ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

Ακριβείς και προσεγγιστικοί υπολογισμοί για ποσότητες με ενδιαφέρον στη θεωρία κινδύνου με χρήση του πακέτου actuar

Δήμητρα Ε. Ντεργιαντέ

Διπλωματική Εργασία

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

Πειραιάς 2018

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή και 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)
- Πιτσέλης Γεώργιος
- Αντζουλάκος Δημήτριος

UNIVERSITY OF PIRAEUS

DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

POSTGRADUATE PROGRAM IN ACTUARIAL SCIENCE
AND RISK MANAGEMENT

Exact and approximate calculations for quantities of interest in risk theory using the actuar package

By
Dimitra E. Ntergiante

MSc Dissertation

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

Piraeus, Greece 2018

Στους γονείς μου Βαγγέλη και Σταυρούλα

στις αδερφές μου Βάσω και Μαρία

ΕΥΧΑΡΙΣΤΙΕΣ

Σε όλη την διάρκεια εκπόνησης της διπλωματικής εργασίας, υπήρξαν πολλά αγαπημένα πρόσωπα, που βοήθησαν και στήριξαν την προσπάθεια αυτή και στάθηκαν δίπλα μου. Ένα θερμό ευχαριστώ σε όλους τους συναδέλφους, τους φίλους και την οικογένειά μου που με την υπομονή τους με βοήθησαν να ανταπεξέλθω σε όλες τις δυσκολίες.

Επίσης, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου Κωνσταντίνο Πολίτη, Αναπληρωτή Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς για τις πολύτιμες γνώσεις και την καθοδήγησή του καθ' όλη την διάρκεια της διπλωματικής εργασίας μου, καθώς και τα άλλα δύο μέλη της τριμελούς εξεταστικής επιτροπής, τους κυρίους Δημήτριο Αντζουλάκο, Αναπληρωτή Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς, και Γεώργιο Πιτσέλη, Επίκουρο Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς, για την επίβλεψή τους.

ΠΕΡΙΛΗΨΗ

Αντικείμενο της παρούσας διπλωματικής είναι η μελέτη της κατανομής των συνολικών αποζημιώσεων ενός χαρτοφυλακίου κινδύνων. Η εν λόγω μελέτη αφορά το συλλογικό μοντέλο της θεωρίας κινδύνου και η ανάλυσή του θα γίνει με ακριβείς υπολογισμούς, μέσω γεννητριών και συνελίξεων συναρτήσεων, όσο και προσεγγιστικά μέσω προσομοίωσης με χρήση του πακέτου actuar της γλώσσας προγραμματισμού R.

Η μελέτη μας θα ξεκινήσει παρουσιάζοντας κάποιες βασικές έννοιες και θεωρήματα στον κλάδο των πιθανοτήτων και της στατιστικής όσο αφορά τις γεννήτριες συναρτήσεις (πιθανογεννήτριες, ροπογεννήτριες και μετασχηματισμό Laplace) και κάποιους βασικούς τύπους υπολογισμού συνελίξεων καθώς και μερικά παραδείγματα για καλύτερη κατανόηση. Στην συνέχεια, κυρίαρχο αντικείμενο θα αποτελέσει το συλλογικό μοντέλο, καθώς και οι κλάσεις κατανομών που μπορεί να το αποτελούν. Στόχος μας είναι η αποδοχή ότι, η ύπαρξη ενός αναλογιστικού πακέτου, όπως είναι το actuar μπορεί να απλοποιήσει τις παραδοσιακές μεθόδους υπολογισμού του προβλήματος της θεωρίας κινδύνου, εκείνο της εύρεσης της κατανομής των συνολικών αποζημιώσεων.

Όλα τα ζητήματα πάνω στα οποία διαπραγματεύεται η παρούσα εργασία, μελετώνται και ερμηνεύονται τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο, μέσω θεωρημάτων και αριθμητικών παραδειγμάτων. Μέλημά μας είναι η αναλυτική παρουσίαση διαφόρων μεθόδων υπολογισμού σύνθετων κατανομών που μπορούν να αποτελούν το συλλογικό μας μοντέλο και η αποδοχή ότι το actuar μπορεί να διευκολύνει την διαδικασία υπολογισμού τους, καθώς μέσα από μερικές εντολές μπορούμε να εξάγουμε το επιθυμητό αποτέλεσμα.

Περιεχόμενα

Κατάλογος Συντομογραφιών

τ.μ.	τυχαία μεταβλητή
βλ.	βλέπε
σ.π.π	συνάρτηση πυκνότητας πιθανότητας
σ.π.	συνάρτηση πιθανότητας
σ.κ.	συνάρτηση κατανομής
σ.δ.ο.	συνάρτηση δεξιάς ουράς
N(· ,·)	Κανονική Κατανομή
G(· , ·)	Κατανομή Γάμμα
G(·)	Γεωμετρική Κατανομή
P(·)	Κατανομή Poisson
Exp(·)	Εκθετική Κατανομή
B(·,·)	Διωνυμική Κατανομή
NB(·,·)	Αρνητική Διωνυμική Κατανομή
U(·,·)	Ομοιόμορφη Κατανομή
Erl(·,·)	Κατανομή Erlang

Κατάλογος Συμβολισμών

- $x \geq 0$: ύψος ατομικής ζημιάς
- $S \geq 0$: ύψος συνολικής ζημιάς
- $N\epsilon\{0,1,2,...\}$: πλήθος ζημιών στον μοντέλο συλλογικού κινδύνου
- $F(x) = Pr(X \leq x)$: σ.κ. της X
- $\bar{F}(x) = 1 - F(x) = 1 - Pr(X \leq x) = Pr(X > x)$: συνάρτηση δεξιάς ουράς
- $f(x) = F^{'}(x)$: σ.π.π της X, αν X συνεχής τ.μ. και $f(x) = F(x) - F(x-1)$: σ.π της X, αν X διακριτή τ.μ.
- $G(x) = Pr(S \leq x)$: σ.κ της S
- $\bar{G}(x) = 1 - G(x) = Pr(S > x)$: συνάρτηση δεξιάς ουράς της S
- $g(x) = G^{'}(x)$: σ.π.π. της S, αν S συνεχής τ.μ. και $g(x) = G(x) - G(x-1)$: σ.π. της S, αν S διακριτή τ.μ.
- $p_n = Pr(N = n), n = 0, 1, ...$: σ.π. της N
- $M_x(t) = E(e^{tx})$: ροπογεννήτρια της τ.μ. X
- $P_x(u) = E(u^x)$: πιθανογεννήτρια της τ.μ. X
- $\hat{f}(s) = E(e^{-sx})$: μετασχηματισμός Laplace της σ.π.π. της f
- $M_N(t) = E(e^{tN})$: ροπογεννήτρια της τ.μ. N
- $P_N(u) = E(u^N)$: πιθανογεννήτρια της τ.μ. N
- $M_S(t) = E(e^{tS})$: ροπογεννήτρια της τ.μ. S
- $P_S(u) = E(u^S)$: πιθανογεννήτρια της τ.μ. S
- $\hat{g}(s) = E(e^{-sS})$: μετασχηματισμός Laplace της σ.π.π. της g

Κεφάλαιο 1

ΕΙΣΑΓΩΓΗ

Ο διεθνής όρος για τον αναλογιστή (actuary) προέρχεται από το λατινικό actuarius, ή αλλιώς τηρητής αρχείων. Παρά το γεγονός ότι η λέξη αυτή μπήκε στην ελληνική γλώσσα τα χρόνια του Βυζαντίου, η σύνδεσή του με τον όρο "άρχων ακτουάριος" την καθιστά κτήμα της εκκλησιαστικής ορολογίας. Ο όρος "αναλογιστής" που επικράτησε στην ελληνική πιθανόν να πρόκειται για νεολογισμό. Μη γνωρίζοντας ποιος πρότεινε τον ελληνικό όρο ή την επικράτηση της παρούσας λέξης, ακατανόητη παραμένει και η σημασία της ως προς το δεύτερο συνθετικό της. Λογιστική είναι η τέλεση πράξεων, καταγραφή και αποτύπωση οικονομικών γεγονότων. Η λογιστική ως συγγενής κλάδος της μαθηματικής επιστήμης καθιστά τον λογιστή ως τελεστή πράξεων και υπολογισμών[6].

Πολλοί ορισμοί έχουν προταθεί για την αναλογιστική επιστήμη, κανένας δεν θεωρήθηκε απόλυτα αποδεκτός. Ένας συνήθης είναι ότι «αναλογιστής είναι ένα άτομο με μαθηματική και οικονομική παιδεία, ικανό να συνδυάσει τις γνώσεις του και να εφαρμόσει κατάλληλα πρότυπα, που έχουν αναπτυχθεί κατά καιρούς στον τομέα του αναλογισμού, με σκοπό την βελτιστοποίηση μακροπρόθεσμων οικονομικών επιπτώσεων που παρουσιάζουν διάφορα αβέβαια γεγονότα». Η ημιμάθεια ως προς το αντικείμενο της αναλογιστικής επιστήμης θεωρείται πιθανό να οδήγησε στην σύνδεση της με εκείνο της λογιστικής (βλ.https://en.wikipedia.org/wiki/Actuary).

Παρά το γεγονός ότι η αναλογιστική επιστήμη είδε το φως πριν από 350 χρόνια, ο θεσμός της ασφάλισης έκανε την εμφάνισή του περίπου τέσσερις χιλιετίες νωρίτερα. Βασικές ασφαλιστικές έννοιες εμφανίζονται σε εδάφια του Κώδικα του Χαμουραμπί

[1], νομοθεσία του 23ου αιώνα π.Χ. Τον 6ο αιώνα π.χ. στην Αθήνα θεσπίζεται νόμος για την λειτουργία ερανικών εταιριών για ασφάλιση δούλων, ως περιουσιακό στοιχείο. Στην Κύπρο τον 5ο αιώνα κάνει την εμφάνισή της η ιατρική κάλυψη με αφορμή τον επικείμενο πόλεμο. Αξιοσημείωτο θεωρείται επίσης το παλαιότερο ιστορικό έγγραφο θαλάσσιας ασφάλισης, που φέρει έτος έκδοσης 1347 και αναφέρεται στο πλοίο Santa Clara [18], καθώς και η δημιουργία ασφάλισης πυρός το 1680 στην Μεγάλη Βρετανία, ως επακόλουθο της Μεγάλης Πυρκαγιάς του 1666 που έπληξε την πόλη (βλ. https://www.museumoflondon.org.uk/discover/how-great-fire-london-created-insurance).

Επιπρόσθετα, η πρώτη εμφάνιση ασφάλισης προσώπου φαίνεται να καταγράφεται σε έναν πίνακα του 220 π.Χ. του Domitius Ulpianus[2], το πρώτο σωζόμενο ασφαλιστήριο συμβόλαιο όμως φαίνεται να εκδόθηκε το 1583 και αναφερόταν σε κάποιον William Gybbons. Κατά τον 17ο και 18ο αιώνα, οι ασφαλίσεις ζωής φαίνεται να κερδίζουν έδαφος και το 1787 ιδρύεται η πρώτη γαλλική εταιρία, η Compagnie Royale d'Assurance. Το 1789 καταγράφεται ο πρώτος αμερικανικός πίνακας θνησιμότητας. Στις αρχές του 19ου αιώνα κυκλοφορεί το πρώτο κείμενο για μαθηματικά των ασφαλίσεων ζωής και το 1843, το βιβλίο του David Jones αποτελεί ορόσημο στον συμβολισμό που χρησιμοποιείται έως σήμερα από τους αναλογιστές ανά τον κόσμο, συμβολισμοί οι οποίοι έγιναν αποδεκτοί το 1898 από το Αναλογιστικό Συνέδριο. Μετά το πέρας του 19ου αιώνα δεν υπάρχουν εξελίξεις στις βασικές αρχές των ασφαλίσεων και του αναλογισμού[6].

Παρά το γεγονός ότι ο θεσμός των γενικών ασφαλίσεων είναι παλαιότερος από εκείνον των ασφαλίσεων ζωής, μόλις το 1747 επιτεύχθηκε η εφαρμογή μαθηματικών μεθόδων σε προβλήματα θαλάσσιας ασφάλισης από τον Corbyn Morris. Η εργασία του χαρακτηρίστηκε πρωτοποριακή, παρά τις μαθηματικές αδυναμίες της, καθώς μελετά την "πιθανότητα του κινδύνου" και τις επιπτώσεις της. Ακολούθησε ο Carl Bremiker με την "κατανομή των συνολικών αποζημιώσεων", καθώς και το "αναγκαίο περιθώριο". Η Θεωρία των Κινδύνων αποτελεί έναν ξεχωριστό κλάδο του αναλογισμού, τα θεμέλια της οποίας τέθηκαν το 1909 με την δημοσίευση του Filip Lundberg με τίτλο "On the theory of

[1]Ο Κώδικας του Χαμουραμπί ήταν εκτεθειμένος δημοσίως ώστε να μπορούν να τον βλέπουν όλοι και να μην υπάρχει άγνοια. Συνολικά υπήρχαν 282 νόμοι, γνωστοί για την ακρίβεια, αλλά και την σκληρότητά τους. Σήμερα βρίσκεται στο Μουσείο του Λούβρου στο Παρίσι.

[2]Το αρχαιότερο γραπτό μνημείο είναι ο πίνακας του Gnaeus Domitius Annius Ulpianus(c. 170 – 223) για την αποτίμηση κληροδοτημάτων με την μορφή ισοβίων συντάξεων.

Reinsurance", η οποία εισήγαγε ένα συνεχές στοχαστικό πρότυπο για τις ασφαλιστικές εταιρίες [6].

Ο ρόλος της θεωρίας κινδύνου στην ανάπτυξη βασικών θεωριών στις στοχαστικές ανελίξεις, των οποίων η χρήση είναι συχνή, θεωρείται κομβικός. Σπουδαίο ρόλο στην ανάπτυξη των θεωριών αυτών κατέχει ο στατιστικός Harald Cramer. Κυρίαρχο θέμα της κλασικής θεωρίας κινδύνου είναι ο προσδιορισμός της κατανομής των συνολικών αποζημιώσεων ενός χαρτοφυλακίου κινδύνων. Το πρόβλημα αυτό πηγάζει από την αδυναμία εύρεσης μαθηματικών τύπων επίλυσης του προβλήματος, σε βαθμό που η προσομοίωση καθίσταται αναγκαία.

Στα πλαίσια της παρούσας διπλωματικής εργασίας το θέμα που θα αναλυθεί αφορά την επίλυση του κυρίαρχου προβλήματος την θεωρίας κινδύνου, αυτό της εύρεσης της κατανομής των συνολικών αποζημιώσεων, τόσο με την χρήση μαθηματικών σχέσεων όσο και με προσομοίωση με την βοήθεια του πακέτου actuar της γλώσσας προγραμματισμού R. Δοθέντων γνωστών κατανομών για το ύψος των αποζημιώσεων και το πλήθος ζημιογόνων γεγονότων που λαμβάνουν χώρα σε ένα χρονικό διάστημα (0,t), θα εξεταστεί η κατανομή των συλλογικών αποζημιώσεων σε διάφορες περιπτώσεις που αυτό μπορεί να παρουσιαστεί.

Για τον λόγο αυτό, στο Κεφάλαιο 2 θα γίνει μια πρώτη αναφορά στις βασικές μεθόδους των συνελίξεων και των γεννητριών συναρτήσεων. Στα πλαίσια της θεωρίας κινδύνου και στις μεθόδους επίλυσης των αριθμητικών προβλημάτων που παρουσιάζονται, οι συνελίξεις και οι γεννήτριες διαδραματίζουν κυρίαρχο ρόλο. Όσο αφορά τις γεννήτριες, οι οποίες αποτελούνται από τις ροπογεννήτριες, πιθανογεννήτριες και μετασχηματισμό Laplace, στο εν λόγω κεφάλαιο θα γίνει μια πρώτη αναφορά στον τύπο της εκάστοτε γεννήτριας καθώς και στις ιδιότητές της, καθώς στην συνέχεια της διπλωματικής θα αναλυθούν εκτενέστερα οι διάφοροι τρόποι εφαρμογής τους. Αντίστοιχα, οι συνελίξεις συναρτήσεων δεν θα αναλυθούν περαιτέρω στα κεφάλαια που θα ακολουθήσουν, καθώς στην πράξη δεν θεωρούνται εύκολες ως προς τον τρόπο υπολογισμού τους, σε περιπτώσεις που έχουμε μεγάλο αριθμό δεδομένων. Παρ' όλα αυτά, καθίσταται απαραίτητη η αναφορά του θεωρητικού μέρους, καθώς και η παρουσίαση κάποιων απλών παραδειγμάτων στην περίπτωση που υπάρχουν διακριτά δεδομένα, καθώς στην συνέχεια της διπλωματικής θα γίνει χρήση συνελίξεων μόνο μέσω προσομοίωσης.

Στην συνέχεια, στο Κεφάλαιο 3 θα παρουσιαστεί εκτενέστερα το κλασικό μοντέλο της θεωρίας κινδύνου. Στην παρούσα διπλωματική, το κύριο πρόβλημα που θα αναλυθεί είναι εκείνο της εύρεσης των συνολικών αποζημιώσεων στο συλλογικό μοντέλο και όχι στο ατομικό, καθώς αποτελεί τμήμα του συλλογικού. Αρχικά θα γίνει μια πρώτη παρουσίαση του τύπου του εν λόγω μοντέλου καθώς και οι διάφορες περιπτώσεις για την μορφή που αυτό έχει, δηλαδή αν είναι πλήρως διακριτό, συνεχές ή μεικτού τύπου με μάζα πιθανότητας στο σημείο μηδέν και συνεχής στο $(0, \infty)$. Στα πλαίσια του κεφαλαίου αυτού θα παρουσιαστούν διάφορα παραδείγματα σχετικά με τις μορφές που μπορεί να έχει η κατανομή των συνολικών αποζημιώσεων και πώς αυτές μπορούν να υπολογισθούν αναλυτικά. Κυρίαρχο ρόλο στο κεφάλαιο αυτό θα αποτελέσουν οι κλάσεις κατανομών $\mathscr{R}(\alpha, b, 0)$ και $\mathscr{R}(\alpha, b, 1)$.

Με βάση τα παραπάνω, η παρούσα διπλωματική χωρίζεται σε δύο τμήματα, το ακριβές υπολογιστικό κομμάτι της κατανομής των συνολικών αποζημιώσεων, και το κομμάτι της προσομοίωσης. Για τον λόγο αυτό, στο κεφάλαιο 4 παρουσιάζεται ένα από τα σημαντικότερα πακέτα με εφαρμογή στην θεωρία κινδύνου και θεωρία χρεοκοπίας στο περιβάλλον της γλώσσα R. Το πακέτο actuar αποτελεί ένα στατιστικό πακέτο με πλήθος λειτουργιών σε διάφορους τομείς του αναλογισμού και της στατιστικής. Στα πλαίσια της παρούσας διπλωματικής θα γίνει αναφορά στο τμήμα εκείνο που αφορά μόνο το κομμάτι της θεωρίας κινδύνου, καθώς και μια μικρή εισαγωγή σε διάφορες εφαρμογές στον κλάδο των ασφαλίσεων. Όπως θα αναφερθεί και στην συνέχεια, το πακέτο αυτό καλύπτει διάφορα πεδία του αναλογισμού και εξελίσσεται διαρκώς.

Τέλος, στο Κεφάλαιο 5 θα παρουσιαστούν διάφορες περιπτώσεις υπολογισμού σύνθετων κατανομών τόσο με ακριβής υπολογισμούς όσο και με την χρήση του πακέτου actuar, καθώς και μια μικρή αναφορά στο κεντρικό οριακό θεώρημα με εφαρμογή στον ασφαλιστικό κλάδο. Στο σημείο αυτό, καθίσταται αναγκαίο να αναφερθεί ότι ο υπολογισμός της κατανομής των συνολικών απαιτήσεων, δηλαδή η κατανομή της τ.μ. S μπορεί να υπολογισθεί με διάφορες μεθόδους στο περιβάλλον της γλώσσα R. Στα πλαίσια της παρούσας διπλωματικής όμως χρησιμοποιήθηκε αποκλειστικά και μόνο το εν λόγω πακέτο, καθώς δημιουργήθηκε με σκοπό την διευκόλυνση των αναλογιστών και όπως θα γίνει εμφανές στην συνέχεια, μπορεί μέσα σε λίγες γραμμές να υπολογίσει πολλές φορές και τα πιο σύνθετα προβλήματα.

Κεφάλαιο 2

ΣΥΝΕΛΙΞΕΙΣ ΚΑΙ ΓΕΝΝΗΤΡΙΕΣ ΣΥΝΑΡΤΗΣΕΙΣ

Στην θεωρία κινδύνου κυρίαρχο ρόλο διαδραματίζουν οι συνελίξεις και οι γεννήτριες συναρτήσεων. Στο παρόν κεφάλαιο θα γίνει μια συνοπτική αναφορά τόσο στην μέθοδο των συνελίξεων όσο και στις μορφές των γεννητριών. Πιο συγκεκριμένα, στην Ενότητα 2.1 θα παρουσιαστούν οι διάφορες μορφές γεννητριών, οι οποίες δεν είναι άλλες από τις ροπογεννήτριες, πιθανογεννήτριες και τον μετασχηματισμό Laplace. Στη συνέχεια, στην Ενότητα 2.2 θα αναφερθεί το θεώρημα της διπλής μέσης τιμής για διακριτές και συνεχείς τυχαίες μεταβλητές. Η Ενότητα 2.3 είναι αφιερωμένη στις συνελίξεις συναρτήσεων και τέλος στην 2.4 γίνεται αναφορά στις γεννήτριες συναρτήσεις ακολουθιών.

Οι τέσσερις εν λόγω ενότητες του παρόντος κεφαλαίου αποτελούν μια εισαγωγή στις μεθόδους επίλυσης προβλημάτων στον κλάδο της θεωρίας κινδύνου. Όπως θα αναφερθεί και στην συνέχεια, για τον ακριβή υπολογισμό της κατανομής των συνολικών απαιτήσεων του μοντέλου μας είναι απαραίτητη η χρήση των συγκεκριμένων εργαλείων.

2.1 Γεννήτριες συναρτήσεις τυχαίων μεταβλητών

Στην ενότητα αυτή θα γίνει αναφορά στις τρεις κυριότερες γεννήτριες συναρτήσεις τυχαίων μεταβλητών που συνήθων χρησιμοποιούνται για την εξέταση διαφόρων ποσοτήτων στην Θεωρία Κινδύνων. Συγκεκριμένα θα γίνει αναφορά στις ροπογεννήτριες συναρτήσεις, τις πιθανογεννήτριες και τον μετασχηματισμό Laplace. Μέσω αυτών των συναρτήσεων μπορούν να υπολογισθούν τα μέτρα θέσης και μεταβλητότητας καθώς και να βρεθούν ακριβείς, αναδρομικοί και ασυμπτωτικοί τύποι για τις κατανομές των τυχαίων μεταβλητών. Ουσιαστικά η επικράτηση της ονομασίας ¨*γεννήτριες συναρτήσεις'* οφείλεται κατά κύριο λόγο στον Laplace, καθώς μέσα από το έργο του *Theorie analytique des probabilites* έθεσε τις βάσεις για μια εκτενέστερη εξέταση των γεννητριών.[15]

2.1.1 Ροπογεννήτριες συναρτήσεις τυχαίων μεταβλητών

ΟΡΙΣΜΟΣ 2.1 *Αν η μέση τιμή $E[e^{tX}]$ της συνάρτησης e^{tx} της τυχαίας μεταβλητής X υπάρχει για $0 \leq t \leq$τ όπου τ$\in (0, +\infty)$, τότε αυτή καλείται ροπογεννήτρια συνάρτηση της τυχαίας μεταβλητής X και συμβολίζεται με $M_X(t)$.*

Δηλαδή ισχύει ότι:

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

□

Αν $F(x) = P(X \leq x)$ είναι η συνάρτηση κατανομής της τυχαίας μεταβλητής X, τότε:

- Αν η τ.μ. X είναι απόλυτα συνεχής, τότε η ροπογεννήτρια συνάρτηση υπολογίζεται ως εξής:
$$M_X(t) = \int_0^{\infty} e^{tx} f_X(x) dx$$

- Αν η τ.μ. X είναι διακριτή, τότε είναι:
$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} f_X(x)$$
όπου $f_X(x) = P(X = x)$ η συνάρτηση πιθανότητας της τ.μ. X.

Πρέπει να τονισθεί ότι η ροπογεννήτρια συνάρτηση δεν υπάρχει πάντοτε. Ορισμένες κατανομές, όπως η κατανομή Pareto, lognormal, loggamma κ.α., δεν έχουν ροπογεννήτρια (έχουν όμως ροπές).

Ιδιότητες ροπογεννητριών συναρτήσεων:

1. Αν $Y = \alpha X + \beta$, τότε ισχύει ότι:

$$M_Y(t) = e^{\beta t} M_X(\alpha t) \tag{2.1}$$

2. Αν οι τ.μ. $X_1, X_2, ..., X_n$ είναι αμοιβαία ανεξάρτητες και $S_n = X_1 + X_2 + ... + X_n$, τότε ισχύει ότι:

$$M_{S_n}(t) = M_{X_1}(t) M_{X_2}(t) ... M_{X_n}(t).$$

   Το αντίστροφο δεν ισχύει πάντοτε.

3. Αν οι τ.μ. $X_1, X_2, ..., X_n$ είναι αμοιβαία ανεξάρτητες και ισόνομες έστω και με μια τ.μ. Χ, και $S_n = X_1 + X_2 + ... + X_n$, τότε:

$$M_{S_n}(t) = (M_X(t))^n .$$

4. Αν για δύο τ.μ. Χ και Υ είναι $M_X(t) = M_Y(t) \; \forall \; t$, τότε ισχύει ότι:

$$X \stackrel{d}{=} Y$$

   όπου το σύμβολο $\stackrel{d}{=}$ σημαίνει ισότητα ως προς την κατανομή (equal in distribution), δηλαδή ότι οι τ.μ. Χ και Υ είναι ισόνομες.

(βλ. Χατζηκωνσταντινίδης, 2015).

2.1.2 Μετασχηματισμοί Laplace

Ένα αρκετά χρήσιμο εργαλείο για συνεχείς και μικτού τύπου τυχαίες μεταβλητές είναι ο μετασχηματισμός Laplace ο οποίος υπάρχει πάντοτε, καθώς ορίζεται γενικότερα για συνεχείς ή και τμηματικά συνεχείς συναρτήσεις και όχι μόνο για συναρτήσεις πυκνότητας πιθανότητας τυχαίων μεταβλητών.

ΟΡΙΣΜΟΣ 2.2 *Έστω $h(x)$ μια συνάρτηση που ορίζεται για $x > 0$. Ο μετασχηματισμός Laplace της $h(x)$ συμβολίζεται με $\hat{s}(x)$ και ορίζεται ως:*

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

$\square$

Για λόγους ευκολίας, στην Θεωρία Κινδύνου θεωρείται ότι το s είναι ένας θετικός πραγματικός αριθμός. Γενικότερα, ο μετασχηματισμός Laplace μπορεί να ορισθεί όταν το s είναι ένας μιγαδικός αριθμός με θετικό πραγματικό μέρος, δηλαδή είναι της μορφής $s = \alpha + bi$ με $\alpha > 0$ και $i = \sqrt{-1}$.

(βλ. Κυπαρισσίδης, 2010).

ΠΑΡΑΔΕΙΓΜΑ 2.1 *Έστω $h(x) = e^{\alpha x}, x > 0$. Τότε,*

$$\hat{h}(s) = \int_0^\infty e^{-sx} e^{\alpha x} dx = \int_0^\infty e^{-(s-\alpha)x} dx,$$

οπότε είναι,

$$\hat{h}(s) = \frac{1}{s - \alpha}, \quad s > \alpha.$$

Για $s < \alpha$ είναι $-(s - \alpha) > 0$, οπότε το παραπάνω ολοκλήρωμα δεν συγκλίνει (απειρίζεται).

$\square$

Από το παραπάνω παράδειγμα προκύπτει εύκολα το συμπέρασμα ότι ο μετασχηματισμός Laplace που αντιστοιχεί στην συνάρτηση $e^{\alpha x}$ είναι ο $1/(s - \alpha)$ και η αντίστροφη συνάρτηση που αντιστοιχεί στον μετασχηματισμό Laplace $1/(s - \alpha)$ είναι η $e^{\alpha x}$.

Οι μετασχηματισμοί Laplace επίσης χρησιμοποιούνται για την επίλυση διαφορικών εξισώσεων, ολοκληρωτικών εξισώσεων και ολοκληρο-διαφορικών εξισώσεων. Αρκετές ποσότητες, που μελετάμε στη Θεωρία Κινδύνων και ιδιαίτερα στα μοντέλα συλλογικού κινδύνου και στη θεωρία χρεοκοπίας, ικανοποιούν τέτοιες εξισώσεις.

Όταν γνωρίζουμε τον μετασχηματισμό Laplace, τότε μπορούμε να βρούμε άμεσα και την ροπογεννήτρια συνάρτηση, και αντίστροφα όταν γνωρίζουμε τη ροπογεννήτρια συνάρτηση μπορούμε να βρούμε και το μετασχηματισμό Laplace ως εξής:

- Αν γνωρίζουμε τον μετασχηματισμό Laplace, τότε εύκολα θέτοντας $s = -t$ παίρνουμε την ροπογεννήτρια συνάρτηση:

$$\hat{f}_X(-t) = E[e^{-(-t)X}] = E[e^{tX}] = M_X(t).$$

- Αντίστοιχα, αν γνωρίζουμε την ροπογεννήτρια συνάρτηση, θέτοντας $t = -s$ προκύπτει ο μετασχηματισμός Laplace:

$$M_X(-s) = E(e^{-sX}) = \hat{f}_X(s).$$

Ιδιότητες μετασχηματισμού Laplace:

1. Αν $Y = \alpha X + \beta$, τότε ισχύει οτι:

$$L_Y(s) = e^{-\beta s} L_X(\alpha s).$$

2. Αν οι τ.μ. $X_1, X_2, ..., x_n$ είναι αμοιβαία ανεξάρτητες και $S_n = X_1 + X_2 + ... + X_n$, τότε ισχύει ότι:

$$L_{S_n}(s) = L_{X_1}(s) L_{X_2}(s) ... L_{X_n}(s).$$

   Το αντίστροφο δεν ισχύει πάντοτε.

3. Αν οι τ.μ. $X_1, X_2, ..., X_n$ είναι αμοιβαία ανεξάρτητες και ισόνομες έστω και με μια τ.μ. Χ, και $S_n = X_1 + X_2 + ... + X_n$, τότε:

$$L_{S_n}(s) = (L_X(s))^n .$$

4. Αν για δύο τ.μ. Χ και ϒ είναι $L_X(s) = L_Y(s) \; \forall \; s$, τότε ισχύει ότι:

$$X \stackrel{d}{=} Y$$

(βλ. Χατζηκωνσταντινίδης, 2015).

2.1.3 Πιθανογεννήτριες συναρτήσεις τυχαίων μεταβλητών

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

ΟΡΙΣΜΟΣ 2.3 *Έστω η μη-αρνητική διακριτή τ.μ. X με συνάρτηση πιθανότητας $f_X(x) = Pr(X = x)$. Τότε, η πιθανογεννήτρια συνάρτηση της τ.μ. X συμβολίζεται με $P_X(u)$ και ορίζεται ως:*

$$P_X(u) = E(u^X) = \sum_{x=0}^{\infty} f_X(x)u^x$$

και η σειρά αυτή συγκλίνει ομοιόμορφα για $|u| \leq 1$.

□

ΠΑΡΑΔΕΙΓΜΑ 2.2 *Αν η τ.μ. $X \sim G_0(p)$, δηλαδή:*

$$f_X(x) = pq^x, \quad x = 0, 1, 2, ..., \quad 0 < p < 1, \ q = 1 - p$$

τότε η πιθανογεννήτρια συνάρτησή της ορίζεται ως,

$$P_X(u) = \sum_{x=0}^{\infty} pq^x u^x = p\sum_{x=0}^{\infty}(qu)^x.$$

Αν $|qu| \geq 1$, τότε το παραπάνω άθροισμα δεν συγκλίνει (απειρίζεται), οπότε η πιθανογεννήτρια συνάρτηση δεν υπάρχει.
Αν $|qu| < 1$, δηλαδή αν $|u| < \frac{1}{q}$, τότε ισχύει ότι:

$$\sum_{x=0}^{\infty}(qu)^x = \frac{1}{1 - qu}$$

οπότε, είναι

$$P_X(u) = \frac{p}{1 - qu}, \quad |u| < \frac{1}{q},$$

που είναι και η ροπογεννήτρια της Γεωμετρικής Κατανομής όταν αυτή ορίζεται στο μηδέν.

(βλ. ΠΑΡΑΡΤΗΜΑ Α)

□

Όταν γνωρίζουμε την πιθανογεννήτρια συνάρτηση τότε μπορούμε να βρούμε άμεσα και τη ροπογεννήτρια συνάρτηση, και αντίστροφα όταν γνωρίζουμε τη ροπογεννήτρια συνάρτηση μπορούμε να βρούμε και την πιθανογεννήτρια συνάρτηση, ως εξής:

- Αν γνωρίζουμε την πιθανογεννήτρια συνάρτηση, τότε θέτοντας $u = e^t$ παίρνουμε την ροπογεννήτρια, ως εξής:

$$P_X(e^t) = E[(e^t)^X] = E(e^{tX}) = M_X(t).$$

- Αντίστοιχα, αν γνωρίζουμε την ροπογεννήτρια, θέτοντας $t = lnu$ παίρνουμε την πιθανογεννήτρια συνάρτηση:

$$M_X(lnu) = E[e^{Xlnu}] = E[e^{lnu^X}] = E(u^X) = P_X(u).$$

Ιδιότητες πιθανογεννητριών συναρτήσεων:

1. Αν $Y = \alpha X + \beta$, τότε ισχύει οτι:

$$P_Y(u) = u^\beta P_X(u^\alpha).$$

2. Αν οι τ.μ. $X_1, X_2, ..., x_n$ είναι αμοιβαία ανεξάρτητες και $S_n = X_1 + X_2 + ... + X_n$, τότε ισχύει ότι:

$$P_{S_n}(u) = P_{X_1}(u)P_{X_2}(u)...P_{X_n}(u).$$

   Το αντίστροφο δεν ισχύει πάντοτε.

3. Αν οι τ.μ. $X_1, X_2, ..., X_n$ είναι αμοιβαία ανεξάρτητες και ισόνομες έστω και με μια τ.μ. Χ, και $S_n = X_1 + X_2 + ... + X_n$, τότε:

$$P_{S_n}(u) = (P_X(u))^n.$$

4. Αν για δύο τ.μ. Χ και Υ είναι $P_X(u) = P_Y(u)\ \forall\ u$, τότε ισχύει ότι:

$$X \stackrel{d}{=} Y.$$

(βλ. Χατζηκωνσταντινίδης, 2015).

2.2 Το θεώρημα της διπλής μέσης τιμής

Έστω ότι Χ είναι μια τυχαία μεταβλητή, για την οποία η μέση τιμή της υπάρχει, δηλαδή ισχύει ότι $E[|X|] < \infty$.

2.2.1 Υπό συνθήκη μέση τιμή

(Α) Έστω ότι η τυχαία μεταβλητή Χ είναι διακριτή

Αν $f_X(x) = P[X = x]$ είναι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Χ, τότε η μέση τιμή της δίνεται από τον τύπο:

$$E(X) = \sum_{x=0}^{\infty} x f_X(x).$$

Έστω ότι $f_{X,Y}(x, y) = P(X = x, Y = y)$ είναι η από κοινού συνάρτηση πιθανότητας των Χ και Υ και $f_{X|Y}(x|y) = P(X = x|Y = y)$ η δεσμευμένη συνάρτηση πιθανότητας της Χ δοθείσης της Υ. Τότε, η υπό συνθήκη μέση τιμή της τυχαίας μεταβλητής Χ δοθείσης της τιμής y της Υ, δίνεται από την σχέση:

$$E[X|Y = y] = \sum_{x} x f_{X|Y}(x|y).$$

ΠΡΟΤΑΣΗ 2.1 *Αν οι Χ και Υ είναι διακριτές τυχαίες μεταβλητές, τότε:*

$$E(X) = E[E(X|Y)].$$

Αν οι X και Υ είναι διακριτές τυχαίες μεταβλητές, τότε για οποιαδήποτε συνάρτηση h για την οποία η $E[h(X)]$ είναι πεπερασμένη, ισχύει ότι:

$$E[h(X)] = E[E(h(X)|Y)].$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

(Β) Έστω ότι η τυχαία μεταβλητή X είναι συνεχής

Έστω $F_X(x) = P(X \leq x)$ και $f_X(x) = F'_X(x)$ η συνάρτηση κατανομής και η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X αντίστοιχα. Κάτω από την υπόθεση ότι η X είναι απόλυτα συνεχής, η μέση τιμή της δίνεται από την σχέση:

$$E(X) = \int x dF_X(x) = \int x f_X(x) dx.$$

Αντίστοιχα, η υπό συνθήκη μέση τιμή της τυχαίας μεταβλητής X δοθείσης της τιμή y της Υ, είναι:

$$E(X|Y = y) = \int x f_{X|Y}(x|y) dx.$$

ΠΡΟΤΑΣΗ 2.2 *Αν οι X και Υ είναι συνεχείς τυχαίες μεταβλητές, τότε για οποιαδήποτε συνάρτηση h για την οποία η $E[h(X)]$ είναι πεπερασμένη, ισχύει ότι:*

$$E[h(X)] = E[h(X)|Y].$$

Επιπλέον, ισχύει ότι:

$$E(X) = E[E(X|Y)].$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

2.2.2 Το θεώρημα της διπλής μέσης τιμής

ΘΕΩΡΗΜΑ 2.1 *Για οποιεσδήποτε τυχαίες μεταβλητές X και Υ, ισχύει ότι:*
(α) $E(X) = E[E(X|Y)]$.
(β) $Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$.

*(**Απόδειξη**: βλ. Dickson, 2005).*

Όσο αφορά την μεταβλητότητα της τυχαίας μεταβλητής X, υπάρχουν δύο πηγές, δηλαδή δύο συνιστώσες που συνεισφέρουν στην ολική μεταβλητότητα της X.

(i) Μεταβλητότητα της X μέσα στις υποομάδες

Η μεταβλητότητα της X μέσα στις υποομάδες είναι οι διακυμάνσεις

$$Var(X|Y=1), Var(X|Y=2), ..., Var(X|Y=n).$$

Επειδή οι τιμές της Var(X|Y=y) μεταβάλλονται καθώς μεταβάλλεται το y, έπεται ότι η Var(X|Y) είναι συνάρτηση της τ.μ. Υ, οπότε η μέση τιμή

$$E[Var(X|Y)] = \sum_{y=1}^{n} Var(X|Y=y)P(Y=y).$$

είναι η μεταβλητότητα μέσα στις υποομάδες, δηλαδή κάθε ακέραια τιμή που παίρνει η τ.μ. Υ.

(i) Μεταβλητότητα μεταξύ των υποομάδων

Παριστάνοντας κάθε υποομάδα από τη μέση τιμή E(X|Y), είναι εφικτό να υπολογιστεί η μεταβλητότητα μεταξύ των υποομάδων υπολογίζοντας την $Var[E(X|Y)]$, που είναι η μεταβλητότητα στους μέσους των υποομάδων ως συνάρτηση της τυχαίας μεταβλητής Υ.

Επομένως, από τις ανωτέρω περιπτώσεις, έπεται ύστερα από πράξεις, ότι:

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)],$$

όπου

$E[Var(X|Y)]$= μέση μεταβλητότητα μέσα στις υποομάδες, και

$Var[E(X|Y)]$= μεταβλητότητα μεταξύ των μέσων των υποομάδων.

2.3 Συνελίξεις Συναρτήσεων

ΟΡΙΣΜΟΣ 2.4 *Έστω οι συναρτήσεις f,g, η συνέλιξη των οποίων ορίζεται ως μια νέα συνάρτηση h, της μορφής h = f * g ως εξής:*

(α) Έστω οι f και g είναι διακριτές συναρτήσεις (ακολουθίες πραγματικών αριθμών), ορισμένες στο $\{0, 1, 2, ...\}$. *Τότε είναι:*

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

(β) Έστω οι f και g είναι συνεχείς συναρτήσεις, ορισμένες στο $[0, \infty)$. *Τότε είναι:*

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

□

Ισχύει ότι:

$$f * g = g * f$$

δηλαδή, η συνέλιξη των συναρτήσεων f και g είναι ίση με την συνέλιξη των συναρτήσεων g και f, οπότε δεν έχει σημασία αν γράφουμε $f * g$ ή $g * f$.

Από τον προηγούμενο ορισμό, προκύπτει το συμπέρασμα ότι:

- Αν οι συναρτήσεις f και g είναι συνεχείς, τότε ισχύει $(f * g)(0) = 0$.

- Ενώ, αν οι συναρτήσεις f και g είναι διακριτές, ισχύει $(f * g)(0) = f(0)g(0)$.

