ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

ΕΞΟΜΑΛΥΝΣΗ ΠΙΝΑΚΩΝ ΕΠΙΒΙΩΣΗΣ ΚΑΙ ΕΦΑΡΜΟΓΗ ΜΕ ΤΑ ΔΕΔΟΜΕΝΑ ΤΟΥ ΠΛΗΘΥΣΜΟΥ ΤΗΣ ΕΛΛΑΔΑΣ ΤΟΥ ΕΤΟΥΣ 2001

Βασίλειος Αντ. Καμούτσης

Μ.Α.Ε. 09011

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 Διοικητική Κινδύνου

Πειραιάς

Οκτώβριος 2011

UNIVERSITY OF PIRAEUS

DEPARTMENT OF STATISTICS & INSURANCE SCIENCE

POSTGRADUATE PROGRAM IN ACTUARIAL SCIENCE & RISK MANAGEMENT

GRADUATION OF MORTALITY TABLES AND AN APPLICATION TO GREECE'S POPULATION DATA OF 2001

Vasileios Ant. Kamoutsis

M.A.E. 09011

Msc Dissertation

submitted to the Department of Statistics & Insurance Science of the University of Piraeus in partial fulfillment of the requirements for the degree of Master of Science in Actuarial Science & Risk Management

Piraeus

October 2011

Στους γονείς μου

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες προς τον επιβλέποντα καθηγητή της παρούσας διπλωματικής εργασίας, κύριο Κλέωνα Τσίμπο, για τη συμπαράσταση και τη βοήθεια που μου προσέφερε κατά τη διάρκεια συγγραφής της. Επίσης, θα ήθελα να ευχαριστήσω τα μέλη της συμβουλευτικής επιτροπής, κύριο Κωνσταντίνο Πολίτη και κυρία Γεωργία Βερροπούλου, για την προσοχή που επέδειξαν κατά την ανάγνωση της παρούσας εργασίας και για τις επισημάνσεις τους στα τυχόν λάθη της.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου και ιδιαίτερα τους γονείς μου, τόσο για την υλική, όσο και για την ηθική τους υποστήριξη κατά τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Για την κατασκευή των δημογραφικών Πινάκων Επιβίωσης, απαιτούνται ληξιαρχικά και απογραφικά δεδομένα ταξινομημένα κατά ηλικία και φύλο. Η ποιότητα των Πινάκων Επιβίωσης, εξαρτάται σε μεγάλο βαθμό από την αξιοπιστία του χρησιμοποιούμενου στατιστικού υλικού. Σε όλες σχεδόν τις περιπτώσεις, παρατηρούνται σφάλματα που αφορούν κυρίως στις διαφυγές βρεφικών και παιδικών θανάτων και στη δήλωση των ηλικιών του πληθυσμού και των θανόντων. Αποτέλεσμα αυτού είναι ότι οι εμπειρικές πιθανότητες θανάτου - που εκτιμώνται από τα ακατέργαστα στοιχεία - παρουσιάζουν τιμές που δεν συμβαδίζουν με τα διεθνώς παρατηρούμενα πρότυπα και τους βιολογικούς νόμους που διέπουν το φαινόμενο της θνησιμότητας, και ως εκ τούτου, για την κατασκευή των Πινάκων εφαρμόζονται τεχνικές που έχουν ως σκοπό την εξομάλυνση τέτοιου είδους σφαλμάτων.

Σκοπός της παρούσας εργασίας είναι η συστηματική βιβλιογραφική αναφορά στις υπάρχουσες στατιστικές προσεγγίσεις, καθώς και η εφαρμογή κάποιας από αυτές τις μεθόδους στα στοιχεία του πληθυσμού της Ελλάδας, τα οποία βασίζονται στην απογραφή του 2001 και τις ληξιαρχικές καταγραφές θανάτων του ίδιου έτους.

ABSTRACT

To create demographic Mortality Tables, we need to have the census data and the deaths sorted by age and sex. The quality of the Mortality Tables, depends mainly on the credibility of the statistical data we use. Almost in every case, we observe mistakes regarding the infant and child deaths, as well as the age of each member of the population and the deceased. As a result the probabilities of death - which are estimated by the raw data – do not go together with the international observed patterns and the biological laws of mortality, therefore, to create the Tables we apply certain techniques that intend to graduate this type of mistakes.

This dissertation intends to refer to the existing statistical approaches, as well as to adapt one of them to Greece's population data, based on the 2001 census and the deaths of the same year.

ΠΕΡΙΕΧΟΜΕΝΑ

ΜΕΡΟΣ 1ο

<u>ΚΕΦΑΛΑΙΟ 1ο</u>

ΘΝΗΣΙΜΟΤΗΤΑ – ΑΠΟΓΡΑΦΕΣ - ΛΗΞΙΑΡΧΙΚΕΣ ΚΑΤΑΓΡΑΦΕΣ

1.1 ΘΝΗΣΙΜΟΤΗΤΑ

1.1.1 ΤΟ ΦΑΙΝΟΜΕΝΟ ΤΗΣ ΘΝΗΣΙΜΟΤΗΤΑΣ

Η *Παγκόσμια Οργάνωση Υγείας* (Π.Ο.Υ.) ορίζει το *θάνατο* ως εξής : ‹‹ *Θάνατος είναι η διαρκής και οριστική εξαφάνιση κάθε ένδειξης ζωής, η οποία επέρχεται σε οποιαδήποτε χρονική στιγμή μετά τη γέννηση ζώντος ανθρωπίνου οργανισμού* ››. Από τον προηγούμενο ορισμό εξαιρείται ο *θάνατος εμβρύου* ή αλλιώς *γέννηση νεκρού. Ως βρεφικός θάνατος ορίζεται ο θάνατος ενός βρέφους κατά τη διάρκεια του πρώτου έτους της ζωής του* (δηλαδή μεταξύ της στιγμής της γέννησής του και της συμπλήρωσης του πρώτου έτους της ηλικίας του). Ακόμη, *καλείται ως μητρικός θάνατος, ο θάνατος γυναίκας που επέρχεται είτε κατά τη διάρκεια της γέννας, είτε κατά τη διάρκεια της εγκυμοσύνης*. Ο θάνατος - ως δημογραφικό γεγονός ή συμβάν - είναι *αναπόφευκτος* (δηλαδή είναι βέβαιο ότι θα συμβεί) και *μη - επαναλαμβανόμενος* (δηλαδή θα συμβεί μία και μόνο φορά). Αν τώρα προσεγγίσουμε το παραπάνω γεγονός σε συνολικό και όχι σε ατομικό επίπεδο, τότε η επέλευση του θανάτου γενικεύεται και αποκτά χαρακτήρα φαινομένου. Αυτό ακριβώς το γενικευμένο φαινόμενο, το οποίο χρίζει μελέτης, καλείται *θνησιμότητα*.

Η μελέτη της θνησιμότητας είναι σχετικά απλούστερη απ’ ότι των υπολοίπων φαινομένων που αφορούν και επηρεάζουν ένα πληθυσμό, καθώς *τα δύο μοναδικά στοιχεία που είναι άγνωστα όσον αφορά το θάνατο είναι ο χρόνος που αυτός θα συμβεί και η αιτία του*. Παράγοντες που παρουσιάζουν ερευνητικό ενδιαφέρον κατά τη μελέτη του παραπάνω φαινομένου είναι :

- *Τα χαρακτηριστικά του θανόντος (ηλικία, φύλο, επάγγελμα, αιτία θανάτου, εκπαίδευση, υπηκοότητα κ.λπ.).*

- *Τα χαρακτηριστικά του περιβάλλοντος (υψόμετρο, τιμές μόλυνσης, διατροφικές συνήθειες, ποιότητα νερού κ.λπ.).*

- *Τα χαρακτηριστικά της καταγραφής του γεγονότος (τόπος θανάτου, ημερομηνία θανάτου, μέρος όπου συνέβη ο θάνατος (ιδιωτική κατοικία, νοσοκομείο), άτομο που πιστοποίησε το θάνατο κ.λπ.).*

Μέσω της μελέτης της θνησιμότητας προκύπτουν αποτελέσματα τα οποία έχουν πολύ μεγάλη σημασία, κυρίως λόγω του ότι μπορούν να χρησιμοποιηθούν αποτελεσματικά από πολλούς επιστημονικούς κλάδους όπως :

- Από τη *Δημογραφία* για την εκτίμηση της τρέχουσας δημογραφικής κατάστασης ενός πληθυσμού, την ανάλυση ιστορικών τάσεων εξέλιξης του πληθυσμού και την επεξεργασία δημογραφικών προβολών και προβλέψεων.

- Από την *Ασφαλιστική - Αναλογιστική Επιστήμη* για την εκτίμηση διαφόρων ποσοτήτων των προγραμμάτων που σχετίζονται με την διάρκεια ζωής ενός ατόμου, όπως οι Ασφαλίσεις Ζωής, τα Συνταξιοδοτικά Προγράμματα κ.ά..

- Από τη *Βιολογία και τις Ιατρικές Επιστήμες* γενικότερα για την αξιολόγηση των διαφόρων νοσημάτων όπως αυτά σχετίζονται με τον κίνδυνο του θανάτου.

1.1.2 ΔΕΙΚΤΕΣ ΘΝΗΣΙΜΟΤΗΤΑΣ

Οι *δείκτες* που αφορούν τη θνησιμότητα και υπολογίζονται με βάση τα διαθέσιμα στατιστικά στοιχεία, χωρίζονται σε *απλούς* και *προτυποποιημένους*. Τα στατιστικά στοιχεία που είναι απαραίτητα για τον υπολογισμό αναλυτικών και προτυποποιημένων δεικτών θνησιμότητας είναι :

- *Ο αριθμός των θανάτων που προέρχεται από τις ληξιαρχικές καταγραφές.*

- *Το μέγεθος του πληθυσμού που προέρχεται είτε από μια απογραφή πληθυσμού, είτε από μια εκτίμησή του από κάποιον φορέα (όπως η Στατιστική Υπηρεσία επί παραδείγματι).*

- *Το μέγεθος ενός πρότυπου πληθυσμού ή μιας σειράς πρότυπων δεικτών θνησιμότητας (για τους προτυποποιημένους δείκτες).*

Παρακάτω, παρατίθενται ενδεικτικά ορισμένοι βασικοί *δείκτες θνησιμότητας*.

- ***Αδρός Δείκτης Θανάτων*** *(όλες οι αιτίες, όλες οι ηλικίες, και τα δύο φύλα)*

$$CDR = \frac{D}{P} * 1.000\ ,$$

όπου D : Το σύνολο των θανάτων ενός ημερολογιακού έτους.

P : Ο συνολικός πληθυσμός στο μέσο του έτους.

Ο δείκτης δείχνει την αναλογία των θανάτων σε 1.000 άτομα για ένα έτος και περιλαμβάνει όλες τις αιτίες θανάτου, όλες τις ηλικίες και για τα δύο φύλα.

- ***Ειδικός κατά Ηλικία Δείκτης Θνησιμότητας***

$$m_x = \frac{D_x}{P_x} * 1.000$$

ή

$${}_nm_x = \frac{{}_nD_x}{{}_nP_x} * 1.000 \quad \text{(για ομάδες ηλικιών } x \text{ έως } x+n \text{),}$$

όπου D_x : Το σύνολο των θανόντων μεταξύ των ηλικιών x και $x+1$.

${}_nD_x$: Το σύνολο των θανόντων μεταξύ των ηλικιών x και $x+n$.

P_x : Το σύνολο των ατόμων ηλικίας x στο μέσο του έτους.

${}_nP_x$: Το σύνολο των ατόμων μεταξύ των ηλικιών x και $x+n$ στο μέσο του έτους.

Ο δείκτης δείχνει την αναλογία των θανάτων σε 1.000 άτομα για μια συγκεκριμένη ηλικία ή για ένα διάστημα ηλικιών περιλαμβάνοντας όλες τις αιτίες θανάτου, για κάθε φύλο ξεχωριστά.

- ***Αδρός κατά Αιτία Δείκτης Θανάτου***

$$CDR_j = \frac{D_j}{P} * 100.000\,,$$

όπου D_j : Το σύνολο των θανάτων ενός ημερολογιακού έτους λόγω του αιτίου j.

P : Ο συνολικός πληθυσμός στο μέσο του έτους.

Ο δείκτης δείχνει την αναλογία των θανάτων σε 100.000 άτομα για ένα έτος, λόγω του αιτίου θανάτου j.

- ***Ειδικός κατά Ηλικία και Αιτία Δείκτης Θανάτου***

$$m_{x,j} = \frac{D_{x,j}}{P_x} * 100.000$$

ή

$${}_n m_{x,j} = \frac{{}_n D_{x,j}}{{}_n P_x} * 100.000\,,$$

όπου $D_{x,j}$: Το σύνολο των θανόντων μεταξύ των ηλικιών x και $x+1$, λόγω του αιτίου j.

${}_n D_{x,j}$: Το σύνολο των θανόντων μεταξύ των ηλικιών x και $x+n$, λόγω του αιτίου j .

P_x : Το σύνολο των ατόμων ηλικίας x στο μέσο του έτους.

${}_{n}P_{x}$: Το σύνολο των ατόμων μεταξύ των ηλικιών x και $x+n$ στο μέσο του έτους.

Ο δείκτης δείχνει την αναλογία των θανάτων σε 100.000 άτομα για μια συγκεκριμένη ηλικία ή για ένα διάστημα ηλικιών για το αίτιο θανάτου j, για κάθε φύλο ξεχωριστά.

- ***Δείκτης Βρεφικής Θνησιμότητας***

$$IMR = \frac{D_{0-365}}{B} * 1.000 ,$$

όπου D_{0-365} : Το σύνολο των βρεφικών θανάτων ενός ημερολογιακού έτους.

B : Ο συνολικός αριθμός γεννήσεων ζώντων του ίδιου έτους.

Ο δείκτης δείχνει την αναλογία των βρεφικών θανάτων σε 1.000 γεννήσεις ζώντων για ένα έτος.

- ***Άμεσα Προτυποποιημένος Δείκτης Θανάτου***

$$DSDR = \frac{\sum_{x=0}^{\infty} {}_{n}m_{x} * {}_{n}P_{x}^{s}}{\sum_{x=0}^{\infty} {}_{n}P_{x}^{s}} * 1.000 ,$$

όπου ${}_{n}P_{x}^{s}$: Το σύνολο των ατόμων μεταξύ των ηλικιών x και $x+n$ ενός πρότυπου πληθυσμού που επιλέγεται για την προτυποποίηση.

${}_{n}m_{x}$: Ο Ειδικός κατά Ηλικία Δείκτης Θνησιμότητας (για την ομάδα ηλικιών από x έως $x+n$).

$\sum_{x=0}^{\infty} {}_nm_x * {}_nP_x^s$: Το σύνολο των θανάτων για όλα τα κλιμάκια των ηλικιών με βάση την ηλικιακή δομή του πρότυπου πληθυσμού.

$\sum_{x=0}^{\infty} {}_nP_x^s$: Το σύνολο των ατόμων του πρότυπου πληθυσμού.

Με αυτόν το δείκτη εφαρμόζουμε άμεση προτυποποίηση, επιλέγοντας έναν πρότυπο πληθυσμό και εφαρμόζοντας την κατά ηλικία δομή του στους δείκτες θνησιμότητας του υπό μελέτη πληθυσμού μας.

- ***Προτυποποιημένος Λόγος Θνησιμότητας***

$$SMR = \frac{\sum_{x=0}^{\infty} {}_nm_x * {}_nP_x}{\sum_{x=0}^{\infty} {}_nm_x^s * {}_nP_x} * 100 = \frac{D}{\sum_{x=0}^{\infty} {}_nm_x^s * {}_nP_x} * 100 ,$$

όπου ${}_nP_x$: Το σύνολο των ατόμων μεταξύ των ηλικιών x και $x + n$.

${}_nm_x^s$: Ο Ειδικός κατά Ηλικία Δείκτης Θνησιμότητας του πρότυπου πληθυσμού (για την ομάδα ηλικιών από x έως $x + n$).

$\sum_{x=0}^{\infty} {}_nm_x^s * {}_nP_x$: Το σύνολο των θανάτων με βάση τις συνθήκες θνησιμότητας του πρότυπου πληθυσμού.

D: Το σύνολο των θανάτων ενός ημερολογιακού έτους.

Με αυτόν το δείκτη εφαρμόζουμε έμμεση προτυποποίηση, επιλέγοντας πρότυπους κατά ηλικία δείκτες θνησιμότητας και εφαρμόζοντάς τους στην πραγματική κατά ηλικία δομή του υπό μελέτη πληθυσμού μας.

1.2 ΑΠΟΓΡΑΦΕΣ ΠΛΗΘΥΣΜΟΥ

Ως απογραφή πληθυσμού, ορίζουμε το σύνολο των ενεργειών και χειρισμών που αποσκοπούν στη συγκέντρωση στατιστικών πληροφοριών σχετικά με το μέγεθος και τα χαρακτηριστικά του πληθυσμού ενός συγκεκριμένου γεωγραφικού χώρου, σε συγκεκριμένο χρόνο. Πληροφορίες συλλέγονται με βάση δημογραφικά, οικονομικά, κοινωνικά, οικιστικά κ.ά. χαρακτηριστικά του πληθυσμού. Η ίδια η διαδικασία της απογραφής είναι *καθολική*, *ονομαστική*, *ταυτόχρονη* και *περιοδική*. Οι απογραφές πραγματοποιούνται σε εκτέλεση νόμου και προϋποθέτουν την καθολική και ταυτόχρονη καταγραφή όλων όσων βρίσκονται κατά την ημέρα διεξαγωγής της απογραφής σε έναν οριοθετημένο γεωγραφικό χώρο, που συνήθως είναι η εθνική επικράτεια και οι γεωγραφικές υποδιαιρέσεις της. Η συλλογή των δεδομένων γίνεται με τη συμπλήρωση ειδικών ερωτηματολογίων που καλούνται *απογραφικά δελτία*. Το απόρρητο των παρεχόμενων πληροφοριών προστατεύεται θεσμικά. Οι απογραφές συνήθως πραγματοποιούνται την περίοδο της άνοιξης είτε Κυριακές, είτε αργίες, ώστε να εξασφαλιστεί η όσο το δυνατόν μικρότερη γεωγραφική κινητικότητα του πληθυσμού και η οργανωτική και διοικητική ευελιξία του συμμετέχοντος προσωπικού. Ως στατιστική μονάδα άντλησης πληροφοριών λαμβάνεται το κάθε *νοικοκυριό*. Εκτός από την ημέρα κατά την οποία πραγματοποιείται η απογραφή, λίγες ημέρες πριν τη διεξαγωγή της γίνονται *προ - απογραφικές δοκιμαστικές έρευνες* δοκιμαστικά. Αντίστοιχα, μετά το πέρας της απογραφής και της συγκέντρωσης των συμπληρωμένων ερωτηματολογίων, διενεργούνται *μετα - απογραφικές έρευνες*, ώστε να εκτιμηθεί η πληρότητα των απογραφικών εγγραφών.

Οι απογραφές ως πηγές στατιστικής πληροφόρησης, είναι γνωστές από την αρχαιότητα. Από τις απογραφές της αρχαιότητας οι πιο σημαντικές είναι αυτές που διενήργησαν οι Αιγύπτιοι, οι Έλληνες, οι Ρωμαίοι και οι Ασσυροβαβυλώνιοι. Απτά στοιχεία από απογραφή υπάρχουν για την απογραφή που έγινε με διαταγή του Κινέζου βασιλιά Γιάο το 2238 π. Χ. και αφορούσε φορολογικούς σκοπούς. Φορολογικοί και στρατιωτικοί ήταν οι λόγοι για τους οποίους πραγματοποιούνταν απογραφές κατά την αρχαιότητα. Κατά τον 5ο π. Χ. αιώνα στην αρχαία Ρώμη οι απογραφές αρχίζουν να ξεχωρίζουν από αυτές παλαιότερων ετών, καθώς εμφανίζονται για πρώτη φορά απογραφείς, στους οποίους οι πολίτες κατέθεταν με ένορκή τους δήλωση, τα ζητούμενα από αυτούς στοιχεία. Στην Ελλάδα η πρώτη απογραφή έγινε το 1828 και περιελάμβανε και την αναδρομική εξακρίβωση του πληθυσμού το 1821. Από το 2^{o} Παγκόσμιο Πόλεμο και ύστερα πραγματοποιούνται απογραφές (αρχής γενομένης από το 1951) κάθε δεκαετία. Η *απογραφή του 2001* πραγματοποιήθηκε στις 18 Μαρτίου.

1.3 ΛΗΞΙΑΡΧΙΚΕΣ ΚΑΤΑΓΡΑΦΕΣ

Οι *ληξιαρχικές καταγραφές* συνιστούν την κύρια πηγή άντλησης πληροφοριών, σχετικά με τη *φυσική κίνηση* ενός πληθυσμού. *Με τον όρο φυσική κίνηση αναφερόμαστε τόσο σε γεγονότα βιολογικής χροιάς (γεννήσεις, θανάτους), όσο και σε γεγονότα κοινωνικής φύσης (διαζύγια, γάμους) και αποτελούν τον κύριο παράγοντα αναφορικά με τη δημογραφική εξέλιξη ενός πληθυσμού.* Για κάθε τέτοιου είδους γεγονός λοιπόν, το σύστημα των ληξιαρχικών καταγραφών προβαίνει σε όλες τις νόμιμες ενέργειες που αποσκοπούν στην καταγραφή, στην καταχώριση, στη συγκέντρωση, στην παρουσίαση και στη διανομή όλων των απαραιτήτων πληροφοριών που σχετίζονται με αυτά τα συμβάντα. Για κάθε γέννηση, θάνατο, γάμο, διαζύγιο κ.ά. συμπληρώνεται, μετά από αίτηση στον αρμόδιο ληξίαρχο, ένα ειδικό νόμιμο έντυπο, που καλείται *ληξιαρχική πράξη*. Με τις ληξιαρχικές πράξεις βεβαιώνεται η επέλευση των διαφόρων δημογραφικών γεγονότων. Οι δηλώσεις όλων των γεγονότων είναι υποχρεωτικές, δεν χρεώνονται και η εμπιστευτικότητα των στοιχείων προστατεύεται με νομοθετική πράξη. Τα στοιχεία μπορούν να χρησιμοποιηθούν μόνο για διοικητικούς και στατιστικούς σκοπούς και σε συλλογικό επίπεδο. Κάθε διοικητική υποδιαίρεση της χώρας αποτελεί μια *ληξιαρχική περιφέρεια* της οποίας προΐσταται ο *ληξίαρχος*. Το σύστημα των ληξιαρχικών καταγραφών είναι απαραίτητο να καλύπτει όλες τις περιοχές και τις πληθυσμιακές ομάδες της χώρας, ώστε να επιτυγχάνεται η συνεχής και μόνιμη καταγραφή και καταχώριση όλων των γεγονότων και των στοιχείων των συμμετεχόντων σε αυτά. Το σύστημα των ληξιαρχικών καταγραφών θεσπίστηκε και λειτουργεί στην Ελλάδα από το 1836. Τα στοιχεία που συλλέγονται από τις κατά τόπους ληξιαρχικές περιφέρειες, διοχετεύονται στον κεντρικό φορέα, την *Εθνική Στατιστική Υπηρεσία Ελλάδος* (Ε.Σ.Υ.Ε.) η οποία τα επεξεργάζεται και τα δημοσιεύει σε ετήσιες και ειδικές εκδόσεις της.

Όσον αφορά ειδικότερα τους θανάτους, αυτοί πρέπει να δηλώνονται υποχρεωτικά στο ληξίαρχο, εντός 24 ωρών από την επέλευσή τους και μετά. Αρμόδιος για αυτή τη δήλωση είναι είτε η Διεύθυνση του νοσοκομείου στο οποίο επήλθε ο θάνατος, είτε ο θεράπων ιατρός, είτε ο πλησιέστερος συγγενής. Με βάση αυτή τη δήλωση καταρτίζεται η *ληξιαρχική πράξη θανάτου* ή *Ιατρικό Πιστοποιητικό Θανάτου* που περιέχει στοιχεία που αφορούν κυρίως τα στοιχεία ταυτότητας του ατόμου που προέβη στη δήλωση (ονοματεπώνυμο, επάγγελμα, διεύθυνση κατοικίας κλπ.), τον τόπο και το χρόνο του θανάτου, τα στοιχεία ταυτότητας του θανόντος (ονοματεπώνυμο, ιθαγένεια, επάγγελμα, διεύθυνση κατοικίας, τόπο και χρόνο γέννησης), το φύλο του θανόντος, το μέρος που συνέβη ο θάνατος, το ονοματεπώνυμο του ή της συζύγου (εάν υπάρχει), το ονοματεπώνυμο και το έτος γέννησης των ανήλικων τέκνων (εάν υπάρχουν) και την αιτία του θανάτου. Ειδικότερα για την αιτία του θανάτου γίνεται κωδικοποίηση σύμφωνα με τη *Στατιστική Ταξινόμησης των Νόσων, Κακώσεων και Αιτιών Θανάτου* που συντάσσει και αναθεωρεί περιοδικά η Παγκόσμια Οργάνωση Υγείας (Π.Ο.Υ.).

ΚΕΦΑΛΑΙΟ 2ο

ΠΙΝΑΚΕΣ ΕΠΙΒΙΩΣΗΣ - ΘΝΗΣΙΜΟΤΗΤΑΣ

2.1 ΟΡΙΣΜΟΣ, ΕΙΔΗ & ΣΥΝΑΡΤΗΣΕΙΣ ΠΙΝΑΚΩΝ ΕΠΙΒΙΩΣΗΣ

Το μέσο, μέσω του οποίου, μπορούμε να εκθέσουμε και να επεξεργαστούμε στοιχεία που αφορούν τη θνησιμότητα ονομάζεται Πίνακας Θνησιμότητας ή Πίνακας Επιβίωσης. Δεν υπάρχει σαφής διαχωρισμός των δύο εννοιών, καθώς με βάση τα στοιχεία της θνησιμότητας και του πληθυσμού είναι δυνατόν να υπολογιστούν όλες οι βιομετρικές ποσότητες οι οποίες εκτίθενται σε έναν Πίνακα Επιβίωσης. Κατά τον *London* είναι καθαρά υποκειμενική η επιλογή του εάν κάποιος θα αποκαλεί τον πίνακα, Πίνακα Επιβίωσης ή Θνησιμότητας και εξαρτάται από το εάν το βλέπει από πεσιμιστική σκοπιά ή όχι. Οι Πίνακες Επιβίωσης χωρίζονται σε *γενεαλογικούς* και *χρονολογικούς*.

Για την κατάρτιση των *γενεαλογικών πινάκων* εξετάζουμε μια *γενεά* (δηλαδή μια ομάδα ατόμων που έχουν γεννηθεί το ίδιο έτος, γεγονός το οποίο ορίζουμε ως αρχικό) από τη δημιουργία της, μέχρι την πλήρη εξαφάνισή της (δηλαδή μέχρι να πεθάνει και το τελευταίο μέλος της), θεωρώντας ότι η παραπάνω γενεά είναι *κλειστή* σε μεταναστευτικές κινήσεις (δηλαδή δεν επιτρέπονται μεταναστευτικές εισροές ή εκροές). Αυτό σημαίνει ότι θα πρέπει η γενεά να παρακολουθηθεί για μεγάλο χρονικό διάστημα (ίσως μεγαλύτερο και του αιώνα) και θα πρέπει γι' αυτό το λόγο να έχουμε διαθέσιμες για κάθε έτος τις ληξιαρχικές καταγραφές των θανάτων αυτής της γενεάς. Το κυριότερο πλεονέκτημα των γενεαλογικών πινάκων είναι ότι αναφέρονται σε μια πραγματική γενεά και αποτελούν ένα ευκόλως κατανοητό και επαρκές μέσο αποτύπωσης της διαχρονικής φθοράς που υφίσταται η γενεά λόγω θανάτου. Από την άλλη πλευρά, για την κατάρτιση ενός γενεαλογικού πίνακα απαιτείται μεγάλο χρονικό διάστημα, το στατιστικό υλικό μπορεί να μην είναι σε περιπτώσεις αξιόπιστο ή διαθέσιμο και το μεγαλύτερο μειονέκτημα είναι ότι στο τέλος της μελέτης έχουμε έναν πίνακα ο οποίος αναφέρεται σε μια γενεά που δημιουργήθηκε πριν 100 και πλέον έτη. Συνεπώς, τα αποτελέσματα του πίνακα όσον αφορά τη θνησιμότητα της γενεάς θα έχουν μόνο ιστορική αξία, καθώς τα πρότυπα επιβίωσης θα έχουν μεταβληθεί, όπως επίσης και οι συνήθειες ζωής και η ποιότητα της ιατροφαρμακευτικής περίθαλψης κ.λπ..

Αντίθετα, με τους *χρονολογικούς* Πίνακες Επιβίωσης περιγράφουμε το ιστορικό επιβίωσης (ή θνησιμότητας) μιας *πλασματικής γενεάς* με βάση τα στατιστικά στοιχεία που έχουμε στη διάθεσή μας από τις ληξιαρχικές καταγραφές των θανάτων και τις απογραφές του πληθυσμού. Επομένως, με τους χρονολογικούς πίνακες πετυχαίνουμε την αποτύπωση και την περιγραφή των συνθηκών θνησιμότητας που επικρατούν αυτή τη στιγμή και εξετάζουμε τι θα συνέβαινε πραγματικά σε μια γενεά, με την προϋπόθεση ότι αυτές οι συνθήκες δεν θα μεταβάλλονταν. Σε αυτήν την περίπτωση για την κατάρτιση των χρονολογικών πινάκων οι *υποθέσεις* στις οποίες στηριζόμαστε είναι οι εξής :

- *Η πλασματική γενεά αποτελείται από σταθερό αριθμό γεννήσεων ο οποίος είναι συνήθως δύναμη του δέκα (100, 1000 κ.ο.κ.). Ο αριθμός αυτός λέγεται ρίζα του πίνακα, αποτελεί τον αρχικό πληθυσμό και συμβολίζεται με* l_0.

- *Η γενεά είναι κλειστή σε μεταναστευτικές κινήσεις.*

- *Ο αρχικός πληθυσμός μειώνεται σταδιακά με την πάροδο των ετών σύμφωνα με αμετάβλητα πρότυπα θνησιμότητας. Η ηλικία κατά την οποία πεθαίνει και το τελευταίο μέλος της πλασματικής γενεάς καλείται οριακή ή έσχατη ηλικία και συμβολίζεται με ω.*

- *Οι θάνατοι ισοκατανέμονται κατά τη διάρκεια όλων των ηλικιών πλην των ηλικιών 0 και 1.*

- *Ο συνολικός αριθμός των θανάτων της πλασματικής γενεάς ισούται με τον συνολικό αριθμό των γεννήσεων, δηλαδή τη ρίζα του πίνακα.*

- *Η γενεά περιλαμβάνει μέλη του ίδιου φύλου, λόγω των διαφορών που υπάρχουν ανάμεσα στα πρότυπα θνησιμότητας των ανδρών και των γυναικών.*

Οι Πίνακες Επιβίωσης, στους οποίους θα παρουσιαστούν οι διάφορες συναρτήσεις οι οποίες σχετίζονται με τη θνησιμότητα, μπορεί να είναι είτε *πλήρεις*, είτε *συνεπτυγμένοι*. Στους *πλήρεις πίνακες* οι συναρτήσεις όλες οι συναρτήσεις αναφέρονται σε κάθε ηλικία ξεχωριστά από την ηλικία 0 έως την οριακή ηλικία. Αντίθετα, στους *συνεπτυγμένους πίνακες* οι συναρτήσεις εκφράζονται για ένα διάστημα ηλικιών (το οποίο είναι συνήθως πενταετές), ενώ συνήθως εξαιρούνται οι δύο πρώτες ηλικίες 0 και 1. Ακολουθούν οι ορισμοί των κυριότερων συναρτήσεων οι οποίες παρουσιάζονται σε έναν Πίνακα Επιβίωσης:

- x *είναι η ακριβής ηλικία του ατόμου στην αρχή του έτους.*
- n *είναι το μέγεθος του διαστήματος μεταξύ των ηλικιών* x *και* $x+n$.
- l_0 *είναι το μέγεθος του αρχικού πληθυσμού ή ρίζα του πίνακα.*
- l_x *είναι ο αριθμός των επιζώντων στην αρχή της ηλικίας* x.
- l_{x+n} *είναι ο αριθμός των επιζώντων στην αρχή της ηλικίας* $x+n$.
- d_x *είναι ο αριθμός των θανόντων μεταξύ των ηλικιών* x *και* $x+1$.
- ${}_nd_x$ *είναι ο αριθμός των θανόντων μεταξύ των ηλικιών* x *και* $x+n$.
- p_x *είναι η πιθανότητα άτομο ηλικίας* x *να επιζήσει μέχρι την αρχή της ηλικίας* $x+1$.
- ${}_np_x$ *είναι η πιθανότητα άτομο ηλικίας* x *να επιζήσει μέχρι την αρχή της ηλικίας* $x+n$.
- q_x *είναι η πιθανότητα άτομο ηλικίας* x *να πεθάνει κάποια στιγμή εντός του έτους, δηλαδή πριν φτάσει στην αρχή της ηλικίας* $x+1$.
- ${}_nq_x$ *είναι η πιθανότητα άτομο ηλικίας* x *να πεθάνει κάποια στιγμή εντός των επόμενων* n *ετών, δηλαδή πριν φτάσει στην αρχή της ηλικίας* $x+n$.
- L_x *είναι οι επιζώντες στο μέσο της ηλικίας* x *(δηλαδή στο μέσο του διαστήματος των ηλικιών* x *και* $x+1$ *).*

- $_nL_x$ *είναι οι επιζώντες στο μέσο του ηλικιακού διαστήματος* x *και* $x + n$.

- T_x *είναι ο συνολικός αριθμός των επιζώντων από το μέσο της ηλικίας* x *και άνω ή ο συνολικός αριθμός των ανθρωπο - ετών από το μέσο της ηλικίας* x *και άνω.*

- $\overset{\circ}{e}_x$ *είναι η προσδοκώμενη ή αναμενόμενη ή μέση διάρκεια ζωής ατόμου ηλικίας* x *(δηλαδή ο μέσος αριθμός ετών που αναμένεται να ζήσει ακόμα, άτομο ηλικίας* x *μέχρι το θάνατό του).*

- $\dot{m}_x$ *είναι ο κεντρικός δείκτης θνησιμότητας για την ηλικία* x.

- $_n\dot{m}_x$ *είναι ο κεντρικός δείκτης θνησιμότητας για το διάστημα ηλικιών* x *και* $x + n$.

Επίσης, μια ακόμη πολύ σημαντική συνάρτηση είναι η μ_x *που καλείται ένταση θνησιμότητας στην ηλικία* x *και εκφράζει την στιγμιαία πιθανότητα θανάτου στην ηλικία* x, *δηλαδή την πιθανότητα άτομο που έχει φτάσει στην ηλικία* x *να πεθάνει εντός διαστήματος* dx.

