

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ
ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

**Λοξές Κατανομές στα Χρηματοοικονομικά και στην Αναλογιστική
Επιστήμη**

ΜΙΧΑΗΛ ΓΚΙΟΚΑΣ

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

Πειραιάς

Δεκέμβριος 2020

UNIVERSITY OF PIRAEUS



DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

POSTGRADUATE PROGRAM IN ACTUARIAL SCIENCE AND RISK
MANAGEMENT

Skewed distributions in finance and actuarial science

MICHAEL GKIOKAS

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

Piraeus,

December 2020

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. 12η/24-6-2019 συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Αντζουλάκος Δημήτριος (Επιβλέπων)
- Πολίτης Κωνσταντίνος
- Ψαρράκος Γεώργιος

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

*Στους γονείς μου,
Δημήτρη και Σωτηρία*

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω ιδιαιτέρως τον κύριο Αντζουλάκο για τη στήριξη και τη καθοδήγηση του κατά τη διάρκεια της εκπόνησης της παρούσας διπλωματικής εργασίας, όπως επίσης και τους επιβλέποντες καθηγητές. Τέλος, θα ήθελα να ευχαριστήσω όλους τους διδάσκοντες είτε μέλη ΔΕΠ είτε εξωτερικούς, για την πολύτιμη γνώση που μας μετέδωσαν καθ' όλη τη διάρκεια του μεταπτυχιακού προγράμματος.

ΠΕΡΙΛΗΨΗ

Στην διπλωματική εργασία που θα ακολουθήσει, εξετάζουμε τις λοξές κατανομές, κατανομές των οποίων η μια ουρά είναι πιο φαρδιά από την άλλη κάτι που συνηθίζεται να παρατηρείται στις μελέτες χαρτοφυλακίων. Έχοντας ως βάση τη μελέτη του Adelchi Azzalini και της Antonella Capitanio (The Skew-Normal and Related Families (2014)) περί των λοξών κατανομών θα μελετήσουμε κατανομές πιθανότητας οι οποίες μπορούν να προβλέψουν μικρής πιθανότητας κινδύνους αλλά με υψηλές ζημιές. Συγκεκριμένα, στο πρώτο κεφάλαιο γίνεται γνωριμία με τις βασικές γνωστές κατανομές πιθανότητας όπου δίνουμε τη συνάρτηση πιθανότητάς τους, βασικές ιδιότητες, μέση τιμή, διακύμανση και διάφορες γραφικές παραστάσεις. Στο ίδιο κεφάλαιο υπάρχει εκτενής αναφορά στη βασικότερη ίσως στατιστική μέθοδο εκτίμησης παραμέτρων, τη μέθοδο μέγιστης πιθανοφάνειας καθώς επίσης γίνεται εισαγωγή στη λοξότητα μέσω κάποιων σημαντικών ορισμών και προτάσεων. Στο δεύτερο κεφάλαιο γίνεται αναλυτική μελέτη και ανάλυση των λοξών κατανομών όπου για κάθε μια δίνεται η συνάρτηση πυκνότητας πιθανότητας, η συνάρτησης κατανομής, βασικά στατιστικά μέτρα όπως μέση τιμή, διακύμανση, λοξότητα, κύρτωση, σημαντικές ιδιότητες που τις διέπουν· διάφορα διαγράμματα αλλά και τύποι εκτίμησης των παραμέτρων της κάθε κατανομής. Οι λοξές κατανομές που θα μελετηθούν είναι η λοξή κανονική, λοξή Students-t, Βήτα λοξή t, λοξή λογιστική και λοξή Cauchy. Τέλος, στο τρίτο κεφάλαιο με τη βοήθεια της γλώσσας προγραμματισμού R θα γίνει εφαρμογή των λοξών κατανομών και σύγκριση τους με άλλες γνωστές κατανομές πάνω σε πραγματικά δεδομένα προκειμένου να μπορέσουμε να εξάγουμε συμπεράσματα.

ABSTRACT

In this diploma thesis that will follow, we look at the skew distributions, distributions of which one tail is wider, something that is usually observed in portfolio analysis. Based on the study of Adelchi Azzalini and Antonella Capitanio (The Skew-Normal and Related Families (2014)) on skew distributions, we will study probability distributions which can predict low probability risks but with high losses. Specifically, in the first chapter we get acquainted with the basic known probability distributions by analyzing the probability function, basic properties, mean value, variance and we present the necessary diagrams. In the same chapter there is an extensive reference of the most important statistical method of parameter estimation, the method of maximum likelihood, as well as an introduction to skewness through some important definitions and propositions. In the second chapter, there is a detailed study and analysis of skewed distributions in terms of probability density function, distribution function, basic descriptive measures such as mean, variance, skewness and kurtosis, important properties that govern them, diagrams and standard functions to estimate the parameters of each skew distribution . The skew distributions to be studied are the Skew Normal, Skew Students-t, Beta Skew t, Skew Logistic, and Skew Cauchy. Lastly, in the third chapter with the help of the programming language R the skew distributions will be applied and compared with other known distributions on real data to be able to draw conclusions.

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	i
Abstract.....	ii
Κατάλογος Πινάκων.....	iv
Κατάλογος Σχημάτων.....	vi
Εισαγωγή	1
1.1 Κανονική Κατανομή.....	1
1.2 Λογαριθμοκανονική Κατανομή.....	3
1.3 Κατανομή t	4
1.4 Κατανομή Βήτα.....	6
1.5 Λογιστική Κατανομή	7
1.6 Κατανομή Cauchy.....	8
1.7 Κατανομή Weibull.....	10
1.8 Μέθοδος Μέγιστης Πιθανοφάνειας	11
1.9 Λοξές συμμετρικές κατανομές	12
Λοξές κατανομές	16
2.1 Γέννηση και ιδιότητες της λοξής κανονικής κατανομής	16
2.2 Γέννηση και ιδιότητες της λοξής t κατανομής.....	23
2.3 Γέννηση και ιδιότητες της Βήτα λοξής t κατανομής	28
2.4 Γέννηση και ιδιότητες της λοξής λογιστικής κατανομής	32
2.5 Γέννηση και ιδιότητες της λοξής Cauchy κατανομής.....	35
Εφαρμογές σε πραγματικά δεδομένα	38
3.1 Εισαγωγή	38
3.2 Τα δεδομένα	38
3.3 Ανάλυση των δεδομένων.....	39
3.4 Συμπεράσματα.....	49
ΠΑΡΑΡΤΗΜΑ	50
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	70

Κατάλογος Πινάκων

Πίνακας 3.1. Περιγραφικά μέτρα των δεδομένων.....	39
Πίνακας 3.2. Περιγραφικά μέτρα των λογαριθμικών δεδομένων.	41
Πίνακας 3.3. Εκτιμητές μεγίστης πιθανοφάνειας λοξής κανονικής και t κατανομής.....	43
Πίνακας 3.4. Εκτιμητές μεγίστης πιθανοφάνειας για 5 επιλεγμένες κατανομές.....	44
Πίνακας 3.5. Αποτελέσματα Kolmogorov - Smirnov τεστ.....	45
Πίνακας 3.6. Τιμές AIC για τα δεδομένα Danish και Expert Advisor	46
Πίνακας 3.7. Τιμές AIC για τα δεδομένα Log Danish και Log Expert Advisor.....	46
Πίνακας 3.8. VaR0.99 και TVaR0.99 για τα δεδομένα Danish και Expert Advisor	48
Πίνακας 3.9. VaR0.99 και TVaR0.99 για τα δεδομένα Log Danish και Log Expert Advisor..	48

Κατάλογος Σχημάτων

Σχήμα 1.1. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $N(0,4)$, $N(1,9)$ και $N(-1,16)$	2
Σχήμα 1.2. Συνάρτηση πυκνότητας πιθανότητας της κατανομής $N(0,1)$	3
Σχήμα 1.3. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $LN(0,0.252)$, $LN(0,0.52)$ και $LN(0,1)$	4
Σχήμα 1.4. Συνάρτηση πυκνότητας πιθανότητας των κατανομών t_2 , t_5 και t_{20}	5
Σχήμα 1.5. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $Beta(0.5,0.5)$, $Beta(2,5)$ και $Beta(5,1)$	6
Σχήμα 1.6. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $logis(2,1)$, $logis(5,2)$ και $logis(9,3)$	8
Σχήμα 1.7. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $Cauchy(0,1)$, $Cauchy(0,2)$ και $Cauchy(-2,0.5)$	9
Σχήμα 1.8. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $Wei(0.5,1)$, $Wei(1,1)$, $Wei(1.5,0.5)$ και $Wei(5,2)$	10
Σχήμα 1.9. Λοξότητα μιας κατανομής.....	13
Σχήμα 2.1. Συνάρτηση πυκνότητας της κατανομής SNa για $a = 0, 1, 3, 10$	18
Σχήμα 2.2. Συνάρτηση πυκνότητας της κατανομής SNa για $a = 0, -1, -3, -10$	19
Σχήμα 2.3. Συνάρτηση πυκνότητας της κατανομής $SN\xi$, $1, 5$ για $\xi = 0, 2, 4, 6$	20
Σχήμα 2.4. Συνάρτηση πυκνότητας της κατανομής $SN0$, $\omega 2, 5$ για $\omega = 1, 2, 3, 4$	20
Σχήμα 2.5. Συνάρτηση πυκνότητας της κατανομής $ST(a, 2)$, για $a = 0, 1, 3, 10$	25
Σχήμα 2.6. Συνάρτηση πυκνότητας της κατανομής $ST(a, 2)$, για $a = 0, -1, -3, -10$	25
Σχήμα 2.7. Συνάρτηση πυκνότητας της κατανομής $ST\xi$, $1, 3, 2$, για $\xi = 0, 2, 4, 6$	26
Σχήμα 2.8. Συνάρτηση πυκνότητας της κατανομής $ST0$, $\omega 2, 3, 2$ για $\omega = 1, 2, 3, 4$	27
Σχήμα 2.9. Συνάρτηση πυκνότητας της κατανομής $BST(5,3, \lambda, 3)$	29
Σχήμα 2.10. Συνάρτηση πυκνότητας της κατανομής $SL(2, \lambda)$ για $\lambda = 0, 1, 2, 5, 10$	33
Σχήμα 2.11. Συνάρτηση πυκνότητας της κατανομής SCa για $a = 0, 1, 3, 10$	37
Σχήμα 2.12. Συνάρτηση πυκνότητας της κατανομής SCa για $a = 0, -1, -3, -10$	37
Σχήμα 3.1. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για τα δεδομένα Danish.....	40
Σχήμα 3.2. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για τα δεδομένα Expert Advisor.....	40
Σχήμα 3.3. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για το λογάριθμο των δεδομένων Danish.....	41
Σχήμα 3.4. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για το λογάριθμο των δεδομένων Expert Advisor.....	42
Σχήμα 3.5. Προσαρμογή της λοξής t κατανομής στα δεδομένα.....	47

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Στο παρόν κεφάλαιο θα παρουσιαστούν περιληπτικά αποτελέσματα από τη θεωρία πιθανοτήτων και τη στατιστική που θα φανούν χρήσιμα στα επόμενα κεφάλαια. Θα γίνει αναφορά σε ορισμένες βασικές συνεχείς κατανομές δίνοντας τη μέση τιμή, διακύμανση και διάφορες γραφικές παραστάσεις. Θα ακολουθήσει εκτενής αναφορά στη βασικότερη στατιστική μέθοδο εκτίμησης παραμέτρων, τη μέθοδο μέγιστης πιθανοφάνειας. Τέλος, θα παρουσιαστεί ένα αποτέλεσμα των Azzalini & Capitano (2014) το οποίο επιτρέπει την μετατροπή μιας συμμετρικής κατανομής σε μία ασύμμετρη/λοξή κατανομή.

1.1 Κανονική Κατανομή

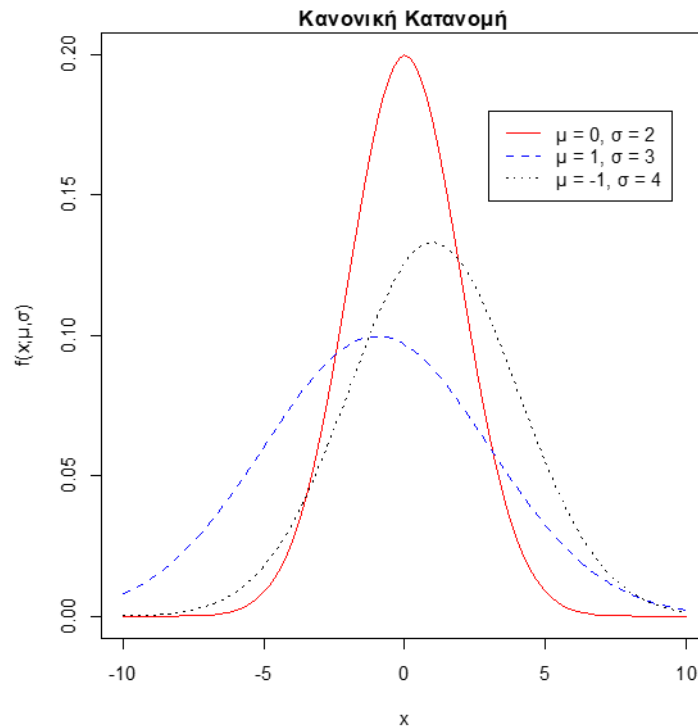
Η κανονική κατανομή εισήχθη από τον Abraham De Moivre ο οποίος προσπάθησε να υπολογίσει προσεγγιστικά πιθανότητες που αφορούν τη διωνυμική κατανομή. Η μελέτη του τελικά οδήγησε στην εύρεση της πιο σπουδαίας κατανομής της θεωρίας πιθανοτήτων και της στατιστικής, την κανονική κατανομή.

Ορισμός. Θα λέμε ότι η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή με παραμέτρους μ και σ^2 ($\mu \in \mathbb{R}$, $\sigma > 0$), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R} \quad (1.1)$$

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

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $N(\mu, \sigma^2)$ για διάφορες τιμές των μ, σ .



Σχήμα 1.1. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $N(0,4)$, $N(1,9)$ και $N(-1,16)$

Ορισμένες βασικές ιδιότητες της κανονικής κατανομής είναι οι ακόλουθες:

- Αν η $X \sim N(\mu, \sigma^2)$, τότε η $Z = \frac{x-\mu}{\sigma} \sim N(0,1)$.
- Αν η $X \sim N(\mu, \sigma^2)$, τότε $E(X) = \mu$ και $V(X) = \sigma^2$
- Η μέση τιμή, η διάμεσος και η κορυφή συμπίπτουν.
- Η συνάρτηση $f(x; \mu, \sigma)$ έχει άξονα συμμετρίας τη ευθεία $x = \mu$.

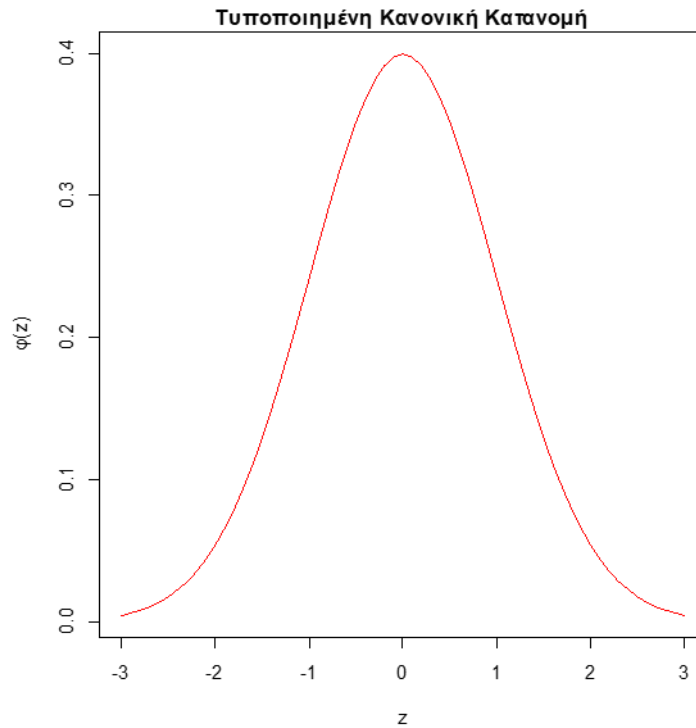
Η κανονική κατανομή $N(0,1)$ ονομάζεται τυποποιημένη κανονική κατανομή. Αν $Z \sim N(0,1)$, τότε η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \quad z \in \mathbb{R} \quad (1.2)$$

και η συνάρτηση κατανομής της συμβολίζεται με $\Phi(z)$. Για τη $\Phi(z)$ ισχύει η ενδιαφέρουσα σχέση

$$\Phi(-z) = 1 - \Phi(z), \quad z \in \mathbb{R}.$$

Η ακόλουθη γραφική παράσταση παρουσιάζει τη συνάρτηση πυκνότητας πιθανότητας της τυποποιημένης κανονικής κατανομής.



Σχήμα 1.2. Συνάρτηση πυκνότητας πιθανότητας της κατανομής $N(0,1)$.

1.2 Λογαριθμοκανονική Κατανομή

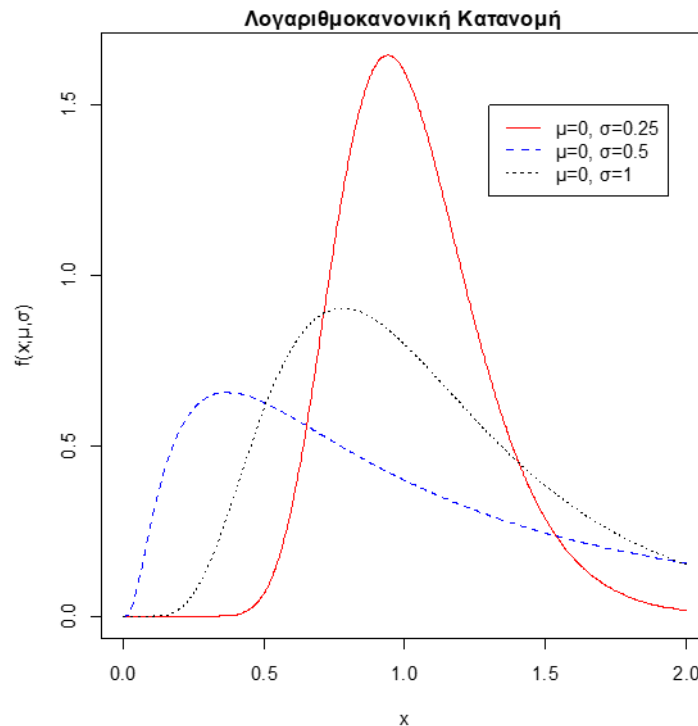
Μια τυχαία μεταβλητή X για την οποία ο λογάριθμός της $\log X$ ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$, θα λέμε ότι έχει τη λογαριθμοκανονική κατανομή με παραμέτρους μ και σ^2 . Για τη συνάρτηση πυκνότητας πιθανότητας της λογαριθμοκανονικής κατανομής έχουμε τον ακόλουθο ορισμό.

Ορισμός. Θα λέμε ότι η τυχαία μεταβλητή X ακολουθεί τη λογαριθμοκανονική κατανομή με παραμέτρους μ ($\mu \in \mathbb{R}$, meanlog parameter) και σ^2 ($\sigma > 0$, sdlog parameter), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

$$f(x; \mu, \sigma) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad x \geq 0 \quad (1.3)$$

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

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $LN(\mu, \sigma^2)$ για διάφορες τιμές των μ, σ .



Σχήμα 1.3. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $LN(0,0.25^2)$, $LN(0,0.5^2)$ και $LN(0,1)$

Για τη μέση τιμή και τη διακύμανση της λογαριθμοκανονικής κατανομής ισχύουν οι ακόλουθοι τύποι:

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad V(X) = \exp(2\mu + \sigma^2) (\exp(\sigma^2) - 1).$$

1.3 Κατανομή t

Η κατανομή του Student, ή απλά κατανομή t , είναι χρήσιμη όταν έχουμε μικρά δείγματα.

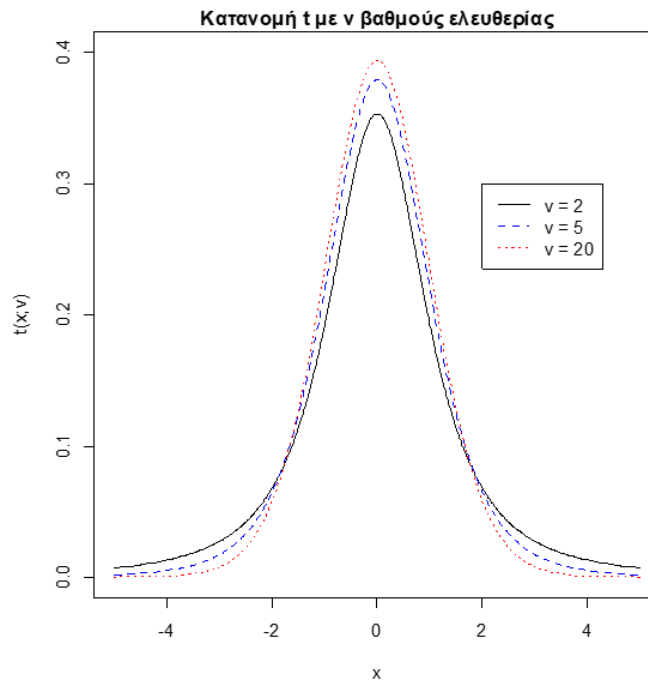
Ορισμός. Αν η τυχαία μεταβλητή $Z \sim N(0,1)$, η τυχαία μεταβλητή $Y \sim \chi^2_\nu$ και οι Z, Y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε η κατανομή της τυχαίας μεταβλητής

$$X = \frac{Z}{\sqrt{\frac{Y}{\nu}}}$$

ονομάζεται κατανομή t με ν βαθμούς ελευθερίας ($\nu > 0$, df parameter) και συμβολίζεται με t_ν . Αν $t(x; \nu)$ είναι η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X τότε

$$t(x; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \quad x \in \mathbb{R} \quad (1.4)$$

Στο ακόλουθο σχήματα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής t_ν για διάφορες τιμές του ν .



Σχήμα 1.4. Συνάρτηση πυκνότητας πιθανότητας των κατανομών t_2 , t_5 και t_{20} .

