

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ
ΣΤΑΤΙΣΤΙΚΗΣ**

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ
ΕΠΙΣΤΗΜΗΣ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ
ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ**

**ΜΕΘΟΔΟΙ ΠΡΟΣΕΓΓΙΣΗΣ ΤΗΣ
ΚΑΤΑΝΟΜΗΣ ΤΩΝ ΣΥΝΟΛΙΚΩΝ
ΑΠΟΖΗΜΙΩΣΕΩΝ**

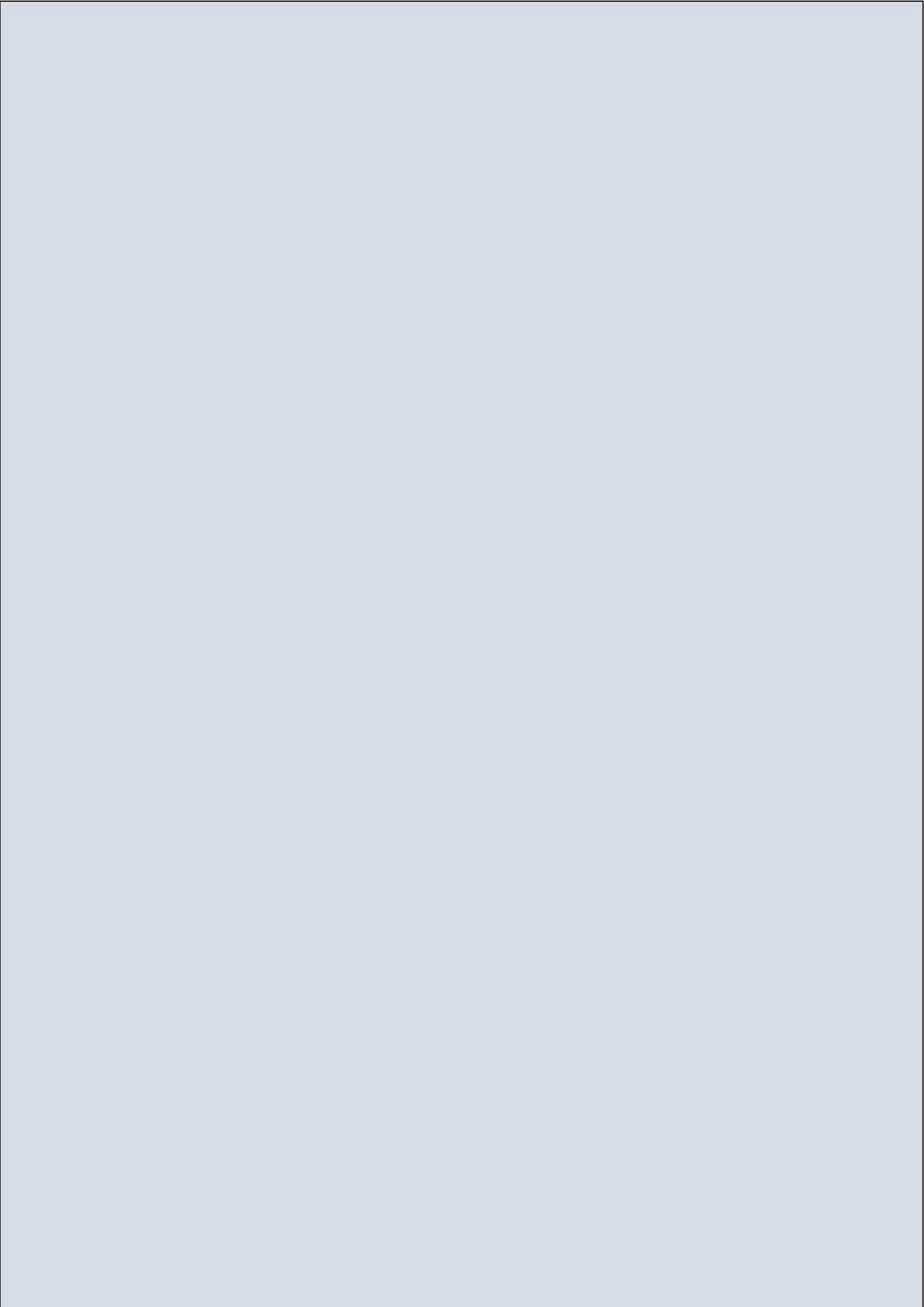
Νάσσος Παναγιώτης

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

Σεπτέμβριος 2021



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ
ΣΤΑΤΙΣΤΙΚΗΣ**

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ
ΕΠΙΣΤΗΜΗΣ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ
ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ**

**ΜΕΘΟΔΟΙ ΠΡΟΣΕΓΓΙΣΗΣ ΤΗΣ
ΚΑΤΑΝΟΜΗΣ ΤΩΝ ΣΥΝΟΛΙΚΩΝ
ΑΠΟΖΗΜΙΩΣΕΩΝ**

Νάσσος Παναγιώτης

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

Σεπτέμβριος 2021

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

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής Αντζουλάκος Δημήτριος (Επιβλέπων)
- Αναπληρωτής Καθηγητής Τζαβελάς Γεώργιος
- Επίκουρος Καθηγητής Πιτσέλης Γεώργιος

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

UNIVERSITY OF PIRAEUS



SCHOOL OF FINANCE AND STATISTICS

**DEPARTMENT OF STATISTICS AND
INSURANCE SCIENCE**

POSTGRADUATE PROGRAM IN

ACTUARIAL SCIENCE

AND RISK MANAGEMENT

**APPROXIMATION METHODS FOR
THE DISTRIBUTION OF AGGREGATE
CLAIMS**

By

Nassos Panagiotis

MSc Dissertation

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

Piraeus, Greece

September 2021

Στους γονείς μου,
Φώτη και Κατερίνα.

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα για την παρούσα διπλωματική εργασία Αναπληρωτή Καθηγητή κ. Δημήτριο Αντζουλάκο για την πολύτιμη καθοδήγηση που μου προσέφερε κατά τη διάρκεια εκπόνησης της εργασίας. Επίσης θα ήθελα να ευχαριστήσω και τα άλλα δύο μέλη της Τριμελούς Εξεταστικής Επιτροπής, τον Αναπληρωτή Καθηγητή κ. Γεώργιο Τζαβελά και τον Επίκουρο Καθηγητή κ. Γεώργιο Πιτσέλη για την τιμή που μου έκαναν να είναι μέλη της επιτροπής. Ακόμη θα ήθελα να ευχαριστήσω το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης που μου έδωσε την δυνατότητα να ασχοληθώ με την εν λόγω εργασία. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου που με στήριξαν σε όλη αυτή την προσπάθεια ολοκλήρωσης του μεταπτυχιακού κύκλου σπουδών.

Περίληψη

Στην παρούσα διπλωματική εξετάζουμε μια σειρά από προσεγγίσεις που έχουν προταθεί και χρησιμοποιούνται στα αναλογιστικά και ασφαλιστικά μαθηματικά και ιδιαιτέρως στη θεωρία συλλογικού κινδύνου. Πιο συγκεκριμένα, στο συλλογικό πρότυπο της θεωρίας κινδύνου το μεγαλύτερο ενδιαφέρον εστιάζεται στην κατανομή των συνολικών αποζημιώσεων για την ασφαλιστική εταιρεία. Με βάση τη βιβλιογραφία, ο υπολογισμός της κατανομής αυτής τις περισσότερες φορές είναι αδύνατον να γίνει με χρήση κάποιου κλειστού τύπου ή έστω μέσω μιας απλής επαναληπτικής διαδικασίας με εξαίρεση ίσως κάποιες πολύ συγκεκριμένες υποθέσεις για την κατανομή του ύψους των ατομικών ζημιών, όπως η Εκθετική κατανομή ή η Γάμμα κατανομή. Για το λόγο αυτό, έχουν προταθεί πολλές προσεγγιστικές μέθοδοι με στόχο την όσο το δυνατόν καλύτερη προσέγγιση της κατανομής που μας ενδιαφέρει. Ορισμένες από τις μεθόδους αυτές έχουν ήδη ενσωματωθεί, με τη μορφή κατάλληλων αλγορίθμων στο πακέτο *actuar* της γλώσσας προγραμματισμού R και έτσι η εφαρμογή και η αξιολόγησή τους καθίσταται εφικτή και πλέον άμεση. Η παρούσα διπλωματική αξιοποιώντας πρότυπα διακριτών και συνεχών κατανομών χρησιμοποιεί τις εξής προσεγγίσεις, προκειμένου να μελετήσει την κατανομή των συνολικών αποζημιώσεων: Κανονική προσέγγιση (NA), Προσέγγιση με μετατοπισμένη Γάμμα κατανομής (TGA), Δυναμοκανονική προσέγγιση (NPA), Προσέγγιση Bowers, Προσέγγιση Haldane, Προσέγγιση Wilson – Hilferty. Τα αποτελέσματα που μελετώνται στις εφαρμογές της παρούσας εργασίας αναδεικνύουν τις ιδιαιτερότητες αλλά και τα ισχυρά και αδύναμα στοιχεία κάθε μεθόδου.

Abstract

In this dissertation we examine a number of approaches that have been proposed and used in the theory of collective risk. More specifically, in the collective model of risk theory, the greatest interest is focused on the distribution of total compensation for the insurance company. Based on the literature, the calculation of this distribution is almost impossible using a closed form or even through a simple iterative process except for specific distributions of individual damages. For this reason, several approximation methods have been proposed for this calculation. Some of these methods have already been integrated, in the form of suitable algorithms in the actuar package of the R programming language and thus their implementation and evaluation becomes possible. The present dissertation utilizing models of discrete and continuous distributions uses the following approaches in order to study the issue of the distribution of total compensation: Normal approach (NA), Approach with Translated Gamma Approach (TGA), Normal Power Approach (NPA), Bowers Approach, Haldane Approach and Wilson-Hilferty Approach. The results highlight the peculiarities but also the strengths and weaknesses of each method.

Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή	1
1.1 Το πρόβλημα	1
1.2 Το μοντέλο συλλογικού κινδύνου	3
1.3 Πρότυπα διακριτών κατανομών	4
1.3.1 Διωνυμική κατανομή.....	4
1.3.2 Κατανομή Poisson	4
1.3.3 Γεωμετρική κατανομή.....	5
1.3.4 Αρνητική Διωνυμική κατανομή	5
1.4 Πρότυπα συνεχών κατανομών.....	6
1.4.1 Εκθετική κατανομή	6
1.4.2 Κατανομή Γάμμα.....	6
1.4.3 Κανονική κατανομή.....	7
1.4.4 Κατανομή Weibull	7
1.4.5 Κατανομή Pareto	8
1.5 Μελέτη της κατανομής των συνολικών αποζημιώσεων.....	9
1.5.1 Μέθοδος των συνελίξεων.....	11
1.5.2 Αναδρομικές σχέσεις	13
Κεφάλαιο 2: Μέθοδοι προσέγγισης της κατανομής των συνολικών αποζημιώσεων	18
2.1 Κανονική προσέγγιση (NA)	19
2.2 Προσέγγιση με μετατοπισμένη Γάμμα κατανομή (TGA)	19
2.3 Δυναμοκανονική προσέγγιση (NPA)	20
2.4 Η προσέγγιση Bowers	22
2.5 Η προσέγγιση Haldane	25
2.6 Η προσέγγιση Wilson - Hilferty	27
2.7 Ανάπτυγμα Edgeworth	28
Κεφάλαιο 3: Εφαρμογές και σύγκριση των μεθόδων προσέγγισης	31
3.1 1 ^η Εφαρμογή	31
3.2 2 ^η Εφαρμογή	48
3.3 3 ^η Εφαρμογή	56
3.4 Συμπεράσματα.....	63
Παράρτημα	64
Βιβλιογραφία	87

Κεφάλαιο 1: Εισαγωγή

1.1 Το πρόβλημα

Ο πιστωτικός κίνδυνος πάντα υπήρξε ένα κύριο πρόβλημα της διαχείρισης κινδύνου στον χρηματοοικονομικό κλάδο. Οι τρεις βασικοί παράγοντες που επηρεάζουν τον πιστωτικό κίνδυνο των χρηματοοικονομικών δανειστών είναι η πιθανότητα αθέτησης υποχρεώσεων, το ποσοστό ανάκτησης και το άνοιγμα κατά την αθέτηση, σύμφωνα με τη Βασιλεία III (BCBS, 2011).

Για τη διαχείριση του κινδύνου, τα χρηματοπιστωτικά ιδρύματα και οι ασφαλιστικές εταιρείες έχουν προσπαθήσει να κατασκευάσουν ευφυή συστήματα ώστε να εκτιμήσουν την πιθανότητα αθέτησης των υποχρεώσεων από κάποιο πελάτη, συνθήκη η οποία μπορεί να οδηγήσει σε χρεωκοπία. Στη σύγχρονη εποχή της θεωρίας κινδύνου και των αναλογιστικών μαθηματικών, η εύρεση και η περιγραφή της κατανομής των συνολικών απαιτήσεων μέχρι να συμβεί η χρεωκοπία είναι πολύ σημαντική για τη μελέτη ασφαλιστικών χαρτοφυλακίων. Στη θεωρία κινδύνου, το μοντέλο συλλογικού κινδύνου περιγράφει την διαδοχική εμφάνιση ζημιών στα ασφαλιστικά χαρτοφυλάκια, μοντελοποιεί τόσο το εκτιμώμενο συνολικό ύψος όσο και τις διακυμάνσεις που έχει το αντίστοιχο απόθεμα της ασφαλιστικής εταιρείας το οποίο φυσικά υπάρχει ώστε να είναι δυνατή η αποπληρωμή των ζημιών που πιθανόν να προκύψουν. Είναι σαφές ότι όταν το συνολικό ύψος των ζημιών ξεπερνά το απόθεμα της εταιρείας τότε επίκειται η χρεωκοπία. Στόχος κάθε εταιρείας είναι η ταυτόχρονη εξασφάλιση και η παροχή αξιοπιστίας στους πελάτες της ότι θα υπάρχει κάλυψη των ζημιών τους σε οτιδήποτε προβλέπει το συμβόλαιο τους. Οι πελάτες από την άλλη πληρώνουν στην εταιρία το συμφωνηθέν ασφάλιστρο το οποίο εξασφαλίζει τη βιωσιμότητα της εταιρίας και βέβαια συμβάλει στη δημιουργία/συντήρηση αποθέματος. Οι ασφαλιστικές εταιρείες καθορίζουν το ασφάλιστρο έτσι ώστε αυτό να είναι γνησίως μεγαλύτερο από το μέσο κόστος των ζημιών σε κάθε χρονική στιγμή, διαφορετικά λόγω του νόμου των μεγάλων αριθμών μακροπρόθεσμα θα οδηγούνταν στην χρεωκοπία.

Όλα τα αναλογιστικά μοντέλα που υπάρχουν μελετούν την σχέση όλων των εμπλεκόμενων μεταβλητών. Οι συνολικές απαιτήσεις είναι ένα τυχαίο άθροισμα της μορφής $S_N = X_1 + X_2 + \dots + X_N$ όπου στη γενική περίπτωση το πλήθος των απαιτήσεων N είναι μια τυχαία μεταβλητή, ενώ οι τυχαίες μεταβλητές των επιμέρους απαιτήσεων X_i είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές. Η κατανομή των συνολικών απαιτήσεων ωστόσο είναι εν γένει από δύσκολο ως αδύνατον να βρεθεί. Υπάρχουν αναλυτικά αποτελέσματα στην ειδική περίπτωση όπου οι αποζημιώσεις ακολουθούν την Εκθετική ή την Γάμμα κατανομή ή την κανονική κατανομή, καθώς σε αυτή την περίπτωση οι υπολογισμοί για την αναγνώριση της κατανομής μέσω ροπογεννητριών μπορούν να γίνουν αλγεβρικά. Πριν την ανάπτυξη των υπολογιστών, η κατανομή μπορούσε να προσεγγιστεί μόνο από κεντρικά οριακά θεωρήματα, ή κάποια άμεση επαναληπτική σχέση που ικανοποιούσε η κατανομή. Με την ανάπτυξη των υπολογιστών αναπτύχθηκαν και τα εργαλεία μελέτης και για την περίπτωση πιο γενικών κατανομών. Έτσι από

το 1981 και έπειτα μέσω της δημοσίευσης του Panjer στο διεθνές αναλογιστικό περιοδικό *Astin Bulletin (Actuarial Studies In Non-Life Insurance)*, διατυπώθηκε ένας αναδρομικός τύπος υπολογισμού της συνάρτησης πιθανότητας των συνολικών απαιτήσεων σε ένα χαρτοφυλάκιο υποθέτοντας ότι αυτό έχει επιμέρους ζημιές οι οποίες είναι διακριτές τυχαίες μεταβλητές. Στην αρχή οι αναδρομικοί τύποι που δόθηκαν αφορούσαν μια κλάση διακριτών κατανομών που περιείχε την κατανομή Poisson, την Διωνυμική αλλά και την Αρνητική Διωνυμική κατανομή. Από τότε αναπτύχθηκαν αρκετές προσεγγιστικές μέθοδοι με σκοπό την προσέγγιση της συνάρτησης κατανομής του ύψους των συνολικών αποζημιώσεων και ιδιαιτέρως της ουράς της κατανομής, δηλαδή την προσέγγιση της συνάρτησης $\bar{G}(x) = P(S > x)$ για μεγάλες τιμές του x , όταν αυτές πληρούν συγκεκριμένες ιδιότητες.

Οι προσεγγιστικές μέθοδοι που θα παρουσιαστούν στην παρούσα εργασία χρησιμοποιούνται ευρέως στην βιβλιογραφία και είναι δυνατόν να εφαρμοστούν σχεδόν για όλες τις κατανομές. Στο παρόν κεφάλαιο γίνεται παρουσίαση των ορισμών και των απαραίτητων εργαλείων των πιθανοτήτων και συγκεκριμένα αναφέρονται τα πρότυπα διακριτών και συνεχών κατανομών, οι συνελίξεις και οι σύνθετες κατανομές. Στο δεύτερο κεφάλαιο παρουσιάζονται όλες οι προσεγγιστικές μέθοδοι μαζί με τα απαραίτητα μέτρα υπολογισμού που απαιτεί η κάθε μέθοδος προκειμένου να εφαρμοστεί και εν συνεχεία, στο τρίτο κεφάλαιο συγκρίνονται οι προσεγγιστικές μέθοδοι μέσω κάποιων αριθμητικών εφαρμογών με σκοπό την εξαγωγή συμπερασμάτων σχετικά με το ποιες μέθοδοι προσεγγίζουν καλύτερα την κατανομή των συνολικών αποζημιώσεων ανά περίπτωση και ποιες όχι.

Το υλικό που παρουσιάζεται σε αυτό το κεφάλαιο βασίζεται κυρίως στα ακόλουθα άρθρα, βιβλία και σημειώσεις: Αντζουλάκος (2017), Κούτρας (2004), Πολίτης (2012), Χατζηκωνσταντινίδης (2018), Kouassi, E., Akpata, E., Pokou, K. (2020), Sundt B., Jewell W., (1981).

1.2 Το μοντέλο συλλογικού κινδύνου

Το μοντέλο συλλογικού κινδύνου βοηθά στην ορθή περιγραφή της εμφάνισης των ζημιών και των συνολικών αποζημιώσεων στα πλαίσια ενός ασφαλιστικού χαρτοφυλακίου σε ένα συγκεκριμένο χρονικό διάστημα. Με τη χρήση τυχαίων μεταβλητών είναι εφικτό να μοντελοποιηθεί και να μελετηθεί το ύψος και η διακύμανση των συνολικών ζημιών και κατά συνέπεια και του ασφαλιστικού αποθέματος έτσι ώστε να καθοριστεί εάν η ασφαλιστική εταιρεία θα μπορεί να είναι συνεπής στην πληρωμή των αποζημιώσεων. Στόχος είναι να μπορεί η εταιρεία να καλύπτει τις υποχρεώσεις της με μη ζημιογόνο τρόπο και παράλληλα οι πελάτες να είναι ευχαριστημένοι από την υπηρεσία. Η βιωσιμότητα της εταιρείας εξαρτάται από την καταβολή των ασφαλίσεων από τους πελάτες, το οποίο χρησιμοποιείται για την αύξηση ή τήρηση του ήδη υπάρχοντος αποθέματος. Θα πρέπει να σημειωθεί ότι το ύψος των ασφαλίσεων θα πρέπει να είναι γνησίως μεγαλύτερο από το μέσο κόστος των ζημιών. Τα αναλογιστικά μοντέλα παρέχουν τη δυνατότητα να μελετηθεί το είδος της σχέσης που έχουν τα μεγέθη μεταξύ τους αλλά επίσης παρέχουν άμεσο τρόπο υπολογισμού του ύψους των ασφαλίσεων, του ασφαλιστικού κινδύνου και των αποθεμάτων.

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

$$S = I(N \geq 1) \sum_{j=1}^N X_j = \begin{cases} 0, & N = 0 \\ \sum_{j=1}^N X_j, & N \geq 1 \end{cases}$$

όπου:

- (i) οι τυχαίες μεταβλητές X_j παριστάνουν τα μεγέθη των ατομικών αποζημιώσεων (δεν μπορούν να πάρουν αρνητικές τιμές),
- (ii) η τυχαία μεταβλητή N , παριστάνει το πλήθος των ζημιών (κινδύνων) ή και απαιτήσεων/αξιώσεων,
- (iii) η τυχαία μεταβλητή S , παριστάνει το ύψος των συνολικών απαιτήσεων.

Στις δύο επόμενες παραγράφους δίνονται οι βασικότερες κατανομές που μπορούν να περιγράψουν την κατανομή των τυχαίων μεταβλητών N και X που θα μας απασχολήσουν στα επόμενα κεφάλαια.

1.3 Πρότυπα διακριτών κατανομών

1.3.1 Διωνυμική κατανομή

Λέμε ότι μια μη αρνητική τυχαία μεταβλητή N ακολουθεί την Διωνυμική κατανομή με παραμέτρους έναν φυσικό αριθμό m και $0 < p < 1$, εάν έχει συνάρτηση πιθανότητας που δίνεται από τον τύπο:

$$P(N = k) = \binom{m}{k} p^k (1-p)^{m-k}, \quad k = 0, 1, 2, \dots, m$$

και γράφουμε ότι $N \sim B(m, p)$.

Η πιθανογεννήτρια συνάρτηση $P_N(t)$ της N είναι

$$P_N(t) = E(t^N) = (1 - p + pt)^m.$$

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της N έχουμε ότι

$$\mu_N = mp, \quad \sigma_N^2 = mp(1-p), \quad \gamma_N = \frac{1-2p}{\sqrt{mp(1-p)}}.$$

1.3.2 Κατανομή Poisson

Λέμε ότι μια μη αρνητική τυχαία μεταβλητή N ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$, εάν έχει συνάρτηση πιθανότητας που δίνεται από τον τύπο

$$P(N = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

και γράφουμε ότι $N \sim P(\lambda)$.

Η πιθανογεννήτρια συνάρτηση $P_N(t)$ της N είναι

$$P_N(t) = E(t^N) = \exp(\lambda(t-1)).$$

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της N έχουμε ότι

$$\mu_N = \lambda, \quad \sigma_N^2 = \lambda, \quad \gamma_N = \frac{1}{\sqrt{\lambda}}.$$

1.3.3 Γεωμετρική κατανομή

Λέμε ότι μια μη αρνητική τυχαία μεταβλητή N ακολουθεί την γεωμετρική κατανομή με παράμετρο $0 < p < 1$, εάν έχει συνάρτηση πιθανότητας που δίνεται από τον τύπο

$$P(N = k) = (1 - p)^k p, \quad k = 0, 1, 2, \dots$$

και γράφουμε ότι $N \sim G_0(p)$.

Αν $Y = N + 1$, η συνάρτηση πιθανότητας της Y δίνεται από τον τύπο

$$P(Y = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \dots$$

όπου σε αυτή την περίπτωση θα γράφουμε $Y \sim G(p)$. Η διαφορά μεταξύ των $G_0(p)$ και $G(p)$ έγκειται στο γεγονός ότι η $G_0(p)$ χρησιμοποιείται όταν μοντελοποιούμε το πλήθος αποτυχιών σε ένα παιχνίδι στο οποίο γίνονται διαδοχικές προσπάθειες μέχρι την επίτευξη της πρώτης επιτυχίας με σταθερή πιθανότητα προσπάθειας $0 < p < 1$, ενώ η $G(p)$ χρησιμοποιείται όταν μοντελοποιούμε το πλήθος των προσπαθειών σε ένα παιχνίδι στο οποίο γίνονται διαδοχικές προσπάθειες μέχρι την επίτευξη της πρώτης επιτυχίας με σταθερή πιθανότητα προσπάθειας $0 < p < 1$.

Η πιθανογεννήτρια συνάρτηση $P_N(t)$ της N είναι

$$P_N(t) = E(t^N) = \frac{p}{1 - qt}, \quad |t| < \frac{1}{q}$$

ενώ για την περίπτωση της Y είναι:

$$P_Y(t) = E(t^Y) = \frac{pt}{1 - qt}, \quad |t| < \frac{1}{q}$$

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της N έχουμε

$$\mu_N = \frac{q}{p}, \quad \sigma_N^2 = \frac{q}{p^2}, \quad \gamma_N = \frac{1 + q}{\sqrt{q}}$$

1.3.4 Αρνητική Διωνυμική κατανομή

Λέμε ότι μια μη αρνητική τυχαία μεταβλητή N ακολουθεί την αρνητική Διωνυμική κατανομή με παραμέτρους r, p ($r > 0, 0 < p < 1$), εάν έχει συνάρτηση πιθανότητας που δίνεται από τον τύπο

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

και γράφουμε ότι $N \sim NB_0(r, p)$.

Αν $Y = N + r$, η συνάρτηση πιθανότητας της Y δίνεται από τον τύπο

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

και τότε γράφουμε ότι $N \sim NB(r, p)$.

Για την τυχαία μεταβλητή N έχουμε ότι

$$\mu_N = \frac{rq}{p}, \quad \sigma_N^2 = \frac{rq}{p^2}, \quad \gamma_N = \frac{1+q}{\sqrt{rq}}.$$

Επίσης η πιθανογεννήτρια συνάρτηση της τ.μ. N είναι

$$P_N(t) = E(t^N) = \left(\frac{p}{1-qt} \right)^r, \quad |t| < \frac{1}{q}.$$

1.4 Πρότυπα συνεχών κατανομών

1.4.1 Εκθετική κατανομή

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την Εκθετική κατανομή με παράμετρο $\theta > 0$, αν η συνάρτηση πυκνότητας της δίνεται από τον τύπο

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

και γράφουμε ότι $X \sim E(\theta)$.

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της X έχουμε ότι

$$\mu_X = \frac{1}{\theta}, \quad \sigma_X^2 = \frac{1}{\theta^2}, \quad \gamma_X = 2.$$

Ο μετασχηματισμός Laplace της $f_X(x)$ είναι:

$$\hat{f}_X(t) = \frac{\theta}{\theta + t}, \quad t > -\lambda.$$

1.4.2 Κατανομή Γάμμα

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Γάμμα με παραμέτρους $a, \theta > 0$, αν η συνάρτηση πυκνότητας της δίνεται από τον τύπο

$$f_X(x) = \frac{x^{a-1}e^{-x/\theta}}{\theta^a \Gamma(a)}, \quad x, a, \theta > 0$$

και γράφουμε ότι $X \sim G(a, \theta)$.

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της X έχουμε ότι:

$$\mu_X = a\theta, \quad \sigma_X^2 = a\theta^2, \quad \gamma_X = \frac{2}{\sqrt{a}}.$$

Ο μετασχηματισμός Laplace της $f_X(x)$ είναι:

$$\hat{f}_X(t) = \left(\frac{1}{1 + \theta t} \right)^a, \quad t > -1/\theta.$$

Επιπρόσθετα ισχύει ότι αν X_1, X_2, \dots, X_n είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με κατανομή $E(\theta)$ με τότε $X_1 + X_2 + \dots + X_n \sim G(n, \theta)$.

1.4.3 Κανονική κατανομή

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή με παραμέτρους $\mu \in \mathbb{R}$, $\sigma > 0$, αν η συνάρτηση πυκνότητάς της δίνεται από τον τύπο

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x, \mu \in \mathbb{R}, \quad \sigma > 0$$

και γράφουμε ότι $X \sim N(\mu, \sigma^2)$.

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της X έχουμε ότι

$$\mu_X = \mu, \quad \sigma_X^2 = \sigma^2, \quad \gamma_X = 0.$$

Η ροπογεννήτρια συνάρτηση της X είναι

$$M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right), \quad t \in \mathbb{R}.$$

1.4.4 Κατανομή Weibull

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την Weibull κατανομή με παραμέτρους $\kappa, \lambda > 0$, αν η συνάρτηση πυκνότητάς της δίνεται από τον τύπο

$$f_X(x) = \frac{\kappa}{\lambda} \cdot \left(\frac{x}{\lambda}\right)^{\kappa-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^\kappa\right\}, \quad x \geq 0.$$

Συμβολικά θα γράφουμε $X \sim Weibull(\kappa, \lambda)$.

Για την κατανομή αυτή ισχύει ότι

$$\begin{aligned}\mu_X &= \lambda \Gamma\left(1 + \frac{1}{\kappa}\right), \\ \sigma_X^2 &= \lambda^2 \left(\Gamma\left(1 + \frac{2}{\kappa}\right) - \Gamma^2\left(1 + \frac{1}{\kappa}\right) \right), \\ \gamma_X &= \frac{\Gamma\left(1 + \frac{3}{\kappa}\right) \lambda^3 - 3\mu_X \sigma_X^2 - \mu_X^3}{\sigma_X^3}.\end{aligned}$$

Αν θέσουμε $b = \lambda^{-\kappa}$ τότε η συνάρτηση πυκνότητας μπορεί να γραφτεί στην εναλλακτική μορφή

$$f_X(x) = kb \cdot x^{k-1} \exp\{-bx^k\}, \quad x \geq 0.$$

Η ροπογεννήτρια συνάρτηση της X είναι

$$M(t) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!} \Gamma\left(1 + \frac{n}{\kappa}\right), \quad k \geq 1.$$

1.4.5 Κατανομή Pareto

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $a, \beta > 0$, αν η συνάρτηση πυκνότητάς της δίνεται από τον τύπο

$$f_X(x) = \frac{a\beta^a}{(\beta+x)^{a+1}}, \quad x \geq 0.$$

και γράφουμε ότι $X \sim Pa(a, \beta)$. Επίσης

$$F_X(x) = 1 - \left(\frac{\beta}{x+\beta}\right)^a, \quad x \geq 0.$$

Για την κατανομή αυτή ισχύει ότι:

$$E(X^k) = \frac{k! \beta^k}{(a-1)(a-2) \cdots (a-k)}, \quad a > k.$$

Έτσι είναι εύκολο να διαπιστωθεί ότι:

$$\begin{aligned}\mu_X &= \frac{\beta}{a-1}, \\ \sigma_X^2 &= \frac{a\beta^2}{(a-1)^2(a-2)}, \quad a > 2,\end{aligned}$$

$$\gamma_X = \frac{2(1+a)}{a-3} \sqrt{\frac{a-2}{a}}, \quad a > 3.$$

1.5 Μελέτη της κατανομής των συνολικών αποζημιώσεων

Η κατανομή των συνολικών αποζημιώσεων S περιγράφεται από το τυχαίο άθροισμα

$$S = I(N \geq 1) \sum_{j=1}^N X_j = \begin{cases} 0, & N = 0 \\ \sum_{j=1}^N X_j, & N \geq 1 \end{cases}$$

όπου:

- (i) η τυχαία μεταβλητή X_j παριστάνει το ύψος της j ατομικής ζημιάς (δεν μπορεί να πάρει αρνητικές τιμές),
- (ii) η τυχαία μεταβλητή N , παριστάνει το πλήθος των ζημιών (κινδύνων) ή και απαιτήσεων/αξιώσεων,
- (iii) η τυχαία μεταβλητή S , παριστάνει το ύψος της συνολικής ζημιάς.

