

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

ΚΑΤΑΝΟΜΕΣ ΤΥΠΟΥ ΦΑΣΕΩΝ ΚΑΙ ΟΙ ΕΦΑΡΜΟΓΕΣ ΤΟΥΣ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ

Γεώργιος Χ. Κουτσογιαννάκης

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

Πειραιάς
Σεπτέμβριος 2019

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμόν 6 /11.06.2018 συνεδρίασή της σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής Δ. Αντζουλάκος (Επιβλέπων)
- Αναπληρωτής Καθηγητής Μ. Μπούτσικας
- Επίκουρος Καθηγητής Γ. Ψαρράκος

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

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

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

**PHASE TYPE DISTRIBUTIONS AND
THEIR APPLICATIONS IN ACTUARIAL
SCIENCE**

Georgios C. Koutsogiannakis

Thesis

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfillment of
the requirements for the degree of Master of Science in
actuarial science and risk management

Piraeus, Greece
September 2019

Στην οικογένειά μου και στους φίλους μου

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές ευχαριστίες και την ιδιαίτερη εκτίμηση μου, προς τον καθηγητή του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς και επιβλέποντα στην διπλωματική μου εργασία κ. Δημήτριο Αντζουλάκο, για την καθοδήγησή του καθ' όλη τη διάρκεια εκπόνησης της εργασίας αυτής καθώς και την καθοριστική συμβολή του στην ολοκλήρωσή της.

Ακόμη, θα ήθελα να ευχαριστήσω και τους καθηγητές κ. Μιχαήλ Μπούτσικα και κ. Γεώργιο Ψαρράκο που δέχτηκαν να είναι μέλη της τριμελούς επιτροπής αξιολόγησης της μεταπτυχιακής μου εργασίας, καθώς επίσης και όλους τους διδάσκοντες του μεταπτυχιακού προγράμματος για τις γνώσεις που μας προσέφεραν.

Τέλος, δεν θα μπορούσα να παραλείψω τις ευχαριστίες μου, αλλά και την ευγνωμοσύνη μου, προς τους γονείς μου για την υποστήριξη, την υπομονή και την βοήθειά τους σε αυτό το μεγάλο ταξίδι των προπτυχιακών και μεταπτυχιακών μου σπουδών.

Περίληψη

Τα τελευταία χρόνια υπάρχει αυξημένο ακαδημαϊκό και επιστημονικό ενδιαφέρον γύρω από τις κατανομές τύπου φάσεων. Τα πιο απλά παραδείγματα κατανομών τύπου φάσεων είναι οι συνελίξεις εκθετικών κατανομών και τα μοντέλα πεπερασμένων μείξεων. Δεδομένου ότι οι διακριτές κατανομές τύπου φάσεων είναι πυκνές στην κλάση των κατανομών στο \mathbb{N}_0 , τότε είναι σαφές ότι οι συνεχείς κατανομές τύπου φάσεων είναι πυκνές στην κλάση των κατανομών στο \mathbb{R}_+ . Αυτή η ιδιότητα τους τις καθιστά ένα πολύ ευπροσάρμοστο εργαλείο μοντελοποίησης.

Στόχος αυτής της διπλωματικής εργασίας είναι η αναλυτική παρουσίαση των κατανομών τύπου φάσεων και ορισμένων εφαρμογών τους στην αναλογιστική επιστήμη.

Αναλυτικότερα, στο πρώτο κεφάλαιο γίνεται αρχικά μία περιληπτική παρουσίαση των Μαρκοβιανών αλυσίδων τόσο σε διακριτό όσο και σε συνεχή χρόνο. Στη συνέχεια εισάγουμε και μελετούμε τις κατανομές τύπου φάσεων, τόσο στη διακριτή όσο και στη συνεχή περίπτωση, παρουσιάζοντας πληθώρα παραδειγμάτων και εφαρμογών.

Συνεχίζουμε με το δεύτερο κεφάλαιο όπου αναλύουμε το μοντέλο συλλογικού κινδύνου το οποίο συνδυάζουμε με τις κατανομές τύπου φάσεων. Μέσα από θεωρήματα, παραδείγματα και εφαρμογές βλέπουμε πως οι κατανομές τύπου φάσεων εφαρμόζονται σε αυτό το κομμάτι της αναλογιστικής επιστήμης.

Στο τρίτο και τελευταίο κεφάλαιο ασχολούμαστε με τη θεωρία χρεοκοπίας. Η εφαρμογή των κατανομών τύπου φάσεων στη θεωρία χρεοκοπίας παρουσιάζει ιδιαίτερο ενδιαφέρον καθώς μπορεί να αποτελέσει ένα επιπλέον εργαλείο για τους αναλογιστές. Μέσα από τις εφαρμογές και με τη χρήση κατάλληλων αποτελεσμάτων που μας δίνει η θεωρία, βρίσκουμε σε ορισμένες περιπτώσεις την πιθανότητα χρεοκοπίας χρησιμοποιώντας κατανομές τύπου φάσεων.

Abstract

In recent years the interest in Phase Type distributions has been greatly increased. The simplest examples of Phase Type distributions are convolutions of exponential distributions and models of finite mixtures. Given that discrete Phase Type distributions are dense in the class of distributions on \mathbb{N}_0 , it is then clear that continuous Phase Type distributions are dense in the class of distributions on \mathbb{R}_+ . This property makes them a very versatile modeling tool.

The purpose of this thesis is to present Phase Type distributions and some of their applications in actuarial science.

More specifically, in the first chapter, a brief presentation of Markov chains in both discrete and continuous time is given. In the sequel we introduce and study Phase Type distributions both in the discrete and continuous case, presenting a variety of examples and applications.

In the second chapter we study the collective risk model in relation to Phase Type distributions. Using theoretical results, examples, and applications, we show that Phase Type distributions can be applied to this field of actuarial science.

Finally, in the third chapter we deal with ruin theory models. The application of Phase Type distributions to ruin theory can be an additional tool for actuaries. Through applications and using appropriate theoretical results, we are able to find the probability of ruin using Phase Type distributions in some cases.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ.....	xiii
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ.....	xiv
ΚΕΦΑΛΑΙΟ 1.....	1
Κατανομές τύπου φάσεων.....	1
1.1 Εισαγωγή.....	1
1.2 Μαρκοβιανές αλυσίδες σε διακριτό χρόνο.....	2
1.2.1 Ορισμός και βασικές ιδιότητες.....	2
1.2.2 Επαναληπτικές και μεταβατικές καταστάσεις.....	3
1.2.3 Απορροφητικές Μαρκοβιανές αλυσίδες.....	7
1.2.4 Εφαρμογές.....	8
1.3 Μαρκοβιανές αλυσίδες σε συνεχή χρόνο.....	11
1.3.1 Ορισμός και βασικές ιδιότητες.....	11
1.3.2 Γεννήτορας πίνακας, εμφυτευμένη Μαρκοβιανή αλυσίδα και είδη καταστάσεων.....	12
1.3.3 Ιδιότητες του γεννήτορα πίνακα και στάσιμη κατανομή.....	14
1.3.5 Εφαρμογές.....	16
1.4 Κατανομές τύπου φάσεων.....	19
1.4.1 Διακριτές κατανομές τύπου φάσεων.....	19
1.4.2 Πρότυπα διακριτών κατανομών τύπου φάσεων.....	31
1.4.3 Αριθμητικό παράδειγμα διακριτής κατανομής τύπου φάσεων.....	33
1.4.4. Συνεχείς κατανομές τύπου φάσεων.....	35
1.4.5 Πρότυπα συνεχών κατανομών τύπου φάσεων.....	41
1.4.6 Αριθμητικό παράδειγμα συνεχούς κατανομής τύπου φάσεων.....	47
ΚΕΦΑΛΑΙΟ 2.....	49
Μοντέλο συλλογικού κινδύνου και κατανομές τύπου φάσεων.....	49
2.1 Εισαγωγή.....	49
2.2 Το μοντέλο συλλογικού κινδύνου.....	50
2.3 Μοντέλο συλλογικού κινδύνου και η κλάση κατανομών $\mathfrak{R}(a, b, 0)$	54
2.4 Μοντέλο συλλογικού κινδύνου και η κλάση κατανομών $\mathfrak{R}(a, b, 1)$	58
2.4 Μοντέλο συλλογικού κινδύνου και κατανομές τύπου φάσεων.....	61

2.5 Εφαρμογές.....	65
ΚΕΦΑΛΑΙΟ 3.....	77
Θεωρία χρεοκοπίας και κατανομές τύπου φάσεων	77
3.1 Εισαγωγή.....	77
3.2 Η στοχαστική ανέλιξη πλεονάσματος.....	78
3.2.1 Ορισμοί και το κλασσικό μοντέλο της θεωρίας κινδύνων	78
3.2.2 Μελέτη της πιθανότητας χρεοκοπίας	81
3.2.3 Οι τύποι του Lundberg	84
3.2.4 Εφαρμογές.....	86
3.3 Κατανομές τύπου φάσεων στη θεωρία χρεοκοπίας.....	90
3.3.1 Phase type renewal theory	91
3.3.2 Εφαρμογές.....	92
3.3.3 Πιθανότητα χρεοκοπίας σε μια δισδιάστατη διαδικασία Poisson	100
ΠΑΡΑΡΤΗΜΑ.....	107
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	117

ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1.1. Διάγραμμα καταστάσεων διαδικασίας Poisson.....	17
Σχήμα 1.2. Διαγραμματική περιγραφή της διαδικασίας Poisson.....	17
Σχήμα 1.3. Διάγραμμα καταστάσεων διαδικασία γεννήσεως θανάτου.....	19
Σχήμα 1.4. Συνάρτηση πιθανότητας των κατανομών $G_4(0.6)$ και $G_4(0.8)$	35
Σχήμα 1.5. Διάγραμμα φάσης εκθετικής κατανομής.....	41
Σχήμα 1.6. Διάγραμμα φάσης υπερεκθετικής κατανομής.....	43
Σχήμα 1.7. Διάγραμμα φάσης της κατανομής Erlang.....	44
Σχήμα 1.8. Διάγραμμα φάσης υποεκθετικής κατανομής.....	45
Σχήμα 1.9. Διάγραμμα φάσης κατανομής Coxian.....	46
Σχήμα 1.10. Διάγραμμα φάσης γενικευμένης κατανομής Coxian.....	46
Σχήμα 1.11. Συνάρτηση πυκνότητας της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$ όπου $\mathbf{a} = (1/3, 2/3)$, $\mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$	48
Σχήμα 2.1. Συνάρτηση πιθανότητας της $S_N \sim PH_d^{(25)}(\boldsymbol{\gamma}, \mathbf{C})$	71
Σχήμα 2.2. Συνάρτηση πυκνότητας της $S_N \sim PH_c^{(2)}(\boldsymbol{\gamma}, \mathbf{C})$	73
Σχήμα 2.3. Συνάρτηση κατανομής της S_N	75
Σχήμα 3.1. Στοχαστική διαδικασία πλεονάσματος.....	79
Σχήμα 3.2. Στοχαστική διαδικασία πλεονάσματος συναρτήσει της τ.μ. του ελλείματος.....	82
Σχήμα 3.3. Μέγιστη σφρευτική απώλεια στη διαδικασία πλεονάσματος.....	83
Σχήμα 3.4. Εφαρμογή 1: Πιθανότητα χρεοκοπίας στο κλασικό μοντέλο.....	94
Σχήμα 3.5. Εφαρμογή 2: Πιθανότητα χρεοκοπίας στο μοντέλο Sparre-Andersen..	96
Σχήμα 3.6. Εφαρμογή 4: Διάγραμμα ανανεωτικής πυκνότητας.....	99
Σχήμα 3.7. Διάγραμμα πιθανότητας χρεοκοπίας για κάθε λ_0	105

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1.1. Συνάρτηση πιθανότητας των κατανομών $G_4(0.6)$ και $G_4(0.8)$	34
Πίνακας 1.2. Παραγοντικές ροπές των κατανομών $G_4(0.6)$ και $G_4(0.8)$	35
Πίνακας 1.3. Συνάρτηση πυκνότητας της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, όπου $\mathbf{a} = (1/3, 2/3)$, $\mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$	47
Πίνακας 1.4. Ροπές της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, όπου $\mathbf{a} = (1/3, 2/3)$, $\mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$	48
Πίνακας 2.1. Η οικογένεια κατανομών του Panjer.....	55
Πίνακας 2.2. Πίνακας με τις κατανομές που ανήκουν στην κλάση κατανομών $\mathfrak{R}(a, b, 1)$	60-61
Πίνακας 2.3. Συνάρτηση πιθανότητας της $S_N \sim PH_d^{(25)}(\boldsymbol{\gamma}, \mathbf{C})$	70
Πίνακας 2.4. Συνάρτηση πυκνότητας της $S_N \sim PH_c^{(2)}(\boldsymbol{\gamma}, \mathbf{C})$	72
Πίνακας 2.5. Συνάρτηση κατανομής της S_N	74
Πίνακας 3.1. Ανανεωτική πυκνότητα $u(s)$ για την κατανομή $Erlang(2,2)$	99
Πίνακας 3.2. Πιθανότητα χρεοκοπίας $\psi(u)$	103

ΚΕΦΑΛΑΙΟ 1

Κατανομές τύπου φάσεων

1.1 Εισαγωγή

Προτού δοθούν οι ορισμοί και οι ιδιότητες των κατανομών τύπου φάσεων θα γίνει μια εκτενής αναφορά στις βασικές ιδιότητες των Μαρκοβιανών αλυσίδων τόσο σε διακριτό όσο και σε συνεχή χρόνο, καθώς οι κατανομές τύπου φάσεων συνδέονται άμεσα με αυτές τις αλυσίδες.

Μια **στοχαστική διαδικασία** (ή στοχαστική ανέλιξη) στο σύνολο $S \subset \mathbb{R}$ είναι μια οικογένεια τυχαίων μεταβλητών $\{X(t)\}_{t \in T}$ με $T \subset \mathbb{R}$, που είναι ορισμένες στον ίδιο χώρο πιθανότητας (Ω, \mathcal{F}, P) . Το σύνολο S στο οποίο παίρνουν τιμές οι τυχαίες μεταβλητές $X(t)$ ονομάζεται **χώρος καταστάσεων** και τα στοιχεία του S **καταστάσεις** της στοχαστικής διαδικασίας.

Αν το σύνολο T είναι αριθμήσιμο, π.χ. όταν $T = \mathbb{N}_0 = \{0, 1, 2, \dots\}$, ή όταν $T = \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, έχουμε **στοχαστική διαδικασία σε διακριτό χρόνο**. Τότε συνήθως γράφουμε $\{X_n\}_{n \geq 0}$, ή $\{X_n\}_{n \in \mathbb{Z}}$, χρησιμοποιώντας το n στη θέση του t και βάζοντάς το ως υποδείκτη. Αν το σύνολο T είναι μη αριθμήσιμο, π.χ. όταν $T = \mathbb{R} = (-\infty, \infty)$, ή όταν $T = [0, \infty)$, έχουμε **στοχαστική διαδικασία σε συνεχή χρόνο**. Τότε συνήθως γράφουμε $\{X(t)\}_{t \in \mathbb{R}}$, ή $\{X(t)\}_{t \geq 0}$. Το σύνολο S των καταστάσεων στις δύο προαναφερθείσες στοχαστικές διαδικασίες μπορεί και αυτό με τη σειρά του να είναι αριθμήσιμο ή μη αριθμήσιμο.

Μια στοχαστική ανέλιξη $\{X(t)\}_{t \in T}$ καλείται **Μαρκοβιανή διαδικασία** (ή Μαρκοβιανή ανέλιξη) αν

$$\begin{aligned} Pr(X(t_\nu) \leq x_\nu \mid X(t_{\nu-1}) = x_{\nu-1}, \dots, X(t_2) = x_2, X(t_1) = x_1) \\ = Pr(X(t_\nu) \leq x_\nu \mid X(t_{\nu-1}) = x_{\nu-1}) \end{aligned}$$

για κάθε επιλογή $t_1 < t_2 < \dots < t_{\nu-1} < t_\nu \in T$, και για κάθε $x_1, x_2, \dots, x_{\nu-1}, x_\nu \in S$.

Μια Μαρκοβιανή διαδικασία με αριθμήσιμο σύνολο καταστάσεων S καλείται **Μαρκοβιανή αλυσίδα**. Στη συνέχεια θα δώσουμε τις βασικές ιδιότητες των Μαρκοβιανών αλυσίδων σε διακριτό και σε συνεχή χρόνο.

1.2 Μαρκοβιανές αλυσίδες σε διακριτό χρόνο

1.2.1 Ορισμός και βασικές ιδιότητες

Χωρίς βλάβη της γενικότητας θεωρούμε ότι σε μια Μαρκοβιανή αλυσίδα διακριτού χρόνου έχουμε $T = \mathbb{N}_0 = \{0, 1, 2, \dots\}$ και φυσικά ο χώρος καταστάσεων S της αλυσίδας είναι αριθμήσιμο σύνολο. Τότε ένας τυπικός ορισμός της Μαρκοβιανής αλυσίδας διακριτού χρόνου είναι ο ακόλουθος.

Ορισμός 1.1

Μια στοχαστική διαδικασία διακριτού χρόνου $\{X_n\}_{n \geq 0}$ με αριθμήσιμο σύνολο καταστάσεων S ονομάζεται Μαρκοβιανή αλυσίδα διακριτού χρόνου αν έχει τη Μαρκοβιανή ιδιότητα, δηλαδή αν

$$Pr(X_{n+1} = i_{n+1} | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = Pr(X_{n+1} = i_{n+1} | X_n = i_n)$$

για κάθε $n = 0, 1, 2, \dots$ και $i_0, i_1, \dots, i_{n+1} \in S$.

Η πιθανότητα $Pr(X_{n+1} = j | X_n = i)$ ονομάζεται **πιθανότητα μετάβασης** (ή μεταπήδησης) **πρώτης τάξης**. Μια Μαρκοβιανή αλυσίδα λέγεται **ομογενής** αν για κάθε $n \geq 1$ και για κάθε $i, j \in S$,

$$Pr(X_{n+1} = j | X_n = i) = Pr(X_1 = j | X_0 = i) = p_{ij}.$$

Στη συνέχεια θα αναφερόμαστε πάντα σε ομογενείς Μαρκοβιανές αλυσίδες.

Ο πίνακας $\mathbf{P} = (p_{ij})$ ονομάζεται πίνακας πιθανοτήτων μετάβασης. Ο πίνακας \mathbf{P} είναι **στοχαστικός** αφού ικανοποιεί τις ιδιότητες:

I1. $p_{ij} \geq 0$, για κάθε $i, j \in S$.

I2. $\sum_{j \in S} p_{ij} = 1$, για κάθε $i \in S$.

Οι πιθανότητες

$$a_i = Pr(X_0 = i), \quad i \in S,$$

ονομάζονται **αρχικές πιθανότητες**. Φυσικά $\sum_{i \in S} a_i = 1$.

Η πιθανότητα μετάβασης από μια κατάσταση i σε μία άλλη κατάσταση j σε n βήματα ονομάζεται **πιθανότητα μετάβασης n – οστής τάξης** και συμβολίζεται ως $p_{ij}^{(n)}$, δηλαδή

$$p_{ij}^{(n)} = Pr(X_{n+m} = j | X_m = i) = Pr(X_n = j | X_0 = i)$$

για κάθε $i, j \in S$. Ο πίνακας που περιέχει τις πιθανότητες μετάβασης n – οστής τάξης συμβολίζεται με $\mathbf{P}^{(n)}$, δηλαδή $\mathbf{P}^{(n)} = (p_{ij}^{(n)})$. Προφανώς $\mathbf{P}^{(1)} = \mathbf{P}$.

Για τον υπολογισμό των πιθανοτήτων μετάβασης n – οστής τάξης μπορούν να χρησιμοποιηθούν οι **εξισώσεις Chapman-Kolmogorov**, σύμφωνα με τις οποίες για κάθε $i, j \in S$, και για κάθε $m, n \geq 0$,

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

όπου $p_{ij}^{(0)} = \delta_{ij}$ (δ_{ij} συμβολίζει το δέλτα του Kronecker). Πράγματι,

$$\begin{aligned} p_{ij}^{(m+n)} &= Pr(X_{m+n} = j | X_0 = i) \\ &= \sum_{k \in S} Pr(X_{m+n} = j | X_m = k, X_0 = i) \times Pr(X_m = k | X_0 = i) \\ &= \sum_{k \in S} Pr(X_{m+n} = j | X_m = k) \times Pr(X_m = k | X_0 = i) = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}. \end{aligned}$$

Από την παραπάνω σχέση προκύπτει άμεσα ότι

$$\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)} \times \mathbf{P}^{(n)}$$

και αφού $\mathbf{P}^{(1)} = \mathbf{P}$, έχουμε ότι

$$\mathbf{P}^{(n)} = \mathbf{P}^n, \quad n = 1, 2, \dots$$

Αν συμβολίζουμε με $p_j^{(n)}$ την πιθανότητα $Pr(X_n = j)$, $n \geq 0$ και $j \in S$, τότε

$$p_j^{(n)} = Pr(X_n = j) = \sum_{i \in S} Pr(X_n = j | X_0 = i) Pr(X_0 = i) = \sum_{i \in S} a_i p_{ij}^{(n)}.$$

1.2.2 Επαναληπτικές και μεταβατικές καταστάσεις

Έστω

$$T_j = \inf \{n \geq 1: X_n = j\}$$

ο χρόνος πρώτης επίσκεψης της κατάστασης $j \in S$. Η πιθανότητα πρώτης επίσκεψης της κατάστασης j , ξεκινώντας από την κατάσταση i , σε n βήματα συμβολίζεται με $f_{ij}^{(n)}$, δηλαδή

$$f_{ij}^{(n)} = Pr(T_j = n | X_0 = i) = Pr(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i), \quad n \geq 1$$

(αξιοματικά ορίζουμε ότι $f_{ij}^{(0)} = 0$, για κάθε $i, j \in S$).

Η πιθανότητα $f_i^{(n)} = f_{ii}^{(n)}$ είναι η **πιθανότητα πρώτης επανόδου** στην κατάσταση i , δηλαδή

$$f_i^{(n)} = Pr(T_i = n | X_0 = i) = Pr(X_n = i, X_{n-1} \neq i, \dots, X_1 \neq i | X_0 = i), \quad n \geq 1.$$

Η πιθανότητα να επισκεφτεί η Μαρκοβιανή αλυσίδα κάποτε την κατάσταση j , ξεκινώντας από την κατάσταση i , συμβολίζεται με f_{ij} , δηλαδή

$$f_{ij} = Pr(T_j < \infty | X_0 = i) = \sum_{n=1}^{\infty} Pr(T_j = n | X_0 = i) = \sum_{n=1}^{\infty} f_{ij}^{(n)}.$$

Η πιθανότητα $f_i = f_{ii}$ είναι η πιθανότητα να επισκεφτεί η Μαρκοβιανή αλυσίδα κάποτε την κατάσταση i , ξεκινώντας από την κατάσταση i , δηλαδή

$$f_{ii} = Pr(T_i < \infty | X_0 = i) = \sum_{n=1}^{\infty} Pr(T_i = n | X_0 = i) = \sum_{n=1}^{\infty} f_{ii}^{(n)}.$$

Η κατάσταση i καλείται **επαναληπτική** (ή επαναφερόμενη ή έμμονη) αν $f_{ii} = 1$, και **μεταβατική** (ή παροδική) αν $f_{ii} < 1$. Συνεπώς, με πιθανότητα 1, η Μαρκοβιανή αλυσίδα επιστρέφει σε μια επαναληπτική κατάσταση i σε πεπερασμένο αριθμό βημάτων, δηλαδή

$$Pr(T_i < \infty | X_0 = i) = 1,$$

ενώ για μια μεταβατική κατάσταση i έχουμε ότι

$$Pr(T_i = \infty | X_0 = i) = 1 - f_{ii} > 0.$$

Ο μέσος χρόνος επανόδου στην κατάσταση i είναι

$$m_i = E(T_i | X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

Επομένως αν m_i είναι ο μέσος αριθμός βημάτων της Μαρκοβιανής αλυσίδας μεταξύ δύο διαδοχικών επισκέψεων της κατάστασης $i \in S$, τότε $1/m_i$ είναι το μέρος του χρόνου που η Μαρκοβιανή αλυσίδα βρίσκεται στην κατάσταση $i \in S$.

Αν η κατάσταση i είναι επαναληπτική και $m_i < \infty$ τότε η κατάσταση i καλείται **θετικά επαναληπτική**, ενώ αν $m_i = \infty$ τότε καλείται **μηδενικά επαναληπτική**.

Οι γεννήτριες συναρτήσεις

$$P_{ij}(s) = \sum_{n=0}^{\infty} P_{ij}^{(n)} s^n, \quad F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^{(n)} s^n, \quad |s| < 1$$

όπου $f_{ij}^{(0)} = 0$ και $p_{ij}^{(0)} = \delta_{ij}$ για κάθε $i, j \in S$, ικανοποιούν τις σχέσεις

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}$$

και

$$P_{ij}(s) = F_{ij}(s)P_{jj}(s) = \frac{F_{ij}(s)}{1 - F_{ii}(s)}, \quad i \neq j.$$

Έστω τώρα N_j η τυχαία μεταβλητή που δηλώνει το συνολικό αριθμό επισκέψεων της Μαρκοβιανής αλυσίδας στην κατάσταση $j \in S$, δηλαδή

$$N_j = \sum_{n=0}^{\infty} 1\{X_n = j\}.$$

Τότε

$$E(N_j | X_0 = i) = \sum_{n=0}^{\infty} E[1\{X_n = j | X_0 = i\}] = \sum_{n=0}^{\infty} Pr(X_n = j | X_0 = i) = \sum_{n=0}^{\infty} p_{ij}^{(n)}.$$

Έτσι

$$E(N_j | X_0 = i) = \sum_{n=0}^{\infty} p_{ij}^{(n)} = \frac{f_{ij}}{1 - f_{ii}}, \quad i \neq j$$

και

$$E(N_i | X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1 - f_{ii}}.$$

Συνεπώς αν η κατάσταση i είναι επαναληπτική (μεταβατική) έχουμε ότι ο αναμενόμενος αριθμός επανόδων $E(N_i | X_0 = i)$ στην κατάσταση i είναι άπειρος (πεπερασμένος).

Επιπρόσθετα, η κατάσταση i είναι επαναληπτική, αν και μόνο αν

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty,$$

ενώ η κατάσταση i είναι μεταβατική, αν και μόνο αν

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty.$$

Επίσης, αν η κατάσταση j είναι μεταβατική, τότε

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$$

για κάθε $i \in S$, το οποίο συνεπάγεται ότι

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0.$$

Η συνάρτηση πιθανότητας $\boldsymbol{\pi} = (\pi_i : i \in S)$ καλείται **στάσιμη κατανομή** για τη Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ με πίνακα πιθανοτήτων μετάβασης \mathbf{P} , αν

$$\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$$

δηλαδή αν

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \quad j \in S.$$

Αν υποθέσουμε ότι η αρχική κατανομή της Μαρκοβιανής αλυσίδας $\{X_n\}_{n \geq 0}$ είναι η στάσιμη κατανομή, και χρησιμοποιώντας τη σχέση $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}^n$, παίρνουμε ότι $Pr(X_n = i) = \pi_i$ για κάθε $n > 0$, δηλαδή η $\{X_n\}_{n \geq 0}$ παραμένει **αναλλοίωτη** στο χρόνο.

Σύμφωνα με το βασικότερο οριακό θεώρημα που ισχύει στις Μαρκοβιανές αλυσίδες έχουμε ότι μια αδιαχώριστη Μαρκοβιανή αλυσίδα με απεριοδικές θετικά επαναληπτικές καταστάσεις ισχύει ότι

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{1}{\sum_{n=0}^{\infty} n f_{ii}^{(n)}} = \frac{1}{m_i} = \pi_i$$

και

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = \lim_{n \rightarrow \infty} p_{ii}^{(n)}, \quad j \in S.$$

(για περισσότερες λεπτομέρειες δείτε Taylor & Karlin (1998)).

1.2.3 Απορροφητικές Μαρκοβιανές αλυσίδες

Μια κατάσταση $i \in S$ μιας Μαρκοβιανής αλυσίδας ονομάζεται **απορροφητική** αν $p_{ii} = 1$. Μια Μαρκοβιανή αλυσίδα που περιέχει μόνο απορροφητικές και μεταβατικές καταστάσεις καλείται **απορροφητική Μαρκοβιανή αλυσίδα**. Οι απορροφητικές Μαρκοβιανές αλυσίδες παίζουν σημαντικό ρόλο στην παρούσα εργασία.

Χωρίς βλάβη της γενικότητας, ας υποθέσουμε ότι μια απορροφητική Μαρκοβιανή αλυσίδα έχει μόνο μία απορροφητική κατάσταση και πεπερασμένο σύνολο καταστάσεων S . Πιο συγκεκριμένα, έστω $\{X_n\}_{n \geq 0}$ μία Μαρκοβιανή αλυσίδα με σύνολο καταστάσεων $S = \{1, 2, \dots, m, m+1\}$, όπου η κατάσταση $\{m+1\}$ είναι η απορροφητική και οι καταστάσεις $\{1, 2, \dots, m\}$ είναι μεταβατικές. Ο πίνακας πιθανοτήτων μετάβασης μιας τέτοιας Μαρκοβιανής αλυσίδας έχει τη μορφή

$$P = \begin{pmatrix} T & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}$$

όπου στον παραπάνω πίνακα $T = (t_{ij})_{m \times m}$ είναι ένας τετραγωνικός πίνακας διάστασης $m \times m$ που περιέχει τις πιθανότητες μετάβασης (πρώτης τάξης) μεταξύ των μεταβατικών καταστάσεων, $\mathbf{t} = (t_1, t_2, \dots, t_m)'$ είναι ένα διάνυσμα στήλη διάστασης $m \times 1$ που περιέχει τις πιθανότητες μετάβασης από τις μεταβατικές καταστάσεις στην απορροφητική κατάσταση, και το $\mathbf{0}$ είναι ένα διάνυσμα γραμμή διάστασης $1 \times m$ με όλα τα στοιχεία του ίσα με μηδέν (0). Προφανώς ισχύει ότι $t_{ij} \geq 0$ και $t_i \geq 0$ για $1 \leq i, j \leq m$, με τουλάχιστον ένα $t_i > 0$.

Για τον πίνακα T έχουμε ότι

$$\lim_{n \rightarrow \infty} T^n = \mathbf{0}$$

αφού όπως έχουμε αναφέρει για μια μεταβατική κατάσταση j ισχύει ότι

$$\lim_{n \rightarrow \infty} t_{ij}^{(n)} = 0$$

για κάθε $i \in S$. Έτσι μπορεί να αποδειχθεί ότι ο πίνακας $I - T$ έχει αντίστροφο, και μάλιστα

$$(\mathbf{I} - \mathbf{T})^{-1} = \mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \mathbf{T}^3 + \dots = \sum_{n=0}^{\infty} \mathbf{T}^n$$

(δείτε, π.χ., Kemeny, Snell & Knapp (1976)).

Ο πίνακας $\mathbf{U} = (u_{ij})_{m \times m} = (\mathbf{I} - \mathbf{T})^{-1}$ ονομάζεται **θεμελιώδης πίνακας**. Τα στοιχεία του θεμελιώδη πίνακα $(\mathbf{I} - \mathbf{T})^{-1}$ έχουν μια πολύ ενδιαφέρουσα ερμηνεία που δίνεται ακολούθως. Ορίζουμε την τ.μ.

$$Y = \inf \{n \geq 0: X_n = m + 1\}$$

η οποία δηλώνει τον αριθμό των βημάτων της Μαρκοβιανής αλυσίδας έως ότου αυτή εισέλθει στην απορροφητική της κατάσταση $\{m + 1\}$. Έστω επίσης E_{ij} ($1 \leq i, j \leq m$) ο μέσος αριθμός επισκέψεων στη μεταβατική κατάσταση j ξεκινώντας από τη μεταβατική κατάσταση i (προτού φυσικά εισέλθει η Μαρκοβιανή αλυσίδα στην απορροφητική της κατάσταση). Τότε

$$\begin{aligned} E_{ij} &= E(N_j | X_0 = i) = \sum_{n=0}^{\infty} E[1\{X_n = j | X_0 = i\}] \\ &= E\left(\sum_{n=0}^{Y-1} 1\{X_n = j | X_0 = i\}\right) = ((\mathbf{I} - \mathbf{T})^{-1})_{ij} = u_{ij} \end{aligned}$$

(για την απόδειξη δείτε, π.χ., Taylor & Karlin (1998)).

Η ύπαρξη του πίνακα $(\mathbf{I} - \mathbf{T})^{-1}$ συνεπάγεται ότι η τ.μ. Y είναι **μη-ελαττωματική** (non-defective) με την έννοια ότι $F_Y(\infty) = 1$ (για περισσότερες λεπτομέρειες δείτε Resnick (1992) και Bladt & Nielsen (2017)).

1.2.4 Εφαρμογές

Εφαρμογή 1 – Δικατάστατες Μαρκοβιανές αλυσίδες: Έστω μια Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ με χώρο καταστάσεων $S = \{0, 1\}$ και πίνακα πιθανοτήτων μετάβασης

$$\mathbf{P} = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}$$

όπου $0 < a, b < 1$. Για τον υπολογισμό του πίνακα $\mathbf{P}^{(n)} = \mathbf{P}^n$ έχουμε ότι

$$\mathbf{P}^n = \mathbf{Q} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \mathbf{Q}^{-1}$$

όπου λ_1 και λ_2 οι ιδιοτιμές του πίνακα \mathbf{P} , και \mathbf{Q} ο πίνακας των ιδιοδιανυσμάτων. Έχουμε ότι $\lambda_1 = 1, \lambda_2 = 1 - a - b$ και

$$\mathbf{Q} = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix}.$$

Έτσι

$$\mathbf{P}^n = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

Στην ειδική περίπτωση $a = 1 - b$, τότε οι γραμμές του \mathbf{P} είναι ίσες, και οι τυχαίες μεταβλητές X_1, X_2, \dots είναι ανεξάρτητες και ισόνομες με $Pr(X_n = 0) = 1 - a$ και $Pr(X_n = 1) = a$, δηλαδή έχουμε μια ακολουθία δοκιμών Bernoulli με πιθανότητα επιτυχίας ίση με a .

Δεδομένου ότι $|1 - a - b| < 1$ όταν $0 < a, b < 1$, τότε $|1 - a - b|^n \rightarrow 0$ όταν το $n \rightarrow \infty$ και επομένως

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}.$$

Έτσι το σύστημα, μετά από αρκετά μεγάλο αριθμό βημάτων, θα βρίσκεται στην κατάσταση 0 με πιθανότητα $\frac{b}{a+b}$ και στην κατάσταση 1 με πιθανότητα $\frac{a}{a+b}$, ανεξάρτητα από την αρχική κατάσταση που ξεκίνησε το σύστημα. Η συνάρτηση πιθανότητας $\boldsymbol{\pi} = (\frac{b}{a+b}, \frac{a}{a+b})$ αποτελεί **στάσιμη κατανομή** για τη Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ με πίνακα πιθανοτήτων μετάβασης \mathbf{P} .

Για παράδειγμα ας υποθέσουμε ότι ένας κατασκευαστής παράγει ειδικά εξαρτήματα για μία βιομηχανία αυτοκινήτων τα οποία κρίνονται ως «ελαττωματικά» ή «μη ελαττωματικά». Κατά πόσο ένα εξάρτημα θεωρείτε «ελαττωματικό» ή «μη ελαττωματικό», εξαρτάται εν μέρει από το εάν ή όχι τα εξαρτήματα που είχαν κατασκευαστεί στο παρελθόν ήταν «ελαττωματικά» ή «μη ελαττωματικά». Έστω X_n η κατηγορία του n -οστού εξαρτήματος, με $X_n = 0$ να δηλώνει «μη ελαττωματικό» και $X_n = 1$ να δηλώνει «ελαττωματικό». Αν η $\{X_n\}_{n \geq 1}$ περιγράφεται από μία Μαρκοβιανή αλυσίδα με πίνακα μετάβασης

$$\mathbf{P} = \begin{pmatrix} 0.99 & 0.01 \\ 0.12 & 0.88 \end{pmatrix}$$

τότε, μακροπρόθεσμα, η πιθανότητα το παραγόμενο εξάρτημα να είναι «μη ελαττωματικό» ισούται με $\frac{b}{a+b} = \frac{0.12}{0.01+0.12} = 0.924 = 92.4\%$, ενώ η πιθανότητα το παραγόμενο εξάρτημα να είναι «ελαττωματικό» ισούται με $\frac{a}{a+b} = \frac{0.01}{0.01+0.12} = 0.076 = 7.6\%$.

Εφαρμογή 2: Έστω $\{b_n\}_{n \geq 1}$ μια ακολουθία ανεξάρτητων δοκιμών Bernoulli με

$$Pr(b_n = 1) = p, \quad Pr(b_n = 0) = q, \quad n = 1, 2, \dots$$

Για $n = 2, 3, \dots$ ορίζουμε την στοχαστική ανέλιξη

$$X_n = \begin{cases} 0, & \text{αν } b_{n-1} = 1, b_n = 1 \\ 1, & \text{αν } b_{n-1} = 1, b_n = 0 \\ 2, & \text{αν } b_{n-1} = 0, b_n = 1 \\ 3, & \text{αν } b_{n-1} = 0, b_n = 0. \end{cases}$$

Η στοχαστική ανέλιξη $\{X_n\}_{n \geq 2}$ έχει χώρο καταστάσεων $S = \{0, 1, 2, 3\}$. Αν γνωρίζουμε την κατάσταση της X_n , που εξαρτάται από τις τυχαίες μεταβλητές (b_{n-1}, b_n) , τότε μπορούμε να προσδιορίσουμε την κατανομή της X_{n+1} που εξαρτάται από τις τυχαίες μεταβλητές (b_n, b_{n+1}) . Άρα η $\{X_n\}_{n \geq 2}$ είναι Μαρκοβιανή αλυσίδα.

Για τις πιθανότητες μετάβασης πρώτης τάξης έχουμε

$$\begin{aligned} p_{00} &= Pr(X_{n+1} = 0 | X_n = 0) = Pr(b_{n+1} = b_n = 1 | b_{n-1} = b_n = 1) \\ &= \frac{Pr(b_{n-1} = b_n = b_{n+1} = 1)}{Pr(b_{n-1} = b_n = 1)} = Pr(b_{n+1} = 1) = p, \end{aligned}$$

$$p_{01} = Pr(X_{n+1} = 1 | X_n = 0) = Pr(b_{n+1} = 1) = q,$$

$$p_{02} = 0, \quad p_{03} = 0.$$

Δουλεύοντας ανάλογα και με τις υπόλοιπους συνδυασμούς βρίσκουμε ότι

$$P = \begin{pmatrix} p & q & 0 & 0 \\ 0 & 0 & p & q \\ p & q & 0 & 0 \\ 0 & 0 & p & q \end{pmatrix}.$$

Χρησιμοποιώντας τον παραπάνω πίνακα και τις εξισώσεις Chapman-Kolmogorov μπορούμε να υπολογίσουμε διάφορες πιθανότητες. Για παράδειγμα,

$$\begin{aligned} p_{00}(3) &= \sum_{k=0}^3 p_{0k} p_{k0}(2) = p_{00} p_{00}(2) + p_{01} p_{10}(2) \\ &= p_{00}(p_{00} p_{00} + p_{01} p_{10}) + p_{01}(p_{12} p_{20} + p_{13} p_{30}) = pp^2 + qp^2 = p^2. \end{aligned}$$

1.3 Μαρκοβιανές αλυσίδες σε συνεχή χρόνο

1.3.1 Ορισμός και βασικές ιδιότητες

Ένας τυπικός ορισμός μιας Μαρκοβιανής αλυσίδας συνεχούς χρόνου με σύνολο καταστάσεων S είναι ο ακόλουθος.

