

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΤΗ ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΩΝ

**Μοντελοποίηση ασφαλιστικών απαιτήσεων: κατασκευή κατανομών
με βαριές ουρές μέσω μετασχηματισμού της κατανομής Weibull**

Λογοθέτης Ιωάννης

Διπλωματική Εργασία

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

Πειραιάς

Ιούνιος 2023

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

Τα μέλη της Επιτροπής ήταν:

- Καθηγητής Μ. Κούτρας (Επιβλέπων)
- Αναπλ. Καθηγητής Ε. Χατζηκωνσταντινίδης
- Αναπλ. Καθηγητής Γ. Ψαράκος

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

UNIVERSITY OF PIRAEUS

SCHOOL OF FINANCE AND STATISTICS



DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

POSTGRADUATE PROGRAM IN ACTUARIAL SCIENCE AND
RISK MANAGEMENT

**Modeling Insurance claims: heavy tailed distributions generation by
transforming the Weibull distribution**

Logothetis Ioannis

MSc Dissertation

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

Piraeus

June 2023

*Στους γονείς μου και τα αδέρφια μου,
Νεκτάριο, Ματίνα, Μιχαήλ, Παναγιώτη*

ΕΥΧΑΡΙΣΤΙΕΣ

Με αφορμή την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κύριο Κούτρα Μάρκο για την καθοδήγησή του για την διεκπεραίωση της διπλωματικής και για το ενδιαφέρον του σε όλη την πορεία μου στο μεταπτυχιακό πρόγραμμα.

Τέλος, οφείλω το μεγαλύτερο ευχαριστώ στην οικογένειά μου για την στήριξή τους σε όλη τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Είναι γνωστό ότι τα ύψη των ασφαλιστικών ζημιών ακολουθούν κατανομές με βαριές ουρές, ιδιότητα την οποία δεν κατέχουν οι κλασικές κατανομές (κανονική, εκθετική κτλ). Για το λόγο αυτό, έχει δοθεί ιδιαίτερη σημασία στην δημιουργία νέων οικογενειών κατανομών που μπορούν να προσεγγίσουν ικανοποιητικά δεδομένα με βαριές ουρές καθώς και στη μελέτη των ιδιοτήτων τους.

Στην παρούσα διπλωματική εργασία, αναφέρουμε αρχικά, κάποιες βασικές ιδιότητες των κατανομών με βαριά ουρά. Στη συνέχεια γίνεται μια παρουσίαση της κατανομής Weibull, η οποία ανήκει στην οικογένεια κατανομών με βαριές ουρές.

Στη συνέχεια, παρουσιάζεται ένας μετασχηματισμός της κατανομής Weibull ώστε να προσαρμόζεται καλύτερα σε δεδομένα με βαριά ουρά, όπως είναι τα ύψη των ασφαλιστικών ζημιών και γίνεται σύγκριση των μέτρων κινδύνου Αξία σε Κίνδυνο (VaR) και Αναμενόμενο Έλλειμμα ($TVaR/Expected Shortfall$) ανάμεσα στη κλασική κατανομή Weibull και το μετασχηματισμό της.

Τέλος, εξετάζεται η αποτελεσματικότητα των τεχνικών αυτών σε προσομοιωμένα δεδομένα και γίνεται εφαρμογή σε πραγματικά δεδομένα.

ABSTRACT

It is well known that the severity of insurance claims follows heavy-tailed distributions, a property that classical distributions (normal, exponential, etc.) do not possess. For this reason, particular attention has been paid to the creation of new families of distributions that can satisfactorily approximate heavy-tailed data and to the study of their properties.

In this thesis, we first report some basic properties of heavy-tailed distributions. Then, we present the Weibull distribution, which belongs to the heavy-tailed distribution family.

Next, we present a transformation of the Weibull distribution which provide a better fit to heavy-tailed data, such as insurance claims, and compare the risk measures Value at Risk (VaR) and Expected Shortfall ($TVaR$ /Expected Shortfall) between the standard Weibull distribution and its transformation.

Finally, the efficiency of these techniques is investigated on simulated data and then an application to real insurance loss data is conducted.

Περιεχόμενα

ΠΕΡΙΛΗΨΗ	XI
ABSTRACT	XIII
ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ	XVII
ΚΑΤΑΛΟΓΟΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ	1
ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΚΑΤΑΝΟΜΕΣ ΜΕ ΒΑΡΙΕΣ ΟΥΡΕΣ	3
1.1 ΕΙΣΑΓΩΓΗ	3
1.2 ΚΑΤΑΝΟΜΕΣ ΜΕ ΒΑΡΙΕΣ ΟΥΡΕΣ	3
1.3 Η ΚΑΤΑΝΟΜΗ WEIBULL	7
ΚΕΦΑΛΑΙΟ 2 Η ΚΑΤΑΝΟΜΗ WEIBULL ΜΕ ΒΑΡΙΑ ΟΥΡΑ	11
2.1 ΕΙΣΑΓΩΓΗ	11
2.2 ΔΗΜΙΟΥΡΓΙΑ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΜΟΝΤΕΛΟΥ	11
2.3 Η ΚΑΤΑΝΟΜΗ NHTD-WEIBULL	15
ΚΕΦΑΛΑΙΟ 3 ΙΔΙΟΤΗΤΕΣ ΤΗΣ ΚΑΤΑΝΟΜΗΣ NFHT ΚΑΙ ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ	21
3.1 ΕΙΣΑΓΩΓΗ	21
3.2 ΜΑΘΗΜΑΤΙΚΕΣ ΙΔΙΟΤΗΤΕΣ	21
3.3 ΕΚΤΙΜΗΤΕΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ	23
3.4 ΠΡΟΣΟΜΟΙΩΣΗ	24
3.5 ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ	32
3.5.1 Αξία σε κίνδυνο	32
3.5.2 Αναμενόμενο Έλλειμα	37
ΚΕΦΑΛΑΙΟ 4 ΕΦΑΡΜΟΓΗ ΔΕΔΟΜΕΝΩΝ	43
4.1 ΕΙΣΑΓΩΓΗ	43
4.2 ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ	43
4.3 ΠΡΟΣΑΡΜΟΓΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΜΕΤΡΩΝ ΚΙΝΔΥΝΟΥ	47
ΣΥΝΟΨΗ.....	51
ΠΑΡΑΡΤΗΜΑ	52
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	70

ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1.1: Σύγκριση συναρτήσεων πυκνότητας εκθετικής και HTD.....	3
Σχήμα 1.2: Τύποι ασυμμετρίας.....	6
Σχήμα 1.3: Διαγράμματα Quantile για την κατανομή Weibull με παραμέτρους $\alpha=0.9, \gamma=1.3$ και $\alpha=1.8, \gamma=1.3$	9
Σχήμα 1.4: Διαγράμματα Quantile για την κατανομή Weibull με παραμέτρους $\alpha=2.7, \gamma=1.3$ και $\alpha=3.5, \gamma=1.3$	9
Σχήμα 2.1: Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=1.2, \gamma=1$ και διάφορες τιμές του σ (αριστερά). Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=1.5, \gamma=1$ και διάφορες τιμές του σ (δεξιά).....	17
Σχήμα 2.2: Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=2, \gamma=1$ και διάφορες τιμές του σ (αριστερά). Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=2.5, \gamma=1$ και διάφορες τιμές του σ (δεξιά).....	17
Σχήμα 2.3: Σχηματική απεικόνιση της συνάρτησης κατανομής της NHTD-Weibull με παραμέτρους $\alpha=1.5, \gamma=1$ και $\sigma=2.5$ (αριστερά). Σχηματική απεικόνιση της συνάρτησης επιβίωσης της NHTD-Weibull με παραμέτρους $\alpha=1.5, \gamma=1$ και $\sigma=2.5$ (δεξιά).....	18
Σχήμα 2.4: Σχηματική απεικόνιση συνάρτησης κινδύνου hrf.....	18
Σχήμα 2.5: Ιστόγραμμα συχνοτήτων για προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.....	19
Σχήμα 2.6: Θηκόγραμμα για τα προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.	20
Σχήμα 3.1: Εκτιμήσεις των παραμέτρων για δείγματα μεγέθους n για $\alpha=0.9, \sigma=1.2, \gamma=0.5$.	27
Σχήμα 3.2: Εκτιμήσεις MSEs για δείγματα μεγέθους n για $\alpha=0.9, \sigma=1.2, \gamma=0.5$	27
Σχήμα 3.3: Εκτιμήσεις απόλυτης μεροληψίας για δείγματα μεγέθους n για $\alpha=0.9, \sigma=1.2, \gamma=0.5$	28
Σχήμα 3.4: Εκτιμήσεις μεροληψίας για δείγματα μεγέθους n για $\alpha=0.9, \sigma=1.2, \gamma=0.5$	28
Σχήμα 3.5: Εκτιμήσεις των παραμέτρων για δείγματα μεγέθους n για $\alpha=1.3, \sigma=1.7, \gamma=1$	30
Σχήμα 3.6: Εκτιμήσεις MSEs για δείγματα μεγέθους n για $\alpha=1.3, \sigma=1.7, \gamma=1$	30
Σχήμα 3.7: Εκτιμήσεις απόλυτης μεροληψίας για δείγματα μεγέθους n για $\alpha=1.3, \sigma=1.7, \gamma=1$	31

Σχήμα 3.8: Εκτιμήσεις μεροληψίας για δείγματα μεγέθους n για $\alpha=1.3$, $\sigma=1.7$, $\gamma=1$	31
Σχήμα 3.9: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους σ	34
Σχήμα 3.10: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους γ	35
Σχήμα 3.11: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους α	35
Σχήμα 3.12: Σύγκριση VaR των κατανομών NHTW - Weibull.	36
Σχήμα 3.13: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW	39
Σχήμα 3.14: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW για διαφορετικές παραμέτρους σ	40
Σχήμα 3.15: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW για διαφορετικές παραμέτρους α	40
Σχήμα 3.16: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW για διαφορετικές παραμέτρους γ	41
Σχήμα 3.17: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW και Weibull.....	41
Σχήμα 4.1: Ιστόγραμμα συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.	44
Σχήμα 4.2: Θηκόγραμμα συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.	44
Σχήμα 4.3: QQ-plot συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.	45
Σχήμα 4.4: Διαγραμματική απεικόνιση των εμπειρικών συναρτήσεων πυκνότητας για τη συνολικές ζημιές ανά εβδομάδα και ανά ημέρα.....	46
Σχήμα 4.5: Σχηματική απεικόνιση ιστογράμματος των συνολικών ζημιών ανά εβδομάδα με εφαρμογή της εμπειρικής συνάρτησης πυκνότητας (αριστερά) και της εμπειρικής συνάρτησης κατανομής (δεξιά).	46
Σχήμα 4.6: Σχηματική απεικόνιση ιστογράμματος των συνολικών ζημιών ανά ημέρα με εφαρμογή της εμπειρικής συνάρτησης πιθανότητας.	47
Σχήμα 4.7: Σχηματική απεικόνιση VaR για τις κατανομές NHTW και Weibull.	49
Σχήμα 4.8: Σχηματική απεικόνιση $TVaR$ για τις κατανομές NHTW και Weibull.	50

ΚΑΤΑΛΟΓΟΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ

1. Κατανομές με βαριές ουρές (heavy tailed distributions)	HTD
2. Υπό συνθήκη μέση υπέρβαση (Conditional Mean Exceedance)	CME
3. New Family of Heavy Tailed Distributions	NFHT
4. Συνάρτηση επιβίωσης (Survival function)	sf
5. Συνάρτηση ρυθμού κινδύνου (Hazard rate function)	hrf
6. Νέα κατανομή Weibull με βαριά ουρά (New Heavy Tailed Weibull Distribution)	NHTW
7. Εκτιμητές ελαχίστων τετραγώνων	ε.ε.τ.
8. Μέσο τετραγωνικό σφάλμα	μ.τ.σ.

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΚΑΤΑΝΟΜΕΣ ΜΕ ΒΑΡΙΕΣ ΟΥΡΕΣ

1.1 ΕΙΣΑΓΩΓΗ

Σε αυτό το κεφάλαιο θα γίνει αναφορά στις κατανομές με βαριά ουρά και σε κάποιες ιδιότητες αυτών των κατανομών καθώς και στην καταλληλότητά τους για την περιγραφή δεδομένων με βαριά ουρά όπως είναι οι ασφαλιστικές ζημιές.

1.2 ΚΑΤΑΝΟΜΕΣ ΜΕ ΒΑΡΙΕΣ ΟΥΡΕΣ

Στη θεωρία πιθανοτήτων, οι κατανομές με βαριά ουρά (heavy tailed distributions, HTD) είναι κατανομές πιθανότητας των οποίων η ουρά συγκλίνει με πιο αργό ρυθμό στο μηδέν συγκριτικά με την εκθετική κατανομή (Wikipedia, Heavy tailed distributions, βλ. Σχήμα 1.1).



Σχήμα 1.1: Σύγκριση συναρτήσεων πυκνότητας εκθετικής και HTD

Υπάρχουν 3 σημαντικές υποκατηγορίες HTD

- Κατανομές με παχιά ουρά,
- Κατανομές με μακριά ουρά,
- Υποεκθετικές κατανομές.

Στην πράξη, όλες οι κοινώς χρησιμοποιούμενες HTD ανήκουν στην υποεκθετική κατηγορία.

Η κατανομή μια τυχαίας μεταβλητής X , με συνάρτηση κατανομής F , καλείται ότι έχει βαριά (δεξιά) ουρά όταν η ροπογεννήτρια συνάρτηση της X , $M_X(\lambda)$ συγκλίνει στο άπειρο για κάθε $\lambda > 0$.

Αυτό σημαίνει,

$$M_X(\lambda) = \int_{-\infty}^{+\infty} e^{\lambda x} dF(x) \rightarrow \infty \text{ για όλα τα } \lambda > 0.$$

και ότι η συνάρτηση δεξιάς ουράς (ή όπως ονομάζεται εναλλακτικά, συνάρτηση επιβίωσης)

$$\bar{F}(x) = Pr(X > x) = 1 - F(x)$$

ικανοποιεί τη συνθήκη

$$\lim_{x \rightarrow +\infty} e^{\lambda x} \bar{F}(x) = \infty \text{ για κάθε } \lambda > 0.$$

Ένας άλλος τρόπος με τον οποίο μπορούμε να διακρίνουμε μια κατανομή με βαριά ουρά είναι μέσω της Αναμενόμενης Μέσης Υπερβάλλουσας Ζημιάς (Mean Excess Loss, MEL). Η MEL εκφράζει τη μέση απόκλιση μιας τυχαίας μεταβλητής X από μια τιμή x , δεδομένου ότι αυτή η απόκλιση είναι θετική δηλαδή

$$MEL(x) = E(X - x | X \geq x).$$

Ισοδύναμα η MEL μπορεί να γραφεί με την ακόλουθη μορφή:

$$MEL = \frac{1}{\bar{F}(x)} \int_x^{+\infty} \bar{F}(t) dt.$$

Όταν μια κατανομή έχει βαριά ουρά θα υπάρχουν πολλές μεγάλες αποκλίσεις πάνω από ένα συγκεκριμένο μέγεθος. Έτσι, η μέση απόκλιση (ο όρος «μέση» αναφέρεται τόσο στο πλήθος όσο και στο μέγεθος των αποκλίσεων) τείνει να αυξάνεται όσο προχωράμε προς την ουρά της κατανομής. Συνεπώς, η ύπαρξη βαριάς ουράς είναι άμεσα συνδεδεμένη με τη μονοτονία της $MEL(x)$ καθώς είναι αύξουσα συνάρτηση του x , εφόσον η τιμή του x είναι αρκετά μεγάλη.

Κάποιες γνωστές κατανομές με βαριά ουρά είναι οι ακόλουθες:

- Η κατανομή Pareto με σ.π.π.

$$f(x; \alpha, \theta) = \frac{\alpha \theta^\alpha}{x^{\alpha+1}}$$

και συνάρτηση δεξιάς ουράς

$$\bar{F}(x) = \left(\frac{\theta}{x}\right)^\alpha.$$

- Η Log-normal κατανομή με

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$$

και συνάρτηση δεξιάς ουράς

$$\bar{F}(x) = 1 - \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{(\ln x - \mu)^2}{\sigma \sqrt{2}} \right) \right],$$

όπου erf είναι η συνάρτηση σφάλματος (γνωστή και ως συνάρτηση σφάλματος Gauss) και ορίζεται ως:

$$\operatorname{erf} z = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

- Η κατανομή Levy με

$$f(x) = \sqrt{\frac{c}{2\pi}} \frac{e^{-\frac{c}{2(x-\mu)}}}{(x-\mu)^{3/2}}$$

και συνάρτηση δεξιάς ουράς

$$\bar{F}(x) = 1 - \operatorname{erfc} \left(\sqrt{\frac{c}{2(x-\mu)}} \right).$$

- Η κατανομή Weibull με παράμετρο σχήματος k μεγαλύτερη από 0 και μικρότερη από 1 η οποία έχει συνάρτηση πυκνότητας

$$f(x; \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}$$

και συνάρτηση δεξιάς ουράς

$$\bar{F}(x; \lambda, k) = e^{-(x/\lambda)^k}.$$

- Η κατανομή Burr με

$$f(x) = ck \frac{x^{c-1}}{(1+x^c)^{k+1}}$$

και συνάρτηση δεξιάς ουράς

$$\bar{F}(x) = (1+x^c)^{-k}.$$

Άλλο ένα σημαντικό χαρακτηριστικό των HTD είναι η ιδιότητα της ομαλής κύμανσης. Μια κατανομή καλείται ομαλής κύμανσης όταν ικανοποιεί τη συνθήκη

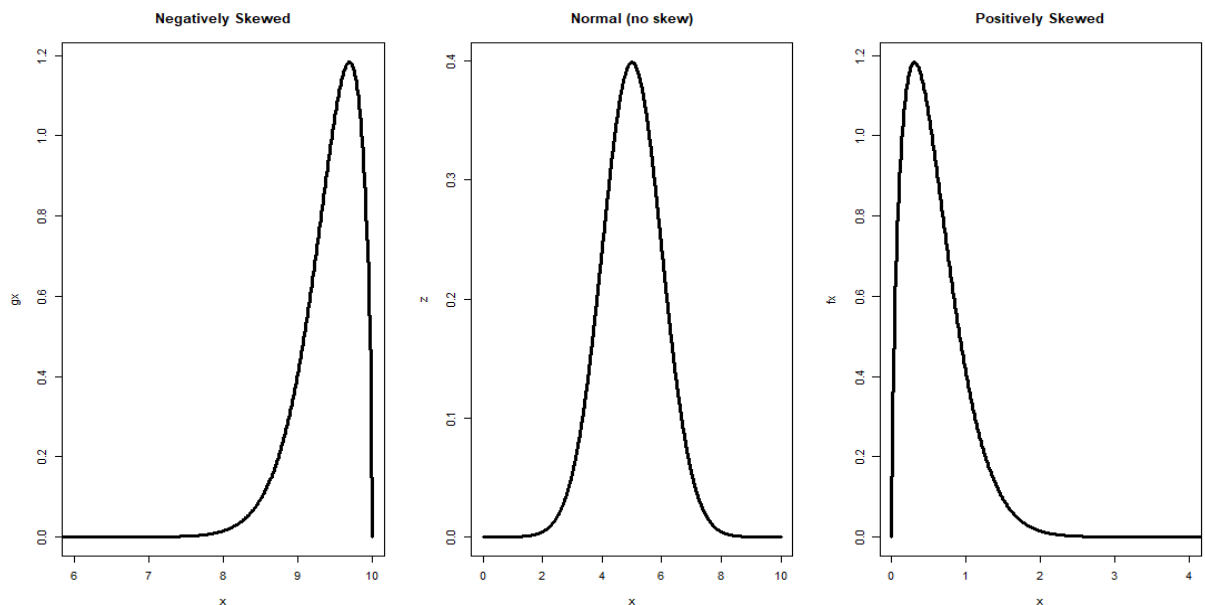
$$\lim_{x \rightarrow \infty} \frac{1 - G(tx)}{1 - G(x)} = t^{-\alpha}, \quad \alpha \in (0, +\infty),$$

όπου το a καλείται δείκτης ομαλής κύμανσης και $G(x)$ είναι η συνάρτηση κατανομής της τυχαίας μεταβλητής X .

Οι κατανομές που έχουν την ιδιότητα της ομαλής κύμανσης είναι αρκετά ανταγωνιστικές για τη μοντελοποίηση δεδομένων με βαριά ουρά, όπως είναι τα δεδομένα ασφαλιστικών ζημιών. Αυτό συμβαίνει διότι, τα δεδομένα ασφαλιστικών ζημιών έχουν θετική ασυμμετρία, υψηλή κορυφή και παχιά δεξιά ουρά.

Αποδεικνύεται ότι για μια κατανομή με βαριά ουρά ισχύει η συνθήκη της ομαλής κύμανσης και αντίστροφα, μια κατανομή για την οποία ισχύει η συνθήκη της ομαλής κύμανσης, θα έχει βαριά ουρά.

Στο Σχήμα 1.2 απεικονίζονται οι τύποι ασυμμετρίας. Κατανομές με θετική ασυμμετρία είναι κατανομές για τις οποίες η μέση τιμή είναι μεγαλύτερη από τη διάμεσο τους και για αυτό βρίσκεται στα δεξιά από τη διάμεσο. Αυτό σημαίνει πως η πλειοψηφία των δεδομένων βρίσκεται αριστερά από τη μέση τιμή αλλά υπάρχουν δεδομένα με αρκετά μεγαλύτερες τιμές δεξιότερα της. Αυτό εξηγεί και την καταλληλότητα αυτών των κατανομών για τη μοντελοποίηση δεδομένων με βαριά ουρά.



Σχήμα 1.2: Τύποι ασυμμετρίας

Γνωστές κατανομές πιθανότητας όπως είναι η κατανομή Log-normal, η Pareto, η Γάμμα και η Weibull είναι πολύ χρήσιμες στη μοντελοποίηση δεδομένων. Παρόλα αυτά, οι κλασικές κατανομές παρουσιάζουν κάποιες αδυναμίες στη μοντελοποίηση δεδομένων ασφαλιστικών

ζημιών καθώς, δεν είναι αρκετά ευέλικτες ώστε να προσαρμοστούν ικανοποιητικά στα σύνολα δεδομένων με βαριά ουρά, που προκύπτουν σε διάφορες περιοχές όπως είναι η ασφάλιση αλλά και η διαχείριση χρηματοοικονομικών κινδύνων. Για παράδειγμα, η Log-normal κατανομή, η Γάμμα και η Βήτα δεν έχουν έκφραση κλειστής μορφής για τη συνάρτηση κατανομής τους και έτσι ο υπολογισμός κάποιων χρήσιμων ποσοτήτων γίνεται εξαιρετικά δύσκολος. Άλλο ένα παράδειγμα είναι, η κατανομή Weibull η οποία ενώ δίνει καλά αποτελέσματα για δεδομένα με μικρές ζημιές, αποτυγχάνει σε δεδομένα με μεγάλες ζημιές.

Η μεγάλη σημασία της ερμηνείας τέτοιων δεδομένων όπως οι ασφαλιστικές ζημιές έχει οδηγήσει πολλούς ερευνητές και αναλογιστές στη αναζήτηση νέων μοντέλων κατανομών με βαριά ουρά, είτε μέσω μίξης δύο ή περισσότερων κατανομών είτε μέσω μετασχηματισμών γνωστών κατανομών, τα οποία είναι πιο ευέλικτα ως προς την προσαρμογή των δεδομένων και δίνουν καλύτερα αποτελέσματα.

Στις επόμενες ενότητες θα παρουσιαστεί μια οικογένεια κατανομών μέσω μετασχηματισμού της κατανομής Weibull όπως προτάθηκε από τους Ahmad, Mahmoudi and Dey (2020).

1.3 Η ΚΑΤΑΝΟΜΗ WEIBULL

Η κατανομή Weibull είναι μια συνεχής κατανομή πιθανότητας και πήρε το όνομα της από τον Σουηδό μαθηματικό Waloddi Weibull, ο οποίος την περιέγραψε λεπτομερώς το 1951.

