ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

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

Η ΣΤΟΧΑΣΤΙΚΗ ΑΠΟΔΟΣΗ ΣΤΑ ΑΣΦΑΛΙΣΤΗΡΙΑ ΖΩΗΣ UNIT LINKED

Πέτρος Ν. Ζεππάτος

Διπλωματική Εργασία

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

Πειραιάς
Σεπτέμβριος 2024

Π Α Ν Ε Π Ι Σ Τ Η Μ Ι Ο Π Ε Ι Ρ Α Ι Ω Σ

Σχολή Χρηματοοικονομικής και Στατιστικής

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

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

Η ΣΤΟΧΑΣΤΙΚΗ ΑΠΟΔΟΣΗ ΣΤΑ ΑΣΦΑΛΙΣΤΗΡΙΑ ΖΩΗΣ UNIT LINKED

Πέτρος Ν. Ζεππάτος

Διπλωματική Εργασία

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

Πειραιάς

Σεπτέμβριος 2024

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ’ αριθμ. …….. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Χατζηκωνσταντινίδης Ευστάθιος Αναπληρωτής Καθηγητής (Επιβλέπων)
- Σεβρόγλου Βασίλειος Αναπληρωτής Καθηγητής
- Τζαβελάς Γεώργιος Αναπληρωτής Καθηγητής

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

UNIVERSITY OF PIRAEUS

School of Finance and Statistics

Department of Statistics and Insurance Science

POSTGRADUATE PROGRAM IN
ACTUARIAL SCIENCE AND RISK
MANAGEMENT

The Stochastic Performance in Unit-Linked Life Insurance Policies

Peter N. Zeppatos

MSc Dissertation

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

Piraeus, Greece

September 2024

Στην μητέρα μου Ελευθερία και

Στα αδέλφια μου Κατερίνα , Άγγελο και Γιώργο

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους μου τους καθηγητές τόσο από το Πανεπιστήμιο Πειραιώς όσο και από το Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών για τις γνώσεις και την αγάπη για την επιστήμη που μου μετέδωσαν όλα αυτά τα χρόνια.

Ιδιαίτερα θα ήθελα να ευχαριστήσω τον κ. Χαντζηκωνσταντινίδη για την προσοχή που μου έδωσε καθώς και την ευκαιρία να εργαστώ σε ένα θέμα που προσωπικά μου κίνησε την περιέργεια ως αποτέλεσμα την ενασχόλησης μου με την ασφαλιστική αγορά.

Ακόμα θα ήθελα να ευχαριστήσω και τους αξιότιμους κ. Σεβρόγλου και κ. Τζαβελά για το πολύτιμο χρόνο τους ως μέλη της επιτροπής.

Τέλος θα ήθελα να ευχαριστήσω την αδελφή μου Κατερίνα για την επί χρόνια στήριξη της σε ακαδημαϊκό και προσωπικό επίπεδο τόσο σε μένα όσο και στη υπόλοιπη οικογένεια μας.

Περιεχόμενα

1 Εισαγωγή

Στο σύγχρονο χρηματοοικονομικό περιβάλλον που ζούμε τα παραδοσιακά ασφαλιστικά προγράμματα ασφαλίσεων ζωής δεν μπορούν να προσαρμοστούν ευκολά στις σύνθετες ανάγκες των καταναλωτών.

Τα *Unite Linked (UL)* είναι ένα από τα προϊόντα της «νέας γενιάς» που τείνουν να αντικαταστήσουν τα παραδοσιακά σε προϊόντα σε αρκετές ευρωπαϊκές χώρες ενώ στη Ελλάδα εμφανίστηκαν μόλις το 1994 και έχουν ήδη κατακτήσει ένα μεγάλο μερίδιο της αγοράς.

Τα UL είναι συγχρόνως **ασφαλιστικά και χρηματοοικονομικά** προϊόντα. Συνδυάζουν με απόλυτα ευκρινή τρόπο ασφαλιστική κάλυψη και την επένδυση – αποταμίευση σε ένα προϊόν. Το βασικό χαρακτηριστικό των Ασφαλιστικών Προγραμμάτων Ζωής Unit Linked όπως άλλωστε υποδηλώνει και ο τίτλος τους, είναι ότι οι παροχές τους συνδέονται με την αξία μεριδίων επενδυτικών κεφαλαίων (ή μονάδες Units) **και την απόδοση των περιουσιακών στοιχείων τους**

Θα εξηγήσουμε συνοπτικά την λειτουργία των UL.

Ο ασφαλισμένος αποφασίζει για το ασφάλιστρο που θα καταβάλει – επενδύσει, σε αντίθεση με τα παραδοσιακά προγράμματα που η ασφαλιστική με βάση της παροχές που έχει ζητήσει ο ασφαλισμένος «επιβάλει» συγκεκριμένο ασφάλιστρο.

Με το ασφάλιστρο αυτό, ή μέρος του ακριβέστερα, αγοράζει μερίδια από ένα fund που είναι συνδεδεμένο με το ασφαλιστήριο, στη *τιμή διάθεσης* .

Το ασφαλιστήριο συμβόλαιο είναι συγκεκριμένης διάρκειας (στη Ελλάδα είναι 10 ή 20 χρόνια) αλλά δίνεται στο πελάτη η δυνατότητα να εξαγοράσει το συμβόλαιο του νωρίτερα αν το επιθυμεί με ενδεχόμενη επιβάρυνση.

Καθώς οι επενδύσεις του fund προχωρούν στο χρόνο οι τιμές των μεριδίων αλλάζουν με βάση την αξία των περιουσιακών στοιχείων .

Ο πελάτης ή όσοι έχει ορίσει σε περίπτωση θανάτου του λαμβάνει/ουν σε χρηματικό πόσο τα μερίδια, με βάση την *τιμή εξαγοράς,* που έχει αγοράσει κατά την περίοδο της επένδυσης του. Αυτό γίνεται σε 3 περιπτώσεις

1. Σε περίπτωση θανάτου του ασφαλισμένου
2. Σε περίπτωση εξαγοράς του συμβολαίου του πριν λήξη
3. Στην λήξη του συμβολαίου

Σε κάθε χρονική στιγμή ο ασφαλισμένος είναι εκτεθειμένος σε δύο κινδύνους

1. Οι αποδόσεις των περιουσιακών στοιχείων του funds να είναι χαμηλές ή και αρνητικές και
2. Ο ίδιος να αποχωρήσει από το ασφαλιστήριο, είτε λόγω θανάτου είτε λόγω εξαγοράς και να μην λάβει τις αποδόσεις του μέλλοντος

Άρα η απόδοση της επένδυσης – αποταμίευσης θα είναι η απόδοση των περιουσιακών στοιχείων του funds δεδομένου ότι ίδιος είναι εν ζωή και δεν έχει εξαγοράσει νωρίτερα το συμβόλαιο του.

Άρα στο έχουμε δύο άγνωστες ποσότητες, δύο τυχαίες μεταβλητές. Η πρώτη είναι η (λογαριθμική) απόδοση του fund σε ένα έτος $\tilde{R}_t$, δηλαδή την απόδοση των μεριδίων, η οποία είναι μια συνεχής τυχαία μεταβλητής που παίρνει τιμές από στο $[-1, +\infty)$. Η δεύτερη είναι οι μήνες μέσα στο έτος που ο ασφαλισμένος παραμένει ενεργός στο πρόγραμμα N δηλαδή είναι εν ζωή και δεν έχει εξαγοράσει το συμβόλαιο του ή ισοδύναμα οι μέρες μέσα στο έτος που επενδύει – αποταμιεύει ο ασφαλισμένος.

Στόχος είναι να υπολογιστεί η πιθανότητα

$$P(\tilde{R}_t \leq r \mid N = n) =$$

$$P(\tilde{\xi}_1 + \tilde{\xi}_2 + \cdots \tilde{\xi}_N \leq r \mid N = n), \qquad n = 0,1, \dots ,12$$

Όπου $\tilde{\xi}_i$ είναι η λογαριθμική απόδοση του fund το μήνα i του έτους t

Για κάθε έτος $t = 1,2, \dots ,20$

2 Χρηματοοικονομικά Μαθηματικά και Θεωρία Χαρτοφυλακίου

2.1 Χρηματοοικονομικά μαθηματικά

2.1.1 Αποτελεσματικό επιτόκιο

Με τον όρο επιτόκιο συνήθως αναφέρεται η αμοιβή που λαμβάνει ένα χρηματοπιστωτικό ίδρυμα για την χορήγηση ενός δανείου από το οφειλέτη που το έλαβε. Υπό ένα διαφορετικό πρίσμα θα το εξετάσουμε εδώ ως η αμοιβή που λαμβάνει ένας επενδυτής για το ρίσκο που λαμβάνει μέσω της επένδυσης.

Στο παρόν κεφάλαιο θα υποθέσουμε ότι το επιτόκιο της επένδυσης είναι σταθερό και ανεξάρτητο από το αρχικό κεφάλαιο και από τις διακυμάνσεις της αγοράς ώστε να εστιάσουμε στο πως επιδρά πάνω στο κεφάλαιο επένδυσης. Ένα τέτοιο επιτόκιο ονομάζεται *ελεύθερο κινδύνου (risk free interest rate)* και θα το συμβολίζουμε με

$$R_f$$

Ας υποθέσουμε ότι την χρονική στιγμή $t_0 = 0$ επενδύονται k χρηματικές μονάδες. Η *συσσωρευμένη αξία (accumulated value)* αυτών μετά από χρόνο $t > 0$ θα την συμβολίζουμε με $A(t)$.

Έτσι ο τόκος που κερδήθηκε στο χρονικό διάστημα $\Delta t = t - (t - 1)$ θα είναι

$$I_t = A(t) - A(t-1)\,, t > 1$$

Με βάση τον ορισμό του τόκου θα ορίσουμε και αριθμητικά το αποτελεσματικό επιτόκιο της περιόδου t ως :

$$R(t) = \frac{A(t) - A(t-1)}{A(t-1)}\,, t > 1 \qquad (1)$$

2.1.2 Μοντέλα ανατοκισμού

Σε κάθε περίοδο επένδυσης επιδρά το επιτόκιο. Ο τρόπος με το οποίο επιδρά διαφέρει ανάλογα με το *μοντέλο ανατοκισμού* . Θα εξετάσουμε δύο δημοφιλή μοντέλα.

1) **Μοντέλο του απλού τόκου:** Σύμφωνα με αυτό το μοντέλο ο ανατοκισμός γίνεται στο αρχικό κεφάλαιο. Με άλλα λόγια το επιτόκιο επιδρά μόνο στην αρχική επένδυση και όχι στη συσσωρευμένο κεφάλαιο. Πιο συγκεκριμένα έχουμε

$$A(t) = A(0)Rt + A(0) = kRt + k = k(Rt + 1)$$

Αν συνδυάσουμε τις εξισώσεις και έχουμε :

$$R(t) = \frac{A(t) - A(t-1)}{A(t)} = \frac{k(Rt+1) - k(R(t-1)+1)}{k(R(t-1)-1)} = \frac{R}{R(t-1)+1}$$
$$\Leftrightarrow R(t) = \frac{R}{R(t-1)+1}$$

Σχόλιο: Σε αυτό το μοντέλο ο τόκος που κερδίζεται κάθε χρόνο είναι σταθερός και ανεξάρτητος του χρόνου. Το επιτόκιο «προσαρμόζεται» βάση της εξίσωσης

2) **Μοντέλου του σύνθετου τόκου:** Σε αντίθεση με το μοντέλο του απλού τόκου, το επιτόκιο επιδρά στο συσσωρεμένο κεφάλαιο της προηγουμένης περιόδου.
Όποτε έχουμε
$$A(t) = A(t-1) + A(t-1)R = A(t-1)(1+R) = A(t-2)(1+R)^2 = \cdots = k(1+R)^t$$

και για το αποτελεσματικό επιτόκιο θα έχουμε
$$R(t) = \frac{A(t) - A(t-1)}{A(t-1)} = \frac{k(1+R)^t - k(1+R)^{t-1}}{k(1+R)^{t-1}} = \frac{(1+R)-1}{1} = R$$
$$\Leftrightarrow R(t) = R$$

Σχόλιο: Συγκρίνοντας τις εξισώσεις για μοντέλο του απλού τόκου και για του σύνθετου για το αποτελεσματικό επιτόκιο έχουμε ότι στη περίπτωση του απλού τόκου το αποτελεσματικό επιτόκιο είναι συνάρτηση του χρόνου ενώ στη περίπτωση του σύνθετου τόκου είναι ανεξάρτητο του t .

Σημείωση: από εδώ και στο εξής αν δεν δηλώνεται διαφορετικά θα θεωρούμε ότι πάντα έχουμε το εν λόγω μοντέλο.

Το αποτελεσματικό επιτόκιο είναι το επιτόκιο που επιδρά στο κεφάλαιο μια φορά το χρόνο στο τέλος του χρόνου. Μια ενδιαφέρουσα περίπτωση επιτοκίου είναι το προεξοφλητικό επιτόκιο. Σε αυτήν την περίπτωση ο τόκος εκταμιεύεται στη αρχή του χρόνου προκαταβολικά.

Το προεξοφλητικό επιτόκιο ορίζεται μαθηματικά ως εξής

$$d(t) = \frac{A(t) - A(t-1)}{A(t)} = \frac{I(t)}{A(t)}$$

Ο ορισμός του αποτελεσματικού προεξοφλητικού επιτοκίου είναι ο λόγος του προεξοφλήματος που «κερδίζεται» στη διάρκεια της επένδυσης προς το πόσο που επενδύθηκε στο τέλος της περιόδου. Η σχέση που ενώνει το αποτελεσματικό επιτόκιο με το προεξοφλητικό είναι

$$\left(1 + R_f\right) = (1-d)^{-1}$$

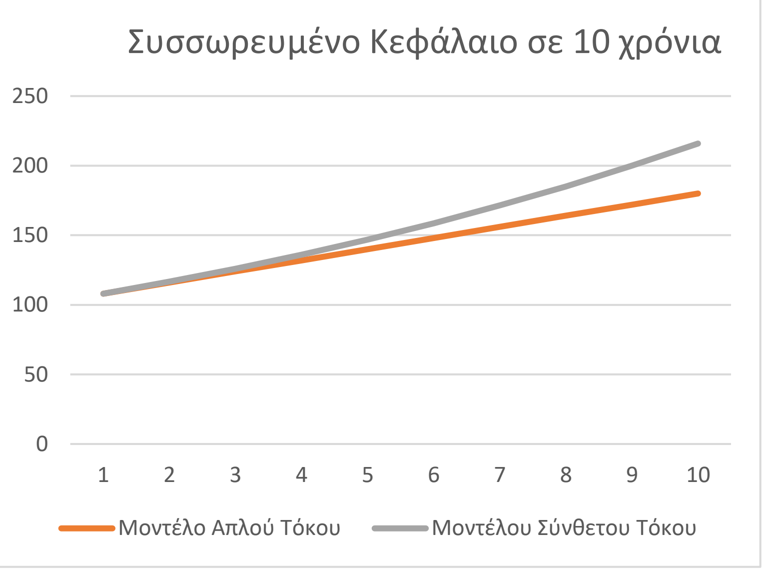
Παράδειγμα 1: Έστω επένδυση 100 νομισματικών μονάδων την χρονική στιγμή $t_0 = 0$ με επιτόκιο χωρίς κίνδυνο $R_f = 8\%$. Να υπολογιστεί η συσσωρευμένη αξία αν η επένδυση διαρκέσει 10 χρόνια.

a) Με την μέθοδο του απλού τόκου και με τη μέθοδο του σύνθετου τόκου

b) Ποια μέθοδος είναι πιο επικερδής για το επενδυτή. Αν η επένδυση διαρκέσει 20 χρόνια ποια θα είναι πιο επικερδής ;

Συμφώνα με τα όσα έχουν ειπωθεί παραπάνω κατασκευάζουμε ένα πίνακα τιμών για να συγκρίνουμε τις δύο μεθόδους.

Χρόνος Επενδυσης	Συσσωρευμένο Κεφάλαιοο	
	Μοντέλο Απλού Τόκου	Μοντέλου Σύνθετου Τόκου
1	108	108
2	116	116.64
3	124	125.9712
4	132	136.048896
5	140	146.9328077
6	148	158.6874323
7	156	171.3824269
8	164	185.093021
9	172	199.9004627
10	180	215.8924997



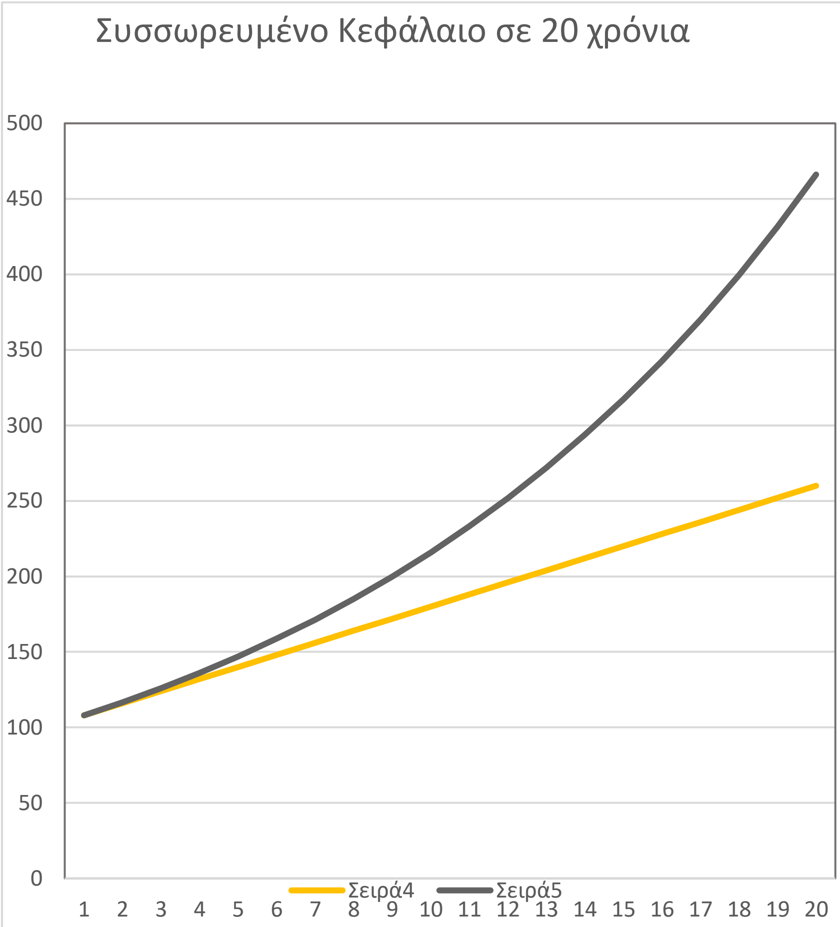
Είναι εμφανές ότι η μέθοδος του σύνθετου τόκου όπου ο ανατοκισμός γίνεται στο συσσωρευμένο της προηγούμενης περιόδου.

Αν επεκτείνουμε το χρόνο της επένδυσης σε 20 χρόνια έχουμε τα εξής αποτελέσματα.

Αναμένουμε τα αποτελέσματα που είχαμε και στα 10 χρόνια μόνο πιο έντονα

Παρατηρούμε ότι ο χρόνος επένδυσης είναι το πιο σημαντικό στοιχείο στη συνάρτηση συσσώρευσης. Ο σύνθετος τόκος αυξάνει εκθετικά την συσσωρευμένη αξία στη πάροδο του χρόνου ενώ το μοντέλο του απλού τόκου γραμμικά.

Χρόνος Επένδυσης	Συσσωρευμένο Κεφάλαιο	
	Μοντέλο Απλού Τόκου	Μοντέλου Σύνθετου Τόκου
1	108	108
2	116	116.64
3	124	125.9712
4	132	136.048896
5	140	146.9328077
6	148	158.6874323
7	156	171.3824269
8	164	185.093021
9	172	199.9004627
10	180	215.8924997
11	188	233.1638997
12	196	251.8170117
13	204	271.9623726
14	212	293.7193624
15	220	317.2169114
16	228	342.5942643
17	236	370.0018055
18	244	399.6019499
19	252	431.5701059
20	260	466.0957144



2.1.3 Περιοδικές Καταβολές

Η επένδυση μπορεί να γίνει με δύο τρόπους. Ο πρώτος είναι με την αγορά ενός χρηματοοικονομικού προϊόντος με εφάπαξ καταβολή είτε σε τμηματικές περιοδικές καταβολές.

Ας υποθέσουμε και πάλι ότι το επιτόκιο είναι σταθερό. Το κεφάλαιο της επένδυσης θα είναι διαιρεμένο το πλήθος των περιοδικών καταβολών από την αρχή μέχρι το τέλος της επένδυσης. Οι τμηματικές καταβολές γνωστές και ως ράντες (*annuities*) έχουν μικρότερο κερδισμένο τόκο στο τέλος της επένδυσης για το ίδιο πόσο σε σχέση με την εφάπαξ καταβολή του πόσου επένδυσης. Αυτό οφείλεται στο ότι οι μεταγενέστερες τις πρώτης καταβολής να έχουν μικρότερο χρονικό ορίζοντα μπροστά τους και άρα λιγότερο κερδισμένο τόκο.

