετα, σε μία **συνεχή** μεταβλητή δύο τιμές μπορούν να διαφέρουν κατά οποιαδήποτε μικρή ποσότητα. Ως παράδειγμα μπορούμε να αναφέρουμε την ηλικία, για την οποία η διαφορά ανάμεσα σε δυο τιμές θα μπορούσε να είναι χρόνια, μήνες, ημέρες, ώρες κλπ. Στην πράξη, συνεχής θεωρείται μια μεταβλητή όταν μπορεί να πάρει όλες τις τιμές σε ένα διάστημα, διαφορετικά θεωρείται διακριτή.

Μέτρα κεντρικής τάσης (measures of central tendency): Σε περίπτωση ανάλυσης ενός δείγματος $x_1, x_2, \dots, x_n$ η μέση τιμή υπολογίζεται σύμφωνα με τη σχέση:

$$\bar{x} = \frac{(x_1 + x_2 + \dots + x_n)}{n} = \frac{1}{n} * \sum_{i=1}^n x_i$$

Μέτρα διασποράς και μεταβλητότητας (measures of variability): Στην περίπτωση που τα δεδομένα αποτελούν ένα δείγμα η διακύμανση συμβολίζεται με s^2 και ορίζεται ως:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

όπου

$\bar{x}$ είναι ο δειγματικός μέσος, δηλαδή η μέση τιμή των παρατηρήσεων του δείγματος.

Η μαθηματική σχέση που δίνει την τυπική απόκλιση του δείγματος είναι:

$$s = (s^2)^{1/2} = \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right]^{1/2}$$

Για την περίπτωση συμμετρικά κατανεμημένου δείγματος σύμφωνα με έναν εμπειρικό κανόνα, προκύπτει ότι:

- Το διάστημα $(-s, +s)$ περιέχει περίπου το 68% των δεδομένων
- Το διάστημα $(-2s, +2s)$ περιέχει περίπου το 95% των δεδομένων
- Το διάστημα $(-3s, +3s)$ περιέχει περίπου το 99% των δεδομένων

Συνδιακύμανση (covariance of the two variables): Αποτελεί ένα μέτρο της σχέσης μεταξύ δύο περιοχών δεδομένων και δίνεται από τη σχέση:

$$\text{Cov}(x, y) = \left[\frac{1}{n-1} \right] \sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]$$

Μέτρα αξιοπιστίας είναι:

- το επίπεδο εμπιστοσύνης: που ορίζεται ως η αναλογία των περιπτώσεων που μια εκτίμηση θα είναι σωστή
- το επίπεδο σημαντικότητας, που ορίζεται ως η αναλογία των περιπτώσεων που ένα συμπέρασμα είναι εσφαλμένο.

Συσχέτιση Μεταβλητών- Συντελεστής Συσχέτισης

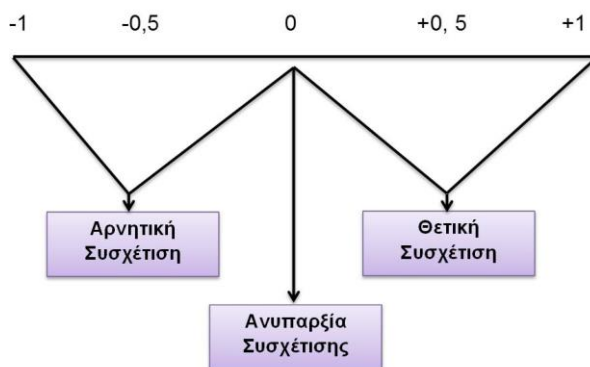
Έστω x, y δυο τυχαίες και συνεχείς μεταβλητές. Ο βαθμός της γραμμικής συσχέτισης των δυο μεταβλητών x και y , οι οποίες έχουν διασπορά σ_x^2 και σ_y^2 αντίστοιχα, και συνδιασπορά $\sigma_{xy} = \text{cov}[x, y]$ καθορίζεται με τον συντελεστή συσχέτισης (correlation coefficient) ρ , ο οποίος ορίζεται ως εξής :

$$\rho = \left(\frac{\sigma_{xy}}{\sigma_x} \right) \left(\frac{1}{\sigma_y} \right)$$

Ο **συντελεστής συσχέτισης ρ** εκφράζει τον βαθμό και τον τρόπο που οι δύο μεταβλητές συσχετίζονται. Δεν εξαρτάται από την μονάδα μέτρησης των x και y και παίρνει τιμές στο διάστημα $[-1, 1]$. Τιμές κοντά στο 1 δηλώνουν ισχυρή θετική συσχέτιση, τιμές κοντά στο -1 δηλώνουν ισχυρή αρνητική συσχέτιση και τιμές κοντά στο 0 δηλώνουν γραμμική ανεξαρτησία των x και y . Η εκτίμηση του συντελεστή συσχέτισης ρ γίνεται με την αντικατάσταση στην ανωτέρω εξίσωση της συνδιασποράς σ_{xy} και των διασπορών σ_x, σ_y , από όπου προκύπτει τελικά η έκφραση της εκτιμήτριας r . Η εκτίμηση του συντελεστή συσχέτισης ρ γίνεται με την αντικατάσταση στην ανωτέρω εξίσωση της συνδιασποράς σ_{xy} και των διασπορών σ_x και σ_y , από όπου προκύπτει τελικά η έκφραση της **εκτιμήτριας r** :

$$r(x, y) = \frac{\left[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]}{\left[\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1/2} \left(\sum_{i=1}^n (y_i - \bar{y})^2 \right)^{1/2} \right]}$$

3. Θεωρητικό Υπόβαθρο



Εικόνα 3.1.: Επεξήγηση πιθανών τιμών του Συντελεστή Συσχέτισης
Πηγή: Σύγγραμμα «Εφαρμογή Μεθόδων Ανάλυσης στην Έρευνα Αγοράς»

3.3 Βασικές Κατανομές

Όπως είναι γνωστό από τη θεωρία της στατιστικής, για να μελετηθούν τα διάφορα στατιστικά μεγέθη πρέπει να είναι γνωστή η μορφή της κατανομής που ακολουθούν οι τιμές τους. Παρακάτω παρατίθενται οι σημαντικότερες στατιστικές κατανομές που χρησιμοποιούνται στην ανάλυση οδικών δεδομένων.

3.3.1 Κανονική Κατανομή

Μία από τις πιο σημαντικές κατανομές πιθανότητας για συνεχείς μεταβλητές είναι η κανονική κατανομή ή κατανομή του Gauss. Η συνάρτηση πυκνότητας της κατανομής αυτής είναι:

$$F(x) = \left(\frac{1}{(2\pi\sigma)^{1/2}} \right) e^{[-(x-\mu)^2 / 2\sigma^2]}$$

όπου μ και σ είναι σταθερές ίσες με την μέση τιμή και την τυπική απόκλιση αντίστοιχα.

3.3.2 Κατανομή Poisson

Είναι γνωστό ότι η πιο κατάλληλη κατανομή για την περιγραφή τελείως τυχαίων διακριτών γεγονότων είναι η **κατανομή Poisson**. Μια τυχαία μεταβλητή x (όπως π.χ. το πλήθος των ατυχημάτων ή των νεκρών από οδικά ατυχήματα) θεωρείται ότι ακολουθεί κατανομή Poisson με παράμετρο λ ($\lambda > 0$), και γράφεται $x \sim P(\lambda)$, όταν έχει συνάρτηση μάζας πιθανότητας την:

$$F(x) = \frac{\mu^x * e^{-\mu}}{x!}$$

όπου $x = 0, 1, 2, 3, \dots$ και $x! = x * (x-1) * \dots * 3 * 2 * 1$

Η μέση τιμή και η διασπορά κατά Poisson είναι $E\{x\} = \mu$ και $\sigma^2\{x\} = \mu$ και είναι ίσες μεταξύ τους.

Η **κατανομή Poisson** αφορά στον αριθμό των “συμβάντων” σε ορισμένο χρονικό ή χωρικό διάστημα. Γενικά, ο αριθμός X των συμβάντων σε χρονικό (ή χωρικό) διάστημα t ακολουθεί την κατανομή Poisson αν **(α)** ο ρυθμός λ , έστω των συμβάντων είναι χρονικά σταθερός και **(β)** οι αριθμοί των συμβάντων σε ξένα διαστήματα αποτελούν ανεξάρτητα ενδεχόμενα (Κοκολάκης και Σπηλιώτης, 1999). Η κατανομή Poisson είναι κατάλληλη για την ανάπτυξη μοντέλων που αφορούν φαινόμενα που εμφανίζονται σπάνια και των οποίων οι εμφανίσεις είναι ανεξάρτητες μεταξύ τους, δηλαδή η εμφάνιση του φαινομένου μια φορά δεν επηρεάζει την επόμενη. Ο αριθμός των παθόντων είναι μία μεταβλητή, η οποία παρουσιάζει όμοιες ιδιότητες με την μεταβλητή του αριθμού των ατυχημάτων και γενικά υποστηρίζεται ότι τα οδικά ατυχήματα ακολουθούν συνήθως κατανομή Poisson ή κανονική κατανομή.

3.3.3 Αρνητική Διωνυμική Κατανομή

Μία άλλη πολύ σημαντική κατανομή που χρησιμοποιείται στην οδική ασφάλεια είναι η **αρνητική διωνυμική κατανομή**. Η χρήση της κατανομής αυτής ενδείκνυται για περιπτώσεις όπου η διακύμανση των στοιχείων του δείγματος είναι μεγαλύτερη από τον μέσο όρο. Αυτό μπορεί να παρατηρηθεί σε φαινόμενα που εμφανίζουν περιοδικές μεταβολές (όπως για παράδειγμα αριθμός αφίξεων οχημάτων που αφορούν σε μικρά χρονικά διαστήματα (π.χ. 10 s) σε κάποιο σημείο μετά από φωτεινό σηματοδότη. Μια τυχαία μεταβλητή X θεωρείται ότι ακολουθεί την αρνητική διωνυμική κατανομή με παραμέτρους k, p (k : θετικός ακέραιος, $0 < p < 1$), και γράφεται $X \sim NB(k, p)$, όταν έχει συνάρτηση μάζας πιθανότητας την:

$$P(x) = \binom{x+k-1}{x} p^k (1-p)^x$$

όπου $x = 0, 1, 2, 3, \dots$

3.4 Μαθηματικά Πρότυπα

3.4.1 Γραμμική Παλινδρόμηση

Ο κλάδος της στατιστικής, ο οποίος εξετάζει τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών ώστε να είναι δυνατή η πρόβλεψη της μίας από τις υπόλοιπες, ονομάζεται **ανάλυση παλινδρόμησης** (regression analysis). Με τον όρο **εξαρτημένη μεταβλητή** εννοείται η μεταβλητή της οποίας η τιμή πρόκειται να προβλεφθεί, ενώ ο όρος **ανεξάρτητη μεταβλητή** αναφέρεται σε εκείνη την μεταβλητή, η οποία χρησιμοποιείται για την πρόβλεψη της εξαρτημένης μεταβλητής. Η ανεξάρτητη μεταβλητή δεν θεωρείται τυχαία αλλά παίρνει καθορισμένες τιμές. Η εξαρτημένη μεταβλητή αντίθετα, θεωρείται τυχαία και “καθοδηγείται” από την ανεξάρτητη μεταβλητή. Προκειμένου να προσδιοριστεί αν μια ανεξάρτητη μεταβλητή ή ένας συνδυασμός ανεξάρτητων μεταβλητών προκαλεί τη μεταβολή της εξαρτημένης μεταβλητής, κρίνεται απαραίτητη η ανάπτυξη μαθηματικών μοντέλων. Η ανάπτυξη ενός μαθηματικού μοντέλου αποτελεί μια στατιστική διαδικασία που συμβάλλει στην ανάπτυξη εξισώσεων που περιγράφουν τη σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης. Επισημαίνεται ότι η επιλογή της μεθόδου ανάπτυξης ενός μοντέλου βασίζεται στο αν η εξαρτημένη μεταβλητή λαμβάνει συνεχείς ή διακριτές τιμές. Στην περίπτωση που η **εξαρτημένη μεταβλητή είναι συνεχές μέγεθος** και ακολουθεί κανονική κατανομή χρησιμοποιείται η μέθοδος της γραμμικής παλινδρόμησης, της οποίας η πιο απλή περίπτωση είναι η

3. Θεωρητικό Υπόβαθρο

απλή γραμμική παλινδρόμηση (Simple Linear Regression). Στην **απλή γραμμική παλινδρόμηση** υπάρχει μόνο μία ανεξάρτητη μεταβλητή X και μία εξαρτημένη μεταβλητή y , η οποία προσεγγίζεται ως μία γραμμική συνάρτηση του X . Η τιμή y_i της μεταβλητής Y , για κάθε τιμή της μεταβλητής X , δίνεται από την σχέση:

$$y_i = a + b * x_i + \varepsilon_i$$

Το πρόβλημα της παλινδρόμησης είναι η εύρεση των παραμέτρων a και b που εκφράζουν όσο το δυνατόν καλύτερα τη γραμμική εξάρτηση της εξαρτημένης μεταβλητής Y από την ανεξάρτητη μεταβλητή X . Κάθε ζεύγος τιμών (a , b) καθορίζει και μία διαφορετική γραμμική σχέση που εκφράζεται γεωμετρικά από ευθεία γραμμή και οι δύο παράμετροι ορίζονται ως εξής:

- Ο σταθερός όρος a είναι η τιμή του y για $x=0$
- Ο συντελεστής b του x είναι η κλίση (slope) της ευθείας ή αλλιώς ο **συντελεστής παλινδρόμησης** (regression coefficient). Εκφράζει τη μεταβολή της μεταβλητής Y όταν η μεταβλητή X μεταβληθεί κατά μία μονάδα.

Η τυχαία μεταβλητή ε_i λέγεται **σφάλμα παλινδρόμησης** (Regression Error) και ορίζεται ως η διαφορά της y_i από τη δεσμευμένη μέση τιμή $E(Y|X=x_i)$ όπου $E(Y|X=x_i) = a + b * x_i$. Για την ανάλυση της γραμμικής παλινδρόμησης γίνονται οι παρακάτω υποθέσεις:

- Η μεταβλητή X είναι ελεγχόμενη για το πρόβλημα υπό μελέτη, δηλαδή είναι γνωστές οι τιμές της χωρίς καμιά αμφιβολία.
- Η εξάρτηση της Y από τη X είναι γραμμική.
- Το σφάλμα παλινδρόμησης έχει μέση τιμή μηδέν για κάθε τιμή της X και η διασπορά του είναι σταθερή και δεν εξαρτάται από τη X , δηλαδή $E(\varepsilon_i) = 0$ και $\text{Var}(\varepsilon_i) = \sigma^2$.

Οι παραπάνω υποθέσεις για γραμμική σχέση και σταθερή διασπορά αποτελούν χαρακτηριστικά πληθυσμών με κανονική κατανομή. Συνήθως, λοιπόν, σε προβλήματα γραμμικής παλινδρόμησης γίνεται η υπόθεση ότι η δεσμευμένη κατανομή της Y είναι κανονική.

Σε περίπτωση που η εξαρτημένη μεταβλητή Y εξαρτάται γραμμικά από περισσότερες από μια ανεξάρτητες μεταβλητές X ($x_1, x_2, x_3, \dots, x_n$) τότε γίνεται αναφορά στην πολλαπλή γραμμική παλινδρόμηση (Multiple Linear Regression). Η εξίσωση η οποία αποτυπώνει τη σχέση ανάμεσα στην εξαρτημένη και τις ανεξάρτητες μεταβλητές έχει τη γενικότερη μορφή:

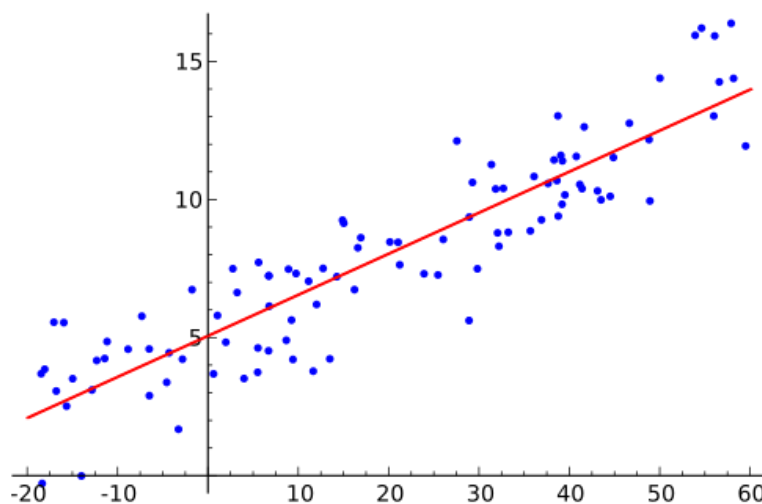
$$y_i = b_0 + b_1 * x_{1i} + b_2 * x_{2i} + \dots + b_n * x_{ni} + \varepsilon_i$$

Οι **υποθέσεις της πολλαπλής γραμμικής παλινδρόμησης** είναι ίδιες με εκείνες της απλής γραμμικής παλινδρόμησης, δηλαδή υποθέτει κανείς ότι τα σφάλματα της παλινδρόμησης (όπως και η τυχαία μεταβλητή Y για κάθε τιμή της X) ακολουθούν κανονική κατανομή με σταθερή διασπορά. Γενικά το πρόβλημα και η εκτίμηση της πολλαπλής γραμμικής παλινδρόμησης δεν διαφέρει ουσιαστικά από εκείνο της απλής γραμμικής παλινδρόμησης. Ένα καινούριο στοιχείο στην πολλαπλή γραμμική παλινδρόμηση είναι ότι πριν προχωρήσει κανείς στην εκτίμηση των παραμέτρων

πρέπει να ελέγξει αν πράγματι πρέπει να συμπεριληφθούν όλες οι ανεξάρτητες μεταβλητές στο μοντέλο. Εκείνο που απαιτείται να εξασφαλιστεί είναι η μηδενική συσχέτιση των ανεξάρτητων μεταβλητών, δηλαδή θα πρέπει να ισχύει: $\rho(X_i, X_j) \neq 0$.

3.4.1.1 Μέθοδος Ελαχίστων Τετραγώνων

Η εκτίμηση των παραμέτρων του μοντέλου της πολλαπλής γραμμικής παλινδρόμησης γίνεται με τη **μέθοδο ελαχίστων τετραγώνων** (Least Squares Method). Σύμφωνα με τη μέθοδο ελαχίστων τετραγώνων, ο προσδιορισμός των συντελεστών b_i δίνει μια προσεγγιστική ευθεία που συνδέει τις τιμές της μεταβλητής Y , δοθέντων των τιμών της μεταβλητής X . Η ευθεία που προκύπτει ονομάζεται ευθεία παλινδρόμησης της Y πάνω στην X . Σκοπός είναι το άθροισμα των τετραγώνων των κατακόρυφων αποστάσεων των σημείων (X, Y) από την ευθεία να είναι ελάχιστο. Παρακάτω δίνεται ένα ενδεικτικό διάγραμμα της ευθείας ελαχίστων τετραγώνων.



Εικόνα 3.2.: Παράδειγμα ευθείας ελαχίστων τετραγώνων
Πηγή: Wikipedia

3.4.2 Λογιστική Ανάλυση Παλινδρόμησης

Η **λογιστική παλινδρόμηση (Logistic Regression)** αποτελεί μία στατιστική μέθοδο κατάλληλη για τις περιπτώσεις που επιδιώκεται πρόβλεψη της επιρροής από την παρουσία ή την απουσία κάποιων χαρακτηριστικών στην επιλογή αναφορικά με κάποιο συγκεκριμένο γεγονός. Οδηγεί στην ανάπτυξη μαθηματικού προτύπου που δίνει τη συνάρτηση χρησιμότητας του συγκεκριμένου γεγονότος, συναρτήσει κάποιων παραγόντων που το επηρεάζουν. Από τη συνάρτηση χρησιμότητας (Utility function) υπολογίζεται εύκολα κατόπιν κατάλληλου μετασχηματισμού, η πιθανότητα που υπάρχει το γεγονός αυτό να πραγματοποιηθεί. Το πρότυπο που δίνει τη συνάρτηση χρησιμότητας είναι γραμμικό συναρτήσει των παραμέτρων που επηρεάζουν τη συγκεκριμένη επιλογή. Αντίθετα η σχέση που συνδέει την πιθανότητα με τη συνάρτηση χρησιμότητας είναι μη γραμμική. Η μορφή της συνάρτησης χρησιμότητας είναι η εξής:

$$U_i = a_0 + a_1x_1 + a_2x_2 + \dots + a_vx_v$$

3. Θεωρητικό Υπόβαθρο

Όπου :

U_i = συνάρτηση χρησιμότητας (Utility function) του γεγονότος i

$x_1, \dots, x_v$ οι μεταβλητές του προβλήματος

a_0 = ο σταθερός όρος ο οποίος δείχνει την επίδραση όλων εκείνων των παραγόντων που επηρεάζουν την επιλογή και δεν έχουν συμπεριληφθεί ως μεταβλητές στο μαθηματικό πρότυπο.

$a_1, \dots, a_v$ οι συντελεστές των μεταβλητών

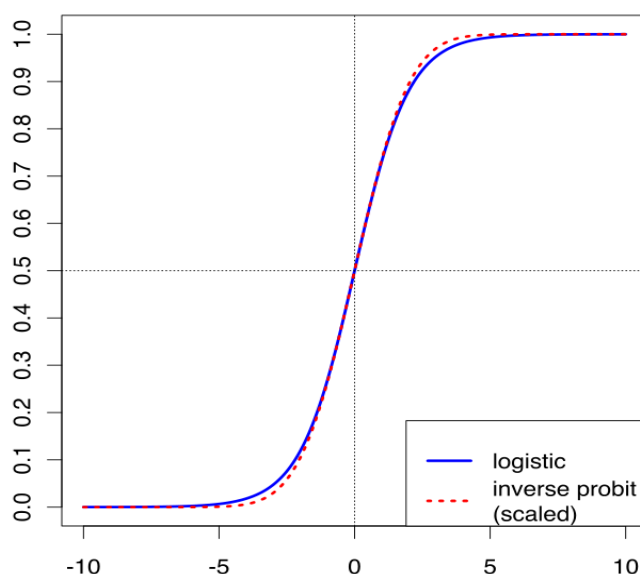
Η πιθανότητα P_i να συμβεί το συγκεκριμένο γεγονός i (εδώ η χρήση κινητού τηλεφώνου) δίνεται από τη σχέση:

$$P_i = \frac{e^{U_i}}{1 + e^{U_i}}$$

ενώ η πιθανότητα να μη συμβεί το συγκεκριμένο γεγονός, όπως είναι ευνόητο είναι $1 - P_i$ και στην περίπτωση της παρούσας Εργασίας ισούται με την πιθανότητα ο οδηγός να μην κάνει χρήση του κινητού τηλεφώνου. Μια διαφορετική σχέση της συνάρτησης χρησιμότητας προκύπτει από την τελευταία σχέση μετά τη εκτέλεση κάποιων αριθμητικών πράξεων :

$$(1 + e^U) * P = e^U \leftrightarrow (1 + e^{-U}) * P = 1 \leftrightarrow e^{-U} = \frac{1 - P}{P} \leftrightarrow e^U = \frac{P}{1 - P} \leftrightarrow U = \ln \frac{P}{1 - P}$$

Επομένως, η εξαρτημένη μεταβλητή U είναι ο νεπέριος λογάριθμος του λόγου $P/(1-P)$ όπου P η πιθανότητα να συμβεί κάποιο συγκεκριμένο γεγονός. Ο λόγος των πιθανοτήτων $P/(1-P)$ στη διεθνή βιβλιογραφία αναφέρεται με τον όρο "odds". Μια σημαντική ιδιότητα του λογιστικού προτύπου είναι ότι μετατρέπει το πρόβλημα της πρόβλεψης πιθανοτήτων σε πρόβλημα πρόβλεψης του λόγου πιθανοτήτων (odds).



Εικόνα 5.3.: Σιγμοειδής καμπύλη λογιστικής παλινδρόμησης
Πηγή: Wikipedia

Αν το P ισούται με 0 ή με 1 τότε οι λόγοι πιθανοτήτων (odds) θα είναι ίσοι με 0 ή θα παρουσιάζουν απροσδιοριστία, με αποτέλεσμα να μην ορίζεται ο νεπέριος λογάριθμός τους, άρα ούτε και η συνάρτηση χρησιμότητας. Για το λόγο αυτό η χρήση της μεθόδου των ελαχίστων τετραγώνων δεν είναι δυνατή. Έτσι ο υπολογισμός των συντελεστών του μαθηματικού προτύπου γίνεται με τη μέθοδο της μέγιστης πιθανοφάνειας (maximum likelihood). Η συγκεκριμένη μέθοδος αποδίδει με ακρίβεια τους παραπάνω συντελεστές, ενώ λειτουργεί με την ίδια ευκολία ακόμη και στην περίπτωση που το δείγμα είναι μεγάλο. Για το λόγο αυτό η συγκεκριμένη μέθοδος αποτελεί ένα εύχρηστο εργαλείο για την εκπόνηση τέτοιου είδους ερευνών.

Με απλό μετασχηματισμό των παραπάνω σχέσεων καταλήγουμε στην εξής νέα εξίσωση του λόγου πιθανοφάνειας:

$$\frac{P_i}{1 - P_i} = e^{a_0} * e^{a_1 x_1} * e^{a_2 x_2} * \dots * e^{a_n x_n}$$

από όπου διαπιστώνουμε ότι για αύξηση κατά μία μονάδα μιας ανεξάρτητης μεταβλητής (δηλαδή από x_i σε x_i+1) διατηρώντας σταθερές όλες τις υπόλοιπες, η πιθανότητα αυξάνεται κατά έναν συντελεστή e^{a_i} .

3.4.3 Μοντέλα πολλαπλών επιπέδων ή μεικτών επιδράσεων ή Μεικτά μοντέλα

Σε περίπτωση στην οποία διατίθενται **επαναλαμβανόμενες μετρήσεις** που αναφέρονται στο ίδιο άτομο, η στατιστική ανάλυση ενδέχεται να παρουσιάσει αρκετά προβλήματα οδηγώντας σε στατιστικά μη έγκυρα ευρήματα. Τα προβλήματα πηγάζουν από το γεγονός ότι οι συγκεκριμένες μετρήσεις είναι συσχετισμένες παραβαίνοντας μία από τις συνηθέστερες προϋποθέσεις διάφορων στατιστικών μοντέλων, την υπόθεση ανεξαρτησίας των παρατηρήσεων του δείγματος. Οι ανεξάρτητες μεταβλητές θεωρείται ότι επιδρούν με **σταθερό (fixed) και τυχαίο (random) τρόπο**. Η πρώτη κατηγορία (fixed) περιλαμβάνει παράγοντες οι οποίοι περιέχουν σημαντική πληροφορία, χαρακτηρίζονται ως *informative* και υπολογίζονται από τα δεδομένα του δείγματος, επηρεάζοντας τη μέση τιμή της εξαρτημένης μεταβλητής. Η δεύτερη κατηγορία περιλαμβάνει παράγοντες οι οποίοι χαρακτηρίζονται ως μη-πληροφοριακοί (*non-informative*) και επηρεάζουν τη μεταβλητότητα του πειράματος. Αποτελούν επέκταση των μοντέλων παλινδρόμησης και μπορούν να συμβολιστούν με μορφή πίνακα ως εξής:

$$y = X * b + Z * u + \varepsilon$$

όπου y ένα γνωστό διάνυσμα παρατηρήσεων με μέση τιμή $E(y) = X*b$,

b ένα άγνωστο διάνυσμα που επιδρά με σταθερό (fixed) τρόπο,

u ένα άγνωστο διάνυσμα που επιδρά με τυχαίο (random) τρόπο με μέση τιμή $E(u)=0$ και πίνακα μεταβλητότητας- συνδιακύμανσης $Var(u)= G$,

ε ένα άγνωστο διάνυσμα τυχαίου σφάλματος με μέση τιμή $E(\varepsilon)= 0$ και συνδιακύμανση $Var(\varepsilon)= R$,

X και Z γνωστοί πίνακες που συνδέουν τις παρατηρήσεις y με το b και το u αντίστοιχα

3.5 Στατιστική Αξιολόγηση και Κριτήρια Αποδοχής Μοντέλου

Τα κριτήρια που χρησιμοποιούνται για την αξιολόγηση ενός μοντέλου μετά τη διαμόρφωσή του είναι τα πρόσημα και οι τιμές των συντελεστών β_i της εξίσωσης, η στατιστική σημαντικότητα, η ποιότητα του μοντέλου και το σφάλμα της εξίσωσης.

- Όσον αφορά στους **συντελεστές της εξίσωσης**, θα πρέπει να υπάρχει δυνατότητα **λογικής ερμηνείας των πρόσημων τους**. Το θετικό πρόσημο του συντελεστή δηλώνει αύξηση της εξαρτημένης μεταβλητής με την αύξηση της ανεξάρτητης. Αντίθετα, αρνητικό πρόσημο συνεπάγεται μείωση της εξαρτημένης μεταβλητής με την αύξηση της ανεξάρτητης. Η τιμή του συντελεστή θα πρέπει και αυτή να ερμηνεύεται λογικά δεδομένου ότι, αύξηση της ανεξάρτητης μεταβλητής (x_i) κατά μία μονάδα επιφέρει αύξηση της εξαρτημένης κατά β_i μονάδες. Στην περίπτωση που η αύξηση αυτή εκφράζεται σε ποσοστά τότε αναφερόμαστε στην ελαστικότητα (elasticity).

- Πληροφορίες σχετικά με το αν **πρέπει μια μεταβλητή να συμπεριληφθεί στο πρότυπο** ή όχι δίνει ένας άλλος στατιστικός δείκτης που υπολογίζεται κατά τη διαδικασία της επεξεργασίας των δεδομένων, ο **t-ratio**.

$$t - ratio = \frac{\beta_i}{s.e.}$$

όπου s.e. το τυπικό σφάλμα των σταθερών παραμέτρων (standard error).

Ο t-ratio αναφέρεται σε καθεμιά από τις μεταβλητές ξεχωριστά. Ο δείκτης αυτός είναι στην ουσία το αποτέλεσμα της διαίρεσης της εκτιμώμενης για το συντελεστή τιμής δια της τυπικής απόκλισής της. Η τυπική απόκλιση είναι ένα μέγεθος που παρουσιάζει τη συνέπεια με την οποία έχει υπολογιστεί η τιμή του συγκεκριμένου συντελεστή. Αυτό σημαίνει ότι η πραγματική τιμή του συντελεστή για διάστημα εμπιστοσύνης (επίπεδο σημαντικότητας) 95% βρίσκεται στο διάστημα με κέντρο την υπολογιζόμενη τιμή του συντελεστή και άκρα την τιμή αυτή $\pm$ την τυπική απόκλιση.

Ο δείκτης t-ratio στην ουσία δείχνει αν η πραγματική τιμή του συγκεκριμένου συντελεστή διαφέρει σημαντικά από το 0 ή όχι. Δείχνει δηλαδή αν η επίδρασή της αντίστοιχης μεταβλητής στο τελικό αποτέλεσμα είναι σημαντική ή όχι. Αν η επιρροή αυτή είναι σημαντική τότε η συγκεκριμένη μεταβλητή πρέπει να συμπεριληφθεί στην ανάπτυξη του μαθηματικού προτύπου. Σε αντίθετη περίπτωση πρέπει να αποκλειστεί. Οι τιμές που μπορεί να πάρει κυμαίνονται από $-\infty$ έως $+\infty$.

Όσο μεγαλύτερη είναι η απόλυτη τιμή του τόσο μεγαλύτερη είναι η επιρροή της συγκεκριμένης μεταβλητής στο τελικό αποτέλεσμα. Ανάλογα με το επίπεδο σημαντικότητας στο οποίο ενδιαφέρει να βρίσκονται τα αποτελέσματα της έρευνας, υπάρχουν πίνακες που δίνουν την τιμή του t-ratio πάνω από την οποία η συγκεκριμένη μεταβλητή πρέπει να συμπεριληφθεί στο πρότυπο. Έτσι για διάστημα εμπιστοσύνης 95%, μια μεταβλητή μπορεί να παραμείνει στο πρότυπο αν η απόλυτη τιμή του t-ratio του συντελεστή της είναι μεγαλύτερη από 1,96.

Πρέπει να σημειωθεί ότι από τη στιγμή που υπάρχει σταθερός όρος, η τελευταία τιμή των κατηγορικών μεταβλητών θεωρείται περιττή και χρησιμοποιείται ως επίπεδο

αναφοράς για τη σύγκριση αυτής με τις άλλες τιμές των κατηγορικών μεταβλητών. Με το t-test λοιπόν καθορίζεται εάν υπάρχει στατιστικά σημαντική διαφορά μεταξύ τους. Το ίδιο συμβαίνει και με τις αλληλεπιδράσεις των κατηγορικών μεταβλητών με τις υπόλοιπες ανεξάρτητες μεταβλητές. Επισημαίνεται ότι οι διακριτές μεταβλητές με κατηγορίες k έχουν k-1 βαθμούς ελευθερίας, ενώ οι συνεχείς έχουν πάντοτε ένα βαθμό ελευθερίας.

Αντί για το t-ratio μπορεί πιο απλά να υπολογιστεί η τιμή του p-value και να συγκριθεί με το αντίστοιχο επίπεδο σημαντικότητας. Συγκεκριμένα, μία συνήθης πρακτική στον έλεγχο στατιστικών υποθέσεων, είναι ο υπολογισμός της τιμής της **πιθανότητας p (probability-value ή p-value)**. Η πιθανότητα p είναι το μικρότερο επίπεδο σημαντικότητας α που οδηγεί στην απόρριψη της **μηδενικής υπόθεσης H_0** έναντι της εναλλακτικής H_1 . Είναι μία σημαντική τιμή, διότι ποσοτικοποιεί τη στατιστική απόδειξη που υποστηρίζει την εναλλακτική υπόθεση. Γενικά, όσο πιο μικρή είναι η τιμή της πιθανότητας p, τόσο περισσότερες είναι οι αποδείξεις για την απόρριψη της μηδενικής υπόθεσης H_0 έναντι της εναλλακτικής H_1 . **Εάν η τιμή p είναι μικρότερη ή ίση του επιπέδου σημαντικότητας α , τότε η μηδενική υπόθεση H_0 απορρίπτεται.**

- Η ποιότητα του μοντέλου εξετάζεται με τον συντελεστή προσαρμογής R^2 . Ο συντελεστής R^2 χρησιμοποιείται ως κριτήριο καλής προσαρμογής των δεδομένων στο γραμμικό μοντέλο και ορίζεται από τη σχέση:

$$R^2 = \frac{SSR}{SST}$$

όπου:

$$SSR = \sum_{i=1}^n (y_i - \hat{y})^2 = \beta^2 * \sum_{i=1}^n (x_i - \bar{x})^2$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

και

$\hat{y}$ η προβλεπόμενη τιμή της εξαρτημένης μεταβλητής από τις ανεξάρτητες.

Τα αρχικά SSR και SST έχουν προέλθει από τις φράσεις υπόλοιπο άθροισμα τετραγώνων (Residual Sum of Squares) και συνολικό άθροισμα τετραγώνων (Total Sum of Squares), αντίστοιχα.

Ο συντελεστής αυτός εκφράζει το ποσοστό της μεταβλητότητας της μεταβλητής Y που εξηγείται από την μεταβλητή X. Λαμβάνει τιμές από 0 έως 1. Όσο πιο κοντά βρίσκεται η τιμή του R^2 στην μονάδα, τόσο πιο ισχυρή γίνεται η γραμμική σχέση εξάρτησης των μεταβλητών Y και X. Ο συντελεστής R^2 έχει συγκριτική αξία, κάτι το οποίο σημαίνει ότι δεν υπάρχει συγκεκριμένη τιμή του R^2 που είναι αποδεκτή ή απορριπτέα, αλλά μεταξύ δύο ή περισσότερων μοντέλων επιλέγεται ως καταλληλότερο εκείνο με τη μεγαλύτερη

τιμή του συντελεστή R^2 . Ο συντελεστής R^2 μπορεί να χρησιμοποιηθεί ως μέτρο ισχυρότητας της γραμμικής σχέσης ανεξάρτητα από το αν το X παίρνει καθορισμένες τιμές ή αν είναι τυχαία μεταβλητή.

- Για τη στατιστική εμπιστοσύνη του μοντέλου της λογιστικής παλινδρόμησης χρησιμοποιείται η **μέθοδος της μεγιστοποίησης της πιθανοφάνειας**. Για να επιτευχθεί υψηλή πιθανοφάνεια πρέπει ο λογάριθμος των συναρτήσεων πιθανοφάνειας $L = -2$ Restricted Log Likelihood να είναι όσο το δυνατόν μικρότερος και γενικά προτιμώνται τα μοντέλα με τον μικρότερο λογάριθμο συνάρτησης πιθανοφάνειας L . Μοντέλα που περιέχουν πολλές μεταβλητές είναι περισσότερο σύνθετα και απαιτείται ένας κανόνας που να αποφασίζει εάν η μείωση του L αξίζει την αυξημένη πολυπλοκότητα. Για τον λόγο αυτό, χρησιμοποιείται το **κριτήριο λόγου πιθανοφάνειας (Likelihood Ratio Test LRT)**. Το μέγεθος $L(0)$ αποτελεί έναν απλό υπολογισμό της πιθανοφάνειας ενός προτύπου στο οποίο για κάθε παρατήρηση (στοιχείο της έρευνας πεδίου), όλες οι εναλλακτικές επιλογές έχουν την ίδια πιθανότητα να επιλεγούν. Το μέγεθος $L(b)$ αποτελεί μια προσέγγιση της πιθανοφάνειας που θα προέκυπτε από ένα πρότυπο στο οποίο όλες οι εναλλακτικές επιλογές εκτός από μία έχουν έναν εναλλακτικό καθορισμένο συντελεστή (alternative specific constant). Οι τιμές των παραπάνω μεγεθών εξεταζόμενες μεμονωμένα δεν δίνουν καμία ουσιαστική πληροφορία. Απεναντίας όμως συγκρινόμενες μεταξύ του και με τις αντίστοιχες τιμές άλλων δοκιμών, δίνουν μια γενική εικόνα για την ποιότητα του συγκεκριμένου προτύπου.

