



Department of Transportation  
Planning and Engineering  
School of Civil Engineering  
National Technical University of Athens

Διερεύνηση της επιρροής της εθνικότητας του οδηγού στην  
οδική ασφάλεια

Διπλωματική εργασία



**Κίμων Π. Αμερικόνας**

Επιβλέπων: Ελεονώρα Παπαδημητρίου, Επ. Καθηγήτρια Ε.Μ.Π

Αθήνα, Μάρτιος 2026

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στην κα. Ελεονώρα Παπαδημητρίου, Επ. Καθηγήτρια της Σχολής Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου, για την εμπιστοσύνη που μου έδειξε αναθέτοντάς μου την παρούσα Διπλωματική Εργασία, καθώς και για την καθοδήγηση, την υποστήριξη και την άριστη συνεργασία μας σε όλα τα στάδια της εκπόνησής της. Την ευχαριστώ ιδιαίτερα για τις πολύτιμες συμβουλές και τις γνώσεις που μου προσέφερε καθ' όλη τη διάρκεια της εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου και τα αδέρφια μου για την αγάπη, την ηθική και υλική υποστήριξή τους καθ' όλη τη διάρκεια των σπουδών μου, καθώς και τους φίλους μου για τις όμορφες και δύσκολες στιγμές που μοιραστήκαμε όλα αυτά τα χρόνια.

# Διερεύνηση της επιρροής της εθνικότητας του οδηγού στην οδική ασφάλεια

Κίμων Αμερικάνος

Επιβλέπων: Ελεονώρα Παπαδημητρίου, Επ. Καθηγήτρια Ε.Μ.Π

## ΣΥΝΟΨΗ

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η διερεύνηση της σχέσης μεταξύ της εθνικότητας του οδηγού και της σοβαρότητας των οδικών ατυχημάτων στην Ελλάδα. Για τον σκοπό αυτό αξιοποιήθηκαν δεδομένα οδικών ατυχημάτων από την ευρωπαϊκή βάση δεδομένων CARE καθώς και από την Ελληνική Στατιστική Αρχή (ΕΛΣΤΑΤ), τα οποία αφορούν την περίοδο 2019–2023. Αρχικά πραγματοποιήθηκε η συλλογή, επεξεργασία και κωδικοποίηση των δεδομένων ώστε να δημιουργηθεί μία βάση δεδομένων κατάλληλη για στατιστική ανάλυση. Στη βάση αυτή περιλαμβάνονται πληροφορίες σχετικά με την εθνικότητα, την ηλικία και το φύλο των οδηγών, καθώς και χαρακτηριστικά του οδικού περιβάλλοντος, όπως ο τύπος οδού, οι συνθήκες φωτισμού και ο τύπος του εμπλεκόμενου οχήματος. Στη συνέχεια πραγματοποιήθηκε στατιστική ανάλυση των δεδομένων με τη χρήση μεθόδων περιγραφικής στατιστικής και ανάλυσης διακύμανσης (ANOVA), καθώς και ανάπτυξη στατιστικών μοντέλων για την ανάλυση της συχνότητας και της σοβαρότητας των οδικών ατυχημάτων. Τα αποτελέσματα της ανάλυσης έδειξαν ότι η εθνικότητα των οδηγών σχετίζεται με διαφοροποιήσεις στη συχνότητα εμπλοκής σε οδικά ατυχήματα, ενώ παράγοντες όπως η ηλικία, το φύλο και ο τύπος του οχήματος επηρεάζουν σημαντικά τη σοβαρότητα των συνεπειών του ατυχήματος.

Λέξεις κλειδιά: Οδική ασφάλεια, οδικά ατυχήματα, εθνικότητα οδηγού, σοβαρότητα ατυχημάτων, στατιστική ανάλυση

# **Investigation of the Impact of Driver Nationality on Road Safety**

Kimon Amerikanos

Supervisor: Eleonora Papadimitriou, A. Professor NTUA

## **ABSTRACT**

The aim of this Diploma Thesis is to investigate the relationship between driver nationality and the severity of road accidents in Greece. For this purpose, road accident data obtained from the European CARE database as well as from the Hellenic Statistical Authority (ELSTAT) were utilized, covering the period 2019–2023. Initially, the data were collected, processed and coded in order to create a database suitable for statistical analysis. The database includes information regarding driver characteristics such as nationality, age and gender, as well as characteristics of the road environment, including road type, lighting conditions and vehicle type. Subsequently, statistical analysis was conducted using descriptive statistics and analysis of variance (ANOVA), while statistical models were also developed for the analysis of accident frequency and severity. The results of the analysis indicate that driver nationality is associated with differences in accident involvement frequency, while factors such as driver age, gender and vehicle type significantly influence accident severity.

Keywords: Road safety, road accidents, driver nationality, accident severity, statistical analysis

## Περιεχόμενα

1. Εισαγωγή .....	7
1.1 Γενική ανασκόπηση.....	7
1.2 Στόχος της Διπλωματικής Εργασίας .....	8
1.3 Μεθοδολογία .....	8
1.4 Δομή της Διπλωματικής Εργασίας.....	10
<b>2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ .....</b>	<b>12</b>
2.1 Εισαγωγή .....	12
2.2 Συναφείς έρευνες και μεθοδολογίες .....	13
2.2.1 Έρευνες που πραγματοποιήθηκαν στην Ελλάδα .....	13
2.2.2 Έρευνες από το εξωτερικό .....	14
2.3 Σύνοψη.....	16
<b>3. Θεωρητικό Υπόβαθρο .....</b>	<b>19</b>
3.1 Εισαγωγή .....	19
3.2 Βασικές Έννοιες Στατιστικής .....	19
3.3 Βασικές Κατανομές .....	21
3.3.1 Κανονική Κατανομή .....	21
3.3.2 Κατανομή Poisson .....	21
3.3.3 Αρνητική Διωνυμική Κατανομή .....	22
3.4 Μαθηματικά Πρότυπα .....	23
3.4.1 Γραμμική Παλινδρόμηση.....	23
3.4.2 Ανάλυση Διακύμανσης (ANOVA).....	23
3.4.3 Πολυωνυμικό Λογιστικό Μοντέλο (Multinomial Logit Model) .....	24
3.5 Έλεγχοι Στατιστικής Σημαντικότητας και Αξιολόγηση Υποδειγμάτων....	25
3.6 Μοντέλα Καταμέτρησης Συμβάντων.....	26
3.7 Λειτουργία του ειδικού στατιστικού λογισμικού.....	28
<b>4. Συλλογή και επεξεργασία δεδομένων .....</b>	<b>30</b>
4.1 Εισαγωγή .....	30
4.2 Συλλογή στοιχείων από βάση δεδομένων .....	30

4.3 Προσδιορισμός πληθυσμού αναφοράς ανά κατηγορία υπηκοότητας.....	32
4.4 Επεξεργασία στοιχείων.....	33
4.5 Περιγραφικά Στατιστικά Στοιχεία.....	36
4.6 Ανάλυση Διακύμανσης (ANOVA) .....	41
5. Ανάπτυξη και Ανάλυση Μαθηματικών Προτύπων .....	44
5.1 Εισαγωγή .....	44
5.2 Ανάπτυξη Μοντέλων Συχνότητας Ατυχημάτων .....	44
5.3 Ανάλυση αποτελεσμάτων του μοντέλου .....	47
5.5 Ανάλυση αποτελεσμάτων του πολυωνυμικού λογιστικού μοντέλου .....	53
5.6 Σύγκριση των αποτελεσμάτων των στατιστικών μοντέλων.....	57
6. Συμπεράσματα.....	60
6.1 Σύνοψη της έρευνας .....	60
6.2 Κύρια αποτελέσματα της ανάλυσης.....	60
6.3 Συζήτηση αποτελεσμάτων .....	62
6.5 Προτάσεις για μελλοντική έρευνα .....	64
7. Βιβλιογραφία .....	65

## 1. Εισαγωγή

### 1.1 Γενική ανασκόπηση

Οι μεταφορές αποτελούν βασικό στοιχείο της σύγχρονης κοινωνίας και της οικονομικής ανάπτυξης, καθώς διευκολύνουν την κινητικότητα ανθρώπων και αγαθών και συμβάλλουν στη λειτουργία της εσωτερικής αγοράς και στη βελτίωση της ποιότητας ζωής των πολιτών. Ωστόσο, η συνεχώς αυξανόμενη κινητικότητα συνοδεύεται και από σημαντικές προκλήσεις, μία από τις σημαντικότερες εκ των οποίων είναι τα οδικά ατυχήματα και οι επιπτώσεις τους στην ανθρώπινη ζωή και την κοινωνία.

Τα οδικά ατυχήματα αποτελούν ένα από τα σημαντικότερα προβλήματα δημόσιας υγείας σε παγκόσμιο επίπεδο. Σύμφωνα με στοιχεία του Παγκόσμιου Οργανισμού Υγείας, εκατομμύρια άνθρωποι τραυματίζονται κάθε χρόνο στους δρόμους, ενώ περισσότεροι από ένα εκατομμύριο χάνουν τη ζωή τους εξαιτίας οδικών ατυχημάτων. Παρά τις σημαντικές προσπάθειες που έχουν πραγματοποιηθεί τις τελευταίες δεκαετίες για τη βελτίωση της οδικής ασφάλειας, τα οδικά ατυχήματα εξακολουθούν να αποτελούν σημαντική αιτία θανάτων και τραυματισμών.

Στην Ευρωπαϊκή Ένωση έχουν τεθεί συγκεκριμένοι στόχοι για τη μείωση των θυμάτων των οδικών ατυχημάτων, μέσω της εφαρμογής πολιτικών οδικής ασφάλειας, της βελτίωσης των οδικών υποδομών, της ενίσχυσης της επιβολής των κανόνων κυκλοφορίας και της προώθησης ασφαλέστερων οχημάτων. Παρά τη σημαντική πρόοδο που έχει επιτευχθεί, ο αριθμός των οδικών ατυχημάτων και των θυμάτων τους εξακολουθεί να παραμένει σημαντικός, γεγονός που καθιστά αναγκαία τη συνεχή διερεύνηση των παραγόντων που επηρεάζουν την οδική ασφάλεια.

Η συχνότητα και η σοβαρότητα των οδικών ατυχημάτων επηρεάζονται από ένα πλήθος παραγόντων που σχετίζονται με τα χαρακτηριστικά του οδηγού, του οχήματος, του οδικού περιβάλλοντος και των κυκλοφοριακών συνθηκών. Παράγοντες όπως η ηλικία και το φύλο του οδηγού, η εμπειρία οδήγησης, η ταχύτητα, οι καιρικές συνθήκες, ο τύπος του οχήματος και οι συνθήκες φωτισμού μπορούν να επηρεάσουν σημαντικά τόσο την πιθανότητα εμφάνισης ενός ατυχήματος όσο και τη σοβαρότητα των συνεπειών του.

Ένας παράγοντας που τα τελευταία χρόνια έχει προσελκύσει αυξανόμενο ερευνητικό ενδιαφέρον είναι η εθνικότητα ή η προέλευση των οδηγών. Σε πολλές χώρες, ιδιαίτερα σε χώρες με έντονη τουριστική δραστηριότητα ή σημαντικές μετακινήσεις πληθυσμών, στο οδικό δίκτυο συνυπάρχουν οδηγοί διαφορετικών εθνικοτήτων με διαφορετικές οδηγικές συνήθειες, επίπεδα εμπειρίας και βαθμό εξοικείωσης με το τοπικό οδικό περιβάλλον. Η αλληλεπίδραση αυτή μπορεί να επηρεάζει τη συμπεριφορά των οδηγών και ενδεχομένως να συμβάλλει στη διαφοροποίηση των χαρακτηριστικών των οδικών ατυχημάτων.

Η Ελλάδα αποτελεί χαρακτηριστικό παράδειγμα χώρας όπου στο οδικό δίκτυο κινούνται τόσο ημεδαποί όσο και αλλοδαποί οδηγοί. Η έντονη τουριστική δραστηριότητα, καθώς και η παρουσία κατοίκων διαφορετικών εθνικοτήτων λόγω αύξησης της μετανάστευσης, δημιουργούν ένα ιδιαίτερα σύνθετο περιβάλλον κυκλοφορίας. Στο πλαίσιο αυτό, η διερεύνηση της σχέσης μεταξύ της εθνικότητας των

οδηγών και των χαρακτηριστικών των οδικών ατυχημάτων παρουσιάζει ιδιαίτερο ενδιαφέρον για την κατανόηση των παραγόντων που επηρεάζουν την οδική ασφάλεια.

Η μελέτη των διαφοροποιήσεων που ενδέχεται να εμφανίζονται μεταξύ διαφορετικών ομάδων οδηγών μπορεί να συμβάλει στη βελτίωση της κατανόησης του φαινομένου των οδικών ατυχημάτων και να υποστηρίξει τον σχεδιασμό αποτελεσματικότερων πολιτικών οδικής ασφάλειας. Για τον λόγο αυτό, η διερεύνηση των χαρακτηριστικών των οδικών ατυχημάτων σε σχέση με την εθνικότητα των οδηγών αποτελεί ένα σημαντικό πεδίο έρευνας στον τομέα της οδικής ασφάλειας.

## **1.2 Στόχος της Διπλωματικής Εργασίας**

Στόχος της παρούσας διπλωματικής εργασίας είναι η διερεύνηση της σχέσης μεταξύ της εθνικότητας των οδηγών και των χαρακτηριστικών των οδικών ατυχημάτων στην Ελλάδα, με ιδιαίτερη έμφαση τόσο στη συχνότητα εμφάνισης των παθόντων όσο και στη σοβαρότητα των συνεπειών τους.

Στο πλαίσιο αυτό εξετάζεται κατά πόσο παράγοντες που σχετίζονται με τα χαρακτηριστικά των οδηγών, του οχήματος και του οδικού περιβάλλοντος επηρεάζουν την πιθανότητα εμφάνισης οδικών ατυχημάτων και τη σοβαρότητα των τραυματισμών. Ιδιαίτερη έμφαση δίνεται στη μεταβλητή της εθνικότητας των οδηγών, προκειμένου να διερευνηθεί αν παρατηρούνται διαφοροποιήσεις μεταξύ οδηγών διαφορετικής προέλευσης όσον αφορά την εμπλοκή τους σε οδικά ατυχήματα.

Τα αποτελέσματα της ανάλυσης μπορούν να συμβάλουν στην καλύτερη κατανόηση του φαινομένου των οδικών ατυχημάτων και να αποτελέσουν χρήσιμη βάση για τον σχεδιασμό πολιτικών και μέτρων που αποσκοπούν στη βελτίωση της οδικής ασφάλειας.

## **1.3 Μεθοδολογία**

Για την εκπόνηση της παρούσας Διπλωματικής Εργασίας και για την επίτευξη του στόχου της ακολουθήθηκε συγκεκριμένη ερευνητική μεθοδολογία, η οποία περιλαμβάνει διαδοχικά στάδια συλλογής, επεξεργασίας και ανάλυσης δεδομένων.

Αρχικά πραγματοποιήθηκε ο καθορισμός του αντικειμένου της έρευνας και του στόχου της Διπλωματικής Εργασίας. Στη συνέχεια ακολούθησε εκτενής βιβλιογραφική ανασκόπηση, κατά την οποία μελετήθηκαν σχετικές επιστημονικές έρευνες τόσο από την ελληνική όσο και από τη διεθνή βιβλιογραφία. Μέσω της βιβλιογραφικής ανασκόπησης κατέστη δυνατή η κατανόηση των βασικών εννοιών που σχετίζονται με την οδική ασφάλεια, καθώς και η επιλογή των κατάλληλων μεθόδων στατιστικής ανάλυσης για τη διερεύνηση του υπό εξέταση ζητήματος.

Στη συνέχεια πραγματοποιήθηκε η συλλογή των απαραίτητων στοιχείων για την ανάλυση των οδικών ατυχημάτων. Τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα εργασία προέρχονται από την ευρωπαϊκή βάση δεδομένων οδικών ατυχημάτων CARE (Community database on Accidents on the Roads in Europe). Η

βάση αυτή περιλαμβάνει αναλυτικά στοιχεία για τα οδικά ατυχήματα που καταγράφονται στα κράτη μέλη της Ευρωπαϊκής Ένωσης. Για τις ανάγκες της παρούσας έρευνας χρησιμοποιήθηκαν δεδομένα που αφορούν οδικά ατυχήματα τα οποία σημειώθηκαν στην Ελλάδα κατά την περίοδο 2019–2023.

Μετά τη συλλογή των δεδομένων πραγματοποιήθηκε η επεξεργασία και η οργάνωσή τους, με σκοπό τη δημιουργία μίας ενιαίας βάσης δεδομένων κατάλληλης για στατιστική ανάλυση. Η επεξεργασία των στοιχείων πραγματοποιήθηκε με τη χρήση του λογισμικού Microsoft Excel, όπου πραγματοποιήθηκε ο καθαρισμός των δεδομένων, η ομαδοποίηση των μεταβλητών και η δημιουργία της τελικής μορφής του συνόλου δεδομένων που χρησιμοποιήθηκε στην ανάλυση.

Στη συνέχεια πραγματοποιήθηκε η στατιστική ανάλυση των δεδομένων με τη χρήση του λογισμικού RStudio. Αρχικά εφαρμόστηκε περιγραφική στατιστική ανάλυση με σκοπό την κατανόηση της κατανομής των βασικών μεταβλητών του συνόλου δεδομένων και την αρχική διερεύνηση των χαρακτηριστικών των οδικών ατυχημάτων.

Σε επόμενο στάδιο εφαρμόστηκε ανάλυση διακύμανσης (ANOVA), προκειμένου να εξεταστεί αν υπάρχουν στατιστικά σημαντικές διαφορές στους δείκτες οδικών ατυχημάτων μεταξύ των διαφορετικών κατηγοριών εθνικότητας οδηγών. Για τον εντοπισμό των ομάδων που παρουσιάζουν στατιστικά σημαντικές διαφοροποιήσεις εφαρμόστηκε επιπλέον η μέθοδος πολλαπλών συγκρίσεων Tukey.

Τέλος, για τη διερεύνηση των παραγόντων που επηρεάζουν τη συχνότητα και τη σοβαρότητα των οδικών ατυχημάτων αναπτύχθηκαν δύο στατιστικά μοντέλα. Συγκεκριμένα, για την ανάλυση της συχνότητας εμφάνισης των ατυχημάτων εκτιμήθηκαν μοντέλα καταμέτρησης. Αρχικά εφαρμόστηκε μοντέλο παλινδρόμησης Poisson, ενώ λόγω της παρουσίας υπερδιασποράς στα δεδομένα εκτιμήθηκε και μοντέλο Αρνητικής Διωνυμικής κατανομής (Negative Binomial). Ενώ για την ανάλυση της σοβαρότητας των συνεπειών των ατυχημάτων χρησιμοποιήθηκε πολυωνυμικό λογιστικό μοντέλο (Multinomial Logit Model). Η εφαρμογή των μοντέλων αυτών επέτρεψε την ποσοτική διερεύνηση της επίδρασης διαφόρων παραγόντων στα οδικά ατυχήματα και την εξαγωγή των αντίστοιχων συμπερασμάτων.

Στη συνέχεια παρουσιάζονται συνοπτικά τα βασικά στάδια της μεθοδολογίας που ακολουθήθηκαν για την εκπόνηση της παρούσας Διπλωματικής Εργασίας:



Διάγραμμα 1.1 Διάγραμμα ροής σταδίων εκπόνησης Διπλωματικής Εργασίας

## 1.4 Δομή της Διπλωματικής Εργασίας

Στην ενότητα αυτή παρουσιάζεται συνοπτικά η δομή της παρούσας Διπλωματικής Εργασίας, προκειμένου να διευκολυνθεί η κατανόηση της οργάνωσης και της λογικής ανάπτυξης των επιμέρους κεφαλαίων.

Το πρώτο κεφάλαιο αποτελεί την εισαγωγή της εργασίας. Στο κεφάλαιο αυτό παρουσιάζεται το γενικό πλαίσιο του προβλήματος των οδικών ατυχημάτων και αναδεικνύεται η σημασία της διερεύνησης των παραγόντων που επηρεάζουν την οδική ασφάλεια. Παράλληλα παρουσιάζεται ο στόχος της Διπλωματικής Εργασίας, η μεθοδολογία που ακολουθήθηκε για την εκπόνησή της και η συνολική δομή της.

Στο δεύτερο κεφάλαιο παρουσιάζεται η βιβλιογραφική ανασκόπηση της έρευνας. Στο κεφάλαιο αυτό αναλύονται σχετικές μελέτες από την ελληνική και τη διεθνή βιβλιογραφία που εξετάζουν τη σχέση μεταξύ των χαρακτηριστικών των οδηγών με έμφαση στην εθνικότητα και των οδικών ατυχημάτων. Μέσα από τη σύγκριση και τη σύνθεση των αποτελεσμάτων των προηγούμενων ερευνών

αναδεικνύονται τα βασικά συμπεράσματα της βιβλιογραφίας καθώς και τα ερευνητικά κενά που επιχειρεί να καλύψει η παρούσα εργασία.

Στο τρίτο κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο της Διπλωματικής Εργασίας. Στο κεφάλαιο αυτό αναλύονται οι βασικές έννοιες της στατιστικής ανάλυσης καθώς και οι στατιστικές κατανομές και τα μαθηματικά πρότυπα που χρησιμοποιούνται στην ανάλυση των οδικών ατυχημάτων. Παράλληλα παρουσιάζονται οι στατιστικοί έλεγχοι που χρησιμοποιούνται για την αξιολόγηση των μοντέλων.

Το τέταρτο κεφάλαιο αφορά τη συλλογή και επεξεργασία των δεδομένων που χρησιμοποιήθηκαν στην παρούσα έρευνα. Περιγράφεται η διαδικασία άντλησης των στοιχείων από τη βάση δεδομένων CARE, καθώς και τα στάδια διαγραφής, κωδικοποίησης και ομαδοποίησης των μεταβλητών που οδήγησαν στη δημιουργία της τελικής βάσης δεδομένων. Επιπλέον παρουσιάζονται τα βασικά περιγραφικά στατιστικά στοιχεία του συνόλου δεδομένων.

Στο πέμπτο κεφάλαιο παρουσιάζεται η ανάπτυξη και η ανάλυση των στατιστικών μοντέλων που χρησιμοποιήθηκαν στην παρούσα εργασία. Αρχικά αναπτύσσεται το μοντέλο παλινδρόμησης Poisson για την ανάλυση της συχνότητας εμφάνισης των οδικών ατυχημάτων. Στη συνέχεια παρουσιάζεται το πολυωνυμικό λογιστικό μοντέλο που χρησιμοποιείται για την ανάλυση της σοβαρότητας των ατυχημάτων. Τέλος πραγματοποιείται σύγκριση των αποτελεσμάτων των δύο μοντέλων.

Στο έκτο κεφάλαιο παρουσιάζονται τα βασικά συμπεράσματα της Διπλωματικής Εργασίας. Στο κεφάλαιο αυτό γίνεται σύνοψη των σημαντικότερων ευρημάτων της ανάλυσης, συζητούνται οι περιορισμοί της έρευνας και διατυπώνονται προτάσεις για μελλοντική έρευνα στον τομέα της οδικής ασφάλειας.

## 2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

### 2.1 Εισαγωγή

Το παρόν κεφάλαιο αφορά στη βιβλιογραφική ανασκόπηση που πραγματοποιήθηκε για τις ανάγκες της παρούσας Διπλωματικής Εργασίας. Παρουσιάζονται αποτελέσματα από συναφείς έρευνες που έχουν διεξαχθεί τόσο στην Ελλάδα όσο και διεθνώς, οι οποίες αφορούν στη συσχέτιση των οδικών ατυχημάτων με τα χαρακτηριστικά του οδηγού και ειδικότερα με την εθνικότητα ή την προέλευσή του. Παράλληλα, εξετάζονται μελέτες που εστιάζουν στη μη εξοικείωση του οδηγού με το τοπικό οδικό περιβάλλον η οποία μπορεί να επηρεάζει την πιθανότητα εμπλοκής σε ατύχημα ή/και τη σοβαρότητα των συνεπειών του.

Για κάθε έρευνα γίνεται συνοπτική παρουσίαση της μεθοδολογίας που χρησιμοποιήθηκε καθώς και των συμπερασμάτων στα οποία κατέληξε. Οι συγκεκριμένες έρευνες επιλέχθηκαν να μελετηθούν καθώς είτε το αντικείμενο είτε η μεθοδολογία τους παρουσιάζουν συνάφεια με τα υπό μελέτη ζητήματα, όπως αυτά τέθηκαν στο προηγούμενο κεφάλαιο.

Τέλος, πραγματοποιείται σύνθεση των αποτελεσμάτων και κριτική αξιολόγησή τους, ώστε να αναδειχθούν οι βασικές τάσεις της βιβλιογραφίας αλλά και τα κενά γνώσης που καθιστούν αναγκαία την περαιτέρω διερεύνηση του ζητήματος στην Ελλάδα.

Στη συνέχεια παρατίθενται οι ακόλουθες έρευνες με την εξής σειρά:

- **Yannis, Golias & Papadimitriou (2002)** — *Accident risk of foreign drivers in various road environments*
- **Alogaili & Mannering (2020)** — *Unobserved heterogeneity and the effects of driver nationality on crash injury severities in Saudi Arabia*
- **Harootunian, Williams, Mackenzie & Noyce (2014)** — *Odds of fault and factors for out-of-state drivers in crashes in four states of the USA*
- **Leviäkangas (1998)** — *Accident risk of foreign drivers: the case of Russian drivers in South-Eastern Finland*
- **Shawky, Kishta, Garib & Almazrouei (2017)** — *Relationship between socio-demographic of drivers and traffic violations and crashes involvements*
- **Nævestad, Phillips & Elvebakk (2017)** — *Internationalisation in road transport of goods in Norway: Safety Outcomes, Risk Factors and Policy Implications*
- Skliami Aikaterini — Διπλωματική εργασία (ΕΜΠ) *Συγκριτική Ανάλυση Χαρακτηριστικών Οδικών Ατυχημάτων ανά Εθνικότητα Οδηγού στην Ευρωπαϊκή Ένωση*

Οι παραπάνω μελέτες αποτελούν τον βασικό κορμό της βιβλιογραφικής ανασκόπησης και αναλύονται στις επόμενες υποενότητες του παρόντος κεφαλαίου.

## **2.2 Συναφείς έρευνες και μεθοδολογίες**

### **2.2.1 Έρευνες που πραγματοποιήθηκαν στην Ελλάδα**

Μία από τις πλέον σημαντικές μελέτες που έχουν πραγματοποιηθεί στην Ελλάδα σχετικά με την επικινδυνότητα οδηγών διαφορετικής προέλευσης είναι αυτή των καθηγητών του Εθνικού Μετσόβιου Πολυτεχνείου το 2002. Η μελέτη βασίστηκε σε δεδομένα από την εθνική βάση οδικών ατυχημάτων της Ελλάδας και στόχευσε στη διερεύνηση του κινδύνου ατυχήματος οδηγών από άλλες χώρες σε σύγκριση με τους ημεδαπούς οδηγούς, σε διαφορετικά οδικά περιβάλλοντα.

Κύρια δυσκολία της ανάλυσης αποτέλεσε η έλλειψη στοιχείων κυκλοφοριακής έκθεσης στον κίνδυνο, όπως ο αριθμός μετακινήσεων ή τα χιλιόμετρα που έχει διανύσει η κάθε ομάδα οδηγών. Για αυτόν τον λόγο εφαρμόστηκε η τεχνική επαγόμενης έκθεσης (induced exposure), η οποία χρησιμοποιείται συχνά όταν δεν υπάρχουν διαθέσιμα άμεσα δεδομένα έκθεσης. Η μέθοδος βασίζεται στην υπόθεση ότι σε ατυχήματα δύο οχημάτων ο μη υπαίτιος οδηγός μπορεί να θεωρηθεί ως τυχαίο δείγμα του πληθυσμού οδηγών που κινούνται στο συγκεκριμένο περιβάλλον.

Στη συνέχεια, εξετάστηκε ο συνδυασμένος ρόλος της προέλευσης του οδηγού και χαρακτηριστικών του οδικού περιβάλλοντος, όπως ο τύπος περιοχής (αστική ή υπεραστική), η ύπαρξη κόμβου και οι συνθήκες φωτισμού. Για αυτόν τον σκοπό εφαρμόστηκε ιεραρχική λογαριθμογραμμική ανάλυση και οι διαφορές μεταξύ των ομάδων εκφράστηκαν με λόγους σχετικών πιθανοτήτων.