2.2 ΚΑΤΑΣΚΕΥΗ ΠΙΝΑΚΩΝ ΕΠΙΒΙΩΣΗΣ

Τα βήματα για την *κατασκευή* ενός πλήρους ή ενός συνεπτυγμένου πίνακα είναι όμοια και η διαδικασία συνοψίζεται ως εξής :

- *Υπολογισμός κεντρικών δεικτών θνησιμότητας (κατά ηλικία ή διάστημα ηλικιών).*

- *Μετατροπή των κεντρικών δεικτών σε πιθανότητες θανάτου.*

- *Εξομάλυνση των εμπειρικών πιθανοτήτων θανάτου.*

- *Εκτίμηση των ανωτέρω συναρτήσεων του πίνακα.*

Αν λοιπόν, για κάποιον πληθυσμό, δεν υποθέσουμε κάποια συνεχή κατανομή για τη τυχαία μεταβλητή X, $X \geq 0$, *που εκφράζει τη συνολική διάρκεια ζωής ενός νεογέννητου ατόμου, θα έχουμε για τις συναρτήσεις του Πίνακα Επιβίωσης:*

- $\boldsymbol{d_x}$

$$d_x = l_x - l_{x+1} \Leftrightarrow l_{x+1} = l_x - d_x \,.$$

- $\boldsymbol{{}_n d_x}$

$$ {}_n d_x = l_x - l_{x+n} = \sum_{t=0}^{n-1} d_{x+t} \Leftrightarrow l_{x+n} = l_x - {}_n d_x \,.$$

- $\boldsymbol{p_x}$

$$p_x = \frac{l_{x+1}}{l_x}.$$

- $\boldsymbol{{}_n p_x}$

$$ {}_n p_x = \frac{l_{x+n}}{l_x} = \prod_{t=0}^{n-1} p_{x+t} \,.$$

- $\boldsymbol{q_x}$

$$q_x = \frac{d_x}{l_x} = 1 - p_x \,.$$

- $\boldsymbol{{}_n q_x}$

$$ {}_n q_x = 1 - {}_n p_x = \frac{{}_n d_x}{l_x}.$$

- $\boldsymbol{L_x}$

$$L_x = l_x - 0{,}5 * d_x = \frac{l_x + l_{x+1}}{2},$$

(για ηλικίες $x \geq 2$ και υποθέτοντας ομοιόμορφη κατανομή των θανάτων κατά τη διάρκεια ενός έτους)

και

$$L_0 = l_0 - 0{,}75 * d_0$$

και

$$L_1 = l_1 - 0{,}65 * d_1 .$$

- $\boldsymbol{{}_nL_x}$

$$_nL_x = \sum_{t=0}^{n-1} L_{x+t} \simeq \frac{n}{2} * (l_x + l_{x+n}),$$

(η δεύτερη ισότητα χρησιμοποιείται για $x \geq 2$).

- $\boldsymbol{T_x}$

$$T_x = \sum_{t=0}^{\omega-1} L_{x+t} = L_x + T_{x+1}$$

ή

$$T_{x+n} = T_x - {}_nL_x .$$

- $\boldsymbol{\mathring{e}_x}$

$$\mathring{e}_x = \frac{T_x}{l_x}.$$

- $\dot{\boldsymbol{m}}_{\boldsymbol{x}}$

$$\dot{m}_x = \frac{d_x}{L_x}.$$

- ${}_n\dot{\boldsymbol{m}}_{\boldsymbol{x}}$

$${}_n\dot{m}_x = \frac{{}_nd_x}{{}_nL_x},$$

(για ηλικίες $x \geq 2$).

Αν τώρα θεωρήσουμε ότι η τυχαία μεταβλητή Χ, όπως αυτή ορίστηκε παραπάνω, ακολουθεί μια συνεχή κατανομή τότε ορίζουμε τις εξής συναρτήσεις για την τυχαία μεταβλητή Χ :

- $F_X(x) = P_r(X \leq x)$ *η συνάρτηση κατανομής της τ.μ. Χ για* $x \geq 0$, *που εκφράζει την πιθανότητα ένα νεογέννητο άτομο να πεθάνει μέχρι την ηλικία* x.
- $S_X(x) = P_r(X > x) = 1 - F_X(x)$ *η συνάρτηση επιβίωσης της τ.μ. Χ για* $x \geq 0$, *που εκφράζει την πιθανότητα ένα νεογέννητο άτομο να επιζήσει μέχρι την ηλικία* x.
- *Ισχύουν προφανώς οι σχέσεις :* $S_X(0) = 1$, $S_X(\omega) = 0$, $F_X(0) = 0$, $F_X(\omega) = 1$.
- $f_X(x) = \frac{d}{dx}F_X(x) = -\frac{d}{dx}S_X(x)$ *η συνάρτηση πυκνότητας πιθανότητας της τ.μ. Χ για* $x \geq 0$.
- μ_x *η ένταση θνησιμότητας ατόμου ηλικίας* x *για* $x \geq 0$, *που εκφράζει το στιγμιαίο ρυθμό θανάτου του ατόμου στη συγκεκριμένη ηλικία.*

Οι τέσσερις αυτές βασικές συναρτήσεις συνδέονται μεταξύ τους με τις παρακάτω σχέσεις του πίνακα :

	$F_X(x) =$	$S_X(x)=$	$f_X(x) =$	$\mu_x=$
$F_X(x)$	-	$1 - F_X(x)$	$\frac{d}{dx}F_X(x)$	$\frac{\frac{d}{dx}F_X(x)}{1 - F_X(x)}$
$S_X(x)$	$1 - S_X(x)$	-	$-\frac{d}{dx}S_X(x)$	$\frac{-\frac{d}{dx}S_X(x)}{S_X(x)}$
$f_X(x)$	$\int_0^x f_X(t)\,dt$	$1 - \int_0^x f_X(t)\,dt$	-	$\frac{f_X(x)}{1 - \int_0^x f_X(t)\,dt}$
μ_x	$1 - e^{-\int_0^x \mu_t\,dt}$	$e^{-\int_0^x \mu_t\,dt}$	$\mu_x * e^{-\int_0^x \mu_t\,dt}$	-

Με βάση λοιπόν μια συγκεκριμένη συνεχή κατανομή, οι συναρτήσεις του Πίνακα Επιβίωσης θα έχουν ως εξής:

- $\boldsymbol{l_x}$

$$l_x = l_0 * S_X(x).$$

- $\boldsymbol{d_x}$

$$d_x = \int_0^1 l_{x+t} * \mu_{x+t}\,dt\,.$$

- ${}_n\boldsymbol{d_x}$

$${}_n d_x = \int_0^n l_{x+t} * \mu_{x+t}\,dt\,.$$

- $\boldsymbol{p_x}$

$$p_x = e^{-\int_0^1 \mu_{x+t}\,dt}\,.$$

- ${}_n\boldsymbol{p}_{\boldsymbol{x}}$

$$ {}_np_x = e^{-\int_0^n \mu_{x+t}\,dt} = \frac{S_X(x+n)}{S_X(x)}. $$

- $\boldsymbol{q}_{\boldsymbol{x}}$

$$ q_x = \int_0^1 {}_tp_x * \mu_{x+t}\,dt\,. $$

- ${}_n\boldsymbol{q}_{\boldsymbol{x}}$

$$ {}_nq_x = \int_0^n {}_tp_x * \mu_{x+t}\,dt = 1 - \frac{S_X(x+n)}{S_X(x)}. $$

- $\boldsymbol{L}_{\boldsymbol{x}}$

$$ L_x = \int_0^1 l_{x+t}\,dt\,. $$

- ${}_n\boldsymbol{L}_{\boldsymbol{x}}$

$$ {}_nL_x = \int_0^n l_{x+t}\,dt\,. $$

- $\boldsymbol{T}_{\boldsymbol{x}}$

$$ T_x = \int_x^\omega l_t\,dt = \int_0^\omega l_{x+t}\,dt\,. $$

- $\mathring{\boldsymbol{e}}_{\boldsymbol{x}}$

$$ \mathring{e}_x = \int_0^\omega p_{x+t}\,dt\,. $$

- $\dot{m}_x$

$$\dot{m}_x = \frac{\int_0^1 {}_tp_x * \mu_{x+t}\, dt}{\int_0^1 {}_tp_x\, dt}.$$

- ${}_n\dot{m}_x$

$${}_n\dot{m}_x = \frac{\int_0^n {}_tp_x * \mu_{x+t}\, dt}{\int_0^n {}_tp_x\, dt}.$$

Θα πρέπει να αναφερθεί ότι σε περιπτώσεις που δεν υπάρχουν διαθέσιμα ή ακριβή στατιστικά στοιχεία που να αφορούν ληξιαρχικές καταγραφές ή απογραφές, χρησιμοποιούνται *Πρότυποι Πίνακες Επιβίωσης*. Τέλος, θα πρέπει να γίνει αναφορά και στους *Ειδικούς Δημογραφικούς Πίνακες* όπως επί παραδείγματι οι *Πίνακες Επιβίωσης κατά Αιτία Θανάτου* ή οι *Πίνακες Επιβίωσης κατά Οικογενειακή Κατάσταση*, οι οποίοι δεν αναλύονται περαιτέρω, καθώς απέχουν από τους στόχους της παρούσας εργασίας.

ΚΕΦΑΛΑΙΟ 3^{ο}

ΕΞΟΜΑΛΥΝΣΗ ΠΙΝΑΚΩΝ ΕΠΙΒΙΩΣΗΣ

3.1 Η ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΕΞΟΜΑΛΥΝΣΗΣ

Πριν εφαρμόσουμε τη διαδικασία της εξομάλυνσης έχουμε λάβει τις *πρωτογενείς τιμές* ή *αρχικές εκτιμήσεις της θνησιμότητας*. Αυτό έχει πραγματοποιηθεί είτε παίρνοντας τα στοιχεία του πληθυσμού και τις ληξιαρχικές καταγραφές των θανάτων, είτε εφαρμόζοντας ένα πρότυπο - μοντέλο επιβίωσης σε έναν υποκείμενο πληθυσμό. Στη συνέχεια έχοντας τις πρωτογενείς τιμές της θνησιμότητας εφαρμόζουμε μια *μέθοδο εξομάλυνσης* που αποσκοπεί στην *αναθεώρηση των πρωτογενών τιμών*. Κρίσιμο ερώτημα αποτελεί το *γιατί πρέπει να εξομαλύνουμε τις αρχικές μας εκτιμήσεις της θνησιμότητας* και να μην δεχτούμε ότι αυτές ακριβώς οι εκτιμήσεις αντικατοπτρίζουν την πραγματική θνησιμότητα στην οποία εκτίθεται ο υπό μελέτη πληθυσμός μας; *Η απάντηση είναι ότι προβαίνουμε στην εφαρμογή μεθόδων εξομάλυνσης, διότι αφενός μεν επιθυμούμε την εξάλειψη του στατιστικού λάθους λόγω του μη επαρκούς δείγματος, αφετέρου δε διότι επιθυμούμε τη διόρθωση ασυνεπειών που διαταράσσουν την ομαλή απεικόνιση της καμπύλης θνησιμότητας (στην περίπτωση που έχουμε στατιστικά στοιχεία από απογραφές και ληξιαρχικές καταγραφές) και προέρχονται είτε από διαφυγές δημογραφικών γεγονότων (κυρίως στις μικρές ηλικίες), είτε από ανακριβείς δηλώσεις ηλικιών (στις μεγάλες ηλικίες).*

Η ίδια λοιπόν η φύση των στοιχείων μας ωθεί στο να προβούμε σε εξομάλυνση. Υπό αυτή την έννοια, αναφερόμενοι πάντα σε εκτιμήσεις θνησιμότητας, *ο κύριος παράγοντας με βάση τον οποίο αποφασίζουμε να εξομαλύνουμε τις αρχικές μας εκτιμήσεις είναι η πρότερη γνώση - γνώμη (prior opinion) που έχουμε για το φαινόμενο που εξετάζουμε, δηλαδή τη θνησιμότητα.* Δεν μπορούμε συνεπώς να παραβλέψουμε την σχέση εξάρτησης που υπάρχει μεταξύ “γειτονικών” τιμών θνησιμότητας σε διάφορες ηλικίες. Συνεπώς, προσπαθούμε να αφαιρέσουμε τα τυχαία σφάλματα που παρουσιάζονται στην καμπύλη θνησιμότητας που προέρχεται από τις πρωταρχικές μας εκτιμήσεις, έχοντας υπόψη την πρότερη γνώση που έχουμε για το φαινόμενο της θνησιμότητας, η οποία έγκειται κυρίως στα εξής :

- *Η θνησιμότητα είναι αύξουσα συνάρτηση της ηλικίας.*

- *Στις δύο πρώτες ηλικίες παρατηρείται αυξημένη θνησιμότητα η οποία ομαλοποιείται κατά τις επόμενες παιδικές ηλικίες.*

- *Παρουσιάζεται μια αύξηση της θνησιμότητας στους άνδρες ηλικίας 18 – 27 (λόγω ατυχημάτων με το αυτοκίνητο κλπ.) η οποία καλείται accident hump.*

- *Η θνησιμότητα παρουσιάζει βαθμιαία αυξητική τάση από την ηλικία των 45 ετών και άνω και για τα δύο φύλα.*

- *Η θνησιμότητα αυξάνεται ραγδαία κατά τις γεροντικές ηλικίες και για τα δύο φύλα.*

- *Οι διαφορές μεταξύ των επιπέδων θνησιμότητας ανάμεσα στα δύο φύλα αμβλύνονται προοδευτικά μετά την ηλικία των 55 ετών.*

Οι παραπάνω παραδοχές (πλην της 1ης) ισχύουν και σε μικρότερο ή μεγαλύτερο βαθμό η κάθε μία, ανάλογα με τη χώρα προέλευσης του υποκείμενου πληθυσμού (π.χ. προφανώς η 3η παραπάνω παραδοχή θα έχει αρκετά μεγάλο αντίκτυπο σε στοιχεία που αφορούν την Ελλάδα).

Άρα λοιπόν, εξομαλύνουμε τις αρχικές εκτιμήσεις της θνησιμότητάς μας, όχι μόνο χρησιμοποιώντας την πρότερη άποψη και γνώση την οποία έχουμε σχετικά με την ακολουθία των τιμών, αλλά και λαμβάνοντας υπόψη και τις ίδιες τις αρχικές μας εκτιμήσεις. *Σημαντικό είναι να κατανοήσουμε ότι και οι εξομαλυμένες ή αναθεωρημένες τιμές οι οποίες θα προκύψουν μετά από την εφαρμογή κάποιας μεθόδου εξομάλυνσης, αποτελούν και πάλι μια (βελτιωμένη) απεικόνιση της πραγματικής θνησιμότητας του υποκείμενου πληθυσμού, η οποία ποτέ δεν θα μας είναι γνωστή στο ακέραιο.* Συνοψίζοντας, έχουμε ότι τα *βήματα της διαδικασίας της εξομάλυνσης* είναι τα εξής (από στατιστικά στοιχεία) :

- *Λαμβάνουμε τα στοιχεία του πληθυσμού και των θανάτων.*

- *Εξάγουμε τις πρωτογενείς τιμές θνησιμότητας.*

- *Εφαρμόζουμε μια μέθοδο εξομάλυνσης και παίρνουμε τις εξομαλυμένες τιμές θνησιμότητας, οι οποίες αποτελούν μια βελτιωμένη προσέγγιση της θνησιμότητας του υπό μελέτη πληθυσμού, αλλά εξακολουθούν να αποτελούν προσέγγιση και όχι τις πραγματικές τιμές της θνησιμότητας.*

Αν θέλουμε να δώσουμε ένα γενικό πλαίσιο που διέπει μια επιτυχημένη εξομάλυνση (για δεδομένα που αφορούν θνησιμότητα πάντα) θα λέγαμε ότι :

Εξομάλυνση (Graduation) είναι η διαδικασία αναθεώρησης των πρωτογενών τιμών θνησιμότητας που αποσκοπεί στην εξάλειψη των τυχαίων σφαλμάτων, στη δημιουργία μιας πιο ομαλής καμπύλης και στην όσο το δυνατόν εγγύτερης προσέγγισης της πραγματικής θνησιμότητας. Ουσιαστικά επιτυγχάνουμε μια καλή εξομάλυνση ισορροπώντας ανάμεσα στην έμφαση στην ομαλότητα (δηλαδή στη δημιουργία μιας πιο ομαλής και "λείας" καμπύλης που να μην παρουσιάζει σημαντικά "άλματα" σε διάφορες ηλικίες) και στην προσέγγιση των πρωτογενών τιμών (δηλαδή στο να μην απομακρυνόμαστε ιδιαίτερα από τις πρωτογενείς τιμές).

Από την παραπάνω ανάλυση κατανοούμε ότι τα δύο βασικά μέτρα που αφορούν την εξομάλυνση είναι :

- *Η ομαλότητα (Smoothness) .*
- *Η καλή εφαρμογή (Fit to data – Fidelity).*

Πριν αναλύσουμε τα δύο παραπάνω μέτρα είναι χρήσιμο να δώσουμε τους συμβολισμούς , καθώς και το βασικό στατιστικό και μαθηματικό υπόβαθρο που διέπει τη διαδικασία της εξομάλυνσης. Σημειώνεται ότι ακολουθείται ο συμβολισμός για μέτρα που σχετίζονται με την εξομάλυνση όπως αυτός ορίστηκε από τον *London* (1985).

3.2 ΒΑΣΙΚΟΣ ΣΥΜΒΟΛΙΣΜΟΣ

- X *η τυχαία μεταβλητή που συμβολίζει την ηλικία ενός μέλους του πληθυσμού.*
- x *η πραγματική ηλικία στην οποία βρίσκεται ένα μέλος του πληθυσμού.*
- t_x *η πραγματική πιθανότητα (που αναφέρεται στην πραγματική θνησιμότητα), άτομο ηλικίας* x *να πεθάνει κάποια στιγμή εντός του επόμενου έτους (μέχρι την αρχή της ηλικίας* $x+1$ *).*
- n_x *το πλήθος του δείγματος ηλικίας* x*, δηλαδή το σύνολο των ζώντων μελών του πληθυσμού ηλικίας* x.
- H_x *η τυχαία μεταβλητή του αριθμού των θανόντων ηλικίας* x *(δηλαδή από την αρχή της ηλικίας* x *έως την αρχή της ηλικίας* $x+1$ *).*
- h_x *ο παρατηρούμενος αριθμός των θανόντων μεταξύ των ηλικιών* x *και* $x+1$.
- U_x *η τυχαία μεταβλητή της πρωτογενούς τιμής θνησιμότητας / πιθανότητας θανάτου ατόμου ηλικίας* x */ αναλογίας θανόντων ηλικίας* x *στο πλήθος του δείγματος ηλικίας* x.
- u_x *η πραγματική τιμή της πρωτογενούς πιθανότητας θανάτου για άτομο ηλικίας* x.
- V_x *η τυχαία μεταβλητή της εξομαλυμένης πιθανότητας θανάτου για άτομο ηλικίας* x.
- v_x *η πραγματοποίηση της τυχαίας μεταβλητής* V_x *ή πραγματική εξομαλυμένη πιθανότητα θανάτου ατόμου ηλικίας* x.
- w_x *η τιμή του βάρους που χρησιμοποιείται για τη στάθμιση διαφόρων μέτρων που αφορούν την ηλικία* x.
- E_x *η τυχαία μεταβλητή του σφάλματος εκτίμησης της πιθανότητας θανάτου για την ηλικία* x.

- e_x *το πραγματικό σφάλμα εκτίμησης της πιθανότητας θανάτου για την ηλικία* x.

- E'_x *η τυχαία μεταβλητή του σφάλματος εκτίμησης που αντιστοιχεί στην εξομαλυμένη πιθανότητα θανάτου για άτομο ηλικίας* x.

- e'_x *το πραγματικό σφάλμα εκτίμησης της εξομαλυμένης πιθανότητας θανάτου για άτομο ηλικίας* x.

3.3 ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΕΞΟΜΑΛΥΝΣΗΣ

Θεωρούμε καταρχήν ότι η τυχαία μεταβλητή H_x ακολουθεί τη *διωνυμική κατανομή* με :

$$H_x \sim Binomial(n_x, t_x),$$

όπου

$$E[H_x] = n_x * t_x$$

και

$$var[H_x] = n_x * t_x * (1 - t_x) .$$

Θεωρούμε στη συνέχεια την τυχαία μεταβλητή U_x που καλείται *διωνυμική αναλογία*, ως το πηλίκο της τυχαίας μεταβλητής των θανόντων προς το σύνολο των ζώντων μελών, δηλαδή :

$$U_x = \frac{H_x}{n_x}.$$

Αν θέλουμε να την εκφράσουμε με βάση την πραγματική πιθανότητα θανάτου t_x και την τυχαία μεταβλητή του σφάλματος εκτίμησης έχουμε :

$$U_x = t_x + E_x .$$

Προφανώς ισχύουν τα εξής :

$$E[U_x] = t_x \, ,$$

δηλαδή η U_x είναι αμερόληπτη εκτιμήτρια της t_x και

$$var[U_x] = \frac{t_x * (1 - t_x)}{n_x} \, .$$

Όσο *μεγαλώνει το δείγμα*, τόσο *μικραίνει η διασπορά της* U_x, συνεπώς μπορούμε να την προσεγγίσουμε με *κανονική κατανομή* για μεγάλο δείγμα (αριθμό ζώντων ηλικίας x δηλαδή) γράφοντας :

$$U_x \sim Normal \left(t_x \, , \frac{t_x * (1 - t_x)}{n_x} \right)$$

ή

$$\frac{U_x - t_x}{\sqrt{\frac{t_x * (1 - t_x)}{n_x}}} \sim Normal(0,1) \, ,$$

όπου $Normal(0,1)$ η τυπική κανονική κατανομή.

Για την πρωτογενή τιμή θνησιμότητας δηλαδή την πιθανότητα θανάτου για άτομο ηλικίας x γράφουμε :

$$u_x = \frac{h_x}{n_x} = \hat{t}_x$$

ή αν θέλουμε να την εκφράσουμε με βάση την πραγματική πιθανότητα θανάτου :

$$u_x = t_x + e_x \, .$$

Ακόμη θα ισχύουν οι παρακάτω σχέσεις :

$$E[E_x] = 0$$

και

$$var[E_x] = var[U_x] = \frac{t_x * (1 - t_x)}{n_x}.$$

Επίσης, ενδιαφέρον παρουσιάζει η μελέτη της τυχαίας μεταβλητής N που εκφράζει τις *αλλαγές προσήμου* που παρατηρούνται μετά την εξομάλυνση, είτε στη διαφορά :

$$u_x - v_x ,$$

είτε στη διαφορά :

$$v_x - u_x .$$

Όλοι οι πιθανοί συνδυασμοί αυτής της διαφοράς, εξετάζοντάς τη μεταξύ *δύο διαδοχικών ηλικιών*, είναι οι εξής :

ΠΡΟΣΗΜΟ ΔΙΑΦΟΡΑΣ 1ου ΖΕΥΓΟΥΣ ΗΛΙΚΙΩΝ	ΠΡΟΣΗΜΟ ΔΙΑΦΟΡΑΣ 2ου ΖΕΥΓΟΥΣ ΗΛΙΚΙΩΝ
+	+
+	-
-	+
-	-

Άρα, η *πιθανότητα αλλαγής προσήμου της παραπάνω διαφοράς μεταξύ δύο διαδοχικών ηλικιών* είναι $\frac{1}{2}$ και εάν έχουμε k παρατηρήσεις συνολικά, ο μέγιστος αριθμός μεταβολών προσήμου που μπορούμε να παρατηρήσουμε είναι $k-1$. Άρα, η τυχαία μεταβλητή N ακολουθεί τη *διωνυμική κατανομή* με :

$$N \sim Binomial(\, k-1\,, \frac{1}{2})$$

και

$$E[N] = \frac{k-1}{2}$$

και

$$var[N] = \frac{k-1}{4}.$$

Σε πολλές περιπτώσεις χρησιμοποιούμε τις τιμές θνησιμότητας από έναν άλλον Πίνακα Θνησιμότητας, αρκεί ο πληθυσμός αυτού του πίνακα να έχει κοινά χαρακτηριστικά με τον υπό εξέταση πληθυσμό μας. Σε αυτή την περίπτωση ορίζουμε τα εξής :

- *s_x η (γνωστή) πιθανότητα θανάτου ατόμου ηλικίας x εντός του επόμενου έτους, όπως αυτή λαμβάνεται από τον γνωστό Πίνακα Θνησιμότητας.*
- *R_x η τυχαία μεταβλητή του λόγου της τυχαίας μεταβλητής της πρωτογενούς πιθανότητας θανάτου προς την πιθανότητα θανάτου του γνωστού πίνακα.*
- *r_x η τιμή του λόγου της πρωτογενούς πιθανότητας θανάτου προς την πιθανότητα θανάτου του γνωστού πίνακα.*

Ισχύουν οι εξής σχέσεις :

$$R_x = \frac{U_x}{s_x}$$

ή

$$r_x = \frac{u_x}{s_x},$$

με

$$var[R_x] = \frac{t_x * (1 - t_x)}{n_x * {s_x}^2} \approx \frac{u_x * (1 - u_x)}{n_x * {s_x}^2}.$$

Η παραπάνω τεχνική χρησιμοποιείται διότι μπορεί να επιθυμούμε να ενισχύσουμε την έννοια της πρότερης γνώσης όσον αφορά τις τιμές της θνησιμότητας που θέλουμε να εκτιμήσουμε. Δηλαδή, εισάγοντας πρόσθετη πληροφορία που προέρχεται από τις γνωστές τιμές θνησιμότητας, καταφέρνουμε να αναπαράγουμε το πρότυπο της θνησιμότητας του γνωστού πίνακα το οποίο θα αποτυπώνεται στις εξομαλυμένες τιμές θνησιμότητας. Βέβαια, είναι αυτονόητο ότι τα χαρακτηριστικά αλλά και η πρότερη γνώμη πάνω στα οποία βασίστηκε η κατάρτιση του γνωστού πίνακα, πρέπει να προσομοιάζουν με τα χαρακτηριστικά του πληθυσμού που μελετάμε εμείς. Σε αντίθετη περίπτωση η όλη διαδικασία μπορεί να μην παράγει τα επιθυμητά αποτελέσματα.

Τέλος, αναφέρουμε ότι κατά τις πρώτες απόπειρες βελτίωσης των εξομαλυμένων τιμών χρησιμοποιήθηκε *γραμμικός μετασχηματισμός* για την εύρεση ακόμη αποτελεσματικότερων τιμών. Πιο συγκεκριμένα, ορίστηκαν οι νέες βελτιωμένες εξομαλυμένες τιμές να έχουν την εξής μορφή :

$$v'_x = a * v_x + b\,.$$

Ζητούνταν να ικανοποιούνται οι παρακάτω συνθήκες :

$$\sum_x n_x * (v'_x - u_x) = 0$$

και

$$\sum_x x * n_x * (v'_x - u_x) = 0\,.$$

Έτσι καταλήγουμε σε δύο σχέσεις με δύο αγνώστους απ' όπου επιλύοντας το σύστημα βρίσκουμε τις τιμές των a και b. Εν συνεχεία, υπολογίζουμε τις βελτιωμένες εξομαλυμένες τιμές των πιθανοτήτων θανάτου για κάθε ηλικία.

3.4 ΠΕΠΕΡΑΣΜΕΝΕΣ ΔΙΑΦΟΡΕΣ (FINITE DIFFERENCES)

Πριν εξετάσουμε το μέτρο ομαλότητας είναι χρήσιμο να παρατεθούν κάποιες βασικές σχέσεις που αφορούν τις *διαφορές κ – τάξης*, όπου το *κ* είναι *μη αρνητικός* και *ακέραιος* αριθμός. Για μια συνάρτηση $f(x)$ ορίζουμε και συμβολίζουμε την *πρώτη διαφορά με βήμα* h ως εξής :

$$\Delta_h f(x) = f(x+h) - f(x) .$$

Αν $h = 1$ τότε γράφουμε :

$$\Delta f(x) = f(x+1) - f(x) .$$

Με βήμα $h = 1$ θα έχουμε για τη *δεύτερη διαφορά* :

$$\begin{aligned} \Delta^2 f(x) = \Delta[\Delta f(x)] &= \Delta[f(x+1) - f(x)] = \Delta[f(x+1)] - \Delta[f(x)] \\ &= [f(x+2) - f(x+1)] - [f(x+1) - f(x)] \\ &= f(x+2) - 2 * f(x+1) + f(x) . \end{aligned}$$

Με βήμα $h = 1$ και την ίδια λογική, η *τρίτη διαφορά* θα είναι :

$$\Delta^3 f(x) = f(x+3) - 3 * f(x+2) + 3 * f(x+1) - f(x) .$$

Κάποιες βασικές παρατηρήσεις είναι οι εξής :

- *Η κ τη τάξη διαφορά θα αποτελείται από κ + 1 όρους στην περίπτωση που υπολογίζεται με βάση μια συνάρτηση.*
- *Στην περίπτωση που έχουμε μία ακολουθία έστω n υπολογισμένων τιμών και θέλουμε να υπολογίσουμε τη z τη τάξη διαφορά, τότε η z τη τάξη διαφορά θα περιλαμβάνει n – z στοιχεία.*
- *Αν η* $f(x)$ *είναι πολυώνυμο βαθμού n τότε η n τη τάξη διαφορά θα είναι σταθερός αριθμός ή η n + 1 τη τάξη διαφορά θα είναι μηδενική.*

Δηλαδή :

$$\Delta^{n} f(x) = c \, ,$$

όπου c σταθερά ή

$$\Delta^{n+1} f(x) = 0 \, .$$

- *Αν n: θετικός και ακέραιος, τότε ισχύει ότι :* $\Delta^{n} x^{n} = n!$.
- $\Delta[f(x) + g(x)] = \Delta[f(x)] + \Delta[g(x)]$.
- $\Delta[c * f(x)] = c * \Delta[f(x)]$.

3.5 ΒΑΣΙΚΑ ΜΕΤΡΑ ΕΞΟΜΑΛΥΝΣΗΣ

3.5.1 ΟΜΑΛΟΤΗΤΑ

Η *ομαλότητα* (δηλαδή το πόσο *"λεία"* ή *ομαλή* είναι η καμπύλη θνησιμότητας που θα προκύψει με την εξομάλυνση) είναι ένα από τα δύο *μέτρα εξομάλυνσης* τα οποία μας δίνουν τη δυνατότητα να εξετάσουμε την αποτελεσματικότητα της εφαρμοσμένης μεθόδου. Ο υπολογισμός ενός μέτρου ομαλότητας από μόνος του δεν σημαίνει τίποτα, γι' αυτό μας ενδιαφέρει κυρίως η σύγκριση μέτρων ομαλότητας διαφορετικών μεθόδων εξομάλυνσης που εφαρμόστηκαν πάνω στα ίδια δεδομένα. Τα μέτρα ομαλότητας υπολογίζονται ως άθροισμα των διαφορών κάποιας τάξης (συνήθως τρίτης ή τέταρτης) των εξομαλυμένων τιμών των πιθανοτήτων θανάτου. Συνεπώς, τα συνηθέστερα μέτρα ομαλότητας που υπολογίζουμε είναι τα εξής :

$$S = \sum_{x} (\Delta^{3} v_{x})^{2} \qquad \text{ή} \qquad S = \sum_{x} (\Delta^{4} v_{x})^{2} \, .$$

Επιδιώκουμε την *ελαχιστοποίηση* των παραπάνω μέτρων, συνεπώς το να υπολογίσουμε ένα μέτρο ομαλότητας χρησιμοποιώντας π.χ. τις διαφορές τέταρτης τάξης έχει την ακόλουθη λογική :

Χρησιμοποιώντας διαφορές τέταρτης τάξης ουσιαστικά υποθέτουμε ότι η πραγματική θνησιμότητα (συνεπώς και οι πρωτογενείς τιμές και οι εξομαλυμένες) προσεγγίζεται από ένα πολυώνυμο τρίτου βαθμού, διότι έτσι θα μηδενίζεται το μέτρο

$$S = \sum_x (\Delta^4 v_x)^2 \ ,$$

σύμφωνα με τις ιδιότητες των διαφορών. Άρα, το να θέλουμε να ελαχιστοποιήσουμε το παραπάνω μέτρο ομαλότητας είναι αντίστοιχο με το να θέλουμε οι εξομαλυμένες μας τιμές να ικανοποιούν ένα πολυώνυμο τρίτου βαθμού. Ανάλογος είναι ο συλλογισμός αν επιλέξουμε διαφορές άλλης τάξης για τις εξομαλυμένες τιμές.

3.5.2 ΚΑΛΗ ΕΦΑΡΜΟΓΗ

Οι πρωτογενείς τιμές της θνησιμότητας που λαμβάνουμε αποτελούν μια αρχική εκτίμηση της πραγματικής θνησιμότητας που διέπει τον υπό εξέταση πληθυσμό. Συνεπώς, η βελτίωση τους (δηλαδή οι εξομαλυμένες τιμές) είναι λογικό να θέλουμε να μην απέχουν αισθητά από τις αρχικές μας εκτιμήσεις. Βέβαια, αναμένουμε, σε περιπτώσεις που το δείγμα σε μια ηλικία είναι αρκετά μεγάλο, οι εξομαλυμένες τιμές να μην απέχουν ιδιαίτερα από τις αρχικές τους εκτιμήσεις. Γι' αυτό και δεν αποκλείουμε σε κάποιες περιπτώσεις οι πρωτογενείς τιμές να ισούνται με τις εξομαλυμένες. Βέβαια, στην περίπτωση που όλες οι πρωτογενείς τιμές θνησιμότητας ισούνται με τις εξομαλυμένες, παύει να αποκτά έννοια η διαδικασία της εξομάλυνσης. Επιθυμούμε λοιπόν την *ελαχιστοποίηση* των *μέτρων καλής εφαρμογής*. Τα συνηθέστερα μέτρα καλής εφαρμογής των εξομαλυμένων τιμών είναι τα εξής :

$$F_1 = \sum_x w_x * (v_x - u_x)$$

ή

$$F_2 = \sum_x w_x * (v_x - u_x)^2$$

ή

$$F_3 = \sum_x x * w_x * (v_x - u_x)^2 \ .$$

Το μέτρο F_1 δεν αποτελεί ασφαλές μέτρο όσον αφορά την καλή εφαρμογή, διότι μπορεί να υπάρξουν μεγάλες αρνητικές και θετικές αποκλίσεις, τέτοιες ώστε να αλληλοεξουδετερωθούν και να προκύψει μηδενικό επί παραδείγματι το εν λόγω μέτρο, οδηγώντας μας σε εσφαλμένα συμπεράσματα σχετικά με την απόκλιση των εξομαλυμένων τιμών από τις πρωτογενείς. *Συνήθως ως βάρος στάθμισης σε κάθε ηλικία χρησιμοποιούμε είτε το πλήθος των ζώντων ατόμων, είτε το αντίστροφο της διασποράς της τυχαίας μεταβλητής* U_x *, η οποία υπολογίζεται προσεγγιστικά χρησιμοποιώντας τις εξομαλυμένες τιμές.* Δηλαδή :

$$w_x = n_x$$

ή

$$w_x = \frac{1}{var[U_x]} \approx \frac{n_x}{v_x * (1 - v_x)}.$$

3.6 ΜΕΘΟΔΟΙ ΕΞΟΜΑΛΥΝΣΗΣ

Πριν αναλυθούν τα *στατιστικά test* που αφορούν τον *έλεγχο της εξομάλυνσης*, κρίνεται χρήσιμο να γίνει μια σύντομη αναφορά στα *είδη των μεθόδων εξομάλυνσης*. Οι μέθοδοι που αφορούν την εξομάλυνση χωρίζονται σε δύο είδη :

- *Στις μη - παραμετρικές μεθόδους εξομάλυνσης.*
- *Στις παραμετρικές μεθόδους εξομάλυνσης.*

Στην περίπτωση των *μη - παραμετρικών μεθόδων* δεν θεωρούμε ότι οι πρωτογενείς τιμές της θνησιμότητας ή της έντασης θνησιμότητας προέρχονται από κάποιο γνωστό παραμετρικό μοντέλο ή νόμο θνησιμότητας. Με συγκεκριμένες τεχνικές τις οποίες εφαρμόζουμε στα δεδομένα του πίνακα, προκύπτουν οι εξομαλυμένες τιμές και με βάση το πώς ενσωματώνουμε κάποια συγκεκριμένα κριτήρια στη μελέτη μας, επιτυγχάνουμε το βαθμό ομαλότητας που επιθυμούμε ο οποίος δεν θεωρείται δεδομένος.