Για παράδειγμα οι έλεγχοι πιθανοφάνειας χ^2 (likelihood ratio χ^2) είναι ένας πολύ συνηθισμένος τρόπος ελέγχου προτύπων που έχουν αναπτυχθεί με βάση τα ίδια δεδομένα. Όταν το ένα πρότυπο είναι γενίκευση του άλλου, τα δύο πρότυπα μπορούν να παρουσιάσουν συμπεριφορά (nested hypotheses) για την οποία το διπλάσιο της διαφοράς του λογαρίθμου της πιθανοφάνειας μπορεί να ελεγχθεί σαν μεταβλητή χ^2 με κατάλληλο αριθμό βαθμών ελευθερίας. Η τελική πιθανοφάνεια των δύο αυτών μοντέλων είναι το κλειδί στα δεδομένα για τον παραπάνω έλεγχο.

Σύμφωνα με το **κριτήριο του λόγου πιθανοφάνειας**, εάν η διαφορά

$$LRT = -2 * (L(b) - L(0)) > \chi^2_{b,0.05}$$

όπου $L(b)$ = λογάριθμος πιθανοφάνειας (μοντέλου με b μεταβλητές) και $L(0)$ = λογάριθμος πιθανοφάνειας (μοντέλου χωρίς τις b μεταβλητές), είναι μεγαλύτερη από την τιμή του κριτηρίου χ^2 για p βαθμούς ελευθερίας σε επίπεδο σημαντικότητας 5%, τότε το μοντέλο είναι στατιστικά προτιμότερο από το μοντέλο χωρίς τις μεταβλητές και γίνεται αποδεκτό. Επισημαίνεται ότι οι διακριτές μεταβλητές με κατηγορίες k έχουν $k-1$ βαθμούς ελευθερίας, ενώ οι συνεχείς έχουν πάντοτε ένα βαθμό ελευθερίας.

- Η **ελαστικότητα** αντικατοπτρίζει την ευαισθησία μιας εξαρτημένης μεταβλητής Y στη μεταβολή μιας ή περισσότερων ανεξάρτητων μεταβλητών. Είναι πολλές φορές ορθότερο να εκφραστεί η ευαισθησία ως ποσοστιαία μεταβολή της εξαρτημένης μεταβλητής που προκαλεί η 1% μεταβολή της ανεξάρτητης. Η ελαστικότητα, για γραμμικά μοντέλα, δίνεται από την παρακάτω σχέση:

$$e_i = \left(\frac{\Delta Y_i}{\Delta X_i} \right) * \left(\frac{X_i}{Y_i} \right) = \beta_i * \left(\frac{X_i}{Y_i} \right)$$

- Για να πραγματοποιηθεί με επιτυχία η μέθοδος της γραμμικής και της λογιστικής παλινδρόμησης (logistic regression) **οι ανεξάρτητες μεταβλητές** του προβλήματος πρέπει να είναι μεταξύ τους **ασυσχέτιστες**. Αν δεν είναι ασυσχέτιστες, στη συντριπτική πλειοψηφία των περιπτώσεων, η διαδικασία οδηγείται σε αποτυχία με αποτέλεσμα να μην είναι δυνατή η ανάπτυξη μαθηματικού προτύπου.

4. Συλλογή και Επεξεργασία Στοιχείων

4.1 Εισαγωγή

Μετά τη βιβλιογραφική ανασκόπηση ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας, αναπτύχθηκε το θεωρητικό υπόβαθρο που οδήγησε στην επιλογή μιας κατάλληλης μεθόδου ανάλυσης. Αφού επιλέχθηκαν η γραμμική και η δυαδική λογιστική παλινδρόμηση ως μέθοδοι στατιστικής ανάλυσης, επόμενο βήμα ήταν η συλλογή δεδομένων από συσκευές smartphone μέσα από το πείραμα σε πραγματικές συνθήκες οδήγησης που πραγματοποιήθηκε. Η στατιστική επεξεργασία των στοιχείων αυτών θα οδηγήσει στην επίτευξη του στόχου της Διπλωματικής Εργασίας, που είναι η διερεύνηση της **επίδρασης της χρήσης κινητού τηλεφώνου στην οδηγική συμπεριφορά**.

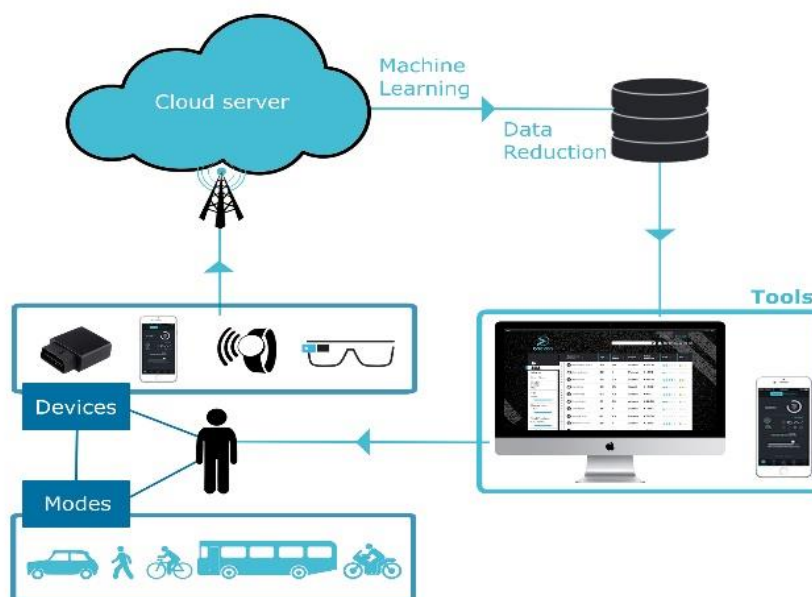
Το κεφάλαιο αυτό, που αφορά **στη συλλογή και επεξεργασία στοιχείων**, περιλαμβάνει δύο υπό-κεφάλαια. Στο εδάφιο που αναφέρεται στη συλλογή στοιχείων, περιγράφεται το πείραμα που πραγματοποιήθηκε για τα στοιχεία που συλλέχθηκαν. Στο δεύτερο εδάφιο το οποίο αφορά στην επεξεργασία των στοιχείων παρατίθενται συγκεντρωτικοί πίνακες στους οποίους περιγράφονται τα χαρακτηριστικά του δείγματος, ακολουθούμενοι από τον απαραίτητο σχολιασμό. Στο τέλος, αναλύεται ο τρόπος εισαγωγής των δεδομένων στο ειδικό λογισμικό στατιστικής ανάλυσης.

4.2 Συλλογή Στοιχείων

Σύμφωνα με τη βιβλιογραφική ανασκόπηση προέκυψε πως έχουν πραγματοποιηθεί αρκετές έρευνες σχετικά με την απόσπαση προσοχής λόγω της χρήσης κινητού τηλεφώνου κατά την οδήγηση. Ωστόσο, η μεθοδολογία συλλογής των στοιχείων περιλάμβανε συνήθως μια συσκευή που βρισκόταν εντός του οχήματος και ήταν συνδεδεμένη με το όχημα. Παραδείγματα τέτοιων συσκευών είναι το **Driving Data Recorder DDR** (Ohta, Tohru, and Shouji Nakajima, 1994), μια συσκευή μικρών διαστάσεων που συνδεόταν με τη μηχανή του αυτοκινήτου ή το **On Board Diagnostics (OBD)** (Zaldivar, Jorge, et al., 2011) το οποίο έχει ως κύριο στόχο την ανίχνευση μηχανικών προβλημάτων στο όχημα, αλλά χρησιμοποιήθηκε για την καταγραφή δεδομένων. Στην παρούσα διπλωματική εργασία η συλλογή των στοιχείων γίνεται μέσω αισθητήρων από το κινητό τηλέφωνο-smartphone.

Πιο αναλυτικά, εφαρμόστηκε ένα καινοτόμο σύστημα συλλογής δεδομένων, μέσα από την εξατομικευμένη καταγραφή της οδηγικής συμπεριφοράς σε πραγματικό χρόνο χρησιμοποιώντας αισθητήρες συσκευών smartphone. 100 οδηγοί συμμετείχαν στο πείραμα (smartphone naturalistic driving experiment) σε χρονικό διάστημα 4 μηνών και δημιουργήθηκε μια μεγάλη βάση δεδομένων με χιλιάδες (18.850) διαδρομές. Οι μετρήσεις πραγματοποιήθηκαν τόσο σε **αστικό** όσο και σε **υπεραστικό περιβάλλον** και **αυτοκινητόδρομο**. Στη συνέχεια, μια ειδική πλατφόρμα που σχεδιάστηκε από την εταιρεία στην οποία ανήκει η εφαρμογή smartphone (OSeven) εξασφάλισε την ομαλή μετάβαση από τη συλλογή στοιχείων στη διαδικασία ανάλυσης των δεδομένων. Τα βήματα που ακολουθούνται κάθε φορά που καταγράφεται μια διαδρομή-ταξίδι απεικονίζονται στην παρακάτω εικόνα 4.1 και περιγράφονται αναλυτικά στην επόμενη παράγραφο.

4. Συλλογή και Επεξεργασία Στοιχείων



Εικόνα 4.1.: Διάγραμμα διαχείρισης δεδομένων

Πηγή: OSeven

4.2.1 Μεθοδολογία καταγραφής, συλλογής, επεξεργασίας και οπτικοποίησης δεδομένων

Η **OSeven** έχει αναπτύξει ένα ολοκληρωμένο σύστημα για την καταγραφή, την συλλογή, την αποθήκευση, την αξιολόγηση και την παρουσίαση των δεδομένων της οδηγικής συμπεριφοράς χρησιμοποιώντας εφαρμογές **έξυπνων κινητών τηλεφώνων (smartphone)** και προηγμένους αλγόριθμους **Machine Learning**. Αυτή η καινοτομία που εφαρμόζεται στον τρόπο συλλογής και ανάλυσης δεδομένων σε μεγάλη κλίμακα, παρουσιάζει νέες προκλήσεις συγκεντρώνοντας μεγάλο όγκο δεδομένων για ανάλυση κατά τη διάρκεια αυτής της έρευνας. Το σύστημα που αναπτύχθηκε ενσωματώνει μια διαδικασία συλλογής και διαβίβασης δεδομένων από τα έξυπνα κινητά τηλέφωνα (Smartphones), των οποίων τα κύρια χαρακτηριστικά περιγράφονται στις επόμενες παραγράφους.

1. Σύστημα καταγραφής δεδομένων

Η καταγραφή των δεδομένων ξεκινάει αυτόματα στις εφαρμογές του κινητού τηλεφώνου όταν αναγνωρίζεται μια κατάσταση οδήγησης και **σταματάει αυτόματα** όταν αναγνωρίζεται κατάσταση μη οδήγησης. Η εγγραφή κατά τη διάρκεια της οδήγησης μένει σε αναμονή για πέντε λεπτά μετά την διακοπή της κατάστασης της οδήγησης, για να ληφθεί υπόψη το γεγονός ότι ο οδηγός συνεχίζει το ταξίδι του με στάση λίγων λεπτών. Όλες οι επιπλέον πληροφορίες που συλλέγονται μετά το πέρας της διαδρομής απορρίπτονται χρησιμοποιώντας τις τεχνικές μηχανικής μάθησης που περιγράφονται παρακάτω.

Τα καταγεγραμμένα δεδομένα προέρχονται από τους διάφορους **αισθητήρες των έξυπνων κινητών τηλεφώνων (smartphone)** και **αλγόριθμους συγχώνευσης (fusion algorithms) δεδομένων** που παρέχονται από το Android (Google) και το iOS (Apple). Η εφαρμογή κινητού τηλεφώνου (mobileApp) που έχει αναπτυχθεί καταγράφει

τη συμπεριφορά του χρήστη χρησιμοποιώντας τους αισθητήρες της συσκευής και χρησιμοποιείται μια ποικιλία APIs για να διαβαστούν τα δεδομένα των αισθητήρων που έχουν καταγραφεί και να αποθηκευτούν προσωρινά στη βάση δεδομένων του έξυπνου κινητού τηλεφώνου (Smartphone) πριν μεταφερθούν στην κεντρική βάση δεδομένων. Μετά τη μετάδοση των δεδομένων στην κεντρική βάση δεδομένων του συστήματος, όλες οι πληροφορίες διαγράφονται από το κινητό τηλέφωνο. Τα πρωτογενή δεδομένα είτε καταγράφονται από τους αισθητήρες των έξυπνων κινητών τηλεφώνων (Smartphone) είτε παρέχονται από iOS και Android με τη μορφή δεδομένων συγχώνευσης.

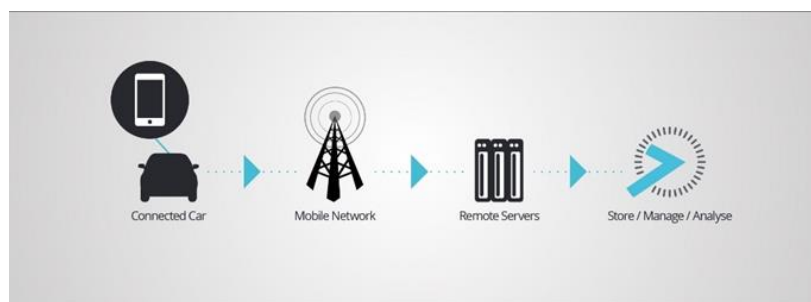
Ενδεικτικά, το κινητό τηλέφωνο ενσωματώνει αισθητήρες τεχνολογίας:

- Επιταχυνσιόμετρο (Accelerometer)¹
- Γυροσκόπιου (Gyroscope)¹
- Μαγνητόμετρο (Magnetometer)
- GPS (ταχύτητα, πορεία, γεωγραφικό μήκος, γεωγραφικό πλάτος)

Τα δεδομένα που παρέχονται από iOS και Android:

- Yaw, Pitch, Roll
- Γραμμική επιτάχυνση¹
- Βαρύτητα¹

Η συχνότητα της καταγραφής των δεδομένων ποικίλλει ανάλογα με τον τύπο του αισθητήρα, με μέγιστη τιμή 1Hz. Είναι αξιοσημείωτο να τονιστεί ότι με αυτόν τον τρόπο συλλέγεται ένα μεγάλο πλήθος δεδομένων των χαρακτηριστικών οδήγησης και αυτό πραγματοποιείται μόνο με την χρήση των εφαρμογών του κινητού τηλεφώνου. Το βασικό πλαίσιο λειτουργίας της ροής δεδομένων φαίνεται στο σχήμα που ακολουθεί.



Εικόνα 4.2.: Σύστημα ροής δεδομένων OSeven
Πηγή: OSeven

2. Μετάδοση δεδομένων

Μετά την ολοκλήρωση της διαδρομής, η εφαρμογή μεταδίδει όλα τα δεδομένα που έχουν συλλεχθεί, στην **κεντρική βάση δεδομένων του OSeven backend office**

¹ x, y, z άξονες

4. Συλλογή και Επεξεργασία Στοιχείων

μέσω ενός κατάλληλου δίαυλου επικοινωνίας, όπως είναι ένα δίκτυο **Wi-Fi** ή ένα **κυψελοειδές δίκτυο** (κατόπιν επιλογής του χρήστη) όπως είναι το δίκτυο 3G / 4G με βάση τις ρυθμίσεις του χρήστη.

Για να επιτευχθεί η διάδραση μεταξύ αυτών των πλευρών, δημιουργήθηκε API το οποίο χρησιμοποιείται για την εισαγωγή των δεδομένων από μια ηλεκτρονική υπηρεσία σε μια εφαρμογή πελάτη. Το API χρησιμοποιείται για τη μεταφορά και τη λήψη δεδομένων μεταξύ συστημάτων, υποστηρίζοντας τη μεταξύ τους διαλειτουργικότητα. Κάνοντας τα δεδομένα προσβάσιμα μέσω του Παγκόσμιου Ιστού με τη χρήση API καθίσταται δυνατή η υποβολή δεδομένων στη βάση δεδομένων από συσκευές τρίτων καθώς επίσης και την πληροφορία άμεσα διαθέσιμη.

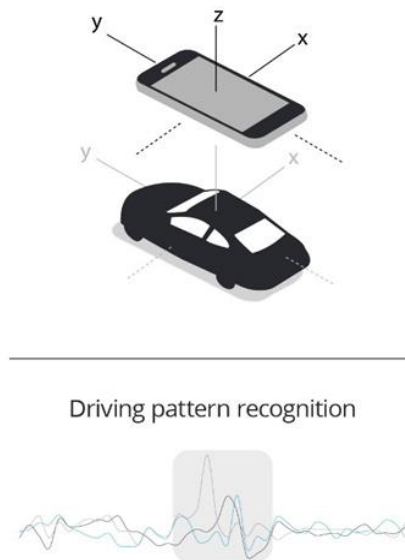
Ο συνολικός όγκος δεδομένων για έναν μέσο οδηγό υπολογίζεται σε 50Mb / μήνα.

3. Αποθήκευση δεδομένων, θέματα ασφαλείας και ιδιωτικότητας

Τα δεδομένα αποθηκεύονται στο σύστημα OSeven backend χρησιμοποιώντας προηγμένες **τεχνικές κρυπτογράφησης και ασφάλειας** δεδομένων, σύμφωνα με τους εθνικούς νόμους και τις οδηγίες της ΕΕ για την προστασία των προσωπικών δεδομένων. Οι εφαρμογές που χρησιμοποιούνται υποστηρίζουν τον έλεγχο ταυτότητας του χρήστη και την κρυπτογράφηση για να αποτρέπεται η πρόσβαση σε μη εξουσιοδοτημένα δεδομένα από τρίτους.

4. Επεξεργασία δεδομένων

Αφού τα δεδομένα αποθηκευτούν στον εξυπηρετητή για να πραγματοποιηθεί η κεντρική επεξεργασία και η μείωση της διάστασης των δεδομένων, μετατρέπονται τα στοιχεία που έχουν συλλεχθεί σε σημαντικές συμπεριφορές και σε συναφείς με την οδική ασφάλεια παραμέτρους (δηλαδή χειρισμό και επεξεργασία μαζικών δεδομένων). Αυτό επιτυγχάνεται χρησιμοποιώντας τις δύο μεγάλες μεθόδους επεξεργασίας δεδομένων που περιλαμβάνουν δύο οικογένειες τεχνικών, **τεχνικές εξόρυξης μεγάλων δεδομένων** και **αλγόριθμους Machine Learning (ML)**.



Εικόνα 4.3.: Yaw, Pitch, Roll
Πηγή: OSeven

Οι **μέθοδοι εκμάθησης μηχανών** (μέθοδοι φιλτραρίσματος, ομαδοποίησης και ταξινόμησης) χρησιμοποιούνται για τον καθαρισμό των δεδομένων από το θόρυβο, τα σφάλματα και την αναγνώριση επαναλαμβανόμενων μοτίβων μέσα στα δεδομένα. Στη συνέχεια, αυτά τα δεδομένα θα υποβληθούν σε επεξεργασία κατάλληλη για μεγάλο όγκο δεδομένων, προκειμένου να υπολογιστούν οι απαραίτητες παράμετροι και να ληφθούν δείκτες συμπεριφοράς που θα χρησιμοποιηθούν στην ανάλυση. Με άλλα λόγια, **τα δεδομένα** χωρικά και χρονικά διαχωρισμένα από το κινητό τηλέφωνο (Smartphone), **επεξεργάζονται προκειμένου να προκύψουν χρήσιμοι δείκτες οδικής ασφάλειας**. Οι μέθοδοι τεχνικής νοημοσύνης επιτρέπουν την **ανίχνευση της επιθετικής συμπεριφοράς** του οδηγού υπό τη μορφή απότομων συμβάντων, την απόσπαση προσοχής του οδηγού από τη χρήση του κινητού του τηλεφώνου, την αναγνώριση των τρόπων οδήγησης, την υπέρβαση του ορίου ταχύτητας, καθώς και ο προσδιορισμός των χρονικών και χωρικών χαρακτηριστικών όλων των παραπάνω.

Η διαδικασία των αλγορίθμων Machine Learning και η τεχνική εξόρυξης των Big Data δίνεται παρακάτω σε απλά βήματα:

- Φιλτράρισμα δεδομένων και ανίχνευση τιμών που αποκλίνουν
- Εξομάλυνση δεδομένων (όπου χρειάζεται)
- Περιοχές υπερβολικής ταχύτητας
- Συμβάντα απότομων επιταχύνσεων
- Συμβάντα απότομων φρεναρισμάτων
- Συμβάντα απότομων στροφών

4. Συλλογή και Επεξεργασία Στοιχείων

- Χρήση κινητού τηλεφώνου (ομιλία, σύνταξη μηνυμάτων, πλοήγηση στο διαδίκτυο)
- Οδήγηση εντός επικίνδυνων ωρών (απόσταση σε επικίνδυνες ώρες)
- Τρόπος μετακίνησης (αυτοκίνητο Ι.Χ., μέσα μαζικής μεταφοράς, ποδήλατο, μοτοποδήλατο)
- Καθορισμός οδηγού ή επιβάτη

Υπολογίζεται μια ποικιλία διαφορετικών μετά-δεδομένων έπειτα από τη διαδικασία Machine Learning (ML) τα οποία είναι χρήσιμα στον χρήστη ή στην αξιολόγηση της οδηγικής συμπεριφοράς, όπως για παράδειγμα:

Δείκτες έκθεσης κινδύνου

- Συνολική απόσταση (απόσταση σε μίλια)
- Διάρκεια οδήγησης
- Τύπος (τύποι) οδικού δικτύου που χρησιμοποιείται (η θέση δίνεται στο Παγκόσμιο Σύστημα Εντοπισμού θέσης (GPS) και ενσωματώνεται σε χορηγούς χαρτών, π.χ. Google, OSM)
- Περίοδος της ημέρας που πραγματοποιήθηκε η οδήγηση (ώρες αιχμής, επικίνδυνες ώρες)
- Καιρικές Συνθήκες
- Σκοπός του ταξιδιού

συνδυασμένα με άλλες πηγές δεδομένων (όρια ταχυτήτων και λεπτομερείς χάρτες ατυχημάτων).



Εικόνα 4.4.: Δείκτες κινδύνου οδήγησης
Πηγή: OSeven

Δείκτες οδηγικής συμπεριφοράς

- ❖ Υπέρβαση του ορίου ταχύτητας (διάρκεια υπέρβασης του ορίου, υπέρβαση του ορίου ταχύτητας κλπ.)

- ❖ Αριθμός και σοβαρότητα απότομων συμβάντων
- ❖ Απότομο φρενάρισμα (επιτάχυνση κατά μήκος)
- ❖ Απότομη επιτάχυνση (επιτάχυνση κατά μήκος)
- ❖ Απότομες στροφές (γωνιακή ταχύτητα, πλευρική επιτάχυνση, πορεία)
- ❖ Οδηγική επιθετικότητα (π.χ. φρενάρισμα, επιτάχυνση)
- ❖ Απόσπαση λόγω χρήσης κινητού τηλεφώνου

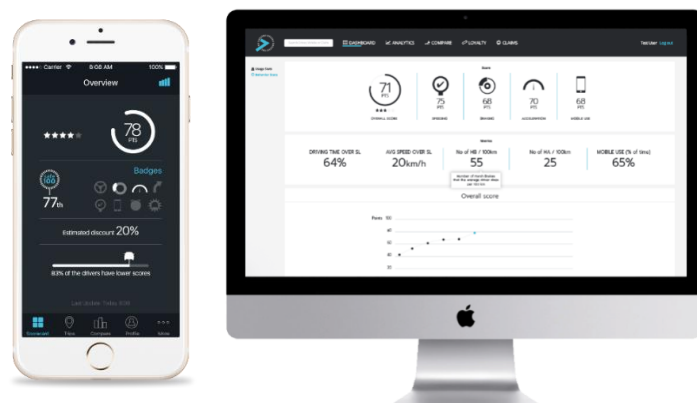
Αυτοί οι δείκτες μαζί με άλλα δεδομένα (π.χ. δεδομένα από χάρτες) αξιοποιούνται διαδοχικά για να εφαρμοστεί η στατιστική ανάλυση ατομικά σε κάθε οδηγό, σε όλους τους τύπους οδών (αστική οδός, αυτοκινητόδρομος κλπ.) και κάτω από διάφορες συνθήκες οδήγησης, επιτρέποντας τη δημιουργία μιας μεγάλης βάσης δεδομένων με οδηγικά χαρακτηριστικά.

Το τελικό βήμα της διαδικασίας επεξεργασίας των δεδομένων είναι η ανάπτυξη του μοντέλου οδηγικής συμπεριφοράς. Συγκεντρωτικά δεδομένα αναλύονται και το σύστημα αξιολόγησης βαθμονομείται με βάση το συνολικό δείγμα. Το μοντέλο οδηγικής συμπεριφοράς αναπτύσσει διάφορους δείκτες για την αξιολόγηση της οδήγησης και τελικώς συγκεντρώνει την όλη διαδικασία σε μια **βάση ανά ταξίδι-διαδρομή για κάθε οδηγό** στο δείγμα με σκοπό να παραχθεί το τελικό σύστημα βαθμολόγησης του κάθε οδηγού. Κάθε διαδρομή και επομένως κάθε οδηγός αξιολογείται με βάση τα χαρακτηριστικά που αναφέρθηκαν παραπάνω. Η τελική αξιολόγηση που παράγεται περιλαμβάνει τόσο συνολική αξιολόγηση όσο και αξιολόγηση ανά χαρακτηριστικό. **Τα επεξεργασμένα δεδομένα μεταφέρονται σε φιλικές προς το χρήστη εφαρμογές έξυπνων κινητών τηλεφώνων ή σε πλατφόρμες στο διαδίκτυο** που προσφέρουν φιλικά προς το χρήστη περιβάλλοντα για να λάβουν τις αναλυτικές αναφορές των διαδρομών τους. Η παρουσίαση της διαδικασίας των δεδομένων περιγράφεται στην επόμενη παράγραφο. Πρέπει να αναφερθεί ότι όλα τα επεξεργασμένα δεδομένα αποθηκεύονται στη βάση δεδομένων της OSeven ακολουθώντας διαδικασίες ανωνυμίας ώστε τα δεδομένα να μη μπορούν να συνδεθούν με συγκεκριμένο χρήστη.

5. Παρουσίαση δεδομένων

Τα αποτελέσματα όλης της προαναφερθείσας διαδικασίας είναι προσβάσιμα στην εφαρμογή έξυπνων κινητών τηλεφώνων (Smartphone) και στη Διαδικτυακή Πύλη, όπου είναι διαθέσιμα για το χρήστη να δει όλα τα συμβάντα που ανιχνεύθηκαν και τη θέση τους στο χάρτη όπως επίσης και όλα τα αποτελέσματα (συνολικά και ανά κατηγορία). Έτσι, προσφέρεται στον οδηγό ένας τρόπος φιλικός στον χρήστη να αντιλαμβάνεται τα τμήματα ταξιδιού με επικίνδυνη συμπεριφορά οδήγησης και να αποφύγει παρόμοιες συμπεριφορές στο μέλλον. Ταυτόχρονα, οι ασφαλιστικές εταιρείες έχουν πρόσβαση σε δεδομένα των πελατών τους χρησιμοποιώντας τη διαδικτυακή πύλη της OSeven. Τα αποτελέσματα οδήγησης χρησιμοποιούνται για τον καθορισμό των ασφάλιστρων ή/και προγραμμάτων αξιοπιστίας.

4. Συλλογή και Επεξεργασία Στοιχείων



Εικόνα 4.5.: Εφαρμογή στο κινητό και διαδικτυακή πύλη
Πηγή: OSeven

Η βάση δεδομένων που προέκυψε δόθηκε σε αρχείο Microsoft Excel και περιλάμβανε τα εξής μεγέθη/στοιχεία:

No: αύξων αριθμός

tripid: κωδική ονομασία διαδρομής

userid: κωδική ονομασία οδηγού

tripstarted: ημερομηνία/ώρα έναρξης διαδρομής

tripended: ημερομηνία/ώρα λήξης διαδρομής

duration: συνολική διάρκεια διαδρομής σε δευτερόλεπτα (s)

time_urban: συνολική διάρκεια οδήγησης σε αστικό περιβάλλον (s)

time_rural: συνολική διάρκεια οδήγησης σε υπεραστικό περιβάλλον (s)

time_rural/duration: ποσοστό χρόνου οδήγησης σε υπεραστικό περιβάλλον

time_highway: συνολική διάρκεια οδήγησης σε αυτοκινητόδρομο (s)

driving_duration: καθαρή διάρκεια οδήγησης χωρίς τον χρόνο για στάσεις σε δευτερόλεπτα (s)

driving_duration_urban: καθαρή διάρκεια οδήγησης σε αστικό περιβάλλον (s)

driving_duration_rural: καθαρή διάρκεια οδήγησης σε υπεραστικό περιβάλλον (s)

driving_duration_highway: καθαρή διάρκεια οδήγησης σε αυτοκινητόδρομο (s)

totaldist: συνολική απόσταση διαδρομής σε χιλιόμετρα (km)

distance_urban: απόσταση διαδρομής σε αστικό περιβάλλον (km)

distance_rural: απόσταση διαδρομής σε υπεραστικό περιβάλλον (km)

distance_highway: απόσταση διαδρομής σε αυτοκινητόδρομο (km)

ha: απότομη επιτάχυνση(απόλυτος αριθμός)

hb: απότομο φρενάρισμα/επιβράδυνση (απόλυτος αριθμός)

hc: απότομη στροφή (απόλυτος αριθμός)

ha_urban: απότομη επιτάχυνση σε αστικό περιβάλλον

ha_rural: απότομη επιτάχυνση σε υπεραστικό περιβάλλον

ha_highway: απότομη επιτάχυνση σε αυτοκινητόδρομο

hb_urban: απότομη επιβράδυνση σε αστικό περιβάλλον

hb_rural: απότομη επιβράδυνση σε υπεραστικό περιβάλλον

hb_highway: απότομη επιβράδυνση σε αυτοκινητόδρομο

hc_urban: απότομη στροφή σε αστικό περιβάλλον

hc_rural: απότομη στροφή σε υπεραστικό περιβάλλον

hc_highway: απότομη στροφή σε αυτοκινητόδρομο

avaccel: μέση επιτάχυνση (km/h/s)

avdecel: μέση επιβράδυνση (km/h/s)

smooth_corner: μέση στροφική ταχύτητα ($^{\circ}/s$)

av_accel_urban: μέση επιτάχυνση σε αστικό περιβάλλον (km/h/s)

av_accel_rural: μέση επιτάχυνση σε υπεραστικό περιβάλλον (km/h/s)

av_accel_highway: μέση επιτάχυνση σε αυτοκινητόδρομο (km/h/s)

av_decel_urban: μέση επιβράδυνση σε αστικό περιβάλλον (km/h/s)

av_decel_rural: μέση επιβράδυνση σε υπεραστικό περιβάλλον (km/h/s)

av_decel_highway: μέση επιβράδυνση σε αυτοκινητόδρομο (km/h/s)

mobileUsage: ποσοστό της καθαρής διάρκειας οδήγησης (driving_duration) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο

mobileUsage_urban: ποσοστό της καθαρής διάρκειας οδήγησης σε αστικό περιβάλλον (driving_duration_urban) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο

mobileUsage_rural: ποσοστό της καθαρής διάρκειας οδήγησης σε υπεραστικό περιβάλλον (driving_duration_rural) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο

4. Συλλογή και Επεξεργασία Στοιχείων

mobileUsage_highway: ποσοστό της καθαρής διάρκειας οδήγησης σε αυτοκινητόδρομο (driving_duration_highway) που ο οδηγός χρησιμοποιεί το κινητό τηλέφωνο

perc_speeding: ποσοστό της καθαρής διάρκειας οδήγησης (driving_duration) με ταχύτητα πάνω από το όριο ταχύτητας

perc_speeding_urban: ποσοστό της καθαρής διάρκειας οδήγησης σε αστικό περιβάλλον (driving_duration_urban) με ταχύτητα πάνω από το όριο ταχύτητας

perc_speeding_rural: ποσοστό της καθαρής διάρκειας οδήγησης σε υπεραστικό περιβάλλον (driving_duration_rural) με ταχύτητα πάνω από το όριο ταχύτητας

perc_speeding_highway: ποσοστό της καθαρής διάρκειας οδήγησης σε αυτοκινητόδρομο (driving_duration_highway) με ταχύτητα πάνω από το όριο ταχύτητας

av_speeding: μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας

av_speeding_urban: μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε αστικό περιβάλλον

av_speeding_rural: μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε υπεραστικό περιβάλλον

av_speeding_highway: μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε αυτοκινητόδρομο

avgspeed: μέση ταχύτητα (km/h)

maxspeed: μέγιστη τιμή ταχύτητας (km/h)

speedq25: τιμή που αντιστοιχεί στο 25% της καταγεγραμμένης ταχύτητας

speedq50: τιμή που αντιστοιχεί στο 50% της καταγεγραμμένης ταχύτητας

speedq75: τιμή που αντιστοιχεί στο 75% της καταγεγραμμένης ταχύτητας

StSpDev: τυπική απόκλιση ταχύτητας (km/h)

avjerk: μέσος όρος ρυθμού μεταβολής της επιτάχυνσης (km/h/s²)

avjerkp: θετικός ρυθμός μεταβολής της επιτάχυνσης (km/h/s²)

avjerkn: μειούμενος ρυθμός μεταβολής της επιτάχυνσης (km/h/s²)

riskyhoursdistance: απόσταση οδήγησης από τις 22:00 μέχρι τις 05:00

Επιπλέον, **προστέθηκαν** στήλες με **καινούργιες μεταβλητές** οι οποίες κρίθηκαν απαραίτητες για την ανάλυση του στόχου της παρούσας Διπλωματικής. Συγκεκριμένα, δημιουργήθηκε η μεταβλητή **htot/totaldist** η οποία αντιπροσωπεύει τον συνολικό αριθμό απότομων γεγονότων (events) προς τη συνολική απόσταση οδήγησης. Η ίδια μεταβλητή δημιουργήθηκε ξεχωριστά και για αστικό και υπεραστικό περιβάλλον και για αυτοκινητόδρομο. Ακόμα, δημιουργήθηκαν δίτιμες κατηγορικές μεταβλητές (0 και 1)

για την πρωινή αιχμή (**driving_morning_rush**) και την απογευματινή αιχμή (**driving_afternoon_rush**) καθώς και για την οδήγηση με φως ημέρας (**driving_daylight**). Η πρωινή αιχμή περιλαμβάνει τις ώρες μεταξύ 06:00 και 10:00, η απογευματινή αιχμή τις ώρες μεταξύ 16:00 και 20:00 και η οδήγηση με φως ημέρας τις ώρες μεταξύ 06:00 και 21:00. Για να προβλεφθεί η χρήση ή η μη χρήση κινητού τηλεφώνου κατά την οδήγηση, η μεταβλητή **mobileUsage** που είναι ποσοστό χρόνου μετατράπηκε σε κατηγορική μεταβλητή (0 για μη χρήση του κινητού τηλεφώνου και 1 για χρήση του κινητού τηλεφώνου), οπότε ορίστηκε η μεταβλητή **mobileUsage/YES-NO**. Η ίδια μεταβλητή ορίστηκε και **ανά τύπο οδού** (**mobileUsage_urban/YES-NO** κ.ο.κ.). Επίσης, δημιουργήθηκαν οι μεταβλητές **mobileUsage/MORE-LESS** για διάφορα ποσοστά του **mobileUsage**. Έτσι, για παράδειγμα, η μεταβλητή **mobileUsage/MORE-LESS0.1** έχει ως τιμή το 0 όταν το ποσοστό του χρόνου που ο οδηγός μιλάει στο κινητό τηλέφωνο (**mobileUsage**) είναι μικρότερο από 0.1 και ως τιμή το 1 όταν η μεταβλητή **mobileUsage** είναι μεγαλύτερη του 0.1. Σχηματίστηκε η μεταβλητή **mobileUsage_time** η οποία αντιπροσωπεύει τον συνολικό χρόνο σε δευτερόλεπτα (εν αντιθέσει με τη μεταβλητή **mobileUsage** που είναι ποσοστό) χρήσης του κινητού τηλεφώνου. Το ίδιο πραγματοποιήθηκε και για κάθε επιμέρους τύπο οδού (**time_mobile_urban**, **time_mobile_rural**, **time_mobile_highway**). Τέλος, δημιουργήθηκαν πρόσθετες μεταβλητές όπως **1/duration**, **avspeed**² κλπ οι οποίες δοκιμάστηκαν για την εξαγωγή καλύτερων μαθηματικών μοντέλων.