Σε αυτό το σημείο είναι σημείο είναι σημαντικό να διαχωρίσουμε τις δύο βασικές κατηγορίες

1) **Ληξιπρόθεσμες ράντες:** Το όνομα τους το παίρνουν από το χρόνο στο οποίο γίνεται η περιοδική καταβολή μέσα σε μια περίοδο. Έτσι η κατάθεση σε αυτή την περίπτωση γίνεται στη *λήξη* της περιόδου. Αυτό το μοντέλο χρησιμοποιείται π.χ. στη αποτίμηση των μελλοντικών μισθών ενός

εργαζομένου που πληρώνεται στο τέλος του μήνα/έτους ή στη αποτίμηση του κέρδους από ένα δάνειο για την τράπεζα που χορηγήθηκε την χρονική στιγμή t=0 και την πρώτη δόση θα την λάβει τη χρονική στιγμή t=1. Την συσσωρευμένη αξία του κεφαλαίου από το της ληξιπρόθεσμες ράντες συμβολίζεται μαθηματικά με

2) **Προκαταβλητέες ράντες:** Και σε αυτή τη περίπτωση οι καταβολή γίνεται στη αρχή της περιόδου, για αυτό και λέγονται προκαταβλητέες καθώς προκαταβάλλονται. Αυτό το μοντέλο είναι ευρύτερα διαδεδομένο στη ασφαλιστική αγορά. Τα ασφάλιστρα καταβάλλονται προκαταβολικά και είναι απαραίτητη προϋπόθεση για την έναρξη του ασφαλιστηρίου συμβολαίου. Στη παρούσα εργασία με αυτό το είδος ραντών θα ασχοληθούμε εκτενέστερα στη συνέχεια. Ο μαθηματικός συμβολισμός της συσσωρευμένης αξίας της επένδυσης των προκαταβλητέων ραντών είναι

$\ddot{s}_{t]}$

Τώρα που περιγράψαμε τα βασικά γνωρίσματα και κατηγορίες των ραντών ας δούμε και την μαθηματική τους κατασκευή. Έστω ότι για χρόνια t θέλουμε να επενδύσουμε k χρηματικές μονάδες. Η απόδοση R_f θεωρούμε είναι σταθερή κατά την διάρκεια αυτών των t ετών. Θα υπολογίσουμε ποιο θα είναι το συσσωρευμένο κεφάλαιο μετά το πέρας των t αυτών χρόνων.

Αρχικά για τις ληξιπρόθεσμες ράντες

$$s_{t\rceil}= \frac{k}{t}\left(1+\left(1+R_f\right)+\left(1+R_f\right)^2+\left(1+R_f\right)^3+\cdots\left(1+R_f\right)^{t-1}\right)$$

$$\frac{k}{t}\,\frac{\left(1+R_f\right)^t-1}{\left(1+R_f\right)-1}=\frac{k}{t}\,\frac{\left(1+R_f\right)^t-1}{R_f(8)}$$

Θα εξηγήσουμε τώρα το πως προέκυψε η εξίσωση μέσω ενός πίνακα

Πίνακας 1.1

Χρονική Στιγμή	Ποσό που επενδύθηκε	Χρόνια μέχρι την λήξη της επένδυσης	Ανατοκισμός της καταβολής την χρονική στιγμή t ως την λήξη	Συσσωρευμένη αξία των πρώτων t καταβολών
0	0	t	$0(1+R_f)^t$	0
1	$\frac{k}{t}$	$t-1$	$\frac{k}{t}(1+R_f)^{t-1}$	$0+\frac{k}{t}(1+R_f)^{t-1}$
2	$\frac{k}{t}$	$t-2$	$\frac{k}{t}(1+R_f)^{t-2}$	$0+\frac{k}{t}(1+R_f)^{t-1}+\frac{k}{t}(1+R_f)^{t-2}$
…	…	…	…	…
$\boldsymbol{t-2}$	$\frac{k}{t}$	2	$\frac{k}{t}(1+R_f)^2$	$0+\frac{k}{t}(1+R_f)^{t-1}+\frac{k}{t}(1+R_f)^{t-2}+\cdots+\frac{k}{t}(1+R_f)^2$
$\boldsymbol{t-1}$	$\frac{k}{t}$	1	$\frac{k}{t}(1+R_f)^1$	$0+\frac{k}{t}(1+R_f)^{t-1}+\frac{k}{t}(1+R_f)^{t-2}+\cdots+\frac{k}{t}(1+R_f)^2+\frac{k}{t}(1+R_f)^1$
$\boldsymbol{t}$	$\frac{k}{t}$	0	$\frac{k}{t}(1+R_f)^0$	$0+\frac{k}{t}(1+R_f)^{t-1}+\frac{k}{t}(1+R_f)^{t-2}+\cdots+\frac{k}{t}(1+R_f)^2+\frac{k}{t}(1+R_f)^1+\frac{k}{t}(1+R_f)^0$

η τελευταία ισότητα στην εξίσωση είναι αποτέλεσμα της γεωμετρικής σειράς

Αντίστοιχα με τις ληξιπρόθεσμες ορίζονται και οι προκαταβλητέες. Η διαφορά μεταξύ των δύο όπως έχει ήδη λεχθεί είναι ότι στις προκαταβλητέες οι καταβολή γίνεται στη αρχή της περιόδου, την χρονική στιγμή 0 και επομένως δεν έχουμε καταβολή στη λήξη, την χρονική στιγμή t. Και στις δύο περιπτώσεις όμως το πλήθος των καταβολών είναι t, όσα τα χρόνια της συνολικής επένδυσης

$$=\frac{k}{t}\left(1+(1+R_f)+(1+R_f)^2+(1+R_f)^3+\cdots(1+R_f)^t\right)$$

$$=\frac{k}{t}\,\frac{(1+R_f)^t-1}{(1+R_f)-1}=\frac{k}{t}\,\frac{(1+R_f)^t-1}{d}(9)$$

που d το προεξοφλητικό επιτόκιο όπως έχει αναφερθεί παραπάνω.

Πίνακας 1.2

Χρονική Στιγμή	Ποσό που επενδύθηκε	Χρόνια μέχρι την λήξη της επένδυσης	Ανατοκισμός της καταβολής την χρονική στιγμή t ως την λήξη	Συσσωρευμένη αξία των πρώτων t καταβολών
0	$\frac{k}{t}$	$t+1$		$\frac{k}{t}(1+R_f)^{t+1}$
1	$\frac{k}{t}$	t	$\frac{k}{t}(1+R_f)^{t}$	$\frac{k}{t}(1+R_f)^{t+1}+\frac{k}{t}(1+R_f)^{t-1}$
2	$\frac{k}{t}$	$t-1$	$\frac{k}{t}(1+R_f)^{t-1}$	$\frac{k}{t}(1+R_f)^{t+1}+\frac{k}{t}(1+R_f)^{t}+\frac{k}{t}(1+R_f)^{t-1}$
…	…	…	…	…
$\boldsymbol{t-2}$	$\frac{k}{t}$	3	$\frac{k}{t}(1+R_f)^{3}$	$\frac{k}{t}(1+R_f)^{t+1}+\frac{k}{t}(1+R_f)^{t}+\frac{k}{t}(1+R_f)^{t-1}+\cdots+\frac{k}{t}(1+R_f)^{3}$
$\boldsymbol{t-1}$	$\frac{k}{t}$	2	$\frac{k}{t}(1+R_f)^{2}$	$\frac{k}{t}(1+R_f)^{t+1}+\frac{k}{t}(1+R_f)^{t}+\frac{k}{t}(1+R_f)^{t-1}+\cdots+\frac{k}{t}(1+R_f)^{3}+\frac{k}{t}(1+R_f)^{2}$
$\boldsymbol{t}$	0	1	$\frac{k}{t}(1+R_f)^{1}$	$\frac{k}{t}(1+R_f)^{t+1}+\frac{k}{t}(1+R_f)^{t}+\frac{k}{t}(1+R_f)^{t-1}+\cdots+\frac{k}{t}(1+R_f)^{3}+\frac{k}{t}(1+R_f)^{2}+\frac{k}{t}(1+R_f)^{1}$

Όπως και στη περίπτωση της ληξιπρόθεσμης

2.2 Θεωρία χαρτοφυλακίου

2.2.1 Οι λογαριθμικές αποδόσεις των περιουσιακών στοιχείων

Στο παρόν κεφάλαιο θα ασχοληθούμε με τη θεωρία του τόκου και ειδικότερα με τις αποδόσεις ενός περιουσιακού στοιχείου που εμπεριέχει κίνδυνο (risky asset). Ειδικότερα θα ασχοληθούμε με την κατανομής της τυχαίας μεταβλητής της απόδοσης σε ένα περιουσιακού στοιχείου που εμπεριέχει κίνδυνο.

Αρχικά να ορίσουμε την έννοια του περιουσιακού στοιχείου χωρίς κίνδυνο (risk-free asset). Είναι ένα asset, συνήθως κρατικό ομόλογο μια αξιόπιστης χώρας όπως οι Ηνωμένες πολιτείες ή η Γερμανία. Αν κάποιος επενδυτής αγοράσει ένα δεκαετές ομόλογο της Γερμανίας μπορούμε με μεγάλη βεβαιότητα να υποθέσουμε ότι θα πάρει πίσω το κεφάλαιο του. Θα ορίσουμε την απόδοση που λαμβάνει ο επενδυτής για αυτή του κίνηση ως R_f και η οποία για πολλούς **αντικατοπτρίζει την αξία του χρήματος που φθίνει στη πάροδο των ετών.**

Αντίθετα ένα risky asset, όπως μια μετοχή, είναι εκτεθειμένη στις μεταβολές των αγορών. Ο επενδυτής δεν είναι βέβαιος για το κεφάλαιο που θα λάβει από αυτή την κίνηση του. Για να ανταμειφθεί για το κίνδυνο αυτό που αναλαμβάνει, η απόδοση R της μετοχής θα πρέπει να είναι $R > R_f$.

Ας δούμε τώρα το πως θα ορίσουμε την απόδοση μια μετοχής μέσα στο χρονικό διάστημα (t-1, t) . Έστω ότι την στιγμή t-1 η μετοχή, ή γενικότερα το risky asset, έχει τιμή S_{t-1} και την χρονική στιγμή t έχει τιμή S_t. Ως απόδοση θα ορίσουμε

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}} \ (1)$$

Σχόλιο: Συγκρίνοντας τον ορισμό της απόδοσης ενός risky asset με τον ορισμό του αποτελεσματικού επιτοκίου θα παρατηρήσουμε ότι έχουν την ίδια «δομή». Η μόνη διαφορά είναι πως τον ρόλο της accumulated value τον χρόνο t, τον έχει η αξία του περιουσιακού στοιχείου τον χρόνο t.

Ως συνάρτηση της αξίας του asset η R_t ορίζεται στο $R_t \in [-1, +\infty)$. Όντως

$$\lim_{S_t \to 0} R_t = \frac{0 - S_{t-1}}{S_{t-1}} = -\frac{S_{t-1}}{S_{t-1}} = -1 \ (2)$$

, ενώ

$$\lim_{S_{t-1} \to 0} R_t = \frac{S_t - 0}{0} = \infty \ \ (3)$$

Όπως είναι ήδη εμφανές κομβικό ρόλο έχει η μέτρηση του χρόνου. Οι συνήθεις μονάδες μέτρησής του χρόνου είναι οι ημέρες, οι μήνες και τα χρόνια. Η σχέση που συνδέει την απόδοση ενός μεγαλύτερου χρονικού διαστήματος όπως ο χρόνος, με ένα μικρότερου θα προκύψει από την θεωρία του τόκου όπως αναλύθηκε στο προηγούμενο κεφάλαιο.

Έχουμε λοιπόν

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1 \ \Leftrightarrow (1 + R_t) = \frac{S_t}{S_{t-1}} \Leftrightarrow$$

$$S_{t-1}(1 + R_t) = S_t$$

Μάλιστα τα ανωτέρω θα ισχύουν για οποιαδήποτε μονάδα μέτρησης του χρόνου. Αν υποθέσουμε δηλαδή ότι R_t είναι η ετήσια απόδοση και $\xi_{t/12}$ είναι η μηνιαίες αποδόσεις για το ίδιο έτος θα έχουμε

$$S_{t-1}(1 + R_t) = S_t \ (5)$$

$$S_{t-1}(1 + \xi_1)(1 + \xi_2) \dots (1 + \xi_{12}) = S_t$$

, όπου ξ_1 είναι συμβολικά η απόδοση του Ιανουαρίου του έτους t και αντίστοιχα οι υπόλοιπες αποδόσεις.

Η αξία του περιουσιακού στοιχείου στη αρχή και στο τέλος τους έτους είναι ανεξάρτητη από την μονάδα μέτρησης του χρόνου.

Διαιρώντας τις δύο τελευταίες

$$\frac{(1+R_t)}{(1+\xi_1)(1+\xi_2)\dots(1+\xi_{12})}=1$$

$$\Leftrightarrow (1+R_t)=(1+\xi_1)(1+\xi_2)\dots(1+\xi_{12})\ (7)$$

Η γενικότερα αν έχουμε την απόδοση R_t μιας μονάδας μέτρησης του χρόνου ή οποία υποδιαιρείται σε k ίσα τμήματα τότε η (7) γίνεται

$$(1+R_t)=(1+\xi_1)(1+\xi_2)\dots(1+\xi_k\,)\ (7a)$$

Οι ποσότητες στο δεξί και αριστερό μέλος της τελευταίας εξίσωσης είναι καλά ορισμένες και μάλιστα με βάση των που αναφέρθηκαν νωρίτερα ξέρουμε ότι είναι και μη αρνητικές ποσότητες.

Παίρνοντας το λογάριθμο και στα δύο μέλη της έχουμε :

$$\ln(1+R_t)=\ln\big((1+\xi_1)(1+\xi_2)\dots(1+\xi_k\,)\big)\Leftrightarrow$$

$$\tilde{R}_t=\ln(1+R_t)=\sum_{i=1}^{k}\ln(1+\xi_i)$$

Η τελευταία εξίσωση μας δίνει μια σχέση μεταξύ της απόδοσης ενός περιουσιακοί στοιχείου στοιχειού την περίοδο t σε σχέση με την απόδοση ξ_i του ίδιου περιουσιακού στοιχείου την υποπερίοδο i της t.

Αν μάλιστα προσεγγιστικά υποθέσουμε ότι $\tilde{\xi}_i$ είναι ο λογάριθμός τους αθροίσματος $(1+\xi_i)$, δηλαδή

$$\ln(1+\xi_i)=\tilde{\xi}_i$$

Η εξίσωση παίρνει την μορφή

$$\tilde{R}_t=\sum_{i=1}^{k}\tilde{\xi}_i$$

Με αυτό τον τρόπο θα ορίζουμε τη στοχαστική ανέλιξη των λογαριθμικών αποδόσεων ενός με ρίσκο περιουσιακού στοιχείου $\{\tilde{R}_t\}_{t\in\mathbb{N}}$.

Όπως σε κάθε στοχαστική ανέλιξη έτσι και στη στοχαστική ανέλιξη των αποδόσεων θα μας απασχολήσουν εξής θέματα της διαχρονικής εξάρτησης, ετερογένειας και κατανομής των αποδόσεων αμοιβαίων κεφαλαίων στη ελληνική αγορά

2.2.2 Το πρόβλημα της ετεροσκεδαστικότητας

Αυτό που έχει παρατηρηθεί είναι ότι οι μεγάλες αποδόσεις συνήθως ακολουθούνται από άλλες μεγάλες αποδόσεις, ενώ οι μικρές αποδόσεις τείνουν να ακολουθούνται από μικρές αποδόσεις. Αυτό το φαινόμενο είναι γνωστό στη βιβλιογραφία ως "δυναμική ετεροσκεδαστικότητα". Ο πρώτος πού παρατήρησε αυτό το φαινόμενο ήταν ο Γάλλος μαθηματικός Benoit Mandelbrot το 1962 σε μιά μελέτη του για τη συμπεριφορά των τιμών του βαμβακιού. Ένα σημαντικό εύρημα της σύγχρονης βιβλιογραφίας είναι ότι η ένταση του φαινομένου της δυναμικής ετεροσκεδαστικότητας αυξάνεται όσο μειώνεται το χρονικό διάστημα των παρατηρήσεων ή, με άλλα λόγια, όσο αυξάνεται η συχνότητα των παρατηρήσεών μας. Επιπλέον, όσο μειώνεται το χρονικό διάστημα των παρατηρήσεών μας, αυξάνεται και ο βαθμός της μη κανονικότητας των αποδόσεων. Ο συντελεστής κύρτωσης για τις ημερήσιες αποδόσεις είναι πάνω από τρεις φορές μεγαλύτερος από τον συντελεστή κύρτωσης για τις τριμηνιαίες αποδόσεις. Αυτό το φαινόμενο αναφέρεται στη βιβλιογραφία ως "παχιά άκρα" (fat tails) της κατανομής των αποδόσεων. Η αντίστοιχη κατανομή των αποδόσεων χαρακτηρίζεται ως "λεπτόκυρτη", με τον βαθμό της λεπτοκύρτωσης να είναι υψηλότερος στις ημερήσιες αποδόσεις.

3 Κατανομή αποδόσεων αμοιβαίων κεφαλαίων στην ελληνική αγορά

3.1 Δεδομένα και επιλογή

Για να προσδιορίσουμε την κατανομή των αμοιβαίων κεφαλαίων στη ελληνική αγορά θα περιοριστούμε στη μελέτη των παρακάτω αμοιβαίων κεφαλαίων για τα ημερολογιακά έτη 2018 και 2019 καθώς κρίθηκε αδόκιμο να χρησιμοποιηθούν οικονομικά δεδομένα των ετών 2020 και 2021 που ήταν επηρεασμένα από το πανδημία του COVID-19 ενός γεγονότος εκτός «κανονικότητας» που επηρέασε ωστόσο την χρηματοοικονομική αγορά παγκοσμίως.

<u>Πίνακας 2.1</u>

Αμοιβαία Κεφάλαια	Ημερες Παρατηρησεων
ALPHA (LUX) Global Equity FoFs EUR	498
ALPHA (LUX) Global Equity FoFs EUR I	498
ALPHA FUND OF FUNDS COSMOS STARS COMMODITIES METOXIKO Classic	497
ALPHA FUND OF FUNDS COSMOS STARS COMMODITIES METOXIKO Institutional	497
ALPHA FUND OF FUNDS COSMOS STARS EUROPE METOXIKO Classic	502
ALPHA FUND OF FUNDS COSMOS STARS EUROPE METOXIKO Institutional	502
ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA METOXIKO Classic	502
ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA METOXIKO Institutional	502
ALPHA FUND OF FUNDS COSMOS STARS USA METOXIKO Classic	502
ALPHA FUND OF FUNDS COSMOS STARS USA METOXIKO Institutional	502
EUROBANK (LF) FUND OF FUNDS- ESG FOCUS	502
EUROBANK $ (LF) FUND OF FUNDS- ESG FOCUS	502
EUROBANK (LF) Fund of Funds-Equity Blend	502
EUROBANK (USD) (LF) Fund of Funds - Equity Blend	502
Eurobank GF Equity Blend Fund of Funds Μετοχικό	502
EUROBANK I (LF) Fund of Funds-Equity Blend	502
EUROBANK (LF) FOF - Global Emerging Markets	502
EUROBANK (USD) (LF) FoF - Global Emerging Markets	502
EUROBANK I (LF) FOF - Global Emerging Markets	502
INTERAMERICAN (LF) FUND OF FUNDS-ESG FOCUS	502
INTERAMERICAN (LF) FOF - Global Emerging Markets	502
NN FUND OF FUNDS ΕΜΠΟΡΕΥΜΑΤΙΚΩΝ ΑΞΙΩΝ	502
NN FUND OF FUNDS ΕΥΡΩΠΑΪΚΟ ΜΕΤΟΧΙΚΟ	502
Optima premium selection fund of funds μετοχικό	502
PRIVATE BANKING (LF) FoF - Global Emerging Markets	502
PRIVATE BANKING CLASS (LF) Fund of Funds - Equity Blend	502
PRIVATE BANKING USD (LF) FOF Equity Blend	502
S&B Συνταξιοδοτικό Global Fund of Funds Μετοχικό	502
ΔΗΛΟΣ SYNTHESIS BEST RED ESG - FUND OF FUNDS ΜΕΤΟΧΙΚΟ	502

Στη δεύτερη στήλη του παραπάνω πίνακα φαίνονται οι ημέρες που τα αμοιβαία αυτά κεφάλαια πραγματοποιούσαν συναλλαγές μέσα στα δύο αυτά χρόνια. Αν αναλογιστούμε ότι οι εργάσιμες ημέρες , πλην των θρησκευτικών ή άλλων αργιών, από 1/1/2018 έως 31/12/2019 ήταν 522 μπορούμε να θεωρήσουμε ότι μας δίνουν μας δίνουν την απαραίτητη εικόνα για τους σκοπούς αυτής της εργασίας.

Ο βασικός λόγος που επιλεχθήκαν τα παραπάνω fund είναι ότι είναι χαρακτηρισμένα ως χαμηλού ρίσκου και προτιμήθηκαν καθώς τα unit linked είναι ασφαλιστικά προγράμματα με στόχο τη επένδυση – αποταμίευση και κυρίως χρησιμοποιούνται από μικρό επενδυτές.

3.2 Στατιστικοί έλεγχοι για τις ημερήσιες αποδόσεις των αμοιβαίων κεφαλαίων

Στη παρούσα παράγραφο θα επαληθεύσουμε τα θεωρητικά ευρήματα που αναφέρθηκαν στη υποπαράγραφο 1.2.2 για την δυναμική ετεροσκεδαστικότητα πως αυτή γίνεται εντονότερη καθώς το πλήθος των παρατηρήσεων αυξάνεται και το πως αυτό. Πριν όμως αυτού θα αναφέρουμε κάποια βασικά στατιστικά τεστ που θα χρησιμοποιηθούν για να είναι ευκολότερο στο αναγνώστη την κατανόηση της στατιστικής ανάλυσης που θα ακολουθηθεί.

3.2.1 Έλεγχοι κανονικότητας ενός πληθυσμού

3.2.1.1 Έλεγχος Kolmogorov – Smirnov

Στη στατιστική, ο δοκιμαστικός Κολμογόροφ-Σμίρνοφ (δοκιμασία Κ-Σ ή KS test) είναι ένας μη παραμετρικός έλεγχος της ισότητας συνεχών (ή ασυνεχών) μονοδιάστατων πιθανοτήτων, που μπορεί να χρησιμοποιηθεί για να ελεγχθεί εάν ένα δείγμα προέρχεται από μια συγκεκριμένη αναφορική πιθανοτική κατανομή (δοκιμασία Κ-Σ μονού δείγματος), ή για να ελεγχθεί εάν δύο δείγματα προέρχονται από την ίδια κατανομή (δοκιμασία Κ-Σ δύο δειγμάτων). Εν γένει, ο δοκιμαστικός παρέχει έναν τρόπο για να απαντήσουμε ποιοτικά στο ερώτημα "Πόσο πιθανό είναι να δούμε μια συλλογή δειγμάτων όπως αυτά, αν είχαν επιλεγεί από εκείνη την πιθανοτική κατανομή;" ή, στη δεύτερη περίπτωση, "Πόσο πιθανό είναι να δούμε δύο σύνολα δειγμάτων όπως αυτά, αν είχαν επιλεγεί από την ίδια (αλλά άγνωστη) πιθανοτική κατανομή;". Ο δοκιμαστικός έχει το όνομά του από τους Αντρέι Κολμογόροφ και Νικολάι Σμίρνοφ

Έστω $X_1, X_2, \dots, X_n$ ένα τυχαίο δείγμα από συνεχή πληθυσμό X με συνάρτηση κατανομής $F(x)$ μας ενδιαφέρει να εξετάσουμε το εξής έλεγχο

$$H_0: F\ (x) = F_0(x) - \ H_1: F(x) \neq F_0(x)$$

Η στατιστική συνάρτηση μετρά την απόσταση μεταξύ της εμπειρικής συνάρτησης κατανομής του δείγματος και της συγκεντρωτικής συνάρτησης κατανομής της αναφερόμενης κατανομής, ή μεταξύ των εμπειρικών συναρτήσεων κατανομής δύο δειγμάτων. Δηλαδή

$$D_n = \sup\{|S_n(x) - F_0(x)|\}, -\infty < x < \infty$$

Όπου

$$S_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x)$$

Η p-value του ελέγχου είναι ίση με $P(D_n > d_n)$.