2.3.1 Αθροίσματα ανεξάρτητων διακριτών τυχαίων μεταβλητών

Έστω οι $X_1, X_2, ..., X_n$ είναι ανεξάρτητες (και όχι απαραίτητα ισόνομες) διακριτές τυχαίες μεταβλητές στο $\{0, 1, ...n\}$, για κάθε $i = 1, 2, ..., n$. Τότε προφανώς και η S_n θα κατανέμεται στο σύνολο των μη-αρνητικών ακέραιων αριθμών.

Αρχικά ας θεωρήσουμε την περίπτωση των δύο ανεξάρτητων τυχαίων μεταβλητών X_1 και X_2. Η συνάρτηση κατανομής $F_{X_1+X_2}$ είναι η συνέλιξη της συνάρτησης κατανομής της τυχαίας μεταβλητής X_1 και της συνάρτησης πιθανότητας της X_2, δηλαδή ισχύει ότι:

$$\begin{aligned} F_{X_1+X_2}(x) &= (F_{X_1} * F_{X_2})(x) \\ &= \sum_{y=0}^{x} Pr(X_2 \le x-y)Pr(X_1 = y) \\ &= \sum_{y=0}^{x} F_{X_2}(x-y) f_{X_1}(y). \end{aligned}$$

Επίσης, ισχύει η δέσμευση της τυχαίας μεταβλητής X_2 ως προς την τυχαία μεταβλητή X_1. Συνεπώς, αντίστοιχα:

$$F_{X_1+X_2} = F_{X_2} * f_{X_1} = f_{X_1} * F_{X_2}.$$

Η συνάρτηση πιθανότητας $f_{X_1} + f_{X_2}$ της τυχαίας μεταβλητής $X_1 + X_2$ είναι η συνέλιξη της συνάρτησης πιθανότητας της X_1 και της συνάρτησης πιθανότητας της X_2, δηλαδή:

$$\begin{aligned} f_{X_1+X_2}(x) &= f_{X_1} * f_{X_2}(x) \\ &= f_{X_2} * f_{X_1}(x) \\ &= \sum_{y=0}^{x} f_{X_1}(x-y) f_{X_2}(y) \\ &= \sum_{y=0}^{x} f_{X_2}(x-y) f_{X_1}(y). \end{aligned}$$

ΠΑΡΑΔΕΙΓΜΑ 2.3 *Έστω οι τ.μ. X_1 και X_2, ανεξάρτητες μεταξύ τους, με συναρτήσεις πιθανότητας που δίνονται στον παρακάτω πίνακα:*

x	$f_1(x)$	$f_2(x)$
0	1/5	
1	2/5	1/4
2	1/5	1/4
3	1/5	1/4
4		1/4

Επειδή η τ.μ $X_1 \in \{0, 1, 2, 3\}$ και η τ.μ. $X_2 \in \{1, 2, 3, 4\}$, έπεται ότι $S_2 \in \{1, 2, ..., 7\}$. Θέλουμε να υπολογίσουμε την συνάρτηση πιθανότητας της τυχαίας μεταβλητής $S_2(x) = X_1 + X_2$. Ισχύει ότι:

$$\begin{aligned} f_{S_2}(x) = Pr(S_2 = x) &= \sum_{y=0}^{x} f_{X_1}(x-y) f_{X_2}(y) \\ &= f_{X_1}(x) f_{X_2}(0) + \sum_{y=1}^{x} f_{X_1}(x-y) f_{X_2}(y) \\ &= \sum_{y=1}^{x} f_{X_1}(x-y) f_{X_2}(y). \end{aligned}$$

Έτσι έχουμε:

$$\begin{aligned} f_{S_2}(1) &= f_{X_1}(0) f_{X_2}(1) = (1/5)(1/4) = 1/20 \\ f_{S_2}(2) &= f_{X_1}(1) f_{X_2}(1) + f_{X_1}(0) f_{X_2}(2) = (2/5)(1/4) + (1/5)(1/4) = 3/20 \\ f_{S_2}(3) &= f_{X_1}(2) f_{X_2}(1) + f_{X_1}(1) f_{X_2}(2) + f_{X_1}(0) f_{X_2}(3) \\ &= (1/5)(1/4) + (2/5)(1/4) + (1/5)(1/4) = 1/5 \\ f_{S_2}(4) &= f_{X_1}(3) f_{X_2}(1) + f_{X_1}(2) f_{X_2}(2) + f_{X_1}(1) f_{X_2}(3) + f_{X_1}(0) f_{X_2}(4) \\ &= (1/5)(1/4) + (1/5)(1/4) + (2/5)(1/4) = 1/4 \\ f_{S_2}(5) &= f_{X_1}(3) f_{X_2}(3) + f_{X_1}(2) f_{X_2}(3) + f_{X_1}(1) f_{X_2}(4) \\ &= (1/5)(1/4) + (1/5)(1/4) + (2/5)(1/4) = 1/5 \\ f_{S_2}(6) &= f_{X_1}(3) f_{X_2}(3) + f_{X_1}(2) f_{X_2}(4) = (1/5)(1/4) + (1/5)(1/4) = 1/10 \\ f_{S_2}(7) &= f_{X_1}(3) f_{X_2}(4) = (1/5)(1/4) = 1/20. \end{aligned}$$

Άρα, η κατανομή της τυχαίας μεταβλητής S_2 μέσω της συνάρτησης πιθανότητας είναι:

x	1	2	3	4	5	6	7
$f_{S_2}(x)$	1/20	3/20	1/5	1/4	1/5	1/10	1/20

□

2.3.2 Αθροίσματα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών

Έστω τώρα ότι οι ανέξαρτητες τυχαίες μεταβλητές είναι και ισόνομες με την τυχαία μεταβλητή X. Τότε η κατανομή της $S_n = X_1 + X_2 + ... + X_n$ είναι η 'n-οστή συνέλιξη' της F_X και η συνάρτηση κατανομής της συμβολίζεται με F_X^{*n}. Δηλαδή είναι:

$$F_X^{*n}(x) = Pr(X_1 + X_2 + ... + X_n \leq x) = \sum_{y=0}^{x} F_X^{*(n-1)}(x-y)f_X(y).$$

Η αντίστοιχη ουρά της κατανομής δίνεται από την σχέση

$$\bar{F}^{*n}(x) = Pr(X_1 + X_2 + ... + X_n > x) = 1 - F_X^{*n}(x) = \sum_{y=0}^{x} f_X^{*(n-1)}(x-y)f_X(y).$$

Αντίστοιχα, η συνάρτηση πιθανότητας της n-οστής συνέλιξης είναι

$$f_X^{*n}(x) = Pr(X_1 + X_2 + ... + X_n = x) = \bar{F}_X(x) + \sum_{y=0}^{x} \bar{F_X}^{*(n-1)}(x-y)f_X(y).$$

Επίσης, η μηδενική συνέλιξη για διακριτές και συνεχείς τυχαίες μεταβλητές, ορίζεται ως εξής:

$$F_X^{*0}(x) = \begin{cases} 1, & \text{x} \geq 0 \\ 0, & \text{x} > 0 \end{cases},$$

$$\bar{F}_X^{*0}(x) = \begin{cases} 0, & \text{x} \geq 0 \\ 1, & \text{x} > 0 \end{cases},$$

$$f_X^{*0}(x) = \begin{cases} 1, & \text{x=0} \\ 0, & \text{x} \neq 0 \end{cases}.$$

Στην θεωρία κινδύνου κυρίαρχο ρόλο αποτελεί η εύρεση της κατανομής των συνολικών αποζημιώσεων, δηλαδή η κατανομή που ακολουθεί η τ.μ. S. Σύμφωνα με τον τύπο της S που θα δοθεί στην συνέχεια (βλ. Κεφάλαιο 3), για τον υπολογισμό της κατανομής της στην περίπτωση όπου οι τ.μ. N και X είναι διακριτές, μπορεί να χρησιμοποιηθεί ο αναδρομικός τύπος του Panjer (βλ. Θεώρημα 3.4), η βάση του οποίου στηρίζεται στο παρακάτω θεώρημα.

ΘΕΩΡΗΜΑ 2.2 *Έστω ότι οι διακριτές τ.μ.* $X_i \in \{0, 1, 2, ...\}$ *είναι ανεξάρτητες και ισόνομες με μια τ.μ.*X *με συνάρτηση πιθανότητας* $f_X^{*n} = Pr(X = x)$. *Τότε, για* $x = 1, 2, 3, ...$ *και* $n = 1, 2, 3, ...n$, *η η-οστή συνέλιξη της συνάρτησης πιθανότητας της τ.μ. X ικανοποιεί την παρακάτω αναδρομική σχέση:*

$$f_X^{*n} = Pr(X_1 + X_2 + ... + X_n = x) = \frac{1}{f_X(0)} \sum_{i=1}^{x} [\frac{(n+1)i}{x} - 1] f_X(i) f_X^{*n}(x - i).$$

και

$$\begin{aligned} f_X^{*n}(0) &= Pr(X_1 + X_2 + ... + X_n = 0) \\ &= Pr(X_1 = 0, X_2 = 0, ..., X_n = 0) \\ &= Pr(X_1 = 0) Pr(X_2 = 0) ... Pr(X_n = 0) \\ &= [Pr(X = 0)]^n = (f_X(0))^n. \end{aligned}$$

Όπως θα αναφερθεί και στην συνέχεια, η τ.μ. S για τιμές $N \geq 1$ είναι ίση με $X_1 + X_2 + .. + X_N$, επομένως ισχύει ότι

$$f^{*n}(x) = Pr(X_1 + X_2 + ... + X_n = x) = Pr(S = x).$$

2.3.3 Αθροίσματα ανεξάρτητων συνεχών τυχαίων μεταβλητών

Έστω οι $X_1, X_2, ..., X_n$ είναι ανεξάρτητες (και όχι απαραίτητα ισόνομες) συνεχείς μη-αρνητικές τυχαίες μεταβλητές, με $X_i \epsilon$ (0,∞), για κάθε $i = 1, 2, ..., n.$ Τότε προφανώς και η $S_n \epsilon (0, \infty).$

Αρχικά ας θεωρήσουμε την περίπτωση των δύο ανεξάρτητων τυχαίων μεταβλητών X_1 και X_2. Η συνάρτηση κατανομής $F_{X_1+X_2}$ είναι η συνέλιξη της συνάρτησης κατανομής της τυχαίας μεταβλητής X_1 και της συνάρτησης πυκνότητας πιθανότητας της X_2, δηλαδή ισχύει ότι $S_2 = X_1 + X_2$. Συνεπώς, έπεται η ακόλουθη πρόταση:

ΠΡΟΤΑΣΗ 2.3 *Έστω οι ανεξάρτητες συνεχείς τυχαίες μεταβλητές (για τις οποίες ισχύει με πιθανότητα ένα)* $X_1 \geq 0$, $X_2 \geq 0$ *και* $S_2 = X_1 + X_2$. *Τότε, για* $x \geq 0$, *ισχύουν τα εξής:*

(α) *Η συνάρτηση κατανομής της τ.μ.* S_2, *είναι:*

$$\begin{aligned} F_{S_2}(x) &= \int_0^x F_{X_1}(x-y) f_{X_2}(y) dy \\ &= \int_0^x F_{X_2}(x-y) f_{X_1}(y) dy \end{aligned}$$

δηλαδή,

$$F_{S_2} = F_{X_1} * f_{X_2} = F_{X_2} * f_{X_1}.$$

(β) *Η συνάρτηση πυκνότητας πιθανότητας της τ.μ.* S_2, *είναι:*

$$\begin{aligned} f_{S_2} &= \int_0^x f_{X_1}(x-y) f_{X_2}(y) dy \\ &= \int_0^x f_{X_2}(x-y) f_{X_1}(y) dy \end{aligned}$$

δηλαδή,

$$f_{S_2} = f_{X_1} * f_{X_2} = f_{X_2} * f_{X_1}.$$

(γ) *Η συνάρτηση δεξιάς ουράς της τ.μ.* S_2, *είναι:*

$$\begin{aligned} \bar{F}_{S_2}(x) &= \bar{F}_{X_2}(x) + \int_0^x \bar{F}_{X_1}(x-y) f_{X_2}(y) dy \\ &= \bar{F}_{X_1}(x) + \int_0^x \bar{F}_{X_2}(x-y) f_{X_1}(y) dy \end{aligned} \tag{2.2}$$

δηλαδή,

$$\bar{F}_{S_2} = \bar{F}_{X_2} + \bar{F}_{X_1} * f_{X_2} = \bar{F}_{X_2} + \bar{F}_{X_2} * f_{X_1}$$

(βλ. Χατζηκωνσταντινίδης, 2015).

2.4 Γεννήτριες συναρτήσεις ακολουθιών

ΟΡΙΣΜΟΣ 2.5 *[5] Έστω μια ακολουθία αριθμών $\{\alpha_x : x = 0, 1, 2, ...\}$, τότε η συνήθης γεννήτρια συνάρτησή της ορίζεται από το άθροισμα:*

$$A(u) = \sum_{x=0}^{\infty} \alpha_x u^x = \alpha_0 + \alpha_1 u + \alpha_2 u^2 + \alpha_3 u^3 + ...$$

□

Στην περίπτωση που η ακολουθία $\{\alpha_x\}_{x=0,1,...}$ είναι πεπερασμένη, δηλαδή υπάρχει φυσικός αριθμός m τέτοιος ώστε να ισχύει

$$\alpha_m \neq 0 \quad \kappa\alpha\iota \quad \alpha_x = 0, \forall x > m,$$

τότε η συνάρτηση $A(u)$ είναι ένα πολυώνυμο βαθμού m, δηλαδή

$$A(u) = \sum_{x=0}^{m} \alpha_x u^x = \alpha_0 + \alpha_1 u + \alpha_2 u^2 + \alpha_3 u^3 + ... + \alpha_m u^m.$$

Είναι προφανές, ότι αν η ακολουθία $\{\alpha_x : x = 0, 1, 2, ...\}$ είναι φραγμένη, τότε η γεννήτρια συνάρτηση $A(u)$ συγκλίνει τουλάχιστον για $|u| < 1$.

Όπως ορίσαμε την συνέλιξη μεταξύ δύο συναρτήσεων, έτσι μπορούμε να ορίσουμε και την συνέλιξη μεταξύ δύο γεννητριών συναρτήσεων ακολουθιών.

ΟΡΙΣΜΟΣ 2.6 *Έστω $\{\alpha_x : x = 0, 1, 2, ...\}$ και $\{\beta_x : x = 0, 1, 2, ...\}$ δύο ακολουθίες πραγματικών αριθμών και $A(u) = \sum_{x=0}^{\infty} \alpha_x u^x$, $B(u) = \sum_{x=0}^{\infty} b_x u^x$ οι αντίστοιχες γεννήτριες* *συναρτήσεις τους. Τότε η συνέλιξη των a_x και b_y συμβολίζεται με $(A*B)(u)$ και ορίζεται ως*

$$(A * B)(u) = \sum_{x=0}^{\infty} \left(\sum_{y=0}^{x} \alpha_y b_{x-y} \right) u^x.$$

□

Επίσης, ισχύει ότι:

$$(A * B)(u) = (B * A)(u) = \sum_{x=0}^{\infty} \left(\sum_{y=0}^{x} \alpha_y b_{x-y} \right) u^x = \sum_{x=0}^{\infty} \left(\sum_{y=0}^{x} \alpha_{x-y} b_y \right) u^x$$

και

$$(A * B)(u) = A(u)B(u).$$

(βλ. Χατζηκωνσταντινίδης, 2015).

Κεφάλαιο 3

ΤΟ ΚΛΑΣΙΚΟ ΜΟΝΤΕΛΟ ΤΗΣ ΘΕΩΡΙΑΣ ΚΙΝΔΥΝΩΝ

Στόχος του παρόντος κεφαλαίου είναι η μελέτη του κλασικού μοντέλου της θεωρίας κινδύνου και η εύρεση της κατανομής των συνολικών απαιτήσεων σε διάφορες περιπτώσεις. Στην θεωρία κινδύνου διακρίνουμε δυο μορφές χαρτοφυλακίου, το ατομικό και το συλλογικό. Παρότι οι διαφορές είναι ελάχιστες και πρακτικά το ατομικό μπορεί να θεωρηθεί υποσύνολο του συλλογικού, για διδακτικούς λόγους εξετάζονται χωριστά. Το ατομικό πρότυπο εξετάζει τα επιμέρους τμήματα του χαρτοφυλακίου χωριστά ως ανεξάρτητες τυχαίες μεταβλητές, σε αντίθεση με το συλλογικό πρότυπο που αντιμετωπίζει το χαρτοφυλάκιο ως ένα ενιαίο σύνολο αποζημιώσεων. Μια κρίσιμη διαφορά των παραπάνω προτύπων αποτελεί το γεγονός ότι οι αποζημιώσεις στο ατομικό δεν είναι απαραίτητο να είναι ισόνομες, εν αντιθέσει με το συλλογικό που είναι ανεξάρτητες και ισόνομες.

Στην παρούσα διπλωματική εργασία θα παραλειφθεί το τμήμα του ατομικού προτύπου καθώς, όπως προαναφέρθηκε, μπορεί να θεωρηθεί μέρος του συλλογικού.

Το εν λόγω κεφάλαιο αποτελεί μια εισαγωγή στο κλασικό μοντέλο της θεωρίας κινδύνου. Συγκεκριμένα, στην Ενότητα 3.1 θα γίνει μια απλή παρουσίαση του μοντέλου συλλογικού κινδύνου, μέσω του τύπου της τυχαίας μεταβλητής των συνολικών ζημιών, την πιθανότητα στο μηδέν, μέση τιμή και διακύμανση της τυχαίας μεταβλητής S καθώς και τις γεννήτριες συναρτήσεις της εν λόγω τυχαίας μεταβλητής.

Στην συνέχεια, οι Ενότητες 3.2 και 3.3 αναφέρονται στις κλάσεις κατανομών $\mathscr{R}(\alpha, b, 0)$ και $\mathscr{R}(\alpha, b, 1)$ αντίστοιχα και στους αναδρομικούς τύπους υπολογισμού της συνάρτησης πιθανότητας της τυχαίας μεταβλητής S.

3.1 Μοντέλο συλλογικού κινδύνου

Θεωρούμε ένα χαρτοφυλάκιο κινδύνων με N το πλήθος αποζημιώσεων και X το ύψος της αποζημίωσης να είναι ανεξάρτητα μεταξύ τους. Στο συλλογικό πρότυπο, στην περίπτωση όπου δεν σημειωθεί καμία απαίτηση, το ποσό που θα καταβάλλει η εταιρία θα είναι μηδέν, δηλαδή

$$N = 0 \Rightarrow S = 0.$$

Συνεπώς, το μοντέλο μας θα έχει την μορφή:

$$S = \begin{cases} 0, & \text{N=0,} \\ \sum_{i=1}^{N} X_i, & \text{N=1,2,...} \end{cases} \tag{3.1}$$

Κύριο μέλημα της Θεωρίας Κινδύνου για το συλλογικό πρότυπο είναι η εύρεση της κατανομής των συνολικών απαιτήσεων, δηλαδή της S. Αυτό μπορεί να επιτευχθεί με δύο τρόπους, είτε μέσω συνελίξεων, είτε μέσω γεννητριών συναρτήσεων.

3.1.1 Η κατανομή της τυχαίας μεταβλητής των συνολικών ζημιών

Κατά την χρήση του συλλογικού μοντέλου ως μέσο περιγραφής του μεγέθους συνολικής ζημιάς χαρτοφυλακίου, βασική προϋπόθεση θεωρείται η εύρεση της μορφής της κατανομής των συνολικών απαιτήσεων S. Παρά το γεγονός ότι η τυχαία μεταβλητή N είναι πάντα μη-αρνητική ακέραια τυχαία μεταβλητή, η κατανομή της τυχαίας μεταβλητής X μπορεί να είναι είτε διακριτή είτε συνεχής. Με βάση αυτή την ιδιαιτερότητα, η μεταβλητή S μπορεί να έχει τις εξής μορφές:

1. Στην περίπτωση που η X είναι διακριτή τυχαία μεταβλητή, η τυχαία μεταβλητή S είναι και αυτή διακριτή, ως άθροισμα διακριτών τυχαίων μεταβλητών.

2. Στην περίπτωση που η X είναι συνεχής, η S μπορεί να είναι συνεχής ή μικτού τύπου τυχαία μεταβλητή ανάλογα με την τιμή της $P(N=0)$:

   - Αν η $P(N=0)=0$ τότε η τυχαία μεταβλητή S θα είναι και αυτή συνεχής τ.μ. ως άθροισμα συνεχών τυχαίων μεταβλητών.
   - Αν η $P(N=0)>0$ τότε η τυχαία μεταβλητή S θα είναι μικτού τύπου με μάζα πιθανότητας στο μηδέν ($P(S=0)=P(N=0)=p_0$) και συνεχής στο διάστημα $(0,+\infty)$.

Η δυσκολία στην μελέτη των τυχαίων αθροισμάτων $S = X_1 + X_2 + ... + X_N$ οφείλεται στον τυχαίο αριθμό όρων που πρέπει να αθροιστούν. Για τον υπολογισμό της συνάρτησης κατανομής της τυχαίας μεταβλητής S όσο και των άλλων ποσοτήτων, σημαντικό ρόλο παίζει η τ.μ. N και όχι η μεταβλητή X. Πράγματι, δοθείσης της τιμής της τ.μ. N γνωρίζουμε και το πλήθος των όρων της τ.μ. S, αφού πλέον το πλήθος των όρων είναι συγκεκριμένο και όχι τυχαίο.

Στο θεώρημα που έπεται δίνονται κάποιοι τύπου υπολογισμού της συνάρτησης κατανομής, της συνάρτησης δεξιάς ουράς και της συνάρτησης πυκνότητας πιθανότητας της τ.μ. S με των συνελίξεων και χωρίς να γίνει κάποια υπόθεση για την κατανομή των τυχαίων μεταβλητών N και X.

ΘΕΩΡΗΜΑ 3.1 *Για $x \geq 0$, ισχύει ότι:*

(α)

$$G(x) = \sum_{n=0}^{\infty} p_n F^{*n}(x) = p_0 + \sum_{n=1}^{\infty} p_n F^{*n}(x), \tag{3.2}$$

όπου $p_n = P(N=n), \quad n = 0, 1, 2, ...$

(β)

$$\bar{G}(x) = \sum_{n=1}^{\infty} p_n \bar{F}^{*n}(x).$$

(γ) *Αν η X είναι μη-ακέραια τυχαία μεταβλητή, τότε για* $x = 0$*, είναι:*

$$P(S=0) = g(0) = \begin{cases} p_0, & \text{αν f(0)=0,} \\ P_N(f(0)), & \text{αν f(0)}\neq\text{0,} \end{cases}$$

όπου $P_N(u)$ *είναι η πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής N. Αν η X είναι θετική ακέραια τυχαία μεταβλητή, τότε ισχύει:*

$$g(0) = p_0 \quad \textbf{και} \quad g(x) = \sum_{n=1}^{\infty} p_n f^{*n}(x), \quad \textbf{για } x = 1, 2, 3, ...$$

(δ) *Αν η X είναι συνεχής τ.μ. και* $P(N=0) = p_0 = 0$*, τότε:*

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

(ε) *Αν η X είναι συνεχής τ.μ. και* $P(N=0) = p_0 > 0$*, τότε:*

$$g(x) = \begin{cases} p_0, & \text{x=0,} \\ \sum_{n=1}^{\infty} p_n f^{*n}(x), & \text{x>0.} \end{cases}$$

(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).

Από το Θεώρημα 3.1 γίνεται κατανοητό με ποιο τρόπο πρέπει να συνδυαστούν η συνάρτηση πιθανότητας της τ.μ. N με αυτή της τ.μ. X για να βρούμε τον ακριβή τύπο υπολογισμού της g(x). Στην πράξη, επειδή αρκετές φορές είναι δύσκολος ο υπολογισμός των συνελίξεων f^{*n}, F^{*n} και $\bar{F}^{*n}$, στην συνέχεια θα παρουσιαστεί μια απλή συνδυαστική μέθοδος υπολογισμού της συνάρτησης πιθανότητας $g(x) = P(S = x)$ για $x = 0, 1, 2, ...$ όταν η X είναι μια θετική ακέραια τυχαία μεταβλητή.

Για παράδειγμα είναι:

$g(0) = p_0$

$$g(1) = P(S = 1, N = 1) = P(X_1 = 1, N = 1) = P(X_1 = 1)P(N = 1) = p_1 f(1)$$

$$\begin{aligned} g(2) &= P(S = 2, N = 2) + P(S = 2, N = 1) = P(X_1 + X_2 = 2, N = 2) + P(X_1 = 2, N = 1) \\ &= P(X_1 = 1)P(X_2 = 1)P(N = 2) + P(X_1 = 2)P(N = 1) \\ &= p_2 f^2(1) + p_1 f(2) \end{aligned}$$

κ.ο.κ.

Π.χ., για $x = 0, 1, 2, 3$ και 4, είναι:

$g(0)=$	p_0
$g(1)=$	$p_1 f(1)$
$g(2) =$	$p_1 f(2) + p_2 f^2(1)$
$g(3) =$	$p_1 f(3) + 2p_2 f(1)f(2) + p_3 f^3(1)$
$g(4) =$	$p_1 f(4) + 2p_2 f(1)f(2) + p_2 f^2(2) + 3p_3 f^2(1)f(2) + p_4 f^4(1)$

3.1.2 Η πιθανότητα στο μηδέν

Στο σημείο αυτό θα γίνει μια αναφορά στον τρόπο υπολογισμού της πιθανότητας της τ.μ. S όταν παίρνει την τιμή μηδέν. Έτσι, έχουμε:

1. Αν η X είναι συνεχής τ.μ., τότε:

   - Αν η τ.μ. N έχει μάζα πιθανότητας στο σημείο μηδέν, δηλαδή αν είναι $P(N = 0) = p_0 > 0$, τότε:

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

   - Αν η τ.μ. N δεν έχει μάζα πιθανότητας στο σημείο μηδέν, δηλαδή αν είναι $P(N = 0) = 0$, τότε:

     $$g(0) = 0.$$

2. Αν η X είναι θετική ακέραια τ.μ., δηλαδή αν είναι $f(0) = P(X = 0) = 0$, τότε:

- Αν η τ.μ. N έχει μάζα πιθανότητας στο σημείο μηδέν, δηλαδή αν είναι $P(N=0)=p_0>0$, τότε:
$$g(0)=p_0.$$
- Αν η τ.μ. N δεν έχει μάζα πιθανότητας στο σημείο μηδέν, δηλαδή αν είναι $P(N=0)=0$, τότε:
$$g(0)=0.$$

3. Αν η X είναι μη-αρνητική ακέραια τ.μ., δηλαδή αν είναι $f(0)=P(X=0)>0$, τότε:
$$g(0)=P_N(f(0))=M_N(lnf(0)). \tag{3.3}$$

ΠΑΡΑΔΕΙΓΜΑ 3.1 *Έστω ένα ασφαλιστήριο συμβόλαιο για το οποίο το πλήθος των απαιτήσεων N σε ένα έτος έχει συνάρτηση πιθανότητας* $p_n=P(N=n)$

n	*0*	*1*	*2*	*3*
p_n	*0.2*	*0.2*	*0.2*	*0.4*

Το μέγεθος της ατομικής απαίτησης X έχει συνάρτηση πιθανότητας $f(x)=P(X=x)$

x	*1*	*2*	*3*
$f(x)$	*0.4*	*0.4*	*0.2*

Επειδή η X είναι θετική ακέραια τυχαία μεταβλητή και η N είναι μη αρνητική ακέραια τ.μ., έπεται ότι η S είναι μη-αρνητική ακέραια τυχαία μεταβλητή. Επίσης, η τ.μ. S είναι:

- $S=0$, *αν* $N=0$
- $S=X_1$, *αν* $N=1$
- $S=X_1+X_2$, *αν* $N=2$ *και*
- $S=X_1+X_2+X_3$, *αν* $N=3$

Επειδή οι τυχαίες μεταβλητές $X_i\in\{1,2,3\}$, *έπεται ότι η τ.μ.*
$S\in\{0,1,2,3,4,5,6,7,8,9\}$.

Επειδή $f(0) = 0$, *είναι*

$$g(0) = P_N[f(0)] = P_N(0) = p_0 = 0.2.$$

Επιπλέον, επειδή $p_n = 0 \ \forall n \geq 4$, *τότε για* $x \geq 1$, *είναι:*

$$g(x) = \sum_{n=1}^{3} p_n f^{*n}(x) = p_1 f^{*1}(x) + p_2 f^{*2}(x) + p_3 f^{*3}(x)$$

δηλαδή η $g(x)$ *θα υπολογιστεί μέσω της σχέσης:*

$$g(x) = 0.2f(x) + 0.2f^{*2}(x) + 0.4f^{*3}(x), \quad x = 1, 2, ..., 9.$$

Επομένως, για τον υπολογισμό της $g(x)$ *είναι απαραίτητο να υπολογίσουμε την δεύτερη και την τρίτη συνέλιξη της τυχαίας μεταβλητής Χ.*

Η δεύτερη συνέλιξη υπολογίζεται από την σχέση:

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

Έτσι, έχουμε

$f^{*2}(0) = f(0)f(0) = 0$

$f^{*2}(1) = f(0)f(1) + f(1)f(0) = 0$

$f^{*2}(2) = f(0)f(2) + f(1)f(1) + f(2)f(0) = f(1)f(1) = (0.4)(0.4) = 0.16$

$f^{*2}(3) = f(0)f(3) + f(1)f(2) + f(2)f(1) + f(3)f(0) = 2f(1)f(2) = 2(0.4)(0.4) = 0.32$

κ.ο.κ.

Η τρίτη συνέλιξη υπολογίζεται από την σχέση:

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

Έτσι, έχουμε

$f^{*3}(0) = f(0)f^{*2}(0) = 0$

$f^{*3}(1) = f(0)f^{*2}(1) + f(1)f^{*2}(0) = 0$

$f^{*3}(2) = f(0)f^{*2}(2) + f(1)f^{*2}(1) + f(2)f^{*2}(0) = 0$

$f^{*3}(3) = f(0)f^{*2}(3) + f(1)f^{*2}(2) + f(2)f^{*2}(1) + f(3)f^{*2}(0) = f(1)f^{*2}(2) =$
$(0.4)(0.16) = 0.064$

κ.ο.κ.

Τότε,

$g(1) = 0.3f(1) + 0.4f^{*2}(1) + 0.2f^{*3}(1) = (0.2)(0.4) + (0.2)(0) + (0.4)(0) = 0.08$

$g(2) = 0.3f(2) + 0.4f^{*2}(2) + 0.2f^{*3}(2) = (0.2)(0.4) + (0.2)(0.16) + (0.4)(0) = 0.112$

$g(3) = 0.3f(3) + 0.4f^{*2}(3) + 0.2f^{*3}(3) = (0.2)(0.2) + (0.2)(0.16) + (0.4)(0.064) = 0.1296$

κ.ο.κ.

Αντίστοιχα, η συνάρτηση κατανομής υπολογίζεται από την σχέση:

$$G(x) = \sum_{n=1}^{3} p_n F^{*n}(x) = p_1 F^{*1}(x) + p_2 F^{*2}(x) + p_3 F^{*3}(x)$$

και επειδή $F^{*0}(x) = 1$, για $x \geq 0$ και $F^{*1}(x) = F(x)$, η G(x) μπορεί να υπολογιστεί από την σχέση:

$$G(x) = 0.2 + 0.2F(x) + 0.2F^{*2}(x) + 0.4F^{*3}(x), \quad x = 0, 1, ..., 9.$$

Επομένως, χρειάζεται να υπολογιστούν οι τιμές των συναρτήσεων κατανομής F της τυχαίας μεταβλητής X, καθώς και οι δεύτερη και τρίτη συνέλιξη αυτής, μέσω των σχέσεων:

$$F(x) = \sum_{y=0}^{x} f(y)$$
$$F^{*2}(x) = \sum_{y=0}^{x} f(y)F(x-y)$$
$$F^{*3}(x) = \sum_{y=0}^{x} f(y)F^{*3}(x-y)$$

Επίσης, οι τιμές της συνάρτησης κατανομής G(x) μπορούν να βρεθούν και από τις τιμές της συνάρτησης πιθανότητας g(x) ως εξής:

$$G(x) = \sum_{y=0}^{x} g(y).$$

Τα αποτελέσματα των υπολογισμών δίνονται στον κάτωθι πίνακα:

x	$f^{*0}(x)$	$f(x)$	$f^{*2}(x)$	$f^{*3}(x)$	$g(x)$	$G(x)$
0	1	0.0	0.00	0.000	0.2000	0.2000
1	0	0.4	0.00	0.000	0.0800	0.2800
2	0	0.4	0.16	0.000	0.1120	0.3920
3	0	0.2	0.32	0.064	0.1296	0.5216
4	0	0.0	0.32	0.192	0.1408	0.6624
5	0	0.0	0.16	0.288	0.1472	0.8096
6	0	0.0	0.04	0.256	0.1104	0.9200
7	0	0.0	0.00	0.144	0.0576	0.9776
8	0	0.0	0.00	0.048	0.0192	0.9968
9	0	0.0	0.00	0.008	0.2000	1.0000

(βλ. ΠΑΡΑΡΤΗΜΑ Δ).

□

3.1.3 Μέση τιμή και διακύμανση

Ο υπολογισμός της μέσης τιμής και της διακύμανσης των συνολικών ζημιών S μπορεί να γίνει με διάφορους τρόπους:

(α) Μέσω του θεωρήματος της διπλής μέσης τιμής, δεσμεύοντας ως προς N και χρησιμοποιώνταε τις σχέσεις:

$$E(S) = [E(S|N)]$$

$$Var(S) = E[Var(S|N)] + Var[E(S|N)]$$

(β) Μέσω των γεννητριών συναρτήσεων της τυχαίας μεταβλητής S.

(γ) Μέσω της συνάρτησης πυκνότητας πιθανότητας (ή της συνάρτησης πιθανότητας) ή και της συνάρτησης δεξιάς ουράς της τυχαίας μεταβλητής S.

ΘΕΩΡΗΜΑ 3.2 *Ισχύει ότι:*

(α) $E(S) = E(N)E(X)$,

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

(**Απόδειξη**: βλ. Dickson, 2005).

ΠΑΡΑΔΕΙΓΜΑ 3.2 *Έστω ότι μια επιχείρηση με 500 υπαλλήλους αποφασίζει να τους ασφαλίσει για μια επίσκεψη στον γιατρό. Κάθε υπάλληλος έχει 20% πιθανότητα να επισκεφθεί τον γιατρό μέσα σε ένα έτος και το κόστος σε ευρώ ακολουθεί ομοιόμορφη κατανομή στο διάστημα (0,2000). Οι υπάλληλοι θεωρούνται ανεξάρτητοι μεταξύ τους. Έστω S ότι είναι το συνολικό κόστος όλων των υπαλλήλων για το επόμενο έτος.*

Αν $N_i, 1 \leq i \leq 100$, *το πλήθος των επισκέψεων του i υπαλλήλου τότε, οι τυχαίες μεταβλητές* $N_1, N_2, ..., N_{100}$ *είναι μεταξύ τους ανεξάρτητες και ισόνομες με συνάρτηση πιθανότητας*

$$P(N_i = 0) = 0.8, \;\; P(N_i = 1) = 0.2, \;\; i = 1, 2, ..., 500,$$

δηλαδή η τυχαία μεταβλητή $N_i \sim Bernoulli(0.20)$

Έστω τώρα N είναι το πλήθος των επισκέψεων και των 500 υπαλλήλων, συνεπώς:

$$N = N_1 + N_2 + ... + N_{500},$$

δηλαδή η τ.μ $N \sim B(500, 0.20)$

Συνεπώς, το συνολικό κόστος S χρησιμοποιώντας το συλλογικό πρότυπο, δίνεται από την σχέση:

$$S = X_1 + X_2 + ... + X_N.$$

Είναι,

$$E(N) = np = 500(0.20) = 100, \qquad Var(N) = np(1-p) = 500(0.20)(1-0.20) = 80$$

$$E(X) = \frac{2000}{2} = 1000, \qquad Var(X) = \frac{2000^2}{12} = 333333,33$$

Τότε, για τις συνολικές αποζημιώσεις προκύπτουν τα εξής:
$E(S) = E(N)E(X) = 100000$
$Var(S) = E(N)Var(X) + Var(N)E^2(X) = 113333333,33$

$\square$

Στην υποενότητα που ακολουθεί θα γίνει αναφορά του τρόπου υπολογισμού των γεννητριών συναρτήσεων της τυχαίας μεταβλητής S, μέσω των αντίστοιχων γεννητριών συναρτήσεων των τυχαίων μεταβλητών N και X.

3.1.4 Γεννήτριες συναρτήσεις

ΘΕΩΡΗΜΑ 3.3 *Ισχύουν τα εξής:*

(α) $M_S(t) = M_N[lnM_X(t)]$,

(β) $M_S(t) = P_N[M_X(t)]$,

(γ) $P_S(u) = P_N[P_X(u)]$,

(δ) $\hat{g}(s) = P_N[\hat{f}(s)]$.

(βλ. Κούτρας, 2005).

Ως γνωστόν, η σύνθεση δύο συναρτήσεων f και g που συμβολίζεται $f \circ g$, ορίζεται ως $f[g(x)]$. Σύμφωνα με το Θεώρημα 3.3 προκύπτει ότι, η ροπογεννήτρια συνάρτηση της τυχαίας μεταβλητής S είναι η σύνθεση της πιθανογεννήτριας συνάρτησης της N και της ροπογεννήτριας συνάρτησης της X (ανάλογα ισχύουν και για τις άλλες περιπτώσεις). Για τον λόγο αυτό έπεται ότι "*η τυχαία μεταβλητή S έχει μια σύνθετη (compound) κατανομή*".

Η ονομασία της κατανομής που ακολουθεί η S, προσδιορίζεται από την κατανομή που ακολουθεί η τυχαία μεταβλητή N. Αν είναι γνωστή και η κατανομή της X, π.χ. αν η $N \sim Geo$ και η $X \sim Exp$, τότε η τυχαία μεταβλητή S ακολουθεί την σύνθετη γεωμετρική-εκθετική κατανομή.

Στην συνέχεια παρατίθεται ένας πίνακας με την μορφή που παίρνουν οι γεννήτριες συναρτήσεις στις διάφορες περιπτώσεις σύνθετων κατανομών, ανάλογα με την κατανομή που ακολουθεί η τ.μ. N.

Πίνακας 3.1.1: Γεννήτριες Συναρτήσεις [9]

	Κατανομή της τ.μ. N	$M_S(t)$	$P_S(u)$	$L_S(s)$
1	Bernoulli(p)	$1-p+pM_x(t)$	$1-p+pP_X(u)$	$1-p+pL_X(s)$
2	B(m,p)	$(1-p+pM_x(t))^m$	$(1-p+pP_X(u))^m$	$(1-p+pL_X(s))^m$
3	P(λ)	$exp\{\lambda(M_X(t)-1)\}$	$exp\{\lambda(P_X(u)-1)\}$	$exp\{\lambda(L_x(s)-1)\}$
4	Geo(p)	$\frac{p}{1-(1-p)M_X(t)}$	$\frac{p}{1-(1-p)P_X(u)}$	$\frac{p}{1-(1-p)L_X(s)}$
5	$Geo_1(p)$	$\frac{pM_X(t)}{1-(1-p)M_X(t)}$	$\frac{pP_X(u)}{1-(1-p)P_X(u)}$	$\frac{pL_X(s)}{1-(1-p)L_X(s)}$
6	$NB(r,p)$	$\left(\frac{p}{1-(1-p)M_X(t)}\right)^r$	$\left(\frac{p}{1-(1-p)P_X(u)}\right)^r$	$\left(\frac{p}{1-(1-p)L_X(s)}\right)^r$
7	$NB_1(r,p)$	$\left(\frac{pM_X(t)}{1-(1-p)M_X(t)}\right)^r$	$\left(\frac{pP_X(u)}{1-(1-p)P_X(u)}\right)^r$	$\left(\frac{pL_X(s)}{1-(1-p)L_X(s)}\right)^r$

3.2 Η κλάση κατανομών $\mathscr{R}(\alpha, b, 0)$

Αν και η κατανομή των συνολικών αποζημιώσεων δεν είναι γενικά εύκολο να υπολογισθεί, στην περίπτωση όπου τα ατομικά μεγέθη των απαιτήσεων παίρνουν ακέραιες και μη αρνητικές τιμές υπάρχει ένας αναδρομικός τύπος που επιτρέπει τον υπολογισμό της κατανομής αυτής. Ο τύπος αυτός μπορεί να χρησιμοποιηθεί κάτω από την προϋπόθεση ότι η κατανομή του πλήθους των απαιτήσεων ανήκει στην ειδική κατηγορία διακριτών κατανομών $\mathscr{R}(\alpha, b, 0)$. [8]

ΟΡΙΣΜΟΣ 3.1 *Η διακριτή τ.μ $N \in \{0,1,2,...\}$, θα ανήκει στην οικογένεια κατανομών $\mathscr{R}(\alpha, b, 0)$, αν η συνάρτηση πιθανότητας $p_n = P(N=n)$ ικανοποιεί την αναδρομική σχέση*

$$p_n = \left(\alpha + \frac{b}{n}\right) p_{n-1}, \quad n = 1, 2, ..., \tag{3.4}$$

όπου τα α και b είναι κατάλληλες σταθερές και $p_n = 0, \ \forall n < 0$.

