ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

ΚΑΤΑΝΟΜΕΣ ΤΥΧΑΙΩΝ ΑΘΡΟΙΣΜΑΤΩΝ ΣΤΟΝ ΑΝΑΛΟΓΙΣΜΟ ΚΑΙ ΤΗ ΔΙΟΙΚΗΤΙΚΗ ΚΥΝΔΙΝΟΥ

Δήμητρα Ν. Σπανού

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

Ιούλιος 2017

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

Τα μέλη της Επιτροπής ήταν:

- Καθηγητής Μ. Κούτρας (Επιβλέπων)
- Αναπληρωτής Καθηγητής Κ. Πολίτης
- Αναπληρωτής Καθηγητής Ε. Χατζηκωνσταντινίδης

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

UNIVERSITY OF PIRAEUS

DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

MASTER PROGRAM IN ACTUARIAL SCIENCE AND RISK MANAGEMENT

DISTRIBUTIONS OF RANDOM SUMS IN ACTUARIAL & RISK MANAGEMENT

Dimitra N. Spanou

Thesis

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

Piraeus, Greece

July 2017

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

Τα μέλη της Επιτροπής ήσαν:

- Καθηγητής Μ. Κούτρας (Επιβλέπων)
- Αναπληρωτής Καθηγητής Κ. Πολίτης
- Αναπληρωτής Καθηγητής Ε. Χατζηκωνσταντινίδης

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

Στον πατέρα μου

Νίκο

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. Μάρκο Κούτρα για την καθοδήγησή του καθ' όλη τη διάρκεια εκπόνησης της εργασίας αυτής καθώς και την καθοριστική συμβολή του στην ολοκλήρωσή της.

Επίσης, τις ευχαριστίες μου θα ήθελα να εκφράσω και στους καθηγητές κ. Ευστάθιο Χατζηκωνσταντινίδη και κ. Κωνσταντίνο Πολίτη που δέχτηκαν να είναι μέλη της τριμελούς επιτροπής αξιολόγησης της μεταπτυχιακής μου εργασίας.

Τέλος, δεν θα μπορούσα να παραλείψω τις ευχαριστίες μου, αλλά και την αγάπη μου, προς τους γονείς μου και τον ξεχωριστώ Γιώργο Κύρου, για την υποστήριξη και την βοήθειά τους.

Περίληψη

Σκοπός της παρούσας διπλωματικής εργασίας είναι η παρουσίαση και ανάλυση δύο μοντέλων πιθανοτήτων συνδεδεμένων με τε ροές επιτυχιών και συναρτήσεις σάρωσης. Η εφαρμογή των δύο αυτών μοντέλων γίνεται χρησιμοποιώντας ένα τυχαίο άθροισμα τυχαίων μεταβλητών που μετράει τον συνολικό χρόνο αναμονής μέχρι να εμφανιστεί μια επικίνδυνη κατάσταση χρεοκοπίας σε μία τράπεζα ή μία ασφαλιστική.

Η μοντελοποίηση γίνεται βάσει των διακριτών κατανομών Phase-type της γεωμετρικής κατανομής τάξης k. Υποθέτουμε ότι η απαριθμήτρια τυχαία μεταβλητή ενός τυχαίου αθροίσματος περιγράφει τον χρόνο αναμονής πάνω σε μία ακολουθία διωνυμικών μεταβλητών, μέχρι εμφάνιση για πρώτη φορά ένα προκαθορισμένο κριτήριο τερματισμού.

Γίνεται παρουσίαση των ρυθμιστικών πλαισίων που ισχύουν σήμερα και καθιστούν την διαχείριση στο εσωτερικό των επιχειρήσεων πιο επίκαιρη από ποτέ.

Abstract

In the present master thesis we study and analyze two probability models associated with run statistics and scan statistics. The implementation of those model is being done by using a random sum of random variables that counts the waiting time until a hazard situation occurs.

Modeling is based on the distinct Phase-type distributions and a geometric distribution of order k. We present some results on the distribution of a compound random variable and how these results can be used in a bank or in an insurance company.

We also give the scope of regulatory frameworks and how these regulations make risk management more opportune than ever before.

ΠΕΡΙΕΧΟΜΕΝΑ

Κατάλογος Πινάκων

Κατάλογος Σχημάτων

Κατάλογος Συντομογραφιών

τ.μ.	τυχαία μεταβλητή
τ.α	τυχαίο άθροισμα
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
σ.π	συνάρτηση πιθανότητας
σ.κ	συνάρτηση κατανομής
σ.δ	στοχαστική διαδικασία
M.I	μαρκοβιανή ιδιότητα
PH	Phase type κατανομή
DPH	Διακριτή phase type
NB	Αρνητική διωνυμική κατανομή
G	Γεωμετρική κατανομή

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Στο κεφάλαιο αυτό, αναφέρονται συνοπτικά βασικές έννοιες από τη θεωρία Πιθανοτήτων, χρήσιμες στην ερμηνεία και κατανόηση προβλημάτων της θεωρίας κινδύνου. Γίνεται εισαγωγή στην έννοια των στοχαστικών διαδικασιών και τέλος παρατίθενται, χωρίς την απόδειξή τους, κάποια γνωστά αποτελέσματα που αφορούν άθροισμα τυχαίου αριθμού τυχαίων μεταβλητών.

1.1 Κατανομές Αθροισμάτων

Ας υποθέσουμε δύο πραγματικές συναρτήσεις f, g . Η συνέλιξη των δύο θα αποτελεί μία νέα συνάρτηση *h,* η οποία ορίζεται ως εξής:

- Για συνεχείς συναρτήσεις *f, g* στο διάστημα $[0,+\infty)$, η σ.π.π. δίνεται από τον τύπο :

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

- Για διακριτές συναρτήσεις *f, g* που ορίζονται στο $(0,1,2,\ldots)$, η σ.π. δίνεται από τον τύπο :

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

Η συνάρτηση $h(x)$ ονομάζεται συνέλιξη των συναρτήσεων *f* και *g* και συμβολίζεται με $h=f*g$.

Ας θεωρήσουμε πως έχουμε δύο ανεξάρτητες και μη-αρνητικές διακριτές τυχαίες μεταβλητές X_1, X_2 . Το άθροισμά τους δημιουργεί μία νέα τυχαία μεταβλητή X_1+X_2 η κατανομή της οποίας καλείται συνέλιξη (convolution) των κατανομών των δύο τυχαίων μεταβλητών. Κάνοντας χρήση του Θεωρήματος Ολικής Πιθανότητας προκύπτει μία σχέση για την συνάρτηση πιθανότητας για την διακριτή συνέλιξη.

Συμβολίζοντας με f_{X_1}, f_{X_2} τις συναρτήσεις πιθανότητας των τ.μ X_1, X_2 αντίστοιχα, η προκύπτουσα σχέση είναι η σ.π. $f(x)$ για το άθροισμα X_1+X_2 και θα δίνεται από τον τύπο:

$$\begin{aligned} f(x)&=P(X_1+X_2=x)\sum_{y=0}^{x} P(X_1+X_2=x|X_2=y)P(X_2=y) \\ &=\sum_{y=0}^{x} P(X_1=x-y)P(X_2=y)=\sum_{y=0}^{x} f_{X_1}(x-y)f_{X_2}(y). \end{aligned} \tag{1.1}$$

Αντίστοιχα, η σ.κ. της X_1+X_2 θα δίνεται από τον τύπο:

$$\begin{aligned} F(x)&=P(X_1+X_2\leq x)=\sum_{y=0}^{x} P(X_1+X_2\leq x|X_2=y)P(X_2=y) \\ &=\sum_{y=0}^{x} P(X_1\leq x-y)P(X_2=y)P(X_2=y)=\sum_{y=0}^{x} F_{X_1}(x-y)f_{X_2}(y). \end{aligned} \tag{1.2}$$

Αντίστοιχα, τύποι υπάρχουν και στην περίπτωση που οι ανεξάρτητες τ.μ X_1, X_2 είναι συνεχείς τ.μ.

Παράδειγμα 1.1: Αν θεωρήσουμε δύο συναρτήσεις $f(x)=\frac{1}{2^x}$, $g(x)=\frac{1}{3^x}$ με τα x να παίρνουν τιμές στο σύνολο $x\in\{1,2,3,\ldots\}$. Η συνέλιξη των παραπάνω συναρτήσεων *f, g* για τις διάφορες τιμές του x θα είναι:

Για $x=0$, $(f*g)(0)=f(0)g(0)=0$

Για $x=1$, $(f*g)(1)=f(0)g(1)+f(1)g(0)=0$

Για $x=2,3,4,\ldots$,

$$(f*g)(x)=\sum_{y=0}^{x} f(y)g(x-y)=f(0)g(x)+\sum_{y=1}^{x-1} f(y)g(x-y)+f(x)g(0)$$
$$=\sum_{y=1}^{x-1} f(y)g(x-y)=\sum_{y=1}^{x-1}\frac{1}{2^y}\frac{1}{3^{x-y}}=\frac{1}{3^x}\sum_{y=1}^{x-1}\left(\frac{3}{2}\right)^y$$
$$=\frac{1}{3^x}\sum_{y=1}^{x-1}\left(\frac{3}{2}\right)^y-1=\frac{1}{3^x}\left[\frac{1-\left(\frac{3}{2}\right)^x}{1-\frac{3}{2}}-1\right]=\frac{1}{3^x}\left[2\left(\frac{3}{2}\right)^x-2-1\right].$$
$$=\frac{1}{3^x}\left(\frac{3^x}{2^{x-1}}-3\right)=\frac{1}{2^{x-1}}-\frac{1}{3^{x-1}}$$

Επομένως, τελικά έχουμε τη συνέλιξη των δύο συναρτήσεων να δίνεται από την συνάρτηση:

$$(f*g)(x)=\frac{1}{2^{x-1}}-\frac{1}{3^{x-1}} \quad ,x=2,3,\dots ,$$

και

$$(f*g)(x)=0 \quad ,x=0,1 .$$

Γενικά, αν υποθέσουμε ότι έχουμε n στο πλήθος από ανεξάρτητες και ισόνομες[1] τ.μ, το άθροισμά τους θα είναι η τ.μ $X_1+X_2+\dots+X_n$. Η συνάρτηση κατανομής του αθροίσματος αποτελεί την n-οστή συνέλιξη της τ.μ X. Συμβολίζεται με F^{*n} όπου F είναι η σ.κ της τ.μ X. Επομένως, έχουμε πως:

- Για συνεχείς τ.μ και για κάθε $n\geq 1$,

$$F^{*n}(x)=P(X_1+X_2+\dots+X_n\leq x)=(F^{*(n-1)}*f)(x)=\int_0^x F^{*(n-1)}(x-y)f(y)dy \quad ,$$

- Για διακριτές τ.μ και για κάθε $n\geq 1$,

$$F^{*n}(x)=P(X_1+X_2+\dots+X_n\leq x)=\sum_{y=0}^{x} F^{*(n-1)}(x-y)f(y) .$$

Δηλαδή, για ανεξάρτητες και ισόνομες διακριτές τ.μ X_i η συνάρτηση κατανομής του αθροίσματος $X_1+X_2+\dots+X_n$ θα ικανοποιεί την παρακάτω σχέση:

$$F^{*n}(x)=\sum_{y=0}^{x} F^{*n-1}(x-y)f(x) . \tag{1.3}$$

Τι συμβαίνει όμως με το παραπάνω άθροισμα στην περίπτωση όπου δεν μπορούμε να γνωρίζουμε πόσες τυχαίες μεταβλητές αθροίζονται συνολικά; Τότε το πλήθος των X_i εκφράζεται από μία νέα τ.μ N. Συνεπώς, έχουμε το άθροισμα των μεταβλητών

[1]ισόνομες καλούνται οι τυχαίες μεταβλητές που ακολουθούν την ίδια κατανομή.

$X_1 + X_2 + ... + X_N$ να είναι τυχαίο άθροισμα, με την τ.μ N να είναι ανεξάρτητη από τις τ.μ $X_1, X_2, ...$.

Εξακολουθεί να μας ενδιαφέρει ποια η πιθανότητα του ενδεχομένου $\{X_1 + X_2 + ... + X_N \le x\}$ και το οποίο μπορεί να εκφραστεί ως ένωση ξένων ενδεχομένων ως εξής:

$$\{X_1 + X_2 + ... + X_N \le x\} = \cup \{(X_1 + X_2 + ... + X_N \le x) \cap (N = n)\} .$$

Μέσω του θεωρήματος ολικής πιθανότητας παίρνει κανείς το παρακάτω αποτέλεσμα για την συνάρτηση κατανομής του τυχαίου πλέον αθροίσματος:

$$F(x) = P(X_1 + X_2 + ... + X_N \le x) = \sum_{n=0}^{\infty} P(X_1 + X_2 + ... + X_N \le x) P(N = n), \tag{1.4}$$

ή

$$F(x) = \sum_{n=0}^{\infty} F^{*n}(x) P(N = n). \tag{1.5}$$

Αντίστοιχα, η σ.π του αθροίσματος αυτού δίνεται από τη σχέση:

$$f(x) = \sum_{n=0}^{\infty} f^{*n}(x) P(N = n). \tag{1.6}$$

Σημειώνεται πως οι σχέσεις (1.5) και (1.6) ισχύουν τόσο για διακριτές τ.μ X_i , όσο και για συνεχείς.

Τέλος, το ενδεχόμενο $\{X_1 + X_2 + ... + X_N > x\}$ έχει πιθανότητα που δίνεται από την σχέση

$$\overline{F}(x) = 1 - F(x) = 1 - \sum_{n=0}^{\infty} F^{*n}(x) P(N = n), \tag{1.7}$$

ή

$$\overline{F}(x) = \sum_{n=1}^{\infty} \overline{F^{*n}}(x) P(N = n),$$

και αποτελεί τη συνάρτηση δεξιάς ουράς της κατανομής του αθροίσματος.

Παράδειγμα 1.2: Ας υποθέσουμε πως μία ασφαλιστική εταιρεία έχει στην κατοχή της ένα ασφαλιστικό συμβόλαιο για το οποίο γνωρίζουμε ότι σε ένα χρόνο, το πλήθος N των απαιτήσεων έχει συνάρτηση πιθανότητας που περιγράφεται από τον πίνακα :

n	0	1	2	3
$P(N=n)$	0.1	0.3	0.4	0.2

Επιπλέον, έχουμε συλλέξει πληροφορίες για τα μεγέθη των ατομικών αποζημιώσεων που συγκεντρώνονται στον πίνακα:

X	1	2	3
$f(x)=P(X=x)$	0.5	0.4	0.1

Η τ.μ S θα παίρνει τιμές στο σύνολο $\{0,1,\ldots,9\}$. Για $X=0$ έχουμε $f(0)=0$ ενώ για $N=0$ έχουμε $P(S=0)=P(N=0)=0.1$ που εκφράζει την πιθανότητα να μην εμφανιστεί καμία ζημιά.

Η συνάρτηση πιθανότητας των ετήσιων συνολικών ζημιών S για το συγκεκριμένο ασφαλιστήριο συμβόλαιο θα δίνεται από τον τύπο:

$$f_S(x)=\sum_{n=1}^{3} P(N=n)\, f^{*n}(x)=P(N=1)\, f^{*1}(x)+P(N=2)\, f^{*2}(x)+P(N=3)\, f^{*3}(x)$$
$$=0.3\, f^{*1}(x)+0.4\, f^{*2}(x)+0.2\, f^{*3}(x).$$

Για τον υπολογισμό των συνελίξεων χρησιμοποιούμε τον τύπο (1.6) και έχουμε ότι :

$$f^{*1}(x)=f(x) \quad \text{και} \quad f^{*2}(x)=\sum_{y=0}^{x} f(y)\, f(x-y) \quad , x=0,1,2,\ldots .$$

Οπότε παίρνουμε τις τιμές:

$$f^{*2}(0)=f(0)\cdot f(0)=0\ ,$$
$$f^{*2}(1)=f(0)\cdot f(1)+f(1)\cdot f(0)=0\ ,$$
$$f^{*2}(2)=f(0)\cdot f(2)+f(1)\cdot f(1)+f(2)\cdot f(0)=0.5\cdot 0.5=0.25\ ,$$
$$f^{*2}(3)=f(0)\cdot f(3)+f(1)\cdot f(2)+f(2)\cdot f(1)+f(3)\cdot f(0)=2(0.5\cdot 0.4)=0.4$$
$$\ldots$$

Επίσης ισχύει:

$$f^{*3}(x)=\sum_{y=0}^{x} f(y)\, f^{*2}(x-y) \quad , x=0,1,2,\ldots$$

Οπότε παίρνουμε τις τιμές:

$$f^{*3}(0)=f(0)\cdot f^{*2}(0)=0\ ,$$
$$f^{*3}(1)=f(0)\cdot f^{*2}(1)+f(1)\cdot f^{*2}(0)=0\ ,$$
$$f^{*3}(2)=f(0)\cdot f^{*2}(2)+f(1)\cdot f^{*2}(1)+f(2)\cdot f^{*2}(0)=0\ ,$$
$$f^{*3}(3)=f(0)\cdot f^{*2}(3)+f(1)\cdot f^{*2}(2)+f(2)\cdot f^{*2}(1)+f(3)\cdot f^{*2}(0)=0.5\cdot 0.25=0.125$$
$$\ldots$$

Από την σχέση για την σ.π $f_S(x)$ και αντικαθιστώντας τιμές για την τ.μ *X*, μπορούμε να εξάγουμε πιθανότητες για την τ.μ *S*.

Πιο συγκεκριμένα έχουμε:

$f_S(1) = 0.3 \cdot f(1) + 0.4 \cdot f^{*2}(1) + 0.2 \cdot f^{*3}(1) = 0.15$,

$f_S(2) = 0.3 \cdot f(2) + 0.4 \cdot f^{*2}(2) + 0.2 \cdot f^{*3}(2) = 0.22$,

$f_S(3) = 0.3 \cdot f(3) + 0.4 \cdot f^{*2}(3) + 0.2 \cdot f^{*3}(3) = 0.215$.

Στον παρακάτω πίνακα παρουσιάζονται συγκεντρωμένα τα προηγούμενα αποτελέσματα.

x	$f^{*0}(x)$	$f(x)$	$f^{*2}(x)$	$f^{*3}(x)$	$f_S(x)$
0	1	0	0	0	0.1
1	0	0.5	0	0	0.25
2	0	0.4	0.25	0	0.47
3	0	0.1	0.4	0.125	0.685
4	0	0	0.26	0.3	0.849
5	0	0	0.08	0.315	0.944
6	0	0	0.01	0.184	0.9848
7	0	0	0	0.063	0.9974
8	0	0	0	0.012	0.9998
9	0	0	0	0.001	1

Μία μικτή κατανομή ορίζεται υποθέτοντας αρχικά πως η μη-αρνητική τ.μ $X \geq 0$ έχει πιθανότητα στο μηδέν ίση με $P(X=0) = f_0 > 0$ και στο υπόλοιπο πεδίο ορισμού της $(0, +\infty)$ έχει πιθανότητα ίση με $P(X>0) = 1 - P(X \leq 0) = 1 - P(X=0) = 1 - f_0$.

Τότε θα ισχύει η εξής σχέση για την τ.μ *X*:

$$P(X \leq x \mid X > 0) = \frac{P(0 < X \leq x)}{P(X > 0)} , \tag{1.8}$$

οπότε η πιθανότητα η τ.μ να μην πάρει την τιμή 0 μπορεί να γραφεί στη μορφή:

$$P(0 < X \leq x) = (1 - f_0) P(X \leq x \mid X > 0) . \tag{1.9}$$

Επομένως, η μικτού τύπου κατανομή της τ.μ X θα έχει συνάρτηση κατανομής που δίνεται από τον τύπο:

$$F_X(x) = P(X \leq x) = P(X=0) + P(0 < X \leq x) . \tag{1.10}$$

Δηλαδή, η κατανομή της τ.μ X μπορεί να γραφεί ως κυρτός[2] συνδυασμός μίας εκφυλισμένης κατανομής[3] (degenerate distribution) στο σημείο μηδέν και της κατανομής της X στο θετικό ημιάξονα:

$$F_X(x)=f_0 \cdot 1+(1-f_0)\cdot P(0<X\leq x) .$$

Παράδειγμα 1.3: Έστω ένα άθροισμα $X_1+X_2+...+X_N$ με $P(N=n)=pq^n$, $n=0,1,2,...$ και ας υποθέσουμε ότι η κατανομή των τ.μ X_i είναι η εκθετική με παράμετρο $\lambda, \lambda>0$. Παρατηρούμε πως η τ.μ N ακολουθεί γεωμετρική με παράμετρο p και $P(N=0)=p>0$. Επίσης, η πιθανότητα το άθροισμα να πάρει την τιμή 0 θα είναι $P(X_1+X_2+...+X_N=0)=P(N=0)=p$. Δεδομένου ότι το άθροισμα παίρνει τιμές στο διάστημα $(0,+\infty)$ ενώ συγχρόνως υπάρχει θετική πιθανότητα αυτό να πάρει την τιμή 0, συμπεραίνουμε ότι η κατανομή του αθροίσματος είναι μικτού τύπου με μάζα πιθανότητας p στο σημείο μηδέν και συνεχής για θετικές τιμές.

1.2 Στοχαστικές Διαδικασίες

Η ανάγκη μαθηματικής περιγραφής και μοντελοποίησης συστημάτων, τα οποία εξελίσσονται στο χρόνο κατά τρόπο τυχαίο (random, stochastic) και όχι ντετερμινιστικό[4] οδήγησε στο να αναπτυχθεί τον τελευταίο αιώνα μία σημαντική περιοχή της θεωρίας Πιθανοτήτων, αυτή της Στοχαστικής Ανάλυσης .

Αφορμή για την ανάπτυξή της στάθηκε το 1827 ένα φαινόμενο από τον χώρο της Φυσικής. O Brown παρατήρησε μία άτακτη κίνηση σε σωματίδιο ενώ αυτό βρισκόταν μέσα σε νερό ή αέριο. Οι πρώτες δημοσιεύσεις είχαν ως κύριο αντικείμενο έρευνας την εξέλιξη, μέσα στο χρόνο, ενός μεταβλητού συστήματος. O Bachelier (1900) μελέτησε τις μεταπτώσεις των τιμών της αγοράς και κατέληξε σε κάποια μορφή στοχαστικής διαδικασίας. O Lundberg (1903) δημοσίευσε τη διατριβή του, εισάγοντας έναν νέο τρόπο με τον οποίο οι ασφαλιστικές θα μπορούσαν να υπολογίζουν τις επικείμενες απώλειες από ζημιές στα χαρτοφυλάκιά τους (σ.δ Poisson).

---

[2] κυρτός καλείται ένας γραμμικός συνδυασμός, του οποίου τα στοιχεία έχουν συντελεστές που αθροίζουν στη μονάδα.

[3] πρόκειται για την κατανομή μίας τ.μ που όλη η πιθανότητά της βρίσκεται συγκεντρωμένη σ΄ένα σημείο, δηλαδή $P(X=a)=1$.

[4] είναι τα γεγονότα εκ των προτέρων καθορισμένα είτε λόγω κάποιων νόμων είτε κάποιων συγκεκριμένων αιτίων.

Μαθηματικά, Στοχαστική διαδικασία ή Στοχαστική Ανέλιξη ορίζεται μία οικογένεια τυχαίων μεταβλητών $\{X(t): t \in T\}$.

Το σύνολο T λέγεται σύνολο δεικτών της διαδικασίας ή παραμετρικός χώρος αυτής. Αν το σύνολο T είναι αριθμήσιμο[5] τότε η διαδικασία $\{X(t): t \in T\}$ λέγεται στοχαστική διαδικασία διακριτού χρόνου. Αν το T είναι σύνολο μη-αριθμήσιμο τότε η διαδικασία $\{X(t): t \in T\}$ λέγεται διαδικασία συνεχούς χρόνου.

