

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

Η κλάση κατανομών $(a,b,0)$ του Panjer: Θεωρία, γενικεύσεις
και εφαρμογές στη θεωρία συλλογικού κινδύνου.

Εμμανουήλ Γ. Μπιλάς

Διπλωματική Εργασία

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

Πειραιάς
Ιούλιος 2018

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:

- Πολίτης Κωνσταντίνος (Επιβλέπων)
- Αντζουλάκος Δημήτριος
- Χατζηκωνσταντινίδης Ευστάθιος

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

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

The Panjer $(a,b,0)$ class of distributions: theory, generalizations
and applications in collective risk theory

Emmanouil G. Bilas

MSc Dissertation

submitted to the Department of Statistics and Insurance

Science of the University of Piraeus in partial fulfilment

of the requirements for the degree of Master of Actuarial Science

and Risk Management

Piraeus, Greece

July 2018

Σελίδα 4 από 81

Στην οικογένεια μου

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή της διπλωματικής εργασίας, κ. Κ. Πολίτη, για την καθοδήγησή του με στόχο το καλύτερο δυνατό αποτέλεσμα.

Θα ήθελα να ευχαριστήσω επίσης την οικογένεια μου και τους φίλους μου για την στήριξή τους καθ' όλη την διάρκεια των σπουδών μου αλλά και σε όλους όσους συνέβαλαν σε αυτήν.

ΠΕΡΙΛΗΨΗ

Μία σημαντική ποσότητα στη θεωρία κινδύνου είναι το συνολικό ύψος των αποζημιώσεων, το οποίο εξαρτάται από το πλήθος και από το ύψος των ζημιών. Στην παρούσα εργασία μελετάμε τις περιπτώσεις όπου το πλήθος των αποζημιώσεων ανήκει σε κάποια από τις κλάσεις κατανομών $R(a, b, 0)$, $R(a, b, 1)$ και $R(a, b, l)$. Ειδικότερα για την κλάση κατανομών $R(a, b, 0)$ μελετήθηκε ένα άρθρο του Fackler στο οποίο δίνεται η συνάρτηση πιθανότητας του πλήθους των αποζημιώσεων με έναν μη-αναδρομικό τύπο. Για αυτές τις περιπτώσεις δίνονται αναδρομικοί αλλά και μη-αναδρομικοί τύποι υπολογισμού του συνολικού ύψους των αποζημιώσεων, όταν το ύψος των ατομικών αποζημιώσεων ακολουθεί είτε κάποια συνεχή είτε κάποια διακριτή κατανομή. Επίσης γίνεται ανάλυση του τρόπου με τον οποίο γίνεται η διακριτοποίηση μιας συνεχούς τ.μ. με τη χρήση της γλώσσας προγραμματισμού R. Τέλος δίνονται κάποια αριθμητικά παραδείγματα για τον υπολογισμό της συνάρτησης πιθανότητας του συνολικού ύψους των αποζημιώσεων με τη χρήση της γλώσσας προγραμματισμού R.

ABSTRACT

A significant quantity in risk theory is the total amount of compensation for a given portfolio, which depends on the number and the amount of the claims. In this dissertation we study the cases where the number of claims belongs to some of the classes of distributions $R(a, b, 0)$, $R(a, b, 1)$ and $R(a, b, l)$. In particular, for the class of distributions $R(a, b, 0)$ we study an article by Fackler (2009), in which the probability function of the number of claims is given with a non-recursive formula. For these cases, both recursive and non-recursive formulas of calculation of the total amount of compensation are given when the amount of individual claims follows either a continuous or a discrete distribution. In addition, we analyze the methods of discretization a continuous random variable using the programming language R. Finally, some numerical examples are given for calculating the probability functions of the total amount of compensation using the programming language R.

Περιεχόμενα

1	Εισαγωγή.....	14
2	Η κλάση κατανομών $R(\mathbf{a}, \mathbf{b}, \mathbf{0})$	17
	2.1 Ο αναδρομικός τύπος του Panjer	17
	2.2 Ο μη-αναδρομικός τύπος του Fackler	28
3	Η κλάση κατανομών $R(\mathbf{a}, \mathbf{b}, \mathbf{1})$	33
	3.1 Ο αναδρομικός τύπος του Panjer και βασικά μεγέθη	33
	3.2 Περικομμένες στο σημείο μηδέν κατανομές	37
	3.3 Τροποποιημένες στο σημείο μηδέν κατανομές.....	39
	3.4 Εκτεταμένη Περικομμένη Αρνητική Διωνυμική Κατανομή.....	42
4	Η κατανομή των συνολικών αποζημιώσεων.....	48
	4.1 Βασικοί Τύποι	48
	4.2 Υπολογισμός της σ.κ. με τη μέθοδο Recursive.....	52
	4.3 Διακριτοποίηση της κατανομής συνεχών τυχαίων μεταβλητών	56
	4.4 Εφαρμογές	60
	4.5 Συμπεράσματα.....	70
5	Η κλάση κατανομών $R(\mathbf{a}, \mathbf{b}, l)$	71
	5.1 Βασικοί ορισμοί και τύποι	71
	5.2 Λέπτυνση (Thinning)	75
6	Συμπεράσματα.....	77
	Παράρτημα I	78
	Βιβλιογραφία	81

1 Εισαγωγή

Ένας από τους πρωταρχικούς στόχους στην θεωρία κινδύνου είναι η αξιολόγηση των κινδύνων που συνδέονται με ένα χαρτοφυλάκιο ασφαλιστικών συμβάσεων με σύντομη διάρκεια, π.χ. ενός έτους. Ασφαλιστήρια συμβόλαια διάρκειας ενός έτους συνήθως είναι η ασφάλεια πυρός, η ασφάλιση αυτοκινήτου, ασφαλιστήρια συμβόλαια υγείας αλλά και ομαδικά ασφαλιστήρια ζωής. Συνεπώς μία ασφαλιστική εταιρία την ενδιαφέρει να μπορεί να εκτιμήσει το συνολικό κόστος των ζημιών που θα επέλθουν προκειμένου να πάρει κάποιες αποφάσεις όπως η εκτίμηση του ασφαλιστρού της για την επόμενη περίοδο ή να προβεί σε τεχνικές μείωσης του κινδύνου όπως η ανασφάλιση.

Σύμφωνα λοιπόν με τα όσα αναφέραμε αυτό που χρειάζεται μία ασφαλιστική εταιρεία είναι να μοντελοποιήσει τόσο το πλήθος των ζημιών που θα επέλθουν όσο και το ύψος των ζημιών που θα συμβούν και κατά συνέπεια το συνολικό κόστος της. Το συνολικό κόστος μίας ασφαλιστικής Εταιρείας μπορεί να υπολογιστεί είτε μέσω του ατομικού προτύπου είτε μέσω του συλλογικού προτύπου. Η βασική διαφορά τους είναι ότι στο ατομικό πρότυπο το χαρτοφυλάκιο είναι «κλειστό» και το συνολικό κόστος υπολογίζεται ως άθροισμα των επιμέρους απαιτήσεων σε αντίθεση με το συλλογικό πρότυπο στο οποίο το χαρτοφυλάκιο είναι «ανοιχτό» και το πλήθος των αποζημιώσεων είναι τυχαία μεταβλητή.

Ο πρώτος αναδρομικός τύπος υπολογισμού του συνολικού ύψους των αποζημιώσεων δόθηκε από τον H.Panjer το 1981. Ο Panjer αρχικά απέδειξε ότι το πλήθος των αποζημιώσεων ικανοποιεί μία αναδρομική σχέση όταν αυτό ακολουθεί μία από τις κατανομές Poisson, διωνυμική ή αρνητική διωνυμική για κατάλληλες σταθερές a, b και η αντίστοιχη οικογένεια κατανομών ονομάστηκε $R(a, b, 0)$.

Με αφορμή τα όσα αναφέρθηκαν παραπάνω στο δεύτερο κεφάλαιο της παρούσας εργασίας μελετάται η κλάση κατανομών $R(a, b, 0)$. Αρχικά υπολογίζονται οι παράμετροι a και b για κάθε μία από τις κατανομές που ανήκει στην οικογένεια αυτή. Στη συνέχεια με την χρήση ροπογεννητριών καταλήγουμε σε κάποια μέτρα κινδύνου όπως η μέση τιμή, η διακύμανση και ο συντελεστής ασυμμετρίας συναρτήσει των παραμέτρων που αναφέραμε.

Ακόμη γίνεται ανάλυση ενός μη-αναδρομικού τύπου για τον υπολογισμό της συνάρτησης πιθανότητας του πλήθους των αποζημιώσεων που δόθηκε από τον M. Fackler το 2009 για τις κατανομές που ανήκουν στην κλάση κατανομών $R(a, b, 0)$.

Τέλος γίνεται μία εφαρμογή στην οποία αποδεικνύεται ότι ο τύπος του Fackler και αυτός του Panjer μας δίνουν τα ίδια αποτελέσματα.

Στη συνέχεια στο τρίτο κεφάλαιο μελετάται μία ευρύτερη κλάση κατανομών, η $R(a, b, 1)$, για την οποία αρχικά θα δούμε ποιες κατανομές ανήκουν σε αυτήν. Επιπλέον δίνεται ο αναδρομικός τύπος υπολογισμού της συνάρτησης πιθανότητας καθώς και οι παράμετροι a, b και στη συνέχεια, συναρτήσει αυτών, δίνονται κάποιες χρήσιμες σχέσεις για τον υπολογισμό της μέσης τιμής, της δεύτερης ροπής και της διακύμανσης. Επιπροσθέτως μελετώνται οι κατανομές που ανήκουν σε αυτή την κλάση και ειδικότερα η εκτεταμένη περικομμένη αρνητική διωνυμική κατανομή αλλά και η λογαριθμική κατανομή.

Στο τέταρτο κεφάλαιο αρχικά δίνονται οι αναδρομικοί τύποι για τον υπολογισμό της σ.π. της τ.μ. S , που εκφράζει το συνολικό μέγεθος των αποζημιώσεων στο διάστημα που εξετάζουμε, στην περίπτωση όπου το πλήθος των ζημιών ανήκει στην κλάση $R(a, b, 0)$ ή στην περίπτωση όπου ανήκει στην κλάση $R(a, b, 1)$ και το ύψος των αποζημιώσεων παίρνει θετικές ακέραιες τιμές. Εκτός όμως από διακριτές κατανομές το ύψος των αποζημιώσεων μπορεί να ακολουθεί και κάποια συνεχή κατανομή. Στην περίπτωση που το πλήθος των αποζημιώσεων ακολουθεί γεωμετρική, διωνυμική ή αρνητική διωνυμική κατανομή και το ύψος των αποζημιώσεων ακολουθεί εκθετική κατανομή υπάρχει ακριβής τύπο της σ.π.π. και της σ.κ. για τις συνολικές αποζημιώσεις.

Δεδομένης της δυσκολίας η οποία προκύπτει από το γεγονός ότι έχουμε να εργαστούμε με αναδρομικούς τύπους, στη συνέχεια της εργασίας γίνεται αναλυτική περιγραφή του τρόπου με τον οποίο υπολογίζεται η σ.κ. της τ.μ. S όταν η κατανομή του πλήθους των αποζημιώσεων ανήκει στην κλάση $R(a, b, 1)$ με τη χρήση της γλώσσας προγραμματισμού R.

Λόγω του ότι ο αναδρομικός τύπος του Panjer που δόθηκε ισχύει μόνο για τις περιπτώσεις που το ύψος των αποζημιώσεων ακολουθεί διακριτή κατανομή, γίνεται ανάλυση του τρόπου με τον οποίο μπορεί κάποιος να διακριτοποιήσει την κατανομή μίας συνεχούς τ.μ., με τη χρήση της γλώσσας προγραμματισμού R, με σκοπό να μπορεί να χρησιμοποιηθεί στους αναδρομικούς τύπους που αναφέρθηκαν. Τέλος γίνονται εφαρμογές στις οποίες μελετώνται οι μέθοδοι διακριτοποίησης προκειμένου να ελέγξουμε αν με κάποια συγκεκριμένη μέθοδο παίρνουμε πάντα καλύτερα αποτελέσματα, αλλά και τις διαφορές μεταξύ των «πραγματικών» αποτελεσμάτων σε σχέση με αυτά που προκύπτουν έπειτα από τη

διακριτοποίηση για την συνάρτηση πιθανότητας του συνολικού ύψους των αποζημιώσεων.

Στο τελευταίο κεφάλαιο δίνονται κάποιοι βασικοί τύποι και ορισμοί για την κλάση κατανομών $R(a, b, l)$. Επίσης παρουσιάζονται οι κατανομές που ανήκουν σε αυτή την κλάση και δίνεται ο αναδρομικός τύπος της σ.π. της τ.μ. S . Κλείνοντας γίνεται αναφορά στην έννοια της λέπτυνσης (thinning) η οποία έχει εφαρμογή σε διάφορους τομείς της ασφάλισης όπως η αντασφάλιση.

2 Η κλάση κατανομών $R(a, b, 0)$

Στο παρόν κεφάλαιο αρχικά θα ορίσουμε την αναδρομική σχέση που ισχύει για την συνάρτηση πιθανότητας για τις κατανομές που ανήκουν στην κλάση κατανομών του Panjer $R(a, b, 0)$. Έπειτα θα αποδείξουμε ποιες κατανομές ανήκουν σε αυτήν την κλάση και θα υπολογίσουμε τις παραμέτρους a και b για κάθε μία από αυτές. Στη συνέχεια θα παρατεθούν τύποι για την μέση τιμή, τη δεύτερη ροπή, τη διακύμανση καθώς και για τον συντελεστή ασυμμετρίας, υπολογιζόμενοι με την χρήση ροπογεννητριών, συναρτήσεων των a και b . Ακόμη με τη βοήθεια των παραπάνω θα υπολογιστούν αντίστοιχα ο συντελεστής μεταβλητότητας και ο συντελεστής διακύμανσης. Επιπλέον θα παρουσιαστεί και θα αναλυθεί ένας μη-αναδρομικός τύπος που προτάθηκε από τον Fackler για τις κλάσεις κατανομών $R(a, b, 0)$ και θα δοθούν οι αντίστοιχοι τύποι για όλες τις παραπάνω ποσότητες. Τέλος, θα γίνει ένα παράδειγμα για κάθε κατανομή που ανήκει σε αυτή την κλάση με τη χρήση της γλώσσας προγραμματισμού R, με σκοπό να δείξουμε ότι και με τους δύο τύπους που προαναφέραμε καταλήγουμε όπως είναι αναμενόμενο στην ίδια συνάρτηση πιθανότητας.

2.1 Ο αναδρομικός τύπος του Panjer

Στη θεωρία κινδύνου, στο μοντέλο συλλογικού κινδύνου η τ.μ.

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

εκφράζει το ύψος των συνολικών ζημιών για ένα χαρτοφυλάκιο κινδύνων. Οι τ.μ $X_i \geq 0$, $i = 1, 2, 3, \dots$ συμβολίζουν το ύψος της ατομικής ζημιάς και η τ.μ. $N \in \{0, 1, 2, \dots\}$ δηλώνει το πλήθος των ζημιών του χαρτοφυλακίου. Θα δείξουμε μία αναδρομική σχέση που ικανοποιείται για συγκεκριμένες κατανομές της διακριτής τ.μ N .

Ορισμός 2.1.1

Μία διακριτή τ.μ. $N \in \{0, 1, 2, \dots\}$ θα λέμε ότι ανήκει στην οικογένεια κατανομών $R(a, b, 0)$, αν η συνάρτηση πιθανότητας της υπολογίζεται από την αναδρομική σχέση

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, \quad n = 1, 2, 3, \dots \quad (2.1.1)$$

όπου a και b είναι σταθερές και $p_n = 0 \quad \forall n < 0$.

■

Οι μοναδικές κατανομές που ανήκουν σε αυτή την κλάση είναι η Poisson (P), η Διωνυμική (B1), η Αρνητική Διωνυμική (NB1) και η Γεωμετρική (ειδική περίπτωση της αρνητικής διωνυμικής για $r = 1$). Για κάθε μία από τις κατανομές που αναφέραμε θα υπολογίσουμε τις σταθερές a , b καθώς και την τιμή της $p_0 = P(N = 0)$ που χρειάζεται για να οριστεί πλήρως η κατανομή της τ.μ. N (βλέπε Χατζηκωνσταντινίδης (2015)).

Poisson

Αν η τ.μ. $N \sim \text{Poisson}(\lambda)$, έπεται ότι

$$\frac{p_n}{p_{n-1}} = \frac{e^{-\lambda} \frac{\lambda^n}{n!}}{e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!}} = \frac{\lambda}{n} \Leftrightarrow p_n = \frac{\lambda}{n} p_{n-1}$$

και εξισώνοντας τους συντελεστές της παραπάνω σχέσης με την (2.1.1), παίρνουμε ότι $a = 0$ και $b = \lambda$. Ακόμη, $p_0 = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda}$.

Διωνυμική

Αν η τ.μ. $N \sim B(m, p)$, έπεται ότι

$$\begin{aligned} \frac{p_n}{p_{n-1}} &= \frac{\binom{m}{n} p^n q^{m-n}}{\binom{m}{n-1} p^{n-1} q^{m-n+1}} = \frac{\frac{m!}{n!(m-n)!} p}{\frac{m!}{(n-1)!(m-n+1)!} q} \\ &= \frac{p(m-n+1)}{nq} = -\frac{p}{q} + \frac{p(m+1)}{qn} \\ \Leftrightarrow p_n &= \left(-\frac{p}{q} + \frac{p(m+1)}{qn} \right) p_{n-1} \end{aligned}$$

και σύμφωνα με την σχέση (2.1.1), προκύπτει ότι $a = -\frac{p}{q}$ και $b = (m+1) \frac{p}{q}$.
Ακόμη $p_0 = \binom{m}{0} p^0 q^{m-0} = q^m$.

Αρνητική Διωνυμική

Αν η τ.μ. $N \sim NB(r, p)$, έπεται ότι

$$\begin{aligned} \frac{p_n}{p_{n-1}} &= \frac{\binom{r+n-1}{n} p^r q^n}{\binom{r+(n-1)-1}{n-1} p^r q^{n-1}} = \frac{\frac{(r+n-1)!}{n!(r+n-1-n)!} q}{\frac{(r+n-2)!}{(n-1)!(r+n-2-n+1)!}} = \frac{r+n-1}{n} q \\ &= q + \frac{r-1}{n} q \\ &\Leftrightarrow p_n = \left(q + \frac{r-1}{n} q \right) p_{n-1} \end{aligned}$$

επομένως από την σχέση (2.1.1), παίρνουμε ότι $a = q$ και $b = (r-1)q$.

Ακόμη $p_0 = \binom{r+0-1}{0} p^r q^0 = p^r$.

Γεωμετρική

Η γεωμετρική κατανομή είναι ειδική περίπτωση της αρνητικής διωνυμικής για $r = 1$, συνεπώς από τα αποτελέσματα που βρήκαμε προηγουμένως προκύπτει ότι $a = q$, $b = 0$ και $p_0 = p$.

■

Αφού έχουμε υπολογίσει τα a, b και p_0 για τις παραπάνω κατανομές, στη συνέχεια θα βρούμε κατάλληλες σχέσεις για τον υπολογισμό της μέσης τιμής, της δεύτερης ροπής, της διακύμανσης καθώς και τον συντελεστή ασυμμετρίας. Γνωρίζουμε ότι η σχέση μεταξύ της ροπογεννήτριας και των ροπών μίας τ.μ. είναι η εξής:

$$E(X^r) = M^{(r)}(0) = \left[\frac{d^r}{dt^r} M(t) \right]_{t=0}, \quad r = 1, 2, \dots \quad (2.1.2)$$

και συνεπώς εύκολα προκύπτουν τα ακόλουθα αποτελέσματα,

$$E(X) = M^{(1)}(0),$$

$$E(X^2) = M^{(2)}(0),$$

$$\text{Var}(X) = M^{(2)}(0) - [M^{(1)}(0)]^2,$$

τα οποία θα τα χρειαστούμε στη συνέχεια.

Αυτό που παρατηρείται είναι ότι οι σχέσεις της μέσης τιμής και της διακύμανσης, για τις κατανομές της κλάσης $(a, b, 0)$, συναρτήσεων των a, b αποδεικνύονται με την

χρήση πιθανογεννητριών. Με αφορμή το γεγονός αυτό θα υπολογιστούν η μέση τιμή, η δεύτερη ροπή, η διακύμανση καθώς και ο συντελεστής ασυμμετρίας με τη χρήση ροπογεννητριών.

Αρχικά στο επόμενο λήμμα θα δειχθεί μία σχέση που θα μας φανεί χρήσιμη, η οποία συνδέει την πρώτη παράγωγο της ροπογεννήτριας με την ροπογεννήτρια.

