ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ
& ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ
ACTUARIAL+RISK MANAGEMENT

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ &
ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

με θέμα

«Προσεγγίσεις Κατανομών Πιθανότητας από μία Μείξη Εκθετικών Κατανομών»

Μιχαλάκης Κωνσταντίνος

ΜΑΕ/11023

Επιβλέπων Καθηγητής:

Πολίτης Κωνσταντίνος (Αναπληρωτής Καθηγητής)

Αθήνα, 2015

Approximations of probability distributions by mixtures of exponentials

By
Michalakis Konstantinos

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Actuarial Science and Risk Management

UNIVERSITY OF PIRAEUS
Department of Statistics and Insurance Science

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον κ. Πολίτη Κωνσταντίνο, επιβλέποντα καθηγητή μου, για την ουσιαστική βοήθεια και τις πολύτιμες συμβουλές που μου παρείχε καθ'όλη τη διάρκεια της εκπόνησης της παρούσας διπλωματικής εργασίας.

Ευχαριστώ επίσης τα υπόλοιπα μέλη της τριμελούς συμβουλευτικής επιτροπής, τον κ. Τζαβελά Γεώργιο και τον κ. Ψαρράκο Γεώργιο.

ΠΕΡΙΛΗΨΗ

Στην Αναλογιστική Επιστήμη, και πιο συγκεκριμένα στη Θεωρία Χρεοκοπίας, ιδιαίτερο ενδιαφέρον παρουσιάζει ο υπολογισμός της Πιθανότητας Χρεοκοπίας (η εύρεση της πιθανότητας, το Πλεόνασμα που διαθέτει μία εταιρεία, να πάρει αρνητική τιμή). Ανάμεσα στις Κατανομές Αποζημιώσεων που παρουσιάζουν κύριο ενδιαφέρον, ανήκουν και οι κατανομές οι οποίες διαθέτουν βαριά ουρά. Παρ'ολα αυτά ο υπολογισμός της Πιθανότητας Χρεοκοπίας για τις παραπάνω κατανομές παρουσιάζει δυσκολίες. Μία περίπτωση στην οποία η συγκεκριμένη Πιθανότητα υπολογίζεται εύκολα, είναι όταν η Κατανομή των Αποζημιώσεων ακολουθεί μία Μείξη Εκθετικών Κατανομών. Στην παρούσα εργασία θα ασχοληθούμε κυρίως με προσεγγίσεις Κατανομών Πιθανότητας από μία Μείξη Εκθετικών Κατανομών, έτσι ώστε να μπορούμε να υπολογίζουμε προσεγγιστικά την Πιθανότητα Χρεοκοπίας διαφόρων συνεχών τυχαίων μεταβλητών, που ακολουθούν κατανομές με βαριά ουρά.

ABSTRACT

In actuarial science and more specifically in ruin theory, it is important to calculate the ruin probability (the calculation of the probability company's surplus gets a negative value). Long tail distributions are among the most studied claim amount distributions. Nevertheless, the calculation of ruin probability in these distributions faces difficulties. A case where this probability can be easily calculated is when the claim amount distribution is a mixture of exponential distributions. In the current dissertation we will mainly focus on the approximations of probability distributions by mixtures of exponentials. This will allow us to calculate approximately the ruin probability of various continuous random variables which follow distributions with long tail.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

ΚΕΦΑΛΑΙΟ 2: ΚΑΤΑΝΟΜΕΣ ΠΙΘΑΝΟΤΗΤΑΣ

ΚΕΦΑΛΑΙΟ 3: ΠΡΟΣΕΓΓΙΣΕΙΣ ΚΑΤΑΝΟΜΩΝ ΠΙΘΑΝΟΤΗΤΑΣ ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

Ιδιαίτερο ενδιαφέρον στη θεωρία χρεοκοπίας παρουσιάζει η εύρεση της πιθανότητας χρεοκοπίας, της πιθανότητας δηλαδή όπου η στοχαστική ανέλιξη του πλεονάσματος θα πάρει τελικά αρνητική τιμή. Στον Αναλογισμό μελετείται κυρίως ο υπολογισμός της πιθανότητας χρεοκοπίας σε κατανομές αποζημιώσεων με βαριά ουρά, όπως είναι οι κατανομές Pareto, Lognormal και Weibull. Στην περίπτωση της κατανομής Weibull, θα διαπιστώσουμε κατά την διάρκεια της εργασίας ότι διαθέτει βαριά ουρά, όταν η δεύτερη παράμετρός της έχει τιμές μικρότερες της μονάδας. Παρότι οι συγκεκριμένες κατανομές εξετάζονται ιδιαίτερα, η πιθανότητα χρεοκοπίας τους δεν είναι εύκολο να υπολογιστεί. Οι περιπτώσεις όπου η πιθανότητα χρεοκοπίας υπολογίζεται χωρίς ιδιαίτερη δυσκολία, είναι όταν η κατανομή των αποζημιώσεων ακολουθεί είτε την Εκθετική κατανομή, είτε μία μείξη εκθετικών κατανομών ή μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο να ισούται με 2.

Στην παρούσα εργασία θα ασχοληθούμε με τις προσεγγίσεις κατανομών πιθανότητας από μία μείξη εκθετικών κατανομών. Αυτή η προσέγγιση θα διαπιστώσουμε ότι είναι ιδιαίτερα χρήσιμη, διότι μπορούμε να υπολογίσουμε προσεγγιστικά την πιθανότητα χρεοκοπίας διαφόρων συνεχών τυχαίων μεταβλητών, οι οποίες δεν ακολουθούν τις κατανομές που προαναφέραμε.

Πιο αναλυτικά η δομή της εργασίας θα είναι η εξής:

Στο δεύτερο κεφάλαιο, θα παρουσιαστούν αναλυτικά οι βασικές έννοιες της θεωρίας πιθανοτήτων και οι σημαντικότερες κατανομές πιθανότητας. Μετά από μία σύντομη εισαγωγή για κάθε κατανομή πιθανότητας, θα παρουσιάσουμε κάποια παραδείγματα για την εύρεση των βασικών τους ποσοτήτων, όπως η μέση τιμή και οι ροπές $\kappa -$τάξης, τις οποίες θα χρησιμοποιούμε στα επόμενα κεφάλαια.

Στο τρίτο κεφάλαιο θα προσεγγίσουμε τις συναρτήσεις πυκνότητας πιθανότητας τριών συνεχών τυχαίων μεταβλητών, οι οποίες ακολουθούν τις κατανομές Pareto, Weibull και Lognormal από μία μείξη δύο εκθετικών κατανομών. Για να το επιτύχουμε αυτό θα χρησιμοποιήσουμε την μέθοδο των ροπών. Έπειτα θα αξιολογήσουμε τις παραπάνω προσεγγίσεις με την χρήση συγκεκριμένων μεθόδων που θα αναλυθούν κατά την διάρκεια του κεφαλαίου. Τέλος, θα προσεγγίσουμε τις ίδιες συναρτήσεις πυκνότητας πιθανότητας από μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο να ισούται με 2 και με τις μεθόδους που χρησιμοποιούμε, θα διαπιστώσουμε για κάθε μία περίπτωση ξεχωριστά ποια από τις δύο προσεγγίσεις είναι περισσότερο ικανοποιητική.

Στο τέταρτο κεφάλαιο θα αναλύσουμε τις σύνθετες κατανομές και θα περιγράψουμε τις βασικές έννοιες της θεωρίας χρεοκοπίας, όπως η πιθανότητα χρεοκοπίας, το περιθώριο ασφαλείας και το πλεόνασμα. Στην συνέχεια θα παρουσιάσουμε την έννοια της μέγιστης σωρευτικής απώλειας, η οποία όπως θα διαπιστώσουμε συνδέεται με την πιθανότητα μη χρεοκοπίας.

Στο τελευταίο κεφάλαιο θα παρουσιάσουμε δύο συναφείς μεθόδους για τον υπολογισμό της πιθανότητας χρεοκοπίας, όταν η κατανομή των αποζημιώσεων ακολουθεί μία μείξη εκθετικών κατανομών και μία μέθοδο η οποία υπολογίζει την πιθανότητα χρεοκοπίας όταν η κατανομή των αποζημιώσεων ακολουθεί μία μείξη Γάμμα κατανομών, με την πρώτη παράμετρο να είναι ίση με 2. Τέλος θα υπολογίσουμε προσεγγιστικά την πιθανότητα χρεοκοπίας των συναρτήσεων πυκνότητας πιθανότητας των κατανομών Pareto, Weibull και Lognormal, που θα προσεγγίσουμε στο τρίτο κεφάλαιο από μία μείξη εκθετικών κατανομών, έχοντας κατανομές των αποζημιώσεων τις συγκεκριμένες μείξεις εκθετικών κατανομών.

ΚΕΦΑΛΑΙΟ 2: ΚΑΤΑΝΟΜΕΣ ΠΙΘΑΝΟΤΗΤΑΣ.

Σε αυτό το κεφάλαιο θα ασχοληθούμε με τις κυριότερες κατανομές πιθανότητας. Θα αναφέρουμε βασικές έννοιες της θεωρίας πιθανοτήτων όπως η συνάρτηση κατανομής, η συνάρτηση πιθανότητας, η μέση τιμή, η διασπορά και η ροπογεννήτρια συνάρτηση. Στην συνέχεια θα δούμε πιο αναλυτικά κάθε μία από τις βασικές κατανομές πιθανότητας, όπως ειναι η εκθετική, η Γάμμα, η γεωμετρική,η λογαριθμοκανονική, η Pareto, η Weibull κ.ά και θα δούμε κάποια παραδείγματα για την πλήρη κατανόηση των βασικών εννοιών και κάποια γραφήματα που θα απεικονίζουν την γραφική παράσταση των συγκεκριμένων κατανομών. Οι πηγές από τις οποίες συλλέξαμε πληροφορίες είναι: Οι σημειώσεις του Πολίτη (2015), το βιβλίο του Πολίτη (2012), το βιβλίο του Κούτρα (2005) και το βιβλίο του Bean (2001).

2.1: ΜΙΑ ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΚΑΤΑΝΟΜΕΣ.

Στην θεωρία πιθανοτήτων μελετώνται οι κατανομές πιθανότητας. Αυτές χωρίζονται σε διακριτές και συνεχείς. Οι διακριτές συνήθως εμφανίζονται όταν η τυχαία μεταβλητή Χ παίρνει τιμές στο σύνολο των φυσικών αριθμών, ισχύει δηλαδή ότι:

$X \in \{1,2, \dots\}$,

ενώ οι συνεχείς εμφανίζονται όταν η τυχαία μεταβλητή Χ μπορεί να πάρει οποιαδήποτε πραγματική τιμή (συνήθως θετική).

Για παράδειγμα σε ένα πείραμα θέλουμε να μελετήσουμε πόσες φορές θα φέρουμε έξι σε κ ρίψεις ενός ζαριού. Ο αριθμός κ θα είναι φυσικός. Επομένως θα χρησιμοποιήσουμε κάποια διακριτή κατανομή.

Αν σε ένα άλλο πείραμα εξετάζουμε ποια χρονική στιγμή θα καεί μία λάμπα, η χρονική στιγμή t που θα συμβεί το ενδεχόμενο που εξετάζουμε μπορεί να πάρει οποιαδήποτε θετική πραγματική τιμή, οπότε και θα χρησιμοποιήσουμε κάποια συνεχή κατανομή.

Κάποιες διακριτές κατανομές, μερικές από τις οποίες θα ασχοληθούμε και στην συνέχεια είναι οι: Poisson, Bernoulli, γεωμετρική και η διωνυμική.

Κάποιες συνεχείς κατανομές είναι οι: Εκθετική, Γάμμα, Κανονική, Weibull, Pareto, λογαριθμοκανονική κ.ά.

Παρακάτω θα παρουσιάσουμε τις βασικές ιδιότητες των συνεχών κατανομών.

Όταν η συνεχής τυχαία μεταβλητή X ακολουθεί μία συνεχή κατανομή, τότε οι: $F_X(x)$, $f_X(x)$ ονομάζονται συνάρτηση κατανομής και συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X αντίστοιχα και ορίζονται από τις σχέσεις:

$$F_X(x) = P(X \leq x), \; F_X(x) = \int_{-\infty}^{x} f_X(\xi) d\xi.$$

Από τον ορισμό της η συνάρτηση κατανομής έχει τις παρακάτω ιδιότητες:

$i)$ Για συνεχή τυχαία μεταβλητή X η $F_X(x)$, είναι σε όλο το πεδίο ορισμού της συνεχής.

$ii)$Με παραγώγιση της συνάρτησης κατανομής ως προς x προκύπτει η συνάρτηση πυκνότητας πιθανότητας, ισχύει συγκεκριμένα ότι:

$$\frac{d}{dx} F_X(x) = f_X(x).$$

$iii)$Η $F_X(x)$, παίρνει τιμές απο 0 έως 1, δηλαδή:

$$0 \leq F_X(x) \leq 1,$$

$iv)$ Ισχύει ότι:

$$\lim_{x \to -\infty} F_X(x) = 0 \text{ και } \quad \lim_{x \to \infty} F_X(x) = 1.$$

Επίσης για την $f_X(x)$ ισχύει:

$$\int_{-\infty}^{\infty} f_X(x) dx = 1.$$

Η παραπάνω σχέση καλείται συνθήκη κανονικοποίησης και ισχύει για οποιαδήποτε συνάρτηση πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής X.

Κάποιες βασικές παράμετροι των συνεχών κατανομών είναι οι παρακάτω:

Μέση τιμή: Ονομάζεται η ποσότητα:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Η ροπή κ-τάξης: Υπολογίζεται από την σχέση:

$$E(X^{\kappa}) = \int_{-\infty}^{\infty} x^{\kappa} f_X(x) dx.$$

Από την τελευταία σχέση παρατηρούμε ότι η μέση τιμή είναι στην ουσία η ροπή πρώτης τάξης.

Διασπορά (ή διακύμανση): Είναι η ποσότητα:

$$Var(X) = E\big(X - E(X)\big)^2 = E(X^2) - E^2(X).$$

Οι βασικές ιδιότητες των διακριτών κατανομών είναι οι εξής:

Όταν η τυχαία μεταβλητή X ακολουθεί μία διακριτή κατανομή, η $f_X(x)$ ονομάζεται συνάρτηση πιθανότητας. Γενικότερα μία τυχαία μεταβλητή μπορεί να πάρει τόσο θετικές όσο και αρνητικές τιμές. Στην παρούσα όμως εργασία θα ασχοληθούμε με μεταβλητές που παίρνουν τιμές στο σύνολο των μη αρνητικών ακεραίων $\{0,1,2,..\}$.

Η $f_X(x)$ αποτελεί συνάρτηση πιθανότητας όταν ισχύουν οι ιδιότητες:

$$i) f_X(x) \geq 0,$$

$$ii) \sum_{x=0}^{\infty} f_X(x) = 1.$$

Ισχύει δηλαδή η συνθήκη κανονικοποίησης όπως και στις συνεχείς κατανομές.

$iii)$ Ορίζεται η συνάρτηση κατανομής της τυχαίας μεταβλητής X:

$$F_X(x) = P(X \leq x).$$

Η μέση τιμή, η ροπή κ-τάξης και η διασπορά υπολογίζονται από τις παρακάτω σχέσεις:

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

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

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

Μία άλλη παράμετρος με κύριο ενδιαφέρον τόσο στις συνεχείς όσο και στις διακριτές κατανομές είναι η ροπογεννήτρια συνάρτηση, η οποία συμβολίζεται ως $M_X(t)$.

Αν η X είναι μία τυχαία μεταβλητή (συνεχής ή διακριτή) η $M_X(t)$ ορίζεται από την σχέση:

$$M_X(t) = E(e^{tX}),$$

είναι δηλαδή η μέση τιμή της ποσότητας e^{tX}.

- Για συνεχή τυχαία μεταβλητή X η $M_X(t)$ είναι:
  $M_X(t) = \int_{-\infty}^{\infty} e^{tX} f_X(x) dx,$
- Για διακριτή τυχαία μεταβλητή η $M_X(t)$ είναι:
  $M_X(t) = \sum_{x=0}^{\infty} e^{tX} f_X(x).$

Η $M_X(t)$ καλείται ροπογεννήτρια διότι η μέση τιμή της X^κ (ροπή κ-τάξης) ισούται με την ν-οστή παράγωγο της ροπογεννήτριας συνάρτησης $M_X(t)$ για $t = 0$. Ισχύει δηλαδή:

$$E(X^\kappa) = [\frac{d^\kappa M_X(t)}{dx}]_{t=0}.$$

Στην συνέχεια θα παρουσιάσουμε κάθε μία απο τις βασικές κατανομές, συνεχείς και διακριτές με τις οποίες θα ασχοληθούμε και θα χρησιμοποιήσουμε περισσότερο στην παρούσα εργασία.

2.2: Η ΚΑΤΑΝΟΜΗ POISSON.

Η διακριτή τυχαία μεταβλητή X ακολουθεί την κατανομή Poisson όταν η συνάρτηση πιθανότητας $f_X(x)$ είναι η:

$$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}\ (\lambda > 0, x = 0,1,2 \dots).$$

Την κατανομή Poisson ακολουθεί π.χ. ο αριθμός εμφανίσεων ενός γεγονότος σε ορισμένο χρονικό διάστημα και η παράμετρος λ είναι ανάλογη του μήκους του διαστήματος του γεγονότος αυτού.

Για να υπολογίσουμε την μέση τιμή της κατανομής Poisson εφαρμόζουμε την σχέση:

$$E(X) = \sum_{x=0}^{\infty} x f_X(x) = \sum_{x=0}^{\infty} x e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty} x\frac{\lambda^x}{x!} =$$

$$= e^{-\lambda}\sum_{\kappa=0}^{\infty} \kappa\frac{\lambda^k}{\kappa!}\ (\gamma\iota\alpha\ x = \kappa).$$

Από τον Απειροστικό Λογισμό γνωρίζουμε ότι:

$$e^x = \sum_{k=0}^{\infty}\frac{x^\kappa}{\kappa!} \Rightarrow \frac{\partial}{\partial x}e^x = \frac{\partial}{\partial x}\sum_{\kappa=0}^{\infty}\frac{x^\kappa}{\kappa!} \Rightarrow e^x = \sum_{\kappa=0}^{\infty}\frac{\kappa x^{\kappa-1}}{\kappa!}\ .$$

Αν πολλαπλασιάσουμε την τελευταία σχέση με x θα έχουμε:

$$\Rightarrow e^x x = \sum_{\kappa=0}^{\infty}\frac{\kappa x^\kappa}{\kappa!}.$$

Για $x = \lambda$ έχουμε:

$$e^\lambda \lambda = \sum_{\kappa=0}^{\infty}\frac{\kappa\lambda^\kappa}{\kappa!} \Rightarrow E(X) = e^{-\lambda}e^\lambda\lambda \Rightarrow E(X) = \lambda.$$

Για να βρούμε την ροπή $2^{\eta\varsigma}$ τάξης $E(X^2)$ θα υπολογίσουμε πρώτα την μέση τιμή της $X(X-1)$, διότι ισχύει:

$$E\big(X(X-1)\big) = E(X^2 - X) = E(X^2) - E(X),$$

και την μέση τιμή την έχουμε υπολογίσει παραπάνω.

Έχουμε:

$$E\big(X(X-1)\big) = \sum_{x=0}^{\infty} x(x-1)f(x) = \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} =$$

$$= \sum_{\kappa=0}^{\infty} \kappa(\kappa-1)\frac{\lambda^\kappa}{\kappa!}\ (\gamma\iota\alpha\ \kappa = x) = e^{-\lambda}\sum_{\kappa=0}^{\infty} \kappa(\kappa-1)\frac{\lambda^\kappa}{\kappa!}$$

Είδαμε προηγουμένως ότι:

$$e^x = \sum_{\kappa=0}^{\infty} \frac{x^\kappa}{\kappa!}$$

Με διπλή παραγώγιση αυτής της σχέσης ως προς x έχουμε:

$$e^x = \sum_{\kappa=2}^{\infty} \kappa(\kappa-1)\frac{x^{\kappa-2}}{\kappa!}$$

Πολλαπλασιάζουμε με x^2 και θα ισχύει:

$$x^2 e^x = x^2 \sum_{\kappa=2}^{\infty} \frac{\kappa(\kappa-1)x^{\kappa-2}}{\kappa!} = \sum_{\kappa=2}^{\infty} \frac{\kappa(\kappa-1)x^{\kappa}}{\kappa!}$$

Θέτουμε $x = \lambda$ και προκύπτει:

$$\lambda^2 e^\lambda = \sum_{\kappa=2}^{\infty} \frac{\kappa(\kappa-1)\lambda^{\kappa}}{\kappa!}$$

Επομένως ισχύει ότι:

$$E\big(X(X-1)\big) = e^{-\lambda}\lambda^2 e^\lambda = \lambda^2 \Rightarrow$$

$$\Rightarrow E(X^2 - X) = \lambda^2 \Rightarrow E(X^2) - E(X) = \lambda^2 \Rightarrow E(X^2) = \lambda^2 + \lambda.$$

$$\Rightarrow Var(X) = E(X^2) - E^2(X) = \lambda^2 + \lambda - \lambda \Rightarrow Var(X) = \lambda.$$

2.3: Η ΓΕΩΜΕΤΡΙΚΗ ΚΑΤΑΝΟΜΗ.

Η Γεωμετρική κατανομή παρουσιάζεται σε πειράματα στα οποία μελετάμε τον αριθμό δοκιμών έως ότου επέλθει η πρώτη επιτυχία. Θεωρούμε $X = x$, όταν μέχρι την $x-1$ δοκιμή έχουμε αποτυχία και στην $x-οστή$ δοκιμή έχουμε επιτυχία. Συμβολίζουμε με p την πιθανότητα επιτυχίας και με q την πιθανότητα αποτυχίας (προφανώς θα ισχύει: $q = 1 - p$).

Αν η τυχαία μεταβλητή ακολουθεί την Γεωμετρική κατανομή με παράμετρο $p\big(X \sim Ge(p)\big)$, τότε η συνάρτηση πιθανότητας της X θα είναι:

$$f_X(x) = (1-p)^{x-1}p \ \ (x = 0,1,2,\dots).$$

Για την μέση τιμή $E(X)$ θα ισχύει:

$$E(X) = \sum_{x=0}^{\infty} x f_X(x) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p =$$

$$= p\sum_{x=1}^{\infty} x(1-p)^{x-1} = p\sum_{\kappa=1}^{\infty} \kappa(1-p)^{\kappa-1} \ (\gamma\iota\alpha\ \kappa = x).$$

Από τον Απειροστικό Λογισμό γνωρίζουμε ότι:

$$\sum_{\kappa=0}^{\infty} x^{\kappa} = \frac{1}{1-x} \Rightarrow \frac{d}{dx}\sum_{\kappa=0}^{\infty} x^{\kappa} = \frac{d}{dx}\frac{1}{1-x} \Rightarrow$$

$$\Rightarrow \sum_{\kappa=1}^{\infty} \kappa x^{\kappa-1} = \frac{1}{(1-x)^2} \ .$$

Επομένως για την μέση τιμή θα ισχύει:

$$E(X) = p\frac{1}{[1-(1-p)]^2} \Rightarrow E(X) = \frac{1}{p}.$$

Θα υπολογίσουμε την ροπή $2^{\eta\varsigma}$ τάξης μέσω της $E\big(X(X-1)\big)$. Έχουμε:

$$E\big(X(X-1)\big) = \sum_{x=0}^{\infty} x(x-1)f_X(x) = \sum_{x=2}^{\infty} x(x-1)(1-p)^{x-1}p.$$

(Σε αυτό το σημείο πρέπει να αναφερθεί ότι το κάτω άκρο του αθροίσματος είναι το 2, διότι για $\kappa = 0{,}1$ το άθροισμα έχει όρους ίσους με 0)

$$E\big(X(X-1)\big) = p\sum_{x=2}^{\infty} x(x-1)(1-p)^{x-1} =$$

$$= p\sum_{\kappa=2}^{\infty} \kappa(\kappa-1)(1-p)^{\kappa-1} \ \ (\gamma\iota\alpha\ x = \kappa) =$$

$$= p(1-p)\sum_{\kappa=2}^{\infty} \kappa(\kappa-1)\frac{(1-p)^{\kappa-1}}{1-p} =$$

$$= p(1-p)\sum_{\kappa=2}^{\infty} \kappa(\kappa-1)(1-p)^{\kappa-2}.$$

Είδαμε προηγουμένως ότι:

$$\sum_{\kappa=0}^{\infty} x^{\kappa} = \frac{1}{1-x} \ ,$$

επομένως θα ισχύει και:

$$\frac{d^2}{dx}\sum_{\kappa=0}^{\infty} x^{\kappa} = \frac{d^2}{dx}\frac{1}{1-x} \Rightarrow \sum_{\kappa=2}^{\infty} \kappa(\kappa-1)x^{\kappa-2} = \frac{2}{(1-x)^3} \ .$$

Άρα θα ισχύει επίσης:

$$E\big(X(X-1)\big) = p(1-p)\frac{2}{[1-(1-p)]^3} = p(1-p)\frac{2}{p^3}$$

$$\Rightarrow E\big(X(X-1)\big) = \frac{2(1-p)}{p^2} \Rightarrow E(X^2 - X) = \frac{2(1-p)}{p^2} \Rightarrow$$

$$\Rightarrow E(X^2) = \frac{2(1-p)}{p^2} + E(X) = \frac{2(1-p)}{p^2} + \frac{1}{p} \Rightarrow$$

$$\Rightarrow E(X^2) = \frac{2-p}{p^2}.$$

Τότε η διακύμανση θα είναι της μορφής:

$$Var(X) = E(X^2) - E^2(X) = \frac{2-p}{p^2} - \frac{1}{p^2} \Rightarrow$$

$$\Rightarrow Var(X) = \frac{1-p}{p^2}.$$

2.4: Η ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ.

Η συνεχής τυχαία μεταβλητή Χ ακολουθεί την εκθετική κατανομή με παράμετρο $\lambda > 0$, όταν η συνάρτηση πυκνότητας πιθανότητάς της είναι της μορφής:

$$f_X(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0, \lambda > 0 \end{cases}$$

Η συνάρτηση κατανομής της εκθετικής κατανομής είναι:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0, \lambda > 0 \end{cases}$$

Με ολοκλήρωση κατά παράγοντες για την μέση τιμή της εκθετικής θα ισχύει:

$$E(X) = \int_0^\infty x f_X(x) dx = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$

Παρόμοια για την ροπή $2^{ης}$ τάξης θα ισχύει:

$$E(X^2) = \int_0^\infty x^2 f_X(x) dx = \int_0^\infty x^2 \lambda e^{-\lambda x} = \frac{2}{\lambda^2}.$$

Η διακύμανση υπολογίζεται από τον τύπο:

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

Επομένως αν η $X \sim Exp(\lambda)$, τότε για την διακύμανσή της θα ισχύει:

$$Var(X) = E(X^2) - E^2(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 \Rightarrow Var(X) = \frac{1}{\lambda^2}.$$

Η ροπογεννήτρια συνάρτηση μίας συνεχούς τυχαίας μεταβλητής X έχουμε δει ότι υπολογίζεται από την σχέση:

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

Επομένως για την ροπογεννήτρια της εκθετικής κατανομής με παράμετρο λ,θα ισχύει:

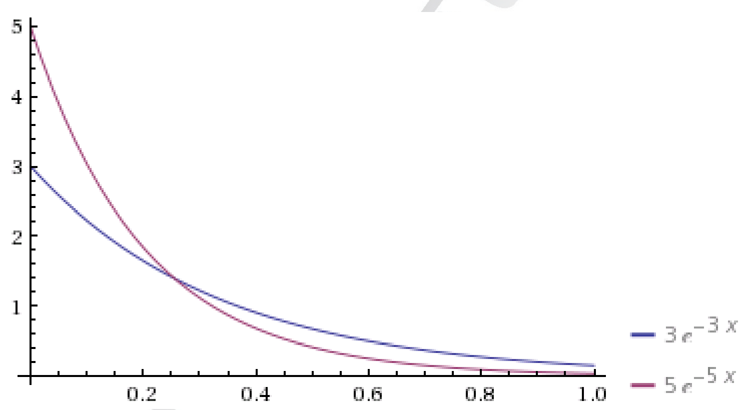
$$M_x(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx =$$

$$= \frac{\lambda}{\lambda - t} \int_0^{\infty} (\lambda - t) e^{-(\lambda - t)x} dx.$$

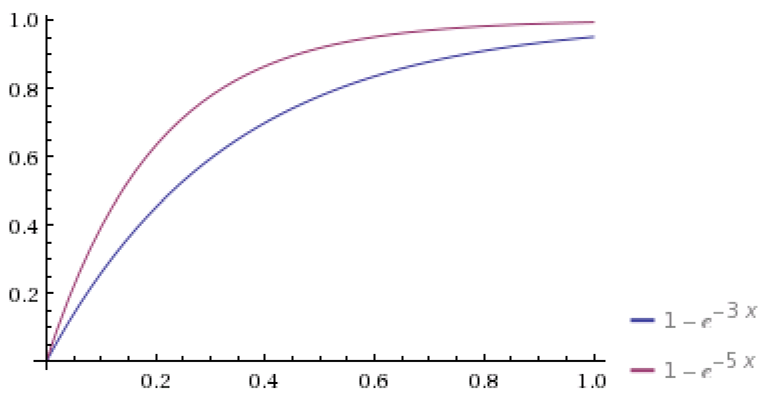
Το συγκεκριμένο ολοκλήρωμα είναι ολοκλήρωμα πυκνότητας εκθετικής κατανομής με παράμετρο $(\lambda - \mathrm{t})$ και σύμφωνα με την συνθήκη κανονικοποίησης ισούται με την μονάδα. Επομένως η ροπογεννήτρια της εκθετικής κατανομής με παράμετρο λ δίνεται από την σχέση:

$$M_x(t) = \frac{\lambda}{\lambda - t} \ (t < \lambda) .$$

Στο παρακάτω πρώτο σχήμα βλέπουμε τις γραφικές παραστάσεις των συναρτήσεων πυκνότητας πιθανότητας εκθετικής κατανομής με παραμέτρους 3 και 5, ενώ στο δεύτερο παρατηρούμε τις συναρτήσεις κατανομής τους.



Σχήμα 2.1: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\lambda = 3$ (μπλε χρώμα) , $\lambda = 5$. (ροζ χρώμα).



Σχήμα 2.2: Διάγραμμα της συνάρτησης κατανομής $F_X(x)$ για $\lambda = 3$ (μπλε χρώμα) , $\lambda = 5$ (ροζ χρώμα).

Στο παρακάτω παράδειγμα θα υπολογίσουμε την μέση τιμή, την διασπορά και την ροπογεννήτρια συνάρτηση μίας συνεχούς τυχαίας μεταβλητής X, η οποία ακολουθεί την Εκθετική κατανομή.

Παράδειγμα 2.4.1

Έστω μία συνεχής τυχαία μεταβλητή X η οποία ακολουθεί την Εκθετική κατανομή με παράμετρο 3. Από τις σχέσεις που αποδείξαμε παραπάνω έχουμε:

$$E(X) = \frac{1}{3},$$

$$Var(X) = \frac{2}{3^2} - \left(\frac{1}{3}\right)^2 = \frac{2}{9} - \frac{1}{9} = \frac{1}{9},$$

$$M_x(t) = \frac{3}{3-t} \ (t < 3).$$

2.5: ΚΑΤΑΝΟΜΗ ΓΑΜΜΑ.

Η συνεχής τυχαία μεταβλητή X ακολουθεί την κατανομή Γάμμα με παραμέτρους α, β, όταν η συνάρτηση πυκνότητας πιθανότητάς της είναι ίση με:

$$f_X(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \ , x \geq 0, \alpha \geq 0, \beta > 0.$$

Ο αριθμός Γ(α) έχει τις παρακάτω ιδιότητες:

$i) \Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$, όταν $\alpha \in R$.

$ii) \Gamma(n) = (n-1)!$, όταν $n \in N$.

Για παράδειγμα θα ισχύει ότι: $\Gamma(4) = (4-1)! = 6$.