Το σύνολο όλων των δυνατών τιμών που μπορεί να πάρει η τυχαία μεταβλητή $X(t)$, $t \in T$ ονομάζεται χώρος καταστάσεων και συνήθως συμβολίζεται με S. Ο χώρος των καταστάσεων μπορεί να είναι ποιοτικός ή ποσοτικός, διακριτός ή συνεχής. Όταν το S είναι αριθμήσιμο η διαδικασία καλείται αλυσίδα.

Σημειώνεται πως ο δείκτης T δεν αντιστοιχεί αναγκαστικά σε χρονικές μονάδες. Ο λόγος που ορίζεται και συνήθως λογίζεται ως χρόνος είναι η δυναμική φύση των υπο μελέτη προβλημάτων, τα οποία συνήθως εξελλίσσονται στο χρόνο.

Οι στοχαστικές διαδικασίες διακρίνονται βάσει του είδους του παραμετρικού χώρου και του χώρου των καταστάσεων στις εξής τέσσερις κατηγορίες :

- Συνεχής με συνεχή παραμετρικό χώρο
- Συνεχής με διακριτό παραμερικό χώρο
- Διακριτή με συνεχή παραμετρικό χώρο
- Διακριτή με διακριτό παραμετρικό χώρο

Μία από τις σημαντικότερες κατηγορίες στοχαστικών διαδικασιών είναι οι διαδικασίες Markov ή Μαρκοβιανές στοχαστικές διαδικασίες. Κάποια από τα σημαντικότερα σήμερα στοχαστικά μοντέλα της Θεωρίας Κινδύνου περιέχουν διαδικασίες Markov. Τα μοντέλα αυτά εφαρμόζονται σε μία πληθώρα περιπτώσεων όπως στην μελέτη των τάσεων της χρηματιστηριακής αγοράς, σε ασφαλιστικές μεθόδους, σε επιχειρηματικούς κινδύνους που αντιμετωπίζουν εταιρείες κ.λ.π. Αποτελούν, δηλαδή, ένα εξαιρετικά χρήσιμο εργαλείο για την πρόβλεψη και διαχείριση των κινδύνων, από ασφαλιστικούς οργανισμούς και χρηματοπιστωτικά ιδρύματα.

Στην συνέχεια για τις Μαρκοβιανές Διαδικασίες που θα χρησιμοποιήσουμε, υποθέτουμε ότι ισχύουν τα εξής:

[5] Αριθμήσιμο είναι ένα σύνολο όταν είναι πεπερασμένο ή ισοδύναμο με τους φυσικούς αριθμούς $\mathbb{N}$

- Ο χώρος των καταστάσεων *S* της διαδικασίας είναι διακριτός. Επίσης, αν η τ.μ $X(t)=j$, για κάποια κατάσταση j, τότε λέμε πως η τ.μ $X(t)$ βρίσκεται στην κατάσταση j σε χρόνο t.
- Η ακολουθία των τ.μ $\{X(t) : t \in T\}$ έχει την αμνήνονα ιδιότητα ή ιδιότητα Markov (Μ.Ι), δηλαδή:

$$P(X_{t+1}=j \mid X_t=i, X_{t-1}=i_{t-1}, \cdots, X_0=i_0)=P(X_{t+1}=j \mid X_t=i).$$

Αν δούμε την χρονική στιγμή t σαν το παρόν, την t+1 σαν το μέλλον και τις 0,1,...,t-1 ως το παρελθόν, τότε η μαρκοβιανή ιδιότητα υποδηλώνει την εξάρτηση του μέλλοντος μόνο από το παρόν κι όχι από το παρελθόν. Σημειώνεται πως δεν είναι απαραίτητο να γνωρίζουμε όλες τις καταστάσεις στο παρελθόν για να εφαρμοστεί η Μ.Ι καθώς επίσης κι πως ως παρελθόν δεν θεωρείται απαραίτητα η προηγούμενη χρονική στιγμή αλλά μπορεί να ληφθεί η πλησιέστερη στο παρόν χρονική στιγμή.

Τέλος, κρίνεται σκόπιμο να αναφερθεί και μία ακόμη ιδιότητα η οποία ικανοποιείται από κάποιες Μαρκοβιανές διαδικασίες. Πιο συγκεκριμένα, η ιδιότητα της ομογένειας του χρόνου. Δεδομένου ότι τη χρονική στιγμή t η διαδικασία βρίσκεται στην κατάσταση i, η πιθανότητα την χρονική στιγμή t+1 να βρεθεί στην κατάσταση j είναι ανεξάρτητη από τη χρονική στιγμή t. Τέτοιες διαδικασίες καλούνται στάσιμες ή ομογενείς (stationery, time homogenous). Στην περίπτωση που υπάρχει κάποιο είδος εξάρτησης ανάμεσα στην κατάσταση στην οποία βρίσκεται η διαδικασία και στο χρόνο που θα γίνει η μετάβαση της σε άλλη κατάσταση τότε η Μαρκοβιανή διαδικασία λέμε ότι είναι μη-ομογενής.

Στην παρούσα εργασία, θα μας απασχολήσουν κυρίως στοχαστικές διαδικασίες των οποίων τόσο παραμετρικός χώρος όσο και ο χώρος των καταστάσεων τους θα είναι διακριτός. Επικεντρώνουμε δηλαδή το ενδιαφέρον μας σε εφαρμογές αλυσίδων διακριτού χρόνου.

1.3 Μαρκοβιανές Αλυσίδες διακριτού χρόνου

Πρόκειται για διακριτές στοχαστικές διαδικασίες, οι οποίες ικανοποιούν την Μαρκοβιανή ιδιότητα.Για τις αλυσίδες αυτές σημαντικό ρόλο παίζουν τα παρακάτω μεγέθη:

- $p_{ij}=P(X_{n+1}=j\,|X_n=i)=P(X_1=j\,|X_0=i)$: Η πιθανότητα μετάβασης σε ένα βήμα.
- $p_{ij}^{(n)}=P(X_n=j\,|X_0=i)$: Η πιθανότητα μετάβασης από την *i* στην *j* σε *n* βήματα.
- $P^{(n)}=(p_{ij}^{(n)})$: Ο πίνακας όλων των πιθανοτήτων μετάβασης σε *n* βήματα.
- $f_{ij}^{(1)}=P(X_1=j\,|X_0=i)=p_{ij}$: Η πιθανότητα μετάβασης (ή επίσκεψης) από την *i* κατάσταση στην *j* σε ένα βήμα.
- $f_{ij}^{(n)}=P(X_n=j\,,X_k\neq j\,,\forall k=1,\dots,n-1\,|X_0=i)$: Η πιθανότητα πρώτης μετάβασης από την *i* κατάσταση στην *j* σε *n* βήματα.
- $T_{ij}=min\{n:X_n=j\,,X_k\neq j\,,\forall k=1,\dots,n-1\,|X_0=i\}$: Η χρονική στιγμή πρώτης μετάβασης από την *i* κατάσταση στην *j*.

Ο πίνακας *P* με γενικό στοιχείο p_{ij} λέγεται πίνακας πιθανοτήτων μετάβασης πρώτης τάξης ή πιο σύντομα απλά πίνακας μετάβασης.

Τέλος, γνωρίζουμε πως για τις ομογενείς Μαρκοβιανές αλυσίδες να ισχύει πως η πιθανότητα μετάβασης σε μία κατάσταση σε *m* βήματα είναι ίδια ανεξαρτήτως της χρονικής στιγμής που ξεκινάει.

Έχουμε δηλαδή ότι,

$$P(X_{n+m}=j\,|X_n=i)=P(X_m=j\,|X_0=i)\,.$$

Η κατάσταση *j* λέγεται προσιτή από την i αν υπάρχει $n\geq 0$ τέτοιο ώστε η πιθανότητα $p_{ij}^{(n)}>0$. Συμβολικά γράφουμε πως $i\Rightarrow j$.

- Μια κατάσταση i λέγεται επαναλαμβανόμενη (recurrent) ή απορροφητική (absorbing)αν $f_{ii}=1$, δηλαδή αν η αλυσίδα βρεθεί στην ίδια κατάσταση δεν απομακρύνεται ποτέ από αυτή.
- Μια κατάσταση i λέγεται μεταβατική (transient) αν $f_{ii}<1$, δηλαδή υπάρχει πιθανότητα $1-f_{ii}>0$ να διαφύγει από αυτή την κατάσταση.

1.4 Αθροισμα τυχαίου αριθμού τυχαίων μεταβλητών

Μία στοχαστική διαδικασία $\{N(t):t\geq 0\}$ ονομάζεται απαριθμήτρια διαδικασία αν και μόνο αν

- $N(t)>0, όταν\ \ N(0)=0$,
- $N(t)$ είναι διακριτή,
- αν $s\leq t$ τότε και $N(s)\leq N(t)$.

Ας υποθέσουμε ότι έχουμε $N(t)$ το πλήθος από ανεξάρτητες και ισόνομες μεταξύ τους τυχαίες μεταβλητές. Αντιμετωπίζουμε δυσκολία στον υπολογισμό του αθροίσματός τους επειδή δεν γνωρίζουμε τον ακριβή αριθμό τους. Υπάρχει μια σειρά από εφαρμογές όπου είναι απαραίτητος ο υπολογισμός της στοχαστικής συμπεριφοράς ενός τέτοιου αθροίσματος. Κλασικό παράδειγμα που αποτέλεσε και το έναυσμα για την ανάπτυξη και τη μοντελοποίηση τέτοιων μεταβλητών είναι ο συνολικός κίνδυνος που διατρέχει μία ασφαλιστική επιχείρηση από ένα σύνολο ασφαλιστηρίων συμβολαίων. Ποια είναι δηλαδή η συνολική υποχρέωση προς τους ασφαλισμένους, που θα σχηματιστεί μέσα σε μία χρονική περίοδο, για την εταιρεία αυτή; Καταλαβαίνουμε πως είναι αδύνατο να γνωρίζει κανείς εκ των προτέρων πόσες ζημιές θα συμβούν αλλά κι ποιο το μέγεθος της κάθε μίας εξ αυτών.

Επομένως, προκειμένου να δοθεί μία απάντηση στο ερώτημα αυτό, θεωρούμε πως το πλήθος των ζημιών που θα προκύψουν την επόμενη χρονική περίοδο ή αλλιώς η συχνότητα εμφάνισης του κινδύνου συμπεριφέρεται ως μία απαριθμήτρια στοχαστική διαδικασία. Επίσης, το μέγεθος του κινδύνου ή αλλιώς το ύψος κάθε μίας από τις ανεξάρτητες τ.μ που απαρτίζει το άθροισμα, είναι κι αυτό άγνωστο και αναφέρεται συνήθως ως σφοδρότητα του κινδύνου. Η σφοδρότητα του κινδύνου περιγράφεται από τυχαίες μεταβλητές που παίρνουν για τιμές τα ύψη των ζημιών που θα προκύψουν.

Ο διαχωρισμός για τον συνολικό κίνδυνο, σε συχνότητα και σφοδρότητα προσφέρει πολλά πλεονεκτήματα. Γύρω από τον διαχωρισμό αυτό αναπτύχθηκε το βασικό μοντέλο συλλογικού κινδύνου για τον υπολογισμό της κατανομής του τυχαίου αθροίσματος τυχαίων μεταβλητών

$$S=\sum_{i=1}^{N(t)} X_i \ \ ,$$

όπου το $N(t)$ μετράει την συχνότητα εμφάνισης της τ.μ X και τα X_i , i=1,2,... παριστάνουν την σφοδρότητα του κινδύνου.

Υπό το συλλογικό πρότυπο κινδύνου υποθέτουμε ανεξαρτησία ανάμεσα στην $N = N(t)$ και τις τ.μ X_i . Παρατίθενται σχετικά με την κατανομή της τ.μ X :

- $F_X(x) = P(X \leq x)$ η συνάρτηση κατανομής της *X*.
- $\overline{F}(x) = P(X > x) = 1 - F(x)$ η συνάρτηση δεξιάς ουράς της *X*.

Για την τ.μ $N(t)$ έχουμε αντίστοιχα:

- $p_n = P(N = n), n = 0,1,2,\ldots$ η συνάρτηση πιθανότητας της *N*.

Με βάση τα παραπάνω προκύπτουν για την κατανομή του *S* τα εξής αποτελέσματα:

- $f_S(x) = \sum_{k=o}^{\infty} f_N(k) f_X^{(*k)}(x)$ η συνάρτηση πιθανότητας του *S*.
- $F_S(x) = \sum_{k=o}^{\infty} f_N(k) F_X^{(*k)}(x)$ η αθροιστική συνάρτηση πιθανότητας του *S*.
- $M_S = M_N(lnM_X(t)) = P_N(M_X(t)) = (P_X \circ M_X)(t)$ η ροπογεννήτρια συνάρτηση της τ.μ *S*.
- $P_S(t) = P_N(P_X(t)) = (P_N \circ P_X)(t)$ η πιθανογεννήτρια συνάρτηση της τ.μ *S*.
- $E(S) = E(X)E(N)$
- $Var(S) = E(N)Var(X) + E(X)^2 Var(N)$

Η τ.μ *S* λέμε ότι ακολουθεί σύνθετη κατανομή επειδή οι γεννήτριες συναρτήσεις της, προκύπτουν από τις γεννήτριες των *X*, *N* μέσω της πράξης της σύνθεσης. Το όνομα της σύνθεσης συνήθως καθορίζεται από την κατανομή της απαριθμήτρια *N*, η οποία καλείται και πρωτογενής κατανομή στο συλλογικό πρότυπο κινδύνου. Η κατανομή της *X* καλείται δευτερογενής. Η κατανομή του *S* μπορεί να είναι τριών ειδών: διακριτή, συνεχής είτε μικτού τύπου.

Αν η τ.μ *X* είναι διακριτή τότε και η κατανομή του *S* είναι διακριτή και το *S* θα αποτελεί ένα τυχαίο άθροισμα διακριτών τυχαίων μεταβλητών. Η $N(t)$ είναι πάντοτε μία διακριτή τ.μ.

Αν η τ.μ *X* είναι συνεχής τότε η κατανομή του *S* μπορεί να είναι συνεχής ή μικτού τύπου. Το είδος της εξαρτάται από την τιμή της πιθανότητας *P*(*N*=0). *Δ*ηλαδή, εξαρτάται από την πιθανότητα να μην προκύψει κανένας κίνδυνος.

- Αν $P(N = 0) = 0$, τότε η κατανομή του *S* είναι συνεχής, ως άθροισμα συνεχών τ.μ
- Αν $P(N = 0) > 0$, τότε η κατανομή του *S* είναι μικτού τύπου. Δηλαδή, το τυχαίο άθροισμα *S* θα έχει κάποια πιθανότητα συγκεντρωμένη στο σημείο μηδέν ίση

με $P(S=0)=P(N=0)=p_0$ και στο υπόλοιπο διάστημα $(0,+\infty)$ θα είναι μία συνεχής κατανομή.

1.5 Οικογένεια Κατανομών Panjer

Προκειμένου να υπολογιστεί η κατανομή του τυχαίου αθροίσματος *S* είχε αναπτυχθεί η βασική μέθοδος ή μέθοδος των συνελίξεων, που για πολλά χρόνια χρησιμοποιείτο στο μοντέλο συλλογικού κινδύνου. Ο Panjer (1981) έθεσε τα θεμέλια μίας νέας προσέγγισης της για τον υπολογισμό της κατανομής αυτής. Προσπαθώντας να βελτιώσει τον τρόπο με τον οποίο αναλογιστές της εποχής εκείνης υπολόγιζαν την κατανομή των συνολικών απαιτήσεων μίας ασφαλιστικής, κατασκεύασε ένα αναδρομικό τύπο υπολογισμού της συνάρτησης πιθανότητας $f_S(x)$ της τ.μ *S* για διακριτές τ.μ *X*, θεωρώντας ότι η κατανομή της τ.μ $N(t)$ ανήκει σε μία ευρεία οικογένεια κατανομών.

Τα μέλη της οικογένειας αυτής είναι κατανομές, η συνάρτηση πιθανότητας των οποίων ικανοποιεί μία ιδιότητα. Όταν αυτό συμβαίνει υπάρχει η δυνατότητα να βρούμε αναδρομικές σχέσεις υπολογισμού της σ.π της *S*.

Μία διακριτή τυχαία μεταβλητή $N\in\{0,1,2,...\}$ λέμε ότι ανήκει στην κλάση $R(a,b,0)$, αν η συνάρτηση πιθανότητας της $p_n=P(N=n)$ ικανοποιεί την παρακάτω αναδρομική σχέση πρώτης τάξης :

$$p_n=(a+\frac{b}{n})\,p_{n-1}\,,\qquad n=1,2,3,...\,, \tag{1.11}$$

όπου *a,b* σταθερές ποσότητες, χαρακτηριστικές για την κάθε κατανομή.

Η παραπάνω κλάση είναι επίσης γνωστή και ως οικογένεια Panjer. Η σχέση (1.11) μαζί με κάποια αρχική συνθήκη p_0 μπορούν να ορίσουν πλήρως την κατανομή του τυχαίου αθροίσματος *S*. Πρέπει να τονιστεί επίσης πως δεν υπάρχουν για όλους τους συνδυασμούς των σταθερών *a,b* συναρτήσεις πιθανότητας που να επαληθεύουν τον τύπο (1.11).

Οι Sundt και Jewell (1981) απέδειξαν πως η κλάση $R(a,b,0)$ περιέχει τις κατανομές: Poisson, Διωνυμική, Αρνητική Διωνυμική και την Γεωμετρική κατανομή.

Στον παρακάτω πίνακα (βλ. Χατζηκωνσταντινίδης (2014)) δίνονται συγκεντρωμένες οι τιμές των σταθερών παραμέτρων *a,b* για τις προαναφερθείσες κατανομές.

Πίνακας 1.1 Μέλη της κλάσης *R(a,b,0)* και οι αντίστοιχες τιμές των παραμέτρων.

Κατανομή	***a***	***b***	$\boldsymbol{P(N=0)}$
Poisson(λ)	0	λ	$e^{-\lambda}$
B(n, p)	$\frac{-p}{q}$	$(n+1)\frac{p}{q}$	q^n
NB(r, q)	q	$(r-1)q$	p^r
G(p)	q	0	p

Αν μία τυχαία μεταβλητή *N* ανήκει στην κλάση $R(\alpha, \beta, 0)$, τότε αποδεικνύεται πως η πιθανογεννήτρια συνάρτησή της $P_N(t)=E(e^{tN})$ επαληθεύει τη παρακάτω διαφορική εξίσωση:

$$P'_N(t)=\frac{a+b}{1-at}P_N(t)\,. \tag{1.12}$$

Με την βοήθεια της (1.12) μπορούν να υπολογιστούν παραγοντικές ροπές για την τυχαία μεταβλητή *N*, άρα και οι αντίστοιχες ροπές για την κατανομή του τυχαίου αθροίσματος *S*.

Πιο συγκεκριμένα έχουμε ότι:

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

Σημειώνουμε πως οι παραπάνω τύποι για τη μέση τιμή και τη διασπορά του *S* ισχύουν για οποιοδήποτε τύπο κι αν έχει η κατανομή της τ.μ *X*, δηλαδή είτε η *X* είναι συνεχής, διακριτή είτε μικτού τύπου τυχαία μεταβλητή.

Όταν η συχνότητα εμφάνισης *N* ενός κινδύνου ανήκει στην οικογένεια Panjer $R(\alpha, \beta, 0)$ και η σφοδρότητα του *X* είναι μία διακριτή τυχαία μεταβλητή με σ.π $f_X(x)=P(X=x)$ με $x=0,1,2,...$ τότε η κατανομή του *S* ικανοποιεί την παρακάτω αναδρομική σχέση:

$$f_S(x)=\frac{1}{1-af(0)}\sum_{y=1}^{x}\left(a+\frac{by}{x}\right)f_X(y)f_S(x-y) \quad , x=1,2,3,... \tag{1.13}$$

με αρχική συνθήκη $f_S(x)=P_N(f(0))$.

Για παράδειγμα, αν η τ.μ N ακολουθεί την κατανομή Poisson(λ), $\lambda>0$ τότε θα έχουμε α=0, b=λ και η συνάρτηση πιθανότητας του αθροίσματος $P(S=x)$ θα ικανοποιεί την αναδρομική σχέση:

$$P(S=x)=\frac{\lambda}{x}\sum_{i=1}^{x} i\, f(i)\, P(S=x-i) \;\; , x=1,2,3,\ldots,$$

με

$$P(S=0)=\mathrm{e}^{\lambda(f(0)-1)}.$$

Εύκολα διαπιστώνει κανείς πως εφόσον η τ.μ N ακολουθεί κατανομή Poisson, θα ανήκει στην οικογένεια Panjer με $\alpha=0$, $b=\lambda$ και $P(N=n)=\mathrm{e}^{\lambda(n-1)}$. Επομένως, με απλή αντικατάσταση στην σχέση (1.12) βρίσκουμε τον αναδρομικό τύπο για την σ.π μίας Poisson.

Ανακεφαλαιώνοντας, έχουμε ότι αν η τ.μ $N\in R(\alpha,\beta,0)$ τότε για την τ.μ S ισχύουν τα εξής:

- Αν η τ.μ X παίρνει μη αρνητικές διακριτές τιμές στο $(0,1,2,\ldots)$ τότε θα ικανοποιεί την αναδρομική σχέση (1.13).
- Αν η τ.μ X παίρνει μόνο θετικές διακριτές τιμές στο $(1,2,3,\ldots)$ τότε $f_X(0)=0$, αφού η πιθανότητα της στο μηδέν θα είναι $P(X=0)=0$ και επομένως η αρχική τιμή για την τ.μ S θα ισούται με $f_S(0)=P_N(f(0))=P_N(0)=P(N=0)=p_0$.
- Αν η τ.μ X είναι συνεχής στο $(0,+\infty)$, η αρχική τιμή θα ισούται με:
  $$f_S(0)=P(S=0)=P(N=0)=p_0.$$
  - Αν $N\in\{0,1,2,\ldots\}$, τότε θα υπάρχει θετική πιθανότητα στο μηδέν $f_S(0)=p_0>0$.

    Επομένως, το S θα έχει μία μικτού τύπου κατανομή με πιθανότητα p_0 στο μηδέν και συνεχής για τις θετικές τιμές της.
  - Αν $N\in\{1,2,3,\ldots\}$, τότε η πιθανότητα να πάρει η τ.μ την τιμή 0 είναι $f_S(0)=p_0=0$. Επομένως, το S θα ακολουθεί μία συνεχή κατανομή στο $(0,+\infty)$

Για σειρά δεκαετιών η βασική μέθοδος υπολογισμού της κατανομής του S βασιζόταν στις συνελίξεις συναρτήσεων πιθανότητας γνωστών κατανομών για την προσέγγιση της τ.μ S. Μετά και το έργο του Panjer, κύρια μέθοδος υπολογισμού είναι ο αναδρομικός τύπος.

Παρόλα αυτά, αξίζει να αναφέρουμε πως από τον Panjer μέχρι και σήμερα έχουν υπάρξει κάποιες γενικεύσεις του έργου του, σημαντικότερη εκ των οποίων είναι του Sundt

(1992, 2003). Στην εργασία του γενίκευσε την (1.10) αντικαθιστώντας την από την αναδρομική σχέση:

$$p_n=\sum_{s=1}^{r\wedge n}(a_s+\frac{b_s}{n})\,p_{n-s}\,,\quad n\geq 1\ .$$

Σε αυτή την περίπτωση η σχέση (1.13) που ικανοποιεί η τ.μ S θα είναι:

$$f_S(x)=\sum_{u=1}^{x} f_S(x-u)\sum_{s=u}^{r\wedge x} f_X^{*s}(u)(\alpha_S+\frac{b_S}{s}\frac{u}{x})\,,\quad x\geq 1\ . \tag{1.14}$$

Επίσης, ο Hess (2002) έδειξε πως η σχέση (1.10) ικανοποιείται και από κατανομές που ανήκουν στην οικογένεια $R(a,b,0)$ μέχρι τάξης k με $n\geq k$, οπότε και άνοιξε το δρόμο για την μελέτη καινούργιων κατανομών σχετικά με την οικογένεια Panjer.