Λήμμα 2.1.1 Αν η τ.μ. N ανήκει στην κλάση κατανομών $(a, b, 0)$, τότε η ροπογεννήτρια συνάρτηση $M_N(t)$ της τ.μ. N , ικανοποιεί τη διαφορική εξίσωση

$$M'_N(t) = \frac{(a+b)e^t}{1-ae^t} M_N(t). \quad (2.1.3)$$

Απόδειξη

Από τον ορισμό της ροπογεννήτριας έχουμε ότι,

$$\begin{aligned} M_N(t) &= E(e^{tN}) = \sum_{n=0}^{\infty} e^{tn} p_n = p_0 + \sum_{n=1}^{\infty} e^{tn} p_n \\ &= p_0 + \sum_{n=1}^{\infty} e^{tn} \left(a + \frac{b}{n}\right) p_{n-1} \\ &= p_0 + a \sum_{n=1}^{\infty} e^{tn} p_{n-1} + b \sum_{n=1}^{\infty} \frac{e^{tn}}{n} p_{n-1} \\ &= p_0 + ae^t M_N(t) + b \sum_{n=1}^{\infty} \frac{e^{tn}}{n} p_{n-1} \\ \Leftrightarrow M_{N(t)}(1 - ae^t) &= p_0 + b \sum_{n=1}^{\infty} \frac{e^{tn}}{n} p_{n-1}. \end{aligned}$$

Επομένως αν παραγωγίσουμε ως προς t την παραπάνω σχέση, παίρνουμε

$$\begin{aligned} M'_N(t)(1 - ae^t) + M_N(t)(-ae^t) &= b \sum_{n=1}^{\infty} e^{tn} p_{n-1} \\ \Leftrightarrow (1 - ae^t)M'_N(t) &= be^t M_N(t) + ae^t M_N(t), \end{aligned}$$

από όπου τελικά προκύπτει ότι,

$$M'_N(t) = \frac{(a+b)e^t}{1-ae^t} M_N(t).$$



Χρησιμοποιώντας το παραπάνω λήμμα θα αποδείξουμε κάποιες σημαντικές ποσότητες που αφορούν την κλάση κατανομών $R(a, b, 0)$.

Πρόταση 2.1.1

Αν η τ.μ. N ανήκει στην κλάση κατανομών $(a, b, 0)$, τότε η μέση τιμή, δεύτερη ροπή, η διακύμανσή της και ο συντελεστής ασυμμετρίας είναι αντίστοιχα:

$$\alpha) E(N) = \frac{a+b}{1-a}$$

$$\beta) E(N^2) = \frac{(a+b)(1+a+b)}{(1-a)^2}$$

$$\gamma) \text{Var}(N) = \frac{a+b}{(1-a)^2}$$

$$\delta) \gamma = \frac{a+1}{\sqrt{a+b}}$$

Απόδειξη

α) Ισχύει ότι $E(N) = M'_N(0)$. Από την (2.1.3) για $t = 0$, έχουμε

$$M'_N(0) = \frac{(a+b)e^0}{1-ae^0} M_N(0)$$

Επίσης είναι $M_N(0) = 1$, οπότε έπεται ότι

$$E(N) = \frac{a+b}{1-a}. \quad (2.1.4)$$

β) Παραγωγίζοντας την (2.1.3) ως προς t παίρνουμε

$$\begin{aligned} M''_N(t) &= (a+b) \frac{(e^t M_N(t) + e^t M'_N(t))(1-ae^t) - e^t M_N(t)(1-ae^t)'}{(1-ae^t)^2} \\ &= (a+b) \frac{e^t M_N(t)(1-ae^t) + (a+b)e^{2t} M_N(t) + ae^{2t} M_N(t)}{(1-ae^t)^2} \\ &= (a+b)e^t \frac{M_N(t) - ae^t M_N(t) + ae^t M_N(t) + be^t M_N(t) + ae^t M_N(t)}{(1-ae^t)^2} \end{aligned}$$

$$= \frac{e^t(a+b)(1+ae^t+be^t)}{(1-ae^t)^2} M_N(t). \quad (2.1.5)$$

Συνεπώς για $t = 0$ από την παραπάνω σχέση προκύπτει ότι

$$M''_N(0) = E(N^2) = \frac{(a+b)e^0 M_N(0)(1+ae^0+be^0)}{(1-a)^2}$$

Επομένως προκύπτει ότι,

$$E(N^2) = \frac{(a+b)(1+a+b)}{(1-a)^2}. \quad (2.1.6)$$

γ) Η διακύμανση δίνεται από τον τύπο

$$Var(N) = E(N^2) - E^2(N).$$

Αντικαθιστώντας στον παραπάνω τύπο τις σχέσεις (2.1.5) και (2.1.3), παίρνουμε

$$\begin{aligned} Var(N) &= \frac{(a+b)(1+a+b)}{(1-a)^2} - \left(\frac{a+b}{1-a}\right)^2 \\ &= \frac{(a+b)(a+b+1-a-b)}{(1-a)^2} \\ &= \frac{a+b}{(1-a)^2}. \end{aligned} \quad (2.1.7)$$

δ) Για να βρούμε τον συντελεστή ασυμμετρίας θα χρησιμοποιήσουμε τον παρακάτω τύπο

$$\begin{aligned} \gamma &= \frac{E(N-\mu)^3}{\sigma^3} = \frac{E(N^3 - 3N^2\mu + 3N\mu^2 - \mu^3)}{\sigma^3} \\ &= \frac{E(N^3) - 3\mu E(N^2) + 3\mu^2 E(N) - \mu^3}{\sigma^3} \\ &= \frac{E(N^3) - 3E(N)E(N^2) + 3E^2(N)E(N) - E^3(N)}{\sigma^3} \\ &= \frac{E(N^3) - E^3(N) - 3E(N)Var(N)}{\frac{3}{2}\sqrt{Var(N)}} \end{aligned} \quad (2.1.8)$$

Αρχικά θα υπολογίσουμε την ποσότητα $M'''_N(0) = E(N^3)$, οπότε παραγωγίζοντας την (2.1.5) ως προς t και θέτοντας τον $t=0$ θα βρούμε την τρίτη ροπή της τ.μ. N .

Για να γίνουν ευκολότερα οι υπολογισμοί θα θέσουμε

$$k(t) = M_N(t)(e^t + ae^{2t} + be^{2t})$$

και αν παραγωγίσουμε αυτή τη συνάρτηση ως προς t έχουμε

$$k'(t) = M'_N(t)(e^t + ae^{2t} + be^{2t}) + M_N(t)(e^t + 2ae^{2t} + 2be^{2t}).$$

Για $t = 0$, οι δύο παραπάνω σχέσεις μας δίνουν αντίστοιχα

$$k(0) = M_N(0)(e^0 + ae^0 + be^0)$$

$$= 1 + a + b$$

$$k'(0) = M'_N(0)(e^0 + ae^0 + be^0) + M_N(0)(e^0 + 2ae^0 + 2be^0)$$

$$= \frac{a+b}{1-a}(1+a+b) + (1+2a+2b).$$

Για την τρίτη παράγωγο της $M_N(t)$, προκύπτει με παραγωγή της σχέσης (2.1.5) ότι

$$\begin{aligned} M'''_N(t) &= (a+b) \frac{k'(t)(1-ae^t)^2 - k(t)[(1-ae^t)^2]'}{(1-ae^t)^4} \\ &= (a+b) \frac{k'(t)(1-ae^t)^2 + 2ae^t(1-ae^t)k(t)}{(1-ae^t)^4} \\ &= (a+b) \frac{k'(t)(1-ae^t) + 2ae^t k(t)}{(1-ae^t)^3} \end{aligned}$$

Από την παραπάνω σχέση για $t = 0$ παίρνουμε

$$M'''_N(0) = E(N^3)$$

$$= (a+b) \frac{k'(0)(1-ae^0) + 2ae^0 k(0)}{(1-ae^0)^3}$$

$$= (a+b) \frac{\left[\frac{a+b}{1-a}(1+a+b) + (1+2a+2b) \right] (1-a) + 2a(1+a+b)}{(1-a)^3}$$

$$= (a+b) \frac{(a+b)(1+a+b) + (1+2a+2b)(1-a) + 2a(1+a+b)}{(1-a)^3}$$

$$\begin{aligned}
&= (a+b) \frac{(1+a+b)(3a+b) + (1+2a+2b)(1-a)}{(1-a)^3} \\
&= \frac{(a+b)}{(1-a)^3} (3a + 3a^2 + 3ab + b + ab + b^2 + 1 - a + 2a - 2a^2 + 2b \\
&\quad - 2ab) \\
&= \frac{(a+b)(4a + a^2 + 2ab + 3b + b^2 + 1)}{(1-a)^3} \tag{2.1.9}.
\end{aligned}$$

Αντικαθιστώντας στην (2.1.8) τις σχέσεις (2.1.4), (2.1.7) και (2.1.9) έχουμε ότι

$$\begin{aligned}
\gamma &= \frac{\frac{(a+b)(4a + a^2 + 2ab + 3b + b^2 + 1)}{(1-a)^3} - \frac{(a+b)^3}{(1-a)^3} - 3 \frac{a+b}{1-a} \frac{a+b}{(1-a)^2}}{\sqrt{\frac{a+b}{(1-a)^2}}} \\
&= \frac{\frac{(a+b)(4a + a^2 + 2ab + 3b + b^2 + 1 - a^2 - 2ab - b^2 - 3a - 3b)}{(1-a)^3}}{\frac{(a+b)\sqrt{a+b}}{(1-a)^3}},
\end{aligned}$$

οπότε μετά από πράξεις προκύπτει ότι

$$\gamma = \frac{a+1}{\sqrt{a+b}},$$

η οποία είναι η ζητούμενη σχέση.

■

Αν τώρα χρησιμοποιήσουμε τα αποτελέσματα που αποδείξαμε στην παραπάνω πρόταση, μπορούμε να υπολογίσουμε δύο ακόμα βασικές ποσότητες που είναι η σχετική διακύμανση και ο συντελεστής μεταβλητότητας.

Σχετική διακύμανση (squared coefficient of variation) καλείται το τετράγωνο του συντελεστή μεταβλητότητας. Η ποσότητα αυτή μας δίνει τη διακύμανση των τιμών ως προς το μέσο και ο τύπος είναι

$$CV^2 = \frac{Var(N)}{E^2(N)}.$$

Συνεπώς για την κλάση κατανομών $(a, b, 0)$ θα έχουμε ότι

$$CV^2 = \frac{\frac{a+b}{(1-a)^2}}{\frac{(a+b)^2}{(1-a)^2}} = \frac{1}{a+b} \quad (2.1.10).$$

Εν συνεχεία θα υπολογίσουμε και το συντελεστή διακύμανσης (dispersion coefficient) που δίνεται από τον παρακάτω τύπο

$$D(N) = \frac{Var(N)}{E(N)} = \frac{\frac{a+b}{(1-a)^2}}{\frac{a+b}{1-a}} = 1-a, \quad (2.1.11)$$

και ο οποίος είναι πάντα μεγαλύτερος του μηδέν διότι το a παίρνει τιμές μικρότερες της μονάδας.

Κλείνοντας την ενότητα αυτή ακολουθεί ένα παράδειγμα στο οποίο θα συγκριθούν τα παραπάνω μεγέθη για όλες τις κατανομές που ανήκουν στην κλάση $R(a, b, 0)$, λαμβάνοντας ως παραδοχή ότι έχουν την ίδια μέση τιμή.

Παράδειγμα 2.1.1

Αρχικά θεωρούμε ότι και οι 4 κατανομές που θα συγκριθούν έχουν την ίδια μέση τιμή, η οποία είναι ίση με 4.

Poisson

Αφού γνωρίζουμε ότι η μέση τιμή είναι ίση με 4 τότε και η παράμετρος λ και κατ' επέκταση το b ισούται με 4. Η παράμετρος a για την Poisson ισούται με 0.

Από τις σχέσεις (2.1.7), (2.1.10), (2.1.10) και (2.1.11) προκύπτουν αντίστοιχα τα εξής,

$$Var(N) = 4$$

$$\gamma = 0.5$$

$$CV^2(N) = 0.25$$

$$D(N) = 1$$

Διωνυμική

Στην περίπτωση όπου η διωνυμική έχει παραμέτρους $m = 20$, $p = 0.2$ τότε η μέση τιμή της ισούται με 4. Αρχικά ισχύει ότι,

$$a = -\frac{p}{q} = -0.25$$

$$b = (m + 1)\frac{p}{q} = 5.25 .$$

Χρησιμοποιώντας τις ίδιες σχέσεις όπως παραπάνω παίρνουμε τα ακόλουθα αποτελέσματα:

$$Var(N) = 3.2$$

$$\gamma = 0.3354$$

$$CV^2(N) = 0.2$$

$$D(N) = 1.25$$

Αρνητική Διωνυμική

Αν για την αρνητική διωνυμική κατανομή πάρουμε σαν παραμέτρους $r = 4$ και $p=0.5$ τότε καταλήγουμε και σε αυτή την περίπτωση ότι η μέση τιμή θα ισούται με 4. Τότε έχουμε,

$$a = q = 0.5$$

$$b = (r - 1)q = 1.5 .$$

Ακολουθώντας την ίδια διαδικασία όπως προηγουμένως προκύπτει ότι,

$$Var(N) = 8$$

$$\gamma = 1.0607$$

$$CV^2(N) = 0.5$$

$$D(N) = 0.5$$

Γεωμετρική

Τέλος για την γεωμετρική κατανομή αν η παράμετρος p ισούται με 0.2 τότε η μέση τιμή της κατανομής ισούται με 4. Η παράμετρος b ισούται με 0, ενώ η παράμετρος a είναι ίση με $q = 0.8$.

Σύμφωνα λοιπόν με όσα προαναφέραμε προκύπτουν οι ακόλουθες ισότητες,

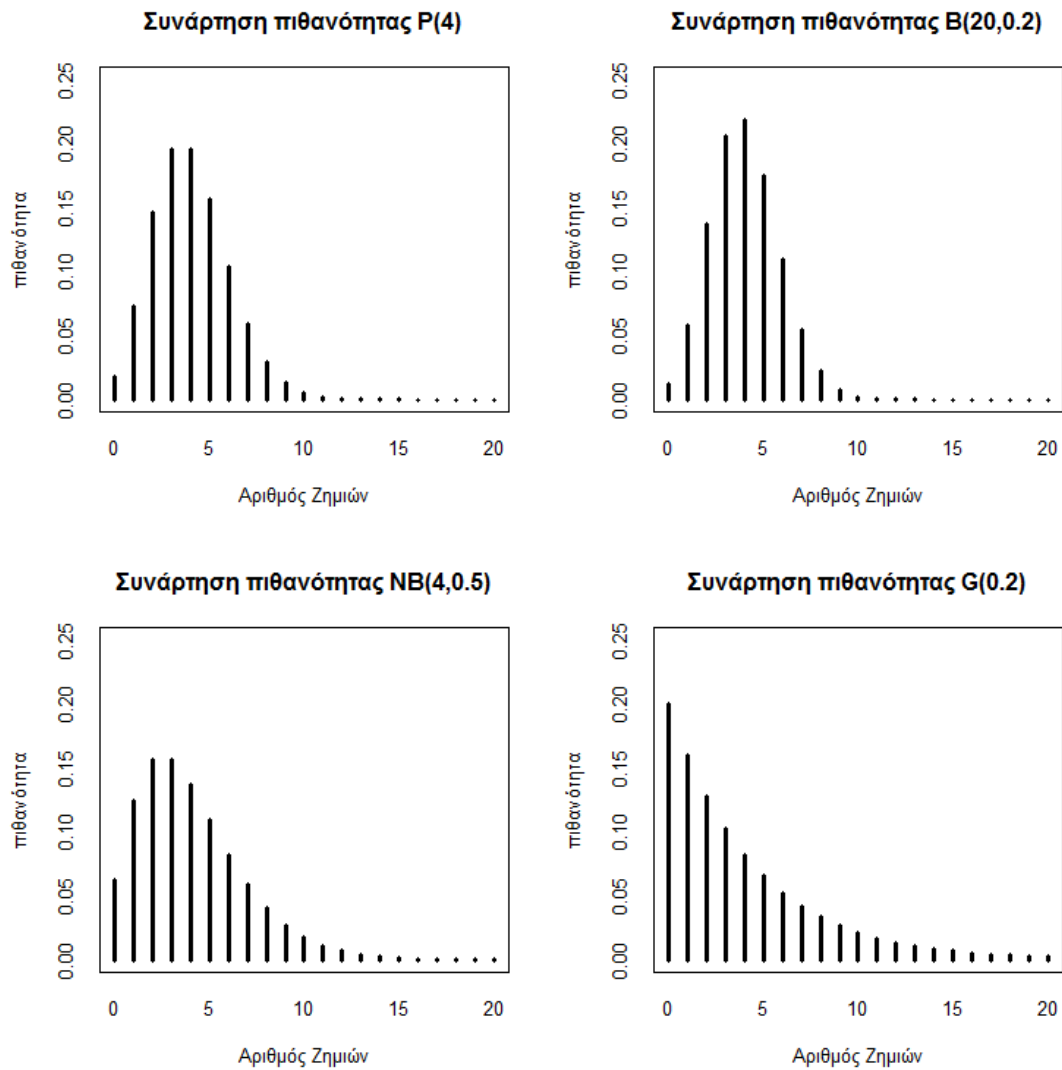
$$Var(N) = 20$$

$$\gamma = 2.012$$

$$CV^2(N) = 1,25$$

$$D(N) = 0.2$$

Πριν προχωρήσουμε στο σχολιασμό των παραπάνω αποτελεσμάτων παρατίθεται για κάθε μία από τις προαναφερόμενες κατανομές η γραφική παράσταση της συνάρτησης πιθανότητας.



Σχήμα 2.1.1

Ο κώδικας που χρησιμοποιήθηκε για την κατασκευή των γραφημάτων υπάρχει στο Παράρτημα Ι.

Αρχικά αυτό που παρατηρείται από τους υπολογισμούς μας είναι ότι μικρότερη διακύμανση εμφανίζει η poisson και μετά ακολουθούν η διωνυμική, η αρνητική διωνυμική και τέλος η γεωμετρική.

Από τα παραπάνω γραφήματα φαίνεται επίσης ότι όλες οι προαναφερθείσες κατανομές έχουν θετική ασυμμετρία, με την poisson και τη διωνυμική να είναι πιο συμμετρικές κάτι το οποίο επαληθεύεται, μέσω των παραπάνω σχέσεων, αφού οι συντελεστές ασυμμετρίας υπολογίστηκαν 0.5 και 0.3354 αντίστοιχα. Επίσης μεγαλύτερο συντελεστή ασυμμετρίας από τις προηγούμενες 2 κατανομές παρουσιάζει η αρνητική διωνυμική και ακόμη πιο ασύμμετρη είναι η γεωμετρική.

2.2 Ο μη-αναδρομικός τύπος του Fackler

Σε αυτή την ενότητα θα παρουσιάσουμε και θα αναλύσουμε έναν άλλο μη-αναδρομικό τύπο που απέδειξε ο Michael Fackler για τις κατανομές που ανήκουν στη κλάση $R(a, b, 0)$. Στη βιβλιογραφία υπάρχουν πολλές διαφορετικές παραμετροποιήσεις για τις κατανομές Panjer. Στο συγκεκριμένο άρθρο ο Fackler (2009) χρησιμοποίησε την αναμενόμενη ή μέση τιμή λ ως παράμετρο, όπου $\lambda > 0$ (την περίπτωση όπου το $\lambda = 0$ δεν θα την εξετάσουμε γιατί όλες οι κατανομές συμπίπτουν).

Ξεκινάμε ορίζοντας τις νέες παραμέτρους για καθμία από τις τρεις κατανομές που ανήκουν στη κλάση $R(a, b, 0)$:

Στην Poisson (P) χρησιμοποιούμε εξ ορισμού τη μέση τιμή ως μόνη παράμετρο. Η συνάρτηση πιθανότητας της είναι

$$p_n = e^{-\lambda} \frac{\lambda^n}{n!}, \quad (2.2.1)$$

η μέση τιμή είναι

$$E(N) = \lambda,$$

ενώ η πιθανογεννήτρια είναι

$$P_N(z) = e^{\lambda(z-1)}.$$

Στην Διωνυμική κατανομή (B1) γενικά χρησιμοποιούμε σαν παραμέτρους τον αριθμό δοκιμών m και την πιθανότητα επιτυχίας p . Επομένως θα ορίσουμε την Διωνυμική κατανομή (B2), όπου το p θα αντικατασταθεί από την αναμενόμενη τιμή $\lambda = mp$, με $p < 1$ και $\lambda < m$. Η συνάρτηση πιθανότητας, η μέση τιμή και η ροπογεννήτρια της B2 είναι αντίστοιχα:

$$p_n = \binom{m}{n} \frac{\lambda^n (m - \lambda)^{m-n}}{m^m} \quad (2.2.2)$$

$$E(N) = \lambda$$

$$P_N(z) = \left(1 + \frac{\lambda}{m}(z-1)\right)^m$$

Στην Αρνητική Διωνυμική κατανομή (NB1) το α ορίζεται ως ο αριθμός των επιτυχιών και το $p < 1$ ως η πιθανότητα επιτυχίας. Για την Αρνητική Διωνυμική κατανομή (NB2) θα αντικαταστήσουμε το p με το $\lambda = \alpha(1-p)/p$. Η συνάρτηση πιθανότητας, η μέση τιμή και η ροπογεννήτρια με της NB2 είναι αντίστοιχα:

$$p_n = \binom{\alpha+n-1}{n} \left(\frac{\alpha}{\alpha+\lambda}\right)^\alpha \left(\frac{\lambda}{\lambda+\alpha}\right)^n \quad (2.2.3)$$

$$E(N) = \lambda$$

$$P_N(z) = \left(1 + \frac{\lambda}{\alpha}(z-1)\right)^{-\alpha}$$

Πρόταση 2.2.1

Αν η τ.μ. N ανήκει στην κλάση κατανομών $R(a, b, 0)$, τότε η συνάρτηση πιθανότητας δίνεται από τον τύπο

$$p_n = \left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa} \frac{\lambda^n}{n!} \prod_{i=0}^{n-1} \frac{\kappa+i}{\kappa+\lambda}, \quad n = 0, 1, 2, \dots \quad (2.2.4)$$

όπου το $\lambda > 0$ είναι μέση τιμή. Επίσης το κ παίρνει τις εξής τιμές:

- i) Αν το $\kappa \in (0, \infty)$, τότε η κατανομή είναι η αρνητική διωνυμική.
- ii) Αν το $\kappa \rightarrow \infty$ ή το $\kappa \rightarrow -\infty$, τότε η κατανομή είναι η Poisson.
- iii) Αν το $\kappa \in (-\infty, \lambda)$, τότε η κατανομή είναι η διωνυμική. Ακόμη η παράμετρος κ περιορίζεται στους ακεραίους $-m$ του διαστήματος αυτού με την συνθήκη που πρέπει να ισχύει να είναι το $m > \lambda$.