Ορισμένες βασικές ιδιότητες της κανονικής κατανομής είναι οι ακόλουθες:

- Η συνάρτηση $t(x; \nu)$ έχει άξονα συμμετρίας τη ευθεία $x = 0$.
- Αν η $X \sim t_\nu$, τότε $E(X) = 0$ για $\nu > 1$, και $V(X) = \frac{\nu}{\nu-2}$ για $\nu > 2$.
- Για $\nu = 1$, η κατανομή t_1 ανάγεται στην κατανομή Cauchy(0,1).
- Καθώς $\nu \rightarrow \infty$, η κατανομή t_ν συγκλίνει στην κατανομή $N(0,1)$.

1.4 Κατανομή Βήτα

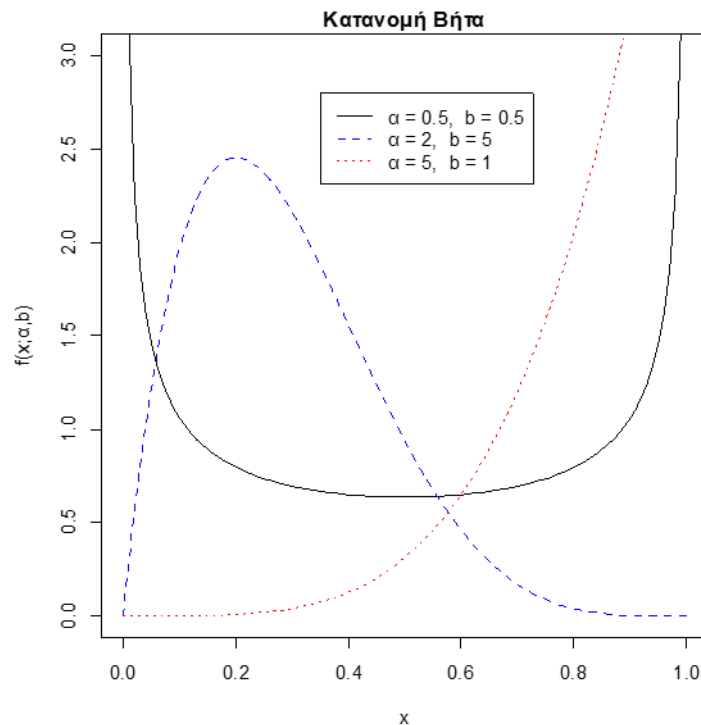
Η κατανομή Βήτα προσφέρει ένα ικανοποιητικό μοντέλο για την περιγραφή ποσοστών.

Ορισμός. Θα λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Βήτα με παραμέτρους a και b ($a, b > 0$), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

$$f(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad 0 \leq x \leq 1 \quad (1.5)$$

και θα γράφουμε $X \sim \text{Beta}(a, b)$.

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $\text{Beta}(a, b)$ για διάφορες τιμές των a, b .



Σχήμα 1.5. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $\text{Beta}(0.5, 0.5)$, $\text{Beta}(2, 5)$ και $\text{Beta}(5, 1)$.

Η μέση τιμή και τη διακύμανση της κατανομής Βήτα είναι

$$E(X) = \frac{a}{a+b}, \quad V(X) = \frac{ab}{(a+b+1)(a+b)^2}$$

1.5 Λογιστική Κατανομή

Η λογιστική κατανομή περιγράφει ικανοποιητικά μοντέλα ανάπτυξης, εμφανίζει πολλές ομοιότητες με την κανονική κατανομή, και μπορεί να χρησιμοποιηθεί αντί αυτής στην περίπτωση που εμφανίζονται ακραία γεγονότα (κατανομή με φαρδιά ουρά).

Ορισμός. Θα λέμε ότι η τυχαία μεταβλητή X ακολουθεί τη λογιστική κατανομή με παραμέτρους a ($a \in \mathbb{R}$, location parameter) και β ($\beta > 0$, scale parameter), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

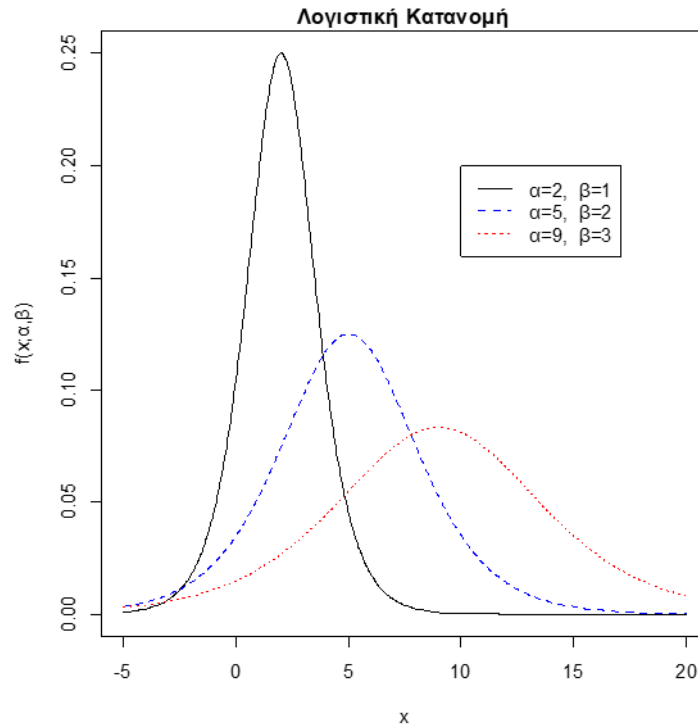
$$f(x; a, \beta) = \frac{1}{\beta} \cdot \frac{\exp\left(\frac{x-a}{\beta}\right)}{\left(1 + \exp\left(\frac{x-a}{\beta}\right)\right)^2}, \quad x \in \mathbb{R} \quad (1.6)$$

και θα γράφουμε $X \sim \text{logis}(a, \beta)$.

Η συνάρτηση κατανομής της δίνεται από το τύπο

$$F(x; a, \beta) = \frac{1}{1 + \exp\left(-\frac{x-a}{\beta}\right)}, \quad x \in \mathbb{R} \quad (1.7)$$

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $\text{logis}(a, \beta)$ για διάφορες τιμές των a, β .



Σχήμα 1.6. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $\text{logis}(2,1)$, $\text{logis}(5,2)$ και $\text{logis}(9,3)$.

Για τη μέση τιμή και τη διακύμανση της λογιστικής κατανομής ισχύουν οι ακόλουθοι τύποι:

$$E(X) = a, \quad V(X) = \frac{\beta^2 \pi^2}{3}.$$

1.6 Κατανομή Cauchy

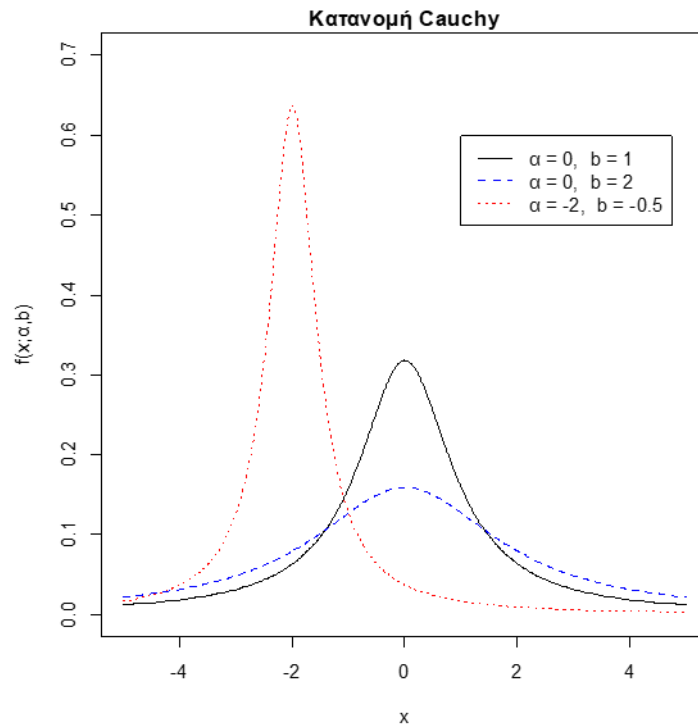
Η κατανομή Cauchy, ή αλλιώς κατανομή Lorentz είχε αρχίσει να μελετάτε από τον 17^ο αιώνα. Η τυπική της μορφή προκύπτει ως ο λόγος δυο ανεξάρτητων τυποποιημένων κανονικών κατανομών, είναι συμμετρική και έχει φαρδιές ουρές. Η χαρακτηριστική ιδιότητά της είναι ότι δεν έχει μέση τιμή.

Ορισμός. Θα λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Cauchy με παραμέτρους a ($a \in R$, location parameter) και b ($b > 0$, scale parameter), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

$$f(x; a, b) = \frac{1}{\pi b \left(1 + \left(\frac{x-a}{b}\right)^2\right)}, \quad x \in \mathbb{R} \quad (1.8)$$

και θα γράφουμε $X \sim \text{Cauchy}(a, b)$.

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $\text{Cauchy}(a, b)$ για διάφορες τιμές των a, b .



Σχήμα 1.7. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $\text{Cauchy}(0,1)$, $\text{Cauchy}(0,2)$ και $\text{Cauchy}(-2,0.5)$.

Η συνάρτηση κατανομής της δίνεται από τον τύπο

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \quad (1.9)$$

Στην ειδική περίπτωση όπου $a = 0$ και $b = 1$ λαμβάνουμε τη τυπική Cauchy κατανομή που έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R} \quad (1.10)$$

1.7 Κατανομή Weibull

Η κατανομή Weibull είναι μια συνεχής κατανομή που πήρε το όνομα της από το Σουηδό μαθηματικό Waloddi Weibull, αν και είχε ήδη αναγνωριστεί παλαιότερα από τον Fréchet.

Ορισμός. Θα λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Weibull με παραμέτρους a ($a > 0$, shape parameter) και b ($b > 0$, scale parameter), αν η συνάρτηση πυκνότητας πιθανότητάς της δίνεται από τον τύπο

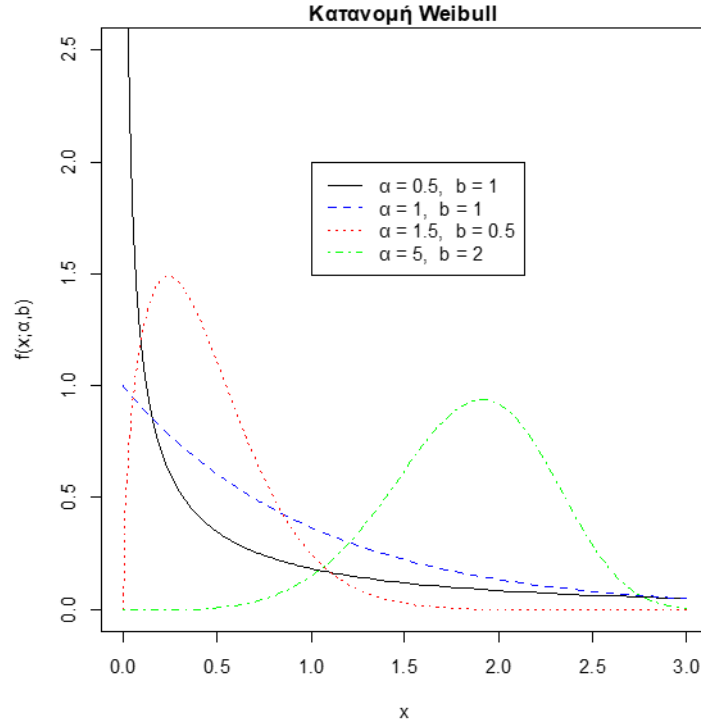
$$f(x; a, b) = \left(\frac{a}{b}\right) \left(\frac{x}{b}\right)^{a-1} \exp\left(-\left(\frac{x}{b}\right)^a\right), \quad x \geq 0 \quad (1.11)$$

και θα γράφουμε $X \sim \text{Wei}(a, b)$.

Η συνάρτηση κατανομής της $F(x; a, b)$ ικανοποιεί τη σχέση

$$F(x; a, b) = 1 - \exp\left(-\left(\frac{x}{b}\right)^a\right), \quad x \geq 0 \quad (1.12)$$

Στο ακόλουθο σχήμα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $\text{Wei}(a, b)$ για διάφορες τιμές των a, b .



Σχήμα 1.8. Συνάρτηση πυκνότητας πιθανότητας των κατανομών $\text{Wei}(0.5,1)$, $\text{Wei}(1,1)$ $\text{Wei}(1.5,0.5)$ και $\text{Wei}(5,2)$.

Για τη μέση τιμή και τη διακύμανση της λογιστικής κατανομής ισχύουν οι ακόλουθοι τύποι:

$$E(X) = b \cdot \Gamma\left(1 + \frac{1}{a}\right), \quad \text{Var}(X) = b^2 \left(\Gamma\left(1 + \frac{2}{a}\right) - \Gamma\left(1 + \frac{1}{a}\right)^2 \right)$$

1.8 Μέθοδος Μέγιστης Πιθανοφάνειας

Ένα συχνό πρόβλημα το οποίο αντιμετωπίζετε στις στατιστικές μελέτες είναι η εύρεση εκτιμητριών των παραμέτρων μιας κατανομής πιθανότητας. Η μέθοδος που εφαρμόζεται πιο συχνά στην πράξη είναι η μέθοδος μέγιστης πιθανοφάνειας.

Έστω $X = (X_1, X_2, \dots, X_n)$ ένα τυχαίο δείγμα από ένα πληθυσμό με συνάρτηση πυκνότητας πιθανότητας (ή συνάρτηση πιθανότητας) $f(x|\theta)$ όπου $\theta = (\theta_1, \theta_2, \dots, \theta_k)'$, $k \geq 1$, ένα διάνυσμα παραμέτρων με $\theta \in \Theta$ (Θ συμβολίζει τον παραμετρικό χώρο, $\Theta \subseteq R^k$). Η από κοινού συνάρτηση πυκνότητας πιθανότητας του δείγματος, $L(\theta) = L(\theta|x)$, η οποία θεωρείται συνάρτηση της παραμέτρου θ ονομάζεται συνάρτηση πιθανοφάνειας, δηλαδή

$$L(\theta) = L(\theta|x) = L(\theta_1, \theta_2, \dots, \theta_k | x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i|\theta) = \prod_{i=1}^n L_i(\theta), \quad \theta \in \Theta$$

Αν για κάθε δειγματοσημείο x το διάνυσμα $\hat{\theta} = \hat{\theta}(x) = (\hat{\theta}_1(x), \hat{\theta}_2(x), \dots, \hat{\theta}_k(x))'$ με $\hat{\theta} \in \Theta$ ικανοποιεί τη σχέση

$$L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

ή ισοδύναμα τη σχέση

$$\ell(\hat{\theta}) = \sup_{\theta \in \Theta} \ell(\theta)$$

όπου

$$\ell(\theta) = \ell(\theta|x) = \log L(\theta) = \sum_{i=1}^n \log f(x_i|\theta) = \sum_{i=1}^n \log L_i(\theta),$$

τότε η στατιστική συνάρτηση $\hat{\theta} = (\hat{\theta}_1(X), \hat{\theta}_2(X), \dots, \hat{\theta}_k(X))'$ ονομάζεται εκτιμητής μέγιστης πιθανοφάνειας (maximum likelihood estimator) της παραμέτρου θ .

Ο εκτιμητής μέγιστης πιθανοφάνειας (EMΠ) προκύπτει συνήθως από τη λύση του συστήματος των εξισώσεων

$$U_r(\theta) = \frac{\partial}{\partial \theta_r} \ell(\theta) = \frac{\partial}{\partial \theta_r} \log L(\theta) = 0, \quad r = 1, 2, \dots, k$$

ή ισοδύναμα

$$U(\theta) = \frac{\partial}{\partial \theta} \ell(\theta) = (U_1(\theta), U_2(\theta), \dots, U_k(\theta))' = 0$$

αρκεί ο πίνακας των δευτέρων μερικών παραγώγων

$$\frac{\partial^2}{\partial \theta^2} \ell(\theta) = \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\theta) \right)_{k \times k} = \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log L(\theta) \right)_{k \times k}$$

υπολογισμένος στη θέση $\hat{\theta}$ να είναι αρνητικά ορισμένος. Η λύση του συστήματος $U(\theta) = 0$ επιτυγχάνεται συνήθως με μεθόδους της αριθμητικής ανάλυσης.

Η προαναφερθείσα μεθοδολογία γίνεται καλύτερα κατανοητή μέσω ενός παραδείγματος. Έστω ένα τυχαίο δείγμα X_1, X_2, \dots, X_n από την κατανομή $N(\theta)$ ($\theta = (\theta_1, \theta_2) = (\mu, \sigma^2)$, $(\theta_1, \theta_2) \in R \times (0, \infty)$). Τότε

$$L(\theta) = L(\theta|x) = \prod_{i=1}^n f(x_i|\theta) = \frac{1}{(2\pi\theta_2)^{n/2}} \exp\left(-\frac{1}{2\theta_2} \sum (x_i - \theta_1)^2\right),$$

$$\ell(\theta) = \ell(\theta|x) = \log L(\theta) = \sum_{i=1}^n \log f(x_i|\theta) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \theta_2 - \frac{1}{2\theta_2} \sum (x_i - \theta_1)^2.$$

Ο εκτιμητής μέγιστης πιθανοφάνειας $\hat{\theta} = (\hat{\theta}_1(X), \hat{\theta}_2(X), \dots, \hat{\theta}_k(X))'$ της παραμέτρου θ προκύπτει από τη λύση του συστήματος

$$U(\theta) = (U_1(\theta), U_2(\theta))' = \left(\frac{1}{\theta_2} \sum (x_i - \theta_1), -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2} \sum (x_i - \theta_1)^2 \right)' = (0, 0)'$$

που είναι η

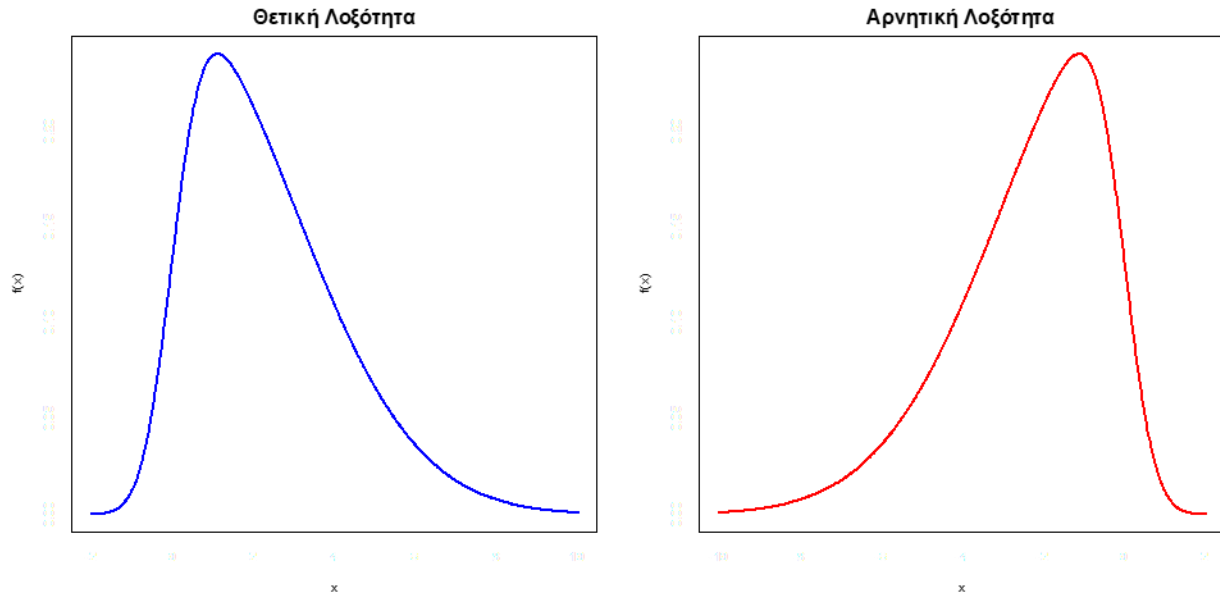
$$\hat{\theta} = (\hat{\theta}_1(X), \hat{\theta}_2(X)) = \left(\bar{X}, \frac{1}{n} \sum (X_i - \bar{X})^2 \right).$$

1.9 Λοξές συμμετρικές κατανομές

Στη θεωρία πιθανοτήτων η λοξότητα είναι ένα μέτρο της ασυμμετρίας της κατανομής πιθανότητας μιας τυχαίας μεταβλητής γύρω από τη μέση τιμή της. Η τιμή της λοξότητας μπορεί να είναι θετική, μηδενική, αρνητική ή απροσδιόριστη. Η λοξότητα (συμβολισμός $\gamma_1(X)$) μιας τυχαίας μεταβλητής X υπολογίζεται από τον τύπο

$$\gamma_1(X) = E \left[\left(\frac{X - \mu_X}{\sigma_X} \right)^3 \right].$$

Για μια μονοκόρυφη κατανομή, αρνητική τιμή της λοξότητας υποδεικνύει ότι η ουρά βρίσκεται στην αριστερή πλευρά της κατανομής και θετική τιμή υποδεικνύει ότι η ουρά βρίσκεται στα δεξιά της πλευρά. Αν η λοξότητα είναι μηδενική, τότε αυτό συνήθως σημαίνει ότι οι ουρές και στις δύο πλευρές της μέσης τιμής είναι ίδιες.



Σχήμα 1.9. Λοξότητα μιας κατανομής

Πολλές από τις κατανομές πιθανότητας που θα μελετήσουμε παρακάτω αποτελούν τροποποιήσεις συμμετρικών συνεχών συναρτήσεων πυκνότητας πιθανότητας. Αυτές καλούνται συνήθως πυκνότητες βάσης (base densities). Επομένως δεδομένου ότι η έννοια της συμμετρικής πυκνότητας πιθανότητας έχει σημαντικό ρόλο, αξίζει να δώσουμε ένα ορισμό για τη μονοδιάστατη περίπτωση.