1.6 Αναδρομικές σχέσεις στη Θεωρία Κινδύνου

Ο κυριότερος λόγος που οδήγησε στην ανάπτυξη και εδραίωση της θεωρίας των κινδύνων είναι η ανάγκη αναλογιστών να προσδιορίσουν την κατανομή των συνολικών αποζημιώσεων που προκύπτουν από ένα χαρτοφυλάκιο συμβολαίων μιας ασφαλιστικής εταιρείας.

Γενικά, θεωρούμε πως οι συνολικές αποζημιώσεις περιγράφονται από ένα άθροισμα τυχαίων αριθμών $S=\sum_{i=1}^{n} X_i$. Ιδανικά θα θέλαμε να προσδιοριστεί πλήρως η κατανομή της μεταβλητής S έτσι ώστε να έχουμε στη διάθεσή μας έναν κλειστό τύπο υπολογισμού για τις πιθανότητες που αφορούν την S. Στην πράξη όμως κάτι τέτοιο σπάνια επιτυγχάνεται.

Για το λόγο αυτό η θεωρία των κινδύνων έχει αναπτύξει δύο βασικά πρότυπα για να αντιμετωπίσει το πρόβλημα προσδιορισμού της κατανομής των συνολικών αποζημιώσεων, το ατομικό πρότυπο και το συλλογικό πρότυπο κινδύνου. Η διαφορά τους εντοπίζεται στον τρόπο με τον οποίο κάθε πρότυπο προσεγγίζει το πρόβλημα.

Έτσι, στο ατομικό μοντέλο βλέπουμε κάθε συμβόλαιο ξεχωριστά και αθροίζουμε ως τυχαίες μεταβλητές τους κινδύνους (τις εν δυνάμει αποζημιώσεις) που μπορεί να προκύψουν για το κάθε συμβόλαιο. Επομένως, n είναι το πλήθος συμβολαίων που θα έχει το υπό μελέτη χαρτοφυλάκιο, ενώ η τ.μ $X_i \;, i=1,\ldots,n$ μετράει το ύψος του κινδύνου (τιμή αποζημίωσης) που θα προκύψει από το *i* συμβόλαιο.

Το συλλογικό πρότυπο αντιλαμβάνεται το τυχαίο άθροισμα σφαιρικά. Το βλέπει σαν μία μηχανή παραγωγής κινδύνων. Το μοντέλο αυτό δεν εξετάζει ξεχωριστά τα επιμέρους συμβόλαια για να ελέγξει ποιο το ύψος της αποζημίωσης που απαιτείται από το καθένα αλλά αθροίζει ως τ.μ κάθε κίνδυνο (αποζημίωση) που θα προκύψει ξεχωριστά, ανεξάρτητα από το συμβόλαιο που την προκαλεί, θεωρώντας πως το πλήθος των κινδύνων σε ένα χρονικό διάστημα $(0,t)$ είναι επίσης μία τυχαία μεταβλητή, έστω $N(t)$. Οι κίνδυνοι εδώ θεωρούνται ανεξάρτητες και ισόνομες τ.μ εν αντιθέσει με το ατομικό πρότυπο όπου δεν είναι απαραίτητα ισόνομοι. Επομένως, το τυχαίο πλέον άθροισμα $S=\sum_{i=1}^{N(t)} X_i$ είναι μία στοχαστική διαδικασία (σ.δ) συνεχούς ή διακριτού χρόνου με την $N(t)$ να απαριθμεί το πλήθος των κινδύνων που συμβαίνουν σε μία συγκεκριμένη χρονική περίοδο και την τ.μ $X_i,\; i=1,\ldots,N(t)$ να εκφράζει το ύψος του *i* κινδύνου.

Γενικά στον αναλογισμό, παίζουν πολύ σημαντικό ρόλο οι κατανομές που έχουν μία συγκέντρωση πιθανότητας στο μηδέν εις βάρος της υπόλοιπης κατανομής. Το γεγονός πως μία αποζημίωση θα προκύψει μηδενική είναι αρκετά συχνό. Παραδείγματος χάριν, σε ένα συμβόλαιο μπορεί να προβλέπεται ένα ποσό απαλλαγής (deductible) και έτσι να μη γίνεται η καταβολή μικρών και μέχρι του ποσού αυτού αποζημιώσεων. Αντίστοιχα, ένας αντασφαλιστής που μετέχει σε ένα ασφαλιστικό σχήμα υπερβάλλουσας ζημιάς (excess of loss) πιθανότατα να μην κληθεί να λάβει μέρος σε μεγάλο αριθμό αποζημιώσεων, καθώς οι περισσότερες θα εμπίπτουν στην ασφάλιση του ασφαλιστή.

Ο αναδρομικός τύπος (1.11), γνωστός κι ως τύπος Panjer δίνει τη δυνατότητα εύκολου υπολογισμού της κατανομής των συνολικών αποζημιώσεων σε ένα χαρτοφυλάκιο. Παρόλα αυτά, μετά τον Panjer και στα 30 χρόνια που ακολούθησαν μελετήθηκαν περιπτώσεις και αναπτύχθηκαν αποτελέσματα και αναδρομικές σχέσεις για αρκετές άλλες οικογένειες κατανομών.

Υπάρχουν οικογένειες απαριθμητριών τ.μ για τις οποιες ισχύουν αναδρομικές σχέσεις ανάλογες με αυτή του Panjer. Δηλαδή, αν η σ.π του πλήθους των αποζημιώσεων ανήκει σε κάποια από τις κλάσεις που έως σήμερα έχουν μελετηθεί κι έχουν βρεθεί αποτελέσματα (με τα ύψη των αποζημιώσεων να είναι διακριτές τ.μ) τότε μπορεί να υπολογιστεί με αναδρομικό τρόπο η κατανομή των συνολικών αποζημιώσεων σε ένα χαρτοφυλάκιο συμβολαίων.

Όπως έχει ήδη αναφερθεί ο Sundt (1992) έδωσε μία αναδρομική σχέση για την σ.π $f_S(x)$ γενικεύοντας τον αρχικό τύπο Panjer (1.11) για την περίπτωση όπου η απαριθμήτρια $N(t)$ ανήκει στην οικογένεια $R(a,b,0)$, ενώ η τ.μ X_i παίρνει θετικές ακέραιες τιμές. Πιο συγκεκριμένα ο Sundt(1992) έδωσε τον παρακάτω αναδρομικό τύπο:

$$p_n=\sum_{i=1}^{k}(a_i+\frac{b_i}{n})p_{n-i}\ ,\ \ n\geq 1\ , \tag{1.15}$$

με $k\in\mathbb{Z},\ \ k>0$ και τους συντελεστές a_i , b_i να είναι γνωστοί αριθμοί.

Μέσα από υπολογισμούς κατέληξε πως η σ.π της σύνθετης τ.μ S δίδεται από τον τύπο:

$$f_S(x)=\sum_{y=1}^{x}f(x-y)\sum_{i=1}^{k}(a_i+\frac{b_i}{i}\frac{y}{x})F_X(y)^{i*}\ ,\ \ x\geq 1\ . \tag{1.16}$$

Οι Panjer και Willmot (1992) αλλά και ο Hesselager λίγο αργότερα μελέτησαν την περίπτωση που η $N(t)$ ανήκει στη κλάση μη-αρνητικών κατανομών με θετική μάζα πιθανότητας στο μηδέν, την οποία και συμβολίζουμε, P_{10} και η τ.μ X_i ανήκει στην κλάση των μη – αρνητικών κατανομών, η οποία συνήθως συμβολίζεται, P_1. Απέδειξαν πως η σύνθετη κατανομή του S ικανοποιεί την εξής αναδρομική σχέση:

$$p_n=\frac{\sum_{j=0}^{t}a(j)n^j}{\sum_{j=0}^{t}b(j)n^j}p_{n-1}\ ,\ \ n=1,2,... \tag{1.17}$$

Οι Sobreno και Wang (1994) έδωσαν μία γενίκευση του τύπου αυτού για την περίπτωση όπου και η συχνότητα με την οποία εμφανίζονται οι κίνδυνοι είναι μη αρνητικές τ.μ, δίνοντας έτσι την παρακάτω γενικευμένη αναδρομική σχέση, όπου για κάθε $n=r+1,r+2,..$ και $r>0$ ισχύει ότι:

$$p_n=\sum_{i=1}^{k}\frac{\sum_{j=0}^{t}a(i,j)n^j}{\sum_{j=0}^{t}b(j)n^j}p_{n-1}\ . \tag{1.18}$$

Ο Ambagaspitiya (1995) κατασκεύασε έναν γενικό αναδρομικό τύπο για κατανομές συχνότητας που ανήκουν στην κλάση P_1 των μη- αρνητικών κατανομών. Έτσι, αν έχουμε μία κατανομή δύο παραμέτρων $p_{a,b} \in P_1$ με a,b σταθερές τότε για την σ.π ισχύει η αναδρομική σχέση:

$$p_{a,b}(n)=(u(a,b)+\frac{v(a,b)}{n})\,p_{a+b,b}(n-1)\,,\quad n=r+1,r+2,\ldots \qquad (1.19)$$

Επομένως, η μείξη από την κατανομή που ικανοποιεί την σχέση (1.19) και μία h κατανομή που περιγράφει τις τιμές που μπορεί να πάρει ένας κίνδυνος X_i, $i=0,1,2,\ldots$ και η οποία ανήκει στην οικογένεια κατανομών με μάζα πιθανότητας στο μηδέν.

Τότε υπάρχει τύπος υπολογισμού της συνάρτησης πιθανότητας:

$$\begin{aligned} f_{a,b}(x)=&\sum_{n=1}^{r}\left(p_{a,b}(n)-(u(a,b)+\frac{v(a,b)}{n})\,p_{a+b,b}(n-1)\right)h^{n*}(x)\\ &+\sum_{y=1}^{x}\left(u(a,b)+v(a,b)\frac{y}{x}\right)h(y)\,f_{a+b,b}(x-y) \end{aligned} \qquad (1.20)$$

με $x=1,2,\ldots$.

Εύκολα διαπιστώνει κανείς πως οι αναδρομικές σχέσεις τέτοιου είδους βοηθούν σημαντικά στην μείωση του χρόνου επεξεργασίας αλλά και των πράξεων που απαιτούνται για τον υπολογισμό τέτοιων ποσοτήτων, ιδιαίτερα αν τις συγκρίνει κανείς με την βασική μέθοδο των συνελίξεων. Το όφελος που αποκομίζει ο χρήστης τους είναι τεράστιο, γι' αυτό και η δημιουργία νέων σχέσεων για περισσότερες κατηγορίες κατανομών, αποτελεί τα τελευταία χρόνια αδήριτη ανάγκη της θεωρίας των κινδύνων.

ΚΕΦΑΛΑΙΟ 2

Κατανομές Phase-type

Οι κατανομές phase type ανήκουν σε μία ευρεία οικογένεια κατανομών. Εισήχθησαν για πρώτη φορά από τον Neuts (1975) και έχουν βρει εφαρμογή σε ένα πλήθος από προβλήματα στοχαστικής μοντελοποίησης. Ενδεικτικά αναφέρουμε τη μοντελοποίηση νοσοκομειακών δεδομένων, τις τηλεπικοινωνίες, τη θεωρία κινδύνου, τη θεωρία αξιοπιστίας και τέλος τα οικονομικά. Στο κεφάλαιο αυτό, γίνεται μια εκτενής αναφορά στις ιδιότητες τους, τα διάφορα αποτελέσματα που έχουν προκύψει μέχρι σήμερα για αυτές.

2.1 Κατανομές Phase-type

Ο κύριος λόγος που οι κατανομές Phase-type (*PH*) είναι τόσο δημοφιλείς είναι επειδή η κλάση τους περιέχει πολλές κατανομές με στήριγμα το σύνολο των μη-αρνητικών πραγματικών αριθμών. Αυτό αποτελεί βασικό πλεονέκτημα καθώς κάθε μία εξ αυτών μπορεί να προσεγγιστεί ικανοποιητικά όπως αναφέρει και ο Nelson (1995), από οποιαδήποτε ζημιοκατανομή.

Οι κατανομές *PH* περιγράφουν τον χρόνο που χρειάζεται για να φτάσει στην κατάσταση απορρόφησης 0 μία ομογενής Μαρκοβιανή διαδικασία με πεπερασμένες στο πλήθος και διακεκριμένες καταστάσεις. Φυσικά, μία τέτοια κατανομή μπορεί να είναι είτε διακριτού είτε συνεχούς χρόνου. Θεωρούμε πως τα βήματα της παραπάνω στοχαστικής διαδικασίας, οι λεγόμενες “φάσεις”, είναι οι καταστάσεις της μαρκοβιανής στοχαστικής διαδικασίας.

Η βασική ιδέα δόθηκε στις αρχές του προηγούμενου αιώνα από τον Erlang (1917), καθώς ήταν ο πρώτος που προσπάθησε να μοντελοποιήσει τη συμπεριφορά μη-αρνητικών τυχαίων μεταβλητών. Για λόγους ιστορικούς αναφέρουμε ότι ο Erlang (1917) γενίκευσε την εκθετική κατανομή κατασκευάζοντας μία μέθοδο, γνωστή ως “μέθοδος των φάσεων”. Αυτό που πρότεινε τότε ήταν να αντιληφθούμε τον τυχαίο χρόνο αναμονής σαν να δημιουργείται από πλήθος (τυχαίων ή μη-τυχαίων) εκθετικά κατανεμημένων βημάτων και αξιοποιώντας τη μαρκοβιανή αλυσίδα που παράγεται, να απλοποιήσουμε την ανάλυσή του. Παρότι η πρώτη ιδέα για την χρήση κατανομών *PH* φαίνεται να έχει αναφερθεί από τον Erlang, η συστηματική θεμελίωση και μελέτη τους έγινε από τον Neuts (1975).

Στην παρούσα εργασία ενδιαφερόμαστε για τον τρόπο συμπεριφοράς μιας απαριθμήτριας τ.μ $N(t)$ έτσι όπως αυτή ορίστηκε στο πρώτο κεφάλαιο. Έτσι, στις επόμενες ενότητες παρουσιάζονται ιδιότητες του τυχαίου αθροίσματος αλλά και αποτελέσματα ήδη αποδεδειγμένα για την περίπτωση όπου η $N(t)$ ακολουθεί μία κατανομή *PH* και συγκεκριμένα μία διακριτή κατανομή *PH.*

Αξίζει να ειπωθεί πως ο Hipp (2004) μελέτησε εκτενώς το τυχαίο άθροισμα για την περίπτωση όπου η τ.μ. που μετράει το μέγεθος ενός κινδύνου ή διαφορετικά τη σφοδρότητά του, να ακολουθεί κατανομή *PH*, κάνοντας την υπόθεση πως η απαριθμήτρια $N(t)$ ή διαφορετικά η συχνότητα εμφάνισης του κινδύνου ανήκε στην οικογένεια Panjer. Απέδειξε ότι ο τύπος του Panjer (1.13) έχει τοπικό χαρακτήρα. Δηλαδή, για τις σύνθετες κατανομές τέτοιων τ.α υπάρχει κάποιο επίπεδο x τέτοιο ώστε μόνο οι πιθανότητες $f_S(x-u)$ για $u=1,\ldots,min(d,x)$ γύρω από το x παίζουν ρόλο. Η ποσότητα d είναι η διάσταση της *PH* κατανομής ή αλλιώς το συνολικό πλήθος των σταδίων της.

Τέλος, στα πλαίσια της στατιστικής συμπερασματολογίας για τις κατανομές *PH*, ο Assmusen (1996) πρώτος έδωσε αποτελέσματα για εκτιμητή πιθανοφάνειας για τις *PH*. Στην θεωρία κινδύνου σημαντικό έργο μπορεί να συναντήσει κανείς ξανά στην εργασία του Assmussen (2000).

2.2 Διακριτές κατανομές phase-type

Μια διακριτή κατανομή *PH* (*DPH*) όπως αναφέρει ο Nielsen (2012) περιγράφει τον χρόνο που απαιτείται μέχρι να φτάσει στο τελικό της βήμα (απορροφητικό στάδιο) μια ομογενής μαρκοβιανή αλυσίδα. Η κατανομή *PH* μπορεί να είναι συνεχής στο διάστημα $(0,+\infty)$ ή διακριτή με τιμές από το σύνολο των φυσικών αριθμών $N_0=\{1,2,\ldots\}$.

Στην περίπτωση των DPH έχουμε μία ομογενή μαρκοβιανή αλυσίδα $(X(t))_{t\in N_0}$ με πλήθος βημάτων από το σύνολο $D=\{0,1,\ldots,d\}$. Ανάλογα με το πλήθος των βημάτων της αλυσίδας προσδιορίζεται και η διάσταση της κατανομής *PH*. Ο πίνακας της παραπάνω αλυσίδας περιέχει όλες τις πιθανότητες μετάβασής της από το ένα στάδιο στο επόμενο. Ο πίνακας αυτός καλείται πίνακας μετάβασης *P*, είναι διάστασης d και περιέχει στοιχεία της μορφής $P=(p_{jk})_{0\leq j,k\leq d}$. Για να οριστεί πλήρως μία διακριτή κατανομή *PH* χρειάζεται ακόμα ένα αρχικό διάνυσμα $\boldsymbol{\pi}=(\pi_{0,}\pi_{1,}\ldots,\pi_d)'=(\pi_{0,}\boldsymbol{\pi})$ με $\pi_0=1-\sum_{k=1}^{d}\pi_k$ να αναφέρεται ως η πιθανότητα μετάβασης στην κατάσταση απορρόφησης της αλυσίδας, όπου εμείς θα θεωρούμε ότι θα είναι πάντα το 0. Επίσης, οι πιθανότητες μετάβασης p_{j0} στο τελικό στάδιο της αλυσίδας από κάθε άλλο στάδιο δίνονται από την σχέση $p_{j0}=1-\sum_{k=1}^{d}p_{jk}$. Συμβολικά γράφουμε ότι μία τ.μ. $T\sim PH_d(\boldsymbol{\pi},P)$.

Σύμφωνα με τον Eisele (2006) η χρήση του θεωρήματος Perron-Frobenius αποδεικνύει ότι ο πίνακας μετάβασης *P* έχει την μεγαλύτερη θετική, πραγματική ιδιοτιμή του απλή[1] και ίση με την μονάδα $\lambda_0=1$ και ενώ ο υποπίνακας $P=(p_{jk})_{1\leq j,k\leq d}$ έχει την μεγαλύτερη θετική και πραγματική του ιδιοτιμή μικρότερη της μονάδας, $(\lambda_1<1)$.

Όπως είπαμε και παραπάνω μία τ.μ. με κατανομή *PH* ορίζεται ως ο "χρόνος ζωής" μιας σ.δ $X(t)$. Αν συμβολίσουμε με *T* μια τέτοια τ.μ. τότε η *T* θα είναι ίση με $T=min\{t\in N_0, X(t)=0\}$. Η κατάσταση 0 όπως χαρακτηριστικά αναφέρει ο Eisele (2006) είναι το "νεκροταφείο" της σ.δ $X(t)$. Εφόσον η τ.μ *T* ακολουθεί μία διακριτή κατανομή *PH* χρειάζονται τρεις πληροφορίες για να οριστεί πλήρως, έναν "χώρο καταστάσεων" *D*, ένα πίνακα μετάβασης *P* και ένα αρχικό διάνυσμα $\boldsymbol{\pi}$. Η τριάδα $(D,P,\boldsymbol{\pi})$ χαρακτηρίζει την υποθεώρηση μαρκοβιανή αλυσίδα $X(t)$ και λέμε ότι αναπαριστά πλήρως μία διακριτή κατανομή *PH*.

Η συνάρτηση πιθανότητας $P_T(t)=P(T=t)$ της τ.μ *T* θα είναι στο σημείο 0 ίση με $P_T(0)=P(T=0)=\boldsymbol{\pi_0}$. Βασική σχέση για τον υπολογισμό του P^t δίνεται από την ισότητα $P^t=(p_{jk}^{(t)})$, η οποία μπορεί να αποδειχθεί εύκολα είτε με επαγωγή είτε αριθμητικά με πράξεις πινάκων, και ισχύει ότι $p_{jk}^t=P(x(t)=k|X(0)=l)$ έχουμε ότι η πιθανότητα η σ.δ

[1]Απλή ιδιοτιμή είναι η ιδιοτιμή γεωμετρικής πολλαπλότητας 1.

$X(t)$ ξεπεράσει τα $t+1$ στάδια, άρα κατ' επέκταση η τ.μ T να μετρήσει "χρόνο ζωής " περισσότερο από $t+1$ στάδια θα δίνεται από την σχέση:

$$P_T(t)=(\sigma\tau\acute{\alpha}\delta\iota\alpha\, \tau\eta\varsigma\, X(t)\geq t+1)=P_T(\{s\in N_{0,}\, s\geq t+1\})=\boldsymbol{\pi}'\, P^t\, \boldsymbol{e} \tag{2.1}$$

όπου το $\boldsymbol{e}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}\in R^d$ είναι ένα διάνυσμα-στήλη για το οποίο ισχύει ότι $P\,\boldsymbol{e}=\boldsymbol{e}$.

Επομένως, για $t\geq 1$, η (2.1) μετασχηματίζεται στην σχέση:

$$P_T(t)=\boldsymbol{\pi}'\, P^{t-1}(I-P)\boldsymbol{e} \tag{2.2}$$

Η συνάρτηση κατανομής μια διακριτής PH , σύμφωνα με τα παραπάνω θα δίνεται από την σχέση:

$$P_T(t)(\{s\in N_{0,}\, s\leq t\})=P_T(t)=1-\boldsymbol{\pi}'\, P^{t-1}\,\boldsymbol{e}\ . \tag{2.3}$$

Παράδειγμα 2.1: Το πιο απλό παράδειγμα διακριτής κατανομής *PH* είναι η περίπτωση που ο πίνακας *P* είναι διάστασης *d*=1. Τότε θα δημιουργηθεί ένας 2x2 πίνακας μετάβασης με:

$$P=\begin{pmatrix}p & 1-p\\ 0 & 1\end{pmatrix} \quad \text{και} \quad \boldsymbol{\pi}=(1)\ .$$

Επί της ουσίας, μια κατανομή *PH* με τέτοια αναπαράσταση είναι μία απλή γεωμετρική κατανομή με παράμετρο *p*.

2.3 Γεννήτριες Κατανομών Phase-type

Από τις γεννήτριες συναρτήσεις κατανομών μπορούμε να παράγουμε όλες τις ροπές μίας κατανομής, κάνοντας διαδοχικές παραγωγίσεις στην γεννήτρια συνάρτηση. Η συνάρτηση $M_X(t)=E(e^{tx})\forall t\in(-\delta,\delta),\delta>0$ έχει την ικανότητα να παράγει οποιαδήποτε ροπή αντιστοιχεί στην κατανομή της τ.μ. *X* και να την προσδιορίζει μονοσήμαντα.

Όπως αναφέρεται στον Hipp (2004), για μία διακριτή τ.μ. που ακολουθεί κατανομή *PH,* η γεννήτρια συνάρτησή της δίνεται από την σχέση:

$$
\begin{aligned}
\Phi_T(z)=E(z^T)&=\boldsymbol{\pi}_{\mathbf{0}}+\sum_{t=1}^{\infty} z^t(\boldsymbol{\pi}' P^{t-1}(I-P)\boldsymbol{e}) \\
&=\boldsymbol{\pi}_{\mathbf{0}}+\boldsymbol{\pi}'(I z^{-1}-P)^{-1}(I-P)\boldsymbol{e}
\end{aligned}, \tag{2.4}
$$

όπου $z\in C\, \mu\varepsilon\, |z|<\frac{1}{\lambda_1}$ και $\lambda_1<1$ η μεγαλύτερη ιδιοτιμή του υποπίνακα $P=(p_{jk})_{1\le j,k\le d}$.