□

Προφανώς, αν είναι γνωστή η τιμή της $p_0 = P(N = 0)$, τότε σύμφωνα με την 3.4 βρίσκονται διαδοχικά οι τιμές

$$p_1 = (\alpha + b)p_0, \;\; p_2 = \left(\alpha + \frac{b}{2}\right)p_1, \;\; p_3 = \left(\alpha + \frac{b}{3}\right)p_2, ...$$

Στον παρακάτω πίνακα δίνονται τα μέλη της κλάσης $\mathscr{R}(\alpha, b, 0)$ και οι αντίστοιχες τιμές των παραμέτρων α, b και p_0.

Πίνακας 3.2.1: Μέλη της κλάσης $\mathscr{R}(\alpha, b, 0)$

	Κατανομή	α	b	p_0
1	Poisson(λ)	0	λ	$e^{-\lambda}$
2	B(m,p)	$-\frac{p}{q}$	$\frac{(m+1)p}{q}$	q^m
3	NB(r,p)	q	$(r-1)q$	p^r

Αν μια μεταβλητή $X \sim G_0(p)$, έτσι ώστε

$$p_n = P(X = n) = pq^n, \quad n = 0, 1, 2, ...$$

με $q = 1 - p$, τότε είναι σαφές ότι για $n = 1, 2, ...$, ισχύει

$$p_n = p_{n-1}q.$$

Δηλαδή η γεωμετρική κατανομή ανήκει στην κλάση κατανομών $\mathscr{R}(\alpha, b, 0)$ ως ειδική περίπτωση της αρνητικής διωνυμικής κατανομής, με

$$\alpha = q = 1 - p, \quad b = 0.$$

(**Απόδειξη**: βλ. Πολίτης, 2012).

3.2.1 Αναδρομικός τύπος υπολογισμού της g(x)

ΛΗΜΜΑ 3.1 *Αν η τ.μ. $N \in \mathscr{R}(\alpha, b, 0)$, τότε η πιθανογεννήτρια συνάρτηση $P_N(u)$ της τ.μ. N, ικανοποιεί την διαφορική εξίσωση*

$$P_N'(u) = \frac{\alpha + b}{1 - \alpha u} P_N(u). \tag{3.5}$$

(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2012).

ΘΕΩΡΗΜΑ 3.4 *(Αναδρομικός τύπος του Panjer)*

Έστω $S = X_1 + X_2 + ... + X_N, N \geq 1$ (με $S = 0$ αν $N = 0$). Αν η X είναι μη-αρνητική ακέραια τυχαία μεταβλητή , δηλαδή $X \in \{0, 1, 2, ...\}$ με συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τυχαία μεταβλητή $N \in \mathscr{R}(\alpha, b, 0)$, τότε η $g(x) = P(S = x)$ ικανοποιεί την αναδρομική σχέση

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

με $g(0) = P_N(f(0))$.

(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).

Από το παραπάνω Θεώρημα, προκύπτουν άμεσα τα εξής πορίσματα, που αφορούν ειδικές περιπτώσεις για την κατανομή της N.

ΠΟΡΙΣΜΑ 3.1 *Αν η τ.μ. $N \in P(\lambda), \lambda > 0$, τότε η $g(x) = P(S = x)$ ικανοποιεί την αναδρομική σχέση,*

$$g(x) = \frac{\lambda}{x} \sum_{y=1}^{x} f(y) g(x - y), \quad x = 1, 2, 3, ...$$

με $g(0) = e^{\lambda(f(0) - 1)}$.

ΠΟΡΙΣΜΑ 3.2 *Αν η τ.μ.* $N \sim B(m,p), 0 < p < 1, q = 1-p$, *τότε η* $g(x) = P(S = x)$ *ικανοποιεί την αναδρομική σχέση,*

$$g(x) = \frac{p}{q + pf(0)} \sum_{y=1}^{x} \left(\frac{(m+1)y}{x} - 1 \right) f(y) g(x-y), \quad x = 1, 2, 3, ...$$

με $g(0) = (q + pf(0))^m$.

Προφανώς, αν η τ.μ. $N \sim Bernoulli(p), 0 < p < 1$, ως ειδική περίπτωση της κατανομής $B(m,p)$, όπου $m = 1$, μπορεί να θεωρηθεί ότι ανήκει στην κλάση κατανομών $\mathscr{R}(\alpha, b, 0)$ με $\alpha = -\frac{p}{q}$ και $b = \frac{2q}{p}$. Συνεπώς, η αναδρομική σχέση που ικανοποιεί η $g(x) = P(S = x)$ όταν η τ.μ. S έχει μια σύνθετη Bernoulli κατανομή, προκύπτει άμεσα από το Πόρισμα 3.2 για $m = 1$.

Επιπλέον, αν $N \sim Bernoulli(p), 0 < p < 1$, τότε η $g(x) = P(S = x)$ είναι:

$$\begin{aligned} g(0) &= q + pf(0), \quad q = 1 - p \\ g(x) &= pf(x), \quad x = 1, 2, 3, ... \end{aligned}$$

Επειδή,

$$P_N(u) = q + pu$$

είναι

$$\begin{aligned} & P_S(u) = P_N(P_X(u)) = q + pP_X(u) \\ \Rightarrow \quad & \sum_{x=0}^{\infty} g(x) u^x = q + p \sum_{x=0}^{\infty} f(x) u^x \\ \Rightarrow \quad & g(0) + \sum_{x=1}^{\infty} g(x) u^x = q + pf(0) + p \sum_{x=1}^{\infty} f(x) u^x \end{aligned}$$

ΠΟΡΙΣΜΑ 3.3 *Αν η τ.μ* $N \sim NB(r,p), r > 0, 0 < p < 1, q = 1-p$, *με συνάρτηση πιθανότητας*

$$p_n = \binom{r+n-1}{n} p^r q^n, \quad n = 0, 1, 2, ...$$

τότε η $g(x) = (S = x)$ *ικανοποιεί την αναδρομική σχέση*

$$g(x) = \frac{q}{1 - qf(0)} \sum_{y=1}^{x} \left(1 + \frac{(r-1)y}{x}\right) f(y)g(x-y)$$

με $g(0) = \left(\dfrac{p}{1 - qf(0)}\right)^r.$

Από το Πόρισμα 3.3 για $r = 1$ προκύπτει άμεσα ότι:

Αν η τ.μ. $N \sim G(p), 0 < p < 1, q = 1 - p$, με συνάρτηση πιθανότητας

$$p_n = pq^n, \quad n = 0, 1, 2, ...$$

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

$$g(x) = \frac{q}{1 - qf(0)} \sum_{y=1}^{x} f(y)g(x-y), \quad x = 1, 2, 3, ...$$

με $g(0) = \dfrac{p}{1 - qf(0)}.$

3.2.2 Ροπές των συνολικών ζημιών

Στην παρούσα υποενότητα θα γίνει μια αναφορά στην μέση τιμή και την διακύμανση των συνολικών ζημιών ενός χαρτοφυλακίου, ανάλογα με την κατανομή που ακολουθεί η εκάστοτε τ.μ. N, με δεδομένο ότι ανήκει στην κλάση $\mathscr{R}(\alpha, b, 0)$.

ΠΟΡΙΣΜΑ 3.4 *Αν η τ.μ.* $N \in \mathscr{R}(\alpha, b, 0)$ *και* $S = X_1 + X_2 + ... + X_N, N \geq 1$ *(με* $S = 0$ *όταν* $N = 0$*), τότε η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής S αντίστοιχα είναι:*

(α) $E(S) = \dfrac{\alpha + b}{1 - \alpha} E(X),$

(β) $Var(S) = \dfrac{\alpha + b}{1 - \alpha} Var(X) + \dfrac{\alpha + b}{(1 - \alpha)^2} E^2(X).$

(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).

Από το Λήμμα 3.1 για $u = 1$, ισχύει ότι

$$E(N) = P_N'(1) = \frac{\alpha + b}{1 - \alpha} P_N(1) = \frac{\alpha + b}{1 - \alpha},$$

επειδή $P_N(1) = 1$.

Επιπλέον, ισχύει ότι

$$P_N^{"}(1) = E[N(N-1)] = E[N^2] - E[N] = Var(N) + E(N^2) - E(N),$$

οπότε είναι

$$Var(N) = P_N^{"}(1) + E(N^2) - E(N).$$

Από το Λήμμα 3.1, παραγωγίζοντας και τα δύο μέλη της Σχέσης (3.5) και θέτοντας $u = 1$, έπεται ότι

$$\begin{aligned} P_N^{"}(1) = E[N(N-1)] &= \frac{\alpha(\alpha + b)}{(1-\alpha)^2} + \frac{\alpha + b}{1 - \alpha} P_N'(1) \\ &= \frac{\alpha(\alpha + b)}{(1-\alpha)^2} + \frac{\alpha + b}{1 - \alpha} \cdot \frac{\alpha + b}{1 - \alpha} = \frac{(\alpha + b)(2\alpha + b)}{(1-\alpha)^2}, \end{aligned}$$

οπότε,

$$Var(N) = \frac{(\alpha + b)(2\alpha + b)}{(1-\alpha)^2} + \frac{\alpha + b}{1 - \alpha} - \left(\frac{\alpha + b}{1 - \alpha} \right)^2,$$

ή

$$Var(N) = \frac{\alpha + b}{(\alpha + b)^2}.$$

Συνεπώς, σύμφωνα με τις τιμές των παραμέτρων α και b του Πίνακα 3.2.1, το Πόρισμα 3.4 και τις σχέσεις

$$E(N) = \frac{\alpha + b}{1 - \alpha} \ , \ Var(N) = \frac{\alpha + b}{(1-\alpha)^2}$$

η μέση τιμή και η διακύμανση των τυχαίων μεταβλητών $N \in \mathscr{R}(\alpha, b, 0)$ και $S = X_1 + X_2 + ... + X_N$ δίνονται απο τις ακόλουθες σχέσεις:

(α) Αν η τ.μ. $N \sim P(\lambda)$, τότε για $\alpha = 0, b = \lambda$, είναι

$$\begin{aligned} &E(N) = Var(N) = \lambda, \\ &E(S) = \lambda E(X), \\ &Var(S) = \lambda E(X^2). \end{aligned}$$

(β) Αν η τ.μ. $N \sim B(m, p)$, τότε για $\alpha = -\frac{p}{q}, b = \frac{(m-1)p}{q}$ είναι

$$\begin{aligned} &E(N) = mp, \quad Var(N) = mpq, \\ &E(S) = mpE(X), \\ &Var(S) = mpVar(X) + mpqE^2(X) = mpE(X^2) - mp^2E^2(X). \end{aligned}$$

(γ) Αν η τ.μ. $N \sim NB(r, p)$, τότε για $\alpha = q, b = (r-1)q$, είναι

$$\begin{aligned} &E(N) = \frac{rq}{p}, \quad Var(N) = \frac{rq}{p^2}, \\ &E(S) = \frac{rq}{p}E(X), \\ &Var(S) = \frac{rq}{p}Var(X) + \frac{rq}{p^2}E^2(X) = \frac{rq}{p}E(X^2) + \frac{rq^2}{p^2}E^2(X). \end{aligned}$$

(δ) Αν η τ.μ. $N \sim G(p)$, τότε για $\alpha = q, b = 0$, είναι

$$\begin{aligned} &E(N) = \frac{q}{p}, \quad Var(N) = \frac{q}{p^2}, \\ &E(S) = \frac{q}{p}E(X), \\ &Var(S) = \frac{q}{p}Var(X) + \frac{q}{p^2}E^2(X) = \frac{q}{p}E(X^2) + \frac{q^2}{p^2}E^2(X). \end{aligned}$$

3.2.3 Αναδρομικοί τύποι υπολογισμού των $G(x)$ και $\bar{G}(x)$

Έστω οι συναρτήσεις

$$Q_S(u) = \sum_{x=0}^{\infty} G(x)u^x \;\; \kappa\alpha\iota \;\; R_S(u) = \sum_{x=0}^{\infty} \bar{G}(x)u^x$$

είναι οι γεννήτριες συναρτήσεις των ακολουθιών

$$\{G(x) : x = 0, 1, 2, ...\} \ \ \kappa\alpha\iota \ \ \{\bar{G}(x) : x = 0, 1, 2, ...\},$$

αντίστοιχα.

Επιπλέον ισχύει ότι:

$$R_S(u) = \frac{1 - P_S(u)}{1 - u}.$$

ΛΗΜΜΑ 3.2 *Αν* $P_S(u) = \sum_{x=0}^{\infty} g(x)u^x$ *είναι η πιθανογεννήτρια συνάρτηση της τ.μ. S, όπου* $g(s) = P(S = x)$*, και* $Q_S(u) = \sum_{x=0}^{\infty} G(x)u^x$*, όπου* $G(x) = P(S \leq x)$*, τότε ισχύει ότι*

$$Q_S(u) = \frac{P_S(u)}{1 - u}.$$

*(**Απόδειξη:** βλ. Χατζηκωνσταντινίδης, 2015).*

3.3 Η κλάση κατανομών $\mathscr{R}(\alpha, b, 1)$

Στην υποενότητα αυτή θα γίνει μια αναφορά στην περίπτωση όπου υπάρχει πρόβλημα μη-ικανοποιητικής προσαρμογής στο σημείο μηδέν. Σε ορισμένες περιπτώσεις, οι κλάσεις κατανομών $\mathscr{R}(\alpha, b, 0)$ δεν περιγράφουν ικανοποιητικά τα χαρακτηριστικά ενός συνόλου δεδομένων είτε λόγω της ύπαρξης μια βαριάς δεξιάς ουράς, είτε εξαιτίας μη-ικανοποιητικής προσαρμογής του σχήματος των δεδομένων στο σημείο μηδέν (αυτή είναι και η διαφορά μεταξύ των $\mathscr{R}(\alpha, b, 0)$ και $\mathscr{R}(\alpha, b, 1)$).

3.3.1 Περικομμένες στο σημείο μηδέν κατανομές

Ένα συνηθισμένο και βασικό πρόβλημα στην εύρεση της κατανομής των συνολικών ζημιών είναι η πιθανότητά της στο σημείο μηδέν. Υπάρχουν περιπτώσεις στις οποίες η ελάχιστη τιμή την οποία παίρνει το πλήθος των ζημιών είναι 1. Για τον λόγο αυτό ένα μεγάλο κομμάτι της θεωρίας κινδύνου όσο αφορά το συλλογικό μοντέλο αφιερώνεται στο κομμάτι των περικομμένων στο σημείο μηδέν κατανομών.

ΟΡΙΣΜΟΣ 3.2 *Η περικομμένη στο σημείο μηδέν τυχαία μεταβλητή (zero-truncated distribution) N^T που αντιστοιχεί στην τυχαία μεταβλητή $N \in \{0, 1, 2, ...\}$, ορίζεται ως η δεσμευμένη τυχαία μεταβλητή*

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

□

Δηλαδή, η τ.μ. N^T έχει συνάρτηση πιθανότητας που προκύπτει με "αποκοπή" του σημείου μηδέν από το πεδίο ορισμού της αρχικής τ.μ. N και επιμερισμό της πιθανότητας $p_n = P(N = n)$ στα σημεία $n = 1, 2, 3, ...$ ανάλογα προς την πιθανότητα που ήδη υπάρχει σ' αυτά τα σημεία. Έστω, $p_n^T = P(N^T = n)$ η συνάρτηση πιθανότητας της τ.μ. N^T.

Τότε,

$$p_n^T = \begin{cases} P(N^T = 0), & \text{n=0} \\ P(N^T = n), & \text{n=1,2,3,...} \end{cases}$$

$$= \begin{cases} 0, & n = 0 \\ P(N = n|N > 0), & n = 1, 2, 3, ... \end{cases}$$

όπου,

$$P(N = n|N > 0) = \frac{P(N = n)}{P(N > 0)} = \frac{P(N = n)}{1 - P(N \leq 0)} = \frac{P(N = n)}{1 - P(N = 0)} = \frac{p_n}{1 - p_0}.$$

Άρα η συνάρτηση πιθανότητας της τ.μ. N^T, είναι

$$p_n^T = \begin{cases} 0, & n = 0 \\ \dfrac{p_n}{1 - p_0}, & n = 1, 2, 3, ... \end{cases} \tag{3.6}$$

Αντίστοιχα, η πιθανογεννήτρια συνάρτηση της τ.μ. N^T, είναι:

$$P_{N^T}(u) = \frac{P_N(u) - p_0}{1 - p_0}. \tag{3.7}$$

3.3.2 Τροποποιημένες στο σημείο μηδέν κατανομές

Εκτός από τις περικομμένες στο μηδέν κατανομές, μια ακόμα βασική κατηγορία κατανομών αποτελούν οι τροποποιημένες στο σημείο μηδέν κατανομές, εξίσου σημαντικές καθώς μπορεί να ορισθεί η πιθανότητα στο εν λόγω σημείο ανάλογα με το χαρτοφυλάκιό κινδύνων.

ΟΡΙΣΜΟΣ 3.3 *Η N^M καλείται τροποποιημένη στο σημείο μηδέν τυχαία μεταβλητή που αντιστοιχεί στην τυχαία μεταβλητή $N \in \{0, 1, 2, ...\}$, αν έχει συνάρτηση πιθανότητας*

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

όπου p_0^M είναι οποιοσδήποτε αριθμός με $0 \leq p_0^M < 1$, $c = \dfrac{1 - p_0^M}{1 - p_0}$ και p_n είναι η συνάρτηση πιθανότητας της τ.μ. N.

□

Άρα, η συνάρτηση πιθανότητας της τ.μ. N^M, είναι

$$p_n^M = \begin{cases} p_0^M, & n = 0 \\ \dfrac{1 - p_0^M}{1 - p_0} p_n, & n = 1, 2, 3, ... \end{cases} \tag{3.8}$$

Από τις σχέσεις (3.6) και (3.8) έπεται ότι:

$$p_n^M = (1 - p_0^M) p_n^T, \quad n = 1, 2, 3, ...$$

Η πιθανογεννήτρια συνάρτηση της τ.μ. N^T, είναι

$$P_{N^M}(u) = \sum_{n=0}^{\infty} p_n^M u^n = p_0^M + \sum_{n=1}^{\infty} p_n^M u^n.$$

Οπότε, λόγω της Σχέσης (3.8), η παραπάνω σχέση γίνεται

$$\begin{aligned}
P_{N^M}(u) &= p_0^M + \frac{1-p_0^M}{1-p_0} \sum_{n=1}^{\infty} p_n u^n \\
&= p_0^M + \frac{1-p_0^M}{1-p_0} \left(\sum_{n=0}^{\infty} p_n u^n - p_0 \right) \\
&= p_0^M + \frac{1-p_0^M}{1-p_0} \left(P_N(u) - p_0 \right) \\
&= p_0^M + \frac{1-p_0^M}{1-p_0} p_0 + \frac{1-p_0^M}{1-p_0} P_N(u) \\
&= \frac{p_0^M - p_0}{1-p_0} + \frac{1-p_0^M}{1-p_0} P_N(u) \\
&= \frac{1-p_0-(1-p_0^M)}{1-p_0} + \frac{1-p_0^M}{1-p_0} P_N(u) \\
&= 1 - \frac{1-p_0^M}{1-p_0} + \frac{1-p_0^M}{1-p_0} P_N(u)
\end{aligned} \tag{3.9}$$

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

Επιπλέον, από τις σχέσεις (3.7) και (3.9) έπεται ότι:

$$P_{N^M}(u) = p_0^M + (1-p_0^M) P_{N^T}(u). \tag{3.10}$$

Έστω ότι στην σχέση 3.10 είναι $p_0^M = 0$, τότε ισχύει

$$P_{N^M}(u) = P_{N^T}(u),$$

δηλαδή οι τυχαίες μεταβλητές N^M και N^T είναι ισόνομες.

Με άλλα λόγια, η περικομμένη στο σημείο μηδέν τυχαία μεταβλητή N^M προκύπτει από την τροποποιημένη στο σημείο μηδέν τυχαία μεταβλητή N^T για $p_0^M = 0$, άρα η πρώτη μπορεί να θεωρηθεί ως ειδική περίπτωση της δεύτερης.

ΠΡΟΤΑΣΗ 3.1 *Έστω N^M η τροποποιημένη στο σημείο μηδέν τυχαία μεταβλητή που αντιστοιχεί στη διακριτή τυχαία μεταβλητή $N \in \{0, 1, 2, ...\}$. Τότε, η τυχαία μεταβλητή N^M έχει μια σύνθετη Bernoulli κατανομή. Δηλαδή:*

- *Είναι,*

$$N^M = N_1 + N_2 + ... + N_K,$$

όπου η τυχαία μεταβλητή

$$K \sim Bernoulli\left(\frac{1 - p_0^M}{1 - p_0}\right),$$

και οι τυχαίες μεταβλητές $N_1, N_2, ..., N_K$ είναι ανεξάρτητες και ισόνομες με την τ.μ. N.

- *Είναι,*

$$N^M = N_1^T + N_2^T + ... + N_L^T,$$

όπου η τυχαία μεταβλητή

$$L \sim Bernoulli(1 - p_0^M),$$

και οι τυχαίες μεταβλητές $N_1^T, N_2^T, ..., N_L^T$ είναι ανεξάρτητες και ισόνομες με την τ.μ. N^T.

ΠΡΟΤΑΣΗ 3.2 *Έστω η τυχαία μεταβλητή $N \in \mathscr{R}(\alpha, b, 0)$ και οι τυχαίες μεταβλητές N^T,N^M είναι η περικομμένη και η τροποποιημένη στο σημείο μηδέν τυχαία μεταβλητή αντίστοιχα. Τότε:*

(α) *Η συνάρτηση πιθανότητας* $p_n^T = P(N^T = n),\ n = 1,2,...$ *ικανοποιεί την αναδρομική σχέση*

$$p_n^T = \left(\alpha + \frac{b}{n}\right) p_{n-1}^T, \quad n = 2,3,4,...$$

(β) *Η συνάρτηση πιθανότητας* $p_n^M = P(N^M = n),\ n = 0,1,2,...$ *ικανοποιεί την αναδρομική σχέση*

$$p_n^M = \left(\alpha + \frac{b}{n}\right) p_{n-1}^M, \quad n = 2,3,4,...$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Στην συνέχεια θα δοθεί ένας ορισμός για μια νέα κλάση διακριτών κατανομών που όπως και η $\mathscr{R}(\alpha, b, 0)$, ορίζεται βάσει αναδρομικού τύπου.

ΟΡΙΣΜΟΣ 3.4 *Η διακριτή τυχαία μεταβλητή* $N \in \{0,1,2,...\}$*, ανήκει στην οικογένεια κατανομών* $\mathscr{R}(\alpha, b, 1)$*, αν η συνάρτηση πιθανότητας* $p_n = P(N = n)$ *ικανοποιεί την αναδρομική σχέση πρώτης τάξης*

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

όπου τα α και b είναι κατάλληλες σταθερές και $p_n = 0,\ \forall n < 0$ *και* $p_1 \neq 0$.

□

Στην συνέχεια θα παρουσιαστεί μια διακριτή κατανομή, η οποία ανήκει στην κλάση $\mathscr{R}(\alpha, b, 1)$.

ΟΡΙΣΜΟΣ 3.5 *Η θετική ακέραια τυχαία μεταβλητή* $N \in \{1,2,3,...\}$ *έχει την λογαριθμική κατανομή με παράμετρο* $q,\ 0 < q < 1$*, αν έχει συνάρτηση πιθανότητας*

$$p_n = \frac{q^n}{-nln(1-q)}, \quad n = 1,2,3,...$$

και συμβολικά γράφεται $N \sim LS(q)$.

Συνεπώς, η λογαριθμική κατανομή ανήκει στην κλάση κατανομών $\mathscr{R}(\alpha, b, 1)$ με $\alpha = q$ και $b = -q$, καθώς είναι ειδική περίπτωση της περικομμένης στο σημείο μηδέν NB(r,p).

□

Στην κλάση κατανομών $\mathscr{R}(\alpha, b, 1)$ ανήκουν όλες οι κατανομές που ανήκουν στην κλάση $\mathscr{R}(\alpha, b, 0)$ καθώς επίσης και οι τροποποιημένες και περικομμένες στο μηδέν Poisson, Διωνυμική, Αρνητική Διωνυμική, Γεωμετρική και λογαριθμική κατανομή.

Σύμφωνα όσα αναφέρθηκαν, προκύπτει ο παρακάτω πίνακας για την κλάση κατανομών $\mathscr{R}(\alpha, b, 1)$.

Πίνακας 3.3.1: Μέλη της κλάσης $\mathscr{R}(\alpha, b, 1)$[9]

	Κατανομή	p_0	α	b	Τιμές παραμέτρων
1	Poisson(λ)	$e^{-\lambda}$	0	λ	$\lambda > 0$
2	ZT P(λ)	0	0	λ	$\lambda > 0$
3	ZM P(λ)	οποιοδήποτε	0	λ	$\lambda > 0$
4	B(m,p)	q^m	$-p/q$	$(m+1)p/q$	$0 < p < 1$
5	ZT B(m,p)	0	$-p/q$	$(m+1)p/q$	$0 < p < 1$
6	ZM B(m,p)	οποιοδήποτε	$-p/q$	$(m+1)p/q$	$0 < p < 1$
7	NB(r,p)	p^r	q	$(r-1)q$	$r > 0, 0 < p < 1$
8	ETNB(r,p)	0	q	$(r-1)q$	$r > -1, 0 < p < 1$
9	ZM ETNB(r,p)	οποιοδήποτε	q	$(r-1)q$	$r > -1, 0 < p < 1$
10	G(p)		q	0	$0 < p < 1$
11	ZT G(p)	0	q	0	$0 < p < 1$
12	ZM G(p)	οποιοδήποτε	q	0	$0 < p < 1$
13	LS(q)	0	q	$-q$	$0 < p < 1$

όπου:

$p_0 = P(N = 0)$,

$q = 1 - p$,

ZT=Zero-Truncated (Περικομμένη στο μηδέν),

ZM=Zero-Modified (Τροποποιημένη στο μηδέν),

ETNB(r,p)= η περικομμένη στο σημείο μηδέν μιας NB(r,p), με $r > -1$,

ZM ETNB(r,p)= η τροποποιημένη στο σημείο μηδέν μιας ETNB(r,p), με $r > -1$.

3.3.3 Αναδρομικός τύπος υπολογισμού της g(x)

Αντίστοιχα με το Λήμμα 3.1 που αφορά την κλάση $\mathscr{R}(\alpha, b, 0)$, για την κλάση $\mathscr{R}(\alpha, b, 1)$ ισχύει το εξής:

ΛΗΜΜΑ 3.3 *Αν η τυχαία μεταβλητή* $N \in \mathscr{R}(\alpha, b, 1)$, *τότε η πιθανογεννήτρια συνάρτηση* $P_N(u)$ *της τ.μ. N, ικανοποιεί την διαφορική εξίσωση*

$$P_N'(u) = \frac{(\alpha+b)P_N(u) + p_1 - (\alpha+b)p_0}{1-\alpha u}.$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Επιπλέον, σε αντιστοιχία με το Θεώρημα 3.4 προκύπτει το παρακάτω Θεώρημα για τον αναδρομικό τύπο του Panjer στην περίπτωση της κλάσης $\mathscr{R}(\alpha, b, 1)$

ΘΕΩΡΗΜΑ 3.5 *(Αναδρομικός τύπος του Panjer)*

Έστω $S = X_1 + X_2 + ... + X_N, \;\; N \geq 1$ *(με* $S = 0$ *όταν* $N = 0$*). Αν η X είναι μη-αρνητική ακέραια τυχαία μεταβλητή, δηλαδή* $X \in \{0, 1, 2, ...\}$, *με συνάρτηση πιθανότητας* $f(x) = P(X = x)$ *και η τ.μ* $N \in \mathscr{R}(\alpha, b, 1)$, *τότε η* $g(x) = P(S = x)$ *ικανοποιεί την αναδρομική σχέση*

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

με $g(0) = P_N(f(0))$.

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Κεφάλαιο 4

ΤΟ ΠΑΚΕΤΟ **ACTUAR**

Στο κεφάλαιο αυτό θα γίνει μια σύντομη παρουσίαση του πακέτου actuar, καθώς στην πορεία της διπλωματικής θα χρησιμοποιηθεί εκτενώς. Αρχικά θα γίνει μια εισαγωγή στο τι περιέχει το πακέτο αυτό και γιατί θεωρείται χρήσιμο στην αναλογιστική επιστήμη και στην συνέχεια θα παρουσιαστούν κάποιες από τις λειτουργίες του. Παρά το γεγονός ότι το πακέτο αυτό καλύπτει αρκετά πεδία της αναλογιστικής επιστήμης, στην παρούσα διπλωματική θα παρουσιαστεί μόνο το κομμάτι της θεωρίας κινδύνου. Πιο συγκεκριμένα, στις Υποενότητες 4.2.1 και 4.2.2 θα γίνει μια πρώτη αναφορά στις μορφές διακριτοποίησης και στον υπολογισμό της κατανομής των συνολικών απαιτήσεων με χρήση της λειτουργίας *aggregateDist*. Τέλος, στην Ενότητα 4.2.2 θα παρουσιαστούν κάποιες εντολές για προσομοίωση ασφαλιστικών δεδομένων. Στη συνέχεια της διπλωματικής εργασίας (Κεφάλαιο 5) θα γίνει περαιτέρω χρήση των παραπάνω εντολών στις διάφορες περιπτώσεις των τυχαίων μεταβλητών.

4.1 Εισαγωγή

Η γλώσσα προγραμματισμού R δεν είναι απλά ένα ακόμα στατιστικό πρόγραμμα, αλλά μια πλήρης και ανεξάρτητη γλώσσα με ισχυρό μαθηματικό προσανατολισμό. Το όλο πρόγραμμα βρίσκεται υπό ενεργή ανάπτυξη και κερδίζει έδαφος με την πάροδο του χρόνου σε όλους τους κλάδους που συνδέονται με την στατιστική. Ένα απο τα δυνατά σημεία του θεωρείται το γεγονός ότι επιτρέπει στην χρήστη την δυνατότητα προσθήκης λειτουργιών μέσω διαφόρων πακέτων. Με απλά λόγια, κάθε πακέτο περιλαμβάνει ένα ευρύ φάσμα λειτουργιών και εντολών, ικανό να καλύψει ένα μεγάλο μέρος από τις απαιτήσεις του χρήστη.

Το πακέτο actuar δημιουργήθηκε με σκοπό την εύκολη και άμεση ανάλυση αναλογιστικών μοντέλων με χρήση προσομοίωσης. Παρά το γεγονός ότι ο κλάδος του αναλογισμού υπάρχει εδώ και αρκετά χρόνια, η επίλυση των επιμέρους προβλημάτων του με τον παραδοσιακό τρόπο θεωρείται αναξιόπιστη, καθώς η εξάλειψη ανθρωπίνου λάθους καθίσταται αδύνατη.

Το actuar πήρε το όνομά του από το actuarial, που σημαίνει αναλογισμός και χωρίζεται σε τέσσερις κατηγορίες: ζημιοκατανομές, θεωρία ακραίων τιμών, θεωρία κινδύνου (περιλαμβάνει και θεωρία χρεοκοπίας) και θεωρία αξιοπιστίας. Οι προγραμματιστές αν και θεώρησαν ορθό να διατηρήσουν την ιδιαιτερότητα των εντολών, η φιλοσοφία του όλου προγράμματος της γλώσσας R επηρέασε την τελική δομή για τα μοντέλα χρήσης του.

Το έργο ξεκίνησε επίσημα το 2005, από τους Goulet και Pigeon, οι οποίοι έδωσαν στον κόσμο ένα πακέτο προγραμματισμού ικανό να επιλύει σύνθετα αναλογιστικά προβλήματα με προσομοίωση. Παρόμοια πακέτα φαίνεται να έχουν αντίστοιχη χρησιμότητα, όπως για παράδειγμα είναι το copula (Yan 2007[17]), Rmetrics (Wuertz 2007) και SuppDists (Wheeler 2008), όμως το actuar φαίνεται να εξυπηρετεί ένα πιο ευρύ φάσμα αναλογιστικών λειτουργιών και δεδομένων. Το πακέτο δημοσιεύθηκε στην επίσημη σελίδα CRAN(https://cran.r-project.org/) τον Φεβρουάριο του 2006 και η τροποποίησή του συνεχίζει από τότε.

Το actuar (C. Dutang, V. Goulet, M. Pigeon, 2008[12]) προσφέρει μια επιπλέον βοήθεια στην αναλογιστική επιστήμη. Παρά το γεγονός ότι αρκετά πακέτα στο Comprehensive R Archive Network (CRAN) παρέχουν λειτουργίες χρήσιμες προς τους αναλογιστές, το actuar κατέχει κυρίαρχη θέση στον αναλογιστικό κλάδο λόγω του μεγάλου πλήθους λειτουργιών και εφαρμογών του.

Η παρούσα έκδοση του εν λόγω πακέτου είναι η 2.3-1 με τίτλο Actuarial Functions and Heavy Tailed Distributions η οποία αναθεωρήθηκε τον Μάρτιο του 2018. Το πακέτο περιλαμβάνει λειτουργίες και ένα σύνολο δεδομένων με εφαρμογή στην αναλογιστική επιστήμη, όπως μεθόδους προσομοίωσης για σύνθετες κατανομές, διακεκριμένα μοντέλα

και μείξεις ιεραρχικών μοντέλων στα πλαίσια της θεωρίας αξιοπιστίας. Επιπλέον, παρέχει ένα ευρύ φάσμα εφαρμογών σε πολλές συναρτήσεις πιθανοτήτων για τον υπολογισμό των ποσών ασφαλιστικής ζημιάς και των κατανομών απώλειας, 19 συνεχείς κατανομές με βαριά ουρά, την Poisson - Inverse Gaussian κατανομή, περικομμένη και τροποποιημένη στο μηδέν κατανομή. Στην ενημέρωση του πακέτου συνέβαλαν οι Vincent Goulet, Sebastien Auclair, Christophe Dutang, Xavier Milhaud, Tommy Ouellet, Alexandre Parent, Mathieu Pigeon, Louis-Philippe Pouliot.
(βλ. https://cran.r-project.org/web/packages/actuar/index.html)

4.2 Θεωρία κινδύνου

Η θεωρία κινδύνου αποτελείται από ένα σύνολο τεχνικών προσαρμογής και μέτρησης των κινδύνων που σχετίζονται με ένα χαρτοφυλάκιο ασφαλιστηρίων συμβολαίων. Σε πρώτο στάδιο κύριος στόχος της είναι η εύρεση της κατανομής των συνολικών απαιτήσεων σε ένα συγκεκριμένο χρονικό διάστημα χρησιμοποιώντας το κλασικό συλλογικό μοντέλο της θεωρίας κινδύνου. Στη συνέχεια, αυτό που έχει σημασία για τον αναλογιστή είναι η ανέλιξη του πλεονάσματος της ασφαλιστικής εταιρίας σε πολλές διαφορετικές χρονικές στιγμές. Στην θεωρία χρεοκοπίας, το ενδιαφέρον στρέφεται στην πιθανότητα το πλεόνασμα αυτό να είναι αρνητικό, που κατά συνέπεια οδηγεί σε τεχνική χρεοκοπία του χαρτοφυλακίου της ασφαλιστικής.

Το πακέτο actuar διαθέτει τέσσερεις λειτουργίες σχετικά με τους παραπάνω τομείς, δύο για την εύρεση της κατανομής των συνολικών απαιτήσεων και δύο για τον υπολογισμό της πιθανότητας χρεοκοπίας του εν λόγω χαρτοφυλακίου της ασφαλιστικής εταιρίας.

4.2.1 Διακριτοποίηση της κατανομής των συνολικών ζημιών

Σε ορισμένες περιπτώσεις, ο υπολογισμός του συνολικού ποσού ζημιάς καθιστά αναγκαία την διακριτοποίηση των συνεχών κατανομών των ποσών απαίτησης. Το πακέτο actuar διαθέτει την λειτουργία *discretize* ή *discritise*, η οποία διακριτοποιεί τις συνεχείς κατανομές. Η εν λόγω λειτουργία παρέχει επιπλέον και μια υποστήριξη στις ήδη διακριτές κατανομές. Με την διακριτοποίηση αυτό που επιτυγχάνουμε είναι να μετατρέψουμε μία συνεχή κατανομή σε διακριτή χωρίζοντας το πεδίο τιμών σε πολλά ίσα τμήματα.

Έστω ότι $F(x)$, δηλώνει την συνάρτηση κατανομής (cdf) της τ.μ που χρήζει διακριτοποίησης σε κάποιο διάστημα (a,b) το οποίο ορίζεται στις μεταβλητές from και to αντίστοιχα, καθώς και το βήμα που θα χωριστεί το επιλεγμένο τμήμα της σ.κ.[14]. Όσο πιο μικρό βήμα επιλεχθεί 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. Η λειτουργία *discretize* διαθέτει τις τέσσερις ακόλουθες μεθόδους διακριτοποίησης:

1. Ανώτερη (upper)διακριτοποίηση της $F(x)$:

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

   για $x = \alpha, \alpha + h, ..., b - h$.
   Η διακριτοποίηση της κατανομής είναι πάντα πάνω από την πραγματική κατανομή.

2. Κατώτερη (lower)διακριτοποίηση της $F(x)$:

$$f_x = \begin{cases} F(\alpha), & \text{x=} \ \alpha, \\ F(x) - F(x-h), & \text{x=} \ \alpha + h, ..., b. \end{cases}$$

   Η διακριτοποίηση της κατανομής είναι πάντα κάτω από την πραγματική τιμή της κατανομής.

3. Στρογγυλοποίηση (rounding)της τ.μ.:

$$f_x = \begin{cases} F(\alpha + \frac{h}{2}), & \text{x=} \ \alpha, \\ F(x + \frac{h}{2}) - F(x - \frac{h}{2}), & \text{x=} \ \alpha + h, ..., b - h \ . \end{cases}$$

Η πραγματική τιμή της κατανομής περνάει στα μέσα του βήματος της διακριτοποιημένης κατανομής.

4. Αμερόληπτη διακριτοποίηση (unbiased) ή μέθοδος της πρώτης ροπής:

$$f_x = \begin{cases} \dfrac{E[X \wedge \alpha] - E[X \wedge \alpha + h]}{h} + 1 - F(\alpha), & \text{x}= \alpha, \\ \dfrac{2E[X \wedge x] - E[X \wedge x - h] - E[X \wedge x + h]}{h}, & \alpha < x < b \ , \\ \dfrac{E[X \wedge b] - E[X \wedge b - h]}{h} - 1 + F(b), & \text{x}= b. \end{cases}$$

Η διακριτοποιημένη και η πραγματική κατανομή έχουν την ίδια πιθανότητα και αναμενόμενη τιμή στο διάστημα (α,b).

