

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ  
ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ  
ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΤΗΝ ΔΙΑΧΕΙΡΙΣΗ  
ΚΙΝΔΥΝΩΝ

**Η κατανομή Burr και οι εφαρμογές της στον αναλογισμό**

**ΚΟΥΝΤΟΥΡΑΚΟΥ ΕΛΕΝΗ-ΧΡΙΣΤΙΝΑ**

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

Πειραιάς

Ιούνιος 2023

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

- Αντζουλάκος Δημήτριος (Επιβλέπων - Αναπληρωτής Καθηγητής)
- Μπούτσικας Μιχαήλ (Αναπληρωτής Καθηγητής)
- Ψαρράκος Γεώργιος (Αναπληρωτής Καθηγητής)

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

UNIVERSITY OF PIRAEUS  
SCHOOL OF FINANCE AND STATISTICS



DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

POSTGRADUATE PROGRAM IN  
**ACTUARIAL SCIENCE AND RISK MANAGEMENT**

**Burr distribution and its applications in actuarial science**

**KOUNTOURAKOU ELENI-CHRISTINA**

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

Piraeus

June 2023



*Αφιερωμένο στο σύζυγό μου,*

*Κωνσταντίνο.*



# ΕΥΧΑΡΙΣΤΙΕΣ

---

Θα ήθελα να ευχαριστήσω ιδιαιτέρως τον αναπληρωτή καθηγητή κύριο Αντζουλάκο Δημήτριο, για την πολύτιμη βοήθειά του κατά την εκπόνηση της διπλωματικής μου εργασίας, όπως και τους επιβλέποντες Μπούτσικα Μιχαήλ και Ψαρράκο Γεώργιο. Τέλος, θα ήθελα να ευχαριστήσω όλους τους διδάσκοντες είτε μέλη ΔΕΠ είτε εξωτερικούς, για την πολύτιμη γνώση που μας μετέδωσαν.





# ΠΕΡΙΛΗΨΗ

---

Στη διπλωματική εργασία που θα ακολουθήσει, εξετάζουμε την οικογένεια κατανομών Burr και την εφαρμογή της στον αναλογισμό. Η προαναφερθείσα οικογένεια κατανομών αποτελείται από δώδεκα βασικές συνεχείς κατανομές, οι οποίες εφαρμόζονται σε πολλές επιστήμες. Έχοντας ως γνώμονα την πολύπλευρη εφαρμογή των κατανομών Burr σε διάφορους επιστημονικούς τομείς, πραγματοποιήθηκε εφαρμογή αυτών τόσο στην επιστήμη του αναλογισμού, όσο και σε άλλες επιστήμες. Η εργασία αποτελείται από τρία κεφάλαια. Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή σε όλες τις βασικές έννοιες που θα χρησιμοποιηθούν στα δύο επόμενα κεφάλαια. Εν συνεχεία, ακολουθεί το δεύτερο κεφάλαιο, στο οποίο λαμβάνουμε ως σημείο αναφοράς το 1942, το έτος κατά το οποίο ο Burr εισήγαγε τις δώδεκα αθροιστικές συναρτήσεις κατανομής. Πραγματοποιείται παρουσίαση και των δώδεκα κατανομών και περαιτέρω ανάλυση των πιο σημαντικών από αυτών. Τέλος, στο τρίτο κεφάλαιο, με τη βοήθεια της γλώσσας προγραμματισμού R, γίνεται προσαρμογή μερικών εκ των κατανομών Burr πάνω σε πραγματικά δεδομένα, καθώς και σύγκριση αυτών με άλλες γνωστές κατανομές με απώτερο σκοπό την διεξαγωγή συμπερασμάτων.



# ABSTRACT

---

In the following dissertation, the family of Burr distributions and its applications in actuarial science will be examined. The aforementioned distribution family consists of twelve basic continuous distributions that are applied in various scientific fields. Considering the versatile application of Burr distributions in multiple scientific disciplines, some of them have been applied in the field of actuarial as well as in other scientific fields. The dissertation consists of three chapters. The first chapter provides an introduction to all the fundamental concepts that will be used in the subsequent chapters. Next, the second chapter focuses on the year 1942 as a reference point when Burr introduced the twelve cumulative distribution functions. A presentation of the twelve distributions is given, followed by a detailed analysis of the most significant ones. Finally, in the third chapter, utilizing the programming language R, we fit some of the Burr distributions to real data and compare them to other well-known distributions with the ultimate goal of drawing conclusions.



# ΠΕΡΙΕΧΟΜΕΝΑ

---

|  |           |
|--|-----------|
| ΠΕΡΙΛΗΨΗ.....  | i         |
| ABSTRACT .....   | iii       |
| ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ.....                                 | ix        |
| ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ .....                               | xii       |
| ΚΕΦΑΛΑΙΟ 1.....  | 1         |
| <b>Εισαγωγή.....</b>                                   | <b>1</b>  |
| <b>1.1 Τυχαίες Μεταβλητές.....</b>                     | <b>1</b>  |
| Διακριτές τυχαίες μεταβλητές.....                      | 1         |
| Συνεχείς τυχαίες μεταβλητές.....                       | 2         |
| <b>1.2 Συνάρτηση Κατανομής.....</b>                    | <b>2</b>  |
| <b>1.3 Συνάρτηση Επιβίωσης.....</b>                    | <b>3</b>  |
| <b>1.4 Συνάρτηση Κινδύνου.....</b>                     | <b>3</b>  |
| <b>1.5 Ροπές μιας τυχαίας μεταβλητής.....</b>          | <b>4</b>  |
| <b>1.6 Ασυμμετρία.....</b>                             | <b>5</b>  |
| <b>1.7 Κύρτωση.....</b>                                | <b>6</b>  |
| <b>1.8 Μέθοδοι Εκτίμησης Παραμέτρων.....</b>           | <b>7</b>  |
| Μέθοδος Μέγιστης Πιθανοφάνειας.....                    | 7         |
| Μέθοδος Εκτίμησης Ελάχιστης Απόστασης.....             | 8         |
| <b>1.9 Κριτήρια επιλογής στατιστικών μοντέλων.....</b> | <b>9</b>  |
| Κριτήριο πληροφορίας Akaike (AIC).....                 | 10        |
| Μπεϋζιανό κριτήριο πληροφορίας (BIC).....              | 10        |
| <b>1.10 Ζημιοκατανομές.....</b>                        | <b>11</b> |
| Κατανομή Pareto.....                                   | 11        |
| Κατανομή Lognormal.....                                | 12        |
| Κατανομή Weibull.....                                  | 13        |
| Κατανομή Γάμμα.....                                    | 13        |

