

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**  
**Σχολή Χρηματοοικονομικής και Στατιστικής**



**Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ  
ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΩΝ**

**ΣΥΓΚΡΙΤΙΚΗ ΜΕΛΕΤΗ ΣΤΟΧΑΣΤΙΚΩΝ  
ΜΟΝΤΕΛΩΝ ΘΝΗΣΙΜΟΤΗΤΑΣ ΚΑΙ  
ΕΦΑΡΜΟΓΕΣ ΜΕ ΣΤΟΙΧΕΙΑ ΤΟΥ  
ΠΛΗΘΥΣΜΟΥ ΤΗΣ ΕΛΛΑΔΑΣ**

**Ηρακλής-Νικόλαος Καραΐσκος**

**Διπλωματική Εργασία**

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής  
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των  
απαιτήσεων για την απόκτηση του Μεταπτυχιακού  
Διπλώματος Ειδίκευσης στην *Αναλογιστική Επιστήμη και  
Διαχείριση Κινδύνων*

Πειραιάς

Μάρτιος 2023

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**  
**Σχολή Χρηματοοικονομικής και Στατιστικής**



**Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ  
ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΩΝ**

**ΣΥΓΚΡΙΤΙΚΗ ΜΕΛΕΤΗ ΣΤΟΧΑΣΤΙΚΩΝ  
ΜΟΝΤΕΛΩΝ ΘΝΗΣΙΜΟΤΗΤΑΣ ΚΑΙ  
ΕΦΑΡΜΟΓΕΣ ΜΕ ΣΤΟΙΧΕΙΑ ΤΟΥ  
ΠΛΗΘΥΣΜΟΥ ΤΗΣ ΕΛΛΑΔΑΣ**

**Ηρακλής-Νικόλαος Καραϊσκος**

**Διπλωματική Εργασία**

*που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής  
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των  
απαιτήσεων για την απόκτηση του Μεταπτυχιακού  
Διπλώματος Ειδίκευσης στην Αναλογιστική Επιστήμη και  
Διαχείριση Κινδύνων*

Πειραιάς

Μάρτιος 2023



Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη Συνέλευση του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. 22/7-7-2021 συνεδρίασή της, σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διαχείριση Κινδύνων.

Τα μέλη της Επιτροπής ήταν:

-Καθηγήτρια κα Γεωργία Βερροπούλου (Επιβλέπων)

-Αναπληρωτής Καθηγητής κ. Γεώργιος Ψαρράκος

- Επίκουρος Καθηγητής κ. Γεώργιος Πιτσέλης

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

**UNIVERSITY OF PIRAEUS**  
**School of Finance and Statistics**



**Department of Statistics and Insurance Science**

**POSTGRADUATE PROGRAM IN  
ACTUARIAL SCIENCE AND RISK MANAGEMENT**

**A COMPARATIVE STUDY OF  
STOCHASTIC MORTALITY MODELS  
AND APPLICATIONS USING GREEK  
POPULATION DATA.**

By  
**Iraklis-Nikolaos Karaiskos**

MSc Dissertation  
submitted to the Department of Statistics and Insurance  
Science of the University of Piraeus in partial fulfilment of  
the requirements for the degree of Master of Science in  
Actuarial Science and Risk Management

Piraeus, Greece  
March 2023



*Στους γονείς μου*  
*Δημήτρη και Αναστασία*





# Ευχαριστίες

Ένα κεφάλαιο κλείνει και δοθείσης της ευκαιρίας, θα ήθελα να ευχαριστήσω τους ανθρώπους που συνέβαλαν σε αυτό.

Αρχικά, ένα ειλικρινές ευχαριστώ στην επιβλέπουσα, Καθηγήτρια κ. Βερροπούλου Γεωργία, για την άψογη συνεργασία, την καθοδήγηση, το χρόνο και το ενδιαφέρον της για την ολοκλήρωση της διπλωματικής μου εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, που μου προσφέρει απλόχερα αγάπη, υποστήριξη και εμπιστοσύνη.



# Περίληψη

Η παρούσα διπλωματική εργασία έχει ως βασικό αντικείμενο πραγμάτευσης τα στοχαστικά μοντέλα θνησιμότητας και την εφαρμογή του μοντέλου Lee-Carter στον ελληνικό πληθυσμό. Το ερευνητικό αντικείμενο προσεγγίστηκε με διττό τρόπο, τόσο βιβλιογραφικά όσο και εμπειρικά. Σε επίπεδο βιβλιογραφικής πραγμάτευσης ορίστηκαν πρωτίστως οι βασικές έννοιες της θνησιμότητας και της πρόβλεψης ενώ παρατέθηκε αναλυτικά το μοντέλο Lee-Carter, το οποίο σχολιάστηκε κριτικά. Εκτός από το εν λόγω μοντέλο δόθηκαν περιγραφές και αναλύθηκαν τα χαρακτηριστικά και άλλων μοντέλων πρόβλεψης θνησιμότητας. Στην συνέχεια ακολούθησε εμπειρική μελέτη της θνησιμότητας του Ελληνικού πληθυσμού. Συγκεκριμένα, για την μελέτη χρησιμοποιήθηκαν δεδομένα θνησιμότητας κατά ηλικία (ηλικίες 0-108 έτη) για την περίοδο 1981-2019 και εκτιμήθηκε η θνησιμότητα για την περίοδο 2020-2039. Η ανάλυση, η μοντελοποίηση και η εξαγωγή των αποτελεσμάτων έγινε με τη χρήση του στατιστικού περιβάλλοντος R, version 4.2.2. Τα αποτελέσματα που εξήχθησαν από την εφαρμογή του στοχαστικού μοντέλου έδειξαν μικρή μείωση των επιπέδων θνησιμότητας και 66.6% εξηγητική ισχύ της μεταβλητότητας για τον γενικό πληθυσμό. Τα μέτρα σφάλματος της μοντελοποίησης ήταν σχετικά χαμηλά ( $MSE=0.001$  και  $MAPE=0.09$ ).

Λέξεις Κλειδιά: Στοχαστικό μοντέλο, θνησιμότητα, Lee-Carter

## **Abstract**

The two core subjects of this thesis are the stochastic mortality models and the application of the Lee-Carter model to the Greek population. The research objectives were approached in a twofold way, both from a bibliographical and an empirical viewpoint. At the level of bibliographical discussion, the basic concepts of mortality and its prediction were initially defined, while the Lee-Carter model was presented in meticulous detail, and, then, was critically commented upon. Further, descriptions of the characteristics of other stochastic models of mortality prediction were also presented. This was followed by the application of the Lee-Carter model on empirical mortality data of the Greek population. More specifically, yearly mortality data by age (ages 0-108) for the period 1981-2019 were used to predict mortality for the period 2020-2039. The analysis, modeling and extraction of the results were implemented using the statistical environment R version 4.2.2. The results derived from the application of the stochastic model, showed a slight decline in mortality for the abovementioned period and a 66.6% explanatory power of the variability for the general population. The measures of modeling error were relatively low (MSE=0.001 and MAPE=0.09).

**Keywords:** Stochastic models, mortality, Lee-Carter method

## Περιεχόμενα

I. Κατάλογος Διαγραμμάτων .....	1
II. Κατάλογος Πινάκων .....	2
III. Κατάλογος Σχημάτων .....	2
Κεφάλαιο 1 <sup>ο</sup> : Εισαγωγή.....	3
Κεφάλαιο 2 <sup>ο</sup> : Θνησιμότητα σύγχρονες και μελλοντικές τάσεις .....	6
2.1 Θνησιμότητα: ορισμός και τρόποι μέτρησης .....	7
2.1.1 Ακαθάριστο ποσοστό θνησιμότητας .....	7
2.1.2 Ποσοστά θνησιμότητας ανάλογα με την ηλικία .....	8
2.1.3 Σφάλματα στις περιοδικές μετρήσεις .....	10
2.1.4 Διαφορές μεταξύ των μετρήσεων.....	11
2.1.5 Η σχέση μεταξύ των δύο τύπων ποσοστού θνησιμότητας.....	13
2.1.6 Πλεονεκτήματα και μειονεκτήματα των δύο ποσοστών θνησιμότητας .....	16
2.2 Πίνακες επιβίωσης .....	17
2.2.1 Η αρχική τιμή .....	18
2.2.2 Η αβεβαιότητα.....	20
2.2.3 Ο πιθανολογικός πίνακας επιβίωσης .....	21
2.2.4 Ο έλεγχος σημαντικότητας του μοντέλου.....	23
2.3 Σύγχρονες τάσεις θνησιμότητας και επιπτώσεις .....	24
2.3.1 Προσδόκιμο ζωής κατά την γέννηση.....	25
2.3.2 Αριθμός θανάτων.....	25
2.3.3 Προσδόκιμο ζωής κατά την γέννα.....	26
2.3.4 Προσδόκιμο ζωής άνω των 65 ετών .....	29
2.3.5 Διαφοροποιήσεις στο προσδόκιμο ζωής βάσει φύλου .....	31
2.3.6 Παιδική θνησιμότητα.....	33
2.4 Μελλοντικές τάσεις θνησιμότητας: Χρησιμότητα και επιπτώσεις στα ασφαλιστικά προγράμματα .....	35

Κεφάλαιο 3 <sup>ο</sup> : Το μοντέλο Lee-Carter.....	40
3.1 Το βασικό μοντέλο Lee-Carter .....	41
3.1.1 Προσαρμογή του μοντέλου.....	41
3.1.2 Διορθώσεις της πρώτης προσαρμογής.....	43
3.1.3 Πρόβλεψη ποσοστών θανάτου ανά ηλικία .....	43
3.1.4 Προσδόκιμο ζωής κατά τη γέννηση .....	44
3.2 Μοντελοποίηση Poisson για τον αριθμό των θανάτων και μεθοδολογία Lee-Carter.....	44
3.2.1 Μοντέλο .....	45
3.2.2 Εκτίμηση μέγιστης πιθανοφάνειας.....	46
3.3 Κριτικές παρατηρήσεις.....	48
Κεφάλαιο 4 <sup>ο</sup> : Προσεγγίσεις πρόβλεψης .....	51
4.1 Προσδοκία .....	51
4.2 Παρέκταση.....	54
4.3 Επεξήγηση .....	57
4.4 Λειτουργίες παραμετροποίησης.....	58
4.4.1 Μοντέλο Heligman-Pollard.....	59
4.4.2 Πολυεκθετικό μοντέλο.....	60
4.4.3 Άλλα μοντέλα.....	60
4.5 Μοντέλα Γενικής Γραμμικής Μοντελοποίησης .....	63
4.6 Μοντέλα Κοορτής.....	65
4.7 Αποσύνθεση των δεδομένων ανά αιτία θανάτου .....	68
4.8 Θνησιμότητα γήρατος .....	70
4.9 Η εκτίμηση της αβεβαιότητας .....	72
4.9.1 Πηγές σφάλματος .....	73
4.9.2 Συσχετίσεις μεταξύ των σφαλμάτων πρόβλεψης .....	74
4.9.3 Εκ των προτέρων εκτίμηση σφαλμάτων βάσει μοντέλου.....	75
4.9.4 Εκ των προτέρων εκτίμηση σφάλματος βάσει εμπειρογνομόνων .....	77

4.9.5 Εκ των υστέρων εκτίμηση σφάλματος .....	78
Κεφάλαιο 5 <sup>ο</sup> : Μεθοδολογία .....	80
5.1 Προδιαγραφές του μοντέλου Lee-Carter.....	80
5.2 Προσαρμογή του μοντέλου .....	81
5.3 Προβλέψεις του μοντέλου .....	82
Κεφάλαιο 6 <sup>ο</sup> : Αποτελέσματα .....	84
6.1 Εφαρμογή στο συνολικό πληθυσμό .....	84
6.1.1 Μελλοντικές Προβλέψεις (Συνολικός Πληθυσμός).....	86
6.2 Εφαρμογή στον Ανδρικό πληθυσμό .....	87
6.2.1 Μελλοντικές Προβλέψεις (Άντρες).....	90
6.3 Εφαρμογή στο Γυναικείο πληθυσμό.....	90
6.3.1 Μελλοντικές Προβλέψεις (Γυναίκες).....	93
6.4 Δείκτες $a_x, b_x$ και $k_t$ βάσει φύλου.....	93
Κεφάλαιο 7 <sup>ο</sup> : Συμπεράσματα.....	95
Βιβλιογραφία.....	97

## I. Κατάλογος Διαγραμμάτων

Διάγραμμα 1: Διάγραμμα Lexis.....	122
Διάγραμμα 2: Παράδειγμα διαγράμματος Lexis.....	13
Διάγραμμα 3: Διάγραμμα Lexis που απεικονίζει τη σχέση μεταξύ δύο ρύπων ποσοστού θνησιμότητας.....	14
Διάγραμμα 4: Αριθμοί Καναδών γυναικών του παρατηρούμενου πληθυσμού και του στάσιμου πληθυσμού υποθετικής κοορτής .....	19
Διάγραμμα 5: Προσδόκιμο ζωής κατά τη γέννηση στην ΕΕ, 2002-2020 .....	25
Διάγραμμα 6: Αριθμός θανάτων στην ΕΕ, 1962-2020.....	266
Διάγραμμα 7: Προσδόκιμο ζωής κατά τη γέννηση, ΕΕ, 2002-2020.....	277
Διάγραμμα 8: Προσδόκιμο ζωής κατά τη γέννηση, αποκλίσεις μεταξύ των φύλων, 2020	32
Διάγραμμα 9: Προσδόκιμο ζωής στην ηλικία των 65 ετών, αποκλίσεις μεταξύ των φύλων, 2020 .....	33
Διάγραμμα 10: Βρεφική θνησιμότητα 2009 και 2020 (θάνατοι ανά 100 γεννήσεις ζώντων) .....	34
Διάγραμμα 11: Δεδομένα θνησιμότητας για τον γενικό πληθυσμό στην Ελλάδα (1981-2019) .....	83
Διάγραμμα 12: Απεικονίσεις συντελεστών $a_x, b_x$ και $k_t$ για τον γενικό πληθυσμό ....	84
Διάγραμμα 13: Παρατηρούμενες τιμές έναντι προβλέψεων στον συνολικό πληθυσμό για 1981 και 2019 .....	85
Διάγραμμα 14: Προβλέψεις για τον γενικό πληθυσμό (2019-2039).....	86
Διάγραμμα 15: Δεδομένα θνησιμότητας για τον αντρικό πληθυσμό στην Ελλάδα (1981-2019) .....	86
Διάγραμμα 16: Απεικονίσεις συντελεστών $a_x, b_x$ και $k_t$ για τον αντρικό πληθυσμό..	87
Διάγραμμα 17: Παρατηρούμενες τιμές έναντι προβλέψεων στον αντρικό πληθυσμό για 1981 και 2019 .....	88
Διάγραμμα 18: Προβλέψεις μοντέλου για αντρικό πληθυσμό (2019-2039).....	89
Διάγραμμα 19: Δεδομένα θνησιμότητας για τον γυναικείο πληθυσμό στην Ελλάδα (1981-2019) .....	90
Διάγραμμα 20: Απεικονίσεις συντελεστών $a_x, b_x$ και $k_t$ για τον γυναικείο πληθυσμό	91



Διάγραμμα 21: Παρατηρούμενες τιμές έναντι προβλέψεων στον αντρικό πληθυσμό για 1981 και 2019 .....	91
Διάγραμμα 22: Προβλέψεις μοντέλου για γυναικείο πληθυσμό (2019-2039) .....	92
Διάγραμμα 23: Δείκτες ax,bx και kt βάσει φύλου για ηλικίες 40-90.....	93

## II. Κατάλογος Πινάκων

Πίνακας 1: Πλεονεκτήματα και μειονεκτήματα των δύο τύπων ποσοστού θνησιμότητας	17
Πίνακας 2: Προσδόκιμο ζωής κατά τη γέννηση 1980-2020 .....	28
Πίνακας 3: Προσδόκιμο ζωής σε ηλικία 65 ετών 1980-2020.....	30
Πίνακας 4: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του συνολικού πληθυσμού (1981-2019).....	84
Πίνακας 5: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του αντρικού πληθυσμού (1981-2019).....	87
Πίνακας 6: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του γυναικείου πληθυσμού (1981-2019).....	90

## III. Κατάλογος Σχημάτων

Σχήμα 1: Προσδόκιμο ζωής κατά τη γέννηση ανά περιοχή NUTS 2, 2020.....	29
Σχήμα 2: Προσδόκιμο ζωής σε ηλικία 65 ετών ανά περιοχή NUTS 2, 2019.....	31

## Κεφάλαιο 1<sup>ο</sup>: Εισαγωγή

Οι προβλέψεις θνησιμότητας είναι απαραίτητες για την πρόβλεψη της μελλοντικής έκτασης της γήρανσης του πληθυσμού, και για τον προσδιορισμό της βιωσιμότητας των συνταξιοδοτικών συστημάτων και των συστημάτων κοινωνικής ασφάλισης. Είναι επίσης χρήσιμες για τον καθορισμό των ασφαλιστρών ασφάλισης ζωής και για να βοηθήσουν τις κυβερνήσεις να σχεδιάσουν τις μεταβαλλόμενες ανάγκες των κοινωνιών τους για υγειονομική περίθαλψη και άλλες υπηρεσίες (Ευρωπαϊκή Επιτροπή 2009). Στην σύγχρονη εποχή, η κοινωνική σημασία των ακριβών προβλέψεων θνησιμότητας είναι μεγαλύτερη από ποτέ άλλοτε. Ως στρατηγική για την αντιμετώπιση της ταχείας πληθυσμιακής γήρανσης, οι πρόσφατες συνταξιοδοτικές μεταρρυθμίσεις σε αρκετές χώρες με χαμηλή θνησιμότητα έχουν κάνει μια ρητή σύνδεση μεταξύ της ηλικίας συνταξιοδότησης και/ή των συνταξιοδοτικών πληρωμών και των παρελθουσών και μελλοντικών αναμενόμενων τιμών θνησιμότητας και προσδόκιμου ζωής (Carone et al. 2016). Λόγω της μεγάλης και αυξανόμενης κοινωνικής σημασίας των ακριβών προβλέψεων θνησιμότητας, ο τομέας των προβλέψεων θνησιμότητας αναπτύσσεται και εξελίσσεται.

Για να αντιμετωπιστούν οι προκλήσεις που συνδέονται με τη γήρανση του πληθυσμού, έχουν αναπτυχθεί τις τελευταίες δεκαετίες πολυάριθμες μέθοδοι για τη μοντελοποίηση και την πρόβλεψη της θνησιμότητας (Tabeau, 2001; Booth & Tickle, 2008; Girosi & King, 2008). Η πλειονότητα των υφιστάμενων σήμερα μεθόδων πρόβλεψης της θνησιμότητας, συμπεριλαμβανομένων των μεθόδων πρόβλεψης της θνησιμότητας που χρησιμοποιούνται από τις στατιστικές υπηρεσίες ανά τον κόσμο, είναι εξωστρεφείς (Booth & Tickle, 2008; Stoeldraijer et al., 2013). Αυτές οι εξωστρεφείς προσεγγίσεις χρησιμοποιούν την κανονικότητα που παρατηρείται τόσο στα ηλικιακά πρότυπα όσο και στις τάσεις της θνησιμότητας με την πάροδο του χρόνου (Booth & Tickle, 2008) και θεωρούνται πιο αντικειμενικές, πιο εύκολες στην εφαρμογή και πιο πιθανό να οδηγήσουν σε ακριβείς προβλέψεις από τους άλλους δύο τύπους προσεγγίσεων για την πρόβλεψη της θνησιμότητας: τις προσεγγίσεις εξήγησης (πρόβλεψη θνησιμότητας με βάση την αιτία θανάτου ή με ένα επεξηγηματικό μοντέλο) και τις προσεγγίσεις εκτίμησης (που βασίζονται στις γνώμες των εμπειρογνομόνων) (Booth & Tickle, 2008).

Για πολλά χρόνια, η μεθοδολογία πρόβλεψης θνησιμότητας Lee-Carter (Lee & Carter, 1992) αποτελεί τη μέθοδο αναφοράς για την εξωστρεφή πρόβλεψη της θνησιμότητας (Stoeldraijer et al., 2013). Το μοντέλο θνησιμότητας Lee-Carter με βάση την ηλικιακή περίοδο αναλύει την ειδική για την ηλικία θνησιμότητα (καταγεγραμμένη) κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου για έναν ενιαίο πληθυσμό στη συνολική χρονική τάση, το ηλικιακό πρότυπο της θνησιμότητας και τις ειδικές για την ηλικία διαφορές στην έκταση της συνολικής χρονικής μεταβολής. Περιλαμβάνεται ένας όρος σφάλματος για να συλλάβει τις επιδράσεις της ηλικιακής περιόδου που δεν καταγράφονται από το μοντέλο. Στη συνέχεια, η θνησιμότητα προβλέπεται με παρέκταση της συνολικής χρονικής τάσης με τη χρήση τυπικών διαδικασιών χρονοσειρών (Lee & Carter, 1992). Μεταξύ των σαφών πλεονεκτημάτων της μεθοδολογίας Lee-Carter είναι ότι περιλαμβάνει ένα απλό στοχαστικό μοντέλο με μία μόνο χρονικά μεταβαλλόμενη παράμετρο- αποδίδει σχετικά καλά όταν οι παρελθούσες τάσεις ήταν γραμμικές, και είναι σε θέση να προβλέψει ένα μεταβαλλόμενο ηλικιακό πρότυπο θνησιμότητας (Booth & Tickle, 2008).

Πολλές παραλλαγές, επεκτάσεις και εναλλακτικές λύσεις του μοντέλου Lee-Carter έχουν αναπτυχθεί από την δημιουργία του. Για παράδειγμα, έχουν προταθεί διάφορες εναλλακτικές διαδικασίες εκτίμησης για τη βελτίωση της προσαρμογής του μοντέλου (Lee & Miller, 2001; Currie et al., 2004; Hatzopoulos & Haberman, 2009). Μια εξ' αυτών των εναλλακτικών προσεγγίσεων είναι η προσέγγιση των λειτουργικών δεδομένων που αναπτύχθηκε ως εναλλακτική λύση στο μοντέλο Lee-Carter. Η προσέγγιση αυτή περιλαμβάνει κύριες συνιστώσες υψηλότερης τάξης και μη παραμετρική εξομάλυνση (Hyndman & Ullah, 2007) και παρέχει ένα πιο ρεαλιστικό μελλοντικό ηλικιακό πρότυπο θνησιμότητας. Επίσης, έχουν προταθεί και άλλες διαδικασίες (Li et al., 2013) για να αντιμετωπιστεί η πολύ επικριθείσα υπόθεση σταθερού ηλικιακού προτύπου μεταβολής στο μοντέλο Lee-Carter (Booth & Tickle, 2008). Επιπλέον, για τη βελτίωση της απόδοσης πρόβλεψης του μοντέλου Lee-Carter, ιδίως όταν οι τάσεις θνησιμότητας στο παρελθόν ήταν μη γραμμικές, έχουν προταθεί και εφαρμοστεί η συμπερίληψη μιας παραμέτρου κοόρτης (Renshaw & Haberman, 2006; Reither et al., 2011) και διάφορες προσεγγίσεις για την ανίχνευση και την αντιμετώπιση της διαρθρωτικής αλλαγής (Booth et al. 2002; van Berkum et al. 2016).

Η παρούσα διπλωματική εργασία εκκινεί από την προαναφερθείσα προβληματική της μελέτης της θνησιμότητας και της ανάλυσης των υφιστάμενων μοντέλων πρόβλεψης και στην συνέχεια αναπτύσσει βάσει του μοντέλου Lee-Carter (1992) μια πρόβλεψη για την θνησιμότητα στον Ελληνικό πληθυσμό. Στο δεύτερο κεφάλαιο της διπλωματικής εργασία αναλύονται οι βασικές

έννοιες που χρησιμοποιούνται στα μοντέλα πρόβλεψης και δίνονται οι βασικοί ορισμοί και οι τρόποι μέτρησης της θνησιμότητας. Τίθεται επίσης υπό επεξεργασία η έννοια των πινάκων επιβίωσης και η σημασία τους στην πρόβλεψη, ενώ στην συνέχεια παρουσιάζονται και οι σύγχρονες τάσεις θνησιμότητας, το προσδόκιμο ζωής κ.ά. μεγέθη για τους πληθυσμούς της ΕΕ.

Το τρίτο κεφάλαιο εστιάζει στο βασικό μοντέλο της μεθόδου Lee-Carter, στην προσαρμογή του μοντέλου και στις βασικές παραμέτρους των προβλέψεων. Επικουρικά παρατίθεται και η μοντελοποίηση Poisson που αποτελεί μια διορθωτική μεθοδολογία στο βασικό μοντέλο, όπου εισάγεται στην πρόβλεψη η σημασία της πιθανοφάνειας, ενώ το κεφάλαιο κλείνει με ορισμένες κριτικές παρατηρήσεις. Το τέταρτο κεφάλαιο δίνει εναλλακτικά μοντέλα πρόβλεψης που χρησιμοποιούνται όπως το μοντέλο Heligman-Pollard, μοντέλα γενικής γραμμικής μοντελοποίησης, μοντέλα κοορτής κ.ά. Στο τέταρτο κεφάλαιο παρατίθεται η μέθοδος Lee-Carter, η οποία χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία, και ο τρόπος με τον οποίο έγινε η επεξεργασία των δεδομένων μέσω του στατιστικού περιβάλλοντος R version 4.2.2. Το πέμπτο κεφάλαιο παρουσιάζει τα αποτελέσματα και τις προβλέψεις σε τρία επίπεδα, δηλαδή στον γενικό πληθυσμό, στους άντρες και τις γυναίκες. Τέλος στο έκτο κεφάλαιο της εργασίας παρατίθενται τα συμπεράσματα της έρευνας.

## Κεφάλαιο 2<sup>ο</sup>: Θνησιμότητα σύγχρονες και μελλοντικές τάσεις

Η καλή ποιότητα και πληρότητα των δεδομένων σχετικά με τον θάνατο αλλά και τις αιτίες θανάτου ενέχουν υψηλή σημασία υπό το πρίσμα της χρήσης τους σε έρευνες και αναλύσεις που διεξάγονται από ιατρικά ιδρύματα, στατιστικές υπηρεσίες, κρατικούς φορείς, την ακαδημαϊκή κοινότητα και μεμονωμένους επαγγελματίες. Η καλή ποιότητα των δεδομένων σχετικά με την θνησιμότητα μπορεί να συνδράμει:

- Στην εκτίμηση του προτύπου πληθυσμιακής θνησιμότητας και στον προσδιορισμό των αλλαγών του με την πάροδο του χρόνου.
- Στον εντοπισμό περιφερειακών διαφορών στα ποσοστά θνησιμότητας και διερεύνηση των αιτιών αυτών των διαφορών.
- Στην παρακολούθηση των τάσεων σε θέματα δημόσιας υγείας όπως η βρεφική και μητρική θνησιμότητα, οι μολυσματικές ασθένειες, τα ατυχήματα και οι αυτοκτονίες.
- Στον εντοπισμό κινδύνων για την υγεία που συνδέονται με περιβαλλοντικούς και επαγγελματικούς παράγοντες και τον τρόπο ζωής.
- Στον καθορισμό προτεραιοτήτων έρευνας για την υγεία και την υγειονομική περίθαλψη και την ορθή κατανομή πόρων.
- Στον σχεδιασμό εγκαταστάσεων υγείας, υπηρεσιών και την κατάρτιση ανθρώπινου δυναμικού.
- Στον σχεδιασμό προγραμμάτων πρόληψης και προσυμπτωματικού ελέγχου και αξιολόγηση των αποτελεσμάτων αυτών των προγραμμάτων.
- Στην ανάπτυξη ασφαλιστικών προγραμμάτων προαγωγής της υγείας και αξιολόγηση των αποτελεσμάτων τους.

Με τις εξελίξεις σε ορισμένους τομείς, συμπεριλαμβανομένης της αυξανόμενης χρήσης διοικητικών δεδομένων και των πιο περίπλοκων συνδυασμών μικτών πηγών (έρευνα και διοικητικά δεδομένα) δεδομένων που χρησιμοποιούνται για την παραγωγή στατιστικών αποτελεσμάτων, δεν είναι πλέον εφικτός ένας ενιαίος κατάλογος βασικών μέτρων ποιότητας. Λόγω αυτού θα εξεταστούν στην συνέχεια οι βασικοί δείκτες θνησιμότητας, οι δείκτες συναρτώμενοι με την ηλικία αλλά και μια συγκριτική αξιολόγηση των δεικτών.

## 2.1 Θνησιμότητα: ορισμός και τρόποι μέτρησης

Οι δείκτες θνησιμότητας, και συγκεκριμένα τα ποσοστά θνησιμότητας, εκφράζονται και υπολογίζονται με διαφορετικούς τρόπους. Το απλούστερο μέτρο θνησιμότητας είναι ο αριθμός των θανάτων. Ωστόσο, το μέτρο αυτό δεν παρουσιάζει μεγάλη χρησιμότητα σε ερευνητικούς και πρακτικούς σκοπούς, καθώς επηρεάζεται σε μεγάλο βαθμό από τον αριθμό των ανθρώπων που κινδυνεύουν να πεθάνουν. Εξαιτίας αυτού, σε έρευνες και μετρήσεις συνήθως η θνησιμότητα μετριέται χρησιμοποιώντας *ποσοστά*. Το ποσοστό θνησιμότητας ορίζεται ως:

$$\text{ποσοστό θνησιμότητα} = \frac{\text{αριθμός θανάτων σε μια συγκεκριμένη χρονική περίοδο}}{\text{αριθμός ανθρώπων υπό κίνδυνο θανάτου την ίδια περίοδο}}$$

Έτσι, για να μετρηθεί η θνησιμότητα, απαιτούνται δεδομένα για τον αριθμό των θανάτων και για τον αριθμό των ατόμων που εκτίθενται στον κίνδυνο θανάτου. Τα δεδομένα για τον αριθμό των θανάτων λαμβάνονται συνήθως από τα εθνικά μητρώα θανάτων και τα δεδομένα για τον αριθμό των ατόμων που εκτίθενται στον κίνδυνο θανάτου λαμβάνονται συνήθως από τις απογραφές του πληθυσμού. Ο δείκτης του αριθμού των θανάτων χρησιμοποιείται ακόμα και σε χώρες όπου η εγγραφή θανάτων είναι ανεπαρκής ή η ποιότητα των δεδομένων της απογραφής παρουσιάζει ανεπάρκειες ή προβλήματα.

### 2.1.1 Ακαθάριστο ποσοστό θνησιμότητας

Το απλούστερο δυνατό ποσοστό θανάτων είναι πιθανώς ο συνολικός αριθμός θανάτων σε μια δεδομένη χρονική περίοδο διαιρεμένος με τον συνολικό πληθυσμό. Αυτό το μέτρο ονομάζεται ακαθάριστο ποσοστό θνησιμότητας ή αδρό ποσοστό θνησιμότητας. Η χρονική περίοδος που είθισται να χρησιμοποιείται για τον υπολογισμό του ακαθάριστου ποσοστού θνησιμότητας είναι συνήθως το ένα ημερολογιακό έτος:

$$\text{ακαθάριστο ποσοστό θνησιμότητας} = \frac{\text{συνολικός αριθμός θανάτων σε ένα έτος}}{\text{συνολικός πληθυσμός}}$$

Ένα ζήτημα που ανακύπτει άμεσα και σχετίζεται με τον δείκτη αφορά τον όρο του *συνολικού πληθυσμού*. Κατά τη διάρκεια του έτους, ο πληθυσμός συνήθως είναι μεταβλητός. Εξ' αυτού πρέπει να οριστεί ειδικώς το χρονικό σημείο κατά το οποίο θα γίνει η μέτρηση. Συμβατικά, το σημείο που επιλέγεται είναι στα μισά του έτους (30 Ιουνίου). Ο πληθυσμός στις 30 Ιουνίου ονομάζεται *πληθυσμός του μέσου έτους*. Χρησιμοποιώντας αυτόν τον ορισμό του πληθυσμού που εκτίθεται στον κίνδυνο θανάτου, η σχέση που προκύπτει είναι η ακόλουθη:

$$\text{ακαθάριστο ποσοστό θνησιμότητας} = \frac{\text{συνολικός αριθμός θανάτων σε ένα έτος}}{\text{πληθυσμός μέσου έτους}}$$

Δηλώνοντας το ακαθάριστο ποσοστό θανάτων του έτους  $t$  με το σύμβολο  $d_t$  τον συνολικό αριθμό θανάτων του έτους  $t$  ως  $\theta_t$ , και τον συνολικό πληθυσμό στις 30 Ιουνίου του έτους  $t$  ως  $P_t$ , μπορούμε να γράψουμε την παραπάνω σχέση:

$$d_t = \frac{\theta_t}{P_t}$$

Επίσης για λόγους απλοποίησης της σχέσης οι δείκτες  $t$ , μπορούν να παραληφθούν, διότι (εκτός και αναφέρεται κάτι διαφορετικό στον ορισμό των ερευνητικών παραμέτρων), η χρονική περίοδος μέτρησης του ακαθάριστου ποσοστού θνησιμότητας είναι το ένα πλήρες ημερολογιακό έτος:

$$d = \frac{\theta}{P}$$

Δεδομένου ότι ο θάνατος είναι ένα σχετικά σπάνιο γεγονός στους περισσότερους πληθυσμούς, το ακαθάριστο ποσοστό θνησιμότητας είναι συχνά μικρό. Για το λόγο αυτό, συχνά εκφράζεται ως  $o$  αριθμός των θανάτων ανά χιλιάδα του πληθυσμού:

$$d = \frac{\theta}{P} \times 1000$$

### 2.1.2 Ποσοστά θνησιμότητας ανάλογα με την ηλικία

Το ακαθάριστο ποσοστό θνησιμότητας δεν παρέχει πολλές πληροφορίες σχετικά με τη θνησιμότητα. Ειδικότερα, ο κίνδυνος θανάτου ποικίλλει πολύ ανάλογα με την ηλικία και ο τύπος του ακαθάριστου ποσοστού θνησιμότητας δεν δείχνει τίποτα σχετικά με αυτή τη διακύμανση. Εξαιτίας αυτού, οι θεωρείται συχνά χρήσιμο να χρησιμοποιούν ποσοστά θνησιμότητας ανάλογα με την ηλικία. Το ποσοστό θνησιμότητας ανάλογα με την ηλικία σε ηλικία  $x$  ετών ορίζεται ως:

*ποσοστό θνησιμότητας ανάλογα με την ηλικία σε αριθμό (ηλικία × έτη)*

$$= \frac{\text{αριθμός θανάτων ατόμων ηλικίας } x \text{ ετών}}{\text{πληθυσμός ηλικίας } x \text{ ετών}}$$

σε ένα δεδομένο ημερολογιακό έτος. Στον τύπο το  $σε$  (ηλικία × έτη), σημαίνει «ηλικία  $x$  τελευταία γενέθλια». Ο παρονομαστής, όπως και πριν, είναι ο πληθυσμός του μέσου έτους.

Δηλώνοντας το ποσοστό θνησιμότητας ανάλογα με την ηλικία στην (ηλικία  $x$  έτη)  $\tau$  με το σύμβολο  $m_x$ , τον αριθμό των θανάτων ατόμων (ηλικίας  $x$  έτη) με  $\theta_x$  και τον πληθυσμό μέσου έτους (ηλικίας  $x$  έτη) ως  $P_x$ , μπορούμε να γράψουμε το ποσοστό θνησιμότητας ανά ηλικία:

$$m_x = \frac{\theta_x}{P_x}$$

ή εναλλακτικά επί της χιλιάδας:

$$m_x = \frac{\theta_x}{P_x} \times 1000$$

Οι δείκτες  $x$  υποδηλώνουν έτη ηλικίας, όχι ημερολογιακά έτη.

Τα ποσοστά θνησιμότητας ανάλογα με την ηλικία μπορούν να υπολογιστούν για ένα συγκεκριμένο έτος ηλικίας ή για ηλικιακές ομάδες, όπως 5-9 ετών, 10-14 και ούτω καθεξής. Επειδή είναι επίσης γνωστό ότι η θνησιμότητα ποικίλλει ανάλογα με το φύλο, τα ποσοστά θνησιμότητας ανάλογα με την ηλικία υπολογίζονται συνήθως ξεχωριστά για τους άνδρες και τις γυναίκες. Όταν υπολογίζονται ποσοστά για συγκεκριμένες ηλικιακές ομάδες, χρησιμοποιείται ένας ειδικός συμβολισμός για να δηλώσει την ακριβή ηλικιακή ομάδα που εξετάζεται. Το σύμβολο  $n\theta_x$  δηλώνει τον αριθμό των θανάτων σε άτομα μεταξύ των ακριβών ηλικιών  $x$  και  $x + n$  έτη. Το σύμβολο  $nP_x$  χρησιμοποιείται για να υποδηλώσει τον πληθυσμό του μέσου έτους των ανθρώπων μεταξύ των ακριβών ηλικιών  $x$  και  $x + n$  ετών και το σύμβολο  $nM_x$  υποδηλώνει το ποσοστό θνησιμότητας ανάλογα με την ηλικία μεταξύ των ακριβών ηλικιών  $x$  και  $x + n$  ετών. Έτσι, για παράδειγμα, το ποσοστό θνησιμότητας ανάλογα με την ηλικία στις ηλικίες 5-9 ετών,  ${}_5m_5$ , υπολογίζεται χρησιμοποιώντας τον τύπο:

$$m_5^5 = \frac{\theta_5^5}{P_5^5}$$

Υπάρχει μία μόνο ηλικιακή ομάδα για την οποία χρησιμοποιείται διαφορετική μέθοδος υπολογισμού των ποσοστών θνησιμότητας ανάλογα με την ηλικία. Αυτή είναι η ηλικιακή ομάδα των ανθρώπων «κάτω του 1 έτους». Για αυτήν την ηλικιακή ομάδα, ως παρονομαστής λαμβάνεται ο αριθμός των ζώντων γεννήσεων στο εν λόγω ημερολογιακό έτος και όχι ο πληθυσμός του μέσου έτους ηλικίας κάτω του 1 έτους.



### 2.1.3 Σφάλματα στις περιοδικές μετρήσεις

Μέχρι στιγμής, παρατέθηκαν δείκτες θνησιμότητας στους οποίους ο παρονομαστής είναι ένας πληθυσμός στα μέσα του έτους και ο αριθμητής είναι ο αριθμός των θανάτων κατά τη διάρκεια ολόκληρου του σχετικού ημερολογιακού έτους. Σε αυτούς τους δείκτες υπάρχουν δύο σφάλματα μέτρησης τα οποία εμφανίζουν δυσαναλογίες υπολογισμού μεταξύ αριθμητή και παρονομαστή. Το πρόβλημα αυτό μπορεί να γίνει κατανοητό καλύτερα στην ειδική ηλικιακή ομάδα «κάτω του 1 έτους». Επί παραδείγματι έστω ένα βρέφος που πέθανε στις 31 Μαρτίου 1997 σε ηλικία εννέα μηνών. Το βρέφος αυτό γεννήθηκε στις 30 Ιουνίου 1996. Ο θάνατός του περιλαμβάνεται στον αριθμητή για το ποσοστό θνησιμότητας ανάλογα με την ηλικία στα 0 τελευταία γενέθλια το 1997, αλλά η γέννησή του περιλαμβάνεται στον παρονομαστή για το ποσοστό θνησιμότητας ανά ηλικία στα 0 τελευταία γενέθλια το 1996.

Για την επίλυση αυτού του σφάλματος μέτρησης απαιτούνται πρόσθετα δεδομένα, δηλαδή πρέπει να είναι γνωστή η ακριβής περίοδος έκθεσης σε κάθε ηλικία κατά τη διάρκεια του συγκεκριμένου έτους για κάθε άτομο που κινδυνεύει να πεθάνει. Έτσι, για παράδειγμα, ένα άτομο του οποίου τα γενέθλια ( $x + 1$ ) ήταν στις 7 Σεπτεμβρίου και πέθανε στις 9 Οκτωβρίου, θα θεωρηθεί ότι συνεισέφερε 250/365 της έκθεσης ενός έτους κατά τη διάρκεια αυτού του έτους στην ηλικία  $x$  από τα τελευταία γενέθλια (καθώς μεσολαβούν 250 ημέρες μεταξύ 1ης Ιανουαρίου και 7 Σεπτεμβρίου) και 32/365 της έκθεσης ενός έτους κατά τη διάρκεια αυτού του έτους στην ηλικία  $x + 1$  από τα τελευταία γενέθλια (καθώς υπάρχουν 32 ημέρες μεταξύ 7 Σεπτεμβρίου και 9 Οκτωβρίου). Η άθροιση αυτών των κλασμάτων ενός έτους σε ολόκληρο τον υπό διερεύνηση πληθυσμό για κάθε έτος ηλικίας και η χρήση του αποτελέσματος στον παρονομαστή, θα έδινε έναν ρυθμό στον οποίο ο αριθμητής και ο παρονομαστής αντιστοιχούν ακριβώς. Το πρόβλημα που ανακύπτει σε αυτή την μεθοδολογία, η οποία επιλύει τα σφάλματα μέτρησης είναι ότι τέτοιες λεπτομερείς πληροφορίες δεν είναι συνήθως διαθέσιμες εκτός από ειδικές (και δαπανηρές) έρευνες που έχουν σχεδιαστεί για να τις συμπεριλάβουν. Ως εκ τούτου, η δημογραφικοί υπολογισμοί βασίζονται σε πληθυσμούς στα μέσα του έτους, λειτουργούν δηλαδή υπό ένα προσεγγιστικό μοντέλο.

Υπάρχει, ωστόσο, μια άλλη προσέγγιση για τη μέτρηση των ποσοστών θνησιμότητας, η οποία δεν οδηγεί σε παραβιάσεις μέτρησης (τουλάχιστον σε «επίπεδο ατόμου»). Σε αυτήν τη δεύτερη προσέγγιση, αυτό που γίνεται είναι να υπολογιστεί ο αριθμός των ατόμων που έχουν τα  $x$  γενέθλιά τους κατά τη διάρκεια μιας δεδομένης περιόδου και στη συνέχεια να καταγραφεί εάν

έκλεισαν άλλον ένα χρόνο ή αν πέθαναν (όποιο συμβεί πρώτο). Διαιρώντας τον αριθμό που πεθαίνουν με τον αρχικό αριθμό που έχουν τα  $x$  γενέθλιά τους, δίνεται ένα συγκεκριμένο ποσοστό θνησιμότητας για την ηλικία  $x$ .

Αυτό το είδος του ποσοστού θνησιμότητας ανάλογα με την ηλικία ονομάζεται συχνά *ρυθμός* τύπου  $q$  και σημαίνεται ως  $q_x$ , για να διακρίνεται από το πρώτο είδος ρυθμού συγκεκριμένης ηλικίας,  $m_x$ , το οποίο ονομάζεται ρυθμός τύπου  $m$ . Στους δημογραφικούς υπολογισμούς χρησιμοποιούνται επίσης οι όροι αρχικές τιμές για τους ρυθμούς τύπου  $q$  και κεντρικές τιμές για τους ρυθμούς τύπου  $m$ . Αυτό συμβαίνει επειδή στα ποσοστά τύπου  $q$  ο εκτεθειμένος σε κίνδυνο ορίζεται στην αρχή ή στην έναρξη του έτους της ηλικίας που ερευνάται (δηλαδή, όταν οι εκτεθειμένοι σε κίνδυνο γιορτάζουν τα  $x$  γενέθλιά τους), ενώ στα ποσοστά τύπου  $m$  η έκθεση σε κίνδυνο είναι μια εκτίμηση του αριθμού των ατόμων ηλικίας  $x$  από τα τελευταία γενεθλίων μέχρι τη στιγμή που έλαβαν χώρα τα γεγονότα. Η μέση ηλικία αυτών των ατόμων είναι  $x + \frac{1}{2}$  έτη: βρίσκονται στα μισά (δηλ. στο «κέντρο») του εν λόγω έτους ηλικίας.

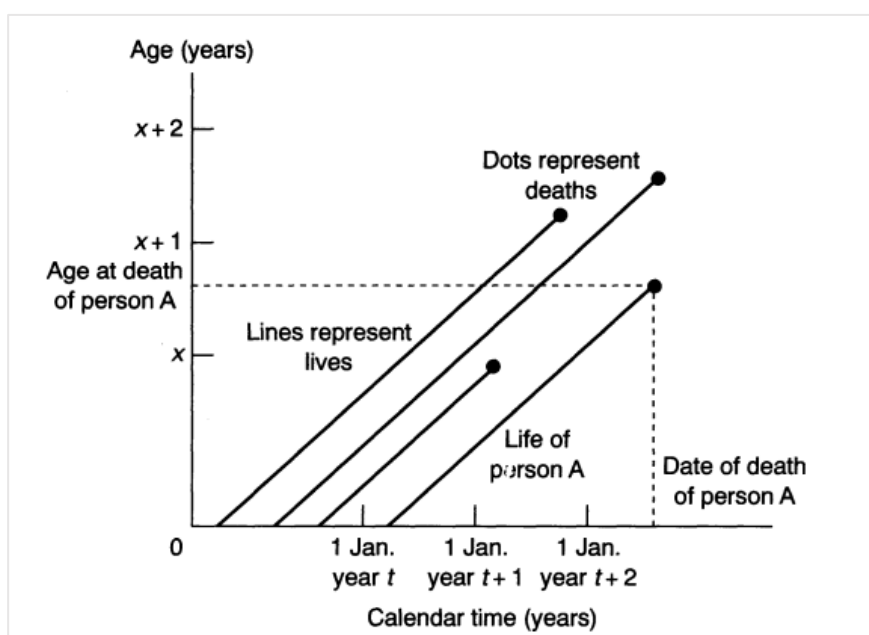
#### 2.1.4 Διαφορές μεταξύ των μετρήσεων

Οι δύο τύποι ποσοστού θνησιμότητας αποτελούν παραδείγματα δύο διαφορετικών προσεγγίσεων για τη μέτρηση των συνιστωσών της πληθυσμιακής αλλαγής. Η μια προσέγγιση υπολογίζει τα ποσοστά με βάση μια συγκεκριμένη ημερολογιακή χρονική περίοδο (ρυθμοί τύπου  $m$ ). Αυτό η μέθοδος είναι γνωστή ως περιοδική προσέγγιση. Η άλλη προσέγγιση υπολογίζει τα ποσοστά με βάση την εμπειρία μιας συγκεκριμένης ομάδας ανθρώπων που γεννήθηκαν κατά τη διάρκεια μιας συγκεκριμένης ημερολογιακής περιόδου (ποσοστά τύπου  $q$ ). Δεδομένου ότι τα ποσοστά τύπου  $q$  βασίζονται σε μια ομάδα ανθρώπων που γιορτάζουν τα  $x$  γενέθλιά τους κατά τη διάρκεια μιας δεδομένης περιόδου, συνεπάγεται ότι όλοι πρέπει να έχουν γεννηθεί κατά τη διάρκεια μιας περιόδου ίδιας διάρκειας  $x$  χρόνια νωρίτερα. Μια τέτοια ομάδα ανθρώπων είναι γνωστή ως *κοορτή γέννησης* και αυτή η προσέγγιση ονομάζεται *προσέγγιση κοορτής*.

Η διαφορά μεταξύ των προσεγγίσεων περιόδου και κοορτής, και ως εκ τούτου μεταξύ των ρυθμών τύπου  $m$  και τύπου  $q$ , μπορεί να απεικονιστεί χρησιμοποιώντας ένα διάγραμμα που ονομάζεται διάγραμμα *Lexis*. Το διάγραμμα *Lexis* έχει έναν κατακόρυφο άξονα που αντιπροσωπεύει την ηλικία και έναν οριζόντιο άξονα που αντιπροσωπεύει τον ημερολογιακό χρόνο. Δεδομένου ότι οι άνθρωποι μεγαλώνουν όσο περνά ο καιρός, η ζωή οποιουδήποτε

ατόμου μπορεί να αναπαρασταθεί σε ένα διάγραμμα Lexis με μια διαγώνια γραμμή που εκτείνεται από τον οριζόντιο άξονα μέχρι ένα σημείο που αντιστοιχεί στην ηλικία του ατόμου κατά το θάνατο, μετρούμενη στον κατακόρυφο άξονα (Διάγραμμα 1).

Σε ένα διάγραμμα Lexis, ο πληθυσμός που είναι ζωντανός και ηλικίας  $x$  από τα τελευταία γενέθλια ( $P_x$ ), σε μια συγκεκριμένη χρονική στιγμή αντιπροσωπεύεται από μια κάθετη γραμμή. Στο Διάγραμμα 2, το ευθύγραμμο τμήμα AB αντιπροσωπεύει όσους ζούσαν και ήταν ηλικίας 30 σύμφωνα με τα τελευταία γενέθλια στις 30 Ιουνίου 1997. Τα κάθετα ευθύγραμμο τμήματα όπως το AB αντιπροσωπεύουν έτσι τους παρονομαστές των ποσοστών θνησιμότητας τύπου  $m$ .