Τα αποτελέσματα της έρευνας έδειξαν ότι οι οδηγοί από άλλες χώρες εμφανίζουν αυξημένο σχετικό κίνδυνο ατυχήματος σε σχέση με τους ημεδαπούς οδηγούς. Παράλληλα, αναδείχθηκε διαφοροποίηση μεταξύ τουριστών και μόνιμων κατοίκων, με τους δεύτερους να παρουσιάζουν χαμηλότερο κίνδυνο, γεγονός που υποδεικνύει ότι η εξοικείωση με το οδικό περιβάλλον και η προσαρμογή στους κανόνες κυκλοφορίας ενδέχεται να αποτελούν κρίσιμους μηχανισμούς μείωσης της επικινδυνότητας. Ιδιαίτερα σημαντικό εύρημα αποτέλεσε η εντονότερη επίδραση του παράγοντα “οδηγός από άλλη χώρα” σε κόμβους, όπου η πολυπλοκότητα του περιβάλλοντος είναι αυξημένη.

Σημαντική συνεισφορά στο υπό εξέταση αντικείμενο αποτελεί η διπλωματική εργασία της Αικατερίνης Σκλιάμη, η οποία παρουσιάζει ανάλυση οδικών ατυχημάτων σε κάποιες Ευρωπαϊκές χώρες με έμφαση στη σοβαρότητα και στη διαφοροποίηση των αποτελεσμάτων ανάλογα με την εθνικότητα και τη καταγωγή των εμπλεκομένων.

Στο πλαίσιο της εργασίας αξιοποιείται η εθνική βάση δεδομένων οδικών ατυχημάτων που περιλαμβάνει ατυχήματα με τραυματισμούς, οι οποίοι κατηγοριοποιούνται ανάλογα με την προέλευσή τους, με βασική διάκριση μεταξύ ημεδαπών και μη ημεδαπών.

Στη συνέχεια, η ανάλυση επεκτείνεται σε επιμέρους μεταβλητές που σχετίζονται με το ατύχημα, όπως οι συνθήκες φωτισμού, ο τύπος οδού και ο χρόνος (π.χ. εποχικότητα), προκειμένου να αποτυπωθούν πιθανά μοτίβα διαφοροποίησης μεταξύ των ομάδων. Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη, καθώς επιτρέπει την κατανόηση του τρόπου με τον οποίο η εθνικότητα και το καθεστώς κατοικίας δεν λειτουργούν ανεξάρτητα, αλλά σε συνδυασμό με το οδικό περιβάλλον.

Τα αποτελέσματα που παρουσιάζονται στην εργασία υποδεικνύουν ότι παρατηρούνται διαφοροποιήσεις στη σοβαρότητα των οδικών ατυχημάτων ανάλογα με την προέλευση των εμπλεκομένων.

### **2.2.2 Έρευνες από το εξωτερικό**

- **Σαουδική Αραβία**

Σε διεθνές επίπεδο, ιδιαίτερη σημασία παρουσιάζει η μελέτη των Alogaili και Mannering το 2020, η οποία εξετάζει τη σοβαρότητα τραυματισμού σε οδικά ατυχήματα σε σχέση με την εθνικότητα του οδηγού στη Σαουδική Αραβία. Η έρευνα βασίστηκε σε δεδομένα ατυχημάτων από την πόλη του Riyadh και εστίασε αποκλειστικά σε ατυχήματα ενός οχήματος, ώστε να απομονωθεί σε μεγαλύτερο βαθμό η συμπεριφορά του οδηγού από την αλληλεπίδραση μεταξύ πολλών οχημάτων.

Για τη μελέτη διαχώρισαν τους οδηγούς σε ημεδαπούς και μη ημεδαπούς, και στη συνέχεια εκτιμήθηκαν ξεχωριστά με μαθηματικά πρότυπα για κάθε ομάδα. Η εξαρτημένη μεταβλητή αφορούσε τη σοβαρότητα τραυματισμού, η οποία εκφράστηκε με κατηγορίες όπως χωρίς τραυματισμό, τραυματισμός, θάνατος.

Η ανάλυση πραγματοποιήθηκε με πολυωνυμικό λογιστικό μοντέλο τυχαίων παραμέτρων, ώστε να αποτυπωθούν διαφοροποιήσεις που δεν είναι άμεσα παρατηρήσιμες στα διαθέσιμα δεδομένα. Η συγκεκριμένη προσέγγιση θεωρείται ιδιαίτερα κατάλληλη σε αναλύσεις οδικής ασφάλειας, καθώς αναγνωρίζει ότι η επίδραση των ίδιων παραγόντων δεν είναι απαραίτητα ίδια για όλους τους οδηγούς.

Τα αποτελέσματα της μελέτης έδειξαν ότι η εθνικότητα συνδέεται με διαφοροποιημένα πρότυπα σοβαρότητας, καθώς οι παράγοντες που επηρεάζουν την έκβαση δεν είναι ίδιοι για τους οδηγούς με την υπηκοότητα της χώρας και μη. Παράλληλα, αναδείχθηκε η σημασία παραγόντων που σχετίζονται με την οδηγική συμπεριφορά, όπως η ταχύτητα και η απροσεξία, με χαρακτηριστικά οχήματος, καθώς και με χωροχρονικές συνθήκες του ατυχήματος. Η μελέτη ενισχύει την υπόθεση ότι οι διαφορές μεταξύ των οδηγών δεν οφείλονται σε έναν μόνο παράγοντα, αλλά προκύπτουν από συνδυασμό δημογραφικών, συμπεριφορών και περιβαλλοντικών παραμέτρων.

- **Ηνωμένες Πολιτείες Αμερικής**

Μία ακόμη ιδιαίτερα σημαντική προσέγγιση διαφορετική από τις προηγούμενες, η οποία δεν εξετάζει την εθνικότητα με την κλασική έννοια αλλά την έλλειψη εξοικείωσης με το τοπικό περιβάλλον, παρουσιάζεται στη μελέτη των Harootunian et al.

Στη συγκεκριμένη εργασία αναλύθηκαν τα δεδομένα ατυχημάτων από τέσσερις Πολιτείες των ΗΠΑ, με στόχο να διερευνηθεί κατά πόσο οι οδηγοί εκτός πολιτείας έχουν αυξημένη πιθανότητα υπαιτιότητας σε σχέση με τους ντόπιους οδηγούς.

Η εξαρτημένη μεταβλητή της ανάλυσης ήταν η υπαιτιότητα, δηλαδή αν ο οδηγός χαρακτηρίστηκε ως υπαίτιος ή μη υπαίτιος στο ατύχημα. Η επιλογή της υπαιτιότητας ως δείκτη επικινδυνότητας είναι ιδιαίτερα σημαντική, καθώς επιτρέπει την αποτύπωση της συμπεριφοράς του οδηγού ανεξάρτητα από την κυκλοφοριακή έκθεση, σε αντίθεση με τους απλούς δείκτες συμμετοχής σε ατυχήματα.

Για την εκτίμηση των παραγόντων εφαρμόστηκαν λογιστικά μοντέλα παλινδρόμησης, ενώ πραγματοποιήθηκε ξεχωριστή ανάλυση για ατυχήματα ενός

οχήματος και ατυχήματα δύο οχημάτων. Παράλληλα, εξετάστηκαν παράγοντες που σχετίζονται με το οδικό περιβάλλον και τις συνθήκες του ατυχήματος, όπως οι καιρικές συνθήκες, ο φωτισμός, η εποχή και ο τύπος της οδού.

Τα αποτελέσματα έδειξαν ότι οι οδηγοί εκτός πολιτείας παρουσιάζουν αυξημένες πιθανότητες υπαιτιότητας, κυρίως σε ατυχήματα ενός οχήματος. Το εύρημα αυτό είναι ιδιαίτερα σημαντικό, καθώς συνδέει άμεσα τη μη εξοικείωση με αυξημένο κίνδυνο, χωρίς να απαιτείται η έννοια της εθνικότητας. Παράλληλα, διαπιστώθηκε ότι οι παράγοντες που αυξάνουν την υπαιτιότητα δεν εμφανίζουν απόλυτη ομοιομορφία μεταξύ πολιτειών, γεγονός που υποδεικνύει ότι οι τοπικές συνθήκες και τα χαρακτηριστικά του οδικού δικτύου μπορούν να επηρεάζουν σημαντικά το τελικό αποτέλεσμα.

Κάτι που είναι πολύ ενδιαφέρον καθώς δείχνει ότι διαφοροποιήσεις μπορούν να προκύψουν ακόμη και μεταξύ οδηγών ίδιας εθνικότητας. Επομένως, η διερεύνηση οδηγών διαφορετικής εθνικότητας αποκτά ακόμη μεγαλύτερο ενδιαφέρον ως προς τη σοβαρότητα των ατυχημάτων.

- **Φινλανδία**

Μία κλασική μελέτη σχετικά με τον κίνδυνο ατυχήματος οδηγών που εισέρχονται σε μία χώρα από γειτονικό κράτος είναι αυτή του Leviäkangas (1998), η οποία εξετάζει τον κίνδυνο ατυχήματος Ρώσων οδηγών στη νοτιοανατολική Φινλανδία. Η έρευνα βασίστηκε σε στατιστικά στοιχεία ατυχημάτων της αστυνομίας, καθώς και σε δεδομένα ροών κυκλοφορίας από τους συνοριακούς σταθμούς.

Το γεγονός ότι χρησιμοποιούνται στοιχεία ροών είναι σημαντικό, επειδή επιτρέπει τη σύγκριση του κινδύνου με βάση την έκθεση και όχι μόνο με βάση τον αριθμό ατυχημάτων. Με τον τρόπο αυτό, η ανάλυση προχωρά και σε σύγκριση δεικτών ατυχημάτων ανά μονάδα κυκλοφορίας.

Τα αποτελέσματα έδειξαν ότι οι οδηγοί που εισέρχονται στη χώρα από τη Ρωσία είναι πιο επικίνδυνοι στο να προκαλέσουν ατύχημα σε σχέση με τους Φινλανδούς οδηγούς. Ιδιαίτερα σημαντικό εύρημα αποτέλεσε η αύξηση του κινδύνου κατά τη χειμερινή περίοδο, γεγονός που υποδεικνύει ότι η προσαρμογή σε διαφορετικές κλιματικές και οδικές συνθήκες μπορεί να αποτελεί κρίσιμο παράγοντα.

Ο συγγραφέας υποστηρίζει ότι οι διαφορές αυτές σχετίζονται με διαφορές στην κυκλοφοριακή κουλτούρα, όπως παραδείγματος χάριν στη γνώση των κανόνων, καθώς και στις δεξιότητες οδήγησης σε χειμερινές συνθήκες. Συνεπώς, η μελέτη ενισχύει την υπόθεση ότι ο αυξημένος κίνδυνος δεν οφείλεται αποκλειστικά στην “εθνικότητα” ως έννοια, αλλά στην αλληλεπίδραση εμπειρίας, περιβάλλοντος και συμπεριφοράς, στοιχεία όμως που διαμορφώνουν την οδηγική συμπεριφορά ενός λαού.

- **Ηνωμένα Αραβικά Εμιράτα**

Η μελέτη των Shawky, Kishta, Garib & Almazrouei το 2017 εξετάζει τη σχέση κοινωνικών και δημογραφικών χαρακτηριστικών οδηγών με τις παραβάσεις και την εμπλοκή σε ατυχήματα. Η έρευνα βασίστηκε σε πολύ μεγάλο αριθμό οδηγών και αξιοποίησε δεδομένα που κάλυπταν χρονικό διάστημα οκτώ ετών.

Η ανάλυση επικεντρώθηκε στη διερεύνηση του τρόπου με τον οποίο παράγοντες όπως η ηλικία, το φύλο, και άλλα χαρακτηριστικά συνδέονται με τη συχνότητα παραβάσεων, καθώς και με την πιθανότητα εμπλοκής σε ατυχήματα. Για τον σκοπό αυτό εφαρμόστηκαν μοντέλα που χρησιμοποιούνται για να προβλέψουν μετρήσεις, τα

οποία είναι κατάλληλα για δεδομένα όπως ο αριθμός παραβάσεων ή ο αριθμός ατυχημάτων.

Τα αποτελέσματα έδειξαν ότι υπάρχει ισχυρή σχέση μεταξύ παραβατικής συμπεριφοράς και εμπλοκής σε ατυχήματα, ενώ τα υπόλοιπα χαρακτηριστικά επηρεάζουν σημαντικά τους δείκτες επικινδυνότητας. Παρότι η συγκεκριμένη μελέτη δεν εστιάζει αποκλειστικά στο ζήτημα οδηγών διαφορετικής εθνικότητας, ενισχύει τη σημασία των χαρακτηριστικών του οδηγού ως κρίσιμο στην ερμηνεία των διαφορών στην οδική ασφάλεια.

- **Νορβηγία**

Η εργασία των Nævestad, Phillips & Elvebakk εξετάζει τις επιπτώσεις της αύξησης της διεθνοποίησης στις οδικές μεταφορές εμπορευμάτων στη Νορβηγία, με έμφαση στην οδική ασφάλεια βαρέων οχημάτων. Η μελέτη αποτελεί μία διαφορετική προσέγγιση σε σχέση με τις υπόλοιπες, καθώς δεν αφορά επιβατικά οχήματα και τουριστική μετακίνηση, αλλά επαγγελματική οδήγηση και μεταφορές, που όμως ένας σημαντικός παράγοντας είναι η εθνικότητα του οδηγού.

Η εργασία βασίστηκε σε συνδυασμό δεδομένων από τα ατυχήματα και από ερωτηματολογία, με στόχο να διερευνηθούν οι παράγοντες που σχετίζονται με την αυξημένη συμμετοχή βαρέων οχημάτων με οδηγούς από άλλες χώρες σε σοβαρά ατυχήματα. Η μελέτη δείχνει ότι, ενώ τα συγκεκριμένα οχήματα αντιστοιχούν σε μικρό ποσοστό της συνολικής δραστηριότητας, εμφανίζουν δυσανάλογα αρκετά αυξημένη συμμετοχή σε ατυχήματα με τραυματισμούς.

Τα συμπεράσματα των συγγραφέων υποδεικνύουν ότι η αυξημένη επικινδυνότητα αποδίδεται κυρίως στο γεγονός ότι οι οδηγοί έχουν έλλειψη εμπειρίας στις ιδιαίτερες χειμερινές συνθήκες της χώρας, τις διαφορετικές πρακτικές εκπαίδευσης και στις απαιτήσεις της επαγγελματικής οδήγησης.

Παρότι το αντικείμενο αυτό δεν είναι ταυτίζεται με το ζήτημα της παρούσας εργασίας, η μελέτη είναι πολύ χρήσιμη καθώς ενισχύει το γενικό συμπέρασμα ότι η εξοικείωση με το περιβάλλον και η προσαρμογή στις τοπικές συνθήκες αποτελούν κρίσιμους παράγοντες του κινδύνου που έχουν οι οδηγοί από άλλες χώρες.

## **2.3 Σύνοψη**

Από τη βιβλιογραφική ανασκόπηση προκύπτει ότι η εθνικότητα όπως και η μη εξοικείωση του οδηγού με το τοπικό οδικό περιβάλλον μπορούν να συνδέονται τόσο με αυξημένο κίνδυνο εμπλοκής σε οδικό ατύχημα, όσο και με αυξημένη σοβαρότητα αυτού. Οι έρευνες συγκλίνουν στο συμπέρασμα ότι οι οδηγοί που κινούνται σε περιβάλλον διαφορετικό από αυτό στο οποίο έχουν αποκτήσει εμπειρία, κυρίως σε άλλες χώρες εμφανίζουν αυξημένη πιθανότητα σφαλμάτων, κυρίως σε περιοχές όπως με αυξημένη κυκλοφορία και συνθήκες μειωμένης ορατότητας.

Παράλληλα, διαπιστώνεται ότι η έννοια της “εθνικότητας” δεν λειτουργεί μεμονωμένα, αλλά σχετίζεται με άλλους κρίσιμους παράγοντες. Η διαφοροποίηση που παρατηρείται μεταξύ τουριστών και μόνιμων κατοίκων σε ελληνικά δεδομένα, αλλά και η αύξηση του κινδύνου σε χειμερινές συνθήκες στη Φινλανδία, υποδεικνύουν ότι

η εξοικείωση με το οδικό δίκτυο, η προσαρμογή στις τοπικές συνθήκες και η γνώση των κανόνων κυκλοφορίας αποτελούν κεντρικούς μηχανισμούς ερμηνείας.

Οι διεθνείς μελέτες που εστιάζουν στη σοβαρότητα τραυματισμού δείχνουν ότι οι διαφορές μεταξύ ομάδων οδηγών μπορούν να είναι αποτέλεσμα σύνθετων αλληλεπιδράσεων, όπου η συμπεριφορά, τα χαρακτηριστικά του οχήματος και οι άλλες συνθήκες να διαμορφώνουν τελικά την έκβαση του ατυχήματος. Ιδιαίτερη σημασία έχει το γεγονός ότι η επίδραση των ίδιων παραγόντων δεν είναι απαραίτητα ίδια για όλες τις ομάδες οδηγών, γεγονός που καθιστά αναγκαία τη χρήση μεθοδολογιών που λαμβάνουν υπόψη τις διαφορές μεταξύ οδηγών.

Τέλος, από τη σύνθεση των παραπάνω ερευνών προκύπτει ότι, παρά τη σημαντική διεθνή βιβλιογραφία, η διερεύνηση του ζητήματος στην Ελλάδα παραμένει περιορισμένη. Η ύπαρξη έντονης τουριστικής δραστηριότητας και η παρουσία οδηγών διαφορετικής προέλευσης καθιστούν αναγκαία την περαιτέρω ανάλυση της σοβαρότητας των ατυχημάτων με βάση την εθνικότητα και το καθεστώς κατοικίας, ώστε να προκύψουν τεκμηριωμένα συμπεράσματα που θα μπορούν να αξιοποιηθούν και σε επίπεδο οδικής ασφάλειας και πρόληψης.

Με σκοπό την καλύτερη οργάνωση και παρουσίαση των βασικών στοιχείων της βιβλιογραφίας, κρίθηκε σκόπιμο να συνοψιστούν συγκριτικά τα κύρια χαρακτηριστικά των επιλεγμένων ερευνών. Στον Πίνακα 2.1 παρουσιάζονται οι μελέτες που χρησιμοποιήθηκαν ως βασικός κορμός της παρούσας βιβλιογραφικής ανασκόπησης, με αναφορά στο πεδίο εφαρμογής τους, στη βάση δεδομένων που αξιοποιήθηκε, στη μεταβλητή που εξετάστηκε ως εξαρτημένη, καθώς και στη μεθοδολογία που εφαρμόστηκε. Παράλληλα, παρατίθενται συνοπτικά τα κύρια συμπεράσματα κάθε μελέτης, ώστε να διευκολυνθεί η σύγκριση και να αναδειχθούν οι κοινές τάσεις και οι διαφοροποιήσεις των αποτελεσμάτων

Μελέτη	Περίοδος μελέτης	Βάση δεδομένων	Εξαρτημένη μεταβλητή	Κόρυφα μεθοδολογία	Βασικά συμπεράσματα
2002 - Accident risk of foreign drivers in various road environments	Ελλάδα	Εθνική βάση δεδομένων	Σχετικός κίνδυνος υπαιτιότητας οδηγού	Επιρροή της τεχνικής επιρροής (induced exposure) και λογαριθμική ανάλυση (odds ratios)	Οι οδηγοί από άλλες χώρες εμφανίζουν αυξημένο κίνδυνο, με εντονότερη επίδραση σε κίβδου. Επιπλέον, προκύπτει διαφοροποίηση μεταξύ τουριστών και μόνιμων κατοίκων, με τους δεύτερους να εμφανίζουν χαμηλότερο κίνδυνο.
2020 - Διπλωματική εργασία (EMT): Ανάλυση σοβαρότητας τραυματισμών σε σχέση με την εθνικότητα και το καθεστώς κατοικίας του οδηγού	Ελλάδα	Εθνική βάση δεδομένων	Σοβαρότητα ατυχημάτων	Στατιστική ανάλυση και ανάπτυξη μεθόδων πρότυπων παλινδρόμησης	Προκύπτει ότι η σοβαρότητα και τα ποσοστά ατυχημάτων διαφοροποιούνται ανάλογα με την εθνικότητα και τον τύπο κατοικίας, με ενδείξεις αυξημένης επικινδυνότητας για τους ξένους.
2020 - Unobserved heterogeneity and the effects of driver nationality on crash injury severities in Saudi Arabia	Σαουδική Αραβία (Riyadh)	Ατυχήματα ενός οχήματος, ώστε να απομονωθεί η συμπεριφορά του οδηγού	Κατηγορία σοβαρότητας τραυματισμού	Ανάπτυξη και εκτίμηση μαθηματικού προτύπου πολυνομομής λογιστικής παλινδρόμησης με τυχασίες κωρυφαίους	Η εθνικότητα σχετίζεται με διαφοροποιημένα πρότυπα σοβαρότητας, καθώς οι παράγοντες που επηρεάζουν την έκβαση δεν είναι ίδιοι για ημεδαπούς και μη ημεδαπούς οδηγούς. Τα αποτελέσματα δείχνουν ότι η μη παρατηρούμενη ετερογένεια παίζει σημαντικό ρόλο.
2017 - Relationship between socio-demographic of drivers and traffic violations and crashes involvements	Ηνωμένα Αραβικά Εμιράτα	Βάση δεδομένων οδηγών για την περίοδο 2008-2015	Συχνότητα παραβάσεων και εμπλοκή σε ατυχήματα	Εκτίμηση μεθόδων πρότυπων αλληλικής διωνομικής παλινδρόμησης	Η εμπλοκή σε ατυχήματα συσχετίζεται ισχυρά με τον αριθμό παραβάσεων, ενώ τα κοινωνικο-δημογραφικά χαρακτηριστικά επηρεάζουν σημαντικά τους δείκτες απινδυνότητας. Η μελέτη λειτουργεί ως τεκμηρίωση της σημασίας των driver-related χαρακτηριστικών.
2017 - Internationalisation in road transport of goods in Norway: safety outcomes, risk factors and policy implications	Νορβηγία	Δεδομένα ατυχημάτων βαρέων οχημάτων και ερευνηματολόγια επαγγελματιών οδηγών	Κίνδυνος εμπλοκής σε σοβαρά ατυχήματα	Συνθετική προσέγγιση που περιλαμβάνει ανάλυση στοιχείων ατυχημάτων, συλλογή δεδομένων μέσω ερωτηματολογίων	Τα βαρέα οχήματα με οδηγούς από άλλες χώρες εμφανίζουν διαφορετικό αυξημένη συμμετοχή σε σοβαρά ατυχήματα. Η αυξημένη επικινδυνότητα συνδέεται κυρίως με περιορισμένη εμπειρία στις ιδιαίτερες χειμερινές συνθήκες της χώρας.
2014 - Odds of fault and factors for out-of-state drivers in crashes in four states of the USA	ΗΠΑ (4 πολιτείες)	Βάσεις δεδομένων ατυχημάτων ανά πολιτεία	Υπαιτιότητα οδηγού	Ανάπτυξη λογιστικών μοντέλων παλινδρόμησης και έλεγχο ανθρώπινων και περιβαλλοντικών παραγόντων	Οι οδηγοί εκτός πολιτείας παρουσιάζουν αυξημένη πιθανότητα υπαιτιότητας, κυρίως σε ατυχήματα ενός οχήματος. Παράλληλα, δεν αντιστοιχεί ένα ενιαίο μοτίβο παραγόντων που να εξηγεί σταθερά την αυξημένη υπαιτιότητα σε όλες τις πολιτείες.
1998 - Accident risk of foreign drivers: the case of Russian drivers in South-Eastern Finland	Φινλανδία	Στατιστικά στοιχεία ατυχημάτων της αυτονομίας και στοιχεία ραβών από συνοριακούς σταθμούς	Δείκτες κίνδυνου ατυχημάτων	Υπόλοιπος δείκτης ατυχημάτων για ημεδαπούς με χρήση δεδομένων κυκλοφορικής έκθεσης από τις συνοριακές ραβές	Οι οδηγοί που εισέρχονται στη χώρα από τη Ρωσία εμφανίζουν υψηλότερους δείκτες ατυχημάτων σε σχέση με τους Φινλανδούς οδηγούς. Η χειμερινή περίοδος αναδεικνύεται ως κρίσιμη για την αύξηση του κινδύνου.

Πίνακας 2.1

Από τη συγκριτική παρουσίαση των ερευνών προκύπτει ότι η εθνικότητα του οδηγού με το τοπικό οδικό περιβάλλον μπορεί να σχετίζεται με αυξημένη επικινδυνότητα, είτε ως προς την πιθανότητα εμπλοκής σε ατύχημα είτε ως προς τη σοβαρότητα του. Επιπλέον, διαπιστώνεται ότι η επίδραση αυτή δεν είναι σταθερή σε όλα τα περιβάλλοντα, αλλά διαφοροποιείται ανάλογα με τις τοπικές συνθήκες, τα χαρακτηριστικά του οδικού δικτύου και τα χαρακτηριστικά των οδηγών.

Παράλληλα, προκύπτει ότι οι διαφορές που αποδίδονται στην εθνικότητα δεν μπορούν να ερμηνευτούν ως αποτέλεσμα ενός και μόνο παράγοντα. Αντίθετα, συνδέονται με ένα σύνολο παραμέτρων. Ιδιαίτερα, οι μελέτες που εξετάζουν τουριστικές περιοχές ή χειμερινές συνθήκες υποδεικνύουν ότι η εξοικείωση με το περιβάλλον αποτελεί κρίσιμο μηχανισμό διαφοροποίησης του κινδύνου.

Συνολικά, ο Πίνακας 2.1 λειτουργεί ως συμπυκνωμένη αποτύπωση των βασικών ερευνητικών προσεγγίσεων που έχουν εφαρμοστεί διεθνώς και στην Ελλάδα. Η συγκριτική παρουσίαση των δεδομένων και των μεθόδων αναδεικνύει τόσο τις ομοιότητες όσο και τις διαφορές μεταξύ των μελετών, ενώ παράλληλα διευκολύνει τον εντοπισμό των μεταβλητών και των παραγόντων που εμφανίζονται σημαντικοί. Με βάση τα παραπάνω, στα επόμενα κεφάλαια αξιοποιούνται τα συμπεράσματα της βιβλιογραφίας ώστε να καθοριστεί το πλαίσιο της ανάλυσης και να επιλεγούν οι κατάλληλες μεταβλητές για την παρούσα έρευνα.

### 3. Θεωρητικό Υπόβαθρο

#### 3.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο που αξιοποιήθηκε για την ανάλυση του αντικειμένου της Διπλωματικής Εργασίας. Αρχικά, περιγράφονται συνοπτικά οι βασικές έννοιες της στατιστικής ανάλυσης καθώς και οι κυριότερες στατιστικές κατανομές που χρησιμοποιούνται στην οδική ασφάλεια και σε πλήθος άλλων εφαρμογών.

Στη συνέχεια, παρουσιάζονται αναλυτικά τα στατιστικά πρότυπα που επιλέχθηκαν για τη διερεύνηση της επίδρασης των χαρακτηριστικών του οδηγού και του οδικού περιβάλλοντος στη σοβαρότητα των οδικών ατυχημάτων, με έμφαση στη συνδυαστική τους επιρροή. Ιδιαίτερη αναφορά γίνεται στη θεωρητική τους θεμελίωση, στις βασικές παραδοχές τους και στη διαδικασία εκτίμησης των παραμέτρων.

Τέλος, αναλύονται οι στατιστικοί έλεγχοι που εφαρμόζονται για την αξιολόγηση της καταλληλότητας των υποδειγμάτων, καθώς και τα κριτήρια αποδοχής και σύγκρισης αυτών, ώστε να διασφαλίζεται η αξιοπιστία και η ερμηνευτική ικανότητα των αποτελεσμάτων.

#### 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 τελειώς 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. Συνηθίζεται το πλήθος των στοιχείων ενός δείγματος να συμβολίζεται με το γράμμα  $n$ .

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). Εάν οι μεταβλητές των οποίων οι δυνατές τιμές είναι κατηγορίες διαφορετικές μεταξύ τους. Η χρήση αριθμών για την παράσταση των τιμών μιας 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 μία συνεχή μεταβλητή δύο τιμές 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.

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 όπου τα δεδομένα αποτελούν ένα δείγμα. Η διακύμανση συμβολίζεται με  $s^2$  και

διαίρεται με  $(n-1)$ :

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

όπου  $\bar{x}$  ο δειγματικός μέσος, δηλαδή η μέση τιμή των παρατηρήσεων στο δείγμα.

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 κατανομημένου δείγματος. Σύμφωνα με έναν εμπειρικό κανόνα προκύπτει ότι το διάστημα:

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

Συνδιακύμανση (covariance of two variables): Αποτελεί ένα μέτρο της σχέσης μεταξύ δύο περιοχών δεδομένων.

$$Cov(X, Y) = \left[ \frac{1}{n-1} \right] \sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]$$

Μέτρα αξιοπιστίας:

- Επίπεδο εμπιστοσύνης: η αναλογία των περιπτώσεων που μια εκτίμηση είναι σωστή.
- Επίπεδο σημαντικότητας: η αναλογία των περιπτώσεων που ένα συμπέρασμα είναι εσφαλμένο.

