

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ**  
**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**  
**ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ**  
**ΤΟΜΕΑΣ ΠΙΘΑΝΟΤΗΤΩΝ, ΣΤΑΤΙΣΤΙΚΗΣ**  
**& ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ**

**ΚΩΝΣΤΑΝΤΙΝΑ ΚΟΥΤΣΟΥΠΙΑ**

**ΜΑΘΗΜΑΤΙΚΟΣ**

**ΜΟΝΟΔΙΑΣΤΑΤΗ ΑΠΟΚΟΜΜΕΝΗ**

**ΚΑΤΑΝΟΜΗ CAUCHY**

**ΙΩΑΝΝΙΝΑ, 2013**



Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική και Επιχειρησιακή Έρευνα, που απονέμει το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

Εγκρίθηκε την 23/05/2013 από την εξεταστική επιτροπή:

**Ζωγράφος Κωνσταντίνος**, Καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

**Λουκάς Σωτήριος**, Καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

**Μπατσίδης Απόστολος**, Λέκτορας του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων (Επιβλέπων)

### **Υπεύθυνη Δήλωση**

«Δηλώνω υπεύθυνα ότι η παρούσα διατριβή εκπονήθηκε κάτω από τους διεθνείς ηθικούς και ακαδημαϊκούς κανόνες δεοντολογίας και προστασίας της πνευματικής ιδιοκτησίας. Σύμφωνα με τους κανόνες αυτούς, δεν έχω προβεί σε ιδιοποίηση ξένου επιστημονικού έργου και έχω πλήρως αναφέρει τις πηγές που χρησιμοποίησα στην εργασία αυτή»

Υπογραφή

Κωνσταντίνα Κουτσουπιά



*Αφιερώνεται*  
*στους γονείς μου*  
*και στον αδελφό μου*



## **ΕΥΧΑΡΙΣΤΙΕΣ**

Η παρούσα διατριβή πραγματοποιήθηκε κατά τη διάρκεια του τελευταίου έτους φοίτησής μου στο Τμήμα Μαθηματικών της Σχολής Θετικών Επιστημών του Πανεπιστημίου Ιωαννίνων για την απόκτηση μεταπτυχιακού διπλώματος ειδίκευσης στον κλάδο «Στατιστική και Επιχειρησιακή Έρευνα».

Στην περάτωση της διατριβής αυτής συνέβαλαν σημαντικά ορισμένα άτομα, τα οποία σπεύδω να ευχαριστήσω.

Πρωτίστως, ευχαριστώ θερμά τον κ. Απόστολο Μπασιδμή, ο οποίος με προθυμία ανέλαβε την επίβλεψη και πρότεινε το θέμα της διατριβής. Η καθοδήγηση, η υπομονή και η ηθική υποστήριξη που μου παρείχε καθ' όλη τη διάρκεια της προσπάθειάς μου, υπήρξαν καθοριστικής σημασίας παράγοντες για την ολοκλήρωση της παρούσας μεταπτυχιακής διατριβής.

Ευχαριστίες απευθύνω στους κ.κ. Κωνσταντίνο Ζωγράφο και Σωτήριο Λουκά, για τη συμμετοχή τους στην Τριμελή Επιτροπή Κρίσης και τις παρατηρήσεις τους και ιδιαίτερα τον κ. Κωνσταντίνο Ζωγράφο, ο οποίος κατηύθυνε τα βήματά μου στις προπτυχιακές μου σπουδές.

Εκτιμώ επίσης τη βοήθεια όλων των καθηγητών του τομέα «Πιθανοτήτων, Στατιστικής και Επιχειρησιακής Έρευνας». Ένα ευχαριστώ χρωστάω στις φίλες μου Κυριακή και Φανουρία καθώς και στους φίλους που απέκτησα κατά την φοίτησή μου στο Μαθηματικό Τμήμα του Πανεπιστημίου Ιωαννίνων και ιδιαίτερα την Φωτεινή, για την ψυχολογική και συναισθηματική τους ενίσχυση κατά τη διάρκεια των σπουδών μου.

Τέλος, επιθυμώ να εκφράσω την ευγνωμοσύνη μου στους γονείς μου Ιωάννα, Παναγιώτη και στον αδελφό μου Ταξιάρχη που δεν έπαψαν να με στηρίζουν πνευματικά, συναισθηματικά, ηθικά και οικονομικά όλα αυτά τα χρόνια. Χωρίς τη βοήθεια και την ενθάρρυνσή τους θα ήταν αδύνατη η ολοκλήρωση των προπτυχιακών και μεταπτυχιακών μου σπουδών.

Κωνσταντίνα Κουτσουπιά

Ιωάννινα, Ιούνιος 2013.





# ΠΕΡΙΕΧΟΜΕΝΑ

## ΚΕΦΑΛΑΙΟ 1

### ΕΙΣΑΓΩΓΗ

1.1 Η Cauchy κατανομή .....	2
1.2 Αποκομμένες κατανομές.....	6
1.3 Σκοπός της μεταπτυχιακής διατριβής και περίληψή της.....	12

## ΚΕΦΑΛΑΙΟ 2

### ΜΟΝΟΔΙΑΣΤΑΤΗ ΔΙΠΛΑ ΑΠΟΚΟΜΜΕΝΗ ΚΑΤΑΝΟΜΗ CAUCHY

2.1 Η Κατανομή .....	15
2.2 Οι ιδιότητες της κατανομής.....	21

## ΚΕΦΑΛΑΙΟ 3

### ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

3.1 Μέθοδος Μέγιστης Πιθανοφάνειας.....	69
3.2 Μελέτη Προσομοίωσης .....	92
3.2.1 Πιθανότητα μη ύπαρξης του Ε.Μ.Π. της παραμέτρου $\sigma$ .....	93

3.2.2 Αξιολόγηση των Ε.Μ.Π. ....	97
3.3 Αριθμητικό Παράδειγμα.....	100
3.4 Μέθοδος των ροπών.....	102

## **ΚΕΦΑΛΑΙΟ 4**

### **ΓΕΝΙΚΕΥΣΕΙΣ ΤΗΣ ΜΟΝΟΔΙΑΣΤΑΤΗΣ ΑΠΟΚΟΜΜΕΝΗΣ ΚΑΤΑΝΟΜΗΣ CAUCHY**

4.1 Αποκομμένη γενικευμένη Cauchy .....	103
4.1.1 Η κατανομή και οι ιδιότητές της.....	103
4.1.2 Εκτίμηση των παραμέτρων.....	112
4.2 Λοξές Γενικεύσεις.....	114
4.2.1 Λοξή αποκομμένη Cauchy κατανομή .....	118
4.2.2 Αποκομμένη λοξή Cauchy κατανομή.....	125
<b>ΕΠΙΛΟΓΟΣ.....</b>	<b>131</b>
<b>ΠΑΡΑΡΤΗΜΑ Α.....</b>	<b>133</b>
<b>ΠΑΡΑΡΤΗΜΑ Β.....</b>	<b>139</b>
<b>ABSTRACT OF THE MSc DISSERTATION.....</b>	<b>143</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ .....</b>	<b>145</b>

# ΚΕΦΑΛΑΙΟ 1

## ΕΙΣΑΓΩΓΗ

Η κατανομή Cauchy αποτελεί μία από τις κατανομές που έχουν εισαχθεί στη βιβλιογραφία και χρησιμοποιούνται σε πρακτικές εφαρμογές με στόχο να περιγραφούν ικανοποιητικά πραγματικά τυχαία φαινόμενα με πιο παχιές ουρές από αυτές της κανονικής κατανομής. Η κατανομή Cauchy έχει εφαρμογές σε πολλούς τομείς, όπως στη βιολογία, σε κλινικές δοκιμές, ως μοντέλο για τις αποδόσεις μετοχών και αλλού (βλέπε Nadarajah, 2011). Όμως, μία βασική αδυναμία της κατανομής Cauchy είναι η μη ύπαρξη πεπερασμένων ροπών. Η αδυναμία αυτή είχε ως αποτέλεσμα τον περιορισμό της χρησιμοποίησής της για τη μοντελοποίηση πραγματικών τυχαίων φαινομένων, για τα οποία έχει αποδειχθεί, από εμπειρικές έρευνες, ότι οι ροπές τους είναι πεπερασμένες.

Με τον όρο αποκομμένες κατανομές Cauchy αναφερόμαστε σε κατανομές που προκύπτουν από την κατανομή Cauchy και επεκτάσεις αυτής με περιορισμό του πεδίου ορισμού τους. Οι κατανομές αυτές εισήχθησαν στη βιβλιογραφία για την αντιμετώπιση του προβλήματος της μη ύπαρξης πεπερασμένων ροπών της κατανομής Cauchy. Επομένως, οι αποκομμένες κατανομές Cauchy ως γενίκευση της κατανομής Cauchy είναι μία επιλογή σε όλες τις πρακτικές καταστάσεις που είτε δεν υπάρχουν πεπερασμένες ροπές είτε οι τιμές της τυχαίας μεταβλητής περιορίζονται σε ένα διάστημα των πραγματικών αριθμών.

Στο εισαγωγικό αυτό κεφάλαιο, για τη διευκόλυνση της μελέτης της διατριβής, παρουσιάζεται αρχικά η κατανομή Cauchy και έπειτα ο πλέον διαδεδομένος τρόπος αποκοπής. Στο πλαίσιο αυτό, παρατίθενται κάποια γενικά αποτελέσματα για την οικογένεια αυτών των αποκομμένων κατανομών. Η

εισαγωγή ολοκληρώνεται με την ενότητα όπου περιγράφεται ο σκοπός της μεταπτυχιακής αυτής διατριβής και δίνεται η περίληψη του περιεχομένου της.

## 1.1 Η Κατανομή Cauchy

Η κανονική κατανομή είναι η πιο ευρέως χρησιμοποιούμενη κατανομή στη Στατιστική. Αυτό συμβαίνει γιατί οι περισσότερες μεταβλητές είναι, ή τείνουν να είναι, κανονικώς κατανεμημένες. Υπάρχουν όμως περιπτώσεις τυχαίων πραγματικών φαινομένων, που η κανονική κατανομή αδυνατεί να μοντελοποιήσει σε ικανοποιητικό βαθμό. Ειδικότερα, όπως αποδεικνύεται από εμπειρικές έρευνες (βλέπε, μεταξύ άλλων, Blattberg and Gonedes (1974) και Kim and Mallick (2003)), η κανονική κατανομή αποτυγχάνει να μοντελοποιήσει, μεταξύ άλλων, τυχαία φαινόμενα των οποίων η κατανομή των δεδομένων έχει πιο παχιές ουρές (heavier tails) από αυτές της κανονικής κατανομής.

Η εισαγωγή νέων κατανομών με την ελπίδα ότι αυτές θα μοντελοποιούν κάποια τυχαία φαινόμενα καλύτερα από τις υπάρχουσες κατανομές, είναι συνήθης στη στατιστική βιβλιογραφία. Η κατανομή Cauchy, που είναι γνωστή και ως κατανομή Lorentz ή Cauchy–Lorentz, αποτελεί μία από τις κατανομές που έχουν εισαχθεί στη βιβλιογραφία και χρησιμοποιούνται σε πρακτικές εφαρμογές, με στόχο να περιγραφούν ικανοποιητικά φαινόμενα με πιο παχιές ουρές από αυτές της κανονικής κατανομής.

### Ορισμός 1.1.1

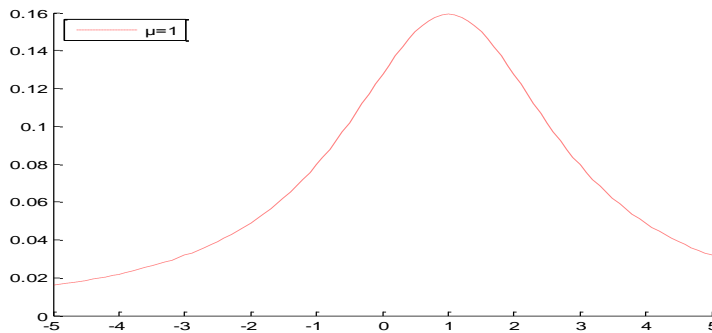
Η τ.μ.  $X$  λέμε ότι ακολουθεί κατανομή Cauchy, με παράμετρο θέσης  $\mu$  και παράμετρο κλίμακας  $\sigma$ , αν η σ.π.π. της δίνεται από τη σχέση:

$$f_X(x; \mu, \sigma) = \frac{1}{\pi\sigma} \left[ 1 + \left( \frac{x - \mu}{\sigma} \right)^2 \right]^{-1}, \quad (1.1.1)$$

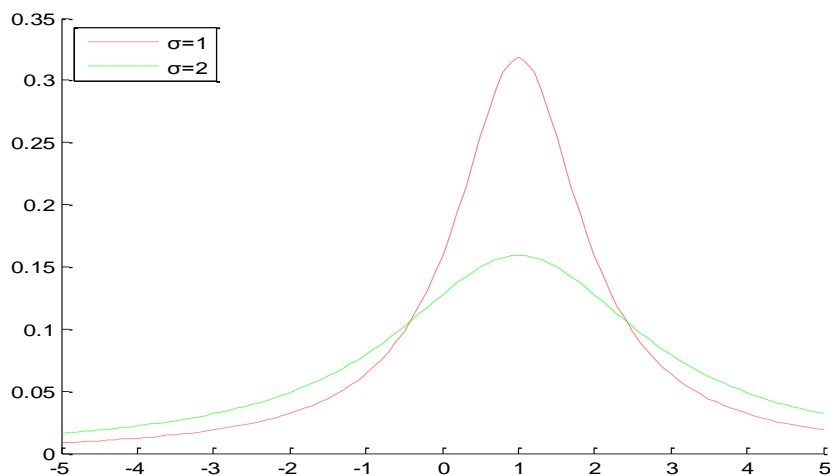
όπου  $x \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$  και  $\sigma > 0$ .

Όταν η τ.μ.  $X$  ακολουθεί την κατανομή Cauchy τότε αυτό θα συμβολίζεται  $X \sim C(\mu, \sigma)$ , όπου  $\mu$  και  $\sigma$ , οι παράμετροι αυτής της κατανομής. Αν  $\mu = 0$  και  $\sigma = 1$ , τότε λέμε ότι η τυχαία μεταβλητή  $X$  ακολουθεί την τυπική κατανομή Cauchy.

Στα γραφήματα που ακολουθούν δίνεται για τιμές του  $x$  στο διάστημα  $[-5, 5]$  η γραφική παράσταση της σ.π. της Κατανομής Cauchy όταν  $\sigma = 2$  και  $\mu = 1$  και όταν  $(\mu, \sigma) \in \{(2, 1), (2, 2)\}$ , αντίστοιχα. Οι γραφικές αυτές παραστάσεις, καθώς και οι υπόλοιπες της μεταπτυχιακής διατριβής, έγιναν δημιουργώντας κάθε φορά κατάλληλους κώδικες της MatLab. Οι κώδικες αυτοί παρατίθενται στο Παράρτημα Α.



Γράφημα 1- Σ.π. της Κατανομής Cauchy όταν  $\mu = 1$  και  $\sigma = 2$ , για  $x \in [-5, 5]$ .



Γράφημα 2- Σ.π.π. της Cauchy όταν  $(\mu, \sigma) \in \{(2,1), (2,2)\}$ , για  $x \in [-5,5]$ .

Στην πρόταση που ακολουθεί προσδιορίζεται η αθροιστική συνάρτηση κατανομής (α.σ.κ.) της κατανομής Cauchy και η αντίστροφή της. Σε όσα ακολουθούν με  $\tan^{-1}(\cdot) = \arctan(\cdot)$  συμβολίζεται η αντίστροφη συνάρτηση της εφαπτομένης που ορίζεται έτσι ώστε  $y = \arctan x \Rightarrow x = \tan y$ .

### Πρόταση 1.1.1

Έστω  $X$  μία τ.μ. που ακολουθεί την  $C(\mu, \sigma)$ . Τότε η α.σ.κ. της  $X$  δίνεται από τη σχέση:

$$F_X(x; \mu, \sigma) = \frac{1}{\pi} \tan^{-1} \left( \frac{x - \mu}{\sigma} \right) + \frac{1}{2}, \quad (1.1.2)$$

όπου  $x \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$  και  $\sigma > 0$ . Επιπλέον, η αντίστροφη συνάρτηση της α.σ.κ. δίνεται από τη σχέση:

$$F_X^{-1}(p) = \mu + \sigma \tan(\pi(p - 0.5)), \quad p \in [0, 1]. \quad (1.1.3)$$

### Απόδειξη

Η α.σ.κ. της τ.μ.  $X$ , δίνεται από τη σχέση:

$$F_X(x) = \int_{-\infty}^x \frac{1}{\pi\sigma} \left[ 1 + \left( \frac{y - \mu}{\sigma} \right)^2 \right]^{-1} dy = \int_{-\infty}^{x-\mu} \frac{1}{\pi\sigma} \left[ 1 + \left( \frac{z}{\sigma} \right)^2 \right]^{-1} dz.$$

Ύστερα από αλγεβρικές πράξεις προκύπτει ότι:

$$F_X(x) = \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\pi(1+w^2)} dw = \frac{1}{\pi} \tan^{-1} \left( \frac{x-\mu}{\sigma} \right) + \frac{1}{2}.$$

Είναι γνωστό ότι  $F_X^{-1}(p)$ , με  $p \in [0, 1]$ , είναι ο μονοσήμαντα ορισμένος πραγματικός αριθμός που ικανοποιεί τη σχέση  $F_X(x) = p$ . Λαμβάνοντας υπόψη τη σχέση (1.1.2) είναι:

$$\frac{1}{\pi} \tan^{-1} \left( \frac{x-\mu}{\sigma} \right) + \frac{1}{2} = p,$$

από όπου, με λίγη άλγεβρα, προκύπτει το επιθυμητό αποτέλεσμα. ■

Η κατανομή Cauchy βρίσκει εφαρμογές σε πολλούς τομείς, όπως στη βιολογία, σε κλινικές δοκιμές, ως μοντέλο για τις αποδόσεις μετοχών και αλλού (βλέπε Nadarajah, 2011). Μία εξαιρετική ιστορική αναδρομή για την κατανομή Cauchy εκπονήθηκε από τον Stigler (1974). Σύμφωνα με τη μελέτη του Stigler (1974), εσφαλμένα αποδίδεται ότι η κατανομή Cauchy έχει εισαχθεί στη βιβλιογραφία από τον Augustin - Louis Cauchy. Ο Stigler (1974) αρχικά επισημαίνει ότι η καμπύλη, που ονομάζουμε συνήθως κατανομή Cauchy, ή



καμπύλες ανάλογες με αυτήν, εμφανίστηκαν για πρώτη φορά στα έργα του Fermat στα μέσα του 17<sup>ου</sup> αιώνα, και στη συνέχεια μελετήθηκαν, μεταξύ άλλων, από τους Newton, Leibniz, Huygens, Guido Grandi, και τη Maria Agnesi. Ουσιαστικά, όπως αναφέρει ο Stigler (1974), ο Poisson (1824) είναι ο πρώτος που χρησιμοποιεί μία τυχαία μεταβλητή με σ.π.π. αυτής της τυπικής κατανομής Cauchy (βλέπε και Prokhorov (1999)), ενώ ο Cauchy (1853) είναι αυτός που αναφέρεται στην ειδική περίπτωση μίας τυχαίας μεταβλητής με σ.π.π. αυτής της σχέσης (1.1.1) με  $\mu = 0$  και  $\sigma = 1/k$ , με  $k > 0$ . Στο πλαίσιο αυτό, ο Poisson (1824) είναι ουσιαστικά ο πρώτος που επισημαίνει τη βασική αδυναμία της κατανομής Cauchy που είναι η μη ύπαρξη πεπερασμένων ροπών.

Η μη ύπαρξη πεπερασμένων ροπών έχει ως αποτέλεσμα τον περιορισμό της χρήσης της κατανομής Cauchy για τη μοντελοποίηση πραγματικών τυχαίων φαινομένων. Αυτό συνέβει καθώς με εμπειρικές έρευνες και μελέτες προσομοιώσεων στους τομείς που χρησιμοποιείται έχει αποδειχθεί, ότι οι ροπές όλων των τάξεων είναι πεπερασμένες. Επομένως, η χρήση της Cauchy δεν είναι κατάλληλη, αφού όλες οι ροπές της είναι μη πεπερασμένες. Ένας τρόπος αντιμετώπισης αυτής της αδυναμίας είναι η εισαγωγή και η χρήση των **αποκομμένων κατανομών (truncated distributions)**. Οι αποκομμένες κατανομές αποτελούν το αντικείμενο μελέτης της επόμενης παραγράφου.

## 1.2 Αποκομμένες κατανομές

Με τον όρο αποκομμένες κατανομές αναφερόμαστε σε κατανομές που προκύπτουν από γνωστές κατανομές με περιορισμό του πεδίου ορισμού τους. Επομένως, μια αποκομμένη κατανομή δεν είναι τίποτε άλλο παρά μία δεσμευμένη κατανομή που προκύπτει από περιορισμό του πεδίου ορισμού μίας υπάρχουσας κατανομής. Οι αποκομμένες κατανομές εκτός του ότι αποτελούν έναν τρόπο αντιμετώπισης της μη ύπαρξης πεπερασμένων ροπών, χρησιμοποιούνται στην

πράξη σε περιπτώσεις που γνωρίζουμε ότι οι τιμές μίας τ.μ. περιορίζονται είτε για μικρότερες ή μεγαλύτερες τιμές από μία δοθείσα ή άγνωστη τιμή είτε εντός ενός διαστήματος. Για παράδειγμα, όπως αναφέρουν οι Nadarajah and Kotz (2006), όταν εξετάζονται οι ημερομηνίες γέννησης των παιδιών σε ένα σχολείο, αυτές συνήθως υπόκεινται σε περιορισμό σε σχέση με τις ηλικίες όλων των παιδιών της περιοχής, δεδομένου ότι το σχολείο δέχεται παιδιά μόνο μιας συγκεκριμένης ηλικιακής ομάδας σε μια συγκεκριμένη ημερομηνία. Επίσης, τα τελικά προϊόντα μιας βιομηχανίας υπόκεινται σε επιθεώρηση ελέγχου πριν από την αποστολή στον πελάτη. Η συνήθης πρακτική είναι ότι εφόσον η απόδοση του προϊόντος περιλαμβάνεται εντός ορισμένων ορίων ανοχής, τότε το προϊόν κρίνεται αποδεκτό και αποστέλλεται στον πελάτη, ενώ αν αποτύχει, το προϊόν απορρίπτεται.

Όπως ήδη έχει αναφερθεί, οι αποκομμένες κατανομές προκύπτουν περιορίζοντας το πεδίο ορισμού γνωστών, υπάρχουσών κατανομών. Διάφοροι τρόποι αποκοπής έχουν προταθεί στη βιβλιογραφία. Στη συνέχεια παρουσιάζονται οι τρεις πλέον διαδεδομένοι τρόποι αποκοπής και παρατίθενται κάποια αποτελέσματα που προκύπτουν για την οικογένεια αυτών των αποκομμένων κατανομών. Χωρίς βλάβη της γενικότητας, τα αποτελέσματα παρουσιάζονται για την περίπτωση που η αρχική τ.μ. είναι συνεχής με πεδίο ορισμού το σύνολο των πραγματικών αριθμών.

### Ορισμός 1.2.1

Έστω μια τυχαία μεταβλητή με σ.π.π.  $g(\cdot)$  και αθροιστική συνάρτηση κατανομής (α.σ.κ.)  $G(\cdot)$ . Τότε η τυχαία μεταβλητή (τ.μ.)  $X$  εκφράζει τη διπλά αποκομμένη μορφή αυτής της κατανομής στο διάστημα  $[a, b]$ , αν η σ.π.π. της δίνεται από τη σχέση:

$$f_X(x) = f_X(x, a, b) = \begin{cases} \frac{g(x)}{G(b) - G(a)}, & \text{εάν } a \leq x \leq b \\ 0, & \text{άλλοι} \end{cases}. \quad (1.2.1)$$

Στην πρόταση που ακολουθεί δίνεται η αθροιστική συνάρτηση κατανομής (α.σ.κ.) της διπλά αποκομμένης κατανομής του Ορισμού 1.2.1 και η αντίστροφή της.

### Πρόταση 1.2.1

Έστω  $X$  μία τ.μ. που ακολουθεί τη διπλά αποκομμένη κατανομή με σ.π.π. που δίνεται από τη σχέση (1.2.1). Τότε η α.σ.κ. της τ.μ.  $X$  δίνεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{για } x < a \\ \frac{G(x) - G(a)}{G(b) - G(a)} & , \text{για } a \leq x < b. \\ 1 & , \text{για } x \geq b \end{cases}. \quad (1.2.2)$$

Επιπλέον, η αντίστροφη συνάρτηση της α.σ.κ., υπό την προϋπόθεση ότι υπάρχει η αντίστροφη της  $G(\cdot)$ , δίνεται από τη σχέση:

$$F_X^{-1}(p) = G^{-1}(G(a) + p(G(b) - G(a))), \quad p \in [0, 1]. \quad (1.2.3)$$

### Απόδειξη

Αν  $x < a$ , τότε προφανώς  $F_X(x) \stackrel{op}{=} P(X \leq x) = 0$ . Επίσης, αν  $a \leq x < b$ , τότε η α.σ.κ. της τ.μ.  $X$  δίνεται από τη σχέση:

$$F_X(x) = \int_{-\infty}^x \frac{g(y)}{G(b)-G(a)} dy = \int_a^x \frac{g(y)}{G(b)-G(a)} dy = \frac{G(x)-G(a)}{G(b)-G(a)}.$$

Τέλος, αν  $b \leq x$ , τότε  $F_X(x) \stackrel{op}{=} P(X \leq x) = 1$ .

Επιπλέον είναι γνωστό ότι  $F_X^{-1}(p)$ ,  $p \in [0,1]$ , είναι ο μονοσήμαντα ορισμένος πραγματικός αριθμός που ικανοποιεί τη σχέση  $F_X(x) = p$ . Λαμβάνοντας υπόψη, για  $a \leq x \leq b$ , τη σχέση (1.2.2) προκύπτει το επιθυμητό αποτέλεσμα μετά από αλγεβρικές πράξεις. ■

**Παρατήρηση 1.2.1** Στο σημείο αυτό είναι απαραίτητο να σημειωθεί ότι στη βιβλιογραφία έχουν εισαχθεί και οι λεγόμενες **περιορισμένες (limited)** κατανομές που προκύπτουν με περιορισμό του πεδίου ορισμού υπαρχουσών κατανομών, αλλά είναι διαφορετικά ορισμένες (βλέπε, μεταξύ άλλων, Jamalizadeh *et al.*, 2009). Ειδικότερα, έστω μια τυχαία μεταβλητή με σ.π.π.  $g(\cdot)$  και αθροιστική συνάρτηση κατανομής (α.σ.κ.)  $G(\cdot)$ . Τότε η τυχαία μεταβλητή (τ.μ.)  $X$  εκφράζει τη limited μορφή αυτής της κατανομής στο διάστημα  $[a,b]$ , αν η α.σ.κ. της δίνεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{για } x < a \\ G(x) & , \text{για } a \leq x < b. \\ 1 & , \text{για } x \geq b \end{cases}$$

### Προσομοίωση δεδομένων από μία διπλά αποκομμένη κατανομή

Από τη σχέση (1.2.3) της Πρότασης 1.2.1 προκύπτει ότι, υπό την προϋπόθεση ότι είναι εφικτός ο προσδιορισμός της  $G^{-1}(\cdot)$ , αν κάποιος θέλει να δημιουργήσει ένα τυχαίο δείγμα (τ.δ.)  $X_1, \dots, X_n$  από μία διπλά αποκομμένη

κατανομή με σ.π.π. όπως αυτή που δίνεται στη σχέση (1.2.1) του Ορισμού 1.2.1 αρκεί να υπολογίσει τις τιμές:

$$X_i = G^{-1}(G(a) + U_i(G(b) - G(a))), \quad i = 1, \dots, n, \quad (1.2.4)$$

όπου  $U_1, \dots, U_n$  είναι ένα τ.δ. από την ομοιόμορφη κατανομή  $U(0,1)$ .

### Αριστερά και δεξιά αποκομμένη κατανομή

Στην ειδική περίπτωση της διπλά αποκομμένης κατανομής με  $b = \infty$ , η κατανομή μελετάται στο διάστημα  $[a, \infty)$ , και τότε λέμε ότι έχουμε την από αριστερά αποκομμένη κατανομή, ενώ όταν  $a = -\infty$  η κατανομή μελετάται στο διάστημα  $(-\infty, b]$ , και τότε λέμε ότι έχουμε την από δεξιά αποκομμένη κατανομή. Λαμβάνοντας υπόψη τις σχέσεις (1.2.1)-(1.2.3) και τις ιδιότητες της α.σ.κ. προκύπτουν άμεσα τα ακόλουθα πορίσματα.

#### Πόρισμα 1.2.1

Έστω μια τυχαία μεταβλητή με σ.π.π.  $g(\cdot)$  και αθροιστική συνάρτηση κατανομής (α.σ.κ.)  $G(\cdot)$ . Η τ.μ.  $X$  εκφράζει την από αριστερά αποκομμένη μορφή αυτής της κατανομής στο διάστημα  $[a, \infty)$ , αν η σ.π.π. της δίνεται από τη σχέση:

$$f_X(x) = f_X(x, a) = \begin{cases} \frac{g(x)}{1 - G(a)} & , \text{εάν } a \leq x \\ 0 & , \text{αλλού} \end{cases} . \quad (1.2.5)$$

Επιπλέον, η α.σ.κ. της τ.μ.  $X$  δίνεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{για } x < a \\ \frac{G(x) - G(a)}{1 - G(a)} & , \text{για } a \leq x \end{cases} , \quad (1.2.6)$$

ενώ η αντίστροφη συνάρτηση της α.σ.κ., υπό την προϋπόθεση ότι υπάρχει η αντίστροφη της  $G(\cdot)$ , δίνεται από τη σχέση:

$$F_X^{-1}(p) = G^{-1}(G(a) + p(1 - G(a))), \quad p \in [0, 1].$$

### Πόρισμα 1.2.2

Έστω μια τυχαία μεταβλητή με σ.π.π.  $g(\cdot)$  και α.σ.κ.  $G(\cdot)$ . Τότε η τ.μ.  $X$  εκφράζει την από δεξιά αποκομμένη μορφή αυτής της κατανομής στο διάστημα  $[-\infty, b)$ , αν η σ.π.π. και η α.σ.κ. δίνονται από τις σχέσεις:

$$f_X(x) = f_X(x, b) = \begin{cases} \frac{g(x)}{G(b)} & , \text{εάν } x \leq b \\ 0 & , \text{αλλού} \end{cases} \quad (1.2.7)$$

και

$$F_X(x) = \begin{cases} \frac{G(x)}{G(b)} & , \text{για } x < b \\ 1 & , \text{για } x \geq b \end{cases} \quad (1.2.8)$$

αντίστοιχα. Επιπλέον, η αντίστροφη συνάρτηση της α.σ.κ., υπό την προϋπόθεση ότι υπάρχει η αντίστροφη της  $G(\cdot)$ , δίνεται από τη σχέση:

$$F_X^{-1}(p) = G^{-1}(pG(b)), \quad p \in [0, 1].$$

Στην παρούσα διατριβή κύριο αντικείμενο μελέτης αποτελεί η διπλά αποκομμένη κατανομή Cauchy, ενώ θα παρουσιασθούν και οι μονοδιάστατες γενικεύσεις αυτής, που έχουν εισαχθεί στη βιβλιογραφία, μέχρι την ώρα

συγγραφής αυτής της μεταπτυχιακής διατριβής. Ο σκοπός της μεταπτυχιακής διατριβής και η περίληψη του περιεχομένου της δίνονται στην επόμενη ενότητα.

### 1.3 Σκοπός της μεταπτυχιακής διατριβής και περίληψή της

Σκοπός αυτής της μεταπτυχιακής διατριβής είναι μια παρουσίαση και κριτική ανασκόπηση της διπλά αποκομμένης κατανομής Cauchy, ενώ απώτερος σκοπός είναι η παρουσίαση των κυριότερων αποτελεσμάτων, που είναι διαθέσιμα για αυτή.

Στο πλαίσιο αυτό, στο **Κεφάλαιο 2** (Μονοδιάστατη διπλά αποκομμένη κατανομή Cauchy), μελετάται η διπλά αποκομμένη κατανομή Cauchy, η οποία πρωτοπαρουσιάστηκε, όσο είμαστε σε θέση να γνωρίζουμε, από τους Daliya *et al.* (2001). Ειδικότερα, σκοπός του κεφαλαίου είναι μια κριτική παράθεση και ανασκόπηση των κυριότερων αποτελεσμάτων που αφορούν τις ιδιότητες αυτής της κατανομής. Στο πλαίσιο αυτό, δίνονται εκφράσεις για τις ροπές, τη μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο, τη ροπογεννήτρια και τη χαρακτηριστική συνάρτηση, την εντροπία του Renyi (1961) και του Shannon (1948), τους συντελεστές λοξότητας και κύρτωσης, καθώς και το μέτρο του σχήματος του Song (2001). Τέλος, προσδιορίζονται οι ασυμπτωτικές κατανομές του μέγιστου και ελάχιστου διατεταγμένου στατιστικού. Τα αποτελέσματα αυτά βασίζονται, κυρίως, στις εργασίες των Nadarajah and Kotz (2006), Nadarajah (2011) και Daliya *et al.* (2001).

Στο **Κεφάλαιο 3** (Εκτίμηση παραμέτρων) αντικείμενο μελέτης είναι η εκτίμηση των τεσσάρων αγνώστων παραμέτρων της διπλά αποκομμένης κατανομής Cauchy. Αρχικά, παρουσιάζονται και αναλύονται εκτενώς τα αποτελέσματα των Daliya *et al.* (2001) σχετικά με την ύπαρξη και τον προσδιορισμό των Εκτιμητών Μέγιστης Πιθανοφάνειας (Ε.Μ.Π.) των αγνώστων παραμέτρων. Ειδικότερα, προκύπτει ότι η ύπαρξη πεπερασμένου εκτιμητή μίας εκ

των παραμέτρων της διπλά αποκομμένης κατανομής Cauchy εξαρτάται από το αν πληρείται ή όχι μία συνθήκη. Για το λόγο αυτό, παρακινούμενοι από τους Daliya *et al.* (2001) διεξάγεται μία μελέτη προσομοίωσης για την εκτίμηση της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή για αυτήν την παράμετρο. Επιπρόσθετα, με χρήση προσομοιωμένων δεδομένων αξιολογείται η απόδοση των Ε.Μ.Π. των αγνώστων παραμέτρων. Τέλος, το κεφάλαιο ολοκληρώνεται με ένα αριθμητικό παράδειγμα, όπου προσαρμόζεται το μοντέλο της διπλά αποκομμένης κατανομής Cauchy, και τα αποτελέσματα που δόθηκαν από τον Nadarajah (2011) και αφορούν στην εκτίμηση των αγνώστων παραμέτρων με τη μέθοδο των ροπών.

Στο **Κεφάλαιο 4** (Γενικεύσεις της μονοδιάστατης διπλά αποκομμένης κατανομής Cauchy) κάποιες μονοδιάστατες γενικεύσεις της υπό μελέτης κατανομής παρουσιάζονται. Ειδικότερα, θα παρουσιασθούν η αποκομμένη γενικευμένη Cauchy (Ateya and Al-Hussaini, 2012), η λοξή αποκομμένη (Nadarajah and Ali, 2004) και η αποκομμένη λοξή κατανομή Cauchy (Jamalizadeh *et al.*, 2009).

Στα Παραρτήματα Α και Β παρατίθενται τα αρχεία της Matlab που δημιουργήθηκαν στα πλαίσια των Κεφαλαίων 1-4 της διατριβής. Ειδικότερα, στο **Παράρτημα Α** παρατίθεται ο τρόπος με τον οποίο δημιουργήθηκαν οι γραφικές παραστάσεις, προσδιορίστηκαν οι τιμές διάφορων χαρακτηριστικών π.χ. μέση τιμή, διακύμανση, ενώ στο **Παράρτημα Β** ο τρόπος με τον οποίο έγινε η μελέτη προσομοίωσης του Κεφαλαίου 3.

Τέλος, η διατριβή ολοκληρώνεται με την περίληψή της στα Αγγλικά (**Abstract of MSc Dissertation**) και τη **Βιβλιογραφία**.





## ΚΕΦΑΛΑΙΟ 2

### ΜΟΝΟΔΙΑΣΤΑΤΗ ΔΙΠΛΑ ΑΠΟΚΟΜΜΕΝΗ ΚΑΤΑΝΟΜΗ CAUCHY

Στο κεφάλαιο αυτό παρουσιάζεται η μονοδιάστατη διπλά αποκομμένη κατανομή Cauchy, η οποία προκύπτει με την εισαγωγή της σ.π.π. και της α.σ.κ. της κατανομής Cauchy (βλέπε (1.1.1) και (1.1.2), αντίστοιχα) στη σχέση (1.2.1). Σκοπός αυτού του κεφαλαίου είναι η παρουσίαση και κριτική ανασκόπηση της μονοδιάστατης διπλά αποκομμένης κατανομής Cauchy, των ιδιοτήτων και των χαρακτηριστικών της, όπως αυτές μελετήθηκαν στη βιβλιογραφία. Στο πλαίσιο αυτό, δίνονται εκφράσεις για τις ροπές, τη μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο, τη ροπογεννήτρια και τη χαρακτηριστική συνάρτηση, τις εντροπίες του Shannon (1948) και του Renyi (1961), τους συντελεστές λοξότητας και κύρτωσης και το μέτρο σχήματος του Song (2001). Τέλος, θα προσδιοριστούν οι ασυμπτωτικές κατανομές του μέγιστου και ελάχιστου διατεταγμένου στατιστικού. Τα αποτελέσματα αυτά βασίζονται, κυρίως, στις εργασίες των Nadarajah and Kotz (2006), Nadarajah (2011) και Daliya *et al.* (2001).