Η συνάρτηση πυκνότητας της κατανομής έχει τη μορφή

$$f(x; \alpha, \gamma) = \alpha \gamma x^{\alpha-1} e^{-\gamma x^\alpha} \text{ για } x \geq 0$$

και

$$f(x; \alpha, \gamma) = 0 \quad \text{για } x < 0$$

όπου $k > 0$ είναι η παράμετρος σχήματος και $\lambda > 0$ είναι η παράμετρος κλίμακας.

Η συνάρτηση κατανομής της Weibull δίνεται από τον τύπο

$$F(x; \alpha, \gamma) = 1 - e^{-\gamma x^\alpha} \text{ για } x \geq 0$$

και

$$F(x; \alpha, \gamma) = 0 \quad \text{για } x < 0.$$

Η μέση τιμή και η διακύμανση της τ.μ. X που ακολουθεί την κατανομή Weibull δίνονται από τους τύπους

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

και

$$Var(X) = \left(\frac{1}{\gamma}\right)^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)^2 \right]$$

Η ροπογεννήτρια συνάρτηση της κατανομής Weibull η οποία είναι

$$M_x(t) = E(e^{tX}) = \int_0^{+\infty} e^{tx} \alpha \gamma x^{\alpha-1} e^{-\gamma x^\alpha} dx$$

δεν μπορεί να εκφρασθεί από έναν κλειστό τύπο, ωστόσο θα μπορούσε να υπολογισθεί μέσω της συνάρτησης Γάμμα από την ακόλουθη έκφραση

$$M_x(t) = E(e^{tX}) = \sum_{n=0}^{\infty} \frac{t^n}{\gamma^n n!} \Gamma\left(1 + \frac{n}{\alpha}\right)$$

Σημειώνοντας τέλος ότι η συνάρτηση ποσοστημορίου (Quantile function) προκύπτει λύνοντας την παρακάτω σχέση

$$Pr(X < x_p) = p$$

ως προς x_p , εύκολα διαπιστώνεται ότι για την κατανομή Weibull αυτή θα δίνεται από τον τύπο:

$$Q(p; \alpha, \gamma) = \frac{1}{\gamma} (-\ln(1-p))^{1/\alpha} \text{ για } 0 < p < 1.$$

Τέλος, η συνάρτηση κινδύνου (hazard rate function)

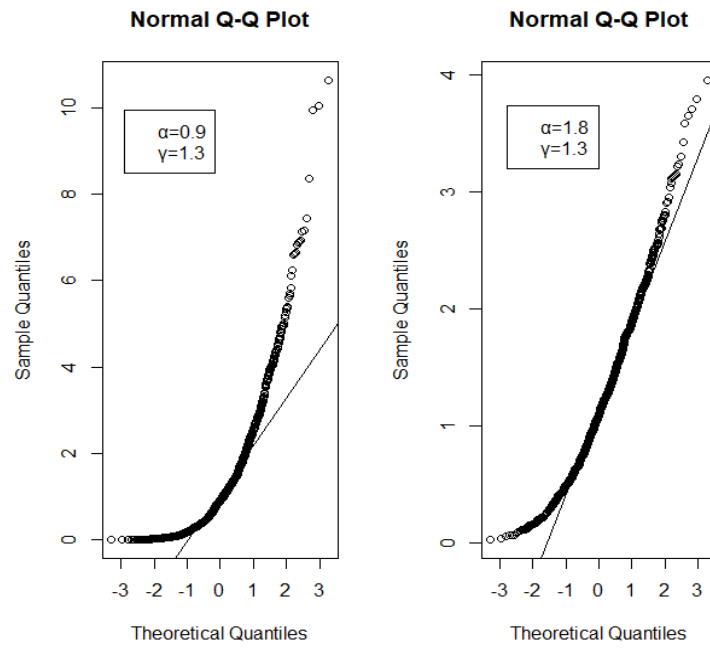
$$h(x) = \frac{f(x)}{\bar{F}(x)}$$

της κατανομής δίνεται από τον τύπο

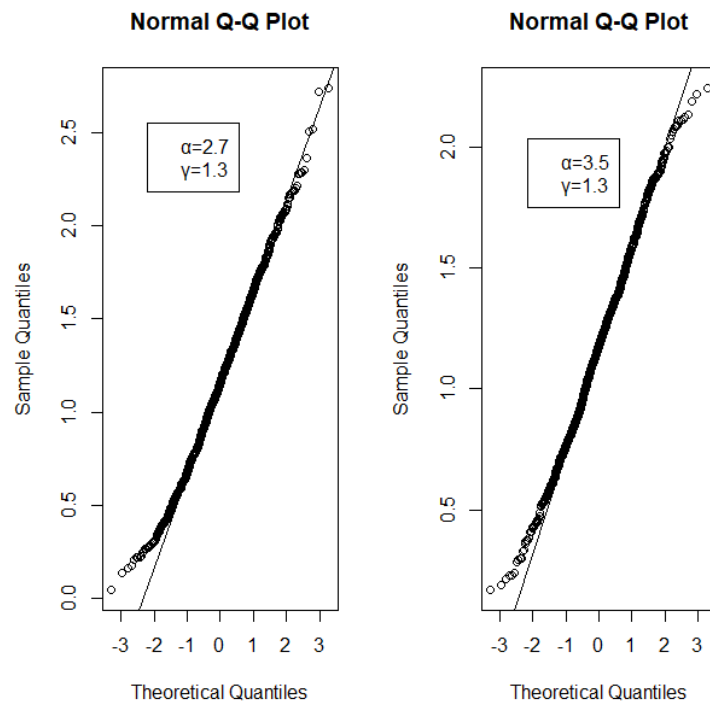
$$h(x; \alpha, \gamma) = \alpha \gamma x^{\alpha-1}.$$

Η κατανομή Weibull όταν αυξάνει η τιμή της παραμέτρου σχήματος α προσεγγίζει την κανονική κατανομή με τον αντίστοιχο μέσο και διασπορά. Για να το διαπιστώσουμε αυτό γραφικά, δημιουργήθηκαν 1000 τυχαίες τιμές από την κατανομή Weibull με διάφορες τιμές της παραμέτρου σχήματος α (0.9, 1.8, 2.7, 3.5) και σταθερή την παράμετρο κλίμακας $\gamma = 1.3$. Στη συνέχεια δημιουργήθηκαν τα αντίστοιχα Q-Q plot για τη κανονική κατανομή. Από τα Σχήματα 1.3, 1.4 προκύπτει, πως όταν η παράμετρος σχήματος α παίρνει τιμές μεγαλύτερες

από 2.7 τα δεδομένα φαίνεται να ακολουθούν την κανονική κατανομή καθώς, συγκλίνουν στην ευθεία $y = x$.



Σχήμα 1.3: Διαγράμματα *Quantile* για την κατανομή Weibull με παραμέτρους $\alpha=0.9$, $\gamma=1.3$ και $\alpha=1.8$, $\gamma=1.3$.



Σχήμα 1.4: Διαγράμματα *Quantile* για την κατανομή Weibull με παραμέτρους $\alpha=2.7$, $\gamma=1.3$ και $\alpha=3.5$, $\gamma=1.3$.

ΚΕΦΑΛΑΙΟ 2

Η ΚΑΤΑΝΟΜΗ WEIBULL ΜΕ ΒΑΡΙΑ ΟΥΡΑ

2.1 ΕΙΣΑΓΩΓΗ

Σε αυτό το κεφάλαιο θα παρουσιαστεί το μοντέλο μιας νέας οικογένειας κατανομών με βαριά ουρά η οποία προκύπτει από την κατανομή Weibull. Το νέο μοντέλο προτάθηκε από τους Ahmad, Mahmoudi and Dey (2020).

2.2 ΔΗΜΙΟΥΡΓΙΑ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΜΟΝΤΕΛΟΥ

Έστω $v(t)$ η συνάρτηση πυκνότητας μιας τυχαίας μεταβλητής T , με $T \in [a_1, a_2]$ για $-\infty \leq a_1 < a_2 \leq +\infty$ και $W[F(x; \xi)]$ η συνάρτηση κατανομής μιας τυχαίας μεταβλητής, έστω X , με παράμετρο ξ , η οποία ικανοποιεί τις παρακάτω συνθήκες :

1. $W[F(x; \xi)] \in [a_1, a_2]$
2. Η $W[F(x; \xi)]$ είναι διαφορίσιμη και μονοτονικά αύξουσα
3. $W[F(x; \xi)] \rightarrow a_1$ όταν $x \rightarrow -\infty$ και $W[F(x; \xi)] \rightarrow a_2$ όταν $x \rightarrow +\infty$.

Η T-X οικογένεια κατανομών περιέχει τυχαίες μεταβλητές των οποίων η αθροιστική συνάρτηση κατανομής ορίζεται ως,

$$G(x) = \int_{a_1}^{W[F(x; \xi)]} v(t) dt, \quad x \in R \quad (2.1.1)$$

Η αντίστοιχη συνάρτηση πιθανότητας δίνεται από τον παρακάτω τύπο:

$$g(x) = \left\{ \frac{\partial}{\partial x} W[F(x; \xi)] v(W[F(x; \xi)]) \right\}, \quad x \in R \quad (2.1.2)$$

Η τεχνική δημιουργίας της T-X οικογένειας προτάθηκε από τους Alzaatreh, Lee, and Famoye (2014). Από αυτή τη προσέγγιση μπορεί να προκύψει ένας μεγάλος αριθμός κατανομών.

Έστω τώρα, ότι η τυχαία μεταβλητή $T \sim \exp(1)$ οπότε η συνάρτηση κατανομής της θα είναι

$$P(t) = 1 - e^{-t}, t \geq 0$$

και αντίστοιχα η συνάρτηση πυκνότητας της είναι

$$p(t) = e^{-t}, t > 0.$$

Χρησιμοποιώντας στην Σχέση (2.1.1) το

$$W[F(x; \xi)] = -\log \left[1 - \left(\frac{\sigma F(x; \xi)}{\sigma - 1 + F(x; \xi)} \right) \right], \text{ με } \sigma > 1$$

προκύπτει πως η συνάρτηση κατανομής της NFHT υπολογίζεται από τον τύπο:

$$G(x; \sigma, \xi) = \frac{\sigma F(x; \xi)}{\sigma - 1 + F(x; \xi)}, \sigma > 1, x \in R \quad (2.1.3)$$

όπου $F(x; \xi)$ είναι η συνάρτηση κατανομής της τυχαίας μεταβλητής X του βασικού μοντέλου.

Η συνάρτηση πυκνότητας όπως προκύπτει από την (2.1.2) δίνεται από τον τύπο:

$$g(x; \sigma, \xi) = \frac{\sigma(\sigma-1)f(x; \xi)}{\{\sigma-1+F(x; \xi)\}^2}, x \in R \quad (2.1.4)$$

όπου $f(x; \xi)$ είναι η συνάρτηση πυκνότητας της τυχαίας μεταβλητής X του βασικού μοντέλου.

Μπορεί να αποδειχθεί πως η συνάρτηση (2.1.3) που ορίσαμε είναι μια έγκυρη συνάρτηση κατανομής.

Πράγματι για τη συνάρτηση κατανομής (2.1.3) εύκολα διαπιστώνεται ότι

$$\lim_{x \rightarrow -\infty} G(x; \sigma, \xi) = 0$$

και

$$\lim_{x \rightarrow +\infty} G(x; \sigma, \xi) = 1.$$

Επίσης είναι προφανές ότι η συνάρτηση $G(x; \sigma, \xi)$ είναι διαφορίσιμη, δεξιά συνεχής και αύξουσα αφού

$$\frac{d}{dx} G(x; \sigma, \xi) = g(x; \sigma, \xi) > 0, \text{ για κάθε } x \in R.$$

Χρησιμοποιώντας την προτεινόμενη μέθοδο μπορεί να παραχθεί ένας σημαντικός αριθμός νέων κατανομών. Ας ονομάσουμε αυτή τη νέα οικογένεια κατανομών ως NFHT (New Family of Heavy Tailed Distributions).

Κάποια σημαντικά πλεονεκτήματα για τη χρήση της NFHT είναι:

- Αποτελεί μια απλή και εύκολη μέθοδο τροποποίησης των ήδη υπάρχοντων κατανομών.

- Βελτιώνει τα χαρακτηριστικά και την ευελιξία των υπαρχόντων κατανομών προσθέτοντας μόνο μία νέα παράμετρο και όχι περισσότερες όπως γίνεται με άλλες μεθόδους.
- Προσφέρει καλύτερη προσαρμογή από άλλα κλασσικά και τροποποιημένα μοντέλα.
- Είναι η καταλληλότερη, συγκριτικά με άλλα ανταγωνιστικά μοντέλα, για την προσαρμογή χρηματοοικονομικών δεδομένων με βαριά ουρά όπως είναι τα δεδομένα ασφαλιστικών ζημιών.

Για μια τυχαία μεταβλητή με συνάρτηση πυκνότητας την (2.1.4) θα χρησιμοποιούμε τον συμβολισμό $X \sim NFHT(x; \sigma, \xi)$. Η συνάρτηση επιβίωσης (survival function ή sf) δίνεται από το παρακάτω τύπο:

$$\begin{aligned}\bar{G}(x; \sigma, \xi) &= 1 - G(x; \sigma, \xi) \\ &= 1 - \frac{\sigma F(x; \xi)}{\sigma - 1 + F(x; \xi)}\end{aligned}$$

ή ισοδύναμα,

$$\bar{G}(x; \sigma, \xi) = \frac{(\sigma-1)(1-F(x;\xi))}{\sigma-1+F(x;\xi)}, \quad x \in R, \quad (2.1.5)$$

ενώ για τη συνάρτηση hrf θα έχουμε

$$h(x; \sigma, \xi) = \frac{g(x; \sigma, \xi)}{\bar{G}(x; \sigma, \xi)} = \frac{\frac{\sigma(\sigma-1)f(x;\xi)}{\{\sigma-1+F(x;\xi)\}^2}}{\frac{(\sigma-1)(1-F(x;\xi))}{\sigma-1+F(x;\xi)}}$$

και τελικά προκύπτει

$$h(x; \sigma, \xi) = \frac{\sigma(\sigma-1)f(x;\xi)}{\bar{F}(x;\xi)(\sigma-1+F(x;\xi))}, \quad x \in R \quad (2.1.6)$$

Όπως αναφέρθηκε στο εισαγωγικό κεφάλαιο, για να έχει μια κατανομή βαριά ουρά θα πρέπει η συνάρτηση δεξιάς ουράς της να ικανοποιεί την συνθήκη:

$$\lim_{x \rightarrow +\infty} e^{\lambda x} \bar{F}(x) = \infty \text{ για κάθε } \lambda > 0.$$

Προκειμένου να ελέγξουμε αν η $\bar{G}(x; \sigma, \xi)$ ικανοποιεί τη συνθήκη αυτή για κάθε $\lambda > 0$, παρατηρούμε αρχικά ότι,

$$e^{\lambda x} \bar{G}(x; \sigma, \xi) = \frac{e^{\lambda x} (\sigma-1)(1-F(x;\xi))}{\sigma-1+F(x;\xi)}.$$

Η $F(x; \xi)$ είναι συνάρτηση κατανομής συνεπώς, όταν το $x \rightarrow +\infty$ θα ισχύει $F(x; \xi) \rightarrow 1$ και $e^{\lambda x} \rightarrow +\infty$ για κάθε $\lambda > 0$.

Επομένως,

$$\lim_{x \rightarrow +\infty} e^{\lambda x} \bar{G}(x; \sigma, \xi) = \frac{e^{\lambda * \infty} (\sigma - 1) (1 - F(\infty; \xi))}{\sigma - 1 + F(\infty; \xi)} = \frac{\infty (\sigma - 1) (1 - 1)}{\sigma - 1 + 1} = \frac{\infty * 0}{\sigma} = \infty * 0,$$

το οποίο είναι απροσδιόριστο. Για αυτό το λόγο, πρέπει να εφαρμοστεί ο κανόνας L'hospital παραγωγίζοντας τόσο τον αριθμητή όσο και τον παρανομαστή για να ελέγξουμε αν προκύπτει το επιθυμητό αποτέλεσμα. Όμως, παραγωγίζοντας τον παρανομαστή, όπου υπάρχει η $F(x; \xi)$, προκύπτει η συνάρτηση πυκνότητας $f(x; \xi)$ για την οποία δεν έχουμε κάποια πληροφορία οπότε δεν γνωρίζουμε πως συμπεριφέρεται όταν το $x \rightarrow +\infty$. Έτσι λοιπόν, ο κανόνας L'hospital αποτυγχάνει να μας δώσει αποτέλεσμα.

Μπορεί να αποδειχθεί ότι αν η $F(x; \xi)$ είναι μια HTD τότε και η $G(x; \sigma, \xi)$ είναι και αυτή HTD.

Πράγματι, γράφοντας την ισότητα

$$e^{\lambda x} \bar{G}(x; \sigma, \xi) = \frac{e^{\lambda x} (\sigma - 1) (1 - F(x; \xi))}{\sigma - 1 + F(x; \xi)}$$

και λαμβάνοντας υπόψη ότι αφού η $F(x; \xi)$ είναι HTD θα ισχύει ότι

$$\lim_{x \rightarrow +\infty} e^{\lambda x} \bar{F}(x; \xi) = \lim_{x \rightarrow +\infty} e^{\lambda x} (1 - F(x; \xi)) = +\infty.$$

συμπεραίνουμε ότι,

$$\begin{aligned} \lim_{x \rightarrow +\infty} e^{\lambda x} \bar{G}(x; \sigma, \xi) &= \lim_{x \rightarrow +\infty} e^{\lambda x} (1 - G(x; \sigma, \xi)) = \\ &= \lim_{x \rightarrow +\infty} \frac{e^{\lambda x} (\sigma - 1) (1 - F(x; \xi))}{\sigma - 1 + F(x; \xi)} = \lim_{x \rightarrow +\infty} \frac{\infty * (\sigma - 1)}{\sigma} = +\infty, \text{ αφού } \sigma > 1. \end{aligned}$$

Μπορεί επίσης να αποδειχθεί ότι η $\bar{G}(x; \sigma, \xi)$ έχει την ιδιότητα της ομαλής κύμανσης όταν η $\bar{F}(x; \xi)$ είναι ομαλής κύμανσης.

Πράγματι, έστω

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}(ax)}{\bar{F}(x)} = g(a)$$

που είναι πεπερασμένο και μη αρνητικό για κάθε $a > 0$.

Παρατηρούμε ότι,

$$\begin{aligned} \frac{\bar{G}(ax; \sigma, \xi)}{\bar{G}(x; \sigma, \xi)} &= \frac{(\sigma - 1)(1 - F(ax; \xi))}{\sigma - 1 + F(ax; \xi)} = \\ &= \frac{(\sigma - 1)(1 - F(x; \xi))}{\sigma - 1 + F(x; \xi)} = \\ &= \frac{(\sigma - 1)(1 - F(ax; \xi))(1 - F(x; \xi))}{(\sigma - 1)(\sigma - 1 + F(ax; \xi))(1 - F(x; \xi))} = \\ &= \frac{\bar{F}(ax; \xi)}{\bar{F}(x; \xi)} * \frac{\sigma - \bar{F}(x; \xi)}{\sigma - \bar{F}(ax; \xi)}. \end{aligned}$$

Αφού η $F(x; \xi)$ είναι συνάρτηση κατανομής τότε όταν $x \rightarrow +\infty$ η $F(x; \xi) \rightarrow 1$, οπότε η $\bar{F}(x; \xi) \rightarrow 0$.

Συνεπώς,

$$\lim_{x \rightarrow +\infty} \frac{\bar{G}(ax; \sigma, \xi)}{\bar{G}(x; \sigma, \xi)} = g(a) * \frac{\sigma - 0}{\sigma - 0} = g(a).$$

Όπως προκύπτει λοιπόν, η $\bar{G}(x; \sigma, \xi)$ είναι ομαλής κύμανσης.

Συνοψίζοντας, σε αυτή την ενότητα παρουσιάστηκε μία μέθοδος παραγωγής κατανομών με βαριά ουρά από ήδη υπάρχουσες κατανομές με βαριά ουρά, όπως προτάθηκε από τους Ahmad, Mahmoudi and Dey (2020). Στην επόμενη ενότητα θα γίνει εφαρμογή της κατανομής Weibull στο μοντέλο που αναπτύχθηκε.

2.3 Η ΚΑΤΑΝΟΜΗ NHTD-WEIBULL

Όπως αναπτύχθηκε στο πρώτο κεφάλαιο η συνάρτηση πυκνότητας της κατανομής Weibull δίνεται από τον τύπο

$$f(x; \alpha, \gamma) = \alpha \gamma x^{\alpha-1} e^{-\gamma x^\alpha}$$

και αντίστοιχα για τη συνάρτησης κατανομής της

$$F(x; \alpha, \gamma) = 1 - e^{-\gamma x^\alpha} \text{ για } x \geq 0 \text{ και } \alpha, \gamma > 0.$$

Αντικαθιστώντας λοιπόν, στη σχέση

$$G(x; \sigma, \alpha, \gamma) = \frac{\sigma F(x; \alpha, \gamma)}{\sigma - 1 + F(x; \alpha, \gamma)}$$

τη συνάρτηση κατανομής της Weibull

$$F(x; \alpha, \gamma) = 1 - e^{-\gamma x^\alpha}$$

προκύπτει η συνάρτηση κατανομής της NHTD-Weibull:

$$G(x; \sigma, \alpha, \gamma) = \frac{\sigma(1 - e^{-\gamma x^\alpha})}{\sigma - 1 + 1 - e^{-\gamma x^\alpha}}$$

ή εναλλακτικά,

$$G(x; \sigma, \alpha, \gamma) = \frac{\sigma(1 - e^{-\gamma x^\alpha})}{\sigma - e^{-\gamma x^\alpha}} \quad (2.2.1)$$

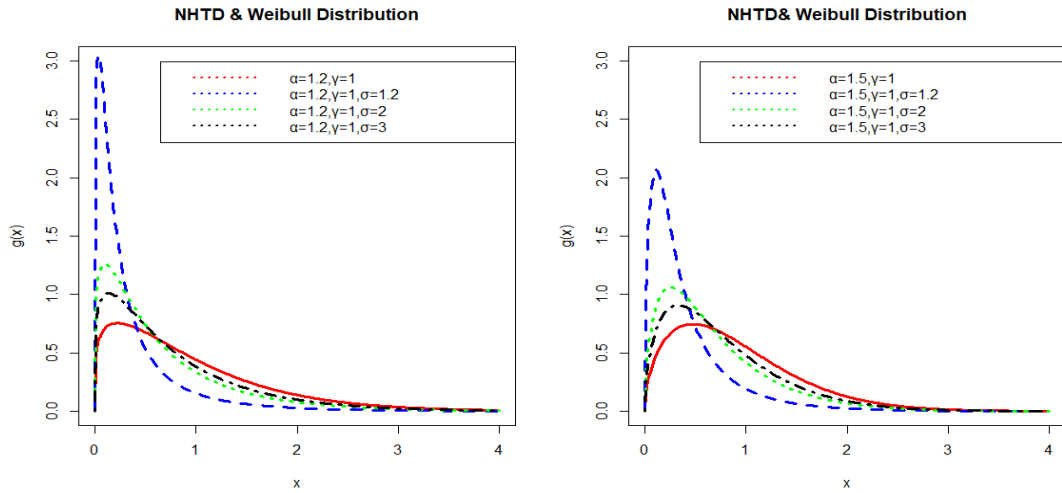
Η συνάρτηση πυκνότητας της NHTD-Weibull και ο τύπος της συνάρτησης κινδύνου hrf θα δίνονται από τους τύπους

$$g(x; \sigma, \alpha, \gamma) = \frac{\sigma(\sigma - 1)\alpha\gamma x^{\alpha-1} e^{-\gamma x^\alpha}}{\{\sigma - e^{-\gamma x^\alpha}\}^2} \quad (2.2.2)$$

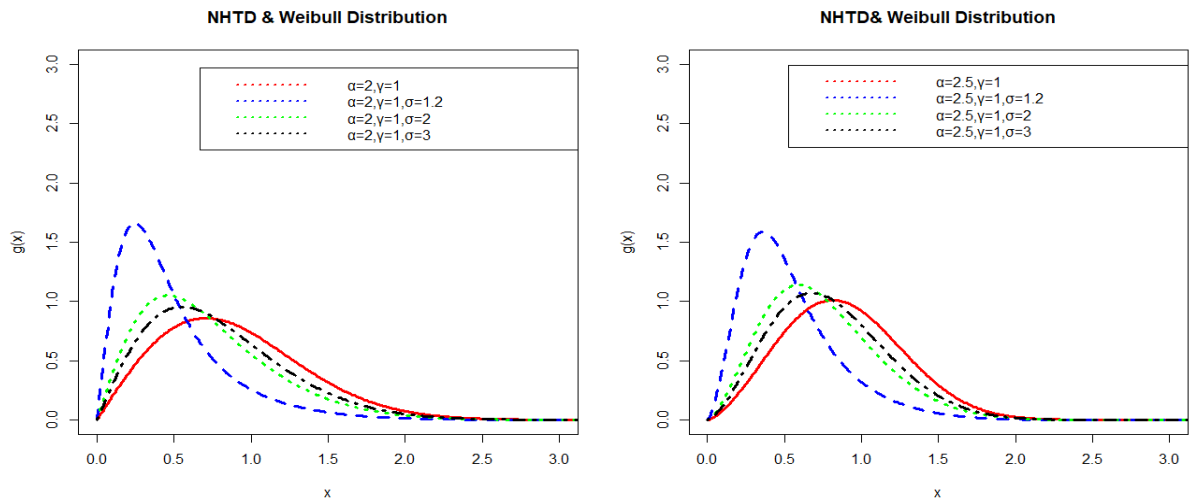
και

$$h(x; \sigma, \alpha, \gamma) = \frac{\sigma(\sigma - 1)\alpha\gamma x^{\alpha-1} e^{-\gamma x^\alpha}}{e^{-\gamma x^\alpha}(\sigma - e^{-\gamma x^\alpha})} \quad (2.2.3)$$

Στα Σχήματα 2.1, 2.2, βλέπουμε την διαγραμματική απεικόνιση των συναρτήσεων πυκνότητας των NHTD-Weibull και Weibull, με χρήση διαφορετικών παραμέτρων. Τόσο στο Σχήμα 2.1 όσο και στο Σχήμα 2.2 γίνεται σύγκριση ανάμεσα στις δύο κατανομές, οι παράμετροι α και γ παραμένουν σταθερές και για τις δύο κατανομές και παρουσιάζεται η συμπεριφορά της NHTD-Weibull για διαφορετικές τιμές της παραμέτρου σ . Παρατηρούμε πως όσο αυξάνεται η τιμή της παραμέτρου σ τόσο πιο βαριά τείνει να γίνεται η ούρα της NHTD-Weibull. Έπειτα, κάνοντας μια σύγκριση για τη συμπεριφορά της NHTD-Weibull για διαφορετικές τιμές της παραμέτρου α παρατηρούμε πως όσο το α αυξάνεται, η τιμή της πιο συχνά παρατηρούμενης τιμής (mode) μειώνεται.