Στο παραπάνω μοντέλο η τυχαία μεταβλητή S μπορεί να είναι διακριτή, συνεχής ή μικτού τύπου. Η τυχαία μεταβλητή N , είναι πάντοτε μια μη αρνητική ακέραια (διακριτή) τυχαία μεταβλητή, με συνάρτηση πιθανότητας

$$p_n = p(n) = P(N = n), \quad n = 0, 1, 2, \dots$$

Έτσι, έχουμε τις ακόλουθες περιπτώσεις:

- ✓ αν η X είναι διακριτή τυχαία μεταβλητή, τότε η S είναι και αυτή διακριτή τυχαία μεταβλητή, επειδή σ' αυτή την περίπτωση η S είναι άθροισμα διακριτών τυχαίων μεταβλητών
- ✓ αν η X είναι συνεχής τυχαία μεταβλητή, τότε η S μπορεί να είναι είτε μια συνεχής είτε μια μικτού τύπου τυχαία μεταβλητή, ανάλογα με την τιμή της πιθανότητας $p_0 = P(N = 0)$
 - αν $P(N = 0) = 0$, τότε προφανώς η τυχαία μεταβλητή S είναι μια συνεχής τυχαία μεταβλητή, ως άθροισμα συνεχών τυχαίων μεταβλητών,
 - αν $P(N = 0) > 0$, τότε η τυχαία μεταβλητή S είναι μικτού τύπου. Η τυχαία μεταβλητή S έχει μάζα πιθανότητας στο σημείο μηδέν, με

$$P(S = 0) = P(N = 0) = p_0$$
 και είναι συνεχής στο διάστημα $(0, +\infty)$.

Μια κλασική υπόθεση για την κατανομή της τυχαίας μεταβλητής N είναι ότι αυτή έχει την κατανομή Poisson με παράμετρο λ ($N \sim P(\lambda)$, $\lambda > 0$). Μια γενίκευση

αυτής της υπόθεσης είναι η περίπτωση που η κατανομή της N είναι μια μικτή κατανομή Poisson, δηλαδή " $N|Λ=λ \sim P(λ)$ " όπου $Λ$ είναι μια τυχαία μεταβλητή με συνάρτηση κατανομής H . Τότε

$$P(N = k) = \int_0^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} dH(\lambda), \quad k = 0, 1, 2, \dots$$

Παράδειγμα 1.1. Αν " $N|Λ=λ \sim P(λ)$ " και $Λ \sim G(a, \theta)$, τότε

$$P(N = k) = \int_0^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot \frac{\lambda^{a-1} e^{-\lambda/\theta}}{\theta^a \Gamma(a)} d\lambda = \frac{1}{k! \theta^a \Gamma(a)} \int_0^{\infty} \lambda^{k+a-1} \cdot e^{-\lambda(1+1/\theta)} d\lambda.$$

Θέτοντας $y = \lambda \left(1 + \frac{1}{\theta}\right) = \lambda \left(\frac{\theta+1}{\theta}\right)$, προκύπτει ότι

$$\begin{aligned} P(N = k) &= \left(\frac{\theta}{\theta+1}\right)^{k+a} \frac{1}{k! \theta^a \Gamma(a)} \int_0^{\infty} y^{k+a-1} \cdot e^{-y} dy \\ &= \frac{\Gamma(k+a)}{k! \Gamma(a)} \left(\frac{1}{1+\theta}\right)^a \left(\frac{\theta}{1+\theta}\right)^k. \end{aligned}$$

Δηλαδή καταλήγουμε στο γνωστό αποτέλεσμα ότι η τυχαία μεταβλητή N έχει την αρνητική Διωνυμική κατανομή με παραμέτρους a και $\frac{1}{1+\theta}$ ($N \sim NB_0(a, \frac{1}{1+\theta})$).

Για τη ροπογεννήτρια $M_S(t)$ της S έχουμε ότι

$$M_S(t) = M_N[\log M_X(t)] = P_N[P_X(t)],$$

όπου $P_N(t)$ η πιθανογεννήτρια της διακριτής τυχαίας μεταβλητής N .

Για τη μέση τιμή, τη διακύμανση και το συντελεστή ασυμμετρίας της S έχουμε ότι

$$\begin{aligned} \mu_S &= E(S) = E(N)E(X) = \mu_N \mu_X \\ \sigma_S^2 &= Var(S) = E[(S - \mu_S)^2] = \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2 \\ \gamma_S &= E\left(\left(\frac{S - \mu_S}{\sigma}\right)^3\right) = \frac{\gamma_N (\sigma_N^2)^{3/2} \mu_X^3 + 3\mu_X \sigma_N^2 \sigma_X^2 + \mu_N \gamma_X (\sigma_X^2)^{3/2}}{(\mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2)^{3/2}} \end{aligned}$$

Παράδειγμα 1.2. Ας υποθέσουμε ότι $N \sim G(p)$, δηλαδή

$$P(N = n) = pq^{n-1}, \quad n = 1, 2, \dots$$

και $X \sim E(\lambda)$, δηλαδή

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

Τότε

$$M_{S_N}(t) = P_N[M_X(t)] = \frac{p \left(\frac{\lambda}{\lambda-t} \right)}{1 - q \left(\frac{\lambda}{\lambda-t} \right)} = \frac{p\lambda}{\lambda - t - q\lambda} = \frac{p\lambda}{p\lambda - t},$$

δηλαδή $S \sim E(p\lambda)$.

1.5.1 Μέθοδος των συνελίξεων

Θεωρητικά ο ακριβής υπολογισμός της συνάρτησης κατανομής της S επιτυγχάνεται με τη μέθοδο των συνελίξεων αφού

$$F_S(s) = P(S \leq s) = \sum_{n=0}^{\infty} P(N = n) F_X^{*n}(s), \quad s \geq 0,$$

όπου για $n \geq 1$

$$F_X^{*n}(s) = \int_0^s F_X^{*(n-1)}(s-t) dF_X(t)$$

$$= \begin{cases} \int_0^s F_X^{*(n-1)}(s-t) f_X(t) dt, & X \text{ συνεχής τ.μ.} \\ \sum_{t=0}^s F_X^{*(n-1)}(s-t) f_X(t), & X \text{ διακριτή τ.μ.} \end{cases}$$

και

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

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

$$f_S(s) = \sum_{n=1}^{\infty} p_n f_X^{*n}(s), \quad s > 0$$

και διακριτό μέρος με πιθανότητα μόνο στο 0 ίση με

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

Στον παραπάνω τύπο

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

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

$$f_S(s) = P r(S = s) = \sum_{n=0}^{\infty} p_n f_X^{*n}(s), \quad s = 0,1,2, \dots$$

Σε αυτή την περίπτωση

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

Ειδικότερα από τον παραπάνω τύπο προκύπτει ότι

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

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

Παράδειγμα 1.3. Ας υποθέσουμε για απλότητα ότι η τυχαία μεταβλητή N λαμβάνει τις τιμές 1,2,3 με ίση πιθανότητα, δηλαδή

$$P(N = 1) = P(N = 2) = P(N = 3) = \frac{1}{3}$$

και επιπλέον ότι

$$X_j \sim E(\theta), \quad j = 1, 2, 3.$$

Τότε, σύμφωνα με τα προηγούμενα θα έχουμε ότι για $s \geq 0$

$$\begin{aligned} F_S(s) &= P(S \leq s) = \sum_{n=1}^3 P(N = n)F_X^{*n}(s) \\ &= \frac{1}{3} \sum_{n=1}^3 F_X^{*n}(s) = \frac{1}{3} [F_X^{*1}(s) + F_X^{*2}(s) + F_X^{*3}(s)]. \end{aligned}$$

Χρησιμοποιώντας το γεγονός ότι το άθροισμα κ ανεξάρτητων Εκθετικών με παράμετρο θ ακολουθεί κατανομή Γάμμα με παραμέτρους κ και θ θα έχουμε:

$$F_S(s) = \frac{1}{3} \left[\int_0^s \theta e^{-\theta t} dt + \int_0^s \frac{\theta^2}{\Gamma(2)} t e^{-\theta t} dt + \int_0^s \frac{\theta^3}{\Gamma(3)} t^2 e^{-\theta t} dt \right], \quad s \geq 0$$

και επομένως η συνάρτηση πυκνότητας $f_S(s)$ θα έχει τύπο

$$f_S(s) = \frac{dF_S(s)}{ds} = \frac{1}{3} \left[\theta e^{-\theta s} + \theta^2 s e^{-\theta s} + \frac{\theta^3}{2} s^2 e^{-\theta s} \right], \quad s \geq 0.$$

1.5.2 Αναδρομικές σχέσεις

- **Κλάση κατανομών $R(a, b, 0)$**

Προκειμένου να υπολογιστεί η κατανομή της S ευκολότερα, ο Panjer (1981) έθεσε τα θεμέλια μιας νέας προσέγγισης για τον υπολογισμό της. Παρουσίασε ένα αναδρομικό τύπο υπολογισμού της συνάρτησης πιθανότητας της τ.μ. S θεωρώντας ότι η κατανομή της τ.μ. N ανήκει σε μία ευρεία οικογένεια κατανομών, την οικογένεια $R(a, b, 0)$.

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

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

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

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

Πίνακας 1.1. Η οικογένεια κατανομών $R(a, b, 0)$.

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

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

Θεώρημα 1.1. Αν η X είναι μια μη αρνητική ακέραια τυχαία μεταβλητή, δηλαδή $R_X = \{0, 1, 2, \dots\}$, με συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τυχαία μεταβλητή N ανήκει στην οικογένεια κατανομών $R(a, b, 0)$, τότε η $g(x) = P(S = x)$ ικανοποιεί την αναδρομική σχέση

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

Παράδειγμα 1.4. Αν η τυχαία μεταβλητή $N \sim B(m, p)$, $0 < p < 1$, $q = 1 - p$, δηλαδή

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

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

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

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

- **Κλάση κατανομών $R(a, b, 1)$**

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

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

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

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

Η κλάση κατανομών $R(a, b, 1)$, περιέχει ως μέλη της όλες τις κατανομές της κλάσης $R(a, b, 0)$. Επιπλέον, περιέχει ως μέλη της όλες τις zero-truncated και όλες τις zero-modified κατανομές, που αντιστοιχούν στην κλάση κατανομών $R(a, b, 0)$.

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

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

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

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

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

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

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

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

Πίνακας 1.2. Πίνακας με τις κατανομές που ανήκουν στην κλάση κατανομών $R(a, b, 1)$

ΚΑΤΑΝΟΜΗ	p_0	p_1	a	b	Περιορισμοί
Geometric [$G_0(p)$] $p_n = (1 - q)q^n$	$1 - q$	$(1 - q)q$	q	0	$0 < q < 1$

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

ZM Logarithmic	$p_0^{(M)}$	$-(1 - p_0^{(M)}) \times \frac{q}{\ln(1 - q)}$	q	$-q$	$0 < q < 1$
----------------	-------------	--	-----	------	-------------

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

Θεώρημα 1.2 Αν η X είναι μια μη αρνητική ακέραια τυχαία μεταβλητή, δηλαδή $R_X = \{0, 1, 2, \dots\}$, με συνάρτηση πιθανότητας $f(x) = P(X = x)$ και η τυχαία μεταβλητή N ανήκει στην οικογένεια κατανομών $R(a, b, 1)$, τότε η

$g(x) = P(S = x)$ ικανοποιεί την αναδρομική σχέση

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

Παράδειγμα 1.5. Αν η τυχαία μεταβλητή $N \sim ZTNB(r, p)$ με $(r > 0, 0 < p < 1,$

$q = 1 - p)$, δηλαδή

$$p_n = P(N = n) = \begin{cases} 0, & n = 0 \\ \frac{1}{1 - p^r} \binom{r + n - 1}{n} p^r q^n, & n = 1, 2, \dots \end{cases}$$

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

$$g(x) = \frac{\left[\frac{rqp^r}{1-p^r}\right] + \sum_{y=1}^x \left(q + (r-1)q \frac{y}{x}\right) f(y) g(x-y)}{1 - qf(0)}, \quad x = 1, 2, 3, \dots$$

με αρχική συνθήκη

$$g(0) = \frac{\left(\frac{p}{1 - (1-p)f(0)}\right)^r - p^r}{1 - p^r}.$$

Τα παραπάνω προκύπτουν με εφαρμογή του Θεωρήματος 2.2 για $a = q,$

$b = (r - 1)q$ και για $p_0 = 0$ και $p_1 = \frac{rqp^r}{1-p^r}.$

Κεφάλαιο 2: Μέθοδοι προσέγγισης της κατανομής των συνολικών αποζημιώσεων

Στο παρόν κεφάλαιο θα παρουσιάσουμε μερικές δημοφιλείς μεθόδους προσέγγισης της κατανομής των συνολικών ζημιών/αποζημιώσεων S . Θυμίζουμε ότι η τ.μ. S δίνεται από τη σχέση

$$S = I(N \geq 1) \sum_{j=1}^N X_j = \begin{cases} 0, & N = 0 \\ \sum_{j=1}^N X_j, & N \geq 1 \end{cases}$$

όπου:

- (i) η τυχαία μεταβλητή X_j παριστάνει το ύψος της j ατομικής ζημιάς (δεν μπορεί να πάρει αρνητικές τιμές),
- (ii) η τυχαία μεταβλητή N , παριστάνει το πλήθος των ζημιών (κινδύνων) ή και απαιτήσεων/αξιώσεων,
- (iii) η τυχαία μεταβλητή S , παριστάνει το ύψος της συνολικής ζημιάς.

Ο ακριβής υπολογισμός της κατανομής της S είναι τις περισσότερες φορές δύσκολος και για αυτό προτιμώνται προσεγγιστικές μέθοδοι. Η χρησιμότητα των προσεγγιστικών μεθόδων έγκειται στο γεγονός ότι μπορούν και προσεγγίζουν αρκούντως ικανοποιητικά την συνάρτηση κατανομής της τυχαίας μεταβλητής S , έστω $F_S(s)$, της οποίας ο ακριβής υπολογισμός είναι τεχνικά εξαιρετικά πολύπλοκη.

Οι μέθοδοι προσέγγισης που θα αναφερθούν στο παρόν κεφάλαιο είναι:

- (1) Η κανονική προσέγγιση (NA)
- (2) Η προσέγγιση με μετατοπισμένη Γάμμα κατανομή (TGA)
- (3) Η Δυναμοκανονική προσέγγιση (NPA)
- (4) Η προσέγγιση Bowers
- (5) Η προσέγγιση Haldane
- (6) Η προσέγγιση Wilson – Hilferty
- (7) Το ανάπτυγμα Edgeworth

Το υλικό που παρουσιάζεται σε αυτό το κεφάλαιο βασίζεται κυρίως στα ακόλουθα άρθρα και βιβλία: Pentikainen (1987), Beard et al. (1984), Bowers (1966), Haldane (1938), Hanspeter Schmidli (2017), Seri & Choirat (2015), Shevchenko, Pavel. (2010), Sundt B. (2004), Teugels J., Wilson & Hilferty (1931).

2.1 Κανονική προσέγγιση (NA)

Η πιο απλή προσέγγιση της συνάρτησης κατανομής $F_S(s)$ προκύπτει με χρήση του Κεντρικού Οριακού Θεωρήματος (NA, Normal approximation). Αφού

$$S = X_1 + X_2 + \dots + X_N$$

μπορούμε να γράψουμε ότι

$$F_S(s) = P(S \leq s) = P\left(\frac{S - \mu_s}{\sigma_s} \leq \frac{s - \mu_s}{\sigma_s}\right) \approx \Phi\left(\frac{s - \mu_s}{\sigma_s}\right)$$

(μ_s και σ_s είναι η μέση τιμή και η τυπική απόκλιση της τυχαίας μεταβλητής S , αντίστοιχα).

Βέβαια η αυστηρή χρήση του Κεντρικού Οριακού Θεωρήματος προϋποθέτει άθροισμα μεγάλου και σταθερού αριθμού τυχαίων μεταβλητών που στη δική μας περίπτωση ο αριθμός αυτός είναι τυχαίος. Γενικά μπορούμε να πούμε ότι η προσέγγιση είναι καλή για μεγάλες τιμές της μέσης τιμής $E(N)$, και στην περίπτωση που η κατανομή των συνολικών αποζημιώσεων S είναι σχεδόν συμμετρική.

Παράδειγμα 2.1. Έστω ότι για την κατανομή των συνολικών αποζημιώσεων S σε ευρώ γνωρίζουμε ότι $\mu_s = 10000$ και $\sigma_s = 1000$. Τότε η πιθανότητα να υπερβούν τις 13000 ευρώ οι συνολικές αποζημιώσεις, σύμφωνα με την προσέγγιση NA είναι

$$P(S > 13000) \approx 1 - \Phi\left(\frac{13000 - 10000}{1000}\right) = 1 - \Phi(3) = 0.00135.$$

2.2 Προσέγγιση με μετατοπισμένη Γάμμα κατανομή (TGA)

Είναι γεγονός ότι στις περισσότερες περιπτώσεις οι συνολικές απαιτήσεις S έχουν συνήθως κατανομή που ταιριάζει με αυτή της κατανομής Γάμμα, δηλαδή παίρνει θετικές τιμές και παρουσιάζει θετική ασυμμετρία $\gamma_s > 0$ (κατανομή λοξή προς τα δεξιά). Μια τέτοια περίπτωση είναι αυτή της σύνθετης κατανομής Poisson. Θυμίζουμε ότι η συνάρτηση πυκνότητας της κατανομής Γάμμα με παραμέτρους a και θ ($S \sim G(a, \theta)$), δίνεται από τον τύπο

$$f_S(s) = \frac{s^{a-1} e^{-s/\theta}}{\theta^a \Gamma(a)}, \quad s, a, \theta > 0.$$

Στην πράξη η κατανομή των συνολικών απαιτήσεων S προσεγγίζεται με μια πιο ευέλικτη κατανομή, την μετατοπισμένη Γάμμα κατανομή (Translated Gamma Distribution).

Αν $S \sim G(a, \theta)$, τότε η τυχαία μεταβλητή $Y = S + k$, $k > 0$, έχει την μετατοπισμένη Γάμμα κατανομή, της οποίας η συνάρτηση πυκνότητας δίνεται από τον τύπο

$$f_Y(y) = \frac{(y-k)^{a-1} e^{-(y-k)/\theta}}{\theta^a \Gamma(a)}, \quad y > k \text{ και } a, \theta > 0$$

(συμβολισμός $Y \sim G(a, \theta, k)$). Είναι εύκολο να διαπιστωθεί ότι η μέση τιμή μ_Y , η διακύμανση σ_Y^2 και ο συντελεστής ασυμμετρίας γ_Y της Y , δίνονται από τους τύπους

$$\mu_Y = a\theta + k, \quad \sigma_Y^2 = a\theta^2, \quad \gamma_Y = \frac{2}{\sqrt{a}}.$$

Έτσι η κατανομή της S προσεγγίζεται από την κατανομή της Y , της οποίας οι παράμετροι προκύπτουν από τη λύση των εξισώσεων

$$\mu_S = a\theta + k, \quad \sigma_S^2 = a\theta^2, \quad \gamma_S = \frac{2}{\sqrt{a}}$$

που δίνουν

$$a = \frac{4}{\gamma_S^2}, \quad \theta = \frac{\gamma_S \sigma_S}{2}, \quad k = \mu_S - \frac{2\sigma_S}{\gamma_S}.$$

Συνεπώς

$$F_S(s) = P(S \leq s) \approx P(Y \leq s).$$

Η παραπάνω μέθοδος προσέγγισης, που είναι γνωστή ως μέθοδος TGA (Translated Gamma Approximation), δίνει καλά αποτελέσματα όταν $\gamma_S > 0$. Αν όμως το γ_S είναι κοντά στο 0 τότε είναι καλύτερο να χρησιμοποιείται η κανονική προσέγγιση (NA).

Παράδειγμα 2.2. Έστω ότι για την κατανομή των συνολικών αποζημιώσεων S σε ευρώ γνωρίζουμε ότι $\mu_S = 10000$, $\sigma_S = 1000$ και $\gamma_S = 1$. Τότε

$$a = \frac{4}{1} = 4, \quad \theta = \frac{1000}{2} = 500, \quad k = 10000 - \frac{2000}{1} = 8000.$$

Η πιθανότητα να υπερβούν τις 13000 ευρώ οι συνολικές αποζημιώσεις, σύμφωνα με την προσέγγιση TGA ($Y \sim G(4, 500, 8000)$), είναι

$$P(S > 13000) \approx 1 - P(Y \leq 13000) = 0.01033605.$$

2.3 Δυναμοκανονική προσέγγιση (NPA)

Αν η κατανομή της τυχαίας μεταβλητής S είναι λοξή τότε η κανονική προσέγγιση προσφέρει φτωχές προσεγγίσεις. Συνηθίζεται σε αυτή την περίπτωση η κανονική προσέγγιση να εφαρμόζεται σε ένα μετασχηματισμό της S που είναι έτσι σχεδιασμένος ώστε να μειώνεται ο βαθμός της λοξότητας.

Η γενική ιδέα (δείτε Beard et al. (1984)) έχει ως εξής: Καταρχήν

$$F_S(s) = P(S \leq s) = P\left(\frac{S - \mu_S}{\sigma_S} \leq \frac{s - \mu_S}{\sigma_S}\right).$$

Βρίσκεται ένας μετασχηματισμός της τ.μ. $\frac{S - \mu_S}{\sigma_S}$, έστω

$$\frac{S - \mu_S}{\sigma_S} = u(Y),$$

τέτοιος ώστε η τ.μ. Y να έχει μέση τιμή 0, διακύμανση 1 και να είναι περίπου συμμετρική. Έτσι σε αυτή την περίπτωση μπορούμε να γράψουμε ότι

$$F_S(s) = P(S \leq s) = P\left(u(Y) \leq \frac{s - \mu_S}{\sigma_S}\right) = F_Y\left(u^{-1}\left(\frac{s - \mu_S}{\sigma_S}\right)\right) \approx \Phi\left(u^{-1}\left(\frac{s - \mu_S}{\sigma_S}\right)\right).$$

Ένας τέτοιος μετασχηματισμός της S (δείτε Beard et al. (1984), σελίδα 110) είναι ο ακόλουθος

$$\frac{S - \mu_S}{\sigma_S} = u(Y) = Y + \frac{\gamma_S}{6}(Y^2 - 1)$$

(μ_S , σ_S και γ_S είναι η μέση τιμή, η τυπική απόκλιση και ο συντελεστής λοξότητας της τυχαίας μεταβλητής S). Θέτοντας

$$u(y) = y + \frac{\gamma_S}{6}(y^2 - 1) = z$$

παίρνουμε ότι

$$y = u^{-1}(z) = -\frac{3}{\gamma_S} + \sqrt{1 + \frac{9}{\gamma_S^2} + \frac{6z}{\gamma_S}}.$$

Επομένως

$$F_S(s) = P(S \leq s) \approx \Phi\left(-\frac{3}{\gamma_S} + \sqrt{1 + \frac{9}{\gamma_S^2} + \frac{6(s - \mu_S)}{\gamma_S \sigma_S}}\right).$$

Η παραπάνω προσέγγιση αναφέρεται στη βιβλιογραφία ως Δυναμοκανονική προσέγγιση (NPA, Normal Power Approximation) και δίνει ικανοποιητικά αποτελέσματα όταν $\gamma_S < 1$ και $s > \mu_S$.

Παράδειγμα 2.3. Έστω ότι για την κατανομή των συνολικών αποζημιώσεων S σε ευρώ γνωρίζουμε ότι $\mu_S = 10000$, $\sigma_S = 1000$ και $\gamma_S = 1$. Τότε η πιθανότητα να υπερβούν τις 13000 ευρώ οι συνολικές αποζημιώσεις, σύμφωνα με την προσέγγιση NPA, είναι

$$\begin{aligned} P(S > 13000) &\approx 1 - \Phi\left(-\frac{3}{1} + \sqrt{1 + \frac{9}{1} + \frac{6(13000 - 10000)}{1000}}\right) \\ &= 1 - \Phi(2.291503) = 0.01096718. \end{aligned}$$

2.4 Η προσέγγιση Bowers

Γνωρίζοντας ότι η κατανομή Γάμμα αποτελεί μια ρεαλιστική προσέγγιση της κατανομής των συνολικών αποζημιώσεων S , ο Bowers (1966) παρουσίασε μια μεθοδολογία προσέγγισης της συνάρτησης κατανομής $F_S(s)$ με χρήση ορθογωνίων πολυωνύμων. Η μεθοδολογία που πρότεινε εφαρμόζεται στην τυχαία μεταβλητή $Y = \beta S$ όπου $\beta = \frac{E(S)}{V(S)}$. Τότε,

$$E(Y) = \beta E(S) = \frac{E^2(S)}{V(S)}$$

και

$$V(Y) = \beta^2 V(S) = \frac{E^2(S)}{V^2(S)} V(S) = \frac{E^2(S)}{V(S)}.$$

Παρατηρούμε ότι η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής Y είναι ίσες. Έστω a αυτή η κοινή τιμή, δηλαδή

$$E(Y) = V(Y) = a.$$

Έστω μ_k η k -οστή κεντρική ροπή της τυχαίας μεταβλητής Y . Ορίζουμε τις ποσότητες

$$A = \frac{\mu_3 - 2a}{3!},$$

$$B = \frac{\mu_4 - 12\mu_3 - 3a^2 + 18a}{4!},$$

$$C = \frac{\mu_5 - 20\mu_4 - (10a - 120)\mu_3 + 60a^2 - 144a}{5!}.$$

Συμβολίζοντας με $F_G(y; a)$ την συνάρτηση κατανομής της κατανομής $G(a, 1)$, δηλαδή

$$F_G(y; a) = \frac{1}{\Gamma(a)} \int_0^y x^{a-1} e^{-x} dx$$

τότε για τη συνάρτηση κατανομής $F_Y(y)$ της τυχαίας μεταβλητής Y , ο Bowers (1966), έδειξε ότι

$$\begin{aligned} F_Y(y) \approx & (1 - A + B - C)F_G(y; a) + (3A - 4B + 5C)F_G(y; a + 1) \\ & + (-3A + 6B - 10C)F_G(y; a + 2) + (A - 4B + 10C)F_G(y; a + 3) \\ & + (B - 5C)F_G(y; a + 4) + (C)F_G(y; a + 4). \end{aligned}$$

Χρησιμοποιώντας τη σχέση

$$F_G(y; a + 1) = F_G(y; a) - \frac{x^a e^{-x}}{\Gamma(a + 1)}$$

προκύπτει ότι

$$F_Y(y) \approx F_G(y; a) - Ae^{-y} \left[\frac{y^a}{\Gamma(a+1)} - \frac{2y^{a+1}}{\Gamma(a+2)} + \frac{y^{a+2}}{\Gamma(a+3)} \right] \\ + Be^{-y} \left[\frac{y^a}{\Gamma(a+1)} - \frac{3y^{a+1}}{\Gamma(a+2)} + \frac{3y^{a+2}}{\Gamma(a+3)} - \frac{y^{a+3}}{\Gamma(a+4)} \right] \\ - Ce^{-y} \left[\frac{y^a}{\Gamma(a+1)} - \frac{4y^{a+1}}{\Gamma(a+2)} + \frac{6y^{a+2}}{\Gamma(a+3)} - \frac{4y^{a+3}}{\Gamma(a+4)} + \frac{y^{a+4}}{\Gamma(a+5)} \right].$$

Συνεπώς η συνάρτηση κατανομής $F_S(s)$ μπορεί να προσεγγιστεί με τη βοήθεια της σχέσης

$$F_S(s) = F_Y(\beta s)$$

η οποία δίνει

$$F_S(s) = P(S \leq s) = P(Y \leq \beta s) = F_Y(\beta s) \\ \approx F_G(\beta s; a) - Ae^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{2(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{(\beta s)^{a+2}}{\Gamma(a+3)} \right] \\ + Be^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{3(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{3(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{(\beta s)^{a+3}}{\Gamma(a+4)} \right] \\ - Ce^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{4(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{6(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{4(\beta s)^{a+3}}{\Gamma(a+4)} \right. \\ \left. + \frac{(\beta s)^{a+4}}{\Gamma(a+5)} \right].$$

Σημειώνουμε ότι για να εφαρμοστεί η συγκεκριμένη μέθοδος θα πρέπει να υπάρχουν οι κεντρικές ροπές της τυχαίας μεταβλητής Y (και επομένως και της τυχαίας μεταβλητής X που δηλώνει το ύψος μιας αποζημίωσης) μέχρι και 5^{ης} τάξης.