Για να υπολογιστεί η πιθανογεννήτρια που δίνει η σχέση (2.4), εκτός από ένα πίνακα μετάβασης P και το αρχικό διάνυσμα $\boldsymbol{\pi}$, απαιτείται και ο υπολογισμός του αντίστροφου πίνακα $(I z^{-1}-P)^{-1}$. Χρησιμοποιώντας τον κανόνα του Cramer για την εύρεση των στοιχείων του αντιστρόφου, δηλαδή τον τύπο $A_{kl}^{-1}=(-1)^{k+l}\frac{det(A(l,k))}{det(A)}$, με A ένας αντιστρέψιμος πίνακας και $A(l,k)$ ο πίνακας A, χωρίς την l- γραμμή και την k- στήλη, μπορούμε να βρούμε όλα τα στοιχεία του πίνακα $(I z^{-1}-P)^{-1}$ σύμφωνα με την παρακάτω σχέση:

$$
\begin{aligned}
(I z^{-1}-P)_{kl}^{-1}&=(-1)^{k+l}\frac{z^{-(d-1)}+a_2(k,l)z^{-(d-2)}+..+a_d(k,l)}{det(I z^{-1}-P)} \\
&=(-1)^{k+l}\frac{z+a_2(k,l)z^2+\ldots+a_d(k,l)z^d}{1+b_1 z^1+b_2 z^2+\ldots+b_d z^d}
\end{aligned} . \tag{2.5}
$$

Οι συντελεστές b_i με $i=1,\ldots,d$ που εμφανίζονται στον παρονομαστή της σχέσης (2.5) είναι οι συντελεστές του χαρακτηριστικού πολυωνύμου του πίνακα $(I z^{-1}-P)$ και δίνονται από την σχέση:

$$
b_i=(-1)^i\sum_{1\le k_1\le\ldots\le k_{d-i}\le d} det(P(k_1,k_2,\ldots,k_{d-i})). \tag{2.6}
$$

Συγκεκριμένα, $b_1=(-1)^1 trace(P), b_d=(-1)^d det(P)$.

Επομένως, η πιθανογεννήτρια συνάρτηση $\Phi_T(z)$, όπως αναφέρεται στον Eisele (2006) μπορεί να δοθεί από την σχέση:

$$
\Phi_T(z)=\pi_0+\frac{a_1 z+\ldots+a_d z^d}{1+b_1 z+\ldots+b_d z^d} . \tag{2.7}
$$

Εφόσον οι συντελεστές b_i υπολογίζονται βάσει της σχέσης (2.6), μπορούμε μέσω αυτών να παράγουμε και τους συντελεστές a_i της σχέσης (2.7) χρησιμοποιώντας αναδρομικό τύπο για τον υπολογισμό τους. Στην σχέση (2.7) έχουμε ότι:

$$
a_1=P_T(1) \tag{2.8}
$$

και για κάθε $i=2,\ldots,d$ οι συντελεστές a_i μπορούν να υπολογιστούν αναδρομικά από την σχέση:

$$a_i = P_T(i) + \sum_{l=1}^{i-1} b_l P_T(i-l). \tag{2.9}$$

Ενδεικτικά, για $i=1$ έχουμε: $a_1 = P_T(1)$ και για $i=2$ έχουμε:

$$a_2 = P_T(2) + b_1 P_T(1) = P_T(2) + det(P) P_T(1) = P_T(2) + det(P) a_1 .$$

Ο αναδρομικός τύπος (2.9) υπολογίζει του συντελεστές a_i για κάθε ένα στοιχείο του πίνακα $(I z^{-1} - P)^{-1}$ και αποδεικνύεται σχετικά εύκολα μέσω αλγεβρικών πράξεων στην σχέση (2.7).

Συγκεκριμένα,

$$\Phi_T(z) = \pi_0 + \sum_{i=1}^{d} a_i z^i + \sum_{i=1}^{d} b_i z^i (\pi_0 - \Phi_T(z)).$$

Επομένως, εφόσον γνωρίζουμε πως $P_T(0) = \pi_0$ θα έχουμε τελικά την σχέση:

$$\begin{aligned}\sum_{t\geq 1} P_T(t) z^t &= \sum_{i=1}^{d} a_i z^i - \sum_{i=1}^{d} b_i z^i \left(\sum_{t\geq 1} P_T(t) z^t\right) \\ &= \sum_{i=1}^{d} a_i z^i - \sum_{t\geq 2} z^t \left(\sum_{i=1}^{min(d,t-1)} b_i P_T(t-i)\right).\end{aligned} \tag{2.10}$$

Τέλος, η ροπή k -τάξης γύρω από το σημείο $t=1$ μίας διακριτής κατανομής PH δίνεται από τον τύπο:

$$E(X(X-1)\cdots(X-(k-1))) = \frac{d^k \Phi_T(z)}{d z^k} = k! \boldsymbol{\pi}' (I-P)^{-k} P^{k-1} \boldsymbol{e}. \tag{2.11}$$

Παράδειγμα 2.3: Ας πάρουμε ως χώρο καταστάσεων μιας σ.δ $X(t)$ τον *D*={0,1}, ένα αρχικό διάνυσμα $\boldsymbol{\pi} = (\pi_{0,} 1-\pi_0)$ και έναν πίνακα μετάβασης τον $P = \begin{pmatrix} 1 & 0 \\ 1-p & p \end{pmatrix}$. Από την (2.6) έχουμε πως ο συντελεστής $b_1 = -p$. Η συνάρτηση πιθανότητας μιας τ.μ. *T* που ακολουθεί κατανομή *PH* που αναπαρίσταται από την τριάδα $(D, P, \boldsymbol{\pi})$ θα είναι :

$$P_T(t) = (1-\pi_0)(1-p) p^{t-1}, t\geq 1 , \qquad P_T(t) = \pi_0,\ t=0 .$$

Επομένως, από την σχέση (2.8) έχουμε πως ο συντελεστής $a_1 = (1-\pi_0)(1-p)$ και από την (2.7) η γεννήτρια συνάρτηση της τ.μ *T* θα είναι: $\Phi_T(z) = \pi_0 + \dfrac{(1-\pi_0)(1-p) z}{1-p z}$. Μία γεωμετρική κατανομή που ξεκινάει από το μηδέν, μπορεί να βρεθεί αρκεί να θέσουμε

$\pi_0 = 1 - p$ οπότε και μπορεί να γίνει υπολογισμός της συνάρτησης πιθανότητας και του γνωστού από τη θεωρία πιθανοτήτων τύπο για την πιθανογεννήτρια συνάρτηση:

$$P_T(t) = (1-p)p^t, \Phi_T(z) = \frac{1-p}{1-zp}.$$

2.4 Κλειστές πράξεις μεταξύ κατανομών Phase-type

Παρά την ευκολία στον χειρισμό τους, οι κατανομές *PH* μπορούν να δημιουργήσουν κάποια προβλήματα όταν χρησιμοποιούμε τα αθροίσματά τους. Παραδείγματος χάριν, ένα τέτοιο πρόβλημα, όπως αναφέρει ο Nielson (2012), είναι το γεγονός ότι δεν υπάρχει, στις περισσότερες των περιπτώσεων, μοναδική αναπαράσταση των κατανομών *PH*.

Όπως έχουμε ήδη αναφέρει υπάρχουν συναρτήσεις πιθανότητας που δεν ικανοποιούν ορισμένες συνθήκες ομαλότητας. Σε αυτή την κατηγορία ανήκουν και οι συναρτήσεις πιθανότητας των διακριτών τ.μ. Phase-type. Αυτό αποτελεί βασικό μειονέκτημα των κατανομών *PH*, ειδικά όταν μας ενδιαφέρει το άθροισμα τους.

Από την άλλη πλευρά, η κλάση των κατανομών *PH* έχει την χαρακτηριστική ιδιότητα να είναι κλειστή ως προς συγκεκριμένες πράξεις. Αυτό ουσιαστικά σημαίνει πως υπό τις πράξεις αυτές μπορούν να παραχθούν ξανά κατανομές *PH*.

Συγκεκριμένα, έχει αποδειχθεί ότι η κλάση τους είναι κλειστή ως προς την πρόσθεση τ.μ. με κατανομή *PH*, την πρόσθεση τυχαίου πλήθους τ.μ. με κατανομή *PH*, τη μείξη κατανομών *PH* και το πεπερασμένο διατεταγμένο δείγμα (order statistics) κατανομών *PH*.

Αυτός είναι και ο κυριότερος λόγος που η οικογένεια αυτή είναι τόσο δημοφιλής στην μοντελοποίηση στοχαστικών προβλημάτων.

Ας υποθέσουμε ότι ενδιαφερόμαστε για το άθροισμα δύο ανεξάρτητων, διακριτών τ.μ. X, Y που ακολουθούν κάποια κατανομή *PH*. Άρα οι τ.μ. θα αναπαρίστανται από τις δυάδες $(\boldsymbol{\alpha}, P)$, $(\boldsymbol{\beta}, S)$ αντίστοιχα, με κοινό χώρο καταστάσεων των υπό θεώρηση σ.δ τους, *D*. Τότε, το άθροισμά τους θα είναι μία νέα τ.μ. $Z = X + Y$ η οποία όπως αναφέραμε, λόγω κλειστότητας, θα ακολουθεί και αυτή μία διακριτή κατανομή *PH*. Η τ.μ. *Z* θα αντιπροσωπεύεται από την δυάδα $(\boldsymbol{\gamma}, L)$ και το σύνολο *D*. Προκειμένου να παραχθεί η τ.μ. *Z* ξεκινάμε από την μαρκοβιανή αλυσίδα που παράγει η τ.μ. *X* και μόλις αυτή φτάσει στη μοναδική ηατάσταση απορρόφησης συνεχίζουμε με την παραγόμενη από την *Y* μαρκοβιανή αλυσίδα. Οπότε δημιουργείται ο στοχαστικός πίνακας μετάβασης της *Z,* ο οποίος είναι το γινόμενο Kronecker δύο πινάκων:

$$\begin{pmatrix} L & \boldsymbol{L^0} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} P_{11} & \cdots & P_{1m} & P_1^0\beta_1 & \cdots & P_1^0\beta_k & P_1^0\beta_{k+1} \\ P_{21} & \cdots & P_{2m} & P_2^0\beta_1 & \cdots & P_2^0\beta_k & P_2^0\beta_{k+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{m1} & \cdots & P_{mm} & P_m^0\beta_1 & \cdots & P_m^0\beta_k & P_1^0\beta_{k+1} \\ 0 & \cdots & 0 & S_{11} & \cdots & S_{1k} & S_1^0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & S_{k1} & \cdots & S_{kk} & S_k^0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} P & P^0\boldsymbol{\beta} & \beta_{k+1}P^0 \\ 0 & S & S^0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.12)$$

με αρχικό διάνυσμα πιθανοτήτων,

$$\boldsymbol{\gamma} = (\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}\beta_1, \alpha_{m+1}\beta_2, \ldots, \alpha_{m+1}\beta_k) = (\boldsymbol{\alpha}, \alpha_{m+1}\boldsymbol{\beta})$$

και

$$\gamma_{m+k+1} = a_{m+1}\beta_{k+1}\,.$$

Το διάνυσμα $\boldsymbol{L^0}$ δίνεται από την σχέση $\boldsymbol{L^0} = (I - L)\boldsymbol{e}$.

Το γινόμενο Kronecker μεταξύ δύο πινάκων $A = (a_{jk})_{n\,x\,m}$, B ορίζεται ως ένας πίνακας της μορφής:

$$A \times B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{pmatrix}.$$

Εναλλακτικά, μπορούμε να καταλήξουμε στο ίδιο αποτέλεσμα μέσω γεννητριών συναρτήσεων:

$$\Phi_Z(z) = \Phi_X(z)\Phi_Y(z) = (\alpha_{m+1} + z\boldsymbol{\alpha}(I - zP)^{-1}P^0)(\beta_{k+1} + z\boldsymbol{\beta}(I - zS)^{-1}S^0)\,, \quad (2.13)$$

$$\Phi_Z(z) = \gamma_{m+1} + z\boldsymbol{\gamma}(I - zL)^{-1}(I - L)\boldsymbol{e}. \quad (2.14)$$

Το άθροισμα $X + Y$ δεν είναι γενικά συμμετρικό ως προς τις παραμέτρους των *X*,*Y* δηλαδή παίζει ρόλο ποια είναι η αλυσίδα που θα θέσουμε ως *X* και ποια ως *Y* και αυτό γιατί ο πίνακας μετάβασης *L* θα είναι διαφορετικός. Τα $F(x)$, $f(x)$ και $\Phi(z)$ όπως αναφέρει ο Nielson (2012) θα παραμείνουν τα ίδια.

Εκτός από την πρόσθεση, και η μείξη κατανομών της οικογένειας *PH* θα παράγει μία νέα τ.μ με κατανομή από την κλάση των *PH.* Σύμφωνα με τον Neuts (1981), αν $f_N(u)$ και $f_X(t)$ είναι συναρτήσεις πιθανότητας δύο διακριτών κατανομών *PH* που ορίζονται από τις δυάδες $(\boldsymbol{\pi}, \Lambda)$, $(\boldsymbol{\rho}, M)$ αντίστοιχα και είναι διάστασης *d* και *c* αντίστοιχα, τότε η μείξη τους μπορεί να υπολογιστεί από σχέση της μορφής:

$$\sum_{u=0}^{\infty} f_N(u) f_X^{*u}(t), \tag{2.15}$$

και θα είναι και αυτή μία διακριτή κατανομή *PH* διάστασης cd. Ο πίνακας μετάβασης απαιτεί τον υπολογισμό του γινομένου Kronecker των δύο επιμέρους πινάκων *Λ*, *Μ* και δίνεται από την σχέση:

$$\Sigma = M \times I_d + u\, p' \times (I_d - a\, \Lambda)^{-1} \Lambda, \tag{2.16}$$

όπου τα α, u ισούνται με $\alpha = 1 - \boldsymbol{\rho}'\mathbf{1}$ και $u = (I_c - M)\mathbf{1}$.

Το αρχικό διάνυσμα υπολογίζεται από τον τύπο:

$$\sigma = \boldsymbol{\rho} \times \boldsymbol{\pi} (I_d - a\, \Lambda)^{-1}. \tag{2.17}$$

Πέρα από τις παραπάνω πράξεις και το διατεταγμένο δείγμα κατανομών Phase-type είναι κλειστή πράξη.

Τέλος, αν ενδιαφερόμαστε για την κατανομή της μεγαλύτερης ή της μικρότερης από δύο ανεξάρτητες τ.μ. *X*,*Y* που ακολουθούν κατανομή *PH*, μπορούμε να ορίσουμε την τ.μ. $Z = min(X, Y)$ και να δημιουργήσουμε μία μαρκοβιανή αλυσίδα που να περιγράφει την ταυτόχρονη εξέλιξη των τ.μ. *X*,*Y* και άρα της *Z*. Η *Z* θα ανήκει κι αυτή στην κλάση των κατανομών *PH* και η μαρκοβιανή αλυσίδα που θα παράγει, ουσιαστικά θα μετράει τον χρόνο αναμονής μέχρι μία από τις αλυσίδες των *X*,*Y* θα φτάσει πρώτη στην κατάσταση απορρόφησης της. Αντίστοιχα αν ορίσουμε την τ.μ. $Z' = max(X, Y)$.

Παράδειγμα 2.4: Το άθροισμα από γεωμετρικά κατανεμημένες τ.μ. γνωρίζουμε ότι ακολουθεί αρνητική διωνυμική κατανομή. Η αρνητική διωνυμική κατανομή μπορεί να οριστεί με τύπους μιας κατανομής *PH*, με *m*= πλήθος των τ.μ. του αθροίσματος. 'Ετσι, θα προκύψει ένας *mxm* πίνακας μετάβασης:

$$P = \begin{pmatrix} p & 1-p & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & p & 1-p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & p & 1-p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & p & 1-p & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & p & 1-p \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \quad \text{και} \quad \boldsymbol{a} = (1, 0, \ldots, 0). \tag{2.18}$$

Παράδειγμα 2.5: Έστω δύο Phase-type τ.μ. *X*~ *NB*(2, p_x) και *Y*~ *NB*(2, p_y) με πίνακες $T_x=\begin{pmatrix} p_x & 1-p_x \\ 0 & p_x \end{pmatrix}$ και $T_y=\begin{pmatrix} p_y & 1-p_y \\ 0 & p_y \end{pmatrix}$ αντίστοιχα και διανύσματα $\boldsymbol{a}_{\boldsymbol{x}}=(1,0)$, $\boldsymbol{a}_y=(1,0)$ αντίστοιχα. Για να μπορέσουμε να βρούμε πως αναπαριστάται η νέα τ.μ $Z=min(X,Y)$ πρέπει να δημιουργήσουμε μία νέα στοχαστική διαδικασία που θα περιγράφει την ταυτόχρονη εξέλιξη των *X*,*Y*. Η διαδικασία αυτή θα έχει τέσσερα δυνατά στάδια, τα οποία θα είναι όλοι οι πιθανοί συνδυασμοί μεταξύ των σταδίων στα οποία μπορούν να βρεθούν οι τ.μ. *X*,*Y*, δηλαδή οι συνδυασμοί (1,1) , (1,2) , (2,1) , (2,2). Στα στάδια (1,1) , (1,2) έχουμε την διαδικασία *X* να παραμένει στο ίδιο στάδιο με πιθανότητα p_x, ενώ την διαδικασία *Y* να μεταπηδά σε άλλο στάδιο με πιθανότητα $1-p_y$.

Η νέα διαδικασία *Z* όπως είπαμε θα περιγράφει την χρόνο αναμονής μέχρι μία από τις *X*,*Y* να φτάσει πρώτη στην κατάσταση απορρόφησης της.

Επομένως, παίρνουμε τον 4x4 πίνακα:

$$T_{min}(X,Y)=\begin{pmatrix} p_x p_y & p_x(1-p_y) & (1-p_x)p_y & (1-p_x)(1-p_y) \\ 0 & p_x p_y & 0 & (1-p_x)p_y \\ 0 & 0 & p_x p_y & p_x(1-p_y) \\ 0 & 0 & 0 & p_x p_y \end{pmatrix},$$

και το διάνυσμα πιθανοτήτων:

$$\boldsymbol{a}_{\boldsymbol{min(X,Y)}}=(1,0,0,0).$$

Δηλαδή παρατηρούμε πως $T_{min(X,Y)}=T_x \times T_y$ και το αρχικό διάνυσμα για την διαδικασία *Z* προκύπτει ως $a_{min(X,Y)}=\alpha_x \times a_y$.

2.5 Αναδρομική σχέση για σύνθετη κατανομή Phase-type

Όπως αναφέρεται στον Eisele (2006) η σύνθετη κατανομή ενός τυχαίου αθροίσματος τυχαίων μεταβλητών με την συχνότητα εμφάνισης των τυχαίων μεταβλητών να ακολουθεί μία διακριτή *PH* και οι $(Y_i)_{i\geq 1}$ να είναι ανεξάρτητες και ισόνομες τ.μ διακριτές στο $\{1,2,3,\ldots\}$ ή συνεχείς στο $(0,\infty)$ με σ.π ή σ.π.π P_Y αντιστοιχεί στο τυχαίο άθροισμα:

$$S=\sum_{i=1}^{T} Y_i, \tag{2.19}$$

όπου η $T \sim DPH$ που ορίζεται από την τριάδα $(D, P, \boldsymbol{\pi})$.

Ας σκεφτούμε μία ασφαλιστική επιχείρηση που ενδιαφέρεται να υπολογίσει ποια είναι η πιθανότητα την επόμενη χρονική περίοδο να πληρώσει ως συνολική αποζημίωση από ζημίες ενός χαρτοφυλακίου της, ένα συγκεκριμένο ποσό *A*. Αν η ασφαλιστική γνωρίζει πως η συχνότητα με την οποία δημιουργούνται οι ζημιές από τα συμβόλαια αυτά ακολουθεί ένα μοτίβο που περιγράφεται ικανοποιητικά από την σ.κ μιας διακριτής κατανομής *PH*, τότε μπορεί να υπολογίσει την πιθανότητα να πληρώσει το ποσό *A* σε κάθε χρονική στιγμή της επόμενης περιόδου.

Ας πάρουμε την περίπτωση για την ασφαλιστική αυτή όπου τα ύψη των επιμέρους ζημιών είναι διακριτά και ανεξάρτητα μεταξύ τους, οπότε οι τ.μ. $(Y_i)_{i\geq 1}$ υποθέτουμε πως είναι διακριτές. Τότε, προκύπτει ότι το άθροισμα S θα ακολουθεί και αυτό μία διακριτή κατανομή η οποία ονομάζεται σύνθετη κατανομή phase-type.

Για την περίπτωση αυτή, ο Eisele (2006) έδωσε αναδρομική σχέση υπολογισμού της συνάρτησης πιθανότητας. Όπως και στην οικογένεια Panjer, η πιθανότητα το παραπάνω τυχαίο άθροισμα να πάρει την τιμή μηδέν, δηλαδή να μην δημιουργηθεί καμία ζημιά από τα συμβόλαια την ασφαλιστικής, είναι $P_S(0) = P_T(0) = \pi_0$ ενώ για κάθε χρονική στιγμή $t \geq 1$ αποδεικνύεται πως ισχύει η αναδρομική σχέση:

$$P_S(t) = \sum_{j=1}^{\min(d,t)} a_j P_Y^{*j}(t) - \sum_{j=1}^{\min(d,t-1)} b_j \left(\sum_{u=1}^{n-1} P_S(u) P_Y^{*j}(t-u) \right), t \geq 1. \tag{2.20}$$

Οι συντελεστές a_j και b_j δίνονται από τους τύπους (2.6) και (2.9). Η σχέση που υπολογίζει την πρώτη συνέλιξη είναι $p^{*1}(t) = P_Y(t)$ ενώ για τις υπόλοιπες k-τάξης συνελίξεις που χρειάζεται η σχέση (2.17), ισχύει αναδρομικός τύπος:

$$P^{*k}(t) = \sum_{l=k-1}^{n-1} P^{*k-1}(l) P_Y(t-l). \tag{2.21}$$

ΚΕΦΑΛΑΙΟ 3

Εσωτερικά Μοντέλα Ελέγχου στην διαχείριση κινδύνου

Στο κεφάλαιο αυτό δίνεται μία σύντομη περιγραφή των υφιστάμενων νομικών πλαισίων που εφαρμόζονται στον ευρωπαϊκό ασφαλιστικό χώρο και τον τραπεζικό κλάδο σε παγκόσμιο επίπεδο. Γίνεται μια ανασκόπηση για το πως τα πλαίσια αυτά αντιλαμβάνονται την διαχείριση των κινδύνων και πως η αντίληψη αυτή ενσωματώνεται στους υπολογισμούς των απαιτούμενων ιδίων κεφαλαίων των επιχειρήσεων αυτών. Τέλος, γίνεται μία εισαγωγή σε δύο νέα μοντέλα διαχείρισης του κινδύνου χρεοκοπίας υπό το καθεστώς Βασιλείας ΙΙΙ αλλά και μία εφαρμογή αυτών στην διαχείρισης του ασφαλιστικού κινδύνου ενός χαρτοφυλακίου.

3.1 Διαχείριση των κινδύνων υπό το καθεστώς της Βασιλείας ΙΙΙ

Η διαχείριση των κινδύνων στις μέρες μας είναι ένας γρήγορα αναπτυσσόμενος κλάδος και ο τρόπος με τον οποίο οφείλει να γίνεται έχει εγείρει πολλές και διαφορετικές απόψεις. Σειρά διαφόρων τεχνικών χρησιμοποιούνται ήδη για την ανάλυση και παρακολούθηση των κινδύνων που απειλούν μία επιχείρηση. Τέτοιες μπορεί να είναι είτε ειδικές για κάθε έναν κίνδυνο είτε ολιστικές, δηλαδή να παρέχουν την δυνατότητα διαχείρισης του συνολικού κινδύνου της εταιρείας ή κάποιου χαρτοφυλακίου της, συνολικά.