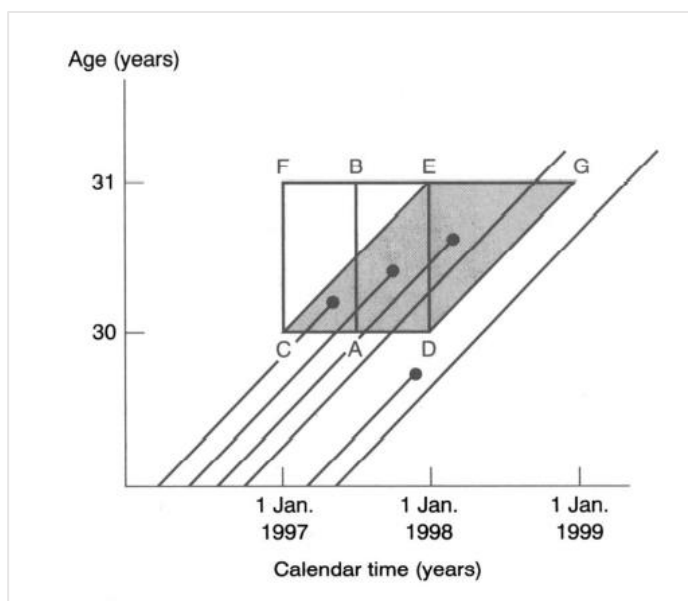


Διάγραμμα 1: Διάγραμμα Lexis

Σημείωση: Οι μεμονωμένες ζωές αντιπροσωπεύονται από διαγώνιες γραμμές που εκτείνονται από κάτω αριστερά προς τα πάνω δεξιά του γραφήματος. Καθώς τα άτομα μεγαλώνουν, οι γραμμές ανεβαίνουν ( $\theta < 90^\circ$ ). Οι συντεταγμένες του άνω άκρου κάθε γραμμής δηλώνουν την ώρα του θανάτου και την ηλικία του ατόμου που πέθανε.

Ένα σύνολο ανθρώπων που γιορτάζουν τα  $x$  γενέθλιά τους κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου (που σημαίνει ότι γεννήθηκαν όλοι κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου και επομένως αποτελούν μια κοορτή γέννησης) αντιπροσωπεύονται από μια οριζόντια γραμμή στο διάγραμμα Lexis. Στο Διάγραμμα 2, το ευθύγραμμο τμήμα CD αντιπροσωπεύει όλα τα άτομα που γιόρτασαν τα 30ά γενέθλιά τους κατά τη διάρκεια του έτους 1997 (και που, ως εκ τούτου, γεννήθηκαν όλοι το 1967). Οι

οριζόντιες γραμμές όπως το ευθύγραμμο τμήμα CD αντιπροσωπεύουν κοορτές γέννησης ή τους παρονομαστές των ποσοστών θνησιμότητας τύπου  $q$ .



Διάγραμμα 2: Παράδειγμα διαγράμματος Lexis

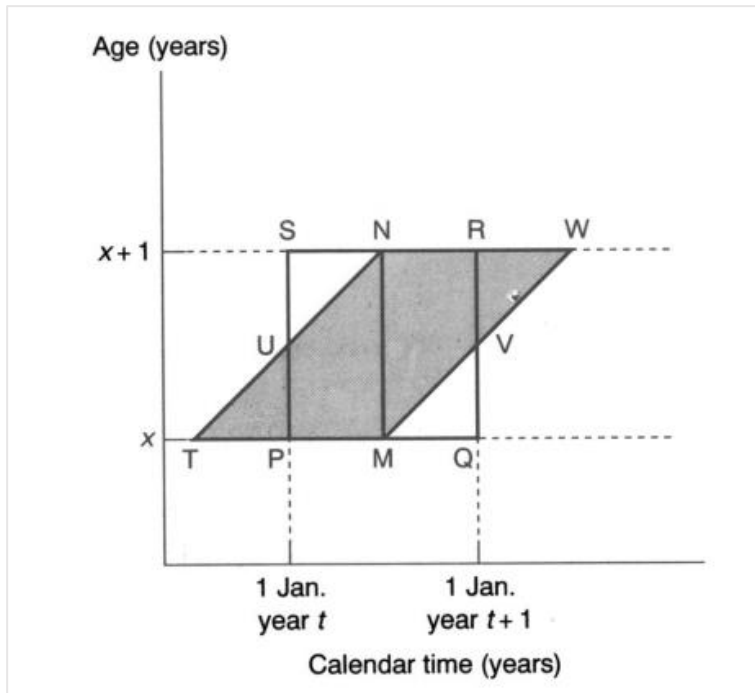
Οι θάνατοι ατόμων ηλικίας  $x$  με βάση τα τελευταία τους γενέθλια κατά την ημερομηνία θανάτου τους και που πέθαναν σε μια συγκεκριμένη χρονική περίοδο, αντιπροσωπεύονται με τετράγωνα. Το τετράγωνο CDEF στο Διάγραμμα 2 αντιπροσωπεύει όλους τους ανθρώπους που πέθαναν κατά τη διάρκεια του ημερολογιακού έτους 1997 και ήταν 30 ετών σύμφωνα με τελευταία γενέθλια όταν πέθαναν. Τετράγωνα σαν αυτό αντιπροσωπεύουν τους αριθμητές των ποσοστών θνησιμότητας τύπου  $m$ .

Οι θάνατοι των ατόμων ηλικίας  $x$  των τελευταίων γενεθλίων κατά την ημερομηνία θανάτου τους και που όλοι γιόρτασαν τα  $x$  γενέθλιά τους κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου, αντιπροσωπεύονται με παραλληλόγραμμα. Στο Διάγραμμα 2, το σκιασμένο παραλληλόγραμμο CDGE αντιπροσωπεύει όλους τους ανθρώπους που πέθαναν σε ηλικία 30 ετών σύμφωνα με τα τελευταία τους γενέθλια όταν πέθαναν, και που γιόρτασαν τα 30ά γενέθλιά τους κατά το ημερολογιακό έτος 1997. Παραλληλόγραμμα όπως αυτό αντιπροσωπεύουν τους αριθμητές των ποσοστών θνησιμότητας τύπου  $q$ .

#### 2.1.5 Η σχέση μεταξύ των δύο τύπων ποσοστού θνησιμότητας

Κατ' αρχήν, δεν υπάρχει κάποια αναγκαία συσχέτιση σχέση μεταξύ των δύο τύπων ποσοστού θνησιμότητας ( $m$  και  $q$ ). Ωστόσο, διατυπώνοντας μια σειρά από υποθέσεις, μπορεί να εξαχθεί μια θεωρητική σχέση.

Βάσει του Διαγράμματος 3 η εικαζόμενη σχέση μεταξύ των δύο τύπων ποσοστού θνησιμότητας μπορεί να καταστεί ορατή. Οι θάνατοι που αντιπροσωπεύουν τον αριθμητή του ποσοστού θνησιμότητας τύπου  $m$  στην ηλικία  $x$  βάσει των τελευταίων γενέθλιων στο ημερολογιακό έτος  $t$  είναι στο τετράγωνο PQRS. Έστω ότι υπάρχουν  $\theta_x$  από αυτούς τους θανάτους. Οι θάνατοι που αντιπροσωπεύουν τον αριθμητή του ποσοστού θνησιμότητας τύπου  $q$  στην ηλικία  $x$  που επικαλύπτεται περισσότερο με τον αριθμητή του ποσοστού τύπου  $m$  για το έτος  $t$  βρίσκονται στο σκιασμένο παραλληλόγραμμο TMWN.



Διάγραμμα 3: Διάγραμμα Lexis που απεικονίζει τη σχέση μεταξύ των δύο τύπων ποσοστού θνησιμότητας

Οι θεωρητικές υποθέσεις που βοηθούν στην σύνδεση των δύο τύπων ( $m$  και  $q$ ) είναι δύο:

1. Η θνησιμότητα ποικίλλει ανάλογα με την ηλικία και όχι με τον ημερολογιακό χρόνο.
2. Οι θάνατοι κατανέμονται ομοιόμορφα σε κάθε έτος ηλικίας.

Αυτές οι δύο υποθέσεις αποτελούν ακρογωνιαίους λίθους στη δημογραφική ανάλυση της θνησιμότητας. Η υπόθεση ότι η θνησιμότητα ποικίλλει μόνο με την ηλικία είναι αρκετά εύκολη στην κατανόηση, αλλά η δεύτερη υπόθεση χρήζει περαιτέρω αποσαφήνισης. Η υπόθεση ότι οι θάνατοι κατανέμονται ομοιόμορφα σε κάθε έτος ηλικίας υπονοεί ότι η μέση ηλικία θανάτου όσων πεθαίνουν μεταξύ των 75ων και 76ων γενεθλίων τους είναι 75 έτη και 6 μήνες. Υπονοεί

ότι ο αριθμός των θανάτων μεταξύ 75 ετών και 75 ετών και 6 μηνών είναι ο ίδιος με τον αριθμό θανάτων μεταξύ 75 ετών και 6 μηνών και 76 ετών ακριβώς.

Βάσει αυτών των δύο υποθέσεων, στο διάγραμμα Lexis, εντός οποιασδήποτε οριζόντιας ζώνης που αντιπροσωπεύει τις ηλικίες μεταξύ  $x$  και  $x + 1$ , οι θάνατοι κατανέμονται ομοιόμορφα. Αυτό σημαίνει ότι ο αριθμός των θανάτων είναι ανάλογος με το εμβαδόν οποιασδήποτε τμήματος του διαγράμματος. Δύο τμήματα του γραφήματος που βρίσκονται και τα δύο εντός της οριζόντιας ζώνης που αντιπροσωπεύει ακριβώς τις ηλικίες μεταξύ  $x$  και  $x + 1$  και τα οποία έχουν την ίδια περιοχή θα αντιπροσωπεύουν τον ίδιο αριθμό θανάτων.

Εφόσον το εμβαδόν του παραλληλογράμμου TMWN στο Διάγραμμα 3 είναι ίσο με το εμβαδόν του τετραγώνου PQRS, τότε ο αριθμός των θανάτων που αντιπροσωπεύεται από το σκιασμένο παραλληλόγραμμο TMWN θα είναι επίσης  $\theta_x$ .

Το ποσοστό θνησιμότητας τύπου  $m$  ( $m_x$ ) δίνεται τότε από:

$$m_x = \frac{\theta_x}{\text{πληθυσμός που αντιπροσωπεύεται από την κάθετη γραμμή MN}} \quad (1)$$

Ενώ το ποσοστό θνησιμότητας τύπου  $q$  ( $q_x$ ) δίνεται από:

$$q_x = \frac{\theta_x}{\text{πληθυσμός που αντιπροσωπεύεται από την οριζόντια γραμμή TM}}$$

Όμως βάσει των δύο υποθέσεων (1,2) που αναφέρθηκαν παραπάνω, ο αριθμός των θανάτων μέσα στο τρίγωνο TMN πρέπει να είναι  $\frac{1}{2}\theta_x$ , δεδομένου ότι όλες οι ζωές που διασχίζουν τη γραμμή TM διασχίζουν επίσης τη γραμμή MN, έχουμε:

$$q_x = \frac{\theta_x}{\text{πληθυσμός που αντιπροσωπεύεται από την οριζόντια γραμμή MN} + \frac{1}{2}\theta_x} \quad (2)$$

αλλά από την εξίσωση (1) έχουμε  $MN = \frac{\theta_x}{m_x}$ , οπότε αντικαθιστώντας αυτόν τον όρο στην εξίσωση (2) έχουμε:

$$q_x = \frac{\theta_x}{\theta_x/m_x + \frac{1}{2}\theta_x}$$

με τους όρους  $\theta_x$  να απλοποιούνται έχουμε:

$$q_x = \frac{1}{1/m_x + \frac{1}{2}} = \frac{m_x}{1 + \frac{1}{2}m_x}$$

ή όπως γράφεται συνήθως:

$$q_x = \frac{2m_x}{2 + m_x}$$

Αν και το αποτέλεσμα εδράζεται στις δύο υποθέσεις που σημειώθηκαν παραπάνω, σε πολλές μετρήσεις η προσέγγιση είναι ικανοποιητική.

Ωστόσο, υπάρχουν ορισμένες ηλικιακές ομάδες στις οποίες η υπόθεση ότι οι θάνατοι κατανέμονται ομοιόμορφα δεν ισχύει. Αυτό το γεγονός άπτεται ιδιαίτερα για τον πρώτο χρόνο της ζωής. Οι περισσότεροι θάνατοι σε βρέφη κατά τη διάρκεια του πρώτου έτους της ζωής λαμβάνουν χώρα κατά τις πρώτες εβδομάδες του ίδιου έτους. Πράγματι, σε πληθυσμούς χαμηλής θνησιμότητας, είναι σύνηθες περισσότεροι από τους μισούς θανάτους σε βρέφη κάτω του 1 έτους να συμβαίνουν κατά τον πρώτο μήνα της ζωής. Οι θάνατοι σε βρέφη κατά τις πρώτες τέσσερις εβδομάδες της ζωής είναι γνωστοί ως νεογνικοί θάνατοι. Οι νεογνικοί θάνατοι μπορούν να μετρηθούν χρησιμοποιώντας το *ποσοστό νεογνικών θανάτων*, που ορίζεται ως

$$\begin{aligned} & \text{νεογνικοί θάνατοι} \\ &= \frac{\text{αριθμός θανάτων σε ένα δεδομένο έτος σε βρέφη ηλικίας 28 ημερών ή κάτω}}{\text{αριθμός γεννήσεων το συγκεκριμένο έτος}} \end{aligned}$$

#### 2.1.6 Πλεονεκτήματα και μειονεκτήματα των δύο ποσοστών θνησιμότητας

Οι δύο τύποι ποσοστού θνησιμότητας έχουν τα πλεονεκτήματα και τα μειονεκτήματά τους (Πίνακας 1). Σε γενικές γραμμές, τα ποσοστά τύπου m έχουν το πλεονέκτημα ότι είναι εύκολο να υπολογιστούν χρησιμοποιώντας απλά δεδομένα. Το μειονέκτημά τους είναι ότι δεν αντικατοπτρίζουν την εμπειρία των «πραγματικών» ανθρώπων και, αν υπολογιστούν χρησιμοποιώντας πληθυσμούς στα μέσα του έτους, παραβιάζουν την αρχή της μέτρησης. Τα πλεονεκτήματα και τα μειονεκτήματα των ποσοστών τύπου q είναι κατά μια έννοια «κατοπτρικές εικόνες» των τιμών τύπου m.

Πίνακας 1: Πλεονεκτήματα και μειονεκτήματα των δύο τύπων ποσοστού θνησιμότητας

Τύπος	Πλεονεκτήματα	Μειονεκτήματα
Τύπος-m	Τα δεδομένα είναι άμεσα διαθέσιμα  Εύκολος υπολογισμός  Μπορεί να εφαρμοστεί σε μια συγκεκριμένη ημερολογιακή χρονική περίοδο	Παραβίαση της αρχής της μέτρησης εάν βασίζεται σε πληθυσμούς στα μέσα του έτους  Δεν αντικατοπτρίζουν την εμπειρία μιας πραγματικής ομάδας ανθρώπων
Τύπος-q	Αντικατοπτρίζουν την εμπειρία πραγματικών ανθρώπων  Δεν παραβιάζετε την αρχή της μέτρησης	Δεν εφαρμόζονται για ένα συγκεκριμένο ημερολογιακό έτος  Τα δεδομένα δεν είναι άμεσα διαθέσιμα  Δύσκολο να υπολογιστούν

## 2.2 Πίνακες επιβίωσης

Ένας πίνακας επιβίωσης περιγράφει τη διαδικασία επιβίωσης μιας υποθετικής κοόρτης που είναι κλειστή ως προς την μετανάστευση και υπόκειται σε γνωστές πιθανότητες θανάτου ανά ηλικία. Στον πρώτο πίνακα ζωής, που συντάχθηκε από τον John Graunt το 1662 (Pollard, 1973), ο αρχικός αριθμός της κοορτής στην ηλικία 0, δηλαδή  $l_0$ , λήφθηκε ως 100 μέσα από την παροχή παραδειγμάτων. Η τρέχουσα τιμή του  $l_0$  συνήθως λαμβάνεται ως 100.000 χωρίς ιδιαίτερο λόγο. Η σημαντικότητα της τιμής  $l_0$  έγκειται στο αν πρέπει να διερευνηθεί η αβεβαιότητα της διαδικασίας επιβίωσης.

Στον παραδοσιακό πίνακα επιβίωσης, εκτός από τον αυθαίρετο αρχικό αριθμό, η διαδικασία επιβίωσης από τη μια ηλικία στην άλλη είναι ντετερμινιστική. Για παράδειγμα, όταν η πιθανότητα θανάτου μεταξύ γέννησης και ηλικίας 1 είναι 0,5, σύμφωνα με τον παραδοσιακό πίνακα επιβίωσης, μεταξύ 100 γεννήσεων, οι 50 θα επιβιώσουν μέχρι την ηλικία του 1 έτους.



Στην πραγματικότητα, ωστόσο, η ρίψη ενός νομίσματος 100 φορές δεν εγγυάται την εμφάνιση της κορώνας ή τον γραμμάτων 50 φορές . Με άλλα λόγια, η διαδικασία επιβίωσης είναι απροσδιόριστη, όχι ντετερμινιστική. Θα πρέπει επίσης να διευκρινιστεί ότι η «υποθετική» κοορτή δεν σημαίνει ότι ακολουθεί μια εξωπραγματική διαδικασία επιβίωσης. Πιο συγκεκριμένα, η υποθετική κοορτή έχει σχεδιαστεί για να αντικατοπτρίζει την πραγματική διαδικασία επιβίωσης σε μια υποθετική κατάσταση, όπου δεν υπάρχουν οι επιπτώσεις της μετανάστευσης ή άλλων γεγονότων «εξόδου» εκτός από τον θάνατο.

Η αυθαίρετη αρχική τιμή και η ντετερμινιστική διαδικασία επιβίωσης είναι ανεπιθύμητα χαρακτηριστικά που σχετίζονται με τον παραδοσιακό πίνακα επιβίωσης. Οι Li και Tuljapurkar (2012, 2013) ανέπτυξαν μια μέθοδο που αφαιρεί τα δύο χαρακτηριστικά και μετατρέπει τον πίνακα ζωής από ντετερμινιστικό σε πιθανολογικό.

### 2.2.1 Η αρχική τιμή

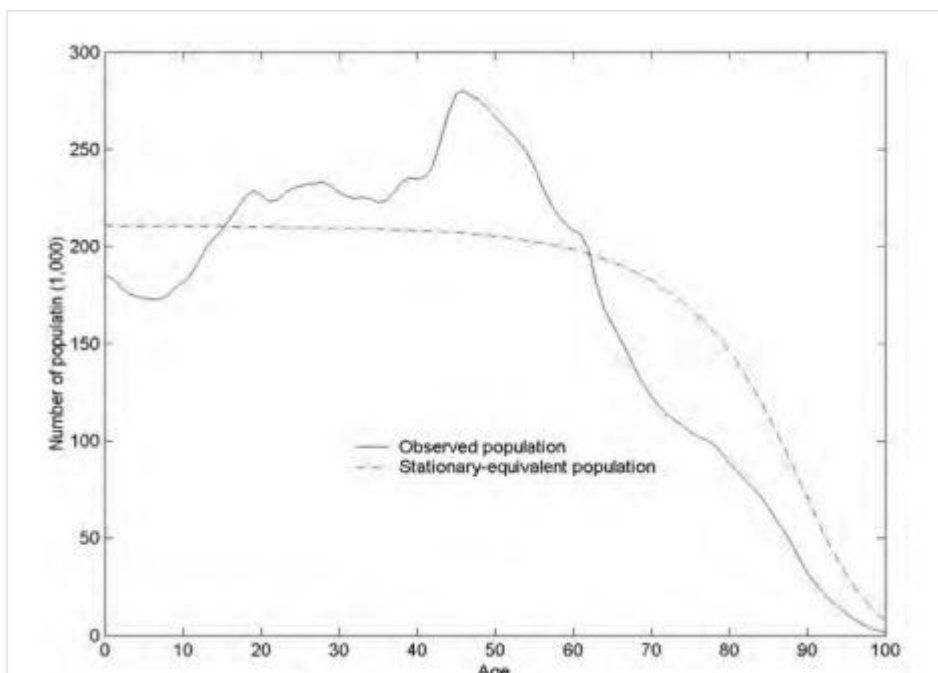
Η διακύμανση της μέσης ηλικίας θανάτου ανά 100 γεννήσεις είναι αισθητικά μεγαλύτερη από αυτή των 100.000 γεννήσεων (Li and Tuljapurkar, 2012). Εξ' αυτής της αισθητικής παρατήρησης θα πρέπει να βρεθεί η αρχική τιμή του  $l_0$  για έναν συγκεκριμένο πληθυσμό σε ένα δεδομένο έτος. Εάν ο πληθυσμός είναι στάσιμος, η τιμή του  $l_0$  είναι προφανώς ο αριθμός των ετήσιων γεννήσεων. Όταν ο παρατηρούμενος πληθυσμός δεν είναι στάσιμος, το  $l_0$  πρέπει να διαφέρει από τον αριθμό των ετήσιων γεννήσεων, ανάλογα με τη διαφορά μεταξύ των ηλικιακών δομών του παρατηρούμενου και του στάσιμου πληθυσμού. Για τους παρατηρούμενους πληθυσμούς που δεν είναι στάσιμοι, οι Li και Tuljapurkar (2012) προτείνουν να εκτιμηθεί ο στάσιμος-ισοδύναμος πληθυσμός, ο οποίος είναι μια υποθετική κοορτή με συγκεκριμένο αρχικό αριθμό που ελαχιστοποιεί τη διαφορά μεταξύ των ατόμων-ετών της κοορτής και του παρατηρούμενου πληθυσμού. Έστω οι άνθρωποι-έτη στο ηλικιακό διάστημα  $(\alpha, \alpha + n)$  να είναι  $p_0(\alpha, \alpha + n)$  και  ${}_nL_\alpha$  για τον παρατηρούμενο και τον στάσιμο πληθυσμό με αρχικό αριθμό 1, αντίστοιχα. Τότε, ο αρχικός αριθμός του στάσιμου-ισοδύναμου πληθυσμού είναι:

$$l_0 = \frac{\sum_{a=0}^{\omega} (\alpha, \alpha + n) {}_nL_a}{\sum_{a=0}^{\omega} {}_nL_a^2}$$

όπου το  $\omega$  υποδηλώνει το κάτω όριο της ανοιχτής ηλικιακής ομάδας. Ο υπολογισμός  $l_0$  δεν απαιτεί δεδομένα πρόσθετα σε σχέση με αυτά ενός παραδοσιακού πίνακα επιβίωσης.

Μεταξύ των υποθετικών κοορτών όλων των μεγεθών, ο στάσιμος-ισοδύναμος πληθυσμός είναι ο πλησιέστερος στον παρατηρούμενο πληθυσμό. Ωστόσο, αναμένονται διαφορές μεταξύ του στάσιμου-ισοδύναμου και των παρατηρούμενων πληθυσμών και αυτός είναι ο λόγος για την εισαγωγή της υποθετικής κοορτής για τη μέτρηση της θνησιμότητας.

Οι διαφορές μεταξύ των ηλικιακών προτύπων της υποθετικής κοορτής και του παρατηρούμενου πληθυσμού μπορούν να θεωρηθούν ως οι επιπτώσεις των ιστορικών δημογραφικών αλλαγών, οι οποίες περιλαμβάνουν τη γονιμότητα, τη θνησιμότητα και τη μετανάστευση και δεν έχουν καμία σχέση με την τρέχουσα θνησιμότητα. Στη συνέχεια, το σκεπτικό της εισαγωγής της υποθετικής κοορτής μπορεί να γίνει κατανοητό υπό την έννοια ότι εξαλείφει τις επιπτώσεις των ιστορικών δημογραφικών αλλαγών στο ηλικιακό πρότυπο του πληθυσμού που παρατηρήθηκε. Σύμφωνα με αυτή την άποψη, το ηλικιακό πρότυπο της υποθετικής κοορτής θα ήταν το ηλικιακό πρότυπο του παρατηρούμενου πληθυσμού, εάν δεν υπήρχαν επιδράσεις ιστορικών (δημογραφικών) αλλαγών που έκαναν το ηλικιακό πρότυπο του παρατηρούμενου πληθυσμού ακανόνιστο.



Διάγραμμα 4: Αριθμοί Καναδών γυναικών του παρατηρούμενου πληθυσμού και του στάσιμου πληθυσμού υποθετικής κοορτής

Σαφώς, ο καθορισμός του στάσιμου-ισοδύναμου πληθυσμού συνεπάγεται την εξάλειψη των επιπτώσεων των ιστορικών γεγονότων στα συγκεκριμένα άτομα-έτη ηλικίας του παρατηρούμενου πληθυσμού. Ως αποτέλεσμα, οι άνθρωποι-έτη του στάσιμου-ισοδύναμου

πληθυσμού θα ήταν τα αντίστοιχα άνθρωποι-έτη του παρατηρούμενου πληθυσμού, εάν δεν υπήρχαν προαναφερθείσες επιπτώσεις. Αντίστοιχα, η απροσδιοριστία της διαδικασίας επιβίωσης του στάσιμου-ισοδύναμου πληθυσμού θα ήταν η απροσδιοριστία της επιβίωσης του παρατηρούμενου πληθυσμού, αν πάλι δεν υπήρχαν προαναφερθέντα αποτελέσματα. Στη συνέχεια, χρησιμοποιώντας τον στάσιμο-ισοδύναμο πληθυσμό, οι απροσδιοριστίες της διαδικασίας επιβίωσης του παρατηρούμενου πληθυσμού μπορούν να χρησιμοποιηθούν για την παραγωγή γνώσεων σχετικά με τη διαδικασία επιβίωσης που δεν μπορούν να γίνουν κατανοητές προηγουμένως. Από την άλλη πλευρά, η διαδικασία επιβίωσης μιας κοορτής με αυθαίρετο αρχικό αριθμό είναι επίσης αβέβαιη, αλλά η αβεβαιότητά της δεν έχει νόημα.

### 2.2.2 Η αβεβαιότητα

Για να κατανοηθεί η αβεβαιότητα της διαδικασίας επιβίωσης, πρέπει να οριστεί η πιθανότητα θανάτου. Στις στατιστικές, η πιθανότητα θανάτου μεταξύ των ηλικιών  $a$  και  $a+n$  δηλαδή  $nq_a$  ορίζεται ως:

$$nq_a = \frac{l_a - l_{a+n}}{l_a} \quad (3)$$

όταν  $l_a$  είναι απείρως μεγάλο (Agresti & Finlay, 1997). Στη δημογραφία, ο ορισμός της πιθανότητας θανάτου είναι ο ίδιος με τον (3) αλλά χωρίς την προϋπόθεση ότι το είναι  $l_a$  απείρως μεγάλο (Chiang, 1984). Δεδομένου ενός πληθυσμού σε ένα συγκεκριμένο έτος, οι περισσότεροι δημογράφοι πιστεύουν ότι υπάρχουν οι πραγματικές τιμές της πιθανότητας θανάτου και μερικές φορές χρειάζεται να κάνουν προσαρμογές των παρατηρούμενων τιμών για να ταιριάζουν καλύτερα στις πραγματικές τιμές. Οι πραγματικές τιμές μπορούν να οριστούν από τον στατιστικό ορισμό. Από την άλλη πλευρά, ο δημογραφικός ορισμός οδηγεί σε δειγματοληπτικές τιμές της πιθανότητας θανάτου.

Με βάση τις στάσιμες-ισοδύναμες γεννήσεις και τον στατιστικό ορισμό της πιθανότητας θανάτου, οι Li και Tuljapurkar (2013) ανέπτυξαν μια στρατηγική που επεκτείνει τον πίνακα επιβίωσης από ντετερμινιστικό σε πιθανολογικό. Αυτή η νέα προσέγγιση προϋποθέτει δύο βασικές υποθέσεις. Η πρώτη υπόθεση είναι ότι οι παρατηρούμενες ειδικές για την ηλικία πιθανότητες θανάτου είναι αμερόληπτες, οι οποίες διαφέρουν από τις πραγματικές τιμές τυχαία με μηδενικό μέσο όρο. Κατά συνέπεια, οι παρατηρούμενες τιμές μπορούν να περιγραφούν ως

οι πραγματικές τιμές συν τυχαίες μεταβλητές με μηδενικό μέσο όρο. και αυτές οι τυχαίες μεταβλητές περιγράφουν την αβεβαιότητα. Ξεκινώντας από τις στάσιμες-ισοδύναμες γεννήσεις και χρησιμοποιώντας τις αμερόληπτες παρατηρούμενες πιθανότητες θανάτου ανά ηλικία, μπορούν να εξαχθούν οι κατανομές πιθανοτήτων των εκτιμώμενων πραγματικών τιμών όλων των μεταβλητών του πίνακα επιβίωσης σε κάθε ηλικία.

Η δεύτερη υπόθεση για την παραγωγή του πιθανολογικού πίνακα είναι ότι τα άτομα επιβιώνουν εξαρτημένα το ένα από το άλλο. Σύμφωνα με αυτήν την υπόθεση, η τιμή δείγματος των επιζώντων σε μια ορισμένη ηλικία μπορεί να υπολογιστεί με την κατανομή Bernoulli, σύμφωνα με την παρατηρούμενη πιθανότητα θανάτου και την τιμή δείγματος των επιζώντων στην προηγούμενη ηλικία, ξεκινώντας από τις στάσιμες-ισοδύναμες γεννήσεις. Επιπλέον, οι τιμές του δείγματος των επιζώντων σε όλες τις ηλικίες συνθέτουν μια κοορτή δείγματος και οι τιμές του δείγματος της πιθανότητας θανάτου ανά ηλικία μιας κοορτής δείγματος μπορούν να υπολογιστούν σύμφωνα με τον δημογραφικό ορισμό και την σχέση (3). Επιπλέον, οι τιμές του δείγματος όλων των άλλων μεταβλητών του πίνακα ζωής υπολογίζονται χρησιμοποιώντας διάφορους τύπους οι οποίοι συνθέτουν ένα δείγμα πίνακα επιβίωσης. Τέλος, πολλαπλοί πίνακες επιβίωσης δειγμάτων συνθέτουν έναν πιθανό πίνακα ζωής.

### 2.2.3 Ο πιθανολογικός πίνακας επιβίωσης

Σε έναν ντετερμινιστικό πίνακα επιβίωσης, όλες οι μεταβλητές εξαρτώνται μόνο από την ηλικία. Σε έναν πιθανολογικό πίνακα επιβίωσης, όλες οι μεταβλητές είναι τυχαίες και επομένως εξαρτώνται από την ηλικία και έναν δείκτη δείγματος, από τον οποίο η μηδενική τιμή «0» προορίζεται για παρατήρηση. Πιο συγκεκριμένα, το  $l_a$  περιγράφει τον αριθμό των επιζώντων στην ηλικία  $a$ , που είναι μια τυχαία μεταβλητή, και το  $l_a(0)$  αντιπροσωπεύει το παρατηρούμενο δείγμα του  $l_a$ . Ομοίως, όλα τα σύμβολα του πίνακα επιβίωσης χρησιμοποιούνται για την περιγραφή των αντίστοιχων τυχαίων μεταβλητών αλλά και για την παρουσίαση μιας τιμής δείγματος όταν επισυνάπτεται ένα ευρετήριο.

Το πρώτο βήμα για την παραγωγή ενός πιθανολογικού πίνακα επιβίωσης είναι ο υπολογισμός των δειγμάτων του  $l_a$  διαφορετικού από το παρατηρούμενο, δηλαδή  $l_a(i)$ , για  $i = 1, 2, 3, \dots, s_n$  όπου  $s_n$  αντιπροσωπεύει τον αριθμό δειγματοληπτικών πινάκων επιβίωσης. Το δεύτερο βήμα είναι ο υπολογισμός των αντίστοιχων δειγμάτων άλλων μεταβλητών του πίνακα επιβίωσης. Χρησιμοποιώντας ένα μεγάλο  $s_n$ , λαμβάνονται οι κατανομές δειγμάτων του  $l_a$  και όλων των άλλων μεταβλητών του πίνακα ζωής.

Έστω  ${}_nq_a(0)$  η αμερόληπτη παρατηρούμενη πιθανότητα θανάτου μεταξύ της ηλικίας  $a$  και  $a+n$ . τότε η διαδικασία επιβίωσης μπορεί να χαρακτηριστεί από την τυχαία συνάρτηση  $\Delta(nQa(0))$  που παίρνει την τιμή 1 για να αντιπροσωπεύει την επιβίωση ενός ατόμου από την ηλικία  $a$  έως την ηλικία  $a+n$  ή 0 για να περιγράψει τον θάνατο μεταξύ των ηλικιών  $a$  και  $a+n$ . Σύμφωνα με τις δύο παραδοχές που σημειώθηκαν παραπάνω, η διαδικασία επιβίωσης ενός ατόμου υπακούει στην κατανομή Bernoulli, με την παράμετρο  $nQa(0)$ . Οι τιμές του  $\Delta(nQa(0))$  μπορούν να δημιουργηθούν από σχεδόν οποιοδήποτε υπολογιστικό λογισμικό. Για παράδειγμα, εάν  $1Q0(0) = 0.5$ , τότε μεταξύ 100 τιμών του  $\Delta(1Q0(0))$ , το 1 θα εμφανιζόταν περίπου 50 φορές, υποδεικνύοντας ότι περίπου 50 θα επιβίωναν στην ηλικία του 1 μεταξύ 100 γεννήσεων. Ο ακριβής αριθμός του  $\Delta(1Q0(0))$  ισούται με 1, ωστόσο, είναι αβέβαιος.

Δηλώνοντας με  $s(a, j)$  την κατάσταση επιβίωσης του  $j$ ου ατόμου στην ηλικία  $a$ , και χρησιμοποιώντας  $s(a, j) = 1$  για να αναπαραστήσουμε ότι είναι ζωντανό στην ηλικία  $a$  και  $s(a, j) = 0$  ως νεκρό πριν από την ηλικία  $a$ , οι τιμές του  $s(a, j)$  μπορούν να υπολογιστούν ως:

$$s(a, j) = s(a - n, j)\Delta(nQ_{a-n}(0))$$

Για κάθε άτομο,  $s(a, j)$  του ξεκινά με 1 στην ηλικία μηδέν, και σε κάποια απρόβλεπτη ηλικία πέφτει στο 0, που σημαίνει θάνατο. Επαναλαμβάνοντας αυτή τη διαδικασία για καθένα από τα  $l_a$  άτομα, λαμβάνεται ένα δείγμα της διαδικασίας επιβίωσης της κοορτής όπως περιγράφεται παρακάτω.

Σε οποιαδήποτε ηλικία  $a$ , αθροίζοντας όλα τα ζωντανά άτομα προκύπτει ο αριθμός των επιζώντων της κοορτής, και επομένως η πρώτη τιμή δείγματος του  $l_a$  λαμβάνεται ως:

$$l_a(1) = \sum_{j=1}^{l_0} s(a, j) \quad (4)$$

Οι τιμές του δεύτερου δείγματος  $l_a(2)$  υπολογίζονται από ανεξάρτητη δημιουργία ενός άλλου συνόλου  $\Delta(nQa(0))$  για καθεμία από τις  $l_0$  γεννήσεις σε όλες τις ηλικίες, και στη συνέχεια χρησιμοποιώντας τις (3) και (4). Επαναλαμβάνοντας αυτή τη διαδικασία, η κατανομή του δείγματος του  $l_a$  λαμβάνεται ως  $l_a(i)$  για  $i = 1, 2, 3, \dots, s_n$ . Στην πραγματικότητα το  $l_a(i)$  θα μπορούσε επίσης να υπολογιστεί σε επίπεδο κοορτής και όχι σε ατομικό επίπεδο. Δεδομένης της τιμής του  $l_{a-n}(i)$ , το  $l_a^a(i)$  θα μπορούσε επίσης να υπολογιστεί ως τυχαίος αριθμός από μια διωνυμική κατανομή  $l_{a-n}(i)$  ιχνών και την πιθανότητα  $nQ_{a-n}(0)$ . Με αυτόν τον τρόπο, ο

υπολογιστικός χρόνος θα μπορούσε να μειωθεί σημαντικά, αλλά η δυνατότητα μελέτης ετερόγονων μοντέλων θα χαθεί (Vaupel et al. 1979).

Δεδομένου του  $i^{\text{ου}}$  δείγματος  $l_a(i)$ , ο πίνακας επιβίωσης του δείγματος  $i^{\text{ου}}$  μπορεί να υπολογιστεί χρησιμοποιώντας τους ίδιους τύπους. Η επανάληψη αυτής της διαδικασίας παράγει  $s_n$  δείγμα πινάκων επιβίωσης, που συνθέτουν έναν πιθανολογικό πίνακα επιβίωσης. Σε έναν πιθανολογικό πίνακα επιβίωσης, το περιεχόμενο μιας μεταβλητής είναι μια κατανομή πιθανότητας της εκτιμώμενης πραγματικής τιμής αυτής της μεταβλητής, σε μια ορισμένη ηλικία. Εν συγκρίσει, με έναν παραδοσιακό πίνακα επιβίωσης, το περιεχόμενο μιας μεταβλητής είναι μόνο ένα δείγμα της πραγματικής τιμής.

Ο αριθμός των στάσιμων-ισοδύναμων γεννήσεων  $l_0$  και ο αριθμός των κοορτών δείγματος  $s_n$  είναι απαραίτητα για τον πιθανολογικό πίνακα επιβίωσης. Όταν η τιμή του  $l_0$  είναι μεγαλύτερη, οι διακυμάνσεις των μεταβλητών του πίνακα επιβίωσης είναι μικρότερες. Από την άλλη πλευρά, όταν η τιμή του  $s_n$  είναι μεγαλύτερη, οι κατανομές δειγμάτων αυτών των μεταβλητών είναι πιο ομαλές και ακριβέστερες.

#### 2.2.4 Ο έλεγχος σημαντικότητας του μοντέλου

Ένας σημαντικός σκοπός της μέτρησης της θνησιμότητας είναι η ανίχνευση της διαφοράς μεταξύ των πληθυσμών. Για την εξυπηρέτηση αυτού του σκοπού, οι ντετερμινιστικοί πίνακες επιβίωσης θα μπορούσαν να προσδιορίσουν εάν μια μεταβλητή του πίνακα για έναν πληθυσμό είναι μεγαλύτερη από αυτή τη μεταβλητή για έναν άλλο πληθυσμό, ενώ μέσω του πιθανολογικού πίνακα επιβίωσης μπορεί να ελεγχθεί περαιτέρω εάν μια τέτοια διαφορά είναι στατιστικά σημαντική ή αν μπορεί να εμφανιστεί απλώς τυχαία.

Συμβολίζουμε μια μεταβλητή του πίνακα επιβίωσης για τους πληθυσμούς 1 και 2 με  $x_1$  και  $x_2$ , και τις αντίστοιχες τιμές δείγματος με  $x_1(i)$  και  $x_2(i)$  για  $i = 1, 2, 3, \dots, s_n$ . Για να ελεγχθεί η σημασία της διαφοράς μεταξύ των μέσων τιμών των  $x_1$  και  $x_2$ , η μηδενική υπόθεση μπορεί να οριστεί ως  $H_0: mean(x_1) = mean(x_2)$ . Σύμφωνα με αυτήν την υπόθεση, το στατιστικό test μπορεί να κατασκευαστεί ως:

$$z = \frac{x_1 - x_2 - mean(x_1 - x_2)}{\sqrt{\hat{\sigma}_1^2 - \hat{\sigma}_2^2}}$$

Όπου τα  $\hat{\sigma}_1^2$  και  $\hat{\sigma}_2^2$  είναι οι εκτιμήσεις τις διακύμανσης των  $x_1$  και  $x_2$  και μπορούν να υπολογιστούν από τις κατανομές δειγμάτων ως:

$$\sigma_h^2 = \frac{\sum_{i=1}^{s_n} \left[ x_h(i) - \frac{\sum_{i=1}^{s_n} x_h(k)}{s_n} \right]^2}{(s_n - 1)}, \quad h = 1, 2$$

Από την άλλη πλευρά, χωρίς τη μηδενική υπόθεση, το στατιστικό test της παρατηρούμενης τιμής  $z$  είναι:

$$z(0) = \frac{x_1(0) - x_2(0)}{\sqrt{\hat{\sigma}_1^2 - \hat{\sigma}_2^2}}$$

Στις συνήθεις εφαρμογές του ελέγχου σημαντικότητας, η αναλυτική κατανομή του  $\mathbf{z}$  είναι γνωστή ως τυπική ή κανονική κατανομή, και επομένως το 95 τοις εκατό διάστημα εμπιστοσύνης του είναι  $[-1,96, 1,96]$ . Εάν το  $z(0)$  πέσει εκτός  $[-1,96, 1,96]$ , τότε η  $H_0$  οδηγεί σε μια παρατήρηση που συμβαίνει με πιθανότητα μικρότερη από 0,05, και ως εκ τούτου απορρίπτεται στο επίπεδο στατιστικής σημαντικότητας 0,05. Στη συνέχεια, το συμπέρασμα είναι ότι η διαφορά μεταξύ των μέσων τιμών του  $x_1$  και  $x_2$  είναι στατιστικά σημαντική. Διαφορετικά, η διαφορά μεταξύ των μέσων των  $x_1$  και  $x_2$  δεν μπορεί να χαρακτηριστεί στατιστικά σημαντική.

Στην υπό εξέταση περίπτωση δεν γνωρίζουμε την αναλυτική κατανομή, αλλά έχουμε την κατανομή δείγματος, του  $z$ . Εφόσον το  $i$  επιλέγεται ανεξάρτητα για το  $x_1(i)$  και το  $x_2(i)$ , η κατανομή δείγματος του  $\mathbf{z}$  μπορεί να υπολογιστεί ως:

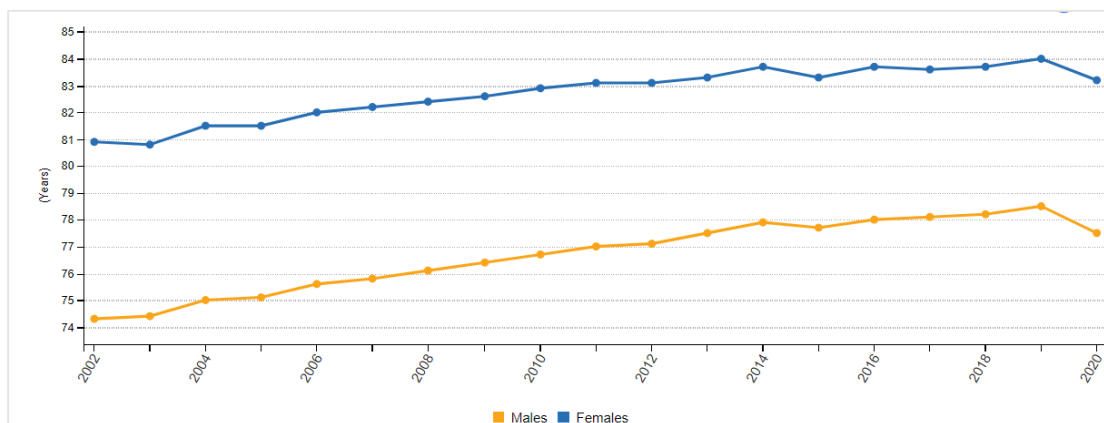
$$z(i) = \frac{x_1(i) - x_2(i) - \frac{\sum_{i=1}^{s_n} (x_1(i) - x_2(i))}{s_n}}{\sqrt{\hat{\sigma}_1^2 - \hat{\sigma}_2^2}}, \quad i = 1 \sim s_n$$

Χρησιμοποιώντας την κατανομή του δείγματος του  $z$ , μπορούμε να βρούμε το 95 τοις εκατό διάστημα εμπιστοσύνης, δηλαδή  $[C_1, C_2]$ , όχι επακριβώς, αλλά κατά προσέγγιση. Στη συνέχεια, ο έλεγχος στατιστικής σημαντικότητας μπορεί να πραγματοποιηθεί ακολουθώντας την ίδια λογική της χρήσης της κανονικής κατανομής.

### 2.3 Σύγχρονες τάσεις θνησιμότητας και επιπτώσεις

### 2.3.1 Προσδόκιμο ζωής κατά την γέννηση

Το προσδόκιμο ζωής κατά τη γέννηση αυξήθηκε ραγδαία τον περασμένο αιώνα λόγω μιας πληθώρας παραγόντων. Οι εν λόγω παράγοντες περιλαμβάνουν τη μείωση της βρεφικής θνησιμότητας, την άνοδο του βιοτικού επιπέδου, τη βελτίωση του τρόπου ζωής και την καλύτερη εκπαίδευση, καθώς και την πρόοδο στην υγειονομική περίθαλψη και την ιατρική. Οι επίσημες στατιστικές καταδεικνύουν ότι το προσδόκιμο ζωής αυξάνεται, κατά μέσο όρο, περισσότερο από δύο χρόνια ανά δεκαετία από τη δεκαετία του 1960. Ωστόσο, τα τελευταία στοιχεία δείχνουν ότι, το 2020, ο δείκτης αυτός μειώθηκε σε 23 από τα 27 κράτη μέλη της ΕΕ, με εξαίρεση τη Δανία, την Εσθονία, τη Φινλανδία και την Κύπρο. Με βάση τα διαθέσιμα στοιχεία θνησιμότητας για το 2020, η Eurostat δημιούργησε έναν ενημερωμένο χάρτη των επιπτώσεων της πανδημίας COVID-19 στο προσδόκιμο ζωής.



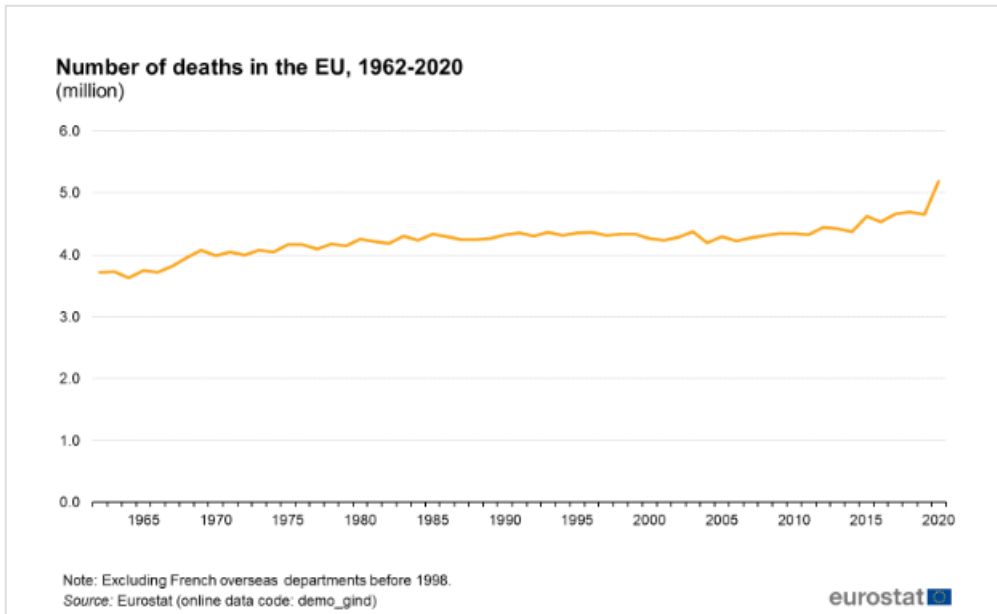
Διάγραμμα 5: Προσδόκιμο ζωής κατά τη γέννηση στην ΕΕ, 2002-2020

Πηγή: (Eurostat, 2022)

### 2.3.2 Αριθμός θανάτων

Το 2020, 5.184.077 άνθρωποι πέθαναν στην ΕΕ, σε σύγκριση με 4.653.033 το 2019, γεγονός που αποτελεί μια απότομη αύξηση άνω των 530.000 θανάτων σε ένα χρόνο (Διάγραμμα 6). Το στρογγυλοποιημένο ποσοστό θανάτων, που είναι ο αριθμός των θανάτων ανά 1.000 άτομα, υπολογίστηκε σε 11,6 στην ΕΕ το 2020 (1,2 περισσότεροι από το 2019).