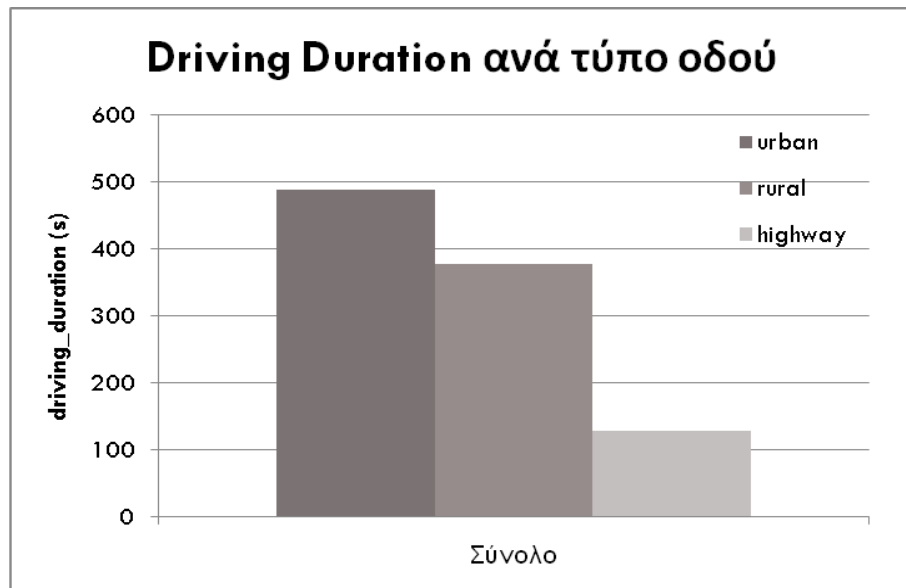
4.3 Επεξεργασία Στοιχείων

Στην παράγραφο αυτή παρατίθενται ορισμένα διαγράμματα που δημιουργήθηκαν στη φάση της προκαταρκτικής ανάλυσης με σκοπό την καλύτερη κατανόηση των δεδομένων. Ακόμα, παρουσιάζεται επιγραμματικά η διαδικασία εισαγωγής των μεταβλητών που προέκυψαν στο λογισμικό της στατιστικής ανάλυσης.

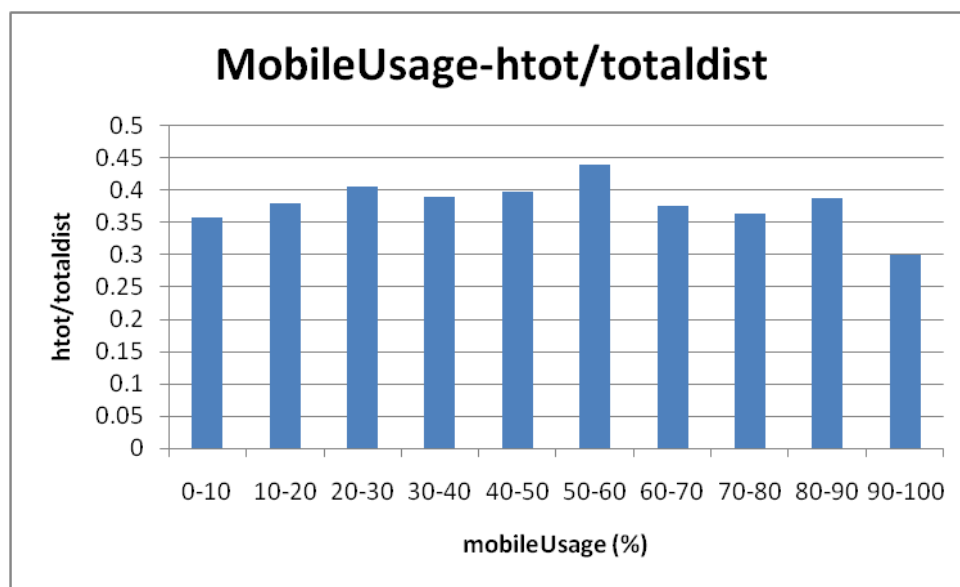
4.3.1 Προκαταρκτική Ανάλυση

Τα παρακάτω διαγράμματα αποτελούν μια προκαταρκτική ανάλυση, η οποία συμβάλλει στην καλύτερη κατανόηση των αποτελεσμάτων και θα χρησιμοποιηθεί για την εξαγωγή ποιοτικότερων συμπερασμάτων.

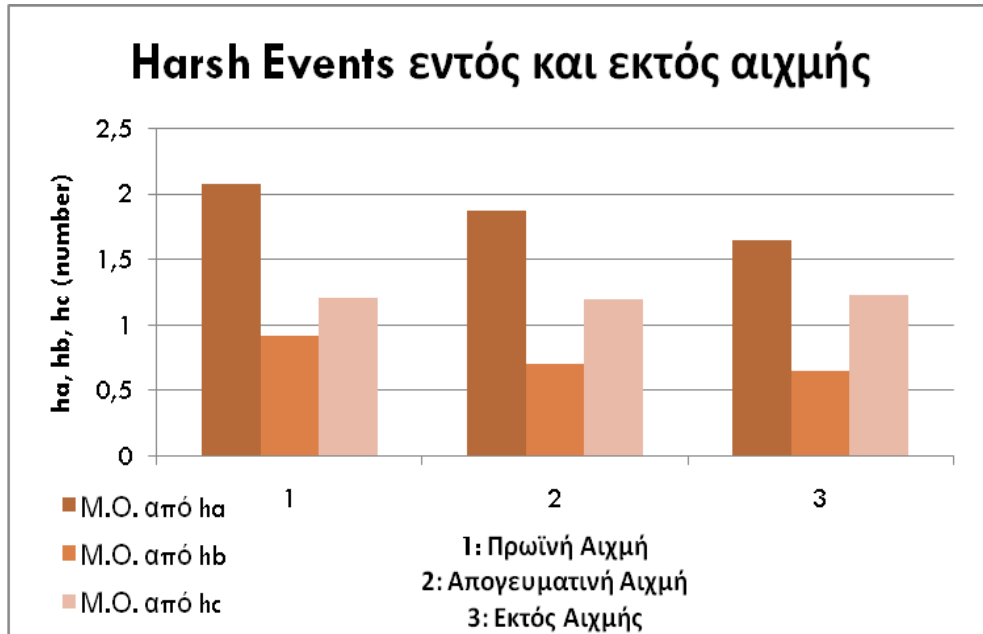
4. Συλλογή και Επεξεργασία Στοιχείων



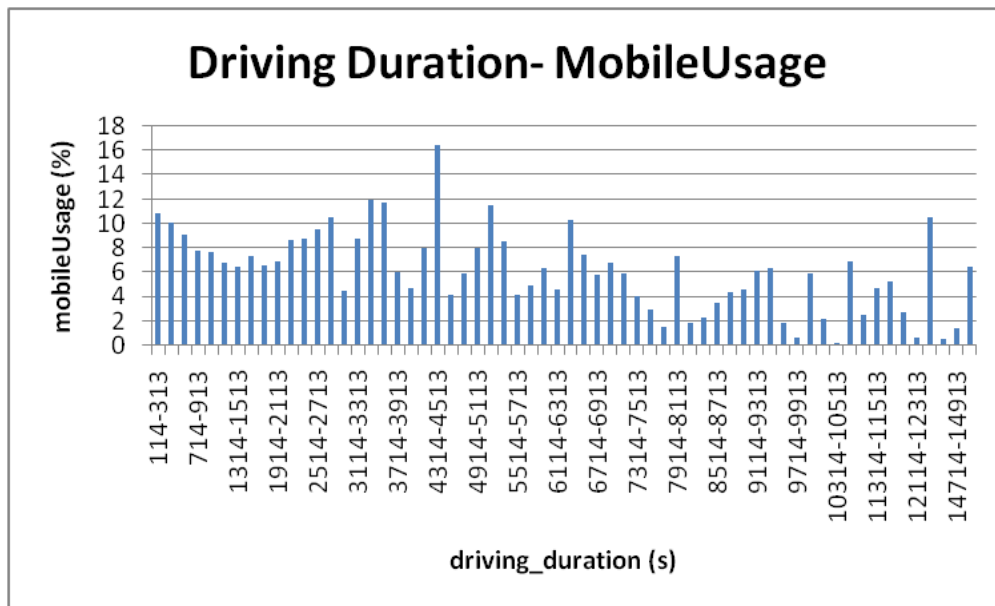
Διάγραμμα 4.1.: Καθαρή διάρκεια οδήγησης ανά τύπο οδού



Διάγραμμα 4.2.: Ποσοστό χρήσης κινητού τηλεφώνου σε συνάρτηση με τα συνολικά απότομα γεγονότα (events) προς τη συνολική απόσταση οδήγησης

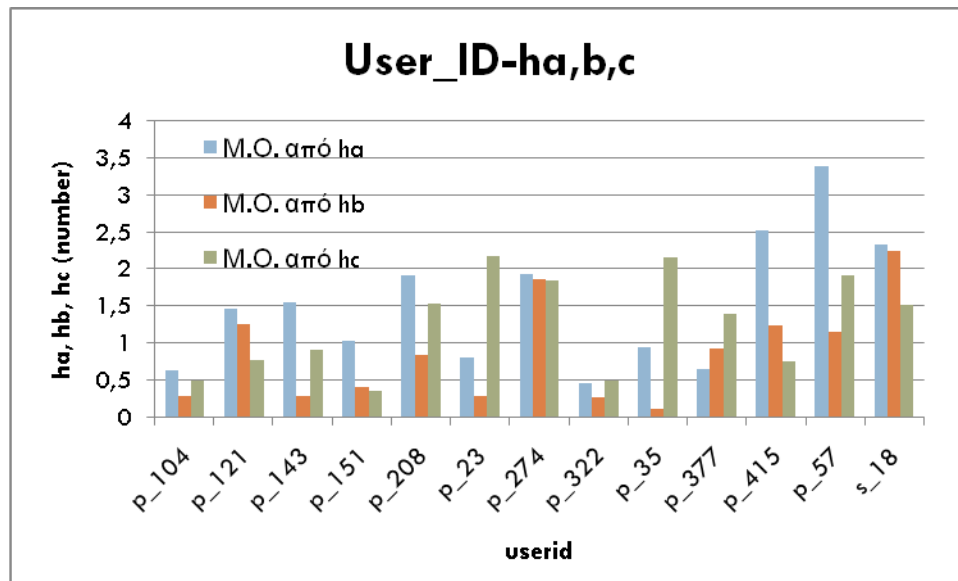


Διάγραμμα 4.3.: Απότομα γεγονότα εντός και εκτός ωρών αιχμής

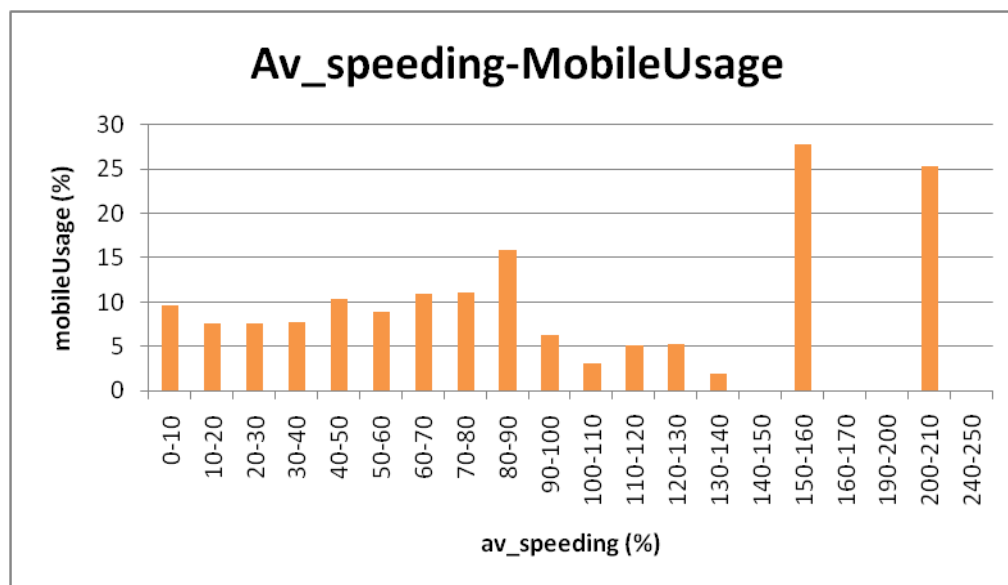


Διάγραμμα 4.4.: Καθαρή διάρκεια οδήγησης σε συνάρτηση με το ποσοστό χρήσης κινητού τηλεφώνου

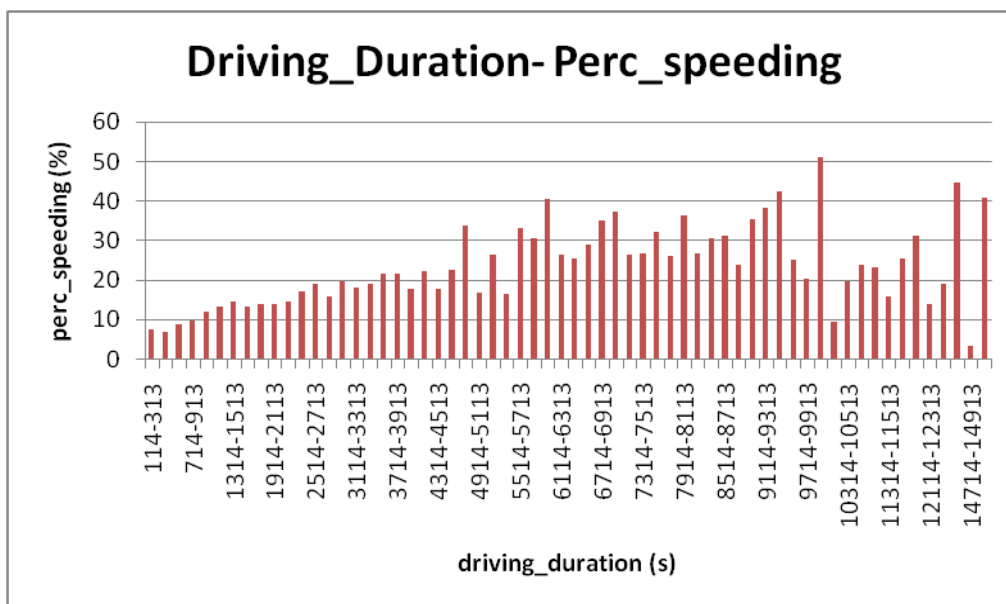
4. Συλλογή και Επεξεργασία Στοιχείων



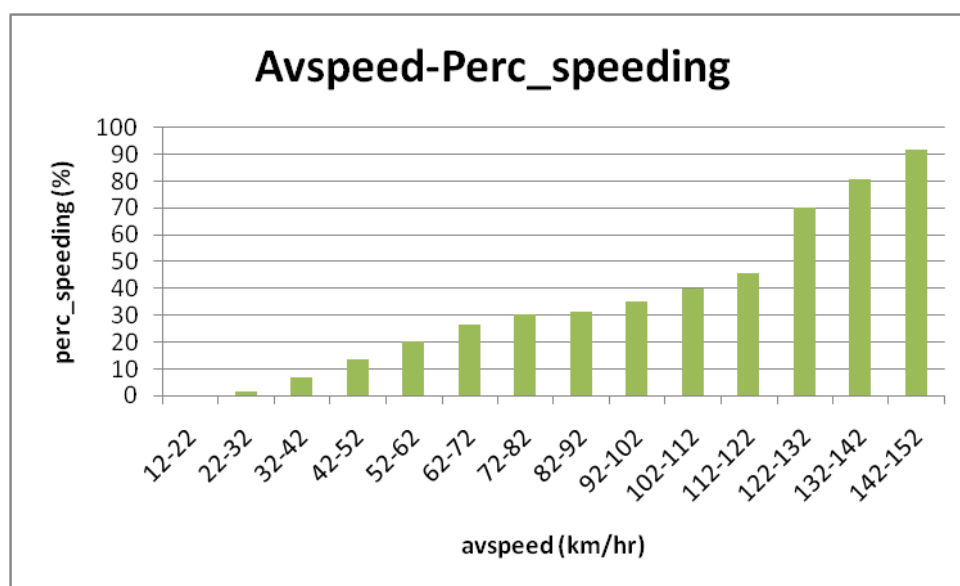
Διάγραμμα 4.5.: Απότομα γεγονότα (events) ανά χρήστη (ενδεικτικά για κάποιους χρήστες)



Διάγραμμα 4.6.: Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε συνάρτηση με το ποσοστό χρήσης του κινητού τηλεφώνου

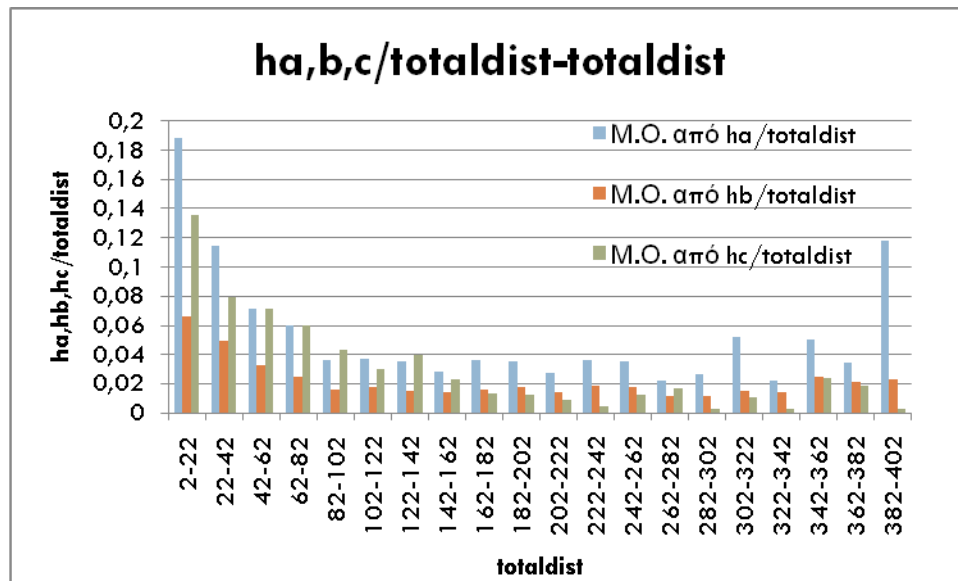


Διάγραμμα 4.7: Καθαρή διάρκεια οδήγησης σε συνάρτηση με το ποσοστό της διάρκειας πάνω από το όριο ταχύτητας

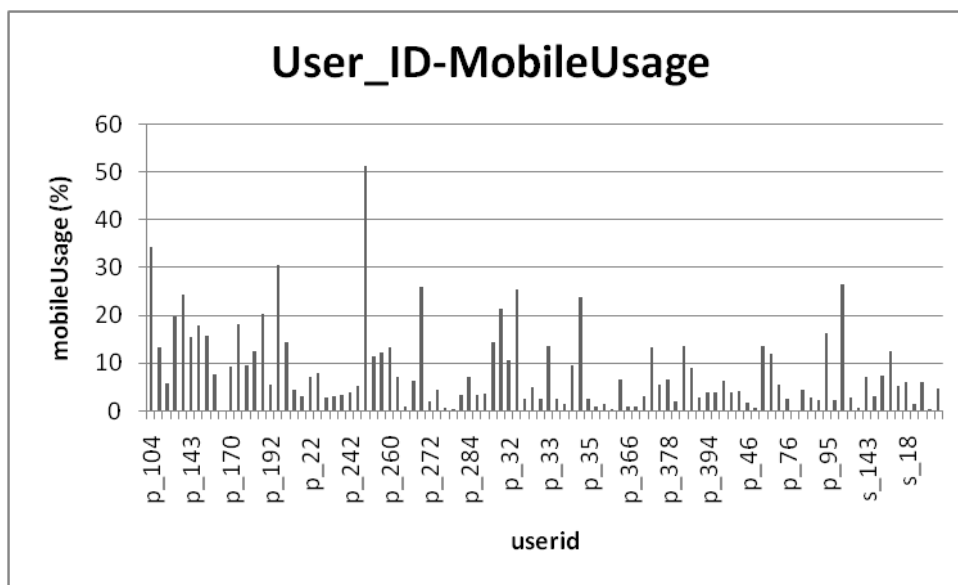


Διάγραμμα 4.8.: Μέση ταχύτητα σε συνάρτηση με το ποσοστό του χρόνου πάνω από το όριο ταχύτητας

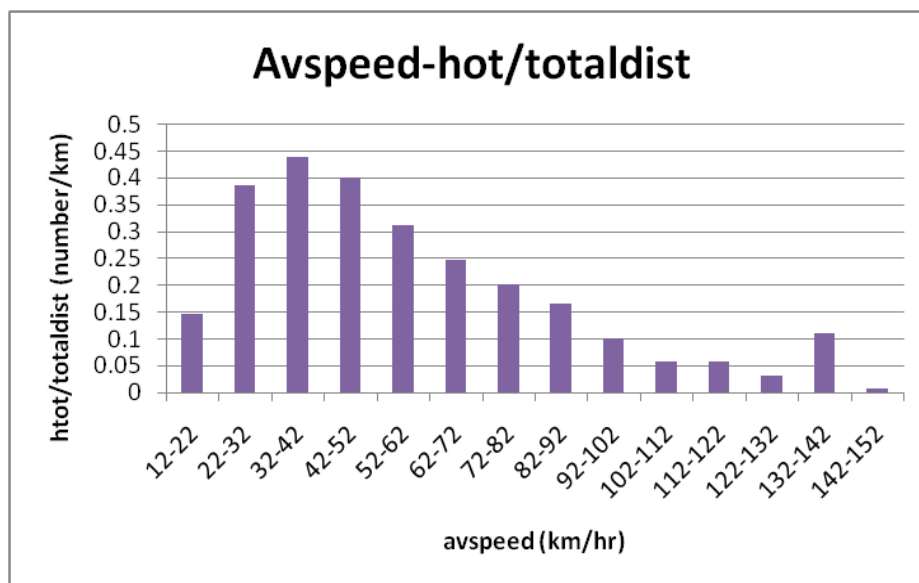
4. Συλλογή και Επεξεργασία Στοιχείων



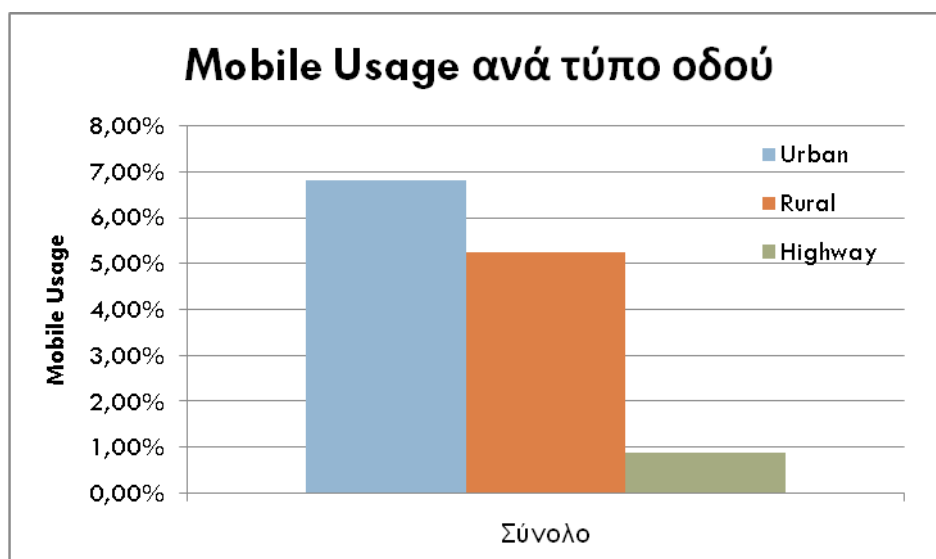
Διάγραμμα 4.9.: Απότομα γεγονότα ανά συνολική απόσταση σε συνάρτηση με τη συνολική απόσταση οδήγησης



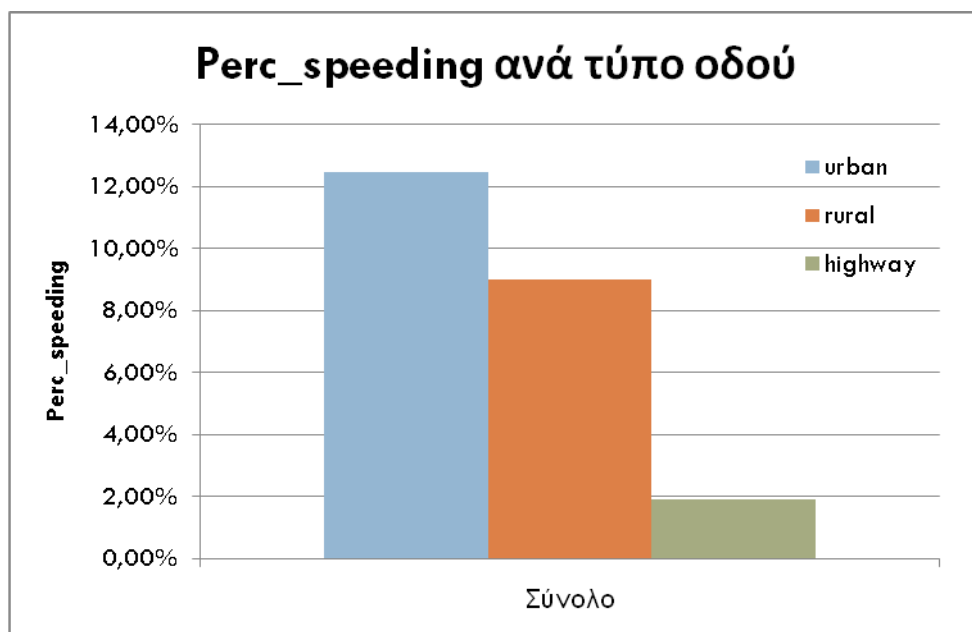
Διάγραμμα 4.10.: Ποσοστό χρήσης κινητού τηλεφώνου ανά οδηγό



Διάγραμμα 4.11.: Μέση ταχύτητα σε συνάρτηση με τα συνολικά απότομα (harsh) γεγονότα προς τη συνολική απόσταση οδήγησης



Διάγραμμα 4.12.: Ποσοστό χρήσης κινητού τηλεφώνου ανά τύπο οδού



Διάγραμμα 4.13.: Ποσοστό χρόνου υπέρβασης του ορίου ταχύτητας ανά τύπο οδού

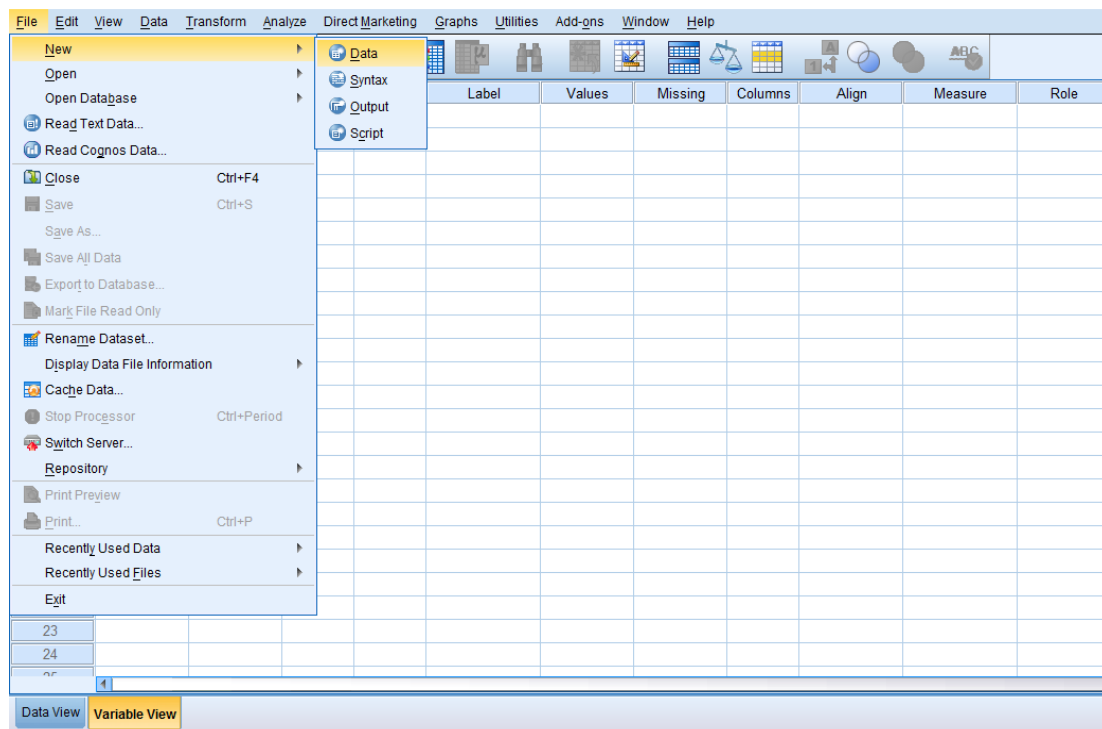
Εξετάζοντας τα ανωτέρω διαγράμματα προκύπτουν οι εξής παρατηρήσεις:

- ✓ Η μεγαλύτερη διάρκεια οδήγησης πραγματοποιείται σε **αστικό περιβάλλον**.
- ✓ Ο αριθμός των συνολικών απότομων γεγονότων (events) προς τη συνολική απόσταση (**htot/totaldist**) κυμαίνεται γύρω από την τιμή 0.4 και αποκτά τη μέγιστη τιμή ίση με 0.45 όταν το ποσοστό του χρόνου χρήσης του κινητού είναι 50-60%.
- ✓ Τα περισσότερα απότομα γεγονότα (harsh events) γίνονται κατά σειρά στην **πρωινή αιχμή**, μετά στην απογευματινή αιχμή και τέλος εκτός αιχμής.
- ✓ Όσο **μικρότερη** είναι η **διάρκεια** του ταξιδιού τόσο **μεγαλύτερο** είναι το **ποσοστό χρήσης** του κινητού τηλεφώνου. Αυτό μπορεί πιθανόν να εξηγηθεί καθώς τα μεγαλύτερης χρονικής διάρκειας ταξίδια γίνονται συνήθως σε αυτοκινητοδρόμους οπότε η χρήση του κινητού είναι μικρότερη λόγω επικινδυνότητας, όπως απεικονίζεται και στο σχετικό διάγραμμα 4.7.
- ✓ Μεγαλύτερη **υπέρβαση του ορίου ταχύτητας** παρατηρείται για ποσοστό χρόνου ομιλίας στο κινητό τηλέφωνο κοντά στο 28%.
- ✓ Όσο **μεγαλύτερη** είναι η **διάρκεια οδήγησης** τόσο μεγαλύτερο είναι και το ποσοστό του χρόνου πάνω από το όριο ταχύτητας, γεγονός που πιθανόν υποδηλώνει ότι στους **αυτοκινητοδρόμους** όπου γίνονται συνήθως τα ταξίδια μεγαλύτερης διάρκειας υπάρχουν **μεγαλύτερες υπερβάσεις**.
- ✓ Το ποσοστό του χρόνου υπέρβασης του ορίου ταχύτητας είναι μεγαλύτερο σε **αστικό περιβάλλον**.
- ✓ Όσο μεγαλύτερη είναι η **μέση ταχύτητα** τόσο μεγαλύτερο είναι το ποσοστό του χρόνου πάνω από το όριο ταχύτητας.

- ✓ Τα **περισσότερα απότομα γεγονότα** (harsh events) ανά απόσταση οδήγησης πραγματοποιούνται στα ταξίδια μικρότερης συνολικής απόστασης, καθώς ο οδηγός δεν έχει προλάβει να προσαρμοστεί επαρκώς στις συνθήκες του οδικού περιβάλλοντος
- ✓ Σε **μικρές και μέσες ταχύτητες οδήγησης** παρατηρούνται τα περισσότερα απότομα γεγονότα (harsh events) ανά απόσταση οδήγησης και συγκεκριμένα γύρω στα 30-50 km/h. Αυτό πιθανόν να μπορούσε να εξηγηθεί από την ύπαρξη κορεσμένου οδικού δικτύου που οδηγεί σε χαμηλές ταχύτητες οδήγησης και σε απότομες επιταχύνσεις και επιβραδύνσεις.

4.3.2 Εισαγωγή βάσης δεδομένων στο ειδικό λογισμικό στατιστικής ανάλυσης, IBM SPSS 21.0

Η βάση δεδομένων δόθηκε σε μορφή αρχείου Excelσυμβατή με το στατιστικό πρόγραμμα. Όπως αναφέρθηκε παραπάνω δημιουργήθηκαν επιπλέον **στήλες για νέες μεταβλητές** (π.χ. driving_morning_rush, driving_daylight, mobileUsage_time, time_mobile_highway, mobileUsage/YES-NO κλπ.), ορισμένες από τις οποίες χρησιμοποιήθηκαν εν τέλει στα μοντέλα. Μετά τη διαμόρφωση του τελικού πίνακα στο λογισμικό Excel πραγματοποιήθηκε εισαγωγή του στο πεδίο δεδομένων (**data view**) του ειδικού πακέτου στατιστικής ανάλυσης, όπως απεικονίζεται παρακάτω.



Εικόνα 4.6.: Τρόπος Εισαγωγής δεδομένων στο SPSS

4. Συλλογή και Επεξεργασία Στοιχείων

No	tripid	userid	tripstarted	tripended	duration	Visible
1	1 B714DAE-6946-45B9-BFA5-3E8ECD7CD52F	p_104	2016-09-29T07:50:21.021+0300	2016-09-29T07:55:20.000+0300	300	
2	2 C6D8FCA5-621C-4FBD-B540-D16B62493491	p_104	2016-09-29T09:08:24.024+0300	2016-09-29T09:30:42.000+0300	1338	
3	3 0B8C5D88-9E03-4553-AA48-EB885EA9A288	p_104	2016-09-29T23:12:49.049+0300	2016-09-29T23:33:37.000+0300	1249	
4	4 EB58D17A-9A3B-47DA-A6DD-B6FCF58BEFFC	p_104	2016-09-30T06:25:04.004+0300	2016-09-30T06:29:49.000+0300	286	
5	5 B25A09C2-462A-4EEF-8B09-A0368E515CEB	p_104	2016-09-30T14:23:43.043+0300	2016-09-30T14:43:52.000+0300	1210	
6	6 A4731369-819F-4037-8D9F-C2B16F5C05F3	p_104	2016-09-30T21:16:35.035+0300	2016-09-30T21:31:42.000+0300	908	
7	7 601A7642-4952-4ED5-9D63-13877C6B0CB	p_104	2016-10-01T00:44:49.049+0300	2016-10-01T00:51:57.000+0300	429	
8	8 7F24480-17F4-49F0-8CB1-3654B178C778	p_104	2016-10-01T19:23:18.018+0300	2016-10-01T19:31:12.000+0300	475	
9	9 0A3EF9E2-9359-4495-A807-F0B987A9D177	p_104	2016-10-01T21:54:04.004+0300	2016-10-01T22:09:53.000+0300	950	
10	10 AC496424-3F79-4D08-8185-51D67AC29217	p_104	2016-10-02T01:29:57.057+0300	2016-10-02T01:38:52.000+0300	536	
11	11 5C409D1E-6677-46A3-99C9-B6964EFEE2B4	p_104	2016-10-02T17:17:03.003+0300	2016-10-02T17:23:11.000+0300	369	
12	12 22A91AC6-5E22-49F1-B39D-DBF03D4651EB	p_104	2016-10-03T00:34:18.018+0300	2016-10-03T00:46:37.000+0300	740	
13	13 EEBD733F-597E-4C74-BD29-6070990B4D83	p_104	2016-10-03T07:33:03.033+0300	2016-10-03T07:33:08.000+0300	1356	
14	14 B113ECDD-83D9-43D4-AAAB-99C8E17D72A6	p_104	2016-10-03T23:44:48.048+0300	2016-10-03T23:53:09.000+0300	502	
15	15 405C1B31-AEEB-48BB-A483-2E42E07DDEC7	p_104	2016-10-04T09:28:29.029+0300	2016-10-04T09:50:20.000+0300	1311	
16	16 2C369686-1C59-4066-A6E9-54EC28F2D55E	p_104	2016-10-04T19:16:20.020+0300	2016-10-04T19:38:10.000+0300	1311	
17	17 367238CA-B8F7-4AF2-A8A3-0C597AD33228	p_104	2016-10-05T07:59:29.029+0300	2016-10-05T08:07:30.000+0300	482	
18	18 83E98955-353A-4CED-92B2-7E38614B86B5	p_104	2016-10-05T09:14:09.009+0300	2016-10-05T09:24:57.000+0300	649	
19	19 55067ACA-807D-4908-A781-9C2436C55414	p_104	2016-10-05T09:42:46.046+0300	2016-10-05T09:55:29.000+0300	764	
20	20 00E3D7EF-7761-4C45-AE1A-49CBF65715EE	p_104	2016-10-05T19:17:00.000+0300	2016-10-05T19:37:47.000+0300	1248	
21	21 A6D431BA-D4A1-447E-A849-F9AFB956C35F	p_104	2016-10-06T08:00:37.037+0300	2016-10-06T08:07:10.000+0300	394	
22	22 853375F9-5968-4B6B-BA44-E9328D85138D	p_104	2016-10-06T09:18:17.017+0300	2016-10-06T10:04:05.000+0300	2749	

Εικόνα 4.7.: Πεδίο Δεδομένων SPSS (data view)

Στη συνέχεια καθορίστηκε το όνομα, ο τύπος, και ο αριθμός των ψηφίων κάθε μεταβλητής στο πεδίο των μεταβλητών (**variable view**). Επισημαίνεται ότι, το πρόγραμμα αυτό αναγνωρίζει μόνο λατινικούς χαρακτήρες και όλες οι στήλες αποτελούνται από αριθμούς και όχι από κείμενο. Επίσης, έγινε διάκριση κάθε μεταβλητής σε συνεχή (**scale**), διατεταγμένη (**ordinal**) και διακριτή (**nominal**). Για παράδειγμα η μεταβλητή `driving_morning_rush` ορίστηκε ως **nominal** καθώς παίρνει μόνο τις τιμές 0 και 1.