Ορισμός. Λέμε ότι η συνάρτηση πυκνότητας πιθανότητας $f_0(x)$ μιας τυχαίας μεταβλητής X είναι συμμετρική γύρω από ένα σημείο, έστω x_0 , αν

$$f_0(x_0 + x) = f_0(x_0 - x), \quad x \in \mathbb{R}$$

Χωρίς βλάβη της γενικότητας από εδώ και πέρα θα θεωρούμε ότι $x_0 = 0$ (κεντρική συμμετρία), δηλαδή

$$f_0(x) = f_0(-x), \quad x \in \mathbb{R}.$$

Σε αυτή την περίπτωση είναι εύκολο να διαπιστωθεί ότι οι τυχαίες μεταβλητές X και $-X$ έχουν την ίδια κατανομή.

Στη συνέχεια δίνουμε μια πρόταση η οποία παίζει κεντρικό ρόλο στην παρούσα διπλωματική εργασία αφού επιτρέπει την κατασκευή λοξών κατανομών από αντίστοιχες συμμετρικές κατανομές και οφείλεται στον στους Azzalini (1985) (δείτε επίσης Azzalini (2005) και Azzalini & Capitano (2014)).

Πρόταση. Έστω

- $f_0(x)$ συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς τυχαίας μεταβλητής τέτοια ώστε $f_0(-x) = f_0(x)$, $x \in \mathbb{R}$,
- $G_0(y)$ συνάρτηση κατανομής μιας συνεχούς τυχαίας μεταβλητής τέτοια ώστε $G_0(-y) = 1 - G_0(y)$, $y \in \mathbb{R}$, και
- $w(x)$ πραγματική συνάρτηση τέτοια ώστε $w(-x) = -w(x)$, $x \in \mathbb{R}$.

Τότε, η συνάρτηση

$$f(x) = 2f_0(x)G_0(w(x)), \quad x \in \mathbb{R}$$

είναι μια συνάρτηση πυκνότητας.

Απόδειξη 1.

Η συνάρτηση

$$g(x) = 2 \left[G_0\{w(x)\} - \frac{1}{2} \right] f_0(x), \quad x \in \mathbb{R},$$

είναι μια περιττή συνάρτηση αφού

$$\begin{aligned} g(-x) &= 2 \left[G_0\{w(-x)\} - \frac{1}{2} \right] f_0(-x) \\ &= 2 \left[G_0\{-w(x)\} - \frac{1}{2} \right] f_0(x) \\ &= 2 \left[1 - G_0\{w(x)\} - \frac{1}{2} \right] f_0(x) \end{aligned}$$

$$\begin{aligned}
&= -2 \left[G_0\{w(x)\} - \frac{1}{2} \right] f_0(x) \\
&= -g(x).
\end{aligned}$$

Επίσης είναι ολοκληρώσιμη στο \mathbb{R} αφού

$$|g(x)| = 2 \left| G_0\{w(x)\} - \frac{1}{2} \right| f_0(x) \leq 2 \cdot \frac{1}{2} \cdot f_0(x) = f_0(x).$$

Έτσι

$$0 = \int_{\mathbb{R}} g(x) dx = \int_{\mathbb{R}} 2 \left[G_0\{w(x)\} - \frac{1}{2} \right] f_0(x) dx = \int_{\mathbb{R}} [2f_0(x)G_0(w(x)) - f_0(x)] dx$$

δηλαδή

$$\int_{\mathbb{R}} 2f_0(x)G_0(w(x)) = 1.$$

Απόδειξη 2.

Έστω Z_0 μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας f_0 (δηλαδή συμμετρική γύρω από το 0) και T μια τυχαία μεταβλητή με συνάρτηση κατανομής G_0 (δηλαδή συμμετρική γύρω από το 0) ανεξάρτητη της Z_0 . Οι τυχαίες μεταβλητές Z_0 και $-Z_0$ έχουν την ίδια κατανομή, αφού

$$F_{Z_0}(-x) = P(Z_0 \leq -x) = P(-Z_0 \geq x) = 1 - F_{-Z_0}(x)$$

απ' όπου παραγωγίζοντας προκύπτει ότι $f_{Z_0}(-x) = f_{-Z_0}(x)$, οπότε $f_{Z_0}(x) = f_{-Z_0}(x)$, $x \in \mathbb{R}$.

Έτσι η τυχαία μεταβλητή $W = w(Z_0)$ είναι μια συμμετρική κατανομή γύρω από το 0 αφού για ένα σύνολο Borel A έχουμε ότι

$$\begin{aligned}
P\{W \in -A\} &= P\{-W \in A\} = P\{-w(Z_0) \in A\} = P\{w(-Z_0) \in A\} \\
&= P\{w(Z_0) \in A\} = P\{W \in A\}.
\end{aligned}$$

Εφόσον η T είναι συμμετρική γύρω από το 0, τότε και η $T - W$ είναι επίσης συμμετρική γύρω από το 0. Συνεπώς

$$\frac{1}{2} = P(T \leq W) = E_{Z_0}\{P\{T \leq w(Z_0) | Z_0 = x\}\} = \int_{\mathbb{R}} G_0\{w(x)\} f_0(x) dx$$

δηλαδή

$$\int_{\mathbb{R}} 2f_0(x)G_0(w(x)) = 1.$$

ΚΕΦΑΛΑΙΟ 2

Λοξές κατανομές

2.1 Γέννηση και ιδιότητες της λοξής κανονικής κατανομής

Αναφέραμε στο προηγούμενο κεφάλαιο ότι αν

- $f_0(x)$ είναι μια συνάρτηση πυκνότητας μιας συνεχούς τυχαίας μεταβλητής στο χώρο \mathbb{R} τέτοια ώστε $f_0(-x) = f_0(x)$,
- $G_0(y)$ είναι μια συνάρτηση κατανομής μιας συνεχούς τυχαίας μεταβλητής τέτοια ώστε $G_0(-y) = 1 - G_0(y)$, και
- $w(x)$ είναι μια πραγματική συνάρτηση τέτοια ώστε $w(-x) = -w(x)$,

τότε η συνάρτηση

$$f(x) = 2f_0(x)G_0(w(x)), \quad x \in \mathbb{R}, \quad (2.1)$$

είναι μια συνάρτηση πυκνότητας.

Θέτοντας στη σχέση (2.1)

- $f_0(z)$ τη συνάρτηση πυκνότητας της τυπικής κανονικής κατανομής $N(0,1)$, δηλαδή

$$f_0(z) = \varphi(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}},$$

- $G_0(z)$ τη συνάρτηση κατανομής της τυπικής κανονικής κατανομής $N(0,1)$, δηλαδή

$$G_0(z) = \Phi(z) = \int_{-\infty}^z \varphi(t)dt = \int_{-\infty}^z \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt,$$

- και $w(z) = az$, όπου $a \in \mathbb{R}$,

προκύπτει η συνάρτηση πυκνότητας

$$\varphi(z; a) = 2\varphi(z)\Phi(az), \quad -\infty < z < \infty. \quad (2.2)$$

Μια συνεχής τυχαία μεταβλητή Z θα λέμε ότι έχει την **τυπική λοξή κατανομή (Standard Skew-Normal distribution)** με παράμετρο a και θα γράφουμε $Z \sim SN(a)$, αν η συνάρτηση πυκνότητας της Z δίνεται από τη σχέση (2.2).

Η τυπική λοξή κατανομή με παράμετρο a έχει ενδιαφέρουσες ιδιότητες. Μερικές από αυτές είναι οι ακόλουθες:

1. Η κατανομή $SN(a)$, για $a = 0$, ανάγεται στην τυπική κανονική κατανομή.
2. Η κατανομή $SN(a)$, για $a \rightarrow \infty$, συγκλίνει στη δεξιά του μηδενός μισή κανονική κατανομή (Half-Normal distribution) ή αλλιώς στην κατανομή χ_1 με ένα βαθμό ελευθερίας (χ_1), αφού

$$\lim_{a \rightarrow \infty} \varphi(z; a) = \begin{cases} 0, & z < 0 \\ 2\varphi(z), & z \geq 0. \end{cases}$$

3. Η κατανομή $SN(a)$, για $a \rightarrow -\infty$, συγκλίνει στην αριστερά του μηδενός μισή κανονική κατανομή (Half-Normal distribution), αφού

$$\lim_{a \rightarrow -\infty} \varphi(z; a) = \begin{cases} 2\varphi(z), & z \leq 0 \\ 0, & z > 0. \end{cases}$$

4. Αν η τυχαία μεταβλητή $Z \sim SN(a)$, τότε η τυχαία μεταβλητή $-Z \sim SN(-a)$.
5. Αν η τυχαία μεταβλητή $Z \sim SN(a)$, τότε η τυχαία μεταβλητή $Z^2 \sim \chi_1^2$.
6. Αν η διδιάστατη τ.μ. (X, Y) έχει τη διδιάστατη κανονική κατανομή με παραμέτρους $\mu_X = \mu_Y = 0$, $\sigma_X^2 = \sigma_Y^2 = 1$ και ρ , τότε η δεσμευμένη κατανομή της τ.μ. Y δοθέντος ότι $X > 0$ είναι η κατανομή $SN(a)$, όπου $a = \rho(1 - \rho^2)^{-1/2}$.
7. Όσο το $|a|$ αυξάνει τόσο η λοξότητα της κατανομής $SN(a)$ αυξάνει.
8. Αν η τ.μ. $Z \sim SN(a)$, τότε

$$E(Z) = \delta\sqrt{2/\pi} \xrightarrow{a \rightarrow \pm\infty} \text{sign}(a) \times 0.79788$$

$$\text{Var}(Z) = 1 - \frac{2\delta^2}{\pi} \xrightarrow{a \rightarrow \pm\infty} 0.36338$$

όπου

$$\delta = \frac{a}{\sqrt{1+a^2}}.$$

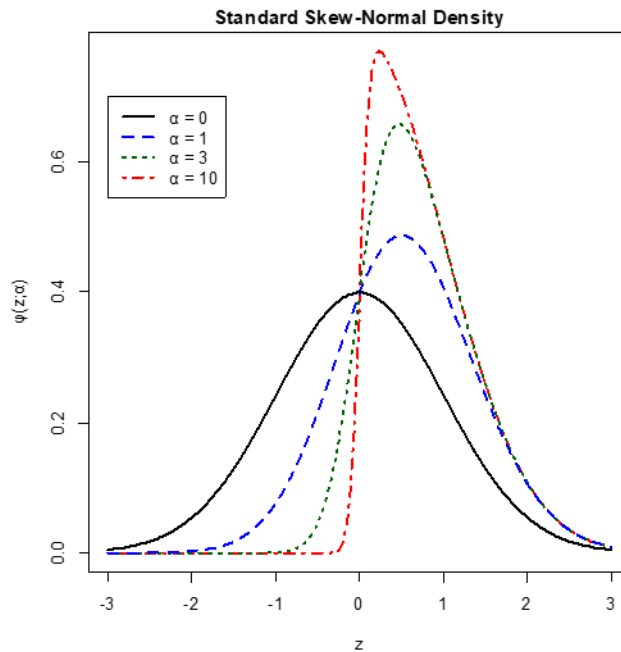
9. Αν η τ.μ. $Z \sim SN(a)$, τότε ο συντελεστής λοξότητας γ_1 και ο συντελεστής κύρτωσης γ_2 δίνονται από τους τύπους

$$\gamma_1(Z) = E \left[\left(\frac{Z - \mu_Z}{\sigma_Z} \right)^3 \right] = \frac{(4 - \pi)(\delta\sqrt{2/\pi})^3}{2 \left(1 - \frac{2\delta^2}{\pi} \right)^{3/2}}$$

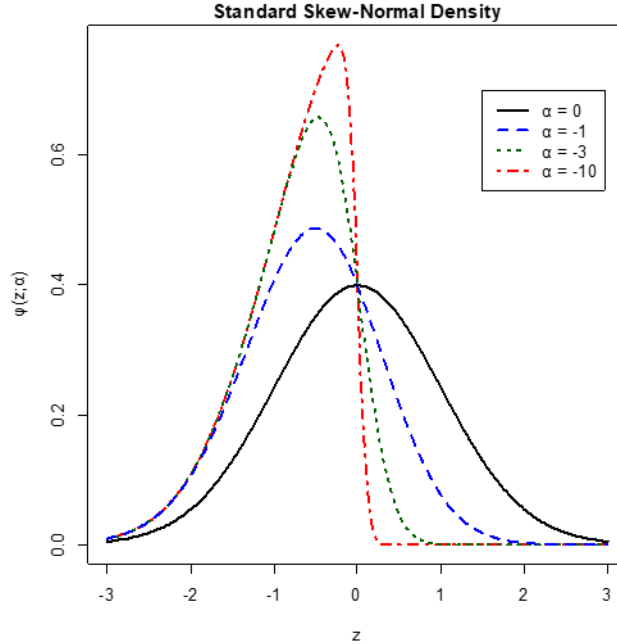
και

$$\gamma_2(Z) = E \left[\left(\frac{Z - \mu_Z}{\sigma_Z} \right)^4 \right] - 3 = 2(\pi - 3) \frac{(\delta\sqrt{2/\pi})^4}{2 \left(1 - \frac{2\delta^2}{\pi} \right)^2}.$$

Στα δύο ακόλουθα σχήματα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας της κατανομής $SN(a)$ για διάφορες τιμές του a .



Σχήμα 2.1. Συνάρτηση πυκνότητας της κατανομής $SN(a)$ για $a = 0, 1, 3, 10$.



Σχήμα 2.2. Συνάρτηση πυκνότητας της κατανομής $SN(a)$ για $a = 0, -1, -3, -10$.

Από τα δύο παραπάνω σχήματα προκύπτει ότι για μικρές αρνητικές τιμές και μεγάλες θετικές τιμές του a αυξάνεται το μέγεθος της λοξότητας της κατανομής.

Ας υποθέσουμε ότι η τυχαία μεταβλητή $Z \sim SN(a)$. Η συνάρτηση πυκνότητας της τυχαίας μεταβλητής

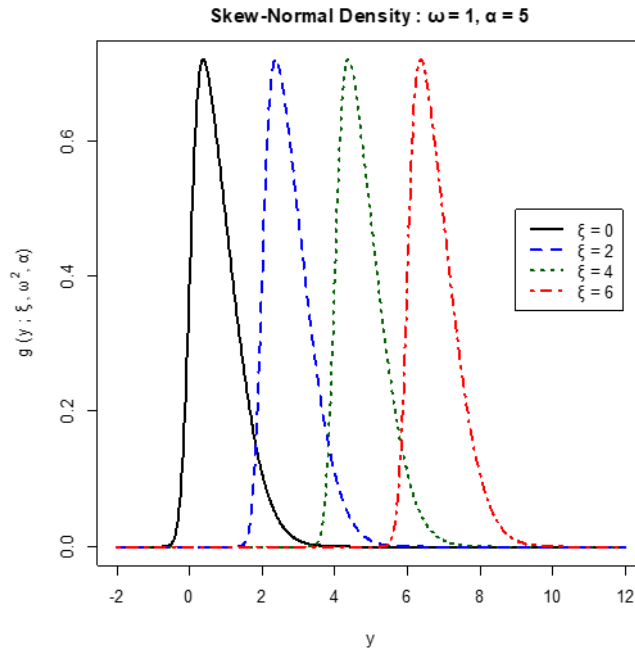
$$Y = \xi + \omega Z$$

μπορεί εύκολα να διαπιστωθεί ότι δίνεται από τον τύπο

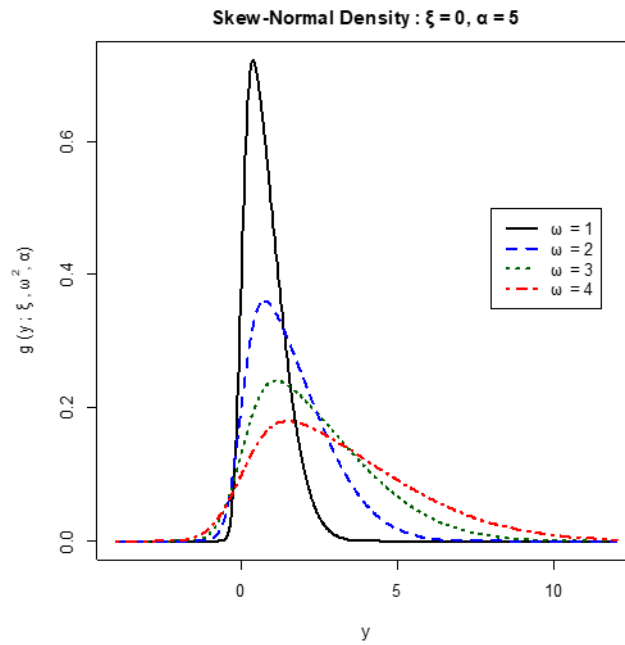
$$g(y; \xi, \omega^2, a) = \frac{2}{\omega} \varphi\left(\frac{y - \xi}{\omega}\right) \Phi\left(a \frac{y - \xi}{\omega}\right) = \frac{1}{\omega} \varphi\left(\frac{y - \xi}{\omega}; a\right). \quad (2.3)$$

Μια συνεχής τυχαία μεταβλητή Y θα λέμε ότι έχει τη **λοξή κανονική κατανομή (Skew-Normal distribution)** με παραμέτρους $\xi \in R$, ω^2 ($\omega \in R^+$) και $a \in R$, και θα γράφουμε $Y \sim SN(\xi, \omega^2, a)$, αν η συνάρτηση πυκνότητας της Y δίνεται από τη σχέση (2.3). Η παράμετρος ξ είναι παράμετρος θέσης, η παράμετρος ω είναι παράμετρος κλίμακας και η παράμετρος a είναι παράμετρος λοξότητας.

Ακολουθούν δύο γραφήματα τα οποία παρουσιάζουν τη λοξή κανονική κατανομή με διαφορετικές παραμέτρους ξ και ω το καθένα.



Σχήμα 2.3. Συνάρτηση πυκνότητας της κατανομής $SN(\xi, 1, 5)$ για $\xi = 0, 2, 4, 6$.



Σχήμα 2.4. Συνάρτηση πυκνότητας της κατανομής $SN(0, \omega^2, 5)$ για $\omega = 1, 2, 3, 4$.

Η ροπογεννήτρια της τυχαίας μεταβλητής $Y \sim SN(\xi, \omega^2, a)$ δίνεται από τον τύπο

$$M_Y(t) = 2 \exp\left(\xi t + \frac{\omega^2 t^2}{2}\right) \Phi(\delta \omega t).$$

Από τον παραπάνω τύπο με παραγωγή προκύπτει ότι

$$E(Y) = \xi + \omega\delta\sqrt{2/\pi}$$

και

$$Var(Y) = \omega^2 \left(1 - \frac{2\delta^2}{\pi}\right).$$

Επίσης για τους συντελεστές λοξότητας και κύρτωσης της τυχαίας μεταβλητής Y έχουμε ότι

$$\gamma_1(Y) = \gamma_1(Z) = \frac{(4 - \pi)(\delta\sqrt{2/\pi})^3}{2\left(1 - \frac{2\delta^2}{\pi}\right)^{3/2}},$$

$$\gamma_2(Y) = \gamma_2(Z) = 2(\pi - 3) \frac{(\delta\sqrt{2/\pi})^4}{2\left(1 - \frac{2\delta^2}{\pi}\right)^2}.$$

Στη συνέχεια θα δώσουμε βασικές σχέσεις με τις οποίες μπορούν να προκύψουν εκτιμητές μέγιστης πιθανοφάνειας για τις παραμέτρους της λοξής κανονικής κατανομής. Έστω λοιπόν Y_1, Y_2, \dots, Y_n ένα τυχαίο δείγμα μεγέθους n από την κατανομή $SN(\xi, \omega^2, a)$. Η συνάρτηση πιθανοφάνειας δίνεται από το τύπο

$$L(\xi, \omega, a) = \left(\frac{2}{\omega}\right)^n \prod_{i=1}^n \varphi\left(\frac{y_i - \xi}{\omega}\right) \Phi\left(a \frac{y_i - \xi}{\omega}\right).$$

Προκειμένου μεγιστοποιήσουμε τη συνάρτηση $L(\xi, \omega, a)$ θα χρησιμοποιήσουμε το λογάριθμό της $l(\xi, \omega, a) = \ln(L(\xi, \omega, a))$. Παρατηρούμε ότι

$$l(\xi, \omega, a) = n \ln\left(\frac{2}{\omega}\right) + \sum_{i=1}^n \ln \varphi\left(\frac{y_i - \xi}{\omega}\right) + \sum_{i=1}^n \ln \Phi\left(a \frac{y_i - \xi}{\omega}\right)$$

ή ισοδύναμα

$$l(\xi, \omega, a) = n \ln\left(\frac{2}{\omega}\right) - \frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \left(\frac{y_i - \xi}{\omega}\right)^2 + \sum_{i=1}^n \Phi\left(a \frac{y_i - \xi}{\omega}\right).$$

Για τη εύρεση των εκτιμητών μέγιστης πιθανοφάνειας θα πρέπει να λυθεί το σύστημα των εξισώσεων

$$\frac{\partial l(\xi, \omega, \alpha)}{\partial a} = \sum_{i=1}^n \frac{\left(\frac{y_i - \xi}{\omega}\right) \varphi\left(a \frac{y_i - \xi}{\omega}\right)}{\Phi\left(a \frac{y_i - \xi}{\omega}\right)} = 0$$

$$\frac{\partial l(\xi, \omega, \alpha)}{\partial \xi} = \sum_{i=1}^n \frac{y_i - \xi}{\omega} - a \sum_{i=1}^n \frac{\varphi\left(a \frac{y_i - \xi}{\omega}\right)}{\Phi\left(a \frac{y_i - \xi}{\omega}\right)} = 0$$

$$\frac{\partial l(\xi, \omega, \alpha)}{\partial \omega} = -n + \sum_{i=1}^n \left(\frac{y_i - \xi}{\omega}\right)^2 - a \sum_{i=1}^n \frac{\left(\frac{y_i - \xi}{\omega}\right) \varphi\left(a \frac{y_i - \xi}{\omega}\right)}{\Phi\left(a \frac{y_i - \xi}{\omega}\right)} = 0$$

Παρατηρούμε ότι η τελευταία εξίσωση μπορεί να γραφτεί στην ισοδύναμη μορφή

$$\frac{1}{n} \sum_{i=1}^n (y_i - \xi)^2 = \omega^2.$$

Αξίζει να σημειωθεί ότι στην ειδική περίπτωση όπου η παράμετρος λοξότητας $a = 0$ (δηλαδή $Y \sim N(\xi, \omega^2)$), από το παραπάνω σύστημα προκύπτουν οι εκτιμητές μέγιστης πιθανοφάνειας

$$\hat{\xi} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \hat{\omega}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2,$$

όπως αναμενόταν.