Ορισμός 1.2

Μια στοχαστική διαδικασία συνεχούς χρόνου $\{X(t)\}_{t \geq 0}$ με αριθμήσιμο σύνολο καταστάσεων S ονομάζεται Μαρκοβιανή αλυσίδα συνεχούς χρόνου αν έχει τη Μαρκοβιανή ιδιότητα, δηλαδή αν για κάθε ακολουθία χρόνων $0 \leq t_1 < t_2 < \dots < t_n$ και για κάθε $i_1, i_2, \dots, i_n \in S$,

$$Pr(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \dots, X(t_1) = i_1) = Pr(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}).$$

Η πιθανότητα $Pr(X(t_n) = i_n | X(t_{n-1}) = i_{n-1})$ ονομάζεται **πιθανότητα μετάβασης** (ή μεταπήδησης). Μια Μαρκοβιανή αλυσίδα λέγεται **ομογενής** αν για κάθε $s, t \geq 0$ και για κάθε $i, j \in S$,

$$Pr(X(s+t) = j | X(s) = i) = Pr(X(t) = j | X(0) = i) = p_{ij}(t).$$

Στη συνέχεια θα αναφερόμαστε πάντα σε ομογενείς Μαρκοβιανές αλυσίδες.

Ο πίνακας $\mathbf{P}(t) = (p_{ij}(t))$ ονομάζεται πίνακας πιθανοτήτων μετάβασης. Ο πίνακας $\mathbf{P}(t)$ είναι **στοχαστικός** πίνακας αφού για κάθε $t \geq 0$ ικανοποιεί τις δύο ακόλουθες ιδιότητες:

- I1. $p_{ij}(t) \geq 0$, για κάθε $i, j \in S$.
- I2. $\sum_{j \in S} p_{ij}(t) = 1$, για κάθε $i \in S$.

Συνήθως ορίζουμε ότι $\mathbf{P}(0) = \mathbf{I}$.

Οι εξισώσεις **Chapman-Kolmogorov** σε συνεχή χρόνο έχουν την παρακάτω μορφή

$$p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s)p_{kj}(t)$$

για κάθε $i, j \in S$ και για κάθε $s, t \geq 0$. Πράγματι

$$\begin{aligned} p_{ij}(s+t) &= Pr(X(s+t) = j | X(0) = i) \\ &= \sum_{k \in S} Pr(X(s+t) = j | X(s) = k, X(0) = i) \times Pr(X(s) = k | X(0) = i) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \in S} Pr(X(s+t) = j | X(s) = k) \times Pr(X(s) = k | X(0) = i) \\
&= \sum_{k \in S} p_{ik}(s) p_{kj}(t)
\end{aligned}$$

Σε μορφή πινάκων οι παραπάνω εξισώσεις, για κάθε $s, t \geq 0$, γράφονται ως

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t).$$

Αν

$$a_i = Pr(X(0) = i), \quad i \in S$$

είναι οι αρχικές πιθανότητες της Μαρκοβιανής αλυσίδας $\{X(t)\}_{t \geq 0}$, τότε

$$Pr(X(t) = j) = \sum_{i \in S} a_i p_{ij}(t), \quad t \geq 0.$$

1.3.2 Γεννήτορας πίνακας, εμφυτευμένη Μαρκοβιανή αλυσίδα και είδη καταστάσεων

Από εδώ και στο εξής θεωρούμε ότι η Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ έχει πεπερασμένο σύνολο καταστάσεων S αφού οι κατανομές τύπου φάσεων που διαπραγματεύονται στην παρούσα εργασία αναφέρονται σε τέτοιου είδους Μαρκοβιανές αλυσίδες. Αυτή η υπόθεση εγγυάται την ισχύ ορισμένων αποτελεσμάτων που θα δοθούν στη συνέχεια.

Ορίζουμε αξιωματικά ότι

$$\lim_{h \downarrow 0} \mathbf{P}(h) = \mathbf{P}(0) = \mathbf{I}$$

δηλαδή

$$\lim_{h \downarrow 0} p_{ij}(t) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Από την παραπάνω σχέση και τις εξισώσεις Chapman-Kolmogorov, συνεπάγεται ότι ο πίνακας $\mathbf{P}(t) = (p_{ij}(t))$ είναι συνεχής για κάθε $t \geq 0$, αφού

$$\lim_{h \rightarrow 0} \mathbf{P}(t+h) = \mathbf{P}(t), \quad t > 0$$

(δείτε, π.χ., Taylor & Karlin (1998) και Norris (1997)). Σε αυτή την περίπτωση ο πίνακας $\mathbf{P}(t)$ ικανοποιεί τη σχέση

$$\lim_{h \downarrow 0} \frac{\mathbf{P}(h) - \mathbf{I}}{h} = \mathbf{P}'(0) = \mathbf{G} = (g_{ij})$$

όπου

$$g_{ii} = -q_i = \lim_{h \downarrow 0} \frac{p_{ii}(h) - 1}{h}, \quad i \in S$$

και

$$g_{ij} = q_{ij} = \lim_{h \downarrow 0} \frac{p_{ij}(h)}{h}, \quad i \neq j, \quad i, j \in S.$$

Ο πίνακας \mathbf{G} ονομάζεται **(απειροστός) γεννήτορας πίνακας** ((infitesimal) generator matrix) της Μαρκοβιανής αλυσίδας. Γενικά έχουμε ότι $0 \leq g_{ij} < \infty$ για $i \neq j$ και $g_{ii} \leq 0$. Επιπλέον

$$\sum_{j \in S} g_{ij} = 0,$$

δηλαδή $\mathbf{G}\mathbf{1} = \mathbf{0}$ (τότε η Μαρκοβιανή αλυσίδα καλείται **συντηρητική**). Αν το σύνολο καταστάσεων S δεν είναι πεπερασμένο τότε $\sum_{j \in S} g_{ij} \leq 0$.

Από τα παραπάνω έχουμε ότι για μικρό h

$$p_{ij}(h) = q_{ij}h + o(h), \quad i \neq j, \quad i, j \in S,$$

$$p_{ii}(h) = 1 - q_i h + o(h), \quad i \in S.$$

Οι πίνακες \mathbf{G} και $\mathbf{P}(t)$ ικανοποιούν τη διαφορική εξίσωση

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{G} = \mathbf{G}\mathbf{P}(t)$$

με αρχική συνθήκη την $\mathbf{P}(0) = \mathbf{I}$. Η πρώτη εξίσωση $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{G}$ αναφέρεται ως **προδρομική εξίσωση Kolmogorov** ενώ η εξίσωση $\mathbf{P}'(t) = \mathbf{G}\mathbf{P}(t)$ ως **οπισθοδρομική εξίσωση Kolmogorov**. Μπορεί ναδειχθεί ότι μοναδική λύση των παραπάνω εξισώσεων είναι η

$$\mathbf{P}(t) = e^{\mathbf{G}t} = \sum_{n=0}^{\infty} \mathbf{G}^n \frac{t^n}{n!}, \quad t \geq 0.$$

Σε κάθε Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{X(t)\}_{t \geq 0}$ με σύνολο καταστάσεων S αντιστοιχεί μια **εμφυτευμένη** Μαρκοβιανή αλυσίδα (embedded or Jump chain) διακριτού χρόνου $\{Y_n\}_{n \geq 0}$ με το ίδιο σύνολο καταστάσεων και τιμές που ορίζονται ως εξής: Ας

υποθέσουμε ότι $0 = t_0 < t_1 < t_2 < \dots < t_n < t_{n+1} < \dots$ είναι οι χρονικές στιγμές που αλλάζει κατάσταση η Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{X(t)\}_{t \geq 0}$. Τότε

$$Y_0 = X(0) = X(t_0), Y_1 = X(t_1), Y_2 = X(t_2), \dots, Y_n = X(t_n), Y_{n+1} = X(t_{n+1}), \dots$$

δηλαδή

$$Y_n = X(s), t_n \leq s < t_{n+1}.$$

Ενδιαφέρουν παρουσιάζουν οι πιθανότητες μετάβασης

$$p_{ij} = Pr(Y_{n+1} = j | Y_n = i), \quad i \neq j$$

της εμφυτευμένης Μαρκοβιανής αλυσίδας που θα δοθούν στην επόμενη παράγραφο. Σημειώνουμε εδώ όμως ότι $p_{ii} = 0$ για κάθε $i \in S$, εκτός αν $g_{ii} = -q_i = 0$, οπότε η κατάσταση i είναι απορροφητική, και τότε ορίζουμε ότι $p_{ii} = 1$ (αν $q_i = 0$ τότε από τη στιγμή που θα εισέλθει η Μαρκοβιανή $\{X(t)\}_{t \geq 0}$ στην κατάσταση i παραμένει σε αυτή για πάντα).

Σε μια Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{X(t)\}_{t \geq 0}$, η κατάσταση $i \in S$ ονομάζεται επαναληπτική αν

$$Pr(\text{το σύνολο } \{t \geq 0: X(t) = i\} \text{ είναι μη φραγμένο}) = 1$$

και μεταβατική αν

$$Pr(\text{το σύνολο } \{t \geq 0: X(t) = i\} \text{ είναι μη φραγμένο}) = 0.$$

Αν η κατάσταση $i \in S$ είναι επαναληπτική (μεταβατική) για την εμφυτευμένη Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$, τότε είναι επαναληπτική (μεταβατική) και για τη Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{X(t)\}_{t \geq 0}$. Έστω

$$T_i = \inf \{t > 0: X(t) = i\}$$

Αν $q_i = 0$, ή ισοδύναμα $P(T_i < \infty | X(0) = i) = 1$, τότε η κατάσταση $i \in S$ είναι επαναληπτική και $\int_0^\infty p_{ii}(t) dt = \infty$, ενώ αν $q_i > 0$ και $P(T_i < \infty | X(0) = i) < 1$, τότε η κατάσταση $i \in S$ είναι μεταβατική και $\int_0^\infty p_{ii}(t) dt < \infty$.

1.3.3 Ιδιότητες του γεννήτορα πίνακα και στάσιμη κατανομή

Ας υποθέσουμε ότι κάποια χρονική στιγμή η Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ εισέρχεται στην κατάσταση $i \in S$ όπου S πεπερασμένο σύνολο. Χωρίς βλάβη της γενικότητας μπορούμε

να υποθέσουμε ότι η είσοδος γίνεται τη χρονική στιγμή $t = 0$, δηλαδή $X(0) = i$. Έστω T_i ο χρόνος παραμονής της Μαρκοβιανής αλυσίδας σε αυτή την κατάσταση. Τότε

$$\begin{aligned} Pr(T_i > s + t | T_i > s) &= Pr(X(u) = i \text{ για } u \in [0, s + t] | X(u) = i \text{ για } u \in [0, s]) \\ &= Pr(X(u) = i \text{ για } u \in (s, s + t] | X(u) = i \text{ για } u \in [0, s]) \\ &= Pr(X(u) = i \text{ για } u \in (s, s + t] | X(s) = i) \\ &= Pr(X(u) = i \text{ για } u \in (0, t] | X(0) = i) = Pr(T_i > t). \end{aligned}$$

Συνεπώς ο χρόνος παραμονής (holding time) σε οποιαδήποτε κατάσταση $i \in S$ ακολουθεί εκθετική κατανομή και μπορεί να αποδειχθεί ότι

$$T_i \sim \text{Exp}(q_i)$$

όπου $q_i = -g_{ii}$. Αφού $\mathbf{G}\mathbf{1} = \mathbf{0}$ έχουμε ότι

$$q_i = \sum_{j \neq i} q_{ij}, \quad i \in S.$$

Επίσης μπορεί να δειχθεί ότι ο χρόνος T_{ij} που δηλώνει το χρόνο που μεσολαβεί από τη χρονική στιγμή που η Μαρκοβιανή αλυσίδα εισέλθει στην κατάσταση $i \in S$ έως ότου εξέλθει από αυτή, λόγω της εισόδου της στην κατάσταση $j \in S$ ($j \neq i$), ακολουθεί εκθετική κατανομή με παράμετρο q_{ij} , δηλαδή

$$T_{ij} \sim \text{Exp}(q_{ij}), \quad i \neq j, \quad i, j \in S.$$

Συνεπώς μπορούμε να γράψουμε ότι

$$T_i = \min_{\substack{j \in S \\ j \neq i}} T_{ij}.$$

Τέλος η πιθανότητα

$$p_{ij} = Pr(X_{n+1} = j | X_n = i) = Pr(X_1 = j | X_0 = i), \quad i \neq j$$

ικανοποιεί τη σχέση

$$p_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & i \neq j, \quad q_i > 0 \\ 0, & i \neq j, \quad q_i = 0. \end{cases}$$

Επίσης

$$p_{ii} = \begin{cases} 0, & q_i > 0 \\ 1, & q_i = 0. \end{cases}$$

Η συνάρτηση πιθανότητας $\boldsymbol{\pi} = (\pi_i: i \in S)$ καλείται **στάσιμη κατανομή** για τη Μαρκοβιανή αλυσίδα $\{X_t\}_{t \geq 0}$ με πίνακα πιθανοτήτων μετάβασης $\mathbf{P}(t)$ αν

$$\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}(t).$$

Η συνάρτηση πιθανότητας $\boldsymbol{\pi} = (\pi_i: i \in S)$ είναι στάσιμη κατανομή αν και μόνο αν

$$\boldsymbol{\pi} \mathbf{G} = \mathbf{0}$$

ή ισοδύναμα

$$\pi_j q_j = \sum_{\substack{i \in S \\ i \neq j}} \pi_i q_{ij}, \quad j \in S.$$

Αν υπάρχει το όριο

$$\pi_j = \lim_{t \rightarrow \infty} P(X(t) = j | X(0) = i) = \lim_{t \rightarrow \infty} p_{ij}(t), \quad i, j \in S$$

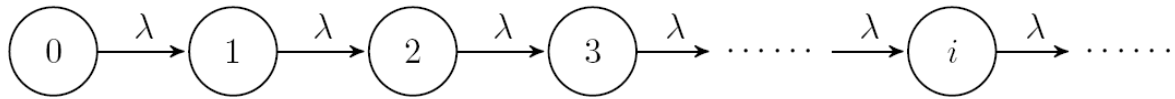
τότε το διάνυσμα $\boldsymbol{\pi} = (\pi_j: j \in S)$ είναι μια **στάσιμη κατανομή** για τη Μαρκοβιανή αλυσίδα $\{X_t\}_{t \geq 0}$. Επιπρόσθετα μια αδιαχώριστη Μαρκοβιανή αλυσίδα $\{X_t\}_{t \geq 0}$ που έχει στάσιμη κατανομή $\boldsymbol{\pi}$ και πίνακα \mathbf{G} διαγωνοποιήσιμο, τότε

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \lim_{t \rightarrow \infty} P(X(t) = j | X(0) = i) = \pi_j, \quad i, j \in S.$$

1.3.5 Εφαρμογές

Εφαρμογή 1 – Διαδικασία Poisson:

Η διαδικασία Poisson με μέσο ρυθμό λ ($0 < \lambda < \infty$) είναι μια Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ συνεχούς χρόνου με σύνολο καταστάσεων $S = \{0, 1, 2, \dots\}$ όπου οι χρόνοι παραμονής $S_j, j \geq 1$, στην κατάσταση $j - 1$ είναι ανεξάρτητες τυχαίες μεταβλητές που κάθε μία ακολουθεί εκθετική κατανομή με παράμετρο λ , και για την εμφυτευμένη Μαρκοβιανή αλυσίδα διακριτού χρόνου $\{Y_n\}_{n \geq 0}$ ισχύει ότι $Y_n = n$. Ενδεικτικό είναι το ακόλουθο σχήμα που περιγράφει την κίνηση της αλυσίδας μεταξύ των καταστάσεων.



Σχήμα 1.1. Διάγραμμα καταστάσεων διαδικασίας Poisson

Ο γεννήτορας πίνακας της αλυσίδας είναι ο ακόλουθος

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ 0 & 0 & -\lambda & \lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

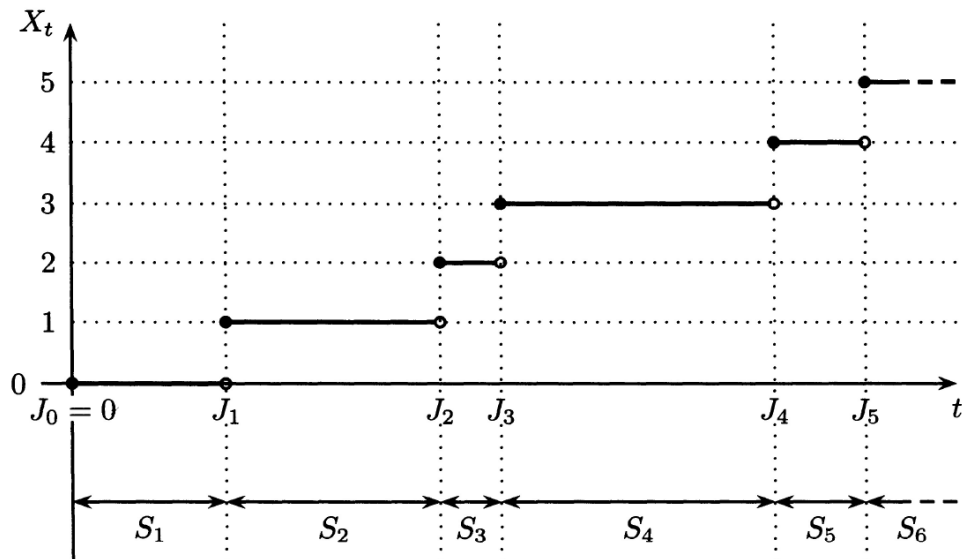
Η τυχαία μεταβλητή

$$J_n = S_1 + S_2 + \dots + S_n, \quad n = 1, 2, \dots$$

που δηλώνει τη χρονική στιγμή κατά την οποία γίνεται n -οστή αλλαγή κατάσταση, ή ισοδύναμα η χρονική στιγμή εμφάνισης του n -οστού συμβάντος, έχει την κατανομή Erlang με παραμέτρους n και λ , δηλαδή έχει συνάρτηση πυκνότητας

$$f_{J_n}(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{(n-1)!}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Αξιοματικά ορίζουμε ότι $J_0 = 0$. Ενδεικτικό είναι το ακόλουθο σχήμα



Σχήμα 1.2. Διαγραμματική περιγραφή της διαδικασίας Poisson

Σχετικά με τον υπολογισμό των πιθανοτήτων $P_{ij}(t)$ από τη διαφορική εξίσωση $P'(t) = QP(t)$, προκύπτει ότι

$$P'_{ij}(t) = -\lambda P_{ij}(t) + \lambda P_{i+1,j}(t), \quad i, j \in S.$$

Από τη λύση του παραπάνω συστήματος προκύπτει ότι

$$P_{ij}(t) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}, & j \geq i, \\ 0, & j < i. \end{cases}$$

Μια στοχαστική διαδικασία $\{X(t)\}_{t \geq 0}$ θα λέμε ότι έχει **ανεξάρτητες προσauξήσεις** αν για κάθε $n = 1, 2, 3, \dots$ και $0 \leq t_0 < t_1 < \dots < t_n$ οι τυχαίες μεταβλητές $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ είναι ανεξάρτητες. Θα λέμε ότι έχει **στάσιμες προσauξήσεις** αν για $h > 0$ η από κοινού κατανομή των τυχαίων μεταβλητών $X_{t_1+h} - X_{t_0+h}, X_{t_2+h} - X_{t_1+h}, \dots, X_{t_n+h} - X_{t_{n-1}+h}$ δεν εξαρτάται από το h .

Το βασικότερο αποτέλεσμα που ισχύει σε μια διαδικασία Poisson είναι η ισοδυναμία των τριών ακόλουθων προτάσεων:

(α) Οι χρόνοι παραμονής $S_j, j \geq 1$, στην κατάσταση $j - 1$ είναι ανεξάρτητες τυχαίες μεταβλητές που κάθε μία ακολουθεί εκθετική κατανομή με παράμετρο λ και για την εμφυτευμένη Μαρκοβιανή αλυσίδα διακριτού χρόνου $\{Y_n\}_{n \geq 0}$ ισχύει ότι $Y_n = n$.

(β) Η Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ έχει ανεξάρτητες προσauξήσεις και

$$P(X_{t+h} - X_t = 0) = 1 - \lambda h + o(h),$$

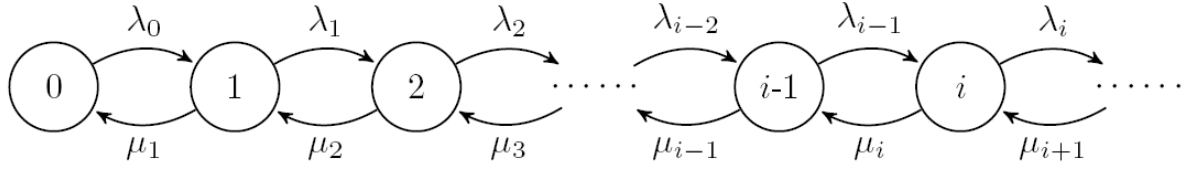
$$P(X_{t+h} - X_t = 1) = \lambda h + o(h).$$

(γ) Η Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ έχει ανεξάρτητες και στάσιμες προσauξήσεις και για κάθε $t > 0$ η τυχαία μεταβλητή X_t έχει κατανομή Poisson με παράμετρο λt .

Εφαρμογή 2 – Διαδικασίες γεννήσεως θανάτου

Πολλές Μαρκοβιανές αλυσίδες συνεχούς χρόνου μπορούν να μεταβούν μόνο σε γειτονικές καταστάσεις, δηλαδή από την κατάσταση i μετάβαση γίνεται είτε στην κατάσταση $i + 1$ (γέννηση), είτε στην κατάσταση $i - 1$ (θάνατος).

Ενώ η αλυσίδα βρίσκεται στην κατάσταση i ο ρυθμός γέννησης δηλώνεται με λ_i , ενώ ο ρυθμός θανάτου με μ_i . Ενδεικτικό είναι το ακόλουθο σχήμα.



Σχήμα 1.3. Διάγραμμα καταστάσεων διαδικασία γεννήσεως θανάτου

Ο γεννήτορας πίνακας της αλυσίδας είναι ο ακόλουθος

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Για τη στάσιμη κατανομή $\pi = (\pi_i; i \in S)$, η εξίσωση $\pi G = \mathbf{0}$ δίνει το σύστημα των εξισώσεων

$$-\lambda_0 \pi_0 + \mu_1 \pi_1 = 0,$$

$$\lambda_{n-1} \pi_{n-1} - (\lambda_n + \mu_n) \pi_n + \mu_{n+1} \pi_{n+1} = 0.$$

Λύνοντας το σύστημα παίρνουμε

$$\pi_0 = \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} \right)^{-1},$$

$$\pi_n = \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} \pi_0, \quad n \geq 1,$$

υπό την προϋπόθεση ότι

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} < \infty.$$

1.4 Κατανομές τύπου φάσεων

1.4.1 Διακριτές κατανομές τύπου φάσεων

Μια διακριτή **κατανομή τύπου φάσεων** (phase type distribution) είναι η κατανομή του αριθμού των βημάτων που απαιτούνται για να εισέλθει στην απορροφητική της κατάσταση μια ομογενής απορροφητική Μαρκοβιανή αλυσίδα διακριτού χρόνου με πεπερασμένο σύνολο

καταστάσεων S (για τις απορροφητικές Μαρκοβιανές αλυσίδες διακριτού χρόνου έχουμε ήδη αναφέρει κάποια αποτελέσματα στην Παράγραφο 1.2.3).

Πιο συγκεκριμένα, έστω $\{X_n\}_{n \geq 0}$ μία ομογενής Μαρκοβιανή αλυσίδα διακριτού χρόνου με σύνολο καταστάσεων $S = \{1, 2, \dots, m, m+1\}$, όπου οι καταστάσεις $\{1, 2, \dots, m\}$ είναι μεταβατικές (παροδικές) και η κατάσταση $\{m+1\}$ είναι απορροφητική. Ο πίνακας πιθανοτήτων μετάβασης μιας τέτοιας Μαρκοβιανής αλυσίδας έχει τη μορφή

$$P = \begin{pmatrix} T & t \\ \mathbf{0} & 1 \end{pmatrix}$$

όπου $T = (t_{ij})_{m \times m}$ είναι ένας τετραγωνικός πίνακας διάστασης $m \times m$ που περιέχει τις πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων, $t = (t_1, t_2, \dots, t_m)'$ είναι ένα διάνυσμα στήλη διάστασης $m \times 1$ που περιέχει τις πιθανότητες μετάβασης από τις παροδικές καταστάσεις στην απορροφητική κατάσταση, και το $\mathbf{0}$ συμβολίζει το διάνυσμα γραμμή διάστασης $1 \times m$ με όλα τα στοιχεία του ίσα με μηδέν (0). Προφανώς ισχύει ότι $t_{ij} \geq 0$ και $t_i \geq 0$ για $1 \leq i, j \leq m$, με τουλάχιστον ένα $t_i > 0$. Επειδή ο πίνακας P είναι στοχαστικός ισχύει η σχέση

$$T\mathbf{1} + t = \mathbf{1}$$

όπου το $\mathbf{1}$ συμβολίζει το διάνυσμα στήλη διάστασης $1 \times m$ με όλα τα στοιχεία του ίσα με ένα (1). Έστω επίσης

$$Pr(X_0 = 1, X_0 = 2, \dots, X_0 = m+1) = (a_1, a_2, \dots, a_m, a_{m+1}) = (\mathbf{a}, a_{m+1})$$

η αρχική κατανομή της Μαρκοβιανής αλυσίδας. Είναι προφανές ότι

$$t = (I - T)\mathbf{1},$$

ή ισοδύναμα

$$(I - T)^{-1}t = \mathbf{1},$$

όπου $I = I_m$ είναι ο ταυτοτικός (μοναδιαίος) πίνακας διάστασης $m \times m$, και

$$a_{m+1} = 1 - \mathbf{a}\mathbf{1}.$$

Οι πιθανότητες που περιέχει το διάνυσμα t αποκαλούνται **πιθανότητες εξόδου** (exit probabilities).

Ορίζουμε

$$Y = \inf \{n \geq 0: X_n = m+1\}.$$

Σύμφωνα με τον παραπάνω ορισμό η τ.μ. Y δηλώνει τον αριθμό των βημάτων της Μαρκοβιανής αλυσίδας $\{X_n\}_{n \geq 0}$ έως ότου αυτή εισέλθει στην απορροφητική της κατάσταση $\{m + 1\}$. Η κατανομή της τ.μ. Y , που είναι διακριτή, ονομάζεται κατανομή τύπου φάσεων με αναπαράσταση (\mathbf{a}, \mathbf{T}) και συμβολίζεται ως $PH_d^{(m)}(\mathbf{a}, \mathbf{T})$. Η διάσταση m του πίνακα \mathbf{T} καλείται **τάξη** της κατανομής, ενώ οι παροδικές καταστάσεις $\{1, 2, \dots, m\}$ καλούνται **φάσεις** (phases) της κατανομής.

Στο επόμενο θεώρημα δίνεται η συνάρτηση κατανομής και η συνάρτηση πιθανότητας μιας διακριτής κατανομής τύπου φάσεων.

Θεώρημα 1.1

Έστω ότι $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$. Τότε η συνάρτηση κατανομής $F_Y(\cdot)$ της τ.μ. Y και η συνάρτηση πιθανότητάς της $f_Y(\cdot)$ δίνονται από τους τύπους

$$F_Y(k) = Pr(Y \leq k) = 1 - \mathbf{aT}^k \mathbf{1}, \quad k = 0, 1, 2, \dots$$

και

$$f_Y(k) = Pr(Y = k) = \begin{cases} a_{m+1}, & k = 0 \\ \mathbf{aT}^{k-1} \mathbf{t}, & k = 1, 2, 3, \dots \end{cases}$$

Απόδειξη

Μπορεί εύκολα να διαπιστωθεί (με επαγωγή) ότι για $k \geq 1$ ισχύει ότι

$$\begin{aligned} \mathbf{P}^{(k)} &= \mathbf{P}^k = \begin{pmatrix} \mathbf{T}^k & \mathbf{t} + \mathbf{Tt} + \mathbf{T}^2 \mathbf{t} + \dots + \mathbf{T}^{k-1} \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{T}^k & (\mathbf{I} - \mathbf{T}^k)(\mathbf{I} - \mathbf{T})^{-1} \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{T}^k & (\mathbf{I} - \mathbf{T}^k) \mathbf{1} \\ \mathbf{0} & 1 \end{pmatrix}. \end{aligned}$$

Τώρα η πιθανότητα του ενδεχομένου $\{Y > k\}$, $k = 1, 2, \dots$, είναι ίση με την πιθανότητα του ενδεχομένου $\{X_k \in \{1, 2, \dots, m\}\}$, αφού μέχρι και το βήμα k η Μαρκοβιανή αλυσίδα δεν είναι δυνατό να έχει εισέλθει στην απορροφητική της κατάσταση $\{m + 1\}$. Έτσι, για $k = 1, 2, \dots$ έχουμε ότι

$$\begin{aligned} Pr(Y > k) &= Pr(X_k \in \{1, 2, \dots, m\}) = \sum_{j=1}^m Pr(X_k = j) \\ &= \sum_{j=1}^m \sum_{i=1}^m Pr(X_k = j | X_0 = i) Pr(X_0 = i) \end{aligned}$$

$$= \sum_{j=1}^m \sum_{i=1}^m a_i p_{ij}^{(k)} = \sum_{j=1}^m \sum_{i=1}^m a_i (\mathbf{P}^k)_{ij} = \mathbf{aT}^k \mathbf{1}.$$

Συνεπώς

$$F_Y(k) = Pr(Y \leq k) = 1 - Pr(Y > k) = 1 - \mathbf{aT}^k \mathbf{1}.$$

Επίσης η πιθανότητα του ενδεχομένου $\{Y \leq 0\} = \{Y = 0\}$ είναι ίση με την πιθανότητα του ενδεχομένου $\{X_0 = m + 1\}$. Συνεπώς

$$F_Y(0) = Pr(Y = 0) = a_{m+1} = 1 - \mathbf{a1} = 1 - \mathbf{aI1} = 1 - \mathbf{aT}^0 \mathbf{1}.$$

Συνεπώς αποδείξαμε ότι

$$F_Y(k) = Pr(Y \leq k) = 1 - \mathbf{aT}^k \mathbf{1}, \quad k = 0, 1, 2, \dots$$

Για τη συνάρτηση πιθανότητας της τ.μ. Y έχουμε ότι

$$Pr(Y = 0) = a_{m+1}$$

και για $k = 1, 2, \dots$, έχουμε ότι

$$\begin{aligned} f_Y(k) &= F_Y(k) - F_Y(k-1) = 1 - \mathbf{aT}^k \mathbf{1} - (1 - \mathbf{aT}^{k-1} \mathbf{1}) \\ &= \mathbf{a(T}^{k-1} - \mathbf{T}^k) \mathbf{1} = \mathbf{aT}^{k-1} (\mathbf{I} - \mathbf{T}) \mathbf{1} = \mathbf{aT}^{k-1} \mathbf{t}. \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Για την πιθανογεννήτρια της τ.μ. Y έχουμε το ακόλουθο θεώρημα.

Θεώρημα 1.2

Έστω ότι $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$. Για την πιθανογεννήτρια συνάρτηση $g_Y(s)$ της τ.μ. Y έχουμε ότι

$$g_Y(s) = a_{m+1} + s\mathbf{a}(\mathbf{I} - s\mathbf{T})^{-1} \mathbf{t} = a_{m+1} + \mathbf{a}(s^{-1}\mathbf{I} - \mathbf{T})^{-1} \mathbf{t}, \quad -1 \leq s \leq 1.$$

Απόδειξη

Κάνοντας χρήση του προηγούμενου θεωρήματος παίρνουμε ότι

$$\begin{aligned} g_Y(s) &= E(s^Y) = \sum_{k=0}^{\infty} s^k f_Y(k) = a_{m+1} + \sum_{k=1}^{\infty} s^k f_Y(k) \\ &= a_{m+1} + \sum_{k=1}^{\infty} s^k \mathbf{aT}^{k-1} \mathbf{t} = a_{m+1} + s\mathbf{a} \left(\sum_{k=0}^{\infty} (s\mathbf{T})^k \right) \mathbf{t} \\ &= a_{m+1} + s\mathbf{a}(\mathbf{I} - s\mathbf{T})^{-1} \mathbf{t} = a_{m+1} + \mathbf{a}(s^{-1}\mathbf{I} - \mathbf{T})^{-1} \mathbf{t}. \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Η σύγκλιση της δυναμοσειράς πινάκων

$$\sum_{k=0}^{\infty} (s\mathbf{T})^k$$

στον πίνακα $(\mathbf{I} - s\mathbf{T})^{-1}$ για $|s| \leq 1$ προκύπτει από το γεγονός ότι

$$\rho(s\mathbf{T}) = \max_{1 \leq i \leq m} |\lambda_i| < 1$$

όπου λ_i είναι οι ιδιοτιμές του πίνακα $s\mathbf{T}$, και $\rho(s\mathbf{T})$ είναι η φασματική ακτίνα του πίνακα (για περισσότερες λεπτομέρειες δείτε Resnick (1992) και Bladt & Nielsen (2017)). Όπως έχουμε ήδη αναφέρει τα στοιχεία του θεμελιώδη πίνακα $\mathbf{U} = (u_{ij})_{m \times m} = (\mathbf{I} - \mathbf{T})^{-1}$ δίνουν το μέσο αριθμό E_{ij} ($1 \leq i, j \leq m$) επανόδων (επισκέψεων) στη μεταβατική κατάσταση j ξεκινώντας από τη μεταβατική κατάσταση i , δηλαδή $E_{ij} = u_{ij}$.

Στο επόμενο θεώρημα δίνεται η καθοδική παραγοντική ροπή r τάξης μιας διακριτής κατανομής τύπου φάσεων. Προτού δώσουμε την απόδειξη θα χρειαστούμε το ακόλουθο λήμμα.

Λήμμα 1.1

Έστω ότι $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$. Τότε

$$(α) \quad \frac{d(\mathbf{I} - s\mathbf{T})^{-1}}{ds} = \mathbf{T}(\mathbf{I} - s\mathbf{T})^{-2}.$$

$$(β) \quad \frac{d^r g_Y(s)}{ds^r} = r! \mathbf{a}\mathbf{T}^{r-1}(\mathbf{I} - s\mathbf{T})^{-r-1}\mathbf{t}, \quad r = 1, 2, \dots$$

Απόδειξη

Έχουμε ότι

$$\begin{aligned} \frac{d(\mathbf{I} - s\mathbf{T})^{-1}}{ds} &= \frac{d}{ds} \left(\sum_{k=0}^{\infty} (s\mathbf{T})^k \right) = \sum_{k=1}^{\infty} k s^{k-1} \mathbf{T}^k \\ &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} s^{k-1} \mathbf{T}^k = \mathbf{T} \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} (s\mathbf{T})^{k-1} = \mathbf{T} \sum_{n=1}^{\infty} (s\mathbf{T})^{n-1} \sum_{k=n}^{\infty} (s\mathbf{T})^{k-n} \\ &= \mathbf{T} \left(\sum_{n=1}^{\infty} (s\mathbf{T})^{n-1} \right) (\mathbf{I} - s\mathbf{T})^{-1} = \mathbf{T}(\mathbf{I} - s\mathbf{T})^{-2}. \end{aligned}$$

Έτσι το (α) αποδείχθηκε.

Θα αποδείξουμε το (β) με επαγωγή. Παρατηρούμε ότι

$$\begin{aligned}
\frac{dg_Y(s)}{ds} &= \frac{d}{ds} (sa(I - sT)^{-1}t) \\
&= a(I - sT)^{-1}t + sa \left(\frac{d}{ds} (I - sT)^{-1} \right) t \\
&= a(I - sT)^{-1}t + saT(I - sT)^{-2}t \\
&= a[(I - sT)^{-1} + sT(I - sT)^{-2}]t \\
&= a[(I - sT)(I - sT)^{-2} + sT(I - sT)^{-2}]t \\
&= a[(I - sT + sT)(I - sT)^{-2}]t \\
&= a(I - sT)^{-2}t
\end{aligned}$$

και επομένως το (β) ισχύει για $r = 1$. Έστω τώρα ότι ο τύπος ισχύει για $r (> 1)$. Τότε

$$\begin{aligned}
\frac{d^{r+1}g_Y(s)}{ds^{r+1}} &= r! aT^{r-1} \left(\frac{d}{ds} (I - sT)^{-r-1} \right) t \\
&= r! aT^{r-1} \{ (r+1)T(I - sT)^{-2}(I - sT)^{-r} \} t \\
&= (r+1)! aT^r (I - sT)^{-r-2} t.
\end{aligned}$$

που σημαίνει ότι ισχύει ο τύπος ισχύει για $r + 1$. Συνεπώς και το (β) ισχύει. ■

Οι πίνακες T και $(I - sT)^{-1}$ είναι αντιμεταθετικοί αφού

$$T(I - sT)^{-1} = T \left(\sum_{k=0}^{\infty} (sT)^k \right) = \sum_{k=0}^{\infty} s^k T^{k+1} = \left(\sum_{k=0}^{\infty} (sT)^k \right) T = (I - sT)^{-1} T.$$

Έτσι μπορούμε να γράψουμε ότι

$$\frac{d^r g_Y(s)}{ds^r} = r! a(I - sT)^{-r-1} T^{r-1} t, \quad r = 1, 2, \dots$$

Θεώρημα 1.3

Έστω ότι $Y \sim PH_d^{(m)}(a, T)$. Η παραγοντική ροπή r τάξης, $\mu_{(r)}$, της τ.μ. Y δίνεται από τον τύπο

$$\mu_{(r)} = E[Y(Y-1) \dots (Y - (r-1))] = r! aT^{r-1}(I - T)^{-r} \mathbf{1}, \quad r = 1, 2, \dots$$

Απόδειξη

Η παραγοντική ροπή r τάξης, $\mu_{(r)}$, της τ.μ. Y δίνεται από τον τύπο

$$\mu_{(r)} = \left. \frac{d^r g_Y(s)}{ds^r} \right|_{s=1}, \quad r = 1, 2, \dots$$

Για $r \geq 1$, από το Λήμμα 1.1 (β) προκύπτει ότι

$$\begin{aligned}
\left. \frac{d^r g_Y(s)}{ds^r} \right|_{s=1} &= r! \mathbf{a} \mathbf{T}^{r-1} (\mathbf{I} - \mathbf{T})^{-r-1} \mathbf{t} \\
&= r! \mathbf{a} \mathbf{T}^{r-1} (\mathbf{I} - \mathbf{T})^{-r} (\mathbf{I} - \mathbf{T})^{-1} \mathbf{t} \\
&= r! \mathbf{a} \mathbf{T}^{r-1} (\mathbf{I} - \mathbf{T})^{-r} \mathbf{1}.
\end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Εναλλακτικά μπορούμε να γράψουμε ότι

$$\mu_{(r)} = r! \mathbf{a} (\mathbf{I} - \mathbf{T})^{-r} \mathbf{T}^{r-1} \mathbf{1}, \quad r = 1, 2, \dots$$

Από το Θεώρημα 1.3 προκύπτει ότι

$$E(Y) = \mu_{(1)} = \mathbf{a} (\mathbf{I} - \mathbf{T})^{-1} \mathbf{1}.$$

Επίσης

$$\mu_{(2)} = E(Y(Y-1)) = 2\mathbf{a} \mathbf{T} (\mathbf{I} - \mathbf{T})^{-2} \mathbf{1}$$

οπότε

$$\begin{aligned}
\text{Var}(Y) &= \mu_{(2)} + \mu_{(1)} - \mu_{(1)}^2 \\
&= 2\mathbf{a} \mathbf{T} (\mathbf{I} - \mathbf{T})^{-2} \mathbf{1} + \mathbf{a} (\mathbf{I} - \mathbf{T})^{-1} \mathbf{1} - (\mathbf{a} (\mathbf{I} - \mathbf{T})^{-1} \mathbf{1})^2.
\end{aligned}$$

Στο επόμενο θεώρημα δίνεται η κατανομή του αθροίσματος δύο ανεξάρτητων διακριτών κατανομών τύπου φάσεων που καταλήγει να είναι και αυτή κατανομή τύπου φάσης.