Παράδειγμα 2.4. Έστω ότι η κατανομή του αριθμού των απαιτήσεων είναι η κατανομή Poisson με παράμετρο $\lambda = 16$, δηλαδή $N \sim P(16)$, και ότι η κατανομή του ύψους μιας αποζημίωσης είναι η Εκθετική κατανομή με παράμετρο $\theta = 1$, δηλαδή $X \sim E(1)$. Τότε είναι γνωστό ότι

$$\mu'_r = E(X^r) = r!, \quad r \geq 0.$$

$$\mu_x = E(X) = 1 \quad \text{και} \quad \mu_N = E(N) = \lambda = 16$$

Για την κατανομή των συνολικών αποζημιώσεων S έχουμε ότι,

$$\mu_S = E(S) = E(N)E(X) = 16$$

και μπορεί εύκολα να διαπιστωθεί η ισχύς των παρακάτω σχέσεων

$$\mu_2 = E((S - \mu_S)^2) = E((S - 16)^2) = \mu'_2 \lambda = 32,$$

$$\mu_3 = E((S - \mu_S)^3) = E((S - 16)^3) = \mu'_3 \lambda = 96,$$

$$\mu_4 = E((S - \mu_S)^4) = E((S - 16)^4) = \mu'_4 \lambda + 3(\mu'_2 \lambda)^2 = 3456,$$

$$\mu_5 = E((S - \mu_S)^5) = E((S - 16)^5) = \mu'_5 \lambda + 10\mu'_2 \mu'_3 \lambda^2 = 32640.$$

Για την τυχαία μεταβλητή $Y = \beta S$ όπου $\beta = \frac{E(S)}{V(S)}$, έχουμε ότι $\beta = \frac{1}{2}$. Επομένως

$$Y = \frac{S}{2}$$

Έτσι

$$a = E(Y) = V(Y) = 8.$$

και μπορεί εύκολα να διαπιστωθεί ότι

$$\mu_2 = E((Y - \mu_Y)^2) = E((Y - 8)^2) = \left(\frac{1}{2}\right)^2 E((S - 16)^2) = \frac{1}{4} 32 = 8,$$

$$\mu_3 = E((Y - \mu_Y)^3) = E((Y - 8)^3) = \left(\frac{1}{2}\right)^3 E((S - 16)^3) = \frac{1}{8} 96 = 12,$$

$$\mu_4 = E((Y - \mu_Y)^4) = E((Y - 8)^4) = \left(\frac{1}{2}\right)^4 E((S - 16)^4) = \frac{1}{16} 3456 = 216,$$

$$\mu_5 = E((Y - \mu_Y)^5) = E((Y - 8)^5) = \left(\frac{1}{2}\right)^5 E((S - 16)^5) = \frac{1}{32} 32640 = 1020.$$

Χρησιμοποιώντας τις παραπάνω κεντρικές ροπές της Y προκύπτει άμεσα ότι

$$A = -\frac{2}{3}, \quad B = 1, \quad C = -\frac{132}{120}.$$

Έτσι για παράδειγμα

$$F_S(25) = F_Y(12.5) \approx 0.93393522.$$

Στον ακόλουθο πίνακα δίνονται τιμές της συνάρτησης κατανομής $F_S(s)$, ακριβείς και προσεγγιστικές.

Πίνακας 2.1 Προσέγγιση της S κατά Bowers ($N \sim P(16)$ και $X \sim E(1)$).

s	$F_S(s)$: Ακριβής τιμή	$F_S(s)$: Προσεγγιστική τιμή
0	0.00000011	0.00000000
5	0.00873117	0.00726645
10	0.13900952	0.14268808
15	0.46445112	0.46126990
20	0.77389312	0.77322055
25	0.93234932	0.93393522
30	0.98479471	0.98503475
35	0.99730117	0.99696568
40	0.99960660	0.99944040
45	0.99995146	0.99993924
50	0.99999481	1.00001625

Σημειώνουμε ότι η ακριβής τιμή $F_S(s)$ υπολογίστηκε με διακριτοποίηση της Εκθετικής κατανομής και εφαρμογή της μεθόδου “recursive” της συνάρτησης “aggregateDist” του πακέτου “actuar” της R.

2.5 Η προσέγγιση Haldane

Η προσέγγιση του Haldane (1938) (δείτε επίσης Pentikainen (1987)) βασίζεται στη λογική της Δυναμοκανονικής προσέγγισης. Ο μετασχηματισμός της S που χρησιμοποιείται είναι ο

$$Y = \left(\frac{S}{\mu_S}\right)^h,$$

όπου η σταθερά h επιλέγεται με τέτοιο τρόπο έτσι ώστε η τυχαία μεταβλητή Y να έχει συντελεστή ασυμμετρίας περίπου ίσο με 0. Έτσι

$$h = 1 - \frac{\gamma_S \mu_S}{3\sigma_S}$$

και τότε

$$\mu_Y = 1 - \frac{\sigma_S^2}{2\mu_S^2} h(1-h) \left(1 - \frac{\sigma_S^2}{4\mu_S^2} (2-h)(1-3h)\right),$$

$$\sigma_Y^2 = \frac{\sigma_S^2}{\mu_S^2} h^2 \left(1 - \frac{\sigma_S^2}{2\mu_S^2} (1-h)(1-3h)\right).$$

Υποθέτουμε ότι η Y ακολουθεί προσεγγιστικά κανονική κατανομή, οπότε

$$F_S(s) = P(S \leq s) = P\left(Y \leq \left(\frac{S}{\mu_S}\right)^h\right) \approx \Phi\left(\frac{\left(\frac{S}{\mu_S}\right)^h - \mu_Y}{\sigma_Y}\right).$$

Στην ειδική περίπτωση που $h = 0$, χρησιμοποιείται η σχέση

$$F_S(s) = P(S \leq s) \approx \Phi\left(\frac{\log\left(\frac{s}{\mu_S}\right) + \frac{\sigma_S^2}{2\mu_S^2} - \frac{\sigma_S^4}{4\mu_S^4}}{\frac{\sigma_S}{\mu_S} \sqrt{1 - \frac{\sigma_S^2}{2\mu_S^2}}}\right).$$

Παράδειγμα 2.5. Συνεχίζοντας το Παράδειγμα 3.4 ($N \sim P(16)$ και $X \sim E(1)$), έχουμε ότι

$$\mu_S = 16, \quad \sigma_S^2 = 32, \quad \gamma_S = 0.5303301.$$

Έτσι

$$h = 0.5, \quad \mu_Y = 0.9840088, \quad \sigma_Y^2 = 0.03173828.$$

Χρησιμοποιώντας τα παραπάνω είναι εύκολο να προκύψουν τα αποτελέσματα του παρακάτω πίνακα.

Πίνακας 2.2 Προσέγγιση της S κατά Haldane ($N \sim P(16)$ και $X \sim E(1)$).

s	$F_S(s)$: Ακριβής τιμή	$F_S(s)$: Προσεγγιστική τιμή
0	0.00000011	0.00000002
5	0.00873117	0.00852676
10	0.13900952	0.13878190
15	0.46445112	0.46474753
20	0.77389312	0.77406659
25	0.93234932	0.93228850
30	0.98479471	0.98471947
35	0.99730117	0.99727017
40	0.99960660	0.99959855
45	0.99995146	0.99994991
50	0.99999481	0.99999457

Υπάρχει στη βιβλιογραφία και μια παραλλαγή της προσέγγισης Haldane, γνωστή ο μετασχηματισμός B του Haldane (δείτε Pentikainen (1987) και Seri & Choirat (2015)) στην οποία δίνεται έμφαση στην ελαχιστοποίηση της ασυμμετρίας αλλά και της κύρτωσης. Ειδικότερα ορίζεται η τυχαία μεταβλητή

$$S' = S + g - \mu_S$$

και ο μετασχηματισμός

$$Y = \left(\frac{S'}{\mu_{S'}}\right)^h = \left(1 + \frac{S - \mu_S}{g}\right)^h,$$

Προτείνονται οι τιμές

$$g = \frac{12\mu_2^2\mu_3}{20\mu_3^2 - 9\mu_2(\mu_4 - 3\mu_2^2)}, \quad h = \frac{16\mu_3^2 - 9\mu_2(\mu_4 - 3\mu_2^2)}{20\mu_3^2 - 9\mu_2(\mu_4 - 3\mu_2^2)}$$

όπου μ_k η k -οστή κεντρική ροπή της S , δηλαδή $\mu_k = E((S - \mu_S)^k)$. Η προσέγγιση της $F_S(s)$ δίνεται από τον τύπο

$$F_S(s) = P(S \leq s) \approx \Phi \left(\frac{\left(1 + b \left(\frac{s - \mu_S}{\sigma_S}\right)\right)^h - \mu_Y}{\sigma_Y} \right),$$

όπου

$$\gamma = \frac{\mu_4}{\mu_2^2} - 3, \quad b = \frac{5}{3}\gamma_S - \frac{3}{4} \cdot \frac{\gamma}{\gamma_S}, \quad c = \frac{4}{3}\gamma_S - \frac{3}{4} \cdot \frac{\gamma}{\gamma_S}, \quad h = \frac{c}{b},$$

$$m_Y = 1 - \frac{1}{2}c(b - c) \left(1 + \frac{1}{4}(2b - c)(3c - b)\right)$$

$$\sigma_Y = |c| \sqrt{1 + \frac{1}{2}(b - c)(3c - b)}.$$

2.6 Η προσέγγιση Wilson - Hilferty

Οι Wilson και Hilferty (δείτε Wilson & Hilferty (1931)) χρησιμοποίησαν έναν παρόμοιο μετασχηματισμό με τον μετασχηματισμό του Haldane όπου η τιμή της σταθεράς h είναι ίση με $\frac{1}{3}$. Τα αποτελέσματα μπορούν να παρουσιαστούν όπως στην περίπτωση της Δυναμοκανονικής προσέγγισης. Εδώ έχουμε ότι

$$\frac{S - \mu_S}{\sigma_S} = u(Y) = \left(\frac{Y - c_1}{c_2}\right)^{1/3} - c_3,$$

όπου

$$c_1 = \frac{\gamma_S}{6} - \frac{6}{\gamma_S}, \quad c_2 = 3 \left(\frac{2}{\gamma_S}\right)^{2/3}, \quad c_3 = \frac{2}{\gamma_S}.$$

Θέτοντας

$$u(y) = \left(\frac{y - c_1}{c_2}\right)^{1/3} - c_3 = z$$

παίρνουμε ότι

$$y = u^{-1}(z) = c_1 + c_2(z + c_3)^{1/3}.$$

Επομένως

$$F_S(s) = P(S \leq s) \approx \Phi \left(c_1 + c_2 \left(\frac{s - \mu_S}{\sigma_S} + c_3 \right)^{\frac{1}{3}} \right).$$

Παράδειγμα 2.6. Συνεχίζοντας το Παράδειγμα 2.4 ($N \sim P(16)$ και $X \sim E(1)$), έχουμε ότι

$$\mu_S = 16, \quad \sigma_S^2 = 32, \quad c_1 = -11.22532, \quad c_2 = 7.268482, \quad c_3 = 3.771236.$$

Χρησιμοποιώντας τα παραπάνω είναι εύκολο να προκύψουν τα αποτελέσματα του παρακάτω πίνακα.

Πίνακας 2.3 Προσέγγιση της S κατά Wilson-Hilferty ($N \sim P(16)$ και $X \sim E(1)$).

s	$F_S(s)$: Ακριβής τιμή	$F_S(s)$: Προσεγγιστική τιμή
0	0.00000011	0.00002083
5	0.00873117	0.00963790
10	0.13900952	0.13765176
15	0.46445112	0.46365701
20	0.77389312	0.77499494
25	0.93234932	0.93277467
30	0.98479471	0.98463535
35	0.99730117	0.99714799
40	0.99960660	0.99954999
45	0.99995146	0.99993754
50	0.99999481	0.99999217

2.7 Ανάπτυγμα Edgeworth

Έστω η τυχαία μεταβλητή $Z = \frac{S - \mu_S}{\sigma_S}$. Η ανάπτυξη σε σειρά Taylor της συνάρτησης $\log M_Z(r) = \log E(e^{rZ})$ γύρω από την τιμή $r = 0$ δίνεται από τον τύπο

$$\log M_Z(r) = a_0 + a_1 r + a_2 \frac{r^2}{2} + a_3 \frac{r^3}{6} + a_4 \frac{r^4}{24} + \dots$$

όπου

$$a_k = \left. \frac{d^k \log M_Z(r)}{dr^k} \right|_{r=0}, \quad k = 0, 1, 2, \dots$$

Από θεωρία ροπογεννητριών συναρτήσεων παίρνουμε ότι

- $a_0 = \log M_Z(0) = 0$
- $a_1 = E[Z] = 0$
- $a_2 = \text{Var}[Z] = 1$
- $a_3 = E[(Z - E[Z])^3] = \gamma_Z$
- $a_4 = E[Z^4] - 4E[Z^3]E[Z] - 3E[Z^2]^2 + 12E[Z^2]E[Z]^2 - 6E[Z]^4$
 $= E[Z^4] - 3 = \frac{E[(S-E[S])^4]}{\text{Var}[S]^2} - 3$

Αποκόπτουμε τη σειρά Taylor μετά τον όρο που περιλαμβάνει το r^4 . Τότε, η ροπογεννήτρια συνάρτηση της Z μπορεί να γραφτεί ως

$$M_Z(r) = \exp(\log M_Z(r)) \approx e^{r^2/2 + a_3 r^3/6 + a_4 r^4/24}$$

$$\approx e^{r^2/2} \left(1 + a_3 \frac{r^3}{6} + a_4 \frac{r^4}{24} + a_3^2 \frac{r^6}{72} \right).$$

Σημειώνουμε ότι ο όρος $\exp\{r^2/2\}$ αντιστοιχεί στη ροπογεννήτρια συνάρτηση της τυπικής κανονικής κατανομής. Ξεκινώντας από τη σχέση

$$\int_{-\infty}^{\infty} e^{rx} \Phi'(x) dx = \exp\{r^2/2\}$$

παίρνουμε ότι

$$r^n e^{r^2/2} = \int_{-\infty}^{\infty} (e^{rx})^{(n)} \Phi'(x) dx = (-1)^n \int_{-\infty}^{\infty} e^{rx} \Phi^{(n+1)}(x) dx.$$

Χρησιμοποιώντας την παραπάνω σχέση για $n = 3, 4, 6$ μπορεί να δειχθεί ότι

$$P[Z \leq z] \approx \Phi(z) - \frac{a_3}{6} \Phi^{(3)}(z) + \frac{a_4}{24} \Phi^{(4)}(z) + \frac{a_3^2}{72} \Phi^{(6)}(z).$$

Για την εφαρμογή της παραπάνω σημειώνουμε ότι

$$\Phi^{(3)}(z) = \frac{1}{\sqrt{2\pi}} (z^2 - 1) e^{-z^2/2},$$

$$\Phi^{(4)}(z) = \frac{1}{\sqrt{2\pi}} (-z^3 + 3z) e^{-z^2/2},$$

και

$$\Phi^{(6)}(z) = \frac{1}{\sqrt{2\pi}} (-z^5 + 10z^3 - 15z) e^{-z^2/2}.$$

Η παραπάνω προσέγγιση είναι γνωστή ως προσέγγιση Edgeworth.

Στην περίπτωση της σύνθετης κατανομής Poisson έχουμε ότι

$$a_k = \lambda \mu'_k (\lambda \mu'_2)^{-k/2}, \quad k \geq 0$$

$$\text{όπου } \mu'_k = E(X^k), \quad k \geq 0.$$

Επομένως

$$\begin{aligned} P(S \leq x) &= P\left(Z \leq \frac{x - \lambda \mu_X}{\sqrt{\lambda \mu'_2}}\right) \\ &\approx \Phi\left(\frac{x - \lambda \mu_X}{\sqrt{\lambda \mu'_2}}\right) - \frac{\mu'_3}{6\sqrt{\lambda}(\mu'_2)^{3/2}} \Phi^{(3)}\left(\frac{x - \lambda \mu_X}{\sqrt{\lambda \mu'_2}}\right) + \frac{\mu'_4}{24\lambda(\mu'_2)^2} \Phi^{(4)}\left(\frac{x - \lambda \mu_X}{\sqrt{\lambda \mu'_2}}\right) \\ &\quad + \frac{(\mu'_3)^2}{72\lambda(\mu'_2)^3} \Phi^{(6)}\left(\frac{x - \lambda \mu_X}{\sqrt{\lambda \mu'_2}}\right). \end{aligned}$$

Παράδειγμα 2.7. Συνεχίζοντας το Παράδειγμα 2.4 ($N \sim P(16)$ και $X \sim E(1)$), είναι προκύπτουν τα ακόλουθα αποτελέσματα του παρακάτω πίνακα.

Πίνακας 2.4 Προσέγγιση της S κατά Edgeworth ($N \sim P(16)$ και $X \sim E(1)$).

s	$F_S(s)$: Ακριβής τιμή	$F_S(s)$: Προσεγγιστική τιμή
0	0.00000011	0.000000000
5	0.00873117	0.008708811
10	0.13900952	0.132079062
15	0.46445112	0.462565583
20	0.77389312	0.783218687
25	0.93234932	0.936996616
30	0.98479471	0.984321571
35	0.99730117	0.997277609
40	0.99960660	0.999778415
45	0.99995146	0.999992692
50	0.99999481	0.999999904

Κεφάλαιο 3: Εφαρμογές και σύγκριση των μεθόδων προσέγγισης

Στην παρούσα παράγραφο θα παρουσιαστούν αναλυτικές εφαρμογές των μεθοδολογιών που παρουσιάστηκαν στο Κεφάλαιο 2.

3.1 1^η Εφαρμογή

Σε αυτή την εφαρμογή θα παρουσιαστούν οι προσεγγιστικές μέθοδοι του Κεφαλαίου 2 στην περίπτωση που η τυχαία μεταβλητή X που παριστάνει το ύψος της ατομικής ζημιάς έχει την κατανομή Pareto ($Pa(\alpha, \beta)$) και η τυχαία μεταβλητή N που παριστάνει το πλήθος των ζημιών ακολουθεί την κατανομή Poisson ($P(\lambda)$).

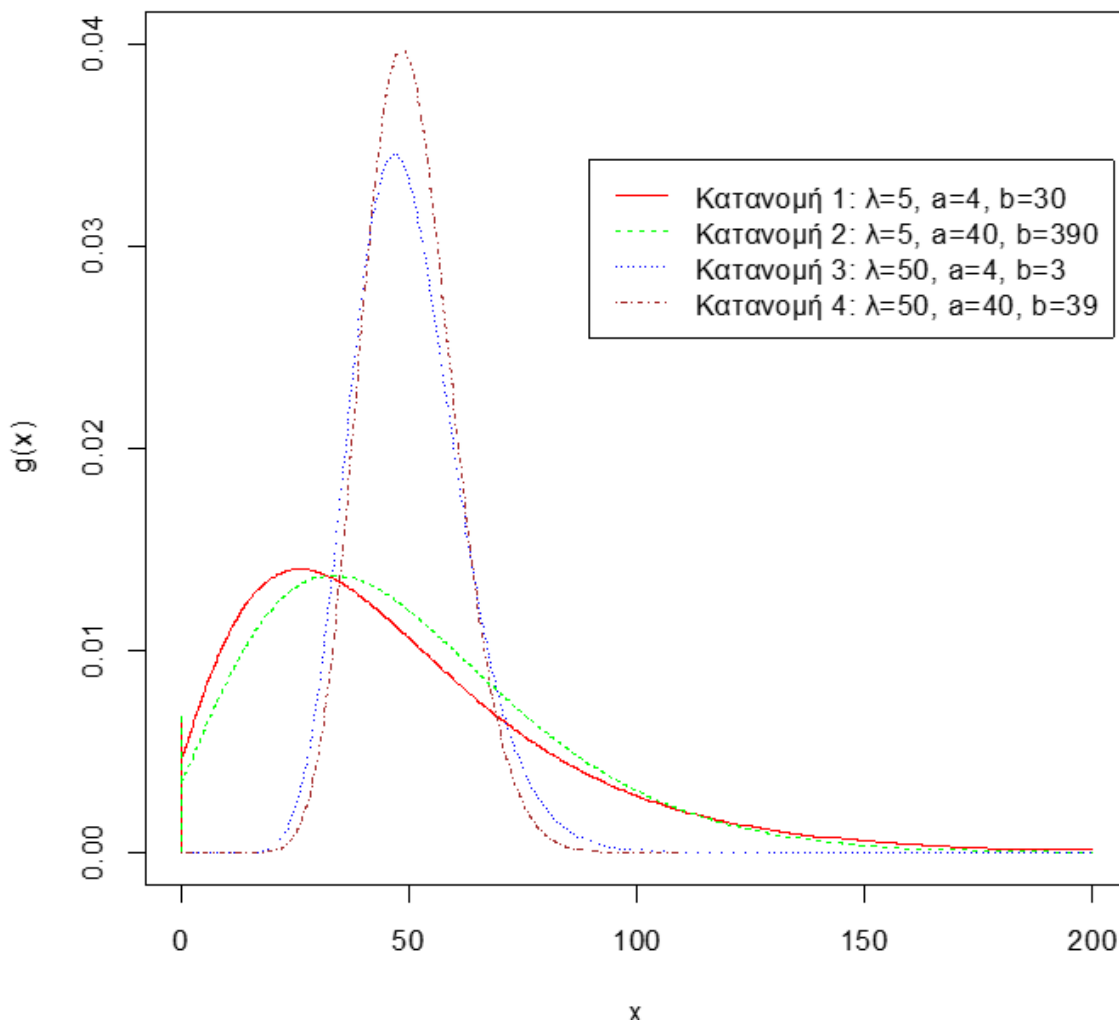
Θα εξεταστούν οι κάτωθι τέσσερις περιπτώσεις για τη σύνθετη κατανομή S :

- Κατανομή 1: $\lambda = 5$, $a = 4$, $\beta = 30$, $\mu_S = 50$, $\sigma_S^2 = 1500$, $\gamma_S = 2.32739$
- Κατανομή 2: $\lambda = 5$, $a = 40$, $\beta = 390$, $\mu_S = 50$, $\sigma_S^2 = 1026.32$, $\gamma_S = 0.98706$
- Κατανομή 3: $\lambda = 50$, $a = 4$, $\beta = 3$, $\mu_S = 50$, $\sigma_S^2 = 150$, $\gamma_S = 0.734847$
- Κατανομή 4: $\lambda = 50$, $a = 40$, $\beta = 39$, $\mu_S = 50$, $\sigma_S^2 = 102.63$, $\gamma_S = 0.31214$

Οι πίνακες και τα γραφήματα που ακολουθούν υλοποιήθηκαν όλα στη γλώσσα προγραμματισμού R . Στο παράρτημα δίνονται οι σχετικοί κώδικες.

Οι συναρτήσεις πυκνότητας της S που παρατίθενται στο επόμενο γράφημα, υπολογίστηκαν χρησιμοποιώντας τον αναδρομικό τύπο του Panjer αφού προηγουμένως έγινε διακριτοποίηση της κατανομής Pareto.

Σύνθετη Poisson-Pareto Κατανομή



Σχήμα 3.1. Συναρτήσεις πυκνότητας της σύνθετης Poisson-Pareto κατανομής

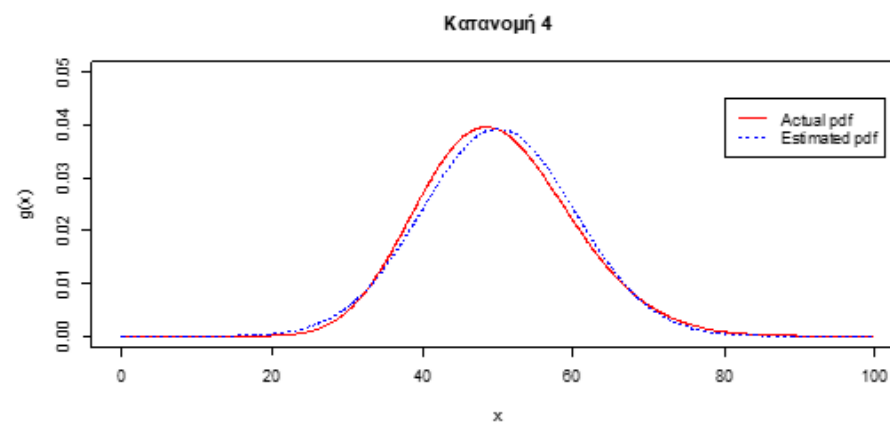
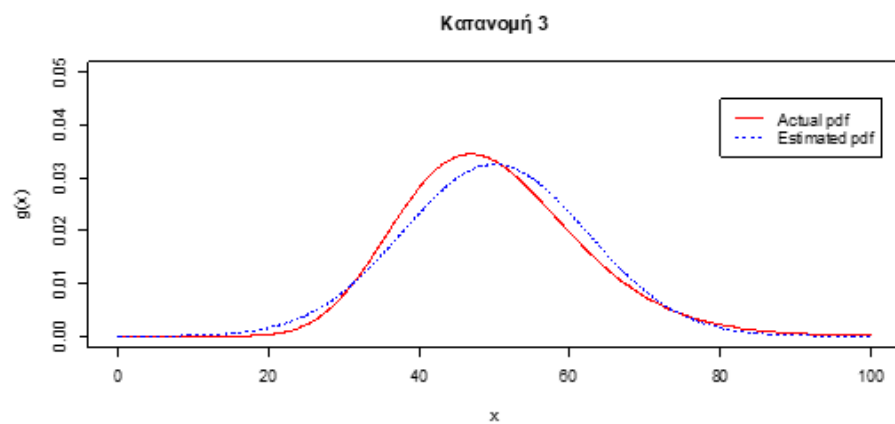
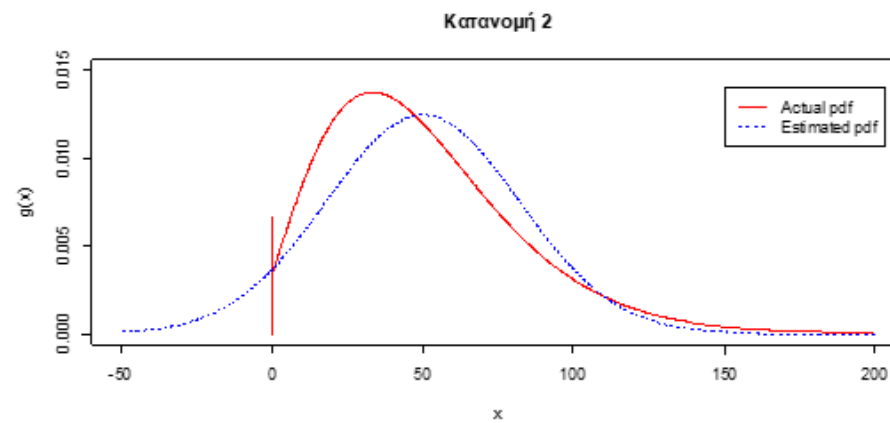
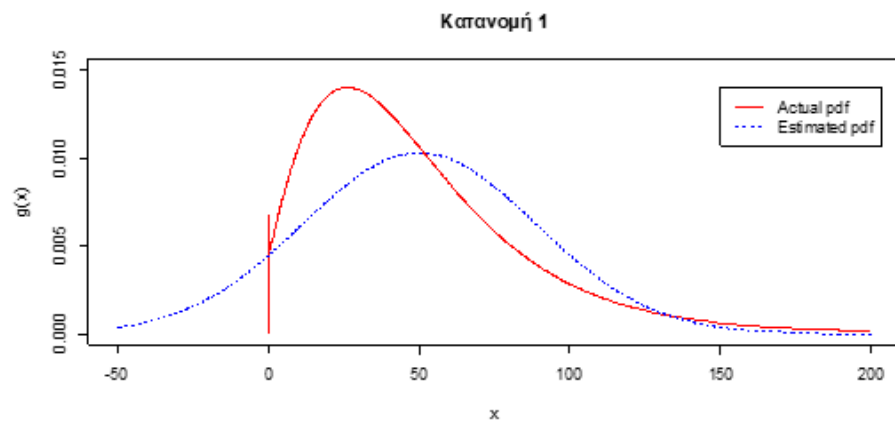
Μπορούμε να παρατηρήσουμε ότι στις περιπτώσεις 1 και 2 υπάρχει υπέρχει πιθανότητα ίση με $0.00674 = e^{-5}$ για μηδενικές αποζημιώσεις (μηδενικό πλήθος ζημιών). Επιπρόσθετα, μπορούμε να παρατηρήσουμε ότι η αλλαγή στην παράμετρο λ της κατανομής του πλήθους των απαιτήσεων N από 5 σε 50 έχει σημαντικότερη επιροή στο σχήμα της κατανομής των συνολικών απαιτήσεων από ότι έχει η αλλαγή της κατανομής του ύψους των απαιτήσεων.

Στη συνέχεια για τις τέσσερις περιπτώσεις προσεγγίζουμε την κατανομή των συνολικών απαιτήσεων μέσω της **κανονικής κατανομής**. Ο λόγος προσέγγισης από την κανονική κατανομή γίνεται σύμφωνα με το Κεντρικό Οριακό Θεώρημα όπου η τυχαία μεταβλητή S αν κανονικοποιηθεί θα τείνει προς την τυπική κανονική ως άθροισμα ανεξάρτητων τυχαίων μεταβλητών.