Στη συνέχεια θεωρούνται δύο τυχαίες μεταβλητές  $X, Y$ . Ο βαθμός της γραμμικής συσχέτισης των δύο αυτών μεταβλητών  $X$  και  $Y$  με διασπορά  $\sigma_X^2$  και  $\sigma_Y^2$  αντίστοιχα και συνδιασπορά  $\sigma_{XY} = Cov [X, Y]$  καθορίζεται με συντελεστή συσχέτισης (correlation coefficient)  $\rho$ , που ορίζεται ως εξής:

$$\rho = \left( \frac{\sigma_{XY}}{\sigma_X} \right) \left( \frac{1}{\sigma_Y} \right)$$

Ο συντελεστής συσχέτισης  $\rho$  εκφράζει το 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  $\rho$  γίνεται με την αντικατάσταση στην ανωτέρω εξίσωση της συνδιασποράς  $\sigma_{XY}$  και των διασπορών  $\sigma_X$ ,  $\sigma_Y$  από όπου προκύπτει τελικά η έκφραση της εκτιμήτριας  $r$

$$r(X, Y) = \frac{[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})]}{[(\sum_{i=1}^n (X_i - \bar{X})^2)^{1/2} (\sum_{i=1}^n (Y_i - \bar{Y})^2)^{1/2}]}$$

### **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. Παρακάτω παρατίθενται οι σημαντικότερες στατιστικές κατανομές που χρησιμοποιούνται στην ανάλυση οδικών ατυχημάτων.

#### **3.3.1 Κανονική Κατανομή**

Από τις σημαντικές κατανομές πιθανότητας η οποία αφορά σε συνεχείς μεταβλητές είναι η κανονική κατανομή ή κατανομή του Gauss. Μια συνεχής τυχαία μεταβλητή  $X$  θεωρείται ότι ακολουθεί την κανονική κατανομή με παραμέτρους  $\mu$ ,  $\sigma$  ( $-\infty < \mu < +\infty$ ,  $\sigma > 0$ ), και γράφεται  $X \sim N(\mu, \sigma^2)$ , όταν έχει συνάρτηση πυκνότητας πιθανότητας την:

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

όπου  $\mu$  και  $\sigma$  είναι σταθερές ίσες με την μέση τιμή και την τυπική απόκλιση αντίστοιχα.

#### **3.3.2 Κατανομή Poisson**

Αντίστοιχα, η κατανομή που εφαρμόζεται συνήθως σε διακριτές μεταβλητές, είναι η κατανομή Poisson. Μια τυχαία μεταβλητή  $X$  θεωρείται ότι ακολουθεί κατανομή Poisson με παράμετρο  $\lambda$  ( $\lambda > 0$ ), και γράφεται  $X \sim P(\lambda)$ , όταν έχει συνάρτηση μάζας πυκνότητας πιθανότητας την:

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

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

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  $\lambda$ , έστω των συμβάντων είναι χρονικά σταθερός και (β) οι αριθμοί των συμβάντων σε ξένα διαστήματα αποτελούν ανεξάρτητα ενδεχόμενα (Κοκολάκης και Σπηλιώτης, 1999).

### 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 παρουσιάζουν περιοδικές μεταβολές (όπως για παράδειγμα αριθμός αφίξεων οχημάτων που αφορούν σε μικρά χρονικά διαστήματα (π.χ. 10 sec) 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:

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

όπου  $x = 0, 1, 2, 3, \dots$

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  $\alpha$  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  $\alpha$ , τότε η μηδενική υπόθεση  $H_0$  απορρίπτεται.

## 3.4 Μαθηματικά Πρότυπα

### 3.4.1 Γραμμική Παλινδρόμηση

Ο κλάδος της στατιστικής που εξετάζει τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών με σκοπό την πρόβλεψη της μίας από τις υπόλοιπες ονομάζεται ανάλυση παλινδρόμησης.

Με τον όρο εξαρτημένη μεταβλητή εννοείται η μεταβλητή της οποίας η τιμή πρόκειται να προβλεφθεί, ενώ ανεξάρτητες μεταβλητές είναι εκείνες που χρησιμοποιούνται για την πρόβλεψη της εξαρτημένης μεταβλητής. Σκοπός της παλινδρόμησης είναι η ανάπτυξη ενός μαθηματικού προτύπου που να περιγράφει τη σχέση μεταξύ των μεταβλητών.

Η απλούστερη μορφή παλινδρόμησης είναι η απλή γραμμική παλινδρόμηση (simple linear regression), όπου η εξαρτημένη μεταβλητή  $Y$  εκφράζεται ως γραμμική συνάρτηση μιας ανεξάρτητης μεταβλητής  $X$  σύμφωνα με τη σχέση:

$$y_i = \alpha + \beta * x_i + \varepsilon_i$$

όπου

$\alpha$  είναι ο σταθερός όρος της εξίσωσης,  $\beta$  είναι ο συντελεστής παλινδρόμησης που εκφράζει τη μεταβολή της μεταβλητής  $Y$  όταν η μεταβλητή  $X$  μεταβληθεί κατά μία μονάδα και  $\varepsilon_i$  είναι ο όρος σφάλματος.

Στην περίπτωση όπου η εξαρτημένη μεταβλητή επηρεάζεται από περισσότερες από μία ανεξάρτητες μεταβλητές χρησιμοποιείται η πολλαπλή γραμμική παλινδρόμηση (multiple linear regression). Η γενική μορφή του μοντέλου είναι:

$$y_i = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \beta_3 * x_{3i} + \dots + \beta_k * x_{ki} + \varepsilon_i$$

Η εκτίμηση των παραμέτρων του μοντέλου πραγματοποιείται συνήθως με τη μέθοδο των ελαχίστων, η οποία επιλέγει τις τιμές των παραμέτρων έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των σφαλμάτων.

Η ανάλυση παλινδρόμησης χρησιμοποιείται ευρέως σε μελέτες οδικής ασφάλειας για τη διερεύνηση της σχέσης μεταξύ των χαρακτηριστικών του οδηγού, του οδικού περιβάλλοντος και των οδικών ατυχημάτων.

### 3.4.2 Ανάλυση Διακύμανσης (ANOVA)

Η ανάλυση διακύμανσης (Analysis of Variance – ANOVA) αποτελεί μία στατιστική μέθοδο η οποία χρησιμοποιείται για τη σύγκριση των μέσων τιμών ενός δείγματος μεταξύ περισσότερων από δύο ομάδων. Σκοπός της μεθόδου είναι να διερευνηθεί αν οι διαφορές που παρατηρούνται μεταξύ των ομάδων είναι στατιστικά σημαντικές ή αν μπορούν να αποδοθούν σε τυχαία μεταβλητότητα των δεδομένων.

Η βασική υπόθεση της ανάλυσης διακύμανσης είναι η μηδενική υπόθεση ( $H_0$ ), σύμφωνα με την οποία οι μέσες τιμές όλων των ομάδων είναι ίσες μεταξύ τους. Η

εναλλακτική υπόθεση ( $H_1$ ) υποστηρίζει ότι τουλάχιστον μία από τις ομάδες παρουσιάζει διαφορετική μέση τιμή.

Η μέθοδος ANOVA βασίζεται στη σύγκριση δύο ειδών διακύμανσης:

- της διακύμανσης μεταξύ των ομάδων, η οποία εκφράζει τη διαφοροποίηση των μέσων τιμών των ομάδων και
- της διακύμανσης εντός των ομάδων, η οποία εκφράζει τη φυσική μεταβλητότητα των δεδομένων μέσα σε κάθε ομάδα.

Η σύγκριση των δύο αυτών ποσοτήτων πραγματοποιείται μέσω του στατιστικού δείκτη F, ο οποίος υπολογίζεται ως λόγος της διακύμανσης μεταξύ των ομάδων προς τη διακύμανση εντός των ομάδων.

$$F = \frac{MS_{between}}{MS_{within}}$$

Μεγάλες τιμές του δείκτη F υποδηλώνουν ότι η μεταβλητότητα μεταξύ των ομάδων είναι σημαντικά μεγαλύτερη από τη μεταβλητότητα εντός των ομάδων και συνεπώς υπάρχουν στατιστικά σημαντικές διαφορές μεταξύ των μέσων τιμών.

Σε περιπτώσεις όπου η ανάλυση ANOVA δείξει ότι υπάρχουν στατιστικά σημαντικές διαφορές μεταξύ των ομάδων, εφαρμόζονται συχνά επιπλέον έλεγχοι πολλαπλών συγκρίσεων (post-hoc tests), όπως η μέθοδος Tukey, προκειμένου να εντοπιστούν τα ζεύγη ομάδων που παρουσιάζουν σημαντικές διαφοροποιήσεις.

Η ανάλυση διακύμανσης χρησιμοποιείται συχνά σε μελέτες οδικής ασφάλειας για τη σύγκριση δεικτών ατυχημάτων μεταξύ διαφορετικών ομάδων οδηγών ή διαφορετικών χαρακτηριστικών του οδικού περιβάλλοντος.

### **3.4.3 Πολυωνυμικό Λογιστικό Μοντέλο (Multinomial Logit Model)**

Το πολυωνυμικό λογιστικό μοντέλο αποτελεί επέκταση της λογιστικής παλινδρόμησης και χρησιμοποιείται για την ανάλυση εξαρτημένων μεταβλητών που είναι κατηγορικές και διαθέτουν περισσότερες από δύο κατηγορίες.

Στα μοντέλα αυτού του τύπου η εξαρτημένη μεταβλητή εκφράζει την πιθανότητα εμφάνισης μιας συγκεκριμένης κατηγορίας αποτελέσματος. Μία από τις κατηγορίες επιλέγεται ως κατηγορία αναφοράς και οι υπόλοιπες συγκρίνονται με αυτήν.

Η βασική μορφή του μοντέλου εκφράζεται ως εξής:

$$\ln \left( \frac{P_i}{P_0} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

όπου

$P_i$  είναι η πιθανότητα εμφάνισης της κατηγορίας  $i$ ,

$P_0$  είναι η πιθανότητα της κατηγορίας αναφοράς και

$X_1, X_2, \dots, X_k$  είναι οι ανεξάρτητες μεταβλητές του μοντέλου.

Οι συντελεστές του μοντέλου εκτιμώνται συνήθως με τη μέθοδο της μέγιστης πιθανοφάνειας (maximum likelihood estimation). Για τη διευκόλυνση της ερμηνείας των αποτελεσμάτων χρησιμοποιούνται συχνά οι εκθετικές τιμές των συντελεστών, οι οποίες αντιστοιχούν στους λόγους πιθανοτήτων (odds ratios).

Το πολυωνυμικό λογιστικό μοντέλο χρησιμοποιείται ευρέως σε μελέτες οδικής ασφάλειας για την ανάλυση της σοβαρότητας των οδικών ατυχημάτων, καθώς επιτρέπει τη διερεύνηση της πιθανότητας εμφάνισης διαφορετικών επιπέδων τραυματισμού σε σχέση με τα χαρακτηριστικά του οδηγού, του οχήματος και των συνθηκών του ατυχήματος.

### **3.5 Έλεγχοι Στατιστικής Σημαντικότητας και Αξιολόγηση Υποδειγμάτων**

Η αξιολόγηση των στατιστικών υποδειγμάτων αποτελεί βασικό στάδιο της ανάλυσης, καθώς επιτρέπει την εκτίμηση της καταλληλότητας του μοντέλου και της αξιοπιστίας των αποτελεσμάτων που προκύπτουν από αυτό. Για τον σκοπό αυτό χρησιμοποιούνται διάφορα στατιστικά κριτήρια και δείκτες αξιολόγησης.

- **Μη συσχέτιση ανεξάρτητων μεταβλητών**

Βασική προϋπόθεση για την ανάπτυξη αξιόπιστων στατιστικών μοντέλων είναι η μη ισχυρή συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών. Οι μεταβλητές που εισάγονται στο μοντέλο πρέπει να είναι κατά το δυνατόν γραμμικά ανεξάρτητες μεταξύ τους, καθώς η ύπαρξη υψηλής συσχέτισης μπορεί να οδηγήσει σε προβλήματα και να επηρεάσει την αξιοπιστία των εκτιμήσεων των παραμέτρων.

- **Λογική ερμηνεία των συντελεστών**

Οι συντελεστές των μεταβλητών που προκύπτουν από την εκτίμηση του μοντέλου θα πρέπει να έχουν πρόσημα και τιμές που μπορούν να ερμηνευθούν λογικά με βάση τη θεωρία και τα χαρακτηριστικά του φαινομένου που μελετάται. Θετικός συντελεστής υποδηλώνει ότι η αύξηση της ανεξάρτητης μεταβλητής οδηγεί σε αύξηση της εξαρτημένης μεταβλητής ή της πιθανότητας εμφάνισης ενός γεγονότος, ενώ αρνητικός συντελεστής υποδηλώνει το αντίθετο.

- **Στατιστική σημαντικότητα των παραμέτρων**

Η στατιστική σημαντικότητα των μεταβλητών αξιολογείται μέσω στατιστικών ελέγχων, όπως οι έλεγχοι  $t$  ή  $z$ , καθώς και μέσω της τιμής πιθανότητας ( $p$ -value). Η τιμή  $p$  εκφράζει την πιθανότητα η παρατηρούμενη επίδραση μιας μεταβλητής να οφείλεται σε τυχαία μεταβολή.

Στις περισσότερες εφαρμογές χρησιμοποιείται επίπεδο σημαντικότητας 5% ( $\alpha=0,05$ ). Εάν η τιμή  $p$  είναι μικρότερη από το επίπεδο αυτό, η μεταβλητή θεωρείται στατιστικά σημαντική και συμβάλλει στην ερμηνεία του μοντέλου.

- **Έλεγχος συνολικής σημαντικότητας του μοντέλου**

Για την αξιολόγηση της συνολικής σημαντικότητας ενός μοντέλου χρησιμοποιούνται έλεγχοι όπως η δοκιμή Omnibus ή ο έλεγχος λόγου πιθανοφάνειας (Likelihood Ratio Test). Οι έλεγχοι αυτοί εξετάζουν αν το μοντέλο με τις ανεξάρτητες

μεταβλητές εξηγεί σημαντικά καλύτερα τα δεδομένα σε σχέση με ένα μοντέλο που περιλαμβάνει μόνο τη σταθερά.

- **Δείκτες προσαρμογής των μοντέλων**

Η ποιότητα προσαρμογής ενός στατιστικού μοντέλου αξιολογείται μέσω διαφόρων δεικτών, οι οποίοι επιτρέπουν την εκτίμηση της ικανότητας του μοντέλου να περιγράψει τη μεταβλητότητα των δεδομένων.

Στην περίπτωση της γραμμικής παλινδρόμησης χρησιμοποιείται συχνά ο συντελεστής προσαρμογής  $R^2$ , ο οποίος εκφράζει το ποσοστό της μεταβλητότητας της εξαρτημένης μεταβλητής που εξηγείται από τις ανεξάρτητες μεταβλητές του μοντέλου. Ο δείκτης αυτός λαμβάνει τιμές μεταξύ 0 και 1 και όσο μεγαλύτερη είναι η τιμή του, τόσο καλύτερη θεωρείται η προσαρμογή του μοντέλου.

Στα γενικευμένα γραμμικά μοντέλα, όπως η παλινδρόμηση Poisson και τα λογιστικά μοντέλα, χρησιμοποιούνται συχνά άλλοι δείκτες αξιολόγησης. Ένας από τους σημαντικότερους είναι η λογαριθμική πιθανοφάνεια (log-likelihood), η οποία εκφράζει την πιθανότητα τα παρατηρούμενα δεδομένα να έχουν παραχθεί από το συγκεκριμένο μοντέλο.

Επιπλέον, χρησιμοποιείται το κριτήριο πληροφοριών Akaike (Akaike Information Criterion – AIC), το οποίο λαμβάνει υπόψη τόσο την καλή προσαρμογή του μοντέλου όσο και την πολυπλοκότητά του. Μικρότερες τιμές του δείκτη AIC υποδηλώνουν καλύτερη ισορροπία μεταξύ ακρίβειας και απλότητας του μοντέλου.

Ένας ακόμη σημαντικός δείκτης είναι η απόκλιση (deviance), η οποία χρησιμοποιείται για τη σύγκριση της προσαρμογής μεταξύ διαφορετικών μοντέλων. Η σύγκριση μεταξύ της null deviance και της residual deviance επιτρέπει την εκτίμηση του βαθμού βελτίωσης του μοντέλου μετά την εισαγωγή των ανεξάρτητων μεταβλητών.

- **Σφάλμα του μοντέλου**

Τέλος, είναι σημαντικό να εξετάζεται η συμπεριφορά του σφάλματος του μοντέλου. Στα γραμμικά μοντέλα θεωρείται ότι τα σφάλματα πρέπει να είναι ανεξάρτητα μεταξύ τους, να ακολουθούν κανονική κατανομή και να έχουν σταθερή διασπορά.

Η εξέταση των παραπάνω προϋποθέσεων συμβάλλει στην αξιολόγηση της αξιοπιστίας του μοντέλου και στη διασφάλιση της εγκυρότητας των συμπερασμάτων που προκύπτουν από την ανάλυση.

### **3.6 Μοντέλα Καταμέτρησης Συμβάντων**

Σε πολλές εφαρμογές της στατιστικής η εξαρτημένη μεταβλητή δεν εκφράζει ένα συνεχές μέγεθος αλλά τον αριθμό εμφανίσεων ενός γεγονότος μέσα σε συγκεκριμένο χρονικό ή χωρικό διάστημα. Τέτοιου είδους δεδομένα ονομάζονται δεδομένα καταμέτρησης (count data).

Στην περίπτωση των τροχαίων ατυχημάτων, η μεταβλητή που εξετάζεται συχνά είναι ο αριθμός των ατυχημάτων ή ο αριθμός των παθόντων που καταγράφονται σε μία συγκεκριμένη περιοχή ή σε μία συγκεκριμένη ομάδα πληθυσμού. Για την ανάλυση

αυτού του τύπου δεδομένων χρησιμοποιούνται ειδικά στατιστικά μοντέλα τα οποία βασίζονται κυρίως στην κατανομή Poisson ή στην κατανομή της αρνητικής διωνυμικής.

- **Μοντέλο παλινδρόμησης Poisson**

Το μοντέλο παλινδρόμησης Poisson χρησιμοποιείται για τη διερεύνηση της σχέσης μεταξύ του αριθμού εμφάνισης ενός γεγονότος και ενός συνόλου ανεξάρτητων μεταβλητών. Η βασική παραδοχή του μοντέλου είναι ότι η εξαρτημένη μεταβλητή ακολουθεί την κατανομή Poisson και ότι η αναμενόμενη τιμή της μεταβλητής είναι ίση με τη διακύμανσή της.

Η γενική μορφή του μοντέλου παλινδρόμησης Poisson εκφράζεται ως εξής:

$$\log(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

όπου

$\mu_i$  είναι η αναμενόμενη τιμή του αριθμού των συμβάντων για την παρατήρηση  $i$ ,

$X_1, X_2, \dots, X_k$  είναι οι ανεξάρτητες μεταβλητές του μοντέλου και

$\beta_0, \beta_1, \dots, \beta_k$  είναι οι παράμετροι που εκτιμώνται.

Σε περιπτώσεις όπου οι ομάδες που εξετάζονται έχουν διαφορετικό μέγεθος πληθυσμού, εισάγεται στο μοντέλο μία μεταβλητή έκθεσης (exposure variable), η οποία επιτρέπει την εκτίμηση του ρυθμού εμφάνισης των συμβάντων αντί του απόλυτου αριθμού τους. Η μεταβλητή αυτή εισάγεται στο μοντέλο μέσω ενός όρου offset.

- **Μοντέλο Αρνητικής Διωνυμικής**

Σε πολλές περιπτώσεις πραγματικών δεδομένων, και ιδιαίτερα σε δεδομένα τροχαίων ατυχημάτων, παρατηρείται ότι η διακύμανση της εξαρτημένης μεταβλητής είναι μεγαλύτερη από τη μέση τιμή της. Το φαινόμενο αυτό είναι γνωστό ως υπερδιασπορά.

Όταν η υπόθεση της ισότητας μεταξύ μέσης τιμής και διακύμανσης που χαρακτηρίζει την κατανομή Poisson δεν ικανοποιείται, η χρήση του μοντέλου Poisson μπορεί να οδηγήσει σε αναξιόπιστες εκτιμήσεις των παραμέτρων. Για την αντιμετώπιση του προβλήματος αυτού χρησιμοποιείται το μοντέλο παλινδρόμησης της Αρνητικής Διωνυμικής κατανομής (Negative Binomial Model), το οποίο αποτελεί γενίκευση του μοντέλου Poisson.

Το μοντέλο αυτό εισάγει μία επιπλέον παράμετρο διασποράς, η οποία επιτρέπει στη διακύμανση της εξαρτημένης μεταβλητής να είναι μεγαλύτερη από τη μέση τιμή της. Με τον τρόπο αυτό καθίσταται δυνατή η καλύτερη προσαρμογή του μοντέλου σε δεδομένα που παρουσιάζουν αυξημένη μεταβλητότητα.

Η γενική μορφή του μοντέλου είναι παρόμοια με αυτή της παλινδρόμησης Poisson:

$$\log(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

όπου οι παράμετροι εκτιμώνται επίσης με τη μέθοδο της μέγιστης πιθανοφάνειας.

Η εκτίμηση των παραμέτρων των μοντέλων καταμέτρησης πραγματοποιείται συνήθως με τη μέθοδο της μέγιστης πιθανοφάνειας (Maximum Likelihood Estimation).

Η μέθοδος αυτή επιλέγει τις τιμές των παραμέτρων που μεγιστοποιούν την πιθανότητα τα παρατηρούμενα δεδομένα να έχουν παραχθεί από το συγκεκριμένο μοντέλο.

Για την αξιολόγηση της προσαρμογής των μοντέλων χρησιμοποιούνται διάφοροι στατιστικοί δείκτες, μεταξύ των οποίων σημαντικό ρόλο διαδραματίζουν η λογαριθμική πιθανοφάνεια (log-likelihood) και το κριτήριο πληροφοριών Akaike (Akaike Information Criterion – AIC).

- **Log-Likelihood**

Η λογαριθμική πιθανοφάνεια (log-likelihood) αποτελεί ένα μέτρο που εκφράζει το πόσο πιθανό είναι τα παρατηρούμενα δεδομένα να έχουν προκύψει από το συγκεκριμένο στατιστικό μοντέλο. Μεγαλύτερες τιμές της log-likelihood υποδηλώνουν καλύτερη προσαρμογή του μοντέλου στα δεδομένα.

Η τιμή αυτή χρησιμοποιείται κυρίως για τη σύγκριση διαφορετικών μοντέλων που εφαρμόζονται στο ίδιο σύνολο δεδομένων.

- **Akaike Information Criterion (AIC)**

Ένας από τους πιο διαδεδομένους δείκτες αξιολόγησης μοντέλων είναι το κριτήριο πληροφοριών Akaike (AIC). Ο δείκτης αυτός λαμβάνει υπόψη τόσο την καλή προσαρμογή του μοντέλου όσο και την πολυπλοκότητά του.

Το AIC ορίζεται ως:

$$AIC = -2 \cdot \log(L) + 2k$$

όπου

- *Λείναι η μέγιστη τιμή της συνάρτησης πιθανοφάνειας του μοντέλου*
- *κείναι ο αριθμός των παραμέτρων του μοντέλου.*

Μικρότερες τιμές του AIC υποδηλώνουν καλύτερη ισορροπία μεταξύ ακρίβειας προσαρμογής και πολυπλοκότητας του μοντέλου. Για τον λόγο αυτό το AIC χρησιμοποιείται ευρέως για τη σύγκριση διαφορετικών στατιστικών μοντέλων και την επιλογή του καταλληλότερου.

Τα μοντέλα Poisson και Αρνητικής Διωνυμικής χρησιμοποιούνται ευρέως σε μελέτες οδικής ασφάλειας για την ανάλυση της συχνότητας των τροχαίων ατυχημάτων, καθώς επιτρέπουν τη διερεύνηση της επίδρασης δημογραφικών, χρονικών και περιβαλλοντικών παραγόντων στη συχνότητα εμφάνισης των παθόντων.

### 3.7 Λειτουργία του ειδικού στατιστικού λογισμικού

Στο παρόν υποκεφάλαιο παρουσιάζονται συνοπτικά τα βασικά στοιχεία που σχετίζονται με τη χρήση του στατιστικού λογισμικού που αξιοποιήθηκε για την επεξεργασία και ανάλυση των δεδομένων της παρούσας διπλωματικής εργασίας.

Η στατιστική ανάλυση των δεδομένων πραγματοποιήθηκε με τη χρήση του λογισμικού RStudio, το οποίο αποτελεί ένα από τα πλέον διαδεδομένα εργαλεία για την ανάλυση δεδομένων και τη στατιστική μοντελοποίηση. Το περιβάλλον RStudio βασίζεται στη γλώσσα προγραμματισμού R και προσφέρει ένα ολοκληρωμένο σύστημα για την επεξεργασία δεδομένων, την εφαρμογή στατιστικών μεθόδων και την οπτικοποίηση αποτελεσμάτων.

Η χρήση του RStudio καθιστά δυνατή την αποτελεσματική επεξεργασία μεγάλων συνόλων δεδομένων και την εφαρμογή σύνθετων στατιστικών μοντέλων, συμβάλλοντας έτσι στην αξιόπιστη διερεύνηση των παραγόντων που σχετίζονται με τα οδικά ατυχήματα.

## **4. Συλλογή και επεξεργασία δεδομένων**

### **4.1 Εισαγωγή**

Στο παρόν κεφάλαιο παρουσιάζονται τα δεδομένα που χρησιμοποιήθηκαν για την ανάλυση της παρούσας διπλωματικής εργασίας, καθώς και η διαδικασία επεξεργασίας και προετοιμασίας τους για τη στατιστική ανάλυση. Αρχικά περιγράφεται η βάση δεδομένων από την οποία προέρχονται τα στοιχεία των οδικών ατυχημάτων, ενώ στη συνέχεια παρουσιάζεται η διαδικασία δημιουργίας του τελικού συνόλου δεδομένων που χρησιμοποιήθηκε στην έρευνα.

Η ανάλυση βασίζεται σε δεδομένα οδικών ατυχημάτων τα οποία περιλαμβάνουν πληροφορίες σχετικά με τα χαρακτηριστικά του ατυχήματος, των εμπλεκόμενων οχημάτων και των εμπλεκόμενων ατόμων. Ιδιαίτερη έμφαση δίνεται στα χαρακτηριστικά των οδηγών και ειδικότερα στη μεταβλητή της εθνικότητας, η οποία αποτελεί τον βασικό παράγοντα διερεύνησης της παρούσας μελέτης.

Στο πλαίσιο αυτό πραγματοποιείται αρχικά περιγραφική στατιστική ανάλυση των δεδομένων, η οποία αποσκοπεί στην κατανόηση της κατανομής των βασικών μεταβλητών και στην αρχική διερεύνηση πιθανών διαφοροποιήσεων μεταξύ των κατηγοριών οδηγών. Η περιγραφική ανάλυση περιλαμβάνει την παρουσίαση πινάκων συχνοτήτων και ποσοστιαίων κατανομών για επιλεγμένες μεταβλητές του συνόλου δεδομένων.

Τα αποτελέσματα της περιγραφικής ανάλυσης αποτελούν τη βάση για την εφαρμογή των στατιστικών μοντέλων που παρουσιάζονται στο επόμενο κεφάλαιο, όπου διερευνάται ποσοτικά η επίδραση διαφόρων παραγόντων στη σοβαρότητα των οδικών ατυχημάτων.

### **4.2 Συλλογή στοιχείων από βάση δεδομένων**

Για να επιτευχθεί ο στόχος της Διπλωματικής Εργασίας ήταν αναγκαία η συλλογή στοιχείων που αφορούν τη σοβαρότητα του τραυματισμού του οδηγού στα οδικά ατυχήματα που έγιναν στην Ελλάδα από το 2019-2023. Η βάση δεδομένων που χρησιμοποιήθηκε για την εύρεση του ήταν η Ευρωπαϊκή Βάση Οδικών Ατυχημάτων, CARE (CARE – Community database on Accidents on the Roads in Europe).

Η βάση αυτή αποτελεί μια πρωτοβουλία της Ευρωπαϊκής Επιτροπής, προκειμένου να δημιουργηθεί μία βάση δεδομένων εξατομικευμένων στοιχείων οδικών ατυχημάτων. Περιέχονται αναλυτικά στοιχεία για κάθε ατύχημα, όπως αυτά συγκεντρώνονται από τα κράτη μέλη της Ε.Ε., παράλληλα η δομή του συστήματος CARE εξασφαλίζει δυνατότητες πολλαπλής επεξεργασίας των πληροφοριών που εισάγονται.