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
tripended_minutes	Numeric	12	0		None	None	12	Right	Scale	Input
tripended_seconds	String	11	0		None	None	11	Left	Nominal	Input
driving_morning_rush	Numeric	12	0		None	None	12	Right	Nominal	Input
driving_afternoon_rush	Numeric	12	0		None	None	12	Right	Nominal	Input
driving_daylight	Numeric	12	0		None	None	12	Right	Nominal	Input
durations	Numeric	12	0		None	None	12	Right	Scale	Input
durationhr	Numeric	17	16		None	None	17	Right	Scale	Input
time_urban	Numeric	12	0		None	None	12	Right	Scale	Input
time_urbanrural	Numeric	17	16		None	None	17	Right	Scale	Input
time_rural	Numeric	12	0		None	None	12	Right	Scale	Input
time_ruralrural	Numeric	17	16		None	None	17	Right	Scale	Input
time_highway	Numeric	12	0		None	None	12	Right	Nominal	Input
time_highwayrural	Numeric	17	16		None	None	17	Right	Scale	Input
driving_duration	Numeric	12	0		None	None	12	Right	Scale	Input
driving_duration_urban	Numeric	12	0		None	None	12	Right	Scale	Input
driving_duration_rural	Numeric	12	0		None	None	12	Right	Scale	Input
driving_duration_highway	Numeric	12	0		None	None	12	Right	Scale	Input
totaldist	Numeric	17	16		None	None	17	Right	Scale	Input
distance_urban	Numeric	17	16		None	None	17	Right	Scale	Input
distance_rural	Numeric	17	16		None	None	17	Right	Scale	Input
distance_highway	Numeric	17	16		None	None	17	Right	Scale	Input
ha	Numeric	12	0		None	None	12	Right	Scale	Input
hatotaldist	Numeric	17	16		None	None	17	Right	Scale	Input
hb	Numeric	12	0		None	None	12	Right	Scale	Input
hbatotaldist	Numeric	17	16		None	None	17	Right	Scale	Input

Εικόνα 4.8.: Πεδίο μεταβλητών SPSS (Variable View)

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

5.1 Εισαγωγή

Ύστερα από τη βιβλιογραφική ανασκόπηση συναφών ερευνών και μεθοδολογιών, έγινε η επιλογή της κατάλληλης μεθοδολογίας. Οι μέθοδοι που επιλέχθηκαν, οι οποίες παρουσιάζονται αναλυτικά στο κεφάλαιο 3 με τίτλο «Θεωρητικό Υπόβαθρο», είναι: το **Μεικτό Μοντέλο Γραμμικής Παλινδρόμησης** (Mixed Linear Regression) και το **Μεικτό Μοντέλο Δυαδικής Λογιστικής Παλινδρόμησης** (Mixed Binary Logistic Regression). Πιο συγκεκριμένα, στο κεφάλαιο αυτό περιγράφονται αναλυτικά τα βήματα που ακολουθήθηκαν κατά την εφαρμογή της μεθοδολογίας και παρουσιάζεται η διαδικασία ανάπτυξης κατάλληλων μοντέλων. Αναλύονται οι δοκιμές που πραγματοποιήθηκαν ακόμα και αυτές που εν τέλει δεν οδήγησαν σε ικανοποιητικά αποτελέσματα, ωστόσο, αποτέλεσαν το υπόβαθρο πάνω στο οποίο στηρίχθηκαν τα τελικά μοντέλα. Αναπόσπαστο μέρος των αποτελεσμάτων αποτελούν οι στατιστικοί έλεγχοι που απαιτούνται για την αποδοχή ή την απόρριψη των μαθηματικών μοντέλων.

Τέλος, παρατίθενται τα αποτελέσματα που προκύπτουν από την εφαρμογή των μεθοδολογιών, η περιγραφή τους, και η ερμηνεία τους με βάση το γενικότερο πλαίσιο της έρευνας.

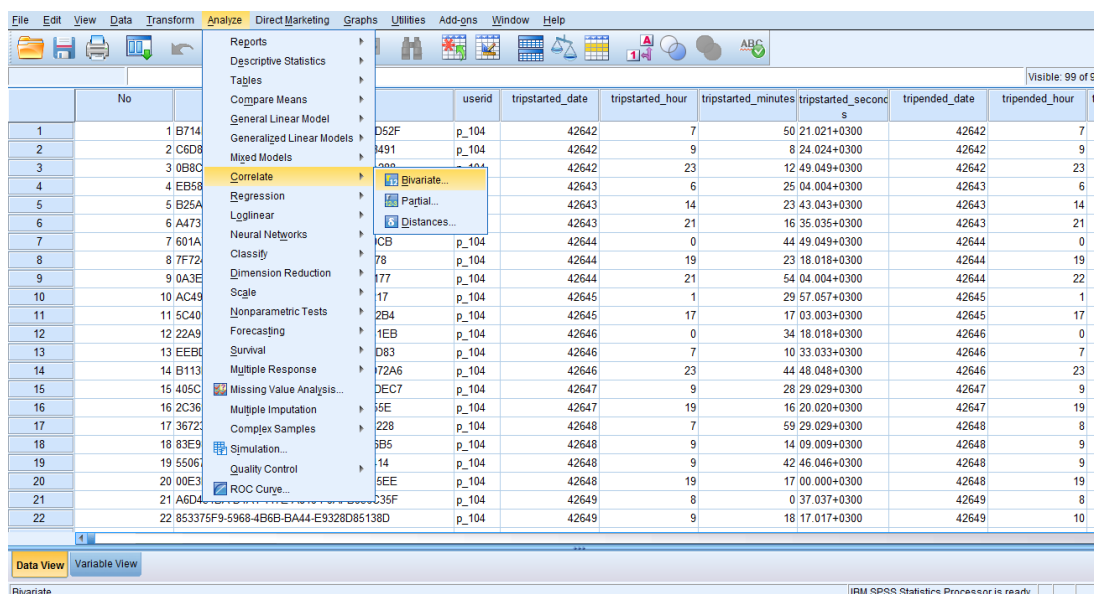
5.2 Έλεγχος Συσχέτισης

Για να ολοκληρωθεί ο στόχος της Διπλωματικής Εργασίας, δηλαδή η δημιουργία μοντέλων για τη διερεύνηση της απόσπασης προσοχής από τη χρήση κινητού τηλεφώνου, εξετάστηκε σε πρώτη φάση η συσχέτιση μεταξύ των μεταβλητών. Εκείνο που επιδιώκεται είναι η **μέγιστη δυνατή συσχέτιση** μεταξύ εξαρτημένης και ανεξάρτητων μεταβλητών και μηδενική συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών. Απόλυτες τιμές των συντελεστών συσχέτισης κοντά στη μονάδα δείχνουν ισχυρή συσχέτιση, ενώ τιμές κοντά στο μηδέν φανερώνουν ανύπαρκτη συσχέτιση μεταξύ των μεταβλητών. Στην πράξη θεωρούμε μικρή συσχέτιση μεταξύ δύο μεταβλητών όταν η απόλυτη του δείκτη συσχέτισης κατά Pearson r είναι μικρότερη ή ίση με **0.5~0.6** ($r \leq 0.5 \sim 0.6$).

Για τον έλεγχο ακολουθήθηκε η παρακάτω διαδικασία στο πρόγραμμα SPSS, όπως απεικονίζεται και στις σχετικές εικόνες.

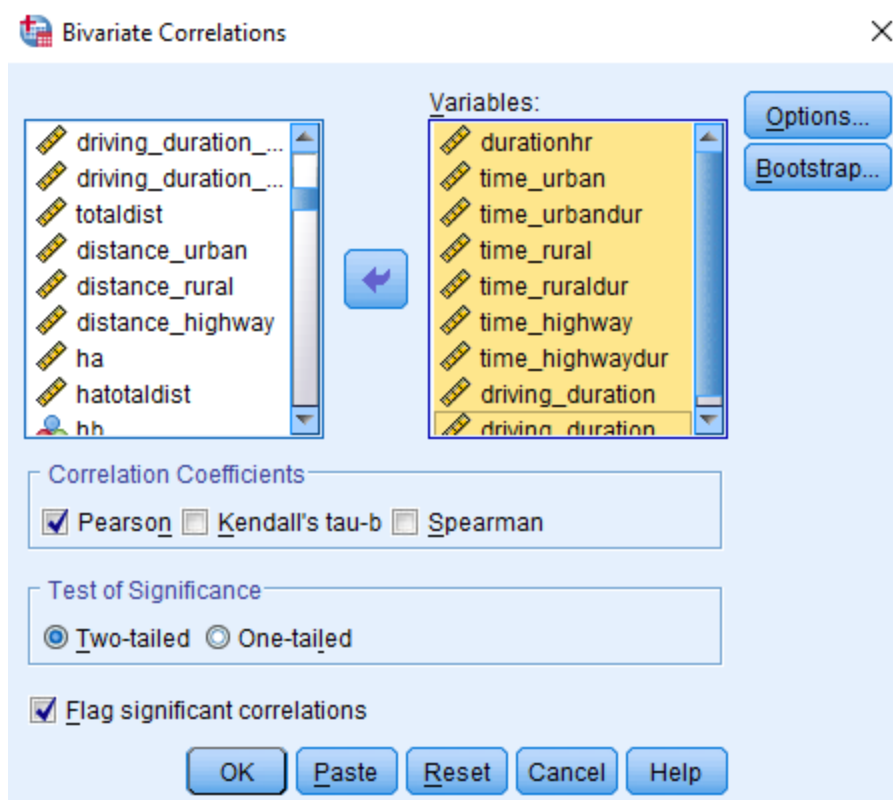
analyze → correlate → bivariate

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



Εικόνα 5.1: Επιλογή εντολής για τον έλεγχο συσχέτισης των μεταβλητών

Οι μεταβλητές που παρουσιάζουν ενδιαφέρον για το αντικείμενο της παρούσας Διπλωματικής Εργασίας εισάγονται στο πεδίο Variables.



Εικόνα 5.2.: Επιλογή παραμέτρων για τον έλεγχο συσχέτισης των μεταβλητών

Στη συνέχεια, το αποτέλεσμα της συσχέτισης μεταβλητών, εισάχθηκε σε φύλλο υπολογισμού Excel με τη μορφή πίνακα, όπου επεξεργάστηκε και χρωματίστηκαν τα κελιά των μεταβλητών που έχουν απόλυτη τιμή $r \leq 0.5$. Απόσπασμα του πίνακα παρουσιάζεται στη συνέχεια στην εικόνα 5.3.

		time_urban	time_rural	durations	htottotaldist	smooth_corner	time_urban	time_rural	time_highway	time_highwaydur	driving_duration	driving_duration_urban
time_urban	Pearson Correlation	1	-.880	-.291	.066	.229	.374	-.594	-.279	-.395	-.318	.337
	Sig. (2-tailed)		0.000	0.000	.000	.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853
time_rural	Pearson Correlation	-.880	1	.109	.015	-.177	-.361	.589	-.066	-.090	.108	-.327
	Sig. (2-tailed)	0.000		.000	.038	.000	0.000	0.000	.000	.000	.000	0.000
	N	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853
durations	Pearson Correlation	-.291	.109	1	-.139	-.130	.555	.723	.690	.401	.972	.562
	Sig. (2-tailed)	0.000	.000		.000	.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853
htottotaldist	Pearson Correlation	.066	.015	-.139	1	.147	-.072	-.069	-.131	-.167	-.159	-.085
	Sig. (2-tailed)	.000	.038	.000		.000	.000	.000	.000	.000	.000	.000
	N	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853	18853

Εικόνα 5.3: Παραδείγματα συσχέτισης μεταβλητών

Οι μεταβλητές οι οποίες δεν παρουσίαζαν συσχέτιση μεταξύ τους συγκεντρώθηκαν και εισάγονταν σταδιακά στα μοντέλα για την πραγματοποίηση δοκιμών. Ενδεικτικά, ασυσχέτιστες μεταβλητές ήταν οι: duration, smooth_corner, avjerkp, StSpDev, avgspeed, speedq25, av_speeding, mobileUsage, riskyhoursdiastance κλπ. Σημειώνεται ότι μεταβλητές οι οποίες σχετίζονται άμεσα με την εξαρτημένη μεταβλητή, όπως για παράδειγμα η ανεξάρτητη smooth_corner (μέση στροφική ταχύτητα) που σχετίζεται άμεσα με την εξαρτημένη htot/totaldist (απότομες επιταχύνσεις, επιβραδύνσεις, ελιγμοί ανά απόσταση), δεν δοκιμάζονται στα μοντέλα καθώς θεωρείται ότι οδηγούν σε μη αντιπροσωπευτικά αποτελέσματα.

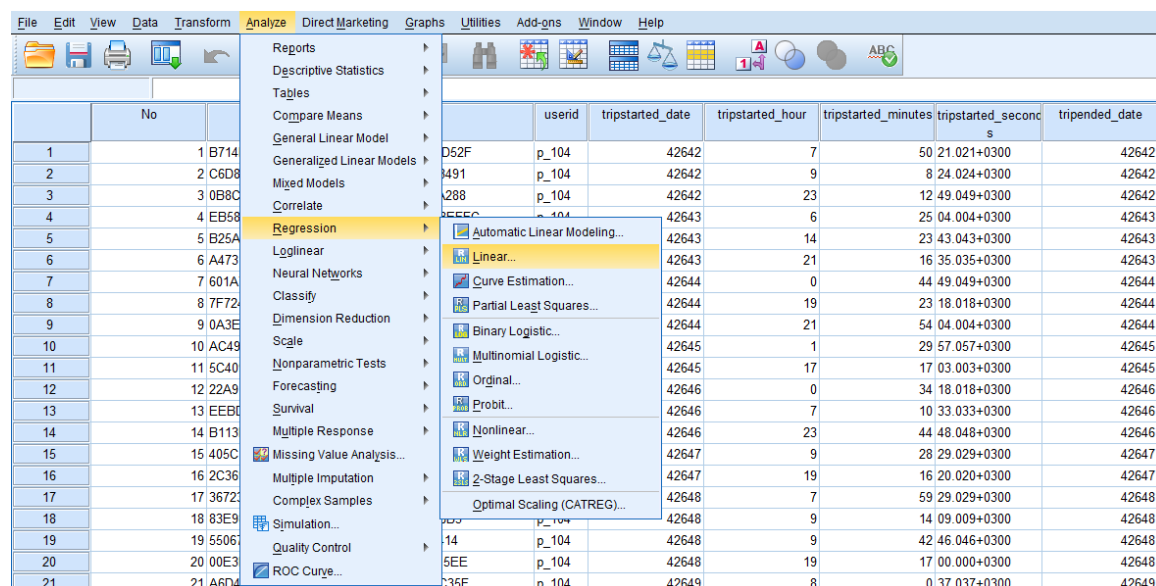
5.3 Γραμμική παλινδρόμηση

5.3.1 Διαδικασία Γραμμικής Παλινδρόμησης

Για τη διερεύνηση της επίδρασης της απόστασης προσοχής μέσω της χρήσης κινητού τηλεφώνου στον αριθμό των απότομων συμβάντων ανά απόσταση (htot/totaldist) δοκιμάστηκε σε πρώτη φάση η μέθοδος της γραμμικής παλινδρόμησης. Αυτή η επιλογή εξηγείται από το γεγονός ότι η μεταβλητή που εξετάζεται (εξαρτημένη) είναι συνεχής. Τα βήματα που εφαρμόζονται μέσω του προγράμματος SPSS είναι τα εξής, όπως απεικονίζονται και στη σχετική εικόνα:

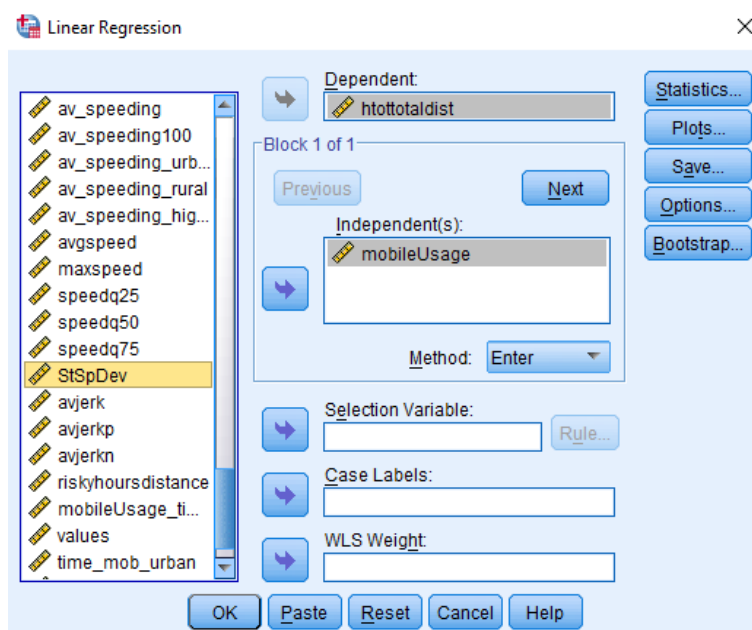
analyze → regression → linear

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



Εικόνα 5.4: Επιλογή εντολής γραμμικής παλινδρόμησης

Στη συνέχεια καθορίζονται η εξαρτημένη και οι ανεξάρτητες μεταβλητές και εισάγονται αντίστοιχα στα πεδία Dependent και Independent(s). Στο πλαίσιο Method μπορεί να επιλεγεί μια μέθοδος για τη βέλτιστη επιλογή επεξηγηματικών μεταβλητών. Αυτή συνήθως αφήνεται «Enter», που σημαίνει ότι στο μοντέλο εισέρχονται όσες μεταβλητές βρίσκονται στο πλαίσιο Independent(s) με τη σειρά που γράφονται εκεί. Σταδιακά εισάγονται πιθανές ανεξάρτητες (επεξηγηματικές) μεταβλητές έως ότου καταλήξουμε σε επιθυμητά αποτελέσματα.



Εικόνα 5.5: Επιλογή ανεξάρτητων και εξαρτημένων μεταβλητών για τη γραμμική παλινδρόμηση

5.3.2 Δοκιμές Γραμμικής Παλινδρόμησης

Στην προσπάθεια εξαγωγής μοντέλων γραμμικής παλινδρόμησης πραγματοποιήθηκαν αρκετές δοκιμές με σκοπό την **αύξηση του R^2** και την ικανοποίηση των συνθηκών σημαντικότητας των μεταβλητών. Ωστόσο, η τιμή του R^2 παρέμενε αρκετά χαμηλή γύρω στο 0.044. Στη συνέχεια, δοκιμάστηκε αφαίρεση των διαδρομών που παρουσίαζαν μηδενικά απότομα συμβάντα. Έτσι, βελτιώθηκε το R^2 , όμως η βελτίωση θεωρήθηκε αμελητέα οπότε προτιμήθηκε η διατήρηση ολόκληρης της βάσης δεδομένων με σκοπό να μη χαθούν πληροφορίες-διαδρομές. Επόμενο βήμα ήταν η δοκιμή αφαίρεσης της σταθεράς με το σκεπτικό ότι θα βελτιωνόταν το μοντέλο. Πράγματι, το R^2 αυξήθηκε σημαντικά στα επίπεδα του $R^2 = 0.64$, όμως και αυτή η εναλλακτική απορρίφθηκε αφού τα μοντέλα με και χωρίς τη σταθερά δεν μπορούν να είναι συγκρίσιμα μεταξύ τους, όπως υποδεικνύεται και σε σχετική σημείωση του SPSS (εικόνα 5.6).

b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

Εικόνα 5.6.: Σημείωση SPSS για την αφαίρεση της σταθεράς

Σε επόμενη φάση εξετάστηκαν γραμμικά μοντέλα ξεχωριστά ανά τύπο οδού, όμως ούτε σε αυτήν την περίπτωση αυξήθηκε το R^2 σε ανεκτά επίπεδα, εφόσον συνέχισε να κυμαίνεται γύρω από το 0.025. Σημειώνεται ότι η μεταβλητή *driving_daylight* απορρίφθηκε από τις δοκιμές των μοντέλων καθώς διαπιστώθηκε ότι υποκαθιστά τη σταθερά μιας και το μεγαλύτερο μέρος των παρατηρήσεων της είναι το ίδιο και συγκεκριμένα ίσο με την τιμή 1.

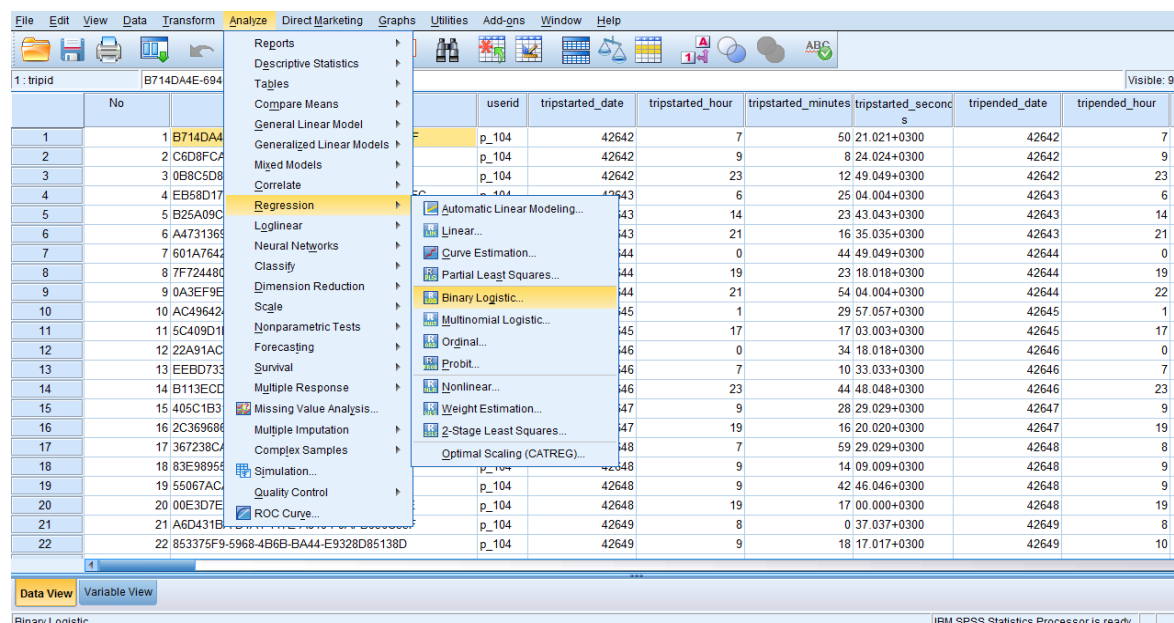
5.4 Λογιστική παλινδρόμηση

5.4.1 Διαδικασία Λογιστικής παλινδρόμησης

Για την πρόβλεψη της χρήσης ή μη χρήσης κινητού τηλεφώνου δοκιμάστηκε σε πρώτη φάση η μέθοδος της δυαδικής λογιστικής παλινδρόμησης. Αυτή η επιλογή εξηγείται από το γεγονός ότι η μεταβλητή που εξετάζεται (*mobileUsage/YES-NO*) είναι κατηγορική και παίρνει δύο τιμές, 0 και 1 (0 όταν δεν μιλάει στο κινητό, 1 όταν μιλάει έστω και ένα δευτερόλεπτο στο κινητό). Τα βήματα που ακολουθούνται στο SPSS για αυτήν την διαδικασία είναι τα εξής, όπως απεικονίζονται και στις εικόνες παρακάτω:

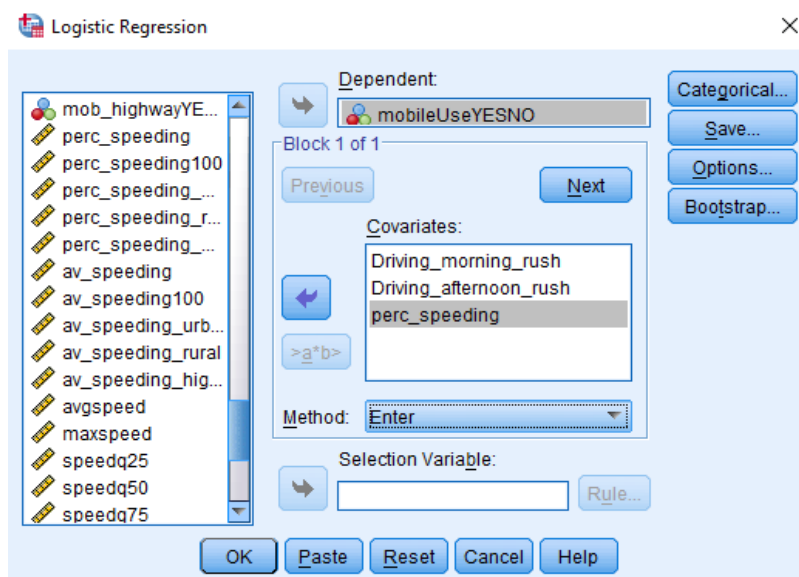
analyze → regression → binary logistic

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



Εικόνα 5.7.: Επιλογή εντολής δυαδικής λογιστικής παλινδρόμησης

Στη συνέχεια καθορίζονται η εξαρτημένη και οι ανεξάρτητες μεταβλητές και εισάγονται αντίστοιχα στα πεδία Dependent και Covariates. Στο πλαίσιο Method μπορεί να επιλεγεί μια μέθοδος για τη βέλτιστη επιλογή επεξηγηματικών μεταβλητών. Αυτή συνήθως αφήνεται «Enter», που σημαίνει ότι στο μοντέλο εισέρχονται όλες μεταβλητές βρίσκονται στο πλαίσιο Covariates με τη σειρά που γράφονται εκεί. Σταδιακά εισάγονται πιθανές ανεξάρτητες (επεξηγηματικές) μεταβλητές έως ότου καταλήξουμε σε επιθυμητά αποτελέσματα.

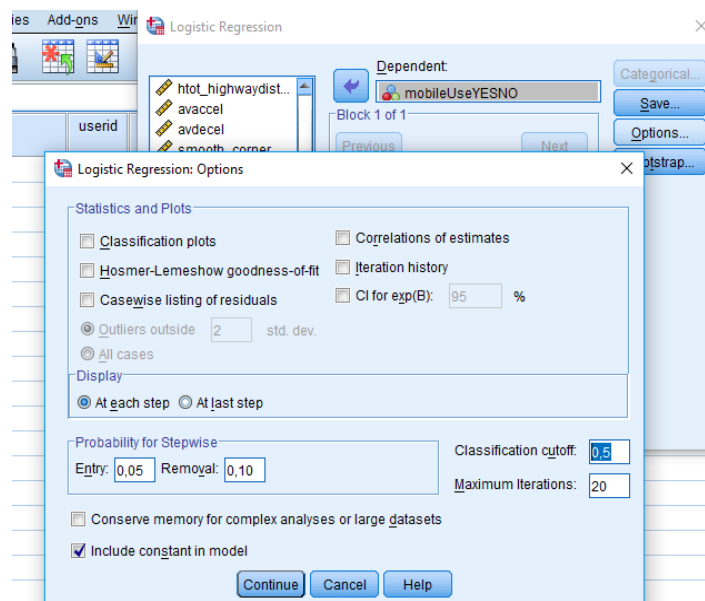


Εικόνα 5.8.: Επιλογή ανεξάρτητων και εξαρτημένων μεταβλητών για τη δυαδική λογιστική παλινδρόμηση

5.4.2 Δοκιμές Δυναμικής Λογιστικής Παλινδρόμησης

Στην προσπάθεια εξαγωγής μοντέλων που θα προβλέπουν τη χρήση ή μη χρήσης του κινητού τηλεφώνου μέσα από την παρατήρηση οδηγικών μεγεθών πραγματοποιήθηκαν αρκετές δοκιμές με κύριο γνώμονα τη βελτιστοποίηση των αποτελεσμάτων του πίνακα ταξινόμησης (**classification table**), την ικανοποίηση των συνθηκών σημαντικότητας των μεταβλητών και την κατά το δυνατόν αύξηση του Nagelkerke R^2 . Συγκεκριμένα, το classification table δείχνει σε τι ποσοστό ταυτίζονται οι παρατηρηθείσες τιμές της εξαρτημένης μεταβλητής με τις προβλεφθείσες και έτσι αξιολογείται η αξιοπιστία του μοντέλου. Επιδιώκονται τιμές πάνω από 60% τόσο για την πρόβλεψη της χρήσης όσο και για την πρόβλεψη της μη χρήσης κινητού τηλεφώνου. Η τιμή του R^2 στη λογιστική παλινδρόμηση δεν είναι τόσο εύχρηστη όσο στη γραμμική παλινδρόμηση, οπότε διαδραματίζει δευτερεύοντα ρόλο ως κριτήριο αξιολόγησης των διαφορών μοντέλων.

Με σκοπό την **αύξηση των ποσοστών του classification table** επιχειρήθηκε αφαίρεση των διαδρομών με μηδενικό ποσοστό χρήσης κινητού τηλεφώνου ($mobileUsage=0$) και επομένως δημιουργήθηκε μεταβλητή ($mobileUsage/MORE-LESS$) που θεωρεί δεδομένη τη χρήση του κινητού και εξετάζει αν η χρήση είναι μικρή ή μεγάλη. Ως όριο καθορισμού μεταξύ μικρής και μεγάλης χρήσης τέθηκαν διάφορες τιμές του ποσοστού χρόνου χρήσης κινητού (π.χ. $mobileUsage=0.05, 0.1, 0.15$ κ.ο.κ.). Ωστόσο, παρατηρήθηκε μόνο μικρή βελτίωση των ποσοστών του classification table και δεν θεωρήθηκε σκόπιμη η μείωση των διαδρομών της βάσης δεδομένων, οπότε διατηρήθηκε η μεταβλητή $mobileUsage/YES-NO$. Στη συνέχεια, και προκειμένου να βελτιωθεί το classification table αυξομειώνουμε την τιμή του classification cut-off, η οποία είναι προκαθορισμένη στο 0.5, έως ότου φτάσουμε σε επιθυμητά αποτελέσματα. Με άλλα λόγια, ορίζεται κάθε φορά το όριο της πιθανότητας το οποίο διαχωρίζει την πρόβλεψη του 0 ή του 1. Η αλλαγή της τιμής του cut-off απεικονίζεται στην επόμενη εικόνα 5.9 και γίνεται επιλέγοντας την εντολή options.



Εικόνα 5.9.: Τρόπος αλλαγής της τιμής cut-off στο SPSS

5.5 Μοντέλα Μεικτών Επιδράσεων ή Μεικτά Μοντέλα (Mixed Models)

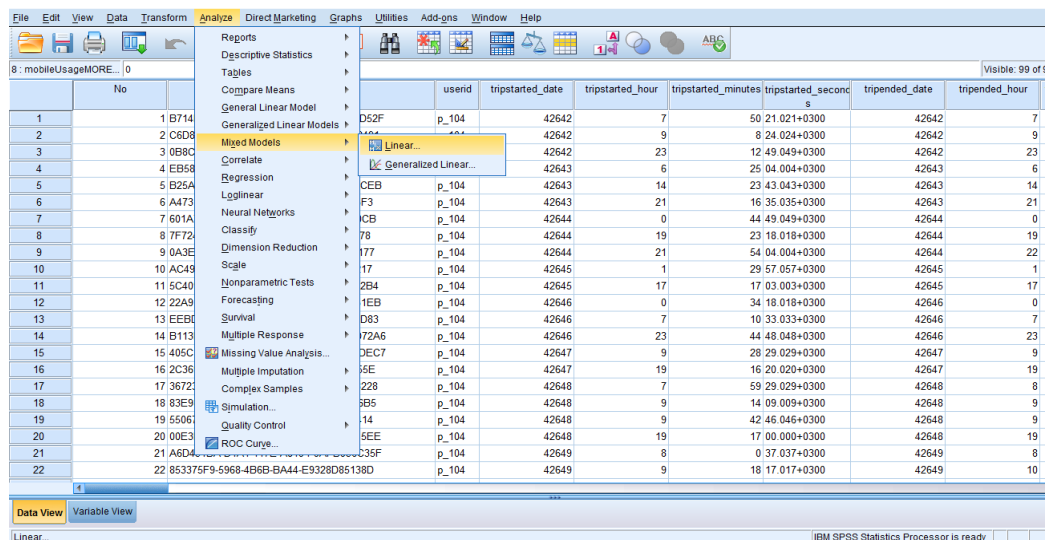
Στη βάση δεδομένων μας παρατηρείται το γεγονός ότι πολλές **διαδρομές** (trips) αναφέρονται στον ίδιο οδηγό-χρήστη (**userid**), άρα πρόκειται για επαναλαμβανόμενες μετρήσεις (**repeated measures**). Ως εκ τούτου οι συγκεκριμένες μετρήσεις είναι συσχετισμένες παραβαίνοντας την υπόθεση ανεξαρτησίας των παρατηρήσεων του δείγματος και έτσι η στατιστική ανάλυση ενδέχεται να παρουσιάσει αρκετά προβλήματα οδηγώντας σε στατιστικά μη έγκυρα ευρήματα. Για τον σκοπό αυτό αποφασίστηκε τόσο στη γραμμική όσο και στη δυαδική λογιστική παλινδρόμηση να θεωρηθεί ότι οι μεταβλητές επιδρούν και με σταθερό (**fixed**) και με τυχαίο (**random**) τρόπο και επομένως να αναζητηθούν mixed μοντέλα για τη διερεύνηση των στόχων της Διπλωματικής. Τα μοντέλα μεικτών επιδράσεων αναπτύχθηκαν στο κεφάλαιο 3. Σημειώνεται ότι η τυχαία επίδραση (random effect) αποδόθηκε στη σταθερά ενώ οι μεταβλητές θεωρήθηκε ότι έχουν σταθερή επίδραση (fixed effect).

5.6 Mixed Linear Model

5.6.1 Διαδικασία Mixed Linear Model

Τα βήματα που εφαρμόζονται μέσω του προγράμματος SPSS είναι τα εξής, όπως απεικονίζονται και στη σχετική εικόνα:

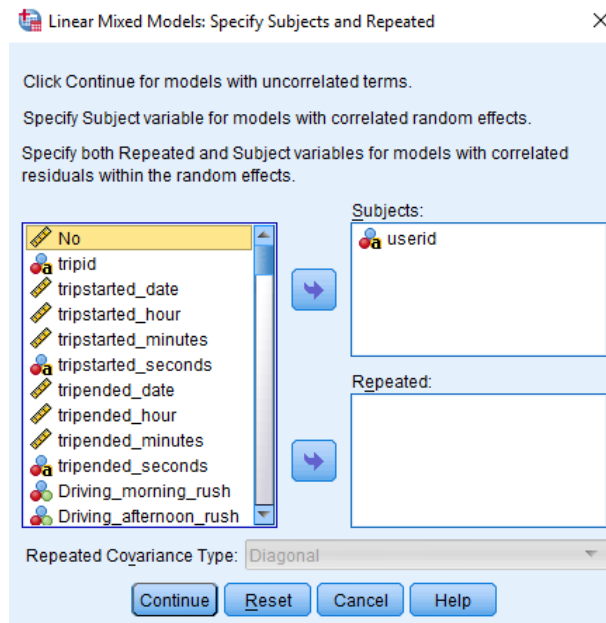
analyze → mixed models → linear



Εικόνα 5.10.:Επιλογή εντολής μεικτής γραμμικής παλινδρόμησης

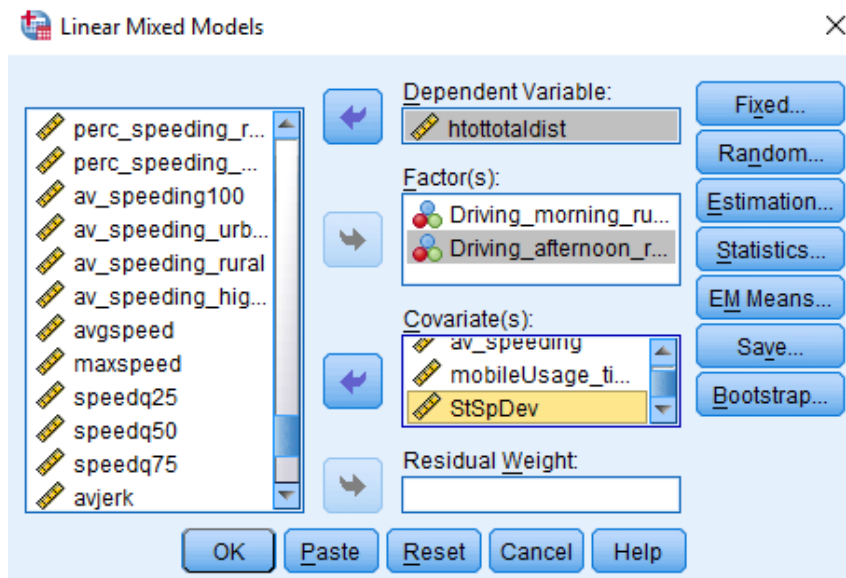
Στη συνέχεια ορίζεται το υποκείμενο για το οποίο έχουμε επαναλαμβανόμενες μετρήσεις, δηλαδή για κάθε οδηγό (userid).