Διάγραμμα 6: Αριθμός θανάτων στην ΕΕ, 1962-2020

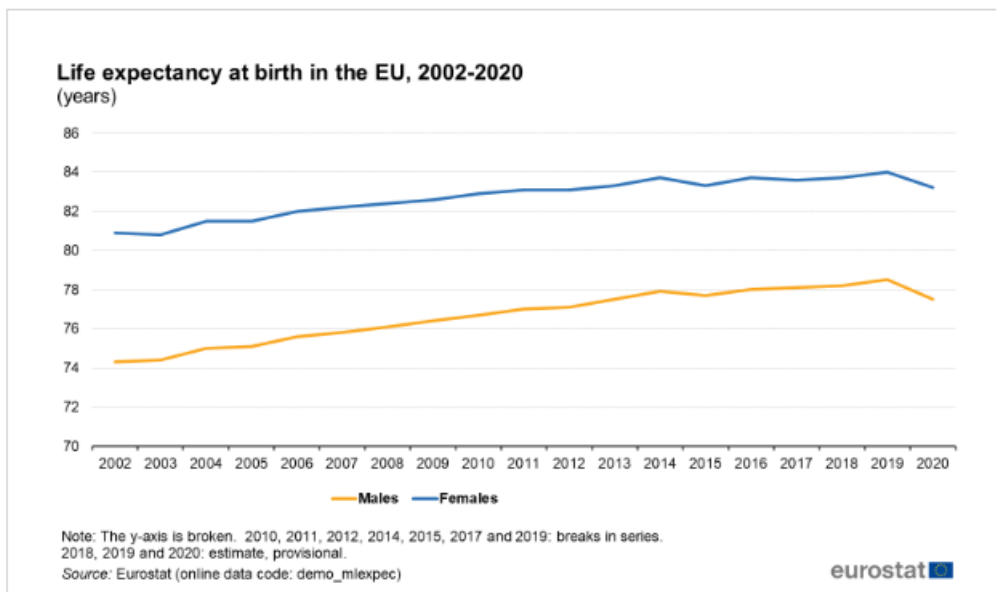
Πηγή: (Eurostat, 2022)

### 2.3.3 Προσδόκιμο ζωής κατά την γέννα

Ο πιο συχνά χρησιμοποιούμενος δείκτης για την ανάλυση της θνησιμότητας είναι το προσδόκιμο ζωής κατά τη γέννηση, δηλαδή ο μέσος αριθμός ετών που ένα άτομο αναμένεται να ζήσει, εάν υποβληθεί σε τρέχουσες συνθήκες θνησιμότητας καθ' όλη τη διάρκεια της υπόλοιπης ζωής του. Είναι ένας απλός αλλά ισχυρός τρόπος απεικόνισης των εξελίξεων στη θνησιμότητα.

Το προσδόκιμο ζωής κατά τη γέννηση στην ΕΕ υπολογίστηκε σε 80,4 έτη το 2020 (0,9 έτη χαμηλότερο από το 2019), φθάνοντας τα 83,2 έτη για τις γυναίκες (0,8 χαμηλότερα από το 2019) και τα 77,5 έτη για τους άνδρες (1 έτος χαμηλότερα από το 2019).

Συνολικά, μεταξύ του 2002 (το πρώτο έτος για το οποίο έγιναν διαθέσιμα στοιχεία για το προσδόκιμο ζωής για όλα τα κράτη μέλη της ΕΕ) και το 2020, το προσδόκιμο ζωής κατά τη γέννηση στην ΕΕ αυξήθηκε κατά 2,8 χρόνια, από 77,6 σε 80,4 έτη. Η αύξηση ήταν 2,4 έτη για τις γυναίκες και 3,2 έτη για τους άνδρες.



Διάγραμμα 7: Προσδόκιμο ζωής κατά τη γέννηση, ΕΕ, 2002-2020

Πηγή: (Eurostat, 2022)

Οι εκτιμήσεις από τη Eurostat με βάση τα στοιχεία του 2020 υποδηλώνουν ότι στα περισσότερα κράτη μέλη σημειώθηκε σημαντική μείωση του προσδόκιμου ζωής το 2020 (Πίνακας 2 και Πίνακας 3). Οι μεγαλύτερες μειώσεις του συνολικού προσδόκιμου ζωής κατά τη γέννηση (άνδρες και γυναίκες) καταγράφηκαν στην Ισπανία (-1,6 σε σύγκριση με το 2019) ακολουθούμενη από τη Βουλγαρία και την Πολωνία (και οι δύο -1,5) και στη συνέχεια η Λιθουανία και η Ρουμανία (και οι δύο -1,4).

Ο Πίνακας 2 δείχνει ότι το 2019 το προσδόκιμο ζωής αυξήθηκε σε 25 κράτη μέλη της ΕΕ σε σύγκριση με το προηγούμενο έτος, με εξαίρεση την Ελλάδα και την Κύπρο. Ωστόσο, όταν συγκρίνουμε το προσδόκιμο ζωής το 2020 με το 2019, μειώθηκε κατά 1 έτος ή περισσότερο σε 9 κράτη μέλη (Βέλγιο, Βουλγαρία, Τσεχία, Ισπανία, Ιταλία, Λιθουανία, Πολωνία, Ρουμανία και Σλοβενία), μειώθηκε κατά λιγότερο από 1 έτος σε 16 κράτη μέλη και αυξήθηκε κατά 0,1 στη Δανία και την Κύπρο.

Όπως φαίνεται στον Πίνακα 2, το 2020, το προσδόκιμο ζωής μειώθηκε επίσης σε όλες τις χώρες της ΕΦΤΑ, εκτός από τη Νορβηγία.

Πίνακας 2: Προσδόκιμο ζωής κατά τη γέννηση 1980-2020

Πηγή: (Eurostat, 2022)

Life expectancy at birth, 1980-2020  
(years)

	Total										Males					Females					
	1980	1990	2000	2010	2018	2019	2020	1980	1990	2000	2010	2018	2019	2020	1980	1990	2000	2010	2018	2019	2020
EU (*)	73.3	76.2	77.9	80.3	81.7	82.1	80.8	69.9	72.7	74.6	77.5	79.4	79.8	78.6	76.7	79.5	81.0	83.0	83.9	84.3	83.0
Belgium	71.1	71.2	71.6	73.8	75.0	75.1	73.6	68.4	68.0	68.4	70.3	71.5	71.6	70.0	73.9	74.7	75.0	77.4	78.6	78.8	77.5
Bulgaria	70.4	71.5	75.1	77.7	79.1	79.3	78.3	66.9	67.6	71.6	74.5	76.2	76.4	75.3	74.0	75.5	78.5	80.9	82.0	82.2	81.3
Czechia	74.2	74.9	76.9	79.3	81.0	81.5	81.6	71.2	72.0	74.5	77.2	79.1	79.5	79.7	77.3	77.8	79.2	81.4	82.9	83.5	83.6
Denmark	73.1	75.4	78.3	80.5	81.0	81.3	81.1	69.6	72.0	75.1	78.0	78.6	79.0	78.7	76.2	78.5	81.2	83.0	83.3	83.7	83.5
Germany	69.5	69.9	71.1	76.0	78.5	79.0	78.9	64.2	64.7	65.6	70.9	74.0	74.5	74.4	74.3	74.9	76.4	80.8	82.7	83.0	83.0
Estonia	74.8	76.6	80.8	82.2	82.8	82.6	82.6	72.1	74.0	78.5	80.4	80.8	80.8	80.8	77.7	79.2	83.1	84.1	84.7	84.4	84.4
Ireland (†)	75.3	77.1	78.6	80.6	81.9	81.7	81.4	73.0	74.7	75.9	78.0	79.3	79.2	78.8	77.5	79.5	81.3	83.3	84.4	84.2	83.9
Greece	75.5	76.9	79.3	82.4	83.5	84.0	82.4	72.3	73.3	75.8	79.2	80.7	81.1	79.6	78.5	80.6	82.8	85.5	86.3	86.7	85.2
Spain	79.2	81.8	82.8	83.0	82.3	82.3	82.3	75.3	78.2	79.7	79.9	79.2	79.2	79.2	83.0	85.3	85.8	85.9	85.9	85.3	85.3
France (†)	76.7	78.2	78.6	77.8	77.8	77.8	77.8	73.4	74.9	75.5	74.7	74.7	74.7	74.7	80.3	82.8	84.7	85.6	85.7	84.5	84.5
Croatia	77.1	79.9	82.2	83.4	83.6	82.3	82.3	73.8	76.9	79.5	81.2	81.4	80.0	80.0	80.1	83.9	84.8	84.4	84.4	84.4	84.4
Italy (†)	73.1	75.1	75.7	75.7	75.5	75.5	75.5	67.9	70.1	70.9	70.6	70.6	70.6	70.6	78.0	79.7	80.1	80.0	80.0	80.0	80.0
Cyprus	70.5	71.5	72.1	73.3	76.0	76.5	75.1	65.4	66.4	66.7	67.6	70.9	71.6	70.1	75.4	76.3	77.4	78.9	80.7	81.2	80.1
Latvia	72.8	75.7	78.0	80.8	82.3	82.7	82.2	70.0	72.4	74.6	77.9	80.1	80.2	79.9	75.6	78.7	81.3	83.5	84.6	85.2	84.5
Lithuania	69.1	69.4	71.9	74.7	76.2	76.5	75.7	65.5	65.2	67.5	70.7	72.7	73.1	72.3	72.8	73.8	76.2	78.6	79.6	79.7	79.0
Luxembourg	70.4	78.5	81.5	82.5	82.9	82.3	82.3	68.0	76.3	79.3	80.4	81.2	80.3	80.3	80.5	83.6	84.6	84.6	84.6	84.5	84.5
Hungary	77.1	78.2	81.0	81.9	82.2	81.4	81.4	73.8	75.6	78.9	80.3	80.6	79.7	79.7	80.2	80.7	83.0	83.4	83.7	83.1	83.1
Malta	72.7	75.8	78.3	80.7	81.8	82.0	81.3	69.0	72.3	75.2	77.8	79.4	79.7	78.9	76.1	79.0	81.2	83.5	84.1	84.2	83.6
Netherlands	70.7	73.8	76.4	77.7	78.0	76.5	76.5	66.3	69.6	72.2	73.7	74.1	72.5	72.5	75.3	78.0	80.7	81.7	81.9	80.7	80.7
Austria	71.5	74.1	76.8	80.1	81.5	81.9	81.1	67.9	70.6	73.3	76.8	78.3	78.7	78.0	74.9	77.5	80.4	83.2	84.5	84.8	84.1
Poland (‡)	69.2	69.9	71.2	73.7	75.3	75.6	74.2	66.6	66.7	67.7	70.0	71.7	71.9	70.4	71.9	73.1	74.8	77.7	79.2	79.5	78.3
Portugal	73.9	76.2	79.8	81.5	81.6	80.6	80.6	69.8	72.2	76.4	78.5	78.7	77.8	77.8	77.8	79.9	83.1	84.4	84.5	84.5	83.4
Romania (‡)	70.4	71.1	73.3	75.6	77.4	77.8	77.0	66.7	66.7	69.2	71.8	73.9	74.3	73.5	74.4	75.7	77.5	79.3	80.8	81.2	80.4
Slovenia	73.7	75.1	77.8	80.2	81.8	82.1	82.0	69.2	71.0	74.2	76.9	79.1	79.3	79.2	78.0	79.0	81.2	83.5	84.5	84.8	84.8
Slovakia	75.8	77.7	79.8	81.6	82.6	83.2	82.4	72.8	74.8	77.4	79.6	80.9	81.5	80.6	79.0	80.5	82.0	83.6	84.3	84.8	84.2
Finland	76.8	78.1	79.7	81.9	82.9	83.2	83.1	73.5	75.5	77.8	79.8	81.3	81.7	81.6	80.4	80.7	81.6	84.1	84.5	84.7	84.6
Sweden	77.0	81.8	83.1	84.3	81.8	81.8	81.8	73.9	79.5	80.7	82.6	80.0	80.0	80.0	79.9	84.3	85.5	85.8	85.8	83.4	83.4
Iceland	75.8	76.6	78.8	81.2	82.8	83.0	83.3	72.4	73.4	76.0	79.0	81.1	81.3	81.6	79.3	79.9	81.5	83.3	84.5	84.7	84.9
Liechtenstein	75.7	77.5	80.0	82.7	83.8	84.0	83.1	72.3	74.0	77.0	80.3	81.9	82.1	81.0	79.0	80.9	82.8	84.9	85.7	85.8	85.1
Norway	76.1	76.9	76.7	75.9	75.9	75.9	75.9	73.6	74.5	74.0	73.2	73.2	73.2	73.2	75.2	77.2	78.6	78.6	76.7	76.7	76.7
Montenegro	73.0	75.0	76.7	76.6	74.4	74.4	74.4	70.8	72.9	74.6	74.7	72.2	72.2	72.2	80.5	80.5	80.5	80.5	80.5	80.5	80.5
North Macedonia	71.6	74.4	75.9	76.0	74.5	74.5	74.5	68.9	71.8	73.5	73.4	71.6	71.6	71.6	74.4	77.0	78.4	78.6	77.5	77.5	77.5
Albania	76.8	78.9	79.1	79.1	79.1	79.1	79.1	74.2	76.2	76.4	76.4	76.4	76.4	76.4	79.4	81.6	81.8	81.8	81.8	81.8	81.8
Serbia	74.2	76.2	76.4	76.4	76.4	76.4	76.4	74.2	76.2	76.4	76.4	76.4	76.4	76.4	79.4	81.6	81.8	81.8	81.8	81.8	81.8
Turkey	74.2	76.2	76.4	76.4	76.4	76.4	76.4	74.2	76.2	76.4	76.4	76.4	76.4	76.4	79.4	81.6	81.8	81.8	81.8	81.8	81.8

(†) 2010, 2015, 2017 and 2019: breaks in series. 2018, 2019 and 2020: provisional.

(‡) 2018, 2019: estimated

(§) 2018, 2019, 2020: provisional

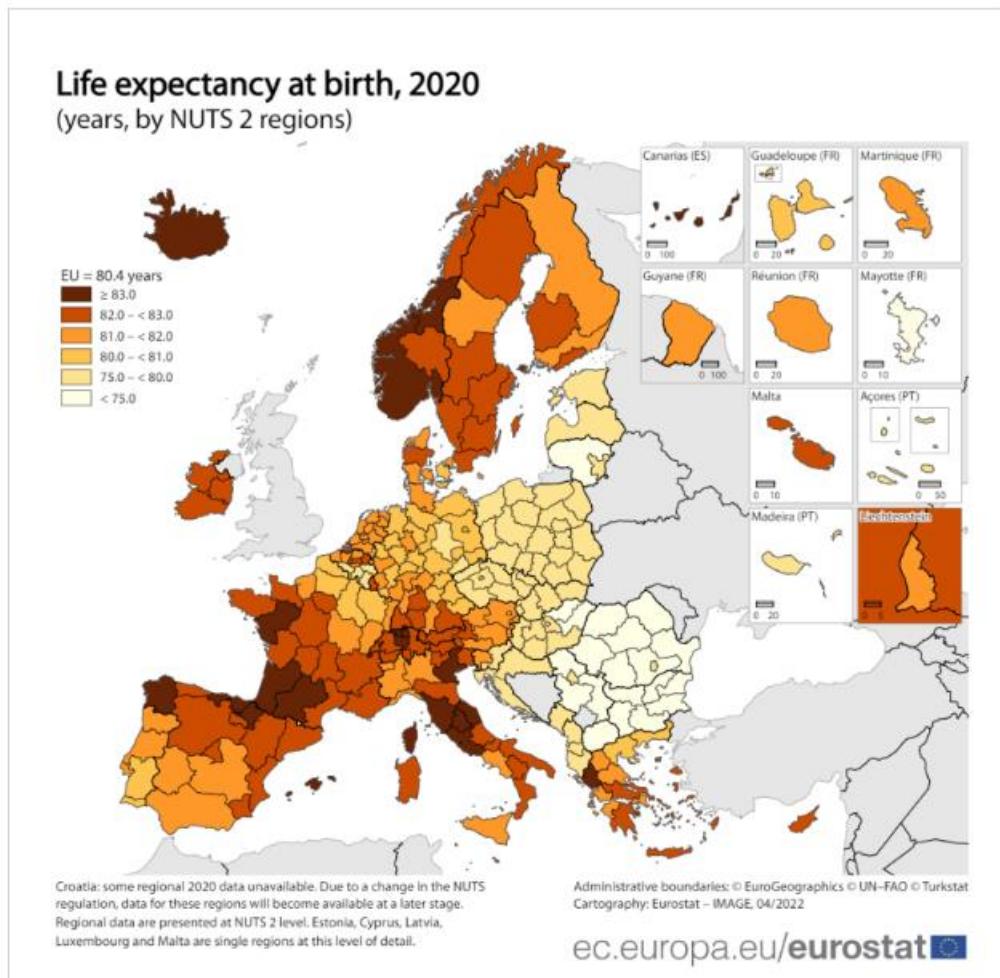
(¶) 2019: break in series

(\*) 2000 and 2010: break in series. 2020: provisional

(†) 2020: estimated

Source: Eurostat (online data code: demo\_mlexpec)

Σε όλες τις περιοχές της ΕΕ, το 2020, το υψηλότερο προσδόκιμο ζωής κατά τη γέννηση ήταν στη γαλλική νήσο Κορσική (84,0), ακολουθούμενες από τις Βαlearίδες Νήσους στην Ισπανία (83,9), την περιοχή της Ηπείρου στην Ελλάδα (83,8), τη Γαλικία στην Ισπανία και την Ούμπρια στην Ιταλία (83,7 και οι δύο) και στις Μιντί Πιρενέ στη Γαλλία (83,6). Οι περιφέρειες της ΕΕ με το χαμηλότερο προσδόκιμο ζωής κατά τη γέννηση ήταν στη Βουλγαρία: Βορειοδυτική (72,1), Βορειοκεντρική (72,8) και Βόρεια και Νοτιοανατολική Βουλγαρία (72,9).



Σχήμα 1: Προσδόκιο ζωής κατά τη γέννηση ανά περιοχή NUTS 2, 2020

Πηγή: Eurostat (Eurostat, 2022)

Όπως φαίνεται στον Πίνακα 3, το προσδόκιο ζωής στην ηλικία των 65 ετών παρουσίασε επίσης απότομη μείωση το 2020. Το 2020, το προσδόκιο ζωής στα 65 υπολογίστηκε σε 19,3 χρόνια (0,9 έτη χαμηλότερο από το προηγούμενο έτος), πέφτοντας στα 21,0 έτη για τις γυναίκες (0,8 έτη χαμηλότερα από το 2019) και 17,4 έτη για τους άνδρες (0,9 έτη χαμηλότερα από το 2019). Το 2020, το προσδόκιο ζωής στην ηλικία των 65 ετών μειώθηκε σημαντικά στα περισσότερα κράτη μέλη της ΕΕ, τόσο για τους άνδρες όσο και για τις γυναίκες (Πίνακας 3).

#### 2.3.4 Προσδόκιο ζωής άνω των 65 ετών

Πίνακας 3: Προσδόκιο ζωής στην ηλικία των 65 ετών 1980-2020

Πηγή: (Eurostat, 2022)

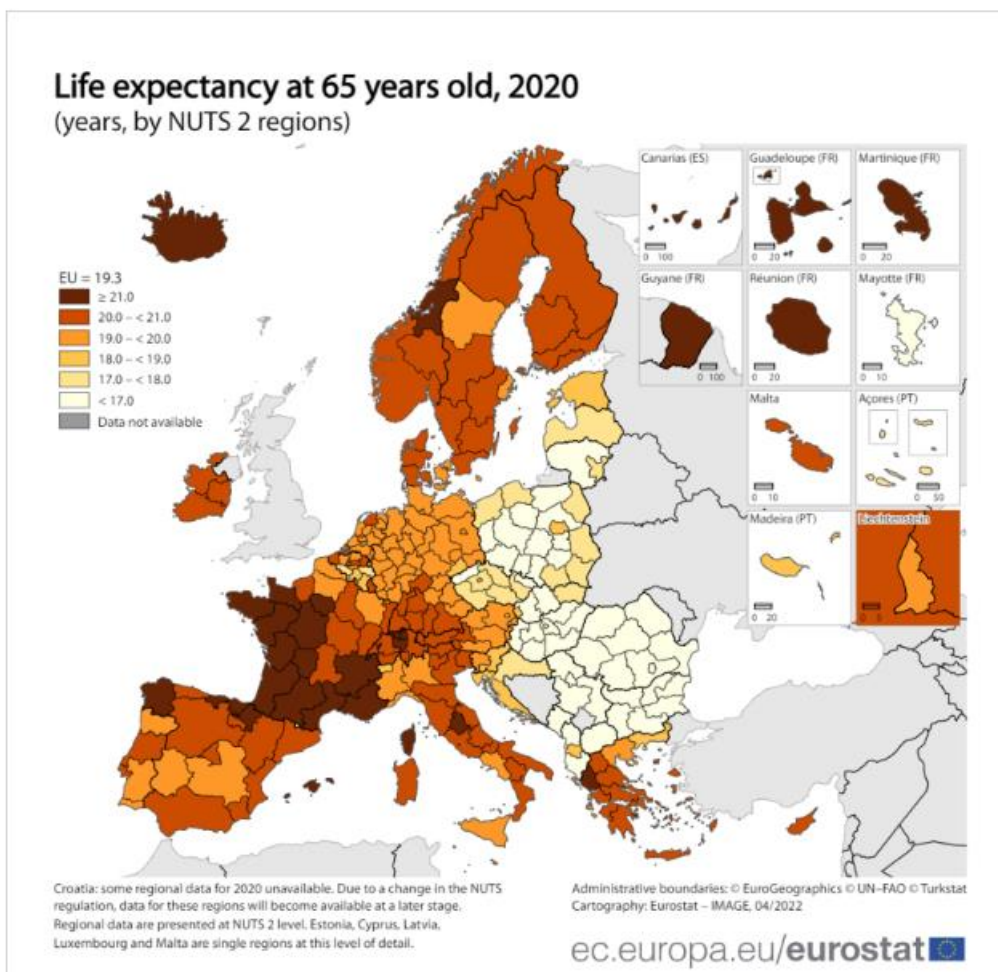
Life expectancy at age 65, 1980-2020 (years)																								
	Total										Males										Females			
	1980	1990	2000	2010	2018	2019	2020	1980	1990	2000	2010	2018	2019	2020	1980	1990	2000	2010	2018	2019	2020			
EU (*)	19.4	20.0	20.2	19.3	17.4	18.1	18.3	17.4	21.0	21.6	21.8	21.0												
Belgium	15.0	16.8	17.8	19.6	20.3	20.6	19.3	12.9	14.3	15.6	17.6	18.6	18.9	17.6	16.8	18.8	19.7	21.3	21.9	22.1	20.8			
Bulgaria	13.6	14.0	14.1	15.6	16.2	16.3	15.1	12.6	12.7	12.7	13.8	14.2	14.2	12.9	14.6	15.2	15.3	17.1	18.0	18.1	17.1			
Czechia	13.0	13.7	15.7	17.4	18.2	18.4	17.3	11.2	11.7	13.7	15.5	16.2	16.4	15.2	14.4	15.3	17.2	19.0	19.8	20.1	19.1			
Denmark	15.7	16.1	16.9	18.4	19.4	19.8	19.8	13.6	14.0	15.2	17.0	18.0	18.4	18.4	17.7	17.9	18.3	19.7	20.7	21.0	21.2			
Germany	14.9	16.3	18.0	19.5	19.6	19.9	19.7	12.8	14.0	15.8	17.8	18.0	18.3	18.0	16.3	17.7	19.6	20.9	21.1	21.4	21.2			
Estonia	14.2	14.4	15.4	17.4	18.6	19.0	19.0	11.8	12.0	12.7	14.3	15.7	15.9	15.9	15.6	15.8	17.1	19.5	20.6	21.1	21.1			
Ireland (†)	15.2	16.4	19.3	20.4	20.8	20.7	13.3	14.6	17.7	19.0	19.4	19.4	17.0	18.0	20.8	21.6	22.1	21.9						
Greece	16.2	17.0	18.0	19.7	20.6	20.4	20.0	15.2	15.7	16.7	18.2	19.1	19.0	18.5	17.0	18.1	19.2	21.0	21.9	21.7	21.4			
Spain	16.4	17.5	18.8	20.9	21.6	22.0	20.5	14.7	15.5	16.6	18.6	19.5	19.8	18.4	17.9	19.2	20.7	22.9	23.5	23.9	22.4			
France (†)	19.3	21.3	21.9	22.0	21.2	16.8	18.9	19.7	19.8	19.0	21.4	23.4	23.8	23.9	23.1									
Croatia	16.7	17.7	17.9	17.1	14.7	15.7	15.9	15.1	18.2	19.3	19.5	18.8												
Italy (†)	17.2	18.9	20.4	21.3	21.4	20.1	15.2	16.7	18.3	19.6	19.7	18.3	18.9	20.7	22.1	22.8	22.9	21.7						
Cyprus	17.2	19.7	20.5	20.3	20.3	15.9	18.3	19.1	18.9	19.1	18.3	21.0	21.8	21.5	21.5									
Latvia	16.1	17.1	17.4	17.0	13.1	14.1	14.4	14.0	18.1	19.0	19.4	19.1												
Lithuania	15.3	15.6	16.1	16.7	17.6	17.9	16.8	13.4	13.3	13.6	13.8	14.5	14.8	13.6	16.6	17.0	17.8	18.8	19.7	20.0	19.1			
Luxembourg	14.7	16.7	18.1	19.6	20.5	20.9	20.2	12.6	14.3	15.5	17.3	18.6	19.2	18.5	16.5	18.5	20.1	21.6	22.1	22.4	21.8			
Hungary	13.3	13.9	15.1	16.5	16.8	16.9	16.2	11.6	12.1	13.0	14.1	14.6	14.8	14.0	14.7	15.4	16.7	18.2	18.5	18.6	17.9			
Malta	17.0	19.9	20.8	21.1	20.5	10.7	15.2	18.5	19.2	19.4	18.9	18.6	21.1	22.3	22.5	22.0								
Netherlands	17.0	17.5	19.5	20.0	20.3	19.5	14.4	15.4	17.7	18.7	19.0	18.2	19.1	19.3	21.0	21.1	21.4	20.7						
Austria	14.9	16.6	18.1	19.8	20.1	20.3	19.6	12.9	14.4	16.0	17.9	18.5	18.7	17.9	16.3	18.1	19.6	21.4	21.6	21.7	21.0			
Poland (†)	14.6	15.8	17.6	18.2	18.5	17.1	12.4	13.5	15.1	15.8	16.1	14.6	16.2	17.5	19.5	20.1	20.4	19.2						
Portugal	14.7	15.7	17.4	19.3	20.3	20.6	19.8	13.1	14.0	15.4	17.2	18.2	18.5	17.8	16.1	17.1	19.1	21.0	22.0	22.3	21.6			
Romania (†)	13.4	14.3	14.8	16.1	16.7	16.9	15.7	12.5	13.2	13.4	14.2	14.7	14.9	13.4	14.2	15.2	15.9	17.6	18.4	18.6	17.7			
Slovenia	15.6	16.9	19.2	20.0	20.1	18.9	13.3	14.2	16.8	17.8	18.1	16.9	17.1	18.7	20.0	21.8	21.8	20.6						
Slovakia	13.7	14.3	15.0	16.3	17.6	17.9	17.1	12.0	12.3	12.9	14.1	15.4	15.7	14.8	15.2	16.0	16.7	18.0	19.3	19.7	18.9			
Finland	15.1	16.2	17.8	19.7	20.4	20.6	20.6	12.6	13.8	15.5	17.5	18.6	18.8	18.8	17.0	17.8	19.5	21.5	22.0	22.3	22.2			
Sweden	16.3	17.4	18.6	19.8	20.4	20.9	20.2	14.3	15.4	16.7	18.3	19.2	19.6	18.9	18.1	19.2	20.2	21.2	21.6	22.1	21.4			
Iceland	17.5	18.1	18.9	19.9	20.9	21.0	21.1	15.7	16.4	17.8	18.3	20.0	20.0	20.3	19.3	19.8	19.8	21.5	21.7	22.0	22.0			
Liechtenstein	17.5	20.7	20.5	21.3	19.8	15.2	19.6	19.0	20.0	18.1	19.5	21.8	21.9	22.5	21.5									
Norway	16.4	16.8	18.1	19.7	20.6	20.8	21.0	14.3	14.6	16.1	18.0	19.4	19.6	19.8	18.2	18.7	19.9	21.2	21.7	21.9	22.1			
Switzerland	16.5	17.7	19.2	20.9	21.7	21.7	20.8	14.3	15.3	17.0	19.0	20.2	20.3	19.3	18.2	19.7	20.9	22.5	23.0	23.0	22.2			
Montenegro	16.3	16.4	16.6	15.7	15.2	15.1	15.1	13.9	17.3	17.6	17.8	17.3												
North Macedonia	14.1	15.0	16.0	15.9	14.4	13.1	13.9	14.9	14.9	13.0	15.1	16.0	17.1	16.9	15.7									
Albania	17.8	18.1	16.6	17.2	17.5	15.6	18.5	18.7	17.7															
Serbia	13.6	15.2	16.0	16.2	15.0	12.5	14.0	14.7	14.8	13.4	14.6	16.2	17.2	17.3	16.5									
Turkey	17.3	18.3	18.4	15.6	16.6	16.6	18.7	19.9	20.0															

(\*) 2010, 2015, 2017 and 2019: breaks in series. 2018, 2019 and 2020: provisional.  
(†) 2019: estimated  
(‡) 2018, 2019, 2020: provisional  
(§) 2019: break in series  
(¶) 2009 and 2010: break in series. 2020: provisional  
(‡) 2020: estimated  
Source: Eurostat (online data code: demo\_mlexpec)

Μεταξύ των χωρών της EFTA, το 2020 η Ελβετία και το Λιχτενστάιν κατέγραψαν επίσης πτώση στο προσδόκιμο ζωής στην ηλικία των 65 ετών, ενώ η Ισλανδία και η Νορβηγία κατέγραψαν μικρή αύξηση. Το 2020, το προσδόκιμο ζωής των ανδρών ηλικίας 65 ετών κυμαινόταν από 18,1 έτη στο Λιχτενστάιν έως 20,3 χρόνια στην Ισλανδία, ενώ το προσδόκιμο ζωής για τις γυναίκες ηλικίας 65 ετών κυμαινόταν από τουλάχιστον 21,5 έτη στο Λιχτενστάιν έως 22,2 έτη κατ' ανώτατο όριο στην Ελβετία.

Κοιτάζοντας τον χάρτη σε περιφερειακό επίπεδο, το 2020 οι περιοχές στη Νότια Γαλλία και τη Βόρεια Ισπανία κατέγραψαν το υψηλότερο προσδόκιμο ζωής στην ηλικία των 65 ετών.



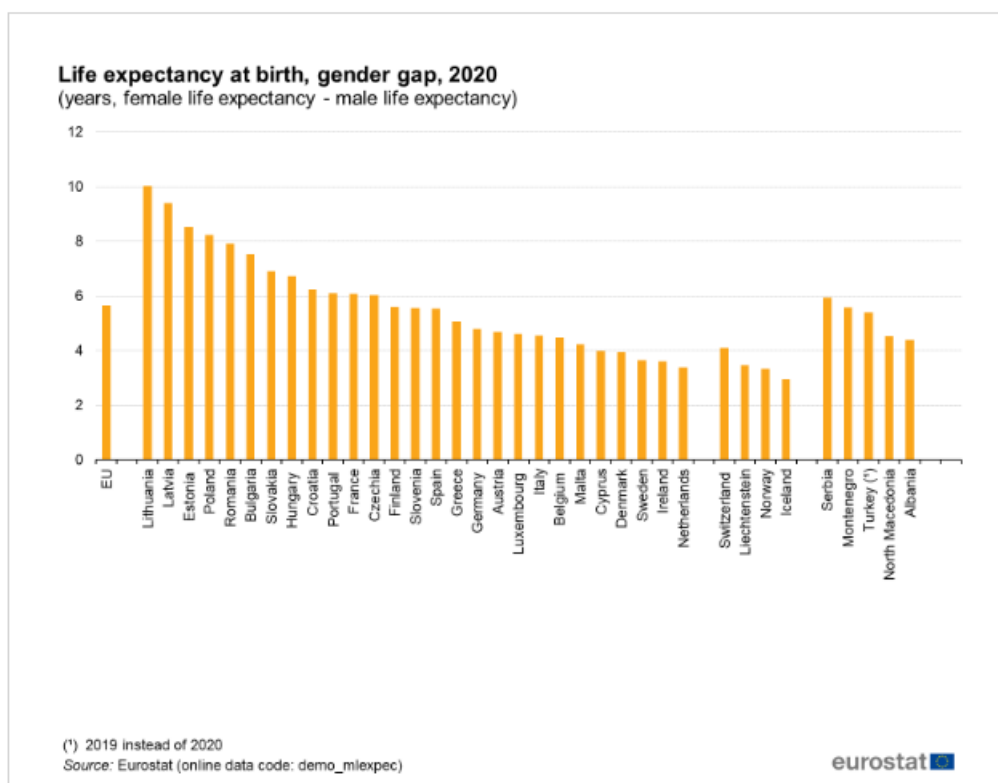


Σχήμα 2: Προσδόκιμο ζωής σε ηλικία 65 ετών ανά περιοχή NUTS 2, 2019

Πηγή: (Eurostat, 2022)

### 2.3.5 Διαφοροποιήσεις στο προσδόκιμο ζωής βάσει φύλου

Εξακολουθούν να υπάρχουν σημαντικές διαφορές μεταξύ των χωρών όσον αφορά το προσδόκιμο ζωής ανδρών και γυναικών. Το 2020, το προσδόκιμο ζωής για τις γυναίκες ήταν ακόμα υψηλότερο από το προσδόκιμο ζωής των ανδρών. Με ένα χάσμα μεταξύ των φύλων 5,6 ετών το 2020, τα νεογέννητα θηλυκά στην ΕΕ θα μπορούσαν γενικά να περιμένουν να επιβιώσουν περισσότερο από τους άνδρες. Επιπλέον, αυτό το χάσμα διέφερε σημαντικά μεταξύ των κρατών μελών της ΕΕ. Το 2020, η μεγαλύτερη διαφορά μεταξύ των φύλων βρέθηκε στη Λιθουανία (10,0 έτη) και η μικρότερη στην Ολλανδία (3,4 έτη)



Διάγραμμα 8: Προσδόκιμο ζωής κατά τη γέννηση αποκλίσεις μεταξύ των φύλων, 2020

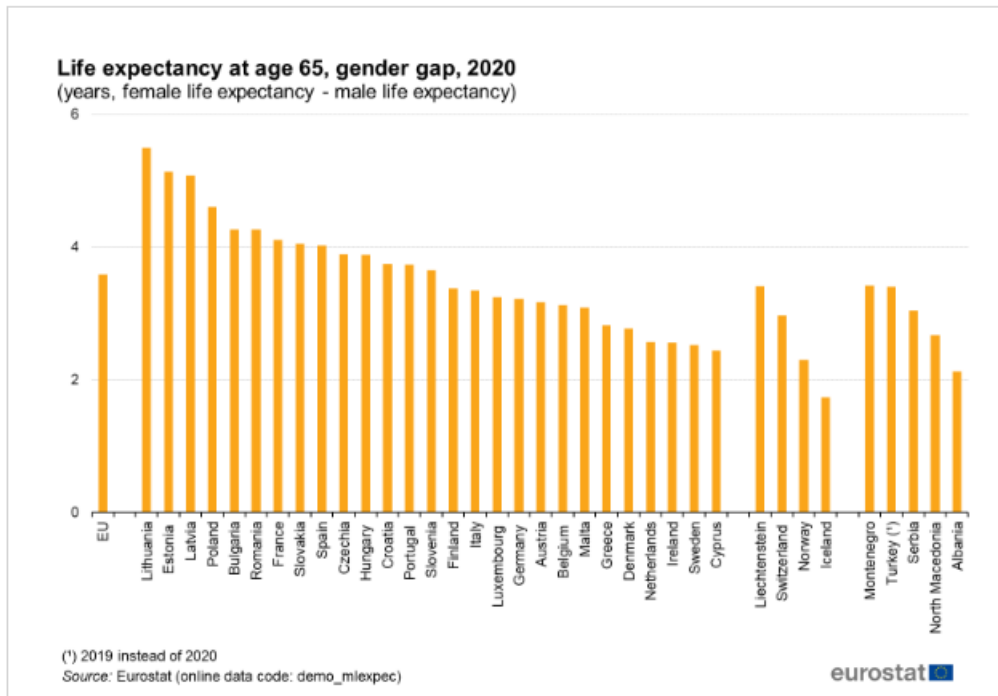
Πηγή: (Eurostat, 2022)

Όσον αφορά τη μείωση του προσδόκιμου ζωής για το 2020, οι άνδρες επηρεάστηκαν ελαφρώς περισσότερο στην πλειονότητα των κρατών μελών της ΕΕ, με τη μεγαλύτερη μείωση του προσδόκιμου ζωής να καταγράφεται στη Βουλγαρία (-1,6 έτη), την Πολωνία (-1,6 έτη), τη Λιθουανία, την Ισπανία και Ρουμανία (όλα -1,5 έτος).

Το 2020, όπως και τα προηγούμενα έτη, οι χώρες της ΕΕΤΑ κατέγραψαν διαφορές κάτω από τον μέσο όρο της ΕΕ μεταξύ του προσδόκιμου ζωής κατά τη γέννηση για άνδρες και γυναίκες, που κυμαίνονται από 4,1 χρόνια στην Ελβετία έως 2,9 στην Ισλανδία.

Εξετάζοντας το προσδόκιμο ζωής στην ηλικία των 65 ετών, φαίνεται ότι το χάσμα μεταξύ των δύο φύλων είναι μικρότερο από το χάσμα κατά τη γέννηση. Το 2020, οι γυναίκες ηλικίας 65 ετών στην ΕΕ θα μπορούσαν γενικά να αναμένουν ότι θα ζήσουν περισσότερο από τους άνδρες κατά 3,6 χρόνια. Η μεγαλύτερη διαφορά μεταξύ των φύλων το 2020 βρέθηκε στη Λιθουανία (5,5 έτη) και η μικρότερη στην Κύπρο (2,4) και τη Σουηδία (2,5)

Το Διάγραμμα 9 δείχνει ότι το 2020 οι γυναίκες ηλικίας 65 ετών στις χώρες της ΕΕΤΑ θα μπορούσαν να περιμένουν να ζήσουν περισσότερο από τους άνδρες κατά τουλάχιστον 1,7 χρόνια στην Ισλανδία έως 3,4 χρόνια κατ' ανώτατο όριο στο Λιχτενστάιν.



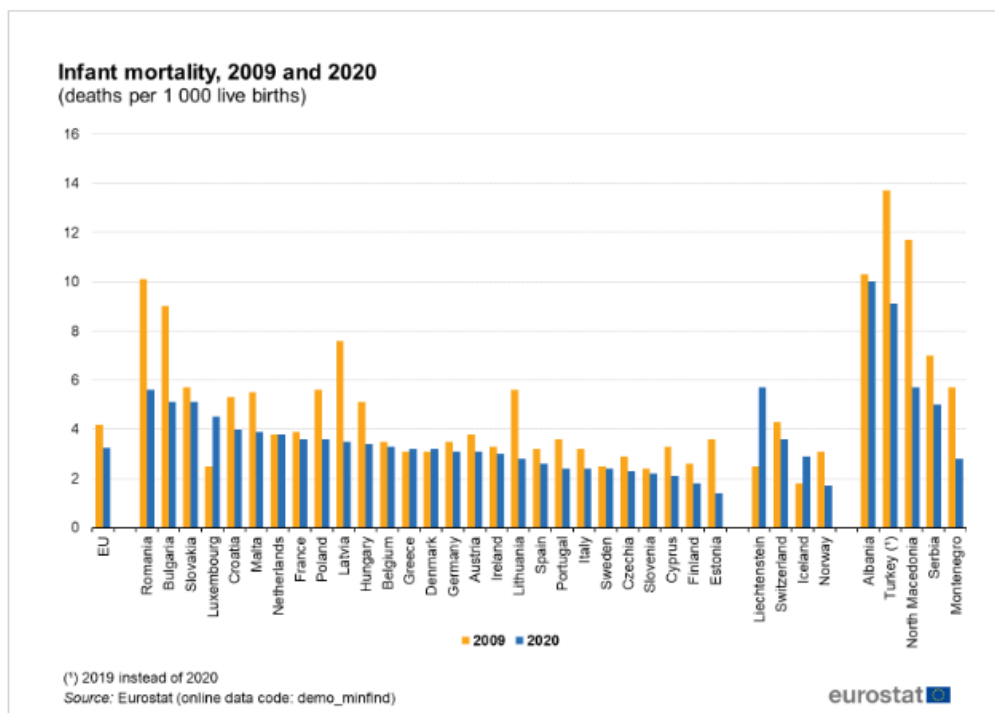
Διάγραμμα 9: Προσδόκιο ζωής στην ηλικία των 65 ετών, αποκλίσεις μεταξύ των φύλων 2020

Πηγή: (Eurostat, 2022)

### 2.3.6 Παιδική θνησιμότητα

Περίπου 13.250 παιδιά πέθαναν πριν συμπληρώσουν το ένα έτος στην ΕΕ το 2020 (849 λιγότερα από το προηγούμενο έτος). Αυτό ισοδυναμούσε με ποσοστό βρεφικής θνησιμότητας 3,3 θανάτων ανά 1.000 γεννήσεις ζώντων.





Διάγραμμα 10: Βρεφική θνησιμότητα 2009 και 2020 (θάνατοι ανά 1000 γεννήσεις ζώντων)

Πηγή: (Eurostat, 2022)

Μία από τις πιο σημαντικές αλλαγές που οδήγησαν σε αύξηση του προσδόκιμου ζωής κατά τη γέννηση ήταν η μείωση των ποσοστών βρεφικής θνησιμότητας. Μεταξύ 2009 και 2020, το ποσοστό βρεφικής θνησιμότητας στην ΕΕ μειώθηκε από 4,2 θανάτους ανά 1.000 γεννήσεις ζώντων σε 3,3 θανάτους ανά 1.000 γεννήσεις. Κατά την επέκταση της ανάλυσης στα τελευταία 20 χρόνια, το ποσοστό βρεφικής θνησιμότητας μειώθηκε σχεδόν στο μισό (6,2 θάνατοι ανά 1.000 το 1999). Οι πιο σημαντικές μειώσεις στη βρεφική θνησιμότητα καταγράφηκαν γενικά στα κράτη μέλη της ΕΕ που έτειναν να καταγράφουν υψηλότερα επίπεδα βρεφικής θνησιμότητας τα προηγούμενα χρόνια σε σύγκριση με τον μέσο όρο της ΕΕ.

Το 2020, τα υψηλότερα ποσοστά βρεφικής θνησιμότητας στην ΕΕ καταγράφηκαν στη Ρουμανία (5,6 θάνατοι ανά 1.000 γεννήσεις ζώντων γεννήσεων), τη Βουλγαρία και τη Σλοβακία (και οι δύο 5,1 θάνατοι ανά 1.000 γεννήσεις ζώντων γεννήσεων), ενώ τα χαμηλότερα στην Εσθονία (1,4 θάνατοι ανά 1 000 ζώντες γεννήσεις) και τη Φινλανδία (1,8 θάνατοι ανά 1 000 γεννήσεις ζώντων).

Στις χώρες της ΕΦΤΑ, τα ποσοστά βρεφικής θνησιμότητας κυμαίνονταν από τουλάχιστον 1,7 θανάτους ανά 1 000 γεννήσεις στη Νορβηγία έως 5,7 θανάτους ανά 1 000 γεννήσεις στο Λιχτενστάιν.

## 2.4 Μελλοντικές τάσεις θνησιμότητας: Χρησιμότητα και επιπτώσεις στα ασφαλιστικά προγράμματα

Καθώς οι προβλέψεις θνησιμότητας έχουν αποκτήσει ολοένα και μεγαλύτερη σημασία, έχουν αναπτυχθεί πολυάριθμα μοντέλα για τη μοντελοποίηση και την πρόβλεψη της θνησιμότητας (Pollard, 1987; Booth & Tickle, 2008). Οι διάφορες μέθοδοι πρόβλεψης της θνησιμότητας μπορούν να χωριστούν σε τρεις προσεγγίσεις: παρέκταση, εξήγηση και προσδοκία (Booth & Tickle 2008). Οι μέθοδοι παρέκτασης χρησιμοποιούν την κανονικότητα που συνήθως εντοπίζεται τόσο στα ηλικιακά πρότυπα όσο και στις τάσεις στο χρόνο. Η εξηγητική προσέγγιση κάνει χρήση δομικών ή επιδημιολογικών μοντέλων θνησιμότητας από ορισμένες αιτίες θανάτου για τις οποίες οι βασικές εξωγενείς μεταβλητές είναι γνωστές και μπορούν να μετρηθούν. Η προσέγγιση των προσδοκιών βασίζεται στις υποκειμενικές γνώμες των εμπειρογνομόνων που περιλαμβάνουν διάφορους βαθμούς τυπικότητας. Πρέπει να σημειωθεί ότι ορισμένες μέθοδοι πρόβλεψης της θνησιμότητας περιλαμβάνουν πτυχές μιας ή περισσότερων προσεγγίσεων.

Στο παρελθόν, οι περισσότερες μέθοδοι ήταν σχετικά απλές και βασίζονταν σε μεγάλο βαθμό στην υποκειμενικότητα (Pollard, 1987). Με την πάροδο του χρόνου, ωστόσο, αναπτύχθηκαν και εφαρμόστηκαν πιο εξελιγμένες μέθοδοι που χρησιμοποιούν όλο και περισσότερο τις συνήθεις στατιστικές μεθόδους (Booth and Tickle 2008). Η πλειονότητα αυτών των μεθόδων μπορεί να ταξινομηθεί στις προσεγγίσεις παρέκτασης, εκ των οποίων η μέθοδος Lee-Carter έχει καταστεί κυρίαρχη. Η μέθοδος αυτή συνοψίζει τη θνησιμότητα ανά ηλικία και περίοδο για έναν ενιαίο πληθυσμό ως μια συνολική χρονική τάση, μια ηλικιακή συνιστώσα και την έκταση της μεταβολής με την πάροδο του χρόνου ανά ηλικία (Lee & Carter, 1992). Ένα από τα δυνατά σημεία της μεθόδου Lee-Carter και των μεθόδων παρέκτασης γενικά είναι η ανθεκτικότητά τους σε καταστάσεις στις οποίες τα ποσοστά λογαριθμικής θνησιμότητας ανά ηλικία έχουν γραμμικές τάσεις (Booth et al., 2006). Ωστόσο, ορισμένες χώρες έχουν λιγότερο γραμμικές τάσεις (π.χ. Booth et al., 2002 για την Αυστραλία, Renshaw & Haberman, 2006 για την Αγγλία και την Ουαλία, Janssen et al., 2007 για τις Κάτω Χώρες). Συνεπώς, είναι σημαντικό να συζητηθεί αν θα πρέπει να χρησιμοποιηθούν απλώς «αντικειμενικές» μέθοδοι γραμμικής παρέκτασης, παρά τη μη γραμμικότητα των τάσεων, ή αν είναι προτιμότερο να προστεθούν πληροφορίες π.χ. με τη συμπίληψη του φαινομένου της κοορτής ή των τάσεων σε άλλες χώρες ή με τη χρήση πιο επεξηγηματικών μοντέλων, παρά την υποκειμενικότητα που αυτό συνεπάγεται.

Ένα παράδειγμα μεθόδου που περιλαμβάνει πρόσθετες πληροφορίες είναι η συνεκτική πρόβλεψη (Li & Lee, 2005). Αυτή η επέκταση του μοντέλου Lee-Carter επιδιώκει να εξασφαλίσει ότι οι προβλέψεις για συναφείς πληθυσμούς διατηρούν ορισμένες διαρθρωτικές σχέσεις με βάση για παράδειγμα ότι οι προβλέψεις για παρόμοιους χώρες δεν είναι ριζικά διαφορετικές. Η μέθοδος Lee-Carter έχει επίσης πρόσφατα επεκταθεί ώστε να συμπεριλάβει τη διάσταση της κοορτής (Renshaw & Haberman, 2006), και άλλες στοχαστικά μοντέλα έχουν εισαχθεί για την ενσωμάτωση της διάστασης της κοορτής στη θνησιμότητα πρόβλεψη (Cairns et al., 2011). Άλλα παραδείγματα είναι οι μέθοδοι πρόβλεψης που χρησιμοποιούν πολύτιμες ιατρικές γνώσεις και πληροφορίες σχετικά με συμπεριφορικές και περιβαλλοντικές αλλαγές, όπως το κάπνισμα ή/και η παχυσαρκία (Olshansky et al., 2005; Janssen et al., 2013). Αν και αυτοί οι νέοι τύποι μεθόδων έχουν πολλά πλεονεκτήματα, οι πιο επεξηγηματικές μέθοδοι εμπεριέχουν ένα μεγάλο στοιχείο υποκειμενικής κρίσης. Επιπλέον, καθώς όλες οι προαναφερθείσες μέθοδοι είναι αρκετά νέες, η ο αριθμός των φορών που έχουν εφαρμοστεί είναι ακόμη σχετικά μικρός.

Επιπλέον, έχουν σημειωθεί και άλλες γενικότερες πρόοδοι στην πρόβλεψη θνησιμότητας που αντιμετωπίζουν μια σημαντική πρόκληση που προκύπτει κατά τη χρήση παρεκβλητικών προσεγγίσεων: δηλαδή τον εντοπισμό της υποκείμενης μακροπρόθεσμης τάσης θνησιμότητας που μπορεί πραγματικά να χρησιμεύσει ως βάση για την παρέκταση (Janssen & Kunst, 2007). Αυτή η διαδικασία δεν είναι απλώς μια καθαρά μαθηματική στοχαστική διαδικασία, αλλά απαιτεί προσεκτική ανάλυση των προηγούμενων εθνικών και διεθνών τάσεων θνησιμότητας και των καθοριστικών παραγόντων της (Janssen & Kunst, 2007).