Όλες αυτές οι μεταβλητές είναι:

- Η περιοχή όπου έγινε το ατύχημα
- Πόσα οχήματα εμπλέκονται στο ατύχημα
- Ο μήνας που έγινε το ατύχημα
- Η ώρα του ατυχήματος σε μορφή 24 ωρών
- Κωδικός γεωγραφικής ταξινόμησης της ΕΕ (περιφερειακή ενότητα)
- Το όνομα της περιοχής NUTS3 (π.χ. Αττική, Θεσσαλονίκη κτλ.)
- Οι καιρικές συνθήκες την ώρα του ατυχήματος
- Συνθήκες φωτισμού
- Τύπος οχήματος
- Η ηλικία του ατόμου
- Το φύλο του ατόμου
- Η υπηκοότητα του οδηγού/ατόμου
- Η σοβαρότητα του ατυχήματος

### **4.3 Προσδιορισμός πληθυσμού αναφοράς ανά κατηγορία υπηκοότητας**

Για τις ανάγκες της παρούσας ανάλυσης κρίθηκε απαραίτητος ο προσδιορισμός ενός πληθυσμού αναφοράς για κάθε κατηγορία υπηκοότητας, προκειμένου να είναι δυνατή η συγκριτική αξιολόγηση των ατυχημάτων μεταξύ διαφορετικών ομάδων πληθυσμού.

Αρχικά συλλέχθηκαν δεδομένα από την Ελληνική Στατιστική Αρχή (ΕΛΣΤΑΤ) σχετικά με τον μόνιμο πληθυσμό της χώρας. Συγκεκριμένα χρησιμοποιήθηκαν τα αποτελέσματα της Απογραφής Πληθυσμού 2021, τα οποία παρέχουν την πιο πρόσφατη και αξιόπιστη καταγραφή του μόνιμου πληθυσμού ανά υπηκοότητα.

Παράλληλα, προκειμένου να ληφθεί υπόψη και η παρουσία μη μόνιμων κατοίκων που επισκέπτονται τη χώρα, συλλέχθηκαν στοιχεία από την ΕΛΣΤΑΤ σχετικά με τις αφίξεις τουριστών ανά χώρα προέλευσης για τα έτη 2015 και 2016. Τα δύο αυτά έτη επιλέχθηκαν ώστε να εξαχθεί μια αντιπροσωπευτική εικόνα της τουριστικής κινητικότητας πριν από την περίοδο ανάλυσης των ατυχημάτων.

Στη συνέχεια, τόσο ο μόνιμος πληθυσμός όσο και οι αφίξεις τουριστών ομαδοποιήθηκαν σε ευρύτερες γεωγραφικές κατηγορίες υπηκοότητας. Οι κατηγορίες που χρησιμοποιήθηκαν στην παρούσα εργασία είναι οι εξής:

- Έλληνες
- Ασία
- Αμερική
- Αυστραλία
- Αφρική

- Βόρεια Ευρώπη
- Νότια Ευρώπη
- Ανατολική Ευρώπη

Η ομαδοποίηση αυτή πραγματοποιήθηκε ώστε να είναι συμβατή με την κατηγοριοποίηση που χρησιμοποιείται στο σύνολο δεδομένων των οδικών ατυχημάτων.

Τέλος, για τις ανάγκες της ανάλυσης θεωρήθηκε ότι ο πληθυσμός κάθε κατηγορίας παραμένει σταθερός κατά την πενταετή περίοδο μελέτης της διπλωματικής εργασίας. Ο πληθυσμός αναφοράς για κάθε κατηγορία υπολογίστηκε ως το άθροισμα του μόνιμου πληθυσμού (απογραφή 2021) και των αφίξεων τουριστών που αντιστοιχούν στην ίδια γεωγραφική κατηγορία.

Με αυτόν τον τρόπο προέκυψε ένας συνολικός πληθυσμός ανά κατηγορία υπηκοότητας, ο οποίος χρησιμοποιήθηκε ως μέτρο έκθεσης στον κίνδυνο για την ανάλυση των οδικών ατυχημάτων που εξετάζονται στην παρούσα μελέτη.

#### 4.4 Επεξεργασία στοιχείων

Αφού ολοκληρώθηκε η συλλογή των στοιχείων επόμενο βήμα είναι η επεξεργασία των δεδομένων. Για αυτό το λόγο δημιουργήθηκε μια ενιαία βάση δεδομένων η οποία πραγματοποιήθηκε με τη χρήση του προγράμματος Microsoft Excel.

Η βάση αυτή περιέχει για κάθε οδικό ατύχημα τα αντίστοιχα δεδομένα που αναφέραμε παραπάνω. Στον παρακάτω πίνακα φαίνεται η μορφή των στοιχείων όπως ακριβώς τα αντλήσαμε από τη βάση δεδομένων CARE σε πίνακα του Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	C-Year	R-X Area	Number of vehicles involved	A-2 Month	A-2 Day Of Week	A-3 Time hour24	A-4 Nuts Level 3	A-4 Nuts Level 3 Description	A-6 Weather condition	A-7 Light Condition	U-2 Transport Mode (TU type group)	P-2 Person Age Group	P-3 Person Gender	P-4 Person Nationality	Fatally Injured (at 30 days)	
1	2019	Motoway	1	2	6	12	EL632	Achaia Dry / Clear	Daylight	car + taxi	50 - 64	Male	Albania	1		
2	2019	Motoway	1	2	6	12	EL632	Achaia Dry / Clear	Daylight	car + taxi	50 - 64	Male	Albania	1		
3	2019	Motoway	1	3	1	18	EL651	Argolida, Ad Dry / Clear	Daylight	car + taxi	50 - 64	Male	Greece	1		
4	2019	Motoway	1	4	6	12	EL303	Kentrikos Ti Rain	Darkness	motor cycle	25 - 49	Male	Greece	1		
5	2019	Motoway	1	6	4	17	EL031	Aitoloukama Dry / Clear	Daylight	motor cycle	50 - 64	Male	Greece	1		
6	2019	Motoway	1	6	6	15	EL305	Anatoliki All Dry / Clear	Daylight	motor cycle	25 - 49	Male	Greece	1		
7	2019	Motoway	1	6	6	17	EL525	Pieria Dry / Clear	Daylight	motor cycle	65+	Male	Greece	1		
8	2019	Motoway	1	6	7	18	EL641	Voiotia Dry / Clear	Daylight	motor cycle	50 - 64	Male	Greece	1		
9	2019	Motoway	1	7	2	11	EL611	Karditsa, Ti Dry / Clear	Daylight	car + taxi	50 - 64	Male	Albania	1		
10	2019	Motoway	1	7	5	23	EL531	Grevena, Ki Dry / Clear	Darkness	car + taxi	18 - 24	Male	Palestinian Terrtc	1		
11	2019	Motoway	1	8	3	11	EL515	Thessalos, Ki Dry / Clear	Daylight	car + taxi	18 - 24	Male	Foreigner, not spe	1		
12	2019	Motoway	1	8	5	12	EL522	Thessaloniki Dry / Clear	Daylight	car + taxi	25 - 49	Male	Greece	1		
13	2019	Motoway	1	9	2	19	EL611	Karditsa, Ti Dry / Clear	Twilight	car + taxi	65+	Female	Greece	1		
14	2019	Motoway	1	9	6	20	EL305	Anatoliki Att Dry / Clear	Darkness	motor cycle	50 - 64	Male	Greece	1		
15	2019	Motoway	1	9	7	12	EL522	Thessaloniki Dry / Clear	Daylight	car + taxi	65+	Male	Greece	1		
16	2019	Motoway	1	9	7	12	EL532	Achaia Dry / Clear	Daylight	motor cycle	25 - 49	Male	Greece	1		
17	2019	Motoway	1	11	1	16	EL431	Iraklio Dry / Clear	Darkness	car + taxi	25 - 49	Male	Greece	1		
18	2019	Motoway	1	11	5	12	EL301	Vomios Tor Dry / Clear	Darkness	motor cycle	25 - 49	Male	Greece	1		
19	2019	Motoway	1	11	7	10	EL522	Thessaloniki Dry / Clear	Daylight	car + taxi	18 - 24	Male	Pakistan	1		
20	2019	Motoway	1	12	6	17	EL652	Korinthia Dry / Clear	Darkness	car + taxi	65+	Female	Greece	1		
21	2019	Motoway	1	12	7	16	EL301	Vomios Tor Dry / Clear	Darkness	motor cycle	18 - 24	Male	Greece	1		
22	2019	Motoway	2	1	1	20	EL526	Samos Dry / Clear	Darkness	car + taxi	50 - 64	Male	Greece	1		
23	2019	Motoway	2	3	4	15	EL651	Argolida, Ad Dry / Clear	Daylight	car + taxi	65+	Male	Greece	1		
24	2019	Motoway	2	3	7	19	EL305	Anatoliki Att Dry / Clear	Darkness	car + taxi	25 - 49	Male	Greece	1		
25	2019	Motoway	2	4	1	14	EL543	Ipsimitia Dry / Clear	Daylight	heavy goods	50 - 64	Male	Greece	1		
26	2019	Motoway	2	4	5	22	EL531	Grevena, Ki, Severe wind	Darkness	lorry, under	65+	Male	Greece	1		
27	2019	Motoway	2	4	6	21	EL644	Fihotida Dry / Clear	Darkness	car + taxi	50 - 64	Female	Netherlands	1		
28	2019	Motoway	2	5	1	11	EL522	Thessaloniki Dry / Clear	Daylight	car + taxi	50 - 64	Male	Albania	1		
Σ		Fatalities														
		Serious injuries														
		Slight injuries														

Πίνακας 4.1: Μορφή των εξαγομένων στοιχείων από τη βάση CARE σε πίνακα EXCEL

Η αρχική κωδικοποίηση που είχαν οι μεταβλητές στη βάση δεδομένων μας είναι εκείνη που αναφέρθηκε προηγουμένως. Κρίθηκε αναγκαίο πως κάποιες από αυτές τις κατηγορίες πρέπει να ομαδοποιηθούν και κάποιες άλλες να διαγραφούν.

Ορισμένες μεταβλητές της αρχικής βάσης δεδομένων δεν συμπεριλήφθηκαν στην τελική ανάλυση, καθώς δεν κρίθηκαν απαραίτητες για τους στόχους της παρούσας μελέτης ή παρουσίαζαν περιορισμένη χρησιμότητα στη στατιστική μοντελοποίηση. Συγκεκριμένα, αφαιρέθηκαν η μεταβλητή που αφορά τον αριθμό των εμπλεκόμενων οχημάτων στο ατύχημα καθώς και ο κωδικός γεωγραφικής ταξινόμησης της Ευρωπαϊκής Ένωσης.

Στον παρακάτω πίνακα 4.2 φαίνεται η αρχική κωδικοποίηση των εξεταζόμενων χαρακτηριστικών των ατυχημάτων, όπως αυτή ήταν από τη βάση δεδομένων CARE, η τελική κωδικοποίηση τους όπως προέκυψε μετά από αρκετές δοκιμές και η αντίστοιχη έτσι ώστε να είναι συμβατές με το λογισμικό Rstudio.

Όπως προκύπτει από τη τελική κωδικοποίηση ορισμένες μεταβλητές παρέμειναν ίδιες όπως το έτος που σημειώθηκε το ατύχημα, το φύλο του οδηγού, η εθνικότητα. Επιπλέον, οι μήνες του έτους ομαδοποιήθηκαν σε εποχές, προκειμένου να εξεταστεί η εποχική διακύμανση των οδικών ατυχημάτων, λαμβάνοντας υπόψη ότι η Ελλάδα αποτελεί χώρα με έντονη τουριστική δραστηριότητα. Παράλληλα, οι ώρες της ημέρας ομαδοποιήθηκαν σε χρονικές περιόδους, ώστε να διακρίνονται τα διαστήματα αυξημένης και μειωμένης κυκλοφοριακής έντασης.

Παράλληλα, μεγάλη ομαδοποίηση χαρακτηριστικών πραγματοποιήσαμε στους τύπους εμπλεκόμενων οχημάτων και στις καιρικές συνθήκες του ατυχήματος έχοντας και στις δύο περιπτώσεις πολύ μικρή ποσότητα δείγματος για πολλές κατηγορίες.

Τέλος, η μεταβλητή της σοβαρότητας του ατυχήματος ομαδοποιήθηκε σε δύο κατηγορίες: KSI ( Killed or Seriously Injured ) και Slight Injury, προκειμένου να χρησιμοποιηθεί τόσο στην περιγραφική στατιστική ανάλυση όσο και στην εφαρμογή της ανάλυσης διακύμανσης (ANOVA).

	Χαρακτηριστικά Ατυχημάτων		
	Κωδικοποίηση CARE	Τελική κωδικοποίηση	Κωδικοποίηση για Rstudio
YEAR	2019	2019	2019
	2020	2020	2020
	2021	2021	2021
	2022	2022	2022
	2023	2023	2023
AREA	MOTORWAY	ΑΥΤΟΚΙΝΗΤΟΔΡΟΜΟΣ	0
	RURAL	ΑΓΡΟΤΙΚΗ ΠΕΡΙΟΧΗ	1
	URBAN	ΑΣΤΙΚΗ ΠΕΡΙΟΧΗ	2
Month	12	ΧΕΙΜΩΝΑΣ	0
	1		
	2		
	3	ΑΝΟΙΞΗ	1
	4		
	5		
	6	ΚΑΛΟΚΑΙΡΙ	2
	7		
	8		
	9	ΦΘΙΝΟΠΩΡΟ	3
	10		
11			
Day of Week	1	Καθημερινή	0
	2		
	3		
	4		
	5		
	6		
	7	Σαββατοκύριακο	1
Time hour	00:00 - 05:59	Νύχτα	0
	06:00 - 09:59	Πρωινή αιχμή	1
	10:00 - 15:59	Μεσημέρι	2
	16:00 - 19:59	Απογευματινή αιχμή	3
	20:00 - 23:59	Βράδυ	4
Nuts level 3	Περιφέρεια Αττικής		0
	Περιφέρεια Κεντρικής Μακεδονίας		1
	Περιφέρεια Δυτικής Μακεδονίας		2
	Περιφέρεια Α. Μακεδονίας - Θράκης		3
	Περιφέρεια Θεσσαλίας		4
	Περιφέρεια Ηπείρου		5
	Περιφέρεια Ιονίων Νήσων		6
	Περιφέρεια Δυτικής Ελλάδας		7
	Περιφέρεια Στερεάς Ελλάδας		8
	Περιφέρεια Πελοποννήσου		9
	Περιφέρεια Αιγαίου		10
Περιφέρεια Κρήτης		11	
WEATHER CONDITION	DRY/CLEAR	ΚΑΛΟΚΑΙΡΙΑ	0
	FOG, MIST, SMOKE	ΚΑΚΟΚΑΙΡΙΑ	1
	RAIN		
	SEVERE WINDS		
	SLEET, HAIL		
	OTHER		
UNKNOWN			
LIGHT CONDITION	DAYLIGHT	ΜΕΡΑ	0
	TWILIGHT	ΣΟΥΡΟΥΠΟ	1
	DARKNESS	ΝΥΧΤΑ	2
TRANSPORT MODE	CAR + TAXI	ΕΠΙΒΑΤΙΚΟ ΟΧΗΜΑ ΚΑΙ ΤΑΞΙ	0
	PEDAL CYCLE	ΠΟΔΗΛΑΤΑ	1
	MOTOR CYCLE	ΜΟΤΟΣΥΚΛΕΤΕΣ / ΜΟΤΟΠΟΔΗΛΑΤΑ	2
	MOPED	ΆΛΛΟ	3
	AGRICULTURAL TRACTOR		
	BUS OR COACH		
	HEAVY GOODS VEHICLE		
	LORRY, UNDER 3.5 TONES		
OTHER			
UNKNOWN			
PERSON AGE GROUP	<15	<17	0
	15 - 17		
	18 - 24		
	25 - 49		
	50 - 64		
	65+		
UNKNOWN	UNKNOWN	5	
PERSON GENDER	MALE	ΑΝΔΡΑΣ	0
	FEMALE	ΓΥΝΑΙΚΑ	1
PERSON NATIONALITY	GREECE	ΕΛΛΑΔΑ	0
	EASTERN EUROPE	ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	1
	NORTHERN EUROPE	ΒΟΡΕΙΑ ΕΥΡΩΠΗ	2
	SOUTHERN EUROPE	ΝΟΤΙΑ ΕΥΡΩΠΗ	3
	AFRICA	ΑΦΡΙΚΗ	4
	AMERICA	ΑΜΕΡΙΚΗ	5
	ASIA	ΑΣΙΑ	6
	AUSTRALIA	ΑΥΣΤΡΑΛΙΑ	7
SEVERITY	FATALLY INJURED ( AT 30 DAYS)	ΘΑΝΑΤΟΣ	0
	SERIOUSLY INJURED	ΣΟΒΑΡΟΣ ΤΡΑΥΜΑΤΙΣΜΟΣ	1
	SLIGHTLY INJURED	ΕΛΑΦΡΥΣ ΤΡΑΥΜΑΤΙΣΜΟΣ	2

**Πίνακας 4.2:** Κωδικοποίηση μεταβλητών στο CARE και στο Rstudio

Μετά απ' αυτή τη διαδικασία δημιουργήθηκε η νέα βάση δεδομένων όπως φαίνεται και στον παρακάτω πίνακα.

Year	Area	Month	Day Of Week	Time hour 24	Nuts Level 3 Description	Weather condition	Light Condition Groups	Transport Mode (TU type group)	Person Age Group	Person Gender	Person Nationality	Severity
2019	0	0	0	1	3	1	2	0	4	0	0	2
2019	0	0	0	3	3	1	0	3	1	0	0	2
2019	0	0	1	3	0	0	1	2	2	0	0	1
2019	0	0	1	4	1	1	2	3	2	0	0	2
2019	0	0	1	4	1	0	2	0	2	0	0	1
2019	0	0	0	1	0	0	0	0	2	0	0	2
2019	0	0	0	0	0	0	2	0	2	0	0	2
2019	0	0	0	4	8	0	2	0	2	0	0	2
2019	0	0	0	4	0	1	2	0	2	0	0	2
2019	0	0	1	2	7	0	0	0	3	0	1	0
2019	0	0	1	1	1	0	2	0	2	0	0	2
2019	0	1	0	3	9	0	0	0	3	0	0	0
2019	0	1	0	2	2	0	0	0	2	1	0	2
2019	0	1	0	2	9	1	0	0	4	1	0	2
2019	0	1	0	1	0	0	0	0	4	0	0	2
2019	0	1	0	3	1	0	0	0	4	0	0	2
2019	0	1	0	4	1	0	2	0	2	0	0	2
2019	0	1	1	1	0	0	0	0	2	0	0	2
2019	0	1	1	2	7	0	0	0	2	1	0	2
2019	0	1	0	1	1	0	0	0	2	0	0	2
2019	0	1	0	4	5	0	2	0	2	1	0	2
2019	0	1	1	0	0	1	2	2	2	0	0	0

**Πίνακας 4.3:** Τελική μορφή στη βάση δεδομένων μετά την επεξεργασία

## 4.5 Περιγραφικά Στατιστικά Στοιχεία

Στο υποκεφάλαιο αυτό παρουσιάζονται τα βασικά περιγραφικά στατιστικά στοιχεία του συνόλου δεδομένων, με στόχο την κατανόηση της κατανομής των μεταβλητών που χρησιμοποιούνται στην παρούσα διπλωματική εργασία. Η περιγραφική ανάλυση πραγματοποιήθηκε στο τελικό σύνολο δεδομένων που προέκυψε μετά τη διαδικασία καθαρισμού και επεξεργασίας των στοιχείων.

Για τον σκοπό αυτό δημιουργήθηκαν συγκεντρωτικοί πίνακες με τη χρήση του λογισμικού Microsoft Excel, στους οποίους παρουσιάζονται οι κατανομές των μεταβλητών ανά κατηγορία υπηκοότητας και ανά επίπεδο σοβαρότητας τραυματισμού. Συγκεκριμένα εξετάζονται δύο κατηγορίες σοβαρότητας ατυχημάτων:

- KSI (Killed or Seriously Injured), δηλαδή άτομα που σκοτώθηκαν ή τραυματίστηκαν σοβαρά.
- Slight Injury

Οι κατηγορίες υπηκοότητας που χρησιμοποιούνται στην ανάλυση είναι οι εξής:

- Ελλάδα
- Ανατολική Ευρώπη
- Βόρεια Ευρώπη
- Νότια Ευρώπη
- Αφρική
- Αμερική
- Ασία
- Αυστραλία

Στη συνέχεια εξετάζεται η κατανομή των ατυχημάτων σε σχέση με διάφορες μεταβλητές που περιγράφουν τα χαρακτηριστικά του ατυχήματος, του οδηγού και του οχήματος.

- Ανάλυση ανά έτος ατυχήματος

Αρχικά εξετάζεται η κατανομή των ατυχημάτων ανά έτος για την περίοδο 2019–2023, τόσο για τα ατυχήματα κατηγορίας KSI όσο και για τα ατυχήματα με ελαφρύ τραυματισμό. Στους πίνακες παρουσιάζεται ο αριθμός των εμπλεκόμενων οδηγών ανά κατηγορία υπηκοότητας για κάθε έτος, καθώς και τα αντίστοιχα ποσοστά συμμετοχής στο σύνολο των ατυχημάτων.

Year - KSI												
	2019	2020	2021	2022	2023	%2019	%2020	%2021	%2022	%2023	%65+	% UNKNOWN
ΕΛΛΑΔΑ	795	677	759	827	827	20,49	17,45	19,56	21,31	21,19	18,87	1,37
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	70	61	68	63	68	21,21	18,48	20,61	19,09	20,61	3,03	0,61
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	20	19	16	15	19	22,47	21,35	17,98	16,85	21,35	12,36	0
ΝΟΤΙΑ ΕΥΡΩΠΗ	4	2	3	5	2	25	12,5	18,75	31,25	12,5	6,25	0
ΑΦΡΙΚΗ	4	3	2	2	5	25	18,75	12,5	12,5	31,25	6,25	0
ΑΜΕΡΙΚΗ	4	1	1	3	4	30,77	7,69	7,69	23,08	30,77	15,38	15,38
ΑΣΙΑ	16	20	17	17	21	17,58	21,98	18,68	18,68	23,08	4,4	2,2
ΑΥΣΤΡΑΛΙΑ	1	0	0	1	0	50	0	0	50	0	0	0

Year - Slight injury												
	2019	2020	2021	2022	2023	%2019	%2020	%2021	%2022	%2023	%65+	% UNKNOWN
ΕΛΛΑΔΑ	6923	6081	6927	7061	6986	20,41	17,93	20,42	20,64	20,6	9,01	2,12
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	521	419	522	498	512	21,08	16,95	21,12	20,15	20,71	1,94	2,22
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	106	51	105	87	72	25,18	12,11	24,94	20,67	17,1	13,54	3,09
ΝΟΤΙΑ ΕΥΡΩΠΗ	42	16	36	26	25	28,97	11,03	24,83	17,93	17,24	8,28	2,07
ΑΦΡΙΚΗ	32	34	35	41	46	17,02	18,09	18,62	21,81	24,47	2,66	4,79
ΑΜΕΡΙΚΗ	10	3	12	12	17	18,52	5,56	22,22	22,22	31,48	7,41	9,26
ΑΣΙΑ	176	179	253	228	231	16,49	16,78	23,71	21,37	21,65	1,03	2,12
ΑΥΣΤΡΑΛΙΑ	2	1	1	3	3	20	10	10	30	30	0	0

**Πίνακας 4.5:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ατυχήματα ανά εθνικότητα οδηγού και έτος (2019–2023)

Οι πίνακες αυτοί επιτρέπουν την παρατήρηση πιθανών μεταβολών στη συμμετοχή διαφορετικών ομάδων υπηκοότητας στα οδικά ατυχήματα κατά τη διάρκεια της εξεταζόμενης χρονικής περιόδου.

- Ανάλυση ανά ημέρα της εβδομάδας

Στη συνέχεια εξετάζεται η μεταβλητή Day of Week, όπου τα ατυχήματα κατηγοριοποιούνται σε δύο βασικές ομάδες:

- Καθημερινές ημέρες
- Σαββατοκύριακο

Day Of Week - KSI				
	ΚΑΘΗΜΕΡΙΝΗ	ΣΑΒΒΑΤΟΚΥΡΙΑΚΟ	% ΚΑΘΗΜΕΡΙΝΗ	% ΣΑΒΒΑΤΟΚΥΡΙΑΚΟ
ΕΛΛΑΔΑ	2625	1255	67,65	32,35
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	211	119	63,94	36,06
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	73	16	82,02	17,98
ΝΟΤΙΑ ΕΥΡΩΠΗ	13	3	81,25	18,75
ΑΦΡΙΚΗ	12	4	75	25
ΑΜΕΡΙΚΗ	10	3	76,92	23,08
ΑΣΙΑ	57	34	62,64	37,36
ΑΥΣΤΡΑΛΙΑ	2	0	100	0

Day Of Week - Slight injury				
	ΚΑΘΗΜΕΡΙΝΗ	ΣΑΒΒΑΤΟΚΥΡΙΑΚΟ	% ΚΑΘΗΜΕΡΙΝΗ	% ΣΑΒΒΑΤΟΚΥΡΙΑΚΟ
ΕΛΛΑΔΑ	25232	8686	74,39	25,61
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	1730	742	69,98	30,02
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	319	102	75,77	24,23
ΝΟΤΙΑ ΕΥΡΩΠΗ	98	47	67,59	32,41
ΑΦΡΙΚΗ	135	53	71,81	28,19
ΑΜΕΡΙΚΗ	33	21	61,11	38,89
ΑΣΙΑ	766	301	71,79	28,21
ΑΥΣΤΡΑΛΙΑ	8	2	80	20

**Πίνακας 4.6:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ατυχήματα ανά εθνικότητα οδηγού και ημέρα της εβδομάδας

Οι πίνακες παρουσιάζουν τον αριθμό και το ποσοστό των ατυχημάτων για κάθε κατηγορία υπηκοότητας τόσο για τα KSI ατυχήματα όσο και για τα ατυχήματα με ελαφρύ τραυματισμό. Η ανάλυση αυτή επιτρέπει την εξέταση πιθανών διαφορών στη χρονική κατανομή των ατυχημάτων κατά τη διάρκεια της εβδομάδας.

- Ανάλυση ανά καιρικές συνθήκες

Η επόμενη μεταβλητή που εξετάζεται αφορά τις καιρικές συνθήκες κατά τη στιγμή του ατυχήματος. Οι συνθήκες αυτές ομαδοποιήθηκαν σε δύο βασικές κατηγορίες:

- καλές καιρικές συνθήκες
- δυσμενείς καιρικές συνθήκες

Weather condition - KSI				
	ΚΑΛΕΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	ΔΥΣΜΕΝΕΙΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	% ΚΑΛΕΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	% ΔΥΣΜΕΝΕΙΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ
ΕΛΛΑΔΑ	3557	323	91,68	8,32
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	303	27	91,82	8,18
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	80	9	89,89	10,11
ΝΟΤΙΑ ΕΥΡΩΠΗ	16	0	100	0
ΑΦΡΙΚΗ	16	0	100	0
ΑΜΕΡΙΚΗ	12	1	92,31	7,69
ΑΣΙΑ	87	4	95,6	4,4
ΑΥΣΤΡΑΛΙΑ	2	0	100	0

Weather condition - Slight injury				
	ΚΑΛΕΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	ΔΥΣΜΕΝΕΙΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	% ΚΑΛΕΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ	% ΔΥΣΜΕΝΕΙΣ ΚΑΙΡΙΚΕΣ ΣΥΝΘΗΚΕΣ
ΕΛΛΑΔΑ	31545	2373	93	7
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	2289	183	92,6	7,4
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	397	24	94,3	5,7
ΝΟΤΙΑ ΕΥΡΩΠΗ	133	12	91,72	8,28
ΑΦΡΙΚΗ	180	8	95,74	4,26
ΑΜΕΡΙΚΗ	53	1	98,15	1,85
ΑΣΙΑ	991	76	92,88	7,12
ΑΥΣΤΡΑΛΙΑ	10	0	100	0

**Πίνακας 4.7:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ανά εθνικότητα οδηγού και καιρικές συνθήκες

Στους πίνακες παρουσιάζεται η κατανομή των ατυχημάτων ανά κατηγορία υπηκοότητας για τις δύο αυτές κατηγορίες, καθώς και τα αντίστοιχα ποσοστά. Η ανάλυση πραγματοποιείται ξεχωριστά για τα KSI ατυχήματα και για τα ατυχήματα με ελαφρύ τραυματισμό.