Θεώρημα 1.4

Έστω Y_1 και Y_2 δύο ανεξάρτητες τυχαίες μεταβλητές με $Y_1 \sim PH_d^{(m_1)}(\mathbf{a}^{(1)}, \mathbf{T}^{(1)})$ και $Y_2 \sim PH_d^{(m_2)}(\mathbf{a}^{(2)}, \mathbf{T}^{(2)})$, τάξεων m_1 και m_2 , αντίστοιχα. Τότε η κατανομή της τ.μ.

$$Y = Y_1 + Y_2$$

είναι η $PH_d^{(m)}(\mathbf{a}, \mathbf{T})$ τάξης $m_1 + m_2$, όπου

$$\mathbf{a} = \left(\mathbf{a}^{(1)}, a_{m_1+1}^{(1)} \mathbf{a}^{(2)} \right),$$

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(1)} & \mathbf{t}^{(1)} \cdot \mathbf{a}^{(2)} \\ \mathbf{0} & \mathbf{T}^{(2)} \end{pmatrix},$$

και

$$(\mathbf{t}^{(1)} \cdot \mathbf{a}^{(2)})_{ij} = t_i^{(1)} a_j^{(2)}, \quad 1 \leq i \leq m_1, \quad 1 \leq j \leq m_2.$$

Απόδειξη

Έστω $\{X_n^1\}_{n \geq 0}$ και $\{X_n^2\}_{n \geq 0}$ οι Μαρκοβιανές αλυσίδες που σχετίζονται με τις τ.μ. Y_1 και Y_2 αντίστοιχα. Ορίζουμε την Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ με σύνολο καταστάσεων

$$S = \{1, 2, \dots, m_1, m_1 + 1, m_1 + 2, \dots, m_1 + m_2, m_1 + m_2 + 1\}$$

ως εξής: Αρχικά η Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ αναπτύσσεται όπως η Μαρκοβιανή αλυσίδα $\{X_n^1\}_{n \geq 0}$ και τη στιγμή της απορροφήσεως η Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ ξεκινά να αναπτύσσεται όπως η Μαρκοβιανή αλυσίδα $\{X_n^2\}_{n \geq 0}$. Ο χρόνος απορρόφησης Y για τη Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ είναι προφανώς ίσος με $Y_1 + Y_2$. Είναι εύκολο να διαπιστωθεί ότι η Μαρκοβιανή αλυσίδα $\{X_n: n \in \mathbb{N}\}$ έχει αρχικές πιθανότητες

$$\Pr(X_0 = 1, X_0 = 2, \dots, X_0 = m_1 + m_2 + 1) = \left(\mathbf{a}^{(1)}, a_{m_1+1}^{(1)} \mathbf{a}^{(2)}, a_{m_1+1}^{(1)} a_{m_2+1}^{(2)} \right)$$

και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} \mathbf{T}^{(1)} & \mathbf{t}^{(1)} \cdot \mathbf{a}^{(2)} & a_{m_2+1}^{(2)} \mathbf{t}^{(1)} \\ \mathbf{0} & \mathbf{T}^{(2)} & \mathbf{t}^{(2)} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}.$$

Το θεώρημα αποδείχθηκε. ■

Στο επόμενο θεώρημα δείχνεται ότι η κατανομή της μείξης δύο ανεξάρτητων διακριτών κατανομών τύπου φάσεων καταλήγει να είναι και αυτή κατανομή τύπου φάσης. Λέμε ότι η διακριτή τ.μ. Y με συνάρτηση κατανομής $F_Y(y)$ αποτελεί μείξη δύο διακριτών κατανομών Y_1 και Y_2 με συναρτήσεις κατανομής $F_{Y_1}(y)$ και $F_{Y_2}(y)$, αντίστοιχα, αν

$$F_Y(y) = pF_{Y_1}(y) + (1 - p)F_{Y_2}(y), \quad 0 < p < 1.$$

Προφανώς αντίστοιχη σχέση ισχύει και για τις συναρτήσεις πιθανότητας, δηλαδή

$$f_Y(y) = pf_{Y_1}(y) + (1 - p)f_{Y_2}(y), \quad y \in \mathbb{R}.$$

Θεώρημα 1.5

Έστω Y_1 και Y_2 δύο τυχαίες μεταβλητές με $Y_1 \sim PH_d^{(m_1)}(\mathbf{a}^{(1)}, \mathbf{T}^{(1)})$ και $Y_2 \sim PH_d^{(m_2)}(\mathbf{a}^{(2)}, \mathbf{T}^{(2)})$, τάξεων m_1 και m_2 , αντίστοιχα. Τότε η κατανομή της τ.μ. Y με συνάρτηση κατανομής

$$F_Y(y) = pF_{Y_1}(y) + (1 - p)F_{Y_2}(y), \quad 0 < p < 1$$

είναι η $PH_d^{(m)}(\mathbf{a}, \mathbf{T})$ τάξης $m_1 + m_2$, όπου

$$\mathbf{a} = (p\mathbf{a}^{(1)}, (1 - p)\mathbf{a}^{(2)})$$

και

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{(2)} \end{pmatrix}.$$

Απόδειξη

Έστω $\{X_n^1\}_{n \geq 0}$ και $\{X_n^2\}_{n \geq 0}$ οι Μαρκοβιανές αλυσίδες που σχετίζονται με τις τ.μ. Y_1 και Y_2 αντίστοιχα. Ορίζουμε την Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ με σύνολο καταστάσεων

$$S = \{1, 2, \dots, m_1, m_1 + 1, m_1 + 2, \dots, m_1 + m_2, m_1 + m_2 + 1\}$$

ως εξής: Με πιθανότητα p η Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ αναπτύσσεται όπως η Μαρκοβιανή αλυσίδα $\{X_n^1\}_{n \geq 0}$ και με πιθανότητα $1 - p$ αναπτύσσεται όπως η Μαρκοβιανή αλυσίδα $\{X_n^2\}_{n \geq 0}$. Ο χρόνος απορρόφησης Y για τη Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ είναι είτε ο χρόνος απορρόφησης Y_1 με πιθανότητα p είτε ο χρόνος απορρόφησης Y_2 με πιθανότητα $1 - p$, και επομένως

$$Y = \begin{cases} Y_1, & \text{με πιθανότητα } p, \\ Y_2, & \text{με πιθανότητα } 1 - p. \end{cases}$$

Είναι εύκολο να διαπιστωθεί ότι η Μαρκοβιανή αλυσίδα $\{X_n\}_{n \geq 0}$ έχει αρχικές πιθανότητες

$$\Pr(X_0 = 1, X_0 = 2, \dots, X_0 = m_1 + m_2 + 1) = (p\mathbf{a}^{(1)}, (1 - p)\mathbf{a}^{(2)}, p\mathbf{a}_{m_1+1}^{(1)} + (1 - p)\mathbf{a}_{m_2+1}^{(2)})$$

και πίνακα πιθανοτήτων μετάβασης

$$\mathbf{P} = \begin{pmatrix} \mathbf{T}^{(1)} & \mathbf{0} & \mathbf{t}^{(1)} \\ \mathbf{0} & \mathbf{T}^{(2)} & \mathbf{t}^{(2)} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}.$$

Το θεώρημα αποδείχθηκε. ■

Το Θεώρημα 1.4 και 1.5 μπορεί εύκολα να γενικευθεί στην περίπτωση που στην άθροιση (μείξη) εμπλέκονται περισσότερες από δύο κατανομές τύπου φάσεων με το αποτέλεσμα να είναι πάλι κατανομή τύπου φάσεων.

Προτού προχωρήσουμε στο τελευταίο θεωρητικό αποτέλεσμα για τις διακριτές κατανομές τύπου φάσεων, υπενθυμίζουμε τον ορισμό του αθροίσματος και του γινομένου Kronecker μεταξύ δύο πινάκων.

Έστω ο πίνακας $\mathbf{A} = (a_{ij})_{(n_1 \times m_1)}$ και ο πίνακας $\mathbf{B} = (b_{ij})_{(n_2 \times m_2)}$. Το **γινόμενο Kronecker** $\mathbf{A} \otimes \mathbf{B}$ μεταξύ των πινάκων \mathbf{A} και \mathbf{B} είναι ένας πίνακας διάστασης $n_1 n_2 \times m_1 m_2$ που ορίζεται ως εξής

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m_1}B \\ a_{21}B & a_{22}B & \cdots & a_{2m_1}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_11}B & a_{n_12}B & \cdots & a_{n_1m_1}B \end{pmatrix}.$$

Αν οι πίνακες A και B είναι τετραγωνικοί, δηλαδή $A = (a_{ij})_{(n_1 \times n_1)}$ και $B = (b_{ij})_{(n_2 \times n_2)}$, το **άθροισμα Kronecker** $A \oplus B$ μεταξύ των πινάκων A και B είναι ένας τετραγωνικός πίνακας διάστασης $(n_1 n_2) \times (n_1 n_2)$ που ορίζεται ως εξής

$$A \oplus B = A \otimes I_{n_2} + I_{n_1} \otimes B.$$

Το ακόλουθο θεωρητικό αποτέλεσμα που αφορά την κατανομή του ελάχιστου και του μέγιστου μεταξύ δύο διακριτών κατανομών τύπου φάσεων θα δοθεί χωρίς απόδειξη (για την απόδειξη δείτε Bladt & Nielsen (2017)).

Θεώρημα 1.6

Έστω Y_1 και Y_2 δύο τυχαίες μεταβλητές με $Y_1 \sim PH_d^{(m_1)}(\mathbf{a}^{(1)}, \mathbf{T}^{(1)})$ και $Y_2 \sim PH_d^{(m_2)}(\mathbf{a}^{(2)}, \mathbf{T}^{(2)})$, τάξεων m_1 και m_2 , αντίστοιχα. Τότε

(i) Για την τ.μ. $Y = \min(Y_1, Y_2)$ ισχύει ότι $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$ όπου $m = m_1 m_2$,

$$\mathbf{a} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}$$

και

$$\mathbf{T} = \mathbf{T}^{(1)} \otimes \mathbf{T}^{(2)}.$$

(ii) Για την τ.μ. $Y = \max(Y_1, Y_2)$ ισχύει ότι $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$ όπου $m = m_1 m_2 + m_1 + m_2$,

$$\mathbf{a} = \left(\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}, \mathbf{a}^{(1)} a_{m_2+1}^{(2)}, a_{m_1+1}^{(1)} \mathbf{a}^{(2)} \right)$$

και

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(1)} \otimes \mathbf{T}^{(2)} & \mathbf{T}^{(1)} \mathbf{t}^{(2)} & \mathbf{t}^{(1)} \mathbf{T}^{(2)} \\ \mathbf{0} & \mathbf{T}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}^{(2)} \end{pmatrix}$$

Απόδειξη

(i) Έστω $\{Y_n^1\}_{n \geq 0}$ και $\{Y_n^2\}_{n \geq 0}$ οι Μαρκοβιανές αλυσίδες που σχετίζονται με τις τ.μ. Y_1 και Y_2 αντίστοιχα. Για να βρούμε την κατανομή της τ.μ. $Y = \min(Y_1, Y_2)$ θα παρατηρήσουμε την εξέλιξη των δύο Μαρκοβιανών αλυσίδων, που συμπεριφέρονται ανεξάρτητα η μια από την άλλη, έως ότου κάποια από αυτές εισέλθει στην απορροφητική της κατάσταση.

Ορίζουμε την Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$ με χώρο καταστάσεων

$$S = \{(i, j): 1 \leq i \leq m_1, 1 \leq j \leq m_2\} \cup \{e\}$$

όπου η κατάσταση e είναι απορροφητική κατάσταση και οι υπόλοιπες καταστάσεις ($m_1 m_2$ το πλήθος) είναι παροδικές. Η Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$ βρίσκεται στην κατάσταση (i, j) ($1 \leq i \leq m_1, 1 \leq j \leq m_2$) όταν η Μαρκοβιανή αλυσίδα $\{Y_n^{(1)}\}_{n \geq 0}$ βρίσκεται στην κατάσταση i και η Μαρκοβιανή αλυσίδα $\{Y_n^{(2)}\}_{n \geq 0}$ βρίσκεται στην κατάσταση j . Επομένως η αρχική κατανομή \mathbf{a} της Μαρκοβιανής αλυσίδα $\{Y_n\}_{n \geq 0}$ δίνεται προφανώς από τη σχέση

$$\mathbf{a} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}$$

δηλαδή

$$Pr(Y_0 = (i, j)) = Pr(Y_0^{(1)} = i) \cdot Pr(Y_0^{(2)} = j) = a_i^{(1)} a_j^{(2)}, \quad 1 \leq i \leq m_1, 1 \leq j \leq m_2.$$

Συνεπώς

$$\begin{aligned} Pr(Y_0 = e) &= 1 - (\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}) \mathbf{1} = 1 - (1 - a_{m_1+1}^{(1)}) (1 - a_{m_2+1}^{(2)}) \\ &= a_{m_1+1}^{(1)} + a_{m_2+1}^{(2)} - a_{m_1+1}^{(1)} a_{m_2+1}^{(2)} \\ &= Pr[(Y_0^{(1)} = m_1 + 1) \cup (Y_0^{(2)} = m_2 + 1)]. \end{aligned}$$

Οι πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων της Μαρκοβιανής αλυσίδα $\{Y_n\}_{n \geq 0}$ είναι

$$\begin{aligned} Pr(Y_{n+1} = (k, l) | Y_n = (i, j)) &= Pr[(Y_{n+1}^{(1)} = k | Y_n^{(1)} = i) \cap (Y_{n+1}^{(2)} = l | Y_n^{(2)} = j)] \\ &= Pr[(Y_{n+1}^{(1)} = k | Y_n^{(1)} = i)] \cdot Pr[(Y_{n+1}^{(2)} = l | Y_n^{(2)} = j)] \\ &= T_{ik}^{(1)} T_{jl}^{(2)} \end{aligned}$$

δηλαδή ο πίνακας πιθανοτήτων μετάβασης \mathbf{T} της Μαρκοβιανής αλυσίδα $\{Y_n\}_{n \geq 0}$ δίνεται προφανώς από τον τύπο

$$\mathbf{T} = \mathbf{T}^{(1)} \otimes \mathbf{T}^{(2)}.$$

Για τις πιθανότητες εξόδου έχουμε ότι

$$\begin{aligned} Pr(Y_{n+1} = e | Y_n = (i, j)) &= Pr[(Y_{n+1}^{(1)} = m_1 + 1) \cup (Y_{n+1}^{(2)} = m_2 + 1) | Y_n^{(1)} = i, Y_n^{(2)} = j] \\ &= t_i^{(1)} + t_j^{(2)} - t_i^{(1)} t_j^{(2)} \\ &= 1 - (1 - t_i^{(1)}) (1 - t_j^{(2)}) \\ &= 1 - (\mathbf{T}^{(1)} \mathbf{1})_i (\mathbf{T}^{(2)} \mathbf{1})_j \\ &= 1 - (\mathbf{T} \mathbf{1})_{(i, j)} \end{aligned}$$

Συνεπώς δείξαμε ότι η κατανομή της τ.μ. $Y = \min(Y_1, Y_2)$ είναι η $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$.

(ii) Για να βρούμε την κατανομή της τ.μ. $Y = \max(Y_1, Y_2)$ θα παρατηρήσουμε την εξέλιξη των δύο Μαρκοβιανών αλυσίδων, που συμπεριφέρονται ανεξάρτητα η μια από την άλλη, έως ότου

και οι δύο αλυσίδες απορροφηθούν. Έτσι ορίζουμε την Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$ με χώρο καταστάσεων

$$S = \{(i, j), (i, m_2 + 1), (m_1 + 1, j) : 1 \leq i \leq m_1, 1 \leq j \leq m_2\} \cup \{e\}$$

που αποτελεί επέκταση του χώρου καταστάσεων που εξετάστηκε στο (i) προσθέτοντας ακόμη $m_1 + m_2$ καταστάσεις που αντιστοιχούν στις περιπτώσεις που μόνο μία από τις δύο αλυσίδες έχει φτάσει στην κατάσταση απορρόφησής. Η Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$ θα εισέλθει στην κατάσταση απορρόφησης e όταν αμφότερες οι αλυσίδες $\{Y_n^1\}_{n \geq 0}$ και $\{Y_n^2\}_{n \geq 0}$ εισέλθουν στην αντίστοιχη απορροφητική τους κατάσταση. Οι αρχικές πιθανότητες εκκίνησης σε οποιαδήποτε μεταβατική κατάσταση (i, j) ($1 \leq i \leq m_1, 1 \leq j \leq m_2$) δίνεται από το διάνυσμα $\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}$, για τις μεταβατικές καταστάσεις της μορφής $(i, m_2 + 1)$ ($1 \leq i \leq m_1$) από το διάνυσμα $\mathbf{a}^{(1)} \mathbf{a}_{m_2+1}^{(2)}$, και για τις μεταβατικές καταστάσεις της μορφής $(m_1 + 1, j)$ ($1 \leq j \leq m_2$) από το διάνυσμα $\mathbf{a}_{m_1+1}^{(1)} \mathbf{a}^{(2)}$. Τέλος η πιθανότητα να ξεκινήσει η Μαρκοβιανή αλυσίδα $\{Y_n\}_{n \geq 0}$ στην κατάσταση απορρόφησης e είναι $a_e = a_{m_1+1}^{(1)} a_{m_2+1}^{(2)}$. Έτσι

$$\mathbf{a} = \left(\mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)}, \mathbf{a}^{(1)} \mathbf{a}_{m_2+1}^{(2)}, \mathbf{a}_{m_1+1}^{(1)} \mathbf{a}^{(2)} \right).$$

Οι πιθανότητες μεταβάσεις μεταξύ των μεταβατικών καταστάσεων (i, j) ($1 \leq i \leq m_1, 1 \leq j \leq m_2$) δίνονται από τον πίνακα $\mathbf{T}^{(1)} \otimes \mathbf{T}^{(2)}$, μεταξύ των μεταβατικών καταστάσεων (i, j) και $(i, m_2 + 1)$ από τον πίνακα $\mathbf{T}^{(1)} \mathbf{t}^{(2)}$, και μεταξύ των μεταβατικών καταστάσεων (i, j) και $(m_1 + 1, j)$ από τον πίνακα $\mathbf{t}^{(1)} \mathbf{T}^{(2)}$. Από τη στιγμή που μια από τις Μαρκοβιανές αλυσίδες $\{Y_n^1\}_{n \geq 0}$ και $\{Y_n^2\}_{n \geq 0}$ εισέλθει στην απορροφητική της κατάσταση η $\{Y_n\}_{n \geq 0}$ συνεχίζει να εξελίσσεται σύμφωνα με την άλλη αλυσίδα. Δηλαδή

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(1)} \otimes \mathbf{T}^{(2)} & \mathbf{T}^{(1)} \mathbf{t}^{(2)} & \mathbf{t}^{(1)} \mathbf{T}^{(2)} \\ \mathbf{0} & \mathbf{T}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}^{(2)} \end{pmatrix}$$

Σημειώνουμε ότι το διάνυσμα των πιθανοτήτων εξόδου είναι το

$$\mathbf{t} = \begin{pmatrix} \mathbf{t}^{(1)} \otimes \mathbf{t}^{(2)} \\ \mathbf{t}^{(1)} \\ \mathbf{t}^{(2)} \end{pmatrix}.$$

Συνεπώς δείξαμε ότι η κατανομή της τ.μ. $Y = \max(Y_1, Y_2)$ είναι η $Y \sim PH_d^{(m)}(\mathbf{a}, \mathbf{T})$.

Το θεώρημα αποδείχθηκε. ■

Κλείνοντας την παρούσα παράγραφο αξίζει να σημειώσουμε ότι οποιαδήποτε διακριτή κατανομή με πεπερασμένο σύνολο τιμών είναι μια διακριτή τύπου φάσεων κατανομή. Πράγματι η τ.μ. X με σύνολο τιμών $R_X = \{0, 1, \dots, n\}$ και συνάρτηση πιθανότητας

$$Pr(X = k) = p_k, \quad k = 0, 1, \dots, n$$

είναι διακριτή κατανομή τύπου φάσεων $PH_d^{(n)}(\mathbf{a}, \mathbf{T})$, με

$$\mathbf{a} = (p_1, p_2, \dots, p_n),$$

και

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{n \times n}$$

(ο πίνακας \mathbf{T} έχει όλα τα στοιχεία του είναι μηδενικά εκτός των στοιχείων κάτω από τη διαγώνιο που είναι ίσα με 1 (δηλαδή $T_{i,i-1} = 1$ για $2 \leq i \leq n$)).

1.4.2 Πρότυπα διακριτών κατανομών τύπου φάσεων

Στην παρούσα παράγραφο δίνουμε ορισμένα παραδείγματα κατανομών τύπου φάσεων που οδηγούν σε γνωστές κατανομές όπως η γεωμετρική και η αρνητική διωνυμική.

- **Γεωμετρική κατανομή**

Μια τ.μ. Y έχει τη γεωμετρική κατανομή με παράμετρο p (συμβ. $Y \sim G(p)$) αν η συνάρτηση πιθανότητας της $f_Y(\cdot)$ δίνεται από τον τύπο

$$f_Y(k) = Pr(Y = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$

Επίσης η συνάρτηση κατανομής της $F_Y(\cdot)$ δίνεται από τον τύπο

$$F_Y(k) = Pr(Y \leq k) = 1 - (1 - p)^k, \quad k = 1, 2, \dots$$

Η τ.μ. Y περιγράφει τον αριθμό των ανεξάρτητων δοκιμών Bernoulli που απαιτούνται μέχρι την πρώτη εμφάνιση μιας επιτυχίας, όπου σε κάθε δοκιμή η πιθανότητα επιτυχίας είναι p και κατά συνέπεια η πιθανότητα αποτυχίας σε μια οποιαδήποτε δοκιμή ισούται με $1 - p$. Η τ.μ. Y μπορεί να θεωρηθεί ως μια διακριτή κατανομή τύπου φάσεων τάξης 1, με

$$\mathbf{a} = a_1 = 1, \quad \mathbf{P} = \begin{pmatrix} 1 - p & p \\ 0 & 1 \end{pmatrix}$$

δηλαδή $Y \sim PH_d^{(1)}(1, 1 - p)$. Η Μαρκοβιανή αλυσίδα ξεκινά με πιθανότητα 1 στη μεταβατική κατάσταση (αποτυχία) και σε κάθε βήμα υπάρχει πιθανότητα $1 - p$ να παραμένει στη

μεταβατική κατάσταση, και πιθανότητα p να εισέλθει στην κατάσταση απορρόφησης (επιτυχία).

- **Μείξη γεωμετρικών κατανομών**

Μία πεπερασμένη μείξη n ανεξάρτητων γεωμετρικών κατανομών $G(p_i)$ με αντίστοιχα βάρη π_i έχει συνάρτηση πιθανότητας

$$f_Y(k) = Pr(Y = k) = \sum_{i=1}^n \pi_i (1 - p_i)^{k-1} p_i, \quad k = 1, 2, \dots,$$

όπου $\pi_i \geq 0$ και $\sum_{i=1}^n \pi_i = 1$. Σύμφωνα με το Θεώρημα 1.5 η μείξη γεωμετρικών κατανομών είναι μια διακριτή κατανομή τύπου φάσεων τάξης n , με $\mathbf{a} = (\pi_1, \pi_2, \dots, \pi_n)$ και

$$\mathbf{P} = \begin{pmatrix} 1 - p_1 & 0 & \dots & 0 & p_1 \\ 0 & 1 - p_2 & \dots & 0 & p_2 \\ & & \ddots & & \\ 0 & 0 & \dots & 1 - p_n & p_n \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}_{(n+1) \times (n+1)} .$$

- **Αρνητική διωνυμική κατανομή**

Μια τ.μ. Y έχει την αρνητική διωνυμική κατανομή με παραμέτρους r και p (συμβ. $Y \sim NB(r, p)$) αν η συνάρτηση πιθανότητάς της $f_Y(\cdot)$ δίνεται από τον τύπο

$$f_Y(k) = Pr(Y = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2, \dots .$$

Η τ.μ. Y περιγράφει τον αριθμό των ανεξάρτητων δοκιμών Bernoulli που απαιτούνται μέχρι την εμφάνιση της r -οστής επιτυχίας, όπου σε κάθε δοκιμή η πιθανότητα επιτυχίας είναι p και η πιθανότητα αποτυχίας $1-p$. Η τ.μ. Y μπορεί να θεωρηθεί ως μια διακριτή κατανομή τύπου φάσεων τάξης r , με $\mathbf{a} = (1, 0, \dots, 0)$ και

$$\mathbf{P} = \begin{pmatrix} 1 - p & p & 0 & \dots & 0 & 0 \\ 0 & 1 - p & p & \dots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 1 - p & p \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}_{(r+1) \times (r+1)} .$$

- **Μετατοπισμένη αρνητική διωνυμική κατανομή**

Μια τ.μ. Y έχει τη μετατοπισμένη αρνητική διωνυμική κατανομή με παραμέτρους r και p (συμβ. $Y \sim NB_0(r, p)$) αν η συνάρτηση πιθανότητάς της $f_Y(\cdot)$ δίνεται από τον τύπο

$$f_Y(k) = Pr(Y = k) = \binom{k+r-1}{k} p^r (1-p)^k, \quad k = 0, 1, 2, \dots$$

Η τ.μ. Y μπορεί να γραφεί στη μορφή $Y = X - r$, όπου $Y \sim NB(r, p)$. Η τ.μ. Y μπορεί να θεωρηθεί ως μια διακριτή κατανομή τύπου φάσεων τάξης r , με $\mathbf{a} = (a_1, a_2, \dots, a_r)$ όπου

$$a_{i+1} = \binom{r}{i} p^i q^{r-i}, \quad i = 0, 1, 2, \dots, r-1$$

(οπότε $a_{r+1} = p^r$) και

$$\mathbf{P} = \begin{pmatrix} 1-p & p & 0 & \dots & 0 & 0 \\ 0 & 1-p & p & \dots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 1-p & p \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}_{(r+1) \times (r+1)}$$

1.4.3 Αριθμητικό παράδειγμα διακριτής κατανομής τύπου φάσεων

Έστω X_1, X_2, \dots μία ακολουθία ανεξάρτητων και ισόνομων δοκιμών Bernoulli με πιθανότητα επιτυχίας p , δηλαδή $Pr(X_i = 1) = p$ για $i = 1, 2, \dots$, και έστω T_k τυχαία μεταβλητή που δηλώνει τον αριθμό των δοκιμών που πρέπει να εκτελεστούν έως ότου εμφανιστούν για πρώτη φορά k (≥ 1) διαδοχικές επιτυχίες. Η κατανομή της τυχαίας μεταβλητής T_k ονομάζεται γεωμετρική κατανομή k -τάξης (συμβ. $T_k \sim G_k(p)$) και έχει μελετηθεί από διάφορους ερευνητές (Philippou & Muwafi (1982), Philippou et al. (1983), Aki et al. (1984), κ.λ.π.). Αν και αρχικώς δεν είχε σαφώς αναφερθεί από τους ερευνητές που μελέτησαν αυτή την κατανομή με χρήση αλυσίδων Markov, η κατανομή $G_k(p)$ είναι μια κατανομή τύπου φάσεων διάστασης k με αναπαράσταση (\mathbf{a}, \mathbf{T}) όπου

$$\mathbf{a} = (1, 0, 0, \dots, 0)$$

και

$$\mathbf{T} = \begin{pmatrix} q & p & 0 & 0 & \dots & 0 \\ q & 0 & p & 0 & \dots & 0 \\ q & 0 & 0 & p & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q & 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{k \times k}$$

(δείτε, π.χ., Balakrishnan & Koutras (2002) και Koutras & Eryilmaz (2017)).

Χρησιμοποιώντας τη σχέση

$$Pr(T_k = n) = \mathbf{aT}^{n-1}\mathbf{t}, \quad n = 1, 2, \dots$$

(δείτε Θεώρημα 1.1) όπου

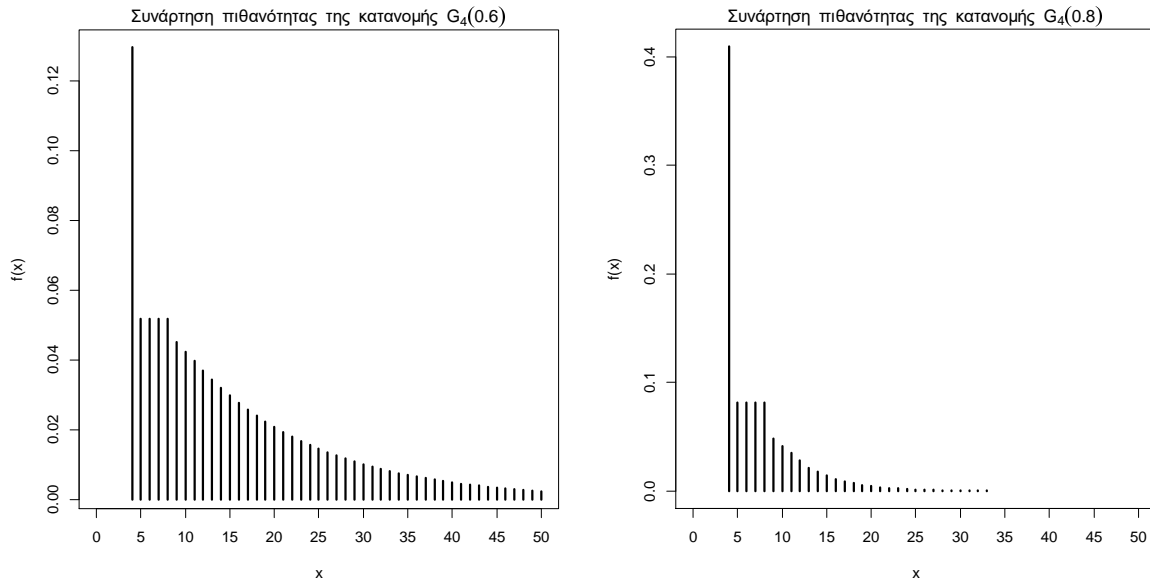
$$\mathbf{t} = (0, 0, \dots, 0, q)$$

μπορούμε εύκολα να υπολογίσουμε τη συνάρτηση πιθανότητας της κατανομής $G_k(p)$. Ενδεικτικά, στον παρακάτω πίνακα, δίνουμε τη συνάρτηση πιθανότητας των κατανομών $G_4(0.6)$ και $G_4(0.8)$.

Πίνακας 1.1. Συνάρτηση πιθανότητας των κατανομών $G_4(0.6)$ και $G_4(0.8)$

$p = 0.6$	n	$Pr(T_4 = n)$	$p = 0.8$	n	$Pr(T_4 = n)$
	0	0		0	0
	1	0		1	0
	2	0		2	0
	3	0		3	0
	4	0.1296		4	0.4096
	5	0.0518		5	0.0819
	6	0.0518		6	0.0819
	7	0.0518		7	0.0819
	8	0.0518		8	0.0819
	9	0.0451		9	0.0484
	10	0.0424		10	0.0417
	11	0.0397		11	0.0349
	12	0.0371		12	0.0282
	13	0.0344		13	0.0215
	14	0.0320		14	0.0176
	15	0.0298		15	0.0141

Ενδεικτικές είναι και οι ακόλουθες δύο γραφικές παραστάσεις.



Σχήμα 1.4. Συνάρτηση πιθανότητας των κατανομών $G_4(0.6)$ και $G_4(0.8)$

Χρησιμοποιώντας το Θεώρημα 1.3 βρίσκουμε τις ακόλουθες παραγοντικές ροπές των κατανομών $G_4(0.6)$ και $G_4(0.8)$.

Πίνακας 1.2. Παραγοντικές ροπές των κατανομών $G_4(0.6)$ και $G_4(0.8)$

$G_4(0.6)$	r	$\mu(r)$	$G_4(0.8)$	r	$\mu(r)$
	1	16.7901		1	7.2070
2	459.8659	2	63.8826		
3	18617.2129	3	768.7898		
4	1004715.5506	4	12287.1686		

1.4.4. Συνεχείς κατανομές τύπου φάσεων

Μια συνεχής κατανομή τύπου φάσεων είναι η κατανομή του χρόνου απορρόφησης μιας απορροφητικής Μαρκοβιανής αλυσίδας συνεχούς χρόνου με πεπερασμένο σύνολο καταστάσεων S .

Πιο συγκεκριμένα, έστω $\{X(t)\}_{t \geq 0}$ μία ομογενής Μαρκοβιανή αλυσίδα συνεχούς χρόνου με σύνολο καταστάσεων $S = \{1, 2, \dots, m, m + 1\}$, όπου οι καταστάσεις $\{1, 2, \dots, m\}$ είναι παροδικές και η κατάσταση $\{m + 1\}$ είναι απορροφητική. Ο απειροστός γεννήτορας πίνακας (infinitesimal generator matrix) πιθανοτήτων μετάβασης μιας τέτοιας Μαρκοβιανής αλυσίδας έχει τη μορφή

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}$$

όπου $\mathbf{T} = (t_{ij})_{m \times m}$ είναι ένας τετραγωνικός πίνακας διάστασης $m \times m$ (το στοιχείο t_{ij} , για $i \neq j$, είναι ο ρυθμός μετάβασης από τη μεταβατική κατάσταση i στη μεταβατική κατάσταση j , $\mathbf{t} = (t_1, t_2, \dots, t_m)'$ είναι ένα διάνυσμα στήλη διάστασης $m \times 1$ που περιέχει τους ρυθμούς εξόδου από τις μεταβατικές καταστάσεις στην απορροφητική κατάσταση, και το $\mathbf{0}$ συμβολίζει το διάνυσμα γραμμή διάστασης $1 \times m$ με όλα τα στοιχεία του ίσα με μηδέν (0). Προφανώς ισχύει ότι $t_{ij} \geq 0$ για $i \neq j$, $t_{ii} < 0$, και $t_i \geq 0$ ($1 \leq i, j \leq m$). Επειδή ο πίνακας \mathbf{Q} είναι γεννήτορας πίνακας που αντιστοιχεί σε πεπερασμένο σύνολο καταστάσεων ισχύει η σχέση

$$\mathbf{T}\mathbf{1} + \mathbf{t} = \mathbf{0}$$

όπου το $\mathbf{1}$ συμβολίζει το διάνυσμα στήλη διάστασης $m \times 1$ με όλα τα στοιχεία του ίσα με ένα (1). Έστω επίσης

$$Pr(X_0 = 1, X_0 = 2, \dots, X_0 = m + 1) = (a_1, a_2, \dots, a_m, a_{m+1}) = (\mathbf{a}, a_{m+1})$$

η αρχική κατανομή της Μαρκοβιανής αλυσίδας. Είναι προφανές ότι

$$\mathbf{t} = -\mathbf{T}\mathbf{1},$$

ή ισοδύναμα

$$-\mathbf{T}^{-1}\mathbf{t} = \mathbf{1},$$

όπου $\mathbf{I} = \mathbf{I}_m$ είναι ο ταυτοτικός (μοναδιαίος) πίνακας διάστασης $m \times m$, και

$$a_{m+1} = 1 - \mathbf{a}\mathbf{1}.$$

Τα στοιχεία που περιέχει το διάνυσμα \mathbf{t} αποκαλούνται **ρυθμοί εξόδου** (exit rates).

Ορίζουμε

$$Y = \inf(t \geq 0: X(t) = m + 1).$$

Σύμφωνα με τον παραπάνω ορισμό η τ.μ. Y δηλώνει το χρόνο που απαιτείται προκειμένου να εισέλθει η Μαρκοβιανή αλυσίδα $\{X(t)\}_{t \geq 0}$ στην απορροφητική της κατάσταση $\{m + 1\}$. Η κατανομή της τ.μ. Y , που είναι συνεχής, ονομάζεται κατανομή τύπου φάσεων με αναπαράσταση (\mathbf{a}, \mathbf{T}) και συμβολίζεται ως $PH_c^{(m)}(\mathbf{a}, \mathbf{T})$. Η διάσταση m του πίνακα \mathbf{T} καλείται **τάξη** της κατανομής, ενώ οι μεταβατικές καταστάσεις $\{1, 2, \dots, m\}$ καλούνται **φάσεις** της κατανομής

Στο επόμενο θεώρημα δίνεται η συνάρτηση κατανομής και η συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς κατανομής τύπου φάσεων.