|  |           |
|--|-----------|
| ΚΕΦΑΛΑΙΟ 2.....  | 15        |
| <b>2.1 Εισαγωγή.....</b>   | <b>15</b> |
| <b>2.2 Γέννηση και μέλη του συστήματος κατανομών του Burr.....</b>     | <b>16</b> |
| <b>2.3 Η κατανομή Burr τύπου XII.....</b>                              | <b>18</b> |
| Ροπές.....   | 21        |
| Συντελεστής ασυμμετρίας και κύρτωσης.....                              | 21        |
| Συμπεριφορά της δεξιάς ουράς.....                                      | 22        |
| Εκτιμητές μέγιστης πιθανοφάνειας.....                                  | 23        |
| Επέκταση της Burr XII κατανομής.....                                   | 24        |
| <b>2.4 Η κατανομή Burr τύπου II.....</b>                               | <b>25</b> |
| Διακριτή κατανομή Burr τύπου II.....                                   | 28        |
| <b>2.5 Η κατανομή Burr τύπου III.....</b>                              | <b>29</b> |
| Εκτιμητές μέγιστης πιθανοφάνειας.....                                  | 31        |
| Διακριτή κατανομή Burr τύπου III.....                                  | 32        |
| Επέκταση της Burr III κατανομής.....                                   | 33        |
| <b>2.6 Η κατανομή Burr τύπου X.....</b>                                | <b>33</b> |
| Η διπαραμετρική Burr X κατανομή.....                                   | 35        |
| <b>2.7 Η Κατανομή Burr τύπου IX.....</b>                               | <b>39</b> |
| Επέκταση της Burr IX κατανομής.....                                    | 40        |
| <b>2.8 Γνωστές κατανομές που προκύπτουν από την κατανομή Burr.....</b> | <b>41</b> |
| Παραλλαγές της Burr III κατανομής.....                                 | 41        |
| Παραλλαγές της τριπαραμετρικής Burr XII κατανομής.....                 | 42        |
| ΚΕΦΑΛΑΙΟ 3.....  | 46        |
| <b>3.1 Εισαγωγή.....</b>   | <b>46</b> |
| <b>3.2 Δεδομένα.....</b>   | <b>46</b> |
| <b>3.3 Ανάλυση δεδομένων.....</b>                                      | <b>47</b> |
| Δεδομένα Danish.....   | 47        |
| Δεδομένα Annual Rainfall Amount.....                                   | 53        |

|  |           |
|--|-----------|
| Δεδομένα Maximum Daily Precipitation ..... | 59        |
| <b>3.4 Συμπεράσματα .....</b>              | <b>66</b> |
| ΠΑΡΑΡΤΗΜΑ.....                             | 67        |
| ΒΙΒΛΙΟΓΡΑΦΙΑ .....                         | 73        |
| Ελληνική.....                              | 73        |
| Ξένη .....                                 | 73        |





# ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

---

|   |    |
|---|----|
| Πίνακας 2. 1 - Δημοφιλείς κατανομές που ανήκουν στο σύστημα κατανομών του Burr .....            | 18 |
| Πίνακας 2. 2 - Ειδικές περιπτώσεις της NMBIII κατανομής.....                                    | 42 |
| Πίνακας 2. 3 - Ειδικές περιπτώσεις της BXII-BXII κατανομής .....                                | 45 |
| Πίνακας 3. 1 - Περιγραφικά μέτρα δεδομένων Danish .....   | 47 |
| Πίνακας 3. 2 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr XII...         | 50 |
| Πίνακας 3. 3 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III ....       | 50 |
| Πίνακας 3. 4 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr IX....         | 50 |
| Πίνακας 3. 5 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal<br>.....  | 51 |
| Πίνακας 3. 6 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto .....        | 51 |
| Πίνακας 3. 7 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Danish .....                         | 51 |
| Πίνακας 3. 8 - Περιγραφικά μέτρα δεδομένων Annual Rainfall Amount.....                          | 54 |
| Πίνακας 3. 9 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr XII...         | 56 |
| Πίνακας 3. 10 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III ..        | 56 |
| Πίνακας 3. 11 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κανονικής κατανομής<br>..... | 56 |
| Πίνακας 3. 12 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal<br>..... | 57 |
| Πίνακας 3. 13 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto ....        | 57 |
| Πίνακας 3. 14 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Annual Rainfall Amount ...          | 57 |
| Πίνακας 3. 15 - Περιγραφικά μέτρα δεδομένων Maximum Daily Precipitation .....                   | 60 |
| Πίνακας 3. 16 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr XII.          | 61 |
| Πίνακας 3. 17 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III ..        | 62 |
| Πίνακας 3. 18 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr IX..          | 62 |
| Πίνακας 3. 19 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal<br>..... | 62 |
| Πίνακας 3. 20 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto ....        | 63 |

Πίνακας 3. 21 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Maximum Daily  
Precipitation.....63



# ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

---

|  |    |
|--|----|
| Σχήμα 1. 1 - Παραδείγματα συναρτήσεων κινδύνου .....   | 4  |
| Σχήμα 1. 2 – Λοξότητα μιας κατανομής .....   | 6  |
| Σχήμα 2. 1 - Συναρτήσεις πυκνότητας διάφορων κατανομών Burr τύπου XII .....                                    | 19 |
| Σχήμα 2. 2 - Συναρτήσεις κατανομής διάφορων κατανομών Burr τύπου XII.....                                      | 19 |
| Σχήμα 2. 3 - Συναρτήσεις επιβίωσης διάφορων κατανομών Burr τύπου XII.....                                      | 20 |
| Σχήμα 2. 4 - Συναρτήσεις κινδύνου διάφορων κατανομών Burr τύπου XII .....                                      | 20 |
| Σχήμα 2. 5 - Συναρτήσεις $g(x)$ για την κατανομή Burr τύπου II .....   | 26 |
| Σχήμα 2. 6 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου II.....                          | 27 |
| Σχήμα 2. 7 - Συναρτήσεις κινδύνου για την κατανομή Burr τύπου II.....  | 27 |
| Σχήμα 2. 8 - Συναρτήσεις πυκνότητας της κατανομής Burr τύπου III.....  | 30 |
| Σχήμα 2. 9 - Συναρτήσεις κατανομής της κατανομής Burr τύπου III.....   | 30 |
| Σχήμα 2. 10 - Συναρτήσεις $g(x)$ για την κατανομή Burr τύπου X.....  | 35 |
| Σχήμα 2. 11 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου X.....                          | 35 |
| Σχήμα 2. 12 - Συναρτήσεις $g(x)$ για την κατανομή Burr τύπου IX.....   | 40 |
| Σχήμα 2. 13 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου IX.....                         | 40 |
| Σχήμα 3. 1 - Θηκόγραμμα δεδομένων Danish.....  | 48 |
| Σχήμα 3. 2 - Ιστόγραμμα και κανονικό Q-Q διάγραμμα δεδομένων Danish .....                                      | 49 |
| Σχήμα 3. 3 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Danish .....                                       | 52 |
| Σχήμα 3. 4 - Ιστόγραμμα πυκνότητας και προσαρμοσμένη κατανομή Burr XII στα δεδομένα Danish .....               | 53 |
| Σχήμα 3. 5 - Θηκόγραμμα δεδομένων Annual Rainfall Amount.....  | 54 |
| Σχήμα 3. 6 - Ιστόγραμμα και κανονικό Q-Q διάγραμμα δεδομένων Annual Rainfall Amount                            | 55 |
| Σχήμα 3. 7 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Annual Rainfall Amount ...                         | 58 |
| Σχήμα 3. 8 - Ιστόγραμμα πυκνότητας και προσαρμοσμένη κατανομή Lognormal στα δεδομένα Annual Rainfall Data..... | 59 |
| Σχήμα 3. 9 - Θηκόγραμμα δεδομένων Maximum Daily Precipitation.....   | 60 |

|   |    |
|---|----|
| Σχήμα 3. 10 - Ιστόγραμμα και κανονικό Q-Q διάγραμμα δεδομένων Maximum Daily Precipitation.....                          | 61 |
| Σχήμα 3. 11 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Maximum Daily Precipitation.....                           | 64 |
| Σχήμα 3. 12 - Ιστόγραμμα πυκνότητας και προσαρμοσμένη κατανομή Lognormal στα δεδομένα Maximum Daily Precipitation ..... | 65 |



# ΚΕΦΑΛΑΙΟ 1

## Εισαγωγικές έννοιες στην αναλογιστική επιστήμη

---

### Εισαγωγή

Στο συγκεκριμένο κεφάλαιο παρατίθενται βασικές εισαγωγικές έννοιες, οι οποίες είναι απαραίτητες στην ανάλυση αποτελεσμάτων και εφαρμογών που αναφέρονται στα επόμενα κεφάλαια.

### 1.1 Τυχαίες Μεταβλητές

Έστω  $\Omega$  ο δειγματικός χώρος ενός πειράματος τύχης. Μια πραγματική συνάρτηση  $X: \Omega \rightarrow \mathbb{R}$  λέγεται τυχαία μεταβλητή, αν για κάθε υποσύνολο  $I$  του  $\mathbb{R}$  ( $I \subseteq \mathbb{R}$ ) το σύνολο (αντίστροφη εικόνα)  $X^{-1}(I)$  είναι ενδεχόμενο του  $\Omega$ , δηλαδή το σύνολο

$$X^{-1}(I) = \{\omega \in \Omega: X(\omega) \in I\} \subseteq \Omega$$

είναι ενδεχόμενο του  $\Omega$ .

Αν στο δειγματικό χώρο  $\Omega$  έχει οριστεί μια πιθανότητα  $P_\Omega$ , τότε μπορούμε να γράψουμε ότι για κάθε  $I \subseteq \mathbb{R}$

$$P(X \in I) = P_\Omega(X^{-1}(I)).$$

### Διακριτές τυχαίες μεταβλητές

Έστω  $X$  μια διακριτή τυχαία μεταβλητή ορισμένη σε ένα δ.χ.  $\Omega$ . Η συνάρτηση  $f$ , που ορίζεται από τη σχέση,

$$f(x) = P(X = x) = P(\{\omega \in \Omega: X(\omega) = x\}), \quad x \in \mathbb{R}$$

ονομάζεται συνάρτηση πιθανότητας (σ.π.) της τυχαίας μεταβλητής  $X$ .

Στην περίπτωση που το σύνολο τιμών της  $X$  είναι απείρως αριθμήσιμο σύνολο, δηλαδή  $R_X = \{x_1, x_2, \dots\}$ , τότε ισχύουν οι εξής ιδιότητες:

$$\text{I. } f(x) = 0, \quad x \notin R_X.$$

$$\text{II. } f(x_i) \geq 0, \quad i = 1, 2, \dots$$

$$\text{III. } \sum_{i=1}^{\infty} f(x_i) = 1.$$

Στην περίπτωση που το σύνολο τιμών της  $X$  είναι πεπερασμένο σύνολο, δηλαδή  $R_X = \{x_1, x_2, \dots, x_n\}$  τότε ισχύουν ανάλογες ιδιότητες.

## Συνεχείς τυχαίες μεταβλητές

Έστω  $X$  μια τυχαία μεταβλητή που ορίζεται σε κάποιο δ.χ.  $\Omega$ . Ας υποθέσουμε επίσης, ότι υπάρχει μια μη αρνητική πραγματική συνάρτηση  $f: R \rightarrow [0, \infty)$  τέτοια ώστε για κάθε υποσύνολο  $I$  του  $R$ , το οποίο μπορεί να γραφεί ως ένωση ενός πεπερασμένου ή απείρως αριθμήσιμου πλήθους διαστημάτων, ισχύει

$$P(X \in I) = \int_I f(x) dx$$

Τότε, η τυχαία μεταβλητή  $X$  λέγεται **απόλυτα συνεχής** (ή απλά συνεχής) και η συνάρτηση  $f$  συνάρτηση πυκνότητας πιθανότητας ή απλά συνάρτηση πυκνότητας της  $X$ .  
Ισχύουν τα εξής:

$$\text{I. } f(x) \geq 0, \quad x \in R.$$

$$\text{II. } \int_{-\infty}^{\infty} f(x) dx = 1.$$

## 1.2 Συνάρτηση Κατανομής

Έστω  $X$  μια τυχαία μεταβλητή, που ορίζεται σε κάποιο δ.χ.  $\Omega$ . Η πραγματική συνάρτηση  $F: R \rightarrow R$  που δίνεται από τον τύπο

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega: X(\omega) \leq x\}), \quad x \in R$$

ονομάζεται (αθροιστική) συνάρτηση κατανομής της  $X$ .