Ο χειρισμός του κινδύνου είναι μία διαδικασία εφαρμογής μέτρων και στρατηγικών για να εξαλειφθεί ο κίνδυνος. Το πιο σημαντικό στοιχείο στον χειρισμό αυτό είναι ο έλεγχος και κατ' επέκταση η μείωσή του ή ακόμα και η αποφυγή του.

Όλα αυτά έχουν σαν αποτέλεσμα την ανάπτυξη μέτρων ελέγχου του κινδύνου στο εσωτερικό της επιχείρησης. Πλέον για τα χρηματοπιστωτικά ιδρύματα και τις ασφαλιστικές εταιρείες προτεραιότητα αποτελεί η δημιουργία μεθόδων και τεχνικών ανάλυσης του κινδύνου, όπου κάθε μία εξ' αυτών θα εξυπηρετεί και διαφορετικό σκοπό.

Τα μοντέλα που αναλύονται στα επόμενα κεφάλαια αποτελούν ένα εσωτερικό μέτρο ελέγχου που επικεντρώνεται στην ανάλυση απειλής (threat analysis) μίας δυνητικής χρεοκοπίας, κάνοντας μία προσπάθεια να μοντελοποιηθεί ο χρόνος αναμονής μέχρι να ικανοποιηθεί ένα προεπιλεγμένο από την εταιρεία κριτήριο.

Έστω ότι έχουμε μία τράπεζα η οποία λειτουργεί και αναπτύσσεται υπό το καθεστώς των κανόνων της Βασιλείας ΙΙΙ[2]. Το ρυθμιστικό πλαίσιο αυτό για την διαχείριση των κινδύνων, απαιτεί οι τράπεζες να υπολογίζουν και να διατηρούν συγκεκριμένη κεφαλαιακή επάρκεια καθώς επίσης και να την αναφέρουν στις εποπτικές αρχές. Τα απαιτούμενα κεφάλαια για τα οποία γίνεται λόγο στον πυλώνα Ι της Βασιλείας ΙΙΙ λειτουργούν ως εξουδετερωτής των μη αναμενόμενων ζημιών.

Ο υπολογισμός τους στηρίζεται στη συχνότητα των ζημιών και την κατανομή μεγέθους των ζημιών αυτών. Η επαρκής κεφαλαιοποίηση μιας τράπεζας είναι πολύ σημαντική, καθώς αυτή αποτελεί την κύρια γραμμή υπεράσπισης κατά αναπάντεχων, μη αναμενόμενων ζημιογόνων γεγονότων. Ακριβώς για τον ίδιο λόγο, η επαρκής ή μη κεφαλαιοποίηση της μπορεί να χρησιμοποιηθεί ως δείκτης προειδοποίησης για επικείμενα πιστωτικά γεγονότα που την αφορούν, όπως π.χ μία χρεοκοπία.

Σύμφωνα με την Βασιλεία ΙΙΙ και το πλαίσιο κανόνων της, μία τράπεζα θεωρείται επαρκώς κεφαλαιοποιημένη εάν διατηρεί ίδια κεφάλαια πάνω από ένα ελάχιστο προκαθορισμένο επίπεδο ανάλογα με τον βαθμό και την φύση των κινδύνων στους οποίους έχει εκτεθεί.

Οι εκάστοτε εποπτικές αρχές υποχρεούνται να υποβάλουν, σε τυχαίο χρόνο, μια σειρά από τεστ αντοχής (stress tests) στον τραπεζικό κλάδο με σκοπό να προσδιορίσουν την κεφαλαιακή επάρκεια. Ουσιαστικά, πρόκειται για προσομοιώσεις μελλοντικών οικονομικών γεγονότων που ως υπόθεσή τους έχουν ένα σενάριο βάσης και ένα αντίξοο οικονομικό σενάριο (base and worst-case/adverse scenarios), σχεδιασμένα έτσι ώστε να αξιολογήσουν αν μία τράπεζα ή ένα χρηματοπιστωτικό ίδρυμα έχει επαρκή κεφάλαια για να αντέξει τις επιπτώσεις δυνητικά αρνητικών εξελίξεων στην οικονομία. Τέτοιου είδους τεστ αντοχής μπορούν να διεξαχθούν συμπληρωματικά και από τις ίδιες τις τράπεζες μέσα στα πλαίσια καλύτερης παρακολούθησης του χαρτοφυλακίου τους αλλά και από ανεξάρτητες εταιρείες

[2]Βασιλεία ΙΙΙ: Διεθνές κανονιστικό πλαίσιο για την εποπτεία των τραπεζών.

που δραστηριοποιούνται στον χρηματοοικονομικό τομέα, για παράδειγμα οι οίκοι αξιολόγησης.

Τα τεστ αντοχής, λοιπόν, ως εργαλείο διαχείρισης κινδύνων στο εσωτερικό μιας τράπεζας αλλά και ως εργαλείο των εποπτικών αρχών, παρέχουν μία ένδειξη για το πόσο επιπλέον κεφάλαιο μπορεί να χρειαστεί προκειμένου να απορροφηθούν ζημιές που θα συμβούν σε περίπτωση μεγάλων ανατροπών στα οικονομικά δεδομένα.

Δηλαδή, όπως ήδη έχουμε αναφέρει, λόγω της Βασιλεία ΙΙΙ οι τράπεζες καλούνται σε τακτά χρονικά διαστήματα να προσδιορίζουν την επαρκή ή μη φερεγγυότητα των κεφαλαίων τους μέσω του υπολογισμού ενός δείκτη που ορίζεται ως ο λόγος των Ιδίων κεφαλαίων της τράπεζας προς τα σταθμισμένα βάσει κινδύνου περιουσιακά της στοιχεία (Capital Ratio). Ο δείκτης αυτός οφείλει να είναι μεγαλύτερος του 8%, προκειμένου η διάρθρωση των κεφαλαίων της τράπεζας να θεωρηθεί επαρκής.

3.2 Διαχείριση των Κινδύνων στα πλαίσια της Φερεγγυότητας ΙΙ

Το ρυθμιστικό πλαίσιο, Φερεγγυότητα ΙΙ (Solvency II), για την ορθή λειτουργία των ασφαλιστικών επιχειρήσεων που δραστηριοποιούνται στον ευρωπαϊκό χώρο, έχει θέσει στο επίκεντρο την διαχείριση των κινδύνων μιας ασφαλιστικής και καθιστά από πλευράς ασφαλιστικών εταιρειών την ανάπτυξη μοντέλων και διαδικασιών αντιμετώπισης των κινδύνων πιο επίκαιρη από ποτέ.

Η Φερεγγυότητα ΙΙ είναι σχεδιασμένη έτσι ώστε να παρέχεται στις ασφαλιστικές εταιρείες ένα ρυθμιστικό πλαίσιο σύμφωνα με το οποίο οι ασφαλιστικές μπορούν να αναγνωρίσουν, να μετρήσουν αλλά και να παρακολουθούν όλους τους κινδύνους που διατρέχουν ως επιχείρηση. Με αυτό τον τρόπο το πλαίσιο διασφαλίζει πως υπάρχουν επαρκή κεφάλαια για να καλύψουν μη αναμενόμενες ζημιές στο 99,5% των περιπτώσεων.

Η ευρωπαϊκή ρυθμιστική επιτροπή EIOPA (European Insurance and Occupational Pensions Authority) θέσπισε τον κανονισμό 2015/35 που βασίζεται κυρίως σε αρχές και κατευθυντήριες γραμμές και όχι τόσο σε κανόνες και παρέχει στις ασφαλιστικές επιχειρήσεις μία Τυποποιημένη μέθοδο υπολογισμού των κεφαλαιακών απαιτήσεων που θα πρέπει οι ασφαλιστικές να εφαρμόσουν, δίνοντας ωστόσο την δυνατότητα ανάπτυξης εσωτερικών μοντέλων πιθανοτήτων από πλευράς επιχειρήσεων που θα καλύπτουν ιδανικά όλους τους κινδύνους στα χαρτοφυλάκια της εκάστοτε ασφαλιστικής και θα λαμβάνουν υπόψη τη αλληλεπίδραση μεταξύ ορισμένων τύπων κινδύνου.

Η Φερεγγυότητα ΙΙ δίνει το έναυσμα στις ευρωπαϊκές ασφαλιστικές επιχειρήσεις να ενσωματώσουν νέες διαδικασίες διαχείρισης κινδύνου και να αναπτύξουν συστήματα ελέγχου των κινδύνων που αντιμετωπίζουν. Το γεγονός αυτό θα τις βοηθήσει να χρησιμοποιούν τα κεφάλαιά τους με ένα πιο αποδοτικό τρόπο και σε βάθος χρόνου να δια κρατούν ίσως και μικρότερα απαραίτητα κεφάλαια.

Ο κανονισμός αυτός είναι οικοδομημένος πάνω σε ένα σύστημα τριών πυλώνων. Ο πυλώνας Ι περιλαμβάνει τις ποσοτικές απαιτήσεις, δηλαδή ποσοτικοποίηση (μέτρηση) και μοντελοποίηση του κάθε κινδύνου που λαμβάνεται υπόψη για τον υπολογισμό των κεφαλαιακών απαιτήσεων. Οι κίνδυνοι που υπολογίζονται στην τυποποιημένη μέθοδο είναι ο κίνδυνος αγοράς, ο κίνδυνος αθέτησης αντισυμβαλλομένου, ο κίνδυνος ανάληψης ασφαλίσεων και ο λειτουργικός κίνδυνος της εταιρείας.

Ο πυλώνας ΙΙ περιλαμβάνει ποιοτικές απαιτήσεις, δηλαδή εποπτικές κεφαλαιακές απαιτήσεις, σύστημα διακυβέρνησης και ποιοτικές απαιτήσεις για την διαχείριση των κινδύνων (Own Risk and Solvency Assessment (ORSA) report, Business plan) και την εποπτική διαδικασία ελέγχου.

Τέλος, ο πυλώνας ΙΙΙ περιλαμβάνει τις απαιτήσεις δημοσίευσης και διαφάνειας, δηλαδή εκθέσεις που περιλαμβάνουν τις πρακτικές που εφαρμόζει η εταιρεία, τους κινδύνους που αντιμετωπίζει, σε ποιες ενέργειες έχει προβεί για να τους αντισταθμίσει και τα συγκεντρωτικά αποτελέσματα των οικονομικών καταστάσεων IFRS και Solvency II της περιόδου. Ορισμένες από αυτές δημοσιεύονται στο ευρύ κοινό προς πληροφόρησή του ενώ άλλες συντάσσονται για σκοπούς εποπτικούς και γνωστοποιούνται στην αρμόδια εποπτική αρχή.

Δύο από τις σημαντικότερες διαδικασίες πλέον μέσα σε μία ασφαλιστική επιχείρηση είναι ο προσδιορισμός της ποσότητας των κεφαλαίων εκείνων που απαιτούνται προκειμένου η ασφαλιστική επιχείρηση να έχει επαρκή κεφάλαια για το 99,5% των μη αναμενόμενων μελλοντικών απωλειών στο χαρτοφυλάκιο της, μέσα σε ένα συγκεκριμένο χρονικό διάστημα, και η ελάχιστη απαιτούμενη κεφαλαιακή απαίτηση.

Η ποσότητα των κεφαλαίων αυτών ορίζεται ως η αξία σε κίνδυνο (Value at Risk) των βασικών ιδίων κεφαλαίων μιας ασφαλιστικής σε επίπεδο εμπιστοσύνης 99,5% σε διάρκεια ενός έτους και ονομάζεται κεφαλαιακές απαιτήσεις φερεγγυότητας (Solvency Capital Requirements, SCR) που έχει η επιχείρηση έναντι των μετρήσιμων κινδύνων που έχει εκτεθεί ή πρόκειται να εκτεθεί.

Η ελάχιστη κεφαλαιακή απαίτηση (Minimum Capital Requirements, MCR) είναι συνάρτηση του SCR και αποτελεί το κατώτερο όριο των απαιτούμενων κεφαλαίων. Αποτελεί το κρίσιμο εκείνο επίπεδο κεφαλαίων για την παρέμβαση των εποπτικών αρχών. Το MCR είναι ένα το 85% επίπεδο εμπιστοσύνης του μέτρου VaR και είναι ένα ποσό μεταξύ του 25% και του 45% του συνολικού SCR.

Βάσει του Solvency II, η ασφαλιστική καλείται να υπολογίσει το SCR για κάθε ένα χαρτοφυλάκιο ξεχωριστά, ανάλογα με τη φύση των κινδύνων που έχει στο κάθε χαρτοφυλάκιο και λαμβάνοντας υπόψη τις τυχόν συσχετίσεις ανάμεσα στους κινδύνους των χαρτοφυλακίων της και να υπολογιστεί τελικά το ολικό SCR της επιχείρησης. Ο υπολογισμός του γίνεται τουλάχιστον μία φορά το χρόνο. Επομένως, τα κεφάλαια της επιχείρησης θα πρέπει να παρακολουθούνται σε συνεχή βάση.

Καταλαβαίνει κανείς ότι σε επίπεδο ασφαλιστικού χαρτοφυλακίου πλέον, η επιχείρηση οφείλει να μετράει συστηματικά τους κινδύνους στους οποίους έχει εκτεθεί. Ένας τρόπος για να το κάνει αυτό είναι δημιουργώντας ένα εσωτερικό σύστημα έγκαιρης προειδοποίησης που θα παρακολουθεί τα ύψη των ζημιών που δημιουργούν τα συμβόλαια του χαρτοφυλακίου. Αν ένας προκαθορισμένος αριθμός αυτών ξεπεράσει ένα συγκεκριμένο κατώφλι ύψους x_0, τότε θα θεωρηθούν ως μη αναμενόμενες ζημιές και θα σηματοδοτηθεί μία κατάσταση υψηλού κινδύνου για το χαρτοφυλάκιο αυτό.

Επομένως, καταλήγουμε ότι στην διαχείριση κινδύνου έναυσμα για την στοχαστική μελέτη τυχαίων μεταβλητών αποτέλεσε η ανάγκη μία τράπεζας ή μίας ασφαλιστικής επιχείρησης να γνωρίζει την υγιή ή μη οικονομική της κατάσταση. Σύμφωνα με το νομικό πλαίσιο, αυτό εξετάζεται παραδείγματος χάριν στην περίπτωση των τραπεζών, μέσω των αποτελεσμάτων από τα τεστ αντοχής (stress tests), ενώ μία ασφαλιστική υποχρεούται να υπολογίζει και να παρακολουθεί τα απαιτούμενα κεφάλαια για κάθε κατηγορία προϊόντων που προσφέρει.

Δύο μοντέλα τα οποία προτείνουν ένα τρόπο καλύτερης διαχείρισης των κινδύνων που αντιμετωπίζουν οργανισμοί σαν κι αυτούς που αναφέραμε περιγράφονται στα επόμενα κεφάλαια, κάνοντας προσπάθεια να μοντελοποιηθεί ο τυχαίος χρόνος αναμονής μέχρι να δοθεί “ένα προκαθορισμένο σήμα κινδύνου”.

3.3 Εφαρμογές στην διαχείριση των Κινδύνων

Λόγω του τρόπου με τον οποίο έχουν διαμορφωθεί τα ρυθμιστικά πλαίσια σχετικά με τον έλεγχο των κινδύνων σε μία ασφαλιστική επιχείρηση (Solvency II), και σε ένα χρηματοπιστωτικό οργανισμό (Βασιλεία III) έχει γεννηθεί η ανάγκη να εστιάσουμε στην μοντελοποίηση τυχαίων διωνυμικών μεταβλητών.

Όπως έχουμε ήδη αναφέρει, τα αποτελέσματα από τα τεστ αντοχής που γίνονται στα ίδια κεφάλαια μίας τράπεζας αντικατοπτρίζουν την ικανότητά τους να απορροφήσουν ζημιές σε περίπτωση αρνητικής μεταβολής των οικονομικών συνθηκών.

Τα πιστωτικά ιδρύματα καλούνται σε τακτά χρονικά διαστήματα να προσδιορίζουν την επαρκή ή μη φερεγγυότητα των κεφαλαίων τους μέσω του υπολογισμού ενός δείκτη που ορίζεται ως ο λόγος των Ιδίων κεφαλαίων της τράπεζας προς τα σταθμισμένα βάσει κινδύνου περιουσιακά της στοιχεία (Capital Ratio). Ο δείκτης αυτός οφείλει να είναι μεγαλύτερος ή ίσος του 8%, προκειμένου η διάρθρωση των κεφαλαίων της τράπεζας να θεωρηθεί επαρκής.

Η αλληλουχία αποτελεσμάτων από τα τεστ αντοχής που υπεβλήθη μία τράπεζα μέσα στο χρόνο μπορεί να χρησιμοποιηθεί ως εργαλείο παρακολούθησης της συνολικής φερεγγυότητας των κεφαλαίων της.

Η ακολουθία αυτή αποτελείται από διαδοχικές επαναλήψεις ενός πειράματος Bernoulli με δύο δυνατά αποτελέσματα S (επιτυχία) όταν η τράπεζα δεν έχει επαρκή κεφαλαιοποίηση, μιας κι αυτό είναι που θέλουμε να εντοπίσουμε, και F (αποτυχία) εναλλακτικά. Αν X_t είναι η κεφαλαιακή επάρκεια της τράπεζας την στιγμή του t τεστ αντοχής, τότε η πιθανότητα να βρεθεί πάνω από το 8% ελάχιστο απαιτούμενο όριο είναι $p=P(X_t>0.08)$ που ισοδυναμεί με το να πάρει την τιμή F (αποτυχία) η t επανάληψη των πειραμάτων Bernoulli. Αντίστοιχα, $q=1-p=P(X_t\leq 0.08)$.

Όλα τα παραπάνω μπορούν να βοηθήσουν να κατασκευάσουμε μοντέλα πιθανοτήτων στα οποία είτε κάνουμε την παραδοχή ότι είναι ανεξάρτητη η κεφαλαιακή επάρκεια την χρονική στιγμή t από την κεφαλαιακή επάρκεια τη χρονική στιγμή $t-1$, είτε πως το αποτέλεσμα εξαρτάται από το προηγούμενο, δηλαδή στην ακολουθία από τις τυχαίες διωνυμικές μεταβλητές (τεστ αντοχής) υπάρχει βαθμός εξάρτησης ή δεν υπάρχει.

Αντιλαμβανόμαστε πως η μοντελοποίηση με την βοήθεια διωνυμικών μεταβλητών που παίρνουν την τιμή 1 (Επιτυχία) ή την τιμή 0 (Αποτυχία) είναι αναγκαία. Είναι δόκιμο επομένως να εστιάσουμε την προσοχή μας σε τυχαίες μεταβλητές που σχετίζονται με την μέτρηση του πλήθους των δοκιμών που απαιτούνται μέχρι την ικανοποίηση ενός προκαθορισμένου κριτηρίου.

Μια ιδέα είναι να χρησιμοποιήσουμε μεταβλητές που μετρούν τις δοκιμές που χρειάζονται μέχρι να δημιουργηθεί μία ροή επιτυχιών (run statistics) πάνω στην διωνυμική ακολουθία, είτε οι μεταβλητές της ακολουθίας είναι ανεξάρτητες είτε όχι. Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε μεταβλητές που μετρούν τις δοκιμές που χρειάζονται μέχρι να δημιουργηθεί ένα μοτίβο από επιτυχίες και αποτυχίες (scan statistics), είτε οι τυχαίες διωνυμικές μεταβλητές είναι εξαρτημένες είτε όχι.

Ας υποθέσουμε πως η παρακάτω ακολουθία από σύμβολα είναι μια τέτοια ακολουθία διωνυμικών αποτελεσμάτων, μέσα σε ένα συγκεκριμένο χρονικό διάστημα, *FFFFSSFFFSSS*. Μπορούμε να διακρίνουμε εύκολα διαδοχικές ροές επιτυχιών και ροές αποτυχιών.

Ροή επιτυχιών καλείται μία ακολουθία συνεχόμενων επιτυχιών (S), των οποίων προηγούνται και έπονται αποτυχίες (F) ή και τίποτα. Μήκος μιας ροής επιτυχιών είναι ο αριθμός των επιτυχιών που περιλαμβάνονται στη ροή. Αντίστοιχα, ορίζεται η ροή των αποτυχιών. Στην παραπάνω ακολουθία μπορούμε να διακρίνουμε μία ροή αποτυχιών μήκους 4 και μία μήκους 3, ενώ αντίστοιχα έχουμε μία ροή επιτυχιών μήκους 2 και μία μήκους 3.

Εναλλακτικά, μπορούμε να παρατηρήσουμε τις επιτυχημένες δοκιμές που απέχουν τουλάχιστον $k-2$ αποτυχημένες δοκιμές μεταξύ τους, δηλαδή να ορίσουμε συναρτήσεις σάρωσης (scan statistics) για διάφορες τιμές του k. Παραδείγματος χάριν, στην παραπάνω ακολουθία υπάρχουν επιτυχημένες δοκιμές που απέχουν $k-2=3$ αποτυχίες μεταξύ τους.

Όπως αναλύσαμε στις προηγούμενες ενότητες η Βασιλεία ΙΙΙ δεν προβλέπει συγκεκριμένη χρονική στιγμή που θα γίνονται τα τεστ αντοχής. Οπότε η μεταβλητή που μετράει τις δοκιμές θα είναι τυχαία. Καταλαβαίνει κανείς πως ερευνητικά ενδιαφέρον έχει τότε το πλήθος της διωνυμικής ακολουθίας τ.μ. $Y_1, Y_2, \ldots$ γιατί είναι κι αυτό τυχαίο. Τότε, η κατανομή του αθροίσματος $S=\sum_{i=1}^{T} Y_i$, είναι σύνθετη και η μοντελοποίηση του βάσει ροής επιτυχιών αλλά και μιας συνάρτησης σάρωσης παρουσιάζεται στα δύο επόμενα κεφάλαια της παρούσας εργασίας.

Στο Solvency II είναι καθορισμένη η χρονική στιγμή που γίνονται έλεγχοι για την φερεγγυότητα των κεφαλαίων μιας ασφαλιστικής. Παρόλα αυτά, μπορούμε να χρησιμοποιήσουμε τις παραπάνω τυχαίες μεταβλητές όταν καλούμαστε να παρακολουθήσουμε την εξέλιξη ενός ασφαλιστικού χαρτοφυλακίου μέσα στο χρόνο. Τότε χρειάζεται να διαχειριστούμε τις υποχρεώσεις από τις διάφορες αποζημιώσεις των συμβολαίων. Για μία συγκεκριμένη χρονική περίοδο, οι συνολικές μελλοντικές αποζημιώσεις όπως είναι γνωστό, περιγράφονται από ένα τυχαίο άθροισμα τυχαίων μεταβλητών της μορφής:

$$S = \sum_{i=1}^{N(t)} Y_i .$$

Στα πλαίσια δημιουργίας ενός μηχανισμού παρακολούθησης του χαρτοφυλακίου, ένα ύψος αποζημίωσης x_0 μπορεί να λειτουργεί σαν κατώφλι πάνω από το οποίο οι αποζημιώσεις θα θεωρούνται μη αναμενόμενες. Το κατώφλι x_0 μπορεί να ενεργοποιεί ένα συναγερμό προειδοποίησης. Δηλαδή, μπορούμε να χρησιμοποιήσουμε το παραπάνω σκεπτικό δημιουργώντας ξανά μία διωνυμική ακολουθία όχι όμως από αποτελέσματα τεστ αντοχής αλλά από τα κατά πόσο η αποζημιώσεις είναι μεγαλύτερες από το ποσό x_0 ή όχι. Έτσι, και στα πλαίσια μίας ασφαλιστικής κατασκευάζουμε εσωτερικά μοντέλα προειδοποίησης.