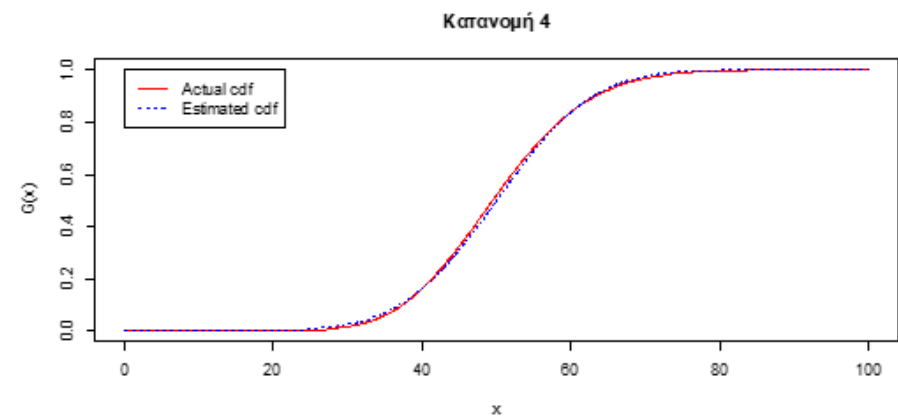
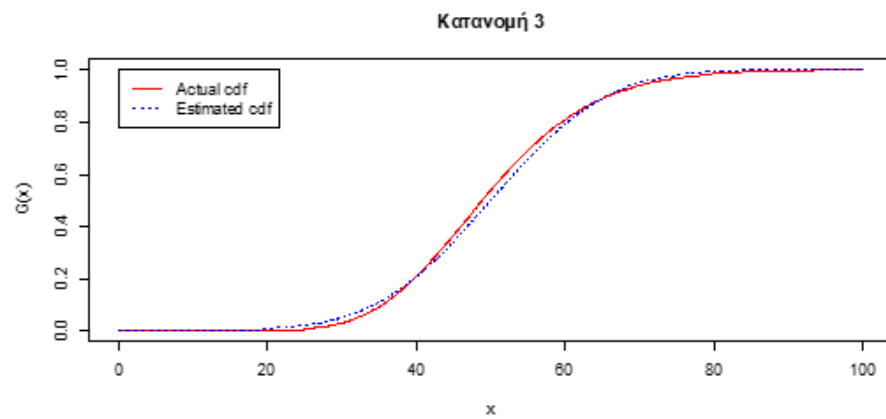
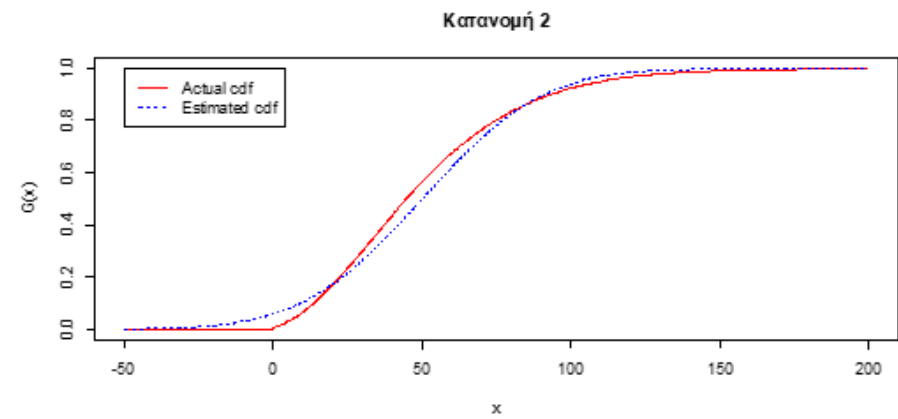
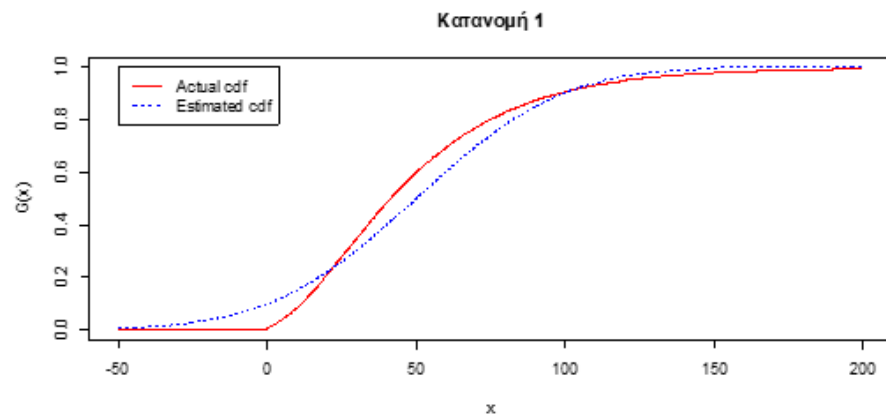
Ωστόσο για να ισχύει αυτό θα πρέπει η μέση τιμή της τυχαίας μεταβλητής N να είναι επαρκώς μεγάλη. Για τον λόγο αυτό θα δούμε ότι η προσέγγιση στις Περιπτώσεις 1 και 2 δεν είναι ικανοποιητική, καθώς $E(N) = 5$ η οποία δεν είναι αρκετά υψηλή. Τα αποτελέσματα για την προσέγγιση της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ φαίνονται στον παρακάτω πίνακα:

Πίνακας 3.1. Κανονική προσέγγιση της πιθανότητας $P(S > \mu_S + 4\sigma_S)$

	$\mu_S + 4\sigma_S$	Πραγματική πιθανότητα	Προσεγγιστική πιθανότητα
Κατανομή 1	204.91933	0.00548864	3.167e-05
Κατανομή 2	178.14466	0.00208920	3.167e-05
Κατανομή 3	98.98979	0.00156217	3.167e-05
Κατανομή 4	90.52290	0.00029018	3.167e-05



Σχήμα 3.2. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω κανονικής κατανομής



Σχήμα 3.3. Πραγματική και προσέγγιση της συνάρτησης κατανομής της S μέσω κανονικής κατανομής

Διαπιστώνεται και γραφικά, όπως ήταν αναμενόμενο, ότι στα Παραδείγματα 3 και 4 η προσαρμογή της κανονικής κατανομής είναι πολύ ικανοποιητική τόσο στη συνάρτηση πυκνότητας όσο και στη συνάρτηση κατανομής ($E(N)=50$).

Αν τώρα προσεγγίσουμε την κατανομή της τυχαίας μεταβλητής S μέσω μετατοπισμένης Γάμμα κατανομής $G(a, \theta, k)$ θα έχουμε τις εξής περιπτώσεις για τις παραμέτρους.

Πίνακας 3.2. Παράμετροι της μετατοπισμένης Γάμμα κατανομής

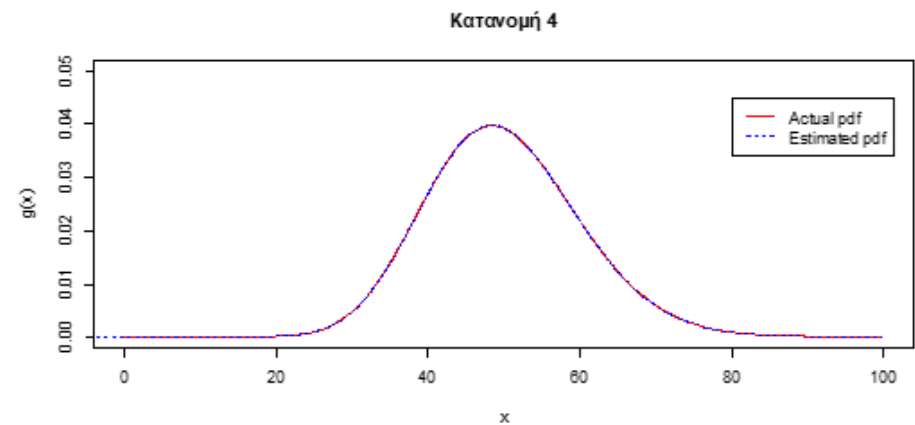
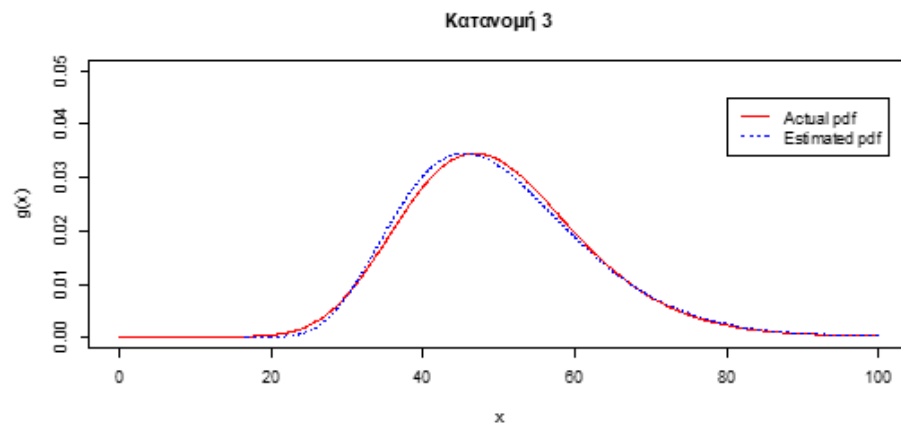
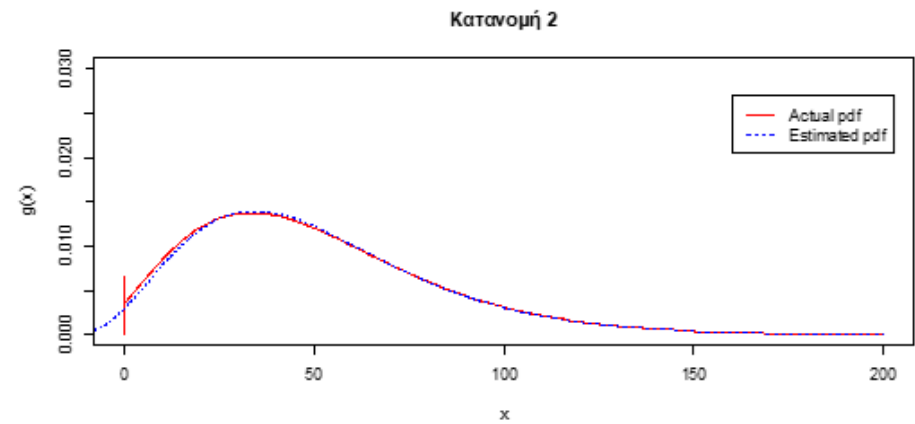
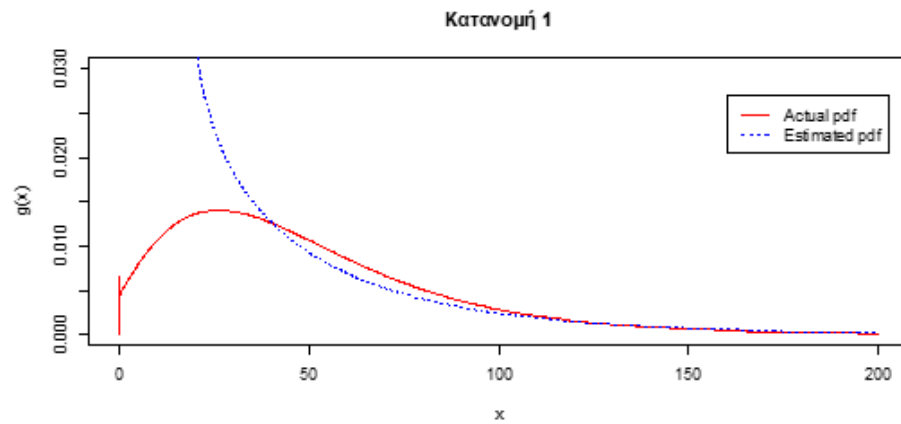
	a	θ	k
Κατανομή 1	0.7407407	45.000000	16.66667
Κατανομή 2	4.1055630	15.810811	-14.91228
Κατανομή 3	7.4074074	4.500000	16.66667
Κατανομή 4	41.0556305	1.581081	-14.91228

Έτσι μπορεί εύκολα να διαπιστωθούν για τα τέσσερα παραδείγματα οι κάτωθι προσεγγιστικές τιμές της πιθανότητας $P(S > \mu_S + 4\sigma_S)$.

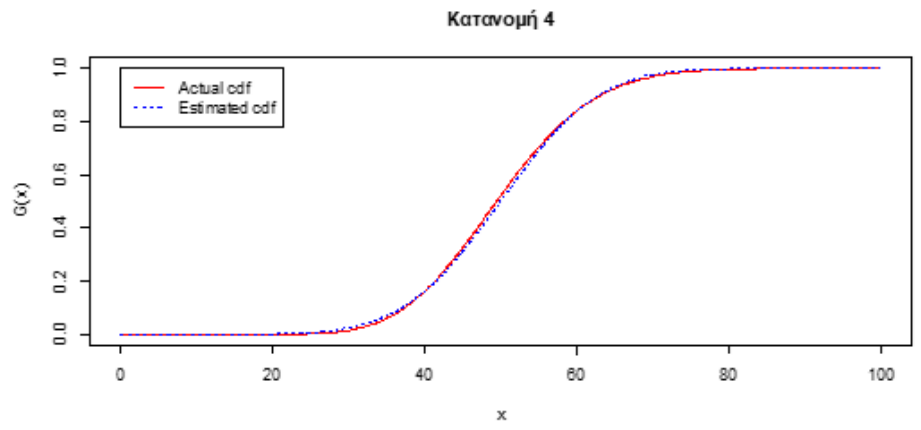
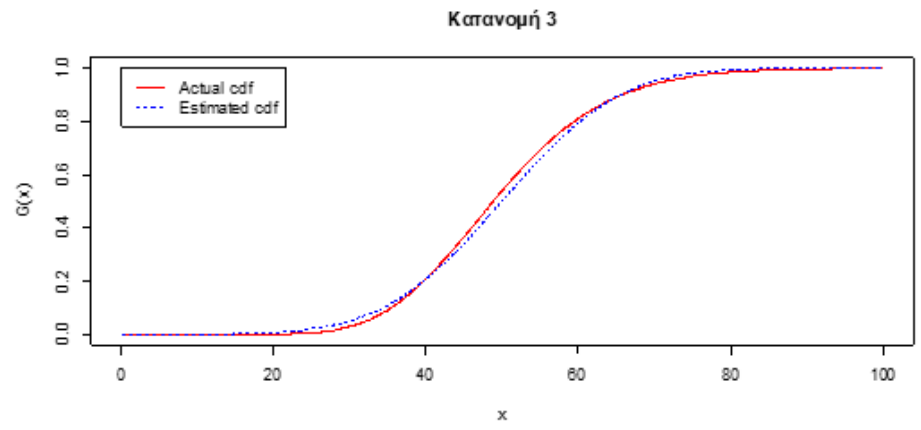
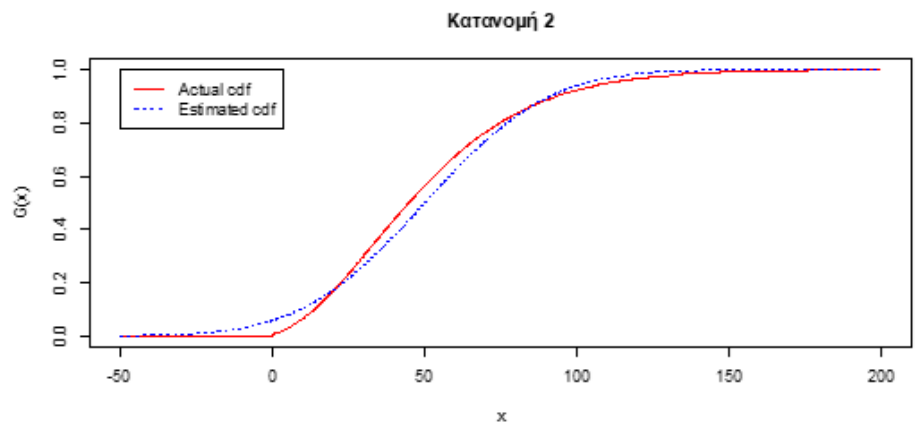
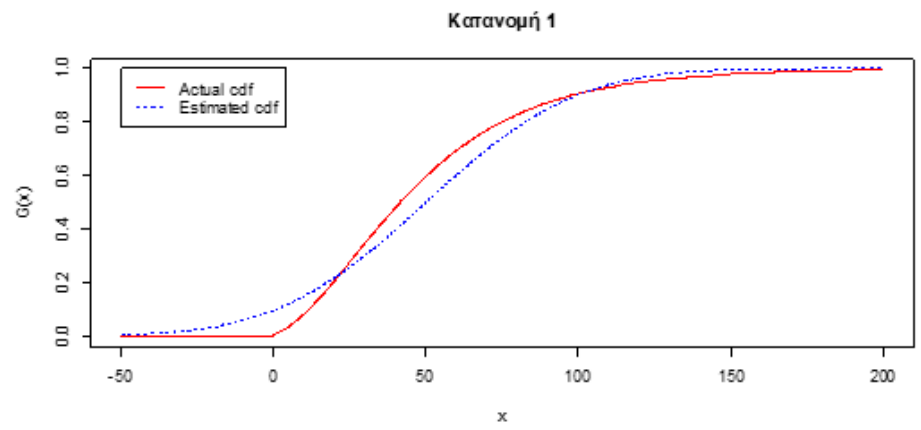
Πίνακας 3.3. Προσέγγιση της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ μέσω μετατοπισμένης Γάμμα κατανομής

	$\mu_S + 4\sigma_S$	Πραγματική πιθανότητα	Προσεγγιστική πιθανότητα
Κατανομή 1	204.91933	0.00548864	0.00807962
Κατανομή 2	178.14466	0.00208920	0.00223998
Κατανομή 3	98.98979	0.00156217	0.00131965
Κατανομή 4	90.52290	0.00029018	0.00030099

Η προσαρμογή της μετατοπισμένης Γάμμα κατανομής δεν είναι καλή για το πρώτο παράδειγμα, ωστόσο είναι ικανοποιητική για τα υπόλοιπα παραδείγματα. Αλλά ακόμη και στο Παράδειγμα 1 η δεξιά ουρά της κατανομής όπως θα δούμε και γραφικά προσεγγίζεται καλά από την μετατοπισμένης Γάμμα κατανομή η οποία είναι καλύτερη από την αντίστοιχη προσέγγιση της κανονικής.



Σχήμα 3.4. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω μετατοπισμένης Γάμμα κατανομής

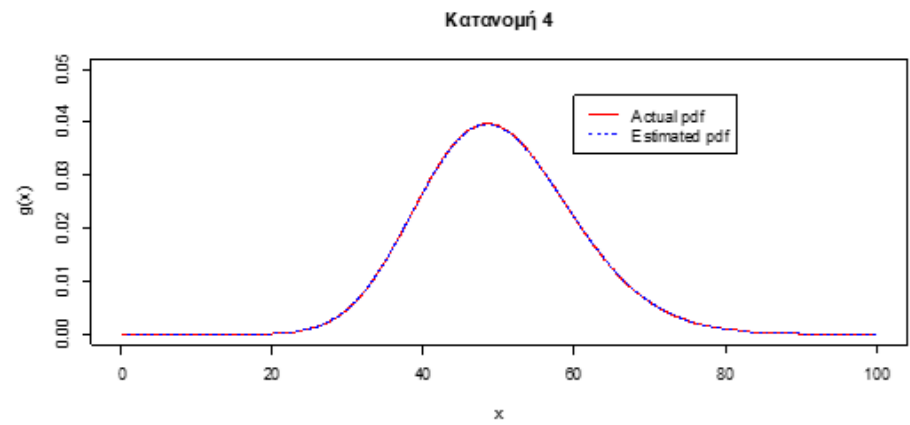
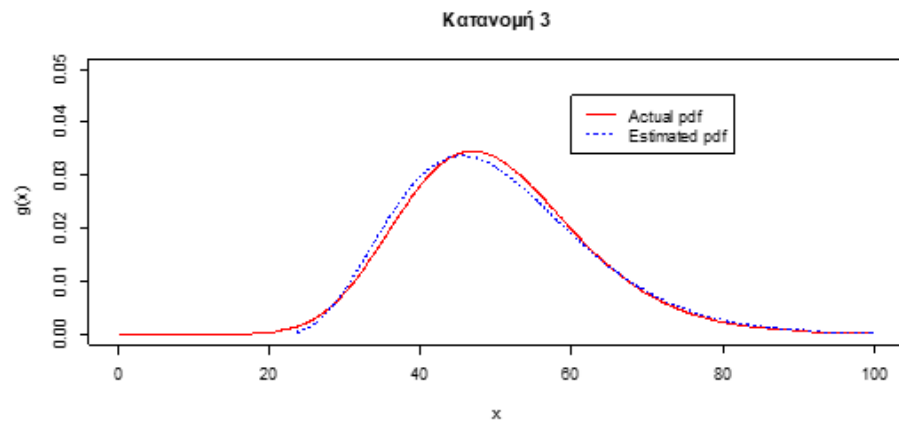
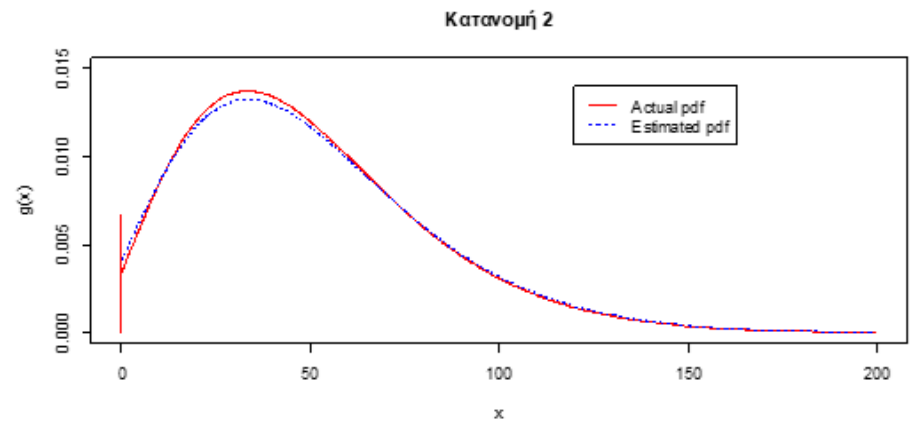
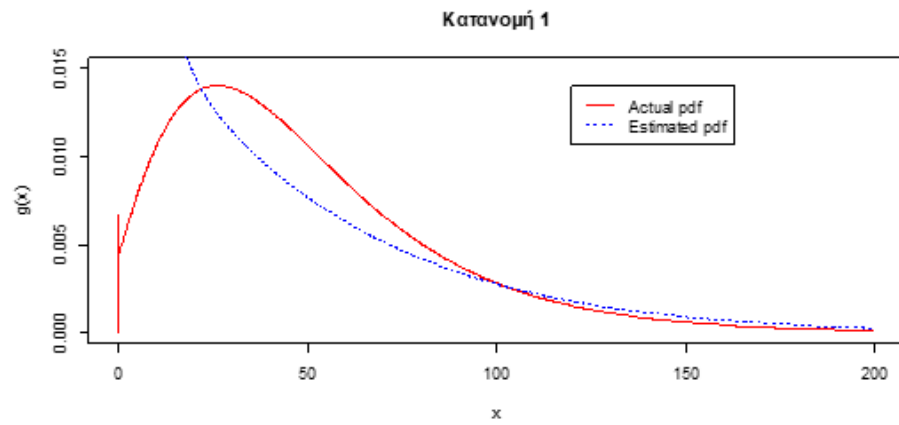


Σχήμα 3.5. Πραγματική και προσέγγιση της συνάρτησης κατανομής της S μέσω μετατοπισμένης Γάμμα κατανομής

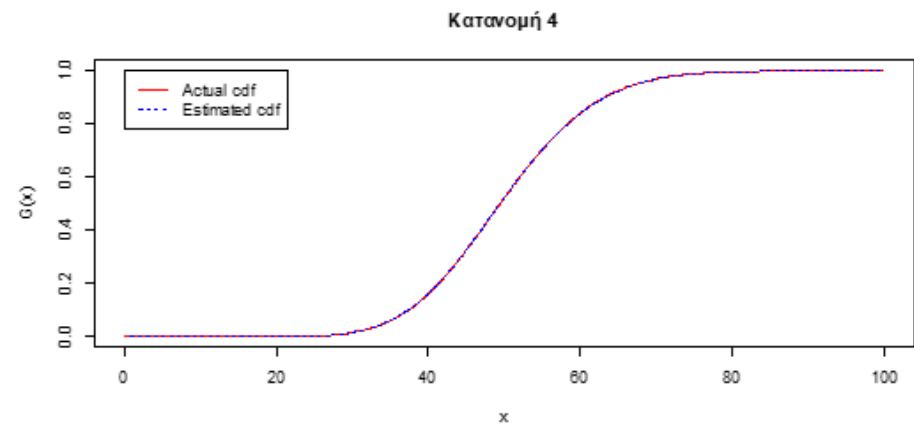
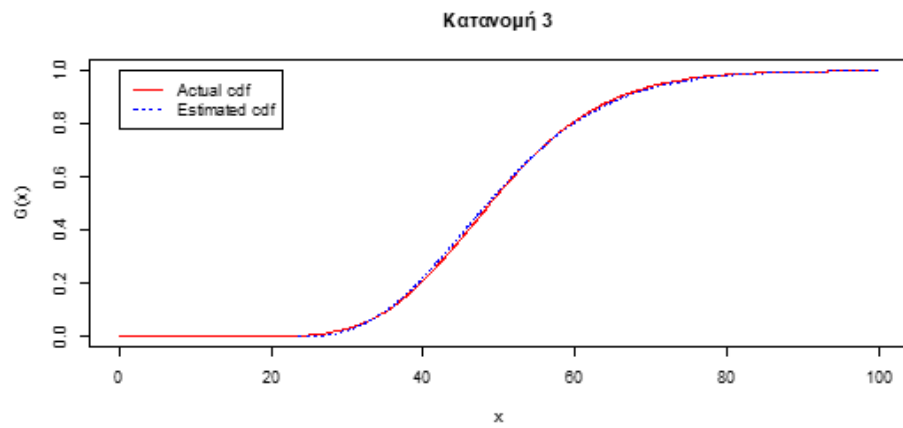
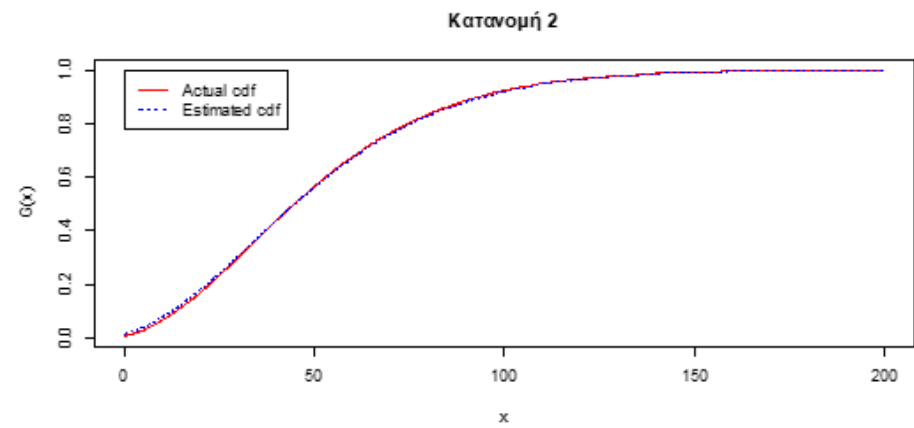
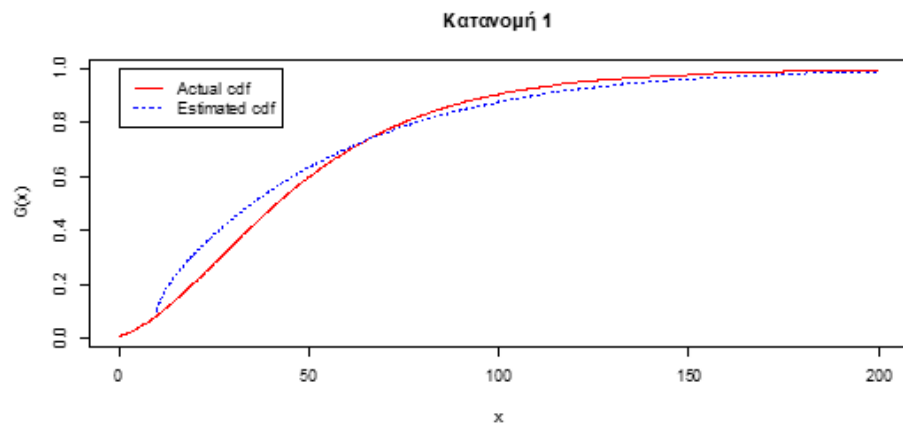
Εφαρμόζοντας τώρα τη μέθοδο της **Δυναμοκανονικής προσέγγισης**, η οποία κατά κύριο λόγο είναι σχεδιασμένη κατά τέτοιο τρόπο ώστε να λειτουργεί καλά για μικρό συντελεστή ασυμμετρίας, καταφέρνουμε να προσεγγίσουμε αρκετά καλύτερα την κατανομή της τυχαίας μεταβλητής S σε σύγκριση με την προσέγγιση της κανονικής κατανομής. Τα αποτελέσματα για την προσέγγιση της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ φαίνονται στον παρακάτω πίνακα.

