ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Διερεύνηση παραγόντων που σχετίζονται με τη σοβαρότητα οδικών ατυχημάτων με μηχανοκίνητα δίκυκλα



Ζώης Ιωάννης Γιαννόπουλος

Επίβλεψη: Γιώργος Γιαννής, Καθηγητής ΕΜΠ

Αθήνα, Οκτώβριος 2026

ΕΥΧΑΡΙΣΤΙΕΣ

Καταρχήν, θα ήθελα να ευχαριστήσω θερμά τον κ. Γιώργο Γιαννή, Καθηγητή της Σχολής Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου, ο οποίος ανέλαβε την ανάθεση και επίβλεψη της παρούσας Διπλωματικής Εργασίας. Θα ήθελα να τον ευχαριστήσω ακόμη, για την καθοδήγησή του καθ΄ όλη την διάρκεια εκπόνησης της Εργασίας αυτής.

Επίσης, θα ήθελα να ευχαριστήσω εξίσου θερμά την Δάφνη Κυπρούλη, ερευνήτρια του Εθνικού Μετσόβιου Πολυτεχνείου, για την συνεχή και σημαντικότατη συνεισφορά της σε όλο το διάστημα της έρευνας και της συγγραφής, καθώς επίσης και για τα εποικοδομητικά σχόλια τα οποία οδήγησαν στην βελτιστοποίηση του τελικού αποτελέσματος της Εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους ανθρώπους που με στήριξαν σε όλη τη διάρκεια των σπουδών μου.

Αθήνα, Οκτώβριος 2026

Ζώης Ιωάννης Γιαννόπουλος

Διερεύνηση παραγόντων που σχετίζονται με τη σοβαρότητα οδικών ατυχημάτων με μηχανοκίνητα δίκυκλα

Ζώης Ιωάννης Γιαννόπουλος

ΣΥΝΟΨΗ

Στόχο της παρούσας Διπλωματικής Εργασίας αποτελεί η διερεύνηση των παραγόντων που σχετίζονται με τη σοβαρότητα των οδικών ατυχημάτων με μηχανοκίνητα δίκυκλα (Powered Two-Wheelers, PTWs) στην Ελλάδα. Η έρευνα αξιοποιεί δεδομένα του Παρατηρητηρίου Οδικής Ασφάλειας του ΕΜΠ και της ΕΛ.ΣΤΑΤ. για το έτος 2023, τα οποία περιλαμβάνουν 6.585 καταγραφές παθόντων με πληροφορίες για τα χαρακτηριστικά του αναβάτη, του οχήματος, του ατυχήματος και του οδικού περιβάλλοντος.

Αναπτύχθηκαν δύο Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων (GLMM) στη γλώσσα προγραμματισμού R, με τη γεωγραφική περιοχή ως τυχαία επίδραση: ένα μοντέλο παλινδρόμησης Poisson για τον αριθμό νεκρών και βαριά τραυματιών (KSI) και ένα μοντέλο Αρνητικής Διωνυμικής παλινδρόμησης για τον αριθμό ελαφρά τραυματιών. Η χωριστή μοντελοποίηση των δύο κατηγοριών σοβαρότητας επέτρεψε τη διάκριση μεταξύ των παραγόντων που επηρεάζουν κυρίως τη συχνότητα και εκείνων που επηρεάζουν κυρίως τη σοβαρότητα των συνεπειών.

Τα αποτελέσματα έδειξαν ότι ο αναμενόμενος αριθμός νεκρών και βαριά τραυματιών είναι υψηλότερος στα ατυχήματα εκτός κατοικημένων περιοχών, στις περιπτώσεις θετικού ελέγχου αλκοόλ, στους αναβάτες ηλικίας 65 ετών και άνω, κατά τη νύχτα, στις εκτροπές και τις προσκρούσεις σε εμπόδια, καθώς και στα μεγάλα και τα μικρά δίκυκλα, ενώ χαμηλότερος για τις γυναίκες και για τα ατυχήματα σε κόμβους. Οι περισσότερες από τις συνθήκες αυτές συνδέονται ταυτόχρονα με λιγότερους ελαφρά τραυματίες, γεγονός που δείχνει ότι μετατοπίζουν την έκβαση των ατυχημάτων προς σοβαρότερες συνέπειες.

Λέξεις κλειδιά: οδική ασφάλεια, μηχανοκίνητα δίκυκλα, σοβαρότητα τραυματισμού, παλινδρόμηση Poisson, Αρνητική Διωνυμική παλινδρόμηση, μοντέλα μικτών επιδράσεων.

Investigation of factors associated with the severity of road crashes involving powered two-wheelers

Zois Ioannis Giannopoulos

ABSTRACT

The objective of this Diploma Thesis is to investigate the factors associated with the severity of road crashes involving powered two-wheelers (PTWs) in Greece. The study uses 2023 data from the NTUA Road Safety Observatory and the Hellenic Statistical Authority (ELSTAT), comprising 6,585 casualty records with information on rider, vehicle, crash and road-environment characteristics.

Two generalized linear mixed models (GLMMs) were developed in the R programming language, with geographical region included as a random effect: a Poisson regression model for the number of killed or seriously injured casualties (KSI) and a negative binomial regression model for the number of slightly injured casualties. Modelling the two severity levels separately made it possible to distinguish between factors that mainly affect the frequency of casualties and factors that mainly affect the severity of the consequences.

The results showed that the expected number of killed or seriously injured casualties is higher for crashes outside built-up areas, for positive alcohol tests, for riders aged 65 and over, at night, for run-off-road crashes and collisions with fixed obstacles, and for both large and small PTWs, while it is lower for female casualties and for crashes at junctions. Most of these conditions are simultaneously associated with fewer slight injuries, indicating that they shift crash outcomes towards more severe consequences.

Keywords: road safety, powered two-wheelers, injury severity, Poisson regression, negative binomial regression, mixed-effects models.

ΠΕΡΙΛΗΨΗ

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η **διερεύνηση των παραγόντων που σχετίζονται με τη σοβαρότητα των οδικών ατυχημάτων με μηχανοκίνητα δίκυκλα στην Ελλάδα**. Για την επίτευξη του στόχου αυτού αξιοποιήθηκαν δεδομένα οδικών ατυχημάτων για το έτος 2023, τα οποία περιλαμβάνουν πληροφορίες για τα χαρακτηριστικά των παθόντων, των οχημάτων, των ατυχημάτων και των συνθηκών υπό τις οποίες συνέβησαν.

Αρχικά, πραγματοποιήθηκε **βιβλιογραφική ανασκόπηση** (Κεφάλαιο 2), με σκοπό την αναγνώριση των βασικών παραγόντων που επηρεάζουν τη σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα. Στη συνέχεια, παρουσιάστηκε το **θεωρητικό υπόβαθρο** (Κεφάλαιο 3), στο οποίο περιγράφονται οι στατιστικές μέθοδοι και οι απαραίτητοι έλεγχοι για την ανάπτυξη και αξιολόγηση των μοντέλων.

Έπειτα, ακολούθησε η **συλλογή και επεξεργασία των δεδομένων** (Κεφάλαιο 4), με τη διαμόρφωση της τελικής βάσης δεδομένων από τα στοιχεία της ΕΛ.ΣΤΑΤ. Η αρχική βάση περιλάμβανε 6.585 καταγραφές παθόντων, ενώ μετά την απαραίτητη επεξεργασία διαμορφώθηκαν οι βάσεις δεδομένων που χρησιμοποιήθηκαν στη στατιστική ανάλυση, με 4.069 εγγραφές.

Τέλος, στο κεφάλαιο της **εφαρμογής της μεθοδολογίας και της ανάλυσης των αποτελεσμάτων** (Κεφάλαιο 5), πραγματοποιήθηκε η ανάπτυξη των τελικών στατιστικών μοντέλων. Συγκεκριμένα, εφαρμόστηκαν δύο **Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων (GLMM)**, με σκοπό τη χωριστή διερεύνηση των παραγόντων που σχετίζονται με τον αναμενόμενο αριθμό παθόντων ανά κατηγορία σοβαρότητας:

- **Μοντέλο παλινδρόμησης Poisson**, με εξαρτημένη μεταβλητή τον αριθμό νεκρών και σοβαρά τραυματιών (KSI).
- **Μοντέλο Αρνητικής Διωνυμικής παλινδρόμησης (Negative Binomial)**, με εξαρτημένη μεταβλητή τον αριθμό ελαφρά τραυματιών.

Και στα δύο μοντέλα συμπεριλήφθηκε η **γεωγραφική περιοχή ως τυχαία επίδραση**, ώστε να ληφθούν υπόψη οι διαφορές μεταξύ των περιοχών της χώρας. Μετά την εφαρμογή των απαραίτητων στατιστικών ελέγχων, προέκυψαν τα τελικά μοντέλα, τα οποία επέτρεψαν την αναγνώριση των σημαντικότερων παραγόντων που συνδέονται με τις διαφορετικές κατηγορίες σοβαρότητας των ατυχημάτων.

Παρακάτω δίνεται ο συγκεντρωτικός πίνακας των αποτελεσμάτων των αναλύσεων για τις μεθόδους που χρησιμοποιήθηκαν και για τα δύο μοντέλα:

Πίνακας: Συγκεντρωτικός πίνακας αποτελεσμάτων των δύο τελικών μοντέλων

Κατηγορία	Μοντέλο 1: Νεκροί και βαριά τραυματίες		Μοντέλο 2: Ελαφρά τραυματίες	
	β	Επίδραση	β	Επίδραση
Κατηγορία αναφοράς: Οδηγός				
Επιβάτης	-	-	-0,344	**Μείωση**
Κατηγορία αναφοράς: Άνδρας				
Γυναίκα	-0,696	**Μείωση**	-0,375	**Μείωση**
Κατηγορία αναφοράς: Μη εθνική οδός				
Εθνική οδός	-	-	-0,250	**Μείωση**
Κατηγορία αναφοράς: Εντός κατοικημένης περιοχής				
Εκτός κατοικημένης περιοχής	0,681	**Αύξηση**	-0,282	**Μείωση**
Κατηγορία αναφοράς: Καθημερινή				
Σαββατοκύριακο	-	-	-0,270	**Μείωση**
Κατηγορία αναφοράς: Ατύχημα εκτός κόμβου				
Ατύχημα σε κόμβο	-0,200	**Μείωση**	0,252	**Αύξηση**
Κατηγορία αναφοράς: Ηλικία 25-49 ετών				
0-14 ετών	-	-	-0,301	*Μη σημαντική*
15-18 ετών	-	-	-0,518	**Μείωση**
≤18 ετών	0,063	*Μη σημαντική*	-	-
19-24 ετών	-0,078	*Μη σημαντική*	-0,366	**Μείωση**
50-64 ετών	0,141	*Μη σημαντική*	-0,283	**Μείωση**
65 ετών και άνω	0,402	**Αύξηση**	-0,675	**Μείωση**
Άγνωστη ηλικία	-0,688	**Μείωση**	-0,739	**Μείωση**
Κατηγορία αναφοράς: Ημέρα				
Νύχτα	0,271	**Αύξηση**	-0,276	**Μείωση**
Κατηγορία αναφοράς: Δύο εμπλεκόμενα οχήματα				
Ένα όχημα	-	-	-0,251	**Μείωση**
Τρία ή περισσότερα οχήματα	-	-	-0,701	**Μείωση**
Κατηγορία αναφοράς: Καλές καιρικές συνθήκες				
Λοιπές καιρικές συνθήκες	-	-	-0,548	**Μείωση**
Κατηγορία αναφοράς: Σύγκρουση με χρήστες της οδού				
Εκτροπή	0,276	**Αύξηση**	-0,281	**Μείωση**
Λοιπές συγκρούσεις	0,447	**Αύξηση**	-0,530	**Μείωση**
Ανατροπή/λοιπά	-0,004	*Μη σημαντική*	-0,434	**Μείωση**
Κατηγορία αναφοράς: Μεσαίο δίκυκλο				
Μικρό δίκυκλο	0,235	**Αύξηση**	-0,246	**Μείωση**
Μεγάλο δίκυκλο	0,343	**Αύξηση**	-0,110	**Μείωση**
Κατηγορία αναφοράς: Αρνητικός/μηδενικός έλεγχος αλκοόλ				
Χαμηλά θετικός έλεγχος αλκοόλ	-	-	-0,450	**Μείωση**
Υψηλά θετικός έλεγχος αλκοόλ	-	-	-0,413	**Μείωση**
Θετικός έλεγχος αλκοόλ	0,648	**Αύξηση**	-	-
Άγνωστος/δεν διενεργήθηκε	1,175	**Αύξηση**	-0,398	**Μείωση**

***Σημείωση**: «-» σημαίνει ότι η μεταβλητή ή η κατηγορία δεν περιλαμβάνεται στο αντίστοιχο τελικό μοντέλο.*

Κατά την εκπόνηση της παρούσας Διπλωματικής Εργασίας προέκυψε μια σειρά συμπερασμάτων, τα οποία συνδέονται άμεσα με τον αρχικό στόχο της διερεύνησης των παραγόντων που σχετίζονται με τη **σοβαρότητα των ατυχημάτων** με μηχανοκίνητα δίκυκλα. Τα σημαντικότερα συμπεράσματα που προέκυψαν από την περιγραφική και τη στατιστική ανάλυση συνοψίζονται ως εξής:

- Η **σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα επηρεάζεται συνδυαστικά από πολλούς διαφορετικούς παράγοντες**, οι οποίοι σχετίζονται με τα χαρακτηριστικά του αναβάτη, τον τύπο του δικύκλου, τις συνθήκες κυκλοφορίας και το οδικό περιβάλλον. Η διεθνής βιβλιογραφία επιβεβαιώνει ότι οι παράγοντες αυτοί συνδέονται με αυξημένη πιθανότητα σοβαρού ή θανατηφόρου τραυματισμού.
- Τα **ατυχήματα εκτός κατοικημένων περιοχών παρουσιάζουν σημαντικά μεγαλύτερη σοβαρότητα** σε σύγκριση με εκείνα εντός κατοικημένων περιοχών. Συγκεκριμένα, σύμφωνα με το μοντέλο Poisson, ο αναμενόμενος αριθμός νεκρών και σοβαρά τραυματιών είναι περίπου **1,98 φορές υψηλότερος** εκτός κατοικημένων περιοχών, όταν οι υπόλοιποι παράγοντες παραμένουν σταθεροί, γεγονός που σχετίζεται με τις υψηλότερες ταχύτητες κυκλοφορίας στο υπεραστικό οδικό δίκτυο.
- Η **ηλικία του αναβάτη αποτελεί σημαντικό παράγοντα σοβαρότητας**, καθώς οι μεγαλύτεροι σε ηλικία αναβάτες παρουσιάζουν αυξημένο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών. Ειδικότερα, η ηλικιακή ομάδα **65 ετών και άνω** εμφανίζει τη μεγαλύτερη αναλογία σοβαρών και θανατηφόρων τραυματισμών, γεγονός που μπορεί να σχετίζεται με τη μεγαλύτερη σωματική ευπάθεια των ηλικιωμένων.
- Η **οδήγηση υπό την επήρεια αλκοόλ συνδέεται με σοβαρότερες συνέπειες** στα ατυχήματα μηχανοκίνητων δικύκλων. Συγκεκριμένα, οι περιπτώσεις με θετικό έλεγχο αλκοόλ παρουσιάζουν περίπου **1,91 φορές υψηλότερο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών** σε σύγκριση με τις περιπτώσεις αρνητικού ή μηδενικού ελέγχου, με τις υπόλοιπες μεταβλητές σταθερές.
- Η **νυχτερινή οδήγηση συνδέεται με αυξημένη σοβαρότητα των ατυχημάτων**. Παρότι κατά τις νυχτερινές και πρώτες πρωινές ώρες καταγράφονται λιγότεροι παθόντες, το ποσοστό των σοβαρών και θανατηφόρων τραυματισμών είναι υψηλότερο. Το εύρημα αυτό ενδέχεται να

σχετίζεται με τη μειωμένη ορατότητα, τις υψηλότερες ταχύτητες και την πιθανή κατανάλωση αλκοόλ.

- Ο **τύπος του ατυχήματος παρουσιάζει σημαντική συσχέτιση με τη σοβαρότητα των τραυματισμών**. Ειδικότερα, οι προσκρούσεις σε σταθερά εμπόδια, οι εκτροπές και οι λοιπές συγκρούσεις εμφανίζουν αυξημένη σοβαρότητα. Το αποτέλεσμα αυτό αναδεικνύει τη σημασία των συνθηκών της σύγκρουσης και των χαρακτηριστικών του οδικού περιβάλλοντος.
- Ο **τύπος και ο κυβισμός του δικύκλου σχετίζονται με τη σοβαρότητα των ατυχημάτων**. Σύμφωνα με τα αποτελέσματα του μοντέλου, τόσο τα μικρά όσο και τα μεγάλα μηχανοκίνητα δίκυκλα εμφανίζουν υψηλότερο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών σε σύγκριση με τα μεσαίου κυβισμού δίκυκλα. Ενδεχομένως αυτές οι δύο κατηγορίες οχημάτων εμφανίζουν τη μεγαλύτερη απόκλιση από τη μέση ταχύτητα κυκλοφορίας.
- Η **γεωγραφική περιοχή αποτελεί σημαντικό παράγοντα διαφοροποίησης των αποτελεσμάτων**, καθώς η συμπερίληψή της ως τυχαίας επίδρασης βελτίωσε σημαντικά και τα δύο στατιστικά μοντέλα. Μάλιστα, οι γεωγραφικές διαφοροποιήσεις ήταν εντονότερες στο μοντέλο των νεκρών και σοβαρά τραυματιών. Το εύρημα αυτό υποδηλώνει ότι υπάρχουν ιδιαίτερα χαρακτηριστικά μεταξύ των περιοχών, τα οποία δεν εξηγούνται πλήρως από τις διαθέσιμες μεταβλητές και απαιτούν περαιτέρω διερεύνηση.
- Τέλος, η ανάλυση ανέδειξε ότι **ο αριθμός των καταγεγραμμένων παθόντων δεν αποτελεί από μόνος του ένδειξη της σοβαρότητας των ατυχημάτων**. Ορισμένες συνθήκες, όπως η οδήγηση εκτός κατοικημένων περιοχών και κατά τις νυχτερινές ώρες, συνδέονται με λιγότερες καταγραφές αλλά υψηλότερη αναλογία σοβαρών και θανατηφόρων τραυματισμών. Συνεπώς, για την αποτελεσματικότερη βελτίωση της οδικής ασφάλειας απαιτείται η χωριστή εξέταση της συχνότητας και της σοβαρότητας των ατυχημάτων.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ..... 1
1.1 Γενική Ανασκόπηση.....1
1.2 Στόχος Διπλωματικής Εργασίας.....3
1.3 Μεθοδολογία.....4
1.4 Δομή Διπλωματικής Εργασίας.....5
2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ..... 7
2.1 Εισαγωγή.....7
2.2 Οδική ασφάλεια δικύκλων.....7
2.3 Παράγοντες που επηρεάζουν τη σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα.....8
2.3.1 Χαρακτηριστικά και συμπεριφορά του αναβάτη.....8
2.3.2 Χαρακτηριστικά οχήματος και προστατευτικός εξοπλισμός.....9
2.3.3 Τύπος ατυχήματος και αριθμός εμπλεκόμενων οχημάτων..... 10
2.3.4 Οδικό περιβάλλον και υποδομή..... 11
2.3.5 Περιβαλλοντικοί και χρονικοί παράγοντες.....12
2.4 Σύνοψη..... 12
3. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....14
3.1 Εισαγωγή.....14
3.2 Βασικές έννοιες στατιστικής.....14
3.2.1 Πληθυσμός και δείγμα.....14
3.2.2 Μεταβλητές..... 14
3.3 Βασικές Μέθοδοι Στατιστικής Ανάλυσης.....16
3.4 Στατιστικοί έλεγχοι και κριτήρια αποδοχής μοντέλου.....18
4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΣΤΟΙΧΕΙΩΝ.....24
4.1 Εισαγωγή.....24
4.2 Συλλογή δεδομένων..... 24
4.2.1 Πηγή δεδομένων.....24
4.2.2 Διαδικασία συλλογής στοιχείων..... 25
4.3 Επεξεργασία δεδομένων..... 27
4.3.1 Διαμόρφωση δεδομένων.....27
4.3.2 Σύνθεση αλγορίθμου.....38
4.3.3 Γεωγραφική κατανομή.....39
4.4 Σύνοψη και παρατηρήσεις.....42
5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ - ΑΠΟΤΕΛΕΣΜΑΤΑ.....44
5.1 Εισαγωγή.....44
5.2 Διερεύνηση κατανομής δεδομένων.....44
5.3 Ανάπτυξη μοντέλων παλινδρόμησης στο λογισμικό R..... 45

5.3.1 Μοντέλα παλινδρόμησης Poisson.....46
5.3.2 Μοντέλα Αρνητικής Διωνυμικής παλινδρόμησης..... 46
5.3.3 Σύγκριση μοντέλων σταθερών και μικτών επιδράσεων..... 47
5.3.4 Επιλογή ανεξάρτητων μεταβλητών..... 48
5.4 Αποτελέσματα τελικών μοντέλων..... 49
5.4.1 Μοντέλο σοβαρών και θανατηφόρων τραυματισμών (KSI)..... 49
5.4.2 Μοντέλο ελαφρών τραυματισμών..... 51
5.4.3 Διαγνωστικός έλεγχος και αξιολόγηση μοντέλων.....54
5.5 Ανάλυση γεωγραφικών διαφοροποιήσεων.....56
5.6 Σύγκριση αποτελεσμάτων των δύο μοντέλων..... 57
5.6.1 Πλαίσιο ερμηνείας της σύγκρισης..... 57
5.6.2 Μεταβλητές με ίδια κατεύθυνση στα δύο μοντέλα.....59
5.6.3 Μεταβλητές με αντίθετη κατεύθυνση στα δύο μοντέλα..... 60
5.6.4 Μεταβλητές σημαντικές μόνο σε ένα από τα δύο μοντέλα.....62
5.6.5 Συνολική αποτίμηση της σύγκρισης.....63
6. ΣΥΜΠΕΡΑΣΜΑΤΑ..... 65
6.1 Σύνοψη αποτελεσμάτων.....65
6.2 Συνολικά συμπεράσματα.....67
6.3 Προτάσεις βελτίωσης της οδικής ασφάλειας.....69
6.4 Προτάσεις για περαιτέρω έρευνα.....70
7. ΒΙΒΛΙΟΓΡΑΦΙΑ..... 71

1. ΕΙΣΑΓΩΓΗ

1.1 Γενική Ανασκόπηση

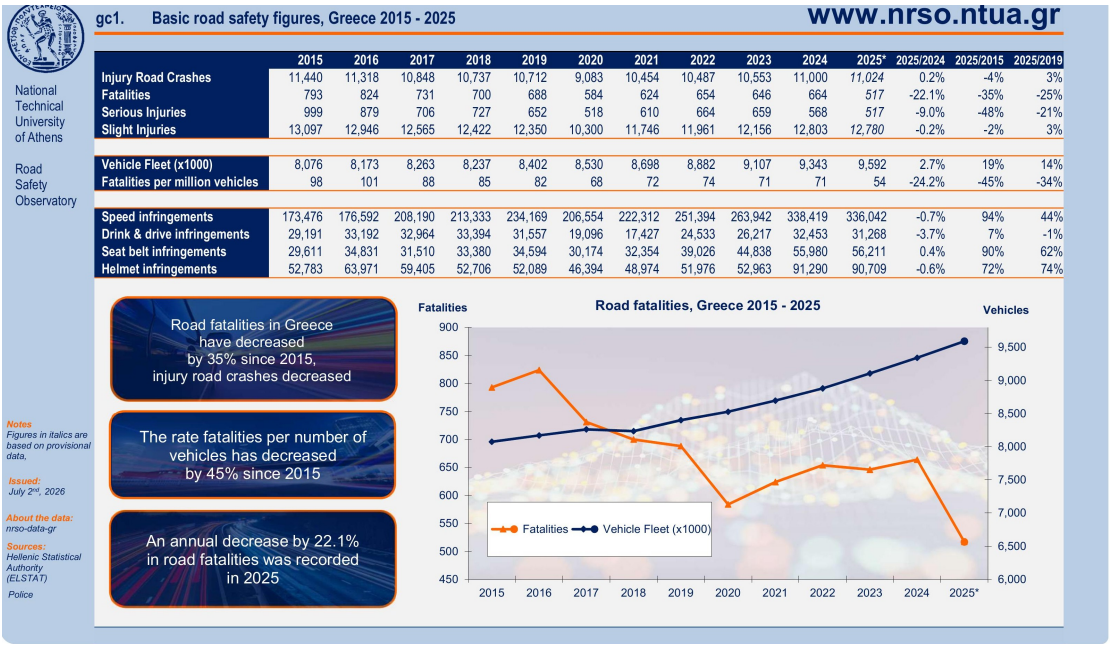
Οι οδικές μετακινήσεις αποτελούν βασικό στοιχείο της καθημερινής ζωής και της οικονομικής και κοινωνικής δραστηριότητας. Παράλληλα, τα οδικά ατυχήματα εξακολουθούν να αποτελούν ένα από τα σημαντικότερα προβλήματα δημόσιας υγείας και οδικής ασφάλειας διεθνώς. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, **περίπου 1,16 εκατομμύρια άνθρωποι χάνουν τη ζωή τους κάθε χρόνο σε οδικά ατυχήματα**, ενώ οι οδικοί τραυματισμοί αποτελούν την κυριότερη αιτία θανάτου για παιδιά και νέους ηλικίας 5 έως 29 ετών. **Περισσότεροι από τους μισούς θανάτους αφορούν ευάλωτους χρήστες της οδού**, στους οποίους περιλαμβάνονται οι **πεζοί**, **οι ποδηλάτες και οι αναβάτες μηχανοκίνητων δικύκλων** (World Health Organization, 2026).

Τα μηχανοκίνητα δίκυκλα (Powered Two-Wheelers, PTWs) παρουσιάζουν ιδιαίτερο ενδιαφέρον από πλευράς οδικής ασφάλειας, καθώς ο αναβάτης διαθέτει **περιορισμένη φυσική προστασία σε περίπτωση σύγκρουσης**. Η αυξημένη τρωτότητα των αναβατών, σε συνδυασμό με τα ιδιαίτερα δυναμικά χαρακτηριστικά των δικύκλων και την αλληλεπίδρασή τους με τα υπόλοιπα οχήματα, οδηγεί σε δυσανάλογα υψηλή συμμετοχή τους στους **σοβαρούς και θανατηφόρους τραυματισμούς**. Για τον λόγο αυτό, η συστηματική παρακολούθηση της εξέλιξης των ατυχημάτων αυτών και η σύγκρισή τους με τις υπόλοιπες κατηγορίες οχημάτων αποτελεί απαραίτητο πρώτο στάδιο για την κατανόηση του προβλήματος.

Στην Ελλάδα, το 2024 καταγράφηκαν **11.000 οδικά ατυχήματα** με θάνατο ή τραυματισμό. Από αυτά προέκυψαν συνολικά 14.035 παθόντες, εκ των οποίων **664 νεκροί**, 568 βαριά τραυματίες και 12.803 ελαφρά τραυματίες (ΕΛ.ΣΤΑΤ., 2026). Για το 2025, τα προσωρινά στοιχεία δείχνουν **σημαντική βελτίωση ως προς τους θανάτους**, οι οποίοι εκτιμώνται σε 517, δηλαδή μειωμένοι κατά 22,1% σε σχέση με το 2024 (NTUA Road Safety Observatory, 2026a). Παρά τη συνολική αυτή θετική εξέλιξη, η επιμέρους εικόνα ανά κατηγορία οχήματος αναδεικνύει ότι **τα μηχανοκίνητα δίκυκλα εξακολουθούν να αποτελούν κρίσιμο ζήτημα οδικής ασφάλειας**.

Στην παρακάτω εικόνα παρουσιάζονται δεδομένα για την εξέλιξη βασικών δεικτών οδικής ασφάλειας στην Ελλάδα την περίοδο 2015 - 2025, σύμφωνα με στοιχεία που συγκεντρώθηκαν από το Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ). Η διαχρονική εξέλιξη των ατυχημάτων, των θυμάτων και των καταγεγραμμένων παραβάσεων επιτρέπει τη

συνολική αποτύπωση των μεταβολών στην οδική ασφάλεια της χώρας και παρέχει το γενικό πλαίσιο για την ειδικότερη διερεύνηση των ατυχημάτων με **μηχανοκίνητα δίκυκλα**.



National Technical University of Athens

Road Safety Observatory

gc1. Basic road safety figures, Greece 2015 - 2025

www.nrso.ntua.gr

	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025*	2025/2024	2025/2015	2025/2019
Injury Road Crashes	11,440	11,318	10,848	10,737	10,712	9,083	10,454	10,487	10,553	11,000	*11,024*	0.2%	-4%	3%
Fatalities	793	824	731	700	688	584	624	654	646	664	*517*	-22.1%	-35%	-25%
Serious Injuries	999	879	706	727	652	518	610	664	659	568	*517*	-9.0%	-48%	-21%
Slight Injuries	13,097	12,946	12,565	12,422	12,350	10,300	11,746	11,961	12,156	12,803	*12,780*	-0.2%	-2%	3%
Vehicle Fleet (x1000)	8,076	8,173	8,263	8,237	8,402	8,530	8,698	8,882	9,107	9,343	9,592	2.7%	19%	14%
Fatalities per million vehicles	98	101	88	85	82	68	72	74	71	71	54	-24.2%	-45%	-34%
Speed infringements	173,476	176,592	208,190	213,333	234,169	206,554	222,312	251,394	263,942	338,419	336,042	-0.7%	94%	44%
Drink & drive infringements	29,191	33,192	32,964	33,394	31,557	19,096	17,427	24,533	26,217	32,453	31,268	-3.7%	7%	-1%
Seat belt infringements	29,611	34,831	31,510	33,380	34,594	30,174	32,354	39,026	44,838	55,980	56,211	0.4%	90%	62%
Helmet infringements	52,783	63,971	59,405	52,706	52,089	46,394	48,974	51,976	52,963	91,290	90,709	-0.6%	72%	74%

Road fatalities in Greece have decreased by 35% since 2015, injury road crashes decreased

The rate fatalities per number of vehicles has decreased by 45% since 2015

An annual decrease by 22.1% in road fatalities was recorded in 2025

Road fatalities, Greece 2015 - 2025

Fatalities

Vehicles

Fatalities

Vehicle Fleet (x1000)

Notes
Figures in italics are based on provisional data,

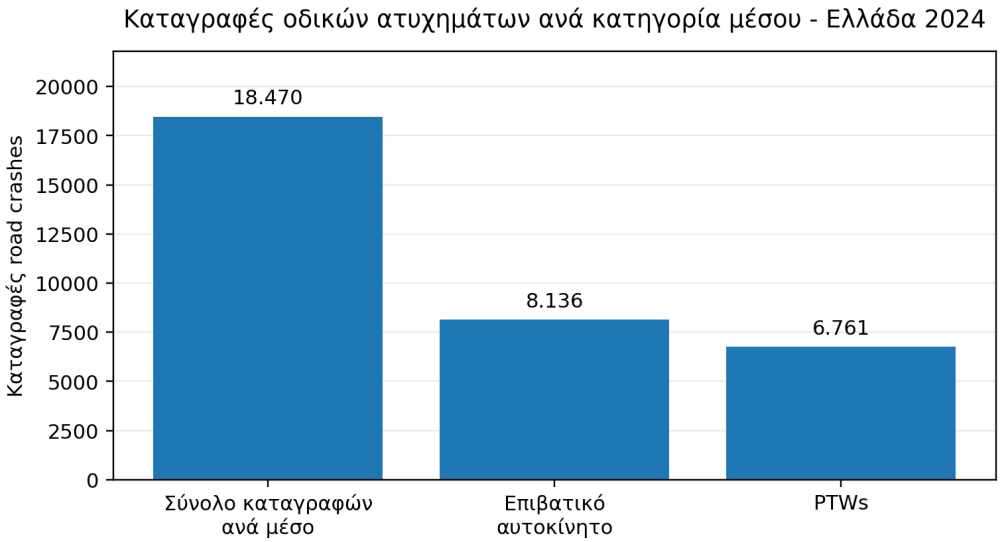
Issued:
July 2nd, 2026

About the data:
nrso-data-gr

Sources:
Hellenic Statistical Authority (ELSTAT)
Police

Εικόνα 1.1: Στατιστικά οδικής ασφάλειας στην Ελλάδα (2015-2025) (Πηγή: NTUA Road Safety Observatory)

Ειδικότερα, σύμφωνα με τα στοιχεία του Παρατηρητηρίου Οδικής Ασφάλειας του ΕΜΠ (NRSO) για το 2024, οι καταγραφές **οδικών ατυχημάτων** ανά κατηγορία μέσου ανέρχονται σε 8.136 για τα επιβατικά αυτοκίνητα και σε **6.761 για τα μηχανοκίνητα δίκυκλα (NTUA Road Safety Observatory, 2026)**. Οι τιμές αυτές αποτελούν καταγραφές ανά μεταφορικό μέσο και δεν ταυτίζονται με τον αριθμό των μοναδικών ατυχημάτων της χώρας, καθώς ένα ατύχημα με περισσότερες κατηγορίες εμπλεκόμενων χρηστών μπορεί να αντιστοιχεί σε περισσότερες από μία καταγραφές ανά μέσο. Η σύγκριση είναι ωστόσο χρήσιμη για την **αποτύπωση της συχνότητας παρουσίας των μηχανοκίνητων δικύκλων στα καταγεγραμμένα ατυχήματα.**



Διάγραμμα 1.1: Καταγραφές οδικών ατυχημάτων ανά κατηγορία μέσου, Ελλάδα 2024 (Πηγή: ΕΛ.ΣΤΑΤ.)

Με βάση τα παραπάνω, καθίσταται **αναγκαία η συστηματική διερεύνηση των παραγόντων που σχετίζονται με τη σοβαρότητα των ατυχημάτων μηχανοκίνητων δικύκλων στην Ελλάδα**. Η αναγνώριση των χαρακτηριστικών που αυξάνουν ή μειώνουν την πιθανότητα σοβαρού ή θανατηφόρου αποτελέσματος μπορεί να συμβάλει στην πληρέστερη κατανόηση του προβλήματος και στην υποστήριξη στοχευμένων μέτρων οδικής ασφάλειας.

1.2 Στόχος Διπλωματικής Εργασίας

Ο κύριος στόχος της παρούσας Διπλωματικής Εργασίας είναι η **διερεύνηση των παραγόντων που σχετίζονται με τις συνέπειες των οδικών ατυχημάτων με εμπλοκή μηχανοκίνητων δικύκλων στην Ελλάδα**, με τη χρήση προηγμένων στατιστικών μοντέλων. Ειδικότερα, επιδιώκεται να **προσδιοριστούν τα χαρακτηριστικά** του χρήστη του δικύκλου, του οχήματος, του ατυχήματος και του περιβαλλοντικού πλαισίου **που συνδέονται με μεταβολές στον** αναμενόμενο αριθμό ελαφρά τραυματιών και, ξεχωριστά, στον αναμενόμενο αριθμό **νεκρών ή σοβαρά τραυματιών** (KSI).

Για την επίτευξη του παραπάνω στόχου, χρησιμοποιήθηκαν δεδομένα οδικών ατυχημάτων με εμπλοκή μηχανοκίνητων δικύκλων στην Ελλάδα, τα οποία περιλαμβάνουν πληροφορίες για τα χαρακτηριστικά των εμπλεκόμενων χρηστών, των οχημάτων και των συνθηκών υπό τις οποίες συνέβη το ατύχημα. Η στατιστική διερεύνηση πραγματοποιήθηκε με **Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων** (Generalized Linear Mixed Models, GLMM), ώστε να εξεταστούν

χωριστά οι παράγοντες που σχετίζονται με τον αναμενόμενο αριθμό παθόντων στις δύο κατηγορίες σοβαρότητας.

Μέσω των αποτελεσμάτων επιδιώκεται η **ανάδειξη των σημαντικότερων παραγόντων που σχετίζονται με αυξημένο ή μειωμένο αναμενόμενο αριθμό παθόντων ανά κατηγορία σοβαρότητας** και η καλύτερη κατανόηση των συνθηκών που συνδέονται με σοβαρότερες συνέπειες. Τα ευρήματα της εργασίας μπορούν να αποτελέσουν χρήσιμη βάση για τη διαμόρφωση στοχευμένων παρεμβάσεων και μέτρων με σκοπό τη μείωση των σοβαρών τραυματισμών και των θανάτων και τη βελτίωση της οδικής ασφάλειας των μηχανοκίνητων δικύκλων.

1.3 Μεθοδολογία

Για την εκπόνηση της παρούσας Εργασίας και την επίτευξη του τελικού της στόχου ακολουθήθηκε η μεθοδολογία, τα στάδια της οποίας παρουσιάζονται στη συνέχεια.