Η εφαρμογή της μεθόδου της μέγιστης πιθανοφάνειας στην περίπτωση της λοξής κανονικής κατανομής παρουσιάζει δύο προβλήματα. Το πρώτο αφορά την περίπτωση των μικρών/μεσαίων μεγεθών δειγμάτων όπου συχνά εμφανίζεται το πρόβλημα να μην φράσσεται από πάνω η πιθανοφάνεια ως προς a , με αποτέλεσμα να τείνει στο άπειρο η τιμή του a . Το δεύτερο πρόβλημα προκύπτει όταν $a = 0$ αφού σε αυτή την περίπτωση ο πίνακας πληροφορίας του Fisher είναι ιδιάζων και επομένως δεν μπορεί να εφαρμοστεί η κλασική ασυμπτωτική θεωρία των εκτιμητών μέγιστης πιθανοφάνειας. Για περισσότερες λεπτομέρειες δείτε Azzalini (2014).

2.2 Γέννηση και ιδιότητες της λοξής t κατανομής

Ας συμβολίσουμε με $t(x; v)$ τη συνάρτηση πυκνότητας της κατανομής t με $v > 0$ βαθμούς ελευθερίας, δηλαδή

$$t(x; v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v} \Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{x^2}{v}\right)^{-(v+1)/2}, \quad -\infty < x < \infty,$$

και με $T(x; n)$ τη συνάρτηση κατανομής της. Θέτοντας στο γνωστό τύπο

$$f(x) = 2f_0(x)G_0(w(x)), \quad x \in \mathbb{R},$$

$f_0(x) = t(x; v)$, $G_0(x) = T(x; n)$ και $w(x) = ax$, προκύπτει η συνάρτηση πυκνότητας

$$2 t(x; v)T(ax; n), \quad -\infty < x < \infty$$

Αν και η παραπάνω συνάρτηση πυκνότητας θα μπορούσε να χρησιμοποιηθεί ως μια ασυμμετρική γενίκευση της κατανομής t_v , στις εφαρμογές χρησιμοποιείται μια άλλη γενίκευσή της.

Μια συνεχής τυχαία μεταβλητή Z θα λέμε ότι έχει την **τυπική λοξή t κατανομή (Standard Skew- t distribution)** με παραμέτρους $v > 0$ και $a \in \mathbb{R}$ και θα γράφουμε $Z \sim ST(a, v)$, αν

$$Z = \frac{Z_0}{\sqrt{V}}$$

όπου $Z_0 \sim SN(a)$, $V \sim \chi_v^2/v$, και Z και V ανεξάρτητες τυχαίες μεταβλητές. Προφανώς, για $a = 0$ η κατανομή $ST(a, v)$ ανάγεται στη κατανομή t με v βαθμούς ελευθερίας, αφού τότε ισχύει ότι $Z_0 \sim N(0,1)$. Η συνάρτηση πυκνότητας $t(x; a, v)$ της κατανομής $ST(a, v)$ δίνεται από τον τύπο

$$t(x; a, v) = 2t(x; v)T\left(ax \sqrt{\frac{v+1}{v+x^2}}; v+1\right), \quad -\infty < x < \infty.$$

Η τυπική λοξή t κατανομή με παραμέτρους a και v έχει μερικές ενδιαφέρουσες ειδικές περιπτώσεις. Μερικές από αυτές είναι οι ακόλουθες:

1. Η κατανομή $ST(a, v)$, για $a = 0$, ανάγεται στην κατανομή t με v βαθμούς ελευθερίας ($Z \sim t_v$).
2. Η κατανομή $ST(a, v)$, για $a \rightarrow \infty$, συγκλίνει στην κατανομή της τυχαίας μεταβλητής $|Z|$ όπου $Z \sim t_v$, αφού

$$\lim_{a \rightarrow \infty} t(x; a, v) = \begin{cases} 0, & x < 0 \\ 2t(x; v), & x \geq 0. \end{cases}$$

3. Η κατανομή $ST(a, v)$, για $a \rightarrow -\infty$, συγκλίνει στην κατανομή της τυχαίας μεταβλητής $-|Z|$ όπου $Z \sim t_v$, αφού

$$\lim_{a \rightarrow -\infty} t(x; a, v) = \begin{cases} 2t(x; v), & x \leq 0 \\ 0, & x > 0. \end{cases}$$

4. Η κατανομή $ST(a, v)$, για $v \rightarrow \infty$, συγκλίνει στην κατανομή $SN(a)$.
5. Αν η τυχαία μεταβλητή $Z \sim ST(a, v)$, τότε η τυχαία μεταβλητή Z^2 έχει την κατανομή $F_{1,v}$.
6. Για τη ροπή k -τάξης της τυχαίας μεταβλητής $Z \sim ST(a, v)$, έχουμε ότι

$$E(Z^k) = E(Z_0^k V^{-k/2}) = E(Z_0^k) E(V^{-k/2}) = E(Z_0^k) \frac{\left(\frac{n}{2}\right)^{k/2} \Gamma\left(\frac{1}{2}(n-k)\right)}{\Gamma\left(\frac{1}{2}n\right)}, \quad n > k.$$

7. Αν η τ.μ. $Z \sim ST(a, v)$, τότε

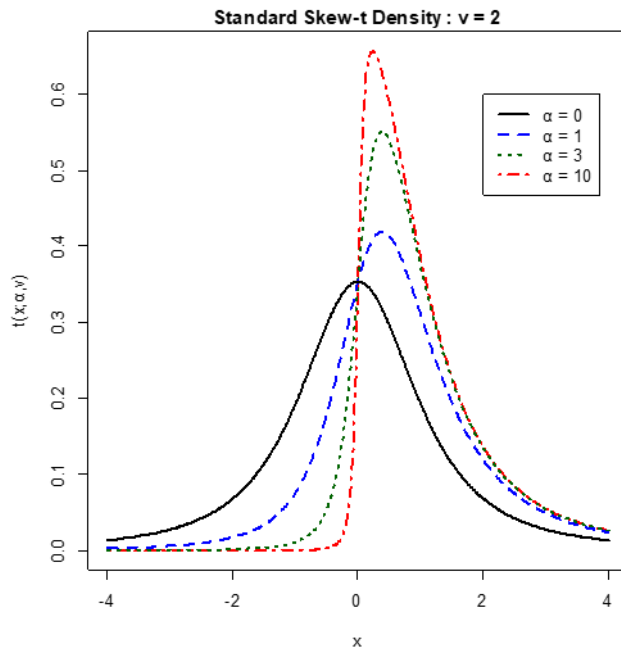
$$E(Z) = b_v \delta$$

$$Var(Z) = \left(\frac{n}{n-2} - (b_v \delta)^2\right)$$

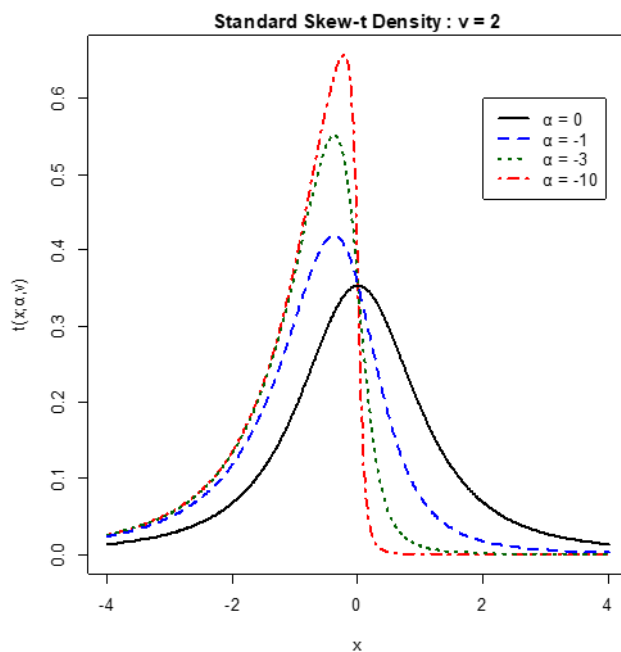
όπου

$$b_v = \left(\frac{v}{\pi}\right)^{1/2} \frac{\Gamma\left(\frac{v-1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)}, \quad v > 1.$$

Στα δύο ακόλουθα σχήματα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας της κατανομής $ST(a, v)$, για διάφορες τιμές του a .



Σχήμα 2.5. Συνάρτηση πυκνότητας της κατανομής $ST(a, 2)$, για $a = 0, 1, 3, 10$.



Σχήμα 2.6. Συνάρτηση πυκνότητας της κατανομής $ST(a, 2)$, για $a = 0, -1, -3, -10$.

Ας υποθέσουμε ότι η τυχαία μεταβλητή $Z \sim ST(a, v)$. Η συνάρτηση πυκνότητας της τυχαίας μεταβλητής

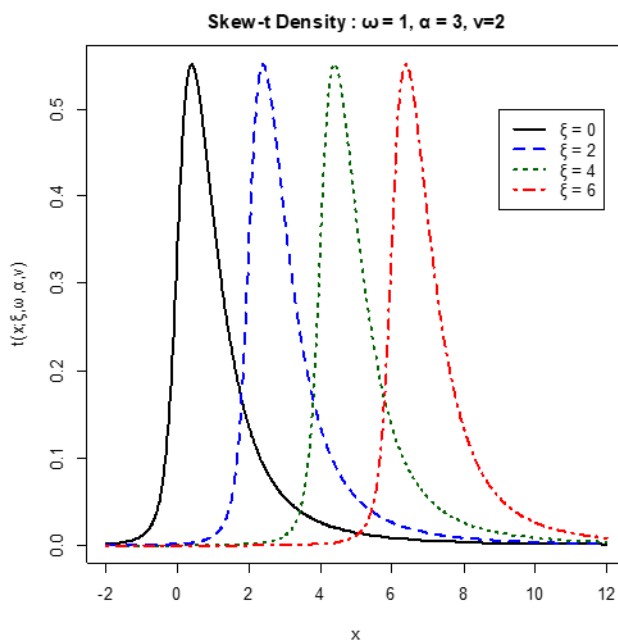
$$Y = \xi + \omega Z$$

μπορεί εύκολα να διαπιστωθεί ότι δίνεται από τον τύπο

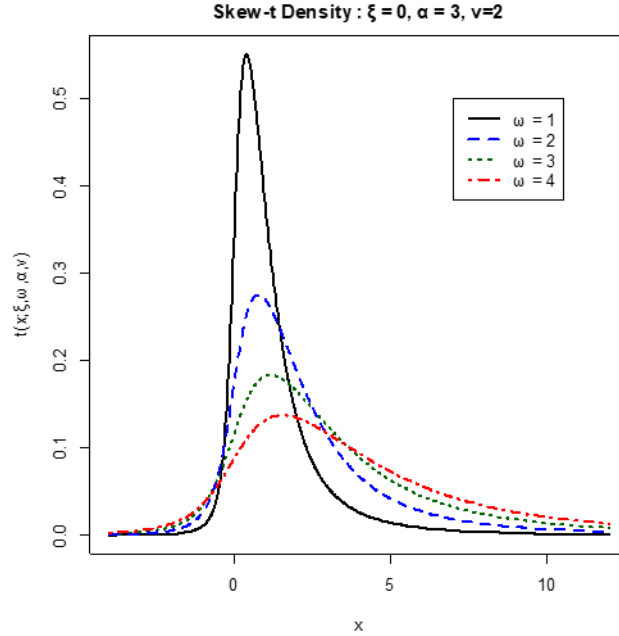
$$t(x; \xi, \omega, a, v) = \frac{1}{\omega} t\left(\frac{x - \xi}{\omega}; a, v\right) \\ = \frac{1}{\omega} 2t\left(\frac{x - \xi}{\omega}; v\right) T\left(a\left(\frac{x - \xi}{\omega}\right) \sqrt{\frac{v + 1}{v + \left(\frac{x - \xi}{\omega}\right)^2}}; v + 1\right) \quad (2.4)$$

Μια συνεχής τυχαία μεταβλητή Y θα λέμε ότι έχει τη **λοξή t κατανομή (Skew- t distribution)** με παραμέτρους $\xi \in R$, ω^2 ($\omega \in R^+$) $a \in R$ και $v > 0$ και θα γράφουμε $Y \sim ST(\xi, \omega^2, a, v)$, αν η συνάρτηση πυκνότητας της Y δίνεται από τη σχέση (2.4). Η παράμετρος ξ είναι παράμετρος θέσης, η παράμετρος ω είναι παράμετρος κλίμακας, η παράμετρος a είναι παράμετρος λοξότητας και v οι βαθμοί ελευθέριας.

Ακολουθούν δύο γραφήματα τα οποία παρουσιάζουν τη λοξή t κατανομή με διαφορετικές παραμέτρους ξ και ω το καθένα.



Σχήμα 2.7. Συνάρτηση πυκνότητας της κατανομής $ST(\xi, 1, 3, 2)$, για $\xi = 0, 2, 4, 6$.



Σχήμα 2.8. Συνάρτηση πυκνότητας της κατανομής $ST(0, \omega^2, 3, 2)$ για $\omega = 1, 2, 3, 4$.

Η λοξή t κατανομή με παραμέτρους $\xi \in R$, ω^2 ($\omega \in R^+$) $a \in R$ και $\nu > 0$ έχει μερικές ενδιαφέρουσες ιδιότητες. Μερικές από αυτές είναι οι ακόλουθες:

1. Η κατανομή $ST(\xi, \omega^2, a, \nu)$, για $\xi = 0$ και $\omega = 1$, ανάγεται στην τυπική λοξή t κατανομή $ST(a, \nu)$.
2. Η κατανομή $ST(\xi, \omega^2, a, \nu)$, για $a = 0$ ανάγεται στην κατανομή της τυχαίας μεταβλητής $\xi + \omega Z$ όπου $Z \sim t_\nu$.
3. Η κατανομή $ST(\xi, \omega^2, a, \nu)$ για $\nu \rightarrow \infty$, συγκλίνει στην κατανομή $SN(\xi, \omega^2, a)$.
4. Αν η τ.μ. $Y \sim ST(\xi, \omega^2, a, \nu)$, τότε

$$E(Y) = \xi + \omega b_\nu \delta = \xi + \omega E(Z), \quad \nu > 1$$

$$Var(Y) = \omega^2 \left(\frac{n}{n-2} - (b_\nu \delta)^2 \right) = \omega^2 Var(Z), \quad \nu > 2$$

$$\gamma_1(Y) = \gamma_1(Z) = \frac{b_\nu \delta}{\left(\frac{n}{n-2} - (b_\nu \delta)^2 \right)^{3/2}} \left(\frac{\nu(3-\delta^2)}{\nu-3} - \frac{3\nu}{\nu-2} + 2(b_\nu \delta)^2 \right), \quad \nu > 3$$

$$\gamma_2 = \frac{b_\nu \delta}{\left(\frac{n}{n-2} - (b_\nu \delta)^2 \right)^2} \times$$

$$\times \left(\frac{3v^2}{(v-2)(v-4)} - \frac{4(b_v\delta)^2 v(3-\delta^2)}{v-3} + \frac{6(b_v\delta)^2 v}{v-2} - 3(b_v\delta)^4 \right) - 3, \quad v > 4$$

Όπως και η κατανομή t_v έτσι και η κατανομή $ST(\xi, \omega^2, a, v)$ δεν έχει ροπογεννήτρια συνάρτηση.

2.3 Γέννηση και ιδιότητες της Βήτα λοξής t κατανομής

Αφού παρουσιάσαμε τα βασικά χαρακτηριστικά της λοξής κατανομής t είμαστε σε θέση να μελετήσουμε μια πιο γενικευμένη μορφή αυτής που βασίζεται στη κατανομή Βήτα. Έστω μια συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής F και συνάρτηση πυκνότητας f . Ο Jones (2004) όρισε τη συνάρτηση πυκνότητας g_F μιας τύπου Βήτα κατανομής που δίνεται από τον τύπο

$$g_F(z; a, b) = \frac{1}{B(a, b)} f(z) (F(z))^{a-1} (1 - F(z))^{b-1}$$

όπου $B(a, b)$ ($a, b > 0$) είναι η συνήθης Βήτα συνάρτηση. Η συνάρτηση κατανομής G_F που αντιστοιχεί στη συνάρτηση πυκνότητας g_F δίνεται από τον τύπο

$$G_F(z; a, b) = \frac{1}{B(a, b)} \int_0^{F(z)} u^{a-1} (1-u)^{b-1} du, \quad 0 \leq F(z) \leq 1.$$

Αν στον παραπάνω τύπο αντικαταστήσουμε τη συνάρτηση κατανομής $F(z)$ με τη συνάρτηση κατανομής $T(z; \lambda, v)$ της τυπικής λοξής t κατανομής $ST(\lambda, v)$ με παραμέτρους λ και v παίρνουμε τη συνάρτηση κατανομής

$$G(z; a, b, \lambda, v) = \frac{1}{B(a, b)} \int_0^{T(z; \lambda, v)} u^{a-1} (1-u)^{b-1} du$$

με αντίστοιχη συνάρτηση πυκνότητας

$$g(z; a, b, \lambda, v) = \frac{1}{B(a, b)} t(z; \lambda, v) (T(z; \lambda, v))^{a-1} (1 - T(z; \lambda, v))^{b-1} \quad (2.5)$$

όπου $t(z; \lambda, v)$ είναι η συνάρτηση πυκνότητας της κατανομής $ST(\lambda, v)$.

Μια συνεχής τυχαία μεταβλητή Z θα λέμε ότι έχει την **(τυπική) Βήτα λοξή t κατανομή ((standard) Beta Skew- t distribution)** με παραμέτρους $a, b, v > 0$ και $\lambda \in R$ και θα γράφουμε $Z \sim BST(a, b, \lambda, v)$, αν η συνάρτηση πυκνότητας της Z δίνεται από τον τύπο (2.5) (δείτε Basalamah et al. (2018)).

Ας υποθέσουμε ότι η τυχαία μεταβλητή $Z \sim BST(a, b, \lambda, v)$. Η τυχαία μεταβλητή

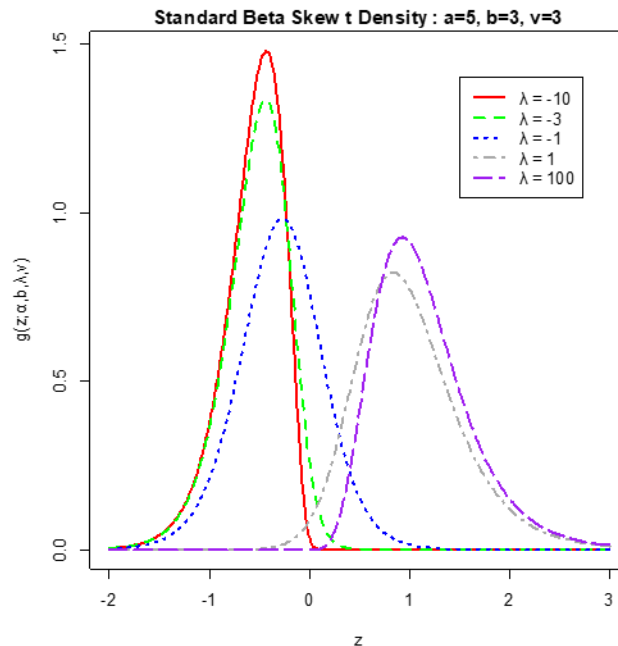
$$Y = \xi + \omega Z$$

θα λέμε ότι έχει την **Βήτα λοξή t κατανομή ((standard) Beta Skew- t distribution)** με παραμέτρους $\xi, \lambda \in R, \omega^2 (\omega \in R^+), a, b, v > 0$ και θα γράφουμε $Z \sim BST(a, b, \xi, \omega^2, \lambda, v)$.

Η τυπική Βήτα λοξή t κατανομή $BST(a, b, \lambda, v)$ έχει ενδιαφέρουσες ιδιότητες. Μερικές από αυτές είναι οι ακόλουθες:

1. Η κατανομή $BST(a, b, \lambda, v)$, για $a = b = 1$, ανάγεται στην τυπική λοξή t κατανομή $ST(\lambda, v)$.
2. Η κατανομή $BST(a, b, \lambda, v)$, για $a = b = 1$ και $\lambda = 0$, ανάγεται στην κατανομή t με v βαθμούς ελευθερίας ($Z \sim t_v$).
3. Η κατανομή $BST(a, b, \lambda, v)$, για $a = b = v = 1$ και $\lambda = 0$, ανάγεται στην κατανομή Cauchy με παραμέτρους $a = 0$ και $b = 1$ ($Z \sim Cauchy(0,1)$).

Στο παρακάτω γράφημα δίνονται κάποιες γραφικές παραστάσεις της συνάρτησης πυκνότητας της κατανομής $BST(a, b, \lambda, v)$.



Σχήμα 2.9. Συνάρτηση πυκνότητας της κατανομής $BST(5,3, \lambda, 3)$ για $\lambda = -10, -3, -1, 1, 100$.

Οι ροπές $E(Y^n)$, $n \geq v$, της κατανομής $BST(a, b, \xi, \omega^2, \lambda, v)$ ικανοποιούν τη σχέση

$$E(Y^n) = \frac{\omega^n}{B(a, b)} \sum_{j=0}^{\infty} \sum_{i=0}^n (-1)^j \binom{b-1}{j} \binom{n}{i} \left(\frac{\xi}{\omega}\right)^i E(X^{n-1}) [T(x; \lambda, v)^{a+j-1} - 1].$$

όπου $X \sim ST(\lambda, v)$.