Μια σημαντική σχετική πρόοδος στην πρόβλεψη της θνησιμότητας είναι η (περαιτέρω) ανάπτυξη συνεκτικών μεθόδων πρόβλεψης πολλαπλών πληθυσμών. Έχουν αναπτυχθεί συνεκτικές προβολές θνησιμότητας πολλών πληθυσμών για την αποφυγή μη ρεαλιστικών διασταυρώσεων ή αποκλίσεων στη μελλοντική θνησιμότητα μεταξύ χωρών ή φύλων, που μπορεί να προκύψουν από την εφαρμογή μοντέλων προβολής θνησιμότητας σε μεμονωμένους πληθυσμούς/χώρες, αυτή η μέθοδος αναφέρεται στην βιβλιογραφία ως ατομική πρόβλεψη θνησιμότητας (Li & Lee, 2005; Enchev et al., 2017). Η κύρια ιδέα πίσω από τη συνεκτική πρόβλεψη είναι ότι οι προβλέψεις θνησιμότητας για πληθυσμούς με παρόμοιες εξελίξεις θνησιμότητας δεν θα αποκλίνουν ριζικά, αλλά ότι οι δομικές διαφορές θα παραμείνουν (για παράδειγμα, σταθερά υψηλότερη θνησιμότητα για τους άνδρες από τις γυναίκες) (Hyndman et al., 2013). Οι συνεκτικές προβλέψεις οδηγούν επίσης σε βελτιωμένη προγνωστική ισχύ επειδή οι τάσεις της δομικής θνησιμότητας μπορούν να προβλεφθούν καλύτερα όταν

χρησιμοποιούνται τάσεις σε παρόμοιους πληθυσμούς (π.χ. χωρών) για τον προσδιορισμό των πιο πιθανών μελλοντικών εξελίξεων (Janssen & Kunst, 2007). Η επέκταση του μοντέλου Lee-Carter σε πολλαπλούς πληθυσμούς από τους Li & Lee (2005) ήταν ο αυτή που έθεσε τις τάσεις. Κατόπιν, αναπτύχθηκαν πρόσθετες συνεκτικές μέθοδοι πρόβλεψης που βασίστηκαν στη δομή Lee-Carter (Li & Hardy, 2011; Russolillo et al., 2011; Li, 2013; Wan et al., 2013; Wang & Yang, 2013; Zhou et al., 2014; Kleinow, 2015), καθώς και σχετικές με τη δομή ηλικιακής περιόδου-κοόρτης (Cairns et al., 2011b; Bdrger & Aleksic, 2014) και την προσέγγιση συνάρτησης δεδομένων (Shang & Hyndman, 2017). Πρόσφατες μελέτες αξιολόγησης έχουν δείξει ότι αυτά τα συνεκτικά μοντέλα πρόβλεψης θνησιμότητας είναι όντως ανώτερα από τα μεμονωμένα μοντέλα πρόβλεψης θνησιμότητας (Shair et al., 2017). Συνεκτικά μοντέλα πρόβλεψης πολυπληθυσμιακής θνησιμότητας έχουν εφαρμοστεί κυρίως σε μια ομάδα χωρών (Torri & Vaupel, 2012), στα δύο φύλα σε έναν πληθυσμό (Hatzopoulos & Haberman, 2013), ή σε περιοχές εντός μιας χώρας (Bergeron-Boucher et al., 2017 και Alexander et al., 2017). Ωστόσο, αυτά τα μοντέλα πρόβλεψης έχουν εφαρμοστεί σπάνια σε κοινωνικο-οικονομικές ομάδες (van Baal et al., 2016; Cairns et al., 2016), παρόλο που υπάρχουν σημαντικές διαφορές στη θνησιμότητα ανά κοινωνικοοικονομική ομάδα (Murtin et al., 2017).

Μια άλλη σημαντική πρόοδος στις μοντελοποιήσεις πρόβλεψης θνησιμότητας είναι η συμπερίληψη πληροφοριών σχετικά με τους καθοριστικούς παράγοντες των τάσεων του παρελθόντος στην πρόβλεψη θνησιμότητας. Αν και η χρήση μιας αντικειμενικής προσέγγισης πρόβλεψης της θνησιμότητας με παρεκβολή έχει σαφή πλεονεκτήματα, δεν λαμβάνει υπόψη σημαντικές πληροφορίες σχετικά με τους καθοριστικούς παράγοντες των τάσεων του παρελθόντος και τις σημαντικές (χρονικές) αποκλίσεις από τη συνολική γραμμική τάση θνησιμότητας (Janssen & Kunst 2007; Olshansky et al., 2009). Κατά συνέπεια, η χρήση γενικών μεθόδων πρόβλεψης της θνησιμότητας, ειδικά σε μια κατάσταση όπου οι τάσεις θνησιμότητας στο παρελθόν ήταν μη γραμμικές, μπορεί να οδηγήσει σε διαφορετικά αποτελέσματα όταν επιλέγονται διαφορετικές ιστορικές περιόδους (Stoeldraijer et al., 2013) και το αποτέλεσμα της επιλογής της ιστορικής περιόδου μπορεί να έχει μεγαλύτερη επίδραση στον τύπο της μεθόδου παρεκβολής πρόβλεψης που επιλέχθηκε (Stoeldraijer et al., 2013).

Επιπλέον, η χρήση γενικών μεθόδων πρόβλεψης της θνησιμότητας με παρεκβολή μπορεί επίσης να δημιουργήσει απίθανα μελλοντικά μοτίβα όταν έχουν σημειωθεί χρονικές αυξήσεις ή μειώσεις της θνησιμότητας, καθώς αυτές θεωρούνται αυτόματα δομικές στις προσεγγίσεις πρόβλεψης της θνησιμότητας (Janssen et al., 2013). Για να ξεπεραστούν αυτά τα ζητήματα, αρκετές πρόσφατες μελέτες έχουν ενσωματώσει επιδημιολογικές πληροφορίες στην πρόβλεψη

θνησιμότητας, κυρίως μέσω πληροφοριών σχετικά με το κάπνισμα ή/και την παχυσαρκία - (Pampel, 2005; Bongaarts, 2006; Wang & Preston, 2009; King & Soneji als, 2011; Janssen et al., 2013; French & O'Hare, 2014). Αυτές οι μελέτες χρησιμοποίησαν διαφορετικές προσεγγίσεις. Πολλές από τις μελέτες που επικεντρώθηκαν στις ΗΠΑ περιελάμβαναν λεπτομερείς πληροφορίες για τον τρόπο ζωής από την εθνική έρευνα για την υγεία στο μοντέλο πρόβλεψης (Wang & Preston, 2009; King & Soneji, 2011; Preston et al., 2012).

Οι French & O'Hare (2014) περιέλαβαν συγκεντρωτικές συμμεταβλητές σχετικά με τη συμπεριφορά υγείας (π.χ. κατανάλωση αλκοόλ και καπνού ηλικίας 15 ετών και άνω) στο ιεραρχικό μοντέλο Girosi και King Bayes (King & Soneji, 2011). Οι Bongaarts (2006) και Janssen & Kunst (2007) επικεντρώθηκαν στην πρόβλεψη των πιο γραμμικών τάσεων του παρελθόντος στη θνησιμότητα που δεν σχετίζεται με το κάπνισμα, παρά στη θνησιμότητα-όλων-των-αιτιών. Ο Pampel (2005) προέβλεψε διαφορές φύλου στη θνησιμότητα-όλων-των-αιτιών συνδυάζοντας μια πρόβλεψη της αναλογίας φύλων στη θνησιμότητα που σχετίζεται με το κάπνισμα - χρησιμοποιώντας ένα μοντέλο παλινδρόμησης που εκφράζει τη σχέση μεταξύ του επιπολασμού του καπνίσματος και της θνησιμότητας που σχετίζεται με το κάπνισμα 25 χρόνια αργότερα - με τις υποτιθέμενες αλλαγές στην αναλογία φύλου σε αιτίες θανάτου που δεν σχετίζονται με το κάπνισμα. Οι Janssen et al. (2013) συνδύασαν την πρόβλεψη της σχετιζόμενης με το κάπνισμα θνησιμότητας μέσω της μοντελοποίησης ηλικιακής περιόδου-κοόρτης με την πρόβλεψη της θνησιμότητας που σχετίζεται με το κάπνισμα μέσω συμβατικών τεχνικών πρόβλεψης της θνησιμότητας με παρεκβολή. Πρόσφατα, ο Bongaarts (2014) έδειξε ότι η χρήση μιας τέτοιας νέας προσέγγισης είναι καλύτερη από την απευθείας εκτέλεση προβολών θνησιμότητας με συγκεκριμένο αίτιο -δηλαδή, παρά απλώς τη χρήση μιας επεξηγηματικής προσέγγισης- καθώς με αυτόν τον τρόπο εκμεταλλεύεται την προβλεψιμότητα της ασυνέχειας στις τάσεις θνησιμότητας ως αποτέλεσμα μιας αιτίας συμπεριφοράς, η οποία απουσιάζει για τακτικές αιτίες θανάτου. Οι Janssen et al. (2013) συνδύασαν τη συμπερίληψη της επιδημίας του καπνίσματος στην πρόβλεψη θνησιμότητας με διεθνώς συνεκτική πρόβλεψη θνησιμότητας.

Άλλες σημαντικοί πρόοδοι της πρόβλεψης θνησιμότητας που έχουν σημειωθεί στην πρόσφατη βιβλιογραφία περιλαμβάνουν:

- (i) Τη χρήση διαφορετικών μέτρων πέρα από τα ποσοστά θνησιμότητας για την πρόβλεψη θνησιμότητας, όπως η χρήση ρυθμών βελτίωσης θνησιμότητας, που είτε εφαρμόζονται στην ατομική πρόβλεψη θνησιμότητας (Haberman & Renshaw,

2012; Haberman & Renshaw, 2013; Mitchell et al., 2013) είτε σε συνεκτική πρόβλεψη θνησιμότητας πολλών πληθυσμών (Schinzinger et al., 2016; Bohk-Ewald & Rau, 2017)

- (ii) Η συμπερίληψη μέτρων παραγωγής πέρα από το προσδόκιμο ζωής κατά τη γέννηση, όπως η ανισότητα στη διάρκεια ζωής (Bohk-Ewald et al., 2017).
- (iii) Η χρήση πολλαπλών προβλέψεων συνδυασμένων αντί για μία μόνο πρόβλεψη (Shang 2012; Kontis et al., 2017).
- (iv) Η χρήση της Bayesian πιθανολογικής πρόβλεψης θνησιμότητας αντί της γενικής στοχαστικής/πιθανολογικής πρόβλεψης θνησιμότητας για να γίνουν δηλώσεις σχετικά με ό,τι δεν είναι γνωστό, υπό την προϋπόθεση ότι είναι γνωστό, που είτε εφαρμόζεται σε ατομικές προβλέψεις (Czado et al., 2005; Pedroza, 2006; Kogure & Kurachi, 2010; Wisniowski et al., 2015; Wong et al., 2017; Hilton et al., 2018) ή σε συνεκτική πρόβλεψη θνησιμότητας πολλών πληθυσμών (Cairns et al., 2011b; Antonio et al., 2011b; Raftery et al., 2013)
- (v) Την επικύρωση του μοντέλου που χρησιμοποιείται για την πρόβλεψη θνησιμότητας μέσω της πρόβλεψης εντός δείγματος (= backtesting) (Cairns et al., 2011a; Hyndman & Athanopoulos, 2018).

## Κεφάλαιο 3<sup>ο</sup>: Το μοντέλο Lee-Carter

Η προσέγγιση των κύριων συνιστωσών για την πρόβλεψη της θνησιμότητας έγινε γνωστή μέσω της εργασίας των Lee & Carter (1992), αν και είχε εφαρμοστεί από νωρίτερα (Bell & Monsell, 1991). Σε αντίθεση με τις παραμετρικές προσεγγίσεις που προσδιορίζουν εκ των προτέρων τη λειτουργική μορφή του ηλικιακού προτύπου της θνησιμότητας, οι προσεγγίσεις των κύριων συνιστωσών εκτιμούν το ηλικιακό πρότυπο από τα δεδομένα. Η πρόβλεψη γίνεται με παρέκταση των παραμέτρων που σχετίζονται με το χρόνο με τη χρήση μεθόδων χρονοσειρών.

Η μέθοδος Lee-Carter είναι πλέον η κυρίαρχη μέθοδος πρόβλεψης της θνησιμότητας. Lee-Carter ή οι παραλλαγές της έχουν εφαρμοστεί στους πληθυσμούς των χωρών G7 (Tuljapurkar et al., 2000), στην Αυστραλία (Booth & Tickle, 2003, 2004), στην Ιαπωνία (Wilmoth, 1996) και στη Σουηδία (Lundstrom & Qvist, 2004) μεταξύ άλλων, και Οι δείκτες αναφοράς Lee-Carter χρησιμοποιούνται από το Γραφείο Απογραφής των ΗΠΑ για τις μακροπρόθεσμες προβλέψεις του προσδόκιμου ζωής (Hollmann et al., 2000).

### 3.1 Το βασικό μοντέλο Lee-Carter

Το μοντέλο των Lee-Carter χαρακτηρίζεται από μεγάλη απλότητα για τις προβλέψεις θνησιμότητας, προσδιορίζοντας μια λογαριθμογραμμική μορφή για τη τάση της θνησιμότητας  $m_x(t)$ . Η μέθοδος είναι στην ουσία ένα σχεσιακό μοντέλο, η βασική παραδοχή του μοντέλου είναι ότι υπάρχει γραμμική σχέση μεταξύ του λογαρίθμου των ποσοστών θανάτου ανά ηλικία  $m_x(t)$  και δύο επεξηγηματικών παραγόντων: του αρχικού ηλικιακού διαστήματος  $x$  και του χρόνου  $t$ . Οι πληροφορίες κατανέμονται σε ηλικιακά διαστήματα, οπότε το διάστημα που αρχίζει με την ηλικία  $x$  ονομάζεται ηλικιακό διάστημα  $x$  (Wang, 2007). Η εξίσωση που περιγράφει αυτό το γεγονός έχει ως εξής:

$$m_{x,t} = \exp(a_x + \beta_x k_t + e_{x,t}) \text{ με } t=1, \dots, n \text{ και } \chi=1, \dots, \omega \quad (1)$$

ή την εφαρμογή του λογαρίθμου:

$$f_{x,t} = \ln(m_{x,t}) = a_x + \beta_x k_t + e_{x,t} \text{ με } t=1, \dots, n \text{ και } \chi=1, \dots, \omega \quad (2)$$

όπου

$m_{x,t}$  : Είναι το ποσοστό θανάτου ανά ηλικία για το διάστημα  $x$  και το έτος  $t$ .

$k_t$  : Είναι ο δείκτης θνησιμότητας κατά το έτος  $t$ .

$a_x$  : Είναι η μέση θνησιμότητα ανά ηλικία.

$\beta_x$  : Είναι μια απόκλιση στη θνησιμότητα που οφείλεται σε μεταβολές του δείκτη  $k_t$

$e_{x,t}$  : Είναι το τυχαίο σφάλμα.

$\omega$ : Είναι η αρχή του τελευταίου διαστήματος ηλικίας.

#### 3.1.1 Προσαρμογή του μοντέλου

Η σχέση (1) δείχνει ότι υπάρχει μια εξίσωση για κάθε χρονικό διάστημα και ηλικιακό διάστημα. Ως αποτέλεσμα, πρέπει να επιλυθεί ένα σύστημα ταυτόχρονων εξισώσεων, ώστε να εκτιμηθούν οι τιμές των  $a_x$ ,  $\beta_x$  και  $k_t$ , οι οποίες αποτελούν τις λύσεις του συστήματος. Μια ομάδα ποσοστών θανάτου με διαφορετικές  $r$  ηλικιακές ομάδες που αναλύθηκαν σε διαφορετικές χρονικές στιγμές δημιουργεί ένα σύστημα εξισώσεων που περιέχει  $2r + n$  άγνωστους παράγοντες που αντιστοιχούν στο άθροισμα των  $r$  τιμών του  $a_x$ ,  $r$  του  $\beta_x$  και  $n$  των  $k_t$  και  $r \times n$  εξισώσεων. Το σύστημα αυτό μπορεί να γραφτεί σε πίνακα ως:



$$M = A + \beta \cdot k, \quad (3)$$

όπου  $M$  είναι ένας πίνακας  $r \times n$  και κάθε στοιχείο  $M_{i,j}$  αντιστοιχεί στον φυσικό λογάριθμο του ηλικιακού ποσοστού θανάτου  $i$  στο έτος  $j$ .

Τα στοιχεία που ανήκουν στις ίδιες κατηγορίες του πίνακα  $A$   $r \times n$  είναι ίσα, για παράδειγμα:  $a_{1j} = a_{2j} = \dots = a_{rj}$  για το ίδιο έτος  $j$ , ενώ το  $\beta$  είναι ένα διάνυσμα  $r \times 1$  και το  $k$  είναι ένα διάνυσμα  $1 \times n$ .

Αρχικά, δεν υπάρχει μια μοναδική λύση για αυτό το σύστημα. Έτσι, είναι απαραίτητο να προστεθούν οι ακόλουθοι δύο περιορισμοί ώστε να ληφθεί μια μοναδική λύση:

$$\sum_{x=1}^{\omega} k_x = 0$$

και

$$\sum_{i=1}^n \beta_x = 1$$

Με τη χρήση αυτών των περιορισμών, οι συντελεστές  $a_x$  είναι ένας απλός αριθμητικός μέσος όρος με την πάροδο του χρόνου για τους λογαρίθμους των συγκεκριμένων ποσοστών. Έτσι, οι παράμετροι  $\beta_x$  και  $k_t$  προσδιορίζονται μοναδικά. Επομένως, οι συντελεστές  $a_x$  υπολογίζονται από την ακόλουθη σχέση:

$$a_x = \frac{\sum_{t=1}^n \ln(m_{x,t})}{n} \quad (4)$$

Αφού προσδιοριστούν οι τιμές του πίνακα  $A$ , το σύστημα (3) μπορεί να ξαναγραφεί ως εξής:

$$M^* = M - A = b \cdot k \quad (5)$$

Το σύστημα αυτό έχει μια μοναδική λύση όταν περιλαμβάνονται αυτοί οι περιορισμοί. Τώρα, οι παράμετροι  $\beta$  και  $k$  πρέπει να προσδιοριστούν με την εφαρμογή της μεθόδου Αποσύνθεσης μοναδιαίας αξίας (SVD)<sup>1</sup>. Η μέθοδος αυτή χρησιμοποιείται για την ακριβή προσαρμογή των ελαχίστων τετραγώνων (Good, 1969). Μέσω της SVD, ο  $M^*$  μπορεί να γραφεί ως το γινόμενο

---

<sup>1</sup> Singular Value Decomposition

δύο πινάκων. Στον  $M$ , το στοιχείο  $(i, j)$  είναι το άθροισμα του γινομένου της γραμμής  $i$  του  $B$  και της γραμμής  $j$  του  $K$ , δηλ:

$$m_{i,j} = \sum_{l=1}^r B_{i,l} K_{j,l}^T$$

Επομένως, η αποσύνθεση δημιουργεί  $r$  όρους που ταιριάζουν ακριβώς στο στοιχείο  $m_{i,j}$  του πίνακα  $M^*$ . Οι Lee και Carter (1992) προτείνουν τη χρήση του  $M$  ως το γινόμενο των διανυσμάτων  $\beta$  και  $k$ . Κατά την εφαρμογή της SVD, έλαβαν υπόψη τους μόνο την προσέγγιση πρώτης τάξης, δηλαδή ο  $M'$  μπορεί να εκφραστεί ως εξής:

$$M' \approx B_1 K_1^T \quad (7)$$

Έτσι, προσδιορίζονται τα  $\beta_1 = B$  και  $k_1 = K$ , πράγμα που σημαίνει ότι υπάρχει μια πρώτη εκτίμηση των παραμέτρων του μοντέλου (Gonzalez Perez & Guerrero Guzman, 2007).

### 3.1.2 Διορθώσεις της πρώτης προσαρμογής

Συνήθως, οι τιμές που λαμβάνονται μέσω της πρώτης εκτίμησης του μοντέλου δεν παρέχουν επαρκή προσαρμογή στα παρατηρούμενα δεδομένα. Εξαιτίας αυτού, μπορεί να υπάρξουν αποκλίσεις στις προβολές, όπως επισημαίνουν οι Lee και Carter (1992) και αργότερα ο Bell (1997). Έτσι, απαιτείται ένα δεύτερο βήμα για την εκτίμηση των παραμέτρων. Αυτό το βήμα χρησιμοποιεί τις τιμές  $a_x$  και  $\beta_x$  από το πρώτο βήμα για να λάβει μια νέα εκτίμηση του  $k$ , ώστε να υπάρχει ένας παρατηρούμενος συνολικός αριθμός θανάτων για το εν λόγω έτος. Ο στόχος είναι να διαπιστωθεί ότι οι τιμές  $k_t$  συμμορφώνονται με την:

$$D_t = \sum_{x=0}^{\omega} N_{x,t} \exp(a_x + \beta_x k_t + e_{x,t}) \quad (8)$$

όπου  $D_t$  υποδηλώνει τον συνολικό αριθμό θανάτων κατά το έτος  $t$ ,  $N_{x,t}$  είναι ο πληθυσμός που ανήκει στο ηλικιακό διάστημα  $x$  κατά το έτος  $t$ , και  $\omega$  είναι η ηλικία έναρξης της τελευταίας παρατηρούμενης ομάδας στους πίνακες θνησιμότητας (Lee, & Carter, 1992).

### 3.1.3 Πρόβλεψη ποσοστών θανάτου ανά ηλικία

Αφού ληφθεί η χρονοσειρά του δείκτη  $k_t$  όπως σημειώθηκε παραπάνω (8), είναι δυνατόν να προβλεφθεί ο εν λόγω δείκτης με τη χρήση των μοντέλων Αυτοπαλίνδρομου Ολοκληρωμένου

Κινητού Μέσου Όρου (ARIMA)<sup>2</sup> και Στατιστικής μοντελοποίησης Σχήματος (SSM)<sup>3</sup>. Μετά την πρόβλεψη του δείκτη  $k_t$ , είναι δυνατόν να ληφθούν τα ποσοστά θανάτου για τα προβλεπόμενα έτη. Οι τιμές πρόβλεψης του  $k_{n+h}$  αντικαθίστανται στον τύπο:

$$\hat{m}_{x,n+h} = \hat{m}_{x,n} \exp\{\widehat{\beta}_x(\hat{k}_{n+h} - \hat{k}_n)\}, h = 1,2, \dots, x = 1,2 \dots \omega \quad (9)$$

όπου  $n$  είναι το τελευταίο έτος από το οποίο υπάρχουν διαθέσιμα στοιχεία,  $h$  είναι ο ορίζοντας πρόβλεψης και το  $x$  αντιπροσωπεύει την ηλικιακή ομάδα. Οι προβλέψεις των ποσοστών θανάτου σε σχέση με το τελευταίο διαθέσιμο ποσοστό θανάτου λαμβάνονται μέσω του τύπου (9). Οι Lee και Carter (1992) προτείνουν τη χρήση ενός προσεγγιστικού διαστήματος πρόβλεψης για την πρόβλεψη των ποσοστών θανάτου. Για τον υπολογισμό του διαστήματος χρησιμοποιούνται οι εκτιμήσεις των  $\beta_x$ , των παραμέτρων και των τυπικών σφαλμάτων των προβλέψεων  $k_t$ .

$$PI: \{m_{x,t} \exp(2\beta_x se_{k_t})\} \{m_{x,t} \exp(-2\beta_x se_{k_t})\} \quad (10)$$

### 3.1.4 Προσδόκιμο ζωής κατά τη γέννηση

Μέσω των προβλεπόμενων ποσοστών θανάτου, εφαρμόζεται η κλασική μέθοδος του Chiang (1984), ώστε να προκύψει το προσδόκιμο ζωής κατά τη γέννηση. Το προσδόκιμο ζωής στο  $x$ ,  $e_x$ , ορίζεται ως εξής:

$$e_x = \frac{T_x}{L_x}$$

όπου  $T_x$  είναι ο αθροιστικός αριθμός ετών που έζησε ο πληθυσμός της κοόρτης στο ηλικιακό διάστημα και σε όλα τα επόμενα ηλικιακά διαστήματα. Η τιμή  $L_x$  αντιπροσωπεύει τον αριθμό των ζώντων ατόμων στην αρχή του ηλικιακού διαστήματος  $x$  από έναν πληθυσμό  $L_0$  νεογέννητων μωρών. Συνήθως ορίζεται ως  $L_0 = 100.000$ .

## 3.2 Μοντελοποίηση Poisson για τον αριθμό των θανάτων και μεθοδολογία Lee-Carter

---

<sup>2</sup> Autoregressive Integrated Moving Average

<sup>3</sup> Statistical Shape Modeling

Η αρχική μέθοδος Lee-Carter χρησιμοποιήθηκε για τη συγκέντρωση και των δύο φύλων των δεδομένων των ΗΠΑ. Οι Lee και Carter (1992) εφάρμοσαν το υπόδειγμά τους για τους άνδρες και τις γυναίκες χωριστά, δείχνοντας ότι οι δύο σειρές είναι καλύτερο να αντιμετωπίζονται ως ανεξάρτητα μειούμενες. Ο Wilmoth (1996) εφάρμοσε τις μεθόδους Lee-Carter για την πρόβλεψη της ιαπωνικής θνησιμότητας και πειραματίστηκε επίσης με παραλλαγές αυτού του μοντέλου. Οι Lee και Nault (1993) εφάρμοσαν τις μεθόδους Lee-Carter για να υποδείξουν τη θνησιμότητα στον Καναδά, οι Lee και Rofman (1994) προσάρμοσαν την εξίσωση (2) στα δεδομένα της Χιλής και οι Brouhns και Denuit (2001b) έκαναν το ίδιο για τα βελγικά στατιστικά στοιχεία.

Θα πρέπει να σημειωθεί ότι η μέθοδος Lee-Carter δεν επιχειρεί να ενσωματώσει υποθέσεις σχετικά με την πρόοδο της ιατρικής επιστήμης ή συγκεκριμένες περιβαλλοντικές αλλαγές, επί της ουσίας δεν λαμβάνονται υπόψη άλλες πληροφορίες εκτός από το προηγούμενο ιστορικό. Αυτό σημαίνει ότι η προσέγγιση αυτή δεν είναι σε θέση να προβλέψει ξαφνικές βελτιώσεις στη θνησιμότητα λόγω της ανακάλυψης νέων ιατρικών θεραπειών ή επαναστατικών θεραπειών, συμπεριλαμβανομένων των αντιβιοτικών. Ομοίως, μελλοντικές επιδεινώσεις που προκαλούνται από επιδημίες, την εμφάνιση νέων ασθενειών ή την επιδείνωση της ρύπανσης δεν μπορούν να εισέλθουν στο μοντέλο. Ο αναλογιστής πρέπει να το έχει αυτό κατά νου όταν καθορίζει το αντασφαλιστικό του πρόγραμμα.

Η μεθοδολογία Lee-Carter είναι μια απλή παρέκταση των τάσεων του παρελθόντος. Όλες οι αμιγώς εξωστρεφείς προβλέψεις υποθέτουν ότι το μέλλον θα είναι κατά κάποιο τρόπο παρόμοιο με το παρελθόν. Ορισμένοι συγγραφείς (Gutterman & Vanderhoof, 2000) άσκησαν δριμύεια κριτική σε αυτή την προσέγγιση, επειδή φαίνεται να αγνοεί τους υποκείμενους μηχανισμούς. Όπως επισημαίνεται από τον Wilmoth (2000), μια τέτοια κριτική είναι έγκυρη μόνο στο βαθμό που οι μηχανισμοί αυτοί κατανοούνται με επαρκή ακρίβεια ώστε να προσφέρουν μια νόμιμη εναλλακτική μέθοδο πρόβλεψης. Καθώς η κατανόηση των πολύπλοκων αλληλεπιδράσεων των κοινωνικών και βιολογικών παραγόντων που καθορίζουν τα επίπεδα θνησιμότητας εξακολουθεί να είναι ασαφής, η εξωστρεφής προσέγγιση της πρόβλεψης είναι ιδιαίτερα επιτακτική στην περίπτωση της ανθρώπινης θνησιμότητας.

### 3.2.1 Μοντέλο

Σύμφωνα με τον Alho (2000), το μοντέλο που περιγράφεται στην εξίσωση (2) δεν είναι κατάλληλο σε αρκετές περιπτώσεις. Το κύριο μειονέκτημα της εκτίμησης μέσω SVD είναι ότι βασίζεται στην υπόθεση της ομοσκεδαστικότητας των σφαλμάτων. Αυτό σχετίζεται με το

γεγονός ότι η εξαγωγή συμπερασμάτων στην πραγματικότητα βασίζεται στην υπόθεση ότι τα σφάλματα κατανέμονται κανονικά, γεγονός που είναι μη ρεαλιστικό. Ο λογάριθμος της παρατηρούμενης έντασης θνησιμότητας είναι πολύ πιο μεταβλητός στις μεγαλύτερες ηλικίες απ' ό,τι στις νεότερες ηλικίες, λόγω του πολύ μικρότερου απόλυτου αριθμού θανάτων στις μεγαλύτερες ηλικίες.

Επειδή ο αριθμός των θανάτων είναι μια τυχαία μεταβλητή καταμέτρησης, σύμφωνα με τον Brillinger (1986), η υπόθεση Poisson φαίνεται να είναι εύλογη. Προκειμένου να παρακαμφθούν τα προβλήματα που συνδέονται με τη κλασική μέθοδο εξετάζεται:

$$D_{xt} \text{Poisson}(E_{xt} m_{xt}) \text{ με } m_{xt} = \exp(a_x + \beta_x k_t) \quad (11)$$

όπου οι παράμετροι εξακολουθούν να υπόκεινται στους περιορισμούς της (2). Η δύναμη της θνησιμότητας θεωρείται έτσι ότι έχει τη λογαριθμογραμμική μορφή  $\ln(m_{x,t}) = a_x + \beta_x k_t + e_{xt}$ . Η σημασία των παραμέτρων  $a_x$ ,  $\beta_x$  και  $k_t$  είναι ουσιαστικά η ίδια με εκείνη του κλασικού μοντέλου Lee-Carter.

### 3.2.2 Εκτίμηση μέγιστης πιθανοφάνειας

Έναντι της χρήσης της SVD για την εκτίμηση των  $a_x$ ,  $\beta_x$  και  $k_t$ , αυτές οι παράμετροι μπορούν να προσδιοριστούν μεγιστοποιώντας τη λογαριθμική πιθανοφάνεια με βάση το μοντέλο (11), η οποία δίνεται από τη σχέση:

$$L(a, \beta, k) = \sum_{x,t} \{D_{xt}(a_x + \beta_x k_t) - E_{x,t} \exp(a_x + \beta_x k_t)\} + \text{σταθερά}$$

Λόγω της παρουσίας του διγραμμικού όρου  $\beta_x k_t$ , δεν είναι δυνατή η εκτίμηση του προτεινόμενου μοντέλου με τα εμπορικά στατιστικά πακέτα που εφαρμόζουν την παλινδρόμηση Poisson. Ωστόσο, το πρόγραμμα LEM (Vermunt, 1997a,b) μπορεί να χρησιμοποιηθεί για το σκοπό αυτό.

Ο αλγόριθμος που εφαρμόζεται στη LEM για την επίλυση των εξισώσεων πιθανότητας είναι μια μονοδιάστατη ή στοιχειώδης μέθοδος Newton. Ο Goodman (1979) ήταν ο πρώτος που πρότεινε αυτή την επαναληπτική μέθοδο για την εκτίμηση λογαριθμογραμμικών μοντέλων με διγραμμικούς όρους. Στο βήμα επανάληψης  $n + 1$ , ενημερώνεται μόνο ένα σύνολο παραμέτρων καθορίζοντας τις άλλες παραμέτρους στις τρέχουσες εκτιμήσεις με το ακόλουθο σύστημα:

$$\hat{\theta}^{(v+1)} = \hat{\theta}^{(v)} - \frac{\partial L^{(v)} / \partial \theta}{\partial^2 L^{(v)} / \partial \theta^2}$$

όπου  $L^{(v)} = L(\hat{\theta}^{(v)})$ .

Σε αυτή την περίπτωση, υπάρχουν τρία σύνολα παραμέτρων, δηλαδή οι όροι  $a_x$ ,  $\beta_x$  και  $k_t$ . Το σύστημα ξεκινά με  $a_x = 0$ ,  $\beta_x = 0$  και  $k_t = 0$  (μπορούν επίσης να χρησιμοποιηθούν τυχαίες τιμές).

$$\hat{\alpha}^{(v+1)} = \hat{\alpha}_x^{(v)} - \frac{\sum_t (D_{xt} - \hat{D}_{xt}^{(v)})}{-\sum_t \hat{D}_{xt}^{(v)}}, \quad \hat{\beta}_x^{(v+1)} = \hat{\beta}_x^{(v)}, \quad \hat{k}_t^{(v+1)} = \hat{k}_t^{(v)},$$

$$\hat{k}_t^{(v+2)} = \hat{k}_t^{(v+1)} - \frac{\sum_t (D_{xt} - \hat{D}_{xt}^{(v+1)}) \hat{\beta}_x^{(v+1)}}{-\sum_t \hat{D}_{xt}^{(v+1)} (\hat{\beta}_x^{(v+1)})^2}, \quad \hat{a}_x^{(v+2)} = \hat{a}_x^{(v+1)}, \quad \hat{\beta}_x^{(v+2)} = \hat{\beta}_x^{(v+1)},$$

$$\hat{\beta}_x^{(v+3)} = \hat{\beta}_x^{(v+2)} - \frac{\sum_t (D_{xt} - \hat{D}_{xt}^{(v+2)}) \hat{k}_t^{(v+2)}}{-\sum_t \hat{D}_{xt}^{(v+2)} (\hat{k}_t^{(v+2)})^2}, \quad \hat{a}_x^{(v+3)} = \hat{a}_x^{(v+2)}, \quad \hat{k}_t^{(v+3)} = \hat{k}_t^{(v+2)}$$

όπου  $\hat{D}_{xt}^{(v)} = E_{xt} \exp(\hat{a}_x^{(v)} + \hat{\beta}_x^{(v)} \hat{k}_t^{(v)})$ , ή ο εκτιμώμενος αριθμός των θανάτων μετά το βήμα επανάληψης  $v$ . Το κριτήριο που χρησιμοποιείται για τη διακοπή της διαδικασίας είναι μια πολύ μικρή αύξηση της συνάρτησης λογαριθμικής πιθανοφάνειας (η προεπιλεγμένη τιμή της LEM είναι  $10^{-6}$ , αλλά μπορεί να συνιστάται να τεθεί το κριτήριο λίγο πιο έντονα, δηλαδή σε  $10^{-10}$ ).

Αφού ενημερωθούν οι παράμετροι  $k_t$ , επιβάλλεται περιορισμός θέσης. Η LEM χρησιμοποιεί τον περιορισμό κεντραρίσματος  $\sum_t \hat{k}_t = 0$ , ο οποίος είναι ο ίδιος περιορισμός όπως και στην παραμετροποίηση Lee-Carter. Μετά την ενημέρωση των παραμέτρων  $\beta_x$ , πρέπει να επιβληθεί ένας περιορισμός κλιμάκωσης. Ο περιορισμός κλιμάκωσης που χρησιμοποιείται από την LEM είναι  $\hat{\beta}_1 = 1$ , ο οποίος διαφέρει από την παραμετροποίηση Lee-Carter. Προκειμένου να ληφθεί η παραμετροποίηση Lee-Carter στην οποία  $\sum_x \hat{\beta}_x = 1$ , πρέπει να διαιρεθούν οι εκτιμήσεις LEM για την  $\beta_x$  με το  $\sum_x \hat{\beta}_x$  και να πολλαπλασιαστούν οι εκτιμήσεις LEM για το  $k_t$  με τον ίδιο αριθμό.

Μια άλλη επιλογή για να ληφθεί υπόψη ο περιορισμός  $\sum_t \hat{k}_t = 0$  συνίσταται στον υπολογισμό των ενημερώσεων για τα  $k_t$  χωρίς περιορισμούς και στην κεντροποίηση των ενημερώσεων πριν από την πραγματική ενημέρωση των  $k_t$ . Αυτή η απλή μέθοδος λειτουργεί επειδή

πρόκειται για περιορισμό ταυτοποίησης (και όχι με έναν περιορισμό μοντέλου). Σε αντίθεση με την κλασική προσέγγιση Lee-Carter (όπου η SVD εφαρμόζεται σε μετασχηματισμένα ποσοστά θνησιμότητας), το σφάλμα εφαρμόζεται απευθείας στον αριθμό των θανάτων στην προσέγγιση παλινδρόμησης Poisson. Συνεπώς, δεν υπάρχει ανάγκη εκτίμησης σε δεύτερο στάδιο.

### 3.3 Κριτικές παρατηρήσεις

Η μέθοδος Lee-Carter έχει πολλά πλεονεκτήματα, το σημαντικότερο εκ των οποίων είναι η απλότητα. Το υποκείμενο μοντέλο έχει αποδειχθεί ότι αντιπροσωπεύει ένα μεγάλο μέρος της μεταβλητότητας των ποσοστών θνησιμότητας στις ανεπτυγμένες χώρες (Tuljapurkar et al., 2000). Οι παράμετροι είναι εύκολα ερμηνεύσιμες και μια απλή πρόβλεψη τυχαίου περιπάτου με διολίσθηση είναι γενικά επαρκής για τη μία και μοναδική προεκτιμώμενη παράμετρο. Η μέθοδος περιλαμβάνει ελάχιστη υποκειμενική κρίση και παράγει στοχαστικές προβλέψεις με πιθανοτικά διαστήματα πρόβλεψης.

Έχουν προταθεί διάφορες βελτιώσεις της βάσης εκτίμησης Lee-Carter, κάποιες από τις οποίες περιγράφονται ακολούθως. Ο Wilmoth (1993) παρουσίασε μια λύση Lee-Carter βασισμένη στη μέθοδο των σταθμισμένων ελαχίστων τετραγώνων. Ανέπτυξε επίσης μια λύση εκτίμησης μέγιστης πιθανοφάνειας με βάση την υπόθεση της κατανομής Poisson των θανάτων. Αυτό το λογαριθμογραμμικό μοντέλο Poisson εφαρμόστηκε στη συνέχεια από άλλους (Brouhns et al., 2002; Czado et al., 2005; Koissi et al., 2006). Οι Koissi & Shapiro (2006) εφάρμοσαν μια ασαφή διατύπωση της Lee-Carter για να αντιμετωπίσουν τις παραβιάσεις της υπόθεσης για σταθερή διακύμανση σφάλματος κατά ηλικία. Ο Wolf (2004) ανέπτυξε μια παραλλαγή του Lee-Carter που εφαρμόζεται στις πρώτες διαφορές των λογαριθμών θνησιμότητας, η οποία ενσωματώνει την εκτίμηση των μοντέλων Lee-Carter και χρονοσειρών. Οι Li & Chan (2005) πρότειναν ένα μοντέλο Lee-Carter προσαρμοσμένο σε ακραίες τιμές το οποίο ενισχύει τη βάση εκτίμησης όταν η χρονοσειρά του δείκτη θνησιμότητας περιέχει ακραίες τιμές από εξωγενή γεγονότα.

Ένα μειονέκτημα της Lee-Carter είναι ότι υποθέτει ότι ο λόγος των ρυθμών μεταβολής της θνησιμότητας στις διάφορες ηλικίες παραμένει σταθερός με την πάροδο του χρόνου (δηλαδή αναλλοίωτο  $\beta_x$ ), ενώ έχουν βρεθεί στοιχεία για σημαντική αλληλεπίδραση ηλικίας-χρόνου. Οι Booth et al. (2002) αντιμετωπίζουν εν μέρει αυτό το πρόβλημα επιλέγοντας μια περίοδο προσαρμογής που ικανοποιεί βέλτιστα τις υποθέσεις του μοντέλου (Carter & Prskawetz, 2001).

Τα ποσοστά πρόβλεψης Lee-Carter στερούνται ομαλότητας μεταξύ των ηλικιών και γίνονται όλο και πιο τεθλασμένα με την πάροδο του χρόνου (Giroso & King, 2006), γεγονός που δεν είναι διαισθητικό και μπορεί να είναι προβληματικό σε πρακτικές εφαρμογές. Οι Czado et al. (2005) χρησιμοποίησαν το λογαριθμογραμμικό μοντέλο Poisson μαζί με την εκτίμηση κατά Bayes για να επιβάλουν ομαλότητα. Οι Hyndman & Ullah (2007) ανέπτυξαν μια προσέγγιση βασισμένη στο παράδειγμα των λειτουργικών δεδομένων, στο οποίο τα δεδομένα εξομαλύνονται κατά την ηλικία πριν από τη μοντελοποίηση με τη χρήση p-splines παλινδρόμησης. Η μέθοδος μοντελοποιεί επίσης όρους υψηλότερης τάξης και ενισχύει τη βάση εκτίμησης ώστε να επιτρέπει τις ακραίες τιμές.

Άλλες πρόσφατες εξελίξεις κάνουν χρήση των μοντέλων χώρου καταστάσεων (Harvey, 1989), τα οποία μπορούν να θεωρηθούν ως ένα πλαίσιο για την από κοινού μοντελοποίηση χρονοσειρών και διατομικών δεδομένων. Ένα πλεονέκτημα αυτών των ολοκληρωμένων μοντέλων είναι ότι οι στατιστικές και διαγνωστικές ιδιότητες ολόκληρης της διαδικασίας εκτίμησης και πρόβλεψης είναι γνωστές και κατανοητές σε σύγκριση με εκείνες του Lee-Carter και άλλων μεθόδων που χρησιμοποιούν μια προσέγγιση σταδιακής εκτίμησης. Οι De Jong & Tickle (2006) προσαρμόζουν το πλαίσιο του χώρου καταστάσεων για να ενσωματώσουν την ηλικιακή ομαλότητα χρησιμοποιώντας b-splines. Σε σύγκριση με την Lee-Carter, η μέθοδος χρησιμοποιεί λιγότερες παραμέτρους λόγω της χρήσης των splines για τη μείωση της διαστατικότητας, παράγει ομαλά ποσοστά πρόβλεψης και προσφέρει τα πλεονεκτήματα της ολοκληρωμένης εκτίμησης και πρόβλεψης.

Οι αξιολογήσεις δείχνουν ότι αυτές οι νεότερες μέθοδοι αποδίδουν τουλάχιστον εξίσου καλά με τις παραλλαγές Lee-Carter όσον αφορά την ακρίβεια των προβλέψεων, ενώ προσφέρουν επίσης, σε διαφορετικό βαθμό, πλεονεκτήματα όσον αφορά την οριζόντια ομαλότητα, τη λιτότητα, την ολοκληρωμένη και εύρωστη εκτίμηση και τη γενικευσιμότητα. Μια μελέτη της ακρίβειας πρόβλεψης σε είκοσι πληθυσμούς διαπίστωσε ότι οι προσεγγίσεις Hyndman-Ullah και de Jong-Tickle έδωσαν χαμηλότερα μέσα σφάλματα στην πρόβλεψη της λογαριθμοποιημένης θνησιμότητας από τις παραλλαγές Lee-Miller και Booth-Maindonald-Smith, αν και όχι σημαντικά. Και οι τέσσερις προσεγγίσεις έδωσαν σημαντικά χαμηλότερα μέσα σφάλματα στην προβλεπόμενη λογαριθμική θνησιμότητα από την αρχική μέθοδο Lee-Carter (Booth et al., 2006).

Οι De Jong & Marshall (2007) πρότειναν μια νέα μέθοδο που συνδυάζει μοντέλα χώρου καταστάσεων και τον μετασχηματισμό Wang, ο οποίος έχει χρησιμοποιηθεί ευρέως σε



τιμολόγηση ασφαλιστικών και χρηματοοικονομικών κινδύνων (Wang, 2000). Η προσέγγιση είναι ανάλογη με την Lee-Carter, αλλά κάνει χρήση ενός μετασχηματισμού που επιτρέπει την υπόθεση ότι ο δείκτης θνησιμότητας επηρεάζει εξίσου όλες τις ηλικίες (ισοδυναμεί με σταθερό  $\beta_x$  στο υπόδειγμα Lee-Carter), δίνοντας μια απλούστερη βάση για την πρόβλεψη. Ο Pedroza (2006) χρησιμοποίησε το πλαίσιο του χώρου καταστάσεων και την εκτίμηση Bayes, καθώς και την ανάπτυξη μιας επέκτασης που επιτρέπει τη διαφορεική μεταβλητότητα μεταξύ των ηλικιών.

## Κεφάλαιο 4<sup>ο</sup>: Προσεγγίσεις πρόβλεψης

Στο παρόν κεφαλαίο θα εξεταστούν ορισμένες μέθοδοι πρόβλεψης πέραν της Lee-Carter.

### 4.1 Προσδοκία

Οι προσδοκίες έχουν συχνά χρησιμοποιηθεί στην πρόβλεψη της θνησιμότητας από πλευράς εμπειρογνομώνων, καθώς προσδιορίζεται μια υποτιθέμενη πρόβλεψη ή ένα σενάριο, το οποίο συχνά συνοδεύεται από εναλλακτικά αναμενόμενα και μη σενάρια. Οι περισσότερες επίσημες στατιστικές υπηρεσίες στρέφονται προς αυτή την προσέγγιση (Waldron, 2005), αν και αρκετές υιοθετούν τα τελευταία χρόνια την εξωστρεφή προσέγγιση. Οι αναλογιστές έχουν επίσης βασιστεί σε μεγάλο βαθμό στην προσδοκία στο παρελθόν, αν και πλέον κινούνται προς πιο εξελιγμένες μεθόδους παρέκτασης (Continuous Mortality Investigation Bureau, 2002, 2004, 2005, 2006, 2007).

Παραδοσιακά, μια ευρέως χρησιμοποιούμενη μέθοδος ήταν η στόχευση του προσδόκιμου ζωής, όπου μια τιμή υποτίθεται για μια μελλοντική ημερομηνία με καθορισμένη πορεία (Pollard, 1987- Olshansky, 1988). Έχουν επίσης χρησιμοποιηθεί πιο λεπτομερείς προσδοκίες. Οι διαχειριστές κοινωνικής ασφάλισης των Ηνωμένων Πολιτειών χρησιμοποιούν διαφορετικές κρίσεις για να προσαρμόσουν τις τάσεις των ποσοστών θανάτου λόγω αιτίας και ηλικίας, ώστε να αντανakλάται μια λογική πορεία βελτίωσης της θνησιμότητας στο μέλλον (Waldron, 2005). Οι προβλέψεις του προσδόκιμου ζωής που προκύπτουν θεωρούνται ευρέως ως πολύ χαμηλές (Lee, 2000- Tuljapurkar et al., 2000- Bongaarts, 2006). Το Γραφείο Απογραφής των ΗΠΑ υποθέτει σύγκλιση σε μακροπρόθεσμους πίνακες ζωής, με τις προβλέψεις προσδόκιμου ζωής Lee-Carter να χρησιμοποιούνται ως σημεία αναφοράς και τη γνώμη των εμπειρογνομώνων να χρησιμοποιείται για τον καθορισμό της σχετικής ταχύτητας μείωσης της θνησιμότητας ανά ηλικία (Hollmann et al., 2000).

Η στόχευση εφαρμόζεται επίσης παραδοσιακά σε συντελεστές μείωσης της θνησιμότητας ανά ηλικία. Το Continuous Mortality Investigation Bureau (CMIB) του Ινστιτούτου Αναλογιστών και της Σχολής Αναλογιστών του Ηνωμένου Βασιλείου έχει προβάλει το ποσοστό των ζώνων ακριβούς ηλικίας  $x$  που πεθαίνουν πριν φτάσουν στην ακριβή ηλικία  $x + 1$ ,  $q_x$ , από το έτος μηδέν πολλαπλασιάζοντας με τον  $t$ -ετή μειωτικό παράγοντα  $RF(x, t)$ , όπου:

$$RF(x, t) = a(x) + [1 - a(x)] \cdot [1 - f_n(x)]^{t/n}$$

και  $a(x)$  και  $f_n(x)$  αντιπροσωπεύουν, αντίστοιχα, τον τελικό συντελεστή μείωσης και το ποσοστό της συνολικής μείωσης  $(1 - a(x))$  που θεωρείται ότι θα συμβεί σε  $n$  έτη. Η προσέγγιση ενσωματώνει μια εκθετική μείωση της θνησιμότητας με την πάροδο του χρόνου στην ασυμπτωτική τιμή  $a(x)$  και χρησιμοποιεί τη γνώμη των εμπειρογνομόνων για τον καθορισμό των στόχων  $a(x)$  και  $f_n(x)$ . Οι πίνακες ζωής των σειρών "80" και "92" ((Continuous Mortality Investigation Bureau, 1990, 1999) βασίζονται σε αυτή τη μέθοδο με διαφορετικές τιμές παραμέτρων. Μια παρόμοια προσέγγιση χρησιμοποιείται από την Αμερικανική Εταιρεία Αναλογιστών, αλλά χωρίς περιοριστικό συντελεστή μείωσης (που ισοδυναμεί με  $a(x) = 0$ ). Οι Wong-Furuy & Haberman (2004) παρέχουν περισσότερες λεπτομέρειες για τις εφαρμογές του Ηνωμένου Βασιλείου και των ΗΠΑ στις ασφαλίσεις ζωής και τις συντάξεις. Το Government Actuary's Department του Ηνωμένου Βασιλείου έχει χρησιμοποιήσει στοχευόμενους συντελεστές μείωσης ανά ηλικία με εκθετική σύγκλιση και μια προσέγγιση κοορτής που εφαρμόζεται σε ορισμένες γενιές (Government Actuary's Department, 2001, 2006). Οι Denton et al. (2005) χρησιμοποίησαν τη στόχευση για να ορίσουν 15 σενάρια που συνδυάζουν πέντε πίνακες στόχων ζωής με διαφορετικές μέγιστες ηλικίες θανάτου και τρία πρότυπα αλλαγής.