$iii) \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Μία σημαντική παρατήρηση που αφορά την κατανομή Γάμμα είναι ότι για $\alpha = 1$ η $f_X(x)$ θα είναι της μορφής:

$$f_X(x) = \frac{\beta^1 x^{1-1} e^{-\beta x}}{\Gamma(1)} = \beta e^{-\beta x}.$$

Παρατηρούμε δηλαδή ότι για $\alpha = 1$ προκύπτει η εκθετική κατανομή με παράμετρο β.

Ο γενικός τύπος των ροπών κ −τάξης θα προκύπτει από την σχέση:

$$E(X^\kappa) = \int_0^\infty x^\kappa f_X(x) dx = \int_0^\infty \frac{x^\kappa \beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\kappa+\alpha-1} e^{-\beta x} dx =$$

$$= \frac{\beta^\alpha \Gamma(\kappa+\alpha)}{\Gamma(\alpha)\beta^{\kappa+\alpha}} \int_0^\infty \frac{x^{\kappa+\alpha-1} \beta^{\kappa+\alpha} e^{-\beta x}}{\Gamma(\kappa+\alpha)} dx.$$

Το ολοκλήρωμα είναι ολοκλήρωμα κατανομής Γάμμα με παραμέτρους $(\kappa + \alpha), \beta$, οπότε ισούται με την μονάδα. Επομένως ισχύει η σχέση:

$$E(X^\kappa) = \frac{\Gamma(\kappa+\alpha)}{\Gamma(\alpha)\beta^\kappa}.$$

Για $\alpha \in R, \kappa = 1$ θα έχουμε:

$$E(X) = \frac{\Gamma(1+\alpha)}{\Gamma(\alpha)} \frac{1}{\beta} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta} \Rightarrow E(X) = \frac{\alpha}{\beta},$$

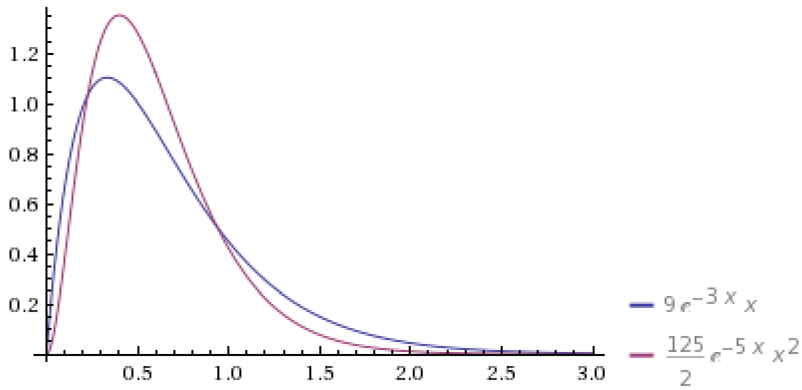
Για $\alpha \in R, \kappa = 2$ βρίσκουμε:

$$E(X^2) = \frac{\Gamma(2+\alpha)}{\Gamma(\alpha)\beta^2} = \frac{\Gamma(\alpha+1+1)}{\Gamma(\alpha)\beta^2} = \frac{(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha)\beta^2} = \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta^2} \Rightarrow E(X^2) = \frac{\alpha(\alpha+1)}{\beta^2},$$

και η διασπορά υπολογίζεται ως εξής :

$$Var(X) = E(X^2) - E^2(X) = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} \Rightarrow Var(X) = \frac{\alpha}{\beta^2}.$$

Στο παρακάτω σχήμα βλέπουμε τις γραφικές παραστάσεις συναρτήσεων πυκνότητας πιθανότητας δύο τυχαίων συνεχών μεταβλητών, οι οποίες ακολουθούν την κατανομή Γάμμα με παραμέτρους $\alpha = 2, \beta = 3$ και $\alpha = 3, \beta = 5$.



Σχήμα 2.3: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\alpha = 2, \beta = 3$ (μπλε χρώμα) και $\alpha = 3, \beta = 5$ (ροζ χρώμα).

2.6: Η ΚΑΤΑΝΟΜΗ WEIBULL.

Η συνεχής τυχαία μεταβλητή Χ ακολουθεί την κατανομή Weibull με παραμέτρους λ, α, όταν η συνάρτηση πυκνότητας πιθανότητάς της γράφεται ως:

$$f_X(x) = \alpha\lambda^{\alpha}x^{\alpha-1}e^{-(\lambda x)^{\alpha}}, x > 0,$$

όπου: $\lambda > 0, \alpha \geq 0$.

Η συνάρτηση κατανομής της Χ, είναι:

$$f_X(x) = 1 - e^{-(\lambda x)^{\alpha}}, x > 0.$$

Αξίζει να σημειωθεί ότι για $\alpha > 1$, η συνάρτηση πυκνότητας πιθανότητας έχει ελαφριά ουρά, ενώ για $\alpha < 1$, έχει βαριά ουρά. Για $\alpha = 1$, η συνάρτηση πυκνότητας πιθανότητας είναι εκθετική κατανομή με παράμετρο λ, διότι:

$$f_X(x) = \lambda x^0 e^{-(\lambda x)^1} = \lambda e^{-\lambda x}, x > 0.$$

Από το βιβλίο του Bean (2001) για τις ροπές κ −τάξεις της Weibull βλέπουμε ότι ισχύει:

$$E(X^{\kappa}) = \lambda^{-\kappa}\Gamma\left(\frac{\kappa+\alpha}{\alpha}\right).$$

Επομένως η μέση τιμή θα είναι:

$$E(X) = \frac{1}{\lambda}\Gamma\left(\frac{1+\alpha}{\alpha}\right).$$

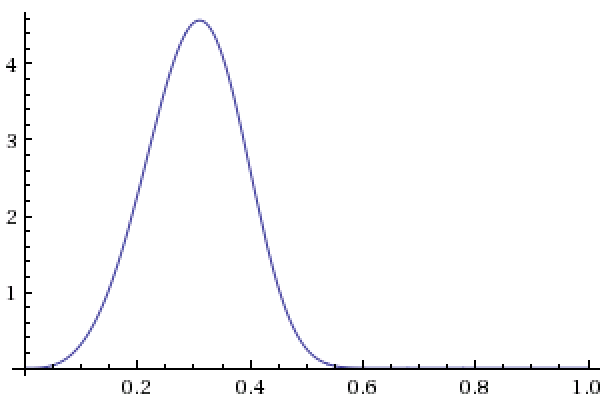
Η ροπή $2^{\eta\varsigma}$ τάξης θα είναι:

$E(X^2) = \frac{1}{\lambda^2}\Gamma\left(\frac{2+\alpha}{\alpha}\right).$

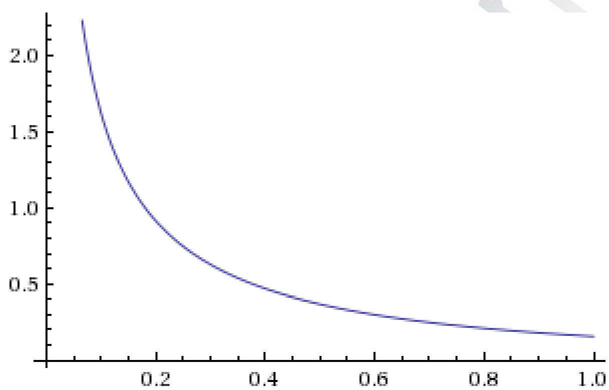
Συνεπώς η διασπορά θα είναι ίση με:

$$Var(X) = E(X^2) - E^2(X) = \frac{1}{\lambda^2}\left(\Gamma\left(1+\frac{2}{\alpha}\right) - \left(\Gamma\left(1+\frac{1}{\alpha}\right)\right)^2\right).$$

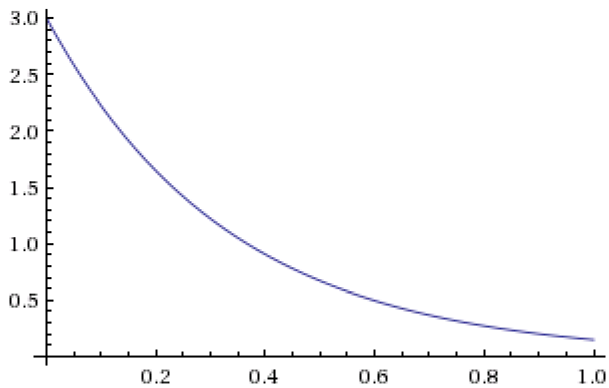
Στα παρακάτω τρία σχήματα βλέπουμε τις γραφικές παραστάσεις της συνάρτησης πυκνότητας κατανομής Weibull με παραμέτρους $\lambda = 3, \alpha = 4,\ \lambda = 3, \alpha = \frac{1}{2}$ και $\lambda = 3, \alpha = 1$ αντίστοιχα και παρατηρούμε ότι η πρώτη έχει ελαφριά ενώ η δεύτερη έχει βαριά ουρά, όπως ήταν αναμενόμενο, διότι $\alpha < 1$ και η τρίτη είναι εκθετική κατανομή με παράμετρο $\lambda = 3$.



Σχήμα 2.4:Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\lambda = 3, \alpha = 4$.



Σχήμα 2.5: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\lambda = 3, \alpha = \frac{1}{2}$.



Σχήμα 2.6: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\lambda = 3, \alpha = 1$.

Ας δούμε ένα παράδειγμα για τον υπολογισμό της διασποράς σε μία μεταβλητή Χ, η οποία ακολουθεί την κατανομή Weibull.

Παράδειγμα 2.6.1

Η συνεχής μεταβλητή Χ ακολουθεί την κατανομή Weibull με παραμέτρους 3,2 ($\lambda = 3, \alpha = 2$). Με βάση τις παραπάνω σχέσεις η διασπορά της θα είναι:

$$Var(X) = \frac{1}{3^2}\left(\Gamma\left(1+\frac{2}{2}\right) - \Gamma\left(1+\frac{1}{2}\right)^2\right) = \frac{1}{9}\left(\Gamma(2) - \left(\Gamma\left(\frac{3}{2}\right)\right)^2\right) =$$

$$= \frac{1}{9}\left((2-1)! - \left(\Gamma\left(\frac{1}{2}+1\right)\right)^2\right) = \frac{1}{9}\left(1 - \left(\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\right)^2\right) =$$

$$= \frac{1}{9}\left(1 - \left(\frac{1}{2}\sqrt{\pi}\right)^2\right) = \frac{4-\pi}{36}$$

2.7: Η ΚΑΤΑΝΟΜΗ PARETO.

Η συνεχής τυχαία μεταβλητή Χ ακολουθεί την κατανομή Pareto με παραμέτρους α, β, όταν η συνάρτηση πυκνότητας πιθανότητάς της είναι:

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

Η συνάρτηση κατανομής της είναι:

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

Από το βιβλίο του Bean (2001) βλέπουμε ότι οι ροπές κ −τάξης υπολογίζονται από την σχέση:

$$E(X^{\kappa}) = \frac{\beta^{\kappa}\kappa!}{(\alpha-1)(\alpha-2)\ldots(\alpha-\kappa)}.$$

Είναι απαραίτητο να αναφερθεί ότι η ροπή κ −τάξης ορίζεται $(E(X^{\kappa}) < \infty)$ όταν και μόνο όταν $\alpha > \kappa$.

Για $\kappa = 1$ έχουμε ότι:

$$E(X) = \frac{\beta}{\alpha-1}.$$

Για $\kappa = 2$:

$$E(X^2) = \frac{\beta^2 2!}{(\alpha-1)(\alpha-2)} = \frac{2\beta^2}{(\alpha-1)(\alpha-2)}.$$

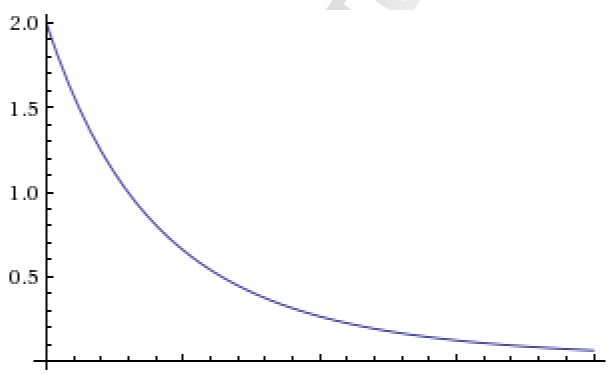
Για $\kappa = 3$:

$$E(X^3) = \frac{\beta^3 3!}{(\alpha-1)(\alpha-2)(\alpha-3)} = \frac{6\beta^3}{(\alpha-1)(\alpha-2)(\alpha-3)}.$$

Η διασπορά λοιπόν θα δίνεται από την σχέση:

$$Var(X) = E(X)^2 - E^2(X) = \frac{2\beta^2}{(\alpha-1)(\alpha-2)} - \frac{\beta^2}{(\alpha-1)^2} \Rightarrow Var(X) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}.$$

Στο παρακάτω σχήμα βλέπουμε την γραφική παράσταση της συνάρτησης πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής Χ που ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 4, \beta = 2$.



Σχήμα 2.7: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ για $\alpha = 4, \beta = 2$.

Στο παράδειγμα που θα κάνουμε παρακάτω θα υπολογίσουμε την διασπορά μίας συνεχούς τυχαίας μεταβλητής X, η οποία ακολουθεί την κατανομή Pareto.

Παράδειγμα 2.7.1

Η συνεχής μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 4, \beta = 2$. Από τα παραπάνω θα ισχύει ότι:

$$E(X) = \frac{2}{4-1} = \frac{2}{3}$$

Επίσης για την ροπή $2^{ης}$ τάξης θα ισχύει:

$$E(X^2) = \frac{2 \cdot 2^2}{(4-1)\cdot(4-2)} = \frac{4}{3}.$$

Η διασπορά επομένως θα είναι:

$$Var(X) = E(X^2) - E^2(X) = \frac{4}{3} - \left(\frac{2}{3}\right)^2 = \frac{8}{9}.$$

2.8: Η ΛΟΓΑΡΙΘΜΟΚΑΝΟΝΙΚΗ (LOGNORMAL) ΚΑΤΑΝΟΜΗ.

Αν η συνεχής τυχαία μεταβλητή Y ακολουθεί την κανονική κατανομή με παραμέτρους μ,σ $\big(Y \sim N(\mu,\sigma)\big)$ τότε η κατανομή της τυχαίας μεταβλητής $X = e^Y$, καλείται λογαριθμοκανονική κατανομή με παραμέτρους μ,σ $\big(X \sim LN(\mu,\sigma)\big)$. Επομένως αν $X \sim LN(\mu,\sigma)$, τότε: $logX \sim N(\mu,\sigma)$. Οι παράμετροι μ, σ, είναι πραγματικοί αριθμοί, οι οποίοι αντιπροσωπεύουν την μέση τιμή και την τυπική απόκλιση του $logX$.

Η τυχαία μεταβλητή X έχει συνάρτηση πυκνότητας πιθανότητας:

$$f_X(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(logx-\mu)^2}{2\sigma^2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Η συνάρτηση πυκνότητας πιθανότητας προκύπτει από παραγώγιση της συνάρτησης κατανομής, η οποία είναι:

$$F_X(x) = \Phi\left(\frac{logx-\mu}{\sigma}\right), x \geq 0.$$

Οι ροπές κ −τάξης βλέπουμε στο βιβλίο του Bean (2001) ότι προκύπτουν από την σχέση:

$$E(X^\kappa) = e^{\mu\kappa + \frac{\sigma^2\kappa^2}{2}}.$$

Έτσι τότε για $\kappa = 1$, θα έχουμε:

$$E(X) = e^{\mu+\frac{\sigma^2}{2}}.$$

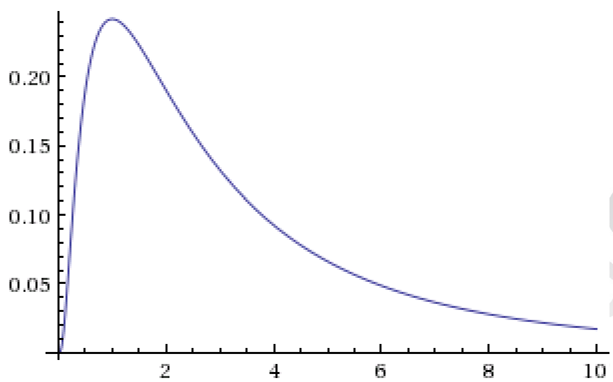
Για $\kappa = 2$:

$$E(X^2) = e^{2(\mu+\sigma^2)}.$$

Συνεπώς η διασπορά θα είναι:

$$Var(X) = E(X^2) - E^2(X) = e^{2(\mu+\sigma^2)} - \left(e^{\mu+\frac{\sigma^2}{2}}\right)^2 = e^{2\mu}\left(e^{\sigma^2}(e^{\sigma^2} - 1)\right).$$

Στα παρακάτω δύο σχήματα βλέπουμε την γραφική παράσταση λογαριθμοκανονικής κατανομής μίας συνεχούς τυχαίας μεταβλητής X, με παραμέτρους $\mu = 1, \sigma = 1$.



Σχήμα 2.8: Διάγραμμα της συνάρτησης πυκνότητας πιθανότητας $f_X(x)$ με παραμέτρους $\mu = 1, \sigma = 1$.

Παράδειγμα 2.8.1

Η τυχαία μεταβλητή X ακολουθεί την λογαριθμοκανονική κατανομή με παραμέτρους 3,2, $(\mu = 3, \sigma = 2)$. Τότε έχουμε:

$$E(X) = e^{3+\frac{4}{2}} = e^5,$$

$$E(X^2) = e^{2(3+4)} = e^{14},$$

$$Var(X) = E(X)^2 - E^2(X) = e^{14} - e^{10} \Rightarrow Var(X) = e^{10}(e^4 - 1).$$

2.9: ΜΕΙΚΤΕΣ ΚΑΤΑΝΟΜΕΣ.

Στην θεωρία πιθανοτήτων οι κατανομές χωρίζονται σε διακριτές και συνεχείς. Υπάρχουν όμως κατανομές οι οποίες δεν είναι ούτε «εξ ολοκλήρου» διακριτές, ούτε «εξ ολοκλήρου» συνεχείς. Αυτές οι κατανομές έχουν ένα διακριτό και ένα συνεχές τμήμα και γι’αυτόν τον λόγο ονομάζονται μεικτές. Από το βιβλίο του Πολίτη (2012) για τις πυκνότητες πιθανότητας στις συνεχείς κατανομές βλέπουμε ότι ισχύει η σχέση:

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

Για τις συναρτήσεις πιθανότητας στις διακριτές κατανομές ισχύει:

$$\sum_{x\in R} f_X(x) = 1.$$

Γενικότερα το πεδίο ορισμού για μία μεικτή κατανομή χωρίζεται σε δύο υποσύνολα, ένα διακριτό και ένα συνεχές. Ας υποθέσουμε ότι η μεικτή τυχαία μεταβλητή X παίρνει τιμές σε ένα σύνολο R, το οποίο χωρίζεται σε δύο υποσύνολα R_d, το οποίο είναι αριθμήσιμο και R_c , στο οποίο η X είναι συνεχής, ισχύει δηλαδή ότι: $\mathrm{R_d} \cup \mathrm{R_c} = \mathrm{R}$. Για την X τότε θα ισχύει:

$$\sum_{x\in R_d} f_1(x) + \int_{R_c} f_2(x)dx = 1,$$

όπου f_1 η συνάρτηση πιθανότητας και f_2 η συνάρτηση πυκνότητας πιθανότητας, έτσι ώστε η συνολική μάζα πιθανότητας να είναι ίση με την μονάδα. Ο γενικός τύπος για τον υπολογισμό συνάρτησης μίας τυχαίας μεταβλητής που ακολουθεί την μεικτή κατανομή δίνεται από την σχέση:

$$E[g(x)] = \sum_{i:x_i\in R_d} g(x_i)f_1(x_i) + \int_{R_c} g(x)f_2(x)dx.$$

όπου $f_1, f_{2,}$η συνάρτηση πιθανότητας και η συνάρτηση πυκνότητας αντίστοιχα.

Αν για παράδειγμα θέλουμε να υπολογίσουμε την διακύμανση της μεικτής κατανομής X, χρησιμοποιούμε την εξίσωση:

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

Σύμφωνα με την προηγούμενη σχέση η ροπή δεύτερης τάξης θα είναι:

$$E(X^2) = \sum_{x\in R_d} x^2 f_1(x) + \int_{x\in R_c} x^2 f_2(x)dx,$$

Ενώ η μέση τιμή θα δίνεται από την σχέση:

$$E(X) = \sum_{x\in R_d} x f_1(x) + \int_{x\in R_c} x f_2(x)dx$$

και υπολογίζουμε την διακύμανση από τις δύο παραπάνω σχέσεις.

2.10: ΚΑΤΑΝΟΜΗ ΑΘΡΟΙΣΜΑΤΟΣ-ΔΙΑΦΟΡΑΣ-ΓΙΝΟΜΕΝΟΥ ΚΑΙ ΠΗΛΙΚΟΥ ΔΥΟ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ.

Σύμφωνα με τις σημειώσεις του Πολίτη (2015) και το βιβλίο του Κούτρα (2005) μπορούμε να δούμε ότι όταν έχουμε δύο συνεχείς τυχαίες μεταβλητές X, Y, με από κοινού συνάρτηση πυκνότητας $f(x, y)$ και επιθυμούμε να υπολογίσουμε την συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής: $U = g(x, y)$, τότε υποθέτουμε ότι η U έχει μοναδική λύση ως προς x και έχουμε ότι:

$$u = g(x, y) \Rightarrow x = g^*(x, y).$$

Εισάγουμε την βοηθητική μεταβλητή $V = Y = h(X, Y)$ και έχουμε:

$$g(x, y) = u \Rightarrow x = g^*(u, y)$$

$$h(x, y) = v \Rightarrow y = v \Rightarrow y = v = h^*(u, v),$$

επομένως θα ισχύει ότι:

$$x = g^*(u, v).$$

Αν πάρουμε την Ιακωβιανή ορίζουσα θα έχουμε:

$$J = \begin{vmatrix} \frac{\partial g^*(u,v)}{\partial u} & \frac{\partial g^*(u,v)}{\partial v} \\ \frac{\partial h^*(u,v)}{\partial u} & \frac{\partial h^*(u,v)}{\partial v} \end{vmatrix} = \frac{\partial g^*(u,v)}{\partial u}.$$

Σε περίπτωση που οι X,Y είναι ανεξάρτητες, για την από κοινού συνάρτηση $f(x, y)$, θα ισχύει:

$$f(x, y) = f_X(x) f_Y(y),$$

και η πυκνότητα της U θα υπολογίζεται από την σχέση:

$$f_U(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u, v)\big) f_Y(v) \left|\frac{\partial g^*(u,v)}{\partial u}\right| dv.$$

Η παραπάνω σχέση γράφεται ισοδύναμα στην μορφή:

$$f_U(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u, y)\big) f_Y(y) \left|\frac{\partial g^*(u,y)}{\partial u}\right| dy.$$

Όσον αφορά την κατανομή αθροίσματος δύο συνεχών τυχαίων μεταβλητών, έχουμε ότι:

Έστω X και Y δύο συνεχείς τυχαίες μεταβλητές με συναρτήσεις πυκνότητας πιθανότητας $f_X(x)$, $f_Y(y)$ αντίστοιχα. Αν οι Χ, Υ είναι ανεξάρτητες και $U = X + Y$ τότε η U έχει συνάρτηση πυκνότητας πιθανότητας η οποία συμβολίζεται ως $(f * g)$ και εισάγοντας τη συνάρτηση $g(x, y) = x + y$, έχουμε ότι:

$$g(x + y) = u \Rightarrow x + y = u \Rightarrow x = u - y = g^*(u, y).$$

Επίσης ισχύει ότι:

$$\frac{\partial g^*(u,y)}{\partial u} = 1.$$

Θα έχουμε συνεπώς ότι:

$$(f * g)(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u, y)\big) f_Y(y) \left|\frac{\partial g^*(u,y)}{\partial u}\right| dy = \int_{-\infty}^{\infty} f_X(u - y) f(y) dy.$$

Εάν έχουμε δύο ανεξάρτητες διακριτές τυχαίες μεταβλητές με συναρτήσεις πιθανότητας $f_X(x), f_Y(y)$ αντίστοιχα, τότε η συνάρτηση πιθανότητας της U θα υπολογίζεται από τον τύπο:

$$f_U(u) = \sum_{y \in R_y} f_X(u - y) f_Y(y),$$

ή ισοδύναμα:

$$f_U(u) = \sum_{x \in R_x} f_Y(u - x) f_X(x),$$

όπου R_X , R_Y είναι τα πεδία ορισμού των f_X , f_Y αντίστοιχα.

Στις συνελίξεις επίσης ισχύουν κανονικά η αντιμεταθετική και η προσεταιριστική ιδιότητα. Αν έχουμε τρεις συναρτήσεις πυκνότητας f, g, h, θα ισχύει:

$$(f * g)(x) = (g * f)(x)\,, (f * h)(x) = (h * f)(x),$$
$$\big((f * g) * h\big) = \big(f * (g * h)\big).$$

Στην συνέχεια θα παρουσιάσουμε ένα παράδειγμα συνέλιξης.

Παράδειγμα 2.10.1

Έστω οτι η μεταβλητή Χ ακολουθεί την εκθετική κατανομή με παράμετρο 4 και η μεταβλητή Υ ακολουθεί την κατανομή Γάμμα με παραμέτρους: $\alpha = 2, \beta = 3$. Τότε η συνάρτηση πυκνότητας της Χ θα είναι της μορφής:

$$f_X(x) = 4e^{-4x}(x > 0).$$

Η συνάρτηση πυκνότητας της Υ θα είναι:

$$f_Y(y) = 9ye^{-3y}(y > 0).$$

Θέλουμε να υπολογίσουμε την συνάρτηση πυκνότητας της μεταβλητής:

$$Z = X + Y.$$

Η συνάρτηση πυκνότητάς της θα υπολογίζεται από την σχέση:

$$f_Z(z) = \int_0^z f_x(z-y) f_y(y) dy.$$

Ο λόγος που τα άκρα του παραπάνω ολοκληρώματος είναι 0 και z είναι διότι οι συγκεκριμένες συνεχείς τυχαίες μεταβλητές X και Y ορίζονται μόνο για θετικές τιμές. Ισχύει δηλαδή ότι: $(x, y \geq 0)$. Επομένως θα ισχύει και:

$$y \geq 0,$$

όπως και:

$$z - y \geq 0 \Rightarrow -y \geq -z \Rightarrow y \leq z,$$

επομένως θα ισχύει:

$$0 \leq y \leq z.$$

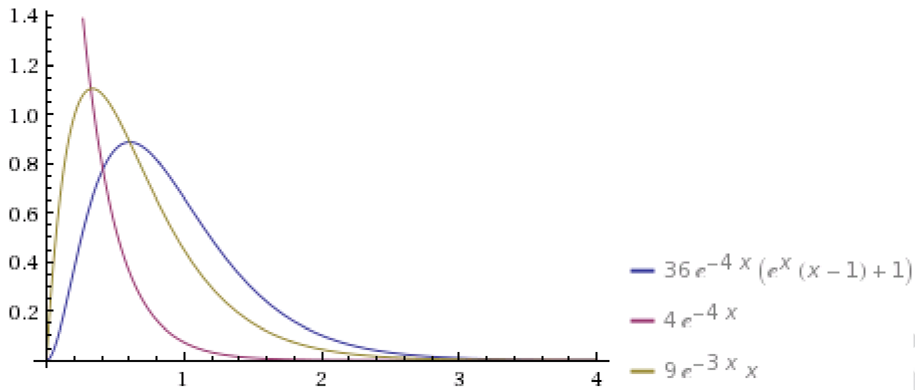
Έχουμε τότε ότι:

$$f_Z(z) = \int_0^z f_x(z-y) f_y(y) dy = \int_0^z 4e^{-4(z-y)} 9ye^{-3y} dy =$$

$$= \int_0^z 4e^{-4z} e^{4y} 9ye^{-3y} dy = 36e^{-4z} \int_0^z e^y y dy$$

$$\Rightarrow f(z) = 36e^{-4z}[e^z(z-1) + 1], \ (z \geq 0).$$

Στο παρακάτω σχήμα βλέπουμε τις γραφικές παραστάσεις των τριών συναρτήσεων και παρατηρούμε ότι η εκθετική κατανομή είναι φθίνουσα συνάρτηση ως προς X (ισχύει δηλαδή $\forall$ $x_1 > x_2 \Rightarrow f_X(x_1) < f_X(x_2)$ και η γραφική παράσταση της συνάρτησης Γάμμα είναι πιο «απότομη» από την γραφική παράσταση της συνάρτησης της συνέλιξης.



Σχήμα 2.9: Διάγραμμα των $f_X(x)$ (ροζ χρώμα), $f_Y(x)$ (κίτρινο χρώμα), $f_Z(x)$ (μπλε χρώμα).

Όσον αφορά την κατανομή της διαφοράς δύο συνεχών τυχαίων μεταβλητών, έχουμε ότι:

Έστω X,Y δύο ανεξάρτητες συνεχείς τυχαίες μεταβλητές με $f_X(x), f_Y(y)$ τις αντίστοιχες συναρτήσεις πυκνότητας πιθανότητάς τους. Τότε η συνάρτηση πυκνότητας της τυχαίας μεταβλητής $U = X - Y$, δίνεται από την σχέση:

$$f_U(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u,y)\big) f_Y(y) \left|\frac{\partial g^*(u,y)}{\partial u}\right| dy.$$

Εισάγοντας την συνάρτηση $g(x,y) = x - y$ έχουμε:

$$g(x,y) = x - y, \quad g(x,y) = u \Rightarrow x - y = u \Rightarrow x = u + y = g^*(u,y).$$

Επομένως θα ισχύει:

$$\frac{\partial g^*(u,y)}{\partial u} = 1,$$

Προφανώς τότε θα είναι:

$$f_U(u) = \int_{-\infty}^{\infty} f_X(y+u) f_Y(y) dy = \int_{-\infty}^{\infty} f_Y(x-u) f_X(x) dx.$$

Για την συνάρτηση πυκνότητας της τυχαίας μεταβλητής $U = XY$ δύο ανεξάρτητων τυχαίων μεταβλητών X,Y, με αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας $f_X(x), f_Y(y)$ αντίστοιχα θα ισχύει:

$$f_U(u) = f_U(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u,y)\big) f_Y(y) \left|\frac{\partial g^*(u,y)}{\partial u}\right| dy.$$

Έχουμε τότε την συνάρτηση:

$$g(x,y) = xy \Rightarrow g(x,y) = u \Rightarrow xy = u \Rightarrow x = \frac{u}{y} = g^*(u,y).$$

Επίσης θα ισχύει ότι:

$$\frac{\partial g^*(u,y)}{\partial u} = \frac{1}{y},$$

επομένως η $f_U(u)$ θα γράφεται στην μορφή:

$$f_U(u) = \int_{-\infty}^{\infty} f_X\left(\frac{u}{y}\right) f_Y(y) \left|\frac{1}{y}\right| dy.$$

Αναλόγως εάν θέλουμε να υπολογίσουμε την κατανομή του λόγου δύο συνεχών τυχαίων μεταβλητών X,Y, θα πρέπει να υπολογίσουμε την συνάρτηση πυκνότητας της τυχαίας μεταβλητής $U = \frac{X}{Y}$. Παρομοίως θα έχουμε:

$$f_U(u) = f_U(u) = \int_{-\infty}^{\infty} f_X\big(g^*(u,y)\big) f_Y(y) \left|\frac{\partial g^*(u,y)}{\partial u}\right| dy.$$

Εισάγουμε την συνάρτηση $g = \frac{x}{y}$ και προκύπτει:

$$g(x,y) = u \Rightarrow \frac{x}{y} = u \Rightarrow x = uy = g^*(u,y),$$

και η $f_U(u)$ γράφεται πλέον στην μορφή:

$$f_U(u) = \int_{-\infty}^{\infty} f_X(uy) f_Y(y) |y| dy.$$

ΚΕΦΑΛΑΙΟ 3: ΠΡΟΣΕΓΓΙΣΕΙΣ ΚΑΤΑΝΟΜΩΝ ΠΙΘΑΝΟΤΗΤΑΣ ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Σε αυτό το κεφάλαιο θα ασχοληθούμε κυρίως με τις προσεγγίσεις κατανομών πιθανότητας από μία μείξη εκθετικών κατανομών. Θα προσεγγίζουμε δηλαδή κατανομές πιθανότητας όπως οι: Weibull, Pareto και Lognormal από μία μορφή μείξης εκθετικών κατανομών. Για να το επιτύχουμε αυτό θα χρησιμοποιήσουμε την μέθοδο των ροπών, σύμφωνα με την οποία έαν έχουμε μία συνεχή τυχαία μεταβλητή Υ, η οποία ακολουθεί μία κατανομή όπως π.χ την Pareto και μία συνεχή τυχαία μεταβλητή Χ, με συνάρτηση πυκνότητας $f_X(x)$, η οποία αποτελεί μείξη εκθετικών κατανομών, τότε για να προσεγγίσουμε την συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ της μεταβλητής Υ από μία μείξη εκθετικών κατανομών $f_X(x)$, για τις ροπές κ −τάξης θα χρησιμοποιούμε τις σχέσεις :

$E(Y) = E(X)$

$E(Y^2) = E(X^2)$

$E(Y^3) = E(X^3)$, κ.ο.κ .