Θεώρημα 1.7

Έστω ότι $Y \sim PH_C^{(m)}(\mathbf{a}, \mathbf{T})$. Τότε η συνάρτηση κατανομής $F_Y(\cdot)$ της τ.μ. Y και η συνάρτηση πυκνότητας πιθανότητας της $f_Y(\cdot)$ δίνονται από τους τύπους

$$F_Y(y) = Pr(Y \leq y) = 1 - \mathbf{a}e^{T y} \mathbf{1}, \quad y \geq 0,$$

και

$$f_Y(y) = \begin{cases} a_{m+1}, & y = 0 \\ \mathbf{a}e^{T y} \mathbf{t}, & y > 0. \end{cases}$$

Απόδειξη

Αρχικά παρατηρούμε ότι

$$\mathbf{Q}^0 = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}^0 = \mathbf{I}_{m+1} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}.$$

Με την επαγωγική μέθοδο μπορεί να διαπιστωθεί ότι

$$\mathbf{Q}^n = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}^n = \begin{pmatrix} \mathbf{T}^n & \mathbf{T}^{n-1} \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{T}^n & -\mathbf{T}^n \mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix}, \quad n = 1, 2, \dots$$

Έτσι

$$\begin{aligned} e^{\mathbf{Q}y} &= \sum_{n=0}^{\infty} \frac{(\mathbf{Q}y)^n}{n!} = \sum_{n=0}^{\infty} \frac{\mathbf{Q}^n y^n}{n!} = \sum_{n=0}^{\infty} \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}^n \frac{y^n}{n!} \\ &= \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} + \begin{pmatrix} \mathbf{T} & -\mathbf{T} \mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix} \frac{y^1}{1!} + \begin{pmatrix} \mathbf{T}^2 & -\mathbf{T}^2 \mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix} \frac{y^2}{2!} + \begin{pmatrix} \mathbf{T}^3 & -\mathbf{T}^3 \mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix} \frac{y^3}{3!} + \dots \\ &= \begin{pmatrix} \mathbf{I}_m + \sum_{n=1}^{\infty} \frac{\mathbf{T}^n y^n}{n!} & -\sum_{n=1}^{\infty} \frac{\mathbf{T}^n y^n}{n!} \mathbf{1} \\ \mathbf{0} & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{T y} & \mathbf{1} - e^{T y} \mathbf{1} \\ \mathbf{0} & 1 \end{pmatrix}. \end{aligned}$$

Τώρα η πιθανότητα του ενδεχομένου $\{Y > y\}$, $y \geq 0$ είναι ίση με την πιθανότητα του ενδεχομένου $\{X_y \in \{1, 2, \dots, m\}\}$, αφού μέχρι και το χρόνο y η Μαρκοβιανή αλυσίδα δεν είναι δυνατό να έχει εισέλθει στην απορροφητική κατάσταση $\{m + 1\}$. Έτσι, για $y \geq 0$ έχουμε ότι

$$\begin{aligned}
Pr(Y > y) &= Pr(X_y \in \{1, 2, \dots, m\}) = \sum_{j=1}^m Pr(X_y = j) \\
&= \sum_{j=1}^m \sum_{i=1}^m Pr(X_y = j | X_0 = i) Pr(X_0 = i) \\
&= \sum_{j=1}^m \sum_{i=1}^m a_i p_{ij}(y) = \mathbf{a} e^{T y} \mathbf{1}
\end{aligned}$$

(αφού $\mathbf{P}(t) = e^{Qt}$). Συνεπώς

$$F_Y(y) = Pr(Y \leq y) = 1 - \mathbf{a} e^{T y} \mathbf{1}, \quad y \geq 0.$$

Για τη συνάρτηση πυκνότητας της τ.μ. Y έχουμε ότι για $y > 0$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d(1 - \mathbf{a} e^{T y} \mathbf{1})}{dy} = -\mathbf{a} e^{T y} \mathbf{T} \mathbf{1} = \mathbf{a} e^{T y} \mathbf{t}.$$

Επίσης

$$f_Y(0) = Pr(Y = 0) = Pr(X(0) = m + 1) = a_{m+1}.$$

Το θεώρημα αποδείχθηκε. ■

Σημειώνουμε ότι η τ.μ. Y έχει μεικτή κατανομή (mixed distribution) με διακριτό μέρος

$$f_1(0) = Pr(Y = 0) = a_{m+1},$$

και συνεχές μέρος

$$f_2(y) = \mathbf{a} e^{T y} \mathbf{t}, \quad y > 0,$$

και επιπροσθέτως είναι μη-ελαττωματική (non-defective) με την έννοια ότι $F_Y(\infty) = 1$.

Επίσης σημειώνουμε ότι ο πίνακας \mathbf{T} είναι αντιστρέψιμος (αφού οι καταστάσεις $\{1, 2, \dots, m\}$ είναι παροδικές), και τα στοιχεία του πίνακα $\mathbf{U} = (u_{ij})_{m \times m} = (-\mathbf{T})^{-1}$ δηλώνουν το μέσο χρόνο E_{ij} που βρίσκεται η Μαρκοβιανή αλυσίδα στη μεταβατική κατάσταση j ξεκινώντας από τη μεταβατική κατάσταση i προτού αυτή απορροφηθεί, δηλαδή

$$E_{ij} = E \left(\int_0^Y 1_{\{X(t) = j | X_0 = i\}} dt \right) = (-\mathbf{T})_{ij}^{-1} = u_{ij}$$

(για περισσότερες λεπτομέρειες δείτε, π.χ., Breuer & Baum (2005) και Bladt & Nielsen (2017)).

Στο επόμενο θεώρημα δίνεται ο μετασχηματισμός Laplace μιας συνεχούς κατανομής τύπου φάσεων.

Θεώρημα 1.8

Έστω ότι $Y \sim PH_c^{(m)}(\mathbf{a}, \mathbf{T})$. Για το μετασχηματισμό Laplace $L_Y(s)$ της τ.μ. Y έχουμε ότι

$$L_Y(s) = \mathbf{a}_{m+1} + \mathbf{a}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}, \quad s \geq 0.$$

Απόδειξη

Ο τύπος για το μετασχηματισμό Laplace $L_Y(s)$ της τ.μ. Y ισχύει για $s = 0$ αφού

$$L_Y(0) = \mathbf{a}_{m+1} + \mathbf{a}(-\mathbf{T})^{-1}\mathbf{t} = \mathbf{a}_{m+1} + \mathbf{a}\mathbf{1} = \mathbf{1}$$

δεδομένου ότι $(-\mathbf{T})^{-1}\mathbf{t} = \mathbf{1}$.

Για $s > 0$ έχουμε ότι

$$\begin{aligned} L_Y(s) &= E(e^{-sY}) = \int_0^\infty e^{-sy} dF_Y(y) = \mathbf{a}_{m+1} + \int_0^\infty e^{-sy} f_Y(y) dy \\ &= \mathbf{a}_{m+1} + \mathbf{a} \left(\int_0^\infty e^{-sy} e^{Ty} dy \right) \mathbf{t} \\ &= \mathbf{a}_{m+1} + \mathbf{a} \left[-\frac{1}{s} \left([e^{-sy} e^{Ty}]_0^\infty - \int_0^\infty e^{-sy} e^{Ty} T dy \right) \right] \mathbf{t}. \end{aligned}$$

Αφού ο πίνακας \mathbf{T} είναι αντιστρέψιμος τότε

$$\lim_{s \rightarrow \infty} e^{Ts} = \mathbf{0}$$

(δείτε, π.χ., Breuer & Baum (2005)). Συνεπώς

$$\int_0^\infty e^{-sy} e^{Ty} dy = \frac{1}{s} \mathbf{I} + \frac{1}{s} \int_0^\infty e^{-sy} e^{Ty} dy \mathbf{T}$$

οπότε

$$\int_0^\infty e^{-sy} e^{Ty} dy (s\mathbf{I} - \mathbf{T}) = \mathbf{I}$$

δηλαδή

$$\int_0^\infty e^{-sy} e^{Ty} dy = (s\mathbf{I} - \mathbf{T})^{-1}.$$

Λαμβάνοντας υπόψη όλα τα παραπάνω, παίρνουμε τελικά ότι

$$L_Y(s) = \mathbf{a}_{m+1} + \mathbf{a}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}$$

Το θεώρημα αποδείχθηκε. ■

Στο παραπάνω θεώρημα χρησιμοποιήθηκε το γεγονός ότι ο πίνακας $(sI - T)$ έχει αντίστροφο (δείτε, π.χ., Bladt & Nielsen (2017)). Στο επόμενο θεώρημα δίνεται η ροπή r τάξης μιας συνεχούς κατανομής τύπου φάσεων.

Θεώρημα 1.9

Έστω ότι $Y \sim PH_c^{(m)}(\mathbf{a}, \mathbf{T})$. Η ροπή r τάξης, μ_r , της τ.μ. Y δίνεται από τον τύπο

$$\mu_r = E(Y^r) = (-1)^r r! \mathbf{aT}^{-r} \mathbf{1}, \quad r = 1, 2, \dots$$

Απόδειξη

Η ροπή r τάξης, μ_r , της τ.μ. Y δίνεται από τον τύπο

$$\mu_r = E(Y^r) = (-1)^r \left. \frac{d^r L_Y(s)}{ds^r} \right|_{s=0}, \quad r = 1, 2, \dots$$

Χρησιμοποιώντας το Θεώρημα 1.7 και τη σχέση

$$\frac{d^r}{ds^r} (sI - T)^{-1} = (-1)^r r! (sI - T)^{-(r+1)}$$

παίρνουμε

$$\frac{d^r L_Y(s)}{ds^r} = \frac{d^r}{ds^r} (\mathbf{a}_{m+1} + \mathbf{a}(sI - T)^{-1} \mathbf{t}) = \mathbf{a} (-1)^r r! (sI - T)^{-(r+1)} \mathbf{t}.$$

Συνεπώς

$$\begin{aligned} \mu_r = E(Y^r) &= (-1)^r \left. \frac{d^r L_Y(s)}{ds^r} \right|_{s=0} = \mathbf{a} r! (-T)^{-(r+1)} \mathbf{t} \\ &= r! \mathbf{a} (-T)^{-r} (-T)^{-1} \mathbf{t} = (-1)^r r! \mathbf{aT}^{-r} \mathbf{1} \end{aligned}$$

από την οποία προκύπτει το ζητούμενο. ■

Από το Θεώρημα 1.9 προκύπτει ότι

$$E(Y) = \mu_1 = -\mathbf{aT}^{-1} \mathbf{1}$$

και

$$\text{Var}(Y) = \mu_2 - \mu_1^2 = 2\mathbf{aT}^{-2} \mathbf{1} - (\mathbf{aT}^{-1} \mathbf{1})^2.$$

Σημειώνουμε ότι οι εκφωνήσεις των Θεωρημάτων 1.4 και 1.5 που αναφέρονται σε άθροισμα και μείξη δύο διακριτών κατανομών τύπου φάσεων, αλλά και του Θεωρήματος 1.6

που αναφέρεται σε γινόμενο Kronecker διακριτών κατανομών τύπου φάσεων, ισχύουν και στη συνεχή περίπτωση.

Τέλος αξίζει να αναφέρουμε ότι η κλάση των συνεχών κατανομών τύπου φάσεων είναι **πυκνή** (dense) με την έννοια της ασθενούς σύγκλισης (δείτε Breuer & Baum (2005)).

1.4.5 Πρότυπα συνεχών κατανομών τύπου φάσεων

Στην παρούσα παράγραφο δίνουμε ορισμένα παραδείγματα συνεχών κατανομών τύπου φάσεων που οδηγούν σε γνωστές κατανομές όπως η εκθετική και η Erlang κατανομή.

- **Εκθετική κατανομή**

Υπενθυμίζουμε ότι μια τ.μ. Y έχει την εκθετική κατανομή με παράμετρο λ (συμβ. $Y \sim \text{Exp}(\lambda)$) αν η συνάρτηση πυκνότητας πιθανότητάς της $f_Y(\cdot)$ δίνεται από τον τύπο

$$f_Y(y) = \lambda e^{-\lambda y}, \quad y \geq 0.$$

Η συνάρτηση κατανομής της ικανοποιεί τη σχέση $F_Y(y) = 1 - e^{-\lambda y}$, $y \geq 0$.

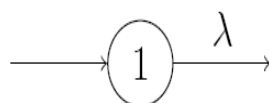
Η εκθετική κατανομή είναι πολύ δημοφιλής στη στοχαστική μοντελοποίηση αφού έχει μόνο μία παράμετρο λ και έχει την ιδιότητα της έλλειψης μνήμης. Η εκθετική κατανομή είναι το συνεχές ανάλογο της γεωμετρικής κατανομής που είναι η μόνη διακριτή διανομή που έχει την ιδιότητα έλλειψης μνήμης. Η εκθετική έχει μέση τιμή $E(Y) = 1/\lambda$ και διακύμανση $\text{Var}(Y) = 1/\lambda^2$.

Η εκθετική κατανομή έχει αναπαράσταση ως μιας συνεχούς κατανομής τύπου φάσεων. Πράγματι έχουμε ότι η $Y \sim PH_c^{(1)}(\mathbf{a}, \mathbf{T})$ τάξης 1, με $\mathbf{a} = 1$, $\mathbf{T} = -\lambda$ και $\mathbf{t} = \lambda$. Επομένως

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda \\ 0 & 0 \end{pmatrix}.$$

Η Μαρκοβιανή αλυσίδα ξεκινά με πιθανότητα 1 στη μοναδική της μεταβατική κατάσταση (φάση), παραμένει σε αυτή χρόνο που κατανέμεται εκθετικά με παράμετρο λ , και εισέρχεται στην κατάσταση απορρόφησης με ρυθμό λ .

Το διάγραμμα φάσεων της εκθετικής κατανομής είναι το ακόλουθο:



Σχήμα 1.5. Διάγραμμα φάσης εκθετικής κατανομής.

- **Υπερεκθετική κατανομή**

Μια τ.μ. Y έχει την υπερεκθετική κατανομή εάν είναι μια πεπερασμένη μείξη n (≥ 2) ανεξάρτητων εκθετικών κατανομών $Exp(\lambda_i)$, $1 \leq i \leq n$. Η τ.μ. Y έχει συνάρτηση πυκνότητας πιθανότητας

$$f_Y(y) = \sum_{i=1}^n \pi_i \lambda_i e^{-\lambda_i y}, \quad y \geq 0,$$

όπου $\pi_i \geq 0$ και $\sum_{i=1}^n \pi_i = 1$, συνάρτηση κατανομής

$$F_Y(y) = 1 - \sum_{i=1}^n \pi_i e^{-\lambda_i y}, \quad y \geq 0,$$

ροπογεννήτρια συνάρτηση

$$M_Y(s) = E(e^{sY}) = \sum_{i=1}^n \pi_i \frac{\lambda_i}{\lambda_i - s}, \quad s < \lambda = \min(\lambda_1, \lambda_2, \dots, \lambda_n),$$

και ροπή r - τάξης

$$\mu_r = E(Y^r) = r! \sum_{i=1}^n \frac{\pi_i}{\lambda_i^r}, \quad r \geq 1.$$

Η μέση τιμή της $E(Y)$ και η διακύμανσή της $Var(Y)$ είναι ίσες με

$$\mu = E(Y) = \sum_{i=1}^n \frac{\pi_i}{\lambda_i}, \quad \sigma^2 = Var(Y) = \sum_{i=1}^n \frac{2\pi_i}{\lambda_i^2} - \left(\sum_{i=1}^n \frac{\pi_i}{\lambda_i} \right)^2.$$

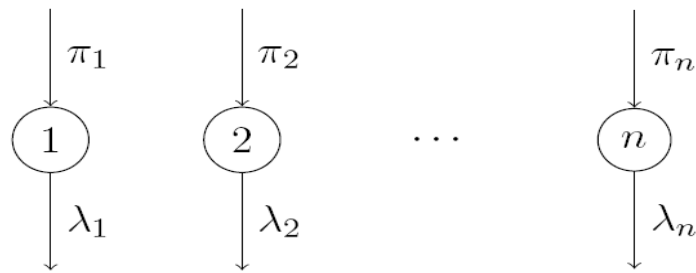
Μπορεί εύκολα να δειχθεί ότι $\left(\frac{\sigma}{\mu}\right)^2 = \frac{Var(Y)}{E^2(Y)} > 1$, και έτσι δικαιολογείται η ονομασία υπερεκθετική κατανομή (για την εκθετική κατανομή έχουμε ότι ο συντελεστής μεταβλητότητας ς είναι 1).

Η υπερεκθετική κατανομή είναι μία συνεχής κατανομή τύπου φάσεων τάξης n , με αναπαράσταση $\mathbf{a} = (\pi_1, \pi_2, \dots, \pi_n)$ και

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & & \\ & \ddots & \\ & & -\lambda_n \end{pmatrix}_{n \times n}, \quad \mathbf{t} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{1 \times n}$$

(τα υπόλοιπα στοιχεία του πίνακα \mathbf{T} είναι μηδέν).

Το διάγραμμα φάσεων της υπερεκθετικής κατανομής είναι το ακόλουθο



Σχήμα 1.6. Διάγραμμα φάσης υπερεκθετικής κατανομής.

- **Κατανομή Erlang**

Μια τ.μ. Y έχει την κατανομή Erlang με παραμέτρους n και λ (συμβ. $Y \sim Erlang(n, \lambda)$) αν είναι η κατανομή του αθροίσματος $n \geq 2$ ανεξάρτητων τ.μ. που κάθε μια έχει την κατανομή $Exp(\lambda)$. Η τ.μ. Y έχει συνάρτηση πυκνότητας πιθανότητας

$$f_Y(y) = \frac{\lambda^n}{(n-1)!} y^{n-1} e^{-\lambda y}, \quad y \geq 0,$$

συνάρτηση κατανομής

$$F_Y(y) = 1 - e^{-\lambda y} \sum_{i=1}^{n-1} \frac{(\lambda y)^i}{i!}, \quad y \geq 0$$

ροπογεννήτρια συνάρτηση

$$M_Y(s) = E(e^{sY}) = \left(\frac{\lambda}{\lambda - s} \right)^n, \quad s < \lambda,$$

και ροπή r - τάξης

$$\mu_r = E(Y^r) = \frac{(n+r-1)!}{(n-1)! \lambda^r}, \quad r \geq 1.$$

Η μέση τιμή $E(Y)$ και η διακύμανση $Var(Y)$ της Y είναι ίσες με

$$\mu = E(Y) = \frac{n}{\lambda}, \quad \sigma^2 = Var(Y) = \frac{n}{\lambda^2}.$$

Επίσης η κατανομή $Erlang(1, \lambda)$ είναι η κατανομή $Exp(\lambda)$.

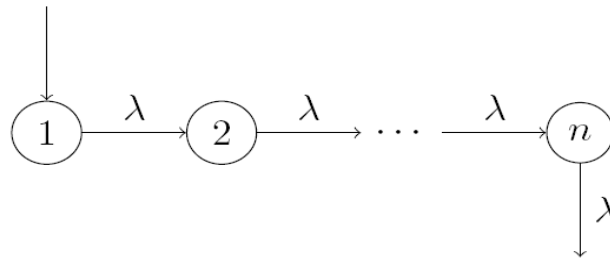
Η κατανομή Erlang είναι μία συνεχής κατανομή τύπου φάσεων τάξης n , με αναπαράσταση $\mathbf{a} = (1, 0, \dots, 0)$ και

$$T = \begin{pmatrix} -\lambda & \lambda & \cdots & & \\ 0 & -\lambda & \cdots & & \\ & & \ddots & & \\ & & & -\lambda & \lambda \\ & & & & -\lambda \end{pmatrix}_{n \times n}, \quad t = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda \end{pmatrix}_{1 \times n}$$

(τα υπόλοιπα στοιχεία του πίνακα T είναι μηδέν).

Η κατανομή Erlang θεωρείται ως ο χρόνος μέχρι την απορρόφηση μιας Μαρκοβιανής αλυσίδας συνεχούς χρόνου με n μεταβατικές καταστάσεις, η οποία ξεκινά από την κατάσταση 1 και μπορεί να μετακινηθεί μόνο στην επόμενη κατάστασή της με σταθερό ρυθμό μετάβασης λ , έως ότου απορροφηθεί.

Το διάγραμμα φάσεων της κατανομής Erlang είναι το ακόλουθο



Σχήμα 1.7. Διάγραμμα φάσης της κατανομής Erlang.

- **Υποεκθετική κατανομή**

Η τυχαία μεταβλητή Y έχει την υποεκθετική κατανομή (ή γενικευμένη κατανομή Erlang) με παραμέτρους $\lambda_1, \lambda_2, \dots, \lambda_n$ αν είναι η κατανομή του αθροίσματος $n \geq 2$ ανεξάρτητων τ.μ. που κάθε μια έχει την κατανομή $Exp(\lambda_i)$. Η μέση τιμή $E(Y)$ και η διακύμανση $Var(Y)$ της τυχαίας μεταβλητής Y είναι ίσες με

$$\mu = E(Y) = \sum_{i=1}^n \frac{1}{\lambda_i}, \quad \sigma^2 = Var(Y) = \sum_{i=1}^n \frac{1}{\lambda_i^2}$$

Η τ.μ. Y έχει ροπογεννήτρια συνάρτηση

$$M_Y(s) = E(e^{sY}) = \prod_{i=1}^n \left(\frac{\lambda_i}{\lambda_i - s} \right)^n, \quad s < \lambda = \min(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Μπορεί εύκολα να δειχθεί ότι $\left(\frac{\sigma}{\mu}\right)^2 = \frac{Var(Y)}{E^2(Y)} < 1$, και έτσι δικαιολογείται η ονομασία υποεκθετική κατανομή (για την εκθετική κατανομή έχουμε ότι ο συντελεστής μεταβλητότητας είναι 1).

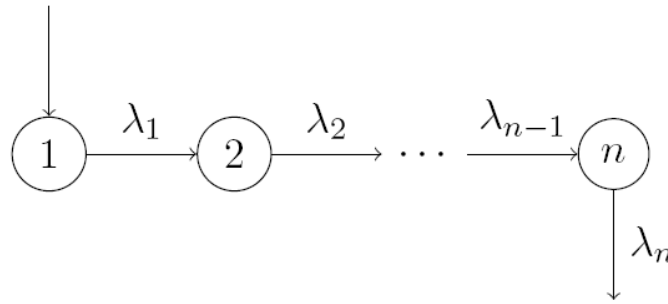
Η υποεκθετική κατανομή είναι μία συνεχής κατανομή τύπου φάσεων τάξης n , με αναπαράσταση $\mathbf{a} = (1, 0, \dots, 0)$ και

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 & \ddots & & & \\ & -\lambda_2 & & \ddots & & \\ & & & & \ddots & \\ & & & & & -\lambda_{n-1} & \lambda_{n-1} \\ & & & & & & -\lambda_n \end{pmatrix}_{n \times n}, \quad \mathbf{t} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{pmatrix}_{1 \times n}.$$

(τα υπόλοιπα στοιχεία του πίνακα \mathbf{T} είναι μηδέν).

Η υποεκθετική κατανομή είναι ο χρόνος μέχρι την απορρόφηση μιας απορροφητικής Μαρκοβιανής αλυσίδας συνεχούς χρόνου με n μεταβατικές καταστάσεις, η οποία ξεκινά από την κατάσταση 1 και μπορεί να μετακινηθεί μόνο στην επόμενη κατάστασή της με μη σταθερό ρυθμό μετάβασης, έως ότου απορροφηθεί.

Το διάγραμμα φάσεων της υποεκθετικής κατανομής είναι το ακόλουθο



Σχήμα 1.8. Διάγραμμα φάσης υποεκθετικής κατανομής.

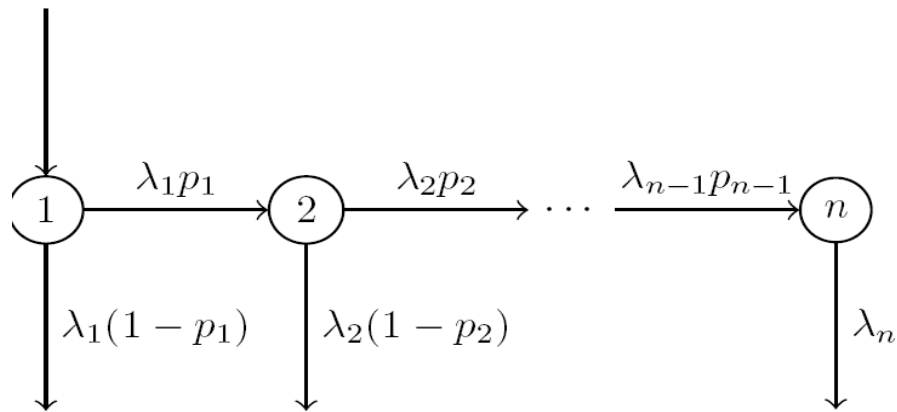
- **Coxian και γενικευμένη Coxian κατανομή**

Η κλάση των κατανομών Coxian είναι δημοφιλής στη βιβλιογραφία των εφαρμοσμένων επιστημών (δείτε, π.χ., Buchholz et al. (2014)). Η Coxian κατανομή ανήκει στη κλάση των συνεχών κατανομών τύπου φάσεων και αποτελεί μία γενίκευση της υποεκθετικής κατανομής με την έννοια ότι υπάρχει δυνατότητα απορρόφησης από κάθε μεταβατική κατάσταση. Η κατανομή Coxian τάξης n έχει αναπαράσταση $\mathbf{a} = (1, 0, \dots, 0)$ και

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 p_1 & \ddots & & & \\ & -\lambda_2 & & \ddots & & \\ & & & & \ddots & \\ & & & & & -\lambda_{n-1} & \lambda_{n-1} p_{n-1} \\ & & & & & & -\lambda_n \end{pmatrix}_{n \times n}, \quad \mathbf{t} = \begin{pmatrix} \lambda_1(1 - p_1) \\ \vdots \\ \lambda_{n-1}(1 - p_{n-1}) \\ \lambda_n \end{pmatrix}_{1 \times n}$$

(τα υπόλοιπα στοιχεία του πίνακα \mathbf{T} είναι μηδέν). Για τα p_i ισχύει ότι $0 \leq p_i \leq 1$ ($i = 1, \dots, n - 1$). Στην περίπτωση που για όλα τα p_i είναι 1 αναγόμαστε στην υποεκθετική κατανομή.

Το διάγραμμα φάσεων της κατανομής Coxian είναι το ακόλουθο

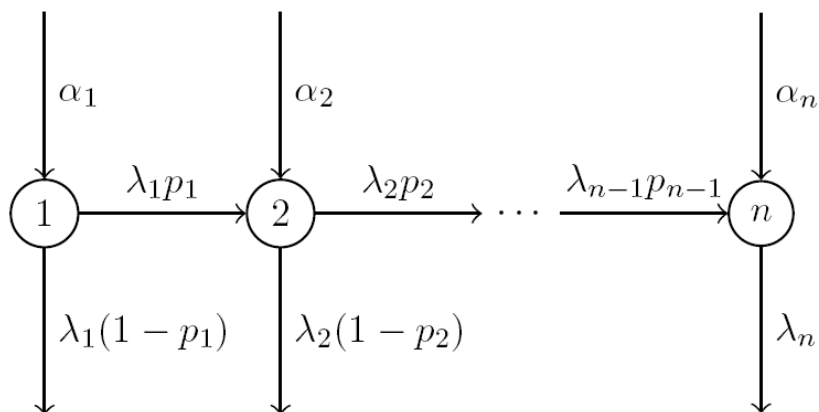


Σχήμα 1.9. Διάγραμμα φάσης κατανομής Coxian

Όσον αφορά τη γενικευμένη Coxian κατανομή η μόνη διαφορά με την Coxian κατανομή είναι ότι η αρχική κατάσταση της αλυσίδας μπορεί να είναι σε οποιαδήποτε από τις μεταβατικές καταστάσεις. Σε αυτή την περίπτωση έχουμε ότι

$$\mathbf{a} = (a_1, a_2, \dots, a_m)$$

και ένα τυπικό διάγραμμα φάσης είναι το ακόλουθο



Σχήμα 1.10. Διάγραμμα φάσης γενικευμένης κατανομής Coxian

1.4.6 Αριθμητικό παράδειγμα συνεχούς κατανομής τύπου φάσεων

Έστω Y μια συνεχής κατανομής τύπου φάσεων διάστασης 2, δηλαδή $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, με αναπαράσταση

$$\mathbf{a} = (1/3, 2/3), \quad \mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix},$$

οπότε

$$\mathbf{t} = (6/7, 2).$$

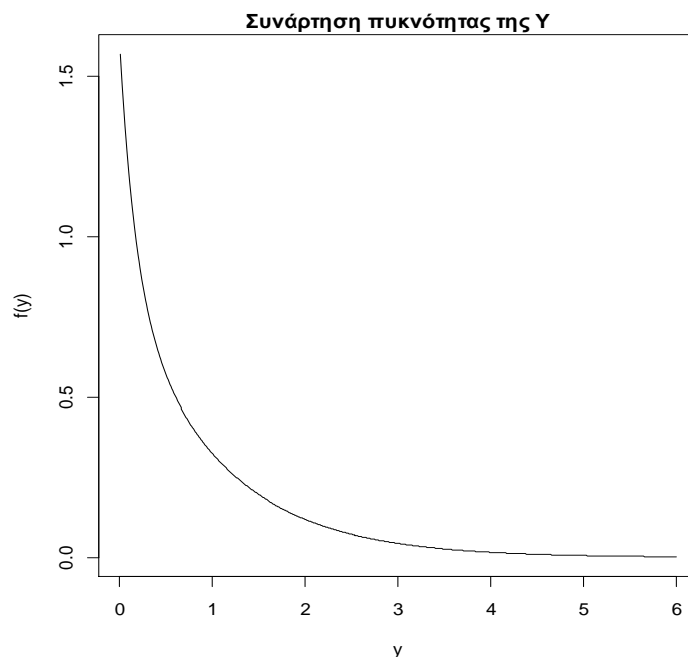
Από το Θεώρημα 1.6 προκύπτει ότι η συνάρτηση πυκνότητας της τυχαίας μεταβλητής Y μπορεί να βρεθεί από τον τύπο $f(y) = \mathbf{a}e^{\mathbf{T}y}\mathbf{t}$, $y > 0$. Χρησιμοποιώντας τη συνάρτηση `drhType` του πακέτου `actuar` της γλώσσας προγραμματισμού R βρίσκουμε τις ακόλουθες τιμές της συνάρτησης $f(y)$.

Πίνακας 1.3. Συνάρτηση πυκνότητας της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$ όπου

$$\mathbf{a} = (1/3, 2/3), \quad \mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$$

y	$f(y)$	y	$f(y)$
0.5	0.568	3.5	0.026
1	0.324	4	0.016
1.5	0.196	4.5	0.010
2	0.119	5	0.006
2.5	0.072	5.5	0.004
3	0.044	6	0.002

Ενδεικτικό είναι και το ακόλουθο σχήμα



Σχήμα 1.11. Συνάρτηση πυκνότητας της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$ όπου

$$\mathbf{a} = (1/3, 2/3), \quad \mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$$

Σημειώνουμε ότι γράφοντας τον πίνακα \mathbf{T} στη διαγώνια μορφή του

$$\mathbf{T} = - \begin{pmatrix} 9/10 & 9/70 \\ 7/10 & 1/10 \end{pmatrix} - 6 \begin{pmatrix} 1/10 & -9/70 \\ -7/10 & 9/10 \end{pmatrix}$$

προκύπτει ότι

$$\begin{aligned} f(y) &= \mathbf{a}e^{\mathbf{T}y\mathbf{t}} \\ &= e^{-y}(1/3, 2/3) \begin{pmatrix} 9/10 & 9/70 \\ 7/10 & 1/10 \end{pmatrix} \begin{pmatrix} 6/7 \\ 2 \end{pmatrix} + e^{-6y}(1/3, 2/3) \begin{pmatrix} 1/10 & -9/70 \\ -7/10 & 9/10 \end{pmatrix} \begin{pmatrix} 6/7 \\ 2 \end{pmatrix} \\ &= \frac{92}{105}e^{-y} + \frac{26}{35}e^{-6y}. \end{aligned}$$

Οι ροπές r τάξης της Y προκύπτουν από το Θεώρημα 1.8. Ενδεικτικός είναι ο ακόλουθος πίνακας.

Πίνακας 1.4. Ροπές της $Y \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, όπου $\mathbf{a} = (1/3, 2/3)$, $\mathbf{T} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}$

r	1	2	3	4	5
$\mu_r = E(Y^r)$	0.8968	1.7593	5.2606	21.0309	105.1448

ΚΕΦΑΛΑΙΟ 2

Μοντέλο συλλογικού κινδύνου και κατανομές τύπου φάσεων

2.1 Εισαγωγή

Στόχος του παρόντος κεφαλαίου είναι η μελέτη του κλασικού μοντέλου της θεωρίας κινδύνου και η εύρεση της κατανομής των συνολικών απαιτήσεων σε διάφορες περιπτώσεις. Το μοντέλο συλλογικού κινδύνου εξετάζει τα επιμέρους τμήματα του χαρτοφυλακίου και τα αντιμετωπίζει ως ένα ενιαίο σύνολο αποζημιώσεων. Επιπροσθέτως είναι σημαντικό να αναφερθεί πως οι αποζημιώσεις είναι ανεξάρτητες και ισόνομες.

Η κατανομή του συνολικού ποσού απαίτησης μιας ασφάλισης του χαρτοφυλακίου, περιγράφεται από το συλλογικό μοντέλο το οποίο παράγει έναν τυχαίο αριθμό N απαιτήσεων σε μια συγκεκριμένη χρονική περίοδο. Οι όροι του μοντέλου αντιστοιχούν στις πραγματικές απαιτήσεις, ενώ στην περίπτωση όπου δεν σημειωθεί καμία απαίτηση, το ποσό S που θα καταβάλλει η εταιρία θα είναι μηδέν. Στην πράξη, η κατανομή Poisson είναι η πιο συχνά χρησιμοποιούμενη σχετικά με τις απαιτήσεις του χαρτοφυλακίου. Λόγω της έλλειψης διαθέσιμων δεδομένων η οποία δυσχεραίνει τους υπολογισμούς, χρησιμοποιούμε την κατανομή Poisson ως πρώτη επιλογή, η οποία απαιτεί την εκτίμηση μόνο μίας παραμέτρου.

Στη συνέχεια του κεφαλαίου θα αναλύσουμε το μοντέλο και θα δούμε πως αυτό σχετίζεται με τις κατανομές τύπου φάσεων και μέσω των στατιστικών πακέτων `actuar` και `PhaseType` του R θα εξετάσουμε τις εφαρμογές μας.

2.2 Το μοντέλο συλλογικού κινδύνου

Η γενική μορφή του μοντέλου συλλογικού κινδύνου περιγράφεται από τη σύνθετη κατανομή

$$S_N = S = \begin{cases} X_1 + X_2 + \dots + X_N, & N = 1, 2, \dots \\ 0, & N = 0 \end{cases}$$

όπου:

- (i) η τυχαία μεταβλητή X_i παριστάνει το ύψος της i ατομικής ζημιάς (δεν μπορεί να πάρει αρνητικές τιμές), με συνάρτηση πιθανότητας/πυκνότητας f και συνάρτηση κατανομής F
- (ii) η τυχαία μεταβλητή N , παριστάνει το πλήθος των ζημιών (κινδύνων) ή και απαιτήσεων/αξιώσεων με συνάρτηση πιθανότητας p
- (iii) η τυχαία μεταβλητή S_N , παριστάνει το ύψος της συνολικής ζημιάς με συνάρτηση πιθανότητας/πυκνότητας g και συνάρτηση κατανομής G .

Συνήθως, γράφουμε

$$g = p \vee f.$$

Θα χρησιμοποιήσουμε τους ακόλουθους συμβολισμούς:

- $F(x) = Pr(X \leq x)$ η συνάρτηση κατανομής της X ,
- $f(x) = Pr(X = x)$ η συνάρτηση πιθανότητας/πυκνότητας της X ,
- $G(x) = Pr(S_N \leq x)$ η συνάρτηση κατανομής της S_N ,
- $g(x) = Pr(S = x)$ η συνάρτηση πιθανότητας/πυκνότητας της S_N ,

Ειδικότερα για την τυχαία μεταβλητή N , που είναι πάντοτε μια μη-αρνητική ακέραια (διακριτή) τυχαία μεταβλητή, συμβολίζουμε τη συνάρτηση πιθανότητάς ως

$$p_n = p(n) = P(N = n), \quad n = 0, 1, \dots$$

Είναι γνωστό ότι για την τ.μ.

$$S_n = X_1 + X_2 + \dots + X_n, \quad n \geq 1$$

ισχύει ότι

$$P(S_n \leq x) = F^{*n}(x) = \begin{cases} \int_0^x F^{*(n-1)}(x-y)f(y)dy, & X \text{ συνεχής τ.μ.} \\ \sum_{y=0}^x F^{*(n-1)}(x-y)f(y), & X \text{ διακριτή τ.μ.} \end{cases}$$

όπου

$$F^{*0}(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0. \end{cases}$$

Στο επόμενο θεώρημα δίνεται η συνάρτηση κατανομής του ύψους των συνολικών ζημιών S_N στο μοντέλο συλλογικού κινδύνου.

Θεώρημα 2.1

Η συνάρτηση κατανομής της τ.μ. S_N ικανοποιεί τη σχέση

$$G(x) = Pr(S_N \leq x) = \sum_{n=0}^{\infty} p_n F^{*n}(x), \quad x \geq 0.$$

Απόδειξη

Για $x \geq 0$ έχουμε

$$\begin{aligned} G(x) &= Pr(S_N \leq x) = \sum_{n=0}^{\infty} Pr(S_N \leq x | N = n) Pr(N = n) \\ &= p_0 Pr(S_N \leq x | N = 0) + \sum_{n=1}^{\infty} p_n Pr(X_1 + X_2 + \dots + X_N \leq x | N = n) \\ &= p_0 F^{*0}(x) + \sum_{n=1}^{\infty} p_n Pr(S_n \leq x) = \sum_{n=0}^{\infty} p_n F^{*n}(x). \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Αν τη τυχαία μεταβλητή X είναι συνεχής, τότε η τυχαία μεταβλητή S_N είναι τυχαία μεταβλητή μικτού τύπου με συνεχές μέρος στο διάστημα $(0, \infty)$ που δίνεται από τον τύπο

$$g(x) = \sum_{n=1}^{\infty} p_n f^{*n}(x), \quad x > 0.$$

και διακριτό μέρος πιθανότητα μόνο στο 0 ίση με

$$g(0) = P(N = 0) = p_0.$$

Στον παραπάνω τύπο

$$f^{*0}(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad f^{*n}(x) = \begin{cases} f(x), & n = 1 \\ \int_0^x f^{*(n-1)}(x-y)f(y)dy, & n = 2, 3, \dots \end{cases}$$

Αν τη τυχαία μεταβλητή X είναι διακριτή τότε και η τυχαία μεταβλητή S_N είναι διακριτή με συνάρτηση πιθανότητας που δίνεται από τον τύπο

$$g(x) = Pr(S_N = x) = \sum_{n=0}^{\infty} p_n f^{*n}(x), \quad x = 0, 1, 2, \dots$$

Ειδικότερα από τον παραπάνω τύπο προκύπτει ότι

$$g(0) = P_N(f(0))$$

όπου $P_N(\cdot)$ είναι η πιθανογεννήτρια της τ.μ. N .

Έστω $M_Y(t)$ και $P_Y(t)$ η ροπογεννήτρια και η πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής Y . Στο ακόλουθο θεώρημα δίνονται διάφορες σχέσεις που ικανοποιούν η ροπογεννήτρια και η πιθανογεννήτρια συνάρτηση του ύψους S_N των συνολικών ζημιών. Η τ.μ. S_N έχει πιθανογεννήτρια συνάρτηση όταν οι ατομικές ζημιές έχουν διακριτή κατανομή.

Θεώρημα 2.2

Έστω $M_{S_N}(\cdot)$ η ροπογεννήτρια συνάρτηση και $P_{S_N}(\cdot)$ η πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής S_N . Τότε ισχύουν οι παρακάτω σχέσεις:

- (i) $M_{S_N}(t) = M_N[\log M_X(t)]$.
- (ii) $M_{S_N}(t) = P_N[M_X(t)]$.
- (iii) $P_{S_N}(t) = P_N[P_X(t)]$.