Προσδιορίζοντας ένα ποσό x_0 δημιουργείται ακολουθία από επιτυχίες και αποτυχίες. Βασιζόμενοι σε ιστορικά δεδομένα ή την κρίση του αναλογιστή μπορούμε να προσδιορίσουμε ένα συγκεκριμένο μοτίβο που αν δημιουργηθεί, τότε το χαρτοφυλάκιο της εταιρείας απειλείται με χρεωκοπία. Δηλαδή, μετράμε τον τυχαίο χρόνο μέχρι να δημιουργηθεί για πρώτη φορά μία ροή από συνεχόμενες αποζημιώσεις που θα είναι μεγαλύτερες από το ποσό x_0 (επιτυχίες) ή αποζημιώσεις πάνω από το ποσό x_0 που θα απέχουν όμως k-2 αποζημιώσεις μικρότερες του x_0 (αποτυχίες).

Στα επόμενα δύο κεφάλαια παρουσιάζονται τα δύο αυτά μοντέλα ροής επιτυχιών και συνάρτησης σάρωσης.

ΚΕΦΑΛΑΙΟ 4

Χρόνος αναμονής μέχρι την εμφάνιση πρώτης ροής επιτυχιών

Στο κεφάλαιο αυτό δίνεται ο ορισμός μιας γεωμετρικής κατανομής *k* τάξης και βασικές σχέσεις που ισχύουν για την κατανομή αυτή. Στην συνέχεια μελετάμε τον χρόνο αναμονής μέχρι να εμφανιστεί για πρώτη φορά ροής συγκεκριμένου μήκους *k,* και δίνεται ένας αναδρομικός τύπος υπολογισμού της συνάρτησης πιθανότητας της σύνθετης κατανομής της ροής επιτυχιών.

4.1 Γεωμετρική Κατανομή τάξης *k*

Αν η τ.μ X περιγράφει το πλήθος των ανεξάρτητων πειραμάτων Bernoulli που απαιτούνται μέχρι να εμφανιστούν k συνεχόμενες επιτυχίες, με πιθανότητα εμφάνισης p, τότε η κατανομή που θα ακολουθεί η τ.μ X καλείται γεωμετρική κατανομή k τάξης με πιθανότητα επιτυχίας p. Συμβολικά γράφουμε, $X\sim\ G_k(x,p)$.

Η συνάρτηση πιθανότητας μιας τέτοια κατανομής δίνεται από την σχέση:

$$P(X=x)=\sum_{x_1,\dots,x_k}\binom{x_1+\dots+x_k}{x_1,\dots,x_k}p^x\left(\frac{q}{p}\right)^{x_1+\dots+x_k},\ x\geq k\ . \tag{4.1}$$

Το άθροισμα γίνεται πάνω σε ακεραίους $x_{1,}\, x_2,\dots,x_k\geq 0$ τέτοιους ώστε $x_1+2x_2+3x_3+\dots+k\,x_k=x$.

Η γεννήτρια συνάρτηση πιθανοτήτων μια γεωμετρικής k τάξης δίνεται από τον τύπο:

$$\Phi_k(z)=\frac{p^k z^k(1-p\,z)}{1-z+q\,p^k z^{k+1}},|z|\leq 1\ . \tag{4.2}$$

Η πρώτη και η δεύτερη παράγωγος της (4.2) στο σημείο 1 θα δώσουν την μέση τιμή και την παραγοντική ροπή δευτέρας τάξης για μια τ.μ που ακολουθεί γεωμετρική κατανομή k τάξης. Οι παρακάτω σχέσεις δίνουν την μέση τιμή κι τη διασπορά της κατανομής:

$$E[X]=\Phi'_k(1)=\frac{(1-p^k)}{q\,p^k}\,, \tag{4.3}$$

$$\sigma^2=Var[X]=\Phi''_k(1)+\Phi'_k(1)-(\Phi'_k(1))^2=\frac{[1-(2k+1)q\,p^k-p^{2k+1}]}{q^2\,p^{2k}}\,. \tag{4.4}$$

Παρατηρούμε ότι για $k=1$ όλες οι παραπάνω σχέσεις μετατρέπονται σε γνωστούς τύπους της Γεωμετρικής κατανομής.

4.2 Σύνθετη Γεωμετρική Κατανομή τάξης *k*

Όπως ήδη έχουμε αναφέρει, ερευνητικό ενδιαφέρον παρουσιάζει το τυχαίο άθροισμα τυχαίων μεταβλητών $S=\sum_{i=1}^{T_k} Y_i$, όπου οι μεταβλητές $Y_1, Y_2, \ldots$ είναι ακολουθία από ανεξάρτητες, θετικές και ισόνομες τυχαίες μεταβλητές, ανεξάρτητες της τ.μ T_k. Το τυχαίο άθροισμα S_k χρησιμοποιείται στην διαχείριση κινδύνου για την περιγραφή του χρόνου αναμονής μέχρι να εμφανιστεί για πρώτη φορά μία προκαθορισμένη και συγκεκριμένου μήκους ροή επιτυχιών.

Η μελέτη τυχαίων μεταβλητών που σχετίζονται με ροές είναι εξαιρετικά δημοφιλής σε πληθώρα επιστημονικών πεδίων. Ήδη από την εποχή του De Moivre (1756) η μελέτη του αριθμού των ροών επιτυχιών προσέλκυσε το ερευνητικό ενδιαφέρον. Στις μέρες μας, βρίσκει εφαρμογές σε πεδία όπως η βιολογία (ακολουθίες έλικας DNA), η αξιοπιστία μηχανικών συστημάτων, ο ποιοτικός έλεγχος και η θεωρία των κινδύνων. Στο κεφάλαιο αυτό περιγράφεται ένα μοντέλο διαχείρισης του κινδύνου χρεοκοπίας μέσα σε ένα χρηματοπιστωτικό ίδρυμα, αναζητώντας κυρίως ροές επιτυχιών μήκους *k*.

Υπάρχει εκτενής βιβλιογραφία που αφορά αποτελέσματα για την σύνθετη γεωμετρική κατανομή ανώτερης τάξης *k*, που είναι η κατανομή που ακολουθεί η τ.μ. T_k . Αναφέρουμε ότι οι γεννήτριες συναρτήσεις των τυχαίων μεταβλητών που παίρνουν μέρος σε ένα τ.α S_k δίνονται από τις σχέσεις:

$$P_Y(z)=\sum_{x=1}^{\infty} P(Y_t=x)z^x ,$$

$$P_{T_k}(z)=\sum_{t=1}^{\infty} P(T_k=t)z^t ,$$

$$P_{S_k}(z)=\sum_{t=1}^{\infty} P(S_k=t)z^t .$$

Σύμφωνα με τους Balakrishman, Koutras (2002) η γεννήτρια συνάρτηση της γεωμετρικής κατανομής τάξης *k*, είναι η ακόλουθη:

$$P_{T_k}(z)=E(z^{T_k})=\frac{(pz)^k-(pz)^{k+1}}{1-z+q p^k z^{k+1}} . \tag{4.5}$$

Επομένως, μέσω της σχέσης (4.5) παράγεται μια έκφραση για την γεννήτρια συνάρτηση του τ.α S_k που ακολουθεί σύνθετη γεωμετρική κατανομή τάξης k, η οποία δίνεται από τον τύπο:

$$P_{S_k}(z)=E(z^{S_k})=\frac{(p P_Y(z))^k-(p P_Y(z))^{k+1}}{1-P_Y(z)+q p^k (P_Y(z))^{k+1}} . \tag{4.6}$$

Όπως είναι γνωστό, από την θεωρία κινδύνου έχουμε σχέσεις για τον υπολογισμό βασικών ροπών, όπως η μέση τιμή και η διασπορά του τυχαίου αθροίσματος. Για την σύνθετη γεωμετρική κατανομή τάξης k του S_k έχουμε τους παρακάτω τύπους:

- $$E(S_k)=E(T_k)E(Y_t)=\frac{1-p^k}{p^k(1-p)}E(Y_t) ,$$
- $$\begin{aligned} Var(S_k)&=E(T_k)Var(Y_t)+(E(Y_t))^2 Var(T_k)=\frac{1-p^k}{p^k(1-p)}Var(Y_t) \\ &+(E(Y_t))^2\frac{1-(2k+1)(1-p)p^k-p^{2k+1}}{(1-p)^2 p^{2k}} \end{aligned} .$$

Για τον υπολογισμό των παραπάνω σχέσεων χρησιμοποιήθηκαν οι τύποι (4.3), (4.4) για την μέση τιμή και την διασπορά μίας γεωμετρικής κατανομής τάξης *k*.

Με εφαρμογή της σχέσης (4.6) θα παράγουμε στην συνέχεια μία αναδρομική σχέση για τη συνάρτηση πιθανότητας μιας σύνθετης γεωμετρικής τάξης $k=2$. Σ΄ αυτή την περίπτωση θα έχουμε ένα τυχαίο άθροισμα που περιγράφει τον τυχαίο χρόνο αναμονής μέχρι να εμφανιστεί μία ροή επιτυχίας μήκους 2. Το άθροισμα αυτό ταυτίζεται, όπως θα δούμε και στο επόμενο κεφάλαιο με την περίπτωση σάρωσης για *k*=2.

Ας υποθέσουμε ότι οι τ.μ. $Y_1, Y_2, \ldots$ που σχηματίζουν τυχαίο άθροισμα και ακολουθούν μία γεωμετρική κατανομή με συνάρτηση πιθανότητας $P\left(Y_i=y\right)=\theta\left(1-\theta\right)^{y-1}$, $y=1,2,\ldots$. Τότε η γεννήτρια συνάρτηση της γεωμετρικής κατανομής, όπως γνωρίζουμε είναι $P_Y(z)=\dfrac{\theta z}{1-(1-\theta)z}$.

Οπότε, μέσω της (4.6) και για $k=2$ μπορούμε να παράγουμε την γεννήτρια συνάρτηση του τυχαίου αθροίσματος S_2 :

$$\begin{aligned} P_{S_2}(z)=E(z^{S_2})&=\frac{(p P_Y(z))^2-(p P_Y(z))^3}{1-P_Y(z)+q p^2 (P_Y(z))^3} \\ &=\frac{(p z\theta)^2}{1-[2-\theta(1+p)]z-[(1+p)\theta-p^2\theta^2-1]z^2} \end{aligned}. \tag{4.7}$$

Η συνάρτηση πιθανότητας $f_2(t)=P(S_2=t)$ του τυχαίου αθροίσματος με κατανομή μία σύνθετη γεωμετρική δευτέρας τάξης θα ικανοποιεί την εξής αναδρομική σχέση:

$$f_2(t)=[2-\theta(1+p)]f_2(t-1)+[(1+p)\theta-p^2\theta^2-1]f_2(t-2), \tag{4.8}$$

με αρχικές συνθήκες

$$f_2(1)=0, \quad f_2(2)=p^2\theta^2. \tag{4.9}$$

Όπως αναφέρεται στους Koutras & Eryilmaz (2016) στην σχέση (4.7) έχουμε το πηλίκο δύο πολυωνύμων ως προς *z*. Η σχέση για την γεννήτρια συνάρτηση ενός τ.α S_2 μπορεί να μετασχηματιστεί σε μία σχέση της μορφής:

$$P_{S_2}(z)=E(z_2^S)=c+\frac{u_1 z+u_0}{u_2 z^2+u_1 z+u_0}=c+\frac{U(z)}{V(z)}, \tag{4.10}$$

όπου *c* ένας αριθμός ανεξάρτητος του *z* και ίσος με $c=\dfrac{(p\theta)^2}{u_2}$ και για τους συντελεστές $u_{0,}\, u_{1,}\, u_2$ ισχύουν τα εξής:

$$u_0=1, u_1=-(1+p)\theta+(p\theta)^2+1, u_2=-2+\theta(1+p), \tag{4.11}$$

$$u_1=-(p\theta)^2\frac{u_1}{u_2}, \quad u_0=-(p\theta)^2\frac{u_0}{u_2}.$$

Τέλος, τα πολυώνυμα ως προς *z* όπως φαίνεται και από την (4.10) δίνονται από τις σχέσεις:

$$U(z)=u_1 z+u_0, \quad V(z)=u_2 z^2+u_{1z}+u_0. \tag{4.12}$$

Αν η παράμετρος της γεωμετρικής τ.μ. T_2 είναι διάφορη της μονάδας, δηλαδή $p\neq 1$ τότε η δευτεροβάθμια εξίσωση ως προς z, $V(z)=0$ θα έχει δύο πραγματικές ρίζες, τις z_1, z_2.

Η διακρίνουσα υπολογίζεται από την γνωστή σχέση και ισούται με την εξής παράσταση:

$$\Delta = u_1^2 - 4u_2 u_0 ,$$

οπότε

$$\sqrt{\Delta} = \theta\sqrt{(1-p)(1+3p)} .$$

Επομένως, οι δύο λύσεις της δευτεροβάθμιας εξίσωσης, $V(z)=0$ θα δίνονται από τους τύπους:

$$z_{1,2} = \frac{-2+\theta(1+p)\pm\theta\sqrt{(1-p)(1+3p)}}{2[(1+p)\theta-(p\theta)^2-1]} . \tag{4.13}$$

Αν προχωρήσουμε σε ανάλυση σε απλά κλάσματα για τα πολυώνυμα $\frac{U(z)}{V(z)}$ της σχέσης (4.10) οδηγούμαστε στον επόμενο τύπο για την συνάρτηση πιθανότητας $f_2(t)$:

$$f_2(t) = P(S_2 = t) = \frac{\rho_1}{z_1^{t+1}}\frac{\rho_2}{z_2^{t+1}}, t>0 . \tag{4.14}$$

Οι συντελεστές ρ_1, ρ_2 μπορούν να βρεθούν από τον τύπο:

$$\rho_i = \frac{-U(z_i)}{V'(z_i)} = \frac{-u_1 z_i + u_0}{2u_2 z_i + u_1}, i=1,2 . \tag{4.15}$$

Γρήγορα καταλαβαίνει κανείς παρατηρώντας την παραπάνω μεθοδολογία πως οι υπολογισμοί μίας τέτοιας διαδικασίας είναι σχετικά χρονοβόροι, με αποτέλεσμα η μέθοδος αυτή να θεωρείται δυσλειτουργική.

Σύμφωνα με Koutras & Eryilmaz (2016) μία καλή προσέγγιση της συνάρτησης πιθανότητας του τυχαίου αθροίσματος S_2 μπορεί να δώσει το επόμενο αποτέλεσμα για μεγάλες τιμές της μεταβλητής *t*. Χρησιμοποιώντας μόνο την μικρότερη κατά απόλυτο τιμή ρίζα από την (4.13) αποδεικνύεται ότι ισχύει η επόμενη ασυμπτωτική σχέση για την $f_2(t)$:

$$f_2(t) = P(S_2 = t) \rightarrow \frac{\rho_1}{z_1^{t+1}} . \tag{4.16}$$

Επομένως, η πιθανότητα ο χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών μήκους 2 μπορεί σε μεγάλα χρονικά διαστήματα να προσεγγιστεί μέσω της ρίζας της (4.13).

4.3 Αναδρομική σχέση για σύνθετη γεωμετρική κατανομή τάξης *k*

Όπως αναπτύξαμε στο προηγούμενο κεφάλαιο, η προσπάθεια μοντελοποίησης του χρόνου αναμονής με σκοπό την αποφυγή χρεοκοπίας μέσα σε μία τράπεζα είναι εξαιρετικά επίκαιρη. Επομένως, ας υποθέσουμε πως μία κατάσταση που θα μπορούσε να σηματοδοτεί τον ερχομό της χρεοκοπίας, είναι η τράπεζα να καταστεί αφερέγγυα *k* συνεχόμενες φορές.

Αν έχουμε μία ακολουθία $Y_1, Y_2, \ldots$ από ανεξάρτητα και ισόνομα πειράματα Bernoulli με πιθανότητα επιτυχίας $P\left(Y_i=1\right)=p$, τότε η κατανομή της τ.μ T_k θα είναι μία γεωμετρική κατανομή τάξης *k*. Συμβολικά, έχουμε ότι:

$$T_k=min\{t : X_{t-k+1}=\ldots=X_t=1\} . \tag{4.17}$$

Παράδειγμα 4.1: Ας υποθέσουμε ότι έχουμε την παρακάτω ακολουθία από *ν=15* ανεξάρτητες δοκιμές Bernoulli: *SFFSFSFSSFFFSSSFFS.* Τότε οι τιμές της τυχαίας μεταβλητής T_k για διάφορες τιμές του ακεραίου *k* θα είναι:

- Για *k*=2 έχουμε: $T_2=9$
- Για *k*=3 έχουμε: $T_3=15$

Αφού τόσες ανεξάρτητες δοκιμές θα χρειαστούν μέχρι να πάρουμε ροές επιτυχιών μήκους 2 και 3 αντίστοιχα.

Δηλαδή, σύμφωνα με το παράδειγμα 4.1, αναζητάμε την εμφάνιση στην ακολουθία των αποτελεσμάτων από τα τεστ αντοχής, μίας ροής επιτυχιών μήκους *k*. Ο συνολικός χρόνος που χρειάζεται μέχρι να εμφανιστεί το μοτίβο των συνεχόμενων επιτυχιών μπορεί να αναπαρασταθεί μέσω ενός τυχαίου αθροίσματος της μορφής, $S=\sum_{i=1}^{T_k} Y_i$.

Γενικά, καταλαβαίνει κανείς πως μπορούμε να ορίσουμε σειρά από τυχαίες μεταβλητές που θα αναφέρονται στα αποτελέσματα ενός πειράματος τύχης Bernoulli, θα μετρούν διαφορετικές ποσότητες και θα βοηθούν στην κατασκευή μοντέλων για την στοχαστική μοντελοποίηση του χρόνου αναμονής.

Το μοτίβο των *k* συνεχόμενων επιτυχιών είναι ένα καθορισμένο εξ' αρχής κριτήριο και ορίζει σε κάθε περίπτωση τον συνδυασμό των γεγονότων εκείνων που εκτιμάται πως μπορεί να προκαλέσουν χρεοκοπία. Τέτοια μοτίβα αναφέρονται ως “μοτίβα ανίχνευσης” (detection patterns).

Ο ενδιάμεσος χρόνος που μεσολαβεί από το $t-1$ τεστ αντοχής μέχρι το t τεστ, μπορεί να παρασταθεί από μία τ.μ $Y_t, t\geq 2$. Το πλήθος των τεστ αντοχής που χρειάζονται μέχρι να εμφανιστεί για πρώτη φορά το καθορισμένο μοτίβο ανίχνευσης, μετράται από την τ.μ T_k με *k* να είναι το μήκος του μοτίβου ανίχνευσης.

Επομένως, θα έχουμε ότι ο ολικός χρόνος μέχρι να εμφανιστεί ένα “σημάδι χρεοκοπίας” (default signal) περιγράφεται από ένα τυχαίο άθροισμα τυχαίων μεταβλητών:

$$S_k=\sum_{t=1}^{T_k} Y_t=Y_1+\ldots+Y_{T_k}. \tag{4.18}$$

Έχουμε αναφερθεί στο κεφάλαιο 2 στην οικογένεια κατανομών PH και την αναδρομική σχέση που παρουσίασε ο Eisele (2006) για την συνάρτηση πιθανότητας ενός τυχαίου αθροίσματος $S=\sum_{i=1}^{T} Y_i$ όταν ακολουθεί σύνθετη κατανομή *PH,* δηλαδή η τ.μ $T\sim DPH_d(\Lambda,\boldsymbol{\pi})$.

Αξίζει να πούμε ότι βάσει των Tank, Eryilmaz (2014) γνωρίζουμε πως η γεωμετρική κατανομή τάξης *k* ανήκει στην οικογένεια κατανομών *PH*. Συγκεκριμένα απέδειξαν πως η τ.μ T_k ενός τ.α S_k όπως έχει οριστεί, ακολουθεί $T_k\sim PH_k(\Lambda,\boldsymbol{\pi})$ με αρχικό διάνυσμα:

$$\boldsymbol{\pi}=(1,0,\ldots,0)',$$

και πίνακα μετάβασης:

$$\Lambda=\begin{pmatrix} 1-p & p & 0 & \cdots & 0 \\ 1-p & 0 & p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1-p & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{4.19}$$

Χρησιμοποιώντας το χαρακτηριστικό πολυώνυμο του παραπάνω πίνακα *Λ*, το οποίο υπολογίζεται βάσει της σχέσης:

$$det(x I_k-\Lambda)=x^k+(p-1)x^{k-1}+\ldots+(p^{k-1}-p^{k-2})x+(p^k-p^{k-1})=x^k+\sum_{i=1}^{k} b_i x^{k-i}, \tag{4.20}$$

μπορούμε να βρούμε τους συντελεστές α_i , b_i για μία γεωμετρική κατανομή τάξης *k*:

$$b_i=p^i-p^{i-1}, i=1,\ldots,k \quad \text{και} \quad \alpha_i=0, i=1,2,\ldots,k-1\ , \quad \alpha_k=p^k\ . \tag{4.21}$$

Βάσει των παραπάνω μπορούμε να παράγουμε αναδρομικές σχέσεις για την σύνθετη γεωμετρική κατανομή τάξης *k.* Ο τύπος του Eisele που δόθηκε στο Κεφάλαιο 2, σε συνδυασμό με τους συντελεστές (4.21) μας παρέχουν αναδρομική σχέση υπολογισμού της συνάρτησης πιθανότητας για μία σύνθετη γεωμετρική κατανομή τάξης *k*.

Η κατανομή αυτή περιγράφει ικανοποιητικά ένα τυχαίο άθροισμα τυχαίων μεταβλητών S_k που ορίσαμε στην αρχή της ενότητας, όπου δηλαδή η $T_k\sim PH_k(\Lambda,\boldsymbol{\pi})$ και η ακολουθία των τ.μ

$Y_1, Y_2, \ldots$ απαρτίζεται από θετικές ανεξάρτητες και ισόνομες, συνεχείς ή διακριτές τυχαίες μεταβλητές με $f_Y^{*j}(t)$ να παριστάνει την j συνέλιξη των τυχαίων μεταβλητών $Y_1, Y_2, \ldots, Y_j$, η οποία υπολογίζεται από μία σχέση της μορφής:

$$f_Y^{*j}(t) = P\left(\sum_{j=1}^{t-1} Y_i = t\right), \, j = 1,2, \ldots .$$

Επομένως, σύμφωνα με τους Koutras & Eryilmaz (2016) η συνάρτηση πιθανότητας $f_k(t) = P(S_k = t), t = 1,2, \ldots$ μιας σύνθετης γεωμετρικής κατανομής τάξης *k,* η οποία όπως είπαμε αναπαριστάται και ως *PH* μέσω της (4.19), θα ικανοποιεί τον παρακάτω αναδρομικό τύπο:

$$f_k(t) = p^k f_Y^{*k}(t) - \sum_{j=1}^{k} (p^j - p^{j-1}) \left(\sum_{u=1}^{t-1} f_k(u) f_Y^{*j}(t-u)\right), t > y_0 k \, . \tag{4.22}$$

Οι αρχικές συνθήκες της σχέσης (4.22) είναι οι εξής:

$$f_k(t) = 0, \; t < y_0 k \quad \text{και} \quad f_k(y_0 k) = [\, p P(Y = y_0)]^k \, . \tag{4.23}$$

Σημειώνεται ότι το στήριγμα των τ.μ. $Y_1, Y_2, \ldots$ θα είναι της μορφής $\{y_{0,} y_0 + 1, \ldots.\}$. Επίσης, παρατηρούμε ότι η (4.22) προκύπτει εφαρμόζοντας τους συντελεστές (4.21) στον τύπο (2.20) του Eisele (2006).

Ενδιαφέρον παρουσιάζει η περίπτωση που και οι τ.μ $Y_1, Y_2, \ldots$ ακολουθούν κατανομή *PH* τάξης *c*, με αναπαράσταση $(M, \boldsymbol{\rho})$. Επομένως, έχουμε ένα τυχαίο άθροισμα στο οποίο και η τ.μ T_k αλλά και η τ.μ $Y_i, i = 1,2, \ldots$ είναι διακριτές κατανομές ανεξάρτητες μεταξύ τους και ανήκουν στην οικογένεια κατανομών *PH*.