Για μια διακριτή τυχαία μεταβλητή που παίρνει τιμές  $x_1, x_2, \dots$  η συνάρτηση κατανομής ισούται με

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i).$$



Για μια συνεχή τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας  $f$  η αντίστοιχη συνάρτηση κατανομής ισούται με

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt.$$

Ισχύουν τα εξής:

- $0 \leq F(x) \leq 1$ , με  $0 = F(-\infty)$  και  $1 = F(+\infty)$ .
- Αν  $x_1 < x_2$ , τότε  $F(x_1) \leq F(x_2)$ .
- Η  $F(x)$  είναι δεξιά συνεχής συνάρτηση.

### 1.3 Συνάρτηση Επιβίωσης

Έστω  $X$  μια συνεχής τ.μ. με σύνολο τιμών στο διάστημα  $[0, \infty)$ . Τότε η συνάρτηση επιβίωσης ορίζεται από τον τύπο

$$S(x) = P(X > x), \quad x \geq 0.$$

Γνωρίζοντας ότι  $F(x) = P(X \leq x)$ , τότε η συνάρτηση επιβίωσης γράφεται ως

$$S(x) = 1 - F(x) = P(X > x) = \int_x^{\infty} f(u) du.$$

Επίσης, ισχύει

$$f(x) = -\frac{dS(x)}{dx}.$$

Η  $S(x)$  είναι μια φθίνουσα συνάρτηση του  $x$  με τις κάτωθι ιδιότητες:

- $S(x) = 1$  για  $x = 0$ .
- $S(x) = 0$  για  $x \rightarrow \infty$ .

Η γραφική παράσταση της  $S(x)$  συναρτήσει του  $x$  ονομάζεται καμπύλη επιβίωσης (survival curve).

### 1.4 Συνάρτηση Κινδύνου

Έστω  $X$  μια συνεχής τ.μ. με σύνολο τιμών στο διάστημα  $[0, \infty)$ . Τότε η συνάρτηση κινδύνου (hazard rate) ορίζεται ως εξής:

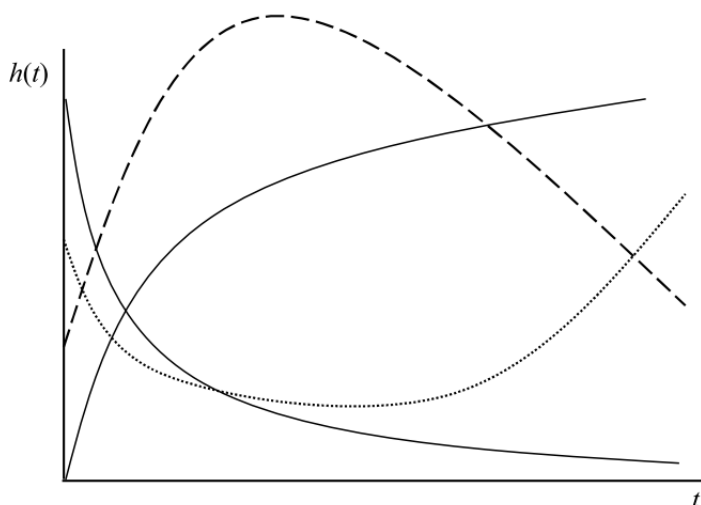
$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x | X > x)}{\Delta x}.$$

Από την παραπάνω σχέση προκύπτει ότι

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)} = \frac{-S'(x)}{S(x)}.$$

Η συνάρτηση κινδύνου είναι γνωστή και ως στιγμιαίος ρυθμός αποτυχίας (instantaneous failure rate) ή ως δεσμευμένος ρυθμός θνησιμότητας (conditional mortality).

Η συμπεριφορά της  $h(x)$  ποικίλει. Μπορεί να αυξάνεται, να μειώνεται, να μένει σταθερή ή να έχει μια πιο περίπλοκη μορφή. Στο παρακάτω σχήμα φαίνονται διάφορες μορφές της συνάρτησης κινδύνου.



Σχήμα 1. 1 - Παραδείγματα συναρτήσεων κινδύνου

(Αντζουλάκος, 2018)

## 1.5 Ροπές μιας τυχαίας μεταβλητής

Η ροπή τάξης  $k$  μιας τ.μ  $X$  (γύρω από το 0) δίνεται από τον τύπο

$$\mu_X^k = E(X^k) = \begin{cases} \sum x^k P(X = x), & \text{αν η } X \text{ είναι διακριτή τ. μ.} \\ \int_{-\infty}^{\infty} x^k f_X(x) dx, & \text{αν η } X \text{ είναι συνεχής τ. μ.} \end{cases}$$

όπου  $f_X(x)$  συμβολίζει την συνάρτηση πυκνότητας της συνεχούς τ.μ.  $X$ . Η ροπή πρώτης τάξης της  $X$  είναι η μέση τιμή της και θα συμβολίζεται απλά ως  $\mu_X^1 = \mu_X = E(X)$ .

Η κεντρική ροπή τάξης  $k$  μιας τ.μ.  $X$  δίνεται από τον τύπο

$$\mu_X'^k = E((X - \mu_x)^k) = \begin{cases} \sum_x (x - \mu_x)^k P(X = x), & \text{αν η } X \text{ είναι διακριτή τ. μ.} \\ \int_{-\infty}^{\infty} (x - \mu_x)^k f_X(x) dx, & \text{αν η } X \text{ είναι συνεχής τ. μ.} \end{cases}$$

Η κεντρική ροπή δεύτερης τάξης είναι η διακύμανση της  $X$ , δηλαδή  $\mu_X'^2 = \sigma_X^2 = Var(X)$ .

Η (καθοδική) παραγοντική ροπή τάξης  $k$  μιας τ.μ.  $X$  δίνεται από τον τύπο

$$\mu_X^{(k)} = E(X(X-1) \dots (X-k+1)) = \begin{cases} \sum_x x(x-1) \dots (x-k+1) P(X=x), & \text{αν η } X \text{ είναι διακριτή τ. μ.} \\ \int_{-\infty}^{\infty} x(x-1) \dots (x-k+1) f_X(x) dx, & \text{αν η } X \text{ είναι συνεχής τ. μ.} \end{cases}$$

Σημειώνουμε ότι  $\mu_X^0 = \mu_X'^0 = \mu_X^{(0)} = 1$ .