- Ανάλυση ανά συνθήκες φωτισμού

Επιπλέον εξετάζονται οι συνθήκες φωτισμού κατά τη στιγμή του ατυχήματος. Οι συνθήκες αυτές κατηγοριοποιούνται σε:

- Ημέρα
- Σούρουπο
- Νύχτα

Light Condition Groups - KSI						
	ΜΕΡΑ	ΣΟΥΡΟΥΠΟ	ΝΥΧΤΑ	% ΜΕΡΑ	% ΣΟΥΡΟΥΠΟ	% ΝΥΧΤΑ
ΕΛΛΑΔΑ	2226	188	1466	57,37	4,85	37,78
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	160	11	159	48,48	3,33	48,18
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	56	8	25	62,92	8,99	28,09
ΝΟΤΙΑ ΕΥΡΩΠΗ	9	1	6	56,25	6,25	37,5
ΑΦΡΙΚΗ	9	0	7	56,25	0	43,75
ΑΜΕΡΙΚΗ	5	0	8	38,46	0	61,54
ΑΣΙΑ	41	3	47	45,05	3,3	51,65
ΑΥΣΤΡΑΛΙΑ	1	0	1	50	0	50

Light Condition Groups - Slight injury						
	ΜΕΡΑ	ΣΟΥΡΟΥΠΟ	ΝΥΧΤΑ	% ΜΕΡΑ	% ΣΟΥΡΟΥΠΟ	% ΝΥΧΤΑ
ΕΛΛΑΔΑ	22435	1731	9752	66,14	5,1	28,75
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	1506	126	840	60,92	5,1	33,98
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	319	23	79	75,77	5,46	18,76
ΝΟΤΙΑ ΕΥΡΩΠΗ	98	5	42	67,59	3,45	28,97
ΑΦΡΙΚΗ	111	6	71	59,04	3,19	37,77
ΑΜΕΡΙΚΗ	38	1	15	70,37	1,85	27,78
ΑΣΙΑ	566	62	439	53,05	5,81	41,14
ΑΥΣΤΡΑΛΙΑ	10	0	0	100	0	0

**Πίνακας 4.8:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ανά εθνικότητα οδηγού και συνθήκες φωτισμού

Στους πίνακες παρουσιάζεται η κατανομή των ατυχημάτων για κάθε κατηγορία υπηκοότητας, τόσο σε απόλυτες τιμές όσο και σε ποσοστά, για τις δύο κατηγορίες σοβαρότητας ατυχημάτων.

- Ανάλυση ανά τύπο οχήματος

Η μεταβλητή Transport Mode αφορά τον τύπο του οχήματος που εμπλέκεται στο ατύχημα. Οι κατηγορίες που χρησιμοποιούνται είναι οι εξής:

- Επιβατικό όχημα και ταξί
- Ποδήλατα
- Μοτοσυκλέτες / Μοτοποδήλατα
- Άλλο όχημα

Transport Mode - KSI								
	ΕΠΙΒΑΤΙΚΟ ΟΧΗΜΑ ΚΑΙ ΤΑΞΙ	ΠΟΔΗΛΑΤΑ	ΜΟΤΟΣΥΚΛΕΤΕΣ / ΜΟΤΟΠΟΔΗΛΑΤΑ	ΆΛΛΟ	% ΕΠΙΒΑΤΙΚΟ ΟΧΗΜΑ ΚΑΙ ΤΑΞΙ	% ΠΟΔΗΛΑΤΑ	% ΜΟΤΟΣΥΚΛΕΤΕΣ / ΜΟΤΟΠΟΔΗΛΑΤΑ	% ΆΛΛΟ
ΕΛΛΑΔΑ	1326	93	2142	339	34,18	2,4	55,21	8,22
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	99	19	187	25	30	5,76	56,67	7,58
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	19	8	56	6	21,35	8,59	62,92	6,74
ΝΟΤΙΑ ΕΥΡΩΠΗ	1	1	13	1	6,25	6,25	81,25	6,25
ΑΦΡΙΚΗ	1	2	13	0	6,25	12,5	81,25	0
ΑΜΕΡΙΚΗ	0	1	11	1	0	7,69	84,62	7,69
ΑΣΙΑ	18	7	62	4	19,78	7,69	68,13	4,4
ΑΥΣΤΡΑΛΙΑ	0	0	2	0	0	0	100	0

Transport Mode - Slight injury								
	ΕΠΙΒΑΤΙΚΟ ΟΧΗΜΑ ΚΑΙ ΤΑΞΙ	ΠΟΔΗΛΑΤΑ	ΜΟΤΟΣΥΚΛΕΤΕΣ / ΜΟΤΟΠΟΔΗΛΑΤΑ	ΆΛΛΟ	% ΕΠΙΒΑΤΙΚΟ ΟΧΗΜΑ ΚΑΙ ΤΑΞΙ	% ΠΟΔΗΛΑΤΑ	% ΜΟΤΟΣΥΚΛΕΤΕΣ / ΜΟΤΟΠΟΔΗΛΑΤΑ	% ΆΛΛΟ
ΕΛΛΑΔΑ	9948	767	21878	1325	29,33	2,26	64,5	3,91
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	660	72	1631	109	26,7	2,91	65,98	4,41
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	95	27	288	16	22,57	5,23	68,41	3,8
ΝΟΤΙΑ ΕΥΡΩΠΗ	25	7	108	5	17,24	4,83	74,48	3,45
ΑΦΡΙΚΗ	23	12	149	4	12,23	6,38	79,26	2,13
ΑΜΕΡΙΚΗ	17	2	34	1	31,48	3,7	62,96	1,85
ΑΣΙΑ	143	86	815	27	13,21	8,06	76,19	2,53
ΑΥΣΤΡΑΛΙΑ	1	0	8	1	10	0	80	10

**Πίνακας 4.9:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ανά εθνικότητα οδηγού και τύπος οχήματος

Η ανάλυση αυτή επιτρέπει την κατανόηση του βαθμού συμμετοχής διαφορετικών τύπων οχημάτων στα οδικά ατυχήματα ανά κατηγορία υπηκοότητας.

- Ανάλυση ανά ηλικιακή ομάδα

Στη συνέχεια εξετάζεται η ηλικιακή ομάδα των εμπλεκόμενων οδηγών. Οι ηλικιακές κατηγορίες που χρησιμοποιούνται είναι οι ακόλουθες:

- <17 ετών
- 18–24 ετών
- 25–49 ετών
- 50–64 ετών
- 65+ ετών
- Άγνωστη ηλικία

Person Age Group - KSI												
	<17	18 - 24	25 - 49	50 - 64	65+	UNKNOWN	%<17	% 18 - 24	% 25 - 49	% 50 - 64	% 65+	% UNKNOWN
ΕΛΛΑΔΑ	124	608	1586	777	732	53	3,2	15,67	40,88	20,03	18,87	1,37
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	11	67	181	59	10	2	3,33	20,3	54,85	17,88	3,03	0,61
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	0	18	40	20	11	0	0	20,22	44,94	22,47	12,36	0
ΝΟΤΙΑ ΕΥΡΩΠΗ	0	4	10	1	1	0	0	25	62,5	6,25	6,25	0
ΑΦΡΙΚΗ	1	7	7	5	1	0	6,25	32,5	48,75	41,25	6,25	0
ΑΜΕΡΙΚΗ	0	2	7	0	2	2	0	15,38	53,85	0	15,38	15,38
ΑΣΙΑ	7	24	48	11	4	7	2,2	26,37	52,75	12,09	4,4	2,2
ΑΥΣΤΡΑΛΙΑ	0	2	0	0	0	0	0	100	0	0	0	0

Person Age Group - Slight Injury												
	<17	18 - 24	25 - 49	50 - 64	65+	UNKNOWN	%<17	% 18 - 24	% 25 - 49	% 50 - 64	% 65+	% UNKNOWN
ΕΛΛΑΔΑ	863	4908	17616	6956	3055	720	1,95	14,47	51,94	20,51	9,01	2,12
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	56	422	1515	376	48	55	2,27	17,07	61,29	15,21	1,94	2,22
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	7	79	171	94	57	13	1,66	18,76	40,62	22,33	13,54	3,09
ΝΟΤΙΑ ΕΥΡΩΠΗ	1	29	77	23	12	3	0,69	20	53,1	15,86	6,28	2,07
ΑΦΡΙΚΗ	1	33	112	28	5	9	0,53	17,55	59,57	14,89	2,66	4,79
ΑΜΕΡΙΚΗ	0	10	24	11	4	5	0	18,52	44,44	20,37	7,41	9,26
ΑΣΙΑ	19	181	734	93	11	29	1,78	16,96	68,79	8,72	1,03	2,72
ΑΥΣΤΡΑΛΙΑ	0	2	6	2	0	0	0	20	60	20	0	0

**Πίνακας 4.10:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ατυχήματα ανά εθνικότητα οδηγού και ηλικίας οδηγού

Οι πίνακες παρουσιάζουν τόσο τον αριθμό όσο και το ποσοστό των ατόμων ανά ηλικιακή κατηγορία και υπηκοότητα για κάθε επίπεδο σοβαρότητας ατυχήματος.

- Ανάλυση ανά φύλο

Τέλος, εξετάζεται η κατανομή των ατυχημάτων ανά φύλο οδηγού, όπου οι κατηγορίες είναι:

- Άνδρας
- Γυναίκα

Person Gender - KSI				
	ΑΝΔΡΑΣ	ΓΥΝΑΙΚΑ	% ΑΝΔΡΑΣ	% ΓΥΝΑΙΚΑ
ΕΛΛΑΔΑ	3569	311	91,98	8,02
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	305	25	92,42	7,58
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	70	19	78,65	21,35
ΝΟΤΙΑ ΕΥΡΩΠΗ	14	2	87,5	12,5
ΑΦΡΙΚΗ	16	0	100	0
ΑΜΕΡΙΚΗ	13	0	100	0
ΑΣΙΑ	89	2	97,8	2,2
ΑΥΣΤΡΑΛΙΑ	2	0	100	0

Person Gender - Slight injury				
	ΑΝΔΡΑΣ	ΓΥΝΑΙΚΑ	% ΑΝΔΡΑΣ	% ΓΥΝΑΙΚΑ
ΕΛΛΑΔΑ	28735	5183	84,72	15,28
ΑΝΑΤΟΛΙΚΗ ΕΥΡΩΠΗ	2157	315	87,26	12,74
ΒΟΡΕΙΑ ΕΥΡΩΠΗ	318	103	75,53	24,47
ΝΟΤΙΑ ΕΥΡΩΠΗ	126	19	86,9	13,1
ΑΦΡΙΚΗ	181	7	96,28	3,72
ΑΜΕΡΙΚΗ	36	18	66,67	33,33
ΑΣΙΑ	1026	41	96,16	3,84
ΑΥΣΤΡΑΛΙΑ	5	5	50	50

**Πίνακας 4.4:** Κατανομή νεκρών ή σοβαρά τραυματισμένων (KSI) και ελαφρά τραυματισμένων σε οδικά ανά εθνικότητα οδηγού και φύλο οδηγού

Οι πίνακες παρουσιάζουν τον αριθμό και το ποσοστό συμμετοχής ανδρών και γυναικών οδηγών στα ατυχήματα για κάθε κατηγορία υπηκοότητας, τόσο για τα KSI ατυχήματα όσο και για τα ατυχήματα με ελαφρύ τραυματισμό.

Οι παραπάνω πίνακες αποτελούν μία πρώτη περιγραφική προσέγγιση των δεδομένων και συμβάλλουν στην καλύτερη κατανόηση των χαρακτηριστικών των οδικών ατυχημάτων. Η ανάλυση αυτή λειτουργεί ως βάση για τα επόμενα στάδια της μελέτης, όπου εφαρμόζονται στατιστικά μοντέλα με σκοπό την διερεύνηση της επίδρασης των διαφόρων μεταβλητών στη σοβαρότητα των οδικών ατυχημάτων.

Από την ανάλυση των περιγραφικών πινάκων που παρουσιάστηκαν προηγουμένως προκύπτουν ορισμένες βασικές παρατηρήσεις σχετικά με τα χαρακτηριστικά των εμπλεκόμενων οδηγών και των συνθηκών των οδικών ατυχημάτων.

Αρχικά, όσον αφορά την ηλικία των εμπλεκόμενων ατόμων, παρατηρείται ότι τα υψηλότερα ποσοστά ατυχημάτων τόσο στην κατηγορία KSI (Killed or Seriously Injured) όσο και στην κατηγορία ελαφρού τραυματισμού (Slight injury) αντιστοιχούν κυρίως στην ηλικιακή ομάδα 25 έως 49 ετών. Η συγκεκριμένη ηλικιακή ομάδα εμφανίζει τη μεγαλύτερη συμμετοχή στα οδικά ατυχήματα, γεγονός που μπορεί να συνδέεται με την αυξημένη καθημερινή μετακίνηση και δραστηριότητα των ατόμων αυτής της ηλικίας.

Παράλληλα, από την ανάλυση της μεταβλητής φύλου προκύπτει ότι οι άνδρες οδηγοί εμπλέκονται σε σημαντικά μεγαλύτερο ποσοστό ατυχημάτων σε σύγκριση με τις γυναίκες οδηγούς. Η τάση αυτή παρατηρείται σε όλες τις κατηγορίες υπηκοότητας που εξετάζονται στην παρούσα μελέτη.

Όσον αφορά τις συνθήκες φωτισμού, τα περισσότερα ατυχήματα καταγράφονται κατά τη διάρκεια της ημέρας. Ωστόσο, αξίζει να σημειωθεί ότι τα ατυχήματα που συμβαίνουν κατά τη διάρκεια της νύχτας αποτελούν επίσης σημαντικό ποσοστό, ιδιαίτερα αν ληφθεί υπόψη ότι η κυκλοφορία κατά τις νυχτερινές ώρες είναι συνήθως μειωμένη.

Τέλος, σε σχέση με τον τύπο οχήματος, προκύπτει ότι το μεγαλύτερο ποσοστό των εμπλεκόμενων οδηγών στα ατυχήματα αφορά οδηγούς επιβατικών αυτοκινήτων

και ταξί, ενώ μικρότερα ποσοστά αντιστοιχούν σε άλλους τύπους οχημάτων, όπως μοτοσυκλέτες, μοτοποδήλατα και ποδήλατα.

Οι παρατηρήσεις αυτές συμβάλλουν στην καλύτερη κατανόηση των βασικών χαρακτηριστικών των οδικών ατυχημάτων και αποτελούν ένα πρώτο βήμα για την περαιτέρω στατιστική διερεύνηση της επίδρασης των μεταβλητών που εξετάζονται στην παρούσα εργασία.

Ωστόσο, η περιγραφική ανάλυση από μόνη της δεν επαρκεί για την εξαγωγή ασφαλών συμπερασμάτων σχετικά με τη στατιστική σημαντικότητα των παρατηρούμενων διαφορών. Για τον λόγο αυτό, στο επόμενο στάδιο της ανάλυσης εφαρμόζεται ανάλυση διακύμανσης (ANOVA), προκειμένου να διερευνηθεί αν οι διαφορές μεταξύ των κατηγοριών μεταβλητών είναι στατιστικά σημαντικές.

## **4.6 Ανάλυση Διακύμανσης (ANOVA)**

Για τη διερεύνηση της σχέσης μεταξύ της εθνικότητας των οδηγών και του δείκτη οδικών ατυχημάτων εφαρμόστηκε ανάλυση διακύμανσης μιας κατεύθυνσης (one-way ANOVA). Η ανάλυση βασίστηκε σε δεδομένα οδικών ατυχημάτων για περίοδο πέντε ετών, ενώ οι δείκτες υπολογίστηκαν ως αριθμός ατυχημάτων ανά 100.000 άτομα πληθυσμού για κάθε κατηγορία εθνικότητας οδηγών. Συνολικά εξετάστηκαν οκτώ κατηγορίες εθνικότητας.

Η μέθοδος ANOVA χρησιμοποιείται για τη σύγκριση των μέσων τιμών ενός δείκτη μεταξύ περισσότερων από δύο ομάδων και βασίζεται στον έλεγχο της μηδενικής υπόθεσης, σύμφωνα με την οποία όλοι οι πληθυσμοί έχουν την ίδια μέση τιμή. Η ανάλυση συγκρίνει τη μεταβλητότητα μεταξύ των ομάδων με τη μεταβλητότητα εντός των ομάδων, μέσω του στατιστικού δείκτη F. Μεγάλες τιμές του δείκτη F υποδηλώνουν ότι οι διαφορές μεταξύ των ομάδων είναι μεγαλύτερες από τη φυσική διασπορά των δεδομένων.

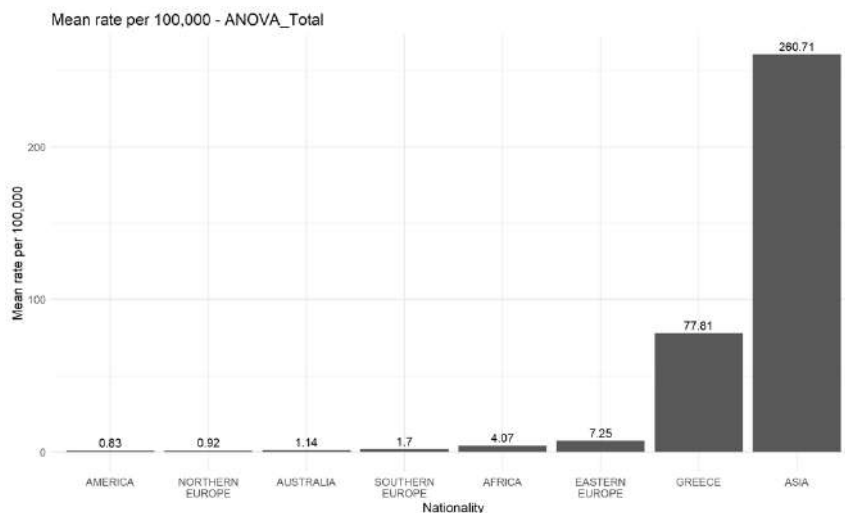
Τα αποτελέσματα της ανάλυσης έδειξαν ότι η εθνικότητα επηρεάζει στατιστικά σημαντικά τον δείκτη οδικών ατυχημάτων. Συγκεκριμένα, για τον συνολικό δείκτη ατυχημάτων προέκυψε τιμή  $F(7,32) = 221$  με επίπεδο σημαντικότητας  $p < 0.001$ , γεγονός που οδηγεί στην απόρριψη της μηδενικής υπόθεσης ισότητας των μέσων τιμών.

Δεδομένου ότι η ANOVA επιτρέπει μόνο τη διαπίστωση της ύπαρξης διαφορών μεταξύ των ομάδων, πραγματοποιήθηκε στη συνέχεια ανάλυση πολλαπλών συγκρίσεων (post-hoc) με τη μέθοδο Tukey, προκειμένου να εντοπιστούν τα ζεύγη ομάδων που παρουσιάζουν στατιστικά σημαντικές διαφορές. Η ανάλυση αυτή συγκρίνει όλα τα πιθανά ζεύγη ομάδων και εξετάζει τη διαφορά των μέσων τιμών καθώς και το αντίστοιχο επίπεδο σημαντικότητας.

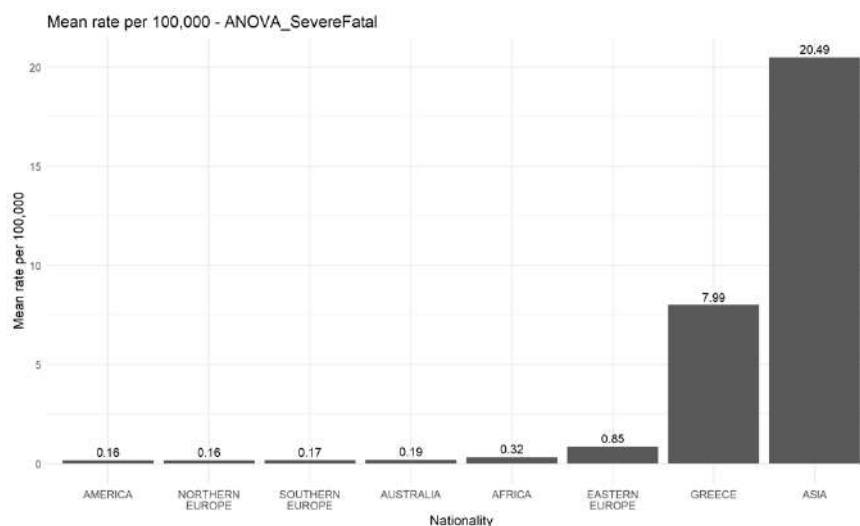
Τα αποτελέσματα των συγκρίσεων έδειξαν ότι ορισμένες κατηγορίες εθνικότητας παρουσιάζουν σημαντικές διαφοροποιήσεις στον δείκτη ατυχημάτων. Ενδεικτικά, σημαντικές διαφορές παρατηρούνται μεταξύ των κατηγοριών ASIA – AFRICA και ASIA – AMERICA, γεγονός που υποδηλώνει διαφορετικά επίπεδα στον εξεταζόμενο δείκτη.

Η ίδια διαδικασία εφαρμόστηκε και για τον δείκτη σοβαρών και θανατηφόρων τραυματισμών (KSI – Killed or Seriously Injured). Και σε αυτή την περίπτωση προέκυψε ισχυρή στατιστική σημαντικότητα, με αποτέλεσμα  $F(7,32) = 327$  και  $p < 0.001$ , γεγονός που δείχνει ότι η διαφοροποίηση μεταξύ των κατηγοριών εθνικότητας αφορά όχι μόνο τη συχνότητα των ατυχημάτων αλλά και τη σοβαρότητα των συνεπειών τους.

Συνολικά, τα αποτελέσματα της ανάλυσης δείχνουν ότι η εθνικότητα των οδηγών συνδέεται με διαφοροποίηση τόσο στη συχνότητα όσο και στη σοβαρότητα των οδικών ατυχημάτων όταν οι τιμές εκφράζονται ως δείκτες ανά πληθυσμό. Η κανονικοποίηση των τιμών ως προς τον πληθυσμό επιτρέπει τη σύγκριση μεταξύ ομάδων διαφορετικού μεγέθους και καθιστά τα αποτελέσματα πιο αντιπροσωπευτικά.



**Εικόνα 4.2 :** Μέσος δείκτης οδικών ατυχημάτων ανά 100.000 κατοίκους ανά κατηγορία εθνικότητας



**Εικόνα 4.3 :** Μέσος δείκτης σοβαρών και θανατηφόρων τραυματισμών (KSI) ανά 100.000 κατοίκους

TOTAL ANOVA					
term	df	sumsq	meansq	statistic	p.value
nationality	7	291980,1724	41711,4532	220,5994194	<0,001
Residuals	32	6050,634704	189,0823345		

TOTAL TUKEY				
comparison	diff	lwr	upr	p adj
AMERICA-AFRICA	-3,246939838	-31,41821012	24,92433045	0.9999
ASIA-AFRICA	256,6379377	228,4666674	284,8092079	0.0000
AUSTRALIA-AFRICA	-2,93112952	-31,10239981	25,24014077	1.0000
EASTERN EUROPE-AFRICA	3,179884809	-24,99138548	31,3511551	0.9999
GREECE-AFRICA	73,73820394	45,56693365	101,9094742	0.0000
NORTHERN EUROPE-AFRICA	-3,148632412	-31,3199027	25,02263787	1.0000
SOUTHERN EUROPE-AFRICA	-2,370965103	-30,54223539	25,80030518	1.0000
ASIA-AMERICA	259,8848775	231,7136072	288,0561478	0.0000
AUSTRALIA-AMERICA	0,315810318	-27,85545997	28,4870806	1.0000
EASTERN EUROPE-AMERICA	6,426824648	-21,74444564	34,59809493	0.9950
GREECE-AMERICA	76,98514378	48,81387349	105,1564141	0.0000
NORTHERN EUROPE-AMERICA	0,098307426	-28,07296286	28,26957771	1.0000
SOUTHERN EUROPE-AMERICA	0,875974735	-27,29529555	29,04724502	1.0000
AUSTRALIA-ASIA	-259,5690672	-287,7403375	-231,3977969	0.0000
EASTERN EUROPE-ASIA	-253,4580529	-281,6293231	-225,2867826	0.0000
GREECE-ASIA	-182,8997337	-211,071004	-154,7284634	0.0000
NORTHERN EUROPE-ASIA	-259,7865701	-287,9578404	-231,6152998	0.0000
SOUTHERN EUROPE-ASIA	-259,0089028	-287,1801731	-230,8376325	0.0000
EASTERN EUROPE-AUSTRALIA	6,111014329	-22,06025596	34,28228462	0.9963
GREECE-AUSTRALIA	76,66933346	48,49806317	104,8406037	0.0000
NORTHERN EUROPE-AUSTRALIA	-0,217502893	-28,38877318	27,95376739	1.0000
SOUTHERN EUROPE-AUSTRALIA	0,560164417	-27,61110587	28,7314347	1.0000
GREECE-EASTERN EUROPE	70,55831913	42,38704884	98,72958942	0.0000
NORTHERN EUROPE-EASTERN EUROPE	-6,328517222	-34,49978751	21,84275306	0.9954
SOUTHERN EUROPE-EASTERN EUROPE	-5,550849912	-33,7221202	22,62042037	0.9980
NORTHERN EUROPE-GREECE	-76,88683635	-105,0581066	-48,71556607	0.0000
SOUTHERN EUROPE-GREECE	-76,10916904	-104,2804393	-47,93789876	0.0000
SOUTHERN EUROPE-NORTHERN EUROPE	0,77766731	-27,39360298	28,9489376	1.0000

Εικόνα 4.4 : Αποτελέσματα ανάλυσης διακύμανσης (ANOVA) και πολλαπλών συγκρίσεων Tukey για τον δείκτη οδικών ατυχημάτων ανά κατηγορία εθνικότητας

SEVFAT ANOVA					
term	df	sumsq	meansq	statistic	p.value
nationality	7	1847,518212	263,9311731	326,9152839	6,60472E-28
Residuals	32	25,83482008	0,807338128		

SEVFAT TUKEY				
comparison	diff	lwr	upr	p adj
AMERICA-AFRICA	-0,159164712	-1,999973522	1,681644098	1.0000
ASIA-AFRICA	20,16819611	18,3273873	22,00900492	0.0000
AUSTRALIA-AFRICA	-0,129134701	-1,969943511	1,711674109	1.0000
EASTERN EUROPE-AFRICA	0,534745183	-1,306063627	2,375553993	0.9793
GREECE-AFRICA	7,667943017	5,827134207	9,508751827	0.0000
NORTHERN EUROPE-AFRICA	-0,158134838	-1,998943648	1,682673972	1.0000
SOUTHERN EUROPE-AFRICA	-0,150303579	-1,991112389	1,690505231	1.0000
ASIA-AMERICA	20,32736083	18,48655202	22,16816964	0.0000
AUSTRALIA-AMERICA	0,030030011	-1,810778799	1,870838821	1.0000
EASTERN EUROPE-AMERICA	0,693909895	-1,146898915	2,534718705	0.9195
GREECE-AMERICA	7,827107729	5,986298919	9,667916539	0.0000
NORTHERN EUROPE-AMERICA	0,001029874	-1,839778936	1,841838684	1.0000
SOUTHERN EUROPE-AMERICA	0,008861133	-1,831947677	1,849669943	1.0000
AUSTRALIA-ASIA	-20,29733082	-22,13813963	-18,45652201	0.0000
EASTERN EUROPE-ASIA	-19,63345093	-21,47425974	-17,79264212	0.0000
GREECE-ASIA	-12,5002531	-14,34106191	-10,65944429	0.0000
NORTHERN EUROPE-ASIA	-20,32633095	-22,16713976	-18,48552214	0.0000
SOUTHERN EUROPE-ASIA	-20,31849969	-22,1593085	-18,47769088	0.0000
EASTERN EUROPE-AUSTRALIA	0,663879884	-1,176928926	2,504688694	0.9351
GREECE-AUSTRALIA	7,797077718	5,956268908	9,637886528	0.0000
NORTHERN EUROPE-AUSTRALIA	-0,029000137	-1,869808947	1,811808673	1.0000
SOUTHERN EUROPE-AUSTRALIA	-0,021168878	-1,861977688	1,819639932	1.0000
GREECE-EASTERN EUROPE	7,133197834	5,292389024	8,974006644	0.0000
NORTHERN EUROPE-EASTERN EUROPE	-0,692880021	-2,533688831	1,147928789	0.9201
SOUTHERN EUROPE-EASTERN EUROPE	-0,685048762	-2,525857572	1,155760048	0.9244
NORTHERN EUROPE-GREECE	-7,826077855	-9,666886665	-5,985269045	0.0000
SOUTHERN EUROPE-GREECE	-7,818246596	-9,659055406	-5,977437786	0.0000
SOUTHERN EUROPE-NORTHERN EUROPE	0,007831259	-1,832977551	1,848640069	1.0000

Εικόνα 4.5 : Αποτελέσματα ανάλυσης διακύμανσης (ANOVA) και πολλαπλών συγκρίσεων Tukey για τον δείκτη σοβαρών και θανατηφόρων τραυματισμών (KSI)

## **5. Ανάπτυξη και Ανάλυση Μαθηματικών Προτύπων**

### **5.1 Εισαγωγή**

Το παρόν κεφάλαιο παρουσιάζει την εφαρμογή της μεθοδολογίας που χρησιμοποιήθηκε στην παρούσα Διπλωματική Εργασία, καθώς και την ανάλυση των αποτελεσμάτων που προέκυψαν από τη στατιστική επεξεργασία των δεδομένων. Μετά την παρουσίαση του θεωρητικού υποβάθρου και την περιγραφή των δεδομένων στα προηγούμενα κεφάλαια, ακολουθεί η ανάπτυξη των στατιστικών μοντέλων που χρησιμοποιήθηκαν για τη διερεύνηση των παραγόντων που σχετίζονται με τα οδικά ατυχήματα.