Στο περιβάλλον της R η εν λόγω λειτουργία ορίζεται ως εξής:

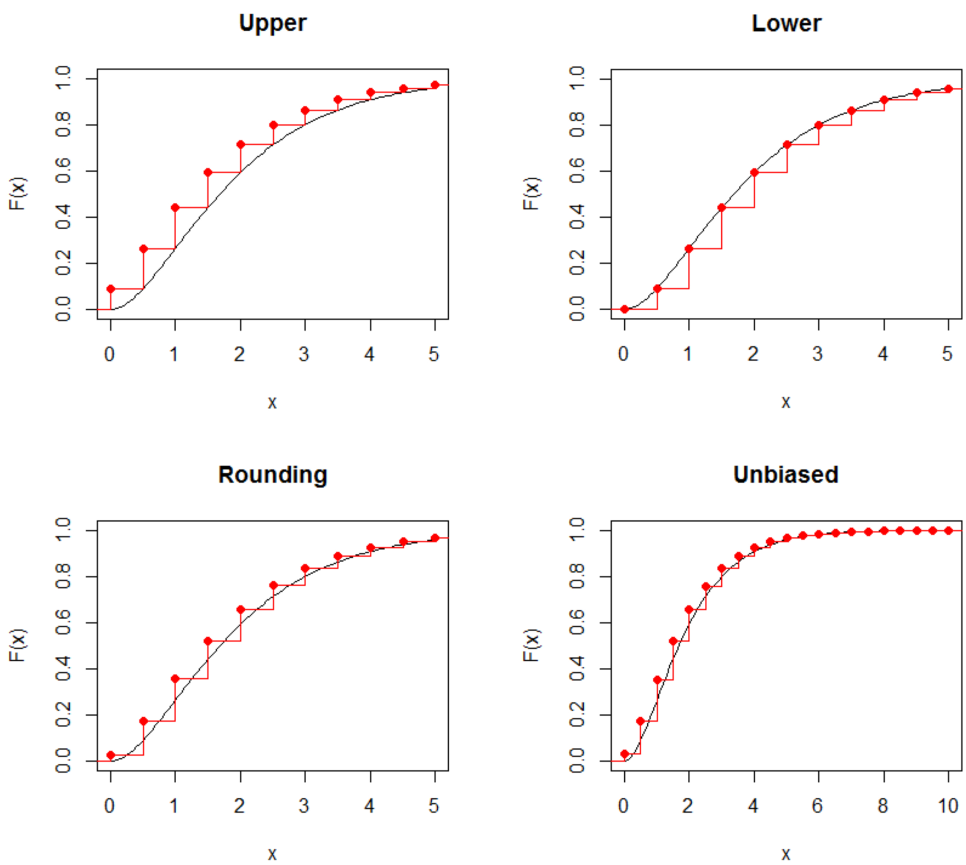
```
discretize(cdf, from, to, step = 1,
method = c("upper", "lower", "rounding", "unbiased"),
lev, by = step, xlim = NULL)
```

(βλ. ΠΑΡΑΡΤΗΜΑ Γ)

Για παράδειγμα, χρησιμοποιώντας τις παραπάνω μεθόδους θα κάνουμε διακριτοποίηση σε μια κατανομή Γάμμα(2,1) στο διάστημα (0,10) με βήμα 0.5. (βλ. Αντζουλάκος, 2016) [1].

```
library(actuar)
a <- 0; b <- 10; h <- 0.5
f1<-discretise(pgamma(x,2,1),method="upper",from=a,to=b,step=h)
f2<-discretise(pgamma(x,2,1),method="lower",from=a,to=b,step=h)
f3<-discretise(pgamma(x,2,1),method="rounding",from=a,to=b,
               step=h)
f4<-discretise(pgamma(x,2,1),method="unbiased",from=a,to=b,
               lev=levgamma(x,2,1),step=h)
m <- cbind(f1[1:10],f2[1:10],f3[1:10],f4[1:20])
colnames(m) <- c("Upper","Lower","Rounding","Unbiased")
round(m,digits=5)


        Upper     Lower  Rounding  Unbiased
[1,]  0.09020   0.00000   0.02650   0.03265
[2,]  0.17404   0.09020   0.14686   0.14197
[3,]  0.17793   0.17404   0.18201   0.18001
[4,]  0.15182   0.17793   0.16676   0.16614
[5,]  0.11871   0.15182   0.13533   0.13531
[6,]  0.08815   0.11871   0.10282   0.10302
[7,]  0.06326   0.08815   0.07494   0.07519
[8,]  0.04431   0.06326   0.05308   0.05332
[9,]  0.03048   0.04431   0.03682   0.03701
[10,] 0.02067   0.03048   0.02514   0.02529
[11,] 0.09020   0.00000   0.02650   0.01706
[12,] 0.17404   0.09020   0.14686   0.01139
[13,] 0.17793   0.17404   0.18201   0.00754
[14,] 0.15182   0.17793   0.16676   0.00496
[15,] 0.11871   0.15182   0.13533   0.00324
[16,] 0.08815   0.11871   0.10282   0.00211
[17,] 0.06326   0.08815   0.07494   0.00136
[18,] 0.04431   0.06326   0.05308   0.00088
```

```
[19,] 0.03048    0.04431    0.03682    0.00056
[20,] 0.02067    0.03048    0.02514    0.00036
```

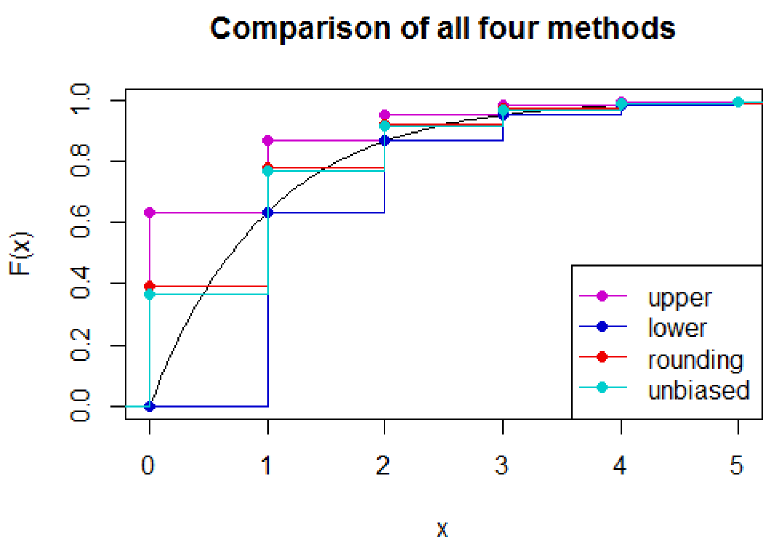


Όπως φαίνεται και στα ανωτέρω διαγράμματα, η "upper και "lower" διακριτοποίηση φαίνεται να βρίσκεται πάνω και κάτω αντίστοιχα από την πραγματική τιμή της συνάρτηση κατανομής της τυχαίας μεταβλητής. Αντίστοιχα, η "rounding" διατρέχει κατά μήκος την καμπύλη, ενώ η "unbiased" φαίνεται να έχει την ίδια πιθανότητα και αναμενόμενη τιμή.

Στο παρακάτω παράδειγμα, παρουσιάζεται η γραφική παράσταση και των τεσσάρων μεθόδων διακριτοποίησης σε ένα διάγραμμα, με την χρήση μιας εκθετικής κατανομής στο διάστημα (0,5).

```
library(actuar)
a<-0;b<-5
fu <- discretize(pexp(x), method = "upper", from = a, to = b)
```

```
fl <- discretize(pexp(x), method = "lower", from = a, to = b)
fr <- discretize(pexp(x), method = "rounding", from = a, to = b)
fb <- discretize(pexp(x), method = "unbiased", from = a, to = b,
                 lev = levexp(x))
curve(pexp(x), from = 0, to = 5,col="black",ylab="F(x)",
      main="Comparison␣of␣all␣four␣methods")
plot(stepfun(0:4,diffinv(fu)),pch=19,add=TRUE,col="magenta3")
plot(stepfun(0:5,diffinv(fl)),pch=19,add=TRUE,col="blue3")
plot(stepfun(0:4,diffinv(fr)),pch=19,add=TRUE,col="red2")
plot(stepfun(0:5,diffinv(fb)),pch=19,add=TRUE,col="cyan3")
legend("bottomright", legend = c("upper","lower","rounding",
       "unbiased"),col = c("magenta3", "blue3", "red2",
       "cyan3"), lty = 1, pch = 19, text.col = "black")
par(col = "black")
```



Όσο αφορά την διακριτοποίηση των συνεχών τ.μ. και κατ' επέκταση του υπολογισμού της σ.κ. της τ.μ. S, προκύπτουν κάποια συμπεράσματα. Αρχικά παρατηρείται ότι σημαντικό ρόλο στην διακριτοποίηση έχει το βήμα που θα επιλεγεί καθώς όσο πιο μικρό είναι τόσο μικρότερες διαφορές προκύπτουν μεταξύ της διακριτοποιημένης κατανομής με την πραγματική. Όσο αφορά την μέθοδο διακριτοποίησης που θα χρησιμοποιηθεί, έχει παρατηρηθεί ότι οι μέθοδοι Rounding και Unbiased προσεγγίζουν καλύτερα την συνεχή κατανομή, όμως η βέλτιστη μέθοδος διακριτοποίησης είναι η Upper καθώς κάνοντας αλλαγή στις παραμέτρους της κατανομής έχει πάντοτε τις μικρότερες αποκλίσεις. Τέλος, ένα σημαντικό αποτέλεσμα που προκύπτει είναι ότι σε κάθε περίπτωση η πραγματική τιμή της σ.κ. της τ.μ. S είναι μεγαλύτερη από την προκύπτουσα κατόπιν της διακριτοποίησης[19].

Παρ' όλα αυτά, στην παρούσα διπλωματική, η μέθοδος που θα χρησιμοποιηθεί είναι η Lower, καθώς δίνει το κατώτατο διακριτοποιημένο σημείο της κατανομής που δεν είναι άλλο από το σημείο μηδέν, στο οποίο η μεικτή κατανομή έχει μάζα πιθανότητας.

4.2.2 Υπολογισμός της κατανομής του συνολικού ποσού απαίτησης

Μια από τις σημαντικότερες εντολές του πακέτου είναι αδιαμφισβήτητα η *aggregateDist*. Η εν λόγω εντολή υπολογίζει το σωρευτικό ποσό της αθροιστικής κατανομής ενός χαρτοφυλακίου σε ένα ορισμένο χρονικό διάστημα, χρησιμοποιώντας μία από τις, ως επί το πλείστον, πέντε μεθόδους (κανονική, αναδρομική, συνέλιξη, προσομοίωση, npower).

Ας δούμε όμως πιο αναλυτικά τις μεθόδους αυτές:

1. *Αναδρομική μέθοδος υπολογισμού (Recursive mentod)*, χρησιμοποιώντας τον αναδρομικό τύπο του Panjer [16]. Αυτό απαιτεί η X να είναι μη-αρνητική ακέραια τ.μ., δηλαδή X∈{0,1,2,...} και η τ.μ. N να ακολουθεί μια διακριτή κατανομή. Προφανώς, επειδή υπάρχουν πάρα πολλές επιλογές για την κατανομή της τ.μ. N (για να μοντελοποιήσουμε το πλήθος των κινδύνων ή ζημιών ή και απαιτήσεων του χαρτοφυλακίου), επομένως, είναι λογικό να θεωρούμε οικογένειες κατανομών για την τ.μ. N τέτοιες ώστε να έχουν μια συγκεκριμένη ιδιότητα. Συνεπώς αυτό που μας ενδιαφέρει για την εν λόγω μέθοδο είναι να εξετάσουμε τις περιπτώσεις εκείνες που ανήκουν στις κλάσεις $\mathscr{R}(\alpha, b, 0)$ και $\mathscr{R}(\alpha, b, 1)$ [14]. Αυτές οι κλάσεις περιλαμβάνουν τις κατανομές Poisson, Διωνυμική, Αρνητική Διωνυμική, Λογαριθμική, και τις τροποποιημένες και αποκομμένες στο μηδέν κατανομές τους που επιτρέπουν ύπαρξη μάζας πιθανότητας στο σημείο μηδέν. Ο γενικός τύπος είναι:

$$f_s(x) = \frac{(p_1 - (\alpha+b)p_0)fx(x) + \sum_{y=1}^{min(x,m)}(a+\frac{by}{x})fx(y)f_s(x-y)}{1-afx(0)}.$$

Η παραπάνω σχέση ξεκινάει από $f_s(0) = P_N(fx(0))$ με $P_N(.)$ να είναι η πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής N.

2. Ακριβής υπολογισμός με την *μέθοδο των συνελίξεων (Convolution method)*, χρησιμοποιώντας το Θεώρημα 3.1. Εξίσου, η εν λόγω μέθοδος απαιτεί η τυχαία μεταβλητή N να ακολουθεί μια διακριτή κατανομή. Το πακέτο απλώς κάνει εφαρμογή της σχέσης 3.2, καθώς οι συνελίξεις απλά υπολογίζονται βάση του μετασχηματισμού Fast Fourier (Fast Fourier Transform)[13]. Παρόλα αυτά, ακόμα και με την χρήση της R, οι υπολογισμοί μπορούν να γίνουν για έναν πεπερασμένο το πλήθος πράξεων, ακόμα και για τους σημερινούς υπολογιστές.

3. *Κανονική προσέγγιση (Normal approximation method)* της κατανομής, η οποία ορίζεται ως:
$$F_S(x) \approx \Phi\left(\frac{x-\mu_S}{\sigma_S}\right)$$
όπου $\mu_s = E[S]$ και $\sigma_s^2 = Var[S]$.

4. *Normal Power II method*, η οποία χρησιμοποιεί την προσέγγιση:
$$F_S(x) \approx \Phi\left(-\frac{3}{\gamma_S} + \sqrt{\frac{9}{\gamma_S^2} + 1 + \frac{6}{\gamma_S}\frac{x-\mu_S}{\sigma_s}}\right),$$
όπου $\gamma_S = \dfrac{E[(S-\mu_S)^3]}{\sigma_S^{3/2}}$ είναι ο συντελεστής ασυμμετρίας της τ.μ. S. Η προσέγγιση ισχύει μόνο όταν $x > \mu_S$ και είναι ικανοποιητική όταν $\gamma_S < 1$ [10] (βλ. Παράδειγμα 5.12).

5. *Μέθοδος προσομοίωσης (Simulation method)* ενός τυχαίου δείγματος της S και προσέγγιση της $F_S(x)$ από την εμπειρική συνάρτηση κατανομής
$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I\{x_j \le x\}.$$
Η μέθοδος της προσομοίωσης μπορεί να γίνει και με την εντολή *simul*.

Σε αυτό το σημείο, να αναφερθεί ότι, η προσθήκη περαιτέρω εντολών στην λειτουργία *aggregateDist* είναι εφικτή, λόγω της απλής αρχικής της μορφής. Τα διαφορετικά αποτελέσματα που προκύπτουν οφείλονται κατά κύριο λόγο στην μέθοδο υπολογισμού που θα χρησιμοποιηθεί.

Επιπρόσθετα, η αναδρομική μέθοδος παρουσιάζει σφάλμα, όταν ο αναμενόμενος αριθμός απαιτήσεων είναι τόσο μεγάλος, ώστε το $f_S(0)$ είναι αριθμητικά ίσο με το μηδέν. Κατά τους Klugman et al.[14], μια εφικτή λύση για το ανωτέρω πρόβλημα αποδείχτηκε η διαίρεση της κατάλληλης παραμέτρου κατά $2n$, τέτοιο ώστε $f_S(0) > 0$ και οι επαναλήψεις να μπορούν να ξεκινήσουν.

Ένα επιπλέον πρόβλημα της αναδρομικής μεθόδου είναι η αδυναμία επίτευξης μιας αθροιστικής συνάρτησης κατανομής που να φτάνει κοντά στην μονάδα. Αυτό κυρίως οφείλεται σε μια γενικευμένη διακριτοποίηση της κατανομής της τυχαίας μεταβλητής N. Μια απλή επίλυση του προβλήματος είναι η χρήση αρκετά μικρού βήματος υπολογισμού διακριτοποίησης και διακριτοποίηση της κατανομής βαρύτητας στην δεξιά ουρά της.

```
aggregateDist(method = c("recursive", "convolution", "normal",
              "npower", "simulation"), model.freq = NULL,
              model.sev = NULL, p0 = NULL,x.scale = 1, convolve = 0,
              moments, nb.simul, ...,tol = 1e-06, maxit = 500,
              echo = FALSE)

## S3 method for class 'aggregateDist'
plot(x, xlim, ylab = expression(F[S](x)),
main = "Aggregate␣Claim␣Amount␣Distribution",
sub = comment(x), ...)
```

(βλ. ΠΑΡΑΡΤΗΜΑ Γ)

ΠΑΡΑΔΕΙΓΜΑ 4.1 *Έστω ότι:*

x	0	1	2	3	4	5	6	7	8	9	10
$f_x = P(X = x)$	0	0.153	0.23	0.235	0.1235	0.0735	0.053	0.053	0.053	0.0253	0.0253

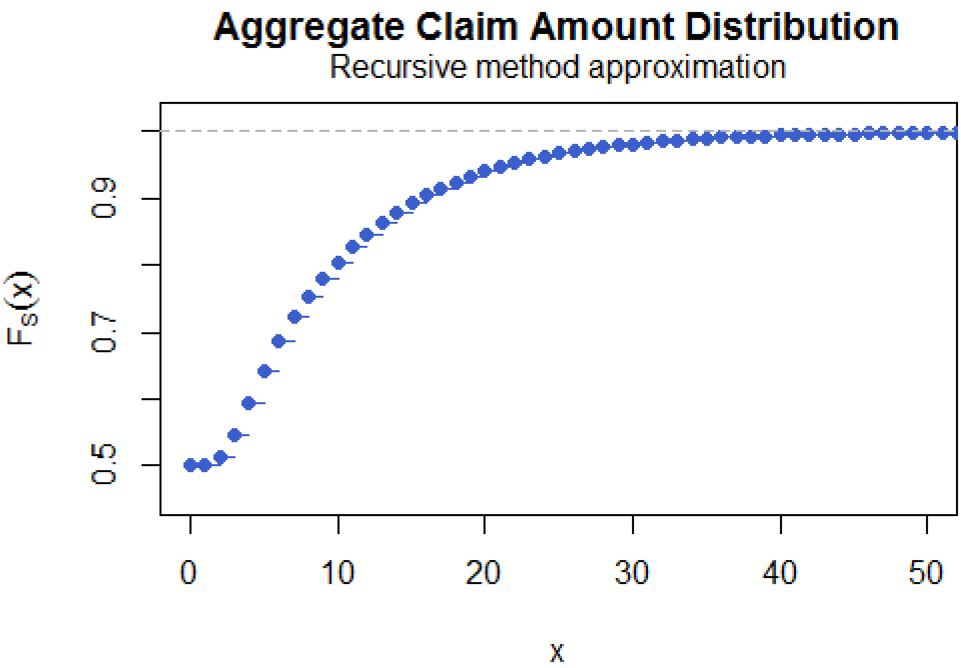
n	0	1	2	3	4	5	6	7	8
$f_n = P(N = n)$	0.0253	0.123	0.1253	0.223	0.2253	0.1253	0.0263	0.0323	0.0213

Ζητούμενο είναι η συνάρτηση πιθανότητας $g_s(x)$ του συνολικού ποσού ζημιών, με χρήση προσεγγιστικής μεθόδου. Στην παρούσα φάση θα χρησιμοποιηθεί η μέθοδος των συνελίξεων.

```
library(actuar)
fx <- c(0, 0.153, 0.23, 0.235, 0.1235, 0.0735, 0.053, 0.053,
        0.053, 0.0253, 0.0253)
fn <- c(0.0253, 0.123, 0.1253, 0.223, 0.2253, 0.1253, 0.0263,
        0.0323, 0.0213)
Gs <- aggregateDist("convolution",model.freq = fn,model.sev=fx)
mode(Gs) ; x <- 0:10 ; g <-diff(Gs)[1:11]
mat<-cbind(g,Gs(x)) ; colnames(mat) <- c("gs(x)","Gs(x)")
round(mat,digits=5)

        gs(x)     Gs(x)
[1,]   0.02530   0.02530
[2,]   0.01882   0.04412
[3,]   0.03122   0.07534
[4,]   0.03852   0.11386
[5,]   0.03455   0.14842
[6,]   0.03717   0.18559
[7,]   0.04160   0.22719
[8,]   0.04681   0.27400
[9,]   0.05140   0.32540
[10,]  0.05200   0.37740
[11,]  0.05459   0.43198
```

ΠΑΡΑΔΕΙΓΜΑ 4.2 *Έστω ότι η N ∼Geom(1/2) και η τυχαία μεταβλητή $X \sim Gamma(5,2)$. Για την προσέγγιση της κατανομής της S αρχικά θα πρέπει να γίνει διακριτοποίηση της Gamma στο διάστημα (0,50) με την μέθοδο "rounding" και βήμα 0.5 και στην συνέχεια μέσω του αναδρομικού τύπου της λειτουργίας aggregateDist:*

```
library(actuar)
fx <- discretize(pgamma(x,5, 2), method = "lower", from = 0,
to = 50, step = 0.5)
Fs <- aggregateDist("recursive", model.freq ="geometric", model.sev=fx,
                    prob=0.5)
x<-0:10 ; fs<- diff(Fs)[1:11]
mat<-cbind(Fs(x),fs) ; colnames(mat) <- c("Fs(x)","fs(x)")
round(mat,digits = 5)

plot(Fs,xlim=c(0,40),ylim=c(0.45,1.02),col="royalblue3")

        Fs(x)   fs(x)
[1,]  0.50000  0.50000
[2,]  0.50091  0.00091
[3,]  0.51316  0.01225
[4,]  0.54623  0.03307
[5,]  0.59326  0.04703
[6,]  0.64213  0.04888
[7,]  0.68565  0.04351
[8,]  0.72235  0.03671
[9,]  0.75355  0.03120
[10,] 0.78083  0.02727
[11,] 0.80513  0.02430
```



□

4.3 Προσομοίωση ασφαλιστικών δεδομένων

Το πακέτο actuar παρέχει διάφορες λειτουργίες για την διευκόλυνση της δημιουργίας τυχαίων παραλλαγών όσο αφορά τις κατανομές των απαιτήσεων και το ύψος ζημιάς. Χρησιμοποιώντας διάφορα μοντέλα πιθανοτήτων με εφαρμογή στην θεωρία κινδύνου είναι ικανό να παράγει αποτελέσματα χρησιμοποιώντας τις εξής λειτουργίες:

- rmixture : για την προσομοίωση από διακριτές μείξεις κατανομών,
- rcompound : για την προσομοίωση σύνθετων μοντέλων (μια πιο απλοποιημένη έκδοση είναι η rcompois για προσομοίωση απλών σύνθετων Poisson κατανομών) και
- rcomphierarc : για την προσομοίωση σύνθετων μοντέλων όπου ο αριθμός των κινδύνων και το μέγεθος της ζημιάς έχουν μια ιεραρχική δομή.

4.3.1 Προσομοίωση από διακριτές μείξεις κατανομών

Μια διακριτή μείξη κατανομών θεωρείται η τυχαία μεταβλητή που αποτελείται από έναν πεπερασμένο αριθμό τυχαίων μεταβλητών και η συνάρτηση πιθανότητας είναι της μορφής

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + ... + p_n f_n(x) = \sum_{i=1}^{n} p_i f_i(x),$$

όπου $pi \geq 0$ είναι πιθανότητες με $\sum_{i=1}^{n} p_i = 1.$

Στο περιβάλλον της R η εν λόγω λειτουργία ορίζεται ως εξής:

```
rmixture(n, probs, models)
```

(βλ. ΠΑΡΑΡΤΗΜΑ Γ)

ΠΑΡΑΔΕΙΓΜΑ 4.3 ***(Μείξη εκθετικών κατανομών)***
Έστω ότι η τ.μ. X έχει συνάρτηση πιθανότητας

$$f(x) = \frac{2}{5}(2e^{-2x}) + \frac{3}{5}(5e^{-5x}) = \frac{4}{5}e^{-2x} + 3e^{-5x}.$$

Η παρακάτω λειτουργία παράγει 5 τ.μ. από την εν λόγω συνάρτηση

```
library(actuar)
rmixture(5, probs = c(2, 3),models = expression(rexp(2), rexp(5)))

[1] 0.25682159  0.12781334  0.38321094  0.14078915  0.09902185
```

□

4.3.2 Προσομοίωση σύνθετων μοντέλων

Σε πολλές περιπτώσεις οι αναλογιστές καλούνται να υπολογίσουν πιο σύνθετα μοντέλα συλλογικού προτύπου. Στην προκειμένη περίπτωση, το πακέτο actuar διαθέτει την λειτουργία rcompound η οποία παράγει παραλλαγές από την τυχαία μεταβλητή S όταν η κατανομή των δύο τυχαίων μεταβλητών N και X δεν είναι ιεραρχική.

Επιπλέον, υπάρχει η λειτουργία rcomppois που αποτελεί μια απλοποιημένη εκδοχή της παραπάνω εντολής και χρησιμοποιείται όταν η τ.μ. N έχει την Poisson κατανομή και επομένως η τ.μ. S έχει την σύνθετη Poisson. Στην προκειμένη περίπτωση, το όρισμα model.freq αντικαθίσταται από την παράμετρο λάμδα της Poisson.

Η παραπάνω λειτουργίες ορίζονται ως εξής:

```
rcompound(n, model.freq, model.sev, SIMPLIFY = TRUE)

rcomppois(n, lambda, model.sev, SIMPLIFY = TRUE)
```

(βλ. ΠΑΡΑΡΤΗΜΑ Γ)

Στην συνέχεια παρουσιάζονται κάποια παραδείγματα για καλύτερη κατανόηση της προαναφερθείσας λειτουργίας.

ΠΑΡΑΔΕΙΓΜΑ 4.4 ***(Σύνθετη Poisson-Γάμμα κατανομή)***
Έστω ότι η τυχαία μεταβλητή $N \sim Poisson(2)$ και η τ.μ. $X \sim G(2,1)$. Τότε οι παραλλαγές της τ.μ. S για $n = 5$ που προκύπτουν με την χρήση των εντολών rcompound και rcompois είναι

```
library(actuar)
rcompound(n=5, model.freq=rpois(2), model.sev=rgamma(2,1),
          SIMPLIFY=TRUE)

[1] 6.4687940  0.7879747  2.2982099  3.4508065  4.5807660

rcomppois(n=5, lambda=2, model.sev=rgamma(2,1), SIMPLIFY=TRUE)

[1] 7.7434561  2.2554002  0.0000000  6.5848813  0.3900632
```

ΠΑΡΑΔΕΙΓΜΑ 4.5 *Έστω η τ.μ. S_1, με $N_1 \sim P(2)$ και $X_2 \sim G(3,1)$ και η τ.μ. S_2, με $N_2 \sim P(1/2)$ και $X_2 \sim G(5,3)$. Για $\lambda = \lambda_1 + \lambda_2 = 2 + 1/2 = 3/2$, ο υπολογισμός της τ.μ. $S = S_1 + S_2$ υπολογίζεται εύκολα ως εξής:*

```
library(actuar)
rcomppois(n=5, lambda=3/2, rmixture(probs=c(2,1/2),
          expression(rgamma(4,1),rgamma(5, 3)))

[1]  9.763473   2.658197   10.831954   0.000000   3.249116
```

4.3.3 Προσομοίωση σύνθετων ιεραρχικών μοντέλων

Τα ιεραρχικά μοντέλα χρησιμοποιούνται για δεδομένα που παρουσιάζουν μια πολυεπίπεδη μορφή στην δομή τους. Το κύριο χαρακτηριστικό τους είναι η ύπαρξη ενός κανόνα πιθανότητας σε κάθε επίπεδο της ταξινόμησής τους το οποίο εξαρτάται από κάποιο προηγούμενο αποτέλεσμα. Μια απλή μορφή ιεραρχικού μοντέλου είναι το κάτωθι:

$$
\begin{aligned}
X_t|\Lambda,\Theta &\sim Poisson(\Lambda), \\
\Lambda|\Theta &\sim Gamma(3,\Theta), \\
\Theta &\sim Gamma(2,2).
\end{aligned}
$$

Στην περίπτωση της προσομοίωσης ιεραρχικών μοντέλων με χρήση της γλώσσα προγραμματισμού R πρέπει να οριστούν τα εν λόγω μοντέλα με τρόπο τέτοιο ώστε να ανταποκρίνονται στα παρακάτω κριτήρια:

1. Πρέπει να είναι απλό και σύμφωνα με τα πρότυπα της εν λόγω έκδοσης της R και προγενέστερων.

2. Να επιτρέπει κάθε επίπεδο αριθμών και κόμβων.

3. Ανεξάρτητα από το επίπεδο να επιτρέπεται η χρήση παραμέτρων υψηλότερων από την ιεραρχική δομή.

(βλ. https://cran.r-project.org/web/packages/actuar/vignettes/simulation.pdf)

Τα ιεραρχικά μοντέλα επιτυγχάνονται μέσω αντικατάστασης μια ή περισσότερων παραμέτρων μιας κατανομής σε ένα δεδομένο επίπεδο από οποιονδήποτε συνδυασμό των παραπάνω επιπέδων. Η εν λόγω λειτουργία μπορεί να προσομοιώσει δεδομένα στην περίπτωση όπου οι τ.μ. N και X είναι ιεραρχικές και να επιστρέψει αποτελέσματα στην μορφή ενός ενιαίου χαρτοφυλακίου σύμφωνα με τα συμβόλαια και τα έτη που δόθηκαν, όπως παρουσιάζεται στο παράδειγμα που έπεται. Η μορφή της στο περιβάλλον της R, είναι:

```
rcomphierarc(nodes,model.freq = NULL,model.sev = NULL,weights = NULL)
```

ΠΑΡΑΔΕΙΓΜΑ 4.6 *Έστω το παρακάτω μοντέλο συλλογικού κινδύνου*

$$S_{ijt} = X_{ijt1} + ... + X_{ijtN_{ijt}},$$

όπου $i = 1, ..., I,\ \ j = 1, ..., J_1,\ \ t = 1, ..., n_{ij}$ *και*

$$\begin{aligned} N_{ijt}|\Lambda_{ij}, K_i &\sim Poisson(w_{ijt}\Lambda_{ij}) \\ \Lambda_{ij}|K_i &\sim Gamma(K_i, 1) \\ K_i &\sim Exp(3) \end{aligned}$$

και

$$\begin{aligned} X_{ijtu}|\Theta_{ij}, P_i &\sim Lognormal(\Theta_{ij}, 1) \\ \Theta_{ij}|P_i &\sim N(P_i, 1) \\ P_i &\sim N(2, 0.5). \end{aligned}$$

Για $I = 2,\ \ J_1 = 3,\ \ J_2 = 4$ *και* $n_{11} = ... = n_{13} = 4,\ \ n_{21} = ... = n_{24} = 5$ *και τα βάρη* $w_{ij} \sim U(0.5, 2.5)$ *είναι:*

```
library(actuar)
nodes <- list(cohort = 2,contract = c(3, 4),
              year = c(4, 4, 4, 5, 5, 5, 5))
mf <- expression(cohort = rexp(3),contract = rgamma(cohort, 1),
                 year = rpois(weights * contract))
ms <- expression(cohort = rnorm(2, sqrt(0.5)),contract =
                 rnorm(cohort, 1),
year = rlnorm(contract, 1))
wijt <- runif(32, 0.5, 2.5)
pf <- rcomphierarc(nodes=nodes,model.freq=mf,model.sev=ms,weights=wijt)
```

```
s<-aggregate(pf)
round(s,digits=3)

     cohort contract year.1  year.2  year.3 year.4  year.5
[1,]      1        1  0.000   0.000   0.000  0.000      NA
[2,]      1        2  0.000   3.597   1.759  0.000      NA
[3,]      1        3 88.654  52.034 176.965 72.433      NA
[4,]      2        1 19.242   1.562   4.277  0.000   1.265
[5,]      2        2  4.298   0.000   0.000 36.136   0.000
[6,]      2        3  0.000 166.534  23.656  0.000   0.000
[7,]      2        4  0.000   0.000  12.114  0.000 122.893
```

Κεφάλαιο 5

ΑΝΑΛΥΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΥΠΟΛΟΓΙΣΜΟΥ ΣΥΝΘΕΤΩΝ ΚΑΤΑΝΟΜΩΝ

Στο κεφάλαιο που έπεται, θα γίνει μια εκτενέστερη αναφορά στις διάφορες περιπτώσεις κατανομών που μπορεί να ακολουθεί η τυχαία μεταβλητή S. Πιο συγκεκριμένα, θα παρουσιαστούν διαφορετικές μέθοδοι υπολογισμού της κατανομής των συνολικών ζημιών θεωρώντας ότι οι απαιτήσεις και το πλήθος απαιτήσεων ακολουθούν διάφορες κατανομές.

Αντίστοιχα, για κάθε περίπτωση του εν λόγω κεφαλαίου, πέρα από τον καθιερωμένο τρόπο υπολογισμού των παραδειγμάτων εύρεσης της κατανομής των συνολικών ζημιών, θα γίνεται και μια αντίστοιχη παρουσίαση επίλυσής τους με χρήση του πακέτου actuar. Με αυτό τον τρόπο θα γίνει μια ακριβής και προσεγγιστική παρουσίαση υπολογισμού της κατανομής της τ.μ. S.

Πιο συγκεκριμένα, στην Ενότητα 5.1 θα παρουσιαστεί η περίπτωση της κατανομής γάμμα, η οποία μαζί με την εκθετική κατανομή αποτελούν μια από τις σημαντικότερες κατανομές στο κομμάτι της Θεωρίας Κινδύνου. Στην Ενότητα 5.2, γίνεται μια αναφορά στις σύνθετες Bernoulli κατανομές και μικτού τύπου τυχαίες μεταβλητές. Στην συνέχεια, η Ενότητα 5.3 είναι αφιερωμένη στις σύνθετες Poisson κατανομές, ως μια από τις πιο κλασικές περιπτώσεις σύνθετης κατανομής.

Μια εξίσου σημαντική περίπτωση στον τομέα της Θεωρίας Κινδύνου είναι αυτή όπου η τυχαία μεταβλητή N ακολουθεί της γεωμετρική κατανομή. Έτσι στις Ενότητες 5.4 και

5.5 γίνεται μια εκτενέστερη παρουσίαση στις περιπτώσεις όπου έχουμε μικτές σύνθετες γεωμετρικές κατανομές και σύνθετες διακριτές γεωμετρικές κατανομές αντίστοιχα.

Στην Ενότητα 5.6 παρουσιάζονται οι διάφορες περιπτώσεις των σύνθετων τροποποιημένων και περικομμένων στο μηδέν γεωμετρικών κατανομών και στην Ενότητα 5.7, η περίπτωση της σύνθετης αρνητικής διωνυμικής κατανομής.

Τέλος, στην Ενότητα 5.8 γίνεται μια αναφορά στην περίπτωση του κεντρικού οριακού θεωρήματος και πως αυτό μπορεί να χρησιμοποιηθεί στον κλάδο των ασφαλίσεων και της θεωρίας κινδύνου.

5.1 Η περίπτωση της γάμμα κατανομής

Μια από τις πιο σημαντικές κατανομές στην Θεωρία Κινδύνου αποτελεί η κατανομή γάμμα, όσο και η εκθετική κατανομή, σαν υποπερίπτωση της γάμμα. Στην ενότητα αυτή θα γίνει αναφορά του γενικού τύπου υπολογισμού της κατανομής των συνολικών ζημιών θεωρώντας ότι οι απαιτήσεις ακολουθούν την γάμμα κατανομή, ενώ η τ.μ. N ακολουθεί οποιαδήποτε κατανομή.

Η κατανομή γάμμα ορίζεται ως εξής:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \quad \alpha > 0.$$

ενώ η μη-πλήρης συνάρτηση γάμμα ορίζεται ως:

$$\Gamma(\alpha, x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt, \quad \alpha > 0, \;\; x > 0.$$

Αν $\alpha \in Z^+$, δηλαδή αν $\alpha \in \{1, 2, 3, ...\}$,τότε ισχύει ότι

$$\Gamma(\alpha, x) = 1 - e^{-x} \sum_{j=0}^{\alpha-1} \frac{x^j}{j!}.$$

ΠΡΟΤΑΣΗ 5.1 *Για το μοντέλο συλλογικού κινδύνου*

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

και η τ.μ. $N \in \{0, 1, 2, ...\}$.

(α) *Έστω η τ.μ.* $X \sim G(\alpha, \lambda), \;\; \alpha > 0, \;\; \lambda > 0.$

- *Αν* $\alpha \notin Z^+$*, τότε*

$$\bar{G}(x) = 1 - p_0 - \sum_{n=1}^{\infty} p_n \Gamma(\alpha n, \lambda x), \quad x \geq 0,$$
$$g(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \frac{\lambda^{\alpha n}}{\Gamma(\alpha n)} x^{\alpha n - 1}, \quad x > 0.$$

- *Αν* $\alpha \in Z^+$*, (οπότε η τ.μ.* $X \sim Erl(\alpha, \lambda)$*) τότε*

$$\bar{G}(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \sum_{j=0}^{\alpha n - 1} \frac{(\lambda x)^j}{j!}, \quad x \geq 0,$$
$$g(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \frac{\lambda^{\alpha n}}{(\alpha n - 1)!} x^{\alpha n - 1}, \quad x > 0.$$

(β) *Έστω ότι η τ.μ.* $X \sim Exp(\lambda), \;\; \lambda > 0.$ *Τότε*

$$\bar{G}(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \sum_{j=0}^{n-1} \frac{(\lambda x)^j}{j!}, \quad x \geq 0,$$
$$g(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \frac{\lambda^n}{(n-1)!} x^{n-1}, \quad x > 0.$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Κάποια απλά παραδείγματα εφαρμογής της Πρότασης 5.1 παρουσιάζονται στην συνέχεια.

(A) *Αν η τ.μ N δεν είναι φραγμένη*, δηλαδή αν $N \in \{0,1,2,...\}$, τότε σύμφωνα με την Πρόταση 5.1 πρέπει να υπολογίσουμε ένα άπειρο πλήθος όρων καθώς το n παίρνει τιμές στο διάστημα $[1,\infty)$.

ΠΑΡΑΔΕΙΓΜΑ 5.1 ***(Σύνθετη Poisson-εκθετική)***
Αν η τ.μ. $N \sim P(\mu)$ και η τ.μ. $X \sim Exp(\lambda)$, με $\mu = 2, \lambda = 2$ τότε η συνάρτηση κατανομής της τ.μ. S δίνεται από την σχέση

$$G(x) = 1 - e^{-(\lambda x+\mu)} \sum_{n=1}^{\infty} \frac{mu^n}{n!} \sum_{j=0}^{n-1} \frac{(\lambda x)^j}{j!}, \quad x \geq 0$$

και η συνάρτηση πυκνότητας πιθανότητας, είναι

$$g(x) = \begin{cases} e^{-\mu}, & x = 0, \\ e^{-(\lambda x+\mu)} \sum_{n=1}^{\infty} \frac{(\lambda\mu)^n}{n!(n-1)!} x^{n-1}, & x > 0. \end{cases}$$

Επειδή $p_0 = P(N=0) \neq 0$, η τ.μ. S είναι μικτού τύπου με μάζα πιθανότητας στο σημείο μηδέν, την $P(S=0) = P(N=0) = e^{-\mu} = e^{-2} = 0.13534$, και συνεχής στο διάστημα $(0,\infty)$.
Συνεπώς,

```
library(actuar)

#Compute a discrete probability mass function from an exponential
#distribution
fx <- discretize(pexp(x, 2),from=0,to=100,step=1,method="lower")

#Compute the aggregate claim amount cumulative distribution function
#of a portfolio
Gs <- aggregateDist("recursive",model.freq="poisson",model.sev=fx,
                    lambda=2)

#Get or set the type or storage mode of an object.
mode(Gs)

#Vector of X
```

```
x <- 0:10

#Compute the probability density function from the cumulative
#distribution function
g <-diff(Gs)[0:11]

#Combine R objects by columns
mat<-cbind(x,g,Gs(x))

#Retrieve or set the row or column names of a matrix-like object
colnames(mat) <- c("x","gs(x)","Gs(x)")

#Rounding of Numbers
round(mat,digits=5)

#Output (x=vector, gs(x)=probability density function,
#Gs(x)=cumulative distribution function)
       x   gs(x)    Gs(x)
[1,]   0  0.13534  0.13534
[2,]   1  0.23404  0.36937
[3,]   2  0.23404  0.60341
[4,]   3  0.17571  0.77913
[5,]   4  0.10949  0.88862
[6,]   5  0.05965  0.94827
[7,]   6  0.02931  0.97758
[8,]   7  0.01326  0.99084
[9,]   8  0.00560  0.99644
[10,]  9  0.00224  0.99868
[11,]  10 0.00085  0.99953
```

□

ΠΑΡΑΔΕΙΓΜΑ 5.2 ***(Σύνθετη γεωμετρική-εκθετική)***

Έστω ότι η τ.μ. $N \sim G_0(p)$ *και η τ.μ.* $X \sim Exp(\lambda)$ *με* $p = 0.5, \lambda = 1$. *Τότε, για* $x \geq 0$ *η συνάρτηση δεξιάς ουράς της τ.μ. S δίνεται από την σχέση*

$$\bar{G}(x) = e^{-\lambda x} \sum_{n=1}^{\infty} p_n \sum_{j=0}^{n-1} \frac{(\lambda x)^j}{j!} = e^{-\lambda x} \sum_{n=1}^{\infty} pq^n \sum_{j=0}^{n-1} \frac{(\lambda x)^j}{j!}$$
$$= pe^{-\lambda x} \sum_{j=0}^{\infty} \sum_{n=j+1}^{\infty} q^n \frac{(\lambda x)^j}{j!} = pe^{-\lambda x} \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} \sum_{n=j+1}^{\infty} q^n.$$

Ύστερα από πράξεις προκύπτει ότι:

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

Επίσης, για $x > 0$ *είναι*

$$g(x) = e^{-\lambda x} \sum_{n=1}^{\infty} pq^n \frac{\lambda^n}{(n-1)!} x^{n-1} = pq\lambda e^{-\lambda x} \sum_{n=1}^{\infty} \frac{(q\lambda x)^{n-1}}{(n-1)!}$$
$$= pq\lambda e^{-\lambda x} \sum_{m=0}^{\infty} \frac{(q\lambda x)^m}{m!} = pq\lambda e^{-\lambda x} e^{q\lambda x} = pq\lambda e^{p\lambda x}.$$

Επειδή $p_0 = P(N = 0) \neq 0$, *η τυχαία μεταβλητή S είναι μικτού τύπου με μάζα πιθανότητας στο σημείο μηδέν, την* $P(S = 0) = P(N = 0) = p$, *και συνεχής στο διάστημα* $(0, \infty)$. *Τότε, είναι*

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

Επομένως, για $p = 0.5$ *και* $\lambda = 1$, *είναι:*

```
library(actuar)
fx <- discretize(pexp(x, 1),from = 0,to = 200,step=1,method ="lower")
Gs <- aggregateDist("recursive", model.freq ="geometric", model.sev=fx,
                    prob=0.5)
mode(Gs) ; x <- 0:10 ; g <-diff(Gs)[0:11]
mat<-cbind(x,g,Gs(x)) ; colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)
```

```
         x    gs(x)    Gs(x)
[1,]     0  0.50000  0.50000
[2,]     1  0.15803  0.65803
[3,]     2  0.10808  0.76611
[4,]     3  0.07392  0.84004
[5,]     4  0.05056  0.89059
[6,]     5  0.03458  0.92517
[7,]     6  0.02365  0.94882
[8,]     7  0.01618  0.96500
[9,]     8  0.01106  0.97606
[10,]    9  0.00757  0.98363
[11,]   10  0.00517  0.98880
```

□

(B) *Αν η τ.μ N είναι φραγμένη,* δηλαδή αν $N \in \{1, 2, ..., M\}$ με $M < \infty$, τότε σύμφωνα με την Πρόταση 5.1 το πρώτο άθροισμα δεν θα εκτείνεται μέχρι το ∞ αλλά μέχρι το M, οπότε για την εφαρμογή των τύπων χρειάζεται ο υπολογισμός ενός πεπερασμένου πλήθους όρων.