Για $a, b \geq 2$, οι ροπές $E(Z^n)$, $v \geq n$, της κατανομής $BST(a, b, \lambda, v)$ ικανοποιούν τη σχέση

$$E(Z^n) = c(a, b)E(X^n)$$

όπου $X \sim ST(\lambda, v)$ και

$$c(a, b) = \frac{1}{B(a, b)} \left[\sum_{i=0}^{b-2} \frac{(-1)^i}{B(i+1, b-i-1)} \left[\frac{1}{a+i} - \frac{(a-1)}{(a+i-1)(b-i-1)} \right] - (-1)^{b-1} \frac{(a-1)}{a+b-2} \right].$$

Από την παραπάνω σχέση προκύπτει ότι

$$E(Y) = \delta \frac{\sqrt{v} \Gamma\left(\frac{v-1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{v}{2}\right)} c(a, b)$$

και

$$Var(Y) = v c(a, b) \left[\frac{1}{v-2} - \frac{\delta^2 \Gamma\left(\frac{v-1}{2}\right)}{\pi \Gamma\left(\frac{v}{2}\right)} c(a, b) \right]$$

όπου

$$\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}$$

Στη συνέχεια θα δώσουμε βασικές σχέσεις με τις οποίες μπορούν να προκύψουν εκτιμητές μέγιστης πιθανοφάνειας για τις παραμέτρους της Βήτα λοξής t κατανομής. Έστω λοιπόν Y_1, Y_2, \dots, Y_n ένα τυχαίο δείγμα μεγέθους n από την κατανομή $BST(a, b, \xi, \omega^2, \lambda, v)$. Για $\theta = (a, b, \xi, \omega, \lambda, v)$. Ο λογάριθμος της συνάρτησης πιθανοφάνειας δίνεται από τον τύπο

$$l(\theta) = n \log(\Gamma(a+b)) - n \log(\Gamma(a)) - n \log(\Gamma(b)) - n \log(\omega) + \sum_{i=1}^n \log(t(z_i; \xi, \omega, \lambda, v))$$

$$+(a-1) \sum_{i=1}^n \log(T(z_i; \xi, \omega, \lambda, \nu)) + (b-1) \sum_{i=1}^n \log(1 - T(z_i; \xi, \omega, \lambda, \nu))$$

όπου $t(z_i; \xi, \omega, \lambda, \nu)$ και $T(z_i; \xi, \omega, \lambda, \nu)$ η συνάρτηση πυκνότητας και η συνάρτηση κατανομής της λοξής t κατανομής $ST(\xi, \omega^2, \lambda, \nu)$, και $z_i = \frac{y_i - \xi}{\omega}$, $i = 1, 2, \dots, n$. Επομένως οι συνιστώσες του διανύσματος σκορ (score) είναι οι ακόλουθες:

$$\frac{\partial l(a, b, \xi, \omega, \lambda, \nu)}{\partial a} = n\psi(a+b) - n\psi(a) + \sum_{i=1}^n \log(T(z_i; \xi, \omega, \lambda, \nu))$$

$$\frac{\partial l(a, b, \xi, \omega, \lambda, \nu)}{\partial b} = n\psi(a+b) - n\psi(b) + \sum_{i=1}^n \log(1 - T(z_i; \xi, \omega, \lambda, \nu))$$

$$\begin{aligned} \frac{\partial l(a, b, \xi, \omega, \lambda, \nu)}{\partial \xi} &= \sum_{i=0}^n \frac{-1}{\omega t\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)} \frac{dt\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)}{d\xi} \\ &\quad - \frac{(a-1)}{\omega} \sum_{i=0}^n \frac{1}{T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)} \frac{dF\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)}{d\xi} \\ &\quad + \frac{(b-1)}{\omega} \sum_{i=0}^n \frac{1}{(1 - T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right))} \frac{d(1 - T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right))}{d\xi} \end{aligned}$$

$$\begin{aligned} \frac{\partial l(a, b, \xi, \omega, \lambda, \nu)}{\partial \omega} &= -\frac{n}{\omega} + \sum_{i=0}^n \frac{1}{\omega t\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)} \frac{dt\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)}{d\omega} \\ &\quad - \frac{(a-1)}{\omega} \sum_{i=0}^n \frac{1}{T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)} \frac{dF\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right)}{d\omega} \\ &\quad + \frac{(b-1)}{\omega} \sum_{i=0}^n \frac{1}{(1 - T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right))} \frac{d(1 - T\left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, \nu\right))}{d\omega} \end{aligned}$$

$$\begin{aligned} \frac{\partial l(a, b, \xi, \omega, \lambda, v)}{\partial \lambda} &= \sum_{i=0}^n \frac{-1}{t \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right)} \frac{dt \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right)}{d\lambda} \\ &\quad - (\alpha - 1) \sum_{i=0}^n \frac{1}{T \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right)} \frac{dF \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right)}{d\lambda} \\ &\quad + (b - 1) \sum_{i=0}^n \frac{1}{(1 - T \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right))} \frac{d(1 - T \left(\frac{x_i - \xi}{\omega}; \xi, \omega, \lambda, v \right))}{d\lambda} \end{aligned}$$

όπου $\psi(x)$ είναι η συνάρτηση δίζαμμα που ορίζεται από τον τύπο $\psi(x) = \frac{d}{dx} \log \Gamma(x)$.

Ένα σημαντικό πλεονέκτημα της κατανομής *BST* είναι ότι μπορεί να προσαρμοστεί ακόμα καλύτερα σε δεδομένα που παρουσιάζουν λοξότητα ή αρκετά βαριές ουρές σε σχέση με την λοξή t κατανομή αφού έχει περισσότερες παραμέτρους και περιέχει τη λοξή t κατανομή ως ειδική περίπτωση.

2.4 Γέννηση και ιδιότητες της λοξής λογιστικής κατανομής

Ας συμβολίσουμε με $f(x; \beta)$ και $F(x; \beta)$ τη συνάρτηση πυκνότητας και τη συνάρτηση κατανομής της λογιστικής κατανομής $\text{logis}(0, \beta)$, δηλαδή

$$f(x; \beta) = \frac{\exp \left(-\frac{x}{\beta} \right)}{\beta \left(1 + \exp \left(-\frac{x}{\beta} \right) \right)^2}, \quad x \in \mathbb{R}$$

και

$$F(x; \beta) = \frac{1}{1 + \exp \left(-\frac{x}{\beta} \right)}, \quad x \in \mathbb{R}.$$

Θέτοντας στο γνωστό τύπο

$$f(x) = 2f_0(x)G_0(w(x)), \quad x \in \mathbb{R},$$

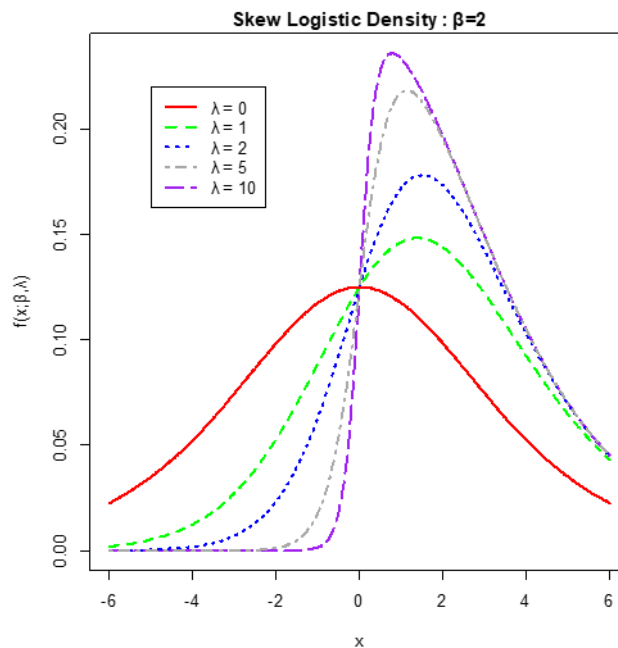
$f_0(x) = f(x; \beta)$, $G_0(x) = \Phi(x; \beta)$ και $w(x) = \lambda x$, προκύπτει η συνάρτηση πυκνότητας

$$f(x; \beta, \lambda) = \frac{2 \exp\left(-\frac{x}{\beta}\right)}{\beta \left(1 + \exp\left(-\frac{x}{\beta}\right)\right)^2 \left(1 + \exp\left(-\frac{\lambda x}{\beta}\right)\right)}, \quad x \in \mathbb{R}, \quad (2.6)$$

όπου $\beta > 0$ και $\lambda \in \mathbb{R}$.

Μια συνεχής τυχαία μεταβλητή X θα λέμε ότι έχει τη **λοξή λογιστική κατανομή (skew logistic distribution)** με παραμέτρους $\beta > 0$ και $\lambda \in \mathbb{R}$ και θα γράφουμε $X \sim SL(\beta, \lambda)$, αν η συνάρτηση πυκνότητας της X δίνεται από τον τύπο (2.6) (δείτε Nadarajah (2009)).

Στο παρακάτω γράφημα δίνονται κάποιες γραφικές παραστάσεις της συνάρτησης πυκνότητας της κατανομής $SL(\beta, \lambda)$.



Σχήμα 2.10. Συνάρτηση πυκνότητας της κατανομής $SL(2, \lambda)$ για $\lambda = 0, 1, 2, 5, 10$.

Μπορούμε να διαπιστώσουμε ότι η λοξή λογιστική κατανομή έχει κορυφή στη θέση $x = -\beta \log(y_0)$, όπου y_0 η λύση της εξίσωσης

$$1 - y + (1 - \lambda)y^\lambda - (1 + \lambda)y^{1+\lambda} = 0.$$

Για τη συνάρτηση κατανομής $F(x; \beta, \lambda)$ της κατανομής $SL(\beta, \lambda)$ έχει προταθεί ο τύπος

$$F(x; \beta, \lambda) = \begin{cases} 1 - 2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \binom{-1}{j} \binom{-2}{k} \frac{1}{1 + \lambda j + k} \exp\left(-\frac{(1 + \lambda j + k)x}{\beta}\right), & x > 0 \\ 2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \binom{-1}{j} \binom{-2}{k} \frac{1}{1 + \lambda + \lambda j + k} \exp\left(-\frac{(1 + \lambda + \lambda j + k)x}{\beta}\right), & x < 0. \end{cases}$$

Για τις ροπές της της κατανομής $SL(\beta, \lambda)$ έχουμε ότι αν το n είναι άρτιος τότε

$$E(X^n) = 2n! \beta^n (1 - 2^{1-n}) \zeta(n)$$

όπου $\zeta(n) = \sum_{j=1}^{\infty} j^{-n}$ είναι η συνάρτηση ζήτα του Riemann.

Αν το n είναι περιττός τότε

$$E(X^n) = 2n! \beta^n \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \binom{-1}{j} \binom{-2}{k} \left[\frac{1}{(1 + \lambda j + k)^{n+1}} + \frac{(-1)^n}{(1 + \lambda + \lambda j + k)^{n+1}} \right].$$

Έτσι

$$E(X) = \frac{\beta}{\lambda^2} \sum_{j=0}^{\infty} (-1)^j (j+1) \left[\sum_{i=0}^{\infty} \frac{1}{\left(\frac{1 + 2\lambda + j}{2\lambda} + i\right)^2} - \sum_{i=0}^{\infty} \frac{1}{\left(\frac{1 + \lambda + j}{2\lambda} + i\right)^2} \right]$$

και

$$E(X^2) = \frac{(\pi\beta)^2}{3}.$$

Μια γενίκευση της λοξής λογιστικής κατανομής με παραμέτρους $\beta > 0$ και $\lambda \in \mathbb{R}$ είναι η «μετατοπισμένη» λοξή λογιστική κατανομή με παραμέτρους $m \in \mathbb{R}$, $\beta > 0$ και $\lambda \in \mathbb{R}$ (συμβολισμός $SL(m, \beta, \lambda)$), με συνάρτηση πυκνότητας που δίνεται από τον τύπο

$$f(x; \beta, \lambda) = \frac{2 \exp\left(-\frac{x-m}{\beta}\right)}{\beta \left(1 + \exp\left(-\frac{x-m}{\beta}\right)\right)^2 \left(1 + \exp\left(-\frac{\lambda(x-m)}{\beta}\right)\right)}, \quad x \in \mathbb{R}.$$

Έστω τώρα ότι έχουμε ένα τυχαίο δείγμα μεγέθους n , το X_1, X_2, \dots, X_n , από την κατανομή $SL(m, \beta, \lambda)$. Ο λογάριθμος της συνάρτησης πιθανοφάνειας δίνεται από το τύπο

$$\log L(m, \beta, \lambda) = n \log 2 - n \log \beta - \frac{1}{\beta} \sum_{j=1}^n (x_j - m) - 2 \sum_{j=1}^n \log \left(1 + \exp \left(-\frac{x_j - m}{\beta} \right) \right) - \sum_{j=1}^n \log \left(1 + \exp \left(-\frac{\lambda(x_j - m)}{\beta} \right) \right)$$

Από τις ακόλουθες εξισώσεις προκύπτουν εκτιμητές μέγιστης πιθανοφάνειας των παραμέτρων της λοξής λογιστικής κατανομής

$$2 \sum_{j=1}^n \frac{\exp \left(-\frac{x_j - m}{\beta} \right)}{1 + \exp \left(-\frac{x_j - m}{\beta} \right)} + \lambda \sum_{j=1}^n \frac{\exp \left(-\frac{\lambda(x_j - m)}{\beta} \right)}{1 + \exp \left(-\frac{\lambda(x_j - m)}{\beta} \right)} = n$$

$$2 \sum_{j=1}^n \frac{(x_j - m) \exp \left(-\frac{x_j - m}{\beta} \right)}{1 + \exp \left(-\frac{x_j - m}{\beta} \right)} = -n\beta + \sum_{j=1}^n (x_j - m)$$

$$\sum_{j=1}^n \frac{(x_j - m) \exp \left(-\frac{\lambda(x_j - m)}{\beta} \right)}{1 + \exp \left(-\frac{\lambda(x_j - m)}{\beta} \right)} = 0$$

2.5 Γέννηση και ιδιότητες της λοξής Cauchy κατανομής

Αν έχουμε δύο ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, έστω X και Y , που κάθε μια έχει την τυπική κανονική κατανομή $N(0,1)$, τότε γνωρίζουμε ότι οι τυχαίες μεταβλητές Y/X και $Y/|X|$ έχουν την κατανομή $Cauchy(0,1)$. Αν αντικαταστήσουμε την τυχαία μεταβλητή Y με την τυχαία μεταβλητή $Z \sim SN(a)$, τότε η κατανομή της τυχαίας μεταβλητής

$$W = \frac{Z}{|X|}$$

ονομάζεται **τυπική λοξή Cauchy κατανομή ((standard) Skew-Cauchy distribution)** με παράμετρο $a \in R$ και θα γράφουμε $W \sim SC(a)$.

Η κατανομή $SC(a)$ αποτελεί ειδική περίπτωση της λοξής t κατανομής που ορίσαμε στην Παράγραφο 2.2. Μπορεί εύκολα να διαπιστωθεί ότι η κατανομή $SC(a)$ προκύπτει από την κατανομή $ST(a, v)$ για $v = 1$.

Για τη συνάρτηση κατανομής $G(w; a)$ της τυχαίας μεταβλητής W έχουμε

$$G(w; a) = P\left(\frac{Z}{|X|} \leq w\right) = P(Z \leq w|X|) = E[\Phi(w|X|; a)] = 2 \int_0^{\infty} \Phi(wx; a) \varphi(x) dx.$$

Με παραγωγή της συνάρτησης κατανομής $G(w; a)$ και μετά με ολοκλήρωση κατά παράγοντες παίρνουμε ότι η συνάρτηση πυκνότητας $g(w; a)$ της τυχαίας μεταβλητής W δίνεται από τον τύπο

$$g(w; a) = \frac{1}{\pi(1+w^2)} \left(1 + \frac{aw}{\sqrt{1+(1+a^2)w^2}}\right), \quad -\infty < w < \infty \quad (2.7)$$

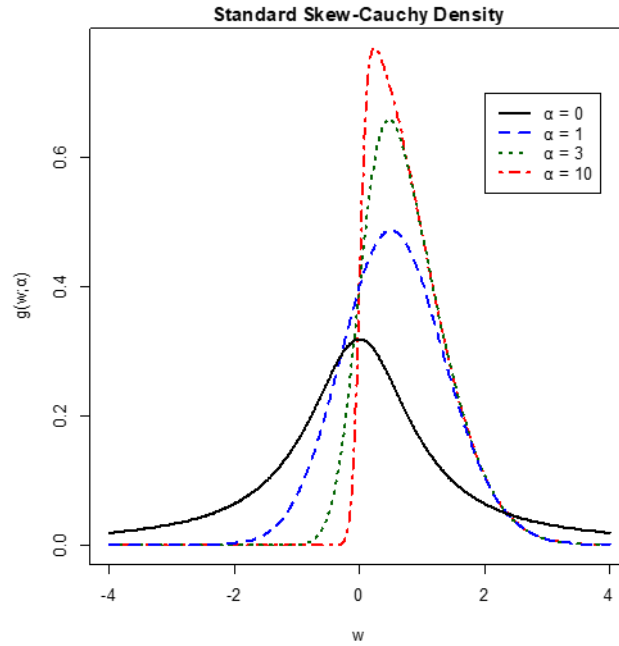
Ένας εναλλακτικός τύπος για τη συνάρτηση κατανομής $G(w; a)$ της τυχαίας μεταβλητής W (δείτε Behboodian et al. (2006)) είναι ο ακόλουθος

$$G(x; a) = \frac{1}{\pi} \left(\arctan w + \arccos \frac{a}{\sqrt{(1+a^2)(1+w^2)}} \right).$$

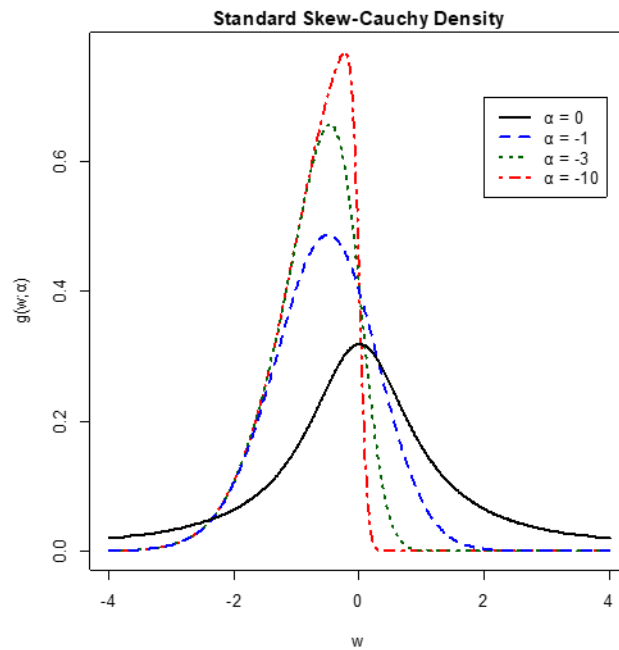
Η τυπική λοξή Cauchy κατανομή $SC(a)$ έχει μερικές ενδιαφέρουσες ειδικές περιπτώσεις. Μερικές από αυτές είναι οι ακόλουθες:

1. Η κατανομή $SC(a)$, για $a = 0$, ανάγεται στην κατανομή $Cauchy(0,1)$.
2. Η κατανομή $SC(a)$, για $a \rightarrow \infty$, συγκλίνει στην κατανομή $Cauchy(0,1)$.
3. Αν $W \sim SC(a)$, τότε $-W \sim SC(-a)$.

Στα δύο ακόλουθα σχήματα δίνονται γραφικές παραστάσεις της συνάρτησης πυκνότητας της κατανομής $SC(a)$ για διάφορες τιμές του a .



Σχήμα 2.11. Συνάρτηση πυκνότητας της κατανομής $SC(a)$ για $a = 0, 1, 3, 10$.



Σχήμα 2.12. Συνάρτηση πυκνότητας της κατανομής $SC(a)$ για $a = 0, -1, -3, -10$.

ΚΕΦΑΛΑΙΟ 3

Εφαρμογές σε πραγματικά δεδομένα

3.1 Εισαγωγή

Στην ενότητα αυτή θα προσαρμόσουμε τη λοξή κανονική κατανομή (skew normal) και λοξή t κατανομή (skew Students t) σε δύο σύνολα δεδομένων ζημιών και ταυτόχρονα θα τις συγκρίνουμε με άλλες γνωστές κατανομές.

Η ανάλυσή μας θα γίνει με τη χρήση της γλώσσας προγραμματισμού R. Τα πιο σημαντικά εργαλεία για τη μελέτη των λοξών κατανομών τα λαμβάνουμε από το πακέτο sn της R. Για συγκρίσεις με άλλες κατανομές και τεστ καλής προσαρμογής θα χρησιμοποιήσουμε μοντέλα που αποτελούν σημεία αναφοράς στη χρηματοοικονομική θεωρία και στην αναλογιστική επιστήμη συγκεκριμένα για την αξιολόγηση της προσαρμογής θα χρησιμοποιήσουμε τον έλεγχο Kolmogorov – Smirnov και το κριτήριο του Akaike . Τέλος θα υπολογίσουμε κάποια σημαντικά μέτρα κινδύνου όπως αυτό της αξία σε κίνδυνο (value at risk – VaR) και της ουράς της αξίας σε κίνδυνο (tail value at risk – TVaR).

3.2 Τα δεδομένα

Το πρώτο σετ δεδομένων (Danish) αποτελείται από ζημιές που προκλήθηκαν λόγω πυρκαγιάς στη Δανία. Τα δεδομένα συλλέχθηκαν από την ανασφαλιστική εταιρεία Copenhagen Re και αφορούν το χρονικό διάστημα από 3 Ιανουαρίου 1980 έως 31 Δεκεμβρίου 1990. Τα δεδομένα αποτελούνται από 2167 ζημιές όπου κάθε ζημιά αφορά μεμονωμένη απώλεια άνω του ενός εκατομμύριου δανέζικων κορόνων. Τα δεδομένα αυτά βρίσκονται στο πακέτο fExtremes της R.

Το δεύτερο σετ δεδομένων (Expert Advisor) αφορά ζημιές σε ένα επενδυτικό χαρτοφυλάκιο από ένα αυτοματοποιημένο σύστημα συναλλαγών σε νομίσματα βασισμένο σε Martingale και με απόδοση ιδίων κεφαλαίων σχεδόν 200%. Τέτοια συστήματα είναι πολύ

συνηθισμένα και τείνουν να αντικαθιστούν τον άνθρωπο τα τελευταία χρόνια αφού μελέτες πάνω στη ψυχολογία των αγορών έχουν αποδείξει ότι συναισθήματα όπως απληστία και φόβος οδηγούν τους επενδυτές στο να λαμβάνουν λανθασμένες αποφάσεις. Έχουμε 210 παρατηρήσεις από ζημιές σε αγγλική λίρα (GBP) με βαριά ουρά που προκλήθηκαν τα τελευταία δυο χρόνια.