Πίνακας 3.4. Υπολογισμός της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ μέσω Δυναμοκανονικής προσέγγισης

	$\mu_S + 4\sigma_S$	Πραγματική πιθανότητα	Προσεγγιστική πιθανότητα
Κατανομή 1	204.91933	0.00548864	0.01033905
Κατανομή 2	178.14466	0.00208920	0.00226433
Κατανομή 3	98.98979	0.00156217	0.00129922
Κατανομή 4	90.52290	0.00029018	0.00029371



Σχήμα 3.6. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω Δυναμοκανονικής προσέγγισης



Σχήμα 3.7. Πραγματική και προσέγγιση της συνάρτησης κατανομής της S μέσω Δυναμοκανονικής προσέγγισης

Για να εφαρμόσουμε τη **μέθοδο Haldane** στην περίπτωση της σύνθετης Poisson-Pareto κατανομής βρίσκουμε αρχικά ότι

$$h = 1 - \frac{\gamma_S \mu_S}{\sigma_S} = \frac{a - 4}{2(a - 3)}$$

και προσεγγίζουμε τη συνάρτηση κατανομή της τυχαίας μεταβλητής S μέσω της σχέσης

$$F_S(s) = P(S \leq s) \approx \Phi \left(\frac{\left(\frac{s}{\mu_S}\right)^h - \mu_Y}{\sigma_Y} \right).$$

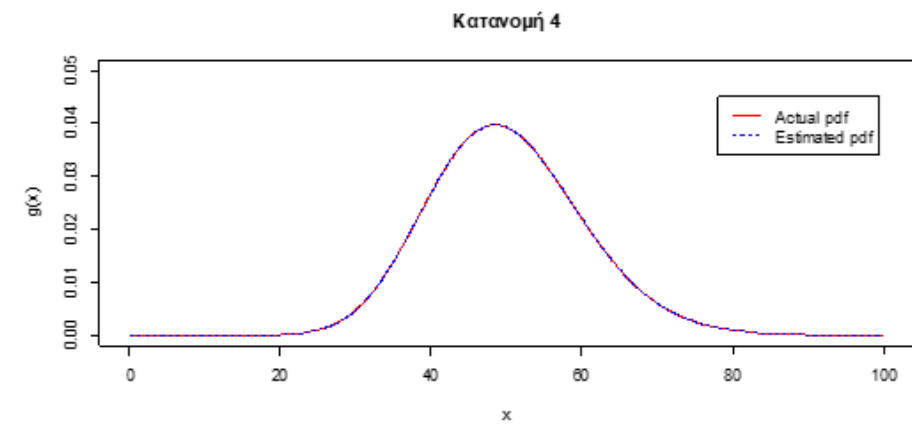
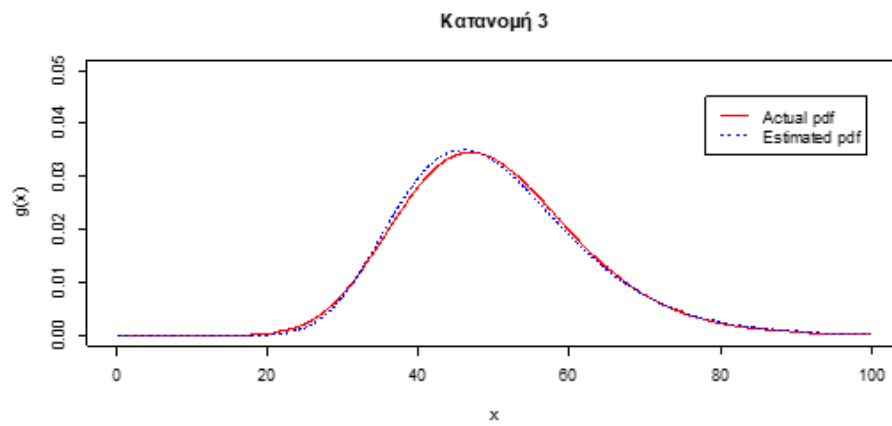
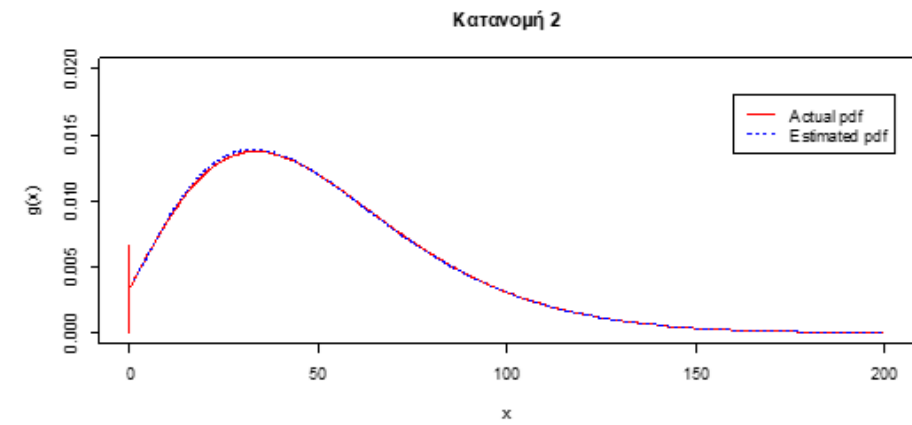
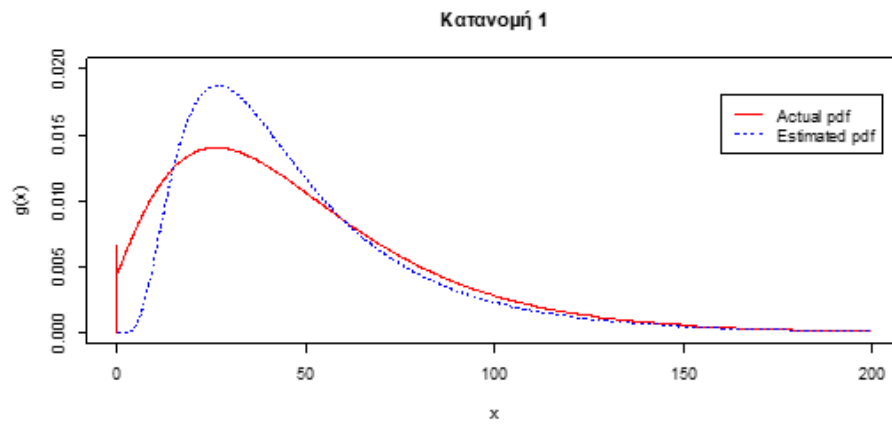
Ειδικότερα για τα Παραδείγματα 1 και 3, επειδή έχουμε $a = 4$, προσεγγίζουμε τη συνάρτηση κατανομή της τυχαίας μεταβλητής S μέσω της σχέσης

$$F_S(s) = P(S \leq s) \approx \Phi \left(\frac{\log \left(\frac{s}{\mu_S} \right) + \frac{\sigma_S^2}{2\mu_S^2} - \frac{\sigma_S^4}{4\mu_S^4}}{\frac{\sigma_S}{\mu_S} \sqrt{1 - \frac{\sigma_S^2}{2\mu_S^2}}} \right).$$

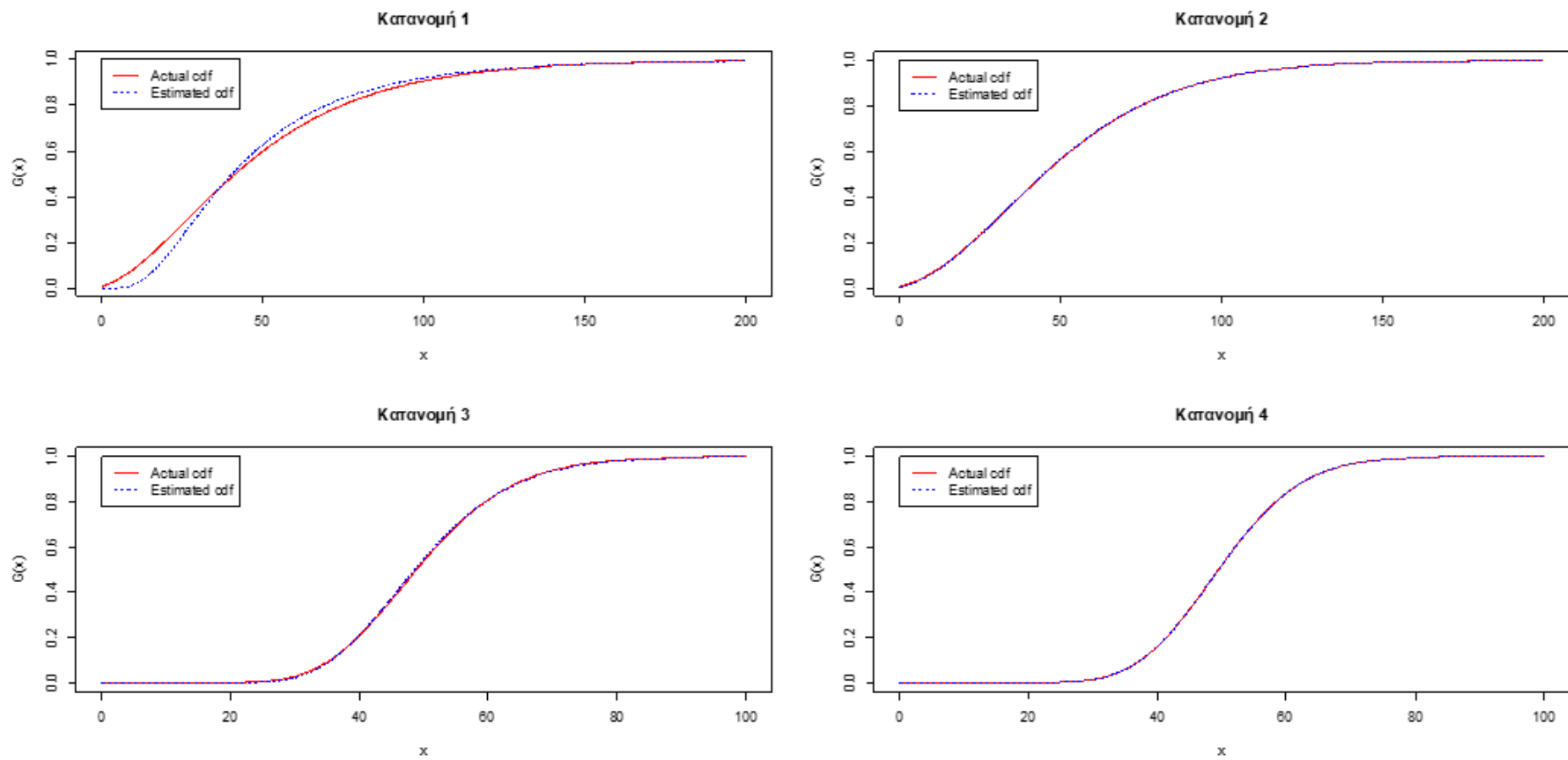
Τα αποτελέσματα για την προσέγγιση κατά Haldane της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ φαίνονται στον παρακάτω πίνακα.

Πίνακας 3.5. Υπολογισμός της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ μέσω προσέγγισης Haldane

	$\mu_S + 4\sigma_S$	Πραγματική πιθανότητα	Προσεγγιστική πιθανότητα
Κατανομή 1	204.91933	0.00548864	0.00619863
Κατανομή 2	178.14466	0.00208920	0.00216971
Κατανομή 3	98.98979	0.00156217	0.00158003
Κατανομή 4	90.52290	0.00029018	0.00029032



Σχήμα 3.8. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω προσέγγισης Haldane



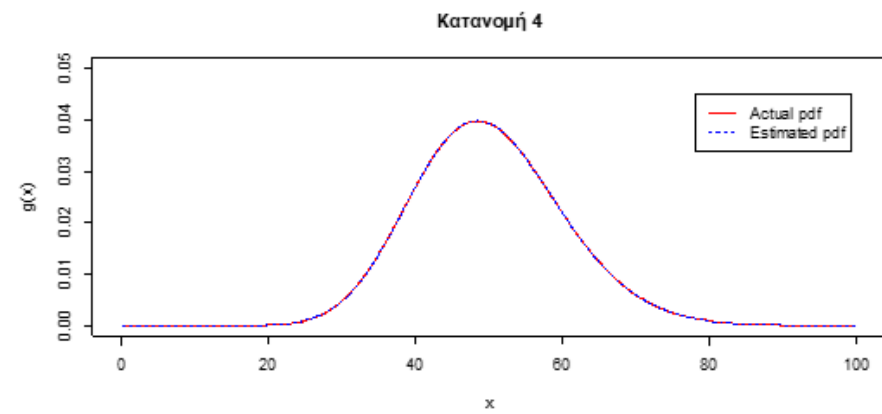
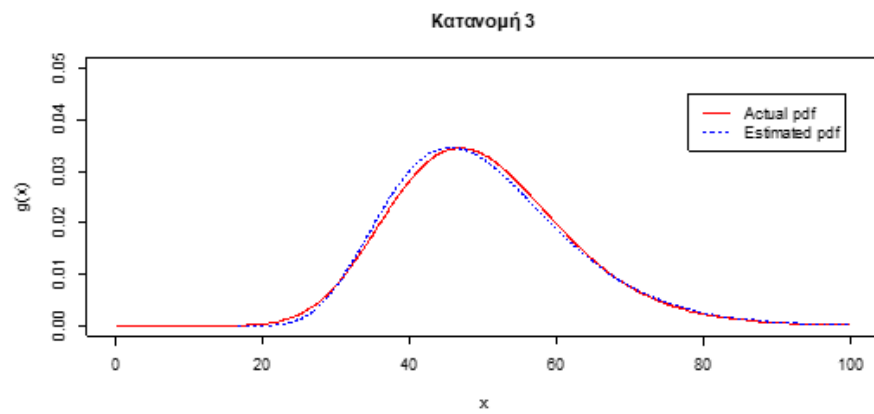
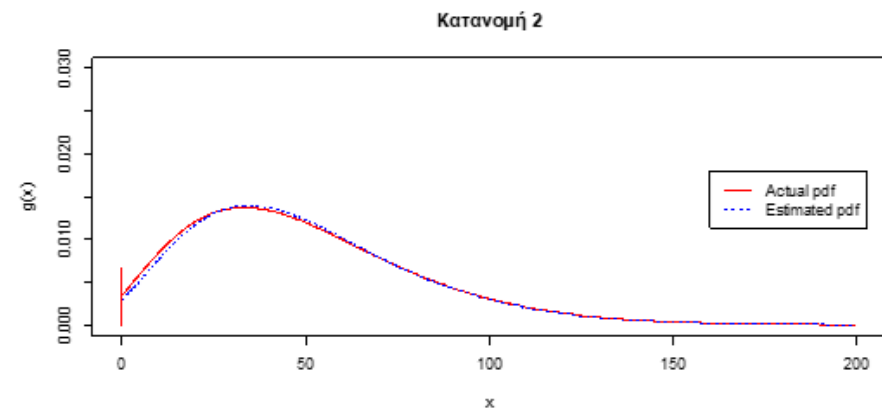
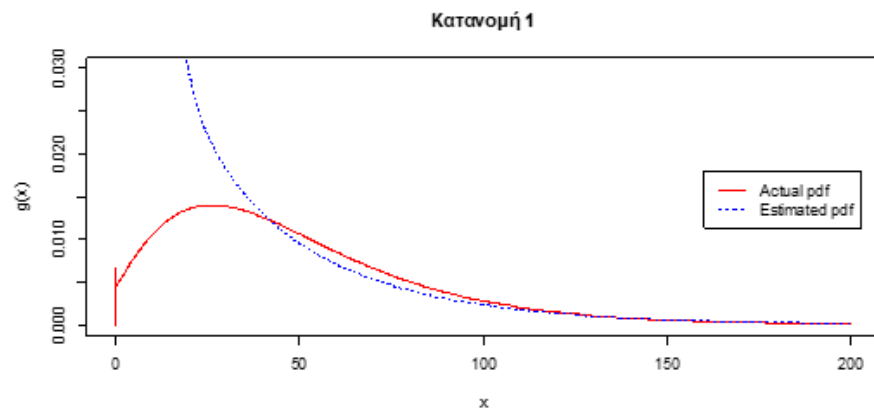
Σχήμα 3.9. Πραγματική και προσέγγιση της συνάρτησης κατανομής της S μέσω προσέγγισης Haldane

Η προσέγγιση Haldane δουλεύει αρκετά καλά στο δεξιό άκρο της κατανομής όπως φαίνεται καθαρά στον Πίνακα 3.5 ακόμη και στο πρώτο παράδειγμα

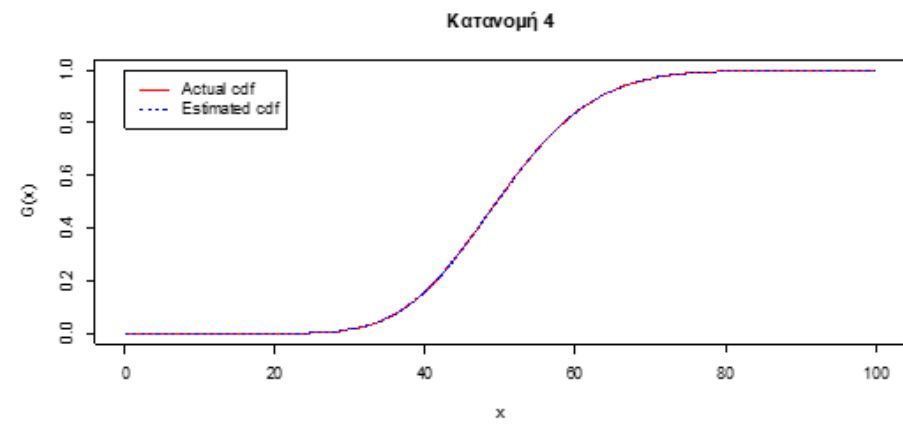
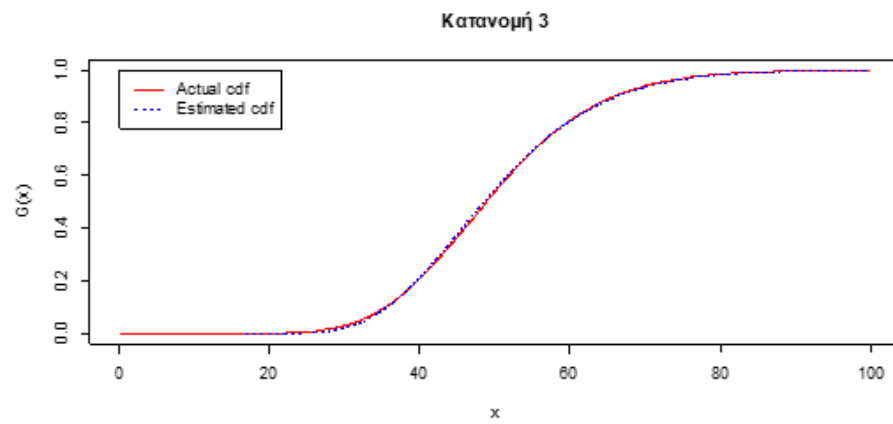
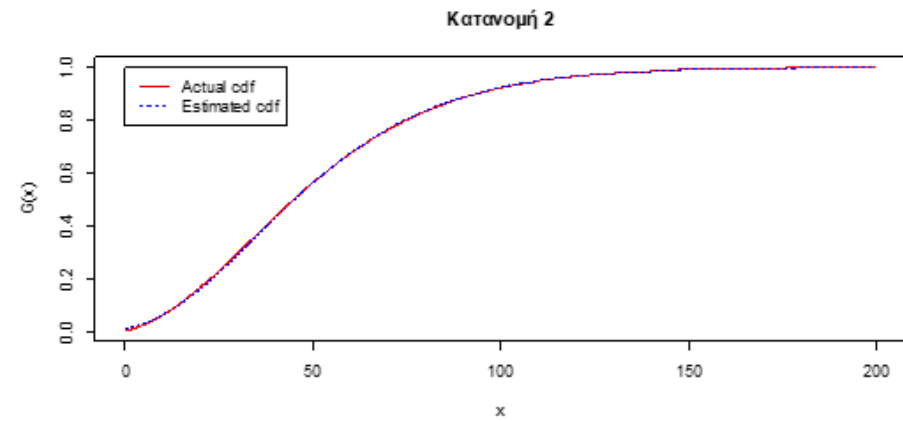
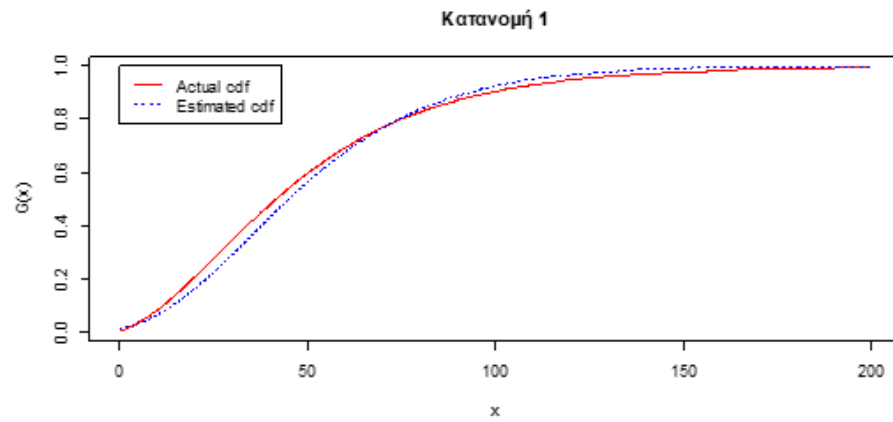
Οι υπολογισμοί που χρειάζονται να γίνουν για την περίπτωση της προσέγγισης **Wilson-Hilferty** είναι αρκετά απλοί, απλούστεροι σε σχέση με τη μέθοδο Haldane, αλλά η προσέγγιση είναι λιγότερο ακριβής. Τα αποτελέσματα για την προσέγγιση της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ φαίνονται στον παρακάτω πίνακα.

Πίνακας 3.6. Υπολογισμός της πιθανότητας $P(S > \mu_S + 4\sigma_S)$ μέσω προσέγγισης Wilson-Hilferty

	$\mu_S + 4\sigma_S$	Πραγματική πιθανότητα	Προσεγγιστική πιθανότητα
Κατανομή 1	204.91933	0.00548864	0.00812226
Κατανομή 2	178.14466	0.00208920	0.00234557
Κατανομή 3	98.98979	0.00156217	0.00137623
Κατανομή 4	90.52290	0.00029018	0.00030695



Σχήμα 3.10. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω προσέγγισης Wilson-Hilferty



Σχήμα 3.11. Πραγματική και προσέγγιση της συνάρτησης πυκνότητας της S μέσω προσέγγισης Wilson-Hilferty

3.2 2^η Εφαρμογή

Στην εφαρμογή αυτή θα θεωρήσουμε την περίπτωση όπου η τυχαία μεταβλητή που παριστάνει το ύψος της ατομικής ζημιάς X_j , $j \geq 1$ ακολουθεί Εκθετική κατανομή με παράμετρο $\theta > 0$, ενώ η τυχαία μεταβλητή N που παριστάνει το πλήθος των ζημιών θα υποτεθεί ότι κατανέμεται σύμφωνα με την κατανομή Poisson(λ) με $\lambda > 0$. Στόχος της εφαρμογής είναι να προσεγγίσουμε την συνάρτηση κατανομής της τυχαίας μεταβλητής $S = I(N \geq 1) \cdot \sum_{j=1}^N X_j$ συγκρίνοντας με αυτό τον τρόπο και τις προσεγγιστικές μεθόδους του Κεφαλαίου 2 μεταξύ τους ως προς την αποτελεσματικότητα, με κριτήριο την μέγιστη απόσταση τιμών μεταξύ της προσεγγιστικής τιμής και της πραγματικής τιμής της συνάρτησης κατανομής.

Παρουσιάζουμε για την κάθε μέθοδο τους βασικούς υπολογισμούς που πρέπει να γίνουν προκειμένου να εφαρμοστεί σωστά η μέθοδος και στη συνέχεια δίνεται μια αριθμητική εφαρμογή στην οποία παρουσιάζονται γραφικά οι προσεγγίσεις της συνάρτησης κατανομής της τυχαίας μεταβλητής S ενώ πραγματοποιείται και σύγκριση των Μέσων απόλυτων ποσοστιαίων σφαλμάτων (δίνεται με τη μορφή πίνακα). Οι πίνακες και τα γραφήματα που ακολουθούν υλοποιήθηκαν όλα στη γλώσσα προγραμματισμού R . Στο παράρτημα δίνονται οι σχετικοί κώδικες.

Συγκεκριμένα, στο παράδειγμα που θα ακολουθήσει για την τυχαία μεταβλητή N θα υποθέσουμε την κατανομή Poisson με $\lambda = 100$.

Ως προς τις βασικές παραμέτρους, με υπολογισμούς μπορεί να δειχθεί ότι:

$$\mu_N = 100, \quad \sigma_N^2 = 100, \quad \gamma_N = 100^{-1/2} = \frac{1}{10} = 0.1.$$

Για την τυχαία μεταβλητή X έχουμε επιλέξει την Εκθετική κατανομή με παράμετρο $\theta = 1$, οπότε

$$\mu_X = 1, \quad \sigma_X^2 = 1, \quad \gamma_X = 2.$$

Σύμφωνα με τους τύπους που έχουμε δώσει για τις σύνθετες κατανομές έχουμε:

$$\mu_S = E(S) = E(N)E(X) = \mu_N \cdot \mu_X = 100,$$

$$\sigma_S^2 = Var(S) = E[(S - \mu_S)^2] = E(V(S|N)) + V(E(S|N)) = \sigma_X^2 \mu_N + \mu_X^2 \sigma_N^2 = 200.$$

Επιπρόσθετα ο συντελεστής λοξότητας της S υπολογίζεται ως εξής:

$$\gamma_S = \frac{\mu_X^3 \sqrt{\gamma_N^3 \sigma_N^6} + 3\mu_X \sigma_N^2 \sigma_X^2 + \mu_X \gamma_X \sigma_X^{3/2}}{\sigma_S^{3/2}} = 0.21.$$

Μέθοδος κανονικής προσέγγιση (NA)

Χρησιμοποιούμε το γεγονός ότι η ποσότητα $E(N) = \lambda$ είναι μεγάλη, οπότε

$$\frac{S - \mu_S}{\sigma_S} \approx N(0,1)$$

Στην περίπτωση όπου $X \sim E(1)$, $N \sim P(100)$ είναι $\mu_S = 100$, $\sigma_S = \sqrt{200}$. Επομένως θα εφαρμόσουμε την προσέγγιση

$$P(S \leq s) \approx \Phi\left(\frac{s - 100}{\sqrt{200}}\right).$$

Μέθοδος της μετατοπισμένης Γάμμα κατανομής (TGA)

Η κατανομή της τυχαίας μεταβλητής S προσεγγίζεται από την κατανομή της $Y \sim G(a, \theta, k)$, της οποίας οι παράμετροι προκύπτουν από τη λύση των εξισώσεων

$$\mu_S = a\theta + k, \quad \sigma_S^2 = a\theta^2, \quad \gamma_S = \frac{2}{\sqrt{a}}.$$

Επειδή έχουμε ότι:

$$\mu_S = 100, \quad \sigma_S^2 = 200, \quad \gamma_S = 0.21$$

προκύπτει το σύστημα

$$\begin{cases} a\theta + k = 100 \\ a\theta^2 = 200 \\ \frac{2}{\sqrt{a}} = 0.21 \end{cases}$$

με λύση

$$a = \frac{4}{\gamma_S^2} = 90.70295, \quad \theta = \frac{\gamma_S \sigma_S}{2} = 1.484924, \quad k = \mu_S - \frac{2\sigma_S}{\gamma_S} = -34.68701.$$

Συνεπώς

$$F_S(s) = P(S \leq s) \approx P(Y \leq s) = F_G(s - k, a, \theta),$$

όπου $F_G(s, a, \theta)$ είναι η συνάρτηση κατανομής της συνήθους Γάμμα κατανομής $G(a, \theta)$.

Δυναμοκανονική μέθοδος (NPA)

Με βάση τη θεωρία έχουμε

$$F_S(s) = P(S \leq s) \approx \Phi \left(-\frac{3}{\gamma_S} + \sqrt{1 + \frac{9}{\gamma_S^2} + \frac{6(s - \mu_S)}{\gamma_S \sigma_S}} \right),$$

οπότε

$$F_S(s) \approx \Phi \left(-\frac{3}{0.21} + \sqrt{1 + \frac{9}{0.21^2} + \frac{6(s - 100)}{0.21\sqrt{200}}} \right).$$

Μέθοδος Bowers

Για τη μέθοδο Bowers έχουμε $Y = \beta S$ όπου $\beta = \frac{E(S)}{V(S)} = \frac{100}{200} = \frac{1}{2}$. Τότε,

$$E(Y) = \beta E(S) = \frac{E^2(S)}{V(S)} = 50$$

και

$$V(Y) = \beta^2 V(S) = \frac{E^2(S)}{V^2(S)} V(S) = \frac{E^2(S)}{V(S)} = 50,$$

οπότε

$$a = E(Y) = V(Y) = 50.$$

Έστω μ_k η k -οστή κεντρική ροπή της τυχαίας μεταβλητής Y . Τότε

$$\mu_1 = \mu_Y = E(Y) = E\left(\frac{S}{2}\right) = \frac{1}{2}E(N)E(X) = \frac{1}{2} \cdot 100 = 50,$$

$$\begin{aligned} \mu_2 &= E((Y - \mu_Y)^2) = E((Y - 50)^2) = E\left(\left(\frac{1}{2}S - 50\right)^2\right) = \frac{1}{4}\sigma_S^2 \\ &= \frac{1}{4}[E(V(S|N)) + V(E(S|N))] = \frac{1}{4}(\sigma_X^2 \mu_N + \mu_X^2 \sigma_N^2) \\ &= \frac{1}{4}(1 \cdot 100 + 1^2 \cdot 100) = \frac{1}{4} \cdot 200 = 50, \end{aligned}$$

$$\begin{aligned} \mu_3 &= E((Y - \mu_Y)^3) = E((Y - 50)^3) = E\left(\left(\frac{1}{2}S - 50\right)^3\right) \\ &= \frac{1}{8}[E(N) \cdot E((X - \mu_X)^3) + 3\sigma_N^2 \sigma_X^2 \mu_X + E((N - \mu_N)^3) \mu_X^3] = \frac{1}{8} \cdot 598.88 \\ &= 74.86, \end{aligned}$$

$$\begin{aligned}\mu_4 &= E((Y - \mu_Y)^4) = E((Y - 50)^4) = E\left(\left(\frac{1}{2}S - 50\right)^4\right) \\ &= \frac{1}{16} [E(N) \cdot E((X - \mu_X)^4) + 4\sigma_N^2 E((X - \mu_X)^3)\mu_X + 3(\sigma_N^2 \\ &\quad + \mu_N(\mu_N - 1))\sigma_X^4 + 6(E(N - \mu_N)^3 + \mu_N\sigma_N^2)\mu_X^2\sigma_X^2 + E(N - \mu_N)^4\mu_X^4] \\ &= \frac{1}{16} \cdot 122386.448 = 7649.15,\end{aligned}$$

$$\mu_5 = E((Y - \mu_Y)^5) = E((Y - 50)^5) = \frac{1}{32} \cdot 121121.76 = 37850.55.$$

Για να εφαρμόσουμε τη μέθοδο Bowers χρειαζόμαστε τις ποσότητες

$$A = \frac{\mu_3 - 2a}{3!} = -4.190131,$$

$$B = \frac{\mu_4 - 12\mu_3 - 3a^2 + 18a}{4!} = 6.30621,$$

$$C = \frac{\mu_5 - 20\mu_4 - (10a - 120)\mu_3 + 60a^2 - 144a}{5!} = -6.551706.$$

Τελικά

$$\begin{aligned}F_S(s) &= P(S \leq s) = P(Y \leq \beta s) = F_Y(\beta s) \\ &\approx F_G(\beta s; a) - Ae^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{2(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{(\beta s)^{a+2}}{\Gamma(a+3)} \right] \\ &\quad + Be^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{3(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{3(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{(\beta s)^{a+3}}{\Gamma(a+4)} \right] \\ &\quad - Ce^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{4(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{6(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{4(\beta s)^{a+3}}{\Gamma(a+4)} \right. \\ &\quad \left. + \frac{(\beta s)^{a+4}}{\Gamma(a+5)} \right].\end{aligned}$$

Μέθοδος Haldane

Θεωρούμε τον μετασχηματισμό-τυποποίηση του Haldane

$$Y = \left(\frac{S}{\mu_S}\right)^h.$$

Η παράμετρος h δίνεται από τον τύπο

$$h = 1 - \frac{\gamma_S \mu_S}{3\sigma_S} = 0.505.$$

Επίσης

$$\mu_Y = 1 - \frac{\sigma_S^2}{2\mu_S^2} h(1-h) \left(1 - \frac{\sigma_S^2}{4\mu_S^2} (2-h)(1-3h) \right) = 0.9974$$

και

$$\sigma_Y^2 = \frac{\sigma_S^2}{\mu_S^2} h^2 \left(1 - \frac{\sigma_S^2}{2\mu_S^2} (1-h)(1-3h) \right) = 0.005.$$

Υποθέτουμε τότε ότι η Y ακολουθεί προσεγγιστικά κανονική κατανομή, οπότε

$$F_S(s) = P\left(Y \leq \left(\frac{s}{\mu_S}\right)^h\right) \approx \Phi\left(\frac{\left(\frac{s}{\mu_S}\right)^h - \mu_Y}{\sigma_Y}\right) = \Phi\left(\frac{\left(\frac{s}{100}\right)^{0.505} - 0.9974}{0.07150876}\right).$$

Μέθοδος Wilson-Hilferty

Σύμφωνα με τη θεωρία οι ποσότητες που χρειαζόμαστε για να εφαρμόσουμε τη μέθοδο είναι