ΠΑΡΑΔΕΙΓΜΑ 5.3 ***(Σύνθετη διωνυμική-εκθετική)***

Έστω ότι η τ.μ. $N \sim B(m, p),\ 0 < p < 1$ και η τ.μ. $X \sim Exp(\lambda),\ \lambda > 0$, τότε είναι $M = m$. Τότε, για $x \geq 0$ η συνάρτηση κατανομής της τ.μ. S σύμφωνα με την Πρόταση 5.1 είναι

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

και

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

Επομένως, για $p = 0.2,\ m = 3$ και $\lambda = 5$, είναι:

```
library(actuar)
fx <- discretize(pexp(x, 5), from = 0, to = 200, step=1,
                method ="lower")
Gs <- aggregateDist("recursive",model.freq ="binomial",model.sev=fx,
                    prob=0.2,size=3)
mode(Gs); x <- 0:5; g <-diff(Gs)[1:6];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

      x   gs(x)    Gs(x)
[1,]  0  0.51200  0.51200
[2,]  1  0.38141  0.89341
[3,]  2  0.09728  0.99069
[4,]  3  0.00913  0.99983
[5,]  4  0.00017  1.00000
[6,]  5  0.00000  1.00000
```

□

ΠΑΡΑΔΕΙΓΜΑ 5.4 *Έστω ότι το μέγεθος ατομικής ζημιάς X έχει συνάρτηση πυκνότητας πιθανότητας*

$$f(x) = 4xe^{-2x}, \quad x > 0$$

και το πλήθος των ζημιών N έχει συνάρτηση πιθανότητας $p_n = 1/3, \ n = 0, 1, 2$. Τότε, από το Πόρισμα 5.1 έπεται ότι η τ.μ. N είναι φραγμένη, άρα η συνάρτηση κατανομής της τ.μ. S υπολογίζεται ως εξής:

$$\begin{aligned}
G(x) = 1 - \bar{G}(x) &= 1 - e^{-2x} \sum_{n=1}^{2} p_n \sum_{j=0}^{2n-1} \frac{(2x)^j}{j!} \\
&= 1 - e^{-2x} \left\{ p_1 \sum_{j=0}^{1} \frac{(2x)^j}{j!} + p_2 \sum_{j=0}^{3} \frac{(2x)^j}{j!} \right\} \\
&= 1 - \frac{1}{3} e^{-2x} \left\{ 1 + 2x + 1 + 2x + \frac{(2x)^2}{2!} + \frac{(2x)^3}{3!} \right\} \\
&= 1 - \frac{2}{3} e^{-2x} \left\{ 1 + 2x + x^2 + \frac{2x^3}{3} \right\}, \quad x \geq 0.
\end{aligned}$$

Επειδή $p_0 \neq 0$ *και η τ.μ.* $X \sim G(2,2)$, *έπεται ότι η τ.μ. S είναι μικτού τύπου τ.μ. με μάζα πιθανότητας στο σημείο μηδέν με* $P(S=0) = P(N=0) = p_0 = 1/3$ *και συνεχής στο διάστημα* $(0,\infty)$.
Επομένως, η συνάρτηση πυκνότητας πιθανότητας της τ.μ. S, είναι

$$g(x) = \begin{cases} \dfrac{1}{3}, & x = 0, \\ \dfrac{4}{3}e^{-2x}\left(x + \dfrac{2x^3}{3}\right), & x > 0. \end{cases}$$

Συνεπώς,

```
library(actuar)
fx <- discretise(pgamma(x,shape=2,scale=2), from=0, to=100, step=1,
                 method="lower")
fn <-c(1/3,1/3,1/3)
Gs <- aggregateDist("convolution", model.freq =fn, model.sev = fx)
mode(Gs); x <- 0:10; g <-diff(Gs)[1:11]
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)


          x    gs(x)     Gs(x)
[1,]     0  0.33333  0.33333
[2,]     1  0.03007  0.36340
[3,]     2  0.06072  0.42413
[4,]     3  0.06978  0.49390
[5,]     4  0.07140  0.56531
[6,]     5  0.06934  0.63465
[7,]     6  0.06469  0.69934
[8,]     7  0.05817  0.75751
[9,]     8  0.05057  0.80808
[10,]    9  0.04264  0.85071
[11,]   10  0.03499  0.88570
```

$\square$

5.2 Σύνθετες Bernoulli κατανομές και μικτού τύπου τυχαίες μεταβλητές

ΠΡΟΤΑΣΗ 5.2 *Αν η κατανομή της τυχαίας μεταβλητής Y είναι η διακριτή μείξη της εκφυλισμένης κατανομής στο σημείο μηδέν και της κατανομής της τυχαίας μεταβλητής Z με βάρη μείξης $1-\theta$ και θ αντίστοιχα, τότε η τ.μ. Υ έχει μια σύνθετη Bernoulli κατανομή, και συγκεκριμένα*

$$Y = Z_1 + Z_2 + ... + Z_N,$$

όπου η τ.μ. $N \sim Bernoulli(\theta)$ και οι τυχαίες μεταβλητές $Z_1, Z_2, Z_3, ...$ είναι ανεξάρτητες και ισόνομες με την τ.μ. Z.

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.5 ***(Σύνθετη Bernoulli-εκθετική)***

Έστω η τ.μ. $X \sim Exp(2)$ και η τ.μ. $N \sim Bernoulli(0.1)$, τότε η συνάρτηση πυκνότητας πιθανότητας και η συνάρτηση κατανομής της τ.μ. S των συνολικών ζημιών, δίνεται από τις σχέσεις:

$$G(x) = q + p(1 - e^{-\lambda x}) = 1 - pe^{-\lambda x}, \quad x \geq 0.$$

και

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

Επομένως,

```
library(actuar)
library(Rlab) #we use Rlab package for Bernoulli distribution

fn <- dbern(x, 0.1, log = FALSE)
fx <- discretize(pexp(x,2),from=0,to=200,method="lower")
Gs <- aggregateDist("convolution",model.freq=fn,model.sev = fx)
mode(Gs); x <- 0:5; g <-diff(Gs)[0:6]
```

```
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

      x    gs(x)     Gs(x)
[1,]  0  0.90000   0.90000
[2,]  1  0.08647   0.98647
[3,]  2  0.01170   0.99817
[4,]  3  0.00158   0.99975
[5,]  4  0.00021   0.99997
[6,]  5  0.00003   1.00000
```

□

Στο παραπάνω παράδειγμα θεωρήθηκε απαραίτητο να γίνει η προσθήκη του πακέτου Rlab για την χρήση της κατανομής Bernoulli, καθώς το πακέτο actuar δεν διαθέτει την εν λόγω κατανομή. Στην πορεία θα θεωρηθεί ότι η κατανομή Bernoulli είναι ειδική περίπτωση της διωνυμικής κατανομής.

ΠΑΡΑΔΕΙΓΜΑ 5.6 ***(Σύνθετη διωνυμική-ομοιόμορφη κατανομή)***

Μια τριμελής οικογένεια ασφαλίζεται με ένα ομαδικό συμβόλαιο ενός έτους. Κάθε μέλος της οικογένειας έχει την δυνατότητα να επισκεφτεί το συμβεβλημένο ιατρικό κέντρο το πολύ μια φορά κατά την διάρκεια του έτους. Αν η πιθανότητα κάποιο μέλος της οικογένειας να επισκεφτεί το ιατρικό κέντρο είναι ίση με 80% και το αντίστοιχο κόστος σε ευρώ κατανέμεται ομοιόμορφα στο διάστημα $(0, 1000)$, αν $S = X_1 + X_2 + ... + X_N$, όπου

$$N = N_1 + N_2 + N_3, \quad N_i \sim Bernoulli(0.8)$$

και $X \sim U(0, 1000)$, τότε

$$N \sim B(3, 0.8).$$

Η συνάρτηση κατανομής της τ.μ. S σύμφωνα με την Σχέση (3.2), είναι

$$G(x) = \sum_{n=0}^{3} p_n F^{*n}(x) = p_0 F^{*0}(x) + p_1 F^{*1}(x) + p_2 F^{*2}(x) + p_3 F^{*3}(x).$$

```
library(actuar)
pn<-dbinom(x,3,0.8,log=FALSE)
fx<-discretise(punif(x,0,1000),from = 0,to=200,step=1,method ="lower")
Gs <- aggregateDist("convolution",model.freq =pn,model.sev=fx)
mode(Gs)
x <- 0:5
mat<-cbind(x,Gs(x))
colnames(mat) <- c("x","Gs(x)")
round(mat,digits=10)

     x      Gs(x)
[1,] 0   0.008000000
[2,] 1   0.008096000
[3,] 2   0.008192384
[4,] 3   0.008289153
[5,] 4   0.008386306
[6,] 5   0.008483845
```

□

ΠΡΟΤΑΣΗ 5.3 *Αν η τ.μ. Υ έχει την σύνθετη Bernoulli κατανομή, και συγκεκριμένα*

$$Y = Z_1 + Z_2 + ... + Z_N,$$

όπου η $N \sim Bernoulli(\theta)$ *και οι τ.μ.* $Z_1, Z_2, Z_3, ...$ *είναι ανεξάρτητες και ισόνομες με την τ.μ. Z και οι τυχαίες μεταβλητές Z και N είναι ανεξάρτητες, τότε ισχύει ότι*

$$Y \stackrel{d}{=} NZ.$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΡΟΤΑΣΗ 5.4 *(Συνέλιξη μικτού τύπου τυχαίων μεταβλητών)*
Έστω οι ανεξάρτητες σύνθετες Bernoulli τυχαίες μεταβλητές

$$S_1 = X_1 + X_2 + ... + X_{N_1} \quad \kappa\alpha\iota \quad S_2 = X_1 + X_2 + ... + X_{N_2}$$

όπου η τ.μ. $N_i \sim Bernoulli(p_i),\ \ 0 < p_i < 1,\ \ i = 1, 2$. *Τότε, για την τ.μ.* $S = S_1 + S_2$*, ισχύει:*

(α) *η συνάρτηση κατανομής* $F(s) = P(S \leq x)$*, είναι*

$$F_S(x) = q_1q_2 + p_1q_2F_X(x) + q_1p_2F_Y(x) + p_1p_2F_{X+Y}(x), \quad x \geq 0$$

όπου $q_i = 1 - p_i,\ \ i = 1, 2$*,*

(β) *η συνάρτηση πυκνότητας πιθανότητας* $f_S(x)$*, είναι:*

$$f_S(x) = \begin{cases} q_1q_2, & x = 0 \\ p_1q_2f_X(x) + q_1p_2f_Y(x) + p_1p_2f_{X+Y}(x), & x > 0, \end{cases}$$

(γ) *ισχύει ότι:*

$$S \stackrel{d}{=} W_1 + W_2 + ... + W_M,$$

όπου η τ.μ. $M \sim Bernoulli(1 - q_1q_2)$ *και οι τυχαίες μεταβλητές* $W_1, W_2, W_3, ...$*, είναι ανεξάρτητες και ισόνομες, έστω με την τ.μ. W που έχει συνάρτηση πυκνότητας πιθανότητας*

$$f_W(x) = \frac{p_1q_2}{1 - q_1q_2}f_X(x) + \frac{q_1pq_2}{1 - q_1q_2}f_Y(x) + \frac{p_1p_2}{1 - q_1q_2}f_{X+Y}(x), \quad x > 0.$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

5.3 Σύνθετες Poisson κατανομές

Στην περίπτωση όπου η τυχαία μεταβλητή $N \sim Poisson(\lambda)$, και το ύψος της ατομικής ζημιάς Χ ακολουθεί μια συνεχή κατανομή ή έχει συνάρτηση πιθανότητας $f(x) = F(x) - F(x-1)$ στην περίπτωση που η Χ είναι διακριτή, τότε η $S \sim CP(\lambda, F)$ ή $S \sim CP(\lambda, f)$, όπου CP είναι τα αρχικά των λέξεων Compound Poisson (Σύνθετη Poisson).

Στην κατανομή Poisson, μια βασική ιδιότητά της είναι ότι η μέση τιμή, η διακύμανση και η τρίτη κεντρική ροπή της είναι ίσες με την παράμετρο της κατανομής, δηλαδή ισχύει ότι:

$$E(N) = Var(N) = E\lfloor(N - (E(N))^3)\rfloor = \lambda.$$

Όταν η $N \sim Poisson(\lambda)$, η συνάρτηση πυκνότητας πιθανότητάς της $p_n = P(N = n)$, είναι:

$$p_n = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2, 3, ...$$

και η πιθανογεννήτρια και ροπογεννήτρια της κατανομής είναι

$$P_N(u) = e^{\lambda(u-1)} \quad \text{και} \quad M_N(t) = e^{\lambda(e^t-1)}.$$

Συνεπώς, όταν $S \sim CP(\lambda, f)$, ισχύει:

$$P_N(u) = P_N(P_X(u)) = e^{\lambda(P_x(u)-1)},$$
$$M_N(t) = P_N(M_X(t)) = e^{\lambda(M_x(t)-1)}.$$

Με βάση τα παραπάνω, μπορεί να αποδειχθεί το ακόλουθο αποτέλεσμα για τις ροπές των συνολικών ζημιών S.

ΠΡΟΤΑΣΗ 5.5 *Ισχύει ότι:*

(α) $E(S) = \lambda E(X)$,

(β) $Var(S) = \lambda E(X^2)$,

(γ) $E[(S - E(S)^3)] = E(S^3) = \lambda E(X^3)$,

(δ) $\gamma_1 = \dfrac{E[(S - E(S)^3)]}{Var(S)^{3/2}} = \dfrac{\lambda E(X^3)}{(\lambda E(X^2))^{3/2}}$.

(βλ. Πολίτης, 2012).

Στη συνέχεια θα παρουσιαστούν οι τρόποι υπολογισμού της κατανομής της τυχαίας μεταβλητής S στην περίπτωση της σύνθετης κατανομής Poisson. Στην εν λόγω περίπτωση υπάρχουν τρεις κύριες μέθοδοι υπολογισμού:

1. με την βασική μέθοδο,

2. μέσω αναδρομικών σχέσεων,

3. με την εναλλακτική μέθοδο.

5.3.1 Βασική μέθοδος

ΠΟΡΙΣΜΑ 5.1 *Αν η τυχαία μεταβλητή $S \sim CP(\lambda, f)$, τότε ισχύει ότι:*

(α) $G(x) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} F^{*n}(x), \quad x \geq 0,$

(β) $\bar{G}(x) = e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \bar{F}^{*n}(x), \quad x \geq 0,$

(γ) *Αν η X είναι μη αρνητική ακέραια τυχαία μεταβλητή, τότε και η S είναι μη αρνητική ακέραια τυχαία μεταβλητή, με*

$$g(x) = \begin{cases} e^{\lambda(f(0)-1)}, & \text{x=0,} \\ e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} f^{*n}(x), & \text{x=1,2,3,...} \end{cases}$$

(δ) *Αν η X είναι θετική ακέραια τυχαία μεταβλητή, τότε η S είναι μη αρνητική ακέραια τυχαία μεταβλητή, με*

$$g(x) = \begin{cases} e^{-\lambda}, & \text{x=0,} \\ e^{-\lambda} \sum_{n=1}^{x} \frac{\lambda^n}{n!} f^{*n}(x), & \text{x=1,2,3,...} \end{cases}$$

(ϵ) *Αν η X είναι μη αρνητική συνεχής τυχαία μεταβλητή, τότε η S είναι μικτού τύπου τυχαία μεταβλητή, με μάζα πιθανότητας στο σημείο μηδέν και συνεχής στο* διάστημα (0,∞), με

$$g(x) = \begin{cases} e^{-\lambda}, & \text{x=0,} \\ e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} f^{*n}(x), & \text{x>0.} \end{cases}$$

ΠΑΡΑΔΕΙΓΜΑ 5.7 *Για ένα χαρτοφυλάκιο κινδύνων, η ασφαλιστική εταιρεία αναμένει να εμφανισθούν 0.2 απαιτήσεις σ' ένα έτος. Οι απαιτήσεις εμφανίζονται σύμφωνα με το μοντέλο της διαδικασίας Poisson, δηλαδή η τυχαία μεταβλητή $N \sim Poisson(0.2)$. Επίσης, η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X, όπου X είναι το ύψος της ατομικής απαίτησης, έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$ και είναι*

x	1	2	3
f(x)	3/5	1/5	1/5

Έστω ότι στόχος του παρόντος παραδείγματος είναι η εύρεση της $g(x) = Pr(S = x)$, για $x = 0, 1, ..., 6$, όπου S είναι οι συνολικές απαιτήσεις του χαρτοφυλακίου.

Επειδή η $N \sim Poi(\lambda) \Leftrightarrow \lambda = 0.2$.

Οι συνολικές απαιτήσεις του χαρτοφυλακίου δίνονται από την σχέση:

$S = X_1 + X_2 + ... + X_N, N \geq 1$ με $S = 0$ αν $N = 0$ και $S \sim CP(\lambda, f)$.

Σύμφωνα με το Πόρισμα 3.1 ισχύει:

$$g(x) = \begin{cases} e^{-0.2}, & \text{x=0} \\ e^{-0.2} \sum_{n=1}^{x} \frac{(0.8)^n}{n!} f^{*n}(x), & \text{x=1,2,3,...,} \end{cases}$$

όπου f^{*n} είναι η n-οστή συνέλιξη της f της τυχαίας μεταβλητής X και υπολογίζεται από την σχέση:

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

όπου $f^{*1}(x) = f(x)$ και $f^{*0}(x) = \begin{cases} 1, & \text{x=0,} \\ 0, & x \neq 0 \end{cases}$.

Για τον υπολογισμό των $g(1), g(2), ..., g(6)$ *πρέπει πρώτα να γίνει υπολογισμός των* f^{*n} *για κάθε* $n \geq 6$ *και* $x \geq 6$. *Τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα*

x	$f^{*0}(x)$	$f^{*1}(x)$	$f^{*2}(x)$	$f^{*3}(x)$	$f^{*4}(x)$	$f^{*5}(x)$	$f^{*6}(x)$
0	1	0	0	0	0	0	0
1	0	0.6	0	0	0	0	0
2	0	0.2	0.36	0	0	0	0
3	0	0.2	0.24	0.216	0	0	0
4	0	0	0.28	0.216	0.1296	0	0
5	0	0	0.08	0.288	0.1728	0.07776	0
6	0	0	0.04	0.152	0.2592	0.12960	0.04666

Για $p_n = e^{-0.2}\dfrac{(0.2)^n}{n!} = 0.81873.$

n	0	1	2	3	4	5	6
p_n	0.81873	0.1637462	0.0163746	0.0010916	0.0000546	0.0000022	0.0000001

Επομένως,

$g(1) = p_1 f^{*1}(1) = 0.09825$

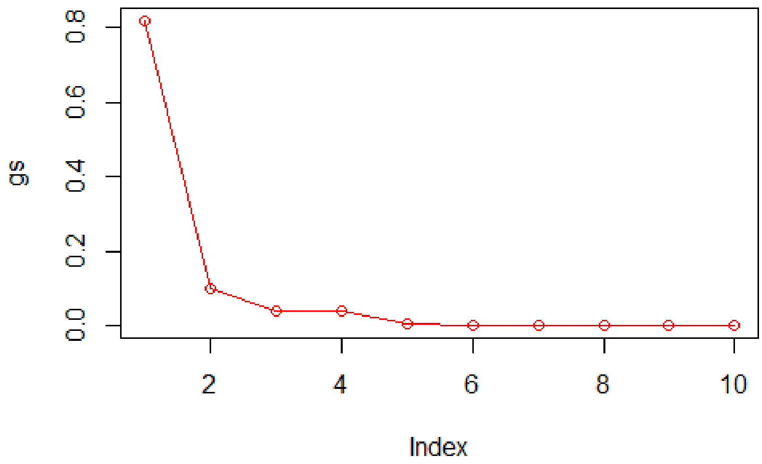
$g(2) = p_1 f^{*1}(2) + p_2 f^{*2}(2) = 0.03864$

$g(3) = p_1 f^{*1}(3) + p_2 f^{*2}(3) + p_3 f^{*3}(3) = 0.03691$

$g(4) = p_1 f^{*1}(4) + p_2 f^{*2}(4) + p_3 f^{*3}(4) + p_4 f^{*4}(4) = 0.00483$

$g(5) = p_1 f^{*1}(5) + p_2 f^{*2}(5) + p_3 f^{*3}(5) + p_4 f^{*4}(5) + p_5 f^{*5}(5) = 0.00163$

$g(6) = p_1 f^{*1}(6) + p_2 f^{*2}(6) + p_3 f^{*3}(6) + p_4 f^{*4}(6) + p_5 f^{*5}(6) + p_6 f^{*6}(6) = 0.00084$



(βλ. ΠΑΡΑΡΤΗΜΑ Δ).

Σχολιασμός: *Σύμφωνα με το διάγραμμα, φαίνεται ότι για $x = 0$, η τιμή που παίρνει το $g(x)$ είναι 0.81873. Από κει και έπειτα, οι τιμές φαίνεται να φθίνουν πλησιάζοντας το μηδέν, έως την τιμή $x = 10$ όπου η $g(x)$ γίνεται μηδέν.*

$\square$

ΠΑΡΑΔΕΙΓΜΑ 5.8 *(Σύνθετη Poisson-Bernoulli)*

Έστω τώρα ότι η $S \sim CP(\lambda, f), \lambda > 0$ και η $X \sim Bernoulli(p), 0 < p < 1$. Χρησιμοποιώντας τις πιθανογεννήτριες συναρτήσεις είναι,

$$P_x(u) = 1 - p + pu,$$

οπότε η πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής S θα είναι

$$P_S(u) = e^{\lambda[P_X(u)-1]} = e^{\lambda(1-p+pu-1)} = e^{\lambda(pu-p)} = e^{\lambda p(u-1)}.$$

Δηλαδή η τυχαία μεταβλητή $S \sim P(\lambda p)$. Συνεπώς, η σύνθετη Poisson(λ)-Bernoulli(p) τ.μ. είναι ισόνομη με την Poisson(λp).

Έστω ότι η τ.μ. $X \sim B(1, 0.8)$ και η τ.μ. $N \sim P(2)$, τότε:

```
library(actuar)
fx<-dbinom(x,1,0.8,log = FALSE)
Gs <- aggregateDist("recursive", model.freq ="poisson", model.sev=fx,
                    lambda=2)
mode(Gs); x <- 0:10; g <-diff(Gs)[1:11];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

       x    gs(x)    Gs(x)
[1,]   0  0.20190  0.20190
[2,]   1  0.32303  0.52493
[3,]   2  0.25843  0.78336
[4,]   3  0.13783  0.92119
[5,]   4  0.05513  0.97632
[6,]   5  0.01764  0.99396
[7,]   6  0.00470  0.99866
[8,]   7  0.00108  0.99974
[9,]   8  0.00022  0.99995
[10,]  9  0.00004  0.99999
[11,] 10  0.00001  1.00000
```

***Σχόλιο**: Το πακέτο actuar δεν αναγνωρίζει την κατανομή Bernoulli, καθώς την θεωρεί ως μια ειδική περίπτωση της διωνυμικής κατανομής με $m = 1$. Για τον λόγο αυτό, στο εν λόγω παράδειγμα θεωρήθηκε ότι η κατανομή που ακολουθεί η τ.μ. S είναι η σύνθετη Poisson(2)-διωνυμική(1,0.8).*

□

5.3.2 Αναδρομικές μέθοδοι

ΘΕΩΡΗΜΑ 5.1 *Αν η τυχαία μεταβλητή $X \in \{0, 1, 2, ...\}$ και η $S \sim CP(\lambda, f)$, τότε η συνάρτηση πιθανότητας $g(x) = P(S = x)$ για $x = 0, 1, 2, ...$, μπορεί να υπολογισθεί μέσω της αναδρομικής σχέσης*

$$g(x) = \frac{\lambda}{x} \sum_{y=1}^{x} y f(y) g(x - y), \quad x = 1, 2, 3, ... \tag{5.1}$$

με $g(0) = e^{\lambda(f(0)-1)}$.

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΘΕΩΡΗΜΑ 5.2 *Αν η τυχαία μεταβλητή* $X \in \{0, 1, 2, ...\}$ *έχει πιθανογεννήτρια συνάρτηση* $P_x(u) = E(u^X)$, *τέτοια ώστε*

$$P'_X(u) = \frac{\sum_{x=0}^{\infty} \alpha_x u^x}{\sum_{x=0}^{\infty} b_x u^x}$$

και η τ.μ $S \sim CP(\lambda, f)$, *τότε η συνάρτηση πυκνότητας πιθανότητας* $g(x) = P(S = x)$ *ικανοποιεί την αναδρομική σχέση*

$$g(x) = \frac{1}{xb_0} \sum_{y=1}^{x} [\lambda \alpha_{y-1} - (x-y) b_y] g(x-y), \quad x = 1, 2, 3, ...$$

με $g(0) = e^{\lambda(f(0)-1)}$.

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.9 *Έστω ότι η τ.μ* $N \sim P(2)$ *και η τ.μ.* $X \epsilon \{1, 2\}$ *με συνάρτηση πυκνότητας πιθανότητας* $f(x) = P(X = x)$,

$$f(1) = 0.5, \qquad f(2) = 0.5,$$

όπου

N=πλήθος απαιτήσεων του χαρτοφυλακίου σε ένα έτος

X=ύψος ατομικής απαίτησης

Αν S είναι οι συνολικές απαιτήσεις του χαρτοφυλακίου σε ένα έτος τότε, ισχύει:

$$S = X_1 + X_2 + ... + X_N, \quad N \geq 1$$

με $S = 0$ *αν* $N = 0$ *και* $S \sim CP(\lambda, f)$.

Η ζητούμενη πιθανότητα είναι η $P(S \leq 6) = G(6)$ *και επειδή* $G(6) = \sum_{x=0}^{6} g(x)$, *αρχικά θα γίνει υπολογισμός των τιμών της* $g(x) = P(S = x)$ *για* $x = 0, 1, 2, ..., 6$ *μέσω* *της αναδρομικής σχέσης του Θεωρήματος 5.1.*

Επειδή, $f(0) = 0$ *και* $f(y) = 0$*, για κάθε* $y \geq 3$*, προκύπτει ότι*

$$g(x) = \frac{2}{x}\{f(1)g(x-1) + 2f(2)g(x-2)\}, \quad x = 1, 2, 3, ...$$

με $g(0) = e^{-\lambda} = e^{-2} = 0.13534$*.*

Άρα,

$g(1) = 2\{f(1)g(0) + 2f(2)g(-1)\} = 0.13534$,

$g(2) = \frac{2}{2}\{f(1)g(1) + 2f(2)g(0)\} = 0.20300$,

$g(3) = \frac{2}{3}\{f(1)g(2) + 2f(2)g(1)\} = 0.15789$,

$g(4) = \frac{2}{4}\{f(1)g(3) + 2f(2)g(2)\} = 0.14097$,

$g(5) = \frac{2}{5}\{f(1)g(4) + 2f(2)g(3)\} = 0.09135$,

$g(6) = \frac{2}{6}\{f(1)g(5) + 2f(2)g(4)\} = 0.06222$.

Άρα η ζητούμενη πιθανότητα είναι

$$G(6) = \sum_{x=0}^{6} g(x) = g(0) + g(1) + ... + g(6) = 0.92611.$$

```
library(actuar)
fx<-c(0,0.5,0.5)
Gs <- aggregateDist("recursive", model.freq ="poisson", model.sev = fx,
                    lambda=2)
mode(Gs) ; x <- 0:18 ; g <-diff(Gs)[0:19]
mat<-cbind(x,g,Gs(x)) ; colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

        x     gs(x)     Gs(x)
[1,]    0   0.13534   0.13534
[2,]    1   0.13534   0.27067
[3,]    2   0.20300   0.47367
[4,]    3   0.15789   0.63156
[5,]    4   0.14097   0.77254
```

```
[6,]    5  0.09135  0.86389
[7,]    6  0.06222  0.92611
[8,]    7  0.03499  0.96110
[9,]    8  0.01993  0.98102
[10,]   9  0.00999  0.99101
[11,] 10  0.00498  0.99600
[12,] 11  0.00227  0.99827
[13,] 12  0.00102  0.99929
[14,] 13  0.00043  0.99971
[15,] 14  0.00018  0.99989
[16,] 15  0.00007  0.99996
[17,] 16  0.00003  0.99999
[18,] 17  0.00001  0.99999
[19,] 18  0.00000  1.00000
```

$\square$

5.3.3 Άθροισμα ανεξάρτητων σύνθετων Poisson τ.μ.

ΘΕΩΡΗΜΑ 5.3 *Έστω $S_1, S_2, ..., S_m$ ανεξάρτητες μεταξύ τους τυχαίες μεταβλητές, με*

$$S_i \sim CP(\lambda_i, f_i), \;\; i = 1, 2, ..., m$$

και

$$S = S_1 + S_2 + ... + S_m.$$

Τότε η τυχαία μεταβλητή $S \sim CP(\lambda, f)$, όπου $\lambda = \lambda_1 + \lambda_2 + ... + \lambda_m$ και

$$f(x) = \frac{\lambda_1}{\lambda} f_1(x) + \frac{\lambda_2}{\lambda} f_2(x) + ... + \frac{\lambda_m}{\lambda} f_m(x).$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Από το παραπάνω Θεώρημα έπεται ότι η συνάρτηση κατανομής της τυχαίας μεταβλητής X είναι ίση με

$$F(x) = \frac{\lambda_1}{\lambda} F_1(x) + \frac{\lambda_2}{\lambda} F_2(x) + ... + \frac{\lambda_m}{\lambda} F_m(x)$$

και η συνάρτηση δεξιάς ουράς είναι ίση με

$$\bar{F}(x) = \frac{\lambda_1}{\lambda}\bar{F}_1(x) + \frac{\lambda_2}{\lambda}\bar{F}_2(x) + ... + \frac{\lambda_m}{\lambda}\bar{F}_m(x),$$

όπου $\bar{F}_i(x) = 1 - F_i(x), \ \ i = 1, 2, ..., m.$

ΠΑΡΑΔΕΙΓΜΑ 5.10 *[9] Έστω ότι μια ασφαλιστική εταιρία ασφαλίζει δύο ανεξάρτητες καλύψεις για τον ίδιο κίνδυνο. Έστω $S_1 \sim CP(\lambda_1, f_1)$ το ύψος της συνολικής απαίτησης της πρώτης κάλυψης, με*

$$\lambda_1 = 0.2, \quad f_1(1) = 0.8, \quad f_1(2) = 0.2$$

και $S_2 \sim CP(\lambda_2, f_2)$ το ύψος της συνολικής απαίτησης της δεύτερης κάλυψης, με

$$\lambda_2 = 0.3, \quad f_2(2) = f_2(4) = 0.5.$$

Αν S είναι οι συνολικές καλύψεις του χαρτοφυλακίου έτσι ώστε να ισχύει $S = S_1 + S_2$, έπεται ότι η τυχαία μεταβλητή $S \sim CP(\lambda, f)$. Δηλαδή,

$$S = X_1 + X_2 + ... + X_N,$$

όπου η τ.μ. $N \sim Poisson(\lambda)$, με $\lambda = \lambda_1 + \lambda_2 = 0.2 + 0.3 = 0.5.$
Αντίστοιχα, η διακριτή τ.μ. X έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$, όπου

$$f(x) = \frac{\lambda_1}{\lambda}f_1(x) + \frac{\lambda_2}{\lambda}f_2(x) = \frac{0.2}{0.5}f_1(x) + \frac{0.3}{0.5}f_2(x) = 0.4f_1(x) + 0.6f_2(x), \ \ \forall x \geq 0.$$

Τότε:

- $x = 1: \quad f(1) = 0.4f_1(1) + 0.6f_2(1) = (0.4)(0.8) = 0.32,$
- $x = 2: \quad f(2) = 0.4f_1(2) + 0.6f_2(2) = (0.4)(0.2) + (0.6)(0.5) = 0.38,$
- $x = 4: \quad f(4) = 0.4f_1(4) + 0.6f_2(4) = (0.6)(0.5) = 0.30,$
- $x \neq 1, 2, 4 \Rightarrow f(x) = 0.$

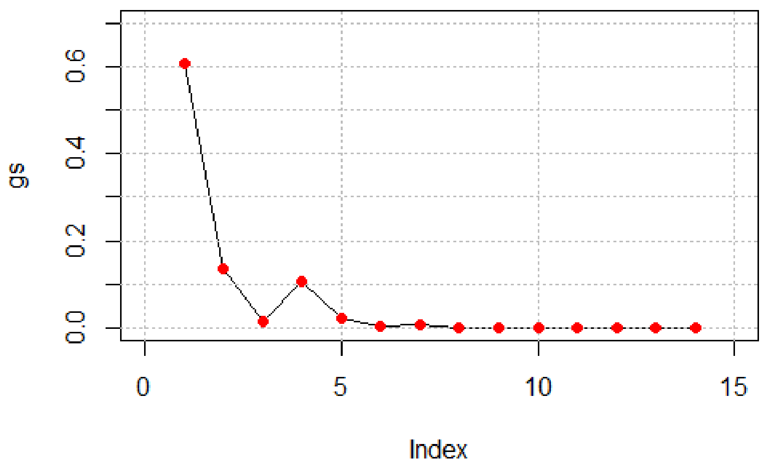
Συνεπώς, η συνάρτηση πιθανότητας $g(x) = P(S = x)$ *της τυχαίας μεταβλητής S, μπορεί να υπολογισθεί σύμφωνα με την αναδρομική σχέση (5.1) ως εξής:*

$$\begin{aligned} g(x) &= \frac{\lambda}{x}\{1f(1)g(x-1) + 2f(2)g(x-2) + 3f(3)g(x-3) + ...\} \\ &= \frac{0.5}{x}\{0.32g(x-1) + 2(0.38)g(x-2) + 3(0)g(x-3) + 4(0.30)g(x-4) + ...\} \\ &= \frac{0.5}{x}\{0.32g(x-1) + 0.76g(x-2) + 1.2g(x-4)\}, \quad x = 1, 2, 3, ... \end{aligned}$$

με αρχική τιμή $g(0) = e^{-0.5} = 0.6065307$ *και* $g(x) = 0 \ \forall x < 0$.

```
l1<-0.2
l2<-0.3
lambda<-l1+l2
f1<-c(0.8,0.2,0,0)
f2<-c(0,0.5,0,0.5)

fx<-(l1/lambda)*f1+(l2/lambda)*f2
fx

[1] 0.32 0.38 0.00 0.30

library(actuar)
Gs <- aggregateDist("recursive", model.freq ="poisson", model.sev = fx,
                    lambda=0.5)
mode(Gs) ; x <- 1:13; g <-diff(Gs)[2:14]
mat<-cbind(x,g) ; colnames(mat) <- c("x","gs(x)")
round(mat,digits=5)

      x    gs(x)
[1,]  1  0.13524
[2,]  2  0.01285
[3,]  3  0.10758
[4,]  4  0.02032
[5,]  5  0.00193
[6,]  6  0.00813
[7,]  7  0.00153
[8,]  8  0.00014
```

```
[9,]    9   0.00041
[10,]  10   0.00008
[11,]  11   0.00001
[12,]  12   0.00002
[13,]  13   0.00000

g0<-exp(-lambda) #g(0)=e^{-0.5}

[1]  0.6065307

gs<-c(g0,g)

# plot(gs(x),type="l" for lines, lwd=controls the line width,
# xlim=axis limits, ylim=axis limits)
plot(gs,type="l",lwd=1,xlim=c(0,15),ylim=c(0,0.7))
# grid(nx=number of cells of the grid in x,
# col=color of the grid lines)
grid(nx=NULL,col="grey")
# Add points to a Plot
# points(x=gs, pch=solid circle, col=color code or name)
points(gs,pch=19,col="red")
```



Σχολιασμός: *Σύμφωνα με το παραπάνω διάγραμμα, για* $x = 3$ *φαίνεται να πλησιάζει την τιμή μηδέν. Αυτό οφείλεται στο γεγονός ότι οι τ.μ.* $X_1, X_2, \ldots$ *δεν* *έχουν τιμή στο συγκεκριμένο σημείο, κάτι που έχει ως αποτέλεσμα η συνάρτηση πιθανότητας της τ.μ. S στο σημείο αυτό να πλησιάζει την τιμή μηδέν, χωρίς όμως να είναι απαραίτητα μηδενική. Ένα επιπλέον συμπέρασμα είναι ότι από την στιγμή που οι μεταβλητές X παίρνουν τιμές μέχρι την τιμή 4, όπως είναι εμφανές και από το γράφημα, μετά την τιμή αυτή αρχίζουν να φθίνουν με γρήγορο ρυθμό, έως ότου πλησιάσουν πολύ κοντά στην τιμή μηδέν.*

$\square$

5.3.4 Η εναλλακτική μέθοδος

Έστω ότι το μέγεθος της ατομικής ζημιάς X είναι μια διακριτή τ.μ. με

$$X \in \{x_1, x_2, ..., x_m\},\ x_1 \neq x_2 \neq ..., \neq x_m,$$

και η τ.μ

$$N_i = \text{το πλήθος τιμών της τ.μ. } X_i \text{ που είναι όλες ίσες με } x_i, i = 1, 2, ..., m.$$

Τότε, επειδή υπάρχουν:

- N_1 τιμές ίσες με x_1, οπότε οι συνολικές τιμές είναι ίσες με $x_1 N_1$,
- N_2 τιμές ίσες με x_2, οπότε οι συνολικές τιμές είναι ίσες με $x_2 N_2$,
- ...
- N_m τιμές ίσες με x_m, οπότε οι συνολικές τιμές είναι ίσες με $x_m N_m$,

έπεται ότι η τ.μ. $S = X_1 + X_2 + ... + X_N$ γράφεται στην μορφή

$$S = x_1 N_1 + x_2 N_2 + ... + x_m N_m,$$

και επίσης ισχύει ότι

$$N = N_1 + N_2 + ... + N_m.$$

Επειδή οι $x_1, x_2, ..., x_m$ είναι οι διαφορετικές τιμές που παίρνει η τ.μ. X, έπεται ότι τα σύνολα $\{x_1\}$, $\{x_2\}$,...,$\{x_m\}$ αποτελούν μια διαμέριση του δειγματοχώρου των τιμών της τ.μ. X, οπότε θα είναι:

$$\pi_i = P(X \in A_i) = P(X = x_i) = f(x_i), \quad \forall i = 1, 2, ..., m.$$

Επομένως, σύμφωνα με τα παραπάνω, προκύπτει το επόμενο πόρισμα.

ΠΟΡΙΣΜΑ 5.2 *Έστω η τ.μ. $S = X_1 + X_2 + ... + X_N$, τέτοια ώστε*

$$S \sim CP(\lambda, f),$$

όπου η διακριτή τ.μ. $X \in \{x_1, x_2, ..., x_m\}$, με $x_1 \neq x_2 \neq ... \neq x_m$.
Αν N_i είναι το πλήθος των τιμών της τ.μ X που είναι όλες ίσες με $x_i, i = 1, 2, ..., m$, τότε:

- *οι τυχαίες μεταβλητές $N_1, N_2, ..., N_m$ είναι αμοιβαία ανεξάρτητες, και*
- *$N_i \sim P(\lambda_i)$, όπου $\lambda_i = \lambda\pi_i = \lambda f(x_i), \ \forall i = 1, 2, ..., m.$*

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Ως συνέπεια του Πορίσματος 5.2, είναι εφικτό να ορισθεί η τυχαία μεταβλητή S μέσω των δυνατών τιμών $x_1, x_2, ..., x_m$ της ατομικής ζημιάς και των παραμέτρων $\lambda_1, \lambda_2, ..., \lambda_m$ των αντίστοιχων Poisson τ.μ. $N_1, N_2, ..., N_m$. Έτσι για κάθε δυνατή τιμή υπάρχει μια αντίστοιχη τ.μ. $N_i \sim P(\lambda_i)$, όπου $\lambda_i = \lambda f(x_i)$.