Επειδή η κατανομή της στατιστικής συνάρτησης του ελέγχου είναι δύσκολο να υπολογιστεί υπάρχουν πίνακες με τα ποσοστιαία σημεία της *D.* Ωστόσο , έχει αποδειχθεί ότι

$$\lim_{n\to\infty} P\left(D_n \leq \frac{d}{\sqrt{n}}\right) = 1 - 2\sum_{i=1}^{\infty}\left((-1)^{i-1}e^{-2i^2d^2}\right)$$

3.2.1.2 Άλλοι έλεγχοι κανονικότητας ενός πληθυσμού

Ο έλεγχος **Shapiro-Wilk** προτιμάται σε σχέση με τον έλεγχο Kolmogorov-Smirnov για μικρά δείγματα ($n < 50$) και έχει γενικά υψηλή ικανότητα ανίχνευσης. Η στατιστική συνάρτηση του ελέγχου αξιολογεί πόσο κοντά βρίσκονται τα εμπειρικά ποσοστιαία του δείγματος στα αντίστοιχα θεωρητικά ποσοστιαία της κανονικής κατανομής.

Ο test αυτό ελέγχει τη μηδενική υπόθεση ότι ένα δείγμα $X_1, X_2, \dots, X_n$ προέρχεται από μια κανονικά κατανεμημένη πληθυσμό. Η στατιστική σταθερά του ελέγχου είναι

$$W = \frac{\left(\sum_{i=1}^{n} a_i X_{(i)}\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

Όπου $X_{(i)}$ με παρενθέσεις γύρω από το δείκτη i αναφέρεται στην i-οστή ταξινομημένη στατιστική, δηλαδή το i-οστό μικρότερο αριθμό στο δείγμα (όχι να μπερδευτεί με X_i)

Ο **έλεγχος Lilliefors** είναι ένας έλεγχος για την κανονικότητα που βασίζεται στον έλεγχο Kolmogorov–Smirnov. Χρησιμοποιείται για να ελέγξει τη μηδενική υπόθεση ότι τα δεδομένα προέρχονται από έναν κανονικά κατανεμημένο πληθυσμό, όταν η μηδενική υπόθεση δεν καθορίζει ποια κανονική κατανομή ακριβώς, δηλαδή δεν καθορίζει την αναμενόμενη τιμή και τη διακύμανση της κατανομής. Ο έλεγχος αυτός πήρε το όνομά του από τον Hubert Lilliefors, καθηγητή στατιστικής στο Πανεπιστήμιο George Washington.

Μια παραλλαγή του ελέγχου μπορεί να χρησιμοποιηθεί για να ελέγξει τη μηδενική υπόθεση ότι τα δεδομένα προέρχονται από έναν εκθετικά κατανεμημένο πληθυσμό, όταν η μηδενική υπόθεση δεν καθορίζει ποια εκθετική κατανομή ακριβώς

Για να εφαρμοστεί ο έλεγχος Kolmogorov-Smirnov, είναι αναγκαίο να είναι εκ των προτέρων γνωστές οι παράμετροι του πληθυσμού. Αν οι παράμετροι δεν είναι γνωστές και χρησιμοποιήσουμε εκτιμήσεις τους ως πραγματικές τιμές των ορισμάτων της συνάρτησης ks.test της γλώσσας προγραμματισμού *R*, το αποτέλεσμα του ελέγχου δεν θα είναι αξιόπιστο. Σε αυτήν την περίπτωση, για τον έλεγχο της κανονικότητας του πληθυσμού, είναι προτιμότερο να χρησιμοποιήσουμε τον ελέγχο Lilliefors του πακέτου nortest της γλώσσας προγραμματισμού *R*, μια παραλλαγή του ελέγχου Kolmogorov-Smirnov, ο οποίος παρέχει ένα προσαρμοσμένο p-value

3.2.2 Q-Q Διαγράμματα

Τα διαγράμματα Q-Q μας δίνουν την δυνατότητα να εξετάσουμε αν ένα τυχαίο δείγμα προέρχεται $X_1, X_2, \ldots, X_n$ από μια γνωστή κατανομή $F(x)$. Το Q-Q διάγραμμα δημιουργεί ζεύγη $\left(F^{-1}(p_i), x_{(i)}\right)$ όπου

$$p_i = \frac{i-a}{n+1-a}, \qquad \alpha = \begin{cases} \frac{3}{8}, & n \leq 10 \\ \frac{1}{2}, & n > 10 \end{cases}$$

και $x_{(i)}$ η i μεγαλύτερη τιμή του δείγματος

Όταν τα ζεύγη αυτά είναι περιμετρικά και όσο γίνεται πιο κοντά στην ευθεία $y = x$ δηλαδή τα θεωρητικά ποσοστιμόρια είναι «ίσα» τότε μπορούμε να ισχυριστούμε ότι το δείγμα προέρχεται από την κατανομή αυτή.

Τέτοια διαγράμματα μπορούμε να σχεδιάσουμε με την γλώσσα προγραμματισμού R χρησιμοποιώντας την εντολή *qqplot(x,y)* όπου το x είναι τα ποσοστιμόρια που μπορούμε να υπολογίσουμε με το ανωτέρω τύπο (σε μορφή διανύσματος) και y το δείγμα μας.

Πιο συγκεκριμένα η R για την κανονική κατανομή μπορεί να αυτόματα να σχεδιάσει αυτά τα σημεία μέσω της εντολής *qqnorm(x)* όπου το x είναι το δείγμα μας σε μορφή διανύσματος και να προσθέσει μια επιπλέον γραμμή μέσω της εντολής *qqline(x)* η οποίο περνάει από το πρώτο και τρίτο ποσοστιμόριο της τυποποιημένης κανονικής.

3.3 Στατιστικά αποτελέσματα

3.3.1 Οι Στατιστικοί έλεγχοι

Με βάση και το κεφάλαιο ένα την διπλωματικής αυτή θα ελέγξουμε αν η λογαριθμικές αποδόσεις των funds που αναφέρθηκαν στη αρχή αυτού του κεφαλαίου, ακολουθούν κανονική κατανομή. Αυτό βεβαία αυτόματα σημαίνει ότι οι αποδόσεις ακολουθούν την λογαριθμοκανονική.

Επειδή ωστόσο βασικές συναρτήσεις της λογαριθμοκανονικής, όπως η ροπογεννήτρια και ο μετασχηματισμός Laplace δεν είναι σε κλειστή μορφή, στα πλαίσια της διπλωματικής αυτής έγιναν οι έλεγχοί που αναφέρθηκαν στις δύο προηγούμενες παραγράφους.

Ακολουθεί ο πίνακας που καταγραφεί τα p-values των ελέγχων σε για κάθε ένα από τα funds.

Πίνακας 2.2

fund_names,	ks_test_results	shapiro_test_results	lillie_test_results
ALPHA (LUX) Global Equity FoFs EUR I,	77.60%	60.38%	38.31%
ALPHA FUND OF FUNDS COSMOS STARS COMMODITIES ΜΕΤΟΧΙΚΟ Classic,	77.32%	60.32%	37.88%
ALPHA FUND OF FUNDS COSMOS STARS COMMODITIES ΜΕΤΟΧΙΚΟ Institutional,	72.96%	19.21%	31.53%
ALPHA FUND OF FUNDS COSMOS STARS EUROPE ΜΕΤΟΧΙΚΟ Classic,	73.41%	19.24%	32.15%
ALPHA FUND OF FUNDS COSMOS STARS EUROPE ΜΕΤΟΧΙΚΟ Institutional,	56.60%	49.17%	14.64%
ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA ΜΕΤΟΧΙΚΟ Classic,	56.32%	49.10%	14.43%
ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA ΜΕΤΟΧΙΚΟ Institutional,	48.08%	9.29%	8.94%
ALPHA FUND OF FUNDS COSMOS STARS USA ΜΕΤΟΧΙΚΟ Classic,	47.65%	9.25%	8.71%
ALPHA FUND OF FUNDS COSMOS STARS USA ΜΕΤΟΧΙΚΟ Institutional,	98.69%	78.88%	91.40%
EUROBANK $ (LF) FUND OF FUNDS- ESG FOCUS,	98.68%	78.85%	91.37%
EUROBANK (LF) FUND OF FUNDS- ESG FOCUS,	94.64%	54.60%	75.82%
EUROBANK (LF) Fund of Funds-Equity Blend,	94.64%	54.57%	75.82%
EUROBANK (USD) (LF) Fund of Funds - Equity Blend,	94.36%	64.03%	74.92%
Eurobank GF Equity Blend Fund of Funds Μετοχικό,	94.36%	64.14%	74.94%
EUROBANK I (LF) Fund of Funds-Equity Blend,	92.69%	62.73%	69.90%
EUROBANK (LF) FOF - Global Emerging Markets,	94.09%	63.87%	74.08%
EUROBANK (USD) (LF) FoF - Global Emerging Markets,	67.59%	17.40%	24.87%
EUROBANK I (LF) FOF - Global Emerging Markets,	67.66%	17.40%	24.95%
INTERAMERICAN (LF) FUND OF FUNDS-ESG FOCUS,	67.08%	17.51%	24.30%
INTERAMERICAN (LF) FOF - Global Emerging Markets,	94.63%	54.57%	75.79%
NN FUND OF FUNDS ΕΜΠΟΡΕΥΜΑΤΙΚΩΝ ΑΞΙΩΝ,	67.59%	17.40%	24.87%
NN FUND OF FUNDS ΕΥΡΩΠΑΪΚΟ ΜΕΤΟΧΙΚΟ,	63.54%	34.22%	20.61%
Optima premium selection fund of funds μετοχικό,	73.23%	57.44%	31.90%
PRIVATE BANKING (LF) FoF - Global Emerging Markets,	81.72%	60.97%	45.27%
PRIVATE BANKING CLASS (LF) Fund of Funds - Equity Blend,	67.53%	17.38%	24.81%
PRIVATE BANKING USD (LF) FOF Equity Blend,	94.36%	64.03%	74.92%
S&B Συνταξιοδοτικό Global Fund of Funds Μετοχικό,	94.36%	64.03%	74.94%
ΔΗΛΟΣ SYNTHESIS BEST RED ESG - FUND OF FUNDS ΜΕΤΟΧΙΚΟ,	79.50%	58.46%	41.40%
ALPHA (LUX) Global Equity FoFs EUR I,	81.28%	65.53%	44.48%

Όπως είναι εμφανές δεν απορρίπτεται η κανονικότητα σε επίπεδα στατιστικής εμπιστοσύνης 1% και 5% ενώ για λίγες περιπτώσεις θα απορρίπταμε την υπόθεση της κανονικότητας με επίπεδο στατιστικής εμπιστοσύνης 10%.

Παρατηρούμε ιδιαίτερα υψηλές μεγάλες τιμές του Kolmogorov – Smirnov τεστ ως ενδείξεις κανονικότητας αλλά όπως αναφέρθηκε και παραπάνω ο εν λόγω έλεγχος «δουλεύει» για μεγάλα δείγματα $n>2.000$.

Καθώς το δείγμα μας για κάθε fund είναι οι είκοσι δύο μήνες, από Φεβρουάριο του 2018 έως και Νοέμβριο του 2019 ο έλεγχος που θα προτείναμε ως πιο αξιόπιστο είναι ο Shapiro – Wilk. Θα μπορούσαμε και εδώ να παρατηρήσουμε υψηλά ποσοστά για τα p – values αλλά θα πρέπει να αναφερθεί πως το **μέσος όρος των τιμών αυτών** είναι στο 46.34% **το χαμηλότερο που παρατηρήθηκε**. Ωστόσο θα απορρίπταμε την κανονικότητα στα συνήθης επίπεδα εμπιστοσύνης μόλις για δύο από τα συνολικά είκοσι εννιά funds.

Τέλος για το έλεγχο Lillieforse, ο οποίος εμφάνισε και τις μικρότερες τιμές που παρατηρήθηκαν (8.94% για το ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA ΜΕΤΟΧΙΚΟ Institutional, και 8.71% για το ALPHA FUND OF FUNDS COSMOS STARS

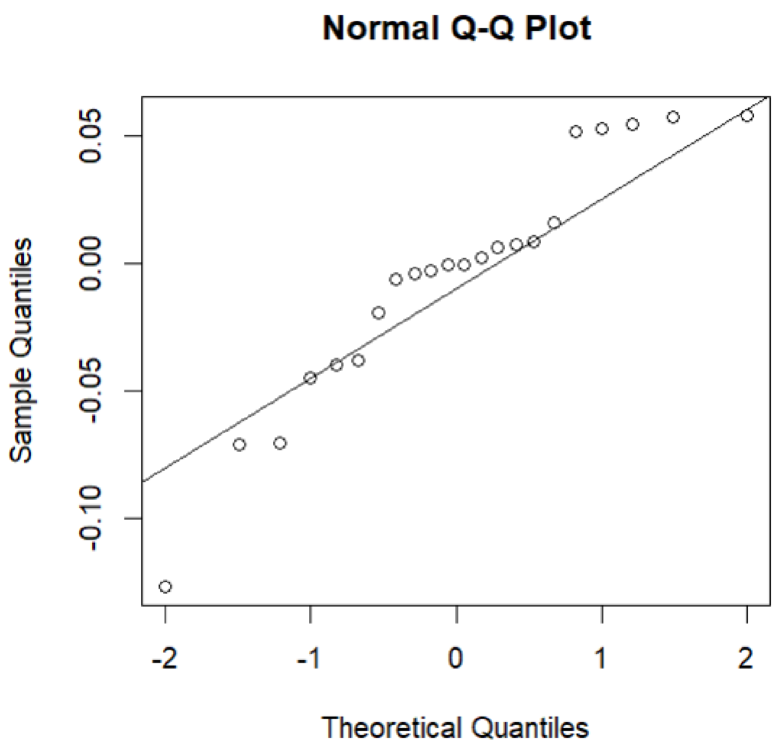
USA METOXIKO Classic) ο μέσος όρος των p – values κυμάνθηκε στα ίδια πεδία με του ελέγχου Shapiro – Wilk, καθώς ήταν 46.48%.

3.3.2 Q-Q Διαγράμματα

Για τα δύο funds όπου εμφανίστηκαν οι χαμηλότερες τιμές στου ελέγχους θα εξετάσουμε την κανονικότητα τους μέσω και γραφικών παραστάσεων *qqplots* καθώς και για το fund με το μεγαλύτερο p – value κατά μέσο όρο.

ALPHA FUND OF FUNDS COSMOS STARS USA METOXIKO Classic,

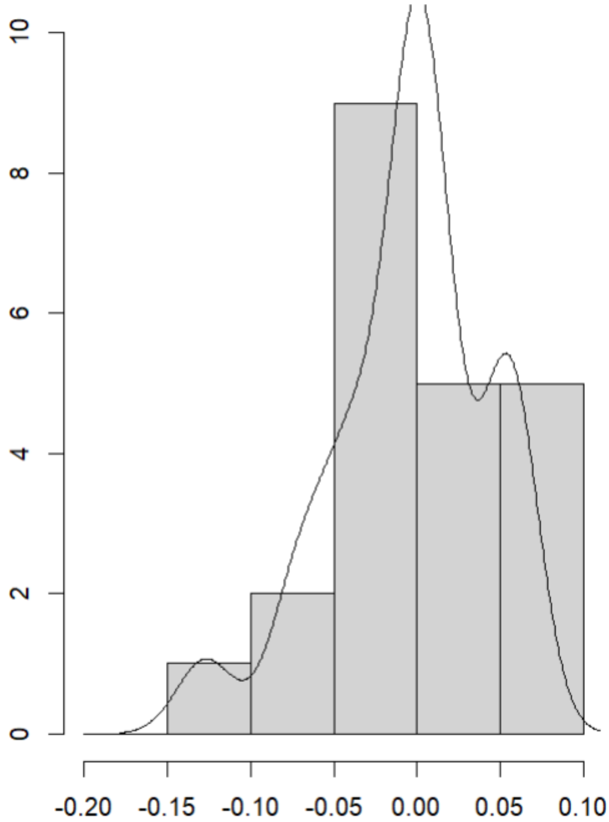
Γράφημα 2.1



Βάση της θεωρίας τα σημεία θα έπρεπε να είναι στη σχεδιασμένη ευθεία. Παρατηρούμε ότι ενώ σχετικά βρίσκονται κοντά στη γραμμή υπάρχει μια τάση να βρίσκονται πάνω από αυτήν. Ακόμα τέσσερις τιμές φαίνονται να μην έχουνε την ίδια κλίση με την ευθεία και αυτό αποτελεί ένδειξη για την μην κανονικότητα του πληθυσμού.

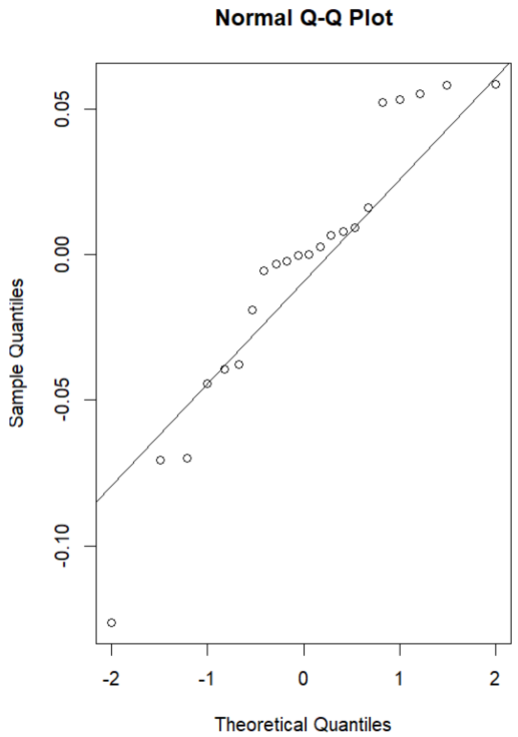
Γράφημα 2.2

Στα ίδια συμπεράσματα θα καταλήγαμε και για με βάση το ιστόγραμμα για τις τιμές αυτές. Προσαρμόζοντας πάνω στις γραμμές μια συνάρτηση πυκνότητας πιθανότητας φαίνεται να μην έχει την ιδιότητες της κανονικής, μονοκόρυφη και συμμετρική. Ωστόσο δεν μπορούμε να απορρίψουμε την υπόθεση της



Την ίδια διαδικασία θα επαναλάβουμε και για το

ALPHA FUND OF FUNDS COSMOS STARS SILK ROUTE ASIA METOXIKO Institutional

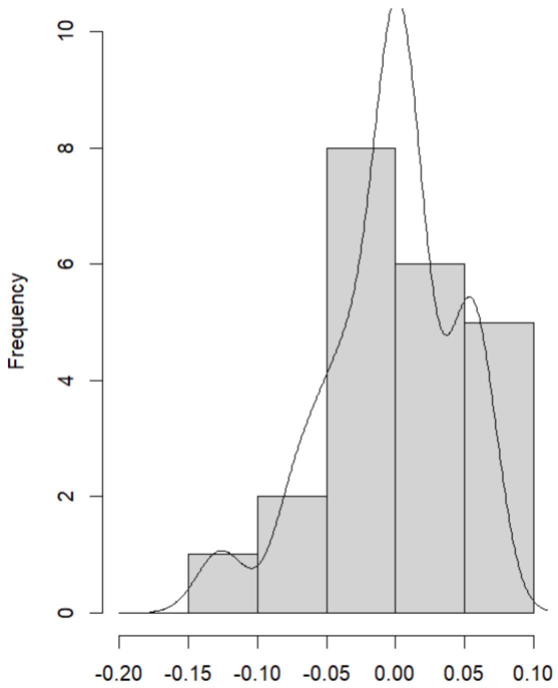


Γράφημα 2.3

Για το *qqplot* διάγραμμα έχουμε, το οποίο είναι παρεμφερές με το αντίστοιχο διάγραμμα το προηγουμένου fund, αλλά οι αποστάσεις από την γραμμή είναι πιο μεγάλες.

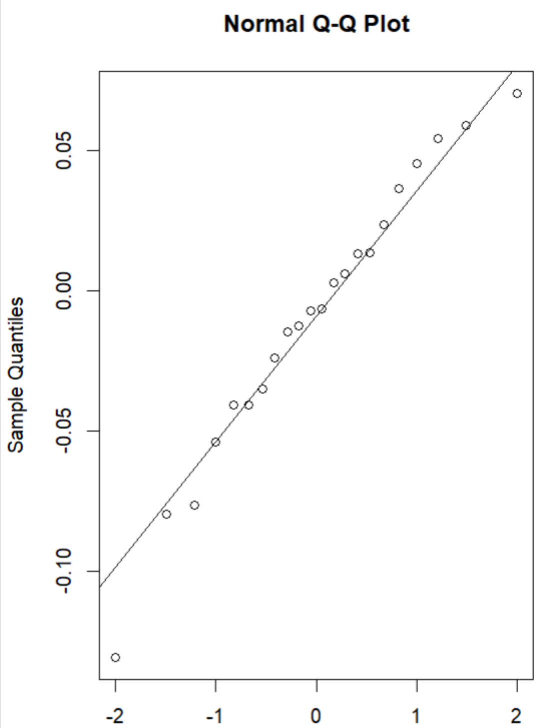
Γράφημα 2.4

Και στο ιστόγραμμα φαίνεται να έχουν την ίδια κατανομή οι λογαριθμικές αποδόσεις των δύο funds, και αυτό το πιο πιθανό είναι να έχει σχέση με το χαρτοφυλάκιο τους το οποίο φαίνεται να επηρεάζεται με το ίδιο τρόπο.