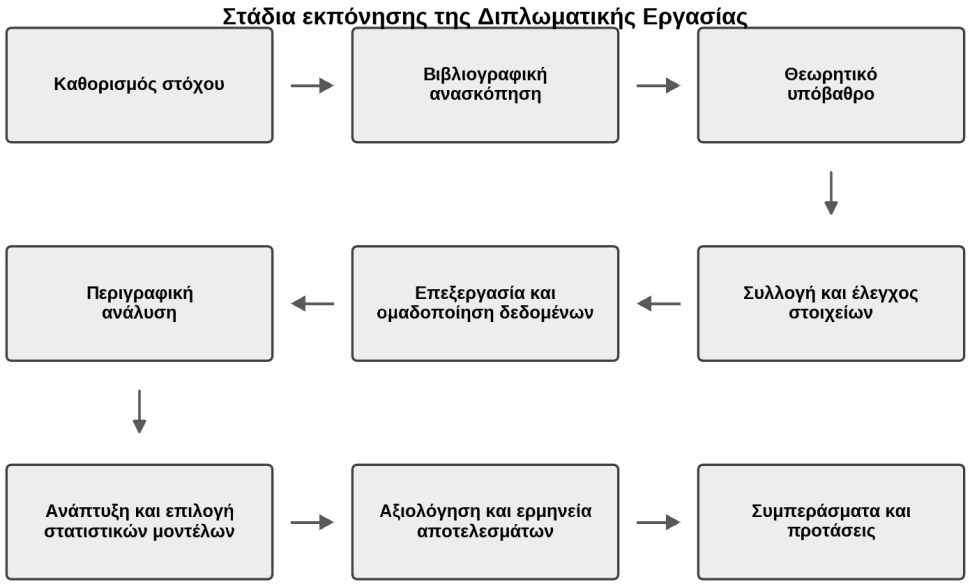
Αρχικά, οριστικοποιήθηκε το **θέμα της εργασίας** καθώς και ο επιδιωκόμενος στόχος της. Έπειτα, πραγματοποιήθηκε **εκτενής βιβλιογραφική ανασκόπηση**, τόσο σε ελληνική όσο και σε διεθνή βιβλιογραφία, μέσα από μελέτη αντίστοιχων ερευνών και επιστημονικών κειμένων, που θα μπορούσαν να φανούν χρήσιμα για τη συγκεκριμένη έρευνα. Μέσα από τη διαδικασία αυτή αποκτήθηκε εξοικείωση με την επεξεργασία αντίστοιχων θεμάτων και επιλέχθηκε η μέθοδος βάσει της οποίας θα πραγματοποιούνταν στη συνέχεια η επεξεργασία των στοιχείων.

Επόμενο βήμα αποτέλεσε η **συλλογή των στοιχείων από** τις βάσεις δεδομένων του **Παρατηρητηρίου Οδικής Ασφάλειας του ΕΜΠ** (NTUA Road Safety Observatory, NRSO) **και της Ελληνικής Στατιστικής Αρχής** (ΕΛ.ΣΤΑΤ.), που αφορούν τα καταγεγραμμένα οδικά ατυχήματα με μηχανοκίνητα δίκυκλα στην Ελλάδα, καθώς και ο καθορισμός των μεταβλητών που χρησιμοποιήθηκαν στη στατιστική ανάλυση.

Ύστερα, **επιλέχθηκαν οι μέθοδοι στατιστικής επεξεργασίας των στοιχείων**, η οποία πραγματοποιήθηκε στο περιβάλλον της γλώσσας προγραμματισμού R. Συγκεκριμένα, επιλέχθηκαν **Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων** (Generalized Linear Mixed Models, GLMM), τα οποία επιτρέπουν τη μοντελοποίηση του αριθμού παθόντων ως μεταβλητής καταμέτρησης. Για τους ελαφρά τραυματίες εφαρμόστηκε μοντέλο Αρνητικής Διωνυμικής (Negative Binomial) παλινδρόμησης, ενώ για τους **νεκρούς και βαριά τραυματίες** εφαρμόστηκε μοντέλο **παλινδρόμησης Poisson**.

Τελικό στάδιο αποτέλεσε η **αξιολόγηση και ερμηνεία** των αποτελεσμάτων με βάση τους **απαραίτητους στατιστικούς ελέγχους**, ώστε να προσδιοριστούν οι παράγοντες

που σχετίζονται με τον αναμενόμενο αριθμό παθόντων ανά κατηγορία σοβαρότητας. Στη συνέχεια διατυπώνονται προτάσεις για την αξιοποίηση των αποτελεσμάτων και για περαιτέρω έρευνα. Τα βασικά στάδια της μεθοδολογίας παρουσιάζονται στο ακόλουθο διάγραμμα ροής.



Διάγραμμα 1.2: Διάγραμμα ροής σταδίων εκπόνησης της Διπλωματικής Εργασίας

1.4 Δομή Διπλωματικής Εργασίας

Στην ενότητα αυτή παρουσιάζεται συνοπτικά η δομή της Διπλωματικής Εργασίας, ώστε να διευκολύνεται η παρακολούθηση της διαδικασίας που ακολουθείται από τη διατύπωση του προβλήματος έως την εξαγωγή των τελικών συμπερασμάτων.

Το **πρώτο κεφάλαιο αποτελεί την εισαγωγή** της Διπλωματικής Εργασίας. Παρουσιάζεται το γενικότερο ζήτημα της οδικής ασφάλειας, με ιδιαίτερη έμφαση στα μηχανοκίνητα δίκυκλα, παρατίθενται βασικά συγκριτικά στατιστικά στοιχεία για ατυχήματα, νεκρούς και τραυματίες και διατυπώνονται ο στόχος, η μεθοδολογία και η δομή της εργασίας.

Στο **δεύτερο κεφάλαιο παρουσιάζεται η βιβλιογραφική ανασκόπηση** ελληνικών και διεθνών ερευνών σχετικών με την οδική ασφάλεια των μηχανοκίνητων δικύκλων και τη σοβαρότητα των ατυχημάτων. Εξετάζονται οι βασικοί παράγοντες και συνοψίζονται τα κύρια ευρήματα προηγούμενων συναφών αναλύσεων.

Το **τρίτο κεφάλαιο περιλαμβάνει το θεωρητικό υπόβαθρο** στο οποίο βασίζεται η στατιστική ανάλυση των δεδομένων. Πιο συγκεκριμένα, ορίζονται οι βασικές στατιστικές έννοιες και περιγράφονται τα μοντέλα παλινδρόμησης για δεδομένα

καταμέτρησης, καθώς και οι στατιστικοί έλεγχοι και τα κριτήρια που χρησιμοποιούνται για την αξιολόγηση και την ερμηνεία τους.

Στο **τέταρτο κεφάλαιο περιγράφεται η διαδικασία συλλογής και επεξεργασίας των δεδομένων** που χρησιμοποιήθηκαν για την επίτευξη του στόχου της παρούσας εργασίας. Παρουσιάζονται η πηγή των δεδομένων, οι μεταβλητές και η κωδικοποίησή τους για την ανάλυση, καθώς και η περιγραφική ανάλυση του τελικού συνόλου δεδομένων.

Στο **πέμπτο κεφάλαιο εφαρμόζονται οι επιλεγμένες στατιστικές μέθοδοι για την κατασκευή των τελικών μοντέλων**, αφού προηγηθεί η επιλογή της κατάλληλης οικογένειας κατανομής, ο έλεγχος της τυχαίας επίδρασης της γεωγραφικής περιοχής και η επιλογή των βασικών επεξηγηματικών μεταβλητών. Παρουσιάζονται οι έλεγχοι στατιστικής σημαντικότητας, τα κριτήρια προσαρμογής, οι διαγνωστικοί έλεγχοι που προβλέπονται από τη μεθοδολογία και τα τελικά αποτελέσματα των μοντέλων.

Στο **έκτο κεφάλαιο αναλύονται τα κυριότερα συμπεράσματα**, τα οποία προέκυψαν από τα προηγούμενα κεφάλαια και πιο συγκεκριμένα από την περιγραφική ανάλυση των στοιχείων και την εξαγωγή των τελικών μοντέλων. Ακόμη, στο ίδιο κεφάλαιο επισημαίνονται προτάσεις σχετικά με την αξιοποίηση των αποτελεσμάτων της Διπλωματικής Εργασίας, με στόχο τη βελτίωση της οδικής ασφάλειας. Δίνονται επίσης προτάσεις για περαιτέρω έρευνες που σχετίζονται με το συγκεκριμένο θέμα.

Τέλος, στο **έβδομο κεφάλαιο παρατίθενται οι βιβλιογραφικές αναφορές** σε αλφαβητική σειρά, οι οποίες χρησιμοποιήθηκαν για την εκπόνηση της Διπλωματικής Εργασίας. Ειδικότερα, πρόκειται για έναν κατάλογο, ο οποίος περιλαμβάνει έρευνες σχετικές με το θέμα της παρούσας εργασίας, καθώς και οποιεσδήποτε άλλες πηγές λήφθηκαν υπόψη κατά τη διαδικασία σύνταξης.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 Εισαγωγή

Το παρόν κεφάλαιο αποτελεί τη **βιβλιογραφική ανασκόπηση**, η οποία πραγματοποιήθηκε στο πλαίσιο της εκπόνησης της παρούσας Διπλωματικής Εργασίας. Περιλαμβάνονται **δημοσιευμένες έρευνες της ελληνικής και διεθνούς βιβλιογραφίας**, με έμφαση στη μεθοδολογία και στο αντικείμενό τους. Συγκεκριμένα, αναζητήθηκαν έρευνες που επικεντρώνονται στη διερεύνηση της σοβαρότητας των οδικών ατυχημάτων με εμπλοκή μηχανοκίνητων δικύκλων. Μέσω της παρουσίασης των ερευνών τεκμηριώνονται ο στόχος της παρούσας εργασίας και η κατάλληλη μεθοδολογία για την επίτευξή του, ενώ εντοπίζονται παράμετροι που δεν έχουν μελετηθεί επαρκώς και μπορούν να αποτελέσουν αντικείμενο έρευνας.

2.2 Οδική ασφάλεια δικύκλων

Τα μηχανοκίνητα δίκυκλα, στα οποία περιλαμβάνονται κυρίως οι μοτοσυκλέτες και τα μοτοποδήλατα, αποτελούν σημαντικό μέσο μετακίνησης σε πολλές ευρωπαϊκές χώρες, ενώ ταυτόχρονα αποτελούν **μία από τις πλέον ευάλωτες κατηγορίες χρηστών** του οδικού δικτύου. Η έλλειψη δομικής προστασίας του αναβάτη, ο μικρός όγκος και η χαμηλή οπτική τους παρουσία σε σχέση με τα υπόλοιπα οχήματα, καθώς και τα ιδιαίτερα δυναμικά χαρακτηριστικά τους (ευστάθεια, επιταχύνσεις, ελιγμοί) τα καθιστούν ιδιαίτερα επιρρεπή σε ατυχήματα με σοβαρές συνέπειες (Vlahogianni et al., 2012). Παρά το γεγονός ότι τα μηχανοκίνητα δίκυκλα **αντιπροσωπεύουν σχετικά μικρό ποσοστό του συνολικού κυκλοφοριακού φόρτου**, **ο αριθμός των θανάτων και των σοβαρών τραυματισμών** που σχετίζονται με αυτά **είναι δυσανάλογα υψηλός** τόσο σε ευρωπαϊκό όσο και σε παγκόσμιο επίπεδο (Theofilatos & Yannis, 2015).

Σύμφωνα με την Παγκόσμια Έκθεση για την Οδική Ασφάλεια του 2023 του Παγκόσμιου Οργανισμού Υγείας, οι **αναβάτες μηχανοκίνητων δικύκλων** και τρικύκλων **αντιστοιχούν** στο **21% των θανάτων από οδικά ατυχήματα παγκοσμίως**, ενώ το σύνολο των ευάλωτων χρηστών της οδού αντιστοιχεί στο 53% (World Health Organization, 2023). Παράλληλα, η ανασκόπηση των Damani και Vedagiri (2021) επισημαίνει ότι οι σχετικές μελέτες εξετάζουν διαφορετικά σύνολα παραγόντων κινδύνου και καταλήγουν συχνά σε διαφορετικά ή και αντικρουόμενα αποτελέσματα, γεγονός που δυσκολεύει τη γενίκευση των ευρημάτων από τη μία χώρα στην άλλη.

Η διεθνής βιβλιογραφία συγκλίνει συνεπώς στην άποψη ότι **τα ατυχήματα με μηχανοκίνητα δίκυκλα δεν πρέπει να αντιμετωπίζονται ως ένα ομοιογενές σύνολο περιστατικών**. Διαφορετικά χαρακτηριστικά του αναβάτη, του οχήματος, του τύπου σύγκρουσης και του οδικού περιβάλλοντος μπορούν να δημιουργούν σημαντικά διαφορετικές συνθήκες κινδύνου. Για τον λόγο αυτό, η αναζήτηση των παραγόντων που σχετίζονται με την εμφάνιση και κυρίως με τη σοβαρότητα των ατυχημάτων αποτελεί κεντρικό αντικείμενο της σχετικής επιστημονικής έρευνας.

2.3 Παράγοντες που επηρεάζουν τη σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα

Η πλειονότητα των ερευνητικών εργασιών που ασχολούνται με την ασφάλεια των δικύκλων υιοθετεί μια **πολυπαραγοντική προσέγγιση**, αναγνωρίζοντας ότι η πρόκληση και η σοβαρότητα ενός ατυχήματος είναι αποτέλεσμα σύνθετης αλληλεπίδρασης ανθρώπου, οχήματος και περιβαλλοντικών παραγόντων. Μία συστηματική ανασκόπηση, η οποία συνέθεσε τα αποτελέσματα 32 μελετών δημοσιευμένων σε περιοδικά με σύστημα κρίσης (peer-reviewed), **ταξινόμησε τους παράγοντες σοβαρότητας σε έξι κατηγορίες**: οδικούς, παράγοντες που σχετίζονται με τη σύγκρουση, οχηματικούς, χρονικούς, ανθρώπινους και περιβαλλοντικούς (Zhang et al., 2026). Η πολυδιάστατη αυτή προσέγγιση είναι σημαντική, διότι ένας παράγοντας δεν έχει απαραίτητα την ίδια επίδραση ανεξάρτητα από τις υπόλοιπες συνθήκες.

2.3.1 Χαρακτηριστικά και συμπεριφορά του αναβάτη

Η ηλικία και η εμπειρία του αναβάτη αποτελούν από τους πλέον μελετημένους και σταθερά σημαντικούς παράγοντες κινδύνου στη διεθνή βιβλιογραφία. Σε μελέτη περιπτώσεων-μαρτύρων που βασίστηκε στα δεδομένα της Μελέτης Αιτιών Ατυχημάτων Μοτοσυκλετών (Motorcycle Crash Causation Study, MCCS) στις ΗΠΑ, **ο κίνδυνος εμπλοκής σε ατύχημα** βρέθηκε αυξημένος στους νεότερους αναβάτες και μειωνόταν με την αύξηση της ηλικίας. Κάθε επιπλέον έτος εμπειρίας οδήγησης συνδέθηκε με μείωση των σχετικών πιθανοτήτων (odds) εμπλοκής σε ατύχημα **κατά περίπου 2,5%,** ενώ η πρόσφατη συμμετοχή σε προγράμματα εκπαίδευσης αναβατών, καθώς και **η αποχή από την κατανάλωση αλκοόλ,** συνδέθηκαν σημαντικά με μειωμένο κίνδυνο (Adeel et al., 2025).

Αντίστοιχα ευρήματα αναφέρονται και σε μελέτες σοβαρότητας τραυματισμού: στην Κίνα, **αναβάτες ηλικίας άνω των 60 ετών** παρουσιάζουν σημαντικά αυξημένη πιθανότητα **θανατηφόρου ή σοβαρού τραυματισμού** (Chang et al., 2016). Παρόμοια, τόσο οι νεαροί όσο και οι ηλικιωμένοι αναβάτες, καθώς και οι άπειροι

οδηγοί, βρίσκονται σε αυξημένο κίνδυνο σοβαρού τραυματισμού (Zhang et al., 2026). Η αυξημένη τρωτότητα των νεότερων αναβατών αποδίδεται συχνά σε ριψοκίνδυνη συμπεριφορά και έλλειψη εμπειρίας στη διαχείριση κρίσιμων καταστάσεων, ενώ η αυξημένη τρωτότητα των μεγαλύτερων σε ηλικία αναβατών αποδίδεται στη μειωμένη φυσική ανθεκτικότητα του οργανισμού στην πρόσκρουση (Theofilatos & Yannis, 2015). Η συμπεριφορά του αναβάτη κατά την οδήγηση αναγνωρίζεται ως κρίσιμος παράγοντας που επηρεάζει τόσο την **πιθανότητα εμπλοκής σε ατύχημα όσο και τη σοβαρότητά** του. Χαρακτηριστικά όπως η υπερβολική ταχύτητα, οι παραβάσεις κυκλοφοριακών κανόνων, οι επικίνδυνες προσπεράσεις, οι αλλαγές λωρίδας και η χρήση κινητού τηλεφώνου κατά την οδήγηση σχετίζονται στενά με δημογραφικά χαρακτηριστικά όπως η ηλικία, το φύλο και το επίπεδο εκπαίδευσης του αναβάτη (Chouhan et al., 2024). Χαρακτηριστικά προσωπικότητας όπως η αναζήτηση συγκινήσεων, η επιθετικότητα, το αντιληπτό επίπεδο κινδύνου και οι στάσεις απέναντι στην οδική ασφάλεια **αποτελούν σταθερούς προγνωστικούς παράγοντες ριψοκίνδυνης συμπεριφοράς** (Theofilatos & Yannis, 2015). Επιπλέον, προηγούμενη εμπλοκή σε ατύχημα φαίνεται να σχετίζεται με επιθετική συμπεριφορά και τάση για αναζήτηση συγκινήσεων, υποδεικνύοντας έναν φαύλο κύκλο επανάληψης του κινδύνου (Theofilatos & Yannis, 2015).

Σε μια μελέτη που πραγματοποιήθηκε στην Πορτογαλία διαπιστώθηκε ότι **το αρσενικό φύλο του αναβάτη, η οδήγηση κατά τις νυχτερινές ώρες (20:00-05:59) και οι μη εργάσιμες ημέρες** σχετίζονται με **αυξημένη πιθανότητα** σοβαρού τραυματισμού, στοιχεία που συνάδουν με την υπόθεση ότι η ψυχαγωγική χρήση του δίκυκλου συνοδεύεται συχνά από μεγαλύτερη ανάληψη κινδύνου (Santos et al., 2021).

Η οδήγηση υπό την επήρεια αλκοόλ αποτελεί έναν από τους π**λέον συνεπείς παράγοντες κινδύνου** που αναφέρονται στη διεθνή βιβλιογραφία για τα ατυχήματα με μηχανοκίνητα δίκυκλα. Οι νηφάλιοι αναβάτες παρουσίαζαν μείωση της πιθανότητας εμπλοκής σε ατύχημα κατά περίπου 75% σε σχέση με όσους είχαν καταναλώσει αλκοόλ, ακόμη και μετά τον έλεγχο για την ηλικία, την εμπειρία και την εκπαίδευση του αναβάτη (Adeel et al., 2025). Στην Πορτογαλία, συγκεκριμένα επίπεδα συγκέντρωσης αλκοόλ στο αίμα (0,5-0,8 g/L) σχετίζονταν με αυξημένη πιθανότητα σοβαρού τραυματισμού (Santos et al., 2021). Η οδήγηση υπό την επήρεια αλκοόλ κατατάσσεται μεταξύ των παραγόντων για τους οποίους υπάρχει ευρεία επιστημονική συναίνεση ως προς την αυξητική τους επίδραση στη σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα (Zhang et al., 2026).

2.3.2 Χαρακτηριστικά οχήματος και προστατευτικός εξοπλισμός

Η χρήση κράνους αποτελεί το πλέον τεκμηριωμένο **μέτρο παθητικής προστασίας** των αναβατών μηχανοκίνητων δικύκλων. Μια κλασική συστηματική ανασκόπηση της Cochrane, η οποία συνέθεσε ευρήματα 61 παρατηρησιακών μελετών, κατέληξε ότι η χρήση κράνους **μειώνει τον κίνδυνο θανάτου κατά** 42% και τον **κίνδυνο κρανιοεγκεφαλικής κάκωσης κατά** 69% στους αναβάτες που εμπλέκονται σε ατύχημα (Liu et al., 2008). Μεταγενέστερη συστηματική ανασκόπηση και μετα-ανάλυση οκτώ μελετών από χώρες χαμηλού και χαμηλότερου μεσαίου εισοδήματος της Αφρικής επιβεβαίωσε τα ευρήματα αυτά, εντοπίζοντας μείωση της πιθανότητας σοβαρής κρανιοεγκεφαλικής κάκωσης κατά 88% στους χρήστες κράνους σε σχέση με τους μη χρήστες (Abdi et al., 2022).

Πέραν της απλής χρήσης, η ορθή και σταθερή στερέωση του κράνους, καθώς και ο τύπος του (πλήρους προσώπου έναντι ανοιχτού τύπου), φαίνεται να επηρεάζουν σημαντικά τον βαθμό προστασίας που παρέχεται, ιδίως όσον αφορά τραυματισμούς στο πρόσωπο. Στην Πορτογαλία, η **μη χρήση κράνους** αναδείχθηκε ως **ένας από τους σημαντικότερους παράγοντες αύξησης της σοβαρότητας τραυματισμού** σε ατυχήματα με μηχανοκίνητα δίκυκλα (Santos et al., 2021), ενώ στην Κίνα η απουσία κράνους σχετίστηκε σημαντικά με αυξημένη πιθανότητα θανατηφόρου ή σοβαρού τραυματισμού (Chang et al., 2016).

2.3.3 Τύπος ατυχήματος και αριθμός εμπλεκόμενων οχημάτων

Ο τύπος ατυχήματος αποτελεί έναν από τους σημαντικότερους παράγοντες της σοβαρότητας. Ο τύπος σύγκρουσης επηρεάζει σημαντικά την πιθανότητα των διαφορετικών επιπέδων τραυματισμού, με **μεγαλύτερη σοβαρότητα** να συνδέεται, μεταξύ άλλων, **με μετωπικές και πλαγιομετωπικές συγκρούσεις**, **συγκρούσεις με σταθερά αντικείμενα, σκοτάδι, κατανάλωση αλκοόλ και μη ασφαλή ταχύτητα** (Savolainen & Mannering, 2007). Η ευρωπαϊκή μελέτη SaferWheels, η οποία εξέτασε σε βάθος 500 ατυχήματα με μηχανοκίνητα δίκυκλα και ποδήλατα σε έξι χώρες, έδειξε ότι συχνό σενάριο ήταν η στροφή ή η διάσχιση του άλλου οχήματος μπροστά από το δίκυκλο, ενώ περίπου το ένα τέταρτο των σοβαρών ή θανατηφόρων τραυματισμών αναβατών μηχανοκίνητων δικύκλων προέκυψε από απώλεια ελέγχου χωρίς εμπλοκή άλλου οχήματος (Brown et al., 2021). Παρόμοια, οι μετωπικές συγκρούσεις και οι εκτροπές από την οδό εντοπίζονται ως κατηγορίες με σχετικά σταθερή συσχέτιση με αυξημένη σοβαρότητα (Zhang et al., 2026).

Σημαντικό εύρημα της διεθνούς βιβλιογραφίας αφορά τον ρόλο των οδηγών των υπόλοιπων οχημάτων, και ιδίως των επιβατικών αυτοκινήτων. Η κλασική μελέτη των

Hurt et al. (1981) έδειξε ότι περίπου τ**α τρία τέταρτα των ατυχημάτων με μοτοσυκλέτες αφορούσαν σύγκρουση με άλλο όχημα, συνήθως επιβατικό,** και ότι στα δύο τρίτα αυτών των συγκρούσεων ο οδηγός του άλλου οχήματος παραβίασε την προτεραιότητα του αναβάτη. Ως κυριότερη αιτία αναγνωρίστηκε το γεγονός ότι **οι οδηγοί δεν εντοπίζουν έγκαιρα το δίκυκλο, ιδιαίτερα σε διασταυρώσεις.** Τα σφάλματα αυτά, γνωστά με τον όρο «κοίταξε αλλά δεν είδε» (Looked But Failed To See, LBFTS), παραμένουν από τις συχνότερες αιτίες ατυχημάτων με μηχανοκίνητα δίκυκλα και στην Ευρώπη (Brown et al., 2021), ενώ νεότερη έρευνα συνδέει μέρος τους με αστοχίες της βραχυπρόθεσμης μνήμης των οδηγών σε κόμβους (Robbins et al., 2019).

Τα δομικά και λειτουργικά χαρακτηριστικά του ίδιου του δίκυκλου επηρεάζουν τόσο την **πιθανότητα εμπλοκής σε ατύχημα όσο και τη σοβαρότητα των συνεπειών του**. Λόγω του μικρού μεγέθους τους και του συχνά ισχυρού κινητήρα, τα μηχανοκίνητα δίκυκλα χαρακτηρίζονται από αυξημένη ευκολία απώλειας ευστάθειας σε χαμηλές ταχύτητες, απώλεια πρόσφυσης σε κακή κατάσταση οδοστρώματος, καθώς και υψηλές δυνατότητες επιτάχυνσης που σχετίζονται με την υπερβολική ταχύτητα και τη δυσκολία στην πέδηση (Vlahogianni et al., 2012). Ο αναβάτης καλείται να συντονίζει ταυτόχρονα την ταχύτητα, τη γωνία κλίσης του σώματος, την πρόσφυση και τον έλεγχο του οχήματος σε ποικίλες επιφάνειες, καμπύλες και συνθήκες κυκλοφορίας, γεγονός που αυξάνει τη γνωστική και κινητική επιβάρυνση σε σύγκριση με την οδήγηση επιβατικού οχήματος. Ο κυβισμός και η κατηγορία του οχήματος αποτελούν επίσης σημαντικούς παράγοντες. Παλαιότερες μελέτες έχουν δείξει ότι τα ποσοστά θνησιμότητας αναβατών σε δίκυκλα σπορ υψηλών επιδόσεων (super sport) ήταν έως τέσσερις φορές υψηλότερα σε σχέση με άλλες κατηγορίες δικύκλων, γεγονός που αποδίδεται στα αυξημένα χαρακτηριστικά επιτάχυνσης και μέγιστης ταχύτητας των οχημάτων αυτής της κατηγορίας (Theofilatos & Yannis, 2015). Σε ελληνικά δεδομένα, **η σοβαρότητα των ατυχημάτων νεαρών αναβατών** βρέθηκε να εξαρτάται από τον **συνδυασμό της ηλικίας του αναβάτη και του κυβισμού του δικύκλου,** ενώ ο κυβισμός δεν είχε σημαντική επίδραση στον κίνδυνο υπαιτιότητας (Yannis et al., 2005).

2.3.4 Οδικό περιβάλλον και υποδομή

Η σχέση μεταξύ οδικής υποδομής και ασφάλειας των μηχανοκίνητων δικύκλων είναι ιδιαίτερα σύνθετη. Η γεωμετρία της οδού, οι καμπύλες, οι κλίσεις, το πλάτος των λωρίδων, οι διασταυρώσεις, η **κατάσταση του οδοστρώματος**, η ύπαρξη πλευρικών εμποδίων και το επίπεδο φωτισμού μπορούν να επηρεάσουν τόσο την **πιθανότητα απώλειας ελέγχου** όσο και τις συνέπειες της σύγκρουσης. Στην Κίνα,

ατυχήματα σε επαρχιακές οδούς με καμπύλη και κλίση χάραξης σχετίζονταν με **αυξημένο κίνδυνο θανατηφόρου έκβασης για τον αναβάτη**, ενώ οι μη σηματοδοτούμενες διασταυρώσεις σχετίζονταν με λιγότερο σοβαρούς τραυματισμούς (Chang et al., 2016).

Σε αστικό περιβάλλον, η κατάσταση του οδοστρώματος (υγρή επιφάνεια, ανομοιομορφίες), η ύπαρξη λωρίδων μονής κατεύθυνσης και οι κεντρικές λωρίδες οδών με πολλαπλές λωρίδες έχουν συσχετιστεί με **ατυχήματα στα οποία ο αναβάτης μηχανοκίνητου δικύκλου δεν φέρει ευθύνη,** κυρίως λόγω απρόβλεπτων μεταβολών στη διαθέσιμη πρόσφυση ή στον διαθέσιμο χώρο ελιγμού (Theofilatos & Yannis, 2015).

2.3.5 Περιβαλλοντικοί και χρονικοί παράγοντες

Οι καιρικές συνθήκες, η ώρα της ημέρας, η εποχή και ο φωτισμός αποτελούν παράγοντες με μεικτά και εν μέρει **αντικρουόμενα ευρήματα** στη διεθνή βιβλιογραφία. **Η νυχτερινή οδήγηση, τα Σαββατοκύριακα/αργίες και η θερινή περίοδος** κατατάσσονται μεταξύ των παραγόντων για τους οποίους υπάρχει ευρεία συναίνεση ως προς την αυξητική τους επίδραση στη σοβαρότητα ατυχήματος, ενώ αντίθετα οι καιρικές συνθήκες αναφέρονται ως **αμφιλεγόμενος παράγοντας**, καθώς οι επιδράσεις τους φαίνεται να διαφέρουν σημαντικά ανάλογα με το γεωγραφικό και οδικό πλαίσιο της κάθε μελέτης (Zhang et al., 2026). Χαρακτηριστικά, στην Πορτογαλία το ξηρό και καθαρό οδόστρωμα σχετιζόταν με αυξημένη πιθανότητα σοβαρού τραυματισμού, εύρημα το οποίο αποδίδεται συχνά στην ανάπτυξη υψηλότερων ταχυτήτων υπό καλές καιρικές συνθήκες, σε αντίθεση με τη διαισθητική υπόθεση ότι η βροχή ή το υγρό οδόστρωμα θα αύξανε τη σοβαρότητα (Santos et al., 2021). Σε έρευνα στην Αττική Οδό, οι μεταβολές της κυκλοφορίας και της ταχύτητας συνδέθηκαν με σοβαρότερους τραυματισμούς επιβαινόντων σε μηχανοκίνητα δίκυκλα, ενώ οι καιρικές παράμετροι δεν είχαν σημαντική επίδραση (Theofilatos & Ziakopoulos, 2018).

2.4 Σύνοψη

Από τη συνολική επισκόπηση της βιβλιογραφίας προκύπτει ότι **τα ατυχήματα με μηχανοκίνητα δίκυκλα** αποτελούν πολυπαραγοντικό φαινόμενο, στο οποίο συνυπάρχουν παράγοντες που **σχετίζονται με τον άνθρωπο, το όχημα, την υποδομή και το περιβάλλον.** Επιπλέον, παράγοντες όπως η νεαρή ή προχωρημένη ηλικία, η έλλειψη εμπειρίας, η οδήγηση υπό την επήρεια αλκοόλ, η εμπλοκή βαρέων οχημάτων και η μη χρήση κράνους εμφανίζουν σταθερή και συνεπή συσχέτιση με αυξημένη σοβαρότητα ατυχήματος σε πολλαπλά γεωγραφικά και πολιτισμικά πλαίσια.

Αντίθετα, **παράγοντες όπως οι καιρικές συνθήκες, η γεωμετρία της οδού και το φύλο του αναβάτη** εμφανίζουν αντικρουόμενα ευρήματα **ανάλογα με τη γεωγραφική περιοχή,** το είδος του δικτύου και τη μεθοδολογία της κάθε μελέτης, υποδεικνύοντας την ανάγκη προσεκτικής ερμηνείας κατά τη μεταφορά συμπερασμάτων από τη μία χώρα στην άλλη.

Ιδιαίτερη σημασία έχει το φαινόμενο LBFTS, το οποίο δείχνει ότι σε σημαντικό μέρος των ατυχημάτων το κρίσιμο σφάλμα αφορά τον οδηγό του άλλου εμπλεκόμενου οχήματος και σχετίζεται με γνωστικούς παράγοντες αντίληψης και μνήμης (Brown et al., 2021; Robbins et al., 2019). Το εύρημα αυτό υποστηρίζει, πέρα από τα παραδοσιακά μέτρα αύξησης της οπτικής αντιληπτότητας του δικύκλου (φωτεινός ρουχισμός, φώτα ημέρας), και την **ανάπτυξη συστημάτων υποβοήθησης οδηγού (Advanced Driver Assistance Systems, ADAS)** που ανιχνεύουν δίκυκλα και προειδοποιούν έγκαιρα τον οδηγό του αυτοκινήτου.

3. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

3.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται το **θεωρητικό υπόβαθρο των στατιστικών μεθόδων** που χρησιμοποιούνται για την ανάλυση δεδομένων οδικών ατυχημάτων. Αρχικά, ορίζονται οι βασικές έννοιες της στατιστικής, όπως ο πληθυσμός, το δείγμα, οι τύποι μεταβλητών και τα βασικά περιγραφικά μέτρα. Στη συνέχεια, περιγράφονται τα **μοντέλα παλινδρόμησης** για δεδομένα καταμέτρησης, δηλαδή η παλινδρόμηση **Poisson** και η **Αρνητική Διωνυμική παλινδρόμηση**, καθώς και τα μοντέλα μικτών επιδράσεων. Τέλος, παρουσιάζονται οι στατιστικοί έλεγχοι, τα κριτήρια σύγκρισης και οι **διαγνωστικοί έλεγχοι** που χρησιμοποιούνται για την αξιολόγηση και την ερμηνεία των μοντέλων.

3.2 Βασικές έννοιες στατιστικής

3.2.1 Πληθυσμός και δείγμα

Ως **πληθυσμός** (population) νοείται το σύνολο των μονάδων ή παρατηρήσεων στις οποίες αναφέρεται μία στατιστική έρευνα. Οι μονάδες αυτές μπορεί να είναι άτομα, οχήματα, ατυχήματα ή οποιοδήποτε άλλο σαφώς ορισμένο σύνολο γεγονότων. Επειδή στις περισσότερες εφαρμογές δεν είναι εφικτή η εξέταση ολόκληρου του πληθυσμού, η ανάλυση βασίζεται συνήθως σε ένα υποσύνολό του, το **δείγμα** (sample) (Agresti & Finlay, 2009). Όλα τα στοιχεία ενός δείγματος ανήκουν και στον πληθυσμό, χωρίς όμως να ισχύει αναγκαστικά και το αντίστροφο. Η δυνατότητα γενίκευσης των συμπερασμάτων εξαρτάται τόσο από την αντιπροσωπευτικότητα του δείγματος όσο και από την ποιότητα των καταγραφών. Στις βάσεις οδικών ατυχημάτων χρειάζεται επίσης να ελέγχονται η πληρότητα των πεδίων, η συνέπεια της κωδικοποίησης και η ύπαρξη ειδικών κατηγοριών, όπως οι άγνωστες τιμές.

3.2.2 Μεταβλητές

Μεταβλητές (variables) ορίζονται τα χαρακτηριστικά που καταγράφονται για κάθε παρατήρηση.

Οι **ποσοτικές μεταβλητές** (quantitative variables) είναι οι μεταβλητές των οποίων οι τιμές είναι αριθμοί με την έννοια της μέτρησης και διακρίνονται σε διακριτές και συνεχείς. Διακριτή θεωρείται μια μεταβλητή όταν η μικρότερη μη μηδενική διαφορά που μπορεί να έχουν δύο τιμές της είναι σταθερή ποσότητα. Συνεχής θεωρείται μια

μεταβλητή όταν μπορεί να λάβει οποιαδήποτε τιμή σε ένα διάστημα τιμών, δηλαδή όταν δύο τιμές της μπορούν να διαφέρουν κατά οποιαδήποτε μικρή ποσότητα.

Οι **ποιοτικές μεταβλητές** (qualitative variables) είναι οι μεταβλητές των οποίων οι δυνατές τιμές είναι κατηγορίες διαφορετικές μεταξύ τους. Η χρήση των αριθμών για την αναπαράσταση των τιμών μιας τέτοιας μεταβλητής είναι αποκλειστικά συμβολική και δεν έχει την έννοια της μέτρησης.

Ως **εξαρτημένη μεταβλητή** (dependent variable) ορίζεται η έκβαση που επιχειρείται να εξηγηθεί ή να προβλεφθεί. Οι **ανεξάρτητες μεταβλητές** (independent variables ή predictors) είναι τα χαρακτηριστικά που χρησιμοποιούνται για την εξήγηση ή την πρόβλεψη της εξαρτημένης μεταβλητής.

Ειδική κατηγορία διακριτής ποσοτικής μεταβλητής αποτελεί η **μεταβλητή καταμέτρησης** (count variable), η οποία λαμβάνει μη αρνητικές ακέραιες τιμές (0, 1, 2, …) και εκφράζει το πλήθος εμφανίσεων ενός γεγονότος σε μία μονάδα παρατήρησης ή σε ένα συγκεκριμένο χρονικό ή χωρικό διάστημα. Χαρακτηριστικά παραδείγματα στον τομέα της οδικής ασφάλειας αποτελούν ο αριθμός ατυχημάτων σε ένα οδικό τμήμα και ο αριθμός παθόντων (νεκρών ή τραυματιών) ανά ατύχημα.

Όταν η **εξαρτημένη μεταβλητή** είναι **μεταβλητή καταμέτρησης**, η φύση της είναι καθοριστική για την επιλογή του κατάλληλου στατιστικού μοντέλου, καθώς η κλασική γραμμική παλινδρόμηση δεν εξασφαλίζει μη αρνητικές προβλέψεις και δεν αποδίδει ορθά τη σχέση μεταξύ μέσης τιμής και διακύμανσης που χαρακτηρίζει τα δεδομένα αυτά.

Μέση τιμή (mean) ορίζεται το πηλίκο του αθροίσματος των τιμών προς το πλήθος τους, σύμφωνα με την παρακάτω σχέση:

$$\bar{x} = \frac{(x1 + x2 + ... + x\nu)}{\nu} = \frac{1}{\nu}\sum_{i=1}^{\nu} xi$$

Διάμεσος (median) είναι η τιμή που χωρίζει ένα σύνολο δεδομένων σε δύο ίσα μέρη, έτσι ώστε το 50% των παρατηρήσεων να βρίσκεται κάτω από αυτήν και το υπόλοιπο 50% πάνω, όταν τα δεδομένα είναι ταξινομημένα σε αύξουσα ή φθίνουσα σειρά.