Τότε, είναι

$$E(N_i) = Var(N_i) = \lambda_i, \ \ i = 1, 2, ..., m.$$

και

$$E(S) = \left(\sum_{i=1}^{m} x_i N_i\right) = \sum_{i=1}^{m} x_i E(N_i) = \sum_{i=1}^{m} x_i \lambda_i = \lambda \sum_{i=1}^{m} x_i f(x_i)$$

Επίσης, λόγω της ανεξαρτησίας των τυχαίων μεταβλητών $N_1, N_2, ..., N_m$, είναι:

$$Var(S) = Var\left(\sum_{i=1}^{m} x_i N_i\right) = \sum_{i=1}^{m} x_i^2 Var(N_i) = \sum_{i=1}^{m} x_i^2 \lambda_i = \lambda \sum_{i=1}^{m} x_i^2 f(x_i).$$

Δηλαδή, αν η σύνθετη Poisson τυχαία μεταβλητή S γραφτεί στη μορφή ενός τυχαίου αθροίσματος τυχαίων μεταβλητών, συγκεκριμένα

$$S = X_1 + X_2 + ... + X_N, \quad N \sim P(\lambda), \quad X \in \{x_1, x_2, ..., x_m\},$$

τότε

$$E(S) = \lambda E(X) = \lambda \sum_{i=1}^{m} x_i f(x_i)$$

και

$$Var(S) = \lambda E(X^2) = \lambda \sum_{i=1}^{m} x_i^2 f(x_i).$$

ΠΟΡΙΣΜΑ 5.3 *Έστω η τ.μ. $S = x_1N_1 + x_2N_2 + ... + x_mN_m$, όπου $x_i \neq x_j$, για $i, j \in \{1, 2, ..., m\}$ είναι γνωστοί αριθμοί και οι τ.μ. $N_1, N_2, ..., N_m$ είναι ανεξάρτητες μεταξύ τους με $N_i \sim P(\lambda_i), \ \forall i = 1, 2, ..., m$. Τότε*

$$S \sim CP(\lambda, f), \quad \lambda = \lambda_1 + \lambda_2 + ... + \lambda_m$$

και

$$f(x) = \begin{cases} \dfrac{\lambda_i}{\lambda}, & x = x_i \ \forall i = 1, 2, ..., m \\ 0, & x \neq x_1, x_2, ..., x_m. \end{cases}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.11 *Έστω οι ανεξάρτητες τ.μ. S_1 και M, όπου $S_1 \sim CP(0.2, f_x)$ με $X \sim G_0(1/2)$ και η τ.μ. $M \sim B(3, 0.8)$. Έστω ότι ισχύει $S = S_1 + M$, τότε η συνάρτηση πιθανότητας και η συνάρτηση κατανομής της τ.μ. S για διάφορες τιμές του x, υπολογίζονται με χρήση του πακέτου actuar ως εξής:*

```
library(actuar)

s1 <- rcomppois(n=100,lambda=0.2,model.sev=rgeom(1/2), SIMPLIFY=TRUE)
m <- dbinom(x, size=3, prob=0.8, log = FALSE)

Gs<-aggregateDist("convolution", model.freq=s1, model.sev=m)
mode(Gs) ; x <- 0:9; g <-diff(Gs)[1:10]
mat<-cbind(x,g,Gs(x)) ; colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=6)

      x    gs(x)     Gs(x)
[1,]  0  0.000001  0.000001
[2,]  1  0.000018  0.000019
[3,]  2  0.000295  0.000314
[4,]  3  0.002753  0.003067
[5,]  4  0.016527  0.019594
[6,]  5  0.066161  0.085755
[7,]  6  0.176833  0.262587
[8,]  7  0.305445  0.568032
[9,]  8  0.315809  0.883841
[10,] 9  0.177210  1.000000
```

□

ΠΑΡΑΔΕΙΓΜΑ 5.12 ***(Σύνθετη Poisson-γάμμα κατανομή)***
Έστω η τ.μ. $S \sim CP(\lambda, f)$, όπου $\lambda = 2$ και η τ.μ. $X \sim Gamma(3, 2)$. Η συνάρτηση κατανομής της τ.μ. S με την μέθοδο npower υπολογίζεται ως εξής:

```
library(actuar)
a<-3;b<-2;lambda=2

fx <- discretize(pgamma(x,shape=a,scale=b), from=0, to=100,
                 method="lower")
pn<-ppois(x,lambda)

m1x<-mgamma(1,a,b) ; m2x<-mgamma(2,a,b) ; m3x<-mgamma(3,a,b)
```

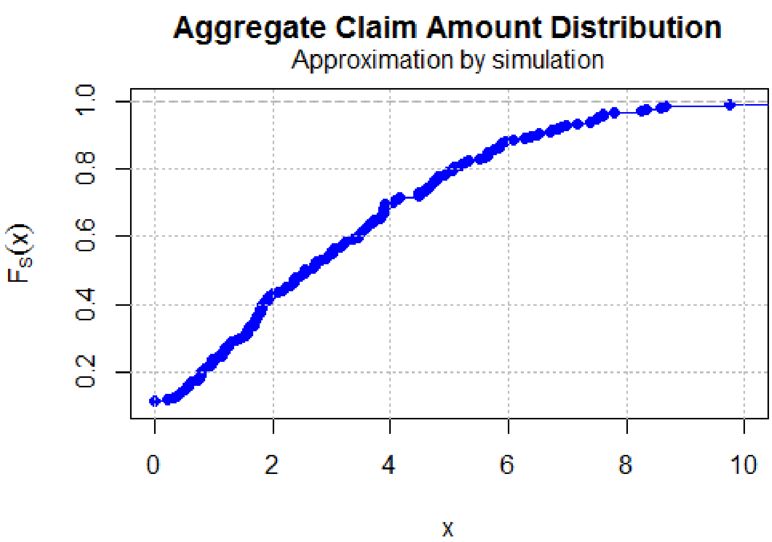
```
varx<-mgamma(2,a,b)-mgamma(1,a,b)^2
m1n<-varn<-lambda
m1S<-m1x*m1n
varS<- m1n*varx+varn*m1x^2
skewS<-(lambda*m3x)/(lambda*m2x)^(3/2)

Gs<-aggregateDist("npower",moments = c(meanS,varS,skewS))
mode(Gs) ; x<-0:20 ;
mat<-cbind(x,Gs(x)) ; colnames(mat) <- c("x","Gs(x)")
round(mat,digits=6)


        x      Gs(x)
 [1,]   0         NA
 [2,]   1         NA
 [3,]   2         NA
 [4,]   3         NA
 [5,]   4   0.702111
 [6,]   5   0.805274
 [7,]   6   0.877843
 [8,]   7   0.926045
 [9,]   8   0.956597
 [10,]  9   0.975216
 [11,] 10   0.986188
```

Όπως είναι εμφανές, η συνάρτηση κατανομής παίρνει τιμές για $x > 3$. Αυτό είναι λογικό, καθώς η προσέγγιση ισχύει μόνο για τιμές όπου $x > E(S)$, στην προκειμένη περίπτωση είναι $E(S) = 3$.

Το ίδιο παράδειγμα θα παρουσιαστεί τώρα με την μέθοδο simulation, ως εξής:

```
library(actuar)
fn <- expression(data = rpois(2))
fx <- expression(data = rgamma(3, 2))

Gs <- aggregateDist("simulation", nb.simul = 200, model.freq=fn,
                    model.sev=fx)
mode(Gs);x<-0:15
g <-diff(Gs)[1:16]
mat<-cbind(x,g,Gs(x))
```

```
colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

        x   gs(x)   Gs(x)
[1,]    0   0.115   0.115
[2,]    1   0.005   0.235
[3,]    2   0.005   0.425
[4,]    3   0.005   0.545
[5,]    4   0.005   0.695
[6,]    5   0.005   0.790
[7,]    6   0.005   0.880
[8,]    7   0.005   0.930
[9,]    8   0.005   0.965
[10,]   9   0.005   0.985
[11,]  10   0.005   0.990
[12,]  11   0.005   0.990
[13,]  12   0.005   0.990
[14,]  13   0.005   0.990
[15,]  14   0.005   0.990
[16,]  15   0.005   1.000

plot(Gs,xlim = c(0,10),ylim =c(0.1,1),col="blue" )
grid(nx=NULL,col="grey")
```



Σχολιασμός: *Σύμφωνα με το διάγραμμα της αθροιστικής συνάρτησης κατανομής της τ.μ. S, οι τιμές που παίρνει η σ.κ. ξεκινάνε από την τιμή 0.115 και αυξάνονται φτάνοντας στην τιμή 1. Αυτό είναι λογικό από την στιγμή που η σ.π.π. της τ.μ. αυξάνεται κατά 0.005, κάποια στιγμή η σ.κ. της να φτάσει στην μονάδα.*

Επιπλέον, όπως είναι εμφανές και από τον ανωτέρω κώδικα, η μέθοδος simulation στις περιπτώσεις όπου υπάρχει συνεχής κατανομή δεν απαιτεί διακριτοποίηση. Η εν λόγω λειτουργία επιστρέφει την εμπειρική συνάρτηση κατανομής των συνολικών απαιτήσεων ενός δείγματος μεγέθους σύμφωνα με την τιμή που ορίστηκε στην εντολή nb.simul.

□

5.4 Μικτές σύνθετες γεωμετρικές κατανομές

Έστω ότι η τ.μ. $N \sim G_0(p),\ \ 0 < p < 1$, δηλαδή η $N \in \{0, 1, 2, ...\}$. Τότε, η τυχαία μεταβλητή $S \sim CG_0(p, f)$, δηλαδή η Σ έχει μια σύνθετη γεωμετρική κατανομή όπου f είναι η συνάρτηση πυκνότητας πιθανότητάς της ή η συνάρτηση πιθανότητας της τ.μ. X Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής N είναι:

$$p_n = P(N = n) = pq^n, \quad n = 0, 1, 2, ..., \quad q = 1 - p$$

και η πιθανογεννήτρια συνάρτησή της, είναι

$$P_N(u) = \frac{p}{1 - qu}, \quad |u| < \frac{1}{q}.$$

Από εδώ και στο εξής θεωρείται ότι η τυχαία μεταβλητή $X \geq 0$ είναι συνεχής. Τοτε, επειδή $p_0 = p \neq 0$, έπεται ότι η σύνθετη γεωμετρική τ.μ. S είναι πάντοτε μικτού τύπου με μάζα πιθανότητας στο σημείο μηδέν και συνεχής στο διάστημα $(0, \infty)$.

Στις υποενότητες που ακολουθούν, θα παρουσιαστούν αναλυτικά αποτελέσματα υπολογισμού της κατανομής της τ.μ. S στο μοντέλο συλλογικού κινδύνου, όταν η τ.μ. $N \sim G_0(p)$ και για συγκεκριμένες κατανομές του μεγέθους ατομικής ζημιάς X.

5.4.1 Πρώτη μέθοδος υπολογισμού της $\bar{G}(x)$

ΘΕΩΡΗΜΑ 5.4 *Θεωρούμε το μοντέλο συλλογικού κινδύνου*

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

όπου η τ.μ. $N \sim G_0(p), \quad 0 < p < 1.$
Τότε ισχύει:

(α) $G(x) = p + qF_Y(x)$ $\;\kappa\alpha\iota\;$ $\bar{G}(x) = q\bar{F}_Y(x)$ ***για*** $x \geq 0,\; q = 1 - p$

(β) $g(x) = \begin{cases} p, & x = 0 \\ qf_Y(x), & x > 0, \end{cases}$ ***με*** $g(0) = P(S = 0),$

όπου Υ είναι μια συνεχής θετική τυχαία μεταβλητή με ροπογεννήτρια συνάρτηση

$$M_Y(t) = \frac{pM_X(t)}{1 - qM_X(t)}. \tag{5.2}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.13 ***(Σύνθετη γεωμετρική–εκθετική κατανομή)***
Έστω ότι το πλήθος των κινδύνων είναι τ.μ. $N \sim G_0(p),\; 0 < p < 1$, *και το μέγεθος της ατομικής ζημιάς είναι τ.μ.* $X \sim Exp(\lambda),\; \lambda > 0.$
Η ροπογεννήτρια συνάρτηση της τ.μ. Χ, είναι

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda,$$

τότε από την Σχέση (5.2) ισχύει ότι

$$M_Y(t) = \frac{pM_X(t)}{1 - qM_X(t)} = \frac{\frac{p\lambda}{\lambda - t}}{1 - \frac{q\lambda}{\lambda - t}} = \frac{p\lambda}{\lambda - t - q\lambda} = \frac{p\lambda}{(1-q)\lambda - t} = \frac{p\lambda}{p\lambda - t}.$$

Επομένως, η τ.μ. $Y \sim Exp(p\lambda)$, οπότε από το Θεώρημα 5.4 έπεται ότι η συνάρτηση κατανομής της τ.μ. S είναι

$$G(x) = p + q(1 - e^{-p\lambda x}) = p + q - qe^{-p\lambda x} = 1 - qe^{-p\lambda x}, \;\; x \geq 0.$$

και η συνάρτηση πυκνότητας πιθανότητας, θα είναι

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

Για $p = 0.5$ και $\lambda = 3$, είναι:

```
library(actuar)
fx <- discretize(pexp(x,3), from=0, to=100, step=1, method="lower")
Gs <- aggregateDist("recursive", model.freq="geometric",
                    model.sev=fx,prob=0.5)
mode(Gs); x <- 0:10; g <-diff(Gs)[1:11]
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

        x   gs(x)    Gs(x)
[1,]    0 0.50000  0.50000
[2,]    1 0.23755  0.73755
[3,]    2 0.12469  0.86224
[4,]    3 0.06545  0.92769
[5,]    4 0.03435  0.96205
[6,]    5 0.01803  0.98008
[7,]    6 0.00946  0.98954
[8,]    7 0.00497  0.99451
[9,]    8 0.00261  0.99712
[10,]   9 0.00137  0.99849
[11,]  10 0.00072  0.99921
```

□

Επομένως, σύμφωνα με το παραπάνω παράδειγμα έπεται ότι

"η σύνθετη γεωμετρική(p)-εκθετική(λ) κατανομή, είναι η διακριτή μείξη της εκφυλισμένης στο σημείο μηδέν κατανομής και της εκθετικής$(p\lambda)$ κατανομής με βάρη μείξης p και $q = 1 - p$ αντίστοιχα. Ισοδύναμα, η σύνθετη γεωμετρική(p)-εκθετική(λ) τ.μ. είναι ισόνομη με μια σύνθετη Bernoulli$(1-p)$-εκθετική$(p\lambda)$ τ.μ.". [9]

5.4.2 Δεύτερη μέθοδος υπολογισμού της $\bar{G}(x)$

ΛΗΜΜΑ 5.1 *Έστω h η συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς μη-αρνητικής τυχαίας μεταβλητής και*

$$\bar{H}(x) = \int_x^{\infty} h(y)dy$$

η συνάρτηση δεξιάς ουράς της.
Επιπλέον,

$$\hat{h}(s) = \int_0^{\infty} e^{-sx} h(x)dx \quad \kappa\alpha\iota \quad \hat{\bar{H}}(s) = \int_0^{\infty} e^{-sx} \bar{H}(x)dx$$

οι μετασχηματισμοί Laplace των συναρτήσεων h και $\bar{H}$ αντίστοιχα.
Τότε, ισχύει ότι

$$\hat{\bar{H}}(s) = \frac{1 - \hat{h}(s)}{s}. \tag{5.3}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης Ε., 2015).*

Έστω,

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} f(x)dx \quad \kappa\alpha\iota \quad \hat{\bar{F}}(s) = \int_0^{\infty} e^{-sx} \bar{F}(x)dx$$

οι μετασχηματισμοί Laplace της συνάρτησης πυκνότητας πιθανότητας και της συνάρτησης δεξιάς ουράς αντίστοιχα του μεγέθους ατομικής ζημιάς X, και

$$\hat{g}(s) = \int_0^{\infty} e^{-sx} g(x)dx \quad \kappa\alpha\iota \quad \hat{\bar{G}}(s) = \int_0^{\infty} e^{-sx} \bar{G}(x)dx$$

οι μετασχηματισμοί Laplace της συνάρτησης πυκνότητας πιθανότητας και της συνάρτησης δεξιάς ουράς αντίστοιχα του μεγέθους συνολικής ζημιάς S για το μοντέλο συλλογικού κινδύνου.

Τότε από το Λήμμα 5.1 έπεται, ότι:

$$\hat{\bar{F}}(s) = \frac{1 - \hat{f}(s)}{s}, \tag{5.4}$$

και

$$\hat{\bar{G}}(s) = \frac{1 - \hat{g}(s)}{s}. \tag{5.5}$$

ΘΕΩΡΗΜΑ 5.5 *Ο μετασχηματισμός Laplace $\hat{\bar{G}}(s)$ της συνάρτησης δεξιάς ουράς $\bar{G}(x)$ της σύνθετης γεωμετρικής τυχαίας μεταβλητής $S = X_1 + X_2 + ... + X_N$ (με $S = 0$ αν $N = 0$) είναι,*

$$\hat{\bar{G}}(s) = \frac{q[1 - \hat{f}(s)]}{s[1 - q\hat{f}(s)]}. \tag{5.6}$$

*(**Απόδειξη:** βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.14 ***(Σύνθετη γεωμετρική-εκθετική κατανομή)***

Το πλήθος των κινδύνων είναι τ.μ. $N \sim G_0(p),\ 0 < p < 1$, και το μέγεθος ατομικής ζημιάς είναι τ.μ. $X \sim Exp(\lambda),\ \lambda > 0$.

Επειδή ο μετασχηματισμός Laplace της συνάρτησης πυκνότητας πιθανότητας της τ.μ. X είναι

$$\hat{f}(s) = \frac{\lambda}{\lambda + s},$$

από την Σχέση (5.6) έπεται ότι

$$\hat{\bar{G}}(s) = \frac{q[1 - \frac{\lambda}{\lambda+s}]}{s[1 - q\frac{\lambda}{\lambda+s}]} = \frac{qs}{s[\lambda + s - q\lambda]} = \frac{q}{\lambda(1-q)s} = \frac{q}{p\lambda + s}.$$

Η παραπάνω σχέση γράφεται ισοδύναμα στην μορφή

$$\int_0^\infty e^{-sx}\bar{G}(x)dx = q\int_0^\infty e^{-sx}e^{-p\lambda x}dx$$
$$\Rightarrow \quad \bar{G}(x) = qe^{-p\lambda x}, \quad x \geq 0.$$

Επειδή η τ.μ. S έχει μάζα πιθανότητας στο σημείο μηδέν με $P(S=0) = P(N=0) = p$ *και είναι συνεχής στο διάστημα* $(0,\infty)$*, η συνάρτηση πυκνότητας πιθανότητας της τ.μ. S είναι*

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

□

ΟΡΙΣΜΟΣ 5.1 *Η κατανομή της συνεχούς τ.μ. X θα ανήκει στην κλασματική ή ρητή οικογένεια κατανομών, αν ο μετασχηματισμός Laplace* $\hat{f}(s)$ *της συνάρτησης πυκνότητας πιθανότητάς της* $f(x)$ *είναι της μορφής*

$$\hat{f}(s) = \frac{h_{n-1}(s)}{q_n(s)},$$

όπου:

$q_n(s)$: *είναι πολυώνυμο ως προς s*

$h_{n-1}(s)$: *είναι πολυώνυμο ως προς s*

$h_{n-1}(0) = q_n(0)$.

Η κλασματική οικογένεια είναι μια ευρεία κλάση κατανομών που μεταξύ άλλων περιλαμβάνει και τις παρακάτω κατανομές:

- Εκθετικές κατανομές
- Erlang κατανομές
- Coxian κατανομές
- Phase-type κατανομές,

καθώς και τις μείξεις αυτών.

Συνεπώς για τον υπολογισμό της συνάρτησης δεξιάς ουράς μιας σύνθετης τυχαίας μεταβλητής αρχικά πρέπει να βρεθεί ο μετασχηματισμός Laplace της συνάρτησης δεξιάς ουράς $\bar{G}(x) = P(S > x)$ μέσω της σχέσης

$$\hat{\bar{G}}(x) = \frac{1 - \hat{g}(s)}{s} = \frac{1 - P_N(\hat{f}(s))}{s}$$

και στην συνέχεια να γραφτεί ως πηλίκο πολυωνύμων ως προς s

$$\hat{\bar{G}}(s) = \frac{A(s)}{B(s)}.$$

Επειδή το $s = 0$ είναι κοινή ρίζα του αριθμητή και του παρονομαστή, το πολυώνυμο γίνεται

$$\hat{\bar{G}}(s) = \frac{A_1(s)}{B_1(s)}.$$

Έστω ότι $p_1, p_2, ..., p_n$ είναι οι ρίζες της εξίσωσης $B_1(s) = 0$ με $p_i \neq p_j$ για $i, j \in \{1, 2, ..., n\}$.

Εφαρμόζοντας την τεχνική μερικών κλασμάτων για το πηλίκο των πολυωνύμων, έπεται

$$\frac{A_1(s)}{B_1(s)} = \frac{\alpha_1}{s - p_1} + \frac{\alpha_2}{s - p_2} + ... + \frac{\alpha_n}{s - p_n}.$$

Άρα σύμφωνα με τα παραπάνω, ισχύει ότι

$$\hat{\bar{G}} = \frac{\alpha_1}{s - p_1} + \frac{\alpha_2}{s - p_2} + ... + \frac{\alpha_n}{s - p_n}$$

και $\dfrac{1}{s - p_i}$ είναι ο μετασχηματισμός Laplace της συνάρτησης $e^{p_i x}$, από την παραπάνω σχέση έπεται τελικά ότι:

$$\bar{G}(x) = \alpha_1 e^{p_1 x} + \alpha_2 e^{p_2 x} + ... + \alpha_n e^{p_n x}, \quad x \geq 0.$$

ΠΑΡΑΔΕΙΓΜΑ 5.15 ***(Σύνθετη γεωμετρική-συνδυασμός εκθετικών)***

Έστω ότι το πλήθος των ζημιών είναι τ.μ. $N \sim G_0(p),\ 0 < p < 1$ και το μέγεθος ατομικής ζημιάς X έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \theta\alpha e^{-\alpha x} + (1-\theta)\beta e^{-\beta x}, \quad x > 0,\ \alpha > \beta,\ \theta > 0.$$

Αν $S = X_1 + X_2 + ... + X_N$ (με $S = 0$ αν $N = 0$) είναι το μέγεθος συνολικής ζημιάς, τότε η συνάρτηση δεξιάς ουράς δίνεται από τον τύπο

$$\bar{G}(x) = \alpha_1 e^{-p_1 x} + \alpha_2 e^{-p_2 x}, \quad x \geq 0$$

και η συνάρτηση πυκνότητας πιθανότητας της τ.μ. S, είναι

$$g(x) = \begin{cases} p, & x = 0, \\ \alpha_1 p_1 e^{-p_1 x} + \alpha_2 p_2 e^{-p_2 x}, & x > 0, \end{cases}$$

όπου,

$$\alpha_1 = q\frac{(1-\theta)\alpha + \theta\beta - p_1}{p_2 - p_1}, \quad \alpha_2 = q\frac{p_2 - (1-\theta)\alpha - \theta\beta}{p_2 - p_1},$$

και $-p_1, -p_2$ είναι οι διακριτές πραγματικές αρνητικές ρίζες της εξίσωσης

$$s^2 + [(1-q\theta)\alpha + (p+q\theta)\beta]s + p\alpha\beta = 0.$$

Έστω ότι $\alpha = 2,\ \beta = 3,\ \theta = 1/2$ και $p = 1/5$, τότε

```
library(actuar)
fx<-rmixture(n=100,probs=1/2, models=expression(rexp(2), rexp(3)))
Gs <- aggregateDist("recursive", model.freq="geometric", model.sev=fx,
                    prob=1/5)
mode(Gs); x <- 0:5; g <-diff(Gs)[1:6]
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)


      x   gs(x)    Gs(x)
[1,]  0  0.24302  0.24302
[2,]  1  0.02170  0.26472
[3,]  2  0.04563  0.31036
```

```
[4,]  3  0.04918  0.35953
[5,]  4  0.06678  0.42632
[6,]  5  0.32017  0.74649
```

$\square$

5.4.3 Τρίτη μέθοδος υπολογισμού της $\bar{G}(x)$

ΘΕΩΡΗΜΑ 5.6 *Θεωρούμε το μοντέλο συλλογικού κινδύνου*

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

όπου η τ.μ. $N \sim G_0(p), \quad 0 < p < 1.$

Τότε η συνάρτηση δεξιάς ουράς $\bar{G} = P(S > x)$ *ικανοποιεί την ολοκληρωτική εξίσωση*

$$\bar{G}(x) = q \int_0^x \bar{G}(x-y) f(y) dy + q\bar{F}(x), \quad x \geq 0. \tag{5.7}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.16 ***(Σύνθετη γεωμετρική-εκθετική κατανομή)***

Το πλήθος των κινδύνων είναι τ.μ. $N \sim G_0(p), \; 0 < p < 1$ *και το μέγεθος ατομικής ζημιάς είναι τ.μ.* $X \sim Exp(\lambda), \; \lambda > 0.$

Ισχύει ότι:

$$\begin{aligned} \bar{G}(x) &= q \int_0^x \bar{G}(t) f(x-t)(dt) + q\bar{F}(x) \\ &= q \int_0^x \bar{G}(t) \lambda e^{-\lambda(x-t)} dt + q e^{-\lambda x}. \end{aligned}$$

Ύστερα από πράξεις, έπεται ότι:

$$\bar{G}(x) = q e^{-\lambda p x}, \quad x \geq 0.$$

Για $p = 0.5$ *και* $\lambda = 3$, *είναι:*

```
library(actuar)
fx <- discretize(pexp(x,3),from=0,to=100,step=1,method="lower")
Gs <- aggregateDist("recursive", model.freq="geometric", model.sev=fx,
                    prob=0.5)
mode(Gs); x <- 0:10
Gbar<-1-Gs(x)[1:11]
mat<-cbind(x,Gbar); colnames(mat) <- c("x","Gbar(x)")
round(mat,digits=5)

       x  Gbar(x)
[1,]   0  0.50000
[2,]   1  0.26245
[3,]   2  0.13776
[4,]   3  0.07231
[5,]   4  0.03795
[6,]   5  0.01992
[7,]   6  0.01046
[8,]   7  0.00549
[9,]   8  0.00288
[10,]  9  0.00151
[11,] 10  0.00079
```

□

ΛΗΜΜΑ 5.2 *Υποθέτουμε ότι η συνάρτηση* $h(x)$ *ικανοποιεί την διαφορική εξίσωση*

$$h^{''}(x) + \alpha h^{'}(x) + \beta h(x) = 0. \tag{5.8}$$

Αν η εξίσωση $p^2 + \alpha p + \beta = 0$ *έχει δύο πραγματικές ρίζες, έστω τις* p_1 *και* p_2, *τότε η λύση της διαφορικής εξίσωσης δίνεται από την σχέση*

$$h(x) = A_1 e^{p_1 x} + A_2 e^{p_2 x},$$

όπου A_1 *και* A_2 *είναι δύο σταθερές..*

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.17 *[9]* ***(Σύνθετη γεωμετρική-μείξη εκθετικών κατανομών)***
Το πλήθος των κινδύνων είναι τ.μ. $N \sim G_0(p),\ 0 < p < 1$ *με* $E(N) = 2$ *και το μέγεθος ατομικής ζημιάς είναι τ.μ.* X *με συνάρτηση πυκνότητας πιθανότητας*

$$\begin{aligned} f(x) &= e^{-3x} + \frac{10}{3}e^{-5x} \\ &= \frac{1}{3}3e^{-3x} + \frac{2}{3}5e^{-5x}, \quad x > 0. \end{aligned}$$

Δηλαδή η τ.μ. X είναι μια διακριτή μείξη εκθετικών κατανομών με $Exp(3)$ *και* $Exp(5)$ *με βάρη μείξης* $1/3$ *και* $2/3$ *αντίστοιχα. Τότε η συνάρτηση δεξιάς ουράς της τ.μ. X δίνεται από τον τύπο*

$$\bar{F}(x) = \frac{1}{3}e^{-3x} + \frac{2}{3}e^{-5x}, \quad x \geq 0.$$

Επειδή $E(N) = \dfrac{q}{p} = 2 \ \Rightarrow\ q = \dfrac{2}{3}$, *τότε ισχύει:*

$$\bar{G}(x) = \frac{2}{3}\int_0^x \bar{G}(t)\left(e^{-3(x-t)} + \frac{10}{3}e^{-5(x-t)}\right)dt + \frac{2}{3}\left(\frac{1}{3}e^{-3x} + \frac{2}{3}e^{-5x}\right)$$

Ύστερα από πράξεις προκύπτει ότι

$$\bar{G}(x) = \frac{\sqrt{31}-5}{3\sqrt{31}}e^{-\frac{23+2\sqrt{31}}{9}x} + \frac{\sqrt{31}+5}{3\sqrt{31}}e^{-\frac{23-2\sqrt{31}}{9}x}, \quad x \geq 0.$$

και η συνάρτηση πυκνότητας πιθανότητας είναι

$$g(x) = \begin{cases} \dfrac{1}{3}, & x = 0, \\ \dfrac{13\sqrt{31}-53}{27\sqrt{31}}e^{-\frac{23+2\sqrt{31}}{9}x} + \dfrac{13\sqrt{31}+53}{27\sqrt{31}}e^{-\frac{23-2\sqrt{31}}{9}x}, & x > 0. \end{cases}$$

```
library(actuar)
fx<-rmixture(n=100, probs=1/3, expression(rexp(3), models=rexp(5)))
Gs <- aggregateDist("recursive", model.freq="geometric", model.sev=fx,
                    prob=1/3)
mode(Gs); x <- 0:2 ; g <-diff(Gs)[1:3]
```

```
mat<-cbind(x,g,Gs(x)) ; colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x    gs(x)     Gs(x)
[1,] 0   0.60733   0.60733
[2,] 1   0.14186   0.74919
[3,] 2   0.25610   1.00000
```

$\square$

ΠΑΡΑΔΕΙΓΜΑ 5.18 *Ένα ανομοιογενές χαρτοφυλάκιο κινδύνων υποδιαιρείται σε δύο ομοιογενή υπόχαρτοφυλάκια, έτσι ώστε*

- *το πλήθος των κινδύνων για το $i-$υποχαρτοφυλάκιο είναι τ.μ. N_i με συνάρτηση πιθανότητας $P(N_i = n) = \dfrac{1}{2^{n+2-i}}$, $n = i-1, i, i+1$ με $i = 1, 2$.*
- *το μέγεθος ατομικής ζημιάς για το $i-$υποχαρτοφυλάκιο είναι τ.μ. X που ακολουθεί την εκθετική κατανομή με παράμετρο $2i$, $i = 1, 2$.*

Ισχύει,

$$S_1 = \begin{cases} 0, & N_1 = 0 \\ X_1^{(1)} + X_2^{(1)} + ... + X_N^{(1)}, & N_1 \geq 1, \end{cases}$$

$$S_2 = \begin{cases} 0, & N_2 = 0 \\ X_1^{(2)} + X_2^{(2)} + ... + X_N^{(2)}, & N_2 \geq 1. \end{cases}$$

Δηλαδή,

$$S = S_1 + S_2.$$

Είναι,

$$P(N_1 = n) = \frac{1}{2^{n+2-1}} = \frac{1}{2^{n+1}} = \frac{1}{2} \cdot \frac{1}{2^n} = \frac{1}{2} \cdot \left(\frac{1}{2}\right)^n \quad \Rightarrow \quad S_1 \sim G_0(1/2)$$
$$P(N_2 = n) = \frac{1}{2^{n+2-2}} = \frac{1}{2^n} = 2 \cdot \frac{1}{2} \cdot \frac{1}{2^n} = \frac{1}{2} \cdot \left(\frac{1}{2}\right)^{n-1} \quad \Rightarrow \quad S_2 \sim G_1(1/2)$$

και

$$X_1 \sim Exp(2)$$
$$X_2 \sim Exp(4).$$

Είναι,

$$\begin{aligned} P_S(u) &= P_{S_1}(u) \cdot P_{S_2}(u) \\ &= P_{N_1}(\hat{f}_{X_1}(s)) \cdot P_{N_2}(\hat{f}_{X_2}(s)) \\ &= ... = \frac{1}{s+1}, \end{aligned}$$

$$\hat{\bar{G}}(s) = \frac{1-\hat{g}(s)}{s} = ... = \frac{1}{s+1}.$$

Ισχύει ότι

$$\begin{aligned} &\frac{1}{s+\alpha} = \int_0^\infty e^{-sx} e^{\alpha x} dx \quad \kappa\alpha\iota \\ &\hat{\bar{G}} = \int_0^\infty e^{-sx} \bar{G}(x) dx \quad \Rightarrow \quad \bar{G}(x) = e^{-x}. \end{aligned}$$

Άρα $G(x) = 1 - \bar{G}(x) = 1 - e^{-x}$*, δηλαδή η τ.μ.* $S \sim Exp(1)$.

```
library(actuar)

s1 <- rcompound(n=10, model.freq=rgeom(1/2), model.sev=rexp(2),
                SIMPLIFY = TRUE)
s 2<- rcompound(n=10, model.freq=rztgeom(1/2), model.sev=rexp(4),
                SIMPLIFY = TRUE)

Gs <- aggregateDist("convolution", model.freq=s1, model.sev=s2)
mode(Gs); x <- 0:1; g <-diff(Gs)[1:2];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x    gs(x)    Gs(x)
[1,] 0   0.43350  0.4335
[2,] 1   0.96138  1.0000
```

□

5.5 Σύνθετες διακριτές γεωμετρικές κατανομές

ΘΕΩΡΗΜΑ 5.7 *Θεωρούμε το μοντέλο συλλογικού κινδύνου*

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

όπου η τυχαία μεταβλητή $N \sim G_0(p),\ \ 0 < p < 1$, *και η τ.μ.* $X \in \{0, 1, 2, ...\}$. *Τότε ισχύει ότι:*

(α**)** $g(x) = \begin{cases} p + qf_Y(0), & x = 0, \\ qf_Y(x) & x = 1, 2, 3, ... \end{cases}$

(β**)** $G(x) = p + qF_Y(x)$ ***και*** $\bar{G}(x) = q\bar{F}_Y(x)$ ***για*** $x = 0, 1, 2, ...,\ \ q = 1 - p$,

όπου Υ είναι μια μη-αρνητική ακέραια τ.μ. με πιθανογεννήτρια συνάρτηση

$$P_Y(u) = \frac{pP_X(u)}{1 - qP_X(u)}. \tag{5.9}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Έστω ότι η συνάρτηση

$$R_S(u) = \sum_{x=0}^{\infty} \bar{G}(x)u^x \tag{5.10}$$

είναι η γεννήτρια συνάρτηση της ακολουθίας αριθμών $\{\bar{G}(x) : x \geq 0\}$.

Τότε, η συνάρτηση που συνδέει την γεννήτρια συνάρτηση της παραπάνω ακολουθίας με την πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής S δίνεται από την σχέση:

$$R_S(u) = \frac{1 - P_S(u)}{1 - u}. \tag{5.11}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΟΡΙΣΜΑ 5.4 *Έστω η διακριτή σύνθετη γεωμετρική τ.μ.* $S = X_1 + X_2 + ... + X_N$ *(με* $S = 0$ *αν* $N = 0$*), όπου η τ.μ.* $N \sim G_0(p),\ 0 < p < 1,\ q = 1 - p$*. Τότε, ισχύει ότι*

$$R_S(u) = \frac{q[1 - P_X(u)]}{(q - u)[q - P_X(u)]}. \tag{5.12}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

ΠΑΡΑΔΕΙΓΜΑ 5.19 ***(Σύνθετη γεωμετρική-γεωμετρική κατανομή)***

Θεωρούμε το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N,\ N \geq 1$ *(με* $S = 0$ *αν* $N = 0$*), όπου η τ.μ.* $N \sim G_0(p),\ 0 < p < 1$ *και η τ.μ.* $X \sim G_0(\theta),\ 0 < \theta < 1$*. Ισχύει ότι*

$$\bar{G}(x) = \frac{q(1-\theta)^{x+1}}{(1-q\theta)^{x+1}}, \quad x = 0, 1, 2, ...$$

και

$$g(x) = \begin{cases} \dfrac{p}{1-q\theta}, & x = 0 \\ \dfrac{pq\theta(1-\theta)^x}{(1-q\theta)^{x+1}}, & \text{ξ=1,2,3...} \end{cases}$$

Για $p = 0.5$ *και* $\theta = 0.2$*, είναι*

```
library(actuar)
fx <- pgeom(x,0.5,lower.tail = TRUE,log.p = FALSE)
Gs <- aggregateDist("recursive", model.freq="geometric", model.sev=fx,
                    prob=0.2)
mode(Gs) ; x <- 0:2 ; g <-diff(Gs)[1:3]
mat<-cbind(x,g,Gs(x)) ; colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x    gs(x)    Gs(x)
[1,] 0  0.33333  0.33333
[2,] 1  0.33333  0.66667
[3,] 2  0.72222  1.00000
```

□

Από τα παραπάνω, προκύπτει ότι:

Η σύνθετη διακριτή γεωμετρική(p) - γεωμετρική(θ) κατανομή είναι η διακριτή μείξη της εκφυλισμένης κατανομής στο σημείο μηδέν και της γεωμετρικής ($p^{'}$) κατανομής με βάρη μείξης p και q αντίστοιχα [9],

όπου $p^{'} = \dfrac{p\theta}{1-q\theta}$

Επίσης, ισχύει ότι:

$$\begin{aligned} P_S(u) &= P_N(P_X(u)) \\ &= \frac{p}{1-qP_X(u)} \\ &= p + \frac{p}{1-qP_X(u)} - p \\ &= p + q\frac{pP_X(u)}{1-qP_X(u)}, \end{aligned}$$

δηλαδή είναι

$$P_S(u) = p + qP_Y(u) \tag{5.13}$$

και έστω ότι η τ.μ. $Y \sim G_0(p^{'})$, έπεται ότι:

Η σύνθετη διακριτή γεωμετρική(p) - γεωμετρική(θ) κατανομή είναι ισόνομη με μια σύνθετη Bernoulli (q) - γεωμετρική ($p^{'}$) κατανομή. [9]

Επιπλέον, είναι

$$\begin{aligned} g(0) &= \frac{p}{1-q\theta} \\ \Rightarrow 1 - g(0) &= 1 - \frac{p}{1-q\theta} = \frac{1-p-q\theta}{1-q\theta} = \frac{q(1-\theta)}{1-q\theta} \end{aligned}$$

Τότε, για $x = 1, 2, 3, ...$, η συνάρτηση πιθανότητας της τ.μ. S δίνεται από την σχέση

$$\begin{aligned} g(x) &= \frac{pq\theta(1-\theta)^x}{(1-q\theta)^{x+1}} = \frac{q(1-\theta)}{1-q\theta}\frac{p\theta}{1-q\theta}\left(\frac{1-\theta}{1-q\theta}\right)^{x-1} \\ \Rightarrow g(x) &= (1-g(0))p^{'}(1-p^{'})^{x-1}. \end{aligned}$$

Άρα,

Η σύνθετη γεωμετρική(p) - γεωμετρική(θ) κατανομή είναι ισόνομη με μια τροποποιημένη στο σημείο μηδέν (zero-modified) γεωμετρική ($p^{'}$) κατανομή. [9]

ΠΑΡΑΔΕΙΓΜΑ 5.20 ***(Σύνθετη γεωμετρική-αρνητική διωνυμική)***
Θεωρούμε το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ...X_N,\ N \geq 1$ *(με* $S = 0$ *αν* $N = 0$*), όπου η τ.μ.* $N \sim G_0(p),\ 0 < p < 1$ *και η τ.μ.* $X \sim NB(2, \theta),\ 0 < \theta < 1.$

Η πιθανογεννήτρια συνάρτηση της τ.μ. X είναι

$$P_X(u) = \left(\frac{\theta}{1-(1-\theta)u}\right)^2, \quad |u| < \frac{1}{1-\theta}$$

Ύστερα από πράξεις προκύπτει ότι

$$R_S(u) = \frac{\frac{q(1-\theta^2)}{(1-\theta)^2} - qu}{u^2 - \frac{2}{1-\theta}u + \frac{1-q\theta^2}{(1-\theta)^2}} \tag{5.14}$$

Χρησιμοποιώντας την τεχνική μερικών κλασμάτων, ύστερα από πράξεις έπεται ότι

$$\bar{G}(x) = \frac{\sqrt{q}+q}{2}\frac{1}{p_2^{x+1}} - \frac{\sqrt{q}-q}{2}\frac{1}{p_1^{x+1}}$$

όπου p_1, p_2 *οι διακριτές πραγματικές ρίζες της Σχέσης 5.14.*

Η συνάρτηση πιθανότητας της τ.μ. S, είναι

$$g(x) = \begin{cases} P_N(f(0)), & x = 0 \\ \bar{G}(x-1) - \bar{G}(x), & x = 1, 2, ... \end{cases}$$

όπου

$$P_N(f(0)) = \frac{p}{1-q\theta^2}$$

και για $x = 1, 2, 3, ...$

$$\bar{G}(x-1) - \bar{G}(x) = \frac{p\theta\sqrt{q}(1-\theta)^x}{2(1-\theta\sqrt{q})^{x+1}} - \frac{p\theta\sqrt{q}(1-\theta)^x}{2(1+\theta\sqrt{q})^{x+1}}.$$

Για $p = 1/2$ και $\theta = 1/2$ είναι:

```
library(actuar)
fx <- pnbinom(x,size=2,prob=1/2)
Gs <- aggregateDist("recursive", model.freq="geometric", model.sev=fx,
                    prob=1/2)
mode(Gs); x <- 0:2; g <-diff(Gs)[1:3];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x   gs(x)    Gs(x)
[1,] 0  0.57143  0.57143
[2,] 1  0.16327  0.73469
[3,] 2  0.27114  1.00000
```

□

ΟΡΙΣΜΟΣ 5.2 *Η κατανομή της διακριτής τ.μ. Χ θα ανήκει στη διακριτή κλασματική ή διακριτή ρητή οικογένεια κατανομών, αν η πιθανογεννήτρια συνάρτηση $P_X(u)$ της συνάρτησης πιθανότητάς της f(x) είναι της μορφής*

$$P_X(u) = \frac{h_{n-1}(u)}{q_n(u)}.$$

□

Από τον ανωτέρω ορισμό έπεται ότι η κατανομή της τ.μ Χ θα ανήκει στην διακριτή κλασματική ή διακριτή ρητή οικογένεια κατανομών αν η πιθανογεννήτρια συνάρτηση $P_X(u)$ είναι πηλίκο πολυωνύμων που έχουν την μονάδα ως κοινή ρίζα (ισχύει ότι $h_{n-1}(1) = q_n(1)$) και ο βαθμός του πολυωνύμου του αριθμητή είναι μικρότερος του βαθμού του πολυωνύμου του παρονομαστή.