Οι παραπάνω εξισώσεις υποθέτουμε ότι ισχύουν για τιμές του κ ανάλογα με πόσες παραμέτρους έχει η μείξη εκθετικών κατανομών που χρησιμοποιούμε. Στα παραδείγματα που θα κάνουμε κατά την διάρκεια του κεφαλαίου θα διαπιστώσουμε για παράδειγμα ότι εάν η μείξη εκθετικών κατανομών είναι μείξη δύο εκθετικών, τότε πρέπει να υπολογίσουμε τρεις αγνώστους και χρειαζόμαστε τρεις εξισώσεις, οπότε και χρησιμοποιούμε μέχρι και την ροπή τρίτης τάξης. Αν έχει τρεις παραμέτρους τότε πρέπει να υπολογίσουμε πέντε αγνώστους, οπότε και χρειαζόμαστε πέντε εξισώσεις και θα χρησιμοποήσουμε μέχρι και την ροπή πέμπτης τάξης.

Με αυτόν τον τρόπο δημιουργούμε ένα σύστημα εξισώσεων, με το οποίο υπολογίζουμε τις παραμέτρους της μείξης εκθετικών κατανομών. Αυτή η προσέγγιση θα διαπιστώσουμε ότι είναι ιδιαίτερα χρήσιμη, διότι μπορούμε για παράδειγμα να υπολογίσουμε την πιθανότητα χρεοκοπίας μίας κατανομής αποζημιώσεων, η οποία δεν ακολουθεί μία μείξη εκθετικώ κατανομών. Με αυτήν την περίπτωση όμως θα ασχοληθούμε αναλυτικότερα στο κεφάλαιο 5.

Στην ενότητα 3.1 θα εξηγήσουμε αναλυτικά ποιες συναρτήσεις αποτελούν μείξεις εκθετικών κατανομών, ενώ στις υπόλοιπες ενότητες του κεφαλαίου θα ασχοληθούμε με την προσέγγιση διαφόρων πυκνοτήτων που αντιστοιχούν σε γνωστές κατανομές από μία μείξη εκθετικών κατανομών, καθώς επίσης και θα παρουσιάσουμε και κάποιες μεθόδους με τις οποίες αξιολογούμε τις προσεγγίσεις αυτές. Στις τρεις τελευταίες Ενότητες του κεφαλαίου θα χρησιμοποιήσουμε την ίδια μέθοδο για να προσεγγίσουμε τις ίδες συναρτήσεις πυκνότητας πιθανότητας, από μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο α να ισούται με 2 και έπειτα θα γίνει μία αξιολόγηση για να συμπεράνουμε ποια από τις δύο προσεγγίσεις είναι περισσότερο ικανοποιητική.

3.1:ΜΕΙΞΕΙΣ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Μία κατανομή με συνάρτηση πυκνότητας η οποία μπορεί να γραφτεί στην μορφή:

$$f_X(x) = \alpha_1\beta_1 e^{-\beta_1 x} + \alpha_2\beta_2 e^{-\beta_2 x} + \cdots + \alpha_n\beta_n e^{-\beta_n x} =$$

$$= \sum_{j=1}^{n} \alpha_j\beta_j e^{-\beta_j x} \left(\alpha_j, \beta_j > 0\right), x \geq 0,$$

για την οποία ισχύει:

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1 \Rightarrow \sum_{j=1}^{n} \alpha_j = 1,$$

ονομάζεται μείξη εκθετικών κατανομών.

Εάν υπάρχουν και κάποια: $\alpha_j < 0$, με:

$$\sum_{j=1}^{n} \alpha_j = 1,$$

τότε η παραπάνω συνάρτηση ονομάζεται συνδυασμός εκθετικών κατανομών.

Κατά την διάρκεια του κεφαλαίου θα ασχοληθούμε κυρίως με μείξεις εκθετικών δύο παραμέτρων, που είναι δηλαδή της μορφής:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x \geq 0.$$

Σύμφωνα με άρθρο των Babier & Chan (1992) για να αποτελεί η $f_X(x)$ συνάρτηση πυκνότητας πιθανότητας πρέπει να ισχύει:

$$0 \leq A \leq \frac{\beta_2}{\beta_2 - \beta_1} (\, \beta_2 > \beta_1),$$

διότι για $A < 0$, η $f_X(x)$ θα είναι αρνητική για μεγάλους αριθμούς x και για $A > \frac{\beta_2}{\beta_2 - \beta_1}$, η $f_X(x)$ θα είναι αρνητική για τιμές του x κοντά στο 0. Επίσης για $\beta_2 = \beta_1$, έχουμε ότι $A = 0$ ή $A = 1$ και έτσι η $f_X(x)$ είναι εκθετική συνάρτηση με μία παράμετρο $\beta_2 = \beta_1$.

Ας δούμε μερικά παραδείγματα:

- Η: $f_X(x) = \frac{1}{2}3e^{-3x} + \frac{1}{4}5e^{-5x} + \frac{1}{4}6e^{-6x}$,
  αποτελεί μία μείξη εκθετικών κατανομών, διότι:
  $\beta_1(=3), \beta_2(=5), \beta_3(=6) > 0$,

  $$\alpha_1\left(=\frac{1}{2}\right) + \alpha\left(=\frac{1}{4}\right) + \alpha_3\left(=\frac{1}{4}\right) = 1.$$

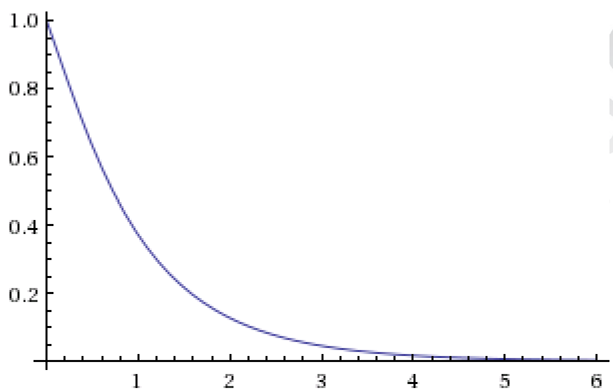
- Η $f_X(x) = 4e^{-12x} + 10e^{-15x}$,

αποτελεί μία μείξη εκθετικών κατανομών, διότι μπορεί να γραφτεί στην μορφή:
$f_X(x) = \frac{1}{3}12e^{-12x} + \frac{2}{3}15e^{-15x}, \mu\varepsilon\ \ \frac{1}{3} + \frac{2}{3} = 1.$

- Η $f_X(x) = 4e^{-2x} + 3e^{-5x} = 2 \cdot 2e^{-2x} + \frac{3}{5}5e^{-5x}$
  δεν αποτελεί μείξη εκθετικών κατανομών διότι:
  $2 + \frac{3}{5} \neq 1.$

- Η $f_X(x) = \frac{7e^{-x}+4e^{-2x}-3e^{-3x}}{8}, x > 0$
  αποτελεί έναν συνδυασμό εκθετικών κατανομών, διότι μπορει να γραφτεί στην μορφή:
  $f_X(x) = \frac{7}{8}e^{-x} + \frac{2}{8} \cdot 2e^{-2x} - \frac{1}{8} \cdot 3e^{-3x}, x > 0$
  Και έχουμε ότι:
  $\alpha_1\left(=\frac{7}{8}\right) + \alpha_2\left(=\frac{2}{8}\right) + \alpha_3\left(=-\frac{1}{8}\right) = 1\,, (\alpha_3 < 0).$

Στο παρακάτω σχήμα βλέπουμε την γραφική παράσταση της $f_X(x)$.



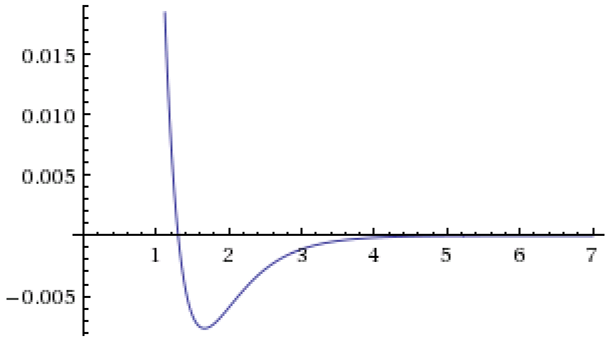
Σχήμα 3.1: Γραφική παράσταση της $f_X(x)$.

Σε αυτό το σημείο είναι απαραίτητο να αναφερθεί ότι για να αποτελεί η συνάρτηση $f_X(x)$ συνδυασμό εκθετικών κατανομών θα πρέπει να ισχύει: $f_X(x) \geq 0, \forall\ x > 0.$ Για παράδειγμα η:

$f_Y(y) = \frac{1}{4}3e^{-3y} + 1 \cdot 4e^{-4y} - \frac{1}{4}2e^{-2y}$, έχει αντίστοιχα βάρη:

$a_1\left(=\frac{1}{4}\right) + a_2(=1) + a_3\left(=-\frac{1}{4}\right),$

όμως στο παρακάτω σχήμα βλέπουμε ότι δεν ισχύει: $f_Y(y) \geq 0, \forall\ y > 0.$



Σχήμα 3.2: Γραφική παράσταση της $f_Y(y)$.

Στην ενότητα 3.2 θα παρουσιάσουμε τέσσερις διαφορετικές μεθόδους για την αξιολόγηση των προσεγγίσεων, που θα πραγματοποιηθούν, ενώ στις ενότητες 3.3, 3.4, 3.5, θα προσεγγίσουμε τις κατανομές Weibull, Pareto, Lognormal αντίστοιχα, από μία μείξη εκθετικών κατανομών.

3.2: ΑΞΙΟΛΟΓΗΣΗ ΠΡΟΣΕΓΓΙΣΕΩΝ.

Έστω ότι θέλουμε να αξιολογήσουμε την προσέγγιση δύο συναρτήσεων πυκνότητας πιθανότητας $f_Y(y)$ και $f_X(x)$ (στην περίπτωση που εξετάζουμε, έστω $f_Y(y)$ η συνάρτηση πυκνότητας πιθανότητας της κατανομής που θέλουμε να προσεγγίσουμε από μία μείξη εκθετικών κατανομών $f_X(x)$), παρουσιάζουμε παρακάτω τέσσερις διαφορετικές μεθόδους για την αξιολόγηση αυτή.

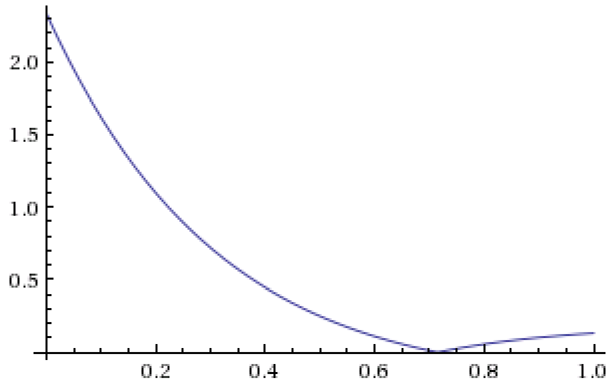
i) Υπολογίζουμε πρώτα τις συναρτήσεις κατανομής $F_Y(y)$ και $F_X(x)$ και βρίσκουμε την μέγιστη διαφορά των δύο συναρτήσεων ως εξής:

$$sup_{x\geq 0}|f_Y(x) - f_X(x)| \quad \text{και} \quad sup_{x\geq 0}|F_Y(x) - F_X(x)|. \qquad \text{(μέθοδος 3.2.1)}$$

Έχουμε για παράδειγμα μία συνεχής τυχαία μεταβλητή X, η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 2, \beta = 3$ και την συνεχή τυχαία μεταβλητή Y, η οποία ακολουθεί την εκθετική κατανομή με παράμετρο $\lambda = 3$. Τότε η μέγιστη διαφορά των δύο συναρτήσεων πυκνότητας πιθανότητας θα είναι:

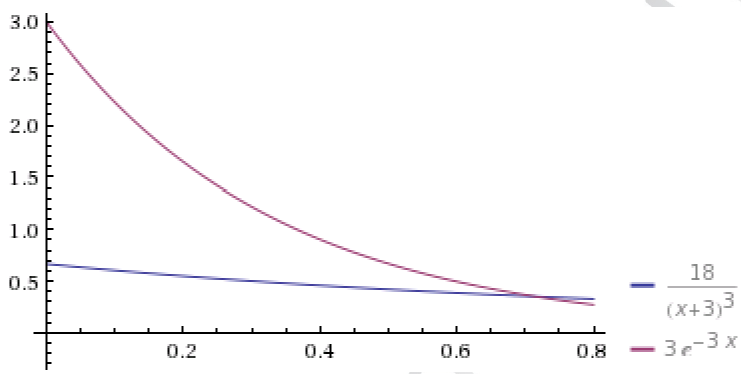
$$sup_{x\geq 0}|f_Y(x) - f_X(x)| = sup_{x\geq 0}\left|\frac{2}{3}\left(1+\frac{x}{3}\right)^{-3} - 3e^{-3x}\right| \approx 2.32501,$$

στο σημείο $x_1 \approx 0$. Στο παρακάτω γράφημα βλέπουμε την γραφική παράσταση του $|f_Y(x) -$
$-fXx.$



Σχήμα 3.3: Διάγραμμα του $|f_Y(x) - f_X(x)|$.

Η συγκεκριμένη απόσταση παρατηρείται και από το παρακάτω σχήμα των γραφικών παραστάσεών τους στο συγκεκριμένο σημείο.



Σχήμα 3.4: Διάγραμμα των $f_X(x)$ (μπλε χρώμα), $f_Y(x)$ (ροζ χρώμα).

ii) Υπολογίζουμε την «απόσταση» των δύο συναρτήσεων πυκνότητας πιθανότητας, δηλαδή το ολοκλήρωμα της απόλυτης διαφοράς τους, ως εξής:

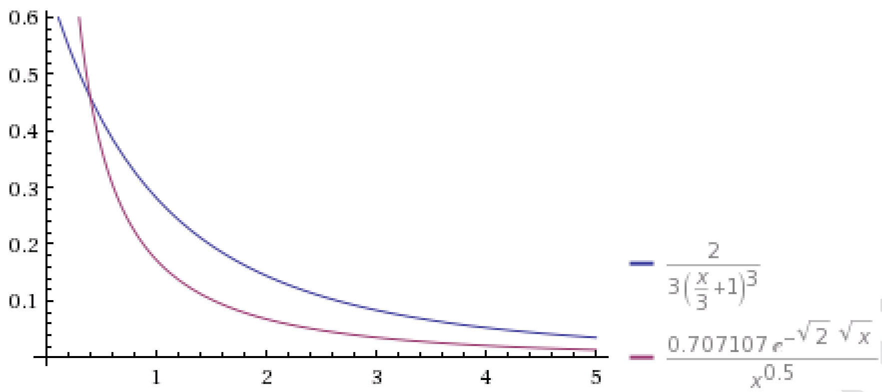
$\int_0^\infty |f_Y(x) - f_X(x)|dx.$ (μέθοδος 3.2.2)

Ας πάρουμε για παράδειγμα δύο τυχαίες συνεχείς μεταβλητές X και Y, οι οποίες ακολουθούν τις κατανομές Pareto με παραμέτρους $\alpha = 2, \beta = 3$ και Weibull με παραμέτρους $\lambda = 0.5, \alpha = 2$ αντίστοιχα. Οι συναρτήσεις πυκνότητας πιθανότητας θα είναι της μορφής:

$f_X(x) = \frac{2}{3}\left(1 + \frac{x}{3}\right)^{-3}, x > 0,$

$f_Y(y) = 0.5\sqrt{2}y^{-0.5}e^{-\sqrt{2y}}, y > 0,$

αντίστοιχα. Οι γραφικές παραστάσεις των $f_X(x)$ και $f_Y(x)$, απεικονίζονται στο παρακάτω σχήμα.



Σχήμα 3.5: Διάγραμμα των $f_X(x)$ (μπλε χρώμα) και $f_Y(x)$ (ροζ χρώμα).

Τότε το ολοκλήρωμα της απόλυτης διαφοράς τους θα είναι:

$$\int_0^{\infty}|f_Y(x)-f_X(x)|dx=\int_0^{\infty}\left|\frac{2}{3}\left(1+\frac{x}{3}\right)^{-3}-0.5\sqrt{2}x^{-0.5}e^{-\sqrt{2x}}\right|dx\approx 0.739.$$

Σε αυτό το σημείο αξίζει να σημειωθεί ότι στην περίπτωση δύο συναρτήσεων πυκνότητας πιθανότητας το παραπάνω ολοκλήρωμα είναι έως 2, διότι από την συνθήκη κανονικοποίησης το ολοκλήρωμα κάθε συνάρτησης πυκνότητας πιθανότητας ισούται με 1, και σύμφωνα με την τριγωνική ανισότητα ισχύει ότι:

$$\int_0^{\infty}|f_Y(x)-f_X(x)|dx\le\int_0^{\infty}|f_Y(x)|dx+\int_0^{\infty}f_X(x)dx=2.$$

Οπότε μία προσέγγιση είναι επιτυχής όταν το συγκεκριμένο ολοκλήρωμα είναι κοντά στο 0 και δεν είναι ικανοποιητική όταν το παραπάνω ολοκλήρωμα παίρνει τιμές κοντά στο 2.

$iii)$ Εξετάζουμε την διαφορά των ροπών ανώτερης τάξης από αυτές που έχουμε χρησιμοποιήσει. Στις επόμενες τρεις Ενότητες θα προσεγγίσουμε τις συναρτήσεις πυκνότητας πιθανότητας τριών συνεχών τυχαίων μεταβλητών, οι οποίες ακολουθούν τις κατανομές Pareto, Weibull και Lognormal από μία μείξη δύο εκθετικών κατανομών και θα χρησιμοποιήσουμε τις σχέσεις:

$$E(Y)=E(X),\ E(Y^2)=E(X^2),\ E(Y^3)=E(X^3).$$

Οπότε θα εξετάσουμε πόσο «κοντά» είναι οι ροπές:

$E(X^4)$ με την $E(Y^4)$, η $E(X^5)$ με την $E(Y^5)$, η $E(X^6)$ με την $E(Y^6)$, κ.ο.κ. (μέθοδος 3.2.3)

$iv)$ Υπολογίζουμε τα άνω ποσοστημόρια των δύο κατανομών και εξετάζουμε πόσο κοντά βρίσκονται μεταξύ τους, δηλαδή χρησιμοποιούμε τις παρακάτω σχέσεις:

$P(Y > a) = F_{Y,a}$, $P(X > a) = F_{X,a}$.

Αυτή η περίπτωση μας ενδιαφέρει ιδιαίτερα σε περιπτώσεις κατανομών με βαριά ουρά (όπως έχουν οι Pareto και Lognormal) να έχουμε μία προσέγγιση που θα είναι ικανοποιητική στο δεξιό άκρο της κατανομής, θέλουμε δηλαδή τα σημεία $F_{Y,a}$, $F_{X,a}$ να είναι κοντά για $\alpha = 0.01$, $\alpha =$

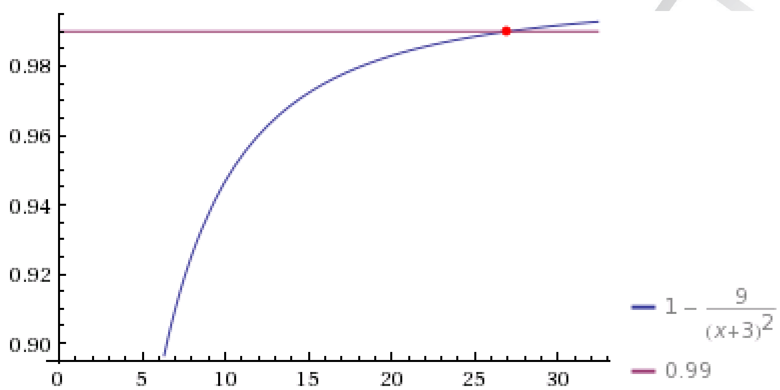
$= 0.02, \alpha = 0.03$ $\kappa.o.\kappa$ (μέθοδος 3.2.4)

Αν πάρουμε για παράδειγμα την συνάρτηση κατανομής μίας συνεχούς τυχαίας μεταβλητής X, η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 2, \beta = 3$, θα ισχύει:

$$F_X(x) = 1 - \left(\frac{3}{x+3}\right)^2, x \geq 0.$$

Για $\alpha = 0.01 \Rightarrow F_X(x) = 0.99 \Rightarrow x = 27,$

Συνεπώς το σημείο που παρατηρούμε στο παρακάτω γράφημα έχει τετμημένη 27 και τεταγμένη 0.99



Σχήμα 3.6: Διάγραμμα της $F_X(x) = 0.99$.

Για $\alpha = 0.02 \Rightarrow F_X(x) = 0.98 \Rightarrow x \approx 18.21$.

Παρατηρούμε ότι όσο αυξάνεται η τιμή του α, η τιμή του x μειώνεται, αυτό συμβαίνει διότι κάθε συνάρτηση κατανομής είναι αύξουσα συνάρτηση ως προς x και ισχύει ότι:

$$\forall f_X(x_1) < f_X(x_2) \Leftrightarrow x_1 < x_2.$$

3.3: ΠΡΟΣΕΓΓΙΣΗ ΚΑΤΑΝΟΜΗΣ PARETO ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Έχουμε μία συνεχή τυχαία μεταβλητή Y η οποία ακολουθεί την κατανομή Pareto με παραμέτρους (α,β). Στην Ενότητα 2.7 έχουμε δει ότι η συνάρτηση πυκνότητας πιθανότητας της Y θα είναι:

$$f_Y(y) = \frac{\alpha}{\beta}\left(1+\frac{y}{\beta}\right)^{-\alpha-1}, y \geq 0.$$

Με την μέθοδο των ροπών θα προσεγγίσουμε την συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ από μία μείξη δύο εκθετικών κατανομών $f_X(x)$, η οποία θα είναι της μορφής:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x \geq 0.$$

Στην ενότητα 2.7 είδαμε επίσης ότι όταν $Y \sim Pa(\alpha,\beta)$, τότε οι ροπές κ −τάξης υπολογίζονται από την σχέση:

$$E(Y^\kappa) = \frac{\beta^\kappa \kappa!}{(\alpha-1)(\alpha-2)\ldots(\alpha-\kappa)}.$$

Στην ενότητα 2.4 είδαμε ότι αν $X \sim Exp(\lambda)$, τότε ισχύει:

$$E(X) = \frac{1}{\lambda},\ \ E(X^2) = \frac{2}{\lambda^2} \text{ και } E(X^3) = \frac{6}{\lambda^3}.$$

Χρειαζόμαστε τότε ένα σύστημα τριών εξισώσεων με το οποίο θα μπορούμε να υπολογίσουμε τους αγνώστους: A, β_1, β_2 της μείξης εκθετικών κατανομών. Σύμφωνα με την μέθοδο των ροπών θα έχουμε ότι:

$$E(Y) = E(X) \Leftrightarrow \frac{\beta}{\alpha-1} = \frac{A}{\beta_1} + \frac{1-A}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow \frac{2\beta^2}{(\alpha-1)(\alpha-2)} = \frac{2A}{{\beta_1}^2} + \frac{2(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow \frac{6\beta^3}{(\alpha-1)(\alpha-2)(\alpha-3)} = \frac{6}{{\beta_1}^3} + \frac{6(1-A)}{{\beta_2}^3}.$$

Υπολογίζουμε τους αγνώστους της μείξης κατανομών και θα έχουμε ότι:

$$f_Y(x) \approx f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x \geq 0.$$

Σε αυτό το σημείο πρέπει να αναφέρουμε ότι η προσέγγιση μίας συνάρτησης πυκνότητας πιθανότητας $f_Y(y)$ βελτιώνεται όσο αυξάνεται το πλήθος των παραμέτρων της μείξης εκθετικών κατανομών $f_X(x)$. Θα μπορούσαμε για παράδειγμα να προσεγγίσουμε την κατανομή Pareto από μία μείξη εκθετικών κατανομών τριών παραμέτρων: $\beta_1, \beta_2, \beta_3$ με αντίστοιχα βάρη A_1, A_2, A_3: $A_1 + A_2 + A_3 = 1$, που είναι δηλαδή της μορφής:

$$f_X(x) = (1 - A_2 - A_3)\beta_1 e^{-\beta_1 x} + A_2 \beta_2 e^{-\beta_2 x} + A_3 \beta_3 e^{-\beta_3 x}, x \geq 0.$$

Έτσι θα είχαμε να υπολογίσουμε πέντε αγνώστους (έναντι τριών στην περίπτωση που χρησιμοποιούμε), τους: $\beta_1, \beta_2, \beta_3, A_2, A_3$ και θα χρειαζόμασταν ένα σύστημα πέντε εξισώσεων,το οποίο προκύπτει από την μέθοδο των ροπών. Επίσης επειδή θα χρησιμοποιούσαμε έως και την πέμπτη ροπή της Pareto,θα έπρεπε να ίσχυε ότι: $a > 5$. Θα είχαμε λοιπόν να λύσουμε το παρακάτω σύστημα.

$$E(Y) = E(X) \Leftrightarrow \frac{\beta}{\alpha-1} = \frac{1-A_2-A_3}{\beta_1} + \frac{A_2}{\beta_2} + \frac{A_3}{\beta_3},$$

$$E(Y^2) = E(X^2) \Leftrightarrow \frac{2\beta^2}{(\alpha-1)(\alpha-2)} = \frac{2(1-A_2-A_3)}{{\beta_1}^2} + \frac{2A_2}{{\beta_2}^2} + \frac{2A_3}{{\beta_3}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow \frac{6\beta^3}{(\alpha-1)(\alpha-2)(\alpha-3)} = \frac{6(1-A_2-A_3)}{{\beta_1}^3} + \frac{6A_2}{{\beta_2}^3} + \frac{6A_3}{{\beta_3}^3},$$

$$E(Y^4) = E(X^4) \Leftrightarrow \frac{24\beta^4}{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)} = \frac{24(1-A_2-A_3)}{{\beta_1}^4} + \frac{24A_2}{{\beta_2}^4} + \frac{24A_3}{{\beta_3}^4},$$

$$E(Y^5) = E(X^5) \Leftrightarrow \frac{120\beta^5}{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)(\alpha-5)} = \frac{120(1-A_2-A_3)}{{\beta_1}^5} + \frac{120A_2}{{\beta_2}^5} + + \frac{120A_3}{{\beta_3}^5}.$$

Με την συγκεκριμένη όμως περίπτωση δε θα ασχοληθούμε στην παρούσα εργασία.

Ας δούμε ένα παράδειγμα προσέγγισης μίας συνεχούς τυχαίας μεταβλητής, η οποία ακολουθεί την κατανομή Pareto, από μία μείξη εκθετικών κατανομών με δύο παραμέτρους.

ΠΑΡΑΔΕΙΓΜΑ 3.3.1

Θέλουμε να προσεγγίσουμε από μία μείξη δύο εκθετικών κατανομών με παραμέτρους $\beta_1, \beta_2 > 0$ και A_1, A_2: $A_1 + A_2 = 1$, την συνάρτηση πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής Y, η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 6, \beta = 2$ $\left(Y \sim Pa(6,2)\right)$.

Έχουμε ότι $Y \sim Pa(6,2)$. Στην ενότητα 2.7 είδαμε ότι η συνάρτηση πυκνότητας πιθανότητας της Y θα είναι:

$$f_Y(y) = 3\left(1 + \frac{y}{2}\right)^{-7}, y \geq 0.$$

Η μείξη εκθετικών κατανομών θα είναι της μορφής:

$$f_X(x) = A_1\beta_1 e^{-\beta_1 x} + A_2\beta_2 e^{-\beta_2 x} = A_1\beta_1 e^{-\beta_1 x} + (1 - A_1)\beta_2 e^{-\beta_2 x}, x \geq 0 \ .$$

Οι ροπές πρώτης, δεύτερης και τρίτης τάξης της Y θα είναι:

$$E(Y) = \frac{2}{6-1} = \frac{2}{5}, \quad E(Y^2) = \frac{2^2 \cdot 2!}{(6-1)(6-2)} = \frac{2}{5}, \quad , \quad E(Y^3) = \frac{2^3 \cdot 3!}{(6-1)(6-2)(6-3)} = \frac{4}{5}.$$

Σύμφωνα με την μέθοδο των ροπών θα έχουμε ότι:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{A_1}{\beta_1} + \frac{1-A_1}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow \frac{2}{5} = \frac{2A_1}{{\beta_1}^2} + \frac{2(1-A_1)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow \frac{4}{5} = \frac{6A_1}{{\beta_1}^3} + \frac{6(1-A_1)}{{\beta_2}^3}.$$

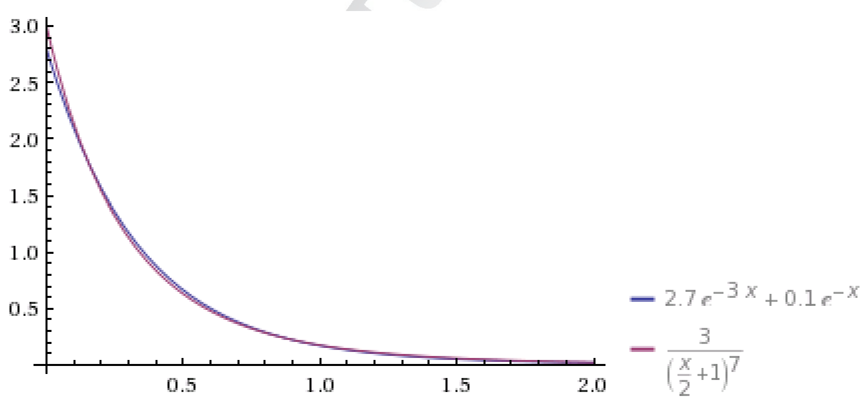
Με χρήση του Mathematica βρίσκουμε ότι:

$$\beta_1 = 1\,, \quad \beta_2 = 3\,, \quad A = 0.1\,, \ \mathrm{A}_2 = 0.9\,.$$

Επομένως θα ισχύει ότι:

$$f_Y(x) \approx f_X(x) = 0.1 \cdot \mathrm{e}^{-x} + 0.9 \cdot 3e^{-3x} = 0.1e^{-x} + 2.7e^{-3x}, x \geq 0.$$

Στο παρακάτω σχήμα βλέπουμε τις γραφικές παραστάσεις των δύο συναρτήσεων πυκνότητας πιθανότητας και διαπιστώνουμε ότι είναι παρόμοιες.



Σχήμα 3.7: Διάγραμμα των $f_X(x)$ (μπλε χρώμα), $f_Y(x)$ (ροζ χρώμα).