Θα κλείσουμε αυτή την υποπαραγραφό με τα γραφήματα του fund που ο μέσος όρος των p – values ήταν ο υψηλότερος 89.66%

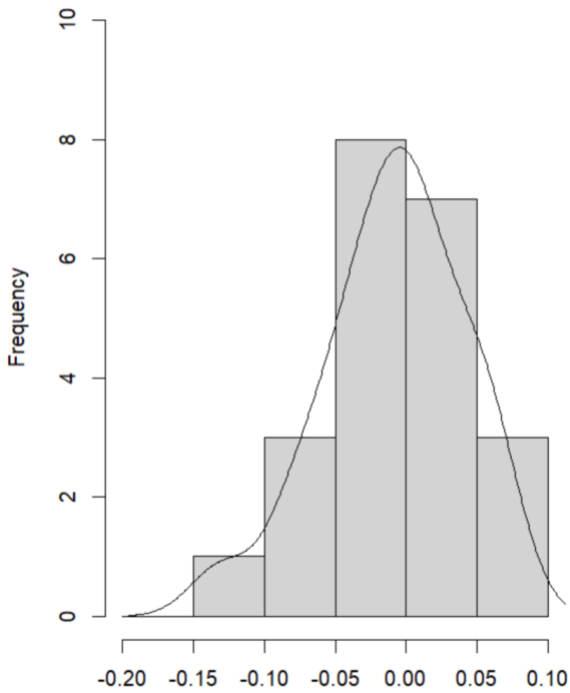
ALPHA FUND OF FUNDS COSMOS STARS USA METOXIKO Institutional



Όπως ήταν αναμενόμενο οι τιμές είναι πολύ κοντά στη ευθεία και οποιασδήποτε τιμές και να επιλέγαμε για να χαράξουμε μια εναλλακτική ευθεία θα ήταν παράλληλη και κοντά στη ευθεία που είναι αυτόματα δοσμένη.

<u>Γράφημα 2.6</u>

Το ιστόγραμμα μας δείχνει μια εικόνα από συμμετρική και μονοκόρυφη κατανομή που σωρευτικά με όλα τα προηγούμενα στοιχεία και ελέγχους που έχουν γίνει μας υποδεικνύει ότι είναι από κανονική κατανομή.



4 Μοντέλα επιβίωσης για ένα άτομο με ένα και πολλαπλά αίτια εξόδου

4.1 Μοντέλα επιβίωσης με ένα αίτιο εξόδου

4.1.1 Συνάρτηση κατανομής και συνάρτηση επιβίωσης

Ας υποθέσουμε ότι έχουμε ένα νεογέννητο άτομο, θα ορίσουμε ως X την τυχαία μεταβλητή των χρόνων που θα ζήσει. Όντως η X είναι καλά ορισμένη ως τυχαία μεταβλητή ως μετρήσιμη συνάρτηση με πεδίο ορισμού κάποιο δειγματικό χώρο Ω στο $\Re=[0,\infty)$ καθώς είναι αβέβαιο το πότε θα αποβιώσει ενώ λόγω της φύσης της δεν παίρνει αρνητικές τιμές.

Ο ορισμός της τυχαίας μεταβλητής X μας οδηγεί και στο ορισμό της **συνάρτησης κατανομής της.** Ποια είναι πιθανότητα να ζήσει ένα νεογέννητο άτομο ως τα 50 του έτη;

$$F_{50} = P(\textit{να ζήσει ένα νεογγένητο άτομο ως τα } 50) = P(X \leq 50)$$

Η γενικότερα

$$F_x = P(\textit{να ζήσει ένα νεογηένητο άτομο ως τα } x \textit{ έτη}) = P(X \leq x)$$

Καθώς ο χρόνος είμαι συνεχής από την θεωρία πιθανοτήτων η παραπάνω εξίσωση ισοδυναμεί με την

$$F_X(x) = P(X \leq x) \ \ (1)$$

Η παραπάνω πιθανότητα είναι η πιθανότητα ένα που μόλις γεννήθηκε, άρα ηλικίας 0, να ζήσει ως τα x του έτη ή αντίστοιχα να πεθάνει πριν τα x γενέθλια. Για αυτό είναι γνωστή και ως συνάρτηση θανάτου.

Ορίζουμε λοιπόν την **συνάρτηση επιβίωσης**, την συμπληρωματική της συνάρτηση θανάτου ως προς το χώρο πιθανότητας, $S(x)$ ως

$$S_X(x) = P(X \geq x) = 1 - P(X \leq x) = 1 - F_X(x) \ (2)$$

Όπως είναι φυσικό για την τυχαία μεταβλητή X ως συνεχή θα ορίζουμε και την **συνάρτηση πυκνότητας πιθανότητας** f όποτε έχουμε

$$f(x) = \frac{d}{dx} F_X(x) = \lim_{dx \to 0} \frac{F_X(x+dx) - F_X(x)}{dx} =$$

$$\lim_{dx \to 0} \frac{P(x < X \leq x + dx)}{dx}$$

Με την βοήθεια της συνάρτησης πυκνότητας μπορούμε τώρα να γράψουμε τη συνάρτηση κατανομής ως :

$$F_X(x) = P(X \leq x) = \int_0^x f(u)\, d(u) \quad (1b)$$

Έστω τώρα ότι θέλουμε να εξετάσουμε την πιθανότητα θανάτου ενός ατόμου νεογέννητου ατόμου να πεθάνει στη ηλικία *x.* Καθώς η Χ είναι συνεχής η πιθανότητα σε ένα σημείο είναι 0. Θωρούμε dx απειροστό του χρόνου π.χ. $dx = \frac{1}{365}$. Δηλαδή $X \in (x\,, x + dx)$. Τότε η πιθανότητα θα είναι

$$P(x < X \leq x + dx) = \int_x^{x+dx} f_X(u)\, du =$$

$$F_X(x + dx) - F_X(x) =$$

$$S_X(x) - S_X(x + dx) \; (4)$$

Αν επιπλέον υποθέσουμε ότι το έχει ήδη φτάσει στη ηλικία x λογώ της δεσμευμένης πιθανότητας θα έχουμε

$$P(x < X \leq x + dx \mid X > X) = P(x < X \leq x + dx)/P(x < X) =$$

$$\frac{F_X(x + dx) - F_X(x)}{1 - F_X(x)} \cong \frac{f_X(x)dx}{1 - F_X(x)} \; (5)$$

Η παραπάνω πιθανότητα είναι βάση για το συλλογισμός του υπολειπόμενου χρόνου ζωής που θα εξετάσουμε στη επόμενη παράγραφο. Τη δεομένη στιγμή απλά θα αφήσουμε το dx να τείνει στο 0 και θα ορίσουμε έτσι την **ένταση θνησιμότητας**. Ως ένταση θνησιμότητας (*force of mortality)* ορίζεται το εξής όριο

$$\mu_x = \lim_{dx \to 0} \frac{P(x < X \leq x + dx \mid X > x\,)}{dx}$$

$$= \lim_{dx \to 0} \frac{P\big((x < X \leq x + dx\,) \cap \; (X > x\,)\big)}{P(X > x\,)dx}$$

Επειδή όμως

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{\lim_{dx \to 0}\big(F_X(x + dx) - F_X(x)\big)}{dx} =$$

$$\lim_{dx \to 0} \frac{P(x < X \leq x + dx\ )}{dx}$$

Καταλήγουμε

$$\mu_x = \frac{f_X(x)}{P(x < X)} = \frac{f_X(x)}{S_X(x)}$$

Δηλαδή

$$\mu_x = -\frac{\frac{d}{dx} S_X(x)}{S_X(x)}$$

ή

$$\mu_x = -\frac{d}{dx} \ln\big(S_X(x)\big)$$

Και ολοκληρώνοντας

$$S_X(x) = e^{-\int_0^x \mu_u du}$$

Κλείνοντας την παράγραφο αυτή παραθέτουμε ένα πίνακα με τις εξισώσεις που συνδέουν τις τέσσερις αυτές συναρτήσεις μεταξύ τους

<u>Πίνακας 4.1</u>

	Συνάρτηση Κατανομής	Συνάρτηση Πυκνότητας Πιθανότητας	Συνάρτηση Επιβίωσης	Ένταση Θνησιμότητας
Συνάρτηση Κατανομής	—	$\frac{dF_X(x)}{dx}$	$1 - F_X(x)$	$\frac{\frac{d}{dx} F_X(x)}{1 - F_X(x)}$
Συνάρτηση Πυκνότητας Πιθανότητας	$\int_0^x f_X(u) du$	—	$\int_x^\infty f_X(u) du$	$\frac{f(x)}{\int_x^\infty f(u) du}$
Συνάρτηση Επιβίωσης	$1 - S(x)$	$-\frac{dS_X(x)}{dx}$	—	$-\frac{\frac{d}{dx} S_X(x)}{1 - S_X(x)}$
Ένταση Θνησιμότητας	$1 - e^{-\int_0^x \mu_u du}$	$\mu_x e^{-\int_0^x \mu_u du}$	$e^{\int_0^x \mu_u du}$	—

Σχόλιο 1: Αν και θεωρητικά η τιμή της X μπορεί να πάρει και άπειρες τιμές στη πράξη ένα άτομο δεν μπορεί να ζήσει για πάντα. Έτσι θεωρούμε ένα ω ως έσχατη ηλικία ή οριακή ηλικία (*the limiting age)* όπου θεωρούμε σίγουρο το θάνατο. Άρα πιο σωστά θα λέγαμε ότι το σύνολο τιμών της X είναι το $\Re$=$[0, \omega]$

Σχόλιο 2 : Είναι ευρέως γνωστός ο αναλογιστικός συμβολισμός *(x)* όταν αναφερόμαστε σε ένα άτομο ηλικίας x .

Σχόλιο 3: Ένας ακόμα αναλογιστικός συμβολισμός που θα ακολουθήσουμε είναι αντί για *f(x)* θα γράφουμε f_x

4.1.2 Υπολειπόμενος χρόνος ζωής

Στα ασφαλιστήρια ζωής έχει σημαντικό ρόλο ο υπολειπόμενος χρόνος ζωής καθώς είναι σπάνιο ένα συμβόλαιο να συναφθεί μόλις γεννηθεί ένα παιδί. Έστω λοιπόν *(x)* που ασφαλίζεται με συμβόλαιο ζωής. Έχει ήδη επιβιώσει τα πρώτα x χρόνια της ζωής του και θα εξετάσουμε την πιθανότητα να ζήσει αλλά t έτη.

Ορίζουμε λοιπόν την τυχαία μεταβλητή.

$$T(x) = ο\ υπολειπόμενος\ χρόνος\ ζωής\ ατόμου\ ηλικίας\ x$$

Δηλαδή X συνολικός χρόνος που θα ζήσει και x τα χρόνια που έχει ήδη ζήσει είναι

$$T(x) = X - x \mid X > x$$

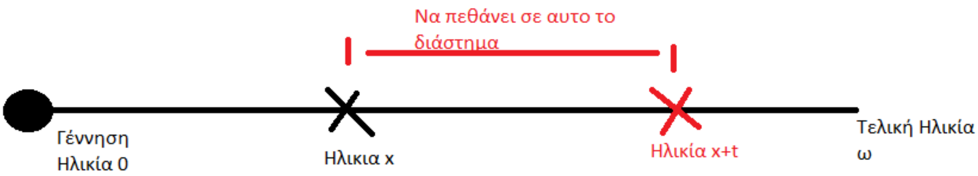
Δεδομένου ότι έχει ζήσει x ποια είναι η πιθανότητα να επιβιώσει έως την ηλικία X .

Στη αναλογιστική επιστήμη η δεσμευμένη πιθανότητα, η **συνάρτηση κατανομής** της τυχαίας μεταβλητής *T(x)* συμβολίζεται ως ${}_{t}q_x$

Και είναι η πιθανότητα

$${}_{t}q_x = P(T \leq x + t \mid T > x) = \frac{P(x < T \leq x + t)}{P(T > x)} = \frac{F(x+t) - F(x)}{1 - F(x)}$$

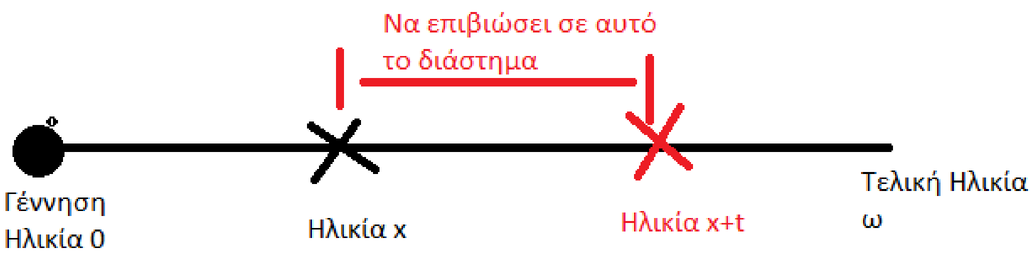
Η παραπάνω πιθανότητα είναι η πιθανότητα ένα άτομο ηλικίας x να πεθάνει μέσα στα επόμενα t χρόνια. Δηλαδή να μην φτάσει στη ηλικία x+t. Πιο περιγραφικά θα φανεί στο παρακάτω σχήμα



Και για την **συνάρτηση επιβίωσης** της T έχουμε τον αναλογιστικό συμβολισμό ${}_{t}p_x$ και την αναλυτική περιγραφή ως :

$$ {}_{t}p_x = P(T > x + t \mid T > x\ ) = = \frac{P(T > x + t)}{P(T > t)} == \frac{S(x + t)}{S(x)} $$

Η παραπάνω πιθανότητα, είναι η πιθανότητα του ενδεχομένου *(x)* να ζήσει τουλάχιστον t ακόμα χρόνια ή ισοδύναμα να φτάσει την ηλικία *x+t*. Το οποίο περιγράφεται στο παρακάτω σχήμα



Σε αντιστοιχία της της προηγουμένης παραγράφου αφού ορίστηκαν οι συναρτήσεις κατανομής και επιβίωσης για το την τυχαία μεταβλητή T θα ορίσουμε και τις **συναρτήσεις πυκνότητας πιθανότητας** και **έντασης θνησιμότητας.** Αρχικά για την συνάρτηση πυκνότητας της τυχαίας μεταβλητής T θα έχουμε

$$ f_t = \frac{d}{dt}\, {}_{t}q_x = \frac{d}{dt} \frac{F_{x+t} - F_x}{S_x} = \frac{f_{x+t}}{S_x} $$

Ενώ λόγω της

$$ S_x = \frac{S_{x+t}}{S_t} $$

Παίρνουμε

$$f_t = \frac{f_{x+t}}{S_x} = \frac{f_{x+t}}{S_{x+t}} S_t$$

Και από την

$$\mu_x = \frac{f_x}{S_x}$$

Καταλήγουμε

$$f_t = {}_t p_x \, \mu_{x+t}$$

Υπενθυμίζουμε ότι σε όλες τις παραπάνω εξισώσεις χρησιμοποιούμε τους αναλογιστικούς συμβολισμούς βάση του **Σχολίου 3** της προηγούμενης παραγράφου .

Τώρα θα μελετήσουμε την ένταση της θνησιμότητας της τυχαίας μεταβλητής T, άρα δεδομένου ότι έχει ζήσει *x+t* πλέον χρόνια να πεθάνει σε ένα μικρό χρονικό διάστημα dx μετά. Δηλαδή έχουμε

$$\mu_{x+t} = \frac{f_t}{{}_t p_x} = \lim_{dt \to 0} \frac{P(x + t \le T \le x + t + dt \mid T > x + t)}{dt}$$

Σχόλιο 1 : Όταν γράφουμε ${}_t p_x$ μιλάμε για την πιθανότητα *(x)* να ζήσει για t ακόμα χρόνια. Δηλαδή να φτάσει στη ηλικία *x+t*. Στη ειδική περίπτωση που ισχύει *t=1* γράφουμε p_x και εκφράζει την πιθανότητα *(x)* να φτάσει στη ηλικία *x+1* . Δηλαδή αν ζήσει ακόμα ένα έτος. Κατά αντιστοιχία αντί ${}_t q_x$ για *t=1* θα γράφουμε q_x και θα αναφερόμαστε στη πιθανότητα *(x)* να μην φτάσει στη ηλικία *x+1.* Δηλαδή να πεθάνει στο επόμενος έτος.

4.1.3 Πίνακες Θνησιμότητάς

Οι **πίνακες θνησιμότητες** είναι το σημαντικότερο εργαλείο για την εκτίμηση της πιθανότητας επιβίωσης / αποβίωσης για *(x)* . Η λογική τους είναι αρκετά απλή. Ξεκινάμε με ένα αυθαίρετο πλήθος νεογέννητων (συνήθως επιλέγεται κάποια δύναμή του 10 πχ 10^4 ή 10^5) το όποιος συμβολίζεται με l_0 και λέγεται **βάση** ή **ρίζα** *(radix)* . Παρατηρούμε λοιπόν μια γενιά *(cohort)* από l_0 νεογέννητα άτομα με τις υποθέσεις ότι δεν υπάρχουν άλλες γεννήσεις - είσοδοι ούτε αποχωρίσεις εκτός από το θάνατο.

Θεωρούμε λοιπόν τη τυχαία μεταβλητή X_j με $j \in \{ 1,2, \dots , l_0 \}$την ηλικία θανάτου ενός νεογέννητου ατόμου από την ρίζα του πίνακα.

Θα ορίσουμε ακόμα την εξής δείκτρια

$$I_j \begin{cases} 1, & \text{αν το ατόμο } j \text{ φθάσει στη ηλικία } x \\ 0, & \text{αλλίως} \end{cases} \Leftrightarrow I_j \begin{cases} 1, & X_j > x \\ 0, & \text{αλλίως} \end{cases}$$

Η κατανομή της I_j λόγω του ορισμού της θα είναι

$P\left(I_j = 1\right) = P\left(X_j > x\right) = S_x$ και $P\left(I_J = 0\right) = P(X \leq x) = 1 - S_x$

Ενώ για την μέση τιμή της θα έχουμε

$$E(X) = 1\, P\left(I_j = 1\right) + 0\, P\left(I_j = 1\right) = S_x$$

Το συνολικό πλήθος που θα βρίσκεται εν ζώη στη ηλικία x από το αρχικό πληθυσμό l_0 θα είναι

$$\mathcal{L}(x) = I_1 + I_2 + \cdots + I_{l_0} \Longrightarrow E\left(\mathcal{L}(x)\right) = E\left(I_1 + I_2 + \cdots + I_{l_0}\right) \Leftrightarrow$$

$$E\left(\mathcal{L}(x)\right) = E(I_1) + E(I_2) + \cdots + E\left(I_{l_0}\right) \Leftrightarrow E\left(\mathcal{L}(x)\right) = l_0 S_x$$

Η μέση τιμή $E\left(\mathcal{L}(x)\right)$ είναι ο αναμενόμενος αριθμός ατόμων της ομάδας που φθάνουν την ηλικία x από αρχικό πληθυσμό l_0 και την συμβολίζουμε με l_x.

Δηλαδή ισχύει

$$E\left(\mathcal{L}(x)\right) \coloneqq l_x$$

Επίσης

$$ {}_{t}p_x = \frac{S(x+t)}{S(x)} = \frac{\frac{l_{x+t}}{l_0}}{\frac{l_x}{l_0}} \Rightarrow {}_{t}p_x = \frac{l_{x+t}}{l_x}$$

Ενώ για την συνάρτηση κατανομής θα έχουμε

$${}_{t}q_x = 1 - {}_{t}p_x = 1 - \frac{l_{x+t}}{l_x} \Rightarrow {}_{t}q_x = \frac{l_x - l_{x+t}}{l_x}$$

Ο αριθμητής στο τελευταίο κλάσμα είναι η διαφορά του αναμενόμενου αριθμού που θα βρίσκεται στη ζωή στη ηλικία x με το αντίστοιχο αριθμό που θα βρίσκεται στη ζωή στη ηλικία $x+t$. Ισοδύναμα είναι ο αναμενόμενος αριθμός θανόντων μεταξύ των ηλικιών x και $x+t$. Ο αναλογιστικός συμβολισμός αυτού του αριθμού είναι ${}_{t}d_x$.

Δηλαδή

$${}_{t}d_x = l_x - l_{x+t}$$

Μέσω των εξισώσεων ${}_{t}p_x = \frac{l_{x+t}}{l_x}$ και ${}_{t}q_x = \frac{l_x - l_{x+t}}{l_x}$ παραπάνω αν μας δοθούν τα l_x και l_{x+t}

Για να γίνει πιο κατανοητή η χρήση ενός πίνακα θνησιμότητας / ζωής θα δώσουμε ένα παράδειγμα.

Παράδειγμα 1 : Δίνεται το παρακάτω τμήμα από ένα πίνακα θνησιμότητας. Να συμπληρωθούν οι γραμμές του

Ηλικία x	Αναμενόμενος αριθμός επιβιωσάντων l_x	Αναμενόμενος αριθμός θανόντων d_x	Πιθανότητα Θανάτου	Πιθανότητα Επιβίωσης
20	30,000	1,200		
21				
22	27,350			
23			0.07	
24	23,900		0.079	

Στόχος του παραδείγματος είναι να δούμε πως από το l_x και d_x μπορούμε να εκτιμήσουμε τις πιθανότητες ${}_{t}p_x$ και ${}_{t}q_x$ και ανάποδα αν είναι γνώστες οι ανωτέρω πιθανότητες πως μπορούμε να εκτιμήσουμε το πλήθος των ατόμων που θα είναι εν ζωή στη ηλικία x έχουμε λοιπόν

$$l_{21} = l_{20} - d_{20} = 30{,}000 - 1{,}200 = 28{,}000 \Leftrightarrow p_{20} = \frac{l_{21}}{l_{20}} = \frac{28{,}800}{30{,}000} = 0.96$$

$$q_{20} = 1 - p_{20} = 1 - 0.96 = 0.04$$

$$d_{21} = l_{21} - l_{22} = 28{,}800 - 27{,}350 = 1{,}450 \Leftrightarrow q_{21} = \frac{d_{21}}{l_{21}} = \frac{1{,}450}{28{,}800} = 0.05$$

Και λόγω την συμπληρωματικότητας των ενδεχομένων θανάτου – ζωής για την πιθανότητα επιβίωσης θα ισχύει

$$p_{21} = 1 - q_{21} = 1 - 0.05 = 0.95$$

Και ακόμα

$$p_{23} = 1 - q_{23} = 1 - 0.07 = 0.93$$

$$p_{24} = 1 - q_{24} = 1 - 0.079 = 0.921$$

Με την βοήθεια των παραπάνω πιθανοτήτων βρίσκουμε τους πληθυσμούς επιζήσαντων ως εξής

$$p_{23} = \frac{l_{24}}{l_{23}} \Leftrightarrow l_{23} = \frac{l_{24}}{p_{23}} = \frac{23{,}900}{0.93} = 25{,}698.92 \approx 25{,}699$$

Το l_{23} είναι η μέση τιμή μιας τυχαίας μεταβλητής της $\mathcal{L}(x)$ βέβαια επειδή συμβολίζει το αναμενόμενο πλήθος επιβιωσάντων στη ηλικία x συχνά θεωρούμε ότι $l_{23} \in \mathbb{N}$ και ειδικότερα όταν το αντικαθιστούμε σε ένα πίνακα θνησιμότητας.