Από την άλλη, στην περίπτωση των *παραμετρικών μεθόδων*, εφαρμόζουμε μια γνωστή μαθηματική συνάρτηση στα πρωτογενή δεδομένα θνησιμότητας, εκτιμώντας τις παραμέτρους των συναρτήσεων μέσω της μεθόδου ελαχίστων τετραγώνων, της μεγίστης πιθανοφάνειας κ.λπ.. Μία συνάρτηση μπορεί να εφαρμοστεί είτε σε ολόκληρο το εύρος των ηλικιών του πίνακα, είτε μπορούμε να εφαρμόσουμε διαφορετικές συναρτήσεις σε επί μέρους τμήματα – εύρη ηλικιών του πίνακα.

Στις παραμετρικές μεθόδους εφόσον χρησιμοποιούνται συναρτήσεις γνωστές, δεν υπάρχει νόημα να αναφερόμαστε στο κριτήριο της ομαλότητας η οποία είναι δεδομένη, βάση της ήδη επιλεγμένης συνάρτησης.

Μια σημαντική διαφορά μεταξύ των δύο μεθόδων πέρα από το κριτήριο της ομαλότητας είναι ότι στις παραμετρικές μεθόδους, θα πρέπει να δοθεί ιδιαίτερη έμφαση στο κριτήριο της καλής εφαρμογής, καθώς καμία συνάρτηση δεν μπορεί να παράγει εξομαλυμένες τιμές που να προσεγγίζουν με απόλυτη ακρίβεια τις πρωτογενείς τιμές θνησιμότητας, ενώ αντίθετα στις μη - παραμετρικές μεθόδους η καλή εφαρμογή μεταξύ των πρωτογενών και των εξομαλυμένων τιμών μπορεί να μειωθεί με διάφορες τεχνικές, όπως για παράδειγμα με τη μείωση της διασποράς της τυχαίας μεταβλητής U_x αν τη λάβουμε ως βάρος στάθμισης για κάποια ηλικία.

3.7 ΣΤΑΤΙΣΤΙΚΟΙ ΕΛΕΓΧΟΙ ΕΞΟΜΑΛΥΝΣΗΣ

Αφού ολοκληρωθεί η εφαρμογή κάποιας μεθόδου εξομάλυνσης, μπορούμε να εφαρμόσουμε κάποια *test* τα οποία θα αφορούν τον *έλεγχο της εξομάλυνσης. Δηλαδή, θέλουμε να αξιολογήσουμε αν και σε τι βαθμό η εφαρμοσμένη μέθοδος εξομάλυνσης παράγει ικανοποιητικά αποτελέσματα. Συνεπώς, ως μηδενική υπόθεση, για τους ελέγχους υποθέσεων που παρουσιάζονται παρακάτω, θεωρούμε την υπόθεση ότι η εφαρμοσμένη μέθοδος εξομάλυνσης παράγει ικανοποιητικά αποτελέσματα. Σε περίπτωση που δεν ικανοποιούνται οι συνθήκες που τίθενται στα παρακάτω test, απορρίπτουμε την μηδενική μας υπόθεση, δηλαδή την εφαρμοσμένη μέθοδο εξομάλυνσης.* Τα κυριότερα test που αφορούν τον έλεγχο μιας εφαρμοσμένης μεθόδου εξομάλυνσης είναι τα εξής :

- ***X^2 test***

Έχοντας εφαρμόσει μια μέθοδο εξομάλυνσης θεωρούμε ότι οι εξομαλυμένες τιμές που έχουμε υπολογίσει, προσεγγίζουν αυτές που αφορούν την πραγματική θνησιμότητα. Έτσι, για την τυχαία μεταβλητή H_x θα έχουμε ότι :

$$H_x \sim Binomial(n_x, v_x)$$

με

$$E[H_x] = n_x * v_x$$

και

$$var[H_x] = n_x * v_x * (1 - v_x).$$

Θα πρέπει για κάθε ηλικία να ικανοποιείται η συνθήκη $n_x * v_x \geq 5$, ώστε να μπορέσουμε να προσεγγίσουμε με μια κανονική κατανομή και να γράψουμε :

$$\frac{H_x - n_x * v_x}{\sqrt{n_x * v_x * (1 - v_x)}} \sim Normal(0,1).$$

Αν δεν ισχύει η παραπάνω συνθήκη τότε ομαδοποιούμε όσες ηλικίες χρειάζεται ώστε να ικανοποιείται η παραπάνω συνθήκη και τις λαμβάνουμε ως μια παρατήρηση, μειώνοντας αντιστοίχως τους βαθμούς ελευθερίας. Υπολογίζουμε για κάθε ηλικία x, $x = 1,2,3, \ldots, n$ την ποσότητα :

$$Z_x = \frac{h_x - n_x * v_x}{\sqrt{n_x * v_x * (1 - v_x)}}.$$

Άρα, υπολογίζοντας την ποσότητα,

$$X^2 = \sum_{x=1}^{n} {Z_x}^2$$

έχουμε ότι :

$$X^2 \sim {X^2}_{k,a} ,$$

όπου k είναι οι *βαθμοί ελευθερίας* της X^2 κατανομής ανάλογα με το πόσο αυστηροί είμαστε και πόσες παραμέτρους έχουμε εκτιμήσει και a το *αντίστοιχο επίπεδο σημαντικότητας* με το οποίο πραγματοποιούμε τον έλεγχο. Θα θεωρήσουμε και για τους επόμενους ελέγχους για το επίπεδο σημαντικότητας ότι $a = 5\%$. Για τους βαθμούς ελευθερίας παίρνουμε αντίστοιχα :

- Στην περίπτωση των *μη - παραμετρικών μεθόδων εξομάλυνσης* :

$$k = n \qquad \text{ή} \qquad k = n - 1 .$$

- Στην περίπτωση των *παραμετρικών μεθόδων εξομάλυνσης* :

$$k = n - πλήθος\ των\ παραμέτρων\ που\ εκτιμούμε$$

ή

$$k = n - 1 - πλήθος\ των\ παραμέτρων\ που\ εκτιμούμε .$$

Απορρίπτουμε την εφαρμοσμένη μέθοδο εξομάλυνσης ως ακατάλληλη σε επίπεδο σημαντικότητας a, δηλαδή τη μηδενική μας υπόθεση, όταν :

$$X^2 \succ {X^2}_{k,a} ,$$

όπου η τιμή της ποσότητας ${X^2}_{k,a}$ βρίσκεται από τους πίνακες της X^2 κατανομής.

- ***Έλεγχος προσήμων (Signs test)***

Για όλες τις ηλικίες x, $x = 1,2,3,\dots,n$ *ορίζουμε την τυχαία μεταβλητή* N_p *ως τη μεταβλητή που εκφράζει των αριθμό των θετικών αποκλίσεων μεταξύ των διαφορών* $h_x - n_x * v_x$. Η πιθανότητα αλλαγής προσήμου μεταξύ των διαφορών δύο διαδοχικών ηλικιών είναι $\frac{1}{2}$, όπως αναλύθηκε παραπάνω. Για τις συνολικές παρατηρήσεις πρέπει $n \succ 20$, έτσι ώστε να έχουμε τα εξής :

$$N_p \sim Binomial\left(n , \frac{1}{2} \right),$$

με

$$E[N_p] = \frac{n}{2}$$

και

$$var[N_p] = \frac{n}{4}.$$

Προσεγγίζοντας με την τυπική κανονική κατανομή έχουμε :

$$\frac{N_p - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \sim Normal(0,1).$$

Υπολογίζουμε την ποσότητα :

$$X = \frac{n_p - \frac{n}{2}}{\frac{\sqrt{n}}{2}}.$$

Ενδιαφερόμαστε η ποσότητα n_p όπως προκύπτει από την εξομάλυνση να μην είναι ούτε πολύ μικρή, ούτε πολύ μεγάλη.

Για επίπεδο σημαντικότητας $a = 5\%$ και $n \prec 20$, απορρίπτουμε την εφαρμοσμένη μέθοδο εξομάλυνσης αν η πιθανότητα να παρατηρήσουμε τόσα ή λιγότερα θετικά (ή αρνητικά) πρόσημα είναι μικρότερη του 2,5% ή αν η πιθανότητα να παρατηρήσουμε τόσα ή περισσότερα είναι μικρότερη του 2,5% . Δηλαδή υπολογίζουμε με βάση το δεδομένο n_p που βρίσκουμε μετά την εξομάλυνση την πιθανότητα (μέσω της διωνυμικής κατανομής) :

$$P_r(N_p \geq n_p) \qquad \text{ή} \qquad P_r(N_p \leq n_p)$$

και απορρίπτουμε την εφαρμοσμένη μέθοδο εξομάλυνσης, δηλαδή τη μηδενική μας υπόθεση, εάν :

$$P_r(N_p \geq n_p) \prec 2{,}5\% \qquad \text{ή} \qquad P_r(N_p \leq n_p) \prec 2{,}5\%.$$

Αν προσεγγίζουμε με την τυπική κανονική κατανομή έχοντας $n \succ 20$, τότε δεν απορρίπτουμε την εφαρμοσμένη μέθοδο εξομάλυνσης για επίπεδο σημαντικότητας $a = 5\%$, εάν :

$$-1{,}96 \prec X = \frac{n_p - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \prec 1{,}96\,.$$

- ***Έλεγχος κανονικότητας αποκλίσεων (Individual standardized deviations)***

Σε αυτόν τον έλεγχο μας ενδιαφέρει ο εντοπισμός μεγάλων διαφορών μεταξύ των παρατηρούμενων και των κανονικοποιημένων αποκλίσεων. Με βάση την προσέγγιση της τυπικής κανονικής κατανομής θα αναμέναμε το πλήθος των κανονικοποιημένων αποκλίσεων για n το σύνολο παρατηρήσεις να έχει την παρακάτω μορφή :

$(-\infty,-3)$	$(-3,-2)$	$(-2,-1)$	$(-1{,}0)$	$(0{,}1)$	$(1{,}2)$	$(2{,}3)$	$(3,+\infty)$
$0\% * n$	$2\% * n$	$14\% * n$	$34\% * n$	$34\% * n$	$14\% * n$	$2\% * n$	$0\% * n$

όπου η πρώτη γραμμή απεικονίζει τα διαστήματα τιμών μέσα στα οποία εμπίπτουν οι κανονικοποιημένες αποκλίσεις και η δεύτερη γραμμή μας δείχνει τον αναμενόμενο αριθμό (Expected number) των κανονικοποιημένων αποκλίσεων επί συνόλου n παρατηρήσεων, που αναμένουμε να εμπίπτει στα αντίστοιχα διαστήματα. Στη συνέχεια, υπολογίζουμε για κάθε ηλικία τις κανονικοποιημένες αποκλίσεις :

$$Z_x = \frac{h_x - n_x * v_x}{\sqrt{n_x * v_x * (1 - v_x)}},$$

και υπολογίζουμε τον *παρατηρούμενο αριθμό (Observed number)* αυτών των αποκλίσεων, όπως αυτές εμπίπτουν στα παραπάνω διαστήματα τιμών των αποκλίσεων του πίνακα. Και σε αυτήν την περίπτωση ζητάμε ο *αναμενόμενος αριθμός (Expected number)* για κάθε διάστημα της δεύτερης γραμμής να ικανοποιεί τη συνθήκη :

$$Expected\ number \succ 5\,.$$

Σε διαφορετική περίπτωση ομαδοποιούμε όσες κλάσεις χρειάζονται ώστε να ικανοποιείται η συνθήκη, τις θεωρούμε ως μία τιμή και μειώνουμε αντιστοίχως τους βαθμούς ελευθερίας. Πραγματοποιούμε ένα X^2 test υπολογίζοντας την τιμή της ποσότητας :

$$X^2 = \sum_x \frac{(Actual\ number - Expected\ number)^2}{Expected}.$$

Για επίπεδο σημαντικότητας $a = 5\%$, δεν θα απορρίψουμε τη μηδενική μας υπόθεση περί καταλληλότητας της εφαρμοσμένης μεθόδου εξομάλυνσης, εάν :

$$X^2 \prec {X^2}_{k,2,5\%}\,,$$

όπου k οι αντίστοιχοι βαθμοί ελευθερίας.

- ***Έλεγχος αθροιστικών αποκλίσεων (Cumulative deviations)***

Σε αυτόν τον έλεγχο μας ενδιαφέρει ο εντοπισμός πολύ μεγάλων ή πολύ μικρών αντίστοιχα σωρευτικών αποκλίσεων. Έχουμε ότι :

$$H_x \sim Binomial(n_x, v_x),$$

με

$$E[H_x] = n_x * v_x$$

και

$$var[H_x] = n_x * v_x * (1 - v_x)\,.$$

Προσεγγίζοντας με κανονική κατανομή και υποθέτοντας ανεξαρτησία μεταξύ των συνολικών παρατηρήσεων, $x = 1,2,3,\ldots,n$ παίρνουμε :

$$H_x \sim Normal(n_x * v_x, n_x * v_x * (1 - v_x))$$

ή

$$H_x - n_x * v_x \sim Normal\big(0, n_x * v_x * (1 - v_x)\big)$$

ή

$$\sum_{x=1}^{n}(H_x - n_x * v_x) \sim Normal(0, \sum_{x=1}^{n} n_x * v_x * (1 - v_x))$$

ή

$$\frac{\sum_{x=1}^{n}(H_x - n_x * v_x)}{\sqrt{\sum_{x=1}^{n} n_x * v_x * (1 - v_x)}} \sim Normal(0,1).$$

Υπολογίζουμε για το σύνολο των παρατηρούμενων ηλικιών την τιμή της ποσότητας :

$$X = \frac{\sum_{x=1}^{n}(h_x - n_x * v_x)}{\sqrt{\sum_{x=1}^{n} n_x * v_x * (1 - v_x)}}$$

και δεν απορρίπτουμε τη μηδενική μας υπόθεση, για επίπεδο σημαντικότητας $a = 5\%$, περί καταλληλότητας της μεθόδου εξομάλυνσης που εφαρμόσαμε, αν ισχύει ότι :

$$-1,96 \prec X \prec 1,96\,.$$

- ***Έλεγχος ομαδοποίησης προσήμων (Steven's test for the grouping of signs)***

Σε αυτόν τον έλεγχο, μας ενδιαφέρει ο εντοπισμός της συγκέντρωσης μεγάλου πλήθους ομόσημων αποκλίσεων της διαφοράς $h_x - n_x * v_x$. Ως μία ομάδα θετικής απόκλισης ορίζεται έστω και ένα μεμονωμένο θετικό πρόσημο της προηγούμενης διαφοράς που θα βρίσκεται ανάμεσα σε δύο αρνητικά πρόσημα της ίδιας διαφοράς. Αν έχουμε για δύο συνεχόμενες ηλικίες και παραπάνω, συνεχόμενα θετικά πρόσημα, τότε αυτά λαμβάνονται επίσης υπ' όψη ως μία ομάδα θετικών αποκλίσεων έως ότου εμφανιστεί αρνητικό πρόσημο της διαφοράς σε κάποια επόμενη ηλικία. Ορίζουμε τις εξής ποσότητες :

- n_p *το πλήθος των θετικών αποκλίσεων της διαφοράς* $h_x - n_x * v_x$.

- n_n *το πλήθος των αρνητικών αποκλίσεων της διαφοράς* $h_x - n_x * v_x$.

- G *η τυχαία μεταβλητή που μετράει το πλήθος των ομάδων των θετικών αποκλίσεων της διαφοράς* $h_x - n_x * v_x$.

Προφανώς, ισχύει ότι το άθροισμα των θετικών αποκλίσεων και των αρνητικών αποκλίσεων της διαφοράς $h_x - n_x * v_x$, ισούται με τον αριθμό των συνολικών παρατηρήσεων, δηλαδή :

$$n_p + n_n = n\,.$$

Τότε, η τυχαία μεταβλητή G ακολουθεί την *υπεργεωμετρική κατανομή* δηλαδή, $G \sim Hypergeometric$ με συνάρτηση πιθανότητας την :

$$P_r(G = g) = \frac{\binom{n_p - 1}{g - 1} * \binom{n_n + 1}{n_n + 1 - g}}{\binom{n_p + n_n}{n_n}}.$$

Με βάση τον δεδομένο αριθμό g του πλήθους των ομάδων των θετικών αποκλίσεων στο σύνολο των παρατηρήσεων, υπολογίζουμε την παραπάνω πιθανότητα και απορρίπτουμε (για επίπεδο σημαντικότητας $a = 5\%$) τη μηδενική μας υπόθεση της καταλληλότητας της μεθόδου εξομάλυνσης που εφαρμόσαμε εάν :

$$P_r(G = g) \prec 5\% \,.$$

Τέλος, αν θέλουμε να προσεγγίσουμε με την τυπική κανονική κατανομή υπολογίζουμε τις ποσότητες :

$$E(G) = \frac{n_n * (n_p - 1)}{n_n + n_p}$$

και

$$\mathrm{var}(G) = \frac{n_p * n_n * (n_p - 1) * (n_n + 1)}{(n_p + n_n)^2 * (n_p + n_n - 1)},$$

όπου

$$\frac{G - E(G)}{\sqrt{var(G)}} \sim Normal(0,1).$$

Υπολογίζουμε την ποσότητα :

$$X = \frac{g - E(G)}{\sqrt{var(G)}},$$

και απορρίπτουμε τη μηδενική μας υπόθεση (για επίπεδο σημαντικότητας $a = 5\%$), της καταλληλότητας της μεθόδου εξομάλυνσης που εφαρμόσαμε εάν :

$$X \prec -1{,}65 \,.$$

<u>ΜΕΡΟΣ 2ο</u>

ΚΕΦΑΛΑΙΟ 4^{ο}

ΜΗ - ΠΑΡΑΜΕΤΡΙΚΕΣ ΜΕΘΟΔΟΙ ΕΞΟΜΑΛΥΝΣΗΣ

4.1 ΕΞΟΜΑΛΥΝΣΗ ΚΙΝΟΥΜΕΝΟΥ ΣΤΑΘΜΙΣΜΕΝΟΥ ΜΕΣΟΥ (MOVING WEIGHTED AVERAGE GRADUATION)

Με βάση αυτή τη μέθοδο, οι εξομαλυμένες τιμές προκύπτουν για κάθε ηλικία ως σταθμισμένος μέσος μιας σειράς πρωτογενών τιμών θνησιμότητας. Ο βασικός τύπος για την εξομαλυμένη τιμή θνησιμότητας σε μια ηλικία είναι :

$$v_x = \sum_{r=-n}^{n} u_{x+r} * a_r \, .$$

Κάποιες βασικές παρατηρήσεις είναι οι εξής :

- *Για την εξομάλυνση μιας τιμής σε συγκεκριμένη ηλικία σταθμίζουμε* n *παρατηρήσεις πριν,* n *παρατηρήσεις μετά και την πρωτογενή τιμή της ίδιας της ηλικίας που εξομαλύνουμε. Άρα, για μια ηλικία χρησιμοποιούμε συνολικά για την εξομάλυνση* $2n + 1$ *παρατηρήσεις.*

- *Οι συντελεστές* a_r *είναι οι συντελεστές στάθμισης και αυξάνονται για ηλικίες που βρίσκονται εγγύτερα στην ηλικία της οποίας εξομαλύνουμε την πρωτογενή τιμή θνησιμότητας. Στον παραπάνω τύπο για την εύρεση της εξομαλυμένης τιμής* v_x*, η αντίστοιχη πρωτογενής τιμή αυτής της ηλικίας πολλαπλασιάζεται με το συντελεστή* a_0 *.*

- *Στις περισσότερες περιπτώσεις θεωρούμε συμμετρικούς τύπους, δηλαδή τύπους για τους οποίους ισχύει* $a_r = a_{-r}$ *για* $r = 1,2, \ldots, n$ *.*

- *Οι ακραίες τιμές θνησιμότητας (σε πολύ μικρές και πολύ μεγάλες ηλικίες) δεν μπορούν να εξομαλυνθούν, λόγω έλλειψης αντίστοιχων στοιχείων προηγουμένως και μετά, με βάση βέβαια τον αριθμό των στοιχείων που χρησιμοποιούμε πριν και μετά αντίστοιχα. Αν για παράδειγμα επιλέγουμε* n *στοιχεία πριν και μετά αντίστοιχα για την εξομάλυνση μιας τιμής* *θνησιμότητας, τότε αν* u_a *και* u_b *είναι η πρώτη και η τελευταία αντίστοιχα τιμή των πρωτογενών τιμών θνησιμότητας, το εύρος των τιμών των εξομαλυμένων τιμών που μπορούμε να υπολογίσουμε θα κινείται μεταξύ των τιμών* v_{a+n} *και* v_{b-n} *.*

Ισχύει ότι :

$$u_{x+r} = t_{x+r} + e_{x+r} \, .$$

Άρα, ο βασικός τύπος της μεθόδου θα γίνει :

$$v_x = \sum_{r=-n}^{n} t_{x+r} * a_r + \sum_{r=-n}^{n} e_{x+r} * a_r \, .$$

Επιθυμούμε την εύρεση τέτοιων συντελεστών στάθμισης a_r, ώστε το πρώτο άθροισμα της παραπάνω ισότητας να προσεγγίζει την πραγματική θνησιμότητα t_x και το δεύτερο άθροισμα να ισούται με e'_x, για το οποίο επιθυμούμε, αφενός μεν την ελαχιστοποίησή του, αφετέρου δε να ισχύει $e'_x \prec e_x$.
Για την τυχαία μεταβλητή E'_x θα ισχύουν τα εξής :

$$E(E'_x) = 0$$

και

$$\mathrm{var}(E'_x) = \mathrm{var}(V_x) = \sum_{r=-n}^{n} {a_r}^2 * \mathrm{var}(U_{x+r}) \, .$$

Υποθέτοντας ότι, οι τυχαίες μεταβλητές U_{x+r} και E_{x+r} είναι ασυσχέτιστες και έχουν διασπορά ίση με σ^2 και ότι :

$$R_0^2 = \sum_{r=-n}^{n} {a_r}^2 \, ,$$

τότε προκύπτει ο τύπος :

$$\mathrm{var}(V_x) = \sigma^2 * R_0^2 \, .$$

Ο προηγούμενος τύπος καλείται *τύπος ελαχίστου* R_0 ή *μεγίστου βάρους*, όπου ως βάρος ορίζουμε την ποσότητα :

$$Βάρος = \frac{1}{R_0^2}.$$

Εφόσον :

$$R_0^2 = \frac{\text{var}(V_x)}{\text{var}(U_x)},$$

στην περίπτωση που υπολογίσουμε $R_0^2 \succ 1$, τότε δεν έχουμε επιτύχει καλή εξομάλυνση.

Στη γενική περίπτωση ζητάμε την ελαχιστοποίηση της ποσότητας R_z^2, που δίνεται από τον *τύπο ελαχίστου* R_z και είναι ο εξής :

$$R_z^2 = \frac{\text{var}(\Delta^z V_x)}{\text{var}(\Delta^z U_x)}.$$

Για τον αριθμητή ισχύει :

$$\text{var}(\Delta^z V_x) = \sigma^2 * \sum_{r=-n-z}^{n} (\Delta^z a_r)^2 .$$

Για τον παρονομαστή ισχύει :

$$\text{var}(\Delta^z U_x) = \sigma^2 * \binom{2z}{z}.$$

Άρα, γενικά για *διαφορές z – τάξης* θα έχουμε :

$$R_z^2 = \frac{\sum_{r=-n-z}^{n} (\Delta^z a_r)^2}{\binom{2z}{z}} ,$$

όπου

$$\binom{n}{k} = \frac{n!}{(n-k)! * k!}.$$

Από τον τύπο :

$$v_x = \sum_{r=-n}^{n} u_{x+r} * a_r = t_x + e'_x ,$$

προκύπτει ότι η εφαρμογή της μεθόδου περιορίζεται στο διάστημα ηλικιών $[x-n, x+n]$, για το οποίο μπορούμε να υποθέσουμε ότι η πραγματική θνησιμότητα t_x προσεγγίζεται από ένα πολυώνυμο μικρού βαθμού (συνήθως 3^{ου} ή 4^{ου}). *Θεωρώντας ότι η t_x προσεγγίζεται από ένα πολυώνυμο 3^{ου} βαθμού ή έχει βαθμό αναπαραγωγής 3 θα πρέπει με βάση τις σχέσεις* :

$$t_{x+r} = a + b * r + c * r^2 + d * r^3$$

και

$$t_x = \sum_{r=-n}^{n} t_{x+r} * a_r ,$$

να ικανοποιούνται οι παρακάτω συνθήκες :

$$\sum_{r=-n}^{n} a_r = 1 , \quad \sum_{r=-n}^{n} r * a_r = 0 , \quad \sum_{r=-n}^{n} r^2 * a_r = 0 , \quad \sum_{r=-n}^{n} r^3 * a_r = 0 .$$

Για πολυώνυμο 4^{ου} βαθμού θα προσθέταμε και τη συνθήκη :

$$\sum_{r=-n}^{n} r^4 * a_r = 0 \; \kappa o \kappa. \; .$$

Για τους συμμετρικούς τύπους για τους οποίους ισχύει $a_r = a_{-r}$ για $r = 1,2,\ldots.,n$ οι παραπάνω συνθήκες θα επαληθεύονται πάντα για τις περιττές δυνάμεις του r . Αντίστροφα λοιπόν, για να βρούμε το βαθμό αναπαραγωγής της συνάρτησης της πραγματικής θνησιμότητας εξετάζουμε μέχρι ποιο βαθμό ισχύουν οι παραπάνω συνθήκες.

Για την εύρεση των συντελεστών στάθμισης a_r, αν και εφόσον αυτοί είναι άγνωστοι, χρησιμοποιούμε το πολύ χρήσιμο αποτέλεσμα του παρακάτω λήμματος :

Έστω συμμετρικός τύπος κινούμενου σταθμισμένου μέσου με $2n + 1$ *όρους, ακριβής για πολυώνυμο βαθμού* d *(όπου* d *περιττός) και ελαχίστου* R_z. *Τότε οι συντελεστές στάθμισης* a_r *είναι τιμές ενός άρτιου πολυωνύμου (πολυώνυμο του οποίου οι συντελεστές των περιττών δυνάμεων είναι μηδενικοί), με βαθμό* $2z + d - 1$.

Τέλος, μια ακόμη χρήσιμη ποσότητα είναι ο *συντελεστής εξομάλυνσης* που προκύπτει εισάγοντας *διαφορές 3ης τάξης* στον τύπο του κινούμενου σταθμισμένου μέσου. Άρα, υπολογίζοντας την ποσότητα R_3^2 προκύπτει και ο συντελεστής εξομάλυνσης ως :

$$R_3 = \sqrt{R_3^2}\,.$$

Γενικά τα μέτρα R_0^2 και R_3^2 αποτελούν σημαντικά κριτήρια, με βάση τα οποία μπορούμε να κρίνουμε αν ο τύπος κινούμενου σταθμισμένου μέσου θα οδηγήσει σε "επιτυχημένες" εξομαλυμένες τιμές. Και για τα δύο μέτρα επιθυμούμε $R_0^2 \prec 1$ και $R_3^2 \prec 1$ ώστε να θεωρήσουμε κατάλληλο τον τύπο που επιλέξαμε. Επίσης, η σύγκριση των δύο μέτρων μας δείχνει ποιος από διάφορους τύπους κινούμενου σταθμισμένου μέσου είναι ο καταλληλότερος (προφανώς όποιος έχει τη μικρότερη τιμή στο αντίστοιχο μέτρο). Ο υπολογισμός του μέτρου R_0^2 αποτελεί κριτήριο σχετικό με τη μείωση της διακύμανσης (reduction of variance), ενώ ο υπολογισμός του μέτρου R_3^2 αποτελεί κριτήριο σχετικό με τη μείωση της "τραχύτητας" (reduction of roughness).

4.2 ΕΞΟΜΑΛΥΝΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ WHITTAKER

Με τη μέθοδο Whittaker, επιθυμούμε την ταυτόχρονη βελτιστοποίηση των μέτρων καλής εφαρμογής και ομαλότητας, F και S αντίστοιχα. Ουσιαστικά, επιδιώκουμε την ελαχιστοποίηση του μέτρου M, όπου :

$$M = F + h * S\,.$$

Ο παραπάνω τύπος παίρνει τις εξής τρείς μορφές :

$$M = F + h * S = \sum_{x=1}^{n} w_x * (v_x - u_x)^2 + h * \sum_{x=1}^{n-z} (\Delta^z v_x)^2\,,$$

που καλείται *τύπος Β του Whittaker*, όπου n το πλήθος των ηλικιών που εξομαλύνουμε και $h \succ 0$ η τιμή μιας σταθεράς που εκφράζει το μέγεθος του βάρους που δίνουμε στο κριτήριο της ομαλότητας S, κατά τη διαδικασία της εξομάλυνσης. Το βάρος για κάθε ηλικία w_x μπορεί να είναι είτε το αντίστροφο της τυχαίας μεταβλητής U_x, είτε το πλήθος των ζώντων ατόμων της ηλικίας κτλ. . Στον *τύπο Α του Whittaker*, θεωρούμε τα ίδια βάρη σε κάθε ηλικία και ίσα με τη μονάδα, δηλαδή $w_x = 1, x = 1,2,\dots,n-1,n$, άρα ο τύπος γίνεται :

$$M = F + h * S = \sum_{x=1}^{n} (v_x - u_x)^2 + h * \sum_{x=1}^{n-z} (\Delta^z v_x)^2\,.$$

Στη *γενικευμένη του μορφή ο τύπος του Whittaker* δίνεται ως εξής :

$$M = \sum_{x=1}^{n} w_x * (v_x - u_x)^2 + h_1 * \sum_{x=1}^{n-1} (\Delta v_x)^2 + h_2 * \sum_{x=1}^{n-2} (\Delta^2 v_x)^2 + .. + h_z * \sum_{x=1}^{n-z} (\Delta^z v_x)^2$$

και συνήθως ο τύπος χρησιμοποιείται μέχρι και την ύπαρξη διαφορών τρίτης ή τέταρτης τάξης.

Στόχος μας είναι η εύρεση των εξομαλυμένων τιμών θνησιμότητας v_x για κάθε ηλικία, που ελαχιστοποιούν την τιμή του μέτρου M . Αυτό μπορεί να επιτευχθεί είτε παίρνοντας τις μερικές παραγώγους του μέτρου M ως προς κάθε ηλικία και θέτοντάς τες ίσες με μηδέν, ώστε να καταλήξουμε σε ένα σύστημα n εξισώσεων με n αγνώστους. Δηλαδή, θέτοντας :

$$\frac{\partial M}{\partial v_r} = 0, \text{για } r = 1, 2, \ldots, n.$$

Αντί της παραπάνω διαδικασίας, επιλέγουμε να εκφράσουμε το μέτρο M με τη μορφή Πινάκων – Μητρών, βρίσκοντας μέσω αυτής της διαδικασίας τις εξομαλυμένες τιμές θνησιμότητας v_x . Ορίζουμε τους εξής Πίνακες :

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix},$$

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{bmatrix},$$

$$\boldsymbol{w} = \begin{bmatrix} w_1 & 0 & . & . & 0 \\ 0 & w_2 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & w_n \end{bmatrix}.$$

Οι Πίνακες - Μήτρες θα συμβολίζονται με έντονα γράμματα. Ο πρώτος $n * 1$ Πίνακας παραπάνω, περιέχει τις πρωτογενείς τιμές θνησιμότητας για κάθε ηλικία. Αντίστοιχα ο δεύτερος $n * 1$ Πίνακας περιέχει τις ζητούμενες εξομαλυμένες τιμές θνησιμότητας για κάθε ηλικία, ώστε να ελαχιστοποιείται το μέτρο M. Ο τρίτος $n * n$ Πίνακας είναι διαγώνιος (περιέχει δηλαδή μη – μηδενικά στοιχεία μόνο στην διαγώνιό του) και έχει ως στοιχεία της διαγωνίου του τα βάρη σε κάθε ηλικία. Επίσης, ορίζουμε τον Πίνακα $\boldsymbol{K_z}$ διαστάσεων $(n - z) * n$, που περιέχει τους διωνυμικούς συντελεστές που προκύπτουν από την :

$$(-1)^{z-k} * \binom{z}{\kappa}, k = 0, 1, \ldots, z.$$

Άρα, ο $\boldsymbol{K_z}$ έχει την παρακάτω μορφή :

$$\begin{bmatrix} (-1)^z * \binom{z}{0} & (-1)^{z-1} * \binom{z}{1} & . & . & . & (-1)^0 * \binom{z}{z} & 0 & . & . & 0 \\ 0 & (-1)^z * \binom{z}{0} & . & . & . & . & (-1)^0 * \binom{z}{z} & . & . & 0 \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 & 0 & . & . & (-1)^0 * \binom{z}{z} \end{bmatrix}.$$

Άρα, αν για παράδειγμα επιλέξουμε τις 3ες τη τάξη διαφορές και έχουμε $n = 7$ ηλικίες να εξομαλύνουμε τότε θα προκύψει ο Πίνακας $\boldsymbol{K_3}$ με 4 γραμμές και 7 στήλες με την παρακάτω μορφή :

$$\boldsymbol{K_3} = \begin{bmatrix} -1 & 3 & -3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 3 & -3 & 1 \end{bmatrix}.$$

Πριν προχωρήσουμε στην αναδιατύπωση μέσω Πινάκων του τύπου του μέτρου M, παραθέτουμε ορισμένα βασικά στοιχεία Γραμμικής Άλγεβρας που αφορούν τους Πίνακες.

- *Αν* $\boldsymbol{A}$ *είναι ένας Πίνακας – Μήτρα* $n * k$ *διαστάσεων, τότε καλείται ανάστροφός του, ο Πίνακας* $k * n$ *διαστάσεων που έχει ως γραμμές του τις στήλες του αρχικού Πίνακα και ως στήλες του τις γραμμές του αρχικού Πίνακα και συμβολίζεται με* $\boldsymbol{A'}$.