Διακύμανση (variance) ορίζεται το μέτρο διασποράς που εκφράζει τον μέσο όρο των τετραγώνων των αποκλίσεων των τιμών από τη μέση τιμή. Για δείγμα ν παρατηρήσεων, η δειγματική διακύμανση υπολογίζεται με παρονομαστή ν − 1, σύμφωνα με την παρακάτω σχέση:

$$S^2 = \frac{1}{\nu\text{-}1}\sum_{i=1}^{\nu}(xi - \bar{x})^2$$

Στα δεδομένα καταμέτρησης ιδιαίτερη σημασία έχει η σχέση μεταξύ διακύμανσης και μέσης τιμής. Όταν η διακύμανση είναι αισθητά μεγαλύτερη από τη μέση τιμή υπάρχει ένδειξη υπερδιασποράς (overdispersion), ενώ όταν είναι μικρότερη υπάρχει ένδειξη υποδιασποράς (underdispersion). Ο λόγος διακύμανσης προς μέση τιμή αποτελεί χρήσιμη αρχική ένδειξη, ωστόσο η τελική επιλογή του μοντέλου βασίζεται και σε διαγνωστικούς ελέγχους της προσαρμογής.

Ο συντελεστής συσχέτισης (correlation coefficient) εκφράζει τον βαθμό και την κατεύθυνση της γραμμικής συσχέτισης μεταξύ δύο μεταβλητών και λαμβάνει τιμές μεταξύ -1 και 1. Τιμές κοντά στο 1 δηλώνουν ισχυρή θετική γραμμική συσχέτιση, τιμές κοντά στο -1 ισχυρή αρνητική γραμμική συσχέτιση, ενώ τιμές κοντά στο 0 υποδηλώνουν απουσία ισχυρής γραμμικής συσχέτισης, χωρίς αυτό να συνεπάγεται κατ' ανάγκη πλήρη στατιστική ανεξαρτησία.

3.3 Βασικές Μέθοδοι Στατιστικής Ανάλυσης

Η ανάλυση παλινδρόμησης (regression analysis) είναι η στατιστική μέθοδος με την οποία εξετάζεται η σχέση μιας εξαρτημένης μεταβλητής με μία ή περισσότερες **ανεξάρτητες μεταβλητές**, ώστε να είναι δυνατή η ερμηνεία ή η πρόβλεψη της πρώτης με βάση τις δεύτερες. Στο κλασικό πλαίσιο της παλινδρόμησης οι ανεξάρτητες μεταβλητές θεωρούνται γνωστές (καθορισμένες) τιμές, ενώ η **εξαρτημένη μεταβλητή** θεωρείται τυχαία. Όταν η εξαρτημένη μεταβλητή δεν ακολουθεί κανονική κατανομή, όπως συμβαίνει με τις μεταβλητές καταμέτρησης, χρησιμοποιούνται τα Γενικευμένα Γραμμικά Μοντέλα (Generalized Linear Models, GLM). Σε αυτά, η αναμενόμενη τιμή της εξαρτημένης μεταβλητής συνδέεται με έναν γραμμικό συνδυασμό των ανεξάρτητων μεταβλητών (γραμμικός προβλεπτικός παράγοντας) μέσω μιας συνάρτησης σύνδεσης (link function) (Agresti, 2018).

Παλινδρόμηση Poisson (Poisson Regression)

Η παλινδρόμηση Poisson αποτελεί στατιστική μέθοδο που χρησιμοποιείται όταν η εξαρτημένη μεταβλητή είναι μεταβλητή καταμέτρησης, δηλαδή λαμβάνει μη αρνητικές ακέραιες τιμές (0, 1, 2, …). Υποθέτει ότι η εξαρτημένη μεταβλητή Y ακολουθεί κατανομή Poisson, στην οποία ο μέσος όρος ισούται με τη διακύμανση (E(Y) = Var(Y) = μ). Το μοντέλο συνδέει την αναμενόμενη τιμή (μέσο όρο) της εξαρτημένης μεταβλητής με τις ανεξάρτητες μεταβλητές X μέσω μιας λογαριθμικής συνάρτησης σύνδεσης (log link):

$$log(\mu) = \beta 0 + \beta 1X1 + \beta 2X2 + ... + \beta \kappa X\kappa$$

μ: Η αναμενόμενη τιμή (μέσος όρος) της εξαρτημένης μεταβλητής.
β0: Σταθερός όρος
β1,β2,…,βκ: Συντελεστές που εκφράζουν τη βαρύτητα των ανεξάρτητων μεταβλητών στη μεταβολή του log(μ)
Χ1,Χ2,…Χκ: Ανεξάρτητες μεταβλητές

Η χρήση της λογαριθμικής συνάρτησης εξασφαλίζει ότι η εκτιμώμενη μέση τιμή παραμένει πάντα θετική, ενώ λόγω της λογαριθμικής μορφής της, οι συντελεστές β ερμηνεύονται πολλαπλασιαστικά (και όχι προσθετικά), μέσω του exp(β).

Αρνητική Διωνυμική Παλινδρόμηση (Negative Binomial Regression)

Η μέθοδος αυτή αποτελεί επέκταση της παλινδρόμησης Poisson και χρησιμοποιείται όταν τα δεδομένα καταμέτρησης παρουσιάζουν υπερδιασπορά (overdispersion), δηλαδή όταν η παρατηρούμενη διακύμανση της εξαρτημένης μεταβλητής είναι μεγαλύτερη από τη μέση της τιμή (Var(Y) > E(Y)), κάτι που παραβιάζει τη βασική υπόθεση της κατανομής Poisson. Σε αντίθεση με την Poisson, η Αρνητική Διωνυμική Παλινδρόμηση εισάγει μία επιπλέον παράμετρο διασποράς θ, η οποία επιτρέπει στη διακύμανση να ξεπερνά τη μέση τιμή (Long, 1997). Παρακάτω δίνεται η εξίσωση της μέσης τιμής (ίδια μορφή με την Poisson) και η σχέση της διακύμανσης:

$$log(\mu) = \beta 0 + \beta 1X1 + \beta 2X2 + ... + \beta \kappa X\kappa$$

$$Var(Y) = \mu + \mu^2 / \theta$$

μ: Η αναμενόμενη τιμή (μέσος όρος) της εξαρτημένης μεταβλητής
θ: Παράμετρος διασποράς (dispersion parameter). Όσο μεγαλύτερη είναι η τιμή της, τόσο περισσότερο η κατανομή προσεγγίζει την Poisson
β0, β1,…,βκ: Συντελεστές παλινδρόμησης, με την ίδια ερμηνεία όπως στην Poisson
Χ1,Χ2,…Χκ: Ανεξάρτητες μεταβλητές

Καθώς η παράμετρος θ αυξάνεται προς το άπειρο, το μοντέλο Αρνητικής Διωνυμικής Παλινδρόμησης συγκλίνει στο μοντέλο Poisson. Η επιλογή μεταξύ των δύο κατανομών γίνεται συνήθως με βάση τον έλεγχο υπερδιασποράς των καταλοίπων του μοντέλου, όπως περιγράφεται και στην ενότητα των διαγνωστικών ελέγχων.

Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων (Generalized Linear Mixed Models, GLMM)

Όταν οι παρατηρήσεις παρουσιάζουν ομαδοποίηση (για παράδειγμα, ατυχήματα που ανήκουν στην ίδια γεωγραφική περιοχή), τα μοντέλα μικτών επιδράσεων επιτρέπουν την ταυτόχρονη εκτίμηση σταθερών επιδράσεων (fixed effects) και τυχαίων επιδράσεων (random effects). Με τον τρόπο αυτό λαμβάνεται υπόψη η πιθανή

μεγαλύτερη ομοιότητα παρατηρήσεων που ανήκουν στην ίδια ομάδα σε σχέση με παρατηρήσεις διαφορετικών ομάδων. Για δεδομένα καταμέτρησης, η γενική μορφή του μοντέλου είναι η ακόλουθη (Bolker et al., 2009):

$$log(\mu ij) = \beta 0 + \beta 1X1ij + ... + \beta \kappa X\kappa ij + uj$$

όπου οι συντελεστές β εκφράζουν τις σταθερές επιδράσεις των ανεξάρτητων μεταβλητών και uj την τυχαία επίδραση που σχετίζεται με την ομάδα j, η οποία θεωρείται ότι ακολουθεί κανονική κατανομή:

$$uj \sim N(0\ ,\ \sigma u^2)$$

Τυχαίος σταθερός όρος και διακύμανση μεταξύ ομάδων

Στα μοντέλα μικτών επιδράσεων η ετερογένεια μεταξύ των ομάδων αποτυπώνεται μέσω της διακύμανσης σu² του τυχαίου σταθερού όρου (random intercept). Ο τυχαίος σταθερός όρος επιτρέπει σε κάθε ομάδα να έχει διαφορετικό σημείο εκκίνησης στην κλίμακα του γραμμικού προβλεπτικού παράγοντα, ενώ οι σταθερές επιδράσεις των επεξηγηματικών μεταβλητών εκτιμώνται από κοινού για όλες τις ομάδες. Όσο μεγαλύτερη είναι η σu², τόσο εντονότερες είναι οι διαφορές μεταξύ των ομάδων. Ένα συνηθισμένο μέτρο του ποσοστού της συνολικής μεταβλητότητας που οφείλεται στις διαφορές μεταξύ ομάδων είναι ο συντελεστής ενδοταξικής συσχέτισης (Intraclass Correlation Coefficient, ICC):

$$ICC = \sigma u^2\ /\ (\sigma u^2 + \delta \iota \alpha \kappa \kappa \mu \alpha \nu \sigma \eta\ \sigma \varepsilon\ \varepsilon \pi \Omega \pi \varepsilon \delta o\ \pi \alpha \rho \alpha \tau \Psi \rho \eta \sigma \eta \varsigma)$$

Στα γραμμικά μικτά μοντέλα ο ICC υπολογίζεται άμεσα. Στα μοντέλα καταμέτρησης (Poisson και Αρνητικής Διωνυμικής), όμως, η διακύμανση σε επίπεδο παρατήρησης δεν είναι σταθερή αλλά εξαρτάται από τη μέση τιμή, με αποτέλεσμα ο ICC να μην ορίζεται με απλό και μονοσήμαντο τρόπο. Για τον λόγο αυτό, στα μοντέλα αυτά η χρησιμότητα της τυχαίας επίδρασης αξιολογείται συνήθως με σύγκριση του μοντέλου σταθερών επιδράσεων (fixed-effects) με το αντίστοιχο μοντέλο μικτών επιδράσεων (mixed-effects), μέσω του ελέγχου λόγου πιθανοφάνειας και των **κριτηρίων AIC και BIC**, σε συνδυασμό με την εκτιμώμενη διακύμανση σu² του **τυχαίου σταθερού όρου**.

3.4 Στατιστικοί έλεγχοι και κριτήρια αποδοχής μοντέλου

Μετά τη διαμόρφωση των μοντέλων απαιτείται συστηματική αξιολόγηση της στατιστικής τους επάρκειας και της λογικής ερμηνείας των εκτιμήσεων. Στο πλαίσιο αυτό εξετάζονται η στατιστική σημαντικότητα των συντελεστών και των μεταβλητών ως συνόλου, η σύγκριση εναλλακτικών προδιαγραφών, η πιθανή πολυσυγγραμμικότητα και, όπου εφαρμόζεται, η συμπεριφορά των καταλοίπων του τελικού μοντέλου.

Παρακάτω παρουσιάζονται οι συνηθέστεροι έλεγχοι και κριτήρια για την αποδοχή ενός μοντέλου παλινδρόμησης:

Συντελεστές παλινδρόμησης

Κριτήριο αποδοχής ενός μοντέλου αποτελεί η λογική ερμηνεία των προσήμων των συντελεστών του. Αν τα πρόσημα δεν μπορούν να ερμηνευθούν λογικά, το μοντέλο πρέπει να επανεξεταστεί ή να απορριφθεί.

Ο έλεγχος αφορά όλες τις ανεξάρτητες μεταβλητές. Θετικό πρόσημο του συντελεστή β συνεπάγεται αύξηση της αναμενόμενης τιμής της εξαρτημένης μεταβλητής όταν αυξάνεται η ανεξάρτητη μεταβλητή ή, στην περίπτωση κατηγορικής μεταβλητής, για τη συγκεκριμένη κατηγορία σε σχέση με την κατηγορία αναφοράς. Αρνητικό πρόσημο συνεπάγεται το αντίθετο, δηλαδή μείωση της εξαρτημένης μεταβλητής, και επομένως αρνητική συσχέτιση μεταξύ των μεταβλητών.

Συσχέτιση Παραμέτρων

Ο συντελεστής συσχέτισης (ενότητα 3.2.2) χρησιμοποιείται για τον έλεγχο της συσχέτισης μεταξύ των ανεξάρτητων μεταβλητών. Υψηλές απόλυτες τιμές (κοντά στο 1 ή στο -1) υποδηλώνουν ότι δύο ανεξάρτητες μεταβλητές μεταφέρουν σε μεγάλο βαθμό την ίδια πληροφορία και ότι η ταυτόχρονη ένταξή τους στο μοντέλο μπορεί να προκαλέσει πολυσυγγραμμικότητα. Τιμές κοντά στο 0 υποδηλώνουν απουσία γραμμικής συσχέτισης.

Δείκτης Πολυσυγγραμμικότητας (VIF)

Η πολυσυγγραμμικότητα μεταξύ των ανεξάρτητων μεταβλητών ελέγχεται πιο συστηματικά μέσω του δείκτη VIF (Variance Inflation Factor). Για κάθε ανεξάρτητη μεταβλητή j ο δείκτης υπολογίζεται ως $VIFj = 1/(1 - Rj^2)$, όπου Rj^2 είναι ο συντελεστής προσδιορισμού της παλινδρόμησης της μεταβλητής j στις υπόλοιπες ανεξάρτητες μεταβλητές. Τιμές VIF μεγαλύτερες του 5 (ή, σύμφωνα με λιγότερο αυστηρά κριτήρια, του 10) υποδηλώνουν προβληματική πολυσυγγραμμικότητα, η οποία διογκώνει τα τυπικά σφάλματα των συντελεστών και καθιστά ασταθείς τις εκτιμήσεις τους (Field, 2018).

Στατιστική σημαντικότητα (Wald z-test)

Για τον έλεγχο της στατιστικής σημαντικότητας κάθε επιμέρους συντελεστή β στα μοντέλα GLM χρησιμοποιείται ο έλεγχος Wald (Wald z-test), αντίστοιχος του κλασικού ελέγχου t (t-test) της γραμμικής παλινδρόμησης. Το στατιστικό z προκύπτει από τη διαίρεση της εκτιμώμενης τιμής του συντελεστή με το τυπικό σφάλμα της:

$$z = \beta i \ / \ SE(\beta i)$$

Σε αντίθεση με τον κλασικό έλεγχο t, όπου η κρίσιμη τιμή εξαρτάται από τους βαθμούς ελευθερίας του δείγματος, ο έλεγχος Wald βασίζεται ασυμπτωτικά στην τυποποιημένη κανονική κατανομή, με κρίσιμη τιμή περίπου 1,96 για αμφίπλευρο έλεγχο στο επίπεδο 5%. Μεγαλύτερη απόλυτη τιμή του z παρέχει ισχυρότερη ένδειξη ότι ο αντίστοιχος συντελεστής διαφέρει από το μηδέν, υπό τις προϋποθέσεις του μοντέλου.

Έλεγχος Λόγου Πιθανοφάνειας (Likelihood-Ratio Test, LRT)

Όταν μια κατηγορική μεταβλητή αντιστοιχεί σε περισσότερες από μία δυαδικές (dummy) μεταβλητές, ο έλεγχος των μεμονωμένων συντελεστών δεν αρκεί για να κριθεί αν η μεταβλητή συνεισφέρει στο μοντέλο και απαιτείται έλεγχος της σημαντικότητάς της ως συνόλου. Για τον σκοπό αυτό χρησιμοποιείται ο **έλεγχος λόγου πιθανοφάνειας** (Likelihood-Ratio Test, LRT), ο οποίος συγκρίνει τη λογαριθμική πιθανοφάνεια (log-likelihood) δύο ένθετων μοντέλων, με και χωρίς την υπό εξέταση μεταβλητή:

$$LR = -2 \cdot (logLik(\mu\varepsilon\iota\omega\mu X\nu o \ \mu o\nu\tau X\lambda o) - logLik(\pi\lambda\Psi\rho\varepsilon\varsigma \ \mu o\nu\tau X\lambda o))$$

Για τη σύγκριση ένθετων μοντέλων ως προς τις σταθερές επιδράσεις, το στατιστικό LR ακολουθεί ασυμπτωτικά κατανομή χ^2 με βαθμούς ελευθερίας ίσους με τη διαφορά του πλήθους των παραμέτρων των δύο μοντέλων. Ο έλεγχος LRT αποτελεί τη βάση της διαδικασίας οπισθοδρομικής απαλοιφής (backward elimination), κατά την οποία, ξεκινώντας από το πλήρες μοντέλο, αφαιρείται σε κάθε βήμα η λιγότερο σημαντική μεταβλητή και η διαδικασία επαναλαμβάνεται έως ότου όλες οι εναπομείνασες μεταβλητές είναι στατιστικά σημαντικές στο προκαθορισμένο επίπεδο. Ειδική περίπτωση αποτελεί ο έλεγχος της διακύμανσης ενός τυχαίου σταθερού όρου (H0: $\sigma u^2 = 0$). Επειδή η τιμή μηδέν βρίσκεται στο όριο του χώρου των παραμέτρων (η διακύμανση δεν μπορεί να είναι αρνητική), η κατανομή χ^2 με έναν βαθμό ελευθερίας δεν είναι ακριβής και το στατιστικό LR ακολουθεί μείγμα $0{,}5 \cdot \chi^2(0) + 0{,}5 \cdot \chi^2(1)$. Στην πράξη, αυτό σημαίνει ότι η τιμή p του κλασικού ελέγχου διαιρείται με το 2.

Κριτήρια Σύγκρισης Μοντέλων (AIC, BIC και λογαριθμική πιθανοφάνεια)

Στα μοντέλα GLM, η σύγκριση εναλλακτικών μοντέλων πραγματοποιείται με βάση τη συνολική ποιότητα προσαρμογής τους στα δεδομένα. Για τον σκοπό αυτό χρησιμοποιούνται κυρίως η λογαριθμική πιθανοφάνεια (log-likelihood) και τα κριτήρια πληροφορίας Akaike (Akaike Information Criterion, AIC) και Bayes (Bayesian Information Criterion, BIC).

Η λογαριθμική πιθανοφάνεια εκφράζει τον βαθμό στον οποίο το εκτιμημένο μοντέλο προσαρμόζεται στα παρατηρούμενα δεδομένα. Μεγαλύτερες τιμές της λογαριθμικής πιθανοφάνειας, δηλαδή τιμές πλησιέστερες στο μηδέν, υποδηλώνουν καλύτερη προσαρμογή. Τα κριτήρια **AIC** και **BIC** υπολογίζονται ως εξής:

$$AIC = -2 \cdot logLik + 2k$$

$$BIC = -2 \cdot logLik + k \cdot ln(n)$$

όπου logLik είναι η τιμή της λογαριθμικής πιθανοφάνειας του μοντέλου, (k) το πλήθος των εκτιμώμενων παραμέτρων και (n) το μέγεθος του δείγματος.

Και τα δύο κριτήρια λαμβάνουν υπόψη τόσο την προσαρμογή του μοντέλου όσο και την πολυπλοκότητά του, επιβάλλοντας ποινή για την προσθήκη επιπλέον παραμέτρων. Το BIC επιβάλλει ισχυρότερη ποινή στην πολυπλοκότητα, ιδιαίτερα σε μεγάλα δείγματα. Κατά τη σύγκριση μοντέλων που έχουν εκτιμηθεί στο ίδιο σύνολο δεδομένων και για την ίδια εξαρτημένη μεταβλητή, προτιμώνται γενικά μικρότερες τιμές AIC και BIC, καθώς υποδηλώνουν καλύτερη ισορροπία μεταξύ προσαρμογής και πολυπλοκότητας (Burnham & Anderson, 2002).

Τιμή p (p-value)

Η τιμή p εκφράζει, υπό τη μηδενική υπόθεση και τις παραδοχές του αντίστοιχου στατιστικού ελέγχου, την πιθανότητα να παρατηρηθεί τιμή του στατιστικού ελέγχου τουλάχιστον τόσο ακραία όσο η τιμή που προέκυψε από τα δεδομένα.

Συνήθως χρησιμοποιείται επίπεδο σημαντικότητας α = 0,05. Στην περίπτωση αυτή, όταν p < 0,05 η μηδενική υπόθεση απορρίπτεται, ενώ όταν p ≥ 0,05 δεν υπάρχει επαρκής στατιστική ένδειξη για την απόρριψή της. Η τιμή p δεν αποτελεί μέτρο του μεγέθους ή της πρακτικής σημασίας μιας σχέσης και πρέπει να ερμηνεύεται μαζί με τον συντελεστή, το **IRR** και το διάστημα εμπιστοσύνης.

Επίπεδο σημαντικότητας

Κατά τον στατιστικό έλεγχο διατυπώνεται συνήθως μία μηδενική υπόθεση H0, η οποία εκφράζει απουσία επίδρασης ή διαφοράς. Η τιμή p αποτελεί μέτρο της συμβατότητας των παρατηρούμενων δεδομένων με τη μηδενική υπόθεση. Εάν η τιμή p είναι μικρότερη από το προκαθορισμένο επίπεδο σημαντικότητας α, η H0 απορρίπτεται.

Στατιστικό σφάλμα

Στατιστικό σφάλμα ονομάζεται η διαφορά μεταξύ μιας τιμής που προκύπτει από μια διαδικασία συλλογής δεδομένων και της πραγματικής τιμής του πληθυσμού. Τα σφάλματα διακρίνονται σε δειγματοληπτικά, τα οποία οφείλονται στο ότι εξετάζεται μόνο ένα μέρος του πληθυσμού, και σε μη δειγματοληπτικά, τα οποία οφείλονται σε λάθη καταγραφής, κωδικοποίησης ή μέτρησης.

Όσο μεγαλύτερο είναι το σφάλμα, τόσο λιγότερο αντιπροσωπευτικά είναι τα δεδομένα για τον πληθυσμό. Το δειγματοληπτικό σφάλμα μειώνεται όσο αυξάνεται το μέγεθος του δείγματος, ενώ τα μη δειγματοληπτικά σφάλματα δεν περιορίζονται απαραίτητα με τη χρήση μεγαλύτερου δείγματος.

Έλεγχοι Επάρκειας Μοντέλου GLM (Model Fit Diagnostics)

Η καταλληλότητα ενός μοντέλου GLM ή **GLMM** μπορεί να ελεγχθεί μέσω προσομοιωμένων καταλοίπων (simulation-based residuals) με τη μέθοδο **DHARMa** (Hartig, 2022). Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη στα μοντέλα καταμέτρησης και στα μοντέλα μικτών επιδράσεων, όπου τα κλασικά κατάλοιπα (κατάλοιπα Pearson ή απόκλισης, deviance) δεν έχουν γνωστή κατανομή αναφοράς και είναι δύσκολο να ερμηνευθούν.

Η μέθοδος αυτή παράγει, μέσω επαναληπτικών προσομοιώσεων από το εκτιμημένο μοντέλο, μία εμπειρική κατανομή αναφοράς για κάθε παρατήρηση και συγκρίνει τις πραγματικές τιμές με αυτήν μέσω τυποποιημένων καταλοίπων (scaled residuals), τα οποία, υπό ορθή προσαρμογή του μοντέλου, ακολουθούν ομοιόμορφη κατανομή στο διάστημα [0,1]. Οι κυριότεροι έλεγχοι που παρέχει η μέθοδος είναι:

- Έλεγχος διασποράς (dispersion test): εξετάζει αν η παρατηρούμενη διασπορά των καταλοίπων αποκλίνει σημαντικά από την αναμενόμενη βάσει του εκτιμημένου μοντέλου, υποδεικνύοντας πιθανή υπερδιασπορά ή υποδιασπορά.
- Έλεγχος μηδενικών (zero-inflation test): εξετάζει αν το παρατηρούμενο πλήθος μηδενικών διαφέρει από εκείνο που αναμένεται βάσει των προσομοιώσεων του εκτιμημένου μοντέλου, δηλαδή αν υπάρχουν περισσότερα ή λιγότερα μηδενικά από τα αναμενόμενα.
- Έλεγχος ακραίων τιμών (outlier test): εντοπίζει παρατηρήσεις με κατάλοιπα εκτός του αναμενόμενου εύρους της προσομοιωμένης κατανομής.
- Έλεγχος ομοιομορφίας (uniformity test, Kolmogorov-Smirnov): εξετάζει αν η κατανομή των τυποποιημένων καταλοίπων αποκλίνει σημαντικά από την

αναμενόμενη ομοιόμορφη κατανομή και αποτελεί έναν γενικό δείκτη της συνολικής προσαρμογής του μοντέλου.

Σε μεγάλα δείγματα οι παραπάνω έλεγχοι καθίστανται ιδιαίτερα ευαίσθητοι και μπορεί να εντοπίσουν **στατιστικά σημαντικές** αλλά πρακτικά αμελητέες αποκλίσεις, οπότε η ερμηνεία τους πρέπει να συνδυάζεται πάντα με την εξέταση του μεγέθους της απόκλισης και όχι μόνο της στατιστικής της σημαντικότητας.

Λόγος Ρυθμών Επίπτωσης (Incidence Rate Ratio, IRR)

Για την ερμηνεία των συντελεστών των μοντέλων καταμέτρησης χρησιμοποιείται ο λόγος ρυθμών επίπτωσης (IRR), ο οποίος προκύπτει από την εκθετοποίηση του αντίστοιχου συντελεστή παλινδρόμησης:

$$IRRi = exp(\beta i)$$

Το IRR εκφράζει τον πολλαπλασιαστικό παράγοντα μεταβολής του αναμενόμενου αριθμού συμβάντων για μία μονάδα αύξησης μίας συνεχούς ανεξάρτητης μεταβλητής ή για μία συγκεκριμένη κατηγορία σε σχέση με την κατηγορία αναφοράς, διατηρώντας τις υπόλοιπες μεταβλητές σταθερές.

- IRR > 1: υποδηλώνει αύξηση του αναμενόμενου αριθμού συμβάντων.
- IRR < 1: υποδηλώνει μείωση του αναμενόμενου αριθμού συμβάντων.
- IRR = 1: υποδηλώνει απουσία επίδρασης ή διαφοροποίησης.

Διαστήματα Εμπιστοσύνης Συντελεστή

Για κάθε συντελεστή β̂i, το 95% διάστημα εμπιστοσύνης υπολογίζεται ως:

$$\beta_{\mathrm{i}} \pm 1,96 \cdot SE(\beta_{\mathrm{i}})$$

Το αντίστοιχο 95% διάστημα εμπιστοσύνης για το IRR προκύπτει με εκθετοποίηση των ορίων του παραπάνω διαστήματος:

$$[\, exp(\beta_{\mathrm{i}} - 1,96 \cdot SE(\beta_{\mathrm{i}})) \;,\; exp(\beta_{\mathrm{i}} + 1,96 \cdot SE(\beta_{\mathrm{i}})) \,]$$

Εάν το διάστημα αυτό περιλαμβάνει την τιμή 1, η επίδραση της αντίστοιχης μεταβλητής δεν θεωρείται στατιστικά σημαντική στο επίπεδο 5%.

4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΣΤΟΙΧΕΙΩΝ

4.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφεται η **διαδικασία συλλογής, οργάνωσης και επεξεργασίας των δεδομένων** που χρησιμοποιήθηκαν για τη **διερεύνηση των παραγόντων** που σχετίζονται **με τον αναμενόμενο αριθμό παθόντων σε δύο κατηγορίες σοβαρότητας τραυματισμού** στα οδικά ατυχήματα με **μηχανοκίνητα δίκυκλα**. Μετά τη βιβλιογραφική ανασκόπηση και την παρουσίαση του θεωρητικού υπόβαθρου των μοντέλων καταμέτρησης, το επόμενο στάδιο αφορά τη διαμόρφωση της βάσης μοντελοποίησης, ώστε οι διαθέσιμες πληροφορίες να χρησιμοποιηθούν με συνεπή και ερμηνεύσιμο τρόπο στη στατιστική ανάλυση.

Αρχικά, παρουσιάζονται η πηγή των **δεδομένων**, η διαδικασία συλλογής των στοιχείων και η δομή της αρχικής βάσης. Στη συνέχεια, περιγράφεται η επεξεργασία των δεδομένων: η κωδικοποίηση και ομαδοποίηση των μεταβλητών, η δημιουργία των δύο βάσεων μοντελοποίησης ανά κατηγορία σοβαρότητας, η επιλογή των κατηγοριών αναφοράς και τα βασικά περιγραφικά χαρακτηριστικά των δεδομένων. Έπειτα, παρουσιάζεται ο αλγόριθμος που αναπτύχθηκε στη γλώσσα R για την εκτίμηση των μοντέλων, καθώς και η γεωγραφική κατανομή των εγγραφών. Το κεφάλαιο ολοκληρώνεται με τη σύνοψη των βασικών παρατηρήσεων.

4.2 Συλλογή δεδομένων

4.2.1 Πηγή δεδομένων

Τα δεδομένα της παρούσας Διπλωματικής Εργασίας αφορούν **οδικά ατυχήματα με εμπλοκή μηχανοκίνητων δικύκλων στην Ελλάδα για το έτος 2023**. Η αρχική βάση προέκυψε από στοιχεία που διατέθηκαν από το Παρατηρητήριο Οδικής Ασφάλειας του Εθνικού Μετσόβιου Πολυτεχνείου (NTUA Road Safety Observatory) και την Ελληνική Στατιστική Αρχή (ΕΛ.ΣΤΑΤ.). Για κάθε καταγραφή περιλαμβάνονται πληροφορίες που αφορούν **τον παθόντα, το όχημα, τον χρόνο και τον τόπο του ατυχήματος, το οδικό περιβάλλον, τον τύπο του ατυχήματος και τον έλεγχο αλκοόλ.**

Η αρχική βάση περιλαμβάνει **6.585 καταγραφές παθόντων**. Από αυτές, 5.981 αφορούν ελαφρά τραυματίες (Slight Injured), 357 βαριά τραυματίες (Serious Injured) και 247 νεκρούς (Killed). Για τις ανάγκες της μοντελοποίησης, οι βαριά τραυματίες και οι νεκροί εξετάζονται από κοινού ως ενιαία κατηγορία νεκρών και βαριά τραυματιών (Killed or Seriously Injured, KSI), η οποία περιλαμβάνει συνολικά 604 παθόντες (9,17%). Τα βασικά μεγέθη της αρχικής βάσης συνοψίζονται στον Πίνακα 4.1.

Η συνένωση αυτή κρίθηκε απαραίτητη λόγω της **έντονης ανισορροπίας των δεδομένων ως προς τη σοβαρότητα τραυματισμού.** Οι ελαφρά τραυματίες αποτελούν το 90,83% των παθόντων, ενώ οι βαριά τραυματίες και οι νεκροί μόλις το 5,42% και το 3,75%, αντίστοιχα. Με τόσο μικρό αριθμό περιστατικών, η χωριστή μοντελοποίηση των νεκρών και των βαριά τραυματιών θα οδηγούσε σε πολύ μεγάλο ποσοστό συνδυασμών χαρακτηριστικών χωρίς κανένα περιστατικό και σε ασταθείς εκτιμήσεις των συντελεστών, ιδιαίτερα για τις λιγότερο συχνές υποκατηγορίες των μεταβλητών. Επιπλέον, ο δείκτης KSI χρησιμοποιείται ευρέως στη διεθνή βιβλιογραφία και στις στατιστικές οδικής ασφάλειας ως μέτρο των σοβαρών συνεπειών των ατυχημάτων, γεγονός που διευκολύνει τη σύγκριση των αποτελεσμάτων με άλλες μελέτες. Έτσι, διερευνώνται χωριστά ο αναμενόμενος αριθμός ελαφρά τραυματιών και ο αναμενόμενος αριθμός νεκρών και βαριά τραυματιών.

Πίνακας 4.1: Βασικά μεγέθη της αρχικής βάσης δεδομένων

Μέγεθος	Αριθμός	Ποσοστό
Σύνολο καταγραφών παθόντων	6.585	100,00%
Ελαφρά τραυματίες (Slight Injured)	5.981	90,83%
Βαριά τραυματίες (Serious Injured)	357	5,42%
Νεκροί (Killed)	247	3,75%

4.2.2 Διαδικασία συλλογής στοιχείων

Η αρχική βάση οργανώθηκε σε επίπεδο παθόντος: **κάθε γραμμή αντιστοιχεί σε ένα άτομο που τραυματίστηκε ή σκοτώθηκε ως χρήστης μηχανοκίνητου δικύκλου σε οδικό ατύχημα του 2023.** Για κάθε παθόντα καταγράφηκαν **25 πεδία**, από τα οποία τα 17 περιέχουν τις αρχικές τιμές της καταγραφής και παρουσιάζονται στον Πίνακα 4.2, ενώ τα υπόλοιπα 8 αποτελούν ομαδοποιημένες εκδοχές που δημιουργήθηκαν για τις ανάγκες της ανάλυσης (Ενότητα 4.3.1).

Τα αρχικά πεδία αφορούν το έτος και τον μήνα του ατυχήματος, τον τύπο χρήστη της οδού (οδηγός ή επιβάτης), το φύλο και την ηλικία, τη σοβαρότητα τραυματισμού, τον τύπο οδού, την ένδειξη εντός ή εκτός κατοικημένης περιοχής, την Περιφερειακή Ενότητα, την ημέρα της εβδομάδας και την ώρα, τον αριθμό των εμπλεκόμενων οχημάτων, τις καιρικές συνθήκες, την ύπαρξη κόμβου, τον λεπτομερή τύπο ατυχήματος, την κατηγορία του δικύκλου βάσει κυβισμού και το αποτέλεσμα του ελέγχου αλκοόλ στον οδηγό. Η άγνωστη ηλικία είναι καταχωρημένη με την τιμή 99, ενώ στο πεδίο του ελέγχου αλκοόλ εμφανίζονται, εκτός από τις αριθμητικές μετρήσεις, ο κωδικός 9,99 και η ένδειξη «άγνωστο» (Unknown) για τις περιπτώσεις χωρίς διαθέσιμη μέτρηση.

Πίνακας 4.2: Πεδία της αρχικής βάσης δεδομένων

Πεδίο	Τιμές
Έτος (YEAR)	2023
Μήνας (MONTH)	Ιανουάριος - Δεκέμβριος
Τύπος χρήστη (ROAD USER TYPE)	Οδηγός (Driver), Επιβάτης (Passenger)
Φύλο (GENDER)	Άνδρας (Male), Γυναίκα (Female)
Ηλικία (AGE)	0 - 98, 99 = άγνωστη
Σοβαρότητα (SEVERITY)	Ελαφρά τραυματίας (Slight Injured), Βαριά τραυματίας (Serious Injured), Νεκρός (Killed)
Τύπος οδού (ROAD TYPE)	Δημοτική (Municipal), επαρχιακή (Regional), παλαιά εθνική (Old National), νέα εθνική (New National), τοπική (Local), λοιπά (Other)
Κατοικημένη περιοχή (URBAN AREA)	Εντός (Inside), Εκτός (Outside)
Περιφερειακή Ενότητα (REGIONAL UNIT)	73 Περιφερειακές Ενότητες
Ημέρα (DAY)	Δευτέρα - Κυριακή
Ώρα (HOUR)	1 - 24
Εμπλεκόμενα οχήματα (VEHICLES INVOLVED)	1 - 7
Καιρικές συνθήκες (WEATHER CONDITIONS)	9 κατηγορίες, π.χ. αίθριος καιρός (Clear sky), ψιχάλα (Drizzle), βροχή (Rain)
Κόμβος (JUNCTION)	Ναι (Yes), Όχι (No)
Τύπος ατυχήματος (ACCIDENT TYPE)	18 κατηγορίες
Τύπος οχήματος (VEHICLE TYPE)	5 κατηγορίες, από μοτοποδήλατο έως 49cc έως μοτοσικλέτα άνω των 730cc
Έλεγχος αλκοόλ (ALCOHOL TEST)	Αριθμητική τιμή, κωδικός 9,99 ή άγνωστο (Unknown)

Η συλλογή και η προετοιμασία των στοιχείων πραγματοποιήθηκαν σε διαδοχικά στάδια. Αρχικά, οι καταγραφές των παθόντων σε οδικά ατυχήματα του 2023 με εμπλοκή μηχανοκίνητου δικύκλου συγκεντρώθηκαν σε ενιαίο φύλλο εργασίας του Microsoft Excel, με μία γραμμή ανά παθόντα και μία στήλη ανά πεδίο. Στη συνέχεια, ε**λέγχθηκαν η πληρότητα και η συνέπεια των τιμών** κάθε πεδίου, ώστε να εντοπιστούν κενές ή μη έγκυρες καταχωρίσεις και ειδικοί κωδικοί, όπως η τιμή 99 στην ηλικία και ο κωδικός 9,99 στον έλεγχο αλκοόλ. Οι κωδικοί αυτοί δεν διαγράφηκαν, αλλά αντιστοιχίστηκαν σε ξεχωριστές κατηγορίες «άγνωστο», ώστε να μη χαθεί καμία καταγραφή.