Ειδικότερα ισχύει ότι,

$$V(X) = \sigma^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu_X^2.$$

Η ποσότητα

$$\sigma = +\sqrt{Var(X)}$$

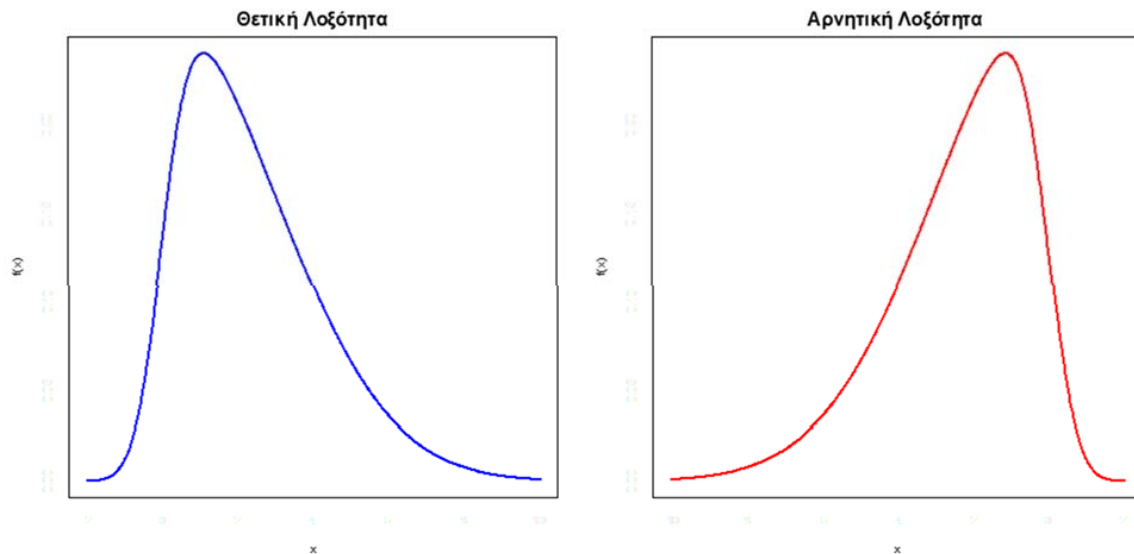
καλείται τυπική απόκλιση (mean square deviation or standard deviation) της  $X$ .

## 1.6 Ασυμμετρία

Η κατανομή ενός πληθυσμού μπορεί να είναι είτε συμμετρική, είτε μη συμμετρική. Στην πρώτη περίπτωση η κορυφή, η διάμεσος και η μέση τιμή συμπίπτουν. Στις άλλες περιπτώσεις, ένα από τα τμήματα στα οποία χωρίζει την κατανομή η κορυφή περιέχει περισσότερες παρατηρήσεις από το άλλο. Υπάρχουν δύο ειδών ασυμμετρίες, η θετική ασυμμετρία (λοξότητα προς τα δεξιά) στην οποία οι περισσότερες παρατηρήσεις, καθώς επίσης και η διάμεσος και η μέση τιμή, βρίσκονται δεξιά της κορυφής και η αρνητική ασυμμετρία (λοξότητα προς τα αριστερά) στην οποία οι περισσότερες παρατηρήσεις, όπως και η διάμεσος με τη μέση τιμή, βρίσκονται αριστερά της κορυφής.

Το συνηθέστερο αριθμητικό μέτρο ασυμμετρίας είναι ο συντελεστής ασυμμετρίας ο οποίος ορίζεται ως

$$\beta = E \left[ \left( \frac{X - \mu}{\sigma} \right)^3 \right].$$



Σχήμα 1. 2 – Λοξότητα μιας κατανομής

## 1.7 Κύρτωση

Μια κατανομή η οποία έχει σχετικά μεγάλη μέγιστη συχνότητα (κορυφή), δηλαδή μεγάλη συγκέντρωση των τιμών της γύρω από το μέσο, λέγεται λεπτόκυρτη (leptokurtic). Σε αντίθετη περίπτωση, αν δηλαδή η μέγιστη συχνότητα της είναι σχετικά μικρή, λέγεται πλατύκυρτη (platykurtic). Οι κατανομές που προσεγγίζονται από την κανονική κατανομή, χαρακτηρίζονται ως μεσόκυρτες (mesokurtic).

Μέτρο για την κύρτωση μιας κατανομής είναι ο συντελεστής κύρτωσης

$$\gamma = E \left[ \left( \frac{X - \mu}{\sigma} \right)^4 \right].$$

Επειδή για κανονικές κατανομές έχουμε  $\gamma = 3$ , συνηθίζεται να μετράμε την κυρτότητα με την διαφορά  $\gamma - 3$ , η οποία για λεπτόκυρτες κατανομές παίρνει θετικές τιμές (θετική κύρτωση), ενώ για πλατύκυρτες κατανομές γίνεται αρνητική (αρνητική κύρτωση).

## 1.8 Μέθοδοι Εκτίμησης Παραμέτρων

Ένα συχνό πρόβλημα που εμφανίζεται στις στατιστικές μελέτες είναι η εύρεση εκτιμητριών των παραμέτρων μιας κατανομής πιθανότητας. Υπάρχουν διάφορες μέθοδοι για την επίλυση αυτού του προβλήματος όπως οι εκτιμητές μέγιστης πιθανοφάνειας (EMΠ), εκτιμητές με τη μέθοδο των Ροπών (EMP), εκτιμητές ελάχιστης απόστασης (ΕΕΑ), εκτιμητές minimax, εκτιμητές Bayes. Θα σταθούμε όμως λίγο παραπάνω στους EMΠ και στους ΕΕΑ, καθώς είναι οι μέθοδοι που χρησιμοποιήθηκαν στην συγκεκριμένη εργασία (δείτε Αντζουλάκος (2018)).

### Μέθοδος Μέγιστης Πιθανοφάνειας

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

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

Αν για κάθε δειγματοσημείο  $\mathbf{x}$  το διάνυσμα

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{x}) = (\hat{\theta}_1(\mathbf{x}), \hat{\theta}_2(\mathbf{x}), \dots, \hat{\theta}_k(\mathbf{x}))$$

με  $\hat{\boldsymbol{\theta}} \in \Theta$ , ικανοποιεί τη σχέση

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

ή ισοδύναμα τη σχέση

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

όπου

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

τότε η στατιστική συνάρτηση  $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{x}) = (\hat{\theta}_1(\mathbf{x}), \hat{\theta}_2(\mathbf{x}), \dots, \hat{\theta}_k(\mathbf{x}))$  ονομάζεται εκτιμητής μεγίστης πιθανοφάνειας (maximum likelihood estimator) της παραμέτρου  $\boldsymbol{\theta}$ .