Στην ενότητα 2.7 είδαμε ότι αν $Y \sim Pa(6,2)$, τότε η συνάρτηση κατανομής της είναι της μορφής:

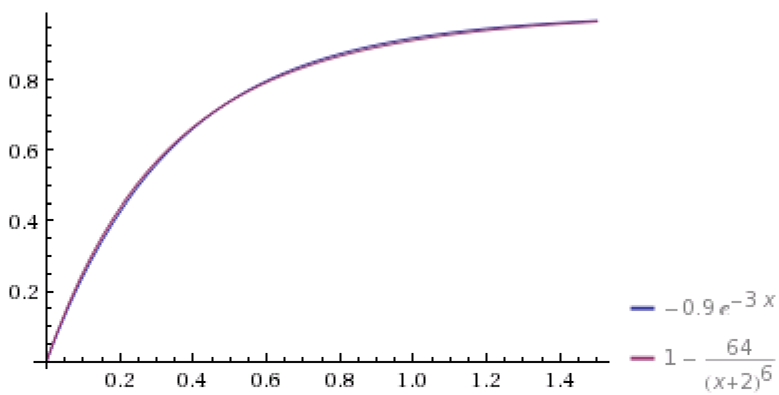
$$F_Y(y) = 1 - \left(\frac{2}{y+2}\right)^6, y \geq 0.$$

Η συνάρτηση κατανομής της μείξης εκθετικών κατανομών υπολογίζεται από την σχέση:

$$F_X(x) = \int_0^x f(\xi)d\xi = \int_0^x (0.1 \cdot e^{-\xi} + 2.7 \cdot e^{-4.312\xi})d\xi \Rightarrow$$

$$\Rightarrow F_X(x) = 1 - 0.1e^{-x} - 0.9e^{-3x}, x \geq 0.$$

Στο σχήμα 3.3.2 παρατηρούμε τις γραφικές παραστάσεις των δύο συναρτήσεων κατανομής $F_X(x), F_Y(x)$.



Σχήμα 3.8: Διάγραμμα των $F_X(x)$ (μπλε χρώμα), $F_Y(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ.

Για να υπολογίσουμε την «απόσταση» των δύο συναρτήσεων πυκνότητας πιθανότητας θα υπολογίσουμε το ολοκλήρωμα της απόλυτης διαφοράς τους και θα ισχύει ότι:

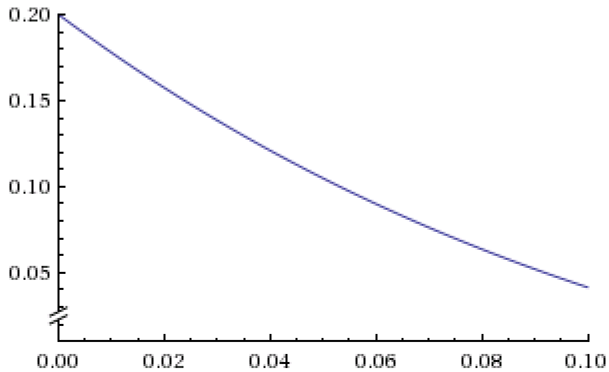
$$\int_0^\infty |f_Y(x) - f_X(x)| dx = \int_0^\infty \left| 3\left(1 + \frac{x}{2}\right)^{-7} - 0.1e^{-x} - 2.7e^{-3x} \right| dx \approx 0.039.$$

Θα υπολογίσουμε την μέγιστη διαφορά των δύο συναρτήσεων από την παρακάτω σχέση:

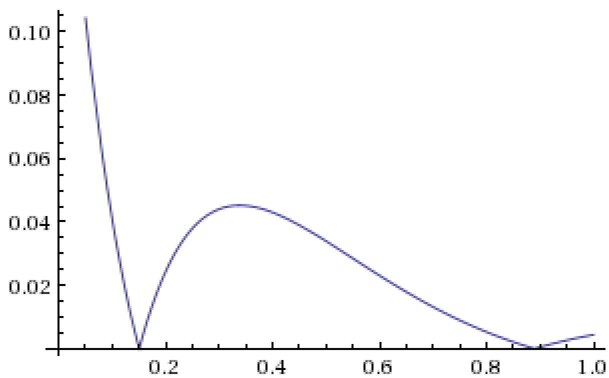
$$sup_{x\geq 0}|f_Y(x) - f_X(x)| = sup_{x\geq 0}\left| 3\left(1 + \frac{x}{2}\right)^{-7} - 0.1e^{-x} - 2.7e^{-3x} \right| = 0.2,$$

στο σημείο $x_0 = 0$.

Στα δύο παρακάτω γραφήματα βλέπουμε την γραφική παράσταση της απόλυτης διαφοράς $|f_Y(x) - f_X(x)|$ για $x \in [0,0.1]$ και για $x \in [0,1]$ αντίστοιχα και παρατηρούμε την παραπάνω διαφορά στο πρώτο γράφημα ενώ στο δεύτερο παρατηρείται ένα άλλο τοπικό μέγιστο το οποίο έχει τιμή 0.045 στο σημείο $x_1 \approx 0.338$. .



Σχήμα 3.9: Διάγραμμα του $|f_Y(x) - f_X(x)|$, για $x \in [0,0.1]$.



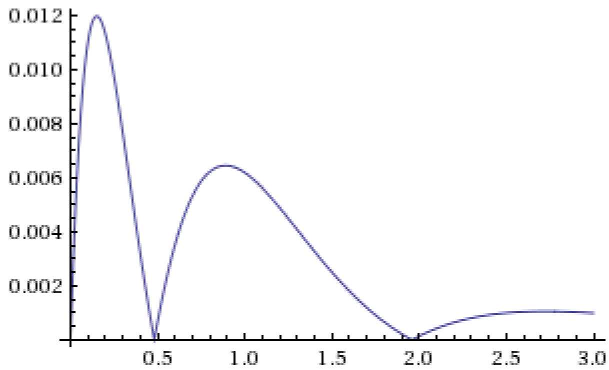
Σχήμα 3.10: Διάγραμμα του $|f_Y(x) - f_X(x)|$, για $x \in [0,1]$.

Επίσης θα υπολογίσουμε και την μέγιστη διαφορά των δύο συναρτήσεων κατανομής $F_Y(x)$ και $F_X(x)$ παρομοίως από την σχέση:

$$sup_{x\geq 0}|F_Y(x) - F_X(x)| = sup_{x\geq 0}\left|1 - \left(\frac{2}{x+2}\right)^6 - 1 + 0.1e^{-x} + 0.9e^{-3x}\right| = 0.0119749,$$

στο σημείο $x_0 \approx 0.150944$.

Στο παρακάτω σχήμα μπορούμε να δούμε την γραφική παράσταση του $|F_Y(x) - F_X(x)|$. Παρατηρούνται επίσης άλλα δύο τοπικά μέγιστα και με την χρήση του Mathematica βρίσκουμε ότι είναι ίσα με 0.00105676 και 0.00645562 στα σημεία $x_1 \approx 2.71916, x_2 \approx 0.889763$ αντίστοιχα.



Σχήμα 3.11: Διάγραμμα του $|F_Y(x) - F_X(x)|$.

Χρησιμοποιώντας την μέθοδο 3.2.3, θα εξετάσουμε πόσο «κοντά» είναι οι ροπές ανώτερης τάξης από αυτές που έχουμε χρησιμοποιήσει, όπως για παράδειγμα οι ροπές τέταρτης τάξης. Έχουμε τότε ότι:

$$E(Y^4) = \frac{24 \cdot 16}{(6-1)(6-2)(6-3)(6-4)} = \frac{16}{5} = 3.2.$$

$$E(X^4) = \int_0^{\infty} x^5 f_X(x) dx = 2.66667.$$

Επίσης χρησιμοποιώντας την μέθοδο 3.2.4 θα υπολογίσουμε τα άνω ποσοστημόρια των δύο συναρτήσεων κατανομής. Έχουμε συνεπώς ότι:

$$P(Y > a) = F_{Y,a}, \qquad P(X > a) = F_{X,a}.$$

Για $\alpha = 0.001$, έχουμε:

$$F_X(x) = 0.999 \qquad F_Y(y) = 0.999,$$

Με χρήση του Mathematica βρίσκουμε ότι: $x \approx 4.60607$, $y \approx 4.32456$. Το ποσοστό σφάλματος βρίσκεται ως εξής:

$$\frac{y-x}{y}\% = \frac{4.32456-4.60607}{4.32456}\% \approx -6.51\%.$$

Για $\alpha = 0.005$, έχουμε:

$$F_X(x) = 0.995, \qquad F_Y(y) = 0.995$$

Και βρίσκουμε ότι: $x \approx 3.01706$, $y \approx 2.83654$. Το ποσοστό σφάλματος θα είναι:

$$\frac{y-x}{y}\% = \frac{2.83654-3.01706}{2.83654}\% \approx -6.36\%.$$

Κατ'αυτόν τον τρόπο συνεπώς δημιουργούμε τον παρακάτω πίνακα:

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό σφάλματος
0.001	0.999	0.999	4.60607	4.32456	−6.51%
0.005	0.995	0.995	3.01706	2.83654	−6.36%
0.01	0.99	0.99	2.37724	2.30887	−2.96%
0.02	0.98	0.98	1.82111	1.83877	0.96%
0.03	0.97	0.97	1.5464	1.58792	2.61%

Πίνακας 3.1: Πίνακας άνω ποσοστημορίων των $F_X(x), F_Y(y)$.

3.4: ΠΡΟΣΕΓΓΙΣΗ ΚΑΤΑΝΟΜΗΣ WEIBULL ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Στην Ενότητα 2.6, είδαμε πως όταν μία συνεχής τυχαία μεταβλητή Y ακολουθεί την κατανομή Weibull με παραμέτρους λ, α, η συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ θα είναι της μορφής:

$$f_Y(y) = \alpha\lambda^\alpha y^{\alpha-1}e^{-(\lambda y)^\alpha},\ y > 0.$$

Με την μέθοδο των ροπών θα προσεγγίσουμε την συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ της μεταβλητής Y από μία μείξη δύο εκθετικών κατανομών $f_X(x)$, η οποία θα είναι της μορφής:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x \geq 0.$$

Οι παράμετροι λ, α της μεταβλητής Y θα είναι γνωστοί, επομένως θα πρέπει να υπολογίσουμε τους αγνώστους: A, β_1, β_2, και χρειαζόμαστε ένα σύστημα τριών εξισώσεων, το οποίο προκύπτει από την μέθοδο των ροπών και είναι το παρακάτω:

$$E(Y) = E(X),$$

$$E(Y^2) = E(X^2),$$

$$E(Y^3) = E(X^3).$$

Στην ενότητα 2.6 έχουμε δει ότι αν $Y \sim Wei(\lambda, \alpha)$, τότε οι ροπές κ −τάξης προκύπτουν από την σχέση:

$$E(Y^\kappa) = \frac{1}{\lambda^\kappa}\Gamma\left(\frac{\kappa+\alpha}{\alpha}\right).$$

Επομένως θα έχουμε:

$$E(Y) = E(X) \Leftrightarrow \frac{1}{\lambda}\Gamma\left(\frac{1+\alpha}{\alpha}\right) = \frac{A}{\beta_1} + \frac{1-A}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow \frac{1}{\lambda^2}\Gamma\left(\frac{2+\alpha}{\alpha}\right) = \frac{2A}{{\beta_1}^2} + \frac{2(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow \frac{1}{\lambda^3}\Gamma\left(\frac{3+\alpha}{\alpha}\right) = \frac{6A}{{\beta_1}^3} + \frac{6(1-A)}{{\beta_2}^3}.$$

Τότε με τον υπολογισμό των παραμέτρων β_1, β_2 , με αντίστοιχα βάρη $A, 1-A,$ η μείξη δύο εκθετικών κατανομών θα είναι:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x > 0 \approx f_Y(y).$$

Στο παρακάτω παράδειγμα θα προσεγγίσουμε την συνάρτηση πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής $Y \sim Wei(\lambda, \alpha)$, από μία μείξη δύο εκθετικών κατανομών.

ΠΑΡΑΔΕΙΓΜΑ 3.4.1

Για να είναι καλύτερη η σύγκριση μεταξύ των προσεγγίσεων Pareto και Weibull, είναι προτιμότερο οι μέσες τιμές των συναρτήσεων πυκνότητάς τους να είναι ίσες. Στο παράδειγμα που είδαμε πριν με την κατανομή Pareto ίσχυε: $E(Y) = \frac{2}{5}$. Θα επιλέξουμε επομένως την ίδια μέση τιμή και την παράμετρο $\alpha = 0.9$ και θα υπολογίσουμε έτσι με χρήση του Mathematica την πρώτη παράμετρο λ και έπειτα θα προσεγγίσουμε την κατανομή Weibull από μία μείξη εκθετικών κατανομών με δύο παραμέτρους β_1, β_2 και αντίστοιχα βάρη $A, 1-A$.

Στην συγκεκριμένη περίπτωση έχουμε μία συνεχή τυχαία μεταβλητή Y, για την οποία ισχύει ότι: $Y \sim Wei(\lambda, 0.9)$ και $E(Y) = \frac{2}{5}$. Επομένως ισχύει:

$$\frac{1}{\lambda}\Gamma\left(\frac{0.9+1}{0.9}\right) = \frac{2}{5} \Rightarrow \lambda = 2.63045.$$

Επίσης για τις ροπές δεύτερης και τρίτης τάξης θα ισχύει αντίστοιχα:

$$E(Y^2) = \frac{1}{2.63045^2}\Gamma\left(\frac{2.9}{0.9}\right) = 0.358, \qquad E(Y^3) = \frac{1}{2.63045^3}\Gamma\left(\frac{3.9}{0.9}\right) = 0.509.$$

Σύμφωνα με την μέθοδο των ροπών οι τρεις εξισώσεις που χρειαζόμαστε για την εύρεση των A, β_1, β_2, είναι οι παρακάτω:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{A}{\beta_1} + \frac{1-A}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow 0.358 = \frac{2A}{{\beta_1}^2} + \frac{2(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow 0.509 = \frac{6A}{{\beta_1}^3} + \frac{6(1-A)}{{\beta_2}^3}.$$

Με την χρήση του Mathematica βρίσκουμε ότι:

$$A \approx 0.324247, \qquad \beta_1 \approx 4.97489, \qquad \beta_2 \approx 2.01824.$$

Προφανώς τότε η μείξη εκθετικών κατανομών θα γράφεται στην μορφή:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x} =$$

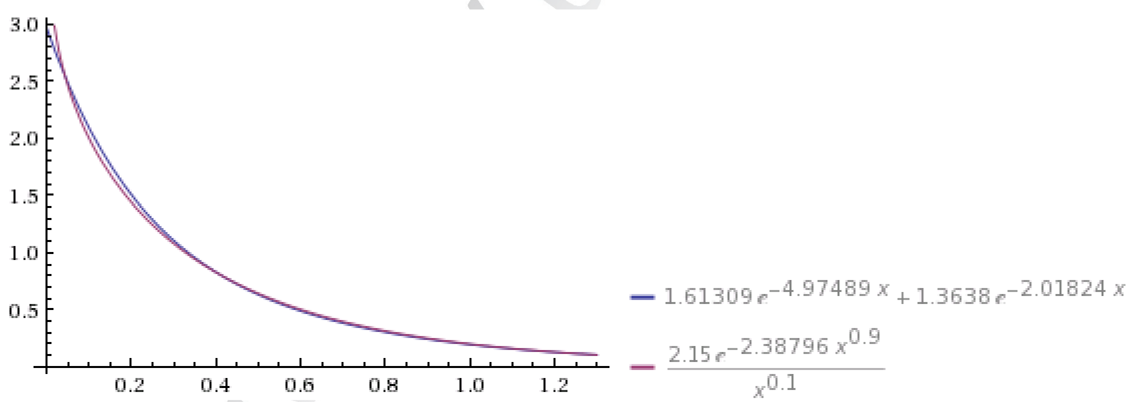
$$\approx 0.324247 \cdot 4.97489 e^{-4.97489x} + 0.675753 \cdot 2.01824 e^{-2.01824x}$$

$$\approx 1.613093 e^{-4.97489x} + 1.3638 e^{-2.01824x}, x > 0.$$

Στην ενότητα 2.6 είδαμε ότι αν $Y \sim Wei(2.63045{,}0.9)$, τότε η συνάρτηση πυκνότητας πιθανότητάς της θα είναι της μορφής:

$$f_Y(y) = 0.9 \cdot 2.63045^{0.9} y^{-0.1} e^{-(2.63045y)^{0.9}} \approx 2.15 y^{-0.1} e^{-(2.63045y)^{0.9}}, y > 0.$$

Στο παρακάτω γράφημα βλέπουμε τις γραφικές παραστάσεις των συναρτήσεων $f_X(x)$ και $f_Y(x)$ και παρατηρούμε ότι είναι πάρα πολύ κοντά η μία στην άλλη.



Σχήμα 3.12: Διάγραμμα των $f_X(x)$ (μπλε χρώμα), $f_Y(x)$ (ροζ χρώμα).

Είναι χρήσιμο να αναφερθεί ότι η $f_Y(x)$, έχει κατακόρυφη ασύμπτωτη τον άξονα $y'y$, διότι ισχύει:

$$\lim_{x \to 0^+} f_Y(x) = \lim_{x \to 0^+} 2.15 x^{-0.1} e^{-(2.63045x)^{0.9}} = \infty.$$

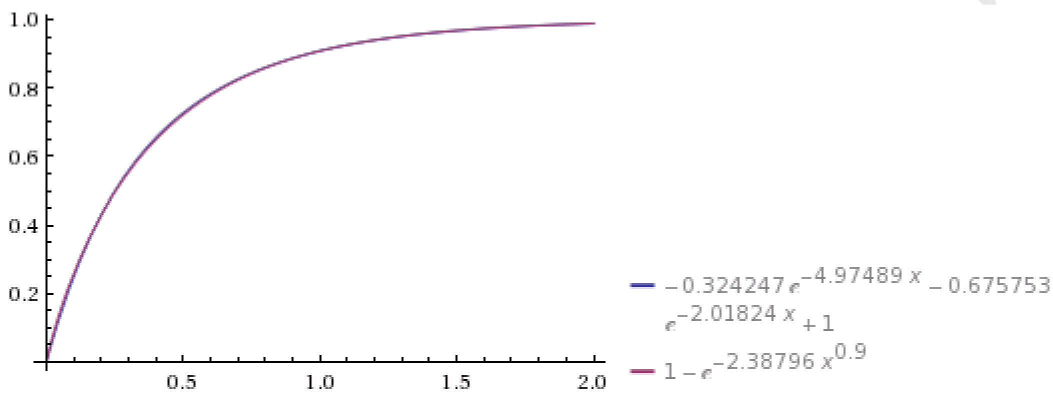
Η συνάρτηση κατανομής της $f_X(x)$ θα είναι:

$F_X(x) \approx 1 - 0.324247e^{-4.97489x} - 0.675753e^{-2.01824x}, x \geq 0.$

Η συνάρτηση κατανομής της $f_Y(y)$ είδαμε στην ενότητα 2.6 ότι θα είναι της μορφής:

$F_Y(y) = 1 - e^{-(2.63045y)^{0.9}}, y > 0.$

Η γραφική παράσταση των $F_X(x), F_Y(y)$ απεικονίζεται στο παρακάτω γράφημα:



Σχήμα 3.13: Διάγραμμα των $F_X(x)$ (μπλε χρώμα), $F_Y(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ ΠΡΟΣΕΓΓΙΣΗΣ

Εξηγήσαμε παραπάνω ότι η $f_Y(x)$ έχει κατακόρυφη ασύμπτωτη την ευθεία $x = 0$, επομένως δεν έχει νόημα να υπολογίσουμε την μέγιστη απόσταση των δύο πυκνοτήτων, διότι θα είναι άπειρη «κοντά» στο 0. Σύμφωνα με την μέθοδο (3.2.2) η «απόσταση» των δύο συναρτήσεων πυκνότητας πιθανότητας θα είναι ίση με:

$\int_0^\infty |f_Y(x) - f_X(x)| dx \approx 0.038.$

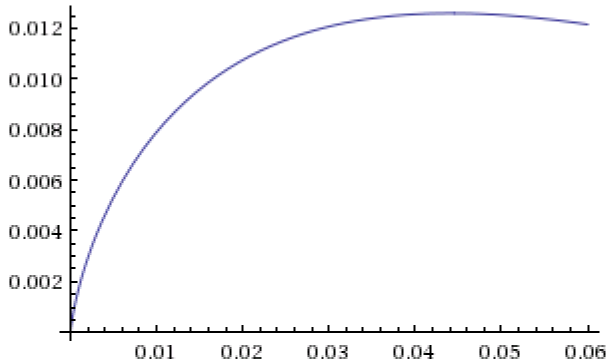
Θα υπολογίσουμε επίσης την μέγιστη διαφορά των δύο συναρτήσεων κατανομής. Συγκεκριμένα με χρήση του Mathematica προκύπτει ότι:

$sup_{x>0}|F_Y(x) - F_X(x)| =$

$= \left|1 - 0.324247e^{-4.97489x} - 0.675753e^{-2.01824x} - 1 + e^{-(2.63045x)^{0.9}}\right| \approx$

≈ 0.0125938 στο σημείο $x_0 \approx 0.044083.$

Στο παρακάτω γράφημα βλέπουμε την γραφική παράσταση του $|F_Y(x) - F_X(x)|$ για $x \in [0,0.06]$ και παρατηρούμε την συγκεκριμένη απόσταση.



Σχήμα 3.14: Διάγραμμα του $|F_Y(x) - F_X(x)|$.

Θα συγκρίνουμε τις ροπές τέταρτης τάξης για να δούμε πόσο κοντά είναι μεταξύ τους και θα ισχύει:

$$E(Y^4) \approx 0.999998, \qquad E(X^4) \approx 0.990126,$$

Έπειτα θα υπολογίσουμε τα άνω ποσοστημόρια των δύο συναρτήσεων κατανομής σύμφωνα με την μέθοδο 3.2.4. Θα ισχύει τότε:

$$P(Y > a) = F_{Y,a}, \qquad P(X > a) = F_{X,a}.$$

Για $\alpha = 0.001$ έχουμε ότι: $F_X(x) = 0.999$, $F_Y(y) = 0.999$ και βρίσκουμε τότε ότι: $x \approx 3.236$, $y \approx 3.255$. Όπως και στην προηγούμενη ενότητα το ποσοστό λάθους θα προκύπτει από την σχέση:

$$\frac{y-x}{y}\% = \frac{3.255-3.228}{3.255}\% \approx 0.83\%.$$

Ο παρακάτω πίνακας περιέχει συνοπτικά αυτά τα αποτελέσματα για διάφορες τιμές του α. Αν συγκρίνουμε τις τιμές των x με των y, με εκείνες του Πίνακα 3.1, θα διαπιστώσουμε ότι παρουσιάζουν μικρότερη «απόκλιση».

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό Σφάλματος
0.001	0.999	0.999	3.228	3.255	0.83%
0.005	0.995	0.995	2.431	2.424	−0.28%
0.01	0.99	0.99	2.088	2.074	−0.67%
0.02	0.98	0.98	1.745	1.731	−0.81%
0.03	0.97	0.97	1.546	1.532	−0.91%

Πίνακας 3.2: Πίνακας άνω ποσοστημορίων των δύο συναρτήσεων κατανομής.

Πρέπει να αναφέρουμε ότι ο λόγος που σε αυτό το παράδειγμα χρησιμοποιήσαμε μία συνεχή μεταβλητή Y, η οποία ακολουθεί την κατανομή Weibull με παραμέτρους $\lambda = 2.63045$, $\alpha = 0.9$, είναι ότι για τιμές της παραμέτρου α μεγαλύτερες της μονάδας, ήταν αρκετά δύσκολο να επιτευχθεί προσέγγιση, διότι κάναμε πολλαπλές δοκιμές στο Mathematica και οι λύσεις των αγνώστων: A, β_1, β_2, ήταν είτε μιγαδικοί αριθμοί, είτε οι τιμές του A δεν ήταν ανάμεσα σε 0 και 1.

3.5: ΠΡΟΣΕΓΓΙΣΗ ΛΟΓΑΡΙΘΜΟΚΑΝΟΝΙΚΗΣ (LOGNORMAL) ΚΑΤΑΝΟΜΗΣ ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Στην λογαριθμοκανονική κατανομή ισχύει ότι αν $Y \sim LN(\mu, \sigma)$, η συνάρτηση πυκνότητας πιθανότητας της συνεχούς τυχαίας μεταβλητής Y, γράφεται στην μορφή:

$$f_Y(y) = \begin{cases} \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\log y - \mu)^2}{2\sigma^2}} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Θα προσεγγίσουμε συνεπώς την $f_Y(y)$ από μία μείξη δύο εκθετικών κατανομών, η οποία θα είναι της μορφής:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1 - A)\beta_2 e^{-\beta_2 x}, x \geq 0.$$

Στην Ενότητα 2.8 περιγράψαμε την λογαριθμοκανονική κατανομή και είδαμε ότι η σχέση:

$$E(Y^\kappa) = e^{\mu\kappa + \frac{1}{2}\sigma^2\kappa^2},$$

υπολογίζει τις ροπές $\kappa -$τάξης της τυχαίας συνεχούς μεταβλητής Y.

Σύμφωνα τότε με την μέθοδο των ροπών θα έχουμε ότι:

$$E(Y) = E(X) \Leftrightarrow e^{\mu + \frac{1}{2}\sigma^2} = \frac{A}{\beta_1} + \frac{1-A}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow e^{2\mu + 2\sigma^2} = \frac{2A}{{\beta_1}^2} + \frac{2(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow e^{3\mu + \frac{9}{2}\sigma^2} = \frac{6A}{{\beta_1}^3} + \frac{6(1-A)}{{\beta_2}^3}.$$

Με τον υπολογισμό των παραμέτρων β_1, β_2 και των αντίστοιχων βαρών $A, 1 - A$, βρίσκουμε τον τύπο της μείξης δύο εκθετικών κατανομών, η οποία αποτελεί και την προσέγγιση της $f_Y(y)$.

ΠΑΡΑΔΕΙΓΜΑ 3.5.1

Όπως και στην προηγούμενη Ενότητα, για την προσέγγιση της κατανομής Weibull, για να είναι εύκολη η σύγκριση των αποτελεσμάτων Pareto, Weibull και Lognormal, θα επιλέξουμε η μέση τιμή της συνεχούς τυχαίας μεταβλητής Y, η οποία ακολουθεί την λογαριθμοκανονική κατανομή, να είναι ίση με $\frac{2}{5}$. Επίσης θα επιλέξουμε η δεύτερη παράμετρος σ, να είναι ίση με 1. Συνεπώς θα ισχύει : $Y \sim LN(\mu, 1)$. Θα προσεγγίσουμε την συνάρτηση πυκνότητας πιθανότητας της $f_Y(y)$, από μία μείξη εκθετικών κατανομών $f_X(x)$, δύο παραμέτρων β_1, β_2, με αντίστοιχα βάρη: $A, 1-A$.

Στην ενότητα 2.8 περιγράψαμε αναλυτικά την λογαριθμοκανονική κατανομή και είδαμε ότι αν $Y \sim LN(\mu, \sigma)$, τότε ροπές κ −τάξης υπολογίζονται από την σχέση:

$$E(Y^\kappa) = e^{\mu\kappa + \frac{1}{2}\sigma^2\kappa^2},$$

επομένως στην περίπτωσή μας θα ισχύει ότι:

$$E(Y) = \frac{2}{5} \Rightarrow e^{\mu + \frac{1}{2}} = \frac{2}{5},$$

Με χρήση του Mathematica βρίσκουμε ότι: $\mu = -1.41629$.

Επίσης για τις ροπές δεύτερης και τρίτης τάξης θα ισχύει ότι:

$$E(Y^2) = e^{-1.41629 \cdot 2 + 2} \approx 0.434926,$$

$$E(Y^3) = e^{-1.27754 \cdot 3 + \frac{9}{2}} \approx 1.28548.$$

Οι τρεις εξισώσεις που θα χρειαστούμε για την εύρεση των παραμέτρων και τα βάρη της μείξης εκθετικών κατανομών, σύμφωνα με την μέθοδο των ροπών είναι οι παρακάτω:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{A}{\beta_1} + \frac{1-A}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow 0.434926 = \frac{2A}{{\beta_1}^2} + \frac{2(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow 1.28548 = \frac{6A}{{\beta_1}^3} + \frac{6(1-A)}{{\beta_2}^3}.$$

Με χρήση του Mathematica βρίσκουμε ότι:

$$A \approx 0.02645, \quad \beta_1 \approx 0.539321, \quad \beta_2 \approx 2.77405,$$

επομένως η μείξη εκθετικών κατανομών θα είναι της μορφής:

$$f_X(x) = A\beta_1 e^{-\beta_1 x} + (1-A)\beta_2 e^{-\beta_2 x}, x > 0 \approx$$

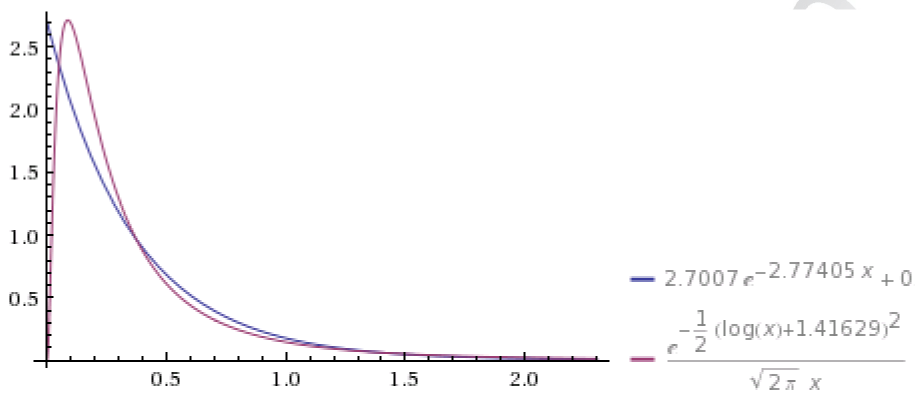
$$\approx 0.02645 \cdot 0.539321 e^{-0.539321x} + (1 - 0.02645) 2.77405 e^{-2.77405x}, x > 0$$

$$\approx 0.0143 e^{-0.539321x} + 2.7007 e^{-2.77405x}, x > 0.$$

Η συνάρτηση πυκνότητας πιθανότητας της Y, από την ενότητα 2.8 βλέπουμε ότι θα γράφεται ως:

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{(logy + 1.41629)^2}{2}\right), y > 0.$$

Στο παρακάτω γράφημα αποτυπώνονται οι γραφικές παραστάσεις των δύο συναρτήσεων πυκνότητας πιθανότητας.



Σχήμα 3.15: : Διάγραμμα των $f_X(x)$ (μπλε χρώμα), $f_Y(x)$ (ροζ χρώμα).

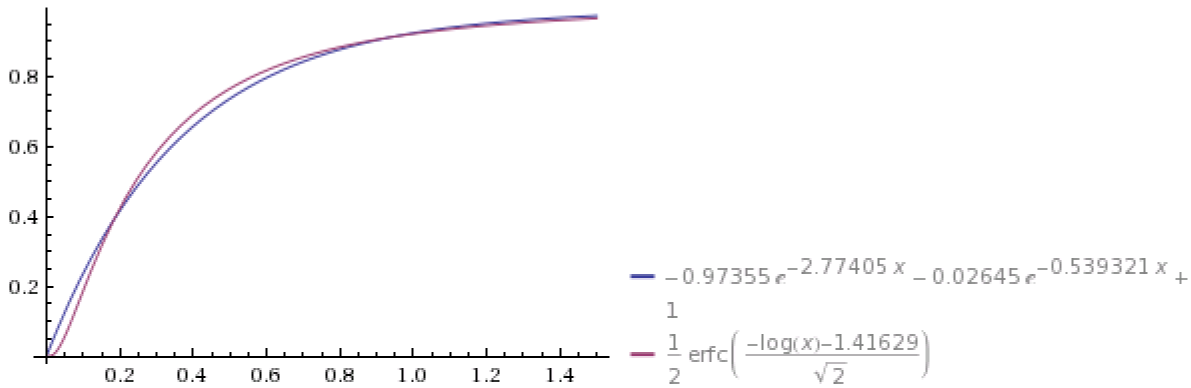
Η συνάρτηση κατανομής της Y, θα είναι:

$$F_Y(y) = \Phi(lny + 1.41629), y > 0,$$

ενώ της μεταβλητής X, θα είναι:

$$F_X(x) \approx 1 - 0.97355 e^{-2.77405x} - 0.02645 e^{-0.539321x}, x > 0.$$

Στο παρακάτω σχήμα βλέπουμε τις γραφικές παραστάσεις τους:



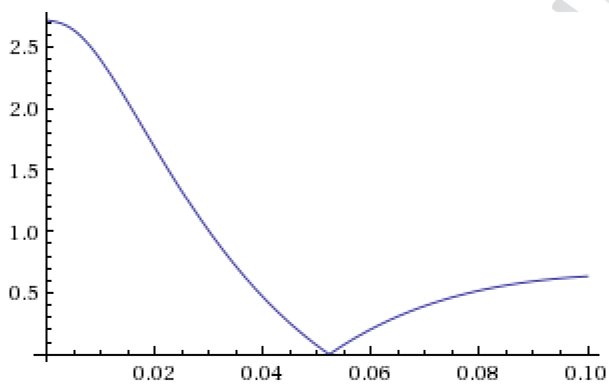
Σχήμα 3.16: Διάγραμμα των $F_X(x)$ (μπλε χρώμα), $F_Y(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ

Η μέγιστη διαφορά των δύο πυκνοτήτων προκύπτει από την σχέση:

$sup_{x>0}|f_Y(x) - f_X(x)| \approx 2.7$ στο σημείο $x_0 \approx 0$.