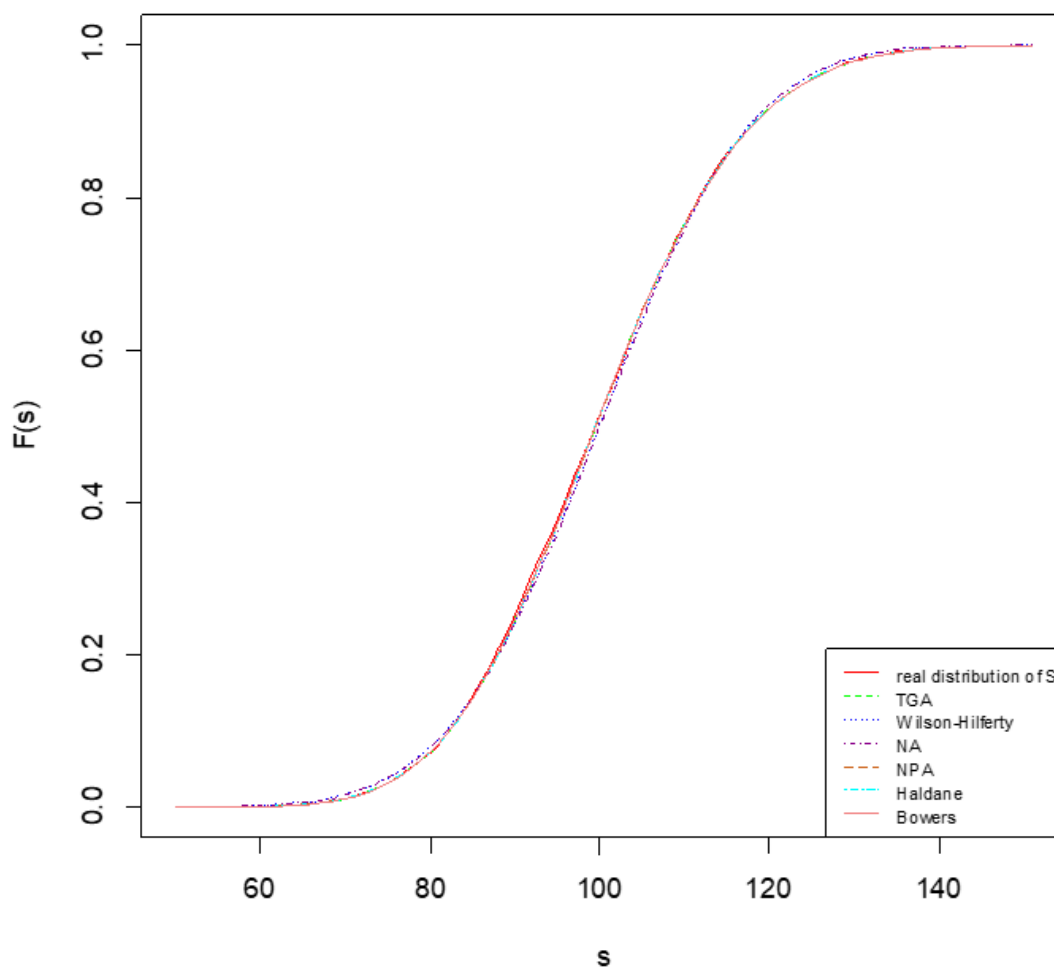
$$c_1 = \frac{\gamma_S}{6} - \frac{6}{\gamma_S} = -28.53643, \quad c_2 = 3\left(\frac{2}{\gamma_S}\right)^{\frac{2}{3}} = 13.47913, \quad c_3 = \frac{2}{\gamma_S} = 9.52381.$$

Τότε

$$\begin{aligned} F_S(s) &= P(S \leq s) \approx \Phi\left(c_1 + c_2 \left(\frac{s - \mu_S}{\sigma_S} + c_3\right)^{\frac{1}{3}}\right) \\ &= \Phi\left(-28.53643 + 13.47913 \left(\frac{s-100}{\sqrt{200}} + 9.52381\right)^{1/3}\right). \end{aligned}$$

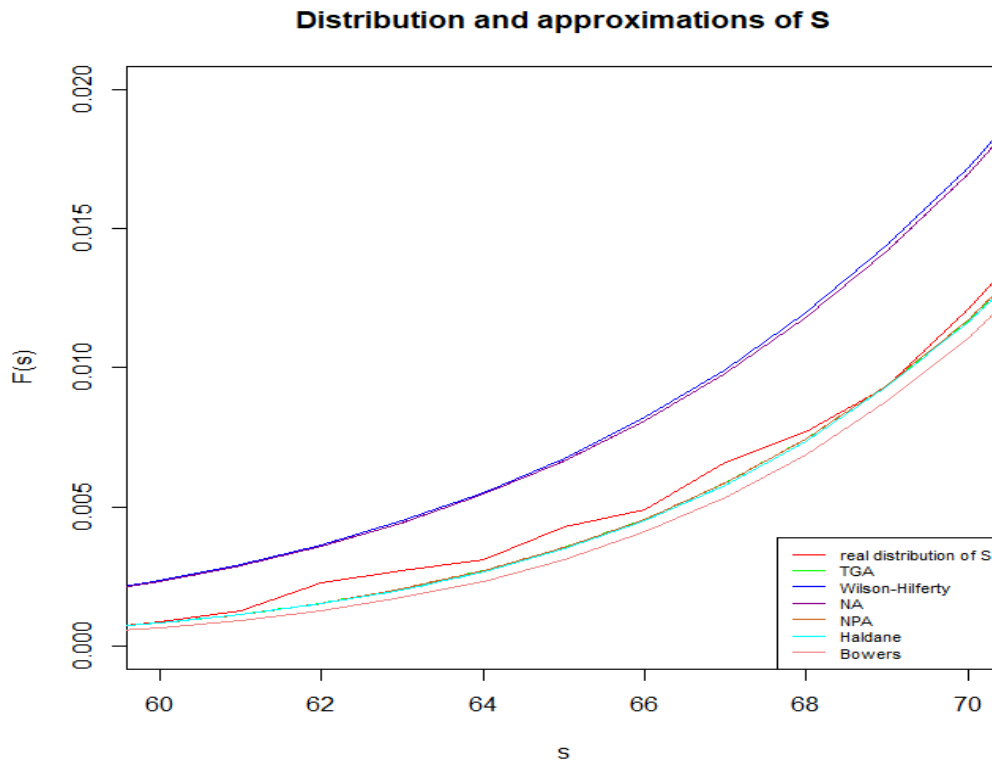
Στο παρακάτω σχήμα δίνεται γραφική παράσταση της συνάρτησης κατανομής της S με τις παραπάνω μεθόδους.

Distribution and approximations of S

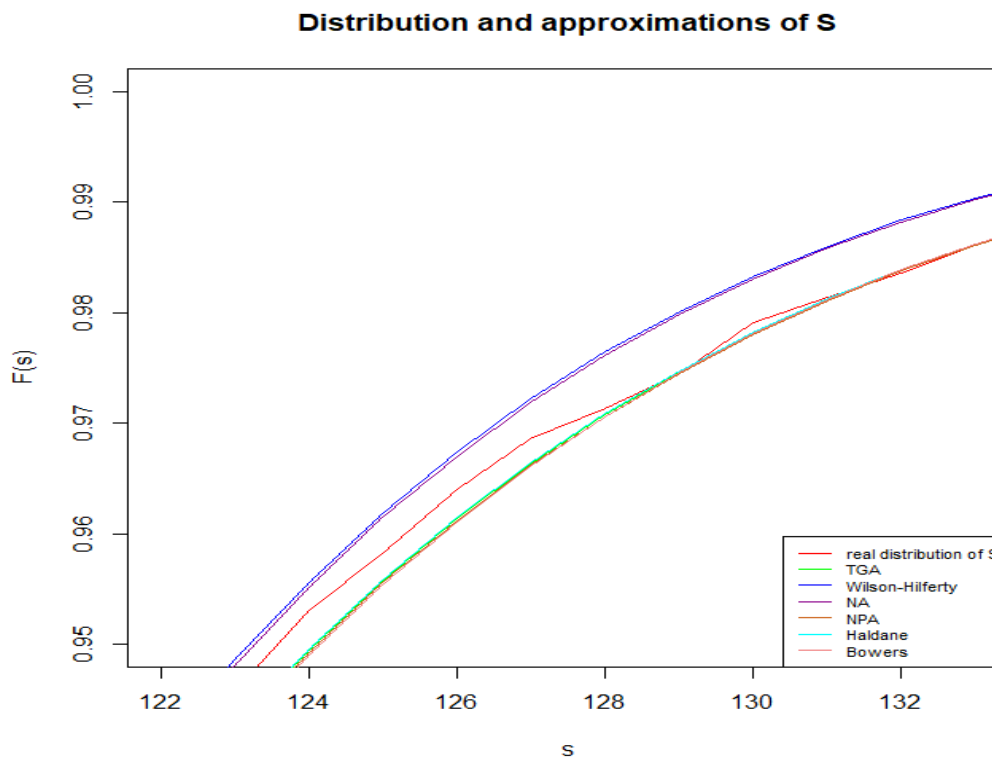


Σχήμα 3.12. Πραγματική και προσεγγίσεις της συνάρτησης κατανομής της S ($N \sim P(100)$, $X \sim Exp(1)$)

Από το παραπάνω γράφημα (αλλά και από άλλα πιο ευδιάκριτα σχήματα που ακολουθούν), μπορούμε να διαπιστώσουμε ότι η μέθοδος της Translated Gamma (TGA), η Δυναμοκανονική μέθοδος (NPA) καθώς επίσης και η μέθοδος Bowers προσεγγίζουν με αρκετά ικανοποιητικό τρόπο τη συνάρτηση κατανομής της S . Παρόμοια αποτελεσματικότητα στην προσέγγιση εμφανίζει και η μέθοδος του Haldane. Συγκριτικά, οι υπόλοιπες δύο μέθοδοι (Κανονική (NA) και Wilson-Hilferty) δίνουν σχετικά λιγότερο καλές προσεγγίσεις.



Σχήμα 3.13. Πραγματική και προσεγγίσεις της συνάρτησης κατανομής της S στο $[60, 70]$ ($N \sim P(100)$, $X \sim Exp(1)$)



Σχήμα 3.14. Πραγματική και προσεγγίσεις της συνάρτησης κατανομής της S στο $[122, 132]$ ($N \sim P(100)$, $X \sim Exp(1)$)

Στη συνέχεια θα συγκρίνουμε τις μεθόδους προσέγγισης με βάση το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα που αντιστοιχεί στην κάθε μία από αυτές.

Υπολογισμός Μέσου Απόλυτου Ποσοστιαίου Σφάλματος

$$R = \frac{100}{n} \sum_{i=1}^n \left| \frac{F(x_i) - F_{approx}(x_i)}{F(x_i)} \right|,$$

όπου F είναι η πραγματική συνάρτηση κατανομής της τυχαίας μεταβλητής S και F_{approx} είναι η εκάστοτε προσέγγιση της συνάρτησης κατανομής που μελετάμε (π.χ. Bowers κλπ.)

Ακολουθεί ο πίνακας Μέσου Απόλυτου Ποσοστιαίου Σφάλματος για τιμές της S στο διάστημα $[50,100]$ (Πίνακας 3.7) όπου βλέπουμε ποιες από τις προσεγγίσεις κρίνονται πιο αποτελεσματικές. Πιο συγκεκριμένα, η προσέγγιση της Κανονικής κατανομής κρίνεται χειρότερη από τις υπόλοιπες με Μέσο απόλυτο ποσοστιαίο σφάλμα 0.0050 ενώ και η μέθοδος Wilson-Hilferty εμφανίζει παρόμοια αποτελεσματικότητα στην προσέγγιση. Η μέθοδος της Translated Gamma (TGA), η Δυναμοκανονική μέθοδος (NPA) καθώς επίσης και οι μέθοδοι Bowers και Haldane προσεγγίζουν με αρκετά ικανοποιητικό τρόπο τη συνάρτηση κατανομής της S .

Πίνακας 3.7. Πίνακας Μέσου Απόλυτου Σφάλματος για τιμές της S στο διάστημα $[50,100]$

Μέθοδος Προσέγγισης	Μέσο Απόλυτο Ποσοστιαίο Σφάλμα
Δυναμοκανονική (NPA)	0.0022
Μετατοπισμένη Γάμμα (TGA)	0.0023
Bowers	0.0023
Haldane	0.0024
Wilson-Hilferty	0.0044
Κανονική προσέγγιση (NA)	0.0050

3.3 3^η Εφαρμογή

Σε αυτή την εφαρμογή θα θεωρήσουμε την περίπτωση όπου η τυχαία μεταβλητή που παριστάνει το ύψος της ατομικής ζημιάς $X_j, j \geq 1$ ακολουθεί την κατανομή Weibull με παραμέτρους $\kappa, \lambda > 0$, ενώ για την τυχαία μεταβλητή N που παριστάνει το πλήθος των ζημιών θα υποθεθεί ότι κατανέμεται σύμφωνα με την αρνητικής διωνυμική κατανομή. Στόχος της εφαρμογής είναι να προσεγγίσουμε την συνάρτηση κατανομής της τυχαίας μεταβλητής $S = I(N \geq 1) \cdot \sum_{j=1}^N X_j$ συγκρίνοντας με αυτό τον τρόπο τις προσεγγιστικές μεθόδους του Κεφαλαίου 2 μεταξύ τους ως προς την αποτελεσματικότητα, με κριτήριο την μέγιστη απόσταση τιμών μεταξύ της προσεγγιστικής τιμής και της πραγματικής τιμής της συνάρτησης κατανομής.

Οι πίνακες και τα γραφήματα που ακολουθούν υλοποιήθηκαν όλα στη γλώσσα προγραμματισμού R . Η επιλογή των παραμέτρων είναι τυχαία και με βάση τον κώδικα του παραρτήματος ο αναγνώστης μπορεί να θέσει τις δικές του τιμές ώστε να κάνει εκ νέου τη σύγκριση των μεθόδων. Παρουσιάζονται για την κάθε μέθοδο οι βασικοί υπολογισμοί που πρέπει να γίνουν προκειμένου να εφαρμοστεί σωστά η μέθοδος και στη συνέχεια δίνεται μια αριθμητική εφαρμογή στην οποία παρουσιάζονται γραφικά οι προσεγγίσεις της συνάρτησης κατανομής της τυχαίας μεταβλητής S ενώ πραγματοποιείται και σύγκριση των Μέσων απόλυτων ποσοστιαίων σφαλμάτων (δίνεται με τη μορφή πίνακα).

Συγκεκριμένα στο παράδειγμα που θα ακολουθήσει, για την κατανομή της τυχαίας μεταβλητής N θα υποθέσουμε ότι $r = 4, p = 0.7$. Έτσι

$$\mu_N = \frac{1.2}{0.7} = 1.714286, \quad \sigma_N^2 = 2.44898, \quad \gamma_N = \frac{1+q}{\sqrt{rq}} = 1.186732.$$

Για την τυχαία μεταβλητή X έχουμε επιλέξει Weibull με παραμέτρους $k = 2, \lambda = 1$. Έτσι

$$\mu_X = \Gamma\left(1 + \frac{1}{2}\right) = 0.886, \quad \sigma_X^2 = \Gamma(2) - \Gamma^2\left(1 + \frac{1}{2}\right) = 0.2146,$$

$$\gamma_X = \frac{\Gamma\left(1 + \frac{3}{2}\right) - 3\mu_X\sigma_X^2 - \mu_X^3}{\sigma_X^3} = 0.6311$$

Σύμφωνα με τους τύπους που έχουμε δώσει για τις σύνθετες κατανομές έχουμε:

$$\mu_S = E(S) = \mu_N \cdot \mu_X = 1.519$$

$$\sigma_S^2 = Var(S) = E[(S - \mu_S)^2] = E(V(S|N)) + V(E(S|N)) = \sigma_X^2\mu_N + \mu_X^2\sigma_N^2 = 2.291$$

Επιπρόσθετα ο συντελεστής λοξότητας υπολογίζεται ως εξής:

$$\gamma_S = \frac{\mu_X^3\sqrt{\gamma_N^3\sigma_N^6} + 3\mu_X\sigma_N^2\sigma_X^2 + \mu_X\gamma_X\sigma_X^{3/2}}{\sigma_S^{3/2}} = 1.346$$

Μέθοδος κανονικής προσέγγιση (NA)

Σύμφωνα με τα παραπάνω αποτελέσματα η μέθοδος NA δίνει

$$P(S \leq s) \approx \Phi\left(\frac{s - \mu_S}{\sigma_S}\right) = \Phi\left(\frac{s - 1.519}{1.513891}\right).$$

Αφού η τιμή μ_N είναι μικρή αναμένουμε φτωχή προσέγγιση της πραγματικής τιμής της συνάρτησης κατανομής.

Μέθοδος της μετατοπισμένης Γάμμα Κατανομής (TGA)

Η κατανομή της τυχαίας μεταβλητής S προσεγγίζεται από την κατανομή της $Y \sim G(a, \theta, k)$, της οποίας οι παράμετροι προκύπτουν από τη λύση των εξισώσεων

$$\mu_S = a\theta + k, \quad \sigma_S^2 = a\theta^2, \quad \gamma_S = \frac{2}{\sqrt{a}}.$$

Επειδή έχουμε ότι

$$\mu_S = 1.519, \quad \sigma_S^2 = 2.291, \quad \gamma_S = 1.346$$

προκύπτει το σύστημα

$$\begin{cases} a\theta + k = 1.519 \\ a\theta^2 = 2.291 \\ \frac{2}{\sqrt{a}} = 1.346 \end{cases}$$

με αντίστοιχη λύση

$$a = \frac{4}{\gamma_S^2} = 2.205895, \quad \theta = \frac{\gamma_S \sigma_S}{2} = 1.019177, \quad k = \mu_S - \frac{2\sigma_S}{\gamma_S} = -0.7289521.$$

Συνεπώς,

$$F_S(s) = P(S \leq s) \approx P(Y \leq s) = F_G(s - k, a, \theta),$$

όπου $F_G(s, a, \theta)$ είναι η συνάρτηση κατανομής της συνήθους Γάμμα κατανομής $G(a, \theta)$.

Δυναμοκανονική μέθοδος NPA

Με βάση τη θεωρία έχουμε:

$$F_S(s) = P(S \leq s) \approx \Phi \left(-\frac{3}{\gamma_S} + \sqrt{1 + \frac{9}{\gamma_S^2} + \frac{6(s - \mu_S)}{\gamma_S \sigma_S}} \right)$$

οπότε

$$F_S(s) \approx \Phi \left(-\frac{3}{1.346} + \sqrt{1 + \frac{9}{1.346^2} + \frac{6(s - 1.519)}{1.346\sqrt{2.291}}} \right).$$

Μέθοδος Bowers

Για τη μέθοδο Bowers έχουμε $Y = \beta S$ όπου $\beta = \frac{E(S)}{V(S)} = \frac{1.519}{2.291} = 0.6630461$. Τότε,

$$E(Y) = \beta E(S) = \frac{E^2(S)}{V(S)} = 1.0071$$

και

$$V(Y) = \beta^2 V(S) = \frac{E^2(S)}{V^2(S)} V(S) = \frac{E^2(S)}{V(S)} = 1.0071,$$

οπότε

$$a = E(Y) = V(Y) = 1.0071$$

Έστω μ_k η k -οστή κεντρική ροπή της τυχαίας μεταβλητής Y . Τότε

$$\mu_1 = \mu_Y = E(Y) = 1.0071,$$

$$\begin{aligned} \mu_2 &= E((Y - \mu_Y)^2) = E\left(\left(\frac{1}{2}S - \mu_Y\right)^2\right) = \frac{1}{4}\sigma_S^2 = \frac{1}{4}[E(V(S|N)) + V(E(S|N))] \\ &= \frac{1}{4}(\sigma_X^2\mu_N + \mu_X^2\sigma_N^2) = \frac{1}{4}(0.2146 \cdot 1.714286 + 0.886^2 \cdot 2.44898) \\ &= \frac{1}{4}(0.3679 + 1.9224) = 0.5726, \end{aligned}$$

$$\begin{aligned} \mu_3 &= E((Y - \mu_Y)^3) = E\left(\left(\frac{1}{2}S - \mu_Y\right)^3\right) \\ &= \frac{1}{8}[E(N) \cdot E((X - \mu_X)^3) + 3\sigma_N^2\sigma_X^2\mu_X + E((N - \mu_N)^3)\mu_X^3] \\ &= \frac{1}{8} \cdot 11.041576 = 1.380197, \end{aligned}$$

$$\begin{aligned}
\mu_4 &= E((Y - \mu_Y)^4) = E\left(\left(\frac{1}{2}S - \mu_Y\right)^4\right) \\
&= \frac{1}{16} [E(N) \cdot E((X - \mu_X)^4) + 4\sigma_N^2 E((X - \mu_X)^3)\mu_X + 3(\sigma_N^2 \\
&\quad + \mu_N(\mu_N - 1)\sigma_X^4) + 6(E(N - \mu_N)^3 + \mu_N\sigma_N^2)\mu_X^2\sigma_X^2 + E(N - \mu_N)^4\mu_X^4] \\
&= \frac{1}{16} \cdot 87.97472 = 5.49842,
\end{aligned}$$

$$\mu_5 = E((Y - \mu_Y)^5) = E\left(\left(\frac{1}{2}S - \mu_Y\right)^5\right) = \frac{1}{32} \cdot 619.68768 = 19.36524.$$

Για να εφαρμόσουμε τη μέθοδο Bowers χρειαζόμαστε τους εξής υπολογισμούς:

$$A = \frac{\mu_3 - 2a}{3!} = -0.1033005,$$

$$B = \frac{\mu_4 - 12\mu_3 - 3a^2 + 18a}{4!} = 0.1640024,$$

$$C = \frac{\mu_5 - 20\mu_4 - (10a - 120)\mu_3 + 60a^2 - 144a}{5!} = -0.5383596.$$

Τελικά

$$\begin{aligned}
F_S(s) &= P(S \leq s) = P(Y \leq \beta s) = F_Y(\beta s) \\
&\approx F_G(\beta s; a) - Ae^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{2(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{(\beta s)^{a+2}}{\Gamma(a+3)} \right] \\
&\quad + Be^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{3(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{3(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{(\beta s)^{a+3}}{\Gamma(a+4)} \right] \\
&\quad - Ce^{-\beta s} \left[\frac{(\beta s)^a}{\Gamma(a+1)} - \frac{4(\beta s)^{a+1}}{\Gamma(a+2)} + \frac{6(\beta s)^{a+2}}{\Gamma(a+3)} - \frac{4(\beta s)^{a+3}}{\Gamma(a+4)} \right. \\
&\quad \left. + \frac{(\beta s)^{a+4}}{\Gamma(a+5)} \right].
\end{aligned}$$

Μέθοδος Haldane

Θεωρούμε τον μετασχηματισμό-τυποποίηση του Haldane

$$Y = \left(\frac{S}{\mu_S}\right)^h.$$

Η παράμετρος h δίνεται από τον τύπο

$$h = 1 - \frac{\gamma_S \mu_S}{3\sigma_S} = 0.5487306$$

Επίσης

$$\mu_Y = 1 - \frac{\sigma_S^2}{2\mu_S^2} h(1-h) \left(1 - \frac{\sigma_S^2}{4\mu_S^2} (2-h)(1-3h) \right) = 0.8484476$$

και

$$\sigma_Y^2 = \frac{\sigma_S^2}{\mu_S^2} h^2 \left(1 - \frac{\sigma_S^2}{2\mu_S^2} (1-h)(1-3h) \right) = 0.3422519.$$

Υποθέτουμε τότε ότι η Y ακολουθεί προσεγγιστικά κανονική κατανομή, οπότε

$$F_S(s) = P\left(Y \leq \left(\frac{s}{\mu_S}\right)^h\right) \approx \Phi\left(\frac{\left(\frac{s}{\mu_S}\right)^h - \mu_Y}{\sigma_Y}\right) = \Phi\left(\frac{\left(\frac{s}{1.519}\right)^{0.5487306} - 0.8484476}{0.585023}\right)$$

Μέθοδος Wilson-Hilferty

Σύμφωνα με τη θεωρία οι ποσότητες που χρειαζόμαστε για να εφαρμόσουμε τη μέθοδο είναι

$$c_1 = \frac{\gamma_S}{6} - \frac{6}{\gamma_S} = -4.222906, \quad c_2 = 3 \left(\frac{2}{\gamma_S}\right)^{\frac{2}{3}} = 3.9, \quad c_3 = \frac{2}{\gamma_S} = 1.48258.$$

Θέτοντας

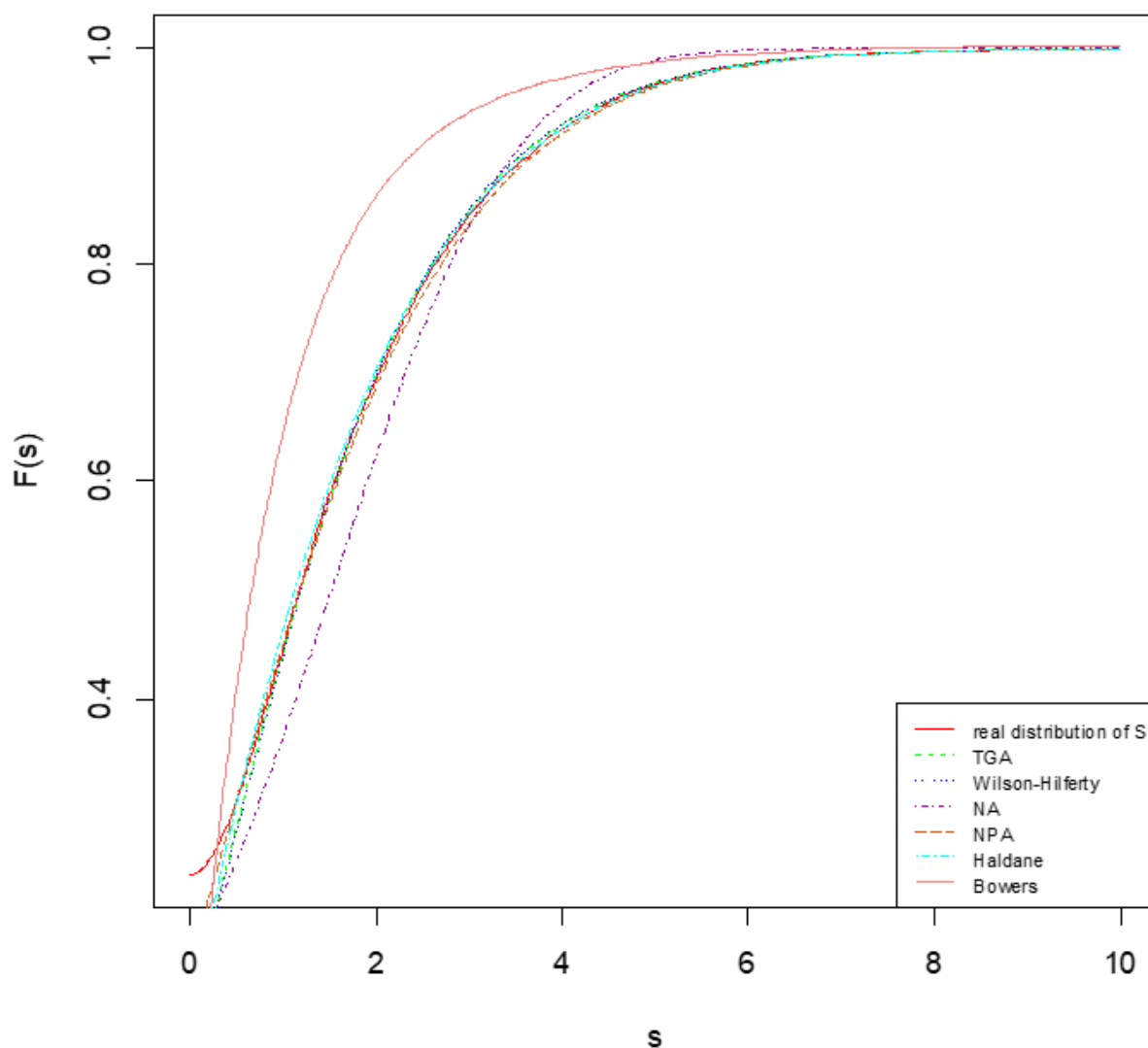
$$u(y) = \left(\frac{y - c_1}{c_2}\right)^{1/3} - c_3 = \left(\frac{y + 4.222906}{3.9}\right)^{\frac{1}{3}} - 1.48258 = z$$

Τότε

$$\begin{aligned} F_S(s) &= P(S \leq s) \approx \Phi\left(c_1 + c_2 \left(\frac{s - \mu_S}{\sigma_S} + c_3\right)^{\frac{1}{3}}\right) \\ &= \Phi\left(-4.222906 + 3.9 \left(\frac{s - 1.519}{\sqrt{2.291}} + 1.48258\right)^{\frac{1}{3}}\right) \end{aligned}$$

Στο παρακάτω σχήμα δίνεται γραφική παράσταση της συνάρτησης κατανομής της S με τις παραπάνω μεθόδους.

Distribution and approximations of S



Σχήμα 3.15. Πραγματική και προσεγγίσεις της συνάρτησης κατανομής της S

Από το παραπάνω γράφημα, μπορούμε να διαπιστώσουμε ότι η μέθοδος της Translated Gamma (TGA), η μέθοδος Wilson-Hilferty και η Δυναμοκανονική μέθοδος (NPA), κρίνονται πιο αποτελεσματικές έναντι των υπολοίπων και είναι πολύ κοντά στην πραγματική κατανομή της τυχαίας μεταβλητής S. Μια σχετικά καλή προσέγγιση δίνει επίσης η μέθοδος Haldane, ενώ η μέθοδος της Κανονικής κατανομής και η μέθοδος Bowers είναι λιγότερο ικανοποιητικές ως προς την προσέγγιση με σημαντική απόκλιση συγκριτικά με τις υπόλοιπες μεθόδους.

Στη συνέχεια θα συγκρίνουμε τις μεθόδους προσέγγισης με βάση το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα που αντιστοιχεί στην κάθε μία από αυτές.

Υπολογισμός Μέσου Απόλυτου Ποσοστιαίου Σφάλματος:

$$R = \frac{100}{n} \sum_{i=1}^n \left| \frac{F(x_i) - F_{approx}(x_i)}{F(x_i)} \right|,$$

όπου F είναι η πραγματική συνάρτηση κατανομής της τυχαίας μεταβλητής S και F_{approx} είναι η εκάστοτε προσέγγιση της συνάρτησης κατανομής που μελετάμε (π.χ. Bowers κλπ.)

Ακολουθεί ο πίνακας Μέσου Απόλυτου Ποσοστιαίου Σφάλματος για τιμές της S στο διάστημα $[0,10]$ (Πίνακας 3.8) όπου βλέπουμε ποιες από τις προσεγγίσεις κρίνονται πιο αποτελεσματικές. Πιο συγκεκριμένα, η προσέγγιση Bowers κρίνεται ως η λιγότερο αποτελεσματική μέθοδος, με σημαντική απόκλιση συγκριτικά με τις υπόλοιπες, με Μέσο απόλυτο ποσοστιαίο σφάλμα 0.1017. Η Δυναμοκανονική μέθοδος (NPA) εμφανίζεται ως η πιο αποτελεσματική ενώ οι μέθοδοι Translated Gamma (TGA), Haldane και Wilson προσεγγίζουν με αρκετά ικανοποιητικό τρόπο τη συνάρτηση κατανομής της S .