Απόδειξη

Έστω ότι η $N \sim NB2$, τότε η συνάρτηση πιθανότητας δίνεται από την σχέση (2.2.3), όπου ισχύουν οι παρακάτω ισότητες,

$$\binom{\kappa+n-1}{n} = \frac{(\kappa+n-1)!}{n! (\kappa+n-1-n)!} = \frac{(\kappa+n-1) \dots (\kappa-1)!}{n! (\kappa-1)!} = \frac{1}{n!} \prod_{i=0}^{n-1} (\kappa+i)$$

$$\left(\frac{\kappa}{\kappa + \lambda}\right)^\kappa = \left(\frac{\lambda + \kappa}{\kappa}\right)^{-\kappa} = \left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa}$$

$$\left(\frac{\lambda}{\kappa + \lambda}\right)^n = \lambda^n \prod_{i=0}^{n-1} \frac{1}{\kappa + \lambda}$$

Συνεπώς αποδείξαμε ότι συνάρτηση πιθανότητας της NB2, ικανοποιεί τη σχέση (2.2.4).

Έστω τώρα ότι η $N \sim P$ τότε σύμφωνα με τον ορισμό το κ θα τείνει στο άπειρο ή στο πλην άπειρο, οπότε θα έχουμε

$$\lim_{\kappa \rightarrow \infty} \prod_{i=0}^{n-1} \frac{\kappa + i}{\kappa + \lambda} = 1 = \lim_{\kappa \rightarrow -\infty} \prod_{i=0}^{n-1} \frac{\kappa + i}{\kappa + \lambda}$$

$$\lim_{\kappa \rightarrow \infty} \left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa} = e^{-\lambda} = \lim_{\kappa \rightarrow -\infty} \left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa}$$

Με βάση τα παραπάνω αποτελέσματα και παρατηρώντας ότι το $\lambda^n/n!$ δεν θα αλλάξει αφού δεν επηρεάζεται από την ποσότητα κ , αποδείξαμε ότι η σχέση του ορισμού μας δίνει την συνάρτηση πιθανότητας της *Poisson*.

Τέλος αν η $N \sim B2$, αν θέσουμε όπου m το $-\kappa$, τότε θα προκύψει

$$\binom{m}{n} \frac{\lambda^n (m - \lambda)^{m-n}}{m^m} = \frac{1}{n!} \frac{\lambda^n (-\kappa - \lambda)^{-\kappa-n}}{m^m} \prod_{i=0}^{n-1} (-\kappa - i) = \frac{\lambda^n}{n!} (-1)^n \frac{\kappa}{(\kappa + \lambda)^{\kappa+n}} (-1)^n$$

άρα αποδείχθη ότι

$$p_n = \left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa} \frac{\lambda^n}{n!} \prod_{i=0}^{n-1} \frac{\kappa + i}{\kappa + \lambda}$$

■

Αφού έχουμε ορίσει την συνάρτηση πιθανότητας μπορούμε αρχικά να βρούμε τις αντίστοιχες παραμέτρους a, b για την κλάση $(a, b, 0)$ και στη συνέχεια κάνοντας χρήση των τύπων που αποδείξαμε παραπάνω να υπολογίσουμε τη μέση τιμή, τη διακύμανση, τη σχετική διακύμανση καθώς και το συντελεστή διακύμανσης.

Από τη σχέση (2.2.4) έχουμε,

$$\begin{aligned} \frac{p_n}{p_{n-1}} &= \frac{\left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa} \frac{\lambda^n}{n!} \prod_{i=0}^{n-1} \frac{\kappa + i}{\kappa + \lambda}}{\left(1 + \frac{\lambda}{\kappa}\right)^{-\kappa} \frac{\lambda^{n-1}}{(n-1)!} \prod_{i=0}^{n-2} \frac{\kappa + i}{\kappa + \lambda}} \\ &= \frac{\lambda \kappa + n - 1}{n \kappa + \lambda} \\ &= \frac{\lambda}{\kappa + \lambda} + \frac{(\kappa - 1)\lambda}{\kappa + \lambda} \frac{1}{n} \end{aligned}$$

άρα τελικά

$$\alpha = \frac{\lambda}{\kappa + \lambda} \quad \text{και} \quad b = \frac{(\kappa - 1)\lambda}{\kappa + \lambda} .$$

Έχοντας πλέον τα α, b θα υπολογίσουμε τις ποσότητες $\alpha + b, 1 - \alpha$ γιατί εμφανίζονται σε όλους τους υπολογισμούς μας στη συνέχεια.

$$\begin{aligned} \alpha + b &= \frac{\lambda}{\kappa + \lambda} + \frac{(\kappa - 1)\lambda}{\kappa + \lambda} = \frac{\kappa\lambda}{\kappa + \lambda} \\ 1 - \alpha &= \frac{\kappa}{\kappa + \lambda} \end{aligned}$$

Από τις σχέσεις (2.1.4), (2.1.7), (2.1.10), (2.1.11) θα πάρουμε αντίστοιχα ότι

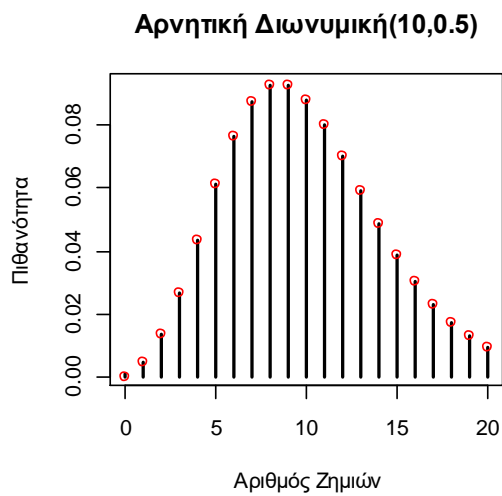
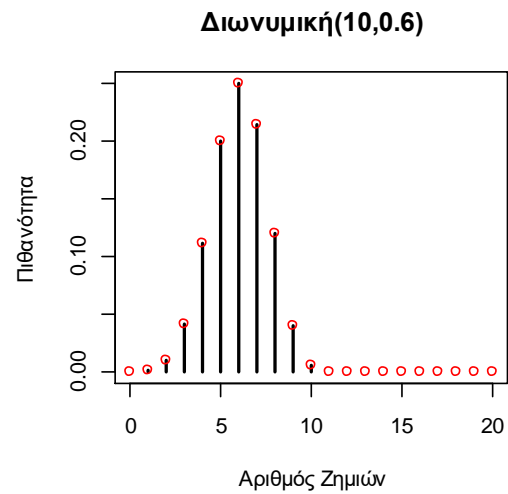
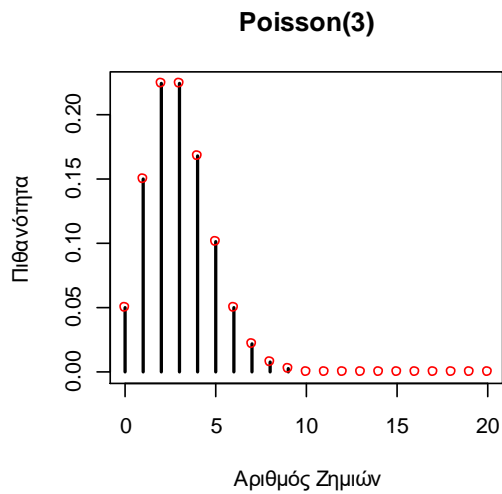
$$\begin{aligned} E(N) &= \frac{\frac{\kappa\lambda}{\kappa + \lambda}}{\frac{\kappa}{\kappa + \lambda}} = \lambda \\ \text{Var}(N) &= \frac{\frac{\kappa\lambda}{\kappa + \lambda}}{\left(\frac{\kappa}{\kappa + \lambda}\right)^2} = \lambda \left(1 + \frac{\lambda}{\kappa}\right) \\ \text{CV}^2(N) &= \frac{1}{\frac{\kappa\lambda}{\kappa + \lambda}} = \frac{1}{\kappa} + \frac{1}{\lambda} \\ D(N) &= \frac{1}{\frac{\kappa}{\kappa + \lambda}} = 1 + \frac{\lambda}{\kappa} . \end{aligned}$$

Τέλος ακολουθεί μία εφαρμογή στην οποία με την βοήθεια της γλώσσας προγραμματισμού R επαληθεύεται για κάθε μία από τις 3 κατανομές η ισότητα των αποτελεσμάτων με τον αναδρομικό τύπο του Panjer και με τον μη-αναδρομικό τύπο του Fackler.

Αναλυτικότερα, χρησιμοποιήθηκε ο κώδικας που υπάρχει στο Παράρτημα I και υπολογίστηκαν οι πιθανότητες για τις κατανομές $Poisson(3), Διωνυμική(10,0.6)$

και την *Αρνητική Διωνυμική*(10,0.5) όταν ο αριθμός των ζημιών παίρνει τιμές από το 0 έως το 20.

Όπως παρατηρούμε από το σχήμα που ακολουθεί και με τον αναδρομικό τύπο του *Ranjer* αλλά και με τον μη αναδρομικό τύπο του *Fackler* καταλήγουμε στα ίδια αποτελέσματα.



3 Η κλάση κατανομών $R(a, b, 1)$

Στο κεφάλαιο αυτό θα αναφερθούμε σε μία ευρύτερη κλάση κατανομών από την $R(a, b, 0)$, η οποία συμβολίζεται με $R(a, b, 1)$. Ξεκινώντας το κεφάλαιο θα δοθεί ένας αναδρομικός τύπος για τον υπολογισμό της συνάρτησης πιθανότητας καθώς και οι κατανομές που ανήκουν σε αυτή την οικογένεια. Στη συνέχεια θα δοθούν σχέσεις για τη μέση τιμή, τη δεύτερη ροπή και τη διακύμανση συναρτήσει των a και b , οι οποίες θα προκύψουν με τη χρήση ροπογεννητριών. Έπειτα θα παρουσιαστούν τόσο για τις περικομμένες όσο και για τις τροποποιημένες στο σημείο μηδέν κατανομές που ανήκουν στην κλάση αυτή η συνάρτηση πιθανότητας, η ροπογεννήτρια και η πιθανογεννήτρια καθώς και πώς όλες αυτές οι ποσότητες συνδέονται με τις αντίστοιχες αρχικές κατανομές τους. Τέλος θα δοθούν κάποια αποτελέσματα για την εκτεταμένη περικομμένη αρνητική διωνυμική κατανομή καθώς και για την λογαριθμική κατανομή.

3.1 Ο αναδρομικός τύπος του Panjer και βασικά μεγέθη

Σε αυτή την ενότητα θα ξεκινήσουμε δίνοντας τον ορισμό της κλάσης κατανομών $R(a, b, 1)$ και στη συνέχεια θα αποδείξουμε κάποιες σημαντικές σχέσεις που θα μας βοηθήσουν να βρούμε έναν κοινό τύπο για τον υπολογισμό της μέσης τιμής, της δεύτερης ροπής και της διακύμανσης για τις κατανομές που ανήκουν σε αυτή την οικογένεια.

Ορισμός 3.1.1

Για να ανήκει μία διακριτή τ.μ. $N \in \{0, 1, 2, \dots\}$ στην οικογένεια κατανομών $R(a, b, 1)$, αρκεί για κατάλληλες σταθερές a, b η συνάρτηση πιθανότητας $p_n = P(N = n)$ να ικανοποιεί την αναδρομική σχέση

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, \quad n = 2, 3, 4, \dots \quad (3.1.1)$$

και $\forall n < 0$ να ισχύει ότι $p_n = 0$.

■

Στην οικογένεια κατανομών $R(a, b, 1)$ ανήκουν όλες οι κατανομές που ανήκουν στην κλάση $R(a, b, 0)$ και επιπλέον ανήκουν οι περικομμένες αλλά και οι τροποποιημένες στο μηδέν Poisson, Διωνυμική, Αρνητική Διωνυμική, Γεωμετρική και η λογαριθμική κατανομή (Klugman et al, 2004).

Λήμμα 3.1.1

Αν η τ.μ. ανήκει στην κλάση κατανομών $(a, b, 1)$, τότε η ροπογεννήτρια συνάρτηση $M_N(t)$ της τ.μ. N , ικανοποιεί τη διαφορική εξίσωση

$$M'_N(t) = \frac{[p_1 + (a + b)(M_N(t) - p_0)]e^t}{1 - ae^t} \quad (3.1.2).$$

Απόδειξη

Σύμφωνα με τον ορισμό της ροπογεννήτριας ισχύει ότι,

$$\begin{aligned} M_N(t) &= E(e^{tN}) = \sum_{n=0}^{\infty} e^{tn} p_n = p_0 + p_1 e^t + \sum_{n=2}^{\infty} \left(a + \frac{b}{n}\right) p_{n-1} e^{tn} \\ &= p_0 + p_1 e^t + a \sum_{n=2}^{\infty} e^{tn} p_{n-1} + b \sum_{n=2}^{\infty} \frac{e^{tn}}{n} p_{n-1} \end{aligned} \quad (3.1.3).$$

Επίσης έχουμε ότι,

$$\begin{aligned} \sum_{n=2}^{\infty} e^{tn} p_{n-1} &= \sum_{n=1}^{\infty} e^{(t+1)p_n} \\ &= -e^t p_0 + e^t \sum_{n=0}^{\infty} e^{tn} p_n \\ &= -e^t p_0 + e^t M_N(t). \end{aligned}$$

Επομένως αντικαθιστώντας το παραπάνω αποτέλεσμα στη σχέση (3.1.3) παίρνουμε

$$M_N(t) = p_0 + p_1 e^t - ae^t p_0 + ae^t M_N(t) + b \sum_{n=2}^{\infty} \frac{e^{tn}}{n} p_{n-1},$$

ή ισοδύναμα,

$$(1 - ae^t)M_N(t) = p_0 + p_1 e^t - ae^t p_0 + b \sum_{n=2}^{\infty} \frac{e^{tn}}{n} p_{n-1}.$$

Στη συνέχεια θα παραγωγίσουμε την παραπάνω σχέση,

$$(1 - ae^t)M'_N(t) - ae^tM_N(t) = p_1e^t - ae^tp_0 + b \sum_{n=2}^{\infty} e^{tn}p_{n-1}$$

$$M'_N(t) = \frac{p_1e^t - ae^tp_0 - be^tp_0 + be^tM_N(t) + ae^tM_N(t)}{(1 - ae^t)}$$

$$M'_N(t) = \frac{[p_1 - (a + b)p_0 + (a + b)M_N(t)]e^t}{1 - ae^t}.$$

Συνεπώς προκύπτει το ζητούμενο που είναι,

$$M'_N(t) = \frac{[p_1 + (a + b)(M_N(t) - p_0)]e^t}{1 - ae^t}.$$

■

Πρόταση 3.1.1

Έστω ότι η τ.μ. N ανήκει στην κλάση κατανομών $R(a, b, 1)$ τότε η μέση τιμή, δεύτερη ροπή και η διακύμανση της δίνονται αντίστοιχα από τους παρακάτω τύπους:

$$\alpha) E(N) = \frac{p_1 + (a + b)(1 - p_0)}{1 - a}$$

$$\beta) E(N^2) = \frac{(a + b + 1)[p_1 + (a + b)(1 - p_0)]}{(1 - a)^2}$$

$$\gamma) Var(N) = \frac{[p_1 + (a + b)(1 - p_0)][1 - p_1 + (a + b)p_0]}{(1 - a)^2}$$

Απόδειξη

α) Για να βρούμε την μέση τιμή ή διαφορετικά μαθηματική ελπίδα, θα θέσουμε όπου t το 0 στη σχέση (3.1.2), δηλαδή

$$\begin{aligned} M'_N(0) &= E(N) \\ &= \frac{p_1e^0 - (a + b)e^0p_0 + (ae^0 + b)M_N(t)}{1 - ae^0} \\ &= \frac{p_1 + (a + b)(1 - p_0)}{1 - a} \end{aligned} \quad (3.1.4)$$

β) Σύμφωνα με τη σχέση (2.1.2) , για να βρούμε τη δεύτερη ροπή αρχικά χρειάζεται να παραγωγίσουμε την σχέση (3.1.2)

$$\begin{aligned}
 M''(t) &= \frac{[p_1 e^t + (a+b)(e^t M_N(t) + e^t M'_N(t) - p_0 e^t)](1 - ae^t)}{(1 - ae^t)^2} \\
 &\quad - \frac{[p_1 e^t + (a+b)(M_N(t)e^t - p_0 e^t)](-ae^t)}{(1 - ae^t)^2} \\
 &= \frac{[p_1 e^t + (a+b)(e^t M_N(t) - p_0 e^t)](1 - ae^t)}{(1 - ae^t)^2} \\
 &\quad - \frac{[p_1 e^t + (a+b)(M_N(t)e^t - p_0 e^t)](-ae^t)}{(1 - ae^t)^2} \\
 &\quad + \frac{(a+b)e^t M'_N(t)(1 - ae^t)}{(1 - ae^t)^2}
 \end{aligned}$$

Έπειτα για $t = 0$, προκύπτει η δεύτερη ροπή που είναι

$$\begin{aligned}
 M''(0) = E(N^2) &= \frac{p_1 + (a+b)(1 - p_0)}{(1 - a)^2} + \frac{(a+b) \frac{p_1 + (a+b)(1 - p_0)}{1 - a}}{1 - a} \\
 &= \frac{p_1 + (a+b)(1 - p_0) + (a+b)[p_1 + (a+b)(1 - p_0)]}{(1 - a)^2} \\
 &= \frac{(a+b+1)[p_1 + (a+b)(1 - p_0)]}{(1 - a)^2} .
 \end{aligned}$$

γ) Αφού λοιπόν έχουμε βρει την πρώτη και τη δεύτερη ροπή της τ.μ. N, έχουμε ό,τι χρειάζεται για να βρούμε την διακύμανση. Συγκεκριμένα, παίρνουμε ότι

$$\begin{aligned}
 Var(N) &= E(N^2) - E^2(N) \\
 &= \frac{(a+b+1)(p_1 + (a+b)(1 - p_0))}{(1 - a)^2} - \frac{[p_1 + (a+b)(1 - p_0)]^2}{(1 - a)^2} \\
 &= \frac{[a+b+1 - p_1 - (a+b)(1 - p_0)][p_1 + (a+b)(1 - p_0)]}{(1 - a)^2} \\
 &= \frac{[p_1 + (a+b)(1 - p_0)][1 - p_1 + (a+b)p_0]}{(1 - a)^2} \tag{3.1.5}.
 \end{aligned}$$

3.2 Περικομμένες στο σημείο μηδέν κατανομές

Ένα συνηθισμένο και βασικό πρόβλημα στην κατανομή του πλήθους των ζημιών είναι η πιθανότητά της στο μηδέν. Υπάρχουν περιπτώσεις στις οποίες η ελάχιστη τιμή την οποία παίρνει το πλήθος των ζημιών είναι 1. Για παράδειγμα, αν μετράμε τον αριθμό των ζημιών από ατυχήματα τα οποία είχαν τουλάχιστον μία ζημιά, η ελάχιστη παρατηρηθείσα τιμή είναι το ένα. Για τον λόγο αυτό θα δείξουμε για τις κατανομές Poisson, Διωνυμική, Αρνητική Διωνυμική και Γεωμετρική κάποιους τύπους που μας δίνουν τις αντίστοιχες περικομμένες στο μηδέν κατανομές τους.

Ορισμός 3.2.1

Έστω N μία τυχαία μεταβλητή η οποία παριστάνει το πλήθος των ζημιών σε ένα χαρτοφυλάκιο, έτσι ώστε $N \in \{0, 1, 2, \dots\}$. Η κατανομή της τ.μ. N^T που ορίζεται από τον τύπο

$$N^T = N | N > 0$$

καλείται περικομμένη στο σημείο μηδέν (zero truncated distribution).

■

Η συνάρτηση πιθανότητας της τ.μ. N^T είναι

$$p_n^T = P(N^T = n) = \begin{cases} 0, & n = 0 \\ \frac{p_n}{1 - p_0}, & n = 1, 2, 3, \dots \end{cases}$$

Απόδειξη

Αρχικά για $n = 1, 2, 3, \dots$, από την θεωρία πιθανοτήτων ισχύει ότι

$$\begin{aligned} p_n^T &= P(N^T = n) = P(N = n | N > 0) \\ &= \frac{P(N = n)}{P(N > 0)} = \frac{P(N = n)}{1 - P(N = 0)} = \frac{p_n}{1 - p_0}. \end{aligned}$$

Τέλος, για $n = 0$

$$p_0^T = P(N^T = 0) = P(N = 0 | N > 0) = 0,$$

κάτι που είναι αναμενόμενο αφού η κατανομή δεν έχει μάζα στο σημείο μηδέν.



Έχοντας λοιπόν την συνάρτηση πιθανότητας για την N^T μπορούμε να βρούμε την πιθανογεννήτρια, την ροπογεννήτρια και κατ' επέκταση τη μέση τιμή και τη διακύμανσή της.

Κάνοντας χρήση του ορισμού της ροπογεννήτριας, θα πάρουμε

$$\begin{aligned} M_{N^T}(t) &= \sum_{n=1}^{\infty} e^{tn} p_n^T = \sum_{n=1}^{\infty} e^{tn} \frac{p_n}{1-p_0} \\ &= \frac{1}{1-p_0} \left\{ \sum_{n=0}^{\infty} e^{tn} p_n - p_0 \right\}, \end{aligned}$$

άρα,

$$M_{N^T}(t) = \frac{M_N(t) - p_0}{1 - p_0} \quad (3.2.1)$$

και αν γίνει η αντικατάσταση $t = \ln t$, προκύπτει η πιθανογεννήτρια που είναι

$$P_{N^T}(t) = \frac{P_N(t) - p_0}{1 - p_0}.$$

Ο υπολογισμός της μέσης τιμής και της διακύμανσης της περικομμένης στο σημείο μηδέν μεταβλητής συναρτήσεως της μέσης και της διακύμανσης της τ.μ N είναι εύκολος τώρα έχοντας υπολογίσει την ροπογεννήτρια.