Τέλος, για κάθε πεδίο **δημιουργήθηκαν** στο ίδιο αρχείο **συγκεντρωτικοί πίνακες συχνοτήτων**, αρχικά για κάθε μεταβλητή μεμονωμένα και στη συνέχεια σε συνδυασμό με τη σοβαρότητα τραυματισμού. Οι πίνακες αυτοί **αποτέλεσαν τη βάση** για την **περιγραφική ανάλυση** και για τις ομαδοποιήσεις των μεταβλητών που περιγράφονται στην Ενότητα 4.3.1.

4.3 Επεξεργασία δεδομένων

4.3.1 Διαμόρφωση δεδομένων

Η χρήση **ομαδοποιημένων μεταβλητών κρίθηκε απαραίτητη** ώστε να περιοριστεί η υπερβολική διάσπαση του δείγματος σε κατηγορίες με πολύ μικρό αριθμό παρατηρήσεων και να διατηρηθεί η ουσιαστική ερμηνεία των αποτελεσμάτων. Παράλληλα, οι ομαδοποιήσεις διευκολύνουν τη σύγκριση κατηγοριών με κοινά λειτουργικά ή κυκλοφοριακά χαρακτηριστικά και ενισχύουν τη σταθερότητα της εκτίμησης των συντελεστών.

Η επεξεργασία των αρχικών στοιχείων πραγματοποιήθηκε με στόχο τη δημιουργία μεταβλητών κατάλληλων για την ανάπτυξη Γενικευμένων Γραμμικών Μοντέλων καταμέτρησης. Ο μήνας μετατράπηκε σε εποχή του έτους (SEASON): χειμώνας (Winter: Δεκέμβριος - Φεβρουάριος), άνοιξη (Spring: Μάρτιος - Μάιος), καλοκαίρι (Summer: Ιούνιος - Αύγουστος) και φθινόπωρο (Autumn: Σεπτέμβριος - Νοέμβριος). Η ημέρα μετατράπηκε σε τύπο ημέρας (DAY TYPE): καθημερινή (Weekday: Δευτέρα - Παρασκευή) ή Σαββατοκύριακο (Weekend). Η ώρα ομαδοποιήθηκε σε περίοδο ημέρας (HOUR DAY/NIGHT): ημέρα (Day: τιμές 7:00 έως 19:00) και νύχτα (Night: τιμές 20:00 έως 24:00 και 1:00 έως 6:00). Στον τύπο οδού (ROAD TYPE GROUP), η παλαιά εθνική (Old National) και η νέα εθνική οδός (New National) συνενώθηκαν στην κατηγορία εθνική οδός (National), ενώ η δημοτική (Municipal), η επαρχιακή (Regional), η τοπική οδός (Local) και η κατηγορία λοιπά (Other) στην κατηγορία μη εθνική οδός (Non-National). Οι καιρικές συνθήκες (WEATHER) ομαδοποιήθηκαν σε καλές (Good), που αντιστοιχούν σε αίθριο καιρό, και λοιπές (Other), που περιλαμβάνουν όλες τις υπόλοιπες συνθήκες (ψιχάλα, βροχή, παγετό, ισχυρό άνεμο, ομίχλη, θύελλα και καταιγίδα).

Η ηλικία ομαδοποιήθηκε σε ηλικιακές ομάδες (AGE GROUP): 0-14, 15-18, 19-24, 25-49, 50-64, 65+ και άγνωστη (Unknown). Ο αριθμός εμπλεκόμενων οχημάτων (VEHICLES INVOLVED) ομαδοποιήθηκε σε ένα όχημα (1 vehicle), δύο οχήματα (2 vehicles) και τρία ή περισσότερα οχήματα (3+ vehicles). Ο τύπος του ατυχήματος (ACCIDENT TYPE) διαμορφώθηκε σε τέσσερις κατηγορίες: σύγκρουση με άλλους χρήστες της οδού (Collision with road users), που περιλαμβάνει την πλαγιομετωπική, την πλευρική και τη μετωπική σύγκρουση, τη σύγκρουση από πίσω και τη σύγκρουση με πεζό, εκτροπή (Diversion) δεξιά, αριστερά ή στο αντίθετο ρεύμα, λοιπές συγκρούσεις (Other collision), όπως πρόσκρουση σε σταθερό εμπόδιο, στύλο ή δέντρο, σε σταθμευμένο ή σταματημένο όχημα, σε ζώο ή σε συρμό, και

ανατροπή/λοιπά (Overturning/Other), που περιλαμβάνει την ανατροπή εντός ή εκτός οδοστρώματος και λοιπούς τύπους.

Ο τύπος του μηχανοκίνητου δικύκλου (VEHICLE TYPE) ομαδοποιήθηκε σε μικρά (Small PTWs: μοτοποδήλατα έως 49cc και μοτοσικλέτες 50-115cc), μεσαία (Medium PTWs: 116-269cc) και μεγάλα δίκυκλα (Large PTWs: 270-730cc και άνω των 730cc). Για τον έλεγχο αλκοόλ (ALCOHOL TEST) δημιουργήθηκαν οι κατηγορίες αρνητικός/μηδενικός (Negative/zero: μέτρηση 0), χαμηλά θετικός (Low positive: 0,02 έως 0,49), υψηλά θετικός (High positive: 0,50 και άνω) και άγνωστος/δεν διενεργήθηκε (Unknown/not tested: κωδικός 9,99 ή ένδειξη Unknown). Η τελευταία κατηγορία διατηρήθηκε ως διακριτό επίπεδο της μεταβλητής και δεν αντιμετωπίστηκε ως ελλιπής τιμή, καθώς η απουσία διαθέσιμης μέτρησης αποτελεί και αυτή χαρακτηριστικό της καταγραφής.

Για την παρουσίαση της δομής του τελικού συνόλου **δεδομένων**, στον ακόλουθο πίνακα παρατίθεται ενδεικτικό απόσπασμα από τον συγκεντρωτικό πίνακα (master table), μετά την ομαδοποίηση των μεταβλητών. Κάθε γραμμή αντιστοιχεί σε έναν παθόντα και κάθε στήλη σε ένα χαρακτηριστικό της καταγραφής.

MONTH	ROAD USER	GENDER	URBAN /	DAY	JUNCTI	SEVERITY	AGE_GI	HOUR_D	VEHICLES_INVC	WEATHER	ACCIDENT_TYPE_	VEHICLE_TYP	ALCOHOL_T	T_4CAT
January	Driver	Male	Inside	Wednesd	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Friday	No	0	25-49	Night	1 vehicle	Good	Diversion	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Friday	No	0	25-49	Day	2 vehicles	Good	Collision with road	Large PTW	Unknown/not tested	
January	Driver	Male	Inside	Friday	No	0	25-49	Day	3+ vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Monday	Yes	0	25-49	Day	2 vehicles	Good	Collision with road	Large PTW	Negative/zero	
January	Driver	Male	Inside	Sunday	No	0	25-49	Day	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Monday	Yes	0	Unknown	Night	2 vehicles	Good	Collision with road	Large PTW	Unknown/not tested	
January	Driver	Male	Inside	Wednesd	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Monday	Yes	0	50-64	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Sunday	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Large PTW	Unknown/not tested	
January	Driver	Male	Inside	Sunday	Yes	0	25-49	Day	2 vehicles	Good	Collision with road	Large PTW	Unknown/not tested	
January	Passenger	Female	Inside	Saturday	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Saturday	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Monday	No	0	25-49	Day	2 vehicles	Good	Collision with road	Large PTW	Negative/zero	
January	Passenger	Male	Inside	Saturday	Yes	0	25-49	Day	3+ vehicles	Good	Collision with road	Large PTW	Unknown/not tested	
January	Driver	Male	Inside	Friday	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Thursday	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Small PTW	Unknown/not tested	
January	Driver	Male	Inside	Wednesd	Yes	0	25-49	Night	2 vehicles	Good	Collision with road	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Tuesday	No	0	25-49	Night	1 vehicle	Good	Other collision	Medium PTW	Unknown/not tested	
January	Driver	Male	Inside	Tuesday	Yes	0	65+	Day	2 vehicles	Good	Collision with road	Small PTW	Unknown/not tested	
January	Driver	Male	Inside	Monday	Yes	0	19-24	Night	2 vehicles	Good	Collision with road	Small PTW	Unknown/not tested	

Τέλος, οι 73 Περιφερειακές Ενότητες ομαδοποιήθηκαν σε **12 ευρύτερες γεωγραφικές περιοχές (REGION)**, οι οποίες παρουσιάζονται αναλυτικά, μαζί με τις Περιφερειακές Ενότητες που περιλαμβάνει η καθεμία, στον Πίνακα 4.5 της Ενότητας 4.3.3. Η συγκεκριμένη μεταβλητή **δεν χρησιμοποιήθηκε ως** μία από τις 14 υποψήφιες **σταθερές επεξηγηματικές μεταβλητές**, αλλά ως **μεταβλητή ομαδοποίησης** για τον τυχαίο σταθερό όρο (random intercept) των μικτών μοντέλων (mixed models). Με τον τρόπο αυτό είναι δυνατή η αποτύπωση μη παρατηρούμενης ετερογένειας μεταξύ

ευρύτερων γεωγραφικών περιοχών. Οι ομαδοποιημένες μεταβλητές που χρησιμοποιήθηκαν στη μοντελοποίηση συνοψίζονται στον Πίνακα 4.3.

Πίνακας 4.3: Ομαδοποιημένες μεταβλητές που χρησιμοποιήθηκαν στη μοντελοποίηση

Μεταβλητή	Κατηγορίες	Περιγραφή
Εποχή (SEASON)	Χειμώνας (Winter), Άνοιξη (Spring), Καλοκαίρι (Summer), Φθινόπωρο (Autumn)	Εποχή ατυχήματος
Τύπος χρήστη (ROAD USER TYPE)	Οδηγός (Driver), Επιβάτης (Passenger)	Ιδιότητα του παθόντα στο δίκυκλο
Φύλο (GENDER)	Άνδρας (Male), Γυναίκα (Female)	Φύλο του παθόντα
Τύπος οδού (ROAD TYPE GROUP)	Μη εθνική (Non-National), Εθνική (National)	Ομαδοποίηση τύπου οδού
Κατοικημένη περιοχή (URBAN AREA)	Εντός (Inside), Εκτός (Outside)	Εντός / εκτός κατοικημένης περιοχής
Τύπος ημέρας (DAY TYPE)	Καθημερινή (Weekday), Σαββατοκύριακο (Weekend)	Τύπος ημέρας
Κόμβος (JUNCTION)	Όχι (No), Ναι (Yes)	Ύπαρξη κόμβου
Ηλικιακή ομάδα (AGE GROUP)	0-14, 15-18, 19-24, 25-49, 50-64, 65+, Άγνωστη (Unknown)	Ηλικιακή ομάδα του παθόντα
Περίοδος ημέρας (HOUR DAY/NIGHT)	Ημέρα (Day), Νύχτα (Night)	Χρονική περίοδος του ατυχήματος
Εμπλεκόμενα οχήματα (VEHICLES INVOLVED)	1 όχημα (1 vehicle), 2 οχήματα (2 vehicles), 3+ οχήματα (3+ vehicles)	Αριθμός εμπλεκόμενων οχημάτων
Καιρικές συνθήκες (WEATHER)	Καλές (Good), Λοιπές (Other)	Ομαδοποιημένες καιρικές συνθήκες
Τύπος ατυχήματος (ACCIDENT TYPE)	Σύγκρουση με χρήστες της οδού (Collision with road users), Εκτροπή (Diversion), Λοιπές συγκρούσεις (Other collision), Ανατροπή/Λοιπά (Overturning/Other)	Ομαδοποιημένος τύπος ατυχήματος
Τύπος δικύκλου (VEHICLE TYPE)	Μικρό (Small PTWs), Μεσαίο (Medium PTWs), Μεγάλο (Large PTWs)	Κατηγορία μηχανοκίνητου δικύκλου
Έλεγχος αλκοόλ (ALCOHOL TEST)	Αρνητικός/μηδενικός (Negative/zero), Χαμηλά θετικός (Low positive), Υψηλά θετικός (High positive), Άγνωστος/δεν διενεργήθηκε (Unknown/not tested)	Κατηγορία ελέγχου αλκοόλ
Γεωγραφική περιοχή (REGION)	12 γεωγραφικές περιοχές (Πίνακας 4.5)	Μεταβλητή ομαδοποίησης για τον τυχαίο σταθερό όρο

Για την εφαρμογή των μοντέλων καταμέτρησης δημιουργήθηκαν **δύο ξεχωριστές βάσεις δεδομένων**, με κοινό σύνολο επεξηγηματικών χαρακτηριστικών αλλά διαφορετική εξαρτημένη μεταβλητή. Η πρώτη αφορά τους νεκρούς και βαριά τραυματίες και η δεύτερη τους ελαφρά τραυματίες. Και οι δύο βάσεις περιλαμβάνουν τις ίδιες **4.069 ομαδοποιημένες εγγραφές και 16 στήλες:** τις 14 υποψήφιες

επεξηγηματικές μεταβλητές, **τη γεωγραφική περιοχή** (REGION) **και την εξαρτημένη μεταβλητή, δηλαδή τον αριθμό παθόντων** (CASUALTIES).

Ο αριθμός παθόντων είναι **μεταβλητή καταμέτρησης**, δηλαδή παίρνει **μόνο ακέραιες μη αρνητικές τιμές** (0, 1, 2, ...). Δείχνει πόσοι παθόντες της υπό εξέταση κατηγορίας σοβαρότητας καταγράφηκαν με τον συγκεκριμένο συνδυασμό χαρακτηριστικών: στη βάση των νεκρών και βαριά τραυματιών μετρά μόνο τους νεκρούς και τους βαριά τραυματίες, ενώ στη βάση των ελαφρά τραυματιών μόνο τους ελαφρά τραυματίες. Αυτή είναι η ποσότητα που επιχειρούν να εξηγήσουν τα μοντέλα του Κεφαλαίου 5.

Η **ομαδοποίηση** πραγματοποιήθηκε ως εξής. Στην αρχική βάση κάθε γραμμή είναι ένας παθόντας. Όταν περισσότεροι παθόντες έχουν ακριβώς τα ίδια χαρακτηριστικά σε όλες τις μεταβλητές, δεν καταχωρίζονται σε ξεχωριστές γραμμές, αλλά συγκεντρώνονται σε μία γραμμή και καταμετρώνται. Έτσι, **κάθε γραμμή** των βάσεων μοντελοποίησης **αντιπροσωπεύει έναν μοναδικό συνδυασμό χαρακτηριστικών** και όχι ένα μεμονωμένο ατύχημα ή έναν μεμονωμένο παθόντα.

Με τον τρόπο αυτό, **οι 6.585 καταγραφές παθόντων συμπτύχθηκαν σε 4.069 γραμμές,** ενώ το άθροισμα του αριθμού παθόντων των δύο βάσεων αναπαράγει το σύνολο των παθόντων της αρχικής βάσης. **Διατηρήθηκαν και οι γραμμές με μηδενικό αριθμό παθόντων** της υπό εξέταση κατηγορίας. Για παράδειγμα, ένας συνδυασμός στον οποίο καταγράφηκαν μόνο ελαφρά τραυματίες εμφανίζεται στη βάση των νεκρών και βαριά τραυματιών με τιμή 0. Τα μηδενικά αυτά αποτελούν ουσιαστικό τμήμα της κατανομής της εξαρτημένης μεταβλητής, καθώς δείχνουν σε ποιους συνδυασμούς δεν προέκυψαν παθόντες της συγκεκριμένης σοβαρότητας.

Κατά τον έλεγχο πληρότητας των πεδίων στις δύο βάσεις δεδομένων **δεν προέκυψε απώλεια εγγραφών** και το τελικό σύνολο παρέμεινε ίσο με 4.069 εγγραφές μοντελοποίησης. Οι περιπτώσεις μη διαθέσιμης ηλικίας ή μη διαθέσιμου αποτελέσματος αλκοτέστ είχαν ήδη κωδικοποιηθεί ως διακριτές κατηγορίες: άγνωστη (Unknown) και άγνωστος/δεν διενεργήθηκε (Unknown/not tested), αντίστοιχα.

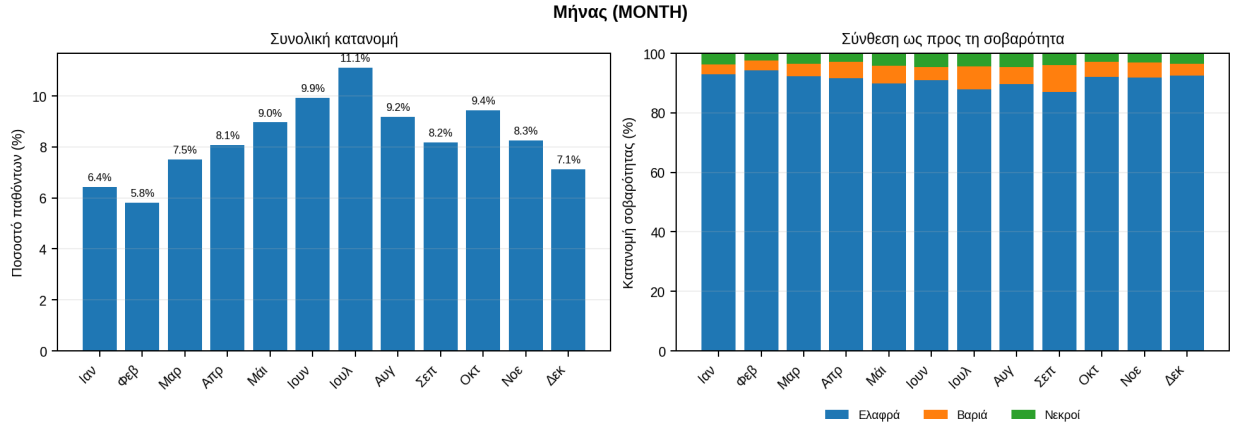
Για κάθε κατηγορική μεταβλητή επιλέχθηκε ως κατηγορία αναφοράς **η συχνότερη κατηγορία** στις 4.069 εγγραφές μοντελοποίησης. Η επιλογή αυτή εξασφαλίζει ότι οι συντελεστές των λιγότερο συχνών κατηγοριών συγκρίνονται με μια κατηγορία που αντιπροσωπεύεται επαρκώς στα δεδομένα. Οι κατηγορίες αναφοράς είναι κοινές στα δύο μοντέλα και παρουσιάζονται στον Πίνακα 4.4.

Πίνακας 4.4: Κατηγορίες αναφοράς των 14 υποψήφιων επεξηγηματικών μεταβλητών

Μεταβλητή	Κατηγορία αναφοράς
Εποχή (SEASON)	Καλοκαίρι (Summer)
Τύπος χρήστη (ROAD USER TYPE)	Οδηγός (Driver)
Φύλο (GENDER)	Άνδρας (Male)
Τύπος οδού (ROAD TYPE GROUP)	Μη εθνική (Non-National)
Κατοικημένη περιοχή (URBAN AREA)	Εντός (Inside)
Τύπος ημέρας (DAY TYPE)	Καθημερινή (Weekday)
Κόμβος (JUNCTION)	Όχι (No)
Ηλικιακή ομάδα (AGE GROUP)	25-49
Περίοδος ημέρας (HOUR DAY/NIGHT)	Ημέρα (Day)
Εμπλεκόμενα οχήματα (VEHICLES INVOLVED)	2 οχήματα (2 vehicles)
Καιρικές συνθήκες (WEATHER)	Καλές (Good)
Τύπος ατυχήματος (ACCIDENT TYPE)	Σύγκρουση με χρήστες της οδού (Collision with road users)
Τύπος δικύκλου (VEHICLE TYPE)	Μεσαίο (Medium PTWs)
Έλεγχος αλκοόλ (ALCOHOL TEST)	Αρνητικός/μηδενικός (Negative/zero)

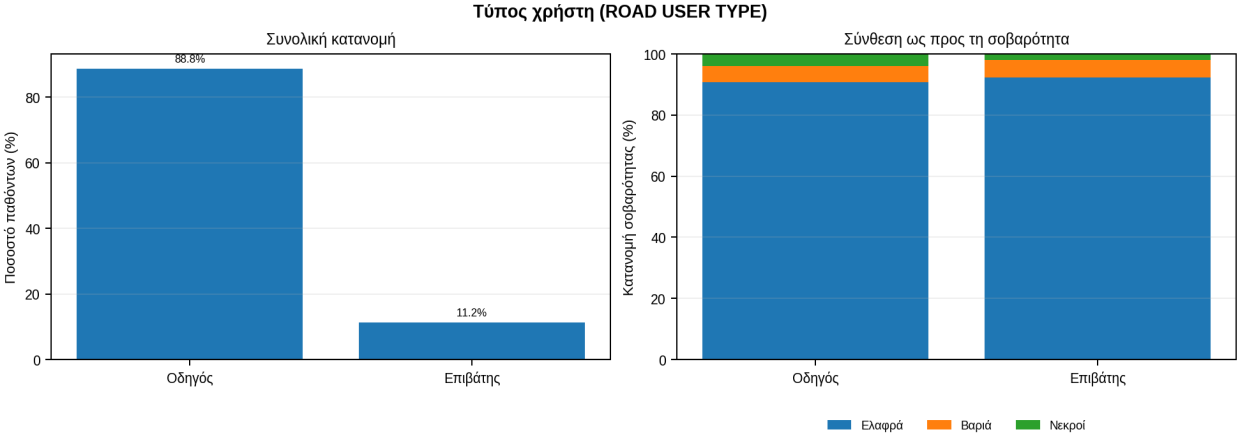
Ειδικά **για το μοντέλο των νεκρών και βαριά τραυματιών (KSI)**, λόγω του μικρού αριθμού περιστατικών σε ορισμένες υποκατηγορίες, εφαρμόστηκαν **δύο επιπλέον συγχωνεύσεις** πριν από τον ορισμό των κατηγοριών αναφοράς και την εκτίμηση του μοντέλου. Η ηλικιακή ομάδα 0-14 περιλάμβανε μόλις 5 νεκρούς και βαριά τραυματίες σε 38 εγγραφές και συνενώθηκε με την ομάδα 15-18 στην κατηγορία ≤18 (334 εγγραφές). Αντίστοιχα, η κατηγορία χαμηλά θετικός (Low positive) του ελέγχου αλκοόλ περιλάμβανε μόλις 11 νεκρούς και βαριά τραυματίες σε 128 εγγραφές και συνενώθηκε με την κατηγορία υψηλά θετικός (High positive) στην κατηγορία θετικός (Positive), με 275 εγγραφές. Οι συγχωνεύσεις περιορίζουν την αστάθεια που μπορεί να προκύψει από πολύ αραιές κατηγορίες, διατηρώντας παράλληλα το ουσιαστικό περιεχόμενο των μεταβλητών. Στο μοντέλο των **ελαφρά τραυματιών οι αρχικές κατηγορίες διατηρήθηκαν,** καθώς όλες περιλαμβάνουν επαρκή αριθμό περιστατικών.

Πριν από τη μοντελοποίηση πραγματοποιήθηκε **περιγραφική στατιστική ανάλυση σε δύο στάδια**. Στο πρώτο στάδιο, για κάθε μεταβλητή υπολογίστηκαν οι απόλυτες και οι σχετικές συχνότητες των υποκατηγοριών της (**μονοδιάστατοι πίνακες συχνοτήτων**), ώστε να αποτυπωθεί η κατανομή των παθόντων και να εντοπιστούν υποκατηγορίες με πολύ μικρό αριθμό παρατηρήσεων. **Στο δεύτερο στάδιο**, **κάθε μεταβλητή εξετάστηκε σε συνδυασμό με τη σοβαρότητα τραυματισμού** (διδιάστατοι πίνακες συχνοτήτων), δηλαδή υπολογίστηκε για κάθε υποκατηγορία ο αριθμός και το ποσοστό των νεκρών, των βαριά τραυματιών και των ελαφρά τραυματιών. Τα αποτελέσματα των δύο σταδίων αξιοποιήθηκαν για την τεκμηρίωση των ομαδοποιήσεων και για τον εντοπισμό των αραιών υποκατηγοριών.



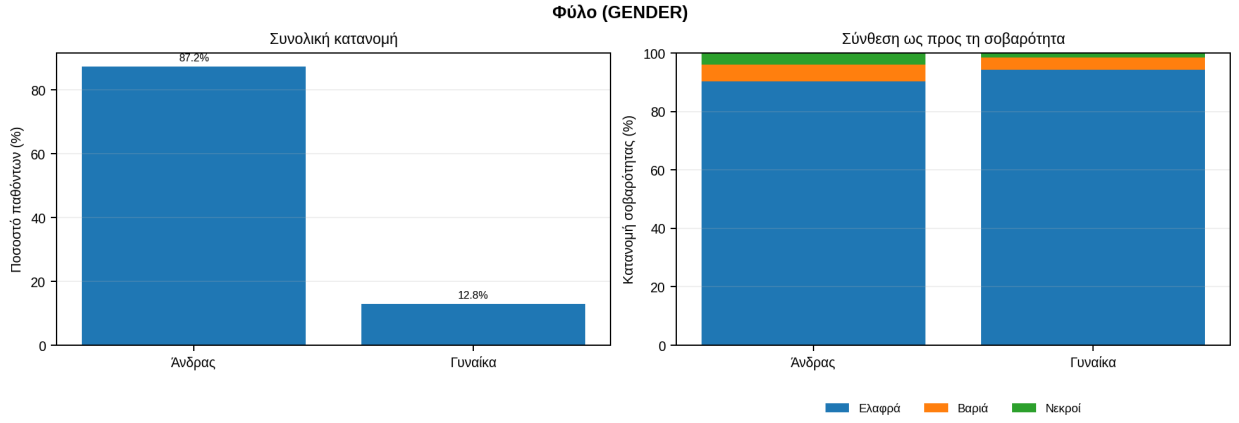
Διάγραμμα 4.1: Κατανομή και σύνθεση σοβαρότητας ανά μήνα.

Ο **Ιούλιος** συγκεντρώνει το **μεγαλύτερο ποσοστό παθόντων** (11,1%), ενώ ο **Φεβρουάριος το μικρότερο** (5,8%). Ως προς τη σοβαρότητα, υψηλότερο ποσοστό KSI εμφανίζεται τον Σεπτέμβριο (13,0%) και τον Ιούλιο (12,2%), χωρίς αυτό να τεκμηριώνει από μόνο του εποχική αιτιώδη επίδραση.



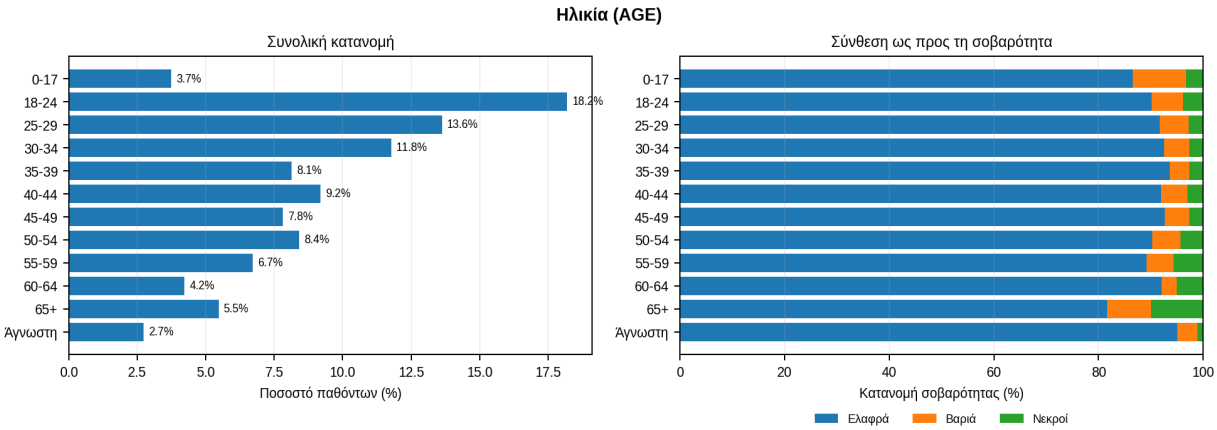
Διάγραμμα 4.2: Κατανομή και σύνθεση σοβαρότητας ανά τύπο χρήστη.

Οι οδηγοί αποτελούν το 88,8% των παθόντων και οι επιβάτες το 11,2%. Η αναλογία KSI είναι υψηλότερη στους οδηγούς (9,4%) από ό,τι στους επιβάτες (7,7%).



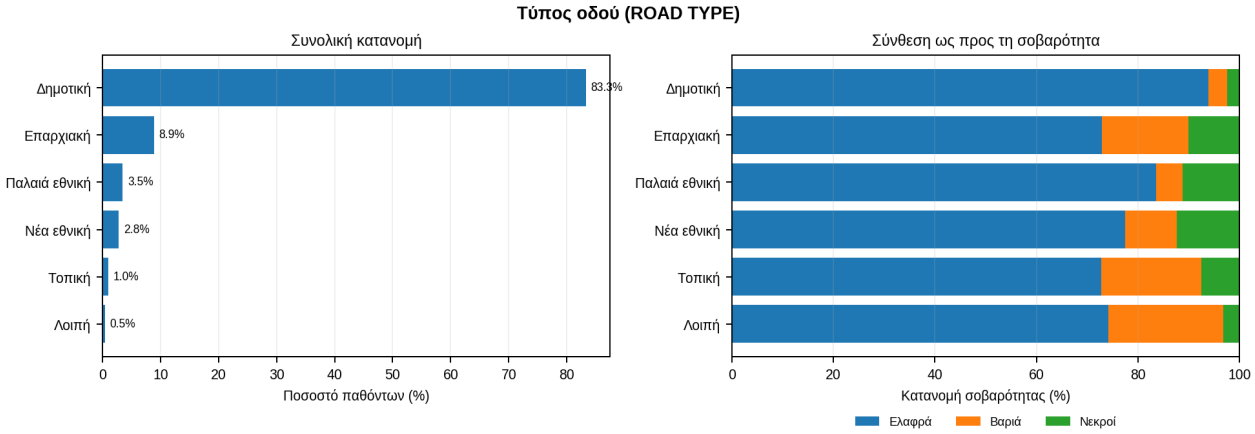
Διάγραμμα 4.3: Κατανομή και σύνθεση σοβαρότητας ανά φύλο.

Οι άνδρες αποτελούν το 87,2% των παθόντων. Η αναλογία KSI είναι 9,7% στους άνδρες και 5,8% στις γυναίκες, διαφορά που εξετάζεται στη συνέχεια πολυμεταβλητά.



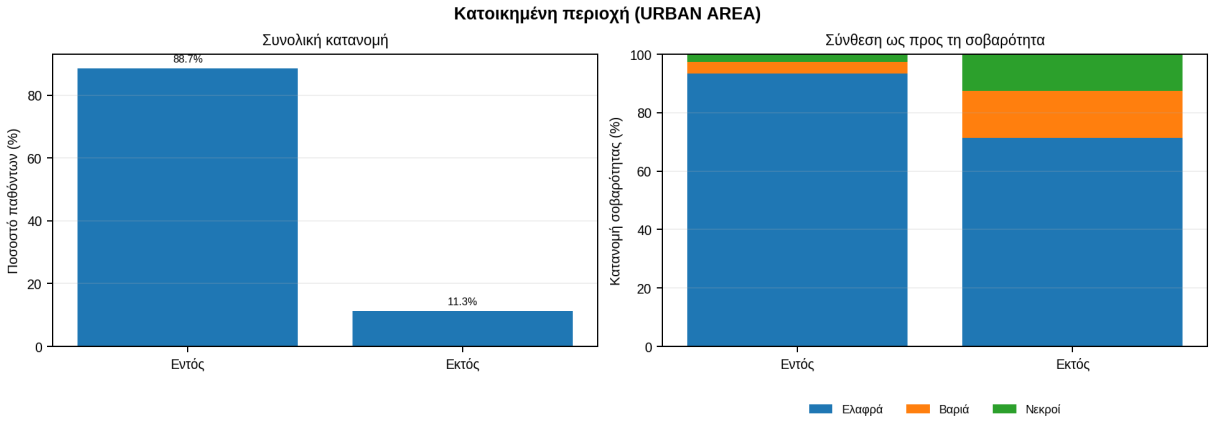
Διάγραμμα 4.4: Κατανομή και σύνθεση σοβαρότητας ανά ηλικιακή ομάδα.

Το μεγαλύτερο μέρος των παθόντων συγκεντρώνεται **στις νεότερες και μέσες ηλικίες.** Η **ομάδα 65+** παρουσιάζει την υψηλότερη αναλογία KSI (**18,3%**), ενώ αυξημένη αναλογία παρατηρείται και στην ομάδα 0-17 (13,4%).



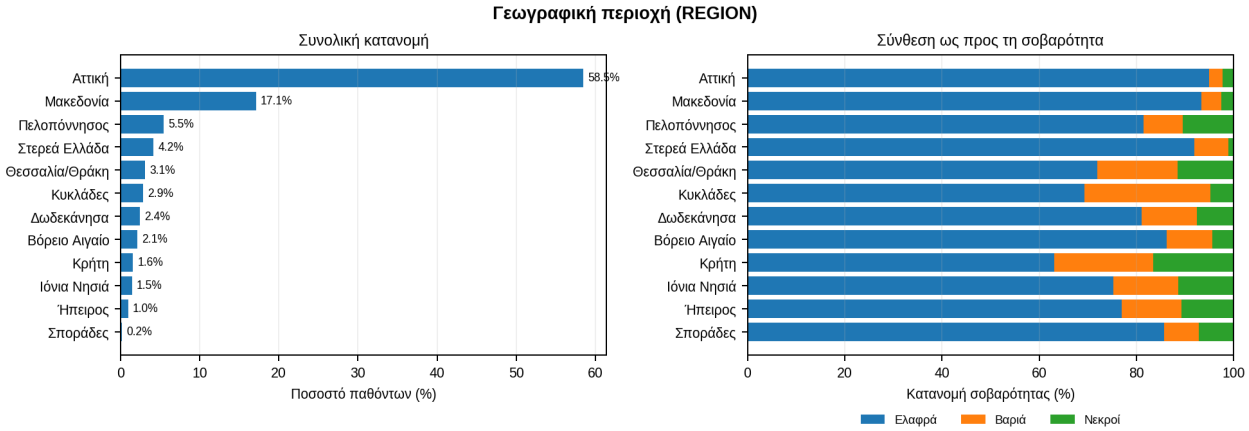
Διάγραμμα 4.5: Κατανομή και σύνθεση σοβαρότητας ανά τύπο οδού.

Η πλειονότητα των παθόντων καταγράφεται σε **δημοτικές οδούς,** λόγω και της μεγάλης συμμετοχής τους στο σύνολο των καταγραφών. Οι λιγότερο συχνές κατηγορίες οδών εμφανίζουν διαφοροποιήσεις στη σύνθεση σοβαρότητας, οι οποίες αξιολογούνται μετά την ομαδοποίηση της μεταβλητής.



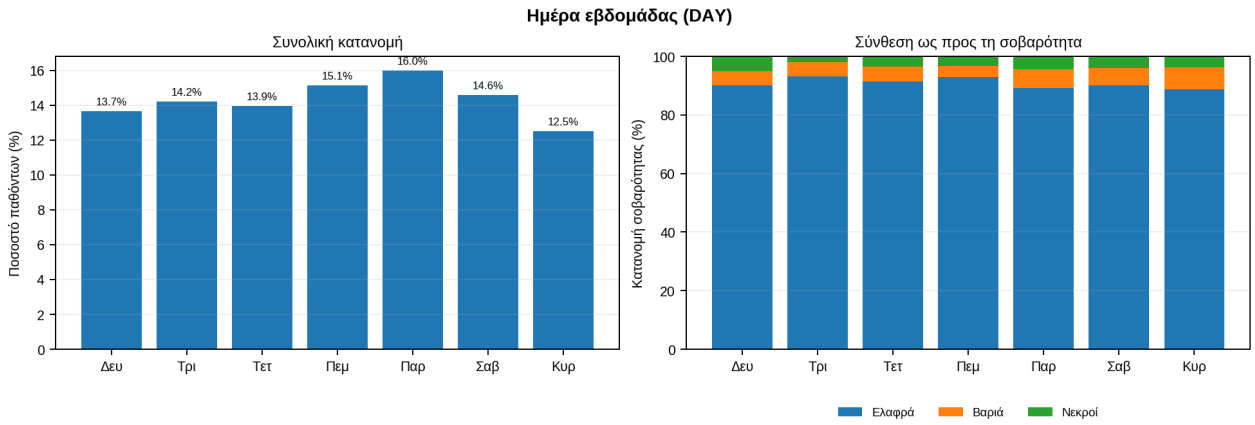
Διάγραμμα 4.6: Κατανομή και σύνθεση σοβαρότητας ανά κατοικημένη περιοχή.