#### 2.1 Η Κατανομή

Η μονοδιάστατη διπλά αποκομμένη κατανομή Cauchy, προκύπτει από τη σχέση (1.2.1) αν λάβουμε υπόψη τη σ.π.π. και την α.σ.κ. της κατανομής Cauchy, όπως αυτές δίνονται στις σχέσεις (1.1.1) και (1.1.2), αντίστοιχα. Μία ειδική περίπτωση αυτής πρωτοεισήχθη, όσο είμαστε σε θέση να γνωρίζουμε, από τους

Johnson and Kotz (1970), αλλά οι Daliya *et al.* (2001) εισάγουν για πρώτη φορά τον ορισμό της μονοδιάστατης διπλά αποκομμένης κατανομής που δίνεται ακολούθως.

### Ορισμός 2.1.1 (Daliya *et al.* , 2001)

Έστω μια τυχαία μεταβλητή που ακολουθεί την κατανομή Cauchy με παραμέτρους θέσης και κλίμακας  $\mu$  και  $\sigma$ , αντίστοιχα. Έστω  $X$  η τυχαία μεταβλητή που εκφράζει τη διπλά αποκομμένη μορφή της κατανομής στο διάστημα  $[a, b]$ , όπου  $-\infty < a < b < \infty$ . Η σ.π.π. της  $X$  δίνεται από τη σχέση:

$$f_X(x) = \frac{\left\{1 + \left(\frac{x - \mu}{\sigma}\right)^2\right\}^{-1}}{\sigma \left[ \tan^{-1}\left(\frac{b - \mu}{\sigma}\right) - \tan^{-1}\left(\frac{a - \mu}{\sigma}\right) \right]}, \quad (2.1.1)$$

ή ισοδύναμα από τη σχέση

$$f_X(x) = \frac{\sigma}{D} \frac{1}{\left[ \sigma^2 + (x - \mu)^2 \right]}, \quad (2.1.2)$$

όπου

$$D = \tan^{-1}\left(\frac{b - \mu}{\sigma}\right) - \tan^{-1}\left(\frac{a - \mu}{\sigma}\right), \quad (2.1.3)$$

$$a \leq x \leq b, \quad \sigma > 0, \quad -\infty < \mu < \infty.$$

Όταν η τ.μ.  $X$  ακολουθεί τη διπλά αποκομμένη κατανομή Cauchy τότε αυτό θα συμβολίζεται  $X \sim TC^{(a,b)}(\mu, \sigma)$  ή  $X \sim TC(\mu, \sigma; a, b)$ . Στην περίπτωση που  $\mu = 0$  και  $\sigma = 1$  λέμε ότι έχουμε τη διπλά αποκομμένη τυπική κατανομή

Cauchy. Για λόγους πληρότητας παραθέτουμε στη συνέχεια τον ορισμό της διπλά αποκομμένης τυπικής κατανομής Cauchy.

### Ορισμός 2.1.2

Έστω μια τυχαία μεταβλητή που ακολουθεί την κατανομή Cauchy με παραμέτρους θέσης και κλίμακας, 0 και 1, αντίστοιχα. Έστω  $X$  η τυχαία μεταβλητή που εκφράζει τη διπλά αποκομμένη μορφή της κατανομής στο διάστημα  $[a, b]$ , όπου  $-\infty < a < b < \infty$ . Η σ.π.π. της  $X$  δίνεται από τη σχέση:

$$f_X(x) = \frac{\{1+x^2\}^{-1}}{\tan^{-1}(b) - \tan^{-1}(a)}. \quad (2.1.4)$$

Τότε γράφουμε ότι  $X \sim TC^{(a,b)}(0,1)$ .

**Παρατήρηση 2.1.1** Η ειδική περίπτωση της διπλά αποκομμένης κατανομής Cauchy που προκύπτει όταν  $a = -b$ ,  $\mu = 0$  και  $\sigma = 1$ , δηλαδή η ειδική περίπτωση της διπλά συμμετρικά αποκομμένης τυπικής κατανομής Cauchy πρωτοπαρουσιάστηκε από τους Johnson and Kotz (1970), ενώ ο Rohatgi (1976) δίνει τη σ.π.π. και τις πρώτες δύο ροπές της διπλά αποκομμένης τυπικής κατανομής Cauchy, δηλαδή της  $TC^{(a,b)}(0,1)$ , χωρίς να κάνει κανένα περαιτέρω σχόλιο. Στην εργασία των Daliya et al. (2001) εισάγεται, όσο είμαστε σε θέση να γνωρίζουμε, για πρώτη φορά η σ.π.π. του Ορισμού 2.1.1.

Εκτός από τις διπλά αποκομμένες κατανομές, όπως ήδη έχει αναφερθεί στο πρώτο κεφάλαιο, στη βιβλιογραφία έχουν εισαχθεί οι από αριστερά και οι από δεξιά αποκομμένες κατανομές ως ειδικές περιπτώσεις, όταν  $b = \infty$  ή  $a = -\infty$ , αντίστοιχα. Ακολούθως και για λόγους πληρότητας δίνουμε τους ορισμούς αυτών στην περίπτωση της κατανομής Cauchy.

Η από αριστερά αποκομμένη κατανομή Cauchy προκύπτει από τη σχέση

$$(1.2.5) \text{ λαμβάνοντας υπόψη ότι } G(a) = \frac{1}{\pi} \tan^{-1} \left( \frac{a - \mu}{\sigma} \right) + \frac{1}{2}.$$

### Ορισμός 2.1.3

Έστω μια τυχαία μεταβλητή που ακολουθεί την κατανομή Cauchy με παραμέτρους θέσης και κλίμακας,  $\mu$  και  $\sigma$ , αντίστοιχα. Έστω  $X$  η τυχαία μεταβλητή που εκφράζει την από αριστερά αποκομμένη μορφή της κατανομής στο διάστημα  $[\alpha, \infty)$ , όπου  $-\infty < \alpha < \infty$ . Η σ.π.π. της  $X$  είναι

$$f_X(x) = \frac{\left\{ 1 + \left( \frac{x - \mu}{\sigma} \right)^2 \right\}^{-1}}{\sigma \left[ \frac{\pi}{2} - \tan^{-1} \left( \frac{a - \mu}{\sigma} \right) \right]}. \quad (2.1.5)$$

Τότε γράφουμε ότι  $X \sim {}_LTC^{(a)}(\mu, \sigma)$  ή  $X \sim {}_LTC(\mu, \sigma, a)$ .

Από την άλλη μεριά η από δεξιά αποκομμένη κατανομή Cauchy προκύπτει

$$\text{από τη σχέση (1.2.7) λαμβάνοντας υπόψη ότι } G(b) = \frac{1}{\pi} \tan^{-1} \left( \frac{b - \mu}{\sigma} \right) + \frac{1}{2}.$$

### Ορισμός 2.1.4

Έστω μια τυχαία μεταβλητή που ακολουθεί την κατανομή Cauchy με παραμέτρους θέσης και κλίμακας,  $\mu$  και  $\sigma$ , αντίστοιχα. Έστω  $X$  η τυχαία μεταβλητή που εκφράζει την από δεξιά αποκομμένη μορφή της κατανομής στο διάστημα  $(-\infty, b]$ , όπου  $-\infty < b < \infty$ . Η σ.π.π. της  $X$  είναι η:

$$f_X(x) = \frac{\left\{1 + \left(\frac{x - \mu}{\sigma}\right)^2\right\}^{-1}}{\sigma \left[\frac{\pi}{2} + \tan^{-1}\left(\frac{b - \mu}{\sigma}\right)\right]} \quad (2.1.6)$$

Τότε γράφουμε ότι  $X \sim {}_R TC^{(b)}(\mu, \sigma)$  ή  $X \sim {}_R TC(\mu, \sigma, b)$ .

Έστω τώρα ότι  $X \sim TC^{(a,b)}(\mu, \sigma)$ . Στην πρόταση που ακολουθεί δίνεται η α.σ.κ. της τ.μ.  $X$ .

### Πρόταση 2.1.1

Έστω  $X$  μια τ.μ. που ακολουθεί τη διπλά αποκομμένη κατανομή Cauchy στο διάστημα  $[a, b]$ , δηλαδή  $X \sim TC^{(a,b)}(\mu, \sigma)$ . Τότε η α.σ.κ. της τ.μ.  $X$  ορίζεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{ για } x < a \\ \frac{\tan^{-1}\left(\frac{x - \mu}{\sigma}\right) - \tan^{-1}\left(\frac{a - \mu}{\sigma}\right)}{D} & , \text{ για } a \leq x < b \\ 1 & , \text{ για } x \geq b \end{cases} \quad (2.1.7)$$

όπου η σταθερά  $D$  προσδιορίστηκε στη σχέση (2.1.3).

### Απόδειξη

Η α.σ.κ. μιας διπλά αποκομμένης κατανομής δίνεται από τη σχέση (1.2.2). Με αντικατάσταση της α.σ.κ. της Cauchy που δίνεται στη σχέση (1.1.2) έχουμε ότι για  $a \leq x \leq b$  είναι:

$$F_X(x) = \frac{\tan^{-1}\left(\frac{x-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)}{D},$$

και η απόδειξη ολοκληρώθηκε. ■

Στην πρόταση που ακολουθεί προσδιορίζεται η α.σ.κ. της διπλά αποκομμένης τυπικής κατανομής Cauchy, η οποία προκύπτει με άμεση εφαρμογή της προηγούμενης πρότασης για  $\mu = 0$  και  $\sigma = 1$ .

### Πρόταση 2.1.2

Έστω  $X$  μια τ.μ. που ακολουθεί τη διπλά αποκομμένη τυπική κατανομή Cauchy, δηλαδή  $X \sim TC^{(a,b)}(0,1)$ . Τότε η α.σ.κ. της τ.μ.  $X$  ορίζεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{για } x < a \\ \frac{\tan^{-1}(x) - \tan^{-1}(a)}{D} & , \text{για } a \leq x < b \\ 1 & , \text{για } x \geq b \end{cases} \quad (2.1.8)$$

Όπως έχει ήδη αναφερθεί η κατανομή Cauchy αποτελεί μία από τις κατανομές που έχουν εισαχθεί στη βιβλιογραφία και χρησιμοποιούνται σε πρακτικές εφαρμογές με στόχο να περιγραφούν ικανοποιητικά φαινόμενα με πιο παχιές ουρές από αυτές της κανονικής κατανομής. Όμως η χρήση της είναι περιορισμένη λόγω του πεδίου ορισμού της που είναι το σύνολο των πραγματικών αριθμών και της μη ύπαρξης πεπερασμένων ροπών. Γίνεται άμεσα αντιληπτό ότι η αποκομμένη κατανομή Cauchy, που δεν είναι τίποτε άλλο παρά η δεσμευμένη κατανομή που προκύπτει από περιορισμό του πεδίου ορισμού της κατανομής Cauchy, χρησιμοποιείται σε πλήθος πρακτικών εφαρμογών για τη μοντελοποίηση

τυχαίων πραγματικών φαινομένων που η κατανομή Cauchy αδυνατεί να περιγράψει ικανοποιητικά λόγω των προαναφερθεισών αδυναμιών. Μεταξύ άλλων (βλέπε Nadarajah (2011)) η αποκομμένη κατανομή Cauchy εφαρμόζεται:

α) στη ρευστομηχανική για τη μελέτη της κλίση της ταχύτητας της τυρβώδους ροής μέσω των αριθμών Reynolds (Jimenez, 1996),

β) στην κρυσταλλογραφία (Mitra and Das, 1989),

γ) στη μελέτη του μεγέθους πακέτου (packet size) σε μοντέλα κυκλοφορίας. Ενδεικτικά αναφέρεται το παράδειγμα για το Finnish University and Research Network, που το μέγεθος πακέτου περιγράφεται ικανοποιητικά από την  $TC^{(a,b)}(\mu, \sigma)$  με  $\mu = 0,8$ ,  $\sigma = 1$ ,  $a = 0$  και  $b = 10$  (βλέπε Ni, 2001),

δ) στη μελέτη των μαγνητικών πεδίων (βλ. Borgia *et al.*, 1996, Fantazzini and Brown, 2005), και τέλος,

ε) σε οικονομικές μελέτες όπου θεωρείται εκ των προτέρων κατανομή (βλέπε Bauwens *et al.*, 1999).

Η πληθώρα των εφαρμογών της αποκομμένης κατανομής Cauchy είχε ως αποτέλεσμα οι ιδιότητες αυτής της κατανομής να αποτελέσουν αντικείμενο μελέτης. Η κριτική ανασκόπηση και μελέτη αυτών των ιδιοτήτων αποτελεί το αντικείμενο μελέτης της επόμενης ενότητας.

## 2.2 Ιδιότητες της κατανομής

Οι ιδιότητες της διπλά αποκομμένης κατανομής Cauchy  $TC^{(a,b)}(\mu, \sigma)$ , έχουν μελετηθεί από τους Nadarajah (2011), Nadarajah and Kotz (2006) και Daliya *et al.* (2001). Αντικείμενο αυτής της παραγράφου είναι η κριτική παρουσίαση, ανασκόπηση και μελέτη αυτών των ιδιοτήτων. Στο πλαίσιο αυτό, δίνονται εκφράσεις για τις ροπές, τη μέση απόκλιση από τη μέση τιμή και τη διάμεσο, τη ροπογεννήτρια και τη χαρακτηριστική συνάρτηση, τις εντροπίες του Shannon και



του Renyi, περιγραφικά μέτρα του σχήματος μιας κατανομής καθώς και για τις ασυμπτωτικές κατανομές του μέγιστου και ελάχιστου διατεταγμένου στατιστικού.

### Σχήμα και κορυφότητα της $TC^{(a,b)}(\mu, \sigma)$

Σε αυτήν την παράγραφο, λαμβάνοντας υπόψη την εργασία του Nadarajah (2011), θα εξεταστεί το σχήμα και η κορυφότητα της διπλά αποκομμένης κατανομής Cauchy,  $TC^{(a,b)}(\mu, \sigma)$ .

#### Πρόταση 2.2.1 (Nadarajah, 2011)

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Η σ.π.π. της τ.μ.  $X$  είναι μονοκόρυφη (unimodal). Ειδικότερα, προκύπτει ότι όταν  $a < \mu < b$  η κορυφή είναι στο  $x_0 = \mu$ , αν  $b < \mu$  η κορυφή είναι στο  $x_0 = b$ , ενώ αν  $\mu < a$  η κορυφή είναι στο  $x_0 = a$ .

#### Απόδειξη

Όταν η τ.μ.  $X \sim TC^{(a,b)}(\mu, \sigma)$  εύκολα προκύπτει ότι η παράγωγος ως προς  $x$  της σ.π.π. της που προσδιορίστηκε στη σχέση (2.1.2) είναι:

$$\frac{df_x(x)}{dx} = -\frac{2\sigma}{D} \frac{x - \mu}{\left[\sigma^2 + (x - \mu)^2\right]^2} \quad (2.2.1)$$

όπου

$$D = \tan^{-1}\left(\frac{b-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right) > 0,$$

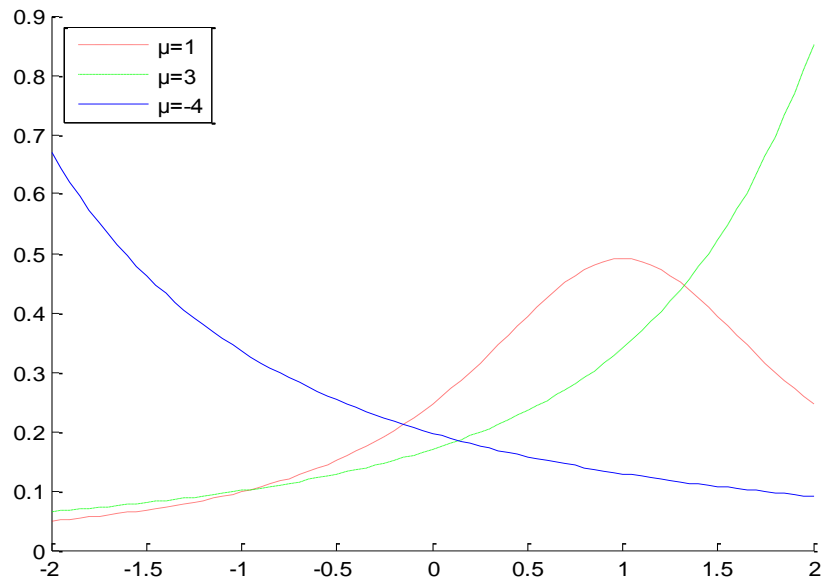
και  $\tan^{-1}(\cdot)$  συμβολίζεται η αντίστροφη συνάρτηση εφαπτομένης. Λαμβάνοντας υπόψη ότι  $D > 0$ , παρατηρούμε τα ακόλουθα:

α) Αν  $a < \mu < b$ , η πρώτη παράγωγος είναι θετική για  $x < \mu$  και αρνητική για  $x > \mu$ . Επομένως, συνεπάγεται ότι η σ.π.π. είναι αύξουσα μέχρι το  $x_0 = \mu$  και μετά φθίνουσα, άρα παρουσιάζει μέγιστο στο  $x_0 = \mu$ .

β) Αν  $\mu > b$ , λαμβάνοντας υπόψη ότι  $a \leq x \leq b$ , προκύπτει ότι η πρώτη παράγωγος είναι θετική για κάθε  $x \in [a, b]$ . Επομένως η σ.π.π. είναι αύξουσα σε όλο το πεδίο ορισμού της. Άρα παίρνει τη μέγιστη τιμή της στο άνω άκρο του πεδίου ορισμού της, δηλαδή για  $x_0 = b$ .

γ) Τέλος, αν  $\mu < a$ , λαμβάνοντας υπόψη ότι  $a \leq x \leq b$ , προκύπτει ότι η πρώτη παράγωγος είναι αρνητική για κάθε  $x \in [a, b]$ . Επομένως, η σ.π.π. είναι φθίνουσα σε όλο το πεδίο ορισμού της. Άρα παίρνει τη μέγιστη τιμή της στο κάτω άκρο του πεδίου ορισμού της, δηλαδή για  $x_0 = a$ . ■

Από την παραπάνω πρόταση γίνεται αντιληπτό ότι η κορυφότητα της διπλά αποκομμένης κατανομής Cauchy δεν επηρεάζεται από την τιμή της παραμέτρου  $\sigma$ . Στη συνέχεια, στο σχήμα που ακολουθεί δίνεται η γραφική παράσταση της σ.π.π. της συμμετρικά διπλά αποκομμένης κατανομής  $TC^{(-2,2)}(\mu, 1)$ , για τιμές της παραμέτρου  $\mu = 1, 3, -4$ .



Γράφημα 3- Σ.π.π. της  $TC^{(-2,2)}(\mu, 1)$  για  $\mu = 1, 3, -4$ , και  $x \in [-2, 2]$ .

Όντως, επιβεβαιώνεται ότι όταν  $\mu = 1$ , η σ.π.π. παρουσιάζει μέγιστο στο  $x_0 = 1$  καθώς τότε  $a < \mu < b$ , όταν  $\mu = 3 > b = 2$ , η σ.π.π. είναι αύξουσα σε όλο το πεδίο ορισμού της και παίρνει τη μέγιστη τιμή της για  $x_0 = 2$ , ενώ όταν  $\mu = -4 < a = -2$ , η σ.π.π. είναι φθίνουσα σε όλο το πεδίο ορισμού της και παίρνει τη μέγιστη τιμή της για  $x_0 = -2$ .

### Προσομοίωση δεδομένων από τη διπλά αποκομμένη κατανομή Cauchy

Η προσομοίωση δεδομένων από τη διπλά αποκομμένη κατανομή Cauchy είναι άμεση λαμβάνοντας υπόψη την επόμενη πρόταση, όπου υπολογίζεται η αντίστροφη της α.σ.κ., και τη σχέση (1.2.4).

**Πρόταση 2.2.2**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Η αντίστροφη συνάρτηση της αθροιστικής συνάρτησης κατανομής της τ.μ.  $X$ , δίνεται από τη σχέση:

$$F_X^{-1}(p) = \mu + \sigma \tan\left(\tan^{-1}\left(\frac{a-\mu}{\sigma}\right) + pD\right), p \in [0,1]. \quad (2.2.2)$$

**Απόδειξη**

Το επιθυμητό αποτέλεσμα προκύπτει συνδυάζοντας τα αποτελέσματα των Πρότασεων 1.1.1 και 1.2.1. Ειδικότερα, η αντίστροφη της α.σ.κ. της διπλά αποκομμένης κατανομής Cauchy σύμφωνα με τη σχέση (1.2.3) δίνεται από τη  $F_X^{-1}(p) = G^{-1}(G(a) + p(G(b) - G(a)))$ , όπου  $G(\cdot)$  και  $G^{-1}(\cdot)$  η α.σ.κ. και η αντίστροφης της α.σ.κ. της Cauchy, που προσδιορίστηκαν στις σχέσεις (1.1.2) και (1.1.3), αντίστοιχα. Εύκολα προκύπτει, με αλγεβρικές πράξεις, ότι:

$$F_X^{-1}(p) = \mu + \sigma \tan\left(\tan^{-1}(A) + pD\right),$$

όπου  $A = \left(\frac{a-\mu}{\sigma}\right)$  και  $B = \frac{b-\mu}{\sigma}$ . ■

Επομένως, σύμφωνα με την παραπάνω πρόταση και τη σχέση (1.2.4), αν κάποιος θέλει να δημιουργήσει ένα τ.δ.  $X_1, \dots, X_n$  από τη διπλά αποκομμένη κατανομή Cauchy με παραμέτρους  $a, b, \mu$  και  $\sigma$ , αρκεί να δημιουργήσει ένα τ.δ.  $U_1, \dots, U_n$  από την ομοιόμορφη στο διάστημα  $[0,1]$ , και να υπολογίσει τις τιμές

$$X_i = \mu + \sigma \tan\left[\tan^{-1}\left(\frac{a-\mu}{\sigma}\right) + U_i\left(\tan^{-1}\left(\frac{b-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)\right)\right],$$

για  $i = 1, \dots, n$ .

### Μέση τιμή, διακύμανση και ροπές της αποκομμένης κατανομής Cauchy

Μια από τις σημαντικότερες ιδιότητες της αποκομμένης κατανομής Cauchy είναι ότι η αποκοπή της εξαλείφει την αποκλίνουσα συμπεριφορά της Cauchy ως προς τη μη ύπαρξη πεπερασμένων ροπών. Αυτό συμβαίνει διότι οι ουρές της κατανομής, που είναι η αιτία του προβλήματος, αφαιρούνται. Αντικείμενο μελέτης σε αυτήν την παράγραφο είναι η  $n$  τάξης ροπή της  $TC^{(a,b)}(\mu, \sigma)$ , η μέση τιμή και η διακύμανσή της.

Οι Nadarajah and Kotz (2006) μας εφοδιάζουν με έναν άμεσο τύπο των ροπών για  $n$  ακέραιο ή πραγματικό, με τη βοήθεια της υπεργεωμετρικής Gauss συνάρτησης. Η υπεργεωμετρική Gauss συνάρτηση ορίζεται ως εξής:

$${}_2F_1(a, b; c; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!} \quad (2.2.3)$$

όπου με  $(z)_k$  παριστάνεται το αυξανόμενο παραγοντικό που ορίζεται από τη σχέση  $(z)_k = z(z+1)\cdots(z+k-1)$ .

Επιπλέον, σε όσα ακολουθούν αν  $z$  είναι ένας μιγαδικός αριθμός της μορφής  $z = x + iy$ , τότε με  $\operatorname{Re}(z)$  συμβολίζεται το πραγματικό μέρος του, δηλαδή  $\operatorname{Re}(z) = x$  και:

$$\arg(z) = \begin{cases} \tan^{-1}(y/x), & \text{όταν } x \neq 0 \\ -\pi/2, & \text{όταν } x = 0, y < 0 \\ \pi/2, & \text{όταν } x = 0, y > 0 \\ \text{δεν ορίζεται,} & \text{όταν } x = 0, y = 0 \end{cases}$$

**Θεώρημα 2.2.1 (Nadarajah and Kotz, 2006)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Τότε η  $n$  τάξης κεντρική ροπή της  $X$ , για  $n \geq 1$ , δίνεται από τη σχέση:

$$E(X^n) = \frac{\sum_{k=0}^n \frac{1}{k+1} \binom{n}{k} \mu^{n-k} \sigma^k \Lambda_k}{D}, \tag{2.2.4}$$

όπου

$$\Lambda_k = B^{k+1} {}_2F_1\left(1, \frac{k+1}{2}; 1 + \frac{k+1}{2}; -B^2\right) - A^{k+1} {}_2F_1\left(1, \frac{k+1}{2}; 1 + \frac{k+1}{2}; -A^2\right),$$

με  $A = \frac{a-\mu}{\sigma}$ ,  $B = \frac{b-\mu}{\sigma}$  και  $D = \tan^{-1}(B) - \tan^{-1}(A)$ .

**Απόδειξη**

Με βάση τον ορισμό της, η  $n$  τάξης κεντρική ροπή της  $X$ , για  $n \geq 1$ , είναι:

$$E(X^n) = \frac{1}{\sigma D} \int_a^b x^n \left\{ 1 + \left( \frac{x-\mu}{\sigma} \right)^2 \right\}^{-1} dx,$$

όπου η σταθερά  $D$  έχει προσδιοριστεί στη σχέση (2.1.3). Θεωρούμε το μετασχηματισμό  $y = (x-\mu)/\sigma$ , οπότε

$$E(X^n) = \frac{1}{D} \int_A^B \frac{(\mu + \sigma y)^n}{1+y^2} dy, \tag{2.2.5}$$

όπου  $A = \frac{a-\mu}{\sigma}$  και  $B = \frac{b-\mu}{\sigma}$ .

Χρησιμοποιώντας το ανάπτυγμα της διωνυμικής σειράς του Νεύτωνα, δηλαδή ότι

$$(\mu + \sigma y)^n = \sum_{k=0}^n \binom{n}{k} \mu^{n-k} \sigma^k y^k,$$

το δεξιό μέλος της σχέσης (2.2.5) γράφεται

$$\frac{\sum_{k=0}^n \binom{n}{k} \mu^{n-k} \sigma^k \int_A^B \frac{y^k}{1+y^2} dy}{D} = \frac{\sum_{k=0}^n \binom{n}{k} \mu^{n-k} \sigma^k \{I_k(B) - I_k(A)\}}{D}, \quad (2.2.6)$$

όπου

$$I_k(c) = \int_0^c \frac{y^k}{1+y^2} dy = \int_0^{y^2=z} \frac{z^{(k-1)/2}}{2(1+z)} dz.$$

Από τη μονογραφία των Gradshteyn and Ryzhik (2000) είναι γνωστό (βλέπε σχέση (3.194.5)) ότι:

$$\int_0^u \frac{x^{\lambda-1}}{1+\beta x} dx = \frac{u^\lambda}{\lambda} {}_2F_1(1, \lambda; 1+\lambda; -\beta u), \quad |\arg(1+\beta u)| < \pi, \quad \operatorname{Re}(\lambda) > 0,$$

όπου με  ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$  συμβολίζεται η υπεργεωμετρική συνάρτηση Gauss, όπως αυτή ορίστηκε στην σχέση (2.2.3). Επομένως, εφαρμόζοντας την σχέση (3.194.5)

των Gradshteyn and Ryzhik (2000) στην ειδική περίπτωση που  $u = c^2$ ,  $\lambda = \frac{k+1}{2}$

και  $\beta = 1$ , για τα οποία  $|\arg(1+\beta u)| = 0 < \pi$ ,  $\operatorname{Re}(\lambda) > 0$ , προκύπτει ότι

$$I_k(c) = \frac{c^{k+1}}{k+1} {}_2F_1\left(1, \frac{k+1}{2}; 1 + \frac{k+1}{2}; -c^2\right). \quad (2.2.7)$$

Συνδυάζοντας τις σχέσεις (2.2.5)-(2.2.7) προκύπτει το ζητούμενο αποτέλεσμα. ■

Λαμβάνοντας υπόψη το αποτέλεσμα του προηγούμενου θεώρηματος για  $n = 1$  είναι εφικτός ο προσδιορισμός της μέσης τιμής (βλέπε Nadarajah and Kotz, 2006).

### Πόρισμα 2.2.1 (Nadarajah and Kotz, 2006)

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Τότε η μέση τιμή της  $X$  δίνεται από τη σχέση:

$$E(X; a, b) = \mu + \frac{\sigma}{2D} \left\{ \log \left[ 1 + \left( \frac{b-\mu}{\sigma} \right)^2 \right] - \log \left[ 1 + \left( \frac{a-\mu}{\sigma} \right)^2 \right] \right\}. \quad (2.2.8)$$

### Απόδειξη

Λαμβάνοντας υπόψη το αποτέλεσμα του προηγούμενου θεώρηματος για  $n = 1$  έχουμε ότι:

$$E(X) = \frac{\mu\Lambda_0 + 0.5\sigma\Lambda_1}{D},$$

όπου

$$\Lambda_0 = B {}_2F_1\left(1, \frac{1}{2}; \frac{3}{2}; -B^2\right) - A {}_2F_1\left(1, \frac{1}{2}; \frac{3}{2}; -A^2\right),$$

και



$$\Lambda_1 = B^2 {}_2F_1(1, 1; 2; -B^2) - A^2 {}_2F_1(1, 1; 2; -A^2),$$

$$\text{με } A = \frac{a-\mu}{\sigma}, B = \frac{b-\mu}{\sigma} \text{ και } D = \tan^{-1}(B) - \tan^{-1}(A).$$

Από τις ιδιότητες της υπεργεωμετρικής συνάρτησης Gauss (βλέπε Gradshteyn and Ryzhik, 2000) έχουμε ότι:

$${}_2F_1\left(1, \frac{1}{2}; \frac{3}{2}; -c^2\right) = \frac{\tan^{-1}(c)}{c},$$

και

$${}_2F_1(1, 1; 2; -c^2) = \frac{1}{c^2} \log(1+c^2).$$

Επομένως, εύκολα προκύπτει ότι:

$$\Lambda_0 = \tan^{-1}(B) - \tan^{-1}(A) = D,$$

και

$$\Lambda_1 = \log(1+B^2) - \log(1+A^2).$$

Άρα,

$$E(X) = \frac{\mu\Lambda_0 + 0.5\sigma\Lambda_1}{D} = \mu + \frac{\sigma}{2D} [\log(1+B^2) - \log(1+A^2)],$$

και η απόδειξη ολοκληρώθηκε. ■

Τα επόμενα δύο πορίσματα δίνουν τους άμεσους τύπους εύρεσης των ροπών μέχρι 4<sup>ης</sup> τάξης, στην περίπτωση της διπλά αποκομμένης τυπικής

κατανομής Cauchy και στην περίπτωση της διπλά συμμετρικά αποκομμένης τυπικής κατανομής Cauchy, αντίστοιχα.

**Πόρισμα 2.2.2 (Nadarajah and Kotz, 2006)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(0,1)$ . Οι τέσσερις πρώτης τάξης κεντρικές ροπές της  $X$  δίνονται από τις σχέσεις:

$$E(X) = \{\log(1+b^2) - \log(1+a^2)\} / (2D),$$

$$E(X^2) = \{\tan^{-1}(a) - \tan^{-1}(b) - a + b\} / (D),$$

$$E(X^3) = \{\log(1+a^2) - \log(1+b^2) - a^2 + b^2\} / (2D),$$

και

$$E(X^4) = \{3 \tan^{-1}(b) - 3 \tan^{-1}(a) - a^3 + b^3 + 3a - 3b\} / (3D).$$

όπου  $D = \tan^{-1}(b) - \tan^{-1}(a)$ .

**Απόδειξη**

Η πρώτη σχέση προκύπτει άμεσα από το αποτέλεσμα του Πορίσματος 2.2.1 λαμβάνοντας υπόψη ότι  $\mu = 0$  και  $\sigma = 1$  και επομένως  $D = \tan^{-1}(b) - \tan^{-1}(a)$ . Σύμφωνα με το Θεώρημα 2.2.1 λαμβάνοντας υπόψη ότι  $\mu = 0$  και  $\sigma = 1$  προκύπτει άμεσα ότι:

$$E(X^2) = \frac{\Lambda_2}{3D} = \frac{b^3 {}_2F_1\left(1, \frac{3}{2}; \frac{5}{2}; -b^2\right) - a^3 {}_2F_1\left(1, \frac{3}{2}; \frac{5}{2}; -a^2\right)}{3D}.$$

Ισχύει ότι:

$${}_2F_1\left(1, \frac{3}{2}; \frac{5}{2}; -c^2\right) = \frac{3}{c^2} - \frac{3 \tan^{-1}(c)}{c^3},$$

οπότε προκύπτει με αλγεβρικές πράξεις ότι

$$E(X^2) = \left\{ \tan^{-1}(a) - \tan^{-1}(b) - a + b \right\} / (D).$$

Επιπλέον, σύμφωνα με το Θεώρημα 2.2.1 λαμβάνοντας υπόψη ότι  $\mu = 0$  και  $\sigma = 1$ , προκύπτει άμεσα ότι:

$$E(X^3) = \frac{\Lambda_3}{4D} = \frac{b^4 {}_2F_1(1, 2; 3; -b^2) - a^4 {}_2F_1(1, 2; 3; -a^2)}{4D}.$$

Είναι:

$${}_2F_1(1, 2; 3; -c^2) = \frac{-2(-c^2 + \log(1+c^2))}{c^4},$$

οπότε με αλγεβρικές πράξεις έχουμε ότι:

$$E(X^3) = \left\{ \log(1+a^2) - \log(1+b^2) - a^2 + b^2 \right\} / (2D).$$

Τέλος, σύμφωνα με το Θεώρημα 2.2.1 λαμβάνοντας υπόψη ότι  $\mu = 0$  και  $\sigma = 1$  προκύπτει άμεσα ότι

$$E(X^4) = \frac{\Lambda_4}{5D} = \frac{b^5 {}_2F_1\left(1, \frac{5}{2}; \frac{7}{2}; -b^2\right) - a^5 {}_2F_1\left(1, \frac{5}{2}; \frac{7}{2}; -a^2\right)}{5D}.$$

Καθώς

$${}_2F_1\left(1, \frac{5}{2}; \frac{7}{2}; -c^2\right) = \frac{5(c^2 - 3)}{3c^4} + \frac{5 \tan^{-1}(c)}{c^5},$$

προκύπτει με αλγεβρικές πράξεις ότι

$$E(X^4) = \{3 \tan^{-1}(b) - 3 \tan^{-1}(a) - a^3 + b^3 + 3a - 3b\} / (3D). \quad \blacksquare$$

**Πόρισμα 2.2.3 (Nadarajah and Kotz, 2006)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(-b,b)}(0,1)$ . Τότε όλες οι ροπές περιττής τάξης της  $X$  μηδενίζονται, ενώ οι δύο πρώτες ροπές ζυγής τάξης δίνονται από τις σχέσεις:

$$E(X^2) = \{b - \tan^{-1}(b)\} / (\tan^{-1}(b)),$$

$$E(X^4) = \{b^3 - 3b + 3 \tan^{-1}(b)\} / (3 \tan^{-1}(b))$$

**Απόδειξη**

Όταν  $a = -b$ , προκύπτει ότι  $\tan^{-1}(a) = \tan^{-1}(-b) = -\tan^{-1}(b)$ . Από το Θεώρημα 2.2.1 για  $n$  περιττό,  $\mu = 0$  και  $\sigma = 1$  έχουμε ότι:

$$E(X^n) = \frac{\Lambda_n}{2(n+1) \tan^{-1}(b)} = 0,$$

καθώς, σε αυτήν την περίπτωση είναι:

$$\Lambda_n = b^{n+1} {}_2F_1\left(1, \frac{n+1}{2}; 1 + \frac{n+1}{2}; -b^2\right) - (-b)^{n+1} {}_2F_1\left(1, \frac{n+1}{2}; 1 + \frac{n+1}{2}; -(-b)^2\right) = 0,$$

διότι  $(-b)^{n+1} = b^{n+1}$ , για κάθε  $n+1$  άρτιο.

Επιπρόσθετα, από το Πρόρισμα 2.2.2, λαμβάνοντας υπόψη ότι  $a = -b$ , επομένως ότι  $\tan^{-1}(a) = \tan^{-1}(-b) = -\tan^{-1}(b)$ , έχουμε ότι οι δύο πρώτες ζυγές ροπές είναι:

$$E(X^2) = \frac{\tan^{-1}(a) - \tan^{-1}(b) - a + b}{D} = \frac{-2 \tan^{-1}(b) + 2b}{\tan^{-1}(b) + \tan^{-1}(b)} = \frac{b - \tan^{-1}(b)}{\tan^{-1}(b)}$$

και

$$E(X^4) = \frac{3 \tan^{-1}(b) - 3 \tan^{-1}(a) - a^3 + b^3 + 3a - 3b}{3D} = \frac{6 \tan^{-1}(b) + 2b^3 - 6b}{6 \tan^{-1}(b)},$$

αντίστοιχα, και η απόδειξη ολοκληρώθηκε. ■

**Παρατήρηση 2.2.2 α)** Σύμφωνα με τους Daliya et al. (2001) για  $p > 1$  ισχύει η ακόλουθη αναδρομική σχέση:

$$E(X^p) = \frac{\sigma}{D} \frac{b^{(p-1)} - a^{(p-1)}}{p-1} + 2\mu E(X^{p-1}) - (\mu^2 + \sigma^2) E(X^{p-2}).$$

Ειδικότερα, για  $p = 2$  προκύπτει ότι

$$E(X^2) = \frac{\sigma}{D} (b - a) + 2\mu E(X) - (\mu^2 + \sigma^2).$$

Λαμβάνοντας υπόψη τις εκφράσεις των  $E(X)$  και  $E(X^2)$  είναι:

$$\text{Var}(X) = \sigma \left[ \frac{b-a}{D} - \sigma \right] - [\mu - E(X)]^2.$$

Επομένως, για σταθεροποιημένο εύρος  $b - a$  εξάγεται το συμπέρασμα ότι η διακύμανση μεγιστοποιείται όταν η αποκοπή είναι συμμετρική, όταν δηλαδή  $a = -b$ .