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



Εικόνα 5.11.: Επιλογή υποκειμένου επαναλαμβανόμενων μετρήσεων

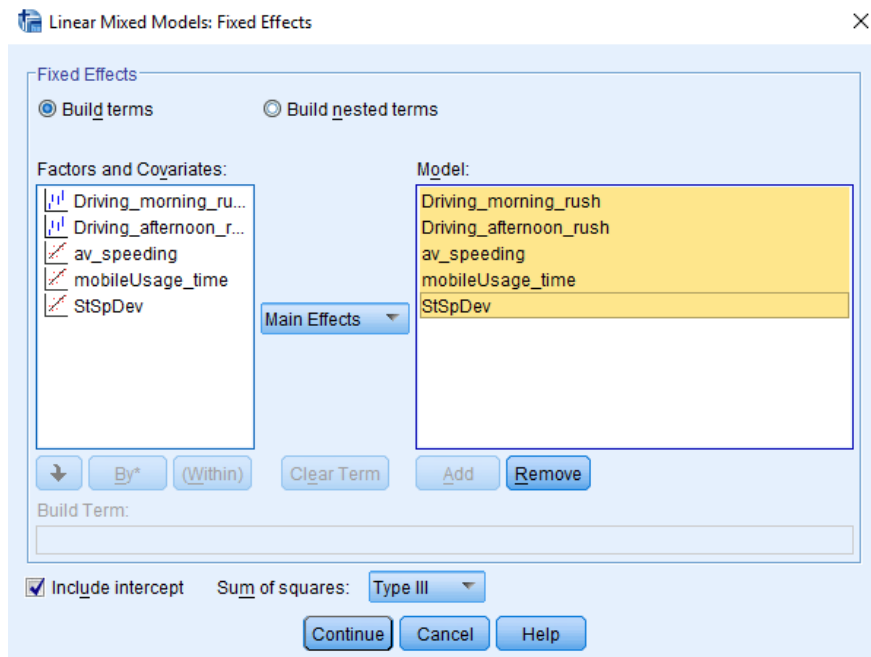
Στη συνέχεια ορίζεται η εξαρτημένη μεταβλητή, εδώ για το συνολικό αριθμό των απότομων συμβάντων από όλους τους τύπους οδού (htot/totaldist), στο πεδίο Dependent Variable. Οι κατηγορικές ανεξάρτητες μεταβλητές εισάγονται στο πεδίο Factor(s) και οι συνεχείς στο πεδίο Covariate(s), όπως απεικονίζεται στην εικόνα 5.12.



Εικόνα 5.12.: Επιλογή εξαρτημένων και ανεξάρτητων μεταβλητών για τη μεικτή γραμμική παλινδρόμηση

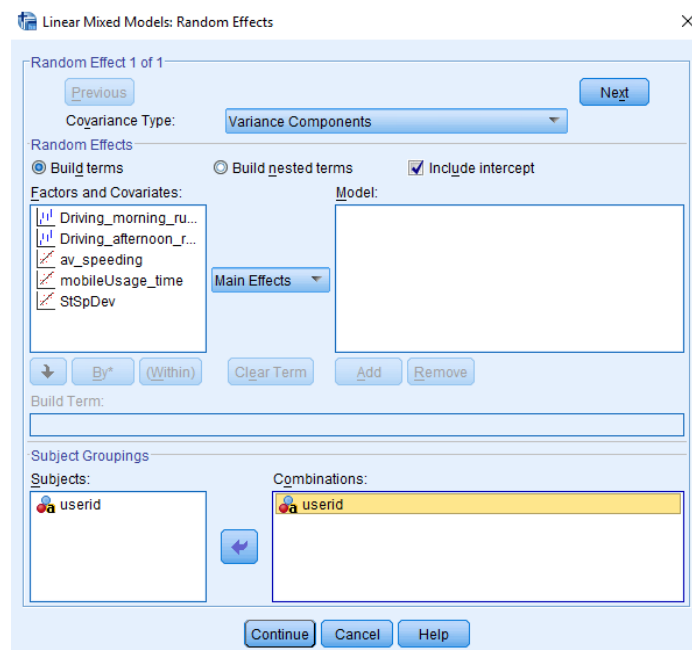
Έπειτα καθορίζεται για ποιες **μεταβλητές** δε θα επιτραπεί random effect αλλά **μόνο fixed effect**. Οι μεταβλητές αυτές τοποθετούνται στο πεδίο "model".

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



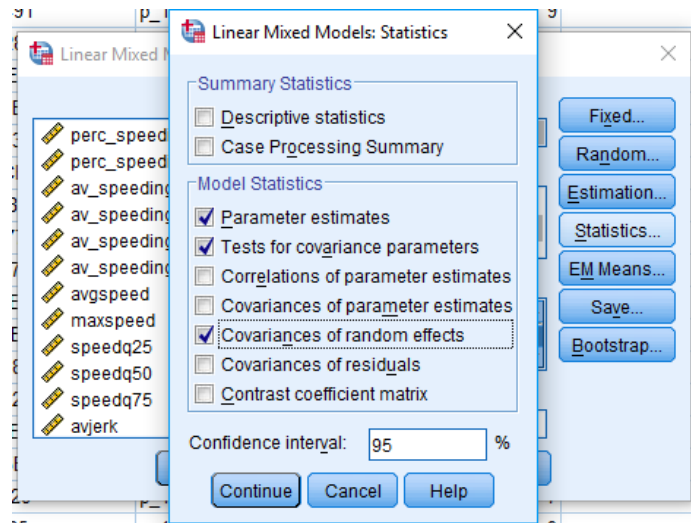
Εικόνα 5.13: Καθορισμός μεταβλητών με fixed effect

Επιλέγεται η εντολή random και στη συνέχεια η εντολή include intercept, οπότε το **random effect** αποδίδεται μόνο **στη σταθερά** και όχι στις υπόλοιπες μεταβλητές, οι οποίες δεν επιλέγεται να τοποθετηθούν στο πεδίο “model”. Τοποθετείται το userid στο πεδίο combinations και έτσι υπολογίζεται διαφορετική σταθερά για κάθε οδηγό.



Εικόνα 5.14.: Καθορισμός random effect

Τέλος, επιλέγοντας την εντολή statistics καθορίζονται ποια μεγέθη θα εμφανιστούν στα αποτελέσματα.



Εικόνα 5.15.: Επιλογή στατιστικών ελέγχων

5.6.2 Αποτελέσματα Mixed Linear Regression

5.6.2.1 Σύνολο των απότομων συμβάντων σε όλους τους τύπους οδού

Αποτελέσματα για το σύνολο των απότομων συμβάντων σε όλους τους τύπους οδού

Πραγματοποιήθηκαν αρκετές διαδοχικές δοκιμές. Οι δοκιμές έγιναν με συστηματικό τρόπο. Συγκεκριμένα, δεν δοκιμάστηκαν μεταβλητές όπου ήταν συσχετισμένες σε μεγάλο βαθμό μεταξύ τους ούτε μεταβλητές που ήταν παρόμοιες με την εξαρτημένη. Τελικώς, προέκυψαν τα εξής αποτελέσματα:

Model Dimension ^a					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Variance Components	1	userid
	Driving_morning_rush	2		1	
	Driving_afternoon_rush	2		1	
	av_speeding	1		1	
	mobileUsage_time	1		1	
	avgspeed	1		1	
	Intercept	1		1	
Random Effects					
Residual				1	
Total		9		8	

a. Dependent Variable: htottotaldist.

Εικόνα 5.16.: Χαρακτηριστικά του μοντέλου μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Η επιλογή variance components ορίζει τη δομή της συνδιακύμανσης του πίνακα των random effects, είναι η προκαθορισμένη από το SPSS και αφήνεται ίδια. Έχει την ακόλουθη μορφή:

$$\begin{bmatrix} \sigma_A^2 & 0 & 0 & 0 \\ 0 & \sigma_B^2 & 0 & 0 \\ 0 & 0 & \sigma_{AB}^2 & 0 \\ 0 & 0 & 0 & \sigma_{AB}^2 \end{bmatrix}$$

Από τον παρακάτω πίνακα ενδιαφέρει η τιμή του **Log Likelihood**. Πραγματοποιείται έλεγχος χ^2 στη διαφορά των λογαρίθμων των συναρτήσεων πιθανοφάνειας για το μοντέλο με τις παραπάνω πέντε μεταβλητές και για το «άδειο» μοντέλο (null model). Η τιμή του Likelihood-ratio test (LRT) υπολογίστηκε σε 754.22 που είναι σημαντική για διάστημα εμπιστοσύνης 95%. Το μοντέλο με τις παραπάνω πέντε μεταβλητές συγκρίθηκε και με μοντέλα με διαφορετικές μεταβλητές. Εν γένει, επιλέγονται μοντέλα με βάση την μικρότερη τιμή Log Likelihood.

Information Criteria	
-2 Restricted Log Likelihood	12779.525
Akaike's Information Criterion (AIC)	12783.525
Hurvich and Tsai's Criterion (AICC)	12783.525
Bozdogan's Criterion (CAIC)	12801.213
Schwarz's Bayesian Criterion (BIC)	12799.213

Εικόνα 5.17.: Κριτήρια ποιότητας μοντέλου μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού

Στον επόμενο πίνακα παρουσιάζονται οι συντελεστές των fixed effects και η σημαντικότητά τους όπως προκύπτει από τα t-ratios.

Estimates of Fixed Effects					
Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.517196	.018199	224.755	28.419	.000
[Driving_morning_rush=0]	-.017937	.006963	18814.501	-2.576	.010
[Driving_morning_rush=1]	0	0			
[Driving_afternoon_rush=0]	.029942	.006001	18788.543	4.989	.000
[Driving_afternoon_rush=1]	0	0			
av_speeding	.294491	.022223	18836.900	13.251	.000
mobileUsage_time	-3.365870E-05	1.299912E-05	18845.800	-2.589	.010
avgspeed	-.004749	.000174	18846.730	-27.359	.000

Εικόνα 5.18.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού

Έτσι προκύπτει το μοντέλο:

$$\text{htot/totaldist} = 0.52 - 0.018 * \text{driving_morning_rush} + 0.03 * \text{driving_afternoon_rush} + 0.29 * \text{av_speeding} - 3.37 * 10^{-5} * \text{mobileUsage_time} - 0.005 * \text{avgspeed}$$

Ο παρακάτω πίνακας παρουσιάζει την **τυχαία διακύμανση του μοντέλου**, καθώς και τη σημαντικότητά του (p-value) πραγματοποιώντας Wald-tests.

Estimates of Covariance Parameters				
Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	.112901	.001166	96.821	0.000
Intercept [subject = userid] Variance	.021095	.003113	6.776	.000

Εικόνα 5.19.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για όλους τους τύπους οδού

Κάθε οδηγός έχει δικό του random effect και επομένως δική του σταθερά. Η κατανομή των τιμών των διαφόρων σταθερών έχει διακύμανση ίση με 0.021. Η επίδραση του random effect για κάθε οδηγό μέσω της σταθεράς του είναι σημαντική αφού στον παραπάνω πίνακα ισχύει για τη σταθερά **p-value < 0.05**.

Σχολιασμός αποτελεσμάτων για το σύνολο των απότομων συμβάντων

- ✓ Η επίδραση του fixed και random effect είναι σημαντική όπως προκύπτει από τα **p-value < 0.05** τόσο για καθεμία από τις fixed μεταβλητές όσο και για την τυχαία (random) διακύμανση.
- ✓ Η τιμή του **Log Likelihood** είναι μικρότερη σε σχέση με αυτή άλλων μοντέλων που δοκιμάστηκαν.
- ✓ Τα **κατάλοιπα** του μοντέλου ελέγχθηκαν και κρίθηκαν ικανοποιητικά (Παράρτημα).
- ✓ Η μεταβλητή **"driving_morning_rush"** έχει αρνητικό συντελεστή που σημαίνει ότι κατά την πρωινή αιχμή είναι λιγότερα τα απότομα συμβάντα στην οδήγηση.
- ✓ Η μεταβλητή **"driving_afternoon_rush"** έχει θετικό συντελεστή που σημαίνει ότι κατά την απογευματινή αιχμή αυξάνονται τα απότομα συμβάντα στην οδήγηση σε σχέση με τις ώρες εκτός απογευματινής αιχμής.
- ✓ Το **"av_speeding"** έχει θετικό συντελεστή άρα όσο αυξάνεται η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας τόσο αυξάνονται τα απότομα συμβάντα, γεγονός το οποίο ερμηνεύεται λογικά (παράγραφος 6.2).
- ✓ Η μεταβλητή **"avgspeed"** έχει αρνητικό συντελεστή που σημαίνει ότι όσο αυξάνεται η ταχύτητα οδήγησης τόσο μειώνονται τα απότομα συμβάντα. Αυτό εξηγείται λογικά καθώς στο υπεραστικό δίκτυο και στους αυτοκινητοδρόμους στους οποίους αναπτύσσονται υψηλές ταχύτητες υπάρχουν λιγότερες

απότομες επιταχύνσεις, επιβραδύνσεις και ελιγμοί σε σχέση με το περιβάλλον εντός πόλης όπου υπάρχουν περισσότεροι κόμβοι, πεζοί κλπ. (παράγραφος 6.2).

- ✓ Η μεταβλητή "**mobileUsage_time**" έχει αρνητικό συντελεστή που σημαίνει ότι όσο αυξάνεται ο χρόνος ομιλίας στο τηλέφωνο τόσο μειώνονται τα απότομα συμβάντα ανά απόσταση. Αυτό εξηγείται από το γεγονός ότι κατά τη διάρκεια χρήσης του κινητού τηλεφώνου, ο οδηγός προβαίνει σε αντισταθμιστική συμπεριφορά ώστε να μειώσει τον κίνδυνο ατυχήματος (παράγραφος 6.2). Η τιμή του συντελεστή αυτού είναι πολύ μικρή εξαιτίας της μονάδας μέτρησης της μεταβλητής σε δευτερόλεπτα.

Σχετική επιρροή των μεταβλητών

Στο σημείο αυτό κρίθηκε σημαντικός ο υπολογισμός του βαθμού επιρροής των ανεξάρτητων μεταβλητών του παραπάνω μοντέλου στην εξαρτημένη μεταβλητή, δηλαδή αυτή του δείκτη συνολικής επίδοσης. **Ο βαθμός επιρροής των ανεξάρτητων μεταβλητών εκφράζεται ποσοτικά μέσω του μεγέθους της σχετικής επιρροής.** Ο υπολογισμός του μεγέθους αυτού βασίστηκε στη θεωρία της ελαστικότητας και αντικατοπτρίζει την ευαισθησία της εξαρτημένης μεταβλητής Y στην μεταβολή μιας ή περισσότερων ανεξάρτητων μεταβλητών (X_i). Η ελαστικότητα είναι ένα αδιάστατο μέγεθος, που σε αντίθεση με τους συντελεστές των μεταβλητών των μοντέλων, δεν εξαρτάται από τις μονάδες μέτρησης των μεταβλητών. Σε συνδυασμό με το πρόσημο των συντελεστών, είναι πιθανό να προσδιοριστεί αν η αύξηση κάποιας ανεξάρτητης μεταβλητής επιφέρει αύξηση ή μείωση στην εξαρτημένη. Είναι πολλές φορές ορθότερο να εκφραστεί η ευαισθησία ως ποσοστιαία μεταβολή της εξαρτημένης μεταβλητής που προκαλεί η κατά 1% μεταβολή της ανεξάρτητης. Η σχετική επιρροή των ανεξάρτητων μεταβλητών του μοντέλου, που αναπτύχθηκε, υπολογίστηκε σύμφωνα με τη σχέση:

$$e_i = \left(\frac{\Delta Y_i}{\Delta X_i} \right) * \left(\frac{X_i}{Y_i} \right) = \beta_i * \left(\frac{X_i}{Y_i} \right)$$

Ο προσδιορισμός της σχετικής επιρροής κάθε ανεξάρτητης μεταβλητής, αποδείχθηκε η πιο απλή και κατάλληλη τεχνική, ικανή να αναδείξει την επιρροή της κάθε μεταβλητής ξεχωριστά, αλλά και να καταστήσει εφικτή τη σύγκριση μεταξύ των επιρροών των διαφορετικών μεταβλητών του ίδιου μοντέλου. Αξίζει να σημειωθεί ότι η έννοια της επιρροής έχει νόημα μόνο για συνεχείς μεταβλητές και όχι για διακριτές μεταβλητές, αλλά στη στατιστική ανάλυση που πραγματοποιήθηκε στην παρούσα Διπλωματική Εργασία υπολογίστηκε η σχετική επιρροή και για τις διακριτές μεταβλητές ως μια θεωρητική έννοια, μόνο για να πραγματοποιηθεί θεωρητικά μια σύγκριση μεταξύ των μεταβλητών του μοντέλου σε ότι αφορά στην επιρροή τους στην εξαρτημένη μεταβλητή. Στη συνέχεια παρουσιάζεται σε μορφή πίνακα η σχετική επιρροή των μεταβλητών.

Πίνακας 5.1.: Ελαστικότητες και σχετική επιρροή μεταβλητών για το γραμμικό μοντέλο

	Ανεξάρτητες Μεταβλητές	Συντελεστές β	Ελαστικότητες	Σχετική Επιρροή ϵ_i
Συνεχείς	Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας	0.290	0.093	2.138
	Μέση ταχύτητα	-0.005	1.272	29.166
	Διάρκεια χρήσης κινητού τηλεφώνου	0.000	0.044	1.000
Διακριτές	Πρωινή αιχμή	-0.018	0.021	1.000
	Απογευματινή αιχμή	0.030	0.023	1.100

Στη στήλη ϵ_i δίνεται ο βαθμός της σχετικής επιρροής των ανεξάρτητων μεταβλητών ως προς την επιρροή εκείνης της μεταβλητής που επηρεάζει λιγότερο την εξαρτημένη μεταβλητή, δηλαδή ως προς την "driving_morning_rush" για τις διακριτές και ως προς την "mobileUsage_time" για τις συνεχείς μεταβλητές

- Η μεταβλητή **"avgspeed"** επηρεάζει περισσότερο από όλες τις υπόλοιπες μεταβλητές την εξαρτημένη μεταβλητή (κατά 29.17 φορές περισσότερο σε σχέση με τη μεταβλητή **"mobileUsage_time"**). Αυτό δείχνει το πόσο σημαντική είναι η ταχύτητα για τον αριθμό των απότομων συμβάντων που πραγματοποιούνται κατά την οδήγηση. Για αύξηση της "avgspeed" κατά 1%, η μεταβλητή "htot/totaldist" μειώνεται κατά 127.2%.
- Η μεταβλητή **"av_speeding"** επηρεάζει την εξαρτημένη μεταβλητή κατά 2.14 φορές περισσότερο σε σχέση με τη μεταβλητή "mobileUsage_time". Για αύξηση της "av_speeding" κατά 1%, η μεταβλητή "htot/totaldist" αυξάνεται κατά 9.3%.
- Για αύξηση της μεταβλητής **"mobileUsage_time"** κατά 1%, η μεταβλητή "htot/totaldist" μειώνεται κατά 4.4%.
- Η μεταβλητή **"driving_afternoon_rush"** παρουσιάζει κατά 1.1 φορές μεγαλύτερη επιρροή σε σχέση με την μεταβλητή "driving_morning_rush".

5.6.2.2 Απότομα συμβάντα σε αστικό περιβάλλον (urban)

Εξετάζεται η δημιουργία μοντέλου mixed γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την $htot_urban/distance_urban$. Πραγματοποιείται η εφαρμογή της διαδικασίας στο SPSS όπως περιγράφηκε στην παράγραφο 5.6.1. Στις επόμενες παραγράφους παρουσιάζονται τα αποτελέσματα και ο σχολιασμός για το συγκεκριμένο μοντέλο. Σημειώνεται ότι κατά τη διενέργεια δοκιμών, δεν εξετάστηκαν μεταβλητές που αφορούν το σύνολο των οδικών δικτύων, όπως η ταχύτητα (avgspeed, StSpDev), αλλά μόνο μεταβλητές που αφορούσαν τον συγκεκριμένο τύπο οδού.

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Αποτελέσματα για τα απότομα συμβάντα σε αστικό περιβάλλον

Η εξίσωση του μοντέλου που προέκυψε είναι η εξής:

$$\text{htot_urban/distance_urban} = 0.55 - 0.095 * \text{driving_morning_rush} - 0.0088 * \text{distance_urban} + 0.083 * \text{mobileUsage_urban} + 0.3 * \text{av_speeding_urban}$$

Model Dimension ^a					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Variance Components	1	userid
	Driving_morning_rush	2		1	
	av_speeding_urban	1		1	
	mobileUsage_urban	1		1	
	distance_urban	1		1	
Random Effects	Intercept ^b	1	Variance Components	1	userid
Residual				1	
Total		7		7	

a. Dependent Variable: htot_urbandistance_urban.

Εικόνα 5.20.: Χαρακτηριστικά του μοντέλου της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον

Information Criteria	
-2 Restricted Log Likelihood	32454.131
Akaike's Information Criterion (AIC)	32458.131
Hurvich and Tsai's Criterion (AICC)	32458.132
Bozdogan's Criterion (CAIC)	32475.820
Schwarz's Bayesian Criterion (BIC)	32473.820

Εικόνα 5.21.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον

Estimates of Fixed Effects					
Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.550616	.024676	151.924	22.314	.000
[Driving_morning_rush=0]	-.094756	.011269	18818.082	-8.409	.000
[Driving_morning_rush=1]	0 ^b	0			
av_speeding_urban	.301796	.026730	18841.859	11.291	.000
mobileUsage_urban	.082789	.029484	18824.815	2.808	.005
distance_urban	-.008794	.000905	18847.244	-9.721	.000

Εικόνα 5.22.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον

Estimates of Covariance Parameters				
Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	.321457	.003320	96.825	0.000
Intercept Variance	.046831	.006969	6.720	.000
[subject = userid]				

Εικόνα 5.23.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για αστικό περιβάλλον

Σχολιασμός αποτελεσμάτων

- ✓ Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως φαίνεται στον πίνακα με τα t-tests (Εικόνα 5.24) όπου το **p-value < 0.05**. Επομένως απορρίπτεται η μηδενική υπόθεση ότι τα fixed effects δεν είναι σημαντικά.
- ✓ Το **random effect** είναι επίσης σημαντικό όπως απεικονίζεται στην εικόνα 5.23 μετά από πραγματοποίηση του Wald-test, εφόσον p-value < 0.05.
- ✓ Η διακύμανση (random effect) ισούται με 0.047.
- ✓ Η μεταβλητή "**driving_morning_rush**" έχει συντελεστή ίσο με -0.095 που σημαίνει ότι κατά τη διάρκεια της πρωινής αιχμής μειώνονται τα απότομα συμβάντα σε σχέση με τις ώρες εκτός πρωινής αιχμής.
- ✓ Η μεταβλητή "**distance_urban**" έχει συντελεστή -0.0088 που σημαίνει ότι όσο αυξάνεται η απόσταση οδήγησης σε αστικό περιβάλλον τόσο μειώνονται τα απότομα συμβάντα. Αυτό εξηγείται καθώς ο οδηγός θα εξοικειωθεί με τις συνθήκες του περιβάλλοντος (παράγραφος 6.2).

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

- ✓ Η μεταβλητή **"mobileUsage_urban"** έχει θετικό συντελεστή 0.083 που σημαίνει ότι όσο αυξάνεται το ποσοστό χρόνου χρήσης του κινητού τόσο αυξάνονται τα απότομα συμβάντα ανά απόσταση.
- ✓ Η μεταβλητή **"av_speeding_urban"** έχει θετικό συντελεστή 0.3 οπότε όσο αυξάνεται η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας στο αστικό περιβάλλον τόσο αυξάνονται τα απότομα συμβάντα.

5.6.2.3 Απότομα συμβάντα σε υπεραστικό περιβάλλον

Εξετάζεται η δημιουργία μοντέλου mixedγραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την `htot_rural/distance_rural`. Πραγματοποιείται η εφαρμογή της διαδικασίας στο SPSS όπως περιγράφηκε στην παράγραφο 5.6.1. Στις επόμενες παραγράφους παρουσιάζονται τα αποτελέσματα και ο σχολιασμός για το συγκεκριμένο μοντέλο. Σημειώνεται ότι κατά τη διενέργεια δοκιμών, δεν εξετάστηκαν μεταβλητές που αφορούν το σύνολο των οδικών δικτύων, όπως η ταχύτητα (`avgspeed`, `StSpDev`), αλλά μόνο μεταβλητές που αφορούσαν στον συγκεκριμένο τύπο οδού.

Αποτελέσματα για τα απότομα συμβάντα σε υπεραστικό περιβάλλον

Το μοντέλο που προέκυψε περιγράφεται από την εξής εξίσωση:

$$\text{htot_rural/distance_rural} = 0.32 - 0.05 * \text{driving_morning_rush} - 0.092 * \text{perc_speeding_rural} + 0.14 * \text{mobileUsage_rural} - 0.0011 * \text{distance_rural}$$

Model Dimension ^a					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Driving_morning_rush	2		1	
	perc_speeding_rural	1		1	
	mobileUsage_rural	1		1	
	distance_rural	1		1	
Random Effects	Intercept ^b	1	Variance Components	1	userid
Residual				1	
Total		7		7	

a. Dependent Variable: `htot_rural/distance_rural`.

Εικόνα 5.24.: Χαρακτηριστικά του μοντέλου της γραμμικής παλινδρόμησης μεικτών επιδράσεων για υπεραστικό περιβάλλον

Information Criteria ^a	
-2 Restricted Log Likelihood	30801.621
Akaike's Information Criterion (AIC)	30805.621
Hurvich and Tsai's Criterion (AICC)	30805.622
Bozdogan's Criterion (CAIC)	30823.310
Schwarz's Bayesian Criterion (BIC)	30821.310

Εικόνα 5.25.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον

Estimates of Fixed Effects					
Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.323355	.017213	203.617	18.786	.000
[Driving_morning_rush=0]	-.051331	.010793	18847.929	-4.756	.000
[Driving_morning_rush=1]	0 ^b	0			
perc_speeding_rural	-.092080	.025243	18639.348	-3.648	.000
mobileUsage_rural	.140284	.027701	18725.129	5.064	.000
distance_rural	-.001145	.000422	18847.739	-2.713	.007

Εικόνα 5.26.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον

Estimates of Covariance Parameters				
Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	.295659	.003053	96.830	0.000
Intercept Variance [subject = userid]	.018512	.002908	6.366	.000

Εικόνα 5.27.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για υπεραστικό περιβάλλον

Σχολιασμός αποτελεσμάτων

- ✓ Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως φαίνεται στον πίνακα με τα t-tests (Εικόνα 5.30) όπου το **p-value < 0.05**. Επομένως απορρίπτεται η μηδενική υπόθεση ότι τα fixed effects δεν είναι σημαντικά.
- ✓ Το **random effect** είναι επίσης σημαντικό όπως απεικονίζεται στην εικόνα 5.27 μετά από πραγματοποίηση του Wald-test, εφόσον p-value < 0.05.
- ✓ Η τυχαία διακύμανση ισούται με 0.019.
- ✓ Η μεταβλητή "**driving_morning_rush**" έχει αρνητικό συντελεστή ίσο με -0.05 που σημαίνει ότι κατά τη διάρκεια της πρωινής αιχμής μειώνονται τα απότομα συμβάντα σε σχέση με τις ώρες εκτός πρωινής αιχμής.
- ✓ Η μεταβλητή "**distance_rural**" έχει αρνητικό συντελεστή -0.0011 που σημαίνει ότι όσο αυξάνεται η απόσταση οδήγησης σε υπεραστικό περιβάλλον τόσο μειώνονται τα απότομα συμβάντα καθώς ο οδηγός θα συνηθίσει την οδήγηση εντός του συγκεκριμένου οδικού περιβάλλοντος (παράγραφος 6.2).
- ✓ Η μεταβλητή "**mobileUsage_rural**" έχει θετικό συντελεστή 0.14 που σημαίνει ότι όσο αυξάνεται το ποσοστό χρόνου χρήσης του κινητού τόσο αυξάνονται τα απότομα συμβάντα ανά απόσταση οδήγησης.
- ✓ Η μεταβλητή "**perc_speeding_rural**" έχει αρνητικό συντελεστή -0.092 οπότε όσο αυξάνεται το ποσοστό χρόνου πάνω από το όριο ταχύτητας στο υπεραστικό περιβάλλον τόσο μειώνονται τα απότομα συμβάντα ανά απόσταση.

5.6.2.4 Απότομα συμβάντα σε αυτοκινητόδρομο

Εξετάζεται η δημιουργία μοντέλου mixed γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την `htot_highway/distance_highway`. Πραγματοποιείται η εφαρμογή της διαδικασίας στο SPSS όπως περιγράφηκε στην παράγραφο 5.6.1. Στις επόμενες παραγράφους παρουσιάζονται τα αποτελέσματα και ο σχολιασμός για το συγκεκριμένο μοντέλο. Σημειώνεται ότι κατά τη διενέργεια δοκιμών, δεν εξετάστηκαν μεταβλητές που αφορούν το σύνολο των οδικών δικτύων, όπως η ταχύτητα (`avgspeed`, `StSpDev`), αλλά μόνο μεταβλητές που αφορούσαν τον συγκεκριμένο τύπο οδού.

Αποτελέσματα για τα απότομα συμβάντα σε αυτοκινητόδρομο

Το μοντέλο που προέκυψε περιγράφεται από την εξής εξίσωση:

$$\text{htot_highway/distance_highway} = 0.016 - 0.007 * \text{driving_morning_rush} - 0.000182 * \text{distance_highway} + 0.37 * \text{av_speeding_highway} + 0.000147 * \text{time_mobile_highway}$$

Model Dimension ^a					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Driving_morning_rush	2		1	
	distance_highway	1		1	
	av_speeding_highway	1		1	
	time_mobile_highway	1		1	
Random Effects	Intercept ^b	1	Variance Components	1	userid
Residual				1	
Total		7		7	

a. Dependent Variable: htot_highwaydistance_highway.

Εικόνα 5.28.: Χαρακτηριστικά του μοντέλου της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο

Information Criteria	
-2 Restricted Log Likelihood	-24856.418
Akaike's Information Criterion (AIC)	-24852.418
Hurvich and Tsai's Criterion (AICC)	-24852.418
Bozdogan's Criterion (CAIC)	-24834.730
Schwarz's Bayesian Criterion (BIC)	-24836.730

Εικόνα 5.29.: Κριτήρια ποιότητας μοντέλου της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Estimates of Fixed Effects					
Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.016202	.002858	391.033	5.668	.000
[Driving_morning_rush=0]	-.007432	.002459	18577.030	-3.022	.003
[Driving_morning_rush=1]	0 ^b	0			
distance_highway	-.000182	5.785310E-05	18841.875	-3.142	.002
av_speeding_highway	.374347	.047986	18085.639	7.801	.000
time_mobile_highway	.000147	1.880347E-05	18844.461	7.826	.000

Εικόνα 5.30.: Πίνακας με τους συντελεστές των ανεξάρτητων μεταβλητών της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο

Estimates of Covariance Parameters				
Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	.015503	.000160	96.828	0.000
Intercept Variance	.000285	5.545236E-05	5.134	.000
[subject = userid]				

Εικόνα 5.31.: Πίνακας εκτίμησης τυχαίας διακύμανσης της μεικτής γραμμικής παλινδρόμησης για αυτοκινητόδρομο

Σχολιασμός αποτελεσμάτων

- ✓ Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως φαίνεται στον πίνακα με τα t-tests (εικόνα 5.36) όπου το **p-value < 0.05**. Επομένως απορρίπτεται η μηδενική υπόθεση ότι τα fixed effects δεν είναι σημαντικά.
- ✓ Το **random effect** είναι επίσης σημαντικό όπως απεικονίζεται στην εικόνα 5.31 μετά από πραγματοποίηση του Wald-test, εφόσον p-value < 0.05.
- ✓ Η τυχαία διακύμανση της σταθεράς λόγω επαναλαμβανόμενων παρατηρήσεων ισούται με 0.00026.
- ✓ Η μεταβλητή "**driving_morning_rush**" έχει συντελεστή ίσο με -0.007 που σημαίνει ότι κατά τη διάρκεια εκτός της πρωινής αιχμής τα απότομα συμβάντα είναι αυξημένα σε σχέση με τις ώρες εντός πρωινής αιχμής.
- ✓ Η μεταβλητή "**distance_highway**" έχει αρνητικό συντελεστή -0.000182 που σημαίνει ότι όσο αυξάνεται η απόσταση οδήγησης σε αυτοκινητόδρομο τόσο

μειώνονται τα απότομα συμβάντα. Αυτό είναι λογικό καθώς ο οδηγός έχει περισσότερο χρόνο (απόσταση) για να προσαρμοστεί στις συνθήκες του συγκεκριμένου οδικού περιβάλλοντος.

- ✓ Η μεταβλητή **"time_mobile_highway"** έχει θετικό συντελεστή 0.000147 που σημαίνει ότι όσο αυξάνεται ο χρόνος χρήσης του κινητού τόσο αυξάνονται τα απότομα συμβάντα ανά απόσταση, όπως εξηγείται και στην παράγραφο 6.2.
- ✓ Η μεταβλητή **"av_speeding_highway"** έχει θετικό συντελεστή 0.37 οπότε όσο αυξάνεται η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας στον αυτοκινητόδρομο τόσο αυξάνονται τα απότομα συμβάντα ανά απόσταση. Αυτό είναι λογικό καθώς όσο μεγαλύτερη ταχύτητα έχει ο οδηγός τόσο περισσότερο εκτελεί απότομες κινήσεις προκειμένου να επαναφέρει την οδηγική του κατάσταση σε περίπτωση αναπάντεχου γεγονότος (παράγραφος 6.2).

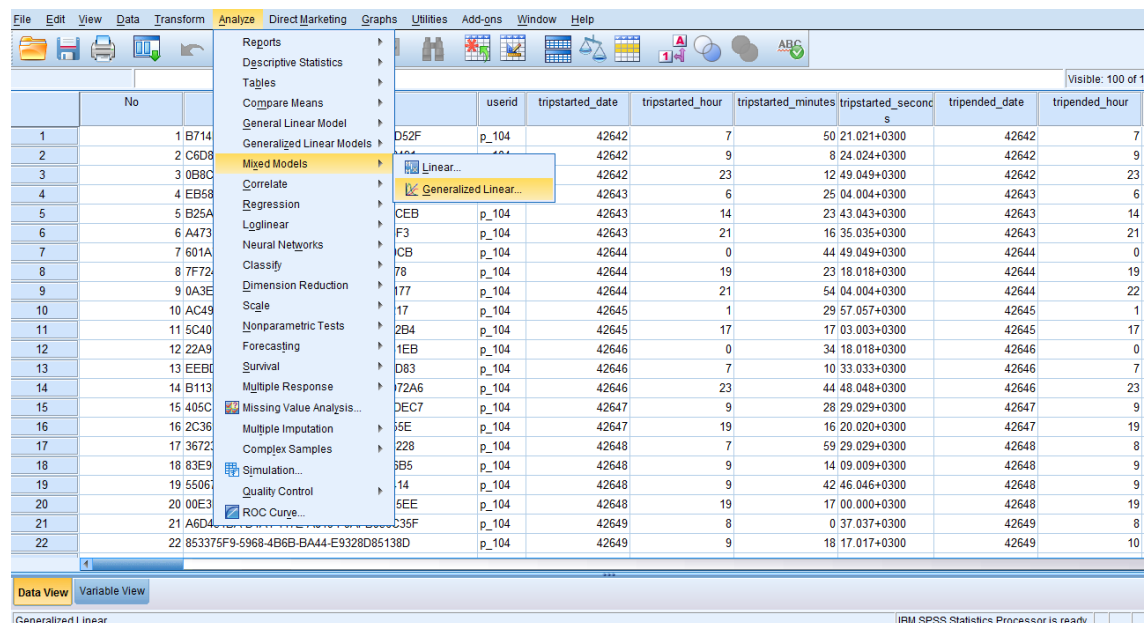
5.7 Mixed Binary Logistic Model

5.7.1 Διαδικασία Mixed Binary Logistic Model

Τα βήματα που ακολουθήθηκαν για την εξαγωγή του μοντέλου μεικτής δυαδικής λογιστικής παλινδρόμησης είναι τα εξής:

Analyze → mixed models → generalized linear

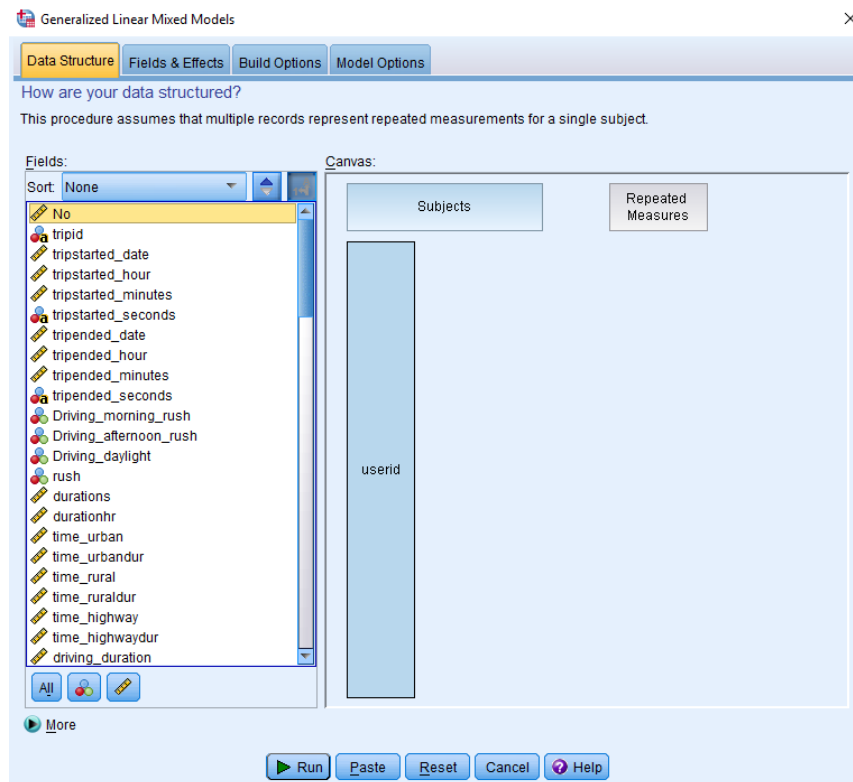
Το generalized linear μοντέλο είναι μια γενίκευση της συνηθισμένης γραμμικής παλινδρόμησης μόνο που η κατανομή των καταλοίπων- σφαλμάτων δεν είναι κανονική, όπως συμβαίνει στην απλή γραμμική παλινδρόμηση. Στην κατηγορία των generalized linear μοντέλων ανήκουν και τα μοντέλα λογιστικής παλινδρόμησης. Τα παραπάνω βήματα συνοψίζονται στην ακόλουθη εικόνα:



Εικόνα 5.32.: Επιλογή εντολής generalized mixed linear μοντέλου

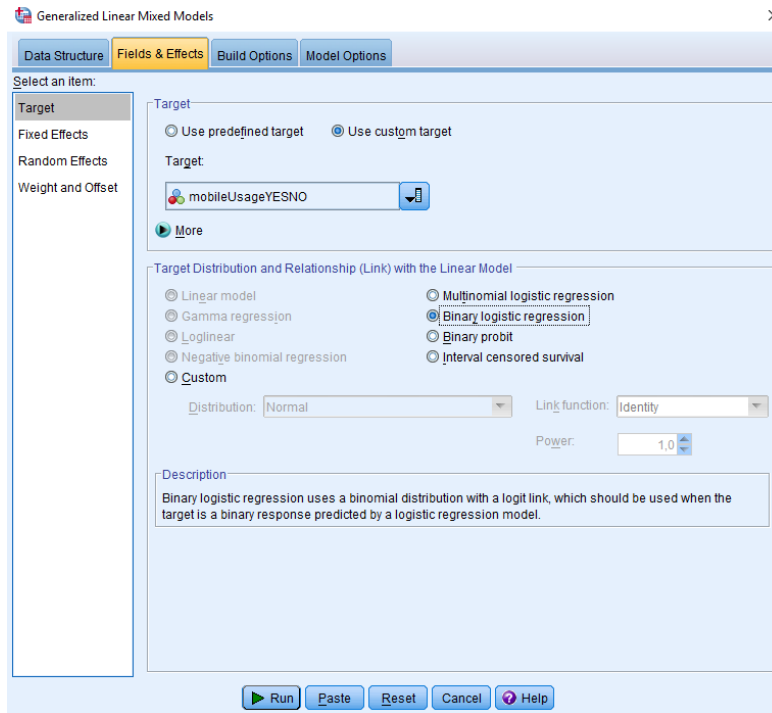
5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Στη συνέχεια, στο πεδίο subject ορίζεται το υποκείμενο για το οποίο έχουμε επαναλαμβανόμενες μετρήσεις, δηλαδή ο "κωδικός" του κάθε οδηγού userid.