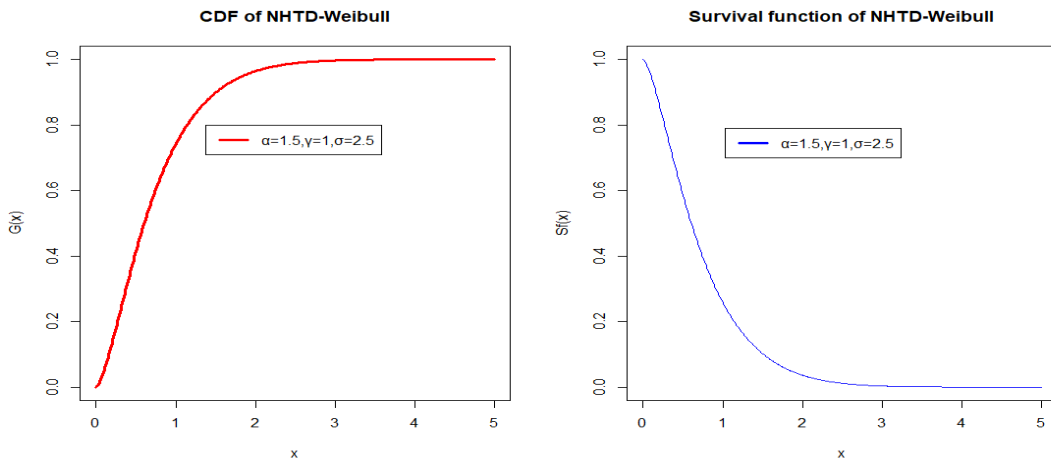


Σχήμα 2.1: Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=1.2$, $\gamma=1$ και διάφορες τιμές του σ (αριστερά). Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=1.5$, $\gamma=1$ και διάφορες τιμές του σ (δεξιά).



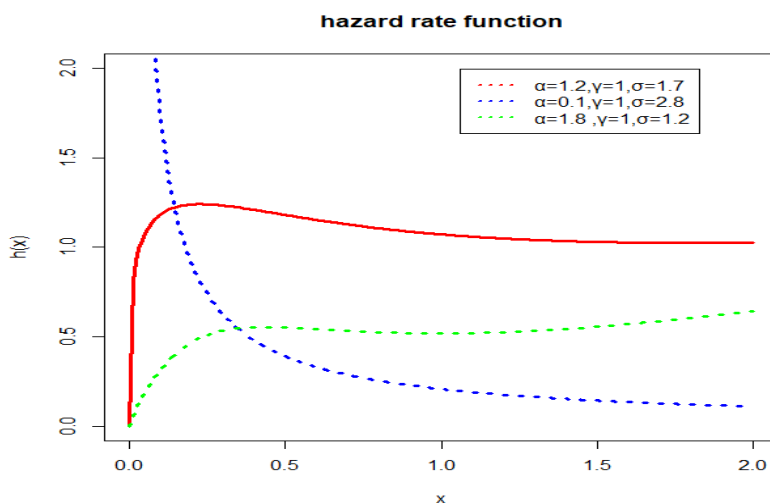
Σχήμα 2.2: Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=2$, $\gamma=1$ και διάφορες τιμές του σ (αριστερά). Σχηματική απεικόνιση συνάρτησης πυκνότητας της NHTD-Weibull και Weibull (κόκκινη καμπύλη) για $\alpha=2.5$, $\gamma=1$ και διάφορες τιμές του σ (δεξιά).

Στο Σχήματα 2.3 απεικονίζεται συνάρτηση κατανομής και η συνάρτηση επιβίωσης της NHTD-Weibull για επιλεγμένες τιμές των παραμέτρων α, γ, σ .



Σχήμα 2.3: Σχηματική απεικόνιση της συνάρτησης κατανομής της NHTD-Weibull με παραμέτρους $\alpha=1.5$, $\gamma=1$ και $\sigma=2.5$ (αριστερά). Σχηματική απεικόνιση της συνάρτησης επιβίωσης της NHTD-Weibull με παραμέτρους $\alpha=1.5$, $\gamma=1$ και $\sigma=2.5$ (δεξιά).

Στο Σχήμα 2.4 απεικονίζεται η συνάρτηση κινδύνου hrf της NHTD-Weibull για επιλεγμένες τιμές των παραμέτρων.



Σχήμα 2.4: Σχηματική απεικόνιση συνάρτησης κινδύνου hrf

Θα κλείσουμε την ενότητα παρουσιάζοντας δύο τεχνικές για τη δημιουργία προσομοιωμένων δεδομένων. Τα δεδομένα αυτά θα χρησιμοποιηθούν σε επόμενο κεφάλαιο ώστε να γίνει εκτίμηση των παραμέτρων της προτεινόμενης κατανομής καθώς και υπολογισμός των μέτρων κινδύνου VaR και $TVaR$.

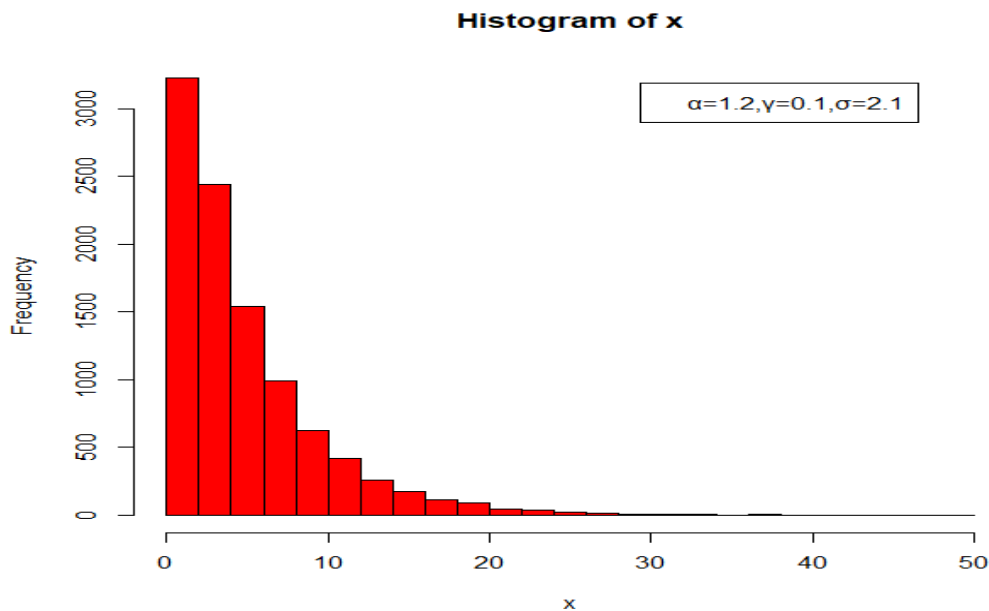
Στο Σχήμα 2.5 απεικονίζεται το ιστόγραμμα συχνοτήτων από προσομοιωμένα δεδομένα, μεγέθους 10000, που δημιουργήθηκαν με τη μέθοδο της αντιστροφής (Μ. Μπούτσικας, 2003). Η διαδικασία έχει ως εξής:

Παράγουμε 10000 τυχαίους αριθμούς $u_1, u_2, \dots, u_{10000} \sim U(0,1)$

Θέτουμε $x = \left[\frac{-\ln\left(\frac{\sigma(u-1)}{u-\sigma}\right)}{\gamma} \right]^{1/\alpha}$, το οποίο προκύπτει ως εξής:

$$G(x; \sigma, \alpha, \gamma) = \frac{\sigma(1 - e^{-\gamma x^\alpha})}{\sigma - e^{-\gamma x^\alpha}} = u \Leftrightarrow e^{-\gamma x^\alpha} = \frac{\sigma(u-1)}{u-\sigma} \Leftrightarrow x^\alpha = -\frac{\ln\left(\frac{\sigma(u-1)}{u-\sigma}\right)}{\gamma} \Leftrightarrow$$

$$x = \left[\frac{-\ln\left(\frac{\sigma(u-1)}{u-\sigma}\right)}{\gamma} \right]^{1/\alpha}.$$



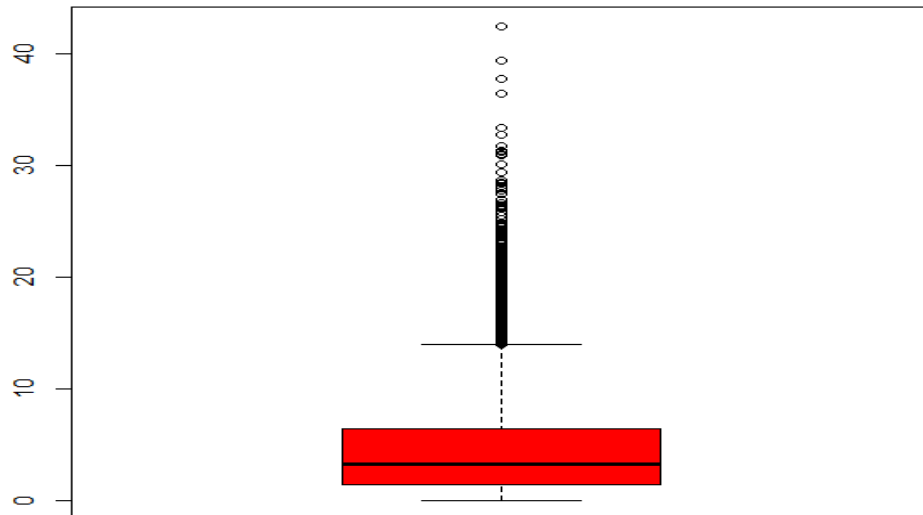
Σχήμα 2.5: Ιστόγραμμα συχνοτήτων για προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.

Να σημειωθεί πως υπάρχει και η μέθοδος της απόρριψης για τη παραγωγή τυχαίων αριθμών, για την οποία ακολουθείται η παρακάτω διαδικασία:

Παράγουμε έναν τυχαίο αριθμό Y από την κατανομή με συνάρτηση πυκνότητας g και στην συνέχεια τον αποδεχόμαστε (θέτουμε $X = Y$) με πιθανότητα ανάλογη του πηλίκου $f(x)/g(x)$. Εάν δεν γίνει αποδεκτός, επαναλαμβάνουμε τη διαδικασία. Έστω ότι κάνουμε N επαναλήψεις, το πλήθος των επαναλήψεων ακολουθεί γεωμετρική κατανομή με πιθανότητα επιτυχίας $\frac{1}{c}$. Συνεπώς, χρειάζονται κατά μέσο όρο $E(N) = c$ επαναλήψεις. Θα πρέπει να πάρουμε το μικρότερο c για το οποίο ισχύει

$$\frac{f(x)}{g(x)} \leq c \text{ για κάθε } x: g(x) \neq 0.$$

Στο Σχήμα 2.6 απεικονίζεται το θηκόγραμμα για τα προσομοιωμένα δεδομένα τα οποία παράχθηκαν με τη μέθοδο της αντιστροφής. Όπως προκύπτει, τα δεδομένα έχουν θετική ασυμμετρία και υπάρχει μεγάλο πλήθος παρατηρήσεων με τιμή αρκετά μεγαλύτερη από τη διάμεσο το οποίο υποδηλώνει πως η κατανομή που ακολουθούν τα δεδομένα έχει βαριά ουρά.



Σχήμα 2.6: Θηκόγραμμα για τα προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.

Και οι δύο μέθοδοι παραγωγής τυχαίων αριθμών μπορούν να χρησιμοποιηθούν τόσο για διακριτές όσο και για συνεχείς κατανομές. Ειδικά, η μέθοδος της απόρριψης είναι πολύ ισχυρή στην συνεχή περίπτωση διότι μας βοηθά να παράγουμε τυχαίους αριθμούς από κατανομές των οποίων η F^{-1} δεν μπορεί να γραφεί σε κλειστή μορφή οπότε δεν θα μπορούσε να γίνει χρήση της μεθόδου της αντιστροφής. Στη δική μας περίπτωση υπάρχει, η αθροιστική συνάρτηση κατανομής έχει κλειστό τύπο οπότε η μέθοδος της αντιστροφής θα χρησιμοποιηθεί για τη παραγωγή προσομοιωμένων δεδομένων, όπως θα παρουσιαστεί στο επόμενο κεφάλαιο.

ΚΕΦΑΛΑΙΟ 3 ΙΔΙΟΤΗΤΕΣ ΤΗΣ ΚΑΤΑΝΟΜΗΣ NFHT ΚΑΙ ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ

3.1 ΕΙΣΑΓΩΓΗ

Σε αυτό το κεφάλαιο θα αναφερθούμε σε κάποιες μαθηματικές ιδιότητες της NFHT κατανομής, θα υπολογιστούν οι εκτιμητές μεγίστης πιθανοφάνειας από προσομοιωμένα δεδομένα και θα δοθούν τύποι υπολογισμού των μέτρων κινδύνου VaR και $TVaR$.

3.2 ΜΑΘΗΜΑΤΙΚΕΣ ΙΔΙΟΤΗΤΕΣ

Η συνάρτηση ποσοστημορίου παράγεται ως εξής:

$$x = Q(u) = G^{-1}(u) = F^{-1}\left(\frac{(\sigma-1)u}{\sigma-u}\right),$$

όπου το $u \in (0,1)$.

Τα μέτρα κυρτότητας και ασυμμετρίας μπορούν να υπολογισθούν μέσω της συνάρτησης ποσοστημορίου.

Η ασυμμετρία Bowley (Bowley,1901) της τ.μ. X δίνεται από τον τύπο

$$\text{Skewness} = \frac{Q(1/4) + Q(3/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)} \quad (3.1.1)$$

ενώ η κυρτότητα του Moor (Moor,1988):

$$\text{Kurtosis} = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)} \quad (3.1.2)$$

Αυτά τα μέτρα είναι λιγότερο ευαίσθητα σε ακραίες παρατηρήσεις. Εξάλλου, υπάρχουν και για κατανομές οι οποίες δεν έχουν ροπές. Οι υπολογισμοί των παραπάνω ποσοτήτων γίνεται πολύ εύκολα αντικαθιστώντας στους τύπους (3.1.1) και (3.1.2) τα $Q(u)$ με χρήση του τύπου που δόθηκε παραπάνω.

Υποθέτοντας ότι η τ.μ. $X \sim \text{NFHT}$ τότε η ροπή r τάξης δίνεται από τον τύπο:

$$\begin{aligned}\mu'_r &= \int_{-\infty}^{+\infty} x^r \frac{\sigma(\sigma-1)f(x; \alpha, \gamma)}{(\sigma-1+F(x; \alpha, \gamma))^2} dx \\ &= \int_{-\infty}^{+\infty} x^r \frac{\sigma f(x; \alpha, \gamma)}{(\sigma-1)\left(1+\frac{F(x; \alpha, \gamma)}{(\sigma-1)}\right)^2} dx\end{aligned}\quad (3.1.3)$$

Αναπτύσσοντας σε σειρά Taylor τη συνάρτηση $\frac{1}{(1+x)^2}$ προκύπτει ότι:

$$\frac{1}{(1+x)^2} = \sum_{n=1}^{+\infty} (-1)^{n-1} n x^{n-1}, \quad |x| < 1.$$

Αντικαθιστώντας το x με $\frac{F(x; \alpha, \gamma)}{(\sigma-1)}$ στην παραπάνω εξίσωση και χρησιμοποιώντας τη Σχέση (3.1.3) έχουμε ότι:

$$\mu'_r = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{n\sigma}{(\sigma-1)^n} \eta_{r,n} \quad (3.1.4)$$

όπου

$$\eta_{r,n} = \int_{-\infty}^{+\infty} x^r f(x; \xi) F(x; \xi)^{n-1} dx.$$

Η ροπογεννήτρια συνάρτηση της κατανομής NFHT, $M_X(t)$ δίνεται από τον τύπο:

$$M_X(t) = \sum_{r=0}^{+\infty} \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{n\sigma t^r}{(\sigma-1)^n r!} \eta_{r,n} \quad (3.1.5)$$

Ο υπολειπόμενος χρόνος ζωής έχει ευρύτερες εφαρμογές τόσο στη θεωρία αξιοπιστίας όσο και στη διαχείριση κινδύνων. Αν συμβολίσουμε με $T(x)$ τον υπολειπόμενο χρόνο ζωής ο υπολογισμός του γίνεται με τον παρακάτω τύπο

$$T(x) = X - x | X > x$$

Είναι μια δεσμευμένη τ.μ. καθώς υπολογίζεται εφόσον έχει επέλθει η ηλικία x .

Η συνάρτηση του υπολειπόμενου χρόνου ζωής της τ.μ. $X \sim \text{NFHT}$ συμβολίζεται με $R_t(x)$ και δίνεται από τον τύπο:

$$R_t(x) = \frac{S(x+t)}{S(t)} \quad (3.1.6)$$

από όπου προκύπτει

$$R_t(x) = \frac{\bar{F}(x+t; \alpha, \gamma)\{\sigma - 1 + F(t; \alpha, \gamma)\}}{\bar{F}(t; \alpha, \gamma)\{\sigma - 1 + F(x+t; \alpha, \gamma)\}}$$

ή εναλλακτικά

$$R_t(x) = \frac{[1 - F(x+t; \alpha, \gamma)]\{\sigma - 1 + F(t; \alpha, \gamma)\}}{[1 - F(t; \alpha, \gamma)]\{\sigma - 1 + F(x+t; \alpha, \gamma)\}} \quad (3.1.7)$$

Ουσιαστικά, η συνάρτηση του υπολειπόμενου χρόνου ζωής εκφράζει την παρακάτω πιθανότητα

$$R_t(x) = \Pr(T(x) > t) = \Pr(X - x > t | X > x)$$

το οποίο μπορεί να αναλυθεί περαιτέρω ως εξής

$$R_t(x) = \Pr(X > x + t | X > x) = \frac{\Pr(X > x) \cap (X > x + t)}{\Pr(X > x)} = \frac{\Pr(X > x + t)}{\Pr(X > x)}$$

και τελικά προκύπτει ο τύπος (3.1.6).

3.3 ΕΚΤΙΜΗΤΕΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ

Η μέθοδος της μέγιστης πιθανοφάνειας είναι μια μέθοδος για την εκτίμηση των παραμέτρων από ένα στατιστικό μοντέλο με χρήση των διαθέσιμων δεδομένων. Η μέθοδος της μέγιστης πιθανοφάνειας επιλέγει τις τιμές των παραμέτρων του μοντέλου που μεγιστοποιούν την από κοινού συνάρτηση πιθανότητας ή πυκνότητας του δείγματος. Διαισθητικά, αυτό μεγιστοποιεί την "συμφωνία" από το επιλεγμένο μοντέλο με τα παρατηρούμενα δεδομένα, και για διακριτές τυχαίες μεταβλητές πραγματικά μεγιστοποιεί την πιθανότητα να προκύπτουν τα παρατηρούμενα δεδομένα όταν λαμβάνεται ένα τυχαίο δείγμα από τη συγκεκριμένη κατανομή. Στη συνέχεια αναλύεται η διαδικασία εκτίμησης των παραμέτρων του μοντέλου NFHT με τη μέθοδο της μέγιστης πιθανοφάνειας.

Έστω x_1, x_2, \dots, x_n παρατηρήσεις από τη συνάρτηση πυκνότητας

$$g(x; \sigma, \alpha, \gamma) = \frac{\sigma(\sigma - 1)f(x; \alpha, \gamma)}{\{\sigma - 1 + F(x; \alpha, \gamma)\}^2}$$

Ο λογάριθμος της συνάρτησης πιθανοφάνειας δίνεται από τον τύπο:

$$L(x_i; \sigma, \alpha, \gamma) = n \log \sigma + n \log(\sigma - 1) + \sum_{i=1}^n \log f(x_i; \alpha, \gamma) - 2 \sum_{i=1}^n \log(\sigma - 1 + F(x_i; \alpha, \gamma))$$

Για να μεγιστοποιήσουμε τη παραπάνω συνάρτηση πρέπει να τη παραγωγίσουμε ως προς τις τρεις παραμέτρους της σ , α , γ . Έτσι βρίσκουμε,

$$\frac{\partial}{\partial \sigma} L(x_i; \sigma, \alpha, \gamma) = \frac{n}{\sigma} + \frac{n}{\sigma - 1} - 2 \sum_{i=1}^n \frac{1}{\sigma - 1 + F(x_i; \alpha, \gamma)} \quad (3.2.1)$$

και

$$\frac{\partial}{\partial \alpha} L(x_i; \sigma, \alpha, \gamma) = \sum_{i=1}^n \frac{\partial f(x_i; \alpha, \gamma) / \partial \alpha}{f(x_i; \alpha, \gamma)} - 2 \sum_{i=1}^n \frac{\partial F(x_i; \alpha, \gamma) / \partial \alpha}{\sigma - 1 + F(x_i; \alpha, \gamma)} \quad (3.2.2)$$

$$\frac{\partial}{\partial \gamma} L(x_i; \sigma, \alpha, \gamma) = \sum_{i=1}^n \frac{\partial f(x_i; \alpha, \gamma) / \partial \gamma}{f(x_i; \alpha, \gamma)} - 2 \sum_{i=1}^n \frac{\partial F(x_i; \alpha, \gamma) / \partial \gamma}{\sigma - 1 + F(x_i; \alpha, \gamma)} \quad (3.2.3)$$

Για να υπολογίσουμε τους εκτιμητές μεγίστης πιθανοφάνειας (εμπ) θέτουμε και τις τρεις παραγώγους ίσες με μηδέν και λύνουμε το σύστημα των 3 εξισώσεων ως προς τις παραμέτρους α , σ , γ .