Πίνακας 3.8. Πίνακας Μέσου Απόλυτου Ποσοστιαίου Σφάλματος για τιμές της S στο $[0,10]$

Μέθοδος Προσέγγισης	Μέσο Απόλυτο Ποσοστιαίο σφάλμα
Δυναμοκανονική (NPA)	0.0115
Μετατοπισμένη Γάμμα (TGA)	0.0166
Haldane	0.0170
Wilson - Hilferty	0.0176
Κανονική προσέγγιση (NA)	0.0479
Bowers	0.1017

3.4 Συμπεράσματα

Συνοψίζουμε τώρα κάποια συμπεράσματα σχετικά με τις προσεγγιστικές μεθόδους που αναπτύξαμε στο κεφάλαιο αυτό.

Σύμφωνα με τα αποτελέσματα που προέκυψαν από τις εφαρμογές θα μπορούσαμε να ισχυριστούμε ότι η προσεγγιστική μέθοδος που δείχνει να υπερέχει έναντι των υπολοίπων έστω και οριακά είναι η **Δυναμοκανονική μέθοδος (NPA)**, τόσο για την αποτελεσματικότητα της, όσο και για την ευκολία στην εφαρμογή της. Μειονέκτημα της η Αναποτελεσματικότητα της για πολύ μεγάλο γ_S .

Ενδιαφέρον παρουσιάζει και η **μέθοδος Haldane**, η οποία είναι επίσης εύκολη στην εφαρμογή της και πέτυχε τις πιο ικανοποιητικές προσεγγίσεις στην 1η εφαρμογή. Επίσης εμφανίζεται αρκετά Ικανοποιητική στην προσαρμογή της δεξιάς ουράς.

Έπειτα, η μέθοδος της **μετατοπισμένης Γάμμα (TGA)** είναι επίσης μια αξιόπιστη μέθοδος, η οποία πάντα βρίσκεται ανάμεσα στις καλύτερες προσεγγίσεις με βάση τις εφαρμογές που έχουμε αναπτύξει. Για να εφαρμοστεί όμως, πρέπει να βρεθούν πρώτα οι παράμετροι α , θ , κ , γεγονός που δημιουργεί μια παραπάνω υπολογιστική πολυπλοκότητα. Αυτό όμως δεν είναι αποτρεπτικός παράγοντας για την χρησιμοποίηση της.

Η μέθοδος **Wilson-Hilferty** απαιτεί ελαφρώς λιγότερες πράξεις από τις προηγούμενες μεθόδους όμως παρουσιάζεται λιγότερο αποτελεσματική στις προσεγγίσεις που έχουμε δει στη εργασία. Είναι αναγκαίο να υπολογιστούν οι σταθερές c_1, c_2, c_3 κατά την εφαρμογή της.

Σχετικά με τη **μέθοδο Bowers** τώρα, απαιτεί περισσότερες πράξεις προκειμένου να υπολογιστούν οι απαραίτητες ποσότητες που χρειάζονται για την εφαρμογή της και στις εφαρμογές που αναπτύχθηκαν, δεν είναι το ίδιο αποτελεσματική με τις προηγούμενες προσεγγίσεις. Πάντως εμφανίζεται καλή στην προσαρμογή της δεξιάς ουράς.

Η **Κανονική μέθοδος (NA)** παρόλο που είναι εξαιρετικά απλή στην εφαρμογή της, εντούτοις δεν κατορθώνει να προσεγγίσει ικανοποιητικά την κατανομή της τυχαίας μεταβλητής S κυρίως στην περίπτωση όπου ο συντελεστής λοξότητας γ_S δεν βρίσκεται κοντά στο 0 και αυτό έχει ως αποτέλεσμα να υπάρχουν μεγάλες αποκλίσεις στην προσέγγιση.

Παράρτημα

Κώδικας για το παράδειγμα 2.4: Προσέγγιση Bowers

```
library(actuar)
ra=1 # X~E(1)
lam=16 # N~P(16)
a=0;b=50;h=0.001
fx <- discretize(pexp(x,rate=ra), from=a, to=b, step=h, method="rounding")
Gs.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev = fx,
lambda = lam, x.scale = h,maxit=100000, tol=0.000001)
s=seq(0,50,5)
Gs.Exact <- Gs.rec(s)
#####
a=8;beta=1/2;A=-2/3;B=1;C=-1.1
y=beta*s
z1 <- pgamma(y, shape=a, scale=1)
z2 <- -A*exp(-y)*( ((y^a)/gamma(a+1))- ((2*y^(a+1))/gamma(a+2))+
((y^(a+2))/gamma(a+3)) )
z3 <- B*exp(-y)*( ((y^a)/gamma(a+1)) - ((3*y^(a+1))/gamma(a+2)) +
((3*y^(a+2))/gamma(a+3)) - ((y^(a+3))/gamma(a+4)) )
z4 <- -C*exp(-y)*( ((y^a)/gamma(a+1)) - ((4*y^(a+1))/gamma(a+2)) +
((6*y^(a+2))/gamma(a+3)) - ((4*y^(a+3))/gamma(a+4)) +
((y^(a+4))/gamma(a+5)))
Gs.Approx <- z1+z2+z3+z4
```

#(Πίνακας 2.1)

```
m <- cbind(s,Gs.Exact,Gs.Approx)
rownames(m) <- rep("", times=length(s))
colnames(m) <- c("s"," F(s): Πραγματική τιμή", " F(s): Προσεγγιστική
τιμή")
round(m, digits=8)
#####
Gs.rec(0)
exp(lam*(fx[1]-1))
#####
w1=(1-A+B-C)*pgamma(y, shape=a, scale=1);w1
w2=(3*A-4*B+5*C)*pgamma(y, shape=a+1, scale=1);w2
w3=(-3*A+6*B-10*C)*pgamma(y, shape=a+2, scale=1);w3
w4=(A-4*B+10*C)*pgamma(y, shape=a+3, scale=1);w4
w5=(B-5*C)*pgamma(y, shape=a+4, scale=1);w5
w6=C*pgamma(y, shape=a+5, scale=1);w6
Gs.Approx <- w1+w2+w3+w4+w5+w6
```

#####Figure - CDFS (Διάγραμμα)

```
s <- seq(0,30,0.01)
plot(s, Gs.rec(s), type="l", col="red", ylab="G(x)", main="Κατανομή S")
y=beta*s
z1 <- pgamma(y, shape=a, scale=1)
z2 <- -A*exp(-y)*( ((y^a)/gamma(a+1))- ((2*y^(a+1))/gamma(a+2))+
((y^(a+2))/gamma(a+3)) )
z3 <- B*exp(-y)*( ((y^a)/gamma(a+1)) - ((3*y^(a+1))/gamma(a+2)) +
((3*y^(a+2))/gamma(a+3)) - ((y^(a+3))/gamma(a+4)) )
z4 <- -C*exp(-y)*( ((y^a)/gamma(a+1)) - ((4*y^(a+1))/gamma(a+2)) +
((6*y^(a+2))/gamma(a+3)) - ((4*y^(a+3))/gamma(a+4)) +
((y^(a+4))/gamma(a+5)))
Gs.Approx <- z1+z2+z3+z4
lines(s,Gs.Approx,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))
```

Κώδικας για το παράδειγμα 2.5: Προσέγγιση Haldane

```
library(actuar)
# Poisson
lam=16
meanN=varN=lam
skewN=1/sqrt(lam)
# Exponential
ra <- 1
meanX=1/ra
varX=1/ra^2
skewX=2
# Compound
(meanS <- meanN*meanX)
(varS <- meanN*varX+varN*meanX^2)
(sdS <- sqrt(varS))
(skewS <-
(skewN*varN^(3/2)*meanX^3+3*varN*meanX*varX+meanN*skewX*varX^(3/2))/(varS^(3
/2)))
# Haldane
h <- 1-(skewS*meanS)/(3*sdS)
mY <- 1- ((varS/(2*meanS^2))*h*(1-h)*(1-(varS/(4*meanS^2))*(2-h)*(1-3*h)))
s2Y <- (varS/meanS^2)*(h^2)*(1-(varS/(2*meanS^2))*(1-h)*(1-3*h))
s <- seq(0,50,5)
z <- ((s/meanS)^h-mY)/sqrt(s2Y)
Gs.Approx <- pnorm(z,0,1)
```

(Πίνακας 2.2)

```
# Table
m <- cbind(s,Gs.Approx)
rownames(m) <- rep("", times=length(s))
colnames(m) <- c("s", " F(s): Προσεγγιστική τιμή")
round(m, digits=8)
```

Κώδικας για το παράδειγμα 2.6: Προσέγγιση Wilson-Hilferty

```
library(actuar)
# Poisson
lam=16
meanN=varN=lam
skewN=1/sqrt(lam)
# Exponential
ra <- 1
meanX=1/ra
varX=1/ra^2
skewX=2
# Compound
(meanS <- meanN*meanX)
(varS <- meanN*varX+varN*meanX^2)
(sdS <- sqrt(varS))
(skewS <-
(skewN*varN^(3/2)*meanX^3+3*varN*meanX*varX+meanN*skewX*varX^(3/2))/(varS^(3
/2)))
# Wilson-Hilferty
c1 <- (skewS/6)-(6/skewS)
c2 <- 3*((2/skewS)^(2/3))
c3 <- 2/skewS
s <- seq(0,50,5)
z <- c1+c2*((s-meanS)/sdS)+c3)^(1/3)
Gs.Approx <- pnorm(z,0,1)
```

(Πίνακας 2.3)

```

# Table
m <- cbind(s,Gs.Approx)
rownames(m) <- rep("", times=length(s))
colnames(m) <- c("s","      F(s): Προσεγγιστική τιμή")
round(m, digits=8)

```

Κώδικας για την εφαρμογή 3.1

```

library(actuar)
# Poisson
lam1=lam2=5
lam3=lam4=50
# Pareto
sh1=4; sc1=30
sh2=40; sc2=390
sh3=4; sc3=3
sh4=40; sc4=39
# Pareto density plots
a=0;b=10
curve(dpareto(x,shape=sh1,scale=sc1), xlim = c(a,b), ylim=c(0,1.2),
col="red", main="Καμπύλες Pareto(a,b)", ylab="f(x)",)
x=seq(a,b,0.01)
lines(x,dpareto(x,shape=sh2,scale=sc2), col="green", lty=2)
lines(x,dpareto(x,shape=sh3,scale=sc3), col="blue", lty=3)
lines(x,dpareto(x,shape=sh4,scale=sc4), col="brown", lty=4)
legend(locator(1),c("a=4, b=30","a=40, b=390","a=4, b=3","a=40,
b=39"),col=c("red","green","blue","brown"), lty=1:4)
#####
#Discretize Pareto Unbiased
#####
au=0; bu=20000
h <- 0.01
fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)
#####
#Compound Poisson-Pareto SN
#####
##### Recursive
Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)
Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)
gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)
gs4.rec <- diff(Gs4.rec)
##### P(SN=0)
exp(-lam1)
gs1.rec[1];Gs1.rec(0);exp(lam1*(fx1[1]-1))
gs2.rec[1];Gs2.rec(0);exp(lam2*(fx2[1]-1))
exp(-lam3)
gs3.rec[1];Gs3.rec(0);exp(lam3*(fx3[1]-1))
gs4.rec[1];Gs4.rec(0);exp(lam4*(fx4[1]-1))
##### Plot density SN

```

```

lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)

```

#Σχήμα 3.1

```

plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.04), main="Σύνθετη Poisson-Pareto Κατανομή")
segments(0,0,0,gs1.rec[1], col="red")
lines(x, (1/h)*gs2.rec[2:up], type="l", col="green", lty=2)
segments(0,0,0,gs2.rec[1], col="green",lty=2)
lines(x, (1/h)*gs3.rec[2:up],type="l", col="blue", lty=3)
segments(0,0,0,gs3.rec[1], col="blue",lty=2)
lines(x, (1/h)*gs4.rec[2:up], type="l", col="brown", lty=4)
segments(0,0,0,gs4.rec[1], col="brown",lty=4)
legend(locator(1),c("Κατανομή 1: λ=5, a=4, b=30","Κατανομή 2: λ=5, a=40,
b=390","Κατανομή 3: λ=50, a=4, b=3","Κατανομή 4: λ=50, a=40,
b=39"),col=c("red","green","blue","brown"), lty=1:4)

```

#Κανονική προσέγγιση

```

library(actuar)
# Poisson
lam1=lam2=5
meanN1=meanN2=varN1=varN2=lam1
lam3=lam4=50
meanN3=meanN4=varN3=varN4=lam3
# Pareto
sh1=4; sc1=30
sh2=40; sc2=390
sh3=4; sc3=3
sh4=40; sc4=39
(meanX1=mpareto(1,sh1,sc1)); (varX1=mpareto(2,sh1,sc1)-
(mpareto(1,sh1,sc1))^2)
(meanX2=mpareto(1,sh2,sc2)); (varX2=mpareto(2,sh2,sc2)-
(mpareto(1,sh2,sc2))^2)
(meanX3=mpareto(1,sh3,sc3)); (varX3=mpareto(2,sh3,sc3)-
(mpareto(1,sh3,sc3))^2)
(meanX4=mpareto(1,sh4,sc4)); (varX4=mpareto(2,sh4,sc4)-
(mpareto(1,sh4,sc4))^2)
# Compound
(meanS1 <- meanN1*meanX1)
(varS1 <- meanN1*varX1+varN1*meanX1^2)
(meanS2 <- meanN2*meanX2)
(varS2 <- meanN2*varX2+varN2*meanX2^2)
(meanS3 <- meanN3*meanX3)
(varS3 <- meanN3*varX3+varN3*meanX3^2)
(meanS4 <- meanN4*meanX4)
(varS4 <- meanN4*varX4+varN4*meanX4^2)
#####
#Discretize Unbiased - Recursive
#####
au=0; bu=20000
h <- 0.01
fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)
Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)

```

```

Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)
gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)
gs4.rec <- diff(Gs4.rec)
#####
# Normal approximation
#####
# (Πίνακας 3.1)

# Table 1
t1 <- meanS1+4*sqrt(varS1)
t2 <- meanS2+4*sqrt(varS2)
t3 <- meanS3+4*sqrt(varS3)
t4 <- meanS4+4*sqrt(varS4)
pa1 <- pnorm(t1,meanS1,sqrt(varS1),lower.tail = FALSE)
pa2 <- pnorm(t2,meanS2,sqrt(varS2),lower.tail = FALSE)
pa3 <- pnorm(t3,meanS3,sqrt(varS3),lower.tail = FALSE)
pa4 <- pnorm(t4,meanS4,sqrt(varS4),lower.tail = FALSE)
(p1 <- 1-Gs1.rec(t1))
(p2 <- 1-Gs2.rec(t2))
(p3 <- 1-Gs3.rec(t3))
(p4 <- 1-Gs4.rec(t4))
t <- c(t1,t2,t3,t4)
pa <- c(pa1,pa2,pa3,pa4)
p <- c(p1,p2,p3,p4)
m <- cbind(t,p,pa)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4
")
colnames(m) <- c("x", " G(x):True Probability", " G(x):Normal
Approximation")
round(m, digits=8)

```

(Σχήματα 3.2 και 3.3)

```

# Figure 2
par(mfrow=c(2,2))
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.015), xlim=c(-50,200), main="Κατανομή 1")
segments(0,0,0,gs1.rec[1], col="red")
x <- seq(-50,lim,h)
lines(x, dnorm(x,meanS1,sqrt(varS1)),col="blue",lty=2)
legend(150,0.014,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs2.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.015), xlim=c(-50,200), main="Κατανομή 2")
segments(0,0,0,gs2.rec[1], col="red")
x <- seq(-50,lim,h)
lines(x, dnorm(x,meanS2,sqrt(varS2)),col="blue",lty=2)
legend(150,0.014,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs3.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 3")
segments(0,0,0,gs3.rec[1], col="red")

```



```

x <- seq(0,lim,h)
lines(x,dnorm(x,meanS3,sqrt(varS3)),col="blue",lty=2)
legend(80,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x,(1/h)*gs4.rec[2:up],type="l",col="red",ylab="g(x)",
ylim=c(0,0.05),main="Κατανομή 4")
segments(0,0,0,gs4.rec[1],col="red")
x <- seq(0,lim,h)
lines(x,dnorm(x,meanS4,sqrt(varS4)),col="blue",lty=2)
legend(80,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
# Figure 2 - CDFS
par(mfrow=c(2,2))
lim <- 200
up <- lim*(1/h)+1
x <- seq(-50,lim,h)
plot(x,Gs1.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 1")
lines(x,pnorm(x,meanS1,sqrt(varS1)),col="blue",lty=2)
legend(-50,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))
lim <- 200
up <- lim*(1/h)+1
x <- seq(-50,lim,h)
plot(x,Gs2.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 2")
lines(x,pnorm(x,meanS2,sqrt(varS2)),col="blue",lty=2)
legend(-50,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))
lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs3.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 3")
lines(x,pnorm(x,meanS3,sqrt(varS3)),col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))
lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs4.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 4")
lines(x,pnorm(x,meanS4,sqrt(varS4)),col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))

```

#Μετατοπισμένη Γάμμα

```

library(actuar)
# Poisson
lam1=lam2=5
meanN1=meanN2=varN1=varN2=lam1
skewN1=skewN2=1/sqrt(lam1)
lam3=lam4=50
meanN3=meanN4=varN3=varN4=lam3
skewN3=skewN4=1/sqrt(lam3)
# Pareto
sh1=4; sc1=30
sh2=40; sc2=390
sh3=4; sc3=3
sh4=40; sc4=39
(meanX1=mpareto(1,sh1,sc1)); (varX1=mpareto(2,sh1,sc1)-
(mpareto(1,sh1,sc1))^2)
(meanX2=mpareto(1,sh2,sc2)); (varX2=mpareto(2,sh2,sc2)-
(mpareto(1,sh2,sc2))^2)
(meanX3=mpareto(1,sh3,sc3)); (varX3=mpareto(2,sh3,sc3)-
(mpareto(1,sh3,sc3))^2)
(meanX4=mpareto(1,sh4,sc4)); (varX4=mpareto(2,sh4,sc4)-
(mpareto(1,sh4,sc4))^2)
(skewX1=(2*(1+sh1)/(sh1-3))*sqrt((sh1-2)/sh1))
(skewX2=(2*(1+sh2)/(sh2-3))*sqrt((sh2-2)/sh2))
(skewX3=(2*(1+sh3)/(sh3-3))*sqrt((sh3-2)/sh3))
(skewX4=(2*(1+sh4)/(sh4-3))*sqrt((sh4-2)/sh4))

```

```

# Compound
(meanS1 <- meanN1*meanX1)
(varS1 <- meanN1*varX1+varN1*meanX1^2)
(skewS1 <-
(skewN1*varN1^(3/2)*meanX1^3+3*varN1*meanX1*varX1+meanN1*skewX1*varX1^(3/2))
/(varS1^(3/2)))
(meanS2 <- meanN2*meanX2)
(varS2 <- meanN2*varX2+varN2*meanX2^2)
(skewS2 <-
(skewN2*varN2^(3/2)*meanX2^3+3*varN2*meanX2*varX2+meanN2*skewX2*varX2^(3/2))
/(varS2^(3/2)))
(meanS3 <- meanN3*meanX3)
(varS3 <- meanN3*varX3+varN3*meanX3^2)
(skewS3 <-
(skewN3*varN3^(3/2)*meanX3^3+3*varN3*meanX3*varX3+meanN3*skewX3*varX3^(3/2))
/(varS3^(3/2)))
(meanS4 <- meanN4*meanX4)
(varS4 <- meanN4*varX4+varN4*meanX4^2)
(skewS4 <-
(skewN4*varN4^(3/2)*meanX4^3+3*varN4*meanX4*varX4+meanN4*skewX4*varX4^(3/2))
/(varS4^(3/2)))
# Translated Gamma
(a1 <- (2/skewS1)^2)
(b1 <- sqrt(varS1/a1))
(k1 <- meanS1-a1*b1)

(a2 <- (2/skewS2)^2)
(b2 <- sqrt(varS2/a2))
(k2 <- meanS2-a2*b2)

(a3 <- (2/skewS3)^2)
(b3 <- sqrt(varS3/a3))
(k3 <- meanS3-a3*b3)

(a4 <- (2/skewS4)^2)
(b4 <- sqrt(varS4/a4))
(k4 <- meanS4-a4*b4)

a <- c(a1,a2,a3,a4)
b <- c(b1,b2,b3,b4)
k <- c(k1,k2,k3,k4)

```

#(Πίνακας 3.2)

```

# Table for Translated Gamma
m <- cbind(a,b,k)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4 ")
colnames(m) <- c("a", " b", " k")
round(m, digits=8)
#####
#Discretize Pareto Unbiased
#####
au=0; bu=20000
h <- 0.01
fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)
#####
#Compounnd Poisson-Pareto SN
#####
##### Recursive

```

```

Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)
Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)

gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)
gs4.rec <- diff(Gs4.rec)

```

(Πίνακας 3.3)

```

# Table 2
t1 <- meanS1+4*sqrt(varS1)
t2 <- meanS2+4*sqrt(varS2)
t3 <- meanS3+4*sqrt(varS3)
t4 <- meanS4+4*sqrt(varS4)
(pa1 <- pgamma(t1-k1, shape=a1, scale=b1, lower.tail = FALSE))
(pa2 <- pgamma(t2-k2, shape=a2, scale=b2, lower.tail = FALSE))
(pa3 <- pgamma(t3-k3, shape=a3, scale=b3, lower.tail = FALSE))
(pa4 <- pgamma(t4-k4, shape=a4, scale=b4, lower.tail = FALSE))

(p1 <- 1-Gs1.rec(t1))
(p2 <- 1-Gs2.rec(t2))
(p3 <- 1-Gs3.rec(t3))
(p4 <- 1-Gs4.rec(t4))
t <- c(t1,t2,t3,t4)
pa <- c(pa1,pa2,pa3,pa4)
p <- c(p1,p2,p3,p4)
m <- cbind(t,p,pa)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4 ")
colnames(m) <- c("x", " G(x):True Probability", " G(x):Translated Gamma Approximation")
round(m, digits=8)

```

(Σχήματα 3.4 και 3.5)

```

##### Plot density SN
# Figure 2
par(mfrow=c(2,2))
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.03), xlim=c(0,200), main="Κατανομή 1")
segments(0,0,0,gs1.rec[1], col="red")
x <- seq(0,lim-k1,h)
lines(x+k1,dgamma(x, shape=a1, scale=b1), col="blue", lty=2)
legend(160,0.027,c("Actual pdf", "Estimated pdf"),
lty=1:2, col=c("red", "blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs2.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.03), xlim=c(0,200), main="Κατανομή 2")
segments(0,0,0,gs2.rec[1], col="red")
x <- seq(0,lim-k2,h)
lines(x+k2,dgamma(x, shape=a2, scale=b2), col="blue", lty=2)
legend(160,0.027,c("Actual pdf", "Estimated pdf"),
lty=1:2, col=c("red", "blue"))

```

```

lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs3.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 3")
segments(0,0,0,gs3.rec[1], col="red")
x <- seq(0,lim-k3,h)
lines(x+k3,dgamma(x,shape=a3,scale=b3),col="blue",lty=2)
legend(80,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs4.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 4")
segments(0,0,0,gs4.rec[1], col="red")
x <- seq(0,lim-k4,h)
lines(x+k4,dgamma(x,shape=a4,scale=b4),col="blue",lty=2)
legend(80,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
# Figure 2 - CDFS
par(mfrow=c(2,2))

lim <- 200
up <- lim*(1/h)+1
x <- seq(-50,lim,h)
plot(x, Gs1.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 1")
lines(x,pnorm(x,meanS1,sqrt(varS1)),col="blue",lty=2)
legend(-50,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(-50,lim,h)
plot(x, Gs2.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 2")
lines(x,pnorm(x,meanS2,sqrt(varS2)),col="blue",lty=2)
legend(-50,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs3.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 3")
lines(x,pnorm(x,meanS3,sqrt(varS3)),col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs4.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 4")
lines(x,pnorm(x,meanS4,sqrt(varS4)),col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

```

#Δυναμοκανονική προσέγγιση

```

library(actuar)
# Poisson
lam1=lam2=5
meanN1=meanN2=varN1=varN2=lam1
skewN1=skewN2=1/sqrt(lam1)

lam3=lam4=50
meanN3=meanN4=varN3=varN4=lam3
skewN3=skewN4=1/sqrt(lam3)

# Pareto
sh1=4; sc1=30
sh2=40; sc2=390

```

```

sh3=4; sc3=3
sh4=40; sc4=39

(meanX1=mpareto(1,sh1,sc1)); (varX1=mpareto(2,sh1,sc1)-
(mpareto(1,sh1,sc1))^2)
(meanX2=mpareto(1,sh2,sc2)); (varX2=mpareto(2,sh2,sc2)-
(mpareto(1,sh2,sc2))^2)
(meanX3=mpareto(1,sh3,sc3)); (varX3=mpareto(2,sh3,sc3)-
(mpareto(1,sh3,sc3))^2)
(meanX4=mpareto(1,sh4,sc4)); (varX4=mpareto(2,sh4,sc4)-
(mpareto(1,sh4,sc4))^2)

(skewX1=(2*(1+sh1)/(sh1-3))*sqrt((sh1-2)/sh1))
(skewX2=(2*(1+sh2)/(sh2-3))*sqrt((sh2-2)/sh2))
(skewX3=(2*(1+sh3)/(sh3-3))*sqrt((sh3-2)/sh3))
(skewX4=(2*(1+sh4)/(sh4-3))*sqrt((sh4-2)/sh4))

# Compound

(meanS1 <- meanN1*meanX1)
(varS1 <- meanN1*varX1+varN1*meanX1^2)
(skewS1 <-
(skewN1*varN1^(3/2)*meanX1^3+3*varN1*meanX1*varX1+meanN1*skewX1*varX1^(3/2))
/(varS1^(3/2)))

(meanS2 <- meanN2*meanX2)
(varS2 <- meanN2*varX2+varN2*meanX2^2)
(skewS2 <-
(skewN2*varN2^(3/2)*meanX2^3+3*varN2*meanX2*varX2+meanN2*skewX2*varX2^(3/2))
/(varS2^(3/2)))

(meanS3 <- meanN3*meanX3)
(varS3 <- meanN3*varX3+varN3*meanX3^2)
(skewS3 <-
(skewN3*varN3^(3/2)*meanX3^3+3*varN3*meanX3*varX3+meanN3*skewX3*varX3^(3/2))
/(varS3^(3/2)))

(meanS4 <- meanN4*meanX4)
(varS4 <- meanN4*varX4+varN4*meanX4^2)
(skewS4 <-
(skewN4*varN4^(3/2)*meanX4^3+3*varN4*meanX4*varX4+meanN4*skewX4*varX4^(3/2))
/(varS4^(3/2)))
#####
#Discretize Unbiased - Recursive
#####
au=0; bu=20000
h <- 0.01
fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)

Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)
Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)

gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)

```

```

gs4.rec <- diff(Gs4.rec)
#####
#NPower
#####
Gs1.npower <- aggregateDist("npower", moments=c(meanS1,varS1,skewS1))
Gs2.npower <- aggregateDist("npower", moments=c(meanS2,varS2,skewS2))
Gs3.npower <- aggregateDist("npower", moments=c(meanS3,varS3,skewS3))
Gs4.npower <- aggregateDist("npower", moments=c(meanS4,varS4,skewS4))

#####
# Normal Power approximation

```

(Πίνακας 3.4)

```

# Table 4

t1 <- meanS1+4*sqrt(varS1)
t2 <- meanS2+4*sqrt(varS2)
t3 <- meanS3+4*sqrt(varS3)
t4 <- meanS4+4*sqrt(varS4)
(pa1 <- 1-Gs1.npower(t1))
(pa2 <- 1-Gs2.npower(t2))
(pa3 <- 1-Gs3.npower(t3))
(pa4 <- 1-Gs4.npower(t4))
(p1 <- 1-Gs1.rec(t1))
(p2 <- 1-Gs2.rec(t2))
(p3 <- 1-Gs3.rec(t3))
(p4 <- 1-Gs4.rec(t4))
t <- c(t1,t2,t3,t4)
pa <- c(pa1,pa2,pa3,pa4)
p <- c(p1,p2,p3,p4)
m <- cbind(t,p,pa)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4 ")
colnames(m) <- c("x", " G(x):True Probability", " G(x):Normal Power Approximation")
round(m, digits=8)

#####
#####

```

(Σχήματα 3.6 και 3.7)

```

# Figure 5

par(mfrow=c(2,2))

lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.015), xlim=c(0,200), main="Κατανομή 1")
segments(0,0,0,gs1.rec[1], col="red")
x <- seq(0,lim,h)
Gs1 <- pnorm((-3/skewS1)+ sqrt((9/skewS1^2)+1+(6*(x-meanS1)/(skewS1*sqrt(varS1))))),0,1)
v2 <- diff(Gs1)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(120,0.014,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs2.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.015), xlim=c(0,200), main="Κατανομή 2")

```

```

segments(0,0,0,gs2.rec[1], col="red")
x <- seq(0,lim,h)
Gs2 <- pnorm((-3/skewS2)+ sqrt((9/skewS2^2)+1+(6*(x-
meanS2)/(skewS2*sqrt(varS2))))),0,1)
v2 <- diff(Gs2)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(120,0.014,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs3.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 3")
segments(0,0,0,gs3.rec[1], col="red")
x <- seq(0,lim,h)
Gs3 <- pnorm((-3/skewS3)+ sqrt((9/skewS3^2)+1+(6*(x-
meanS3)/(skewS3*sqrt(varS3))))),0,1)
v2 <- diff(Gs3)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(60,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs4.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 4")
segments(0,0,0,gs4.rec[1], col="red")
x <- seq(0,lim,h)
Gs4 <- pnorm((-3/skewS4)+ sqrt((9/skewS4^2)+1+(6*(x-
meanS4)/(skewS4*sqrt(varS4))))),0,1)
v2 <- diff(Gs4)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(60,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))
#####
# Figure - CDFS

par(mfrow=c(2,2))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs1.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 1")
Gs1 <- pnorm((-3/skewS1)+ sqrt((9/skewS1^2)+1+(6*(x-
meanS1)/(skewS1*sqrt(varS1))))),0,1)
lines(x,Gs1,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs2.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 2")
Gs2 <- pnorm((-3/skewS2)+ sqrt((9/skewS2^2)+1+(6*(x-
meanS2)/(skewS2*sqrt(varS2))))),0,1)
lines(x,Gs2,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs3.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 3")
Gs3 <- pnorm((-3/skewS3)+ sqrt((9/skewS3^2)+1+(6*(x-
meanS3)/(skewS3*sqrt(varS3))))),0,1)