- *Αν* $\boldsymbol{A}$ *είναι ένας Πίνακας – Μήτρα* $n * n$ *διαστάσεων, τότε καλείται αντίστροφός του, ο Πίνακας* $n * n$ *διαστάσεων που συμβολίζεται με* $\boldsymbol{A^{-1}}$.

- *Ο Πίνακας* $\boldsymbol{A}$ *καλείται ομαλός, αν και μόνο αν υπάρχει ο* $\boldsymbol{A^{-1}}$.

- *Ο Πίνακας* $\boldsymbol{A}$ *καλείται αντιστρέψιμος, αν και μόνο αν υπάρχει ο* $\boldsymbol{A^{-1}}$.

- *Ο Πίνακας* $\boldsymbol{A}$ *καλείται μη - ομαλός, αν και μόνο αν δεν υπάρχει ο* $\boldsymbol{A^{-1}}$.

- *Ο Πίνακας $\boldsymbol{A}$ καλείται μη – αντιστρέψιμος, αν και μόνο αν δεν υπάρχει ο $\boldsymbol{A^{-1}}$.*

- *Ισχύει η σχέση : $\boldsymbol{A^{-1}} * \boldsymbol{A} = \boldsymbol{A} * \boldsymbol{A^{-1}} = \boldsymbol{I_n}$, όπου $\boldsymbol{I_n}$ η Μοναδιαία Μήτρα (Μήτρα με τα στοιχεία της κύριας διαγωνίου της όλα ίσα με μονάδα και όλα τα υπόλοιπα στοιχεία της ίσα με μηδέν).*

Άρα, ο αρχικός τύπος του Whittaker (τύπος B), μπορεί να γραφεί με τη μορφή Μητρών ως εξής :

$$M = F + h * S = (\boldsymbol{v} - \boldsymbol{u})' * \boldsymbol{w} * (\boldsymbol{v} - \boldsymbol{u}) + h * (\boldsymbol{K_z} * \boldsymbol{v})' * (\boldsymbol{K_z} * \boldsymbol{v}).$$

Το ζητούμενο διάνυσμα $\boldsymbol{v}$ που εμπεριέχει τις εξομαλυμένες τιμές θνησιμότητας και ελαχιστοποιεί ταυτοχρόνως το μέτρο M, αποδεικνύεται ότι προκύπτει μέσω της σχέσης :

$$\boldsymbol{c} * \boldsymbol{v} = \boldsymbol{w} * \boldsymbol{u},$$

όπου ο $\boldsymbol{c}$ είναι ένας αντιστρέψιμος Πίνακας διαστάσεων $n * n$, για τον οποίο ισχύει η σχέση :

$$\boldsymbol{c} = \boldsymbol{w} + h * (\boldsymbol{K_z}' * \boldsymbol{K_z}).$$

Στον *τύπο B του Whittaker* θεωρούμε ότι η πραγματική θνησιμότητα t_x, προσεγγίζεται από ένα πολυώνυμο βαθμού $z - 1$, ώστε να ελαχιστοποιείται το μέτρο :

$$\sum_{x=1}^{n-z} (\Delta^z v_x)^2 .$$

Γενικά, αν θεωρήσουμε ότι η πραγματική μας θνησιμότητα προσεγγίζεται από ένα συνδυασμό ενός πολυωνύμου βαθμού $z - 2$ και ενός εκθετικού όρου της μορφής $B * C^x$, δηλαδή :

$$t_x = a_0 + a_1 * x + a_2 * x^2 + \ldots + a_{z-2} * x^{z-2} + B * C^x ,$$

τότε το κριτήριο ομαλότητας S ελαχιστοποιείται, έχοντας την ακόλουθη μορφή :

$$\sum_{x=1}^{n-z}(\Delta^z v_x - r * \Delta^{z-1} v_x)^2 \,,$$

όπου

$$r = c - 1 \,.$$

4.3 ΕΞΟΜΑΛΥΝΣΗ ΚΑΤΑ BAYES

4.3.1 Η ΔΙΑΔΙΚΑΣΙΑ ΕΞΟΜΑΛΥΝΣΗΣ ΚΑΤΑ BAYES

Η προσέγγιση κατά Bayes της διαδικασίας της εξομάλυνσης, εισάγει στο μέγιστο βαθμό το στοιχείο της πρότερης γνώμης (prior opinion), όσον αφορά την κατανομή της πραγματικής θνησιμότητας, σε σχέση με τις υπόλοιπες μη – παραμετρικές μεθόδους. *Η διαδικασία της εξομάλυνσης κατά Bayes συνίσταται στα εξής τέσσερα βήματα :*

- *Διατύπωση της πρότερης γνώμης σχετικά με τη συνάρτηση πυκνότητας πιθανότητας $f_T(t)$ της τυχαίας μεταβλητής T που εκφράζει την πραγματική θνησιμότητα.*

- *Επιλέγουμε το μοντέλο (δηλαδή την κατανομή) του πειράματος, που αφορά τη δεσμευμένη κατανομή $f_{U\,/T}(u\ /\ t)$ των πρωτογενών τιμών θνησιμότητας με βάση την υπόθεση που έχουμε κάνει για την πραγματική θνησιμότητα.*

- *Εξάγουμε την εκ των υστέρων κατανομή $f_{T\,/U}(t\ /\ u)$ μέσω του τύπου του Bayes, που αφορά την πραγματική θνησιμότητα με βάση το πείραμα που εκτελέσαμε και λάβαμε τις πρωτογενείς τιμές θνησιμότητας.*

- *Επιλέγουμε τις εξομαλυμένες τιμές θνησιμότητας από την αναμενόμενη τιμή ή την επικρατούσα τιμή ή τη διάμεσο της εκ των υστέρων κατανομής.*

Άρα, ο βασικός τύπος με τον οποίο δουλεύουμε και περιγράφει συνοπτικά την ανωτέρω διαδικασία είναι ο :

$$f_{T\,/U}(t\ /\ u) = \frac{f_{U\,/T}(u\ /\ t) * f_T(t)}{f_U(u)} \,.$$

Ο παραπάνω τύπος καθιστά ξεκάθαρο το βαθμό της *υποκειμενικότητας* που υπεισέρχεται στην όλη διαδικασία της εξομάλυνσης κατά Bayes και αφορά την επιλογή της συνάρτησης πυκνότητας πιθανότητας της πραγματικής θνησιμότητας. *Επίσης, τις περισσότερες φορές είναι χρήσιμο να επιλέγεται μια εκ των προτέρων κατανομή για την πραγματική θνησιμότητα, τέτοια ώστε σε συνδυασμό με την κατανομή του πειράματος να παράγει μια εκ των υστέρων κατανομή που να ανήκει στην ίδια οικογένεια και να έχει τις ίδιες ιδιότητες με την πρότερη κατανομή.* Ορισμένοι τέτοιοι συνδυασμοί παρουσιάζονται στον παρακάτω πίνακα και ακολουθεί το παράδειγμα του συνδυασμού της πρότερης κατανομής Beta, με την κατανομή Binomial ως κατανομή του πειράματος, που καταλήγει σε Beta εκ των υστέρων κατανομή.

ΠΡΟΤΕΡΗ ΚΑΤΑΝΟΜΗ $f_T(t)$	ΔΕΣΜΕΥΜΕΝΗ ΚΑΤΑΝΟΜΗ ΠΕΙΡΑΜΑΤΟΣ $f_{U/T}(u/t)$	ΥΣΤΕΡΗ ΚΑΤΑΝΟΜΗ $f_{T/U}(t/u)$
NORMAL	NORMAL	NORMAL
GAMMA	POISSON	GAMMA
GAMMA	GAMMA	GAMMA
BETA	BINOMIAL	BETA
DIRICHLET	MULTINOMIAL	DIRICHLET
GAMMA	NORMAL	GAMMA

Θεωρούμε ότι πρότερη κατανομή της τυχαίας μεταβλητής που εκφράζει την πραγματική θνησιμότητα είναι η Beta με παραμέτρους $a \succ 0$ και $b \succ 0$, δηλαδή ισχύει :

$$T \sim \text{Beta}(a, b)\,,$$

με

$$f_T(t) = \frac{\Gamma(a+b)}{\Gamma(a) * \Gamma(b)} * t^{a-1} * (1-t)^{b-1}\,, \qquad 0 \le t \le 1$$

και

$$E(T) = \frac{a}{a+b}$$

και

$$\text{var}(T) = \frac{a * b}{(a+b)^2 * (a+b+1)}\,.$$

Η δεσμευμένη κατανομή του πειράματος, που αφορά την τυχαία μεταβλητή των θανάτων που παρατηρούνται, θεωρούμε ότι είναι Binomial, με παραμέτρους n, t. Δηλαδή :

$$H\ / T \sim \text{Binomial}(n\,, t),$$

με

$$f_{H\ /T}(h\ / t) = \binom{n}{h} * t^h * (1-t)^{n-h}\,, \qquad h = 0, 1, 2, \dots, n\,.$$

Η περιθώρια κατανομή της H προκύπτει από τον τύπο :

$$f_H(h) = \int_0^1 f_{H\ /T}(h\ / t) * f_T(t)\, dt\,.$$

Συνεπώς, για την ύστερη κατανομή της πραγματικής θνησιμότητας θα έχουμε :

$$f_{T\ /H}(t\ / h) = \frac{\binom{n}{h} * t^h * (1-t)^{n-h} * \frac{\Gamma(a+b)}{\Gamma(a) * \Gamma(b)} * t^{a-1} * (1-t)^{b-1}}{\int_0^1 \binom{n}{h} * t^h * (1-t)^{n-h} * \frac{\Gamma(a+b)}{\Gamma(a) * \Gamma(b)} * t^{a-1} * (1-t)^{b-1}\, dt} =$$

$$\frac{t^{a+h-1} * (1-t)^{b+n-h-1}}{\frac{\Gamma(a+h) * \Gamma(b+n-h)}{\Gamma(a+b+n)} * \int_0^1 \frac{\Gamma(a+b+n)}{\Gamma(a+h) * \Gamma(b+n-h)} * t^{a+h-1} * (1-t)^{b+n-h-1}\, dt}\,.$$

Όμως το εσωτερικό του ολοκληρώματος είναι η συνάρτηση πυκνότητας πιθανότητας μιας κατανομής Beta με παραμέτρους $a+h$ και $b+n-h$, επομένως το ολοκλήρωμα κάνει μονάδα. Άρα, καταλήγουμε στο :

$$f_{T\ /H}(t\ /h) = \frac{\Gamma(a+b+n)}{\Gamma(a+h) * \Gamma(b+n-h)} * t^{a+h-1} * (1-t)^{b+n-h-1},$$

που σημαίνει ότι η εκ των υστέρων κατανομή της θνησιμότητας με βάση το πείραμα είναι :

$$T\ /H \sim \mathrm{Beta}(\, a+h\,, b+n-h\,)\,.$$

Οι εξομαλυμένες τιμές της θνησιμότητας προκύπτουν ως οι μέσες τιμές της ύστερης κατανομής, δηλαδή της :

$$E(T\ /H) = \frac{a+h}{a+b+n} = \frac{a}{a+b} * \frac{a+b}{a+b+n} + \frac{h}{n} * \frac{n}{a+b+n} \Leftrightarrow$$

$$E(T\ /H) = E(T) * \frac{a+b}{a+b+n} + u * \frac{n}{a+b+n}\ .$$

Δηλαδή παρατηρούμε ότι η εξομαλυμένες τιμές είναι ένας σταθμισμένος μέσος της μέσης τιμής της πρότερης κατανομής που υποθέσαμε για την πραγματική θνησιμότητα και της παρατηρούμενης – πρωτογενούς τιμής θνησιμότητας που προέκυψε μέσω του πειράματος.

4.3.2 Η ΜΕΘΟΔΟΣ KIMELDORF - JONES

Στην ίδια λογική της εξομάλυνσης κατά Bayes, στηρίζεται και η μέθοδος εξομάλυνσης που ανέπτυξαν οι Kimeldorf και Jones. *Στη συγκεκριμένη μέθοδο εφαρμόζουμε εξομάλυνση συγχρόνως σε όλο το εύρος των ηλικιών και θα θεωρήσουμε ότι το πλήθος του συνόλου των ηλικιών που θα εξομαλύνουμε είναι* n .

Για την *τυχαία μεταβλητή που εκφράζει την πραγματική θνησιμότητα σε κάθε ηλικία* i, για $i = 1, 2, \dots, n,$ θα θεωρήσουμε ότι ακολουθεί την *κανονική κατανομή* με μέση τιμή m_i και διακύμανση a_{ii}. Δηλαδή :

$$T_i \sim Normal\,(\,m_i\,,\,a_{\text{ii}}\,)\,,$$

με

$$f_{T_i}(t_i) = \frac{1}{\sqrt{2 * \pi * a_{\text{ii}}}} * e^{-\frac{(t_i - m_i)^2}{2 * a_{\text{ii}}}}\,.$$

Τότε το *διάνυσμα της πραγματικής θνησιμότητας* $\boldsymbol{T}$ που περιλαμβάνει όλες τις ηλικίες, ακολουθεί την *πολυδιάστατη κανονική κατανομή* με παραμέτρους $\boldsymbol{m}$ και $\boldsymbol{A}$. Ορίζουμε το *συντελεστή συσχέτισης* μεταξύ δύο ηλικιών i και j, για $i, j = 1, 2, \dots, n$, ως εξής :

$$\sigma_{\text{ij}} = \frac{\operatorname{cov}(\,T_i\,,\,T_j\,)}{\sqrt{a_{\text{ii}} * a_{\text{jj}}}}\,.$$

Έχουμε συνεπώς :

$$f_T(t) = k_1 * e^{-\frac{(\boldsymbol{t}-\boldsymbol{m})' * \boldsymbol{A}^{-1} * (\boldsymbol{t}-\boldsymbol{m})}{2}},$$

όπου

$$k_1 = \frac{1}{\sqrt{(2 * \pi)^n * \;|\,\boldsymbol{A}\,|}}\,.$$

Ο Πίνακας $\boldsymbol{m}$ είναι διαστάσεων $n * 1$ και περιέχει τις μέσες τιμές m_i για την κάθε ηλικία, ενώ ο Πίνακας $\boldsymbol{A}$ είναι διαστάσεων $n * n$, περιέχει τις *συνδιακυμάνσεις* μεταξύ των θνησιμοτήτων κάθε ηλικίας και έχει την εξής μορφή :

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & c_{12} * \sqrt{a_{11} * a_{22}} & . & . & c_{1n} * \sqrt{a_{11} * a_{\text{nn}}} \\ c_{21} * \sqrt{a_{22} * a_{\text{nn}}} & a_{22} & . & . & c_{2n} * \sqrt{a_{22} * a_{\text{nn}}} \\ . & . & . & . & . \\ . & . & . & . & . \\ c_{\text{n1}} * \sqrt{a_{\text{nn}} * a_{11}} & . & . & . & a_{\text{nn}} \end{bmatrix},$$

με $|\,\boldsymbol{A}\,|$ την *ορίζουσα* της συγκεκριμένης τετραγωνικής Μήτρας . Η μορφή του $\boldsymbol{m}$ είναι :

$$\boldsymbol{m} = \begin{bmatrix} m_1 \\ m_2 \\ . \\ . \\ m_n \end{bmatrix}.$$

Για τη *δεσμευμένη κατανομή των πρωτογενών τιμών της θνησιμότητας* έχουμε για την κάθε ηλικία ξεχωριστά (για μεγάλο πλήθος ζώντων ατόμων n_i) :

$$U_i \sim Normal\left(t_i, \frac{t_i * (1 - t_i)}{n_i}\right),$$

που καταλήγει για το συνδυασμό όλων των ηλικιών μαζί :

$$f_{U\ /T}(u\ /\ t) = k_2 * e^{-\frac{(\boldsymbol{u}-\boldsymbol{t})' * \boldsymbol{B}^{-1} * (\boldsymbol{u}-\boldsymbol{t})}{2}},$$

με

$$k_2 = \frac{1}{\sqrt{(2 * \pi)^n * \ |\boldsymbol{B}|}}\ .$$

Ο Πίνακας $\boldsymbol{t}$ είναι διαστάσεων $n * 1$ και έχει την παρακάτω μορφή :

$$\boldsymbol{t} = \begin{bmatrix} t_1 \\ t_2 \\ . \\ . \\ t_n \end{bmatrix}.$$

Ο Πίνακας $\boldsymbol{B}$ είναι ένας διαγώνιος Πίνακας διαστάσεων $n * n$ και περιέχει στη διαγώνιό του τις διακυμάνσεις της τυχαίας μεταβλητής U_i, όπου,

$$\operatorname{var}(U_i) = \frac{t_i * (1 - t_i)}{n_i}$$

και

$$\boldsymbol{B} = \begin{bmatrix} \operatorname{var}(U_1) & 0 & . & . & 0 \\ 0 & \operatorname{var}(U_2) & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \operatorname{var}(U_n) \end{bmatrix},$$

με $|\boldsymbol{B}|$ την ορίζουσα της συγκεκριμένης τετραγωνικής Μήτρας και έχοντας υποθέσει ανεξαρτησία μεταξύ των τυχαίων μεταβλητών U_i και T_i .

Μετά από μια σειρά βημάτων και εφαρμόζοντας τον τύπο του Bayes καταλήγουμε στο ότι η *ύστερη κατανομή* είναι :

$$f_{T\ /\ U}(t\ /\ u) = k_6 * e^{-\frac{(\boldsymbol{t}-\boldsymbol{v})' * \boldsymbol{C}^{-1} * (\boldsymbol{t}-\boldsymbol{v})}{2}},$$

όπου η ποσότητα k_6 δεν εμπεριέχει το $\boldsymbol{t}$ και $\boldsymbol{v}$ ο $n*1$ Πίνακας που περιέχει τις εξομαλυμένες τιμές θνησιμότητας κάθε ηλικίας (που προκύπτουν όπως παρατηρούμε ως οι μέσες τιμές της εκ των υστέρων κατανομής) και δίνεται από τη σχέση :

$$\boldsymbol{v} = \boldsymbol{C} * (\boldsymbol{B}^{-1} * \boldsymbol{u} + \boldsymbol{A}^{-1} * \boldsymbol{m}),$$

όπου

$$\boldsymbol{C} = (\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1}\ .$$

ΚΕΦΑΛΑΙΟ 5ο

ΠΑΡΑΜΕΤΡΙΚΕΣ ΜΕΘΟΔΟΙ ΕΞΟΜΑΛΥΝΣΗΣ

5.1 ΝΟΜΟΙ ΘΝΗΣΙΜΟΤΗΤΑΣ & ΑΛΛΕΣ ΜΟΡΦΕΣ ΣΥΝΑΡΤΗΣΕΩΝ

5.1.1 ΤΥΠΟΙ & ΒΑΣΙΚΕΣ ΣΧΕΣΕΙΣ

Οι *Νόμοι Θνησιμότητας*, αποτελούν βασικές συναρτήσεις οι οποίες έχουν χρησιμοποιηθεί ευρέως όσον αφορά δεδομένα θνησιμότητας. Παρακάτω θα αναπτυχθούν οι Νόμοι Θνησιμότητας των *Gompertz*, *Makeham* και *Weibull* και θα δοθούν και οι τύποι άλλων συναρτήσεων που έχουν διατυπωθεί για δεδομένα που αφορούν θνησιμότητα.

- ***Gompertz***

Ο βασικός τύπος που ανέπτυξε ο *Gompertz* με αναφορά στην *ένταση θνησιμότητας* είναι :

$$\mu_x = B * C^x ,$$

όπου

$$B \succ 0 \text{ και } C \succ 1 .$$

Προκύπτει ότι :

$$p_x = e^{-\int_0^1 \mu_{x+t}\,dt} = g^{C^x*(C-1)} ,$$

όπου

$$g = e^{-\frac{B}{\ln C}} .$$

- ***Makeham***

Ο βασικός τύπος που ανέπτυξε ο *Makeham* με αναφορά στην *ένταση θνησιμότητας* είναι :

$$\mu_x = A + B * C^x,$$

όπου Α $\succ -B$, μια σταθερά που εκφράζει την επίδραση ενός τυχαίου γεγονότος.

Προκύπτει ότι :

$$p_x = e^{-\int_0^1 \mu_{x+t}\, dt} = S * g^{C^x*(C-1)} ,$$

όπου

$$S = e^{-A} .$$

- ***Weibull***

Ο βασικός τύπος που ανέπτυξε ο *Weibull* με αναφορά στην *ένταση θνησιμότητας* είναι :

$$\mu_x = k * x^n,$$

όπου

$$k \succ 0 \;\; \text{και} \;\; n \succ -1 .$$

Προκύπτει ότι :

$$p_x = e^{-\int_0^1 \mu_{x+t}\, dt} = e^{-\frac{k}{n+1}*[(x+1)^{n+1}-x^{n+1}]},$$

όπου

$$(x+1)^{n+1} - x^{n+1} = \Delta\, x^{n+1} .$$

Τέλος, άλλες μορφές συναρτήσεων που έχουν διατυπωθεί κατά καιρούς είναι οι παρακάτω :

- **2ος Νόμος Makeham**

$$\mu_x = A + H * x + B * C^x .$$

- **Logistic (Perks)**

$$\mu_x = \frac{c + a_1 * e^{b*x}}{1 + a_2 * e^{b*x}} .$$

- **Log – Quadratic (Coale and Kisker)**

$$ln\, q_x = a_0 + a_1 * x + a_2 * x^2 \,.$$

- **Heligman – Pollard**

$$q_x = \frac{a * e^{b*x}}{1 + a * e^{b*x}} \,.$$

- **Kannisto**

$$\mu_x = \frac{c + a * e^{b*x}}{1 + a * e^{b*x}} \,.$$

5.1.2 ΕΠΙΛΟΓΗ ΣΥΝΑΡΤΗΣΗΣ & ΕΥΡΕΣΗ ΠΑΡΑΜΕΤΡΩΝ

Αυτό που μας ενδιαφέρει είναι, αφενός μεν να εξετάσουμε αν τα πρωτογενή μας δεδομένα μπορούν να προσεγγιστούν από κάποιο Νόμο Θνησιμότητας και αφετέρου δε να υπολογίσουμε εν συνεχεία τις τιμές των παραμέτρων των συναρτήσεων, όπως αυτές ορίζονται από τον κάθε Νόμο Θνησιμότητας. Το πρώτο επιτυγχάνεται είτε μέσω της *γραφικής μεθόδου*, είτε μετασχηματίζοντας κατάλληλα μέτρα του κάθε Νόμου ώστε να προκύψουν *σταθερές*. Το δεύτερο επιτυγχάνεται μέσω μεθόδων όπως η *Μέθοδος Ελαχίστων Τετραγώνων*. Πιο συγκεκριμένα θα έχουμε για κάθε Νόμο Θνησιμότητας τα εξής :

- ***Gompertz***

Για να δούμε αν οι πρωτογενείς τιμές της θνησιμότητας μπορούν να προσεγγιστούν από το *Νόμο του Gompertz*, αρκεί να εξετάσουμε αν η τιμή του μέτρου :

$$\frac{ln(1 - u_{x+1})}{ln(1 - u_x)},$$

είναι μια σταθερά, καθώς για το *Νόμο του Gompertz* ισχύει :

$$\frac{ln\, p_{x+1}}{ln\, p_x} = C\ .$$

Για την εύρεση των τιμών των παραμέτρων B και C, εφαρμόζοντας τη Μέθοδο Ελαχίστων Τετραγώνων, αποσκοπούμε στην *ελαχιστοποίηση* του μέτρου :

$$SS = \sum_{x=1}^{n} w_x * ({u_x}^* - a - b * x^*)^2\ .$$

Για να επιτευχθεί αυτό, αρκεί να εκφράσουμε κάποιο μέτρο θνησιμότητας μέσω του *Νόμου του Gompertz* μέσω μιας *γραμμικής σχέσης* και να θέσουμε τις μερικές παραγώγους ως προς a και b, του μέτρου SS, ίσες με μηδέν. Από το Νόμο του *Gompertz* προκύπτει η εξής *γραμμική σχέση* με αναφορά στην ένταση θνησιμότητας :

$$ln\, \mu_x = ln\, B + x * ln\, C\ .$$

Άρα, αντιστοιχίζοντας τα μέτρα $ln\, \mu_x$, $ln\, B$, x και $ln\, C$ με τα μέτρα ${u_x}^*$, a , x^* και b στον τύπο των Ελαχίστων Τετραγώνων, παίρνουμε :

$$\frac{\partial SS}{\partial a} = 0 \qquad \text{και} \qquad \frac{\partial SS}{\partial b} = 0\ ,$$

απ' όπου λύνοντας το σύστημα που προκύπτει υπολογίζονται οι τιμές των παραμέτρων B και C . Έχοντας σαν πρωτογενή δεδομένα πιθανότητες θανάτου θα χρησιμοποιούσαμε τη σχέση :

$$ln[-ln(1 - q_x)] = ln\left[\frac{(C-1) * B}{ln\, C}\right] + x * ln\, C\ ,$$

εφαρμόζοντας ομοίως τις αντιστοιχίες στον τύπο των Ελαχίστων Τετραγώνων και λύνοντας.

- ***Makeham***

Για να δούμε αν οι πρωτογενείς τιμές της θνησιμότητας μπορούν να προσεγγιστούν από το *Νόμο του Makeham*, αρκεί να εξετάσουμε αν η τιμή του μέτρου :

$$\frac{\Delta\, ln(1-u_{x+1})}{\Delta\, ln(1-u_x)},$$

είναι μια σταθερά, καθώς για το *Νόμο του Makeham* ισχύει :

$$\frac{\Delta\, ln\, p_{x+1}}{\Delta\, ln\, p_x} = C\,.$$

Για την εύρεση των τιμών των παραμέτρων B, C και A εφαρμόζοντας τη Μέθοδο Ελαχίστων Τετραγώνων, αποσκοπούμε και πάλι στην *ελαχιστοποίηση* του μέτρου :

$$SS = \sum_{x=1}^{n} w_x * ({u_x}^* - a - b * x^*)^2\,.$$

Από το *Νόμο του Makeham* προκύπτει η εξής *γραμμική σχέση* με αναφορά στην ένταση θνησιμότητας :

$$ln\,\Delta\mu_x = ln[B*(C-1)] + x * ln\,C\,.$$

Άρα, αντιστοιχίζοντας τα μέτρα $ln\,\Delta\mu_x$, $ln[B*(C-1)]$, x και $ln\,C$ με τα μέτρα ${u_x}^*$, a , x^* και b στον τύπο των Ελαχίστων Τετραγώνων, παίρνουμε :

$$\frac{\partial SS}{\partial a} = 0 \qquad \text{και} \qquad \frac{\partial SS}{\partial b} = 0\,,$$

απ' όπου λύνοντας το σύστημα που προκύπτει υπολογίζονται οι τιμές των παραμέτρων B, C και A .

- ***Weibull***

Για να δούμε αν οι πρωτογενείς τιμές της θνησιμότητας μπορούν να προσεγγιστούν από το *Νόμο του Weibull*, αρκεί να εξετάσουμε αν οι τιμές των μέτρων :

$$\Delta^n ln(1 - u_x)$$

ή

$$\Delta^{n+1}(ln\, n_x)\,,$$

είναι σταθερές, καθώς για το *Νόμο του Weibull* ισχύουν αντίστοιχα οι σχέσεις :

$$\Delta^n ln\, p_x = -k * n!\,,$$

και

$$\Delta^{n+1}(ln\, l_x) = -k * n!\,.$$

Παρατηρούμε ότι μπορούμε να εξάγουμε άμεσα τις τιμές των παραμέτρων k και n από τις παραπάνω σχέσεις, για ένα περιορισμένο διάστημα ηλικιών. Εμάς μας ενδιαφέρει να λάβουμε υπόψη μας σύνολο των ηλικιών για την εύρεση των τιμών των παραμέτρων k και n. Εφαρμόζοντας συνεπώς και πάλι τη Μέθοδο Ελαχίστων Τετραγώνων, αποσκοπούμε εκ νέου στην *ελαχιστοποίηση* του μέτρου :

$$SS = \sum_{x=1}^{n} w_x * ({u_x}^* - a - b * x^*)^2\,.$$

Από το *Νόμο του Weibull* προκύπτει η εξής *γραμμική σχέση* με αναφορά στην ένταση θνησιμότητας :

$$ln\, \mu_x = ln\, k + n * ln\, x.$$

Άρα, αντιστοιχίζοντας τα μέτρα $ln\, \mu_x$, $ln\, k$, $ln\, x$ και n με τα μέτρα ${u_x}^*$, a , x^* και b στον τύπο των Ελαχίστων Τετραγώνων, παίρνουμε :

$$\frac{\partial SS}{\partial a} = 0 \qquad \text{και} \qquad \frac{\partial SS}{\partial b} = 0\,,$$

απ’ όπου λύνοντας το σύστημα που προκύπτει υπολογίζονται οι τιμές των παραμέτρων k και n. *Προς αποφυγή παρανοήσεων, αναφέρεται η διαφορά μεταξύ της παραμέτρου του Νόμου του Weibull* n *και του συνόλου των ηλικιών που θα εξομαλυνθούν που συμβολίζεται επίσης με* n.

Υπολογίζοντας συνεπώς την τιμή των μεταβλητών για κάθε συναρτησιακή μορφή (είτε Νόμου Θνησιμότητας, είτε διαφορετικής) μέσω των πρωτογενών τιμών θνησιμότητας, υπολογίζουμε με βάση τις ευρεθείσες τιμές των παραμέτρων, τις εξομαλυμένες τιμές θνησιμότητας για κάθε ηλικία.

5.2 ΕΞΟΜΑΛΥΝΣΗ ΜΕ SPLINES

Στην εξομάλυνση με τη χρήση *splines* (*τόξων*), χωρίζουμε το σύνολο των ηλικιών σε υπο - ομάδες στις οποίες εφαρμόζουμε ξεχωριστά εξομάλυνση, με τη χρήση ενός πολυωνύμου. Τα πολυώνυμα που εφαρμόζουμε σε κάθε τμήμα των ηλικιών είναι του *ίδιου βαθμού* και υπολογίζονται με τη *Μέθοδο Ελαχίστων Τετραγώνων*. Τα σημεία στα οποία χωρίζουμε τις ηλικίες καλούνται *κόμβοι* (*knots*) και συμβολίζονται με K_i. Οι κόμβοι θα πρέπει να *επαληθεύουν* τις εξισώσεις δύο διαδοχικών τόξων και να ικανοποιούν συγκεκριμένες *συνθήκες*. Πιο συγκεκριμένα, αν θεωρήσουμε εμείς ότι οι εξομαλυμένες τιμές θα προκύψουν μέσω ενός πολυωνύμου *τρίτου βαθμού*, τότε για *δύο τόξα – splines*, τα $i-1$ και i, που ενώνονται στον κόμβο K_i, θα πρέπει να ικανοποιούνται οι εξής *συνθήκες* :

$$p_{i-1}(K_i) = p_i(K_i)\,,$$
$$p_{i-1}'(K_i) = p_i'(K_i)\,,$$
$$p_{i-1}''(K_i) = p_i''(K_i)\,.$$

Αν θεωρούσαμε πολυώνυμο τετάρτου βαθμού, θα προσθέταμε και τη συνθήκη :

$$p_{i-1}'''(K_i) = p_i'''(K_i) \text{ κ.ο.κ. .}$$

Στην περίπτωση λοιπόν που το εύρος των ηλικιών που θέλουμε να εξομαλύνουμε, με τη χρήση *δύο τόξων*, κυμαίνεται από την ηλικία a έως την ηλικία b, θεωρώντας *κόμβο* k στον οποίο ενώνονται τα δύο τόξα και θεωρώντας πολυώνυμο *τρίτου βαθμού* για τις εξομαλυμένες μας τιμές, η διαδικασία εξομάλυνσης συνοψίζεται στα εξής :

- *Ορίζουμε τις εξομαλυμένες τιμές για κάθε ηλικία με βάση τα δύο τόξα*

$$v_x = \begin{cases} p_0(x), & a \leq x \leq k \\ p_1(x), & k \leq x \leq b \end{cases},$$

όπου τα πολυώνυμα $p_0(x)$ και $p_1(x)$ είναι τρίτου βαθμού και ικανοποιούν τις συνθήκες :

$$p_0(k) = p_1(k)\,,$$
$$p_0{}'(k) = p_1{}'(k)\,,$$
$$p_0{}''(k) = p_1{}''(k)\,.$$

- *Υπολογίζουμε τις παραμέτρους των δύο πολυωνύμων*

Με τη Μέθοδο Ελαχίστων Τετραγώνων επιδιώκουμε την ελαχιστοποίηση του μέτρου:

$$SS = \sum_{x=a}^{b} w_x * (u_x - v_x)^2 = \sum_{x=a}^{h} w_x * (u_x - p_0(x))^2 + \sum_{x=h+1}^{b} w_x * (u_x - p_1(x))^2\,,$$

όπου

$$p_0(x) = c_1 + c_2 * x + c_3 * x^2 + c_4 * x^3$$

και

$$p_1(x) = c_5 + c_6 * x + c_7 * x^2 + c_8 * x^3$$

ή

$$p_1(x) = p_0(x) + c_5 * (x - k)^3\,.$$

Τα αθροίσματα χωρίζονται στις τιμές h *και* $h+1$ *αντίστοιχα, οι οποίες είναι οι τιμές πριν και μετά τον κόμβο* k *και αποτελούν ηλικίες στις οποίες έχει υπολογιστεί πρωτογενής τιμή θνησιμότητας. Αυτό σημαίνει ότι ο κόμβος* k *δεν αποτελεί απαραίτητα πρωτογενή τιμή θνησιμότητας. Για την εύρεση των παραμέτρων των δύο πολυωνύμων είτε παίρνουμε όλες τις μερικές παραγώγους του μέτρου* SS *και τις θέτουμε ίσες με μηδέν, δηλαδή* :

$$\frac{\partial SS}{\partial c_1} = 0\,, \qquad \frac{\partial SS}{\partial c_2} = 0\,, \ldots\ldots\,, \qquad \frac{\partial SS}{\partial c_5} = 0\,,$$

είτε λύνουμε την παρακάτω εξίσωση των Πινάκων – Μητρών :

$$\boldsymbol{x}' * \boldsymbol{w} * \boldsymbol{x} * \boldsymbol{c} = \boldsymbol{x}' * \boldsymbol{w} * \boldsymbol{u}\,.$$

Οι Πίνακες της παραπάνω εξίσωσης παρουσιάζονται παρακάτω :

$$\boldsymbol{x} = \begin{bmatrix} 1 & a & a^2 & a^3 & 0 \\ 1 & a+1 & (a+1)^2 & (a+1)^3 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & h & h^2 & h^3 & 0 \\ 1 & h+1 & (h+1)^2 & (h+1)^3 & (h+1-k)^3 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & b & b^2 & b^3 & (b-k)^3 \end{bmatrix},$$

$$\boldsymbol{w} = \begin{bmatrix} w_a & 0 & . & . & 0 \\ 0 & w_{a+1} & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & w_b \end{bmatrix},$$

$$\boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ c_5 \end{bmatrix},$$

$$\boldsymbol{u} = \begin{bmatrix} u_a \\ u_{a+1} \\ . \\ . \\ u_b \end{bmatrix}.$$

*Όσον αφορά τον Πίνακα **x**, ο αριθμός των γραμμών του ισούται με το πλήθος των ηλικιών που εξομαλύνουμε και ο αριθμός των στηλών του ισούται με το άθροισμα του βαθμού των πολυωνύμων και του πλήθους των τόξων – splines στα οποία χωρίζουμε το εύρος των ηλικιών.* Οι Πίνακες $\boldsymbol{u}$ και $\boldsymbol{w}$ ορίζονται αναλόγως, όπως ορίστηκαν στη μέθοδο Whittaker και ο Πίνακας $\boldsymbol{c}$ είναι Πίνακας μιας στήλης που έχει τόσες γραμμές, όσες και οι παράμετροι που θέλουμε να εκτιμήσουμε.