Σύμφωνα με τον Neuts (1981), η μείξη δύο διακριτών κατανομών *PH* τάξης *k* και *c* αντίστοιχα δίνεται από την σχέση:

$$\sum_{u=0}^{\infty} f_{T_k}(u) f_Y^{*u}(t)$$

και ακολουθεί επίσης κατανομή *PH* τάξης *ck*, με αναπαράσταση $(\Sigma, \boldsymbol{\sigma})$ και παραμέτρους που προσδιορίζονται από τις σχέσεις:

$$\Sigma = M \times I_k + \boldsymbol{u}\, \boldsymbol{p}' \times (I_k - a\Lambda)^{-1} \Lambda \, , \quad \boldsymbol{\sigma} = \boldsymbol{\rho} \times \boldsymbol{\pi} (I_k - a\Lambda)^{-1} \, , \tag{4.24}$$

και με αρχικά διανύσματα,

$$\alpha = 1 - \boldsymbol{\rho}' 1 \, , \;\; u = (I_c - M)\mathbf{1} \, . \tag{4.25}$$

Γνωρίζουμε ότι αν οι τ.μ T_k και $Y_i, i = 1,2, \ldots$ είναι ανεξάρτητες μεταξύ τους τ.μ, τότε η συνάρτηση πιθανότητας της σύνθετης κατανομής του τ.α S_k θα δίνεται από την σχέση:

$$P(S_k=t)=\sum_{j=1}^{\infty} P(\sum_{i=1}^{j} Y_i=t)P(T_k=j)=\sum_{j=0}^{\infty} f_{T_k}(j) f_Y^{*j}(t).$$

Δηλαδή, αν τα Y_i ακολουθούν κατανομή $PH_c(M,\boldsymbol{\rho})$, τότε η συνάρτηση πιθανότητας του τ.α S_k μπορεί να υπολογιστεί και αναδρομικά, σύμφωνα με Koutras & Eryilmaz (2016), βάσει της σχέσης:

$$f_k(t)=P(S_k=t)=\boldsymbol{\sigma}' \Sigma^{t-1}(I_{ck}-\Sigma)\mathbf{1},\ t=1,2,\dots, \tag{4.26}$$

με

$$\Sigma=M\times I_k+\boldsymbol{u}\,\boldsymbol{p}'\times(I_k-a\Lambda)^{-1}\Lambda,$$
$$\boldsymbol{\sigma}=\boldsymbol{\rho}\times\boldsymbol{\pi}(I_k-a\Lambda)^{-1}.$$

Αναφέρουμε ότι η πράξη $A\times B$ μεταξύ 2 πινάκων Α, Β παριστάνει το γινόμενο Kronecker.

Παράδειγμα 4.2: Ας υποθέσουμε πως έχουμε τ.μ $T_k \sim PH_k(\Lambda,\boldsymbol{\pi})$ και μία ακολουθία απ ό τυχαίες μεταβλητές $Y_1, Y_2, \dots$ που ακολουθούν αρνητική διωνυμική με παραμέτρους $r=2$ και *p*, δηλαδή η συνάρτηση πιθανότητας της δίνεται από τον τύπο $f_Y(y)=P(Y_i=y)=(y-1)\theta^2(1-\theta)^{y-2}, y=2,3,\dots$.

Η αρνητική διωνυμική είναι μία κατανομή *PH* τάξης $c=2$ με αναπαράσταση $(M,\boldsymbol{\rho})$ με:

$$\boldsymbol{\rho}=(1,0)', \qquad M=\begin{pmatrix} 1-\theta & \theta \\ 0 & 1-\theta \end{pmatrix}.$$

Επομένως, από τις σχέσεις (4.21) θα προκύψει $\alpha=1-\boldsymbol{\rho}'\mathbf{1}=0$. Υπολογίζεται ότι ο πίνακας μετάβασης Σ και το αρχικό διάνυσμα $\boldsymbol{\sigma}$ που περιγράφει την μείξη των δύο κατανομών *PH*, με αντικατάσταση από τις σχέσεις (4.19) δίνονται από τις σχέσεις:

$$\boldsymbol{\sigma}=\boldsymbol{\rho}\times\boldsymbol{e}=(1,0,0,\dots,0)'_{1\times(2k)}, \qquad \boldsymbol{u}=(I_2-M)\mathbf{1}=(0,\theta)'$$

$$\Sigma=M\times I_k+\boldsymbol{u}\,\boldsymbol{p}'\times\Lambda=\begin{pmatrix} 1-\theta & \theta \\ 0 & 1-\theta \end{pmatrix}\times I_k+\begin{pmatrix} 1 & 0 \\ \theta & 0 \end{pmatrix}\times\Lambda=\begin{pmatrix} (1-\theta)I_k & \theta I_k \\ R & (1-\theta)I_k \end{pmatrix},$$

με

$$R=\begin{pmatrix} \theta(1-p) & \theta p & 0 & \cdots & 0 \\ \theta(1-p) & 0 & \theta p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \theta(1-p) & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Τελικά, αφού έχουμε ότι η συνάρτηση πιθανότητας της σύνθετης γεωμετρικής κατανομής τάξης *k* θα ικανοποιεί βάσει της (4.22), την εξής αναδρομική σχέση:

$$f_k(t)=P(S_k=t)=\boldsymbol{\sigma}'\Sigma^{t-1}(I_{2k}-\Sigma)\mathbf{1}=(\theta p)\boldsymbol{\sigma}'\begin{pmatrix}(1-\theta)I_k & \theta I_k \\ R & (1-\theta)I_k\end{pmatrix}^{t-1}\boldsymbol{\tau}',\, t=1,2,\ldots,$$

με

$$(I_{2k}-\Sigma)\mathbf{1}=(0,0,\ldots,\theta p)'_{2k+1}=\theta p\,\boldsymbol{\tau}'$$

και $\boldsymbol{\tau}'=(0,0,\ldots,1)$ να είναι διάνυσμα του πίνακα R^{2k}.

ΚΕΦΑΛΑΙΟ 5

Χρόνος αναμονής μέχρι την επίτευξη προκαθορισμένης σάρωσης

Στο κεφάλαιο αυτό παρουσιάζεται η σύνθετη κατανομή ενός τυχαίου αθροίσματος τυχαίων μεταβλητών που σχετίζεται με μία συνάρτηση σάρωσης (scan statistic). Γίνεται η υπόθεση ότι η τ.μ T_k μετράει τον χρόνο αναμονής μέχρι να εμφανιστούν δύο επιτυχίες που απέχουν το πολύ k-2 αποτυχίες μεταξύ τους. Δίνονται αποτελέσματα που ισχύουν για την περίπτωση που τα πειράματα Bernoulli είναι ανεξάρτητα μεταξύ τους, ως ειδική περίπτωση ενός πιο γενικού μοντέλου στο οποίο η διωνυμική συνάρτηση σάρωσης συνδέεται με μία ομογενή εξαρτημένη Μαρκοβιανή σ.δ δύο σταδίων.

5.1 Κατασκευή συνάρτησης σάρωσης

Ο κυριότερος λόγος που πολλά στοχαστικά μοντέλα μέχρι σήμερα δεν αποτυπώνουν ικανοποιητικά την πραγματικότητα είναι η υπόθεση ανεξαρτησίας μεταξύ των μεταβλητών που συμμετέχουν σε αυτά. Σε διάφορες εφαρμογές είτε στην οικονομία, είτε στην φύση ελάχιστες ποσότητες είναι ανεξάρτητες μεταξύ τους. Αυτό έχει οδηγήσει μερίδα του ερευνητικού κόσμου να στρέψει το ενδιαφέρον του στην ανάπτυξη μοντέλων, οι υποθέσεις των οποίων προβλέπουν κάποιο βαθμό εξάρτησης μεταξύ των μεταβλητών. Τα γενικευμένα μοντέλα τέτοιου τύπου αποτυπώνουν μία κατάσταση πιο κοντά στην πραγματικότητα και δίνουν καλύτερα προσεγγιστικά αποτελέσματα.

Ένα επαρκές μοντέλο για να περιγράψει την εξέλιξη του Capital Ratio της τράπεζας τη χρονική στιγμή t είναι μία εξαρτημένη μαρκοβιανή αλυσίδα. Δηλαδή, μπορούμε να υποθέτουμε ότι η τ.μ. X_t Όπως την έχουμε ορίσει στα προηγούμενα κεφάλαια, εξαρτάται από το μέγεθος του προηγούμενου Capital Ratio τη χρονική στιγμή $t-1$, δηλαδή εξαρτάται από την τιμή της τ.μ. X_{t-1}.

Τα αποτελέσματα από τα τεστ αντοχής κάθε χρονική στιγμή t δημιουργούν μία ακολουθία διωνυμικών αποτελεσμάτων ξ_t, $t=1, 2, \ldots$ μιας ομογενούς μαρκοβιανής αλυσίδας δύο σταδίων (0-1) από εξαρτημένες δοκιμές με πιθανότητα:

$$P[\xi_t=j|\xi_{t-1}=i]\,p_{ij}, \quad t=2, 3, \ldots,$$

με αρχικές πιθανότητες $P(\xi_1=0)=P(X_t\geq 0{,}08)=p_0$ και $P(\xi_1=1)=P(X_t<0{,}08)=1-p_0=p_1$, με $i, j\in\{0,1\}$.

Εφόσον έχουμε υποθέσει ότι η τράπεζα δεν είναι σε θέση να γνωρίζει την χρονική στιγμή που θα λάβει χώρα το επόμενο τεστ αντοχής,αφού κάτι τέτοιο δεν προβλέπεται στο ρυθμιστικό πλαίσιο της Βασιλείας ΙΙΙ, τότε και ο ενδιάμεσος χρόνος ανάμεσα στο t και $t-1$ τεστ αντοχής είναι τυχαίος και θα περιγράφεται από μία τ.μ. $Y_t,\quad t\geq 2$.

Επίσης, ας υποθέσουμε ότι η τράπεζα έχει εκτιμήσει ότι μια επικίνδυνη κατάσταση για εκείνη είναι να επιτύχει σε δύο κοντινά μεταξύ τους, το πολύ $k-2$, τεστ αντοχής. Άρα, ο συνολικός χρόνος αναμονής μέχρι να σηματοδοτηθεί μία επικίνδυνη κατάσταση μπορεί να περιγραφεί από ένα τυχαίο άθροισμα τυχαίων μεταβλητών $S_k=\sum_{i=1}^{T_k} Y_i$.

Η απαριθμήτρια τ.μ. T_k μετράει τον χρόνο αναμονής μέχρι στην ακολουθία των διωνυμικών αποτελεσμάτων να εμφανιστούν για πρώτη φορά δύο επιτυχημένα τεστ αντοχής το πολύ k τεστ μακριά. Σαρώνοντας την ακολουθία των αποτελεσμάτων περιμένουμε μέχρι να εμφανιστεί ένα συγκεκριμένος συνδυασμός πλήθους επιτυχιών και αποτυχιών όπου είναι το ''προκαθορισμένο μοτίβο ανίχνευσης''. Οι προσθετέοι στο παραπάνω τυχαίο άθροισμα προσδιορίζονται από μία απλή συνάρτηση σάρωσης T_k και γι΄ αυτό το τ.α $S_k=\sum_{i=1}^{T_k} Y_i$ λέμε ότι ακολουθεί σύνθετη συνάρτηση σάρωσης.

Σημειώνουμε ότι η ακολουθία $\xi_t,\ t=1, 2, \ldots$ οι διωνυμικές μεταβλητές $\xi_{1,}\xi_{2,}\ldots$ παίρνουν την τιμή 1 (επιτυχία, S) και 0 διαφορετικά, όπου 1 θα θεωρούμε πάντα την μη επαρκή κεφαλαιοποίηση της τράπεζας.

Οπότε ένα ικανό κριτήριο αξιολόγησης για τον έλεγχο της τράπεζας θα μπορούσε να είναι σε αυτή την ακολουθία δύο επιτυχίες απέχουν το πολύ *k*-2, $k\geq 2$ αποτυχίες μεταξύ τους. Τότε η τ.μ. T_k μετράει τις δοκιμές που θα χρειαστούν μέχρι να εμφανιστούν για πρώτη φορά μοτίβα της μορφής: *SS*, *SFS*, *SFFS*, *SFFFS* κ.λ.π.

Παράδειγμα 5.1: Έστω ότι έχουμε την παρακάτω ακολουθία διωνυμικών μεταβλητών FFSFFFSFFFFSFFSFSS. Τότε,

- Για *k*=2, περιμένουμε το μοτίβο SS οπότε έχουμε $T_2=18$.
- Για *k*=3, περιμένουμε το μοτίβο της μορφής SFS οπότε έχουμε $T_3=17$.

Παρατηρούμε ότι για την περίπτωση όπου k=2, δηλαδή η συνάρτηση σάρωσης T_k μετράει τις δοκιμές που θα χρειαστούν μέχρι να εμφανιστεί το μοτίβο *SS*, των επιτυχιών που απέχουν $k-2=0$ αποτυχίες μεταξύ τους. Αυτή η περίπτωση ταυτίζεται με το μοντέλο ροής επιτυχιών που περιγράψαμε στο προηγούμενο κεφάλαιο. Επομένως σε αυτή την περίπτωση η τ.μ T_k περιγράφεται από μία γεωμετρική κατανομή τάξης k=2.

Κάνοντας χρήση των κατανομών *PH* και συγκεκριμένα των διακριτών *PH,* μπορούμε να κατασκευάσουμε ένα μοντέλο διαχείρισης κινδύνου με την διακριτή κατανομή *PH* τάξης d να περιγράφει την κατανομή της τ.μ T_k που μετράει τον χρόνο μέχρι να φτάσει στην απορροφητική της κατάσταση μία μαρκοβιανή αλυσίδα d βημάτων. Όπως αναφέραμε και στο 2ο κεφάλαιο ο μεταβατικός πίνακας πιθανοτήτων της μαρκοβιανής αλυσίδας, Λ_0 διάστασης $(d+1)\times(d+1)$ και το αρχικό διάνυσμα πιθανοτήτων $\boldsymbol{\pi_0}=(\pi_1,\pi_2,...,\pi_{d+1})'$ περιγράφουν την κατανομή αυτή και η συνάρτηση πιθανότητας της διακριτής *PH* δίνεται από τον τύπο:

$$P(T=t)=\boldsymbol{\pi}'\Lambda^{t-1}\boldsymbol{u},\ t=1,2,...,$$

όπου Λ ο $d\times d$ υποπίνακας του Λ_0.

5.2 Σύνθετης συνάρτησης σάρωσης σε ανεξάρτητες διωνυμικές τυχαίες μεταβλητές

Όπως έχουμε αναφέρει ο Eisele (2006) παρουσίασε μία αναδρομική σχέση (2.20) για την εύρεση της συνάρτησης πιθανότητας του τυχαίου αθροίσματος $S=\sum_{i=1}^{T}Y_i$. Η σχέση αυτή αναφέρεται στην περίπτωση που η τ.μ T ακολουθεί μια διακριτή κατανομή *PH* διάστασης d. Για να εφαρμοστεί ο τύπος του Eisele δεν χρειάζεται οι τ.μ. Y_i να ακολουθούν συγκεκριμένη κατηγορία κατανομών, αρκεί να είναι θετικές, ανεξάρτητες και ισόνομες τυχαίες μεταβλητές.

Η τ.μ T_k που μετράει τις ανεξάρτητες διωνυμικές μεταβλητές που θα χρειαστούν μέχρι την εμφάνιση επιτυχιών που απέχουν k-2 αποτυχίες μεταξύ τους, ακολουθεί μία διακριτή κατανομή *PH* διάστασης $k+1$ με αρχικό διάνυσμα $\boldsymbol{\pi}=(1,0,..,0)'$ και ο πίνακας μετάβασης $(k+1)\times(k+1)$ της μορφής:

$$
\Lambda=\begin{pmatrix}
1-p & p & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1-p & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 1-p & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1-p & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1-p \\
1-p & p & 0 & 0 & \cdots & 0 & 0
\end{pmatrix}.
$$

Συμβολικά γράφουμε ότι η $T_k \sim PH_{k+1}(\boldsymbol{\pi}, \Lambda)$.

Για να εφαρμοστεί ο αναδρομικός τύπος (2.20) και να έχουμε ένα αποτέλεσμα για την συνάρτηση πιθανότητας του τ.α S_k πρέπει να υπολογιστούν οι συντελεστές a_i, b_i. Οι συντελεστές αυτοί βρίσκονται από τις σχέσεις (2.6) και (2.9) αντίστοιχα.

Επομένως, για την τ.μ T_k σύμφωνα με Κούτρας, Κούτρας, Yalcin (2015) οι συντελεστές θα είναι:

$$b_1 = p-1\,, \quad b_k = -p(1-p)^{k-1} \ \text{ και } \ b_2 = \ldots = b_{k+1} = 0\,.$$

Οι συντελεστές a_i βάσει της σχέσεις (2.9) και για $t = 1,2,3,\ldots k+1$ έχουμε ότι:

$$a_1 = P(T=1) = 0 \ \text{ και } \ a_t = P(T=t) + \sum_{i=1}^{t-1} b_i P(T=t-i).$$

Επομένως, για τη τ.μ. T_k ισχύει ότι:

$$a_1 = 0 = a_{k+1} \ \text{ και } \ a_i = p^2(1-p)^{i-2}, \quad i = 2,3,\ldots,k\,.$$

Οι παραπάνω σχέσεις αποδεικνύονται με χρήση της (2.9) και του τύπου για την συνάρτηση πιθανότητας τ.μ. T_k, ο οποίος είναι:

$$P(T_k = i) = (i-1)\,p^2 q^{i-2}, \quad 0 < i \leq k\,. \tag{5.1}$$

Τελικά, εφαρμόζοντας τα παραπάνω αποτελέσματα για τους συντελεστές a_i, b_i στον αναδρομικό τύπο του Eisele (2006) θα έχουμε μια αναδρομική σχέση για την συνάρτηση πιθανότητας του τ.α S_k που εκφράζει τον συνολικό χρόνο που θα χρειαστεί να περιμένουμε έως την εμφάνιση δύο επιτυχίες που απέχουν το πολύ k-2 αποτυχίες, για την περίπτωση ανεξάρτητων διωνυμικών μεταβλητών.

Ο αναδρομικός τύπος της συνάρτησης πιθανότητας του τυχαίου αθροίσματος τυχαίων μεταβλητών $Y_1, Y_2, \ldots$ με στήριγμα το σύνολο $\{y_{0,}\, y_0+1, \ldots\}$ δίνεται για κάθε τιμή του t από τις σχέσεις:

$$f_k(t)=\sum_{j=2}^{k} p^2(1-p)^{j-2} f_y^{*j}(t)-(p-1)\sum_{u=1}^{t-1} f_k(u)P(Y=t-u) \\ -p(1-p)^{k-1}\sum_{u=1}^{t-1} f_k(u) f_Y^{*k}(t-u), \quad \text{όταν}\, t>y_0 k \tag{5.2}$$

$$f_k(t)=\sum_{j=2}^{k} p^2(1-p)^{j-2} f_y^{*j}(t)-(p-1)\sum_{u=1}^{t-1} f_k(u)P(Y=t-u), \quad \text{όταν}\, 0<t\le y_0 k\ . \tag{5.3}$$

Οι αρχικές συνθήκες που χρειάζονται είναι $f_k(0)=f_k(1)=0$.

Ενδιαφέρον παρουσιάζει η περίπτωση του τ.α με τ.μ. $Y_1, Y_2, \ldots$ να ακολουθούν κατανομή από την οικογένεια *PH*. Από τους τύπους (2.14), (2.15) και (2.16) για διακριτές κατανομές *PH* έχουμε ότι στην περίπτωση που οι τ.μ. $Y_1, Y_2, \ldots$ ακολουθούν διακριτή *PH* διάστασης *c*, με αναπαράσταση $(M, \boldsymbol{\rho})$ και η τ.μ. T_k είναι όπως έχουμε αναφέρει είναι μία διακριτή *PH* διάστασης $k+1$ με αναπαράσταση $(\Lambda, \boldsymbol{\pi})$, τότε το τυχαίο άθροισμα $S_k=\sum_{t=1}^{T_k} Y_t$ με αναπαράσταση $(\Sigma, \boldsymbol{\sigma})$, έχει παραμέτρους σύμφωνα με τον Neuts (1981)

$$\boldsymbol{\sigma}=\boldsymbol{\rho}\times\boldsymbol{\pi}(I_{k+1}-a\Lambda)^{-1}\ ,$$

και

$$\Sigma=M\times I_{k+1}+\boldsymbol{u}\,\boldsymbol{p}'\times(I_{k+1}-a\Lambda)^{-1}\Lambda\ ,$$

με τα $\alpha=1-\rho'\mathbf{1}$ και $u=(I_c-M)\mathbf{1}$.

Επομένως, προκύπτει πως η συνάρτηση πιθανότητας του συνολικού χρόνου αναμονής $S_k=\sum_{t=1}^{T_k} Y_t$ μπορεί να γραφεί στη μορφή:

$$f_k(t)=P(S_k=t)=\boldsymbol{\sigma}'\Sigma^{t-1}(I_{c(k+1)}-\Sigma)\mathbf{1}, \qquad t=1,2,\ldots\ . \tag{5.4}$$

5.3 Γεννήτριες συναρτήσεις της σύνθετης συνάρτησης σάρωσης σε εξαρτημένες τυχαίες μεταβλητές

Όπως έχουμε ορίσει ήδη, σύμφωνα με τον Bowers (1997), η πιθανογεννήτρια συνάρτηση της σύνθετης κατανομής ενός τυχαίου αθροίσματος τυχαίων μεταβλητών $S_k = \sum_{i=1}^{T_k} Y_i$, ικανοποιεί την σχέση $P_{S_k}(z) = P_{T_k}(P_Y(z))$.

Η ροπογεννήτρια συνάρτηση του τ.α μπορεί να εκφραστεί βάσει του ορισμού της ροπογεννήτριας:

$$M_{S_k}(z) = E(e^{zS_k}) \sum_{t=1}^{\infty} P[S_k = t] e^{zt} = \sum_{t=1}^{\infty} f_k(t) e^{zt}, \tag{5.5}$$

και η ροπογεννήτρια συνάρτηση της T_k δίνεται από την σχέση:

$$M_{T_k}(z) = E(e^{zT_k}) \sum_{t=1}^{\infty} P[T_k = t] e^{zt}. \tag{5.6}$$

Σύμφωνα με Balakrishman, Koutras (2002), αποδεικνύεται ότι η πιθανογεννήτρια της συνάρτησης σάρωσης σε εξαρτημένες διωνυμικές δοκιμές, δίνεται από μία σχέση της μορφής:

$$P_{T_k}(z) = z[p_1 + (p_0 p_{01} - p_1 p_{00}) z] \frac{p_{11} z + p_{01} p_{10} z^2 \sum_{j=3}^{k} p_{00} z^{j-3}}{1 - p_{00} z - p_{01} p_{10} p_{00}^{k-2} z^k}. \tag{5.7}$$

Επομένως, μπορούμε να δώσουμε σχέσεις για την πιθανογεννήτρια και ροπογεννήτρια συνάρτηση της σύνθετης συνάρτησης σάρωσης (scan statistics). Οι σχέσεις αυτές θα έχουν την μορφή:

$$P_{S_k}(z) = P_Y(z)[p_1 + (p_0 p_{01} - p_1 p_{00}) P_Y(z)] \frac{p_{11} P_Y(z) + p_{01} p_{10} (P_Y(z))^2 \sum_{j=3}^{k} (p_{00} P_Y(z))^{j-3}}{1 - p_{00} P_Y(z) - p_{01} p_{10} p_{00}^{k-2} (P_Y(z))^k} \tag{5.8}$$

και

$$M_{S_k}(z) = M_Y(z)[p_1 + (p_0 p_{01} - p_1 p_{00}) M_Y(z)] \frac{p_{11} M_Y(z) + p_{01} p_{10} (M_Y(z))^2 \sum_{j=3}^{k} (p_{00} M_Y(z))^{j-3}}{1 - p_{00} M_Y(z) - p_{01} p_{10} p_{00}^{k-2} (M_Y(z))^k}.$$

Για την ροπή πρώτης τάξης και την διασπορά έχουμε πως θα υπακούν στις γνωστές από την θεωρία κινδύνου παρακάτω σχέσεις:

$$E(S_k) = E(\sum_{t=1}^{T_k} Y_t) = E(T_k) E(Y_t),$$

και

$$Var(S_k) = Var(\sum_{t=1}^{T_k} Y_t) = E(T_k) Var(Y_t) + (E(Y_t))^2 Var(T_k).$$

Παραγωγίζοντας την σχέση (5.7) μπορούμε να παράγουμε τις σχέσεις για την μέση τιμή και την διασπορά για την συνάρτηση σάρωσης T_k .