Τώρα έχουμε για το d_{22} , d_{23} και d_{24}

$$d_{22} = l_{22} - l_{23} = 27.350 - 25{,}699 = 1{,}651$$

$$d_{23} = l_{23}q_{23} = 25{,}699 * 0.07 = 1{,}798.93 \approx 1{,}799$$

$$d_{24} = l_{23}q_{24} = 23{,}900 * 0.079 = 1{,}888.1 \approx 1{,}888$$

Και για να ολοκληρωθεί ο πίνακας θα υπολογίσουμε και τα p_{22} και q_{22} με το βασικό τους τύπο.

$$p_{22} = \frac{l_{23}}{l_{22}} = \frac{25{,}699}{27{,}350} = 0.94$$

$$q_{22} = \frac{d_{22}}{l_{22}} = \frac{1{,}651}{27{,}350} = 0.06$$

Ηλικία x	Αναμενόμενος αριθμός επιβιωσάντων l_x	Αναμενόμενος αριθμός θανόντων d_x	Πιθανότητα Θανάτου	Πιθανότητα Επιβίωσης
20	30,000	1,200	0.04	0.96
21	28,800	1,450	0,05	0.95
22	27,350	1,651	0.06	0.94
23	25,699	1,799	0.07	0.93
24	23,900	1,888	0.079	0.921

4.1.4 Συναρτήσεις θνησιμότητας για κλασματικές ηλικίες

Οι πίνακες θνησιμότητας είναι ένα καλό εργαλείο για την εκτίμηση των πιθανοτήτων θανάτου και επιβίωσης ενός ατόμου. Βέβαια οι πίνακες δύναται να εκτιμήσουν τις εν λόγω πιθανότητες στις ακέραιες ηλικίες αλλά δεν μπορούμε να εκτιμήσουμε την πιθανότητα ένα άτομο ηλικία 30 χρόνων να πεθάνει μέσα στο επόμενο εξάμηνο ή την πιθανότητα ένα άτομο 30 χρονών και 3 μηνών να ζήσει μέχρι τα τριακοστά πρώτα γενέθλια του.

Για το λόγο αυτό θα χρησιμοποιήσουμε κάποιες τεχνικές αριθμητικής ανάλυσης. Θα κάνουμε 3 υποθέσεις για την S_x ή ισοδύναμα l_x στο διάστημα $[x, x+1]$ δηλαδή θα εκτιμήσουμε την S_{x+t} και l_{x+t} για $t \in (0,1)$

4.1.4.1 Γραμμική παρεμβολή (UDD μέθοδος)

Η πρώτη μέθοδος θεωρεί ότι η $S_x = \frac{l_x}{l_0}$ ή ισοδύναμα η l_x είναι γραμμική στο διάστημα $[x, x+1]$. Η ανωτέρω υπόθεση ισοδυναμεί με την υπόθεση ότι η l_x είναι γραμμική του συνάρτηση του t στο $[x, x+1]$. Δηλαδή

$$l_{x+t} = a + bt \quad a, b \in \mathbb{R} \; και \; t \in (0,1)$$

Τότε έχουμε για $t = 0$ ότι $l_x = a$ ενώ για $t = 1$ είναι

$$l_{x+1} = a + b \Leftrightarrow l_{x+1} = l_x + b \Leftrightarrow b = l_{x+1} - l_x \Leftrightarrow b = -d_x$$

Άρα η αρχική μας εξίσωση γίνεται

$$l_{x+t} = l_x - d_x t$$

Η τελευταία εξίσωση ότι από το αρχικό πληθυσμό l_x στη αρχή του έτος στη ηλικία xθα έχει αποβιώσει $d_x t$ στο t τμήμα του έτος. Ανάλογα με το ποσοστό του χρόνου που θα έχει περάσει θα έχει αποβιώσει το ίδιο ποσοστό από το σύνολο των d_x θανάτων μέσα στο έτος. Καθώς υπάρχει αυτή η αναλογία μεταξύ του τμήματος του έτους που έχει εκπνεύσει και το ποσοστό των θανάτων που έχει πραγματοποιηθεί η γραμμική παρεμβολή, γνωστή και ως **ομοιόμορφη κατανομή των θανάτων** *(uniform distribution of deaths ή UDD),* τείνει να προτιμάται έναντι όλων των υπολοίπων μεθόδων στις σύγχρονες εφαρμογές.

Θα δούμε τώρα κάτω από αύτη τη υπόθεση πως διαμορφώνονται οι πιθανότητες θανάτου / επιβίωσης που ήταν και ο αρχικός μας στόχος. Αν διαιρέσουμε με l_x τα δύο μέλη της τελευταίας εξίσωσης έχουμε.

$$l_{x+t} = l_x - d_x t \Leftrightarrow \frac{l_{x+t}}{l_x} = 1 - \frac{d_x}{l_x} t = {}_t p_x = 1 - t q_x$$

και λόγω συμπληρωματικότητας έχουμε

$${}_t q_x = 1 - {}_t p_x \Leftrightarrow {}_t q_x = 1 - 1 + q_x t \Leftrightarrow {}_t q_x = t q_x$$

Για την συνάρτηση πυκνότητας κάτω από τη υπόθεση της ομοιόμορφής κατανομής των θανάτων θα έχουμε

$$f_t = \frac{d}{dt} {}_t q_x = \frac{d}{dt} t q_x = q_x$$

Αρά η f_t είναι ανεξάρτητη του t, άρα είναι σταθερή. Άρα η T ακολουθεί την ομοιόμορφη κατανομή στο *(0,1)*.

Τέλος για την ένταση θνησιμότητας θα έχουμε

$$f_t = {}_tp_x\, \mu_{x+t} \Leftrightarrow \mu_{x+t} = \frac{f_t}{{}_tp_x} = \frac{q_x}{1-q_x}$$

Η παραπάνω μορφή της έντασης θνησιμότητας φανερώνει ότι είναι **αύξουσα συνάρτηση του χρόνου**. Το οποίο δεν είναι πάντοτε επιθυμητή ιδιότητα.

4.1.4.2 Εκθετική Παρεμβολή (CFM μέθοδος)

Η δεύτερη υπόθεση για την μορφή της l_x ως συνεχή συνάρτηση στο $[x, x+1]$ είναι ότι είναι εκθετική δηλαδή

$$l_{x+t} = ab^t$$

Τότε έχουμε για $t = 0$ ότι $l_x = a$ ενώ για $t = 1$ είναι

$$l_{x+1} = ab \Leftrightarrow l_{x+1} = l_x b \Leftrightarrow b = \frac{l_{x+1}}{l_x} \Leftrightarrow b = p_x$$

Άρα η αρχική εξισώση γίνεται

$$l_{x+t} = l_x p_x^t$$

Και με διαίρεση και στα δύο μέλη με l_x γίνεται

$$\frac{l_{x+t}}{l_x} = p_x^t \Leftrightarrow {}_tp_x = p_x^t$$

Και για την πιθανότητα θανάτου θα έχουμε

$${}_tq_x = 1 - {}_tp_x = 1 - p_x^t = 1 - (1-q_x)^t$$

Διατηρώντας την συλλογιστική πορεία των προηγούμενων παραγράφων για την συνάρτηση πυκνότητας κάτω από την συγκεκριμένη υπόθεση θα έχουμε

$$f_t = \frac{d}{dt}\,{}_tq_x = \frac{d}{dt}(1 - p_x^t) \Rightarrow f_t = \frac{d}{dt}p_x^t \Leftrightarrow f_t = -\ln p_x\, p_x^t$$

Και για την ένταση θνησιμότητας θα έχουμε

$$f_t = {}_tp_x\, \mu_{x+t} \Leftrightarrow \mu_{x+t} = \frac{f_t}{{}_tp_x} \Leftrightarrow \frac{-\ln p_x\, p_x^t}{p_x^t} \Leftrightarrow \mu_{x+t} = -\ln p_x$$

Αρά κάτω από την δεύτερη μέθοδο η ένταση θνησιμότητας **είναι σταθερή μέσα σε κάθε έτος ηλικίας** *(constant force of mortality ή CFM)*

4.1.4.3 Αρμονική παρεμβολή (Balducci μέθοδος)

Η τελευταία από τις μεθόδους που θα μελετήσουμε, που είναι γνωστή ως υπόθεση Balducci, υποθέτει ότι η αντιστροφή συνάρτηση της συνάρτησης επιβίωσης της απομένουσας ζωής *(x)* είναι γραμμική συνάρτηση του $t \in (0,1)$

Έτσι έχουμε

$$\frac{1}{S_{x+t}} = (1-t)\frac{1}{S_x} + t\frac{1}{S_{x+1}}$$

Ή ισοδύναμα

$$\frac{1}{l_{x+t}} = (1-t)\frac{1}{l_x} + t\frac{1}{l_{x+1}} = \frac{1}{l_x} + \left(\frac{1}{l_{x+1}} - \frac{1}{l_x}\right)t$$

Πολλαπλασιάζοντας με l_x η παραπάνω εξίσωση γίνεται

$$\frac{1}{{}_tp_x} = 1 + \left(\frac{1}{p_x} - 1\right)t = \frac{p_x + tq_x}{p_x} \Leftrightarrow {}_tp_x = \frac{p_x}{p_x + tq_x}$$

Και επομένως για την συνάρτηση κατανομής θα έχουμε

$${}_tq_x = \frac{tq_x}{p_x + tq_x}$$

Επίσης από την γνωστή σχέση

$$f_t = \frac{d}{dt}\,{}_tq_x = \frac{p_xq_x}{(p_x + tq_x)^2}$$

Ενώ για την ένταση θνησιμότητας θα έχουμε

$$f_t = {}_tp_x\mu_{x+t} \Leftrightarrow \frac{p_xq_x}{(p_x + tq_x)^2} = \frac{p_x}{p_x + tq_x}\,\mu_{x+t} \Leftrightarrow \mu_{x+t} = \frac{q_x}{p_x + tq_x}$$

Η τελευταία ισότητα μας δείχνει ότι με την εν λόγω υπόθεση η ένταση θνησιμότητας είναι **φθίνουσα συνάρτηση του t**.

Συνοψίζοντας τα παραπάνω σε ένα πίνακα θα έχουμε

Πίνακας 4.2

Συναρτήσεις	Υπόθεση UDD	Υπόθεση CFM	Υπόθεση Balducci
${}_{t}q_x$	tq_x	$1-p_x^t$	$\frac{tq_x}{p_x+tq_x}$
${}_{t}p_x$	$1-tq_x$	p_x^t	$\frac{p_x}{p_x+tq_x}$
f_t	q_x	$-\ln p_x\, p_x^t$	$\frac{p_x q_x}{(p_x+tq_x)^2}$
μ_{x+t}	$\frac{q_x}{1-q_x}$	$-\ln p_x$	$\frac{q_x}{p_x+tq_x}$

4.2 Μοντέλα με πολλαπλά αίτια εξόδου

4.2.1 Εισαγωγή Ορισμοί

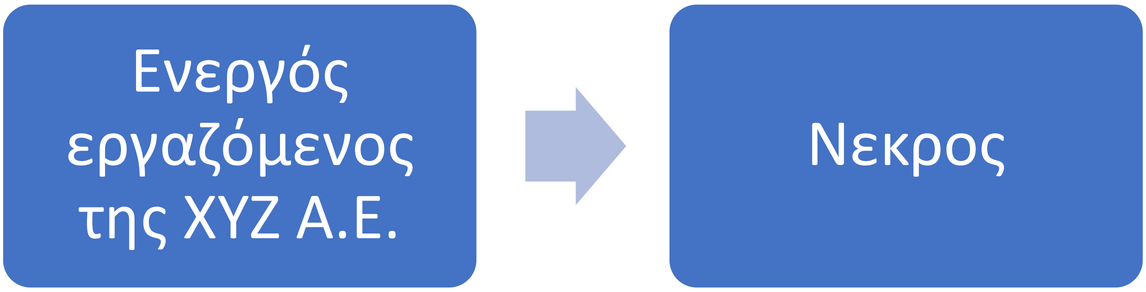
Μέχρι στιγμής έχουμε εξετάσει την έξοδο ενός ατόμου από το παρατηρούμενο πληθυσμό μόνο εξαιτίας του θανάτου του. Το μοντέλο αυτό ήταν αρκετά απλό με δύο πιθανές καταστάσεις «ζωντανός» και «νεκρός». Βέβαια σε πολλές αναλογιστικές μελέτες οι πιθανές καταστάσεις είναι περισσότερες, καθώς ένα άτομο μπορεί να αποχωρήσει για από το παρατηρούμενο πληθυσμό για διαφορετικούς λόγους. Σε αυτή τη παράγραφο θα μελετήσουμε *από κοινού τα συνδυασμένα αίτια εξόδου* του παρατηρούμενου πληθυσμού και να δημιουργήσουμε αναλογιστικές συναρτήσεις που θα επεκτείνουν τις συναρτήσεις που εξετάσαμε στη προηγουμένη παράγραφο ενός αιτίου εξόδου.

Για το σκοπό αυτό θα κάνουμε δύο **υποθέσεις**

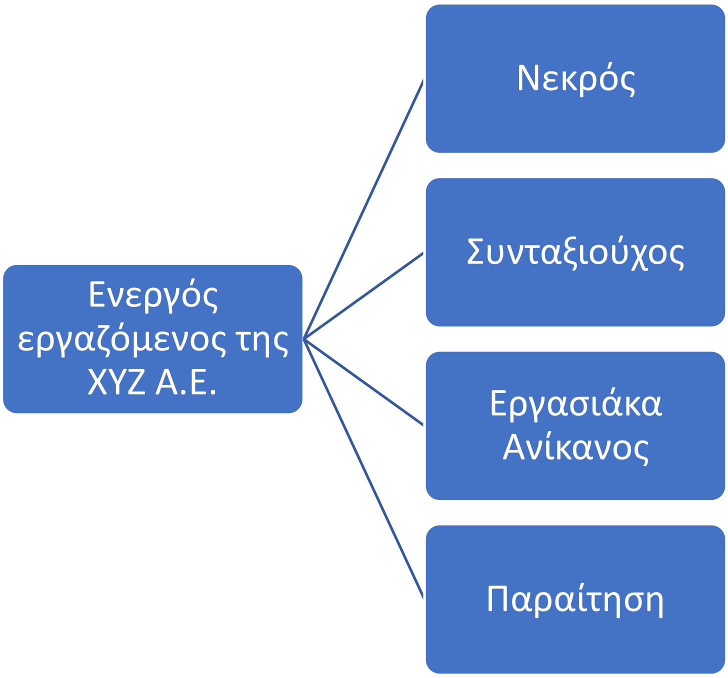
1. Κάθε έξοδος από το παρατηρούμενο πληθυσμό οφείλεται αποκλειστικά σε ένα.
2. Κάθε άτομο μέσα στο πληθυσμό μπορεί να αποχωρήσει από αυτόν για οποιοδήποτε αίτιο είναι διαθέσιμο στο πληθυσμό.

Αρχικά θα δούμε ένα παράδειγμα που θα αντιπαραβάλει όσα αναφέρθηκαν στη προηγουμένη παράγραφο με αυτά που έχουμε στόχο να εξετάσουμε στην παρούσα.

Παράδειγμα: έστω ότι ο παρατηρούμενος πληθυσμός είναι οι εργαζόμενοι της εταιρίας XYZ A.E. Με βάση όσα έχουν ειπωθεί *(x)* μπορεί να βρίσκεται σε δύο καταστάσεις



το οποίο είναι παράλογο καθώς για ένα άτομο υπάρχουν διαφορετικοί «θάνατοι» για την εργασιακή του ζωή όπως η συνταξιοδότηση ή η παραίτηση και φυσικά ο φυσικός θάνατος. Άρα το παρακάτω σχήμα είναι πιο αντιπροσωπευτικό



Αν θέλουμε να εξέτασουμε ένα συνταξιοδοτικό πρόγραμμα για τη XYZ θα πρέπει να λάβουμε υπόψη μας όλα τα παραπάνω *αίτια εξόδου* από το πληθυσμό για *(x)* .

Ορίζουμε λοιπόν την τυχαία μεταβλητή Τ ως το **υπολειπόμενο χρόνο που το *(x)*** θα παραμείνει στη κατάσταση ενεργός, εντός του εξεταζόμενου πληθυσμού.

Ακόμα ορίζουμε X_j την τυχαία μεταβλητή του υπολειπόμενου χρόνου ζωής από το *j* – αίτιο για *j=1,2,…m* (όταν έχουμε *m* αίτια εξόδου). Επομένως έχουμε $T = \min(X_1, X_2, \dots X_m)$. Συμφώνα με το παραπάνω το T είναι ο μόνος χρόνος που παρατηρούμε.

Η τελευταία τυχαία μεταβλητή που θα ορίσουμε είναι η *J*. Η *J* είναι μια διακριτή τυχαία μεταβλητή ή οποία παίρνει τιμές στο σύνολο *{1,2,3, … , m}* (όταν έχουμε *m* αίτια εξόδου) και παίρνει την τιμή *j* τότε και μόνο τότε αν το άτομο αποχωρήσει από το *j* αίτιο εξόδου. Έτσι

$$J = \begin{cases} j, & \alpha\nu\ T = X_j \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

4.2.2 Η συνάρτηση επιβίωσης της τυχαίας μεταβλητής T

Η **από κοινού (πολυμεταβλητή) συνάρτηση επιβίωσης** ορίζεται ως

$$S_{X_1,X_2,\ldots,X_m}(x_1,x_2,\ldots,x_m) = P\left\{\bigcap_{j=1}^{m}\{(X_j > x_j)\}\right\}$$

Και έτσι μπορούμε να ορίσουμε και τις **περιθώριες συναρτήσεις επιβίωσης** ως

$$S_{X_j}(x_j) = S_{X_1,X_2,\ldots,X_m}(0,\ldots,x_j,\ldots,0)\ \gamma\iota\alpha\ j = 1,2,\ldots,m$$

Με βάση τη υπόθεση 2 στη αρχή του κεφαλαίου αυτού μπορούμε να διαπιστώσουμε ότι $S_{X_j}(x_j)$ είναι συναρτήσεις επιβίωσης με βάση τον ορισμό στης παράγραφο 3.1.

Καθώς είναι αδύνατο να παρατηρήσουμε το τυχαίο διάνυσμα $(X_1,X_2,\ldots,X_m)$ από κοινού θα παρατηρήσουμε το T που είναι ο χρόνος τερματισμού του της αρχικής μας κατάστασης από οποιοδήποτε αίτιο. Έτσι έχουμε

$$S_T = P(T > t) = P\left\{\bigcap_{j=1}^{m}\{(X_j > x_j)\}\right\} = S_{X_1,X_2,\ldots,X_m}(t,t,\ldots,t)$$

Η συνάρτηση αυτή είναι η **ολική συνάρτηση επιβίωσης από όλες τις m- αίτιες**

4.2.3 Η από κοινού συνάρτηση μάζας πιθανότητας των J και T

Θα ορίσουμε την συνάρτηση μάζας πιθανότητας της τυχαίας μεταβλητής (T,J) ως $f_{T,J}(t,j)$ με πεδίο ορισμού το σύνολο

$$\Re_{T,J} \subseteq [0,\infty] \times \{1,2,\ldots,m\}$$

Αρχικά λόγω της φύσης των τυχαίων μεταβλητών T,J καθώς είναι καλά ορισμένες τυχαίες μεταβλητές θα ισχύει

$$\sum_{j=1}^{m} f_J(j) = 1\ \kappa\alpha\iota\ \int_0^{\infty} f_T(t)dt = 1$$

Επομένως και η $f_{T,J}(t,j)$ θα ικανοποιεί τη σχέση

$$\sum_{j=1}^{m}\int_0^{\infty} f_{T,J}(t,j)dt = 1$$

Η ποσότητα $f_{T,J}(t,j)dt \cong P(t < T < t+dt, J=j)$ σε αντιστοιχία με το $f_T(t)dt \cong P(t < T < t+dt)$ εκφράζει (προσεγγιστικά) τη πιθανότητα ένα άτομο ηλικία x να αποχωρήσει από το σύστημα στο διάστημα $(t,t+dt]$ εξαιτίας του αιτίου j

4.2.4 Η περιθώριες κατανομές των τυχαίων μεταβλητων *T, J*

Για να βρούμε τις περιθώριες κατανομές από θεωρία των πιθανοτήτων, είναι γνωστό ότι σταθεροποιούμε την μια (ή τις περισσότερες αν υπάρχουν) τυχαία μεταβλητή σε μια τιμή και αφήνουμε την άλλη να «τρέξει» σε όλο το πεδίο ορισμού της.

Έτσι θα υιοθετήσουμε και σε αυτό το σημείο τους αναλογιστικούς συμβολισμούς ώστε να είναι σε αντιστοιχία με το προηγουμένη παράγραφο.

Έχουμε λοιπόν ότι η **περιθώρια συνάρτηση πιθανότητας** της J είναι

$$ {}_{\infty}q_x^{(j)} := f_J(j) = \int_0^{\infty} f_{T,J}(t,j)dt \ , j = 1,2,\dots,m $$

Και μας δίνει τη πιθανότητα ένος *(x)* που βρίσκονταν σε μια αρχική κατάσταση να αποχωρήσει από αύτη λόγω του αιτίου j.