Στο παρακάτω γράφημα βλέπουμε την γραφική παράσταση του $|f_Y(x) - f_X(x)|$ και απεικονίζεται και η παραπάνω διαφορά στο συγκεκριμένο σημείο .

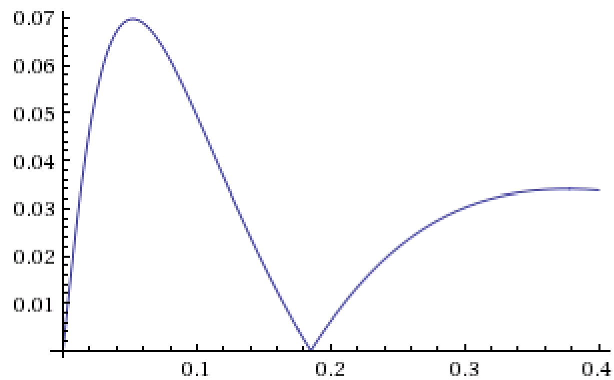


Σχήμα 3.17: Γραφική παράσταση της απόλυτης διαφοράς των δύο συναρτήσεων πυκνότητας πιθανότητας.

Αντίστοιχα η μέγιστη διαφορά των δύο συναρτήσεων κατανομής θα είναι:

$sup_{x>0}|F_Y(x) - F_X(x)| \approx 0.069755,$

Στο σημείο $x_0 \approx 0.0522828$. Στο παρακάτω σχήμα βλέπουμε την γραφική παράσταση του $|F_Y(x) - F_X(x)|$:



Σχήμα 3.18: Γραφική παράσταση της απόλυτης διαφοράς των δύο συναρτήσεων κατανομής.

Θα υπολογίσουμε το ολοκλήρωμα της απόλυτης διαφοράς των δύο συναρτήσεων πυκνότητας πιθανότητας, όπως και στις δύο προηγούμενες ενότητες και θα έχουμε ότι:

$$\int_0^\infty |f_Y(x) - f_X(x)| dx =$$

$$= \int_0^\infty \left| \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{(logx + 1.41629)^2}{2}\right) - 0.0143e^{-0.539321x} - 2.7007e^{-2.77405x} \right| dx \approx$$

$$\approx 0.223.$$

Προφανώς οι ροπές πρώτης, δεύτερης και τρίτης τάξης είναι ίσες. Ας δούμε όμως την διαφορά των ροπών τέταρτης τάξης. Έχουμε συνεπώς ότι:

$$E(Y^4) = e^{-1.41629 \cdot 4 + 8} \approx 10.3278,$$

$$E(X^4) = \int_0^\infty x^4 (0.0143e^{-0.539321x} + 2.7007e^{-2.77405x}) dx \approx 7.91617.$$

Όπως και στις δύο προηγούμενες Ενότητες, θα κατασκευάσουμε έναν πίνακα για να συγκρίνουμε τα άνω ποσοστημόρια των δύο συναρτήσεων κατανομής. Θα έχουμε τότε ότι:

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό Σφάλματος
0.001	0.999	0.999	6.073	5.333	−13.9%
0.005	0.995	0.995	3.147	3.188	1.27%
0.01	0.99	0.99	2.224	2.484	10.5%
0.02	0.98	0.98	1.677	1.891	11.3%
0.03	0.97	0.97	1.442	1.591	9.4%

Πίνακας 3.3: Πίνακας άνω ποσοστημορίων των $F_X(x), F_Y(y)$.

Όπως και στην προηγούμενη ενότητα για την προσέγγιση της κατανομής Weibull, έτσι και εδώ πρέπει να αναφέρουμε ότι ο λόγος που η μεταβλητή Y ακολουθεί την λογαριθμοκανονική κατανομή με παραμέτρους $\mu = -1.41629, \sigma = 1,$ είναι ότι έπειτα από αρκετές δοκιμές στο Mathematica, για τιμές της παραμέτρου σ μικρότερες από 0.8 η προσέγγιση παρουσίαζε προβλήματα (οι τιμές των A, β_1, β_2 που έδινε το Mathematica είτε ήταν συνήθως μιγαδικοί αριθμοί, είτε το A δεν ήταν ανάμεσα σε 0 και 1) καθώς και για τιμές μεγαλύτερες από 1 η προσέγγιση δεν ήταν ικανοποιητική καθώς το ολοκλήρωμα της απόλυτης διαφοράς των δύο συναρτήσεων πυκνότητας πιθανότητας είχε τιμές πιο κοντά στο 2.

3.6: ΜΕΙΞΕΙΣ ΓΑΜΜΑ ΚΑΤΑΝΟΜΩΝ ΜΕ ΠΑΡΑΜΕΤΡΟΥΣ $\alpha = 2, \beta_j$.

Στην ενότητα 2.5 περιγράψαμε αναλυτικά την κατανομή Γάμμα και είδαμε ότι αν η συνεχής τυχαία μεταβλητή X ακολουθεί την κατανομή Γάμμα με παραμέτρους: α, β, η συνάρτηση πυκνότητας πιθανότητάς της είναι ίση με:

$$f_X(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, x \geq 0.$$

Εάν έχουμε μία συνάρτηση:

$$f_X(x) = \alpha_1 {\beta_1}^2 x e^{-\beta_1 x} + a_2 {\beta_2}^2 x e^{-\beta_2 x} + \cdots + a_n {\beta_n}^2 x e^{-\beta_n x} \ (x, \beta_j > 0),$$

για την οποία ισχύει:

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n \Rightarrow \sum_{j=1}^{n} \alpha_j = 1,$$

τότε θα είναι μία μείξη Γάμμα κατανομών με παραμέτρους $2, \beta_j$.

Για παράδειγμα η:

$$f_X(x) = \frac{9xe^{-3x}}{4} + 12xe^{-4x}, x > 0,$$

είναι μία μείξη Γάμμα κατανομών διότι μπορεί να γραφτεί και ως:

$$f_X(x) = \frac{1}{4} 3^2 x e^{-3x} + \frac{3}{4} 4^2 x e^{-4x}, x > 0.$$

Με παραμέτρους:

$$\beta_1(= 3), \beta_2(= 4) > 0,$$

και αντίστοιχα βάρη:

$$\alpha_1 \left(=\frac{1}{4}\right) + \alpha_2 \left(=\frac{3}{4}\right) = 1.$$

Στις επόμενες τρεις Ενότητες θα προσεγγίσουμε τις ίδιες συναρτήσεις πυκνότητας πιθανότητας των κατανομών Pareto, Weibull και Lognormal από μία μείξη Γάμμα $(2, \beta_j)$ κατανομών για να αξιολογήσουμε για κάθε μία περίπτωση ποια από τις δύο προσεγγίσεις είναι περισσότερο ικανοποιητική.

3.7: ΠΡΟΣΕΓΓΙΣΗ ΚΑΤΑΝΟΜΗΣ PARETO ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΓΑΜΜΑ $(2, \beta_j)$ ΚΑΤΑΝΟΜΩΝ.

Στο παράδειγμα 3.3.1 προσεγγίσαμε την συνάρτηση πυκνότητας πιθανότητας μίας τυχαίας συνεχούς μεταβλητής Y, η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 6, \beta = 2,$ από μία μείξη εκθετικών κατανομών. Σε αυτή την ενότητα θα προσεγγίσουμε την ίδια συνάρτηση πυκνότητας πιθανότητας από μία μείξη Γάμμα κατανομών $f_X(x)\ (2, \beta_j)$, με παραμέτρους β_1, β_2 και έπειτα θα δούμε ποια προσέγγιση ειναι πιο ικανοποιητική.

Στο παράδειγμα 3.3.1, είδαμε ότι η συνάρτηση πυκνότητας πιθανότητας της Y, με τις συγκεκριμένες παραμέτρους γράφεται στην μορφή:

$$f_Y(y) = 3\left(1 + \frac{y}{2}\right)^{-7}, y \geq 0.$$

Η μείξη Γάμμα κατανομών θα γράφεται στην μορφή:

$$f_X(x) = A{\beta_1}^2 x e^{-\beta_1 x} + (1 - A){\beta_2}^2 x e^{-\beta_2 x}, x \geq 0.$$

Στην Ενότητα 2.5 είδαμε ότι αν η συνεχής τυχαία μεταβλητή X ακολουθεί την κατανομή Γάμμα με παραμέτρους $\alpha = 2, \beta$, τότε οι ροπές $\kappa -$τάξης προκύπτουν από την σχέση:

$$E(X^\kappa) = \frac{\Gamma(\kappa+2)}{\Gamma(2)\beta^\kappa} = \frac{(\kappa+1)!}{\beta^\kappa}.$$

Σύμφωνα με την μέθοδο των ροπών θα ισχύει:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{2A}{\beta_1} + \frac{2(1-A)}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow \frac{2}{5} = \frac{6A}{{\beta_1}^2} + \frac{6(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow \frac{4}{5} = \frac{24A}{{\beta_1}^3} + \frac{24(1-A)}{{\beta_2}^3}.$$

Με την χρήση του Mathematica, βρίσκουμε:

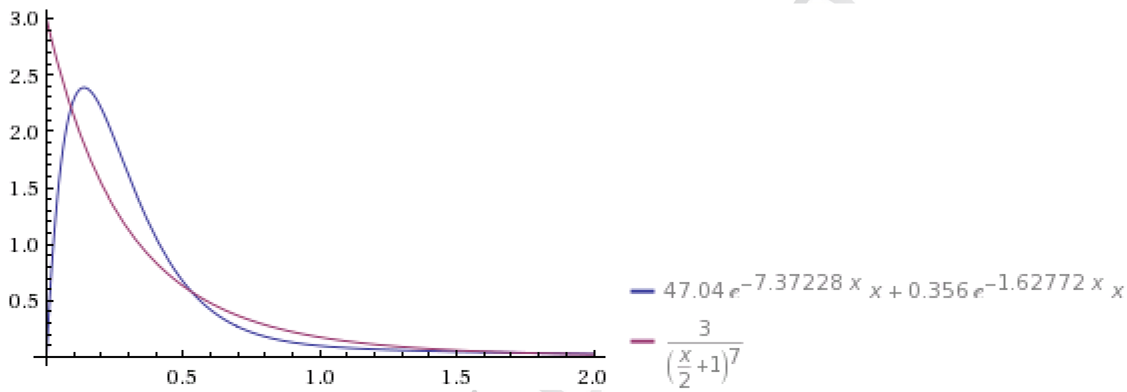
$$A \approx 0.865563, \quad \beta_1 \approx 7.37228, \quad \beta_2 \approx 1.62772,$$

επομένως η μείξη Γάμμα κατανομών θα είναι της μορφής:

$$f_X(x) \approx 0.865563 \cdot 7.37228^2 x e^{-7.37228x} + (1 - 0.865563) \cdot 1.62772^2 x e^{-1.62772x}, x > 0$$

$$\Rightarrow f_X(x) \approx 47.04 x e^{-7.37228x} + 0.356 x e^{-1.62772x}, x > 0.$$

Στο παρακάτω σχήμα παρατηρούμε τις γραφικές παραστάσεις των δύο συναρτήσεων πυκνότητας πιθανότητας:



Σχήμα 3.19: Γραφική παράσταση των δύο συναρτήσεων πυκνότητας πιθανότητας (με μπλε χρώμα της μείξης Γάμμα και με ροζ χρώμα της Pareto).

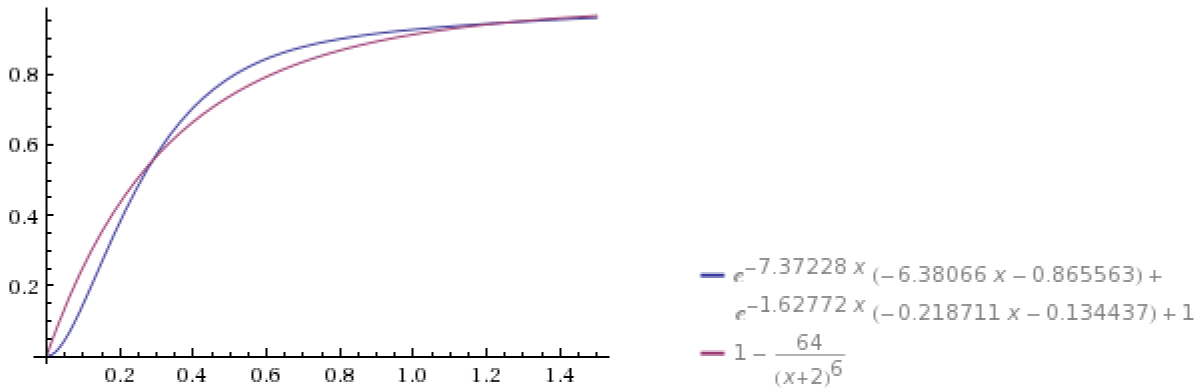
Η συνάρτηση κατανομής της Y, θα είναι:

$$F_Y(y) = 1 - \frac{64}{(y+2)^6}, y \geq 0,$$

ενώ της X, η οποία υπολογίστηκε με χρήση του Mathematica, θα γράφεται στην μορφή:

$$F_X(x) = e^{-7.37228x}(-6.38066x - 0.865563) + e^{-1.62772x}(-0.218711x - 0.134437) + 1.$$

Οι γραφικές τους παραστάσεις απεικονίζονται στο παρακάτω γράφημα:



Σχήμα 3.20: Γραφική παράσταση των $F_X(x)$ (μπλε χρώμα), $F_Y(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ ΠΡΟΣΕΓΓΙΣΗΣ

Η απόσταση των δύο συναρτήσεων πυκνότητας πιθανότητας, θα είναι:

$$\int_0^\infty |f_X(x) - f_Y(x)|dx = \int_0^\infty \left|47.04xe^{-7.37228x} + 0.356xe^{-1.62772x} - 3\left(1+\frac{x}{2}\right)^{-7}\right|dx \approx$$

$$\approx 0.334.$$

Υπολογίζοντας την μέγιστη διαφορά των δύο συναρτήσεων βρίσκουμε ότι:

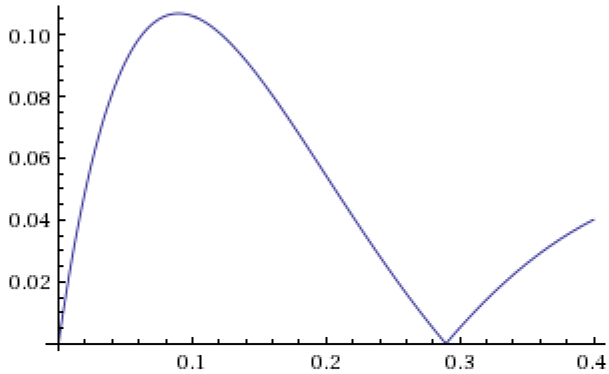
$$sup_{x\geq 0}|f_X(x) - f_Y(x)| \approx 3,$$

στο σημείο $x_0 = 0$.

Αντίστοιχα η μέγιστη διαφορά των δύο συναρτήσεων κατανομής προκύπτει από την σχέση:

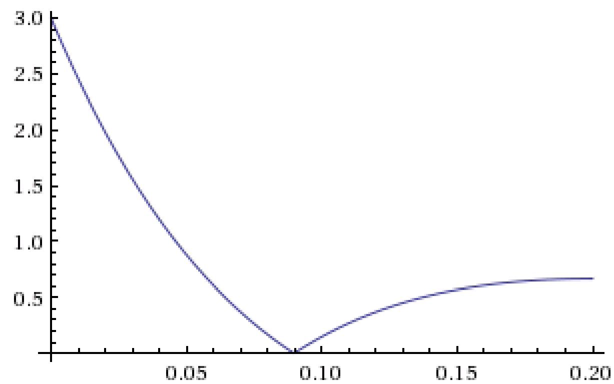
$$sup_{x\geq 0}|F_X(x) - F_Y(x)| \approx 0.106794,$$

Στο σημείο $x_0 \approx 0.0897424$. Αυτή η «απόσταση» απεικονίζεται στο παρακάτω σχήμα:



Σχήμα 3.21: Διάγραμμα του $|F_X(x) - F_Y(x)|$.

Στο παρακάτω γράφημα βλέπουμε την γραφική παράσταση του $|f_X(x) - f_Y(x)|$.



Σχήμα 3.22: Διάγραμμα του $|f_X(x) - f_Y(x)|$.

Για τις ροπές τέταρτης τάξης θα ισχύει:

$$E(Y^4) = 3.2, \qquad E(X^4) = 2.33212.$$

Στο παράδειγμα 3.3.1 είδαμε ότι το ολοκλήρωμα της απόλυτης διαφοράς της συνάρτησης πυκνότητας πιθανότητας με την μείξη εκθετικών κατανομών ήταν περίπου ίσο με 0.039, ενώ το ολικό μέγιστο της απόλυτης διαφοράς τους ήταν ίσο με 0.2 και των κατανομών τους ήταν ίσο με 0.0119749 . Στο συγκεκριμένο παράδειγμα το ολοκλήρωμα της απόλυτης διαφοράς της συνάρτησης πυκνότητας πιθανότητας της κατανομής Pareto με την μείξη Γάμμα κατανομών είναι περίπου ίσο με 0.334, ενώ το ολικό μέγιστο της απόλυτης διαφοράς τους είναι περίπου ίσο με 3 και των συναρτήσεων κατανομής τους είναι περίπου ίσο με 0.106794. Επίσης η ροπή τέταρτης τάξης της Y είναι ίση με 3.2, της μείξης Γάμμα είναι ίση με 2.33212. Η ροπή τέταρτης τάξης της μείξης εκθετικών κατανομών ήταν ίση με 2.66667. Όπως και στο παράδειγμα 3.3.1, έτσι και εδώ θα υπολογίσουμε τα άνω ποσοστημόρια των δύο συναρτήσεων κατανομής για να διαπιστώσουμε πόσο «κοντά» βρίσκονται μεταξύ τους. Έχουμε τότε τον παρακάτω πίνακα:

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό σφάλματος
0.001	0.999	0.999	4.286	4.324	0.88%
0.005	0.995	0.995	3.133	2.836	−10.47%
0.01	0.99	0.99	2.616	2.308	−13.34%
0.02	0.98	0.98	2.078	1.839	−12.99%
0.03	0.97	0.97	1.750	1.588	−10.20%

Πίνακας 3.4: Πίνακας άνω ποσοστημορίων των $F_X(x), F_Y(y)$.

Παρατηρούμε επίσης ότι οι τιμές των x από εκείνες του y, έχουν μεγαλύτερη «απόκλιση», σε σχέση με την προσέγγιση από μία μείξη εκθετικών κατανομών.

Διαπιστώνουμε τότε ότι για την κατανομή Pareto με τις συγκεκριμένες παραμέτρους η προσέγγιση από μία μείξη εκθετικών κατανομών είναι αρκετά πιο ικανοποιητική από αυτής της προσέγγισης από μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο να είναι ίση με 2.

3.8: ΠΡΟΣΕΓΓΙΣΗ ΚΑΤΑΝΟΜΗΣ WEIBULL ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΓΑΜΜΑ $(2, \beta_j)$ ΚΑΤΑΝΟΜΩΝ.

Κατά την προσέγγιση της συνάρτησης πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής Y με παραμέτρους $\lambda = 2.63045, \alpha = 0.9$ από μία μείξη δύο εκθετικών κατανομών στο παράδειγμα 3.4.1 είδαμε ότι η $f_Y(y)$ γράφεται στην μορφή:

$$f_Y(y) = 2.15y^{-0.1}e^{-(2.63045y)^{0.9}}, y > 0.$$

Οι ροπές πρώτης, δεύτερης και τρίτης τάξης ήταν ίσες με:

$$E(Y) = \frac{2}{5}, \quad E(Y^2) = 0.358, \quad E(Y^3) = 0.509.$$

Στην παρούσα Ενότητα θα προσεγγίσουμε την συγκεκριμένη συνάρτηση πυκνότητας πιθανότητας από μία μείξη δύο Γάμμα $(2, \beta_j)$ κατανομών $f_X(x)$, για να αξιολογήσουμε ποια προσέγγιση είναι πιο ικανοποιητική. Η $f_X(x)$, θα είναι της μορφής:

$$f_X(x) = A{\beta_1}^2 xe^{-\beta_1 x} + (1 - A){\beta_2}^2 xe^{-\beta_2 x}, x > 0.$$

Σύμφωνα με την μέθοδο των ροπών θα έχουμε:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{2A}{\beta_1} + \frac{2(1-A)}{\beta_2},$$

$$E(Y^2) = E(X^2) \iff 0.358 = \frac{6A}{{\beta_1}^2} + \frac{6(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \iff 0.509 = \frac{24A}{{\beta_1}^3} + \frac{24(1-A)}{{\beta_2}^3}.$$

Τα αποτελέσματα που προκύπτουν για τους αριθμούς A, β_1, β_2, είναι τα εξής:

$$A \approx 0.376309, \quad \beta_1 \approx 2.62783, \quad \beta_2 \approx 10.9807.$$

Τότε η $f_X(x)$, θα είναι:

$$f_X(x) \approx 0.376309(2.62783)^2 x e^{-2.62783x} + (1 - 0.376309)(10.9807)^2 x e^{-10.9807x}, x > 0 \Rightarrow$$

$$\Rightarrow f_X(x) \approx 2.5986 x e^{-2.62783x} + 75.202 x e^{-10.9807x}, x > 0.$$

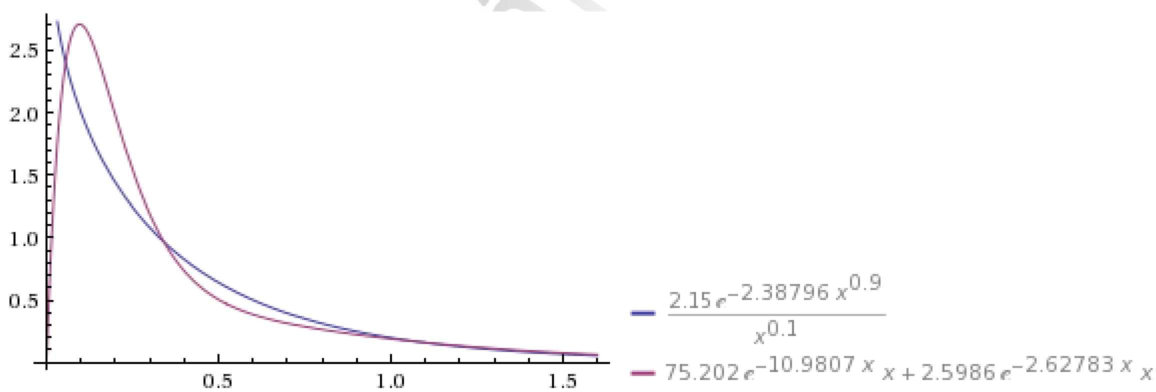
Η συνάρτηση κατανομής της, με χρήση του Mathematica βρίσκουμε ότι είναι:

$$F_X(x) \approx e^{-10.9807x}(-6.84856x - 0.623691) + e^{-2.62783x}(-0.988877x - 0.376309) + 1, x > 0.$$

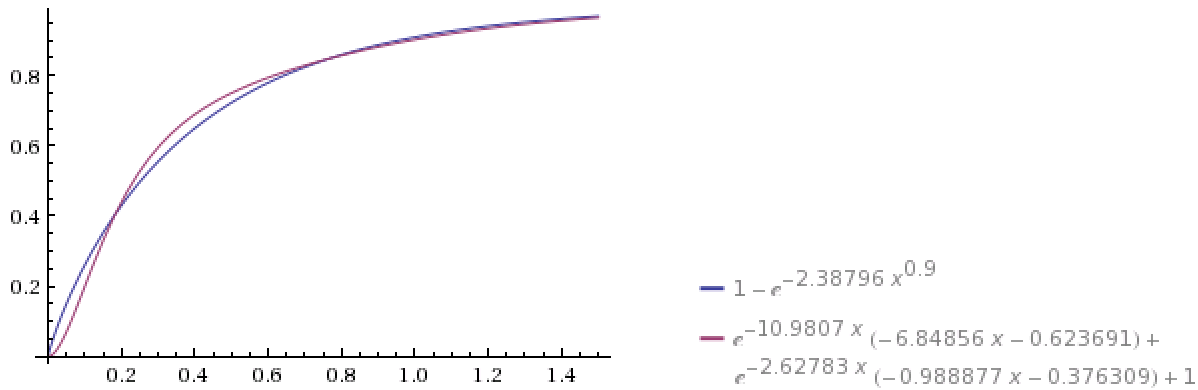
Η συνάρτηση κατανομής της Υ, θα γράφεται ως:

$$F_Y(y) = 1 - e^{-(2.63045y)^{0.9}}, y > 0.$$

Τα παρακάτω δύο γραφήματα απεικονίζουν τις γραφικές παραστάσεις των δύο συναρτήσεων πυκνότητας πιθανότητας και των δύο συναρτήσεων κατανομής αντίστοιχα.



Σχήμα 3.23: Διάγραμμα των $f_Y(x)$ (μπλε χρώμα) και $f_X(x)$ (ροζ χρώμα).



Σχήμα 3.24: Διάγραμμα των $F_Y(x)$ (μπλε χρώμα) και $F_X(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ ΠΡΟΣΕΓΓΙΣΗΣ

Το ολοκλήρωμα της απόλυτης διαφοράς των $f_X(x), f_Y(x)$ με χρήση του Mathematica βρίσκουμε ότι είναι ίσο με:

$$\int_0^\infty |f_X(x) - f_Y(x)| dx \approx 0.26.$$

Το αντίστοιχο ολοκλήρωμα στην προσέγγιση από μία μείξη εκθετικών κατανομών ήταν περίπου ίσο με 0.038.

Στο παράδειγμα 3.4.1 υπολογίσαμε την ροπή τέταρτης για την μείξη εκθετικών $f_X(x)$ και ίσχυε ότι:

$$E(X^4) \approx 0.990126,$$

ενώ η αντίστοιχη ροπή της Weibull $f_Y(y)$ ήταν ίση με:

$$E(Y^4) \approx 0.999998.$$

Θα υπολογίσουμε τώρα την ίδια ροπή για την μείξη Γάμμα $(2, \beta_j)$ κατανομών $f_X(x)$. Ισχύει τότε ότι:

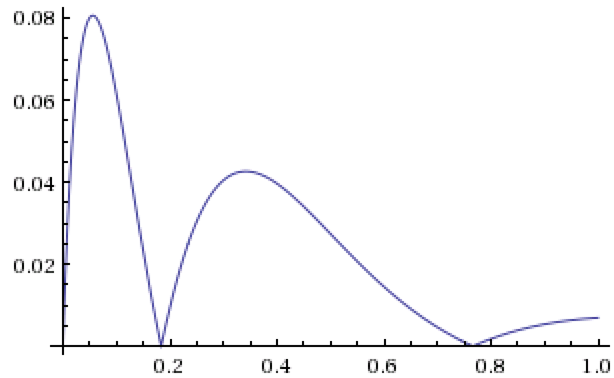
$$E(X^4) \approx 0.95212,$$

Παρατηρούμε ότι η ροπή τέταρτης τάξης της μείξης εκθετικών κατανομών είναι πιο «κοντά» στην ροπή της Weibull απ'ότι είναι η συγκεκριμένη ροπή της μείξης Γάμμα.

Η μέγιστη διαφορά των δύο συναρτήσεων κατανομής είναι:

$$sup_{x>0}|F_X(x) - F_Y(x)| \approx 0.0805891,$$

στο σημείο $x_0 \approx 0.0558638$, η οποία απεικονίζεται στο παρακάτω γράφημα.



Σχήμα 3.25: Διάγραμμα του $|F_X(x) - F_Y(x)|$.

Στο παράδειγμα 3.4.1 αυτή η διαφορά ήταν περίπου ίση με 0.0125938.

Θα υπολογίσουμε επίσης τα άνω ποσοστημόρια των δύο συναρτήσεων κατανομής, όπως και στο παράδειγμα 3.4.1 και ο πίνακας που προκύπτει είναι ο παρακάτω:

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό Σφάλματος
0.001	0.999	0.999	3.099	3.255	4.79%
0.005	0.995	0.995	2.401	2.424	0.94%
0.01	0.99	0.99	2.092	2.074	−0.86%
0.02	0.98	0.98	1.777	1.731	−2.65%
0.03	0.97	0.97	1.587	1.532	−3.59%

Πίνακας 3.5: Πίνακας άνω ποσοστημορίων των δύο συναρτήσεων κατανομής.

Σε αυτόν τον πίνακα επίσης παρατηρούμε σε σχέση με εκείνον της μείξης εκθετικών ότι οι τιμές των x από των y διαφέρουν περισσότερο και το ποσοστό σφάλματος έχει μεγαλύτερες τιμές. Διαπιστώνουμε τότε με όλα τα παραπάνω ότι η προσέγγιση της συγκεκριμένης συνάρτησης πυκνότητας πιθανότητας από μία μείξη Γάμμα $(2, \beta_j)$ κατανομών δεν είναι τόσο ικανοποιητική, όσο εκείνη από μία μείξη εκθετικών κατανομών.

3.9: ΠΡΟΣΕΓΓΙΣΗ ΛΟΓΑΡΙΘΜΟΚΑΝΟΝΙΚΗΣ ΚΑΤΑΝΟΜΗΣ ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΓΑΜΜΑ $(2, \beta_j)$ ΚΑΤΑΝΟΜΩΝ.

Στο παράδειγμα 3.5.1 είδαμε την προσέγγιση της συνάρτησης πυκνότητας πιθανότητας μίας τυχαίας συνεχούς μεταβλητής Υ, η οποία ακολουθεί την λογαριθμοκανονική κατανομή με παραμέτρους $\mu = -1.41629,\ \sigma = 1$, από μία μείξη δύο εκθετικών κατανομών. Στην παρούσα Ενότητα θα εξετάσουμε την προσέγγιση της ίδιας συνάρτησης πυκνότητας πιθανότητας από μία μείξη δύο Γάμμα $(2, \beta_j)$ κατανομών $f_X(x)$, με παραμέτρους $(2, \beta_1), (2, \beta_2)$ και αντίστοιχα βάρη $A, 1 - A$, για να διαπιστώσουμε ποια από τις δύο μείξεις είναι πιο «κοντά» στην συνάρτηση πυκνότητας πιθανότητας της συνεχούς τυχαίας μεταβλητής Υ.