**β)** Σε μία πρόσφατη εργασία (βλέπε Kim 2008) προσδιορίστηκαν οι  $n$  τάξης ροπές για την περίπτωση της διπλά αποκομμένης  $t$  κατανομής με  $\nu$  βαθμούς ελευθερίας, η οποία περικλείει ως ειδική περίπτωση τη διπλά αποκομμένη Cauchy, για  $\nu = 1$ . Ο προσδιορισμός έγινε χρησιμοποιώντας ότι η υπό μελέτη κατανομή δεν είναι τίποτε άλλο παρά μίξη κλίμακας μιας διπλά αποκομμένης κανονικής κατανομής. Όμως, όπως επισημαίνει ο Nadarajah (2008), ο προσδιορισμός αυτός γίνεται υπό περιορισμούς που δεν επιτρέπουν για παράδειγμα την εύρεση της μέσης τιμής της διπλά αποκομμένης κατανομής Cauchy. Επιπλέον, οι σχέσεις αυτές δεν είναι άμεσες.

Στον Πίνακα 1 που ακολουθεί παρουσιάζονται οι τιμές των τεσσάρων πρώτων ροπών και της διακύμανσης της  $TC^{(a,b)}(\mu, \sigma)$  για επιλεγμένες τιμές των παραμέτρων. Παρατηρώντας τα αποτελέσματα του Πίνακα 1 προκύπτουν τα ακόλουθα:

α) για την διπλά συμμετρικά αποκομμένη στο  $(-5, 5)$  κατανομή Cauchy ισχύει ότι αν  $X \sim TC^{(-5,5)}(\mu, \sigma)$  και  $Y \sim TC^{(-5,5)}(-\mu, \sigma)$ , τότε οι ροπές άρτιας τάξης αυτών είναι ίδιες, ενώ οι περιττές είναι αντίθετες. Επίσης, όταν  $X \sim TC^{(-5,5)}(\mu, \sigma)$  όσο αυξάνουν κατά απόλυτη τιμή οι τιμές της παραμέτρου  $\mu$  αυξάνουν κατά απόλυτη τιμή και οι τιμές των τεσσάρων πρώτων ροπών και της διακύμανσης. Τέλος, για σταθεροποιημένη την τιμή της παραμέτρου  $\mu$ ,  $\mu > 0$ , όσο αυξάνει η τιμή της παραμέτρου  $\sigma$  αυξάνει η τιμή της διακύμανσης. Το ίδιο ισχύει και για  $\mu < 0$ , από αποτελέσματα που δεν παρατίθενται, αλλά είναι διαθέσιμα.

β) Για τη διπλά αποκομμένη στο  $(0, 10)$  προκύπτει ότι οι περιττής τάξης ροπές είναι θετικές, ενώ για τη διπλά αποκομμένη στο  $(-5, 0)$  είναι αρνητικές. Επιπλέον, εξάγεται το συμπέρασμα ότι για σταθερή τιμή της παραμέτρου  $\sigma$ , όσο

αυξάνουν οι τιμές της παραμέτρου  $\mu$ , στην περίπτωση αποκοπής στο  $(0,10)$  μειώνεται η τιμή της διακύμανσης, ενώ στο  $(-5,0)$  αυξάνεται. Επιπλέον, και στις δύο περιπτώσεις, για σταθεροποιημένη την τιμή της παραμέτρου  $\mu$ ,  $\mu > 0$ , όσο αυξάνει η τιμή της παραμέτρου  $\sigma$  αυξάνει η τιμή της διακύμανσης. Το ίδιο συμπέρασμα προκύπτει και για  $\mu < 0$ , από αποτελέσματα που δεν παρατίθενται, αλλά είναι διαθέσιμα.

**Πίνακας 1:** Οι τέσσερις πρώτες τάξης ροπές περί το μηδέν και η διακύμανση, της τ.μ.  $X$ , με σ.π.π. της μορφής (2.2.1), για διάφορες τιμές των παραμέτρων  $\mu, \sigma, a$  και  $b$ .

$(a, b)$	$\sigma$	$\mu$	$\alpha_1 = EX$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$Var(X)$
$a = 0$ $b = 10$	1	-3	2.7826	14.1245	91.5270	670.2563	6.3817
		-2	2.4248	11.5817	72.9531	526.3046	5.7023
		-1	1.9586	8.4768	51.0989	360.6457	4.6407
		0	1.5686	5.7975	32.4190	220.7860	3.3371
		1	1.8269	6.1070	30.8267	197.8819	2.7695
		2	2.5022	8.9250	42.7690	256.9863	2.6638
	2	3	3.3005	13.5372	66.8892	390.4370	2.6439
		2	3.0137	13.5279	77.3683	517.0238	4.4457
		3	3.4379	17.4180	108.3114	762.3580	5.5988
		4	3.7651	20.5251	134.1228	974.8149	6.3493
10	4.6113	29.1060	210.1574	1635.6372	7.8423		
$a = -5$ $b = 5$	1	-3	-2.4978	8.9027	-28.4385	114.2377	2.6638
		-2	-1.6995	5.5322	-13.6313	57.9827	2.6439
		-1	-0.8576	3.3763	-5.0374	33.8307	2.6408
		0	0	2.6406	0	27.6977	2.6406
		1	0.8576	3.3763	5.0374	33.8307	2.6408
		2	1.6995	5.5322	13.6313	57.9827	2.6439
	2	3	2.4978	8.9027	28.4385	114.2377	2.6638
		2	1.3823	6.3195	14.2191	79.5714	4.4086
		3	1.1005	6.7765	12.7990	91.2207	5.5653
		4	0.8727	7.0873	10.8962	98.4777	6.3258
10	0.2675	7.9125	3.8253	116.0909	7.8409		
$a = -5$ $b = 0$	1	-3	-2.8529	9.2395	-32.2132	118.5679	1.1004
		-2	-2.1471	5.7104	-17.4114	58.7774	1.1004
		-1	-1.5068	3.3820	-9.6710	32.3140	1.1114
		0	-1.1861	2.6406	-7.9154	27.6977	1.2337
		1	-1.3521	3.3571	-10.7348	38.9935	1.5289
		2	-1.5782	4.2271	-14.0504	52.1627	1.7364
	2	3	-1.7412	4.8824	-16.6179	62.5507	1.8504
		2	-1.7288	4.8050	-16.2501	60.8936	1.8164
		3	-1.8817	5.4292	-18.7112	70.8758	1.8884
		4	-2.0090	5.9774	-20.9438	80.1223	1.9413
10	-2.3495	7.5706	-27.7919	109.5583	2.0505		

**Ροπογεννήτρια και χαρακτηριστική συνάρτηση της  $TC^{(a,b)}(\mu, \sigma)$**

Σε αυτήν την ενότητα, λαμβάνοντας υπόψη την εργασία των Nadarajah and Kotz (2006), θα δοθεί η ργ.σ. της διπλά αποκομμένης κατανομής Cauchy καθώς επίσης και η χαρακτηριστική της συνάρτηση (χ.σ.). Σε όσα ακολουθούν  $i^2 = -1$ , δηλαδή  $i$  είναι η φανταστική μονάδα και με  $Ei(x)$  συμβολίζεται το

$$Ei(x) = \int_{-\infty}^x \frac{\exp(t)}{t} dt.$$

**Θεώρημα 2.2.2 (Nadarajah, 2011)**

Έστω  $X$  μια τ.μ. η οποία ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Τότε η ργ.σ. της  $X$  δίνεται από τη σχέση:

$$M(t) = \frac{\exp(\mu t - \sigma t/i)}{2Di} \left\{ Ei\left(B\sigma t + \frac{\sigma t}{i}\right) - Ei\left(A\sigma t + \frac{\sigma t}{i}\right) \right\} - \frac{\exp(\mu t + \sigma t/i)}{2Di} \left\{ Ei\left(B\sigma t - \frac{\sigma t}{i}\right) - Ei\left(A\sigma t - \frac{\sigma t}{i}\right) \right\}, \tag{2.2.9}$$

όπου  $A = \frac{a - \mu}{\sigma}$  και  $B = \frac{b - \mu}{\sigma}$ .

**Απόδειξη**

Με βάση τον ορισμό της ροπογεννήτριας συνάρτησης είναι  $M(t) = E[\exp(tX)]$

και αντικαθιστώντας τον τύπο της  $f_X(x)$  από τη σχέση (2.1.1) έχουμε ότι:

$$M(t) = \frac{1}{\sigma D} \int_a^b \exp(tx) \left\{ 1 + \left( \frac{x - \mu}{\sigma} \right)^2 \right\}^{-1} dx,$$



όπου η σταθερά  $D$  έχει προσδιοριστεί στη σχέση (2.1.3). Θεωρούμε έπειτα το μετασχηματισμό  $y = (x - \mu)/\sigma$ , οπότε

$$M(t) = \frac{\exp(\mu t)}{D} \int_A^B \exp(\sigma ty) \{1 + y^2\}^{-1} dy,$$

όπου  $A = \frac{a - \mu}{\sigma}$ ,  $B = \frac{b - \mu}{\sigma}$  και  $D = \tan^{-1}(B) - \tan^{-1}(A)$ .

Λαμβάνοντας υπόψη ότι:

$$\frac{1}{1 + y^2} = \frac{1}{\left(y - \frac{1}{i}\right)\left(y + \frac{1}{i}\right)} = \frac{i}{2\left(y - \frac{1}{i}\right)} - \frac{i}{2\left(y + \frac{1}{i}\right)} = \frac{1}{-i2\left(y - \frac{1}{i}\right)} + \frac{1}{2i\left(y + \frac{1}{i}\right)},$$

προκύπτει ότι:

$$M(t) = \frac{\exp(\mu t)}{2Di} \left[ \int_A^B \frac{\exp(\sigma ty)}{\left(y + \frac{1}{i}\right)} dy - \int_A^B \frac{\exp(\sigma ty)}{\left(y - \frac{1}{i}\right)} dy \right].$$

Από τη σχέση (3.352.1) των Gradshteyn and Ryzhik (2000) ισχύει ότι:

$$\int_0^u \frac{\exp(-mx)}{x + \beta} dx = \exp(m\beta) [Ei(-mu - m\beta) - Ei(-m\beta)], \quad |\arg \beta| < \pi.$$

Εφαρμόζοντας την παραπάνω σχέση για  $(m, \beta) = \left(\sigma t, \frac{1}{i}\right)$  και

$(m, \beta) = \left(\sigma t, -\frac{1}{i}\right)$ , αντίστοιχα, προκύπτει ύστερα από λίγη άλγεβρα, το επιθυμητό αποτέλεσμα. ■

Παρόμοια προσδιορίζεται, στο επόμενο θεώρημα, η χ.σ. λαμβάνοντας υπόψη ότι  $\varphi(t) = E[\exp(itX)]$ .

**Θεώρημα 2.2.3 (Nadarajah, 2011)**

Έστω  $X$  μια τ.μ. η οποία ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Τότε η χ.σ. της  $X$  δίνεται από τη σχέση:

$$\begin{aligned} \varphi(t) = & \frac{\exp(\mu it - \sigma t)}{2Di} \{Ei(B\sigma it + \sigma t) - Ei(A\sigma it + \sigma t)\} \\ & - \frac{\exp(\mu it + \sigma t)}{2Di} \{Ei(B\sigma it - \sigma t) - Ei(A\sigma it - \sigma t)\} \end{aligned}, \quad (2.2.10)$$

όπου  $A = \frac{a - \mu}{\sigma}$  και  $B = \frac{b - \mu}{\sigma}$ .

**Μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο**

Η μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο, αντίστοιχα, αποτελούν μέτρα της διασποράς της κατανομής και σκοπός τους είναι να δώσουν μια εικόνα του πως είναι κατανεμημένες οι πιθανότητες στις διάφορες τιμές της τ.μ. Ο Nadarajah (2011) μας εφοδιάζει με τις μαθηματικές εκφράσεις για τη μέση

απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο όταν η τ.μ.  $X \sim TC^{(a,b)}(\mu, \sigma)$ . Σε όσα ακολουθούν συμβολίζεται με  $E(X; a, b)$  και  $D(a, b)$  η μέση τιμή και η σταθερά της  $TC^{(a,b)}(\mu, \sigma)$ , αντίστοιχα.

#### Θεώρημα 2.2.4 (Nadarajah 2011)

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$  με μέση τιμή έστω  $\theta$  και διάμεσο  $M$ . Τότε η μέση απόλυτη απόκλιση από τη μέση τιμή, έστω  $\delta_1(X)$ , και τη διάμεσο, έστω  $\delta_2(X)$ , δίνονται από τις σχέσεις:

$$\delta_1(X) = 2 \left\{ \theta F_X(\theta) - \frac{D(a, \theta) E(X; a, \theta)}{D(a, b)} \right\} \quad (2.2.11)$$

και

$$\delta_2(X) = E(X; a, b) - 2 \frac{D(a, M) E(X; a, M)}{D(a, b)}, \quad (2.2.12)$$

αντίστοιχα, όπου  $F_X(\cdot)$ ,  $E(X; a, b)$  και  $D(a, b)$  η α.σ.κ., η μέση τιμή και η σταθερά κανονικοποίησης της τ.μ.  $X$ .

#### Απόδειξη

Με βάση τον ορισμό της μέσης απόλυτης απόκλισης από τη μέση τιμή είναι:

$$\begin{aligned} \delta_1(X) &= \int_a^b |x - \theta| f_X(x) dx = \int_a^\theta (\theta - x) f_X(x) dx + \int_\theta^b (x - \theta) f_X(x) dx \\ &= \theta \int_a^\theta f_X(x) dx - \int_a^\theta x f_X(x) dx + \int_\theta^b x f_X(x) dx - \theta \int_\theta^b f_X(x) dx \\ &= \theta F_X(\theta) - \int_a^\theta x f_X(x) dx + \int_\theta^b x f_X(x) dx - \theta [1 - F_X(\theta)]. \end{aligned}$$

Το ζητούμενο αποτέλεσμα προκύπτει εύκολα λαμβάνοντας υπόψη ότι:

$$\int_{\theta}^b xf_X(x) dx = \int_a^b xf_X(x) dx - \int_a^{\theta} xf_X(x) dx = E(X; a, b) - \int_a^{\theta} xf_X(x) dx$$

και

$$\begin{aligned} \int_a^{\theta} xf_X(x) dx &= \int_a^{\theta} \frac{x}{\sigma} \frac{\left\{1 + \left(\frac{x-\mu}{\sigma}\right)^2\right\}^{-1}}{\tan^{-1}\left(\frac{b-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)} dx \\ &= \frac{\tan^{-1}\left(\frac{\theta-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)}{\tan^{-1}\left(\frac{b-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)} \int_a^{\theta} \frac{x}{\sigma} \frac{\left\{1 + \left(\frac{x-\mu}{\sigma}\right)^2\right\}^{-1}}{\tan^{-1}\left(\frac{\theta-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)} dx, \end{aligned}$$

ή ισοδύναμα,

$$\int_a^{\theta} xf_X(x) dx = \frac{D(a, \theta)}{D(a, b)} E(X; a, \theta). \quad (2.2.13)$$

Με παρόμοιο τρόπο:

$$\begin{aligned} \delta_2(X) &= MF_X(M) - \int_a^M xf_X(x) dx + \int_M^b xf_X(x) dx - M[1 - F_X(M)] \\ &= \int_a^b xf_X(x) dx - 2 \int_a^M xf_X(x) dx, \end{aligned}$$

καθώς, αφού  $M$  η διάμεσος, ισχύει ότι  $F_X(M) = 0.5$ , και

$$\int_M^b xf_X(x) dx = \int_a^b xf_X(x) dx - \int_a^M xf_X(x) dx. \text{ Επομένως,}$$

$$\delta_2(X) = E(X; a, b) - 2 \int_a^M xf_X(x) dx.$$

Όμως, λαμβάνοντας υπόψη τη σχέση (2.2.13)

$$\int_a^M xf_X(x) dx = \frac{D(a, M)E(X; a, M)}{D(a, b)},$$

οπότε προκύπτει το ζητούμενο αποτέλεσμα. ■

Για επιλεγμένες τιμές των παραμέτρων της  $TC^{(a,b)}(\mu, \sigma)$  υπολογίζουμε τη μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο. Οι τιμές αυτές παρουσιάζονται στον Πίνακα 2. Από τα αποτελέσματα του Πίνακα 2 προκύπτει ότι όταν  $X \sim TC^{(0,10)}(\mu, \sigma)$ , τότε  $\delta_1(X) > \delta_2(X)$ . Επιπλέον, ισχύει το αναμενόμενο αποτέλεσμα ότι για τη διπλά συμμετρικά αποκομμένη στο  $(-5, 5)$  κατανομή Cauchy ισχύει ότι αν  $X \sim TC^{(-5,5)}(\mu, \sigma)$  και  $Y \sim TC^{(-5,5)}(-\mu, \sigma)$ , τότε οι μέσες απόλυτες αποστάσεις αυτών από τη μέση τιμή και τη διάμεσο είναι ίσες.

**Πίνακας 2:** Μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο τυχαίας μεταβλητής  $X$ , με σ.π.π. της μορφής (2.2.1), για διάφορες τιμές των παραμέτρων  $\mu, \sigma, a$  και  $b$ .

$(a, b)$	$\sigma$	$\mu$	$\theta$	M	$\delta_1(X)$	$\delta_2(X)$
$a = 0$ $b = 10$	1	-3	2.7826	1.9519	2.0577	1.9580
		-2	2.4248	1.5661	1.8956	1.7725
		-1	1.9586	1.1350	1.6326	1.4873
		0	1.5686	0.9050	1.2955	1.1619
		1	1.8269	1.3508	1.1639	1.0839
		2	2.5022	2.1713	1.1608	1.1212
		3	3.3005	3.0902	1.1734	1.1573
	2	2	3.0137	2.5540	1.6447	1.6003
	3		3.4379	2.9676	1.9217	1.8858
	4		3.7651	3.3333	2.0911	2.0650
10		4.6113	4.4331	2.4049	2.4015	
$a = -5$ $b = 5$	1	-3	-2.4978	-2.8287	1.1608	1.1212
		-2	-1.6995	-1.9098	1.1734	1.1573
		-1	-0.8576	-0.9601	1.1828	1.1790
		0	0	0	1.1861	1.1861
		1	0.8576	0.9601	1.1828	1.1790
		2	1.6995	1.9098	1.1734	1.1573
	3	2.4978	2.8287	1.1608	1.1212	
	2	2	1.3823	1.6878	1.6522	1.6325
	3		1.1005	1.4223	1.9259	1.9091
	4		0.8727	1.1722	2.0940	2.0815
10	0.2675		0.3900	2.4056	2.4040	

### Εντροπία του Renyi και του Shannon

Στην ενότητα αυτή θα παρουσιαστούν τα αποτελέσματα του Nadarajah (2011) που αφορούν την εντροπία του Renyi (1961) και του Shannon (1968). Το ενδιαφέρον επικεντρώνεται στον προσδιορισμό της εντροπίας καθώς αυτή αποτελεί ένα μέτρο της αβεβαιότητας και επιπλέον δίνει χρήσιμες πληροφορίες σχετικά με το σχήμα της κατανομής.

Η εντροπία του Renyi (1961) ορίζεται ως εξής.

#### Ορισμός 2.2.1 (Renyi, 1961)

Έστω  $X$  μια συνεχής τ.μ. με σ.π.π. την  $f(\cdot)$ . Η εντροπία του Renyi ορίζεται από τη σχέση:

$$H_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \int_{-\infty}^{+\infty} (f(x))^\gamma \right\} \quad (2.2.14)$$

όπου  $\gamma \in \mathbb{R}^+ - \{1\}$ .

Στην επόμενη πρόταση προσδιορίζεται συναρτήσει της υπεργεωμετρικής συνάρτησης Gauss η εντροπία του Renyi (1961) όταν η τ.μ.  $X$  ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από τη σχέση (2.1.1).

#### Πρόταση 2.2.3 (Nadarajah 2011)

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από τη σχέση (2.1.1). Η εντροπία του Renyi δίνεται από τη σχέση:

$$H_R(\gamma) = \frac{1}{1-\gamma} \log \{ C^{-\gamma} \sigma H(\gamma) \}, \quad (2.2.15)$$

όπου

$$H(\gamma) = B {}_2F_1\left(\frac{1}{2}, \gamma; \frac{3}{2}; -B^2\right) - A {}_2F_1\left(\frac{1}{2}, \gamma; \frac{3}{2}; -A^2\right), \quad (2.2.16)$$

$$A = \frac{a-\mu}{\sigma}, \quad B = \frac{b-\mu}{\sigma}, \quad C = \sigma D \quad \text{και} \quad D = \tan^{-1}(B) - \tan^{-1}(A).$$

### Απόδειξη

Σύμφωνα με τον Ορισμό 2.1.1 η εντροπία του Renyi για τη διπλά αποκομμένη κατανομή Cauchy που η σ.π.π. της δίνεται από τη σχέση (2.1.1) είναι η εξής:

$$H_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \int_a^b \frac{\left[ 1 + \left( \frac{x-\mu}{\sigma} \right)^2 \right]^{-1}}{\sigma D} dx \right\}^{\gamma}. \quad (2.2.17)$$

Μετά από αλγεβρικές πράξεις η σχέση (2.2.17) γράφεται ισοδύναμα,

$$H_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \left( \frac{1}{\sigma D} \right)^{\gamma} \int_a^b \left[ 1 + \left( \frac{x-\mu}{\sigma} \right)^2 \right]^{-\gamma} dx \right\}, \quad (2.2.18)$$



όπου  $D$  η σταθερά που έχει προσδιοριστεί στη σχέση (2.1.3).

Κάνοντας χρήση του μετασχηματισμού  $y = \frac{x-\mu}{\sigma}$  και θέτοντας  $C = \sigma D$ ,

$A = \frac{a-\mu}{\sigma}$  και  $B = \frac{b-\mu}{\sigma}$  προκύπτει ότι:

$$H_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \sigma C^{-\gamma} \int_A^B [1+y^2]^{-\gamma} dy \right\}. \quad (2.2.19)$$

Έπειτα, με χρήση του μετασχηματισμού  $z = y^2$ , η σχέση (2.2.19) γράφεται ως εξής:

$$\begin{aligned} H_R(\gamma) &= \frac{1}{1-\gamma} \log \left\{ \frac{1}{2} \sigma C^{-\gamma} \int_{A^2}^{B^2} \frac{z^{-1/2}}{(1+z)^\gamma} dz \right\} \\ &= \frac{1}{1-\gamma} \log \left\{ 2^{-1} C^{-\gamma} \sigma [J(B) - J(A)] \right\}, \end{aligned} \quad (2.2.20)$$

όπου  $J(c) = \int_0^c \frac{z^{-1/2}}{(1+z)^\gamma} dz$ .

Από τη μονογραφία των Gradshteyn and Ryzhik (2000) είναι γνωστό (βλέπε σχέση (3.194.1)) ότι:

$$J(c) = 2c {}_2F_1\left(\frac{1}{2}, \gamma; \frac{3}{2}; -c^2\right), \quad (2.2.21)$$

όπου με  ${}_2F_1(\cdot, \cdot; \cdot)$  η υπεργεωμετρική συνάρτηση Gauss (βλέπε σχέση 2.2.3).

Συνδυάζοντας τις (2.2.20) και (2.2.21) προκύπτει το ζητούμενο. ■

Η εντροπία του Shannon (1948) ορίζεται ως εξής.

**Ορισμός 2.2.2 (Shannon, 1948)**

Έστω  $X$  μια συνεχής τ.μ. με σ.π.π. την  $f_X(\cdot)$ . Η εντροπία του Shannon ορίζεται από τη σχέση:

$$H_{Sh}(x) = E[-\log(f_X(x))]. \tag{2.2.22}$$

Η εντροπία του Shannon για μια τ.μ.  $X$  αποτελεί ειδική περίπτωση της εντροπίας του Renyi, για  $\gamma \rightarrow 1$ . Παίρνοντας το όριο για  $\gamma \rightarrow 1$  της εντροπίας του Renyi, όπως αυτή προσδιορίστηκε στην Πρόταση 2.2.3, εμφανίζεται απροσδιοριστία. Χρησιμοποιώντας τον κανόνα του L'Hospital και ύστερα από αρκετές αλγεβρικές πράξεις, ο Nadarajah (2011) προσδιόρισε την εντροπία του Shannon συναρτήσει της  $\psi(\cdot)$  συνάρτησης, η οποία ορίζεται ως εξής

$$\psi(x) = \frac{d \log \Gamma(x)}{dx} \tag{2.2.23}$$

με  $\Gamma(\cdot)$  να είναι η συνάρτηση Γάμμα, και της σταθεράς  $\delta$  του Euler, η οποία ορίζεται από τη σχέση  $\delta = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \ln(n) \right)$ . Η αριθμητική τιμή της σταθεράς του Euler, με προσέγγιση 50 δεκαδικών ψηφίων, είναι 0.57721566490153286060651209008240243104215933593992.

**Πρόταση 2.2.4 (Nadarajah 2011)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$  με σ.π.π. που δίνεται από τη σχέση (2.1.1). Η εντροπία του Shannon δίνεται από τη σχέση:

$$\begin{aligned}
H_{Sh}(x) = \log C - \delta - \frac{B}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k)}{(3/2)_k} (-B^2)^k \\
- \frac{A}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k)}{(3/2)_k} (-A^2)^k
\end{aligned} \tag{2.2.24}$$

όπου  $C = \sigma D$ ,  $D$  η σταθερά όπως έχει οριστεί στη σχέση (2.1.3),  $A = \frac{a-\mu}{\sigma}$ ,  $B = \frac{b-\mu}{\sigma}$ ,  $\delta$  η σταθερά του Euler και  $\psi(\cdot)$  η συνάρτηση όπως αυτή ορίστηκε στη σχέση (2.2.23).

### Μέτρα Σχήματος

Στην ενότητα αυτή το ενδιαφέρον επικεντρώνεται στον προσδιορισμό τριών δημοφιλών μέτρων του σχήματος μιας κατανομής. Παρότι στη βιβλιογραφία έχουν εισαχθεί διάφορα πληθυσμιακά μέτρα του σχήματος μιας κατανομής εξέχουσα θέση κατέχουν οι συντελεστές λοξότητας και κύρτωσης, ενώ τα τελευταία έτη το ενδιαφέρον επικεντρώνεται και στο μέτρο σχήματος που προτάθηκε από τον Song (2001). Αντικείμενο αυτής της παραγράφου είναι ο προσδιορισμός αυτών όταν  $X \sim TC^{(a,b)}(\mu, \sigma)$ .

Έστω  $X$  μια τ.μ. που ακολουθεί με πεπερασμένες  $k$ -τάξης ροπές, και έστω  $\alpha_k = EX^k$  και  $\beta_k = E(X - EX)^k$ . Τότε, είναι γνωστό, ότι ο συντελεστής λοξότητας δίνεται από τη σχέση:

$$\gamma_1 = \frac{\beta_3}{(\text{Var}X)^{3/2}} = \frac{\alpha_3 - 3\alpha_1\alpha_2 + 2(\alpha_1)^3}{(\alpha_2 - (\alpha_1)^2)^{3/2}},$$

ενώ ο συντελεστής κύρτωσης δίνεται από τη σχέση:

$$\gamma_2 = \frac{\beta_4}{(\text{Var}X)^2} = \frac{\alpha_4 - 4\alpha_1\alpha_3 + 6(\alpha_1)^2\alpha_2 - 3(\alpha_1)^4}{(\alpha_2 - (\alpha_1)^2)^2}.$$

Επιπλέον, γνωρίζουμε ότι αν  $\gamma_1 > 0$  τότε η κατανομή της τ.μ.  $X$  είναι λοξή προς τα δεξιά, ενώ αν  $\gamma_1 < 0$  τότε η κατανομή της τ.μ.  $X$  είναι λοξή προς τα αριστερά. Τέλος αν η κατανομή της τ.μ.  $X$  είναι συμμετρική γύρω από τη μέση της τιμή τότε  $\gamma_1 = 0$ . Επιπρόσθετα, είναι γνωστό ότι μεγάλες τιμές του συντελεστή κύρτωσης δείχνουν ότι η αντίστοιχη κατανομή είναι πιο κυρτή ή λιγότερη επίπεδη γύρω από τη μέση της τιμή. Μικρές τιμές για το συντελεστή κύρτωσης δείχνουν μεγαλύτερη επιπεδότητα της αντίστοιχης κατανομής (βλέπε Σωκράτους (2009)).

Είναι φανερό ότι για τον υπολογισμό των συντελεστών κύρτωσης και λοξότητας μιας τ.μ. που ακολουθεί μια διπλά αποκομμένη κατανομή Cauchy αρκεί να προσδιοριστούν οι ροπές της  $\alpha_k = EX^k$  για  $k = 1, \dots, 4$ . Ο προσδιορισμός αυτός έχει επιτευχθεί στο Θεώρημα 2.2.1 και για την ειδική περίπτωση της διπλά αποκομμένης τυπικής κατανομής Cauchy στο Πρόρισμα 2.2.1. Ανακαλώντας τα αποτελέσματα αυτά έχουμε ότι:

α) αν  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , τότε:

$$a_k = \frac{\sum_{l=0}^k \frac{1}{l+1} \binom{k}{l} \mu^{k-l} \sigma^l \Lambda_l}{D},$$

όπου

$$\Lambda_l = B^{l+1} {}_2F_1\left(1, \frac{l+1}{2}; 1 + \frac{l+1}{2}; -B^2\right) - A^{l+1} {}_2F_1\left(1, \frac{l+1}{2}; 1 + \frac{l+1}{2}; -A^2\right),$$

$$\text{με } A = \frac{a-\mu}{\sigma}, B = \frac{b-\mu}{\sigma} \text{ και } D = \tan^{-1}(B) - \tan^{-1}(A),$$

β) ενώ αν  $X$  μια τ.μ. που ακολουθεί την  $TC^{(a,b)}(0,1)$ , τότε

$$\alpha_1 = \left\{ \log(1+b^2) - \log(1+a^2) \right\} / (2D),$$

$$\alpha_2 = \left\{ \tan^{-1}(a) - \tan^{-1}(b) - a + b \right\} / (D),$$

$$\alpha_3 = \left\{ \log(1+a^2) - \log(1+b^2) - a^2 + b^2 \right\} / (2D),$$

και

$$\alpha_4 = \left\{ 3 \tan^{-1}(b) - 3 \tan^{-1}(a) - a^3 + b^3 + 3a - 3b \right\} / (3D).$$

όπου  $D = \tan^{-1}(b) - \tan^{-1}(a)$ .

Για επιλεγμένες τιμές των παραμέτρων της  $TC^{(a,b)}(\mu, \sigma)$  υπολογίζουμε τον συντελεστή λοξότητας και κύρτωσης. Οι τιμές αυτές παρουσιάζονται στον Πίνακα 3. Από τα αποτελέσματα αυτά προκύπτει ότι η  $X \sim TC^{(0,10)}(\mu, \sigma)$  είναι λοξή δεξιά, η  $X \sim TC^{(-5,0)}(\mu, \sigma)$  είναι λοξή αριστερά, ενώ η  $X \sim TC^{(-5,5)}(\mu, \sigma)$  είναι λοξή δεξιά (αριστερά) για αρνητικές (θετικές) τιμές της παραμέτρου  $\mu$ , ενώ για  $\mu = 0$  είναι συμμετρική.

**Πίνακας 3:** Συντελεστής κύρτωσης και λοξότητας της τυχαίας μεταβλητής  $X$ , με σ.π.π. της μορφής (2.2.1) για τις διάφορες τιμές των παραμέτρων  $\mu, \sigma, a$  και  $b$ .

$(a, b)$	$\sigma$	$\mu$	$\gamma_1$	$\gamma_2$
$a = 0, b = 10$	1	-3	1.0365	3.1393
		-2	1.2644	3.8009
		-1	1.6323	5.1670
		0	2.1089	7.6152
		1	2.0721	8.0168
		2	1.6344	6.5661
		3	1.1066	5.1732
	2	2	1.0459	3.7489
	3		0.7498	2.8399
	4		0.5646	2.4256
10	0.1647		1.8902	
$a = -5, b = 5$	1	-3	1.6344	6.5661
		-2	1.1066	5.1732
		-1	0.5565	4.2771
		0	0	3.9723
		1	-0.5565	4.2771
		2	-1.1066	5.1732
		3	-1.6344	6.5661
	2	2	-0.7243	3.2131
	3		-0.5262	2.5739
	4		-0.3978	2.2763
10	-0.1133		1.8767	
$a = -5, b = 0$	1	-3	0.3683	2.8374
		-2	-0.3683	2.8374
		-1	-1.0459	3.7489
		0	-1.3550	4.2678
		1	-1.0903	3.3079
		2	-0.8297	2.6619
		3	-0.6644	2.3502
	2	2	-0.6795	2.3914
	3		-0.5350	2.1791
	4		-0.4196	2.0431
10	-0.1258		1.8315	

Ο Song (2001) παρατήρησε ότι η κλίση της εντροπίας του Renyi  $H'_R(\gamma) = (d/d\gamma)H_R(\gamma)$  σχετίζεται με το λογάριθμο της πιθανοφάνειας ως εξής  $H'_R(1) = -(1/2)Var[\log(f(X))]$ . Η σχέση αυτή και το γεγονός ότι η τιμή  $-H'_R(1)$  παραμένει σταθερή σε κάθε είδους μετατροπές, ώθησε τον Song να προτείνει το ακόλουθο μέτρο

$$S(f) = -2H'_R(1) = Var[\log(f(X))]$$

για τη μελέτη του σχήματος της κατανομής της τ.μ..

Αναλυτικές εκφράσεις του μέτρου αυτού για διάφορες συνήθεις μονοδιάστατες (βλέπε Song, 2001 και Nadarajah and Zografos, 2003), διδιάστατες και πολυδιάστατες κατανομές (βλέπε Nadarajah and Zografos, 2005 και Zografos and Nadarajah, 2005) είναι διαθέσιμες. Επιπλέον, ο Zografos (2008) διεξάγει μια συγκριτική μελέτη του μέτρου σχήματος του Song (2001) και του συντελεστή κύρτωσης του Mardia στο πλαίσιο των ελλειπτικών κατανομών και προσδιορίζει τις εκφράσεις τους για ειδικές περιπτώσεις αυτών.

Ο Nadarajah (2011) παρακινούμενος από τις παραπάνω μελέτες προσδιορίζει το μέτρο σχήματος  $S(f)$  όταν η τ.μ. ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ . Είναι φανερό ότι για τον υπολογισμό του  $S(f)$  μιας τ.μ. που ακολουθεί μια διπλά αποκομμένη κατανομή αρκεί να προσδιοριστεί αρχικά η κλίση της εντροπίας του Renyi  $H'_R(\gamma) = (d/d\gamma)H_R(\gamma)$ , όπου ο προσδιορισμός της  $H_R(\gamma)$  έγινε στη (2.2.20). Είναι:

$$\begin{aligned}
 H'_R(\gamma) &= (1-\gamma)^{-1} \left\{ -\log C + \frac{1}{H(\gamma)} \left[ B \frac{\partial}{\partial \gamma} {}_2F_1\left(\frac{1}{2}, \gamma; \frac{3}{2}; -B^2\right) \right. \right. \\
 &\quad \left. \left. - A \frac{\partial}{\partial \gamma} {}_2F_1\left(\frac{1}{2}, \gamma; \frac{3}{2}; -A^2\right) \right] \right\} \\
 &\quad + (1-\gamma)^{-2} \{ -\gamma \log C + \log \sigma + \log H(\gamma) \},
 \end{aligned}$$

όπου

$$\begin{aligned}
 \frac{\partial^2}{\partial b^2} {}_2F_1(a, b; c; x) &= \sum_{k=0}^{\infty} \frac{(a)_k [\Gamma(b)\Gamma''(b+k) - \Gamma'(b)(b+k)] x^k}{(c)_k \Gamma^2(b) k!} \\
 &\quad - \psi(b) \sum_{k=0}^{\infty} \frac{(a)_k (b)_k \psi(b+k) x^k}{(c)_k k!} + [\psi^2(b) - \psi'(b)] {}_2F_1(a, b; c; x)
 \end{aligned}$$

και εφαρμόζοντας τον κανόνα L'Hospital προκύπτει ότι:

$$\begin{aligned}
 S(f) &= \delta^2 - \frac{\pi^2}{6} + \frac{B}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k [\Gamma''(1+k) + \delta \Gamma'(1+k)] (-B^2)^k}{(3/2)_k k!} \\
 &\quad - \frac{A}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k [\Gamma''(1+k) + \delta \Gamma'(1+k)] (-A^2)^k}{(3/2)_k k!} \\
 &\quad + \frac{B\delta}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k) (-B^2)^k}{(3/2)_k} - \frac{A\delta}{D} \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k) (-A^2)^k}{(3/2)_k} \\
 &\quad - \frac{1}{D^2} \left[ B \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k) (-B^2)^k}{(3/2)_k} - A \sum_{k=0}^{\infty} \frac{(1/2)_k \psi(1+k) (-A^2)^k}{(3/2)_k} + \delta D \right]^2.
 \end{aligned}$$



### Διατεταγμένες στατιστικές συναρτήσεις

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την  $TC^{(a,b)}(0,1)$ , δηλαδή υποθέτουμε, χωρίς βλάβη της γενικότητας και για απλούστευση των συμβολισμών και των πράξεων, ότι  $\mu = 0$  και  $\sigma = 1$ . Επομένως, υποθέτουμε ότι το τ.δ.  $X_1, \dots, X_n$  προέρχεται από τη διπλά αποκομμένη τυπική κατανομή Cauchy.