Λόγω της φύσης τη J δεν υπάρχει λόγος να ορίσουμε συνάρτηση κατανομής καθώς δεν η J είναι μια κατηγορική τυχαία μεταβλητή χωρίς ιεραρχία άρα η πιθανότητα

$$ P(J \leq j) $$

ως ορισμός συνάρτησης κατανομής δεν έχει νόημα αφού η αιτία εξόδου δεν έχει ευθεία αντιστοίχιση με αριθμό που τις έχει αποδοθεί.

Αντίθετα για τη **περιθώρια συνάρτηση κατανομή** της T θα έχουμε

$$ {}_{t}q_x^{(\tau)} = P(T \leq t) = \int_0^{t} f_T(s)\,ds = \sum_{j=1}^{m} \int_0^{t} f_{T,J}(s,j)ds $$

Η ποσότητα ${}_{\infty}q_x^{(\tau)}$ δηλώνει την πιθανότητα εξόδου *(x)* από την κατάσταση που βρισκόταν αρχικά πριν από τη χρονική στιγμή t ανεξαρτήτως αιτίου εξόδου. Για την **συνάρτηση επιβίωσης** της τ.μ. Τ έχουμε το συμβολισμό.

Σε αντίθεση η

$$ {}_{t}p_x^{(\tau)} = P(T > t) = 1 - {}_{t}q_x^{(\tau)} $$

Ενώ για την **συνολική ένταση εξόδου** (σε αντιστοιχία με τη ένταση θνησιμότητας) θα έχουμε

$$ \mu_{x+t}^{(\tau)} = \frac{f_t}{1-F_t} = \frac{1}{{}_{t}p_x^{(\tau)}} \frac{d}{dt} {}_{t}q_x^{(\tau)} = -\frac{1}{{}_{t}p_x^{(\tau)}} \frac{d}{dt} {}_{t}p_x^{(\tau)} = -\frac{d}{dt} \ln {}_{t}p_x^{(\tau)} $$

άρα συνεπάγεται

$$ {}_{t}p_x^{(\tau)} = \exp\left(-\int_0^{t} \mu_{x+s}^{(\tau)} ds\right) $$

Τέλος ορίζουμε την $\mu_{x+t}^{(j)}$ ως την ένταση εξόδου στη ηλικία $x+t$ λόγω του αιτίου j και θα ισχύει

$$\mu_{x+t}^{(j)} = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t, J = j \mid T > t)}{\Delta t} = \frac{f_{T,J}(t,j)}{{}_{t}p_x^{(\tau)}}$$

Ένας άλλος τρόπος να γραφεί η συνολική ένταση θνησιμότητας είναι ως άθροισμα των επιμέρους εντάσεων.

Δηλαδή

$$\mu_{x+t}^{(\tau)} = \sum_{j=1}^{m} \mu_{x+t}^{(j)} = \mu_{x+t}^{(1)} + \mu_{x+t}^{(2)} + \cdots + \mu_{x+t}^{(m)}$$

Το οποίο μας δίνει και άλλο ένα χρήσιμο αποτέλεσμα για τη πιθανότητα παραμονής στη αρχική κατάσταση. Καθώς έχουμε

$${}_{t}p_x^{(\tau)} = \exp\left(-\int_0^t \mu_{x+s}^{(\tau)} ds\right) = \exp\left(-\int_0^t \sum_{j=1}^{m} \mu_{x+t}^{(j)}\ ds\right) =$$

$$\exp\left(-\sum_{j=1}^{m} \int_0^t \mu_{x+t}^{(j)}\, ds\right) =$$

$$\exp\left(-\int_0^t \mu_{x+s}^{(1)} ds\right) \exp\left(-\int_0^t \mu_{x+s}^{(2)} ds\right) \ldots \exp\left(-\int_0^t \mu_{x+s}^{(m)} ds\right)$$

Και τέλος για την από κοινού συνάρτηση πυκνότητας πιθανότητας θα ισχύει

$$f_{T,J}(t,j) = {}_{t}p_x^{(\tau)}\, \mu_{x+t}^{(j)}$$

Στα συνέχεια θα εστιάσουμε στις εξής συναρτήσεις ${}_{t}p_x^{(\tau)}$, ${}_{t}q_x^{(\tau)}$, ${}_{t}q_x^{(j)}$, $\mu_{x+t}^{(\tau)}$ και $\mu_{x+t}^{(j)}$

4.2.5 Πίνακες με πολλαπλά αίτια εξόδου

Θεωρούμε ότι έχουμε σε μια αρχική ηλικία α , $l_{\alpha}^{(\tau)}$ πλήθος ατόμων σε μια κατάσταση. Για κάθε άτομο υποθέτουμε ότι η από κοινού κατανομή του χρόνου ως τη έξοδο και το αίτιο της εξόδου προσδιορίζεται απότ τη συνάρτηση μάζας πιθανότητας

$f_{T,J}(t,j) = {}_{t}p_x^{(\tau)}\, \mu_{x+t}^{(j)}$ για t ≥ 0 και $j = 1,2, \ldots, m$. Κατά αναλογία με του πίνακες θνησιμότητας θα ορίσουμε την δείκτρια

$$I_i^{(j)} = \begin{cases} 1, & \text{αν το } i \text{ άτομο αποχωρήσει από αυτή για } J = j \text{ και } x < T \leq x + t \\ 0, & \text{αλλιώς} \end{cases}$$

Όπου παίρνει την τιμή 1 αν και μόνο αν το άτομο i της ομάδας αποχωρήσει από αυτή λόγω του αιτίου εξόδου J μεταξύ των ηλικιών x και $x+n$

Έστω ${}_tD_x^{(j)}$ το πλήθος των ατόμων που αποχώρησαν από το την αρχική κατάσταση μεταξύ των ηλικιών x και $x+t$ εξαιτίας του αιτίου j. Τότε καθώς

$${}_tD_x^{(j)} = I_1^{(j)} + I_2^{(j)} + \cdots I_{l_\alpha^{(\tau)}}^{(j)}$$

Από την φύση του και την δομή του το ${}_tD_x^{(j)}$ είναι τυχαία μεταβήτη. Θέτουμε λοιπόν

$${}_td_x^{(j)} = E\left({}_tD_x^{(j)}\right) = E\left(I_1^{(j)}\right) + E\left(I_2^{(j)}\right) + \cdots + E\left(I_{l_\alpha^{(\tau)}}^{(j)}\right) =$$

$$l_\alpha^{(\tau)} P\left(I_i^{(j)} = 1\right) = l_\alpha^{(\tau)} \int_{x-a}^{x+t-a} f_{T,J}(t,j)dt \Leftrightarrow$$

$${}_td_x^{(j)} = \int_{x-a}^{x+t-a} {}_tp_a^{(\tau)}\, \mu_{a+t}^{(j)} dt$$

Τώρα μπορούμε να ορίσουμε ${}_tD_x^{(\tau)}$ ως το πλήθος όσων αποχωρήσαν από τη αρχική κατάσταση για ένα από τα οποιαδήποτε m αίτια και αντίστοιχα η μέση τιμή της τελευταίας τυχαίας μεταβλητής θα είναι το ${}_td_x^{(\tau)}$. Μάλιστα θα ισχύει

$${}_tD_x^{(\tau)} = \sum_{j=1}^{m} {}_tD_x^{(j)}$$

Και

$${}_td_x^{(\tau)} = \sum_{j=1}^{m} {}_td_x^{(j)}$$

Τώρα θα ορίσουμε τις εξής δείκτριες για $i = \left\{1,2, \ldots\ ,\ l_\alpha^{(\tau)}\right\}$

$$I_i = \begin{cases} 1, & \text{αν το } i \text{ άτομο της ομαδάς φθάσει την ηλικία } x \\ 0, & \text{αλλιως} \end{cases}$$

ακολουθώντας την ίδια συλλογιστική πορεία με το πίνακα θνησιμότητας θα ορίσουμε $\mathcal{L}^{(\tau)}(x)$ το πλήθος των ατόμων που παραμένουν στην αρχική ομάδα στη ηλικία x ξεκινώντας από την ηλικία a . Τότε έχουμε

$$\mathcal{L}^{(\tau)}(x) = I_1 + I_2 + \cdots + I_{l_a^{(\tau)}}$$

Και η μέση τιμή της της παραπάνω τυχαίας μεταβλητής θα είναι

$$l_x^{(\tau)} = E\left(\mathcal{L}^{(\tau)}(x)\right) = E(I_1) + E(I_2) + \cdots + E\left(I_{l_a^{(\tau)}}\right) =$$

$$l_a^{(\tau)} P(I_i = 1) = l_a^{(\tau)}\,{}_{t}p_x^{(\tau)} \Leftrightarrow$$

$$l_x^{(\tau)} = l_a^{(\tau)}\,{}_{x-\alpha}p_x^{(\tau)} = l_a^{(\tau)}\,{}_{x-\alpha}p_x^{(\tau)}\,p_{x-1}^{(\tau)} = l_{x-1}^{(\tau)} p_{x-1}^{(\tau)}$$

$$\Leftrightarrow p_{x-1}^{(\tau)} = \frac{l_x^{(\tau)}}{l_{x-1}^{(\tau)}}$$

Και ακόμα καθώς το $d_x^{(j)}$ είναι η εκτίμηση για τα άτομα από το συνολικό πληθυσμό που θα αποχωρήσουν από την ομάδα λόγω του αιτίου j στη ηλικία x έχουμε

$$d_x^{(j)} = l_a^{(\tau)}\,{}_{x-\alpha}p_x^{(\tau)} q_x^{(j)} = l_x^{(\tau)} q_x^{(j)} \Leftrightarrow q_x^{(j)} = \frac{d_x^{(j)}}{l_x^{(\tau)}}$$

Η παραπάνω ισότητα μας δίνει ένα εύχρηστο τύπο για το υπολογισμό της αποχώρησης από τη αρχική ομάδα λόγω του αιτίου j το οποίο θα είναι το (εκτιμώμενο) πλήθος που θα αποχωρήσουν εξαιτίας αυτού του αιτίου στη ηλικία x προς όλο το πληθυσμό στη ίδια ηλικία

Όπως επίσης θα ισχύει

$$q_x^{(j)} = \frac{d_x^{(\tau)}}{l_x^{(\tau)}}$$

Οι δύο ισότητες μας δείχνουν και τη αναλογία με όσα μελετήθηκαν στη προηγούμενο κεφάλαιο. Εξάλλου οι πίνακες θνησιμότητας / ζωής είναι ειδική περίπτωση των πινάκων με πολλαπλά αίτια εξόδου.

Παράδειγμα 1 : Δίνεται ο παρακάτω πίνακας τριπλής εξόδου με αίτια:

1. Θάνατο
2. Αποχώρηση
3. Μετάθεση

από ένα καθεστώς που αντιπροσωπεύει τους εργαζομένους σε έναν οργανισμό.

Ηλικία x	Αριθμός εργαζομένων $l_x^{(\tau)}$	Θάνατοι $d_x^{(1)}$	Αποχωρήσεις $d_x^{(2)}$	Μεταθέσεις $d_x^{(3)}$
20	100,000	452	5,517	2,569
21	91,462	433	4,780	2,431
22	83,818	414	4,136	2,303
23	76,966	396	3,574	2,183

Η πιθανότητα ότι ένας εργαζόμενος ηλικίας 20 ετών θα είναι στο οργανισμό στα επόμενα τρία έτη θα είναι

$$ {}_{t}p_x^{(\tau)} = \frac{l_{x+t}^{(\tau)}}{l_x^{(\tau)}} \Rightarrow {}_{3}p_{20}^{(\tau)} = \frac{l_{23}^{(\tau)}}{l_{20}^{(\tau)}} = \frac{76{,}966}{100{,}000} \Leftrightarrow {}_{3}p_{20}^{(\tau)} = 0{,}76966 $$

Εναλλακτικά θα μπορούσαμε να πούμε ότι

$$ {}_{t}p_x^{(\tau)} = p_x^{(\tau)} p_{x+1}^{(\tau)} p_{x+2}^{(\tau)} \Rightarrow $$

$$ {}_{3}p_{20}^{(\tau)} = p_{20}^{(\tau)} p_{21}^{(\tau)} p_{22}^{(\tau)} = \frac{\cancel{91{,}462}}{100{,}000} \frac{\cancel{83{,}818}}{\cancel{91{,}462}} \frac{76{,}966}{\cancel{83{,}818}} $$

Η πιθανότητα ότι ένας εργαζόμενος ηλικίας 21 ετών δεν θα βρίσκεται στο οργανισμό μετά από ένα έτος είναι

$$ q_x^{(\tau)} = \frac{l_x^{(\tau)} - l_{x+1}^{(\tau)}}{l_x^{(\tau)}} \Rightarrow q_{21}^{(\tau)} = \frac{l_{21}^{(\tau)} - l_{22}^{(\tau)}}{l_{21}^{(\tau)}} \Leftrightarrow $$

$$ q_{21}^{(\tau)} = \frac{91{,}462 - 83{,}818}{91{,}462} \Leftrightarrow q_{21}^{(\tau)} = 0.08358 $$

Ενώ η πιθανότητα ένας εργαζόμενος ηλικίας 21 ετών να πάρει μετάθεση πριν την ηλικία των 23 ετών είναι

$$ {}_{t}q_x^{(j)} = \frac{{}_{t}d_x^{(j)}}{l_x^{(\tau)}} \Rightarrow {}_{2}q_{21}^{(3)} = \frac{{}_{2}d_{21}^{(3)}}{l_{21}^{(\tau)}} \Leftrightarrow $$

$$ {}_{2}q_{21}^{(3)} = \frac{d_{21}^{(3)} + d_{22}^{(3)}}{l_{21}^{(\tau)}} \Leftrightarrow {}_{2}q_{21}^{(3)} = \frac{2{,}431 + 2{,}302}{91{,}462} = 0.05175 $$

4.2.6 Συναφείς πίνακες με ένα αίτιο εξόδου

Στη προηγούμενη ενότητα αναφερθήκαμε στη πιθανότητα ${}_{t}q_x^{(j)}$ ως η *πιθανότητα εξόδου από τη αρχική κατάσταση ενός ατόμου ηλικίας x λόγω του j αιτίου εξόδου.* Βέβαια η πιθανότητα αυτή είναι μέσα στο πλαίσιο ότι υπάρχουν ακόμα *m-1* ακόμα αίτια εξόδου. Όλα αυτά τα αίτια λειτουργούν ταυτόχρονα και *ανταγωνιστικά*. Αν για παράδειγμα η ένταση εξόδου ενός αιτίου αυξηθεί απότομα αυτό θα οδηγήσει σε συρρίκνωση του πληθυσμού και σε μείωση της πιθανότητας αποχώρησης από άλλα αίτια. Ορίζουμε λοιπόν τη συνάρτηση χρησιμοποιώντας αποκλειστικά την ένταση εξόδου $\mu_{x+s}^{(j)}$

$$ {}_{t}p_x'^{(j)} = \exp\left(-\int_0^t \mu_{x+s}^{(j)} ds\right) $$

Και αντίστοιχα ${}_{t}q_x^{j}$ ως **καθαρή πιθανότητα εξόδου του αιτίου εξόδου** ***j*** *(net probability of decrement)* ως

$$ {}_{t}q_x^{j} = 1 - {}_{t}p_x^{j} $$

Μια βασική σχέση που ισχύει στη θεωρία των συναφών πινάκων με ένα αίτιο εξόδου είναι η

$$ {}_{t}p_x^{(\tau)} = \exp\left(-\int_0^t \mu_{x+s}^{(\tau)} ds\right) = \exp\left(-\int_0^t \sum_{j=1}^{m} \mu_{x+t}^{(j)}\ ds\right) = $$

$$ \exp\left(-\sum_{j=1}^{m} \int_0^t \mu_{x+t}^{(j)}\ ds\right) = $$

$$ \exp\left(-\int_0^t \mu_{x+s}^{(1)} ds\right) \exp\left(-\int_0^t \mu_{x+s}^{(2)} ds\right) \dots \exp\left(-\int_0^t \mu_{x+s}^{(m)} ds\right) = $$

$$ {}_{t}p_x^{1}\,{}_{t}p_x^{2} \dots {}_{t}p_x^{j} \Leftrightarrow {}_{t}p_x^{(\tau)} = \prod_{j=1}^{m} {}_{t}p_x^{j} $$

και ακόμα

$$ {}_{t}q_x^{(\tau)} = 1 - {}_{t}p_x^{(\tau)} = 1 - \prod_{j=1}^{m} {}_{t}p_x^{j} = 1 - \prod_{j=1}^{m} \left(1 - {}_{t}q_x^{j}\right) $$

Με βάση την σχέση

$$\frac{d}{dt}\,{}_tp_x^{j} = \frac{d}{dt}\exp\left(-\int_0^t \mu_{x+s}^{(j)}ds\right) = -{}_tp_x^{j}\mu_{x+s}^{(j)}$$

παίρνουμε

$$\int_0^1 d\left({}_tp_x^{j}\right) = -\int_0^1 {}_tp_x^{j}\mu_{x+s}^{(j)}dt \Leftrightarrow 1 - p_x^{j} = \int_0^1 {}_tp_x^{j}\mu_{x+s}^{(j)}dt \ \Leftrightarrow$$

$$q_x^{j} = \int_0^1 {}_tp_x^{j}\mu_{x+t}^{(j)}dt$$

Τέλος είναι σημαντικό να αποδείξουμε μια ανισότητα η οποία εξηγεί και τη διαφορετική οπτική που εξετάζουμε το ίδιο πρόβλημα τώρα

$${}_tp_x^{j} \geq {}_tp_x^{(\tau)} \Leftrightarrow$$

$${}_tp_x^{j}\mu_{x+t}^{(j)} \ \geq \ {}_tp_x^{(\tau)}\ \mu_{x+t}^{(j)} \Leftrightarrow$$

$$\int_0^1 {}_tp_x^{j}\mu_{x+t}^{(j)}dt \geq \int_0^1 {}_tp_x^{(\tau)}\ \mu_{x+t}^{(j)}dt \ \Leftrightarrow$$

$$q_x^{j} \ \geq \ q_x^{(j)}$$

Η παραπάνω ανισοτική σχέση είναι αναμενόμενη αφού η καθαρή πιθανότητα q_x^{j} έχει υπολογιστεί χρησιμοποιώντας **αποκλειστικά** τη επίδραση του αιτίου εξόδου j ως αίτιο αποχώρησης του ατόμου από την αρχική κατάσταση σε, αντιδιαστολή με την πιθανότητα $q_x^{(j)}$ που έχει υπολογιστεί με το να είναι το άτομο εκτεθειμένο σε όλα τα αίτια εξόδου.

4.2.7 Σταθερή ένταση εξόδου για κάθε αίτιο εξόδου.

Είναι σημαντικό να καταλάβουμε ότι αν είναι γνωστές οι καθαρές πιθανότητες q_x^{j} για κάθε αίτιο είναι εύκολο να υπολογιστούν οι $q_x^{(j)}$ και κατ' επέκταση να καταρτιστεί πίνακας με πολλαπλά αίτια εξόδου.

Υπενθυμίζουμε ότι q_x^{j} είναι η πιθανότητα αποχώρησης από το πληθυσμό εξαιτίας του j αιτίου αν υπάρχει μόνο το j αίτιο. Ουσιαστικά είναι η πιθανότητα «θανάτου» που μελετήσαμε στη προηγούμενη παράγραφό ως μοναδικό αίτιο εξόδου.

Θα υποθέσουμε λοιπόν ότι η ένταση θνησιμότητας **μέσα σε ένα έτος** για το j αίτιο και αργότερα για κάθε j είναι σταθερή , δηλαδή

$$\mu_{x+t}^{(j)} = \mu_x^{(j)}, \qquad 0 \leq t \leq 1$$

επίσης θα υποθέσουμε ότι και η συνολική ένταση θνησιμότητας είναι δηλαδή

$$\mu_{x+t}^{(\tau)} = \mu_x^{(\tau)}, \qquad 0 \le t \le 1$$

Έτσι έχουμε για $0 \le s \le 1$

$${}_{s}q_x^{(j)} = \int_0^s {}_{t}p_x^{(\tau)} \mu_{x+t}^{(j)} dt = \int_0^s {}_{t}p_x^{(\tau)} \frac{\mu_{x+t}^{(j)}}{\mu_{x+t}^{(\tau)}} \mu_{x+t}^{(\tau)} dt =$$

$$\frac{\mu_x^{(j)}}{\mu_x^{(\tau)}} \int_0^s {}_{t}p_x^{(\tau)}\, \mu_{x+t}^{(\tau)} dt = \frac{\mu_x^{(j)}}{\mu_x^{(\tau)}}\, {}_{s}q_x^{(\tau)} \Leftrightarrow {}_{s}q_x^{(j)} = \frac{\mu_x^{(j)}}{\mu_x^{(\tau)}}\, {}_{s}q_x^{(\tau)}$$

Η τελευταία ισότητα θα μπορούσε να ερμηνευθεί ως ότι ${}_{s}q_x^{(j)}$ είναι το πόσοστο της έντασης εξόδου που συμβάλει το j αίτιο στη συνολική πιθανότητα εξόδου ${}_{s}q_x^{(\tau)}$

Από την αρχή αυτής της παραγράφου είχαμε αποδείξει την ισότητα

$${}_{t}p_x^{(\tau)} = \exp\left(-\int_0^t \mu_{x+s}^{(\tau)} ds\right)$$

Και όταν $0 \le t \le 1$ είναι

$$p_x^{(\tau)} = \exp\left(-\int_0^1 \mu_{x+t}^{(\tau)} dt\right)$$

Η οποία λόγω της $\mu_{x+t}^{(\tau)} = \mu_x^{(\tau)}$ γίνεται

$$p_x^{(\tau)} = \exp\left(-\mu_x^{(\tau)}\right) \Leftrightarrow \mu_x^{(\tau)} = -\ln p_x^{(\tau)} \Leftrightarrow r\mu_x^{(\tau)} = -\ln\left(p_x^{(\tau)}\right)^r \Leftrightarrow$$

$$r\mu_x^{(\tau)} = \ln {}_{r}p_x^{(\tau)}, \qquad \gamma\iota\alpha\ 0 < r < 1$$

Με την ίδια συλλογιστική πορεία μπορούμε να αποδείξουμε και την

$$r\mu_x^{(j)} = \ln {}_{r}p_x^{j}$$

Και επομένως

$$\frac{\mu_x^{(j)}}{\mu_x^{(\tau)}} = \frac{\ln {}_{r}p_x^{j}}{\ln {}_{r}p_x^{(\tau)}}$$

Ενώ με την χρήση της ${}_{s}q_x^{(j)} = \frac{\mu_x^{(j)}}{\mu_x^{(\tau)}}\, {}_{s}q_x^{(\tau)}$πολλαπλασιάζοντας με ${}_{s}q_x^{(\tau)}$θα έχουμε

$${}_{s}q_x^{(j)} = \frac{\ln {}_{r}p_x^{j}}{\ln {}_{r}p_x^{(\tau)}}\, {}_{s}q_x^{(\tau)} \Leftrightarrow \ln {}_{r}p_x^{j} = \frac{{}_{s}q_x^{(j)}}{{}_{s}q_x^{(\tau)}} \ln {}_{r}p_x^{(\tau)} \Leftrightarrow$$

$${}_{r}p_x^{j} = \left(p_x^{(\tau)}\right)^{\frac{{}_{s}q_x^{(j)}}{{}_{s}q_x^{(\tau)}}}$$

αν $r \longrightarrow 1$ τότε προκύπτει άμεσα ότι

$$q_x^{j} = 1 - \left(p_x^{(\tau)}\right)^{\frac{{}_{s}q_x^{(j)}}{{}_{s}q_x^{(\tau)}}}$$

Ενώ αν επιπροσθέτως ισχύει και $s \longrightarrow 1$

$$q_x^{j} = 1 - \left(p_x^{j}\right)^{\frac{q_x^{(j)}}{q_x^{(\tau)}}}$$

και

$$q_x^{(j)} = q_x^{(\tau)} \frac{\ln\left(1 - q_x^{j}\right)}{\ln p_x^{(\tau)}}$$

Έτσι αν γνωρίζουμε των ποσότητων q_x^{j} για $j = 1,2, \dots, m$ τότε χρησιμοποιώντας τη σχέση

$${}_{t}p_x^{(\tau)} = \prod_{j=1}^{m} {}_{t}p_x^{j} = {}_{t}p_x^{1}\, {}_{t}p_x^{2} \dots {}_{t}p_x^{m}$$

Μπορούμε ευκολά να υπολογίσουμε τις ποσότητες ${}_{t}q_x^{(j)}$ για $j = 1,2, \dots, m$.