Στην παρούσα εργασία εφαρμόστηκαν δύο διαφορετικά στατιστικά μοντέλα, τα οποία επιτρέπουν τη διερεύνηση διαφορετικών πτυχών του φαινομένου των οδικών ατυχημάτων. Συγκεκριμένα, για την ανάλυση της συχνότητας εμφάνισης των ατυχημάτων χρησιμοποιήθηκε μοντέλο παλινδρόμησης Poisson, ενώ για την ανάλυση της σοβαρότητας των ατυχημάτων χρησιμοποιήθηκε πολυωνυμικό λογιστικό μοντέλο.

Η χρήση των δύο αυτών μοντέλων επιτρέπει την εξέταση διαφορετικών διαστάσεων του φαινομένου των οδικών ατυχημάτων. Το μοντέλο Poisson χρησιμοποιείται για την ανάλυση δεδομένων καταμέτρησης και επιτρέπει τη διερεύνηση των παραγόντων που επηρεάζουν τη συχνότητα εμφάνισης των παθόντων. Αντίστοιχα, το πολυωνυμικό λογιστικό μοντέλο χρησιμοποιείται για την ανάλυση κατηγορικών μεταβλητών και επιτρέπει τη διερεύνηση των παραγόντων που επηρεάζουν τη σοβαρότητα των συνεπειών των οδικών ατυχημάτων.

Στο κεφάλαιο αυτό παρουσιάζονται τα στατιστικά μοντέλα που αναπτύχθηκαν, η μαθηματική τους διατύπωση καθώς και τα αποτελέσματα που προέκυψαν από την εφαρμογή τους. Αρχικά παρουσιάζεται το μοντέλο καταμέτρησης ατυχημάτων, το οποίο χρησιμοποιείται για την ανάλυση της συχνότητας εμφάνισης των ατυχημάτων. Στη συνέχεια παρουσιάζεται το μοντέλο σοβαρότητας ατυχημάτων, το οποίο χρησιμοποιείται για την ανάλυση της σοβαρότητας των συνεπειών τους.

Η παρουσίαση των αποτελεσμάτων περιλαμβάνει τόσο την περιγραφή των στατιστικών εκτιμήσεων των μοντέλων όσο και την ερμηνεία τους με βάση τη σχετική βιβλιογραφία και τη γενικότερη κατανόηση του φαινομένου των οδικών ατυχημάτων.

### **5.2 Ανάπτυξη Μοντέλων Συχνότητας Ατυχημάτων**

Για την ανάλυση της συχνότητας των οδικών ατυχημάτων χρησιμοποιήθηκε μοντέλο παλινδρόμησης Poisson. Η επιλογή του συγκεκριμένου μοντέλου βασίστηκε στη φύση της εξαρτημένης μεταβλητής, η οποία αφορά τον αριθμό των παθόντων και συνεπώς αποτελεί μεταβλητή καταμέτρησης. Τα μοντέλα παλινδρόμησης Poisson χρησιμοποιούνται ευρέως για την ανάλυση δεδομένων καταμέτρησης, καθώς επιτρέπουν τη διερεύνηση της σχέσης μεταξύ της συχνότητας εμφάνισης ενός γεγονότος και ενός συνόλου ανεξάρτητων μεταβλητών.

Στην παρούσα εργασία το μοντέλο χρησιμοποιείται για την εκτίμηση της συχνότητας εμφάνισης οδικών ατυχημάτων σε σχέση με τα χαρακτηριστικά των

οδηγών και τις συνθήκες του ατυχήματος. Η εξαρτημένη μεταβλητή του μοντέλου είναι ο αριθμός των παθόντων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος.

Στο μοντέλο συμπεριλήφθηκε επίσης μεταβλητή έκθεσης (population exposure), η οποία εκφράζει το μέγεθος του πληθυσμού που αντιστοιχεί σε κάθε κατηγορία οδηγών. Η μεταβλητή αυτή εισήχθη στο μοντέλο μέσω όρου offset, ώστε η εκτίμηση του μοντέλου να αφορά τον ρυθμό εμφάνισης ατυχημάτων και όχι τον απόλυτο αριθμό τους. Με τον τρόπο αυτό λαμβάνεται υπόψη η διαφορετική έκθεση στον κίνδυνο ατυχήματος μεταξύ των κατηγοριών του πληθυσμού.

Ως ανεξάρτητες μεταβλητές στο μοντέλο χρησιμοποιήθηκαν χαρακτηριστικά που σχετίζονται με τον οδηγό, το όχημα και τις συνθήκες του ατυχήματος. Συγκεκριμένα εξετάστηκαν οι ακόλουθες μεταβλητές:

- η εθνικότητα του οδηγού
- το φύλο του οδηγού
- η ηλικιακή ομάδα
- ο μήνας του ατυχήματος
- το έτος του ατυχήματος
- ο τύπος του οχήματος
- οι συνθήκες φωτισμού
- η ημέρα της εβδομάδας
- οι καιρικές συνθήκες
- η περιοχή του ατυχήματος
- η ώρα του ατυχήματος

Οι περισσότερες από τις παραπάνω μεταβλητές είναι κατηγορικές. Για τον λόγο αυτό εισήχθησαν στο μοντέλο με τη μορφή διακριτών μεταβλητών, με μία κατηγορία να χρησιμοποιείται ως κατηγορία αναφοράς, ως προς την οποία συγκρίνονται οι υπόλοιπες κατηγορίες της μεταβλητής.

Το μαθηματικό πρότυπο της παλινδρόμησης Poisson μπορεί να εκφραστεί ως εξής:

$$\log(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \log(\text{Population}_i)$$

όπου  $\mu_i$  είναι η αναμενόμενη τιμή του αριθμού των ατυχημάτων για την παρατήρηση  $i$ ,  $X_1$  έως  $X_k$  είναι οι ανεξάρτητες μεταβλητές του μοντέλου και  $\text{Population}_i$  είναι η μεταβλητή έκθεσης που εισάγεται στο μοντέλο μέσω του όρου offset.

Η εκτίμηση των παραμέτρων του μοντέλου πραγματοποιήθηκε με τη μέθοδο της μέγιστης πιθανοφάνειας (maximum likelihood estimation). Η αξιολόγηση της προσαρμογής του μοντέλου βασίστηκε σε στατιστικούς δείκτες όπως το log-likelihood και το κριτήριο πληροφοριών Akaike (AIC), τα οποία χρησιμοποιούνται ευρέως για τη σύγκριση και αξιολόγηση στατιστικών μοντέλων.

Ωστόσο, ένα βασικό χαρακτηριστικό του μοντέλου Poisson είναι ότι υποθέτει ισότητα μεταξύ της μέσης τιμής και της διακύμανσης της εξαρτημένης μεταβλητής. Στην πράξη, και ιδιαίτερα στα δεδομένα τροχαίων ατυχημάτων, η διακύμανση συχνά υπερβαίνει τη μέση τιμή, φαινόμενο το οποίο είναι γνωστό ως υπερδιασπορά (overdispersion). Στις περιπτώσεις αυτές η χρήση του μοντέλου Poisson μπορεί να οδηγήσει σε υποεκτίμηση των τυπικών σφαλμάτων και σε λιγότερο αξιόπιστα

αποτελέσματα. Για τον λόγο αυτό εκτιμήθηκε και μοντέλο παλινδρόμησης Αρνητικής Διωνυμικής κατανομής, το οποίο αποτελεί γενίκευση του μοντέλου Poisson και επιτρέπει την καλύτερη προσαρμογή σε δεδομένα που παρουσιάζουν υπερδιασπορά.

Για την αντιμετώπιση του φαινομένου της υπερδιασποράς εκτιμήθηκε επίσης μοντέλο παλινδρόμησης Αρνητικής Διωνυμικής κατανομής (Negative Binomial regression). Το μοντέλο αυτό χρησιμοποιείται ευρέως στη διεθνή βιβλιογραφία της οδικής ασφάλειας για την ανάλυση δεδομένων καταμέτρησης, όταν η διακύμανση της εξαρτημένης μεταβλητής είναι μεγαλύτερη από τη μέση τιμή της.

Στην παρούσα εργασία το μοντέλο της Αρνητικής Διωνυμικής χρησιμοποιήθηκε για την εκτίμηση της συχνότητας εμφάνισης οδικών ατυχημάτων σε σχέση με τα χαρακτηριστικά των οδηγών και τις συνθήκες του ατυχήματος, χρησιμοποιώντας την ίδια εξαρτημένη μεταβλητή και το ίδιο σύνολο ανεξάρτητων μεταβλητών που χρησιμοποιήθηκαν και στο μοντέλο Poisson.

Η εξαρτημένη μεταβλητή του μοντέλου είναι ο αριθμός των παθόντων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος. Όπως και στο μοντέλο Poisson, στο μοντέλο της Αρνητικής Διωνυμικής συμπεριλήφθηκε μεταβλητή έκθεσης (population exposure), η οποία εκφράζει το μέγεθος του πληθυσμού που αντιστοιχεί σε κάθε κατηγορία οδηγών. Η μεταβλητή αυτή εισήχθη στο μοντέλο μέσω όρου offset, ώστε η εκτίμηση να αφορά τον ρυθμό εμφάνισης ατυχημάτων και όχι τον απόλυτο αριθμό τους.

Οι ανεξάρτητες μεταβλητές που χρησιμοποιήθηκαν στο μοντέλο είναι οι ίδιες με εκείνες που χρησιμοποιήθηκαν στο μοντέλο Poisson και περιλαμβάνουν χαρακτηριστικά που σχετίζονται με τον οδηγό, το όχημα και τις συνθήκες του ατυχήματος. Οι περισσότερες από τις μεταβλητές αυτές είναι κατηγορικές και για τον λόγο αυτό εισήχθησαν στο μοντέλο ως διακριτές μεταβλητές με μία κατηγορία να χρησιμοποιείται ως κατηγορία αναφοράς.

Το μαθηματικό πρότυπο του μοντέλου της Αρνητικής Διωνυμικής μπορεί να εκφραστεί με παρόμοιο τρόπο με το μοντέλο Poisson:

$$\log(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \log(\text{Population}_i)$$

όπου  $\mu_i$  είναι η αναμενόμενη τιμή του αριθμού των ατυχημάτων για την παρατήρηση  $i$ ,  $X_1$  έως  $X_k$  είναι οι ανεξάρτητες μεταβλητές του μοντέλου και  $\text{Population}_i$  είναι η μεταβλητή έκθεσης που εισάγεται στο μοντέλο μέσω του όρου offset.

Η εκτίμηση των παραμέτρων του μοντέλου πραγματοποιήθηκε επίσης με τη μέθοδο της μέγιστης πιθανοφάνειας (maximum likelihood estimation). Η αξιολόγηση της προσαρμογής του μοντέλου βασίστηκε σε στατιστικούς δείκτες όπως το log-likelihood και το κριτήριο πληροφοριών Akaike (AIC), τα οποία επιτρέπουν τη σύγκριση των διαφορετικών μοντέλων και την επιλογή του καταλληλότερου.

### **5.3 Ανάλυση αποτελεσμάτων του μοντέλου**

Στην παρούσα ενότητα παρουσιάζονται και αξιολογούνται τα αποτελέσματα του μοντέλου παλινδρόμησης Poisson που αναπτύχθηκε για τη διερεύνηση των παραγόντων που σχετίζονται με τη συχνότητα εμφάνισης των οδικών ατυχημάτων. Η εξαρτημένη μεταβλητή του μοντέλου είναι ο αριθμός των ατυχημάτων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος.

Στο μοντέλο συμπεριλήφθηκαν μεταβλητές που σχετίζονται με τα χαρακτηριστικά του οδηγού, τις συνθήκες του ατυχήματος, τον τύπο του οχήματος καθώς και χωρικές και χρονικές μεταβλητές. Συγκεκριμένα εξετάστηκαν η εθνικότητα, το φύλο και η ηλικιακή ομάδα του οδηγού, οι καιρικές συνθήκες και οι συνθήκες φωτισμού, ο τύπος του οχήματος, καθώς και μεταβλητές που σχετίζονται με την περιοχή και τον χρόνο εμφάνισης του ατυχήματος, όπως η περιφερειακή ενότητα, η ημέρα της εβδομάδας, το έτος, ο μήνας και η ώρα του ατυχήματος. Στο μοντέλο εισήχθη επίσης μεταβλητή έκθεσης μέσω του όρου  $\text{offset}(\log(\text{Population}))$ , ώστε η εκτίμηση του μοντέλου να αφορά τον ρυθμό εμφάνισης ατυχημάτων σε σχέση με το μέγεθος του πληθυσμού.

Τα αποτελέσματα του μοντέλου παρουσιάζονται στον Πίνακα 5.1, όπου παρατίθενται οι εκτιμήσεις των συντελεστών, τα τυπικά σφάλματα και οι αντίστοιχοι λόγοι επίπτωσης (Incidence Rate Ratios – IRR). Οι λόγοι επίπτωσης προκύπτουν από τον εκθετικό μετασχηματισμό των συντελεστών του μοντέλου και επιτρέπουν την ερμηνεία της επίδρασης των ανεξάρτητων μεταβλητών στη συχνότητα εμφάνισης των παθόντων.

Variable	Estimate	Std_Error	z_value	p_value	IRR	CI_low	CI_high
(Intercept)	-17,0161	0,0432	-393,6959	0,0000	0,0000	0,0000	0,0000
`Person Nationality`1	-1,6564	0,0137	-121,3241	0,0000	0,1908	0,1858	0,1960
`Person Nationality`2	-1,3557	0,0422	-32,1087	0,0000	0,2578	0,2370	0,2797
`Person Nationality`3	-0,1023	0,0763	-1,3403	0,1801	0,9028	0,7744	1,0446
`Person Nationality`4	1,8426	0,0676	27,2602	0,0000	6,3131	5,5133	7,1868
`Person Nationality`5	0,4791	0,1270	3,7714	0,0002	1,6147	1,2453	2,0508
`Person Nationality`6	-0,3739	0,0193	-19,3648	0,0000	0,6881	0,6624	0,7144
`Person Nationality`7	2,9857	0,2889	10,3346	0,0000	19,8009	10,5986	33,2243
`Person Gender`1	-1,2747	0,0083	-154,0763	0,0000	0,2795	0,2750	0,2841
`Person Age Group`1	0,6005	0,0316	19,0241	0,0000	1,8231	1,7147	1,9406
`Person Age Group`2	1,7933	0,0311	57,6900	0,0000	6,0095	5,6577	6,3910
`Person Age Group`3	0,9418	0,0313	30,0968	0,0000	2,5646	2,4134	2,7284
`Person Age Group`4	0,1870	0,0327	5,7187	0,0000	1,2056	1,1314	1,2862
`Weather condition`1	-1,5966	0,0143	-111,8099	0,0000	0,2026	0,1970	0,2083
`Light Condition Groups`1	-1,3821	0,0137	-100,5427	0,0000	0,2511	0,2444	0,2579
`Light Condition Groups`2	-0,8153	0,0120	-68,0584	0,0000	0,4425	0,4322	0,4530
`Transport Mode (TU type group)`1	-0,7980	0,0280	-28,5112	0,0000	0,4502	0,4260	0,4754
`Transport Mode (TU type group)`2	1,0843	0,0061	178,2685	0,0000	2,9575	2,9225	2,9930
`Transport Mode (TU type group)`3	-0,6373	0,0223	-28,5256	0,0000	0,5287	0,5059	0,5522
Area1	1,1581	0,0280	41,3419	0,0000	3,1839	3,0149	3,3648
Area2	1,8898	0,0258	73,1671	0,0000	6,6180	6,2940	6,9646
`Day Of Week`1	-1,0692	0,0049	-216,1792	0,0000	0,3433	0,3400	0,3466
`Nuts Level 3 Description`1	-1,2182	0,0056	-216,1757	0,0000	0,2958	0,2925	0,2990
`Nuts Level 3 Description`10	-2,0515	0,0162	-126,7123	0,0000	0,1285	0,1245	0,1327
`Nuts Level 3 Description`11	-2,1818	0,0360	-60,5499	0,0000	0,1128	0,1051	0,1210
`Nuts Level 3 Description`2	-1,9516	0,0652	-29,9189	0,0000	0,1420	0,1247	0,1610
`Nuts Level 3 Description`3	-2,1792	0,0228	-95,5808	0,0000	0,1131	0,1082	0,1183
`Nuts Level 3 Description`4	-2,1362	0,0294	-72,7152	0,0000	0,1181	0,1114	0,1250
`Nuts Level 3 Description`5	-2,3000	0,0403	-57,0681	0,0000	0,1003	0,0925	0,1084
`Nuts Level 3 Description`6	-2,3355	0,0324	-72,0025	0,0000	0,0968	0,0907	0,1030
`Nuts Level 3 Description`7	-2,1424	0,0191	-112,1670	0,0000	0,1174	0,1130	0,1218
`Nuts Level 3 Description`8	-2,0100	0,0171	-117,7292	0,0000	0,1340	0,1296	0,1385
`Nuts Level 3 Description`9	-2,0590	0,0216	-95,2754	0,0000	0,1276	0,1223	0,1331
Year2020	0,0304	0,0046	6,6261	0,0000	1,0309	1,0216	1,0402
Year2021	0,1268	0,0044	28,8343	0,0000	1,1352	1,1255	1,1450
Year2022	0,0389	0,0046	8,5150	0,0000	1,0397	1,0304	1,0490
Year2023	-0,0118	0,0047	-2,5371	0,0112	0,9883	0,9793	0,9973
Month	0,0549	0,0013	42,1231	0,0000	1,0564	1,0537	1,0591
`Time hour 24`1	-0,5303	0,0143	-36,9532	0,0000	0,5884	0,5721	0,6052
`Time hour 24`2	0,4895	0,0141	34,7848	0,0000	1,6314	1,5870	1,6770
`Time hour 24`3	-0,1887	0,0137	-13,7557	0,0000	0,8281	0,8061	0,8506
`Time hour 24`4	0,4933	0,0082	59,8568	0,0000	1,6377	1,6115	1,6644

Πίνακας 5.1: Αποτελέσματα μοντέλου Poisson

Παρατηρείται ότι αρκετές από τις μεταβλητές εμφανίζονται στατιστικά σημαντικές, γεγονός που υποδηλώνει ότι οι παράγοντες που σχετίζονται με τα χαρακτηριστικά των οδηγών, τις συνθήκες του ατυχήματος και τις χωρικές και χρονικές διαφοροποιήσεις επηρεάζουν τη συχνότητα εμφάνισης των οδικών ατυχημάτων. Το αποτέλεσμα αυτό είναι αναμενόμενο σε αναλύσεις που βασίζονται σε μεγάλα σύνολα

δεδομένων, καθώς το μεγάλο μέγεθος δείγματος οδηγεί σε μικρότερα τυπικά σφάλματα των εκτιμήσεων και συνεπώς αυξάνει τη στατιστική ισχύ των ελέγχων.

Για την αξιολόγηση της συνολικής προσαρμογής του μοντέλου εξετάστηκαν διάφοροι δείκτες προσαρμογής. Η τιμή του log-likelihood του μοντέλου είναι  $-127481,15$  και εκφράζει την πιθανότητα τα δεδομένα να έχουν παραχθεί από το συγκεκριμένο μοντέλο. Η τιμή αυτή χρησιμοποιείται κυρίως για τη σύγκριση μεταξύ διαφορετικών στατιστικών μοντέλων.

Το κριτήριο πληροφοριών Akaike (AIC) του μοντέλου είναι  $255046,31$ . Ο δείκτης αυτός λαμβάνει υπόψη τόσο την καλή προσαρμογή του μοντέλου όσο και την πολυπλοκότητά του και χρησιμοποιείται για τη σύγκριση εναλλακτικών μοντέλων, όπου μικρότερες τιμές υποδηλώνουν καλύτερη ισορροπία μεταξύ ακρίβειας και πολυπλοκότητας.

Η τιμή της null deviance είναι  $1052899,50$  και αντιστοιχεί σε ένα μοντέλο που περιλαμβάνει μόνο τη σταθερά, χωρίς ανεξάρτητες μεταβλητές. Αντίθετα, η residual deviance του εκτιμημένου μοντέλου είναι  $121565,03$ . Η σημαντική μείωση της απόκλισης μεταξύ των δύο αυτών τιμών δείχνει ότι η εισαγωγή των ανεξάρτητων μεταβλητών βελτιώνει σημαντικά την ικανότητα του μοντέλου να εξηγεί τη μεταβλητότητα των δεδομένων.

Οι βαθμοί ελευθερίας του μοντέλου είναι  $df\_null = 41818$  και  $df\_residual = 41777$  και προκύπτουν από τον αριθμό των παρατηρήσεων και των εκτιμώμενων παραμέτρων του μοντέλου.

LogLik	AIC	Null deviance	Residual deviance	df_null	df_residual	McFadden_R2
-127481,1542	255046,3083	1052899,5038	121565,0308	41818,0000	41777,0000	0,7851

**Πίνακας 5.2:** Στατιστικά προσαρμογής μοντέλου Poisson

Τέλος, για την αξιολόγηση της συνολικής προσαρμογής του μοντέλου υπολογίστηκε ο δείκτης  $R^2$ , ο οποίος αποτελεί έναν από τους συχνότερα χρησιμοποιούμενους δείκτες προσαρμογής στα γενικευμένα γραμμικά μοντέλα. Η τιμή του δείκτη στο παρόν μοντέλο είναι  $0,7851$ , γεγονός που υποδηλώνει πολύ καλή προσαρμογή του μοντέλου στα δεδομένα.

Συνολικά, οι δείκτες αξιολόγησης δείχνουν ότι το μοντέλο παλινδρόμησης Poisson παρουσιάζει ικανοποιητική προσαρμογή και μπορεί να χρησιμοποιηθεί για τη διερεύνηση των παραγόντων που σχετίζονται με τον αριθμό παθόντων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος. Ωστόσο, δεδομένου ότι στα δεδομένα παθόντων ενδέχεται να εμφανίζεται το φαινόμενο της υπερδιασποράς, εκτιμήθηκε επιπλέον και μοντέλο παλινδρόμησης Αρνητικής Διωνυμικής κατανομής προκειμένου να εξεταστεί εάν παρουσιάζει καλύτερη προσαρμογή στα δεδομένα.

Για το λόγο αυτό στην παρούσα ενότητα παρουσιάζονται και αξιολογούνται και τα αποτελέσματα του μοντέλου παλινδρόμησης Αρνητικής Διωνυμικής κατανομής (Negative Binomial), το οποίο εκτιμήθηκε για τη διερεύνηση των παραγόντων που

σχετίζονται με τον αριθμό των παθόντων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος. Η εξαρτημένη μεταβλητή του μοντέλου είναι ο αριθμός των παθόντων που καταγράφονται για κάθε συνδυασμό χαρακτηριστικών των οδηγών και των συνθηκών του ατυχήματος.

Στο μοντέλο συμπεριλήφθηκαν οι ίδιες ανεξάρτητες μεταβλητές που χρησιμοποιήθηκαν και στο μοντέλο Poisson. Συγκεκριμένα εξετάστηκαν η εθνικότητα, το φύλο και η ηλικιακή ομάδα του οδηγού, οι καιρικές συνθήκες και οι συνθήκες φωτισμού, ο τύπος του οχήματος, καθώς και μεταβλητές που σχετίζονται με την περιοχή και τον χρόνο εμφάνισης του ατυχήματος, όπως η περιφερειακή ενότητα, η ημέρα της εβδομάδας, το έτος, ο μήνας και η ώρα του ατυχήματος. Όπως και στο προηγούμενο μοντέλο, εισήχθη μεταβλητή έκθεσης μέσω του όρου  $\text{offset}(\log(\text{Population}))$ , ώστε η εκτίμηση του μοντέλου να αφορά τον ρυθμό εμφάνισης παθόντων σε σχέση με το μέγεθος του πληθυσμού.

Τα αποτελέσματα του μοντέλου παρουσιάζονται στον Πίνακα 5.3, όπου παρατίθενται οι εκτιμήσεις των συντελεστών, τα τυπικά σφάλματα και οι αντίστοιχοι λόγοι επίπτωσης (Incidence Rate Ratios – IRR). Όπως και στο μοντέλο Poisson, οι λόγοι επίπτωσης προκύπτουν από τον εκθετικό μετασχηματισμό των συντελεστών και επιτρέπουν την ερμηνεία της επίδρασης των ανεξάρτητων μεταβλητών στον αριθμό των παθόντων.

Variable	Estimate	Std_Error	z_value	p_value	IRR	CI_low	CI_high
(Intercept)	-15,8983	0,0546	-290,9561	0,0000	0,0000	0,0000	0,0000
`Person Nationality` 1	-1,2754	0,0189	-67,4916	0,0000	0,2793	0,2684	0,2907
`Person Nationality` 2	-0,9043	0,0494	-18,3103	0,0000	0,4048	0,3629	0,4510
`Person Nationality` 3	0,4515	0,0876	5,1555	0,0000	1,5707	1,2901	1,9039
`Person Nationality` 4	2,2082	0,0781	28,2857	0,0000	9,0996	7,7143	10,6979
`Person Nationality` 5	0,9879	0,1444	6,8432	0,0000	2,6855	1,9338	3,6817
`Person Nationality` 6	-0,0109	0,0272	-0,3991	0,6898	0,9892	0,9353	1,0461
`Person Nationality` 7	2,8184	0,3262	8,6394	0,0000	16,7505	8,2175	31,7937
`Person Gender` 1	-0,9281	0,0128	-72,3028	0,0000	0,3953	0,3844	0,4065
`Person Age Group` 1	0,4136	0,0381	10,8669	0,0000	1,5122	1,3921	1,6437
`Person Age Group` 2	1,3525	0,0370	36,5793	0,0000	3,8670	3,5669	4,1950
`Person Age Group` 3	0,6312	0,0376	16,7718	0,0000	1,8799	1,7320	2,0417
`Person Age Group` 4	0,1399	0,0397	3,5271	0,0004	1,1502	1,0551	1,2545
`Weather condition` 1	-1,0558	0,0192	-54,9209	0,0000	0,3479	0,3333	0,3631
`Light Condition Groups` 1	-1,0827	0,0218	-49,5644	0,0000	0,3387	0,3240	0,3540
`Light Condition Groups` 2	-0,6675	0,0204	-32,7815	0,0000	0,5130	0,4908	0,5362
`Transport Mode (TU type group)` 1	-0,6157	0,0341	-18,0643	0,0000	0,5403	0,5007	0,5827
`Transport Mode (TU type group)` 2	0,7758	0,0099	78,0188	0,0000	2,1723	2,1250	2,2206
`Transport Mode (TU type group)` 3	-0,3948	0,0271	-14,5948	0,0000	0,6738	0,6344	0,7155
Area1	0,9263	0,0329	28,1498	0,0000	2,5250	2,3453	2,7196
Area2	1,2891	0,0304	42,4617	0,0000	3,6296	3,3884	3,8896
`Day Of Week` 1	-0,8284	0,0090	-91,5395	0,0000	0,4367	0,4288	0,4449
`Nuts Level 3 Description` 1	-1,0767	0,0104	-103,5195	0,0000	0,3407	0,3337	0,3479
`Nuts Level 3 Description` 10	-1,8412	0,0213	-86,3709	0,0000	0,1586	0,1517	0,1658
`Nuts Level 3 Description` 11	-1,9712	0,0418	-47,1630	0,0000	0,1393	0,1274	0,1521
`Nuts Level 3 Description` 2	-1,7164	0,0742	-23,1199	0,0000	0,1797	0,1528	0,2107
`Nuts Level 3 Description` 3	-1,8811	0,0280	-67,1689	0,0000	0,1524	0,1436	0,1617
`Nuts Level 3 Description` 4	-1,9091	0,0345	-55,3127	0,0000	0,1482	0,1377	0,1594
`Nuts Level 3 Description` 5	-1,9559	0,0465	-42,0369	0,0000	0,1414	0,1278	0,1563
`Nuts Level 3 Description` 6	-2,0564	0,0397	-51,8579	0,0000	0,1279	0,1177	0,1390
`Nuts Level 3 Description` 7	-1,8886	0,0239	-79,0222	0,0000	0,1513	0,1439	0,1590
`Nuts Level 3 Description` 8	-1,7139	0,0219	-78,3334	0,0000	0,1802	0,1720	0,1887
`Nuts Level 3 Description` 9	-1,7721	0,0266	-66,6795	0,0000	0,1700	0,1606	0,1799
Year2020	-0,0073	0,0116	-0,6307	0,5282	0,9927	0,9703	1,0156
Year2021	0,0825	0,0112	7,3907	0,0000	1,0860	1,0624	1,1101
Year2022	0,0207	0,0113	1,8355	0,0664	1,0210	0,9985	1,0439
Year2023	0,0035	0,0113	0,3087	0,7575	1,0035	0,9813	1,0262
Month	0,0522	0,0033	15,7164	0,0000	1,0536	1,0467	1,0606
`Time hour 24` 1	-0,5027	0,0246	-20,4356	0,0000	0,6049	0,5741	0,6373
`Time hour 24` 2	0,3497	0,0245	14,2529	0,0000	1,4187	1,3465	1,4948
`Time hour 24` 3	-0,2044	0,0233	-8,7787	0,0000	0,8151	0,7753	0,8570
`Time hour 24` 4	0,3328	0,0154	21,6561	0,0000	1,3949	1,3523	1,4388

**Πίνακας 5.3: Αποτελέσματα μοντέλου Αρνητικής Διωνυμικής**

Παρατηρείται ότι αρκετές από τις μεταβλητές εμφανίζονται στατιστικά σημαντικές και στο μοντέλο της Αρνητικής Διωνυμικής, γεγονός που επιβεβαιώνει ότι οι παράγοντες που σχετίζονται με τα χαρακτηριστικά των οδηγών, τις συνθήκες του ατυχήματος καθώς και τις χωρικές και χρονικές μεταβλητές επηρεάζουν τον αριθμό των παθόντων.