και  $a(x)$  και  $f_n(x)$  αντιπροσωπεύουν, αντίστοιχα, τον τελικό συντελεστή μείωσης και το ποσοστό της συνολικής μείωσης  $(1 - a(x))$  που θεωρείται ότι θα συμβεί σε  $n$  έτη. Η προσέγγιση ενσωματώνει μια εκθετική μείωση της θνησιμότητας με την πάροδο του χρόνου στην ασυμπτωτική τιμή  $a(x)$  και χρησιμοποιεί τη γνώμη των εμπειρογνομόνων για τον καθορισμό των στόχων  $a(x)$  και  $f_n(x)$ . Οι πίνακες ζωής των σειρών "80" και "92" ((Continuous Mortality Investigation Bureau, 1990, 1999) βασίζονται σε αυτή τη μέθοδο με διαφορετικές τιμές παραμέτρων. Μια παρόμοια προσέγγιση χρησιμοποιείται από την Αμερικανική Εταιρεία Αναλογιστών, αλλά χωρίς περιοριστικό συντελεστή μείωσης (που ισοδυναμεί με  $a(x) = 0$ ). Οι Wong-Furuy & Haberman (2004) παρέχουν περισσότερες λεπτομέρειες για τις εφαρμογές του Ηνωμένου Βασιλείου και των ΗΠΑ στις ασφαλίσεις ζωής και τις συντάξεις. Το Government Actuary's Department του Ηνωμένου Βασιλείου έχει χρησιμοποιήσει στοχευόμενους συντελεστές μείωσης ανά ηλικία με εκθετική σύγκλιση και μια προσέγγιση κοορτής που εφαρμόζεται σε ορισμένες γενιές (Government Actuary's Department, 2001, 2006). Οι Denton et al. (2005) χρησιμοποίησαν τη στόχευση για να ορίσουν 15 σενάρια που συνδυάζουν πέντε πίνακες στόχων ζωής με διαφορετικές μέγιστες ηλικίες θανάτου και τρία πρότυπα αλλαγής.

Το πλεονέκτημα της γνώμης των εμπειρογνομόνων είναι η ενσωμάτωση δημογραφικών, επιδημιολογικών και άλλων σχετικών γνώσεων, τουλάχιστον με ποιοτικό τρόπο. Το μειονέκτημα είναι η υποκειμενικότητά της και η πιθανότητα μεροληψίας. Η συντηρητικότητα της γνώμης των εμπειρογνομόνων όσον αφορά τη μείωση της θνησιμότητας είναι ευρέως διαδεδομένη, καθώς οι εμπειρογνώμονες δεν είναι γενικά πρόθυμοι να προβλέψουν τη μακροπρόθεσμη συνέχιση των τάσεων, σχετικά με τα όρια του προσδόκιμου ζωής. Στις Η.Π.Α., η γνώμη των εμπειρογνομόνων έχει βρεθεί ότι συστηματικά προβλέπει μικρότερες μειώσεις από ό,τι προβλέπουν τόσο οι εξωστρεφείς μέθοδοι όσο και αυτό που πραγματικά συνέβη (Alho & Spencer, 1990; Lee & Miller, 2001; Waldron, 2005). Οι προβλέψεις του Ηνωμένου Βασιλείου έχουν επίσης υποεκτιμήσει τις βελτιώσεις της θνησιμότητας (Shaw, 1994; Murphy, 1995; ; Continuous Mortality Investigation Bureau, 2004) . Στην Αυστραλία, η χρήση μειωμένων ποσοστών μείωσης της θνησιμότητας μετά τη βραχυπρόθεσμη περίοδο έχει επανειλημμένα οδηγήσει σε υποεκτίμηση του προσδόκιμου ζωής μακροπρόθεσμα.

Στην πραγματικότητα, έχει διαπιστωθεί ότι η γνώμη των εμπειρογνομόνων υπόκειται σε "υποθετική αντίσταση", όπου οι προσδοκίες τείνουν να υστερούν παρά να προηγούνται της πραγματικής εμπειρίας (Ahlburg & Vaupel, 1990). Ένας τρόπος για να μειωθεί αυτή η μεροληψία και να αξιοποιηθεί η ευρύτερη εμπειρογνομοσύνη είναι η χρήση του μέσου όρου των απόψεων πολλών εμπειρογνομόνων με τη χρήση μεθόδων πρόβλεψης Delphi (π.χ. Lutz & Scherbov, 1998). Η μεροληψία μπορεί να εξακολουθεί να υφίσταται λόγω της "συρροής εμπειρογνομόνων", η οποία προκύπτει από την κοινή βάση πληροφοριών και την επιθυμία συμμόρφωσης (Alho, 2005). Πρόσφατα δόθηκε μεγαλύτερη έμφαση στα επιχειρήματα των εμπειρογνομόνων (και όχι στη γνώμη) για να αυξηθεί η ποικιλομορφία των προσδοκιών (Lutz et al., 1999, 2004).

Οι προσδοκίες των εμπειρογνομόνων αναφέρονται σε συμπεριφορές σε επίπεδο πληθυσμού που βασίζονται σε ιστορικά δεδομένα και έρευνες και είναι κατάλληλες για μακροπρόθεσμες προβλέψεις. Για πολύ βραχυπρόθεσμες προβλέψεις, θα μπορούσαν να εξεταστούν οι ατομικές προσδοκίες, μια τέτοια προσέγγιση έχει χρησιμοποιηθεί στην πρόβλεψη της γονιμότητας, αν και με περιορισμένη επιτυχία (Booth, 2006). Οι προσδοκίες των ατόμων σχετικά με την κατάσταση της υγείας τους έχει αποδειχθεί ότι αποτελούν καλό προγνωστικό παράγοντα για τον θάνατο (Idler & Benyamini, 1997), και οι προσωπικές εκτιμήσεις για την πιθανότητα επιβίωσης μέχρι μια ορισμένη ηλικία έχει αποδειχθεί ότι συμφωνούν με τις εκτιμήσεις του πίνακα ζωής (Hauser & Willis, 2005). Αν και η προσέγγιση αυτή είναι σαφώς περιορισμένη, η

ατομική ετερογένεια των προσδοκιών μπορεί να παρέχει χρήσιμες πληροφορίες για τη μοντελοποίηση και την πιθανολογική πρόβλεψη.

## 4.2 Παρέκταση

Όπως έχει ήδη σημειωθεί, η παρέκταση αποτελεί τη βάση των περισσότερων μεθόδων πρόβλεψης της θνησιμότητας. Οι μέθοδοι παρέκτασης υποθέτουν ότι οι μελλοντικές τάσεις θα είναι ουσιαστικά η συνέχεια του παρελθόντος. Στην πρόβλεψη της θνησιμότητας, αυτή είναι συνήθως μια λογική υπόθεση λόγω των ιστορικών κανονικοτήτων. Ωστόσο, υπάρχουν εξαιρέσεις, παραδείγματα αποτελούν οι προσωρινές αυξήσεις στη θνησιμότητα των νεαρών ενηλίκων ανδρών λόγω των θανάτων από AIDS, οι πρόσφατες διακυμάνσεις της ρωσικής θνησιμότητας και η περίοδος στασιμότητας της θνησιμότητας στην Αυστραλία τη δεκαετία του 1960. Αυτή η συζήτηση για τις μεθόδους παρέκτασης προχωρά από την απλή γραμμική παρέκταση σε μεθόδους που βασίζονται σε μοντέλα δύο παραγόντων (ηλικία-περίοδος).

Μεταξύ των απλούστερων μεθόδων είναι η γραμμική συνέχιση της τάσης του προσδόκιμου ζωής. Σύμφωνα με τον White (2002), οι ιστορικές τάσεις υποδηλώνουν ότι μια τέτοια μέθοδος αναμένεται να είναι ακριβής. Παρά το γεγονός αυτό, η συνεχιζόμενη γραμμική τάση σπάνια έχει προβλεφθεί πέραν του βραχυπρόθεσμου μέλλοντος. Η ανεξάρτητη γραμμική παρέκταση των λογαρίθμων των ειδικών κατά ηλικία ποσοστών, που ισοδυναμεί με την παραδοχή σταθερών συντελεστών μείωσης της θνησιμότητας κατά ηλικία, έχει συχνά εφαρμοστεί (Pollard, 1987). Οι συντελεστές μείωσης ή τα κλάσματα των συντελεστών μείωσης μπορούν να προεκτεθούν άμεσα (Goss et al., 1998). Οι Denton et al. (2005) χρησιμοποίησαν λογαριθμογραμμική προεκβολή. Οι Golularati et al. (1984) προέβλεψαν τη θνησιμότητα της αυστραλιανής κοορτής με γραμμική παρέκταση των μετασχηματισμών Brass logit (Pollard, 1987). Η μη γραμμική παρέκταση μπορεί να χρησιμοποιηθεί με την παραδοχή μιας κατάλληλης λειτουργικής μορφής, όπως η λογιστική.

Οι μέθοδοι χρονολογικών σειρών (Box et al., 1994) χρησιμοποιούνται συνήθως στις εξωστρεφείς προβλέψεις. Τα μοντέλα χρονοσειρών έχουν το πλεονέκτημα ότι είναι στοχαστικά, επιτρέποντας τον υπολογισμό του πιθανολογικού διαστήματος πρόβλεψης για την τιμή πρόβλεψης. Στην περίπτωση ενός υποκείμενου μοντέλου μηδενικού παράγοντα, όπως μια χρονοσειρά προσδόκιμου ζωής ή ενός ποσοστού ανά ηλικία, μοντελοποιείται και προβλέπεται η ίδια η σειρά. Για μοντέλα ενός και δύο παραγόντων, οι χρονοσειρές είναι οι προσαρμοσμένες παράμετροι του υποκείμενου μοντέλου. Για παράδειγμα οι οκτώ παράμετροι του μοντέλου Heligman-Pollard ή η μία χρονική παράμετρος του μοντέλου Lee-Carter. Στην πρόβλεψη της

θνησιμότητας χρησιμοποιείται συνήθως η μονόπλευρη μοντελοποίηση ARIMA. Για πολλαπλές παραμέτρους, η ταυτόχρονη πρόβλεψη των παραμέτρων λαμβάνοντας υπόψη τις αλληλεξαρτήσεις είναι το η ιδανική συνθήκη, που επιτυγχάνεται με την πολυμεταβλητή μοντελοποίηση ARIMA. Αν και αναπτύχθηκαν για βραχυπρόθεσμες προβλέψεις, οι μέθοδοι χρονοσειρών χρησιμοποιούνται όλο και περισσότερο για μακροπρόθεσμες προβλέψεις θνησιμότητας. Οι Alho & Spencer (2005) συνιστούν ότι οι σειρές δεδομένων πρέπει να είναι πάντοτε μεγαλύτερες από τον ορίζοντα πρόβλεψης και κατά προτίμηση δύο έως τρεις φορές μεγαλύτερες.

Οι προβλέψεις χρονολογικών σειρών που βασίζονται σε υποκείμενα υποδείγματα μηδενικού παράγοντα είναι σχετικά σπάνιες. Οι Denton et al. (2005) προβλέπουν τη συνολική θνησιμότητα με αυτόν τον τρόπο. Ο Bell (1997) εφάρμοσε ένα απλό μοντέλο τυχαίου περιπάτου με διολίσθηση χωριστά στους λογαρίθμους κάθε ποσοστού ανά ηλικία και διαπίστωσε ότι η ακρίβεια των προβλέψεων ήταν τουλάχιστον εξίσου καλή με τις πιο σύνθετες μεθόδους, όπως η Lee-Carter. Ένα σημαντικό μειονέκτημα μιας τέτοιας προσέγγισης είναι ότι η ανεξάρτητη πρόβλεψη της θνησιμότητας σε κάθε ηλικία μπορεί να οδηγήσει σε απίθανα ηλικιακά πρότυπα, ιδίως μακροπρόθεσμα. Για να αντιμετωπιστεί αυτό το πρόβλημα, οι Denton κ.ά. (2005) προέβλεψαν τις πρώτες διαφορές των ειδικών για κάθε ηλικία λογαρίθμων θανάτου ως ένα σύστημα διασταυρούμενων μοντέλων χρονοσειρών. Η εκτίμηση επιτεύχθηκε με τη χρήση μιας επαναληπτικής μορφής του εκτιμητή της φαινομενικά ασύνδετης παλινδρόμησης (SURE).

Μεταξύ των μονοπαραγοντικών μοντέλων που χρησιμοποιούνται στην πρόβλεψη της θνησιμότητας είναι οι συναρτήσεις παραμετροποίησης και το σχεσιακό μοντέλο Brass (Brass, 1971, 1974), το οποίο αναπτύχθηκε για σκοπούς εκτίμησης σε περιβάλλοντα με φτωχά δεδομένα, αλλά μπορεί να χρησιμοποιηθεί ευρύτερα. Το σχεσιακό μοντέλο δύο παραμέτρων εκφράζει τον μετασχηματισμό logit της παρατηρούμενης συνάρτησης επιβίωσης ως γραμμική συνάρτηση του logit ενός εμπειρικού πρότυπου προγράμματος. Αποτελεί επομένως μια ειδική περίπτωση της διαβάθμισης με αναφορά σε έναν πρότυπο πίνακα (Benjamin & Pollard, 1980). Η προσαρμογή σε ετήσια δεδομένα δίνει χρονοσειρές των δύο παραμέτρων που μπορούν να προβλεφθούν με τη χρήση μεθόδων χρονοσειρών. Ο Keyfitz (1991) εφάρμοσε τη μέθοδο σε καναδικά δεδομένα με περιορισμένη επιτυχία (βλέπε επίσης Pollard, 1987). Στις επεκτάσεις του μοντέλου περιλαμβάνονται οι Zaba (1979) και Ewbank κ.ά. (1983), οι οποίοι ανέπτυξαν ξεχωριστά μοντέλα τεσσάρων παραμέτρων για να αυξήσουν την ευελιξία στις νεότερες και τις παλαιότερες ηλικίες- βλέπε επίσης Zaba & Paes (1995). Ο Congdon (1993) προσάρμοσε το

μοντέλο Zaba στη θνησιμότητα για την ευρύτερη περιοχή του Λονδίνου κατά την περίοδο 1971-1990, λαμβάνοντας σχετικά σαφείς τάσεις για τρεις από τις τέσσερις παραμέτρους, και προέβλεψε όλες τις παραμέτρους χρησιμοποιώντας μονομεταβλητά μοντέλα ARIMA.

Ένα μοντέλο δύο παραγόντων που χρησιμοποιείται στην πρόβλεψη της θνησιμότητας είναι το μοντέλο Lee-Carter (Lee & Carter, 1992). Αυτός ο τύπος μοντέλου μπορεί να εκτιμηθεί με τη χρήση κύριων συνιστωσών, με την οποία η αποσύνθεση πινάκων προσδιορίζει ανεξάρτητες συνιστώσες θνησιμότητας ή ηλικιακά πρότυπα και τη σημασία τους με την πάροδο του χρόνου. Στην πρόβλεψη, οι εκτιμώμενες παράμετροι ηλικίας θεωρούνται σταθερές και χρησιμοποιούνται μέθοδοι χρονοσειρών για την παρέκταση των χρονικά μεταβαλλόμενων παραμέτρων. Η μέθοδος Lee-Carter και οι συναφείς προσεγγίσεις εξετάζονται λεπτομερώς στο υποκεφάλαιο 4.5. Σε μια παρόμοια προσέγγιση, το μοντέλο δεδομένων πάνελ των Babel et al. (2008) εκφράζει τον λογάριθμο του συντελεστή μείωσης της θνησιμότητας ως το άθροισμα στοχαστικά ανεξάρτητων επιδράσεων της ηλικίας και του χρόνου- οι μέθοδοι χρονοσειρών χρησιμοποιούνται για τη μοντελοποίηση της επίδρασης του χρόνου, ενώ η επίδραση της ηλικίας θεωρείται σταθερή.

Ένα εναλλακτικό μοντέλο δύο παραγόντων είναι το διμεταβλητό μοντέλο παλινδρόμησης (ηλικία και χρόνος) στο πλαίσιο του GLM. Η ουσιαστική διαφορά μεταξύ των μοντέλων GLM και Lee-Carter είναι ότι στο GLM ο χρόνος είναι μια ρητή συνδιακύμανση, ενώ στο Lee-Carter ο χρόνος είναι ένας παράγοντας που αντιπροσωπεύεται από τη σειρά. Η πρόβλεψη στην προσέγγιση GLM επιτυγχάνεται απλώς με τη χρήση κατάλληλων χρονικών τιμών. Η προσέγγιση GLM συζητείται στο υποκεφάλαιο 4.6.

Οι δυναμικές συναρτήσεις παραμετροποίησης αποτελούν μια ειδική περίπτωση δύο παραγόντων στο πλαίσιο της GLM. Τέτοια μοντέλα υιοθετούν μια αρχική στατική βασική συνάρτηση παραμετροποίησης του ηλικιακού προτύπου (βλ. υποκεφάλαιο 4.4), μειώνοντας έτσι τον αριθμό των παραμέτρων που πρέπει να εκτιμηθούν, και αντικαθιστούν κάθε παράμετρο με μια συνάρτηση του χρόνου. Για παράδειγμα, στο δυναμικό μοντέλο παραμετροποίησης των Tabeau et al. (2001b), οι παράμετροι του τροποποιημένου μοντέλου Heligman-Pollard είναι γραμμικές ή μη γραμμικές συναρτήσεις του χρόνου. Οι ακριβείς συναρτήσεις τάσης καθορίζονται κατά την προσαρμογή του μοντέλου. Σε ορισμένες περιπτώσεις, παρόμοια μοντέλα συνολικής θνησιμότητας προσαρμόζονται σχεδόν εξίσου καλά, αλλά οι προβλέψεις τους διαφέρουν σημαντικά- πολλά εξαρτώνται από τις παραδοχές

που γίνονται κατά την παρέκταση των συναρτήσεων τάσης, ιδίως σε ιδιαίτερα μη γραμμικές συναρτήσεις παραμετροποίησης (Tabeau et al., 2001).

### 4.3 Επεξήγηση

Οι επεξηγηματικές μέθοδοι πρόβλεψης της θνησιμότητας βασίζονται σε δομικά ή αιτιώδη επιδημιολογικά μοντέλα ορισμένων αιτιών θανάτου που περιλαμβάνουν διαδικασίες νόσου και γνωστούς παράγοντες κινδύνου. Με τον τρόπο αυτό γίνεται χρήση πολύτιμων ιατρικών γνώσεων και πληροφοριών σχετικά με τη συμπεριφορά και τις περιβαλλοντικές αλλαγές (Guterman & Vanderhoof, 1998). Το κύριο πλεονέκτημα της επεξηγηματικής προσέγγισης είναι ότι μπορούν να ληφθούν υπόψη οι μηχανισμοί ανατροφοδότησης και οι περιοριστικοί παράγοντες.

Στην πράξη, η επεξηγηματική προσέγγιση των προβλέψεων δεν έχει ακόμη αναπτυχθεί πλήρως. Οι σχέσεις μεταξύ των παραγόντων κινδύνου και της θνησιμότητας είναι ατελώς κατανοητές, γεγονός που καθιστά τη χρήση τους στην πρόβλεψη λιγότερο από αξιόπιστη. Η κύρια χρήση των υφιστάμενων επιδημιολογικών μοντέλων είναι η προσομοίωση της επίδρασης στη νοσηρότητα και τη θνησιμότητα των αλλαγών πολιτικής που επηρεάζουν τους παράγοντες κινδύνου και όχι η πρόβλεψη καθεαυτή.

Πολλά από τα μοντέλα που χρησιμοποιούνται στις επεξηγηματικές προβλέψεις βασίζονται στην παλινδρόμηση και, επομένως, εντάσσονται στο πλαίσιο του GLM.<sup>4</sup> Διαφέρουν από τα μοντέλα παρέκτασης με βάση την παλινδρόμηση στο ότι ενσωματώνουν επεξηγηματικές μεταβλητές ή παράγοντες κινδύνου, οι οποίες είτε έχουν χρονική υστέρηση είτε είναι προβλεπόμενες. Όταν τα δομικά μοντέλα βασίζονται αποκλειστικά σε εξωγενείς παράγοντες κινδύνου με υστέρηση, οι ορίζοντες πρόβλεψης περιορίζονται στον μικρότερο χρόνο υστέρησης. Ενώ η παρέκταση των παραγόντων κινδύνου επεκτείνει τον ορίζοντα πρόβλεψης, η πρόβλεψή τους μπορεί να μην είναι λιγότερο δύσκολη από την απευθείας πρόβλεψη της θνησιμότητας. Έτσι, οι επεξηγηματικές μέθοδοι περιορίζονται γενικά σε βραχυπρόθεσμες προβλέψεις.

---

<sup>4</sup> Τα δυναμικά μοντέλα πολλαπλών καταστάσεων ή μακροπροσομοίωσης εμπίπτουν επίσης στην επεξηγηματική προσέγγιση. Αυτά λαμβάνουν υπόψη τη δομή του πληθυσμού για την πρόβλεψη του αριθμού των συμβάντων και μπορούν να χρησιμοποιούν ως δεδομένα εισόδου είτε στατικούς είτε προβλεπόμενους ρυθμούς, οι οποίοι λαμβάνονται από ξεχωριστό μοντέλο. Για παραδείγματα, βλέπε Van Den Berg Jeths et al. (2001) και Van Genugten et al. (2001).



Σε ορισμένες περιπτώσεις, ένα επεξηγηματικό μοντέλο χρησιμοποιείται σε συνδυασμό με την προσδοκία κατά τον προσδιορισμό των μελλοντικών σεναρίων, για παράδειγμα, η χρήση υποτιθέμενων μελλοντικών προτύπων καπνίσματος σε ένα μοντέλο καρκίνου του πνεύμονα (Alderson & Ashwood, 1985). Τα επεξηγηματικά μοντέλα που αναπτύχθηκαν στο πλαίσιο του WHO Global Burden of Disease Study έχουν επίσης χρησιμοποιηθεί στην πρόβλεψη σεναρίων (Murray & Lopez, 1997).

Δομικά μοντέλα που περιλαμβάνουν πολυμεταβλητή αυτοπαλινδρόμηση έχουν χρησιμοποιηθεί για να συσχετίσουν τον τρόπο ζωής και άλλους παράγοντες κινδύνου με τη λειτουργική κατάσταση και τη θνησιμότητα σε μεγαλύτερες ηλικίες (Manton et al., 1991; Manton et al., 1992). Η προσέγγιση αυτή πέτυχε κάποια βελτίωση βραχυπρόθεσμα σε σχέση με τις μεθόδους που περιλάμβαναν παραδοσιακές χρονοσειρές και τη γνώμη των εμπειρογνομόνων. Μακροπρόθεσμα, ωστόσο, οι δυνατότητες πρόβλεψης περιορίζονται από τις σύντομες χρονοσειρές των παραγόντων κινδύνου, τον μεγάλο αριθμό παραμέτρων και την παρουσία μη γραμμικών αλληλεπιδράσεων.

Οι Girosi & King (2006) επεκτείνουν τη χρήση δομικών μοντέλων στην πρόβλεψη της θνησιμότητας στις αναπτυσσόμενες χώρες, χρησιμοποιώντας μεθόδους Bayes. Χρησιμοποιούνται γραμμικά μοντέλα παλινδρόμησης για την ηλικία και την περίοδο για να συσχετιστεί η θνησιμότητα με την αιτία θανάτου με γνωστές συνδιακυμάνσεις. Οι συγκυριακές μεταβλητές μπορεί να διαφέρουν ανάλογα με την ηλικία και μπορεί να είναι καθυστερημένες, παρατηρούμενες, εξωγενείς ή προβλέψεις αξιόπιστων μεγαλύτερων χρονοσειρών. Οι γνωστές ομοιότητες μεταξύ των (υπο)πληθυσμών και η ομαλότητα τόσο κατά την ηλικία όσο και κατά το χρόνο ενσωματώνονται ως προτεραιότητες, με βάση εμπειρικά στοιχεία ή τεκμηριωμένη κρίση.

#### 4.4 Λειτουργίες παραμετροποίησης

Οι συναρτήσεις παραμετροποίησης, γνωστές και ως νόμοι θνησιμότητας, είναι μονοπαραγοντικά μοντέλα που επιδιώκουν να αναπαραστήσουν με φειδώ το ηλικιακό πρότυπο θνησιμότητας. Το μοντέλο Heligman-Pollard είναι ένα πολύ γνωστό παράδειγμα. Αν και οι συναρτήσεις παραμετροποίησης έχουν εφαρμοστεί ευρέως στη διαβάθμιση της θνησιμότητας (Forfar et al., 1988), η χρήση τους στην πρόβλεψη της θνησιμότητας είναι πιο περιορισμένη. Ένα πλεονέκτημα των συναρτήσεων παραμετροποίησης για την πρόβλεψη είναι η ομαλότητα των ποσοστών πρόβλεψης σε όλες τις ηλικίες. Ο Bell (1997) σημείωσε την εξοικείωση και την

ερμηνευσιμότητα ως πρόσθετα πλεονεκτήματα αυτής της προσέγγισης. Στις περισσότερες εφαρμογές, τα μοντέλα αυτά αντιμετωπίστηκαν ως ντετερμινιστικά.

Η πρόβλεψη με συναρτήσεις παραμετροποίησης περιλαμβάνει επαναλαμβανόμενη προσαρμογή σε ετήσια δεδομένα, με αποτέλεσμα να λαμβάνεται μια χρονοσειρά κάθε παραμέτρου του μοντέλου. Αυτές οι σειρές παραμέτρων προβλέπονται με μεθόδους χρονοσειρών. Ιδανικά, χρησιμοποιούνται πολυμεταβλητές μέθοδοι χρονοσειρών για να ληφθούν υπόψη τυχόν αλληλεξαρτήσεις μεταξύ των παραμέτρων. Στην πράξη, οι ακανόνιστες χρονοσειρές παραμέτρων και οι αλληλεξαρτήσεις (υπερ-παραμετροποίηση) παρουσιάζουν σοβαρές δυσκολίες για την πρόβλεψη και περιορίζουν τους ορίζοντες πρόβλεψης: τα πρότυπα ηλικίας πρόβλεψης μπορεί να είναι απίθανα και, εάν δεν ληφθούν υπόψη οι αλληλεξαρτήσεις και τα διαστήματα πρόβλεψης θα είναι ανακριβή. Το μέγεθος του σφάλματος προσέγγισης αποτελεί επίσης μειονέκτημα (Bell, 1997).

#### 4.4.1 Μοντέλο Heligman-Pollard

Μεταξύ των πολλών συναρτήσεων παραμετροποίησης που έχουν προταθεί, το μοντέλο Heligman-Pollard (Heligman & Pollard, 1980) είναι από τα πιο γνωστά. Το μοντέλο οκτώ παραμέτρων έχει τρεις όρους που αποτυπώνουν το ηλικιακό πρότυπο θνησιμότητας στην παιδική ηλικία, τη νεαρή ενηλικίωση (η καμπούρα του ατυχήματος) και τη γήρανση, αντίστοιχα:

$$\frac{q_x}{p_x} = A^{(x+B)^C} + D \exp[-E(\ln x - \ln F)^2] + GH^x$$

όπου  $p_x = 1 - q_x$ , και οι A έως H είναι οι οκτώ παράμετροι. Κάθε όρος παίρνει θετικές τιμές μόνο στις σχετικές ηλικίες και ολόκληρη η συνάρτηση εκτιμάται σε ένα βήμα. Το μοντέλο έχει βρεθεί ότι ταιριάζει καλά σε μια σειρά πληθυσμών (Mode & Busby, 1982; Hartmann, 1987; Kostaki, 1988; Rogers & Gard, 1991).

Ωστόσο, μόνο οι Forfar & Smith (1987) βρήκαν το υπόδειγμα Heligman-Pollard χρήσιμο για την πρόβλεψη. Ένας λόγος γι' αυτό μπορεί να βρίσκεται στις υψηλές συσχετίσεις μεταξύ των παραμέτρων, ιδίως για την ανδρική θνησιμότητα (Hartmann, 1987). Οι συσχετίσεις αυτές θέτουν επίσης σε κίνδυνο τη διαπροσωπεία των παραμέτρων. Δυσκολίες έχουν επίσης παρουσιαστεί στον προσδιορισμό της καλύτερης περιόδου βάσης για την πρόβλεψη των παραμέτρων όταν χρησιμοποιούνται απλές μέθοδοι (Pollard, 1987; Keyfitz, 1991). Χρησιμοποιώντας μονομεταβλητά μοντέλα χρονοσειρών για τις οκτώ σειρές παραμέτρων που προέκυψαν από την προσαρμογή σε δεδομένα των ΗΠΑ για την περίοδο 1900 έως 1975, οι

McNown & Rogers (1989) προέβλεψαν τη θνησιμότητα για την περίοδο 1976 έως 1985 με "λογική ακρίβεια" (McNown & Rogers, 1992).

#### 4.4.2 Πολυεκθετικό μοντέλο

Το πολυεκθετικό μοντέλο αναπτύχθηκε για πρώτη φορά για τη μοντελοποίηση της μετανάστευσης. Οι Rogers & Planck (1983) εφάρμοσαν την εκδοχή τεσσάρων όρων και εννέα παραμέτρων στη θνησιμότητα. Το υπόδειγμα περιλαμβάνει μια σταθερή, εκθετικά φθίνουσα παιδική θνησιμότητα, μια διπλή εκθετική καμπούρα ατυχήματος<sup>5</sup> (η συνάρτηση Coale & McNeil (1972)) και μια γεροντική θνησιμότητα Gompertzian:

$$mx = a_0 + a_1 \exp(-a_1 x) + a_2 \exp\{-a_2 (x - \mu) - \exp[-\lambda(x - \mu)]\} + a_3 \exp(-a_3 x)$$

όπου:  $mx$  είναι το κεντρικό ποσοστό θνησιμότητας στην ηλικία  $x$ , οι  $a_0$ ,  $a_1$ ,  $a_2$  και  $a_3$  είναι παράμετροι κλίμακας ή επιπέδου. Οι  $a_1$ ,  $a_2$ ,  $a_3$ ,  $\mu$  και  $\lambda$  ορίζουν το προφίλ θνησιμότητας.

Οι Rogers & Little (1994) βρήκαν ελαφρώς φτωχότερη προσαρμογή από το μοντέλο Heligman-Pollard. Όπως και το Heligman-Pollard, το πολυεκθετικό μοντέλο έχει περιορισμένη χρήση στην πρόβλεψη. Οι McNown, Rogers & Little (1995) διαπίστωσαν αλληλεξαρτήσεις μεταξύ των παραμέτρων στην πρόβλεψη. Πέντε παράμετροι ήταν σταθερές και τέσσερις μοντελοποιήθηκαν από τρεις μονομεταβλητές διαδικασίες ARIMA και ένα διμεταβλητό αυτοπαλίνδρομο μοντέλο. Οι McNown & Rogers (1992) προέβλεψαν συνολικά πέντε θνησιμότητες ανά αιτία καθορίζοντας έξι παραμέτρους και μοντελοποιώντας μόνο τις παραμέτρους επιπέδου,  $a_1$ ,  $a_2$  και  $a_3$ , με μονομεταβλητά μοντέλα ARIMA, δίνοντας αρκετά ακριβείς προβλέψεις για την επόμενη δεκαετία.

#### 4.4.3 Άλλα μοντέλα

Διάφορες παραμετροποιήσεις έχουν αποδειχθεί ότι παρέχουν καλύτερη προσαρμογή στην ηλικιακή κατανομή της θνησιμότητας από ό,τι τα μοντέλα Heligman-Pollard ή τα πολυεκθετικά μοντέλα, αλλά, εξ όσων γνωρίζουμε, δεν έχουν χρησιμοποιηθεί σε προβλέψεις.

---

<sup>5</sup> Το φαινόμενο των αυξημένων ποσοστών θνησιμότητας για τα άτομα στα τέλη της εφηβείας και στις αρχές της τρίτης δεκαετίας σε έναν δεδομένο πληθυσμό.

Τέτοια μοντέλα περιλαμβάνονται εδώ, καθώς η υιοθέτησή τους θα μπορούσε να οδηγήσει σε αύξηση της ακρίβειας των προβλέψεων.

Ο Carriere (1992) πρότεινε ένα μοντέλο τριών όρων με οκτώ ερμηνεύσιμες παραμέτρους:

$$s_x = \psi_1 \exp\left\{-\left(\frac{x}{m_1}\right)^{\frac{m_1}{\sigma_1}}\right\} + \psi_2 \left[1 - \exp\left\{-\left(\frac{x}{m_2}\right)^{\frac{m_2}{\sigma_2}}\right\}\right] - \psi_3 \exp\left\{e^{-\frac{m_3}{\sigma_3}} - e^{-\frac{x-m_3}{\sigma_3}}\right\}$$

όπου:  $s_x$  είναι η πιθανότητα επιβίωσης μέχρι την ακριβή ηλικία  $x$ ,  $x > 0$ ,  $\psi_1$  είναι η πιθανότητα θανάτου από "παιδικά αίτια",  $\psi_2$  είναι η πιθανότητα θανάτου από "εφηβικά αίτια",  $\psi_3 = 1 - \psi_1 - \psi_2$  είναι η πιθανότητα θανάτου από "ενήλικα αίτια" και  $m_1, m_2, m_3, \sigma_1, \sigma_2$  και  $\sigma_3$  είναι οι παράμετροι θέσης και διασποράς.

Αυτό το μοντέλο προσαρμόστηκε καλύτερα στη θνησιμότητα στις ΗΠΑ από ό,τι το μοντέλο Heligman-Pollard. Αν και παρόμοιο με το μοντέλο Heligman-Pollard, το μοντέλο Carriere έχει αρκετά σημαντικά πλεονεκτήματα. Οι παράμετροι έχουν σαφέστερη ερμηνεία και οι παράμετροι της θέσης και της κλίμακας παρέχουν διορατικές πληροφορίες. Η χρήση ενός μείγματος συναρτήσεων επιβίωσης είναι ισοδύναμη με έναν πίνακα ζωής πολλαπλών μειώσεων και το γενικευμένο πλαίσιο μπορεί να φιλοξενήσει επιπλέον όρους και άλλες συναρτήσεις. Ο Carriere (1992) διαπίστωσε ότι για τη θνησιμότητα των γυναικών στις ΗΠΑ ένα μοντέλο τεσσάρων όρων με έντεκα παραμέτρους έδωσε σημαντικές βελτιώσεις σε σχέση με ένα μειωμένο μοντέλο οκτώ παραμέτρων. Η ευελιξία αυτού του μοντέλου έχει πιθανά πλεονεκτήματα στην πρόβλεψη.

Ο Siler (1983) πρότεινε ένα μοντέλο πέντε παραμέτρων το οποίο βρέθηκε να ταιριάζει καλύτερα από τα μοντέλα Heligman-Pollard ή Mode-Busby οκτώ παραμέτρων (Gage & Mode, 1993). Το μοντέλο Siler έχει τρεις ανεξάρτητους όρους που περιγράφουν τη θνησιμότητα κατά την νεανική ηλικία, την ενηλικίωση και τη γήρανση. Οι πιθανότητες επιβίωσης είναι πολλαπλασιαστικές:

$$s_x = \exp\left\{\left(-\frac{a_1}{b_1}\right)[1 - \exp(-b_1x)]\right\} \cdot \exp(-a_2x) \cdot \exp\left(\frac{a_3}{b_3}\right)[1 - \exp(b_3x)]$$

όπου  $s_x$  είναι η πιθανότητα επιβίωσης στην ακριβή ηλικία  $x$ ,  $x > 0$ , και  $a_1, a_2, a_3, b_1$  και  $b_3$  είναι παράμετροι θέσης και διασποράς. Ο Siler προσάρμοσε με επιτυχία το μοντέλο σε ιστορικά δεδομένα σε μεγάλο βαθμό για την περίοδο πριν από την εμφάνιση της «καμπούρας του ατυχήματος» και με ευρέως μεταβαλλόμενες διάρκειες ζωής.

Ο Hannerz (1999; 2001a) πρότεινε επίσης ένα μοντέλο πέντε παραμέτρων το οποίο ταίριαζε καλύτερα στην πρόσφατη θνησιμότητα των γυναικών στη Σουηδία από ό,τι τα μοντέλα που περιλάμβαναν έναν όρο για την «καμπούρα του ατυχήματος». Το γυναικείο μοντέλο είναι το εξής:

$$1 - s_x = \exp(G_x)$$

όπου  $s_x$  είναι η πιθανότητα επιβίωσης μέχρι την ακριβή ηλικία  $x$ ,  $x > 0$ , και

$$G_x = \alpha_0 - \frac{\alpha_1}{x} + \frac{\alpha_2}{x} x^2 + \frac{\alpha_3}{c} [\exp(cx)]$$

όπου  $G_x$  είναι μια εμπειρικά παραγόμενη αναπαράσταση του λογαριθμικού μετασχηματισμού του  $(1 - s_x)$ ,  $\alpha_0$  είναι μια παράμετρος επιπέδου και  $\alpha_1$ ,  $\alpha_2$  και  $\alpha_3$  είναι παράμετροι προφίλ. Για τους άνδρες, η επιμονή της «καμπούρας ατυχήματος» επέβαλε έναν πρόσθετο όρο που περιλαμβάνει τρεις παραμέτρους για να συλλάβει την υπερβολική θνησιμότητα στις νεαρές ηλικίες ενηλίκων (Hannerz, 2001b). Το ανδρικό μοντέλο είναι:

$$1 - s_x = a \frac{\exp(G_x^1)}{1 + \exp(G_x^1)} + (1 - a) \frac{\exp(G_x^2)}{1 + \exp(G_x^2)}$$

όπου  $a$  είναι μια παράμετρος κλίμακας και

$$G_x^1 = a_0 - \frac{a_1}{x} + \frac{a_2}{2} x^2 + \frac{a_3}{c} [\exp(cx)]$$

$$G_x^2 = a_4 - \frac{a_5}{x} + \frac{a_2}{2} x^2 + \frac{a_3}{c} [\exp(cx)]$$

όπου  $a_4$  είναι μια παράμετρος επιπέδου και  $a_5$  είναι μια παράμετρος προφίλ.

Ο Hannerz (2001c) συνδύασε τα χαρακτηριστικά των σχεσιακών μοντέλων με αυτές τις συναρτήσεις παραμετροποίησης και τους πίνακες ζωής των μοντέλων σε ένα μοντέλο παλινδρόμησης. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι δεν υπάρχει ανάγκη επιλογής μεταξύ αυτών των μεθόδων, αντίθετα γίνεται χρήση των χαρακτηριστικών και των τριών.

#### 4.5 Μοντέλα Γενικής Γραμμικής Μοντελοποίησης

Ο Renshaw (1991) παρατήρησε ότι πολλές από τις συναρτήσεις παραμετροποίησης ενός παράγοντα αποτελούν ειδικές περιπτώσεις της Γενικής Γραμμικής Μοντελοποίησης (GLM)<sup>6</sup>. Η τοποθέτηση τέτοιων μοντέλων στο πλαίσιο της GLM παρέχει ένα ενοποιημένο πλαίσιο ανάλυσης και διευκολύνει τις γενικεύσεις, όπως την επέκταση των μοντέλων ώστε να περιλαμβάνουν περίοδο (χρόνο) για πρόβλεψη. Το δυναμικό μοντέλο παραμετροποίησης (που εξετάζεται στην ενότητα 3.2) είναι ένα παράδειγμα.

Οι Renshaw et al. (1996) πρότειναν ένα μοντέλο δύο παραγόντων που βασίζεται σε μια υπερδιασκορπισμένη κατανομή Poisson των θανάτων και περιλαμβάνει έναν όρο διαβάθμισης Gompertz-Makeham σε συνδυασμό με έναν πολλαπλασιαστικό όρο προσαρμογής της τάσης ανά ηλικία:

$$\log \mu_{x,t} = \beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r a_i t'^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x') t'^i$$

όπου:  $\mu_{x,t}$  είναι η ένταση της θνησιμότητας στην ηλικία  $x$  κατά το έτος  $t$ ,  $a_t$ ,  $\beta_j$  και  $\gamma_{ij}$  είναι άγνωστες παράμετροι, ο όρος  $L_j(x')$  δηλώνει πολυώνυμο Legendre βαθμού  $j$  και  $x'$  και  $t'$  είναι  $x$  και  $t$  αναπροσαρμοσμένα στο διάστημα  $[-1, 1]$ . Οι Renshaw et al. (1996) διαπίστωσαν ότι το μοντέλο παρείχε επαρκή προσαρμογή στη θνησιμότητα για ασφαλισμένους άνδρες του Ηνωμένου Βασιλείου ηλικίας 22 έως 89 ετών και διάρκειας πέντε ετών και άνω. Οι Sithole et al. (2000) χρησιμοποίησαν την ίδια δομή μοντέλου για να προβλέψουν τη θνησιμότητα των συνταξιούχων του Ηνωμένου Βασιλείου με σημαντική επιτυχία- οι βέλτιστα προσαρμοσμένες παράμετροι του μοντέλου δεν παρήγαγαν απαραίτητα αληθοφανείς προβλέψεις, γι' αυτό απαιτήθηκε η χρήση πολυώνυμων χαμηλότερης τάξης.

Άλλοι ερευνητές έχουν επίσης διαπιστώσει ότι απαιτούνται πολυώνυμα δευτέρου ή υψηλότερου βαθμού στο χρόνο για να επιτευχθεί ικανοποιητική προσαρμογή του μοντέλου, αλλά οδηγούν σε μη αξιόπιστες προβλέψεις. Οι Heathcote & Higgins (2001a, 2001b) μοντελοποίησαν τις λογαριθμικές πιθανότητες θνησιμότητας παραλληλογράμμου-Lexis στις ηλικίες 40+ με μια παλινδρόμηση δύο παραγόντων που περιλάμβανε το χρόνο, το τετράγωνο της ηλικίας, μια αλληλεπίδραση ηλικίας-χρόνου και έως 24 πρόσθετους προγνωστικούς παράγοντες, αλλά η παρέκταση εμφάνισε μη εφικτή απόκλιση ανά φύλο, καθιστώντας

---

<sup>6</sup> Generalized Linear Modeling

αναγκαία τη χρήση υποκειμενικά προσαρμοσμένων συντελεστών. Σε τέτοιες περιπτώσεις, μπορεί να είναι δυνατή η χρήση γραμμικών splines<sup>7</sup> για τις υποπεριόδους, αλλά αυτή η λύση περιορίζεται στη βραχυπρόθεσμη πρόβλεψη, επειδή δεν μπορούν να εντοπιστούν σημεία καμπής (De Beer, 1989).

Οι Renshaw & Haberman (2000) χρησιμοποίησαν GLM για να μοντελοποιήσουν τους παράγοντες μείωσης της θνησιμότητας και προσδιόρισαν τις συνθήκες υπό τις οποίες η υποκείμενη δομή του μοντέλου τους είναι ίδια με εκείνη του μοντέλου Lee-Carter. Ανέπτυξαν μια προσέγγιση βασισμένη στη GLM η οποία παραλληλίζεται με τη μέθοδο Lee-Carter, συμπεριλαμβανομένης της αντιστοίχισης παρατηρούμενων και αναμενόμενων συνολικών θανάτων (Renshaw & Haberman, 2003b), και προσάρμοσαν τη μέθοδο Lee-Carter για χρήση στην πρόβλεψη των παραγόντων μείωσης της θνησιμότητας (Renshaw & Haberman, 2003c). Οι Renshaw & Haberman (2003a) επέκτειναν το υπόδειγμα Lee-Carter ώστε να συμπεριλάβει τον δεύτερο όρο αποσύνθεσης μοναδιαίων τιμών (SVD)<sup>8</sup> για να επιτρέψει την "ηλικιακή ενίσχυση" και συνέκριναν την πρόβλεψή του με παρόμοια ενισχυμένες GLM και λογαριθμογραμμικές προβλέψεις Poisson.

Σε μια εναλλακτική προσέγγιση που χρησιμοποιεί το πλαίσιο GLM, οι Currie *et al.* (2004) χρησιμοποίησαν p-splines (penalised b-splines)<sup>9</sup> για να επιβάλουν ομαλότητα στην επιφάνεια της έντασης της θνησιμότητας ανά ηλικία και περίοδο. Οι ρυθμοί πρόβλεψης εκτιμώνται ταυτόχρονα με την προσαρμογή της επιφάνειας θνησιμότητας και έχουν «καλή συμπεριφορά» λόγω της ποινής που επιβάλλεται στην έλλειψη ομαλότητας. Το CMIB (2005, 2006) αναλύει περαιτέρω τα χαρακτηριστικά της μεθόδου. Η εφαρμογή της σε ασφαλισμένους άνδρες στο Ηνωμένο Βασίλειο παρήγαγε ποσοστά πρόβλεψης που μειώνονται πιο αργά από ό,τι στην αντίστοιχη πρόβλεψη Lee-Carter (Currie *et al.*, 2004).

---

<sup>7</sup> Στα μαθηματικά, η spline είναι μια ειδική συνάρτηση που ορίζεται κατά τμήματα από πολυώνυμα.

<sup>8</sup> Singular value decomposition: Στη γραμμική άλγεβρα, η αποσύνθεση μοναδιαίων τιμών (SVD) είναι μια παραγοντοποίηση ενός πραγματικού ή μιγαδικού πίνακα.

<sup>9</sup> Η p-spline ή spline βάσης είναι μια συνάρτηση spline που έχει ελάχιστη υποστήριξη σε σχέση με έναν δεδομένο βαθμό, ομαλότητα και κατάτμηση τομέα. Οποιαδήποτε συνάρτηση spline δεδομένου βαθμού μπορεί να εκφραστεί ως γραμμικός συνδυασμός p-spline αυτού του βαθμού.

#### 4.6 Μοντέλα Κοορτής

Τα μονοπαραγοντικά και διπαραγοντικά μοντέλα που συζητήθηκαν στο κεφάλαιο 2 (Lee-Carter) και στις ενότητες 3.4 και 3.5 μπορούν να εφαρμοστούν σε δεδομένα κοορτής για την παραγωγή μοντέλων θνησιμότητας κατά τη διάρκεια της ζωής. Τα μοντέλα αυτά και τα μοντέλα τριών παραγόντων για την ηλικία-περίοδο-κοορτή (APC)<sup>10</sup> συζητούνται στην παρούσα ενότητα.

Στην περίπτωση των μοντέλων ενός και δύο παραγόντων, η προσέγγιση της κοορτής έχει το πλεονέκτημα ότι είναι απαλλαγμένη από τις στρεβλώσεις του «ρυθμού» που προκαλούνται από τις μεταβολές στο χρόνο- όταν η μέση ηλικία θανάτου αυξάνεται (δηλαδή όταν η θνησιμότητα μειώνεται), το μέτρο του προσδόκιμου ζωής της περιόδου υπερεκτιμά το προσδόκιμο ζωής (Bongaarts & Feeney, 2002). Θεωρητικά, τα μοντέλα κοορτής παρέχουν καλύτερη βάση για την πρόβλεψη από τα μοντέλα περιόδου. Στην πράξη, ωστόσο, διάφοροι παράγοντες έχουν περιορίσει τη χρήση τους.

Το κύριο μειονέκτημα όλων των μοντέλων κοορτής είναι ότι συνεπάγονται μεγάλες απαιτήσεις σε δεδομένα- εάν ενδιαφέρει ολόκληρο το ηλικιακό φάσμα, μόνο τα δεδομένα που καλύπτουν έναν αιώνα περίπου παρέχουν μια πλήρη κοορτή και απαιτείται αντίστοιχα μια πολύ μεγαλύτερη σειρά ετήσιων δεδομένων για την πρόβλεψη. Όταν υπάρχουν διαθέσιμα δεδομένα, όπως σε ορισμένες ευρωπαϊκές χώρες, τα μοντέλα ολόκληρου του ηλικιακού φάσματος θα είναι περιορισμένα, δεδομένου ότι εξαρτώνται από την (ακατάλληλη) εμπειρία των κοορτών που γεννήθηκαν τον 19ο αιώνα (π.χ. Tabeau *et al.*, 2001b). Αυτά τα προβλήματα διαθεσιμότητας και δυνατότητας εφαρμογής των δεδομένων μειώνονται όταν ενδιαφέρει μόνο ένα μέρος του ηλικιακού εύρους.