Παραγωγίζοντας ως προς t , τη σχέση (3.2.1) θα έχουμε

$$M'_{N^T}(t) = \frac{M'_N(t)}{1 - p_0} \quad (3.2.2)$$

και για $t = 0$, θα πάρουμε τη μέση τιμή που είναι

$$E(N^T) = \frac{E(N)}{1 - p_0}.$$

Για να βρούμε τώρα την δεύτερη ροπή που θα μας χρειαστεί για τον υπολογισμό της διακύμανσης θα παραγωγίσουμε τη σχέση (3.2.2) ως προς t και άμεσα παίρνουμε ότι

$$M''_{N^T}(t) = \frac{M''_N(t)}{1 - p_0}.$$

Αν τώρα στη θέση του t θέσουμε το μηδέν θα πάρουμε τη δεύτερη ροπή που είναι

$$E[(N^T)^2] = \frac{E(N^2)}{1 - p_0}.$$

Έχοντας τώρα τα παραπάνω αποτελέσματα, η διακύμανση ισούται με

$$\begin{aligned} \text{Var}(N^T) &= E[(N^T)^2] - [E(N^T)]^2 \\ &= \frac{E(N^2)}{1 - p_0} - \left(\frac{E(N)}{1 - p_0}\right)^2 \\ &= \frac{(1 - p_0)E(N^2) - [E(N)]^2}{(1 - p_0)^2} \end{aligned}$$

από όπου προκύπτει ο παρακάτω τύπος

$$\text{Var}(N^T) = \frac{\text{Var}(N) - p_0 E(N^2)}{(1 - p_0)^2}.$$

3.3 Τροποποιημένες στο σημείο μηδέν κατανομές

Εκτός από τις περικομμένες στο μηδέν κατανομές, υπάρχει και άλλη μία βασική κατηγορία, οι τροποποιημένες στο μηδέν κατανομές που είναι εξίσου σημαντικές διότι μπορούμε να ορίσουμε εμείς την πιθανότητα στο σημείο μηδέν ανάλογα με το χαρτοφυλάκιό μας. Συνεπώς για τις κατανομές που ανήκουν στην κλάση $R(a, b, 0)$, θα δώσουμε έναν ενιαίο τύπο που θα μας δίνει την συνάρτηση πιθανότητας καθώς και κάποια άλλα βασικά μεγέθη.

Πρόταση 3.3.1

Έστω η τ.μ. $N \in \{0, 1, 2, \dots\}$, τότε η N^M θα λέγεται τροποποιημένη στο σημείο μηδέν τ.μ. (zero modified distribution) όταν η συνάρτηση πιθανότητάς της θα είναι της μορφής,

$$p_n^M = \begin{cases} p_0^M, & n = 0 \\ cp_n, & n = 1, 2, 3, \dots \end{cases}$$

όπου το p_0^M είναι ένας οποιοσδήποτε αριθμός που ανήκει στο διάστημα $[0, 1)$ και το c μία σταθερά μεγαλύτερη του 0, η οποία ισούται με $\frac{1-p_0^M}{(1-p_0)}$.

Απόδειξη

Γνωρίζουμε από την θεωρία πιθανοτήτων ότι για να είναι η p_n^M συνάρτηση πιθανότητας θα πρέπει να ισχύει ότι

$$\sum_{n=0}^{\infty} p_n^M = 1$$

$$\Leftrightarrow \sum_{n=1}^{\infty} p_n^M + p_0^M = 1$$

$$\Leftrightarrow cp_n + p_0^M = 1$$

$$\Leftrightarrow c(1 - p_0) + p_0^M = 1.$$

Από την σχέση αυτή τελικά προκύπτει ότι,

$$c = \frac{1 - p_0^M}{(1 - p_0)}.$$

Οι άλλες 2 συνθήκες που πρέπει να ικανοποιούνται για να είναι η p_n^M συνάρτηση πιθανότητας ισχύουν εξ' ορισμού.

■

Θα εργαστούμε όπως και στις περικομμένες τ.μ., με σκοπό να βρούμε τη ροπογεννήτρια, την πιθανογεννήτρια και τις ροπές των τροποποιημένων στο μηδέν τ.μ.

Ξεκινάμε με τη ροπογεννήτρια και έχουμε

$$\begin{aligned} M_{N^M}(t) &= \sum_{n=0}^{\infty} e^{tn} p_n^M = p_0^M + \sum_{n=1}^{\infty} e^{tn} p_n^M \\ &= p_0^M + \sum_{n=1}^{\infty} e^{tn} \frac{1 - p_0^M}{1 - p_0} p_n \end{aligned}$$

$$\begin{aligned}
&= p_0^M + \frac{1 - p_0^M}{1 - p_0} \left\{ \sum_{n=0}^{\infty} e^{tn} p_n - p_0 \right\} \\
&= p_0^M - \frac{1 - p_0^M}{1 - p_0} p_0 + \frac{1 - p_0^M}{1 - p_0} M_N(t) \\
&= \frac{p_0^M(1 - p_0) - (1 - p_0^M)p_0}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} M_N(t) \\
&= \frac{p_0^M - p_0}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} M_N(t) \\
&= 1 - \frac{1 - p_0^M}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} M_N(t) \tag{3.2.3}.
\end{aligned}$$

Με απλή αντικατάσταση του t με lnt θα πάρουμε την αντίστοιχη πιθανογεννήτρια συνάρτηση της που είναι,

$$P_{N^M}(t) = 1 - \frac{1 - p_0^M}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} P_N(t).$$

Στη συνέχεια θα παραγωγίσουμε τη σχέση (3.2.3),

$$M'_{N^M}(t) = \frac{1 - p_0^M}{1 - p_0} M'_N(t).$$

Άρα εύκολα προκύπτει ότι η μέση τιμή είναι

$$E(N^M) = \frac{1 - p_0^M}{1 - p_0} E(N).$$

Είναι προφανές ότι με τον ίδιο τρόπο η δεύτερη ροπή της είναι

$$E[(N^M)^2] = \frac{1 - p_0^M}{1 - p_0} E(N^2)$$

οπότε βρίσκουμε ότι η διακύμανση ισούται με,

$$\begin{aligned}
Var(N^M) &= \frac{1 - p_0^M}{1 - p_0} E(N^2) - \left[\frac{1 - p_0^M}{1 - p_0} E(N) \right]^2 \\
&= \frac{1 - p_0^M}{1 - p_0} \left[E(N^2) - \frac{1 - p_0^M}{1 - p_0} E^2(N) \right].
\end{aligned}$$

Παρατηρούμε λοιπόν από τα αποτελέσματά μας ότι υπάρχουν σχέσεις οι οποίες συνδέουν τις ροπογεννήτριες και κατ' επέκταση και τις ροπές των δύο παραπάνω κατανομών. Ενδεικτικά κάποιες από αυτές είναι οι παρακάτω:

- 1) $M_{NM}(t) = 1 - \frac{1-p_0^M}{1-p_0} + p_0 M_N(t)$
- 2) $E(N^M) = (1 - p_0^M)E(N^T)$
- 3) $E[(N^M)^2] = (1 - p_0^M)E[(N^T)^2]$.

3.4 Εκτεταμένη Περικομμένη Αρνητική Διωνυμική Κατανομή

Στο κεφάλαιο 2 αναφέραμε ότι η Αρνητική Διωνυμική κατανομή ορίζεται για $r \in (0, \infty)$. Αν επεκτείνουμε το πεδίο τιμών του r , δηλαδή $r \in (-1, \infty)$, τότε προκύπτει μία νέα κατανομή που ονομάζεται Εκτεταμένη Περικομμένη Αρνητική Διωνυμική κατανομή (Extended Truncated Negative Binomial Distribution). Η κατανομή αυτή έχει βρει πρόσφατα σημαντική εφαρμογή στη Θεωρία Χρεοκοπίας για το πλήθος των απαιτήσεων έως τη χρεοκοπία στο άρθρο των Frostig et al (2012). Σύμφωνα με το άρθρο του Willmot (1988) έχουμε ότι η συνάρτηση πιθανότητας της ETNB είναι

$$p_n = \frac{-r\Gamma(n+r)}{n!\Gamma(1+r)} \frac{p^n}{1 - (1-p)^{-r}} \quad (3.4.1)$$

όπου $-1 < r < 0$ και $0 < p < 1$.

Γνωρίζοντας την συνάρτηση πιθανότητας, θα αποδείξουμε ότι η πιθανογεννήτρια συνάρτηση δίνεται από τον ακόλουθο τύπο

$$P(t) = \frac{1 - (1-pt)^{-r}}{1 - (1-p)^{-r}}.$$

Από τον ορισμό της πιθανογεννήτριας έχουμε ότι

$$\begin{aligned} P(t) &= \sum_{n=1}^{\infty} t^n p_n \\ &= \sum_{n=1}^{\infty} t^n \frac{-r\Gamma(n+r)}{n!\Gamma(1+r)} \frac{p^n}{1 - (1-p)^{-r}} \\ &= \frac{-r}{1 - (1-p)^{-r}} \sum_{n=1}^{\infty} \frac{(tp)^n (n+r-1)!}{n! r!} \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{1 - (1 - p)^{-r}} \sum_{n=1}^{\infty} \frac{(tp)^n (n + r - 1)!}{n! (r - 1)!} \\
&= -\frac{1}{1 - (1 - p)^{-r}} \sum_{n=1}^{\infty} \binom{n + r - 1}{n} (tp)^n. \quad (3.4.2)
\end{aligned}$$

Επίσης γνωρίζουμε ότι μία άλλη έκφραση της Αρνητικής Διωνυμικής κατανομής είναι

$$p_n = \binom{r + n - 1}{n} p^r (1 - p)^n, \quad n = 0, 1, 2, \dots$$

συνεπώς θα ισχύει η παρακάτω ισότητα

$$\sum_{n=0}^{\infty} \binom{r + n - 1}{n} p^r (1 - p)^n = 1,$$

από όπου τελικά προκύπτει ότι

$$\sum_{n=1}^{\infty} \binom{r + n - 1}{n} (1 - p)^n = \frac{1 - p^r}{p^r}. \quad (3.4.3)$$

Από τις σχέσεις (3.3.2) και (3.3.3) τελικά προκύπτει η πιθανογεννήτρια συνάρτηση δηλαδή,

$$\begin{aligned}
P(t) &= -\frac{1}{1 - (1 - p)^{-r}} \frac{1 - (1 - pt)^r}{(1 - pt)^r} \\
&= \frac{1 - (1 - pt)^{-r}}{1 - (1 - p)^{-r}}.
\end{aligned}$$

Με δεδομένο ότι πλέον γνωρίζουμε την συνάρτηση πιθανότητας της ETNB, θα αποδείξουμε ότι ανήκει στην κλάση $R(a, b, 1)$ και θα βρούμε τα a, b .

Από την σχέση (3.3.1) του ορισμού έχουμε ότι

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

$$\begin{aligned}
\frac{p_n}{p_{n-1}} &= \frac{\frac{-r\Gamma(n+r)}{n!\Gamma(1+r)} \frac{p^n}{1-(1-p)^r}}{\frac{-r\Gamma(n-1+r)}{(n-1)!\Gamma(1+r)} \frac{p^{n-1}}{1-(1-p)^r}} \\
&= \frac{p(n+r-1)!}{n(n+r-2)!} \\
&= p + \frac{(r-1)}{n}p.
\end{aligned}$$

Τελικά καταλήξαμε στο συμπέρασμα ότι ανήκει στην κλάση $R(a, b, 1)$ με $a = p$ και $b = (r-1)p$.

Παρατηρούμε λοιπόν ότι αν βρούμε την p_1 και κάνοντας χρήση των αποτελεσμάτων της πρότασης 3.1.1 θα έχουμε και την μέση τιμή και τη διακύμανση της ETNB.

Ξεκινάμε υπολογίζοντας την p_1 .

$$p_1 = \frac{-r\Gamma(1+r)}{1!\Gamma(1+r)} \frac{1}{1-(1-p)^{-r}} = \frac{-rp}{1-(1-p)^{-r}}$$

Από την σχέση (3.1.4) προκύπτει ότι,

$$\begin{aligned}
E(N) &= \frac{\frac{-rp}{1-(1-p)^{-r}} + p + (r-1)p}{1-p} \\
&= \frac{\frac{-rp + rp[1-(1-p)^{-r}]}{1-(1-p)^{-r}}}{1-p} \\
&= \frac{rp(1-p)^{-r-1}}{1-(1-p)^{-r}}
\end{aligned}$$

■

Έχοντας τώρα την συνάρτηση πιθανότητας της Εκτεταμένης Περικομμένης Αρνητικής Διωνυμικής κατανομής αν υποθέσουμε ότι το $r \rightarrow 0$ τότε η προκύπτουσα κατανομή είναι η λογαριθμική κατανομή - logarithmic distribution (Willmot 1988).

Ορισμός 3.4.1

Μία τ.μ $N \in \{0,1,2, \dots\}$ ακολουθεί τη λογαριθμική κατανομή και θα την συμβολίζουμε με $LS(q)$, αν έχει την ακόλουθη συνάρτηση πιθανότητας

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

όπου η παράμετρος q ισούται με $1-p$ και επίσης $0 < p < 1$.

■

Προφανώς και η κατανομή αυτή ανήκει στην κλάση κατανομών $R(\alpha, b, 1)$ με $\alpha = p$, $b = -p$ αφού είναι ειδική περίπτωση της ETNB.

Αρχικά, θα υπολογιστεί ο τύπος της πιθανογεννήτριας της λογαριθμικής κατανομής μέσω του ορισμού της πιθανογεννήτριας.

Ξεκινώντας έχουμε ότι,

$$\begin{aligned} P_N(t) &= E(t^N) = \sum_{n=1}^{\infty} t^n \frac{p^n}{-n \ln(1-p)} \\ &= -\frac{1}{\ln(1-p)} \sum_{n=1}^{\infty} \frac{(pt)^n}{n} = -\frac{1}{\ln(1-p)} [-\ln(1-pt)] \end{aligned}$$

για $t < p^{-1}$, έτσι ώστε να συγκλίνει η σειρά, και τελικά καταλήγουμε ότι,

$$P_N(t) = \frac{\ln(1-pt)}{\ln(1-p)}.$$

Στη συνέχεια αφού έχουμε τον τύπο της πιθανογεννήτριας μπορούμε μέσω αυτού να υπολογίσουμε τη μέση τιμή, τη δεύτερη ροπή και τη διακύμανση της λογαριθμικής κατανομής.

Αρχικά για την εύρεση του τύπου της μέση τιμής θα παραγωγίσουμε τη πιθανογεννήτρια, οπότε έχουμε για $t < p^{-1}$,

$$P'_N(t) = \frac{1}{\ln(1-p)} \frac{(1-p)'}{(1-pt)}$$

και τελικά ισχύει ότι,

$$P'_N(t) = -\frac{p}{(1-pt)\ln(1-p)}.$$

Αρκεί τώρα να αντικαταστήσουμε όπου $t = 1$, για να καταλήξουμε ότι,

$$E(N) = P'_N(1) = -\frac{p}{(1-p)\ln(1-p)}.$$

Έπειτα θα υπολογιστεί η δεύτερη παράγωγος της πιθανογεννήτριας κάτι που θα μας βοηθήσει να βρούμε τη δεύτερη ροπή.

Παραγωγίζοντας λοιπόν 2 φορές την πιθανογεννήτρια παίρνουμε,

$$\begin{aligned} P''_N(t) &= -\frac{p}{\ln(1-p)} \left(-\frac{(1-pt)'}{(1-pt)^2} \right) \\ &= -\frac{p^2}{(1-pt)^2 \ln(1-p)}. \end{aligned}$$

Αν τώρα στη βάλουμε όπου $t = 1$ από γνωστή ιδιότητα της πιθανογεννήτριας προκύπτει ότι,

$$P''_N(1) = E(N(N-1))$$

άρα η δεύτερη ροπή ισούται με

$$\begin{aligned} E(N^2) &= P''_N(1) + E(N) \\ &= -\frac{p^2}{(1-p)^2 \ln(1-p)} - \frac{p}{(1-p)\ln(1-p)} \end{aligned}$$

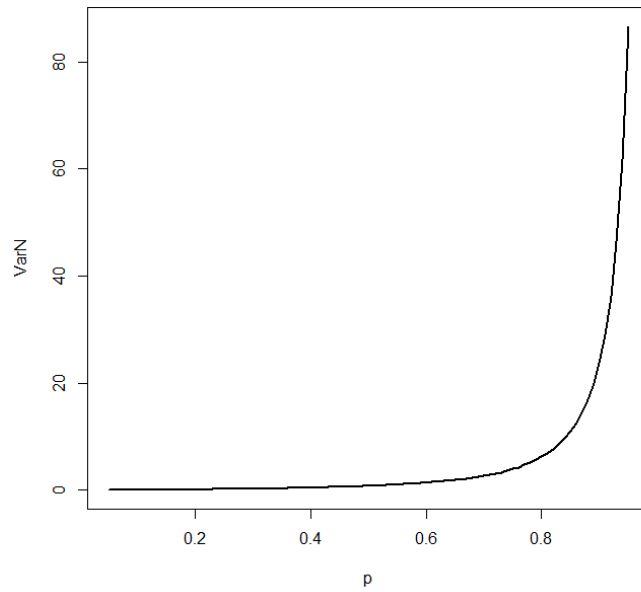
και τελικά καταλήγουμε ότι,

$$E(N^2) = -\frac{p}{(1-p)^2 \ln(1-p)}.$$

Έχοντας υπολογίσει όλα τα παραπάνω καταλήγουμε στο συμπέρασμα ότι η διακύμανση ισούται με

$$Var(N) = -\frac{p \ln(1-p) + p^2}{(1-p)^2 (\ln(1-p))^2}.$$

Αφού δείξαμε πως υπολογίζεται η διακύμανση, στη συνέχεια θα δούμε σχηματικά πώς μεταβάλλονται οι τιμές τις όταν μεταβάλλεται η παράμετρος p .



Σχήμα 3.3.1

Το συμπέρασμα που προκύπτει από το παραπάνω σχήμα είναι ότι όταν αυξάνεται η παράμετρος ρ , η διακύμανση της λογαριθμικής κατανομής αυξάνεται με πολύ γρήγορο ρυθμό.

4 Η κατανομή των συνολικών αποζημιώσεων

Στο παρόν κεφάλαιο αρχικά θα δοθούν κάποιοι χρήσιμοι αναδρομικοί τύποι με τους οποίους υπολογίζουμε την συνάρτηση πιθανότητας της τ.μ. S . Έπειτα για τις περιπτώσεις όπου η τ.μ. N ακολουθεί είτε Γεωμετρική, είτε Διωνυμική, είτε Αρνητική Διωνυμική κατανομή και η τ.μ. X ακολουθεί Εκθετική κατανομή θα δοθεί ακριβής τύπος για τον υπολογισμό της σ.κ. αλλά και της σ.π.π. Στη συνέχεια θα γίνει αναλυτική περιγραφή του τρόπου με τον οποίο υπολογίζεται η αθροιστική συνάρτηση κατανομή της τ.μ. S με τη χρήση της γλώσσας προγραμματισμού R για τις κατανομές που ανήκουν στην κλάση κατανομών $(a, b, 0)$ και $(a, b, 1)$ και κατ' επέκταση θα δοθεί ο υπολογισμός της συνάρτησης πιθανότητας. Οι περιπτώσεις οι οποίες θα αναλυθούν είναι πρώτον όταν και η τ.μ. X αλλά και η N ακολουθούν διακριτές κατανομές και δεύτερον όταν η τ.μ. X ακολουθεί συνεχή κατανομή και η N ακολουθεί διακριτή. Στην δεύτερη περίπτωση θα δειχθεί πως γίνεται μέσω της R να διακριτοποιήσουμε την τ.μ. X με τέσσερις τρόπους, έτσι ώστε να μπορεί να χρησιμοποιηθεί στον αναδρομική σχέση του Panjer. Επιπροσθέτως θα παρουσιαστούν παραδείγματα στα οποία θα συγκριθούν τα αποτελέσματα τα οποία προέκυψαν με τον ακριβή τύπο σε σχέση με αυτά που προκύπτουν όταν γίνεται διακριτοποίηση στην συνεχή κατανομή. Τέλος θα δειχθεί πώς αυτά τα αποτελέσματα επηρεάζονται από την μέθοδο διακριτοποίησης, το βήμα αλλά και τις παραμέτρους των κατανομών.

4.1 Βασικοί Τύποι

Αρχικά θα αναφέρουμε 2 βασικές προτάσεις που μας δίνουν τη σ.π. της τ.μ. S στις περιπτώσεις που η τ.μ. $N \in R(a, b, 0)$ και $N \in R(a, b, 1)$. Οι αποδείξεις παραλείπονται καθώς είναι αρκετά εκτενείς και επίσης υπάρχουν στις πανεπιστημιακές σημειώσεις (βλέπε Χατζηκωνσταντινίδης (2015)).