Σύμφωνα τώρα με το Πόρισμα 5.4, μπορεί να βρεθεί μια αναδρομική σχέση υπολογισμού της συνάρτησης δεξιάς ουράς μιας σύνθετης γεωμετρικής τυχαίας μεταβλητής S, σύμφωνα με το ακόλουθο Πόρισμα.

ΠΟΡΙΣΜΑ 5.5 *Έστω η διακριτή σύνθετη γεωμετρική τ.μ.* $S = X_1 + X_2 + ... + X_N$ *(με* $S = 0$ *αν* $N = 0$*), όπου η τ.μ.* $N \sim G_0(p),\ 0 < p < 1,\ q = 1 - p$*. Τότε, η συνάρτηση δεξιάς ουράς ικανοποιεί την αναδρομική σχέση*

$$\bar{G}(x) = \frac{q}{1 - qf(0)} \sum_{y=1}^{x} \bar{G}(x-y)f(y) + \frac{q}{1 - qf(0)} \bar{F}(x), \quad x = 0, 1, 2, ... \tag{5.15}$$

*(**Απόδειξη**: βλ. Χατζηκωνσταντινίδης, 2015).*

Για $x = 0$, η Σχέση (5.15) γίνεται:

$$\bar{G}(0) = \frac{q\bar{F}(0)}{1 - q\bar{f}(0)}.$$

Προφανώς, η τιμή μπορεί να υπολογισθεί και μέσω της

$$\begin{aligned} \bar{G}(0) &= 1 - G(0) = 1 - g(0) = 1 - P_N(f(0)) = 1 - \frac{p}{1 - qf(0)} \\ &= \frac{q[1 - f(0)]}{1 - qf(0)} = \frac{q[1 - F(0)]}{1 - qf(0)} = \frac{q\bar{F}(0)}{1 - qf(0)}. \end{aligned}$$

5.6 Σύνθετες τροποποιημένες και περικομμένες στο μηδέν γεωμετρικές κατανομές

ΘΕΩΡΗΜΑ 5.8 *Έστω* $0 \leq p_0 < 1,\ 0 < p < 1$*. Η διακριτή τ.μ.* $N \in \{0, 1, 2, ...\}$ *θα λέμε ότι έχει την τροποποιημένη στο σημείο μηδέν γεωμετρική κατανομή (zero-modified) με παραμέτρους* p_0 *και p, αν η συνάρτηση πιθανότητάς της, είναι*

$$P(N = n) = \begin{cases} p_0, & n = 0, \\ (1 - p_0)pq^{n-1}, & n = 1, 2, 3, ... \end{cases}$$

όπου $q = 1 - p$

Τότε, συμβολικά γράφουμε ότι η τ.μ. $N \sim ZMG(p_0, p)$.

Η τροποποιημένη στο σημείο μηδέν γεωμετρική κατανομή αποτελεί γενίκευση της γεωμετρικής κατανομής $G_0(p)$ και της περικομμένης στο σημείο μηδέν γεωμετρικής κατανομής $G_1(p)$ ή $ZTG(p)$. Με άλλα λόγια:

(α) Για $p_0 = p$, η συνάρτηση πυκνότητας πιθανότητας της τ.μ. N είναι

$$P(N=n) = \begin{cases} p, & n=0, \\ pq^n, & n=1,2,... \end{cases}$$

οπότε ισχύει ότι

$$P(N=n) = pq^n, \quad n=0,1,2,...$$

Επομένως, η τ.μ. $N \sim G_0(p)$.

(β) Για $p_0 = 0$, η συνάρτηση πυκνότητας πιθανότητας της τ.μ. N είναι

$$P(N=n) = \begin{cases} 0, & n=0, \\ pq^{n-1}, & n=1,2,... \end{cases}$$

οπότε ισχύει ότι

$$P(N=n) = pq^{n-1}, \quad n=1,2,...$$

Επομένως, η τ.μ. $N \sim G_1(p)$.

5.6.1 Σύνθετες τροποποιημένες στο μηδέν γεωμετρικές κατανομές

Θεωρούμε το μοντέλο συλλογικού κινδύνου

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

όπου η τυχαία μεταβλητή $N \sim ZMG(p_0, p)$ με $0 \leq p_0 < 1,\ 0 < p < 1$.

Τότε θα λέμε ότι η S ακολουθεί την σύνθετη τροποποιημένη στο μηδέν γεωμετρική κατανομή και θα γράφεται συμβολικά $S \sim CZMG(p_0, p; f)$, όπου CZMG είναι τα αρχικά των λέξεων Compound Zero Modified Geometric.

1. **Η μη-αρνητική τ.μ. X είναι συνεχής**

Αν $p_0 \neq 0$, επειδή $P(N=0) = p_0 \neq 0$, έπεται ότι η η σύνθετη τροποποιημένη στο μηδέν γεωμετρική τ.μ. S είναι μικτού τύπου, η οποία έχει μάζα πιθανότητας στο σημείο μηδέν ίση με p_0 και συνεχής στο διάστημα $(0, \infty)$.

Έστω ότι το ενδιαφέρον στρέφεται στην εύρεση μιας σχέσης που να συνδέει την σύνθετη τροποποιημένη στο μηδέν γεωμετρική κατανομή με αυτή της σύνθετης γεωμετρικής κατανομής. Ως εκ τούτου, έστω ότι η τ.μ. $S \sim CZMG(p_0, p; f)$ και η τ.μ. $S \sim CG(p; f)$, δηλαδή είναι

$$S_0 = \begin{cases} 0, & N_0 = 0, \\ X_1 + X_2 + ... + X_{N_0}, & N_0 \geq 1, \end{cases}$$

όπου η τ.μ. $N_0 \sim G_0(p)$.
Έστω ότι $\bar{H}_0(x)$ είναι η συνάρτηση δεξιάς ουράς της S_0 και $h_0(x)$ είναι η συνάρτηση πυκνότητας πιθανότητάς της. Τότε, με βάση το Θεώρημα 5.5, είναι

$$\hat{\bar{H}}_0(s) = \frac{q[1 - \hat{f}(s)]}{s[1 - q\hat{f}(s)]}. \tag{5.16}$$

Επιπλέον, ισχύει ότι:

$$\hat{\bar{G}}(s) = (1 - p_0)\frac{1 - \hat{f}(s)}{s[1 - q\hat{f}(s)]}, \tag{5.17}$$

οπότε από την Σχέση (5.17) προκύπτει ότι

$$\hat{\bar{G}}(s) = \frac{1 - p - 0}{q}\hat{\bar{H}}_0(s).$$

Επειδή η τελευταία σχέση γράφεται ισοδύναμα ως

$$\int_0^\infty e^{-sx}\bar{G}(x)dx = \frac{1-p_0}{q}\int_0^\infty e^{-sx}\bar{H}_0(x)dx$$

έπεται ότι

$$\bar{G}(x) = \frac{1-p_0}{q}\bar{H}_0(x), \quad x \geq 0. \tag{5.18}$$

ΠΑΡΑΔΕΙΓΜΑ 5.21 ***(Σύνθετη τροποποιημένη στο μηδέν γεωμετρική-εκθετική κατανομή)***

Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N,\ N \geq 1$ (με $S = 0$ αν $N = 0$), όπου η τ.μ. $N \sim ZMG(p_0, p)$ και η τ.μ. $X \sim Exp(\lambda)$.

Επειδή

$$\hat{f}(s) = \frac{\lambda}{\lambda + s},$$

από την Σχέση (5.17) έπεται ότι

$$\begin{aligned} & \hat{\bar{G}}(s) = (1-p_0)\frac{1-\frac{\lambda}{\lambda+s}}{1-q\frac{\lambda}{\lambda+s}} = (1-p_0)\frac{s}{s(\lambda+s-q\lambda)} \\ \Rightarrow \quad & \hat{\bar{G}}(s) = (1-p_0)\frac{1}{p\lambda+s}. \end{aligned}$$

Η τελευταία σχέση ισοδύναμα γράφεται

$$\begin{aligned} & \int_0^\infty e^{-sx}\bar{G}(x)dx = (1-p_0)\int_0^\infty e^{-sx}e^{-p\lambda x}dx \\ \Rightarrow \quad & \bar{G}(x) = (1-p_0)e^{-p\lambda x}, \quad x \geq 0. \end{aligned}$$

και

$$g(x) = \begin{cases} p_0, & x = 0, \\ (1-p_0)p\lambda e^{-p\lambda x}, & x > 0. \end{cases}$$

Έστω ότι η τ.μ. $N \sim ZMG(0.1, 0.5)$ και η τ.μ. $X \sim Exp(5)$, τότε

```r
library(actuar)
fx <- discretize(pexp(x,5),from=0,to=200,step=1,method="lower")
Gs <- aggregateDist("recursive",model.freq="zero-modified geometric",
                    model.sev=fx,prob=0.5,p0=0.1)
mode(Gs); x <- 0:18; g <-diff(Gs)[1:19];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

       x   gs(x)   Gs(x)
[1,]   0 0.10000 0.10000
[2,]   1 0.44697 0.54697
[3,]   2 0.22499 0.77196
[4,]   3 0.11325 0.88521
[5,]   4 0.05701 0.94222
[6,]   5 0.02870 0.97091
[7,]   6 0.01444 0.98536
[8,]   7 0.00727 0.99263
[9,]   8 0.00366 0.99629
[10,]  9 0.00184 0.99813
[11,] 10 0.00093 0.99906
[12,] 11 0.00047 0.99953
[13,] 12 0.00023 0.99976
[14,] 13 0.00012 0.99988
[15,] 14 0.00006 0.99994
[16,] 15 0.00003 0.99997
[17,] 16 0.00002 0.99998
[18,] 17 0.00001 0.99999
[19,] 18 0.00000 1.00000
```

□

2. Η μη-αρνητική τ.μ. X είναι διακριτή

Έστω ότι η τυχαία μεταβλητή $X \in \{0, 1, 2, ...\}$. Τότε, σύμφωνα με την σχέση

$$P_N(u) = p_0 + (1 - p_0)\frac{pu}{1 - qu}, \quad |u| < \frac{1}{q} \tag{5.19}$$

έπεται ότι η πιθανογεννήτρια συνάρτηση της διακριτής σύνθετης τροποποιημένης στο μηδέν γεωμετρικής κατανομής, είναι

$$P_S(u) = P_N(P_X(u)) = p_0 + (1-p_0)\frac{pP_X(u)}{1-qP_X(u)} \tag{5.20}$$

οπότε από την Σχέση (5.11) έπεται ότι η ροπογεννήτρια συνάρτηση $R_S(u) = \sum_{x=0}^{\infty} \bar{G}(x)u^x$ των αριθμών $\{\bar{G}(x) : x \geq 0\}$, είναι

$$\begin{aligned} R_S(u) &= \frac{1-P_S(u)}{1-u} \\ &= \frac{1-p_0-(1-p_0)\frac{pP_X(u)}{1-qP_X(u)}}{1-u} \\ &= (1-p_0)\frac{1-\frac{pP_X(u)}{1-qP_X(u)}}{1-u} \\ &= (1-p-0)\frac{1-pP_X(u)-pP_X(u)}{(1-u)[1-qP_X(u)]} \end{aligned}$$

ή

$$R_S(u) = (1-p_0)\frac{1-P_X(u)}{(1-u)[1-qP_X(u)]} \tag{5.21}$$

ΠΑΡΑΔΕΙΓΜΑ 5.22 ***(Σύνθετη τροποποιημένη στο μηδέν γεωμετρική-μείξη γεωμετρικών κατανομών)*** *[9]*

Έστω ένα χαρτοφυλάκιο κινδύνων για το οποίο η τ.μ. $N \sim ZMG(p_0, p)$, όπου $p_0 = 1/3$ και $p = 1/2$ και το μέγεθος ατομικής ζημιάς X έχει συνάρτηση πυκνότητας πιθανότητας $f(x) = P(X = x)$, όπου

$$f(x) = \frac{3^{x-1}}{2^{2x+1}} + \frac{1}{3}\cdot\frac{1}{2^{x+1}} = \frac{2}{3}\cdot\frac{1}{4}\left(\frac{3}{4}\right)^x + \frac{1}{3}\cdot\frac{1}{2}\left(\frac{1}{2}\right)^x, \quad x = 0, 1, 2, ...$$

Δηλαδή η τ.μ. X είναι η διακριτή μείξη των γεωμετρικών κατανομών $G_0(1/4)$ και $G_0(1/2)$ με βάρη μείξης $2/3$ και $1/3$ αντίστοιχα.

Επομένως, η πιθανογεννήτρια συνάρτησή της, είναι

$$\begin{aligned} P_X(u) &= \frac{2}{3}\cdot\frac{\frac{1}{4}}{1-\frac{3}{4}u}+\frac{1}{3}\cdot\frac{\frac{1}{2}}{1-\frac{1}{2}u} \\ &= \frac{2}{3}\cdot\frac{1}{4-3u}+\frac{1}{3}\cdot\frac{1}{2-u} \\ &= \frac{8-5u}{3(4-3u)(2-u)}, \quad |u|<\frac{1}{2}. \end{aligned}$$

Ύστερα από πράξεις προκύπτει ότι η συνάρτηση πιθανότητας της τ.μ. S είναι

$$g(x) = \begin{cases} \frac{7}{15}, & x=0 \\ \frac{\sqrt{145}+9}{3\sqrt{145}}\cdot\frac{19-\sqrt{145}}{55-\sqrt{145}}\left(\frac{36}{55-\sqrt{145}}\right)^x+\frac{\sqrt{145}-9}{3\sqrt{145}}\cdot\frac{19+\sqrt{145}}{55+\sqrt{145}}\left(\frac{36}{55+\sqrt{145}}\right)^x, & x\geq 1 \end{cases}$$

Συνεπώς,

```
library(actuar)
fx<-rmixture(100,2/3, expression(rgeom(1/4), rgeom(1/2)))
Gs <- aggregateDist("recursive",model.freq ="zero-modified␣geometric",
model.sev=fx,prob=1/2,p0=1/3)
mode(Gs); x <- 0:3; g <-diff(Gs)[1:4];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x   gs(x)    Gs(x)
[1,] 0   0.41935  0.41935
[2,] 1   0.06466  0.48400
[3,] 2   0.50168  0.98568
[4,] 3   0.38377  1.00000
```

□

5.6.2 Σύνθετες περικομμένες στο μηδέν γεωμετρικές κατανομές

Αν η τυχαία μεταβλητή $N \sim ZMG(p_0, p)$, με $p_0 = 0$, τότε η τ.μ. N έχει την περικομμένη στο σημείο μηδέν (zero-truncated) γεωμετρική κατανομή, δηλαδή είναι $N \sim$

$G_1(p)$. Τότε, η πιθανογεννήτρια συνάρτηση της τ.μ. Ν, δίνεται από την σχέση

$$P_N(u) = \frac{pu}{1-qu}, \quad |u| < \frac{1}{q} \tag{5.22}$$

και επειδή η $N \in \{1, 2, 3, ...\}$, ισχύει ότι στο μοντέλο συλλογικού κινδύνου

$$S = X_1 + X_2 + ...X_N, \quad N \geq 1$$

και η τ.μ. $N \sim G_1(p)$.

Τότε, η τ.μ. S έχει την σύνθετη περικομμένη στο σημείο μηδέν γεωμετρική κατανομή (συμβολικά γράφεται $S \sim CG_1(p)$ ή $S \sim CZTG(p;f)$), όπου CZTG είναι τα αρχικά των λέξεων Compound Zero Truncated Geometric.

1. Η μη-αρνητική τ.μ. Χ είναι συνεχής

Έστω ότι η μη-αρνητική τ.μ. Χ είναι συνεχής, και επειδή η τ.μ. $N \geq 1$, έπεται ότι η τυχαία μεταβλητή S είναι θετική και συνεχής στο διάστημα $(0, \infty)$, χωρίς όμως να είναι μικτού τύπου (δεν έχει μάζα πιθανότητας στο σημείο μηδέν).

Σύμφωνα με την Σχέση (5.22), η ροπογεννήτρια συνάρτηση της τ.μ. S δίνεται από τον τύπο

$$M_S(t) = P_N(M_x(t)) = \frac{pM_X(t)}{1-qM_X(t)}. \tag{5.23}$$

Σύμφωνα με την Σχέση (5.2) και την Σχέση (5.23) προκύπτει άμεσα ότι οι μεταβλητές Υ και S είναι ισόνομες.

Συνεπώς,

"Η σύνθετη γεωμετρική κατανομή (p) είναι η διακριτή μείξη της εκφυλισμένης στο σημείο μηδέν κατανομής και της σύνθετης περικομμένης στο μηδέν γεωμετρικής (p) κατανομής" [9].

Ο μετασχηματισμός Laplace της δεξιάς ουράς της τυχαίας μεταβλητής S δίνεται από την σχέση

$$\hat{\bar{G}}(s) = \frac{1-\hat{f}(s)}{s[1-q\hat{f}(s)]}. \tag{5.24}$$

ΠΑΡΑΔΕΙΓΜΑ 5.23 ***(Σύνθετη περικομμένη στο μηδέν γεωμετρική-εκθετική κατανομή)***

Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N$, $N \geq 1$, *όπου η τ.μ.* $N \sim ZTG(p)$ *και η τ.μ.* $X \sim Exp(\lambda)$.

Επειδή ο μετασχηματισμός Laplace της τ.μ. X είναι

$$\hat{f}(s) = \frac{\lambda}{\lambda+s}, \quad s > 0,$$

από την Σχέση (5.24) έπεται ότι

$$\hat{\bar{G}}(s) = \frac{1-\frac{\lambda}{\lambda+s}}{s\left(1-q\frac{\lambda}{\lambda+s}\right)} = \frac{s}{s(\lambda+s-q\lambda)} = \frac{1}{p\lambda+s},$$

ή ισοδύναμα

$$\begin{aligned} &\int_0^\infty e^{-sx}\bar{G}(x)dx = \int_0^\infty e^{-sx}e^{-p\lambda x}dx \\ \Rightarrow \quad &\bar{G}(x) = e^{-p\lambda x} \end{aligned}$$

και η συνάρτηση πυκνότητας πιθανότητας είναι

$$g(x) = -\bar{G}'(x) = p\lambda e^{-p\lambda x}, \quad x \geq 0.$$

Για $p = 0.5$ *και* $\lambda = 3$ *είναι,*

```
library(actuar)

fx <- discretize(pexp(x,3),from = 0,to = 200, step=1,method ="lower")

p<-1/2
q<-1-p
p0<-0
```

```
p1<-p
a<-q
b<-0
pmfZTgeom=function(n){
  if(n==0)return(p0)
  if(n==1)return(p1)
  return((a+b/n)*pmfZTgeom(n-1))
}
ZTGeom=sapply(n,pmfZTgeom)

Gs <- aggregateDist("convolution",model.freq =ZTGeom,model.sev=fx)
mode(Gs); x <- 0:7; g <-diff(Gs)[1:8];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x   gs(x)    Gs(x)
[1,] 0  0.00000  0.00000
[2,] 1  0.47511  0.47511
[3,] 2  0.24938  0.72449
[4,] 3  0.13090  0.85538
[5,] 4  0.01776  0.87314
[6,] 5  0.00171  0.87485
[7,] 6  0.00014  0.87499
[8,] 7  0.00001  0.87500
```

□

ΠΑΡΑΔΕΙΓΜΑ 5.24 ***(Σύνθετη περικομμένη στο μηδέν γεωμετρική-μείξη εκθετικών)*** *[9]*

Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N,\ N \geq 1$, *όπου η τ.μ.* $N \sim ZTG(p)$ *με* $E(N) = 3$ *και η τ.μ. X έχει συνάρτηση πυκνότητας πιθανότητας*

$$f(x) = e^{-2x} + \frac{3}{2}e^{-3x} = \frac{1}{2}2e^{-2x} + \frac{1}{2}3e^{-3x}, \quad x > 0,$$

δηλαδή η τ.μ. X είναι η διακριτή μείξη της $Exp(2)$ και της $Exp(3)$ κατανομής με βάρη μείξης $1/2$ και $1/2$ αντίστοιχα. Τότε, ο μετασχηματισμός Laplace είναι:

$$\hat{f}(s) = \frac{1}{2} \cdot \frac{2}{2+s} + \frac{1}{2} \cdot \frac{3}{3+s} = \frac{5s+12}{2(s+2)(s+3)}, \quad s > 0.$$

Επειδή, $E(N) = 1/p = 3 \;\Rightarrow\; p = 1/3.$

Ύστερα από πράξεις έπεται ότι

$$\bar{G}(x) = \left(\frac{15}{4\sqrt{7}} + \frac{1}{2}\right) e^{-\left(\frac{5-\sqrt{7}}{3}\right)^{x}} - \left(\frac{15}{4\sqrt{7}} - \frac{1}{2}\right) e^{-\left(\frac{5+\sqrt{7}}{3}\right)^{x}}, \quad x \geq 0$$

και

$$g(x) = -\bar{G}'(x) = \frac{73-5\sqrt{7}}{12\sqrt{7}} e^{-\left(\frac{5-\sqrt{7}}{3}\right)^{x}} - \frac{73+5\sqrt{7}}{12\sqrt{7}} e^{-\left(\frac{5+\sqrt{7}}{3}\right)^{x}}, \quad x > 0.$$

```
library(actuar)
fx<-rmixture(100,1/2, expression(rexp(2), rexp(3)))

p<-1/3
q<-1-p
p0<-0
p1<-p
a<-q
b<-0
pmfZTgeom=function(n){
  if(n==0)return(p0)
  if(n==1)return(p1)
  return((a+b/n)*pmfZTgeom(n-1))
}
ZTGeom=sapply(n,pmfZTgeom)

Gs <- aggregateDist("convolution",model.freq =ZTGeom,model.sev=fx)
mode(Gs); x <- 0:4; g <-diff(Gs)[1:5];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

      x    gs(x)    Gs(x)
[1,]  0  0.42018  0.42018
[2,]  1  0.08915  0.50934
```

```
[3,] 2   0.26519   0.77452
[4,] 3   0.14894   0.92347
[5,] 4   0.12828   1.00000
```

2. Η μη-αρνητική τ.μ. X είναι διακριτή

Έστω τώρα ότι η τ.μ. $X \in \{0, 1, 2, ...\}$ και η τ.μ. $N \neq 0$,τότε:

- Αν $f(0) \neq 0 \quad \Rightarrow \quad$ η $S \in \{0, 1, 2, ...\}$ με $g(0) = P(S = 0) = P(f(0))$.

  Οπότε, από την Σχέση (5.22), προκύπτει ότι:

$$g(0) = \frac{pf(0)}{1 - qf(0)}$$

- Αν $f(0) = 0 \quad \Rightarrow \quad S \in \{1, 2, 3, ...\}$ με $g(0) = 0$.

Από την Σχέση (5.22) έπεται ότι, η ροπογεννήτρια συνάρτηση της εν λόγω κατανομής είναι:

$$P_S(u) = P_N(P_X(u)) = \frac{pP_X(u)}{1 - qP_X(u)}. \tag{5.25}$$

Οπότε, σύμφωνα με την Σχέση (5.11), η γεννήτρια συνάρτηση $R_S(u) = \sum_{x=0}^{\infty} \bar{G}(x)u^x$, είναι

$$R_S(u) = \frac{1 - P_X(u)}{(1 - u)[1 - qP_X(u)]}. \tag{5.26}$$

ΠΑΡΑΔΕΙΓΜΑ 5.25 ***(Σύνθετη περικομμένη στο μηδέν γεωμετρική-γεωμετρική κατανομή)***

Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N$, $N \geq 1$, όπου η τ.μ. $N \sim ZTG(p)$ και η τ.μ. $X \sim G(\theta)$.

Η πιθανογεννήτρια συνάρτηση της τ.μ. X, είναι

$$P_X(u) = \frac{\theta}{1 - (1 - \theta)u}, \quad |u| < \frac{1}{1 - \theta}.$$

Από την Σχέση (5.26), έπεται ότι

$$R_S(u) = \frac{1 - \frac{\theta}{1-(1-\theta)u}}{(1-u)\left(1 - \frac{q\theta}{1-(1-\theta)u}\right)} = ... = \frac{1-\theta}{1-q\theta-(1-\theta)u}$$
$$\Rightarrow \quad R_S(u) = \frac{\frac{1-\theta}{1-q\theta}}{1 - \frac{1-\theta}{1-q\theta}u}.$$

Επειδή $0 < p < 1$, *είναι*

$$0 < \frac{1-\theta}{1-q\theta} < 1.$$

Συνεπώς, η προηγούμενη σχέση ισοδύναμα γίνεται

$$\sum_{x=0}^{\infty} \bar{G}(x)u^x = \frac{1-\theta}{1-q\theta}\sum_{x=0}^{\infty}\left(\frac{1-\theta}{1-q\theta}\right)^x u^x = \sum_{x=0}^{\infty}\left(\frac{1-\theta}{1-q\theta}\right)^{x+1} u^x$$
$$\Rightarrow \quad \bar{G}(x) = \left(\frac{1-\theta}{1-q\theta}\right)^{x+1}, \quad x = 0, 1, 2, ...$$

Επειδή $f(0) = \theta \neq 0$, *είναι*

$$g(x) = \begin{cases} P_N(f(0)), & x = 0, \\ \bar{G}(x-1) - \bar{G}(x), & x = 1, 2, 3, ..., \end{cases}$$

όπου,

$$P_N(f(0)) = P_N(\theta) = \frac{q\theta}{1-q\theta},$$

Συνεπώς,

$$g(x) = \begin{cases} \dfrac{q\theta}{1-q\theta}, & x = 0, \\ \dfrac{p\theta}{1-q\theta}\left(\dfrac{1-\theta}{1-q\theta}\right)^x, & x = 1, 2, 3, ... \end{cases}$$

από την οποία έπεται τελικά ότι

$$g(x) = \frac{p\theta}{1-q\theta}\left(\frac{1-\theta}{1-q\theta}\right)^x, \quad x = 0, 1, 2, ...$$

Έστω ότι $p = 1/2$ *και* $\theta = 1/2$, *τότε είναι:*

```
library(actuar)
fx <- pgeom(x,1/3,lower.tail = TRUE,log.p = FALSE)
```

```
p<-1/2
q<-1-p
p0<-0
p1<-p
a<-q
b<-0
pmfZTgeom=function(n){
  if(n==0)return(p0)
  if(n==1)return(p1)
  return((a+b/n)*pmfZTgeom(n-1)
}
ZTGeom=sapply(n,pmfZTgeom)

Gs <- aggregateDist("convolution",model.freq =ZTGeom,model.sev=fx)
mode(Gs); x <- 0:2; g <-diff(Gs)[1:3];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

     x    gs(x)    Gs(x)
[1,] 0  0.19907  0.19907
[2,] 1  0.39352  0.59259
[3,] 2  0.61420  1.00000
```

□

Από το παραπάνω παράδειγμα έπεται ότι,

"η διακριτή σύνθετη περικομμένη στο μηδέν γεωμετρική(p) - γεωμετρική(θ) κατανομή είναι η γεωμετρική $\left(\frac{p\theta}{1-q\theta}\right)$ κατανομή, δηλαδή η τ.μ. $S \sim G_0\left(\frac{p\theta}{1-q\theta}\right)$" [9].

5.7 Σύνθετες αρνητικές διωνυμικές κατανομές

Μια ακόμα χρήσιμη κατανομή για την μοντελοποίηση της συχνότητας των ζημιών σε ένα χαρτοφυλάκιο κινδύνων είναι η αρνητική διωνυμική κατανομή, η οποία αποτελεί μια γενίκευση της γεωμετρικής κατανομής.

Η συνάρτηση πιθανότητας της κατανομής $N \sim NB(r,p)$ ορίζεται ως:

$$p_n = \binom{r+n-1}{n} p^r q^n, \quad n = 0, 1, 2, ..., \quad q = 1-p.$$

Έστω το μοντέλο συλλογικού κινδύνου

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

όπου η $N \sim NB(r,p)$.

Τότε, η τυχαία μεταβλητή S έχει την σύνθετη αρνητική διωνυμική κατανομή ($S \sim CNB(r,p;f)$), όπου τα CNB είναι τα αρχικά των λέξεων Compound Negative Binomial.

5.7.1 Η τυχαία μεταβλητή X είναι συνεχής

Επειδή $P(N=0) = p^r \neq 0 \quad \Rightarrow \quad P(S=0) = P(N=0) = p^r$, οπότε η τυχαία μεταβλητή S είναι μικτού τύπου με μάζα πιθανότητας στο σημείο μηδέν ίση με p^r και συνεχής στο διάστημα $(0, \infty)$.

ΠΑΡΑΔΕΙΓΜΑ 5.26 ***(Σύνθετη αρνητική διωνυμική-εκθετική κατανομή)***
Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N$, $N \geq 1$ (με $S = 0$ αν $N = 0$), όπου η τ.μ. $N \sim NB(2,p)$, $0 < p < 1$ και η τ.μ. $X \sim Exp(\lambda)$, $\lambda > 0$.
Η πιθανογεννήτρια συνάρτηση της τ.μ. N είναι

$$P_N(u) = \left(\frac{p}{1-qu}\right)^2, \quad |u| < \frac{1}{q},$$

και η ροπογεννήτρια της τ.μ. X αντίστοιχα είναι

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

Συνεπώς, η ροπογεννήτρια συνάρτηση της τ.μ. S είναι

$$\begin{aligned} M_S(t) &= P_N(M_X(t)) \\ &= \left(\frac{p}{1 - qM_X(t)}\right)^2 \\ &= \left(p + \frac{p}{1 - qM_X(t)} - p\right)^2 \\ &= ... = p^2 + 2pq\frac{p\lambda}{p\lambda - t} + q^2\left(\frac{p\lambda}{p\lambda - t}\right)^2. \end{aligned}$$

Η παραπάνω σχέση ισοδύναμα γράφεται

$$M_S(t) = p^2 M_{Y_1}(t) + 2pqM_{Y_2}(t) + qM_{Y_3}(t),$$

όπου:

- *η τ.μ. Y_1 έχει την εκφυλισμένη κατανομή στο σημείο μηδέν, με $M_{Y_1}(t) = 1$,*
- *η τ.μ. $Y_2 \sim Exp(p\lambda)$, με $M_{Y_2}(t) = \frac{p\lambda}{p\lambda - t}$,*
- *η τ.μ. $Y_3 \sim Erl(2, p\lambda)$, με $M_{Y_3}(t) = \left(\frac{p\lambda}{p\lambda - t}\right)^2$.*

Συνεπώς η σ.κ. της τυχαίας μεταβλητής S δίνεται από την σχέση

$$G(x) = p^2 F_{Y_1}(x) + 2pqF_{Y_2}(x) + q^2 F_{Y_3}(x), \quad x \geq 0.$$

Ύστερα από πράξεις έπεται ότι

$$\begin{aligned} G(x) &= 1 - (2pq + q^2)e^{-p\lambda x} - p\lambda q^2 x e^{-p\lambda x} \\ &= 1 - (1 - p^2 + p\lambda q^2 x)e^{-p\lambda x}, \quad x \geq 0. \end{aligned}$$

Επειδή $P(N=0)=p^2$, *η συνάρτηση πυκνότητας πιθανότητας της S είναι*

$$g(x)=\begin{cases} p^2, & x=0,\\ p^2 q\lambda(2+q\lambda x)e^{-p\lambda x}, & x>0. \end{cases}$$

Έστω ότι, η τ.μ. $N\sim NB(2,1/3)$ *και η τ.μ.* $X\sim Exp(4)$, *τότε είναι:*

```
library(actuar)
fx <- discretize(pexp(x,4),from = 0,to = 200, step=1,method ="lower")

Gs <- aggregateDist("recursive",model.freq ="negative␣binomial",
                    model.sev=fx,size=2,prob=1/3)
mode(Gs); x <- 0:24; g <- diff(Gs)[1:25];
mat <- cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=3)

      x   gs(x)  Gs(x)
[1,]  0   0.111  0.111
[2,]  1   0.145  0.257
[3,]  2   0.145  0.402
[4,]  3   0.130  0.532
[5,]  4   0.109  0.641
[6,]  5   0.088  0.729
[7,]  6   0.069  0.797
[8,]  7   0.053  0.850
[9,]  8   0.040  0.890
[10,] 9   0.030  0.920
[11,] 10  0.022  0.942
[12,] 11  0.016  0.958
[13,] 12  0.012  0.970
[14,] 13  0.009  0.979
[15,] 14  0.006  0.985
[16,] 15  0.004  0.989
[17,] 16  0.003  0.992
[18,] 17  0.002  0.995
[19,] 18  0.002  0.996
[20,] 19  0.001  0.997
[21,] 20  0.001  0.998
```

```
[22,] 21   0.001   0.999
[23,] 22   0.000   0.999
[24,] 23   0.000   0.999
[25,] 24   0.000   1.000
```

□

5.7.2 Η τυχαία μεταβλητή X είναι διακριτή

Έστω τώρα ότι η τυχαία μεταβλητή $X \in \{0, 1, 2, ...\}$. Τότε, κάτω από την υπόθεση ότι η κατανομή της τυχαίας μεταβλητής X ανήκει στην διακριτή οικογένεια κατανομών, μπορούν να βρεθούν ανάλογα αναλυτικά αποτελέσματα για την κατανομή της διακριτής σύνθετης αρνητικής διωνυμικής τυχαίας μεταβλητής S.

Ισχύει ότι:

$$g(0) = P_N(f(0)) = \begin{cases} p^r, & \text{αν } f(0) = 0, \\ \left(\dfrac{p}{1 - qf(0)}\right)^r, & \text{αν } f(0) \neq 0. \end{cases}$$

ΠΑΡΑΔΕΙΓΜΑ 5.27 ***(Σύνθετη αρνητική διωνυμική-γεωμετρική κατανομή)***
Έστω το μοντέλο συλλογικού κινδύνου $S = X_1 + X_2 + ... + X_N$, $N \geq 1$ (με $S = 0$ αν $N = 0$), όπου η τ.μ. $N \sim NB(2, p)$, $0 < p < 1$ και η τ.μ. $X \sim G_0(\theta)$, $0 < \theta < 1$.
Η πιθανογεννήτρια συνάρτηση της τ.μ. N είναι

$$P_N(u) = \left(\frac{p}{1 - qu}\right)^2, \quad |u| < \frac{1}{q}$$

και της τ.μ. X αντίστοιχα είναι

$$P_X(u) = \frac{\theta}{1 - (1 - \theta)u}, \quad |u| < \frac{1}{1 - \theta}.$$

Συνεπώς, η πιθανογεννήτρια συνάρτηση της τ.μ. S είναι

$$P_S(u) = P_N(P_X(u)) = \left(\frac{p}{1 - qP_X(u)}\right)^2.$$

Τότε, σύμφωνα με την Σχέση (5.11), ισχύει ότι

$$R_S(u) = \frac{1-P_S(u)}{1-u} = ... = \frac{\frac{(1-q\theta)^2+p^2}{(1-\theta)^2} - (1-p^2)u}{\left(\frac{1-q\theta}{1\theta} - u\right)^2}.$$

Έστω ότι,

$$\alpha_1 = \frac{(1-q\theta)^2 - p^2}{(1-\theta)^2}, \quad \alpha_2 = 1-p^2, \quad k = \frac{1-q\theta}{1-\theta} > 1.$$

Επομένως, η παραπάνω συνάρτηση γράφεται

$$R_S(u) = \frac{\alpha_1 - \alpha_2 u}{(k-u)^2}.$$

Χρησιμοποιώντας τεχνική μερικών κλασμάτων η εν λόγω συνάρτηση γίνεται

$$R_S(u) = A\frac{1}{k-u} + B\frac{1}{(k-u)^2},$$

όπου $A = 1-p^2$ *και* $B = \dfrac{(1-q\theta)^2 - p^2}{(1-\theta)^2} - \dfrac{(1-q\theta)(1-p^2)}{1-\theta}$.
Ισχύει,

$$\begin{aligned} \sum_{x=0}^{\infty} \bar{G}(x)u^x &= A\sum_{x=0}^{\infty} \frac{1}{k^{x+1}}u^x + B\sum_{x=0}^{\infty}(x+1)\frac{1}{k^{x+2}}u^x \\ \Rightarrow \quad \bar{G}(x) &= A\frac{1}{k^{x+1}} + B(x+1)\frac{1}{k^{x+2}}. \end{aligned}$$

Ύστερα από πράξεις έπεται ότι

$$\bar{G}(x) = \frac{(1-p^2)(1-\theta)^{x+1}}{(1-q\theta)^{x+1}} + (x+1)\frac{pq^2\theta(1-\theta)^{x+1}}{(1-q\theta)^{x+2}}.$$

Έστω ότι η τ.μ. $N \sim NB(2, 1/3)$ *και η τ.μ.* $X \sim G_0(1/5)$*, τότε είναι:*

```
library(actuar)
fx <- pgeom(x,1/5,lower.tail = TRUE,log.p = FALSE)
Gs <- aggregateDist("recursive",model.freq ="negative␣binomial",
                    model.sev=fx,size=2,prob=1/3)
mode(Gs); x <- 0:5;
Gbar <- (1-Gs(x))[1:6]
mat<-cbind(x,Gs(x),Gbar); colnames(mat) <- c("x","Gs(x)","Gbar(x)")
round(mat,digits=10)
```

```
     x       Gs(x)       Gbar(x)
[1,] 0   0.1479290   0.852071006
[2,] 1   0.2298589   0.770141102
[3,] 2   0.3749519   0.625048143
[4,] 3   0.6141484   0.385851605
[5,] 4   0.9967697   0.003230291
[6,] 5   1.0000000   0.000000000
```

□

5.8 Το κεντρικό οριακό θεώρημα

• Έστω $S_n = X_1 + X_2 + ... + X_n$, όπου οι τ.μ. $X_i,\ 1 < i < n$ είναι ανεξάρτητες και ισόνομες με

$$E(X_i) = \mu \quad , \quad Var(X_i) = \sigma^2 < \infty.$$

Τότε, $\dfrac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \to N(0,1)$ για μεγάλο n.

Ισχύει,

$$E(S_n) = n\mu \quad , \quad Var(S_n) = n\sigma^2,$$

άρα η τυχαία μεταβλητή $S_n \sim N(E(S_n), Var(S_n))$ (ασυμπτωτικά).

• Έστω τώρα ότι

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

▷ Αν η τ.μ. S είναι συνεχής, τότε:

$$\begin{aligned} P\left[\alpha < S < \beta\right] &= P\left[\frac{\alpha - E(S)}{\sqrt{Var(S)}} < \frac{S - E(S)}{\sqrt{Var(S)}} < \frac{\beta - E(S)}{\sqrt{Var(S)}}\right] \\ &= P\left[\frac{\alpha - E(S)}{\sqrt{Var(S)}} < Z < \frac{\beta - E(S)}{\sqrt{Var(S)}}\right] \\ &= \Phi\left[\frac{\beta - E(S)}{\sqrt{Var(S)}}\right] - \Phi\left[\frac{\alpha - E(S)}{\sqrt{Var(S)}}\right], \end{aligned}$$

όπου $Z = \dfrac{S - E(S)}{\sqrt{Var(S)}} \to N(0,1)$ και $\Phi(z) = P(Z = z)$.