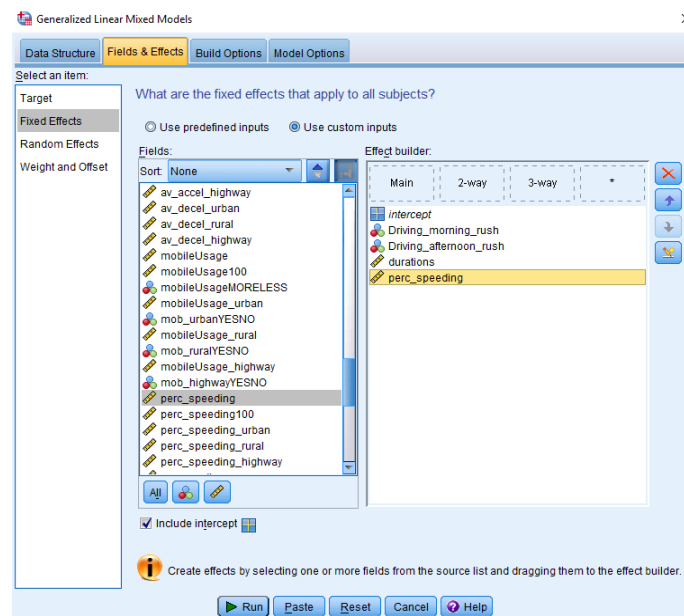
Εικόνα 5.33.: Επιλογή υποκειμένου επαναλαμβανόμενων μετρήσεων

Στο πεδίο Fields&Effects ορίζεται η εξαρτημένη μεταβλητή (target), εδώ το mobileUsage/YES-NO, καθώς και το ότι πρόκειται για μοντέλο δυαδικής λογιστικής παλινδρόμησης (Binary Logistic Regression).



Εικόνα 5.34.: Καθορισμός εξαρτημένης μεταβλητής και τύπου κατανομής της στη μεικτή δυαδική λογιστική παλινδρόμηση

Έπειτα, ορίζονται ποιες μεταβλητές θα έχουν fixed effect (Εικόνα 5.35).

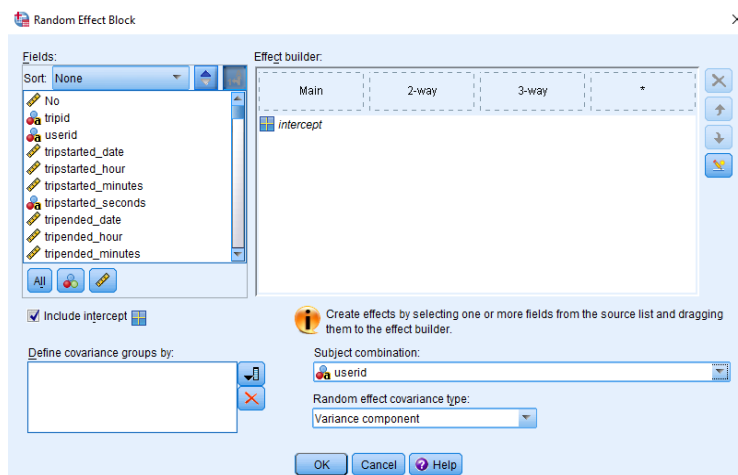


Εικόνα 5.35.: Επιλογή μεταβλητών που έχουν fixed effects στη μεικτή δυαδική λογιστική παλινδρόμηση

Τέλος, καθορίζεται ότι το random effect θα ληφθεί από τη σταθερά του μοντέλου, οπότε επιλέγεται η εντολή include intercept, καθώς και ότι για κάθε οδηγό θα υπολογιστεί ξεχωριστή σταθερά άρα ξεχωριστό random effect, οπότε επιλέγεται η μεταβλητή userid στο subject combination. Ο τύπος διακύμανσης του random effect αφήνεται στο

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

προκαθορισμένο Variance component. Στην παράγραφο 5.6.2.1 έχει αναφερθεί ο τρόπος ορισμού αυτού του τύπου διακύμανσης.



Εικόνα 5.36. Καθορισμός πεδίου random effect στη μεικτή δυαδική λογιστική παλινδρόμηση

5.7.2 Αποτελέσματα Mixed Binary Logistic Regression

5.7.2.1 Χρήση ή μη κινητού τηλεφώνου σε όλους τους τύπους οδού

Αποτελέσματα για το σύνολο των απότομων συμβάντων από όλους τους τύπους οδού

Πραγματοποιήθηκαν δοκιμές με μεθοδικό τρόπο προσέχοντας οι επεξηγηματικές μεταβλητές να είναι κατά το δυνατόν ασυσχέτιστες μεταξύ τους. Προέκυψε η εξής συνάρτηση χρησιμότητας U ή αλλιώς Utility Function:

$$U = 1.094 - 0.13 * \text{driving_morning_rush} + 0.262 * \text{driving_afternoon_rush} - 0.064 * \text{htot} - 0.004 * \text{avgspeed} - 0.058 * \text{smooth_corner} - 0.334 * \text{av_speeding}$$

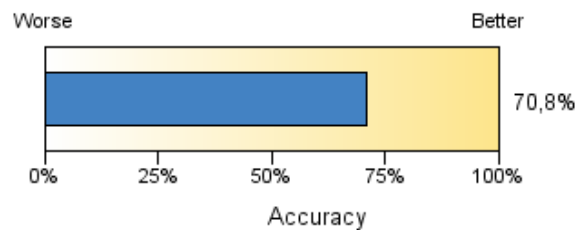
Πρώτα εμφανίζεται μια περίληψη των χαρακτηριστικών του μοντέλου. Η εικόνα 5.45 εμφανίζει τον στόχο- εξαρτημένη μεταβλητή (εδώ mobileUsage/YES-NO), την κατανομή (εδώ δυαδική) καθώς και το κριτήριο ποιότητας του μοντέλου που αξιοποιήθηκε (Log pseudolikelihood). Η τιμή Log pseudolikelihood του τελικού μοντέλου με τις έξι μεταβλητές είναι μικρότερη από την τιμή του αντίστοιχου κριτηρίου σχετικών μοντέλων (nested), καθώς και από το μοντέλο χωρίς καμία μεταβλητή, παρά μόνο με τη σταθερά (null model). Συγκεκριμένα, εξετάζεται η διαφορά των λογαρίθμων των συναρτήσεων πιθανοφάνειας του τελικού και του «άδειου» μοντέλου και προκύπτει Likelihood-ratio test (LRT) ίσο με 1555.862 το οποίο είναι σημαντικό για διάστημα εμπιστοσύνης 95%. Επίσης απεικονίζεται η **ακρίβεια** του μοντέλου η οποία εδώ φτάνει το 70.8%.

Model Summary

Target: mobileUsageYESNO

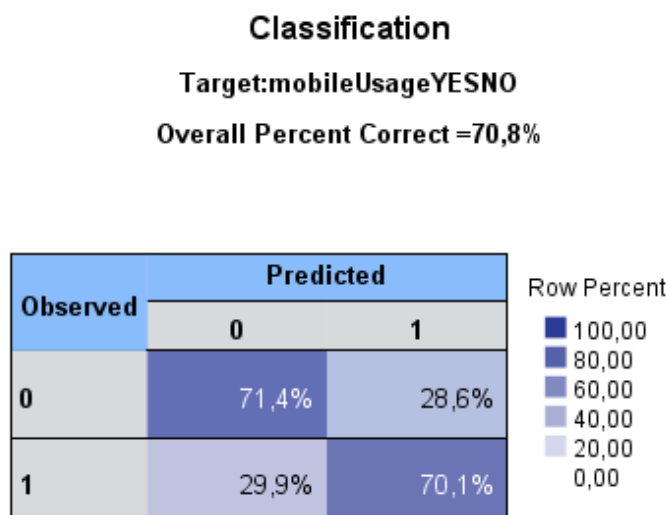
Target	mobileUsageYESNO
Probability Distribution	Binomial
Link Function	Logit
Akaike Corrected Information Criterion	86.301,336
Bayesian	86.309,179

Information criteria are based on the -2 log pseudo likelihood (86.299,335) and are used to compare models. Models with smaller information criterion values fit better. When comparing models using pseudo likelihood values, caution should be used because different data transformations may be used across the models.



Εικόνα 5.37.:Χαρακτηριστικά του μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού

Η επόμενη εικόνα πρόκειται για τον πίνακα ταξινόμησης (**classification table**). Μας ενδιαφέρουν κυρίως οι τιμές **True Positive** (TP) και **True Negative** (TN), δηλαδή ποιες τιμές ήταν 0 και προβλέφθηκαν ως 0 και ποιες τιμές ήταν 1 και προβλέφθηκαν ως 1. Έχουμε αντίστοιχα ποσοστά 71.4% και 70.1%, που θεωρούνται αρκετά ικανοποιητικά για την πρόβλεψη της χρήσης ή μη κινητού τηλεφώνου. Στη συνέχεια παρουσιάζεται επεξηγηματικός πίνακας του classification table. Η ευαισθησία (sensitivity) του μοντέλου μπορεί να εκτιμηθεί από τον δείκτη $TP / (Observed\ positive) = 71.4\% / 101.3\% = 0.7$ και η εξειδίκευση (specificity) του μοντέλου από τον δείκτη $TN / (Observed\ negative) = 70.1\% / 98.7\% = 0.71$.



Εικόνα 5.38.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού

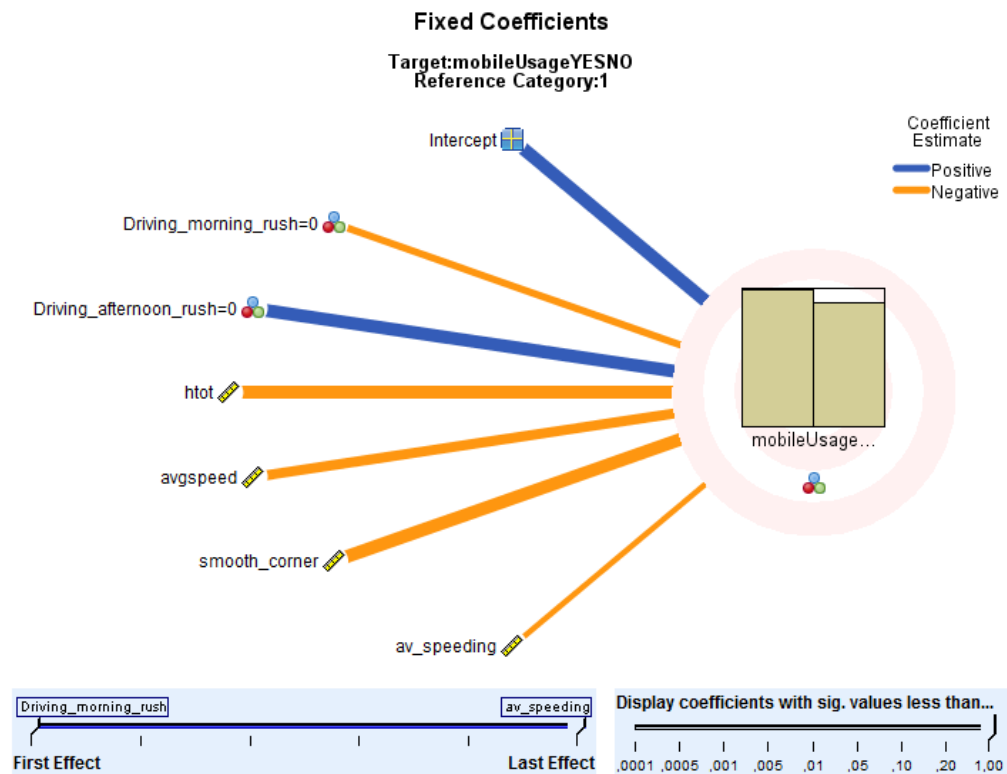
		predicted condition			
		prediction positive	prediction negative	Prevalence $= \frac{\Sigma \text{ condition positive}}{\Sigma \text{ total population}}$	
true condition	condition positive	True Positive (TP)	False Negative (FN) (type II error)	True Positive Rate (TPR), Sensitivity, Recall, Probability of Detection $= \frac{\Sigma TP}{\Sigma \text{ condition positive}}$	False Negative Rate (FNR), Miss Rate $= \frac{\Sigma FN}{\Sigma \text{ condition positive}}$
	condition negative	False Positive (FP) (Type I error)	True Negative (TN)	False Positive Rate (FPR), Fall-out, Probability of False Alarm $= \frac{\Sigma FP}{\Sigma \text{ condition negative}}$	True Negative Rate (TNR), Specificity (SPC) $= \frac{\Sigma TN}{\Sigma \text{ condition negative}}$
		Accuracy $= \frac{\Sigma TP + \Sigma TN}{\Sigma \text{ total population}}$	Positive Predictive Value (PPV), Precision $= \frac{\Sigma TP}{\Sigma \text{ prediction positive}}$	False Omission Rate (FOR) $= \frac{\Sigma FN}{\Sigma \text{ prediction negative}}$	Positive Likelihood Ratio (LR+) $= \frac{TPR}{FPR}$
		False Discovery Rate (FDR) $= \frac{\Sigma FP}{\Sigma \text{ prediction positive}}$	Negative Predictive Value (NPV) $= \frac{\Sigma TN}{\Sigma \text{ prediction negative}}$	Negative Likelihood Ratio (LR-) $= \frac{FNR}{TNR}$	Diagnostic Odds Ratio (DOR) $= \frac{LR+}{LR-}$

Εικόνα 5.39.: Επεξηγηματικός πίνακας για τον πίνακα ταξινόμησης
 Πηγή: Wikipedia

Η επόμενη εικόνα αποτελεί οπτικοποίηση του παραδοσιακού πίνακα με τα fixed effects. **Το πάχος της γραμμής έχει σχέση με τη σημαντικότητα του fixed effect κάθε μεταβλητής.** Όσο "μεγαλύτερη" είναι η σημαντικότητα μιας μεταβλητής τόσο πιο παχιά είναι η γραμμή. Έτσι για τη μεταβλητή "av_speeding" η γραμμή είναι η πιο λεπτή καθώς όπως παρουσιάζεται και στον πιο κάτω πίνακα έχει τη "μικρότερη" σημαντικότητα $p\text{-value} = 0.027$ (πιο κοντά στο 0.05), το οποίο για διάστημα εμπιστοσύνης 95%, όπως λήφθηκε εδώ, είναι αποδεκτό. Πορτοκαλί χρώμα σημαίνει ότι ο συντελεστής της συγκεκριμένης μεταβλητής είναι αρνητικός, ενώ αν είναι μπλε σημαίνει ότι είναι θετικός.

Έπειτα, σειρά έχει ο παραδοσιακός πίνακας σημαντικότητας των fixed effects όπως προκύπτει από τα t-tests και οι συντελεστές των μεταβλητών.

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα



Εικόνα 5.40.: Οπτικοποίηση υπολογισμού των παραμέτρων των *fixed effects* της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού

Model Term	Coefficient ▼	Std.Error	t	Sig.	Exp(Coefficient)	95% Confidence Interval for Exp (Coefficient)	
						Lower	Upper
Intercept	1,094	0,158	6,925	,000	2,986	2,191	4,069
Driving_morning_rush=0	-0,130	0,047	-2,744	,006	0,878	0,800	0,964
Driving_morning_rush=1	0,000 ^a						
Driving_afternoon_rush=0	0,262	0,041	6,419	,000	1,299	1,199	1,407
Driving_afternoon_rush=1	0,000 ^a						
htot	-0,064	0,005	-14,107	,000	0,938	0,930	0,946
avgspeed	-0,004	0,001	-3,390	,001	0,996	0,993	0,998
smooth_corner	-0,058	0,007	-8,085	,000	0,943	0,930	0,957
av_speeding	-0,334	0,151	-2,213	,027	0,716	0,532	0,963

Probability distribution: Binomial
Link function: Logit

^aThis coefficient is set to zero because it is redundant.

Εικόνα 5.41.: Πίνακας συντελεστών *fixed effects* της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού

Covariance Parameters

Target:mobileUsageYESNO

Covariance Parameters	Residual Effect	0
	Random Effects	1
Design Matrix Columns	Fixed Effects	9
	Random Effects	1 ^a
Common Subjects		100

Common subjects are based on the subject specifications for the residual and random effects and are used to chunk the data for better performance.

^aThis is the number of columns per common subject.

Random Effect	Estimate	Std.Error	Z	Sig.	95% Confidence Interval	
					Lower	Upper
Var(Intercept)	1,261	0,190	6,628	,000	0,938	1,695

Covariance Structure:Variance components
Subject Specification:userld

Εικόνα 5.42.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για όλους τους τύπους οδού

Σε αυτόν τον πίνακα παρουσιάζεται η τιμή της διακύμανσης (random effect). Εδώ έχει τιμή 1.261 και είναι σημαντική αφού $p\text{-value} < 0.05$ για διάστημα εμπιστοσύνης 95%.

Σχολιασμός αποτελεσμάτων

- Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως προκύπτει από τα t-tests αφού **p-value < 0.05** για διάστημα εμπιστοσύνης 95%.
- Το **random effect** ήταν σημαντικό αφού ισχύει και πάλι $p\text{-value} < 0.05$
- Η διακύμανση του random effect ισούται με 1.261.
- Η τιμή **Log pseudolikelihood** είναι μικρότερη συγκριτικά με σχετικά μοντέλα (nested models).
- Η μεταβλητή **"driving_morning_rush"** έχει συντελεστή -0.13 που σημαίνει ότι κατά την πρωινή αιχμή ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά -0.13. Ή αλλιώς ο λόγος σχετικών πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.13}$. Όπου αναφέρεται ο λόγος σχετικών πιθανοτήτων νοείται ο λόγος της πιθανότητας ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο προς την πιθανότητα να μη χρησιμοποιεί το κινητό τηλέφωνο.
- Η μεταβλητή **"driving_afternoon_rush"** έχει συντελεστή 0.262 που σημαίνει ότι κατά την απογευματινή αιχμή (οπότε θα ισχύει η τιμή 1 της ανεξάρτητης μεταβλητής), ο λογάριθμος του λόγου πιθανοτήτων αυξάνεται κατά 0.262. Στη λογιστική παλινδρόμηση η εξαρτημένη μεταβλητή ισούται με τον λογάριθμο του λόγου πιθανοτήτων οπότε στην ουσία αυξάνεται η εξαρτημένη μεταβλητή κατά 0.262.
- Η μεταβλητή **"htot"** έχει συντελεστή -0.064 οπότε για αύξηση του "htot" κατά μία μονάδα, ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά 0.064.
- Το **"av_speeding"** έχει συντελεστή -0.334 οπότε για αύξηση της ανεξάρτητης μεταβλητής κατά μία μονάδα διατηρώντας σταθερή την τιμή των υπόλοιπων μεταβλητών, μειώνεται κατά 0.334 μονάδες η εξαρτημένη μεταβλητή και επομένως ο λογάριθμος του λόγου της πιθανότητας ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο προς την πιθανότητα να μη χρησιμοποιεί το κινητό τηλέφωνο.
- Η μεταβλητή **"avgspeed"** έχει συντελεστή -0.004 οπότε για αύξηση της μεταβλητής κατά μία μονάδα διατηρώντας σταθερές τις υπόλοιπες μεταβλητές, ο λόγος πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.004} = 0.996$.
- Η μεταβλητή **"smooth_corner"** έχει συντελεστή ίσο με -0.058 οπότε για αύξηση του "smooth_corner" κατά μία μονάδα, διατηρώντας σταθερές τις υπόλοιπες μεταβλητές, η εξαρτημένη μεταβλητή Y μειώνεται κατά 0.058. Η μονάδα μέτρησης του "smooth_corner" είναι σε %/s.

Σχετική επιρροή των μεταβλητών

Η σχετική επιρροή των μεταβλητών στα μοντέλα που αφορούν την πιθανότητα ο χρήστης να μιλάει στο κινητό τηλέφωνο υπολογίστηκαν με τη **θεωρία της ελαστικότητας** με σχέσεις που αφορούν στο λογιστικό μοντέλο παλινδρόμησης.

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Με την ελαστικότητα υπολογίζεται η επιρροή που έχουν μικρές αλλαγές των ανεξάρτητων μεταβλητών πάνω στην πιθανότητα ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο (1). Η ελαστικότητα είναι αδιάστατο μέγεθος και η τιμή της ερμηνεύεται ως το ποσοστό επί τις εκατό της μεταβολής της πιθανότητας $P(i)$ που προκαλείται από μία μεταβολή της ανεξάρτητης μεταβλητής (x_{ki}) κατά 1%.

Από τον επόμενο πίνακα, προκύπτει **το είδος και το μέγεθος της επιρροής** της κάθε ανεξάρτητης μεταβλητής στην εξαρτημένη.

Πίνακας 5.2.: Ελαστικότητες και σχετική επιρροή μεταβλητών για το μοντέλο δυαδικής λογιστικής παλινδρόμησης

	Ανεξάρτητες μεταβλητές	Συντελεστές β	Ελαστικότητα	Σχετική επιρροή e_i
Συνεχείς	Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας	-0.334	0.019	1.000
	Μέση ταχύτητα	-0.004	0.078	4.150
	Μέση στροφική ταχύτητα	-0.058	0.251	13.438
	Συνολικά απότομα συμβάντα	-0.064	0.104	5.589
Διακριτές	Πρωινή αιχμή	-0.130	0.058	1.000
	Απογευματινή αιχμή	0.262	0.122	2.082

- Η μεταβλητή **"avgspeed"** επηρεάζει κατά 4.15 φορές περισσότερο την πιθανότητα ο χρήστης να μιλάει στο κινητό τηλέφωνο σε σχέση με τη μεταβλητή **"av_speeding"**. Για αύξηση του **"avgspeed"** κατά 1% η πιθανότητα ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο μειώνεται κατά 7.8%.
- Η μεταβλητή **"smooth_corner"** επηρεάζει κατά 13.4 φορές περισσότερο την πιθανότητα ο χρήστης να μιλάει στο κινητό τηλέφωνο σε σχέση με τη μεταβλητή **"av_speeding"**. Για αύξηση του **"smooth_corner"** κατά 1% η πιθανότητα ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο μειώνεται κατά 25.1%.
- Η μεταβλητή **"htot"** επηρεάζει κατά 5.6 φορές περισσότερο την πιθανότητα ο χρήστης να μιλάει στο κινητό τηλέφωνο σε σχέση με τη μεταβλητή **"av_speeding"**. Για αύξηση του **"htot"** κατά 1% η πιθανότητα ο χρήστης να χρησιμοποιεί το κινητό τηλέφωνο μειώνεται κατά 10.4%.
- Για αύξηση της μεταβλητής **"av_speeding"** κατά 1%, η πιθανότητα χρήσης του κινητού τηλεφώνου κατά την οδήγηση μειώνεται κατά 1.9%.
- Η μεταβλητή **"driving_afternoon_rush"** επηρεάζει κατά 2.08 φορές περισσότερο την πιθανότητα ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο σε σχέση με τη μεταβλητή **"driving_morning_rush"**.

5.7.2.2 Χρήση ή μη κινητού τηλεφώνου σε αστικό περιβάλλον

Εφαρμόστηκαν τα βήματα της Mixed Binary Logistic Regression όπως πραγματοποιήθηκε για τη χρήση ή μη του κινητού τηλεφώνου σε όλους τους τύπους

οδού (παράγραφος 5.7.1). Έπειτα από διαδοχικές δοκιμές προέκυψε η ακόλουθη συνάρτηση χρησιμότητας ή Utility Function:

$$U = 1.488 - 0.233 * \text{driving_morning_rush} + 0.111 * \text{driving_afternoon_rush} - 0.034 * \text{htot_urban} - 0.001 * \text{time_urban} + 0.395 * \text{perc_speeding_urban}$$

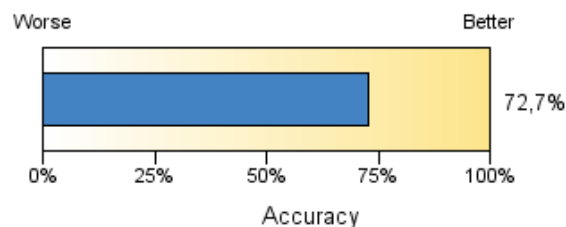
Στη συνέχεια, παρουσιάζονται με μορφή εικόνων τα αποτελέσματα του μοντέλου.

Model Summary

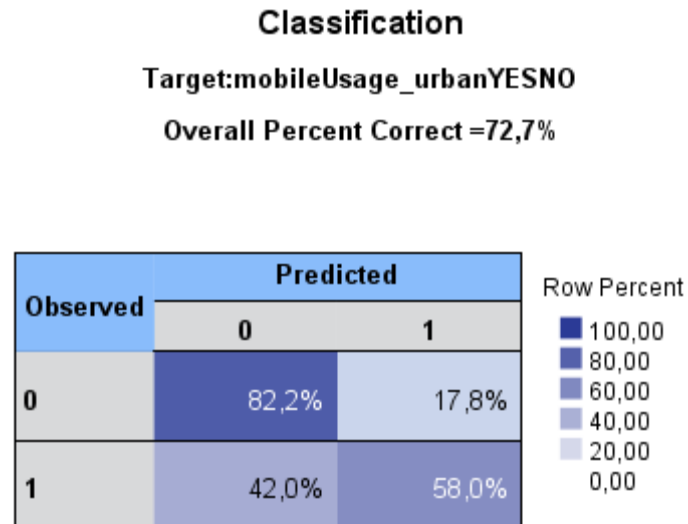
Target: mobileUsage_urbanYESNO

Target	mobileUsage_urbanYESNO
Probability Distribution	Binomial
Link Function	Logit
Information Criterion	
Akaike Corrected	88.375,633
Bayesian	88.383,477

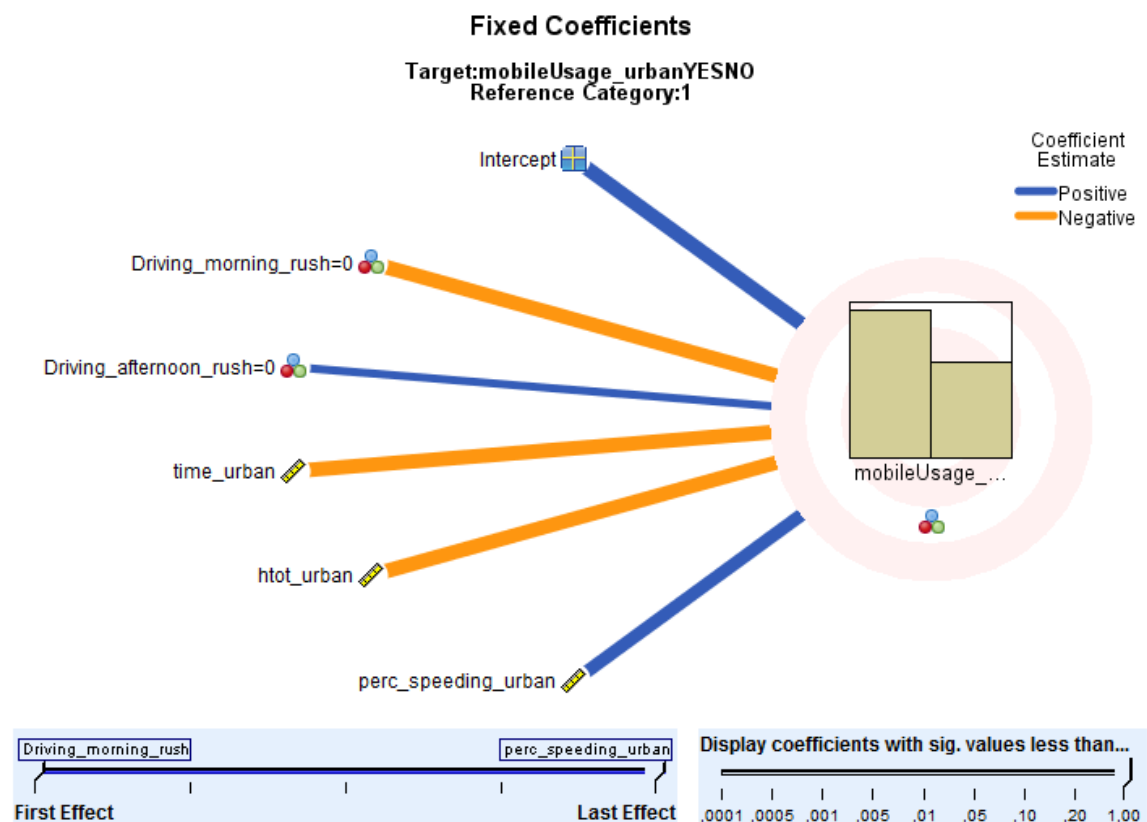
Information criteria are based on the -2 log pseudo likelihood (88.373,633) and are used to compare models. Models with smaller information criterion values fit better. When comparing models using pseudo likelihood values, caution should be used because different data transformations may be used across the models.



Εικόνα 5.43.: Χαρακτηριστικά του μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον



Εικόνα 5.44.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον



Εικόνα 5.45.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Model Term	Coefficient ▼	Std.Error	t	Sig.	Exp(Coefficient)	95% Confidence Interval for Exp (Coefficient)	
						Lower	Upper
Intercept	1,488	0,136	10,935	,000	4,429	3,392	5,784
Driving_morning_rush=0	-0,233	0,050	-4,692	,000	0,792	0,719	0,873
Driving_morning_rush=1	0,000 ^a						
Driving_afternoon_rush=0	0,111	0,042	2,649	,008	1,118	1,029	1,214
Driving_afternoon_rush=1	0,000 ^a						
time_urban	-0,001	0,000	-25,616	,000	0,999	0,999	0,999
htot_urban	-0,034	0,008	-4,409	,000	0,966	0,952	0,981
perc_speeding_urban	0,395	0,112	3,536	,000	1,484	1,192	1,847

Probability distribution:Binomial
Link function:Logit

^aThis coefficient is set to zero because it is redundant.

Εικόνα 5.46.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον

Covariance Parameters Target:mobileUsage_urbanYESNO

Covariance Parameters	Residual Effect	0
	Random Effects	1
Design Matrix Columns	Fixed Effects	8
	Random Effects	1 ^a
Common Subjects		100

Common subjects are based on the subject specifications for the residual and random effects and are used to chunk the data for better performance.

^aThis is the number of columns per common subject.

Random Effect	Estimate	Std.Error	Z	Sig.	95% Confidence Interval	
					Lower	Upper
Var(Intercept)	1,346	0,203	6,627	,000	1,002	1,809

Covariance Structure:Variance components
Subject Specification:userld

Εικόνα 5.47.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αστικό περιβάλλον

Σχολιασμός αποτελεσμάτων

- Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως προκύπτει από τα t-tests αφού **p-value < 0.05** για διάστημα εμπιστοσύνης 95%.
- Το **random effect** είναι σημαντικό αφού ισχύει και πάλι $p\text{-value} < 0.05$.
- Η διακύμανση του random effect ισούται με 1.349.
- Η τιμή **Log pseudolikelihood** είναι μικρότερη συγκριτικά με σχετικά μοντέλα (nested models).
- Τα **κατάλοιπα** του μοντέλου ελέγχθηκαν και κρίθηκαν ικανοποιητικά (Παράρτημα).
- Η **ακρίβεια** του μοντέλου είναι 72.8%.
- Το **True Negative (TN)**, δηλαδή το ποσοστό των τιμών που ήταν 1 (χρήση κινητού τηλεφώνου) και προβλέφθηκαν ως 1 ισούται με 58% και το **True Positive (TP)**, δηλαδή το ποσοστό των τιμών που ήταν 0 και προβλέφθηκαν ως 0 (μη χρήση κινητού τηλεφώνου) ισούται με 82.4%.
- Η μεταβλητή **"driving_morning_rush"** έχει συντελεστή -0.233 που σημαίνει ότι κατά την πρωινή αιχμή ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά -0.233. Ή αλλιώς ο λόγος σχετικών πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.233}$.
- Η μεταβλητή **"driving_afternoon_rush"** έχει συντελεστή 0.111 που σημαίνει ότι κατά την απογευματινή αιχμή (οπότε θα ισχύει η τιμή 1 της ανεξάρτητης μεταβλητής), ο λογάριθμος του λόγου πιθανοτήτων αυξάνεται κατά 0.111. Στη λογιστική παλινδρόμηση η εξαρτημένη μεταβλητή ισούται με τον λογάριθμο του λόγου πιθανοτήτων οπότε στην ουσία αυξάνεται η εξαρτημένη μεταβλητή κατά 0.111.
- Η μεταβλητή **"time_urban"** έχει συντελεστή -0.001 οπότε για αύξηση του "time_urban" κατά μία μονάδα, ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά 0.001.
- Το **"perc_speeding_urban"** έχει συντελεστή 0.395 οπότε για αύξηση της ανεξάρτητης μεταβλητής κατά μία μονάδα διατηρώντας σταθερή την τιμή των υπόλοιπων μεταβλητών, αυξάνεται κατά 0.395 μονάδες η εξαρτημένη μεταβλητή και επομένως ο λογάριθμος του λόγου της πιθανότητας ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο προς την πιθανότητα να μη χρησιμοποιεί το κινητό τηλέφωνο.
- Η μεταβλητή **"htot_urban"** έχει συντελεστή -0.034 οπότε για αύξηση της μεταβλητής κατά μία μονάδα διατηρώντας σταθερές τις υπόλοιπες μεταβλητές, ο λόγος πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.034} = 0.966$.

5.7.2.3 Χρήση ή μη κινητού τηλεφώνου σε υπεραστικό περιβάλλον

Εφαρμόστηκαν τα βήματα της Mixed Binary Logistic Regression όπως πραγματοποιήθηκε για τη χρήση ή μη του κινητού τηλεφώνου σε όλους τους τύπους οδού (παράγραφος 5.7.1). Έπειτα από διαδοχικές δοκιμές προέκυψε η ακόλουθη συνάρτηση χρησιμότητας ή Utility Function:

$$U = 2.443 - 0.11 * \text{driving_morning_rush} + 0.206 * \text{driving_afternoon_rush} - 0.938 * \text{perc_speeding_rural} - 0.084 * \text{htot_rural} - 2.784 * \text{time_rural/duration}$$

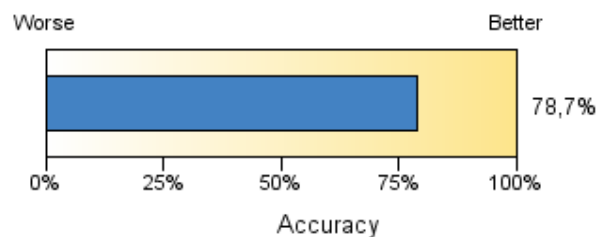
Στη συνέχεια, παρουσιάζονται με μορφή εικόνων τα αποτελέσματα του μοντέλου.

Model Summary

Target: mobileUsage_ruralYESNO

Target	mobileUsage_ruralYESNO
Probability Distribution	Binomial
Link Function	Logit
Akaike Corrected	93.314,919
Information Criterion	
Bayesian	93.322,763

Information criteria are based on the -2 log pseudo likelihood (93.312,919) and are used to compare models. Models with smaller information criterion values fit better. When comparing models using pseudo likelihood values, caution should be used because different data transformations may be used across the models.