Το 88,7% των παθόντων καταγράφεται εντός κατοικημένων περιοχών. Ωστόσο, **εκτός** κατοικημένης περιοχής **η αναλογία KSI ανέρχεται σε 28,7%,** έναντι 6,7% εντός, εύρημα που υποδηλώνει **σαφή διαφοροποίηση της σοβαρότητας.**



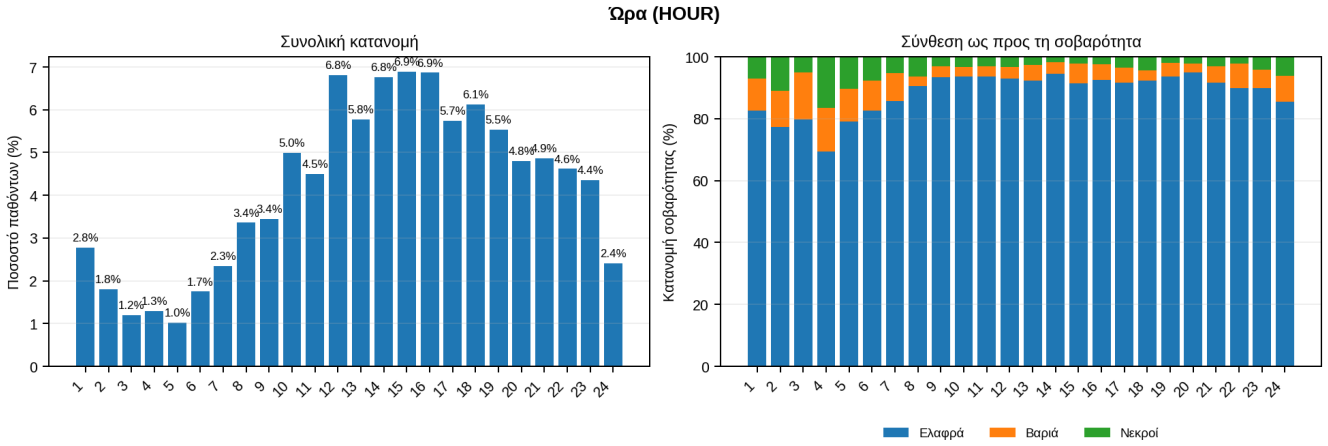
Διάγραμμα 4.7: Κατανομή και σύνθεση σοβαρότητας ανά γεωγραφική περιοχή.

Η Αττική συγκεντρώνει το 58,5% των παθόντων και η Μακεδονία το 17,1%. Οι αναλογίες **KSI διαφέρουν μεταξύ περιοχών**, όμως οι μικρές συχνότητες σε ορισμένες περιοχές απαιτούν προσεκτική ερμηνεία και αιτιολογούν την εξέταση τυχαίας γεωγραφικής επίδρασης.



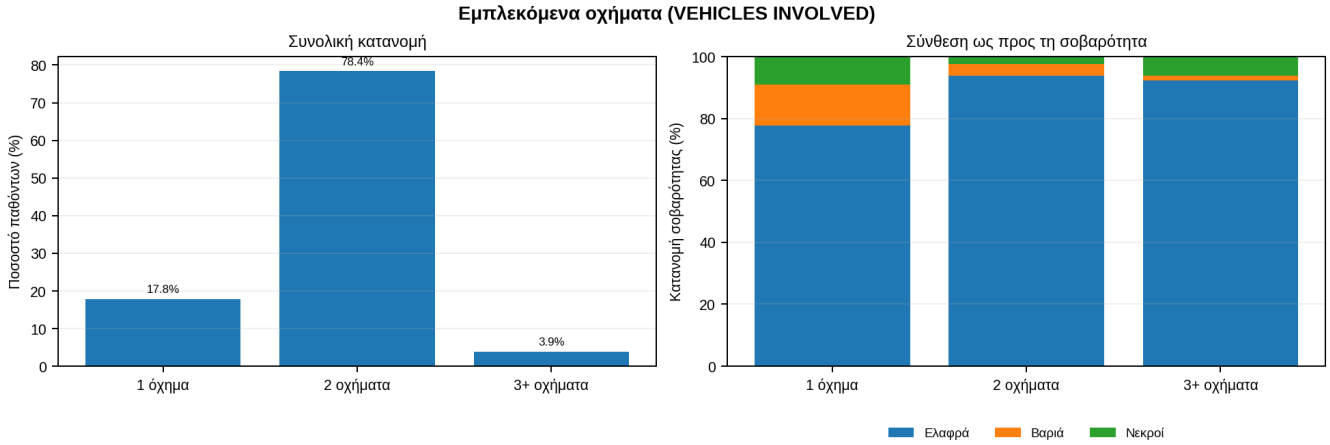
Διάγραμμα 4.8: Κατανομή και σύνθεση σοβαρότητας ανά ημέρα εβδομάδας.

Η κατανομή των παθόντων είναι **σχετικά ομοιόμορφη μέσα στην εβδομάδα**, με **υψηλότερη** συχνότητα την **Παρασκευή**. Οι διαφοροποιήσεις της σοβαρότητας ανά ημέρα είναι μικρότερες από εκείνες που παρατηρούνται σε άλλες μεταβλητές.



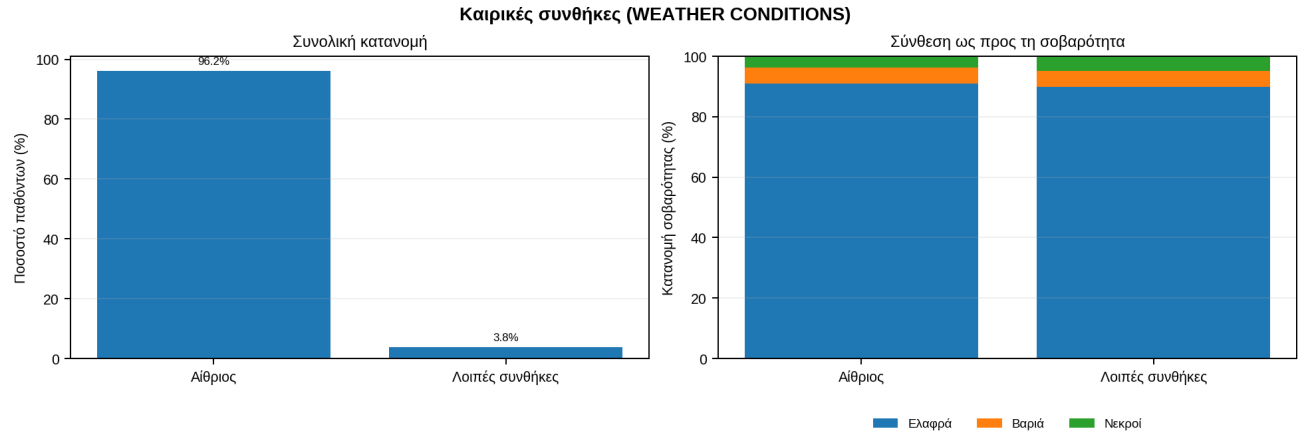
Διάγραμμα 4.9: Κατανομή και σύνθεση σοβαρότητας ανά ώρα.

Οι **περισσότερες καταγραφές** εμφανίζονται από το **μεσημέρι έως το απόγευμα.** Αντίθετα, τις **μεταμεσονύκτιες και πρώτες πρωινές ώρες** οι καταγραφές είναι λιγότερες, αλλά **η αναλογία** KSI είναι υψηλότερη, στοιχείο που υποστηρίζει την ομαδοποίηση της ώρας σε ημέρα και νύχτα.



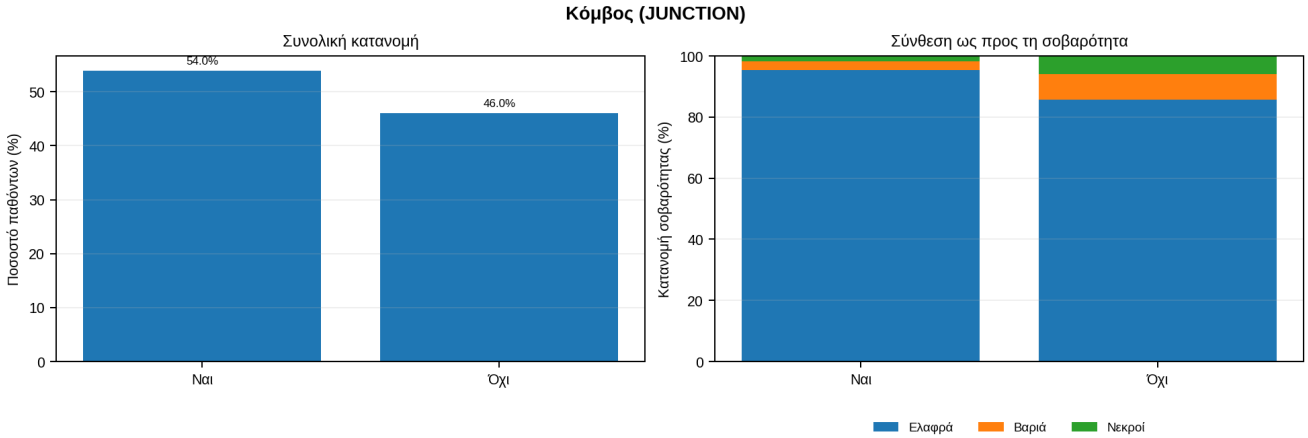
Διάγραμμα 4.10: Κατανομή και σύνθεση σοβαρότητας ανά αριθμό εμπλεκόμενων οχημάτων.

Τα ατυχήματα με δύο εμπλεκόμενα οχήματα αντιστοιχούν στο 78,4% των παθόντων. Στις καταγραφές με ένα όχημα η αναλογία KSI είναι σημαντικά υψηλότερη (22,2%) από ό,τι στις καταγραφές με δύο οχήματα (6,3%).



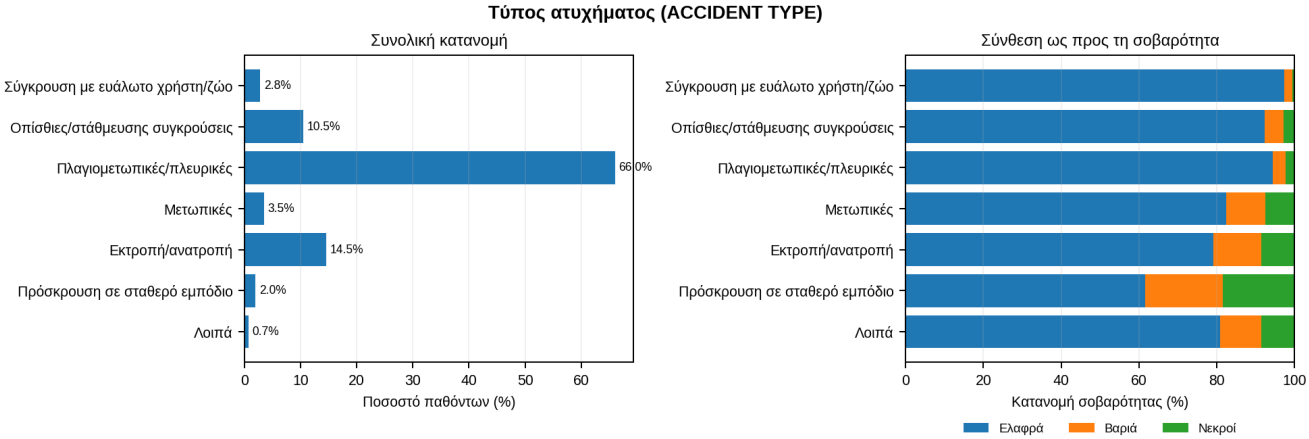
Διάγραμμα 4.11: Κατανομή και σύνθεση σοβαρότητας ανά καιρικές συνθήκες.

Σχεδόν όλες οι καταγραφές αφορούν αίθριο καιρό. Οι λοιπές καιρικές συνθήκες έχουν μικρή συχνότητα, γεγονός που περιορίζει την ασφαλή εξαγωγή συμπερασμάτων μόνο από την περιγραφική σύγκριση.



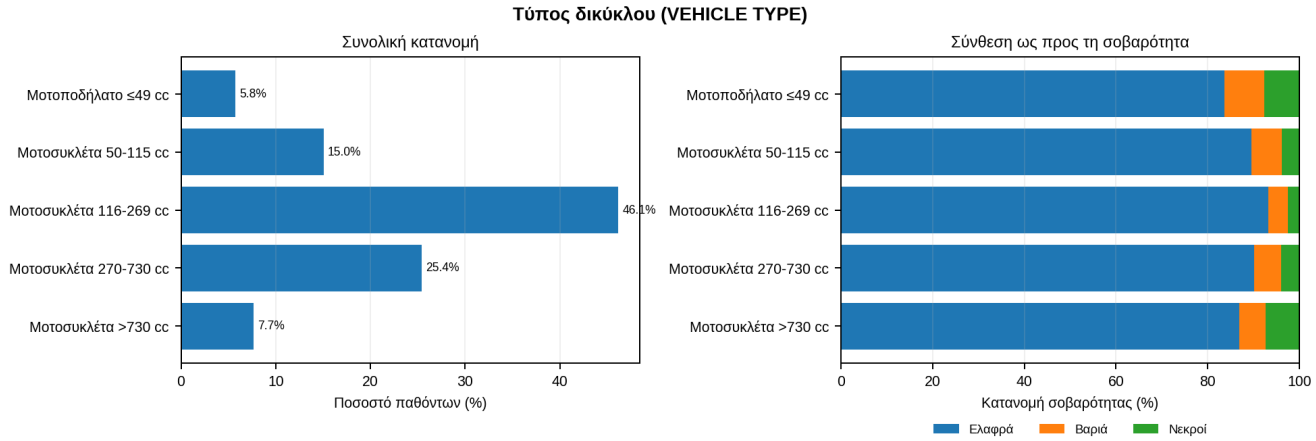
Διάγραμμα 4.12: Κατανομή και σύνθεση σοβαρότητας ανά ύπαρξη κόμβου.

Οι καταγραφές κατανέμονται σχετικά ισόρροπα ως προς την ύπαρξη κόμβου. Η αναλογία **KSI είναι υψηλότερη εκτός κόμβου** (14,4%) από ό,τι σε κόμβο (4,7%).



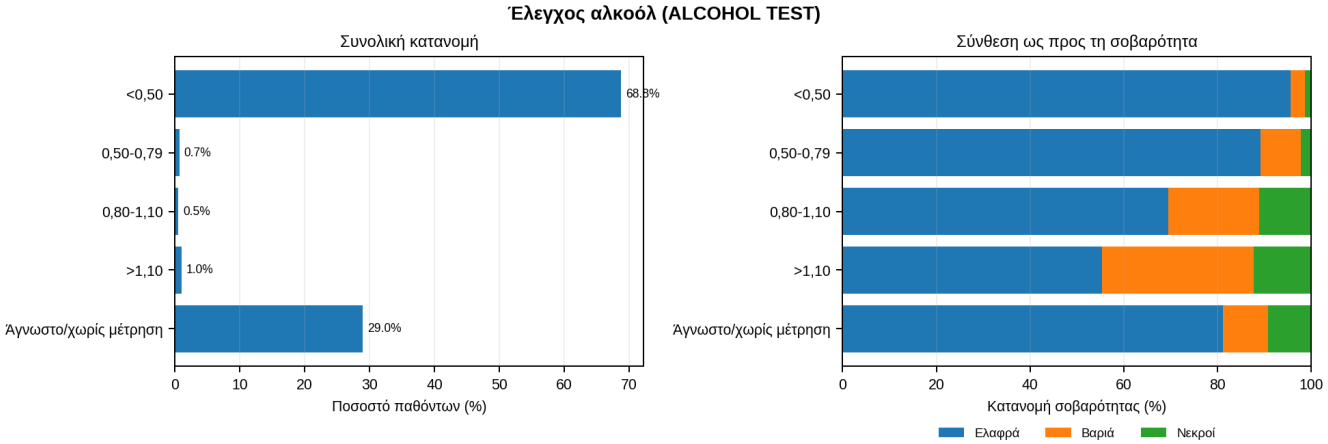
Διάγραμμα 4.13: Κατανομή και σύνθεση σοβαρότητας ανά τύπο ατυχήματος.

Οι πλαγιομετωπικές και πλευρικές συγκρούσεις αποτελούν τη συχνότερη κατηγορία (66,0%). Υψηλές αναλογίες KSI εμφανίζονται στις **προσκρούσεις σε σταθερό εμπόδιο** (**38,5%**), στις εκτροπές/ανατροπές (20,9%) και στις μετωπικές συγκρούσεις (17,5%).



Διάγραμμα 4.14: Κατανομή και σύνθεση σοβαρότητας ανά κατηγορία δικύκλου.

Η κατηγορία 116-269 cc συγκεντρώνει το 46,1% των παθόντων. Μεγαλύτερη αναλογία KSI εμφανίζεται στα μοτοποδήλατα έως 49 cc (16,4%) και στις μοτοσυκλέτες άνω των 730 cc (13,1%).



Διάγραμμα 4.15: Κατανομή και σύνθεση σοβαρότητας ανά αποτέλεσμα ελέγχου αλκοόλ.

Η πλειονότητα των καταγραφών βρίσκεται κάτω από 0,50, ενώ 29,0% ανήκει στην κατηγορία άγνωστο/χωρίς διαθέσιμη μέτρηση. Η αναλογία **KSI αυξάνεται στις υψηλότερες καταγεγραμμένες τιμές αλκοόλ,** αλλά η κατηγορία άγνωστο/χωρίς μέτρηση πρέπει να ερμηνεύεται ως χαρακτηριστικό της διαδικασίας καταγραφής και όχι ως επίπεδο έκθεσης.

4.3.2 Σύνθεση αλγορίθμου

Για την προετοιμασία των δεδομένων και την ανάπτυξη των μοντέλων γράφτηκαν στη γλώσσα προγραμματισμού R **δύο ξεχωριστά προγράμματα (scripts)**: **ένα για τη**

βάση των νεκρών και βαριά τραυματιών και ένα για τη βάση των ελαφρά τραυματιών. Οι δύο βάσεις είχαν ήδη διαμορφωθεί στο Microsoft Excel μετά την ομαδοποίηση των αρχικών καταγραφών. Επομένως, η R δεν χρησιμοποιήθηκε για τη δημιουργία των ομαδοποιημένων γραμμών, αλλά για την εισαγωγή, τον έλεγχο, την κωδικοποίηση και την προετοιμασία των δεδομένων πριν από τη **στατιστική ανάλυση**. Τα βήματα που ακολουθήθηκαν, με τη σειρά που εκτελούνται στον κώδικα, είναι τα εξής:

Βήμα 1: Ρύθμιση του περιβάλλοντος εργασίας. Στην αρχή κάθε προγράμματος ελέγχεται αν είναι εγκατεστημένα όλα τα απαραίτητα πακέτα. Στη συνέχεια φορτώνονται τα πακέτα readxl (εισαγωγή αρχείων Excel), dplyr (επεξεργασία δεδομένων) και glmmTMB (Brooks et al., 2017) για τα μικτά μοντέλα, performance, writexl (εξαγωγή αποτελεσμάτων σε Excel) και ggplot2 (διαγράμματα). Στο πρόγραμμα των **ελαφρά τραυματιών** φορτώνεται επιπλέον το πακέτο MASS για την Αρνητική Διωνυμική παλινδρόμηση (Venables & Ripley, 2002), ενώ στο πρόγραμμα των νεκρών και βαριά τραυματιών το πακέτο **DHARMa** για τους διαγνωστικούς ελέγχους (Hartig, 2022). Τέλος, ορίζονται το αρχείο δεδομένων και ο φάκελος στον οποίο αποθηκεύονται τα αποτελέσματα.

Βήμα 2: Εισαγωγή των δεδομένων. Με τη συνάρτηση read_excel διαβάζεται το αντίστοιχο φύλλο του αρχείου Excel. Κάθε γραμμή του φύλλου αντιστοιχεί σε έναν μοναδικό συνδυασμό χαρακτηριστικών, όπως περιγράφηκε στην Ενότητα 4.3.1.

Βήμα 3: Τυποποίηση των ονομάτων των στηλών. Με μια βοηθητική συνάρτηση αφαιρούνται τα κενά στην αρχή και στο τέλος των ονομάτων, οι μη αλφαριθμητικοί χαρακτήρες αντικαθίστανται με κάτω παύλα και όλα τα ονόματα μετατρέπονται σε κεφαλαίους λατινικούς χαρακτήρες. Έτσι οι μεταβλητές καλούνται με ενιαίο τρόπο σε όλο τον κώδικα.

Βήμα 4: Ορισμός των μεταβλητών της ανάλυσης. Ορίζονται ρητά η **εξαρτημένη μεταβλητή**, δηλαδή ο αριθμός παθόντων (CASUALTIES), η μεταβλητή ομαδοποίησης για τον τυχαίο σταθερό όρο, δηλαδή η **γεωγραφική περιοχή** (REGION), και οι 14 υποψήφιες επεξηγηματικές μεταβλητές του Πίνακα 4.3.

Βήμα 5: Μετατροπή του τύπου των μεταβλητών. Ο αριθμός παθόντων μετατρέπεται σε αριθμητική μεταβλητή, ώστε να χρησιμοποιηθεί ως μεταβλητή καταμέτρησης, ενώ η γεωγραφική περιοχή μετατρέπεται σε κατηγορική μεταβλητή (factor).

Βήμα 6: Κωδικοποίηση των επεξηγηματικών μεταβλητών και ορισμός των κατηγοριών αναφοράς. Καθεμία από τις 14 υποψήφιες μεταβλητές μετατρέπεται σε

κατηγορική και ορίζεται ως κατηγορία αναφοράς αυτόματα η συχνότερη κατηγορία της στη συγκεκριμένη βάση. Οι κατηγορίες αναφοράς που προκύπτουν παρουσιάζονται στον Πίνακα 4.4.

Βήμα 7: Έλεγχος πληρότητας. Διατηρούνται μόνο οι γραμμές που έχουν διαθέσιμη τιμή στην **εξαρτημένη μεταβλητή**, στη **γεωγραφική περιοχή** και σε όλες τις υποψήφιες επεξηγηματικές μεταβλητές. Επειδή οι άγνωστες τιμές είχαν ήδη κωδικοποιηθεί ως ξεχωριστές κατηγορίες, στο βήμα αυτό δεν αφαιρέθηκε καμία γραμμή και παρέμειναν και στις δύο βάσεις 4.069 εγγραφές.

Βήμα 8: Ορισμός των εξισώσεων των μοντέλων. Δημιουργούνται δύο εξισώσεις: η πλήρης εξίσωση σταθερών επιδράσεων, με τις 14 υποψήφιες μεταβλητές, και η αντίστοιχη εξίσωση μικτών επιδράσεων, η οποία περιλαμβάνει επιπλέον τυχαίο σταθερό όρο για τη γεωγραφική περιοχή.

Βήμα 9: Περιγραφική εξέταση της εξαρτημένης μεταβλητής. Υπολογίζονται το μέγεθος του δείγματος, το πλήθος και το ποσοστό των μηδενικών τιμών, η μέση τιμή, η διακύμανση και ο λόγος διακύμανσης προς μέση τιμή του αριθμού παθόντων, ενώ με το πακέτο ggplot2 δημιουργείται και αποθηκεύεται το ιστόγραμμα της κατανομής του. Τα μεγέθη αυτά δίνουν την πρώτη ένδειξη για την ύπαρξη ή όχι υπερδιασποράς.

Με την ολοκλήρωση των παραπάνω βημάτων **οι δύο βάσεις είναι έτοιμες για τη στατιστική ανάλυση.** Τα επόμενα στάδια του κώδικα, τα οποία παρουσιάζονται αναλυτικά στο Κεφάλαιο 5, είναι με τη σειρά: ο έλεγχος υπερδιασποράς με πλήρες μοντέλο Poisson σταθερών επιδράσεων και η επιλογή της κατανομής, η εκτίμηση του πλήρους μοντέλου σταθερών επιδράσεων και του πλήρους μικτού μοντέλου, ο **έλεγχος της χρησιμότητας της γεωγραφικής περιοχής** με τον έλεγχο λόγου πιθανοφάνειας και το κριτήριο AIC (Ενότητα 4.3.3), **η επιλογή των μεταβλητών** με προς τα πίσω απαλοιφή, **η εκτίμηση του τελικού μοντέλου**, οι διαγνωστικοί έλεγχοι και, τέλος, η αποθήκευση των αποτελεσμάτων μαζί με τα σχετικά διαγράμματα.

4.3.3

Η **γεωγραφική περιοχή περιλαμβάνει 12 κατηγορίες**, οι οποίες προέκυψαν από την ο**μαδοποίηση των 73 Περιφερειακών Ενοτήτων** της αρχικής βάσης. Στον Πίνακα 4.5 παρουσιάζονται για κάθε περιοχή οι Περιφερειακές Ενότητες που περιλαμβάνει, καθώς και ο αριθμός των εγγραφών μοντελοποίησης και των παθόντων που αντιστοιχούν σε αυτήν.

Πίνακας 4.5: Γεωγραφικές περιοχές και Περιφερειακές Ενότητες που περιλαμβάνουν

Γεωγραφική περιοχή	Περιφερειακές Ενότητες	Εγγραφές (n, %)	Παθόντες (n, %)
Αττική	Ανατολική Αττική, Βόρειος Τομέας Αθηνών, Δυτική Αττική, Δυτικός Τομέας Αθηνών, Κεντρικός Τομέας Αθηνών, Νήσων, Νότιος Τομέας Αθηνών, Πειραιάς	1.698 (41,7%)	3.851 (58,5%)
Μακεδονία	Γρεβενά, Δράμα, Ημαθία, Θάσος, Θεσσαλονίκη, Καβάλα, Καστοριά, Κιλκίς, Κοζάνη, Πέλλα, Πιερία, Σέρρες, Φλώρινα, Χαλκιδική	855 (21,0%)	1.127 (17,1%)
Πελοπόννησος	Αιτωλοακαρνανία, Αργολίδα, Αρκαδία, Αχαΐα, Ηλεία, Κορινθία, Λακωνία, Μεσσηνία	336 (8,3%)	361 (5,5%)
Στερεά Ελλάδα	Βοιωτία, Εύβοια, Ευρυτανία, Φθιώτιδα, Φωκίδα	251 (6,2%)	274 (4,2%)
Θεσσαλία/Θράκη	Έβρος, Καρδίτσα, Λάρισα, Μαγνησία, Ξάνθη, Ροδόπη, Τρίκαλα	200 (4,9%)	207 (3,1%)
Κυκλάδες	Άνδρος, Θήρα, Κέα, Κύθνος, Μήλος, Μύκονος, Νάξος, Πάρος, Σύρος, Τήνος	174 (4,3%)	189 (2,9%)
Δωδεκάνησα	Κάλυμνος, Κάρπαθος, Κως, Ρόδος	153 (3,8%)	159 (2,4%)
Βόρειο Αιγαίο	Ικαρία, Λέσβος, Λήμνος, Σάμος, Χίος	135 (3,3%)	138 (2,1%)
Κρήτη	Ηράκλειο, Λασίθι, Ρέθυμνο, Χανιά	102 (2,5%)	103 (1,6%)
Ιόνια Νησιά	Ζάκυνθος, Κέρκυρα, Κεφαλληνία, Λευκάδα	89 (2,2%)	97 (1,5%)
Ήπειρος	Άρτα, Θεσπρωτία, Ιωάννινα, Πρέβεζα	62 (1,5%)	65 (1,0%)
Σποράδες	Σποράδες	14 (0,3%)	14 (0,2%)
Σύνολο	**73 Περιφερειακές Ενότητες**	**4.069 (100,0%)**	**6.585 (100,0%)**

Η μεγαλύτερη συγκέντρωση εγγραφών μοντελοποίησης παρατηρείται στην Αττική με 1.698 εγγραφές, δηλαδή 41,7% του τελικού συνόλου, και **ακολουθεί η Μακεδονία** με 855 εγγραφές ή 21,0%. **Σε επίπεδο παθόντων** η συγκέντρωση είναι ακόμη εντονότερη, καθώς η Αττική αντιστοιχεί σε 3.851 από τους 6.585 παθόντες (58,5%) και η Μακεδονία σε 1.127 (17,1%). Στο άλλο άκρο, **η μικρότερη κατηγορία είναι οι Σποράδες** με μόλις 14 εγγραφές.

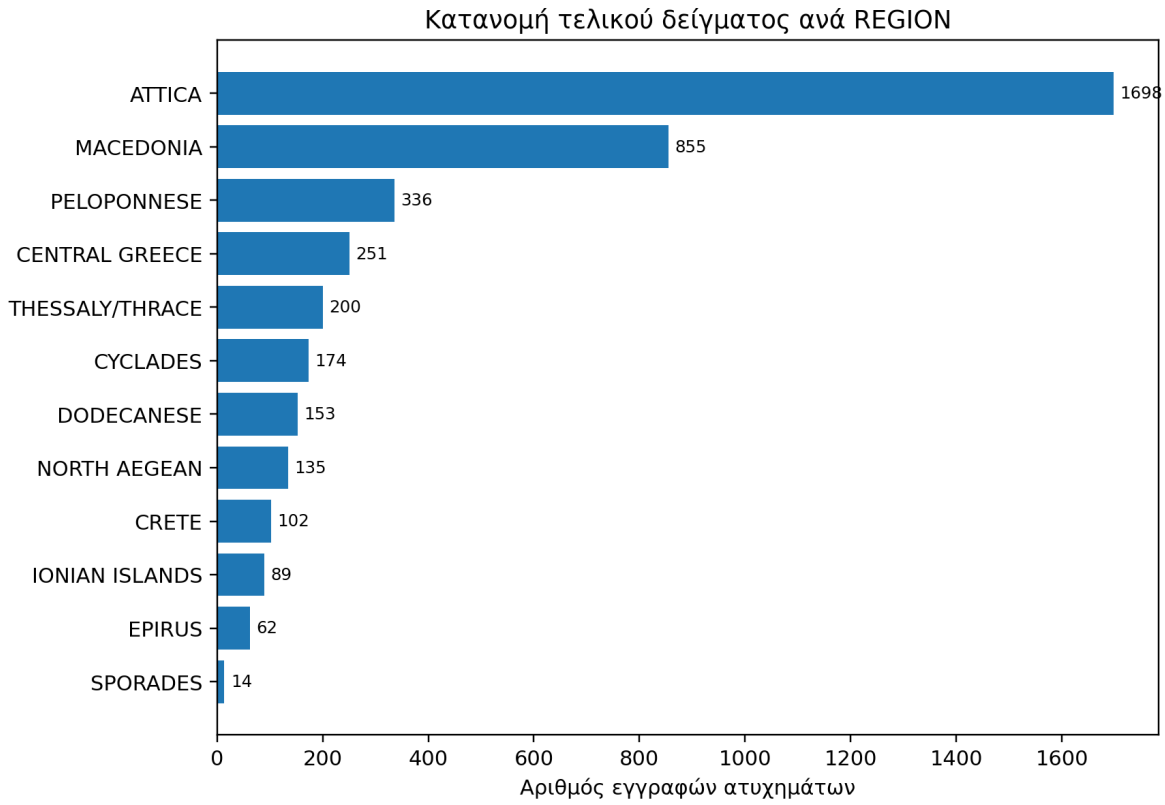
Η ομαδοποίηση των 73 Περιφερειακών Ενοτήτων σε 12 ευρύτερες περιοχές κρίθηκε απαραίτητη **για λόγους στατιστικής επάρκειας**. Σε επίπεδο Περιφερειακής Ενότητας τα δεδομένα θα ήταν ιδιαίτερα αραιά: οι 604 νεκροί και βαριά τραυματίες αντιστοιχούν κατά μέσο όρο σε λιγότερους από 9 παθόντες ανά ενότητα, ενώ πολλές ενότητες θα περιλάμβαναν ελάχιστες εγγραφές. Ενδεικτικά, οι τέσσερις Περιφερειακές Ενότητες της Ηπείρου συγκεντρώνουν συνολικά μόλις 62 εγγραφές και 15 νεκρούς και βαριά τραυματίες, δηλαδή λιγότερους από 4 ανά ενότητα, ενώ η Περιφερειακή Ενότητα των Σποράδων περιλαμβάνει μόλις 14 εγγραφές. Με τόσο μικρές ομάδες, η διακύμανση

του τυχαίου σταθερού όρου θα εκτιμούνταν με μεγάλη αβεβαιότητα, οι επιμέρους εκτιμήσεις ανά ομάδα θα βασίζονταν σε ελάχιστα περιστατικά και θα αυξανόταν ο κίνδυνος προβλημάτων σύγκλισης του μικτού μοντέλου, ιδιαίτερα στο μοντέλο των νεκρών και βαριά τραυματιών. Για τον λόγο αυτό, οι Περιφερειακές Ενότητες συνενώθηκαν **με κριτήριο τη γεωγραφική τους συνάφεια**, ώστε κάθε περιοχή να αποτελεί ενιαίο και ερμηνεύσιμο γεωγραφικό σύνολο (π.χ. Κρήτη, Κυκλάδες, Δωδεκάνησα) και να διαθέτει επαρκή αριθμό εγγραφών για την εκτίμηση της τυχαίας επίδρασης.

Η κατανομή των εγγραφών δεν είναι ομοιόμορφη μεταξύ των περιοχών, γεγονός που καθιστά σκόπιμη τη διερεύνηση πιθανής περιφερειακής ετερογένειας. Η γεωγραφική περιοχή δεν εισήχθη ως σταθερή επεξηγηματική μεταβλητή, καθώς κάτι τέτοιο θα απαιτούσε 11 επιπλέον παραμέτρους, ορισμένες από τις οποίες θα εκτιμώνταν από πολύ λίγες εγγραφές. Αντί αυτού, εισήχθη ως **τυχαίος σταθερός όρος (random intercept)**, ο οποίος επιτρέπει σε **κάθε περιοχή** να έχει **διαφορετικό επίπεδο αναμενόμενου αριθμού παθόντων**, εκτιμώντας μόνο μία επιπλέον παράμετρο: τη διακύμανση μεταξύ των περιοχών.

Στον αλγόριθμο, η χρησιμότητα της γεωγραφικής περιοχής ελέγχεται πριν από την επιλογή των επεξηγηματικών μεταβλητών. Αρχικά εκτιμώνται δύο μοντέλα με όλες τις 14 υποψήφιες μεταβλητές: ένα μοντέλο **σταθερών επιδράσεων** (fixed effects), **χωρίς τη γεωγραφική περιοχή**, και ένα μικτό μοντέλο (mixed effects), που περιλαμβάνει επιπλέον τον τυχαίο σταθερό όρο ανά περιοχή. Τα δύο μοντέλα συγκρίνονται με τον έλεγχο λόγου πιθανοφανειών (Likelihood Ratio Test, LRT). Επειδή υπό τη μηδενική υπόθεση η διακύμανση μεταξύ των περιοχών είναι μηδενική, δηλαδή βρίσκεται στο όριο των επιτρεπτών τιμών της, αφού μια διακύμανση δεν μπορεί να είναι αρνητική, η τιμή p υπολογίζεται με διόρθωση ορίου, ως το μισό της τιμής p της κατανομής χ^2 με έναν βαθμό ελευθερίας. Παράλληλα, συγκρίνεται το κριτήριο πληροφορίας του Akaike (AIC) των δύο μοντέλων.

Η **γεωγραφική περιοχή διατηρείται μόνο** εφόσον ισχύουν ταυτόχρονα **$p < 0,05$ και χαμηλότερο AIC** για το μικτό μοντέλο, διαφορετικά η εκτέλεση του αλγορίθμου διακόπτεται. Εφόσον η τυχαία επίδραση κριθεί χρήσιμη, όλα τα επόμενα βήματα, δηλαδή η επιλογή των μεταβλητών και η εκτίμηση του τελικού μοντέλου, πραγματοποιούνται στο μικτό μοντέλο. Τα αποτελέσματα του ελέγχου για τα δύο μοντέλα παρουσιάζονται στο Κεφάλαιο 5.



Διάγραμμα 4.16: Κατανομή των 4.069 εγγραφών μοντελοποίησης ανά γεωγραφική περιοχή (REGION)

4.4 Σύνοψη και παρατηρήσεις

Η περιγραφική ανάλυση των δεδομένων ανέδειξε **σημαντικές διαφοροποιήσεις** τόσο στην κατανομή των παθόντων όσο και στη σύνθεση της σοβαρότητας μεταξύ των κατηγοριών των μεταβλητών. Συγκεκριμένα παρατηρήθηκε ότι:

- Η **πλειονότητα των παθόντων** (90,83%) **τραυματίστηκε ελαφρά**, ενώ το 9,17% έχασε τη ζωή του ή τραυματίστηκε σοβαρά (KSI).
- Οι **περισσότεροι παθόντες ήταν οδηγοί** (88,8%) και **άνδρες** (87,2%). Το ποσοστό των νεκρών και σοβαρά τραυματιών ήταν υψηλότερο στους άνδρες (9,7%) σε σχέση με τις γυναίκες (5,8%), καθώς και στους οδηγούς (9,4%) σε σχέση με τους επιβάτες (7,7%).
- **Οι αναβάτες ηλικίας 65 ετών και άνω παρουσίασαν το υψηλότερο ποσοστό νεκρών και σοβαρά τραυματιών** (18,3%), παρότι τα περισσότερα θύματα ανήκαν σε μικρότερες ηλικίες.
- Τα **περισσότερα θύματα** (88,7%) καταγράφηκαν **εντός κατοικημένων περιοχών**. Ωστόσο, τα ατυχήματα **εκτός κατοικημένων περιοχών ήταν σοβαρότερα**, καθώς το 28,7% των παθόντων έχασε τη ζωή του ή τραυματίστηκε σοβαρά, έναντι 6,7% εντός κατοικημένων περιοχών.