▷ Αν η τ.μ. S παίρνει τιμές α,b με $\alpha < b$ και καμία τιμή στο (α,b), τότε:

$$\bullet\{S > \alpha\} = \{S \geq b\} = \{S > c\} = \{S > \frac{\alpha + b}{2}\}, \quad \forall c \in (\alpha, b),$$

$$\bullet\{S \leq \alpha\} = \{S < b\} = \{S < c\} = \{S < \frac{\alpha + b}{2}\}, \quad \forall c \in (\alpha, b).$$

ΠΑΡΑΔΕΙΓΜΑ 5.28 *[9] Για ένα ομαδικό ασφαλιστήριο συμβόλαιο, ο αριθμός των απαιτήσεων ακολουθεί την διωνυμική κατανομή με μέση τιμή 100 και διακύμανση 20. Το ύψος X κάθε απαίτησης έχει συνάρτηση πιθανότητας $f(x) = P(X = x)$, με*

x	1	2	3	4
f(x)	0.5	0.35	0.1	0.05

Χρησιμοποιώντας κανονική προσέγγιση θα υπολογισθεί η πιθανότητα οι συνολικές απαιτήσεις να είναι μεγαλύτερες από 180 νομισματικές μονάδες.

Αν $N =$ ο αριθμός των απαιτήσεων, τότε

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

είναι οι συνολικές απαιτήσεις του χαρτοφυλακίου.

Επειδή $N \sim B(m,p)$ με $E(N) = mp = 100$ και $Var(N) = mp(1-p) = 20$, χρησιμοποιώντας κανονική προσέγγιση για την S ισχύει ότι:

$$S \sim N(E(S), Var(S)).$$

Επομένως,

$$\begin{aligned}
&E(X) = \sum_{x=1}^{4} xf(x) = 1(0.5) + 2(0.35) + 3(0.1) + 4(0.05) = 1.7,\\
&E(X^2) = \sum_{x=1}^{4} x^2 f(x) = 1^2(0.5) + 2^2(0.35) + 3^2(0.1) + 4^2(0.05) = 3.6,\\
&Var(X) = E(X^2) - E^2(X) = 3.6 - 1.7^2 = 0.71,\\
&E(S) = E(N)E(X) = 100(1.7) = 170,\\
&Var(S) = E(N)Var(X) + E^2(X)Var(N) = 100(0.71) + 1.7^2(20) = 128.80.
\end{aligned}$$

Άρα η $S \sim N(170, 128.80)$.

Έστω ότι ζητείται η $P[S > 180]$:

- Χωρίς διόρθωση συνέχειας:

$$\begin{aligned}
P[S > 180] &= P\left[\frac{S - E(S)}{\sqrt{Var(S)}} > \frac{180 - E(S)}{\sqrt{Var(S)}}\right]\\
&= P\left[Z > \frac{180 - 170}{\sqrt{128.8}}\right] = P[Z > 0.88]\\
&= 1 - \Phi(0.88) = 0.1894.
\end{aligned}$$

- *Με διόρθωση συνέχειας:*

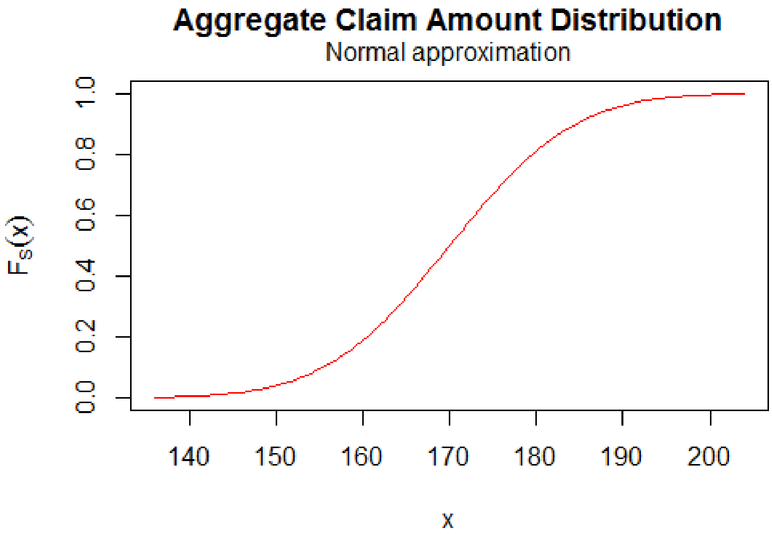
Η τ.μ. S είναι διακριτή με $S \in \{0,1,2,...\}$*, τότε:*

$$\begin{aligned}
P[S>180] &= P\left[S>\frac{180+181}{2}\right] = P\left[S>\frac{180.5}{2}\right] \\
&= P\left[\frac{S-E(S)}{\sqrt{Var(S)}} > \frac{180.5-E(S)}{\sqrt{Var(S)}}\right] \\
&= P\left[Z > \frac{180.5-E(S)}{\sqrt{Var(S)}}\right] \\
&= P\left[Z > \frac{180.5-170}{\sqrt{128.8}}\right] = P[Z>0.93] \\
&= 1-\Phi(0.93) = 0.1762.
\end{aligned}$$

***Παρατήρηση**: Ο δεύτερος τρόπος δίνει καλύτερη προσέγγιση από τον πρώτο και είναι προτιμότερος για διακριτές κατανομές (βλ. Χατζηκωνσταντινίδης, 2015).*

Όμως, σύμφωνα με το πακέτο actuar, προκύπτει το εξής:

```
library(actuar)
meanN<-100 ; varN<-20
x <-c(0,1,2,3,4) ;fx <-c(0,0.5,0.35,0.1,0.05)
meanx <- sum(x*fx) ; meanx2 <-sum((x^2)*fx)
varx <- meanx2-(meanx^2)
means <- meanx*meanN
vars <- meanN*varx+varN*meanx^2

Gs<-aggregateDist("normal", moments = c(means,vars))
Gsbar<- 1-Gs(180)

[1] 0.1891226

plot(Gs,col="red")
```



Συμπέρασμα: *Προφανώς, το εν λόγω πακέτο χρησιμοποιεί την μέθοδο της κανονικής προσέγγισης χωρίς διόρθωση συνέχειας, όπως φαίνεται άλλωστε και από το αποτέλεσμα που προέκυψε. Στην περίπτωση όπου έχουμε διόρθωση συνέχειας, το αποτέλεσμα που προκύπτει οφείλεται στο γεγονός ότι εφόσον η τ.μ. S παίρνει διακριτές τιμές, η επόμενη τιμή μετά το 180 θα είναι το 181. Επομένως αν θέλουμε την* $P(S > 180)$, *τιμή αυτή θα είναι ανάμεσα στο 180 και το 181.*

$\square$

ΘΕΩΡΗΜΑ 5.9 *Έστω ότι η τ.μ.* $S \sim CP(\lambda, f)$ *με* $E(S) = \lambda E(X)$ *και* $Var(S) = \lambda E(X^2)$ *και έστω* $m_k = E(X^k)$, $k \geq 1$. *Τότε, για την τ.μ. Z όταν* $\lambda \to \infty$, *ισχύει ότι*

$$Z = \frac{S - \lambda m_1}{\sqrt{\lambda m_2}} \to N(0,1).$$

Απόδειξη: *Αρκεί να δειχθεί ότι*

$$\lim_{\lambda \to \infty} M_Z(t) = e^{t^2/2}.$$

Είναι

$$Z = \frac{1}{\sqrt{\lambda m_2}} S - \frac{\lambda m_1}{\sqrt{\lambda m_2}}.$$

Έστω ότι $\alpha = \frac{1}{\sqrt{\lambda m_2}}$, $\beta = \frac{\lambda m_1}{\sqrt{\lambda m_2}}$, *τότε από την Σχέση (2.1), προκύπτει ότι*

$$M_Z(t) = M_S\left[\frac{t}{\sqrt{\lambda m_2}}\right]e^{-\frac{\lambda m_1 t}{\sqrt{\lambda m_2}}}$$

όμως,

$$M_S(t) = e^{\lambda[M_X(t)-1]},$$

άρα

$$\begin{aligned} M_Z(t) &= e^{\lambda[M_X(t)[\frac{t}{\sqrt{\lambda m_2}}]-1]}e^{-\frac{\lambda m_1 t}{\sqrt{\lambda m_2}}} \\ &= e^{\lambda\{M_X(t)[\frac{t}{\sqrt{\lambda m_2}}]-1-\frac{m_1 t}{\sqrt{\lambda m_2}}\}} \\ &= ... = M_Z(t) = e^{t^2/2}. \end{aligned} \tag{5.27}$$

Άρα η τ.μ. $Z \sim N(0,1)$.

ΠΑΡΑΔΕΙΓΜΑ 5.29 *[9] Για μια ασφαλιστική κάλυψη, ο αριθμός των απαιτήσεων ακολουθεί την κατανομή Poisson και το μέγεθος ατομικής απαίτησης* $X \sim Pareto(3, 3000)$. *Χρησιμοποιώντας την κανονική προσέγγιση, η πιθανότητα ότι οι συνολικές απαιτήσεις είναι μεγαλύτερες από 4000 νομισματικές μονάδες είναι ίση με 0,2743. Να υπολογισθεί η αναμενόμενη συνολική απαίτηση της ασφαλιστικής κάλυψης.*

Ζητείται η $E(S) = \lambda E(N)$.
Επειδή η τ.μ. $X \sim Pa(3, 3000)$, *ισχύει*

$$\begin{aligned} E(X) &= \frac{3000}{3-1} = 1500 \\ E(X^2) &= \frac{2(3000)^2}{(3-1)(3-2)} = 3000^2 \end{aligned}$$

Συνεπώς,

$$\begin{aligned} & P[S > 4000] = 0.2743 \\ \Leftrightarrow\ & P[\frac{S-E(S)}{\sqrt{Var(S)}} > \frac{4000-E(S)}{\sqrt{Var(S)}}] = 0.2743 \\ \Leftrightarrow\ & P[Z > \frac{4000-1500\lambda}{3000\sqrt{\lambda}}] = 0.2743. \end{aligned}$$

Επειδή, $P[Z > 0.6] = 0.2743$, συνεπάγεται ότι:

$$\frac{4000 - 1500\lambda}{3000\sqrt{\lambda}} = 0.6 \;\Leftrightarrow\; 4000 - 1500\lambda = 1800\sqrt{\lambda}$$
$$\Leftrightarrow\; 1500\lambda + 1800\sqrt{\lambda} - 4000 = 0.$$

Ύστερα από πράξεις προκύπτει ότι $\lambda = 1.299$. Συνεπώς,

$$E(S) = \lambda E(X) = 1.299(1500) = 1948.5.$$

Το εν λόγω πρόβλημα δεν είναι εφικτό να υπολογισθεί με την χρήση του πακέτου actuar καθώς είναι αναγκαία η δήλωση αριθμητικών μεταβλητών για τις παραμέτρους των κατανομών. Στην περίπτωση αυτή ο υπολογισμός της συνολικής απαίτησης θα ήταν αρκετά εύκολος. Έστω τώρα ότι ο στόχος του παραδείγματος ήταν η εύρεση της πιθανότητας οι συνολικές απαιτήσεις να είναι μεγαλύτερες από 4000, η μορφή του θα ήταν η εξής:

```
library(actuar)
meanN<-varN<-1.299
a<-3;b<-3000
fx<-dpareto2(x,a,b,log=FALSE)
meanx<-(b/(a-1)); meanx2<-((a-1)*(b^(a-1)))/((a-1)*(a-2))
varx <-meanx2-(meanx^2)
means <-meanx*meanN
vars <- meanN*varx+varN*meanx^2
means

[1] 1948.5

Gs<-aggregateDist("normal",moments = c(means,vars))
Gsbar<-1-Gs(4000)
Gsbar

[1] 0.2742557
```

Συμπέρασμα: Όπως είναι αντιληπτό και από το ανωτέρω πρόβλημα, το πακέτο actuar καθιστά εφικτό τον υπολογισμό των συνολικών απαιτήσεων καθώς και των *παραμέτρων που χρειάζονται για τον υπολογισμό αυτό μόνο στην περίπτωση που τα δεδομένα είναι αριθμητικά. Με άλλα λόγια δεν εμφανίζει αποτελέσματα συναρτήσει κάποιον τυχαίων μεταβλητών αλλά μόνο αριθμητικά δεδομένα.*

□

ΠΑΡΑΔΕΙΓΜΑ 5.30 *[9] Ο ετήσιος αριθμός ζημιών για κάθε ασφαλισμένο ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$, η οποία μεταβάλλεται για κάθε ασφαλισμένο σύμφωνα με μια κατανομή Gamma με μέση τιμή $1/2$ και διακύμανση $1/2$ αλλά δεν μεταβάλλεται ανά έτος για κάθε ασφαλισμένο. Υπάρχουν 1500 ασφαλισμένοι. Χρησιμοποιώντας κανονική προσέγγιση, θα υπολογισθεί η πιθανότητα να εμφανισθούν περισσότερες από 1600 ζημιές σε δύο έτη.*

Ο αριθμός των ζημιών σε δύο έτη για κάθε ασφαλισμένο, ακολουθεί κατανομή Poisson με $\Lambda = 2\lambda$. Επομένως, είναι

$$E(\Lambda) = 2E(\lambda) = 2(1/2) = 1,$$
$$Var(\Lambda) = 4Var(\lambda) = 4(1/2) = 2.$$

Έστω τώρα ότι,

$N =$ *ο αριθμός των ζημιών για ένα τυχαία επιλεγμένο οδηγό για περίοδο 2 ετών και*

$N_{o\lambda} =$ *ο συνολικός αριθμός ζημιών για 2 έτη για τους 1500 ασφαλισμένους.*

$$E(N_{o\lambda}) = 1500E(N),$$
$$Var(N_{o\lambda}) = 1500Var(N).$$

Σύμφωνα με το Θεώρημα 2.1 ισχύει ότι

$$E(N) = E[E(N|\Lambda)] = E(\Lambda) = 1$$
$$Var(N) = E[Var(N|\Lambda)] + Var[E(N|\Lambda)] = E(\Lambda) + Var(\Lambda) = 1 + 2 = 3$$

Επομένως, είναι

$$
\begin{aligned}
P[N_{o\lambda} > 1600] &= P\left[N_{o\lambda} > \frac{1600-1601}{2}\right] \\
&= P\left[N_{o\lambda} > 1600.5\right] \\
&= P\left[\frac{N_{o\lambda}-E(N_{o\lambda})}{\sqrt{Var(N_{o\lambda})}} > \frac{1600.5-E(N_{o\lambda})}{\sqrt{Var(N_{o\lambda})}}\right] \\
&= P\left[Z > \frac{1600.5-1500}{\sqrt{4500}}\right] \\
&= P[Z > 1.5] = 1 - P[Z \leq 1.5] \\
&= 1 - \Phi(1.5) = 0.0668.
\end{aligned}
$$

```
library(actuar)

years<-2
n<-1500

meanx<-years*(1/2)
varx<-(years^2)*(1/2)

meann<-meanx
varn<-meanx+varx

meanN<-n*meann
varN<-n*varn

Gs <- aggregateDist("normal",moments = c(meanN,varN))
Gsbar <- 1-Gs(1600)
Gsbar

[1] 0.06801856
```

□

ΠΑΡΑΡΤΗΜΑ Α

Βασικές κατανομές

(βλ. Ηλιόπουλος, 2013 [2] και Κούτρας, 2004 [3]).

Α.1 Βασικές διακριτές κατανομές

Διωνυμική (Binomial)

Συμβολισμός: $X \sim B(n,p),\ \ n \in N,\ \ 0 < p < 1.$

Ερμηνεία: Απαριθμεί τον αριθμό των επιτυχιών σε n ανεξάρτητα πειράματα Bernoulli που το κάθε ένα έχει πιθανότητα επιτυχίας p.

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1$$

Μέση τιμή και διασπορά: $E(X) = np, \quad Var(X) = np(1-p).$

Σχόλια:

- Αν $X_1 \sim B(n_1, p), ..., X_m \sim B(n_m, p)$ είναι ανεξάρτητες διωνυμικές τ.μ. (με το ίδιο p), τότε $\sum_{i=1}^{m} X_i \sim B(\sum_{i=1}^{m} n_i, p).$
- Ειδικά για $n = 1$ η διωνυμική κατανομή λέγεται κατανομή Bernoulli.

Αρνητική Διωνυμική (Negative Binomial)

Συμβολισμός: $X \sim NB_0(r,p),\ \ r > 0,\ \ 0 < p < 1.$

Ερμηνεία: Αν $r \in \mathbb{N}$, τότε η X απαριθμεί τον αριθμό των αποτυχιών που προηγούνται της r-οστής επιτυχίας σε μία ακολουθία ανεξάρτητων πειραμάτων Bernoulli που το κάθε ένα έχει πιθανότητα επιτυχίας p.
Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \frac{\Gamma(r+x)}{\Gamma(r)} \frac{p^r(1-p)^x}{x!}, \quad x = 0, 1, 2, ...,$$

όπου $\Gamma(y) = \int_0^\infty t^{y-1}e^{-t}dy, \ y > 0$, είναι η συνάρτηση Γάμμα. Επειδή, αν ο y είναι θετικός ακέραιος ισχύει $\Gamma(y) = (y-1)!$, αν ο r είναι θετικός ακέραιος,

$$f(x) = \binom{r+x-1}{x} p^r(1-p)^x, \quad x = 0, 1, 2, ...$$

Μέση τιμή και διασπορά: $E(X) = r(1-p)/p, Var(X) = r(1-p)/p^2$.

Σχόλια:

- Αν $X_1 \sim NB(r_1, p), ..., X_m \sim NB(r_m, p)$ είναι ανεξάρτητες αρνητικές διωνυμικές τ.μ. (με το ίδιο p), τότε $\sum_{i=1}^{m} X_i \sim NB(\sum_{i=1}^{m} r_i, p)$.

- Ειδικά για $r = 1$ η αρνητική διωνυμική κατανομή λέγεται γεωμετρική κατανομή.

- Σε ορισμένες περιπτώσεις, η αρνητική διωνυμική κατανομή ορίζεται μόνο στο $r \in \mathbb{N} := \{1, 2, ...\}$ με μια διαφορετική πυκνότητα πιθανότητας και με τιμές στο σύνολο $\{r, r+1, ...\}$. Για $r \in \mathbb{N}$ οι δύο ορισμοί διαφέρουν ελάχιστα, γιατί αν Υ είναι μια τυχαία μεταβλητή με εκείνη την κατανομή, τότε ισχύει $Y = X + r$. Επί πλέον, δύο πλεονεκτήματα του εν λόγω ορισμού είναι ότι για κάθε r το σύνολο τιμών της X παραμένει το $\mathbb{Z}_+$ ενώ η κατανομή ορίζεται για κάθε $r \in (0, \infty)$ (και όχι μόνο για ακέραιο). Σε αυτή την περίπτωση ισχύουν τα κάτωθι:

  Συμβολισμός: $X \sim NB_1(r, p), \ r > 0, \ 0 < p < 1$.
  Ερμηνεία: Αν $r \in \mathbb{N}$, τότε η X απαριθμεί τον αριθμό των δοκιμών μέχρι την εμφάνιση της πρώτης επιτυχίας σε μία ακολουθία ανεξάρτητων πειραμάτων Bernoulli που το κάθε ένα έχει πιθανότητα επιτυχίας p.

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, ...$$

Μέση τιμή και διασπορά: $E(X) = r/p, Var(X) = r(1-p)/p^2$.

Γεωμετρική(Geometric)

Συμβολισμός: $X \sim G_0(p), \ 0 < p < 1$.

Ερμηνεία: Απαριθμεί τον αριθμό των αποτυχιών που προηγούνται της πρώτης επιτυχίας σε μια ακολουθία ανεξάρτητων πειραμάτων Bernoulli που το κάθε ένα έχει πιθανότητα επιτυχίας p.

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = p(1-p)^x, \quad x = 0, 1, 2, ...$$

Μέση τιμή και διασπορά: $E(X) = (1-p)/p \ \ Var(X) = (1-p)/p^2$.

Σχόλια:

- Αν $X_1, ..., X_m$ ανεξάρτητες και ισόνομες τ.μ. με την κατανομή $G_0(p)$, τότε $\sum_{i=1}^{m} X_i \sim NB(m, p)$.

Στην περίπτωση τώρα όπου η Γεωμετρική Κατανομή δεν ορίζεται στο μηδέν, αλλά ξεκινάει από την μονάδα ισχύουν τα παρακάτω:

Συμβολισμός: $G_1(p), \ 0 < p < 1$.

Ερμηνεία: Απαριθμεί τον αριθμό των δοκιμών μέχρι την εμφάνιση της πρώτης επιτυχίας σε μια ακολουθία ανεξάρτητων πειραμάτων Bernoulli που το κάθε ένα έχει πιθανότητα επιτυχίας p.

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = p(1-p)^{x-1}, \quad x = 1, 2, ...$$

Μέση τιμή και διασπορά: $E(X) = 1/p, \quad Var(X) = (1-p)/p^2$.

Poisson

Συμβολισμός: $X \sim P(\lambda), \quad \lambda > 0.$

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, ...$$

Μέση τιμή και διασπορά: $E(X) = \lambda, \quad Var(X) = \lambda.$

Σχόλια:

- Αν $X_1 \sim P(\lambda_1), ...X_m \sim P(\lambda_m)$ είναι ανεξάρτητες Poisson τ.μ., τότε $\sum_{i=1}^{m} X_i \sim P(\sum_{i=1}^{m} \lambda_i)$.

A.2 Βασικές συνεχείς κατανομές

Ομοιόμορφη (Uniform)

Συμβολισμός: $X \sim U(\alpha, \beta), \ \alpha < \beta \in \mathbb{R}.$

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \frac{1}{\beta - \alpha}, \quad x \in [\alpha, \beta].$$

Μέση τιμή και διασπορά: $E(X) = (\alpha + \beta)/2, \quad Var(X) = (\alpha - \beta)/12.$

Γάμμα (Gamma)

Συμβολισμός: $X \sim G(\alpha, \beta), \ \alpha, \beta > 0.$

Συνάρτηση πυκνότητας πιθανότητας:

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

όπου $\Gamma(\cdot)$ η συνάρτηση γάμμα που ορίζεται ως $\Gamma(y) := \int_0^\infty t^{y-1}e^{-t}dt, \ y > 0.$

Μέση τιμή και διασπορά: $E(X) = \alpha\beta, \quad Var(X) = \alpha\beta^2.$

Σχόλια:

- Αν $X_1 \sim G(\alpha_1, \beta), ..., X_m \sim G(\alpha_m, \beta)$ είναι ανεξάρτητες τ.μ. γάμμα (με το ίδιο β), τότε $\sum_{i=1}^{m} X_i \sim G(\sum_{i=1}^{m} \alpha_i, \beta)$.
- Ειδικά για $\alpha = 1$ η κατανομή γάμμα λέγεται (αρνητική) εκθετική κατανομή.

Εκθετική (Exponential)

Συμβολισμός: $X \sim Exp(\lambda),\ \ \lambda > 0.$

Συνάρτηση πυκνότητας πιθανότητας:

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

Μέση τιμή και διασπορά: $E(X) = \lambda, Var(X) = \lambda^2$

Σχόλια:

- Αν $X_1, ..., X_m$ ανεξάρτητες και ισόνομες τ.μ. με κατανομή $Exp(\lambda)$, τότε $\sum_{i=1}^{m} X_i \sim G(m, \lambda)$.
- Η κατανομή $Exp(\lambda)$ συμπίπτει με την $G(1, \lambda)$.

Βήτα (Beta)

Συμβολισμός: $X \sim Beta(\alpha, \beta),\ \ \alpha, \beta > 0.$

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in (0, 1).$$

όπου $B(\cdot, \cdot)$ η συνάρτηση βήτα που ορίζεται ως
$B(u, v) := \int_0^1 t^{u-1}(1-t)^{u-1} dt, \ \ u, v > 0.$
Μέση τιμή και διασπορά: $E(X) = \frac{\alpha}{\alpha+\beta}, \ \ Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Σχόλια:

- Αν $X_1 \sim G(\alpha_1, \beta), X_2 \sim G(\alpha_2, \beta)$ είναι ανεξάρτητες τ.μ. γάμμα (με το ίδιο β), τότε $Y = X_1/(X_1 + X_2) \sim Beta(\alpha_1, \alpha_2)$ και είναι ανεξάρτητη από το άθροισμα $X_1 + X_2$.

- Για $\alpha = \beta = 1$ η κατανομή συμπίπτει με την ομοιόμορφη $U(0, 1)$.

Κανονική (Normal)

Συμβολισμός: $X \sim N(\mu, \sigma^2), \ \ \mu \in \mathbb{R}, \ \ \sigma > 0.$

Συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty.$$

Μέση τιμή και διασπορά: $E(X) = \mu, \ \ Var(X) = \sigma^2.$

Σχόλια:

- Αν $X_1 \sim N(\mu_1, \sigma_1^2), ..., X_m \sim N(\mu_m, \sigma_m^2)$ είναι ανεξάρτητες κανονικές τ.μ. και $\alpha_1, ..., \alpha_m, \ \ b_1, ..., b_m \in \mathbb{R}$ σταθερές, τότε

$$\sum_{i=1}^{m} (\alpha_i X_i + b_i) \sim N(\sum_{i=1}^{m} \alpha_i \mu_i + \sum_{i=1}^{m} b_i, \ \ \sum_{i=1}^{m} \alpha_i^2 \sigma_i^2).$$

- Ειδικά για $\mu = 0, \ \ \sigma = 1$, η κατανομή λέγεται τυπική κανονική κατανομή. Σε αυτή την περίπτωση η συνάρτηση κατανομής της συμβολίζεται συνήθως με Φ και η πυκνότητα πιθανότητάς της με φ.
- Αν $X \sim N(\mu, \sigma^2)$ τότε $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$ (τυπική κανονική κατανομή).

ΠΑΡΑΡΤΗΜΑ Β

Πίνακες βασικών κατανομών

(βλ. Χατζηκωνσταντινίδης Ε., 2015). [9]

Πίνακας 5.8.1: Διακριτές Κατανομές

Κατανομή	Συνάρτηση πιθανότητας	Πιθανογεννήτρια $P_x(u)$	Ροπογεννήτρια $M_x(t)$
Εκφυλισμένη στο α	$f_X(\alpha) = 1$	u^{α}	$e^{\alpha t}$
Bernoulli(p)	$p^x q^{1-x}, x = 0, 1$	$q + pu$	$q + pe^t$
B(n,p)	$\binom{n}{x} p^x (1-p)^{n-x}$	$(q + pu)^m$	$(q + pe^t)^m$
Poisson(λ)	$e^{-\lambda} \frac{\lambda^x}{x!}$	$e^{\lambda(u-1)}$	$e^{\lambda(e^t-1)}$
$G_0(p)$	$pq^x, x = 0, 1, ...$	$\frac{p}{1-qu}$	$\frac{p}{1-qe^t}$
$G_1(p)$	$pq^{x-1}, x = 1, 2, ...$	$\frac{pu}{1-qu}$	$\frac{pe^t}{1-qe^t}$
NB(r,p)	$\binom{r+x-1}{x} p^r q^x, x = 0, 1, ...$	$(\frac{p}{1-qu})^r$	$(\frac{p}{1-qe^t})^r$
$NB_1(r,p)$	$\binom{x-1}{r-1} p^r q^{x-r}, x = r, r+1, ...$	$(\frac{pu}{1-qu})^r$	$(\frac{pe^t}{1-qe^t})^r$

Πίνακας 5.8.2: Συνεχείς Κατανομές

Κατανομή	Συνάρτηση πιθανότητας	Μετασχηματισμός Laplace $\hat{f}_x(s)$	Ροπογεννήτρια $M_x(t)$
U(α, β)	$\frac{1}{\beta - \alpha}$	$\frac{e^{-\alpha s} - e^{-\beta s}}{s(\beta - \alpha)}$	$\frac{e^{-\beta t} - e^{-\alpha t}}{t(\beta - \alpha)}, t \neq 0$
$Exp(\lambda)$	$\lambda e^{-\lambda x}$	$\frac{\lambda}{\lambda + s}$	$\frac{\lambda}{\lambda - t}$
$G(\alpha, \lambda)$	$\frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$	$(\frac{\lambda}{\lambda + s})^{\alpha}$	$(\frac{\lambda}{\lambda - t})^{\alpha}$
$Erl(n, \lambda)$	$\frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$	$(\frac{\lambda}{\lambda + s})^n$	$(\frac{\lambda}{\lambda - t})^n$
N(μ,σ^2)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{-\mu s} + \frac{\sigma^2 s^2}{2}$	$e^{-\mu t} + \frac{\sigma^2 t^2}{2}$

ΠΑΡΑΡΤΗΜΑ Γ

Το πακέτο "actuar"

(βλ. Goulet V., et. al, 2018)[12]

---

discretize   Διακριτοποίηση μιας συνεχής κατανομής

Περιγραφή: Υπολογίζει την αθροιστική συνάρτηση μιας συνεχής κατανομής, χρησιμοποιώντας διάφορες μεθόδους.

Παρουσίαση:

```
discretize(cdf, from, to, step = 1, method = c("upper",
           "lower", "rounding", "unbiased"),
lev, by = step, xlim = NULL)
```

cdf	μια έκφραση ορισμένη ως συνάρτηση του x, η οποία μας ορίζει την κατανομή την οποία θέλουμε να διακριτοποιήσουμε
from, to	το εύρος γύρω από το οποίο θα γίνει η διακριτοποίηση
step	ένας αριθμός ο οποίος ορίζει το βήμα που θα γίνει η διακριτοποίηση
method	μεθοδος διακριτοποίησης που θα χρησιμοποιηθεί
lev	υπολογίζει την ελάχιστη αναμενόμενη τιμή που αντιστοιχεί στην κατανομή. Χρησιμοποιείται μόνο στην μέθοδο της αμερόληπτης διακριτοποίησης.
by	ένα ψευδώνυμο για το step
xlim	αριθμητικό μέτρο. Αν ορίζεται, χρησιμεύει ως προεπιλογή για το x(from,to)

aggregateDist	Υπολογισμός της κατανομής των συνολικών απαιτήσεων

Περιγραφή: Υπολογίζει την κατανομή των συνολικών απαιτήσεων σε ένα χαρτοφυλάκιο κινδύνων σε μια ορισμένη χρονική περίοδο χρησιμοποιώντας μια από τις πέντε μεθόδους.

Παρουσίαση:

```
aggregateDist(method = c("recursive", "convolution",
              "normal", "npower", "simulation"),
              model.freq = NULL,model.sev = NULL,
              p0 = NULL, x.scale = 1, convolve = 0,
              moments, nb.simul, ...,tol = 1e-06,
              maxit = 500, echo = FALSE)

## S3 method for class 'aggregateDist'
plot(x, xlim, ylab = expression(F[S](x)),
     main = "Aggregate␣Claim␣Amount␣Distribution",
     sub = comment(x), ...)
```

method	η μέθοδος που θα χρησιμοποιηθεί
method	η μέθοδος που θα χρησιμοποιηθεί
model.freq	για την αναδρομική μέθοδο: μια σειμβολοσειρά χαρακτήρων που δίνει το όνομα στις κλάσεις κατανομών (α,b,0) ή (α,b,1)
	για την μέθοδο των συνελίξεων: είναι ένας φορέας του πλήθους των πιθανών απαιτήσεων
	για την μέθοδο προσομοίωσης: ένα μοντέλο της συχνότητας προσομοίωσης
	για την κανονική και npower μέθοδο: αγνοείται
model.sev	για την αναδρομική και την μέθοδο συνελίξεων: ένας φορέας του πλήθους πιθανών ποσών
	για την μέθοδο προσομοίωσης: η ένταση του μοντέλου προσομοίωσης
	κανονική και npower μέθοδος: αγνοείται
p0	αυθαίρετη πιθανότητα στα μηδέν για την συχνότητα πιθανοτήτων. Δημιουργεί μια τροποποιημένη ή περικομμένη στο μηδέν κατανομή αν δεν είναι άκυρο. Χρησιμοποιείται μόνο στην αναδρομική μέθοδο.
x.scale	αξία ενός ποσού στο αυστηρό μοντέλο. Χρησιμοποιείται μόνο στην αναδρομική και την μέθοδο συνελίξεων.
convolve	αριθμός συνελίξεων που επιθυμούμαι να πραγματοποιηθεί. Χρησιμοποιείται μόνο στην μέθοδο συνελίξεων.
moments	διάνυσμα πραγματικών στιγμών στο συλλογικό μοντέλο απαιτήσεων. Χρησιμοποιείται μόνο στην κανονική και npower μέθοδο.
nb.simul	αριθμός προσομοιώσεων. Ισχύει μόνο για την μέθοδο προσομοίωσης.
...	παράμετροι της κατανομής συχνοτήτων για την αναδρομική μέθοδο
tol	η αθροιστική κατανομή θα μειωθεί έτσι ώστε να απομακρυνθεί από την μονάδα
maxit	μέγιστος αριθμός επαναλήψεων στην αναδρομική μέθοδο.
echo	λογικός τελεστής, επαλαμβάνει τις επαναλήψεις στην αναδρομική μέθοδο.
x, object	ένα αντικείμενο της τάξης *aggregateDist*
xlim	αριθμητικό μήκος. Τα όρια του x για το γράφημα
ylab	τίτλος για τον άξονα των y
main	κυρίως τίτλος γραφήματος
sub	υπότιτλος γραφήματος

rmixture	Προσομοίωση από διακριτές μείξεις κατανομών

Περιγραφή: Παράγει τυχαίες μεταβλητές από μια διακριτή στο πλήθος μείξη κατανομών.

Παρουσίαση:

```
rmixture(n, probs, models)
```

n	Αριθμός τυχαίων μεταβλητών που παράγονται.
probs	Ένας μη-αρνητικός αριθμός που ορίζει την πιθανότητα για κάθε μοντέλο. Αθροίζει στην μονάδα.
models	Ένα διάνυσμα που ορίζει τα μοντέλα προσομοίωσης.

rcompound	Προσομοίωση από σύνθετα μοντέλα κατανομών

Περιγραφή: Η rcompound παράγει τυχαίες μεταβλητές από ένα σύνθετο μοντέλο και η πιο απλή μορφή της η rcomppois χρησιμεύει στην περίπτωση όπου η τ.μ. $N \sim Poisson(\lambda)$.

Παρουσίαση:

```
rcompound(n, model.freq, model.sev, SIMPLIFY = TRUE)

rcomppois(n, lambda, model.sev, SIMPLIFY = TRUE)
```

n	Αριθμός παρατηρήσεων.
model.freq, model.sev	Εκφράσεις που αναφέρονται στις τυχαίες μεταβλητές X και N.
lambda	Η παράμετρος της Poisson.

rcomphierarc	Προσομοίωση από σύνθετα ιεραρχικά μοντέλα

Περιγραφή: Υπολογίζει με προσομοίωση ιεραρχικά μοντέλα.

Παρουσίαση:

```
rcomphierarc(nodes, model.freq = NULL, model.sev = NULL,
             weights = NULL)
```

nodes	Ένα διάνυσμα ή μια λίστα που δίνει τον αριθμό των κόμβων σε κάθε επίπεδο του ιεραρχικού μοντέλου ξεκινώντας από την κορυφή του χαρτοφυλακίου και προχωρώντας προς τα κάτω.
model.freq, model.sev	Εκφράσεις που αναφέρονται στις τυχαίες μεταβλητές X και N.
weights	Ένα διάνυσμα για τα βάρη.

ΠΑΡΑΡΤΗΜΑ Δ

Παρουσίαση εντολών με χρήση της γλώσσας προγραμματισμού R

(βλ. Αντζουλάκος Δ., 2016). [1]

ΠΑΡΑΔΕΙΓΜΑ 3.1, ΣΕΛ. 48

```
pn<-c(0.2,0.2,0.2,0.4)
fx0<-c(1,0,0,0,0,0,0,0,0,0)
fx<-c(0,0.4,0.4,0.2,0,0,0,0,0)
fx2<-convolve(fx,rev(fx),type = "open")
fx3<-convolve(fx,rev(fx2),type = "open")

library(actuar)
Gs<-aggregateDist("convolution",model.freq = pn,model.sev = fx)
knots(Gs); gs<-diff(Gs)[0:9]; x1<-knots(Gs)
mat1<-cbind(x1,fx0,fx,fx2,fx3,gs,Gs(x))
colnames(mat1)<-c("x","fx0","fx","fx2","fx3","gs","Gs")
round(mat1,digits = 4)
```

ΠΑΡΑΔΕΙΓΜΑ 5.7, ΣΕΛ. 104

```
fx0<-c(1,0,0,0,0,0,0,0)
fx1<-c(0,3/5,1/5,1/5,0,0,0,0)
fx2<-convolve(fx1,rev(fx1),type = "open")
fx3<-convolve(fx2,rev(fx1),type = "open")
fx4<-convolve(fx3,rev(fx1),type = "open")
fx5<-convolve(fx4,rev(fx1),type = "open")
fx6<-convolve(fx5,rev(fx1),type = "open")
```

```
x<-c(0:7)
mat<-cbind(x,fx0,fx1,fx2,fx3,fx4,fx5,fx6)
colnames(mat)<-c("x","f*0(x)","f*1(x)","f*2(x)","f*3(x)","f*4(x)",
                 "f*5(x)","f*6(x)")
round(mat,digits = 5)

pn<-dpois(x,0.2,log = FALSE)

library(actuar)
Gs<-aggregateDist("convolution",model.freq = pn,model.sev = fx1)
mode(Gs); x <- 0:10; g <-diff(Gs)[1:11];
mat<-cbind(x,g,Gs(x)); colnames(mat) <- c("x","gs(x)","Gs(x)")
round(mat,digits=5)

plot(gs,col="red",type="o")
```

Βιβλιογραφία

Ελληνική

[1] Αντζουλάκος Δ., 2016, *Σημειώσεις Π.Μ.Σ. «Αναλογιστική Επιστήμη & Διοικητική Κινδύνου»*, Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.

[2] Ηλιόπουλος Γ., 2013, *Βασικές Μέθοδοι Εκτίμησης Παραμέτρων με σημείο και με διάστημα*, Εκδόσεις Αθ. Σταμούλης.

[3] Κούτρας Μ., 2004, *Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές*, Μέρος Ι, Εκδόσεις Αθ. Σταμούλης.

[4] Κούτρας Μ., 2005, *Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές*, Μέρος ΙΙ, Εκδόσεις Αθ. Σταμούλης.

[5] Κούτρας Μ., 2006, *Εισαγωγή στη Συνδυαστική*, Εκδόσεις Αθ. Σταμούλης.

[6] Κουτσόπουλος Κ.Ι., 1999, *Αναλογιστικά Μαθηματικά*, Μέρος Ι, Θεωρία των Κινδύνων, Εκδόσεις Συμμετρία

[7] Κυπαρισσίδης Κ., 2010, *Δυναμική Ανάλυση των Φυσικών και Χημικών Συστημάτων*, Εκδόσεις Α. Τζιόλα & Υιοί ΑΕ.

[8] Πολίτης Κ., 2012, *Εισαγωγή στην Θεωρία Συλλογικού Κινδύνου, Το Συλλογικό Πρότυπο και Θεωρία Χρεοκοπίας*, Εκδόσεις Αθ. Σταμούλης.

[9] Χατζηκωνσταντινίδης Ε., 2015, *Σημειώσεις Π.Μ.Σ. «Αναλογιστική Επιστήμη & Διοικητική Κινδύνου»*, Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.

Ξένη

[10] Daykin C.D., Pentikäinen T., Pesonen M., 1994, *Practical Risk Theory for Actuaries*, Chapman & Hall, London.

[11] Dickson D.C.M., 2005, *Insurance Risk and Ruin*, Centre for Actuarial Studies, Department of Economics, University of Melbourne.

[12] Goulet V., Auclair S., Dutang C., Milhaud X., Ouellet T., Parent A., Pigeon M., Pouliot L.P., 2018, *Package "actuar", Actuarial Functions and Heavy Tailed Distributions*, CRAN.

[13] Heideman M.T., Johnson D. H., Burrus C. S., 1985, *Gauss and the History of the Fast Fourier Transform*, Archive for History of Exact Sciences (Springer).

[14] Klugman, S. A., Panjer, H. H., Willmot, G. E., 2012, *Loss Models, From Data to Decisions*, Fourth Edition, Wiley.

[15] Laplace P.S., 1812, *Theorie analytique des probabilites*, Paris.

[16] Panjer H.H., 1981, *Recursive evaluation of a family of compound distributions* ASTIN Bulletin.

[17] Yan J., 2007, *"Enjoy the Joy of Copulas: With a Package copula"*, Journal of Statistical Software, University of Connecticut.

Διπλωματικές Εργασίες

[18] Καγκελάρη Α., 2010, *Ναυτασφαλιστικές απαιτήσεις σε περίπτωση ολικής απώλειας πλοίου*, Διπλωματική Εργασία, Τμήμα Ναυτιλιακών Σπουδών, ΠΜΣ στην Ναυτιλία, Πανεπιστήμιο Πειραιώς.

[19] Μπιλάς Ε., 2018, *Η κλάση κατανομών (α,β,0) του Panjer: Θεωρία, γενικεύσεις και εφαρμογές στη θεωρία συλλογικού κινδύνου*, Διπλωματική Εργασία, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, ΠΜΣ στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου, Πανεπιστήμιο Πειραιώς.

Ιστοσελίδες (Ημερομηνία τελευταίας επίσκεψης 24.10.2018)

https://cran.r-project.org/web/packages/actuar/index.html

https://cran.r-project.org/

https://cran.r-project.org/web/packages/actuar/vignettes/simulation.pdf

https://www.museumoflondon.org.uk/discover/how-great-fire-london-created-insurance

https://en.wikipedia.org/wiki/Actuary