Εικόνα 5.48.: Χαρακτηριστικά μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Classification

Target:mobileUsage_ruralYESNO

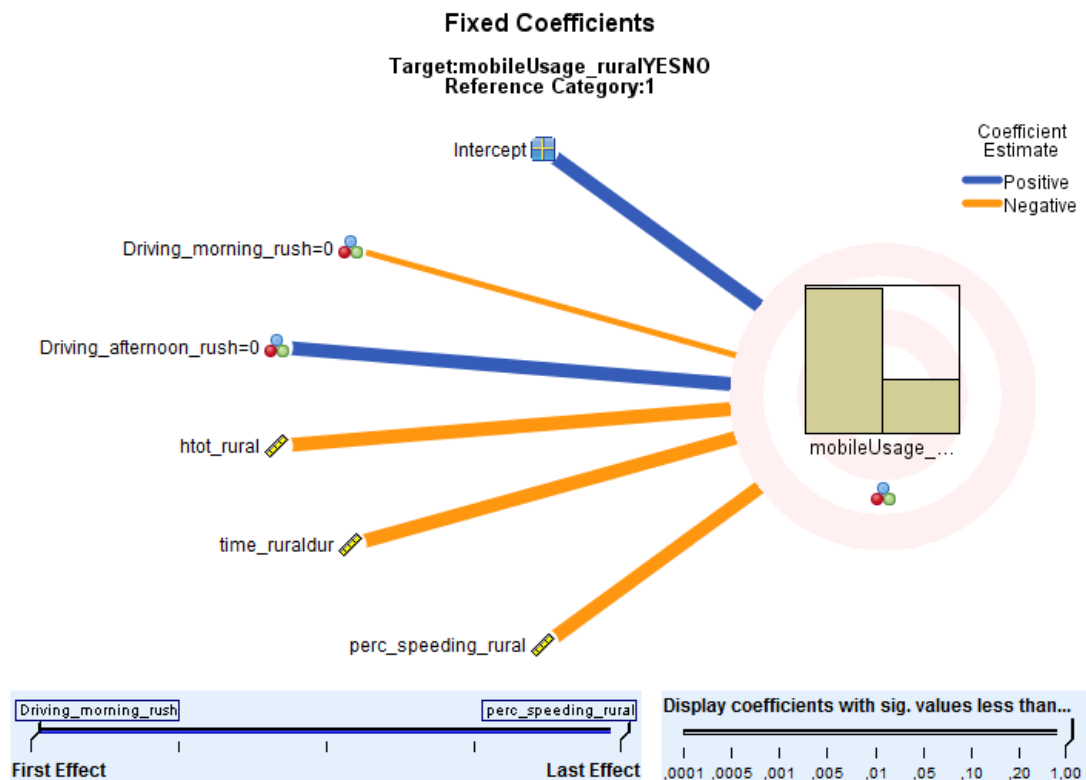
Overall Percent Correct =78,7%

Observed	Predicted	
	0	1
0	91,8%	8,2%
1	56,6%	43,4%

Row Percent

- 100,00
- 80,00
- 60,00
- 40,00
- 20,00
- 0,00

Εικόνα 5.49.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον



Εικόνα 5.50.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Model Term	Coefficient ▼	Std.Error	t	Sig.	Exp(Coefficient)	95% Confidence Interval for Exp (Coefficient)	
						Lower	Upper
Intercept	2,443	0,133	18,421	,000	11,502	8,870	14,916
Driving_morning_rush=0	-0,110	0,055	-2,012	,044	0,896	0,805	0,997
Driving_morning_rush=1	0,000 ^a						
Driving_afternoon_rush=0	0,206	0,046	4,488	,000	1,229	1,123	1,345
Driving_afternoon_rush=1	0,000 ^a						
htot_rural	-0,084	0,008	-10,761	,000	0,919	0,905	0,934
time_ruraldur	-2,784	0,086	-32,258	,000	0,062	0,052	0,073
perc_speeding_rural	-0,938	0,109	-8,563	,000	0,392	0,316	0,485

Probability distribution: Binomial
Link function: Logit

^aThis coefficient is set to zero because it is redundant.

Εικόνα 5.51.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον

Covariance Parameters

Target: mobileUsage_ruralYESNO

Covariance Parameters	Residual Effect	0
	Random Effects	1
Design Matrix Columns	Fixed Effects	8
	Random Effects	1 ^a
Common Subjects		100

Common subjects are based on the subject specifications for the residual and random effects and are used to chunk the data for better performance.

^aThis is the number of columns per common subject.

Random Effect	Estimate	Std.Error	Z	Sig.	95% Confidence Interval	
					Lower	Upper
Var(Intercept)	1,153	0,180	6,398	,000	0,849	1,566

Covariance Structure: Variance components
Subject Specification: userid

Εικόνα 5.52.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για υπεραστικό περιβάλλον

Σχολιασμός αποτελεσμάτων

- Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως προκύπτει από τα t-tests αφού **p-value < 0.05** για διάστημα εμπιστοσύνης 95%.
- Το **random effect** ήταν σημαντικό αφού ισχύει και πάλι $p\text{-value} < 0.05$.
- Η διακύμανση του random effect ισούται με 1.153.
- Η τιμή **Log pseudolikelihood** είναι μικρότερη συγκριτικά με σχετικά μοντέλα (nested models).
- Η **ακρίβεια** του μοντέλου είναι 78.7%.
- Το **True Negative (TN)**, δηλαδή το ποσοστό των τιμών που ήταν 1 (χρήση κινητού τηλεφώνου) και προβλέφθηκαν ως 1 ισούται με 43.4% και το **True Positive (TP)**, δηλαδή το ποσοστό των τιμών που ήταν 0 και προβλέφθηκαν ως 0 (μη χρήση κινητού τηλεφώνου) ισούται με 91.8%.
- Η μεταβλητή "**driving_morning_rush**" έχει συντελεστή -0.110 που σημαίνει ότι κατά την πρωινή αιχμή ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά -0.110. Ή αλλιώς ο λόγος σχετικών πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.110}$.
- Η μεταβλητή "**driving_afternoon_rush**" έχει συντελεστή 0.206 που σημαίνει ότι κατά την απογευματινή αιχμή (οπότε θα ισχύει η τιμή 1 της ανεξάρτητης μεταβλητής), ο λογάριθμος του λόγου πιθανοτήτων αυξάνεται κατά 0.206. Στη λογιστική παλινδρόμηση η εξαρτημένη μεταβλητή ισούται με τον λογάριθμο του λόγου πιθανοτήτων οπότε στην ουσία αυξάνεται η εξαρτημένη μεταβλητή κατά 0.206.
- Η μεταβλητή "**time_rural/duration**" έχει συντελεστή -2.784 οπότε για αύξηση του "time_urban" κατά μία μονάδα, ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά 2.784.
- Το "**perc_speeding_rural**" έχει συντελεστή -0.938 οπότε για αύξηση της ανεξάρτητης μεταβλητής κατά μία μονάδα διατηρώντας σταθερή την τιμή των υπόλοιπων μεταβλητών, μειώνεται κατά 0.938 μονάδες η εξαρτημένη μεταβλητή και επομένως ο λογάριθμος του λόγου της πιθανότητας ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο προς την πιθανότητα να μη χρησιμοποιεί το κινητό τηλέφωνο.
- Η μεταβλητή "**htot_rural**" έχει συντελεστή -0.084 οπότε για αύξηση της μεταβλητής κατά μία μονάδα διατηρώντας σταθερές τις υπόλοιπες μεταβλητές, ο λόγος πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.084} = 0.919$.

5.7.2.4 Χρήση ή μη κινητού τηλεφώνου σε αυτοκινητόδρομο

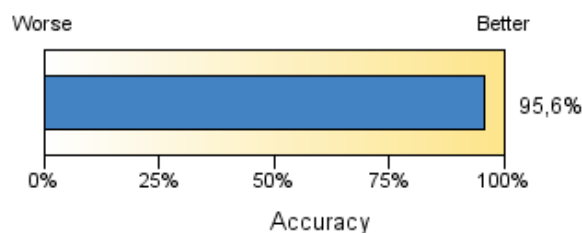
Εφαρμόστηκαν τα βήματα της Mixed Binary Logistic Regression όπως πραγματοποιήθηκε για τη χρήση ή μη του κινητού τηλεφώνου σε όλους τους τύπους οδού (παράγραφος 5.7.1). Έπειτα από διαδοχικές δοκιμές προέκυψε η ακόλουθη συνάρτηση χρησιμότητας ή Utility Function:

$$U = 4.138 - 0.373 * \text{driving_morning_rush} + 0.161 * \text{driving_afternoon_rush} - 2.512 * \text{perc_speeding_highway} + 0.148 * \text{htot_highway} - 0.001 * \text{time_highway}$$

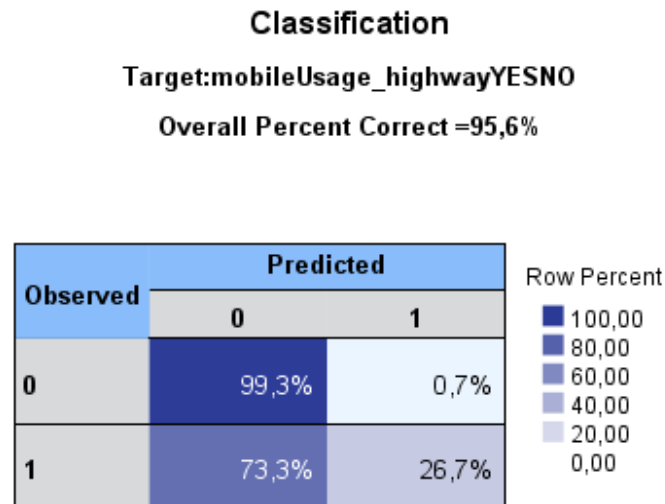
Στη συνέχεια, παρουσιάζονται με μορφή εικόνων τα αποτελέσματα του μοντέλου.

Model Summary	
Target: mobileUsage_highwayYESNO	
Target	mobileUsage_highwayYESNO
Probability Distribution	Binomial
Link Function	Logit
Akaike Corrected	142.526,599
Information Criterion	
Bayesian	142.534,443

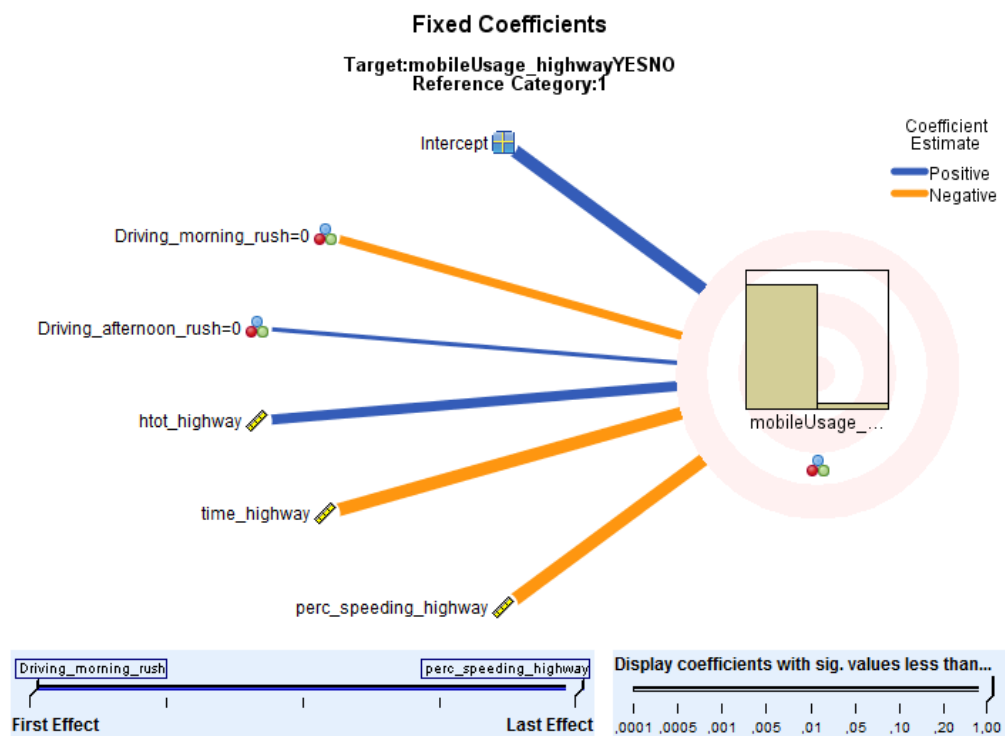
Information criteria are based on the -2 log pseudo likelihood (142.524,599) and are used to compare models. Models with smaller information criterion values fit better. When comparing models using pseudo likelihood values, caution should be used because different data transformations may be used across the models.



Εικόνα 5.53.: Χαρακτηριστικά μοντέλου της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο



Εικόνα 5.54.: Πίνακας ταξινόμησης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο



Εικόνα 5.55.: Οπτικοποίηση υπολογισμού των παραμέτρων των fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο

5. Εφαρμογή μεθοδολογίας- Αποτελέσματα

Model Term	Coefficient ▼	Std.Error	t	Sig.	Exp(Coefficient)	95% Confidence Interval for Exp (Coefficient)	
						Lower	Upper
Intercept	4,138	0,196	21,092	,000	62,682	42,671	92,077
Driving_morning_rush=0	-0,373	0,123	-3,041	,002	0,689	0,542	0,876
Driving_morning_rush=1	0,000 ^a						
Driving_afternoon_rush=0	0,161	0,092	1,744	,081	1,175	0,980	1,408
Driving_afternoon_rush=1	0,000 ^a						
htot_highway	0,148	0,043	3,472	,001	1,160	1,067	1,261
time_highway	-0,001	0,000	-21,698	,000	0,999	0,998	0,999
perc_speeding_highway	-2,512	0,241	-10,407	,000	0,081	0,051	0,130

Probability distribution: Binomial
Link function: Logit

^aThis coefficient is set to zero because it is redundant.

Εικόνα 5.56.: Πίνακας συντελεστών fixed effects της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο

Covariance Parameters Target: mobileUsage_highwayYESNO

Covariance Parameters	Residual Effect	0
	Random Effects	1
Design Matrix Columns	Fixed Effects	8
	Random Effects	1 ^a
Common Subjects		100

Common subjects are based on the subject specifications for the residual and random effects and are used to chunk the data for better performance.

^aThis is the number of columns per common subject.

Random Effect	Estimate	Std.Error	Z	Sig.	90% Confidence Interval	
					Lower	Upper
Var(Intercept)	1,429	0,283	5,059	,000	1,032	1,978

Covariance Structure: Variance components
Subject Specification: userid

Εικόνα 5.57.: Πίνακας υπολογισμού τυχαίας διακύμανσης της μεικτής δυαδικής λογιστικής παλινδρόμησης για αυτοκινητόδρομο

Σχολιασμός αποτελεσμάτων

- Οι μεταβλητές που διατηρήθηκαν στο μοντέλο είχαν όλες σημαντικό fixed effect όπως προκύπτει από τα t-tests αφού **p-value < 0.1** για διάστημα εμπιστοσύνης 90%.
- Το **random effect** ήταν σημαντικό αφού ισχύει και πάλι $p\text{-value} < 0.1$ (για διάστημα εμπιστοσύνης 90%, όπως επιλέχθηκε εδώ).
- Η διακύμανση του random effect ισούται με 1.429.
- Η τιμή **Log pseudolikelihood** είναι μικρότερη συγκριτικά με σχετικά μοντέλα (nested models).
- Η **ακρίβεια** του μοντέλου είναι 95.6%, η υψηλότερη από όλα τα προηγούμενα μοντέλα.
- Το **True Negative (TN)**, δηλαδή το ποσοστό των τιμών που ήταν 1 (χρήση κινητού τηλεφώνου) και προβλέφθηκαν ως 1 ισούται με 26.7% και το **True Positive (TP)**, δηλαδή το ποσοστό των τιμών που ήταν 0 και προβλέφθηκαν ως 0 (μη χρήση κινητού τηλεφώνου) ισούται με 99.3%. Η τιμή του TP είναι πάρα πολύ υψηλή, ωστόσο η χαμηλή τιμή του TN δεν μας επιτρέπει να πούμε ότι το συγκεκριμένο μοντέλο προβλέπει το ίδιο ικανοποιητικά και την πιθανότητα ο χρήστης να μιλάει στο κινητό τηλέφωνο .
- Η μεταβλητή **"driving_morning_rush"** έχει συντελεστή -0.373 που σημαίνει ότι κατά την πρωινή αιχμή ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά -0.373. Ή αλλιώς ο λόγος σχετικών πιθανοτήτων μειώνεται κατά τον συντελεστή $e^{-0.373}$.
- Η μεταβλητή **"driving_afternoon_rush"** έχει συντελεστή 0.161 που σημαίνει ότι κατά την απογευματινή αιχμή (οπότε θα ισχύει η τιμή 1 της ανεξάρτητης μεταβλητής), ο λογάριθμος του λόγου πιθανοτήτων αυξάνεται κατά 0.161. Στη λογιστική παλινδρόμηση η εξαρτημένη μεταβλητή ισούται με τον λογάριθμο του λόγου πιθανοτήτων οπότε στην ουσία αυξάνεται η εξαρτημένη μεταβλητή κατά 0.161.
- Η μεταβλητή **"time_highway"** έχει συντελεστή -0.001 οπότε για αύξηση του "time_highway" κατά μία μονάδα, ο λογάριθμος του λόγου πιθανοτήτων μειώνεται κατά 0.001.
- Το **"perc_speeding_highway"** έχει συντελεστή -2.512 οπότε για αύξηση της ανεξάρτητης μεταβλητής κατά μία μονάδα διατηρώντας σταθερή την τιμή των υπόλοιπων μεταβλητών, μειώνεται κατά 2.512 η εξαρτημένη μεταβλητή και επομένως ο λογάριθμος του λόγου της πιθανότητας ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο προς την πιθανότητα να μη χρησιμοποιεί το κινητό τηλέφωνο.

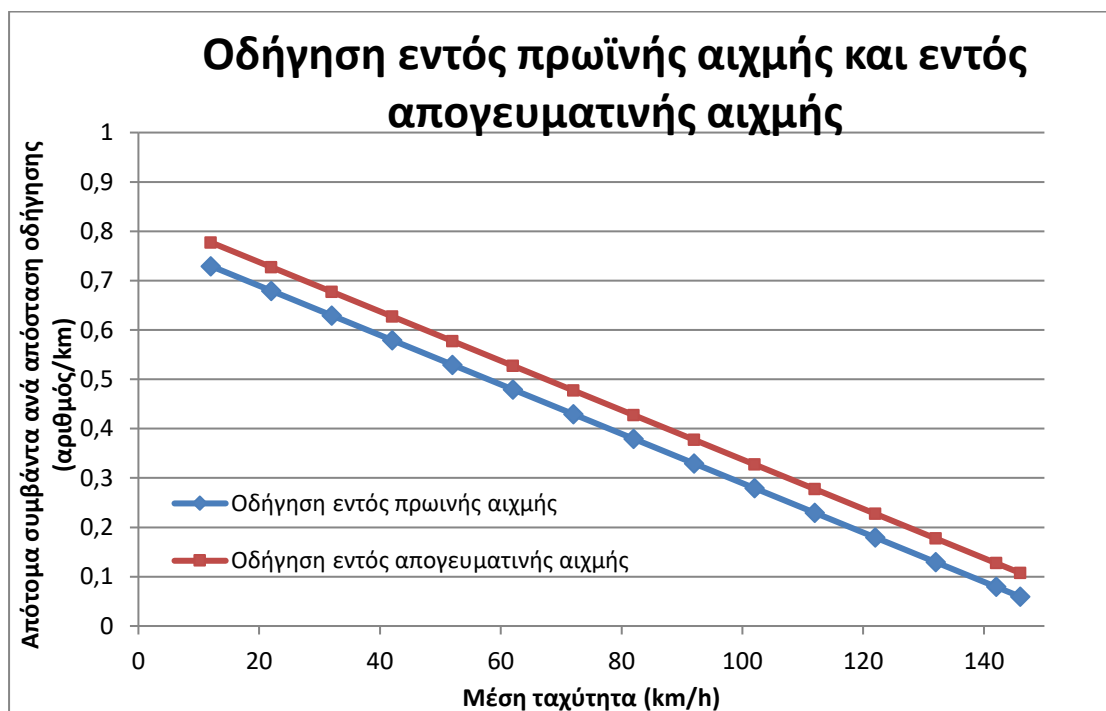
- Η μεταβλητή "**htot_highway**" έχει συντελεστή 0.148 οπότε για αύξηση της μεταβλητής κατά μία μονάδα διατηρώντας σταθερές τις υπόλοιπες μεταβλητές, ο λόγος πιθανοτήτων αυξάνεται κατά τον συντελεστή $e^{0.148} = 1.160$.

5.8 Ανάλυση ευαισθησίας

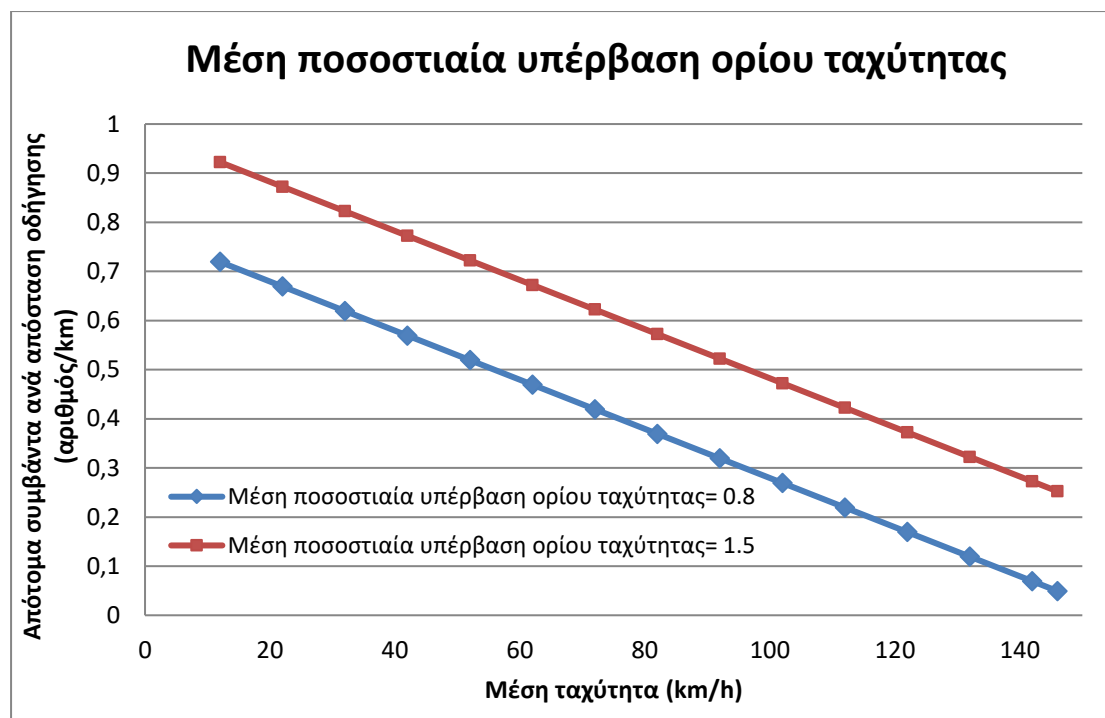
Με στόχο την καλύτερη κατανόηση της επιρροής των ανεξάρτητων μεταβλητών στην εξαρτημένη μεταβλητή, που προβλέπει το μοντέλο της γραμμικής και της δυαδικής λογιστικής παλινδρόμησης, αναπτύχθηκαν τα διαγράμματα ευαισθησίας. Το συγκεκριμένο διάγραμμα στη γραμμική παλινδρόμηση περιγράφει την ευαισθησία της εξεταζόμενης εξαρτημένης μεταβλητής htot/totaldist όταν μεταβάλλεται μια εκ των ανεξάρτητων συνεχών μεταβλητών και οι υπόλοιπες παραμένουν σταθερές. Στη συγκεκριμένη περίπτωση η μεταβλητή που μεταβάλλεται είναι η συνεχής μεταβλητή που περιγράφει την μέση ταχύτητα οδήγησης (avgspeed). Στη λογιστική παλινδρόμηση το διάγραμμα ανάλυσης ευαισθησίας περιγράφει την ευαισθησία της πιθανότητας ο οδηγός να κάνει χρήση του κινητού τηλεφώνου όταν μεταβάλλεται και πάλι η μεταβλητή avgspeed. Τέλος αναφέρονται μερικές γενικές παρατηρήσεις. Υπενθυμίζεται ότι η πιθανότητα χρήσης του κινητού τηλεφώνου κατά την οδήγηση δίνεται από την παρακάτω σχέση (όπου U η συνάρτηση χρησιμότητας):

$$P = \frac{e^U}{e^U + 1}$$

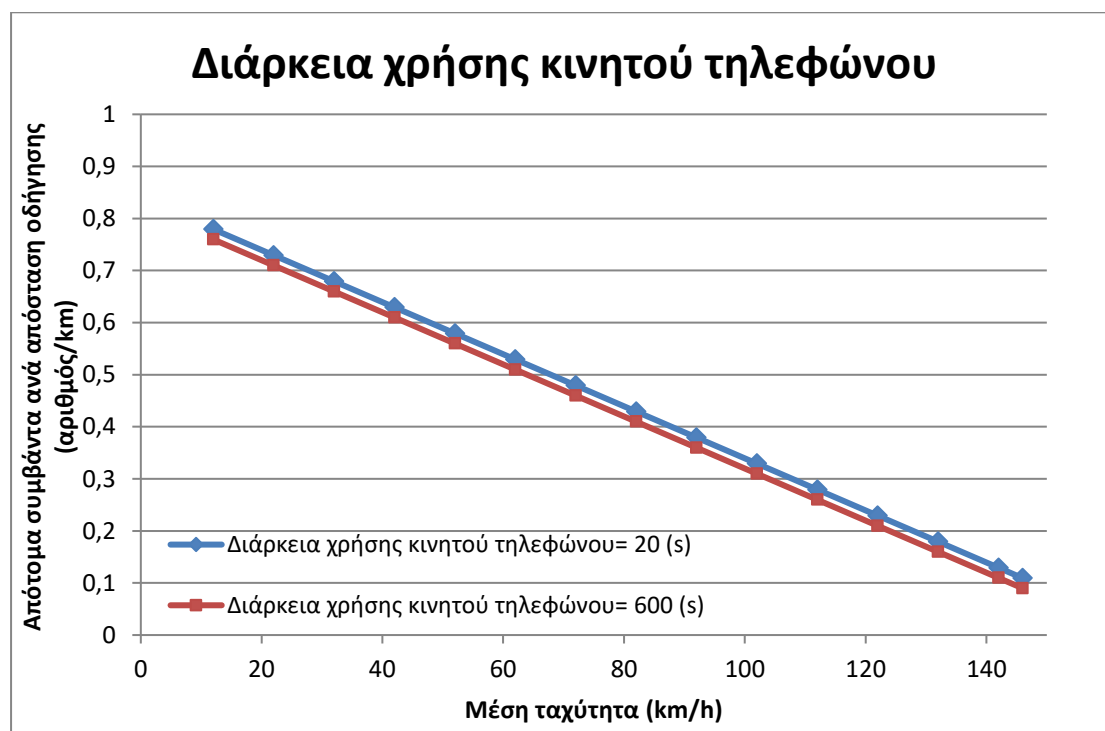
5.8.1 Mixed Γραμμική Παλινδρόμηση για το σύνολο των απότομων συμβάντων από όλους τους τύπους οδού



Διάγραμμα 5.1.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για τις ώρες εντός πρωινής και απογευματινής αιχμής



Διάγραμμα 5.2.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης ποσοστιαίας υπέρβασης του ορίου ταχύτητας

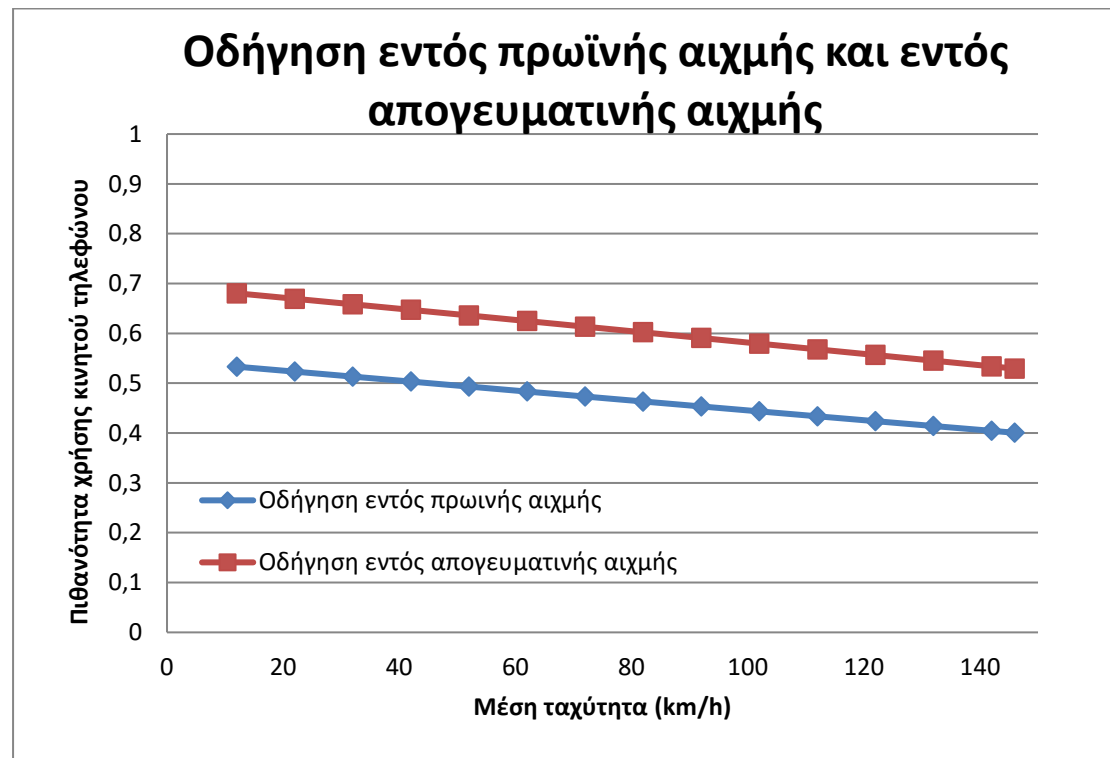


Διάγραμμα 5.3.: Μεταβολή των συνολικών συμβάντων ανά απόσταση με τη μέση ταχύτητα για δυο πιθανές τιμές της διάρκειας χρήσης του κινητού τηλεφώνου

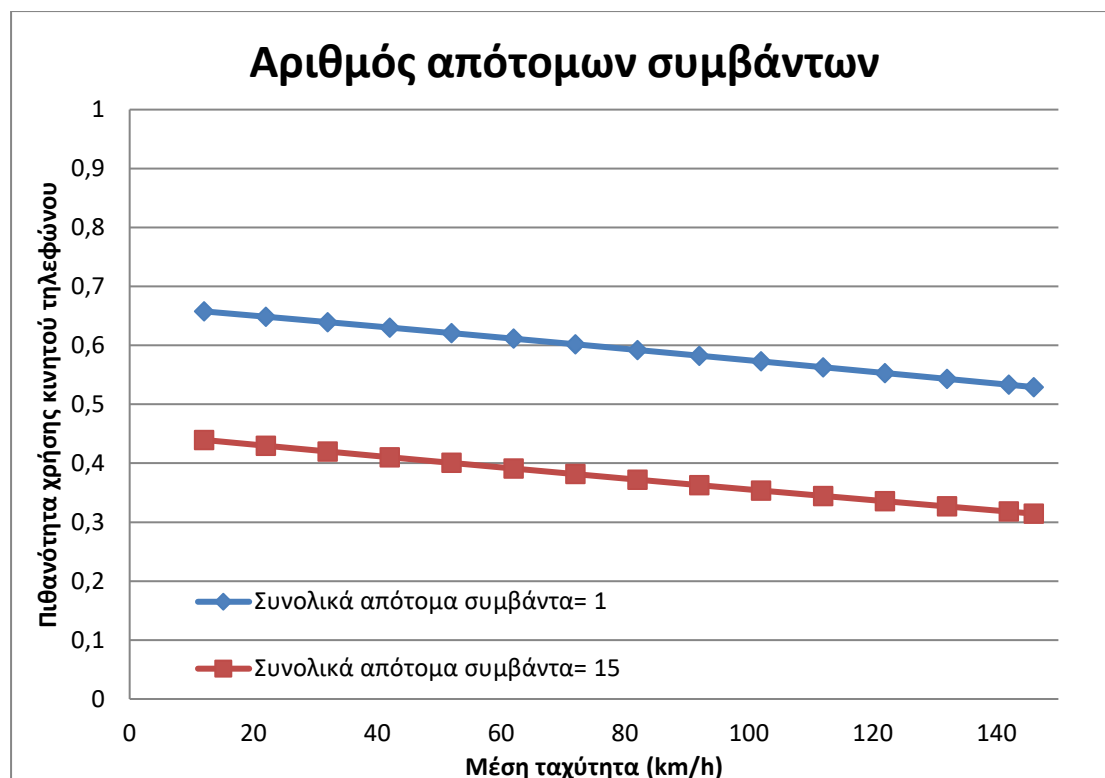
Διαπιστώνουμε ότι σε όλες τις παραπάνω περιπτώσεις, όσο **αυξάνεται η μέση ταχύτητα** οδήγησης τόσο **μειώνονται τα απότομα συμβάντα** ανά απόσταση οδήγησης. Το ίδιο συμβαίνει και για χαμηλές και για υψηλές τιμές της εξεταζόμενης

σταθερής μεταβλητής. Για παράδειγμα, όπως διαπιστώνουμε από το τελευταίο διάγραμμα, για αύξηση της ταχύτητας κατά $146-12=134$ km/h, κατά την απογευματινή αιχμή και για μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας ίση με 1 (100%), ο αριθμός των απότομων συμβάντων ανά απόσταση μειώνεται περίπου κατά 0.65. Επιπλέον, παρατηρείται ότι εντός απογευματινής αιχμής, για χαμηλότερη τιμή *mobileUsage_time* (77 s αντί για 600 s) και υψηλότερη τιμή *av_speeding* (1.5 αντί για 0.8), με σταθερή τιμή μέσης ταχύτητας *avspeed* και των υπόλοιπων εκάστοτε μεταβλητών, τα απότομα συμβάντα ανά απόσταση οδήγησης είναι περισσότερα.