Στην παράγραφο που ακολουθεί αρχικά παραθέτουμε τα αποτελέσματα του Nadarajah (2011), που αφορούν τη σ.π.π. του  $r$ -οστού διατεταγμένου στατιστικού,  $r = 1, 2, \dots, n$ . Έπειτα, προσδιορίζονται οι ασυμπτωτικές κατανομές του ελαχίστου και του μέγιστου διατεταγμένου στατιστικού,  $X_{(1)} = \min_i X_i$  και  $X_{(n)} = \max_i X_i$ , αντίστοιχα.

#### Θεώρημα 2.2.5 (Nadarajah, 2011)

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(0,1)$ , με σ.π.π. που δίνεται από την σχέση (2.1.4). Η σ.π.π. του  $r$ -οστού διατεταγμένου στατιστικού,  $Y = X_{(r)}$  δίνεται από την σχέση:

$$\begin{aligned}
 f_Y(y) &= \frac{n!}{(r-1)!(n-r)!D^n} \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} \binom{r-1}{i} \binom{n-r}{j} (-1)^{r-1-i+j} \\
 &\quad \times (\tan^{-1}(a))^{r-1-i} (\tan^{-1}(b))^{n-r-j} \\
 &\quad \times \sum_{k_1=0}^{\infty} \dots \sum_{k_{i+j}=0}^{\infty} \frac{(-1)^{k_1+\dots+k_{i+j}} y^{2(k_1+\dots+k_{i+j})+i+j}}{(2k_1+1)\dots(2k_{i+j}+1)(1+y^2)},
 \end{aligned} \tag{2.2.25}$$

όπου  $a \leq y \leq b$ ,  $D$  η σταθερά όπως αυτή ορίστηκε στη σχέση (2.1.3) και  $r = 1, 2, \dots, n$ .

### Απόδειξη

Είναι γνωστό ότι η κατανομή του  $r$ -οστού διατεταγμένου στατιστικού,  $Y = X_{(r)}$ ,  $r = 1, 2, \dots, n$  από πληθυσμό που ακολουθεί κατανομή με σ.π.π.  $f(\cdot)$  και α.σ.κ.  $F(\cdot)$ , δίνεται από τη σχέση,

$$f_Y(y) = \frac{n!}{(r-1)!(n-r)!} F^{r-1}(y) \{1-F(y)\}^{n-r} f(y). \quad (2.2.26)$$

Με αντικατάσταση των  $f(\cdot)$ ,  $F(\cdot)$  της σχέσης (2.2.26) με τις αντίστοιχες της  $TC^{(a,b)}(0,1)$ , που δίνονται στις σχέσεις (2.1.4) και (2.1.8) αντίστοιχα, προκύπτει ότι:

$$f_Y(y) = \binom{n}{r} \frac{\{\tan^{-1}(y) - \tan^{-1}(a)\}^{r-1} \{\tan^{-1}(b) - \tan^{-1}(y)\}^{n-r}}{D^n (1+y^2)}. \quad (2.2.27)$$

Χρησιμοποιώντας το ανάπτυγμα της διωνυμικής σειράς του Νεύτωνα, η σχέση (2.2.27) γράφεται

$$f_Y(y) = \frac{n!}{(r-1)!(n-r)! D^n} \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} \binom{r-1}{i} \binom{n-r}{j} (-1)^{r-1-i+j} \times (\tan^{-1}(a))^{r-1-i} (\tan^{-1}(b))^{n-r-j} (\tan^{-1}(y))^{i+j} (1+y^2)^{-1}. \quad (2.2.28)$$

Επιπρόσθετα, καθώς ισχύει ότι

$$\tan^{-1}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1},$$

προκύπτει ότι η σχέση (2.2.28) γράφεται ως,

$$\begin{aligned}
f_Y(y) &= \frac{n!}{(r-1)!(n-r)!D^n} \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} \binom{r-1}{i} \binom{n-r}{j} (-1)^{r-1-i+j} \\
&\quad \times (\tan^{-1}(a))^{r-1-i} (\tan^{-1}(b))^{n-r-j} \\
&\quad \times \sum_{k_1=0}^{\infty} \cdots \sum_{k_{i+j}=0}^{\infty} \frac{(-1)^{k_1+\dots+k_{i+j}} y^{2(k_1+\dots+k_{i+j})+i+j}}{(2k_1+1)\cdots(2k_{i+j}+1)(1+y^2)},
\end{aligned}$$

και η απόδειξη ολοκληρώθηκε. ■

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από τη σχέση (2.1.2). Η θεωρία των ακραίων τιμών (extreme value theory) πρωτίστως ασχολείται με την κατανομή του ελαχίστου και μεγίστου διατεταγμένου στατιστικού,  $X_{(1)}$  και  $X_{(n)}$ , αντίστοιχα. Όμως, όπως γίνεται ίσως αντιληπτό από το Θεώρημα 2.2.5, πολλές φορές η σ.π.π. των  $X_{(1)}$  και  $X_{(n)}$  δεν έχει απλή και εύκολα διαχειρίσιμη μορφή. Σε τέτοιες περιπτώσεις το ενδιαφέρον επικεντρώνεται στην εύρεση της ασυμπτωτικής κατανομής, δηλαδή στην εύρεση της κατανομής όταν το μέγεθος του δείγματος  $n$  είναι μεγάλο. Όπως επισημαίνουν οι Mann and Singpurwalla (2006) το σημείο κλειδί που κάνει τις ασυμπτωτικές κατανομές των  $X_{(1)}$  και  $X_{(n)}$  ενδιαφέρουσες είναι ότι για κάποιες σταθερές  $a_n, b_n > 0, c_n$  και  $d_n > 0$ , οι ποσότητες  $(X_{(n)} - a_n)/b_n$  και  $(X_{(1)} - c_n)/d_n$  γίνονται ολοένα και πιο ανεξάρτητες από το μέγεθος του τ.δ.  $n$ .

Ο σκοπός της ασυμπτωτικής θεωρίας των ακραίων τιμών είναι ο καθορισμός των συνθηκών για την ύπαρξη των σταθερών που προαναφέρθηκαν και η εύρεση της ασυμπτωτικής κατανομής. Στην κατεύθυνση αυτή, τα θεμελιώδη

αποτελέσματα εμφανίστηκαν για πρώτη φορά στη βιβλιογραφία από τον Frechet (1927), αλλά τυποποιήθηκαν από τον Gnedenko (1943). Το κυριότερο αποτέλεσμα που προκύπτει είναι ότι, ανεξάρτητα από τη σ.π.π. του πληθυσμού, αν υπάρχει η ασυμπτωτική κατανομή του  $X_{(n)} = \max_i X_i$ , άρα και του  $X_{(1)} = \max_i (-X_i)$ , είναι μία εκ τριών πιθανών.

Στα θεώρηματα που ακολουθούν συνοψίζονται τα αποτελέσματα που αφορούν την ασυμπτωτική κατανομή του ελαχίστου και μεγίστου διατεταγμένου στατιστικού,  $X_{(1)}$  και  $X_{(n)}$ , αντίστοιχα. Σε όσα ακολουθούν  $a(F) = \inf\{x : F(x) > 0\}$  και  $w(F) = \sup\{x : F(x) < 1\}$ .

**Θεώρημα 2.2.6 (Gnedenko, 1947, Galambos, 1978)**

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό με α.σ.κ.  $F(\cdot)$  και  $X_{(n)} = \max_i X_i$  το μέγιστο διατεταγμένο στατιστικό.

α) Έστω  $w(F) = +\infty$ . Υποθέτουμε ότι υπάρχει μία σταθερά  $\tau > 0$  τέτοια ώστε για κάθε  $x > 0$  καθώς το  $t \rightarrow \infty$ ,  $\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\tau}$ . Τότε υπάρχει μια ακολουθία  $b_n > 0$  τέτοια ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(n)} < b_n x] = H^1(x),$$

όπου

$$H^1(x) = \begin{cases} \exp[-x^{-\tau}], & x > 0 \\ 0 & , \text{ αλλού} \end{cases}.$$

Η σταθερά  $b_n$  μπορεί να επιλεγεί να είναι:  $b_n = \inf\{x : 1 - F(x) \leq 1/n\}$ .

β) Έστω  $w(F) < \infty$ . Υποθέτουμε ότι υπάρχει μία σταθερά  $\tau > 0$  τέτοια ώστε για

κάθε  $x > 0$  καθώς το  $t \rightarrow \infty$ ,  $\lim_{t \rightarrow \infty} \frac{1 - F^*(tx)}{1 - F^*(t)} = x^{-\tau}$ , με

$F^*(x) = F(w(F) - 1/x)$ ,  $x > 0$ . Τότε υπάρχουν ακολουθίες  $a_n$  και  $b_n > 0$ ,

τέτοιες ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(n)} < a_n + b_n x] = H^2(x),$$

όπου

$$H^2(x) = \begin{cases} \exp[-(-x)^\tau], & x < 0 \\ 1 & , \text{ αλλιού} \end{cases}.$$

Οι σταθερές  $a_n$  και  $b_n > 0$  μπορούν να επιλεγούν να είναι:  $a_n = w(F)$  και

$b_n = w(F) - \inf\{x : 1 - F(x) \leq 1/n\}$ .

γ) Υποθέτουμε ότι για κάποιο πεπερασμένο  $a$  ισχύει ότι

$$\int_a^{w(F)} (1 - F(y)) dy < +\infty.$$

Επιπλέον, για  $a(F) < t < w(F)$  ορίζουμε την

$$R(t) = (1 - F(t))^{-1} \int_t^{w(F)} (1 - F(y)) dy.$$

Υποθέτουμε ότι για όλους τους πραγματικούς  $x$ , καθώς το  $t \rightarrow w(F)$ , ισχύει ότι:

$$\lim \frac{1 - F(t + xR(t))}{1 - F(t)} = e^{-x}.$$

Τότε υπάρχουν ακολουθίες  $a_n$  και  $b_n > 0$ , τέτοιες ώστε

$$\lim_{n \rightarrow \infty} P[X_{(n)} < a_n + b_n x] = H^3(x),$$

όπου  $H^3(x)$  είναι η α.σ.κ. της Gumbel κατανομής, δηλαδή  $H^3(x) = \exp(-\exp(-x))$ ,  $-\infty < x < \infty$ . Οι σταθερές  $a_n$  και  $b_n$  μπορούν να επιλεγούν να είναι:  $a_n = \inf\{x : 1 - F(x) \leq 1/n\}$  και  $b_n = R(a_n)$ .

Έστω τώρα  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από την σχέση (2.1.2). Εφαρμόζοντας το προηγούμενο θεώρημα μπορεί να προσδιοριστεί η ασυμπτωτική κατανομή του μέγιστου διατεταγμένου στατιστικού.

**Θεώρημα 2.2.7 (Nadarajah, 2011)**

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με α.σ.κ. που δίνεται από την σχέση (2.1.7). Η ασυμπτωτική κατανομή του μεγίστου διατεταγμένου στατιστικού  $X_{(n)}$  αν  $b = \infty$  δίνεται από τη σχέση:

$$\lim_{n \rightarrow \infty} P[X_{(n)} < b_n x] = \exp(-1/x), \tag{2.2.29}$$

όπου  $b_n$  μπορεί να επιλεγεί να είναι:

$$b_n = F_X^{-1}(1-1/n) = \mu + \sigma \cot\left(\frac{\pi}{2n} - \frac{1}{n} \tan^{-1}(A)\right),$$

ενώ αν  $b < \infty$

$$\lim_{n \rightarrow \infty} P\left[X_{(n)} < a_n + b_n x\right] = \exp(x), \quad (2.2.30)$$

όπου  $a_n = w(F) = b$ , και

$$\begin{aligned} b_n &= w(F) - \inf\{x : 1 - F(x) \leq 1/n\} \\ &= b - \mu - \sigma \tan\left(\frac{1}{n} \tan^{-1}(A) - \frac{n-1}{n} \tan^{-1}(B)\right) \end{aligned}$$

$$\text{με } A = \frac{a - \mu}{\sigma} \text{ και } B = \frac{b - \mu}{\sigma}.$$

### Απόδειξη

Όταν  $b = \infty$  εύκολα προκύπτει ότι  $w(F) = \sup\{x : F_X(x) < 1\} = +\infty$ . Επιπλέον, προκύπτει, με αντικατάσταση της  $F_X(\cdot)$  από τη σχέση (2.1.7), και εφαρμόζοντας τον κανόνα L' Hospital, ότι:

$$\lim_{t \rightarrow \infty} \frac{1 - F_X(tx)}{1 - F_X(t)} = \lim_{t \rightarrow \infty} \frac{x \left[ 1 + \left( \frac{t - \mu}{\sigma} \right)^2 \right]}{1 + \left( \frac{tx - \mu}{\sigma} \right)^2} = \lim_{t \rightarrow \infty} \frac{2x \left( \frac{t - \mu}{\sigma} \right) \frac{1}{\sigma}}{2 \left( \frac{tx - \mu}{\sigma} \right) \frac{x}{\sigma}} = \lim_{t \rightarrow \infty} \frac{t - \mu}{tx - \mu} = \frac{1}{x}.$$

Επομένως, σύμφωνα με το Θεώρημα 2.2.6 α), αφού υπάρχει μία σταθερά  $\tau = 1 > 0$  τέτοια ώστε για κάθε  $x > 0$ ,  $\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1}$ , συμπεραίνουμε ότι

υπάρχει μια ακολουθία  $b_n = \inf\{x : 1 - F_X(x) \leq 1/n\} = F_X^{-1}(1 - 1/n)$  τέτοια ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(n)} < b_n x] = \exp[-x^{-1}], x > 0.$$

Λαμβάνοντας υπόψη ότι  $b = \infty$  και τη σχέση (2.1.7) προκύπτει ύστερα από λίγη άλγεβρα ότι:

$$b_n = F_X^{-1}(1 - 1/n)\mu + \sigma \cot\left(\frac{\pi}{2n} - \frac{1}{n} \tan^{-1}(A)\right).$$

Όταν  $b \neq \infty$  εύκολα προκύπτει ότι  $w(F) = \sup\{x : F_X(x) < 1\} = b < \infty$ . Τότε,

$$F^*(x) = F_X(w(F) - 1/x) = F_X(b - 1/x),$$

και

$$\lim_{t \rightarrow \infty} \frac{1 - F^*(tx)}{1 - F^*(t)} = \lim_{t \rightarrow 0} \frac{1 - F_X(b - tx)}{1 - F_X(b - t)} = \lim_{t \rightarrow 0} \frac{x \left[ 1 + \left( \frac{t - \mu}{\sigma} \right)^2 \right]}{1 + \left( \frac{tx - \mu}{\sigma} \right)^2} = x^{-1}.$$

Επομένως, σύμφωνα με το Θεώρημα 2.2.6 β), αφού υπάρχει μία σταθερά

$\tau = 1 > 0$  τέτοια ώστε για κάθε  $x > 0$ ,  $\lim_{t \rightarrow \infty} \frac{1 - F^*(tx)}{1 - F^*(t)} = x^{-1}$ , υπάρχουν

ακολουθίες  $a_n = w(F)$  και  $b_n = w(F) - \inf\{x : 1 - F(x) \leq 1/n\}$ , τέτοιες ώστε καθώς το  $n \rightarrow \infty$



$$\lim_{n \rightarrow \infty} P[X_{(n)} < a_n + b_n x] = \exp[-(-x)^1] = \exp[x].$$

$$\text{Είναι } a_n = b \text{ και } b_n = b - F_x^{-1}\left(\frac{n-1}{n}\right) = b - \mu - \sigma \tan\left(\tan^{-1}\left(\frac{a-\mu}{\sigma}\right) + \frac{n-1}{n} D\right).$$

Το Θεώρημα 2.2.6 μπορεί να διαμορφωθεί ανάλογα και για το ελάχιστο διατεταγμένο στατιστικό καθώς  $X_{(1)} = \max_i(-X_i)$ . Για λόγους πληρότητας δίνεται αρχικά για τη γενική περίπτωση και στη συνέχεια εφαρμόζεται για την περίπτωση τυχαίου δείγματος από τη διπλά αποκομμένη κατανομή Cauchy.

### Θεώρημα 2.2.8 (Gnedenco, 1947, Galambos, 1978)

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό με α.σ.κ.  $F(\cdot)$  και  $X_{(1)} = \min_i X_i$  το ελάχιστο διατεταγμένο στατιστικό.

α) Έστω  $a(F) = -\infty$ . Υποθέτουμε ότι υπάρχει μία σταθερά  $\tau > 0$  τέτοια ώστε για κάθε  $x > 0$  καθώς το  $t \rightarrow -\infty$ ,  $\lim_{t \rightarrow -\infty} \frac{F(tx)}{F(t)} = x^{-\tau}$ . Τότε υπάρχει μια ακολουθία  $d_n > 0$  τέτοια ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(1)} < d_n x] = L^1(x),$$

όπου

$$L^1(x) = \begin{cases} 1 & , x > 0 \\ 1 - \exp[-(-x)^{-\tau}] & , x < 0 \end{cases}.$$

Η σταθερά  $d_n$  μπορεί να επιλεγεί να είναι:  $d_n = \sup\{x : F(x) \leq 1/n\}$ .

β) Έστω  $a(F) < \infty$ . Υποθέτουμε ότι υπάρχει μία σταθερά  $\tau > 0$  τέτοια ώστε για

κάθε  $x > 0$  καθώς το  $t \rightarrow -\infty$ ,  $\lim_{t \rightarrow -\infty} \frac{F^*(tx)}{F^*(t)} = x^{-\tau}$ , με

$F^*(x) = F(a(F) - 1/x)$ ,  $x > 0$ . Τότε υπάρχουν ακολουθίες  $c_n$  και  $d_n > 0$ , τέτοιες ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(1)} < c_n + d_n x] = L^2(x),$$

όπου

$$L^2(x) = \begin{cases} 0 & , x < 0 \\ 1 - \exp[-x^\tau] & , x > 0 \end{cases}.$$

Οι σταθερές  $c_n$  και  $d_n > 0$  μπορούν να επιλεγούν να είναι:  $c_n = a(F)$  και  $d_n = \sup\{x : F(x) \leq 1/n\} - a(F)$ .

γ) Υποθέτουμε ότι για κάποιο πεπερασμένο  $a$  ισχύει ότι

$$\int_{a(F)}^a F(y) dy < +\infty.$$

Επιπλέον, για  $a(F) < t$  ορίζουμε την

$$r(t) = (F(t))^{-1} \int_{a(F)}^t F(y) dy.$$

Υποθέτουμε ότι για όλους τους πραγματικούς  $x$ , καθώς το  $t \rightarrow a(F)$ , ισχύει ότι:

$$\lim_{t \rightarrow a(F)} \frac{F(t + xr(t))}{F(t)} = e^x.$$

Τότε υπάρχουν ακολουθίες  $c_n$  και  $d_n > 0$ , τέτοιες ώστε

$$\lim_{n \rightarrow \infty} P[X_{(1)} < c_n + d_n x] = L^3(x),$$

όπου  $L^3(x) = 1 - \exp(-\exp(x))$ ,  $-\infty < x < \infty$ . Οι σταθερές  $c_n$  και  $d_n$  μπορούν να επιλεγούν να είναι:  $c_n = \sup\{x : F(x) \leq 1/n\}$  και  $d_n = r(c_n)$ .

Έστω τώρα  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από την σχέση (2.1.2). Εφαρμόζοντας το προηγούμενο θεώρημα μπορεί να προσδιοριστεί η ασυμπτωτική κατανομή του ελαχίστου διατεταγμένου στατιστικού.

### Θεώρημα 2.2.9 (Nadarajah, 2011)

Έστω  $X_1, \dots, X_n$  τ.δ. από έναν πληθυσμό που ακολουθεί την  $TC^{(a,b)}(\mu, \sigma)$  με α.σ.κ. που δίνεται από την σχέση (2.1.7). Η ασυμπτωτική κατανομή του ελαχίστου διατεταγμένου στατιστικού  $X_{(1)} = \min_i X_i$  αν  $a = -\infty$  δίνεται από τη σχέση:

$$\lim_{n \rightarrow \infty} P[X_{(1)} < d_n x] = 1 - \exp(1/x), \quad (2.2.31)$$

όπου  $d_n = F_X^{-1}(1/n) = \mu - \sigma \cot\left(\frac{\pi}{2n} + \frac{1}{n} \tan^{-1}(B)\right)$ , ενώ αν  $a > -\infty$

$$\lim_{n \rightarrow \infty} P \left[ X_{(1)} < c_n + d_n x \right] = 1 - \exp(-x), \quad (2.2.32)$$

όπου  $c_n = a(F) = a$ , και

$$d_n = F_X^{-1}(1/n) - a = -a + \mu + \sigma \tan \left( \frac{1}{n} \arctan(B) + \frac{n-1}{n} \arctan(A) \right)$$

με  $A = \frac{a - \mu}{\sigma}$  και  $B = \frac{b - \mu}{\sigma}$ .

**Απόδειξη**

Όταν  $a = -\infty$  εύκολα προκύπτει ότι  $a(F) = \inf\{x : F_X(x) > 0\} = -\infty$ . Επιπλέον, προκύπτει, με αντικατάσταση της  $F_X(\cdot)$  από τη σχέση (2.1.7) και εφαρμόζοντας τον κανόνα L' Hospital, ότι:

$$\lim_{t \rightarrow -\infty} \frac{F_X(tx)}{F_X(t)} = \lim_{t \rightarrow -\infty} \frac{x \left[ 1 + \left( \frac{t - \mu}{\sigma} \right)^2 \right]}{1 + \left( \frac{tx - \mu}{\sigma} \right)^2} = \lim_{t \rightarrow -\infty} \frac{2x \left( \frac{t - \mu}{\sigma} \right) \frac{1}{\sigma}}{2 \left( \frac{tx - \mu}{\sigma} \right) \frac{x}{\sigma}} = \lim_{t \rightarrow -\infty} \frac{t - \mu}{tx - \mu} = \frac{1}{x}.$$

Επομένως, σύμφωνα με το Θεώρημα 2.2.8 α), αφού υπάρχει μία σταθερά  $\tau = 1 > 0$  τέτοια ώστε για κάθε  $x > 0$ ,  $\lim_{t \rightarrow -\infty} \frac{F(tx)}{F(t)} = x^{-1}$ , συμπεραίνουμε ότι υπάρχει μια ακολουθία  $d_n = \sup\{x : F_X(x) \leq 1/n\} = F_X^{-1}(1/n)$  τέτοια ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P \left[ X_{(n)} < d_n x \right] = 1 - \exp \left[ x^{-1} \right].$$

$$\text{Επομένως είναι } d_n = F_X^{-1}(1/n) = \mu - \sigma \cot\left(\frac{\pi}{2n} + \frac{1}{n} \tan^{-1}(B)\right).$$

Όταν  $a \neq -\infty$  εύκολα προκύπτει ότι  $a(F) = \inf\{x: F_X(x) > 0\} = a < \infty$ . Τότε,

$$F^*(x) = F_X(a(F) - 1/x) = F_X(a - 1/x),$$

και

$$\lim_{t \rightarrow \infty} \frac{F^*(tx)}{F^*(t)} = \lim_{t \rightarrow 0} \frac{F_X(a+tx)}{F_X(a+t)} = \lim_{t \rightarrow 0} \frac{x \left[ 1 + \left( \frac{t-\mu}{\sigma} \right)^2 \right]}{1 + \left( \frac{tx-\mu}{\sigma} \right)^2} = x^{-1}.$$

Επομένως, σύμφωνα με το Θεώρημα 2.2.8 β), αφού υπάρχει μία σταθερά

$\tau = 1 > 0$  τέτοια ώστε για κάθε  $x > 0$ ,  $\lim_{t \rightarrow \infty} \frac{F^*(tx)}{F^*(t)} = x^{-1}$ , υπάρχουν ακολουθίες

$c_n = a(F)$  και  $d_n = \sup\{x: F(x) \leq 1/n\} - a(F)$ , τέτοιες ώστε καθώς το  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P[X_{(1)} < c_n + d_n x] = 1 - \exp[-x].$$

Επομένως

$$c_n = a(F) = a$$

και

$$\begin{aligned} d_n &= F_X^{-1}(1/n) - a = -a + \mu + \sigma \tan\left(\tan^{-1}\left(\frac{a-\mu}{\sigma}\right) + (1/n)D\right) \\ &= -a + \mu + \sigma \tan\left(\frac{1}{n} \arctan(B) + \frac{n-1}{n} \arctan(A)\right) \end{aligned}$$

$$\text{με } A = \frac{a - \mu}{\sigma} \text{ και } B = \frac{b - \mu}{\sigma} .$$

■



## ΚΕΦΑΛΑΙΟ 3

### Εκτίμηση παραμέτρων

Στο προηγούμενο κεφάλαιο αναφέρθηκαν οι πιο σημαντικές ιδιότητες της μονοδιάστατης διπλά αποκομμένης κατανομής Cauchy. Αντικείμενο αυτού του κεφαλαίου αρχικά είναι η εκτίμηση των παραμέτρων με τη μέθοδο της μέγιστης πιθανοφάνειας. Ειδικότερα, θα παρουσιαστούν τα αποτελέσματα των Daliya *et al.* (2001) που αφορούν στην ύπαρξη και τον προσδιορισμό των Ε.Μ.Π. της διπλά αποκομμένης Κατανομής Cauchy.

Προκύπτει ότι η ύπαρξη πεπερασμένου εκτιμητή της παραμέτρου  $\sigma$  της διπλά αποκομμένης κατανομής Cauchy εξαρτάται από το αν πληρείται ή όχι μία συνθήκη. Για το λόγο αυτό, παρακινούμενοι από τους Daliya *et al.* (2001), διεξάγουμε μία μελέτη προσομοίωσης για την εκτίμηση της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή. Επιπρόσθετα, με χρήση προσομοιωμένων δεδομένων, αξιολογείται η απόδοση των Ε.Μ.Π. των άγνωστων παραμέτρων.

Τέλος, το κεφάλαιο θα ολοκληρωθεί με ένα αριθμητικό παράδειγμα, όπου προσαρμόζεται το μοντέλο της διπλά αποκομμένης κατανομής Cauchy σε ένα γνωστό σύνολο δεδομένων, και με την παρουσίαση των αποτελεσμάτων που δόθηκαν από τον Nadarajah (2011) για την εκτίμηση των αγνώστων παραμέτρων με τη μέθοδο των ροπών.

#### 3.1 Μέθοδος Μέγιστης Πιθανοφάνειας

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό που ακολουθεί την διπλά αποκομμένη κατανομή Cauchy  $TC^{(a,b)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από τη σχέση:



$$f_X(x) = \frac{\sigma}{D} \frac{1}{[\sigma^2 + (x - \mu)^2]},$$

όπου

$$D = \tan^{-1}\left(\frac{b - \mu}{\sigma}\right) - \tan^{-1}\left(\frac{a - \mu}{\sigma}\right),$$

$$a \leq x \leq b, \sigma > 0.$$

Σε όσα ακολουθούν υποθέτουμε ότι  $\mu \in [a, b]$ , καθώς, σύμφωνα με τους Daliya et al. (2001) η περίπτωση που  $\mu \notin [a, b]$  δεν έχει πρακτική αξία. Εύκολα προκύπτει ότι η συνάρτηση πιθανοφάνειας είναι η:

$$L_n(a, b, \mu, \sigma) = \left(\frac{\sigma}{D}\right)^n \frac{1}{\prod_{i=1}^n [\sigma^2 + (x_i - \mu)^2]}, a \leq X_{(1)} \leq X_{(n)} \leq b, \quad (3.1.1)$$

όπου  $X_{(1)} = \min_i X_i$  και  $X_{(n)} = \max_i X_i$  είναι το ελάχιστο και το μέγιστο διατεταγμένο στατιστικό, αντίστοιχα, και  $D = \tan^{-1}\left(\frac{b - \mu}{\sigma}\right) - \tan^{-1}\left(\frac{a - \mu}{\sigma}\right)$ .

Για την εύρεση των Ε.Μ.Π. απαιτείται η μεγιστοποίηση της συνάρτησης πιθανοφάνειας ως προς τις άγνωστες παραμέτρους. Ο συνηθισμένος τρόπος αντιμετώπισης του παραπάνω θέματος είναι η παραγωγή της συνάρτησης πιθανοφάνειας ως προς τις τέσσερις παραμέτρους  $a, b, \mu$  και  $\sigma$  και η επίλυση του συστήματος των εξισώσεων που προκύπτουν θέτοντας τις μερικές παραγώγους ίσες με το μηδέν. Όμως αυτό δε σημαίνει ότι πάντοτε θα οδηγούμαστε στην εύρεση ολικού μεγίστου, καθώς μπορεί να έχουμε σημείο είτε τοπικού μεγίστου ή ελαχίστου είτε ολικού μεγίστου ή ελαχίστου. Στις επόμενες ενότητες θα παρουσιαστεί η μελέτη των Daliya et al. (2001) ως προς την ύπαρξη των Ε.Μ.Π.

Στο σημείο αυτό είναι απαραίτητο να αναφέρουμε ότι ο Nadarajah (2011) παραθέτει το σύστημα των εξισώσεων, από την επίλυση των οποίων με κάποια αριθμητική μέθοδο, όπως η Newton-Raphson, προκύπτουν οι Ε.Μ.Π. των παραμέτρων στην περίπτωση που τα σημεία αποκοπής είναι είτε άγνωστα είτε γνωστά, χωρίς όμως να εξασφαλίζει και να μελετά την ύπαρξή τους.

Σε όσα ακολουθούν όταν θα διαπραγματευόμαστε τη μεγιστοποίηση της συνάρτησης πιθανοφάνειας ως προς την παράμετρο π.χ.  $\mu$  θα γράφουμε για χάριν συντομίας  $L_n(\mu)$ .

### Εκτιμητές μέγιστης πιθανοφάνειας των σημείων αποκοπής $a$ και $b$

Η μεγιστοποίηση της (3.1.1) ως προς  $a$  και  $b$  ανάγεται σε ένα πρόβλημα ελαχιστοποίησης της συνάρτησης  $D = \tan^{-1}\left(\frac{b-\mu}{\sigma}\right) - \tan^{-1}\left(\frac{a-\mu}{\sigma}\right)$  ως προς  $a$  και  $b$ . Λαμβάνοντας υπόψη ότι η αντίστροφη συνάρτηση της εφαπτομένης,  $\tan^{-1}(\cdot)$  είναι αυστηρά αύξουσα περιττή συνάρτηση προκύπτει εύκολα ότι η συνάρτηση  $D$  ως προς  $a$  είναι φθίνουσα, ενώ ως προς  $b$  η συνάρτηση  $D$  είναι αύξουσα. Επομένως, η  $D$  ελαχιστοποιείται στο άνω άκρο του πεδίου ορισμού της παραμέτρου  $a$  που είναι το διατεταγμένο στατιστικό  $X_{(1)}$  και στο κάτω άκρο του πεδίου ορισμού της παραμέτρου  $b$  που είναι το διατεταγμένο στατιστικό  $X_{(n)}$ . Επομένως οι Ε.Μ.Π. των σημείων αποκοπής  $a$  και  $b$  είναι τα διατεταγμένα στατιστικά  $X_{(1)}$  και  $X_{(n)}$  αντίστοιχα.

Καθώς το πεδίο ορισμού της συνάρτησης πιθανοφάνειας ως προς τις παραμέτρους αποκοπής  $a$  και  $b$  είναι  $a \leq X_{(1)} \leq X_{(n)} \leq b$ , άμεσα προκύπτει ότι οι Ε.Μ.Π. των σημείων αποκοπής δεν είναι αμερόληπτοι. Στη συνέχεια,

ακολουθώντας την τεχνική που προτάθηκε από τους Robson and Whitlock (1964), θα υπολογιστεί η μεροληψία του  $X_{(n)}$ , του Ε.Μ.Π. της παραμέτρου  $b$ , ενώ με παρόμοιο τρόπο υπολογίζεται η μεροληψία του  $X_{(1)}$ .

Είναι γνωστό (βλέπε σχέση 1.1.5) ότι η α.σ.κ. της τυχαίας μεταβλητής  $X$  δίνεται από τη σχέση:

$$F_X(x) = \frac{G(x) - G(a)}{G(b) - G(a)}, \text{ για } a \leq x \leq b,$$

όπου  $G(\cdot)$  είναι η α.σ.κ. της κατανομής Cauchy με παραμέτρους  $\mu$  και  $\sigma$ . Έστω ο ακόλουθος μετασχηματισμός:

$$Y_n = \frac{G(X_{(n)}) - G(a)}{G(b) - G(a)}, \quad (3.1.2)$$

ή ισοδύναμα

$$X_{(n)} = G^{-1} \left[ Y_n (G(b) - G(a)) + G(a) \right] = H(Y_n), \quad (3.1.3)$$

όπου, λαμβάνοντας υπόψη την (1.2.3),  $H(\cdot) = F_X^{-1}(\cdot)$ . Το ανάπτυγμα Taylor της παραπάνω σχέσης ως προς το σημείο 1 δίνει:

$$X_{(n)} = H(1) + (Y_n - 1)H'(1) + \frac{1}{2!}(Y_n - 1)^2 H''(1) + \frac{1}{3!}(Y_n - 1)^3 H'''(1) + \dots \quad (3.1.4)$$

Ύστερα από λίγες πράξεις έχουμε μεταξύ άλλων ότι:

$$H(1) = b,$$

$$H'(1) = \frac{G(b) - G(a)}{G'(b)},$$

και

$$H''(1) = -\frac{G''(b)}{G'(b)} \left[ \frac{G(b) - G(a)}{G'(b)} \right]^2.$$

Επιπρόσθετα, λαμβάνοντας υπόψη ότι  $Y_n = F_X(X_{(n)})$  προκύπτει ότι η κατανομή του  $1 - Y_n$  είναι αυτή του ελάχιστου διατεταγμένου στατιστικού, έστω  $U_{(1)}$ , μιας ομοιόμορφης κατανομής στο διάστημα  $[0,1]$ . Έτσι προκύπτει ότι:

$$E(Y_n - 1)^k = (-1)^k E U_{(1)}^k = \frac{(-1)^k k!}{(n+1) \cdot \dots \cdot (n+k)} = (-1)^k k! n^{(-k)}, \quad (3.1.5)$$

$$\text{όπου } n^{(-k)} = \frac{1}{(n+1) \cdot \dots \cdot (n+k)}.$$

Συνδυάζοντας τα παραπάνω αποτελέσματα προκύπτει ότι:

$$E X_{(n)} = b - n^{(-1)} H'(1) + n^{(-2)} H''(1) - n^{(-3)} H'''(1) + \dots \quad (3.1.6)$$

γεγονός που αποδεικνύει ότι ο Ε.Μ.Π. της παραμέτρου  $b$  είναι μεροληπτικός. Επιπλέον, από τη σχέση (3.1.6) προσδιορίζεται και το ποσό μεροληψίας.

Οι Daliya et al. (2001), παρακινούμενοι από τους Robson και Whitlock (1964), θέλησαν να περιορίσουν τη μεροληψία των  $X_{(1)}$  και  $X_{(n)}$ . Για να επιτευχθεί ο στόχος αυτός τροποποίησαν τη μεθοδολογία που έχει προταθεί στη βιβλιογραφία από τον Quenouille (1964). Σύμφωνα με τη μεθοδολογία του Quenouille (1964), αν  $T_n$  είναι ένας εκτιμητής της άγνωστης παραμέτρου  $\theta$  τέτοιος ώστε:

$$E T_n = \theta + \frac{a_1}{n} + \frac{a_2}{n^2} + \frac{a_3}{n^3} + \dots$$

δηλαδή είναι μεροληπτικός τάξεως τουλάχιστον  $1/n$ , τότε ο εκτιμητής  $T_n^*$  που προσδιορίζεται από τη σχέση:  $T_n^* = nT_n - (n-1)T_{n-1}$ , είναι μεροληπτικός της τάξης του  $1/n^2$  μόνο, καθώς  $ET_n^* = \theta - \frac{a_2}{n^2} - \frac{a_2 + a_3}{n^3} - \dots$

Προσαρμόζοντας κατάλληλα την παραπάνω μέθοδο οι Daliya et al. (2001) προτείνουν τους ακόλουθους εκτιμητές, οι οποίοι περιορίζουν την αμεροληψία της τάξης  $1/n$  στους Ε.Μ.Π. των  $a$  και  $b$ ,

$$\tilde{a} = \left(\frac{2n-1}{n}\right)X_{(1)} - \left(\frac{n-1}{n}\right)X_{(2)} \quad (3.1.7)$$

και

$$\tilde{b} = \left(\frac{2n-1}{n}\right)X_{(n)} - \left(\frac{n-1}{n}\right)X_{(n-1)}, \quad (3.1.8)$$

όπου  $X_{(i)}$  το  $i$ -οστό διατεταγμένο στατιστικό.