- Τα **ατυχήματα με ένα μόνο όχημα παρουσίασαν υψηλότερο ποσοστό** νεκρών και σοβαρά τραυματιών (22,2%), σε σύγκριση με εκείνα στα οποία εμπλέκονταν δύο οχήματα (6,3%). Ωστόσο, τ**α ατυχήματα με δύο οχήματα** συγκέντρωσαν τους **περισσότερους παθόντες** (78,4%).
- Τα **ατυχήματα εκτός κόμβων και διασταυρώσεων παρουσίασαν μεγαλύτερη σοβαρότητα**, με ποσοστό νεκρών και σοβαρά τραυματιών 14,4%, έναντι 4,7% στα ατυχήματα εντός κόμβων.
- Οι **πλαγιομετωπικές και πλευρικές συγκρούσεις ήταν οι συχνότερες** (66,0%). Ωστόσο, τα υψηλότερα ποσοστά νεκρών και σοβαρά τραυματιών καταγράφηκαν στις **προσκρούσεις σε σταθερά εμπόδια** (38,5%), στις **εκτροπές και ανατροπές** (20,9%) και στις **μετωπικές συγκρούσεις** (17,5%).
- Τα **περισσότερα θύματα** (46,1%) καταγράφηκαν σε **δίκυκλα κυβισμού 116-269 cc.** Ωστόσο, **μεγαλύτερη σοβαρότητα** παρουσίασαν τα **ατυχήματα με μοτοποδήλατα έως 49 cc** (16,4%) και μοτοσυκλέτες άνω των 730 cc (13,1%).
- Κατά τη διάρκεια της **νύχτας και τις πρώτες πρωινές ώρες, τα θύματα ήταν λιγότερα**, αλλά **τα ποσοστά σοβαρών και θανατηφόρων τραυματισμών υψηλότερα**. Ο Ιούλιος συγκέντρωσε τα περισσότερα θύματα (11,1%) και ο Φεβρουάριος τα λιγότερα (5,8%).
- Στο 29,0% των καταγραφών δεν υπήρχαν διαθέσιμα στοιχεία για το αλκοόλ. Στις περιπτώσεις όπου υπήρχε μέτρηση, παρατηρήθηκε ότι **όσο υψηλότερη ήταν η συγκέντρωση αλκοόλ, τόσο μεγαλύτερο ήταν το ποσοστό νεκρών και σοβαρά τραυματιών.**
- Η συντριπτική **πλειονότητα** των ατυχημάτων καταγράφηκε **με αίθριο καιρό**, ενώ οι υπόλοιπες καιρικές συνθήκες εμφανίστηκαν πολύ σπανιότερα.
- Οι **περισσότεροι παθόντες** καταγράφηκαν στην **Αττική** (58,5%) και στη **Μακεδονία** (17,1%). Αντίθετα, σε άλλες περιοχές, όπως οι Σποράδες, οι καταγραφές ήταν ελάχιστες.
- Συνολικά, **τα περισσότερα θύματα δεν καταγράφονται απαραίτητα στις συνθήκες όπου τα ατυχήματα έχουν τις σοβαρότερες συνέπειες**.

5. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ - ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφονται τα βήματα που ακολουθήθηκαν κατά την **εφαρμογή της μεθοδολογίας**, η διαδικασία ανάπτυξης των στατιστικών μοντέλων και τα αποτελέσματα της ανάλυσης. Στόχος είναι να **προσδιοριστούν οι παράγοντες που σχετίζονται με τον** αναμενόμενο αριθμό **παθόντων σε δύο κατηγορίες σοβαρότητας τραυματισμού**: τους νεκρούς και βαριά τραυματίες (KSI) και τους ελαφρά τραυματίες. Για τον σκοπό αυτό αναπτύχθηκαν **δύο ανεξάρτητα μοντέλα καταμέτρησης**, με κοινό αρχικό σύνολο 14 υποψήφιων επεξηγηματικών μεταβλητών.

Αρχικά, διερευνάται η κατανομή των δύο εξαρτημένων μεταβλητών και ελέγχεται η ύπαρξη υπερδιασποράς, ώστε να επιλεγεί η κατάλληλη κατανομή. Στη συνέχεια, περιγράφεται η ανάπτυξη των μοντέλων στο λογισμικό R: **τα μοντέλα Poisson και Αρνητικής Διωνυμικής** κατανομής, η σύγκριση μοντέλων σταθερών και μικτών επιδράσεων και η επιλογή των ανεξάρτητων μεταβλητών. Έπειτα, παρουσιάζονται τα αποτελέσματα των τελικών μοντέλων και οι **διαγνωστικοί έλεγχοι**, στη συνέχεια αναλύονται οι γεωγραφικές διαφοροποιήσεις και, τέλος, **συγκρίνονται τα αποτελέσματα των δύο μοντέλων.**

5.2 Διερεύνηση κατανομής δεδομένων

Πριν από την ανάπτυξη των μοντέλων **εξετάστηκε η κατανομή των δύο εξαρτημένων μεταβλητών,** δηλαδή του αριθμού νεκρών και βαριά τραυματιών και του αριθμού ελαφρά τραυματιών **ανά εγγραφή μοντελοποίησης.** Βασικό κριτήριο για την επιλογή της κατανομής αποτελεί **η σχέση μέσης τιμής και διακύμανσης:** η κατανομή Poisson προϋποθέτει ότι οι δύο είναι ίσες (ισοδιασπορά), ενώ όταν η διακύμανση είναι σημαντικά μεγαλύτερη από τη μέση τιμή (υπερδιασπορά) καταλληλότερη είναι η **Αρνητική Διωνυμική κατανομή** (Negative Binomial).

Οι δύο μεταβλητές παρουσιάζουν ουσιαστικά διαφορετική κατανομή (Πίνακας 5.1). Στον **αριθμό νεκρών και βαριά τραυματιών**, η μέση τιμή είναι 0,15 και η διακύμανση 0,13, με λόγο διακύμανσης προς μέση τιμή 0,89. Η εικόνα αυτή **δεν υποδηλώνει** υπερδιασπορά. Αντίθετα, στον αριθμό **ελαφρά τραυματιών** η κατανομή παρουσιάζει έντονη ασυμμετρία προς τα δεξιά: 130 εγγραφές (3,20%) έχουν περισσότερους από 5 ελαφρά τραυματίες και η μέγιστη τιμή φτάνει τους 56. Η μέση τιμή είναι 1,47 και η διακύμανση 8,10, με λόγο διακύμανσης προς μέση τιμή 5,51 γεγονός που αποτελεί **σαφή ένδειξη υπερδιασποράς.**

Πίνακας 5.1: Περιγραφικά στοιχεία των δύο εξαρτημένων μεταβλητών

Εξαρτημένη μεταβλητή	Σύνολο παθόντων	Μέση τιμή	Διακύμανση	Λόγος διακύμανσης/ μέσης τιμής
Νεκροί και βαριά τραυματίες (KSI)	604	0,1484	0,1328	0,8948
Ελαφρά τραυματίες	5.981	1,4699	8,1046	5,5137

Ο λόγος διακύμανσης προς μέση τιμή των ακατέργαστων δεδομένων δεν λαμβάνει υπόψη την επίδραση των επεξηγηματικών μεταβλητών. Για τον λόγο αυτό, **για κάθε εξαρτημένη μεταβλητή** εκτιμήθηκε αρχικά ένα πλήρες μοντέλο Poisson **σταθερών επιδράσεων** με τις 14 υποψήφιες μεταβλητές και υπολογίστηκε ο λόγος διασποράς Pearson (Pearson dispersion ratio), δηλαδή το άθροισμα των τετραγώνων των καταλοίπων Pearson διαιρεμένο με τους βαθμούς ελευθερίας του μοντέλου. Τιμές κοντά στη μονάδα είναι συμβατές με την κατανομή Poisson, ενώ τιμές σημαντικά μεγαλύτερες υποδηλώνουν υπερδιασπορά. Ως κριτήριο υπερδιασποράς ορίστηκε λόγος μεγαλύτερος του 1,5 με στατιστικά σημαντικό έλεγχο ($p < 0,05$).

Στον **αριθμό νεκρών και βαριά τραυματιών** ο λόγος διασποράς ήταν 0,86, χωρίς στατιστική ένδειξη υπερδιασποράς, και επομένως **διατηρήθηκε η κατανομή Poisson**. Η τιμή μικρότερη της μονάδας υποδηλώνει ήπια τάση υποδιασποράς, η οποία εξετάστηκε περαιτέρω στους διαγνωστικούς ελέγχους του τελικού μοντέλου (Ενότητα 5.4.3). Στον αριθμό **ελαφρά τραυματιών** ο λόγος διασποράς ήταν 1,95 και ο έλεγχος υπερδιασποράς ιδιαίτερα σημαντικός ($p < 0,001$). Η υπόθεση ισότητας μέσης τιμής και διακύμανσης δεν ήταν συνεπώς επαρκής και χρησιμοποιήθηκε η **Αρνητική Διωνυμική κατανομή τύπου NB2** (Πίνακας 5.2).

Πίνακας 5.2: Έλεγχος διασποράς και επιλογή κατανομής

Μοντέλο	Λόγος διακύμανσης /μέσης τιμής	Λόγος διασποράς Pearson	Τιμή p	Τελική κατανομή
Νεκροί και βαριά τραυματίες (KSI)	0,8948	0,8557	≈ 1,000	Poisson
Ελαφρά τραυματίες	5,5137	1,9446	< 0,001	Αρνητική Διωνυμική Κατανομή

5.3 Ανάπτυξη μοντέλων παλινδρόμησης στο λογισμικό R

Τα μοντέλα αναπτύχθηκαν στη γλώσσα προγραμματισμού R, με χωριστό κώδικα για κάθε **εξαρτημένη μεταβλητή**. Χρησιμοποιήθηκαν **Γενικευμένα Γραμμικά Μοντέλα** (Generalized Linear Models, GLM) και **Γενικευμένα Γραμμικά Μικτά Μοντέλα**

(Generalized Linear Mixed Models, GLMM) με λογαριθμική συνάρτηση σύνδεσης (log link). Τα μοντέλα **σταθερών επιδράσεων** εκτιμήθηκαν με τις συναρτήσεις glm και glm.nb (πακέτο MASS, Venables & Ripley, 2002), ενώ τα μικτά μοντέλα με το πακέτο glmmTMB (Brooks et al., 2017), το οποίο εκτιμά τις παραμέτρους με τη μέθοδο της μέγιστης πιθανοφάνειας μέσω προσέγγισης Laplace. Για τους διαγνωστικούς ελέγχους χρησιμοποιήθηκε το πακέτο DHARMa (Hartig, 2022).

Και στα δύο μοντέλα η γενική μορφή του μικτού μοντέλου είναι $\log(\mu_{ij}) = \beta_0 + \beta_1 \cdot x_{1ij} + \ldots + \beta_k \cdot x_{kij} + u_j$, όπου μ_{ij} ο αναμενόμενος αριθμός παθόντων της εγγραφής i στη **γεωγραφική περιοχή** j, β οι συντελεστές των επεξηγηματικών μεταβλητών x και u_j ο τυχαίος σταθερός όρος της περιοχής j, ο οποίος ακολουθεί κανονική κατανομή με μέση τιμή μηδέν και διακύμανση σ^2_u. Στο μοντέλο **σταθερών επιδράσεων** ο όρος u_j απουσιάζει. Οι δύο μορφές διαφέρουν ως προς την κατανομή της εξαρτημένης μεταβλητής, όπως περιγράφεται στις Ενότητες 5.3.1 και 5.3.2.

5.3.1 Μοντέλα παλινδρόμησης Poisson

Το μοντέλο **παλινδρόμησης Poisson** υποθέτει ότι η διακύμανση της εξαρτημένης μεταβλητής ισούται με τη μέση τιμή της, δηλαδή $\mathrm{Var}(Y) = \mu$. Με βάση τα αποτελέσματα της Ενότητας 5.2, η κατανομή Poisson επιλέχθηκε για τον αριθμό νεκρών και βαριά τραυματιών, για τον οποίο **δεν διαπιστώθηκε υπερδιασπορά**. Ο κώδικας περιλαμβάνει έλεγχο, ώστε η διαδικασία να συνεχίζεται με την κατανομή Poisson μόνο εφόσον δεν διαπιστώνεται υπερδιασπορά.

Πριν από την εκτίμηση, στην ηλικιακή ομάδα και στον έλεγχο αλκοόλ εφαρμόστηκαν οι συγχωνεύσεις αραιών κατηγοριών που περιγράφηκαν στο Κεφάλαιο 4 (κατηγορίες ≤18 και θετικός). Στη συνέχεια εκτιμήθηκαν ένα πλήρες μοντέλο Poisson **σταθερών επιδράσεων** με τη συνάρτηση glm και ένα πλήρες μικτό μοντέλο Poisson με τυχαίο σταθερό όρο ανά **γεωγραφική περιοχή** με τη συνάρτηση glmmTMB, τα οποία συγκρίνονται στην Ενότητα 5.3.3. Για τον αριθμό ελαφρά τραυματιών, το μοντέλο Poisson χρησιμοποιήθηκε μόνο ως αρχικό εργαλείο για τον έλεγχο υπερδιασποράς και απορρίφθηκε.

5.3.2 Μοντέλα Αρνητικής Διωνυμικής παλινδρόμησης

Η **Αρνητική Διωνυμική κατανομή** (Negative Binomial) τύπου NB2 επιτρέπει στη διακύμανση να υπερβαίνει τη μέση τιμή, μέσω της σχέσης $\mathrm{Var}(Y) = \mu + \mu^2/\theta$, όπου θ η παράμετρος διασποράς. Όσο μικρότερη είναι η τιμή του θ, τόσο μεγαλύτερη είναι η **υπερδιασπορά**, ενώ όσο το θ αυξάνεται η κατανομή προσεγγίζει την Poisson. Η

κατανομή αυτή επιλέχθηκε για τον αριθμό ελαφρά τραυματιών, λόγω της σαφούς υπερδιασποράς που διαπιστώθηκε στην Ενότητα 5.2.

Το πλήρες μοντέλο **σταθερών επιδράσεων** εκτιμήθηκε με τη συνάρτηση glm.nb, με εναλλακτική την εκτίμηση μέσω glmmTMB σε περίπτωση μη σύγκλισης, και το πλήρες μικτό μοντέλο με τη συνάρτηση glmmTMB. Στο **τελικό μοντέλο** η παράμετρος θ εκτιμήθηκε ίση με 4,992, γεγονός που επιβεβαιώνει ότι η διακύμανση υπερβαίνει τη μέση τιμή. Για τον αριθμό νεκρών και βαριά τραυματιών η Αρνητική Διωνυμική δεν εφαρμόστηκε, καθώς δεν υπήρχε υπερδιασπορά που να τη δικαιολογεί.

5.3.3 Σύγκριση μοντέλων σταθερών και μικτών επιδράσεων

Μετά την επιλογή της κατάλληλης κατανομής, για κάθε εξαρτημένη μεταβλητή συγκρίθηκε το πλήρες μοντέλο **σταθερών επιδράσεων** (fixed effects) με το αντίστοιχο πλήρες μικτό μοντέλο (mixed effects), το οποίο περιλαμβάνει **επιπλέον τυχαίο σταθερό όρο** (random intercept) ανά **γεωγραφική περιοχή** (REGION). Η σύγκριση βασίστηκε στον έλεγχο λόγου πιθανοφανειών (Likelihood Ratio Test, LRT) με διόρθωση ορίου, καθώς η μηδενική τιμή της διακύμανσης του τυχαίου όρου βρίσκεται στο όριο του χώρου των παραμέτρων, και παράλληλα στη μεταβολή του κριτηρίου πληροφορίας Akaike (AIC). Πριν από τη συνέχιση της διαδικασίας ελέγχθηκε και η ορθή σύγκλιση του μικτού μοντέλου.

Και στα δύο μοντέλα η προσθήκη της γεωγραφικής περιοχής **βελτίωσε ουσιαστικά την προσαρμογή** (Πίνακας 5.3). Στο μοντέλο των νεκρών και βαριά τραυματιών η στατιστική LRT ήταν 70,00 με $p < 0,001$ και το AIC μειώθηκε από 3.169,34 στο μοντέλο **σταθερών επιδράσεων** σε 3.101,33 στο μικτό μοντέλο. Στο μοντέλο των ελαφρά τραυματιών η στατιστική LRT ήταν 340,49 με $p < 0,001$ και το AIC μειώθηκε από 11.941,14 σε 11.602,65. Συνεπώς, η **γεωγραφική περιοχή διατηρήθηκε ως τυχαίος σταθερός όρος και στα δύο μοντέλα** και όλα τα επόμενα βήματα πραγματοποιήθηκαν στα μικτά μοντέλα.

Πίνακας 5.3: Σύγκριση πλήρων μοντέλων σταθερών και μικτών επιδράσεων

Μοντέλο	AIC σταθερών επιδράσεων	AIC μικτού μοντέλου	Τιμή p	Απόφαση
Νεκροί και βαριά τραυματίες (KSI)	3.169,34	3.101,33	< 0,001	Διατήρηση γεωγραφικής περιοχής
Ελαφρά τραυματίες	11.941,14	11.602,65	< 0,001	Διατήρηση γεωγραφικής περιοχής

5.3.4 Επιλογή ανεξάρτητων μεταβλητών

Η **επιλογή των τελικών ανεξάρτητων μεταβλητών** πραγματοποιήθηκε με τη μέθοδο της προς τα πίσω απαλοιφής (backward elimination), απευθείας στα μικτά μοντέλα και σε επίπεδο ολόκληρης μεταβλητής. Σε κάθε βήμα αφαιρούνταν προσωρινά καθεμία από τις εναπομείνασες μεταβλητές, μαζί με όλες τις κατηγορίες της, και το μειωμένο μοντέλο συγκρινόταν με το μοντέλο του συγκεκριμένου βήματος μέσω LRT. Η μεταβλητή με τη μεγαλύτερη μη στατιστικά σημαντική **τιμή p αφαιρούνταν οριστικά και η διαδικασία επαναλαμβανόταν, μέχρι όλες οι μεταβλητές που παρέμειναν να είναι στατιστικά σημαντικές** σε επίπεδο α = 0,05.

Στο **μοντέλο των νεκρών και βαριά τραυματιών** αφαιρέθηκαν διαδοχικά η εποχή (SEASON), ο τύπος ημέρας (DAY TYPE), οι καιρικές συνθήκες (WEATHER), ο αριθμός εμπλεκόμενων οχημάτων (VEHICLES INVOLVED), ο τύπος οδού (ROAD TYPE GROUP) και ο τύπος χρήστη (ROAD USER TYPE), με αποτέλεσμα να παραμείνουν **οκτώ μεταβλητές**. Στο μοντέλο των **ελαφρά τραυματιών** αφαιρέθηκε μόνο η εποχή, με αποτέλεσμα να παραμείνουν **13 από τις 14** αρχικές μεταβλητές (Πίνακες 5.4 και 5.5). Ο μεγαλύτερος αριθμός σημαντικών μεταβλητών στο μοντέλο των ελαφρά τραυματιών σχετίζεται και με το πολύ μεγαλύτερο πλήθος περιστατικών (5.981 έναντι 604), το οποίο αυξάνει τη στατιστική ισχύ των ελέγχων.

Πίνακας 5.4: Ιστορικό της προς τα πίσω απαλοιφής

Βήμα	Μεταβλητή που αφαιρέθηκε	Τιμή p (LRT)
Μοντέλο νεκρών και βαριά τραυματιών (KSI): Poisson		
1	Εποχή (SEASON)	0,8139
2	Τύπος ημέρας (DAY TYPE)	0,6430
3	Καιρικές συνθήκες (WEATHER)	0,2958
4	Εμπλεκόμενα οχήματα (VEHICLES INVOLVED)	0,3105
5	Τύπος οδού (ROAD TYPE GROUP)	0,1034
6	Τύπος χρήστη (ROAD USER TYPE)	0,0895
Μοντέλο ελαφρά τραυματιών: Αρνητική Διωνυμική (NB2)		
1	Εποχή (SEASON)	0,5857

Πίνακας 5.5: Μεταβλητές που διατηρήθηκαν ή αφαιρέθηκαν σε κάθε μοντέλο και τιμή p του ελέγχου LRT

Μεταβλητή	Νεκροί και βαριά τραυματίες	Ελαφρά τραυματίες
Εποχή (SEASON)	✗ Αφαιρέθηκε (p = 0,8139)	✗ Αφαιρέθηκε (p = 0,5857)
Τύπος χρήστη (ROAD USER TYPE)	✗ Αφαιρέθηκε (p = 0,0895)	✓ Διατηρήθηκε (p < 0,001)
Φύλο (GENDER)	✓ Διατηρήθηκε (p < 0,001)	✓ Διατηρήθηκε (p < 0,001)

Μεταβλητή	Νεκροί και βαριά τραυματίες	Ελαφρά τραυματίες
Τύπος οδού (ROAD TYPE GROUP)	✗ Αφαιρέθηκε (p = 0,1034)	✓ Διατηρήθηκε (p < 0,001)
Κατοικημένη περιοχή (URBAN AREA)	✓ Διατηρήθηκε (p < 0,001)	✓ Διατηρήθηκε (p < 0,001)
Τύπος ημέρας (DAY TYPE)	✗ Αφαιρέθηκε (p = 0,6430)	✓ Διατηρήθηκε (p < 0,001)
Κόμβος (JUNCTION)	✓ Διατηρήθηκε (p = 0,0450)	✓ Διατηρήθηκε (p < 0,001)
Ηλικιακή ομάδα (AGE GROUP)	✓ Διατηρήθηκε (p = 0,0065)	✓ Διατηρήθηκε (p < 0,001)
Περίοδος ημέρας (HOUR DAY/NIGHT)	✓ Διατηρήθηκε (p = 0,0018)	✓ Διατηρήθηκε (p < 0,001)
Εμπλεκόμενα οχήματα (VEHICLES INVOLVED)	✗ Αφαιρέθηκε (p = 0,3105)	✓ Διατηρήθηκε (p < 0,001)
Καιρικές συνθήκες (WEATHER)	✗ Αφαιρέθηκε (p = 0,2958)	✓ Διατηρήθηκε (p < 0,001)
Τύπος ατυχήματος (ACCIDENT TYPE)	✓ Διατηρήθηκε (p = 0,0031)	✓ Διατηρήθηκε (p < 0,001)
Τύπος δικύκλου (VEHICLE TYPE)	✓ Διατηρήθηκε (p = 0,0016)	✓ Διατηρήθηκε (p < 0,001)
Έλεγχος αλκοόλ (ALCOHOL TEST)	✓ Διατηρήθηκε (p < 0,001)	✓ Διατηρήθηκε (p < 0,001)
Μεταβλητές τελικού μοντέλου	**8 από 14**	**13 από 14**

5.4 Αποτελέσματα τελικών μοντέλων

Στην ενότητα αυτή παρουσιάζονται οι **εκτιμήσεις των δύο τελικών μοντέλων**. Για κάθε κατηγορία δίνονται ο συντελεστής β, ο λόγος IRR με το διάστημα εμπιστοσύνης 95% και η τιμή p του ελέγχου Wald, σε σύγκριση πάντοτε με την αντίστοιχη κατηγορία αναφοράς.

5.4.1 Μοντέλο σοβαρών και θανατηφόρων τραυματισμών (KSI)

Το τελικό μοντέλο των νεκρών και βαριά τραυματιών είναι μικτό μοντέλο **παλινδρόμησης Poisson** με τυχαίο σταθερό όρο ανά **γεωγραφική περιοχή**. Οι οκτώ μεταβλητές που παρέμειναν είναι το φύλο (GENDER), η κατοικημένη περιοχή (URBAN AREA), ο κόμβος (JUNCTION), η ηλικιακή ομάδα (AGE GROUP), η περίοδος ημέρας (HOUR DAY/NIGHT), ο τύπος ατυχήματος (ACCIDENT TYPE), ο τύπος δικύκλου (VEHICLE TYPE) και ο έλεγχος αλκοόλ (ALCOHOL TEST). Η λογαριθμική πιθανοφάνεια του τελικού μοντέλου είναι -1.528,73, το AIC 3.093,46 και το BIC 3.207,06, ενώ η εκτιμώμενη διακύμανση του τυχαίου όρου είναι 0,1857.

Το AIC του τελικού μοντέλου είναι χαμηλότερο από εκείνο του πλήρους μικτού μοντέλου (3.101,33), παρότι το **τελικό μοντέλο περιλαμβάνει λιγότερες**

παραμέτρους. Αυτό υποστηρίζει την επιλογή μιας πιο φειδωλής μορφής χωρίς ουσιαστική απώλεια προσαρμογής. Οι συντελεστές του τελικού μοντέλου παρουσιάζονται στον Πίνακα 5.6. Οι κατηγορίες ομαδοποιούνται ανά μεταβλητή, με την κατηγορία αναφοράς στη γραμμή τίτλου κάθε ομάδας, ενώ **η στήλη «Επίδραση» δείχνει αν η κατηγορία συνδέεται με αύξηση ή μείωση του αναμενόμενου αριθμού παθόντων σε σχέση με την κατηγορία αναφοράς,** σε επίπεδο σημαντικότητας 5%.

Πίνακας 5.6: Συντελεστές του τελικού μοντέλου νεκρών και βαριά τραυματιών (KSI)

Κατηγορία	β	IRR	95% ΔΕ του IRR	Τιμή p (Wald)	Επίδραση
Κατηγορία αναφοράς: Άνδρα (Male)					
Γυναίκα (Female)	-0,696	**0,498**	0,371 - 0,670	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Εντός (Inside)					
Εκτός (Outside)	0,681	**1,975**	1,634 - 2,389	< 0,001	**Αύξηση**
Κατηγορία αναφοράς: Όχι (No)					
Ναι (Yes)	-0,200	**0,819**	0,672 - 0,997	0,0467	**Μείωση**
Κατηγορία αναφοράς: 25-49					
19-24	-0,078	**0,925**	0,733 - 1,168	0,5122	*Μη σημαντική*
50-64	0,141	**1,152**	0,926 - 1,432	0,2051	*Μη σημαντική*
65+	0,402	**1,496**	1,130 - 1,980	0,0049	**Αύξηση**
≤18	0,063	**1,065**	0,776 - 1,461	0,6983	*Μη σημαντική*
Άγνωστη (Unknown)	-0,688	**0,503**	0,257 - 0,983	0,0446	**Μείωση**
Κατηγορία αναφοράς: Ημέρας (Day)					
Νύχτα (Night)	0,271	**1,311**	1,107 - 1,552	0,0017	**Αύξηση**
Κατηγορία αναφοράς: Σύγκρουσης με χρήστες της οδού (Collision with road users)					
Εκτροπή (Diversion)	0,276	**1,318**	1,071 - 1,621	0,0090	**Αύξηση**
Λοιπές συγκρούσεις (Other collision)	0,447	**1,564**	1,188 - 2,059	0,0014	**Αύξηση**
Ανατροπή/Λοιπά (Overturning/Other)	-0,004	**0,996**	0,732 - 1,356	0,9806	*Μη σημαντική*
Κατηγορία αναφοράς: Μεσαίου (Medium PTWs)					
Μεγάλο (Large PTWs)	0,343	**1,410**	1,162 - 1,710	< 0,001	**Αύξηση**
Μικρό (Small PTWs)	0,235	**1,264**	1,027 - 1,557	0,0271	**Αύξηση**
Κατηγορία αναφοράς: Αρνητικού/μηδενικού (Negative/zero)					
Θετικός (Positive)	0,648	**1,911**	1,401 - 2,608	< 0,001	**Αύξηση**
Άγνωστος/δεν διενεργήθηκε (Unknown/not tested)	1,175	**3,239**	2,704 - 3,880	< 0,001	**Αύξηση**

Ως προς το φύλο, οι γυναίκες (Female) εμφανίζουν IRR = 0,498 σε σχέση με τους άνδρες (Male), δηλαδή **περίπου 50,2% χαμηλότερο αναμενόμενο αριθμό νεκρών και βαριά τραυματιών**, με τις υπόλοιπες μεταβλητές σταθερές. **Τα ατυχήματα εκτός κατοικημένης περιοχής** (Outside) παρουσιάζουν IRR = 1,975 έναντι εκείνων εντός

(Inside), που αντιστοιχεί σε περίπου 97,5% υψηλότερο αναμενόμενο αριθμό. Αντίθετα, τα **ατυχήματα σε κόμβο συνδέονται** με περίπου 18,1% **χαμηλότερο αναμενόμενο αριθμό σε σχέση με εκείνα εκτός κόμβου.**

Η ηλικιακή ομάδα είναι συνολικά **στατιστικά σημαντική** σύμφωνα με τον έλεγχο LRT, παρότι δεν είναι σημαντικές όλες οι επιμέρους συγκρίσεις. Ειδικότερα, **η ομάδα 65+** παρουσιάζει IRR = 1,496 σε σχέση με την ομάδα 25-49, δηλαδή περίπου 49,6% υψηλότερο **αναμενόμενο αριθμό**. **Οι ομάδες 19-24, 50-64 και ≤18 δεν διαφέρουν στατιστικά** από την κατηγορία αναφοράς σε επίπεδο 5%. Η κατηγορία άγνωστη (Unknown) εμφανίζει χαμηλότερη εκτίμηση, αλλά η ερμηνεία της αφορά πρωτίστως την καταγραφή της ηλικίας και όχι κάποιο ουσιαστικό ηλικιακό χαρακτηριστικό.

Η νύχτα (Night) παρουσιάζει IRR = 1,311 έναντι της ημέρας (Day) και συνδέεται επομένως με περίπου 31,1% **υψηλότερο αναμενόμενο αριθμό νεκρών και βαριά τραυματιών.** Σε σχέση με τις συγκρούσεις με άλλους χρήστες της οδού (Collision with road users), οι εκτροπές (Diversion) παρουσιάζουν **αύξηση περίπου 31,8% και οι λοιπές συγκρούσεις** (Other collision) **περίπου 56,4%**, ενώ η κατηγορία ανατροπή/λοιπά (Overturning/Other) δεν διαφοροποιείται στατιστικά. Παρά το μη σημαντικό αυτό επιμέρους αποτέλεσμα, ο τύπος ατυχήματος είναι συνολικά σημαντικός σύμφωνα με τον έλεγχο LRT, γεγονός που αναδεικνύει τη διαφορά μεταξύ του ελέγχου ολόκληρης της μεταβλητής και του ελέγχου μιας μεμονωμένης κατηγορίας.

Σε σχέση με τα μεσαία δίκυκλα (Medium PTWs), τα μεγάλα (Large PTWs) εμφανίζουν περίπου 41,0% και τα μικρά (Small PTWs) περίπου 26,4% υψηλότερο αναμενόμενο αριθμό νεκρών και βαριά τραυματιών. Τέλος, **ο θετικός έλεγχος αλκοόλ** (Positive) συνδέεται με IRR = 1,911, δηλαδή με **περίπου 91,1% υψηλότερο αναμενόμενο αριθμ**ό σε σχέση με τον αρνητικό/μηδενικό (Negative/zero), ενώ η κατηγορία άγνωστος/δεν διενεργήθηκε (Unknown/not tested) έχει IRR = 3,239. Η τελευταία σχέση πρέπει να ερμηνεύεται με ιδιαίτερη προσοχή, καθώς **η μη διενέργεια ή μη καταγραφή του ελέγχου** μπορεί να σχετίζεται με τις **συνθήκες του ατυχήματος και με τη διαδικασία καταγραφής**, για παράδειγμα με τη μεταφορά του παθόντα σε νοσοκομείο, και δεν αποτελεί από μόνη της μέτρηση κατανάλωσης αλκοόλ.

5.4.2 Μοντέλο ελαφρών τραυματισμών

Το τελικό μοντέλο των ελαφρά τραυματιών είναι μικτό μοντέλο παλινδρόμησης **Αρνητικής Διωνυμικής** τύπου NB2 με τυχαίο σταθερό όρο ανά **γεωγραφική περιοχή**. Από τις 14 αρχικές μεταβλητές **αφαιρέθηκε μόνο η εποχή (SEASON)** και παρέμειναν οι εξής: τύπος χρήστη (ROAD USER TYPE), φύλο (GENDER), τύπος

οδού (ROAD TYPE GROUP), κατοικημένη περιοχή (URBAN AREA), τύπος ημέρας (DAY TYPE), κόμβος (JUNCTION), ηλικιακή ομάδα (AGE GROUP), περίοδος ημέρας (HOUR DAY/NIGHT), εμπλεκόμενα οχήματα (VEHICLES INVOLVED), καιρικές συνθήκες (WEATHER), τύπος ατυχήματος (ACCIDENT TYPE), τύπος δικύκλου (VEHICLE TYPE) και έλεγχος αλκοόλ (ALCOHOL TEST).

Η λογαριθμική πιθανοφάνεια του τελικού μοντέλου είναι -5.772,29, το AIC 11.598,59 και το BIC 11.768,99. Η παράμετρος θ της Αρνητικής Διωνυμικής είναι 4,992 και η διακύμανση του τυχαίου όρου 0,0734. Σε σχέση με το πλήρες μικτό μοντέλο (AIC 11.602,65), η αφαίρεση της εποχής οδηγεί σε ελαφρά μείωση του AIC, γεγονός που υποστηρίζει την τελική, πιο φειδωλή μορφή. Οι συντελεστές του τελικού μοντέλου παρουσιάζονται στον Πίνακα 5.7.

Πίνακας 5.7: Συντελεστές του τελικού μοντέλου ελαφρά τραυματιών

Κατηγορία	β	IRR	95% ΔΕ του IRR	Τιμή p (Wald)	Επίδραση
Κατηγορία αναφοράς: Οδηγού (Driver)					
Επιβάτης (Passenger)	-0,344	**0,709**	0,639 - 0,786	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Άνδρα (Male)					
Γυναίκα (Female)	-0,375	**0,687**	0,625 - 0,756	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Μη εθνικής (Non-National)					
Εθνική (National)	-0,250	**0,778**	0,677 - 0,895	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Εντός (Inside)					
Εκτός (Outside)	-0,282	**0,754**	0,671 - 0,847	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Καθημερινής (Weekday)					
Σαββατοκύριακο (Weekend)	-0,270	**0,763**	0,713 - 0,816	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Όχι (No)					
Ναι (Yes)	0,252	**1,287**	1,206 - 1,374	< 0,001	**Αύξηση**
Κατηγορία αναφοράς: 25-49					
0-14	-0,301	**0,740**	0,505 - 1,085	0,1228	*Μη σημαντική*
15-18	-0,518	**0,596**	0,516 - 0,689	< 0,001	**Μείωση**
19-24	-0,366	**0,694**	0,637 - 0,755	< 0,001	**Μείωση**
50-64	-0,283	**0,753**	0,695 - 0,817	< 0,001	**Μείωση**
65+	-0,675	**0,509**	0,443 - 0,584	< 0,001	**Μείωση**
Άγνωστη (Unknown)	-0,739	**0,478**	0,399 - 0,572	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Ημέρας (Day)					
Νύχτα (Night)	-0,276	**0,759**	0,710 - 0,811	< 0,001	**Μείωση**
Κατηγορία αναφοράς: 2 οχημάτων (2 vehicles)					
1 όχημα (1 vehicle)	-0,251	**0,778**	0,693 - 0,874	< 0,001	**Μείωση**
3+ οχήματα (3+ vehicles)	-0,701	**0,496**	0,427 - 0,577	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Καλών (Good)					
Λοιπές (Other)	-0,548	**0,578**	0,496 - 0,674	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Σύγκρουσης με χρήστες της οδού (Collision with road users)					
Εκτροπή (Diversion)	-0,281	**0,755**	0,658 - 0,867	< 0,001	**Μείωση**

Κατηγορία	β	IRR	95% ΔΕ του IRR	Τιμή p (Wald)	Επίδραση
Λοιπές συγκρούσεις (Other collision)	-0,530	**0,588**	0,497 - 0,697	< 0,001	**Μείωση**
Ανατροπή/Λοιπά (Overturning/Other)	-0,434	**0,648**	0,559 - 0,752	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Μεσαίου (Medium PTWs)					
Μεγάλο (Large PTWs)	-0,110	**0,896**	0,836 - 0,961	0,0022	**Μείωση**
Μικρό (Small PTWs)	-0,246	**0,782**	0,722 - 0,846	< 0,001	**Μείωση**
Κατηγορία αναφοράς: Αρνητικού/μηδενικού (Negative/zero)					
Υψηλά θετικός (High positive)	-0,413	**0,661**	0,531 - 0,824	< 0,001	**Μείωση**
Χαμηλά θετικός (Low positive)	-0,450	**0,638**	0,518 - 0,784	< 0,001	**Μείωση**
Άγνωστος/δεν διενεργήθηκε (Unknown/not tested)	-0,398	**0,671**	0,626 - 0,720	< 0,001	**Μείωση**

Στο μοντέλο των ελαφρά τραυματιών, οι επιβάτες (Passenger) εμφανίζουν περίπου 29,1% χαμηλότερο αναμενόμενο αριθμό **σε σχέση με τους οδηγούς** (Driver), ενώ οι γυναίκες **περίπου 31,3% χαμηλότερο σε σχέση με τους άνδρες**. Οι εθνικές οδοί (National) και τα ατυχήματα εκτός κατοικημένης περιοχής (Outside) παρουσιάζουν αντίστοιχα περίπου 22,2% και 24,6% χαμηλότερο αναμενόμενο αριθμό ελαφρά τραυματιών. Το Σαββατοκύριακο (Weekend) συνδέεται επίσης με περίπου **23,7% χαμηλότερο** αναμενόμενο αριθμό **σε σχέση με τις καθημερινές** (Weekday).

Τα ατυχήματα σε κόμβο έχουν θετικό συντελεστή: το IRR = 1,287 αντιστοιχεί σε περίπου 28,7% **υψηλότερο αναμενόμενο αριθμό ελαφρά τραυματιών σε σχέση με τα ατυχήματα εκτός κόμβου**. Η ηλικιακή ομάδα είναι συνολικά ιδιαίτερα σημαντική. Σε σχέση με την ομάδα 25-49, οι ομάδες 15-18, 19-24, 50-64 και 65+ παρουσιάζουν χαμηλότερο αναμενόμενο αριθμό, **με τη μεγαλύτερη μείωση στην ομάδα 65+** (IRR = 0,509). Η ομάδα 0-14 δεν διαφέρει στατιστικά από την κατηγορία αναφοράς σε επιμέρους επίπεδο.