Η $f_Y(y)$ είδαμε ότι είναι:

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{(logy + 1.41629)^2}{2}\right), y > 0,$$

τις ροπές πρώτης, δεύτερης και τρίτης τάξης τις έχουμε υπολογίσει στο παράδειγμα 3.5.1 και είναι ίσες με:

$$E(Y) = \frac{2}{5}, \quad E(Y^2) = 0.434926, \quad E(Y^3) = 1.28548.$$

Με την μέθοδο των ροπών θα υπολογίσουμε τις τιμές των σταθερών A, β_1, β_2, από το σύστημα:

$$E(Y) = E(X) \Leftrightarrow \frac{2}{5} = \frac{2A}{\beta_1} + \frac{2(1-A)}{\beta_2},$$

$$E(Y^2) = E(X^2) \Leftrightarrow 0.434926 = \frac{6A}{{\beta_1}^2} + \frac{6(1-A)}{{\beta_2}^2},$$

$$E(Y^3) = E(X^3) \Leftrightarrow 1.28548 = \frac{24A}{{\beta_1}^3} + \frac{24(1-A)}{{\beta_2}^3}.$$

Βρίσκουμε τότε:

$$A \approx 0.0439103, \quad \beta_1 \approx 0.960562, \quad \beta_2 \approx 6.19683.$$

Έτσι η μείξη Γάμμα $(2, \beta_j)$ κατανομών θα γράφεται στην μορφή:

$$f_X(x) = A{\beta_1}^2 x e^{-\beta_1 x} + (1 - A){\beta_2}^2 x e^{-\beta_2 x} =$$

$$\approx 0.0439103 \cdot 0.960562^2 x e^{-0.960562x} +$$

$$+(1 - 0.0439103)6.19683^2 x e^{-6.19683x}, x > 0$$

$$\Rightarrow \quad f_X(x) \approx 0.0405 \cdot xe^{-0.960562x} + 36.7145 \cdot xe^{-6.19683x}, x > 0,$$

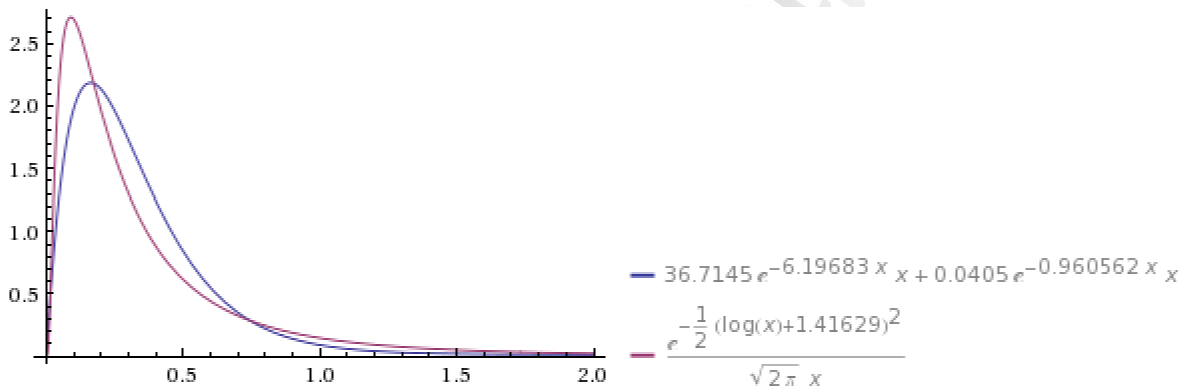
και με χρήση του Mathematica, η συνάρτηση κατανομής της γράφεται ως:

$$F_X(x) \approx 1 + e^{-6.19683x}(-0.9560897 - 5.92472x) +$$
$$+e^{-0.960562x}(-0.0439103 - 0.0421628x), x > 0$$

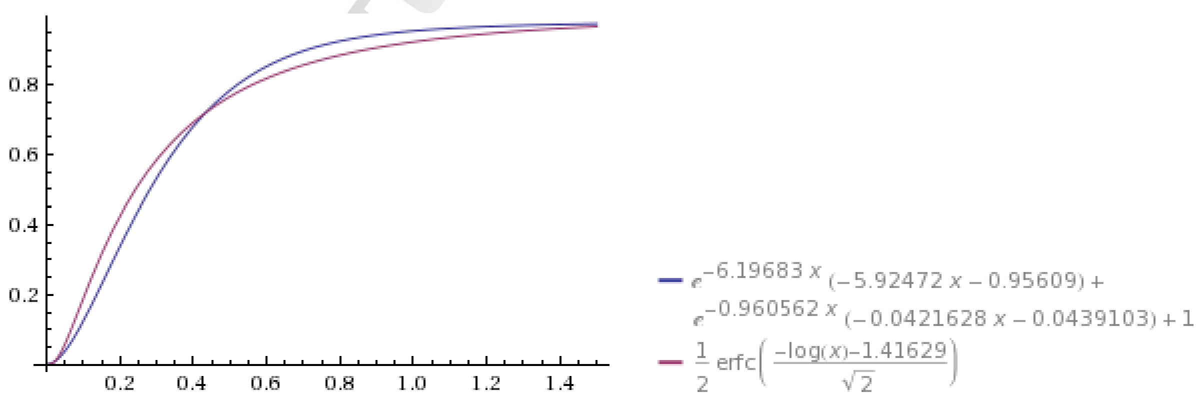
Η συνάρτηση κατανομής της λογαριθμοκανονικής θα είναι:

$$F_Y(y) = (\Phi(lny + 1.41629), y > 0.$$

Στα παρακάτω δύο γραφήματα παρατηρούμε τις γραφικές παραστάσεις των $f_X(x), f_Y(x)$ και $F_Y(x), F_Y(y)$ αντίστοιχα .



Σχήμα 3.26: Διάγραμμα των $f_X(x)$ (μπλε χρώμα),$f_Y(x)$ (ροζ χρώμα).



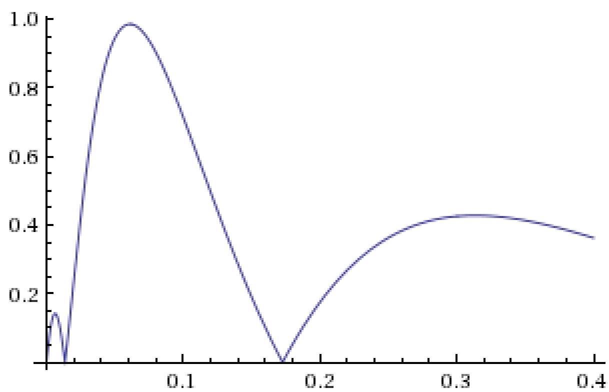
Σχήμα 3.27: Διάγραμμα των $F_X(x)$ (μπλε χρώμα),$F_Y(x)$ (ροζ χρώμα).

ΑΞΙΟΛΟΓΗΣΗ ΠΡΟΣΕΓΓΙΣΗΣ

Θα υπολογίσουμε την μέγιστη «διαφορά» των δύο πυκνοτήτων και των δύο κατανομής αντίστοιχα. Συγκεκριμένα στο Mathematica βρίσκουμε ότι:

$$sup_{x>0}|f_X(x) - f_Y(x)| \approx 0.984734,$$

στο σημείο $x_0 \approx 0.0610846$. Στο παρακάτω γράφημα βλεπουμε την γραφική παράσταση του $|f_X(x) - f_Y(x)|$ και παρατηρούμε και το ολικό μέγιστο στο συγκεκριμένο σημείο.



Σχήμα 3.28: Γραφική παράσταση του $|f_X(x) - f_Y(x)|$.

Στο παράδειγμα 3.5.1 ίσχυε ότι:

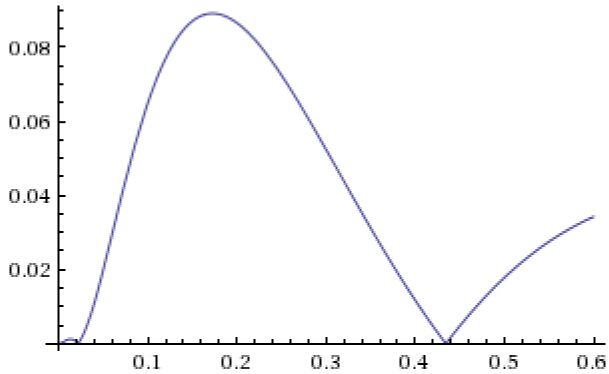
$$sup_{x>0}|f_Y(x) - f_X(x)| \approx 2.7,$$

όπου $f_X(x)$ η μείξη εκθετικών κατανομών.

Αντίστοιχα για τις συναρτήσεις κατανομής ισχύει:

$$sup_{x>0}|F_X(x) - F_Y(x)| \approx 0.089113,$$

στο σημείο $x_0 \approx 0.172692$ και η γραφική παράσταση του $|F_X(x) - F_Y(x)|$ απεικονίζεται στο παρακάτω σχήμα:



Σχήμα 3.29: Γραφική παράσταση του $|F_X(x) - F_Y(x)|$.

Στο παράδειγμα 3.5.1 η συγκεκριμένη «διαφορά» ήταν περίπου ίση με 0.069.

Για την ροπή τέταρτης τάξης της $f_Y(y)$ ισχύει:

$E(Y^4) \approx 10.3278.$

Η ροπή τέταρτης τάξης της μείξης γάμμα κατανομών είναι περίπου ίση με 6.26486, ενώ στο παράδειγμα 3.5.1, η ροπή τέταρτης τάξης της μείξης εκθετικών ήταν περίπου ίση με 7.91617.

Σύμφωνα με την μέθοδο (3.2.2) η «απόσταση» των δύο συναρτήσεων πυκνότητας πιθανότητας θα είναι ίση με:

$$\int_0^{\infty} |f_X(x) - f_Y(x)| dx \approx 0.27,$$

Ενώ εκείνο της προσέγγισης με την μείξη εκθετικών κατανομών είχε περίπου την τιμή 0.223. Τέλος όπως και στις προηγούμενες προσεγγίσεις θα υπολογίσουμε τα άνω ποσοστημόρια των δύο κατανομών και έτσι προκύπτει ο παρακάτω πίνακας:

α	$F_X(x)$	$F_Y(y)$	x	y	Ποσοστό Σφάλματος
0.001	0.999	0.999	5.914	5.333	−10.89%
0.005	0.995	0.995	3.878	3.188	1.85%
0.01	0.99	0.99	2.935	2.484	−21.64%
0.02	0.98	0.98	1.906	1.891	−0.79%
0.03	0.97	0.97	1.337	1.591	15.9%

Πίνακας 3.6: Πίνακας άνω ποσοστημορίων των $F_X(x), F_Y(y)$.

Συνεπώς από όλες τις μεθόδους αξιολόγησης που χρησιμοποιήσαμε, συμπεραίνουμε ότι η προσέγγιση της λογαριθμοκανονικής κατανομής με τις συγκεκριμένες παραμέτρους από μία μείξη δύο εκθετικών κατανομών είναι περισσότερο ικανοποιητική από την προσέγγιση από μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο να ισούται με 2.

ΚΕΦΑΛΑΙΟ 4: ΣΥΝΘΕΤΕΣ ΚΑΤΑΝΟΜΕΣ ΚΑΙ ΠΙΘΑΝΟΤΗΤΑ ΧΡΕΟΚΟΠΙΑΣ.

Σε αυτό το κεφάλαιο θα δούμε αναλυτικά τις σύνθετες κατανομές και θα αναφερθούμε στην εύρεση των βασικών τους ποσοτήτων όπως είναι η μέση τιμή, η διασπορά και η ροπογεννήτρια συνάρτησή τους. Στην συνέχεια θα περιγράψουμε τις βασικές έννοιες της θεωρίας χρεοκοπίας, όπως είναι η πιθανότητα χρεοκοπίας, το πλεόνασμα, το αρχικό αποθεματικό, το περιθώριο ασφαλείας κ.ά.

Το πλεόνασμα μίας ασφαλιστικής εταιρείας, προκύπτει από την διαφορά των εσόδων, τα οποία προέρχονται κυρίως από την είσπραξη των ασφαλίστρων, με των εξόδων της, τα οποία δαπανούνται λόγω των απαιτήσεών της.

Επίσης το αρχικό αποθεματικό, δηλώνει το αρχικό κεφάλαιο, το οποίο διαθέτει η εταιρεία κατά την έναρξη των εργασιών της και αποτελεί το πλεόνασμά της την χρονική στιγμή $t = 0$.

Κεντρικό αντικείμενο στην θεωρία χρεοκοπίας, αποτελεί η εύρεση της πιθανότητας, η στοχαστική ανέλιξη του πλεονάσματος να πάρει για πρώτη φορά αρνητική τιμή. Αυτή συνεπώς η πιθανότητα καλείται πιθανότητα χρεοκοπίας. Αξίζει ωστόσο να σημειωθεί πως στην πραγματικότητα, όταν το πλεόνασμα θα πάρει αρνητική τιμή, δεν επέρχεται απαραίτητα χρεοκοπία για την εταιρεία, διότι δεν κατέχει ένα μόνο χαρτοφυλάκιο.

Στην συνέχεια θα αναλύσουμε την έννοια της μέγιστης σωρευτικής απώλειας, η οποία όπως και θα εξηγήσουμε κατά την διάρκεια του κεφαλαίου, ακολουθεί μία σύνθετη κατανομή και συνδέεται με την πιθανότητα μη χρεοκοπίας.

4.1: ΣΥΝΘΕΤΕΣ ΚΑΤΑΝΟΜΕΣ.

Σε πολλά προβλήματα των εφαρμοσμένων πιθανοτήτων, όταν χρησιμοποιούμε αθροίσματα ανεξάρτητων τυχαίων μεταβλητών, το πλήθος των προσθετέων σε ένα τέτοιο άθροισμα δεν είναι γνωστό, αλλά αποτελεί και αυτό μία διακριτή τυχαία μεταβλητή.

Έστω τότε $\{X_i\}$ μία ακολουθία από ανεξάρτητες και ισόνομες μεταβλητές και μία ακέραια μεταβλητή N, ανεξάρτητη από τις X_i και

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

Τοτε λέμε ότι η S ακολουθεί μία σύνθετη κατανομή.

Από το βιβλίο του Dickson (2005) μπορούμε να δούμε ότι για την μέση τιμή της S, ισχύει ότι:

$$E(S) = E(N)E(X_i),$$

για την διακύμανση της S, ισχύει η σχέση:

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

Σε αυτό το σημείο αξίζει να αναφερθεί ότι, στην ειδική περίπτωση που η τυχαία μεταβλητή N ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0,$ ισχύει ότι:

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

οπότε η μέση τιμή και η διακύμανση της S, θα γράφονται στην μορφή:

$$E(S) = \lambda E(X_i),$$

$$Var(S) = \lambda\big(Var(X_i) + E^2(X_i)\big).$$

Μία άλλη βασική ιδιότητα για την ροπογεννήτρια συνάρτηση $M_S(t)$ της σύνθετης κατανομής S, είναι ότι γράφεται στην μορφή:

$$M_s(t) = M_N\big(lnM_X(t)\big),$$

όπου $M_N(t), M_X(t)$ οι ροπογεννήτριες συναρτήσεις των κατανομών N και X αντίστοιχα.

Ας δούμε ένα παράδειγμα για τον υπολογισμό της διακύμανσης μίας σύνθετης κατανομής.

Παράδειγμα 4.1.1

Υποθέτουμε ότι οι ισόνομες και ανεξάρτητες κατανομές X_i, ακολουθούν την κατανομή Γάμμα, με παραμέτρους $(\alpha, \beta), \alpha, \beta > 0$. Η διακριτή τυχαία μεταβλητή N ακολουθεί την κατανομή Poisson με παράμετρο λ. Τότε θα ισχύει:

$$V\,ar(N) = \lambda,\ E(N) = \lambda,\ Var(X_i) = \frac{\alpha}{\beta^2},\ E(X_i) = \frac{\alpha}{\beta},$$

επομένως από τον παραπάνω τύπο για την διακύμανση της S, έχουμε ότι:

$$Var(S) = \lambda \cdot \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2} \cdot \lambda = \frac{\alpha\lambda(1+\alpha)}{\beta^2}.$$

Ας δούμε τώρα άλλο ένα παράδειγμα για τον υπολογισμό της ροπογεννήτριας συνάρτησης μίας σύνθετης κατανομής S.

Παράδειγμα 4.1.2

Υποθέτουμε ότι η N ακολουθεί την διωνυμική κατανομή με παραμέτρους n, p και η X ακολουθεί την εκθετική κατανομή με παράμετρο λ. Τότε ισχύει:

$$M_N(t) = (pe^t + q)^n,\ \forall t \in R,\ M_X(t) = \frac{\lambda}{\lambda - t}, t < \lambda,$$

$$lnM_X(t) = ln\frac{\lambda}{\lambda - t} = ln\lambda - l\,n(\lambda - t),$$

$$M_S(t) = M_N\big(lnM_X(t)\big) = (pe^{ln\lambda - ln\,(\lambda - t)} + q)^n = \left(p\frac{\lambda}{\lambda - t} + q\right)^n, t < \lambda.$$

Αν η N ακολουθεί την Γεωμετρική κατανομή με παράμετρο $0 < p < 1$ και η X ακολουθεί την εκθετική κατανομή με παράμετρο λ, τότε θα ισχύει:

$$M_N(t) = \frac{pe^t}{1 - qe^t}, q = 1 - p,$$

$$M_S(t) = M_N\big(lnM_X(t)\big) = \frac{pe^{ln\lambda - \ln(\lambda - t)}}{1 - q^{ln\lambda - \ln(\lambda - t)}} = \frac{p\lambda}{\lambda - t - q\lambda}\ , t < \lambda.$$

4.2: ΜΙΑ ΕΙΣΑΓΩΓΗ ΣΤΗ ΘΕΩΡΙΑ ΧΡΕΟΚΟΠΙΑΣ.

Στην εισαγωγή του κεφαλαίου διατυπώσαμε τον ορισμό του πλεονάσματος και του αρχικού αποθεματικού. Εάν τότε η στοχαστική ανέλιξη του πλεονάσματος ενός χαρτοφυλακίου συμβολίζεται με: $\{U(t), t > 0\}$ και με u, το αρχικό αποθεματικό θα ισχύει η παρακάτω ισότητα:

$$U(t) = u + P(t) - S(t) = u + ct - \sum_{i=1}^{N(t)} X_i,$$

όπου:

$P(t)$: Το συνολικό ασφάλιστρο στο διάστημα $[0, \mathrm{t}]$. Στην απλούστερη περίπτωση που θα εξετάσουμε εδώ, ισούται με το γινόμενο ct, όπου:

c: Τα έσοδα των ασφαλίστρων ανά μονάδα χρόνου.

Επίσης είδαμε ότι το αρχικό αποθεματικό, αποτελεί και το αρχικό κεφάλαιο της εταιρείας την χρονική στιγμή $t = 0$. Θα ισχύει τότε ότι $u = U(0)$.

$S(t)$: Μία σύνθετη ανέλιξη για τις αποζημιώσεις στο διάστημα $[0, \mathrm{t}]$

$$S(t) = \sum_{i=1}^{N(t)} X_i = \begin{cases} X_1 + X_2 + \cdots + X_{N(t)}, & N(t) = 1,2 \ldots \\ 0, & N(t) = 0. \end{cases}$$

Όπου:

X_i Μία τυχαία μεταβλητή που δηλώνει το ύψος της αποζημίωσης από τον κίνδυνο i.

$N(t)$: Μία στοχαστική ανέλιξη Poisson που δηλώνει τον αριθμό των αποζημιώσεων στο χρονικό διάστημα $[0, \mathrm{t}]$.

Είναι προφανές ότι για να έχει «φερεγγυότητα» ένα χαρτοφυλάκιο πρέπει τα έσοδα που προσφέρει να είναι περισσότερα από τα έξοδα που απαιτεί, πρέπει δηλαδή να ισχύει:

$$c > \lambda\mu_1 \tag{4.2.1}$$

όπου:

λ: Είναι η ένταση της ανέλιξης Poisson που δηλώνει το αναμενόμενο πλήθος των απαιτήσεων στην μονάδα του χρόνου.

μ_1: Η μέση αποζημίωση που προκύπτει απο την γενικότερη σχέση:

$$\mu_\kappa = E(X^\kappa) = \int_0^\infty x^\kappa dF(x) = \int_0^\infty x^\kappa f(x)dx.$$

Επομένως για $\kappa = 1$ έχουμε ότι:

$$\mu_1 = E(X) = \int_0^\infty x f(x)dx.$$

Η F(x) δηλώνει την κατανομή των αποζημιώσεων και η $f(x)$ την πυκνότητά της.

Το γινόμενο $\lambda\mu_1$ δηλώνει το αναμενόμενο συνολικό μέγεθος αποζημιώσεων στην μονάδα του χρόνου, δηλαδή τα έξοδα. Είναι λοιπόν προφανές ότι αν δεν ισχύει η σχέση (4.2.1), η χρεοκοπία είναι βέβαιη.

Εισάγουμε τώρα μία νέα σταθερά θ, που ονομάζεται περιθώριο ασφαλείας, η οποία υπολογίζεται από την σχέση:

$$\theta = \frac{\text{έσοδα}}{\text{αναμενόμενα έξοδα}} - 1 = \frac{c}{\lambda\mu_1} - 1.$$

(θ: αναμενόμενο ποσοστό κέρδους).

Εξηγήσαμε επίσης ότι η χρεοκοπία εμφανίζεται όταν για πρώτη φορά το πλεόνασμα θα πάρει αρνητική τιμή. Η χρονική στιγμή εκείνη ονομάζεται χρόνος χρεοκοπίας και συμβολίζεται με:

$$T = inf(t{:}\, U(t) < 0)\,,$$

και η πιθανότητα χρεοκοπίας είναι η πιθανότητα η στοχαστική ανέλιξη του πλεονάσματος να γίνει αρνητική. Επομένως θα ισχύει ότι:

$$\psi(u) = P(U(t) < 0(\text{για κάποιο } t > 0)/U(0) = u),$$

και είναι φθίνουσα συνάρτηση ως προς u. Ισχύει δηλαδή ότι:

$$u_1 < u_2 \Leftrightarrow \psi(u_1) > \psi(u_2) \text{ και } \lim_{u\to\infty} \psi(u) = 0.$$

Αντίστοιχα ορίζουμε και την πιθανότητα μη χρεοκοπίας ως:

$$\delta(u) = 1 - \psi(u),$$

η οποία είναι αύξουσα συνάρτηση ως προς u, οπότε και θα ισχύει:

$$u_1 < u_2 \Leftrightarrow \delta(u_1) < \delta(u_2) \text{ και } \lim_{u\to\infty} \delta(u) = 1.$$

Από τις σημειώσεις του Πολίτη (2015), βλέπουμε ότι η συνάρτηση μη χρεοκοπίας ικανοποιεί την εξίσωση:

$$\delta(u) = \delta(0) + \tfrac{\lambda}{c}\int_0^u \delta(u-x)\overline{F}(x)dx, u \geq 0,$$

όπου $\overline{F}(x)$ η ουρά της κατανομής των αποζημιώσεων, η οποία γράφεται στην μορφή:

$$\overline{F}(x) = 1 - F(x).$$

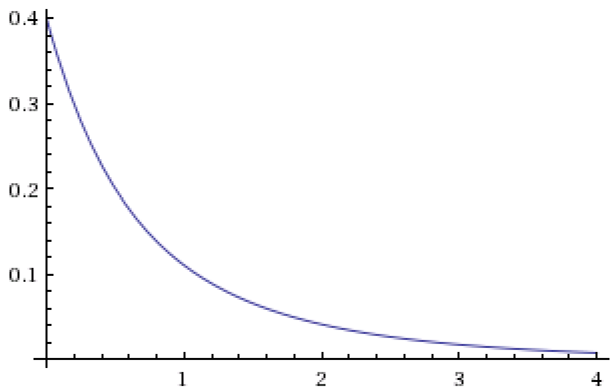
Η πιθανότητα χρεοκοπίας $\psi(u)$ για $u = 0$, είναι ίση με:

$$\psi(0) = \tfrac{1}{1+\theta}\,,$$

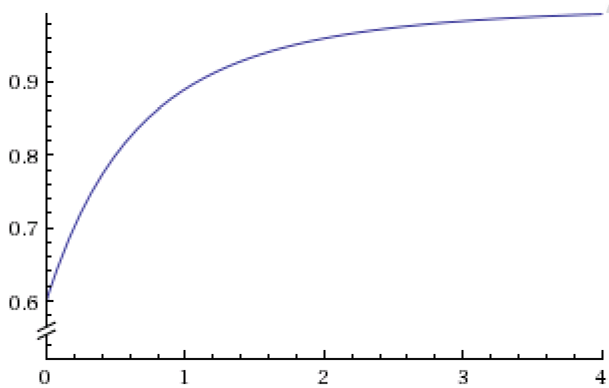
επομένως και η πιθανότητα μη χρεοκοπίας για $u = 0$ θα είναι ίση με:

$$\delta(0) = 1 - \psi(0) = 1 - \frac{1}{1+\theta} = \frac{\theta}{1=\theta} \ .$$

Στα δύο παρακάτω γραφήματα απεικονίζονται οι γραφικές παραστάσεις της πιθανότητας χρεοκοπίας και της μη χρεοκοπίας αντίστοιχα, μίας κατανομής αποζημιώσεων, η οποία αποτελεί μία μείξη εκθετικών κατανομών.



Σχήμα 4.1: Διάγραμμα πιθανότητας χρεοκοπίας.



Σχήμα 4.2: Διάγραμμα πιθανότητας μη χρεοκοπίας.

Παρατηρούμε από το πρώτο σχήμα ότι, η πιθανότητα χρεοκοπίας είναι φθίνουσα συνάρτηση ως προς u και ισχύει ότι:

$$\psi(0) = 0.4 \Rightarrow \frac{1}{1+\theta} = 0.4 \Rightarrow \theta = \frac{3}{2}.$$

Το οποίο μπορούμε να το διαπιστώσουμε και από την γραφική παράσταση της πιθανότητας μη χρεοκοπίας, διότι ισχύει:

$$\delta(0) = 0.6 \ \Rightarrow \frac{\theta}{1+\theta} = 0.6 \Rightarrow \theta = \frac{3}{2}.$$

Επίσης μία ποσότητα με κύριο ενδιαφέρον είναι ο συντελεστής προσαρμογής R, διότι ισχύει ότι:

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

Το R υπολογίζεται ως εξής: Έστω $f(x)$ η συνάρτηση πυκνότητας των αποζημιώσεων και $\mathrm{M(r)}$ η ροπογεννήτριά της, τότε το R είναι η θετική ρίζα της εξίσωσης:

$$-\frac{1}{r}+\frac{1}{r}M_r=\frac{c}{\lambda}\ .$$

Μία άλλη ισοδύναμη μέθοδο για τον υπολογισμό του συντελεστή προσαρμογής, την οποία θα χρησιμοποιούμε κυρίως στο επόμενο κεφάλαιο, αποτελεί η παρακάτω εξίσωση:

$$1+(1+\theta)E(X)r=M_X(r),$$

όπου $\theta, E(X), M_X(r)$, το περιθώριο ασφαλείας, η μέση τιμή και η ροπογεννήτρια συνάρτηση των αποζημιώσεων αντίστοιχα. Ο συντελεστής προσαρμογής είναι η μικρότερη θετική ρίζα της παραπάνω εξίσωσης.

Σε περίπτωση όπου η τυχαία μεταβλητή X ακολουθεί την εκθετική κατανομή με παράμετρο $\alpha > 0$, αποδεικνύεται ότι ο συντεστής προσαρμογής, μπορεί να γραφτεί στην μορφή:

$$R=\frac{\alpha\theta}{1+\theta}\ .$$

Το κυριότερο πρόβλημα στην θεωρία χρεοκοπίας όπως προαναφέραμε στην εισαγωγή του κεφαλαίου είναι ο υπολογισμός της πιθανότητας χρεοκοπίας. Είναι αναγκαίο να αναφερθεί ότι για κατανομές με βαριά ουρά, όπως είναι για παράδειγμα η Pareto, η πιθανότητα χρεοκοπίας δεν μπορεί να υπολογιστεί, παρά μόνο προσεγγιστικά και είναι αυτό με το οποίο θα ασχοληθούμε στο επόμενο κεφάλαιο.

4.3: ΚΛΙΜΑΚΩΤΑ ΥΨΗ ΚΑΙ ΜΕΓΙΣΤΗ ΣΩΡΕΥΤΙΚΗ ΑΠΩΛΕΙΑ.

Μία μεταβλητή, η οποία όπως θα διαπιστώσουμε έχει ιδιαίτερη χρησιμότητα, είναι το μέγεθος της πτώσης του πλεονάσματος κάτω από το αρχικό αποθεματικό u. Η μεταβλητή αυτή στην περίπτωση όπου μελετάμε την πτώση του πλεονάσματος κατά απόλυτη τιμή μπορεί να πάρει αυστηρά μόνο θετικές τιμές και συμβολίζεται με L_1.

Συγκεκριμένα έστω ότι η πρώτη πτώση του πλεονάσματος κάτω από την τιμή u, πραγματοποιείται την χρονική στιγμή t_1 με πλεόνασμα $u_1 = U(t_1)$. Τότε για την τιμή που θα έχει η τυχαία μεταβήτή L_1 ισχύει:

$$L_1 = u - u_1.$$

Αντίστοιχα μία τυχαία μεταβλητή L_2 θα συμβολίζει την πρώτη πτώση του πλεονάσματος κάτω από την τιμή u_1 και η τιμή που θα παίρνει θα είναι:

$$L_2 = u_1 - u_2.$$

Κατ'αυτόν τον τρόπο επαγωγικά μπορούμε να ορίσουμε μία ακολουθία $L_1, L_2, \ldots$ όπου:

$$L_\kappa = u_{\kappa-1} - u_\kappa.$$

Οι μεταβλητές $L_1, L_2, \ldots, L_n, \ldots$ ονομάζονται και κλιμακωτά ύψη που συνδέονται με την ανέλιξη $\{U(t): t \geq 0\}$, και απεικονίζουν την σταδιακή πτώση του πλεονάσματος από την αρχική τιμή u έως την στιγμή της χρεοκοπίας, ή σε περίπτωση που δεν επέλθει χρεοκοπία, έως την ελάχιστη τιμή που μπορεί να πάρει η ανέλιξη

$$\{U(t): t \geq 0\}.$$

Ορίζουμε τώρα μία τυχαία μεταβλητή K, η οποία δηλώνει το πλήθος των κλιμακωτών υψών (το πλήθος δηλαδή των L_i) σε μία ανέλιξη πλενάσματος. Αυτό το πλήθος είναι αντιληπτό ότι είναι πεπερασμένο, επομένως και η μεταβλητή K θα είναι διακριτή και θα παίρνει ακέραιες και μη αρνητικές τιμές. Είναι επίσης χρήσιμο να αναφερθεί ότι η τυχαία μεταβλητή K ακολουθεί την γεωμετρική κατανομή, διότι κάθε φορά που εμφανίζεται ένα νέο L_i, η πιθανότητα να υπάρξει και νέα πτώση είναι ίση με $\psi(0)$. Τότε αν θεωρήσουμε «αποτυχία» την εμφάνιση ενός L_i, η μεταβλητή K θα μετρά τον αριθμό των αποτυχιών έως ότου επέλθει η πρώτη επιτυχία και η κατανομή της δίνεται από την σχέση:

$$P(K=\kappa) = [\psi(0)]^\kappa \delta(0), \kappa = 0,1,2, \ldots,$$

η παραπάνω σχέση μπορεί να γραφτεί και στην μορφή:

$$P(K=\kappa) = \left(\frac{1}{1+\theta}\right)^\kappa \frac{\theta}{1+\theta},$$

διότι είδαμε στην προηγούμενη ενότητα ότι ισχύει:

$$\psi(0) = \frac{1}{1+\theta}\ ,\ \delta(0) = \frac{\theta}{1+\theta}\ .$$

Θεωρούμε τώρα την σύνθετη τυχαία μεταβλητή:

$$L = L_1 + L_2 + \cdots + L_K = \sum_{i=1}^{K} L_i \quad (L = 0 \text{ όταν } K = 0),$$

Η L ονομάζεται μέγιστη σωρευτική απώλεια και απεικονίζει την συνολική πτώση του πλεονάσματος κάτω από το αρχικό αποθεματικό u. Αναφέραμε πριν ότι η K ακολουθεί την γεωμετρική κατανομή, επομένως η L θα είναι μία σύνθετη γεωμετρική κατανομή, για την οποία θα ισχύει ότι:

$$P(L=0) = P(K=0) = \delta(0).$$

Επίσης θα ισχύει και:

$P(L > u) = \psi(u)$ και $P(L \leq u) = \delta(u)$.