**Παρατήρηση 3.1.1** Προφανώς η παραπάνω μεθοδολογία θα μπορούσε να εφαρμοστεί κατάλληλα για να περιορίσει και τη μεροληψία της τάξης του  $1/n^2$ , αλλά όπως επισημαίνεται από τον David (1981), κάτι τέτοιο δεν έχει μεγάλη αξία, καθώς μια τέτοια βελτίωση είναι πολύ μικρή.

### Εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου θέσης $\mu$

Στην ενότητα αυτή παρατίθενται τα αποτελέσματα των Daliya et al. (2001) που αφορούν στην ύπαρξη και τον προσδιορισμό του Ε.Μ.Π. της παραμέτρου θέσης  $\mu$ .

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό που ακολουθεί τη διπλά αποκομμένη κατανομή Cauchy,  $TC^{(a,b)}(\mu, \sigma)$ , με συνάρτηση πιθανοφάνειας που δίνεται στη σχέση (3.1.1). Για να μελετήσουν οι Daliya *et al.* (2001) το πρόβλημα ύπαρξης των Ε.Μ.Π. των παραμέτρων  $\mu$  και  $\sigma$ , θεωρούν ότι  $\hat{a} = -1$  και  $\hat{b} = 1$ . Η θεώρηση αυτή είναι εφικτή μετασχηματίζοντας τα αρχικά δεδομένα  $X_1, \dots, X_n$  στα

$$\tilde{X}_i = \frac{2X_i - (X_{(n)} + X_{(1)})}{(X_{(n)} - X_{(1)})}, i = 1, \dots, n, \quad (3.1.9)$$

για τα οποία προφανώς ισχύει ότι οι Ε.Μ.Π. των σημείων αποκοπής είναι  $\tilde{X}_{(1)} = -1$  και  $\tilde{X}_{(n)} = 1$ . Επομένως, αν  $\tilde{\mu}$  και  $\tilde{\sigma}$  είναι οι Ε.Μ.Π. των παραμέτρων  $\mu$  και  $\sigma$  που προκύπτουν με τα μετασχηματισμένα δεδομένα  $\tilde{X}_i$  της σχέσης (3.1.9), οι Ε.Μ.Π. των παραμέτρων  $\mu$  και  $\sigma$  με τα αρχικά δεδομένα δίνονται από τις σχέσεις:

$$\hat{\mu} = \frac{(X_{(n)} - X_{(1)})\tilde{\mu} + (X_{(n)} + X_{(1)})}{2}, \quad (3.1.10)$$

και

$$\hat{\sigma} = \frac{(X_{(n)} - X_{(1)})}{2} \tilde{\sigma}, \quad (3.1.11)$$

αντίστοιχα. Ανακεφαλαιώνοντας σε όσα ακολουθούν υποθέτουμε, χωρίς βλάβη της γενικότητας, ότι  $\hat{a} = -1$  και  $\hat{b} = 1$ , επομένως  $X_{(1)} = -1$  και  $X_{(n)} = 1$ .

Η μεγιστοποίηση της (3.1.1) ως προς  $\mu$ , ανάγεται σε ελαχιστοποίηση της συνάρτησης  $H_n(\cdot)$  ως προς  $\mu$ , με

$$H_n(\mu) = D^n \prod_{i=1}^n \left[ (x_i - \mu)^2 + \sigma^2 \right], \quad -1 \leq X_{(1)} \leq X_{(n)} \leq 1, \quad (3.1.12)$$

όπου, τώρα έχοντας υποθέσει ότι  $\hat{a} = -1$  και  $\hat{b} = 1$ , και λαμβάνοντας υπόψη τη σχέση (2.1.3), προκύπτει ότι:

$$D = \tan^{-1} \left( \frac{1-\mu}{\sigma} \right) + \tan^{-1} \left( \frac{1+\mu}{\sigma} \right). \quad (3.1.13)$$

Καθώς η  $H_n(\cdot)$  είναι μια συνεχής συνάρτηση ορισμένη στο συμπαγές διάστημα  $[-1,1]$  συμπεραίνουμε ότι η  $H_n(\cdot)$  έχει ολικό ελάχιστο. Επομένως, υπάρχει ο Ε.Μ.Π. της παραμέτρου θέσης  $\mu$  με πιθανότητα 1. Ωστόσο, η παράγωγος του λογάριθμου της συνάρτησης μέγιστης πιθανοφάνειας  $L_n(\mu)$ , άρα και της συνάρτησης  $H_n(\mu)$ , μπορεί να έχει πολλές ρίζες. Γίνεται κατανοητό ότι θα ήταν επιθυμητός ο προσδιορισμός των συνθηκών εκείνων για τις τιμές ή καλύτερα το εύρος των τιμών των παραμέτρων  $\mu$  και  $\sigma$  που θα εξασφάλιζε την ύπαρξη μοναδικού ολικού ελαχίστου ως προς  $\mu$  της συνάρτησης  $H_n(\mu)$  της σχέσης (3.1.12). Στη βιβλιογραφία οι συνθήκες αυτές, όσο είμαστε σε θέση να γνωρίζουμε, δεν έχουν προσδιοριστεί θεωρητικά. Παρόλα αυτά οι Daliya *et al.* (2001) ισχυρίζονται ότι αριθμητικά αποτελέσματα προτείνουν, συνιστούν ότι αν μόνο 2% της πλήρους κατανομής Cauchy αποκόπτεται, πολλαπλά ελάχιστα είναι πολύ σπάνια και ότι όσο το ποσοστό της αποκοπής αυξάνεται, μια τέτοια συμπεριφορά καθίσταται σχεδόν ανύπαρκτη.

### Εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου κλίμακας

Στην ενότητα αυτή παρατίθενται τα αποτελέσματα των Daliya *et al.* (2001) για την ύπαρξη και εύρεση του Ε.Μ.Π. της παραμέτρου κλίμακας  $\sigma$ . Το συμπέρασμα της έρευνας των Daliya *et al.* (2001), το οποίο διατυπώνεται στο

θεώρημα που ακολουθεί είναι ότι ο Ε.Μ.Π. αποτυγχάνει να υφίσταται με θετική πιθανότητα. Ειδικότερα, οι Daliya *et al.* (2001) προσδιορίζουν, αρχικά, μια επαρκή συνθήκη έτσι ώστε ο Ε.Μ.Π. να υφίσταται. Σε όσα ακολουθούν με παρόμοιο σκεπτικό όπως πρωτύτερα υποθέτουμε ότι  $\hat{a} = -1$  και  $\hat{b} = 1$ . Η θεώρηση αυτή είναι εφικτή χρησιμοποιώντας το μετασχηματισμό της σχέσης (3.1.9).

### Θεώρημα 3.1.1 (Daliya *et al.*, 2001)

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό που ακολουθεί την  $TC^{(-1,1)}(\mu, \sigma)$  με σ.π.π. που δίνεται από τη σχέση (2.1.2). Η ικανή συνθήκη για την ύπαρξη του Ε.Μ.Π. της παραμέτρου  $\sigma$  είναι η εξής:

$$\frac{\sum_{i=1}^n (x_i - \mu)^2}{n} < \left( \mu^2 + \frac{1}{3} \right). \quad (3.1.14)$$

### Απόδειξη

Η συνάρτηση πιθανοφάνειας είναι

$$\begin{aligned} L_n(\sigma) &= \left( \frac{\sigma}{D} \right)^n \frac{1}{\prod_{i=1}^n \left( (x_i - \mu)^2 + \sigma^2 \right)} \\ &= \left( \frac{\sigma}{D} \right)^n \frac{1}{\prod_{i=1}^n (w_i + \sigma^2)}, \quad -1 \leq X_{(1)} \leq X_{(n)} \leq 1, \end{aligned} \quad (3.1.15)$$

όπου  $w_i = (x_i - \mu)^2$  και  $D = \tan^{-1} \left( \frac{1 - \mu}{\sigma} \right) + \tan^{-1} \left( \frac{1 + \mu}{\sigma} \right)$ .

Έστω ότι ισχύει η συνθήκη που δίνεται στη σχέση (3.1.14), δηλαδή



$$\frac{\sum_{i=1}^n w_i}{n} = \bar{w} < (\mu^2 + 1/3).$$

Θα αποδείξουμε ότι η  $L_n(\sigma)$  λαμβάνει μέγιστο ως προς  $\sigma$  σε ένα πεπερασμένο σημείο  $\sigma$ . Έστω η συνάρτηση

$$\varphi_1(\sigma) = \sigma^2 \left( \frac{2}{D\sigma} - 1 \right), \quad (3.1.16)$$

όπου

$$D = \tan^{-1} \left( \frac{1-\mu}{\sigma} \right) + \tan^{-1} \left( \frac{1+\mu}{\sigma} \right). \quad (3.1.17)$$

Ισχύει ότι

$$\lim_{\sigma \rightarrow \infty} \sigma D = \lim_{\sigma \rightarrow \infty} \sigma \left[ \tan^{-1} \left( \frac{1-\mu}{\sigma} \right) + \tan^{-1} \left( \frac{1+\mu}{\sigma} \right) \right],$$

ή ισοδύναμα θέτοντας  $u = 1/\sigma$ ,

$$\lim_{\sigma \rightarrow \infty} \sigma D = \lim_{u \rightarrow 0} \left[ \frac{\tan^{-1} [(1-\mu)u] + \tan^{-1} [(1+\mu)u]}{u} \right],$$

και εφαρμόζοντας τον κανόνα D' Hospital προκύπτει ότι:

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} \sigma D &= \lim_{u \rightarrow 0} \left[ \frac{1-\mu}{1+(1-\mu)^2 u^2} + \frac{1+\mu}{1+(1+\mu)^2 u^2} \right] \\ &= (1-\mu) + (1+\mu) = 2. \end{aligned}$$

Άρα,

$$\lim_{\sigma \rightarrow \infty} D \sigma = 2. \quad (3.1.18)$$

Επομένως, λαμβάνοντας υπόψη την (3.1.18), έχουμε ότι για το

$$\lim_{\sigma \rightarrow \infty} \varphi_1(\sigma) = \lim_{\sigma \rightarrow \infty} \left[ \frac{2(D\sigma)^{-1} - 1}{1/\sigma^2} \right],$$

προκύπτει απροσδιόριστη μορφή και

εφαρμόζοντας τον κανόνα D' Hospital:

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} \varphi_1(\sigma) &= \lim_{\sigma \rightarrow \infty} \left[ \frac{-2D^{-2}\sigma^{-1} \frac{\partial D}{\partial \sigma} - 2D^{-1}\sigma^{-2}}{-2/\sigma^3} \right] \\ &= \lim_{\sigma \rightarrow \infty} \left[ \frac{1}{(D\sigma)^2} \frac{D + \sigma \frac{\partial D}{\partial \sigma}}{1/\sigma^3} \right]. \end{aligned} \quad (3.1.19)$$

Κάνοντας χρήση του μετασχηματισμού  $u = 1/\sigma$ , προκύπτει ότι το

$$\lim_{\sigma \rightarrow \infty} \frac{[D + \sigma(\partial D/\partial \sigma)]}{1/\sigma^3},$$

είναι

$$\lim_{u \rightarrow 0} \left\{ \frac{1}{u^3} \left[ \tan^{-1}[(1-\mu)u] + \tan^{-1}[(1+\mu)u] - \left[ \frac{u(1-\mu)}{1+(1-\mu)^2 u^2} + \frac{u(1+\mu)}{1+(1+\mu)^2 u^2} \right] \right] \right\}.$$

Εφαρμόζοντας τώρα τον κανόνα L'Hospital και έπειτα από αλγεβρικές πράξεις προκύπτει ότι:

$$\lim_{\sigma \rightarrow \infty} \frac{[D + \sigma(\partial D/\partial \sigma)]}{1/\sigma^3} = \lim_{u \rightarrow 0} \frac{1}{3} \left[ \frac{2(1-\mu)^3}{[1+(1-\mu)^2 u^2]^2} + \frac{2(1+\mu)^3}{[1+(1+\mu)^2 u^2]^2} \right],$$

και επομένως

$$\lim_{\sigma \rightarrow \infty} \frac{[D + \sigma(\partial D / \partial \sigma)]}{1/\sigma^3} = \frac{4}{3}(1 + 3\mu^2). \quad (3.1.20)$$

Από τις σχέσεις (3.1.18), (3.1.19) και (3.1.20) συνεπάγεται ότι

$$\lim_{\sigma \rightarrow \infty} \varphi_1(\sigma) = \mu^2 + \frac{1}{3}. \quad (3.1.21)$$

Επομένως, καθώς από τη σχέση (3.1.14) ισχύει ότι  $\bar{w} < (\mu^2 + 1/3)$  και

$\lim_{\sigma \rightarrow \infty} \varphi_1(\sigma) = \mu^2 + \frac{1}{3}$ , θα υπάρξει  $\sigma$  τέτοιο ώστε,  $\bar{w} < \varphi_1(\sigma)$ , που σημαίνει ότι

$$(\bar{w}/\sigma^2 + 1)^n < (2/\sigma D)^n \quad (3.1.22)$$

Από τη σχέση (3.1.15) έχουμε ότι:

$$L_n(\sigma) = \left(\frac{\sigma}{D}\right)^n \frac{1}{\prod_{i=1}^n (w_i + \sigma^2)}.$$

Σύμφωνα με την ανισότητα του αριθμητικού γεωμετρικού μέσου είναι γνωστό ότι

αν  $a = (a_1, \dots, a_n)^t$  και  $b = (b_1, \dots, b_n)^t$ , όπου  $a_i \geq 0, b_i > 0$  και  $\sum_{i=1}^n b_i = 1$ , τότε

$\prod_{i=1}^n a_i^{b_i} \leq \sum_{i=1}^n a_i b_i$ . Λαμβάνοντας υπόψη τη σχέση (3.1.22) και με ειδική εφαρμογή

της ανισότητας του αριθμητικού γεωμετρικού μέσου, για  $a_i = \frac{w_i}{\sigma^2} + 1, b_i = 1/n$ ,

προκύπτει ότι:

$$\prod_{i=1}^n \left[ \left( \frac{w_i}{\sigma^2} + 1 \right)^{\frac{1}{n}} \right] \leq \left[ \sum_{i=1}^n \left( \frac{w_i}{\sigma^2} + 1 \right) \frac{1}{n} \right]^n = \left( \frac{\bar{w}}{\sigma^2} + 1 \right)^n < \left( \frac{2}{\sigma D} \right)^n \quad (3.1.23)$$

ή ισοδύναμα

$$\prod_{i=1}^n (w_i + \sigma^2) \leq \left( \frac{2\sigma}{D} \right)^n. \quad (3.1.24)$$

Επομένως

$$\frac{1}{\prod_{i=1}^n (w_i + \sigma^2)} \geq \left( \frac{D}{2\sigma} \right)^n \Rightarrow L_n(\sigma) \geq \frac{1}{2^n}.$$

Επιπλέον, η συνάρτηση  $L_n(\sigma)$  είναι συνεχής και εύκολα αποδεικνύονται, λαμβάνοντας υπόψη τη σχέση (3.1.18), ότι

$$\lim_{\sigma \rightarrow \infty} \sigma^2 \frac{\partial D}{\partial \sigma} = -2,$$

$$\lim_{\sigma \rightarrow 0} L_n(\sigma) = \prod_{i=1}^n \lim_{\sigma \rightarrow 0} \frac{\sigma}{D[\sigma^2 + (x_i - \mu)^2]} = 0$$

και

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} L_n(\sigma) &= \prod_{i=1}^n \lim_{\sigma \rightarrow \infty} \frac{\sigma}{D[\sigma^2 + (x_i - \mu)^2]} \\ &= \prod_{i=1}^n \lim_{\sigma \rightarrow \infty} \frac{1}{\frac{\partial D}{\partial \sigma}[\sigma^2 + (x_i - \mu)^2] + 2D\sigma} = 2^{-n}. \end{aligned}$$

Επομένως, υπό τη συνθήκη (3.1.14), η συνάρτηση  $L_n(\sigma)$  παίρνει μέγιστη τιμή σε πεπερασμένη τιμή της παραμέτρου  $\sigma$ . Επομένως υπάρχει ο Ε.Μ.Π. της  $\sigma$ . ■

Σύμφωνα με το αποτέλεσμα του προηγούμενου θεώρηματος μια επαρκής συνθήκη για την ύπαρξη του Ε.Μ.Π. της παραμέτρου κλίμακας όταν  $X_1, \dots, X_n$  είναι ένα τ.δ. από την  $TC^{(-1,1)}(\mu, \sigma)$ , με  $-1 < \mu < 1$ , είναι η ακόλουθη:

$$\frac{\sum_{i=1}^n (x_i - \mu)^2}{n} < \left( \mu^2 + \frac{1}{3} \right).$$

Η συνθήκη αυτή αποδεικνύεται από τους Daliya *et al.* (2001) ότι είναι και ικανή για την ύπαρξη του Ε.Μ.Π. της παραμέτρου κλίμακας για μεγάλου μεγέθους δείγματα. Πριν την απόδειξη αυτού του αποτελέσματος είναι απαραίτητο να διατυπωθούν κάποιοι ορισμοί και να αποδειχθούν κάποια χρήσιμα αποτελέσματα.

### Λήμμα 3.1.1 (Daliya *et al.*, 2001)

Η συνάρτηση  $\varphi_1(\sigma)$  είναι αυστηρά αύξουσα από το 0 μέχρι το  $\mu^2 + 1/3$ , καθώς η παράμετρος κλίμακας  $\sigma$  τείνει στο άπειρο.

### Απόδειξη

Το επιθυμητό αποτέλεσμα θα αποδειχθεί για την περίπτωση που  $\mu = 0$ . Ομοίως αποδεικνύεται και στη γενική περίπτωση. Είναι τότε:

$$\varphi_1(\sigma) = \frac{\sigma}{\tan^{-1}(1/\sigma)} - \sigma^2, \quad (3.1.25)$$

και

$$\frac{\partial \varphi_1(\sigma)}{\partial \sigma} = \frac{1}{\tan^{-1}(1/\sigma)} + \frac{1}{[\tan^{-1}(1/\sigma)]^2} \frac{\sigma}{1+\sigma^2} - 2\sigma \quad (3.1.26)$$

Θέτοντας  $u = 1/\sigma$ , η σχέση (3.1.26) είναι αυστηρά αύξουσα για  $\sigma > 0$ , αν και μόνο αν ισχύει

$$\frac{u}{\tan^{-1}(u)} + \frac{1}{[\tan^{-1}(u)]^2} \frac{u}{1+u^2} > 2, \quad (3.1.27)$$

με  $u > 0$ . Η σχέση (3.1.27) θέτοντας  $v = \tan^{-1}(u)$ , γράφεται ισοδύναμα ως εξής:

$$\frac{\tan(v)}{v} + \frac{1}{v^2} \frac{\tan^2(v)}{\sec^2(v)} > 2, \quad (3.1.28)$$

όπου  $0 < v < \pi/2$  και με  $\sec(\cdot) = \cos^{-1}(\cdot)$  συμβολίζεται η αντίστροφη συνάρτηση του συνημιτόνου.

Η σχέση (3.1.28) λόγω της σχέσης  $\tan(v) = \frac{\sin(v)}{\cos(v)}$  γράφεται ισοδύναμα

$$\frac{\sin(v)}{v} \frac{1}{\cos(v)} + \frac{\sin^2(v)}{v^2} > 2. \quad (3.1.29)$$

Λαμβάνοντας υπόψη ότι (βλέπε, μεταξύ άλλων, Mitrinovic, 1970)

$$\frac{1}{\cos(v)} > \frac{v^3}{\sin^3(v)}, \text{ για } 0 < v < \pi/2,$$

προκύπτει ότι

$$\frac{\sin(v)}{v} \frac{1}{\cos(v)} + \frac{\sin^2(v)}{v^2} > \frac{v^2}{\sin^2(v)} + \frac{\sin^2(v)}{v^2} \geq 2, \text{ για } 0 < v < \pi/2.$$

Άρα, ισχύει η σχέση (3.1.29), επομένως ισχύει ότι  $\frac{\partial \varphi_1(\sigma)}{\partial \sigma} > 2$  και καταλήγουμε στο επιθυμητό αποτέλεσμα. ■

### Ορισμός 3.1.1 (Marshall and Olkin (1974))

Το  $p$ -διάστατο διάνυσμα  $a$  λέγεται ότι majorized (υπερκαλύπτεται) από το  $p$ -διάστατο διάνυσμα  $b$ , και γράφουμε  $a \prec b$ , αν, έπειτα από αναδιάταξη των συνιστωσών τους έτσι ώστε  $a_1 \geq a_2 \geq \dots \geq a_p$  και  $b_1 \geq b_2 \geq \dots \geq b_p$  ισχύει ότι:

$$\sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i, \quad k=1, \dots, p-1, \text{ και } \sum_{i=1}^p a_i = \sum_{i=1}^p b_i. \quad (3.1.30)$$

Επιπλέον, οι συναρτήσεις  $h$  για τις οποίες  $a \prec b$  συνεπάγεται ότι  $h(a) \leq h(b)$  λέγονται Schur-convex, ενώ οι συναρτήσεις  $h$  για τις οποίες  $a \prec b$  συνεπάγεται ότι  $h(a) \geq h(b)$  λέγονται Schur-concave. Προκύπτει ότι (βλέπε Marshall and Olkin (1974) και αναφορές εκεί) ότι μία αναγκαία και επαρκής συνθήκη έτσι ώστε μια διαφορίσιμη συνάρτηση  $f(w_1, w_2, \dots, w_p)$  να είναι Schur-concave (Schur-convex) είναι η ακόλουθη

$$\left( \frac{\partial f}{\partial w_i} - \frac{\partial f}{\partial w_j} \right) (w_i - w_j) \leq (\geq) 0, \forall i \neq j.$$

Στο ακόλουθο θεώρημα διατυπώνεται η ικανή και αναγκαία συνθήκη για την ύπαρξη πεπερασμένου Ε.Μ.Π. της παραμέτρου κλίμακας  $\sigma$  όταν  $X_1, \dots, X_n$  ένα τ.δ. από την  $TC^{(-1,1)}(\mu, \sigma)$ , με  $-1 < \mu < 1$ , με το  $n$  να θεωρείται αρκετά μεγάλο.

### Θεώρημα 3.1.2 (Daliya et al., 2001)

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα μεγέθους  $n$ , με το  $n$  να θεωρείται αρκετά μεγάλο, από έναν πληθυσμό που ακολουθεί την  $TC^{(-1,1)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από την σχέση (2.1.2). Η ικανή και αναγκαία συνθήκη για την ύπαρξη του Ε.Μ.Π. της παραμέτρου  $\sigma$  είναι η εξής:

$$\frac{\sum_{i=1}^n (x_i - \mu)^2}{n} < \left( \mu^2 + \frac{1}{3} \right). \quad (3.1.31)$$

### Απόδειξη

Έχουμε δείξει ότι η συνθήκη αυτή είναι επαρκής, θα δείξουμε ότι είναι και ικανή για μεγάλο μέγεθος δείγματος. Υπενθυμίζεται ότι  $X_{(n)} = 1$  και  $X_{(1)} = -1$  και ότι η διερεύνηση γίνεται υπό την υπόθεση ότι  $\mu \in (a, b) = (-1, 1)$ . Έστω  $w = (w_1, w_2, \dots, w_n)^t$  είναι το  $n$ -διάστατο διάνυσμα με συνιστώσες που δίνονται από τις σχέσεις  $w_i = (X_i - \mu)^2$ , και έστω  $a$  το  $n$ -διάστατο διάνυσμα

$$a = (1, (n\bar{w}-1)/(n-1), \dots, (n\bar{w}-1)/(n-1)),$$



με  $\bar{w} = \sum_{i=1}^n w_i / n$ . Τότε διαπιστώνεται εύκολα ότι το διάνυσμα  $a$  majorizes το διάνυσμα  $w$ , καθώς ικανοποιείται η σχέση (3.1.30) του Ορισμού 3.1.1.

Η συνάρτηση  $\prod_{i=1}^n (w_i / \sigma^2 + 1)$  ως συνάρτηση των  $(w_1, w_2, \dots, w_n)$  είναι Schur-concave καθώς εύκολα προκύπτει ότι

$$\left( \frac{\partial f}{\partial w_i} - \frac{\partial f}{\partial w_j} \right) (w_i - w_j) \leq 0, \forall i \neq j, i, j = 1, \dots, n.$$

Επομένως, έχουμε ότι:

$$\left( \frac{n\bar{w}-1}{n-1} \frac{1}{\sigma^2} + 1 \right)^{n-1} \left( \frac{1}{\sigma^2} + 1 \right) \leq \prod_{i=1}^n \left( \frac{w_i}{\sigma^2} + 1 \right). \quad (3.1.32)$$

Προφανώς, καθώς  $\lim_{\sigma \rightarrow \infty} L_n(\sigma) = 2^{-n}$ , για να υπάρχει ο Ε.Μ.Π. της παραμέτρου  $\sigma$  θα πρέπει να υπάρχει μια πεπερασμένη τιμή της παραμέτρου έστω  $\sigma_0$ , τέτοια ώστε  $L_n(\sigma_0) > 1/2^n$  ή ισοδύναμα  $\prod_{i=1}^n (w_i + \sigma_0^2) < (2/\sigma_0 D_0)^n$ , όπου  $D_0 = D(\sigma_0)$ . Επομένως, λόγω της (3.1.32) θα πρέπει

$$\left( \frac{n\bar{w}-1}{n-1} \frac{1}{\sigma_0^2} + 1 \right)^{n-1} \left( \frac{1}{\sigma_0^2} + 1 \right) \leq \left( \frac{2}{\sigma_0 D_0} \right)^n. \quad (3.1.33)$$

Μετά από αλγεβρικές πράξεις η σχέση (3.1.33) γράφεται ως εξής

$$\frac{n}{n-1} \bar{w} < \sigma_0^2 \left[ \left( \frac{2}{\sigma_0 D_0} \right)^{n/(n-1)} \left( \frac{\sigma_0^2}{1 + \sigma_0^2} \right)^{1/(n-1)} - 1 \right] + \frac{1}{n-1},$$

ή ισοδύναμα

$$\bar{w} < \sigma_0^2 \left[ \left( \frac{2}{\sigma_0 D_0} \right)^{n/(n-1)} - 1 \right] = \varphi_1(\sigma_0).$$

Χρησιμοποιώντας το Λήμμα 3.1.1, καταλήγουμε ότι:  $\bar{w} < \left( \mu^2 + \frac{1}{3} \right)$ , για  $n \geq n_0$ .

Από τα προηγούμενα θεωρήματα προκύπτει ότι αν  $X_1, \dots, X_n$  είναι ένα τυχαίο δείγμα μεγέθους  $n$  από έναν πληθυσμό που ακολουθεί την  $TC^{(-1,1)}(\mu, \sigma)$ , τότε δεν είναι εξασφαλισμένη πάντοτε η ύπαρξη πεπερασμένου Ε.Μ.Π. της παραμέτρου  $\sigma$ . Επομένως, προκύπτει εύλογα το ερώτημα του προσδιορισμού της πιθανότητας μη ύπαρξης του Ε.Μ.Π. της παραμέτρου  $\sigma$ . Ο προσδιορισμός αυτός αποτελεί αντικείμενο μελέτης της ενότητας 3.2.1.

Στην παράγραφο που ακολουθεί, το ενδιαφέρον επικεντρώνεται στον προσδιορισμό του ασυμπτωτικού πίνακα διακυμάνσεων συνδιακυμάνσεων των Ε.Μ.Π., καθώς είναι ιδιαίτερα χρήσιμος για την εκτίμηση σε διάστημα των αγνώστων παραμέτρων και τον έλεγχο στατιστικών υποθέσεων, που εκφράζονται συναρτήσει των αγνώστων παραμέτρων.

#### **Ασυμπτωτικός πίνακας διακυμάνσεων συνδιακυμάνσεων των Ε.Μ.Π.**

Αρχικά, χωρίς βλάβη της γενικότητας, ο προσδιορισμός διεξάγεται για την περίπτωση της  $TC^{(-1,1)}(\mu, \sigma)$ , δηλαδή υποθέτοντας ότι τα σημεία αποκοπής είναι  $a = -1$  και  $b = 1$  και στη συνέχεια αποδεικνύεται ότι η ασυμπτωτική κατανομή των Ε.Μ.Π. είναι ίδια είτε τα σημεία αποκοπής θεωρηθούν γνωστά είτε άγνωστα.

**Θεώρημα 3.1.3 (Daliya et al., 2001, Nadarajah, 2011)**

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα μεγέθους  $n$ , με το  $n$  να θεωρείται αρκετά μεγάλο, από έναν πληθυσμό που ακολουθεί την  $TC^{(-1,1)}(\mu, \sigma)$ , με σ.π.π. που δίνεται από τη σχέση

$$f_X(x) = \frac{\sigma}{D} \frac{1}{\left[ \sigma^2 + (x - \mu)^2 \right]}, \quad -1 \leq x \leq 1, \quad \mu \in (-1, 1),$$

και  $D = \tan^{-1}\left(\frac{1-\mu}{\sigma}\right) + \tan^{-1}\left(\frac{1+\mu}{\sigma}\right)$ . Ο πίνακας διακυμάνσεων

συνδιακυμάνσεων των Ε.Μ.Π. είναι ο  $(V^{-1}/n)$ , όπου  $V$  είναι ο συμμετρικός  $2 \times 2$

πίνακας πληροφορίας του Fisher,  $V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$ , με στοιχεία  $v_{ij}$ ,  $i, j = 1, 2$ , τα

οποία προσδιορίζονται από τις ακόλουθες σχέσεις:

$$v_{11} = \frac{1}{D} \frac{\partial^2 D}{\partial \mu^2} - \frac{1}{D^2} \left( \frac{\partial D}{\partial \mu} \right)^2 + \frac{2}{\sigma^2} I(1,0) - \frac{4}{\sigma^2} I(2,2),$$

$$v_{12} = v_{21} = \frac{1}{D} \frac{\partial^2 D}{\partial \sigma \partial \mu} - \frac{1}{D^2} \frac{\partial D}{\partial \sigma} \frac{\partial D}{\partial \mu} + \frac{4}{\sigma^2} I(1,1) - \frac{4}{\sigma^2} I(2,3),$$

$$v_{22} = -\frac{1}{\sigma^2} - \frac{1}{D^2} \left( \frac{\partial D}{\partial \sigma} \right)^2 + \frac{1}{D} \frac{\partial^2 D}{\partial \sigma^2} + \frac{6}{\sigma^2} I(1,2) - \frac{4}{\sigma^2} I(2,4),$$

όπου

$$I(k, l) = E \left[ \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-k} \left( \frac{X - \mu}{\sigma} \right)^l \right],$$

ή ισοδύναμα

$$I(k, l) = \frac{\sigma}{l+1} \left[ B^{l+1} {}_2F_1 \left( \frac{l+1}{2}, k+1; \frac{l+3}{2}; -B^2 \right) - A^{l+1} {}_2F_1 \left( \frac{l+1}{2}, k+1; \frac{l+3}{2}; -A^2 \right) \right]$$

με  $A = \frac{a - \mu}{\sigma}$ ,  $B = \frac{b - \mu}{\sigma}$  και  ${}_2F_1(\cdot; \cdot; \cdot)$  την υπεργεωμετρική συνάρτηση

Gauss.

### Απόδειξη

Αρχικά πρέπει να υπολογιστούν οι μερικές παράγωγοι του λογαρίθμου της σ.π.π. της  $TC^{(-1,1)}(\mu, \sigma)$ , που δίνεται από τη σχέση:

$$\log f_X(X) = \log \sigma - \log D - \log \left[ \sigma^2 + (X - \mu)^2 \right],$$

ή ισοδύναμα

$$\log f_X(X) = -\log \sigma - \log D - \log \left[ 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right].$$

Με αλγεβρικές πράξεις έχουμε ότι:

$$\frac{\partial \log f_X(X)}{\partial \mu} = -\frac{1}{D} \frac{\partial D}{\partial \mu} + \frac{2}{\sigma^2} (X - \mu) \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-1},$$

$$\begin{aligned} \frac{\partial^2 \log f_X(x)}{\partial \mu^2} &= \frac{1}{D^2} \left( \frac{\partial D}{\partial \mu} \right)^2 - \frac{1}{D} \frac{\partial^2 D}{\partial \mu^2} \\ &\quad - \frac{2}{\sigma^2} \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-1} + \frac{4}{\sigma^2} (X - \mu)^2 \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \log f_X(x)}{\partial \mu \partial \sigma} &= \frac{1}{D^2} \frac{\partial D}{\partial \mu} \frac{\partial D}{\partial \sigma} - \frac{1}{D} \frac{\partial^2 D}{\partial \mu \partial \sigma} - \frac{4}{\sigma^3} (X - \mu) \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-1} \\ &\quad + \frac{4}{\sigma^2} \left( \frac{X - \mu}{\sigma} \right)^3 \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-2}, \end{aligned}$$

$$\frac{\partial \log f_X(x)}{\partial \sigma} = -\frac{1}{\sigma} - \frac{1}{D} \frac{\partial D}{\partial \sigma} - \frac{2}{\sigma} \left( \frac{X - \mu}{\sigma} \right)^2 \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-1},$$

και

$$\begin{aligned} \frac{\partial^2 \log f_X(x)}{\partial \sigma^2} &= \frac{1}{\sigma^2} + \frac{1}{D^2} \left( \frac{\partial D}{\partial \sigma} \right)^2 - \frac{1}{D} \frac{\partial^2 D}{\partial \sigma^2} - \frac{6}{\sigma^2} \left( \frac{X - \mu}{\sigma} \right)^2 \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-1} \\ &\quad + \frac{4}{\sigma^2} \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-2} \left( \frac{X - \mu}{\sigma} \right)^4. \end{aligned}$$

Επομένως, εύκολα προκύπτει, θέτοντας

$$I(k, l) = E \left[ \left( 1 + \left( \frac{X - \mu}{\sigma} \right)^2 \right)^{-k} \left( \frac{X - \mu}{\sigma} \right)^l \right]$$

ότι:

$$v_{11} = -\frac{1}{D^2} \left( \frac{\partial D}{\partial \mu} \right)^2 + \frac{1}{D} \frac{\partial^2 D}{\partial \mu^2} + \frac{2}{\sigma^2} I(1,0) - \frac{4}{\sigma^2} I(2,2),$$

$$v_{12} = v_{21} = -\frac{1}{D^2} \frac{\partial D}{\partial \mu} \frac{\partial D}{\partial \sigma} + \frac{1}{D} \frac{\partial^2 D}{\partial \mu \partial \sigma} + \frac{4}{\sigma^2} I(1,1) - \frac{4}{\sigma^2} I(2,3),$$

και

$$v_{22} = -\frac{1}{\sigma^2} - \frac{1}{D^2} \left( \frac{\partial D}{\partial \sigma} \right)^2 + \frac{1}{D} \frac{\partial^2 D}{\partial \sigma^2} + \frac{6}{\sigma^2} I(1,2) - \frac{4}{\sigma^2} I(2,4).$$

Από τον ορισμό της μέσης τιμής προκύπτει ότι:

$$\begin{aligned} I(k,l) &= \frac{1}{\sigma D} \int_a^b \left( 1 + \left( \frac{X-\mu}{\sigma} \right)^2 \right)^{-(k+1)} \left( \frac{X-\mu}{\sigma} \right)^l dX \\ &= \frac{1}{2D} \int_{A^2}^{B^2} (1+y)^{-(k+1)} (y)^{l/2-1/2} dy \end{aligned}$$

ή ισοδύναμα

$$I(k,l) = \frac{1}{2D} \left[ \int_0^{B^2} (1+y)^{-(k+1)} (y)^{l/2-1/2} dy - \int_0^{A^2} (1+y)^{-(k+1)} (y)^{l/2-1/2} dy \right].$$

Από τη μονογραφία των Gradshteyn and Ryzhik (2000) είναι γνωστό (βλέπε σχέση (3.194.1)) ότι:

$$\int_0^u \frac{x^{\lambda-1}}{(1+\beta x)^\nu} dx = \frac{u^\lambda}{\lambda} {}_2F_1(\nu, \lambda; 1-\lambda; -\beta u), \quad |\arg(1+\beta u)| < \pi, \quad \operatorname{Re}(\lambda) > 0,$$

όπου με  ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$  συμβολίζεται η υπεργεωμετρική συνάρτηση Gauss, όπως αυτή ορίστηκε στη σχέση (2.2.3). Επομένως, εφαρμόζοντας τη σχέση (3.194.1) των Gradshteyn and Ryzhik (2000) στην ειδική περίπτωση που  $u = c^2, \lambda = \frac{k+1}{2}$ ,  $\nu = k+1$  και  $\beta = 1$ , για  $c = A, B$ , προκύπτει το ζητούμενο αποτέλεσμα. ■

Η ασυμπτωτική κατανομή των Ε.Μ.Π. των παραμέτρων  $\mu$  και  $\sigma$  είναι ίδια είτε τα σημεία αποκοπής  $a$  και  $b$  είναι γνωστά είτε άγνωστα. Αυτό προκύπτει διότι (βλέπε Daliya *et al.*, 2001)  $\sqrt{n}(\hat{a}-a) \xrightarrow{P} 0$  και  $\sqrt{n}(\hat{b}-b) \xrightarrow{P} 0$ .