Τα ελλιπή δεδομένα της κοορτής, δηλαδή τα δεδομένα που έχουν περικοπεί στην τρέχουσα ηλικία της κοορτής, μπορούν να χρησιμοποιηθούν στην πρόβλεψη εάν η μορφή του μοντέλου μπορεί να καθοριστεί εκ των προτέρων. Οι συναρτήσεις παραμετροποίησης ενός παράγοντα και τα σχεσιακά μοντέλα μπορούν να προσαρμοστούν σε ελλιπή δεδομένα κοορτής για να παρέχουν μια εκτίμηση της εμπειρίας της κοορτής πέραν της ηλικίας που έχουν περικοπεί τα δεδομένα- επειδή η ηλικία και ο χρόνος είναι ισοδύναμοι για μια κοορτή η εκτίμηση είναι επίσης μια πρόβλεψη. Η προσέγγιση αυτή χρησιμοποιείται στη μοντελοποίηση της

---

<sup>10</sup> Age-Period-Cohort



θνησιμότητας στις μεγαλύτερες ηλικίες (ενότητα 3.8). Ομοίως, τα μοντέλα δύο παραγόντων για τις ηλικιακές κοορτές μπορούν, θεωρητικά, να χρησιμοποιηθούν με μερικώς ελλιπή δεδομένα για τις κοορτές, καθώς ο πρόσθετος παράγοντας επιτρέπει την αξιοποίηση των πλήρων παρατηρήσεων για τις παλαιότερες κοορτές για την εκτίμηση της πλήρους εμπειρίας των νεότερων κοορτών. Ωστόσο, τα ελλιπή δεδομένα οδηγούν σε στατιστικά προβλήματα στη μοντελοποίηση και σε πρακτικά προβλήματα για την πρόβλεψη (Brass, 1974; Bell, 1997).

Δεδομένων των περιορισμών των δεδομένων κοορτής, μπορεί να είναι προτιμότερο να γίνουν προσαρμογές στις προβλέψεις περιόδου για να ληφθούν τιμές κοορτής. Ο Pollard (1998) πρότεινε μια απλή προσαρμογή του προσδόκιμου ζωής περιόδου με βάση τη θνησιμότητα Gompertz και ένα σταθερό ρυθμό μεταβολής κατά την ηλικία και το χρόνο. Πρέπει επίσης να λαμβάνονται υπόψη οι στρεβλώσεις του ρυθμού στο προσδόκιμο ζωής της περιόδου- οι Bongaarts & Feeney (2006) πρότειναν μια προσαρμογή για τις εν λόγω στρεβλώσεις, με επιπτώσεις στις προβλέψεις. Μια βελτίωση θα ήταν να υιοθετηθεί η προσέγγιση του παραλληλογράμμου-Lexi των Heathcote & Higgins (2001a) για τον καθορισμό των ποσοστών περιόδου. Μπορούν να εξεταστούν και άλλα συγκεντρωτικά μέτρα θνησιμότητας (Bongaarts & Feeney, 2003; Guillot, 2003; Sanderson & Scherbov, 2005).

Μια πραγματική πρόβλεψη κοορτής επιτεύχθηκε από το CMIB (2006) με τη χρήση της μεθόδου παλινδρόμησης p-spline (Currie *et al.*, 2004) για την εξομάλυνση των δεδομένων ηλικίας- κοορτής (όπου η ποινή εφαρμόζεται στην ηλικία και το έτος γέννησης) στις ηλικίες 20+. Πραγματοποιήθηκε επίσης μια ανάλογη πρόβλεψη για τις ηλικιακές περιόδους. Οι δοκιμές εκτός δείγματος έδειξαν ότι η μέθοδος "θα λειτουργούσε καλά τα τελευταία χρόνια" (CMIB, 2006).

Τα μοντέλα APC τριών παραγόντων περιλαμβάνουν την κοορτή ως τρίτο παράγοντα. Ωστόσο, η διάκριση μεταξύ των επιδράσεων της ηλικίας, της περιόδου και της κοορτής είναι δύσκολη λόγω του προβλήματος ταυτοποίησης- οποιοσδήποτε παράγοντας εξαρτάται γραμμικά από τους υπόλοιπους δύο (Tabeau, 2001). Η λύση σε αυτό το πρόβλημα είναι η χρήση τριπλής ταξινόμησης (επίσης γνωστή ως διπλά ταξινομημένα δεδομένα ή μη επικαλυπτόμενες κοορτές) κατά τη συλλογή δεδομένων, έτσι ώστε οι θάνατοι να μπορούν να κατανεμηθούν σε παραλληλόγραμμα-Lexi, οπότε η εξάρτηση δεν υφίσταται (Willekens & Baydar, 1986). Εναλλακτικά, μπορούν να επιβληθούν περιορισμοί στις παραμέτρους του μοντέλου (Van Hoorn & De Beer, 2001). Για να ξεπεραστεί το πρόβλημα ταυτοποίησης ελλείπει τριπλής ταξινόμησης δεδομένων, ο Wilmoth (1990) ανέπτυξε ένα περιγραφικό μοντέλο δύο

παραγόντων που περιλαμβάνει προσθετικές επιδράσεις ηλικίας και περιόδου και διάφορους πολλαπλασιαστικούς όρους αλληλεπίδρασης, στους οποίους μπορεί επίσης να προστεθεί ένας διαγώνιος όρος που αντιπροσωπεύει τη μέση επίδραση της κοορτής.

Το μοντέλο APC έχει εφαρμοστεί με χρησιμότητα για την περιγραφή του παρελθόντος, αλλά έχει θεωρηθεί λιγότερο χρήσιμο για τις μελλοντικές προβλέψεις. Σε μια ανασκόπηση των μοντέλων APC, ο Tabeau (2001) κατέληξε στο συμπέρασμα ότι η πρόβλεψη της θνησιμότητας με βάση τα μοντέλα APC δεν είναι εφικτή λόγω της δυσκολίας να υποθέσει κανείς τις επιδράσεις της μελλοντικής περιόδου (αν και οι επιδράσεις της ηλικίας και της κοορτής μπορούν να θεωρηθούν σταθερές) και ότι μόνο στις προβλέψεις συγκεκριμένων ασθενειών θα ήταν διαθέσιμη επαρκής επιδημιολογική γνώση για τον προσδιορισμό της περιόδου.

Πιο πρόσφατα, οι Renshaw & Haberman (2006) επέκτειναν το μοντέλο Lee- Carter για να δημιουργήσουν το μοντέλο APC:

$$\ln \mu_{x,t} = a_x + b^1_x k_t + b^2_x k_{t-x} + z_{x,t}$$

όπου:  $\mu_{x,t}$  είναι η δύναμη της θνησιμότητας στην ηλικία  $x$  το έτος  $t$ ,  $a_x$  είναι ο μέσος όρος των  $\ln \mu_{x,t}$  με την πάροδο του χρόνου, τα  $b^1_x$  και  $b^2_x$  μετρούν την απόκριση στην ηλικία  $x$  στις μεταβολές των  $k_t$  και  $t-x$  αντίστοιχα, ο όρος  $k_t$  αντιπροσωπεύει το συνολικό επίπεδο θνησιμότητας το έτος  $t$ , το  $k_{t-x}$  αντιπροσωπεύει το συνολικό επίπεδο θνησιμότητας για τη συνομοταξία που γεννήθηκε το έτος  $(t - x)$ , και το  $z_{x,t}$  είναι το υπόλοιπο. Οι προσαρμοσμένες παράμετροι  $k_t$  και  $k_{t-x}$  προβλέφθηκαν με τη χρήση μονομεταβλητών μοντέλων χρονοσειρών-θα μπορούσε επίσης να χρησιμοποιηθεί μια πολυμεταβλητή προσέγγιση. Το υπόδειγμα κατέγραψε επιτυχώς και τις τρεις κύριες επιδράσεις στη θνησιμότητα στο Ηνωμένο Βασίλειο, αντιπροσωπεύοντας σημαντική βελτίωση σε σχέση με τα υποδείγματα ηλικιακής περιόδου και ηλικιακής συνομοταξίας. Η δοκιμή εκτός δείγματος από το CMIB (2007) έδειξε ότι οι προβλέψεις Lee- Carter για τις ηλικιακές περιόδους θεωρήθηκαν γρήγορα ξεπερασμένες, τουλάχιστον εν μέρει, λόγω της έλλειψης αναγνώρισης των επιδράσεων των κοορτών που καταγράφονται από το μοντέλο APC. Μένει να δοκιμαστεί πλήρως το μοντέλο APC, ιδίως η σταθερότητά του ως απάντηση σε αλλαγές στην ηλικιακή κλίμακα ή στην περίοδο προσαρμογής.

Αν και οι επιδράσεις της κοορτής είναι γενικά πολύ μικρότερες από τις επιδράσεις της περιόδου, για ορισμένες χώρες είναι ωστόσο σαφώς διακριτές. Στο Ηνωμένο Βασίλειο, έχουν εντοπιστεί αρνητικές επιδράσεις κοορτής (δηλ. χαμηλότερη θνησιμότητα) στις ηλικίες των

ενηλίκων για τις κοορτές που γεννήθηκαν μεταξύ 1925 και 1945 (Willets, 2004; Willets *et al.*, 2004) και στις ηλικίες 65+ για τις προηγούμενες κοορτές που γεννήθηκαν μετά το 1900 (Government Actuary's Department, 2001). Από το 1992, οι επίσημες προβολές του Ηνωμένου Βασιλείου (οι οποίες προκύπτουν με τη χρήση στόχων) έχουν ενσωματώσει μια προσέγγιση κοορτής, αρχικά για τις κοορτες που γεννήθηκαν πριν από το 1947 και επεκτάθηκαν σε πιο πρόσφατες προβολές στις κοορτές που γεννήθηκαν πριν από το 1960 (Government Actuary's Department, 2001, 2006). Το αναλογιστικό επάγγελμα του Ηνωμένου Βασιλείου έχει εστιάσει πρόσφατα σε μοντέλα που ενσωματώνουν την κοορτή, όπως περιγράφεται παραπάνω (Continuous Mortality Investigation Bureau, 2005, 2006, 2007). Επιδράσεις κοορτής έχουν επίσης εντοπιστεί σε δεδομένα των ΗΠΑ (McNown & Rogers, 1989; Lee & Carter, 1992), σε γαλλικά δεδομένα (Wilmoth, 1990) και σε ιαπωνικά δεδομένα (Willets, 2004), μεταξύ άλλων.

#### 4.7 Αποσύνθεση των δεδομένων ανά αιτία θανάτου

Η πρόβλεψη της θνησιμότητας ανά ξεχωριστή αιτία θανάτου έχει υποστηριχθεί από θεωρητική άποψη ως μέσο για την αύξηση της ακρίβειας στη συνολική πρόβλεψη της θνησιμότητας (Crimmins, 1981), αλλά η μετέπειτα εμπειρία έχει συχνά αποδείξει το αντίθετο. Οι McNown & Rogers (1992) δεν διαπίστωσαν κανένα σταθερό διακριτό κέρδος στην ακρίβεια από την κατηγοριοποίηση των αιτιών θανάτου με τη χρήση του πολυεκθετικού μοντέλου. Ο Wilmoth (1995) απέδειξε ότι, για τα μοντέλα αναλογικών ρυθμών μεταβολής, συμπεριλαμβανομένων των Lee-Carter, οι προβλέψεις θνησιμότητας που βασίζονται στο άθροισμα των προβλέψεων για τις συγκεκριμένες αιτίες θα είναι πάντοτε υψηλότερες από εκείνες που βασίζονται στη συνολική θνησιμότητα, επειδή οι αιτίες θανάτου που μειώνονται αργά έρχονται να κυριαρχήσουν μακροπρόθεσμα. Χρησιμοποιώντας μη γραμμικά μοντέλα APC για τις ηλικίες 60 έως 85+, ο Caselli (1996) διαπίστωσε ότι αυτό ισχύει για τις γυναίκες, αλλά αντιστρέφεται για τους άνδρες. Χρησιμοποιώντας μη γραμμικά μοντέλα ηλικιακής περιόδου, οι Tabeau *et al.* (2001b) διαπίστωσαν επίσης αυτή τη διαφορά μεταξύ των δύο φύλων στις ηλικίες 40, 60 και 80 ετών για τη Γαλλία, την Ιταλία και τις Κάτω Χώρες, αλλά όχι για τη Νορβηγία. Ο Alho (1991) εξετάζει τις συνθήκες υπό τις οποίες η αποσύνθεση (decomposition) μειονεκτεί και εκείνες υπό τις οποίες δίνει παρόμοια αποτελέσματα με την πρόβλεψη της συνολικής θνησιμότητας.

Στην πράξη, η αποσύνθεση με βάση την αιτία θανάτου έχει γενικά περιορισμένη αξία. Οι λόγοι γι' αυτό περιλαμβάνουν την ομοιότητα των ηλικιακών προτύπων για τις κύριες αιτίες, την αναξιπιστία της αναφοράς της αιτίας θανάτου στις μεγαλύτερες ηλικίες όπου συμβαίνουν οι

περισσότεροι θάνατοι και το γεγονός ότι η μείωση των αιτιών θανάτου μπορεί να έχει ελάχιστη επίδραση στη συνολική θνησιμότητα (Murphy, 1995). Η λανθασμένη εξειδίκευση του μοντέλου και η παρουσία πρόδρομων δεικτών (όπου οι αλλαγές σε μια αιτία προηγούνται συστηματικά των αλλαγών σε μια άλλη αιτία) μπορεί επίσης να οδηγήσει σε μειωμένη ακρίβεια πρόβλεψης του μοντέλου (Alho, 1991). Περαιτέρω, η υπόθεση ότι οι διακριτές ασθένειες οδηγούν τη συνολική θνησιμότητα είναι αμφισβητήσιμη, ιδίως στις μεγαλύτερες ηλικίες, και ο κατάλληλος τρόπος συνδυασμού των συσχετιζόμενων μέτρων συγκεκριμένων-αιτιών δεν είναι σαφής (Stoto & Durch, 1993). Πρόσθετοι περιορισμοί προκύπτουν επειδή οι διαθέσιμες χρονοσειρές δεδομένων για τις συγκεκριμένες αιτίες είναι συχνά πολύ σύντομες για την πρόβλεψη με τη χρήση μεθόδων (ARIMA) και οι προς πρόβλεψη παράμετροι είναι συχνά λιγότερο σταθερές για τις αιτίες θανάτου απ' ό,τι για τη συνολική θνησιμότητα (Wilmoth, 1995). Η διαστατικότητα αυξάνεται επίσης με την αποσύνθεση. Στις εργασίες των Alho (1991), Stoto & Durch (1993), Tuljapurkar (1998) και Wilmoth (2005) τίθενται αυτά και άλλα σχετικά ζητήματα.

Ο περιορισμός της αποσύνθεσης των αιτιών θανάτου σε ηλικίες κάτω των 90 ετών (ή περίπου), όπως συνιστάται από τους Tabeau κ.ά. (2001b) λόγω των προβλημάτων στον προσδιορισμό της αιτίας θανάτου σε μεγαλύτερες ηλικίες, αποτελεί σημαντικό περιορισμό της προσέγγισης (οι περισσότεροι θάνατοι συμβαίνουν σε πολύ μεγάλες ηλικίες). Στις νεότερες ηλικίες, οι συγγραφείς υποστηρίζουν την επιλογή των κατάλληλων αιτιών θανάτου για τα διάφορα ηλικιακά τμήματα και προειδοποιούν κατά της πρόβλεψης του υπολοίπου (μιας ομάδας αιτιών) λόγω της μεροληψίας που προκύπτει από την εσφαλμένη ταξινόμηση των θανάτων ανά αιτία. Με άλλα λόγια, ιδίως για τις μακροπρόθεσμες προβλέψεις, δεν υποστηρίζουν την αποσύνθεση ως μέσο πρόβλεψης της συνολικής θνησιμότητας, συνιστώντας μάλλον την πρόβλεψη του υπολοίπου ως τη διαφορά μεταξύ των προβλέψεων για τη συνολική και την κύρια αιτία θανάτου.

Δεδομένων αυτών των περιορισμών, οι περισσότεροι επίσημοι οργανισμοί και αναλογιστές δεν έχουν χρησιμοποιήσει την αποσύνθεση ανά αιτία θανάτου στην πρόβλεψη της θνησιμότητας. Το Τμήμα Αναλογιστών της Κυβέρνησης του Ηνωμένου Βασιλείου (2001) συνέστησε να μην γίνεται αποσύνθεση βάσει αιτίας θανάτου, ενώ το CMIB (2004) δήλωσε ότι οι συνολικές προβλέψεις είναι πιο κατάλληλες για τους σκοπούς του. Μια αξιοσημείωτη εξαίρεση αποτελεί

η αμερικανική υπηρεσία κοινωνικής ασφάλισης (SSA)<sup>11</sup> (Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds, 2004), αλλά οι τεχνικές επιτροπές της SSA του 1999 και του 2003 τάχθηκαν κατά της αποσύνθεσης με βάση την αιτία θανάτου (Waldron, 2005). Σε γενικές γραμμές, τα αντιληπτά πλεονεκτήματα της αποσύνθεσης αντισταθμίζονται από τους περιορισμούς που συνεπάγεται.

Ένας τομέας στον οποίο έχει χρησιμοποιηθεί η αποσύνθεση με βάση την αιτία θανάτου είναι τα μοντέλα (μερικής) εξάλειψης αιτιών και καθυστέρησης αιτιών (Manton *et al.*, 1980; Olshansky, 1987, 1988; Kunst *et al.*, 2002- Le Bras, 2005). Η βάση αυτών των μοντέλων μπορεί να είναι η προσδοκία μέσω της στόχευσης και της τεκμηριωμένης γνώμης, η οποία έχει συχνά οδηγήσει σε συντηρητικές προβλέψεις για τη μείωση της θνησιμότητας, ή η εξήγηση μέσω επιδημιολογικών μοντέλων ορισμένων αιτιών.

Ο Bonggaarts (2006) προέβλεψε το προσδόκιμο ζωής αναλύοντας τη θνησιμότητα σε τρεις συνιστώσες: νεανική, ηλικίες 25+ και υπόβαθρο και γήρανση. Αυτές δεν αντιστοιχούν σε τυποποιημένες αιτίες θανάτου- τέτοια δεδομένα δεν χρησιμοποιούνται. Η γεροντική θνησιμότητα διασπάται περαιτέρω με βάση την απόδοση στο κάπνισμα, με βάση τα διαχρονικά εξαρτώμενα μέσα ποσοστά στις ανεπτυγμένες χώρες. Η κύρια κινητήρια δύναμη της πρόβλεψης είναι η θνησιμότητα γήρανσης που δεν αποδίδεται στο κάπνισμα- η νεανική θνησιμότητα και η θνησιμότητα υποβάθρου είναι αρκετά σταθερές και σχετικά ασήμαντες, ενώ η θνησιμότητα που οφείλεται στο κάπνισμα μπορεί να διαμορφωθεί. Η προσέγγιση αποφεύγει πολλούς από τους περιορισμούς που συνεπάγεται η αποσύνθεση των αιτιών θανάτου, ενώ παράλληλα αποτυπώνει τις τάσεις στις κύριες συνιστώσες της θνησιμότητας.

#### 4.8 Θνησιμότητα γήρατος

Καθώς η θνησιμότητα μειώνεται, δίνεται ολοένα και μεγαλύτερη σημασία στη συνάρτηση που αντιπροσωπεύει τη θνησιμότητα στις μεγάλες ηλικίες, στις οποίες συμβαίνουν οι περισσότεροι θάνατοι. Η συνάρτηση Gompertz έχει χρησιμοποιηθεί ευρέως για την αναπαράσταση της θνησιμότητας λόγω γήρανσης (θνησιμότητα λόγω γήρανσης) (Pollard & Valkovics, 1992; Olshansky & Carnes, 1997), αλλά στις μεγαλύτερες ηλικίες υπερεκτιμά τα παρατηρούμενα ποσοστά. Αυτή η μεροληψία μπορεί να αντιμετωπιστεί με την ενσωμάτωση του Gompertz σε

---

<sup>11</sup> Οι προβλέψεις για την ηλικία, το φύλο και την αιτία βασίζονται στη γνώμη των εμπειρογνομόνων.

μια λογιστική συνάρτηση, περιλαμβάνοντας έτσι μια ασύμπτωτη- για μια σύγκριση των μοντέλων (Thatcher *et al.* 1998; Thatcher, 1999).

Ο Bongaarts (2005) χρησιμοποίησε ένα μεταβαλλόμενο λογιστικό μοντέλο για την πρόβλεψη της θνησιμότητας ανά ηλικία στις ηλικίες 25+. Το μοντέλο βασίζεται στην παραμετροποίηση ενός παράγοντα:

$$\mu_x = \frac{ae^{\beta x}}{1 + ae^{\beta x}} + \gamma$$

όπου  $\mu_x$  είναι η ένταση της θνησιμότητας στην ηλικία  $x$ , και  $a$ ,  $\beta$  και  $\gamma$  είναι παράμετροι που αντιπροσωπεύουν, αντίστοιχα, το επίπεδο της γκομπερτζιανής θνησιμότητας, τον εκθετικό ρυθμό αύξησης της γκομπερτζιανής θνησιμότητας με την ηλικία και το επίπεδο της θνησιμότητας υποβάθρου που δεν μεταβάλλεται με την ηλικία. Ο πρώτος όρος αντιπροσωπεύει την ένταση της γεροντικής θνησιμότητας. Η παράμετρος κλίσης  $\beta$  έχει αποδειχθεί ότι είναι (σχεδόν) σταθερή με την πάροδο του χρόνου, αν και η τιμή της ποικίλλει μεταξύ των πληθυσμών (Bongaarts, 2005). Η μετατόπιση της λογιστικής χρησιμοποιεί αυτή τη σταθερή κλίση, με άλλα λόγια την ισοδυναμία μεταξύ της μείωσης της γεροντικής θνησιμότητας με την πάροδο του χρόνου και της μετατόπισης του χρονοδιαγράμματος της γεροντικής δύναμης της θνησιμότητας σε μεγαλύτερες ηλικίες. Κατά την πρόβλεψη, οι ετήσιες σειρές των χρονικά μεταβαλλόμενων παραμέτρων  $a$  και  $\gamma$  προεκτείνονται ανεξάρτητα και θεωρείται σταθερό  $\beta$ . Με τη μοντελοποίηση των βελτιώσεων της γεροντικής θνησιμότητας με μια σταθερή μετατόπιση του προγράμματος θνησιμότητας με την ηλικία, αντί για ένα σταθερό πρότυπο μείωσης των ειδικών για την ηλικία ποσοστών (όπως στην Lee-Carter), το ηλικιακό πρότυπο θνησιμότητας παραμένει αληθοφανές ακόμη και για πιο μακροπρόθεσμες προβλέψεις. Μια μετατοπισμένη λογιστική έχει χρησιμοποιηθεί σε συνδυασμό με τη μέθοδο Lee-Carter σε μακροπρόθεσμες προβλέψεις του ΟΗΕ (United Nations, 2004; Buettner & Zlotnik, 2005).

Εκτός από τις εκτιμήσεις της μοντελοποίησης που πρέπει να ληφθούν υπόψη, δυσκολίες προκύπτουν επειδή τα δεδομένα στις μεγαλύτερες ηλικίες είναι συχνά προβληματικά (Kannisto, 1994; Kannisto *et al.*, 1994; Tuljapurkar & Boe, 1998). Όταν τα δεδομένα είναι αποσπασματικά, η μοντελοποίηση των παλαιότερων ηλικιών είναι απαραίτητη. Για το σκοπό αυτό, οι υπάρχοντες πρότυποι πίνακες ζωής είναι περιορισμένης χρησιμότητας για πληθυσμούς με χαμηλή θνησιμότητα, επειδή το εύρος των ηλικιών δεν εκτείνεται πέραν των 100 ετών (αν αυτό συμβαίνει) και τα μέγιστα προσδόκιμα ζωής έχουν ξεπεραστεί (για παράδειγμα, 76,6 έτη

για τους άνδρες και 80,0 έτη για τις γυναίκες στους πίνακες ζωής του Princeton (Coale & Guo, 1989)).

Η μέθοδος Coale & Kisker (1990) για την επέκταση της παρατηρούμενης ή προβλεπόμενης θνησιμότητας στις μεγαλύτερες ηλικίες υποθέτει ότι τα ποσοστά θνησιμότητας αυξάνονται με γραμμικά φθίνοντα ρυθμό μέχρι κάποια ηλικία-στόχο. Σε μια εξέταση της προσαρμογής 11 συναρτήσεων στις ηλικίες 80 έως 110 ετών, οι Boleslawki & Tabeau (2001) διαπίστωσαν ότι η μέθοδος Coale- Kisker είναι ανώτερη όταν υπάρχουν αξιόπιστα δεδομένα τουλάχιστον μέχρι την ηλικία των 90 ετών, αλλά συνέστησαν πολυωνυμική παρέκταση όταν υπάρχουν αξιόπιστα δεδομένα μόνο μέχρι την ηλικία των 85 ετών. Οι Buettner & Zlotnik (2005) διαπίστωσαν ότι το λογιστικό μοντέλο, το οποίο δεν απαιτεί τον προσδιορισμό ενός ανώτερου ορίου, είναι λιγότερο περιοριστικό και συνεπώς προτιμότερο από το Coale-Kisker για πολύ μακροπρόθεσμες προβλέψεις. Μια εναλλακτική προσέγγιση είναι το σχεσιακό μοντέλο που αναπτύχθηκε από τους Himes et al. (1994), το οποίο επεκτείνει την παρατηρούμενη θνησιμότητα έως την ηλικία των 99 ετών μέσω ενός γραμμικού σχεσιακού μοντέλου που βασίζεται στον μετασχηματισμό logit<sup>12</sup> του  $\mu_x$  από την ηλικία των 45 ετών- η μέθοδος αυτή δεν κάνει επίσης καμία υπόθεση σχετικά με το ανώτατο όριο. Η μέθοδος των Coale & Guo (1989), η οποία όντως χρησιμοποιεί ηλικία-στόχο, υποστηρίχθηκε από τη μοντελοποίηση των Renshaw & Haberman (2003b).

#### 4.9 Η εκτίμηση της αβεβαιότητας

Ενώ η σημειακή πρόβλεψη της θνησιμότητας είναι πρωταρχικού ενδιαφέροντος, είναι επίσης σημαντικό να παρέχεται μια ένδειξη του πιθανού εύρους των μελλοντικών τιμών. Παραδοσιακά, αυτό αντιμετωπίζεται μέσω σεναρίων, όπου τα υψηλά και τα χαμηλά σενάρια παρέχουν το εύρος γύρω από το μεσαίο σενάριο (ή την πρόβλεψη). Αυτή η υποκειμενική προσέγγιση έχει αρκετές ανεπάρκειες (Lee & Tuljapurkar, 1994; Booth, 2004), ιδίως ότι δεν είναι δυνατόν να προσδιοριστεί η πιθανότητα του υψηλού-χαμηλού εύρους. Πιο πρόσφατες μέθοδοι εκτιμούν την αβεβαιότητα της πρόβλεψης ώστε να παρέχουν πιθανοτικά διαστήματα πρόβλεψης.

---

<sup>12</sup> Η συνάρτηση logit είναι η συνάρτηση ποσοστημορίων που σχετίζεται με την τυπική λογιστική κατανομή. Έχει πολλές χρήσεις στην ανάλυση δεδομένων και τη μηχανική μάθηση, ιδίως σε μετασχηματισμούς δεδομένων

Η εκτίμηση της αβεβαιότητας μιας σημειακής πρόβλεψης ενός ποσοστού θνησιμότητας ανά ηλικία ή του προσδόκιμου ζωής αποτελεί αντικείμενο συνεχιζόμενης έρευνας (Lutz & Goldstein, 2004). Έχουν χρησιμοποιηθεί τρεις γενικές προσεγγίσεις:

1. Εκτίμηση εκ των προτέρων σφαλμάτων βάσει μοντέλου
2. Εκτίμηση εκ των προτέρων σφαλμάτων βάσει εμπειρογνομόνων
3. Εκτίμηση εκ των υστέρων σφαλμάτων

Η εκτίμηση ποικίλλει ανάλογα με τη μέθοδο που χρησιμοποιείται και με το βαθμό στον οποίο λαμβάνονται υπόψη οι διάφορες πηγές σφάλματος. Επιπλέον, κατά την εκτίμηση της αβεβαιότητας των μέτρων που αθροίζουν ή υπολογίζουν κατά μέσο όρο μεμονωμένα μεγέθη πρόβλεψης, είναι απαραίτητο να λαμβάνονται υπόψη οι σχετικές συσχετίσεις μεταξύ των μεμονωμένων σφαλμάτων πρόβλεψης.

#### 4.9.1 Πηγές σφάλματος

Υπάρχουν πολλές πηγές σφάλματος στην πρόβλεψη των δημογραφικών ποσοστών. Ο Alho (1990) ταξινομήσε το σφάλμα σε τέσσερις διαφορετικές, αλλά συναφείς, πηγές:

- Εσφαλμένη εξειδίκευση του μοντέλου
- Εκτίμηση των παραμέτρων
- Σφάλματα τεκμηριωμένης κρίσης και
- Εγγενής τυχαία διακύμανση (Keilman, 1990)

Η πρώτη, η λανθασμένη εξειδίκευση του μοντέλου, μπορεί να συμβεί σε σχέση με το υποτιθέμενο υποκείμενο μοντέλο των δεδομένων (π.χ. μια συνάρτηση παραμετροποίησης) ή το μοντέλο πρόβλεψης (π.χ. ένα μοντέλο χρονοσειρών) ή και τα δύο. Οι αποκλίσεις και οι συνδιακυμάνσεις των εκτιμήσεων των παραμέτρων (η δεύτερη πηγή σφάλματος) προκύπτουν είτε με τυπική εκτίμηση είτε με bootstrapping - επαναδειγματοληψία από τα αρχικά δεδομένα για τη δημιουργία επαναληπτικών συνόλων δεδομένων από τα οποία μπορούν να εκτιμηθούν οι αποκλίσεις και οι συνδιακυμάνσεις. Όσον αφορά την τρίτη πηγή, η τεκμηριωμένη κρίση στη στατιστική μοντελοποίηση αναφέρεται στις «εκ των προτέρων» πεποιθήσεις σχετικά με τις παραμέτρους του μοντέλου και τη βαρύτητα που τους δίνεται στην πρόβλεψη (Alho, 1992; Alho, 2005). Όσον αφορά την τελευταία πηγή, οι παρατηρούμενοι θάνατοι έχουν συχνά θεωρηθεί ότι υπόκεινται σε μεταβλητότητα Poisson (Brillinger, 1986), αν και αυτό μπορεί να υποτιμά την πραγματική διακύμανση. Οι προσεγγίσεις GLM χρησιμοποιούν συχνά την υπερδιασκορπισμένη Poisson (Renshaw & Haberman, 2003b; Renshaw & Haberman, 2006),



ενώ οι Li et al. (2012) χρησιμοποιούν την αρνητική διωνυμική για να λάβουν υπόψη την ατομική ετερογένεια. Εκτός από αυτή την τυχαία διακύμανση, τα δεδομένα υπόκεινται σε μη τυχαία σφάλματα μέτρησης, συμπεριλαμβανομένων των λανθασμένων δηλώσεων ηλικίας και των παραλείψεων (Stoto, 1988).

Ιδανικά, όλες οι πηγές σφάλματος θα πρέπει να λαμβάνονται υπόψη κατά την παρουσίαση των διαστημάτων πρόβλεψης για τις προβλέψεις. Τα διαστήματα πρόβλεψης που παραλείπουν μία ή περισσότερες πηγές αβεβαιότητας θα δώσουν παραπλανητική εντύπωση εμπιστοσύνης στην πρόβλεψη. Οι εκτιμήσεις της αβεβαιότητας που βασίζονται μόνο στην αβεβαιότητα των παραμέτρων της χρονοσειράς θα είναι υπερβολικά στενές, ιδίως όταν το μοντέλο είναι ιδιαίτερα δομημένο, όπως το μοντέλο Lee-Carter με πρόβλεψη τυχαίου περιπάτου με διολίσθηση. Ο Cairns (2000) σημειώνει ότι η αβεβαιότητα των παραμέτρων και τα σφάλματα που οφείλονται σε λανθασμένη εξειδίκευση του μοντέλου παραμελούνται σχετικά στην αναλογιστική επιστήμη.

#### 4.9.2 Συσχετίσεις μεταξύ των σφαλμάτων πρόβλεψης

Τρεις τύποι συσχέτισης μεταξύ των σφαλμάτων πρόβλεψης μπορεί να έχουν σημασία για την εκτίμηση της αβεβαιότητας στην πρόβλεψη της θνησιμότητας. Πρόκειται για τη διαχρονική ή σειριακή συσχέτιση, τη συσχέτιση μεταξύ των ηλικιών και τη συσχέτιση μεταξύ υποπληθυσμών. Αυτές οι συσχετίσεις επηρεάζουν το βαθμό στον οποίο τα (θετικά και αρνητικά) σφάλματα ακυρώνονται κατά τη μέση τιμή ή τη συνάθροιση- όσο ισχυρότερη είναι η θετική συσχέτιση, τόσο μικρότερος είναι ο βαθμός ακύρωσης και τόσο ευρύτερο είναι το διάστημα πρόβλεψης του συνολικού μέτρου (Lee, 1999). Οι συσχετίσεις μεταξύ των σφαλμάτων λαμβάνονται υπόψη στην πρόβλεψη είτε αναλυτικά (Lee & Tuljapurkar, 2000) είτε, συνηθέστερα, μέσω προσομοίωσης. Όταν χρησιμοποιείται η τελευταία, όλες οι πηγές σφάλματος μπορούν να ληφθούν υπόψη και να συνδυαστούν κατάλληλα.

Είναι αναμενόμενο ότι τα προβλεπόμενα, ειδικά για την ηλικία, ποσοστά θανάτου συσχετίζονται θετικά με την πάροδο του χρόνου, οι Lee & Tuljapurkar (2000) παρατήρησαν πολύ υψηλή αυτοσυσχέτιση στους ρυθμούς μεταβολής. Το προσδόκιμο ζωής συσχετίζεται ομοίως. Οι μέθοδοι χρονοσειρών, οι οποίες λαμβάνουν υπόψη την αυτοσυσχέτιση με ειδική μοντελοποίηση της σχέσης μεταξύ των χρονικά εξαρτώμενων παραμέτρων σε διαδοχικά έτη, τείνουν συνεπώς να παράγουν μεγάλα διαστήματα πρόβλεψης. Οι προσεγγίσεις χρονολογικών σειρών επιτρέπουν την προσομοίωση μεμονωμένων δειγματικών διαδρομών, οι οποίες μπορούν στη συνέχεια να χρησιμοποιηθούν ως είσοδος σε ένα πιο σύνθετο στοχαστικό μοντέλο

που περιλαμβάνει άλλες πηγές σφάλματος. Αξίζει να σημειωθεί ότι οι μέθοδοι που βασίζονται στην παλινδρόμηση δεν είναι γενικά χρήσιμες για τη δημιουργία μελλοντικών δειγματικών διαδρομών.

Οι θετικές συσχετίσεις μεταξύ των ηλικιών προκύπτουν από την κανονικότητα των μοντέλων ηλικίας και επειδή οι συνθήκες που οδηγούν σε υψηλά ή χαμηλά ετήσια ποσοστά θνησιμότητας συχνά ισχύουν για ολόκληρο το ηλικιακό φάσμα ή για μια ευρεία ηλικιακή ομάδα. Οι υψηλές θετικές συσχετίσεις οδηγούν σε ευρύτερα διαστήματα πρόβλεψης για μέτρα που αθροίζουν ή υπολογίζουν κατά μέσο όρο την ηλικία, όπως το προσδόκιμο ζωής. Ο Alho (1998) βρήκε συσχετίσεις της τάξης του 0,95 μεταξύ κοντινών ηλικιών για τα ποσοστά θνησιμότητας. Οι Lee & Tuljapurkar (2000) κατέληξαν στο συμπέρασμα ότι μπορεί να υποτεθεί τέλεια συσχέτιση, με αποτέλεσμα κάποια υπερεκτίμηση του διαστήματος πρόβλεψης.

Σε ορισμένες περιπτώσεις, είναι απαραίτητο να ληφθούν υπόψη οι συσχετίσεις μεταξύ των υποπληθυσμών- για παράδειγμα, κατά το συνδυασμό των προβλέψεων θνησιμότητας των υποπληθυσμών. Οι διακρατικές συσχετίσεις δεν μπορούν να αγνοηθούν στις ευρωπαϊκές προβλέψεις (Alho et al., 2006). Οι διακρατικές συσχετίσεις μπορούν να αντιμετωπιστούν με μεταγενέστερη επεξεργασία των αποτελεσμάτων των προβλέψεων (Alho & Spencer, 2005). Προκύπτουν επίσης θετικές συσχετίσεις μεταξύ των δύο φύλων- η συνήθης υπόθεση της τέλει συσχέτισης οδηγεί σε κάποια υπερεκτίμηση της αβεβαιότητας των πληθυσμιακών προβλέψεων (Keilman et al., 2002).

#### 4.9.3 Εκ των προτέρων εκτίμηση σφαλμάτων βάσει μοντέλου

Τα στατιστικά μοντέλα παρέχουν μια βάση για την εκτίμηση του σφάλματος από τις διάφορες πηγές.<sup>13</sup> Η παρούσα ενότητα εξετάζει αρχικά την εκτίμηση του σφάλματος σε σχέση με τις δύο κύριες μεθόδους παρέκτασης, Lee-Carter και GLM. Τα μοντέλα αυτά επιτρέπουν την εκτίμηση της αβεβαιότητας των παραμέτρων και του τυχαίου σφάλματος, αλλά δεν καλύπτουν τα σφάλματα που προκύπτουν από τον εσφαλμένο προσδιορισμό του μοντέλου και την τεκμηριωμένη κρίση. Αυτές οι πρόσθετες πηγές σφάλματος μπορούν να μοντελοποιηθούν ειδικά σε ένα πιο ολοκληρωμένο πλαίσιο, όπως αναλύεται παρακάτω.

---

<sup>13</sup> Οι Denton κ.ά. (2005) παρουσιάζουν επίσης μια μη παραμετρική μέθοδο χωρίς μοντέλα που βασίζεται σε διαδικασίες bootstrap και μια μερικώς παραμετρική μέθοδο που βασίζεται στο αυτοπαλίνδρομο bootstrap.

Στις προσεγγίσεις Lee-Carter και στις συναφείς προσεγγίσεις χρονοσειρών, η αβεβαιότητα πρόβλεψης προκύπτει τόσο από τα υποκείμενα μοντέλα όσο και από τα μοντέλα χρονοσειρών. Κατά την εκτίμηση της αβεβαιότητας, οι Lee & Carter (1992) συμπεριέλαβαν μόνο την αβεβαιότητα στην παρέκταση, επειδή αυτή συνήθως κυριαρχεί μακροπρόθεσμα. Οι Booth et al. (2002) έλαβαν επιπλέον υπόψη το σφάλμα στην εκτίμηση του  $d$ . Αυτό αύξησε το τυπικό σφάλμα μιας πρόβλεψης ενός έτους κατά λιγότερο από 2%, αλλά μιας πρόβλεψης 51 ετών κατά 25% έως 63% ανάλογα με το μήκος της περιόδου προσαρμογής.

Η ακριβής εκτίμηση των διαστημάτων πρόβλεψης για τις μεθόδους Lee-Carter και τις συναφείς μεθόδους αποτέλεσε αντικείμενο πρόσφατης έρευνας. Μέθοδοι που ενσωματώνουν όλες τις πηγές αβεβαιότητας των παραμέτρων έχουν αναπτυχθεί στο πλαίσιο του λογαριθμογραμμικού μοντέλου Poisson από τους Koissi et al. (2006) χρησιμοποιώντας υπολειμματικό (μη παραμετρικό) bootstrapping, από τους Brouhns et al. (2005) χρησιμοποιώντας παραμετρικό bootstrapping και από τους Li et al. (2012) χρησιμοποιώντας παραμετρικό bootstrapping μαζί με μια χαλάρωση της δομής Lee-Carter για να επιτραπεί η ατομική ετερογένεια σε κάθε κελί ηλικιακής περιόδου. Οι Koissi & Shapiro (2006) χρησιμοποίησαν μια ασαφή διατύπωση μοντέλου. Ο Pedroza (2006) χρησιμοποίησε ένα πλαίσιο χώρου καταστάσεων σε συνδυασμό με την εκτίμηση κατά Bayes για την παραγωγή διαστημάτων πρόβλεψης προβλέψεων που ενσωματώνουν όλες τις πηγές σφάλματος.

Όταν χρησιμοποιούνται προσεγγίσεις GLM, παρέχονται εκτιμήσεις των αποκλίσεων και των συνδιακυμάνσεων των εκτιμήσεων των παραμέτρων και, έμμεσα, της εγγενούς τυχαίας διακύμανσης. Από αυτές, η αβεβαιότητα των προβλέψεων μπορεί να προκύψει, είτε αναλυτικά είτε με προσομοίωση, με τη μορφή της κατανομής της πρόβλεψης για κάθε έτος. Τα εκατοστημόρια μπορούν να χρησιμοποιηθούν για τον προσδιορισμό των διαστημάτων πρόβλεψης. Αυτές οι παραγόμενες κατανομές μπορούν να χρησιμοποιηθούν σε πιο σύνθετα μοντέλα που περιλαμβάνουν άλλες πηγές σφάλματος. Καθώς η αυτοσυσχέτιση δεν μοντελοποιείται ειδικά, δεν δημιουργούνται μεμονωμένες δειγματικές διαδρομές.

Μια διαφορετική προσέγγιση που βασίζεται σε μοντέλα είναι η μοντελοποίηση των σφαλμάτων πρόβλεψης καθαυτών. Οι Alho & Spencer (1985) χρησιμοποίησαν ένα παραμετρικό υπόδειγμα των κύριων πηγών σφάλματος για την κατασκευή διαστημάτων πρόβλεψης για βραχυπρόθεσμους ορίζοντες. Το υπόδειγμα τους περιλαμβάνει την εσφαλμένη εξειδίκευση του υποδείγματος, την τεκμηριωμένη κρίση και την εκτίμηση των παραμέτρων. Οι εκτιμήσεις της αβεβαιότητας που βασίζονται σε ιστορικές αφελείς προβλέψεις παρέχουν

μια συντηρητική βάση για την αξιολόγηση των σφαλμάτων πρόβλεψης. Ο Alho (1998, 2003) συσχέτισε τα σφάλματα που παρήγαγαν οι ιστορικές αφελείς προβλέψεις για τη Φινλανδία με το μήκος της πρόβλεψης και το χρησιμοποίησε για να ενημερώσει τα διαστήματα πρόβλεψης για τις νέες φινλανδικές προβλέψεις. Όταν δεν υπάρχουν ιστορικά δεδομένα για έναν πληθυσμό, οι πληροφορίες σχετικά με τη διακύμανση και την αυτοσυσχέτιση των σφαλμάτων μπορούν να δανειστούν από έναν άλλο κατάλληλο πληθυσμό. Η προσέγγιση αυτή εξετάζεται λεπτομερώς στο Alho & Spencer (2005).

Ο Cairns (2000) χρησιμοποίησε το παράδειγμα του Bayes για να αναπτύξει μια συνεκτική προσέγγιση για την εκτίμηση όλων των πηγών αβεβαιότητας. Η εκτίμηση της αβεβαιότητας λόγω εσφαλμένης εξειδίκευσης του μοντέλου απαιτεί τη γνώση των πιθανών μοντέλων και τον αριθμητικό προσδιορισμό των σχετικών πιθανοτήτων- μια πιο εφικτή προσέγγιση είναι να χρησιμοποιηθεί ο έλεγχος ευαισθησίας για να εκτιμηθεί η επίδραση της επιλογής του μοντέλου στην τελική πρόβλεψη (CMIB, 2004).

#### 4.9.4 Εκ των προτέρων εκτίμηση σφάλματος βάσει εμπειρογνομόνων

Όταν η πρόβλεψη βασίζεται στις προσδοκίες μιας ομάδας εμπειρογνομόνων, η αβεβαιότητα μπορεί να καθοριστεί με τη μορφή προσδοκιών των εμπειρογνομόνων σχετικά με τα υψηλά και χαμηλά όρια του στόχου, που αντιστοιχούν σε ένα συγκεκριμένο διάστημα πρόβλεψης πιθανότητας. Οι μέσοι όροι για κάθε όριο υποτίθεται στη συνέχεια ότι αντιστοιχούν στο καθορισμένο διάστημα πρόβλεψης μιας κανονικής κατανομής (Lutz *et al.*, 1996b, 1997- Lutz & Scherbov, 1998), παρέχοντας έτσι μια εκτίμηση του τυπικού σφάλματος. Η μέθοδος έχει αποδειχθεί ότι είναι ανθεκτική σε αλλαγές τόσο στην πιθανότητα όσο και στη μορφή της κατανομής (Lutz & Scherbov, 1998- Lutz *et al.*, 1999). Ωστόσο, έχει αμφισβητηθεί κατά πόσον οι εμπειρογνώμονες λαμβάνουν υπόψη τα όρια των ετήσιων σημειακών προβλέψεων ή των διαχρονικών μέσων όρων και κατά πόσον είναι σε θέση να προσδιορίσουν τιμές που αντιπροσωπεύουν τα όρια πιθανότητας με λογική ακρίβεια (Lee, 1999, σ. 187).

Στις εφαρμογές, η διαδρομή του δείγματος για τη θνησιμότητα επιλέγεται βάσει ενός τυχαίου αριθμού που καθορίζει το εκατοστημόριο της κατανομής στόχου- υποθέτοντας χρονική γραμμικότητα, αυτό το εκατοστημόριο καθορίζει τη διαδρομή του δείγματος, υποδηλώνοντας τέλεια αυτοσυσχέτιση. Η μέθοδος δεν είναι κατάλληλη για βραχυπρόθεσμες προβλέψεις, όπου η διακύμανση μπορεί να είναι μεγαλύτερη από ό,τι μπορεί να καταγραφεί από τις συνδυασμένες υποθέσεις της γραμμικότητας και της κανονικότητας (Lutz *et al.*, 1996b)- οι Alders & De Beer (1998, όπ. αναφ. στο De Beer & Alders, 1999, σ. 4)- δείχνουν ότι το

πρόβλημα αυτό δεν είναι αμελητέο και μακροπρόθεσμα. Η υπόθεση της γραμμικότητας και της τέλει αυτοσυσχέτισης χαλαρώθηκε αργότερα με την προσθήκη ενός τυχαίου όρου (Lutz *et al.*, 1999, 2001). Η κριτική ότι οι προσδοκίες των εμπειρογνομόνων είναι υπερβολικά σίγουρες και παράγουν υπερβολικά στενά διαστήματα πρόβλεψης, ιδίως όταν η πρόσφατη τάση είναι αμετάβλητη (Keilman, 1990), συνέβαλε στη μεγαλύτερη έμφαση στα επιχειρήματα των εμπειρογνομόνων αντί της γνώμης (Lutz *et al.*, 2004). Οι Lutz κ.ά. (1996a) υποστηρίζουν ότι η προσέγγιση με βάση τους εμπειρογνώμονες είναι ο μόνος ουσιαστικός τρόπος για την καταγραφή της μελλοντικής αβεβαιότητας.

Η εμπειρογνωμοσύνη χρησιμοποιήθηκε επίσης από τους Alders & De Beer (2005) για τον προσδιορισμό της αβεβαιότητας της πρόβλεψης θνησιμότητας που προέρχεται από ένα ντετερμινιστικό μοντέλο. Η προσέγγισή τους, η οποία περιγράφεται ως "πρόβλεψη βάσει επιχειρημάτων" (Alders & De Beer, 2005, σ. 57), περιελάμβανε τη χρήση κρίσης με βάση τέσσερις εκτιμήσεις: τάσεις καπνίσματος ανά φύλο, δυνατότητες μείωσης της θνησιμότητας στις νεαρές ηλικίες, συνεχιζόμενη ορθογωνοποίηση της καμπύλης επιβίωσης και ο ρόλος των ιατρικών εξελίξεων.

#### 4.9.5 Εκ των υστέρων εκτίμηση σφάλματος

Η τρίτη προσέγγιση για την εκτίμηση της αβεβαιότητας περιλαμβάνει τη χρήση των εμπειρικών σφαλμάτων των προηγούμενων προβλέψεων για την εκτίμηση της μεταβλητότητας των μελλοντικών προβλέψεων. Υποτίθεται ότι το μέλλον δεν θα διαφέρει από το παρελθόν και ότι τα σφάλματα είναι κανονικά κατανομημένα ή μπορούν να μοντελοποιηθούν. Η αβεβαιότητα κατά την περίοδο πρόβλεψης προκύπτει από τη συνολική εκτίμηση της διακύμανσης σύμφωνα με τον εκτιμώμενο (ή υποτιθέμενο) βαθμό αυτοσυσχέτισης. Η διακύμανση μπορεί να καθοριστεί για το τελευταίο έτος της πρόβλεψης, για παράδειγμα ως διάστημα πρόβλεψης-στόχος, και η παρεμβολή να επιτευχθεί για τα ενδιάμεσα έτη με την παραδοχή ενός μοντέλου τυχαίου περιπάτου. Εναλλακτικά, μπορούν να γίνουν εξωστρεφείς προβλέψεις των εκ των υστέρων σφαλμάτων (De Beer, 1997). Γενικά θεωρείται ότι υπάρχει τέλεια συσχέτιση μεταξύ των ηλικιών.

Τα εκ των υστέρων σφάλματα χρησιμοποιούνται για την εξαγωγή διαστημάτων πρόβλεψης για μια ξεχωριστά παραγόμενη (ντετερμινιστική) πρόβλεψη, με την οποία θα πρέπει να είναι συνεπή (Van Hoorn & De Beer, 2001). Η προσέγγιση είναι πλεονεκτική όταν προτιμάται μια υποκειμενική σημειακή πρόβλεψη, για παράδειγμα λόγω πρόσφατης διαρθρωτικής αλλαγής.