Απόδειξη

Έχουμε ότι

$$\begin{aligned} M_{S_N}(t) &= E(e^{tS_N}) = E[E(e^{tS_N} | N)] \\ &= p_0 E(e^{t \cdot 0}) + \sum_{n=1}^{\infty} p_n E(e^{t(X_1 + X_2 + \dots + X_n)} | N = n) \\ &= p_0 + \sum_{n=1}^{\infty} p_n E(e^{t(X_1 + X_2 + \dots + X_n)}) = \sum_{n=0}^{\infty} p_n (M_X(t))^n \\ &= E[(M_X(t))^N] = P_N[M_X(t)]. \end{aligned}$$

Ακόμη

$$M_{S_N}(t) = E[(M_X(t))^N] = E[e^{\log(M_X(t))^N}] = E[e^{N \log(M_X(t))}] = M_N[\log M_X(t)].$$

Επίσης, αφού

$$M_{S_N}(t) = P_N[M_X(t)]$$

θέτοντας

$$t = \log u$$

παίρνουμε

$$P_{S_N}(u) = P_N[P_X(u)].$$

Το θεώρημα αποδείχθηκε. ■

Θεώρημα 2.3

Για τη μέση τιμή και τη διακύμανση της τυχαίας μεταβλητής S_N ισχύει ότι

$$E(S_N) = E(N)E(X)$$

και

$$Var(S_N) = E(N)Var(X) + Var(N)E^2(X).$$

Απόδειξη

Για τη μέση τιμή έχουμε ότι

$$\begin{aligned} E(S_N) &= E[E(S_N|N)] = \sum_{n=0}^{\infty} p_n E(S_N|N = n) \\ &= p_0 E(S_N|N = 0) + \sum_{n=1}^{\infty} p_n E(X_1 + X_2 + \dots + X_N|N = n) \\ &= p_0(0 \cdot Pr(S_N = 0)) + \sum_{n=1}^{\infty} p_n E(X_1 + X_2 + \dots + X_n) \\ &= \sum_{n=1}^{\infty} p_n (nE(X)) = E(NE(X)) = E(N)E(X). \end{aligned}$$

Για την διακύμανση έχουμε

$$\begin{aligned} Var(S_N) &= E[Var(S_N|N)] + Var[E(S_N|N)] \\ &= E[NVar(X)] + Var[NE(X)] \\ &= E(N)Var(X) + Var(N)E^2(X). \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Κλείνοντας την παρούσα παράγραφο δίνουμε ένα παράδειγμα εύρεσης της συνάρτησης πυκνότητας πιθανότητας και της συνάρτησης κατανομής των συνολικών ζημιών ενός χαρτοφυλακίου. Έστω ότι η συνάρτηση πυκνότητας της ατομικής ζημιάς είναι η

$$f(x) = \frac{1}{2} e^{-x/2}, \quad x > 0,$$

δηλαδή $X \sim \text{Exp}\left(\frac{1}{2}\right)$. Για το πλήθος των ζημιών N ισχύει ότι

$$p_n = \frac{2^{n-1}}{3^n} = \frac{1}{3} \left(\frac{2}{3}\right)^{n-1}, \quad n = 1, 2, \dots$$

δηλαδή $N \sim G\left(\frac{1}{3}\right)$. Οι συνολικές ζημιές του χαρτοφυλακίου δίνονται από τον τύπο

$$S_N = X_1 + X_2 + \dots + X_N, \quad N = 1, 2, \dots$$

και η τ.μ. S_N είναι συνεχής. Σύμφωνα με το Θεώρημα 2.2, έχουμε

$$\begin{aligned} M_{S_N}(t) &= P_N[M_X(t)] = \frac{p\left(\frac{\lambda}{\lambda-t}\right)}{1 - q\left(\frac{\lambda}{\lambda-t}\right)} = \frac{\frac{1}{3} \left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)}{1 - \frac{2}{3} \left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)} = \frac{\frac{1}{3} \left(\frac{1}{1-2t}\right)}{1 - \frac{2}{3} \left(\frac{1}{1-2t}\right)} \\ &= \frac{\frac{1}{3(1-2t)}}{\frac{3(1-2t) - 2}{3(1-2t)}} = \frac{1}{3(1-2t) - 2} = \frac{1}{1-6t} = \frac{\frac{1}{6}}{\frac{1}{6} - t}. \end{aligned}$$

Συνεπώς $S_N \sim \text{Exp}\left(\frac{1}{6}\right)$, δηλαδή

$$g(x) = \frac{1}{6} e^{-x/6}, \quad x > 0,$$

και

$$G(x) = P(S \leq X) = 1 - e^{-x/6} \quad x > 0.$$

2.3 Μοντέλο συλλογικού κινδύνου και η κλάση κατανομών $\mathfrak{R}(a, b, 0)$

Κάνοντας μία αναδρομή στο παρελθόν, βλέπουμε πως οι αναλογιστές χρησιμοποιούσαν τη βασική μέθοδο, δηλαδή αυτή των συνελίξεων, για να υπολογίσουν την κατανομή των συνολικών απαιτήσεων. Προκειμένου να υπολογιστεί η κατανομή του τυχαίου αθροίσματος S_N ευκολότερα, ο Panjer (1981) έθεσε τα θεμέλια μιας νέας προσέγγισης για τον υπολογισμό της κατανομής αυτής. Παρουσίασε ένα αναδρομικό τύπο υπολογισμού της συνάρτησης πιθανότητας $g(x)$ της τ.μ. S_N (η κατανομή των ατομικών αποζημιώσεων είναι διακριτή τ.μ.) θεωρώντας ότι η κατανομή της τ.μ. N ανήκει σε μία ευρεία οικογένεια κατανομών, την οικογένεια $\mathfrak{R}(a, b, 0)$.

Ορισμός 2.1

Μία διακριτή τ.μ. N με σύνολο τιμών $R_N = \{0,1,2, \dots\}$ λέμε ότι ανήκει στην κλάση κατανομών $\mathfrak{R}(a, b, 0)$, αν η συνάρτηση πιθανότητάς της $p_n = Pr(N = n)$ ικανοποιεί την παρακάτω αναδρομική σχέση

$$p_n = \left(a + \frac{b}{n}\right)p_{n-1}, \quad n \geq 1,$$

όπου a, b κατάλληλες σταθερές και $p_n = 0$ για κάθε $n < 0$. Η παραπάνω αναδρομική σχέση μαζί με κάποια αρχική πιθανότητα p_0 ορίζει πλήρως την κατανομή της τ.μ. N . Η παραπάνω κλάση κατανομών είναι γνωστή και ως **οικογένεια κατανομών του Panjer**.

Σύμφωνα με τους Sundt & Jewell (1981) η κλάση $\mathfrak{R}(a, b, 0)$ περιέχει μόνο τις κατανομές Poisson, διωνυμική, αρνητική διωνυμική και τη γεωμετρική κατανομή. Στον πίνακα που ακολουθεί δίνονται τα μέλη της κλάσης $\mathfrak{R}(a, b, 0)$.

Πίνακας 2.1. Η οικογένεια κατανομών του Panjer.

ΚΑΤΑΝΟΜΗ	a	b	p_0	p_n
$P(\lambda)$	0	λ	$e^{-\lambda}$	$e^{-\lambda} \frac{\lambda^n}{n!}$
$B(m, p)$	$\frac{-p}{q}$	$\frac{(m+1)p}{q}$	q^m	$\binom{m}{n} p^n (1-p)^{m-n}$
$NB_0(r, p)$	q	$(r-1)q$	p^r	$\binom{r+n-1}{n} (1-p)^n p^r$
$G_0(p)$	q	0	p	$(1-p)^n p$

Αν η τ.μ. N ανήκει στην κλάση κατανομών $\mathfrak{R}(a, b, 0)$, τότε η πιθανογεννήτρια συνάρτησή της, ικανοποιεί τη διαφορική εξίσωση

$$P'_N(u) = \frac{a+b}{1-au} P_N(u).$$

Από την παραπάνω σχέση προκύπτει ότι

$$E(N) = P'_N(1) = \frac{a+b}{1-a}$$

$$Var(N) = P''_N(1) + P'_N(1) - (P'_N(1))^2 = \frac{a+b}{(1-a)^2}.$$

Συνεπώς η μέση τιμή και η διασπορά του τυχαίου αθροίσματος S_N είναι

$$E(S) = \frac{a+b}{1-a} E(X)$$

και

$$\text{Var}(S) = \frac{a+b}{1-a} \text{Var}(X) + \frac{a+b}{(1-a)^2} E^2(X).$$

Για την κατανομή της τυχαίας μεταβλητής S_N στην περίπτωση που η κατανομή των ζημιών περιγράφεται από μια διακριτή τ.μ. X και το πλήθος ζημιών περιγράφεται από μια τ.μ. N που ανήκει στην κλάση κατανομών $\mathfrak{R}(a, b, 0)$ ισχύει το ακόλουθο αποτέλεσμα.

Θεώρημα 2.4

Αν η X είναι μια μη αρνητική ακέραια τυχαία μεταβλητή, δηλαδή $R_X = \{0, 1, 2, \dots\}$, με συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τυχαία μεταβλητή N ανήκει στην οικογένεια κατανομών $\mathfrak{R}(a, b, 0)$, τότε η $g(x) = P(S_N = x)$ ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{1}{1-af(0)} \sum_{y=1}^x \left(a + b \frac{y}{x}\right) f(y)g(x-y), & x = 1, 2, \dots \\ P_N(f(0)) & x = 0. \end{cases}$$

Στη συνέχεια δίνουμε όλες τις ειδικές περιπτώσεις του Θεωρήματος 2.4.

- $N \sim P(\lambda)$

Αν η τυχαία μεταβλητή $N \sim P(\lambda)$, $\lambda > 0$, δηλαδή

$$p_n = e^{-\lambda} \frac{\lambda^n}{n!}, \quad n = 0, 1, \dots$$

τότε η συνάρτηση πιθανότητας $g(x) = P(S = x)$ ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{\lambda}{x} \sum_{y=1}^x y f(y) g(x-y), & x = 1, 2, \dots \\ e^{\lambda(f(0)-1)}, & x = 0 \end{cases}$$

Για την εξαγωγή του παραπάνω τύπου χρησιμοποιήσαμε τις τιμές $a = 0$, $b = \lambda$ και το γεγονός ότι $P_N(u) = e^{\lambda(u-1)}$.

- $N \sim B(m, p)$

Αν η τυχαία μεταβλητή $N \sim B(m, p)$, $0 < p < 1$, $q = 1 - p$, δηλαδή

$$p_n = \binom{m}{n} p^n (1-p)^{m-n}, \quad n = 0, 1, \dots, m$$

τότε η συνάρτηση πιθανότητας $g(x) = P(S = x)$, ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{p}{q - pf(0)} \sum_{y=1}^x \left\{ (m+1) \frac{y}{x} - 1 \right\} f(y) g(x-y), & x = 1, 2, \dots \\ (q + pf(0))^m, & x = 0. \end{cases}$$

Για την εξαγωγή του παραπάνω τύπου χρησιμοποιήσαμε τις τιμές $a = -p/(1-p)$, $b = (m+1)p/(1-p)$ και το γεγονός ότι $P_N(u) = (q + pu)^m$.

- $N \sim NB_0(r, p)$

Αν η τυχαία μεταβλητή $N \sim NB_0(r, p)$, $r > 0$, $0 < p < 1$, $q = 1 - p$, δηλαδή

$$p_n = \binom{r+n-1}{n} p^r (1-p)^n, \quad n = 0, 1, \dots$$

τότε η συνάρτηση πιθανότητας $g(x) = P(S = x)$, ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{q}{1 - qf(0)} \sum_{y=1}^x \left\{ (r-1) \frac{y}{x} + 1 \right\} f(y) g(x-y), & x = 1, 2, \dots \\ \left(\frac{p}{1 - qf(0)} \right)^r, & x = 0. \end{cases}$$

Για την εξαγωγή του παραπάνω τύπου χρησιμοποιήσαμε τις τιμές $a = q$, $b = (r-1)q$ και το γεγονός ότι $P_N(u) = \left(\frac{p}{1-qu} \right)^r$.

- $N \sim G_0(p)$

Αν η τυχαία μεταβλητή $N \sim G_0(p)$, $0 < p < 1$, $q = 1 - p$, δηλαδή

$$p_n = p(1-p)^n, \quad n = 0, 1, \dots$$

τότε η συνάρτηση πιθανότητας $g(x) = P(S = x)$, ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{q}{1 - qf(0)} \sum_{y=1}^x f(y) g(x-y), & x = 1, 2, \dots \\ \frac{p}{1 - qf(0)}, & x = 0. \end{cases}$$

Για την εξαγωγή του παραπάνω τύπου χρησιμοποιήσαμε τις τιμές $a = q, 0$ και το γεγονός ότι $P_N(u) = \frac{p}{1-qu}$.

Για επιπλέον λεπτομέρειες και παραδείγματα (δείτε Χατζηκωνσταντινίδης, Ε. Πανεπιστημιακές Σημειώσεις στη Θεωρία Κινδύνων I, ΠΜΣ «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου»).

2.4 Μοντέλο συλλογικού κινδύνου και η κλάση κατανομών $\mathfrak{R}(a, b, 1)$

Οι Sundt και Jewell (1981), θεώρησαν την κλάση κατανομών $\mathfrak{R}(a, b, 1)$ που αποτελεί γενίκευση της κλάσης κατανομών $\mathfrak{R}(a, b, 0)$.

Ορισμός 2.2

Μία διακριτή τ.μ. N με σύνολο τιμών $R_N = \{0, 1, 2, \dots\}$ λέμε ότι ανήκει στην κλάση κατανομών $\mathfrak{R}(a, b, 1)$, αν η συνάρτηση πιθανότητάς της $p_n = Pr(N = n)$ ικανοποιεί την παρακάτω αναδρομική σχέση

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, \quad n \geq 2,$$

όπου a, b κατάλληλες σταθερές και $p_n = 0$ για κάθε $n < 0$. Η παραπάνω αναδρομική σχέση μαζί με κάποιες αρχικές πιθανότητες p_0 και p_1 ορίζει πλήρως την κατανομή της τ.μ. N .

Η κλάση κατανομών $\mathfrak{R}(a, b, 1)$, περιέχει ως μέλη της όλες τις κατανομές της κλάσης $\mathfrak{R}(a, b, 0)$, δηλαδή περιέχει ως μέλη της τις κατανομές

- Poisson,
- Διωνυμική,
- Αρνητική Διωνυμική,
- Γεωμετρική.

Επιπλέον, περιέχει ως μέλη της όλες τις zero-truncated και όλες τις zero-modified κατανομές, που αντιστοιχούν στην κλάση κατανομών $\mathfrak{R}(a, b, 0)$. Έτσι, η κλάση κατανομών $\mathfrak{R}(a, b, 1)$ περιέχει ως μέλη της και τις ακόλουθες κατανομές:

- Zero – truncated Διωνυμική,
- Zero – truncated Poisson,
- Zero – truncated Αρνητική Διωνυμική,

- Zero – truncated Γεωμετρική.
- Zero – modified Διωνυμική,
- Zero – modified Poisson,
- Zero – modified Αρνητική Διωνυμική,
- Zero – modified Γεωμετρική.

Σημειώνουμε, ότι για μια διακριτή τ.μ. X με συνάρτηση πιθανότητας

$$p_k = P(X = k), \quad k = 0, 1, 2, \dots$$

η αντίστοιχη zero-truncated κατανομή (κατανομή αποκομμένη στο 0), $X^{(T)}$, έχει συνάρτηση πιθανότητας που ικανοποιεί τη σχέση

$$p_k^{(T)} = P(X^{(T)} = k) = \begin{cases} 0, & k = 0 \\ \frac{1}{1 - p_0} p_k, & k = 1, 2, \dots \end{cases}$$

ενώ, η αντίστοιχη zero-modified κατανομή (κατανομή τροποποιημένη στο 0), $X^{(M)}$, έχει συνάρτηση πιθανότητας, που ικανοποιεί τη σχέση

$$p_k^{(M)} = P(X^{(M)} = k) = \begin{cases} p_0^{(M)}, & k = 0 \\ \frac{1 - p_0^{(M)}}{1 - p_0} p_k = (1 - p_0^{(M)}) p_k^{(T)}, & k = 1, 2, \dots \end{cases}$$

Ειδικότερα για την αποκομμένη στο μηδέν αρνητική διωνυμική κατανομή και την τροποποιημένη στο μηδέν αρνητική διωνυμική κατανομή, με την επέκταση του παραμετρικού χώρου της παραμέτρου r από $r > 0$ σε $r > -1$, χρησιμοποιείται ο όρος εκτεταμένη αρνητική διωνυμική κατανομή και τροποποιημένη στο μηδέν εκτεταμένη αρνητική διωνυμική κατανομή.

Ακολουθεί ο συνοπτικός πίνακας με τα μέλη της κλάσης $\mathfrak{R}(a, b, 1)$.

Πίνακας 2.2. Πίνακας με τις κατανομές που ανήκουν στην κλάση κατανομών $\mathfrak{R}(a, b, 1)$

ΚΑΤΑΝΟΜΗ	p_0	p_1	a	b	Περιορισμοί
<i>Geometric</i> [$G_0(p)$] $p_n = (1 - q)q^n$	$1 - q$	$(1 - q)q$	q	0	$0 < q < 1$
<i>ZT Geometric</i>	0	$1 - q$	q	0	$0 < q < 1$
<i>ZM Geometric</i>	p_0^M	$(1 - p_0^M)(1 - q)$	q	0	$0 < q < 1$
<i>Negative binomial</i> [$Nb_0(r, p)$] $p_n = \binom{n+r-1}{n} \times (1 - q)^r q^n$	$(1 - q)^r$	$rq(1 - q)^r$	q	$(r - 1)q$	$r > 0,$ $0 < q < 1$
<i>ZT Negative binomial</i> or <i>ETNB</i>	0	$\frac{rq(1 - q)^r}{1 - (1 - q)^r}$	q	$(r - 1)q$	$r > -1,$ $0 < q < 1$
<i>ZM Negative binomial</i> or <i>ZM ETNB</i>	p_0^M	$\frac{(1 - p_0^M)rq(1 - q)^r}{1 - (1 - q)^r}$	q	$(r - 1)q$	$r > -1,$ $0 < q < 1$
<i>Poisson</i> [$P(\lambda)$] $p_n = e^{-\lambda} \frac{\lambda^n}{n!}$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	0	λ	$\lambda > 0$
<i>ZT Poisson</i>	0	$\frac{\lambda}{e^\lambda - 1}$	0	λ	$\lambda > 0$
<i>ZM Poisson</i>	p_0^M	$(1 - p_0^M) \frac{\lambda}{e^\lambda - 1}$	0	λ	$\lambda > 0$
<i>Binomial</i> [$B(m, p)$] $p_n = \binom{m}{n} (1 - q)^n q^{m-n}$	q^m	$m(1 - q)q^{m-1}$	$-\frac{(1 - q)}{q}$	$\frac{(m + 1)(1 - q)}{q}$	$0 < q < 1$
<i>ZT Binomial</i>	0	$\frac{m(1 - q)q^{m-1}}{1 - q^m}$	$-\frac{(1 - q)}{q}$	$\frac{(m + 1)(1 - q)}{q}$	$0 < q < 1$
<i>ZM Binomial</i>	p_0^M	$(1 - p_0^M) \times \frac{m(1 - q)q^{m-1}}{1 - q^m}$	$-\frac{(1 - q)}{q}$	$\frac{(m + 1)(1 - q)}{q}$	$0 < q < 1$

Logarithmic [$LS(q)$]	0	$-\frac{q}{\ln(1-q)}$	q	$-q$	$0 < q < 1$
$p_n = -\frac{q^n}{n \ln(1-q)}$					
ZM Logarithmic	p_0^M	$-(1 - p_0^M) \frac{q}{\ln(1-q)}$	q	$-q$	$0 < q < 1$

Για την κλάση κατανομών $\mathfrak{R}(a, b, 1)$ ισχύει το ακόλουθο αποτέλεσμα.

Θεώρημα 2.5

Αν η X είναι μια μη αρνητική ακέραια τυχαία μεταβλητή, δηλαδή $R_X = \{0, 1, 2, \dots\}$, με συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τυχαία μεταβλητή N ανήκει στην οικογένεια κατανομών $\mathfrak{R}(a, b, 1)$, τότε η $g(x) = P(S_N = x)$ ικανοποιεί την αναδρομική σχέση

$$g(x) = \begin{cases} \frac{p_1 - (a+b)p_0}{1 - af(0)} f(x) + \frac{1}{1 - af(0)} \sum_{y=1}^x \left(a + b \frac{y}{x}\right) f(y) g(x-y), & x = 1, 2, \dots \\ P_N(f(0)) & x = 0. \end{cases}$$

Γενικεύσεις του παραπάνω αναδρομικού σχήματος έχουν προταθεί από πολλούς ερευνητές. Μια από τις σημαντικότερες γενικεύσεις είναι αυτή του Sundt (1992, 2003). Ο Sundt (1992) θεώρησε ότι η τ.μ. N ικανοποιεί την αναδρομική σχέση

$$p_n = \sum_{i=1}^{r \wedge n} \left(a_i + \frac{b_i}{n}\right) p_{n-1}, \quad n = 1, 2, \dots$$

για κάποιο ακέραιο αριθμό r και παραμέτρους a_i και b_i ($i = 1, 2, \dots, k$). Τότε, η συνάρτηση πιθανότητας $g(x) = P(S_N = x)$ ικανοποιεί το αναδρομικό σχήμα

$$g(x) = \begin{cases} P_N(f(0)), & x = 0, \\ \frac{1}{1 - \sum_{i=1}^k a_i f^i(0)} \sum_{y=1}^x g(x-y) \sum_{i=1}^k \left(a_i + \frac{b_i y}{i x}\right) f^{*i}(y), & x = 1, 2, \dots \end{cases}$$

2.4 Μοντέλο συλλογικού κινδύνου και κατανομές τύπου φάσεων

Στην παρούσα παράγραφο θα εξετάσουμε τη σύνθετη κατανομή

$$S_N = S = \begin{cases} X_1 + X_2 + \dots + X_N, & N = 1, 2, \dots \\ 0, & N = 0 \end{cases}$$

όταν η τυχαία μεταβλητή N , που παριστάνει το πλήθος των ζημιών (κινδύνων) έχει μια διακριτή κατανομή τύπου φάσεων, και/ή η τυχαία μεταβλητή X_i που παριστάνει το ύψος της i ατομικής ζημιάς έχει μια διακριτή ή συνεχή κατανομή τύπου φάσεων. Θα παρουσιάσουμε ακριβείς τύπους αλλά και αναδρομικούς για τη σ.π. της σύνθετης κατανομής S_N .

Ο Eisele (2006) παρουσίασε αναδρομικές σχέσεις για τον υπολογισμό της σύνθετης κατανομής S_N στην περίπτωση που η κατανομή της τυχαίας μεταβλητής N είναι μια διακριτή κατανομή τύπου φάσεων και οι τυχαίες μεταβλητές $(X_i)_{i \geq 1}$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, ανεξάρτητες από την τυχαία μεταβλητή N , οι οποίες είτε είναι διακριτές με σύνολο τιμών $R_X = \{1, 2, 3, \dots\}$, είτε συνεχείς με σύνολο τιμών $R_X = (0, +\infty)$

Ας υποθέσουμε ότι η τυχαία μεταβλητή N έχει τη διακριτή κατανομή τύπου φάσεων $PH_d(\mathbf{a}, \mathbf{T})$. Από το Θεώρημα 1.2 έχουμε ότι η πιθανογεννήτρια της τ.μ. N δίνεται από τη σχέση

$$P_N(s) = a_{m+1} + \mathbf{a}(s^{-1}\mathbf{I} - \mathbf{T})^{-1}\mathbf{t} = a_{m+1} + \mathbf{a}(s^{-1}\mathbf{I} - \mathbf{T})^{-1}(\mathbf{I} - \mathbf{T})\mathbf{1}.$$

Από τον παραπάνω τύπο προκύπτει ότι η $P_N(s)$ είναι μια ρητή συνάρτηση που έχει ως παρονομαστή την ορίζουσα του πίνακα $(s^{-1}\mathbf{I} - \mathbf{T})^{-1}$. Η ορίζουσα του πίνακα $(s\mathbf{I} - \mathbf{T})$ είναι ένα πολυώνυμο με βαθμό m (είναι το χαρακτηριστικό πολυώνυμο του πίνακα \mathbf{T}), έστω

$$\det(s\mathbf{I} - \mathbf{T}) = s^m + b_1s^{m-1} + b_2s^{m-2} + \dots + b_{m-1}s + b_m.$$

Ειδικότερα

$$b_1 = -\text{trace}(\mathbf{T}), \quad b_m = (-1)^m \det(\mathbf{T}).$$

Έτσι η πιθανογεννήτρια της τ.μ. N μπορεί να γραφτεί στη μορφή

$$P_N(s) = a_{m+1} + \frac{c_1s + c_2s^2 + \dots + c_ms^m}{1 + b_1s + b_2s^2 + \dots + b_ms^m}$$

για κατάλληλες τιμές των συντελεστών c_i ($1 \leq i \leq m$) (δείτε Asmussen (2000), Eisele (2006) και Hipp, C. (2006)). Είναι εύκολο να δειχθεί ότι

$$c_1 = Pr(N = 1)$$

και για $i = 2, 3, \dots, m$,

$$c_i = Pr(N = i) + \sum_{k=1}^{i-1} b_k Pr(N = i - k).$$

Στο επόμενο θεώρημα δίνεται η κατανομή του τυχαίου αθροίσματος S_N στην περίπτωση που η κατανομή της τυχαίας μεταβλητής X είναι διακριτή.

Θεώρημα 2.6

Έστω ότι η τυχαία μεταβλητή N έχει κατανομή $PH_d^{(m)}(\mathbf{a}, \mathbf{T})$ και η τυχαία μεταβλητή X είναι διακριτή με σύνολο τιμών $R_X = \{1, 2, 3, \dots\}$. Τότε η συνάρτηση πιθανότητας $g(t)$ της τυχαίας μεταβλητής S_N ικανοποιεί το αναδρομικό σχήμα

$$g(t) = Pr(S_N = t) = \sum_{j=1}^{m\wedge t} c_j f^{*j}(t) - \sum_{j=1}^{m\wedge(t-1)} b_j \left(\sum_{u=1}^{t-1} g(u) f^{*j}(t-u) \right), \quad t = 1, 2, \dots$$

όπου

$$g(0) = Pr(S_N = 0) = Pr(N = 0) = a_{m+1}.$$

Απόδειξη

Για την πιθανογεννήτρια $P_{S_N}(z)$ της τυχαίας μεταβλητής S_N έχουμε ότι

$$P_{S_N}(z) = P_N(P_X(z)) = a_{m+1} + \frac{\sum_{j=1}^m c_j P_X^j(z)}{1 + \sum_{j=1}^m b_j P_X^j(z)}$$

οπότε

$$P_{S_N}(z) = a_{m+1} + \sum_{j=1}^m c_j P_X^j(z) - \sum_{j=1}^m b_j (P_{S_N}(z) - a_{m+1}) P_X^j(z).$$

Αφού $g(0) = a_{m+1}$ και $P_X^j(z) = P_{X^{*j}}(z) = \sum_{t=1}^{\infty} z^t f^{*j}(t)$, παίρνουμε ότι

$$\begin{aligned} \sum_{t \geq 1} z^t g(t) &= \sum_{t \geq 1} z^t \left(\sum_{j=1}^m c_j f^{*j}(t) \right) - \sum_{t \geq 1} z^t \left(\sum_{u=1}^{t-1} g(u) \left(\sum_{j=1}^m b_j f^{*j}(t-u) \right) \right) \\ &= \sum_{t \geq 1} z^t \left[\sum_{j=1}^{m\wedge t} c_j f^{*j}(t) - \sum_{j=1}^{m\wedge(t-1)} b_j \left(\sum_{u=1}^{t-1} g(u) f^{*j}(t-u) \right) \right]. \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Στο επόμενο θεώρημα δίνεται μια αναδρομική σχέση για την κατανομή του τυχαίου αθροίσματος S_N στην περίπτωση που η κατανομή της τυχαίας μεταβλητής X είναι συνεχής. Σε αυτή την περίπτωση η τυχαία μεταβλητή S_N είναι (ενδεχομένως) μικτού τύπου με μάζα πιθανότητας στο 0 ίση με

$$g(0) = Pr(S_N = 0) = Pr(N = 0) = a_{m+1}$$

και συνεχές μέρος στο διάστημα $(0, +\infty)$ που δίνεται στο επόμενο θεώρημα.

Θεώρημα 2.7

Έστω ότι η τυχαία μεταβλητή N έχει κατανομή $PH_a^{(m)}(\mathbf{a}, \mathbf{T})$ και η τυχαία μεταβλητή X είναι συνεχής στο $(0, +\infty)$. Τότε το συνεχές μέρος $g(t)$ στο $(0, +\infty)$, της τυχαίας μεταβλητής της τυχαίας μεταβλητής S_N ικανοποιεί το αναδρομικό σχήμα

$$g(t) = \sum_{j=1}^m c_j f^{*j}(t) - \sum_{j=1}^m b_j \int_0^t g(u) f^{*j}(t-u) du.$$

Απόδειξη

Για την απόδειξη θα χρησιμοποιήσουμε το μετασχηματισμό Laplace της τυχαίας μεταβλητής της τυχαίας μεταβλητής S_N που δίνεται από τον τύπο

$$L_{S_N}(r) = E(e^{rS_N}) = a_{m+1} + \int_0^{\infty} e^{rt} g(t) dt.$$

Όμως $L_{S_N}(r) = P_N(L_X(r))$, οπότε

$$L_{S_N}(r) = \pi_0 + \frac{\sum_{j=1}^m c_j L_X^j(r)}{1 + \sum_{j=1}^m b_j L_X^j(r)}.$$

Έτσι

$$L_{S_N}(r) = a_{m+1} + \sum_{j=1}^m c_j L_X^j(r) - \sum_{j=1}^m b_j (L_{S_N}(r) - a_{m+1}) L_X^j(r)$$

ή ισοδύναμα

$$\begin{aligned} \int_0^{\infty} e^{rt} g(t) dt &= \\ &= \sum_{j=1}^m c_j \int_0^{\infty} e^{rt} f^{*j}(t) dt - \sum_{j=1}^m b_j \int_0^{\infty} e^{rt} \left(\int_0^t g(u) f^{*j}(t-u) du \right) dt \end{aligned}$$

Το θεώρημα αποδείχθηκε. ■

Στη συνέχεια δίνουμε δύο θεωρήματα που αφορούν την ακριβή κατανομή του τυχαίου αθροίσματος S_N . Το Θεώρημα 2.8 (αντ. 2.9) αφορά την περίπτωση που η τυχαία μεταβλητή X έχει μια συνεχή (αντ. διακριτή) κατανομή τύπου φάσεων.

Θεώρημα 2.8

Έστω ότι η τυχαία μεταβλητή N έχει κατανομή $PH_d^{(n)}(\boldsymbol{\beta}, \mathbf{B})$, $\beta_{n+1} = 1 - \boldsymbol{\beta}\mathbf{1}$ και $\mathbf{b} = \mathbf{1} - \mathbf{B}\mathbf{1}$.

Έστω επίσης ότι η τυχαία μεταβλητή X έχει κατανομή $PH_c^{(m)}(\mathbf{a}, \mathbf{T})$, $a_{m+1} = 1 - \mathbf{a}\mathbf{1}$ και $\mathbf{t} = -\mathbf{T}\mathbf{1}$. Τότε η τυχαίας μεταβλητής S_N έχει την κατανομή $PH_c^{(mn)}(\boldsymbol{\gamma}, \mathbf{C})$, όπου

$$\boldsymbol{\gamma} = \boldsymbol{\alpha} \otimes \boldsymbol{\beta} (\mathbf{I} - a_{m+1} \mathbf{B})^{-1}$$

και

$$\mathbf{C} = \mathbf{T} \otimes \mathbf{I} + (1 - a_{m+1}) \mathbf{t} \mathbf{a} \otimes (\mathbf{I} - a_{m+1} \mathbf{B})^{-1} \mathbf{B}.$$

Η αρχική πιθανότητα εκκίνησης στην κατάσταση απορρόφησης $mn + 1$ δίνεται από τον τύπο

$$\gamma_{mn+1} = 1 - \boldsymbol{\gamma}\mathbf{1} = \beta_{n+1} + a_{m+1} \boldsymbol{\beta} (\mathbf{I} - a_{m+1} \mathbf{B})^{-1} \mathbf{b}$$

και το διάνυσμα των ρυθμών εξόδου δίνεται από τον τύπο

$$\mathbf{c} = -\mathbf{C}\mathbf{1} = \mathbf{t} \otimes (\mathbf{I} - a_{m+1} \mathbf{B})^{-1} \mathbf{b}.$$

Απόδειξη

Δείτε Latouche & Ramaswami (1987). ■

Θεώρημα 2.9

Έστω ότι η τυχαία μεταβλητή N έχει κατανομή $PH_d^{(n)}(\boldsymbol{\beta}, \mathbf{B})$, $\beta_{n+1} = 1 - \boldsymbol{\beta}\mathbf{1}$ και $\mathbf{b} = \mathbf{1} - \mathbf{B}\mathbf{1}$.

Έστω επίσης ότι η τυχαία μεταβλητή X έχει κατανομή $PH_d^{(m)}(\mathbf{a}, \mathbf{T})$, $a_{m+1} = 1 - \mathbf{a}\mathbf{1}$ και $\mathbf{t} = \mathbf{1} - \mathbf{T}\mathbf{1}$. Τότε η τυχαίας μεταβλητής S_N έχει την κατανομή $PH_d^{(mn)}(\boldsymbol{\gamma}, \mathbf{C})$, όπου

$$\boldsymbol{\gamma} = \boldsymbol{\alpha} \otimes \boldsymbol{\beta} (\mathbf{I} - a_{m+1} \mathbf{B})^{-1}$$

και

$$\mathbf{C} = \mathbf{T} \otimes \mathbf{I} + \mathbf{t} \mathbf{a} \otimes (\mathbf{I} - a_{m+1} \mathbf{B})^{-1} \mathbf{B}.$$

Απόδειξη

Δείτε Neuts (1981). ■

2.5 Εφαρμογές

Εφαρμογή 1 – Η μέθοδος του Eisele (2006)

Έστω ότι $X \sim PH_d^{(2)}(\mathbf{a}, \mathbf{T})$, $\mathbf{a} = (a_1, a_2)$, $a_3 = 1 - \mathbf{a}\mathbf{1} = 1 - (a_1 + a_2)$,

$$\mathbf{T} = \begin{pmatrix} p_1 & 1-p_1 \\ 0 & p_2 \end{pmatrix}, \quad \mathbf{t} = \mathbf{1} - \mathbf{T}\mathbf{1} = \begin{pmatrix} 0 \\ 1-p_2 \end{pmatrix},$$

με $p_1, p_2 \in (0,1)$.

Αρχικά βρίσκουμε το χαρακτηριστικό πολυώνυμο το οποίο είναι

$$p(s) = (s - p_1)(s - p_2) = s^2 - (p_1 + p_2)s + p_1p_2 = s^2 + b_1s + b_2$$

όπου $b_1 = -(p_1 + p_2)$ και $b_2 = p_1p_2$.

Επίσης, παρατηρούμε ότι για $k \geq 1$ έχουμε

$$\mathbf{T}^k = \begin{pmatrix} p_1^k & (1-p_1) \sum_{i=0}^{k-1} p_1^{k-1-i} p_2^i \\ 0 & p_2^k \end{pmatrix}.$$

Εφαρμόζοντας το Θεώρημα 1.1 παίρνουμε

$$Pr(X = 0) = a_3, \quad Pr(X = 1) = \mathbf{aT}^0\mathbf{t} = a_2(1-p_2).$$

και για $k \geq 2$ έχουμε

$$\begin{aligned} Pr(X = k) &= \mathbf{aT}^{k-1}\mathbf{t} = (a_1, a_2) \begin{pmatrix} p_1^{k-1} & (1-p_1) \sum_{i=0}^{k-2} p_1^{k-2-i} p_2^i \\ 0 & p_2^{k-1} \end{pmatrix} \begin{pmatrix} 0 \\ 1-p_2 \end{pmatrix} \\ &= a_1(1-p_1)(1-p_2) \sum_{i=0}^{k-2} p_1^{k-2-i} p_2^i + a_2(1-p_2)p_2^{k-1}. \end{aligned}$$

Επομένως

$$c_1 = Pr(X = 1) = a_2(1-p_2)$$

και

$$\begin{aligned} c_2 &= Pr(X = 2) + b_1Pr(X = 1) \\ &= a_1(1-p_1)(1-p_2) + a_2(1-p_2)p_2 - a_2(p_1 + p_2)(1-p_2) \\ &= a_1(1-p_1)(1-p_2) - a_2p_1(1-p_2) \end{aligned}$$

Συνεπώς, η πιθανογεννήτρια της τ.μ. X δίνεται από τον τύπο

$$P_X(s) = a_3 + \frac{c_1s + c_2s^2}{1 + b_1s + b_2s^2}$$

$$= a_3 + \frac{a_2(1 - p_2)s + [a_1(1 - p_1)(1 - p_2) - a_2p_1(1 - p_2)]s^2}{(1 - p_1s)(1 - p_2s)}$$

Στην ειδική περίπτωση που $p_1 = p_2 = p$, $a_1 = 2p(1 - p)^2$, $a_2 = 2p$ και $a_3 = (1 - p)^2$ έχουμε ότι

$$P_X(z) = \left(\frac{1 - p}{1 - ps} \right)^2$$

και

$$Pr(X = k) = (k + 1)(1 - p)^2 p^k = \binom{1 + k}{k} (1 - p)^2 p^k.$$

Επομένως $X \sim NB_0(r, 1 - p)$.

Εφαρμογή 2 – N, X: Διακριτές κατανομές τύπου φάσεων

Θα εφαρμόσουμε το Θεώρημα 2.9 στην περίπτωση που για την κατανομή της τ.μ. X (κατανομή των αποζημιώσεων) ισχύει ότι $X \sim PH_d^{(5)}(\mathbf{a}, \mathbf{T})$, όπου

$$a_{i+1} = \binom{5}{i} 0.25^i 0.75^{r-i}, \quad i = 0, 1, 2, 3, 4, 5$$

και

$$\mathbf{T} = \begin{pmatrix} 0.75 & 0.25 & 0 & 0 & 0 \\ 0 & 0.75 & 0.25 & 0 & 0 \\ 0 & 0 & 0.75 & 0.25 & 0 \\ 0 & 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0 & 0 & 0.75 \end{pmatrix}, \quad \mathbf{t} = \mathbf{1} - \mathbf{T}\mathbf{1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0.25 \end{pmatrix}.$$

Δηλαδή η $X \sim NB_0(5, 0.25)$ (δείτε Παράγραφο 1.4.2).