### 3.2 Μελέτη Προσομοίωσης

Στην ενότητα αυτή παρακινούμενοι από κάποια σχόλια των Daliya *et al.* (2001) με χρήση προσομοιωμένων δεδομένων εκτιμάται η πιθανότητα μη ύπαρξης πεπερασμένου Ε.Μ.Π. για την παράμετρο  $\sigma$ . Έπειτα διεξάγεται μία μικρή σε έκταση μελέτη προσομοίωσης για την αξιολόγηση της απόδοσης των Ε.Μ.Π. των αγνώστων παραμέτρων, για την περίπτωση της  $TC^{(-1,1)}(\mu, \sigma)$ , δηλαδή υποθέτοντας ότι τα σημεία αποκοπής είναι  $a = -1$  και  $b = 1$ , γνωστά. Οι προσομοιώσεις έγιναν χρησιμοποιώντας τη Matlab και οι κώδικες που χρησιμοποιήθηκαν παρατίθενται στο Παράρτημα Β.

### 3.2.1 Πιθανότητα μη ύπαρξης του Ε.Μ.Π. της παραμέτρου $\sigma$

Λαμβάνοντας υπόψη τα αποτελέσματα των Θεωρημάτων 3.1.1 και 3.1.2 οι Daliya *et al.* (2001) συμπεραίνουν ότι η πιθανότητα μη ύπαρξης πεπερασμένου Ε.Μ.Π. της παραμέτρου κλίμακας  $\sigma$  εξαρτάται από τις τιμές των παραμέτρων  $\mu, \sigma$  και  $n$ . Η πιθανότητα αυτή, για  $n=1$  είναι εύκολο να διαπιστωθεί ότι μεγιστοποιείται για  $\mu=0$  και είναι αύξουσα συνάρτηση της παραμέτρου κλίμακας  $\sigma$ , λόγω των ιδιοτήτων της αντίστροφης συνάρτησης της εφαπτομένης. Αν τώρα θεωρήσουμε ότι  $\mu=0$  σύμφωνα με τη σχέση (3.1.14) η πιθανότητα μη ύπαρξης πεπερασμένου Ε.Μ.Π. της παραμέτρου κλίμακας  $\sigma$ , για μεγάλο μέγεθος

δείγματος  $n$ , ανάγεται στον υπολογισμό της πιθανότητας  $P\left(\sum_{i=1}^n X_i^2 \geq \frac{n}{3}\right)$ . Η

πιθανότητα αυτή δύναται να υπολογιστεί προσεγγιστικά χρησιμοποιώντας το Κεντρικό Οριακό Θεώρημα. Έτσι, αν π.χ.  $n=5$  και  $\sigma=1$ , χρησιμοποιώντας τα αποτελέσματα του Πορίσματος 2.2.3 προκύπτει ότι προσεγγιστικά ισούται με  $P(Z > 3.1789)$ , όπου  $Z \sim N(0,1)$ . Επομένως, σε αυτήν την περίπτωση η πιθανότητα μη ύπαρξης πεπερασμένου Ε.Μ.Π. της παραμέτρου κλίμακας  $\sigma$ , ακόμα και για τ.δ. μικρού μεγέθους, είναι υπερβολικά μικρή. Οι Daliya *et al.* (2001) αναφέρουν ότι παρόμοιο συμπέρασμα ισχύει και για άλλες περιπτώσεις, και ειδικότερα για μέτρια σε μέγεθος τ.δ.  $TC^{(-1,1)}(\mu, \sigma)$ . Παρόλα αυτά δεν παραθέτουν επιπλέον αποτελέσματα παραπέμποντας τον αναγνώστη στη διδακτορική διατριβή του Staneski (1990), η πρόσβαση στην οποία δεν μας ήταν εφικτή.

Στη συνέχεια, παρακινούμενοι από τους Daliya *et al.* (2001) προσομοιώνουμε δεδομένα από την  $TC^{(-1,1)}(\mu, \sigma)$ , για διάφορες τιμές των



παραμέτρων  $\mu$ ,  $\sigma$  και  $n$ , με  $\mu \in (-1,1)$  και  $\sigma > 0$ . Για τις επιλεγμένες τιμές των παραμέτρων  $\mu = 0, 0.25, 0.5, 0.75$ ,  $\sigma = 0.5, 1, 1.5, 2, 2.5, 3, 4$ , και  $n = 20, 30, 50, 100, 200$ , με 10.000 δειγματοληψίες, εκτιμήθηκε για κάθε συνδυασμό των παραμέτρων αυτών η πιθανότητα

$$P\left[\sum_{i=1}^n (x_i - \mu)^2 \geq n\left(\mu^2 + \frac{1}{3}\right)\right],$$

δηλαδή η πιθανότητα μη ύπαρξης πεπερασμένου εκτιμητή για την παράμετρο  $\sigma$ .

Τα αποτελέσματα παρουσιάζονται στους Πίνακες 4-7 που ακολουθούν. Από τους πίνακες αυτούς προκύπτουν τα ακόλουθα συμπεράσματα:

- α) Για σταθεροποιημένες τιμές των παραμέτρων  $\mu$  και  $n$  προκύπτει ότι η προσομοιωμένη τιμή της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή για την παράμετρο  $\sigma$  αυξάνει καθώς αυξάνει η τιμή της παραμέτρου  $\sigma$ .
- β) Για σταθεροποιημένες τιμές των παραμέτρων  $\mu$  και  $\sigma$  προκύπτει ότι η προσομοιωμένη τιμή της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή για την παράμετρο  $\sigma$  μειώνεται καθώς αυξάνει το μέγεθος του δείγματος.
- γ) Τέλος, για σταθεροποιημένες τιμές των παραμέτρων  $n$  και  $\sigma$  προκύπτει ότι η προσομοιωμένη τιμή της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή για την παράμετρο  $\sigma$  μειώνεται καθώς αυξάνει η τιμή της παραμέτρου  $\mu$ .

**Πίνακας 4:** Προσομοιωμένες τιμές, βάση 10.000 δειγματοληψιών, από την  $TC^{(-1)}(0, \sigma)$ ,

των πιθανοτήτων  $P\left[\sum_{i=1}^n (x_i - \mu)^2 \geq n\left(\mu^2 + \frac{1}{3}\right)\right]$  για διάφορους συνδυασμούς των παραμέτρων  $n$  και  $\sigma$ .

$n \setminus \sigma$	0.5	1	1.5	2	2.5	3	4
20	0.0120	0.1738	0.2918	0.3730	0.4142	0.4246	0.4505
30	0.0031	0.1198	0.2614	0.3563	0.3977	0.4172	0.4479
50	0.0002	0.0678	0.2119	0.3101	0.3746	0.4096	0.4460
100	0	0.0013	0.1318	0.2517	0.3247	0.3680	0.4207
200	0	0.0011	0.0582	0.1698	0.2624	0.3191	0.3936

**Πίνακας 5:** Προσομοιωμένες τιμές, βάση 10.000 δειγματοληψιών, από την  $TC^{(-1)}(0.25, \sigma)$

των πιθανοτήτων  $P\left[\sum_{i=1}^n (x_i - \mu)^2 \geq n\left(\mu^2 + \frac{1}{3}\right)\right]$  για διάφορους συνδυασμούς των παραμέτρων  $n$  και  $\sigma$ .

$n \setminus \sigma$	0.5	1	1.5	2	2.5	3	4
20	0.0065	0.1114	0.2479	0.3339	0.3834	0.4086	0.4398
30	0.0020	0.0739	0.2082	0.3012	0.3545	0.3936	0.4352
50	0.0001	0.0342	0.1534	0.2556	0.3212	0.3767	0.4234
100	0	0.0057	0.0772	0.1741	0.2689	0.3257	0.3921
200	0	0.0001	0.0201	0.0971	0.1970	0.2650	0.3599

**Πίνακας 6:** Προσομοιωμένες τιμές, βάση 10.000 δειγματοληψιών από την  $TC^{(4,1)}(0.5, \sigma)$ ,

των πιθανοτήτων  $P\left[\sum_{i=1}^n (x_i - \mu)^2 \geq n\left(\mu^2 + \frac{1}{3}\right)\right]$  για διάφορους συνδυασμούς των παραμέτρων  $n$  και  $\sigma$ .

$n \setminus \sigma$	0.5	1	1.5	2	2.5	3	4
20	0.0019	0.0561	0.1732	0.2572	0.3282	0.3749	0.4126
30	0	0.0297	0.1228	0.2258	0.3282	0.3539	0.4012
50	0	0.0065	0.0720	0.1669	0.2511	0.3110	0.3837
100	0	0.0002	0.0203	0.0902	0.1856	0.2543	0.3405
200	0	0	0.0024	0.0324	0.0966	0.1717	0.2850

**Πίνακας 7:** Προσομοιωμένες τιμές, βάση 10.000 δειγματοληψιών από την  $TC^{(4,1)}(0.75, \sigma)$

των πιθανοτήτων  $P\left[\sum_{i=1}^n (x_i - \mu)^2 \geq n\left(\mu^2 + \frac{1}{3}\right)\right]$  για διάφορους συνδυασμούς των παραμέτρων  $n$  και  $\sigma$ .

$n \setminus \sigma$	0.5	1	1.5	2	2.5	3	4
20	0.0002	0.0255	0.1106	0.2150	0.2855	0.3423	0.3983
30	0	0.0083	0.0682	0.1598	0.2397	0.3055	0.3711
50	0	0.0004	0.0277	0.1073	0.1893	0.2540	0.3395
100	0	0	0.0036	0.0424	0.1068	0.1875	0.2939
200	0	0	0	0.0057	0.0415	0.1046	0.2304

### 3.2.2 Αξιολόγηση των Ε.Μ.Π.

Οι Daliya *et al.* (2001) θέλοντας να αξιολογήσουν τις ιδιότητες των εκτιμητών μέγιστης πιθανοφάνειας των παραμέτρων  $\mu$  και  $\sigma$  διεξάγουν μία μελέτη προσομοίωσης, με βάση 1.000 επαναλήψεις (δειγματοληψίες) από την  $TC^{(-1,1)}(\mu, \sigma)$ , μεγέθους  $n = 20$ , για διάφορους συνδυασμούς των παραμέτρων  $\mu, \sigma$ . Στο πλαίσιο αυτό παραθέτουν τα αποτελέσματα που αφορούν την εκτιμώμενη μεροληψία (biased) και το εκτιμώμενο μέσο τετραγωνικό σφάλμα (mean square error) των εκτιμητών. Παρακινούμενοι από αυτήν την εργασία στη συνέχεια παρουσιάζουμε τα αποτελέσματα μιας ανάλογης μελέτης προσομοίωσης θεωρώντας διάφορους συνδυασμούς για τις παραμέτρους  $\mu, \sigma$ , αλλά στηριζόμενοι σε 5.000 επαναλήψεις από τ.δ. μεγέθους  $n = 20, 50$ . Ειδικότερα στα πλαίσια αυτής της μεταπτυχιακής διατριβής παρουσιάζονται τα αποτελέσματα που προκύπτουν θεωρώντας τους συνδυασμούς των ακόλουθων τιμών για τις παραμέτρους:  $\mu = 0, 0.2, 0.5$ ,  $\sigma = 0.05, 0.10, 15, 0.2, 0.25$ .

Πριν προχωρήσουμε στην παράθεση και ερμηνεία των αποτελεσμάτων, για λόγους πληρότητας, στην παρατήρηση που ακολουθεί εξηγούμε τον τρόπο υπολογισμού των εκτιμητών της μεροληψίας και του μέσου τετραγωνικού σφάλματος.

**Παρατήρηση 3.2.2.1** Έστω  $X_1^{(B)}, \dots, X_n^{(B)}$  τ.δ. μεγέθους  $n$ , όπου π.χ.  $B = 1, \dots, 5.000$ , από έναν πληθυσμό που ακολουθεί την  $TC^{(-1,1)}(\mu, \sigma)$ . Συμβολίζουμε με  $\theta = (\theta_1, \theta_2) = (\mu, \sigma)$  το διάνυσμα των άγνωστων παραμέτρων. Χρησιμοποιώντας τα τ.δ.  $X_1^{(B)}, \dots, X_n^{(B)}$  προσδιορίζονται  $B$  το πλήθος εκτιμητές  $\hat{\theta}^{(B)} = (\hat{\theta}_1^{(B)}, \hat{\theta}_2^{(B)})$ . Τότε ο εκτιμητής της μεροληψίας της  $\theta_j$  προσδιορίζεται ως:

$$\frac{1}{B} \sum_{i=1}^B (\hat{\theta}_j^{(i)} - \theta_j),$$

ενώ ο εκτιμητής του μέσου τετραγωνικού σφάλματος υπολογίζεται από τη σχέση:

$$\frac{1}{B} \sum_{i=1}^B (\hat{\theta}_j^{(i)} - \theta_j)^2.$$

Από τα αποτελέσματα που δίνονται στους Πίνακες 8 και 9 εξάγονται τα ακόλουθα συμπεράσματα.

α) Για  $n = 50$ , με βάση τα προσομοιωμένα δεδομένα, το μέσο τετραγωνικό σφάλμα των εκτιμητών των παραμέτρων είναι μικρότερο από το αντίστοιχο για  $n = 20$ . Το γεγονός αυτό επιβεβαιώνει το αναμενόμενο αποτέλεσμα ότι έχουμε καλύτερους εκτιμητές όσο το μέγεθος του δείγματος μεγαλώνει.

β) Όσο αυξάνεται η τιμή της παραμέτρου  $\sigma$ , ενώ οι υπόλοιπες παράμετροι παραμένουν σταθερές, αυξάνεται το μέσο τετραγωνικό σφάλμα των εκτιμητών των παραμέτρων. Αυτό έχει ως αποτέλεσμα να έχουμε καλύτερους εκτιμητές όταν η πραγματική τιμή της παραμέτρου είναι μικρή.

γ) Όμοια με το β) προκύπτει ότι καθώς αυξάνεται η τιμή της παραμέτρου  $\mu$ , ενώ οι υπόλοιπες παράμετροι παραμένουν σταθερές, κάποιες φορές, αυξάνεται το μέσο τετραγωνικό σφάλμα των εκτιμητών των παραμέτρων.

Το β) συμφωνεί απόλυτα με τα συμπεράσματα των Daliya *et al.* (2001, σελ. 1749), οι οποίοι δικαιολογούν αυτή τη συμπεριφορά καθώς «οι αποκομμένες κατανομές με τέτοιες τιμές για τις παραμέτρους είναι πολύ παραμορφωμένες σε σχέση με την ολοκληρωμένη μορφή τους κάνοντας τις παραμέτρους να

διακρίνονται λιγότερο εύκολα από ένα δείγμα από μια τέτοια σοβαρά αποκομμένη κατανομή.»

**Πίνακας 8:** Προσομοιωμένη μεροληψία (μέσο τετραγωνικό σφάλμα, στη δεύτερη γραμμή) των Ε.Μ.Π. της παραμέτρου  $\mu$ , βάση 5.000 δειγματοληψιών, από την  $TC^{(-1)}(\mu, \sigma)$ , για διάφορους συνδυασμούς των παραμέτρων  $n$ ,  $\mu$  και  $\sigma$ .

		$\mu = 0$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		-0.000066	0.000081	0.000232	0.000690	0.001204
		0.000280	0,001047	0,002387	0,004277	0,008923
50		-0.000074	0.000208	0.000406	-0.000489	-0.000674
		0.000100	0.000417	0.000880	0.001547	0.002351
		$\mu = 0.2$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		0.000116	-0.000073	-0.000738	0.001228	0.002077
		0.000285	0.001071	0.002392	0.004042	0.009644
50		-0.000065	-0.000309	-0.000334	0.000185	0.000451
		0.000103	0.000391	0.000873	0.001517	0.002305
		$\mu = 0.5$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		0.000155	-0.000536	0.000963	0.002785	0.008058
		0.000274	0.001041	0.002430	0.004725	0.008804
50		-0.000037	0.000256	0.000478	0.001235	0.001246
		0.000102	0.000389	0.000855	0.001538	0.002442

**Πίνακας 9 :** Προσομοιωμένη μεροληψία (μέσο τετραγωνικό σφάλμα, στη δεύτερη γραμμή) των Ε.Μ.Π. της παραμέτρου  $\sigma$ , βάση 5.000 δειγματοληγιών, από την  $TC^{(c,d)}(\mu, \sigma)$ , για διάφορους συνδυασμούς των παραμέτρων  $n$ ,  $\mu$  και  $\sigma$ .

		$\mu = 0$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		-0.000625	-0.000504	0.000944	0.002726	0.004749
		0.000336	0.001227	0.003097	0.006138	0.012751
50		-0.000037	-0.000427	0.000804	0.000390	0.001654
		0.000115	0.000462	0.001187	0.002017	0.003617
		$\mu = 0.2$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		-0.001513	0.000040	0.000149	0.000548	0.005741
		0.000480	0.001252	0.002923	0.006290	0.012766
50		-0.000568	0.000459	-0.000187	0.000331	0.001021
		0.000159	0.000489	0.001097	0.002174	0.003628
		$\mu = 0.5$				
$n \setminus \sigma$		0.05	0.10	0.15	0.20	0.25
20		-0.000107	-0.000369	0.000523	0.000059	0.003184
		0.000302	0.001311	0.003316	0.006845	0.011764
50		0.000050	0.000139	0.000522	0.000992	0.001338
		0.000115	0.000488	0.001162	0.002344	0.003889

### 3.3 Αριθμητικό Παράδειγμα

Στην ενότητα αυτή, παρακινούμενοι από διάφορους συγγραφείς (βλέπε, μεταξύ άλλων, Nadarajah and Kotz, 2007), θα θεωρήσουμε ένα πραγματικό σύνολο δεδομένων και θα προσαρμόσουμε σε αυτό τη διπλά αποκομμένη κατανομή Cauchy. Το σύνολο δεδομένων που θεωρούμε αφορά τη

συναλλαγματική ισοτιμία της λίρας του Ηνωμένου Βασιλείου με το δολάριο των Η.Π.Α. από το 1800-2003. Τα δεδομένα αυτά είναι διαθέσιμα στον ιστότοπο <http://www.globalfinancialdata.com/> και παρουσιάζονται π.χ. στον Πίνακα 1 της εργασίας των Jamalizadeh et al. (2009), ενώ δίνονται έμμεσα και στο Παράρτημα Β της μεταπτυχιακής διατριβής, όπου παρατίθεται ο κώδικας της Matlab για την υλοποίηση όσων ακολουθούν.

Τα αρχικά δεδομένα ακολουθώντας μια πάγια τακτική, μέθοδο χειρισμού δεδομένων που αφορούν συναλλαγματικές ισοτιμίες, μετασχηματίζονται χρησιμοποιώντας το μετασχηματισμό του λογαρίθμου και έστω  $Z_t$  τα δεδομένα που προκύπτουν για  $t=1, \dots, 204$ . Έπειτα θεωρούμε τα δεδομένα  $X_t = Z_t - Z_{t-1}$ ,  $t=1, \dots, 203$ , στα οποία θα εφαρμόσουμε όσα προηγούμενα αναφέρθηκαν.

Λαμβάνοντας υπόψη τα αποτελέσματα της ενότητας 3.1 τα δεδομένα επιπρόσθετα μετασχηματίζονται στα

$$\tilde{X}_t = \frac{2X_t - (X_{(203)} + X_{(1)})}{(X_{(203)} - X_{(1)})}, t=1, \dots, 203,$$

όπου  $X_{(1)} = -0.4513$  και  $X_{(203)} = 0.4309$  η ελάχιστη και μέγιστη τιμή, αντίστοιχα, των  $X_t$ ,  $t=1, \dots, 203$ .

Προκύπτει τότε ότι οι Ε.Μ.Π.  $\tilde{\mu}$  και  $\tilde{\sigma}$  που προκύπτουν με τα μετασχηματισμένα δεδομένα  $\tilde{X}_t$  της πιο πάνω σχέσης είναι οι:  $\tilde{\mu}=0.0218$  και  $\tilde{\sigma}=0.0294$ , ενώ  $\log L=165.2519$ .

Χρησιμοποιώντας τις σχέσεις (3.1.10) και (3.1.11) έχουμε ότι οι Ε.Μ.Π. των παραμέτρων  $\mu$  και  $\sigma$  με τα αρχικά δεδομένα, δίνονται από τις σχέσεις:



$$\hat{\mu} = \frac{(X_{(n)} - X_{(1)})\tilde{\mu} + (X_{(n)} - X_{(1)})}{2} = -0.0024,$$

και

$$\hat{\sigma} = \frac{(X_{(n)} - X_{(1)})}{2} \tilde{\sigma} = 0.0519,$$

αντίστοιχα.

### 3.4 Μέθοδος Ροπών

Στην παράγραφο αυτή θα παρουσιαστούν τα αποτελέσματα που δόθηκαν από τον Nadarajah (2011) και αφορούν στην εκτίμηση των αγνώστων παραμέτρων με τη μέθοδο των ροπών.

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό με σ.π.π. που δίνεται από την σχέση (2.1.1) ή εναλλακτικά από τη σχέση (2.1.2). Εξισώνοντας τις τέσσερις πρώτες κεντρικές πληθυσμιακές ροπές, οι οποίες προκύπτουν με απλές αντικαταστάσεις στη σχέση (2.2.4), με τις αντίστοιχες δειγματικές ροπές οι εκτιμητές των παραμέτρων  $a, b, \mu$  και  $\sigma$  με τη μέθοδο των ροπών προκύπτουν από την επίλυση του ακόλουθου συστήματος εξισώσεων:

$$\frac{\sum_{k=0}^m \frac{1}{k+1} \binom{m}{k} \mu^{m-k} \sigma^k \Lambda_k}{D} = \frac{1}{n} \sum_{i=1}^n X_i^m.$$

## ΚΕΦΑΛΑΙΟ 4

### ΓΕΝΙΚΕΥΣΕΙΣ ΤΗΣ ΜΟΝΟΔΙΑΣΤΑΤΗΣ ΑΠΟΚΟΜΜΕΝΗΣ ΚΑΤΑΝΟΜΗΣ CAUCHY

Αντικείμενο του κεφαλαίου αυτού είναι μία σύντομη ανασκόπηση των σημαντικότερων γενικεύσεων της μονοδιάστατης αποκομμένης κατανομής Cauchy που έχουν πρόσφατα εισαχθεί στη βιβλιογραφία και έχουν εν μέρει μελετηθεί. Στο πλαίσιο αυτό θα παρουσιαστούν η αποκομμένη γενικευμένη Cauchy (Ateya and Al-Hussaini, 2012), η λοξή αποκομμένη (Nadarajah and Ali, 2004) και η αποκομμένη λοξή κατανομή Cauchy (Jamalizadeh *et al.*, 2009).

#### 4.1. Αποκομμένη γενικευμένη Cauchy

Στην ενότητα αυτή παρουσιάζεται η μονοδιάστατη από αριστερά αποκομμένη γενικευμένη κατανομή Cauchy, η οποία προκύπτει με εισαγωγή της σ.π.π. και της α.σ.κ. της γενικευμένης κατανομής Cauchy στη σχέση (1.2.5). Επιπρόσθετα, θα παρουσιαστούν οι βασικές της ιδιότητες, θα αναφερθούν οι κυριότερες εφαρμογές της και θα εκτιμηθούν οι άγνωστες παράμετροί της.

##### 4.1.1 Η κατανομή και οι ιδιότητές της

Η γενικευμένη Κατανομή Cauchy αποτελεί μία από τις γενικεύσεις της κατανομής Cauchy που έχουν εισαχθεί στη βιβλιογραφία και χρησιμοποιούνται σε πρακτικές εφαρμογές. Η κατανομή αυτή πρωτοπαρουσιάστηκε στη βιβλιογραφία στην εργασία του Rider (1957). Στη συνέχεια και για λόγους πληρότητας δίνεται ο ορισμός της γενικευμένης κατανομής Cauchy.

**Ορισμός 4.1.1.1 (Rider, 1957)**

Η τ.μ.  $X$  λέμε ότι ακολουθεί γενικευμένη κατανομή Cauchy με παραμέτρους  $\mu$ ,  $\delta$  και  $\omega$  αν η σ.π.π. της δίνεται από τη σχέση:

$$g_X(x; \mu, \delta, \omega) = \frac{\delta \Gamma(\omega)}{2\Gamma(1/\delta)\Gamma(\omega-1/\delta)} \left[ 1 + |x - \mu|^\delta \right]^{-\omega}, \quad (4.1.1)$$

όπου  $-\infty < x < \infty$ ,  $-\infty < \mu < \infty$ ,  $\delta, \omega > 0$  και  $\delta\omega > 1$ . Τότε γράφουμε ότι  $X \sim GC(\mu, \delta, \omega)$ .

Εύκολα προκύπτει ότι για  $\delta = 2$  και  $\omega = 1$ , η σχέση (4.1.1) δίνει την σ.π.π. της κατανομής Cauchy, ενώ για  $\delta = 2$ ,  $x - \mu = t/\sqrt{\nu}$  και  $\omega = (\nu + 1)/2$  δίνει την σ.π.π. της  $t$  κατανομής με  $\nu$  βαθμούς ελευθερίας.

Οι Ateya and Al-Hussaini (2012) παρουσιάζουν και μελετούν την ειδική περίπτωση της από αριστερά αποκομμένης στο σημείο  $\mu$  γενικευμένης κατανομής Cauchy, που προκύπτει από τη γενικευμένη Cauchy για  $\delta = 2$  και  $\omega = \alpha + 1/2$ , με  $\alpha > 0$ . Δηλαδή, θεωρούν την από αριστερά αποκομμένη μορφή μιας τυχαίας μεταβλητής  $Y$  με σ.π.π.

$$g_Y(y; \mu, \alpha) = \frac{\Gamma(\alpha + 1/2)}{\sqrt{\pi}\Gamma(\alpha)} \left[ 1 + (y - \mu)^2 \right]^{-(\alpha + 1/2)},$$

και α.σ.κ.

$$G_Y(y; \mu, \alpha) = \frac{\Gamma(\alpha + 1/2)}{\sqrt{\pi}\Gamma(\alpha)} \int_0^{\tan^{-1}(y-\mu)} (\cos \phi)^{2\alpha-1} d\phi. \quad (4.1.2)$$

Τότε, η μονοδιάστατη από αριστερά αποκομμένη γενικευμένη κατανομή Cauchy προκύπτει αρχικά με την εισαγωγή της σ.π.π. και της α.σ.κ. της γενικευμένης κατανομής Cauchy στη σχέση (1.2.5) και έπειτα εισάγοντας σε αυτήν την παράμετρο κλίμακας  $\sigma$  καταλήγουμε στη σ.π.π. που ορίστηκε από τους Ateya and Al-Hussaini (2012).

#### Ορισμός 4.1.1.2 (Ateya and Al-Hussaini, 2012)

Η τ.μ.  $X$  λέμε ότι ακολουθεί την αριστερά αποκομμένη στο σημείο  $\mu$  γενικευμένη κατανομή Cauchy με παραμέτρους  $\mu, \sigma$  και  $\alpha$ , αν η σ.π.π. της είναι:

$$f_X(x; \mu, \alpha) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha + 1/2)}{\sigma \Gamma(\alpha)} \left[ 1 + \left( \frac{x - \mu}{\sigma} \right)^2 \right]^{-\alpha - 1/2}, \quad (4.1.3)$$

όπου  $x \geq \mu$  και  $\mu \in R, \sigma > 0, \alpha > 0$ . Τότε γράφουμε ότι  $X \sim {}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$ .

Εύκολα προκύπτει ότι αν  $\alpha = 0.5$ , η  $f_X(\cdot)$  παίρνει τη μορφή της σ.π.π. της από αριστερά αποκομμένης κατανομής Cauchy, που περιγράφεται στη σχέση (2.1.5) με σημείο αποκοπής  $a = \mu$ .

Στην πρόταση που ακολουθεί δίνεται η α.σ.κ. της τ.μ.  $X$  η οποία ακολουθεί την  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$ . Η απόδειξη προκύπτει άμεσα από τη σχέση (1.2.6), όπου προσδιορίζεται η α.σ.κ. μιας από αριστερά αποκομμένης κατανομής, με αντικατάσταση της α.σ.κ. της γενικευμένης Cauchy που δίνεται στη σχέση (4.1.2) και λαμβάνοντας υπόψη ότι  $G_X(\mu) = 0$ .

**Πρόταση 4.1.1.1**

Έστω  $X \sim {}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$ , με  $\mu \in R, \sigma > 0, \alpha > 0$ . Τότε η α.σ.κ. της τ.μ.

$X$  ορίζεται από τη σχέση:

$$F_X(x) = \begin{cases} 0 & , \text{για } x < \mu \\ \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \int_0^{\tan^{-1}\left(\frac{x-\mu}{\sigma}\right)} (\cos \phi)^{2\alpha-1} d\phi & , \text{για } \mu \leq x \end{cases} .$$

Οι ιδιότητες της  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$  έχουν μελετηθεί από τους Ateya and Al-Hussaini (2012). Στη συνέχεια παρουσιάζονται αυτές οι ιδιότητες.

**Σχήμα και κορυφότητα της  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$** 

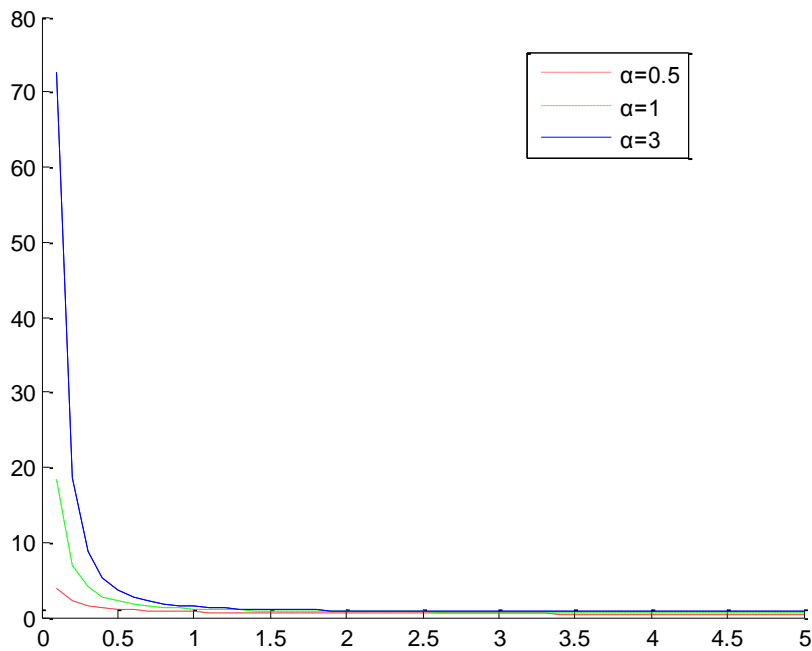
Σε αυτήν την παράγραφο, λαμβάνοντας υπόψη την εργασία των Ateya and Al-Hussaini (2012), θα εξεταστεί το σχήμα και η κορυφότητα της  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$ . Καθώς η σ.π.π. της  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$  είναι γνησίως φθίνουσα στο διάστημα  $[\mu, \infty)$ , εύκολα προκύπτουν τα ακόλουθα.

**Πρόταση 4.1.1.2 (Ateya and Al-Hussaini, 2012)**

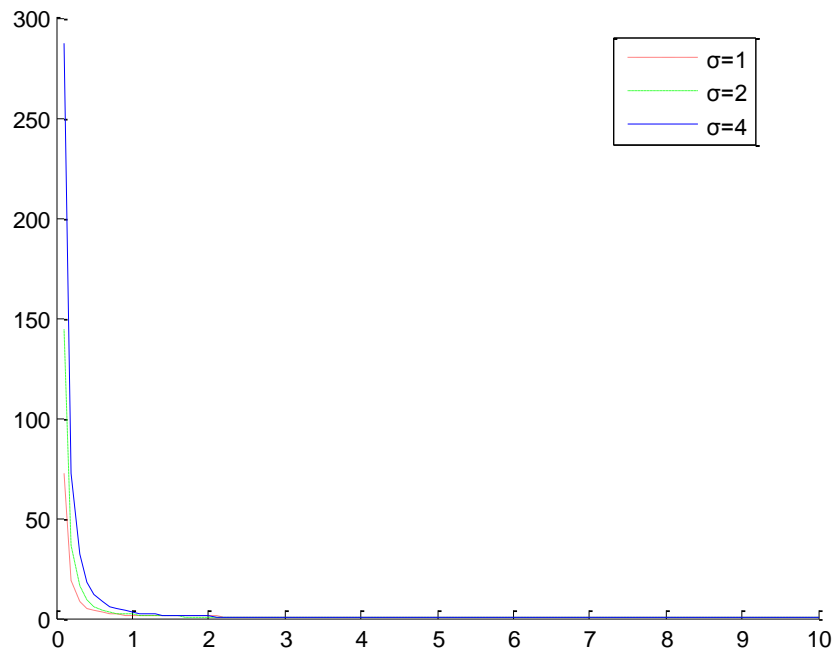
Έστω  $X$  μια τ.μ. που ακολουθεί την  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$ . Η σ.π.π. της τ.μ.  $X$  είναι μονοκόρυφη. Ειδικότερα προκύπτει ότι η κορυφή είναι στο  $x_0 = \mu$  με  $f_X(\mu) = a\Gamma(\alpha+1/2)/[\sqrt{\pi}\sigma\Gamma(\alpha)]$ .

Από την παραπάνω πρόταση γίνεται αντιληπτό ότι η κορυφότητα της δεν επηρεάζεται από την τιμή των παραμέτρων  $\sigma$  και  $\alpha$ . Στη συνέχεια, στα

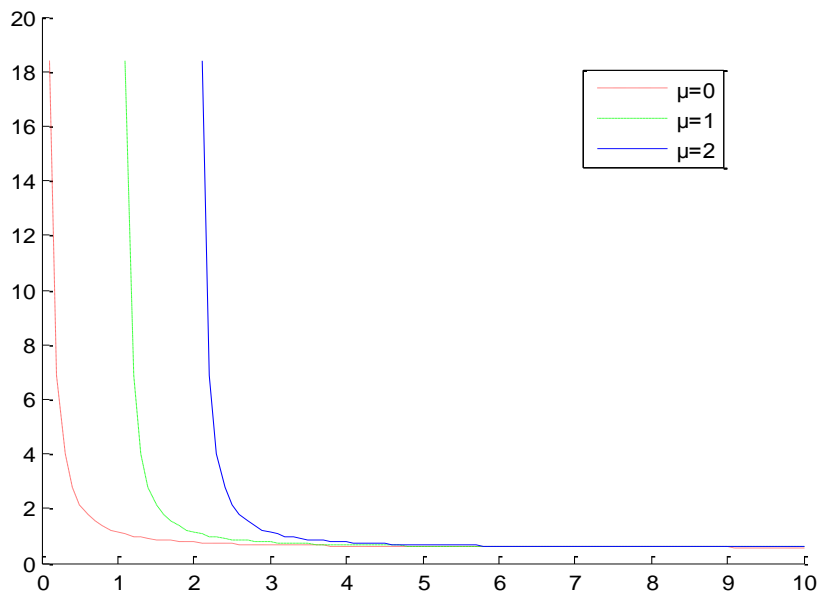
γραφήματα που ακολουθούν δίνεται η γραφική παράσταση της σ.π.π. της  ${}_L TGC^{(\mu)}(\mu, \sigma, \alpha)$  για διάφορες τιμές των παραμέτρων.



**Γράφημα 4-** Σ.π.π. της αριστερά αποκομμένης γενικευμένης κατανομής Cauchy για  $x \in [\mu, 5]$ ,  
όταν  $\mu = 0$ ,  $\sigma = 1$  και  $a \in \{0.5, 1, 1.5\}$ .



**Γράφημα 5-** Σ.π.π. της αριστερά αποκομμένης γενικευμένης κατανομής Cauchy για  $x \in [\mu, 10]$ , όταν  $\mu = 0$ ,  $\alpha = 1.5$  και  $\sigma \in \{1, 2, 4\}$ .



**Γράφημα 6-** Σ.π.λ. της αριστερά αποκομμένης γενικευμένης Κατανομής Cauchy για  $x \in [\mu, 10]$ , όταν  $\sigma = 1$ ,  $\alpha = 1$  και  $\mu \in \{0, 1, 2\}$ .

#### Ροπογεννήτρια και χαρακτηριστική συνάρτηση της ${}_{L}TGC^{(\mu)}(\mu, \sigma, \alpha)$

Σε αυτήν την ενότητα, λαμβάνοντας υπόψη την εργασία των Ateya and Al-Hussaini (2012), θα δοθεί η ργ.σ. και χ.σ. της  ${}_{L}TGC^{(\mu)}(\mu, \sigma, \alpha)$ .

#### Πρόταση 4.1.1.3 (Ateya and Al-Hussaini, 2012)

Έστω  $X$  μια τ.μ. που ακολουθεί την  ${}_{L}TGC^{(\mu)}(\mu, \sigma, \alpha)$ . Η ροπογεννήτρια και η χαρακτηριστική συνάρτηση της  $X$ , δίνονται από τις σχέσεις:



$$M_x(t) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \int_0^{\pi/2} \exp[(\mu + \sigma \tan \phi)t] (\cos \phi)^{2\alpha-1} d\phi, \quad (4.1.4)$$

και

$$\varphi_x(t) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \int_0^{\pi/2} \exp[(\mu + \sigma \tan \phi)it] (\cos \phi)^{2\alpha-1} d\phi,$$

αντίστοιχα, όπου  $x \geq \mu$ ,  $\mu \in R$ ,  $\sigma > 0$ ,  $\alpha > 0$  και  $i$  είναι η φανταστική μονάδα.