3.3 Ανάλυση των δεδομένων

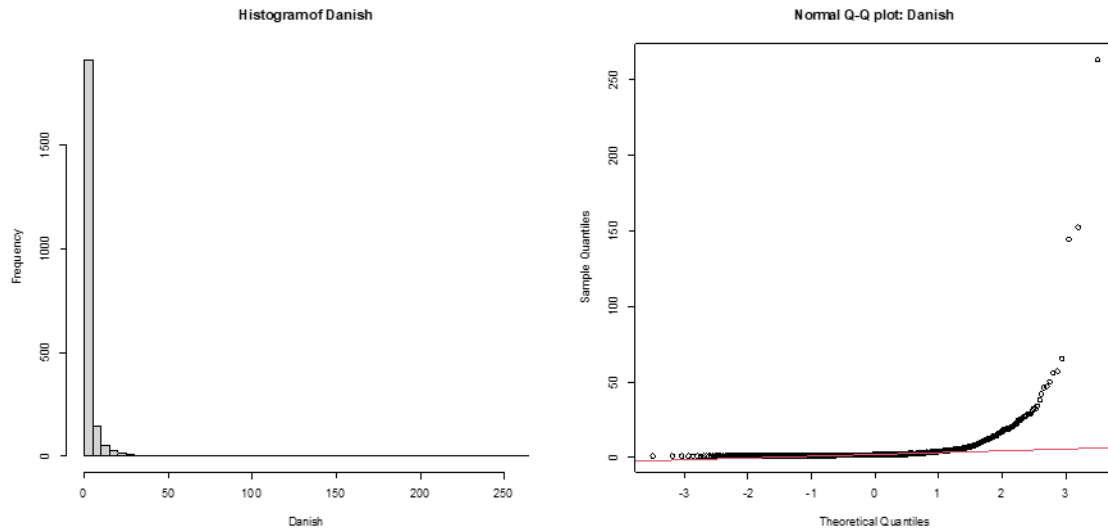
Στο Πίνακα 3.1 παρουσιάζονται βασικά περιγραφικά μέτρα των δυο σετ δεδομένων.

	Danish	Expert Advisor
Μέση Τιμή	3,385	1284,598
Τυπική Απόκλιση	8,507	2729,768
Λοξότητα	18,737	5,295
Κύρτωση	482,198	32,323
Ελάχιστο	1,00	31,1
Μέγιστο	263,250	21809,6
99% Ποσοστημόριο ($VaR_{0.99}$)	26,043	15164,13
$E(X X > VaR_{0.99})$ ($TVaR_{0.99}$)	58,586	19623,96

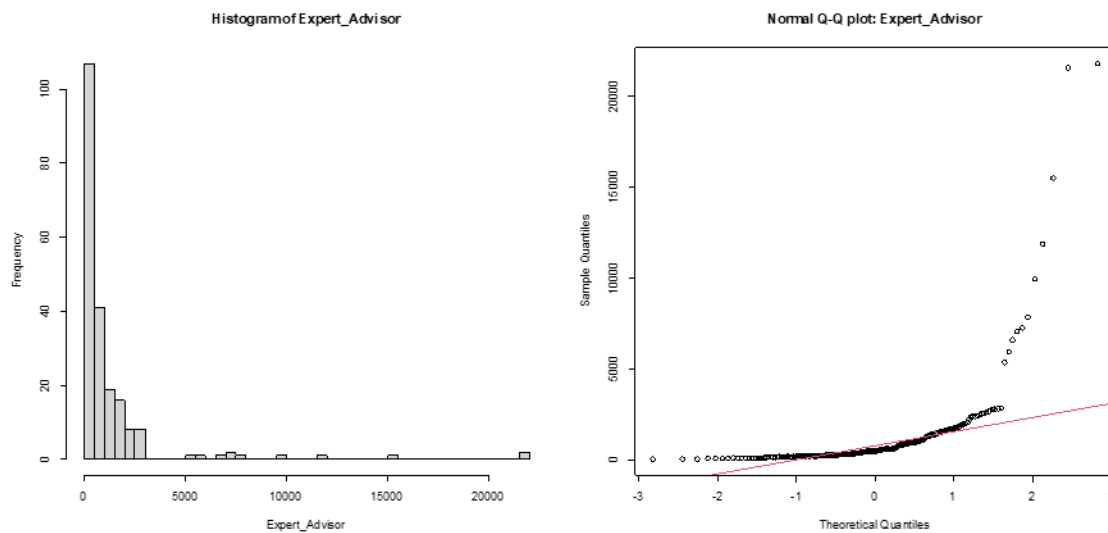
Πίνακας 3.1. Περιγραφικά μέτρα των δεδομένων

Από το παραπάνω πίνακα είναι εμφανές ότι η λοξότητα και η κύρτωση των δεδομένων Danish είναι πολύ πιο μεγάλη από τις αντίστοιχες τιμές των δεδομένων Expert Advisor. Την παρατήρηση αυτή την επιβεβαιώνει και η πολύ μεγάλη ποσοστιαία διαφορά που υπάρχει μεταξύ του $VaR_{0.99}$ και του $TVaR_{0.99}$ ($E(X|X > VaR_{0.99})$) των κάθε σετ δεδομένων.

Ακολούθως δίνονται ιστογράμματα και κανονικά διαγράμματα πιθανότητας για τα δύο σύνολα δεδομένων όπου είναι εμφανής η έντονη λοξότητα προς τα δεξιά των δεδομένων και η μη κανονικότητά τους. Παρατηρούνται πολλές μικρού ύψους ζημιές και λίγες μεγάλου μεγέθους.



Σχήμα 3.1. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για τα δεδομένα Danish



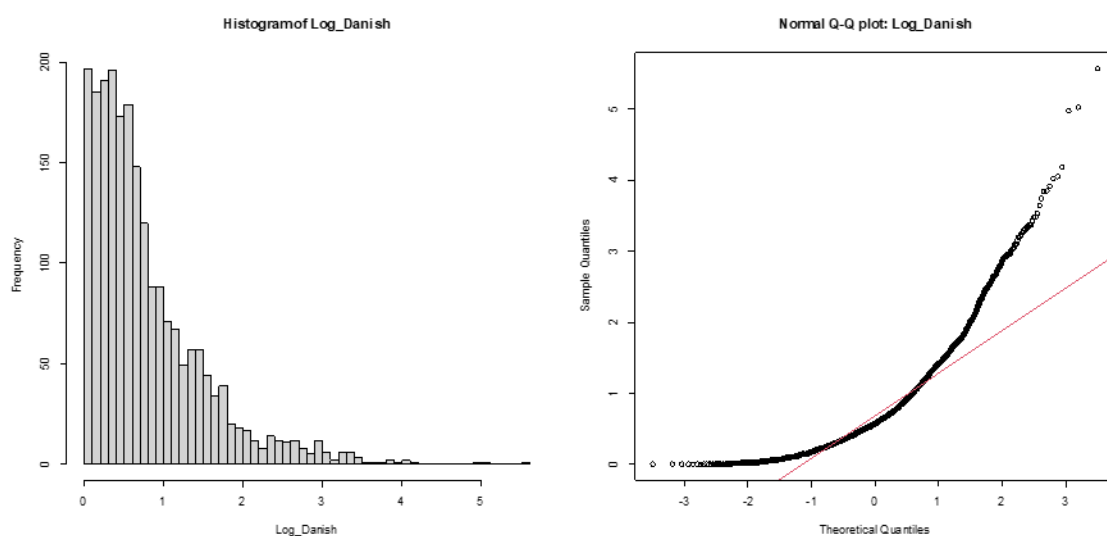
Σχήμα 3.2. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για τα δεδομένα Expert Advisor

Μια πολύ γνωστή τεχνική μοντελοποίησης της στατιστικής επιστήμης προκειμένου να ελαττωθούν ακραίες τιμές της κύρτωσης και της λοξότητας των δεδομένων (και όχι μόνο) είναι η λογαριθμοποίηση των δεδομένων. Βασικά περιγραφικά μέτρα των λογαριθμημένων δεδομένων δίνονται στον Πίνακα 3.2.

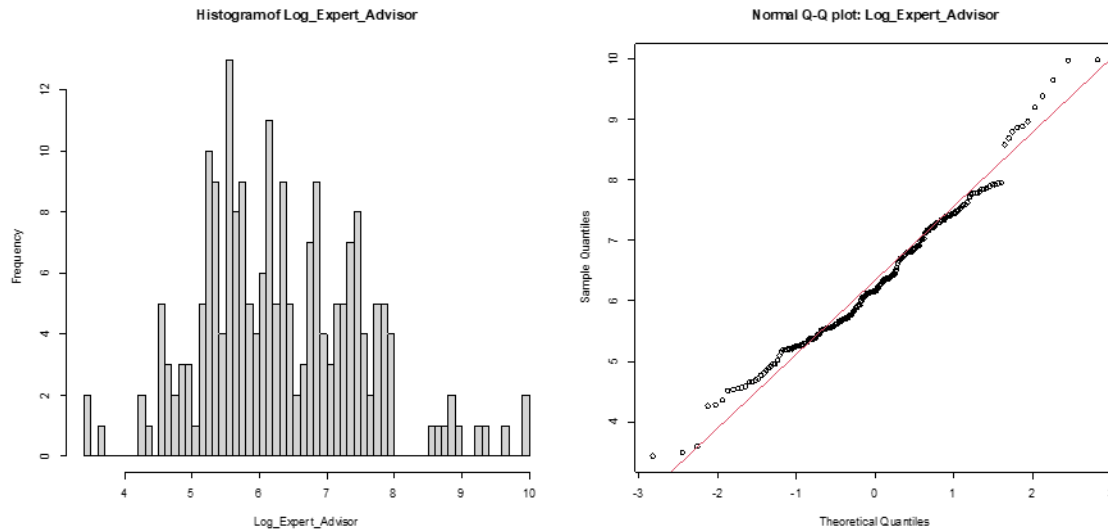
	Log Danish	Log Expert Advisor
Μέση Τιμή	0,787	6,328
Τυπική Απόκλιση	0,717	1,181
Λοξότητα	1,761	0,471
Κύρτωση	4,179	0,360
Ελάχιστο	0,000	3,437
Μέγιστο	5,573	9,990
99% Ποσοστημόριο (VaR _{0.99})	3,260	9,624
E(X X> VaR _{0.99}) (TVaR _{0.99})	3,820	9,872

Πίνακας 3.2. Περιγραφικά μέτρα των λογαριθμημένων δεδομένων.

Η λοξότητα και η κύρτωση στα δυο σετ δεδομένων έχει εξομαλυνθεί σε μεγάλο βαθμό όπως επίσης και η ποσοστιαία διαφορά του VaR_{0.99} και του TVaR_{0.99} έχει ελαττωθεί σημαντικά. Ενδεικτικά είναι τα δύο επόμενα διαγράμματα.



Σχήμα 3.3. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για το λογάριθμο των δεδομένων Danish



Σχήμα 3.4. Ιστόγραμμα και κανονικό Q-Q -διάγραμμα για το λογάριθμο των δεδομένων Expert Advisor

Σημειώνουμε ότι το κανονικό διάγραμμα πιθανότητας του λογάριθμου των δεδομένων Expert Advisor ταιριάζει περισσότερο με την εικόνα μιας κανονικής κατανομής. Ωστόσο η εντύπωση αυτή δεν επιβεβαιώνεται με ελέγχους υποθέσεων που δοκιμάστηκαν (ks.test, ad.test, shapiro.test, κλπ.)

Συνεχίζουμε βρίσκοντας εκτιμητές μεγίστης πιθανοφάνειας για τις παραμέτρους της λοξής κανονικής και της λοξής t κατανομής για τα αρχικά αλλά και για τα λογαριθμημένα δεδομένα.

	Danish	Expert Advisor	Log Danish	Log Expert Advisor
Λοξή Κανονική				
Παράμετρος Θέσης (ξ)	0,909	10,463	-0,0027	5,120
Παράμετρος Κλίμακας (ω)	8,844	3001,768	1,065	1,687
Παράμετρος Λοξότητας (α)	183,45	183,446	183,44	2,039
Λοξή t				
Παράμετρος Θέσης (ξ)	0,999	70,506	-3,54 x 10 ⁻⁸	5,180
Παράμετρος Κλίμακας (ω)	0,824	456,621	0,819	1,580
Παράμετρος Λοξότητας (α)	2912758	15,573	304766300	1,877
Βαθμοί Ελευθερίας (ν)	1,10	1,11	4,6	25,150

Πίνακας 3.3. Εκτιμητές μεγίστης πιθανοφάνειας λοξής κανονικής και t κατανομής

Παρατηρούμε ότι οι βαθμοί ελευθερίας της λοξής t κατανομής είναι πολύ χαμηλοί για τα αρχικά δεδομένα. Πιο συγκεκριμένα για τα δεδομένα Danish Fire είναι 1,1 ενώ για τα Expert Advisor είναι 1,11. Αυτό σημαίνει ότι δεν υπάρχουν οι ροπές αφού γνωρίζουμε ότι θα έπρεπε οι βαθμοί ελευθερίας να είναι τουλάχιστον 2 προκειμένου να υπάρχουν (τουλάχιστον 3 βαθμοί ελευθερίας για να υπάρχει ο συντελεστής ασυμμετρίας, τουλάχιστον 4 βαθμοί ελευθερίας για να υπάρχει ο συντελεστής κύρτωσης, κλπ.). Έτσι εξαιτίας του μικρού αριθμού βαθμών ελευθερίας στην περίπτωση που το προσαρμοσμένο μοντέλο χρησιμοποιηθεί σε μελέτες προσομοίωσης, οι προκύπτοντες τυχαίοι αριθμοί δεν θα οδηγούν σε σταθερές τιμές για τη μέση τιμή, την τυπική απόκλιση το συντελεστή λοξότητας, κτλ. Θα επηρεαστούν και οι ουρές των κατανομών αυτών που θα είναι πιο φαρδιές και άρα έμμεσα και οι τιμές VaR και TVaR. Τέτοιου είδους προβλήματα συνήθως αντιμετωπίζονται θέτοντας μια σταθερή τιμή στους βαθμούς ελευθερίας, έστω 4, πράγμα που οδηγεί όμως σε μείωση της καλής προσαρμογής του νέου προσαρμοσμένου μοντέλου έναντι του αρχικού.

Ωστόσο τα παραπάνω προβλήματα δεν υφίστανται για τα λογαριθμημένα δεδομένα καθώς οι βαθμοί ελευθερίας και στα δυο σετ είναι περισσότεροι από 4. Θυμίζουμε ότι όσο αυξάνονται οι βαθμοί ελευθερίας στη κατανομή t τόσο αυτή τείνει προς τη κανονική κατανομή.

Στη συνέχεια, για λόγους σύγκρισης, προσαρμόζουμε και άλλες γνωστές κατανομές στα δεδομένα μας με τη μέθοδο μεγίστης πιθανοφάνειας.

	Danish	Expert Advisor	Log Danish	Log Expert Advisor
Lognormal				
Μέση Τιμή	0,787	6,328	-0.793	1,828
Τυπική Απόκλιση	0,717	1,178	1.691	0,186
Κανονική				
Μέση Τιμή	3,385	1284,598	0,787	6,328
Τυπική Απόκλιση	8,505	2723,260	0,717	1,178
Weibull				
Παράμετρος Θέσης	0,959	0,766	1,056	5,498
Παράμετρος Κλίμακας	3,292	1034,687	0,803	6,822
Cauchy				
Παράμετρος Θέσης	1,604	380,582	0,512	6,146
Παράμετρος Κλίμακας	0,506	285,343	0,313	0,723
Λογιστική				
Παράμετρος Θέσης	2,351	803,068	0,682	6,273
Παράμετρος Κλίμακας	1,594	785,776	0,366	0,666

Πίνακας 3.4. Εκτιμητές μεγίστης πιθανοφάνειας για 5 επιλεγμένες κατανομές.

Προκειμένου να μπορέσουμε να καταλάβουμε αν η θεωρητική κατανομή εφαρμόζει καλά και ταιριάζει στα εμπειρικά δεδομένα θα εφαρμόσουμε το τεστ καλής προσαρμογής Kolmogorov – Smirnov χρησιμοποιώντας επίπεδο σημαντικότητας 0,05 στο οποίο αντιστοιχεί η προσεγγιστική κρίσιμη τιμή $1.36/\sqrt{n}$ (όπου n ο αριθμός των παρατηρήσεων). Αν η τιμή της στατιστικής συνάρτησης πέσει κάτω από την τιμή $1.36/\sqrt{n}$ δεν μπορούμε να απορρίψουμε την υπόθεση ότι τα δεδομένα προέρχονται από τη συγκεκριμένη κατανομή. Στα συγκεκριμένα δεδομένα η κρίσιμη τιμή για τα δεδομένα Danish είναι 0.0292 ενώ για τα Expert Advisor 0.0938. Στον ακόλουθο πίνακα παρουσιάζονται τα αποτελέσματα του τεστ όπου με παχιά γράμματα σημειώνουμε τις τιμές που είναι μικρότερες από την εκάστοτε κρίσιμη τιμή.

Κατανομή	Danish	Expert Advisor	Log Danish	Log Expert Advisor
Λοξή Κανονική	0.574	0.454	0.098	0.046
Λοξή t	0.002	0.057	0.024	0.047
Lognormal	0.137	0.067	0.993	0.049
Κανονική	0.390	0.323	0.137	0.067
Weibull	0.273	0.117	0.038	0.082
Cauchy	0.222	0.223	0.175	0.105
Λογιστική	0.299	0.272	0.134	0.065

Πίνακας 3.5. Αποτελέσματα Kolmogorov – Smirnov τεστ.

Από τα αποτελέσματα του τεστ Kolmogorov – Smirnov προκύπτει ότι δεν μπορούμε να απορρίψουμε ότι

(α) τα δεδομένα Danish προέρχονται από την λοξή κατανομή t

(β) τα δεδομένα Expert Advisor προέρχονται είτε από λοξή κατανομή t , είτε από λογαριθμική,

(γ) τα δεδομένα Log Danish προέρχονται από την λοξή κατανομή t , και

(δ) τα δεδομένα Log Expert Advisor προέρχονται από κάθε μια από τις κατανομές πλην της κατανομής Cauchy.

Ενώ το τεστ Kolmogorov – Smirnov μας δίνει μια εικόνα για το αν ταιριάζει μια θεωρητική κατανομή στα δεδομένα μας, η σύγκριση μεταξύ των θεωρητικών κατανομών και η επιλογή της πλέον κατάλληλης από αυτές μπορεί να στηριχθεί στην τιμή του κριτηρίου του Akaike. Υπενθυμίζουμε ότι

$$AIC = -2\log\hat{L} + 2 \cdot K$$

όπου K ο αριθμός των παραμέτρων της κατανομής που προσαρμόζεται στα δεδομένα. Η πλέον κατάλληλη κατανομή είναι αυτή με τη μικρότερη τιμή AIC . Αναλυτικά αποτελέσματα παρουσιάζονται στους επόμενους πίνακες όπου με παχιά γράμματα σημειώνουμε τη μικρότερη τιμή AIC .

Κατανομή	Danish		Expert Advisor	
	Log-likelihood	AIC	Log-likelihood	AIC
Λοξή Κανονική	-6301,18	12608,36	-1834,47	3674,94
Λοξή t	-3337,53	6683,05	-1659,97	3327,95
Lognormal	-4057,90	8119,80	-1661,27	3326,53
Κανονική	-7713,76	15431,52	-1958,99	3921,98
Λοξή Κανονική	-4803,62	9611,24	-1693,70	3391,39
Cauchy	-4118,09	8240,17	-1730,34	3464,67
Λογιστική	-5737,86	11479,71	-1855,14	3714,28

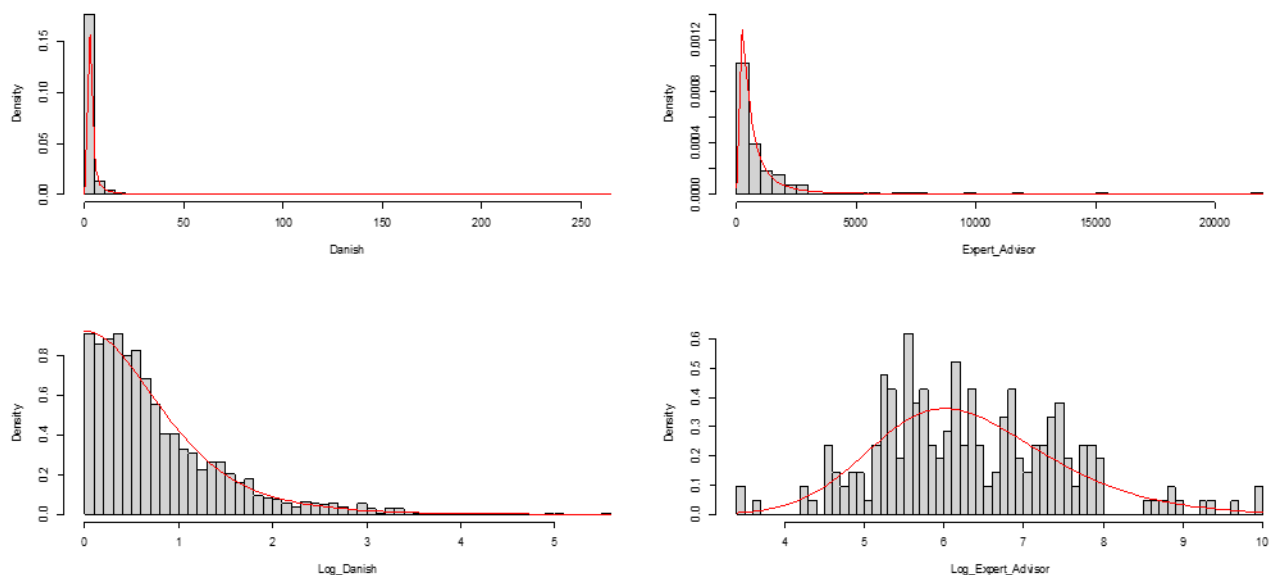
Πίνακας 3.6. Τιμές AIC για τα δεδομένα Danish και Expert Advisor

Κατανομή	Log Danish		Log Expert Advisor	
	Log-likelihood	AIC	Log-likelihood	AIC
Λοξή Κανονική	-1717,74	3441,49	-328,45	662,89
Λοξή t	-1633,10	3274,22	-328,28	664,56
Lognormal	-2495,59	4995,18	-329,12	662,24
Κανονική	-2352,57	4709,15	-332,40	668,80
Weibull	-1642,68	3289,36	-342,68	689,35
Cauchy	-2292,69	4589,38	-368,81	741,61
Λογιστική	-2208,59	4421,17	-332,61	669,23

Πίνακας 3.7. Τιμές AIC για τα δεδομένα Log Danish και Log Expert Advisor

Τα συμπεράσματα που προκύπτουν είναι ότι οι λοξές κατανομές έχουν ιδιαίτερα καλά αποτελέσματα συγκριτικά με τις υπόλοιπες και κυρίως η λοξή t κατανομή η οποία δίνει με διαφορά τα καλύτερα αποτελέσματα και στα δυο σετ δεδομένων. Ομοίως στα λογαριθμημένα δεδομένα η λοξή t έχει τα καλύτερα αποτελέσματα. Παρατηρούμε όμως ότι στο λογαριθμημένο σετ δεδομένων με τις ζημιές χαρτοφυλακίου η κατανομή με τα βέλτιστα αποτελέσματα είναι η λοξή κανονική.