5 Άθροισμα ανεξάρτητών τυχαίων μεταβλητών – Μοντέλο του Συλλογικού Κινδύνου

Στο παρόν κεφάλαιο θα εξετάσουμε την κατανομή μιας τυχαίας μεταβλητής, S η οποία είναι άθροισμα άλλων γνωστών ανεξάρτητων μεταβλητών X_i. Στο πρώτο μέρος του κεφαλαίου το πλήθος των προσθετέων X_i θα είναι μια γνωστή σταθερά n ενώ στο δεύτερο μέρος θα είναι μια τυχαία διακριτή τυχαία μεταβλητή N ανεξάρτητη των X_i.

5.1 Άθροισμα ανεξάρτητών τυχαίων μεταβλητών

5.1.1 Μέθοδος των γεννητριών συναρτήσεων

Αρχικά **ροπογεννήτρια** μια τυχαίας μεταβλητής X ονομάζεται η συνάρτηση, αν υπάρχει η μέση τιμή της συνάρτησης e^{tX} . Δηλαδή

$$M_X(t) = E(e^{tX})$$

Η οποία, υπάρχει όταν η παραπάνω μέση τιμή δεν απειρίζεται, και είναι συνάρτηση της πραγματικής μεταβλητής t. Όταν η τυχαία μεταβλητή X είναι **απόλυτα συνεχής** δηλαδή ισχύει ότι

$$M_X(t) = \int e^{tx} dF(x)dx = \int e^{tx} f_X(x)dx$$

Ενώ στη περίπτωση που η τυχαία μεταβλητή είναι διακριτή αντί για ολοκλήρωμα στις δύο ισότητες έχουμε άθροισμα.

Ένα δεύτερο και πιο χρήσιμο εργαλείο για τις περιπτώσεις που το ολοκλήρωμα δεν συγκλίνει και επομένως η μέση τιμή δεν υπάρχει είναι **ο μετασχηματισμός Laplace**. Έστω λοιπόν ότι έχουμε $h(x)$ μια συνάρτηση που ορίζεται για $x > 0$ τότε ο μετασχηματισμός Laplace συμβολίζεται με $\widehat{h}\,(x)$ ως εξής

$$\widehat{h}\,(x) = \int_0^\infty e^{-sx} h(x)dx \ , \qquad s > 0$$

Οι ροπογεννητρίες συναρτήσεις και οι μετασχηματισμοί Laplace χρησιμοποιούνται για συνεχείς ή και μικτού τύπου τυχαίες μεταβλητές. Στις περιπτώσεις όπου η τυχαία μεταβλητή είναι διακριτή, τότε ευκολότερο να χρησιμοποιούμε την πιθανογεννήτρια συνάρτηση

Έστω η διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας $f_X(x) = P(X = x)$. Τότε η πιθανογεννήτρια συνάρτηση της X συμβολίζεται με

$$P_X(u) = E(u^X) = \sum_x f_X(x)u^x$$

Όπου το άθροισμα επεκτείνεται για όλες τις τιμές x για τις οποίες είναι

$$f_X(x) \neq 0$$

Ακολουθεί ένας πίνακας με τις βασικές κατανομές των τυχαίων μεταβλητών

ΣΥΝΕΧΕΙΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

Πίνακας 5.1

Κατανομή	Συνάρτηση Πυκνότητας Πιθανότητας	Μετασχηματισμός Laplace	Ροπογεννητρια
Ομοιόμορφη	$f_X(x) = \frac{1}{b-a}$ $, a \leq x \leq b$	$\frac{e^{-as} - e^{-bs}}{s(b-a)}$	$\frac{e^{bt} - e^{at}}{t(b-a)}$ $, t \neq 0$
Εκθετική	$f_X(x) = \lambda e^{-\lambda x}$ $, x > 0, \lambda > 0$	$\frac{\lambda}{\lambda + s}$	$\frac{\lambda}{\lambda - t}$
Γάμμα	$f_X(x) = \frac{\lambda^n}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ $, x > 0, \alpha > 0, \lambda > 0$	$\left(\frac{\lambda}{\lambda + s}\right)^{\alpha}$	$\left(\frac{\lambda}{\lambda - t}\right)^{\alpha}$
Erlang	$f_X(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$ $, x > 0, \lambda > 0, n \in N$	$\left(\frac{\lambda}{\lambda + s}\right)^{n}$	$\left(\frac{\lambda}{\lambda - t}\right)^{n}$
Κανονική	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$	$\exp\left\{-\mu s + \frac{\sigma^2 s^2}{2}\right\}$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$

ΔΙΑΚΡΙΤΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

Πίνακας 5.2

Κατανομή	Συνάρτηση Πιθανότητας	Πιθανογεννήτρια	Ροπογεννήτρια
Διωνυμική	$f_X(x) = \binom{n}{x} p^x q^{n-x}$, $x = 0,1,2,\dots,n$	$(q+pu)^n$	$(q+pe^t)^n$
Poisson	$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}$, $x = 0,1,2,\dots$	$e^{\lambda(u-1)}$	$e^{\lambda(e^t-1)}$
Γεωμετρική G_0	$f_X(x) = pq^x$, $x = 0,1,2,\dots$	$\frac{p}{1-qu}$	$\frac{p}{1-qe^t}$
ZTP	$f_X(x) = pq^{x-1}$, $x = 1,2,\dots$	$\frac{pu}{1-qu}$	$\frac{pe^t}{1-qe^t}$
Αρνητική διωνυμική	$f_X(x) = \binom{r+x-1}{x} p^r q^x$, $x = 0, 1, 2$	$\left(\frac{p}{1-qu}\right)^r$	$\left(\frac{p}{1-qe^t}\right)^r$
Αρνητική διωνυμική $NB_1(r,p)$	$f_X(x) = \binom{x-1}{r-1} p^r q^{x-r}$	$\left(\frac{pu}{1-qu}\right)^r$	$\left(\frac{pe^t}{1-qe^t}\right)^r$

Με βάση της τις γεννήτριες και τις ιδιότητες που προκύπτουν από το ορισμό τους μπορούμε να έχουμε έστω κάποια αποτελέσματα για την S την τυχαία μεταβλητή που αποτελεί και το άθροισμα τους. Καθώς

$$S = X_1 + X_2 + \cdots X_n \Leftrightarrow$$

$$e^{tS} = e^{t(X_1+X_2+\cdots X_n)} \Leftrightarrow$$

$$E(e^S) = E(e^{tX_1+tX_2+\cdots tX_n}) \Leftrightarrow$$

$$E(e^S) = E(e^{tX_1})E(e^{tX_2}) \dots E(e^{tX_n}) \Leftrightarrow$$

$$M_S(t) = M_{X_1}(t)M_{X_2}(t) \dots \ M_{X_n}(t)$$

Αν πέρα της ανεξαρτησίας που χρησιμοποιήθηκε στη τρίτη ισοδυναμία υποθέσουμε και την ισονομία μεταξύ των X_i η τελευταία εξίσωση θα μπορούσε να γραφτεί και ως

$$M_S(t) = \left(M_{X_1}(t)\right)^n$$

Καθώς οι ροπογεννητρίες καθορίζουν με μοναδικό τρόπο την κατανομή της τυχαίας μεταβλητής με βάση μπορούμε να αποδείξουμε σχετικά εύκολα το παρακάτω πίνακα που μας δίνει την κατανομή της S όταν αθροίζουμε ανεξάρτητες και ισόνομες τυχαίες μεταβλητές από γνωστές κατανομές

Πίνακας 5.3

Κατανομή της X_i	Κατανομή της S
Bernoulli(p)	Διωνυμική(n,p)
Διωνυμική(m,p)	Διωνυμική(nm,p)
Poisson (λ)	Poisson ($\sum \lambda_i$)
Γεωμετρική (p)	Αρνητική Διωνυμική (n,p)
Μετατοπισμένη στο Μηδέν Γεωμετρική (p)	Αρνητική Διωνυμική (n,p)
Αρνητική Διωνυμική (r,p)	Αρνητική Διωνυμική (nr,p)
Μετατοπισμένη στο Μηδέν Αρνητική Διωνυμική (r,p)	Μετατοπισμένη στο Μηδέν Αρνητική Διωνυμική (nr,p)
Εκθετική (λ)	Erlang (n,λ)
Γάμμα (α, λ)	Γάμμα (nα, λ)

5.1.2 Μέθοδος των συνελίξεων

Η μέθοδος των συνελίξεων είναι αν και είναι πιο δυσχρήστη σε σχέση με την μέθοδο των ροπογεννητριών έχει τα δύο μεγάλα προτερήματα.

1. Δεν απαιτεί να είναι κάποια ονομαστή η κατανομή των X_i όπως για παράδειγμα η εκθετική για να έχει ροπογεννητρία
2. Δίνει άμεσα αποτελέσματα για την κατανομή της S

Γενικά συνέλιξη των συναρτήσεων f και g είναι μια συνάρτηση h, που συμβολίζεται με

$$h(x) = (f * g)(x) = \int_0^x f(y)g(x-y)dy$$

Όταν οι f και g είναι δύο συνεχείς συναρτήσεις με πεδίο ορισμού $[0, +\infty)$ ενώ για την περίπτωση που η f και g είναι διακριτές η παραπάνω σχέση γράφεται

$$h(x) = (f * g)(x) = \sum_{y=0}^{x} f(y)g(x-y)$$

χρησιμοποιώντας της συνελίξεις των συναρτήσεων θα υπολογίσουμε την κατανομή της Sόπως αυτή έχει οριστεί σε παραπάνω. Αρχικά θα υποθέσουμε ότι το πλήθος των προσθετέων είναι 2. Τότε αν υποθέσουμε ότι $X_1 \sim F_{X_1}(x)$ και $X_2 \sim F_{X_2}(x)$ και οι αντίστοιχες συναρτήσεις πιθανότητας είναι $f_{X_1}(x), f_{X_2}(x)$ για τις διακριτές τυχαίες μεταβλητές σε επίπεδο ενδεχομένων έχουμε.

$$F_S(x) = P(X_1 + X_2 \leq x)$$

$$= P\left(\bigcup_{y=0}^{x} \left\{(X_1 \leq x - y) \bigcap (X_2 = y)\right\}\right)$$

$$= \sum_{y=0}^{x} P(X_1 + X_2 \leq x \mid X_2 = y)$$

$$= \sum_{y=0}^{x} P(X_1 + X_2 \leq x \mid X_2 = y)\, P(X_2 = y)$$

$$= \sum_{y=0}^{x} P(X_1 \leq x - y \mid X_2 = y)\, P(X_2 = y)$$

$$= \sum_{y=0}^{x} P(X_1 \leq x - y\ )\, P(X_2 = y)$$

$$= \sum_{y=0}^{x} F_{X_1}(x)\, f_{X_2}(y)$$

Το τελευταίο άθροισμα αν το συγκρίνουμε με το άθροισμα για την που μας δίνει την συνέλιξη των f και g θα παρατήρησουμε ότι η συνάρτηση κατανομής της S είναι η συνέλιξη των συναρτήσεων κατανομής της. Ενώ με την ίδια λογική μπορούμε να καταλήξουμε ότι η συνάρτηση πιθανότητας της Sείναι συνέλιξη των επιμέρους συναρτήσεων πιθανότητας. Δηλαδή

$$f_S = f_{X_1} * f_{X_2}$$

Το επόμενο βήμα είναι να βρούμε την κατανομή της S όταν το πλήθος των τυχαίων μεταβλητών είναι μεγαλύτερο ή ίσο του τρία. Αυτό γίνεται αναδρομικά εξής. Αφού βρούμε την κατανομή του αθροισμάτων των δύο πρώτων τυχαίων μεταβλητών, την χρησιμοποιούμε για να κάνουμε την συνέλιξη της με την συνάρτηση πιθανότητας της τρίτης τυχαίας μεταβλητής. Η διαδικασίας επαναλαμβάνεται με το ίδιο σκεπτικό μέχρι να γίνει η συνέλιξη και με την n-οστή μεταβλητή με την τυχαία μεταβλητη των n-1 πρώτων.

Αυτή η διαδικασία με την χρήση των ηλεκτρονικών υπολογιστών και στατιστικών προγραμμάτων όπως η R έχει απλοποιηθεί αρκετά μέσω πακέτων όπως το *actuar* και συναρτήσεων όπως η *aggregatedist* που δίνει αποτέλσμα τη συνάρτηση κατανομής της *S*.

5.2 Μοντέλο του Συλλογικού Κινδύνου

Στο κεφάλαιο αυτό θα εξετάσουμε το **μοντέλο του συλλογικού κινδύνου**. Το μοντέλο αυτό έχει πολλές εφαρμογές στον αναλογισμό στο γενικό κλάδο ασφαλίσεων όπως στο κλάδο αυτοκινήτων. Η βασική ιδέα είναι ότι ένας ασφαλιστής έχει αβεβαιότητα στο χαρτοφυλάκιο του για δύο (τυχαίες) μεταβλητές

1. Το πλήθος των ζημίων που θα προκύψουν μέσα σε ένα διάστημα, την διακριτή τυχαία μεταβλητή N, από έδω και στο εξής
2. Το μέγεθος της κάθε ζημίας την συνέχεις ή διακριτή τυχαία μεταβλητή X_i

Έτσι προκύπτει η βασική εξίσωση για το μοντέλο, όπου δίνει το μέγεθος της ζημίας S για το χαρτοφυλάκιο, ως εξής

$$S = \begin{cases} X_1 + X_2 + \cdots X_N, & N \geq 1 \\ 0, & N = 0 \end{cases}$$

<u>Υποθέσεις του μοντέλου</u>

1. Όπως και στη προηγούμενη παράγραφο έτσι και τώρα τα X_i θα υποτεθούν ως ανεξάρτητα και ισόνομα μεταξύ τους και αυτό είναι εφικτό με το διαχωρισμό του χαρτοφυλακίου σε υποχαρτοφυλάκια
2. Η ακολουθία $\{X_i\}_{i\geq 0}$ και το N είναι επίσης ανεξάρτητα μεταξύ τους

<u>Σχόλιο 1:</u> Στην παρούσα παράγραφο τα X_i , $i=1,2,\ldots,N$ θα υποτεθούν ως μη αρνητικές τυχαίες μεταβλητές. Αργότερα στο **κεφάλαιο 5** όταν και θα τροποποιήσουμε την εξίσωση για τους σκοπούς αυτής της διπλωματικής οι τιμές θα είναι από μείον ένα έως το άπειρο.

<u>Σχόλιο 2:</u> Από τη δομή της η S είναι ξεκάθαρο ότι δεν είναι ούτε διακριτή ούτε συνεχής τυχαία μεταβλητή. Οι τιμές που παίρνει είναι στο $[0, +\infty)$ και υπάρχει ένα σημείο , το μηδέν, που έχει πιθανότητα ενώ οι συνεχείς τυχαίες μεταβλητές η πιθανότητα σε ένα σημείο είναι μηδέν. Τέτοιες συναρτήσεις λέγονται **μικτού τύπου** και η κατανομή τους είναι καλά ορισμένη όταν έχουμε και την πιθανότητα στο μηδέν.

<u>Συμβολισμοί:</u>

1. $F(x) = P(X \leq x)$ η πιθανότητα μια ζημία να μην ξεπεράσει το x
2. $f(x)\begin{cases} F'^{(x)} \\ P(X = x) \end{cases}$ η συνάρτηση πυκνότητας πιθανότητας
3. $\bar{F}(x) = (X > x)$ η συνάρτηση της δεξίας ούρας της X
4. $p_n = P(N = n)$ η συνάρτηση πιθανότηατς της N

5. $G(x) = P(S \leq x)$
6. $g(x)\begin{cases} G'(x) \\ P(S = x) \end{cases}$

5.2.1 Μέθοδος των γεννητριών συναρτήσεων

Όπως και στην προηγούμενη παράγραφο όπου το πλήθος το προσθετέων ήταν καθορισμένο εξαρχής και ήταν μια γνωστή σταθερά έτσι και τώρα που το πλήθος είναι N τυχαία μεταβλητή ο πιο εύκολος τρόπος να χειριστούμε το πρόβλημα είναι με της γεννήτριες συναρτήσεις.

Η ροπογεννήτρια της S δίνεται είναι

$$M_S(t) = E\big(E(e^{tS}|N)\big) = \sum_{n=0}^{\infty} E(e^{tS}|N = n)p_n$$

Όμως

$$E(e^{tS}|N = n) = E\big(e^{t(X_1+X_2+\cdots+X_n)}|N = n\big) =$$

$$E\big(e^{t(X_1+X_2+\cdots+X_n)}\big) = [M_X(t)]^n$$

Άρα η ροπογεννητρία της S δίνεται από

$$M_S(t) = \sum_{n=0}^{\infty} [M_X(t)]^n p_n$$

Όμως η πιθανογεννήτρια της N είναι

$$P_N(u) = \sum_{n=0}^{\infty} u^n p_n \Rightarrow P_N\big(M_X(t)\big) = \sum_{n=0}^{\infty} [M_X(t)]^n p_n$$

Εξισώντας τις δύο τελευταίες ισότητες έχουμε

$$M_S(t) = P_N\big(M_X(t)\big)$$

Και λόγω του μονοσήμαντου το ροπογεννητριών αν είναι σε κλειστή μορφή η ροπογεννήτρια της X και η πιθανογεννήτρια της N μπορούμε να βρούμε σε κλειστή μορφή της κατανομή της S αν όπου η είναι η ροπογεννήτρια της X αντικαταστήσουμε με την κατανομή της.

5.2.2 Η Μέθοδος των συνελίξεων

Η μέθοδος των συνελίξεων είναι η βασική μέθοδος και μέχρι και το πρόσφατο παρελθόν η μοναδική. Το πλεονέκτημα της έναντι της μεθόδου των γεννητριών είναι ότι μπορεί να χρησιμοποιηθεί για οποιαδήποτε κατανομή των N και X.

Με βάση τους συμβολισμούς που ορίσαμε στη αρχή αυτής της υποπαραγράφου έχουμε

$$G(x) = P(S \leq x) = \sum_{n=0}^{\infty} P(X_1 + X_2 + \cdots + X_N \leq x | N = n)\, P(N = n)$$

$$= \sum_{n=0}^{\infty} p_n P(X_1 + X_2 + \cdots + X_N \leq x | N = n)$$

$$= \sum_{n=0}^{\infty} p_n P(X_1 + X_2 + \cdots + X_n \leq x | N = n) \Leftrightarrow$$

$$G(x) = \sum_{n=0}^{\infty} p_n\, P(X_1 + X_2 + \cdots + X_n \leq x) \Leftrightarrow G(x) = \sum_{n=0}^{\infty} p_n F^{*n}(x)$$

Όπου η $F^{*n}(x)$ είναι η ν-οστή συνέλιξη των τυχαίων ανεξάρτητών και ισόνομων τυχαίων μεταβλητών. Ακόμα αν λάβουμε υπόψη μας ότι $F^{*0}(x) = 1$ η τελευταία εξίσωση γίνεται

$$G(x) = \sum_{n=0}^{\infty} p_n F^{*n}(x) = p_0 + \sum_{n=1}^{\infty} p_n F^{*n}(x)$$

Τα μειονεκτήματα της μεθόδου είναι, όπως αναφέρθηκε και νωρίτερα η εύρεση της ν-οστής συνέλιξης και το αν η τελευταία σειρά συγκλίνει.