Ο εκτιμητής μεγίστης πιθανοφάνειας (ΕΜΠ) προκύπτει συνήθως από τη λύση του συστήματος των εξισώσεων

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

ή ισοδύναμα

$$\mathbf{U}(\boldsymbol{\theta}) = (U_1(\boldsymbol{\theta}), U_2(\boldsymbol{\theta}), \dots, U_k(\boldsymbol{\theta})) = \mathbf{0}$$

αρκεί ο πίνακας των δευτέρων μερικών παραγώγων

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

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

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

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

και

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

Ο εκτιμητής μεγίστης πιθανοφάνειας  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1(\mathbf{X}), \hat{\theta}_2(\mathbf{X}))$  της παραμέτρου  $\boldsymbol{\theta}$  προκύπτει από τη λύση του συστήματος

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

που είναι η

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

## Μέθοδος Εκτίμησης Ελάχιστης Απόστασης

Με αυτή τη γενική μέθοδο ελαχιστοποιείται μια επιλεγμένη «απόσταση» μεταξύ της θεωρητικής και της εμπειρικής συνάρτησης κατανομής. Έστω  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  ένα τυχαίο δείγμα από πληθυσμό με συνάρτηση πυκνότητας  $f(x; \boldsymbol{\theta})$  και συνάρτηση κατανομής  $F(x; \boldsymbol{\theta})$ , όπου  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ .

Η μέθοδος Cramer-von Mises (CvM), βρίσκει το  $\hat{\boldsymbol{\theta}}(\mathbf{x}) = (\hat{\theta}_1(\mathbf{x}), \hat{\theta}_2(\mathbf{x}), \dots, \hat{\theta}_k(\mathbf{x}))$  που ελαχιστοποιεί τη στατιστική συνάρτηση

$$W_n^2 = n \int_{-\infty}^{\infty} [F(x; \boldsymbol{\theta}) - F_n(x)]^2 dF(x; \boldsymbol{\theta}).$$

Η μέθοδος Anderson-Darling (AD) βρίσκει το  $\hat{\boldsymbol{\theta}}(\mathbf{x})$  που ελαχιστοποιεί την στατιστική συνάρτηση

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{[F(x; \boldsymbol{\theta}) - F_n(x)]^2}{F(x; \boldsymbol{\theta})[1 - F(x; \boldsymbol{\theta})]} dF(x; \boldsymbol{\theta}).$$

Η μέθοδος Kolmogorov-Smirnov (KS) βρίσκει το  $\hat{\boldsymbol{\theta}}(\mathbf{x})$  που ελαχιστοποιεί τη στατιστική συνάρτηση

$$D_n = \sup_{x \in \mathbb{R}} |F(x; \boldsymbol{\theta}) - F_n(x)|.$$

Στους παραπάνω τύπους  $F_n(x)$  δηλώνει την εμπειρική συνάρτηση κατανομής.

## 1.9 Κριτήρια επιλογής στατιστικών μοντέλων

Τα στατιστικά μοντέλα είναι ουσιαστικά κατανομές πιθανοτήτων, όπου τα παρατηρούμενα δεδομένα χρησιμοποιούνται για την προσέγγιση της πραγματικής κατανομής από την οποία προέρχονται. Επομένως, οποιοσδήποτε διαθέτει ένα σύνολο δεδομένων, μπορεί να επιλέξει ανάμεσα σε πολλά μοντέλα, αλλά πρέπει να χρησιμοποιήσει το πιο κατάλληλο από αυτά, ώστε να μπορεί να τα περιγράψει σωστά.

Είναι σαφές ότι οι μέθοδοι που χρησιμοποιούνται για την επιλογή στατιστικών μοντέλων διαφέρουν ανάλογα με τις ανάγκες, τους στόχους και τις χρήσεις των μοντέλων που επιλέγονται σε κάθε περίπτωση. Τα στατιστικά μοντέλα στην πράξη επιλέγονται συνήθως με βάση την τιμή του κατάλληλου κριτηρίου πληροφορίας. Το κριτήριο πληροφορίας ορίζεται ως ένας μηχανισμός που δίνει σε κάθε υποψήφιο μοντέλο μια "βαθμολογία" χρησιμοποιώντας τα διαθέσιμα δεδομένα. Έτσι δημιουργείται ένας κατάλογος των καλύτερων μοντέλων ξεκινώντας με το καλύτερο μοντέλο και τελειώνοντας με το χειρότερο μοντέλο. Για να προχωρήσουμε στην επιλογή του κατάλληλου μοντέλου είναι απαραίτητο να εξεταστεί ποιος είναι ο σκοπός της στατιστικής μοντελοποίησης και πού θα

χρησιμοποιηθεί το μοντέλο. Παρακάτω παρατίθενται κάποια από τα πιο βασικά κριτήρια επιλογής στατιστικών μοντέλων.

### Κριτήριο πληροφορίας Akaike (AIC)

Το κριτήριο πληροφορίας Akaike προτάθηκε από τον Akaike (1974) και είναι μία από τις πιο σημαντικές και δημοφιλείς μεθόδους επιλογής μοντέλων. Βρίσκει εφαρμογή σε όλες τις περιπτώσεις προβλημάτων γραμμικής παλινδρόμησης και μοντελοποίησης χρονοσειρών, αναλύσεις χρόνων επιβίωσης και γενικά σε παραμετρικές συγκρίσεις μοντέλων. Ενδεικτικό είναι το γεγονός ότι τα περισσότερα στατιστικά πακέτα χρησιμοποιούν το AIC. Ο τύπος υπολογισμού του AIC έχει συνήθως την εξής μορφή

$$AIC = -2\log L(\hat{\theta}) + 2\dim(\theta)$$

όπου  $L(\hat{\theta})$  είναι η μέγιστη τιμή της πιθανοφάνειας και  $\dim(\theta)$  η διάσταση του διανύσματος των παραμέτρων  $\theta$ .

Στην πράξη, αφού πρώτα υπολογιστεί η τιμή AIC κάθε υποψήφιου μοντέλου, επιλέγεται αυτό με τη χαμηλότερη τιμή, επειδή είναι λογικό να είναι το πιο κοντινό στο πραγματικό μοντέλο.