Για την κατανομή της τ.μ. N (κατανομή του πλήθους των αποζημιώσεων) θα θεωρήσουμε ότι είναι μια αυθαίρετη κατανομή τύπου φάσεων, δηλαδή $N \sim PH_d^{(5)}(\boldsymbol{\beta}, \mathbf{B})$, όπου

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) = (0.1, 0.2, 0.5, 0.05, 1), \quad \beta_6 = 0$$

και

$$\mathbf{B} = \begin{pmatrix} 0.2 & 0.4 & 0 & 0.4 & 0 \\ 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 0.5 \end{pmatrix}, \quad \mathbf{b} = \mathbf{1} - \mathbf{B}\mathbf{1} = \begin{pmatrix} 0 \\ 0 \\ 0.6 \\ 0 \\ 0.5 \end{pmatrix}.$$

Για την κατανομή των συνολικών αποζημιώσεων έχουμε ότι η τυχαία μεταβλητής S_N έχει την κατανομή $S_N \sim PH_d^{(25)}(\boldsymbol{\gamma}, \mathbf{C})$, όπου

Η συνάρτηση πιθανότητας του τυχαίου αθροίσματος δίνεται στον παρακάτω πίνακα

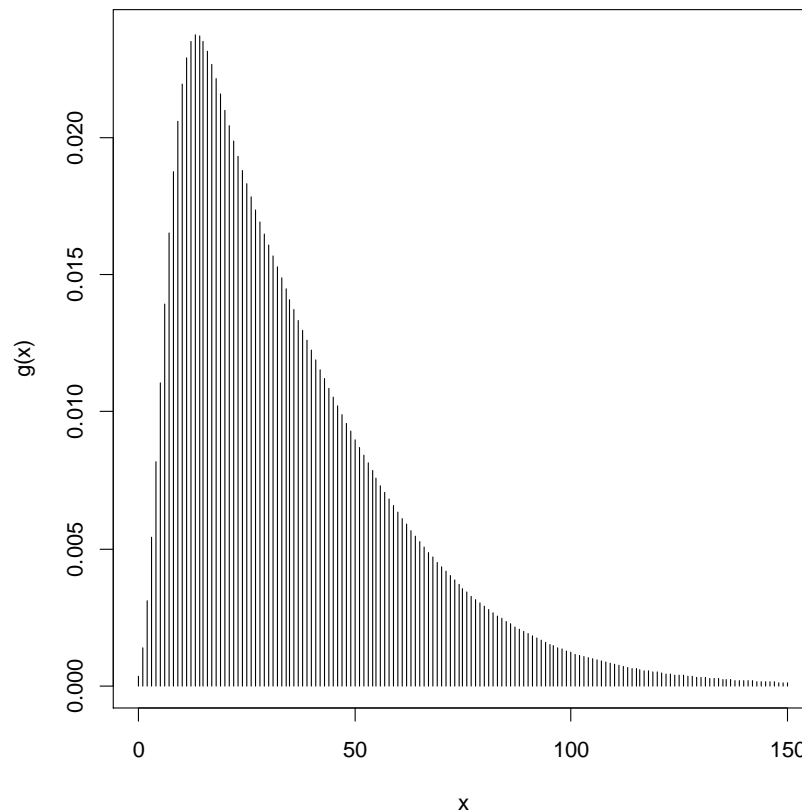
Πίνακας 2.3. Συνάρτηση πιθανότητας της $S_N \sim PH_d^{(25)}(\gamma, C)$

x	$g(x)$	x	$g(x)$	x	$g(x)$
1	0.0013751632	51	0.0086851955	101	0.0011748252
2	0.0030976310	52	0.0083986662	102	0.0011238525
3	0.0054305256	53	0.0081186153	103	0.0010749874
4	0.0081675744	54	0.0078450711	104	0.0010281507
5	0.0110687731	55	0.0075780522	105	0.0009832653
6	0.0139105712	56	0.0073175677	106	0.0009402568
7	0.0165148008	57	0.0070636168	107	0.0008990530
8	0.0187600097	58	0.0068161891	108	0.0008595839
9	0.0205805428	59	0.0065752646	109	0.0008217820
10	0.0219584129	60	0.0063408138	110	0.0007855819
11	0.0229119313	61	0.0061127980	111	0.0007509204
12	0.0234838520	62	0.0058911694	112	0.0007177364
13	0.0237307023	63	0.0056758715	113	0.0006859710
14	0.0237141382	64	0.0054668398	114	0.0006555673
15	0.0234945760	65	0.0052640019	115	0.0006264703
16	0.0231269810	66	0.0050672781	116	0.0005986271
17	0.0226584930	67	0.0048765821	117	0.0005719867
18	0.0221274871	68	0.0046918217	118	0.0005464999
19	0.0215636643	69	0.0045128988	119	0.0005221192
20	0.0209888105	70	0.0043397106	120	0.0004987990
21	0.0204179267	71	0.0041721500	121	0.0004764954
22	0.0198605049	72	0.0040101060	122	0.0004551662
23	0.0193217917	73	0.0038534644	123	0.0004347705
24	0.0188039386	74	0.0037021082	124	0.0004152694
25	0.0183069830	75	0.0035559182	125	0.0003966251
26	0.0178296424	76	0.0034147735	126	0.0003788016
27	0.0173699213	77	0.0032785518	127	0.0003617640
28	0.0169255513	78	0.0031471298	128	0.0003454790
29	0.0164942882	79	0.0030203838	129	0.0003299145
30	0.0160740957	80	0.0028981898	130	0.0003150397
31	0.0156632413	81	0.0027804240	131	0.0003008250
32	0.0152603305	82	0.0026669628	132	0.0002872422
33	0.0148643001	83	0.0025576837	133	0.0002742639
34	0.0144743848	84	0.0024524646	134	0.0002618640
35	0.0140900729	85	0.0023511851	135	0.0002500175
36	0.0137110562	86	0.0022537256	136	0.0002387003
37	0.0133371834	87	0.0021599685	137	0.0002278895
38	0.0129684171	88	0.0020697975	138	0.0002175629
39	0.0126047986	89	0.0019830984	139	0.0002076993
40	0.0122464197	90	0.0018997589	140	0.0001982785
41	0.0118934008	91	0.0018196686	141	0.0001892810
42	0.0115458754	92	0.0017427194	142	0.0001806882
43	0.0112039798	93	0.0016688054	143	0.0001724823
44	0.0108678457	94	0.0015978229	144	0.0001646462
45	0.0105375963	95	0.0015296706	145	0.0001571635
46	0.0102133442	96	0.0014642498	146	0.0001500187
47	0.0098951908	97	0.0014014638	147	0.0001431966
48	0.0095832259	98	0.0013412187	148	0.0001366830
49	0.0092775281	99	0.0012834229	149	0.0001304641
50	0.0089781654	100	0.0012279873	150	0.0001245267

Σημειώνουμε ότι

$$\begin{aligned} g(0) &= P(S_N = 0) = P_N(P(X = 0)) = P_N(0.25^5) \\ &= 0.25^5 + 0.25^5 \boldsymbol{\beta}(\mathbf{I} - 0.25^5 \mathbf{B})^{-1} \mathbf{b} = 0.0003664605 \end{aligned}$$

Η γραφική παράσταση της συνάρτησης πιθανότητας του τυχαίου αθροίσματος S_N δίνεται στο ακόλουθο σχήμα.



Σχήμα 2.1. Συνάρτηση πιθανότητας της $S_N \sim PH_d^{(25)}(\boldsymbol{\gamma}, \mathbf{C})$

Εφαρμογή 3 – X: Συνεχής κατανομή τύπου φάσεων, N: Διακριτή κατανομή τύπου φάσεων

Θα εφαρμόσουμε το Θεώρημα 2.8 στην περίπτωση που για την κατανομή της τ.μ. X (κατανομή των αποζημιώσεων) ισχύει ότι $X \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, όπου

$$\mathbf{a} = (a_1, a_2) = (1, 0), \quad a_3 = 0$$

και

$$\mathbf{T} = \begin{pmatrix} -3.5 & 3.5 \\ 0 & -3.5 \end{pmatrix}, \quad \mathbf{t} = -\mathbf{T}\mathbf{1} = \begin{pmatrix} 0 \\ 3.5 \end{pmatrix}.$$

Δηλαδή η $X \sim Erlang(2, 3.5)$ (δείτε Παράγραφο 1.4.5).

Για την κατανομή της τ.μ. N (κατανομή του πλήθους των αποζημιώσεων) θα θεωρήσουμε ότι είναι μια κατανομή τύπου φάσεων, δηλαδή $N \sim PH_d^{(1)}(\boldsymbol{\beta}, \mathbf{B})$, όπου

$$\boldsymbol{\beta} = \beta_1 = 0.75, \quad \mathbf{B} = 0.75, \quad \mathbf{b} = 0.25.$$

Δηλαδή η $N \sim G_0(0.25)$ (δείτε Παράγραφο 1.4.5).

Για την κατανομή των συνολικών αποζημιώσεων έχουμε ότι η τυχαία μεταβλητής S_N έχει την κατανομή $S_N \sim PH_c^{(2)}(\boldsymbol{\gamma}, \mathbf{C})$, όπου

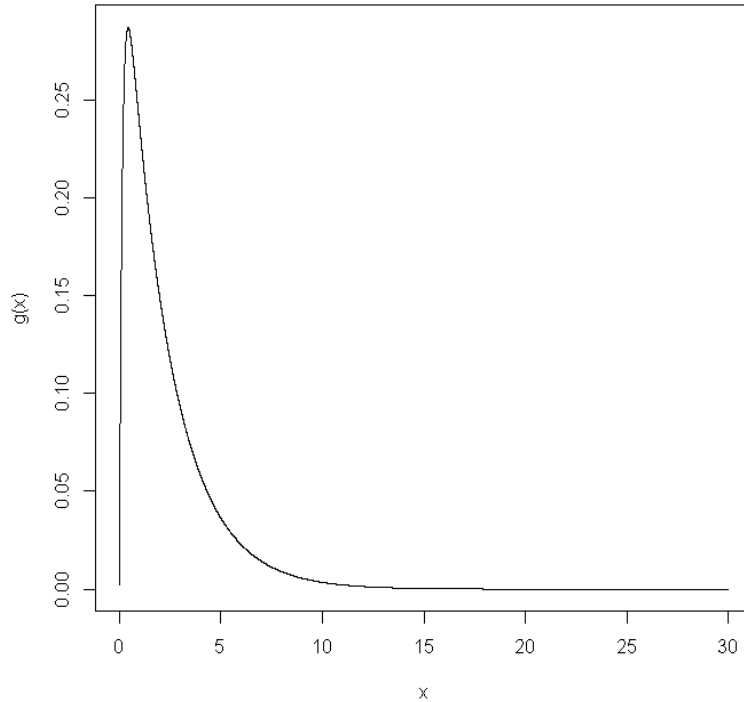
$$\boldsymbol{\gamma} = (0.75, 0), \quad \gamma_3 = 0.25, \quad \mathbf{C} = \begin{pmatrix} -3.5 & 3.5 \\ 2.625 & -3.5 \end{pmatrix}.$$

Ορισμένες τιμές της συνάρτησης πυκνότητας του τυχαίου αθροίσματος δίνεται στον παρακάτω πίνακα

Πίνακας 2.4. Συνάρτηση πυκνότητας της $S_N \sim PH_c^{(2)}(\boldsymbol{\gamma}, \mathbf{C})$

x	$g(x)$	x	$g(x)$
0.5	0.2852350352	8	0.0088990718
1	0.2365104855	8.5	0.0070391782
1.5	0.1874958315	9	0.0055679998
2	0.1483253349	9.5	0.0044042956
2.5	0.1173261532	10	0.0034838040
3	0.0928051744	10.5	0.0027556938
3.5	0.0734090236	11	0.0021797577
4	0.0580666403	11.5	0.0017241914
4.5	0.0459307937	12	0.0013638378
5	0.0363313221	12.5	0.0010787976
5.5	0.0287381266	13	0.0008533304
6	0.0227318983	13.5	0.0006749855
6.5	0.0179809633	14	0.0005339145
7	0.0142229671	14.5	0.0004223271
7.5	0.0112503869	15	0.0003340613

Η γραφική παράσταση του συνεχούς μέρους του τυχαίου αθροίσματος S_N δίνεται στο ακόλουθο σχήμα.



Σχήμα 2.2. Συνάρτηση πυκνότητας της $S_N \sim PH_c^{(2)}(\gamma, C)$

Σημειώνουμε ότι

$$g(0) = P(S_N = 0) = \gamma_3 = 0.25.$$

Εφαρμογή 4 – X: Συνεχής κατανομή τύπου φάσεων, N: Διακριτή κατανομή

Εδώ θα παρουσιάσουμε μια περίπτωση που δεν καλύπτει η γενική θεωρία που παρουσιάστηκε στην προηγούμενη παράγραφο. Για την κατανομή της τ.μ. X (κατανομή των αποζημιώσεων) έχουμε ότι $X \sim PH_c^{(3)}(\mathbf{a}, \mathbf{T})$, όπου

$$\mathbf{a} = (a_1, a_2) = (1, 0), \quad a_3 = 0$$

και

$$\mathbf{T} = \begin{pmatrix} -2 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -2 \end{pmatrix}, \quad \mathbf{t} = -\mathbf{T}\mathbf{1} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

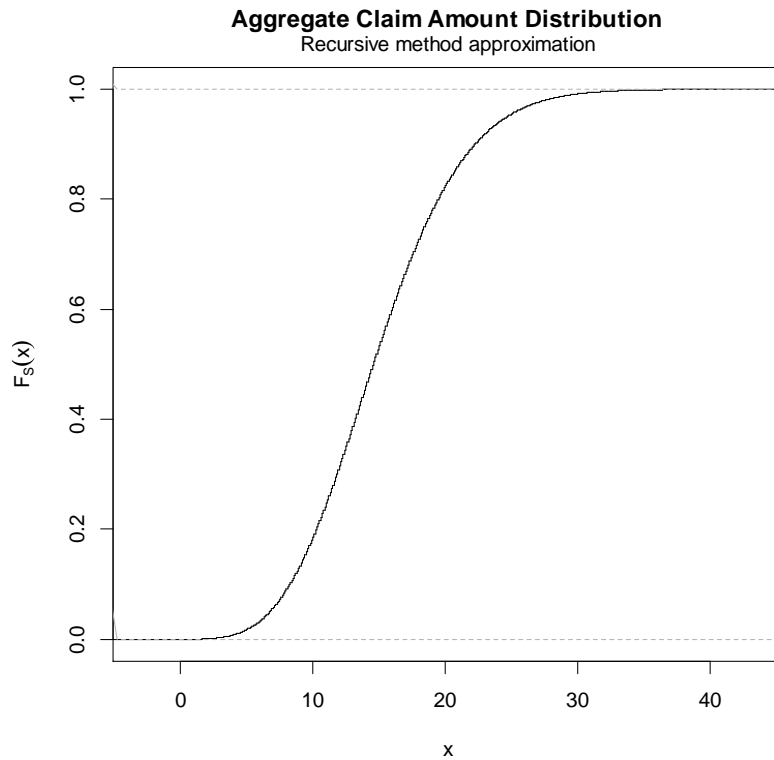
Δηλαδή η $X \sim Erlang(3, 2)$ (δείτε Παράγραφο 1.4.5).

Για την κατανομή της τ.μ. N (κατανομή του πλήθους των αποζημιώσεων) θα θεωρήσουμε ότι ακολουθεί την κατανομή Poisson με παράμετρο $\lambda = 10$. Η κατανομή Poisson δεν είναι κατανομή τύπου φάσεων και επομένως θα αναπτύξουμε νέα μεθοδολογία για την εύρεση της κατανομής των συνολικών αποζημιώσεων. Η απάντηση θα δοθεί με προσεγγιστική μέθοδο.

Αρχικά θα διακριτοποιήσουμε τη συνεχή κατανομή τύπου φάσεων με χρήση της συνάρτησης discretize της R, και στη συνέχεια θα χρησιμοποιήσουμε τη συνάρτηση aggregateDist για να υπολογίσουμε τη συνάρτηση κατανομής της τ.μ. S_N . Τα αποτελέσματα δίνονται στον ακόλουθο πίνακα.

Πίνακας 2.5. Συνάρτηση κατανομής της S_N

x	$G(x)$	x	$G(x)$
0	0.0000454702	16	0.6039540485
1	0.0002537808	17	0.6690575462
2	0.0011303916	18	0.7280358479
3	0.0034975916	19	0.7801255541
4	0.0086467857	20	0.8250505510
5	0.0182412196	21	0.8629408597
6	0.0341142703	22	0.8942329052
7	0.0579772903	23	0.9195673343
8	0.0911062001	24	0.9396960887
9	0.1340819042	25	0.9554057414
10	0.1866431784	26	0.9674599548
11	0.2476794141	27	0.9765607710
12	0.3153561720	28	0.9833264167
13	0.3873389877	29	0.9882822739
14	0.4610660948	30	0.9918614229
15	0.5340195753		



Σχήμα 2.3. Συνάρτηση κατανομής της S_N

Για τη διακριτοποίηση της τ.μ. X (κατανομή των αποζημιώσεων) χρησιμοποιήθηκε η μέθοδος "rounding" με $\text{step} = 0.1$, ενώ στη συνάρτηση `aggregateDist` χρησιμοποιήθηκε η μέθοδος "recursive". Ο σχετικός κώδικας δίνεται στο Παράρτημα, ενώ για επιπλέον λεπτομέρειες σχετικά με τη διακριτοποίηση συνεχών κατανομών μέσω της R (δείτε Dutang et al. (2008)).

ΚΕΦΑΛΑΙΟ 3

Θεωρία χρεοκοπίας και κατανομές τύπου φάσεων

3.1 Εισαγωγή

Στο συλλογικό πρότυπο της θεωρίας κινδύνου χρησιμοποιούμε την τυχαία μεταβλητή $S_N = \sum_{i=1}^N X_i$, η οποία έχει μία σύνθετη κατανομή, για να περιγράψουμε τις συνολικές απαιτήσεις που προέρχονται από ένα χαρτοφυλάκιο μιας ασφαλιστικής εταιρείας. Η τυχαία μεταβλητή N παριστάνει τον αριθμό των απαιτήσεων για αποζημιώσεις σε ένα προκαθορισμένο χρονικό διάστημα που μας ενδιαφέρει, και οι μεταβλητές X_i παριστάνουν τα ύψη των αποζημιώσεων και θεωρούνται ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, και ανεξάρτητες της N .

Στη θεωρία χρεοκοπίας εξετάζουμε όχι μόνο τις συνολικές αποζημιώσεις (έξοδα) ενός χαρτοφυλακίου όπως αυτές εξελίσσονται στο χρόνο αλλά και τα έσοδα της ασφαλιστικής εταιρείας. Για το λόγο αυτό χρησιμοποιούμε μία στοχαστική ανέλιξη, $\{S(t): t \geq 0\}$, και όχι μία τυχαία μεταβλητή για να περιγράψουμε στο χρόνο τις συνολικές αποζημιώσεις.

Επίσης η μεταβλητή N αντικαθίσταται από μία στοχαστική ανέλιξη ή αλλιώς απαριθμητρία ανέλιξη $\{N(t): t \geq 0\}$, η οποία μετρά τον αριθμό των απαιτήσεων προς την εταιρεία στο χρόνο.

Έτσι, η $\{S(t): t \geq 0\}$ ορίζεται για κάθε t από τη σχέση

$$S(t) = \begin{cases} \sum_{i=1}^{N(t)} X_i, & N(t) \geq 1 \\ 0, & N(t) = 0 \end{cases}$$

και είναι μία σύνθετη ανέλιξη. Το σημαντικότερο παράδειγμα μιας τέτοιας ανέλιξης, είναι αυτή στην οποία η $\{N(t)\}$ είναι μία ανέλιξη Poisson, οπότε τότε λέμε ότι η $\{S(t): t \geq 0\}$ ακολουθεί μία σύνθετη ανέλιξη Poisson.

3.2 Η στοχαστική ανέλιξη πλεονάσματος

3.2.1 Ορισμοί και το κλασσικό μοντέλο της θεωρίας κινδύνων

Θεωρούμε μία συνάρτηση $P(t)$ η οποία δηλώνει τα συνολικά ασφάλιστρα που εισρέουν στην εταιρεία στο διάστημα $[0, t]$. Αξίζει να σημειωθεί ότι τα ασφάλιστρα καθορίζονται με ακρίβεια από τον ασφαλιστή, δηλαδή, και σύμφωνα με τον Grandell (1991), δεν υπάρχει αβεβαιότητα ως προς την εξέλιξή τους στο χρόνο, για το λόγο αυτό η $P(t)$ είναι μία αύξουσα (μαθηματική) συνάρτηση και όχι μία στοχαστική ανέλιξη.

Στην παρούσα διπλωματική εργασία, η $P(t)$ είναι μία γραμμική συνάρτηση. Πριν προχωρήσουμε τον ορισμό της στοχαστικής ανέλιξης πλεονάσματος ας δώσουμε κάποιους συμβολισμούς.

- I. Συμβολίζουμε με c το ασφάλιστρο που εισρέει στην εταιρεία στη μονάδα χρόνου, έτσι ώστε το σύνολο των ασφαλίσεων που εισπράττει η εταιρεία στο χρονικό διάστημα $[0, t]$ είναι $P(t) = ct$.
- II. Θεωρούμε ότι για κάθε χαρτοφυλάκιο, η εταιρεία έχει ένα αποθεματικό u τη χρονική στιγμή 0, για να καλύψει την πιθανότητα να έρθει κάποια πολύ μεγάλη αποζημίωση στην αρχή της λειτουργίας του χαρτοφυλακίου.

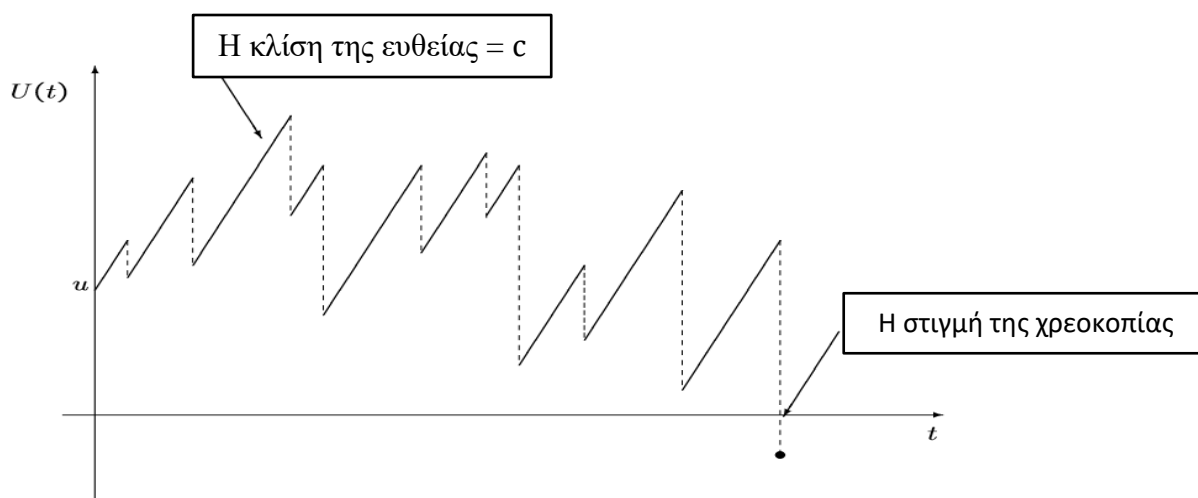
Ορισμός 3.1

Η στοχαστική ανέλιξη του πλεονάσματος $\{U(t), t \geq 0\}$ ορίζεται για κάθε $t \geq 0$ από τη σχέση

$$U(t) = u + P(t) - S(t)$$

όπου u είναι το αρχικό αποθεματικό, $P(t)$ τα συνολικά ασφάλιστρα που εισρέουν στη εταιρεία το χρονικό διάστημα $[0, t]$ και $S(t)$ η σύνθετη ανέλιξη για τις συνολικές αποζημιώσεις στο ίδιο χρονικό διάστημα.

Το $U(t)$ καλείται **αποθεματικό ή πλεόνασμα** τη χρονική στιγμή t , ενώ το $U(0) = u$ ($u \geq 0$) λέγεται **αρχικό αποθεματικό**.



Σχήμα 3.1. Στοχαστική διαδικασία πλεονάσματος

Στο κλασικό μοντέλο της θεωρίας κινδύνου θεωρούμε ότι τα ασφάλιστρα που εισπράττονται από τους ασφαλισμένους είναι η μοναδική πηγή εσόδων της ασφαλιστικής εταιρείας. Επίσης, ως μοναδική πηγή εξόδων θεωρούνται οι αποζημιώσεις που θα κληθεί να καταβάλλει η εταιρεία στους δικαιούχους σε περίπτωση κινδύνου.

Για το κλασικό μοντέλο της θεωρίας κινδύνων θεωρούμε ότι ισχύουν οι ακόλουθες υποθέσεις:

- ◇ Η συνάρτηση $P(t)$ του συνολικού ποσού των καταβληθέντων ασφαλίσεων στο διάστημα $[0, t]$ δίνεται από τη σχέση $P(t) = ct$, όπου $c > 0$ είναι μία σταθερά που εκφράζει το ασφάλιστρο που πληρώνεται στη μονάδα του χρόνου και ονομάζεται **ένταση του ασφαλίστρου (premium rate)**.
- ◇ Η απαριθμήτρια ζημιών $\{N(t): t \geq 0\}$ είναι μία ανέλιξη Poisson με ένταση λ , όπου η σταθερά λ εκφράζει τον αναμενόμενο αριθμό ζημιών στη μονάδα του χρόνου. Αυτό σημαίνει ότι η πιθανότητα να εμφανιστεί ένα ζημιογόνο γεγονός σε ένα διάστημα είναι ανάλογο του μήκους του διαστήματος. Συνεπώς, $N(t) \sim P(\lambda t)$, δηλαδή $Pr(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$, έτσι ώστε η $\{S(t): t \geq 0\}$ να είναι μία σύνθετη ανέλιξη Poisson.
- ◇ Οι τυχαίες μεταβλητές $X_i, i = 1, 2, \dots$, που δηλώνουν τα μεγέθη των αποζημιώσεων, είναι ανεξάρτητες και ισόνομες και είναι επίσης ανεξάρτητες από τον αριθμό των αποζημιώσεων $N(t)$ σε ένα διάστημα $[0, t]$.

Όταν οι τυχαίες μεταβλητές X_i είναι συνεχείς έχουμε τους εξής συμβολισμούς (αντίστοιχοι συμβολισμοί ισχύουν και στη διακριτή περίπτωση):

$$F(x) = Pr(X \leq x) = \int_0^x f(y)dy,$$

$$\bar{F}(x) = 1 - F(x) = 1 - Pr(X \leq x) = \int_x^{\infty} f(y)dy,$$

$$\mu_k = E(X^k) = \int_0^{\infty} y^k f(y)dy = \int_0^{\infty} \bar{F}(y)dy.$$

Ειδικότερα

$$\mu = \mu_1 = E(X) = \int_0^{\infty} yf(y)dy = \int_0^{\infty} \bar{F}(y)dy.$$

Μία βασική υπόθεση σχετικά με το κλασσικό μοντέλο της θεωρίας κινδύνων είναι ότι τα αναμενόμενα έξοδα της επιχείρησης δεν θα πρέπει να υπερβαίνουν τα έσοδα της. Απαιτούμε δηλαδή σε κάθε χρονική στιγμή τα ασφάλιστρα που εισπράττονται να είναι μεγαλύτερα κατά μέσο όρο από τις αποζημιώσεις που καταβάλλονται προς τους ασφαλισμένους. Δηλαδή, απαιτούμε να ισχύει η συνθήκη καθαρού κέρδους

$$ct > E(S(t)) = E(N(t))E(X) = \lambda t\mu$$

ή ισοδύναμα

$$c > \lambda\mu.$$

Από την παραπάνω βασική υπόθεση που κάνουμε πάντα στο κλασσικό μοντέλο, προκύπτει ότι το c ικανοποιεί μια σχέση της μορφής

$$c = (1 + \theta)\lambda\mu, \quad \theta > 0.$$

Το θ καλείται **περιθώριο ασφαλείας ή συντελεστής ασφαλείας (premium loading factor)**, και δίνεται προφανώς από τον τύπο

$$\theta = \frac{c}{\lambda\mu} - 1.$$

Το θ παίρνει συνήθως τιμές ανάμεσα σε 0 και 1 προκειμένου το χαρτοφυλάκιο να είναι ανταγωνιστικό, και αποτελεί ένα μέτρο έκφρασης του αναμενόμενου ποσοστού κέρδους της

ασφαλιστικής επιχείρησης για ένα συγκεκριμένο χαρτοφυλάκιο. Εκφράζει πόσο μεγαλύτερα είναι κατά μέσο όρο τα έσοδα της ασφαλιστικής εταιρείας σε σχέση με τα έξοδά της.

Σύμφωνα με τα παραπάνω η στοχαστική ανέλιξη του πλεονάσματος $\{U(t), t \geq 0\}$

$$U(t) = u + P(t) - S(t)$$

μπορεί να γραφεί στην ισοδύναμη μορφή

$$U(t) = u + (1 + \theta)\lambda\mu t - S(t).$$

3.2.2 Μελέτη της πιθανότητας χρεοκοπίας

Σκοπός μας σε αυτή την ενότητα είναι να δούμε πως ορίζονται κάποια θεμελιώδη μέτρα χρεοκοπίας της θεωρίας κινδύνων. Αρχικά δίνουμε ορισμούς για το χρόνο αλλά και την πιθανότητα χρεοκοπίας.

Ορισμός 3.2

Για $t \geq 0$, ο χρόνος T κατά τον οποίο για πρώτη φορά η στοχαστική διαδικασία του πλεονάσματος γίνεται αρνητική ορίζεται ως

$$T = \begin{cases} \inf\{t \geq 0: U(t) < 0\} \\ \infty, & U(t) \geq 0 \text{ για κάθε } t \geq 0. \end{cases}$$

Είναι προφανές ότι ο χρόνος χρεοκοπίας είναι μια ελαττωματική τυχαία μεταβλητή με την έννοια ότι

$$Pr(T = \infty) > 0.$$

Ορισμός 3.3

Για $u \geq 0$, η πιθανότητα χρεοκοπίας σε άπειρο χρόνο ορίζεται ως

$$\psi(u) = Pr(T < \infty | U(0) = u) = Pr(U(t) < 0 \text{ για κάποιο } t \geq 0 | U(0) = u).$$

Η πιθανότητα μη χρεοκοπίας ορίζεται ως

$$\delta(u) = 1 - \psi(u) = Pr(T = \infty | U(0) = u) = Pr(U(t) \geq 0 \text{ για κάθε } t \geq 0 | U(0) = u).$$

Παρατήρηση

Στο κλασικό μοντέλο, αν δεν ισχύει η σχέση $c > \lambda\mu$, τότε $\psi(u) = 1$ για κάθε u , ενώ αν ισχύει η σχέση $c > \lambda\mu$ τότε $\psi(u) < 1$ για κάθε u . Ακόμη η πιθανότητα χρεοκοπίας είναι φθίνουσα συνάρτηση του u και η πιθανότητα μη χρεοκοπίας είναι αύξουσα συνάρτηση του u . Επομένως

$$\lim_{u \rightarrow \infty} \psi(u) = 0$$

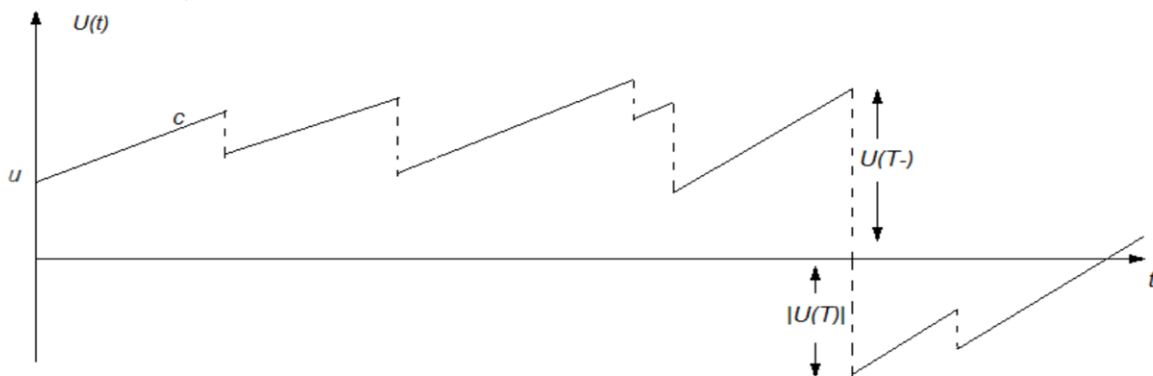
και

$$\lim_{u \rightarrow \infty} \delta(u) = 1$$

Τώρα, με $U(T)$ συμβολίζουμε την τυχαία μεταβλητή που δηλώνει το έλλειμμα τη στιγμή της χρεοκοπίας, δηλαδή δηλώνει το μέγεθος της πτώσης του πλεονάσματος κάτω από το μηδέν τη χρονική στιγμή $t = T$. Με $U(T-)$ συμβολίζουμε την τυχαία μεταβλητή που δηλώνει το πλεόνασμα λίγο πριν τη χρεοκοπία και δίνεται από τον τύπο

$$U(T-) = \lim_{t \rightarrow T-} U(t).$$

Η τυχαία μεταβλητή $U(T-)$ παίρνει μόνο θετικές τιμές.



Σχήμα 3.2. Στοχαστική διαδικασία πλεονάσματος συναρτήσει της τ.μ. του ελλείματος.

Σύμφωνα με το παραπάνω σχήμα το πλεόνασμα $U(t)$ εμφανίζει άλματα (προς τα κάτω) κατά τις χρονικές στιγμές επέλευσης των ζημιογόνων γεγονότων. Τα άλματα αυτά είναι του ίδιου μεγέθους με τα αντίστοιχα άλματα (προς τα πάνω) της διαδικασίας των συνολικών αποζημιώσεων. Ακόμη είναι φανερό ότι η διαδικασία πλεονάσματος μπορεί να γίνει αρνητική. Στην αναλογιστική ορολογία το ενδεχόμενο αυτό ονομάζεται χρεοκοπία και η πιθανότητα αυτού του ενδεχομένου πιθανότητα χρεοκοπίας.

Ορίζουμε τις τυχαίες μεταβλητές L_i ($i = 1, 2, \dots$) που εκφράζουν τις σταδιακές πτώσεις του πλεονάσματος από την τιμή του αρχικού αποθεματικού u μέχρι τη στιγμή της χρεοκοπίας T . Στην περίπτωση μη χρεοκοπίας οι τ.μ. L_i εκφράζουν το μέγεθος της πτώσης του πλεονάσματος

από την τιμή του αρχικού αποθεματικού u μέχρι την ελάχιστη τιμή που παίρνει η στοχαστική διαδικασία πλεονάσματος $\{U(t), t \geq 0\}$.

Συμβολίζουμε με W την διακριτή τ.μ. η οποία εκφράζει το πλήθος των σταδιακών πτώσεων του πλεονάσματος από την τιμή του αρχικού αποθεματικού u μέχρι τη στιγμή της χρεοκοπίας T , δηλαδή το πλήθος των μεταβλητών L_i σε μία ανέλιξη πλεονάσματος. Η κατανομή της W στο κλασικό μοντέλο της θεωρίας κινδύνων είναι η γεωμετρική κατανομή με πιθανότητα επιτυχίας $\delta(0)$, δηλαδή

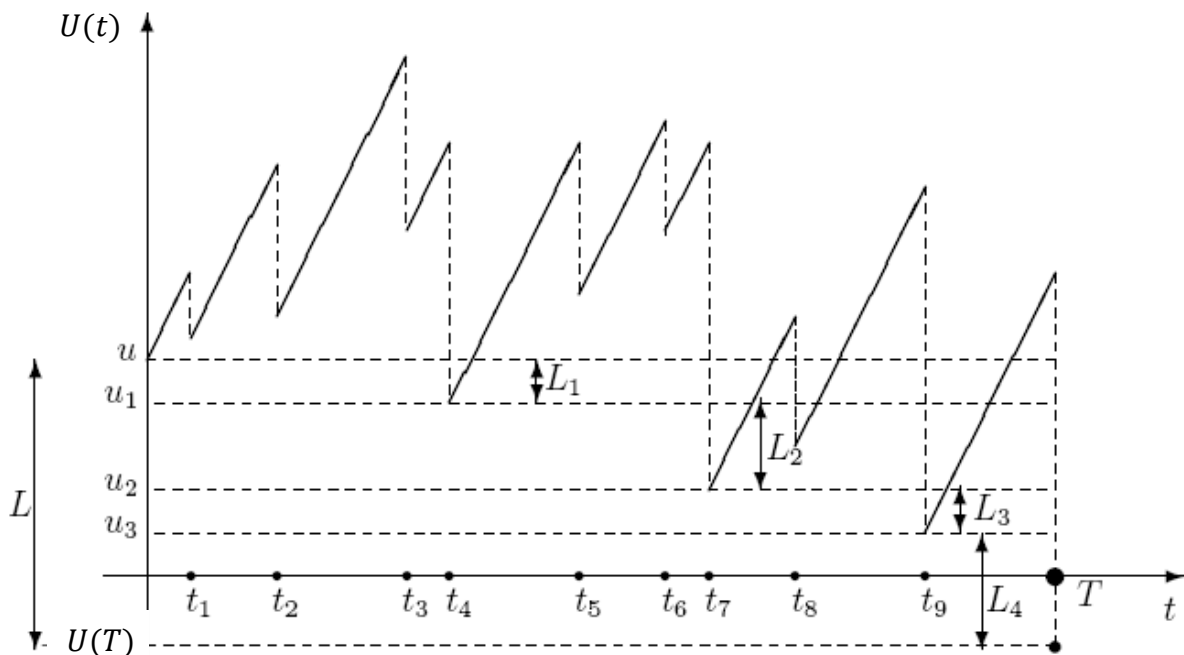
$$Pr(W = k) = (\psi(0))^k \delta(0) = \left(\frac{1}{1+\theta}\right)^k \left(\frac{\theta}{1+\theta}\right), \quad k = 0, 1, \dots$$

Όταν οι L_i ($i = 1, 2, \dots$) είναι μεταξύ τους ανεξάρτητες και ισόνομες και ανεξάρτητες της τ.μ. W τότε η σύνθετη τυχαία μεταβλητή L που ορίζεται από τον τύπο

$$L = \begin{cases} 0, & W = 0 \\ L_1 + L_2 + \dots + L_W, & W \geq 1 \end{cases}$$

ονομάζεται **μέγιστη σωρευτική απώλεια**.

Η τυχαία μεταβλητή L εκφράζει τη συνολική πτώση του πλεονάσματος κάτω από την τιμή του αρχικού αποθεματικού u έως την ελάχιστη τιμή της στοχαστικής διαδικασίας του πλεονάσματος $\{U(t), t \geq 0\}$. Ενδεικτικό είναι το ακόλουθο σχήμα.



Σχήμα 3.3. Μέγιστη σωρευτική απώλεια στη διαδικασία πλεονάσματος

Η κατανομή της τ.μ. L είναι μια σύνθετη γεωμετρική κατανομή με

$$Pr(L = 0) = Pr(W = 0) = \delta(0) = \frac{\theta}{1 + \theta}$$

και

$$Pr(L \leq u) = \delta(u),$$

$$Pr(L > u) = \psi(u) = \bar{F}_L(u).$$

Για την κατανομή των συνεχών τυχαίων μεταβλητών L_i έχουμε ότι

$$F_{L_i}(x) = \frac{1}{\mu} \int_0^x F_X(y) dy, \quad i = 1, 2, \dots$$

3.2.3 Οι τύποι του Lundberg

Στο κλασικό μοντέλο της θεωρίας κινδύνων ο **συντελεστής προσαρμογής** R ορίζεται να είναι η μικρότερη θετική ρίζα της εξίσωσης

$$M_X(r) = 1 + (1 + \theta)\mu r$$

η οποία είναι γνωστή ως **εξίσωση του Lundberg** ή **εξίσωση του συντελεστή προσαρμογής**.

Η ανισότητα Lundberg δίνει πληροφορίες για την πιθανότητα χρεοκοπίας με τη βοήθεια ενός άνω φράγματος συναρτήσεως του αρχικού αποθεματικού και του συντελεστή προσαρμογής. Στην περίπτωση που δεν υπάρχει αναλυτικός τύπος για τον υπολογισμό της πιθανότητας χρεοκοπίας η ανισότητα Lundberg εγγυάται ότι

$$\psi(u) \leq e^{-Ru}, \quad u \geq 0.$$

Παρατηρήσεις

- I. Για δεδομένη τιμή του R , όσο μεγαλώνει το αρχικό αποθεματικό τόσο μικραίνει η πιθανότητα χρεοκοπίας.
 - II. Για δεδομένη τιμή του αρχικού αποθεματικού, όσο μεγαλώνει το R τόσο μικραίνει η πιθανότητα χρεοκοπίας.
- Ένα άνω φράγμα για τον συντελεστή προσαρμογής R

Στο κλασικό μοντέλο της θεωρίας χρεοκοπίας εντοπίζουμε περιπτώσεις όπου ο συντελεστής προσαρμογής υπάρχει, ωστόσο δεν μπορεί να υπολογιστεί με ακρίβεια. Έτσι

για την διευκόλυνση των υπολογισμών υπολογίζουμε ένα άνω φράγμα για το R .