Ανάλογη με την παραπάνω διαδικασία είναι η προσέγγιση της γενικής περίπτωσης στην οποία επιλέγουμε να χωρίσουμε το εύρος των ηλικιών σε n κόμβους, δημιουργώντας έτσι $n+1$ υπο – ομάδες. Θα έχουμε συνεπώς για το εύρος ηλικιών από την ηλικία a έως την ηλικία b, ορίζοντας αναλόγως τους κόμβους $k_1, k_2, \dots, k_n$ και τις τιμές $h_1, h_2, \dots, h_n$, τα εξής :

$$v_x = \begin{cases} p_0(x), & a \leq x \leq k_1 \\ \quad . & \\ \quad . & \\ p_i(x) & k_i \leq x \leq k_{i+1}\,, \\ \quad . & \\ \quad . & \\ p_n(x) & k_n \leq x \leq b \end{cases}$$

για $i = 1, 2, \dots, n$. Θεωρώντας πολυώνυμα τρίτου βαθμού (που είναι η συνήθης περίπτωση), θα πρέπει να ισχύουν οι συνθήκες :

$$\begin{gathered} p_{i-1}(k_i) = p_i(k_i)\,, \\ p_{i-1}{}'(k_i) = p_i{}'(k_i)\,, \\ p_{i-1}{}''(k_i) = p_i{}''(k_i), \end{gathered}$$

για $i = 1, 2, \dots, n$. Άρα, για τη εύρεση των παραμέτρων των τριτοβάθμιων πολυωνύμων είτε θα χρησιμοποιήσουμε την ελαχιστοποίηση του μέτρου SS της Μεθόδου Ελαχίστων Τετραγώνων, παραγωγίζοντας το μέτρο για κάθε παράμετρο, θέτοντάς το ίσο με μηδέν και λύνοντας το σύστημα, είτε θα χρησιμοποιήσουμε τη λύση μέσω Μητρών της εξίσωσης :

$$\boldsymbol{x}' * \boldsymbol{w} * \boldsymbol{x} * \boldsymbol{c} = \boldsymbol{x}' * \boldsymbol{w} * \boldsymbol{u}\,.$$

Το μέτρο SS αναπτύσσεται ως ακολούθως :

$$\mathrm{SS} = \sum_{x=a}^{h_1} w_x * (u_x - p_0(x))^2 + \sum_{x=h_1+1}^{h_2} w_x * (u_x - p_1(x))^2 + \cdots$$

$$+ \sum_{x=h_n+1}^{b} w_x * (u_x - p_n(x))^2 \ .$$

ΜΕΡΟΣ 3ο

ΚΕΦΑΛΑΙΟ 6^{o}

ΕΞΟΜΑΛΥΝΣΗ ΣΤΟΙΧΕΙΩΝ ΠΛΗΘΥΣΜΟΥ ΕΛΛΑΔΟΣ 2001

6.1 ΕΙΣΑΓΩΓΗ

Στο τελευταίο κεφάλαιο της εργασίας, ακολουθεί η εφαρμογή μίας μεθόδου εξομάλυνσης, πάνω στα στοιχεία που αφορούν τον πληθυσμό της Ελλάδος του έτους 2001. Η εξομάλυνση γίνεται για κάθε ηλικία, ξεχωριστά για κάθε φύλο. Τα στοιχεία του μεγέθους του πληθυσμού για κάθε ηλικία προέρχονται από την απογραφή πληθυσμού, η οποία διενεργήθηκε στις 18 Μαρτίου του 2001. Το μέγεθος του πληθυσμού ανά ηλικία και φύλο, σε συνδυασμό με τις ληξιαρχικές καταγραφές θανάτων ανά ηλικία και φύλο του ίδιου έτους, μας δίνουν τις πρωτογενείς τιμές θνησιμότητας οι οποίες και θα εξομαλυνθούν. Έχει επιλεγεί η μέθοδος εξομάλυνσης του *Κινούμενου Σταθμισμένου Μέσου* για αμφότερα τα φύλα, η εφαρμογή και οι λεπτομέρειες της οποίας αναπτύσσονται διεξοδικότερα στην 3^{η} παράγραφο του παρόντος κεφαλαίου. Στην παράγραφο που ακολουθεί, δίδονται τα στοιχεία πριν την εξομάλυνση και τα διαγράμματα των πρωτογενών τιμών θνησιμότητας, από την ηλικία 5 έως την ηλικία 64 και για τα δύο φύλα. Στην 4^{η} παράγραφο παρουσιάζονται τα συγκριτικά αποτελέσματα που αφορούν τις πρωτογενείς και τις εξομαλυμένες τιμές θανάτου. Τα στοιχεία της απογραφής του πληθυσμού πάρθηκαν από την *Ε.Σ.Υ.Ε.*, ενώ τα αντίστοιχα στοιχεία των ληξιαρχικών καταγραφών θανάτων από την *Eurostat*. Όσον αφορά την προσβασιμότητα σε στοιχεία που αφορούν μέγεθος πληθυσμού και θανάτους, προτείνονται τα site των ανωτέρω υπηρεσιών, καθώς και το site του *Εργαστηρίου Δημογραφικών και Κοινωνικών Αναλύσεων*. Αναλυτικότερα οι ηλεκτρονικές διευθύνσεις των 3 υπηρεσιών είναι :

http://www.e-demography.gr/index.cfm (Ε.Δ.Κ.Α.) .

http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/ (Eurostat) .

http://www.statistics.gr/portal/page/portal/ESYE (Ε.Σ.Υ.Ε.) .

6.2 ΠΡΟ ΕΞΟΜΑΛΥΝΣΗΣ – ΠΡΩΤΟΓΕΝΕΙΣ ΤΙΜΕΣ ΘΝΗΣΙΜΟΤΗΤΑΣ

6.2.1 ΑΝΔΡΕΣ

Ακολουθεί η ανάλυση των πρωτογενών τιμών θνησιμότητας που αφορούν τους άνδρες ηλικίας 5 έως 64 ετών. Ωστόσο, σε αυτό το κεφάλαιο (όμοια και για τις γυναίκες), παρουσιάζονται τα στοιχεία των πρωτογενών τιμών θνησιμότητας από την ηλικία 0 έως την ηλικία 69, καθώς οι 5 πρώτες ηλικίες (0 – 4) και οι 5 τελευταίες (65 – 69) μας χρειάζονται στη διαδικασία εξομάλυνσης, παρά το γεγονός ότι οι πρωτογενείς τους τιμές δεν εξομαλύνονται. Παρατίθενται ο πίνακας που περιλαμβάνει τους ζώντες ανά ηλικία και τους θανόντες (στοιχεία από τα οποία εξάγουμε τις πρωτογενείς τιμές θνησιμότητας), διαγράμματα για όλο το εύρος των παραπάνω ηλικιών, διαγράμματα ανά κλάσεις ηλικιών καθώς και σχολιασμός τους. Ανάλογη είναι η προσέγγιση που ακολουθείται στην επόμενη υπο – παράγραφο που αφορά τις γυναίκες, στην οποία παρουσιάζονται τα αντίστοιχα στοιχεία.

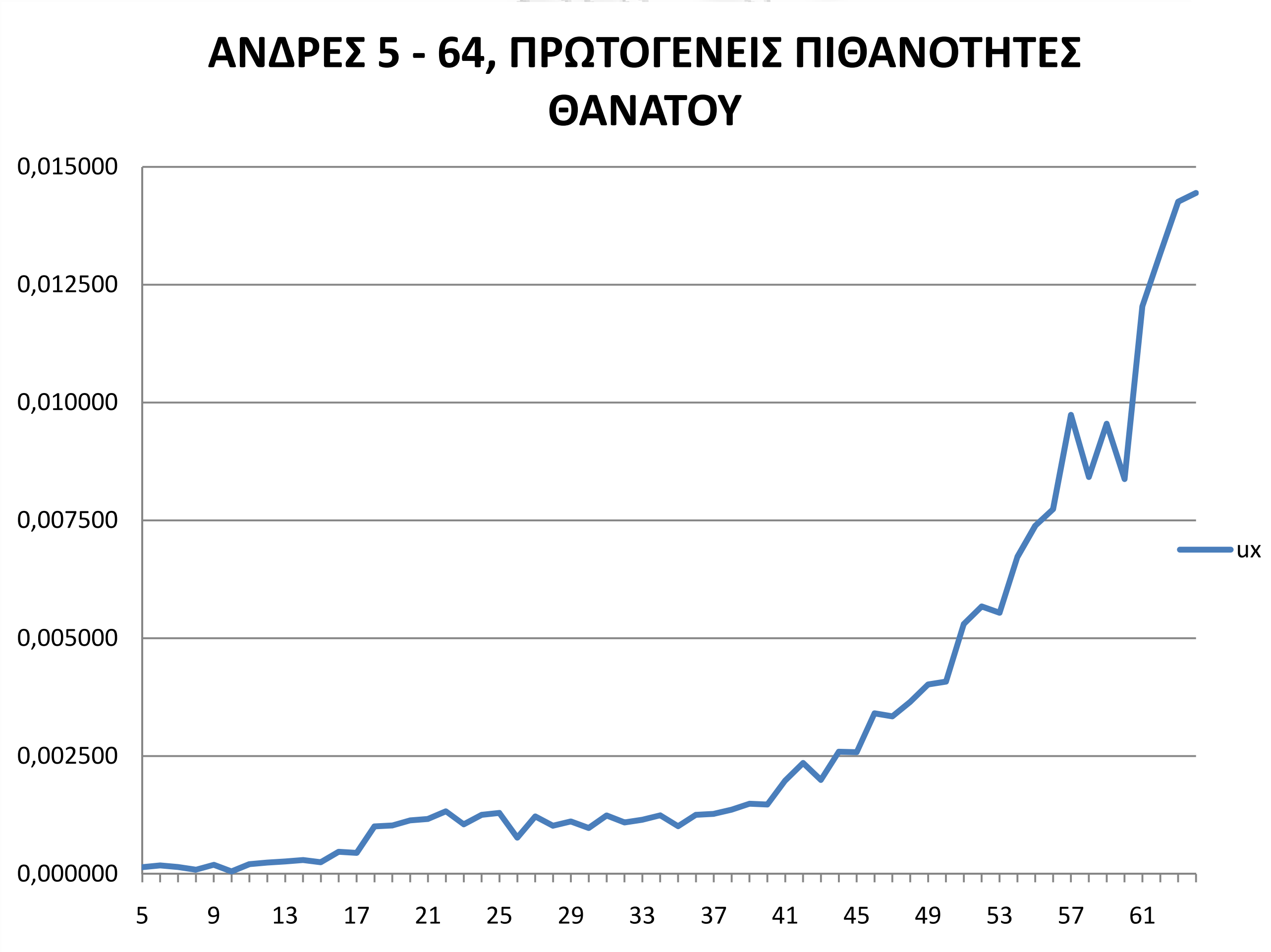
ΠΙΝΑΚΑΣ ΠΡΩΤΟΓΕΝΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 0 – 69

x	n_x	d_x	u_x
0	62.897	307	0,004881
1	51.153	17	0,000332
2	51.051	8	0,000157
3	52.815	10	0,000189
4	53.390	6	0,000112
5	55.761	8	0,000143
6	55.725	10	0,000179
7	54.870	8	0,000146
8	56.891	5	0,000088
9	57.237	11	0,000192
10	59.932	3	0,000050
11	57.779	12	0,000208
12	62.068	15	0,000242
13	60.711	16	0,000264
14	64.914	19	0,000293
15	68.704	17	0,000247
16	72.468	34	0,000469
17	76.185	34	0,000446
18	80.490	81	0,001006
19	82.653	85	0,001028
20	89.081	101	0,001134
21	84.142	98	0,001165
22	86.698	115	0,001326

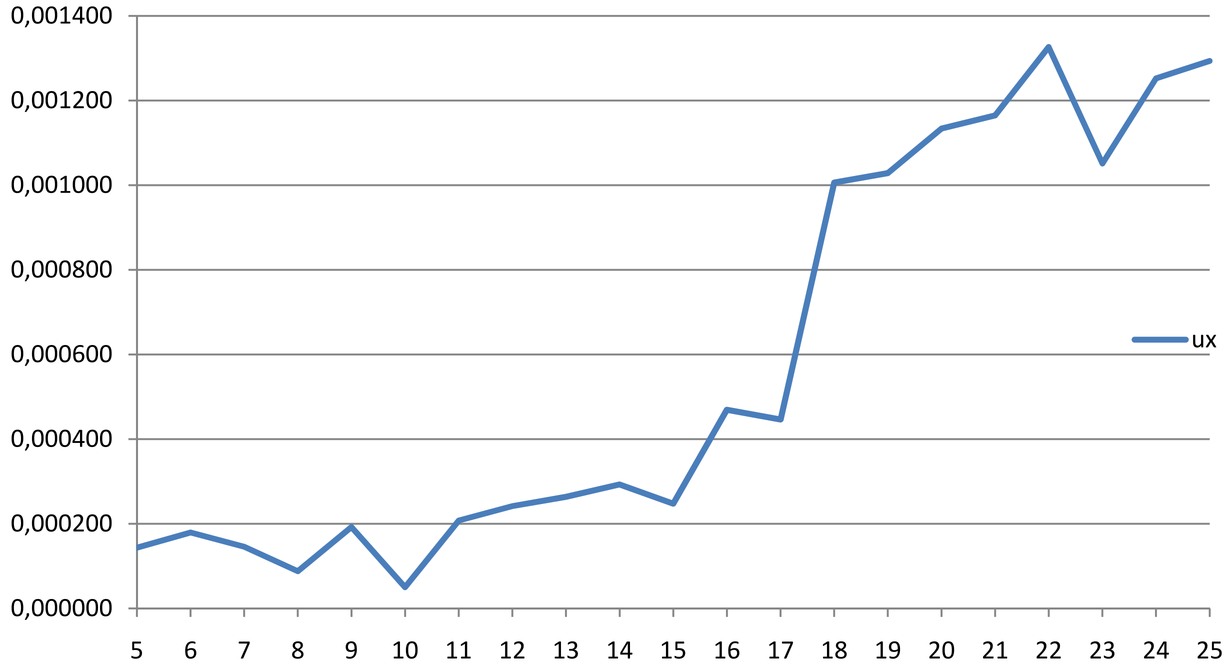
23	88.473	93	0,001051
24	88.624	111	0,001252
25	88.134	114	0,001293
26	89.891	69	0,000768
27	85.414	104	0,001218
28	87.165	89	0,001021
29	85.595	95	0,001110
30	91.474	89	0,000973
31	86.384	107	0,001239
32	88.084	96	0,001090
33	88.807	102	0,001149
34	86.794	108	0,001244
35	83.240	84	0,001009
36	79.812	100	0,001253
37	76.917	98	0,001274
38	76.494	104	0,001360
39	75.863	113	0,001490
40	84.320	124	0,001471
41	77.342	153	0,001978
42	75.683	178	0,002352
43	74.707	149	0,001994
44	75.592	196	0,002593
45	75.939	196	0,002581
46	71.939	245	0,003406
47	68.531	229	0,003342
48	69.329	253	0,003649
49	70.397	283	0,004020
50	72.828	297	0,004078
51	60.939	323	0,005300
52	63.089	358	0,005675
53	69.689	386	0,005539
54	71.545	481	0,006723
55	65.107	481	0,007388
56	59.852	463	0,007736
57	49.295	480	0,009737
58	50.236	423	0,008420
59	46.605	445	0,009548
60	64.104	537	0,008377
61	57.148	688	0,012039
62	59.114	778	0,013161
63	57.720	823	0,014258
64	60.095	868	0,014444
65	60.264	982	0,016295
66	62.730	1.078	0,017185
67	58.462	1.186	0,020287
68	54.604	1.220	0,022343
69	55.540	1.219	0,021948

Ο παραπάνω πίνακας παρουσιάζει στην 1η στήλη του τις ηλικίες **x**, που κυμαίνονται από 0 έως 69 έτη. Στη 2η στήλη παρουσιάζονται για το αντίστοιχο εύρος των ηλικιών, οι ζώντες της κάθε ηλικίας $\mathbf{n_x}$, όπως αυτοί προέκυψαν από τη διαδικασία της απογραφής. Στην 3η στήλη καταγράφονται οι θάνατοι κάθε ηλικίας $\mathbf{d_x}$ (ή $\mathbf{h_x}$ σύμφωνα με το βασικό συμβολισμό της εξομάλυνσης), όπως αυτοί προέκυψαν από τις ληξιαρχικές καταγραφές θανάτων του έτους 2001. Στην τελευταία στήλη προκύπτουν οι πρωτογενείς τιμές θνησιμότητας $\mathbf{u_x}$ ως το πηλίκο της 3ης στήλης προς τη 2η και οι οποίες θα εξομαλυνθούν σε όλο τους το εύρος, πλην των ηλικιών 0 – 4 και 65 – 69. Ακολουθούν τα διαγράμματα των πρωτογενών τιμών θνησιμότητας, που αφορούν μόνο τις ηλικίες που θα εξομαλυνθούν 5 – 64, για όλο το εύρος των ηλικιών και για τις αντίστοιχες υπο – ομάδες ηλικιών 5 – 25, 26 – 45 και 46 – 64.

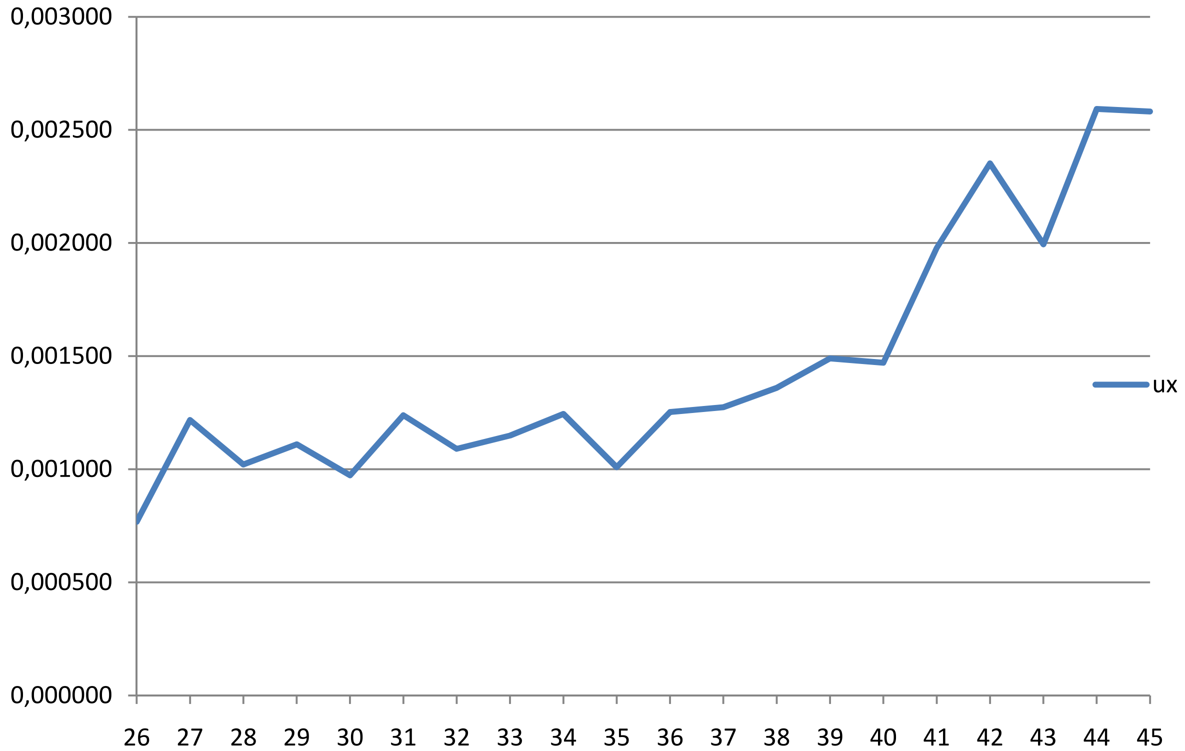
<u>ΔΙΑΓΡΑΜΜΑΤΑ ΠΡΩΤΟΓΕΝΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 5 - 64</u>

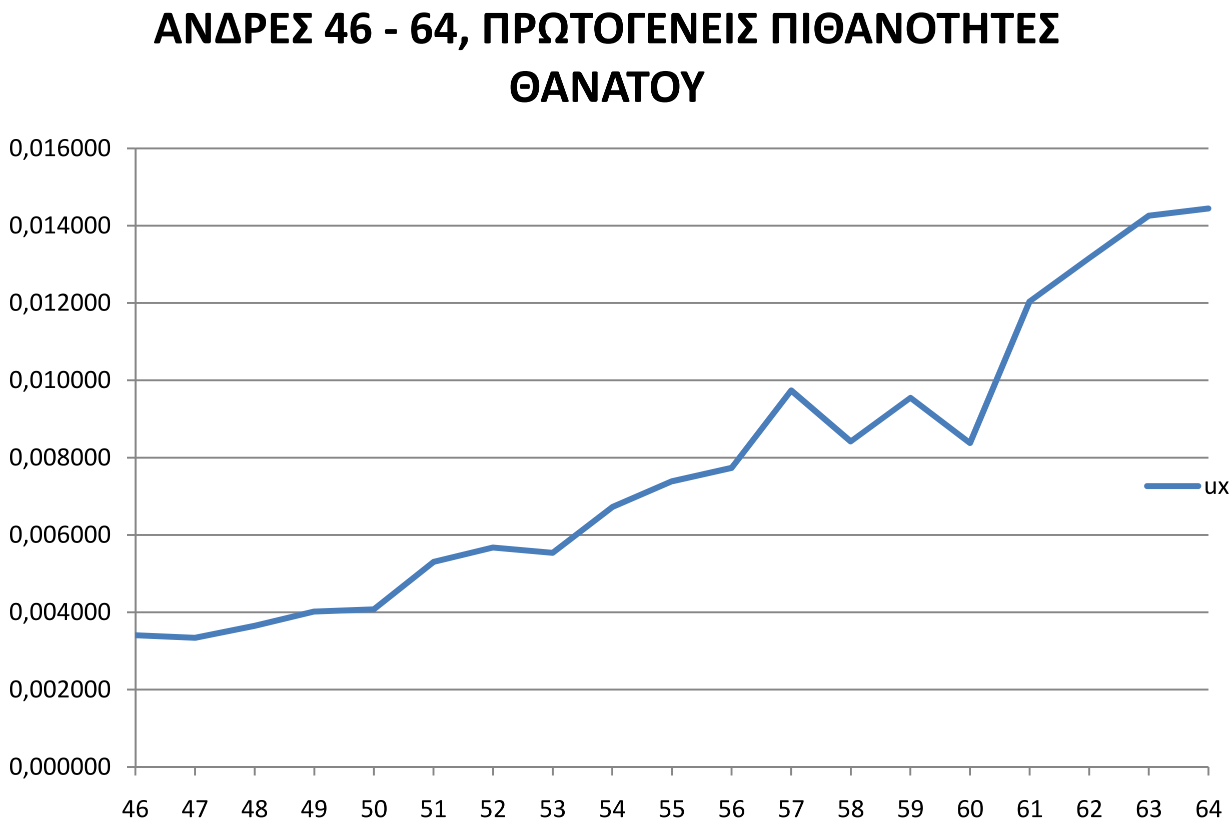


ΑΝΔΡΕΣ 5 - 25, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ



ΑΝΔΡΕΣ 26 - 45, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ





<u>ΣΧΟΛΙΑΣΜΟΣ</u>

Από τα παραπάνω γραφήματα, καθώς και τον πίνακα, παρατηρούμε την αυξημένη θνησιμότητα των δύο πρώτων ηλικιών 0 και 1 (ιδιαίτερα της ηλικίας 0). Στις επόμενες ηλικίες οι πιθανότητες ομαλοποιούνται αρκετά, ωστόσο παρατηρούμε ότι στις ηλικίες 8 και ιδιαίτερα στην ηλικία 10 έχουμε αξιοσημείωτα μικρές πρωτογενείς τιμές θνησιμότητας. Εν συνεχεία, άξια παρατήρησης είναι η αύξηση των πιθανοτήτων θανάτου μεταξύ των ηλικιών 18 – 27 (πλην της ηλικίας 26), το λεγόμενο accident hump (όπως έχει ήδη αναφερθεί στο $3^{ο}$ Κεφάλαιο της παρούσας εργασίας). Παρατηρούμε, τέλος, ότι από την ηλικία των 40 και μετά, η θνησιμότητα παρουσιάζει βαθμιαία αυξητική τάση με εξαίρεση κάποια ‘‘άλματα’’ των πιθανοτήτων στις ηλικίες 43 και 60. Μια ομαλότερη εικόνα της θνησιμότητας παρατηρούμε μέσω των διαγραμμάτων στις ηλικίες 11 – 14, 36 – 40 και ιδιαίτερα στο εύρος ηλικιών 44 – 56 (εκτός των ηλικιών 50 και 53).

6.2.2 ΓΥΝΑΙΚΕΣ

Όμοια για τις γυναίκες έχουμε :

ΠΙΝΑΚΑΣ ΠΡΩΤΟΓΕΝΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 0 - 69

x	n_x	d_x	u_x
0	59.126	215	0,003636
1	48.605	14	0,000288
2	48.380	10	0,000207
3	50.580	2	0,000040
4	51.402	5	0,000097
5	52.139	9	0,000173
6	52.417	9	0,000172
7	52.290	6	0,000115
8	53.450	7	0,000131
9	54.325	1	0,000018
10	55.400	6	0,000108
11	52.906	6	0,000113
12	56.940	7	0,000123
13	56.121	6	0,000107
14	59.624	12	0,000201
15	62.544	8	0,000128
16	65.884	17	0,000258
17	69.194	16	0,000231
18	73.459	20	0,000272
19	74.593	21	0,000282
20	81.066	29	0,000358
21	76.642	21	0,000274
22	79.065	24	0,000304
23	80.763	25	0,000310
24	80.909	23	0,000284
25	82.151	11	0,000134
26	84.552	20	0,000237
27	80.740	25	0,000310
28	82.627	31	0,000375
29	81.158	35	0,000431
30	86.271	27	0,000313
31	83.364	31	0,000372
32	86.003	36	0,000419
33	87.550	40	0,000457
34	85.201	50	0,000587
35	81.521	42	0,000515
36	79.549	36	0,000453
37	77.161	52	0,000674
38	77.433	51	0,000659
39	75.423	48	0,000636
40	85.632	55	0,000642

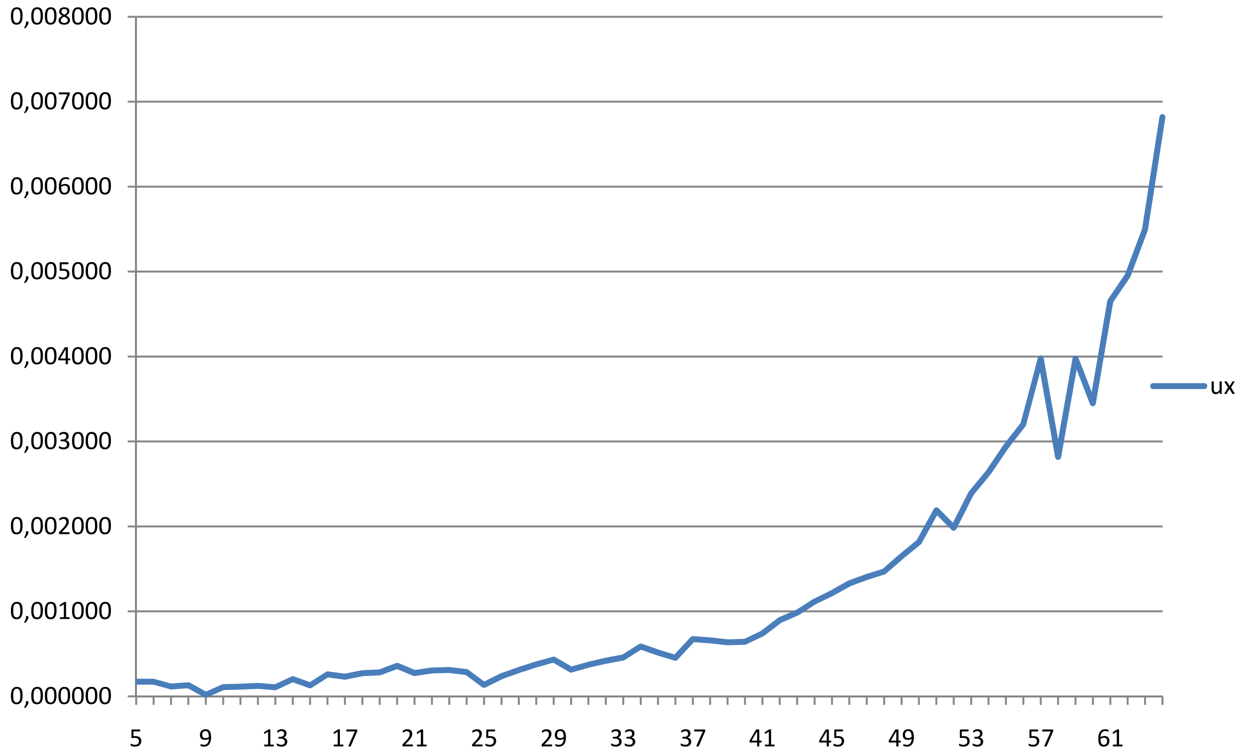
41	78.373	58	0,000740
42	76.956	69	0,000897
43	76.161	75	0,000985
44	77.177	86	0,001114
45	76.528	93	0,001215
46	72.972	97	0,001329
47	68.355	96	0,001404
48	70.778	104	0,001469
49	69.207	114	0,001647
50	74.333	135	0,001816
51	62.108	136	0,002190
52	66.537	132	0,001984
53	71.986	172	0,002389
54	74.295	196	0,002638
55	67.651	199	0,002942
56	63.731	204	0,003201
57	51.535	205	0,003978
58	54.640	154	0,002818
59	51.563	205	0,003976
60	74.205	256	0,003450
61	63.207	294	0,004651
62	68.431	339	0,004954
63	65.516	360	0,005495
64	70.534	481	0,006819
65	68.786	524	0,007618
66	70.468	567	0,008046
67	68.164	595	0,008729
68	65.581	662	0,010094
69	58.646	729	0,012431

ΣΧΟΛΙΑΣΜΟΣ

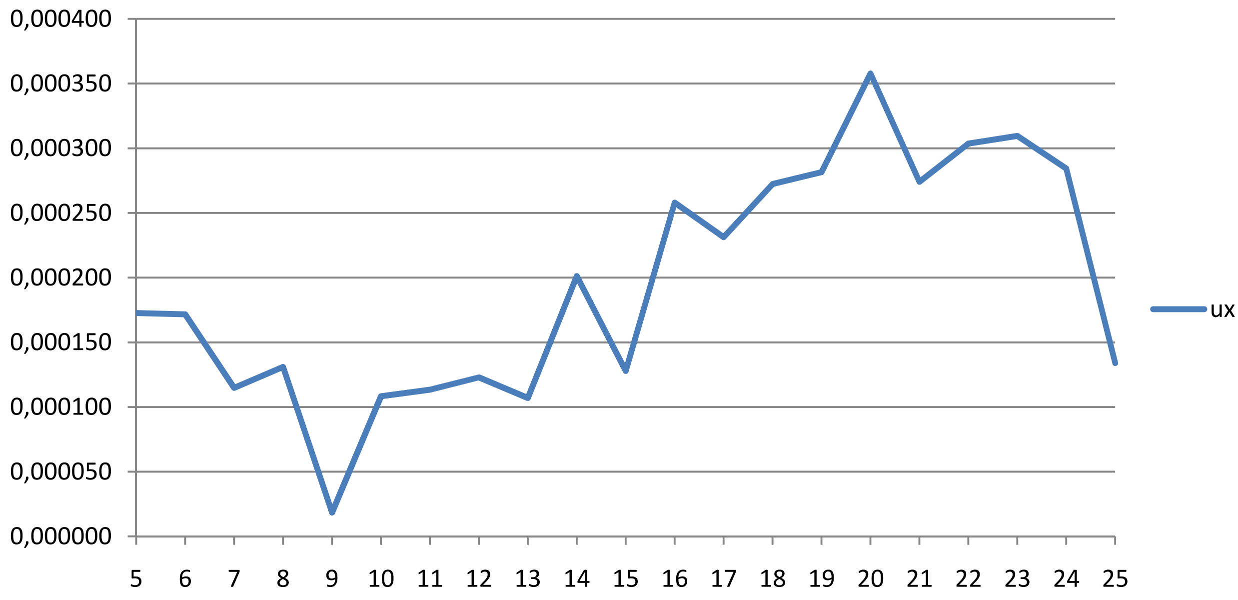
Από τον πίνακα καθώς και τα διαγράμματα που ακολουθούν, παρατηρούμε την αυξημένη θνησιμότητα των τριών πρώτων ηλικιών των γυναικών. Στη συνέχεια η θνησιμότητα μειώνεται δραστικά, ενώ αξιοσημείωτη είναι η σχεδόν αμελητέα πρωτογενής τιμή θνησιμότητας της ηλικίας 9. Μέχρι την ηλικία των 25 (και ιδιαίτερα σε αυτήν την ηλικία), παρατηρούμε συνεχείς αυξομειώσεις των πρωτογενών τιμών οι οποίες προκαλούν και ‘‘άλματα’’ στο αντίστοιχο διάγραμμα ηλικιών 5 – 25. Στη συνέχεια η θνησιμότητα ομαλοποιείται αρκετά παρουσιάζοντας αυξομειώσεις (ηλικίες 30, 36), ενώ από την ηλικία των 40 και μετά παρατηρούμε μια ομαλά αυξητική τάση της θνησιμότητας, πλην των ηλικιών 58 – 60 όπου παρατηρούμε σημαντικές αυξομειώσεις.