```

```

lines(x, Gs3, col="blue", lty=2)
legend(0,1,c("Actual cdf", "Estimated cdf"), lty=1:2, col=c("red", "blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs4.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 4")
Gs4 <- pnorm((-3/skewS4)+ sqrt((9/skewS4^2)+1+(6*(x-
meanS4)/(skewS4*sqrt(varS4))))), 0,1)
lines(x, Gs4, col="blue", lty=2)
legend(0,1,c("Actual cdf", "Estimated cdf"), lty=1:2, col=c("red", "blue"))

```

#Μέθοδος Haldane

```

library(actuar)
# Poisson
lam1=lam2=5
meanN1=meanN2=varN1=varN2=lam1
skewN1=skewN2=1/sqrt(lam1)

lam3=lam4=50
meanN3=meanN4=varN3=varN4=lam3
skewN3=skewN4=1/sqrt(lam3)

# Pareto
sh1=4; sc1=30
sh2=40; sc2=390
sh3=4; sc3=3
sh4=40; sc4=39

(meanX1=mpareto(1, sh1, sc1)); (varX1=mpareto(2, sh1, sc1)-
(mpareto(1, sh1, sc1))^2)
(meanX2=mpareto(1, sh2, sc2)); (varX2=mpareto(2, sh2, sc2)-
(mpareto(1, sh2, sc2))^2)
(meanX3=mpareto(1, sh3, sc3)); (varX3=mpareto(2, sh3, sc3)-
(mpareto(1, sh3, sc3))^2)
(meanX4=mpareto(1, sh4, sc4)); (varX4=mpareto(2, sh4, sc4)-
(mpareto(1, sh4, sc4))^2)

(skewX1=(2*(1+sh1)/(sh1-3))*sqrt((sh1-2)/sh1))
(skewX2=(2*(1+sh2)/(sh2-3))*sqrt((sh2-2)/sh2))
(skewX3=(2*(1+sh3)/(sh3-3))*sqrt((sh3-2)/sh3))
(skewX4=(2*(1+sh4)/(sh4-3))*sqrt((sh4-2)/sh4))

# Compound

(meanS1 <- meanN1*meanX1)
(varS1 <- meanN1*varX1+varN1*meanX1^2)
(sdS1 <- sqrt(varS1))
(skewS1 <-
(skewN1*varN1^(3/2)*meanX1^3+3*varN1*meanX1*varX1+meanN1*skewX1*varX1^(3/2))
/(varS1^(3/2)))
(hS1 <- 1-(skewS1*meanS1)/(3*sdS1))

(meanS2 <- meanN2*meanX2)
(varS2 <- meanN2*varX2+varN2*meanX2^2)
(sdS2 <- sqrt(varS2))
(skewS2 <-
(skewN2*varN2^(3/2)*meanX2^3+3*varN2*meanX2*varX2+meanN2*skewX2*varX2^(3/2))
/(varS2^(3/2)))
(hS2 <- 1-(skewS2*meanS2)/(3*sdS2))

(meanS3 <- meanN3*meanX3)
(varS3 <- meanN3*varX3+varN3*meanX3^2)
(sdS3 <- sqrt(varS3))

```



```

(skewS3 <-
(skewN3*varN3^(3/2)*meanX3^3+3*varN3*meanX3*varX3+meanN3*skewX3*varX3^(3/2))
/(varS3^(3/2)))
(hs3 <- 1-(skewS3*meanS3)/(3*sdS3))

(meanS4 <- meanN4*meanX4)
(varS4 <- meanN4*varX4+varN4*meanX4^2)
(sdS4 <- sqrt(varS4))
(skewS4 <-
(skewN4*varN4^(3/2)*meanX4^3+3*varN4*meanX4*varX4+meanN4*skewX4*varX4^(3/2))
/(varS4^(3/2)))
(hs4 <- 1-(skewS4*meanS4)/(3*sdS4))

#####
#Discretize Unbiased - Recursive
#####
au=0; bu=20000
h <- 0.01
fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)

Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)
Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)

gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)
gs4.rec <- diff(Gs4.rec)

#####
# Haldane Example Distribution Function - Table 5
#####

# Example 1
(t1 <- meanS1+4*sqrt(varS1))
n1 <- log(t1/meanS1)+(varS1/(2*meanS1^2))-((varS1)^2)/(4*meanS1^4)
n2 <- (sdS1/meanS1)*sqrt(1-(varS1/(2*meanS1^2)))
(pa1 <- 1-pnorm(n1/n2,0,1))

# Example 2
(t2 <- meanS2+4*sqrt(varS2))
mY2 <- 1-(varS2/(2*meanS2^2))*hS2*(1-hS2)*(1-(varS2/(4*meanS2^2))*(2-
hS2)*(1-3*hS2))
varY2 <- (varS2/meanS2^2)*hS2^2*(1-(varS2/(2*meanS2^2))*(1-hS2)*(1-3*hS2))
(pa2 <- 1-pnorm(((t2/meanS2)^hS2-mY2)/sqrt(varY2),0,1))

# Example 3
(t3 <- meanS3+4*sqrt(varS3))
n1 <- log(t3/meanS3)+(varS3/(2*meanS3^2))-((varS3)^2)/(4*meanS3^4)
n2 <- (sdS3/meanS3)*sqrt(1-(varS3/(2*meanS3^2)))
(pa3 <- 1-pnorm(n1/n2,0,1))

# Example 4
(t4 <- meanS4+4*sqrt(varS4))
mY4 <- 1-(varS4/(2*meanS4^2))*hS4*(1-hS4)*(1-(varS4/(4*meanS4^2))*(2-
hS4)*(1-3*hS4))

```

```
varY4 <- (varS4/meanS4^2)*hS4^2*(1-(varS4/(2*meanS4^2))*(1-hS4)*(1-3*hS4))
(pa4 <- 1-pnorm(((t4/meanS4)^hS4-mY4)/sqrt(varY4),0,1))
```

(Πίνακας 3.5)

```
# Table 5

(p1 <- 1-Gs1.rec(t1))
(p2 <- 1-Gs2.rec(t2))
(p3 <- 1-Gs3.rec(t3))
(p4 <- 1-Gs4.rec(t4))
t <- c(t1,t2,t3,t4)
pa <- c(pa1,pa2,pa3,pa4)
p <- c(p1,p2,p3,p4)
m <- cbind(t,p,pa)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4 ")
colnames(m) <- c("x", " G(x):True Probability", " G(x):Haldane's Approximation")
round(m, digits=8)
```

(Σχήματα 3.8 και 3.9)

```
#####
# Figure 6

par(mfrow=c(2,2))

# Example 1
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.02), xlim=c(0,200), main="Κατανομή 1")
segments(0,0,0,gs1.rec[1], col="red")
x <- seq(0,lim,h)
n1 <- log(x/meanS1)+(varS1/(2*meanS1^2))-((varS1)^2)/(4*meanS1^4)
n2 <- (sdS1/meanS1)*sqrt(1-(varS1/(2*meanS1^2)))
Gs1 <- pnorm(n1/n2,0,1)
v2 <- diff(Gs1)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(160,0.018,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

# Example 2
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs2.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.02), xlim=c(0,200), main="Κατανομή 2")
segments(0,0,0,gs2.rec[1], col="red")
x <- seq(1,lim,h)
Gs2 <- pnorm(((x/meanS2)^hS2-mY2)/sqrt(varY2),0,1)
v2 <- diff(Gs2)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(160,0.018,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

# Example 3
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs3.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 3")
segments(0,0,0,gs3.rec[1], col="red")
```

```

x <- seq(0,lim,h)
n1 <- log(x/meanS3)+(varS3/(2*meanS3^2))-((varS3)^2)/(4*meanS3^4)
n2 <- (sdS3/meanS3)*sqrt(1-(varS3/(2*meanS3^2)))
Gs3 <- pnorm(n1/n2,0,1)
v2 <- diff(Gs3)
v3 <- c(NA,v2)
lines(x,(1/h)*v3,col="blue",lty=2)
legend(78,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

# Example 4
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x,(1/h)*gs4.rec[2:up],type="l",col="red",ylab="g(x)",
ylim=c(0,0.05),main="Κατανομή 4")
segments(0,0,0,gs4.rec[1],col="red")
x <- seq(0,lim,h)
Gs4 <- pnorm(((x/meanS4)^hS4-mY4)/sqrt(varY4),0,1)
v2 <- diff(Gs4)
v3 <- c(NA,v2)
lines(x,(1/h)*v3,col="blue",lty=2)
legend(78,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

#####
# Figure - CDFS

par(mfrow=c(2,2))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs1.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 1")
n1 <- log(x/meanS1)+(varS1/(2*meanS1^2))-((varS1)^2)/(4*meanS1^4)
n2 <- (sdS1/meanS1)*sqrt(1-(varS1/(2*meanS1^2)))
Gs1 <- pnorm(n1/n2,0,1)
lines(x,Gs1,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs2.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 2")
Gs2 <- pnorm(((x/meanS2)^hS2-mY2)/sqrt(varY2),0,1)
lines(x,Gs2,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs3.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 3")
n1 <- log(x/meanS3)+(varS3/(2*meanS3^2))-((varS3)^2)/(4*meanS3^4)
n2 <- (sdS3/meanS3)*sqrt(1-(varS3/(2*meanS3^2)))
Gs3 <- pnorm(n1/n2,0,1)
lines(x,Gs3,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x,Gs4.rec(x),type="l",col="red",ylab="G(x)",main="Κατανομή 4",)
Gs4 <- pnorm(((x/meanS4)^hS4-mY4)/sqrt(varY4),0,1)
lines(x,Gs4,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"),lty=1:2,col=c("red","blue"))

```

#Προσέγγιση Wilson-Hilferty

```
library(actuar)
# Poisson
lam1=lam2=5
meanN1=meanN2=varN1=varN2=lam1
skewN1=skewN2=1/sqrt(lam1)

lam3=lam4=50
meanN3=meanN4=varN3=varN4=lam3
skewN3=skewN4=1/sqrt(lam3)

# Pareto
sh1=4; sc1=30
sh2=40; sc2=390
sh3=4; sc3=3
sh4=40; sc4=39

(meanX1=mpareto(1,sh1,sc1)); (varX1=mpareto(2,sh1,sc1)-
(mpareto(1,sh1,sc1))^2)
(meanX2=mpareto(1,sh2,sc2)); (varX2=mpareto(2,sh2,sc2)-
(mpareto(1,sh2,sc2))^2)
(meanX3=mpareto(1,sh3,sc3)); (varX3=mpareto(2,sh3,sc3)-
(mpareto(1,sh3,sc3))^2)
(meanX4=mpareto(1,sh4,sc4)); (varX4=mpareto(2,sh4,sc4)-
(mpareto(1,sh4,sc4))^2)

(skewX1=(2*(1+sh1)/(sh1-3))*sqrt((sh1-2)/sh1))
(skewX2=(2*(1+sh2)/(sh2-3))*sqrt((sh2-2)/sh2))
(skewX3=(2*(1+sh3)/(sh3-3))*sqrt((sh3-2)/sh3))
(skewX4=(2*(1+sh4)/(sh4-3))*sqrt((sh4-2)/sh4))

# Compound

(meanS1 <- meanN1*meanX1)
(varS1 <- meanN1*varX1+varN1*meanX1^2)
(sdS1 <- sqrt(varS1))
(skewS1 <-
(skewN1*varN1^(3/2)*meanX1^3+3*varN1*meanX1*varX1+meanN1*skewX1*varX1^(3/2))
/(varS1^(3/2)))

(meanS2 <- meanN2*meanX2)
(varS2 <- meanN2*varX2+varN2*meanX2^2)
(sdS2 <- sqrt(varS2))
(skewS2 <-
(skewN2*varN2^(3/2)*meanX2^3+3*varN2*meanX2*varX2+meanN2*skewX2*varX2^(3/2))
/(varS2^(3/2)))

(meanS3 <- meanN3*meanX3)
(varS3 <- meanN3*varX3+varN3*meanX3^2)
(sdS3 <- sqrt(varS3))
(skewS3 <-
(skewN3*varN3^(3/2)*meanX3^3+3*varN3*meanX3*varX3+meanN3*skewX3*varX3^(3/2))
/(varS3^(3/2)))

(meanS4 <- meanN4*meanX4)
(varS4 <- meanN4*varX4+varN4*meanX4^2)
(sdS4 <- sqrt(varS4))
(skewS4 <-
(skewN4*varN4^(3/2)*meanX4^3+3*varN4*meanX4*varX4+meanN4*skewX4*varX4^(3/2))
/(varS4^(3/2)))

#####
#Discretize Unbiased - Recursive
#####
au=0; bu=20000
h <- 0.01
```

```

fx1 <- discretize(ppareto(x, shape=sh1, scale=sc1), method = "unbiased", lev
= levpareto(x, shape=sh1, scale=sc1), from = au, to = bu, step = h)
fx2 <- discretize(ppareto(x, shape=sh2, scale=sc2), method = "unbiased", lev
= levpareto(x, shape=sh2, scale=sc2), from = au, to = bu, step = h)
fx3 <- discretize(ppareto(x, shape=sh3, scale=sc3), method = "unbiased", lev
= levpareto(x, shape=sh3, scale=sc3), from = au, to = bu, step = h)
fx4 <- discretize(ppareto(x, shape=sh4, scale=sc4), method = "unbiased", lev
= levpareto(x, shape=sh4, scale=sc4), from = au, to = bu, step = h)

Gs1.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx1, lambda = lam1, x.scale = h, maxit=5000000)
Gs2.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx2, lambda = lam1, x.scale = h, maxit=5000000)
Gs3.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx3, lambda = lam3, x.scale = h, maxit=5000000)
Gs4.rec <- aggregateDist("recursive", model.freq = "poisson", model.sev =
fx4, lambda = lam3, x.scale = h, maxit=5000000)

gs1.rec <- diff(Gs1.rec)
gs2.rec <- diff(Gs2.rec)
gs3.rec <- diff(Gs3.rec)
gs4.rec <- diff(Gs4.rec)
#####
# Wilson-Hilferty Example Distribution Function - Table 6
#####

# Example 1
(t1 <- meanS1+4*sqrt(varS1))
(c1 <- (skewS1/6)-(6/skewS1))
(c2 <- 3*((2/skewS1)^(2/3)))
(c3 <- 2/skewS1)
(tt <- (t1-meanS1)/sdS1)
(pa1 <- 1-pnorm(c1+c2*((tt+c3)^(1/3)),0,1))

# Example 2
(t2 <- meanS2+4*sqrt(varS2))
(c1 <- (skewS2/6)-(6/skewS2))
(c2 <- 3*((2/skewS2)^(2/3)))
(c3 <- 2/skewS2)
(tt <- (t2-meanS2)/sdS2)
(pa2 <- 1-pnorm(c1+c2*((tt+c3)^(1/3)),0,1))

# Example 3
(t3 <- meanS3+4*sqrt(varS3))
(c1 <- (skewS3/6)-(6/skewS3))
(c2 <- 3*((2/skewS3)^(2/3)))
(c3 <- 2/skewS3)
(tt <- (t3-meanS3)/sdS3)
(pa3 <- 1-pnorm(c1+c2*((tt+c3)^(1/3)),0,1))

# Example 4
(t4 <- meanS4+4*sqrt(varS4))
(c1 <- (skewS4/6)-(6/skewS4))
(c2 <- 3*((2/skewS4)^(2/3)))
(c3 <- 2/skewS4)
(tt <- (t4-meanS4)/sdS4)
(pa4 <- 1-pnorm(c1+c2*((tt+c3)^(1/3)),0,1))

#(Πίνακας 3.6)

# Table 5

(p1 <- 1-Gs1.rec(t1))
(p2 <- 1-Gs2.rec(t2))
(p3 <- 1-Gs3.rec(t3))
(p4 <- 1-Gs4.rec(t4))
t <- c(t1,t2,t3,t4)
pa <- c(pa1,pa2,pa3,pa4)

```

```

p <- c(p1,p2,p3,p4)
m <- cbind(t,p,pa)
rownames(m) <- c("Κατανομή 1 ", "Κατανομή 2 ", "Κατανομή 3 ", "Κατανομή 4 ")
colnames(m) <- c("x", " G(x):True Probability", " G(x):Wilson-Hilferty Approximation")
round(m, digits=8)

```

#(Σχήματα 3.10 και 3.11)

```

#####
# Figure 6

```

```

par(mfrow=c(2,2))

```

```

# Example 1
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs1.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.03), xlim=c(0,200), main="Κατανομή 1")
segments(0,0,0,gs1.rec[1], col="red")
x <- seq(0,lim,h)
c1 <- (skewS1/6)-(6/skewS1)
c2 <- 3*((2/skewS1)^(2/3))
c3 <- 2/skewS1
y <- (x-meanS1)/sdS1
Gs1 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
v2 <- diff(Gs1)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(160,0.018,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

```

```

# Example 2
lim <- 200
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs2.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.03), xlim=c(0,200), main="Κατανομή 2")
segments(0,0,0,gs2.rec[1], col="red")
x <- seq(0,lim,h)
c1 <- (skewS2/6)-(6/skewS2)
c2 <- 3*((2/skewS2)^(2/3))
c3 <- 2/skewS2
y <- (x-meanS2)/sdS2
Gs2 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
v2 <- diff(Gs2)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(160,0.018,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

```

```

# Example 3
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs3.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 3")
segments(0,0,0,gs3.rec[1], col="red")
x <- seq(0,lim,h)
c1 <- (skewS3/6)-(6/skewS3)
c2 <- 3*((2/skewS3)^(2/3))
c3 <- 2/skewS3
y <- (x-meanS3)/sdS3
Gs3 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
v2 <- diff(Gs3)
v3 <- c(NA,v2)

```

```

lines(x, (1/h)*v3,col="blue",lty=2)
legend(78,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

# Example 4
lim <- 100
up <- lim*(1/h)+1
x <- seq(h,lim,h)
plot(x, (1/h)*gs4.rec[2:up], type="l", col="red", ylab="g(x)",
ylim=c(0,0.05), main="Κατανομή 4")
segments(0,0,0,gs4.rec[1], col="red")
x <- seq(0,lim,h)
c1 <- (skewS4/6)-(6/skewS4)
c2 <- 3*((2/skewS4)^(2/3))
c3 <- 2/skewS4
y <- (x-meanS4)/sdS4
Gs4 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
v2 <- diff(Gs4)
v3 <- c(NA,v2)
lines(x, (1/h)*v3,col="blue",lty=2)
legend(78,0.045,c("Actual pdf","Estimated pdf"),
lty=1:2,col=c("red","blue"))

#####
# Figure - CDFS

par(mfrow=c(2,2))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs1.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 1")
c1 <- (skewS2/6)-(6/skewS2)
c2 <- 3*((2/skewS2)^(2/3))
c3 <- 2/skewS2
y <- (x-meanS2)/sdS2
Gs1 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
lines(x,Gs1,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 200
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs2.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 2")
c1 <- (skewS2/6)-(6/skewS2)
c2 <- 3*((2/skewS2)^(2/3))
c3 <- 2/skewS2
y <- (x-meanS2)/sdS2
Gs2 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
lines(x,Gs2,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs3.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 3")
c1 <- (skewS3/6)-(6/skewS3)
c2 <- 3*((2/skewS3)^(2/3))
c3 <- 2/skewS3
y <- (x-meanS3)/sdS3
Gs3 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
lines(x,Gs3,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

lim <- 100
up <- lim*(1/h)+1
x <- seq(0,lim,h)
plot(x, Gs4.rec(x), type="l", col="red", ylab="G(x)", main="Κατανομή 4,")

```

```

c1 <- (skewS4/6) - (6/skewS4)
c2 <- 3 * ((2/skewS4)^(2/3))
c3 <- 2/skewS4
y <- (x-meanS4)/sdS4
Gs4 <- pnorm(c1+c2*((y+c3)^(1/3)),0,1)
lines(x,Gs4,col="blue",lty=2)
legend(0,1,c("Actual cdf","Estimated cdf"), lty=1:2,col=c("red","blue"))

```

Κώδικας για την εφαρμογή 3.2

```

lambda=100
theta=1
draws=10000
xx=c(50:151)
b=0.6582217
A=-4.19031
B=6.30621
C=-6.551706
r2=0.02
r3=0.0005311495
r4=0.001224484
h=0.505
a=50
F=Fgamma=Fwilson=Fclt=Fnpa=Fhaldane=Fbowers=c(50:151)
library(pracma)
for(i in 1:102){
d=-79822.24+2673.764*((xx[i]-100)/sqrt(200)+26607.41)^(1/3)
for(j in 1:151){}
Fclt[i]=pnorm((xx[i]-100)/sqrt(200),0,1)
Fwilson[i]=pnorm(d,0,1)
d=0.5*xx[i]
Fbowers[i]=pgamma(d,a,1)-A*exp(-d)*(d^a/gamma(a+1)-
2*d^(a+1)/gamma(a+2)+d^(a+2)/gamma(a+3))
+B*exp(-d)*(d^a/gamma(a+1)-3*d^(a+1)/gamma(a+2)+3*d^(a+2)/gamma(a+3)-
d^(a+3)/gamma(a+4))
+C*exp(-d)*(d^a/gamma(a+1)-4*d^(a+1)/gamma(a+2)+6*d^(a+2)/gamma(a+3)-
4*d^(a+3)/gamma(a+4)+d^(a+4)/gamma(a+5))
Fgamma[i]=pgamma(xx[i]+34.68701,90.70295,1/1.484924)
d=-3/0.21+sqrt(1+9/0.21^2+6*(xx[i]-100)/(0.21*sqrt(200)))
Fnpa[i]=pnorm(d,0,1)
d=((xx[i]/100)^h-0.9974)/0.0715
Fhaldane[i]=pnorm(d,0,1)

F[i]=0}
S=c(1:draws)
m4=0
for(i in 1:draws){
N=rpois(1,lambda)
x=rexp(N,theta)
S[i]=sum(x)
m4=m4+(S[i]-100)^4
for(j in 1:102){
if(S[i]<=xx[j]){F[j]=F[j]+1}}
}
F=F/draws
y1 <- F
y2 <- Fgamma
y3<-Fwilson
y4<-Fclt
y5<-Fnpa
y6<-Fhaldane
y7<-Fbowers

### (Σχήμα 3.12)

```



```

plot(xx,y1,type="l",col="red",main="Distribution and approximations of S",
xlab="s", ylab="F(s)")
lines(xx,y2,col="green", lty=2)
lines(xx,y3,col="blue",lty=3)
lines(xx,y4,col="darkmagenta",lty=4)
lines(xx,y5,col="chocolate3",lty=5)
lines(xx,y6,col="cyan",lty=6)
lines(xx,y7,col="lightcoral",lty=7)

legend("bottomright", legend=c("real distribution of S","TGA","Wilson-
Hilferty","NA","NPA","Haldane","Bowers"),
      col=c("red","green","blue", "darkmagenta",
"chocolate3","cyan","lightcoral"), lty=1:7, cex=0.7)

###error calculation (Πίνακας 3.7)

ebowers=abs((F[50:102]-Fbowers[50:102])/F[50:102])
enpa=abs((F[50:102]-Fnpa[50:102])/F[50:102])
egamma=abs((F[50:102]-Fgamma[50:102])/F[50:102])
ewilson=abs((F[50:102]-Fwilson[50:102])/F[50:102])
eclt=abs((F[50:102]-Fcilt[50:102])/F[50:102])
ehaldane=abs((F[50:102]-Fhaldane[50:102])/F[50:102])
error_bowers=mean(ebowers)
error_npa=mean(enpa)
error_gamma=mean(egamma)
error_wilson=mean(ewilson)
error_cilt=mean(ecilt)
error_haldane=mean(ehaldane)
m <-
rbind(error_bowers,error_npa,error_gamma,error_wilson,error_cilt,error_haldan
e)
round(m, digits=6)

```

Κώδικας για την εφαρμογή 3.3

```

k=2
r=4
p=0.7
m4=0
lambda=1
draws=10000
xx=seq(0,10,0.01)
xx
n <- length(xx)
h=0.5442966
a=1
b=0.6623802
A=-0.1033005
B=0.1640024
C=-0.189846

F=Fgamma=Fwilson=Fcilt =Fhaldane=Fbowers=Fnpa=c(1:n)
for(i in 1:n){
d=-4.222906+3.9*((xx[i]- 1.519)/ 1.513891+1.48258)^(1/3)
Fcilt[i]=pnorm((xx[i]-1.519)/1.513891,0,1)
Fwilson[i]=pnorm(d,0,1)
Fgamma[i]=pgamma(xx[i]+0.7252236,2.198043,1/1.020927)
d=b*xx[i]
Fbowers[i]=pgamma(xx[i],a,1)-A*exp(-d)*(d^a/gamma(a+1)-
2*d^(a+1)/gamma(a+2)+d^(a+2)/gamma(a+3))
+B*exp(-d)*(d^a/gamma(a+1)-3*d^(a+1)/gamma(a+2)+3*d^(a+2)/gamma(a+3)-
d^(a+3)/gamma(a+4))
+C*exp(-d)*(d^a/gamma(a+1)-4*d^(a+1)/gamma(a+2)+6*d^(a+2)/gamma(a+3)-
4*d^(a+3)/gamma(a+4)+d^(a+4)/gamma(a+5))
d=((xx[i]/1.519)^h-0.8484476)/0.585023

```

```

Fhaldane[i]=pnorm(d,0,1)
d=-3/1.349+sqrt(1+9/1.349^2+6*(xx[i]-1.519)/(1.349*sqrt(2.291)))
Fnpa[i]=pnorm(d,0,1)
F[i]=0}

```

```

S=c(1:draws)

```

```

for(i in 1:draws){
N=rrnbinom(1,r,p)
x=rweibull(N,k,lambda)
S[i]=sum(x)
for(j in 1:n){
if(S[i]<=xx[j]){F[j]=F[j]+1}}
}
F=F/draws
y1 <-F
y2 <-Fgamma
y3<-Fwilson
y4<-Fclt
y5<-Fnpa
y6<-Fhaldane
y7<-Fbowers

```

(Σχήμα 3.15)

```

plot(xx,y1,type="l",col="red",main="Distribution and approximations of S",
xlab="s", ylab="F(s)")
lines(xx,y2,col="green", lty=2)
lines(xx,y3,col="blue",lty=3)
lines(xx,y4,col="darkmagenta",lty=4)
lines(xx,y5,col="chocolate3",lty=5)
lines(xx,y6,col="cyan",lty=6)
lines(xx,y7,col="lightcoral",lty=7)
legend("bottomright", legend=c("real distribution of S","TGA","Wilson-
Hilferty","NA","NPA","Haldane","Bowers"),col=c("red","green","blue",
"darkmagenta", "chocolate3","cyan","lightcoral"), lty=1:7, cex=0.7)

```

###error calculation (Πίνακας 3.8)

```

ebowers=abs((F[1:n]-Fbowers[1:n])/F[1:n])
egamma=abs((F[1:n]-Fgamma[1:n])/F[1:n])
ewilson=abs((F[1:n]-Fwilson[1:n])/F[1:n])
ehaldane=abs((F[1:n]-Fhaldane[1:n])/F[1:n])
eclt=abs((F[1:n]-Fclt[1:n])/F[1:n])
enpa=abs((F[1:n]-Fnpa[1:n])/F[1:n])
error_bowers=mean(ebowers)
error_gamma=mean(egamma)
error_wilson=mean(ewilson)
error_haldane=mean(ehaldane)
error_clt=mean(eclt)
error_npa=mean(enpa)

m <-
rbind(error_bowers,error_npa,error_gamma,error_wilson,error_clt,error_haldan
e)
round(m, digits=6)

```

Βιβλιογραφία

Ελληνική

Αντζουλάκος Δ. (2017), Πανεπιστημιακές Σημειώσεις Π.Μ.Σ. «Αναλογιστική Επιστήμη, Διοικητική Κινδύνου» Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης Πανεπιστήμιο Πειραιώς Στο Μάθημα Γλώσσα Προγραμματισμού R Με Εφαρμογές Στον Αναλογισμό, Πειραιάς.

Κούτρας Μ. (2004), Εισαγωγή στις πιθανότητες. Σταμούλη Α.Ε. Πειραιάς.

Πολίτης Κ. (2012), Εισαγωγή Στη Θεωρία Συλλογικού Κινδύνου - Το Συλλογικό Πρότυπο Και Θεωρία Χρεοκοπίας, Εκδόσεις Σταμούλης, Αθήνα.

Χατζηκωνσταντινίδης Ε. (2018), Πανεπιστημιακές Σημειώσεις Π.Μ.Σ. «Αναλογιστική Επιστήμη & Διοικητική Κινδύνου» Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης Πανεπιστήμιο Πειραιώς Στο Μάθημα Θεωρία Κινδύνου I, Πειραιάς.

Ξένη

BCBS - Basel Committee on Banking Supervision. (2011), Basel III counterparty credit risk frequently asked questions. Basel: Bank for International Settlements.

Beard, R.E., Pentikäinen, T., and Pesonen E. (1984), Risk Theory: The Stochastic Basis of Insurance. 3rd ed. London: Chapman and Hall.

Bowers, N.L. (1966), Expansion of probability density functions as a sum of gamma densities with applications in risk theory, Transactions of the Society of Actuaries XVIII, 125–139.

Haldane, J.B.S. (1938), The approximate normalization of a class of frequency distributions, *Biometrika* 29, 392–404.

Hanspeter Schmidli, (2017), *Risk Theory*, Springer, Cham.

Kouassi, E., Akpata, E., Pokou, K. (2020), A Note on Laplace Transforms of Some Common Distributions Used in Counting Processes Analysis. *Applied Mathematics*, 11, 67-75.

Pentikäinen, T. (1987), Approximative Evaluation of the Distribution Function of Aggregate Claims. *ASTIN Bulletin*, 17(1), 15-39.

Schmidt K. (1996), *Lectures on Risk Theory*. B. G. Teubner Stuttgart.

Seri, R., Choirat, C. (2015), COMPARISON OF APPROXIMATIONS FOR COMPOUND POISSON PROCESSES. *ASTIN Bulletin*, 45(3), 601-637.

Shevchenko, Pavel. (2010), Calculation of aggregate loss distributions. *Journal of Operational Risk*. 5, 3-40.

Sundt B. (2002), Review: Recursive evaluation of aggregate claims distributions. *Insurance: Mathematics and Economics*, 30, 297-322.

Sundt B., Jewell W. (1981), Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, 12, 27-39.

Teugels J., Sundt B. (2004), *Encyclopedia of actuarial science*. Wiley

Wilson, E.B., Hilferty, M. (1931), The distribution of chi-square, *Proceedings of the National Academy of Science* 17, 684–688.