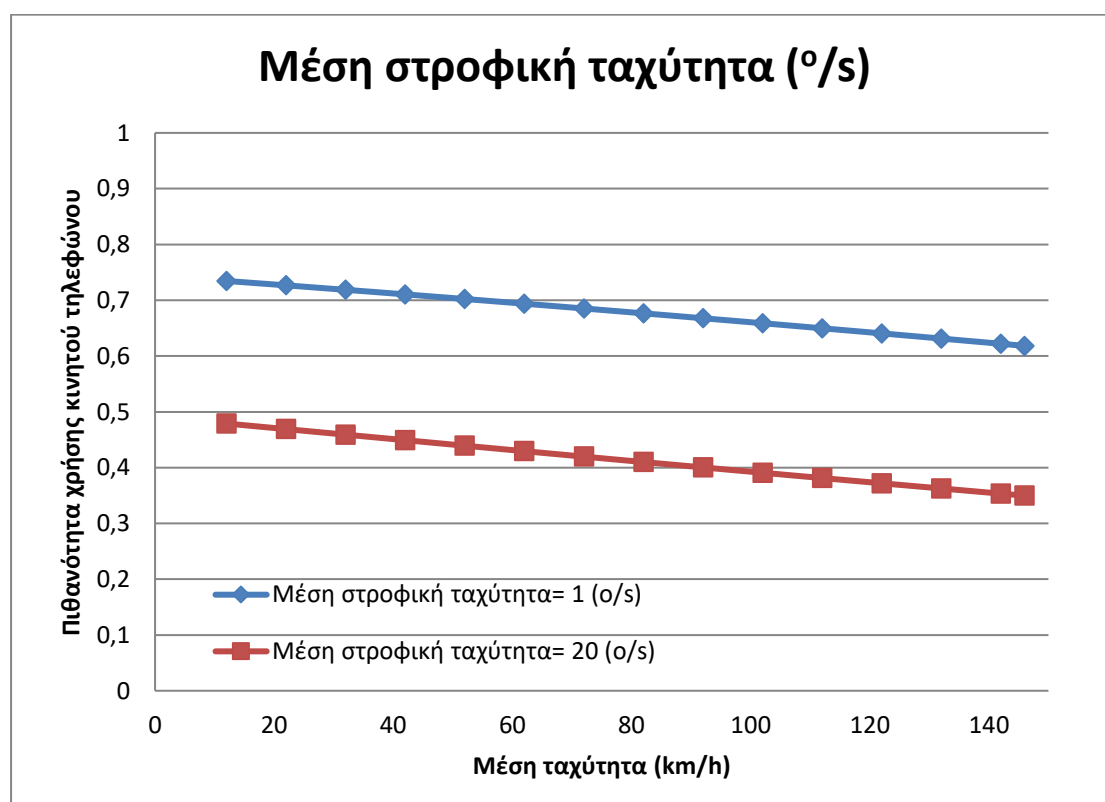
5.8.2 Mixed Δυαδική Λογιστική Παλινδρόμηση για την πιθανότητα χρήσης του κινητού τηλεφώνου κατά την οδήγηση



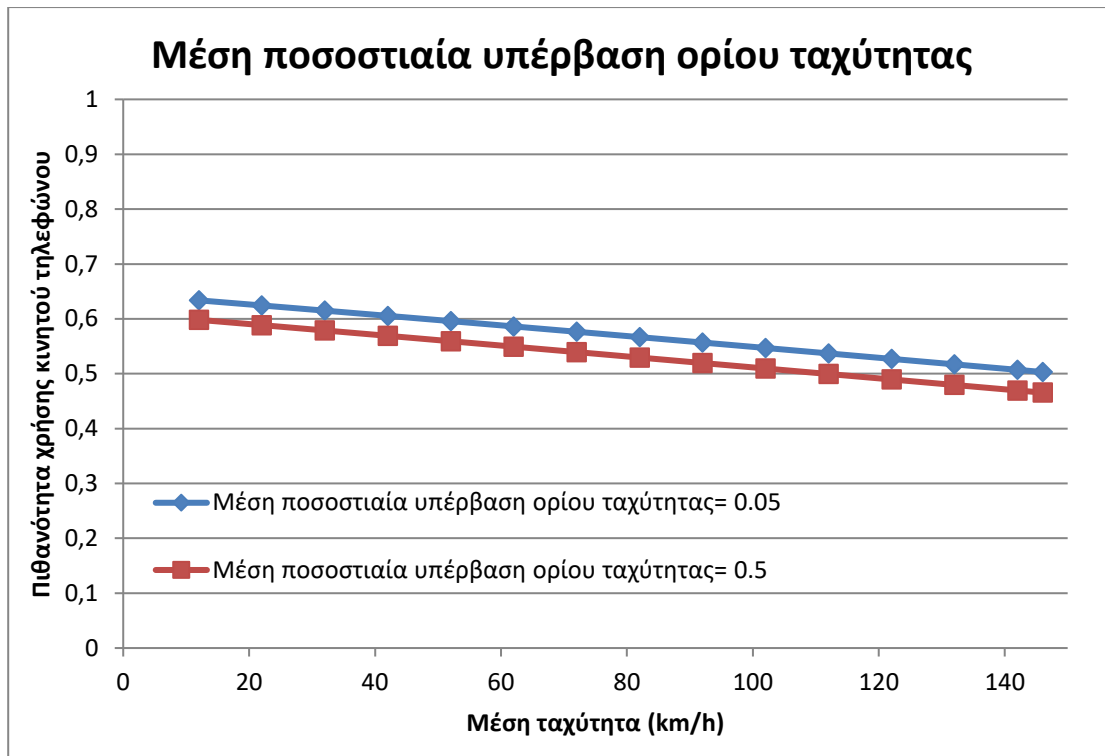
Διάγραμμα 5.4.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για τις ώρες εντός πρωϊνής και απογευματινής αιχμής



Διάγραμμα 5.5.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές του αριθμού απότομων συμβάντων



Διάγραμμα 5.6.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης στροφικής ταχύτητας



Διάγραμμα 5.7.: Μεταβολή της πιθανότητας χρήσης κινητού τηλεφώνου στην οδήγηση με τη μέση ταχύτητα για δυο πιθανές τιμές της μέσης ποσοστιαίας υπέρβασης του ορίου ταχύτητας

Διαπιστώνεται ότι σε όλες τις παραπάνω περιπτώσεις, όσο **αυξάνεται η μέση ταχύτητα** οδήγησης τόσο **μειώνεται η πιθανότητα χρήσης του κινητού τηλεφώνου** κατά την οδήγηση. Το ίδιο συμβαίνει και για χαμηλές και για υψηλές τιμές της εξεταζόμενης σταθερής μεταβλητής. Για παράδειγμα, όπως διαπιστώνουμε από το τελευταίο διάγραμμα, κατά την απογευματινή αιχμή, για αριθμό απότομων συμβάντων ίσο με 3, μέση στροφική ταχύτητα ίση με 9.5 °/s και μέση ποσοστιαία υπέρβαση ορίου ταχύτητας ίση με 0.05 (5%), η πιθανότητα χρήσης του κινητού τηλεφώνου μειώνεται κατά 0.15 για αύξηση της μέση ταχύτητας κατά 134 km/h. Επιπλέον, εντός απογευματινής αιχμής, για χαμηλότερη τιμή `smooth_corner` (1 °/s αντί για 20 °/s), `av_speeding` (0.05 αντί για 0.5) και αριθμού απότομων συμβάντων (1 αντί για 15) με σταθερή τιμή μέσης ταχύτητας `avspeed` και των υπόλοιπων εκάστοτε μεταβλητών, η πιθανότητα χρήσης του κινητού τηλεφώνου είναι μεγαλύτερη.

6. Συμπεράσματα

6.1 Σύνοψη Αποτελεσμάτων

Αντικείμενο της παρούσας Διπλωματικής Εργασίας αποτέλεσε η διερεύνηση και η **προτυποποίηση της επιρροής της χρήσης κινητού τηλεφώνου στη συμπεριφορά του οδηγού αξιοποιώντας λεπτομερή δεδομένα από αισθητήρες έξυπνων κινητών τηλεφώνων.**

Συγκεκριμένα, εξετάστηκαν **δύο επιμέρους στόχοι**. Πρώτον, κατά πόσο χαρακτηριστικά οδήγησης που καταγράφηκαν από τους αισθητήρες έξυπνων κινητών τηλεφώνων επηρεάζουν τα απότομα συμβάντα κατά την οδήγηση και δεύτερον, η δυνατότητα πρόβλεψης της χρήσης ή μη του κινητού τηλεφώνου κατά την οδήγηση μέσα από την καταγραφή σχετικών μεγεθών.

Μετά τον καθορισμό του επιδιωκόμενου στόχου, ξεκίνησε η **βιβλιογραφική αναζήτηση** ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας σε χώρες της Ευρωπαϊκής Ένωσης, αλλά και παγκοσμίως, καθώς και των διαθέσιμων στοιχείων που ήταν απαραίτητα για τη συγκεκριμένη διερεύνηση.

Αφού μελετήθηκαν τα στοιχεία που συγκεντρώθηκαν από τη βιβλιογραφική ανασκόπηση, σειρά είχε η ανάλυση της βάσης δεδομένων που συλλέχθηκαν από 100 οδηγούς σε **πείραμα σε πραγματικές συνθήκες οδήγησης** μέσω των αισθητήρων έξυπνων κινητών τηλεφώνων. Κατά τη στατιστική επεξεργασία αποφασίστηκε η εξαγωγή μαθηματικών μοντέλων τόσο για το σύνολο των οδικών δικτύων (αστικό περιβάλλον, υπεραστικό περιβάλλον, αυτοκινητόδρομοι) όσο και ξεχωριστά για κάθε τύπο οδού.

Για την εκπλήρωση του πρώτου επιμέρους στόχου εφαρμόστηκε **γραμμική παλινδρόμηση** καθώς η εξαρτημένη μεταβλητή (αριθμός συμβάντων ανά απόσταση) είναι συνεχής μεταβλητή. Για την εκπλήρωση του δεύτερου επιμέρους στόχου εφαρμόστηκε η **δυναμική λογιστική παλινδρόμηση** καθώς η εξαρτημένη μεταβλητή παίρνει δύο μόνο τιμές (0 ή 1) ανάλογα αν ο οδηγός χρησιμοποιεί ή όχι το κινητό τηλέφωνο. Σημειώνεται ότι καθότι υπάρχουν πολλές μετρήσεις-διαδρομές που αντιστοιχούν σε κάθε οδηγό εφαρμόζονται μοντέλα μεικτών επιδράσεων (**mixed effects models**) ή αλλιώς μεικτά μοντέλα ώστε να αποφευχθούν σφάλματα λόγω της επαναληπτικότητας των δεδομένων.

Από την παραπάνω επεξεργασία προέκυψαν συγκεντρωτικά τα ακόλουθα μοντέλα που συνοψίζονται στον πίνακα 6.1:

6. Συμπεράσματα

Πίνακας 6.1.: Συγκεντρωτικός πίνακας στατιστικών μοντέλων

	Απότομα συμβάντα ανά απόσταση οδήγησης				Πρόβλεψη χρήσης ή μη κινητού τηλεφώνου			
	Mixed Γραμμική Παλινδρόμηση				Mixed Δυναμική Λογιστική Παλινδρόμηση			
	Όλοι οι τύποι οδού	Αστικό περιβάλλον	Υπεραστικό περιβάλλον	Αυτοκινητό-δρομος	Όλοι οι τύποι οδού	Αστικό περιβάλλον	Υπεραστικό περιβάλλον	Αυτοκινητό-δρομος
Σταθερά	0.520	0.550	0.320	0.016	1.094	1.488	2.443	4.138
Πρωινή Αιχμή	-0.018	-0.095	-0.050	-0.007	-0.130	-0.233	-0.110	-0.373
Απογευματινή Αιχμή	0.030				0.262	0.111	0.206	0.161
Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας	0.290				-0.334			
Μέση ταχύτητα	-0.005				-0.004			
Διάρκεια χρήσης κινητού τηλεφώνου	0.000							
Μέση στροφική ταχύτητα					-0.058			
Συνολικά απότομα συμβάντα					-0.064			
Μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας σε αστικό περιβάλλον		0.300						
Απόσταση οδήγησης σε αστικό περιβάλλον		-0.009						
Ποσοστό χρόνου χρήσης του κινητού τηλεφώνου σε αστικό περιβάλλον		0.083						
Απότομα συμβάντα σε αστικό περιβάλλον						-0.034		
Διάρκεια οδήγησης σε αστικό περιβάλλον						-0.001		
Ποσοστό χρόνου υπέρβασης του ορίου ταχύτητας σε αστικό περιβάλλον						0.395		
Απόσταση οδήγησης σε υπεραστικό περιβάλλον			-0.001					
Ποσοστό χρόνου υπέρβασης του ορίου ταχύτητας σε υπεραστικό περιβάλλον			-0.092				-0.938	
Ποσοστό χρόνου χρήσης του κινητού τηλεφώνου σε υπεραστικό περιβάλλον			0.140					
Απότομα συμβάντα σε υπεραστικό περιβάλλον							-0.084	
Ποσοστό χρόνου οδήγησης σε υπεραστικό περιβάλλον							-2.784	
Μέση ποσοστιαία υπέρβαση ορίου ταχύτητας σε αυτοκινητόδρομο				0.370				
Χρόνος χρήσης του κινητού τηλεφώνου σε αυτοκινητόδρομο				0.000				
Απόσταση οδήγησης σε αυτοκινητόδρομο				0.000				
Ποσοστό χρόνου υπέρβασης του ορίου ταχύτητας σε αυτοκινητόδρομο								-2.512
Απότομα συμβάντα σε αυτοκινητόδρομο								0.148
Διάρκεια οδήγησης σε αυτοκινητόδρομο								-0.001

6.2 Συμπεράσματα

Από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας προέκυψαν αποτελέσματα άμεσα συνδεδεμένα με το αρχικό ερώτημα και στόχο της εργασίας. Στο υποκεφάλαιο αυτό επιχειρείται να δοθεί μια απάντηση στα συνολικά ερωτήματα της έρευνας με σύνθεση των αποτελεσμάτων των προηγούμενων κεφαλαίων. Τα γενικά συμπεράσματα που προέκυψαν συνοψίζονται ως εξής:

- Στις έρευνες που έχουν διεξαχθεί έως τώρα παγκοσμίως αναφορικά με τη μελέτη της συμπεριφοράς των οδηγών, η συλλογή των δεδομένων γινόταν είτε μέσω του συστήματος διαγνωστικών του οχήματος (OBD) είτε μέσω άλλων πολύπλοκων συνδυασμών οργάνων καταγραφής που συνδέονταν με το όχημα (naturalistic driving, κλπ.) ή με προσομοιωτές οδήγησης και ερωτηματολόγια. Στην παρούσα Διπλωματική Εργασία εξετάζεται **για πρώτη φορά** η παρακολούθηση της συμπεριφοράς των οδηγών με **δεδομένα που συλλέγονται από αισθητήρες έξυπνων κινητών τηλεφώνων** που αφορούν στον οδηγό και όχι στο όχημα. Επιπλέον, επιχειρείται η διερεύνηση της επιθετικής οδήγησης (απότομα συμβάντα) και της απόσπασης προσοχής με επιπλέον μεταβλητές πέραν εκείνων που έχουν χρησιμοποιηθεί στην πλειονότητα παρόμοιων ερευνών.
- Με την ανάλυση διαπιστώθηκε ότι **είναι δυνατή η πρόβλεψη της χρήσης ή μη κινητού τηλεφώνου** με αρκετά υψηλή ακρίβεια κατά την οδήγηση εάν είναι γνωστά ορισμένα μεγέθη, όπως η μέση ταχύτητα οδήγησης, η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας, η μέση στροφική ταχύτητα, η διάρκεια χρήσης του κινητού τηλεφώνου, ο αριθμός απότομων συμβάντων και η περίοδος οδήγησης σε σχέση με την πρωινή και την απογευματινή αιχμή. Πιο συγκεκριμένα:

-Η πρόβλεψη της χρήσης ή μη του κινητού τηλεφώνου πραγματοποιείται με **μεγαλύτερη ακρίβεια** κατά σειρά πρώτα για την οδήγηση σε **αυτοκινητόδρομο** και στη συνέχεια για υπεραστικό δίκτυο, αστικό δίκτυο και τέλος για το σύνολο των οδικών δικτύων.

-Το μέγεθος που επηρεάζει περισσότερο σε σχέση με τους υπόλοιπους παράγοντες τη δυνατότητα πρόβλεψης της χρήσης ή μη του κινητού τηλεφώνου συνολικά για όλους τους τύπους οδού είναι **η μέση στροφική ταχύτητα**.

-**Όσο μειώνεται η μέση ταχύτητα οδήγησης τόσο μεγαλύτερη γίνεται η πιθανότητα ο οδηγός να μιλάει στο κινητό τηλέφωνο.** Το ίδιο παρατηρείται για τη μέση στροφική ταχύτητα και τα απότομα συμβάντα τόσο για όλους τους τύπους οδού όσο και για αστικό και υπεραστικό περιβάλλον. Αυτό ενδεχομένως εξηγείται από το γεγονός ότι ο οδηγός προβαίνει σε **αντισταθμιστική συμπεριφορά** (risk compensation) και έτσι οδηγεί πιο προσεκτικά, με χαμηλότερες ταχύτητες και λιγότερο νευρικά προκειμένου να μειώσει τον κίνδυνο ατυχήματος.

-Στον **αυτοκινητόδρομο**, παρατηρείται ότι η αύξηση των απότομων συμβάντων οδηγεί σε αύξηση της πιθανότητας ο οδηγός να μιλάει στο κινητό τηλέφωνο, ενδεχομένως διότι η χρήση του κινητού τηλεφώνου που αποσπά την προσοχή του οδηγού είναι μια βασική αιτία που μπορεί να οδηγήσει σε

μεγαλύτερες διακυμάνσεις στη συμπεριφορά του οδηγού και περισσότερα απότομα συμβάντα σε ένα τέτοιο περιβάλλον.

-Όσο μεγαλώνει η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας συνολικά σε όλους τους τύπους οδού τόσο μειώνεται η πιθανότητα ο οδηγός να κάνει χρήση του κινητού τηλεφώνου. Αυτή η διαπίστωση ενδεχομένως εξηγείται καθώς ο οδηγός δε συνηθίζει να μιλάει στο κινητό τηλέφωνο σε υψηλές ταχύτητες εφόσον απαιτείται μεγαλύτερη συγκέντρωση στο έργο της οδήγησης.

- **Ο αριθμός των απότομων συμβάντων** ανά απόσταση οδήγησης εξαρτάται από την περίοδο οδήγησης σε σχέση με την πρωινή και την απογευματινή αιχμή, τη μέση ταχύτητα οδήγησης, τη διάρκεια χρήσης του κινητού τηλεφώνου και επομένως την απόσπαση προσοχής και τη μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας. Πιο συγκεκριμένα:

-Το μέγεθος που επηρεάζει περισσότερο σε σχέση με τους υπόλοιπους παράγοντες την πραγματοποίηση απότομων συμβάντων είναι η **μέση ταχύτητα οδήγησης**.

-Όσο αυξάνεται η **διάρκεια χρήσης του κινητού τηλεφώνου τόσο μειώνονται τα απότομα συμβάντα** (ανά απόσταση) ίσως για λόγους αντισταθμιστικής συμπεριφοράς του οδηγού.

-Όσο αυξάνεται η μέση ποσοστιαία υπέρβαση του ορίου ταχύτητας συνολικά για όλους τους τύπους οδού τόσο αυξάνονται τα απότομα συμβάντα ενδεχομένως επειδή ο οδηγός σε περίπτωση απρόσμενου γεγονότος (π.χ. στροφή, εμπόδιο) δεν έχει το χρόνο να επαναφέρει ομαλά την οδηγική του κατάσταση οπότε προβαίνει σε απότομες κινήσεις. Το ίδιο ισχύει και για την οδήγηση σε αστικό περιβάλλον.

-Όσο αυξάνεται η μέση ταχύτητα οδήγησης τόσο μειώνονται τα απότομα συμβάντα ανά διανυθείσα απόσταση, γεγονός που ίσως σχετίζεται με την οδήγηση σε αυτοκινητόδρομο και στο υπεραστικό δίκτυο στον οποίο οι ταχύτητες είναι εν γένει μεγαλύτερες ενώ τα απότομα συμβάντα λιγότερα καθώς δεν υπάρχουν συχνές διασταυρώσεις, σηματοδότες κλπ.

-Τέλος, όσο μεγαλύτερη η απόσταση οδήγησης εντός αστικού, υπεραστικού δικτύου και αυτοκινητοδρόμου τόσο χαμηλότερος είναι ο αντίστοιχος αριθμός απότομων συμβάντων ανά διανυθείσα απόσταση αφού ενδεχομένως ο οδηγός εξοικειώνεται με τις συνθήκες του περιβάλλοντος.

- Τα στατιστικά μοντέλα που αναπτύχθηκαν είναι κατάλληλα για τις συγκεκριμένες αναλύσεις και την επίτευξη των στόχων της Διπλωματικής Εργασίας αυτής. Πιο συγκεκριμένα, επιλέχθηκε η **γραμμική παλινδρόμηση** για την εξαρτημένη μεταβλητή "απότομα συμβάντα ανά απόσταση" επειδή είναι συνεχής, ενώ για την μεταβλητή που αναφέρεται στη χρήση ή μη του κινητού τηλεφώνου και η οποία είναι κατηγορική (0 ή 1) εφαρμόστηκε η **δυναμική λογιστική παλινδρόμηση**. Και στις δύο κατηγορίες μοντέλων, επειδή υπήρχαν επαναλαμβανόμενες μετρήσεις-

διαδρομές για κάθε οδηγό, αναπτύχθηκαν μοντέλα **Μεικτών Επιδράσεων** (mixed-effects models).

- Τα σαφή και ερμηνεύσιμα αποτελέσματα, τα οποία μάλιστα συμφωνούν με τη διεθνή βιβλιογραφία καταδεικνύουν την **καταλληλότητα της μεθόδου** ανάλυσης για την πρόβλεψη τόσο της πιθανότητας χρήσης του κινητού τηλεφώνου όσο και των απότομων συμβάντων κατά την οδήγηση. Η αποτελεσματικότητα της μεθόδου επιτρέπει την εφαρμογή της σε έρευνες με παρόμοιο αντικείμενο έπεται από κατάλληλη προσαρμογή των συνθηκών. Βέβαια, οι μεταβλητές θα πρέπει να επιλέγονται ανάλογα με την εξεταζόμενη περίπτωση.

6.3 Προτάσεις για βελτίωση της Οδικής Ασφάλειας

Σύμφωνα με τα αποτελέσματα και τα συνολικά συμπεράσματα που εξήχθησαν κατά την εκπόνηση της Εργασίας αυτής, επιχειρείται η παράθεση μιας σειράς προτάσεων, οι οποίες ενδεχομένως να συμβάλουν στη βελτίωση της διαχείρισης της κυκλοφορίας καθώς και στη βελτίωση του επιπέδου οδικής ασφάλειας.

- ✓ Με την **αυστηρότερη και συχνότερη αστυνόμευση** θα καταστεί δυνατός ο περιορισμός της χρήσης του κινητού τηλεφώνου κατά την οδήγηση και θα δημιουργηθούν ασφαλέστερες συνθήκες οδήγησης στους δρόμους τόσο για τους υπόλοιπους οδηγούς (ΙΧ, μοτοσικλετών, ποδηλάτων) όσο και για τους πεζούς. Τα αποτελέσματα της παρούσας εργασίας μπορούν να αξιοποιηθούν από τους φορείς αστυνόμευσης για την ανίχνευση της χρήσης του κινητού μέσα από την παρατήρηση διαφόρων οδηγικών μεγεθών.
- ✓ Πρώτα από όλα, απαιτείται ένα σχέδιο δράσης, μέσω **εκστρατειών ενημέρωσης** σε όλα τα μέσα ενημέρωσης και το διαδίκτυο, ώστε να επιτευχθεί η αλλαγή της νοοτροπίας των οδηγών και να αναδειχθούν οι κίνδυνοι που εγκυμονεί η χρήση κινητού τηλεφώνου κατά την οδήγηση.
- ✓ Ακόμα, θα πρέπει να υπάρξει συντονισμένη αντιμετώπιση του προβλήματος και από την πολιτεία αλλά και πολύ περισσότερο από τους **εκπαιδευτικούς φορείς**, έτσι ώστε να δημιουργηθεί το υπόβαθρο για τη δημιουργία υπεύθυνων και υποδειγματικών οδηγών από μικρή ηλικία. Είναι, λοιπόν, απαραίτητη η ανάπτυξη εκπαιδευτικών προγραμμάτων αλλά και η ένταξη μαθημάτων οδικής ασφάλειας στο πλαίσιο των σχολικών δραστηριοτήτων με ειδική αναφορά στους κινδύνους της χρήσης του κινητού τηλεφώνου κατά την οδήγηση.
- ✓ Τέλος, καλό θα ήταν οι ασφαλιστικές εταιρείες οχημάτων να επιβραβεύουν τους προσεκτικούς οδηγούς με **μειωμένα ασφάλιστρα** για όσους δεν εμπλέκονται σε οδικά ατυχήματα. Αυτή η τακτική ενδεχομένως θα δώσει οικονομικό κίνητρο σε όλους τους οδηγούς με στόχο τη μείωση της απόσπασης προσοχής και συνεπώς τη μείωση των ατυχημάτων.

6.4 Προτάσεις για περαιτέρω έρευνα

Για την περαιτέρω μελέτη του αντικειμένου της παρούσας Διπλωματικής Εργασίας, ενδιαφέρον θα παρουσίαζε η διερεύνηση των παρακάτω:

6. Συμπεράσματα

- ✓ Η παρατήρηση των ίδιων μεταβλητών σε **μεγαλύτερο δείγμα** οδηγών, θα παρουσίαζε αρκετό ενδιαφέρον. Συγκεκριμένα, όσο περισσότεροι οδηγοί συμμετέχουν στο πείραμα, τόσο πιο αξιόπιστα αποτελέσματα προκύπτουν και ίσως αυτό να έδινε τη δυνατότητα ανάπτυξης μοντέλων με ισχυρότερη συσχέτιση μεταξύ των μεταβλητών.
- ✓ Όσον αφορά στη μεθοδολογία ανάλυσης ενδιαφέρον θα παρουσίαζε η εφαρμογή **διαφορετικών μεθόδων στατιστικής ανάλυσης**, όπως η ανάλυση με χρονοσειρές ή η ανάλυση ομαδοποίησης και η ανάλυση παραγόντων με σκοπό την ομαδοποίηση των οδηγών ανάλογα με τη συμπεριφορά τους.
- ✓ Ενδιαφέρον θα παρουσίαζε μία ανάλυση η οποία θα στηριζόταν σε ακόμα **περισσότερα δεδομένα, όπως το φύλλο των οδηγών, η ηλικία, η ψυχολογική τους κατάσταση και τα χαρακτηριστικά του οχήματος**(κινητήριος δύναμη, ίπποι, ηλικία κλπ.). Ακόμα, θα ήταν σκόπιμη η απόκτηση στοιχείων για την ταχύτητα των οδηγών και στους επιμέρους τύπους οδού. Έτσι, αυτά τα επιπλέον στοιχεία φαίνεται ότι θα οδηγήσουν σε πολύ πιο αξιόπιστα και αντικειμενικά μοντέλα.
- ✓ Ενδιαφέρουσες, επίσης, θα ήταν και έρευνες αντίστοιχες της παρούσας οι οποίες θα πραγματοποιηθούν σε **διαφορετικές συνθήκες κυκλοφορίας και σε άλλα οδικά περιβάλλοντα** (υψηλή/χαμηλή κυκλοφορία, ημέρα/νύχτα, διάφορες καιρικές συνθήκες κα.), αλλά και διάφορες ομάδες οδηγών (νέοι, ηλικιωμένοι, κλπ.).
- ✓ Θα μπορούσε να πραγματοποιηθεί μια αντίστοιχη διερεύνηση για τους **οδηγούς μοτοσικλετών** για τους οποίους μάλιστα ο κίνδυνος από τη χρήση του κινητού τηλεφώνου είναι πολλαπλάσιος καθώς αποτελούν πιο ευάλωτους χρήστες της οδού σε σχέση με τους οδηγούς Ι.Χ.
- ✓ Τέλος, θα μπορούσε να διερευνηθεί η επιρροή που έχει στην συμπεριφορά του οδηγού η **παρουσία ή μη συνοδηγού και άλλων ατόμων (ειδικότερα παιδιών) εντός οχήματος, η χρήση ή μη ζώνης ασφαλείας καθώς και η κατανάλωση αλκοόλ** πριν χρησιμοποιήσει το όχημά του.

7. Βιβλιογραφία

1. Agathe Backer-Grøndahl, Fridulv Sagberg (2010), " Driving and telephoning: Relative accident risk when using hand-held and hands-free mobile phones"
2. Barbara E. Sabey and Harold Taylor (1980), "The known risks we run: The Highway", Transport and Road Research Laboratory Crowthorne, England
3. Βασιλική Αγαθαγγέλου (2016), "Συσχέτιση χαρακτηριστικών και επιδόσεων ασφαλείας του οδηγού", Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Ιούλιος 2016
4. Γιώργος Βασιλάκης (2016), "Η στάση των ευρωπαίων οδηγών απέναντι στην απόσπαση προσοχής", Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Ιούλιος 2016
5. Σιώμκος, Γ., and Α. Ι. Βασιλικοπούλου. "Εφαρμογή μεθόδων ανάλυσης στην έρευνα αγοράς." *Εκδόσεις Σταμούλης, Αθήνα, Κεφάλαιο 7* (2005): 285-312.
6. David Shinar, Richard Compton (2004), "Aggressive driving: an observational study of driver, vehicle, and situational variables"
7. Dingus, Thomas A, Hankey, Jonathan M, Antin, Jonathan F, Lee, Suzanne E, Eichelberger, Lisa, Stulce, Kelly E, McGraw, Doug, Perez, Miguel, Stowe, Loren (2015), " Naturalistic Driving Study: Technical Coordination and Quality Control", Virginia Tech Transportation Institute (VTTI)
8. D.I. Tselentis, G. Yannis, E.I. Vlahogianni (2017), "Innovative motor insurance schemes: A review of current practices and emerging challenges." *Accident Analysis & Prevention*, p:139-148
9. Eric R. Dahlen, Bryan D. Edwards, Travis Tubré, Michael J. Zyphur, Christopher R. Warren (2012), "Taking a look behind the wheel: An investigation into the personality predictors of aggressive driving"
10. G. Yannis, Papadimitriou E., Papantoniou P. (2014), " Distracted driving and mobile phone use: Overview of impacts and countermeasures"
11. Jan Törnros, Anne Bolling (2006), " Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments"
12. Κοκολάκης Γ., Σπηλιώτης Ι., (1999), "Εισαγωγή στη Θεωρία Πιθανοτήτων και Στατιστική", *Εκδόσεις Συμεών, Αθήνα*
13. Maria Papadakaki, Georgia Tzamalouka, Charalampos Gnardellis, Timo Juhani Lajunen, Joannes Chliaoutakis (2016), " Driving performance while using a mobile phone: A simulation study of Greek professional drivers"
14. Md. Mazharul Haque, Simon Washington (2015), " The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model"

7. Βιβλιογραφία

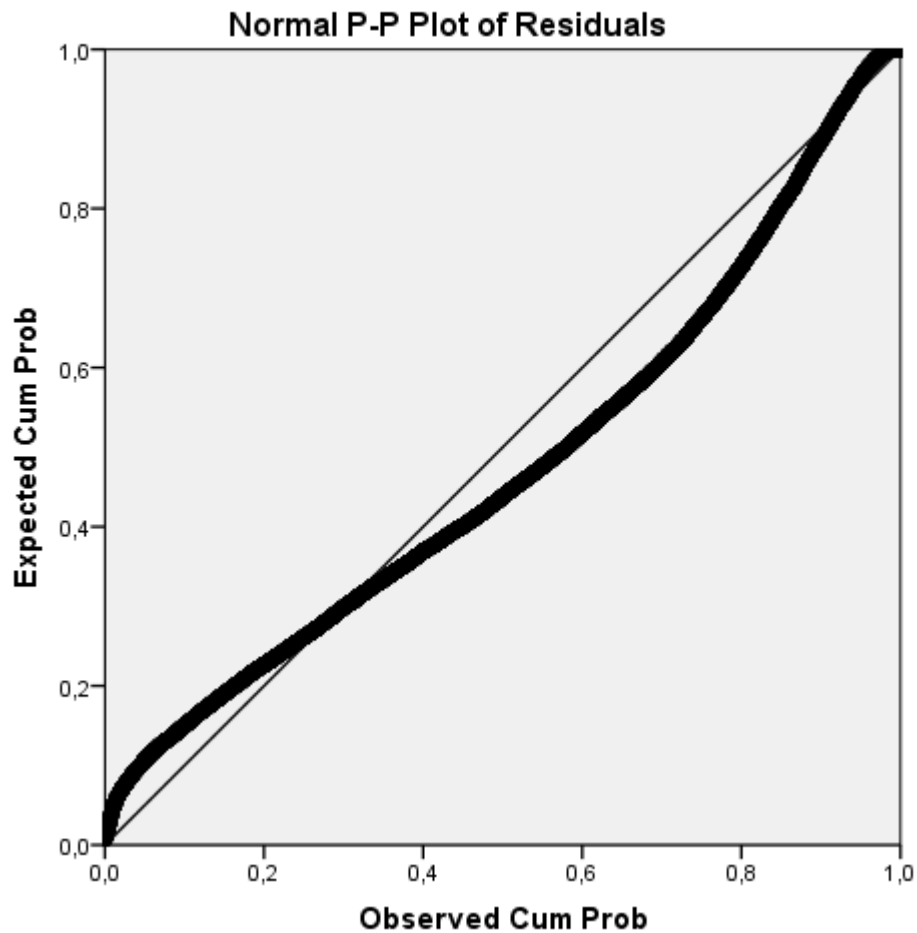
15. Mohammad Saifuzzaman, Md. Mazharul Haque, Zuduo Zheng, Simon Washington (2015), "Impact of mobile phone use on car-following behaviour of young drivers"
16. Ohta, Tohru, and Shouji Nakajima (1994), "Development of a driving data recorder." JSAE Review, p:255-258
17. Patten CJ, Kircher A, Ostlund J, Nilsson L (2004), "Using mobile telephones: cognitive workload and attention resource allocation"
18. Prato, Carlo Giacomo, et al. (2010), "Modeling the behavior of novice young drivers during the first year after licensure." Accident Analysis & Prevention, p:480-486
19. Rakauskas, M.E., Gugerty, L.J., & Ward, N.J. (2004), Effects of naturalistic cell phone conversations on driving performance
20. Toledo, Tomer, Oren Musicant, and Tsippy Lotan (2008), "In-vehicle data recorders for monitoring and feedback on drivers' behavior." Transportation Research Part C: Emerging Technologies, p: 320-331
21. Vaiana, Rosolino, et al. (2014), "Driving behavior and traffic safety: an acceleration based safety evaluation procedure for smartphones." Modern Applied Science, p:88
22. West, Robert, et al. (1993) "Direct observation of driving, self reports of driver behaviour, and accident involvement." Ergonomics 36.5: 557-567
23. Zaldivar, Jorge, et al. (2011), "Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones." Local Computer Networks (LCN)
24. NTUA Road Safety Observatory (NRSO), 2017, www.nrso.ntua.gr/european-commission-road-safety-2016-encouraging-results-2017 (ανάκτηση 17.5.2017)
25. Ευρωπαϊκή Επιτροπή, 2017, ec.europa.eu/transport/road_safety (ανάκτηση 17.5.2017)
26. Ελληνική Στατιστική Υπηρεσία (ΕΛΣΤΑΤ), www.statistics.gr/el/statistics/-/publication/SDT03/ (ανάκτηση 17.5.2017)
27. Ελληνική Αστυνομία, www.astynomia.gr/index.php?option=ozo_content&perform=view&id=70427&Itemid=86&lang= (ανάκτηση 17.5.2017)
28. National Highway Transport Safety Administration (NHTSA), 2017, www.nhtsa.gov/risky-driving/distracted-driving (ανάκτηση 13.4.2017)
29. Ιστοσελίδα Γ. Ανδρουλάκη, Πανεπιστήμιο Πατρών, andrulakis.bma.upatras.gr/mediawiki/index.php (ανάκτηση 20.5.2017)
30. OSeven, 2017, (www.oseven.io) (ανάκτηση 29.5.2017)

31. Statnouns, Academic Innovation for Development, Μοντέλα Ανάμικτων Επιδράσεων με την R, www.statnouns.com/assets/Mixed_Effects_Models_with_R_GR.pdf (ανάκτηση 10.5.2017)
32. Wikipedia, en.wikipedia.org/wiki/Confusion_matrix (ανάκτηση 10.5.2017)
33. Wikipedia, [el.wikipedia.org/wiki/Παλινδρόμηση_\(στατιστική\)](http://el.wikipedia.org/wiki/Παλινδρόμηση_(στατιστική)) (ανάκτηση 10.5.2017)
34. The United Nations and Road Safety, 2015, www.un.org (ανάκτηση 13.4.2017)
35. WHO (World Health Organization), Global Status Report on Road Safety, 2015, www.who.int/violence_injury_prevention/road_safety_status/2015/en (ανάκτηση 17.5.2017)

8. Παράρτημα

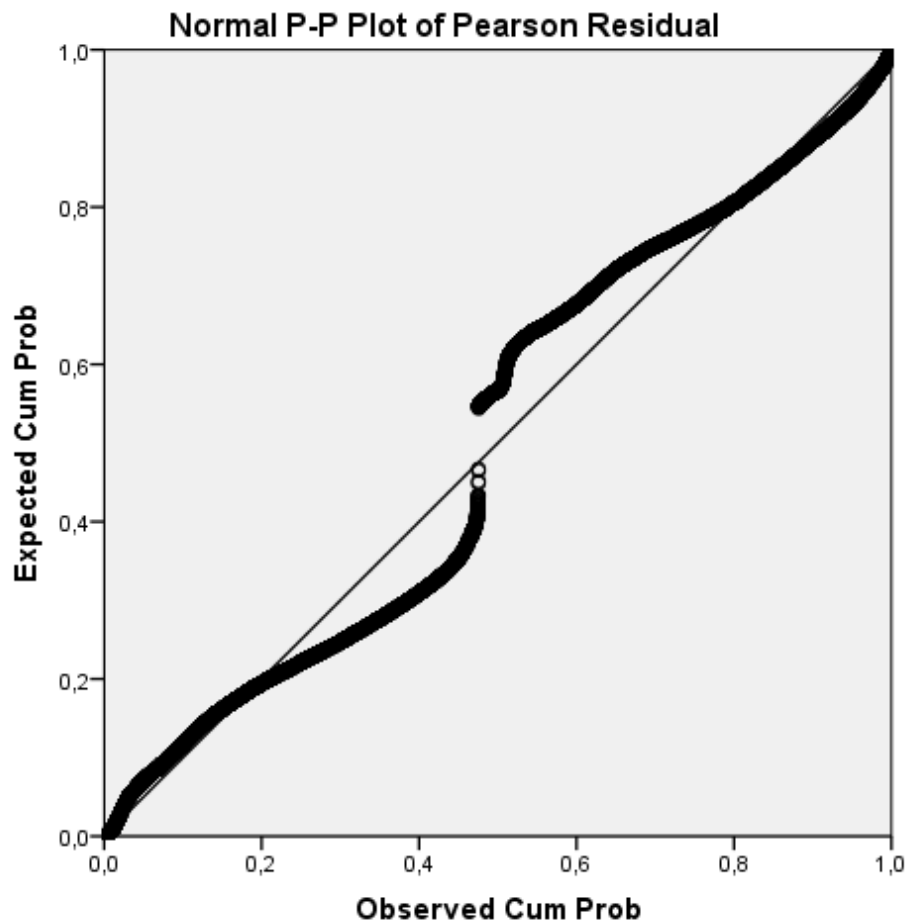
Διαγράμματα ελέγχου κανονικότητας καταλοίπων

Εξαρτημένη μεταβλητή: htot/totaldist



Διάγραμμα Π1. Διάγραμμα ελέγχου κανονικότητας καταλοίπων για το μοντέλο μεικτής γραμμικής παλινδρόμησης

Πιθανότητα ο οδηγός να χρησιμοποιεί το κινητό τηλέφωνο



Διάγραμμα Π2. Διάγραμμα ελέγχου κανονικότητας καταλοίπων για το μοντέλο μεικτής
δυσαδικής λογιστικής παλινδρόμησης