Για την αξιολόγηση της προσαρμογής του μοντέλου εξετάστηκαν οι ίδιοι δείκτες προσαρμογής με το μοντέλο Poisson. Η τιμή της λογαριθμικής πιθανοφάνειας (log-likelihood) του μοντέλου είναι  $-106111,19$ , ενώ το κριτήριο πληροφοριών Akaike (AIC) είναι  $212308,37$ .

Η τιμή της null deviance είναι  $231943,66$ , ενώ η residual deviance του εκτιμημένου μοντέλου είναι  $40281,73$ . Η σημαντική μείωση της απόκλισης υποδηλώνει ότι η εισαγωγή των ανεξάρτητων μεταβλητών βελτιώνει σημαντικά την ικανότητα του μοντέλου να εξηγεί τη μεταβλητότητα των δεδομένων.

Οι βαθμοί ελευθερίας του μοντέλου είναι  $df_{\text{null}} = 41818$  και  $df_{\text{residual}} = 41777$ .

Τέλος, για την αξιολόγηση της συνολικής προσαρμογής του μοντέλου υπολογίστηκε ο δείκτης McFadden  $R^2$ , ο οποίος χρησιμοποιείται συχνά στα γενικευμένα γραμμικά μοντέλα για την εκτίμηση της προσαρμογής του μοντέλου στα δεδομένα. Η τιμή του δείκτη για το μοντέλο Αρνητικής Διωνυμικής είναι 0,2419, γεγονός που υποδηλώνει ικανοποιητική προσαρμογή του μοντέλου στα δεδομένα.

LogLik	AIC	Null_deviance	residual_deviance	df_null	df_residual	McFadden_R2
-106111,1863	212308,3726	231943,6607	40281,7330	41818,0000	41777,0000	0,2419

Πίνακας 5.4: Στατιστικά προσαρμογής μοντέλου Αρνητικής Διωνυμικής

Για την αξιολόγηση της καταλληλότητας των δύο μοντέλων πραγματοποιήθηκε σύγκριση των βασικών δεικτών προσαρμογής. Το μοντέλο Poisson παρουσίασε τιμή  $AIC = 255046,31$  και  $\log\text{-likelihood} = -127481,15$ , ενώ το μοντέλο Αρνητικής Διωνυμικής παρουσίασε  $AIC = 212308,37$  και  $\log\text{-likelihood} = -106111,19$ .

Η σημαντικά μικρότερη τιμή του AIC στο μοντέλο Αρνητικής Διωνυμικής υποδηλώνει καλύτερη προσαρμογή στα δεδομένα. Το αποτέλεσμα αυτό υποδεικνύει την παρουσία υπερδιασποράς (overdispersion) στα δεδομένα των παθόντων, γεγονός που καθιστά το μοντέλο Αρνητικής Διωνυμικής πιο κατάλληλο για την ανάλυση του αριθμού των παθόντων.

Συνεπώς, το μοντέλο Αρνητικής Διωνυμικής θεωρείται το καταλληλότερο για την ερμηνεία των αποτελεσμάτων της παρούσας ανάλυσης, ενώ το μοντέλο Poisson χρησιμοποιείται ως μοντέλο σύγκρισης για την αξιολόγηση της προσαρμογής.

## **5.4 Ανάπτυξη μοντέλου σοβαρότητας ατυχημάτων**

Στη δεύτερη φάση της ανάλυσης εξετάζεται η σχέση μεταξύ των χαρακτηριστικών των οδηγών, των συνθηκών του ατυχήματος και της σοβαρότητας των οδικών ατυχημάτων. Σε αντίθεση με το προηγούμενο μοντέλο, όπου η εξαρτημένη μεταβλητή ήταν ο αριθμός των παθόντων, στην περίπτωση αυτή η είναι η σοβαρότητα και εκφράζει το επίπεδο της στο ατύχημα.

Για την ανάλυση τέτοιου τύπου δεδομένων χρησιμοποιήθηκε πολυωνυμικό λογιστικό μοντέλο (Multinomial Logit Model – MNL). Το μοντέλο αυτό αποτελεί επέκταση της δυαδικής λογιστικής παλινδρόμησης και επιτρέπει την ανάλυση κατηγορικών μεταβλητών με περισσότερες από δύο κατηγορίες.

Στο πλαίσιο του μοντέλου μία από τις κατηγορίες της εξαρτημένης μεταβλητής επιλέγεται ως κατηγορία αναφοράς. Οι υπόλοιπες κατηγορίες συγκρίνονται με αυτήν, επιτρέποντας την εκτίμηση της πιθανότητας εμφάνισης κάθε επιπέδου σοβαρότητας σε σχέση με την κατηγορία αναφοράς.

Η εκτίμηση του μοντέλου πραγματοποιήθηκε με τη μέθοδο της μέγιστης πιθανοφάνειας (Maximum Likelihood Estimation), η οποία χρησιμοποιείται ευρέως σε μοντέλα διακριτών επιλογών.

Στο μοντέλο συμπεριλήφθηκαν μεταβλητές που σχετίζονται με τα χαρακτηριστικά των οδηγών, τις συνθήκες του ατυχήματος καθώς και χωρικές και χρονικές μεταβλητές. Συγκεκριμένα εξετάστηκαν:

- η εθνικότητα του οδηγού
- η ηλικιακή ομάδα
- το φύλο του οδηγού
- η περιοχή του ατυχήματος
- η ημέρα της εβδομάδας
- οι καιρικές συνθήκες
- οι συνθήκες φωτισμού
- ο τύπος του οχήματος
- η περιφερειακή ενότητα
- το έτος του ατυχήματος
- ο μήνας του ατυχήματος
- η ώρα του ατυχήματος

Οι κατηγορικές μεταβλητές εισήχθησαν στο μοντέλο με τη μορφή εικονικών μεταβλητών, με μία κατηγορία να λειτουργεί ως κατηγορία αναφοράς για κάθε μεταβλητή.

Το πολυωνυμικό λογιστικό μοντέλο μπορεί να εκφραστεί μαθηματικά ως:

$$\log \left( \frac{P_i}{P_0} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

όπου,

$P_i$  είναι η πιθανότητα εμφάνισης μιας συγκεκριμένης κατηγορίας σοβαρότητας και  $P_0$  η πιθανότητα της κατηγορίας αναφοράς.

## **5.5 Ανάλυση αποτελεσμάτων του πολυωνυμικού λογιστικού μοντέλου**

Στην παρούσα ενότητα παρουσιάζονται και αναλύονται τα αποτελέσματα του πολυωνυμικού λογιστικού μοντέλου που εφαρμόστηκε για τη διερεύνηση των παραγόντων που επηρεάζουν τη σοβαρότητα των οδικών ατυχημάτων. Η ανάλυση βασίστηκε σε σύνολο 41.819 παρατηρήσεων.

Η εξαρτημένη μεταβλητή του μοντέλου είναι η σοβαρότητα του ατυχήματος. Στο πλαίσιο του πολυωνυμικού λογιστικού μοντέλου μία από τις τρεις κατηγορίες της σοβαρότητας χρησιμοποιείται ως κατηγορία αναφοράς, ενώ οι υπόλοιπες συγκρίνονται με αυτήν.

Τα αποτελέσματα του μοντέλου παρουσιάζονται στον Πίνακα 5.3 και περιλαμβάνουν τις εκτιμήσεις των συντελεστών, τα τυπικά σφάλματα, τις τιμές του

στατιστικού ελέγχου  $z$  καθώς και τα αντίστοιχα  $p$ -values. Παράλληλα παρουσιάζονται και οι εκθετικές τιμές των συντελεστών, οι οποίες αντιστοιχούν στους λόγους πιθανοτήτων (Odds Ratios – OR) και επιτρέπουν την ευκολότερη ερμηνεία της επίδρασης των ανεξάρτητων μεταβλητών στη σοβαρότητα των ατυχημάτων.

outcome_vs_baseline	term	beta	se	z	p_value	OR	OR_CI_low	OR_CI_high
1	(Intercept)	0,3526	0,3174	1,1111	0,2665	1,4228	0,7638	2,6505
1	person_nationality1	-0,0516	0,1174	-0,4396	0,6602	0,9497	0,7545	1,1954
1	person_nationality2	0,3488	0,2277	1,5321	0,1255	1,4174	0,9072	2,2146
1	person_nationality3	0,5937	0,5460	1,0874	0,2769	1,8107	0,6210	5,2798
1	person_nationality4	0,3295	0,5206	0,6328	0,5268	1,3902	0,5011	3,8567
1	person_nationality5	0,3087	0,6343	0,4868	0,6264	1,3617	0,3928	4,7204
1	person_nationality6	0,1045	0,2188	0,4775	0,6330	1,1102	0,7229	1,7048
1	person_nationality7	-0,5545	1,4201	-0,3905	0,6962	0,5744	0,0355	9,2894
1	person_age_group1	-0,6632	0,2123	-3,1237	0,0018	0,5152	0,3398	0,7811
1	person_age_group2	-0,7739	0,2045	-3,7844	0,0002	0,4612	0,3089	0,6886
1	person_age_group3	-1,0603	0,2108	-5,0291	0,0000	0,3463	0,2291	0,5236
1	person_age_group4	-1,3011	0,2151	-6,0487	0,0000	0,2722	0,1786	0,4150
1	person_gender1	0,4066	0,1155	3,5199	0,0004	1,5018	1,1975	1,8834
1	area1	-0,1537	0,1489	-1,0324	0,3019	0,8575	0,6405	1,1481
1	area2	0,3655	0,1483	2,4650	0,0137	1,4413	1,0778	1,9273
1	day_of_week1	-0,0200	0,0660	-0,3036	0,7615	0,9802	0,8613	1,1155
1	weather_condition1	0,0398	0,1145	0,3475	0,7282	1,0406	0,8315	1,3023
1	light_condition_groups1	-0,0920	0,1614	-0,5703	0,5685	0,9121	0,6648	1,2514
1	light_condition_groups2	-0,1256	0,1482	-0,8474	0,3968	0,8820	0,6596	1,1793
1	transport_mode_tu_type_group1	0,0077	0,1944	0,0398	0,9683	1,0078	0,6885	1,4750
1	transport_mode_tu_type_group2	0,4509	0,0766	5,8865	0,0000	1,5697	1,3509	1,8240
1	transport_mode_tu_type_group3	-0,3473	0,1308	-2,6562	0,0079	0,7066	0,5468	0,9130
1	nuts_level_3_description1	-0,0518	0,1078	-0,4807	0,6307	0,9495	0,7686	1,1730
1	nuts_level_3_description1_0	0,5377	0,1170	4,5965	0,0000	1,7120	1,3612	2,1531
1	nuts_level_3_description1_1	0,2028	0,1395	1,4542	0,1459	1,2249	0,9319	1,6100
1	nuts_level_3_description2	-0,0747	0,2493	-0,2996	0,7645	0,9280	0,5694	1,5127
1	nuts_level_3_description3	0,5492	0,1452	3,7819	0,0002	1,7318	1,3029	2,3021
1	nuts_level_3_description4	0,2841	0,1380	2,0583	0,0396	1,3285	1,0137	1,7412
1	nuts_level_3_description5	0,0269	0,1878	0,1430	0,8863	1,0272	0,7109	1,4842
1	nuts_level_3_description6	0,2118	0,1916	1,1055	0,2689	1,2359	0,8490	1,7991
1	nuts_level_3_description7	0,1477	0,1309	1,1280	0,2593	1,1591	0,8968	1,4981
1	nuts_level_3_description8	0,4249	0,1327	3,2018	0,0014	1,5294	1,1791	1,9837
1	nuts_level_3_description9	0,0133	0,1346	0,0986	0,9214	1,0134	0,7784	1,3192
1	year2020	-0,1720	0,0994	-1,7312	0,0834	0,8419	0,6929	1,0230
1	year2021	-0,0126	0,0965	-0,1304	0,8963	0,9875	0,8173	1,1932
1	year2022	0,1117	0,0949	1,1768	0,2393	1,1182	0,9284	1,3468
1	year2023	0,0572	0,0947	0,6033	0,5463	1,0588	0,8794	1,2749
1	month	-0,0248	0,0294	-0,8432	0,3991	0,9755	0,9210	1,0333
1	time_hour_241	-0,1997	0,1670	-1,1962	0,2316	0,8190	0,5904	1,1360
1	time_hour_242	-0,0003	0,1751	-0,0015	0,9988	0,9997	0,7093	1,4090
1	time_hour_243	-0,0155	0,1606	-0,0968	0,9229	0,9846	0,7187	1,3488
1	time_hour_244	0,1700	0,1044	1,6282	0,1035	1,1853	0,9659	1,4546
2	(Intercept)	2,5903	0,2498	10,3694	0,0000	13,3331	8,1714	21,7553
2	person_nationality1	-0,1980	0,0884	-2,2385	0,0252	0,8204	0,6898	0,9757
2	person_nationality2	0,2545	0,1890	1,3461	0,1783	1,2898	0,8904	1,8681
2	person_nationality3	1,0200	0,4653	2,1921	0,0284	2,7731	1,1140	6,9032
2	person_nationality4	0,1704	0,4252	0,4009	0,6885	1,1858	0,5153	2,7287
2	person_nationality5	0,0096	0,5384	0,0178	0,9858	1,0096	0,3514	2,9006
2	person_nationality6	0,1565	0,1688	0,9272	0,3538	1,1694	0,8400	1,6281
2	person_nationality7	-0,3849	1,0866	-0,3543	0,7232	0,6805	0,0809	5,7248
2	person_age_group1	-0,3561	0,1861	-1,9131	0,0557	0,7004	0,4863	1,0088
2	person_age_group2	-0,2779	0,1801	-1,5432	0,1228	0,7573	0,5321	1,0780
2	person_age_group3	-0,6019	0,1833	-3,2837	0,0010	0,5478	0,3824	0,7845
2	person_age_group4	-1,2528	0,1851	-6,7691	0,0000	0,2857	0,1988	0,4106
2	person_gender1	0,8355	0,0883	9,4606	0,0000	2,3059	1,9394	2,7417
2	area1	-0,1188	0,1017	-1,1684	0,2426	0,8880	0,7275	1,0838
2	area2	1,0293	0,1020	10,0885	0,0000	2,7990	2,2917	3,4187
2	day_of_week1	-0,1777	0,0494	-3,5951	0,0003	0,8372	0,7599	0,9224
2	weather_condition1	0,0901	0,0827	1,0900	0,2757	1,0943	0,9306	1,2867
2	light_condition_groups1	-0,0425	0,1188	-0,3581	0,7203	0,9584	0,7593	1,2096
2	light_condition_groups2	-0,2838	0,1073	-2,6462	0,0081	0,7529	0,6102	0,9290
2	transport_mode_tu_type_group1	-0,3546	0,1441	-2,4610	0,0139	0,7015	0,5289	0,9304
2	transport_mode_tu_type_group2	0,0744	0,0560	1,3291	0,1838	1,0772	0,9653	1,2021
2	transport_mode_tu_type_group3	-0,3071	0,0836	-3,6731	0,0002	0,7356	0,6244	0,8666
2	nuts_level_3_description1	-0,1930	0,0761	-2,5346	0,0113	0,8245	0,7102	0,9572
2	nuts_level_3_description1_0	-0,9423	0,0932	-10,1116	0,0000	0,3897	0,3247	0,4678
2	nuts_level_3_description1_1	-2,0300	0,1133	-17,9115	0,0000	0,1313	0,1052	0,1640
2	nuts_level_3_description2	-1,7571	0,1771	-9,9205	0,0000	0,1726	0,1219	0,2442
2	nuts_level_3_description3	-0,6687	0,1131	-5,9123	0,0000	0,5124	0,4105	0,6396
2	nuts_level_3_description4	-1,4173	0,1062	-13,3506	0,0000	0,2424	0,1968	0,2984
2	nuts_level_3_description5	-1,3101	0,1364	-9,6027	0,0000	0,2698	0,2065	0,3525
2	nuts_level_3_description6	-1,0100	0,1478	-6,8340	0,0000	0,3642	0,2726	0,4866
2	nuts_level_3_description7	-0,8857	0,0959	-9,2401	0,0000	0,4124	0,3418	0,4977
2	nuts_level_3_description8	-0,5120	0,0996	-5,1392	0,0000	0,5993	0,4930	0,7285
2	nuts_level_3_description9	-1,0936	0,0962	-11,3623	0,0000	0,3350	0,2774	0,4046
2	year2020	-0,1380	0,0728	-1,8954	0,0580	0,8711	0,7552	1,0047
2	year2021	0,0098	0,0721	0,1353	0,8923	1,0098	0,8768	1,1630
2	year2022	0,0641	0,0715	0,8958	0,3703	1,0662	0,9267	1,2266
2	year2023	-0,0069	0,0712	-0,0968	0,9229	0,9931	0,8637	1,1419
2	month	0,0042	0,0217	0,1922	0,8476	1,0042	0,9623	1,0479
2	time_hour_241	0,3439	0,1213	2,8346	0,0046	1,4105	1,1120	1,7892
2	time_hour_242	0,8375	0,1284	6,5214	0,0000	2,3106	1,7964	2,9719
2	time_hour_243	0,6566	0,1183	5,5523	0,0000	1,9283	1,5293	2,4313
2	time_hour_244	0,7518	0,0808	9,3004	0,0000	2,1208	1,8100	2,4849

Πίνακας 5.3: Αποτελέσματα μοντέλου Multinomial Logit

Η εξέταση των αποτελεσμάτων δείχνει ότι ορισμένες από τις μεταβλητές που συμπεριλήφθηκαν στο μοντέλο εμφανίζουν στατιστικά σημαντική επίδραση στη σοβαρότητα των οδικών ατυχημάτων, ενώ άλλες δεν παρουσιάζουν σημαντική συσχέτιση.

Όσον αφορά τα χαρακτηριστικά των οδηγών, η ηλικιακή ομάδα αποτελεί έναν από τους σημαντικότερους παράγοντες του μοντέλου. Συγκεκριμένα, οι ηλικιακές ομάδες 18–24 ετών, 25–49 ετών και 50–64 ετών εμφανίζουν στατιστικά σημαντικές διαφοροποιήσεις σε σχέση με την κατηγορία αναφοράς. Το αποτέλεσμα αυτό υποδηλώνει ότι η ηλικία επηρεάζει σημαντικά την πιθανότητα εμφάνισης σοβαρότερων συνεπειών σε ένα οδικό ατύχημα.

Το φύλο του οδηγού εμφανίζεται επίσης στατιστικά σημαντικό στο μοντέλο. Οι άνδρες οδηγοί παρουσιάζουν υψηλότερη πιθανότητα εμπλοκής σε σοβαρότερα ατυχήματα σε σχέση με τις γυναίκες οδηγούς. Το εύρημα αυτό είναι συμβατό με αποτελέσματα προηγούμενων ερευνών, σύμφωνα με τα οποία οι άνδρες οδηγοί εμφανίζουν συχνότερα πιο επικίνδυνη οδηγική συμπεριφορά.

Σε σχέση με την εθνικότητα του οδηγού, τα αποτελέσματα δείχνουν ότι μόνο ορισμένες κατηγορίες εμφανίζουν στατιστικά σημαντική διαφοροποίηση σε σχέση με την κατηγορία αναφοράς που αντιστοιχεί στους οδηγούς ελληνικής εθνικότητας. Για παράδειγμα, οδηγοί από την Ανατολική Ευρώπη εμφανίζουν σημαντικά διαφορετική πιθανότητα εμπλοκής σε σοβαρότερα ατυχήματα σε σχέση με τους Έλληνες οδηγούς. Αντίθετα, για αρκετές άλλες κατηγορίες εθνικότητας, όπως η Βόρεια Ευρώπη, η Νότια Ευρώπη ή η Ασία, δεν παρατηρείται στατιστικά σημαντική διαφοροποίηση.

Σημαντική επίδραση εμφανίζουν επίσης ορισμένες μεταβλητές που σχετίζονται με το περιβάλλον του ατυχήματος. Η περιοχή στην οποία συμβαίνει το ατύχημα παρουσιάζει διαφοροποιήσεις στη σοβαρότητα των οδικών συμβάντων. Συγκεκριμένα, τα ατυχήματα που λαμβάνουν χώρα σε αστικές περιοχές εμφανίζουν διαφορετική πιθανότητα σοβαρότητας σε σχέση με τα ατυχήματα που συμβαίνουν σε αυτοκινητοδρόμους.

Ο τύπος του εμπλεκόμενου οχήματος εμφανίζει επίσης σημαντική επίδραση στη σοβαρότητα των ατυχημάτων. Για παράδειγμα, τα ατυχήματα στα οποία εμπλέκονται μοτοσυκλέτες ή μοτοποδήλατα παρουσιάζουν μεγαλύτερη πιθανότητα σοβαρών συνεπειών σε σχέση με τα ατυχήματα που αφορούν επιβατικά οχήματα.

Η ώρα του ατυχήματος παρουσιάζει επίσης σημαντικές διαφοροποιήσεις σε ορισμένες χρονικές κατηγορίες της ημέρας. Το εύρημα αυτό υποδηλώνει ότι οι κυκλοφοριακές συνθήκες και τα επίπεδα κυκλοφορίας κατά τη διάρκεια της ημέρας μπορούν να επηρεάζουν τη σοβαρότητα των συνεπειών ενός οδικού ατυχήματος.

Αντίθετα, ορισμένες μεταβλητές δεν εμφανίζουν στατιστικά σημαντική επίδραση στο μοντέλο. Μεταβλητές όπως οι καιρικές συνθήκες, οι συνθήκες φωτισμού, η ημέρα της εβδομάδας και σε αρκετές περιπτώσεις ο μήνας του έτους δεν παρουσιάζουν σημαντική διαφοροποίηση στην πιθανότητα εμφάνισης διαφορετικών επιπέδων σοβαρότητας.

Για την αξιολόγηση της προσαρμογής του μοντέλου χρησιμοποιήθηκαν διάφοροι δείκτες καταλληλότητας. Η τιμή της λογαριθμικής πιθανοφάνειας του μοντέλου (Log-Likelihood) είναι -15093,79, ενώ το κριτήριο πληροφοριών Akaike (AIC) λαμβάνει τιμή 30355,59. Οι δείκτες αυτοί χρησιμοποιούνται ευρέως για την αξιολόγηση στατιστικών μοντέλων και επιτρέπουν τη σύγκριση μεταξύ εναλλακτικών προτύπων.

Η υπολειμματική απόκλιση (Residual Deviance) του μοντέλου είναι 30187,59, γεγονός που δείχνει ότι η εισαγωγή των ανεξάρτητων μεταβλητών βελτιώνει σημαντικά την προσαρμογή του μοντέλου σε σχέση με το μηδενικό μοντέλο.

Επιπλέον, εξετάστηκε ο δείκτης  $R^2$ , ο οποίος στην παρούσα ανάλυση λαμβάνει τιμή 0,115. Σε αντίθεση με τον κλασικό συντελεστή  $R^2$  της γραμμικής παλινδρόμησης, οι τιμές του  $R^2$  είναι συνήθως μικρότερες. Στη διεθνή βιβλιογραφία τιμές μεταξύ 0,10 και 0,20 θεωρούνται ενδεικτικές ικανοποιητικής προσαρμογής του μοντέλου, γεγονός που υποδηλώνει ότι το μοντέλο παρουσιάζει αποδεκτή ερμηνευτική ικανότητα.

Συνολικά, τα αποτελέσματα του πολυωνυμικού λογιστικού μοντέλου δείχνουν ότι η σοβαρότητα των οδικών ατυχημάτων επηρεάζεται από έναν συνδυασμό δημογραφικών χαρακτηριστικών των οδηγών, χαρακτηριστικών των οχημάτων και περιβαλλοντικών παραγόντων. Κάτι που καθιστά απαραίτητη την εκτίμηση πολλών διαφορετικών μεταβλητών προκειμένου να κατανοηθεί καλύτερα η δυναμική των οδικών ατυχημάτων.

Metric	Value
N observations	41819
LogLikelihood	-15093,7939
AIC	30355,58783
Residual Deviance	30187,58783
McFadden_R2	0,114962926

Πίνακας 5.2: Στατιστικά προσαρμογής Multinomial Logit

## 5.6 Σύγκριση των αποτελεσμάτων των στατιστικών μοντέλων

Στο πλαίσιο της παρούσας εργασίας εφαρμόστηκαν δύο διαφορετικές στατιστικές προσεγγίσεις με σκοπό τη διερεύνηση των παραγόντων που σχετίζονται με την οδική ασφάλεια. Το πρώτο μοντέλο αφορά μοντέλο καταμέτρησης παθόντων (Poisson και Αρνητικής Διωνυμικής), το οποίο χρησιμοποιήθηκε για την ανάλυση του αριθμού των παθόντων που παρατηρούνται για κάθε συνδυασμό χαρακτηριστικών οδηγών και συνθηκών ατυχήματος. Το δεύτερο μοντέλο βασίζεται στην πολυωνυμική λογιστική παλινδρόμηση και χρησιμοποιήθηκε για τη διερεύνηση των παραγόντων που σχετίζονται με τη σοβαρότητα των συνεπειών των ατυχημάτων.

Η χρήση των δύο αυτών μοντέλων επιτρέπει την εξέταση δύο διαφορετικών διαστάσεων του φαινομένου της οδικής ασφάλειας. Από τη μία πλευρά εξετάζεται ο αριθμός των παθόντων που συνδέεται με συγκεκριμένα χαρακτηριστικά των οδηγών και των συνθηκών του ατυχήματος, ενώ από την άλλη διερευνάται η πιθανότητα εμφάνισης σοβαρότερων συνεπειών για τους εμπλεκόμενους χρήστες του οδικού δικτύου.

Από τη σύγκριση των αποτελεσμάτων των μοντέλων προκύπτει ότι ορισμένοι παράγοντες εμφανίζουν σημαντική επίδραση τόσο στον αριθμό των παθόντων όσο και στη σοβαρότητα των συνεπειών. Μεταβλητές όπως η ηλικιακή ομάδα και το φύλο του οδηγού εμφανίζονται σημαντικές και στα δύο μοντέλα, γεγονός που υποδηλώνει ότι τα

δημογραφικά χαρακτηριστικά των οδηγών σχετίζονται τόσο με τον αριθμό των παθόντων όσο και με τη σοβαρότητα των τραυματισμών.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η επίδραση του τύπου του εμπλεκόμενου οχήματος. Τα αποτελέσματα δείχνουν ότι τα δίκυκλα οχήματα συνδέονται με υψηλότερη πιθανότητα σοβαρότερων συνεπειών σε σχέση με τα επιβατικά οχήματα, γεγονός που μπορεί να αποδοθεί στο χαμηλότερο επίπεδο προστασίας που παρέχουν στους οδηγούς και τους επιβάτες τους.

Αντίθετα, ορισμένες μεταβλητές φαίνεται να επηρεάζουν κυρίως τον αριθμό των παθόντων και όχι απαραίτητα τη σοβαρότητα των τραυματισμών. Για παράδειγμα, οι καιρικές συνθήκες και οι συνθήκες φωτισμού εμφανίζουν σημαντική επίδραση στο μοντέλο καταμέτρησης παθόντων, γεγονός που υποδηλώνει ότι οι περιβαλλοντικές συνθήκες επηρεάζουν την πιθανότητα εμπλοκής σε περιστατικά με παθόντες, χωρίς όμως να επηρεάζουν με τον ίδιο τρόπο τη σοβαρότητα των συνεπειών.