### Απόδειξη

Με βάση τον ορισμό της ροπογεννήτριας συνάρτησης  $M(t) = E[\exp(tX)]$  και αντικαθιστώντας τον τύπο της  $f_x(x)$  από τη σχέση (4.1.3) έχουμε ότι

$$M(t) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\sigma \Gamma(\alpha)} \int_{\mu}^{\infty} \exp(tx) \left\{ 1 + \left( \frac{x-\mu}{\sigma} \right)^2 \right\}^{-\alpha-1/2} dx.$$

Θεωρούμε έπειτα το μετασχηματισμό  $z = (x - \mu)/\sigma$ , οπότε

$$M(t) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \int_0^{\infty} \exp[(\mu + \sigma z)t] (1+z^2)^{-\alpha-1/2} dz.$$

Θεωρώντας τον μετασχηματισμό  $\phi = \tan^{-1}(z)$  και λαμβάνοντας υπόψη ότι

$$\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}, \quad \frac{d}{dx} \tan x = \frac{1}{\cos^2 x} \quad \text{και} \quad \tan^{-1}(\infty) = \pi/2,$$

αλγεβρικές πράξεις προκύπτει ότι:

$$M_x(t) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \int_0^{\pi/2} \exp[(\mu + \sigma \tan \phi)t] (\cos \phi)^{2\alpha-1} d\phi.$$

Η χ.σ. προκύπτει με παρόμοιο τρόπο λαμβάνοντας υπόψη ότι  $\varphi(t) = E[\exp(itX)]$ . ■

**Παρατήρηση 4.1.1.1 α)** Από τη σχέση (4.1.4) στην ειδική περίπτωση που  $r = 1, 2, \dots$ , τέτοιο ώστε  $r < 2\alpha$ , προκύπτει ότι:

$$E(X^r) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)} \sum_{i=0}^r \binom{r}{i} \sigma^i \mu^{r-i} B\left(\alpha - \frac{i}{2}, \frac{i+1}{2}\right), \quad (4.1.5)$$

$$\text{όπου } B(a, b) = \int_0^1 z^{a-1} (1-z)^{b-1} dz.$$

Συγκεκριμένα για  $r = 1$ , η μέση τιμή της  $LTGC^{(\mu)}(\mu, \sigma, \alpha)$  δίνεται από τη σχέση:

$$E(X) = \mu + \frac{\sigma \Gamma(\alpha-1/2)}{\sqrt{\pi}\Gamma(\alpha)}, \quad 1 < 2\alpha \quad (4.1.6)$$

ενώ για  $r = 2$ , η δεύτερη ροπή της κατανομής είναι,

$$E(X^2) = \mu^2 + 2\sigma\mu \frac{\Gamma(\alpha-1/2)}{\sqrt{\pi}\Gamma(\alpha)} + \frac{\sigma^2}{2(\alpha-1)}, \quad 1 < \alpha.$$

Λαμβάνοντας υπόψη τις εκφράσεις των  $E(X)$  και  $E(X^2)$  προκύπτει ότι:

$$\text{Var}(X) = \sigma^2 \left\{ \frac{1}{2(\alpha-1)} - \left[ \frac{\Gamma(\alpha-1/2)}{\sqrt{\pi}\Gamma(\alpha)} \right]^2 \right\}, \quad 1 < \alpha.$$

**β)** Είναι φανερό ότι για συγκεκριμένες τιμές των παραμέτρων η μέση τιμή και η διακύμανση μπορούν εύκολα να υπολογιστούν. Το ίδιο ισχύει και για την τρίτη και τέταρτη τάξης ροπή. Αυτό οδηγεί στο συμπέρασμα ότι είναι εύκολος και ο προσδιορισμός των συντελεστών λοξότητας και κύρτωσης, όταν  $1.5 < \alpha$  και  $2 < \alpha$ , αντίστοιχα. Τέλος, η τιμή της διακύμανσης δεν εξαρτάται από την τιμή της παραμέτρου  $\mu$ .

#### 4.1.2 Εκτίμηση των παραμέτρων

Στην παράγραφο αυτή, σύμφωνα με την εργασία των Ateya and Al-Hussaini (2012) παρουσιάζεται η εκτίμηση των παραμέτρων της  $IGC^{(\mu)}(\mu, \sigma, \alpha)$  με τη μέθοδο της μέγιστης πιθανοφάνειας. Για την εύρεση των Ε.Μ.Π. απαιτείται η μεγιστοποίηση της συνάρτησης πιθανοφάνειας ως προς τις άγνωστες παραμέτρους  $\mu, \sigma$  και  $\alpha$ .

Έστω  $X_1, \dots, X_n$  ένα τυχαίο δείγμα από πληθυσμό με κατανομή που δίνεται από την (4.1.3). Η συνάρτηση πιθανοφάνειας είναι η εξής:

$$L_n(\mu, \sigma, \alpha) = \left( \frac{2}{\sqrt{\pi}} \frac{\Gamma(\alpha + 1/2)}{\sigma \Gamma(\alpha)} \right)^n \frac{1}{\prod_{i=1}^n \left[ \sigma^2 + (x_i - \mu)^2 \right]^{\alpha + 1/2}}, \quad (4.1.2.1)$$

$0 < \mu \leq X_{(1)} \leq X_{(n)} < \infty$ , όπου  $X_{(1)} = \min_i X_i$  και  $X_{(n)} = \max_i X_i$  είναι το ελάχιστο και το μέγιστο διατεταγμένο στατιστικό.

Για την εύρεση των Ε.Μ.Π. απαιτείται η μεγιστοποίηση της συνάρτησης πιθανοφάνειας ως προς τις άγνωστες παραμέτρους. Η συνάρτηση πιθανοφάνειας είναι γνησίως αύξουσα ως προς την παράμετρο θέσης  $\mu$ . Επομένως, μεγιστοποιείται στο άνω άκρο του πεδίου ορισμού της παραμέτρου  $\mu$  που είναι το διατεταγμένο στατιστικό  $X_{(1)}$ . Άρα, ο Ε.Μ.Π. της παραμέτρου  $\mu$  είναι το διατεταγμένο στατιστικό  $X_{(1)}$ . Για τον προσδιορισμό των Ε.Μ.Π. των υπόλοιπων παραμέτρων αρκεί να μεγιστοποιηθεί η συνάρτηση πιθανοφάνειας ή ο λογάριθμος

της συνάρτησης πιθανοφάνειας που προκύπτει αφού αντικατασταθεί στη (4.1.2.1) η παράμετρος  $\mu$  από τον Ε.Μ.Π. της  $\hat{\mu}$ .

Επομένως απαιτείται η μεγιστοποίηση της

$$\ln L_n(\hat{\mu}, \sigma, \alpha) = n \ln(2) - \frac{n}{2} \ln(\pi) + n \ln \Gamma(\alpha + 1/2) - n \ln(\Gamma(\alpha)) - n \ln(\sigma) - (\alpha + 1/2) \sum_{i=1}^n \ln \left[ 1 + \left( \frac{x_i - \hat{\mu}}{\sigma} \right)^2 \right]. \quad (4.1.2.2)$$

ως προς  $\sigma$  και  $\alpha$ . Οι Ε.Μ.Π. των παραμέτρων προσδιορίζονται από την επίλυση του συστήματος εξισώσεων που προκύπτει θέτοντας τις μερικές παραγώγους της παραπάνω συνάρτησης ως προς  $\sigma$  και  $\alpha$  ίσες με το μηδέν.

Τότε προκύπτει το ακόλουθο σύστημα εξισώσεων:

$$\begin{cases} -n/\hat{\sigma} + (2\hat{\alpha} + 1) \sum_{i=1}^n (x_i - \hat{\mu})^2 / \left\{ \hat{\sigma} \left[ \hat{\sigma}^2 + (x_i - \hat{\mu})^2 \right] \right\} = 0 \\ -n\psi(\hat{\alpha}) + n\psi(\hat{\alpha} + 1/2) - \sum_{i=1}^n \ln \left[ 1 + \left( [x_i - \hat{\mu}] / \hat{\sigma} \right)^2 \right] = 0 \end{cases} \quad (4.1.2.3)$$

όπου με  $\psi(\cdot)$  συμβολίζεται η συνάρτηση που ορίζεται  $\psi(z) = \frac{d}{dz} \ln \Gamma(z)$ . Οι

εξισώσεις (4.1.2.3) παριστάνουν δύο μη γραμμικές εξισώσεις, οι οποίες μπορούν να επιλυθούν με αριθμητικές μεθόδους όπως η Newton-Raphson.

Είναι ευνόητο ότι στη βιβλιογραφία δεν έχει εμφανιστεί, μέχρι την ώρα συγγραφής αυτής της μεταπτυχιακής διατριβής, μελέτη για την ύπαρξη των Ε.Μ.Π. αυτής της κατανομής και παραμένει ένα ανοικτό προς έρευνα θέμα.

**Παρατήρηση 4.1.2.1** Στο σημείο αυτό θα ήταν παράλειψη να μην αναφερθεί ότι οι Ateya and Al-Hussaini (2012) εκτός από την εκτίμηση με την μέθοδο της μέγιστης πιθανοφάνειας προσδιόρισαν τους εκτιμητές των άγνωστων παραμέτρων

με μεθόδους της Μπευζιανής Στατιστικής. Στο πλαίσιο αυτό θεώρησαν ότι η συνάρτηση απώλειας είναι είτε το τετραγωνικό σφάλμα, οπότε ο εκτιμητής Bayes είναι η μέση τιμή της εκ των υστέρων κατανομής, είτε ότι η συνάρτηση απώλειας είναι η συνάρτηση απώλειας που είναι γνωστή ως γραμμικοεκθετική (linear-exponential) LINEX (βλέπε Varian, 1975). Επιπρόσθετα, για τον υπολογισμό της εκ των υστέρων κατανομής υπέθεσαν τόσο ότι οι εκ των προτέρων κατανομές είναι Γάμμα όσο και ότι είναι μη πληροφοριακές. Τα αποτελέσματα αυτά δεν παρουσιάζονται σε αυτήν τη μεταπτυχιακή διατριβή, καθώς περιοριζόμαστε στην ανασκόπηση και παρουσίαση μεθόδων εκτίμησης στα πλαίσια της κλασικής Στατιστικής Συμπερασματολογίας.

#### 4.2 Λοξές γενικεύσεις

Οι λοξές γενικεύσεις κατανομών άρχισαν να μελετώνται ευρέως τα τελευταία χρόνια αποτελώντας γενικεύσεις υπάρχουσών συμμετρικών κατανομών για τη μοντελοποίηση τυχαιών πραγματικών φαινομένων στα οποία παρατηρείται λοξότητα. Στην ενότητα αυτή θα παρουσιαστούν δύο μη ισοδύναμοι τρόποι ορισμού μιας λοξής κατανομής Cauchy. Οι κατανομές που θα προκύψουν θα γενικευτούν με αποκοπή αυτών.

Στη βιβλιογραφία έχουν προταθεί διάφοροι τρόποι επίτευξης κατασκευής μίας λοξής κατανομής από μία συμμετρική. Ο κυριότερος από αυτούς δόθηκε από τον Azzalini (1985), και δίνεται στο επόμενο λήμμα.

##### Λήμμα 4.2.1 (Azzalini, 1985)

Έστω  $Y_1$  μια τ.μ. με σ.π.π.  $f(\cdot)$  συμμετρική γύρω από το μηδέν και  $Y_2$  μια τ.μ. με απόλυτα συνεχή α.σ.κ.  $G(\cdot)$  τέτοια ώστε η  $G'(\cdot)$  να είναι συμμετρική γύρω από το μηδέν. Τότε για κάθε πραγματικό αριθμό  $\lambda$  η

$$g(y/\lambda) = 2f(y)G(\lambda y), -\infty < y < \infty, \quad (4.2.1)$$

είναι σ.π.π. μιας τυχαίας μεταβλητής  $Y$ .

Οι ιδιότητες της τ.μ. με σ.π.π. αυτή της σχέσης (4.2.1) μελετήθηκαν μεταξύ άλλων από τους Azzalini (1986) και Henze (1986). Η ειδική περίπτωση όπου  $f(\cdot) = \varphi(\cdot)$  και  $G(\cdot) = \Phi(\cdot)$ , με  $\varphi(\cdot)$  και  $\Phi(\cdot)$  τη σ.π.π. και την α.σ.κ. της τυπικής κανονικής κατανομής αντιστοιχεί στη λοξή κανονική κατανομή, η οποία ορίστηκε από τον Azzalini (1985).

Στη συνέχεια, για λόγους πληρότητας, δίνεται ο ορισμός της λοξής κανονικής κατανομής. Για μια κριτική ανασκόπηση και μελέτη των ιδιοτήτων της λοξής κανονικής κατανομής παραπέμπουμε στη μεταπτυχιακή διατριβή της Σωκράτους (2009).

#### Ορισμός 4.2.1

Αν  $\varphi$  και  $\Phi$  είναι αντίστοιχα η σ.π.π. και η α.σ.κ. της τυπικής κανονικής κατανομής  $N(0,1)$  τότε η τυχαία μεταβλητή  $Z_\lambda$  λέμε ότι ακολουθεί λοξή κανονική κατανομή με παράμετρο  $\lambda$  αν η σ.π.π. της δίνεται από τη σχέση:

$$g_{Z_\lambda}(z/\lambda) = 2\varphi(z)\Phi(\lambda z), -\infty < z < \infty,$$

και συμβολίζεται  $Z_\lambda \sim SN(\lambda)$ .

Στο Λήμμα που ακολουθεί δίδονται κάποιες σημαντικές και ενδιαφέρουσες ιδιότητες της τ.μ.  $Y$  με σ.π.π. αυτή της σχέσης (4.2.1).

**Λήμμα 4.2.2 (Gupta et al., 2002)**

Έστω ότι  $Y_1$ ,  $Y_2$  και  $Y$  είναι τρεις τ.μ. όπως αυτές ορίστηκαν στο Λήμμα 4.2.1. Τότε οι ζυγής τάξης ροπές της  $Y$  είναι ανεξάρτητες από την παράμετρο  $\lambda$  και είναι ίδιες με αυτές της τ.μ.  $Y_1$ . Επιπλέον, οι τ.μ.  $Y_1^2$  και  $Y^2$  ακολουθούν την ίδια κατανομή.

**Απόδειξη**

Έστω  $\psi_Y(t)$  η χ.σ. της τ.μ.  $Y$ . Τότε είναι

$$\begin{aligned}\psi_Y(-t) &= \int_{-\infty}^{\infty} e^{-ity} 2f(y)G(\lambda y)dy \stackrel{x=-y}{=} \int_{-\infty}^{\infty} -e^{itx} [2f(-x)G(-\lambda x)]dx \\ &\stackrel{\text{συμμετρία}}{=} \int_{-\infty}^{\infty} e^{itx} 2f(x)[1-G(\lambda x)]dx.\end{aligned}$$

Επομένως  $g(t) = \psi_Y(t) + \psi_Y(-t) = 2\psi_{Y_1}(t)$  είναι ανεξάρτητη της παραμέτρου  $\lambda$ . Επιπλέον προκύπτει ότι αν  $(-1)^n g^{(n)}(0)/2$  και  $(-1)^n \psi_{Y_1}^{(2n)}(t)$  υπάρχουν τότε είναι ίδιες. Τέλος, καθώς αυτές είναι οι  $2n$  ροπές των  $Y$  και  $Y_1$  και οι  $n$  ροπές των  $Y^2$  και  $Y_1^2$ , αντίστοιχα, προκύπτει το επιθυμητό αποτέλεσμα. ■

Η μονοδιάστατη λοξή τυπική κατανομή Cauchy, προκύπτει από τη σχέση (4.2.1) αν λάβουμε υπόψη τη σ.π.π. και την α.σ.κ. της τυπικής κατανομής Cauchy, όπως αυτές δίνονται στις σχέσεις (1.1.1) και (1.1.2), αντίστοιχα, για  $\mu = 0$  και  $\sigma = 1$ . Ειδικότερα, προκύπτει ο ακόλουθος ορισμός.

**Ορισμός 4.2.2 (Gupta et al., 2002)**

Λέμε ότι η τ.μ.  $X$  ακολουθεί λοξή τυπική κατανομή Cauchy με συντελεστή λοξότητας  $\lambda$  αν η σ.π.π. της δίνεται από τη σχέση:

$$f_x(x) = \frac{1}{\pi(1+x^2)} \left[ 1 + \frac{2}{\pi} \tan^{-1}(\lambda x) \right] \quad (4.2.2)$$

όπου  $x \in R$  και  $\lambda \in R$ .

Στη βιβλιογραφία έχουν εισαχθεί διάφοροι τρόποι ορισμού ισοδύναμων ή μη λοξών μονοδιάστατων κατανομών Cauchy με αυτή που δόθηκε παραπάνω (βλέπε Gupta *et al.*, 2002 και Huang and Chen, 2007). Στη συνέχεια, θα παρουσιαστεί ένας από αυτούς, καθώς χρησιμοποιώντας την κατανομή προκύπτει, θα οριστεί σε επόμενη παράγραφο, μια ακόμη αποκομμένη κατανομή Cauchy.

Ένας άλλος τρόπος ορισμού της λοξής κατανομής Cauchy, ο οποίος δεν οδηγεί σε ισοδύναμο ορισμό, είναι αυτός που προτάθηκε από τους Azzalini and Capitanio (2003) στην πολυδιάστατη περίπτωση. Ειδικότερα, εφαρμόζοντας την πρόταση των Azzalini and Capitanio (2003) για τη μονοδιάστατη περίπτωση, η λοξή κατανομή Cauchy ορίζεται ως η κατανομή της τ.μ.  $W_\lambda$  που ακολουθεί την ίδια κατανομή με το πηλίκο  $Z_\lambda / |Y|$ , όπου  $Z_\lambda \sim SN(\lambda)$  και  $Y \sim N(0,1)$ , με  $Z_\lambda$  και  $Y$  ανεξάρτητες. Κάποιες ιδιότητες αυτής μελετήθηκαν μεταξύ άλλων από τους Behboodian *et al.* (2006), οι οποίοι προσδιόρισαν και την σ.π.π. και την α.σ.κ. αυτής της λοξής κατανομής, την οποία για να τη διαχωρίζουμε από αυτή του Ορισμού 4.2.2 θα την αποκαλούμε λοξή AC (από τα αρχικά των επιθέτων Azzalini and Capitanio) κατανομή Cauchy. Τα αποτελέσματα αυτά παρατίθενται στη συνέχεια.

#### **Ορισμός 4.2.3 (Behboodian *et al.*, 2006)**

Λέμε ότι η τυχαία μεταβλητή  $X_\lambda$  ακολουθεί τη λοξή AC κατανομή Cauchy με παράμετρο  $\lambda \in R$  αν η σ.π.π. της δίνεται από τη σχέση:



$$g(x; \lambda) = \frac{1}{\pi(1+x^2)} \left[ 1 + \frac{\lambda x}{\sqrt{1+(1+\lambda^2)x^2}} \right], \quad (4.2.3)$$

όπου  $x, \lambda \in R$ .

Στην πρόταση που ακολουθεί δίνεται η α.σ.κ. της λοξής AC κατανομής Cauchy του Ορισμού 4.2.3.

**Πρόταση 4.2.1 (Behboodian *et al.*, 2006)**

Αν η τυχαία μεταβλητή  $X_\lambda$  ακολουθεί τη λοξή AC κατανομή Cauchy με παράμετρο  $\lambda \in R$ , η α.σ.κ. της δίνεται από τη σχέση:

$$G(x; \lambda) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left[ \frac{x\sqrt{1+(1+\lambda^2)x^2} - \lambda}{\sqrt{1+(1+\lambda^2)x^2} + \lambda x} \right], \quad (4.2.4)$$

όπου  $x, \lambda \in R$ .

Η λοξή κατανομή Cauchy και η λοξή AC κατανομή Cauchy, που προηγούμενα παρουσιάστηκαν, χρησιμεύουν ως βάση για την δημιουργία νέων αποκομμένων κατανομών. Η μελέτη αυτών των κατανομών είναι το αντικείμενο των επόμενων δύο ενοτήτων.

**4.2.1 Λοξή αποκομμένη Κατανομή Cauchy**

Στην ενότητα αυτή παρουσιάζεται η μονοδιάστατη λοξή διπλά αποκομμένη κατανομή Cauchy, η οποία παρουσιάστηκε και μελετήθηκε από τον Nadarajah και τους συνεργάτες του σε τρεις εργασίες (βλέπε Nadarajah and Ali (2004), Nadarajah and Gupta (2005), Nadarajah and Kotz (2007 a)). Η κατανομή αυτή προκύπτει με εισαγωγή της σ.π.π. και της α.σ.κ. της συμμετρικά διπλά

αποκομμένης τυπικής κατανομής Cauchy (βλέπε (2.1.4) και (2.1.8), αντίστοιχα) στη σχέση (4.2.1) του Λήμματος 4.2.1.

Ειδικότερα, σε όσα ακολουθούν υποθέτουμε, χωρίς βλάβη της γενικότητας, ότι  $\lambda \geq 0$ , καθώς για  $\lambda < 0$  μπορούμε να χρησιμοποιούμε την ιδιότητα  $G(\lambda x) = 1 - G(-\lambda x)$ , όπου  $G(\cdot)$  η α.σ.κ. της  $TC^{(-b,b)}(0,1)$ . Τότε προκύπτει ο ακόλουθος ορισμός.

#### Ορισμός 4.2.1.1 (Nadarajah and Ali, 2004)

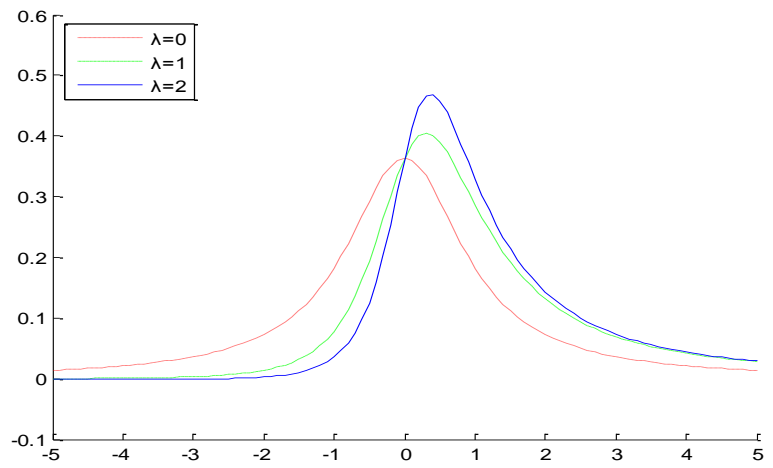
Λέμε ότι η τυχαία μεταβλητή  $X$  ακολουθεί τη λοξή διπλά συμμετρικά αποκομμένη τυπική κατανομή Cauchy, στο διάστημα  $[-b, b]$ , όπου  $b > 0$  αν η σ.π.π. της  $X$  δίνεται από τη σχέση:

$$g_x(x) = \frac{1}{2 \tan^{-1}(b)(1+x^2)} \left\{ 1 + \frac{\tan^{-1}(\lambda x)}{\tan^{-1}(b)} \right\}, \quad (4.2.1.1)$$

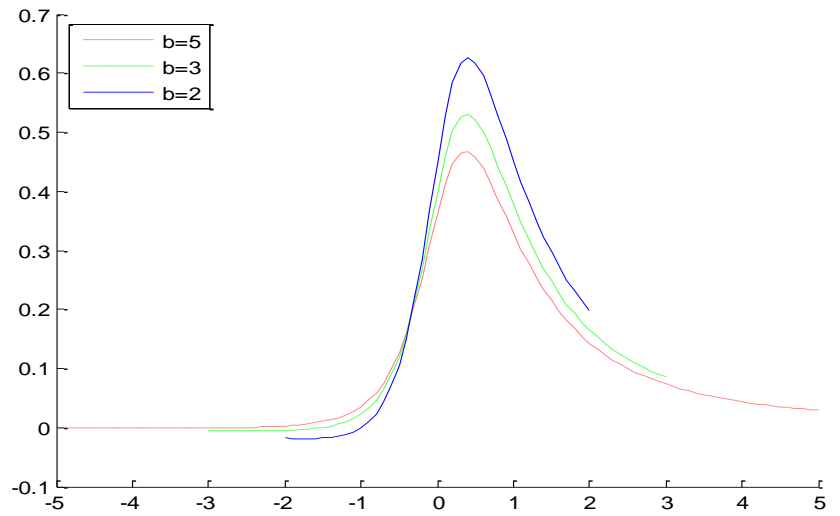
όπου  $-b \leq x \leq b$ ,  $b > 0$ , και  $\lambda \geq 0$ . Τότε γράφουμε  $X \sim STC^{(-b,b)}(\lambda)$ .

Παρατηρούμε ότι για  $\lambda = 0$ , η σχέση (4.2.1.1) ανάγεται στη σ.π.π. της  $TC^{(-b,b)}(0,1)$ .

Στη συνέχεια, στα γραφήματα που ακολουθούν δίνεται η γραφική παράσταση της σ.π.π. της  $STC^{(-b,b)}(\lambda)$  για διάφορες τιμές των παραμέτρων.



Γράφημα 7- Σ.π.π. της  $STC^{(-5,5)}(\lambda)$  όταν  $\lambda \in \{0, 1, 2\}$ .



Γράφημα 8- Σ.π.π. της  $STC^{(-b,b)}(2)$  όταν  $b \in \{2, 3, 5\}$ .

Στην πρόταση που ακολουθεί δίνεται η α.σ.κ. της τ.μ.  $X$  η οποία ακολουθεί την  $STC^{(-b,b)}(\lambda)$ .

**Πρόταση 4.2.1.1 (Nadarajah and Kotz, 2007 a)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $STC^{(-b,b)}(\lambda)$ . Τότε η α.σ.κ. της τ.μ.  $X$  ορίζεται από τη σχέση:

$$G_X(x) = \begin{cases} 0, & \text{για } x \leq -\delta \\ \frac{\tan^{-1}(\delta) + \tan^{-1}(x)}{2 \tan^{-1}(b)} - \frac{I(-x)}{2(\tan^{-1}(b))^2}, & \text{για } -\delta < x \leq 0 \\ \frac{\tan^{-1}(\delta) + \tan^{-1}(x)}{2 \tan^{-1}(b)} - \frac{I(x)}{2(\tan^{-1}(b))^2}, & \text{για } 0 < x \leq \delta \\ \frac{\tan^{-1}(x)}{\tan^{-1}(b)}, & \text{για } \delta < x < b \\ 1, & \text{για } x \geq b, \end{cases} \quad (4.2.1.2)$$

όπου

$$I(y) = \int_y^\delta \frac{\tan^{-1}(\lambda z)}{1+z^2} dz,$$

και  $-b \leq x \leq b$ , με  $b > 0$ , και  $\delta = \min(b/\lambda, b)$ .

Μία από τις σημαντικότερες ιδιότητες της λοξής συμμετρικά αποκομμένης τυπικής κατανομής Cauchy είναι ότι η αποκοπή της εξαλείφει την αποκλίνουσα συμπεριφορά της λοξής Cauchy ως προς τη μη ύπαρξη πεπερασμένων ροπών. Αντικείμενο μελέτης σε αυτήν την παράγραφο είναι ο υπολογισμός της  $n$  τάξης ροπής της λοξής συμμετρικά αποκομμένης κατανομής Cauchy.

**Θεώρημα 4.2.1.1 (Nadarajah and Kotz, 2007 a)**

Έστω  $X$  μια τ.μ. που ακολουθεί τη  $STC^{(-b,b)}(\lambda)$  με σ.π.π. που δίνεται από την σχέση (4.2.1.1). Τότε η  $n$  άρτιας τάξης ροπή της  $X$ , δίνεται από τη σχέση:

$$E(X^n) = \frac{b^{n+1}}{(n+1)\tan^{-1}(b)} {}_2F_1\left(1, \frac{n+1}{2}; 1 + \frac{n+1}{2}; -b^2\right), \quad (4.2.1.3)$$

ενώ η  $n$  περιττής τάξης ροπή της  $X$ , δίνεται από τη σχέση:

$$E(X^n) = \frac{b^{n+1} {}_2F_1\left(1, \frac{n+1}{2}; 1 + \frac{n+1}{2}; -b^2\right) - \delta^{n+1} {}_2F_1\left(1, \frac{n+1}{2}; 1 + \frac{n+1}{2}; -\delta^2\right)}{(n+1)\tan^{-1}(b)}$$

$$- \frac{J_n(\delta)}{(\tan^{-1}(b))^2} \quad (4.2.1.4)$$

όπου  $b > 0$ ,  $\delta = \min(b/\lambda, b)$  και  $J_n(y) = \int_0^y \frac{z^n \tan^{-1}(\lambda z)}{1+z^2} dz$ .

**Απόδειξη**

Από το Λήμμα 4.2.2 προκύπτει ότι η  $n$  άρτιας τάξης ροπή της τ.μ.  $X$  ισούται με τη  $n$  τάξης ροπή της  $X \sim TC^{(-b,b)}(0,1)$ . Επομένως, η  $n$  άρτιας τάξης ροπή της τ.μ.  $X$  προκύπτει από τη σχέση (2.2.4) για  $\mu = 0$  και  $\sigma = 1$ .

Έστω τώρα  $n$  περιττός. Με βάση τον ορισμό, η  $n$  τάξης ροπή της τ.μ.  $X$  με σ.π.π. που δίνεται από την σχέση (4.2.1.1), είναι:

$$\begin{aligned}
E(X^n) &= \frac{1}{2 \tan^{-1}(b)} \int_{\delta}^b \frac{y^n}{1+y^2} dy + \frac{1}{2(\tan^{-1}(b))^2} \int_{\delta}^b \frac{y^n \tan^{-1}(\lambda y)}{1+y^2} dy \\
&= \frac{1}{2 \tan^{-1}(b)} \left[ \int_0^b \frac{y^n}{1+y^2} dy - \int_0^{\delta} \frac{y^n}{1+y^2} dy \right] + \frac{1}{2(\tan^{-1}(b))^2} \int_{\delta}^b \frac{y^n \tan^{-1}(\lambda y)}{1+y^2} dy
\end{aligned}$$

Το ζητούμενο προκύπτει χρησιμοποιώντας το μετασχηματισμό  $z = y^2$  στα δύο πρώτα ολοκληρώματα και εφαρμόζοντας έπειτα τη σχέση (3.194.5) των Gradshteyn and Ryzhik (2000) (βλέπε απόδειξη Θεωρήματος 2.2.1). ■

Χρησιμοποιώντας τις ιδιότητες της υπεργεωμετρικής συνάρτησης είναι εύκολο να δοθούν άμεσοι τύποι υπολογισμού των δέκα πρώτων ροπών. Ειδικότερα προκύπτουν τα ακόλουθα.

#### Πόρισμα 4.2.1.1 (Nadarajah and Kotz, 2007 a)

Έστω  $X$  μια τ.μ. που ακολουθεί την λοξή συμμετρικά αποκομμένη τυπική κατανομή Cauchy με σ.π.π. που δίνεται από την σχέση (4.2.1.1). Τότε οι πέντε πρώτες περιττές τάξης ροπές της  $X$ , δίνονται από τις σχέσεις:

$$E(X) = \frac{\log(1+b^2) - \log(1+\delta^2)}{2 \tan^{-1}(b)} - \frac{J_1(\delta)}{(\tan^{-1}(b))^2},$$

$$E(X^3) = \frac{b^2 - \delta^2 - \log(1+b^2) + \log(1+\delta^2)}{2 \tan^{-1}(b)} - \frac{J_3(\delta)}{(\tan^{-1}(b))^2},$$

$$E(X^5) = \frac{b^4 - 2b^2 - \delta^4 + 2\delta^2 + 2\log(1+b^2) - 2\log(1+\delta^2)}{4 \tan^{-1}(b)} - \frac{J_5(\delta)}{(\tan^{-1}(b))^2},$$

$$E(X^7) = \frac{2b^6 - 3b^4 + 6b^2 - 2\delta^6 + 3\delta^4 - 6\delta^2 - 6\log(1+b^2) + 6\log(1+\delta^2)}{12 \tan^{-1}(b)} - \frac{J_7(\delta)}{(\tan^{-1}(b))^2},$$

και

$$E(X^9) = \frac{3b^8 - 4b^6 + 6b^4 - 12b^2 - 3\delta^8 + 4\delta^6}{24 \tan^{-1}(b)} - \frac{6\delta^4 - 12\delta^2 - 12\log(1+b^2) + 12\log(1+\delta^2)}{24 \tan^{-1}(b)} - \frac{J_9(\delta)}{(\tan^{-1}(b))^2},$$

$$\text{όπου } b > 0, \delta = \min(b/\gamma, b) \text{ και } J_n(\delta) = \int_0^\delta \frac{z^n \tan^{-1}(\gamma z)}{1+z^2} dz.$$

#### Πόρισμα 4.2.1.2 (Nadarajah and Kotz, 2007 a)

Έστω  $X$  μια τ.μ. που ακολουθεί την λοξή συμμετρική αποκομμένη κατανομή Cauchy με σ.π.π. που δίνεται από την σχέση (4.2.1.1). Τότε οι πέντε πρώτες άρτιας τάξης ροπές της  $X$ , δίνονται από τις σχέσεις:

$$E(X^2) = \frac{b - \tan^{-1}(b)}{\tan^{-1}(b)},$$

$$E(X^4) = \frac{b^3 - 3b + 3 \tan^{-1}(b)}{3 \tan^{-1}(b)},$$

$$E(X^6) = \frac{3b^5 - 5b^3 + 15b - 15 \tan^{-1}(b)}{15 \tan^{-1}(b)},$$

$$E(X^8) = \frac{15b^7 - 21b^5 + 35b^3 - 105b + 105 \tan^{-1}(b)}{105 \tan^{-1}(b)},$$

$$E(X^{10}) = \frac{35b^9 - 45b^7 + 63b^5 - 105b^3 + 315b - 315 \tan^{-1}(b)}{315 \tan^{-1}(b)}.$$

#### 4.2.2 Αποκομμένη λοξή Κατανομή Cauchy

Στην ενότητα αυτή θα παρουσιαστεί η μονοδιάστατη διπλά αποκομμένη λοξή κατανομή, που πρωτοπαρουσιάστηκε από τους Jamalizadeh *et al.* (2009). Η κατανομή αυτή προκύπτει από τη σχέση (1.2.1) αν λάβουμε υπόψη τη σ.π.π. και την α.σ.κ. της λοξής AC κατανομής Cauchy, όπως αυτές δίνονται στις σχέσεις (4.2.3) και (4.2.4), αντίστοιχα.

##### Ορισμός 4.2.2.1 (Jamalizadeh *et al.*, 2009)

Λέμε ότι η τυχαία μεταβλητή  $X$  ακολουθεί τη διπλά αποκομμένη λοξή κατανομή Cauchy με παράμετρο  $\lambda \in R$  αν η σ.π.π. της δίνεται από τη σχέση:

$$f_x(x; \lambda, a, b) = \frac{\frac{1}{\pi(1+x^2)} \left[ 1 + \frac{\lambda x}{\sqrt{1+(1+\lambda^2)x^2}} \right]}{G(b; \lambda) - G(a; \lambda)}, \quad (4.2.2.1)$$

όπου  $a \leq x \leq b$ ,  $\lambda \in R$ , με

$$G(x; \lambda) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left[ \frac{x \sqrt{1+(1+\lambda^2)x^2} - \lambda}{\sqrt{1+(1+\lambda^2)x^2} + \lambda x} \right].$$



Τότε γράφουμε ότι  $X \sim TSC^{(a,b)}(\lambda)$ .

Έστω τώρα ότι  $X \sim TSC^{(a,b)}(\lambda)$ . Στην πρόταση που ακολουθεί δίνεται η α.σ.κ. της τ.μ.  $X$ , σύμφωνα με τη σχέση (1.2.2), όπου ορίστηκε η α.σ.κ. της αποκομμένης μορφής μιας κατανομής.

#### Πρόταση 4.2.2.1

Έστω  $X$  μια τ.μ. που ακολουθεί τη διπλά αποκομμένη λοξή κατανομή Cauchy στο διάστημα  $[a, b]$ , δηλαδή  $X \sim TSC^{(a,b)}(\lambda)$ . Τότε η α.σ.κ. της τ.μ.  $X$  ορίζεται από τη σχέση:

$$F_X(x; \lambda, a, b) = \begin{cases} 0 & , \text{για } x < a \\ \frac{G_X(x; \lambda) - G_X(a; \lambda)}{G_X(b; \lambda) - G_X(a; \lambda)} & , \text{για } a \leq x < b \\ 1 & , \text{για } x \geq b \end{cases} \quad (4.2.2.2)$$

όπου  $\lambda \in R$  και  $G(\cdot, \cdot)$  η α.σ.κ. της λοξής AC κατανομής κατανομής, με

$$G(x; \lambda) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left[ \frac{x \sqrt{1 + (1 + \lambda^2)x^2 - \lambda}}{\sqrt{1 + (1 + \lambda^2)x^2 + \lambda x}} \right].$$

Η μονοδιάστατη διπλά αποκομμένη λοξή κατανομή Cauchy ορίστηκε σε ένα πεπερασμένο διάστημα. Αυτό έχει ως αποτέλεσμα να είναι πεπερασμένες οι ροπές κάθε τάξης. Αντικείμενο μελέτης σε αυτήν την παράγραφο είναι η  $n$  τάξης ροπή της  $TSC^{(a,b)}(\lambda)$ , η μέση τιμή και η διακύμανσή της.

Οι Jamalizadeh *et al.* (2009) μας εφοδιάζουν με έναν άμεσο τύπο των ροπών για  $n$  φυσικό, με τη βοήθεια των υπεργεωμετρικών συναρτήσεων Gauss

και Appell. Η υπεργεωμετρική Gauss έχει οριστεί στη σχέση (2.2.3), ενώ η υπεργεωμετρική Appell συνάρτηση ορίζεται από τη σχέση:

$$F_1(a; b_1, b_2; c; x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n} (b_1)_m (b_2)_n}{m! n! (c)_{m+n}} x^m y^n, \quad (4.2.2.3)$$

με  $(z)_k = z(z+1) \cdots (z+k-1)$ .