Χρησιμοποιούμε την αρχική εξίσωση της παραγράφου σε συνδυασμό με την εξίσωση της έντασης ασφαλιστρού, και έχουμε ότι

$$\lambda + cR = \lambda M_X(R)$$

όπου υποθέτουμε ότι οι αποζημιώσεις έχουν πυκνότητα f . Συνεπώς

$$\begin{aligned} \lambda + cR &= \lambda \int_0^{\infty} e^{Rx} f(x) dx > \lambda \int_0^{\infty} \left(1 + Rx + \frac{1}{2} R^2 x^2\right) f(x) dx \\ &= \lambda \left[\int_0^{\infty} f(x) dx + \int_0^{\infty} Rxf(x) dx + \int_0^{\infty} \frac{1}{2} R^2 x^2 f(x) dx \right] \\ &= \lambda \left(1 + R\mu + \frac{1}{2} R^2 \mu_2\right) \\ &\Rightarrow 2(c + \lambda\mu)R > \lambda R^2 \mu_2 \Leftrightarrow R < \frac{2(c + \lambda\mu)}{\lambda\mu_2}. \end{aligned}$$

Και χρησιμοποιώντας τη σχέση $c = (1 + \theta)\lambda\mu$, $\theta > 0$ καταλήγουμε στο άνω φράγμα του συντελεστή προσαρμογής που είναι

$$R < \frac{2\theta\mu}{\mu_2}.$$

- **Θεμελιώδης εξίσωση του Lundberg**

Η εξίσωση

$$cs + \lambda \hat{f}(s) - (\lambda + \delta) = 0$$

ορίζεται ως η **θεμελιώδης εξίσωση του Lundberg**, όπου $\hat{f}(s)$ είναι ο μετασχηματισμός Laplace της πυκνότητας f των αποζημιώσεων,

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} f(x) dx$$

και για $\delta > 0$ έχει μία θετική ρίζα $\rho = \rho(\delta)$ η οποία είναι αύξουσα.

Παρατηρήσεις

- Για $\delta = 0$ έχουμε $\rho = 0$.
- Για $\delta = 0$ θέτουμε ότι $l(s) = \lambda \hat{f}(s)$, με την συγκεκριμένη σχέση να είναι ισοδύναμη με την εξίσωση του Lundberg, με βάση την οποία ορίζεται ο συντελεστής προσαρμογής στο κλασσικό μοντέλο.

iii. Για $\delta = 0$, ο συντελεστής προσαρμογής είναι η μοναδική αρνητική ρίζα της εξίσωσης $l(s) = \lambda \hat{f}(s)$.

- **Ασυμπτωτικός τύπος των Cramer-Lundberg**

Ο ασυμπτωτικός τύπος των Cramer-Lundberg, παρέχει μία προσέγγιση για την πιθανότητα χρεοκοπίας στο κλασικό πρότυπο όταν το αρχικό αποθεματικό u παίρνει πολύ μεγάλες τιμές. Σύμφωνα με αυτόν τον τύπο, υπό την προϋπόθεση ότι

$$\int_0^{\infty} x e^{Rx} \bar{F}(x) dx < \infty$$

η πιθανότητα χρεοκοπίας $\psi(u)$ ικανοποιεί την ακόλουθη ασυμπτωτική σχέση

$$\psi(u) \sim C e^{-Ru}$$

καθώς $u \rightarrow \infty$, ή ισοδύναμα

$$\lim_{u \rightarrow \infty} \frac{\psi(u)}{e^{-Ru}} = C$$

όπου C μία θετική σταθερά που δίνεται από τον τύπο

$$C = \frac{\theta \mu}{R \int_0^{\infty} x e^{Rx} \bar{F}(x) dx}.$$

3.2.4 Εφαρμογές

Εφαρμογή 1 - Εύρεση συντελεστή προσαρμογής με κατανομή αποζημιώσεων Erlang

Έστω ότι η $X \sim \text{Erlang}(2, \beta)$, δηλαδή

$$f(x) = \beta^2 x e^{-\beta x}, \quad \beta > 0, \quad x \geq 0.$$

Η ροπογεννήτρια της X είναι

$$M(r) = \frac{\beta^2}{(\beta - r)^2}, \quad r < \beta.$$

Η εξίσωση για το συντελεστή προσαρμογής γίνεται

$$\frac{\beta^2}{(\beta - r)^2} = 1 + (1 + \theta)r\mu$$

ή ισοδύναμα

$$\beta^2 = (\beta - r)^2 [1 + (1 + \theta)r\mu].$$

Αυτό είναι μια εξίσωση τρίτου βαθμού ως προς r , όπου η μία ρίζα είναι το μηδέν (απορρίπτεται) οπότε καταλήγουμε στην παρακάτω εξίσωση δευτέρου βαθμού

$$r^2(1 + \theta)\mu - r(2\beta(1 + \theta)\mu - 1) - (2\beta + (1 + \theta)\mu\beta^2) = 0$$

η οποία έχει δύο ρίζες

$$r_1 = \frac{2\beta\mu(1 + \theta) - 1 - \sqrt{1 + 4\beta\mu(1 + 2\beta\mu + \theta + 4\beta\mu\theta + 2\beta\mu\theta^2)}}{2\mu(1 + \theta)}$$

και

$$r_2 = \frac{2\beta\mu(1 + \theta) - 1 + \sqrt{1 + 4\beta\mu(1 + 2\beta\mu + \theta + 4\beta\mu\theta + 2\beta\mu\theta^2)}}{2\mu(1 + \theta)},$$

όπου η μικρότερη είναι ο συντελεστής προσαρμογής.

Εφαρμογή 2 - Συντελεστής προσαρμογής για σταθερές αποζημιώσεις

Έστω ότι οι αποζημιώσεις παίρνουν μόνο μία τιμή, έστω $X_i = 10$ με πιθανότητα 1. Η εξίσωση για το συντελεστή προσαρμογής

$$M_X(r) = 1 + (1 + \theta)r\mu$$

παίρνει τη μορφή

$$e^{10r} = 1 + 10(1 + \theta)r.$$

Αυτή είναι μία μη γραμμική εξίσωση ως προς r η οποία δε μπορεί να λυθεί αναλυτικά. Παρατηρούμε δηλαδή ότι αν η κατανομή των αποζημιώσεων είναι εκφυλισμένη, ο συντελεστής προσαρμογής δε μπορεί να υπολογιστεί με ακρίβεια.

Ωστόσο $\mu_1 = \mu = 10$ και $\mu_2 = 100$, οπότε το άνω φράγμα για το R είναι

$$R < \frac{2\theta\mu}{\mu_2} \Leftrightarrow R < \frac{20\theta}{100} = \frac{\theta}{5}.$$

Εφαρμογή 3 - Πιθανότητα χρεοκοπίας για μείξη εκθετικών κατανομών

Θέλουμε να βρούμε την πιθανότητα χρεοκοπίας στην περίπτωση που η κατανομή των αποζημιώσεων είναι μια μείξη εκθετικών κατανομών, δηλαδή

$$f(x) = \sum_{i=1}^n \pi_i b_i e^{-b_i x}, \quad x \geq 0,$$

όπου $\pi_i \geq 0$ και $\sum_{i=1}^n \pi_i = 1$. Σε αυτή την περίπτωση η πιθανότητα χρεοκοπίας (Gerber et. al (1987)) δίνεται από τον τύπο

$$\psi(u) = \sum_{k=1}^n C_k e^{-r_k u}, \quad u \geq 0$$

όπου r_1, r_2, \dots, r_n οι ρίζες της εξίσωσης του Lundberg, και

$$C_k = \sum_{j=1}^n \frac{C_{jk}}{b_j}, \quad C_{jk} = \frac{\pi_j / (b_j - r_k)}{\sum_{m=1}^n \frac{\pi_m}{(b_m - r_k)^2}}.$$

Για παράδειγμα έστω ότι η συνάρτηση πυκνότητας της τυχαίας μεταβλητής X είναι η

$$f(x) = \left(\frac{1}{2}\right) 3e^{-3x} + \left(\frac{1}{2}\right) 7e^{-7x}, \quad x \geq 0,$$

δηλαδή $(\pi_1, \pi_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$, και $(b_1, b_2) = (3, 7)$. Η μέση τιμή και η ροπογεννήτρια συνάρτηση των αποζημιώσεων είναι

$$\mu = \left(\frac{1}{2}\right) \left(\frac{1}{3}\right) + \left(\frac{1}{2}\right) \left(\frac{1}{7}\right) = \frac{5}{21}$$

και

$$M_X(r) = \left(\frac{1}{2}\right) \left(\frac{3}{3-r}\right) + \left(\frac{1}{2}\right) \left(\frac{7}{7-r}\right).$$

Συνεπώς από την εξίσωση του Lundberg, $M_X(r) = 1 + (1 + \theta)r\mu$, για περιθώριο ασφαλείας $\theta = 2/5$, προκύπτει ότι

$$1 + \left(1 + \frac{2}{5}\right) \left(\frac{5}{21}\right) r = \left(\frac{1}{2}\right) \left(\frac{3}{3-r}\right) + \left(\frac{1}{2}\right) \left(\frac{7}{7-r}\right)$$

ή ισοδύναμα

$$2r^3 + 14r^2 + 12r = 0.$$

Η μία ρίζα είναι το μηδέν η οποία απορρίπτεται και οι άλλες δύο ρίζες είναι $r_1 = 1$ και $r_2 = 6$.

Για τα C_1, C_2 έχουμε

$$C_1 = \frac{24}{35}, \quad C_2 = \frac{1}{35}.$$

Έτσι η πιθανότητα χρεοκοπίας είναι

$$\psi(u) = \frac{24}{35} e^{-u} + \frac{1}{35} e^{-6u}, \quad u \geq 0.$$

Το συγκεκριμένο παράδειγμα θα επαναδιαπραγματευτεί στην επόμενη παράγραφο κάτω από το πρίσμα των κατανομών τύπου φάσεων.

Εφαρμογή 4 - Εύρεση πλήθους ριζών της θεμελιώδους εξίσωσης του Lundberg

Στόχος της παρούσας εφαρμογής είναι να υπολογίσουμε το πλήθος των λύσεων που έχει η θεμελιώδης εξίσωση του Lundberg για $\delta = 1$, όταν η κατανομή των αποζημιώσεων έχει συνάρτηση πυκνότητας

$$f(x) = \frac{e^{-2x}}{2} + \frac{9e^{-3x}}{4},$$

δοθέντος $\lambda = \frac{3}{2}$ και $c = 1$.

Αρχικά παρατηρούμε πως η πυκνότητα των αποζημιώσεων είναι μία μείξη εκθετικών κατανομών με παραμέτρους $\lambda_1 = 2$, $\lambda_2 = 3$ και αντίστοιχα βάρη $w_1 = 1/4$, $w_2 = 3/4$. Οπότε ο μετασχηματισμός Laplace της f είναι

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} f(x) dx = \frac{1}{2(2+s)} + \frac{9}{4(3+s)}, \quad s > -2$$

ενώ παρατηρούμε ότι για $s \leq -2$ η $\hat{f}(s)$ απειρίζεται.

Λύνοντας το σύστημα των εξισώσεων

$$\begin{cases} cs + \lambda \hat{f}(s) - (\lambda + \delta) = 0 \\ l(s) = \lambda \hat{f}(s) \end{cases}$$

έχουμε ότι

$$l(s) = \delta + \lambda - cs = \frac{5}{2} - s.$$

Άρα

$$\begin{aligned} l(s) = \lambda \hat{f}(s) &\Leftrightarrow \frac{5}{2} - s = \left(\frac{3}{2}\right) \frac{1}{2(2+s)} + \frac{9}{4(3+s)} = \frac{3}{4(2+s)} + \frac{27}{8(3+s)} \\ &\Leftrightarrow 4(5-2s)(2+s)(3+s) = 6(3+s) + 27(3+s) \\ &\Leftrightarrow 8s^3 + 20s^2 - 19s - 48 = 0, \quad s > -2. \end{aligned}$$

Θέτουμε $L(s) = 8s^3 + 20s^2 - 19s - 48$ και εξετάζουμε τη συνάρτηση σε επιμέρους διαστήματα του πεδίου ορισμού της. Στο διάστημα $(-2, 0)$ είναι $L(-2) = 6 > 0$ και $L(0) = -48 < 0$ συνεπώς η $L(s) = 0$ έχει περιττό πλήθος ριζών και αφού η εξίσωση του Lundberg δεν μπορεί να έχει περισσότερες από μία αρνητικές ρίζες, τότε η $L(s) = 0$ έχει ακριβώς μία ρίζα. Ομοίως και στο διάστημα $(0, 2)$ όπου $L(2) = 58 > 0$ υπάρχει ακριβώς μία ρίζα. Στο διάστημα $(2, +\infty)$ είναι $\lim_{s \rightarrow \infty} L(s) = +\infty$, δηλαδή έχουμε άρτιο πλήθος ριζών, όμως σε αυτό το διάστημα δεν υπάρχει ρίζα αφού οι δύο από τις τρεις της εξίσωσης δεν ανήκουν στο $(2, +\infty)$. Η τρίτη ρίζα η οποία όμως δεν γίνεται αποδεκτή λόγω του πεδίου ορισμού της συνάρτησης ανήκει αναγκαστικά στο διάστημα $(-\infty, -2)$. Συμπερασματικά έχουμε σύνολο τρεις ρίζες εκ των οποίων η μία απορρίπτεται.

Για περισσότερες λεπτομέρειες στη θεωρία χρεοκοπίας (ανατρέξτε Πολίτης, Κ. Πανεπιστημιακές Σημειώσεις στο μάθημα Θεωρία Χρεοκοπίας, ΠΜΣ «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου»).

3.3 Κατανομές τύπου φάσεων στη θεωρία χρεοκοπίας

Στην παρούσα παράγραφο θα παρουσιάσουμε δύο αποτελέσματα για τον ακριβή υπολογισμό της πιθανότητας χρεοκοπίας στην περίπτωση που η κατανομή των αποζημιώσεων είναι μια κατανομή τύπου φάσεων και η στοχαστική ανέλιξη του πλεονάσματος δίνεται από τον τύπο

$$U(t) = u + ct - S(t), \quad t \geq 0.$$

Θεώρημα 3.1.

Έστω το κλασικό μοντέλο της θεωρίας κινδύνων με $N(t) \sim P(\lambda t)$ και $X \sim PH_c^{(m)}(\mathbf{a}, \mathbf{T})$. Τότε

$$\psi(u) = \mathbf{a}_+ e^{(\mathbf{T} + t\mathbf{a}_+)u} \mathbf{1}, \quad u \geq 0,$$

όπου

$$\mathbf{a}_+ = -\frac{\lambda}{c} \mathbf{a} \mathbf{T}^{-1}.$$

Στη συνέχεια δίνουμε ένα αποτέλεσμα για το ανανεωτικό πρότυπο της θεωρίας κινδύνων (Sparre-Andersen model). Η βασική διαφορά με το κλασικό μοντέλο της θεωρίας κινδύνων είναι ότι πλέον η ανέλιξη των χρόνων αναμονής δεν είναι πλέον μια ανέλιξη Poisson. Έτσι

υποθέτουμε ότι οι ενδιάμεσοι χρόνοι μεταξύ των αποζημιώσεων είναι απλά ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με κατανομή αυθαίρετη που έχει όμως σύνολο τιμών στο διάστημα $[0, \infty)$.

Θεώρημα 3.2.

Έστω το Sparre-Andersen μοντέλο της θεωρίας κινδύνων όπου η κατανομή των ενδιάμεσων χρόνων αναμονής είναι μια κατανομή τύπου φάσεων $PH_c^{(n)}(\mathbf{v}, \mathbf{S})$ και η κατανομή των αποζημιώσεων είναι και αυτή μια κατανομή τύπου φάσεων $PH_c^{(m)}(\mathbf{a}, \mathbf{T})$. Τότε

$$\psi(u) = \mathbf{a}_+ e^{(\mathbf{T} + \mathbf{t}\mathbf{a}_+)u} \mathbf{1}, \quad u \geq 0,$$

όπου

$$\mathbf{a}_+ = \frac{\mathbf{1}'(\mathbf{Q} - \mathbf{T})}{c\mathbf{1}'\mathbf{t}},$$

ο πίνακας \mathbf{Q} είναι λύση της μη γραμμικής εξίσωσης πινάκων

$$\mathbf{Q} = \mathbf{T} - \mathbf{t}\mathbf{a}[(\mathbf{I} \otimes \mathbf{v})(\mathbf{Q} \oplus \mathbf{S})^{-1}(\mathbf{I} \otimes \mathbf{s})]$$

και $\mathbf{t} = -\mathbf{T}\mathbf{1}$, $\mathbf{s} = -\mathbf{S}\mathbf{1}$.

3.3.1 Phase type renewal theory

Στην συγκεκριμένη παράγραφο θα παρουσιάσουμε ένα αποτέλεσμα για τον υπολογισμό της ανανεωτικής πυκνότητας μίας ανανεωτικής διαδικασίας της οποίας οι χρόνοι ακολουθούν κατανομή τύπου φάσεων.

Έστω ότι η απαριθμήτρια ανέλιξη $\{N(s): s \geq 0\}$ είναι μία ανανεωτική διαδικασία. Ορίζουμε την ανανεωτική συνάρτηση $U(s) = E(N(s))$ η οποία εκφράζει τον αναμενόμενο αριθμό ζημιών στη μονάδα του χρόνου.

Δεδομένου ότι $N(s) \leq n$ αν και μόνο αν $S_n > s$ με S_n τη σύνθετη ανέλιξη για τις συνολικές αποζημιώσεις στο χρονικό διάστημα $[0, s]$, έχουμε το ακόλουθο αποτέλεσμα

$$\begin{aligned} U(s) &= \sum_{n=0}^{\infty} Pr(N(s) > n) \\ &= \sum_{n=0}^{\infty} Pr(S_n \leq s) \end{aligned}$$

$$\begin{aligned}
&= 1 + \sum_{n=1}^{\infty} F^{*n}(s) \\
&= \sum_{n=0}^{\infty} F^{*n}(s),
\end{aligned}$$

όπου F μία συνάρτηση κατανομής στο \mathbb{R}_+ με $F(0) = 0$.

Σύμφωνα με (Bladt & Nielsen (2017) και Bladt (2005)) αποδεικνύεται ότι για κάθε $s > 0$ το $U(s) < \infty$ και $F^{*n}(s) \rightarrow 0$ όσο $n \rightarrow \infty$.

Ορισμός 3.4

Αν η F είναι απολύτως συνεχής τότε η παράγωγος $u(s) = U'(s)$ καλείται η ανανεωτική πυκνότητα της ανανεωτικής διαδικασίας, και ορίζεται ως

$$u(s) = U'(s) = \sum_{n=1}^{\infty} f^{*n}(s).$$

Εδώ αξίζει να σημειωθεί ότι η πυκνότητα ανανέωσης δεν είναι μια πυκνότητα πιθανότητας δεδομένου ότι το ολοκλήρωμά της πάει στο άπειρο.

Ακολουθεί το θεώρημα της ανανεωτικής πυκνότητας της οποίας οι χρόνοι ακολουθούν κατανομή τύπου φάσεων.

Θεώρημα 3.3

Η ανανεωτική πυκνότητα u μίας ανανεωτικής διαδικασίας της οποίας οι χρόνοι είναι $PH_c^{(n)}(\mathbf{a}, \mathbf{T})$ δίνεται από τον τύπο

$$u(s) = \mathbf{a}e^{(\mathbf{T}+\mathbf{t}\mathbf{a})s}\mathbf{t}, \quad s \geq 0.$$

Απόδειξη

Δείτε (Asmussen & Albrecher (2010))

■

3.3.2 Εφαρμογές

Εφαρμογή 1 – Κλασικό μοντέλο με κατανομή αποζημιώσεων μείξη εκθετικών κατανομών

Εδώ θα διαπραγματευτούμε την Εφαρμογή 3 της προηγούμενης παραγράφου υπό το πρίσμα των κατανομών τύπου φάσεων. Έχουμε δηλαδή ότι η συνάρτηση πυκνότητας της τυχαίας μεταβλητής X που περιγράφει το ύψος των αποζημιώσεων είναι η

$$f(x) = \left(\frac{1}{2}\right) 3e^{-3x} + \left(\frac{1}{2}\right) 7e^{-7x}, \quad x \geq 0.$$

Επομένως $X \sim PH_c^{(2)}(\mathbf{a}, \mathbf{T})$ όπου

$$\mathbf{a} = (\pi_1, \pi_2) = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \mathbf{T} = \begin{pmatrix} -3 & 0 \\ 0 & -7 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

Για περιθώριο ασφαλείας $\theta = 2/5$ έχουμε ότι $\lambda/c = 3$, οπότε από το Θεώρημα 3.1 παίρνουμε

$$\mathbf{a}_+ = -\frac{\lambda}{c} \mathbf{a} \mathbf{T}^{-1} = -3 \left(\frac{1}{2}, \frac{1}{2}\right) \begin{pmatrix} -3 & 0 \\ 0 & -7 \end{pmatrix}^{-1} = \left(\frac{1}{2}, \frac{3}{14}\right)$$

και

$$\mathbf{T} + \mathbf{t} \mathbf{a}_+ = \begin{pmatrix} -3 & 0 \\ 0 & -7 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{3}{14} \end{pmatrix} = \begin{pmatrix} -3/2 & 9/14 \\ 7/2 & -11/2 \end{pmatrix}.$$

Επομένως αν θεωρήσουμε την τυχαία μεταβλητή Y με κατανομή $PH_c^{(2)}(\mathbf{a}_+, \mathbf{T} + \mathbf{t} \mathbf{a}_+)$, τότε η πιθανότητα χρεοκοπίας $\psi(u)$ που δίνεται από τον τύπο

$$\psi(u) = 1 - F_Y(u) = \mathbf{a}_+ e^{(\mathbf{T} + \mathbf{t} \mathbf{a}_+) u} \mathbf{1}, \quad u \geq 0.$$

Για τον υπολογισμό της πιθανότητας χρεοκοπίας έχουμε μέσω της R έχουμε δύο επιλογές. Είτε μέσω της συνάρτησης `pphtype` του πακέτου `actuar` που υπολογίζει τη συνάρτηση κατανομής μιας συνεχούς κατανομής τύπου φάσεων, είτε μέσω της συνάρτησης `ruin` του ίδιου πακέτου που υπολογίζει απευθείας την πιθανότητα χρεοκοπίας. Για λεπτομέρειες (δείτε π.χ. Kaas et al. (2008) και Zhang (2006)).

Έτσι, για τον υπολογισμό της πιθανότητας χρεοκοπίας $\psi(u)$, για $u = 0, 1, 2, \dots, 10$, έχουμε τα ακόλουθα αποτελέσματα:

- Συνάρτηση `pphtype`

```
> library(actuar)
> a <- c(1/2, 3/14)
> T <- matrix(c(-3/2, 7/2, 9/14, -11/2), nrow=2)
> u <- 0:10
> (prob.ruin <- 1-pphype(u, a, T))
[1] 7.142857e-01 2.523310e-01 9.280151e-02 3.413970e-02
[5] 1.255930e-02 4.620307e-03 1.699716e-03 6.252905e-04
[9] 2.300315e-04 8.462387e-05 3.113138e-05
```

- Συνάρτηση ruin

```

> psi <- ruin(claims = "exponential",
+ par.claims = list(rate = c(3, 7), w = c(1/2,1/2)),
+ wait = "exponential", par.wait = list(rate = 3), pre = 1)
> (psi(u))
[1] 7.142857e-01 2.523310e-01 9.280151e-02 3.413970e-02
[5] 1.255930e-02 4.620307e-03 1.699716e-03 6.252905e-04
[9] 2.300315e-04 8.462387e-05 3.113138e-05
> plot(psi, xlim=c(0,10))

```

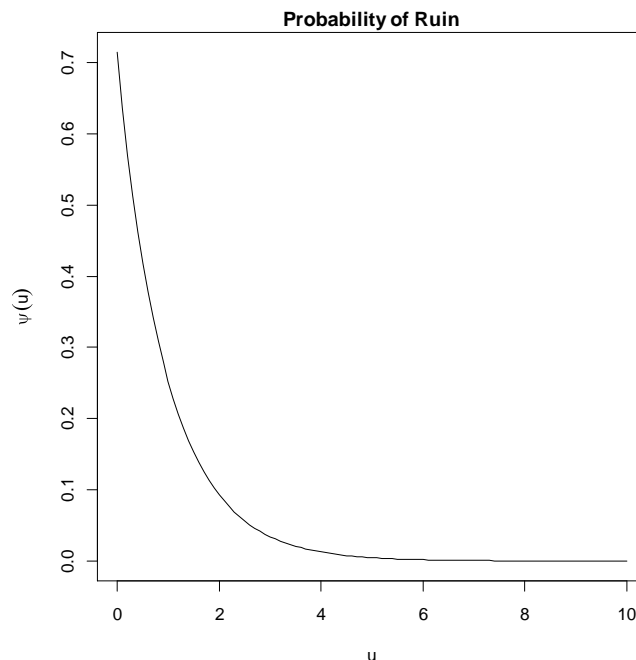
Στην εφαρμογή της συνάρτησης ruin υποθέσαμε ότι $\lambda = 3$, οπότε $c = \lambda\mu(1 + \theta) = 1$ (pre = 1). Σε κάθε περίπτωση μπορεί να διαπιστωθεί ότι

$$\psi(u) = \frac{24}{35}e^{-u} + \frac{1}{35}e^{-6u}, \quad u \geq 0.$$

Αξίζει να σημειώσουμε ότι

$$\psi(0) = \frac{1}{1 + \theta} = \frac{1}{1 + (2/5)} = \frac{5}{7} = 0.71429.$$

Στο ακόλουθο σχήμα δίνεται η γραφική παράσταση της πιθανότητας χρεοκοπίας.



Σχήμα 3.4. Εφαρμογή 1: Πιθανότητα χρεοκοπίας στο κλασικό μοντέλο.

Εφαρμογή 2 – Sparre-Andersen μοντέλο με κατανομή αποζημιώσεων τύπου φάσεων και κατανομή ενδιάμεσων χρόνων μείξη εκθετικών κατανομών

Έστω ότι η κατανομή του ύψους των αποζημιώσεων είναι η μια κατανομή τύπου φάσεων $PH_c^{(2)}(\mathbf{a}, \mathbf{T})$, όπου

$$\mathbf{a} = (0.5614, 0.4386), \quad \mathbf{T} = \begin{pmatrix} -8.640 & 1.997 \\ 0.101 & -1.095 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} 6.643 \\ 0.994 \end{pmatrix}.$$

Θεωρούμε επίσης ότι η κατανομή των ενδιάμεσων χρόνων αναμονής είναι κατανομή τύπου φάσεων $PH_c^{(2)}(\mathbf{v}, \mathbf{S})$, όπου

$$\mathbf{v} = (0.4, 0.6), \quad \mathbf{S} = \begin{pmatrix} -5 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{s} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}.$$

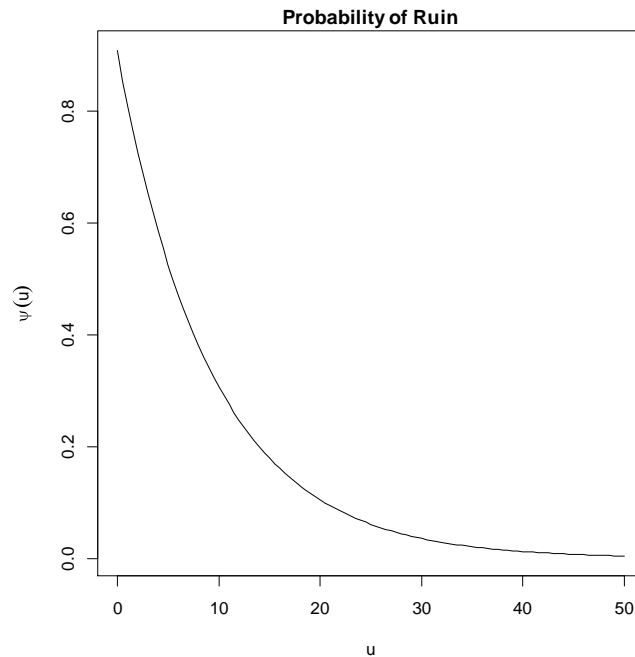
Προφανώς η κατανομή των ενδιάμεσων χρόνων αναμονής είναι μείξη εκθετικών κατανομών με συνάρτηση πυκνότητας

$$g(x) = (0.4)5e^{-5x} + (0.6)e^{-x}, \quad x \geq 0.$$

Για τον υπολογισμό της πιθανότητας χρεοκοπίας $\psi(u)$, για $u = 0, 1, 2, \dots, 10$ και $c = 1$, μέσω της συνάρτησης `ruin` του πακέτου `actuar` έχουμε τα ακόλουθα αποτελέσματα:

```
> library(actuar)
> prob <- c(0.5614, 0.4386)
> rates <- matrix(c(-8.64, 0.101, 1.997, -1.095), 2, 2)
> psi <- ruin(claims = "phase-type",
+ par.claims = list(prob = prob, rates = rates),
+ wait = "exponential", par.wait = list(rate = c(5, 1)),
+ weights = c(0.4, 0.6)), pre=1)
> psi(0:10)
 [1] 0.9081011 0.8082824 0.7257102 0.6515757 0.5850144
 [6] 0.5252526 0.4715957 0.4234201 0.3801659 0.3413303
[11] 0.3064619
> plot(psi, xlim=c(0, 50))
```

Στο ακόλουθο σχήμα δίνεται η γραφική παράσταση της πιθανότητας χρεοκοπίας.



Σχήμα 3.5. Εφαρμογή 2: Πιθανότητα χρεοκοπίας στο μοντέλο Sparre-Andersen.

Εφαρμογή 3 – Πιθανότητα χρεοκοπίας με προσομοίωση

Στην παρούσα εφαρμογή θα παρουσιάσουμε κώδικα στην R για τον υπολογισμό της πιθανότητας χρεοκοπίας μέσω προσομοίωσης. Θεωρούμε το κλασικό μοντέλο όπου οι ενδιάμεσοι χρόνοι αναμονής ακολουθούν εκθετική κατανομή με παράμετρο $\lambda = 1$, και η κατανομή των αποζημιώσεων είναι η κατανομή *Erlang*(2,2). Ακόμη υποθέτουμε ότι το αρχικό αποθεματικό είναι $u = 5$ και το περιθώριο ασφαλείας είναι $\theta = 0.3$. Η ακριβής τιμή της πιθανότητας χρεοκοπίας $\psi(5) = 0.1630571$. Με προσομοίωση βρέθηκε ίση με $\psi(5) = 0.16405$. Ο κώδικας είναι ο ακόλουθος.

```
> ### Crammer Lundberg model
> ### Erlang Distribution for claims
>
> lam <- 1                # N(t)~Poisson(lam*t)
> shape <- 2
> scale <- 0.5            # X~Erlang(shape=2, scale=0.5)
> mean <- shape*scale     # E(X)
> theta <- 0.3           # Περιθώριο Ασφαλείας - Premium Loading Factor
> c <- (1+theta)*lam*mean # Ένταση Ασφαλίστρου - Premium Rate
> u <- 5                  # Αρχικό Αποθεματικό - Initial Surplus
> nSim <- 100000         # Αριθμός Προσομοιώσεων
> n <- 1000              # Αριθμός απαιτήσεων (έναν μεγάλο αριθμό)
#####
> N <- 0
> for (k in 1:nSim){
+   Wi <- rexp(n, rate=lam)
```

```

+   Ti <- cumsum(Wi)
+   Xi <- rgamma(n, shape=2, scale=0.5)
+   Si <- cumsum(Xi)
+   Ui <- u + Ti*c - Si
+   ruin <- !all(Ui>=0)
+   if (ruin) N <- N+1
+ }
> prob.ruin <- N/nSim
> prob.ruin
[1] 0.16405
#####
> library(actuar)
> psi <-ruin(claims = "Erlang",
+ par.claims = list(shape=2, scale=0.5),
+ wait = "exponential", par.wait = list(rate = 1), pre=c)
> psi(u)
[1] 0.1630571

```

Εφαρμογή 4 –Εύρεση της ανανεωτικής πυκνότητας

Έστω ότι έχουμε μία ανανεωτική διαδικασία τύπου φάσης όπου ο πίνακας \mathbf{Q} περιέχει την πινακο-εκθετική μορφή με παράμετρο $\lambda = -q_1 - q_2$ και έχει τύπο $\mathbf{Q} = \begin{pmatrix} -q_1 & q_1 \\ q_2 & -q_2 \end{pmatrix}$. Σκοπός μας είναι να υπολογίσουμε την ανανεωτική πυκνότητα και στη συνέχεια να εισάγουμε δεδομένα έτσι ώστε μέσω του R να εξάγουμε διαφορετικά αποτελέσματα.

Γνωρίζουμε από τη θεωρία ότι $\mathbf{Q} = \mathbf{T} + \mathbf{ta}$. Αναλύοντας τη συγκεκριμένη σχέση έχουμε τα ακόλουθα

$$\mathbf{Q} = \begin{pmatrix} t_{11} + t_1 a_1 & t_{12} + t_1 a_2 \\ t_{12} + t_2 a_1 & t_{22} + t_2 a_2 \end{pmatrix} = \begin{pmatrix} -q_1 & q_1 \\ q_2 & -q_2 \end{pmatrix}.$$

Επιπλέον, σύμφωνα με την Εφαρμογή 1 της παραγράφου 1.2.4 υπολογίζουμε την στάσιμη κατανομή,

$$\boldsymbol{\pi} = (\pi_1 \quad \pi_2) = \left(\frac{q_2}{q_1 + q_2} \quad \frac{q_1}{q_1 + q_2} \right).$$

Οπότε για την συνάρτηση ανανεωτικής πυκνότητας έχουμε:

Για $s \geq 0$,

$$\begin{aligned} u(s) &= \mathbf{a}e^{(\mathbf{T}+\mathbf{ta})s}\mathbf{t} = \mathbf{a}e^{(\mathbf{Q})s}\mathbf{t} \\ &= (a_1 \quad a_2) \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} + e^{\lambda s} (a_1 \quad a_2) \begin{pmatrix} \pi_1 & -\pi_2 \\ -\pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \\ &= (\pi_1 \quad \pi_2) \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} + e^{\lambda s} (a_1 \quad a_2) \begin{pmatrix} \pi_1(t_1 - t_2) \\ \pi_2(t_1 - t_2) \end{pmatrix} \end{aligned}$$

$$= \pi_1 t_1 + \pi_2 t_2 + e^{\lambda s} (a_1 \pi_2 - a_2 \pi_1) (t_1 - t_2).$$

Για τον υπολογισμό των αριθμητικών αποτελεσμάτων, χρησιμοποιήσαμε την κατανομή *Erlang*($\delta = 2,2$) όπου $\mathbf{T} = \begin{pmatrix} -\delta & \delta \\ 0 & -\delta \end{pmatrix}$ και $\mathbf{a} = (1 \ 0)$. Τότε ο \mathbf{t} είναι ίσος με

$$\mathbf{t} = -\mathbf{T}\mathbf{1} = -\begin{pmatrix} -\delta & \delta \\ 0 & -\delta \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \delta \end{pmatrix}$$

Συνεπώς για τον πίνακα \mathbf{Q} έχουμε:

$$\mathbf{Q} = \mathbf{T} + \mathbf{t}\mathbf{a} = \begin{pmatrix} -\delta & \delta \\ 0 & -\delta \end{pmatrix} + \begin{pmatrix} 0 \\ \delta \end{pmatrix} (1 \ 0) = \begin{pmatrix} -\delta & \delta \\ \delta & -\delta \end{pmatrix},$$

και σύμφωνα με την ανάλυση για την ανανεωτική πυκνότητα όπως την παρουσιάσαμε παραπάνω προκύπτει για $\lambda = -2\delta$ ότι

$$\begin{aligned} u(s) &= \mathbf{a}e^{(\mathbf{T}+\mathbf{t}\mathbf{a})s}\mathbf{t} = \mathbf{a}e^{(\mathbf{Q})s}\mathbf{t} \\ &= \frac{\delta}{2}(1 - e^{\lambda s}). \end{aligned}$$

Ακολουθεί το output του προγράμματος στη γλώσσα R καθώς και ο πίνακας με τα αποτελέσματα.

```
library(actuar)
e<-matrix(c(1,1),nrow=2);
T <- cbind(c(-2, 0), c(2, -2));
a <- c(1,0);
t <- -T%*%e;
Q<-T+t%*%a;Q
      [,1] [,2]
[1,]  -2   2
[2,]   2  -2

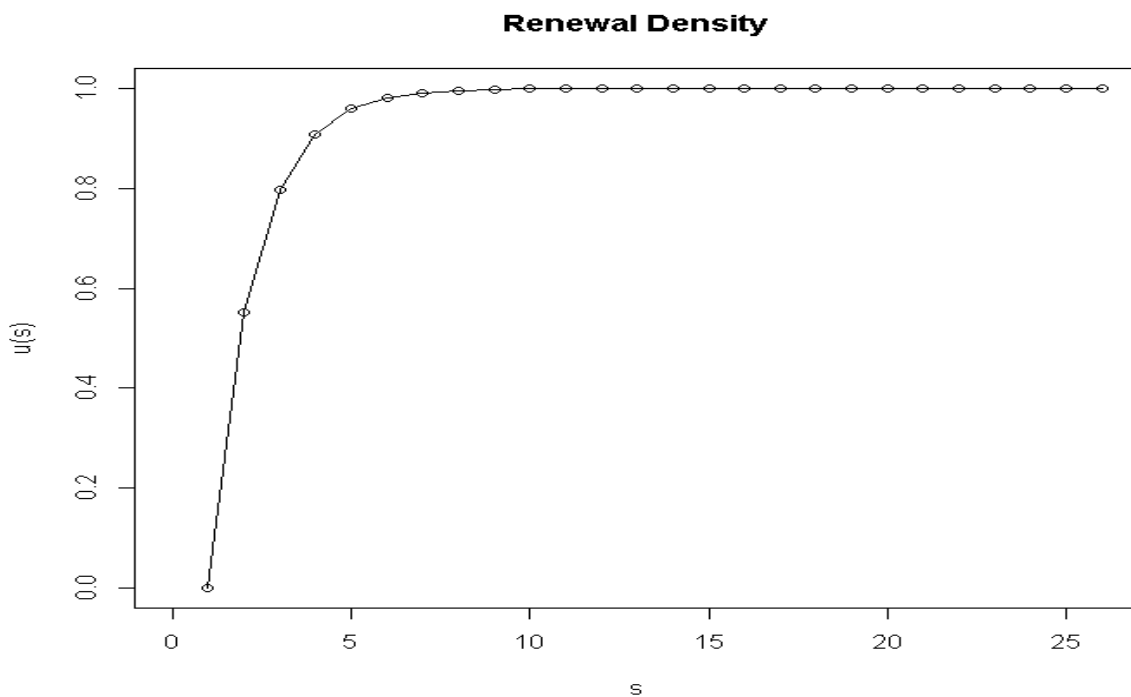
delta<--Q[1];
x <- seq(0,5,0.2)
l<--2*delta
u<-(delta/2)*(1-exp(l*x));u
[1] 0.0000000 0.5506710 0.7981035 0.9092820 0.9592378 0.9816844 0.9917703
0.9963021 0.9983384 0.9992534 0.9996645 0.9998493 0.9999323
[14] 0.9999696 0.9999863 0.9999939 0.9999972 0.9999988 0.9999994
0.9999997 0.9999999 0.9999999 1.0000000 1.0000000 1.0000000 1.0000000

plot(u,xlim=c(0,26),main="Renewal Density",xlab="s",ylab="u(s)")
lines(u)
```

Πίνακας 3.1. Ανανεωτική πυκνότητα $u(s)$ για την κατανομή *Erlang*(2,2)

s	$u(s)$	s	$u(s)$
0	0.0000000	2.6	0.9999696
0.2	0.5506710	2.8	0.9999863
0.4	0.7981035	3	0.9999939
0.6	0.9092820	3.2	0.9999972
0.8	0.9592378	3.4	0.9999988
1	0.9816844	3.6	0.9999994
1.2	0.9917703	3.8	0.9999997
1.4	0.9963021	4	0.9999999
1.6	0.9983384	4.2	0.9999999
1.8	0.9992534	4.4	1.0000000
2	0.9996645	4.6	1.0000000
2.2	0.9998493	4.8	1.0000000
2.4	0.9999323	5	1.0000000

Στο ακόλουθο σχήμα δίνεται η γραφική παράσταση της ανανεωτικής πυκνότητας για την *Erlang*(2,2).