Παράδειγμα 5.2: Αν τώρα ισχυριστούμε ότι σε μία ασφαλιστική ο ενδιάμεσος χρόνος έλευσης των ζημιών είναι ανεξάρτητος, τότε μπορούμε και πάλι βάσει των παραπάνω αποτελεσμάτων, να προσδιορίσουμε τις πιθανογεννήτριες και ροπογεννήτριες συναρτήσεις για την σύνθετη κατανομή μιας συνάρτησης σάρωσης. Αν η πιθανότητα αποτυχίας ήταν $q=1-p=P(\xi_t=0)$ για κάθε $t=1, 2, ...$ και η πιθανότητα επιτυχίας θα είναι αντίστοιχα $p=P(\xi_t=1)$, τότε όλες οι πιθανότητες μετάβασης ισούνται με $p=p_1=p_{11}=p_{01}$ και $q=1-p=p_0=p_{10}=p_{00}$.

Αντικαθιστώντας στην (5.8) προκύπτουν οι σχέσεις:

$$P_{S_k}(z)=E(z^{S_k})=\frac{(pP_Y(z))^2}{1-qP_Y(z)-pq^{k-1}P_Y(z)^k}\frac{1-(qP_Y(z))^{k-1}}{1-qP_Y(z)} ,$$

$$M_{S_k}(z)=E(z^{S_k})=\frac{(pM_Y(z))^2}{1-qM_Y(z)-pq^{k-1}M_Y(z)^k}\frac{1-(qM_Y(z))^{k-1}}{1-qM_Y(z)} ,$$

και για την μέση τιμή και διασπορά έχουμε ότι:

$$E(S_k)=E(T_k)E(Y_t)=\frac{2-q^{k-1}}{p(1-q^{k-1})}E(Y_t) ,$$

$$\begin{aligned}Var(S_k)=E(T_k)Var(Y_t)+(E(Y_t))^2 Var(T_k)=\frac{2-q^{k-1}}{p(1-q^{k-1})}Var(Y_t)\\ +(E(Y_t))^2[\frac{q}{p^2}+(2\mathrm{k}-1)\frac{q^{k-1}}{p(1-q^{k-1})^2}+\frac{q}{p^2(1-q^{k-1})^2}]\end{aligned} .$$

Παράδειγμα 5.3: Ας υποθέσουμε ότι μία τράπεζα εκτιμά πως κινδυνεύει να υποστεί κάποια χρεοκοπία στο χαρτοφυλάκιο της εάν αποτύχει σε δύο τεστ αντοχής που απέχουν το πολύ μία επιτυχία μεταξύ τους. Δηλαδή εάν το $k=3$ άρα και το $k-2=1$. Το τυχαίο άθροισμα τυχαίων μεταβλητών S_k έτσι όπως το ορίσαμε, θα μετράει τον τυχαίο χρόνο αναμονής μέχρι να εμφανιστεί στη αλληλουχία των αποτελεσμάτων της τράπεζας στα τεστ αντοχής, για πρώτη φορά ένα μοτίβο με δύο αποτυχίες που ανάμεσά τους θα μεσολαβεί μία επιτυχία. Επίσης, ας κάνουμε την υπόθεση ότι οι ενδιάμεσοι χρόνοι Y_t μεταξύ των τεστ αντοχής κατανέμονται εκθετικά με συνάρτηση πυκνότητας πιθανότητας $f_Y(y)=\theta e^{-\theta y},\ \ y>0$ και ροπογεννήτρια $M_Y(z)=E(e_t^{zY})=\frac{\theta}{\theta-z},\ \ z<\theta$.

- Αν τα τεστ αντοχής θεωρηθούν ανεξάρτητα μεταξύ τους, τότε η ροπογεννήτρια του τ.α S_3 θα είναι ίση με:

$$M_{S_3}(z)=\frac{p\theta^2[(p-2)\theta+z]}{p^2(p-2)\theta^3+(2\mathrm{p}+1)\theta^{2z}-(p+2)\theta z^2},$$

- Ενώ αν η έκβαση ενός τεστ αντοχής εξαρτάται από το προηγούμενο τότε έχουμε μία εξαρτημένη Μαρκοβιανή αλυσίδα, όπου η αντίστοιχη ροπογεννήτρια για το τ.α S_3 θα είναι ίση με:

$$M_{S_3}(z)=\frac{\theta^2(zp_1-\theta p_{01})[zp_{11}-\theta(p_{11}-p_{01}p_{10})]}{z^4-z^3\theta(p_{01}+3)+3\mathrm{z}^2\theta^2(p_{01}+1)-z\theta^3((p_{01}+1)^2+p_{01}p_{10}p_{00})+\theta^4(p_{01}p_{10}+p_{11})}.$$

Παράδειγμα 5.4: Αν υποθέσουμε τώρα ότι σε μία τράπεζα οι ενδιάμεσοι χρόνοι των τεστ αντοχής κατανέμονται βάσει μίας διακριτής κατανομής, έστω μίας γεωμετρικής με συνάρτηση πιθανότητας $P(Y_t=y)=\theta(1-\theta)^{y-1},\ y=1,\ 2,...$ και πιθανογεννήτρια $P_Y(z)=\frac{\theta z}{1-(1-\theta)z}$. Τότε για την πιθανογεννήτρια συνάρτηση του τ.α S_3 έχουμε τα παρακάτω αποτελέσματα για την περίπτωση εξαρτημένης Μαρκοβιανής αλυσίδας και ανεξάρτητης αλυσίδας.

- Στην περίπτωση των ανεξάρτητων διωνυμικών αποτελεσμάτων θα έχουμε ότι:

$$P_{S_3}(z)=\frac{(p\theta z)^3[A(z)^2-(1-p)^2\theta^2z^2]}{((p\theta-1)z+1)[A(z)^3-p(1-p)^2\theta^3z^3-(1-p)\theta z A(z)^2]},$$

  με $A(z)=(\theta-1)z+1$.

- Στην περίπτωση αποτελεσμάτων των τεστ αντοχής με εξάρτηση θα έχουμε πως η πιθανογεννήτρια του τ.α S_3 θα δίνεται από την σχέση:

$$P_{S_3}(z)=\frac{(\theta z)^2[p_1(1-z)+\theta zp_{01}][A(z)\,p_{11}+\theta zp_{01}p_{10}]}{\theta A(z)^4-A(z)^3\theta zp_{00}-A(z)(\theta z)^3p_{01}p_{00}p_{10}}.$$

5.4 Αποτελέσματα για την σύνθετη συνάρτηση σάρωσης σε εξαρτημένες διωνυμικές τυχαίες μεταβλητές

Στην ενότητα αυτή παρουσιάζονται κάποια αποτελέσματα για σύνθετη συνάρτηση σάρωσης. Όπως έχουμε αναφέρει, η περίπτωση που παρουσιάζει ενδιαφέρον είναι όταν η τ.μ T_k είναι διακριτή και ακολουθεί μία κατανομή *PH* ανώτερης τάξης *d*, δηλαδή $T \sim PH_d(\Lambda, \boldsymbol{\pi})$. Επίσης, οι τ.μ $Y_1, Y_2, \ldots$ είναι διακριτές ή συνεχείς θετικές τυχαίες μεταβλητές.

Για να μπορέσει να εφαρμοστεί ο αναδρομικός τύπος του Eisele (2006) πρέπει να υπολογιστούν οι δύο διαφορετικοί συντελεστές α_i, b_i. Οι συντελεστές b_i είναι απλώς οι συντελεστές του χαρακτηριστικού πολυωνύμου του πίνακα *Λ*. Δηλαδή,

$$det(xI_d - \Lambda) = x^d + \sum_{i=1}^{d} b_i x^{d-i}, \tag{5.9}$$

ενώ οι α_i υπολογίζονται βάσει της σχέσης:

$$\alpha_1 = P(T=1), \quad \alpha_t = P(T=t) + \sum_{i=1}^{t-1} b_i x^{d-i}. \tag{5.10}$$

Η τ.μ T_k ως μεταβλητή του ακολουθεί κατανομή *PH* θα έχει συνάρτηση πιθανότητας που θα είναι της μορφής:

$$P(T_k = n) = \boldsymbol{\pi_0}' \Lambda_0^{n-2} \sum_{i=2}^{k} \boldsymbol{e_i}, \tag{5.11}$$

με $\boldsymbol{e_i} = (0, \ldots, 0,1,0, \ldots, 0)', i=2, \ldots, k$ και αρχικό διάνυσμα $\boldsymbol{\pi_0} = (p_0, p_1, 0, \ldots, 0)'$.

Τέλος, ο πίνακας μετάβασης Λ_0 ο οποίος σύμφωνα με τους Balakrishnan, Κούτρας (2002) θα είναι της μορφής:

$$\Lambda_0 = \begin{pmatrix} p_{00} & p_{01} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{10} & 0 & \ldots & 0 & 0 & 0 & p_{11} \\ 0 & 0 & 0 & p_{00} & \ldots & 0 & 0 & 0 & p_{01} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & p_{00} & 0 & p_{01} \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & p_{00} & p_{01} \\ p_{00} & p_{01} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \end{pmatrix}_{(k+2)\times(k+2)}.$$

Η τ.μ. T_k για την περίπτωσή μας, δηλαδή όταν περιγράφει τον χρόνο αναμονής μέχρι να εμφανιστεί για πρώτη φορά δύο επιτυχίες που απέχουν μεταξύ τους το πολύ $k-2$ αποτυχίες, θα ακολουθεί μία διακριτή κατανομή *PH* τάξης $k+1$ με πίνακα μετάβασης τον υποπίνακα *Λ* του Λ_0 και αρχικό διάνυσμα το $\boldsymbol{\pi} = p_0 \boldsymbol{e_1} + p_1 \boldsymbol{e_2} = (p_0, p_1, \ldots, 0)'$. Ο υποπίνακας *Λ* είναι τάξης $(k+1)\times(k+1)$ και θα έχει την μορφή:

$$
\Lambda = \begin{pmatrix}
p_{00} & p_{01} & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\
0 & 0 & p_{10} & 0 & \dots & 0 & 0 & 0 & p_{11} \\
0 & 0 & 0 & p_{00} & \dots & 0 & 0 & 0 & p_{01} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & 0 & p_{00} & 0 & p_{01} \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & p_{00} & p_{01} \\
p_{00} & p_{01} & 0 & 0 & \dots & 0 & 0 & 0 & 0
\end{pmatrix}. \tag{5.12}
$$

Σύμφωνα με Κούτρας, Κούτρας, Yalcin (2016) αν τα αποτελέσματα ξ_t, $t=1, 2, \dots$ δημιουργούν μία ακολουθία εξαρτημένων διωνυμικών δοκιμών μίας ομογενούς μαρκοβιανής αλυσίδας δύο καταστάσεων $i, j \in \{0,1\}$ με πιθανότητα $P[\xi_t = j | \xi_{t-1} = i] p_{ij}$, $t=2, 3, \dots$ και αρχικές πιθανότητες $P(\xi_1 = 0) = p_0$ και $P(\xi_1 = 1) = 1 - p_0 = p_1$, τότε οι συντελεστές b_i του Eisele (2006) θα δίνονται από τις σχέσεις:

$$
b_1 = -p_{00},\ b_2 = \dots = b_{k-1} = b_{k+1} = 0,\ b_k = -p_{00}^{k-2} p_{01} p_{10}, \tag{5.13}
$$

οι οποίοι είναι οι συντελεστές του χαρακτηριστικού πολυωνύμου του πίνακα (5.12), προκύπτουν δηλαδή από την σχέση $det(xI_{k+1} - \Lambda) = x^{k+1} - p_{00} x^k - p_{00}^{k-2} p_{01} p_{10} x$.

Ο υπολογισμός των συντελεστών α_i βασίζεται στην σχέση (5.10), οπότε για να τους υπολογίζω χρειάζεται να παράγω τιμές της συνάρτησης πιθανότητας μιας τ.μ. $T_k \sim PH_{k+1}(\Lambda, \boldsymbol{\pi})$. Άρα, μέσω της (5.11) θα έχουμε ότι:

$$
\begin{aligned}
&P(T_k = 0) = P(T_k = 1) = 0, \\
&P(T_k = 2) = p_1 p_{11}, \\
&P(T_k = 3) = p_{01}(p_1 p_{10} + p_0 p_{11}), \\
&P(T_k = i) = p_{00}^{i-4} p_{01}[(i-3) p_0 p_{01} p_{10} + p_1 p_{00} p_{10} + p_0 p_{00} p_{11}], \quad 4 \le i \le k \\
&P(T_k = k+1) = p_0 p_{00}^{k-3} p_{01}[(k-2) p_{01} p_{10} + p_{00} p_{11}], \\
&P(T_k = i) = p_{00} P(T_k = i-1) + p_{01} p_{10} p_{00}^{k-2} P(T_k = i-k), \quad i > k+1.
\end{aligned}
$$

Άρα και οι συντελεστές α_i μέσω των (5.10), (5.13) και των παραπάνω τιμών της συνάρτησης πιθανότητας της τ.μ T_k παίρνουν τις παρακάτω τιμές:

$$
\begin{aligned}
&\alpha_1 = 0, \\
&\alpha_2 = p_1 p_{11}, \\
&\alpha_3 = p_{01}(p_1 p_{10} + p_0 p_{11}) - p_{00} p_1 p_{11}, \\
&\alpha_i = p_{00}^{i-4} p_{10} p_{01}^2 p_0, \quad 4 \le i \le k,
\end{aligned}
$$

$$\alpha_i = p_{00}^{k-3} p_{10} p_{01}^2 p_0 - p_{00}^{k-2} p_{10} p_{01} p_1, \quad i = k+1.$$

Επομένως, η αναδρομική σχέση του Eisele (2006), (2.20) για την παραγωγή τιμών από την συνάρτηση πιθανότητας του τ.α $S_k = \sum_{i=1}^{T_k} Y_i$ μπορεί σύμφωνα με τους Κούτρας, Κούτρας, Yalcin (2016) να εφαρμοστεί και στην περίπτωση που η T_k είναι μία διακριτή *PH* όπως την περιγράψαμε παραπάνω και οι ακολουθία των $Y_1, Y_2, \ldots$ είναι θετικές, ανεξάρτητες μεταξύ τους και ισόνομες τ.μ που όμως έχουν εξάρτηση. Τότε, η αναδρομική σχέση της συνάρτησης πιθανότητας θα είναι της μορφής:

$$f_k = \sum_{j=1}^{min(d,t)} a_j f_Y^{*j}(t) - \sum_{j=1}^{min(d,t-1)} b_j \left(\sum_{u=1}^{t-1} P(S_k = u) f_Y^{*j}(t-u) \right) = \sum_{j=1}^{min(k+1,t)} b_j \left(\sum_{u=1}^{t-1} f_k(u) f_Y^{*j}(t-u) \right).$$

Δηλαδή, αν το στήριγμα των τ.μ. $Y_1, Y_2, \ldots$ είναι το σύνολο $\{y_0, y_0+1, \ldots\}$, τότε για $t > y_0 k$ θα έχουμε ότι:

$$\begin{aligned} f_k(t) = {} & p_1 p_{11} f_Y^{*2}(t) + (p_1 p_{01} p_{10} + p_0 p_{01} p_{11} - p_1 p_{00} p_{11}) f_Y^{*3}(t) I_{\{k \geq 3\}} + \\ & + p_0 p_{01}^2 p_{10} \sum_{j=4}^{k} p_{00}^{j-4} f_Y^{*j}(t) + (p_0 p_{00}^{k-3} p_{01}^2 p_{10} - p_1 p_{00}^{k-2} p_{01} p_{10}) f_Y^{*k+1}(t) + \\ & + p_{00} \sum_{u=1}^{t-1} f_k(u) P(Y = t-u) + p_{00}^{k-2} p_{01} p_{10} \sum_{u=1}^{t-1} f_k(u) f_Y^{*k}(t-u) \end{aligned}, \quad (5.14)$$

και για $1 < t \leq y_0 k$ έχουμε ότι:

$$\begin{aligned} f_k(t) = {} & p_1 p_{11} f_Y^{*2}(t) + (p_1 p_{01} p_{10} + p_0 p_{01} p_{11} - p_1 p_{00} p_{11}) f_Y^{*3}(t) I_{\{k \geq 3\}} \\ & + p_0 p_{01}^2 p_{10} \sum_{j=4}^{k} p_{00}^{j-4} f_Y^{*j}(t) + p_{00} \sum_{u=1}^{t-1} f_k(u) P(Y = t-u) \end{aligned}. \quad (5.15)$$

Φυσικά αν η τ.μ. $T_k \sim PH_{k+1}(\Lambda, \boldsymbol{\pi})$ και τα διωνυμικά αποτελέσματα $\xi_t,\ t = 1, 2, \ldots$ θεωρούνται ανεξάρτητα μεταξύ τους, τότε μπορούμε να χρησιμοποιήσουμε τα παραπάνω αποτελέσματα ξανά με πιθανότητες $p = p_1 = p_{11} = p_{01}$ και $q = 1 - p = p_0 = p_{10} = p_{00}$ και να καταλήξουμε σε αναδρομική σχέση για την συνάρτηση πιθανότητας της σύνθετης κατανομής του τυχαίο άθροισμα S_k όπως αυτή δόθηκε στην ενότητα 5.2.

Παράδειγμα 5.5: Για το τέλος ας δούμε την περίπτωση όπου τα δύο μοντέλα ταυτίζονται. Όπως καταλαβαίνουμε για *k*=2 η τ.μ. T_k συνάρτηση σάρωσης ακολουθεί γεωμετρική κατανομή τάξης *k* και ταυτίζεται με την ροή επιτυχιών μήκους *k*=2.

Αν η τ.μ T_k για $k = 2$ ακολουθεί Γεωμετρική κατανομή τάξης *2* με παράμετρο *p*. Η πιθανογεννήτρια συνάρτηση της θα δίνεται από την σχέση:

$$P_{T_k}(z)=E(z^{T_k})=\frac{(pz)^2 A(z)}{1-qz-pq^{k-1}z^k},$$

με

$$A(z)=\sum_{i=0}^{k-2}(qz)^i=\frac{1-(qz)^{k-1}}{1-qz}, \quad |z|\leq 1.$$

Η πιθανογεννήτρια συνάρτηση της τ.μ. $Y_t, t\geq 2$ είναι ίση με:

$$P_Y(z)=E[z^Y]=\sum_{x=1}^{\infty}P(Y_t=x)z^x.$$

Το τυχαίο άθροισμα που περιγράφει τον συνολικό χρόνο αναμονής, έχει γεννήτρια συνάρτηση που υπολογίζεται από την σχέση, $P_{S_k}(z)=P_{T_k}(P_Y(z))$. Άρα, η πιθανογεννήτρια συνάρτηση του τ.α S_k, θα είναι μία σχέση της μορφής:

$$P_{S_k}(z)=E(z_k^S)=\frac{(pP_Y(z))^2}{1-qP_Y(z)-pq^{k-1}P_Y(z)^k}\frac{1-(qP_Y(z))^{k-1}}{1-qP_Y(z)}. \tag{5.16}$$

Στην περίπτωση που οι χρόνοι αναμονής μέχρι να συμβεί το επόμενο τεστ αντοχής περιγράφονται από μία συνεχή κατανομή, η ροπογεννήτρια του τ.α S_k είναι $M_{S_k}(z)=P_{T_k}(M_Y(z))$ Οπότε αντίστοιχα βρίσκουμε την ροπογεννήτρια της σύνθετης κατανομής του S_k :

$$M_{S_k}(z)=E(z_k^S)=\frac{(pM_Y(z))^2}{1-qM_Y(z)-pq^{k-1}M_Y(z)^k}\frac{1-(qM_Y(z))^{k-1}}{1-qM_Y(z)}.$$

Βιβλιογραφία

1. Ambagaspitiya, R.S., 1995. A family of discrete distributions. Insurance: Mathematics and Economics 16, 107–127.
2. Asmussen, S., 1996. Fitting phase-type distributions via the EM algorithm. Scandinavian Journal of Statistics, 419-441
3. Asmussen, S., Nerman, O., Olsson , M., 2000. Ruin Probabilities. World Scientific
4. Balakrishnan, N., Koutras, M. V., 2002. Runs and scans with applications. Wiley, New Yorκ
5. Bowers, N. L., Hickman, J. C., Gerber, H. U., Nesbitt, C. J. and Jones, D. A., 1997. Actuarial Mathematics (2nd Edition), Society of Actuaries, Schaumburg, Illinois.
6. Eisele, K-T., 2006. Recursions for compound phase distributions. Insurance: Math Econ 38:149–156
7. Hess, K. Th., Liewald, A., Schmidt, K., 2002. An extension of Panjer's recursion. ASTIN Bulletin 32, 283–297.
8. Hipp, C., 2004. Speedy Panjer for phase type claims. IME Congress Rome.
9. Koutras, V. M., Koutras, M. V., Femin Yalcin, 2015. A simple compound scan statistic useful for modeling insurance and risk management problems. Department of Statistics and Insurance Science, University of Piraeus, Greece
10. Koutras, V. M., Koutras, M. V., Femin Yalcin, 2015. A Compound Scan Statistic Distribution with Applications in Risk Management. 16th ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece
11. Koutras, M. V., Eryilmaz, S., 2016. Compound geometric distribution of order k. Methodology and Computing in Applied Probability.
12. Neuts, M. F., 1975. Probability distributions of phase type. In Liber Amicorum Prof. Emeritus H. Florin, pages 173–206. University of Louvain, Belgium.
13. Neuts, M. F., 1981. Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach, The Johns Hopkins University Press, Baltimore.
14. Nielsen, B. F., 2012. Lecture notes on phase-type distributions for 02407 Stochastic Processes, MMI-DTU
15. Panjer, H. H., 1981. Recursive evaluation of a family of compound distributions. ASTIN Bulletin 12, 22–26
16. Panjer, H. H., Willmot, G. E., 1992. Insurance Risk Models. Society of Actuaries, Schaumburg, IL.
17. Sundt, B., Jewell, W.S., 1981. Further results on recursive evaluation of compound distributions. ASTIN Bulletin 12, 27–39
18. Sundt, B., 1992. On some extensions of Panjer's class of counting distributions. ASTIN Bulletin 22, 61–80.
19. Sundt, B., 2002. Recursive evaluation of aggregate claims distributions. Insurance: Mathematics and Economics 30, 297–322
20. Tank F., Eryilmaz S., 2014. The distributions of sum, minima and maxima of generalized geometric random 430 variables, Statistical Papers.

21. Wang, S., Sobrero, M., 1994. Further results on Hesselager's recursive procedure for calculation of some compound distributions. ASTIN Bulletin 16, 161–166.
22. Χατζηκωνσταντινίδης, Ε., 2014. Μοντέλα Συλλογικού Κινδύνου. Σημειώσεις Π.Μ.Σ. «Αναλογιστική Επιστήμη & Διοικητική Κινδύνου», Τμήμα Στατιστικής&Ασφαλιστικής Επιστήμης Πανεπιστήμιο Πειραιώς