Πρόταση 4.1.1

Αν η τ.μ. $X \in \{0, 1, 2, \dots\}$ έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τ.μ. $N \in R(a, b, 0)$, τότε η $S = X_1 + X_2 + \dots + X_N$ για $N \geq 1$ και με $S = 0$ αν $N = 0$ έχει σ.π. $g(x) = P(S = x)$ που ικανοποιεί την παρακάτω αναδρομική σχέση

$$g(x) = \frac{1}{1 - af(0)} \sum_{y=1}^x \left(a + \frac{by}{x}\right) f(y)g(x-y), \quad x = 1, 2, 3, \dots$$

και

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

Πρόταση 4.1.2

Αν η τ.μ. $X \in \{0,1,2, \dots\}$ έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τ.μ. $N \in R(a, b, 1)$, τότε η $S = X_1 + X_2 + \dots + X_N$ για $N \geq 1$ και με $S = 0$ αν $N = 0$ έχει σ.π. $g(x) = P(S = x)$ που ικανοποιεί την παρακάτω αναδρομική σχέση

$$g(x) = \frac{1}{1 - af(0)} \sum_{y=1}^x \left(a + \frac{by}{x}\right) f(y)g(x-y) + \frac{p_1 - (a+b)p_0}{1 - af(0)} f(x), \quad x = 1,2,3, \dots$$

και

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

Επίσης εξίσου σημαντικό θα ήταν να ορίσουμε τη μέση τιμή και τη διακύμανση του συνολικού ύψους των ζημιών S .

Πρόταση 4.1.3

Έστω $S = X_1 + X_2 + \dots + X_N$, όπου η τ.μ. $N \in R(a, b, 1)$ και η τ.μ. $X \in \{0,1,2, \dots\}$ τότε η μέση τιμή και η διακύμανση της S είναι,

$$1) E(S) = \frac{p_1 + (a+b)(1-p_0)}{1-a} E[X]$$

$$2) Var[S] = \frac{p_1 + (a+b)(1-p_0)}{1-a} \left[Var[X] + E[X]^2 \frac{[1 - p_1 + (a+b)p_0]}{(1-a)} \right]$$

Απόδειξη

1) Γενικά ισχύει ότι

$$E(S) = E_N[E[X_1 + X_2 + \dots + X_N]]$$

και λόγω του ότι τα X, N είναι ανεξάρτητες τ.μ. προκύπτει ότι,

$$E[S] = E[N]E[X]$$

και από τη σχέση (3.1.4) καταλήγουμε στο συμπέρασμα ότι,

$$E[S] = \frac{p_1 + (a+b)(1-p_0)}{1-a} E[X].$$

2) Γνωρίζουμε ότι,

$$\text{Var}[S] = E[\text{Var}[S|N]] + \text{Var}[E[S|N]]$$

και από την ανεξαρτησία των X, N συνεπάγεται ότι (Dickson, 2006)

$$\text{Var}[S] = E[N\text{Var}[X]] + \text{Var}[NE[X]]$$

$$\text{Var}[S] = E[N]\text{Var}[X] + E[X]^2\text{Var}[N]$$

και χρησιμοποιώντας τις σχέσεις (3.1.4), (3.1.5) καταλήγουμε στο ζητούμενο που είναι,

$$\text{Var}[S] = \frac{p_1 + (a+b)(1-p_0)}{1-a} \left[\text{Var}[X] + E[X]^2 \frac{[1-p_1 + (a+b)p_0]}{(1-a)} \right].$$

■

Σε κάποιες περιπτώσεις όπου η τ.μ. N ακολουθεί κάποια κατανομή από την κλάση $R(a, b, 0)$ και η τ.μ. X ακολουθεί κάποια συνεχή κατανομή υπάρχει ακριβής τύπος για τον υπολογισμό της συνάρτησης κατανομής και της συνάρτησης πιθανότητας για την τ.μ. S . Στη συνέχεια θα δώσουμε κάποια αποτελέσματα των οποίων οι αποδείξεις υπάρχουν στις πανεπιστημιακές σημειώσεις του κ. Χατζηκωνσταντινίδη (2015).

Γεωμετρική - Εκθετική

Έστω ότι το πλήθος των ζημιών είναι τ.μ. και ακολουθεί την γεωμετρική κατανομή με συνάρτηση πιθανότητας

$$p_n = pq^n, \text{ για } n = 0, 1, 2, \dots \text{ και } 0 < p < 1,$$

το ύψος των αποζημιώσεων είναι τ.μ. για την οποία ισχύει ότι $X \sim \text{Exp}(\lambda)$ και τέλος το συνολικό ύψος των ζημιών είναι τ.μ. και συμβολίζεται με S . Ισχύει ότι η συνάρτηση κατανομής δίνεται από τη σχέση

$$G(x) = 1 - qe^{-p\lambda x}, x \geq 0 \quad (4.1.1)$$

και η συνάρτηση πυκνότητας πιθανότητας από τη σχέση

$$g(x) = \begin{cases} p, & x = 0 \\ qp\lambda e^{-p\lambda x}, & x > 0 \end{cases}$$

Διωνυμική - Εκθετική

Αν ο αριθμός των ζημιών ακολουθεί την διωνυμική κατανομή, δηλαδή $N \sim B(m, p)$ και το ύψος των ζημιών την εκθετική, $X \sim Exp(\lambda)$, τότε για την τ.μ. $S = X_1 + \dots + X_N$, η συνάρτηση κατανομής της είναι

$$G(x) = 1 - e^{-\lambda x} \sum_{n=1}^m \binom{m}{n} p^n q^{m-n} \sum_{j=0}^{n-1} \frac{(\lambda x)^j}{j!}, \quad x \geq 0 \quad (4.1.2)$$

Η συνάρτηση πιθανότητας της τ.μ. S είναι μικτού τύπου, διότι έχει μάζα πιθανότητας στο σημείο μηδέν και είναι συνεχής στο διάστημα $(0, +\infty)$ και δίνεται από την σχέση

$$g(x) = \begin{cases} q^m, & x = 0 \\ e^{-\lambda x} \sum_{n=1}^m \binom{m}{n} p^n q^{m-n} \frac{\lambda^n}{(n-1)!} x^{n-1}, & x > 0 \end{cases}$$

Αρνητική Διωνυμική - Εκθετική

Όταν η τ.μ. N που εκφράζει το ύψος των ζημιών ακολουθεί αρνητική διωνυμική κατανομή με συνάρτηση πιθανότητας

$$p_n = \binom{r+n-1}{n} p^r q^n, \quad \text{για } x = 0, 1, 2, \dots \text{ και } 0 < p < 1$$

και η τ.μ. X , η οποία εκφράζει το ύψος των αποζημιώσεων, ακολουθεί εκθετική με παράμετρο λ τότε το συνολικό ύψος των ζημιών έχει συνάρτηση κατανομής την

$$G(x) = 1 - e^{-p\lambda x} \sum_{j=0}^{r-1} \frac{(p\lambda x)^j}{j!} \sum_{n=j+1}^r \binom{r}{n} q^n p^{r-n}, \quad x \geq 0 \quad (4.1.3)$$

και αντίστοιχα συνάρτηση πυκνότητας πιθανότητας την

$$g(x) = \begin{cases} p^r, & x = 0 \\ \sum_{t=1}^r \binom{r}{t} \frac{p^r (q\lambda)^t}{(t-1)!} x^{t-1} e^{-p\lambda x}, & x > 0 \end{cases}$$

4.2 Υπολογισμός της σ.κ. με τη μέθοδο Recursive

Όπως παρατηρείται, ο υπολογισμός της συνάρτησης πιθανότητας αλλά και της συνάρτησης κατανομής είναι αρκετά χρονοβόρος και επίπονος λόγω των αρκετών πράξεων αφού οι σχέσεις είναι αναδρομικές. Για τον λόγο αυτό στη συνέχεια θα δειχθεί πώς μπορεί να υπολογιστεί η συνάρτηση κατανομής της τ.μ. S με τη χρήση της γλώσσας προγραμματισμού R όταν η κατανομή της τ.μ. N ανήκει στη κλάση κατανομών $R(a, b, 1)$.

Αρχικά η συνάρτηση που θα χρησιμοποιηθεί είναι η `aggregateDist`, η οποία μας δίνει την αθροιστική συνάρτηση κατανομής της τ.μ. S και βρίσκεται στο πακέτο "actuar". Η συνάρτηση αυτή υπολογίζει την αθροιστική συνάρτηση κατανομής ενός χαρτοφυλακίου για μία περίοδο με 5 διαφορετικούς τρόπους ανάλογα με την κατανομή που ακολουθεί το πλήθος των ζημιών καθώς και το ύψος των ζημιών.

Στην περίπτωση που το πλήθος των ζημιών ακολουθεί κάποια κατανομή από την κλάση κατανομών $R(a, b, 1)$ συνεπώς και από την κλάση $R(a, b, 0)$ τότε χρησιμοποιείται η μέθοδος "recursive". Η συνάρτηση που μας δίνει την αθροιστική συνάρτηση κατανομής είναι

$$\text{aggregateDist}(\text{"recursive"}, \text{model.freq} = \text{"poisson"}, \text{model.sev} = fx, \\ \text{lambda} = \lambda, \text{x.scale} = h, \text{maxit} = k)$$

και στην συνέχεια θα εξηγήσουμε τις παραμέτρους της συνάρτησης καθώς και τι ορίζει κάθε μία από αυτές.

Αρχικά αν η κατανομή του πλήθους των αποζημιώσεων ακολουθεί κάποια κατανομή από την κλάση $R(a, b, 0)$ τότε το `model.freq` θα πρέπει να πάρει μία από τις τιμές

- poisson,
- binomial,
- geometric,
- negative binomial.

Στην περίπτωση που χρειάζεται να χρησιμοποιήσουμε μία περικομμένη στο σημείο μηδέν κατανομή τότε το `model.freq` θα πρέπει να πάρει συμπληρωθεί με ένα από τα

- zero-truncated binomial,
- zero truncated geometric,

- zero truncated negative binomial,
- zero truncated Poisson,
- logarithmic.

Τέλος, αν η κατανομή που ακολουθεί το πλήθος των ζημιών είναι τροποποιημένη στο σημείο μηδέν, τότε το *model.freq* θα πάρει μία από τις τιμές

- zero modified binomial,
- zero modified geometric,
- zero modified Poisson,
- zero modified negative binomial,
- zero modified logarithmic.

Επίσης πρέπει να οριστεί και η πιθανότητα στο σημείο μηδέν δηλαδή να ορίσουμε μία τιμή στο p_0 από 0 έως 1. Επίσης θα πρέπει να οριστούν και οι παράμετροι της τ.μ. N , δηλαδή έστω ότι η N ακολουθεί Poisson με παράμετρο λ , τότε θα πρέπει να το ορίσουμε ως $\lambda = 2$.

Στην συνέχεια η παράμετρος *model.sev* θα παριστάνεται με μία $f(x)$ την οποία θα την έχουμε ορίσει έτσι ώστε να μας δίνει τις πιθανότητες για κάθε τιμή του x . Δηλαδή εάν η $X \sim G(0.6)$ τότε η $f(0) = 0$, $f(1) = 0.6$ και ούτω καθεξής.

Τέλος το *x.scale* θα το ορίσουμε, στην περίπτωση όπου η τ.μ. X δεν ακολουθεί κάποια διακριτή κατανομή και θα πρέπει να κάνουμε διακριτοποίηση όπως θα δούμε παρακάτω, ίσο με το βήμα. Στην περίπτωση που η τ.μ. X ακολουθεί κάποια διακριτή κατανομή τότε το *x.scale* = 1.

Για να γίνουν πιο κατανοητά τα όσα αναφέρθηκαν ακολουθεί το παράδειγμα 4.8 από το βιβλίο του Dickson (2006), Insurance Risk and Ruin, στο οποίο θα γίνει υπολογισμός της $g(x) = P(S = x)$ μέσω της R.

Παράδειγμα 4.2.1

Έστω ότι $N \sim P(2)$ και $f(x) = 0.6(0.4^{x-1})$ για $x = 1, 2, 3, \dots$. Θα υπολογιστούν οι τιμές της $g(x)$ για $x = 1, 2, 3$.

Λύση

Για την επίλυση του παραδείγματος θα χρησιμοποιήσουμε τη σχέση που μας δίνει η πρόταση 4.1.2.

Αρχικά έχουμε ότι

$$g(0) = e^{-2} = 0.1353.$$

Στη συνέχεια αφού η τ.μ. ακολουθεί την κατανομή Poisson με παράμετρο $\lambda = 2$, σύμφωνα με αυτά που προαναφέρθηκαν στην ενότητα 2.1 συμπεραίνουμε ότι $a = 0$ και $b = 2$. Οπότε από την πρόταση 4.1.1 προκύπτει ότι,

$$g(x) = \frac{2}{x} \sum_{y=1}^x y f(y) g(x-y).$$

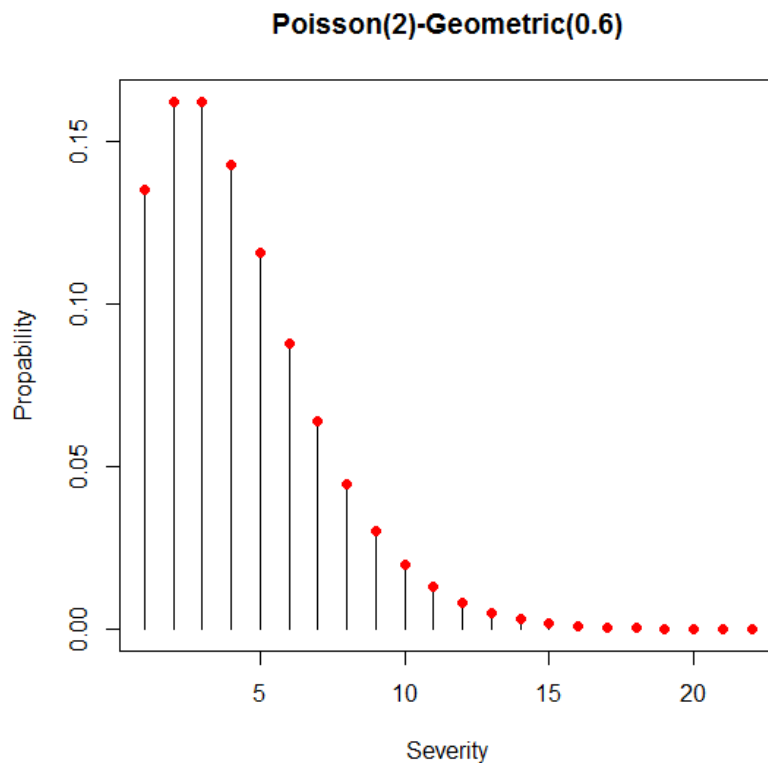
Άρα για $x = 1, 2, 3$ προκύπτουν τα εξής,

$$g(1) = 2f(1)g(0) = 0.1624,$$

$$g(2) = f(1)g(1) + 2f(2)g(0) = 0.1624,$$

$$g(3) = \frac{2}{3}(f(1)g(2) + 2f(2)g(1) + 3f(3)g(0)) = 0.1429.$$

Στο Παράρτημα Ι παρατίθεται ο κώδικας που χρησιμοποιήθηκε στην γλώσσα προγραμματισμού R, με τον οποίο επαληθεύτηκαν τα αποτελέσματα. Βάσει του κώδικα αυτού μπορούμε να υπολογίσουμε γρηγορότερα την πιθανότητα για κάθε τιμή του x . Στη συνέχεια παρατίθεται σχήμα στο οποίο παρουσιάζεται η συνάρτηση πιθανότητας της τ.μ. S .



Αυτό που παρατηρούμε από το σχήμα αλλά και από τους υπολογισμούς μας παραπάνω είναι ότι η πιθανότητα το x να πάρει τις τιμές 1 και 2 είναι ίσες και στη συνέχεια όσο αυξάνει το x τόσο μικρότερες πιθανότητες έχουμε να συμβεί το συγκεκριμένο ενδεχόμενο. Τέλος για τιμές του x μεγαλύτερες του 16 η πιθανότητα τείνει στο μηδέν.

4.3 Διακριτοποίηση της κατανομής συνεχών τυχαίων μεταβλητών

Το επόμενο πρόβλημα που πρέπει να αντιμετωπιστεί είναι στην περίπτωση που η τ.μ. X ακολουθεί συνεχή κατανομή. Με την R λοιπόν μας δίνεται η δυνατότητα στις κλάσεις κατανομών $(a, b, 0)$ και $(a, b, 1)$ για την κατανομή του ύψους των ζημιών εκτός από τις διακριτές κατανομές να χρησιμοποιήσουμε και συνεχείς, αφού πρώτα τις διακριτοποιήσουμε. Με την διακριτοποίηση αυτό που επιτυγχάνουμε είναι να μετατρέψουμε μία συνεχή κατανομή σε διακριτή χωρίζοντας το πεδίο τιμών σε πολλά ίσα τμήματα.

Η συνάρτηση αυτή ονομάζεται *discretize* ή *discritise*, βρίσκεται στο πακέτο *actuar* και ορίζεται ως εξής,

discretize(cdf, from, to, step, method) .

Οι μέθοδοι με τις οποίες γίνεται η διακριτοποίηση στην *R* είναι η “upper”, η “lower”, η “rounding” και η “unbiased”. Αρχικά λοιπόν στην παραπάνω συνάρτηση πρέπει να ορίσουμε τη σ.κ. (*cdf*) που θέλουμε να διακριτοποιήσουμε. Έπειτα επιλέγουμε το τμήμα της κατανομής το οποίο θέλουμε να διακριτοποιήσουμε με τις μεταβλητές *from, to* και τέλος το βήμα (*step*) που θα χωρίσουμε το επιλεγμένο τμήμα της σ.κ.

Στη συνέχεια θα δείξουμε με ποιο τρόπο γίνεται η διακριτοποίηση με την κάθε μέθοδο. Αρχικά θα ορίσουμε κάποιες μεταβλητές οι οποίες θα χρειαστούν για τις μεθόδους που αναφέρθηκαν. Με $F(x)$ θα συμβολίζεται η συνάρτηση κατανομής της τ.μ. X , με p_x η μάζα πιθανότητας για την τιμή x , με a, b το κάτω άκρο και το άνω άκρο αντίστοιχα των τιμών που μπορεί να πάρει το x και με h το βήμα κατά το οποίο θα αυξάνεται το κάτω άκρο μέχρι να φτάσει στο άνω(βήμα διακριτοποίησης).

1) Μέθοδος “Upper”

Με τη μέθοδο αυτή το p_x προκύπτει από την παρακάτω σχέση,

$$p_x = F(x + h) - F(x)$$

για $x = a, a + h, \dots, b - h$.

2) Μέθοδος “Lower”

Η μεθοδολογία αυτή είναι παρεμφερής με την *upper* και το p_x υπολογίζεται από τη σχέση,

$$p_x = F(x) - F(x - h)$$

για $x = a + h, \dots, b$ και $p_a = F(a)$.

3) Μέθοδος "Rounding"

Σύμφωνα με τη μέθοδο αυτή η μάζα πιθανότητας στο x υπολογίζεται σύμφωνα με τον παρακάτω τύπο

$$p_x = F\left(x + \frac{h}{2}\right) - F\left(x - \frac{h}{2}\right)$$

όπου $x = a, a + h, \dots, b - h$ και $p_a = F\left(a + \frac{h}{2}\right)$.

4) Μέθοδος "Unbiased"

Η μέθοδος αυτή συνδυάζει την πρώτη ροπή της διακριτοποιημένης κατανομής και της πραγματικής κατανομής της X και συνεπώς οι πιθανότητες υπολογίζονται ως εξής:

$$p_a = \frac{E[\min(X, a)] - E[\min(X, a + h)]}{h} + 1 - F(a)$$

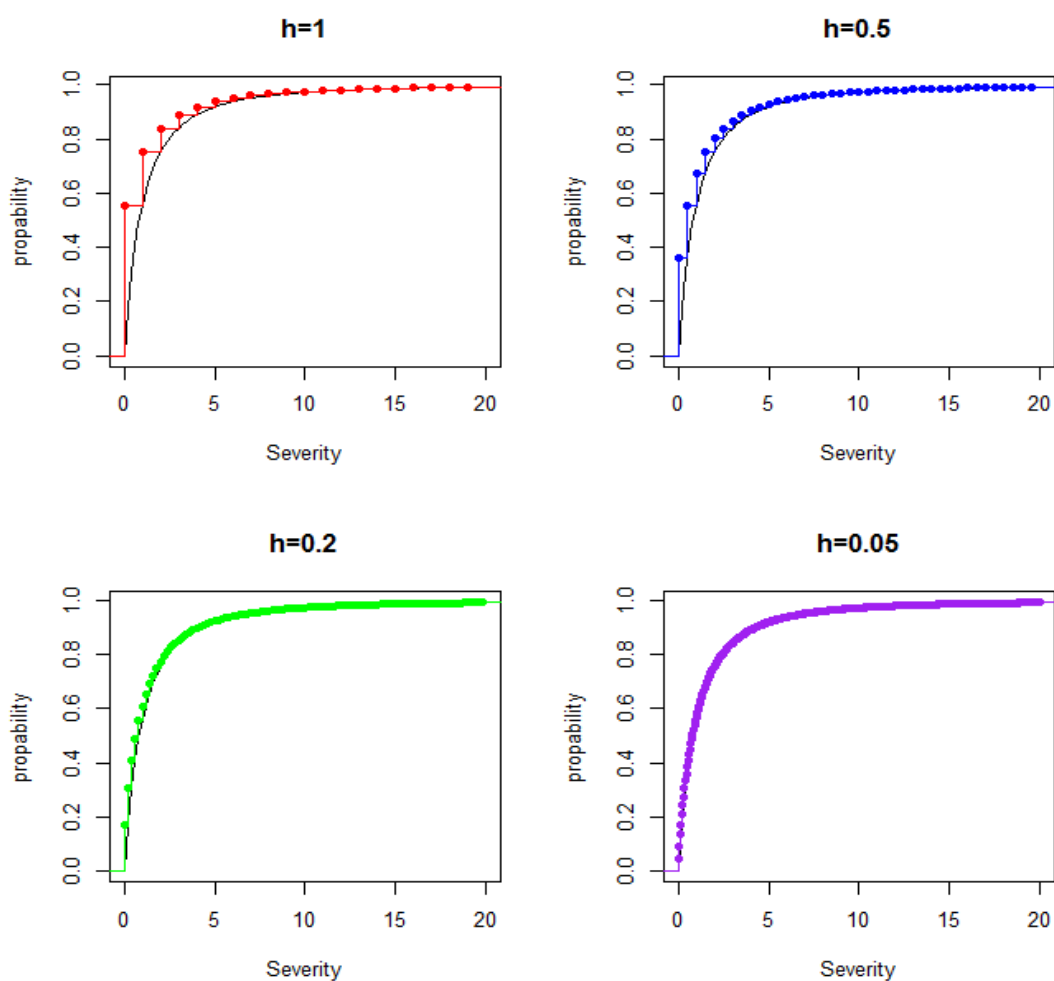
$$p_x = \frac{2E[\min(X, x)] - E[\min(X, x - h)] - E[\min(X, x + h)]}{h}, a < x < b$$

$$p_b = \frac{E[\min(X, b)] - E[\min(X, b - h)]}{h} - 1 + F(a).$$

Συνεπώς ενδεικτικά αναφέρεται ένα παράδειγμα για την καλύτερη κατανόηση του τρόπου με τον οποίο πρέπει να γράφεται ο κώδικας σε κάθε μία από τις παραπάνω περιπτώσεις. Ως μοντέλο ατομικών ζημιών θα χρησιμοποιηθεί η συνεχής κατανομή Pareto(2,2). Θα δειχθεί επίσης σχηματικά τι διαφορές προκύπτουν αν ορίσουμε μικρότερο βήμα διακριτοποίησης (step) καθώς και τι διαφορές προκύπτουν ανάλογα με τη μέθοδο. Η συγκεκριμένη κατανομή έχει άπειρη διακύμανση και πολύ βαριά ουρά και δεν χρησιμοποιείται συχνά στον αναλογισμό αλλά στη συγκεκριμένη περίπτωση την χρησιμοποιούμε για να αναδείξουμε τις διαφορές ανάλογα με τη μέθοδο διακριτοποίησης.