Είναι προφανές ότι οι εξισώσεις (3.2.1), (3.2.2), (3.2.3) έχουν αρκετά πολύπλοκη μορφή οπότε δεν είναι εφικτή η εύρεση κλειστών τύπων για τα α , σ , γ . Για την επίλυση του συστήματος μπορούμε να κάνουμε χρήση αριθμητικών επαναληπτικών μεθόδων.

3.4 ΠΡΟΣΟΜΟΙΩΣΗ

Σε αυτή την ενότητα, αξιολογείται η συμπεριφορά των εκτιμητών μεγίστης πιθανοφάνειας από ένα πεπερασμένο δείγμα μεγέθους n . Η προσομοίωση πραγματοποιείται με βάση την κατανομή NHTW, δημιουργώντας τυχαίους αριθμούς από την κατανομή με τη βοήθεια της γλώσσας προγραμματισμού R. Τα παρακάτω βήματα επεξηγούν τη διαδικασία που ακολουθήθηκε:

- Παράγονται 1000 δείγματα μεγέθους $n=25, 50, 75, \dots, 1000$ από την NHTW κατανομή.
- Υπολογίζονται οι εκτιμητές μεγίστης πιθανοφάνειας για τις παραμέτρους του μοντέλου.

- Υπολογίζονται οι εκτιμήσεις των Mean Square Error (MSEs) και η μεροληψία σύμφωνα με τους ακόλουθους τύπους:

$$MSE(w) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{w}_i - w)^2,$$

και

$$Bias(w) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{w}_i - w)$$

και

$$Absolute\ Bias(w) = \frac{1}{1000} \sum_{i=1}^{1000} |\hat{w}_i - w|$$

για $w \in (\alpha, \sigma, \gamma)$, όπου \hat{w}_i είναι ο εκτιμητής μεγίστης πιθανοφάνειας του w κατά την i -οστή επανάληψη.

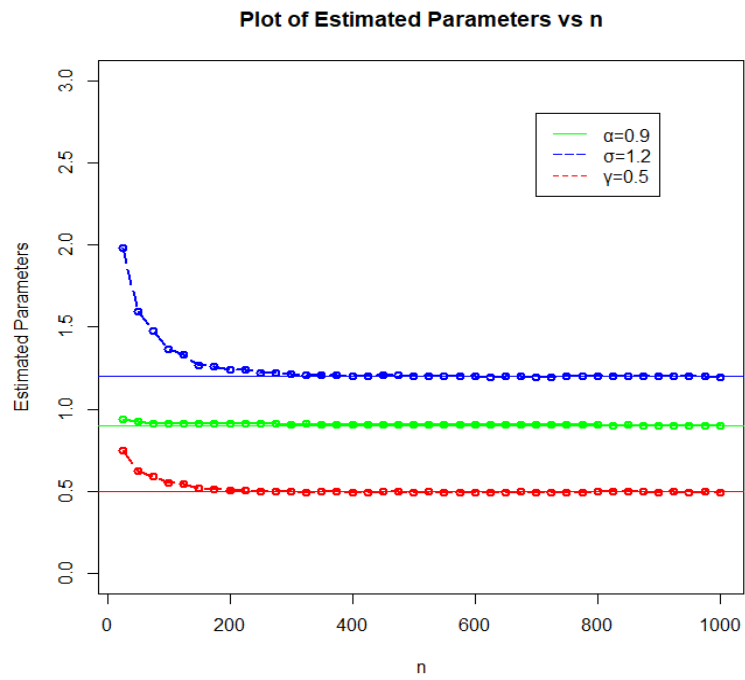
Για διάφορες τιμές των παραμέτρων τα αποτελέσματα της προσομοίωσης φαίνονται στα Σχήματα 3.1-3.8. Από τα αποτελέσματα που προκύπτουν μπορούν να εξαχθούν τα ακόλουθα συμπεράσματα:

- Οι εκτιμητές είναι αρκετά σταθεροί και το σημαντικότερο είναι πως βρίσκονται αρκετά κοντά στις πραγματικές τιμές των παραμέτρων.
- Οι εκτιμητές μεροληψίας (bias) μειώνονται όσο αυξάνεται το μέγεθος του δείγματος.
- Οι εκτιμήσεις απόλυτης μεροληψίας (absolute bias) μειώνονται όσο αυξάνεται το μέγεθος του δείγματος.
- Οι εκτιμήσεις των MSEs τείνουν προς το μηδέν όσο αυξάνεται το μέγεθος του δείγματος.

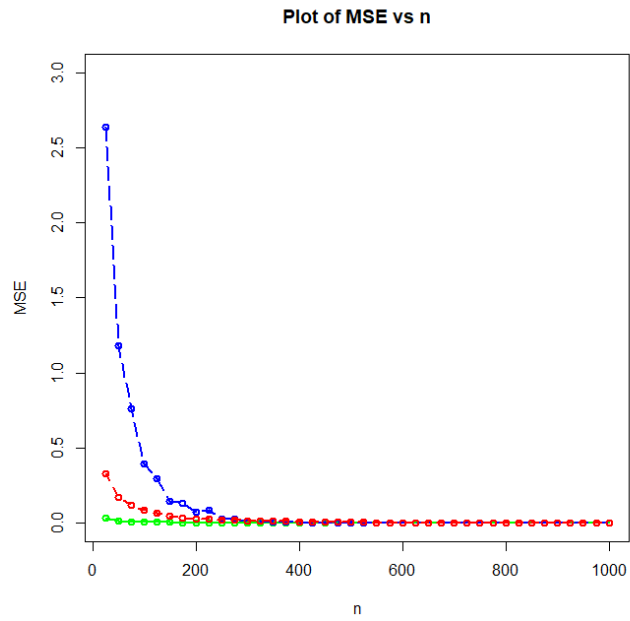
Τα αριθμητικά αποτελέσματα που παρουσιάζονται μέσω των σχημάτων αποδεικνύουν την συνέπεια των εκτιμητών μεγίστης πιθανοφάνειας.

$\alpha=0.9, \sigma=1.2, \gamma=0.5$									
n	MLE			MSE			Bias		
	α	σ	γ	α	σ	γ	α	σ	γ
25	0.94067	1.98033	0.74672	0.03673	2.63360	0.32913	0.04067	0.78033	0.24672
50	0.92309	1.59171	0.62585	0.01789	1.18153	0.17457	0.02309	0.39171	0.12585
75	0.91109	1.47876	0.59030	0.01156	0.76119	0.11708	0.01109	0.27876	0.09030
100	0.91562	1.36650	0.55499	0.00973	0.39224	0.08652	0.01562	0.16650	0.05499
125	0.90989	1.33374	0.54669	0.00696	0.29588	0.06712	0.00989	0.13374	0.04669
150	0.91385	1.26902	0.51827	0.00548	0.14296	0.04588	0.01385	0.06902	0.01827
175	0.91002	1.25992	0.51388	0.00437	0.13286	0.03711	0.01002	0.05992	0.01388
200	0.91379	1.23917	0.50655	0.00409	0.07473	0.02934	0.01379	0.03917	0.00655
225	0.91305	1.23861	0.50319	0.00387	0.08526	0.02819	0.01305	0.03861	0.00319
250	0.91294	1.22150	0.50049	0.00326	0.02646	0.02186	0.01294	0.02150	0.00049
275	0.91117	1.21862	0.49729	0.00277	0.02988	0.01931	0.01117	0.01862	-0.00271
300	0.90796	1.21389	0.50284	0.00219	0.01522	0.01612	0.00796	0.01389	0.00284
325	0.91045	1.20762	0.49516	0.00216	0.01171	0.01353	0.01045	0.00762	-0.00484
350	0.90919	1.20972	0.49674	0.00202	0.01089	0.01400	0.00919	0.00972	-0.00326
375	0.90803	1.20731	0.49785	0.00171	0.00925	0.01210	0.00803	0.00731	-0.00215
400	0.90747	1.20375	0.49548	0.00136	0.00575	0.00934	0.00747	0.00375	-0.00452
425	0.90746	1.20037	0.49177	0.00124	0.00417	0.00777	0.00746	0.00037	-0.00823
450	0.90616	1.20489	0.49711	0.00109	0.00626	0.00827	0.00616	0.00489	-0.00289
475	0.90621	1.20463	0.49749	0.00106	0.00502	0.00804	0.00621	0.00463	-0.00251
500	0.90732	1.19975	0.49327	0.00100	0.00362	0.00615	0.00732	-0.00025	-0.00673
525	0.90468	1.20292	0.49760	0.00082	0.00366	0.00681	0.00468	0.00292	-0.00240
550	0.90481	1.19893	0.49318	0.00079	0.00238	0.00485	0.00481	-0.00108	-0.00682
575	0.90550	1.19844	0.49451	0.00073	0.00198	0.00430	0.00550	-0.00156	-0.00549
600	0.90505	1.20043	0.49464	0.00066	0.00287	0.00503	0.00505	0.00043	-0.00536
625	0.90611	1.19737	0.49265	0.00062	0.00154	0.00380	0.00611	-0.00263	-0.00735
650	0.90376	1.20010	0.49646	0.00055	0.00181	0.00383	0.00376	0.00010	-0.00354
675	0.90327	1.20015	0.49735	0.00043	0.00190	0.00320	0.00327	0.00015	-0.00265
700	0.90464	1.19749	0.49327	0.00044	0.00160	0.00318	0.00464	-0.00251	-0.00673
725	0.90424	1.19693	0.49315	0.00038	0.00098	0.00249	0.00424	-0.00307	-0.00685
750	0.90388	1.19887	0.49563	0.00040	0.00194	0.00303	0.00388	-0.00113	-0.00437
775	0.90348	1.19821	0.49429	0.00042	0.00114	0.00244	0.00348	-0.00179	-0.00571
800	0.90364	1.19978	0.49725	0.00038	0.00127	0.00226	0.00364	-0.00022	-0.00275
825	0.90293	1.20024	0.49784	0.00032	0.00108	0.00224	0.00293	0.00024	-0.00216
850	0.90305	1.19918	0.49688	0.00027	0.00082	0.00162	0.00305	-0.00082	-0.00312
875	0.90234	1.19946	0.49784	0.00028	0.00080	0.00166	0.00234	-0.00054	-0.00216
900	0.90256	1.19863	0.49643	0.00021	0.00062	0.00139	0.00256	-0.00137	-0.00357
925	0.90276	1.19966	0.49737	0.00026	0.00138	0.00162	0.00276	-0.00034	-0.00263
950	0.90244	1.19832	0.49656	0.00020	0.00044	0.00114	0.00244	-0.00168	-0.00344
975	0.90216	1.19916	0.49666	0.00020	0.00079	0.00158	0.00216	-0.00084	-0.00334
1000	0.90288	1.19737	0.49470	0.00020	0.00048	0.00117	0.00288	-0.00263	-0.00530

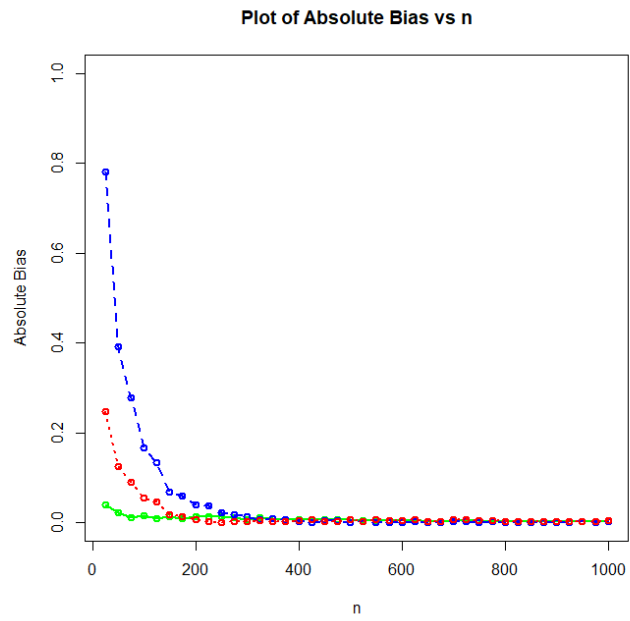
Πίνακας 1: Παρουσίαση αποτελεσμάτων προσομοίωσης για την κατανομή NHTW με $\alpha=0.9, \sigma=1.2, \gamma=0.5$.



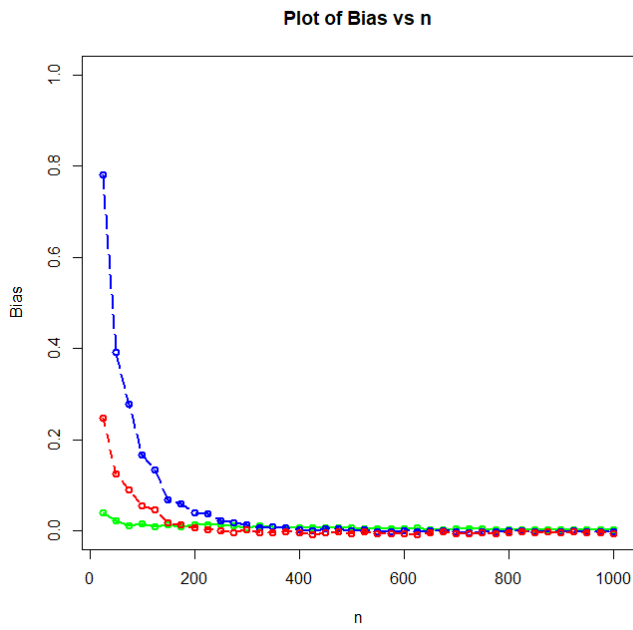
Σχήμα 3.1: Εκτιμήσεις των παραμέτρων για δείγματα μεγέθους n για $\alpha=0.9$, $\sigma=1.2$, $\gamma=0.5$.



Σχήμα 3.2: Εκτιμήσεις MSEs για δείγματα μεγέθους n για $\alpha=0.9$, $\sigma=1.2$, $\gamma=0.5$.



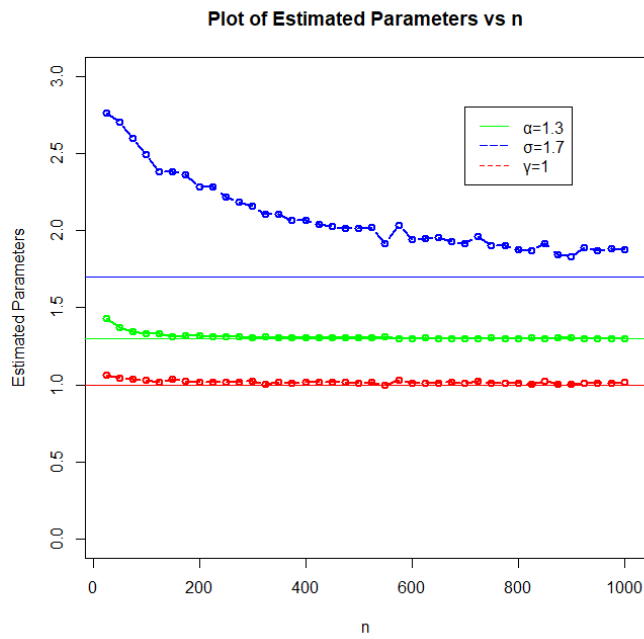
Σχήμα 3.3: Εκτιμήσεις απόλυτης μεροληψίας για δείγματα μεγέθους n για $\alpha=0.9$, $\sigma=1.2$, $\gamma=0.5$.



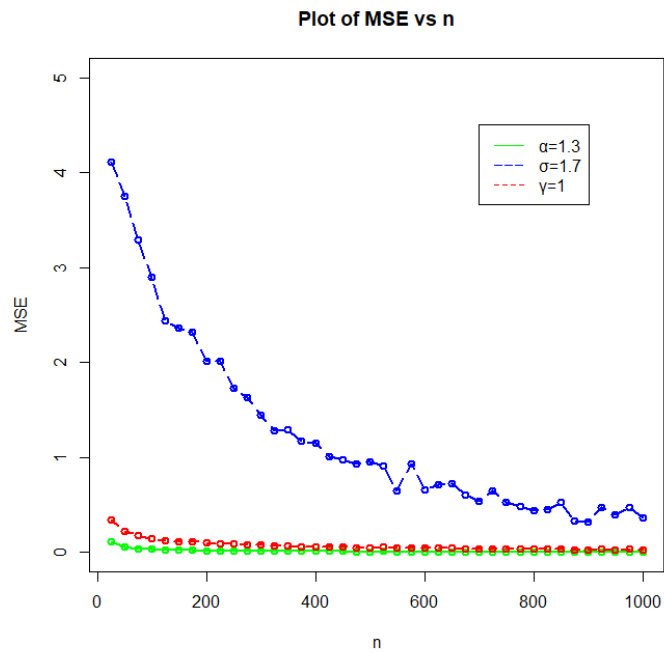
Σχήμα 3.4: Εκτιμήσεις μεροληψίας για δείγματα μεγέθους n για $\alpha=0.9$, $\sigma=1.2$, $\gamma=0.5$.

$\alpha=1.3, \sigma=1.7, \gamma=1$									
n	MLE			MSE			Bias		
	α	σ	γ	α	σ	γ	α	σ	γ
25	1.43259	2.75777	1.06369	0.11217	4.11324	0.34385	0.13259	1.05777	0.06369
50	1.36864	2.70369	1.04079	0.06271	3.75391	0.22275	0.06864	1.00368	0.04079
75	1.34377	2.59869	1.03752	0.03916	3.29151	0.18011	0.04377	0.89869	0.03752
100	1.33103	2.49305	1.03210	0.03099	2.89537	0.14983	0.03103	0.79305	0.03210
125	1.33223	2.38377	1.01719	0.02690	2.44173	0.12733	0.03223	0.68377	0.01719
150	1.31270	2.38342	1.03682	0.02255	2.35746	0.11321	0.01270	0.68341	0.03682
175	1.31907	2.35844	1.02604	0.02123	2.31829	0.10908	0.01907	0.65844	0.02604
200	1.31638	2.27971	1.01618	0.01912	2.01656	0.09677	0.01638	0.57971	0.01618
225	1.31460	2.28324	1.02070	0.01660	2.01159	0.09354	0.01459	0.58324	0.02070
250	1.31160	2.21477	1.01528	0.01563	1.73344	0.08619	0.01160	0.51477	0.01528
275	1.31291	2.18339	1.01627	0.01377	1.63204	0.08087	0.01291	0.48339	0.01627
300	1.30674	2.15951	1.02474	0.01290	1.44807	0.07545	0.00674	0.45951	0.02474
325	1.30942	2.10379	1.00662	0.01250	1.28612	0.06671	0.00942	0.40379	0.00662
350	1.30673	2.10756	1.01634	0.01120	1.29248	0.06845	0.00672	0.40755	0.01634
375	1.30700	2.06701	1.01084	0.01032	1.17238	0.06043	0.00700	0.36701	0.01084
400	1.30639	2.06721	1.01489	0.00997	1.14806	0.06167	0.00639	0.36721	0.01489
425	1.30702	2.04181	1.01561	0.00945	1.00913	0.05857	0.00702	0.34181	0.01561
450	1.30591	2.02885	1.01541	0.00917	0.97501	0.05425	0.00591	0.32885	0.01541
475	1.30287	2.01692	1.01862	0.00821	0.92886	0.05177	0.00287	0.31692	0.01862
500	1.30349	2.01546	1.01175	0.00830	0.95324	0.05091	0.00349	0.31546	0.01175
525	1.30557	2.01894	1.01735	0.00905	0.91472	0.05240	0.00557	0.31894	0.01735
550	1.30982	1.91300	0.99679	0.00736	0.64396	0.04530	0.00982	0.21300	-0.00321
575	1.30028	2.03125	1.02791	0.00721	0.93277	0.04862	0.00028	0.33125	0.02791
600	1.30185	1.93998	1.00949	0.00717	0.65755	0.04256	0.00185	0.23998	0.00949
625	1.30459	1.94735	1.01100	0.00684	0.71142	0.04316	0.00459	0.24735	0.01100
650	1.30253	1.95503	1.00963	0.00661	0.72186	0.04190	0.00253	0.25503	0.00963
675	1.30252	1.92816	1.01548	0.00654	0.59889	0.03947	0.00252	0.22816	0.01548
700	1.29982	1.91321	1.01190	0.00568	0.53664	0.03650	-0.00019	0.21321	0.01190
725	1.29838	1.96037	1.02471	0.00611	0.65058	0.03967	-0.00162	0.26037	0.02471
750	1.30328	1.90189	1.00910	0.00586	0.52247	0.03710	0.00328	0.20189	0.00910
775	1.30046	1.90177	1.01362	0.00522	0.48572	0.03329	0.00046	0.20177	0.01362
800	1.30195	1.87898	1.01241	0.00499	0.43580	0.03218	0.00195	0.17898	0.01241
825	1.30470	1.87290	1.00727	0.00544	0.44817	0.03305	0.00470	0.17290	0.00727
850	1.29857	1.91582	1.02259	0.00510	0.52258	0.03240	-0.00143	0.21582	0.02259
875	1.30296	1.84491	1.00464	0.00480	0.32913	0.03018	0.00296	0.14491	0.00464
900	1.30414	1.82980	1.00304	0.00454	0.31746	0.02839	0.00414	0.12980	0.00304
925	1.30193	1.88896	1.01315	0.00478	0.47480	0.03187	0.00193	0.18896	0.01315
950	1.30037	1.86855	1.01197	0.00448	0.39560	0.02957	0.00037	0.16855	0.01197
975	1.30128	1.88132	1.01202	0.00478	0.47363	0.03057	0.00128	0.18132	0.01202
1000	1.29809	1.87712	1.01827	0.00453	0.36207	0.02983	-0.00191	0.17712	0.01827

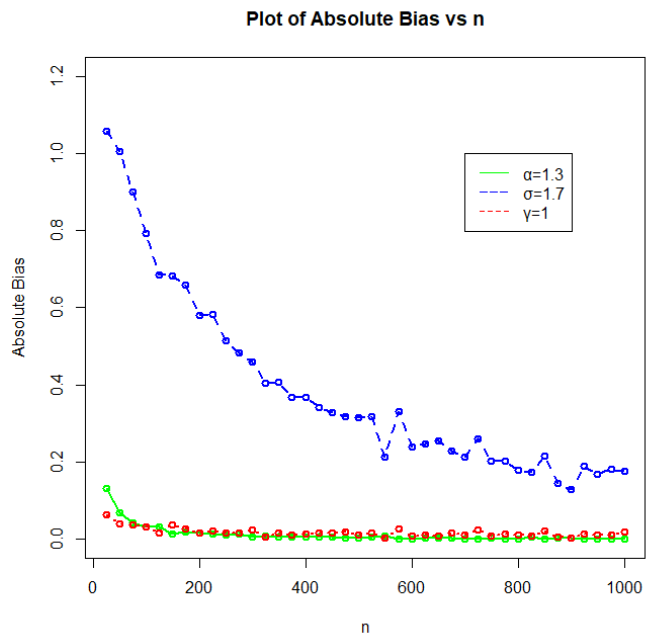
Πίνακας 2: Παρουσίαση αποτελεσμάτων προσομοίωσης για την κατανομή NHTW με $\alpha=1.3, \sigma=1.7, \gamma=1$.



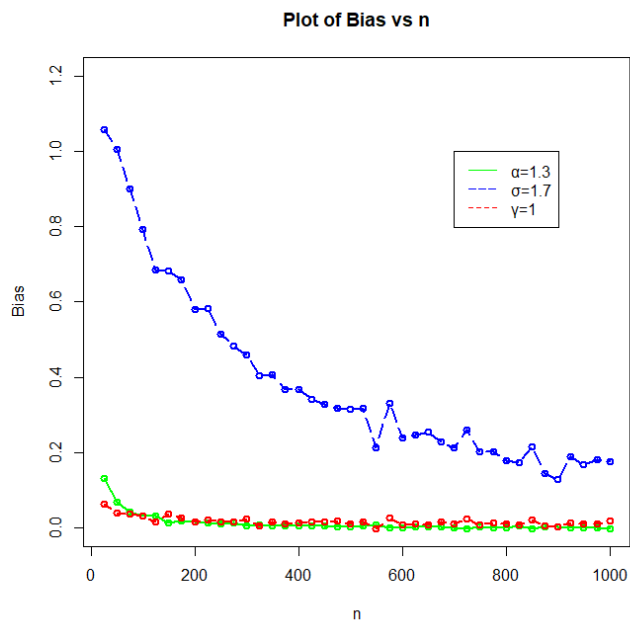
Σχήμα 3.5: Εκτιμήσεις των παραμέτρων για δείγματα μεγέθους n για $\alpha=1.3$, $\sigma=1.7$, $\gamma=1$.