Ιδιαίτερη σημασία για την παρούσα εργασία έχει η μεταβλητή της εθνικότητας του οδηγού, καθώς αποτελεί το βασικό αντικείμενο διερεύνησης της μελέτης. Ως κατηγορία αναφοράς στο μοντέλο χρησιμοποιήθηκαν οι οδηγοί ελληνικής εθνικότητας, γεγονός που σημαίνει ότι οι εκτιμήσεις των υπόλοιπων κατηγοριών συγκρίνονται με τους Έλληνες οδηγούς.

Τα αποτελέσματα του μοντέλου καταμέτρησης παθόντων δείχνουν ότι υπάρχουν σημαντικές διαφοροποιήσεις μεταξύ των κατηγοριών εθνικότητας ως προς τον αριθμό των παθόντων που καταγράφονται για κάθε συνδυασμό χαρακτηριστικών οδηγών και συνθηκών ατυχήματος.

Συγκεκριμένα, οι οδηγοί από Νότια Ευρώπη ( $IRR = 1,57$ ) εμφανίζουν περίπου 1,6 φορές μεγαλύτερο αριθμό παθόντων σε σχέση με τους Έλληνες οδηγούς, για παρόμοιους συνδυασμούς χαρακτηριστικών και συνθηκών ατυχήματος. Ακόμη μεγαλύτερη διαφοροποίηση παρατηρείται για οδηγούς από Αφρική ( $IRR = 9,10$ ), για τους οποίους ο αναμενόμενος αριθμός παθόντων είναι σημαντικά υψηλότερος σε σχέση με την κατηγορία αναφοράς.

Παρόμοια εικόνα εμφανίζουν και οι οδηγοί από Αμερική ( $IRR = 2,69$ ) καθώς και από Αυστραλία ( $IRR = 16,75$ ), οι οποίοι παρουσιάζουν σημαντικά υψηλότερο αριθμό παθόντων σε σχέση με τους Έλληνες οδηγούς. Αντίθετα, οι οδηγοί από Ανατολική Ευρώπη ( $IRR = 0,28$ ) και Βόρεια Ευρώπη ( $IRR = 0,40$ ) εμφανίζουν μικρότερο αριθμό παθόντων σε σχέση με την κατηγορία αναφοράς.

Για τους οδηγούς από Ασία ( $IRR = 0,99$ ) η μεταβλητή δεν εμφανίζεται στατιστικά σημαντική, γεγονός που υποδηλώνει ότι δεν παρατηρείται ουσιαστική διαφοροποίηση στον αριθμό των παθόντων σε σχέση με τους Έλληνες οδηγούς.

Στην περίπτωση της ανάλυσης της σοβαρότητας των τραυματισμών χρησιμοποιήθηκαν οι λόγοι πιθανοτήτων ( $Odds Ratios - OR$ ), οι οποίοι προκύπτουν από τον εκθετικό μετασχηματισμό των συντελεστών του μοντέλου πολυωνυμικής λογιστικής παλινδρόμησης. Οι δείκτες αυτοί εκφράζουν τη μεταβολή της πιθανότητας εμφάνισης μιας κατηγορίας σοβαρότητας σε σχέση με την κατηγορία αναφοράς. Στο παρόν μοντέλο ως κατηγορία αναφοράς χρησιμοποιήθηκε ο θανατηφόρος τραυματισμός ( $severity = 0$ ), ενώ ως κατηγορία αναφοράς για τη μεταβλητή της εθνικότητας χρησιμοποιήθηκαν οι οδηγοί ελληνικής εθνικότητας.

Με βάση τους λόγους πιθανοτήτων που εκτιμήθηκαν από το μοντέλο, εξετάστηκε κατά πόσο οι κατηγορίες εθνικότητας διαφοροποιούνται ως προς την

πιθανότητα εμφάνισης σοβαρού ή ελαφρού τραυματισμού σε σχέση με θανατηφόρο τραυματισμό. Τα αποτελέσματα δείχνουν ότι η εθνικότητα του οδηγού δεν εμφανίζει στατιστικά σημαντική επίδραση στην πιθανότητα σοβαρού τραυματισμού σε σχέση με θάνατο. Αντίθετα, για την κατηγορία του ελαφρού τραυματισμού παρατηρούνται στατιστικά σημαντικές διαφοροποιήσεις για ορισμένες κατηγορίες εθνικότητας. Συγκεκριμένα, οι οδηγοί από Ανατολική Ευρώπη και Νότια Ευρώπη εμφανίζουν στατιστικά σημαντική διαφοροποίηση στην πιθανότητα ελαφρού τραυματισμού σε σχέση με θανατηφόρο τραυματισμό, σε σύγκριση με τους οδηγούς ελληνικής εθνικότητας.

## **6. Συμπεράσματα**

### **6.1 Σύνοψη της έρευνας**

Η παρούσα διπλωματική εργασία είχε ως στόχο τη διερεύνηση της σχέσης μεταξύ της εθνικότητας των οδηγών και των χαρακτηριστικών των οδικών ατυχημάτων στην Ελλάδα. Ιδιαίτερη έμφαση δόθηκε τόσο στον αριθμό των παθόντων που καταγράφονται στα τροχαία συμβάντα όσο και στη σοβαρότητα των συνεπειών τους, εξετάζοντας τον ρόλο που διαδραματίζουν τα δημογραφικά χαρακτηριστικά των οδηγών, τα χαρακτηριστικά των οχημάτων και οι συνθήκες του ατυχήματος.

Για την πραγματοποίηση της ανάλυσης χρησιμοποιήθηκαν δεδομένα οδικών ατυχημάτων που προέρχονται από την ευρωπαϊκή βάση δεδομένων CARE (Community database on Accidents on the Roads in Europe) και την ΕΛΣΤΑΤ. Το σύνολο δεδομένων περιλαμβάνει ατυχήματα που καταγράφηκαν στην Ελλάδα κατά την περίοδο 2019–2023 και περιέχει πληροφορίες σχετικά με τα χαρακτηριστικά του ατυχήματος, των εμπλεκόμενων οχημάτων και των οδηγών.

Στο πλαίσιο της επεξεργασίας των δεδομένων πραγματοποιήθηκε αρχικά καθαρισμός και ομαδοποίηση των μεταβλητών, ώστε να δημιουργηθεί ένα συνεκτικό σύνολο δεδομένων κατάλληλο για στατιστική ανάλυση. Στη συνέχεια πραγματοποιήθηκε περιγραφική στατιστική ανάλυση με σκοπό την αρχική διερεύνηση της κατανομής των βασικών μεταβλητών και την κατανόηση των χαρακτηριστικών των οδικών ατυχημάτων.

Σε επόμενο στάδιο εφαρμόστηκε ανάλυση διακύμανσης (ANOVA), προκειμένου να διερευνηθεί αν υπάρχουν στατιστικά σημαντικές διαφορές στους δείκτες οδικών ατυχημάτων μεταξύ των διαφορετικών κατηγοριών εθνικότητας οδηγών. Για τον εντοπισμό των ομάδων που παρουσιάζουν σημαντικές διαφοροποιήσεις εφαρμόστηκε επιπλέον η μέθοδος πολλαπλών συγκρίσεων Tukey.

Τέλος, για τη διερεύνηση των παραγόντων που επηρεάζουν τον αριθμό των παθόντων και τη σοβαρότητα των συνεπειών των τροχαίων συμβάντων αναπτύχθηκαν διαφορετικά στατιστικά μοντέλα. Συγκεκριμένα, χρησιμοποιήθηκαν μοντέλα παλινδρόμησης καταμέτρησης (Poisson και Αρνητικής Διωνυμικής) για την ανάλυση του αριθμού των παθόντων, καθώς και μοντέλο πολωνυμικής λογιστικής παλινδρόμησης για την ανάλυση της σοβαρότητας των τραυματισμών. Η συνδυαστική εφαρμογή των μοντέλων αυτών επέτρεψε την ολοκληρωμένη διερεύνηση των παραγόντων που σχετίζονται με τα τροχαία συμβάντα και την εξαγωγή συμπερασμάτων σχετικά με τον ρόλο της εθνικότητας των οδηγών στην οδική ασφάλεια.

### **6.2 Κύρια αποτελέσματα της ανάλυσης**

Η στατιστική ανάλυση που πραγματοποιήθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας ανέδειξε μια σειρά από σημαντικά ευρήματα σχετικά με τους παράγοντες που σχετίζονται με τον αριθμό των παθόντων και τη σοβαρότητα των συνεπειών των τροχαίων συμβάντων στην Ελλάδα.

Αρχικά, από την περιγραφική ανάλυση των δεδομένων προέκυψε ότι ορισμένα δημογραφικά χαρακτηριστικά των οδηγών παρουσιάζουν έντονες διαφοροποιήσεις ως προς τη συμμετοχή τους στα οδικά ατυχήματα. Συγκεκριμένα, παρατηρήθηκε ότι το μεγαλύτερο ποσοστό των εμπλεκόμενων οδηγών ανήκει στην ηλικιακή ομάδα 25–49 ετών, γεγονός που μπορεί να αποδοθεί στην αυξημένη κινητικότητα και καθημερινή δραστηριότητα της συγκεκριμένης ηλικιακής ομάδας. Παράλληλα, διαπιστώθηκε ότι οι άνδρες οδηγοί εμπλέκονται σε σημαντικά μεγαλύτερο ποσοστό τροχαίων συμβάντων σε σχέση με τις γυναίκες οδηγούς, εύρημα το οποίο συμφωνεί με αποτελέσματα προηγούμενων ερευνών στον τομέα της οδικής ασφάλειας.

Η ανάλυση διακύμανσης (ANOVA) που εφαρμόστηκε για τη διερεύνηση της σχέσης μεταξύ της εθνικότητας των οδηγών και των δεικτών οδικών ατυχημάτων έδειξε ότι υπάρχουν στατιστικά σημαντικές διαφοροποιήσεις μεταξύ των κατηγοριών εθνικότητας. Συγκεκριμένα, τα αποτελέσματα έδειξαν ότι τόσο ο συνολικός δείκτης οδικών ατυχημάτων όσο και ο δείκτης σοβαρών και θανατηφόρων τραυματισμών (KSI) παρουσιάζουν σημαντικές διαφορές μεταξύ των εξεταζόμενων ομάδων οδηγών. Η εφαρμογή της μεθόδου πολλαπλών συγκρίσεων Tukey επέτρεψε τον εντοπισμό των κατηγοριών εθνικότητας που παρουσιάζουν τις σημαντικότερες διαφοροποιήσεις στους δείκτες αυτούς.

Στη συνέχεια, η εφαρμογή των μοντέλων παλινδρόμησης καταμέτρησης (Poisson και Αρνητικής Διωνυμικής) για την ανάλυση του αριθμού των παθόντων ανέδειξε ότι αρκετοί από τους παράγοντες που εξετάστηκαν εμφανίζουν στατιστικά σημαντική επίδραση στον αναμενόμενο αριθμό παθόντων. Μεταξύ των σημαντικότερων παραγόντων συγκαταλέγονται η ηλικιακή ομάδα και το φύλο του οδηγού, καθώς και ο τύπος του εμπλεκόμενου οχήματος. Τα αποτελέσματα των μοντέλων δείχνουν ότι τα δημογραφικά χαρακτηριστικά των οδηγών καθώς και τα χαρακτηριστικά των οχημάτων επηρεάζουν σημαντικά τον αριθμό των παθόντων που καταγράφονται στα τροχαία συμβάντα.

Ιδιαίτερο ενδιαφέρον παρουσιάζει το αποτέλεσμα που αφορά τη μεταβλητή της εθνικότητας του οδηγού, η οποία αποτελεί και το βασικό αντικείμενο διερεύνησης της παρούσας εργασίας. Από την ανάλυση προκύπτει ότι η εθνικότητα συνδέεται με διαφοροποιήσεις στον αριθμό των παθόντων που καταγράφονται στα τροχαία συμβάντα. Πιο συγκεκριμένα, οι οδηγοί ελληνικής εθνικότητας χρησιμοποιήθηκαν ως κατηγορία αναφοράς στα μοντέλα και οι υπόλοιπες κατηγορίες συγκρίνονται με αυτούς.

Τα αποτελέσματα δείχνουν ότι ορισμένες κατηγορίες οδηγών εμφανίζουν σημαντικά υψηλότερο αναμενόμενο αριθμό παθόντων σε σχέση με τους Έλληνες οδηγούς. Ειδικότερα, οι οδηγοί από την Αφρική, την Αμερική και την Αυστραλία παρουσιάζουν σημαντικά υψηλότερες τιμές λόγων επίπτωσης (IRR), γεγονός που υποδηλώνει μεγαλύτερο αριθμό παθόντων για παρόμοιους συνδυασμούς χαρακτηριστικών οδηγών και συνθηκών ατυχήματος. Παράλληλα, οι οδηγοί από τη Νότια Ευρώπη εμφανίζουν επίσης αυξημένο αριθμό παθόντων σε σχέση με την κατηγορία αναφοράς.

Αντίθετα, οι οδηγοί από την Ανατολική Ευρώπη και τη Βόρεια Ευρώπη εμφανίζουν χαμηλότερες τιμές λόγων επίπτωσης σε σχέση με τους Έλληνες οδηγούς, γεγονός που υποδηλώνει μικρότερο αναμενόμενο αριθμό παθόντων. Για τους οδηγούς από την Ασία δεν παρατηρείται στατιστικά σημαντική διαφοροποίηση σε σχέση με την κατηγορία αναφοράς. Τα αποτελέσματα αυτά υποδηλώνουν ότι η εθνικότητα των οδηγών συνδέεται με διαφορετικά επίπεδα εμπλοκής σε περιστατικά με παθόντες,

γεγονός που ενδέχεται να σχετίζεται με παράγοντες όπως η εξοικείωση με το οδικό περιβάλλον, η εμπειρία οδήγησης ή οι διαφοροποιήσεις στην οδηγική συμπεριφορά.

Αντίθετα, τα αποτελέσματα του πολυωνυμικού λογιστικού μοντέλου που εφαρμόστηκε για την ανάλυση της σοβαρότητας των συνεπειών δείχνουν ότι η επίδραση της εθνικότητας δεν εμφανίζεται με τον ίδιο τρόπο σε όλες τις κατηγορίες οδηγών. Συγκεκριμένα, μόνο ορισμένες κατηγορίες εθνικότητας εμφανίζουν στατιστικά σημαντική διαφοροποίηση σε σχέση με την κατηγορία αναφοράς των οδηγών ελληνικής εθνικότητας. Ενδεικτικά, οδηγοί από την Ανατολική Ευρώπη και τη Νότια Ευρώπη παρουσιάζουν στατιστικά σημαντική διαφοροποίηση ως προς την πιθανότητα εμφάνισης ελαφρού τραυματισμού σε σχέση με θανατηφόρο τραυματισμό, ενώ για αρκετές άλλες κατηγορίες εθνικότητας δεν παρατηρείται στατιστικά σημαντική διαφοροποίηση.

Επιπλέον, από την ανάλυση της σοβαρότητας των τραυματισμών προέκυψε ότι σημαντικό ρόλο διαδραματίζουν και άλλοι παράγοντες, όπως η ηλικιακή ομάδα και το φύλο του οδηγού, καθώς και ο τύπος του εμπλεκόμενου οχήματος. Συγκεκριμένα, τα ατυχήματα στα οποία εμπλέκονται δίκυκλα οχήματα παρουσιάζουν αυξημένη πιθανότητα σοβαρών συνεπειών σε σχέση με τα ατυχήματα που αφορούν επιβατικά οχήματα. Το εύρημα αυτό μπορεί να αποδοθεί στο χαμηλότερο επίπεδο προστασίας που προσφέρουν τα δίκυκλα στους οδηγούς τους.

Συνολικά, τα αποτελέσματα της ανάλυσης δείχνουν ότι ο αριθμός των παθόντων και η σοβαρότητα των συνεπειών των τροχαίων συμβάντων επηρεάζονται από έναν συνδυασμό δημογραφικών, χρονικών και περιβαλλοντικών παραγόντων. Παράλληλα, η εθνικότητα των οδηγών φαίνεται να σχετίζεται περισσότερο με τον αριθμό των παθόντων που καταγράφονται στα τροχαία συμβάντα παρά με τη σοβαρότητα των συνεπειών τους.

### **6.3 Συζήτηση αποτελεσμάτων**

Τα αποτελέσματα της παρούσας έρευνας επιβεβαιώνουν ότι τα οδικά ατυχήματα αποτελούν ένα σύνθετο φαινόμενο, το οποίο επηρεάζεται από ένα σύνολο δημογραφικών, περιβαλλοντικών και συμπεριφορικών παραγόντων. Η ανάλυση που πραγματοποιήθηκε δείχνει ότι τόσο ο αριθμός των παθόντων που καταγράφονται στα τροχαία συμβάντα όσο και η σοβαρότητα των συνεπειών τους δεν εξαρτώνται από έναν μόνο παράγοντα, αλλά από την αλληλεπίδραση πολλών διαφορετικών χαρακτηριστικών.

Ένα από τα βασικά ευρήματα της παρούσας εργασίας αφορά τον ρόλο των δημογραφικών χαρακτηριστικών των οδηγών. Συγκεκριμένα, η ηλικία και το φύλο του οδηγού εμφανίζονται ως σημαντικοί παράγοντες τόσο στον αριθμό των παθόντων όσο και στη σοβαρότητα των συνεπειών των τροχαίων συμβάντων. Οι άνδρες οδηγοί παρουσιάζουν υψηλότερη συμμετοχή σε τροχαία συμβάντα, ενώ συγκεκριμένες ηλικιακές ομάδες εμφανίζουν μεγαλύτερη πιθανότητα εμπλοκής σε σοβαρότερα ατυχήματα. Τα αποτελέσματα αυτά είναι συμβατά με τα ευρήματα της διεθνούς βιβλιογραφίας, σύμφωνα με τα οποία οι νεότεροι και οι άνδρες οδηγοί εμφανίζουν συχνότερα πιο επιθετική ή επικίνδυνη οδηγική συμπεριφορά.

Παράλληλα, ιδιαίτερη σημασία φαίνεται να έχει και ο τύπος του εμπλεκόμενου οχήματος. Τα αποτελέσματα της ανάλυσης έδειξαν ότι τα ατυχήματα στα οποία

εμπλέκονται μοτοσυκλέτες ή μοτοποδήλατα συνδέονται με μεγαλύτερη πιθανότητα σοβαρών τραυματισμών σε σχέση με τα ατυχήματα που αφορούν επιβατικά οχήματα. Το εύρημα αυτό είναι αναμενόμενο, καθώς οι οδηγοί δικύκλων είναι περισσότερο εκτεθειμένοι σε κίνδυνο λόγω της περιορισμένης προστασίας που προσφέρει το όχημα σε περίπτωση σύγκρουσης.

Ιδιαίτερο ενδιαφέρον παρουσιάζει επίσης η επίδραση της εθνικότητας των οδηγών. Από την ανάλυση του αριθμού των παθόντων προέκυψε ότι η εθνικότητα συνδέεται με διαφοροποιήσεις στη συμμετοχή των οδηγών σε τροχαία συμβάντα. Συγκεκριμένα, ορισμένες κατηγορίες οδηγών εμφανίζουν υψηλότερο αναμενόμενο αριθμό παθόντων σε σχέση με τους οδηγούς ελληνικής εθνικότητας, ενώ άλλες εμφανίζουν χαμηλότερες τιμές. Η διαφοροποίηση αυτή μπορεί να σχετίζεται με διάφορους παράγοντες, όπως η εξοικείωση των οδηγών με το οδικό περιβάλλον της χώρας, η εμπειρία οδήγησης σε διαφορετικά κυκλοφοριακά συστήματα ή ακόμη και διαφορές στην κυκλοφοριακή κουλτούρα μεταξύ διαφορετικών χωρών.

Ωστόσο, όσον αφορά τη σοβαρότητα των συνεπειών, τα αποτελέσματα δείχνουν ότι η επίδραση της εθνικότητας είναι πιο περιορισμένη. Μόνο ορισμένες κατηγορίες εθνικότητας εμφανίζουν στατιστικά σημαντική διαφοροποίηση σε σχέση με τους Έλληνες οδηγούς. Ενδεικτικά, οδηγοί από την Ανατολική Ευρώπη και τη Νότια Ευρώπη εμφανίζουν διαφοροποίηση ως προς την πιθανότητα εμφάνισης ελαφρού τραυματισμού σε σχέση με θανατηφόρο τραυματισμό, ενώ για πολλές άλλες κατηγορίες εθνικότητας δεν παρατηρείται στατιστικά σημαντική επίδραση.

Το εύρημα αυτό υποδηλώνει ότι η εθνικότητα ενδέχεται να σχετίζεται περισσότερο με τον αριθμό των παθόντων που καταγράφονται στα τροχαία συμβάντα παρά με το επίπεδο σοβαρότητας των συνεπειών τους. Το αποτέλεσμα αυτό είναι σε μεγάλο βαθμό συμβατό με τα ευρήματα της διεθνούς βιβλιογραφίας, σύμφωνα με τα οποία οι διαφορές που παρατηρούνται μεταξύ οδηγών διαφορετικής προέλευσης δεν οφείλονται αποκλειστικά στην εθνικότητα ως δημογραφικό χαρακτηριστικό, αλλά σχετίζονται με ένα ευρύτερο σύνολο παραγόντων, όπως η εξοικείωση με το τοπικό οδικό δίκτυο, οι οδηγικές συνήθειες και οι κυκλοφοριακές συνθήκες της χώρας στην οποία κινούνται.

## **6.4 Περιορισμοί της έρευνας**

Παρότι η παρούσα ανάλυση παρέχει χρήσιμα συμπεράσματα σχετικά με τους παράγοντες που επηρεάζουν τον αριθμό των παθόντων και τη σοβαρότητα των συνεπειών των τροχαίων συμβάντων, θα πρέπει να ληφθούν υπόψη ορισμένοι περιορισμοί της έρευνας.

Ένας από τους βασικούς περιορισμούς αφορά τη διαθεσιμότητα και τη φύση των δεδομένων που χρησιμοποιήθηκαν. Η ανάλυση βασίστηκε στη βάση δεδομένων CARE, η οποία περιλαμβάνει μόνο τα οδικά ατυχήματα που έχουν καταγραφεί από τις αρμόδιες αρχές. Συνεπώς, ατυχήματα μικρής σοβαρότητας που δεν δηλώθηκαν ή δεν καταγράφηκαν ενδέχεται να μην περιλαμβάνονται στο σύνολο των δεδομένων, γεγονός που μπορεί να επηρεάζει την πληρότητα της βάσης δεδομένων.

Επιπλέον, ορισμένες μεταβλητές που θα μπορούσαν να επηρεάζουν σημαντικά τον αριθμό των παθόντων ή τη σοβαρότητα των συνεπειών των οδικών ατυχημάτων δεν ήταν διαθέσιμες στο σύνολο δεδομένων. Παράγοντες όπως η οδηγική συμπεριφορά, η κατανάλωση αλκοόλ, η ταχύτητα, η χρήση κινητού τηλεφώνου ή η κόπωση του οδηγού δεν περιλαμβάνονται στα διαθέσιμα στοιχεία και επομένως δεν ήταν δυνατό να ληφθούν υπόψη στη στατιστική ανάλυση.

Ένας ακόμη περιορισμός αφορά το γεγονός ότι σε ορισμένες κατηγορίες εθνικότητας το διαθέσιμο δείγμα είναι σχετικά μικρό. Κατηγορίες με πολύ μικρό δείγμα (<50 παρατηρήσεις) ερμηνεύονται με προσοχή λόγω της υψηλής αβεβαιότητας των εκτιμήσεων. Το γεγονός αυτό μπορεί να επηρεάζει την αξιοπιστία των αποτελεσμάτων για συγκεκριμένες ομάδες οδηγών.

Τέλος, στην παρούσα εργασία χρησιμοποιήθηκε πληθυσμός αναφοράς που προέκυψε από τον συνδυασμό δεδομένων μόνιμου πληθυσμού και τουριστικών αφίξεων. Η προσέγγιση αυτή επιτρέπει την εκτίμηση συγκριτικών δεικτών μεταξύ διαφορετικών κατηγοριών υπηκοότητας, ωστόσο αποτελεί μία προσεγγιστική εκτίμηση της πραγματικής κυκλοφοριακής έκθεσης των οδηγών στο οδικό δίκτυο.

## **6.5 Προτάσεις για μελλοντική έρευνα**

Η παρούσα διπλωματική εργασία συμβάλλει στη διερεύνηση της σχέσης μεταξύ της εθνικότητας των οδηγών και των χαρακτηριστικών των οδικών ατυχημάτων στην Ελλάδα. Ωστόσο, υπάρχουν αρκετές δυνατότητες περαιτέρω ανάπτυξης της έρευνας στο συγκεκριμένο αντικείμενο.

Αρχικά, θα ήταν ιδιαίτερα χρήσιμη η ενσωμάτωση πρόσθετων μεταβλητών που σχετίζονται με τη συμπεριφορά των οδηγών, όπως η υπερβολική ταχύτητα, η κατανάλωση αλκοόλ, η χρήση κινητού τηλεφώνου κατά την οδήγηση ή η μη χρήση ζώνης ασφαλείας. Η συμπερίληψη τέτοιων μεταβλητών θα μπορούσε να συμβάλει στην πληρέστερη κατανόηση των παραγόντων που επηρεάζουν τόσο τον αριθμό των παθόντων όσο και τη σοβαρότητα των συνεπειών των οδικών ατυχημάτων.

Επιπλέον, θα μπορούσε να πραγματοποιηθεί ανάλυση που να λαμβάνει υπόψη πιο αναλυτικά δεδομένα σχετικά με την κυκλοφοριακή έκθεση των οδηγών, όπως τα χιλιόμετρα που διανύονται από κάθε κατηγορία οδηγών ή τη συχνότητα χρήσης του οδικού δικτύου. Η χρήση τέτοιων δεδομένων θα επέτρεπε την ακριβέστερη εκτίμηση των δεικτών επικινδυνότητας.

Μία ακόμη ενδιαφέρουσα κατεύθυνση για μελλοντική έρευνα θα ήταν η σύγκριση των αποτελεσμάτων της παρούσας μελέτης με αντίστοιχες αναλύσεις σε άλλες ευρωπαϊκές χώρες. Μια τέτοια συγκριτική προσέγγιση θα μπορούσε να συμβάλει στην καλύτερη κατανόηση του τρόπου με τον οποίο οι διαφορές στο οδικό δίκτυο, στο κυκλοφοριακό περιβάλλον και στις οδηγικές συνθήκες επηρεάζουν τα πρότυπα οδικών ατυχημάτων.

Συνολικά, η περαιτέρω διερεύνηση των παραγόντων που σχετίζονται με τα οδικά ατυχήματα μπορεί να συμβάλει σημαντικά στον σχεδιασμό αποτελεσματικότερων πολιτικών οδικής ασφάλειας και στη λήψη στοχευμένων μέτρων για τη μείωση των οδικών ατυχημάτων και των συνεπειών τους.

## **7. Βιβλιογραφία**

1. Yannis, G., Golias, J., & Papadimitriou, E. (2002). Accident risk of foreign drivers in various road environments. *Journal of Safety Research*, 38(4), pp.471-480.
2. Skliami, A. (2020). Συγκριτική ανάλυση χαρακτηριστικών οδικών ατυχημάτων ανά εθνικότητα οδηγού στην Ευρωπαϊκή Ένωση. Διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Πολιτικών Μηχανικών, Αθήνα.
3. Alogaiali, A., & Mannering, F. (2020). Unobserved heterogeneity and the effects of driver nationality on crash injury severities in Saudi Arabia. *Accident Analysis and Prevention*, 144, 105618.
4. Harootunian, K., Lee, B. H. Y., & Aultman-Hall, L. (2014). Odds of fault and factors for out-of-state drivers in crashes in four states of the USA. *Accident Analysis and Prevention*, 72, 32–43.
5. Leviäkangas, P. (1998). Accident risk of foreign drivers—the case of Russian drivers in South-Eastern Finland. *Accident Analysis and Prevention*, 30(2), 245–254.
6. Shawky, M., Kishta, E., Garib, A., & Almazrouei, A. (2017). Relationship between socio-demographic of drivers and traffic violations and crashes involvements. *World Congress on Civil, Structural, and Environmental Engineering*.
7. Nævestad, T.-O., Phillips, R. O., & Elvebakk, B. (2017). Internationalisation in road transport of goods in Norway: Safety outcomes, risk factors and policy implications. *Safety*, 3(3).
8. European Commission. (2024). CARE Database – Community Database on Accidents on the Roads in Europe. Available at: <https://road-safety.transport.ec.europa.eu>
9. Hellenic Statistical Authority (ELSTAT). (2024). Road Traffic Accidents Database. Available at: <https://www.statistics.gr>