Ωστόσο, θα πρέπει να σημειωθεί ότι τα εκ των υστέρων σφάλματα για περιόδους διαρθρωτικών αλλαγών μπορεί να μην έχουν σημασία για μελλοντικές προβλέψεις (De Beer & Alders, 1999).

Λεπτομερείς μελέτες των σφαλμάτων στις επίσημες προβλέψεις θνησιμότητας έχουν χρησιμοποιήσει πολυμεταβλητές μεθόδους για να διαχωρίσουν τις επιδράσεις του μήκους του ορίζοντα πρόβλεψης, της δυσκολίας της συγκεκριμένης περιόδου και του έτους βάσης (Keilman, 1990; De Beer, 1997; Keilman, 1997). Οι πρόσθετες επιδράσεις περιλαμβάνουν την αλληλεπίδραση μεταξύ της δυσκολίας και της μεθόδου πρόβλεψης, τον ρόλο της τύχης και το σκεπτικό πίσω από την επιλογή της μεθόδου, καθώς και τις αιτίες του σφάλματος πρόβλεψης (για παράδειγμα, λανθασμένη εξειδίκευση του μοντέλου ή πραγματική αλλαγή της τάσης) (De Beer, 2000). Η μεροληψία προς τις παλαιότερες προβλέψεις, λόγω των μεγαλύτερων σειρών σφάλματος, πρέπει επίσης να λαμβάνεται υπόψη. Ο Keilman (1990, 1997) δεν βρήκε σαφείς ενδείξεις βελτίωσης με την πάροδο του χρόνου στην πρόβλεψη των ποσοστών θνησιμότητας.

## Κεφάλαιο 5<sup>ο</sup>: Μεθοδολογία

### 5.1 Προδιαγραφές του μοντέλου Lee-Carter

Εφόσον στο κεφάλαιο 2 της εργασίας έγινε μια εισαγωγή στο μοντέλο Lee-Carter, η παρούσα ενότητα θα ασχοληθεί με τις μαθηματικές παραδοχές που διέπουν το μοντέλο και με τον τρόπο που εφαρμόστηκε το μοντέλο στην μελέτη του ελληνικού πληθυσμού. Όπως αναφέρθηκε παραπάνω, το μοντέλο Lee-Carter υποθέτει επιδράσεις ηλικίας και περιόδου και προβάλλει τη θνησιμότητα βάσει αυτών με ένα στοιχείο προστιθέμενου θορύβου. Οι εισοδοί στο μοντέλο Lee-Carter είναι πίνακες παρατηρούμενων θανάτων σε έναν πληθυσμό και ένας πίνακας εκθέσεων του εν λόγω πληθυσμού. Οι εκθέσεις χρησιμεύουν για να μας πληροφορήσουν για την πιθανότητα ένα τμήμα του πληθυσμού να επιβιώσει μετακινούμενο από κάποιο έτος  $t$  στο  $t+1$ . Και για τους δύο πίνακες, η σειρά είναι τέτοια ώστε οι ηλικίες να παρατάσσονται ως γραμμές, ενώ οι στήλες ως έτη.

Το αρχικό μοντέλο χρησιμοποιούσε SVD για να μειώσει το μέγεθος του πίνακα θνησιμότητας για μια πιο αποτελεσματική διαδικασία. Μπορούμε να υποθέσουμε ότι τα αρχικά δεδομένα είναι ιδιαίτερα συσχετισμένα, και έτσι η εκτέλεση οποιουδήποτε είδους ανάλυσης είναι από μόνη της αρκετά δύσκολη. Η διαδικασία SVD, μειώνει τη διαστατικότητα, έτσι ώστε η περαιτέρω πρόβλεψη να γίνεται σχετικά εύκολη. Η διαδικασία εκτίμησης των παραμέτρων είναι καλά τεκμηριωμένη στους Lee και Carter (1992) και θα δοθεί μια εξήγηση των μεταβλητών που χρησιμοποιήθηκαν:

- **x**: υποδηλώνει τις ηλικιακές ομάδες στον πίνακα θνησιμότητας
- **t**: δηλώνει το έτος του πίνακα θνησιμότητας
- **$m_{x,t}$** : είναι ο κεντρικός ρυθμός θνησιμότητας που είναι ο λόγος των θανάτων προς τις εισερχόμενες εκθέσεις
- **$a_x$** : συμβολίζει τον ειδικό συντελεστή θνησιμότητας για την ηλικία
- **$b_x$** : υποδηλώνει τον δεύτερο ειδικό συντελεστή για την ηλικία που είναι ευαίσθητος στο  $\ln(m_{x,t})$
- **$k_t$** : είναι ο ειδικός για το χρόνο συντελεστής που δείχνει τη χρονική τάση της θνησιμότητας για το έτος  $t$
- **$\varepsilon_{x,t}$** : είναι ο συντελεστής θορύβου που εισάγει θόρυβο/αβεβαιότητα στο πλαίσιο

Ορισμένες παρατηρήσεις σχετικά με τις μεταβλητές του μοντέλου, το  $b_x$  δείχνει πώς τα ποσοστά θνησιμότητας μεταβάλλονται με ευαισθησία στο  $k_t \left( \frac{d \ln \left( \frac{m_{x,t}}{at} \right)}{dt} \right) = b_x dk/dt$  όπως διευκρινίζεται από τον Lee-Carter (1992). Όταν το  $k_t$  είναι γραμμικό, η θνησιμότητα σε κάθε ηλικιακή ομάδα μεταβάλλεται εκθετικά και τα αρνητικά ποσοστά θνησιμότητας δεν επιτρέπονται σε αυτό το μοντέλο. Κάθε ηλικιακό ποσοστό μηδενίζεται καθώς το  $k_t$  τείνει στο μείον άπειρο. Μέρος της εκτίμησης των παραμέτρων βασίζεται στην εκτίμηση του δείκτη  $k_t$  για κάθε χρόνο. Οι Lee και Carter προτείνουν ότι αυτό μπορεί να γίνει για να καθοριστεί ένα σύνολο κεντρικών ποσοστών θανάτου. Αυτό καθίσταται εφικτό με τη χρήση της ακόλουθης σχέσης:

$$D(t) = \sum [E(x, t)e^{a_x + b_x k_t}]$$

Κάνουν επίσης ειδική αναφορά στο γεγονός ότι αυτό δεν μπορεί να γίνει αναλυτικά, αλλά οι απαιτούμενες τιμές του  $k_t$  πρέπει να αναζητηθούν. Αφού συγκεντρωθούν αυτές οι τιμές, χρησιμεύουν ως βάση με την οποία μπορούν να προβλεφθούν τα ποσοστά θνησιμότητας.

## 5.2 Προσαρμογή του μοντέλου

Στο αρχικό προτεινόμενο μοντέλο Lee-Carter (1992) χρησιμοποιείται μια λύση ελαχίστων τετραγώνων στην παραπάνω εξίσωση για το μοντέλο:

$$\ln[m(x, t)] = a_x + b_x k_t + \varepsilon_{x,t}$$

Οι συγγραφείς θέτουν το ζήτημα της ταυτοποίησης με το μοντέλο. Ορισμένοι περιορισμοί πρέπει να επιβληθούν στο μοντέλο για ακριβή αποτελέσματα. Οι περιορισμοί που τέθηκαν ήταν οι εξής:

$$\sum b_x = 1$$

$$\sum k_t = 0$$

Παρεμπιπτόντως, αυτό δεν ισχύει μόνο για αυτό το μοντέλο. Πολλές επεκτάσεις αυτού του μοντέλου υποφέρουν επίσης από προβλήματα αναγνωρισιμότητας και πρέπει να τεθούν παρόμοιοι περιορισμοί. Η λύση των ελαχίστων τετραγώνων στην παραπάνω εξίσωση μπορεί να επιτευχθεί με τη SVD που εφαρμόζεται στον πίνακα θνησιμότητας. Ακολουθεί μια σύντομη



εισαγωγή στη διαδικασία SVD και η διαδικασία προσαρμογής δίνεται παρακάτω. Αναφορικά με τους περιορισμούς που εφαρμόζονται στο μοντέλο που περιγράφηκε παραπάνω, η SVD μπορεί να εφαρμοστεί στον πίνακα που δίνεται από τη σχέση  $Z_{x,t} = \ln(m_{x,t}) - \hat{a}_x$ .

Αυτό με τη σειρά του παράγει τους πίνακες  $UDV'$ , όπου  $U$  αντιπροσωπεύει τους συντελεστές ανά ηλικία,  $D$  είναι ένας διαγώνιος πίνακας που περιέχει τις μοναδικές τιμές και  $V'$  αντιπροσωπεύει την αντιμετάθεση του συντελεστή ανά χρόνο. Η ποσότητα  $\hat{a}_x$  αντιπροσωπεύει τον μέσο όρο του λογαρίθμου των ποσοστών θνησιμότητας, δηλαδή:

$$\hat{a} = \frac{1}{T} \sum_t \ln(m_{x,t})$$

Η παράμετρος  $\hat{b}_x$  μπορεί να προκύψει από τον δεξιό μοναδιαίο πίνακα  $U$  και αντιστοιχεί στην πρώτη στήλη του πίνακα, δηλαδή  $\hat{b} = U_{x1}$ . Ενώ το  $\hat{k}_t$  μπορεί να εκτιμηθεί από τον αριστερό singular πίνακα  $V$ , συγκεκριμένα δίνεται από το γινόμενο των πρώτων στηλών του διαγώνιου πίνακα  $D$  και  $V$ , δηλαδή  $\hat{k}_t = D_1 V_{t1}$ . Ένας σύνθετος πίνακας  $\hat{Z}_{x,t} = \hat{b}_x \hat{k}_t$  μπορεί στη συνέχεια να παραχθεί από το γινόμενο των αναφερόμενων παραμέτρων. Οι εκτιμήσεις για το λογάριθμο των ποσοστών θνησιμότητας μπορούν στη συνέχεια να υπολογιστούν ως εξής:

$$\ln(\hat{m}_{x,t}) = \hat{a}_x + \hat{Z}_{x,t} = \hat{a}_x + \hat{b}_x \hat{k}_t$$

### 5.3 Προβλέψεις του μοντέλου

Το μοντέλο Lee-Carter προσεγγίζει την πρόβλεψη των ποσοστών θνησιμότητας με τη χρήση του χρονικά-συγκεκριμένου δείκτη  $\hat{k}_t$ . Πιο συγκεκριμένα, προσεγγίζεται από μια οπτική γωνία στη μονομεταβλητή χρονική ανάλυση, με τις προβλέψεις να γίνονται με έναν τυχαίο περίπατο με παράμετρο διολίσθησης. Αυτό μπορεί να ενθυλακωθεί με την ακολουθία εξισώσεων που δίνεται παρακάτω:

$$\hat{k}_t = \hat{k}_{t-1} + \theta + \xi_t$$

$$\xi_t \sim N(0, \sigma^2)$$

όπου  $\theta$  είναι γνωστή ως παράμετρος ολίσθησης και  $\xi_t$  είναι η παράμετρος σφάλματος που εισάγει αβεβαιότητα στην πρόβλεψη, διαφορετικά η πρόβλεψη ακολουθεί ευθεία γραμμή. Η παράμετρος διολίσθησης μπορεί να εκτιμηθεί με τη μέγιστη πιθανότητα που εξαρτάται μόνο από τον πρώτο και τον τελευταίο όρο της. Η εξίσωση της δίνεται από τη σχέση:

$$\hat{\theta} = \frac{\hat{K}_T - \hat{K}_1}{T - 1}$$

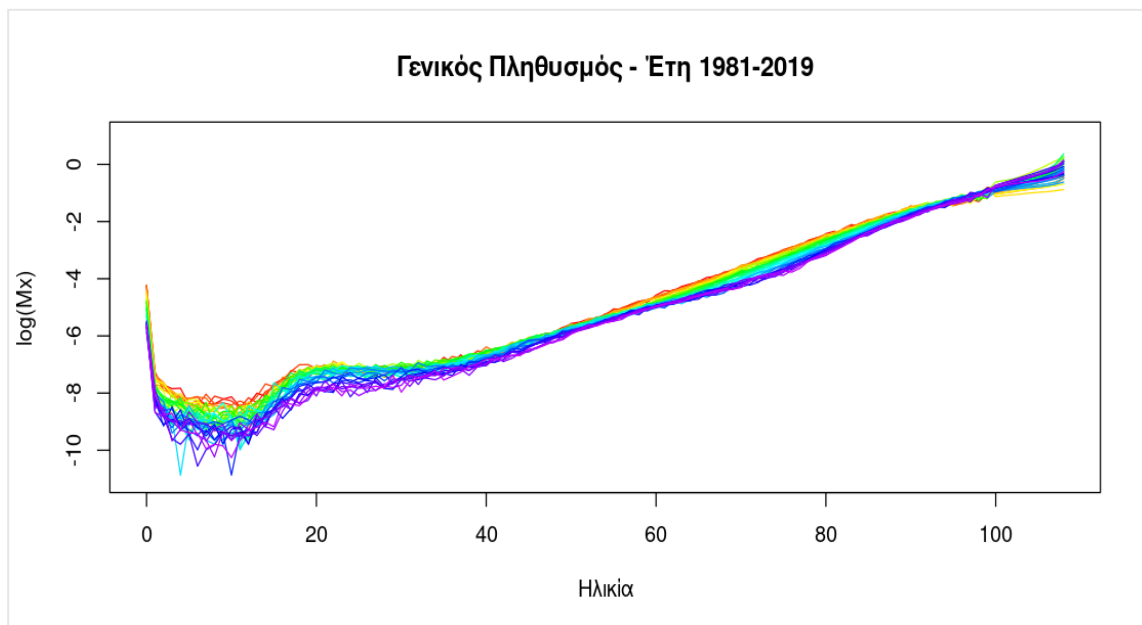
Αυτή η παράμετρος μπορεί στη συνέχεια να εισαχθεί ξανά στην εξίσωση για τον δείκτη θνησιμότητας, αντικαθιστώντας  $\theta$  με  $\hat{\theta}$  για να προκύψουν οι εκτιμήσεις των σημείων πρόβλεψης. Αυτή η διαδικασία που γίνεται επαναληπτικά μπορεί να δώσει μια σειρά από τιμές που μπορούν να βοηθήσουν στη διαδικασία πρόβλεψης. Τώρα, μια πιο γενική προσέγγιση πρόβλεψης για μια μελλοντική χρονική περίοδο  $T$ , η εκτίμηση της παραμέτρου ολίσθησης μπορεί να γραφεί ως εξής:

$$\hat{k}_{T+\Delta T} = \hat{K}_T + (\Delta t)\hat{\theta}$$

## Κεφάλαιο 6<sup>ο</sup>: Αποτελέσματα

Το μοντέλο Lee-Carter (1992) προσαρμόστηκε για την προσομοίωση κι εκτίμηση των δεικτών θνησιμότητας στην Ελλάδα. Στην ανάλυση χρησιμοποιήθηκαν τα δεδομένα από το Human Mortality Database (<https://www.mortality.org/>) για τους δείκτες θνησιμότητας ανά ηλικία (0-108) στην Ελλάδα για τα έτη 1981 έως και 2019, καθώς και τα πληθυσμιακά δεδομένα για τα αντίστοιχα έτη. Ο ορίζοντας προβλέψεων εκτάθηκε στα 20 χρόνια μέχρι το έτος 2039. Η ανάλυση πραγματοποιήθηκε στο στατιστικό περιβάλλον R version 4.2.2 και με χρήση των βιβλιοθηκών «demography» (Hyndman, 2023) και «StMoMo» (Villegas et al., 2018).

### 6.1 Εφαρμογή στο συνολικό πληθυσμό

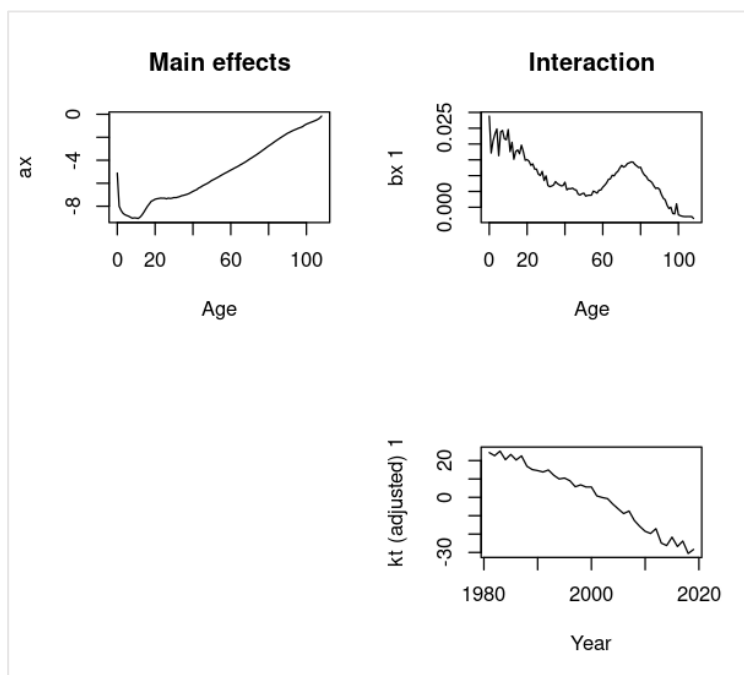


Διάγραμμα 11: Δεδομένα θνησιμότητας για τον γενικό πληθυσμό στην Ελλάδα (1981-2019)

Στο Διάγραμμα 11 απεικονίζονται τα δεδομένα θνησιμότητας για την Ελλάδα, με εναρκτήρια ημερομηνία το 1981 μέχρι και το 2019, τα δεδομένα αφορούν τις ηλικίες εύρους 0-108. Στο μοντέλο φαίνεται πως με το πέρασμα των ετών η τιμή του παρατηρούμενου δείκτη θνησιμότητας συνολικά μειώνεται, η παρατήρηση αυτή αφορά και την βρεφική θνησιμότητα. Για τα δεδομένα του γενικού πληθυσμού το ποσοστό μεταβλητότητας του μοντέλου ανέρχονταν στο 66,6%.

Πίνακας 4: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του συνολικού πληθυσμού (1981-2019)

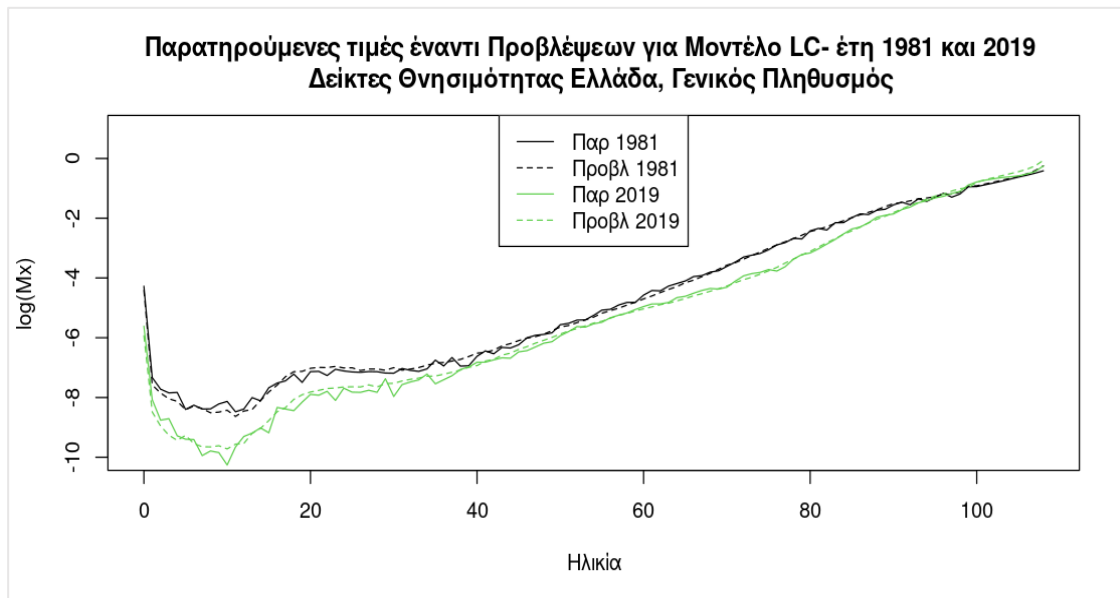
<b>Μέτρα σφάλματος στα ποσοστά θνησιμότητας του συνολικού πληθυσμού (1981-2019)</b>				
	ME*	MSE*	MPE*	MAPE*
Μέσοι όροι ηλικιών	-0.00081	0.00114	0.01304	0.09574
*ME=Mean Error, MSE=Mean Squared Error, MPE=Mean Percentage Error, MAPE=Mean Absolute Percentage Error				



Διάγραμμα 12: Απεικονίσεις συντελεστών  $a_x$ ,  $b_x$  και  $k_t$  για τον γενικό πληθυσμό

Στο Διάγραμμα 12 απεικονίζονται οι συντελεστές  $a_x$ ,  $b_x$  και  $k_t$ . Ο πρώτος δείχνει τον ειδικό συντελεστή μέσης θνησιμότητας ανάλογα με την ηλικία. Όπως φαίνεται στο Διάγραμμα ο συντελεστής ακολουθεί αυξητική τάση ανάλογα με την ηλικία μετά την ενηλικίωση. Ο συντελεστής  $b_x$ , είναι ο ειδικός συντελεστής που είναι ευαίσθητος στον κεντρικό ρυθμό θνησιμότητας και εκφράζει την απόκριση ηλικίας  $x$  κατά τη συνολική μεταβολή στα επίπεδα θνησιμότητας σε βάθος χρόνου. Μια υψηλή απόλυτη τιμή  $b_x$  ειδοποιεί ότι η αντίστοιχη

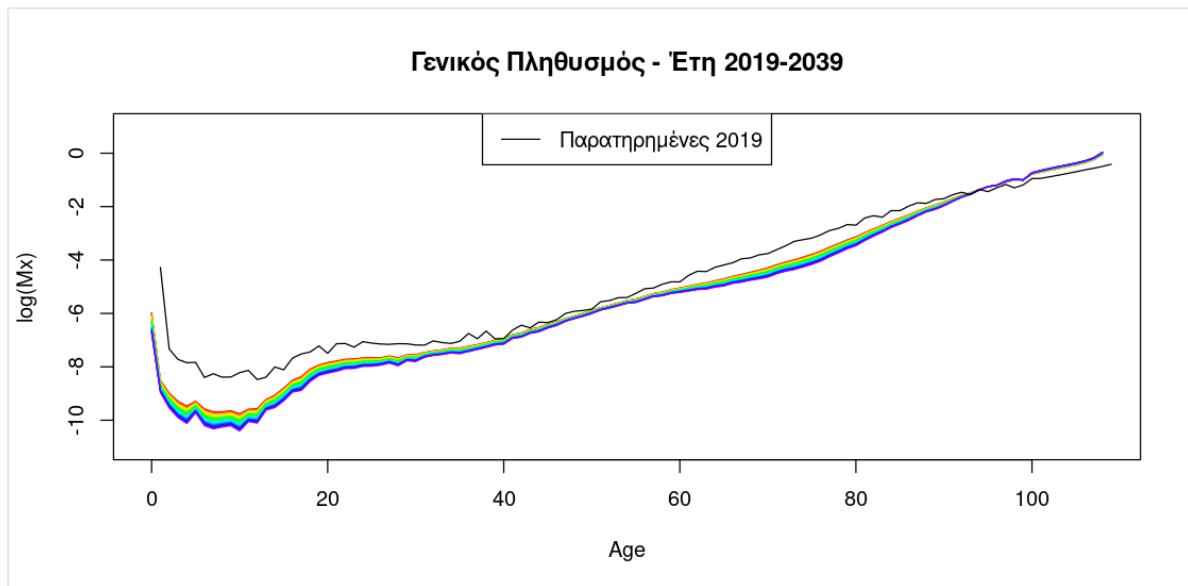
ηλικιακή ομάδα έχει υψηλή ευαισθησία σε μεταβολές των συνολικών επιπέδων θνησιμότητας. Από το συγκεκριμένο διάγραμμα φαίνεται πως οι μικρότερες ηλικίες εμφανίζουν μεγαλύτερες τιμές  $b_x$ . Τέλος ο ειδικός συντελεστής  $k_t$  δείχνει την χρονική τάση της θνησιμότητας, δηλαδή το γενικό επίπεδο της θνησιμότητας ανά τα έτη (1981-2019), η οποία έχει καθοδική πορεία για το σύνολο του πληθυσμού, ενώ οι τιμές της παραμέτρου είναι έντονα αρνητικές προς τα τέλη της περιόδου που μελετήθηκε.



Διάγραμμα 13: Παρατηρούμενες τιμές έναντι προβλέψεων στον συνολικό πληθυσμό για 1981 και 2019

Στο διάγραμμα 13 εμφανίζονται οι παρατηρούμενες τιμές του δείκτη θνησιμότητας σε σχέση με τις προβλέψεις του μοντέλου Lee-Carter, φαίνεται ότι εμφανίζονται σημαντικές αποκλίσεις μεταξύ των ηλικιών 0-30, ενώ στην συνέχεια υπάρχει μεγαλύτερη σύγκλιση μεταξύ των γραμμών.

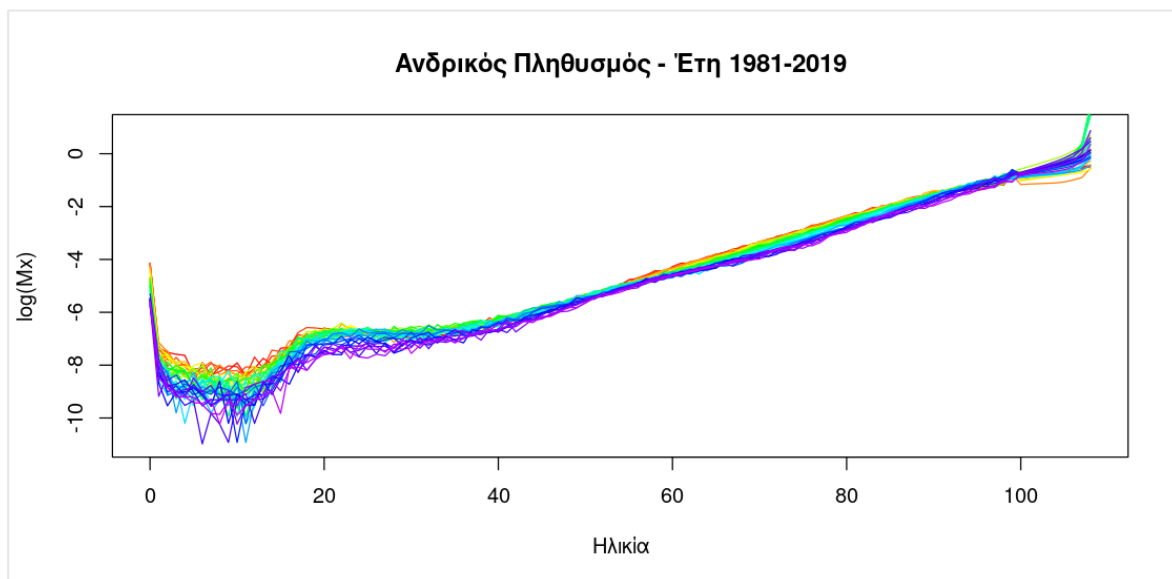
#### 6.1.1 Μελλοντικές Προβλέψεις (Συνολικός Πληθυσμός)



Διάγραμμα 14: Προβλέψεις για τον γενικό πληθυσμό (2019-2039)

Στο Διάγραμμα 14 απεικονίζονται οι μελλοντικές προβλέψεις του μοντέλου επί του συνολικού πληθυσμού για την περίοδο 2019-2039. Από το διάγραμμα φαίνεται πως αναμένεται μείωση του ρυθμού θνησιμότητας του Ελληνικού πληθυσμού, κυρίως στις ηλικίες μέχρι 20 ετών και , λιγότερο εκτεταμένες, στις ηλικίες 20-40 και 60-85 .

## 6.2 Εφαρμογή στον Ανδρικό πληθυσμό

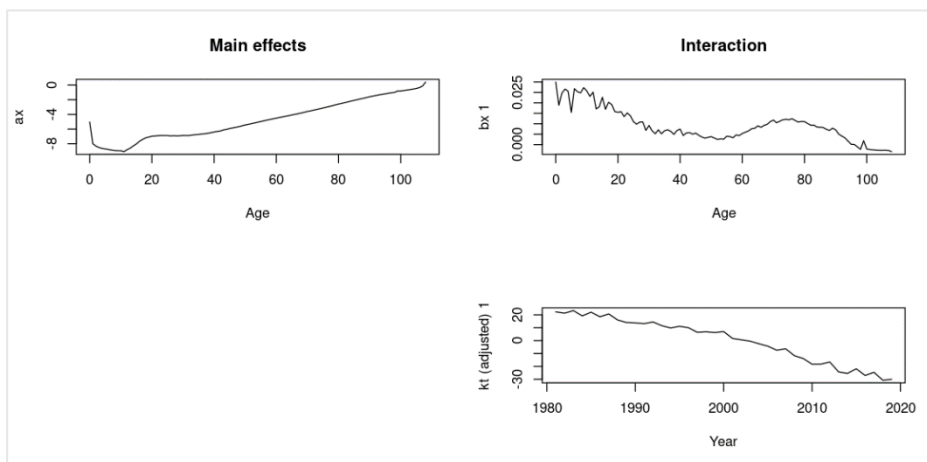


Διάγραμμα 15: Δεδομένα θνησιμότητας για τον αντρικό πληθυσμό στην Ελλάδα (1981-2019)

Στο παραπάνω διάγραμμα έχουμε την απεικόνιση του ρυθμού θνησιμότητας κατά ηλικία (0-108) στον ελληνικό πληθυσμό για τα έτη 1981-2019, φαίνεται από το διάγραμμα πως προοδευτικά με το πέρασ των ετών ο ρυθμός αυτός μειώνεται για όλες τις ηλικίες ανεξαιρέτως. Για τα δεδομένα του αντρικού πληθυσμού το ποσοστό μεταβλητότητας του μοντέλου ανέρχονταν στο 53,3%.

Πίνακας 5: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του αντρικού πληθυσμού (1981-2019)

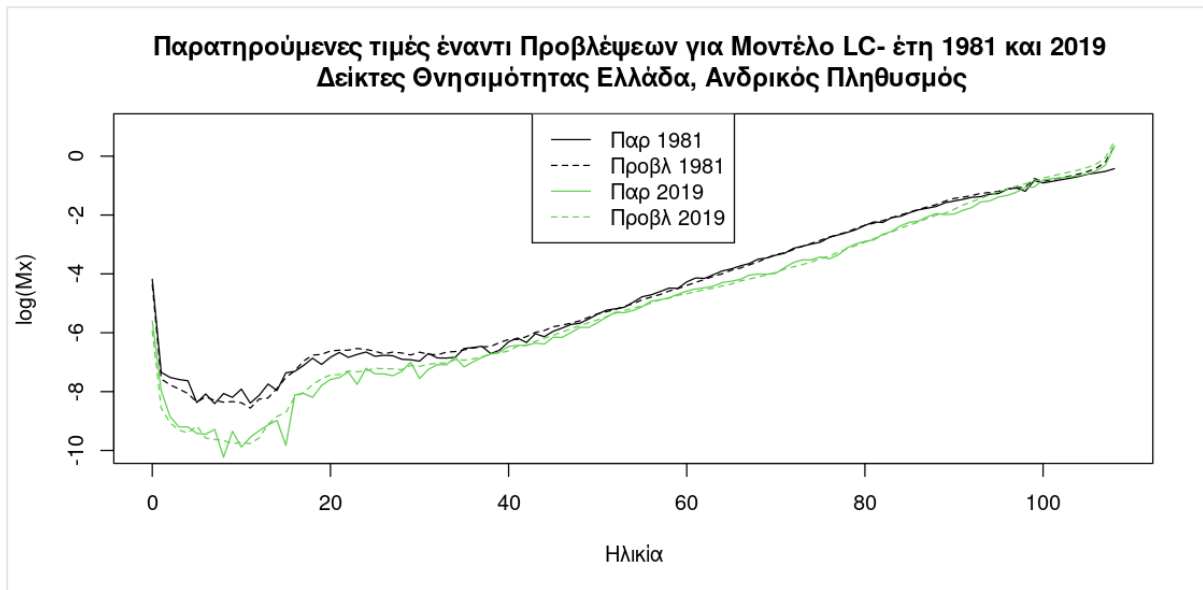
<b>Μέτρα σφάλματος στα ποσοστά θνησιμότητας του αντρικού πληθυσμού (1981-2019)</b>				
	ME*	MSE*	MPE*	MAPE*
Μέσοι όροι ηλικιών	-0.00534	0.02886	0.02064	0.11980
*ME=Mean Error, MSE=Mean Squared Error, MPE=Mean Percentage Error, MAPE=Mean Absolute Percentage Error				



Διάγραμμα 16: Απεικονίσεις συντελεστών  $a_x$ ,  $b_x$  και  $k_t$  για τον αντρικό πληθυσμό

Στο Διάγραμμα 16 απεικονίζονται ορισμένοι σημαντικοί συντελεστές που ( $a_x$ ,  $b_x$  και  $k_t$ ) που βοηθούν στην κατανόηση του μοντέλου. Συγκεκριμένα, ο πρώτος δείχνει τον ειδικό συντελεστή μέσης θνησιμότητας ανάλογα με την ηλικία, όπως φαίνεται στο διάγραμμα ο

συντελεστής ακολουθεί την κλασσική αυξητική πορεία μετά τα 20 χρόνια. Ο συντελεστής  $b_x$ , είναι ο ειδικός συντελεστής που εκφράζει την απόκλιση από την μέση θνησιμότητα, φαίνεται από το διάγραμμα πως οι αποκλίσεις είναι αρκετά μικρές και οι μεγαλύτερες τιμές σημειώνονται στις μικρότερες ηλικίες. Τέλος ο ειδικός συντελεστής  $k_i$ , δείχνει το γενικό επίπεδο της θνησιμότητας ανά τα έτη (1981-2019), η οποία έχει καθοδική πορεία.

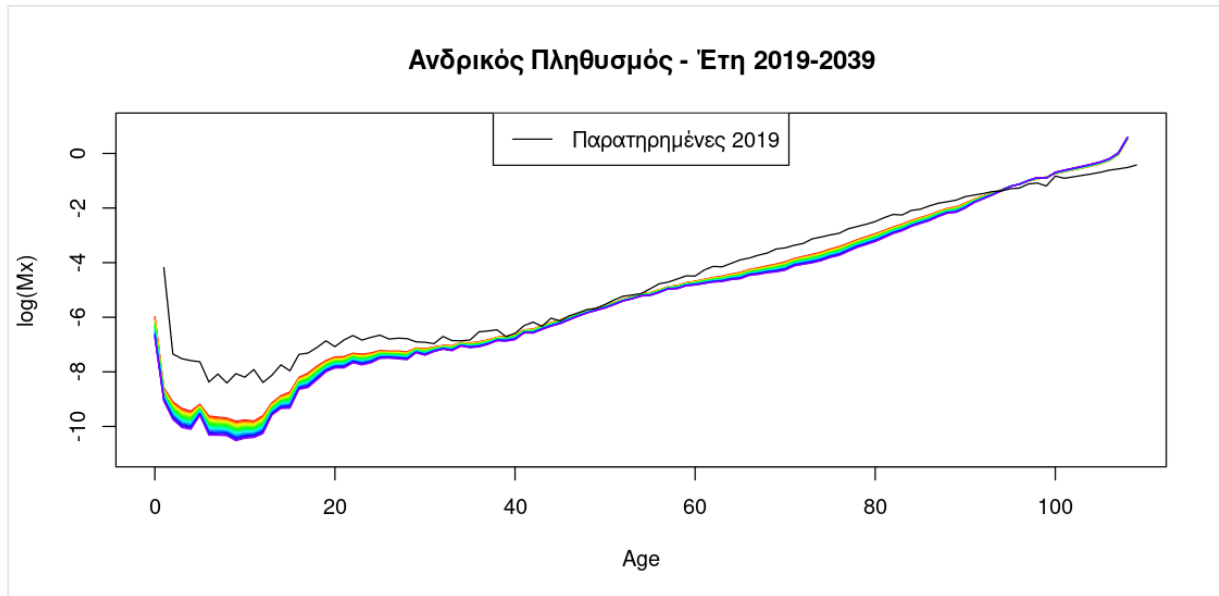


Διάγραμμα 17: Παρατηρούμενες τιμές έναντι προβλέψεων στον αντρικό πληθυσμό για 1981 και 2019

Στο διάγραμμα 17 εμφανίζονται οι παρατηρούμενες τιμές και οι προβλέψεις για δύο διαφορετικά έτη 1981 και 2019, σύμφωνα με το διάγραμμα συνολικά ο δείκτης θνησιμότητας είναι χαμηλότερος για το 2019 έναντι του 1981 και υπάρχει μεγάλη σύμπτωση μεταξύ προβλέψεων και πραγματικών τιμών.



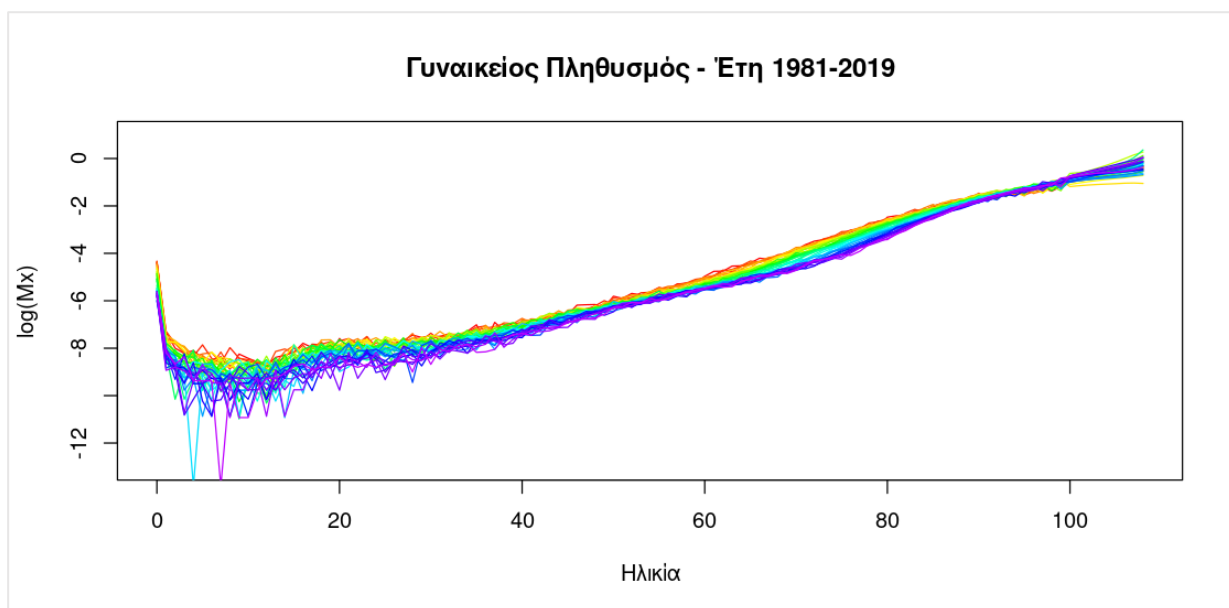
### 6.2.1 Μελλοντικές Προβλέψεις (Άντρες)



Διάγραμμα 18: Προβλέψεις μοντέλου για αντρικό πληθυσμό (2019-2039)

Το διάγραμμα 18, απεικονίζει τις προβλέψεις του μοντέλου Lee-Carter για την εικοσαετία 2019-2039 στον αντρικό πληθυσμό. Σύμφωνα με το διάγραμμα αναμένεται πτώση στον δείκτη θνησιμότητας για τις ηλικίες μέχρι και 40 ετών, πιο έντονη για τις ηλικίες μέχρι 15 ετών, καθώς και μια μείωση του δείκτη για τις ηλικίες 60-85. Για τις υπόλοιπες ηλικίες οι δείκτες φαίνονται να είναι πιο κοντά με αυτούς του 2019.

### 6.3 Εφαρμογή στο Γυναικείο πληθυσμό

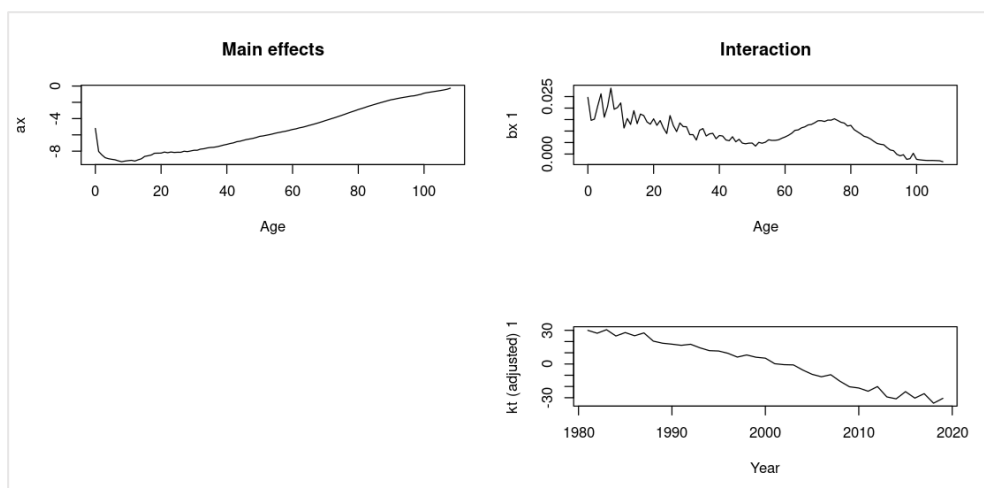


Διάγραμμα 19: Δεδομένα θνησιμότητας για τον γυναικείο πληθυσμό στην Ελλάδα (1981-2019)

Στο παραπάνω διάγραμμα έχουμε την απεικόνιση του ρυθμού θνησιμότητας κατά ηλικία (0-108) στον ελληνικό γυναικείο πληθυσμό για τα έτη 1981-2019. Γενικότερα φαίνεται πως με την πάροδο των χρόνων ο δείκτης μειώνεται, στο διάγραμμα εμφανίζονται ορισμένες απότομες διακυμάνσεις στις μικρές ηλικίες. Για τα δεδομένα του γυναικείου πληθυσμού το ποσοστό μεταβλητότητας του μοντέλου ανέρχονταν στο 50,4%.

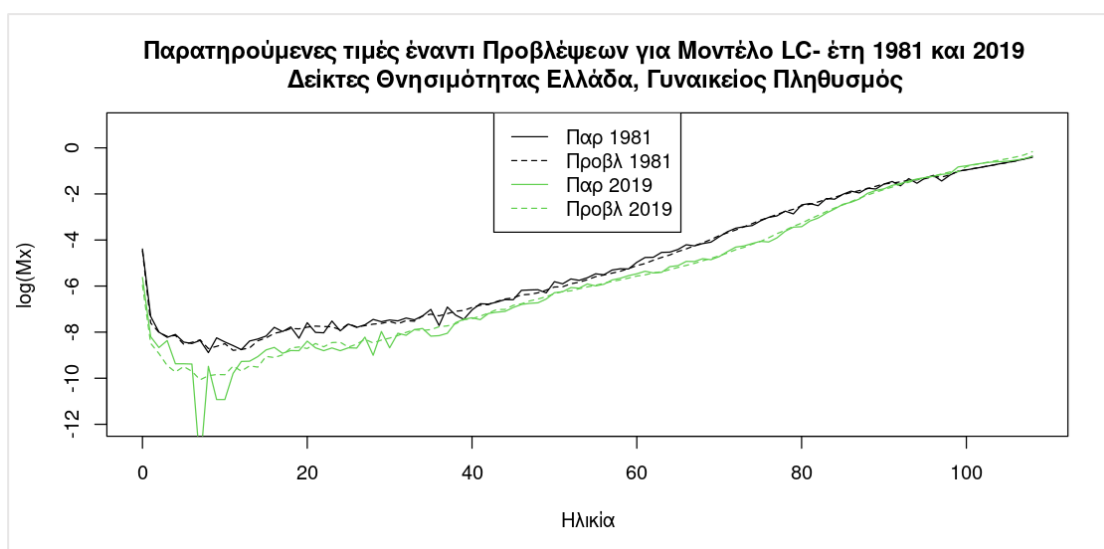
Πίνακας 6: Μέτρα σφάλματος στα ποσοστά θνησιμότητας του γυναικείου πληθυσμού (1981-2019)

<b>Μέτρα σφάλματος στα ποσοστά θνησιμότητας του γυναικείου πληθυσμού (1981-2019)</b>				
	ME*	MSE*	MPE*	MAPE*
Μέσοι όροι ηλικιών	-0.00085	0.00112	0.05973	0.17225
*ME=Mean Error, MSE=Mean Squared Error, MPE=Mean Percentage Error, MAPE=Mean Absolute Percentage Error				



Διάγραμμα 20: Απεικονίσεις συντελεστών  $a_x$ ,  $b_x$  και  $k_t$  για τον γυναικείο πληθυσμό

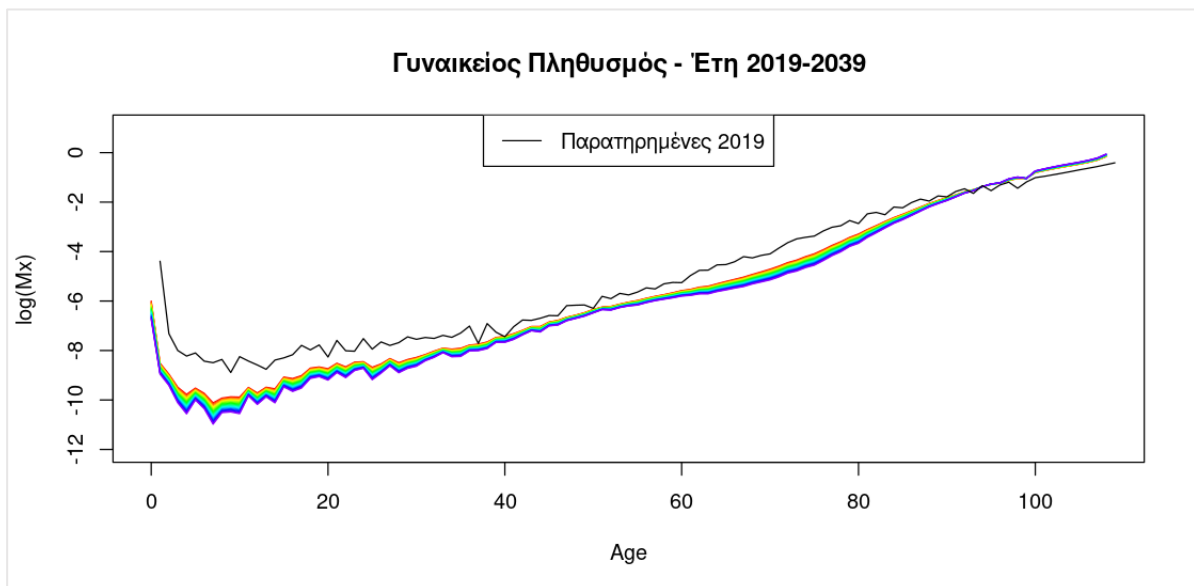
Στο Διάγραμμα 20 απεικονίζονται οι συντελεστές που  $a_x$ ,  $b_x$  και  $k_t$ . Ο  $a_x$  εκφράζει την μέση θνησιμότητα του γυναικείου πληθυσμού, όπου παρατηρείται μια αρκετά ομαλή ανιούσα πορεία με την αύξηση των ηλικιών. Ο συντελεστής  $b_x$ , εκφράζει την απόκλιση από την μέση θνησιμότητα και φαίνεται από το διάγραμμα πως παρατηρούνται έντονες διακυμάνσεις στις αποκλίσεις για τις ηλικίες 0-30. Τέλος ο ειδικός συντελεστής  $k_t$ , δείχνει το γενικό επίπεδο της θνησιμότητας ανά τα έτη για τον γυναικείο πληθυσμό, όπου αποτυπώνεται σημαντική μείωση με το πέρασ των ετών.



Διάγραμμα 21: Παρατηρούμενες τιμές έναντι προβλέψεων στον αντρικό πληθυσμό για 1981 και 2019

Στο διάγραμμα 21 εμφανίζονται οι παρατηρούμενες τιμές και οι προβλέψεις του μοντέλου L-C για δύο διαφορετικά έτη 1981 και 2019. Σύμφωνα με το διάγραμμα συνολικά ο δείκτης θνησιμότητας είναι χαμηλότερος για το 2019 έναντι του 1981 αν και η απόκλιση είναι μικρή. Αυτό που παρουσιάζει ενδιαφέρον είναι πως για το έτος 1981, οι παρατηρούμενες τιμές και οι προβλέψεις σχεδόν ταυτίζονται, αλλά για το έτος 2019 ενώ υπάρχει συμπίεση των δύο δεικτών, το μοντέλο L-C εμφανίζει διακυμάνσεις στις μικρές ηλικίες.