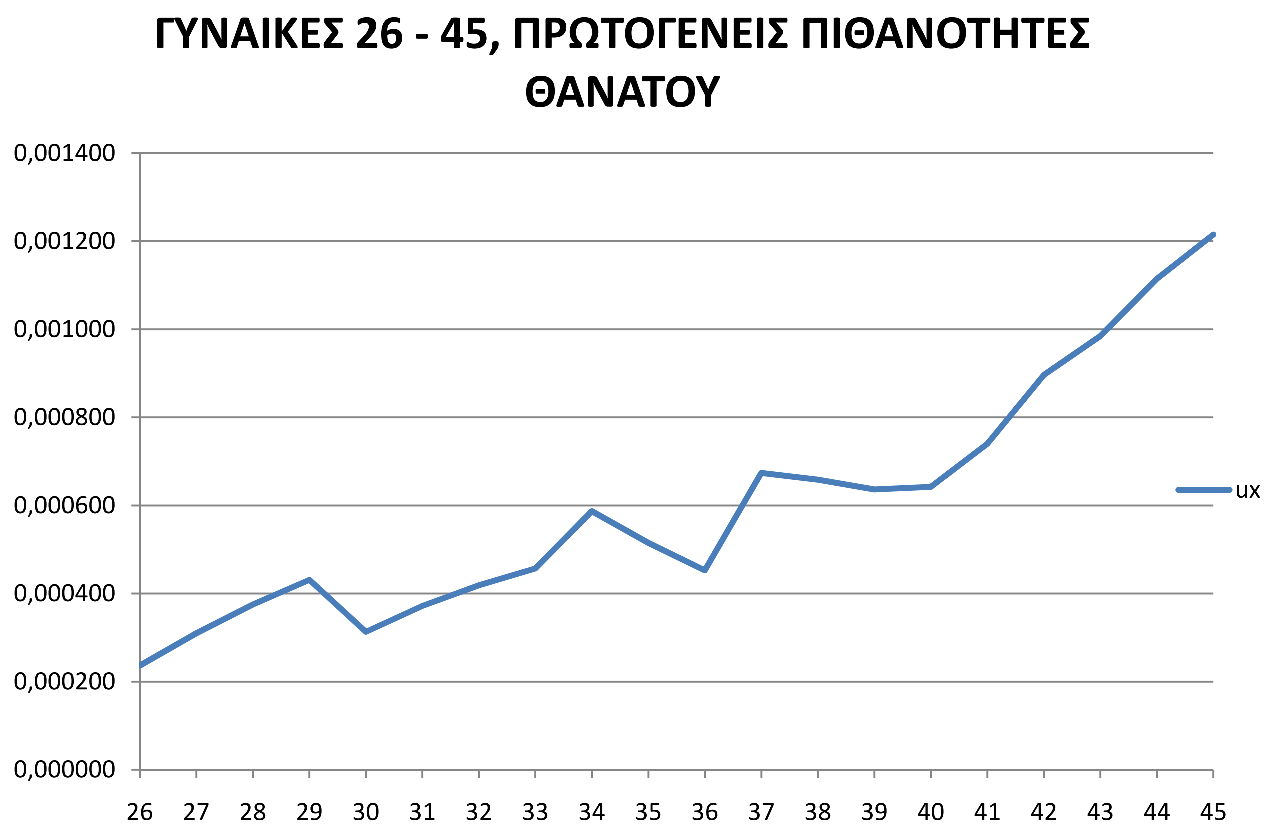
ΔΙΑΓΡΑΜΜΑΤΑ ΠΡΩΤΟΓΕΝΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 5 - 64

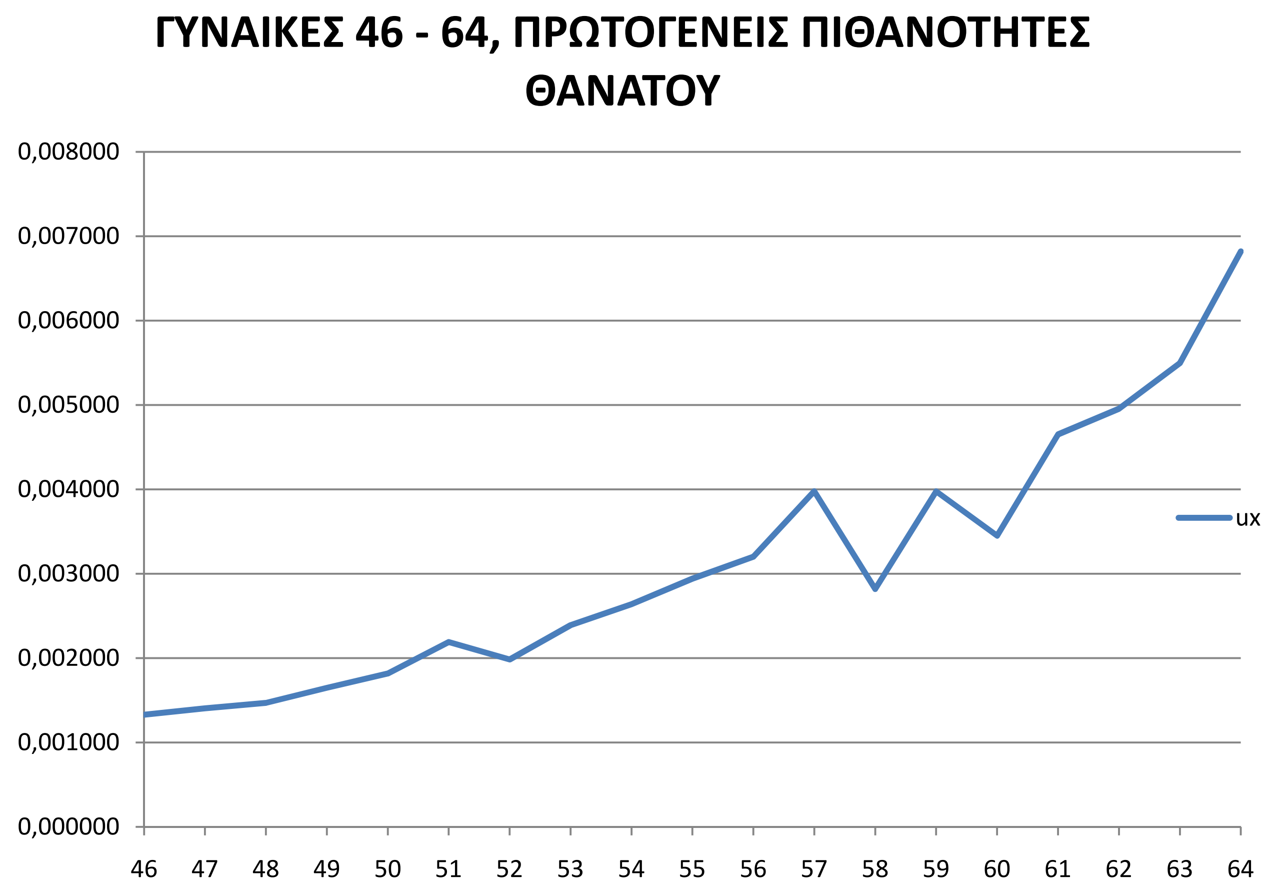
ΓΥΝΑΙΚΕΣ 5 - 64, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ



ΓΥΝΑΙΚΕΣ 5 - 25, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ




ΓΥΝΑΙΚΕΣ 26 - 45, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ
0,001400
0,001200
0,001000
0,000800
0,000600
0,000400
0,000200
0,000000
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
ux


ΓΥΝΑΙΚΕΣ 46 - 64, ΠΡΩΤΟΓΕΝΕΙΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ
0,008000
0,007000
0,006000
0,005000
0,004000
0,003000
0,002000
0,001000
0,000000
46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64
ux

6.3 ΕΞΟΜΑΛΥΝΣΗ – ΕΞΟΜΑΛΥΜΕΝΕΣ ΤΙΜΕΣ ΘΝΗΣΙΜΟΤΗΤΑΣ

6.3.1 ΠΛΑΙΣΙΟ & ΠΑΡΑΔΟΧΕΣ ΜΕΘΟΔΟΥ ΕΞΟΜΑΛΥΝΣΗΣ

Όπως αναφέρθηκε στην εισαγωγή του παρόντος Κεφαλαίου, έχει επιλεγεί η χρήση της μη – παραμετρικής μεθόδου εξομάλυνσης *Κινούμενου Σταθμισμένου Μέσου (Moving Weighted Average)*, για την εξομάλυνση των πρωτογενών τιμών θνησιμότητας *αμφότερων* των φύλων. Στο συγκεκριμένο υπο – κεφάλαιο παρατίθεται το γενικό πλαίσιο της μεθόδου καθώς και οι *παραδοχές* στις οποίες στηρίχθηκε η όλη εφαρμογή της συγκεκριμένης μεθόδου. Το πλαίσιο της μεθόδου έχει δοθεί στο 4^{o} Κεφάλαιο (συγκεκριμένα στο 4.1) και η *εφαρμογή της μεθόδου στηρίζεται στο λήμμα που παρουσιάζεται προς το τέλος του 4.1*. Παρατίθενται ακολούθως, οι παραδοχές στις οποίες προβήκαμε για την εφαρμογή της μεθόδου. *Σημειώνεται ότι οι παραδοχές και τα αποτελέσματα που προκύπτουν μέσω αυτών, ισχύουν εξίσου για τα δύο φύλα.*

Θεωρούμε, λοιπόν, συμμετρικό τύπο Κινούμενου Σταθμισμένου Μέσου, για τον οποίο η κάθε εξομαλυμένη τιμή προκύπτει ως στάθμιση 11 ($2n+1$ *) πρωτογενών τιμών θνησιμότητας (το οποίο σημαίνει n = 5, 5 πριν, 5 μετά και μία της ίδιας της ηλικίας που εξομαλύνουμε)*. Συνεπώς θα έχουμε :

$$v_x = \sum_{r=-n}^{n} u_{x+r} * a_r = \sum_{r=-5}^{5} u_{x+r} * a_r \,.$$

Η συμμετρία του τύπου δηλώνει ότι $a_r = a_{-r}$ *για* $r = 1,2,3,4,5$ *. Γι' αυτόν το λόγο η εξομάλυνση περιορίζεται στο εύρος των ηλικιών 5 έως 64 (σύνολο 60 παρατηρήσεων), καθώς για τις ηλικίες 0 – 4 και 65 – 69 έχουμε έλλειψη αντίστοιχων στοιχείων πριν και μετά. Βασιζόμενοι στο λήμμα που αναφέρθηκε παραπάνω και εφαρμόζοντάς το, κάνουμε τις εξής παραδοχές. Θεωρούμε ότι ο παραπάνω συμμετρικός τύπος των εξομαλυμένων τιμών είναι ακριβής για πολυώνυμο* $3^{ου}$ *βαθμού, το οποίο σημαίνει ότι* $d = 3$ *. Επίσης, θεωρούμε* $z = 3$*. Από το λήμμα προκύπτει ότι οι συντελεστές στάθμισης* $a_r = a_{-r}$ *για* $r = 1,2,3,4,5$ *που θα χρησιμοποιηθούν για την εύρεση των εξομαλυμένων τιμών, αποτελούν τιμές ενός άρτιου πολυωνύμου βαθμού* $2z + d - 1 = 8$. *Συνεπώς, θα έχουμε την εξής σχέση για τα* a_r *:*

$$a_r = a + b * r^2 + c * r^4 + d * r^6 + e * r^8, \qquad r = 1,2,3,4,5\,.$$

Επίσης, ταυτόχρονα ισχύουν οι εξής συνθήκες όσον αφορά τα a_r :

$$\sum_{r=-5}^{5} a_r = 1, \quad \sum_{r=-5}^{5} r^2 * a_r = 0, \quad \sum_{r=-5}^{5} r^4 * a_r = 0, \quad \sum_{r=-5}^{5} r^6 * a_r = 0,$$
$$\sum_{r=-5}^{5} r^8 * a_r = 0\,.$$

Αναπτύσσοντας τις παραπάνω συνθήκες καταλήγουμε στο εξής σύστημα 5 εξισώσεων με 5 αγνώστους :

$$\boldsymbol{x} * \boldsymbol{y} = \boldsymbol{r}\,,$$

όπου οι αντίστοιχες Μήτρες είναι οι εξής :

$$\boldsymbol{x} = \begin{bmatrix} 11 & 110 & 1958 & 41030 & 925958 \\ 110 & 1958 & 41030 & 925958 & 21748550 \\ 1958 & 41030 & 925958 & 21748550 & 522906758 \\ 41030 & 925958 & 21748550 & 522906758 & 12753500870 \\ 925958 & 21748550 & 522906758 & 12753500870 & 313851940358 \end{bmatrix},$$

$$\boldsymbol{y} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix},$$

$$\boldsymbol{r} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Λύνοντας το παραπάνω σύστημα προκύπτουν οι εξής τιμές των παραμέτρων :

$$a = 0{,}6562818, \quad b = -0{,}4416809, \quad c = 0{,}07641576,$$
$$d = -0{,}004676213, \quad e = 0{,}000091374\,.$$

Εν συνεχεία, προκύπτουν μέσω του άρτιου πολυωνύμου, βαθμού 8 οι τιμές των συντελεστών στάθμισης ως εξής :

$$a_{-5} = a_5 = 0{,}0013553937\,,$$
$$a_{-4} = a_4 = -0{,}0136423293\,,$$
$$a_{-3} = a_3 = 0{,}0613775685\,,$$
$$a_{-2} = a_2 = -0{,}1636754589\,,$$
$$a_{-1} = a_1 = 0{,}2864318213\,,$$
$$a_0 = 0{,}6562818\,.$$

Παρατηρούμε ότι όσο πιο κοντά πλησιάζουμε στην τιμή που θα εξομαλύνουμε, τόσο αυξάνεται η αντίστοιχη βαρύτητα των συντελεστών στάθμισης. Έτσι, έχοντας βρει τους συντελεστές στάθμισης, προχωράμε δίνοντας στο επόμενο υπο – κεφάλαιο, τις εξομαλυμένες τιμές όπως αυτές προκύπτουν από το συμμετρικό τύπο της μεθόδου Κινούμενου Σταθμισμένου Μέσου. Προηγείται η ανάλυση των ανδρών και έπεται η αντίστοιχη παράθεση των εξομαλυμένων πιθανοτήτων και των διαγραμμάτων για τις γυναίκες. Στη συνέχεια, αναλύονται κάποια σημαντικά μέτρα, όπως αυτά προκύπτουν μέσα από την εφαρμοσμένη εξομάλυνση και γίνονται και δύο στατιστικοί έλεγχοι σχετικά με την εφαρμοσμένη μέθοδο εξομάλυνσης. Ο σχολιασμός όσον αφορά τις εξομαλυμένες τιμές, παρατίθεται στη συγκριτική ανάλυση στο τέλος του Κεφαλαίου.

6.3.2 ΑΝΔΡΕΣ

Ακολουθεί ο πίνακας των εξομαλυμένων τιμών και τα αντίστοιχα διαγράμματα με αυτά που παρουσιάστηκαν για τις πρωτογενείς τιμές. Σχολιασμός των αποτελεσμάτων παρατίθεται στη συγκριτική ανάλυση των πρωτογενών και των εξομαλυμένων τιμών που έπεται στο $5^{ο}$ υπο - κεφάλαιο. Όμοια ισχύει και για τις γυναίκες.

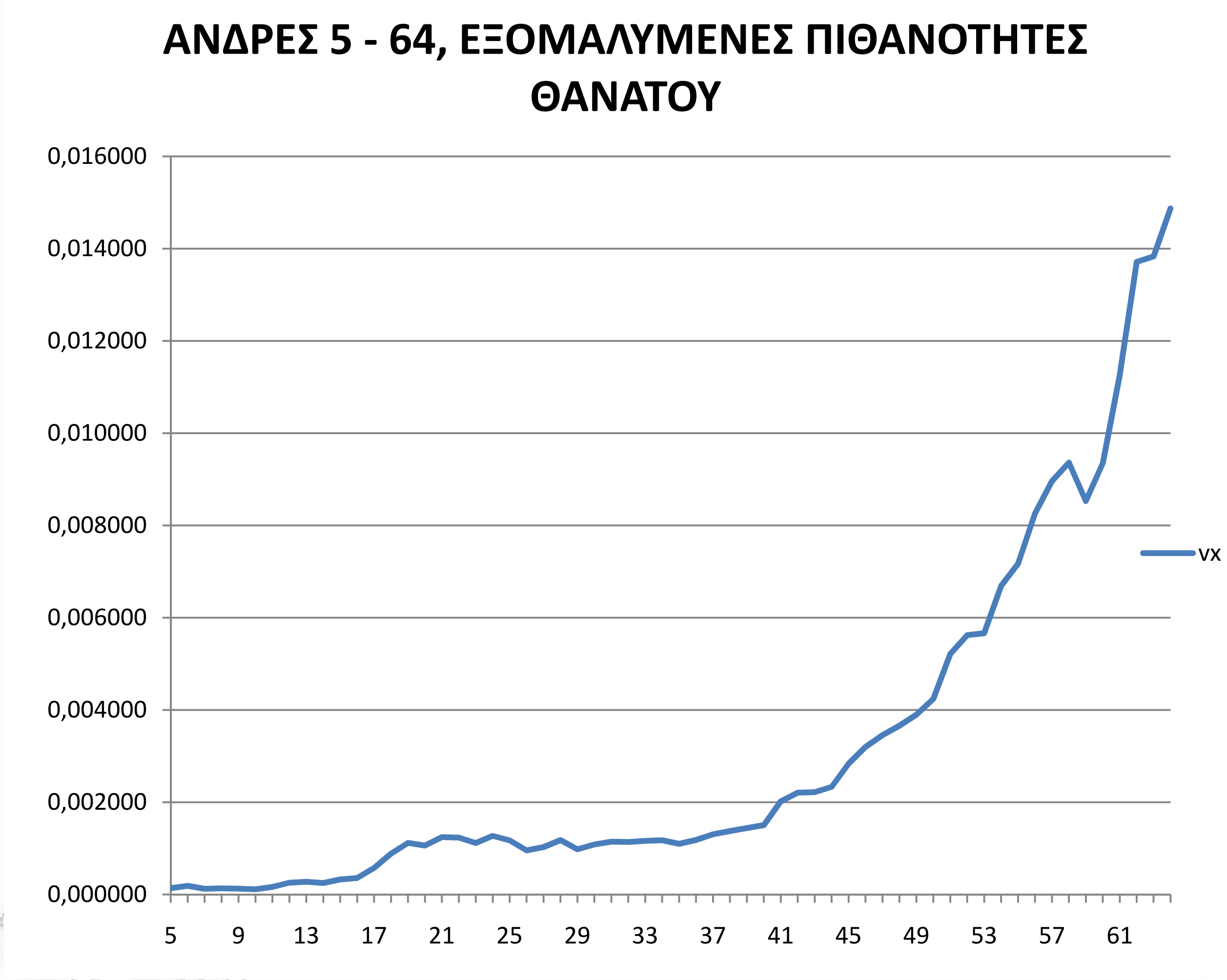
ΠΙΝΑΚΑΣ ΕΞΟΜΑΛΥΜΕΝΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 5 - 64

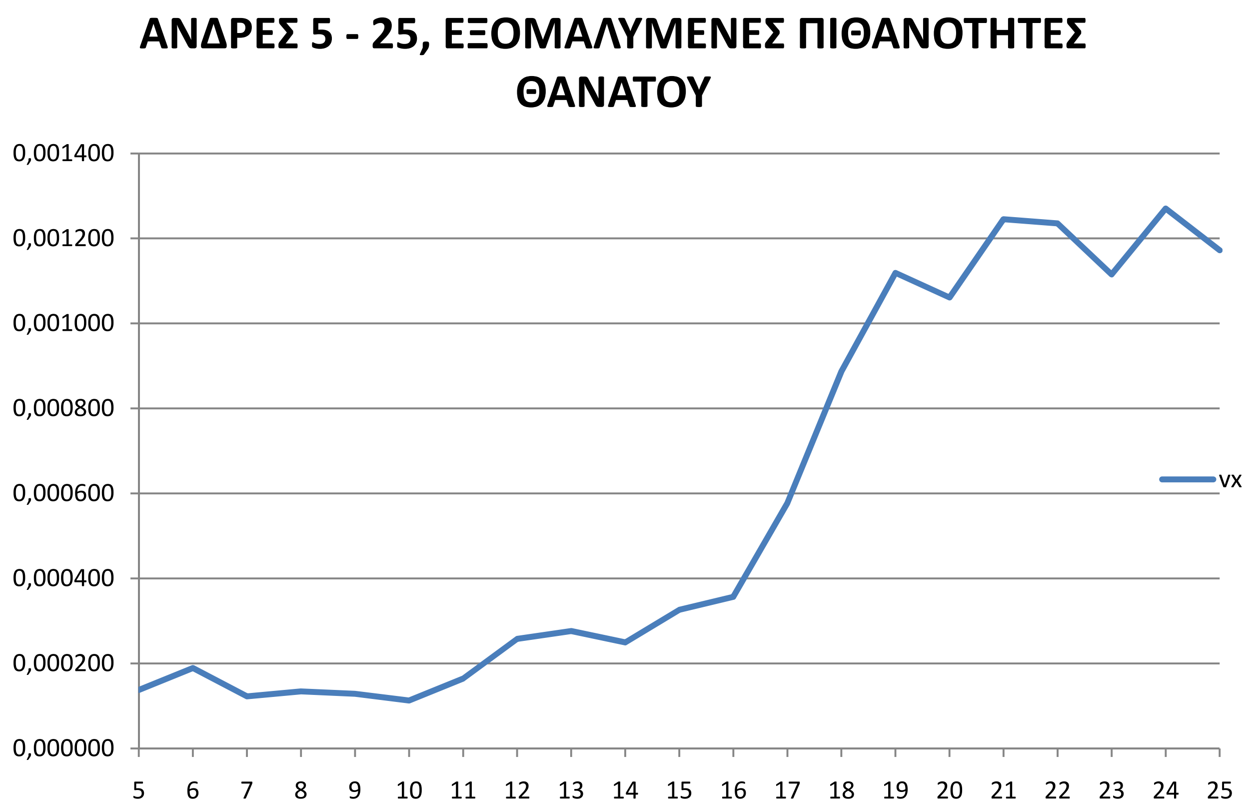
x	**v_x**
5	0,000137
6	0,000189
7	0,000122

8	0,000134
9	0,000129
10	0,000113
11	0,000164
12	0,000258
13	0,000276
14	0,000249
15	0,000326
16	0,000356
17	0,000577
18	0,000887
19	0,001119
20	0,001061
21	0,001245
22	0,001235
23	0,001115
24	0,001271
25	0,001172
26	0,000955
27	0,001027
28	0,001178
29	0,000982
30	0,001087
31	0,001146
32	0,001139
33	0,001162
34	0,001177
35	0,001097
36	0,001181
37	0,001304
38	0,001377
39	0,001441
40	0,001504
41	0,002019
42	0,002207
43	0,002219
44	0,002334
45	0,002833
46	0,003201
47	0,003454
48	0,003660
49	0,003896
50	0,004240
51	0,005213
52	0,005624
53	0,005663
54	0,006691
55	0,007169

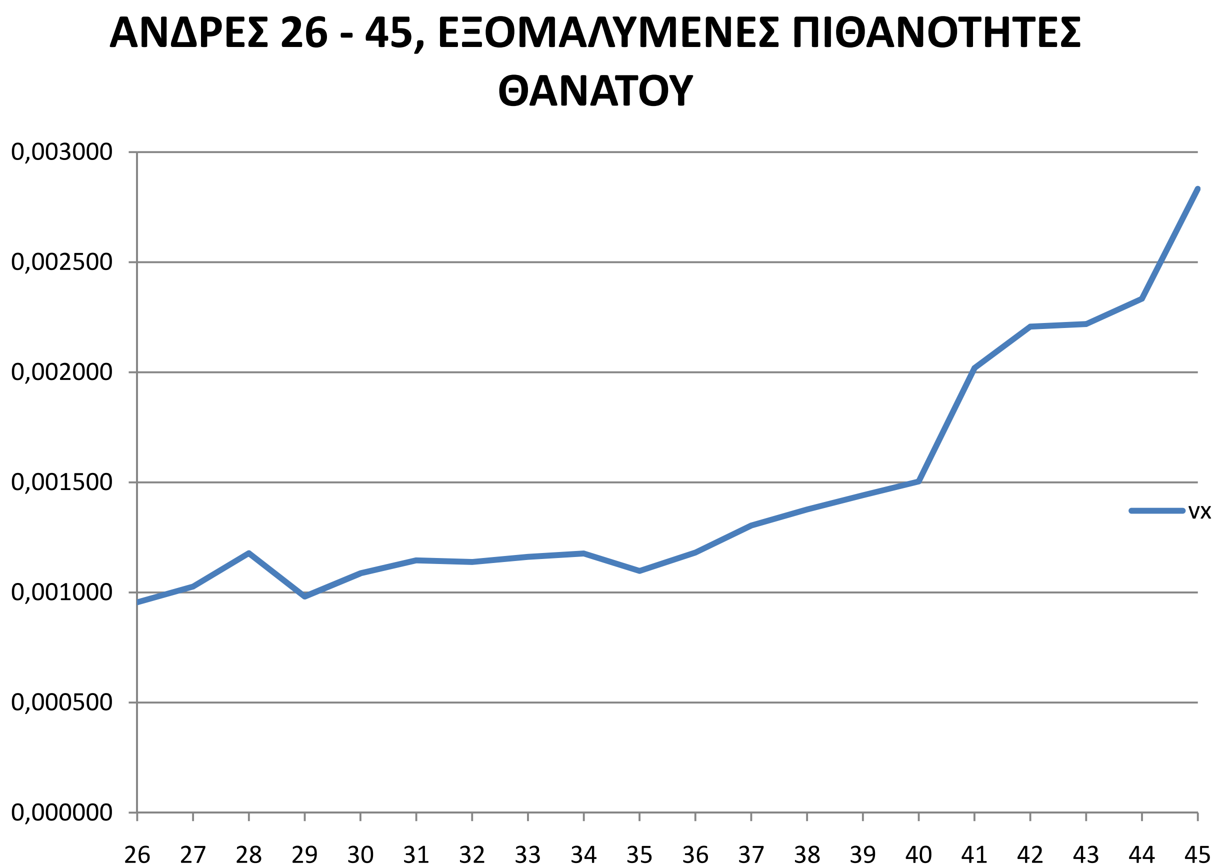
56	0,008261
57	0,008959
58	0,009361
59	0,008529
60	0,009351
61	0,011255
62	0,013712
63	0,013831
64	0,014871

<u>ΔΙΑΓΡΑΜΜΑΤΑ ΕΞΟΜΑΛΥΜΕΝΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 5 - 64</u>

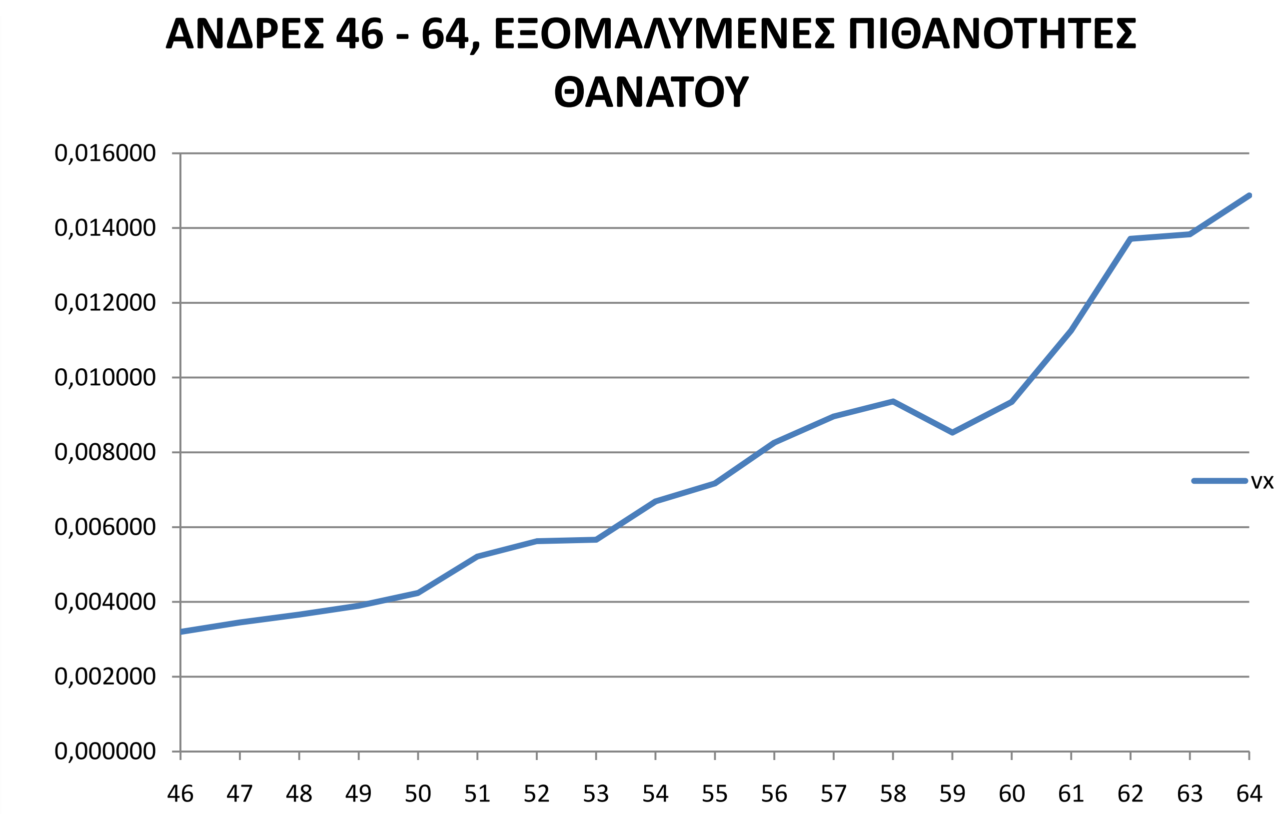




ΑΝΔΡΕΣ 5 - 25, ΕΞΟΜΑΛΥΜΕΝΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ
0,001400
0,001200
0,001000
0,000800
0,000600
0,000400
0,000200
0,000000
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
vx



ΑΝΔΡΕΣ 26 - 45, ΕΞΟΜΑΛΥΜΕΝΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ
0,003000
0,002500
0,002000
0,001500
0,001000
0,000500
0,000000
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
vx



6.3.3 ΓΥΝΑΙΚΕΣ

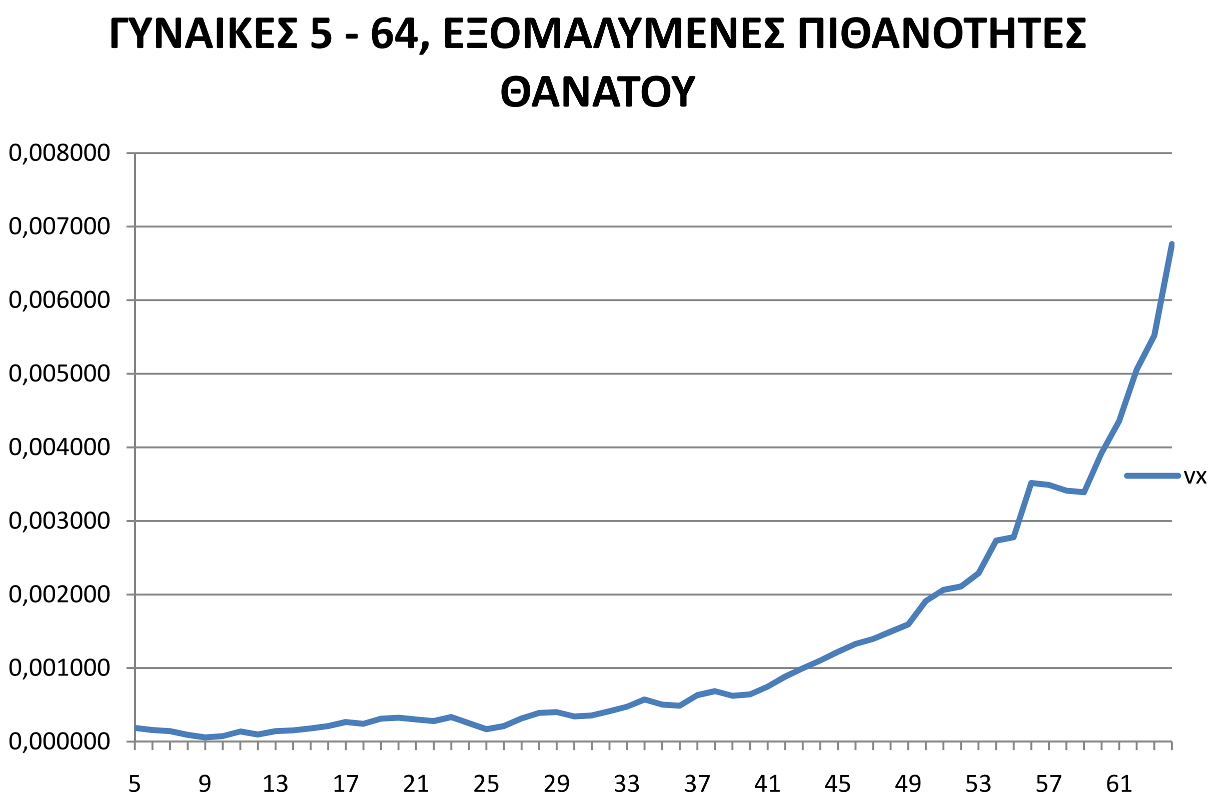
Όμοια για τις γυναίκες έχουμε για τις εξομαλυμένες τιμές και τα αντίστοιχα διαγράμματα :

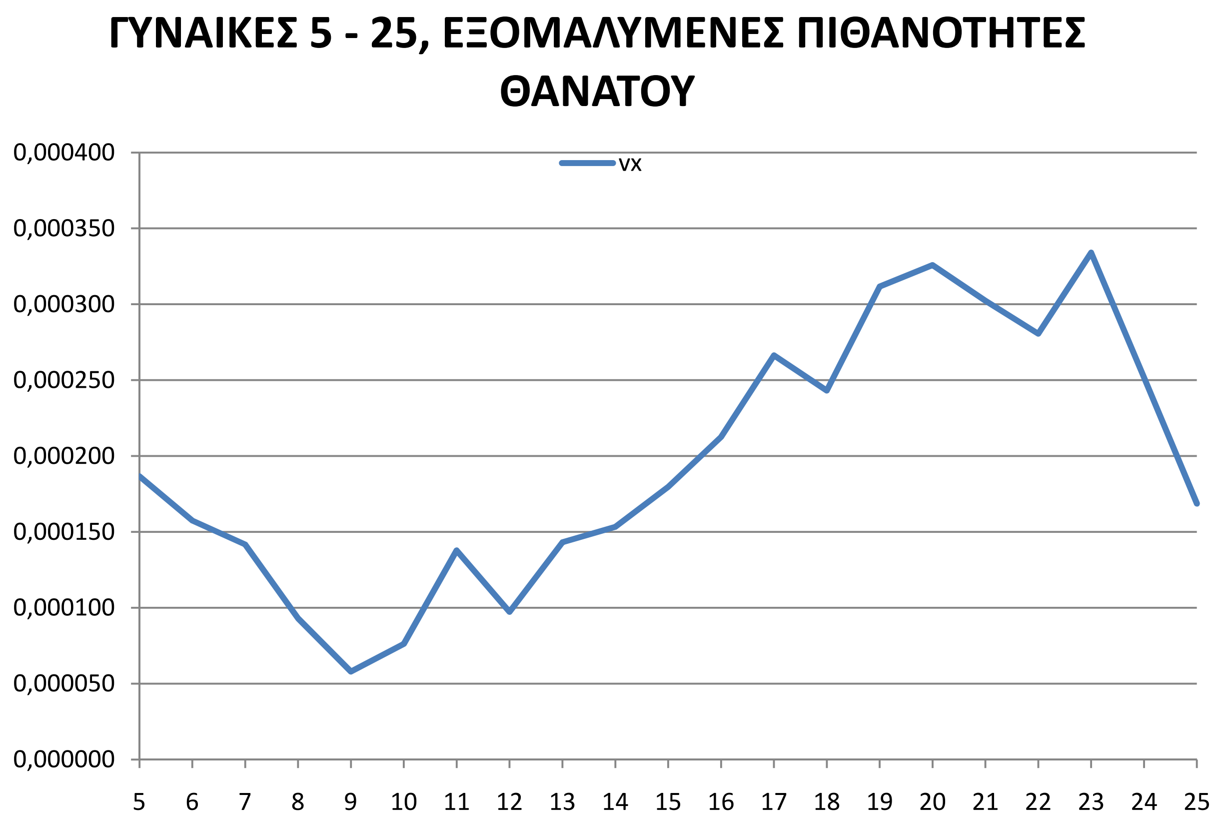
ΠΙΝΑΚΑΣ ΕΞΟΜΑΛΥΜΕΝΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 5 - 64

x	v_x
5	0,000187
6	0,000157
7	0,000142
8	0,000093
9	0,000058
10	0,000076
11	0,000138
12	0,000097
13	0,000143
14	0,000153
15	0,000180
16	0,000213
17	0,000266
18	0,000243