Ειδικά για τη λοξή t κατανομή δίνουμε τα ακόλουθα σχήματα που δείχνουν την προσαρμογή της στα 4 σετ δεδομένων.



Σχήμα 3.5. Προσαρμογή της λοξής t κατανομής στα δεδομένα

Στη συνέχεια δίνονται εκτιμήσεις για τις ποσότητες $VaR_{0.99}$ και $TVaR_{0.99}$ για κάθε κατανομή. Στα χρηματοοικονομικά, η αξία σε κίνδυνο (VaR) είναι μια μέτρηση που δηλώνει πώς η αξία ενός χαρτοφυλακίου περιουσιακών στοιχείων είναι πιθανόν να μειωθεί στη διάρκεια μιας συγκεκριμένης χρονικής περιόδου υπό συγκεκριμένες συνθήκες. Οι τέσσερις παράμετροι που ορίζουν την αξία σε κίνδυνο είναι

- Ο χρονικός ορίζοντας ή περίοδος διακράτησης στο οποίο αναλύουμε το χαρτοφυλάκιο. Η περίοδος διακράτησης που θα μελετήσουμε είναι μία ημέρα.
- Το διάστημα εμπιστοσύνης στο οποίο σκοπεύουμε να κάνουμε την εκτίμηση. Η μελέτη μας θα γίνει για διάστημα εμπιστοσύνης 99%.
- Τη νομισματική μονάδα που θα χρησιμοποιηθεί στον υπολογισμό της αξίας σε κίνδυνο. Στα δεδομένα Danish Fire είναι σε εκατομμύρια δανέζικες κορώνες ενώ τα δεδομένα Expert Advisor είναι σε αγγλική λίρα.

Η ουρά της αξίας σε κίνδυνο ($TVaR_{0.99}$) υπολογίζει την αναμενομένη ζημιά που μπορεί να προκύψει δοθέντος των παραπάνω τεσσάρων παραμέτρων και ότι αυτή είναι μεγαλύτερη του $VaR_{0.99}$.

Προκειμένου να εκτιμήσουμε την αξία σε κίνδυνο θα προσομοιώσουμε ένα εκατομμύριο τυχαίους αριθμούς βάση της εκάστοτε κατανομής και των παραμέτρων της, όπως έχουν προκύψει με τη μέθοδο μεγίστης πιθανοφάνειας, και έπειτα θα γίνουν οι εκτιμήσεις βάση της μεθόδου των ποσοστημορίων. Τα αποτελέσματα παρουσιάζονται στους ακόλουθους πίνακες.

Κατανομή	Danish		Expert Advisor	
	VaR	TVaR	VaR	TVaR
Λοξή Κανονική	23,60	26,37	7761,83	8695,06
Λοξή t	37,86	248,19	19799,26	119631,40
Lognormal	11,66	15,32	8691,97	14161,50
Κανονική	23,16	26,04	7623,77	8539,47
Weibull	16,26	19,96	7623,63	9890,21
Cauchy	17,70	220,36	25294,50	275781,80
Λογιστική	9,69	11,31	4415,35	5204,88
Εμπειρική Τιμή	26,04	58,59	15164,13	19623,96

Πίνακας 3.8. VaR_{0.99} και TVaR_{0.99} για τα δεδομένα Danish και Expert Advisor

Κατανομή	Log Danish		Log Expert Advisor	
	VaR	TVaR	VaR	TVaR
Λοξή Κανονική	2,73	3,07	9,46	10,00
Λοξή t	3,44	4,57	9,57	10,22
Lognormal	23,19	49,8	9,61	10,26
Κανονική	2,42	2,66	9,06	9,46
Weibull	3,42	4,11	8,99	9,30
Cauchy	10,46	143,10	29,13	310,39
Λογιστική	2,33	2,69	9,32	9,99
Εμπειρική Τιμή	3,25	3,82	9,62	9,87

Πίνακας 3.9. VaR_{0.99} και TVaR_{0.99} για τα δεδομένα Log Danish και Log Expert Advisor

Στον Πίνακα 3.8 και 3.9 επιβεβαιώνεται η πρόταση ότι οι θεωρητικές κατανομές τείνουν να υποεκτιμούν την εμπειρική. Οι αποκλίσεις μεταξύ της εμπειρικής τιμής και της τιμής που προκύπτει από τη προσομοίωση ενώ στο πίνακα 3.8 είναι πολύ μεγάλες στο πίνακα 3.9 τείνουν να εξομαλύνονται/μειώνονται λόγω του λογάριθμου και να λαμβάνουμε αποτελέσματα πιο κοντά στα πραγματικά. Παρατηρούμε ότι η λοξή t τείνει να δίνει εκτιμήσεις πολύ κοντά η και μεγαλύτερες από τα εμπειρικά δεδομένα στο πίνακα 3.8 ενώ στο πίνακα 3.9 μαζί με τη λοξή κανονική δίνουν χωρίς αμφιβολία τα καλύτερα αποτελέσματα τόσο στο Log Danish όσο και στο Log Expert Advisor.

3.4 Συμπεράσματα

Στόχος αυτού του κεφαλαίου ήταν το να προσαρμόσει τις λοξές κατανομές πάνω σε δεδομένα και να αναδείξει τη χρησιμότητα και την αξία τους. Τόσο η λοξή κανονική κατανομή όσο και η λοξή Student έχουν εξαιρετικά καλά αποτελέσματα στο σετ δεδομένων με τις ασφαλιστικές ζημίες και στο σετ με τις ζημίες χαρτοφυλακίου. Η λοξότητα των δεδομένων είναι ένα πολύ συχνό φαινόμενο στις αναλύσεις χαρτοφυλακίων, δηλαδή δεν ισχύει η παραδοχή ότι οι παρατηρήσεις μας κατανέμονται περίπου συμμετρικά γύρω από τη μέση τιμή. Ακόμα και όταν δουλέψαμε με τον λογάριθμο των δεδομένων μας προκειμένου να μετριάσουμε τη λοξότητα και τη κύρτωση και πάλι οι λοξές κατανομές έδιναν αποτελέσματα καλύτερα των υπολοίπων για αυτό και θα πρέπει να είναι στις υποψήφιες κατανομές που θα χρησιμοποιεί στην ερευνά του ένας αναλογιστής η διαχειριστής κινδύνου.

ΠΑΡΑΡΤΗΜΑ

ΚΕΦΑΛΑΙΟ 1

Κώδικας για την εξαγωγή των σχημάτων 1.1 έως 1.9

```
par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(-10,10,0.2)
y0<-dnorm(x,mean=0,sd=2)
y1<-dnorm(x,mean=1,sd=2)
y2<-dnorm(x,mean=-1,sd=2)
plot(x,y0,lty=2,col="red",type="l", lwd=3,xlab="x",ylab="f(x)",main="Density
Function")
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y1,lty=1,col="black", type="l",lwd=3)
legend(5,0.15,c("μ=0","μ=1","μ = -1"),col=c("red","blue","black"), lty=1:3,
cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(-3,3,0.1)
y0<-dnorm(x,mean=0,sd=1)
plot(x,y0,lty=1,col="black",type="l", lwd=3,xlab="x",ylab="f(x)",main="Density
Function")
legend(2,0.25,c("μ=0","σ=1"))

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(-3,8,0.01)
y0<-dt(x,df=2,ncp=1)
y1<-dt(x,df=3,ncp=1)
y2<-dt(x,df=5,ncp=1)
plot(x,y0,lty=1,col="black",type="l", lwd=3,xlab="x",ylab="f(x)",main="Density
Function",ylim=c(0,0.4))
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)

legend(5,0.25,c("v = 2","v = 3","v = 5"),col=c("black","blue","red"), lty=1:3,
cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(0,1,0.01)
y0<-dbeta(x,0.5,0.5)
y1<-dbeta(x,2,5)
y2<-dbeta(x,5,1)
plot(x,y0,lty=1,col="black",type="l", lwd=3,xlab="x",ylab="f(x)",main="Density
Function",xlim=c(0,1.2))
```

```

lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)
legend(1,3,c("α=b=0.5","α=2 & b=5","α=5 & b=1"),col=c("red","blue","black"),
lty=1:3, cex=0.7)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(-3,8,0.01)
y0<-dlogis(x,1,5)
y1<-dlogis(x,3,2)
y2<-dlogis(x,10,1)
plot(x,y0,lty=1,col="black",type="l",lwd=3,xlab="x",ylab="f(x)",main="Density
Function",xlim=c(-3,8),ylim=c(0,0.15))
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)
legend(5,0.14,c("α=1 & β=5","α=3 & β=2","α=10 &
β=1"),col=c("black","blue","red"), lty=1:3, cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(-3,8,0.01)
y0<-dcauchy(x,1,5)
y1<-dcauchy(x,2,3)
y2<-dcauchy(x,0.5,4)
plot(x,y0,lty=1,col="black",type="l",lwd=3,xlab="x",ylab="f(x)",main="Density
Function",xlim=c(-3,8),ylim=c(0,0.11))
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)
legend(5,0.10,c("α=1 & β=5","α=3 & β=2","α=10 &
β=1"),col=c("black","blue","red"), lty=1:3, cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(0,3,0.01)
y0<-dlnorm(x,mean=0,sd=0.25)
y1<-dlnorm(x,mean=0,sd=0.5)
y2<-dlnorm(x,mean=0,sd=1)
plot(x,y0,lty=1,col="black",type="l", lwd=3,xlab="x",ylab="f(x)",main="Density
Function")
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y1,lty=2,col="red", type="l",lwd=3)
legend(1.6,0.8,c("σ = 0.25","σ = 0.5","σ = 1"),col=c("black","blue","red"),
lty=1:3, cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(0,10,0.01)
y0<-dweibull(x,1,0.5)
y1<-dweibull(x,2,0.3)

```

```

y2<-dweibull(x,0.5,1)
plot(x,y0,lty=1,col="black",type="l",lwd=3,xlab="x",ylab="f(x)",main="Density
Function",xlim=c(0,1.5),ylim=c(0,4))
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)
legend(1.1,1.5,c("α=1 & b=0.5","α=2 & b=0.03","α=0.5 &
b=1"),col=c("black","blue","red"), lty=1:3, cex=0.8)

par(mfrow=c(1,1))

par(mai=c(1, 1, 0.3, 0.3))
x <- seq(0,5,0.01)
y0<-dgamma(x,1,0.5)
y1<-dgamma(x,2,1)
y2<-dgamma(x,3,2)
plot(x,y0,lty=1,col="black",type="l",lwd=3,xlab="x",ylab="f(x)",main="Density
Function",xlim=c(0,5),ylim=c(0,0.7))
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=2,col="red", type="l",lwd=3)
legend(3,0.50,c("α=1 & β=0.5","α=2 & β=1","α=3 &
β=2"),col=c("black","blue","red"), lty=1:3, cex=0.8)

```

ΚΕΦΑΛΑΙΟ 2

Κώδικας για την εξαγωγή των σχημάτων 2.1 έως 2.8

```

par(mfrow=c(1,2))

par(mai=c(1, 1, 0.3, 0.3))
a1 <- 0
a2 <- 1
a3 <- 3
a4 <- 10
x <- seq(-3,3,0.01)
y1 <-2*dnorm(x,0,1)*pnorm(a1*x,0,1)
y2 <-2*dnorm(x,0,1)*pnorm(a2*x,0,1)
y3 <-2*dnorm(x,0,1)*pnorm(a3*x,0,1)
y4 <-2*dnorm(x,0,1)*pnorm(a4*x,0,1)
par(mai=c(1, 1, 0.3, 0.3))
plot(x,y4,lty=4,col="red",type="l", lwd=3, ylab="φ(z;α)", xlab="z",
main="Standard Skew-Normal Density")
lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y1,lty=2,col="blue", type="l",lwd=3)
lines(x,y2,lty=1,col="black", type="l",lwd=3)
legend(-3,0.7,c("α = 0","α = 1","α = 3","α =
10"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

a1 <- 0
a2 <- -1
a3 <- -3
a4 <- -10

```



```

x <- seq(-3,3,0.01)
y1 <-2*dnorm(x,0,1)*pnorm(a1*x,0,1)
y2 <-2*dnorm(x,0,1)*pnorm(a2*x,0,1)
y3 <-2*dnorm(x,0,1)*pnorm(a3*x,0,1)
y4 <-2*dnorm(x,0,1)*pnorm(a4*x,0,1)
par(mai=c(1, 1, 0.3, 0.3))
plot(x,y4,lty=4,col="red",type="l", lwd=3, ylab="φ(z;α)", xlab="z",
main="Standard Skew-Normal Density")
lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y1,lty=1,col="black", type="l",lwd=3)
legend(1.5,0.7,c("α = 0","α = -1","α = -3","α = -
10"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

library(sn)
loc=xi=3
sc=omega=2
a=alpha=3
x <- seq(0,10,0.01)
y <- dsn(x,xi=loc, omega=sc, alpha=a)
plot(x,y,type="l")
ynew <- (2/sc)*dnorm((x-loc)/sc,0,1)*pnorm(a*((x-loc)/sc),0,1)
lines(x,ynew,lty=2,col="blue", type="l",lwd=3)

loc1=xi=0
loc2=xi=2
loc3=xi=4
loc4=xi=6
sc=omega=1
a=alpha=5

#par(mfrow=c(1,2))
par(mai=c(1, 1, 0.5, 0.3))
x <- seq(-2,12,0.01)
y1 <- (2/sc)*dnorm((x-loc1)/sc,0,1)*pnorm(a*((x-loc1)/sc),0,1)
y2 <- (2/sc)*dnorm((x-loc2)/sc,0,1)*pnorm(a*((x-loc2)/sc),0,1)
y3 <- (2/sc)*dnorm((x-loc3)/sc,0,1)*pnorm(a*((x-loc3)/sc),0,1)
y4 <- (2/sc)*dnorm((x-loc4)/sc,0,1)*pnorm(a*((x-loc4)/sc),0,1)
plot(x,y1,lty=1,col="black", type="l",lwd=3, ylab=(expression(list("g (y ; ξ
",ω^2,"α)"))), xlab="y", main="Skew-Normal Density : ω = 1, α = 5")

lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y4,lty=4,col="red",type="l", lwd=3)
legend(9,0.5,c("ξ = 0","ξ = 2","ξ = 4","ξ =
6"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

##### Αλλαγή του ω
par(mai=c(1, 1, 0.5, 0.3))
loc=xi=0
sc1=1
sc2=2
sc3=3
sc4=4

```

```

a=alpha=5

x <- seq(-4,12,0.01)
y1 <- (2/sc1)*dnorm((x-loc)/sc1,0,1)*pnorm(a*((x-loc)/sc1),0,1)
y2 <- (2/sc2)*dnorm((x-loc)/sc2,0,1)*pnorm(a*((x-loc)/sc2),0,1)
y3 <- (2/sc3)*dnorm((x-loc)/sc3,0,1)*pnorm(a*((x-loc)/sc3),0,1)
y4 <- (2/sc4)*dnorm((x-loc)/sc4,0,1)*pnorm(a*((x-loc)/sc4),0,1)
plot(x,y1,lty=1,col="black", type="l",lwd=3, ylab=(expression(list("g (y ;  $\xi$ 
", $\omega^2$ ," $\alpha$ "))), xlab="y", main="Skew-Normal Density :  $\xi = 0, \alpha = 5$ ")
lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y4,lty=4,col="red",type="l", lwd=3)
legend(8,0.5,c(" $\omega = 1$ "," $\omega = 2$ "," $\omega = 3$ "," $\omega = 4$ "),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

library(sn)
#?dst
n=5
a <- 3
x <- seq(-4, 4, 0.1)
y <- dst(x,xi=0, omega=1, alpha=a, nu=n)
plot(x,y,type="l")

ynew <- 2*dt(x,df=5)*pt(a*x*sqrt((n+1)/(n+x^2)),df=n)
lines(x,ynew,lty=2,col="blue", type="l",lwd=3)

####t distribution alpha=0
x <- seq(-4, 4, 0.1)
y <- dst(x,xi=0, omega=1, alpha=0, nu=n)
plot(x,y,type="l")
lines(x,dt(x,df=5),lty=2,col="blue", type="l",lwd=3)

par(mai=c(1, 1, 0.3, 0.3))
n=2
a1 <- 0
a2 <- 1
a3 <- 3
a4 <- 10
x <- seq(-4,4,0.01)
y1 <- dst(x,xi=0, omega=1, alpha=a1, nu=n)
y2 <- dst(x,xi=0, omega=1, alpha=a2, nu=n)
y3 <- dst(x,xi=0, omega=1, alpha=a3, nu=n)
y4 <- dst(x,xi=0, omega=1, alpha=a4, nu=n)

par(mai=c(1, 1, 0.3, 0.3))
plot(x,y4,lty=4,col="red",type="l", lwd=3, ylab="t(x; $\alpha$ ,v)", xlab="x",
main="Standard Skew-t Density : v = 2")
lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y1,lty=1,col="black", type="l",lwd=3)

```

```

legend(2,0.6,c("α = 0","α = 1","α = 3","α =
10"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

a1 <- 0
a2 <- -1
a3 <- -3
a4 <- -10
x <- seq(-4,4,0.01)
y1 <-dst(x,xi=0, omega=1, alpha=a1, nu=n)
y2 <-dst(x,xi=0, omega=1, alpha=a2, nu=n)
y3 <-dst(x,xi=0, omega=1, alpha=a3, nu=n)
y4 <-dst(x,xi=0, omega=1, alpha=a4, nu=n)

par(mai=c(1, 1, 0.3, 0.3))
plot(x,y4,lty=4,col="red",type="l", lwd=3, ylab="t(x;α,v)", xlab="x",
main="Standard Skew-t Density : v = 2")
lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y1,lty=1,col="black", type="l",lwd=3)
legend(2,0.6,c("α = 0","α = -1","α = -3","α = -
10"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

library(sn)
loc1=xi=0
loc2=xi=2
loc3=xi=4
loc4=xi=6
n=2

par(mai=c(1, 1, 0.5, 0.3))
x <- seq(-2,12,0.01)

y1 <-dst(x, xi=loc1, omega=1, alpha=3, nu=n)
y2 <-dst(x, xi=loc2, omega=1, alpha=3, nu=n)
y3 <-dst(x, xi=loc3, omega=1, alpha=3, nu=n)
y4 <-dst(x, xi=loc4, omega=1, alpha=3, nu=n)

plot(x,y1,lty=1,col="black", type="l",lwd=3, ylab="t(x;ξ,ω,α,v)", main="Skew-t
Density : ω = 1, α = 3, v=2")

lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y4,lty=4,col="red",type="l", lwd=3)
legend(9,0.5,c("ξ = 0","ξ = 2","ξ = 4","ξ =
6"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

par(mai=c(1, 1, 0.5, 0.3))
loc=xi=0
sc1=1
sc2=2
sc3=3
sc4=4

```

```

x <- seq(-4,12,0.01)
y1 <-dst(x, xi=0, omega=sc1, alpha=3, nu=n)
y2 <-dst(x, xi=0, omega=sc2, alpha=3, nu=n)
y3 <-dst(x, xi=0, omega=sc3, alpha=3, nu=n)
y4 <-dst(x, xi=0, omega=sc4, alpha=3, nu=n)
plot(x,y1,lty=1,col="black", type="l",lwd=3, ylab="t(x;ξ,ω,α,ν)", main="Skew-t
Density : ξ = 0, α = 3, ν=2")

lines(x,y3,lty=3,col="darkgreen", type="l",lwd=3)
lines(x,y2,lty=2,col="blue", type="l",lwd=3)
lines(x,y4,lty=4,col="red",type="l", lwd=3)
legend(7,0.5,c("ω = 1","ω = 2","ω = 3","ω =
4"),lty=c(1,2,3,4),col=c("black","blue","darkgreen","red"), lwd=c(3,3,3,3))

```

ΚΕΦΑΛΑΙΟ 3

Κώδικας για την εξαγωγή των σχημάτων 3.1 έως 3.6 και πινάκων 3.1 έως 3.9

```

library(copula)
library(fExtremes)
library(sn)
library(ghyp)
library(fitdistrplus)
library(MASS)

#####
Danish Data - Variable Danish
#####
data(danishClaims)
Danish<- danishClaims[,2] # Variable
mean(Danish)
summary(Danish)
sd(Danish)
skewness(Danish)
kurtosis(Danish)
quantile(Danish,0.99)
mean(Danish[Danish >= quantile(Danish,0.99)])
hist(Danish, breaks=50)
qqnorm(Danish); qqline(Danish, col = 2)

par(mfrow=c(1,2))
hist(Danish, breaks=50)
qqnorm(Danish, main="Normal Q-Q plot: Danish"); qqline(Danish, col = 2)

#####
Log of Danish Data - Variable LDanish
#####
LDanish<- log(Danish)
mean(LDanish)