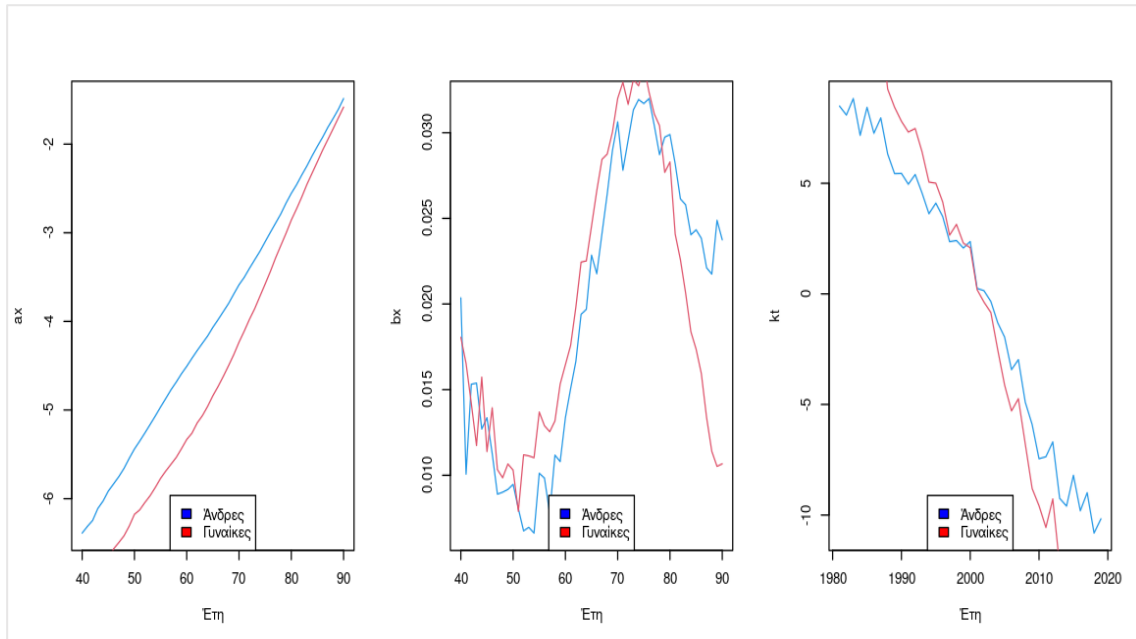
### 6.3.1 Μελλοντικές Προβλέψεις (Γυναίκες)



Διάγραμμα 22: Προβλέψεις μοντέλου για γυναικείο πληθυσμό (2019-2039)

Το διάγραμμα 22, απεικονίζει τις προβλέψεις του μοντέλου Lee-Carter για την εικοσαετία 2019-2039 στον γυναικείο πληθυσμό. Σύμφωνα με το διάγραμμα αναμένεται πτώση στον δείκτη θνησιμότητας ιδιαίτερα στις ηλικίες 0-40, πιο έντονες στον παιδικό πληθυσμό (ηλικίες 0-14). Επίσης, κάποιες μειώσεις στα επίπεδα θνησιμότητας διαφαίνονται και για τις υπόλοιπες ηλικίες μέχρι τα 85 έτη.

### 6.4 Δείκτες $a_x, b_x$ και $k_t$ βάσει φύλου



Διάγραμμα 23: Δείκτες  $a_x$ ,  $b_x$  και  $k_t$  βάσει φύλου για ηλικίες 40-90

Τα παραπάνω γραφήματα παρουσιάζουν ενδεικτικά τις σχετικές με την ηλικία παραμέτρους του μοντέλου Lee-Carter  $a_x$  και  $b_x$  και τη σχετική με το χρόνο παράμετρο  $k_t$  για τον ανδρικό και το γυναικείο πληθυσμό τις Ελλάδας, για τις ηλικίες 40-90. Είναι σαφές -ιδιαίτερα από το πρώτο γράφημα  $a_x$ , πως οι δείκτες θνησιμότητας για τους άνδρες είναι υψηλότεροι. Η παράμετρος  $a_x$  που μπορεί να ερμηνευτεί ως η μέση θνησιμότητα σε μια δεδομένη ηλικιακή ομάδα, βρίσκεται εμφανώς σε υψηλότερα επίπεδα για τους άντρες αν και φαίνεται πως σε μεγαλύτερες ηλικίες τείνει να υπάρχει εξισορρόπηση. Στο τρίτο γράφημα παρατηρούμε τη σχετική με το χρόνο παράμετρο  $k_t$ , που εκφράζει την τάση (trend) του δείκτη θνησιμότητας. Παρατηρούμε πως τα τελευταία χρόνια οι δείκτες θνησιμότητας στον πληθυσμό (και άνδρες και γυναίκες), παρουσιάζουν μείωση και ειδικά η γυναικείοι δείκτες, με περισσότερο απότομη κλίση.

## Κεφάλαιο 7<sup>ο</sup>: Συμπεράσματα

Η παρούσα εργασία είχε ως κεντρικό θέμα την πραγμάτευση των στοχαστικών μοντέλων θνησιμότητας και την εφαρμογή τους στον ελληνικό πληθυσμό. Το πρώτο μέρος της εργασίας είχε θεωρητικό χαρακτήρα και έλαβε χώρα η πραγμάτευση των βασικών εννοιών που είναι απαραίτητες για τις διαδικασίες πρόβλεψης, δηλαδή ο ορισμός τη θνησιμότητας, των πινάκων επιβίωσης αλλά και οι σύγχρονες τάσεις σε Ευρωπαϊκό επίπεδο. Στην συνέχεια συντελέστηκε η ανάλυση του πιο βασικού στοχαστικού μοντέλου πρόβλεψης θνησιμότητας, του μοντέλου Lee-Carter, ενώ συμπεριλήφθηκε και η διορθωτική μοντελοποίηση Poisson αλλά και ορισμένες κριτικές παρατηρήσεις σχετικά με το μοντέλο. Το επόμενο κεφάλαιο αφορούσε λοιπά μοντέλα που χρησιμοποιούνται για την πρόβλεψη της θνησιμότητας και μια σύγκριση των χαρακτηριστικών τους.

Στο δεύτερο μέρος της εργασίας έλαβε χώρα η εμπειρική μελέτη που αφορούσε τον Ελληνικό πληθυσμό. Για την διενέργεια της μελέτης επιλέχθηκε να εφαρμοστεί το μοντέλο Lee-Carter καθώς αξιολογείται ως αρκετά αξιόπιστο στις εφαρμογές του από την βιβλιογραφία. Στην εφαρμογή του μοντέλου έγινε χρήση δεδομένων ηλικιακού εύρους (0-108) για τα έτη 1981 μέχρι και το 2019 τα οποία αντλήθηκαν από το Human Mortality Database. Αν και η δύναμη πρόβλεψης του μοντέλου γενικά αξιολογείται ως υψηλή στην περίπτωση της μελέτης το ποσοστό μεταβλητότητας επί του γενικού πληθυσμού ανήλθε στο 66.6%, ενώ τα ποσοστά για άντρες και γυναίκες ήταν 53.3% και 50.4%, αντίστοιχα. Η δύναμη της πρόβλεψης για τον γενικό πληθυσμό μπορεί να χαρακτηριστεί ως μέτρια και πιθανώς η τιμή αυτή να καθορίστηκε γενικά από τις ελλείψεις του πίνακα δεδομένων για την χώρα μας. Τα ποσοστά για άντρες και γυναίκες, όπως σημειώθηκε και παραπάνω ήταν αρκετά χαμηλά. Από τα παραπάνω και από το γεγονός πως το μοντέλο εξηγεί μόνο το 66% της μεταβλητότητας των δεδομένων εγείρονται ορισμένα ζητήματα εμπιστοσύνης για το μοντέλο.

Μια άλλη παράμετρος αξιολόγησης της δύναμης του μοντέλου σχετίζεται με τα μέτρα σφάλματος που ανέκυψαν κατά την ανάλυση, αξίζει να σταθούμε στους δείκτες MSE και MAPE, ο πρώτος εκφράζει το μέσο τετράγωνο της διαφοράς ανάμεσα στα πραγματικά επίπεδα του δείκτη θνησιμότητας με τα αντίστοιχα προβαλλόμενα και ο δεύτερος την ακρίβεια της πρόβλεψης του μοντέλου και για τους δύο δείκτες όσο μικρότερες οι τιμές τόσο καλύτερη η πρόβλεψη. Στην ανάλυση, οι τιμές που προέκυψαν ήταν μικρές επί του γενικού πληθυσμού (MSE=0.001 και MAPE=0.09), αν και θα μπορούσαν να είναι μικρότερες, λόγω του ότι πρόκειται για πρόβλεψη θνησιμότητας.

Επιπρόσθετα εκτιμήθηκαν οι συντελεστές ( $a_x, b_x$  και  $k_t$ ), του μονέλου. Η μείωση του συντελεστή  $k_t$  δηλώνει μια καθαρή τάση μείωσης του γενικού επιπέδου θνησιμότητας στον γενικό πληθυσμό με τα χρόνια η οποία εμφανίζει μια σχετική γραμμικότητα. Όσον αφορά τις αντίστοιχες τιμές του συντελεστή σε αντρικό και γυναικείο πληθυσμό η τάση είναι και σε αυτές τις περιπτώσεις καθοδική, με εξαίρεση βέβαια πως οι τιμές του δείκτη για τις γυναίκες εκκινούν από υψηλότερο σημείο. Συνολικά βέβαια οι δείκτες θνησιμότητας για τους άντρες ( $a_x$ ) ήταν υψηλότεροι σε σχέση με αυτούς των γυναικών, αλλά με την αύξηση της ηλικίας η διαφοράς αυτές συνέκλιναν. Επίσης για τον ειδικό συντελεστή  $b_x$  που είναι ευαίσθητος στον κεντρικό ρυθμό θνησιμότητας φαίνεται πως η γυναίκες σημειώνουν μεγαλύτερη ευαισθησία στις μεταβολές. Σε επίπεδο μελλοντικών προβλέψεων που στην συγκεκριμένη δοκιμή έγιναν σε βάθος εικοσαετίας (2019-2039), στον γενικό πληθυσμό, στους άντρες και στις γυναίκες αναμένεται πτώση του ρυθμού θνησιμότητας για τα επόμενα χρόνια.

Κλείνοντας, αν και γενικά η ανασκόπηση της βιβλιογραφίας δείχνει την ανωτερότητα του Lee-Carter, στην συγκεκριμένη εμπειρική έρευνα στον Ελληνικό πληθυσμό, το μοντέλο εμφάνισε μέτρια εξηγητική ισχύ της μεταβλητότητας. Όπως σημειώθηκε παραπάνω πιθανότατα αυτό να οφείλεται σε ελλείψεις και ακολουθίες στις τιμές των δεδομένων. Σε επίπεδο μελλοντικής πρότασης, πιθανώς η χρήση μιας διαφορετικής μοντελοποίησης ή ακόμα και κάποιας διόρθωσης, όπως η Poisson θα μπορούσε να δώσει πιο ικανοποιητικά αποτελέσματα.

## Βιβλιογραφία

- Agresti, A., and B. Finlay. (1997). *Statistical Methods for the Social Sciences*. Upper Saddle River NJ: Prentice Hall.
- Ahlburg, D.A. & Vaupel, J.W. (1990). Alternative projections of the U.S. population. *Demography*, 27(4), 639-652.
- Alders, M. & De Beer, J. (2005). An expert knowledge approach to stochastic mortality forecasting in the Netherlands. In (N. Keilman, ed.) *Perspectives on mortality forecasting* (Vol. II. Probabilistic models, pp39-64). Swedish Social Insurance Agency, Stockholm.
- Alderson, M. & Ashwood, F. (1985). Projection of mortality rates for the elderly. *Population Trends*, 42, 22-29.
- Alexander, M., Zagheni, E., & Barbieri, M. (2017). A flexible Bayesian model for estimating subnational mortality. *Demography*, 54(6), 2025–2041.
- Alho, J.M. & Spencer, B.D. (1985). Uncertain population forecasting. *Journal of the American Statistical Association*, 80, 306-314.
- Alho, J.M. & Spencer, B.D. (1990). Error models for official mortality forecasts. *Journal of the American Statistical Association*, 85(411), 609-616.
- Alho, J.M. & Spencer, B.D. (2005). *Statistical Demography and Forecasting*. Springer, New York.
- Alho, J.M. (1991). Effect of aggregation on the estimation of trend mortality. *Mathematical Population Studies*, 2, 53-67.
- Alho, J.M. (1992). Estimating the strength of expert judgement: the case of U.S. mortality forecasts. *Journal of Forecasting*, 11, 157-167.
- Alho, J.M. (1998). A stochastic forecast of the population of Finland (Reviews No. 1998/4). *Statistics Finland*, Helsinki.
- Alho, J.M. (2003). Experiences from forecasting mortality in Finland. In (T. Bengtsson & N. Keilman, eds.) *Perspectives on mortality forecasting* (Vol. I. Current practice, 29-40). Swedish Social Insurance Agency, Stockholm. 34 Mortality Modelling and Forecasting: a Review of Methods



- Alho, J.M., 2000. Discussion of Lee (2000). *North American Actuarial Journal* 4, 91–93
- Alho, J.M., Alders, M., Cruijsen, H., Keilman, N., Nikander, T. & Pham, D.Q. (2006). New forecast: population decline postponed in Europe. *Statistical Journal for the United Nations Economic Commission for Europe*, 23(1), 1-10.
- Antonio, K., Bardoutsos, A., & Ouburg, W. (2015). Bayesian Poisson log-bilinear models for mortality projections with multiple populations. *European Actuarial Journal*, 5(2), 245–281.
- Babel, B., Bomsdorf, E. & Schmidt, R. (2008). Forecasting German mortality using panel data procedures. *Journal of Population Economics*, 21(3), 541-555.
- Bell, W. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, 13(3), 279-303.
- Bell, W.R. & Monsell, B.C. (1991). *Using principal components in time series modeling and forecasting of age-specific mortality rates*. Paper presented at the American Statistical Association 1991 Proceedings of the Social Statistics Section.
- Benjamin, B. & Pollard, J.H. (1980). The analysis of mortality and other actuarial statistics. Heinemann, London. Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (2004). Annual report. Washington DC.
- Bergeron-Boucher, M., Canudas-Romo, V., Oeppen, J., & Vaupel, J. W. (2017). Coherent forecasts of mortality with compositional data analysis. *Demographic Research*, 37(17), 527–566.
- Bohk-Ewald, C., & Rau, R. (2017). Probabilistic mortality forecasting with varying age-specific survival improvements. *Genus*, 73(1), 1.
- Bohk-Ewald, C., Ebeling, M., & Rau, R. (2017). Lifespan disparity as an additional indicator for evaluating mortality forecasts. *Demography*, 54(4), 1559–1577.
- Bongaarts, J. & Feeney, G. (2002). How long do we live? *Population and Development Review*, 28(1), 13-29.
- Bongaarts, J. & Feeney, G. (2003). Estimating mean lifetime. *Proceedings of the National Academy of Sciences*, 100(23), 13127-13133.
- Bongaarts, J. & Feeney, G. (2006). The quantum and tempo of life-cycle events. *Vienna Yearbook of Population Research* 2006, 115-151.

- Bongaarts, J. (2006). How long will we live? *Population and Development Review*, 32(4), 605–628.
- Booth, H. (2004). On the importance of being uncertain: forecasting population futures for Australia. *People and Place*, 12(2), 1-12.
- Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22, 547-581.
- Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: a review of methods. *Annals of Actuarial Science*, 3(1–2), 3–43.
- Booth, H., Hyndman, R.J., Tickle, L. & De Jong, P. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research*, 15(9), 289-310.
- Booth, H., Maindonald, J. & Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56(3), 325-336. Mortality Modelling and Forecasting: a Review of Methods 35
- Booth, H., Tickle, L. & Smith, L. (2005). Evaluation of the variants of the Lee-Carter method of forecasting mortality: a multi-country comparison. In (A. Dharmalingam & I. Pool, eds.) *New Zealand Population Review, Special Issue on Stochastic Population Projections*, 31(1), 13-37.
- Box, G., Jenkins, G.M. & Reinsel, G.C. (1994). *Time series analysis: forecasting and control* (3rd ed.). Prentice Hall, Englewood, NJ.
- Brass, W. (1971). On the scale of mortality. In (W. Brass, ed.) *Biological aspects of demography* (69-110). Taylor and Francis, London.
- Brass, W. (1974). Perspectives in population prediction: illustrated by the statistics of England and Wales. *Journal of the Royal Statistical Society, Series A*, 137(4), 532-583.
- Brillinger, D.R. (1986). The natural variability of vital rates and associated statistics. *Biometrics*, 42(December), 693-734.
- Brillinger, D.R., 1986. The natural variability of vital rates and associated statistics. *Biometrics* 42, 693–73

- Brouhns, N., Denuit, M. & Van Keilegom, I. (2005). Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, 3, 212-224.
- Brouhns, N., Denuit, M. & Vermunt, J.K. (2002). A Poisson log-bilinear regression approach to the construction of projected life-tables. *Insurance: Mathematics and Economics*, 31, 373-393.
- Brouhns, N., Denuit, M., (2001). Risque de longévité et rentes viagères. II. Tables de mortalité prospectives pour la population belge. *Discussion Paper*. Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgique.
- Buettner, T. & Zlotnik, H. (2005). Prospects for increasing longevity as assessed by the United Nations. *Genus*, LXI(1), 213-233.
- Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and *calibration*. *Journal of Risk and Insurance*, 73(4), 687–718.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., & Khalaf-Allah, M. (2011b). Bayesian stochastic mortality modelling for two populations. *Astin Bulletin*, 41(1), 29–59.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., & Khalaf-Allah, M. (2011a). Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3), 355–367.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., & Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1), 1–35.
- Cairns, A.J.G. (2000). A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics*, 27, 313-330.
- Cairns, A.J.G., Kallestrup-Lamb, M., Rosenskjold, C.P.T., Blake, D., & Dowd, K. (2016) Modelling socio-economic differences in the mortality of Danish males using a new affluence index. *CREATES Research Paper* 2016–14.
- Carone, G., Eckefeldt, P., Giamboni, L., Laine, V., & Pamies, S. (2016). Pension reforms in the EU since the early 2000's: achievements and challenges ahead. *European Economy Discussion*

*Papers* 42. Brussels: Directorate-General for Economic and Financial Affairs (European Commission).

Carriere, J.F. (1992). Parametric models for life tables. *Transactions of the Society of Actuaries*, 44, 77-99.

Carter, L.R. & Prskawetz, A. (2001). Examining structural shifts in mortality using the Lee-Carter method (*MPIDR Working Paper No. WP 2001-007*). Max Planck Institute for Demographic Research.

Carter, L.R. & Prskawetz, A. (2001). Examining structural shifts in mortality using the Lee\_Carter method (*MPIDR Working Paper No. WP 2001-007*). Max Planck Institute for Demographic Research

Caselli, G. (1996). Future longevity among the elderly. In (G. Caselli & A. Lopez, eds.) *Health and mortality among elderly populations*. Clarendon Press, Oxford.

Chiang, C. (1984). *The Life Table and its Applications*. Malabar (FL). Robert E. Krieger Publ. Co

Coale, A.J. & Kisker, E.E. (1990). Defects in data on old-age mortality in the United States: new procedures for calculating schedules and life tables at the highest ages. *Asian and Pacific Population Forum*, 4(1), 1-31.

Coale, A.J. & McNeil, D.R. (1972). The distribution by age of the frequency of first marriage in a female cohort. *Journal of the American Statistical Association*, 67(340), 743- 749.

Coelho, E., & Nunes, L. C. (2011). Forecasting mortality in the event of a structural change. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3), 713–736.

Congdon, P. (1993). Statistical graduation in local demographic analysis and projection. *Journal of the Royal Statistical Society*, 156(2), 237-270.

Continuous Mortality Investigation Bureau (1990). Standard tables of mortality based on the 1979-1982 experiences, *CMI report no. 10*. London: Institute of Actuaries and Faculty of Actuaries.

Continuous Mortality Investigation Bureau (1999). Standard tables of mortality based on the 1991-1994 experiences, *CMI report no. 17*. London: Institute of Actuaries and Faculty of Actuaries.

Continuous Mortality Investigation Bureau (2002). An interim basis for adjusting the “92” series mortality projections for cohort effects, *CMI Working paper no. 1*. London: Institute of Actuaries and Faculty of Actuaries.

Continuous Mortality Investigation Bureau (2004). Projecting future mortality: a discussion paper, *CMI Working paper no. 3*. London: Institute of Actuaries and Faculty of Actuaries. 36  
Mortality Modelling and Forecasting: a Review of Methods

Continuous Mortality Investigation Bureau (2005). Projecting future mortality: towards a proposal for a stochastic methodology, *CMI Working paper no. 15*. London: Institute of Actuaries and Faculty of Actuaries.

Continuous Mortality Investigation Bureau (2006). Stochastic projection methodologies: further progress and p-spline model features, example results and implications, *CMI Working paper no. 20*. London: Institute of Actuaries and Faculty of Actuaries.

Continuous Mortality Investigation Bureau (2007). Stochastic projection methodologies: Lee-Carter model features, example results and implications, *CMI Working paper no. 25*. London: Institute of Actuaries and Faculty of Actuaries.

Crimmins, E.M. (1981). The changing pattern of American mortality decline, 1947-1977, and its implication for the future. *Population and Development Review*, 7(2), 229-254.

Currie, D., Durban, M. & Eilers, P.H.C. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, 4, 279-298.

Czado, C., Delwarde, A. & Denuit, M. (2005). Bayesian Poisson logbilinear mortality projections. *Insurance: Mathematics and Economics*, 36, 260-284.

De Beer, J. & Alders, M. (1999). Probabilistic population and household forecasts for the Netherlands. *Working paper no. 45*. Paper presented at the Joint ECE-Eurostat Work Session on Demographic Projections, 3-7 May 1999, Perugia, Italy.

De Beer, J. (1989). Projecting age-specific fertility rates by using time-series methods. *European Journal of Population*, 5(1989), 315-346.

De Beer, J. (1997). The effect of uncertainty of migration on national population forecasts: the case of the Netherlands. *Journal of Official Statistics*, 13(3), 227-243.

- De Beer, J. (2000). *Dealing with uncertainty in population forecasting*. Statistics Netherlands, Department of Population, Voorburg.
- De Jong, P. & Marshall, C. (2007). Forecasting mortality using the Wang transform. *ASTIN Bulletin*, 37(1), 149-162.
- De Jong, P. & Marshall, C. (2007). Forecasting mortality using the Wang transform. *ASTIN Bulletin*, 37(1), 149-162.
- De Jong, P. & Tickle, L. (2006). Extending Lee-Carter mortality forecasting *Mathematical Population Studies*, 13(1), 1-18.
- Denton, F., Feaver, C. & Spencer, B. (2005). Time series analysis and stochastic forecasting: an econometric study of mortality and life expectancy. *Journal of Population Economics*, 18(2), 203-227.
- Enchev, V., Kleinow, T., & Cairns, A. J. (2017). Multi-population mortality models: fitting, forecasting and comparisons. *Scandinavian Actuarial Journal*, 2017(4), 319–342
- European Commission. (2009). *Dealing with the impact of an ageing population in the EU. 2009 Ageing Report*. Brussels: European Commission.
- Ewbank, D.C., Gomez de Leon, J.C. & Stoto, M.A. (1983). A reducible four-parameter system of model life tables. *Population Studies*, 37(1), 105-127.
- Forfar, D.O., McCutcheon, J.J. & Wilkie, A.D. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, 115(1), 97-245.
- French, D., & O'Hare, C. (2014). Forecasting death rates using exogenous determinants. *Journal of Forecasting*, 33(8), 640–650.
- Gage, T.B. & Mode, C.J. (1993). Some laws of mortality: how well do they fit? *Human Biology*, 65(3), 445-461.
- Giroi, F. & King, G. (2006). *Demographic forecasting*. Cambridge University Press, Cambridge.
- Giroi, F., & King, G. (2008). *Demographic forecasting*. Princeton, NJ: Princeton University Press.
- Haberman, S., & Renshaw, A. (2012). Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, 50(3), 309–333.

- Golulapati, R., De Ravin, J.W. & Trickett, P.J. (1984). Projections of Australian mortality rates, 1981-2020 (Occasional paper no. 1983/2). Australian Bureau of Statistics.
- Goodman, L.A., 1979. Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association* 74, 537–552.
- Goss, S.C., Wade, A., Bell, F. & Dussault, B. (1998). Historical and projected mortality for Mexico, Canada, and the United States. *North American Actuarial Journal*, 2(4), 108-128.
- Government Actuary's Department (2001). *National population projections: review of methodology for projecting mortality*. Government Actuary's Department: London.
- Government Actuary's Department (2006). *National population projections 2004-based*. Government Actuary's Department: London.
- Guillot, M. (2003). The cross-sectional average length of life (CAL): a period mortality measure that reflects the experience of cohorts. *Population Studies*, 57(1), 41-54. *Mortality Modelling and Forecasting: a Review of Methods* 37
- Gutterman, S. & Vanderhoof, I.T. (1998). Forecasting changes in mortality: a search for a law of causes and effects. *North American Actuarial Journal*, 2(4), 135-138.
- Gutterman, S., Vanderhoof, I.T., 2000. Forecasting changes in mortality: a search for a law of causes and effects. *North American Actuarial Journal* 2, 135–138.
- Haberman, S., & Renshaw, A. (2012). Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, 50(3), 309–333.
- Haberman, S., & Renshaw, A. (2013). Modelling and projecting mortality improvement rates using a cohort perspective. *Insurance: Mathematics and Economics*, 53(1), 150–168.
- Hannerz, H. (1999). *Methodology and applications of a new law of mortality*. Department of Statistics, University of Lund, Sweden, Lund.
- Hannerz, H. (2001a). Presentation and derivation of a five-parameter survival function intended to model mortality in modern female populations. *Scandinavian Actuarial Journal*, 2001(2), 176-187.
- Hannerz, H. (2001b). Manhood trials and the law of mortality. *Demographic Research*, 4(Article 7), 185-202.

- Hannerz, H. (2001c). An extension of relational methods in mortality estimation. *Demographic Research*, 4(Article 10), 337-367.
- Hartmann, M. (1987). Past and recent attempts to model mortality at all ages. *Journal of Official Statistics*, 3(1), 19-36.
- Harvey, A. (1989). Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.
- Hatzopoulos, P., & Haberman, S. (2009). A parameterized approach to modeling and forecasting mortality. *Insurance: Mathematics and Economics*, 44(1), 103–123.
- Hatzopoulos, P., & Haberman, S. (2013). Common mortality modeling and coherent forecasts. An empirical analysis of worldwide mortality data. *Insurance: Mathematics and Economics*, 52(2), 320–337.
- Hauser, R.M. & Willis, R.J. (2005). Survey design and methodology in the Health and Retirement Study and the Wisconsin Longitudinal Study. In (L.J. Waite, ed.) *Aging, health, and public policy: demographic and economic perspectives*. Population council. Supplement to Population and Development Review, 30 (2004), New York.
- Heathcote, C. & Higgins, T. (2001a). A regression model of mortality, with applications to the Netherlands. In (E. Tabeau, A. Van Den Berg Jeths & C. Heathcote, eds.) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (59-82)*. Kluwer Academic Publishers, Dordrecht.
- Heathcote, C. & Higgins, T. (2001b). Forecasting mortality from regression models: the case of the Netherlands. In (E. Tabeau, A. Van Den Berg Jeths & C. Heathcote, eds.) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (83-103)*. Kluwer Academic Publishers, Dordrecht.
- Heligman, L. & Pollard, J.H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, 107(1, No 434), 49-80.
- Hilton, J., Dodd, E., Forster, J. J., & Smith, P. W. F. (2018). Projecting UK mortality by using Bayesian generalized additive models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*. <https://doi.org/10.1111/rssc.12299>



- Hollmann, F.W., Mulder, T.J. & Kallan, J.E. (2000). Methodology and assumptions for the population projections of the United States: 1999 to 2100. *Working paper 38*, Population Division, U.S. Bureau of the Census.
- Hollmann, F.W., Mulder, T.J. & Kallan, J.E. (2000). Methodology and assumptions for the population projections of the United States: 1999 to 2100. *Working paper 38*, Population Division, U.S. Bureau of the Census.
- Hyndman R (2023). demography: Forecasting Mortality, Fertility, Migration and Population Data. <https://pkg.robjhyndman.com/demography/>, <https://github.com/robjhyndman/demography>.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
- Hyndman, R. J., & Ullah, M. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics & Data Analysis*, 51(10), 4942–4956.
- Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography*, 50(1), 261–283
- Hyndman, R.J. & Ullah, S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis*, 51, 4942-4956.
- Hyndman, R.J. & Ullah, S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis*, 51, 4942-4956.
- Idler, E.L. & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38(1), 21-37.
- Ishii, F. (2008). *Mortality projection model for Japan with age-shifting structure*. Paper presented at 2008 Annual Meeting of Population Association of America (New Orleans).
- Janssen, F., & Kunst, A. E. (2007). The choice among past trends as a basis for the prediction of future trends in old-age mortality. *Population Studies*, 61(3), 315–326.
- Janssen, F., van Wissen, L. J., & Kunst, A. E. (2013). Including the smoking epidemic in internationally coherent mortality projections. *Demography*, 50(4), 1341–1362.
- Kannisto, V. (1994). Development of oldest-old mortality, 1950-1990: evidence from 28 developed countries. Odense University Press, Odense, Denmark.

- Kannisto, V., Lauritsen, J., Thatcher, A.R. & Vaupel, J.W. (1994). Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review*, 20(4), 793-810.
- Keilman, N. (1990). Uncertainty in population forecasting: issues, backgrounds, analyses, recommendations. Swets & Zeitlinger, Amsterdam.
- Keilman, N. (1997). Ex-post errors in official population forecasts in industrialized countries, 13(3), 245-277.
- Keilman, N., Pham, D.Q. & Hetland, A. (2002). Why population forecasts should be probabilistic – illustrated by the case of Norway. *Demographic Research*, 6(15), 409-453.
- Keyfitz, N. (1982). Choice of function for mortality analysis: effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology*, 21(3), 329-352. 38
- Mortality Modelling and Forecasting: a Review of Methods
- Keyfitz, N. (1991). Experiments in the projection of mortality. *Canadian Studies in Population*, 18(2), 1-17.
- King, G., & Soneji, S. (2011). The future of death in America. *Demographic Research*, 25, 1–38.
- Kleinow, T. (2015). A common age effect model for the mortality of multiple populations. *Insurance: Mathematics and Economics*, 63, 147–152.
- Kogure, A., & Kurachi, Y. (2010). A Bayesian approach to pricing longevity risk based on risk-neutral predictive distributions. *Insurance: Mathematics and Economics*, 46(1), 162–172.
- Koissi, M.-C. & Shapiro, A.F. (2006). Fuzzy formulation of the Lee-Carter model for mortality forecasting. *Insurance: Mathematics and Economics*, 39, 287-309.
- Koissi, M.-C., Shapiro, A.F. & Hognas, G. (2006). Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38, 1-20.
- Koissi, M.-C., Shapiro, A.F. & Hognas, G. (2006). Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38, 1-20

Kontis, V., Bennett, J. E., Mathers, C. D., Li, G., Foreman, K., & Ezzati, M. (2017). Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *Lancet*, 389(10076), 1323–1335

Kostaki, A. (1988). Contributions to the methodology and application of the Heligman-Pollard formula. University of Lund, Lund.

Kunst, A.E., Mackenbach, J.P., Lautenbach, H., Oei, F.B. & Bijlsma, F. (2002). Gains in life expectancy by eliminating major causes of death: revised estimates taking into account competing causes of death. In (G. Wunsch, M. Mouchart & J. Duchene, eds.) *The life table: modelling survival and death* (191-207). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Le Bras, H. (2005). Mortality tempo versus removal of causes of mortality: opposite views leading to different estimations of life expectancy. *Demographic Research*, 13(25), 615- 640.

Lee, R. D., & Carter, L. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419), 659–671.

Lee, R.D. & Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38(4), 537-549.

Lee, R.D. & Tuljapurkar, S. (2000). Population forecasting for fiscal planning: issues and innovations. In (A. Auerbach & R. Lee, eds.) *Population and fiscal policy* (7-57). Cambridge University Press, Cambridge.

Lee, R.D. (1999). Probabilistic approaches to population forecasting. In (W. Lutz, J.W. Vaupel & D.A. Ahlburg, eds.) *Frontiers of population forecasting* (156-190). Population Council. A supplement to *Population and Development Review*, 24, 1998, New York.

Lee, R.D. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4(1), 80-93.

Lee, R.D., Nault, F., 1993. Modeling and forecasting provincial mortality in Canada. Paper presented at the World Congress of International Union for the Scientific Study of Population, Montreal

- Lee, R.D., Rofman, R., 1994. Modelacion y proyeccion de la mortalidad in Chile. *Natas* 22 (59), 182–313 (available in English from the authors, titled “Modeling and forecasting mortality in Chile”).
- Li, H., & O’Hare, C. (2017). Semi-parametric extensions of the Cairns–Blake–Dowd model: a one-dimensional kernel smoothing approach. *Insurance: Mathematics and Economics*, 77(3), 166–176.
- Li, J. (2013). A Poisson common factor model for projecting mortality and life expectancy jointly for females and males. *Population Studies*, 67(1), 111–126.
- Li, J. S., & Hardy, M. R. (2011). Measuring basis risk in longevity hedges. *North American Actuarial Journal*, 15(2), 177–200
- Li, N., & Lee, R. (2005). Coherent mortality forecasts for a group of populations: an extension of the lee-carter method. *Demography*, 42(3), 575–594.
- Li, N., and S. Tuljapurkar. 2012. *On the Accuracy of Life Expectancy*. Paper presented at the annual meeting of the Population Association of America, San Francisco.
- Li, N., and S. Tuljapurkar. 2013. *The Probabilistic Life Table and Its Applications*. Paper presented at the annual meeting of the Population Association of America, New Orleans.
- Li, S.-H. & Chan, W.-S. (2005). Outlier analysis and mortality forecasting: the United Kingdom and Scandinavian countries. *Scandinavian Actuarial Journal*, 3, 187-211.
- Li, S.-H., Hardy, M.R. & Tan, K.S. (2012). Uncertainty in mortality forecasting: an extension to the classical Lee-Carter approach. *ASTIN Bulletin*.
- Lundstrom, H. & Qvist, J. (2004). Mortality forecasting and trend shifts: an application of the Lee-Carter model to Swedish mortality data. *International Statistical Review*, 72(1), 37-50.
- Lutz, W. & Scherbov, S. (1998). An expert-based framework for probabilistic national population projections: the example of Austria. *European Journal of Population*, 14(1), 1-17.
- Lutz, W., Goldstein, J.R. & Prinz, C. (1996a). Alternative approaches to population projection. In (W. Lutz, ed.) *The future population of the world. What can we assume today?* (14-44). Earthscan, London. Mortality Modelling and Forecasting: a Review of Methods 39

- Lutz, W., Sanderson, W. & Scherbov, S. (1996b). Probabilistic population projections based on expert opinion. In (W. Lutz, ed.) *The future population of the world: what can we assume today?* (397-428). Earthscan, London.
- Lutz, W., Sanderson, W. & Scherbov, S. (1997). Doubling of world population unlikely. *Nature*, 387, 803-805.
- Lutz, W., Sanderson, W. & Scherbov, S. (1999). Expert-based probabilistic population projections. In (W. Lutz, J.W. Vaupel & D.A. Ahlburg, eds.) *Frontiers of population forecasting* (139-155). Population Council. A Supplement to Population and Development Review, 24 (1998), New York.
- Lutz, W., Sanderson, W. & Scherbov, S. (2001). The end of world population growth. *Nature*, 412, 543-545.
- Lutz, W., Sanderson, W. & Scherbov, S. (2004). The end of world population growth. In (W. Lutz & W. Sanderson, eds.) *The end of world population growth in the 21st century: new challenges for human capital formation and sustainable development*. Earthscan, London.
- Manton, K.G., Patrick, C.H. & Stallard, E. (1980). Mortality model based on delays in progression of chronic diseases: alternative to cause elimination model. *Public Health Reports*, 95(6), 580-588.
- Manton, K.G., Stallard, E. & Singer, B. (1992). Projecting the future size and health status of the U.S. elderly population. *International Journal of Forecasting*, 8(3), 433-458.
- Manton, K.G., Stallard, E. & Tolley, H.D. (1991). Limits to human life expectancy: evidence, prospects, and implications. *Population and Development Review*, 17(4), 603-637.
- McNown, R. & Rogers, A. (1989). Forecasting mortality: a parameterized time series approach. *Demography*, 26(4), 645-660.
- McNown, R. & Rogers, A. (1992). Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting*, 8(3), 413-432.
- McNown, R., Rogers, A. & Little, J. (1995). Simplicity and complexity in extrapolative population forecasting models. *Mathematical Population Studies*, 5(3), 235-257.
- Mitchell, D., Brockett, P., Mendoza-Arriaga, R., & Muthuraman, K. (2013). Modeling and forecasting mortality rates. *Insurance: Mathematics and Economics*, 52(2), 275–285.

- Mode, C.J. & Busby, R.C. (1982). An eight-parameter model of human mortality ö the single decrement case. *Bulletin of Mathematical Biology*, 44(5), 647-659.
- Murphy, M.J. (1995). The prospect of mortality: England and Wales and the United States of America, 1962-1989. *British Actuarial Journal*, 1, 331-350.
- Murray, C. & Lopez, A. (1997). Global mortality, disability, and the contribution of risk factors: global burden of disease study. *The Lancet*, 349(9063), 1436-1442.
- Murtin, F., Mackenbach, J., Jasilionis, D., & d'Ercole, M. M. (2017). Inequalities in longevity by education in OECD countries: insights from new OECD estimates. *OECD Statistics Working Paper No. 78*. Paris: OECD Publishing
- Olshansky, S. J., Carnes, B. A., & Mandell, M. S. (2009). Future trends in human longevity: implications for investments, pensions, and the global economy. *Pensions International Journal*, 14(3), 149–163.
- Olshansky, S.J. & Carnes, B.A. (1997). Ever since Gompertz. *Demography*, 34(1), 1-15.
- Olshansky, S.J. (1988). On forecasting mortality. *The Milbank Quarterly*, 66(3), 482-530.
- Pampel, F. (2005). Forecasting sex differences in mortality in high income nations: the contribution of smoking. *Demographic Research*, 13(18), 455–484
- Pedroza, C. (2006). A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics*, 7(4), 530-550.
- Plat, R. (2009). On stochastic mortality models. *Insurance: Mathematics and Economics*, 45(3), 393–404.
- Pollard, J.H. & Valkovics, E.J. (1992). The Gompertz distribution and its applications. *Genus*, 48(3-4), 15-27.
- Pollard, J.H. (1987). Projection of age-specific mortality rates Population Bulletin of the United Nations 21-22, 55-69. 40 *Mortality Modelling and Forecasting: a Review of Methods*
- Pollard, J.H. (1998). Keeping abreast of mortality change. *Actuarial and Demography Research Paper Series*, No. 002/98.
- Pollard, J.H. 1973. *Mathematical Models for the Growth of Human Population*. Cambridge MA: Cambridge University Press.

- Preston, S. M., Stokes, A., Mehta, N. K., & Cao, B. (2012). Projecting the effect of changes in smoking and obesity on future life expectancy in the United States. National Bureau of Economic Research: *Working Paper 18407*
- Raftery, A. E., Chunn, J. L., Gerland, P., & Sevcikova, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography*, 50(3), 777–801.
- Reither, E. N., Olshansky, S. J., & Yang, Y. (2011). New forecasting methods indicates more disease and earlier mortality ahead for today's young Americans. *Health Affairs*, 30(8), 1562–1568.
- Renshaw, A. E., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3), 556–570.
- Renshaw, A.E. & Haberman, S. (2000). Modelling for mortality reduction factors (Actuarial Research Paper No. 127). *Department of Actuarial Science and Statistics*, City University, London.
- Renshaw, A.E. & Haberman, S. (2003a). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.
- Renshaw, A.E. & Haberman, S. (2003b). Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Applied Statistics*, 51(1), 119-137.
- Renshaw, A.E. & Haberman, S. (2003c). On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics*, 32, 379-401.
- Renshaw, A.E. & Haberman, S. (2006). A cohort-based extension of the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38, 556-570.
- Renshaw, A.E. (1991). Actuarial graduation practice and generalised linear and non-linear models. *Journal of the Institute of Actuaries*, 118, 295-312.
- Renshaw, A.E., Haberman, S. & Hatzoupoulos, P. (1996). The modelling of recent mortality trends in United Kingdom male assured lives. *British Actuarial Journal*, 2, 449-477.
- Rogers, A. & Gard, K. (1991). Applications of the Heligman/Pollard model mortality schedule. *Population Bulletin of the United Nations*, 30, 79-105.

- Rogers, A. & Little, J.S. (1994). Parameterizing age patterns of demographic rates with the multiexponential model schedule. *Mathematical Population Studies*, 4(3), 175-195.
- Rogers, A. & Planck, F. (1983). Model: a general program for estimating parameterized model schedules of fertility, mortality, migration, and marital and labor force status transitions. International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Russolillo, M., Giordano, G., & Haberman, S. (2011). Extending the Lee-Carter model: a three-way decomposition. *Scandinavian Actuarial Journal*, 2011(2), 96–117.
- Sanderson, W.C. & Scherbov, S. (2005). Average remaining lifetimes can increase as human populations age. *Nature*, 435(7043), 811-813.
- Schinzinger, E., Denuit, M. M., & Christiansen, M. C. (2016). A multivariate evolutionary credibility model for mortality improvement rates. *Insurance: Mathematics and Economics*, 69(2), 70–81.
- Shair, S., Purcal, S., & Parr, N. (2017). Evaluating extensions to coherent mortality forecasting models. *Risks*, 5(1), 1–20.
- Shang, H. L. (2012). Point and interval forecasts of age-specific life expectancies: a model averaging approach. *Demographic Research*, 27(21), 593–644.
- Shang, H. L., & Hyndman, R. J. (2017). Grouped functional time series forecasting: an application to age-specific mortality rates. *Journal of Computational and Graphical Statistics*, 26(2), 330–343.
- Shang, H. L., Booth, H., & Hyndman, R. (2011). Point and interval forecasts of mortality rates and life expectancy: a comparison of ten principal component methods. *Demographic Research*, 25(5), 173–214.
- Shaw, C. (1994). Accuracy and uncertainty of the national population projections for the United Kingdom. *Population Trends*, 77, 24-32.
- Siler, W. (1983). Parameters of mortality in human populations with widely varying life spans. *Statistics in Medicine*, 2, 373-380.
- Sithole, T.Z., Haberman, S. & Verrall, R.J. (2000). An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data. *Insurance: Mathematics and Economics*, 27, 285-312.



- Stoeldraijer, L., van Duin, C., van Wissen, L., & Janssen, F. (2013). Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: the case of the Netherlands. *Demographic Research*, 29(13), 323–354.
- Stoto, M.A. & Durch, J.S. (1993). Forecasting survival, health, and disability: report on a workshop. *Population Development and Review*, 19(3), 557-581.
- Stoto, M.A. (1988). Dealing with uncertainty: statistics for an aging population. *The American Statistician*, 42(2), 103-110.
- Tabeau, E. (2001). A review of demographic forecasting models for mortality. In (E. Tableau, A. Van Den Berg Jeths & C. Heathcote, eds.) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (1-32)*. Kluwer Academic Publishers, Dordrecht.
- Tabeau, E., Ekamper, P., Huisman, C. & Bosch, A. (2001b). Predicting mortality from period, cohort or cause-specific trends: a study of four European countries. In (E. Tableau, A. Van Den Berg Jeths & C. Heathcote, eds.) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (159-187)*. Kluwer Academic Publishers, Dordrecht.
- Thatcher, A., Kannisto, V. & Vaupel, J. (1998). The force of mortality at ages 80 to 120. In *Odense Monographs on Population Aging 5*. Odense University Press, Odense, Denmark.
- Thatcher, A.R. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society, Series A, Statistics in Society*, 162(1), 5-43.
- Torri, T., & Vaupel, J. W. (2012). Forecasting life expectancy in an international context. *International Journal of Forecasting*, 28(2), 519–531.
- Tuljapurkar, S. & Boe, C. (1998). Mortality change and forecasting: how much and how little do we know? *North American Actuarial Journal*, 2(4), 13-47.
- Tuljapurkar, S. (1998). Forecasting mortality change: questions and assumptions. *North American Actuarial Journal*, 2(4), 127-134.
- Tuljapurkar, S., Li, N. & Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Nature*, 405, 789-792.
- United Nations (2004). *World population to 2300*. United Nations, New York.

- van Baal, P., Peters, F., Mackenbach, J., & Nusselder, W. (2016). Forecasting differences in life expectancy by education. *Population Studies*, 70(2), 201–216.
- van Berkum, F., Antonio, K., & Vellekoop, M. (2016). The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal*, 2016(7), 581–603
- Van Hoorn, W. & De Beer, J. (2001). Projecting mortality in population forecasts in the Netherlands. In (E. Tabeau, A. Van Den Berg Jeths & C. Heathcote, eds.) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (205-226)*. Kluwer Academic Publishers, Dordrecht.
- Vaupel, J.W., K.G. Manton, and E. Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16(3):439–54.
- Vermunt, J.K., 1997a. *Log-linear Models for Event Histories*. Sage, Thousand Oakes.
- Vermunt, J.K., 1997b. *LEM: A General Program for the Analysis of Categorical Data*. Users' manual. Tilburg University, Tilburg, The Netherlands. <http://www.kub.nl/mto>.
- Villegas, A. M., Millossovich, P., & Kaishev, V. K. (2018). StMoMo: Stochastic mortality modeling in R. *Journal of Statistical Software*, 84(3), 1–38.
- Waldron, H. (2005). Literature review of long-term mortality projections. *Social Security Bulletin*, 66(1), 16-30.
- Wan, C., Bertschi, L., & Yang, Y. (2013). Coherent mortality forecasting for small populations: an application to Swiss mortality data. Working paper presented at the International Actuarial Association Colloquium, June 23–26 2013, Lyon.
- Wang, C., & Yang, S. S. (2013). Pricing survivor derivatives with cohort mortality dependence under the Lee–Carter framework. *Journal of Risk and Insurance*, 80(4), 1027–1056.
- Wan, J. (2007). Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model. Mathematical Statistics Stockholm University.
- Wang, H., & Preston, S. H. (2009). *Forecasting United States mortality using cohort smoking histories*. Proceedings of the National Academy of Sciences, 106(2), 393–398.
- Wang, S.S. (2000). A class of distortion operators for pricing financial and insurance risks. *The Journal of Risk and Insurance*, 67(2), 15-36.

- Wang, S.S. (2000). A class of distortion operators for pricing financial and insurance risks. *The Journal of Risk and Insurance*, 67(2), 15-36
- White, K.M. (2002). Longevity advanced in high-income countries, 1955-96. *Population and Development Review*, 28(1), 59-76.
- Willekens, F.J. & Baydar, N. (1986). Age-period-cohort models for forecasting fertility (No. 45). NIDI, The Hague.
- Willets, R.C. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, 10, 833-877.
- Willets, R.C., Gallop, A.P., Leandro, P.A., Lu, J.L.C., Macdonald, A.S., Miller, K.A., Richards, S.J., Robjohns, N., Ryan, J.P. & Waters, H.R. (2004). Longevity in the 21st century. *British Actuarial Journal*, 10, 685-898.
- Wilmoth, J.R. (1990). Variation in vital rates by age, period and cohort. *Sociological Methodology*, 20, 295-335.
- Wilmoth, J.R. (1993). *Computational methods for fitting and extrapolating the Lee-Carter model of mortality change (Technical report)*. Department of Demography, University of California, Berkeley. 42 Mortality Modelling and Forecasting: a Review of Methods
- Wilmoth, J.R. (1995). Are mortality projections always more pessimistic when disaggregated by cause of death? *Mathematical Population Studies*, 5(4), 293-319.
- Wilmoth, J.R. (2005). Some methodological issues in mortality projection, based on an analysis of the U.S. social security system. *Genus*, LXI(1), 179-212.
- Wilmoth, J.R., 1993. *Computational methods for fitting and extrapolating the Lee-Carter model of mortality change*. Technical report. Department of Demography. University of California, Berkeley.
- Wiśniowski, A., Smith, P. W. F., Bijak, J., Raymer, J., & Forster, J. (2015). Bayesian population forecasting: extending the Lee Carter method. *Demography*, 52(3), 1035–1059
- Wolf, D.A. (2004). Another variation on the Lee-Carter model. Paper presented at the Annual meeting of the Population Association of America, Boston.

Wong, J. S. T., Forster, J. J., & Smith, P. W. F. (2017). Bayesian mortality forecasting with overdispersion. *Insurance: Mathematics and Economics*.  
<https://doi.org/10.1016/j.insmatheco.2017.09.023>.

Wong-Fupuy, C. & Haberman, S. (2004). Projecting mortality trends: recent developments in the United Kingdom and the United States. *North American Actuarial Journal*, 8(2), 56-83.

Zaba, B. & Paes, N. (1995). An alternative procedure for fitting relational model life tables. *Genus*, LI(1-2), 19-43

Zaba, B. (1979). The four-parameter logit life table system. *Population Studies*, 33(1), 79-100.

Zhou, R., Wang, Y., Kaufhold, K., Li, J. S. H., & Tan, K. S. (2014). Modeling mortality of multiple populations with vector error correction models: applications to solvency II. *North American Actuarial Journal*, 18, 150–167.