Σχήμα 3.6: Εκτιμήσεις MSEs για δείγματα μεγέθους n για $\alpha=1.3$, $\sigma=1.7$, $\gamma=1$.



Σχήμα 3.7: Εκτιμήσεις απόλυτης μεροληψίας για δείγματα μεγέθους n για $\alpha=1.3$, $\sigma=1.7$, $\gamma=1$.



Σχήμα 3.8: Εκτιμήσεις μεροληψίας για δείγματα μεγέθους n για $\alpha=1.3$, $\sigma=1.7$, $\gamma=1$.

3.5 ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ

Ένας από τους πιο σημαντικούς στόχους της αναλογιστικής επιστήμης είναι ο υπολογισμός της έκθεσης στον κίνδυνο αγοράς σε ένα χαρτοφυλάκιο επενδύσεων, ο οποίος προκύπτει από μεταβολές στις υποκείμενες μεταβλητές όπως είναι οι τιμές των μετοχών, των επιτοκίων και των συναλλαγματικών ισοτιμιών. Σε αυτή την ενότητα, θα υπολογιστούν τα μέτρα κινδύνου αξία σε κίνδυνο (VaR) καθώς και το expected shortfall (ονομάζεται αλλιώς και $TVaR$) για την NHTW κατανομή.

3.5.1 Αξία σε κίνδυνο

Η αξία σε κίνδυνο (Value at risk ή VaR) είναι ένα μέτρο κινδύνου το οποίο εκτιμά το ποσό που μπορεί να χαθεί σε ένα σύνολο επενδύσεων (με δεδομένη πιθανότητα), δεδομένων των συνθηκών της αγοράς, σε μια καθορισμένη χρονική περίοδο, όπως μια ημέρα, ένας μήνας, ένας χρόνος. Η VaR χρησιμοποιείται συνήθως από τις επιχειρήσεις και τις ρυθμιστικές αρχές του χρηματοπιστωτικού κλάδου για τη μέτρηση του ποσού των περιουσιακών στοιχείων που απαιτούνται για την κάλυψη πιθανών ζημιών. Επίσης, ορίζεται ως το μέτρο του ποσοστιαίου κινδύνου με βάση πάντα ένα επίπεδο εμπιστοσύνης (π.χ. 90%, 95% ή 99%) και αντιπροσωπεύει την ποσοστιαία απώλεια που θα συμβεί στην αξία του χαρτοφυλακίου με πιθανότητα $100(1-q)\%$, όπου το $q \in (0,1)$. Το VaR μιας τ.μ. X είναι το q -οστό ποσοστημόριο της συνάρτησης κατανομής της, βλ. Artzner (1999). Αν η τ.μ. X έχει σ.π.π. την (2.2.2) τότε το VaR δίνεται από τη παρακάτω σχέση

$$VaR_q(X) = F^{-1}\left(\frac{(1-q)(\sigma-1)}{\sigma-1+q}\right).$$

Για παράδειγμα, για να υπολογιστεί το VaR σε επίπεδο εμπιστοσύνης 95%, ακολουθείται η παρακάτω διαδικασία. Ζητάμε να ισχύει

$$G(x; \sigma, \alpha, \gamma) = 0.95$$

οπότε θα έχουμε

$$\frac{\sigma F(x; \alpha, \gamma)}{\sigma - 1 + F(x; \alpha, \gamma)} = \frac{\sigma(1 - e^{-\gamma x^\alpha})}{\sigma - 1 + 1 - e^{-\gamma x^\alpha}} = \frac{\sigma(1 - e^{-\gamma x^\alpha})}{\sigma - e^{-\gamma x^\alpha}} = 0.95.$$

Θέτοντας $e^{-\gamma x^\alpha} = z$ προκύπτει,

$$\frac{\sigma - \sigma * z}{\sigma - z} = 0.95$$

οπότε

$$\sigma - \sigma * z = 0.95 * \sigma - 0.95 * z$$

και βρίσκουμε

$$z = \frac{0.05 * \sigma}{\sigma - 0.95} \Rightarrow e^{-\gamma x^\alpha} = \frac{0.05 * \sigma}{\sigma - 0.95}$$

Λογαριθμίζοντας και τα δύο μέλη προκύπτει,

$$-\gamma x^\alpha = \ln\left(\frac{0.05 * \sigma}{\sigma - 0.95}\right)$$

οπότε

$$x = \left(-\frac{\ln\left(\frac{\sigma * 0.05}{\sigma - 0.95}\right)}{\gamma}\right)^{\frac{1}{\alpha}}$$

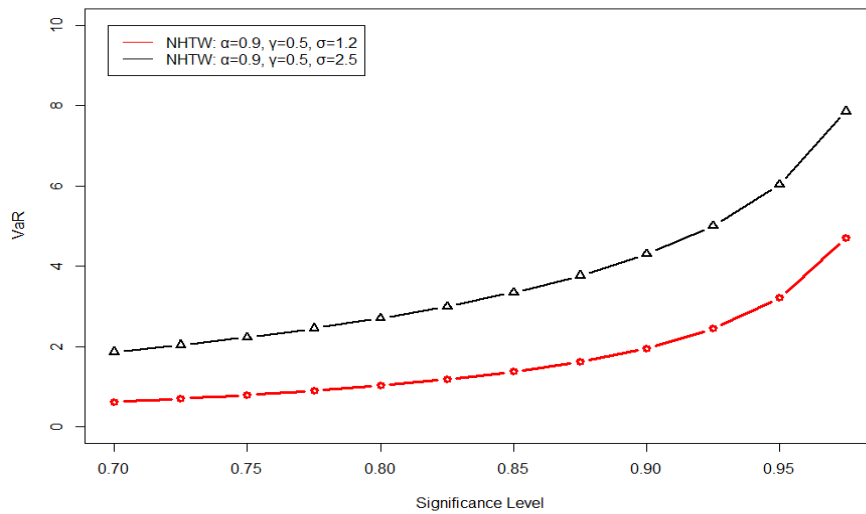
Συνεπώς, το VaR σε επίπεδο εμπιστοσύνης 95% υπολογίζεται από τον τύπο:

$$VaR_{5\%} = \left(-\frac{\ln\left(\frac{\sigma * 0.05}{\sigma - 0.95}\right)}{\gamma}\right)^{\frac{1}{\alpha}}$$

Με ανάλογη διαδικασία μπορούμε να λάβουμε και το γενικό τύπο

$$VaR_{q\%} = \left(-\frac{\ln\left(\frac{\sigma * q}{\sigma - 1 + q}\right)}{\gamma}\right)^{\frac{1}{\alpha}}$$

Στο Σχήμα 3.9 γίνεται απεικόνιση του VaR για τη κατανομή NHTW με διαφορετικές τιμές της παραμέτρου σ και διαφορετικά επίπεδα εμπιστοσύνης. Παρατηρούμε πως, όσο μεγαλύτερη είναι η τιμή της παραμέτρου σ , τόσο μεγαλύτερο είναι το ποσό του κεφαλαίου που πρέπει να διακρατεί μια εταιρία ώστε να μπορεί να ανταπεξέλθει στις υποχρεώσεις της. Αυτό εύκολα μπορεί να επεξηγηθεί αφού, όπως αναφέρθηκε σε προηγούμενη ενότητα, όσο αυξάνεται η τιμή της παραμέτρου σ τόσο πιο «βαριά» τείνει να γίνεται η ουρά της κατανομής NHTW, το οποίο με απλά λόγια σημαίνει πως υπάρχει μια σημαντική πιθανότητα να συμβούν πολύ μεγάλες ζημιές στην εταιρία για αυτό τον λόγο πρέπει να έχει στην κατοχή της μεγαλύτερο κεφάλαιο.

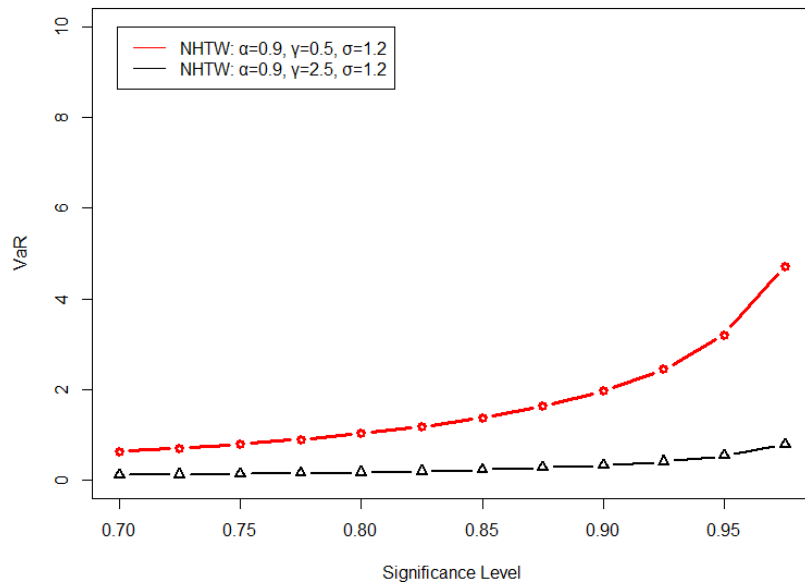


Σχήμα 3.9: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους σ .

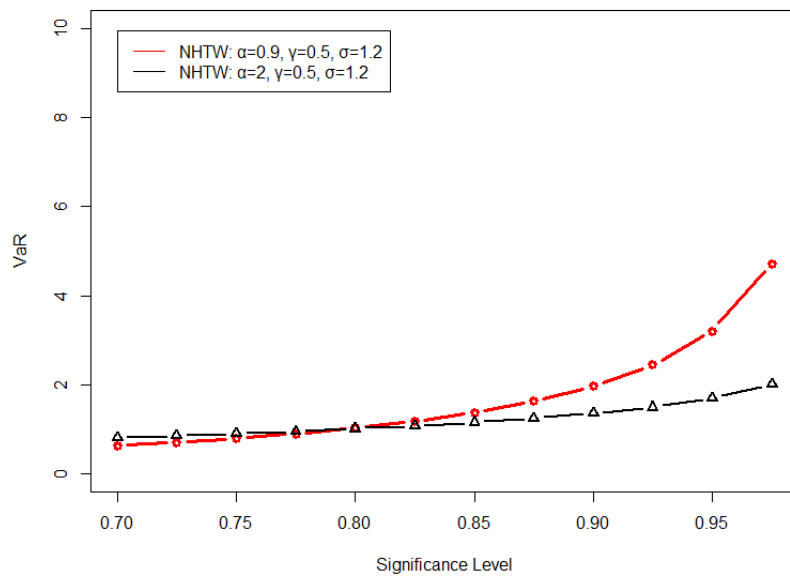
	htwVaR1	htwVaR2
Conf. level	$\alpha=0.9, \sigma=1.2, \gamma=0.5$	$\alpha=0.9, \sigma=2.5, \gamma=0.5$
0.700	0.6270	1.8634
0.725	0.7032	2.0369
0.750	0.7923	2.2313
0.775	0.8976	2.4514
0.800	1.0241	2.7035
0.825	1.1789	2.9964
0.850	1.3728	3.3433
0.875	1.6231	3.7647
0.900	1.9602	4.2951
0.925	2.4426	4.9997
0.950	3.2070	6.0261
0.975	4.7048	7.8512

Πίνακας 3: Αποτελέσματα VaR για τη NHTW με διαφορετικές τιμές παραμέτρων και επίπεδα εμπιστοσύνης.

Στα Σχήματα 3.10 και 3.11 παρουσιάζεται η συμπεριφορά του μέτρου κινδύνου VaR καθώς μεταβάλλεται μόνο η τιμή της παραμέτρου γ και μόνο η τιμή της παραμέτρου α αντίστοιχα. Παρατηρείται πως η τιμή του VaR είναι μικρότερη για μεγαλύτερη τιμή της παραμέτρου γ καθώς και ότι η διαφορά στις τιμές των VaR αυξάνεται σημαντικά σε μεγαλύτερα επίπεδα εμπιστοσύνης. Το ίδιο συμπέρασμα μπορεί να εξαχθεί για το VaR όταν μεταβάλλεται μόνο η τιμή της παραμέτρου α όπως απεικονίζεται στο Σχήμα 3.11.

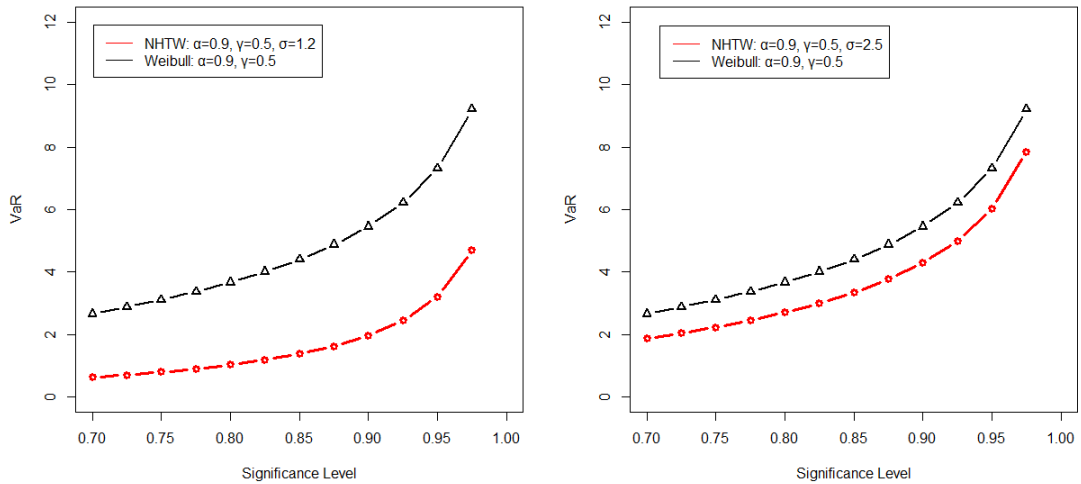


Σχήμα 3.10: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους γ .



Σχήμα 3.11: Σχηματική απεικόνιση VaR της κατανομής NHTW με διαφορετικές παραμέτρους α .

Στο Σχήμα 3.12 γίνεται σύγκριση του μέτρου κινδύνου Value at Risk ανάμεσα στις κατανομές NHTW και Weibull.



Σχήμα 3.12: Σύγκριση VaR των κατανομών NHTW - Weibull.

Όπως ήταν αναμενόμενο το VaR για την κατανομή Weibull είναι μεγαλύτερο σε κάθε επίπεδο εμπιστοσύνης από το VaR της κατανομής NHTW διότι η κατανομή Weibull έχει πιο βαριά ουρά από την κατανομή NHTW. Το οποίο αποδεικνύεται σύμφωνα με τα παρακάτω.

Όπως έχουμε δείξει το VaR της NHTW δίνεται από τη σχέση

$$VaR_q^{NHTW} = \left(\frac{-\ln\left(\frac{\sigma * q}{\sigma - 1 + q}\right)^{\frac{1}{\alpha}}}{\gamma} \right)$$

ενώ το VaR της κατανομής Weibull δίνεται από τη σχέση

$$VaR_q^{Weibull} = \left(\frac{-\ln(q)^{\frac{1}{\alpha}}}{\gamma} \right)$$

και ισχύει διαδοχικά ότι

$$q \leq 1 \Rightarrow q - 1 \leq 0 \Rightarrow \sigma + q - 1 \leq \sigma.$$

Επίσης ισχύει πως $\sigma + q - 1 > 0$, διαιρώντας και τα δυο μέλη της ανισότητας με $\sigma + q - 1$ έχουμε ότι

$$\begin{aligned} \frac{\sigma}{\sigma + q - 1} \geq 1 &\Rightarrow \frac{\sigma q}{\sigma + q - 1} \geq q \Rightarrow \ln\left(\frac{\sigma q}{\sigma + q - 1}\right) \geq \ln(q) \Leftrightarrow \\ \Leftrightarrow -\ln\left(\frac{\sigma q}{\sigma + q - 1}\right) &\leq -\ln(q) \Rightarrow \left(-\ln\left(\frac{\sigma q}{\sigma + q - 1}\right)\right)^{\frac{1}{\alpha}} \leq (-\ln(q))^{\frac{1}{\alpha}} \Leftrightarrow \end{aligned}$$

$$\Leftrightarrow \left(\frac{\left(-\ln \left(\frac{\sigma q}{\sigma + q - 1} \right) \right)^{\frac{1}{a}}}{\gamma} \right) \leq \left(\frac{(-\ln(q))^{\frac{1}{a}}}{\gamma} \right) \Rightarrow VaR_q^{NHTW} \leq VaR_q^{Weibull}.$$

3.5.2 Αναμενόμενο Έλλειμα

Άλλο ένα σημαντικό μέτρο κινδύνου είναι το Αναμενόμενο Έλλειμα ή Expected Shortfall (ES) ή όπως ονομάζεται διαφορετικά Tail Value at Risk ($TVaR$). Το ES μιας τ.μ. X σε επίπεδο εμπιστοσύνης q είναι η αναμενόμενη απώλεια δοθέντος ότι η απώλεια υπερβαίνει το q ποσοστιαίο σημείο της κατανομής της X . Με άλλα λόγια, είναι η μέση τιμή των απωλειών που είναι μεγαλύτερες από το VaR . Αξίζει να σημειωθεί ότι το ES δίνει μέγιστη πληροφορία για την κεφαλαιακή επάρκεια καθώς συμπεριλαμβάνει τις ακραίες ζημιές οι οποίες είναι μεγαλύτερες του VaR . Σύμφωνα με τον ορισμό του το VaR θα δίνεται από τον επόμενο τύπο:

$$TVaR_q(X) = E(X|X > VaR_q) = \frac{\int_{VaR_q}^{\infty} xf(x)dx}{1 - q}.$$

Έστω ότι η τ.μ. X ακολουθεί την NHTW κατανομή. Τότε το $TVaR$ υπολογίζεται από τον παρακάτω τύπο:

$$TVaR_q(X) = \frac{1}{1 - q} \int_{VaR_q}^{\infty} x \frac{\sigma(\sigma - 1)f(x; \alpha, \gamma)}{(\sigma - 1 + F(x; \alpha, \gamma))^2} dx,$$

ή εναλλακτικά

$$TVaR_q(X) = \frac{1}{1 - q} \int_{VaR_q}^{\infty} x \frac{\sigma f(x; \alpha, \gamma)}{(\sigma - 1) \left(1 + \frac{F(x; \alpha, \gamma)}{(\sigma - 1)} \right)^2} dx,$$

Από την παραπάνω σχέση, σε συνδυασμό με το ανάπτυγμα Taylor,

$$\frac{1}{(1 + x)^2} = \sum_{n=1}^{+\infty} (-1)^{n-1} nx^{n-1}, |x| < 1,$$

προκύπτει ο ακόλουθος τύπος για τον υπολογισμό του $TVaR$

$$TVaR_q(X) = \frac{\sigma}{1 - q} \sum_{n=1}^{\infty} \frac{n(-1)^{n-1}}{(\sigma - 1)^n} \int_{VaR_q}^{\infty} xf(x; \alpha, \gamma)F(x; \alpha, \gamma)^{n-1} dx.$$

Αντικαθιστώντας όπου $f(x; \alpha, \gamma)$ και $F(x; \alpha, \gamma)$ τη σ.π.π. και τη συνάρτηση κατανομής της κατανομής Weibull αντίστοιχα, προκύπτει το ακόλουθο αποτέλεσμα:

$$\begin{aligned} TVaR_q(X) &= \frac{\sigma}{1-q} \sum_{n=1}^{\infty} \frac{n(-1)^{n-1}}{(\sigma-1)^n} \int_{VaR_q}^{\infty} \alpha \gamma x^\alpha e^{-\gamma x^\alpha} (1 - e^{-\gamma x^\alpha})^{n-1} dx \\ &= \frac{\sigma}{1-q} \sum_{n=1}^{\infty} \frac{n(-1)^{n-1}}{(\sigma-1)^n} \int_{VaR_q}^{\infty} \alpha \gamma x^\alpha e^{-\gamma x^\alpha} \sum_{i=0}^{\infty} -1^i \binom{n-1}{i} e^{-i\gamma x^\alpha} dx \\ &= \frac{\sigma}{1-q} \sum_{n=1}^{\infty} \sum_{i=0}^{\infty} \binom{n-1}{i} \frac{n(-1)^{i+n-1}}{(\sigma-1)^n} \int_{VaR_q}^{\infty} \alpha \gamma x^\alpha e^{-\gamma(1+i)x^\alpha} dx. \end{aligned}$$

Από τον παραπάνω τύπο προκύπτει ότι

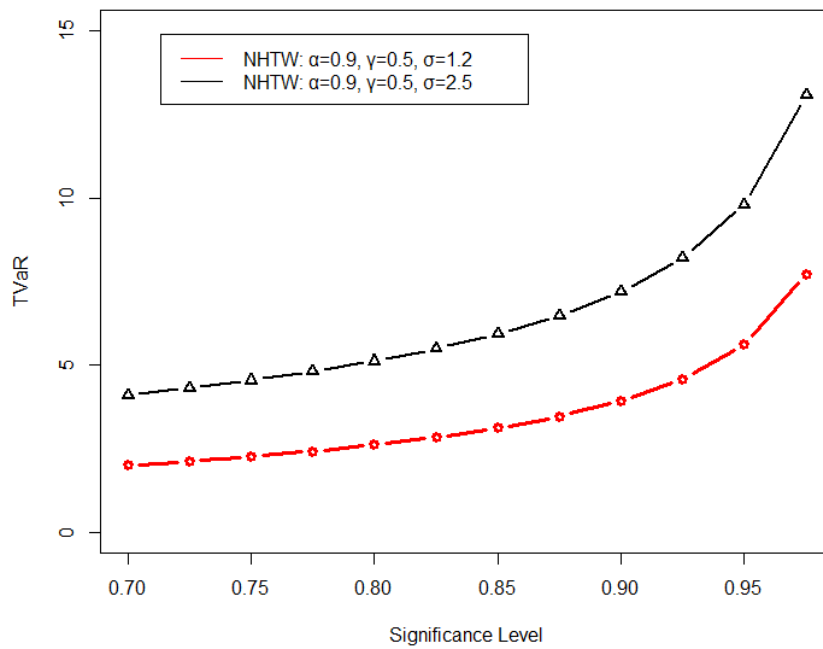
$$\lim_{\sigma \rightarrow \infty} ES_\sigma = \lim_{\sigma \rightarrow \infty} \frac{\sigma}{(\sigma-1)} \int_{VaR_q}^{\infty} \alpha \gamma x^\alpha e^{-\gamma x^\alpha} dx = \int_{VaR_q}^{\infty} \alpha \gamma x^\alpha e^{-\gamma x^\alpha} dx.$$

Για τον υπολογισμό του ES χρησιμοποιήθηκε ο παρακάτω τύπος

$$TVaR_q(X) = \frac{\int_{VaR_q}^{\infty} x f(x) dx}{1-q}$$

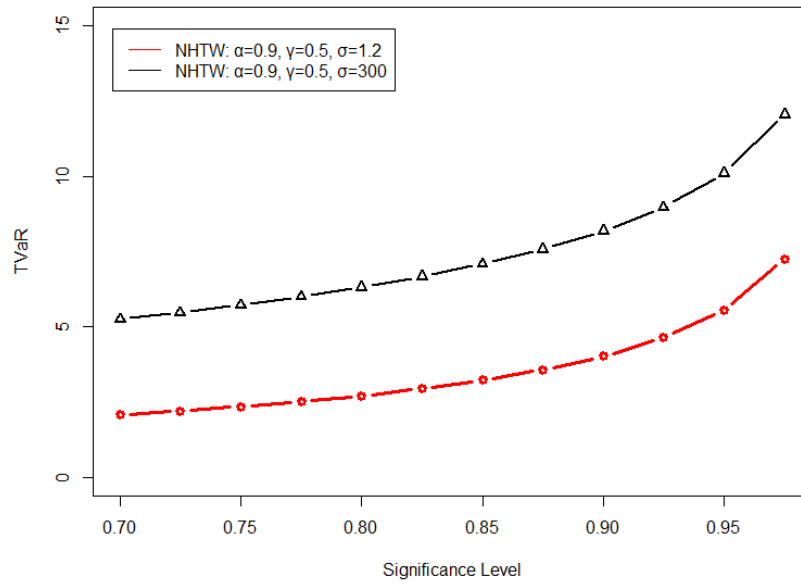
τόσο για τα προσομοιωμένα δεδομένα, όσο και στην εφαρμογή πραγματικών δεδομένων, η οποία θα παρουσιαστεί στο επόμενο κεφάλαιο. Πιο συγκεκριμένα, πρώτα υπολογίστηκε το VaR σε διαφορετικά επίπεδα εμπιστοσύνης (από 0.7-0.975 με βήμα 0.025) και στη συνέχεια για κάθε τιμή του VaR , υπολογίστηκε το ολοκλήρωμα που αποτυπώνεται στον παραπάνω τύπο με τη χρήση της συνάρτησης integrate της γλώσσας προγραμματισμού R, κάθε φορά με διαφορετική τιμή του κάτω ορίου, η οποία είναι ίση με την εκάστοτε τιμή του VaR . Τα αποτελέσματα του υπολογισμού του ES ή $TVaR$ για τα προσομοιωμένα δεδομένα απεικονίζονται στα Σχήματα 3.11 και 3.12.