Παράδειγμα 4.3.1

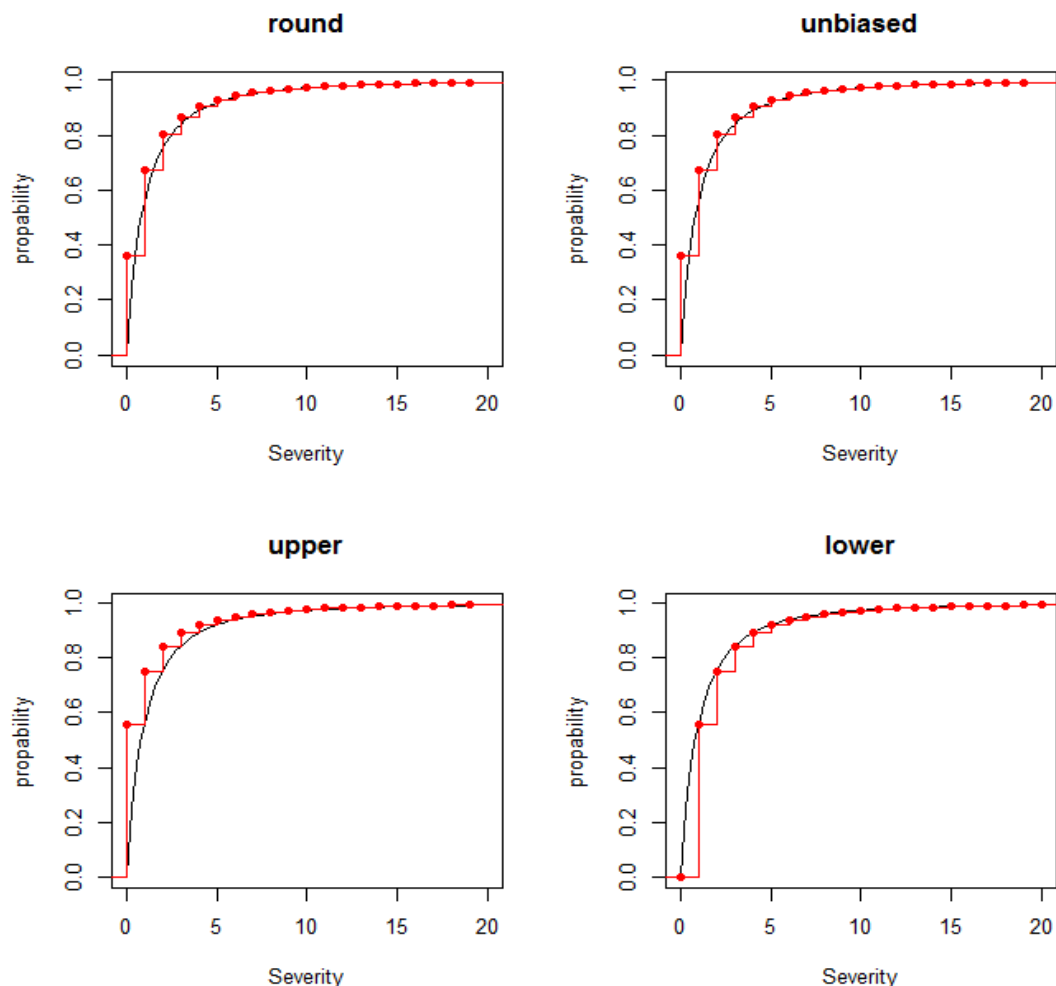
Στο παράδειγμα αυτό θα διακριτοποιήσουμε μία $\text{Pareto}(2,2)$ για τιμές του x από 0 έως 20 και κάθε φορά θα αλλάζουμε το βήμα με σκοπό να εξετάσουμε τις διαφορές που προκύπτουν σχηματικά.



Σχήμα 4.3.1

Όπως παρατηρείται από το παραπάνω σχήμα αυτό που προκύπτει, όπως ήταν αναμενόμενο, όσο πιο μικρό βήμα (*step*) επιλεγεί τόσο πιο κοντά έρχεται η διακριτοποιημένη κατανομή στη συνεχή. Συνεπώς θα πρέπει να επιλέγεται το δυνατότερο μικρό βήμα έτσι ώστε οι αποκλίσεις να είναι αμελητέες μεταξύ των δύο κατανομών.

Στη συνέχεια στην κατανομή Pareto(2,2) θα εφαρμόσουμε όλες τις μεθόδους διακριτοποίησης προκειμένου να εξετάσουμε διαγραμματικά τις αποκλίσεις μεταξύ της πραγματικής σ.κ. με την σ.κ. που προέκυψε κατόπιν διακριτοποίησης της τ.μ. X .



Σχήμα 4.3.2

Αυτό που αξίζει να σημειωθεί είναι ότι με τη μέθοδο round και την unbiased φαίνεται ότι προσεγγίζεται καλύτερα η συνεχής κατανομή σε αυτή την περίπτωση. Αυτό γίνεται εμφανές για τις τιμές το x από 0 έως 5 διότι στα υπόλοιπα σημεία οι αποκλίσεις είναι αρκετά μικρές και δεν φαίνονται διαγραμματικά.

Στην επόμενη ενότητα θα αναλυθούν διεξοδικότερα κάποιες εφαρμογές και θα δειχθεί από ποιες παραμέτρους επηρεάζεται η καταλληλότητα των προαναφερθεισών μεθόδων.

4.4 Εφαρμογές

Στην παρούσα ενότητα θα παρουσιάσουμε τρεις εφαρμογές, στις οποίες αρχικά θα υπολογίσουμε την πραγματική συνάρτηση κατανομής σύμφωνα με τους τύπους που ορίσαμε στην ενότητα 4.1. Έπειτα θα υπολογίσουμε τη σ.κ. που θα προκύψει αφού διακριτοποιήσουμε τη συνεχή κατανομή με κάθε μία από τις μεθόδους που προαναφέραμε και θα σχολιαστούν οι αποκλίσεις αλλά και ποία μέθοδος προσαρμόζεται καλύτερα. Τέλος εξετάζεται πόσο επηρεάζει το βήμα που επιλέγεται κατά τη διαδικασία της διακριτοποίησης.

Εφαρμογή 4.4.1

Στην εφαρμογή αυτή το πλήθος των ζημιών ακολουθεί την γεωμετρική κατανομή με συνάρτηση πιθανότητας

$$p_n = (0.1)(0.9)^n, n = 0,1,2, \dots$$

και το ύψος ζημιών ακολουθεί εκθετική κατανομή με συνάρτηση πυκνότητας πιθανότητας

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

Αφού λοιπόν γνωρίζουμε τις κατανομές των τυχαίων μεταβλητών N, X αρχικά θα υπολογιστεί, με την χρήση της σχέσης (4.1.1), η συνάρτηση κατανομής των συνολικών απαιτήσεων S η οποία είναι η

$$G(x) = 1 - 0.9e^{-0.2x}, x \geq 0.$$

Στη συνέχεια χρησιμοποιώντας την γλώσσα προγραμματισμού R θα υπολογίσουμε τις τιμές της συνάρτησης κατανομής των συνολικών απαιτήσεων S χρησιμοποιώντας τις τέσσερις διαφορετικές μεθόδους διακριτοποίησης για την τ.μ. X η οποία είναι συνεχής. Η διακριτοποίηση έγινε με βήμα $h = 0.001$ για να έχουμε όσο το δυνατόν ακριβέστερα αποτελέσματα. Τέλος αφού υπολογίσαμε όλα τα παραπάνω παρατίθεται πίνακας στον οποίο παρουσιάζονται συγκριτικά τα αποτελέσματα.

Σύγκριση μεθόδων διακριτοποίησης					
x	G(x)	Upper	Lower	Round	Unbiased
0	0,1000000	0,1000000	0,1000000	0,1000000	0,1000000
1	0,2631423	0,2630097	0,2625985	0,2628041	0,2628040
2	0,3967120	0,3964949	0,3959414	0,3962181	0,3962180
3	0,5060695	0,5058029	0,5051722	0,5054875	0,5054874
4	0,5956039	0,5953128	0,5946509	0,5949819	0,5949818
5	0,6689085	0,6686105	0,6679494	0,6682800	0,6682800
6	0,7289252	0,7286324	0,7279935	0,7283130	0,7283130
7	0,7780627	0,7777831	0,7771799	0,7774816	0,7774815
8	0,8182931	0,8180314	0,8174720	0,8177518	0,8177518
9	0,8512310	0,8509899	0,8504782	0,8507342	0,8507342
10	0,8781982	0,8779789	0,8775160	0,8777476	0,8777475

Πίνακας 4.4.1.1

Λόγω του ότι οι διαφορές μεταξύ των αποτελεσμάτων φαίνονται να είναι σχετικά μικρές ακολουθεί πίνακας στον οποίο παρουσιάζονται οι αποκλίσεις της κάθε μεθόδου.

Αποκλίσεις				
x	Upper	Lower	Round	Unbiased
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0001326	0,0005438	0,0003382	0,0003383
2	0,0002171	0,0007706	0,0004939	0,0004940
3	0,0002666	0,0008973	0,0005820	0,0005821
4	0,0002911	0,0009530	0,0006220	0,0006221
5	0,0002980	0,0009591	0,0006285	0,0006285
6	0,0002928	0,0009317	0,0006122	0,0006122
7	0,0002796	0,0008828	0,0005811	0,0005812
8	0,0002617	0,0008211	0,0005413	0,0005413
9	0,0002411	0,0007528	0,0004968	0,0004968
10	0,0002193	0,0006822	0,0004506	0,0004507

Πίνακας 4.4.1.2

Αρχικά βλέπουμε ότι για $x = 0$ δεν προκύπτει κάποια διαφορά κάτι το οποίο ήταν αναμενόμενο αφού η εκθετική κατανομή δεν χρησιμοποιείται στη συγκεκριμένη περίπτωση και η συνάρτηση κατανομής εξαρτάται μόνο από την κατανομή του πλήθους των ζημιών.

Έπειτα βλέπουμε ότι οι μεγαλύτερες αποκλίσεις προκύπτουν από την μέθοδο round και μετά ακολουθούν ο μέθοδοι lower και unbiased. Η μέθοδος που φαίνεται να προσεγγίζει ακριβέστερα τα πραγματικά δεδομένα είναι η upper, της οποίας οι αποκλίσεις κυμαίνονται από 0,0001326 έως 0,0002980. Όμως αυτό που πρέπει να

σημειωθεί είναι ότι, όπως φαίνεται, όλες οι μέθοδοι έχουν αρκετά μικρές αποκλίσεις.

Ακόμη αυτό που αξίζει να ελεγχθεί είναι το πόσο επηρεάζει το βήμα που επιλέγουμε κατά τη διαδικασία της διακριτοποίησης. Για τον λόγο αυτό για κάθε μία από τις παραπάνω μεθόδους αλλάζουμε το βήμα από 0.001 σε 0.01, 0.1 και 0.2. Αυτό που παρατηρήσαμε από την αλλαγή του βήματος είναι ότι στην συγκεκριμένη περίπτωση αλλάζοντας το βήμα συνεχίζει η μέθοδος Uppser να αντιπροσωπεύει καλύτερα την κατανομή μας.

Στη συνέχεια ακολουθεί πίνακας στον οποίο παρουσιάζονται οι αποκλίσεις που προέκυψαν ύστερα από την αλλαγή του βήματος με τη μέθοδο Uppser.

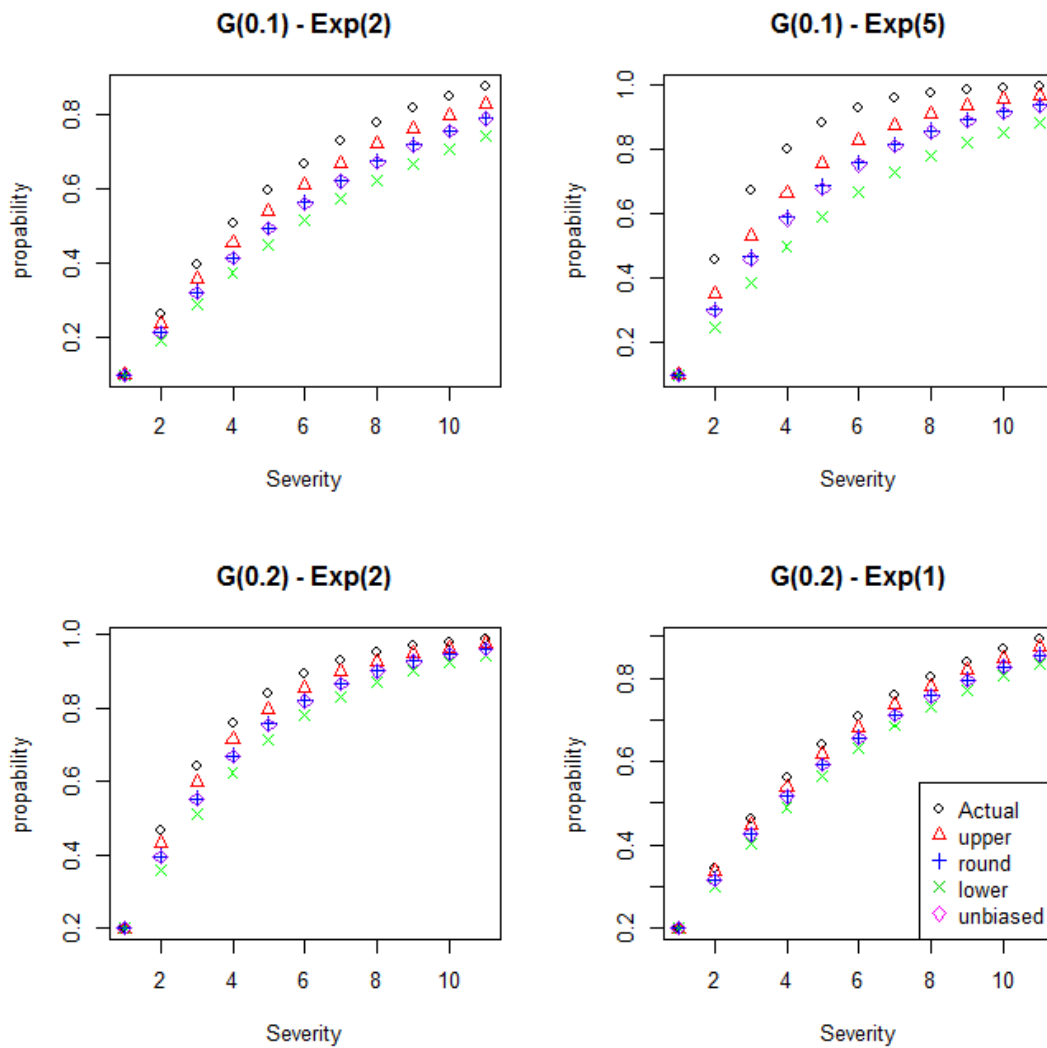
Αποκλίσεις				
x	h=0,001	h=0,01	h=0,1	h=0,2
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0001326	0,0013204	0,0126844	0,0242511
2	0,0002171	0,0021642	0,0209490	0,0403638
3	0,0002666	0,0026602	0,0259494	0,0503907
4	0,0002911	0,0029065	0,0285727	0,0559235
5	0,0002980	0,0029773	0,0294955	0,0581899
6	0,0002928	0,0029277	0,0292309	0,0581313
7	0,0002796	0,0027990	0,0281647	0,0564646
8	0,0002617	0,0026214	0,0265841	0,0537310
9	0,0002411	0,0024167	0,0247009	0,0503353
10	0,0002193	0,0022004	0,0226682	0,0465760

Πίνακας 4.4.1.3

Σύμφωνα λοιπόν με τον παραπάνω πίνακα όσο αυξάνουμε το βήμα τόσο μεγαλώνουν οι αποκλίσεις μας από τα πραγματικά αποτελέσματα. Αυτό είναι λογικό και αναμενόμενο γιατί κατά τη διαδικασία της διακριτοποίησης όσο μικρότερο είναι το βήμα τόσο πιο κοντά «πλησιάζει» η διακριτοποιημένη με την συνεχή κατανομή. Επίσης παρατηρούμε ότι για $x = 10$, ενώ η απόκλιση ήταν 0.0002193 με βήμα 0.001 αυξήθηκε σε 0.0465760 με βήμα 0.2 κάτι το οποίο μπορεί να κάνει την διακριτοποιημένη κατανομή μη αξιόπιστη καθώς το σφάλμα όλο και μεγαλώνει αυξάνοντας το βήμα.

Τέλος, παρατίθεται το ακόλουθο σχήμα με το οποίο θα δειχθεί εάν αλλάζοντας τις παραμέτρους των κατανομών θα συνεχίσει η μέθοδος Uppser να είναι καταλληλότερη σε σχέσεις με τις υπόλοιπες μεθόδους. Να σημειωθεί ότι για να μπορούμε να συγκρίνουμε καλύτερα τα αποτελέσματα επιλέχθηκε βήμα

διακριτοποίησης 0.2, διότι διαφορετικά λόγω των πολύ μικρών διαφορών δεν θα είναι ευδιάκριτες οι διαφορές σχηματικά.



Σχήμα 4.4.1.1

Σύμφωνα με το παραπάνω σχήμα αυτό που παρατηρείται είναι ότι αλλάζοντας τις παραμέτρους των κατανομών N, X και κρατώντας σταθερό το βήμα αλλάζουν σε κάθε περίπτωση οι αποκλίσεις αλλά πάντοτε η μέθοδος Υπερπλησιάζει πιο «κοντά» στην πραγματική τιμή.

Εφαρμογή 4.4.2

Σε αυτήν την εφαρμογή θα εξετάσουμε την περίπτωση όπου το πλήθος των ζημιών ακολουθεί την διωνυμική κατανομή με συνάρτηση πιθανότητας

$$p_n = \binom{100}{n} (0.1)^n (0.9)^{100-n}, \quad 0 \leq n \leq 100$$

και το ύψος ζημιών ακολουθεί εκθετική κατανομή με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = 3e^{-3x}, \quad x > 0.$$

Σύμφωνα λοιπόν με την σχέση (4.1.2) μπορούμε να υπολογίσουμε με ακρίβεια την συνάρτηση κατανομή της τυχαίας μεταβλητής S από τη σχέση

$$G(x) = 1 - e^{-3x} \sum_{n=1}^{100} \binom{100}{n} 0.1^n 0.9^{100-n} \sum_{j=0}^{n-1} \frac{(3x)^j}{j!}, \quad x \geq 0$$

Στον πίνακα που ακολουθεί παρουσιάζονται τα αποτελέσματα για κάθε μία από τις μεθόδους διακριτοποίησης έχοντας επιλέξει βήμα $h=0.001$.

Σύγκριση μεθόδων διακριτοποίησης					
x	G(x)	Upper	Lower	Round	Unbiased
0	0,0000266	0,0000266	0,0000266	0,0000266	0,0000266
1	0,0268918	0,0267071	0,0262678	0,0264868	0,0264867
2	0,1815328	0,1807159	0,1788645	0,1797891	0,1797888
3	0,4512293	0,4499858	0,4472215	0,4486039	0,4486035
4	0,7063965	0,7053086	0,7029152	0,7041135	0,7041132
5	0,8708035	0,8701361	0,8686771	0,8694083	0,8694081
6	0,9517232	0,9514052	0,9507128	0,9510601	0,9510600
7	0,9842646	0,9841395	0,9838678	0,9840042	0,9840042
8	0,9954299	0,9953876	0,9952960	0,9953420	0,9953420
9	0,9987972	0,9987846	0,9987573	0,9987710	0,9987710
10	0,9997093	0,9997059	0,9996986	0,9997023	0,9997023

Πίνακας 4.4.2.1

Παρατηρώντας λοιπόν τα δεδομένα του παραπάνω πίνακα παρατηρούμε ότι όλες οι μέθοδοι έχουν μικρές αποκλίσεις μεταξύ τους άρα με την διακριτοποίηση επιλέγοντας μικρό βήμα τα αποτελέσματα που προκύπτουν φαίνονται αξιόπιστα. Επειδή οι διαφορές είναι αρκετά μικρές στη συνέχεια ακολουθεί πίνακας με τις αποκλίσεις.