Τότε αν

$$A(x; n, \lambda) = \int \frac{x^n}{(1+x^2)} \left[ 1 + \frac{\lambda x}{\sqrt{1+(1+\lambda^2)x^2}} \right] dx, \quad (4.2.2.4)$$

όπου

$$A(x; n, \lambda) = \frac{x^{n+1} \left[ -(n+1)x\Lambda_{n,\lambda,x} + \lambda(n+2)M_{n,x} + (1+\lambda^2)(n+1)xK_{n,\lambda,x} \right]}{\lambda(n^2 + 3n + 2)},$$

με

$$\Lambda_{n,\lambda,x} = F_1\left(\frac{n+2}{2}; -\frac{1}{2}, 1; \frac{n+4}{2}; -(1+\lambda^2)x^2, -x^2\right),$$

$$M_{n,x} = {}_2F_1\left(\frac{n+1}{2}, 1; \frac{n+3}{2}; -x^2\right),$$

και

$$K_{n,\lambda,x} = {}_2F_1\left(\frac{n+2}{2}, \frac{1}{2}; \frac{n+4}{2}; -(1+\lambda^2)x^2\right),$$

προκύπτει το ακόλουθο Θεώρημα για την απόδειξη του οποίου παραπέμπουμε στους Jamalizadeh *et al.* (2009).

**Θεώρημα 4.2.2.1 (Jamalizadeh *et al.*, 2009)**

Έστω  $X$  μια τ.μ. που ακολουθεί την  $TSC^{(a,b)}(\lambda)$ . Τότε η  $n$  τάξης κεντρική ροπή της  $X$ , για  $n \in N$ , δίνεται από τη σχέση:

$$E(X^n) = \frac{1}{\pi} \left[ \frac{A(b;n,\lambda) - A(a;n,\lambda)}{G(b;\lambda) - G(a;\lambda)} \right], \quad (4.2.2.5)$$

όπου  $a \leq x \leq b$ ,  $\lambda \in R$ , η συνάρτηση  $A(x;n,\lambda)$  όπως έχει οριστεί στη σχέση (4.2.2.4), ενώ  $G(\cdot, \cdot)$  η α.σ.κ. της λοξής AC κατανομής κατανομής, με

$$G(x;\lambda) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left[ \frac{x\sqrt{1+(1+\lambda^2)x^2 - \lambda}}{\sqrt{1+(1+\lambda^2)x^2 + \lambda x}} \right].$$

Επιπλέον, προκύπτει ότι

$$E(X) = \frac{B(a,b,\lambda)}{\pi[G(b;\lambda) - G(a;\lambda)]}, \quad (4.2.2.6)$$

όπου

$$B(a,b,\lambda) = \frac{1}{2} \log \left[ \frac{1+b^2}{1+a^2} \right] + \frac{\lambda}{\sqrt{1+\lambda^2}} \log \left[ \frac{b\sqrt{1+\lambda^2} + \sqrt{1+(1+\lambda^2)b^2}}{a\sqrt{1+\lambda^2} + \sqrt{1+(1+\lambda^2)a^2}} \right] - \frac{1}{2} \log \frac{(\sqrt{1+(1+\lambda^2)b^2} + \lambda b)(\sqrt{1+(1+\lambda^2)a^2} - \lambda a)}{(\sqrt{1+(1+\lambda^2)b^2} - \lambda b)(\sqrt{1+(1+\lambda^2)a^2} + \lambda a)}.$$

Για περισσότερες εκφράσεις που αφορούν τις ροπές της μονοδιάστατης διπλά αποκομμένης λοξής κατανομής, παραπέμπουμε στους Jamalizadeh *et al.*

(2009), όπου δίδεται και μία γενίκευση αυτής, ενώ χρήζουν περαιτέρω μελέτης άλλες ιδιότητες αυτής της κατανομής.



## ΕΠΙΛΟΓΟΣ

Στη μεταπτυχιακή αυτή διατριβή έγινε μία προσπάθεια να μελετηθεί η μονοδιάστατη αποκομμένη κατανομή Cauchy και να παρουσιαστούν τα κυριότερα, μέχρι την ώρα συγγραφής, αποτελέσματα και γενικεύσεις αυτής. Στο πλαίσιο αυτό, στο **Κεφάλαιο 1** δόθηκε μια σύντομη εισαγωγή στην κατανομή Cauchy, στον πλέον διαδεδομένο τρόπο αποκοπής και παρουσιάστηκε η διάρθρωση της μεταπτυχιακής διατριβής. Στη συνέχεια, στο **Κεφάλαιο 2** παρουσιάστηκε η μονοδιάστατη διπλά αποκομμένη κατανομή Cauchy και μια κριτική ανασκόπηση των ιδιοτήτων και των χαρακτηριστικών της. Η εκτίμηση των παραμέτρων της μονοδιάστατης διπλά αποκομμένης κατανομής Cauchy αποτέλεσε το αντικείμενο μελέτης του **Κεφαλαίου 3**. Ειδικότερα, παρουσιάστηκαν τα αποτελέσματα των Daliya *et al.* (2001) που αφορούν στην ύπαρξη και τον προσδιορισμό των Εκτιμητών Μέγιστης Πιθανοφάνειας (Ε.Μ.Π.), καθώς και τα αποτελέσματα του Nadarajah (2011) για την εκτίμηση των παραμέτρων με τη μέθοδο των ροπών. Καθώς προκύπτει ότι η ύπαρξη πεπερασμένου Ε.Μ.Π. της παραμέτρου  $\sigma$  της διπλά αποκομμένης κατανομής Cauchy εξαρτάται από το αν πληρείται μία συνθήκη, παρακινούμενοι από τους Daliya *et al.* (2001) διεξάγουμε μία μελέτη προσομοίωσης για την εκτίμηση της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή. Επιπρόσθετα, με χρήση προσομοιωμένων δεδομένων αξιολογείται η απόδοση των Ε.Μ.Π. των άγνωστων παραμέτρων. Τέλος, προσαρμόζεται το μοντέλο της διπλά αποκομμένης κατανομής Cauchy σε ένα γνωστό σύνολο δεδομένων.

Στη στατιστική βιβλιογραφία είναι γνωστό ότι συνεχώς εισάγονται νέες κατανομές ως γενικεύσεις υπάρχουσων για τη μοντελοποίηση δεδομένων που οι τελευταίες αδυνατούν να περιγράψουν ικανοποιητικά. Στο **Κεφάλαιο 4**, το

ενδιαφέρον επικεντρώθηκε στις μονοδιάστατες γενικεύσεις της μονοδιάστατης διπλά αποκομμένης κατανομής που έχουν εισαχθεί στη βιβλιογραφία.

Στα περιορισμένα πλαίσια μιας μεταπτυχιακής διατριβής το ενδιαφέρον επικεντρώθηκε στη μονοδιάστατη αποκομμένη κατανομή Cauchy και σε μονοδιάστατες γενικεύσεις αυτής. Στο σημείο αυτό θα ήταν παράλειψη να μην αναφερθεί ότι τα τελευταία χρόνια στη βιβλιογραφία έχουν παρουσιαστεί και πολυδιάστατες αποκομμένες κατανομές Cauchy. Ειδικότερα, οι Nadarajah and Kotz (2007 a) εισήγαγαν τη διδιάστατη αποκομμένη κατανομή Cauchy, υπολόγισαν τις περιθώριες κατανομές της κατανομής, άμεσους τύπους των ροπών της και μεθόδους εκτίμησης των παραμέτρων της διπλά αποκομμένης διδιάστατης κατανομής Cauchy. Από την άλλη μεριά, οι Ateya and Madhagi (2012) παρουσίασαν τη λεγόμενη πολυδιάστατη γενικευμένη Κατανομή Cauchy και μελέτησαν τις ιδιότητές της και μεθόδους εκτίμησης. Τέλος, σε μία πρόσφατη εργασία οι Ho *et al.* (2012) ασχολήθηκαν με την εύρεση των δύο πρώτων ροπών της πολυδιάστατης αποκομμένης  $t$  κατανομής που περικλείει την Cauchy ως ειδική περίπτωση.

## ΠΑΡΑΡΤΗΜΑ Α

Στο Παράρτημα Α, παρατίθενται οι πηγαίοι κώδικες (m\_files) της MatLab που χρησιμοποιήθηκαν για τις γραφικές παραστάσεις των πυκνοτήτων των διάφορων κατανομών που παρουσιάστηκαν στη μεταπτυχιακή αυτή διατριβή. Στο σημείο αυτό θα πρέπει να αναφερθεί ότι στο πλαίσιο της διπλά αποκομμένης κατανομής Cauchy στη βιβλιογραφία υπάρχουν (βλέπε π.χ. Nadarajah and Kotz, 2007 b) διαθέσιμοι κώδικες στη γλώσσα προγραμματισμού R για τον υπολογισμό της σ.π.π., της α.σ.κ., της μέσης τιμής, της διακύμανσης, της αντίστροφης της α.σ.κ. και για την προσομοίωση δεδομένων.

### Δημιουργία γραφικών παραστάσεων

```
%'Δημιουργία γραφικής παραστάσης της σ.π.π. της Κατανομής  
Cauchy, για x π.χ. στο [-5,5]'  
%'Δηλώνουμε τις τιμές των παραμέτρων μ και σ'  
s=1;m=1;x = -5:0.1:5;  
pdfcauchy=(s/pi)*(1./(s^2.+(x-m).^2));  
plot(x, pdfcauchy, ':r')  
legend('μ=1, σ=1',2)
```

```
%'Δημιουργία γραφικής παραστάσης της σ.π.π. της διπλά  
αποκομμένης Κατανομής Cauchy, για x στο [a,b]'  
%'Δηλώνουμε τις τιμές των σημείων αποκοπής καθώς και των  
παραμέτρων μ και σ'  
a=-3;b=3;s=1;m=1;d=atan((b-m)/s)-atan((a-m)/s);  
x = a:0.1:b;  
pdftruc=(s/d)*(1./(s^2.+(x-m).^2));  
plot(x, pdftruc, ':r')  
legend('μ=1, σ=1',2)
```

```
%'Δημιουργία γραφικής παραστάσης της σ.π.π. της κατανομής της  
παραγράφου 4.1.1, δηλαδή της αριστερά αποκομμένης  
γενικευμένης Κατανομής Cauchy'  
%'Δηλώνουμε τις τιμές των παραμέτρων α,μ, σ'  
a=0.5;m=0.5;s=1;  
x = 0.5:0.1:5;  
pdfgentruc=(2.*gamma(a+0.5)./(pi.*s.*gamma(a))).*(1.+((x-  
m)./s).^(-a-0.5));
```



```
plot(x, pdfgentruc, ':r')
legend('a=0.5,μ=0.5, σ=1',2)
```

```
%'Δημιουργία γραφικής παράστασης της σ.π.π. της κατανομής της
 παραγράφου 4.2.1, δηλαδή της λοξής αποκομμένης Κατανομής
 Cauchy '
%'Δηλώνουμε τις τιμές των b και λ'
b=3;l=2;d=2*atan(b);
x = -b:0.1:b;
pdfskewtruc=(1./(d.*(1+(x.^2)))).*(1.+atan(1.*x)./atan(b));
plot(x, pdfskewtruc, ':r')
legend('λ=2', 2)
```

Στη συνέχεια δίδεται ο τρόπος υπολογισμού των 4 πρώτων τάξης ροπών περί το μηδέν, της μέση τιμής, της διακύμανσης, της λοξότητας, της κύρτωσης, της διαμέσου, της μέση απόλυτης απόκλισης από τη μέση τιμή και τη διάμεσο. Διακρίνουμε τις περιπτώσεις που η τ.μ. ακολουθεί διπλά αποκομμένη: α) μη τυπική και β) τυπική κατανομή Cauchy.

#### Περίπτωση μη τυπικής

```
%'Δώσε τις παραμέτρους a,b, μ και σ για τη διπλά αποκομμένη
 μη τυπική Κατανομή Cauchy '

```

```
a=0;b=10;m=1;s=2;
```

```
%'Υπολόγισε τα A και B και τη σταθερά D '

```

```
A=(a-m)/s;B=(b-m)/s;D=atan(B)-atan(A);
```

```
%'Υπολόγισε τα  $\Lambda_{\{i\}}$   $i=0,1,\dots,4$ '

```

```
l0=B*hypergeom([1,0.5],1.5,-(B^2))-A*hypergeom([1,0.5],1.5,-
(A^2));
```

```
l1=(B^2)*hypergeom([1,1],2,-(B^2))-(A^2)*hypergeom([1,1],2,-
(A^2));
```

```
l2=(B^3)*hypergeom([1,1.5],2.5,-(B^2))-
(A^3)*hypergeom([1,1.5],2.5,-(A^2));
```

```

l3=(B^4)*hypergeom([1,2],3,-(B^2))-(A^4)*hypergeom([1,2],3,-
(A^2));

l4=(B^5)*hypergeom([1,2.5],3.5,-(B^2))-
(A^5)*hypergeom([1,2.5],3.5,-(A^2));

%'Υπολόγισε τις 4 πρώτες τάξης ροπές περί το μηδέν, a1-a4'

a1=(1/D)*(m*10+0.5*s*11);

a2=(1/D)*(10*(m^2)+m*s*11+(1/3)*(s^2)*12);

a3=(1/D)*(10*(m^3)+1.5*(m^2)*s*11+m*(s^2)*12+(1/4)*(s^3)*13);

a4=(1/D)*(10*(m^4)+2*(m^3)*s*11+2*(m^2)*(s^2)*12+m*(s^3)*13+(
1/5)*(s^4)*14);

%'Υπολόγισε τη μέση τιμή (thita), τη διακύμανση (var), τη
λοξότητα (skew), την κύρτωση (kurt), τη διάμεσο (median)'

thita=a1;

var=a2-(a1)^2;

skew=(a3-3*a1*a2+2*(a1)^3)/(var^(3/2));

kurt=(a4-4*a1*a3+6*(a1^2)*a2-3*(a1^4))/(var^(4/2));

median=m+s*tan(atan(A)+0.5*D);

%'Απαραίτητοι υπολογισμοί για την εύρεση της μέση απόλυτης
απόκλισης από τη μέση τιμή, έστω d1. Αρχικά υπολόγίζεται η
τιμή της α.σ.κ. αθροιστική για x=μέση τιμή'

if a1<a  fal=0;

elseif a<=a1<=b fal=(1/D)*(atan((a1-m)/s)-atan(A));

else fal=1;

end

fal;

B1=(a1-m)/s;

```

```

D1=atan(B1)-atan(A);

l01=B1*hypergeom([1,0.5],1.5,-(B1^2))-
A*hypergeom([1,0.5],1.5,-(A^2));

l11=(B1^2)*hypergeom([1,1],2,-(B1^2))-
(A^2)*hypergeom([1,1],2,-(A^2));

a11=(1/D1)*(m*l01+0.5*s*l11);

d1=2*(a1*fa1-(D1/D)*a11);

%'Απαραίτητοι υπολογισμοί για την εύρεση της μέση απόλυτης
απόκλισης από τη διάμεσο, έστω d2.'

B2=(median-m)/s;

D2=atan(B2)-atan(A);

l02=B2*hypergeom([1,0.5],1.5,-(B2^2))-
A*hypergeom([1,0.5],1.5,-(A^2));

l12=(B2^2)*hypergeom([1,1],2,-(B2^2))-
(A^2)*hypergeom([1,1],2,-(A^2));

a12=(1/D2)*(m*l02+0.5*s*l12);

d2=a1-2*D2*a12/D;

'Σύνοψη των αποτελεσμάτων: 4 πρώτες τάξης ροπές περί το
μηδέν, διακύμανση, λοξότητα, κύρτωση, διάμεσο, τη μέση
απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο'

[a1 a2 a3 a4 var skew kurt median d1 d2]

```

### Περίπτωση τυπικής

```

%'Δώσε τις παραμέτρους a,b για τυπική διπλά αποκομμένη
κατανομή Cauchy '

A=-5;B=0;

%'Υπολόγισε τη σταθερά D'

```

```
D=atan(B)-atan(A);

%'Υπολόγισε τις 4 πρώτες τάξης ροπές περί το μηδέν, a1-a4'

a1=(1/(2*D))*(log(1+B^2)-log(1+A^2));

a2=(1/D)*(atan(A)-atan(B)-A+B);

a3=(1/(2*D))*(log(1+A^2)-log(1+B^2)-A^2+B^2);

a4=(1/(3*D))*(3*atan(B)-3*atan(A)-A^3+B^3+3*A-3*B);

%'Υπολόγισε τη μέση τιμή (thita), τη διακύμανση (var), τη
λοξότητα (skew), την κύρτωση (kurt), τη διάμεσο (median)'

thita=a1;

var=a2-(a1)^2;

skew=(a3-3*a1*a2+2*(a1)^3)/(var^(3/2));

kurt=(a4-4*a1*a3+6*(a1^2)*a2-3*(a1^4))/(var^(4/2));

median=tan(atan(A)+0.5*D)

%'Απαραίτητοι υπολογισμοί για την εύρεση της μέση απόλυτης
απόκλισης από τη μέση τιμή, έστω d1. Αρχικά υπολογίζεται η
τιμή της α.σ.κ. αθροιστική για x=μέση τιμή'

if a1<A fa1=0;

elseif A<=a1<=B fa1=(1/D)*(atan(a1)-atan(A));

else fa1=1;

end

fa1;

B1=a1;

D1=atan(B1)-atan(A);

a11=(1/(2*D1))*(log(1+B1^2)-log(1+A^2));

d1=2*(a1*fa1-(D1/D)*a11);
```

%'Απαραίτητοι υπολογισμοί για την εύρεση της μέσης απόλυτης απόκλισης από τη διάμεσο, έστω d2.'

B2=median;

D2=atan(B2)-atan(A);

a12=(1/(2\*D2))\*(log(1+B2^2)-log(1+A^2));

d2=a1-2\*D2\*a12/D;

Σύνοψη των αποτελεσμάτων: 4 πρώτες τάξης ροπές περί το μηδέν, διακύμανση, λοξότητα, κύρτωση, διάμεσο, τη μέση απόλυτη απόκλιση από τη μέση τιμή και τη διάμεσο'

[a1 a2 a3 a4 var skew kurt median d1 d2]

## ΠΑΡΑΡΤΗΜΑ Β

Στο Παράρτημα Β, παρατίθενται οι πηγαίοι κώδικες (m\_files) της MatLab που χρησιμοποιήθηκαν α) για τη μελέτη προσομοίωσης και β) το αριθμητικό παράδειγμα του Κεφαλαίου 3.

### Μελέτη Προσομοίωσης

#### α) Εκτίμηση της πιθανότητας μη ύπαρξης πεπερασμένου εκτιμητή για την παράμετρο $\sigma$ .

% Προσομοίωση Β το πλήθος δειγμάτων μεγέθους N από τη διπλά αποκομμένη στο (-1,1) Κατανομή Cauchy με παραμέτρους  $\mu$  και  $\sigma$

```
B = 10000;  
N = 200;
```

```
m = 0.75; % Πρέπει  $-1 < \mu = m < 1$   
s = 4.0; % Πρέπει  $\sigma = s > 0$ 
```

```
%Υπολογισμός χρήσιμων ποσοτήτων
```

```
d1 = atan((-1 - m) / s);  
d2 = atan(( 1 - m) / s) - d1;
```

```
%Έναρξη επαναλήψεων και υπολογισμός του  $\sum_{i=1}^n (x_i - \mu)^2$ 
```

```
for iter = 1 : B  
    data(:, :) = rand(N, 1); % Δημιουργία τ.δ. μεγέθους N από  
    την ομοιόμορφη στο (0,1)  
    x(:, :, iter) = m + s .* tan(d1 + d2 .* data(:, :)); %  
    Μετασχηματισμός για να δημιουργηθούν τα δεδομένα από την  
    διπλά αποκομμένη κατανομή  
    statistic(iter) = (x(:, :, iter) - m)' * (x(:, :, iter) - m); %  
end
```

%Εύρεση της συχνότητας (count) πραγματοποίησης του

$$\frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \geq \left( \mu^2 + \frac{1}{3} \right) \text{ και εκτίμηση της πιθανότητας ως τη}$$

σχετική συχνότητα

```
count=0;
for i=1:B if statistic(i)>=(N*(m^2+1/3)) count=count+1;
end
end
count;
failtoexist=count/B
```

### β) Αξιολόγηση Ε.Μ.Π.

% Γράφουμε τη σ.π.π. της διπλά αποκομμένης στο (-1,1) Κατανομής Cauchy με παραμέτρους μ και σ άγνωστες

```
pdf_trunczcauchy = @(x, mu, si) (si / (atan((1 - mu) / si) - atan((-1 - mu) / si))) ./ (si ^ 2 + (x - mu) .^ 2);
```

% Προσομοίωση Β το πλήθος τ.δ. μεγέθους N από την παραπάνω κατανομή για κάποιες τιμές των άγνωστων παραμέτρων. Για κάθε ένα από αυτά εύρεση των Ε.Μ.Π.

```
B = 5000;
N = 20
m = 0.2 % Assuming -1<m < 1!
s = 0.10
d1 = atan((-1 - m) / s); % Χρήσιμα για τους υπολογισμούς
d2 = atan(( 1 - m) / s) - d1;

for iter = 1 : B
    data = rand(N, 1); % Δημιουργία τ.δ. μεγέθους N από την
    μοιόμορφη στο (0,1)
    x = m + s .* tan(d1 + d2 .* data); % Μετασχηματισμός για να
    δημιουργηθούν τα δεδομένα από την διπλά αποκομμένη κατανομή
    start = [m, s]; % Αρχικοί εκτιμητές
    apotelesma(:, iter) = mle(x, 'pdf', pdf_trunczcauchy,
    'start', start); % Ε.Μ.Π.
end

% Οι Ε.Μ.Π. των μ και σ είναι αντίστοιχα

mues = apotelesma(1, :);
```

```

sest = apotelesma(2, :);

% Μεροληψία και Μέσο τετραγωνικό Σφάλμα. Σύγκριση με τις
αληθινές τιμές από όπου προσομοιώσαμε τα δεδομένα

%Μεροληψία και μέσο τετραγωνικό σφάλμα για το μ
biasedm = mean(mues - m)
msedm=sum((mues - m).^2)/B

%Μεροληψία και μέσο τετραγωνικό σφάλμα για το σ

biaseds = mean(sest - s);

mseds=sum((sest - s).^2)/B;

```

### γ) Αριθμητικό παράδειγμα

```

% Γράφουμε τη σ.π.π. της διπλά αποκομμένης στο (-1,1)
Κατανομής Cauchy με παραμέτρους μ και σ άγνωστες

pdf_truncCauchy = @(x, mu, si) (si / (atan((1 - mu) / si) -
atan((-1 - mu) / si))) ./ (si ^ 2 + (x - mu) .^ 2);

%Δίνονται στη συνέχεια τα αρχικά δεδομένα
data =
[4.4623;4.363;4.4743;4.662;4.529;4.3956;4.3956;4.4623;4.8828;
4.4405;4.1068;3.73;3.7736;3.8956;4.4843;4.7506;4.5725;4.529;4
.3403;4.529;4.6232;4.9285;4.985;4.7893;4.8614;4.845;4.9652;4.
9432;4.8662;4.8614;4.7281;4.8757;4.8286;4.662;4.7125;4.8403;4
.7893;5.0736;4.89;4.8497;4.8239;4.845;4.7237;4.8239;4.8828;4.
7962;4.717;4.89;4.8054;4.8239;4.8614;4.9068;4.89;4.8497;4.812
3;4.817;4.845;4.8473;4.8709;4.8638;4.6168;4.8444;6.5025;7.391
4;11.0905;7.0621;6.5208;6.5086;6.5701;5.8583;5.3958;5.3105;5.
4563;5.3773;5.4686;5.5051;5.2132;5.0062;4.875;4.845;4.845;4.8
5;4.85;4.855;4.85;4.895;4.85;4.87;4.895;4.84;4.845;4.855;4.88
5;4.885;4.895;4.91;4.88;4.86;4.855;4.885;4.86;4.875;4.875;4.8
55;4.88;4.856;4.84;4.8425;4.871;4.8675;4.865;4.87;4.853;4.855
;4.8525;4.7362;4.7556;4.7512;4.7581;3.75;3.525;4.2063;4.6325;
4.3187;4.7225;4.8481;4.8481;4.8762;4.8488;4.8744;4.8516;3.377
5;3.3275;5.12;4.9363;4.93;4.9088;4.9969;4.6363;3.9537;4.035;
4.0325;4.0325;4.02;4.02;4.025;4.025;4.0331;4.0319;2.8006;2.80
13;2.7814;2.8096;2.8106;2.7856;2.8025;2.7856;2.8093;2.8021;2.
7995;2.8038;2.8081;2.8025;2.7969;2.79;2.8025;2.79;2.4067;2.38
49;2.3989;2.3938;2.552;2.348;2.3225;2.347;2.0242;1.7025;1.92;

```



```

2.0435;2.2145;2.395;1.917;1.618;1.4515;1.158;1.439;1.4819;1.8
867;1.8089;1.611;1.932;1.865;1.51;1.4765;1.566;1.55;1.712;1.6
47;1.6539;1.6176;1.4957;1.4541;1.611;1.785];
% Τα δεδομένα μετασχηματίζονται λογαριθμικά.
y=log(data);
k=size(y); %Δηλώνει τη διάσταση των δεδομένων
% Υπολογίζονται οι σχετικές μεταβολές

x1=zeros(k(1)-1,1);
for j=1:k(1)-1 x1(j)=y(j+1)-y(j);
end
x1;
x=zeros(k(1)-1,1);
for j=1:k(1)-1 x(j)=(2*x1(j)-(max(x1)+min(x1)))/(max(x1)-
min(x1));
end
x;

start = [mean(x), std(x)]; % Αρχικοί εκτιμητές
apotelesma=mle(x, 'pdf', pdf_truncauchy, 'start', start) %
E.M.Π.

sum(log(pdf_truncauchy(x,apotelesma(1), apotelesma(2))))
%logL
%Αρχικοί εκτιμητές
mestimator=(max(x1)-
min(x1))*apotelesma(1)+(max(x1)+min(x1))/0.5
sestimator=(max(x1)-min(x1))*apotelesma(2)/0.5

```

# **ABSTRACT OF THE MSc DISSERTATION**

## **Title: Univariate truncated Cauchy distribution**

The Cauchy distribution has been studied in the statistics literature for over five centuries, since it can be considered as an alternative model whenever one need a density function with heavier tails than the normal distribution allows. However, a main problem, weakness of the Cauchy distribution is that it has no moments. This weakness was overcome by introducing a truncated version. In statistics, a truncated distribution is a distribution that results from restricting the domain of an existing one. This MSc thesis is concerned with the univariate truncated Cauchy distribution. Our main aim is to review the available results in the literature. In this frame, the present dissertation consists of four chapters and the bibliography. We briefly present the subject-matter of each one in the sequel.

**Chapter 1 (Introduction)** explains the importance of the truncated Cauchy distribution. Also, a short summary of this dissertation is presented.

In **Chapter 2 (Univariate doubly truncated Cauchy distribution)**, the doubly truncated version of the Cauchy distribution, which was initially introduced to the best of our knowledge by Daliya et al. (2001), is presented. The rest of the chapter is concerned with a critical review of the properties of this distribution. In this context, expressions are given for the moments, mean deviations, moment generating function, characteristic function, entropies, descriptive measures of the shape of the distribution and the asymptotic distribution of the minimum and

maximum order statistics. These results are mainly based on Nadarajah and Kotz (2006), Nadarajah (2011) and Daliya *et al.* (2001).

In **Chapter 3 (Estimation)**, the double truncated Cauchy distribution with four unknown parameters is considered and the results of Daliya *et al.* (2001) related to the derivation and the existence of the Maximum Likelihood Estimators (m.l.e.) are presented. In this context, a sufficient condition for the m.l.e. of the scale parameter to be finite, which is also proved to be necessary for sufficiently large samples, is given. Motivated by Daliya *et al.* (2001) based on Monte Carlo study we estimate the probability of the non existence of the m.l.e of the scale parameter and we study the small samples properties of the m.l.e. Moreover, estimation by the method of moments is considered (see Nadarajah (2011)). Finally, motivated among others by Jamalizadeh *et al.* (2009), we illustrate the use of the double truncated Cauchy distribution in fitting a real-data on the exchange rate between UK pound and US dollar.

In **Chapter 4 (Generalizations)** some univariate generalizations of the double truncated Cauchy distribution are presented.

In **Appendices A and B** the m\_files created in Matlab are presented. The files of the Appendix A were used in Chapters 1,2,4, while the files of the Appendix B were used in Chapter 3.

Finally, the **Bibliography** on which the dissertation was based is provided.

## BIBΛΙΟΓΡΑΦΙΑ

**Ateya, S.F. and AL-Hussaini, E.K.** (2012). On truncated generalized Cauchy distribution. *J. Math. Comput. Sci.*, **2**, 289-304.

**Ateya, S.F. and Madhagi, E.A.** (2012). On multivariate truncated generalized Cauchy distribution. *Statistical Papers*, Accepted, DOI10.1007/s00362-012-0467-9

**Azzalini, A.** (1985). A class of distributions which includes the normal ones. *Scand. J. Statist.* **12**, 171–178.

**Azzalini, A.** (1986). Further results on a class of distributions which includes the normal ones. *Statistica* **46**, 199–208.

**Azzalini, A. and Capitanio, A.** (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **65**, 367–389.

**Bauwens, L., Lubrano, M. and Richard, J. F.** (1999). *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, New York.

**Behboodian, J., Jamalizadeh, A., Balakrishnan, N.** (2006). A new class of skew-Cauchy distributions. *Statist. Probab. Lett.* **76**, 1488–1493.

**Blattberg, R. C. and Gonedes, N. J.** (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *Journal of Business*, **47**, 244-280.

**Borgia, G. C., Brown, R. J. S. and Fantazzini, P.** (1996). The effect of diffusion and susceptibility differences on T-2 measurements for fluids in porous media and biological tissues. *Magnetic Resonance Imaging*, **14**, 731–736.

**Cauchy, A. L.** (1853). Sur les resultats moyens d'observations de mneme nature, et sur les resultats les plus probables. *C.r. hebd. Se' anc. Acad. Sci., Paris* **37**, 198-206.

**Daliya, Ram C., Staneski, Paul G. and Chaganty, N. Rao** (2001). Maximum likelihood estimation of parameters of the truncated Cauchy distribution, *Communications in Statistics - Theory and Methods*, **30**, 1737—1750.

**David, H.A.** (1981). *Order Statistics, 2nd ed.* Wiley-Interscience, New York.

**Fantazzini, P. and Brown, R. J. S.** (2005). Initially linear echo-spacing dependence of I/T-2 measurements in many porous media with pore-scale inhomogeneous fields. *Journal of Magnetic Resonance*, **177**, 228–235.

**Frechet, M.** (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Pol. Math. Cracovie*, **6**, 93–116.

**Galambos, Janos** (1978). *The asymptotic theory of extreme order statistics. Wiley Series in Probability and Mathematical Statistics.* John Wiley & Sons, New York-Chichester-Brisbane.

**Gnedenko, B.** (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann. of Math. (2)*, **44**, 423–453.

**Gradshteyn, I.S., Ryzhik, I.M.** (2000). *Table of Integrals, Series, and Products*, 6th edn. Academic Press, San Diego.

**Gupta, A. K., Chang, F. C. and Huang, W. J.** (2002). Some skew-symmetric models, *Random Operators and Stochastic Equations*, **10**, 133–40.

**Henze, N.** (1986). A probabilistic representation of the "skew-normal" distribution. *Scand. J. Statist.* **13**, No. **4**, 271–275.

**Ho, Hsiu J., Lin, Tsung-I., Chen, Hsuan-Yu and Wang, Wan-Lun.** (2012). Some results on the truncated multivariate t distribution. *J. Statist. Plann. Inference*, **142**, 25–40.

**Huang, W.-J. and Chen, Y.-H.** (2007). Generalized skew-Cauchy distribution. *Statistics and Probability Letters*, **77**, 1137–1147.

**Jamalizadeh, A., Pourmousa, R. and Balakrishnan** (2009). Truncated and Limited Skew-Normal and Skew-t Distributions: Properties and an Illustration, *Communications in Statistics - Theory and Methods*, **38**, 16-17, 2653-2668

**Jimenez, J.** (1996). Algebraic probability density tails in decaying isotropic two-dimensional turbulence. *Journal of Fluid Mechanics*, **313**, 223–240.

**Johnson, N. L. and Kotz, S.** (1970). *Continuous Univariate Distributions*, Volume I, John Wiley and Sons, New York.

**Kim, H.J.** (2008). Moments of truncated Student-t distribution. *J.Korean Statist.Soc.* **37**, 81–87.

**Kim, H.-M. and B. Mallick.** (2003). A note on Bayesian spatial prediction using the elliptical distribution, *Statistics and Probability Letters* **64**, 271–276.

**Mann, N. and Singpurwalla, N.** (2006). *Extreme value distributions*. In *Encyclopedia of Statistical Sciences*. Eds Kotz et al., John Wiley & Sons, New Jersey.

**Marshall, A.W.; Olkin, I. Inequalities** (1979). *Theory of Majorization and Its Applications*. Academic Press, New York,

**Mitra, G. B. and Das, S.** (1989). Cauchy distribution, intensity statistics and phases of reflections from crystal planes. *Acta Crystallographica, A*, **45**, 314–319.

**Mitrinovic, D.S.** (1970). *Analytic Inequalities*. Springer-Verlag, New York.

**Nadarajah, S.** (2008). Letter to the editor. *Journal of the Korean Statistical Society*, **37**, 291–292.

**Nadarajah, S.** (2011). Making the Cauchy Work. *Braz. J. Probab. Stat.* Volume **25**, 99-120.

**Nadarajah, S. and Masoom Ali, M.** (2004). A skewed truncated t distribution. *Math. Comput. Modelling*, **40**, 935–939.

**Nadarajah, S. and Gupta, A.** (2005). A skewed truncated Pearson type VII distribution. *J. Japan Statist. Soc.* **35**, 61–71.

**Nadarajah, S. and Kotz, S.** (2006). A truncated Cauchy distribution. *International Journal of Mathematical Education in Science and Technology*, **37**, 605–608.

**Nadarajah, S. and Kotz, S.** (2007 a). A skewed truncated Cauchy distribution with applications in economics, *Applied Economics Letters*, Volume **14**, No. **13**, 957 -961.

**Nadarajah, S. and Kotz, S.** (2007 b). Programs in R for Computing Truncated t Distributions. *Qual. Reliab. Engng. Int.*, **23**, 273–278.

**Nadarajah, S. and Zografos, K.** (2003). Formulas for Rényi information and related measures for univariate distributions. *Information Sciences*, **155**, 119-138.

**Nadarajah, S. and Zografos, K.** (2005). Expressions for Rényi and Shannon entropies for bivariate distributions. *Information Sciences*, **170**, 173-189.

**Ni, S.** (2001). Network capacity and quality of service management in F/TDMA cellular systems. *Acta Polytechnica Scandinavica*, Electrical Engineering Series No. **105**, Espoo, 69 pp. Finnish Academies of Technology, Finland.

**Prokhorov, A. V.** (1999). The Cauchy distribution. In *Encyclopedia of Probability and Mathematical Statistics*, Yu. V. Prokhorov, ed. Moscow, Russia.

**Poisson, S. D.** (1824). Sur la probabilité des résultats moyens des observations. *Connaissance des Temps pour l'an 1827*, 273-302.

**Quenouille, M. H.** (1956). Notes on bias in estimation. *Biometrika*, **43**, 353-60.

**Rényi, A.** (1961). On measures of entropy and information. In: *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, pp. 547–561. University of California Press, Berkeley.

**Rider, P. R.** (1957). Generalized Cauchy distribution. *Ann. Instit. Statist. Math.* **9**, 215-223.

**Robson, D.S. and Whitlock, J.H.** (1964). Estimation of truncation point. *Biometrika*, **51**, 33–39.

**Rohatgi, V. K.** (1976). *An Introduction to Probability Theory and Mathematical Statistics*. Wiley and Sons, New York.

**Shannon, C.E.** (1948). A mathematical theory of communication. *Bell System Tech. J.*, **27**, 379-432.

**Song, K. -S.** (2001). Rényi information, loglikelihood and an intrinsic distribution measure. *Journal of Statistical Planning and Inference*, **93**, 51–69.

**Staneski, P.G.** (1990). The truncated Cauchy distribution, estimation of parameters and application to stock returns. Unpublished PhD dissertation, Dept. of Math. and Stat., Old Dominion University, Norfolk, VA.

**Stigler, Stephen M.** (1974). Studies in the history of probability and statistics. XXXIII. Cauchy and the witch of Agnesi: an historical note on the Cauchy distribution. *Biometrika*, **61**, 375–380.

**Σωκράτους, Κ.** (2009). *Μονοδιάστατη λοξή-κανονική κατανομή*. Μετ. Διατριβή, Τμήμα Μαθηματικών, Πανεπιστήμιο Ιωαννίνων.

**Varian, H.** (1975). A Bayesian approach to real estate assessment. In Fienberg, S.E. and Zellner, A. (Eds.): *Studies in Bayesian Econometrics and Statistics*. In honour of Leonard J.Savage, North Holland, Amestrdam, 195-208.

**Zografos, K. and Nadarajah, S.** (2005). Expressions for Rényi and Shannon entropies for multivariate distributions. *Statistics and Probability Letters*, **71**, 71-84.

**Zografos, K.** (2008). On Mardia's and Song's measures of kurtosis in elliptical distributions. *Journal of Multivariate Analysis*, **99**, 858-879.