Η νύχτα συνδέεται με περίπου 24,1% **χαμηλότερο αναμενόμενο αριθμό** ελαφρά τραυματιών από την ημέρα. Σε σχέση με τα ατυχήματα με δύο εμπλεκόμενα οχήματα, η κατηγορία του ενός οχήματος έχει IRR = 0,778 και η κατηγορία τριών ή περισσότερων οχημάτων IRR = 0,496. Οι λοιπές **καιρικές συνθήκες** (Other) παρουσιάζουν **επίσης χαμηλότερο** αναμενόμενο αριθμό από τις καλές (Good), με IRR = 0,578.

Και οι τρεις εναλλακτικές κατηγορίες του τύπου ατυχήματος παρουσιάζουν IRR μικρότερο της μονάδας σε σχέση με τις συγκρούσεις με άλλους χρήστες της οδού. Αντίστοιχα, τα μεγάλα και τα μικρά δίκυκλα εμφανίζουν χαμηλότερο αναμενόμενο αριθμό ελαφρά τραυματιών από τα μεσαία. Για τον έλεγχο αλκοόλ, **οι κατηγορίες**

υψηλά θετικός, χαμηλά θετικός και άγνωστος/δεν διενεργήθηκε **παρουσιάζουν** IRR **μικρότερο** της μονάδας σε σχέση με τον αρνητικό/μηδενικό.

Τα αποτελέσματα αυτά **δεν πρέπει να ερμηνευθούν ως ένδειξη ότι οι συγκεκριμένες συνθήκες είναι ασφαλέστερες**. Η εξαρτημένη μεταβλητή αφορά αποκλειστικά τον αριθμό ελαφρά τραυματιών ανά εγγραφή και επηρεάζεται και από το πόσο συχνός είναι κάθε συνδυασμός χαρακτηριστικών στα δεδομένα: οι συνδυασμοί που περιλαμβάνουν λιγότερο συχνές κατηγορίες συγκεντρώνουν κατά κανόνα λιγότερους παθόντες ανά εγγραφή, με αποτέλεσμα οι κατηγορίες αυτές να τείνουν να εμφανίζουν IRR μικρότερο της μονάδας σε σχέση με τη συχνότερη κατηγορία αναφοράς.

5.4.3 Διαγνωστικός έλεγχος και αξιολόγηση μοντέλων

Η αξιολόγηση των **τελικών μοντέλων βασίστηκε στα στατιστικά προσαρμογής τους** (Πίνακας 5.8) και, για το μοντέλο των νεκρών και βαριά τραυματιών, σε διαγνωστικούς ελέγχους με προσομοιωμένα κατάλοιπα. Και τα δύο τελικά μοντέλα συνέκλιναν κανονικά και παρουσιάζουν χαμηλότερο AIC τόσο από τα αντίστοιχα μοντέλα σταθερών επιδράσεων όσο και από τα πλήρη μικτά μοντέλα. Σημειώνεται ότι οι τιμές AIC και BIC συγκρίνονται μόνο μεταξύ μοντέλων της ίδιας εξαρτημένης μεταβλητής και όχι μεταξύ των δύο μοντέλων.

Πίνακας 5.8: Στατιστικά προσαρμογής των τελικών μοντέλων

Στατιστικό	Νεκροί και βαριά τραυματίες	Ελαφρά τραυματίες
Κατανομή	Poisson	Αρνητική Διωνυμική (NB2)
Αριθμός μεταβλητών τελικού μοντέλου	8	13
Λογαριθμική πιθανοφάνεια (log-likelihood)	-1.528,73	-5.772,29
AIC τελικού μοντέλου	3.093,46	11.598,59
BIC τελικού μοντέλου	3.207,06	11.768,99
AIC πλήρους μικτού μοντέλου	3.101,33	11.602,65
AIC πλήρους μοντέλου σταθερών επιδράσεων	3.169,34	11.941,14
Διακύμανση τυχαίου όρου (σ^2_u)	0,1857	0,0734
Παράμετρος διασποράς θ	-	4,992

Για το **τελικό μικτό μοντέλο Poisson των νεκρών και βαριά τραυματιών εφαρμόστηκαν επιπλέον διαγνωστικοί έλεγχοι με το πακέτο DHARMa**, με 1.000 προσομοιώσεις. Οι έλεγχοι αυτοί κρίθηκαν απαραίτητοι για δύο λόγους. Πρώτον, στα μοντέλα καταμέτρησης, και ιδιαίτερα στα μικτά μοντέλα, τα συνήθη κατάλοιπα

(κατάλοιπα Pearson ή απόκλισης, deviance) δεν ακολουθούν γνωστή κατανομή και είναι δύσκολο να ερμηνευθούν, ειδικά όταν η εξαρτημένη μεταβλητή παίρνει κυρίως τις τιμές 0 και 1. Δεύτερον, το μοντέλο των νεκρών και βαριά τραυματιών βασίζεται στην κατανομή Poisson, η οποία προϋποθέτει ισότητα μέσης τιμής και διακύμανσης, ενώ τα δεδομένα παρουσιάζουν πολύ υψηλό ποσοστό μηδενικών τιμών (85,45%) και ήπια ένδειξη υποδιασποράς. Ήταν επομένως αναγκαίο να ελεγχθεί αν η κατανομή Poisson αναπαράγει επαρκώς το πλήθος των μηδενικών, τη διασπορά και τις ακραίες τιμές των δεδομένων, ή αν θα απαιτούνταν εναλλακτική μορφή μοντέλου, όπως μοντέλο με διόγκωση μηδενικών (zero-inflated).

Η μέθοδος συγκρίνει τις παρατηρούμενες τιμές με τιμές που προσομοιώνονται από το εκτιμημένο μοντέλο και παράγει τυποποιημένα κατάλοιπα (scaled residuals), τα οποία ακολουθούν ομοιόμορφη κατανομή όταν το μοντέλο είναι κατάλληλο. Ο έλεγχος μηδενικών τιμών (zero-inflation) **δεν έδειξε σημαντική απόκλιση στο πλήθος των μηδενικών** (p = 0,888), ενώ ο έλεγχος ακραίων τιμών (outliers) **δεν έδειξε υπερβολικό αριθμό ακραίων παρατηρήσεων** (p = 0,214).

Ο έλεγχος διασποράς (dispersion test) έδωσε p = 0,024 και είναι συμβατός με **ήπια υποδιασπορά**, σε συμφωνία με τον λόγο διασποράς Pearson μικρότερο της μονάδας. Ο έλεγχος ομοιομορφίας Kolmogorov-Smirnov έδωσε p = 0,0048, υποδεικνύοντας μικρή αλλά **στατιστικά σημαντική** απόκλιση των καταλοίπων από την ομοιόμορφη κατανομή. Επιπλέον, η ήπια υποδιασπορά σε μοντέλο Poisson οδηγεί σε ελαφρώς συντηρητικά τυπικά σφάλματα, δηλαδή σε διαστήματα εμπιστοσύνης ελαφρώς ευρύτερα από τα απαιτούμενα (Πίνακας 5.9).

Πίνακας 5.9: Διαγνωστικοί έλεγχοι του τελικού μοντέλου νεκρών και βαριά τραυματιών (KSI)

Έλεγχος	Αποτέλεσμα	Συμπέρασμα
Λόγος διασποράς Pearson	0,8557	Δεν προκύπτει υπερδιασπορά
DHARMa: έλεγχος διασποράς	p = 0,024	Ήπια ένδειξη υποδιασποράς
DHARMa: έλεγχος μηδενικών τιμών	p = 0,888	Δεν προκύπτει πρόβλημα μηδενικών τιμών
DHARMa: έλεγχος ακραίων τιμών	p = 0,214	Δεν προκύπτει σημαντικό πρόβλημα ακραίων τιμών
DHARMa: έλεγχος ομοιομορφίας	p = 0,0048	Μικρή αλλά στατιστικά σημαντική απόκλιση από την ομοιομορφία

Για το μοντέλο των **ελαφρά τραυματιών δεν εφαρμόστηκαν έλεγχοι DHARMa.** Η επάρκειά του αξιολογήθηκε με βάση τα στατιστικά προσαρμογής του Πίνακα 5.8 και τη σύγκριση με το αντίστοιχο μοντέλο σταθερών επιδράσεων και το πλήρες μικτό μοντέλο.

5.5 Ανάλυση γεωγραφικών διαφοροποιήσεων

Η **γεωγραφική περιοχή** εισήχθη στα μοντέλα ως τυχαίος σταθερός όρος. Ο όρος αυτός αποτυπώνει διαφορές στον **αναμενόμενο αριθμό** παθόντων μεταξύ των 12 περιοχών, οι οποίες δεν εξηγούνται από τις επεξηγηματικές μεταβλητές του μοντέλου και μπορεί να οφείλονται σε μη καταγεγραμμένους παράγοντες, όπως τα χαρακτηριστικά του οδικού δικτύου, οι κυκλοφοριακοί φόρτοι, η αστυνόμευση, η πρόσβαση σε υπηρεσίες υγείας ή η πρακτική καταγραφής των ατυχημάτων. Όπως φάνηκε στην Ενότητα 5.3.3, **η προσθήκη του όρου βελτίωσε σημαντικά την προσαρμογή και των δύο μοντέλων.**

Η εκτιμώμενη διακύμανση του τυχαίου όρου είναι μεγαλύτερη στο μοντέλο των νεκρών και βαριά τραυματιών (0,1857) από ό,τι στο μοντέλο των ελαφρά τραυματιών (0,0734), γεγονός που υποδηλώνει **ισχυρότερη μη παρατηρούμενη γεωγραφική ετερογένεια για τις σοβαρές συνέπειες**. Για να γίνει πιο κατανοητό το μέγεθος των διαφορών, η διακύμανση μπορεί να μετατραπεί σε πολλαπλασιαστικό συντελεστή (Πίνακας 5.10). Στο μοντέλο των νεκρών και βαριά τραυματιών η τυπική απόκλιση του τυχαίου όρου είναι 0,431. Επομένως, μια περιοχή που βρίσκεται μία τυπική απόκλιση πάνω από τον μέσο όρο έχει περίπου 1,54 φορές μεγαλύτερο αναμενόμενο αριθμό από μια μέση περιοχή με τα ίδια χαρακτηριστικά, ενώ το εύρος που καλύπτει περίπου το 95% των περιοχών κυμαίνεται από 0,43 έως 2,33 φορές. Στο μοντέλο των ελαφρά τραυματιών η τυπική απόκλιση είναι 0,271 και τα αντίστοιχα μεγέθη είναι 1,31 φορές και από 0,59 έως 1,70 φορές.

Πίνακας 5.10: Μέγεθος της γεωγραφικής ετερογένειας στα δύο τελικά μοντέλα

Μέγεθος	Νεκροί και βαριά τραυματίες (KSI)	Ελαφρά τραυματίες
Διακύμανση τυχαίου όρου ($\sigma^2 u$)	0,1857	0,0734
Τυπική απόκλιση τυχαίου όρου (σu)	0,431	0,271
Πολλαπλασιαστικός συντελεστής για +1 τυπική απόκλιση, $\exp(\sigma u)$	1,54	1,31
Εύρος για περίπου 95% των περιοχών, $\exp(\pm 1{,}96 \cdot \sigma u)$	0,43 - 2,33	0,59 - 1,70

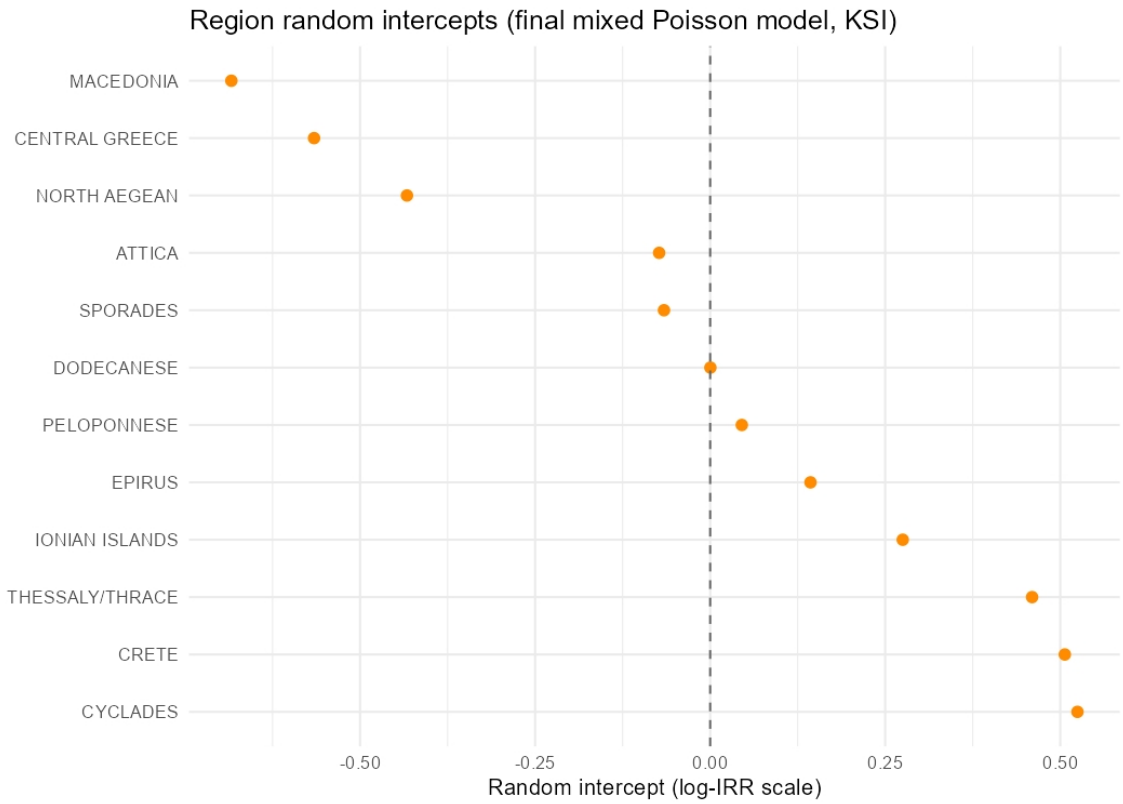
Η μεγαλύτερη ετερογένεια στο μοντέλο των νεκρών και βαριά τραυματιών συμφωνεί με την εικόνα των ακατέργαστων δεδομένων. Το ποσοστό των νεκρών και βαριά τραυματιών επί του συνόλου των παθόντων διαφέρει έντονα μεταξύ των περιοχών (Πίνακας 5.11): είναι χαμηλό στην Αττική (5,1%) και στη Μακεδονία (6,7%), όπου κυριαρχούν τα αστικά κέντρα της Αθήνας και της Θεσσαλονίκης, και πολύ υψηλότερο

στην Κρήτη (36,9%), στις Κυκλάδες (30,7%) και στη Θεσσαλία/Θράκη (28,0%). **Οι διαφορές αυτές αντανακλούν εν μέρει χαρακτηριστικά** που ήδη περιλαμβάνονται στα μοντέλα, **όπως το μερίδιο των ατυχημάτων εκτός κατοικημένης περιοχής, ενώ το υπόλοιπο μέρος τους αποδίδεται στον τυχαίο όρο.**

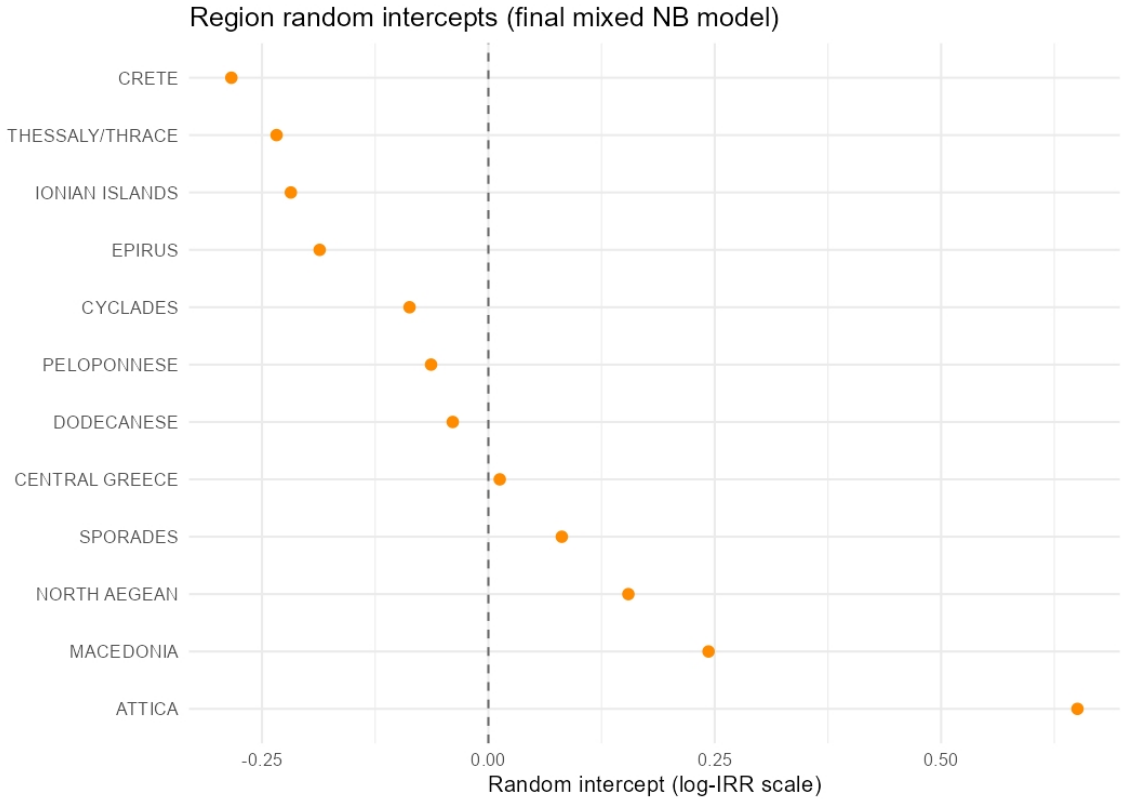
Πίνακας 5.11: Παθόντες και ποσοστό νεκρών και βαριά τραυματιών ανά γεωγραφική περιοχή

Γεωγραφική περιοχή	Νεκροί και βαριά τραυματίες	Ελαφρά τραυματίες	Σύνολο παθόντων	% νεκρών και βαριά τραυματιών
Αττική	196	3.655	3.851	5,1%
Μακεδονία	75	1.052	1.127	6,7%
Στερεά Ελλάδα	22	252	274	8,0%
Βόρειο Αιγαίο	19	119	138	13,8%
Σποράδες	2	12	14	14,3%
Πελοπόννησος	67	294	361	18,6%
Δωδεκάνησα	30	129	159	18,9%
Ήπειρος	15	50	65	23,1%
Ιόνια Νησιά	24	73	97	24,7%
Θεσσαλία/Θράκη	58	149	207	28,0%
Κυκλάδες	58	131	189	30,7%
Κρήτη	38	65	103	36,9%
Σύνολο	**604**	**5.981**	**6.585**	**9,2%**

Οι εκτιμώμενοι **τυχαίοι σταθεροί όροι κάθε περιοχής** για τα δύο τελικά μοντέλα παρουσιάζονται στα Διαγράμματα 5.1 και 5.2. Κάθε σημείο εκφράζει την απόκλιση της περιοχής από τον μέσο όρο σε λογαριθμική κλίμακα (log-IRR), αφού ληφθούν υπόψη οι επεξηγηματικές μεταβλητές του μοντέλου. **Θετικές τιμές υποδηλώνουν υψηλότερο και αρνητικές τιμές χαμηλότερο αναμενόμενο αριθμό παθόντων** σε σχέση με μια μέση περιοχή με τα ίδια χαρακτηριστικά, ενώ η εκθετοποίηση της τιμής δίνει τον αντίστοιχο πολλαπλασιαστικό συντελεστή.



Διάγραμμα 5.1: Τυχαίοι σταθεροί όροι ανά γεωγραφική περιοχή στο τελικό μοντέλο νεκρών και βαριά τραυματιών (KSI)



Διάγραμμα 5.2: Τυχαίοι σταθεροί όροι ανά γεωγραφική περιοχή στο τελικό μοντέλο ελαφρά τραυματιών

Στο **μοντέλο των νεκρών και βαριά τραυματιών** (Διάγραμμα 5.1), τους **υψηλότερους** τυχαίους όρους εμφανίζουν οι **Κυκλάδες, η Κρήτη και η Θεσσαλία/Θράκη** (περίπου 1,6 έως 1,7 φορές υψηλότερος αναμενόμενος αριθμός από μια μέση περιοχή), ενώ τους **χαμηλότερους** η **Μακεδονία, η Στερεά Ελλάδα και το Βόρειο Αιγαίο**. Η Αττική βρίσκεται κοντά στο μηδέν, γεγονός που υποδηλώνει ότι το χαμηλό ποσοστό νεκρών και βαριά τραυματιών που καταγράφεται στα ακατέργαστα δεδομένα (5,1%) αποδίδεται σε μεγάλο βαθμό σε χαρακτηριστικά που ήδη περιλαμβάνονται στο μοντέλο, όπως το μεγάλο μερίδιο ατυχημάτων εντός κατοικημένης περιοχής. Στο **μοντέλο των ελαφρά τραυματιών** (Διάγραμμα 5.2) η εικόνα είναι σε μεγάλο βαθμό **αντίστροφη**: τον **υψηλότερο** τυχαίο όρο εμφανίζει η **Αττική** (περίπου +0,66) και ακολουθεί η **Μακεδονία** (περίπου +0,24), ενώ τους **χαμηλότερους** η **Κρήτη**, η **Θεσσαλία/Θράκη και τα Ιόνια Νησιά** (περίπου -0,22 έως -0,29). Η αντίθεση αυτή συμφωνεί με τη σύγκριση των δύο μοντέλων της Ενότητας 5.6: στις περιοχές των μεγάλων αστικών κέντρων καταγράφονται περισσότεροι ελαφρά τραυματίες, ενώ σε περιοχές όπως οι Κυκλάδες, η Κρήτη και η Θεσσαλία/Θράκη η έκβαση των ατυχημάτων μετατοπίζεται προς σοβαρότερες συνέπειες. Οι εκτιμήσεις για περιοχές με λίγες εγγραφές, όπως οι Σποράδες, πρέπει να ερμηνεύονται με επιφύλαξη.

5.6 Σύγκριση αποτελεσμάτων των δύο μοντέλων

Η παράλληλη εξέταση των δύο τελικών μοντέλων επιτρέπει την ανάδειξη διαφορετικών προτύπων μεταξύ των ελαφρών και των σοβαρότερων συνεπειών.

5.6.1 Πλαίσιο ερμηνείας της σύγκρισης

Τα δύο τελικά μοντέλα εκτιμήθηκαν στις ίδιες 4.069 εγγραφές, με κοινό αρχικό σύνολο μεταβλητών και τις ίδιες κατηγορίες αναφοράς. Για τον λόγο αυτό, για κάθε κατηγορία που εμφανίζεται και στα δύο μοντέλα μπορεί να συγκριθεί το IRR του μοντέλου των νεκρών και βαριά τραυματιών με το IRR του μοντέλου των ελαφρά τραυματιών. Από τη σύγκριση προκύπτουν τρεις περιπτώσεις: **ίδια κατεύθυνση** στα δύο μοντέλα, **αντίθετη κατεύθυνση** και **στατιστική σημαντικότητα μόνο στο ένα από τα δύο** μοντέλα.

Όταν μια κατηγορία έχει την **ίδια κατεύθυνση** και στα δύο μοντέλα, συνδέεται με μεταβολή του συνολικού αριθμού παθόντων, δηλαδή με το πόσο συχνά προκύπτουν παθόντες στη συγκεκριμένη συνθήκη, και όχι απαραίτητα με το πόσο σοβαροί είναι οι τραυματισμοί. Αν τα δύο IRR έχουν και παρόμοιο μέγεθος, η σύνθεση της σοβαρότητας παραμένει περίπου σταθερή. Αντίθετα, όταν μια κατηγορία έχει **αντίθετη**

κατεύθυνση, δηλαδή συνδέεται με περισσότερους νεκρούς και βαριά τραυματίες αλλά με λιγότερους ελαφρά τραυματίες (ή το αντίστροφο), τότε η ίδια συνθήκη μετατοπίζει την έκβαση των ατυχημάτων προς σοβαρότερες (ή ελαφρύτερες) συνέπειες. Οι μεταβλητές αυτής της κατηγορίας μπορούν να χαρακτηριστούν **παράγοντες σοβαρότητας**, σε αντιδιαστολή με τους **παράγοντες συχνότητας**, οι οποίοι επηρεάζουν κυρίως τον αριθμό των παθόντων.

Για να αποτυπωθεί η μετατόπιση αυτή με έναν αριθμό, υπολογίστηκε ο **λόγος των δύο IRR** (IRR του μοντέλου KSI προς IRR του μοντέλου ελαφρά τραυματιών). Επειδή και τα δύο μοντέλα χρησιμοποιούν λογαριθμική συνάρτηση σύνδεσης, ο λόγος αυτός εκφράζει κατά προσέγγιση πόσες φορές μεταβάλλεται ο αναμενόμενος αριθμός νεκρών και βαριά τραυματιών ανά ελαφρά τραυματία στη συγκεκριμένη κατηγορία, σε σχέση με την κατηγορία αναφοράς. Τιμές μεγαλύτερες της μονάδας δείχνουν μετατόπιση προς σοβαρότερες συνέπειες, τιμές μικρότερες της μονάδας μετατόπιση προς ελαφρύτερες, ενώ τιμές κοντά στη μονάδα δείχνουν ότι η σύνθεση της σοβαρότητας δεν μεταβάλλεται. Ο λόγος έχει **ενδεικτικό χαρακτήρα**, καθώς τα δύο μοντέλα εκτιμήθηκαν χωριστά, περιλαμβάνουν διαφορετικό αριθμό μεταβλητών και ο λόγος δεν συνοδεύεται από στατιστικό έλεγχο.

Κατά την ανάγνωση της σύγκρισης πρέπει επίσης να λαμβάνεται υπόψη ότι, όπως αναφέρθηκε στην Ενότητα 5.4.2, στο μοντέλο των ελαφρά τραυματιών οι λιγότερο συχνές κατηγορίες τείνουν να εμφανίζουν IRR μικρότερο της μονάδας, επειδή οι αντίστοιχοι συνδυασμοί χαρακτηριστικών συγκεντρώνουν λιγότερους παθόντες ανά εγγραφή. Επομένως, ένα IRR μικρότερο της μονάδας στο μοντέλο των ελαφρά τραυματιών δεν σημαίνει από μόνο του ότι η συνθήκη είναι ασφαλέστερη. Η ουσιαστική πληροφορία προκύπτει από τη **σχετική θέση των δύο εκτιμήσεων**, όπως συνοψίζεται στον Πίνακα 5.12.

Πίνακας 5.12: Σύγκριση των IRR των δύο τελικών μοντέλων για τις κοινές κατηγορίες

Κατηγορία (κατηγορία αναφοράς)	IRR μοντέλου KSI	IRR μοντέλου ελαφρά τραυματιών	Λόγος IRR (KSI / ελαφρά)	Μορφή συσχέτισης
Φύλο: γυναίκα (έναντι άνδρα)	0,498*	0,687*	0,72	Ίδια κατεύθυνση (↓ / ↓)
Ηλικιακή ομάδα: άγνωστη (έναντι 25-49)	0,503*	0,478*	1,05	Ίδια κατεύθυνση (↓ / ↓)
Κατοικημένη περιοχή: εκτός (έναντι εντός)	1,975*	0,754*	2,62	Αντίθετη (↑ / ↓)
Ηλικιακή ομάδα: 65+ (έναντι 25-49)	1,496*	0,509*	2,94	Αντίθετη (↑ / ↓)
Περίοδος ημέρας: νύχτα (έναντι ημέρας)	1,311*	0,759*	1,73	Αντίθετη (↑ / ↓)

Κατηγορία (κατηγορία αναφοράς)	IRR μοντέλου KSI	IRR μοντέλου ελαφρά τραυματιών	Λόγος IRR (KSI / ελαφρά)	Μορφή συσχέτισης
Τύπος ατυχήματος: εκτροπή (έναντι σύγκρουσης με χρήστες της οδού)	1,318*	0,755*	1,75	Αντίθετη (↑ / ↓)
Τύπος ατυχήματος: λοιπές συγκρούσεις	1,564*	0,588*	2,66	Αντίθετη (↑ / ↓)
Τύπος δικύκλου: μεγάλο (έναντι μεσαίου)	1,410*	0,896*	1,57	Αντίθετη (↑ / ↓)
Τύπος δικύκλου: μικρό (έναντι μεσαίου)	1,264*	0,782*	1,62	Αντίθετη (↑ / ↓)
Έλεγχος αλκοόλ: θετικός (έναντι αρνητικού/μηδενικού)	1,911*	0,661* (υψηλά θετικός) / 0,638* (χαμηλά θετικός)	2,89 / 3,00	Αντίθετη (↑ / ↓)
Έλεγχος αλκοόλ: άγνωστος/δεν διενεργήθηκε	3,239*	0,671*	4,83	Αντίθετη (↑ / ↓)
Κόμβος: ναι (έναντι όχι)	0,819*	1,287*	0,64	Αντίθετη (↓ / ↑)
Ηλικιακή ομάδα: 19-24 (έναντι 25-49)	0,925	0,694*	1,33	Σημαντική μόνο στο μοντέλο ελαφρά τραυματιών
Ηλικιακή ομάδα: 50-64 (έναντι 25-49)	1,152	0,753*	1,53	Σημαντική μόνο στο μοντέλο ελαφρά τραυματιών
Ηλικιακή ομάδα: ≤18 (KSI) / 0-14 και 15-18 (ελαφρά)	1,065	0,740 / 0,596*	-	Σημαντική μόνο στο μοντέλο ελαφρά τραυματιών (15-18)
Τύπος ατυχήματος: ανατροπή/λοιπά	0,996	0,648*	1,54	Σημαντική μόνο στο μοντέλο ελαφρά τραυματιών

5.6.2 Μεταβλητές με ίδια κατεύθυνση στα δύο μοντέλα

Ίδια κατεύθυνση και στα δύο μοντέλα παρουσιάζουν **μόνο δύο κατηγορίες.** Οι **γυναίκες** εμφανίζουν χαμηλότερο αναμενόμενο αριθμό τόσο νεκρών και βαριά τραυματιών (IRR = 0,498) όσο και ελαφρά τραυματιών (IRR = 0,687) σε σχέση με τους άνδρες. Το πρώτο σκέλος του αποτελέσματος αντανακλά κυρίως τη μικρότερη παρουσία των γυναικών στα δεδομένα, αφού μόλις το 12,8% των παθόντων είναι γυναίκες, και επομένως αφορά τη συχνότητα. Ωστόσο, η μείωση είναι εντονότερη στο μοντέλο KSI, με λόγο IRR 0,72. Αυτό σημαίνει ότι, πέρα από τη μικρότερη συχνότητα, και η σύνθεση της σοβαρότητας είναι ευνοϊκότερη για τις γυναίκες. Πιθανές εξηγήσεις είναι η λιγότερο ριψοκίνδυνη συμπεριφορά και η χρήση οχημάτων μικρότερης ισχύος, χωρίς όμως τα διαθέσιμα δεδομένα να επιτρέπουν τη διάκριση μεταξύ των εξηγήσεων αυτών.

Η κατηγορία **άγνωστης ηλικίας** εμφανίζει σχεδόν ίδια μείωση στα δύο μοντέλα (IRR 0,503 και 0,478), με λόγο IRR περίπου ίσο με τη μονάδα (1,05). Η κατηγορία αυτή επηρεάζει επομένως μόνο τον αριθμό των καταγραφών και όχι τη σοβαρότητα, κάτι αναμενόμενο αφού πρόκειται για χαρακτηριστικό της διαδικασίας καταγραφής. Η περίπτωση αυτή επιβεβαιώνει και την ερμηνεία του πλαισίου της σύγκρισης: όταν μια συνθήκη επηρεάζει μόνο τη συχνότητα, τα δύο IRR κινούνται μαζί και ο λόγος τους πλησιάζει τη μονάδα.

5.6.3 Μεταβλητές με αντίθετη κατεύθυνση στα δύο μοντέλα

Η **πλειονότητα των κατηγοριών** που είναι σημαντικές και στα δύο μοντέλα (10 από τις 12) παρουσιάζει **αντίθετη κατεύθυνση**. Το εύρημα αυτό είναι **το πιο ουσιαστικό αποτέλεσμα της σύγκρισης**, διότι δείχνει ότι οι ίδιες συνθήκες που συνδέονται με λιγότερους ελαφρά τραυματίες συνδέονται ταυτόχρονα με περισσότερους νεκρούς και βαριά τραυματίες. Με άλλα λόγια, στις συνθήκες αυτές τα ατυχήματα δεν είναι απαραίτητα περισσότερα, αλλά όταν συμβαίνουν έχουν **σοβαρότερη έκβαση**. Οι κατηγορίες αυτές αποτελούν τους κύριους παράγοντες σοβαρότητας της παρούσας ανάλυσης.

Ο ισχυρότερος παράγοντας σοβαρότητας που σχετίζεται με το οδικό περιβάλλον είναι τα **ατυχήματα εκτός κατοικημένης περιοχής**, με IRR 1,975 στο μοντέλο KSI και 0,754 στο μοντέλο των ελαφρά τραυματιών (λόγος IRR 2,62). Τα ατυχήματα αυτά είναι λιγότερα, καθώς η **κυκλοφορία των δικύκλων είναι πολύ μεγαλύτερη μέσα στις πόλεις**, όμως οι **υψηλότερες ταχύτητες,** η απουσία φωτισμού σε πολλά τμήματα, τα παρόδια εμπόδια και ο μεγαλύτερος χρόνος απόκρισης των υπηρεσιών έκτακτης ανάγκης **αυξάνουν την ενέργεια της σύγκρουσης και τις συνέπειές της.** Το εύρημα συμφωνεί με τη διεθνή βιβλιογραφία, η οποία συνδέει τις αγροτικές περιοχές και τις εθνικές οδούς με αυξημένη σοβαρότητα (Santos et al., 2021).

Από τα χαρακτηριστικά του παθόντα, ο υψηλότερος λόγος IRR εμφανίζεται στην **ηλικιακή ομάδα 65+** (λόγος 2,94). Οι μεγαλύτεροι σε ηλικία αναβάτες **καταγράφονται σπανιότερα ως παθόντες,** πιθανότατα λόγω μικρότερης έκθεσης, όμως όταν εμπλέκονται σε ατύχημα **τραυματίζονται σοβαρότερα.** Η εξήγηση που δίνεται συνήθως στη βιβλιογραφία είναι η **μειωμένη φυσική αντοχή του οργανισμού στην πρόσκρουση** (Theofilatos & Yannis, 2015), ενώ αυξημένη πιθανότητα σοβαρού ή θανατηφόρου τραυματισμού για αναβάτες άνω των 60 ετών έχει καταγραφεί και σε άλλες χώρες (Chang et al., 2016).

Η νύχτα (λόγος 1,73) **συνδέεται με λιγότερους ελαφρά τραυματίες** αλλά περισσότερους νεκρούς και βαριά τραυματίες. Ο **μικρότερος κυκλοφοριακός**

φόρτος της νύχτας **μειώνει τον αριθμό των ατυχημάτων, αλλά** επιτρέπει υψηλότερες ταχύτητες, ενώ η μειωμένη ορατότητα, η κόπωση και η συχνότερη **κατανάλωση αλκοόλ** επιβαρύνουν την έκβαση.

Ως προς τον τύπο ατυχήματος, οι εκτροπές (λόγος 1,75) και κυρίως οι **λοιπές συγκρούσεις** (λόγος 2,66), οι οποίες περιλαμβάνουν τις προσκρούσεις σε σταθερό εμπόδιο, στύλο ή δέντρο, **μετατοπίζουν έντονα την έκβαση προς σοβαρότερες συνέπειες.** Στους τύπους αυτούς ο αναβάτης συνήθως **προσκρούει σε άκαμπτο αντικείμενο** ή στο οδόστρωμα χωρίς καμία προστασία, γεγονός **που εξηγεί και την πολύ υψηλή αναλογία KSI** στις προσκρούσεις σε σταθερό εμπόδιο (38,5%). Το αποτέλεσμα συμφωνεί με τα ευρήματα των Savolainen και Mannering (2007) για τις συγκρούσεις με σταθερά αντικείμενα.

Για τον **τύπο δικύκλου**, τόσο τα μεγάλα (λόγος 1,57) όσο και τα μικρά δίκυκλα (λόγος 1,62) εμφανίζουν σοβαρότερη σύνθεση συνεπειών σε σχέση με τα μεσαία. Η σχέση αυτή έχει μορφή «U»: **στα μεγάλα δίκυκλα η σοβαρότητα συνδέεται κυρίως με τις υψηλότερες ταχύτητες και επιταχύνσεις,** ενώ στα μικρά δίκυκλα, όπου περιλαμβάνονται και τα μοτοποδήλατα, μπορεί να σχετίζεται με τα χαρακτηριστικά των χρηστών και με τη μικρότερη σταθερότητα και προστασία του οχήματος. Τα ακατέργαστα δεδομένα δείχνουν την ίδια εικόνα, με τις υψηλότερες αναλογίες KSI στα μοτοποδήλατα έως 49 cc (16,4%) και στις μοτοσυκλέτες άνω των 730 cc (13,1%).