Αποκλίσεις				
x	Upper	Lower	Round	Unbiased
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0001847	0,0006240	0,0004050	0,0004051
2	0,0008170	0,0026683	0,0017438	0,0017440
3	0,0012435	0,0040078	0,0026254	0,0026258
4	0,0010879	0,0034813	0,0022830	0,0022833
5	0,0006674	0,0021264	0,0013952	0,0013954
6	0,0003181	0,0010105	0,0006632	0,0006632
7	0,0001251	0,0003968	0,0002604	0,0002604
8	0,0000423	0,0001339	0,0000879	0,0000879
9	0,0000126	0,0000399	0,0000262	0,0000262
10	0,0000034	0,0000107	0,0000070	0,0000070

Πίνακας 4.4.2.2

Αυτό που προκαλεί εντύπωση είναι ότι και σε αυτήν την περίπτωση οι μικρότερες διαφορές προκύπτουν με τη μέθοδο Upper. Σε αυτή την περίπτωση όμως αμέσως μετά προσαρμόζεται καλύτερα η μέθοδος Round και η Unbiased, οι οποίες έχουν ελάχιστη διαφορά μεταξύ τους και τέλος η Lower η οποία παρουσιάζει τριπλάσιο σφάλμα σε σχέση με την Upper.

Επιπροσθέτως μελετήθηκε τι αποτελέσματα παίρνουμε όταν αλλάξουμε το βήμα. Το αποτέλεσμα που λάβαμε ήταν ότι και πάλι η μέθοδος Upper έχει τις μικρότερες αποκλίσεις όπως φαίνεται και στον ακόλουθο πίνακα.

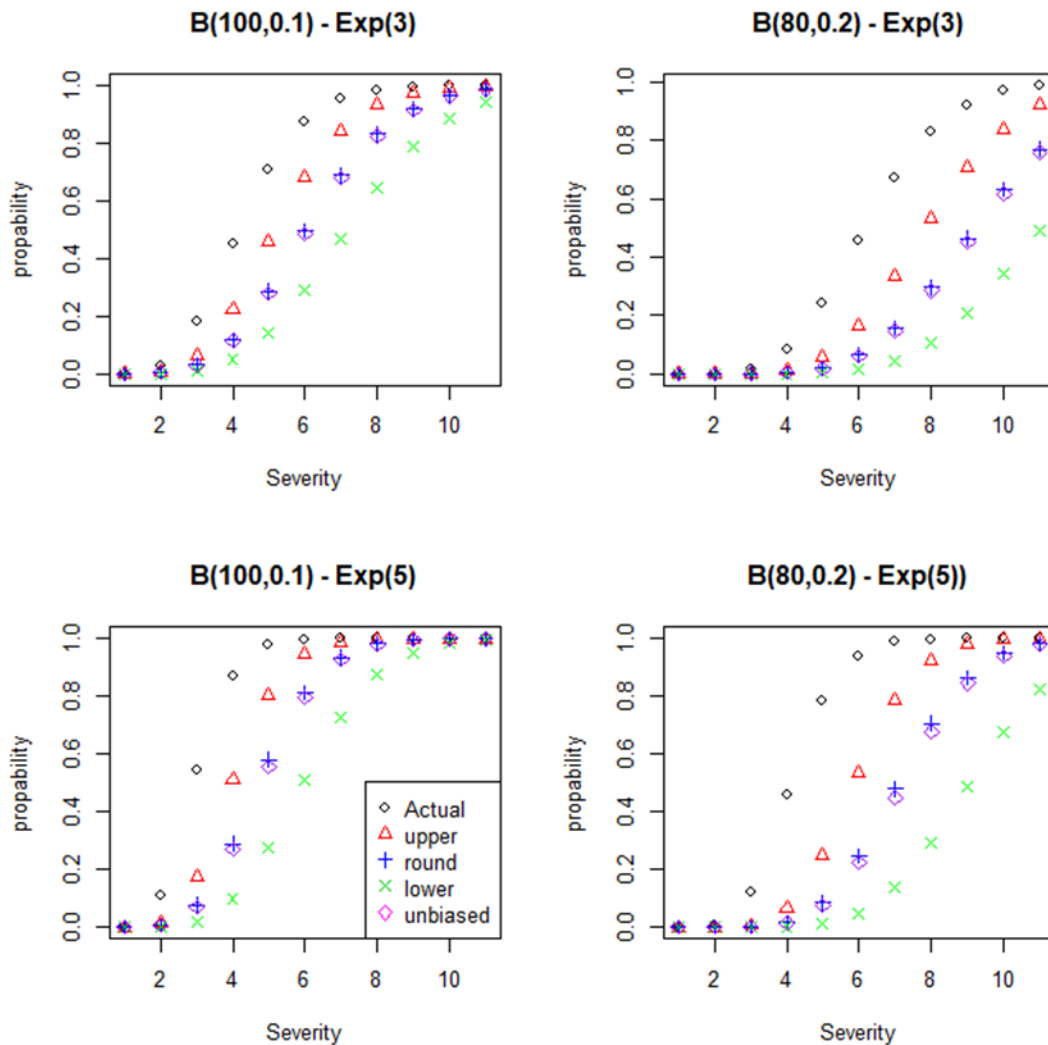
Αποκλίσεις				
x	h=0,001	h=0,01	h=0,1	h=0,2
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0001847	0,0017962	0,0135798	0,0202294
2	0,0008170	0,0080683	0,0700813	0,1169614
3	0,0012435	0,0124290	0,1212800	0,2267434
4	0,0010879	0,0109837	0,1188391	0,2487462
5	0,0006674	0,0067984	0,0808381	0,1892838
6	0,0003181	0,0032658	0,0423957	0,1109322
7	0,0001251	0,0012943	0,0182484	0,0532883
8	0,0000423	0,0004404	0,0067152	0,0218513
9	0,0000126	0,0001323	0,0021735	0,0078681
10	0,0000034	0,0000358	0,0006317	0,0025395

Πίνακας 4.4.2.3

Όπως και στην προηγούμενη εφαρμογή όσο μικρότερο είναι το βήμα που επιλέγεται τόσο πιο μικρές αποκλίσεις παρατηρούνται μεταξύ της πραγματικής και της διακριτοποιημένης κατανομής. Επίσης αξίζει να σημειωθεί ότι καθώς το βήμα

διακριτοποίησης δεκαπλασιάζεται παρατηρούμε ότι και το σφάλμα δεκαπλασιάζεται.

Στη συνέχεια αλλάζουμε τις παραμέτρους των κατανομών N, X με τέσσερις διαφορετικούς συνδυασμούς. Δεδομένου ότι έχουμε λάβει βήμα $h = 0.2$ παρατίθενται στη συνέχεια τα αντίστοιχα διαγράμματα για να συγκρίνουμε τις μεθόδους διακριτοποίησης σε κάθε μία περίπτωση.



Σχήμα 4.4.2.1

Σύμφωνα λοιπόν με το παραπάνω σχήμα φαίνεται σε όλες τις περιπτώσεις να είναι πιο κοντά στην πραγματική κατανομή, η κατανομή που έχει προκύψει κατόπιν διακριτοποίησης με τη μέθοδο Upper. Επιπλέον παρατηρείται ότι σε όλες τις περιπτώσεις μεγαλύτερες αποκλίσεις έχει η μέθοδος lower το οποίο έρχεται σε αντίθεση με τα όσα αναφέραμε όταν είχαμε βήμα $h = 0.001$. (Πίνακας 4.4.2.2)

Εφαρμογή 4.4.3

Στην εφαρμογή αυτή θα εξετάσουμε την περίπτωση όπου το πλήθος των ζημιών ακολουθεί την αρνητική διωνυμική κατανομή με σ.π.π. την

$$p_n = \binom{n+9}{10} (0.4)^{10} (0.6)^n, \quad n = 0, 1, 2, \dots$$

και το ύψος ζημιών ακολουθεί εκθετική κατανομή με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = 3e^{-3x}, \quad x > 0.$$

Αρχικά κάνοντας χρήση της σχέσης (4.1.3) υπολογίζουμε για $x = 0, 1, 2, \dots, 10$ την συνάρτηση κατανομής της τυχαίας μεταβλητής S η οποία είναι

$$G(x) = 1 - e^{-1.2x} \sum_{j=0}^9 \frac{(1.2x)^j}{j!} \sum_{n=j+1}^{10} \binom{10}{n} 0.6^n 0.4^{10-n}, \quad x \geq 0.$$

Στη συνέχεια υπολογίζουμε την συνάρτηση κατανομής της S με τον αναδρομικό τύπο του Panjer έχοντας διακριτοποιήσει την κατανομή του ύψους των ζημιών με βήμα $h = 0.001$. Τα αποτελέσματα που προέκυψαν παρουσιάζονται στον παρακάτω πίνακα.

Σύγκριση μεθόδων διακριτοποίησης					
x	G(x)	Upper	Lower	Round	Unbiased
0	0,0001049	0,0001049	0,0001049	0,0001049	0,0001049
1	0,0156573	0,0155758	0,0153788	0,0154771	0,0154770
2	0,0829344	0,0825693	0,0817466	0,0821572	0,0821571
3	0,2125452	0,2117911	0,2101387	0,2109640	0,2109638
4	0,3805772	0,3795239	0,3772517	0,3783874	0,3783871
5	0,5513263	0,5501757	0,5477173	0,5489469	0,5489465
6	0,6980314	0,6969724	0,6947247	0,6958496	0,6958493
7	0,8093920	0,8085344	0,8067221	0,8076296	0,8076294
8	0,8862004	0,8855716	0,8842472	0,8849108	0,8849106
9	0,9352685	0,9348429	0,9339487	0,9343970	0,9343969
10	0,9647041	0,9644345	0,9638690	0,9641527	0,9641526

Πίνακας 4.4.3.1

Όπως παρατηρούμε από τον παραπάνω πίνακα με όλες τις μεθόδους τα αποτελέσματα που προκύπτουν είναι πάρα πολύ κοντά. Για να δούμε με ποια μέθοδο διακριτοποίησης είμαστε πιο κοντά στην πραγματική κατανομή ακολουθεί πίνακας στον οποίο παρουσιάζονται οι διαφορές μεταξύ της πραγματικής και της διακριτοποιημένης κατανομής.

Αποκλίσεις				
x	Upper	Lower	Round	Unbiased
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0000815	0,0002785	0,0001803	0,0001803
2	0,0003651	0,0011878	0,0007772	0,0007773
3	0,0007541	0,0024065	0,0015812	0,0015814
4	0,0010532	0,0033255	0,0021898	0,0021901
5	0,0011506	0,0036090	0,0023795	0,0023798
6	0,0010589	0,0033067	0,0021818	0,0021821
7	0,0008576	0,0026699	0,0017624	0,0017626
8	0,0006288	0,0019532	0,0012896	0,0012898
9	0,0004255	0,0013197	0,0008714	0,0008715
10	0,0002696	0,0008351	0,0005514	0,0005515

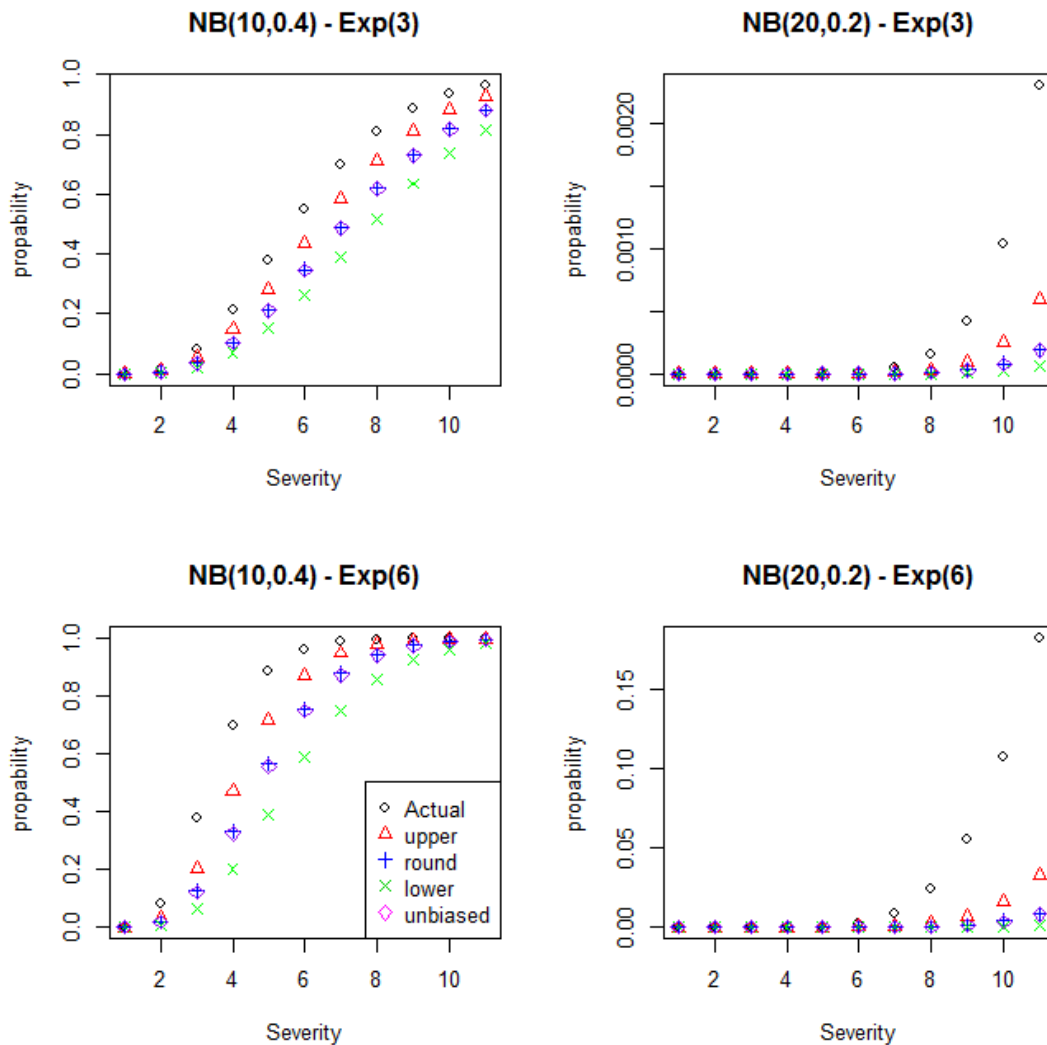
Πίνακας 4.4.3.2

Η μέθοδος που προσαρμόζεται καλύτερα είναι η Upper και μετά ακολουθούν η Round, η Unbiased και τέλος η Lower. Στη συνέχεια αλλάζουμε το βήμα διακριτοποίησης από 0.001 σε 0.01, 0.01, 0.2 και αυτό που παρατηρούμε είναι ότι η μέθοδος Upper προσαρμόζεται καλύτερα σε όλες τις περιπτώσεις και έπειτα η Round, η Unbiased και τέλος η Lower. Στον πίνακα που παρουσιάζεται στη συνέχεια φαίνονται οι διαφορές με την μέθοδο Upper σε κάθε μία περίπτωση. Όσο το βήμα αυξάνεται βλέπουμε ότι τα αποτελέσματα αποκλίνουν αρκετά από τις πραγματικές τιμές.

Αποκλίσεις				
x	h=0,001	h=0,01	h=0,1	h=0,2
0	0,0000000	0,0000000	0,0000000	0,0000000
1	0,0000815	0,0007953	0,0062464	0,0097766
2	0,0003651	0,0035852	0,0298883	0,0489949
3	0,0007541	0,0074534	0,0657605	0,1131878
4	0,0010532	0,0104720	0,0976517	0,1768489
5	0,0011506	0,0115064	0,1133393	0,2164298
6	0,0010589	0,0106492	0,1107853	0,2235608
7	0,0008576	0,0086726	0,0953004	0,2036910
8	0,0006288	0,0063932	0,0742290	0,1684270
9	0,0004255	0,0043499	0,0533868	0,1288931
10	0,0002696	0,0027704	0,0359604	0,0925902

Πίνακας 4.4.3.3

Τέλος ακολουθεί σχήμα στο οποίο σε κάθε μία από τις τέσσερις περιπτώσεις έχουμε αλλάξει τις παραμέτρους των κατανομών των τυχαίων μεταβλητών N και X . Το βήμα που χρησιμοποιήθηκε σε κάθε περίπτωση είναι 0.2 για να είναι ευδιάκριτες οι διαφορές σχηματικά.



Σχήμα 4.4.3.1

Παρατηρούμε λοιπόν ότι αλλάζοντας τις παραμέτρους σε αυτήν την περίπτωση δεν αλλάζει η σειρά με την οποία οι μέθοδοι αντιπροσωπεύουν καλύτερα την αρχική συνάρτηση κατανομής. Αυτό όμως που αξίζει να σημειωθεί είναι ότι στις περιπτώσεις όπου η $N \sim NB(20,0.2)$ η διακριτοποίηση δεν είναι καθόλου καλή καθώς από τα διαγράμματα παρατηρούμε πολύ μεγάλες αποκλίσεις σε σχέση με την πραγματική κατανομή.

Συνοψίζοντας τα συμπεράσματα από κάθε μία από τις 3 εφαρμογές καταλήγουμε στο συμπέρασμα ότι όσο μικρότερο βήμα επιλέξουμε τόσο πιο κοντά θα είμαστε στην πραγματική κατανομή. Επίσης αν το βήμα είναι πολύ μικρό, όπως είδαμε 0.001, τότε οι διαφορές που προκύπτουν είναι αρκετά μικρές συνεπώς θεωρούμε την κατανομή μετά την διακριτοποίηση αξιόπιστη. Τέλος, όσον αφορά τις μεθόδους

διακριτοποίησης δεν μπορούμε να πούμε ότι κάποια είναι καλύτερη καθώς όπως είδαμε αυτό εξαρτάται από το βήμα αλλά και από τις παραμέτρους των κατανομών.

4.5 Συμπεράσματα

Κλείνοντας το κεφάλαιο και έχοντας κάνει κάποιες εφαρμογές στο ενότητα 4.4 προκύπτουν κάποια χρήσιμα συμπεράσματα όσο αφορά την διακριτοποίηση των συνεχών τ.μ. και κατ' επέκταση του υπολογισμού της σ.κ. της τ.μ. S . Αρχικά παρατηρούμε ότι σημαντικό ρόλο στην διακριτοποίηση έχει το βήμα που επιλέγουμε καθώς όσο πιο μικρό είναι τόσο μικρότερες διαφορές προκύπτουν μεταξύ της διακριτοποιημένης κατανομής με την πραγματική. Επίσης αυτό που παρατηρούμε είναι ότι δεν υπάρχει βέλτιστη μέθοδος διακριτοποίησης καθώς ανάλογα με τις παραμέτρους της κατανομής που θέλουμε να διακριτοποιήσουμε αλλάζει και η βέλτιστη μέθοδος. Τέλος ένα σημαντικό αποτέλεσμα που προέκυψε είναι ότι σε κάθε περίπτωση η πραγματική τιμή της σ.κ. της τ.μ. S ήταν μεγαλύτερη από την προκύπτουσα κατόπιν της διακριτοποίησης.

5 Η κλάση κατανομών $R(a, b, l)$

Στο κεφάλαιο αυτό θα εξετάσουμε αρχικά μία ευρύτερη κλάση κατανομών, που περιλαμβάνει τις δύο προηγούμενες κλάσεις που αναφέρθηκαν, η οποία μας δίνει μία αναδρομική σχέση για την κατανομή του πλήθους των ζημιών και ονομάζεται $R(a, b, l)$. Αρχικά δίνεται ο αναδρομικός σχέση που ισχύει για τον υπολογισμό της συνάρτησης πιθανότητας του πλήθους των ζημιών. Στη συνέχεια αναφέρονται οι κατανομές που ανήκουν σε αυτήν την κλάση και παρουσιάζονται κάποια σημαντικά αποτελέσματα. Τέλος κλείνοντας το κεφάλαιο θα αναλύσουμε την έννοια της λέπτυνσης η οποία έχει σημαντικές εφαρμογές τόσο στην ασφάλιση όσο και στην αντασφάλιση όπως για παράδειγμα στον υπολογισμό της συνάρτησης πυκνότητας πιθανότητας για τις συνολικές αποζημιώσεις όταν υπάρχει ίδια κράτηση.

5.1 Βασικοί ορισμοί και τύποι

Ξεκινώντας την ενότητα δίνεται η αναδρομική σχέση η οποία ισχύει για το πλήθος των ζημιών στην περίπτωση όπου η τ.μ. N ανήκει στην οικογένεια κατανομών $R(a, b, l)$ (Sundt and Vernic, 2009).

Ορισμός 5.1.1

Η διακριτή τ.μ. N για $n = 0, 1, 2, \dots$ θα ανήκει στην οικογένεια κατανομών $R(a, b, l)$ αν η συνάρτηση πιθανότητας ικανοποιεί την αναδρομική σχέση

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

όπου a, b είναι σταθερές, το l είναι θετικός ακέραιος αριθμός και $p_n = 0 \quad \forall n < 0$.