### Μπεϋζιανό κριτήριο πληροφορίας (BIC)

Το μπεϋζιανό κριτήριο πληροφορίας (BIC) ή κριτήριο Schwarz, αποτελεί προϊόν της μπεϋζιανής Στατιστικής και αναπτύχθηκε από τον Schwarz (1978). Όπως το κριτήριο AIC, έτσι και το κριτήριο BIC θεωρείται ένα από τα πιο δημοφιλή και ευρέως χρησιμοποιούμενα κριτήρια. Ο τύπος υπολογισμού του BIC έχει συνήθως την εξής μορφή

$$BIC = -2\log L(\hat{\theta}) + \log n \cdot \dim(\theta)$$

όπου  $L(\hat{\theta})$  είναι η μέγιστη τιμή της πιθανοφάνειας,  $n$  το μέγεθος του δείγματος και  $\dim(\theta)$  η διάσταση του διανύσματος των παραμέτρων  $\theta$ . Όμοια με την περίπτωση του κριτηρίου AIC, επιλέγεται το μοντέλο με την μικρότερη τιμή BIC, καθώς αυτό προσεγγίζει καλύτερα τα δεδομένα μας.

Στην εργασία χρησιμοποιήθηκαν τα παραπάνω δύο κριτήρια πληροφορίας. Άλλα κριτήρια πληροφορίας είναι (α) το κριτήριο πληροφορίας απόκλισης (deviance information criterion) που αναπτύχθηκε από τους Spiegelhalter et al. (2002) και χρησιμοποιείται συχνά στη μπεϋζιανή ανάλυση σε πολύπλοκα ιεραρχικά μοντέλα (β) το extended κριτήριο



πληροφορίας που προέκυψε από την εφαρμογή της μεθόδου Bootstrap (Efron (1983)), (γ) το Mallows  $C_p$  που προτάθηκε από τον Mallows (1973) και χρησιμοποιείται συχνά ως δείκτης για την αξιολόγηση της καλής προσαρμογής μοντέλων παλινδρόμησης (δ) το κριτήριο πληροφορίας Hannan-Quinn που αναπτύχθηκε από τους Hannan & Quinn (1979) και αρχικά αποσκοπούσε στον προσδιορισμό της τάξης των αυτοπαλινδρομων χρονοσειρών, και (ε) το τελικό σφάλμα πρόβλεψης (final prediction error (FPE)) που αναπτύχθηκε αρχικά από τον Akaike (1969, 1970) στο πλαίσιο της ανάλυσης χρονοσειρών με σκοπό την επιλογή της τάξης ενός αυτοπαλινδρομου μοντέλου.

## 1.10 Ζημιοκατανομές

Με τον όρο ζημιοκατανομές (ή κατανομές απώλειας), αναφερόμαστε στις θετικές κατανομές, δηλαδή στις κατανομές που έχουν πεδίο τιμών το  $(0, \infty)$ . Ένα κοινό στοιχείο που διαθέτουν όλες οι κατανομές απώλειας είναι ότι παρουσιάζουν θετική ασυμμετρία. Αυτό το φαινόμενο παρατηρείται όταν η πλειοψηφία των παρατηρήσεων βρίσκονται δεξιά της κορυφής. Ένα ακόμη κοινό χαρακτηριστικό των κατανομών αυτών είναι ότι συνήθως έχουν βαριές ουρές.

Οι ζημιοκατανομές είναι στενά συνδεδεμένες με την επιστήμη του αναλογισμού. Μέσω των ζημιοκατανομών παρατηρείται η κατανομή του μεγέθους των απωλειών, δεδομένου ότι έχει επέλθει ο ασφαλιστικός κίνδυνος. Ο αναλογιστής πρέπει να γνωρίζει την πιθανότητα εμφάνισης ζημιών, τον αναμενόμενο αριθμό των ζημιών, όπως και το αναμενόμενο ύψος των ζημιών. Κάποιες από τις πιο γνωστές κατανομές ζημιών είναι η Pareto, η Lognormal, η Burr, η Weibull και άλλες (δείτε Τζαβελάς (2022)).

### Κατανομή Pareto

Η κατανομή ονομάστηκε έτσι εξαιτίας ενός Ιταλού πολιτικού μηχανικού, οικονομολόγου και κοινωνιολόγου, ονόματι Vilfredo Pareto. Πρόκειται για μια κατανομή πιθανοτήτων, η οποία χρησιμοποιείται για την περιγραφή κοινωνικών, επιστημονικών, αναλογιστικών και άλλων τύπων φαινομένων.

Έστω  $X$  μια τ.μ με κατανομή Pareto (συμβ  $P(a, s)$ ), τότε η πιθανότητα το  $X$  να είναι μεγαλύτερο από το  $x$ , δηλαδή η συνάρτηση επιβίωσης, δίνεται από τον τύπο

$$S(x) = \Pr(X > x) = \begin{cases} \left(\frac{s}{x}\right)^a, & x \geq s \\ 1, & x < s \end{cases}$$

όπου  $s$  είναι η ελάχιστη δυνατή τιμή του  $X$  και  $a$  είναι μια θετική παράμετρος.

Η συνάρτηση κατανομής της κατανομής Pareto, δίνεται από τον ακόλουθο τύπο

$$F(x) = \Pr(X \leq x) = \begin{cases} 1 - \left(\frac{s}{x}\right)^a, & x \geq s \\ 0, & x < s \end{cases}.$$

Αυτό συνεπάγεται ότι η συνάρτηση πυκνότητας πιθανότητας είναι

$$f(x) = \begin{cases} \frac{as^a}{x^{a+1}}, & x > s \\ 0, & x < s \end{cases}.$$

Για τη  $n$ -οστή ροπή μια τυχαίας μεταβλητής  $X$  που έχει την κατανομή Pareto, έχουμε

$$E(X^n) = \frac{as^n}{a-n}.$$

Έτσι η μέση τιμή και η διακύμανση είναι

$$\mu = \frac{as}{a-1}$$

$$\sigma^2 = \frac{as^2}{(a-1)^2(a-2)}.$$

## Κατανομή Lognormal

Στη θεωρία πιθανοτήτων, μια λογαριθμοκανονική κατανομή, είναι μια συνεχής κατανομή πιθανότητας μιας τυχαίας μεταβλητής, της οποίας ο λογάριθμος κατανέμεται κανονικά. Έτσι, εάν μια τ.μ  $X$  ακολουθεί λογαριθμοκανονική κατανομή, τότε η τ.μ  $Y = \ln X$  ακολουθεί κανονική κατανομή. Ισοδύναμα, εάν η τ.μ  $Y$  ακολουθεί κανονική κατανομή, τότε η τ.μ  $X = \exp(Y)$  ακολουθεί λογαριθμοκανονική κατανομή.