ΚΕΦΑΛΑΙΟ 5: ΠΡΟΣΕΓΓΙΣΕΙΣ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΑΠΟ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Στο προηγούμενο κεφάλαιο είδαμε βασικές έννοιες της θεωρίας χρεοκοπίας και εξηγήσαμε ότι η πιθανότητα χρεοκοπίας στις περισσότερες περιπτώσεις δεν είναι εύκολο να βρεθεί. Οι περιπτώσεις στις οποίες είναι εύκολο να βρεθεί είναι συνήθως όταν η κατανομή των αποζημιώσεων ακολουθεί είτε την εκθετική κατανομή είτε μία μείξη εκθετικών κατανομών, είτε μία μείξη Γάμμα κατανομών με την πρώτη παράμετρο να ισούται με 2. Στο Κεφάλαιο 3 προσεγγίσαμε την συνάρτηση πυκνότητας πιθανότητας τριών τυχαίων συνεχών μεταβλητών από μία μείξη εκθετικών κατανομών. Σε αυτό το κεφάλαιο θα αναφέρουμε 2 μεθόδους για την εύρεση της πιθανότητας χρεοκοπίας, όταν η κατανομή των αποζημιώσεων είναι μία μείξη εκθετικών κατανομών και μία μέθοδο όταν είναι μία μείξη Γάμμα κατανομών με παραμέτρους $(2, \beta_j)$ και στην συνέχεια θα υπολογίσουμε προσεγγιστικά την πιθανότητα χρεοκοπίας των συναρτήσεων πυκνότητας πιθανότητας των κατανομών Pareto, Weibull και Lognormal, έχοντας κατανομές των αποζημιώσεων τις αντίστοιχες μείξεις εκθετικών κατανομών με τις οποίες τις προσεγγίσαμε στα παραδείγματα 3.3.1, 3.4.1 και 3.5.1 αντίστοιχα.

5.1: ΕΥΡΕΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΓΙΑ ΜΙΑ ΜΕΙΞΗ ΕΚΘΕΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ.

Σε μία μείξη εκθετικών κατανομών η πιθανότητα χρεοκοπίας σύμφωνα με άρθρο των Gerber et.al (1987), υπολογίζεται απο την σχέση:

$$\psi(u) = \sum_{\kappa=1}^{n} C_{\kappa} e^{-r_{\kappa} u}. \qquad (5.1.1)$$

Όπου :

- Οι r_{κ} είναι ρίζες της εξίσωσης:
  $\frac{\lambda}{c}\sum_{j=1}^{n}\frac{\alpha_j}{\beta_j - r} = 1 \quad (r_1 < r_2 < \cdots < r_k)$ (εξίσωση του Lundberg) (5.1.2)
  Εδώ αξίζει να σημειωθεί ότι $r_1 = R$ (ο συντελεστής προσαρμογής που προαναφέραμε).

- Οι $C_{j\kappa}$ είναι αριθμοί που υπολογίζονται από την σχέση:
  $$C_{j\kappa} = \frac{\frac{\alpha_j}{\beta_j - r_{\kappa}}}{\sum_{j=1}^{n}\frac{\alpha_j}{(\beta_j - r_{\kappa})^2}}., \qquad (5.1.3)$$
- Οι C_{κ} είναι σταθερές που υπολογίζονται από την σχέση:
  $$C_{\kappa} = \sum_{j=1}^{n}\frac{C_{j\kappa}}{\beta_j}. \qquad (5.1.4)$$

Επίσης σύμφωνα με τις σημειώσεις του Πολίτη (2015) ένας άλλος τρόπος υπολογισμού της πιθανότητας χρεοκοπίας, στην ειδική περίπτωση που η κατανομή των αποζημιώσεων ακολουθεί μία μείξη δύο εκθετικών κατανομών ($\kappa = 2$) με παραμέτρους $\beta_1, \beta_2 > 0$, είναι ο εξής:

$$\psi(u) = C_1 e^{-r_1 u} + C_2 e^{-r_2 u}. \qquad (5.1.5)$$

Οι r_1, r_2, είναι οι ρίζες που προκύπτουν από την εξίσωση του Lundberg:

$$1 + (1+\theta)E(X)r = M_X(r), \qquad (5.1.6)$$

όπου $E(X)$, $M_X(r)$, αντιπροσωπεύουν την μέση τιμή και την ροπογεννήτρια των αποζημιώσεων αντίστοιχα. Αν υποθέσουμε, χωρίς βλάβη της γενικότητας ότι: $\beta_1 < \beta_2$, τότε για τις ρίζες: r_1, r_2 $(r_1 = R < r_2)$, θα πρέπει να ισχύει ότι:

$$0 < r_1 < \beta_1 < r_2 < \beta_2. \qquad (5.1.7)$$

Στο κεφάλαιο 4 είδαμε ότι για $u = 0$, η πιθανότητα χρεοκοπίας είναι:

$$\psi(0) = \frac{1}{1+\theta},$$

όπου θ, το περιθώριο ασφαλείας που γράφεται στην μορφή:

$$\theta = \frac{c}{\lambda E(X)} - 1,$$

όπου $E(X)$ η κατανομή των αποζημιώσεων.

Έτσι η σχέση (5.1.5) θα γράφεται στην μορφή:

$$\psi(0) = C_1 + C_2,$$

και είναι προφανές ότι θα ισχύει:

$$C_1 + C_2 = \frac{1}{1+\theta}, \quad (5.1.8)$$

επίσης σύμφωνα με τις σημειώσεις του Πολίτη (2015), ισχύει η παρακάτω εξίσωση:

$$C_1 r_1 + C_2 r_2 = \frac{\theta}{(1+\theta)^2 E(X)}. \quad (5.1.9)$$

Οι σταθερές C_1, C_2, προκύπτουν από την επίλυση του συστήματος των δύο παραπάνω εξισώσεων και με αυτό τον τρόπο υπολογίζουμε την πιθανότητα χρεοκοπίας.

Παράδειγμα 5.1.1

Υποθέτουμε ότι:

$$\lambda = c \ \ (\lambda, c \neq 0),$$

και έχουμε την συνάρτηση:

$$f_X(x) = 4e^{-12x} + 10e^{-15x}, x \geq 0,$$

η οποία συμπεραίνουμε ότι γράφεται στην μορφή :

$$f_X(x) = \frac{1}{3}12e^{-12x} + \frac{2}{3}15e^{-15x}, x \geq 0.$$

Παρατηρούμε ότι:

$$\alpha_1 = \frac{1}{3}, \beta_1 = 12, \alpha_2 = \frac{2}{3}, \beta_2 = 15.$$

Θα υπολογίσουμε τις ρίζες r_1, r_2, από την εξίσωση (5.1.6) και θα ισχύει ότι:

$$1 + (1+\theta)E(X)r = M_X(r).$$

Η μέση τιμή της $f_X(x)$ και το περιθώριο ασφαλείας θα είναι:

$$E(X) = \frac{1}{3} \cdot \frac{1}{12} + \frac{2}{3} \cdot \frac{1}{15} \approx 0.072,$$

$$\theta = \frac{c}{\lambda E(X)} - 1 \approx \frac{1}{0.072} - 1 \approx 12.89.$$

Η ροπογεννήτρια των αποζημιώσεων, θα είναι της μορφής:

$$M_X(r) = a_1 \frac{\beta_1}{\beta_1 - r} + a_2 \frac{\beta_2}{\beta_2 - r} = \frac{1}{3} \cdot \frac{12}{12-r} + \frac{2}{3} \cdot \frac{15}{15-r}.$$

Συνεπώς θα ισχύει:

$$1 + (1 + 12.89)0.072r = \frac{1}{3} \cdot \frac{12}{12-r} + \frac{2}{3} \cdot \frac{15}{15-r}.$$

Οι ρίζες που προκύπτουν από την παραπάνω εξίσωση είναι οι :

$r_0 = 0$ (Απορρίπτεται λόγω της ανισότητας (5.1.7), $r_1 = 11.5858, r_2 = 14.4143$.

Θα υπολογίσουμε τις σταθερές C_1, C_2, από το σύστημα των εξισώσεων (5.1.8) και (5.1.9) και θα ισχύει ότι:

$$C_1 + C_2 = \frac{1}{1+\theta} = \frac{1}{1+12.89} = 0.072,$$

$$C_1 r_1 + C_2 r_2 = \frac{\theta}{(1+\theta)^2 E(X)} = \frac{12.89}{(1+12.89)^2 0.072} = 0.9279.$$

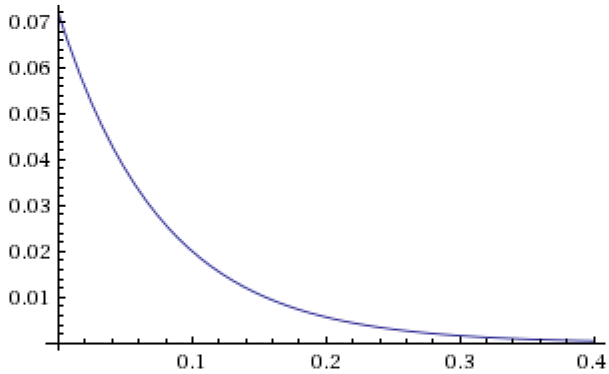
Οι λύσεις που προκύπτουν από το παραπάνω σύστημα είναι οι:

$$C_1 = 0.0389, \quad C_2 = 0.033.$$

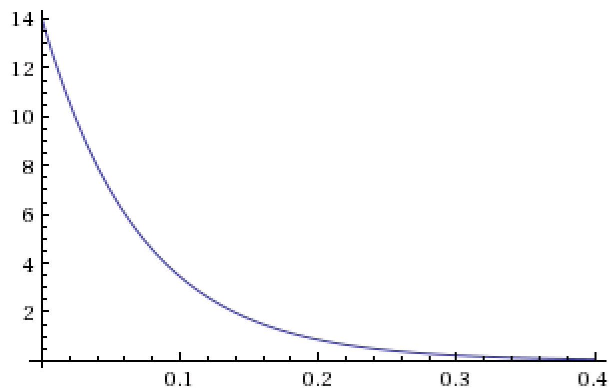
Επομένως από την σχέση (5.1.5), η πιθανότητα χρεοκοπίας θα είναι της μορφής:

$$\psi(u) = C_1 e^{-r_1 u} + C_2 e^{-r_2 u} = 0.0389 e^{-11.5858u} + 0.033 e^{-14.4143u}, u \geq 0.$$

Στα παρακάτω σχήματα βλέπουμε τις γραφικές παραστάσεις της πιθανότητας χρεοκοπίας $\psi(u)$ και της αρχικής συνάρτησης πυκνότητας $f(x)$ αντίστοιχα.



Σχήμα 5.1 : Διάγραμμα της πιθανότητας χρεοκοπίας.



Σχήμα 5.2: Διάγραμμα της κατανομής των αποζημιώσεων.

5.2: ΕΥΡΕΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΓΙΑ ΜΙΑ ΜΕΙΞΗ ΓΑΜΜΑ $(2, \beta_j)$ ΚΑΤΑΝΟΜΩΝ.

Στην Ενότητα 3.6 περιγράψαμε αναλυτικά τη μείξη Γάμμα κατανομών με παραμέτρους $2, \beta_j$, και είδαμε ότι όταν μία συνάρτηση μπορεί να γραφτεί στην μορφή:

$$f_X(x) = \sum_{j=1}^{n} \alpha_j {\beta_j}^2 x e^{-\beta_j x}, (\beta_j, x > 0) ,$$

για την οποία ισχύει ότι:

$$\sum_{j=1}^{n} \alpha_j = 1,$$

τότε αποτελεί μία μείξη Γάμμα κατανομών με παραμέτρους $2, \beta_j$. Στην παρούσα Ενότητα θα περιγράψουμε την μέθοδο υπολογισμού της πιθανότητας χρεοκοπίας για μία μείξη Γάμμα κατανομών με παραμέτρους $2, \beta_j$.

Η μέθοδος για τον υπολογισμό της πιθανότητας χρεοκοπίας σε μία μείξη Γάμμα κατανομών είναι παρόμοια με εκείνης της μείξης εκθετικών κατανομών και σύμφωνα επίσης με άρθρο των Gerber et.al (1987), η πιθανότητα χρεοκοπίας υπολογίζεται από την σχέση:

$$\psi(u) = \sum_{\kappa=1}^{2n} C_\kappa e^{-r_\kappa u} \ . \tag{5.2.1}$$

- *Όπου* r_κ ειναι οι ρίζες που προκύπτουν από την σχέση:
  $$1 - \frac{\lambda}{c}\sum_{j=1}^{n} \alpha_j \frac{2\beta_j - r}{(\beta_j - r)^2} = 0. \tag{5.2.2}$$
- Όπου C_κ είναι αριθμοί που προκύπτουν από την σχέση:

$$C_\kappa = \frac{\sum_{j=1}^{n} \alpha_j \frac{3-2\frac{r_k}{\beta_j}}{(\beta_j - r_\kappa)^2}}{\sum_{j=1}^{n} \alpha_j \frac{3\beta_j - r_\kappa}{(\beta_j - r_\kappa)^3}} \ . \tag{5.2.3}$$

Παράδειγμα 5.2.1

Έχουμε: $\lambda = 1, c = 2$ και την συνάρτηση:

$$f_X(x) = \frac{1}{2} x e^{-x} + \frac{1}{2} 25 x e^{-5x} = \frac{1}{2} 1^2 x e^{-x} + \frac{1}{2} 5^2 x e^{-5x} \ \ (x > 0).$$

η οποία αποτελεί μία μείξη Γάμμα $\left(2, \beta_j\right)$ κατανομών με:

$$a_1 \left(= \frac{1}{2}\right) + a_2 \left(= \frac{1}{2}\right) = 1.$$

και:

$$\beta_1 = 1, \beta_2 = 5.$$

Επιθυμούμε να υπολογίσουμε την πιθανότητα χρεοκοπίας, οπότε θα υπολογίσουμε πρώτα τις ρίζες r_κ από την σχέση (5.2.2). Έτσι θα ισχύει:

$$1 - \frac{1}{2}\left(\alpha_1 \frac{2\beta_1 - r}{(\beta_1 - r)^2} + \alpha_2 \frac{2\beta_2 - r}{(\beta_2 - r)^2}\right) = 0 \Rightarrow 1 - \frac{1}{2}\left(0.5 \frac{2 - r}{(1 - r)^2} + 0.5 \frac{10 - r}{(5 - r)^2}\right) = 0.$$

Με χρήση του Mathematica οι ρίζες της παραπάνω τεταρτοβάθμιας εξίσωσης βρίσκουμε ότι είναι οι:

$$r_1 = R \approx 0.31, \ \ r_2 \approx 1.42, \ \ r_3 \approx 3.79, \ \ r_4 \approx 5.99.$$

Θα υπολογίσουμε τους αριθμούς C_κ από την εξίσωση (5.2.3), οπότε θα ισχύει:

$$C_1 = \frac{\alpha_1 \frac{3-2\frac{r_1}{\beta_1}}{(\beta_1-r_1)^2} + \alpha_2 \frac{3-2\frac{r_1}{\beta_2}}{(\beta_2-r_1)^2}}{\alpha_1 \frac{3\beta_1-r_1}{(\beta_1-r_1)^3} + \alpha_2 \frac{3\beta_2-r_1}{(\beta_2-r_1)^3}} = \frac{\frac{1}{2}\frac{3-2\frac{0.31}{1}}{(1-0.31)^2} + \frac{1}{2}\frac{3-2\frac{0.31}{5}}{(5-0.31)^2}}{\frac{1}{2}\frac{3\times1-0.31}{(1-0.31)^3} + \frac{1}{2}\frac{3\times5-0.31}{(5-0.31)^3}} \Rightarrow$$

$$\Rightarrow C_1 \approx 0.616,$$

$$C_2 = \frac{\alpha_1 \frac{3-2\frac{r_2}{\beta_1}}{(\beta_1-r_2)^2} + \alpha_2 \frac{3-2\frac{r_2}{\beta_2}}{(\beta_2-r_2)^2}}{\alpha_1 \frac{3\beta_1-r_2}{(\beta_1-r_2)^3} + \alpha_2 \frac{3\beta_2-r_2}{(\beta_2-r_2)^3}} = \frac{\frac{1}{2}\frac{3-2\frac{1.42}{1}}{(1-1.42)^2} + \frac{1}{2}\frac{3-2\frac{1.42}{5}}{(5-1.42)^2}}{\frac{1}{2}\frac{3\times1-1.42}{(1-1.42)^3} + \frac{1}{2}\frac{3\times5-1.42}{(5-1.42)^3}} \Rightarrow$$

$$\Rightarrow C_2 = -0.052,$$

$$C_3 = \frac{\alpha_1 \frac{3-2\frac{r_3}{\beta_1}}{(\beta_1-r_3)^2} + \alpha_2 \frac{3-2\frac{r_3}{\beta_2}}{(\beta_2-r_3)^2}}{\alpha_1 \frac{3\beta_1-r_3}{(\beta_1-r_3)^3} + \alpha_2 \frac{3\beta_2-r_3}{(\beta_2-r_3)^3}} = \frac{\frac{1}{2}\frac{3-2\frac{3.79}{1}}{(1-3.79)^2} + \frac{1}{2}\frac{3-2\frac{3.79}{5}}{(5-3.79)^2}}{\frac{1}{2}\frac{3\times1-3.79}{(1-3.79)^3} + \frac{1}{2}\frac{3\times5-3.79}{(5-3.79)^3}} \Rightarrow$$

$$\Rightarrow C_3 \approx 0.067,$$

$$C_4 = \frac{\alpha_1 \frac{3-2\frac{r_4}{\beta_1}}{(\beta_1-r_4)^2} + \alpha_2 \frac{3-2\frac{r_4}{\beta_2}}{(\beta_2-r_4)^2}}{\alpha_1 \frac{3\beta_1-r_4}{(\beta_1-r_4)^3} + \alpha_2 \frac{3\beta_2-r_4}{(\beta_2-r_4)^3}} = \frac{\frac{1}{2}\frac{3-2\frac{5.99}{1}}{(1-5.99)^2} + \frac{1}{2}\frac{3-2\frac{5.99}{5}}{(5-5.99)^2}}{\frac{1}{2}\frac{3\times1-5.99}{(1-5.99)^3} + \frac{1}{2}\frac{3\times5-5.99}{(5-5.99)^3}} \Rightarrow$$

$$\Rightarrow C_4 \approx -0.027.$$

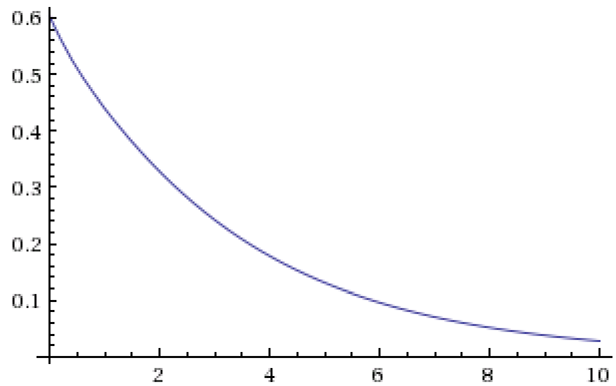
Τέλος θα υπολογίσουμε την πιθανότητα χρεοκοπίας σύμφωνα με την σχέση (5.2.1), η οποία θα γράφεται στην μορφή:

$$\psi(u) = \sum_{\kappa=1}^{2n} C_\kappa e^{-r_\kappa u} = \sum_{\kappa=1}^{4} C_\kappa e^{-r_\kappa u} =$$

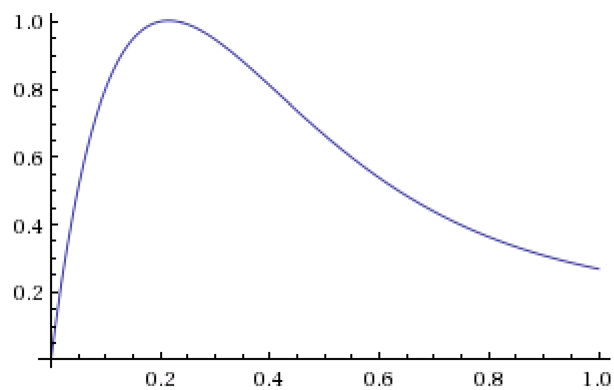
$$= C_1 e^{-r_1 u} + C_2 e^{-r_2 u} + C_3 e^{-r_3 u} + C_4 e^{r_4 u} \Rightarrow$$

$$\Rightarrow \psi(u) \approx 0.616 e^{-0.31u} - 0.052 e^{-1.41u} + 0.067 e^{-3.79u} - 0.027 e^{-5.99u}, u \geq 0.$$

Στα παρακάτω σχήματα βλέπουμε τις γραφικές παραστάσεις της πιθανότητας χρεοκοπίας $\psi(u)$ και της αρχικής συνάρτησης $f_Y(y)$ αντίστοιχα.



Σχήμα 5.3: Διάγραμμα πιθανότητας χρεοκοπίας.



Σχήμα 5.4: Διάγραμμα της κατανομής των αποζημιώσεων.

Από το γράφημα της πιθανότητας χρεοκοπίας βλέπουμε ότι ισχύει: $\psi(0) = 0.6$. Συνεπώς θα πρέπει να ισχύει και

$$\frac{1}{1+\theta} = 0.6.$$

Το περιθώριο ασφαλείας γνωρίζουμε ότι υπολογίζεται ως εξής:

$$\theta = \frac{c}{\lambda E(X)} - 1,$$

Η μέση τιμή της κατανομής των αποζημιώσεων θα είναι:

$$E(X) = \frac{1}{2} \cdot \frac{2}{1} + \frac{1}{2} \cdot \frac{2}{5} = \frac{6}{5},$$

Επομένως θα ισχύει ότι:

$$\theta = \frac{2}{1 \cdot \frac{6}{5}} - 1 = \frac{2}{3} \Rightarrow \frac{1}{1+\theta} = 0.6 = \psi(0).$$

5.3: ΠΡΟΣΕΓΓΙΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΓΙΑ ΜΙΑ ΚΑΤΑΝΟΜΗ ΑΠΟΖΗΜΙΩΣΕΩΝ, Η ΟΠΟΙΑ ΑΚΟΛΟΥΘΕΙ ΤΗΝ ΚΑΤΑΝΟΜΗ PARETO.

Στην εισαγωγή του κεφαλαίου εξηγήσαμε ότι ο υπολογισμός της πιθανότητας χρεοκοπίας σε μία κατανομή αποζημιώσεων, είναι εύκολο να επιτευχθεί όταν η συγκεκριμένη κατανομή ακολουθεί είτε την εκθετική κατανομή είτε μία μείξη εκθετικών (είτε μία μείξη γάμμα με την πρώτη παράμετρο α να είναι ίση με 2). Αν όμως ακολουθεί μία άλλη κατανομή, όπως για παράδειγμα την Pareto, τότε την προσεγγίζουμε από μία μείξη εκθετικών κατανομών όπως είδαμε στο κεφάλαιο 3 και έπειτα υπολογίζουμε (προσεγγιστικά) την πιθανότητα χρεοκοπίας όπως εξηγήσαμε στην ενότητα 5.1. Στο παράδειγμα 3.3.1 προσεγγίσαμε την συνάρτηση πυκνότητας πιθανότητας μίας τυχαίας συνεχούς μεταβλητής Y, η οποία ακολουθεί την κατανομή Pareto με παραμέτρους $\alpha = 6, \beta = 2$ $\big(Y \sim Pa(6,2)\big)$, από μία μείξη δύο εκθετικών κατανομών:

$$f_X(x) = 0.1e^{-x} + 2.7e^{-3x}, x \geq 0,$$

με παραμέτρους:

$$\beta_1 = 1, \quad \beta_2 = 3$$

και αντίστοιχα βάρη:

$$A = \alpha_1 = 0.1, \quad 1 - A = \alpha_2 = 0.9.$$

Θέλουμε τότε να υπολογίσουμε την πιθανότητα χρεοκοπίας αν $\lambda = c$, έχοντας κατανομή των αποζημιώσεων την $f_X(x)$.

Στο παράδειγμα 3.3.1 είδαμε ότι η μέση τιμή της μείξης εκθετικών κατανομών έχει τιμή ίση με 0.4. Το περιθώριο ασφαλείας τότε θα έχει την τιμή:

$$\theta = \frac{c}{\lambda E(X)} - 1 = \frac{1}{0.4} - 1 = \frac{3}{2}.$$

Επομένως για $u = 0$, στην πιθανότητα χρεοκοπίας θα πρέπει να ισχύει:

$$\psi(0) = \frac{1}{1+\theta} = \frac{2}{5},$$

Καθώς επίσης για την μέγιστη σωρευτική απώλεια, την οποία αναφέραμε στο προηγούμενο κεφάλαιο θα ισχύει:

$$P(L = 0) = \delta(0) = \frac{\theta}{1+\theta} = \frac{3}{5}.$$

Θα υπολογίσουμε τις ρίζες r_1, r_2, από την εξίσωση (5.1.2) και θα ισχύει ότι:

$$\frac{\lambda}{c}\left(\frac{a_1}{\beta_1 - r} + \frac{\alpha_2}{\beta_2 - r}\right) = 1 \Rightarrow \frac{0.1}{1-r} + \frac{0.9}{3-r} = 1.$$

Οι ρίζες που προκύπτουν από την παραπάνω δευτεροβάθμια εξίσωση είναι οι:

$$r_1 = R = 0.82918, \qquad r_2 = 2.17082$$

Από την εξίσωση (5.1.3) θα υπολογίσουμε τις σταθερές $C_{11}, C_{12}, C_{21}, C_{22}$, ως εξής:

$$C_{11} = \frac{\frac{\alpha_1}{\beta_1 - r_1}}{\frac{\alpha_1}{(\beta_1 - r_1)^2} + \frac{\alpha_2}{(\beta_2 - r_1)^2}} = \frac{\frac{0.1}{1 - 0.82918}}{\frac{0.1}{(1-0.82918)^2} + \frac{0.9}{(3-0.82918)^2}} = 0.1618$$

$$C_{12} = \frac{\frac{\alpha_1}{\beta_1 - r_2}}{\frac{\alpha_1}{(\beta_1 - r_2)^2} + \frac{\alpha_2}{(\beta_2 - r_2)^2}} = \frac{\frac{0.1}{1 - 2.17082}}{\frac{0.1}{(1-2.17082)^2} + \frac{0.9}{(3-2.17082)^2}} = -0.0618$$

$$C_{21} = \frac{\frac{\alpha_2}{\beta_2 - r_1}}{\frac{\alpha_1}{(\beta_1 - r_1)^2} + \frac{\alpha_2}{(\beta_2 - r_1)^2}} = \frac{\frac{0.9}{3 - 0.82918}}{\frac{0.1}{(1-0.82918)^2} + \frac{0.9}{(3-0.82918)^2}} = 0.1146$$

$$C_{22} = \frac{\frac{\alpha_2}{\beta_2 - r_2}}{\frac{\alpha_1}{(\beta_1 - r_2)^2} + \frac{\alpha_2}{(\beta_2 - r_2)^2}} = \frac{\frac{0.9}{3 - 2.17082}}{\frac{0.1}{(1-2.17082)^2} + \frac{0.9}{(3-2.17082)^2}} = 0.7854$$

Σύμφωνα με την σχέση (5.1.4), για τις σταθερές C_1, C_2, θα ισχύει:

$$C_1 = \frac{C_{11}}{\beta_1} + \frac{C_{21}}{\beta_2} = \frac{0.1618}{1} + \frac{0.1146}{3} = \frac{1}{5},$$

$$C_2 = \frac{C_{12}}{\beta_1} + \frac{C_{22}}{\beta_2} = \frac{-0.0618}{1} + \frac{0.7854}{3} = \frac{1}{5}.$$

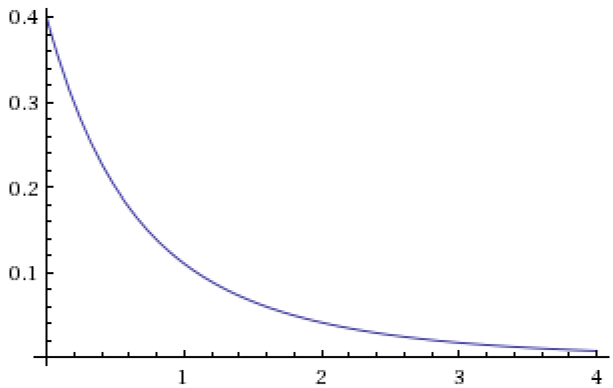
Συνεπώς η πιθανότητα χρεοκοπίας από την σχέση (5.1.1), θα είναι της μορφής:

$$\psi(u) = C_1 e^{-r_1 u} + C_2 e^{-r_2 u} = \frac{1}{5} e^{-0.82918u} + \frac{1}{5} e^{-2.17082u}, u \geq 0.$$

Παρατηρούμε από την παραπάνω σχέση για $u = 0$, ότι ισχύει:

$$\psi(0) = \frac{2}{5},$$

όπως ήταν αναμενόμενο. Στο παρακάτω διάγραμμα αποτυπώνεται η γραφική παράσταση της πιθανότητας χρεοκοπίας.



Σχήμα 5.5: Διάγραμμα της πιθανότητας χρεοκοπίας.

5.4: ΠΡΟΣΕΓΓΙΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΓΙΑ ΜΙΑ ΚΑΤΑΝΟΜΗ ΑΠΟΖΗΜΙΩΣΕΩΝ, Η ΟΠΟΙΑ ΑΚΟΛΟΥΘΕΙ ΤΗΝ ΚΑΤΑΝΟΜΗ WEIBULL.

Στο κεφάλαιο 3, προσεγγίσαμε τις συναρτήσεις πυκνότητας πιθανότητας τριών τυχαίων μεταβλητών, οι οποίες ακολουθούν τις κατανομές: Pareto, Weibull και lognormal, από μία μείξη δύο εκθετικών κατανομών και φροντίσαμε έτσι ώστε να έχουν και οι τρεις την ίδια μέση τιμή. Κατά την διάρκεια του κεφαλαίου, θα διαπιστώσουμε ότι είναι ιδιαίτερα σημαντικό να έχουν την ίδια μέση τιμή, διότι το περιθώριο ασφαλείας, είναι:

$$\theta = \frac{c}{\lambda E(X)} - 1.$$

Συνεπώς για δύο (ή και περισσότερες) κατανομές αποζημιώσεων $f_X(x), f_Y(y)$, εφόσον έχουν την ίδια μέση τιμή, ισχύει $c = \lambda$ και θ_1, θ_2, είναι τα περιθώρια ασφαλείας αντίστοιχα, θα ισχύει και:

$$\theta_1 = \frac{c}{\lambda E(X)} - 1, \quad \theta_2 = \frac{c}{\lambda E(Y)} - 1,$$

επομένως θα ισχύει ότι: $\theta_1 = \theta_2$ και για $u = 0$, η πιθανότητα χρεοκοπίας θα έχει την ίδια τιμή, διότι:

$$\psi(0) = \frac{1}{1+\theta}\,.$$

Στην ενότητα 3.4 προσεγγίσαμε μία συνεχή τυχαία μεταβλητή Y, η οποία ακολουθεί την κατανομή Weibull με παραμέτρους $\lambda = 2.63045, \alpha = 0.9$, από μία μείξη δύο εκθετικών κατανομών $f_X(x)$, για την οποία είδαμε ότι ισχύει:

$$A = a_1 = 0.324247, \quad 1 - A = a_2 = 0.675753, \quad \beta_1 = 4.97489, \quad \beta_2 = 2.01824.$$

Σε αυτή την Ενότητα θα υπολογίσουμε την πιθανότητα χρεοκοπίας της $f_X(x)$, για $\lambda = c$ $(\lambda, c \neq 0)$. Η μέση τιμή της $E(X)$ είναι ίση με $\frac{2}{5}$, οπότε το περιθώριο ασφαλείας θ, θα έχει την ίδια τιμή, με εκείνο της προηγούμενης ενότητας, όπως και η πιθανότητα χρεοκοπίας για $u = 0$.