Σχήμα 3.6. Εφαρμογή 4: Διάγραμμα ανανεωτικής πυκνότητας.

3.3.3 Πιθανότητα χρεοκοπίας σε μια διδιάστατη διαδικασία Poisson

Στην παρούσα παράγραφο θα παρουσιάσουμε μια «διδιάστατη» εφαρμογή στη θεωρία χρεοκοπίας (Heilpern 2009)). Θεωρούμε ότι υπάρχουν δύο είδη αποζημιώσεων X_i και Y_i ($i = 1, 2, \dots$) που ανήκουν σε δύο διαφορετικές κλάσεις, και θεωρούμε τη στοχαστική ανέλιξη πλεονάσματος

$$U(t) = u + ct - S(t), \quad t \geq 0$$

όπου

$$S(t) = S_1(t) + S_2(t),$$

$$S_1(t) = \sum_{i=1}^{N_1(t)} X_i,$$

$$S_2(t) = \sum_{i=1}^{N_2(t)} Y_i,$$

$$N_1(t) = M_1(t) + M_0(t),$$

$$N_2(t) = M_2(t) + M_0(t).$$

Θεωρούμε ότι οι στοχαστικές ανελιξίες $M_j(t)$, $j = 0, 1, 2$ είναι ανεξάρτητες ανελιξίες Poisson με εντάσεις λ_i , $i = 0, 1, 2$, αντίστοιχα, οπότε οι στοχαστικές ανελιξίες $N_1(t)$, $N_2(t)$ είναι εξαρτημένες. Στην ειδική περίπτωση $\lambda_0 = 0$, οι στοχαστικές ανελιξίες $N_1(t)$, $N_2(t)$ είναι ανεξάρτητες. Επίσης κάνουμε την υπόθεση ότι οι αποζημιώσεις X_i και Y_i είναι ανεξάρτητες μεταξύ τους αλλά και ανεξάρτητες από τις ανελιξίες $M_j(t)$, $j = 0, 1, 2$.

Ο Ambagaspitya (1998) έδειξε ότι η κατανομή του αθροίσματος $S(t)$ είναι μια σύνθετη κατανομή Poisson $CP(\lambda, H)$, όπου $\lambda = \lambda_0 + \lambda_1 + \lambda_2$ και η συνάρτηση κατανομής H δίνεται από τον τύπο

$$H(x) = \frac{1}{\lambda} (\lambda_1 F_X(x) + \lambda_2 F_Y(x) + \lambda_0 (F_X * F_Y)(x)).$$

Ως εκ τούτου, μπορούμε να εξαγάγουμε την πιθανότητα χρεοκοπίας του μοντέλου μας χρησιμοποιώντας την θεωρία του κλασικού μοντέλου κινδύνου όπου οι αποζημιώσεις Z_1, Z_2, \dots έχουν συνάρτηση κατανομής $H(x)$.

Η αναμενόμενη τιμή των αποζημιώσεων Z_i είναι ίση με

$$\mu_Z = E(Z_i) = \frac{\lambda_1 + \lambda_0}{\lambda} \mu_X + \frac{\lambda_2 + \lambda_0}{\lambda} \mu_Y,$$

η συνάρτηση κατανομής των σταδιακών πτώσεων του πλεονάσματος L_i είναι

$$\begin{aligned} F_{L_i}(y) &= \frac{1}{\mu_Z} \int_0^y (1 - H(x)) dx \\ &= \frac{1}{(\lambda_1 + \lambda_0)\mu_X + (\lambda_2 + \lambda_0)\mu_Y} \int_0^y (\lambda - \lambda_1 F_X(x) - \lambda_2 F_Y(x) - \lambda_0 (F_X * F_Y)(x)) dx, \end{aligned}$$

και η κατανομή του πλήθους W των σταδιακών πτώσεων του πλεονάσματος από την τιμή του αρχικού αποθεματικού u μέχρι τη στιγμή της χρεοκοπίας T έχει τη γεωμετρική κατανομή με πιθανότητα επιτυχίας $\delta(0)$ που δίνεται από τον τύπο

$$\delta(0) = 1 - \psi(0) = \frac{(\lambda_1 + \lambda_0)\mu_X + (\lambda_2 + \lambda_0)\mu_Y}{c}.$$

Για να εξετάσουμε την επίδραση του βαθμού της εξάρτησης των ανελιξων $N_1(t)$ και $N_2(t)$ στην πιθανότητα χρεοκοπίας θεωρούμε τα δύο ακόλουθα μοντέλα:

$$(1) \quad U(t) = u + ct - S(t), \quad t \geq 0$$

$$S(t) = S_1(t) + S_2(t) = \sum_{i=1}^{N_1(t)} X_i + \sum_{i=1}^{N_2(t)} Y_i$$

$$N_1(t) = M_1(t) + M_0(t), \quad N_2(t) = M_2(t) + M_0(t), \quad M_j(t) \sim P(\lambda_j t).$$

$$(2) \quad U'(t) = u + ct - S'(t), \quad t \geq 0$$

$$S'(t) = S'_1(t) + S'_2(t) = \sum_{i=1}^{N'_1(t)} X'_i + \sum_{i=1}^{N'_2(t)} Y'_i$$

$$N'_1(t) = M'_1(t) + M'_0(t), \quad N'_2(t) = M'_2(t) + M'_0(t), \quad M'_j(t) \sim P(\lambda'_j t).$$

Υποθέτουμε ότι

$$\lambda_1 + \lambda_0 = \lambda'_1 + \lambda'_0, \quad \lambda_2 + \lambda_0 = \lambda'_2 + \lambda'_0.$$

Σε αυτή την περίπτωση ο Heilpern (2009) έδειξε ότι αν $\lambda_0 \leq \lambda'_0$, τότε $\psi(u) \leq \psi'(u)$.

Εφαρμογή – Πιθανότητα χρεοκοπίας σε δισδιάστατη διαδικασία Poisson μέσω των κατανομών τύπου φάσεων

Θεωρούμε ότι για τα ύψη των αποζημιώσεων $X \sim \text{Exp}(1/\mu_X)$ και $Y \sim \text{Exp}(1/\mu_Y)$ αντίστοιχα. Η κατανομή της $S(t)$ είναι μια σύνθετη κατανομή Poisson $CP(\lambda, H)$, όπου $\lambda = \lambda_0 + \lambda_1 + \lambda_2$ και η συνάρτηση κατανομής H δίνεται από τον τύπο

$$H(x) = \frac{1}{\lambda} (\lambda_1 F_X(x) + \lambda_2 F_Y(x) + \lambda_0 (F_X * F_Y)(x)).$$

Η παραπάνω συνάρτηση κατανομής αντιστοιχεί σε συνάρτηση κατανομής μείξης τριών κατανομών: Δύο ανεξάρτητων εκθετικών κατανομών που περιγράφονται από τις τυχαίες μεταβλητές X και Y και μιας τρίτης κατανομής που περιγράφεται από την τυχαία μεταβλητή $X + Y$ (υποεκθετική κατανομή). Συνεπώς για την τ.μ. Z με συνάρτηση κατανομής $H(x)$ έχουμε ότι $Z \sim PH_c^{(4)}(\mathbf{a}, \mathbf{T})$ με

$$\mathbf{a} = \begin{pmatrix} \lambda_1/\lambda \\ \lambda_2/\lambda \\ \lambda_0/\lambda \\ 0 \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} -1/\mu_X & 0 & 0 & 0 \\ 0 & -1/\mu_Y & 0 & 0 \\ 0 & 0 & -1/\mu_X & 1/\mu_X \\ 0 & 0 & 0 & -1/\mu_X \end{pmatrix}.$$

Η πιθανότητα χρεοκοπίας προκύπτει από το Θεώρημα 3.1.

Στη συνέχεια θεωρούμε ότι $l_1 = \lambda_0 + \lambda_1 = 1$, $l_2 = \lambda_0 + \lambda_2 = 1$, $\lambda = l_1 + l_2 - \lambda_0 = 2 - \lambda_0$ και ότι η ένταση του ασφαλιστρού είναι ίση με $c = 2.2$. Επίσης θεωρούμε ότι $X \sim \text{Exp}(1)$ και $Y \sim \text{Exp}(1)$, οπότε $F(x) = F_X(x) = F_Y(x) = 1 - e^{-x}$. Θα εξετάσουμε την πιθανότητα χρεοκοπίας στις περιπτώσεις όπου $\lambda_0 = 0, 0.25, 0.5, 0.75$ και 1 .

Καταρχήν έχουμε ότι η διαδικασία $S(t) \sim CP(\lambda, H)$ με $\lambda = l_1 + l_2 - \lambda_0 = 2 - \lambda_0$ και

$$H(x) = \frac{\lambda_1 + \lambda_2}{\lambda} F(x) + \frac{\lambda_0}{\lambda} (F * F)(x) = \frac{2 - 2\lambda_0}{2 - \lambda_0} F(x) + \frac{\lambda_0}{2 - \lambda_0} (F * F)(x).$$

Η παραπάνω συνάρτηση κατανομής αντιστοιχεί σε συνάρτηση κατανομής μείξης δύο κατανομών, της $\text{Exp}(1)$ και της $\text{Erlang}(2,1)$. Συνεπώς η τ.μ. Z με συνάρτηση κατανομής $H(x)$ είναι κατανομή τύπου φάσης $PH_c^{(3)}(\mathbf{a}, \mathbf{T})$, όπου

$$\mathbf{a} = \left(\frac{2 - 2\lambda_0}{2 - \lambda_0}, \frac{\lambda_0}{2 - \lambda_0}, 0 \right)$$

και

$$\mathbf{T} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Η πιθανότητα χρεοκοπίας δίνεται από το Θεώρημα 3.1, δηλαδή

$$\psi(u) = \mathbf{a}_+ e^{(\mathbf{T} + \mathbf{t}\mathbf{a}_+)u} \mathbf{1}, \quad u \geq 0,$$

όπου

$$\mathbf{a}_+ = -\frac{\lambda}{c} \mathbf{a} \mathbf{T}^{-1}.$$

Έχουμε

$$\mathbf{T}^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix},$$

$$\frac{\lambda}{c} = \frac{2 - \lambda_0}{2.2},$$

οπότε

$$\mathbf{a}_+ = \frac{1}{1.1} \left(\frac{2 - 2\lambda_0}{2}, \frac{\lambda_0}{2}, \frac{\lambda_0}{2} \right)$$

Στη συνέχεια παρουσιάζουμε αριθμητικά αποτελέσματα για την πιθανότητα χρεοκοπίας με μεθόδους προσομοίωσης για διαφορετικά λ_0 (ο σχετικός κώδικας δίνεται στο παράρτημα).

Πίνακας 3.2. Πιθανότητα χρεοκοπίας $\psi(u)$

$\lambda_0 = 0$		$\lambda_0 = 0.25$		$\lambda_0 = 0.5$		$\lambda_0 = 0.75$		$\lambda_0 = 1$	
u	$\psi(u)$	u	$\psi(u)$	u	$\psi(u)$	u	$\psi(u)$	u	$\psi(u)$
0	0.9090909	0	0.9090909	0	0.9090909	0	0.9090909	0	0.9090909
1	0.8300916	1	0.8399903	1	0.8512567	1	0.8649204	1	0.8830155
2	0.7579572	2	0.7752117	2	0.7930987	2	0.812877	2	0.8368857
3	0.6920913	3	0.7151268	3	0.73777	3	0.7614495	3	0.7885993
4	0.631949	4	0.6596008	4	0.6859726	4	0.7126363	4	0.7420717
5	0.5770331	5	0.6083543	5	0.6377172	5	0.6667893	5	0.6980568
6	0.5268893	6	0.5610789	6	0.5928292	6	0.6238502	6	0.6565999
7	0.481103	7	0.517474	7	0.5510929	7	0.5836656	7	0.6175931
8	0.4392955	8	0.4772567	8	0.5122927	8	0.5460667	8	0.5809009
9	0.4011211	9	0.4401647	9	0.4762237	9	0.5108892	9	0.5463881
10	0.3662639	10	0.4059554	10	0.4426939	10	0.4779777	10	0.5139256
11	0.3344359	11	0.3744047	11	0.4115248	11	0.4471863	11	0.4833918

Σε κάθε περίπτωση μπορεί να διαπιστωθεί ότι ισχύουν τα κάτωθι αποτελέσματα:

- Για $\lambda_0 = 0$

$$S(t) \sim CP(2, F)$$

$$\psi(u) = 0.9091e^{-u/11},$$

με $\psi(0) = 0.9091000$, $\psi(5) = 0.5770389$ και $\psi(11) = 0.3344392$.

- Για $\lambda_0 = 0.25$

$$S(t) \sim CP\left(1.75, \frac{6}{7}F + \frac{1}{7}F * F\right),$$

$$\psi(u) = 0.9103e^{-0.0809u} + 0.0012e^{-1.1236u}$$

με $\psi(3) = 0.7140953$, $\psi(4) = 0.6586247$ και $\psi(6) = 0.5602436$.

- Για $\lambda_0 = 0.5$

$$S(t) \sim CP\left(1.5, \frac{2}{3}F + \frac{1}{3}F * F\right),$$

$$\psi(u) = 0.9128e^{-0.073u} + 0.0037e^{-1.2452u},$$

με $\psi(1) = 0.8474745$, $\psi(2) = 0.7884965$ και $\psi(7) = 0.5475839$.

- Για $\lambda_0 = 0.75$

$$S(t) \sim CP\left(1.25, \frac{2}{5}F + \frac{3}{5}F * F\right),$$

$$\psi(u) = 0.9159e^{-0.0666u} + 0.0068e^{-1.3652u},$$

με $\psi(8) = 0.5375958$, $\psi(9) = 0.5029582$ και $\psi(10) = 0.4705523$.

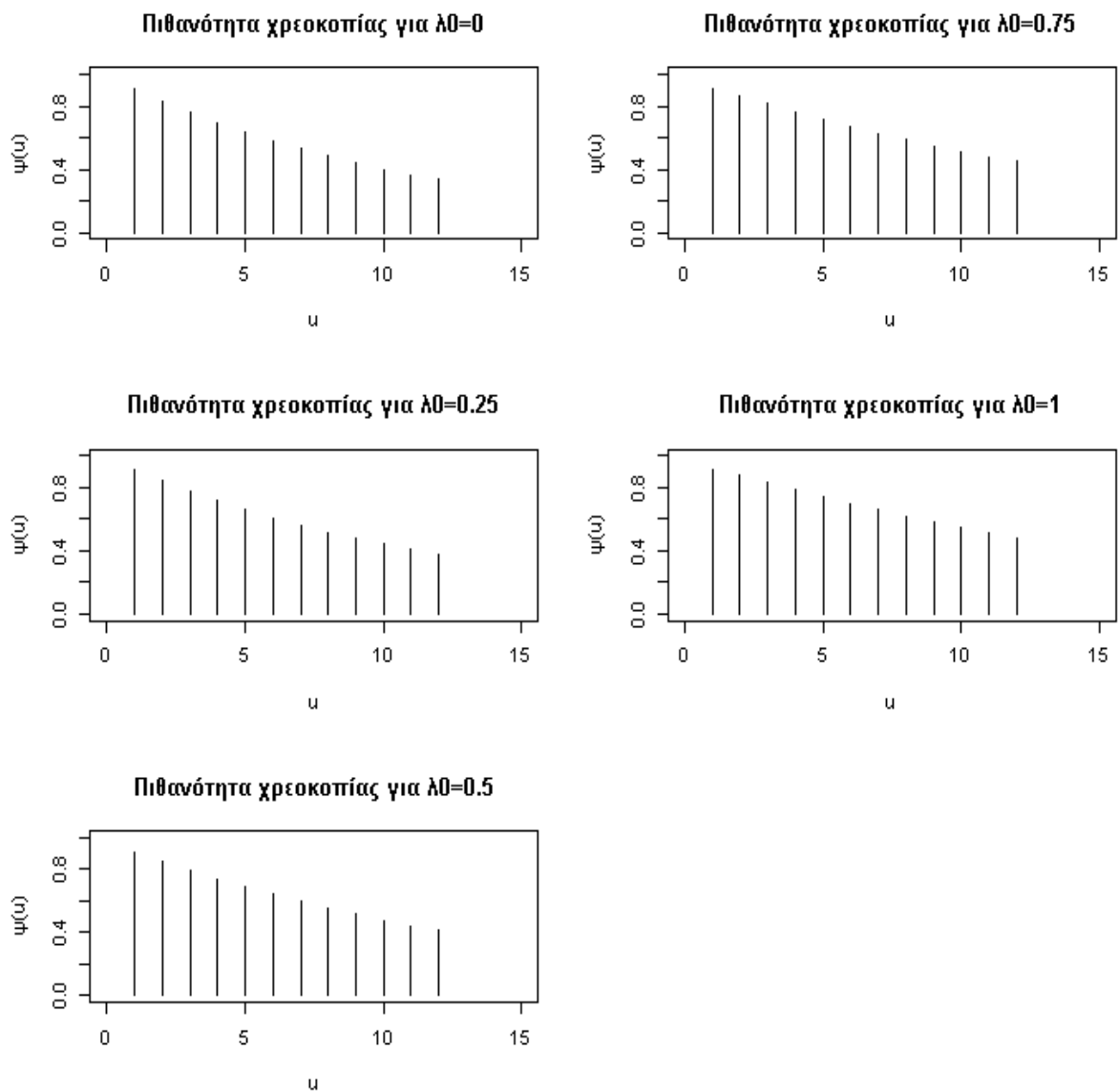
- Για $\lambda_0 = 1$

$$S(t) \sim CP(1, F * F),$$

$$\psi(u) = 0.9192e^{-0.0613u} + 0.01009e^{-1.4842u},$$

με $\psi(0) = 0.9091100$, $\psi(5) = 0.6765422$ και $\psi(11) = 0.4683443$.

Στο ακόλουθο σχήμα δίνεται η γραφική παράσταση της πιθανότητας χρεοκοπίας.



Σχήμα 3.7. Διάγραμμα πιθανότητας χρεοκοπίας για κάθε λ_0 .

ΠΑΡΑΡΤΗΜΑ

Στο παράρτημα παραθέτουμε όλους τους κώδικες σε γλώσσα προγραμματισμού R που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία, εκτός από αυτούς που παρουσιάζονται στο κυρίως κείμενο.

Κώδικας για την Παράγραφο 1.4.3

```
## Geometric distribution of order k ##
#####
# Recurrence Relation 1
# Geometric distribution of order 4
#####
pdf4 <- function(n,p) {
  if (n==0) return(0)
  if (n==1) return(0)
  if (n==2) return(0)
  if (n==3) return(0)
  if (n==4) return(p^4)
  if (n==5) return((1-p)*(p^4))
  if (n==6) return((1-p)*(p^4))
  if (n==7) return((1-p)*(p^4))
  if (n==8) return((1-p)*(p^4))
  return(pdf4(n-1,p) - (1-p)*(p^4)*pdf4(n-5,p))
}
pdfn <- Vectorize(pdf4)
pdfn(1:11,0.8)      #p=0.8
rm(list=ls(all=TRUE))

#####
# Complete Recurrence Relation, General k
#####
pdf <- function(n,p,k) {
  if (n < k-1) return(0)
  if (n == k) return(p^k)
  if (n > k & n < 2*k+1) return((1-p)*(p^k))
  return(pdf(n-1,p,k) - (1-p)*(p^k)*pdf(n-k-1,p,k))
}
pdfN <- Vectorize(pdf)
pdfN(1:11,0.8,4)
rm(list=ls(all=TRUE))

#####
#####
# Phase type
# Pmf Geometric distribution of order 4
#####
library(expm)
a <- c(1,0,0,0)
pdfPhase <- function(x,p) {
  t <- c(0,0,0,p)
  T <- matrix(c(1-p,1-p,1-p,1-p,p,0,0,0,0,p,0,0,0,0,p,0), nrow=4)
```

```

return(a**%(T**%(x-1))**%t)
}
pdfPhaseN <- Vectorize(pdfPhase)
pdfPhaseN(1:15,0.8)
sum(pdfPhaseN(1:15,0.8))

#####
#Graph
max <- 15
k <- 4
x <- seq(k,max,1)
y <- pdfPhaseN(x,0.8)
y
plot(x,y,type="h", xlim=c(0,max),lab=c(16, 5, 7), lwd=3, ylab="f(x)")
title(expression(Συνάρτηση~πιθανότητας~της~κατανομής~G[4](0.8)))

#####
#Factorial moments
#####
library(expm)
a <- c(1,0,0,0)
e <- c(1,1,1,1)

momentsPhase <- function(r,p) {
  T <- matrix(c(1-p,1-p,1-p,1-p,p,0,0,0,0,p,0,0,0,0,p,0), nrow=4)
  return(factorial(r)*a**%(T**%(r-1))**%(solve(diag(4)-T))**%r)**%e)
}
momentsPhaseN <- Vectorize(momentsPhase)
m <- momentsPhaseN(1:4,0.8)
m
#####
m[1] # Mean
### Verification
p <- 0.8
q <- 1-p
k <-4
(1-p^k)/(q*p^k) # mean
#####
m[2]+m[1]-(m[1])^2 #variance
### Verification
n1 <- 1-(2*k+1)*(q*p^k)-p^(2*k+1)
d1 <- (q*p^k)^2
n1/d1 # variance

```

Κώδικας για την Παράγραφο 1.4.6

```

##### Asmussen 2000, page 220

pi <- c(1/2,1/2) # το α
#t <- c(6/7, 2)
T <- matrix(c(-3/2, 7/2, 9/14, -11/2), nrow=2);T
s <- seq(0.1,6,0.1)
library(actuar)
dens1 <-dphtype(s,pi,T) # density
dens1
dens2 <- (32/35)*exp(-s)+(18/35)*exp(-6*s)
dens2

```

```

curve(dphtype(x, pi, T), from = 0.01, to = 6)
m <- mphtype(1, pi, T) # expected value
m
# x <- seq(0.01,10,0.01)
# lines(x,dexp(x,rate=1/m), col="red")
# points(s,dens1,cex=1)
#####
#####
Παραλλαγή με pi <- c(1/3,2/3)

pi <- c(1/3,2/3)
#t <- c(6/7, 2)
T <- matrix(c(-3/2, 7/2, 9/14, -11/2), nrow=2);T
s <- seq(0.1,6,0.1)
library(actuar)
dens1 <-dphtype(s,pi,T) # density
dens1
dens2 <- (92/105)*exp(-s)+(26/35)*exp(-6*s)
dens2

par(mai=c(1, 1, 0.3, 0.3))
s <- seq(0.01,6,0.01)
plot(s,dphtype(s, pi, T), type="l", xlab="y",ylab="f(y)", main="Συνάρτηση
πυκνότητας της Y")
curve(dphtype(x, pi, T), from = 0.01, to = 6)

m <- mphtype(1, pi, T) # expected value

#####
## ΡΟΠΕΣ
#####

library(expm)
pi <- c(1/3,2/3)
T <- matrix(c(-3/2, 7/2, 9/14, -11/2), nrow=2);T
e <- c(1,1)

momentsPhase <- function(r) {
  return(((1)^r)*factorial(r)*pi%%(solve(T))^r))%%e)
}
momentsPhaseN <- Vectorize(momentsPhase)
m <- momentsPhaseN(1:6)
round(m,digits=4)
mphtype(1, pi, T) # expected value
mphtype(2, pi, T)
mphtype(3, pi, T)
mphtype(4, pi, T)

#####
#####

## Erlang(3, 2) distribution
T <- cbind(c(-2, 0, 0), c(2, -2, 0), c(0, 2, -2))
pi <- c(1,0,0)
x <- 0:10
dphtype(x, pi, T) # density
dgamma(x, 3, 2) # same
pphtype(x, pi, T) # cdf
pgamma(x, 3, 2) # same

```

```

rphtype(10, pi, T) # random values
mphtype(1, pi, T) # expected value
curve(mgfphtype(x, pi, T), from = -10, to = 1)

```

Κώδικας για την Εφαρμογή 2 του 2^{ου} Κεφαλαίου

```

#####
# Random variable X - Negative binomial, r=5, p=0.25, distribution
R={0,1,2,...}
# a, T, t
# Θεώρημα 2.8
#####
p <- 0.25
a1 <- (1-p)^5
a2 <- 5*p*(1-p)^4
a3 <- 10*p^2*(1-p)^3
a4 <- 10*p^3*(1-p)^2
a5 <- 5*p^4*(1-p)^1
a6 <- p^5

a <- c(a1,a2,a3,a4,a5)

t <- c(0,0,0,0,p)

v1 <- c(1-p,0,0,0,0)
v2 <- c(p,1-p,0,0,0)
v3 <- c(0,p,1-p,0,0)
v4 <- c(0,0,p,1-p,0)
v5 <- c(0,0,0,p,1-p)

T <- matrix(c(v1,v2,v3,v4,v5), nrow=5)

#####
# Random variable N - General phase type distribution
# β=b, B, b=bt
# Θεώρημα 2.8
#####
b1 <- 0.1
b2 <- 0.2
b3 <- 0.5
b4 <- 0.05
b5 <- 0.15
b6 <- 0

b <- c(b1,b2,b3,b4,b5)

w1 <- c(0.2,0,0,0,0)
w2 <- c(0.4,0.3,0,0,0)
w3 <- c(0,0.7,0.4,0,0)
w4 <- c(0.4,0,0,0.2,0)
w5 <- c(0,0,0,0.8,0.5)

B <- matrix(c(w1,w2,w3,w4,w5), nrow=5)

```

```

bt <- c(0,0,0.6,0,0.5)

#####
#### Random Sum SN
#### Initial vector  $\gamma=g$  (1 x mn)
#####
cM1 <- matrix(b%*(solve(diag(5)-a6*B)),ncol=1)
ca <- matrix(a,ncol=1)
g <- ca%x%cM1
round(g, digits=5)

#####
#### Random Sum SN
#### Matrix C (mn x mn)
#####

M1 <- T%x%diag(5)
ct <- matrix(t,ncol=1)
ra <- matrix(a,ncol=5)
M2 <- ct%x%ra
M3 <- (solve(diag(5)-a6*B))%*%B
M4 <- M2%x%M3
C <- M1+M4
round(C, digits=5)

#####
#### Random Sum
#### ARRAY c
#####

v <- rep(1, times=25)
c <- v-C%x%v
sum(c)

#####
PMF of random sum
Table 3 in WU LI 2010
#####
rg <- as.vector(g)

library(expm)
pdfPhase <- function(x) {
  return(rg%*(C%^(x-1))%*%c)
}

pdfSN <- Vectorize(pdfPhase)

#####
P(SN=0)=PN(f(0))
#####
f0 <- dnbinom(0,5,p)
g0 <- b6+b%*(solve((1/f0)*diag(5)-B))%*%bt
g0
g0 <- b6+f0*b%*(solve(diag(5)-f0*B))%*%bt
g0

```

```

u1 <- pdfSN(1:150)
u <- c(g0,u1)
sum(u)
m <-cbind(0:150,u)
rownames(m) <-rep("",nrow(m))
colnames(m) <-c("y"," g(y)=P(Y=y)")

round(m, digits=10)
m

#####
#Graph
par(mai=c(1, 1, 0.3, 0.3))
x <- 0:150
y <- u
plot(x,y,type="h", ylab="g(x)")

```

Κώδικας για την Εφαρμογή 3 του 2^{ου} Κεφαλαίου

```

#####
# Random variable X - Erlang(2, 3.5)}
# a, T, t
# Θεώρημα 2.8
# m=2
#####
L=3.5
a <- c(1,0)
# a3=0
t <- c(0,L)
T <- cbind(c(-L, 0), c(L, -L))
#a;T;t
library(actuar)
x <- 1:10
dphtype(x, a, T) # density
dgamma(x, 2, 3.5) # same
#####
# Random variable N - Geometric Shifted
# β=b, B, b=bt
# Θεώρημα 2.7
# n=1
#####
p <- 0.25
b <- 1-p
B <- 1-p
bt <- p

#####
#### Random Sum SN
#### Initial vector γ=g (1 x mn) = (1x2)
#####

g <- a*b
g

#####
#### Random Sum SN

```

```

#### Matrix C (mn x mn)
#####

ct <- matrix(t,ncol=1)
ct
dim(ct)
ra <- matrix(a,ncol=2)
ra
dim(ra)
M2 <- ct%*%ra
M2

M4 <- M2%x%B
M4

C <- T+M4
C
round(C, digits=5)

#####
x <- seq(0, 15, 0.5)
dphtype(x, g, C)
par(mai=c(1, 1, 0.3, 0.3))
x <- seq(0.001, 30, 0.001)
plot(x,dphtype(x, g, C), type="l", ylab="g(x)")

#####
#Confirmation Compound Geometric - Erlang
#####
gg <- function(x){
  h <- (1/2)*p*L*(sqrt(1-p))*(exp(-L*x))*(exp(L*x*sqrt(1-p))-exp(-
L*x*sqrt(1-p)))
  return(h)
}
x <- seq(0.5, 15, 0.5)
gg(x)
#####

```

Κώδικας για την Εφαρμογή 4 του 2^{ου} Κεφαλαίου

```

#####
#####
library(actuar)
a <- 0; b <- 20; h <- 0.1
pi <- c(1,0,0)
t <- c(0,0,2)
T <- matrix(c(-2,0,0,2,-2,0,0,2,-2), nrow=3);T
fx <- discretize(pphtype(x,pi,T),method = "rounding", from = a, to = b,
step = h)

#### RECURSIVE METHOD
lam <- 10

```



```

Gs <- aggregateDist("recursive", model.freq = "poisson", model.sev = fx,
lambda = lam, x.scale = h)
gs <- diff(Gs)

c <- 30
x <- seq(0,c,h)
d <- (c/h)+1
mat <- cbind(x, gs[1:d],Gs(x))
rownames(mat) <- rep("", nrow(mat))
colnames(mat) <- c("x","g(x)","G(x)")
round(mat,digits=10)

plot(Gs, do.points=FALSE, verticals=TRUE)
plot(knots(Gs), gs, type="h")

#####
###
### Simulation Confirmation P(S<=10)

# n τυχαίες τιμές από την SN
n<-100000
S<-rep(0,n)
N<-rpois(n,10)
for(i in 1:n){
  V<-rphtype(N[i],pi,T)
  S[i]<-sum(V)
}

# Συνάρτηση κατανομής
F <- rep(0,d)
for(i in 1:d){
  m <- S[S<=x[i]]
  F[i] <- length(m)/n
}

mat <- cbind(x,Gs(x),F[1:d])
rownames(mat) <- rep("", nrow(mat))
colnames(mat) <- c("x"," G(x)- recursive"," G(x)-Simulation")
round(mat,digits=5)

```

Κώδικας για την Παράγραφο 3.3.4

```

library(actuar)
l1<-1;l2<-1; # l1 και l2 προσδιορίζουν τις εντάσεις των τ.μ. X και Y
c<-2.2; #H ένταση του ασφαλιστρου

mx<-1;my<-1; #Οι μέσες τιμές για τις τ.μ. X~Exp(1) και Y~Exp(1)
lamda_0<-0 #To λ0
for(lamda_0 in seq(0,1,0.25)){
  e<-matrix(c(1,1,1),nrow=3); #To μοναδιαίο κάθετο διάνυσμα τριών
στοιχείων

  p<-(l1*mx+l2*my)/c; #δ(0)=1-ψ(0)=p το οποίο προσδιορίζει την κατανομή
του W που είναι γεωμετρική με πιθανότητα επιτυχίας δ(0)

  mz<-(l1*mx+l2*my)/(l1+l2-lamda_0);mz #H αναμενόμενη τιμή της τυχαίας
μεταβλητής Z

```

```

#H τυχαία μεταβλητή Z έχει κατανομή τύπου φάσης με στοιχεία PH(a,T)
a<-c(((2-2*lamda_0)/(2-lamda_0)),(lamda_0/(2-lamda_0)),0);

T<-matrix(c(-1,0,0,0,-1,0,0,1,-1),nrow=3);

a_s<--(a%*%solve(T)/mz); #Το διάνυσμα a^s
t<--T%*%e;t #Το διάνυσμα t

#Ονομάζω Q το εσωτερικό του εκθετικού
Q<-T+p*t%*%a_s

#Για την πιθανότητα χρεοκοπίας έχουμε
for (k in seq(0,11,3)){
  u<-c(k,k+1,k+2);
  prob.ruin<-p*(1-pphype(u,a,Q))
  cat("Για λ0 =",lamda_0,"και u =",u,"το ψ(u) =",prob.ruin,"\n")
  cat("\n")
}
}

#PLOT
d1<-c(0.9090909,0.8300916,0.7579572);d2<-
c(0.6920913,0.631949,0.5770331);d3<-c(0.5268893,0.481103,0.4392955);d4<-
c(0.4011211,0.3662639,0.3344359);
d<-c(d1,d2,d3,d4);
r1<-c(0.9090909,0.8399903,0.7752117);r2<-
c(0.7151268,0.6596008,0.6083543);r3<-c(0.5610789,0.517474,0.4772567);r4<-
c(0.4401647,0.4059554,0.3744047);
r<-c(r1,r2,r3,r4);
n1<-c(0.9090909,0.8512567,0.7930987);n2<-
c(0.73777,0.6859726,0.6377172);n3<-c(0.5928292,0.5510929,0.5122927);n4<-
c(0.4762237,0.4426939,0.4115248);
n<-c(n1,n2,n3,n4);
f1<-c(0.9090909,0.8649204,0.812877);f2<-
c(0.7614495,0.7126363,0.6667893);f3<-
c(0.6238502,0.5836656,0.5460667);f4<-c(0.5108892,0.4779777,0.4471863);
f<-c(f1,f2,f3,f4);
g1<-c(0.9090909,0.8830155,0.8368857);g2<-
c(0.7885993,0.7420717,0.6980568);g3<-
c(0.6565999,0.6175931,0.5809009);g4<-c(0.5463881,0.5139256,0.4833918);
g<-c(g1,g2,g3,g4);

par(mfcol=c(3,2))
plot(d,xlim=c(0,15),ylim=c(0,1),xlab="u",ylab="ψ(u)",type="h",col="black"
,main="Πιθανότητα χρεοκοπίας για λ0=0")
plot(r,xlim=c(0,15),ylim=c(0,1),xlab="u",ylab="ψ(u)",type="h",col="black"
,main="Πιθανότητα χρεοκοπίας για λ0=0.25")
plot(n,xlim=c(0,15),ylim=c(0,1),xlab="u",ylab="ψ(u)",type="h",col="black"
,main="Πιθανότητα χρεοκοπίας για λ0=0.5")
plot(f,xlim=c(0,15),ylim=c(0,1),xlab="u",ylab="ψ(u)",type="h",col="black"
,main="Πιθανότητα χρεοκοπίας για λ0=0.75")
plot(g,xlim=c(0,15),ylim=c(0,1),xlab="u",ylab="ψ(u)",type="h",col="black"
,main="Πιθανότητα χρεοκοπίας για λ0=1")

```


ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

1. Πολίτης, Κ. Πανεπιστημιακές Σημειώσεις στο μάθημα Θεωρία Χρεοκοπίας, ΠΜΣ «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου».
2. Χατζηκωνσταντινίδης, Ε. Πανεπιστημιακές Σημειώσεις στη Θεωρία Κινδύνων Ι, ΠΜΣ «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου».

Ξένα

1. Asmussen, S. & Albrecher, H. (2010). *Ruin Probabilities*, 3rd ed., World Scientific, London.
2. Balakrishnan, N. & Koutras, M. V. (2002). *Runs and Scans with Applications*, 1st ed., John Wiley & Sons, New York.
3. Baum, D. & Breuer, L. (2005). *An Introduction to Queueing Theory : And Matrix-Analytic Methods*, 1st ed., Springer, New York
4. Bladt, M. (2005). A review on phase-type distributions and their use in risk theory, *ASTIN Bulletin*, **35**, 145–167.
5. Bladt, M. & Nielsen. B. F. (2017). *Matrix-Exponential Distributions in Applied Probability*, Volume 81, Springer, New York
6. Buchholz, P., Kriege, J. & Felko, I. (2014). *Input Modeling with Phase-Type Distributions and Markov Models*, 1st ed., Springer, New York
7. Dutang, C., Goulet, V. & Pigeon, M. (2008). An R Package for Actuarial Science, *Journal of Statistical Software*, **25**, 1-37
8. Eisele, K. T. (2006). Recursions for compound phase distributions, *Mathematics and Economics*, **38**, 149–156
9. Grandell, J. (1991). *Aspects of Risk Theory*, 1st ed, Springer-Verlag, New York
10. Heilpern, S. (2009). Probability of ruin for a dependent, two-dimensional Poisson process, *Operations Research and Decisions*, **1**, 77-90.
11. Hipp, C. (2006). Speedy convolution algorithms and Panjer recursions for phase-type distribution, *Mathematics and Economics*, **38**, 176–188
12. Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M. (2008). *Modern Actuarial Risk Theory Using R*, 2st ed., Springer, New York
13. Karlin, S. & Taylor, H. M. (1998). *An Introduction to Stochastic Modeling*, 3rd ed., Academic Press, California.
14. Kemeny, J. G., Snell J. L. & Knapp, A. W. (1976). *Denumerable Markov Chains*, 1st ed., Springer-Verlag, New York

15. Koutras, M. V. & Eryilmaz, S. (2017). Compound Geometric Distribution of Order k , *Methodol Comput Appl Probab* , **19**, 377–393.
16. Latouche, G. & Ramaswami, V. (1987). *Introduction to matrix analytic methods in stochastic modeling*, 1st ed., ASA-SIAM Series on Statistics and Applied Probability, Virginia.
17. Norris, J. R. (1997). *Markov Chains*, 1st ed., Cambridge University Press, Cambridge.
18. Philippou, A. N. (1983). The Poisson and compound Poisson distributions of order k and some of their properties, *Journal of Soviet Mathematics*, **27**, 3294-3297.
19. Philippou, A. N. & Muwafi, A. A. (1982). Waiting for the K -th consecutive success and the Fibonacci sequence of order K , *Fibonacci Quarterly*, **20**, 43-54.
20. Resnick, S. I. (1992). *Adventures in Stochastic Processes*, 1st ed., Birkhäuser, Basel.
21. Zhang, X. (2006). On the ruin problem in a Markov-modulated risk model, *Methodology and Computing in Applied Probability*, **20**, 225-238.