Στο Σχήμα 3.13 γίνεται σύγκριση του $TVaR$ για την κατανομή NHTW με σταθερές τις παραμέτρους α και γ αλλά με διαφορετική παράμετρο σ . Παρατηρούμε πως όπως ήταν αναμενόμενο, όσο αυξάνεται η τιμή της παραμέτρου σ τόσο πιο μεγάλη γίνεται η τιμή του $TVaR$ για δεδομένο επίπεδο εμπιστοσύνης.

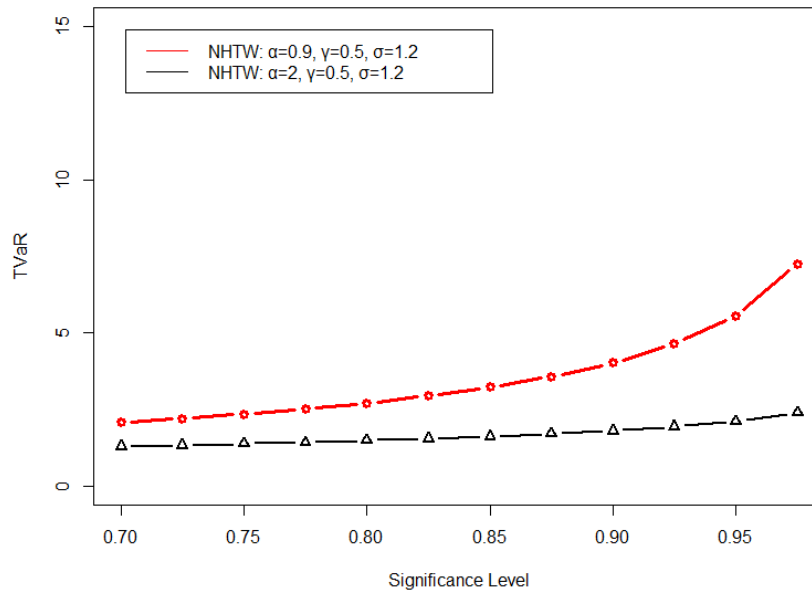


Σχήμα 3.13: Σχηματική απεικόνιση TVaR για τη κατανομή NHTW

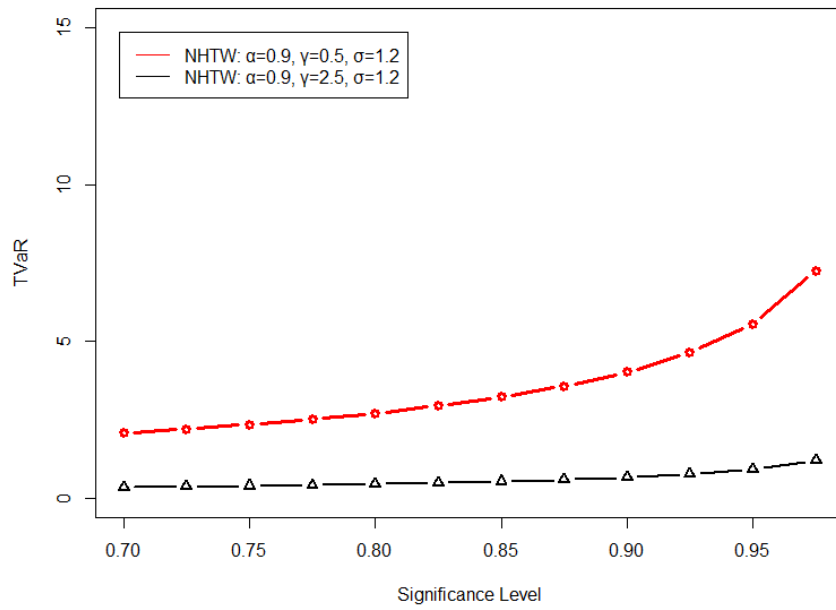
Στα Σχήματα 3.14-3.16 παρουσιάζεται η συμπεριφορά του TVaR της NHTW για διαφορετικές τιμές των παραμέτρων σ , α , γ αντιστοίχως. Από το Σχήμα 3.14 παρατηρείται ότι με την αύξηση της τιμής της παραμέτρου σ υπάρχει μια αύξηση στο μέτρο κινδύνου TVaR η οποία είναι σταθερή σε κάθε επίπεδο εμπιστοσύνης. Η αντίθετη συμπεριφορά παρατηρείται όταν αυξάνεται η εκάστοτε τιμή των παραμέτρων α και γ , όπως παρουσιάζεται στα Σχήματα 3.15 και 3.16 αντίστοιχα, καθώς υψηλότερες τιμές αυτών των παραμέτρων οδηγούν σε χαμηλότερη τιμή του TVaR για κάθε επίπεδο εμπιστοσύνης και μάλιστα η διαφορά αυτή φαίνεται πως αυξάνεται σημαντικά σε υψηλότερα επίπεδα εμπιστοσύνης.



Σχήμα 3.14: Σχηματική απεικόνιση TVaR για τη κατανομή NHTW για διαφορετικές παραμέτρους σ .

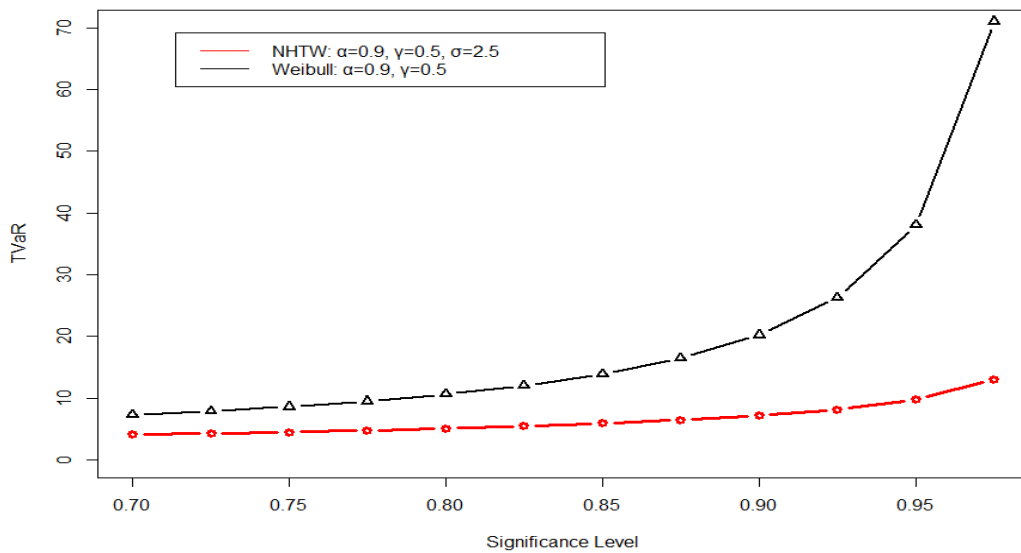


Σχήμα 3.15: Σχηματική απεικόνιση TVaR για τη κατανομή NHTW για διαφορετικές παραμέτρους α .



Σχήμα 3.16: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW για διαφορετικές παραμέτρους γ .

Στο Σχήμα 3.17 γίνεται σύγκριση του $TVaR$ ανάμεσα στη κατανομή NHTW και τη κατανομή Weibull. Παρατηρούμε ότι το $TVaR$ για τη κατανομή Weibull είναι μεγαλύτερο από αυτό της NHTW κατανομής, καθώς η κατανομή Weibull προκύπτει από την NHTW όταν η παράμετρος $\sigma \rightarrow +\infty$.



Σχήμα 3.17: Σχηματική απεικόνιση $TVaR$ για τη κατανομή NHTW και Weibull.

ΚΕΦΑΛΑΙΟ 4

ΕΦΑΡΜΟΓΗ ΔΕΔΟΜΕΝΩΝ

4.1 ΕΙΣΑΓΩΓΗ

Σε αυτό το κεφάλαιο γίνεται παρουσίαση των δεδομένων μιας μεγάλης Ελληνικής ασφαλιστικής εταιρίας τα οποία αναφέρονται σε ζημιές περιουσίας σε ημερήσιο επίπεδο χρονικής διάρκειας δύο ετών 01/01/2007-31/12/2008. Έπειτα γίνεται προσαρμογή των δεδομένων στην NHTW για την εκτίμηση των τιμών των παραμέτρων της, καθώς και ο υπολογισμός των μέτρων κινδύνου *VaR* και *TVaR* τόσο για την NHTW όσο και για την Weibull σε διαφορετικά επίπεδα εμπιστοσύνης.

4.2 ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

Αναφορικά με την ανάλυση των δεδομένων της ασφαλιστικής εταιρίας, έχουν γίνει δύο ειδών ομαδοποιήσεις:

- Συνολικές ζημιές ανά ημέρα πραγματοποίησης της ζημιάς
- Συνολικές ζημιές ανά βδομάδα

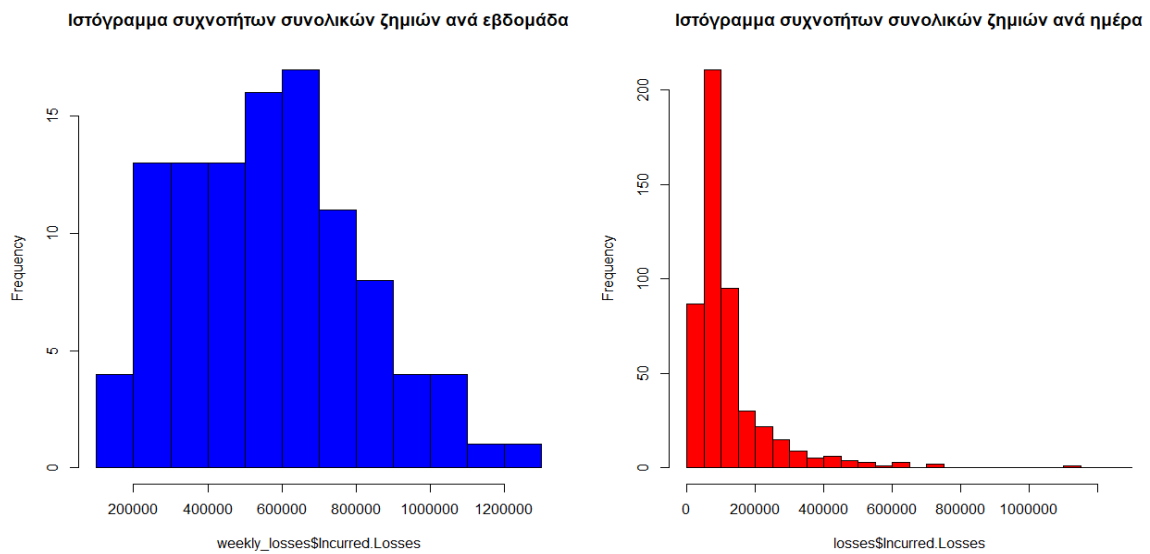
Στον παρακάτω πίνακα φαίνονται κάποια περιγραφικά στοιχεία και για τις δυο αναλύσεις των δεδομένων.

	Συνολικές ζημιές ανά ημέρα	Συνολικές ζημιές ανά εβδομάδα
Ελάχιστη τιμή	382,20	143.698,00
1ο Ποσοστημόριο	58.102,70	369.400,00
Διάμεσος	85.776,70	561.319,00
Μέση τιμή	121.216,70	570.296,00
3ο Ποσοστημόριο	133.803,20	730.113,00
Μέγιστη τιμή	1.143.227,40	1.291.955,00
Διακύμανση	13.715.599.078,00	60.392.845.693,00
Τυπική απόκλιση	117.113,60	245.749,60
Ασυμμετρία	3,32	0,43
Κυρτότητα	19,58	2,84

Κάποια συμπεράσματα που προκύπτουν από τα παραπάνω είναι ότι και για τις δύο αναλύσεις φαίνεται πως τα δεδομένα έχουν θετική ασυμμετρία, ειδικά για τις συνολικές ζημιές

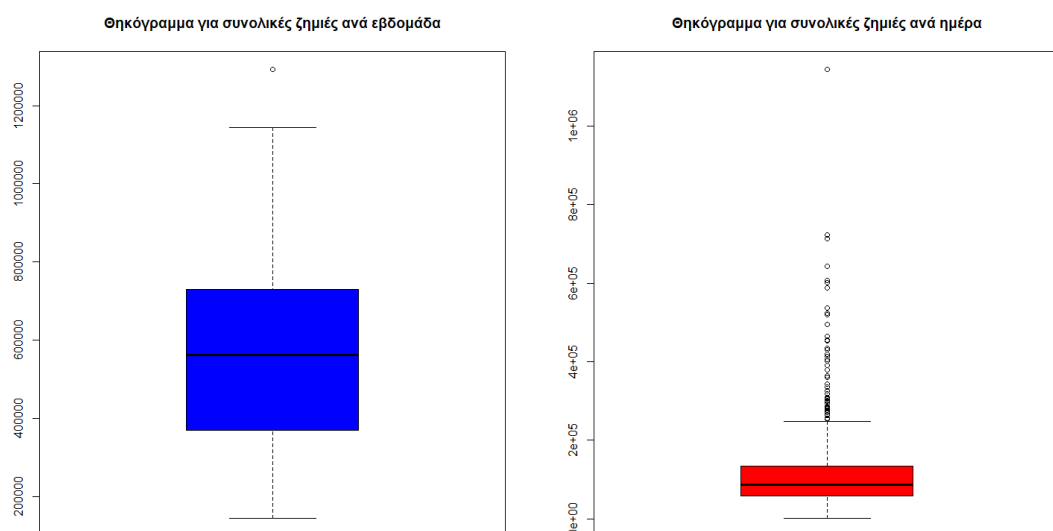
ανά μέρα ο συντελεστής ασυμμετρίας είναι αρκετά μεγαλύτερος σε σύγκριση με τον αντίστοιχο συντελεστή των συνολικών ζημιών ανά εβδομάδα.

Στο Σχήμα 4.1 παρουσιάζονται τα ιστογράμματα συχνοτήτων των δεδομένων και για τις δύο αναλύσεις αντίστοιχα. Φαίνεται πως οι συνολικές ζημιές ανά ημέρα ακολουθούν κατανομή με βαριά ουρά, ενώ οι ζημιές ανά εβδομάδα φαίνεται πως προσεγγίζουν καλύτερα την κανονική κατανομή παρά μια κατανομή με βαριά ουρά.



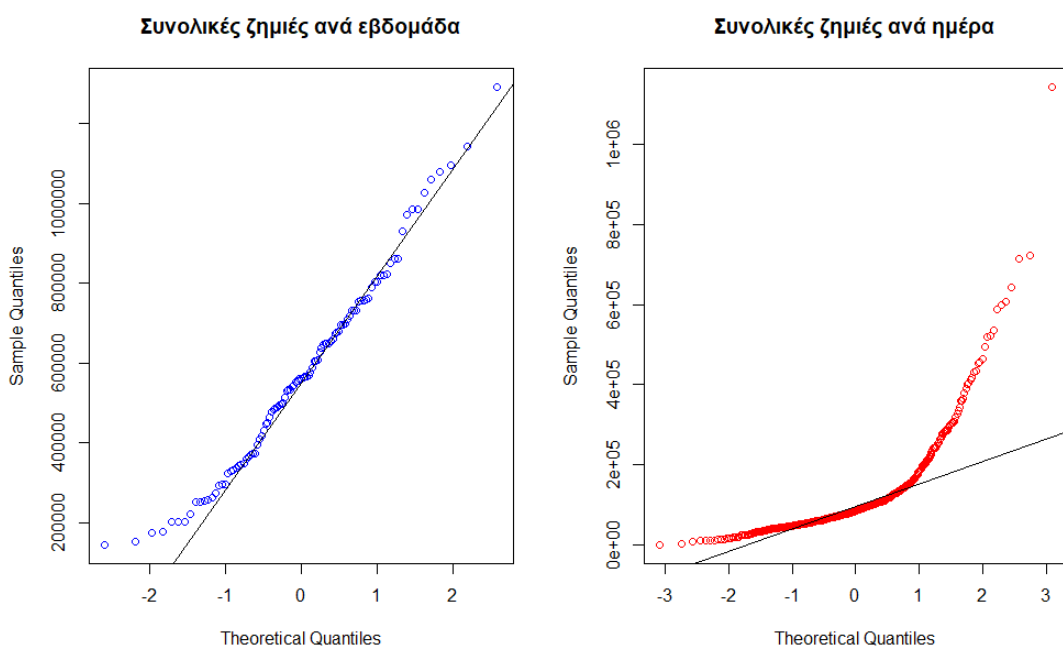
Σχήμα 4.1: Ιστογράμμο συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.

Για να ελεγχθεί ο παραπάνω ισχυρισμός ακολουθεί απεικόνιση των θηκογραμμάτων των δεδομένων και τα αντίστοιχα Q-Q-plots.



Σχήμα 4.2: Θηκόγραμμα συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.

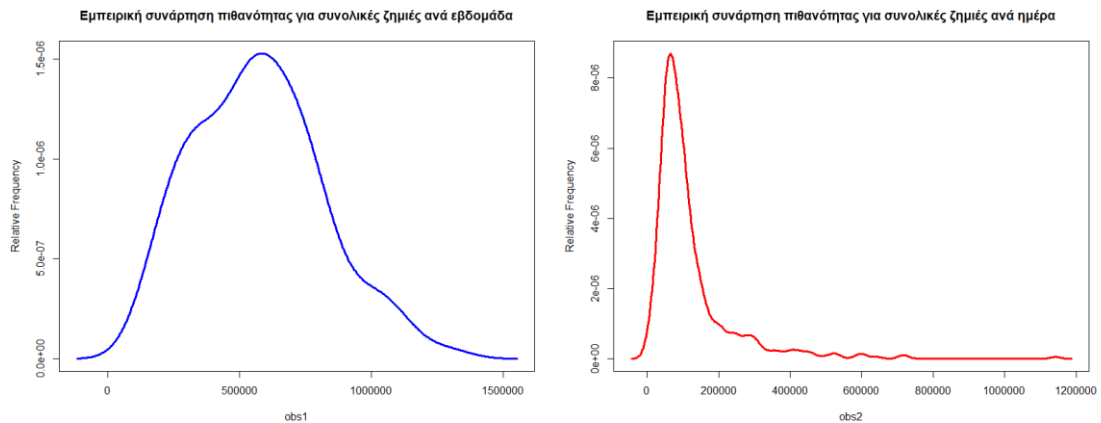
Από τα θηκογράμματα παρατηρείται ότι για τις συνολικές ζημιές ανά εβδομάδα, η διάμεσος βρίσκεται σχεδόν στη μέση του 1^{ου} και 3^{ου} τεταρτημρίου το οποίο αποτελεί χαρακτηριστικό δεδομένων που προσεγγίζονται από την κανονική κατανομή καθώς και ότι υπάρχει μόνο μία ακραία παρατήρηση (outlier). Για τα δεδομένα των συνολικών ζημιών ανά ημέρα η εικόνα είναι αρκετά διαφορετική. Συγκεκριμένα, παρατηρείται ότι η διάμεσος είναι πιο κοντά στο 1^ο τεταρτημριο από ότι στο 3^ο: καθώς και ότι υπάρχει πληθώρα ακραίων παρατηρήσεων, τα οποία χαρακτηριστικά υποδηλώνουν ότι τα δεδομένα ακολουθούν κατανομή με βαριά ουρά.



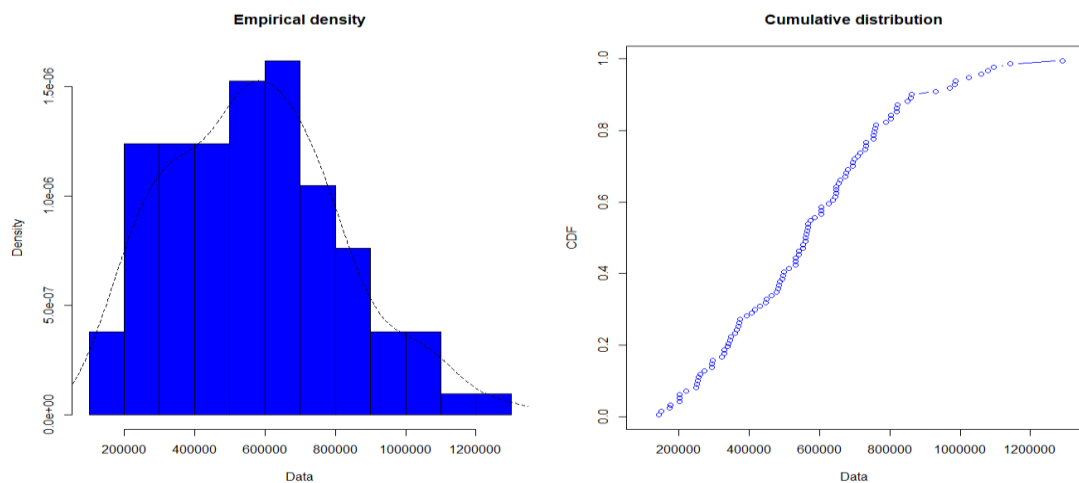
Σχήμα 4.3: *Q-Q-plot συνολικών ζημιών ανά εβδομάδα και ανά ημέρα.*

Από το Σχήμα 4.3 επιβεβαιώνεται ο παραπάνω ισχυρισμός καθώς φαίνεται ότι οι συνολικές ζημιές ανά εβδομάδα προσεγγίζουν σχετικά ικανοποιητικά την κανονική κατανομή αφού συγκλίνουν στην ευθεία με μικρές αποκλίσεις κυρίως στο αριστερό μέρος της κατανομής το οποίο από πρακτική άποψη δεν παρουσιάζει ιδιαίτερο ενδιαφέρον για την ασφαλιστική εταιρία, σε αντίθεση με τις συνολικές ζημιές ανά ημέρα που αποκλίνουν σημαντικά από την ευθεία. Σε αυτό το σημείο διενεργήθηκε έλεγχος κανονικότητας Kolmogorov-Smirnov ώστε να ελεγχθεί αν τα δεδομένα συνολικών ζημιών ανά εβδομάδα και ανά ημέρα προέρχονται από την κανονική κατανομή. Το αποτέλεσμα του p-value ήταν πολύ μικρό ($p\text{-value} < 2.2e-16$) και για τις δύο περιπτώσεις οπότε η μηδενική υπόθεση της κανονικότητας απορρίπτεται.