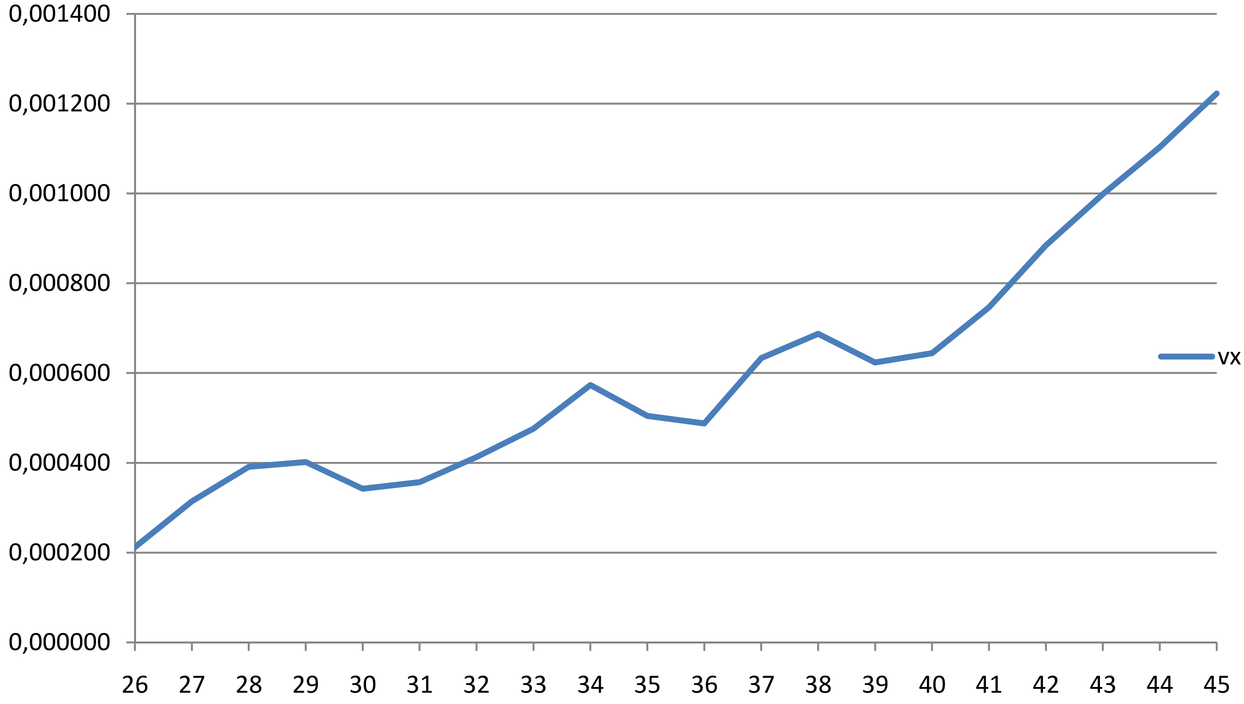
19	0,000312
20	0,000326
21	0,000302
22	0,000281
23	0,000334
24	0,000252
25	0,000169
26	0,000212
27	0,000314
28	0,000391
29	0,000402
30	0,000342
31	0,000357
32	0,000413
33	0,000476
34	0,000573
35	0,000504
36	0,000488
37	0,000633
38	0,000687
39	0,000624
40	0,000644
41	0,000747
42	0,000884
43	0,000998
44	0,001103
45	0,001222
46	0,001328
47	0,001395
48	0,001495
49	0,001593
50	0,001910
51	0,002064
52	0,002110
53	0,002289
54	0,002733
55	0,002779
56	0,003515
57	0,003490
58	0,003412
59	0,003390
60	0,003922
61	0,004358
62	0,005057
63	0,005520
64	0,006762

ΔΙΑΓΡΑΜΜΑΤΑ ΕΞΟΜΑΛΥΜΕΝΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 5 - 64

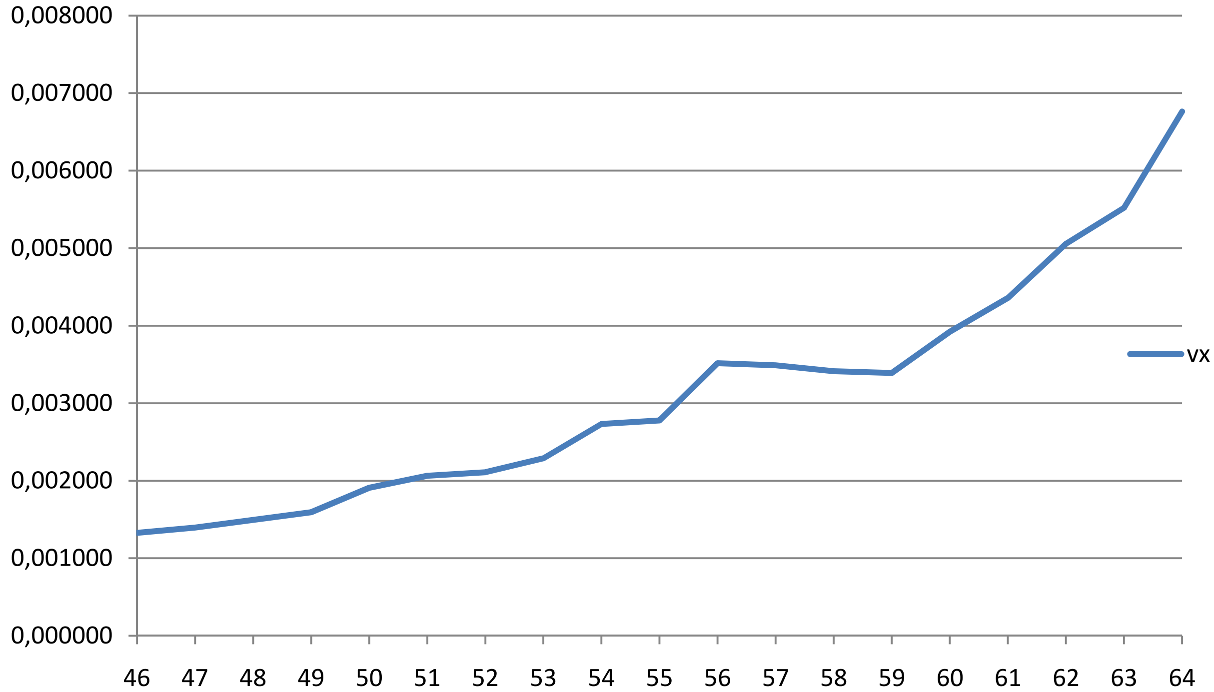




ΓΥΝΑΙΚΕΣ 26 - 45, ΕΞΟΜΑΛΥΜΕΝΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ



ΓΥΝΑΙΚΕΣ 46 - 64, ΕΞΟΜΑΛΥΜΕΝΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ



6.4 ΜΕΤΡΑ ΕΞΟΜΑΛΥΝΣΗΣ & ΣΤΑΤΙΣΤΙΚΟΙ ΕΛΕΓΧΟΙ

6.4.1 ΜΕΤΡΑ ΜΕΘΟΔΟΥ ΚΙΝΟΥΜΕΝΟΥ ΣΤΑΘΜΙΣΜΕΝΟΥ ΜΕΣΟΥ

Και για τα δύο φύλα εξετάζονται τα *μέτρα* που μπορούν να εξαχθούν από τη μέθοδο Κινούμενου Σταθμισμένου Μέσου όπως τα :

$$R_0^2 = \sum_{r=-n}^{n} {a_r}^2 ,$$

$$R_3^2 = \frac{\sum_{r=-n-3}^{n} (\Delta^3 a_r)^2}{\binom{6}{3}} ,$$

$$R_3 = \sqrt{R_3^2} ,$$

και

$$Βάρος = \frac{1}{R_0^2} .$$

(όπως αυτά αναλύονται στο 4.1).

Κρίνεται σκόπιμο να δοθούν τώρα οι τιμές των πρώτων τεσσάρων μέτρων, καθώς εξαρτώνται μόνο από τις τιμές των συντελεστών στάθμισης που όπως τονίστηκε είναι κοινοί και για τα δύο φύλα, καθώς χρησιμοποιούμε το ίδιο μοντέλο. Και για τους άνδρες και για τις γυναίκες έχουμε τις εξής τιμές και συμπεράσματα :

$$R_0^2 = \sum_{r=-5}^{5} {a_r}^2 = 0{,}543494 .$$

Επιθυμούμε το παραπάνω μέτρο μικρότερο της μονάδας, οπότε η μέθοδος, με τις συγκεκριμένες παραδοχές, δείχνει να παράγει θετικά αποτελέσματα. Αντίστοιχα, έχουμε :

$$Βάρος = \frac{1}{R_0^2} = 1{,}839947\ ,$$

τιμή η οποία επίσης είναι μέσα στα αποδεκτά όρια, εφόσον επιθυμούμε το R_0^2 μικρότερο της μονάδας. Όσον αφορά το μέτρο R_3^2 το οποίο επιθυμούμε αντίστοιχα μικρότερο της μονάδος και το μέτρο R_3 που καλείται συντελεστής εξομάλυνσης έχουμε :

$$R_3^2 = \frac{\sum_{r=-8}^{5} (\Delta^3 a_r)^2}{\binom{6}{3}} = 0{,}211459$$

και

$$R_3 = \sqrt{R_3^2} = 0{,}459846\ .$$

Επίσης έχουμε :

$$\binom{6}{3} = \frac{6!}{3! * 3!} = 20\ .$$

Και τα δύο αυτά μέτρα μας δείχνουν ότι η εφαρμοσμένη μέθοδος κινείται μέσα στα αποδεκτά πλαίσια ώστε να την θεωρήσουμε κατάλληλη. Παρακάτω παρατίθεται ενδεικτικά ο πίνακας των υπολογισμών για το μέτρο R_3^2.

r	a_r	Δa_r	$\Delta^2 a_r$	$\Delta^3 a_r$	$(\Delta^3 a_r)^2$
-8	0	0,0000000000	0,0000000000	0,0013553937	0,0000018371
-7	0	0,0000000000	0,0013553937	-0,0177085105	0,0003135913
-6	0	0,0013553937	-0,0163531168	0,1063707376	0,0113147338
-5	0,0013553937	-0,0149977230	0,0900176208	-0,3900905459	0,1521706340
-4	-0,0136423293	0,0750198978	-0,3000729251	0,9752332326	0,9510798580
-3	0,0613775685	-0,2250530273	0,6751603075	-0,7554176089	0,5706557639
-2	-0,1636754589	0,4501072802	-0,0802573014	-0,6594426560	0,4348646166
-1	0,2864318213	0,3698499787	-0,7396999575	0,6594426560	0,4348646166
0	0,6562818000	-0,3698499787	-0,0802573014	0,7554176089	0,5706557639
1	0,2864318213	-0,4501072802	0,6751603075	-0,9752332326	0,9510798580
2	-0,1636754589	0,2250530273	-0,3000729251	0,3900905459	0,1521706340
3	0,0613775685	-0,0750198978	0,0900176208		
4	-0,0136423293	0,0149977230			
5	0,0013553937				
					4,229172

6.4.2 ΜΕΤΡΑ ΚΑΛΗΣ ΕΦΑΡΜΟΓΗΣ – ΟΜΑΛΟΤΗΤΑΣ ΕΞΟΜΑΛΥΝΣΗΣ

Επίσης, υπολογίζονται και για τα δύο φύλα τα *μέτρα ομαλότητας* και *καλής εφαρμογής*,

$$S = \sum_x (\Delta^4 v_x)^2$$

και

$$F_2 = \sum_x w_x * (v_x - u_x)^2 \,,$$

(όπως αυτά αναλύονται στο 3.5), για πλήθος 60 συνολικά παρατηρήσεων (ηλικίες 5 – 64). Και για τα δύο φύλα λαμβάνουμε ως βάρη στάθμισης σε κάθε ηλικία το πλήθος των ζώντων ατόμων της ηλικίας n_x. Δηλαδή $w_x = n_x$ για $x = 5, 6, 7, \dots, 63, 64$. Για τους ***άνδρες*** παίρνουμε τις εξής τιμές του μέτρου καλής εφαρμογής και του μέτρου ομαλότητας :

$$F_2 = \sum_x w_x * (v_x - u_x)^2 = 0{,}324386$$

και

$$S = \sum_x (\Delta^4 v_x)^2 = 0{,}000109\,.$$

Αντίστοιχα για τις ***γυναίκες***, τα ανωτέρω μέτρα παίρνουν τις εξής τιμές :

$$F_2 = \sum_x w_x * (v_x - u_x)^2 = 0{,}086770$$

και

$$S = \sum_x (\Delta^4 v_x)^2 = 0{,}000023\,.$$

Μια απλή ερμηνεία των αποτελεσμάτων μας δείχνει ότι η κοινά εφαρμοσμένη μέθοδος και στα δύο φύλα, παράγει καλύτερα αποτελέσματα στην περίπτωση των γυναικών απ’ ότι στους άνδρες, εφόσον και για τα δύο μέτρα επιθυμούμε την ελαχιστοποίησή τους. Βέβαια, σε καμία περίπτωση δεν τίθεται θέμα σύγκρισης, καθώς αναφερόμαστε σε δύο διακριτούς πληθυσμούς. Παρακάτω παρατίθενται ενδεικτικά οι πίνακες για κάθε μέτρο και για τα δύο φύλα. *Επίσης, πρέπει να τονίσουμε ότι λογικά αναμέναμε τιμές κοντά στο μηδέν (όπως και προέκυψαν) αναφορικά με το μέτρο ομαλότητας, καθώς υποθέσαμε πολυώνυμο τρίτου βαθμού για τις εξομαλυμένες μας τιμές (βλ. ανάλυση υπο - παραγράφου 3.5.1).*

6.4.2.1 ΑΝΔΡΕΣ

ΠΙΝΑΚΑΣ ΜΕΤΡΟΥ F_2 ΑΝΔΡΕΣ

x	n_x	u_x	v_x	$w_x*(v_x-u_x)^2$
5	55.761	0,000143	0,000137	0,000002
6	55.725	0,000179	0,000189	0,000005
7	54.870	0,000146	0,000122	0,000030
8	56.891	0,000088	0,000134	0,000122
9	57.237	0,000192	0,000129	0,000231
10	59.932	0,000050	0,000113	0,000235
11	57.779	0,000208	0,000164	0,000110
12	62.068	0,000242	0,000258	0,000016
13	60.711	0,000264	0,000276	0,000010
14	64.914	0,000293	0,000249	0,000121
15	68.704	0,000247	0,000326	0,000422
16	72.468	0,000469	0,000356	0,000922
17	76.185	0,000446	0,000577	0,001300
18	80.490	0,001006	0,000887	0,001149
19	82.653	0,001028	0,001119	0,000676
20	89.081	0,001134	0,001061	0,000473
21	84.142	0,001165	0,001245	0,000546
22	86.698	0,001326	0,001235	0,000722
23	88.473	0,001051	0,001115	0,000361
24	88.624	0,001252	0,001271	0,000029
25	88.134	0,001293	0,001172	0,001299
26	89.891	0,000768	0,000955	0,003171
27	85.414	0,001218	0,001027	0,003114
28	87.165	0,001021	0,001178	0,002156
29	85.595	0,001110	0,000982	0,001409
30	91.474	0,000973	0,001087	0,001182
31	86.384	0,001239	0,001146	0,000749
32	88.084	0,001090	0,001139	0,000209
33	88.807	0,001149	0,001162	0,000016
34	86.794	0,001244	0,001177	0,000392
35	83.240	0,001009	0,001097	0,000650
36	79.812	0,001253	0,001181	0,000408
37	76.917	0,001274	0,001304	0,000069
38	76.494	0,001360	0,001377	0,000022
39	75.863	0,001490	0,001441	0,000176
40	84.320	0,001471	0,001504	0,000094
41	77.342	0,001978	0,002019	0,000130
42	75.683	0,002352	0,002207	0,001580
43	74.707	0,001994	0,002219	0,003774
44	75.592	0,002593	0,002334	0,005067
45	75.939	0,002581	0,002833	0,004810
46	71.939	0,003406	0,003201	0,003010
47	68.531	0,003342	0,003454	0,000869

48	69.329	0,003649	0,003660	0,000008
49	70.397	0,004020	0,003896	0,001077
50	72.828	0,004078	0,004240	0,001918
51	60.939	0,005300	0,005213	0,000467
52	63.089	0,005675	0,005624	0,000164
53	69.689	0,005539	0,005663	0,001076
54	71.545	0,006723	0,006691	0,000075
55	65.107	0,007388	0,007169	0,003117
56	59.852	0,007736	0,008261	0,016528
57	49.295	0,009737	0,008959	0,029899
58	50.236	0,008420	0,009361	0,044435
59	46.605	0,009548	0,008529	0,048385
60	64.104	0,008377	0,009351	0,060807
61	57.148	0,012039	0,011255	0,035076
62	59.114	0,013161	0,013712	0,017954
63	57.720	0,014258	0,013831	0,010572
64	60.095	0,014444	0,014871	0,010993
			F_2	**0,324386**

ΠΙΝΑΚΑΣ ΜΕΤΡΟΥ *S* ΑΝΔΡΕΣ

x	v_x	Δv_x	$\Delta^2 v_x$	$\Delta^3 v_x$	$\Delta^4 v_x$	$(\Delta^4 v_x)^2$
5	0,000137	0,000052	-0,000119	0,000197	-0,000293	0,00000009
6	0,000189	-0,000067	0,000079	-0,000096	0,000103	0,00000001
7	0,000122	0,000012	-0,000017	0,000007	0,000071	0,00000000
8	0,000134	-0,000006	-0,000010	0,000078	-0,000103	0,00000001
9	0,000129	-0,000016	0,000067	-0,000025	-0,000092	0,00000001
10	0,000113	0,000051	0,000042	-0,000117	0,000147	0,00000002
11	0,000164	0,000094	-0,000075	0,000030	0,000118	0,00000001
12	0,000258	0,000019	-0,000045	0,000148	-0,000297	0,00000009
13	0,000276	-0,000027	0,000103	-0,000149	0,000384	0,00000015
14	0,000249	0,000076	-0,000046	0,000236	-0,000336	0,00000011
15	0,000326	0,000031	0,000190	-0,000100	-0,000067	0,00000000
16	0,000356	0,000221	0,000089	-0,000167	-0,000045	0,00000000
17	0,000577	0,000310	-0,000078	-0,000212	0,000744	0,00000055
18	0,000887	0,000232	-0,000290	0,000532	-0,000969	0,00000094
19	0,001119	-0,000058	0,000242	-0,000437	0,000521	0,00000027
20	0,001061	0,000184	-0,000194	0,000084	0,000302	0,00000009
21	0,001245	-0,000010	-0,000110	0,000386	-0,000916	0,00000084
22	0,001235	-0,000120	0,000276	-0,000530	0,000666	0,00000044
23	0,001115	0,000156	-0,000254	0,000136	0,000270	0,00000007
24	0,001271	-0,000098	-0,000118	0,000406	-0,000614	0,00000038
25	0,001172	-0,000217	0,000288	-0,000208	-0,000221	0,00000005
26	0,000955	0,000071	0,000080	-0,000429	0,001079	0,00000116
27	0,001027	0,000152	-0,000348	0,000650	-0,000998	0,00000100

28	0,001178	-0,000197	0,000302	-0,000348	0,000328	0,00000011
29	0,000982	0,000105	-0,000046	-0,000020	0,000116	0,00000001
30	0,001087	0,000059	-0,000066	0,000096	-0,000134	0,00000002
31	0,001146	-0,000007	0,000030	-0,000038	-0,000049	0,00000000
32	0,001139	0,000023	-0,000008	-0,000087	0,000346	0,00000012
33	0,001162	0,000015	-0,000095	0,000259	-0,000383	0,00000015
34	0,001177	-0,000080	0,000164	-0,000125	0,000036	0,00000000
35	0,001097	0,000084	0,000039	-0,000089	0,000131	0,00000002
36	0,001181	0,000123	-0,000050	0,000042	-0,000037	0,00000000
37	0,001304	0,000072	-0,000008	0,000006	0,000449	0,00000020
38	0,001377	0,000065	-0,000002	0,000455	-0,001235	0,00000152
39	0,001441	0,000063	0,000453	-0,000780	0,000931	0,00000087
40	0,001504	0,000515	-0,000327	0,000151	0,000129	0,00000002
41	0,002019	0,000188	-0,000176	0,000279	0,000002	0,00000000
42	0,002207	0,000012	0,000103	0,000281	-0,000795	0,00000063
43	0,002219	0,000115	0,000384	-0,000514	0,000529	0,00000028
44	0,002334	0,000499	-0,000130	0,000015	0,000053	0,00000000
45	0,002833	0,000368	-0,000115	0,000068	0,000010	0,00000000
46	0,003201	0,000253	-0,000047	0,000078	-0,000002	0,00000000
47	0,003454	0,000206	0,000031	0,000077	0,000444	0,00000020
48	0,003660	0,000237	0,000107	0,000521	-0,001711	0,00000293
49	0,003896	0,000344	0,000628	-0,001190	0,001381	0,00000191
50	0,004240	0,000972	-0,000562	0,000191	0,001168	0,00000137
51	0,005213	0,000411	-0,000371	0,001359	-0,002896	0,00000839
52	0,005624	0,000040	0,000988	-0,001537	0,002701	0,00000729
53	0,005663	0,001028	-0,000549	0,001163	-0,002172	0,00000472
54	0,006691	0,000478	0,000614	-0,001009	0,001109	0,00000123
55	0,007169	0,001092	-0,000395	0,000100	-0,001039	0,00000108
56	0,008261	0,000697	-0,000295	-0,000939	0,003825	0,00001463
57	0,008959	0,000402	-0,001234	0,002886	-0,003456	0,00001195
58	0,009361	-0,000831	0,001653	-0,000570	0,000039	0,00000000
59	0,008529	0,000822	0,001083	-0,000531	-0,002359	0,00000557
60	0,009351	0,001905	0,000552	-0,002890	0,006151	0,00003784
61	0,011255	0,002457	-0,002338	0,003261		
62	0,013712	0,000118	0,000923			
63	0,013831	0,001041				
64	0,014871					
					S	**0,000109**

6.4.2.2 ΓΥΝΑΙΚΕΣ

ΠΙΝΑΚΑΣ ΜΕΤΡΟΥ F_2 ΓΥΝΑΙΚΕΣ

x	n_x	u_x	v_x	$w_x*(v_x-u_x)^2$
5	52.139	0,000173	0,000187	0,000010
6	52.417	0,000172	0,000157	0,000011
7	52.290	0,000115	0,000142	0,000038
8	53.450	0,000131	0,000093	0,000077
9	54.325	0,000018	0,000058	0,000085
10	55.400	0,000108	0,000076	0,000057
11	52.906	0,000113	0,000138	0,000032
12	56.940	0,000123	0,000097	0,000038
13	56.121	0,000107	0,000143	0,000074
14	59.624	0,000201	0,000153	0,000137
15	62.544	0,000128	0,000180	0,000167
16	65.884	0,000258	0,000213	0,000136
17	69.194	0,000231	0,000266	0,000085
18	73.459	0,000272	0,000243	0,000062
19	74.593	0,000282	0,000312	0,000068
20	81.066	0,000358	0,000326	0,000083
21	76.642	0,000274	0,000302	0,000062
22	79.065	0,000304	0,000281	0,000042
23	80.763	0,000310	0,000334	0,000048
24	80.909	0,000284	0,000252	0,000084
25	82.151	0,000134	0,000169	0,000099
26	84.552	0,000237	0,000212	0,000050
27	80.740	0,000310	0,000314	0,000002
28	82.627	0,000375	0,000391	0,000021
29	81.158	0,000431	0,000402	0,000070
30	86.271	0,000313	0,000342	0,000074
31	83.364	0,000372	0,000357	0,000019
32	86.003	0,000419	0,000413	0,000003
33	87.550	0,000457	0,000476	0,000032
34	85.201	0,000587	0,000573	0,000015
35	81.521	0,000515	0,000504	0,000010
36	79.549	0,000453	0,000488	0,000099
37	77.161	0,000674	0,000633	0,000129
38	77.433	0,000659	0,000687	0,000063
39	75.423	0,000636	0,000624	0,000013
40	85.632	0,000642	0,000644	0,000000
41	78.373	0,000740	0,000747	0,000003
42	76.956	0,000897	0,000884	0,000012
43	76.161	0,000985	0,000998	0,000014
44	77.177	0,001114	0,001103	0,000010
45	76.528	0,001215	0,001222	0,000004
46	72.972	0,001329	0,001328	0,000000
47	68.355	0,001404	0,001395	0,000006

48	70.778	0,001469	0,001495	0,000046
49	69.207	0,001647	0,001593	0,000203
50	74.333	0,001816	0,001910	0,000656
51	62.108	0,002190	0,002064	0,000985
52	66.537	0,001984	0,002110	0,001054
53	71.986	0,002389	0,002289	0,000723
54	74.295	0,002638	0,002733	0,000663
55	67.651	0,002942	0,002779	0,001796
56	63.731	0,003201	0,003515	0,006301
57	51.535	0,003978	0,003490	0,012287
58	54.640	0,002818	0,003412	0,019278
59	51.563	0,003976	0,003390	0,017659
60	74.205	0,003450	0,003922	0,016543
61	63.207	0,004651	0,004358	0,005432
62	68.431	0,004954	0,005057	0,000729
63	65.516	0,005495	0,005520	0,000042
64	70.534	0,006819	0,006762	0,000230
			F_2	**0,086770**

ΠΙΝΑΚΑΣ ΜΕΤΡΟΥ *S* ΓΥΝΑΙΚΕΣ

x	**v_x**	**Δv_x**	**$\Delta^2 v_x$**	**$\Delta^3 v_x$**	**$\Delta^4 v_x$**	**$(\Delta^4 v_x)^2$**
5	0,000187	-0,000029	0,000014	-0,000046	0,000093	0,00000001
6	0,000157	-0,000016	-0,000033	0,000047	-0,000007	0,00000000
7	0,000142	-0,000049	0,000014	0,000040	-0,000050	0,00000000
8	0,000093	-0,000035	0,000053	-0,000010	-0,000136	0,00000002
9	0,000058	0,000018	0,000043	-0,000146	0,000334	0,00000011
10	0,000076	0,000062	-0,000102	0,000189	-0,000311	0,00000010
11	0,000138	-0,000041	0,000086	-0,000122	0,000174	0,00000003
12	0,000097	0,000046	-0,000036	0,000052	-0,000061	0,00000000
13	0,000143	0,000010	0,000016	-0,000010	0,000024	0,00000000
14	0,000153	0,000026	0,000007	0,000014	-0,000112	0,00000001
15	0,000180	0,000033	0,000021	-0,000098	0,000267	0,00000007
16	0,000213	0,000054	-0,000077	0,000169	-0,000315	0,00000010
17	0,000266	-0,000023	0,000092	-0,000146	0,000163	0,00000003
18	0,000243	0,000069	-0,000054	0,000017	0,000022	0,00000000
19	0,000312	0,000014	-0,000038	0,000039	0,000034	0,00000000
20	0,000326	-0,000023	0,000002	0,000074	-0,000284	0,00000008
21	0,000302	-0,000022	0,000075	-0,000211	0,000345	0,00000012
22	0,000281	0,000053	-0,000136	0,000134	-0,000006	0,00000000
23	0,000334	-0,000082	-0,000001	0,000128	-0,000197	0,00000004
24	0,000252	-0,000083	0,000127	-0,000069	-0,000015	0,00000000
25	0,000169	0,000044	0,000058	-0,000084	0,000043	0,00000000
26	0,000212	0,000102	-0,000025	-0,000041	0,000036	0,00000000
27	0,000314	0,000077	-0,000066	-0,000005	0,000150	0,00000002
28	0,000391	0,000011	-0,000071	0,000145	-0,000178	0,00000003

29	0,000402	-0,000060	0,000074	-0,000033	-0,000001	0,00000000
30	0,000342	0,000015	0,000041	-0,000034	0,000062	0,00000000
31	0,000357	0,000056	0,000007	0,000027	-0,000228	0,00000005
32	0,000413	0,000063	0,000034	-0,000201	0,000420	0,00000018
33	0,000476	0,000097	-0,000166	0,000219	-0,000110	0,00000001
34	0,000573	-0,000069	0,000052	0,000109	-0,000362	0,00000013
35	0,000504	-0,000017	0,000162	-0,000253	0,000226	0,00000005
36	0,000488	0,000145	-0,000091	-0,000027	0,000229	0,00000005
37	0,000633	0,000054	-0,000118	0,000202	-0,000204	0,00000004
38	0,000687	-0,000064	0,000084	-0,000002	-0,000045	0,00000000
39	0,000624	0,000020	0,000082	-0,000047	-0,000012	0,00000000
40	0,000644	0,000103	0,000035	-0,000059	0,000073	0,00000001
41	0,000747	0,000138	-0,000024	0,000014	0,000009	0,00000000
42	0,000884	0,000114	-0,000009	0,000024	-0,000052	0,00000000
43	0,000998	0,000105	0,000015	-0,000028	0,000002	0,00000000
44	0,001103	0,000119	-0,000013	-0,000026	0,000097	0,00000001
45	0,001222	0,000106	-0,000039	0,000072	-0,000106	0,00000001
46	0,001328	0,000067	0,000033	-0,000034	0,000254	0,00000006
47	0,001395	0,000100	-0,000001	0,000220	-0,000602	0,00000036
48	0,001495	0,000098	0,000219	-0,000382	0,000438	0,00000019
49	0,001593	0,000317	-0,000163	0,000055	0,000186	0,00000003
50	0,001910	0,000154	-0,000108	0,000241	-0,000110	0,00000001
51	0,002064	0,000046	0,000133	0,000131	-0,000792	0,00000063
52	0,002110	0,000179	0,000264	-0,000662	0,001750	0,00000306
53	0,002289	0,000444	-0,000398	0,001088	-0,002542	0,00000646
54	0,002733	0,000046	0,000691	-0,001453	0,002165	0,00000469
55	0,002779	0,000737	-0,000763	0,000711	-0,000605	0,00000037
56	0,003515	-0,000026	-0,000051	0,000107	0,000392	0,00000015
57	0,003490	-0,000077	0,000055	0,000498	-0,001147	0,00000132
58	0,003412	-0,000022	0,000553	-0,000649	0,001007	0,00000101
59	0,003390	0,000532	-0,000095	0,000358	-0,000857	0,00000073
60	0,003922	0,000436	0,000263	-0,000499	0,001514	0,00000229
61	0,004358	0,000699	-0,000236	0,001015		
62	0,005057	0,000463	0,000779			
63	0,005520	0,001242				
64	0,006762					
					S	**0,000023**

6.4.3 ΣΤΑΤΙΣΤΙΚΟΙ ΕΛΕΓΧΟΙ ΜΕΘΟΔΟΥ ΕΞΟΜΑΛΥΝΣΗΣ

Τέλος, για αμφότερα τα φύλα πραγματοποιούνται δύο στατιστικοί έλεγχοι, αναφορικά με την καταλληλότητα της εφαρμοσμένης μεθόδου εξομάλυνσης. Πιο συγκεκριμένα πραγματοποιούμε ένα X^2 *test και* έναν *Έλεγχο προσήμων (Signs test)* (όπως αυτοί αναλύθηκαν στο 3.7), πάντα για επίπεδο σημαντικότητας $a = 5\%$.