Βέβαια αν η N είναι φραγμένη, έστω από ένα M μη αρνητικό ακέραιο τότε θα ισχύει $p_n = 0\ \forall\ n \geq M + 1$ άρα η παραπάνω εξίσωση θα γίνει

$$G(x) = \sum_{n=0}^{M} p_n F^{*n}(x) = p_0 + \sum_{n=1}^{M} p_n F^{*n}(x)$$

Για την εύρεση της συνάρτησης της δεξιάς ουράς εξ ορισμού είναι

$$\bar{G}(x) = 1 - G(x)$$

6 Στοχαστική Απόδοση των Unit – Linked προϊόντων

6.1 Κατασκευή της τυχαία μεταβλητής $\tilde{R}_t$

Ένας ασφαλισμένος με ένα ασφαλιστήριο ζωής Unit – Linked είναι εκτεθειμένος σε δύο κινδύνους

1. Στη απόδοση των επενδύσεων του funds με το οποίο είναι συνδεδεμένο το ασφαλιστήριο ζωής και καταβάλλεται ο τόκος στη λήξη της περιόδου και
2. Το κίνδυνο αποχώρησης από το πρόγραμμα, λόγω θανάτου ή εξαγοράς του

Όπως αποδείχθηκε στο κεφάλαιο ένα η ετήσια λογαριθμική απόδοση ενός χρηματοοικονομικού προϊόντος μπορεί να γραφτεί μέσω αθροίσματος των λογαριθμικών αποδόσεων μικρότερων χρονικών διαστημάτων. Έτσι αν υποθέσουμε ότι

$$\tilde{R}_t = \sum_{i=1}^{k} \tilde{\xi}_i \quad , \qquad \kappa = 1,2,\dots 12$$

όπου το κ είναι το πλήθος των μηνών σε ένα έτος.

Βέβαια λόγω του δεύτερου κινδύνου το οποίο διατρέχει ο ασφαλισμένος το κ δεν μπορεί να υποτεθεί ως σταθερά.

Θα αντικαταστήσουμε το κ με την τυχαία μεταβλητή με τιμές

$$N = \{1,2,\dots\}$$

Άρα η προηγουμένη εξίσωση γίνεται

$$\tilde{R}_t = \sum_{i=1}^{N} \tilde{\xi}_i \quad , \qquad N = 1,2,\dots 12$$

Αυτό μας παραπέμπει στο μοντέλο του συλλογικού κινδύνου και θα εξετάσουμε αν αυτό είναι εφαρμόσιμο.

6.2 Μελέτη της τυχαίας μεταβλητής N

6.2.1 Η συνάρτηση κατανομής της N

Στο τέλος κάθε μήνα ο ασφαλισμένος είναι δυνατόν είτε να παραμένει στο πρόγραμμα είτε να έχει αποχωρήσει. Άρα τα πιθανά ενδεχόμενα είναι δύο

$$A = \{(x)\ \text{έχει επιβιώσει και δεν έχει εξαγοράσει το συμβόλαιο μέχρι τον μήνα}\ n\}$$

$$A = \{(x)\ \text{είναι ενεργός στο ασφαλίστηριο μέχρι το μήνα}\ n\}$$

και

$$A' = \{(x)\ \text{δεν είναι ενεργός στο ασφαλίστηριο μέχρι το μήνα}\ n\}$$

Τότε, η πιθανότητα του ενδεχομένου A, είναι

$$P(A) = P(N \leq n).$$

Αυτό θυμίζει δοκιμή Bernoulli καθώς υπό την εκθετική παρεμβολή μέσα σε ένα έτος η ένταση θανάτου – αποχώρησης είναι σταθερή, όπως αποδείχθηκε στο κεφάλαιο τρία, λόγω την αμνημονής ιδιότητας της εκθετικής.

Άρα για κάθε μήνα θα ισχύει ότι η πιθανότητα να παραμείνει ενεργός ο ασφαλισμένος, κάτω από την υπόθεση της σταθερής έντασης θανάτου , είναι

$$ {}_{t}p_x^{(\tau)} = p \,, \qquad t = \frac{1}{12} $$

το οποίο είναι λογικό όταν αναφερόμαστε σε ένα έτος καθώς *(30)* και *(30.5)* είναι εκτεθειμένοι στους ίδιους κίνδυνους αποχώρησης από το ασφαλιστήριο.

Αυτή η υπόθεση δεν θα μπορούσε να επεκταθεί καθ’ όλη την διάρκεια του ασφαλιστηρίου, που συνηθώς είναι είκοσι χρόνια αλλά μέσα σε ένα έτος είναι χρήσιμη

Άρα την N ως τυχαία μεταβλητή έχουμε ότι είναι επαναλήψεις πειραμάτων Bernoulli που επαναλαμβάνεται μέχρι να εμφανιστεί η πρώτη «επιτυχία» την αποχώρηση του ασφαλισμένου, έστω q ,και ότι είναι φραγμένη καθώς λαμβάνει τιμές στο $n = \{1,2, \ldots 12\}$.

Αυτό μας οδηγεί στο συμπέρασμα ότι η N είναι μια **λογοκριμένη τυχαία μεταβλητή.** Δηλαδή έχουμε

$Y \sim Geom(q)$ η είναι $N = Y \wedge 12 = min\{Y, 12\}$ με

$q = 1 - p = 1 - {}_{t}p_x^{(\tau)} = {}_{t}q_x^{(\tau)}$.

Για την συνάρτηση κατανομής της N έχουμε

$$ F_N(n) = P(N \leq n) = \begin{cases} F_Y(n) \,, & n < 12 \\ 1 \,, & n \geq 12 \end{cases} $$

όπου

$$ F_Y(n) = P(Y \leq n) = \sum_{i=0}^{n} f_Y(i) = \sum_{i=0}^{n} q p^i = \cancel{q} \frac{1 - p^{n+1}}{\cancel{1 - p}} = 1 - p^{n+1} $$

Άρα

$$ F_N(n) = P(N \leq n) = \begin{cases} 1 - p^{n+1} \,, & n < 12 \\ 1 \,, & n \geq 12 \end{cases} $$

6.2.2 Η συνάρτηση πιθανότητας της N

Για την συνάρτηση πιθανότητας έχουμε

$$ f_N(n) = P(N = n) = \begin{cases} f_Y(n), & n < 12 \\ \bar{F}_N(n) \,, & n \geq 12 \end{cases} $$

Και

$$\bar{F}_N(n) = P(N \geq n) = 1 - P(N < n) = 1 - P(N \leq n-1)$$

$$\bar{F}_N(n) \Leftrightarrow 1 - F_N(n-1) \Leftrightarrow \bar{F}_N(n) = p^n$$

Άρα παίρνουμε

$$p_n = f_N(n) = \begin{cases} qp^n, & n < 12 \\ p^{12}, & n \geq 12 \end{cases}$$

Ακολουθεί πίνακας με την συνάρτηση κατανομής και την συνάρτηση πιθανότητας της N για ασφαλιστήριο συμβόλαιο για ένα έτος

<u>**Πίνακας 6.1**</u>

	No	Prob(N<=n)	Prob(N=n)
[1,]	0	0.004987521	0.004987521
[2,]	1	0.009950166	0.004962645
[3,]	2	0.014888060	0.004937894
[4,]	3	0.019801327	0.004913266
[5,]	4	0.024690088	0.004888761
[6,]	5	0.029554466	0.004864378
[7,]	6	0.034394584	0.004840117
[8,]	7	0.039210561	0.004815977
[9,]	8	0.044002518	0.004791957
[10,]	9	0.048770575	0.004768057
[11,]	10	0.053514852	0.004744277
[12,]	11	0.058235466	0.004720614
[13,]	12	1.000000000	0.941764534

Για την παραγωγή του πίνακα έχουμε υποθέσει ότι

$$\mu_{x+t}^{(j)} = \mu_x^{(j)} = 0.03, \qquad j = 1,2$$

Επομένως

$$\mu_{x+t}^{(\tau)} = \mu_x^{(\tau)} = \mu_x^{(1)} + \mu_x^{(2)} = 0.06$$

Και η πιθανότητα «επιτυχίας» είναι

$${}_{t}q_x^{(\tau)} = 1 - {}_{t}p_x^{(\tau)} = 1 - exp\left\{\int_0^t \mu_x^{(\tau)} ds\right\} = 1 - exp\left\{-\mu_x^{(\tau)} t\right\}$$

$$= 1 - exp\left\{\frac{-\mu_x^{(\tau)}}{12}\right\} = 0.004987521$$

6.3 Έλεγχος των προϋποθέσεών του μοντέλου του Συλλογικού Κινδύνου

Όπως αναφέρθηκε και στο ακριβώς προηγούμενο κεφάλαιο το μοντέλο του συλλογικού κίνδυνου το μοντέλο ισχύει κάτω από δύο προϋποθέσεις:

1. Οι συνεχείς τυχαίες μεταβλητές $\tilde{\tilde{\xi}}_i$ είναι μεταξύ τους ανεξάρτητες και ισόνομες
2. Οι συνεχείς τυχαίες μεταβλητές $\tilde{\tilde{\xi}}_i$ είναι ανεξάρτητες με το N.

Θα αρχίσουμε με την δεύτερη υπόθεση που είναι ευκολότερη τα ενδεχόμενα του θανάτου ενός ασφαλισμένου και των αποδόσεων είναι ανεξάρτητα μεταξύ τους. Θα μπορούσε ωστόσο να υπάρχει μια συσχέτιση μεταξύ των αποδόσεων και των εξαγορών των συμβολαίων καθώς σε περίοδο ύφεσης, άρα χαμηλών αποδόσεων, είναι πιθανό ένας ασφαλισμένος να «σπάσει το κουμπάρα» του για να ανταπεξέλθει σε πιο βραχυπρόθεσμες υποχρεώσεις.

Αυτό το φαινόμενο ονομάζεται **insurance runs** και θεωρείται λιγότερο πιθανό σε σχέση με το **bank panic** (ο λογαριασμός που άμεσα δίνει ρευστότητα σε ένα καταθέτη) καθώς στα ασφαλιστήρια ζωής υπάρχουν **penalty** εξαγοράς. Με βάση αυτό τα $\tilde{\tilde{\xi}}_i$ θεωρούνται ανεξάρτητα.

Για την πρώτη προϋπόθεση του μοντέλου έχουμε δύο συνθήκες να ελέγξουμε

1. Την Ισονομία των $\tilde{\tilde{\xi}}_i$
2. Την ανεξαρτησία μεταξύ τους

Για την ισονομία των μηνιαίων λογαριθμικών αποδόσεων, συνεπάγεται λόγω τις δομής των funds που είναι αυστηρά καθορισμένη από τη πρώτη στιγμή. Η δομή του χαρτοφυλακίου που επενδύει ένα fund δεν δύναται να αλλάξει από μήνα σε μήνα. Άρα οι μεταβλητές που ενδεχομένως να επηρεάζουν τις αποδόσεις τους όπως ο πληθωρισμός μπορεί να θεωρηθεί σταθερός μέσα σε ένα έτος. Ακόμα και σε μεγάλες οικονομικές κρίσεις η άνοδος ή η πτώση οικονομικών δεικτών γίνεται σταδιακά μέσα στο έτος.

Για την ανεξαρτησία λόγω του προβλήματος της δυναμικής ετεροσκεδαστικότητας είναι δύσκολο να αποδειχθεί. Ωστόσο στη βιβλιογραφία έχει αποδειχθεί ότι οι αποδόσεις μεταξύ τους είναι γραμμικά ανεξάρτητες που με επιπλέον υπόθεση την κανονικότητα, που στο κεφάλαιο δύο είδαμε ότι έχουμε ισχυρές ενδείξεις ύπαρξης, μας οδηγεί στο συμπέρασμα να μπορούμε να υποθέσουμε την ανεξαρτησία.

6.4 Εφαρμογή του μοντέλου και αριθμητικά αποτελέσματα.

6.4.1 Εφαρμογή του μοντέλου

Για την εφαρμογή του μοντέλου αρχικά θα διευκρινίσουμε ότι στη περίπτωση που εξετάζουμε έχει μάζα πιθανότητας το μηδέν και αυτό ως αποτέλεσμα της δομής της τυχαίας μεταβλητής που ο τόκος καταβάλλεται στο τέλος της περιόδου. Αν ο ασφαλισμένος αποχωρήσει πριν το πρώτο μήνα δεν λαμβάνει απόδοση.

Άρα εξετάζουμε την πιθανότητα

$$G_{\tilde{R}_t} = P(\tilde{R}_t \le r) = \sum_{n=0}^{\infty} P(\tilde{\xi}_1 + \tilde{\xi}_2 + \dots + \tilde{\xi}_n \le r | N = n) p_n$$

Αυτό το τυχαίο άθροισμα με βάση το κεφάλαιο τέσσερα και ειδικότερα την μέθοδο των συνελίξεων είναι

$$G_{\tilde{R}_t}(r) = P(\tilde{R}_t \le r) = \sum_{n=0}^{12} p_n F_{\tilde{\xi}}{}^{*n}(r)$$

Δηλαδή

$$G_{\tilde{R}_t}(r) = q \sum_{n=0}^{12} p^n F_{\tilde{\xi}}{}^{*n}(r) + p^{12} F_{\tilde{\xi}}{}^{*12}(r)$$

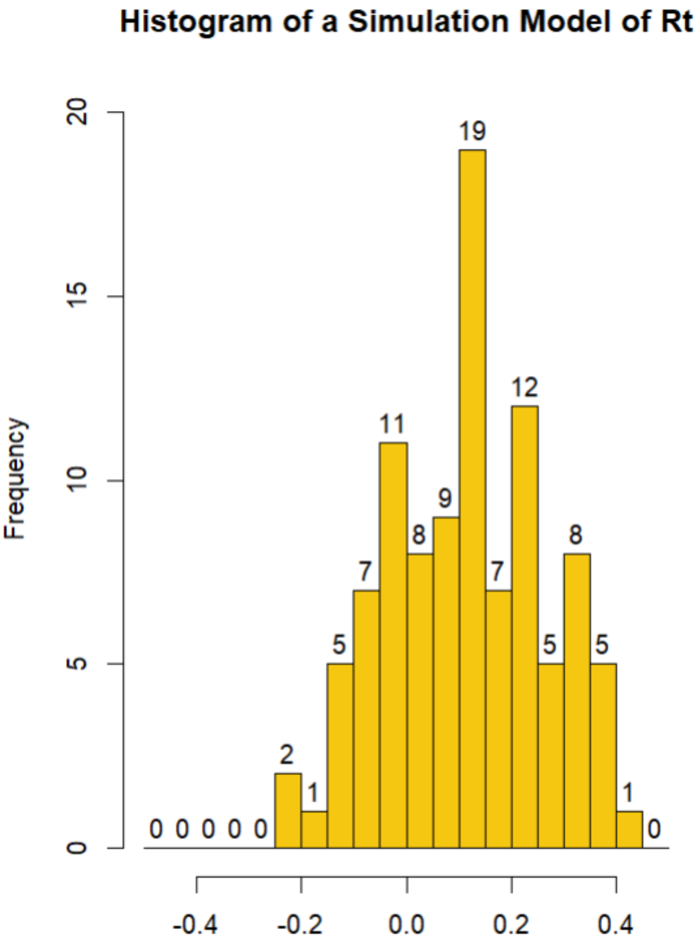
Όπου $F_{\tilde{\xi}}{}^{*n}$ είναι μια συνέλιξη από *n* ανεξάρτητες και ισόνομες κανονικές κατανομές.

6.4.2 Αριθμητικά αποτελέσματα

Υποθέτουμε ότι τα $\tilde{\xi}_i \sim N(0.01, 0.0025)$,όπου είναι η μέση τιμή και η διακύμανση του fund με τα μεγαλύτερα p-values στους ελέγχους υποθέσεων του κεφαλαίου δυο. Θα παραχθεί ένα δείγμα από 100 παρατηρήσεις της $\tilde{R}_t$. Στη συνέχεια θα κατασκευαστεί το ιστόγραμμα συχνότητας του δείγματος και θα γίνουν οι έλεγχοι υποθέσεων της κανονικότητας του πληθυσμού.

Με αυτό το τρόπο θα προσεγγίσουμε ένα δείγμα 100 ασφαλισμένων και τις αποδόσεις που θα λάβουν μέσα στο χρόνο από το ασφαλιστήριο

Παρατηρώντας το ιστόγραμμα μιας προσομοίωσης της $\tilde{R}_t$ για υποθετικά 100 ασφαλισμένους μέσα σε ένα έτος ανεξαρτήτου αν αποχώρησαν από το ασφαλιστήριο μπορούμε να υποστηρίξουμε τα εξής.



Από το ιστόγραμμα ακόμα θα μπορούσαμε να υποθέσουμε ότι και η κατανομή της $\tilde{R}_t$ δείχνει κάποιες ομοιότητες με την κανονική αλλά αυτό μπορεί και να αποδοθεί και στο πλήθος των μετρήσεων και το κεντρικό οριακό θεώρημα.

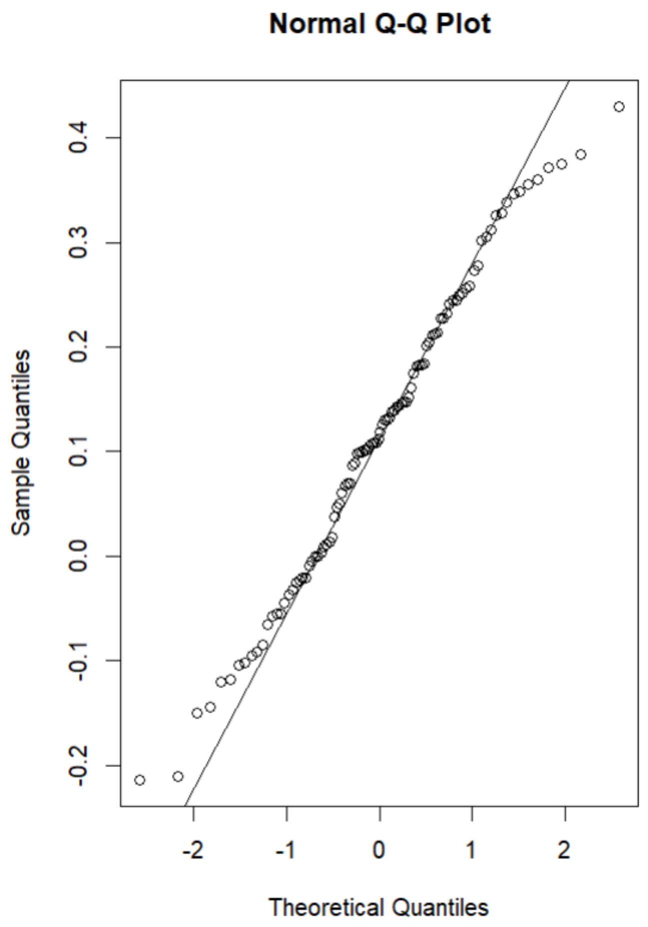
```
> summary(S)
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.2138129 -0.0000453  0.1149714  0.1154454  0.2268824  0.4294342
```

Ακόμα τ α βασικά μέτρα θέσης και διασποράς για το δείγμα μας δίνουν εικόνα κανονικής καθώς η μέση τιμή με την διάμεσο είναι πολύ κοντά σαν ένδειξη μονοκόρυφης κατανομής.

Με αυτά σαν ενδείξεις θα προχωρήσουμε στους ελέγχους κανονικότητας που πραγματοποιηθήκαν στο 2ο κεφάλαιο για την

Παρατηρώντας το qq plot διάγραμμα θα δούμε ότι οι κεντρικές τιμές τείνουν να είναι κοντά στη γραμμή. Οι τιμές άνω του 0.3 και κάτω -0.1 ωστόσο τείνουν να απομακρύνονται και μάλιστα με τάση άνω της γραμμής για τα μικρές τιμές και κάτω τις γραμμής για τις μεγαλύτερες.

Τέλος κάνοντας και τους ελέγχους Kolmogorov – Smirnov, Lillieforce και Shapiro – Wilk λαμβάνουμε τα εξής αποτελέσματα




```
> shapiro.test(S)

        Shapiro-Wilk normality test

data:  S
W = 0.98651, p-value = 0.4054

> library(nortest)
> lillie.test(S)

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  S
D = 0.056035, p-value = 0.6154

> ks.test(S,"pnorm",mean= mean(S),sd=sd(S))

        One-sample Kolmogorov-Smirnov test

data:  S
D = 0.056035, p-value = 0.912
alternative hypothesis: two-sided
```

και με τα p values να είναι ιδιαίτερα ψηλά μπορούμε να επιφύλαξη να υποστηρίξουμε ότι η

$$G_{\tilde{R}_t}(r) = q \sum_{n=0}^{12} p^n F_{\tilde{\xi}}^{*n}(r) + p^{12} F_{\tilde{\xi}}^{*12}(r)$$

Μπορεί έστω να προσεγγιστεί από κανονική κατανομή.

6.5 Συμπεράσματα

Όπως έχει ήδη ειπωθεί η αποδόσεις ενός με ρίσκου περιουσιακού στοιχείου είναι μια στοχαστική ανέλιξη στο χρόνο. Δηλαδή μια ακολουθία τυχαίων μεταβλητών συνάρτηση του χρόνου.

Τα ασφαλιστήρια ζωής Unit – Linked είναι ασφαλιστήρια ζωής με στόχο την αποταμίευση – επένδυση και εμπεριέχουν ρίσκο. Η στοχαστική λογαριθμική τους απόδοση ως στοχαστική ανέλιξη είναι η ακολουθία $\{\widetilde{R}_t\}_{t\geq 1}$

Η παραπάνω ακολουθία μπορεί να υποτεθεί ότι αποτελείται από κανονικές κατανομές η οποίες όμως επηρεάζονται τόσο από το οικονομικό περιβάλλον όσο και από τη ηλικία και τρόπο ζωής του ασφαλισμένου τα οποία μπορούν να το οδηγήσουν σε εξαγορά ή θάνατο.

7 Βιβλιογραφία

7.1 Ελληνική

1. Αντζουλάκος Δ. (2018) Σημειώσεις παραδόσεων για το μάθημα Γλώσσα Προγραμματισμού R με εφαρμογές στον Αναλογισμό, Πανεπιστήμιο Πειραιά
2. Αντζουλάκος Δ. (2022) Σημειώσεις παραδόσεων για το μάθημα Συμβάντα Ζωής και Θανάτου ΙΙ, Πανεπιστήμιο Πειραιά
3. Χατζηκωνσταντινίδης Ευστ. Σημειώσεις παραδόσεων για το μάθημα Θεωρία Κινδύνου Ι, Πανεπιστήμιο Πειραιά
4. Χατζόπουλος Π. (2011) Μαθηματικά Ασφαλίσεων Ζωής, Συμμετρία

7.2 Ξενόγλωση

1. John C. Hull (2014) Options Futures and Others Derivatives *Ninth Edition*, Pearson