■

Στην ειδική περίπτωση όπου το l ισούται με 1 τότε καταλήγουμε στην σχέση

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, \quad \text{για } n = 2, 3, \dots$$

δηλαδή στη σχέση (3.1.1) η οποία καλείται και ως κλάση των Sundt-Jewell.

Η κλάση $R(a, b, l)$ είναι μία ευρύτερη κλάση σε σχέση με την κλάση $R(a, b, 0)$ και την $R(a, b, 1)$ συνεπώς αυτό που περιμένουμε είναι να ανήκουν περισσότερες κατανομές σε αυτή.

Οι βασικές κατανομές που ανήκουν στην ευρύτερη κλάση $R(a, b, l)$ είναι οι εξής:

- Διωνυμική Κατανομή
- Poisson
- Αρνητική Διωνυμική
- Λογαριθμική Κατανομή
- Εκτεταμένη αρνητική Διωνυμική Κατανομή
- Εκτεταμένη Λογαριθμική Κατανομή

Αξίζει σε αυτό το σημείο να αναφέρουμε ότι η εκτεταμένη λογαριθμική κατανομή, η οποία συμβολίζεται με $N \sim ELog(m, p)$, έχει συνάρτηση πυκνότητας πιθανότητας που δίνεται από τη σχέση,

$$p_n = \frac{\binom{n}{m}^{-1} p^n}{\sum_{j=m}^{\infty} \binom{j}{m}^{-1} p^j},$$

όπου $n \geq m$ και $0 < p \leq 1$.

Αν στην παραπάνω σχέση θέσουμε το $m = 1$ και χρησιμοποιώντας την γνωστή σχέση

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots,$$

καταλήγουμε στην συνάρτηση πιθανότητας της λογαριθμικής κατανομής.

Επίσης η ροπογεννήτρια συνάρτηση της εκτεταμένης λογαριθμικής κατανομής δίνεται από τη σχέση

$$P_N(t) = \frac{\sum_{n=m}^{\infty} \binom{n}{m}^{-1} (pt)^n}{\sum_{n=m}^{\infty} \binom{n}{m}^{-1} p^n}.$$

Εκτός από τις βασικές κατανομές που αναφέρθηκαν για τον αριθμό των ζημιών στην κλάση $R(a, b, l)$ ανήκουν και οι l -περικομμένες βασικές κατανομές.

Σε αυτό το σημείο θα ορίσουμε τότε μία κατανομή είναι l -περικομμένη δίνοντας τον ορισμό της.

Ορισμός 5.1.2

Η κατανομή της τ.μ. N^l καλείται l -περικομμένη (l -truncation of N) και ορίζεται ως

$$N^l = N | N > l$$

όπου η τ.μ. $N \in \{0, 1, 2, \dots\}$. Η συνάρτηση πιθανότητας της τ.μ. N^l είναι

$$p_n^l = P(N^l = n) = \begin{cases} 0, & n = 0, 1, \dots, l \\ \frac{p_n}{1 - \sum_{j=0}^{l-1} p_j}, & n = l + 1, l + 2, \dots \end{cases}$$

■

Απόδειξη

Αρχικά για $n = 0, 1, 2, \dots$ από την θεωρία πιθανοτήτων ισχύει ότι

$$\begin{aligned} p_n^l &= P(N^l = n) = P(N = n | N > l) \\ &= \frac{P(N = n)}{P(N > l)} = \frac{P(N = n)}{1 - \sum_{j=0}^{l-1} P(N = j)} = \frac{p_n}{1 - \sum_{j=0}^{l-1} p_j}. \end{aligned}$$

Τέλος, για $n = 0, 1, \dots, l$, $p_n^l = P(N^l = n) = P(N = n | N > l) = 0$,

κάτι που είναι αναμενόμενο αφού η κατανομή δεν έχει μάζα στα σημεία.

Αφού ορίσαμε τη συνάρτηση πιθανότητας στην περίπτωση όπου η τ.μ. $N \sim l$ -truncated κατανομή από τον ορισμό της πιθανογεννήτριας προκύπτει ότι

$$\begin{aligned} P_{N^l}(t) &= \sum_{n=l}^{\infty} t^n p_n^l = \sum_{n=l}^{\infty} t^n \frac{p_n}{1 - \sum_{j=0}^{l-1} p_j} \\ &= \frac{1}{1 - \sum_{j=0}^{l-1} p_j} \left(\sum_{n=0}^{\infty} t^n p_n - \sum_{j=0}^{l-1} p_j t^j \right) \end{aligned}$$

και τελικά καταλήγουμε στη σχέση

$$P_{N^l}(t) = \frac{P_N(t) - \sum_{j=0}^{l-1} p_j t^j}{1 - \sum_{j=0}^{l-1} p_j}$$

Αφού αναφέραμε ποιες κατανομές ανήκουν στην κλάση $R(a, b, l)$ στη συνέχεια δίνεται ο αναδρομικός τύπος με τον οποίο υπολογίζεται το συνολικό ύψος των αποζημιώσεων όταν το N ανήκει σε αυτήν την κλάση κατανομών και η τ.μ. X

Πρόταση 5.1.1

Αν η τ.μ. $X \in \{1, 2, \dots\}$ έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τ.μ. $N \in R(a, b, l)$, τότε η $S = X_1 + X_2 + \dots + X_N$ για $N \geq 1$ και με $S = 0$ αν $N = 0$ έχει σ.π. $g(x) = P(S = x)$ που ικανοποιεί την παρακάτω αναδρομική σχέση

$$g(x) = \sum_{n=1}^l \left(p_n - \left(a + \frac{b}{n} \right) p_{n-1} \right) f^{n*}(x) + \sum_{y=1}^x \left(a + b \frac{y}{x} \right) f(y) g(x-y)$$

$$g(x) = p_1 f(x) + \sum_{n=2}^l \left(p_n - \left(a + \frac{b}{n} \right) p_{n-1} \right) f^{n*}(x) + \sum_{y=1}^{x-1} \left(a + b \frac{y}{x} \right) h(y) f(x-y) \text{ για } x = 1, 2, \dots \quad (5.1.1)$$

5.2 Λέπτυνση (Thinning)

Στην ενότητα αυτή γίνεται ανάλυση της έννοιας της λέπτυνσης της τ.μ. S . Η λέπτυνση μπορεί να έχει εφαρμογή σε πολλούς τομείς της ασφάλισης και της αντασφάλισης καθώς αν το συμβόλαιο έχει κάποια ίδια κράτηση (retention) μπορούμε να υπολογίσουμε τη συνάρτηση πιθανότητας της S για τον πρωτασφαλιστή ή τον αντασφαλιστή αντίστοιχα.

Έστω ότι με την τ.μ. N συμβολίζουμε τον αριθμό των απαιτήσεων και με την τ.μ. X_i το ύψος της i -οστής παρατήρησης από αυτές τις απαιτήσεις. Επίσης υποθέτουμε ότι τα X_i για $i = 1, 2, 3, \dots$ είναι αμοιβαία ανεξάρτητα και ισόνομα ακολουθώντας μία κατανομή με σ.π.π. $f(x)$ αλλά και ανεξάρτητα από την τ.μ. N η οποία ακολουθεί μία κατανομή με σ.π. p_n . Τότε για $n \geq 1$, η τ.μ. $S = X_1 + X_2 + \dots + X_N$ θα ακολουθεί μία κατανομή με σ.π.π $g(x)$.

Ας υποθέσουμε τώρα ότι μας ενδιαφέρει ο αριθμός των απαιτήσεων που ικανοποιούν ένα συγκεκριμένο κριτήριο. Στην ασφάλιση για παράδειγμα θα μπορούσε να είναι ο αριθμός των παρατηρήσεων που υπερβαίνουν κάποιο όριο. Στη συνέχεια θα ορίσουμε την τ.μ. Y_i ως μία δείκτρια τ.μ. η οποία θα παίρνει την τιμή 1 όταν η παρατήρηση ικανοποιεί το κριτήριο που ορίσαμε και την τιμή 0 διαφορετικά.

Σύμφωνα λοιπόν με τα παραπάνω η $f_y(x)$ είναι μία Bernoulli κατανομή με παράμετρο p για $0 < p < 1$. Αφού η $f \sim \text{Bernoulli}(p)$ τότε για $n = 1, 2, \dots$ η n -οστή συνέλιξη της $f_y(x)$, δηλαδή η $f_y(x)^{n*}$, θα ακολουθεί Διωνυμική κατανομή με παραμέτρους n, p .

Αν για την τ.μ. S ισχύουν όσα αναφέραμε παραπάνω και η τ.μ. N ανήκει στην κλάση $R(a, b, l)$, τότε η συνάρτηση πιθανότητας της δίνεται από την σχέση,

$$g(x) = \frac{1}{1 - a(1 - p)} \left(\sum_{n=x}^l \left(p_n - \left(a + \frac{b}{n} \right) p_{n-1} \right) \binom{n}{x} p^x (1 - p)^{n-x} + \left(a + \frac{b}{x} \right) p g(x - 1) \right)$$

η οποία είναι ισοδύναμη με την

$$g(x) = \frac{1}{1 - a(1 - p)} \sum_{n=x}^l \left(p_n - \left(a + \frac{b}{n} \right) p_{n-1} \right) \binom{n}{x} p^x (1 - p)^{n-x} + \left(a_p + \frac{b_p}{x} \right) g(x - 1)$$

όπου

$$a_p = \frac{ap}{1 - a + ap}$$

$$b_p = \frac{bp}{1 - a + ap}$$

Στην περίπτωση όπου ισχύει ότι $x > l$, τότε ο πρώτος όρος της $f(x)$ εξαλείφεται και καταλήγουμε στη σχέση

$$g(x) = \left(a_p + \frac{b_p}{x} \right) g(x - 1) \text{ για } x = l + 1, l + 2, \dots$$

Αυτό που αξίζει να σημειωθεί είναι ότι τα σημεία (a_p, b_p) είναι μία γραμμή μεταξύ των σημείων $(0,0)$ και (a, b) .

Για να γίνουν καλύτερα κατανοητά όσα αναφέρθηκαν παραπάνω στη συνέχεια ακολουθεί μία εφαρμογή.

Έστω ότι σε μία ασφάλιση εμφανίζονται σε ένα έτος N ζημιές με ύψος X_i , όπου όλες οι ζημιές θεωρούνται ανεξάρτητες. Η κατανομή του πλήθους των ζημιών ακολουθεί *Poisson* με παράμετρο λ ενώ το ύψος της ζημιάς έχει κατανομή

x	100	200	500	1000
$P(X = x)$	0,4	0,3	0,2	0,1

Επίσης έστω ότι ισχύει απαλλαγή 100ν.μ., τις οποίες σε περίπτωση επέλευσης της ζημιάς τις πληρώνει ο ασφαλισμένος και η ασφαλιστική εταιρεία πληρώνει από 100ν.μ. και πάνω.

Στην περίπτωση όπου δεν είχαμε την ίδια κράτηση η παράμετρος a θα πάρει την τιμή 0 ενώ η παράμετρος b θα πάρει την τιμή λ . Στη συνέχεια αφού εφαρμοσθεί η ίδια κράτηση οι τιμές των a_p και b_p χρησιμοποιώντας τις σχέσεις που αναφέρθηκαν παραπάνω θα πάρουν τις τιμές 0 και $0,4\lambda$.

6 Συμπεράσματα

Αρχικά, μελετήθηκε ο αναδρομικός τύπος του Panjer για την κλάση κατανομών $(a,b,0)$. Στη βιβλιογραφία, που αναφέρεται στην παρούσα εργασία, ο υπολογισμός των μέτρων κινδύνων γίνεται με τη χρήση πιθανογεννητριών. Εναλλακτικά παρουσιάζεται αναλυτικός υπολογισμός των μέτρων κινδύνου με την χρήση ροπογεννητριών. Έπειτα, πραγματοποιήθηκε ανάλυση του μη-αναδρομικού τύπου του Fackler για την κλάση κατανομών $(a,b,0)$ και εκτιμήθηκαν οι παράμετροι του με τις οποίες καταλήξαμε στις ίδιες πιθανότητες.

Στη συνέχεια, για τη κλάση κατανομών $(a,b,1)$ αποδείχθηκαν με την χρήση ροπογεννητριών κάποια σημαντικά μέτρα κινδύνου. Από την ανάλυση εξάχθηκαν χρήσιμες σχέσεις οι οποίες συνδέουν τα μέτρα κινδύνου της αρχικής κατανομής, της τροποποιημένης κατανομής της και της περικομμένης κατανομής της.

Στο τέταρτο κεφάλαιο έχοντας κάνει κάποιες εφαρμογές προκύπτουν κάποια χρήσιμα συμπεράσματα όσο αφορά την διακριτοποίηση των συνεχών τ.μ. και κατ' επέκταση του υπολογισμού της σ.κ. της τ.μ. S . Αρχικά παρατηρούμε ότι σημαντικό ρόλο στην διακριτοποίηση έχει το βήμα που επιλέγουμε καθώς όσο πιο μικρό είναι τόσο μικρότερες διαφορές προκύπτουν μεταξύ της διακριτοποιημένης κατανομής με την πραγματική. Ειδικότερα όταν το βήμα διακριτοποίησης δεκαπλασιάζεται, περίπου αντίστοιχη συμπεριφορά παρουσιάζει και το σφάλμα.

Επίσης αυτό που παρατηρούμε είναι ότι η βέλτιστη μέθοδος διακριτοποίησης είναι η Upper καθώς αλλάζοντας τις παραμέτρους της κατανομής που θέλουμε να διακριτοποιήσουμε έχει πάντοτε τις μικρότερες αποκλίσεις. Τέλος ένα σημαντικό αποτέλεσμα που προέκυψε είναι ότι σε κάθε περίπτωση η πραγματική τιμή της σ.κ. της τ.μ. S είναι μεγαλύτερη από την προκύπτουσα κατόπιν της διακριτοποίησης.

Παράρτημα Ι

Εφαρμογή 2.2.1

```
par(mfrow=c(2,2))
#Poisson-Panjer
a<-0;b<-3
Panjer.pfpois<-function(k){if(k==0){return(exp(-b))}
return((a+b/k)*Panjer.pfpois(k-1))}
y<-sapply(0:20,Panjer.pfpois)
plot(0:20,y,type="h",lwd=2,xlab="Αριθμός Ζημιών",
ylab="Πιθανότητα",main="Poisson(3)")

#Poisson-Fackler
a<-10000000000;l<-3
Fackler.pfpois<-function(k){
h<-((1+l/a)^(-a))*((1^k)/factorial(k))
for(i in 0:k){y<-((a+i)/(a+1))}
return(h*y)}
y<-sapply(0:20,Fackler.pfpois)
points(0:20,y,col="red")

#Binomial-Panjer
p<-0.6;q<-0.4;m<-10
a<-(-p/q)
b<-(m+1)*p/q
Panjer.pfbinom<-function(k){if(k==0) return(q^m)
return((a+b/k)*Panjer.pfbinom(k-1))}
y<-sapply(0:20,Panjer.pfbinom)
plot(0:20,y,type="h",lwd=2,xlab="Αριθμός Ζημιών",
ylab="Πιθανότητα",main="Διωνυμική(10,0.6)")

#Binomial-Fackler
m<-10;p<-0.6;a<-(-m);l<-m*p;k<-4;i<-0;y<-1
Fackler.pfbinom<-function(k){
h<-((1+l/a)^(-a))*((1^k)/factorial(k))
for(i in 0:(k-1)){y<-((a+i)*y/(a+1))}
return(h*y)}
y<-sapply(0:20,Fackler.pfbinom)
points(0:20,y,col="red")

#NBinom Panjer
a<-q<-0.5
r<-10
b<-(r-1)*q
Panjer.pfnbinom<-function(k){if(k==0) return((1-a)^r)
return((a+b/k)*Panjer.pfnbinom(k-1))}
y<-sapply(0:20,Panjer.pfnbinom)
plot(0:20,y,type="h",lwd=2,xlab="Αριθμός Ζημιών",
ylab="Πιθανότητα",main="Αρνητική Διωνυμική(10,0.5)")

#NBinom Fackler
a<-10;p<-0.5;l<-a*(1-p)/p;k<-10;y<-1;i<-0
Fackler.pfnbinom<-function(k){
h<-((1+l/a)^(-a))*((1^k)/factorial(k))
for(i in 0:(k-1)){y<-((a+i)*y/(a+1))}
return(h*y)}
y<-sapply(0:20,Fackler.pfnbinom)
points(0:20,y,col="red")
```

Παράδειγμα 4.2.1

```
ft<-c(0,dgeom(seq(0,by=1,length=100),prob=0.6))
ft
Gt<-aggregateDist("recursive", model.freq = "poisson",
model.sev = ft, lambda = 2)
g0<-Gt(0);g0
g1<-Gt(1)-Gt(0);g1
g2<-Gt(2)-Gt(1);g2
g3<-Gt(3)-Gt(2);g3

k<-0:20
gt<-c(Gt(0),Gt(k+1)-Gt(k))
plot(gt,type="h",lwd=1,xlab="Severity",ylab="Propability",
main="Poisson(2)-Geometric(0.6)")
points(1:length(gt),gt,pch=16,col="red")
```

Εφαρμογή 4.2.1

```
##NB(10,0.4) - Exp(3)
#Μη αναδρομικός τύπος
p<-0.4;q<-1-p;n<-10;l<-3;x<-0:10;s1<-0;s2<-0
for(i in 0:(n-1)){for(j in (i+1):n){s2<-choose(n,j)*(q^j)*(p^(n-j))+s2}
s1<-s1+s2*((p*l*x)^i)/factorial(i);s2<-0}
Gt1<-1-(exp(-p*l*x))*s1
round(Gt1,digits=10)

#Με διακριτοποίηση (upper)
p<-0.4;q<-1-p;n<-10;l<-3
a<-0; b<-100; h<-0.1
x<-seq(a,b,h)
fe<-c(0,discretize(pexp(x,rate=l,lower.tail=TRUE,log.p=FALSE),
from=a,to=b,step=h,method="upper"))
Gt2<-aggregateDist("recursive", model.freq = "negative binomial",
model.sev = fe,size=n,prob=p,x.scale=h,maxit=1000000)
round(Gt2(0:10),digits=10)

#Με διακριτοποίηση (round)
p<-0.4;q<-1-p;n<-10;l<-3
a<-0; b<-100; h<-0.1
x<-seq(a,b,h)
fe<-c(0,discretize(pexp(x,rate=l,lower.tail=TRUE,log.p=FALSE),
from=a,to=b,step=h,method="round"))
Gt3<-aggregateDist("recursive", model.freq = "negative binomial",
model.sev = fe,size=n,prob=p,x.scale=h,maxit=1000000)
Gt3(0:10)

#Με διακριτοποίηση (lower)
p<-0.4;q<-1-p;n<-10;l<-3
a<-0; b<-100; h<-0.1
x<-seq(a,b,h)
fe<-c(0,discretize(pexp(x,rate=l,lower.tail=TRUE,log.p=FALSE),
from=a,to=b,step=h,method="lower"))
Gt4<-aggregateDist("recursive", model.freq = "negative binomial",
model.sev = fe,size=n,prob=p,x.scale=h,maxit=1000000)
Gt4(0:10)

#Με διακριτοποίηση (unbiased)
p<-0.4;q<-1-p;n<-10;l<-3
a<-0; b<-100; h<-0.1
x<-seq(a,b,h)
fe<-c(0,discretize(pexp(x,rate=l,lower.tail=TRUE,log.p=FALSE)
,from=a,to=b,step=h,method="unbiased",lev=levexp(x,rate=l)))
Gt5<-aggregateDist("recursive", model.freq = "negative binomial",
model.sev = fe,size=n,prob=p,x.scale=h,maxit=1000000)
```


Βιβλιογραφία

Ελληνική

- [1] Αντζουλάκος Δ. (2015) Σημειώσεις στο μάθημα Γλώσσα Προγραμματισμού R, Πανεπιστήμιο Πειραιώς, Π.Μ.Σ. «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου»
- [2] Κούτρας Μ.Β. (2004) Εισαγωγή στις Πιθανότητες Θεωρία και Εφαρμογές, μέρος Ι, Β' Έκδοση, Εκδόσεις Αθ. Σταμούλης.
- [3] Κούτρας Μ.Β. (2005) Εισαγωγή στις Πιθανότητες Θεωρία και Εφαρμογές, μέρος ΙΙ, Β' Έκδοση, Εκδόσεις Αθ. Σταμούλης.
- [4] Χατζηκωνσταντινίδης Ε. (2015) Σημειώσεις στο μάθημα Θεωρία Κινδύνου Ι, Πανεπιστήμιο Πειραιώς, Π.Μ.Σ. «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου»

Ξένη

- [5] Dickson D. (2005) Insurance Risk and Ruin, Cambridge United Press
- [6] Klugman S., Panjer H., Wilmot G., (2004) Loss Models From Data to Decisions, Wiley, Second Edition
- [7] Klaus Th. Hess, Anett Liewald and Klaus D. Schmidt (2002), An Extension of Panjer's Recursion. ASTIN Bulletin, 32, 283-297.
- [8] Sundt B., Vernic R., (2009) Recursions for Convolutions and Compound Distributions with Insurance Applications. Springer
- [9] Willmot G.E. (1988) Sundt and Jewell's family of discrete distributions, ASTIN Bulletin, 18, 17-29
- [10] Fackler M., (2009) Panjer class United - One Formula for the Poisson, Binomial and Negative Binomial distribution. Proceedings of 39th ASTIN Colloquium, Helsinki, June 1-4, 2009.
- [11] Frostig E., Pitts S.M., Politis K. (2012) The time to ruin and the number of claims until ruin for phase-type claims. Insurance: Mathematics and Economics, 51, 19-25
- [12] Gerhold S., Schmock U., Warnung R. (2010) A generalization of Panjer's recursion and numerically stable risk aggregation. Finance and Stochastics, vol. 14, issue 1, 81-121