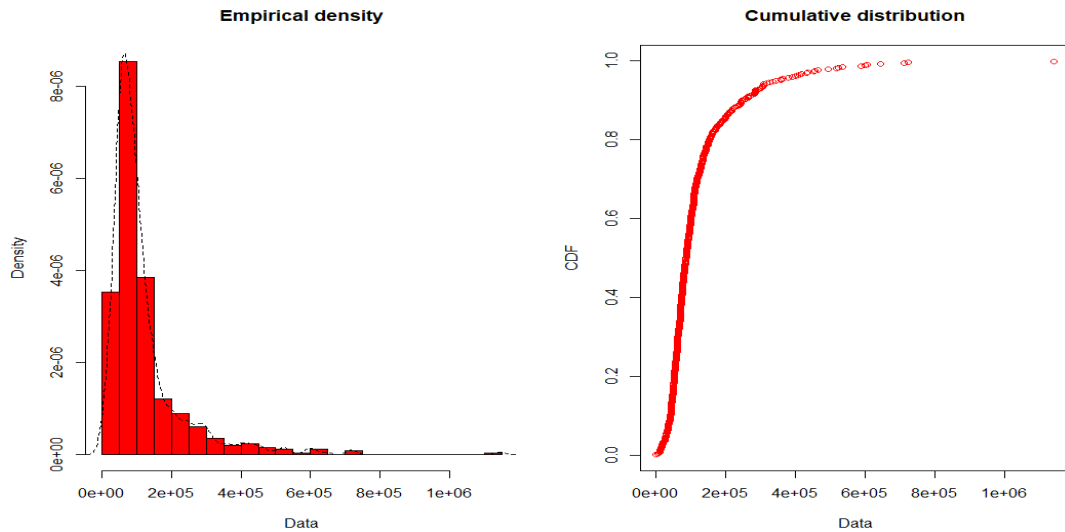
Στο Σχήμα 4.4 παρακάτω, παρουσιάζεται η διαγραμματική απεικόνιση της εμπειρικής συνάρτησης πυκνότητας και για τις δυο ομαδοποιήσεις των συνολικών ζημιών αντίστοιχα. Οι εμπειρικές συναρτήσεις πυκνότητας ανήκουν σε μια μεγάλη κατηγορία μη παραμετρικών συναρτήσεων πυκνότητας εκτιμητών, ένα πεδίο που εξετάστηκε από τον Wegman (1972). Αποτελούν μια απλή τροποποίηση των ιστογραμμάτων και επομένως εισάγονται εύκολα, και οι ιδιότητες σύγκλισής τους είναι ισοδύναμες με εκείνες της ευρύτερης κατηγορίας εκτιμητών.



Σχήμα 4.4: Διαγραμματική απεικόνιση των εμπειρικών συναρτήσεων πυκνότητας για τη συνολικές ζημιές ανά εβδομάδα και ανά ημέρα.



Σχήμα 4.5: Σχηματική απεικόνιση ιστογράμματος των συνολικών ζημιών ανά εβδομάδα με εφαρμογή της εμπειρικής συνάρτησης πυκνότητας (αριστερά) και της εμπειρικής συνάρτησης κατανομής (δεξιά).



Σχήμα 4.6: Σχηματική απεικόνιση ιστογράμματος των συνολικών ζημιών ανά ημέρα με εφαρμογή της εμπειρικής συνάρτησης πιθανότητας.

Στη συνέχεια θα γίνει προσαρμογή των δεδομένων στον μετασχηματισμό της κατανομής Weibull NHTW για να γίνει εκτίμηση των παραμέτρων α, γ, σ .

4.3 ΠΡΟΣΑΡΜΟΓΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΜΕΤΡΩΝ ΚΙΝΔΥΝΟΥ

Σε αυτό το υποκεφάλαιο γίνεται προσαρμογή των συνολικών ζημιών ανά ημέρα στην κατανομή Weibull αλλά και στον μετασχηματισμό της NHTW καθώς και υπολογισμός των μέτρων κινδύνου VaR και $TVaR$.

Η εκτίμηση των παραμέτρων γίνεται με τη βοήθεια της γλώσσας προγραμματισμού R και συγκεκριμένα με την εντολή `optim` η οποία υπολογίζει τους εκτιμητές των παραμέτρων. Στον παρακάτω πίνακα αποτυπώνονται τα αποτελέσματα:

Αποτελέσματα	NHTW	Weibull
α	0.61	0.61
σ	827.54	
γ	0.001	0.001
AIC	-11.54	-13.54
Τιμή	6,430.61	-6,430.46

Πίνακας 4: Αποτελέσματα προσαρμογής των συνολικών ζημιών ανά ημέρα στη κατανομή Weibull και NHTW.

Επίσης υπολογίστηκε ο πληροφοριακός δείκτης AIC των δεδομένων για την κανονική κατανομή ίσος με 12935.76, πολύ μεγαλύτερος σε σχέση με τη κατανομή NHTW και τη κατανομή Weibull. Γενικά, το μοντέλο με το μικρότερο AIC θεωρείται καταλληλότερο και

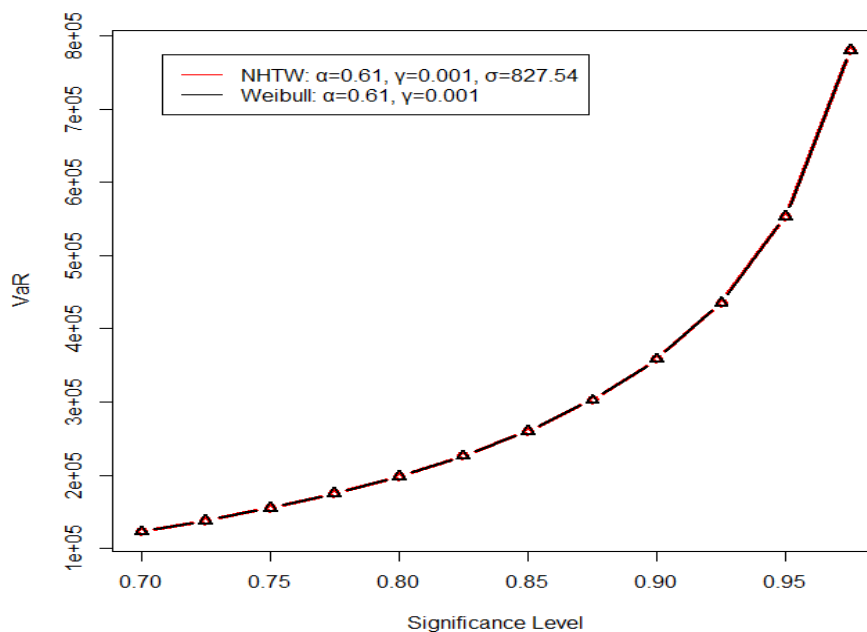
επιλέγεται. Από τα παραπάνω αποτελέσματα γίνεται αντιληπτό πως τα δεδομένα προσαρμόζονται καλύτερα στην Weibull από ότι στη NHTW αφού ο πληροφοριακός δείκτης AIC έχει μικρότερη τιμή υποδηλώνοντας την ανώτερη ποιότητα του μοντέλου. Παρόλα αυτά η διαφορά είναι μικρή και αυτό μπορεί να εξηγηθεί λαμβάνοντας υπόψιν τη τιμή της παραμέτρου σ που εκτιμήθηκε για τη κατανομή NHTW η οποία πήρε πολύ μεγάλη τιμή ($\sigma=827.54$), το οποίο δηλώνει πως η κατανομή NHTW ουσιαστικά προσεγγίζει τη κατανομή Weibull.

Στη συνέχεια θα γίνει υπολογισμός των μέτρων κινδύνου VaR και $TVaR$ για τις παραμέτρους που εκτιμήθηκαν.

Conf. level	VaR for NHTW	VaR for Weibull
0.700	122,732	122,732
0.725	137,734	137,726
0.750	154,941	154,924
0.775	174,881	174,852
0.800	198,285	198,241
0.825	226,203	226,139
0.850	260,211	260,121
0.875	302,823	302,697
0.900	358,386	358,210
0.925	435,361	435,109
0.950	553,641	553,263
0.975	780,924	780,282

Πίνακας 5: Αποτελέσματα VaR για τις Weibull και NHTW με τις παραμέτρους που εκτιμήθηκαν.

Από τον Πίνακα 5 φαίνεται πως η αξία σε κίνδυνο με βάση τη NHTW είναι περίπου ίδια με αυτή της Weibull για κάθε επίπεδο εμπιστοσύνης.



Σχήμα 4.7: Σχηματική απεικόνιση VaR για τις κατανομές NHTW και Weibull.

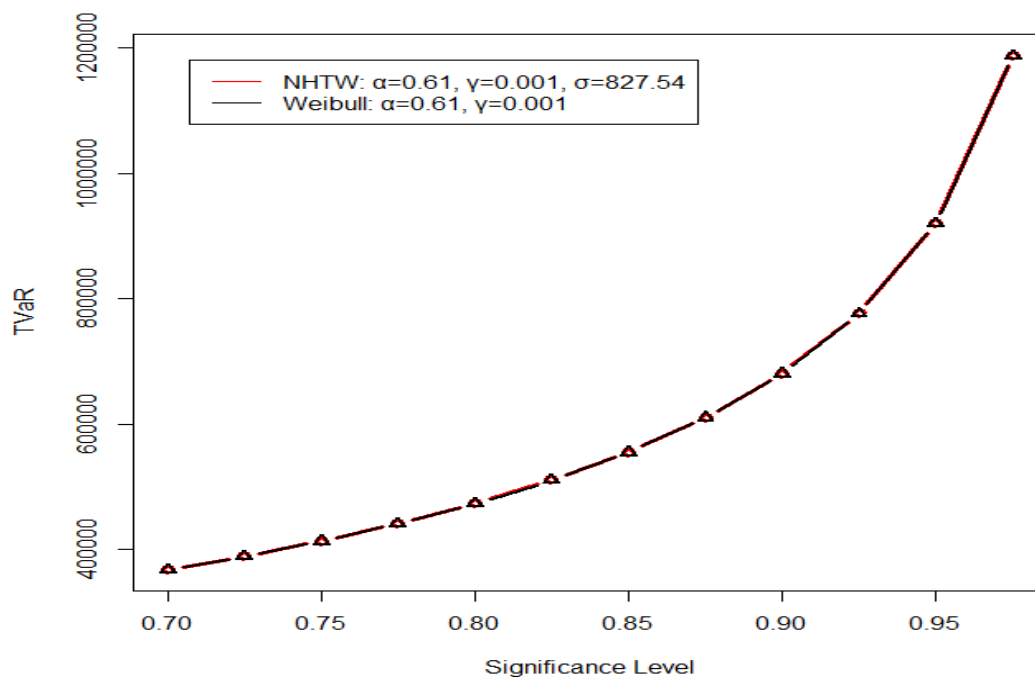
Conf. level	TVaR for NHTW	TVaR for Weibull
0.700	367,956	367,736
0.725	389,582	389,342
0.750	413,927	413,664
0.775	441,624	441,335
0.800	473,545	473,224
0.825	510,935	510,576
0.850	555,655	555,249
0.875	610,655	610,188
0.900	681,003	680,457
0.925	776,491	775,833
0.950	919,944	919,112
0.975	1,188,177	1,187,003

Πίνακας 6: Αποτελέσματα TVaR για τις Weibull και NHTW με τις παραμέτρους που εκτιμήθηκαν.

Όπως έχει αναφερθεί σε προηγούμενο κεφάλαιο, ο υπολογισμός του TVaR γίνεται με τον τύπο:

$$TVaR_q(X) = \frac{\int_{VaR_q}^{\infty} xf(x)dx}{1 - q}.$$

Από τον Πίνακα 6 παρατηρείται ότι το TVaR της NHTW είναι περίπου ίδιο με αυτό της Weibull για κάθε επίπεδο εμπιστοσύνης. Αυτό φαίνεται γραφικά και στο Σχήμα 4.8.



Σχήμα 4.8: Σχηματική απεικόνιση TVaR για τις κατανομές NHTW και Weibull.

Από τα αριθμητικά αποτελέσματα των μέτρων κινδύνου του προτεινόμενου μοντέλου NHTW και της κατανομής Weibull που παρατίθενται στους πίνακες 5 και 6, είναι σαφές ότι η προτεινόμενη κατανομή έχει βαριά ουρά και μπορεί να χρησιμοποιηθεί ως καλό υποψήφιο μοντέλο για τη μοντελοποίηση ασφαλιστικών δεδομένων. Παρουσιάζει, για μεγάλες τιμές της παραμέτρου σ , παρόμοια συμπεριφορά και ποιότητα προσαρμογής με τη Weibull έχει όμως μεγαλύτερη ευελιξία λόγω της επιπλέον παραμέτρου σ , και σε περιπτώσεις που η ουρά της Weibull είναι βαρύτερη από την ουρά των προς εξέταση δεδομένων, μπορεί να τα περιγράψει πιο ικανοποιητικά από εκείνη.

ΣΥΝΟΨΗ

Η επιλογή μιας συγκεκριμένης στατιστικής κατανομής είναι θεμελιώδης κατά την εφαρμογή ανάλυσης κινδύνου και έχει τεράστιο αντίκτυπο στα συνήθη μέτρα εκτίμησης κινδύνου. Στο αυτό το πλαίσιο, οι Zubair Ahmad, Eisa Mahmoudi & Sanku Dey (2020) πρότειναν μια ευέλικτη κατανομή τριών παραμέτρων με βαριά ουρά χρησιμοποιώντας τη νέα οικογένεια κατανομών με βαριά ουρά (NFHT) που επιτρέπει εκφράσεις κλειστού τύπου για ορισμένα μέτρα ασφάλισης και μελέτησαν ορισμένες ιδιότητες των κατανομών NFHT. Η εφαρμοσιμότητα αυτής της νέας οικογένειας κατανομής Weibull με βαριά ουρά παρουσιάστηκε με τη χρήση ενός γνωστού συνόλου δεδομένων από τον τομέα των ασφαλίσεων και το μοντέλο αποδίδει αρκετά καλά σε σύγκριση με ορισμένες δημοφιλείς κατανομές με βαριά ουρά. Αυτή η νέα εξέλιξη, η οποία έχει μια πολλά υποσχόμενη προσέγγιση για τη μοντελοποίηση δεδομένων στον αναλογιστικό τομέα, μπορεί να είναι πολύ χρήσιμη για τους επαγγελματίες που χειρίζονται μεγάλες αποζημιώσεις. Για το λόγο αυτό, μπορεί να θεωρηθεί ως εναλλακτική λύση της κατανομής Weibull.

ΠΑΡΑΡΤΗΜΑ

Για τα αριθμητικά αποτελέσματα που παρουσιάστηκαν στα προηγούμενα κεφάλαια χρησιμοποιήθηκε η γλώσσα προγραμματισμού R. Δίνονται στη συνέχεια οι εντολές που χρησιμοποιήθηκαν.

Π1. Τύποι Ασυμμετρίας

```
x<-seq(0,10,0.01);a<-1.5;g<-2;z<-dnorm(x,5,1)
fx<-a*g*x^(a-1)*exp(-g*x^a)
gx<-a*g*(10-x)^(a-1)*exp(-g*(10-x)^a)
par(mfcol=c(1,3))
plot(x,gx,type="l",lwd=3,xlim=c(6,10),main="Negatively Skewed")
plot(x,z,type="l",lwd=3,main="Normal (no skew)")
plot(x,fx,type="l",lwd=3,xlim=c(0,4),main="Positively Skewed")
```

Π2. Plot για συνάρτηση πυκνότητας NHTD-Weibull

```
par(mfcol=c(1,2));
a1<-1.5; g<-1;x<-seq(0,4,0.01)
s<-1.2;s1<-1.4;s2<-2;s3<-3
a<-1.2;a2<-2;a3<-2.5

nhtd<-(a*s*(s-1)*g*x^(a-1)*exp(-g*x^a))/(s-exp(-g*x^a))^2
nhtd1<-(a*s1*(s1-1)*g*x^(a-1)*exp(-g*x^a))/(s1-exp(-g*x^a))^2
nhtd2<-(a*s2*(s2-1)*g*x^(a-1)*exp(-g*x^a))/(s2-exp(-g*x^a))^2
nhtd3<-(a*s3*(s3-1)*g*x^(a-1)*exp(-g*x^a))/(s3-exp(-g*x^a))^2

fx<-a*g*x^(a-1)*exp(-g*x^a)
fx1<-a1*g*x^(a1-1)*exp(-g*x^a1)
fx2<-a2*g*x^(a2-1)*exp(-g*x^a2)
fx3<-a3*g*x^(a3-1)*exp(-g*x^a3)

plot(x,fx,type="l",lwd=3,col="red",ylab="g(x)",xlim=c(0,4),ylim=c(0,3),main="NHTD &
Weibull Distribution")
```

```

lines(x,nhtd,lty=2,col="blue",lwd=3)
lines(x,nhtd2,lty=3,col="green",lwd=3)
lines(x,nhtd3,lty=4,lwd=3)
legend(locator(1),c("α=1.2,γ=1","α=1.2,γ=1,σ=1.2","α=1.2,γ=1,σ=2","α=1.2,γ=1,σ=3"),lty=3,
lwd=2,col=c("red","blue","green","black"))
nhtd4<-(a1*s*(s-1)*g*x^(a1-1)*exp(-g*x^a1))/(s-exp(-g*x^a1))^2
nhtd5<-(a1*s1*(s1-1)*g*x^(a1-1)*exp(-g*x^a1))/(s1-exp(-g*x^a1))^2
nhtd6<-(a1*s2*(s2-1)*g*x^(a1-1)*exp(-g*x^a1))/(s2-exp(-g*x^a1))^2
nhtd7<-(a1*s3*(s3-1)*g*x^(a1-1)*exp(-g*x^a1))/(s3-exp(-g*x^a1))^2

plot(x,fx1,type="l",lwd=3,col="red",ylab="g(x)",xlim=c(0,4),ylim=c(0,3),main="NHTD&
Weibull Distribution")
lines(x,nhtd4,lty=2,col="blue",lwd=3)
lines(x,nhtd6,lty=3,col="green",lwd=3)
lines(x,nhtd7,lty=4,lwd=3)
legend(locator(1),c("α=1.5,γ=1","α=1.5,γ=1,σ=1.2","α=1.5,γ=1,σ=2","α=1.5,γ=1,σ=3"),lty=3,
lwd=2,col=c("red","blue","green","black"))

par(mfcol=c(1,2));
a1<-1.5; g<-1;x<-seq(0,4,0.01)
s<-1.2;s1<-1.4;s2<-2;s3<-3
a<-1.2;a2<-2;a3<-2.5

nhtd<-(a2*s*(s-1)*g*x^(a2-1)*exp(-g*x^a2))/(s-exp(-g*x^a2))^2
nhtd1<-(a2*s1*(s1-1)*g*x^(a2-1)*exp(-g*x^a2))/(s1-exp(-g*x^a2))^2
nhtd2<-(a2*s2*(s2-1)*g*x^(a2-1)*exp(-g*x^a2))/(s2-exp(-g*x^a2))^2
nhtd3<-(a2*s3*(s3-1)*g*x^(a2-1)*exp(-g*x^a2))/(s3-exp(-g*x^a2))^2

fx<-a*g*x^(a-1)*exp(-g*x^a)
fx1<-a1*g*x^(a1-1)*exp(-g*x^a1)
fx2<-a2*g*x^(a2-1)*exp(-g*x^a2)
fx3<-a3*g*x^(a3-1)*exp(-g*x^a3)

```



```

plot(x,fx2,type="l",lwd=3,col="red",ylab="g(x)",xlim=c(0,3),ylim=c(0,3),main="NHTD &
Weibull Distribution")
lines(x,nhtd,lty=2,col="blue",lwd=3)
lines(x,nhtd2,lty=3,col="green",lwd=3)
lines(x,nhtd3,lty=4,lwd=3)
legend(locator(1),c("α=2,γ=1","α=2,γ=1,σ=1.2","α=2,γ=1,σ=2","α=2,γ=1,σ=3"),lty=3,lwd=2,
col=c("red","blue","green","black"))
nhtd4<-(a3*s*(s-1)*g*x^(a3-1)*exp(-g*x^a3))/(s-exp(-g*x^a3))^2
nhtd5<-(a3*s1*(s1-1)*g*x^(a3-1)*exp(-g*x^a3))/(s1-exp(-g*x^a3))^2
nhtd6<-(a3*s2*(s2-1)*g*x^(a3-1)*exp(-g*x^a3))/(s2-exp(-g*x^a3))^2
nhtd7<-(a3*s3*(s3-1)*g*x^(a3-1)*exp(-g*x^a3))/(s3-exp(-g*x^a3))^2

```

```

plot(x,fx3,type="l",lwd=3,col="red",ylab="g(x)",xlim=c(0,3),ylim=c(0,3),main="NHTD&
Weibull Distribution")
lines(x,nhtd4,lty=2,col="blue",lwd=3)
lines(x,nhtd6,lty=3,col="green",lwd=3)
lines(x,nhtd7,lty=4,lwd=3)
legend(locator(1),c("α=2.5,γ=1","α=2.5,γ=1,σ=1.2","α=2.5,γ=1,σ=2","α=2.5,γ=1,σ=3"),lty=3,
lwd=2,col=c("red","blue","green","black"))

```

Π3. Plot για συνάρτηση κατανομής και συνάρτηση επιβίωσης

```

par(mfcol=c(1,2))
x<-seq(0,5,0.01);a<-1.5;g<-1;s<-2.5
Gnhtd<-s*(1-exp(-g*x^a))/(s-exp(-g*x^a))
plot(x,Gnhtd,type="l",lwd=3,col="red",ylab="G(x)",main="CDF of NHTD-Weibull")
legend(locator(1),c("α=1.5,γ=1,σ=2.5"),col="red",lwd=3)
SF<-1-Gnhtd
plot(x,SF,type="l",col="blue",ylab="Sf(x)",main="Survival function of NHTD-Weibull")
legend(locator(1),c("α=1.5,γ=1,σ=2.5"),col="blue",lwd=3)

```

Π4. Plot για συνάρτηση κινδύνου hrf

```

a<-1.2;a1<-0.1;a2<-1.8;g<-1;s<-1.7;s1<-2.8;s2<-1.2
x<-seq(0,2,0.01)
hrf<-((s*(s-1)*a*g*(x^(a-1))*exp(-g*x^a))/(exp(-g*x^a)*(s-exp(-g*x^a))))
hrf1<-((s1*(s1-1)*a1*g*(x^(a1-1))*exp(-g*x^a1))/(exp(-g*x^a1)*(s1-exp(-g*x^a1))))
hrf2<-((s2*(s2-1)*a2*g*(x^(a2-1))*exp(-g*x^a2))/(exp(-g*x^a2)*(s2-exp(-g*x^a2))))
plot(x,hrf,type="l",lwd=3,col="red",ylab="h(x)",xlim=c(0,2),ylim=c(0,2),main="hazard rate
function")
lines(x,hrf1,lty=3,col="blue",lwd=3)
lines(x,hrf2,lty=3,col="green",lwd=3)
legend(locator(1),c("α=1.2,γ=1,σ=1.7","α=0.1,γ=1,σ=2.8","α=1.8
,γ=1,σ=1.2"),lty=3,lwd=2,col=c("red","blue","green","black"))

```

Π5.

```

set.seed(0)
x<-rweibull(1000,0.9,1.3)
y<-rweibull(1000,1.8,1.3)
z<-rweibull(1000,2.7,1.3)
k<-rweibull(1000,3.5,1.3)
par(mfcol=c(1,2))
qqnorm(x);qqline(x);legend(locator(1),c("α=0.9","γ=1.3"))
qqnorm(y);qqline(y);legend(locator(1),c("α=1.8","γ=1.3"))
qqnorm(z);qqline(z);legend(locator(1),c("α=2.7","γ=1.3"))
qqnorm(k);qqline(k);legend(locator(1),c("α=3.5","γ=1.3"))

```

Π6. Ιστόγραμμα συχνοτήτων για προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.

```

u<-runif(10000,min=0,max=1)
s<-2.1;g<-0.1;a<-1.2
x<-((-log(s*(u-1)/(u-s)))/g)^(1/a)
hist(x,breaks=seq(0,50,2),col="red")
legend(locator(1),c("α=1.2,γ=0.1,σ=2.1"))

```

Π7. Θηκόγραμμα για τα προσομοιωμένα δεδομένα με τη μέθοδο της αντιστροφής.

```
u<-runif(10000,min=0,max=1)
s<-2.1;g<-0.1;a<-1.2
x<-(-log(s*(u-1)/(u-s))/g)^(1/a)
boxplot(x,col="red")
```