Ο **θετικός έλεγχος αλκοόλ** έχει από τους **υψηλότερους λόγους** IRR (περίπου 2,9 έως 3,0), ανάλογα με την κατηγορία σύγκρισης του μοντέλου των ελαφρά τραυματιών. Η κατανάλωση αλκοόλ επομένως **δεν συνδέεται απλώς με περισσότερα ατυχήματα, αλλά κυρίως με σοβαρότερη έκβασή τους,** κάτι που συμφωνεί με την ευρεία συναίνεση της βιβλιογραφίας (Zhang et al., 2026). Ο υψηλότερος λόγος του Πίνακα 5.12 αφορά την κατηγορία **άγνωστος/δεν διενεργήθηκε** (4,83). Το αποτέλεσμα αυτό δεν πρέπει να ερμηνευθεί ως επίδραση του αλκοόλ, αλλά ως ένδειξη ότι ο έλεγχος συχνά δεν πραγματοποιείται ή δεν καταγράφεται όταν ο παθόντας τραυματίζεται σοβαρά και μεταφέρεται σε νοσοκομείο. Επισημαίνει έτσι ένα σημαντικό κενό στην καταγραφή του αλκοόλ ακριβώς στα σοβαρότερα περιστατικά.

Τέλος, ο κόμβος **είναι η μόνη μεταβλητή με την αντίστροφη μορφή**, αφού συνδέεται με περισσότερους ελαφρά τραυματίες (IRR 1,287) και λιγότερους νεκρούς και βαριά τραυματίες (IRR 0,819), με λόγο IRR 0,64. **Στους κόμβους τα ατυχήματα είναι συχνά**, καθώς εκεί διασταυρώνονται οι κινήσεις των οχημάτων και εμφανίζονται τα σφάλματα αντίληψης των οδηγών που περιγράφηκαν στην Ενότητα 2.3.3 (Brown et al., 2021), ό**μως οι ταχύτητες είναι συνήθως χαμηλότερες και οι συγκρούσεις**

κυρίως πλευρικές. Οι κόμβοι αποτελούν επομένως κυρίως πρόβλημα συχνότητας, ενώ τα τμήματα εκτός κόμβου και εκτός κατοικημένης περιοχής κυρίως **πρόβλημα σοβαρότητας**.

Η εικόνα των μοντέλων επιβεβαιώνεται από τα ακατέργαστα δεδομένα (Πίνακας 5.13): για όλα τα χαρακτηριστικά με αντίθετη κατεύθυνση, το ποσοστό των νεκρών και βαριά **τραυματιών** επί του συνόλου των παθόντων είναι υψηλότερο στην κατηγορία που εμφανίζει IRR μεγαλύτερο της μονάδας στο μοντέλο KSI.

Πίνακας 5.13: Διαφορές κατεύθυνσης μεταξύ των δύο μοντέλων

Χαρακτηριστικό	Μοντέλο KSI	Μοντέλο ελαφρά τραυματιών
Κατοικημένη περιοχή: εκτός (Outside)	↑ (IRR 1,975)	↓ (IRR 0,754)
Ηλικιακή ομάδα: 65+	↑ (IRR 1,496)	↓ (IRR 0,509)
Περίοδος ημέρας: νύχτα (Night)	↑ (IRR 1,311)	↓ (IRR 0,759)
Τύπος δικύκλου: μεγάλο (Large PTWs)	↑ (IRR 1,410)	↓ (IRR 0,896)
Έλεγχος αλκοόλ: θετικός (Positive)	↑ (IRR 1,911)	↓ (υψηλά θετικός 0,661 / χαμηλά θετικός 0,638)
Κόμβος: ναι (Yes)	↓ (IRR 0,819)	↑ (IRR 1,287)

5.6.4 Μεταβλητές σημαντικές μόνο σε ένα από τα δύο μοντέλα

Ορισμένες κατηγορίες είναι **στατιστικά σημαντικές μόνο στο μοντέλο των ελαφρά τραυματιών**. Οι ηλικιακές ομάδες 15-18, 19-24 και 50-64, καθώς και ο τύπος ατυχήματος ανατροπή/λοιπά, εμφανίζουν λιγότερους ελαφρά τραυματίες από την κατηγορία αναφοράς, ενώ στο μοντέλο KSI δεν διαφέρουν σημαντικά. Οι κατηγορίες αυτές **επηρεάζουν** δηλαδή τη **συχνότητα** των παθόντων, **χωρίς να μειώνεται αντίστοιχα ο αριθμός των σοβαρών συνεπειών.** Αξίζει να σημειωθεί ότι για την ομάδα 50-64 ο λόγος IRR είναι 1,53, εικόνα που, μαζί με το αποτέλεσμα της ομάδας 65+, υποδηλώνει ότι η σοβαρότητα αυξάνεται σταδιακά με την ηλικία, παρότι για την ομάδα 50-64 η ένδειξη αυτή δεν είναι στατιστικά τεκμηριωμένη.

Πέντε μεταβλητές παρέμειναν μόνο στο μοντέλο των ελαφρά τραυματιών: ο **τύπος χρήστη** (επιβάτης, IRR 0,709), ο τύπος οδού (εθνική, IRR 0,778), ο τύπος ημέρας (Σαββατοκύριακο, IRR 0,763), ο **αριθμός εμπλεκόμενων οχημάτων** (ένα όχημα, IRR 0,778, και τρία ή περισσότερα, IRR 0,496) και οι **καιρικές συνθήκες** (λοιπές, IRR 0,578). Οι μεταβλητές αυτές συνδέονται με το πόσοι ελαφρά τραυματίες καταγράφονται, δηλαδή κυρίως **με την έκθεση και τη συχνότητα των αντίστοιχων συνθηκών,** αλλά δεν βελτίωσαν σημαντικά το μοντέλο KSI όταν οι υπόλοιπες μεταβλητές λήφθηκαν υπόψη. Η απουσία τους από το μοντέλο KSI δεν σημαίνει

απαραίτητα ότι δεν επηρεάζουν τη σοβαρότητα. Το μοντέλο KSI βασίζεται σε πολύ λιγότερα περιστατικά (604 έναντι 5.981), με αποτέλεσμα μικρότερη στατιστική ισχύ, ενώ ο τύπος χρήστη (p = 0,0895) και ο τύπος οδού (p = 0,1034) αφαιρέθηκαν με τιμές p κοντά στο όριο σημαντικότητας.

Ιδιαίτερη περίπτωση αποτελεί ο **αριθμός εμπλεκόμενων οχημάτων**. Στα δεδομένα η αναλογία KSI είναι 22,2% στα ατυχήματα με ένα όχημα έναντι 6,3% στα ατυχήματα με δύο οχήματα, όμως η μεταβλητή δεν παρέμεινε στο μοντέλο KSI. Η πιθανότερη εξήγηση είναι ότι **τα ατυχήματα ενός οχήματος ταυτίζονται σε μεγάλο βαθμό με τις εκτροπές, τις προσκρούσεις σε εμπόδια και τα ατυχήματα εκτός κατοικημένης περιοχής,** οπότε η επίδρασή τους αποτυπώνεται ήδη μέσω του τύπου ατυχήματος και της κατοικημένης περιοχής. Τέλος, η **εποχή δεν ήταν σημαντική** σε κανένα από τα δύο μοντέλα.

5.6.5 Συνολική αποτίμηση της σύγκρισης

Η σύγκριση των δύο μοντέλων οδηγεί σε μια σαφή διάκριση μεταξύ δύο ομάδων παραγόντων. Η πρώτη ομάδα περιλαμβάνει τους **παράγοντες σοβαρότητας**, δηλαδή τα ατυχήματα εκτός κατοικημένης περιοχής, την ηλικία 65+, τη νύχτα, τις εκτροπές και τις προσκρούσεις σε εμπόδια, τα μεγάλα και τα μικρά δίκυκλα και τον θετικό έλεγχο αλκοόλ. Οι συνθήκες αυτές δεν είναι οι συχνότερες, αλλά όταν συμβαίνει ατύχημα η πιθανότητα σοβαρού ή θανατηφόρου τραυματισμού είναι σημαντικά αυξημένη. Η δεύτερη ομάδα περιλαμβάνει τους **παράγοντες συχνότητας**, όπως οι κόμβοι, ο τύπος ημέρας, ο τύπος χρήστη και ο αριθμός των εμπλεκόμενων οχημάτων, οι οποίοι επηρεάζουν κυρίως το πόσοι παθόντες καταγράφονται.

Η διάκριση αυτή αναδεικνύει και τη σημασία της χωριστής μοντελοποίησης των δύο κατηγοριών σοβαρότητας. Αν είχε εκτιμηθεί ένα μόνο μοντέλο για το σύνολο των παθόντων, οι αντίθετες επιδράσεις θα αλληλοαναιρούνταν σε μεγάλο βαθμό, αφού οι ελαφρά τραυματίες αποτελούν το 90,8% του συνόλου, και οι παράγοντες σοβαρότητας θα εμφανίζονταν ως ασήμαντοι ή ακόμη και ως προστατευτικοί. Για την οδική ασφάλεια, το συμπέρασμα είναι ότι απαιτούνται **δύο διαφορετικές κατευθύνσεις παρεμβάσεων**: μέτρα που μειώνουν τη συχνότητα των ατυχημάτων εκεί όπου συγκεντρώνεται η κυκλοφορία, όπως στους αστικούς κόμβους, και μέτρα που μειώνουν τη σοβαρότητα εκεί όπου η ενέργεια της σύγκρουσης είναι μεγαλύτερη, όπως στο υπεραστικό δίκτυο, τη νύχτα και στους αναβάτες που οδηγούν υπό την επήρεια αλκοόλ.

Συνοψίζοντας, η ανάλυση οδήγησε σε δύο μικτά μοντέλα καταμέτρησης: ένα μοντέλο Poisson για τους νεκρούς και βαριά τραυματίες και ένα μοντέλο Αρνητικής Διωνυμικής

για τους ελαφρά τραυματίες, με τη γεωγραφική περιοχή ως τυχαίο σταθερό όρο και στα δύο. **Αυξημένος αναμενόμενος αριθμός νεκρών και βαριά τραυματιών** συνδέεται κυρίως με τα **ατυχήματα εκτός κατοικημένης περιοχής, την ηλικιακή ομάδα 65+, τη νύχτα**, **τις εκτροπές και τις λοιπές συγκρούσεις, τα μεγάλα και τα μικρά δίκυκλα και τον θετικό έλεγχο αλκοόλ, ενώ μειωμένος με τις γυναίκες παθούσες και τα ατυχήματα σε κόμβο.**

Η ερμηνεία των αποτελεσμάτων πρέπει να γίνεται **πάντοτε σε σχέση με τις αντίστοιχες κατηγορίες αναφοράς** και με τις υπόλοιπες μεταβλητές του μοντέλου σταθερές. Επιπλέον, οι σχέσεις που προέκυψαν είναι **στατιστικές συσχετίσεις** και δεν τεκμηριώνουν από μόνες τους αιτιώδεις επιδράσεις. Τα αποτελέσματα του παρόντος κεφαλαίου αποτελούν τη βάση για τη διατύπωση των συνολικών συμπερασμάτων και των προτάσεων της Διπλωματικής Εργασίας στο επόμενο κεφάλαιο.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Σύνοψη αποτελεσμάτων

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η **διερεύνηση των παραγόντων που σχετίζονται με τη σοβαρότητα των οδικών ατυχημάτων με μηχανοκίνητα δίκυκλα στην Ελλάδα**. Για την επίτευξη του στόχου αυτού αξιοποιήθηκαν δεδομένα οδικών ατυχημάτων για το έτος 2023, τα οποία περιλαμβάνουν πληροφορίες για τα χαρακτηριστικά των παθόντων, των οχημάτων, των ατυχημάτων και των συνθηκών υπό τις οποίες συνέβησαν.

Αρχικά, πραγματοποιήθηκε **βιβλιογραφική ανασκόπηση** (Κεφάλαιο 2), με σκοπό την αναγνώριση των βασικών παραγόντων που επηρεάζουν τη σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα. Στη συνέχεια, παρουσιάστηκε το **θεωρητικό υπόβαθρο** (Κεφάλαιο 3), στο οποίο περιγράφονται οι στατιστικές μέθοδοι και οι απαραίτητοι έλεγχοι για την ανάπτυξη και αξιολόγηση των μοντέλων.

Έπειτα, ακολούθησε η **συλλογή και επεξεργασία των δεδομένων** (Κεφάλαιο 4), με τη διαμόρφωση της τελικής βάσης δεδομένων από τα στοιχεία της ΕΛ.ΣΤΑΤ. Η αρχική βάση περιλάμβανε 6.585 καταγραφές παθόντων, ενώ μετά την απαραίτητη επεξεργασία διαμορφώθηκαν οι βάσεις δεδομένων που χρησιμοποιήθηκαν στη στατιστική ανάλυση, με 4.069 εγγραφές.

Τέλος, στο κεφάλαιο της **εφαρμογής της μεθοδολογίας και της ανάλυσης των αποτελεσμάτων** (Κεφάλαιο 5), πραγματοποιήθηκε η ανάπτυξη των τελικών στατιστικών μοντέλων. Συγκεκριμένα, εφαρμόστηκαν δύο **Γενικευμένα Γραμμικά Μοντέλα Μικτών Επιδράσεων (GLMM)**, με σκοπό τη χωριστή διερεύνηση των παραγόντων που σχετίζονται με τον αναμενόμενο αριθμό παθόντων ανά κατηγορία σοβαρότητας:

- **Μοντέλο παλινδρόμησης Poisson**, με εξαρτημένη μεταβλητή τον αριθμό νεκρών και σοβαρά τραυματιών (KSI).
- **Μοντέλο Αρνητικής Διωνυμικής παλινδρόμησης (Negative Binomial)**, με εξαρτημένη μεταβλητή τον αριθμό ελαφρά τραυματιών.

Και στα δύο μοντέλα συμπεριλήφθηκε η **γεωγραφική περιοχή ως τυχαία επίδραση**, ώστε να ληφθούν υπόψη οι διαφορές μεταξύ των περιοχών της χώρας. Μετά την εφαρμογή των απαραίτητων στατιστικών ελέγχων, προέκυψαν τα τελικά μοντέλα, τα οποία επέτρεψαν την αναγνώριση των σημαντικότερων παραγόντων που συνδέονται με τις διαφορετικές κατηγορίες σοβαρότητας των ατυχημάτων.

Παρακάτω δίνεται ο συγκεντρωτικός πίνακας των αποτελεσμάτων των αναλύσεων για τις μεθόδους που χρησιμοποιήθηκαν και για τα δύο μοντέλα:

Πίνακας 6.1: Συγκεντρωτικός πίνακας αποτελεσμάτων των δύο τελικών μοντέλων

Κατηγορία	Μοντέλο 1: Νεκροί και βαριά τραυματίες (KSI)		Μοντέλο 2: Ελαφρά τραυματίες	
	B	Επίδραση	β	Επίδραση
Κατηγορία αναφοράς: Οδηγός				
Επιβάτης	-	-	-0,344	**Μείωση**
Κατηγορία αναφοράς: Άνδρας				
Γυναίκα	-0,696	**Μείωση**	-0,375	**Μείωση**
Κατηγορία αναφοράς: Μη εθνική οδός				
Εθνική οδός	-	-	-0,250	**Μείωση**
Κατηγορία αναφοράς: Εντός κατοικημένης περιοχής				
Εκτός κατοικημένης περιοχής	0,681	**Αύξηση**	-0,282	**Μείωση**
Κατηγορία αναφοράς: Καθημερινή				
Σαββατοκύριακο	-	-	-0,270	**Μείωση**
Κατηγορία αναφοράς: Ατύχημα εκτός κόμβου				
Ατύχημα σε κόμβο	-0,200	**Μείωση**	0,252	**Αύξηση**
Κατηγορία αναφοράς: Ηλικία 25-49 ετών				
0-14 ετών	-	-	-0,301	*Μη σημαντική*
15-18 ετών	-	-	-0,518	**Μείωση**
≤18 ετών	0,063	*Μη σημαντική*	-	-
19-24 ετών	-0,078	*Μη σημαντική*	-0,366	**Μείωση**
50-64 ετών	0,141	*Μη σημαντική*	-0,283	**Μείωση**
65 ετών και άνω	0,402	**Αύξηση**	-0,675	**Μείωση**
Άγνωστη ηλικία	-0,688	**Μείωση**	-0,739	**Μείωση**
Κατηγορία αναφοράς: Ημέρα				
Νύχτα	0,271	**Αύξηση**	-0,276	**Μείωση**
Κατηγορία αναφοράς: Δύο εμπλεκόμενα οχήματα				
Ένα όχημα	-	-	-0,251	**Μείωση**
Τρία ή περισσότερα οχήματα	-	-	-0,701	**Μείωση**
Κατηγορία αναφοράς: Καλές καιρικές συνθήκες				
Λοιπές καιρικές συνθήκες	-	-	-0,548	**Μείωση**
Κατηγορία αναφοράς: Σύγκρουση με χρήστες της οδού				
Εκτροπή	0,276	**Αύξηση**	-0,281	**Μείωση**
Λοιπές συγκρούσεις	0,447	**Αύξηση**	-0,530	**Μείωση**
Ανατροπή/λοιπά	-0,004	*Μη σημαντική*	-0,434	**Μείωση**
Κατηγορία αναφοράς: Μεσαίο δίκυκλο				
Μικρό δίκυκλο	0,235	**Αύξηση**	-0,246	**Μείωση**
Μεγάλο δίκυκλο	0,343	**Αύξηση**	-0,110	**Μείωση**
Κατηγορία αναφοράς: Αρνητικός/μηδενικός έλεγχος αλκοόλ				
Χαμηλά θετικός έλεγχος αλκοόλ	-	-	-0,450	**Μείωση**
Υψηλά θετικός έλεγχος αλκοόλ	-	-	-0,413	**Μείωση**
Θετικός έλεγχος αλκοόλ	0,648	**Αύξηση**	-	-
Άγνωστος/δεν διενεργήθηκε	1,175	**Αύξηση**	-0,398	**Μείωση**

***Σημείωση**: «-» σημαίνει ότι η μεταβλητή ή η κατηγορία δεν περιλαμβάνεται στο αντίστοιχο τελικό μοντέλο.*

6.2 Συνολικά συμπεράσματα

Κατά την εκπόνηση της παρούσας Διπλωματικής Εργασίας προέκυψε μια σειρά συμπερασμάτων, τα οποία συνδέονται άμεσα με τον αρχικό στόχο της διερεύνησης των παραγόντων που σχετίζονται με τη **σοβαρότητα των ατυχημάτων** με μηχανοκίνητα δίκυκλα. Τα σημαντικότερα συμπεράσματα που προέκυψαν από την περιγραφική και τη στατιστική ανάλυση συνοψίζονται ως εξής:

- Η **σοβαρότητα των ατυχημάτων με μηχανοκίνητα δίκυκλα επηρεάζεται συνδυαστικά από πολλούς διαφορετικούς παράγοντες**, οι οποίοι σχετίζονται με τα χαρακτηριστικά του αναβάτη, τον τύπο του δικύκλου, τις συνθήκες κυκλοφορίας και το οδικό περιβάλλον. Η διεθνής βιβλιογραφία επιβεβαιώνει ότι οι παράγοντες αυτοί συνδέονται με αυξημένη πιθανότητα σοβαρού ή θανατηφόρου τραυματισμού.
- Τα **ατυχήματα εκτός κατοικημένων περιοχών παρουσιάζουν σημαντικά μεγαλύτερη σοβαρότητα** σε σύγκριση με εκείνα εντός κατοικημένων περιοχών. Συγκεκριμένα, σύμφωνα με το μοντέλο Poisson, ο αναμενόμενος αριθμός νεκρών και σοβαρά τραυματιών είναι περίπου **1,98 φορές υψηλότερος** εκτός κατοικημένων περιοχών, όταν οι υπόλοιποι παράγοντες παραμένουν σταθεροί, γεγονός που σχετίζεται με τις υψηλότερες ταχύτητες κυκλοφορίας στο υπεραστικό οδικό δίκτυο.
- Η **ηλικία του αναβάτη αποτελεί σημαντικό παράγοντα σοβαρότητας**, καθώς οι μεγαλύτεροι σε ηλικία αναβάτες παρουσιάζουν αυξημένο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών. Ειδικότερα, η ηλικιακή ομάδα **65 ετών και άνω** εμφανίζει τη μεγαλύτερη αναλογία σοβαρών και θανατηφόρων τραυματισμών, γεγονός που μπορεί να σχετίζεται με τη μεγαλύτερη σωματική ευπάθεια των ηλικιωμένων.
- Η **οδήγηση υπό την επήρεια αλκοόλ συνδέεται με σοβαρότερες συνέπειες** στα ατυχήματα μηχανοκίνητων δικύκλων. Συγκεκριμένα, οι περιπτώσεις με θετικό έλεγχο αλκοόλ παρουσιάζουν περίπου **1,91 φορές υψηλότερο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών** σε σύγκριση με τις περιπτώσεις αρνητικού ή μηδενικού ελέγχου, με τις υπόλοιπες μεταβλητές σταθερές.
- Η **νυχτερινή οδήγηση συνδέεται με αυξημένη σοβαρότητα των ατυχημάτων**. Παρότι κατά τις νυχτερινές και πρώτες πρωινές ώρες καταγράφονται λιγότεροι παθόντες, το ποσοστό των σοβαρών και θανατηφόρων τραυματισμών είναι υψηλότερο. Το εύρημα αυτό ενδέχεται να

σχετίζεται με τη μειωμένη ορατότητα, τις υψηλότερες ταχύτητες και την πιθανή κατανάλωση αλκοόλ.

- Ο **τύπος του ατυχήματος παρουσιάζει σημαντική συσχέτιση με τη σοβαρότητα των τραυματισμών**. Ειδικότερα, οι προσκρούσεις σε σταθερά εμπόδια, οι εκτροπές και οι λοιπές συγκρούσεις εμφανίζουν αυξημένη σοβαρότητα. Το αποτέλεσμα αυτό αναδεικνύει τη σημασία των συνθηκών της σύγκρουσης και των χαρακτηριστικών του οδικού περιβάλλοντος.
- Ο **τύπος και ο κυβισμός του δικύκλου σχετίζονται με τη σοβαρότητα των ατυχημάτων**. Σύμφωνα με τα αποτελέσματα του μοντέλου, τόσο τα μικρά όσο και τα μεγάλα μηχανοκίνητα δίκυκλα εμφανίζουν υψηλότερο αναμενόμενο αριθμό νεκρών και σοβαρά τραυματιών σε σύγκριση με τα μεσαίου κυβισμού δίκυκλα. Ενδεχομένως αυτές οι δύο κατηγορίες οχημάτων εμφανίζουν τη μεγαλύτερη απόκλιση από τη μέση ταχύτητα κυκλοφορίας.
- Η **γεωγραφική περιοχή αποτελεί σημαντικό παράγοντα διαφοροποίησης των αποτελεσμάτων**, καθώς η συμπερίληψή της ως τυχαίας επίδρασης βελτίωσε σημαντικά και τα δύο στατιστικά μοντέλα. Μάλιστα, οι γεωγραφικές διαφοροποιήσεις ήταν εντονότερες στο μοντέλο των νεκρών και σοβαρά τραυματιών. Το εύρημα αυτό υποδηλώνει ότι υπάρχουν ιδιαίτερα χαρακτηριστικά μεταξύ των περιοχών, τα οποία δεν εξηγούνται πλήρως από τις διαθέσιμες μεταβλητές και απαιτούν περαιτέρω διερεύνηση.
- Τέλος, η ανάλυση ανέδειξε ότι **ο αριθμός των καταγεγραμμένων παθόντων δεν αποτελεί από μόνος του ένδειξη της σοβαρότητας των ατυχημάτων**. Ορισμένες συνθήκες, όπως η οδήγηση εκτός κατοικημένων περιοχών και κατά τις νυχτερινές ώρες, συνδέονται με λιγότερες καταγραφές αλλά υψηλότερη αναλογία σοβαρών και θανατηφόρων τραυματισμών. Συνεπώς, για την αποτελεσματικότερη βελτίωση της οδικής ασφάλειας απαιτείται η χωριστή εξέταση της συχνότητας και της σοβαρότητας των ατυχημάτων.

6.3 Προτάσεις βελτίωσης της οδικής ασφάλειας

Με βάση τα **αποτελέσματα** της στατιστικής ανάλυσης και τα συμπεράσματα της βιβλιογραφικής ανασκόπησης, διατυπώνονται οι ακόλουθες προτάσεις για τη βελτίωση της οδικής ασφάλειας των μηχανοκίνητων δικύκλων, με στόχο τη μείωση των σοβαρών και θανατηφόρων τραυματισμών:

- **Στοχευμένη διαχείριση της ταχύτητας και βελτίωση των οδικών υποδομών:** Δεδομένου ότι τα ατυχήματα εκτός κατοικημένων περιοχών παρουσιάζουν

υψηλότερα ποσοστά σοβαρών και θανατηφόρων τραυματισμών, προτείνεται η ενίσχυση των ελέγχων ταχύτητας και η βελτίωση της ασφάλειας σε αυτές τις οδούς. Ειδικότερα, η τοποθέτηση κατάλληλης σήμανσης, η βελτίωση της διαγράμμισης και του φωτισμού, καθώς και η εγκατάσταση προστατευτικών συστημάτων φιλικών προς τους αναβάτες δικύκλων μπορούν να περιορίσουν τις σοβαρές συνέπειες των ατυχημάτων.

- **Ενίσχυση της χρήσης κράνους και προστατευτικού εξοπλισμού:** Η πραγματοποίηση συστηματικών ελέγχων για τη χρήση εγκεκριμένου και σωστά δεμένου κράνους, σε συνδυασμό με ενημερωτικές εκστρατείες, μπορεί να συμβάλει στη μείωση των σοβαρών τραυματισμών. Παράλληλα, προτείνεται η προώθηση της χρήσης κατάλληλου προστατευτικού εξοπλισμού, όπως ειδικά μπουφάν, γάντια και προστατευτικά υποδήματα, ιδιαίτερα στους νεότερους αναβάτες.
- **Στοχευμένα προγράμματα εκπαίδευσης και ενημέρωσης αναβατών:** Η ανάπτυξη εκπαιδευτικών προγραμμάτων για την ασφαλή οδήγηση δικύκλων μπορεί να συμβάλει στη βελτίωση της οδηγικής συμπεριφοράς. Τα προγράμματα μπορούν να επικεντρώνονται στην αναγνώριση επικίνδυνων καταστάσεων, στη σωστή διαχείριση της ταχύτητας και στην ασφαλή αντιμετώπιση έκτακτων συνθηκών.
- **Βελτίωση της ορατότητας των δικύκλων και αξιοποίηση νέων τεχνολογιών:** Η βελτίωση των συνθηκών ορατότητας στις διασταυρώσεις και στους οδικούς κόμβους, μέσω κατάλληλης σήμανσης και απομάκρυνσης εμποδίων που περιορίζουν το οπτικό πεδίο, μπορεί να συμβάλει στην αποφυγή συγκρούσεων. Παράλληλα, η αξιοποίηση **προηγμένων συστημάτων υποβοήθησης οδηγού (ADAS)**, τα οποία μπορούν να εντοπίζουν μηχανοκίνητα δίκυκλα και να προειδοποιούν τους οδηγούς για πιθανή σύγκρουση, αποτελεί σημαντική κατεύθυνση για την πρόληψη των ατυχημάτων.
- **Στοχευμένες παρεμβάσεις σε περιοχές υψηλού κινδύνου:** Η αξιοποίηση των γεωγραφικών **δεδομένων** των ατυχημάτων μπορεί να συμβάλει στον εντοπισμό περιοχών όπου απαιτούνται παρεμβάσεις για τη βελτίωση της οδικής ασφάλειας. Προτείνεται η συστηματική καταγραφή και αξιολόγηση των ατυχημάτων ανά περιοχή, σε συνδυασμό με στοιχεία κυκλοφορίας και χαρακτηριστικά του οδικού δικτύου, ώστε να καθορίζονται οι προτεραιότητες για την υλοποίηση μέτρων, όπως η βελτίωση της οδικής υποδομής, η ενίσχυση της αστυνόμευσης και η αντιμετώπιση επικίνδυνων σημείων.

6.4 Προτάσεις για περαιτέρω έρευνα

Για την επέκταση της παρούσας Διπλωματικής μου Εργασίας και την περαιτέρω διερεύνηση του αντικειμένου της, ενδιαφέρον παρουσιάζουν οι ακόλουθες προτάσεις:

- **Επέκταση της ανάλυσης σε περισσότερα έτη:** Η αξιοποίηση δεδομένων οδικών ατυχημάτων από μεγαλύτερη χρονική περίοδο θα μπορούσε να επιτρέψει τη μελέτη της διαχρονικής εξέλιξης των σοβαρών και θανατηφόρων τραυματισμών. Παράλληλα, θα ήταν δυνατή η εξέταση της σταθερότητας των παραγόντων που επηρεάζουν τη **σοβαρότητα των ατυχημάτων**, καθώς και ο εντοπισμός πιθανών μεταβολών με την πάροδο του χρόνου.
- **Ενσωμάτωση στοιχείων κυκλοφορίας και χρήσης δικύκλων:** Η προσθήκη **δεδομένων**, όπως ο κυκλοφοριακός φόρτος, τα διανυόμενα οχηματοχιλιόμετρα και ο αριθμός κυκλοφορούντων δικύκλων ανά περιοχή, θα μπορούσε να οδηγήσει σε πιο αξιόπιστες συγκρίσεις. Με αυτόν τον τρόπο, θα λαμβάνεται υπόψη η διαφορετική κυκλοφορία δικύκλων μεταξύ των περιοχών και θα είναι δυνατή η καλύτερη εκτίμηση του κινδύνου σοβαρού ατυχήματος.
- **Εξέταση επιπλέον παραγόντων οδικής ασφάλειας:** Η ενσωμάτωση πρόσθετων μεταβλητών, όπως η **χρήση κράνους, η ταχύτητα κίνησης, οι συνθήκες φωτισμού και η κατάσταση του οδοστρώματος**, θα μπορούσε να συμβάλει στην πληρέστερη κατανόηση των συνθηκών που επηρεάζουν τη **σοβαρότητα των ατυχημάτων**. Επιπλέον, η εξέταση της γεωμετρίας της οδού, της ύπαρξης παρόδιων εμποδίων και του χρόνου παροχής ιατρικής βοήθειας θα επέτρεπε τη διερεύνηση περισσότερων παραγόντων, εφόσον υπάρχουν διαθέσιμα στοιχεία.
- **Περαιτέρω διερεύνηση της επίδρασης της γεωγραφικής περιοχής:** Η λεπτομερέστερη ανάλυση των γεωγραφικών διαφορών στη **σοβαρότητα των ατυχημάτων** με μηχανοκίνητα δίκυκλα θα μπορούσε να συμβάλει στην καλύτερη κατανόηση των παραγόντων που τις προκαλούν. Ειδικότερα, προτείνεται η εξέταση χαρακτηριστικών όπως η ποιότητα των οδικών υποδομών, ο κυκλοφοριακός φόρτος, η μορφολογία του εδάφους και η πρόσβαση σε υπηρεσίες επείγουσας ιατρικής βοήθειας. Με τον τρόπο αυτό, θα μπορούσαν να εντοπιστούν οι αιτίες των διαφοροποιήσεων μεταξύ των περιοχών και να προταθούν πιο στοχευμένα μέτρα βελτίωσης της οδικής ασφάλειας.

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Abdi, N., Robertson, T., Petrucka, P., & Crizzle, A. M. (2022). Do motorcycle helmets reduce road traffic injuries, hospitalizations and mortalities in low and lower-middle income countries in Africa? A systematic review and meta-analysis. *BMC Public Health, 22*, 824. https://doi.org/10.1186/s12889-022-13138-4
2. Adeel, M., Ahmad, N., Wali, B., & Khattak, A. J. (2025). Exploring the influence of rider age and experience on motorcycle crash risk: Evidence from a case-control study. *Journal of Transportation Safety & Security, 17*(5). https://doi.org/10.1080/19439962.2024.2435297
3. Agresti, A. (2018). *An introduction to categorical data analysis* (3rd ed.). Wiley.
4. Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences* (4th ed.). Pearson Prentice Hall.
5. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01
6. Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*(3), 127-135. https://doi.org/10.1016/j.tree.2008.10.008
7. Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal, 9*(2), 378-400. https://doi.org/10.32614/RJ-2017-066
8. Brown, L., Morris, A., Thomas, P., Ekambaram, K., Margaritis, D., Davidse, R., Usami, D. S., Robibaro, M., Persia, L., Buttler, I., Ziakopoulos, A., Theofilatos, A., Yannis, G., Martin, A., & Wadji, F. (2021). Investigation of accidents involving powered two wheelers and bicycles: A European in-depth study. *Journal of Safety Research, 76*, 135-145. https://doi.org/10.1016/j.jsr.2020.12.015
9. Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer. https://doi.org/10.1007/b97636
10. Chang, F., Li, M., Xu, P., Zhou, H., Haque, M. M., & Huang, H. (2016). Injury severity of motorcycle riders involved in traffic crashes in Hunan, China: A mixed ordered logit approach. *International Journal of Environmental Research and Public Health, 13*(7), 714. https://doi.org/10.3390/ijerph13070714
11. Chouhan, S. S., Kathuria, A., & Chalumuri, R. S. (2024). Powered two-wheeler riding behavior and strategies to improve safety: A review. *Journal of Traffic and Transportation Engineering (English Edition), 11*(6). https://doi.org/10.1016/j.jtte.2024.03.006

12. Damani, J., & Vedagiri, P. (2021). Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors. *Journal of Traffic and Transportation Engineering (English Edition), 8*(1), 35-56. https://doi.org/10.1016/j.jtte.2020.12.003

13. Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.

14. Hartig, F. (2022). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models* (R package version 0.4.6). https://CRAN.R-project.org/package=DHARMa

15. Hurt, H. H., Ouellet, J. V., & Thom, D. R. (1981). *Motorcycle accident cause factors and identification of countermeasures, Volume 1: Technical report*. Traffic Safety Center, University of Southern California, for the National Highway Traffic Safety Administration, U.S. Department of Transportation.

16. Liu, B. C., Ivers, R., Norton, R., Boufous, S., Blows, S., & Lo, S. K. (2008). Helmets for preventing injury in motorcycle riders. *Cochrane Database of Systematic Reviews*, 2008(1), CD004333. https://doi.org/10.1002/14651858.CD004333.pub3

17. Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. SAGE Publications.

18. NTUA Road Safety Observatory. (2026a). *Basic road safety figures, Greece 2015-2025*. https://www.nrso.ntua.gr/wp-content/uploads/nrso-GR_RoadSafety_basic_figures_2015-2025-2.html

19. NTUA Road Safety Observatory. (2026b). *Data warehouse and dashboards: Greece road safety data*. https://www.nrso.ntua.gr/data/

20. Robbins, C. J., Allen, H. A., Miller, K. A., & Chapman, P. (2019). The 'Saw but Forgot' error: A role for short-term memory failures in understanding junction crashes? *PLOS ONE, 14*(9), e0222905. https://doi.org/10.1371/journal.pone.0222905

21. Santos, K., Dias, J. P., Amado, C., Sousa, J., & Francisco, P. (2021). Risk factors associated with the increase of injury severity of powered two wheelers road accidents victims in Portugal. *Traffic Injury Prevention, 22*(8), 646-650. https://doi.org/10.1080/15389588.2021.1987421

22. Savolainen, P., & Mannering, F. (2007). Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. *Accident Analysis & Prevention, 39*(5), 955-963. https://doi.org/10.1016/j.aap.2006.12.016

23. Theofilatos, A., & Yannis, G. (2015). A review of powered-two-wheeler behaviour and safety. *International Journal of Injury Control and Safety Promotion, 22*(4), 284-307. https://doi.org/10.1080/17457300.2014.908224

24. Theofilatos, A., & Ziakopoulos, A. (2018). Examining injury severity of moped and motorcycle occupants with real-time traffic and weather data. *Journal of Transportation Engineering, Part A: Systems, 144*(11). https://doi.org/10.1061/JTEPBS.0000193

25. Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.

26. Vlahogianni, E. I., Yannis, G., & Golias, J. C. (2012). Overview of critical risk factors in Power-Two-Wheeler safety. *Accident Analysis & Prevention, 49*, 12-22. https://doi.org/10.1016/j.aap.2012.04.009

27. World Health Organization. (2023). *Global status report on road safety 2023*. World Health Organization.

28. World Health Organization. (2026). *Road traffic injuries* [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

29. Yannis, G., Golias, J., & Papadimitriou, E. (2005). Driver age and vehicle engine size effects on fault and severity in young motorcyclists accidents. *Accident Analysis & Prevention, 37*(2), 327-333. https://doi.org/10.1016/j.aap.2004.10.003

30. Zhang, Y., Zhang, G., Xu, J., Li, S., & Liu, Y. (2026). Which factors influence rider injury severity in powered two-wheeler accidents? A systematic review. *Traffic Injury Prevention*. Advance online publication. https://doi.org/10.1080/15389588.2025.2592108

31. Ελληνική Στατιστική Αρχή (ΕΛ.ΣΤΑΤ.). (2026). *Οδικά τροχαία ατυχήματα: Έτος 2024* [Δελτίο Τύπου]. https://www.statistics.gr/documents/20181/873e9edb-2f7c-523b-5647-47010bd751ad