Από την εξίσωση του Lundberg (5.1.2) θα ισχύει ότι:

$$\frac{\lambda}{c}\left(\frac{a_1}{\beta_1 - r} + \frac{\alpha_2}{\beta_2 - r}\right) = 1 \Rightarrow \frac{0.324247}{4.97489 - r} + \frac{0.675753}{2.01824 - r} = 1.$$

Οι τιμές που ικανοποιούν την παραπάνω εξίσωση με χρήση του Mathematica βρίσκουμε ότι είναι οι:

$$r_1 = R = 1.27753, \quad r_2 = 4.7156.$$

Για τις σταθερές $C_{11}, C_{12}, C_{21}, C_{22}$, θα ισχύει ότι:

$$C_{11} = \frac{\frac{\alpha_1}{\beta_1 - r_1}}{\frac{\alpha_1}{(\beta_1 - r_1)^2} + \frac{\alpha_2}{(\beta_2 - r_1)^2}} = \frac{\frac{0.324247}{4.97489 - 1.27753}}{\frac{0.324247}{(4.97489 - 1.27753)^2} + \frac{0.675753}{(2.01824 - 1.27753)^2}} =$$

$$= 0.06985,$$

$$C_{12} = \frac{\frac{\alpha_1}{\beta_1 - r_2}}{\frac{\alpha_1}{(\beta_1 - r_2)^2} + \frac{\alpha_2}{(\beta_2 - r_2)^2}} = \frac{\frac{0.324247}{4.97489 - 4.7156}}{\frac{0.324247}{(4.97489 - 4.7156)^2} + \frac{0.675753}{(2.01824 - 4.7156)^2}} =$$

$$= 0.2544,$$

$$C_{21} = \frac{\frac{\alpha_2}{\beta_2 - r_1}}{\frac{\alpha_1}{(\beta_1 - r_1)^2} + \frac{\alpha_2}{(\beta_2 - r_1)^2}} = \frac{\frac{0.675753}{2.01824 - 1.27753}}{\frac{0.324247}{(4.97489 - 1.27753)^2} + \frac{0.675753}{(2.01824 - 1.27753)^2}} =$$

$$= 0.7267,$$

$$C_{22} = \frac{\frac{\alpha_2}{\beta_2 - r_2}}{\frac{\alpha_1}{(\beta_1 - r_2)^2} + \frac{\alpha_2}{(\beta_2 - r_2)^2}} = \frac{\frac{0.675753}{2.01824 - 4.7156}}{\frac{0.324247}{(4.97489 - 4.7156)^2} + \frac{0.675753}{(2.01824 - 4.7156)^2}} =$$

$$= -0.051,$$

Και για τις C_1, C_2, θα ισχύει ότι:

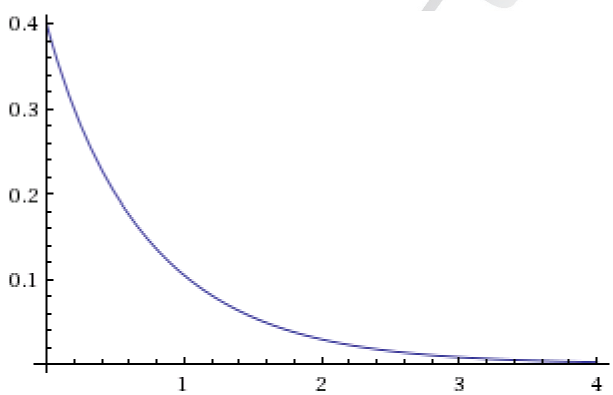
$$C_1 = \frac{C_{11}}{\beta_1} + \frac{C_{21}}{\beta_2} = \frac{0.06985}{4.97489} + \frac{0.7267}{2.01824} = 0.3741,$$

$$C_2 = \frac{C_{12}}{\beta_1} + \frac{C_{22}}{\beta_2} = \frac{0.2544}{4.97489} - \frac{0.051}{2.01824} = 0.02587,$$

και η πιθανότητα χρεοκοπίας θα είναι:

$$\psi(u) = C_1 e^{-r_1 u} + C_2 e^{-r_2 u} = 0.3741 e^{-1.27753u} + 0.02587 e^{-4.7156u}, u \geq 0.$$

Την γραφική παράσταση της $\psi(u)$, μπορούμε να την παρατηρήσουμε στο παρακάτω γράφημα.



Σχήμα 5.6: Διάγραμμα της πιθανότητας χρεοκοπίας.

5.5: ΠΡΟΣΕΓΓΙΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΧΡΕΟΚΟΠΙΑΣ ΓΙΑ ΜΙΑ ΚΑΤΑΝΟΜΗ ΑΠΟΖΗΜΙΩΣΕΩΝ, Η ΟΠΟΙΑ ΑΚΟΛΟΥΘΕΙ ΤΗΝ ΛΟΓΑΡΙΘΜΟΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ.

Στην ενότητα 3.5 προσεγγίσαμε την συνάρτηση πυκνότητας πιθανότητας μίας συνεχούς τυχαίας μεταβλητής Y, η οποία ακολουθεί την λογαριθμοκανονική κατανομή με παραμέτρους $\mu = -1.41629, \sigma = 1$, από μία μείξη εκθετικών κατανομών $f_X(x)$ με παραμέτρους:

$$\beta_1 = 0.539321, \;\; \beta_2 = 2.77405$$

και αντίστοιχα βάρη:

$$A = \alpha_1 = 0.02645, \; 1 - A = \alpha_2 = 0.97355.$$

Όπως και στις δύο προηγούμενες ενότητες η μέση τιμή είναι $E(X) = \frac{2}{5}$. Θα υπολογίσουμε την πιθανότητα χρεοκοπίες της $f_X(x)$ για $\lambda = c$ $(\lambda, c \neq 0)$.

Οι ρίζες r_1, r_2, θα προκύπτουν από την παρακάτω εξίσωση :

$$\frac{\lambda}{c}\left(\frac{\alpha_1}{\beta_1 - r} + \frac{\alpha_2}{\beta_2 - r}\right) = 1 \Rightarrow \frac{0.02645}{0.539321 - r} + \frac{0.97355}{2.77405 - r} = 1.$$

Οι τιμές των r_1, r_2, είναι:

$$r_1 = R = 0.493174, \qquad r_2 = 1.8202.$$

Οι σταθερές $C_{11}, C_{12}, C_{21}, C_{22}$, θα είναι της μορφής:

$$C_{11} = \frac{\frac{\alpha_1}{\beta_1 - r_1}}{\frac{\alpha_1}{(\beta_1 - r_1)^2} + \frac{\alpha_2}{(\beta_2 - r_1)^2}} = \frac{\frac{0.02645}{0.539321 - 0.493174}}{\frac{0.02645}{(0.539321 - 0.493174)^2} + \frac{0.97355}{(2.77405 - 0.493174)^2}} =$$

$$= 0.04546,$$

$$C_{12} = \frac{\frac{\alpha_1}{\beta_1 - r_2}}{\frac{\alpha_1}{(\beta_1 - r_2)^2} + \frac{\alpha_2}{(\beta_2 - r_2)^2}} = \frac{\frac{0.02645}{0.539321 - 1.8202}}{\frac{0.02645}{(0.539321 - 1.8202)^2} + \frac{0.97355}{(2.77405 - 1.8202)^2}} =$$

$$= -0.019,$$

$$C_{21}=\frac{\frac{\alpha_2}{\beta_2-r_1}}{\frac{\alpha_1}{(\beta_1-r_1)^2}+\frac{\alpha_2}{(\beta_2-r_1)^2}}=\frac{\frac{0.97355}{2.77405-0.493174}}{\frac{0.02645}{(0.539321-0.493174)^2}+\frac{0.97355}{(2.77405-0.493174)^2}}=$$

$$=0.03385,$$

$$C_{22}=\frac{\frac{\alpha_2}{\beta_2-r_2}}{\frac{\alpha_1}{(\beta_1-r_2)^2}+\frac{\alpha_2}{(\beta_2-r_2)^2}}=\frac{\frac{0.97355}{2.77405-1.8202}}{\frac{0.02645}{(0.539321-1.8202)^2}+\frac{0.97355}{(2.77405-1.8202)^2}}=$$

$$=0.9397.$$

Επίσης οι τιμές των C_1, C_2, θα προκύπτουν από τις παρακάτω σχέσεις :

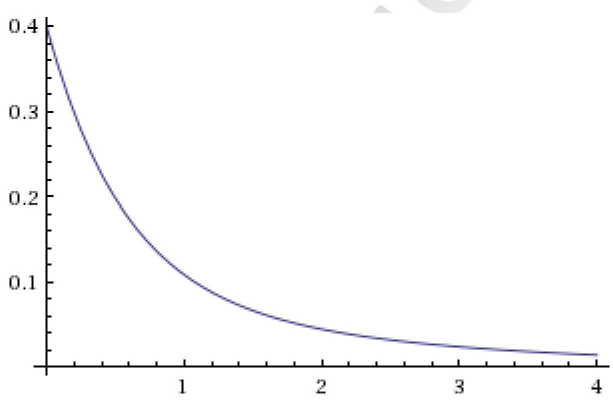
$$C_1=\frac{C_{11}}{\beta_1}+\frac{C_{21}}{\beta_2}=\frac{0.04546}{0.539321}+\frac{0.03385}{2.77405}=0.09649,$$

$$C_2=\frac{C_{12}}{\beta_1}+\frac{C_{22}}{\beta_2}=-\frac{0.019}{0.539321}+\frac{0.9397}{2.77405}=0.3035.$$

Τότε για την πιθανότητα χρεοκοπίας θα ισχύει ότι :

$$\psi(u)=C_1e^{-r_1u}+C_2e^{-r_2u}=0.09649e^{-0.493174u}+0.3035e^{-1.8202u}, u\geq 0.$$

Στο σχήμα 5.5.1, βλέπουμε την γραφική παράσταση της πιθανότητας χρεοκοπίας.



Σχήμα 5.7: Διάγραμμα της $\psi(u)$.

5.6: ΣΥΓΚΡΙΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΣΥΝΑΡΤΗΣΕΩΝ ΠΥΚΝΟΤΗΤΑΣ ΚΑΙ ΚΑΤΑΝΟΜΗΣ, ΧΡΕΟΚΟΠΙΑΣ ΚΑΙ ΜΗ ΧΡΕΟΚΟΠΙΑΣ ΤΩΝ ΤΡΙΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ.

Κατά την διάρκεια του κεφαλαίου 3 προσεγγίσαμε τις συναρτήσεις πυκνότητας πιθανότητας των κατανομών Pareto, Weibull και Lognormal με παραμέτρους:

$$\alpha = 6, \beta = 2, \quad \lambda = 0.63945, \alpha = 0.9 \text{ και } \mu = -1.41629, \sigma = 1,$$

αντίστοιχα από μείξεις δύο εκθετικών κατανομών και φροντίσαμε έτσι ώστε οι τρεις πυκνότητες να έχουν την ίδια μέση τιμή, συνεπώς και οι τρεις μείξεις εκθετικών κατανομών, οι οποίες αποτελούν τις προσεγγίσεις τους θα έχουν την ίδια μέση τιμή. Επίσης οι ροπές δεύτερης τάξης είναι σχετικά «κοντά» μεταξύ τους. Τότε σύμφωνα με την μέθοδο των ροπών και οι γραφικές παραστάσεις των συναρτήσεων πυκνότητας πιθανότητας των προσεγγίσεων, όπως και των συναρτήσεων κατανομής θα πρέπει να είναι παρόμοιες.

Η συνάρτηση πυκνότητας πιθανότητας της προσέγγισης της Pareto, είδαμε ότι είναι της μορφής:

$$f_X(x) = 0.1e^{-x} + 2.7e^{-3x}, x \geq 0,$$

Ενώ η συνάρτηση κατανομής της είναι:

$$F_X(x) = 1 - 0.1e^{-x} - 0.9e^{-3x}, x \geq 0.$$

Αντίστοιχα η συνάρτηση πυκνότητας πιθανότητας της προσέγγισης της Weibull γράφεται ως:

$$f_Y(y) = 1.613093e^{-4.97489y} + 1.3638e^{-2.01824y}, y \geq 0,$$

Και η συνάρτηση κατανομής της ως:

$$F_Y(y) = 1 - 0.324247e^{-4.97489y} - 0.675753e^{-2.01824y}, y \geq 0.$$

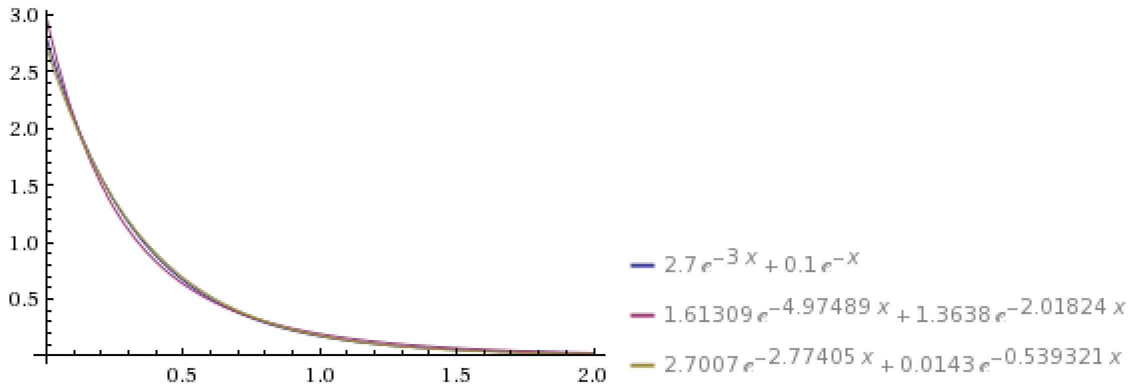
Επίσης η προσέγγιση της Lognormal είδαμε ότι είναι:

$$f_Z(z) = 0.0143e^{-0.539321z} + 2.7007e^{-2.77405z}, z \geq 0,$$

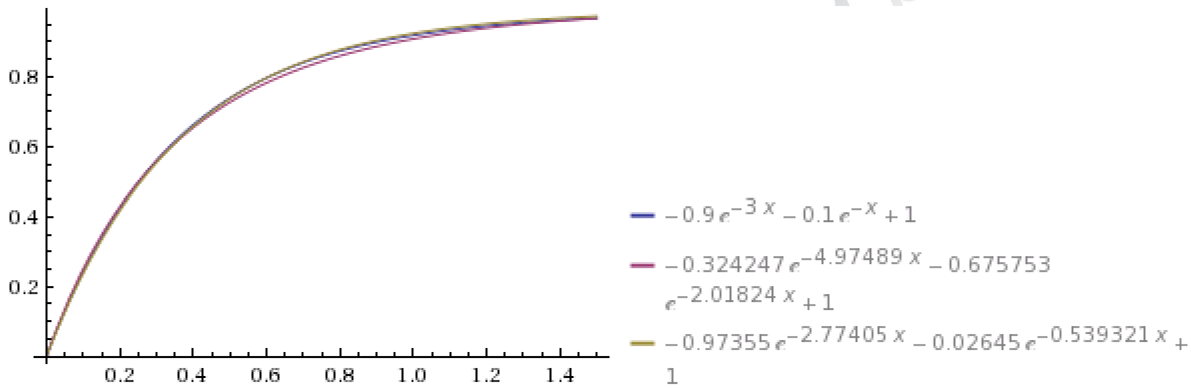
Και η $F_Z(z)$ γράφεται ως:

$$F_Z(z) = 1 - 0.97355e^{-2.77405z} - 0.02645e^{-0.539321z}, z \geq 0.$$

Στα παρακάτω δύο γραφήματα παρατηρούμε τις γραφικές παραστάσεις των τριών πυκνοτήτων και των τριών συναρτήσεων κατανομής αντίστοιχα και διαπιστώνουμε και στις 2 περιπτώσεις ότι είναι αρκετά παρόμοιες.



Σχήμα 5.8: Διάγραμμα των $f_X(x)$ (μπλε χρώμα), $f_Y(x)$ (ροζ χρώμα), $f_Z(x)$ (κίτρινο χρώμα).



Σχήμα 5.9: Διάγραμμα των $F_X(x)$ (μπλε χρώμα), $F_Y(x)$ (ροζ χρώμα), $F_Z(x)$ (κίτρινο χρώμα).

Σύμφωνα με το ολοκλήρωμα της απόλυτης διαφοράς, οι «αποστάσεις» των 3 πυκνοτήτων θα είναι:

$$\int_0^\infty |f_X(x) - f_Y(x)| dx \approx 0.0537,$$

$$\int_0^\infty |f_X(x) - f_Z(x)| dx \approx 0.0283,$$

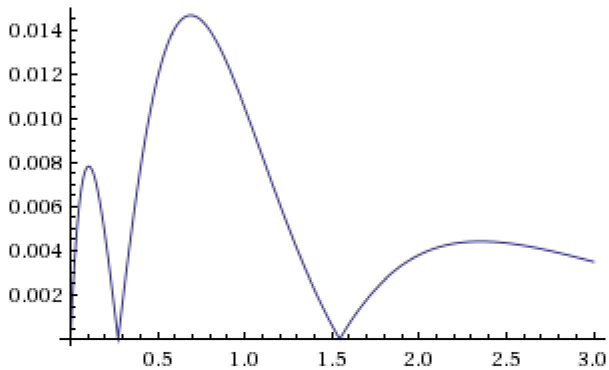
$$\int_0^\infty |f_Y(x) - f_Z(x)| dx \approx 0.0705.$$

Διαπιστώνουμε τότε ότι οι πυκνότητες των προσεγγίσεων Pareto και Lognormal είναι πιο «κοντά» μεταξύ τους απ΄ότι με την κατανομή Weibull.

Επίσης θα υπολογίσουμε την «μέγιστη διαφορά» των συναρτήσεων κατανομής των τριών προσεγγίσεων σύμφωνα με την μέθοδο 3.2.1 και θα δούμε για κάθε μία περίπτωση και το γράφημα της απόλυτης διαφοράς τους. Θα ισχύει τότε ότι:

$$sup_{x \geq 0} |F_X(x) - F_Y(x)| \approx 0.0146,$$

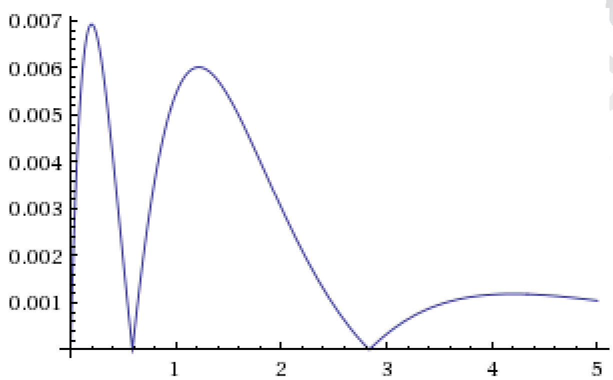
στο σημείο $x_0 \approx 0.689506$. Επίσης παρουσιάζονται άλλα δύο τοπικά μέγιστα, τα οποία είναι ίσα με 0.0078 και 0.0044 στα σημεία $x_1 \approx 0.1046, x_2 \approx 2.352$ αντίστοιχα.



Σχήμα 5.10: Διάγραμμα του $|F_X(x) - F_Y(x)|$.

$sup_{x\geq 0}|F_X(x) - F_Z(x)| \approx 0.0069,$

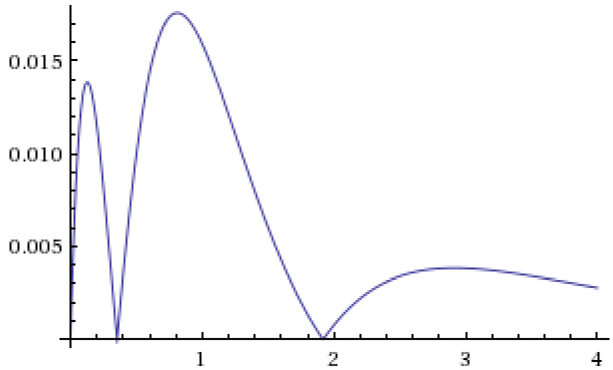
στο σημείο $x_0 \approx 0.2008$. Οι άλλες δύο μέγιστες «αποστάσεις» που παρατηρούναι στην απόλυτη διαφορά των δύο συναρτήσεων κατανομής είναι ίσες με $0.006, 0.0012$ στα σημεία $x_0 \approx 1.217$, και $x_2 \approx 4.2006$.



Σχήμα 5.11: Διάγραμμα του $|F_X(x) - F_Z(x)|$.

$sup_{x\geq 0}|F_Y(x) - F_Z(x)| \approx 0.0175,$

στο σημείο $x_0 \approx 0.8108$. Τα άλλα δύο τοπικά μέγιστα που αποτυπώνονται και στο παρακάτω γράφημα είναι ίσα με $0.01383, 0.0039$ στα σημεία $x_1 \approx 0.1278, x_2 \approx 2.917$.



Σχήμα 5.12: Διάγραμμα του $|F_Y(x) - F_Z(x)|$

Η πιθανότητα χρεοκοπίας της προσέγγισης της Pareto είναι η:

$$\psi_X(x) = \frac{1}{5}e^{-0.82918x} + \frac{1}{5}e^{-2.17082x}, x \geq 0.$$

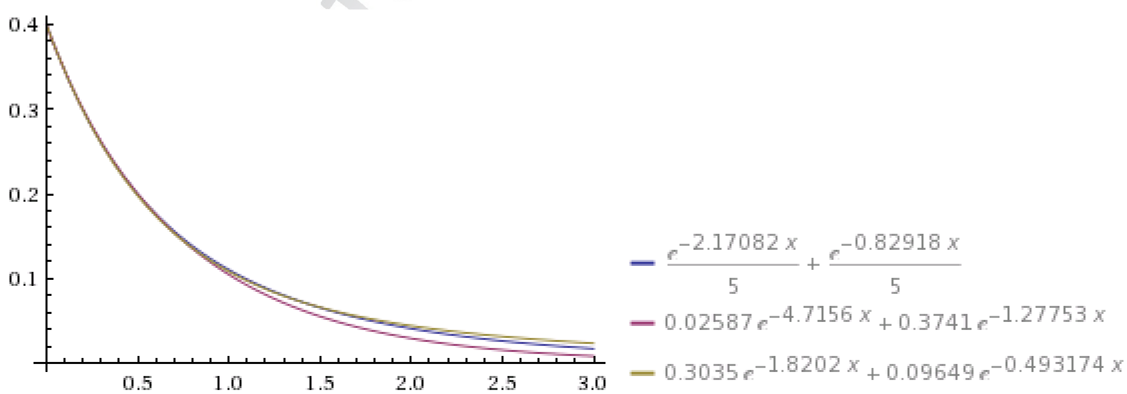
Της προσέγγισης της Weibull γράφεται στην μορφή:

$$\psi_Y(y) = 0.3741e^{-1.27753y} + 0.02587e^{-4.7156y}, y \geq 0.$$

Ενώ της προσέγγισης της Lognormal είναι:

$$\psi_Z(z) = 0.09649e^{-0.493174z} + 0.3035e^{-1.8202z}, z \geq 0.$$

Επίσης και οι γραφικές παραστάσεις των τριών πιθανοτήτων χρεοκοπίας θα πρέπει να είναι παρόποιες και να ισχύει $\psi(0) = \frac{1}{1+\theta} = 0.4$, διότι οι τρεις κατανομές των αποζημιώσεων, είναι οι συναρτήσεις πυκνότητας πιθανότητας των προσεγγίσεων που αναλύσαμε παραπάνω και έχουν την ίδια μέση τιμή, επομένως και το ίδιο περιθώριο ασφαλείας. Βλέπουμε τότε στο παρακάτω γράφημα την απεικόνιση των γραφικών παραστάσεων των τριών πιθανοτήτων χρεοκοπίας.



Σχήμα 5.13: Διάγραμμα των $\psi_X(x)$ (μπλε χρώμα), $\psi_Y(x)$ (ροζ χρώμα), $\psi_Z(x)$ (κίτρινο χρώμα).

Σύμφωνα με το ολοκλήρωμα της απόλυτης διαφοράς, οι «αποστάσεις» των τριών παραπάνω συναρτήσεων θα είναι:

$$\int_0^{\infty} |\psi_X(x) - \psi_Y(x)| dx \approx 0.035,$$

$$\int_0^{\infty} |\psi_X(x) - \psi_Z(x)| dx \approx 0.033,$$

$$\int_0^{\infty} |\psi_Y(x) - \psi_Z(x)| dx \approx 0.066.$$

Αντίστοιχα οι προσεγγίσεις των πιθανοτήτων μη χρεοκοπίας θα είναι ίσες με:

$$\delta_X(x) = 1 - \frac{1}{5} e^{-0.82918x} - \frac{1}{5} e^{-2.17082x}, x \geq 0,$$

της Pareto.

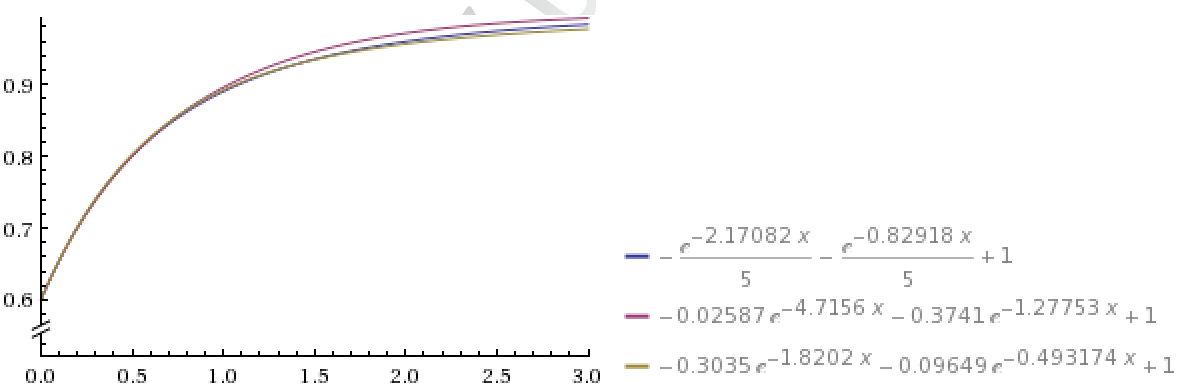
Της προσέγγισης της Weibull γράφεται στην μορφή:

$$\delta_Y(y) = 1 - 0.3741 e^{-1.27753y} - 0.02587 e^{-4.7156y}, y \geq 0.$$

Ενώ της προσέγγισης της Lognormal είναι:

$$\delta_Z(z) = 1 - 0.09649 e^{-0.493174z} - 0.3035 e^{-1.8202z}, z \geq 0.$$

Το γράφημά τους είναι το παρακάτω:



Σχήμα 5.14: Διάγραμμα των $\delta_X(x)$ (μπλε χρώμα), $\delta_Y(x)$ (ροζ χρώμα), $\delta_Z(x)$ (κίτρινο χρώμα).

Γνωρίζουμε ότι κάθε πιθανότητα μη χρεοκοπίας είναι αύξουσα συνάρτηση, μπορούμε επομένως να δημιουργήσουμε έναν πίνακα άνω ποσοστημορίων. Δεν έχει όμως νόημα να εξετάσουμε και το ποσοστό σφάλματος όπως στο κεφάλαιο 3, διότι δεν έχουμε να συγκρίνουμε συναρτήσεις πυκνότητας πιθανότητας με τις προσεγγίσεις τους, αλλά τρεις προσεγγίσεις πιθανότητας μη χρεοκοπίας μεταξύ τους. Θα έχουμε τότε ότι:

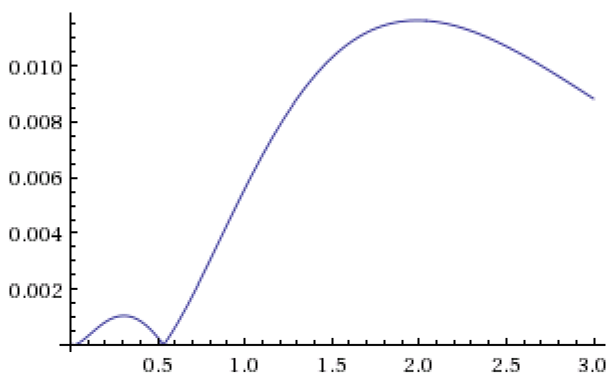
a	$\delta_X(x)$	$\delta_Y(y)$	$\delta_Z(z)$	x	y	z	$\lvert x-y\rvert$	$\lvert x-z\rvert$	$\lvert y-z\rvert$
0.01	0.99	0.99	0.99	3.62	2.83	4.61	0.79	0.99	1.78
0.02	0.98	0.98	0.98	2.80	2.29	3.27	0.51	0.47	0.98
0.03	0.97	0.97	0.97	2.34	1.97	2.57	0.37	0.23	0.60
0.04	0.96	0.96	0.96	2.02	1.75	2.13	0.27	0.11	0.38

Πίνακας 5.1: Πίνακας άνω ποσοστημορίων των τριών προσεγγίσεων πιθανότητας μη χρεοκοπίας.

Όπως και προηγουμένως με τις συναρτήσεις κατανομής των τριών προσεγγίσεων, έτσι και τώρα θα υπολογίσουμε την μέγιστη διαφορά των τριών πιθανοτήτων χρεοκοπίας και μη χρεοκοπίας και για κάθε περίπτωση θα παρουσιάσουμε και το γράφημα της απόλυτης διαφοράς τους. Θα έχουμε τότε ότι:

$$sup_{x\geq 0}|\delta_X(x)-\delta_Y(x)| = sup_{x\geq 0}|\psi_X(x)-\psi_Y(x)| \approx 0.01163,$$

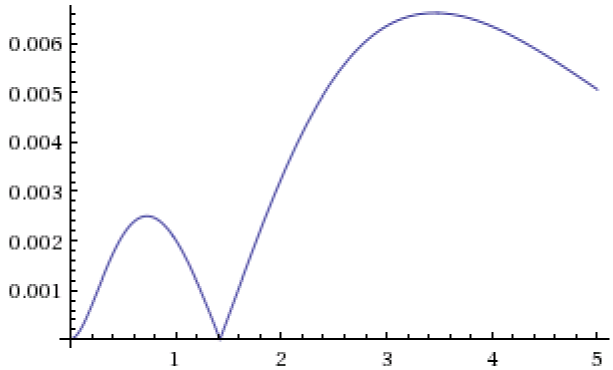
στο σημείο $x_0 \approx 1.98915$.



Σχήμα 5.15: Γραφική παράσταση των $|\delta_X(x)-\delta_Y(x)| = |\psi_X(x)-\psi_Y(x)|$.

$$sup_{x\geq 0}|\delta_X(x)-\delta_Z(x)| = sup_{x\geq 0}|\psi_X(x)-\psi_Z(x)| 0.0066,$$

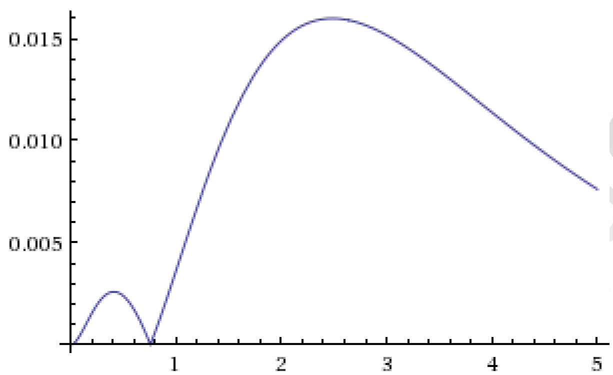
στο σημείο $x_0 \approx 3.4557$.



Σχήμα 5.16: Γραφική παράσταση των $|\delta_X(x) - \delta_Z(x)| = |\psi_X(x) - \psi_Z(x)|$.

$sup_{x\geq 0}|\delta_Y(x) - \delta_Z(x)| = sup_{x\geq 0}|\delta_Y(x) - \delta_Z(x)| \approx 0.0159$,

στο σημείο $x_0 \approx 2.486$.



Σχήμα 5.17: Γραφική παράσταση των $|\delta_Y(x) - \delta_Z(x)| =.|\psi_Y(x) - \psi_Z(x)|$.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΕΛΛΗΝΙΚΗ

1. Κούτρας, Μ. (2005), « Εισαγωγή στις Πιθανότητες: Θεωρία και Εφαρμογές, Β΄ Τόμος». Εκδόσεις Σταμούλη.

2. Πολίτης, Κ. (2012), «Θεωρία Συλλογικού Κινδύνου: Το Συλλογικό Πρότυπο και Θεωρία Χρεοκοπίας». Εκδόσεις Σταμούλη.

3. Σημειώσεις Κ.Πολίτη (2015), μάθημα «Θεωρία Χρεοκοπίας», Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιά.

ΞΕΝΟΓΛΩΣΣΗ

1. Babier, J. & Chan, B., (1992), « Approximations of Ruin Probability by Di-Atomic or Di-Exponential Claims». Astin Bulletin 22, 235-246.

2. Bean, M.A., (2001), «The Science of Uncertainty: With Applications to Investments, Insurance and Engineering». Ed.Brooks Cole.

3. Dickson, D.C.M, (2005) «Insurance Risk and Ruin (International Series on Actuarial Science)». Cambridge University Press.

4. Gerber, H., Goovaerts,M.J., Kaas,R., (1987), «On the Probability and Severity of Ruin». Astin Bulletin 17, 151-163.