Π8.

```
library(rootSolve)
library(AdequacyModel)
rNHTWeibull=function(par,n){
a=par[1];s=par[2];g=par[3]
u=runif(n)
x=c()
for(i in 1:n){
f=function(x) (1-exp(-g*x^a))*(s-u[i]*(s-1+(1-exp(-g*x^a))))
x[i]=rootSolve::uniroot.all(f, interval=c(0,100000))
}
return(x)
}
dNHT_Weibull<-function(par,x){
a=par[1]
s=par[2]
g=par[3]
(a*g*s*(s-1)*(x^(a-1))*exp(-g*x^a)/((s-exp(-g*x^a))^2))
}
pNHT_Weibull<-function(par,x){
a=par[1]
s=par[2]
g=par[3]
(s*(1-exp(-g*x^a))/(s-exp(-g*x^a)))
}
a<-0.9;s<-1.2;g<-0.5
```

```

llnhtw<-function(par) {
a=par[1]
s=par[2]
g=par[3]
aux=(s-exp(-g*x^a))
if(a>0&&g>0&&min(x)>0&&min(aux)>0)
{
ll=log(a)+log(s)+log(g)+log(s-1)+(a-1)*log(x)-(g*x^a)-2*log(aux)
return(sum(ll))
}
else {
return(-9999999.99)
}
}
par<-c(0.9,1.2,0.5);a<-0.9;s<-1.2;g<-0.5
numrep<-1000
matri_par<-matrix(0,40,3)
matri_bias<-matrix(0,40,3)
matri_MSE<-matrix(0,40,3)
matri_std<-matrix(0,40,3)
colnames(matri_par)<-c("a","s","g")
colnames(matri_bias)<-c("a","s","g")
colnames(matri_MSE)<-c("a","s","g")
colnames(matri_std)<-c("a","s","g")
cont=1;n=25
while(n<=1000) {
par_mean<-c(0,0,0)
std_mean<-c(0,0,0)
bias<-c(0,0,0)
MSE<-c(0,0,0)
replica<-1
while(replica<=numrep){

```

```

print(paste("n= ",n,"replica= ",replica))
x<-rNHTWeibull(par,n)
Data<-x

```

Π9. `###simulation results###`

```

result<-optim(c(a, s, g), llhwtw, hessian=F,control=list(fnscale=-1),
method="L-BFGS-B", lower=c(0.001,0.001,0.001), upper=c(5,5,5))
if (class(result)!= "try-error" && result$convergence==0){
par_mean<-par_mean+result$par
bias=bias+(result$par-par)
MSE=MSE+(result$par-par)^2
replica<-replica+1
}
}
par_mean=par_mean/numrep
bias<-bias/numrep
MSE=MSE/numrep
matri_par[cont,]=par_mean
matri_std[cont,]=std_mean
matri_bias[cont,]=bias
matri_MSE[cont,]=MSE
print("mean=")
print(par_mean)
print("bias=")
print(bias)
print("MSE=")
print(MSE)
n<-n+25
cont=cont+1
}
print(matri_par)
print(matri_MSE)

```

```
print(matri_bias)
n<-seq(25,1000,25)
```

```
Π10. ###plot of Estimated parameters###
```

```
plot(n,(matri_par[,1]),type="o",col="green",lty=1,lwd=2,xlab="n",
ylab="Estimated Parameters",ylim=c(0,3))
```

```
lines(n,(matri_par[,2]),col="blue",lty=5,lwd=2,type="o")
```

```
lines(n,(matri_par[,3]),col="red",lty=5,lwd=2,type="o")
```

```
abline(h=0.9,col="green")
```

```
abline(h=1.2,col="blue")
```

```
abline(h=0.5,col="red")
```

```
title("Plot of Estimated Parameters vs n")
```

```
legend(700,2.8,legend=c("α=0.9","σ=1.2","γ=0.5"),lty=c(1,5,8),cex=1,col=c("green","blue","red"))
```

```
Π11. ###plot of MSEs###
```

```
plot(n,matri_MSE[,1],col="green",lty=1,lwd=2,type="o",xlab="n",ylab="MSE",ylim=c(0,3))
```

```
lines(n,matri_MSE[,2],col="blue",lty=5,lwd=2,type="o")
```

```
lines(n,matri_MSE[,3],col="red",lty=8,lwd=2,type="o")
```

```
title("Plot of MSE vs n")
```

```
legend(700,5.5,legend=c("α=0.9","σ=1.2","γ=0.5"),lty=c(1,5,8),cex=1,col=c("green","blue","red"))
```

```
###plot of absolute bias###
```

```
plot(n,abs(matri_bias[,1]),type="o",col="green",lty=1,lwd=2,xlab="n",ylab="Absolute Bias",ylim=c(0,1))
```

```
lines(n,abs(matri_bias[,2]),type="o",col="blue",lty=2,lwd=2)
```

```
lines(n,abs(matri_bias[,3]),type="o",col="red",lty=3,lwd=2)
```

```
title("Plot of Absolute Bias vs n")
```

```
legend(700,1.4,legend=c("α=0.9","σ=1.2","γ=0.5"),lty=c(1,5,8),cex=1,col=c("green","blue","red"))
```

Π12. ###plot of bias###

```
plot(n,(matri_bias[,1]),type="o",col="green",lty=1,lwd=2,xlab="n",ylab="Bias",ylim=c(0,1))
lines(n,(matri_bias[,2]),type="o",col="blue",lty=5,lwd=2)
lines(n,(matri_bias[,3]),type="o",col="red",lty=8,lwd=2)
title("Plot of Bias vs n")
legend(700,1.4,legend=c("α=0.9","σ=1.2","γ=0.5"),lty=c(1,5,8),cex=1,col=c("green","blue","red"))
```

Π13. Σχηματική απεικόνιση VaR για τη κατανομή NHTW με διαφορετικές παραμέτρους

```
a<-0.9;g<-0.5;cl<-seq(0.7,0.975,0.025);s1<-1.2;s2<-2.5
```

```
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a)
```

```
htwVaR2<-(-log(s2*(1-cl)/(s2-cl))/g)^(1/a)
```

```
plot(cl,htwVaR1,ylab="VaR","b",xlab="Significance Level",ylim=c(0,10),lwd=3,col="red",pch=1)
```

```
lines(cl,htwVaR2,"b",lwd=2.5,col="black",pch=2)
```

```
legend(locator(1),c("NHTW: α=0.9, γ=0.5, σ=1.2","NHTW: α=0.9, γ=0.5, σ=2.5"),col=c("red","black"),lty=c(1,1))
```

```
a<-0.9;g1<-0.5;g2<-2.5;cl<-seq(0.7,0.975,0.025);s1<-1.2
```

```
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g1)^(1/a)
```

```
htwVaR2<-(-log(s1*(1-cl)/(s1-cl))/g2)^(1/a)
```

```
plot(cl,htwVaR1,ylab="VaR","b",xlab="Significance Level",ylim=c(0,10),lwd=3,col="red",pch=1)
```

```
lines(cl,htwVaR2,"b",lwd=2.5,col="black",pch=2)
```

```
legend(locator(1),c("NHTW: α=0.9, γ=0.5, σ=1.2","NHTW: α=0.9, γ=2.5, σ=1.2"),col=c("red","black"),lty=c(1,1))
```

```
a1<-0.9;a2<-2;g1<-0.5;cl<-seq(0.7,0.975,0.025);s1<-1.2
```

```
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g1)^(1/a1)
```

```
htwVaR2<-(-log(s1*(1-cl)/(s1-cl))/g1)^(1/a2)
```

```
plot(cl,htwVaR1,ylab="VaR","b",xlab="Significance Level",ylim=c(0,10),lwd=3,col="red",pch=1)
```

```
lines(cl,htwVaR2,"b",lwd=2.5,col="black",pch=2)
```

```
legend(locator(1),c("NHTW:    α=0.9,    γ=0.5,    σ=1.2","NHTW:    α=2,    γ=0.5,
σ=1.2"),col=c("red","black"),lty=c(1,1))
```

Π14. Σύγκριση VaR των κατανομών NHTW και Weibull

```
a<-0.9;g<-0.5;cl<-seq(0.7,1,0.025);s1<-1.2;s2<-2.5
```

```
par(mfcol=c(1,2))
```

```
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a)
```

```
htwVaR2<-(-log(s2*(1-cl)/(s2-cl))/g)^(1/a)
```

```
wVaR<-(-log(1-cl)/g)^(1/a)
```

```
plot(cl,htwVaR1,ylab="VaR","b",xlab="Significance
```

```
Level",ylim=c(0,12),lwd=3,col="red",pch=1)
```

```
lines(cl,wVaR,"b",lwd=2.5,col="black",pch=2)
```

```
legend(locator(1),c("NHTW:    α=0.9,    γ=0.5,    σ=1.2","Weibull:    α=0.9,
γ=0.5"),col=c("red","black"),lty=c(1,1))
```

```
plot(cl,htwVaR2,ylab="VaR","b",xlab="Significance
```

```
Level",ylim=c(0,12),lwd=3,col="red",pch=1)
```

```
lines(cl,wVaR,"b",lwd=2.5,col="black",pch=2)
```

```
legend(locator(1),c("NHTW:    α=0.9,    γ=0.5,    σ=2.5","Weibull:    α=0.9,
γ=0.5"),col=c("red","black"),lty=c(1,1))
```

Π15. Σχηματική απεικόνιση TVaR για τη κατανομή NHTW

```
cl<-0.7;j<-1;a<-0.9;g<-0.5;s1<-1.2;s2<-2.5
```

```
vec<-vector(mode="numeric",length=12)
```

```
mec<-vector(mode="numeric",length=12)
```

```
while(cl<1){
```

```
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a)
```

```
htwVaR2<-(-log(s2*(1-cl)/(s2-cl))/g)^(1/a)
```

```
fx1<-function(x)(x*((a*s1*(s1-1)*g*x^(a-1)*exp(-g*x^a))/(s1-exp(-g*x^a))^2)/(1-cl))
```

```
fx2<-function(x)(x*((a*s2*(s2-1)*g*x^(a-1)*exp(-g*x^a))/(s2-exp(-g*x^a))^2)/(1-cl))
```

```
vec[j]<-integrate(fx1,lower=htwVaR1,upper=Inf)
```

```
mec[j]<-integrate(fx2,lower=htwVaR2,upper=Inf)
```

```
cl<-cl+0.025
```



```

j<-j+1
}
m<-as.vector(vec)
d<-as.vector(mec)
cl<-seq(0.7,0.975,0.025)
plot(cl,m,type="b",ylab="TVaR",xlab="Significance Level",col="red",lwd=3,ylim=c(0,15))
lines(cl,d,type="b",col="black",lwd=2.5,pch=2)
legend(locator(1),c("NHTW:   α=0.9,   γ=0.5,   σ=1.2","NHTW:   α=0.9,   γ=0.5,
σ=2.5"),col=c("red","black"),lty=c(1,1))
cl<-0.7;j<-1;a1<-0.9;a2<-2;g<-0.5;s1<-1.2
vec<-vector(mode="numeric",length=12)
mec<-vector(mode="numeric",length=12)
while(cl<1){
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a1)
htwVaR2<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a2)
fx1<-function(x)(x*((a1*s1*(s1-1)*g*x^(a1-1)*exp(-g*x^a1))/(s1-exp(-g*x^a1))^2)/(1-cl))
fx2<-function(x)(x*((a2*s1*(s1-1)*g*x^(a2-1)*exp(-g*x^a2))/(s1-exp(-g*x^a2))^2)/(1-cl))
vec[j]<-integrate(fx1,lower=htwVaR1,upper=Inf)
mec[j]<-integrate(fx2,lower=htwVaR2,upper=Inf)
cl<-cl+0.025
j<-j+1
}
m<-as.vector(vec)
d<-as.vector(mec)
cl<-seq(0.7,0.975,0.025)
plot(cl,m,type="b",ylab="TVaR",xlab="Significance Level",col="red",lwd=3,ylim=c(0,15))
lines(cl,d,type="b",col="black",lwd=2.5,pch=2)
legend(locator(1),c("NHTW:   α=0.9,   γ=0.5,   σ=1.2","NHTW:   α=2,   γ=0.5,
σ=1.2"),col=c("red","black"),lty=c(1,1))
cl<-0.7;j<-1;a1<-0.9;g1<-0.5;g2<-2.5;s1<-1.2
vec<-vector(mode="numeric",length=12)
mec<-vector(mode="numeric",length=12)

```

```

while(cl<1){
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g1)^(1/a1)
htwVaR2<-(-log(s1*(1-cl)/(s1-cl))/g2)^(1/a1)
fx1<-function(x)(x*((a1*s1*(s1-1)*g1*x^(a1-1)*exp(-g1*x^a1))/(s1-exp(-g1*x^a1))^2)/(1-cl))
fx2<-function(x)(x*((a1*s1*(s1-1)*g2*x^(a1-1)*exp(-g2*x^a1))/(s1-exp(-g2*x^a1))^2)/(1-cl))
vec[j]<-integrate(fx1,lower=htwVaR1,upper=Inf)
mec[j]<-integrate(fx2,lower=htwVaR2,upper=Inf)
cl<-cl+0.025
j<-j+1
}
m<-as.vector(vec)
d<-as.vector(mec)
cl<-seq(0.7,0.975,0.025)
plot(cl,m,type="b",ylab="TVaR",xlab="Significance Level",col="red",lwd=3,ylim=c(0,15))
lines(cl,d,type="b",col="black",lwd=2.5,pch=2)
legend(locator(1),c("NHTW:   α=0.9,   γ=0.5,   σ=1.2","NHTW:   α=0.9,   γ=2.5,
σ=1.2"),col=c("red","black"),lty=c(1,1))

```

Π16. Σύγκριση TVaR των κατανομών NHTW και Weibull

```

cl<-0.7;j<-1;a<-0.9;g<-0.5;s1<-2.5;j<-1
vec<-vector(mode="numeric",length=12)
mec<-vector(mode="numeric",length=12)
while(cl<1){
htwVaR1<-(-log(s1*(1-cl)/(s1-cl))/g)^(1/a)
wVaR<-(-log(1-cl)/g)^(1/a)
fx1<-function(x)(x*((a*s1*(s1-1)*g*x^(a-1)*exp(-g*x^a))/(s1-exp(-g*x^a))^2)/(1-cl))
fx2<-function(x)(x*(a*g*x^(a-1)*exp(-g*x^a))/(1-cl))
vec[j]<-integrate(fx1,lower=htwVaR1,upper=Inf)

```

```

mec[j]<-integrate(fx2,lower=wVaR,upper=Inf)
cl<-cl+0.025
j<-j+1
}
m<-as.vector(vec)
d<-as.vector(mec)
cl<-seq(0.7,0.975,0.025)
plot(cl,m,type="b",ylab="TVaR",xlab="Significance Level",col="red",lwd=3)
lines(cl,d,type="b",col="black",lwd=2.5,pch=2)
legend(locator(1),c("NHTW:       $\alpha=0.9,$        $\gamma=0.5,$        $\sigma=2.5$ ", "Weibull:       $\alpha=0.9,$ 
 $\gamma=0.5$ "),col=c("red", "black"),lty=c(1,1))

```

Π17. Ανάλυση δεδομένων

```

install.packages("moments")
install.packages("stringr")
library(moments)
library(stringr)
library("EnvStats")
library(fitdistrplus)
library(stats4)
library(MASS)
library(survival)
library(actuar)
library(distrMod)
weekly_losses<-read.delim(file.choose())
weekly_losses[2]<-as.numeric(weekly_losses$Incurred.Losses)
sum(weekly_losses[2])

```

```

summary(weekly_losses)
var(weekly_losses$Incurred.Losses)
sd(weekly_losses$Incurred.Losses)
skewness(weekly_losses$Incurred.Losses)
kurtosis(weekly_losses$Incurred.Losses)

losses<-read.delim(file.choose())
losses[2]<-str_replace(losses$Incurred.Losses,",",".")
losses[2]<-as.numeric(losses$Incurred.Losses)
summary(losses)
var(losses$Incurred.Losses)
sd(losses$Incurred.Losses)
skewness(losses$Incurred.Losses)
kurtosis(losses$Incurred.Losses)
sum(losses[2])
par(mfcol=c(1,2))
boxplot(weekly_losses$Incurred.Losses,col="blue",main="Θηκόγραμμα για συνολικές ζημιές
ανά εβδομάδα")
boxplot(losses$Incurred.Losses,col="red",main="Θηκόγραμμα για συνολικές ζημιές ανά
ημέρα")

hist(weekly_losses$Incurred.Losses,breaks=15,col="blue",main="Ιστόγραμμα  συχνοτήτων
συνολικών ζημιών ανά εβδομάδα")
hist(losses$Incurred.Losses,breaks=seq(0,1300000,50000),col="red",main="Ιστόγραμμα
συχνοτήτων συνολικών ζημιών ανά ημέρα")
qqnorm(weekly_losses$Incurred.Losses,col="blue",main="Συνολικές  ζημιές  ανά
εβδομάδα");qqline(weekly_losses$Incurred.Losses)

```

```

qqnorm(losses$Incurred.Losses,col="red",main="Συνολικές          ζημιές          ανά
ημέρα");qqline(losses$Incurred.Losses)

install.packages("EnvStats")

library("EnvStats")

obs1=weekly_losses$Incurred.Losses

obs2=losses$Incurred.Losses

par(mfcol=c(1,2))

epdfPlot(obs1,epdf.col="blue",main="Εμπειρική συνάρτηση πιθανότητας για συνολικές
ζημιές ανά εβδομάδα")

epdfPlot(obs2,epdf.col="red",main="Εμπειρική συνάρτηση πιθανότητας για συνολικές ζημιές
ανά ημέρα")

plotdist(weekly_losses$Incurred.Losses,histo=T,breaks=12,demp=T,col="blue",type="b")

plotdist(losses$Incurred.Losses,histo=T,breaks=20,demp=T,col="red")

```

Π18. Προσαρμογή των συνολικών ζημιών ανά ημέρα και εκτίμηση των παραμέτρων για τις NHTW και Weibull.

```

library(fitdistrplus)
library(stringr)
x <- read.delim(file.choose())
x<-str_replace(x$Incurred.Losses,",",".")
x<-as.numeric(x)

llnhtw<-function(par) {
a=par[1]
s=par[2]
g=par[3]
aux=(s-exp(-g*x^a))
if(a>0&&s>1&&g>0&&min(x)>0&&min(aux)>0)
{
ll=log(a)+log(s)+log(g)+log(s-1)+(a-1)*log(x)-(g*x^a)-2*log(aux)
return(sum(ll))
}
}

```

```

}
else {
return(-9999999.99)
}
}
a<-0.1;s<-1;g<-0.01
result<-optim(c(a, s, g), llnhtw, hessian=F,control=list(fnscale=-1),
method="L-BFGS-B", lower=c(0.001,1,0.001), upper=c(Inf,Inf,Inf))
result
set.seed(42)
u<-runif(494)
s<-result$par[2];g<-result$par[3];a<-result$par[1]
y<-(-log(s*(u-1)/(u-s)))/g^(1/a)
hist(y,col="red",breaks=25)
legend(locator(1),c("α=0.61,γ=0.001,σ=827.54"))
summary(y)
summary(x)
var(y);sd(y)
plot(y)
par<-c(result$par[1],result$par[2],result$par[3])
llnhtw(par)
k1<-length(par)
AIC1<-(-2*log(-llnhtw(par))+2*k1)
AIC1

llw<-function(par) {
a1=par[1]
g1=par[2]
if(a1>0&&g1>0&&min(x)>0)
{
ll=log(a1)+log(g1)+(a1-1)*log(x)-(g1*x^a1)
return(sum(ll))
}
}

```

```

}
else {
return(-9999999.99)
}
}
a1<-0.1;g1<-0.01
result1<-optim(c(a1,g1), llw, hessian=F,control=list(fnscale=-1),
method="L-BFGS-B", lower=c(0.001,0.001), upper=c(Inf,Inf))
result1
par1<-c(result1$par[1],result1$par[2])
llw(par1)
k2<-length(par1)
AIC2<-(-2*log(-llw(par1))+2*k2)
AIC2
library(MASS)
norm<-fitdistr(x,list("normal"))
AIC(norm)

a<-result$par[1];s<-result$par[2];g<-result$par[3];cl<-seq(0.7,0.975,0.025)
aa<-result1$par[1];gg<-result1$par[2]
htwVaR1<-(-log(s*(1-cl)/(s-cl))/g)^(1/a)
wVaR<-(-log(1-cl)/gg)^(1/aa)
plot(cl,htwVaR1,ylab="VaR", "b",xlab="Significance Level",lwd=3,col="red",pch=1)
lines(cl,wVaR,"b",lwd=2.5,col="black",pch=2)
legend(locator(1),c("NHTW:    α=0.61,    γ=0.001,    σ=827.54","Weibull:    α=0.09,
γ=0.369"),col=c("red","black"),lty=c(1,1))

vec<-vector(mode="numeric",length=12)
mec<-vector(mode="numeric",length=12)
conlev<-0.7;j<-1
while(conlev<1){
htwVaR1<-(-log(s*(1-conlev)/(s-conlev))/g)^(1/a)

```

```

wVaR<-(-log(1-conlev)/gg)^(1/aa)
fx1<-function(x)(x*((a*s*(s-1)*g*x^(a-1)*exp(-g*x^a))/(s-exp(-g*x^a))^2))
fx2<-function(x)(x*(aa*gg*x^(aa-1)*exp(-gg*x^aa)))
vec[j]<-integrate(fx1,lower=htwVaR1,upper=10^10)[[1]]/(1-conlev)
mec[j]<-integrate(fx2,lower=wVaR,upper=10^10)[[1]]/(1-conlev)
conlev<-conlev+0.025
j<-j+1
}
m<-as.vector(vec)
d<-as.vector(mec)
cl<-seq(0.7,0.975,0.025)
plot(cl,m,type="b",ylab="TVaR",xlab="Significance Level",col="red",lwd=3)
lines(cl,d,type="b",col="black",lwd=2.5,pch=2)
legend(locator(1),c("NHTW:    α=0.61,    γ=0.001,    σ=827.54","Weibull:    α=0.61,
γ=0.001"),col=c("red","black"),lty=c(1,1))

```


BIBΛΙΟΓΡΑΦΙΑ

1. Zubair Ahmad, Eisa Mahmoudi & Sanku Dey (2020). A new family of heavy tailed distributions with an application to the heavy tailed insurance loss data, *Communication in Statistics: Simulation and Computation*, **51**, 2-24.
2. Eisa Mahmoudi, Zubair Ahmad & Gholamhossein Hamedani (2020). New Methods to Define Heavy-Tailed Distributions with Applications to Insurance Data, *Journal of Taibah University for Science*, **14**, 10-15.
3. Maurice C. Bryson (1974). Heavy-Tailed Distributions: Properties and Tests, *Technometrics*, **16**, 61-62.
4. Jin Zhao, Zubair Ahmad , Eisa Mahmoudi , E. H. Hafez, and Marwa M. Mohie El-Din (2021). A New Class of Heavy-Tailed Distributions: Modeling and Simulating Actuarial Measures, *Complexity*, **2021**, 16-18.
5. Yen Liang Tung, Zubair Ahmad, and G. G. Hamedani (2021). Modeling the Medical Care Insurance Data via a New Statistical Model, *Computers, Materials & Continua*, **66**, 113-126.
6. Zubair Ahmad, Eisa Mahmoudi, Morad Alizadeh , Rasool Roozegar , and Ahmed Z. Afify (2021). The Exponential T-X Family of Distributions: Properties and an Application to Insurance Data, *Journal of Mathematics*, **2021**, 1-11.
7. Muhammad Arif, Dost Muhammad Khan, Saima Khan Khosa, Muhammad Aamir, Adnan Aslam, Zubair Ahmad and Wei Gao (2021). Modelling Insurance Losses with a New Family of Heavy-Tailed Distributions, *Computers, Materials & Continua*, **66**, 537-550.
8. Μιχαήλ Μπούτσικας (2002). Σημειώσεις Παραδόσεων: «Μέθοδοι Προσομοίωσης και Υπολογιστικές Στατιστικές Τεχνικές», *Πανεπιστήμιο Πειραιώς*.
9. M.S. Waterman and D.E. Whiteman (1978). Estimation of probability densities by empirical density functions. *Int. J. Math. Educ. Sci. Technol.*, **9**, 127-137.