6.4.3.1 ΑΝΔΡΕΣ

Για τους άνδρες και το $\boldsymbol{X^2}$ ***test*** σχηματίζουμε την εξής ποσότητα:

$$Z_x = \frac{h_x - n_x * v_x}{\sqrt{n_x * v_x * (1 - v_x)}},$$

όπου $h_x = d_x$. Υπολογίζουμε το στατιστικό :

$$X^2 = \sum_{x=1}^{n} {Z_x}^2,$$

για τις 60 τιμές που εξομαλύναμε στο εύρος των ηλικιών 5 – 64. *Για 60 βαθμούς ελευθερίας και έχουμε επίπεδο σημαντικότητας* $a = 5\%$*, βρίσκουμε από τους πίνακες της* X^2 *κατανομής την τιμή 79,1.* Ο πίνακας των ανδρών είναι ο εξής :

ΠΙΝΑΚΑΣ X^2 TEST ΑΝΔΡΕΣ

x	n_x*v_x	h_x	$(h_x-n_x*v_x)^2$	var (H_x)	Z_x^2
5	7,663404	8	0,113297	7,662350	0,014786
6	10,541861	10	0,293613	10,539867	0,027857
7	6,716937	8	1,646250	6,716115	0,245119
8	7,638169	5	6,959938	7,637144	0,911327
9	7,362098	11	13,234328	7,361151	1,797861
10	6,751925	3	14,076944	6,751165	2,085113
11	9,484081	12	6,329847	9,482525	0,667528
12	15,992776	15	0,985605	15,988655	0,061644
13	16,766384	16	0,587345	16,761754	0,035041
14	16,195140	19	7,867237	16,191100	0,485899
15	22,381870	17	28,964522	22,374578	1,294528
16	25,826517	34	66,805826	25,817313	2,587637
17	43,950546	34	99,013363	43,925191	2,254136

18	71,381285	81	92,519673	71,317982	1,297284
19	92,474775	85	55,872261	92,371311	0,604866
20	94,511027	101	42,106776	94,410755	0,445996
21	104,779584	98	45,962760	104,649105	0,439208
22	107,089846	115	62,570541	106,957568	0,585003
23	98,648736	93	31,908220	98,538741	0,323814
24	112,600542	111	2,561735	112,457478	0,022780
25	103,299435	114	114,502087	103,178361	1,109749
26	85,883672	69	285,058372	85,801617	3,322296
27	87,692055	104	265,949055	87,602025	3,035878
28	102,708168	89	187,913862	102,587145	1,831749
29	84,018906	95	120,584419	83,936434	1,436616
30	99,397787	89	108,113976	99,289779	1,088873
31	98,954488	107	64,730261	98,841134	0,654892
32	100,286703	96	18,375820	100,172523	0,183442
33	103,174367	102	1,379139	103,054501	0,013383
34	102,167532	108	34,017680	102,047268	0,333352
35	91,353924	84	54,080205	91,253666	0,592636
36	94,292079	100	32,580367	94,180679	0,345935
37	100,311655	98	5,343747	100,180833	0,053341
38	105,304791	104	1,702480	105,159824	0,016189
39	109,345683	113	13,354032	109,188077	0,122303
40	126,811417	124	7,904063	126,620701	0,062423
41	156,171368	153	10,057574	155,856022	0,064531
42	167,064306	178	119,589398	166,695525	0,717412
43	165,790320	149	281,914843	165,422397	1,704212
44	176,429876	196	382,989748	176,018093	2,175854
45	215,112370	196	365,282688	214,503021	1,702926
46	230,284308	245	216,551595	229,547144	0,943386
47	236,715054	229	59,522056	235,897409	0,252322
48	253,730529	253	0,533673	252,801926	0,002111
49	274,291635	283	75,835626	273,222897	0,277560
50	308,817807	297	139,660561	307,508305	0,454168
51	317,662511	323	28,488784	316,006602	0,090152
52	354,781989	358	10,355595	352,786867	0,029354
53	394,660247	386	74,999870	392,425221	0,191119
54	478,690776	481	5,332517	475,487968	0,011215
55	466,754544	481	202,933018	463,408363	0,437914
56	494,452400	463	989,253461	490,367605	2,017371
57	441,609273	480	1.473,847921	437,653116	3,367617
58	470,246658	423	2.232,246731	465,844797	4,791825
59	397,513264	445	2.254,990051	394,122710	5,721543
60	599,433526	537	3.897,945219	593,828251	6,564095
61	643,228232	688	2.004,511181	635,988389	3,151805
62	810,578289	778	1.061,344925	799,463542	1,327571
63	798,297446	823	610,216191	787,256579	0,775117
64	893,702460	868	660,616474	880,411769	0,750349

				X^2	**67,912015**

Άρα, εφόσον $X^2 \prec 79{,}1$, *δεν απορρίπτουμε τη μηδενική μας υπόθεση περί καταλληλότητας της εφαρμοσμένης μεθόδου εξομάλυνσης.*

Αντίστοιχα με τον ***Έλεγχο προσήμων***, *σημειώνουμε το πλήθος των θετικών διαφορών της διαφοράς* $h_x - n_x * v_x$, *το οποίο για τους άνδρες είναι* $n_p = 28$. *Επειδή οι συνολικές μας παρατηρήσεις είναι* $n = 60 \succ 20$, *υπολογίζουμε την ποσότητα :*

$$X = \frac{n_p - \frac{n}{2}}{\frac{\sqrt{n}}{2}},$$

παίρνοντας $\boldsymbol{X = -0,516398}$. Παρατηρούμε ότι για επίπεδο σημαντικότητας $a\ = 5\%$,

$$-1{,}96 \prec X = -0{,}516398\ \prec 1{,}96\,,$$

άρα δεν απορρίπτουμε τη μηδενική μας υπόθεση, περί καταλληλότητας της εφαρμοσμένης μεθόδου εξομάλυνσης.

6.4.3.2 ΓΥΝΑΙΚΕΣ

Λειτουργώντας αντίστοιχα για τις γυναίκες, λαμβάνουμε τον εξής πίνακα από τα δεδομένα μας για το $\boldsymbol{X^2}$ ***test*** :

ΠΙΝΑΚΑΣ X^2 TEST ΓΥΝΑΙΚΕΣ

x	n_x*v_x	h_x	$(h_x-n_x*v_x)^2$	var (H_x)	Z_x^2
5	9,734010	9	0,538771	9,732193	0,055360
6	8,252513	9	0,558737	8,251214	0,067716
7	7,409986	6	1,988061	7,408936	0,268333
8	4,971289	7	4,115669	4,970826	0,827965
9	3,147941	1	4,613653	3,147759	1,465694
10	4,221518	6	3,162998	4,221196	0,749313
11	7,291184	6	1,667157	7,290180	0,228685
12	5,538123	7	2,137084	5,537584	0,385924
13	8,033852	6	4,136552	8,032701	0,514964

14	9,143809	12	8,157824	9,142407	0,892306
15	11,236297	8	10,473617	11,234278	0,932291
16	14,001361	17	8,991839	13,998385	0,642348
17	18,427944	16	5,894914	18,423037	0,319975
18	17,858752	20	4,584941	17,854411	0,256796
19	23,248821	21	5,057194	23,241574	0,217593
20	26,411120	29	6,702300	26,402515	0,253851
21	23,171706	21	4,716307	23,164700	0,203599
22	22,183222	24	3,300682	22,176998	0,148834
23	26,978445	25	3,914245	26,969433	0,145136
24	20,386570	23	6,830017	20,381433	0,335110
25	13,855092	11	8,151551	13,852755	0,588443
26	17,948521	20	4,208565	17,944711	0,234530
27	25,373517	25	0,139515	25,365543	0,005500
28	32,306409	31	1,706704	32,293777	0,052849
29	32,619065	35	5,668852	32,605955	0,173859
30	29,522559	27	6,363303	29,512456	0,215614
31	29,756769	31	1,545624	29,746147	0,051960
32	35,514665	36	0,235551	35,499999	0,006635
33	41,669514	40	2,787278	41,649682	0,066922
34	48,851611	50	1,318798	48,823601	0,027011
35	41,119967	42	0,774459	41,099225	0,018844
36	38,810860	36	7,900934	38,791925	0,203675
37	48,839895	52	9,986261	48,808982	0,204599
38	53,206665	51	4,869370	53,170105	0,091581
39	47,026891	48	0,946941	46,997569	0,020149
40	55,146987	55	0,021605	55,111472	0,000392
41	58,512399	58	0,262553	58,468714	0,004490
42	68,052692	69	0,897393	67,992512	0,013198
43	76,030969	75	1,062897	75,955068	0,013994
44	85,129670	86	0,757475	85,035768	0,008908
45	93,547601	93	0,299867	93,433249	0,003209
46	96,932857	97	0,004508	96,804095	0,000047
47	95,373977	96	0,391905	95,240904	0,004115
48	105,805314	104	3,259160	105,647147	0,030849
49	110,254816	114	14,026405	110,079167	0,127421
50	141,982598	135	48,756672	141,711398	0,344056
51	128,178265	136	61,179537	127,913731	0,478287
52	140,375773	132	70,153579	140,079617	0,500812
53	164,784303	172	52,066283	164,407093	0,316691
54	203,020333	196	49,285070	202,465554	0,243424
55	187,976285	199	121,522297	187,453970	0,648278
56	224,039024	204	401,562475	223,251440	1,798700
57	179,835914	205	633,231235	179,208361	3,533492
58	186,455062	154	1.053,331049	185,818798	5,668593
59	174,824338	205	910,570553	174,231597	5,226208
60	291,036949	256	1.227,587828	289,895483	4,234588
61	275,470963	294	343,325221	274,270396	1,251776

62	346,062847	339	49,883810	344,312771	0,144879
63	361,650096	360	2,722818	359,653778	0,007571
64	476,968840	481	16,250247	473,743456	0,034302
				X^2	**35,512245**

Και στην περίπτωση των γυναικών παρατηρούμε ότι $X^2 \prec 79{,}1$, *επομένως δεν απορρίπτουμε τη μηδενική μας υπόθεση, περί καταλληλότητας της εφαρμοσμένης μεθόδου εξομάλυνσης.*

Στον ***Έλεγχο προσήμων*** των γυναικών αντίστοιχα, παρατηρούμε $n_p = 30$ θετικά πρόσημα της διαφοράς $h_x - n_x * v_x$. Άρα,

$$X = \frac{n_p - \frac{n}{2}}{\frac{\sqrt{n}}{2}} = 0\,,$$

και

$$-1{,}96 \prec \boldsymbol{X = 0} \prec 1{,}96\,,$$

Άρα και σ' αυτήν την περίπτωση δεν απορρίπτουμε τη μηδενική μας υπόθεση, περί καταλληλότητας της εφαρμοσμένης μεθόδου εξομάλυνσης.

6.5 ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΟΜΑΛΥΝΣΗΣ

6.5.1 ΑΝΔΡΕΣ

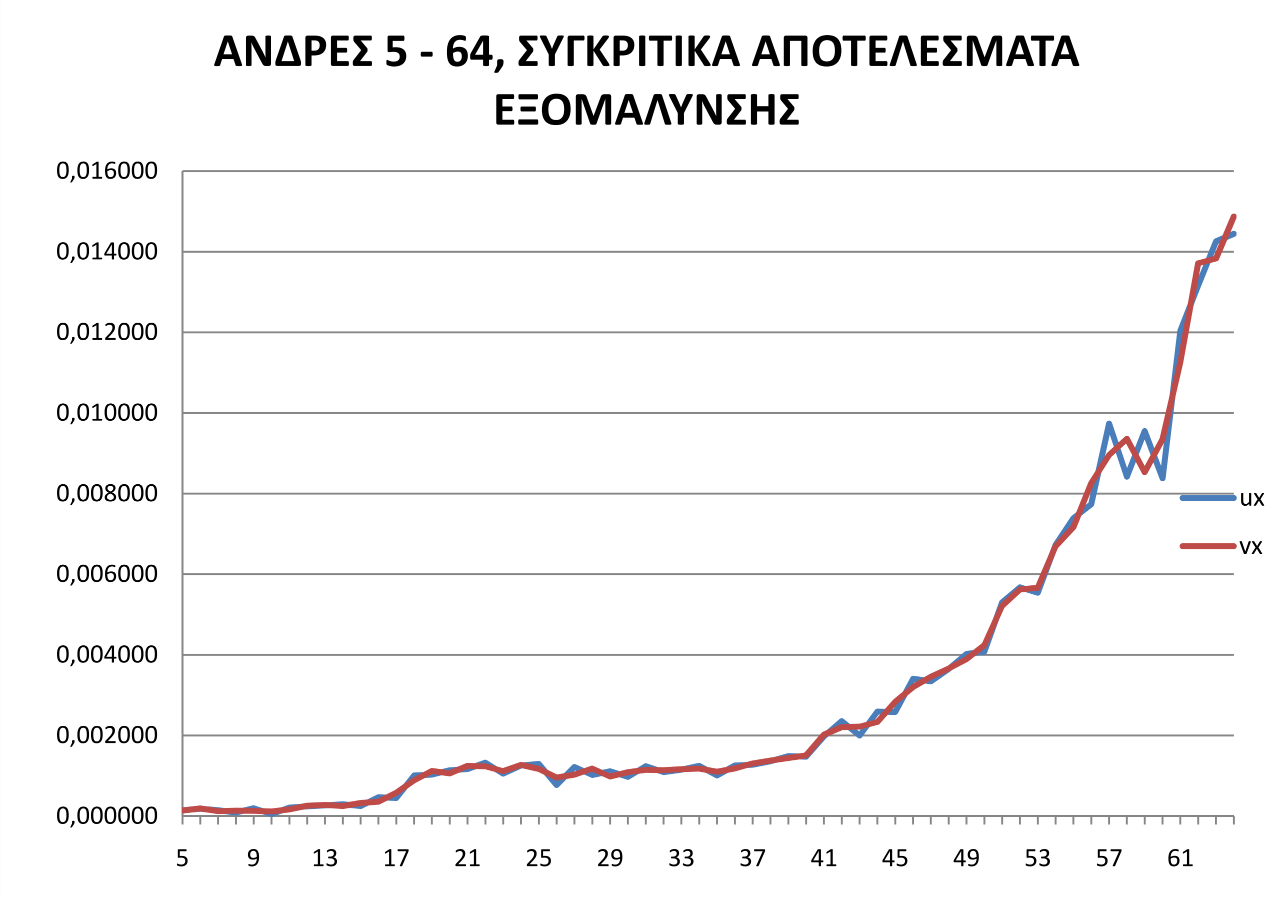
Ακολούθως παρατίθενται τα συγκριτικά αποτελέσματα των πρωτογενών και των εξομαλυμένων τιμών θνησιμότητας για τους άνδρες, ενώ παρουσιάζονται επίσης με σχολιασμό και τα συγκριτικά διαγράμματα όλου του εύρους των ηλικιών που εξομαλύνθηκαν, καθώς και των τριών υπο – ομάδων που παρουσιάστηκαν παραπάνω. Αντίστοιχη προσέγγιση ακολουθείται στην υπο – παράγραφο που ακολουθεί και αφορά τις γυναίκες. Ο πίνακας των συγκριτικών αποτελεσμάτων εξομαλυμένων και πρωτογενών τιμών έχει ως ακολούθως :

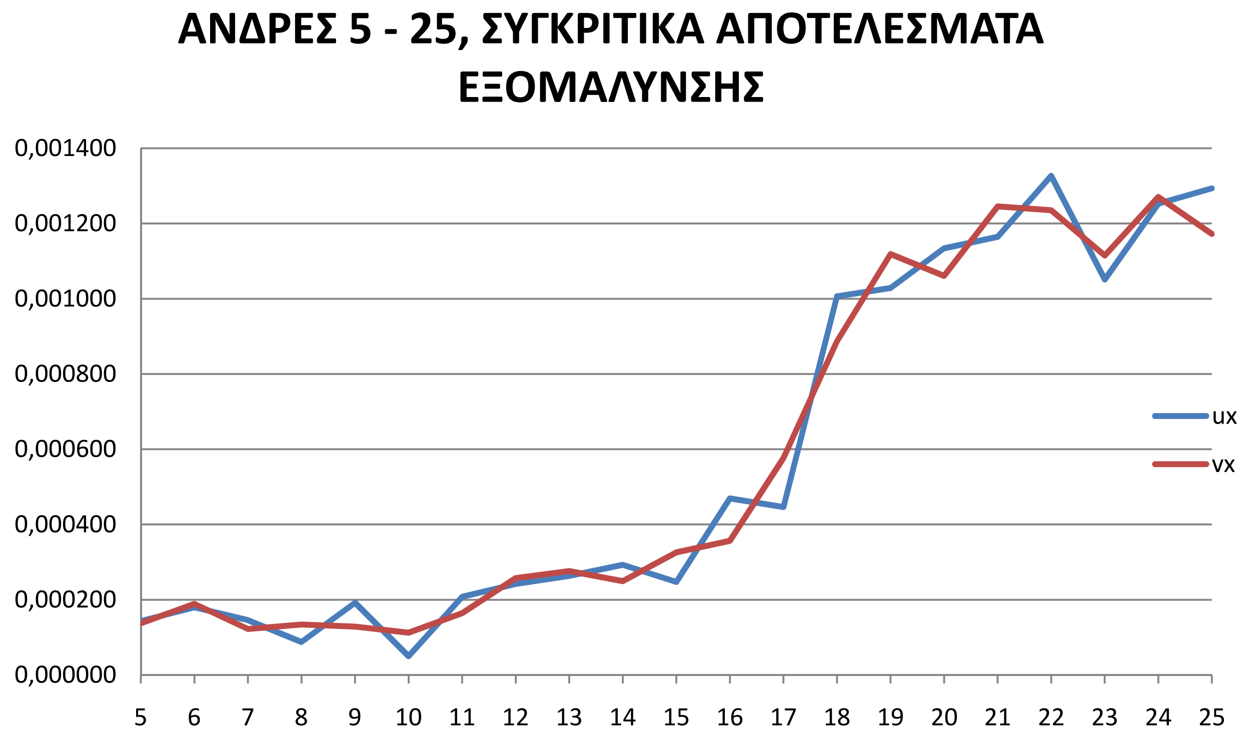
ΠΙΝΑΚΑΣ ΣΥΓΚΡΙΤΙΚΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 5 - 64

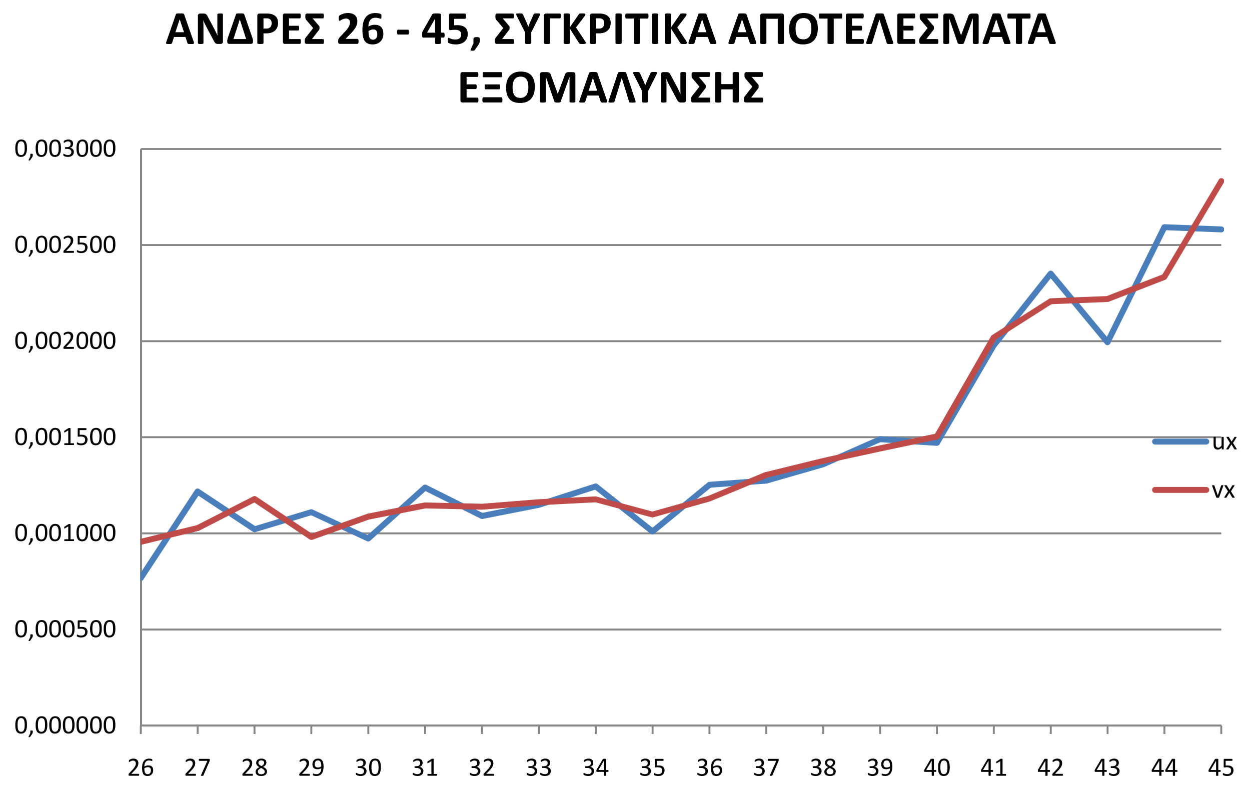
x	u_x	v_x
5	0,000143	0,000137
6	0,000179	0,000189
7	0,000146	0,000122
8	0,000088	0,000134
9	0,000192	0,000129
10	0,000050	0,000113
11	0,000208	0,000164
12	0,000242	0,000258
13	0,000264	0,000276
14	0,000293	0,000249
15	0,000247	0,000326
16	0,000469	0,000356
17	0,000446	0,000577
18	0,001006	0,000887
19	0,001028	0,001119
20	0,001134	0,001061
21	0,001165	0,001245
22	0,001326	0,001235
23	0,001051	0,001115
24	0,001252	0,001271
25	0,001293	0,001172
26	0,000768	0,000955
27	0,001218	0,001027
28	0,001021	0,001178
29	0,001110	0,000982
30	0,000973	0,001087
31	0,001239	0,001146
32	0,001090	0,001139
33	0,001149	0,001162
34	0,001244	0,001177
35	0,001009	0,001097
36	0,001253	0,001181
37	0,001274	0,001304
38	0,001360	0,001377
39	0,001490	0,001441
40	0,001471	0,001504
41	0,001978	0,002019
42	0,002352	0,002207
43	0,001994	0,002219
44	0,002593	0,002334
45	0,002581	0,002833
46	0,003406	0,003201
47	0,003342	0,003454
48	0,003649	0,003660
49	0,004020	0,003896
50	0,004078	0,004240

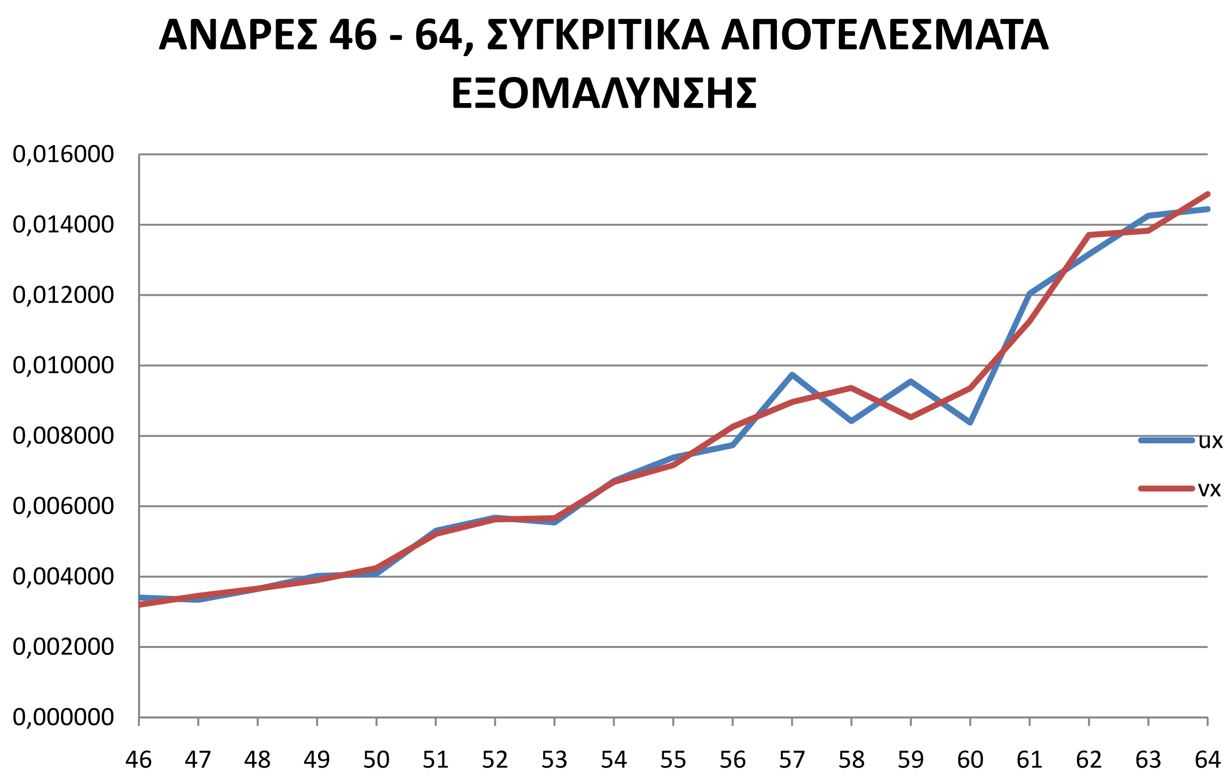
51	0,005300	0,005213
52	0,005675	0,005624
53	0,005539	0,005663
54	0,006723	0,006691
55	0,007388	0,007169
56	0,007736	0,008261
57	0,009737	0,008959
58	0,008420	0,009361
59	0,009548	0,008529
60	0,008377	0,009351
61	0,012039	0,011255
62	0,013161	0,013712
63	0,014258	0,013831
64	0,014444	0,014871

ΔΙΑΓΡΑΜΜΑΤΑ ΣΥΓΚΡΙΤΙΚΩΝ ΤΙΜΩΝ ΑΝΔΡΩΝ 5 - 64




ΑΝΔΡΕΣ 5 - 25, ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΟΜΑΛΥΝΣΗΣ
0,001400
0,001200
0,001000
0,000800
0,000600
0,000400
0,000200
0,000000
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
ux
vx


ΑΝΔΡΕΣ 26 - 45, ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΟΜΑΛΥΝΣΗΣ
0,003000
0,002500
0,002000
0,001500
0,001000
0,000500
0,000000
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
ux
vx



<u>ΣΧΟΛΙΑΣΜΟΣ</u>

Από τα διαγράμματα που προηγήθηκαν, παρατηρούμε ότι η καμπύλη των εξομαλυμένων τιμών είναι πιο ομαλή σχεδόν σε όλο το φάσμα των ηλικιών. Ωστόσο, παράγει σε σημεία ασυνέπειες και ‘‘άλματα’’, όπως π.χ. στις ηλικίες 19, 20 και 59. Ειδικότερα παρατηρούμε, την εξομάλυνση και την ομαλότητα που επιτυγχάνεται μεταξύ των ηλικιών 30 και 57, όπου η εφαρμοσμένη μέθοδος δείχνει να επιτυγχάνει θετικά αποτελέσματα. Τέλος, παρατηρείται ακόμη και στις εξομαλυμένες τιμές, ασυνέπεια στο εύρος ηλικιών 18 – 27 (accident hump), όπου υπάρχει αξιοσημείωτη επιρροή των θνησιμότητας στη χώρα μας, ιδιαίτερα από τα τροχαία ατυχήματα κ.λπ..

6.5.3 ΓΥΝΑΙΚΕΣ

Όμοια με τους άνδρες, παίρνουμε για τις γυναίκες τα εξής :

ΠΙΝΑΚΑΣ ΣΥΓΚΡΙΤΙΚΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 5 - 64

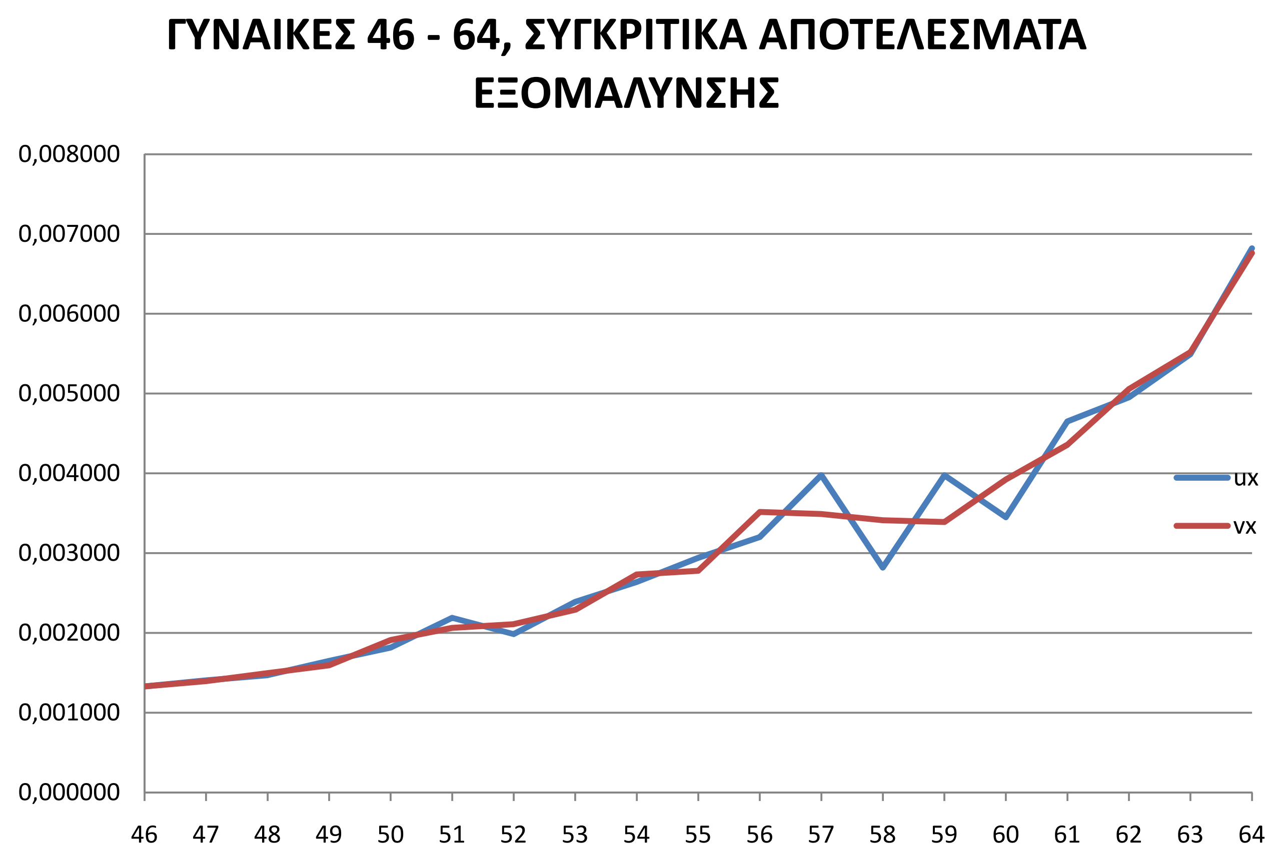
x	u_x	v_x
5	0,000173	0,000187
6	0,000172	0,000157
7	0,000115	0,000142
8	0,000131	0,000093
9	0,000018	0,000058
10	0,000108	0,000076
11	0,000113	0,000138
12	0,000123	0,000097
13	0,000107	0,000143
14	0,000201	0,000153
15	0,000128	0,000180
16	0,000258	0,000213
17	0,000231	0,000266
18	0,000272	0,000243
19	0,000282	0,000312
20	0,000358	0,000326
21	0,000274	0,000302
22	0,000304	0,000281
23	0,000310	0,000334
24	0,000284	0,000252
25	0,000134	0,000169
26	0,000237	0,000212
27	0,000310	0,000314
28	0,000375	0,000391
29	0,000431	0,000402
30	0,000313	0,000342
31	0,000372	0,000357
32	0,000419	0,000413
33	0,000457	0,000476
34	0,000587	0,000573
35	0,000515	0,000504
36	0,000453	0,000488
37	0,000674	0,000633
38	0,000659	0,000687
39	0,000636	0,000624
40	0,000642	0,000644
41	0,000740	0,000747
42	0,000897	0,000884
43	0,000985	0,000998
44	0,001114	0,001103
45	0,001215	0,001222

46	0,001329	0,001328
47	0,001404	0,001395
48	0,001469	0,001495
49	0,001647	0,001593
50	0,001816	0,001910
51	0,002190	0,002064
52	0,001984	0,002110
53	0,002389	0,002289
54	0,002638	0,002733
55	0,002942	0,002779
56	0,003201	0,003515
57	0,003978	0,003490
58	0,002818	0,003412
59	0,003976	0,003390
60	0,003450	0,003922
61	0,004651	0,004358
62	0,004954	0,005057
63	0,005495	0,005520
64	0,006819	0,006762

ΔΙΑΓΡΑΜΜΑΤΑ ΣΥΓΚΡΙΤΚΩΝ ΤΙΜΩΝ ΓΥΝΑΙΚΩΝ 5 - 64


ΓΥΝΑΙΚΕΣ 5 - 25, ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΟΜΑΛΥΝΣΗΣ
0,000400
0,000350
0,000300
0,000250
0,000200
0,000150
0,000100
0,000050
0,000000
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
ux
vx


ΓΥΝΑΙΚΕΣ 26 - 45, ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΟΜΑΛΥΝΣΗΣ
0,001400
0,001200
0,001000
0,000800
0,000600
0,000400
0,000200
0,000000
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
ux
vx



<u>ΣΧΟΛΙΑΣΜΟΣ</u>

Τα παραπάνω διαγράμματα, σε συνδυασμό με τα αποτελέσματα του πίνακα, δείχνουν μια πιο ομαλή καμπύλη των εξομαλυμένων τιμών, στην πλειοψηφία των ηλικιών. Ωστόσο, στο διάστημα των ηλικιών 5 – 25, αν και η εφαρμοσμένη μέθοδος επιτυγχάνει να ομαλοποιήσει τα ‘‘άλματα’’ των πρωτογενών τιμών από την ηλικία 5 έως 9 και 12 έως την ηλικία 17, παρατηρούμε εκ νέου ασυνέπειες και των εξομαλυμένων τιμών στις ηλικίες 12, 18, 22 και 25. Στη συνέχεια η καμπύλη των εξομαλυμένων τιμών είναι αρκετά πιο ομαλή έως την τελευταία ηλικία που εξομαλύναμε και παράγει αποτελέσματα αρκετά συναφή με αυτά των πρωτογενών τιμών, επιτυγχάνοντας μάλιστα μια πιο ομαλή διόρθωση των ‘‘αλμάτων’’ των ηλικιών 57 έως 60.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΕΛΛΗΝΙΚΗ

- Μιχαήλ Παπαδάκης, Κλέων Τσίμπος (2004) *Δημογραφική Ανάλυση Αρχές – Μέθοδοι – Υποδείγματα*, Εκδόσεις Αθ. Σταμούλης, Αθήνα
- Παναγιώτης Α. Καραδήμας (1984) *Δημογραφία β΄ έκδοση*, Εκδόσεις Καραμπερόπουλος
- Νίκος Π. Μπλέσιος (1998) *Μαθηματικά Ασφαλίσεων Ζωής*, Εκδοτικές Επιχειρήσεις ‹‹ Το Οικονομικό ››, Κ. & Π. Σμπίλιας Α.Ε.Β.Ε., Αθήνα
- Μάρκου Β. Κούτρα (2004) *Εισαγωγή στις Πιθανότητες Μέρος Ι Β΄ Έκδοση*, Εκδόσεις Αθ. Σταμούλης, Αθήνα
- Μάρκου Β. Κούτρα (2004) *Εισαγωγή στις Πιθανότητες Μέρος ΙΙ*, Εκδόσεις Αθ. Σταμούλης, Αθήνα
- Μ.Α. Γεωργιακόδης, Π.Ν. Γεωργιάδης (2001) *Γραμμική Άλγεβρα*, Εκδόσεις Αθ. Σταμούλης, Αθήνα
- Μανώλη Λουκάκη (1995), *Μαθηματικά Οικονομικών Επιστημών Τόμος Α΄*
- Μανώλη Λουκάκη (1996) *Μαθηματικά Οικονομικών Επιστημών Τόμος Β΄ Β΄ έκδοση*
- Τάκη Παπαϊωάννου, Κοσμά Φερεντίνου (2000) *Μαθηματική Στατιστική*, Εκδόσεις Αθ. Σταμούλης, Αθήνα
- Χρυσούλα Ζαχαροπούλου (1996) *Στατιστική Μέθοδοι – Εφαρμογές*
- Δημητρίου Γ. Καφφέ (1991) *Μαθήματα Ανάλυσης Παλινδρόμησης*, Εκδόσεις Αθ. Σταμούλης, Πειραιάς
- Κ. Α. Αθανασιάδου (1957) *Στατιστική Μέρος Β΄*, Εκδόσεις Παπαζήση

ΞΕΝΗ

- Dick London (1985) *Graduation : The Revision of Estimates*, ACTEX Publications, Winsted, Connecticut

- Dick London (1988) *Survival Models and their Estimation*, ACTEX Publications, Winsted, Connecticut

- Nathan Keyfitz, Hal Caswell (1985) *Applied Mathematical Demography*, Springer Science & Business Media, Inc, New York

- Chin Long Chiang (1978), *Life Table and Mortality Analysis*, World Health Organization, Geneva