```

```

summary(LDanish)
sd(LDanish)
skewness(LDanish)
kurtosis(LDanish)
quantile(LDanish,0.99)
mean(LDanish[LDanish >= quantile(LDanish,0.99)])

par(mfrow=c(1,2))
Log_Danish <- log(Danish)
hist(Log_Danish, breaks=50)
qqnorm(Log_Danish, main="Normal Q-Q plot: Log_Danish"); qqline(Log_Danish, col
= 2)

#####
Expert Advisor Data - Variable Expert_Advisor
#####

huri<-read.table("EAtest.txt",header=TRUE)
attach(huri)
Expert_Advisor <- damages
mean(Expert_Advisor)
summary(Expert_Advisor)
sd(Expert_Advisor)
skewness(Expert_Advisor)
kurtosis(Expert_Advisor)
quantile(Expert_Advisor,0.99)
mean(Expert_Advisor[Expert_Advisor >= quantile(Expert_Advisor,0.99)])

par(mfrow=c(1,2))
hist(Expert_Advisor, breaks=50)
qqnorm(Expert_Advisor, main="Normal Q-Q plot: Expert_Advisor");
qqline(Expert_Advisor, col = 2)

#####
Log Expert Advisor Data - Variable Log_Expert_Advisor
#####

Log_Expert_Advisor <- log(damages)
mean(Log_Expert_Advisor)
summary(Log_Expert_Advisor)
sd(Log_Expert_Advisor)
skewness(Log_Expert_Advisor)
kurtosis(Log_Expert_Advisor)
quantile(Log_Expert_Advisor,0.99)
mean(Log_Expert_Advisor[Log_Expert_Advisor >=
quantile(Log_Expert_Advisor,0.99)])

par(mfrow=c(1,2))
hist(Log_Expert_Advisor, breaks=50)
qqnorm(Log_Expert_Advisor, main="Normal Q-Q plot: Log_Expert_Advisor");
qqline(Log_Expert_Advisor, col = 2)

```

```

library(nortest)
ad.test(Log_Expert_Advisor)
shapiro.test(Log_Expert_Advisor)
ks.test(Log_Expert_Advisor, "pnorm")

#####
Danish and Log Danish Data - Skew Normal fitting -Skew t fitting
#####

b1<-sn.mple(y=Danish)
B1 <- cp2dp(c(b1[2]$cp[[1]], b1[2]$cp[[2]], b1[2]$cp[[3]]), "SN")
B1 # parameters SN

b2 <- st.mple(y=Danish)
B2 <- b2$dp
B2 # parameters St

Lb1<-sn.mple(y=LDanish)
LB1 <- cp2dp(c(Lb1[2]$cp[[1]], Lb1[2]$cp[[2]], Lb1[2]$cp[[3]]), "SN")
LB1 # parameters SN

Lb2 <- st.mple(y=LDanish)
LB2 <- Lb2$dp
LB2 # parameters St

#####
Expert and Log Expert Data - Skew Normal fitting -Skew t fitting
#####

c1<-sn.mple(y=Expert_Advisor)
C1 <- cp2dp(c(c1[2]$cp[[1]], c1[2]$cp[[2]], c1[2]$cp[[3]]), "SN")
C1 # parameters SN

c2 <- st.mple(y=Expert_Advisor)
C2 <- c2$dp
C2 # parameters St

Lc1<-sn.mple(y=Log_Expert_Advisor)
LC1 <- cp2dp(c(Lc1[2]$cp[[1]], Lc1[2]$cp[[2]], Lc1[2]$cp[[3]]), "SN")
LC1 # parameters SN

Lc2 <- st.mple(y=Log_Expert_Advisor)
LC2 <- Lc2$dp
LC2 # parameters St

#####
DANISH, LDANISH - OTHER DISTRIBUTIOS
#####

a3<- fitdist(data= Danish ,distr = "lnorm")
summary(a3)
A3<-a3$estimate

```

```

A3 # parameter lognormal

a4<- fitdist(data= Danish,distr = "norm")
A4<-a4$estimate
A4 # parameter normal

a5<- fitdist(data= Danish,distr = "weibull")
A5<-a5$estimate
A5 # parameter weibull
?dweibull

a6<- fitdist(data= Danish,distr = "cauchy")
A6<-a6$estimate
A6 # parameter cauchy

a7<- fitdist(data= Danish,distr = "logis")
A7<-a7$estimate
A7 # parameter logistic

a8<-fitdist(data= Danish, distr = "gamma")
A8<-a8$estimate
A8

#####

DDN <- ifelse(LDanish==0, 0.00000001, LDanish)
La3<- fitdist(data= DDN ,distr = "lnorm")
LA3<-La3$estimate
LA3 # parameter lognormal
mean(DDN)

La4<- fitdist(data= LDanish ,distr = "norm")
LA4<-La4$estimate
LA4 # parameter normal

La5<- fitdist(data= DDN,distr = "weibull")
LA5<-La5$estimate
LA5 # parameter weibull

La6<- fitdist(data= LDanish,distr = "cauchy")
LA6<-La6$estimate
LA6 # parameter cauchy

La7<- fitdist(data= LDanish,distr = "logis")
LA7<-La7$estimate
LA7 # parameter logistic

La8<-fitdist(data= LDanish, distr = "gamma")
LA8<-La8$estimate
LA8
#####
EXPERT, Log EXPERT
#####

```

```

g3<- fitdist(data= Expert_Advisor ,distr = "lnorm")
G3<-g3$estimate
G3 # parameter lognormal

g4<- fitdist(data= Expert_Advisor,distr = "norm")
G4<-g4$estimate
G4 # parameter normal

g5<- fitdist(data= Expert_Advisor,distr = "weibull")
G5<-g5$estimate
G5 # parameter weibull

g6<- fitdist(data= Expert_Advisor,distr = "cauchy")
G6<-g6$estimate
G6 # parameter cauchy

g7<- fitdist(data= Expert_Advisor,distr = "logis")
G7<-g7$estimate
G7 # parameter logistic

g8<-fitdist(data= Expert_Advisor, distr = "gamma")
G8<-g8$estimate
G8

#####

Lg3<- fitdist(data= Log_Expert_Advisor ,distr = "lnorm")
LG3<-Lg3$estimate
LG3 # parameter lognormal

Lg4<- fitdist(data= Log_Expert_Advisor ,distr = "norm")
LG4<-Lg4$estimate
LG4 # parameter normal

Lg5<- fitdist(data= Log_Expert_Advisor,distr = "weibull")
LG5<-Lg5$estimate
LG5 # parameter weibull

Lg6<- fitdist(data= Log_Expert_Advisor,distr = "cauchy")
LG6<-Lg6$estimate
LG6 # parameter cauchy

Lg7<- fitdist(data= Log_Expert_Advisor,distr = "logis")
LG7<-Lg7$estimate
LG7 # parameter lo

Lg8<-fitdist(data= Log_Expert_Advisor, distr = "gamma")
LG8<-Lg8$estimate
LG8

```



```

#####
##
#### Kolmogorov Smirnov Goodness of Fit test - Danish data
#####
##

ks.test(Danish, "psn", xi=B1[[1]], omega=B1[[2]], alpha=B1[[3]])
s$statistic

ks.test(Danish,"pst", xi=B2[[1]], omega=B2[[2]], alpha=B2[[3]], nu=B2[[4]])
s$statistic

ks.test(Danish,"plnorm", meanlog=A3[[1]], sdlog=A3[[2]])
s$statistic

ks.test(Danish,"pnorm", mean=A4[[1]], sd=A4[[2]])
s$statistic

ks.test(Danish,"pweibull", shape=A5[[1]], scale=A5[[2]])
s$statistic

ks.test(Danish,"pcauchy", location=A6[[1]], scale=A6[[2]])
s$statistic

ks.test(Danish,"plogis", location=A7[[1]], scale=A7[[2]])
s$statistic

ks.test(Danish,"pgamma", location=A8[[1]], scale=A8[[2]])
s$statistic

#####
##
#### Kolmogorov Smirnov Goodness of Fit test - Log Danish data
#####
##

s <- ks.test(LDanish, "psn", xi=LB1[[1]], omega=LB1[[2]], alpha=LB1[[3]])
s$statistic

s <- ks.test(LDanish,"pst", xi=LB2[[1]], omega=LB2[[2]], alpha=LB2[[3]],
nu=LB2[[4]])
s$statistic

s <- ks.test(LDanish,"plnorm", meanlog=LA3[[1]], sdlog=LA3[[2]])
s

s <- ks.test(LDanish,"pnorm", mean=LA4[[1]], sd=LA4[[2]])
s$statistic

ks.test(LDanish,"pweibull", shape=LA5[[1]], scale=LA5[[2]])
s <- ks.test(LDanish,"pweibull", shape=LA5[[1]], scale=LA5[[2]])
s$statistic

```

```

s <- ks.test(LDanish,"pcauchy", location=LA6[[1]], scale=LA6[[2]])
s$statistic

s <- ks.test(LDanish,"plogis", location=LA7[[1]], scale=LA7[[2]])
s$statistic

#####
##
#### Kolmogorov Smirnov Goodness of Fit test - Expert Advisor data
#####
##

s <- ks.test(Expert_Advisor, "psn", xi=C1[[1]], omega=C1[[2]], alpha=C1[[3]])
s$statistic

s <- ks.test(Expert_Advisor,"pst", xi=C2[[1]], omega=C2[[2]], alpha=C2[[3]],
nu=C2[[4]])
s$statistic

s <- ks.test(Expert_Advisor,"plnorm", meanlog=G3[[1]], sdlog=G3[[2]])
s$statistic

s <- ks.test(Expert_Advisor,"pnorm", mean=G4[[1]], sd=G4[[2]])
s$statistic

s <- ks.test(Expert_Advisor,"pweibull", shape=G5[[1]], scale=G5[[2]])
s$statistic

s <- ks.test(Expert_Advisor,"pcauchy", location=G6[[1]], scale=G6[[2]])
s$statistic

s <- ks.test(Expert_Advisor,"plogis", location=G7[[1]], scale=G7[[2]])
s$statistic

#####
##
#### Kolmogorov Smirnov Goodness of Fit test - Log Expert Advisor data
#####
##

ks.test(Log_Expert_Advisor, "psn", xi=LC1[[1]], omega=LC1[[2]],
alpha=LC1[[3]])
s$statistic

ks.test(Log_Expert_Advisor,"pst", xi=LC2[[1]], omega=LC2[[2]], alpha=LC2[[3]],
nu=LC2[[4]])
s$statistic

ks.test(Log_Expert_Advisor,"plnorm", meanlog=LG3[[1]], sdlog=LG3[[2]])
s$statistic

ks.test(Log_Expert_Advisor,"pnorm", mean=LG4[[1]], sd=LG4[[2]])
s$statistic

```

```

ks.test(Log_Expert_Advisor,"pweibull", shape=LG5[[1]], scale=LG5[[2]])
s$statistic

ks.test(Log_Expert_Advisor,"pcauchy", location=LG6[[1]], scale=LG6[[2]])
s$statistic

ks.test(Log_Expert_Advisor,"plogis", location=LG7[[1]], scale=LG7[[2]])
s$statistic

#####
Log-likelihood, AIC for Danish
#####

LLSN <- function(loc,sc,sh) {
  dens <- dsn(Danish, xi=loc, omega=sc, alpha=sh)
  return(sum(log(dens)))
}
LL <- LLSN(B1[[1]],B1[[2]],B1[[3]]); LL
AIC <- -2*LL+2*3;AIC

LLST <- function(loc,sc, sh, fr) {
  dens <- dst(Danish, xi=loc, omega=sc, alpha=sh, nu=fr)
  return(sum(log(dens)))
}
LL <- LLST(B2[[1]],B2[[2]],B2[[3]],B2[[4]]); LL
AIC <- -2*LL+2*4;AIC

LLLN <- function(m,s) {
  dens <- dlnorm(Danish, meanlog=m, sdlog=s)
  return(sum(log(dens)))
}
LL <- LLLN(A3[[1]],A3[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLN <- function(m,s) {
  dens <- dnorm(Danish, mean=m, sd=s)
  return(sum(log(dens)))
}
LL <- LLN(A4[[1]],A4[[2]]); LL
AIC <- -2*LL+2*2; AIC
LLW <- function(m,s) {
  dens <- dweibull(Danish, shape=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLW(A5[[1]],A5[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLC <- function(m,s) {
  dens <- dcauchy(Danish, location=m, scale=s)
  return(sum(log(dens)))
}

```

```

LL <- LLC(A6[[1]],A6[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLL <- function(m,s) {
  dens <- dlogis(Danish, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLL(A7[[1]],A7[[2]]); LL
AIC <- -2*LL+2*2; AIC

#####
Log-likelihood, AIC for Log Danish
#####

LLSN <- function(loc,sc,sh) {
  dens <- dsn(LDanish, xi=loc, omega=sc, alpha=sh)
  return(sum(log(dens)))
}
LL <- LLSN(LB1[[1]],LB1[[2]],LB1[[3]]); LL
AIC <- -2*LL+2*3;AIC

LLST <- function(loc,sc, sh, fr) {
  dens <- dst(LDanish, xi=loc, omega=sc, alpha=sh, nu=fr)
  return(sum(log(dens)))
}
LL <- LLST(LB2[[1]],LB2[[2]],LB2[[3]],LB2[[4]]); LL
AIC <- -2*LL+2*4;AIC

DDN <- ifelse(LDanish==0, 0.00000001, LDanish)
# DD2 <- LDanish[LDanish>0.01]
LLLN <- function(m,s) {
  dens <- dlnorm(DDN, meanlog=m, sdlog=s)
  return(sum(log(dens)))
}
LL <- LLLN(LA3[[1]],LA3[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLN <- function(m,s) {
  dens <- dnorm(LDanish, mean=m, sd=s)
  return(sum(log(dens)))
}
LL <- LLN(LA4[[1]],LA4[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLW <- function(m,s) {
  dens <- dweibull(DDN, shape=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLW(LA5[[1]],LA5[[2]]); LL
AIC <- -2*LL+2*2; AIC

```

```

LLC <- function(m,s) {
  dens <- dcauchy(LDanish, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLC(LA6[[1]],LA6[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLL <- function(m,s) {
  dens <- dlogis(LDanish, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLL(LA7[[1]],LA7[[2]]); LL
AIC <- -2*LL+2*2; AIC

#####
Log-likelihood, AIC for Expert_Advisor
#####

LLSN <- function(loc,sc,sh) {
  dens <- dsn(Expert_Advisor, xi=loc, omega=sc, alpha=sh)
  return(sum(log(dens)))
}
LL <- LLSN(C1[[1]],C1[[2]],C1[[3]]); LL
AIC <- -2*LL+2*3;AIC

LLST <- function(loc,sc, sh, fr) {
  dens <- dst(Expert_Advisor, xi=loc, omega=sc, alpha=sh, nu=fr)
  return(sum(log(dens)))
}
LL <- LLST(C2[[1]],C2[[2]],C2[[3]],C2[[4]]); LL
AIC <- -2*LL+2*4;AIC

LLLN <- function(m,s) {
  dens <- dlnorm(Expert_Advisor, meanlog=m, sdlog=s)
  return(sum(log(dens)))
}
LL <- LLLN(G3[[1]],G3[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLN <- function(m,s) {
  dens <- dnorm(Expert_Advisor, mean=m, sd=s)
  return(sum(log(dens)))
}
LL <- LLN(G4[[1]],G4[[2]]); LL
AIC <- -2*LL+2*2; AIC

```

```

LLW <- function(m,s) {
  dens <- dweibull(Expert_Advisor, shape=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLW(G5[[1]],G5[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLC <- function(m,s) {
  dens <- dcauchy(Expert_Advisor, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLC(G6[[1]],G6[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLL <- function(m,s) {
  dens <- dlogis(Expert_Advisor, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLL(G7[[1]],G7[[2]]); LL
AIC <- -2*LL+2*2; AIC

#####
Log-likelihood, AIC for Log_Expert_Advisor
#####

LLSN <- function(loc,sc,sh) {
  dens <- dsn(Log_Expert_Advisor, xi=loc, omega=sc, alpha=sh)
  return(sum(log(dens)))
}
LL <- LLSN(LC1[[1]],LC1[[2]],LC1[[3]]); LL
AIC <- -2*LL+2*3;AIC

LLST <- function(loc,sc, sh, fr) {
  dens <- dst(Log_Expert_Advisor, xi=loc, omega=sc, alpha=sh, nu=fr)
  return(sum(log(dens)))
}
LL <- LLST(LC2[[1]],LC2[[2]],LC2[[3]],LC2[[4]]); LL
AIC <- -2*LL+2*4;AIC

LLLN <- function(m,s) {
  dens <- dlnorm(Log_Expert_Advisor, meanlog=m, sdlog=s)
  return(sum(log(dens)))
}
LL <- LLLN(LG3[[1]],LG3[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLN <- function(m,s) {
  dens <- dnorm(Log_Expert_Advisor, mean=m, sd=s)

```

```

    return(sum(log(dens)))
}
LL <- LLN(LG4[[1]],LG4[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLW <- function(m,s) {
  dens <- dweibull(Log_Expert_Advisor, shape=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLW(LG5[[1]],LG5[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLC <- function(m,s) {
  dens <- dcauchy(Log_Expert_Advisor, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLC(LG6[[1]],LG6[[2]]); LL
AIC <- -2*LL+2*2; AIC

LLL <- function(m,s) {
  dens <- dlogis(Log_Expert_Advisor, location=m, scale=s)
  return(sum(log(dens)))
}
LL <- LLL(LG7[[1]],LG7[[2]]); LL
AIC <- -2*LL+2*2; AIC

#####
VAR TVAR DANISH
#####
n <- 1000000
s <- rsn(n,B1[[1]],B1[[2]],B1[[3]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rst(n,B2[[1]],B2[[2]],B2[[3]],B2[[4]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlnorm(n, A3[[1]], A3[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rnorm(n, A4[[1]], A4[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rweibull(n,A5[[1]],A5[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rcauchy(100000000, A6[[1]], A6[[2]])
quantile(s,0.99)

```

```

mean(s[s>=quantile(s,0.99)])

s <- rlogis(n,location=A7[[1]], scale=A7[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

#####
VAR TVAR Log DANISH
#####
n <- 1000000
s <-rsn(n,LB1[[1]],LB1[[2]],LB1[[3]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rst(n,LB2[[1]],LB2[[2]],LB2[[3]],LB2[[4]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlnorm(n, LA3[[1]], LA3[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rnorm(n, LA4[[1]], LA4[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rweibull(n,LA5[[1]],LA5[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rcauchy(100000000, LA6[[1]], LA6[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlogis(n,location=LA7[[1]], scale=LA7[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

#####
VAR TVAR Expert Advisor
#####
n <- 1000000
s <-rsn(n,C1[[1]],C1[[2]],C1[[3]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rst(n,C2[[1]],C2[[2]],C2[[3]],C2[[4]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlnorm(n, A3[[1]], A3[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

```



```

s <- rnorm(n, A4[[1]], A4[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rweibull(n,A5[[1]],A5[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rcauchy(100000000, A6[[1]], A6[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlogis(n,location=A7[[1]], scale=A7[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

#####
VAR TVAR Log Expert Advisor
#####
n <- 1000000
s <- rsn(n,LC1[[1]],LC1[[2]],LC1[[3]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rst(n,LC2[[1]],LC2[[2]],LC2[[3]],LC2[[4]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlnorm(n, LG3[[1]], LG3[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rnorm(n, LG4[[1]], LG4[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rweibull(n, LG5[[1]], LG5[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rcauchy(100000000, LG6[[1]], LG6[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

s <- rlogis(n,location=LG7[[1]], scale=LG7[[2]])
quantile(s,0.99)
mean(s[s>=quantile(s,0.99)])

```

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

Αντζουλάκος Δ. (2019). Γλώσσα Προγραμματισμού R με Εφαρμογές στον Αναλογισμό, ΠΜΣ «Αναλογιστική Επιστήμη και Διοικητική Κινδύνου. Πανεπιστήμιο Πειραιώς.

Ηλιόπουλος Γ. (2013). Βασικές Μέθοδοι Εκτίμησης Παραμέτρων με σημείο και με διάστημα, Β' Έκδοση. Εκδόσεις Σταμούλη Α.Ε..

Κούτρας Β. Μ. (2012). Εισαγωγή στη Θεωρία Πιθανοτήτων και Εφαρμογές. Εκδόσεις Σταμούλη Α.Ε.

Ξένη

Adcock C., Eling M. & Loperfido N. (2012). Skewed distributions in finance and actuarial science: a review. *The European Journal of Finance*, **14**, 1253-1281.

Azzalini A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171–178.

Azzalini A., Capitanio A. (2014). The Skew-Normal and Related Families. *Cambridge University Press*.

Azzalini A. (2005). The skew-normal distribution and related multivariate families (with discussion). *Scandinavian Journal of Statistics*, **32**, 159–188.

Arnold B.C. & Beaver R. J. (1999). The skew-Cauchy distribution. *Statistics & Probability Letters*, **49**, 285-290.

Basalamah D. (2017). Statistical Inference for a New Class of Skew T Distribution and Its Related Properties. *Thesis, Bowling Green State University*.

Basalamah D., Ning W. & Gupta A. (2018). The beta skew-t distribution and its properties. *Journal of statistical theory and practice*.

Behboodian J., Jamalizadeh A. & Balakrishnan N (2006). A new class of skew-Cauchy distributions. *Statistics & Probability Letters*, **76**, 1488–1493.

Eling M. (2011). Fitting Insurance Claims to Skewed Distributions: Are the Skew-Normal and Skew-Student Good Models?., *University of St. Gallen, Working Papers on Risk Management and Insurance*, No. 98.

Jones M.C. and Faddy M.J. (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society, Series B, Vol. 65*.

Kazemi R. & Noorizadeh M. (2015). A Comparison Between Skew-logistic and Skew-normal Distributions. *Matematika, Volume 31, Number 1, 15–24*.

Ibrahim A. H. S.(2009). On The Skew-t Distribution. *Thesis, Cairo University Institute of Statistical Studies and Research*.

Nadarajah S. (2009). The skew logistic distribution. *Springer, 93, 187–203*.