Η τυχαία μεταβλητή  $X$  ακολουθεί τη λογαριθμοκανονική κατανομή (lognormal) με παραμέτρους  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ , αν η συνάρτηση πυκνότητάς της δίνεται από τον τύπο

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

και γράφουμε ότι  $X \sim LN(\mu, \sigma)$ . Για τη μέση τιμή, τη διακύμανση, και τη ροπή τάξης  $k$  της  $X$  έχουμε

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

$$E(X^k) = \exp\left(k\mu + \frac{k^2\sigma^2}{2}\right).$$

### Κατανομή Weibull

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

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right), \quad x > 0,$$

και γράφουμε ότι  $X \sim W(k, \lambda)$ . Για τη μέση τιμή, τη διακύμανση, και τη ροπή τάξης  $r$  της  $X$  έχουμε

$$\mu = \lambda \Gamma\left(1 + \frac{1}{k}\right), \quad \sigma^2 = \lambda^2 \left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right),$$

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

### Κατανομή Γάμμα

Η τυχαία μεταβλητή  $X$  ακολουθεί την κατανομή Γάμμα (Gamma) με παραμέτρους  $a, \theta > 0$ , αν η συνάρτηση πυκνότητάς της δίνεται από τον τύπο

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

και γράφουμε ότι  $X \sim G(a, \theta)$ . Για τη μέση τιμή, τη διακύμανση, και τη ροπή τάξης  $k$  της  $X$  έχουμε ότι

$$\mu = a\theta, \quad \sigma^2 = a\theta^2, \quad E(X^k) = \frac{\theta^k \Gamma(a+k)}{\Gamma(a)}.$$



# ΚΕΦΑΛΑΙΟ 2

## Σύστημα κατανομών Burr

---

### 2.1 Εισαγωγή

Η κανονική κατανομή  $N(\mu, \sigma^2)$  είναι η πιο σημαντική και ευρέως χρησιμοποιούμενη κατανομή στη θεωρία πιθανοτήτων και τη στατιστική αφού έχει σημαντικές ιδιότητες και εφαρμογές. Οι δύο παράμετροί της δηλώνουν τη μέση τιμή και τη διακύμανσή της. Η συνάρτηση πυκνότητας πιθανότητάς της είναι συμμετρική και έχει σχήμα καμπάνας. Ωστόσο, είναι κατάλληλη μόνο για την ανάλυση δεδομένων τα οποία είναι συμμετρικά ή περίπου συμμετρικά. Στην πραγματικότητα, τα δεδομένα που λαμβάνονται στην πράξη, και ειδικότερα στην αναλογιστική επιστήμη, δεν είναι συμμετρικά. Κατά συνέπεια, η κανονική κατανομή δεν συνιστάται για την ανάλυση μη συμμετρικών δεδομένων, επειδή μπορεί να οδηγήσει σε ανακριβή αποτελέσματα.

Στον αναλογισμό και τη διαχείριση κινδύνων η επιλογή της κατάλληλης κατανομής για την ανάλυση μη συμμετρικών δεδομένων όχι μόνο λαμβάνει υπόψη τη λοξότητα της κατανομής αλλά και τη συμπεριφορά της ουράς της. Μια κατανομή που μπορεί να είναι μια εναλλακτική λύση για την ανάλυση λοξών προς τα δεξιά δεδομένων με βαριά ουρά είναι η κατανομή Burr. Χρησιμοποιείται ευρέως και ειδικότερα στη μοντελοποίηση του χρόνου επιβίωσης σε ιατρικές εφαρμογές. Εάν θέλουμε να μοντελοποιήσουμε δεδομένα χρησιμοποιώντας την κατανομή Burr, οι πληροφορίες για τις παραμέτρους είναι σημαντικές για να συμπεράνουμε τις ιδιότητες της κατανομής. Ωστόσο, στην πράξη, οι παράμετροι είναι άγνωστες. Επομένως, πρέπει να τις εκτιμήσουμε από παρατηρήσεις που έχουμε στη διάθεσή μας χρησιμοποιώντας μια κατάλληλη στατιστική μεθοδολογία.

Ο Burr (1942) εισήγαγε δώδεκα αθροιστικές συναρτήσεις κατανομής που θα μπορούσαν να ταιριάζουν σε πραγματικά δεδομένα. Στο παρόν κεφάλαιο θα παρουσιαστούν οι σημαντικότερες κατανομές που ανήκουν στην οικογένεια κατανομών του Burr. Πιο συγκεκριμένα, στην Παράγραφο 2.2 παρατίθεται η διαφορική εξίσωση βάσει της οποίας γεννήθηκε η οικογένεια κατανομών Burr. Παρακάτω, στην Παράγραφο 2.3 αναλύεται η βασικότερη κατανομή της οικογένειας των κατανομών Burr, η κατανομή Burr XII. Στην συνέχεια ακολουθεί η Παράγραφος 2.4 στην οποία γίνεται ανάλυση της Burr II κατανομής και αμέσως μετά στην Παράγραφο 2.5 παρουσιάζεται η Burr III κατανομή. Η Παράγραφος 2.6 είναι αφιερωμένη στην κατανομή Burr X. Στην Παράγραφο 2.7

αναλύεται η κατανομή Burr IX και τέλος στην Παράγραφο 2.8 παρουσιάζονται κάποιες γνωστές κατανομές που προκύπτουν από την κατανομή Burr.

## 2.2 Γέννηση και μέλη του συστήματος κατανομών του Burr

Ο Burr (1942) εισήγαγε τη διαφορική εξίσωση

$$\frac{dF(x)}{dx} = F(x)(1 - F(x))g(x, F(x))$$

όπου  $F(x)$  είναι μια συνάρτηση κατανομής,  $g(x, F(x))$  είναι μια συνάρτηση που είναι θετική για  $0 \leq F(x) \leq 1$  και το εύρος τιμών του  $x$  είναι το επιθυμητό.

Για  $g(x, F(x)) = (c_1 + c_2x + c_3x^2)^{-1}$ , και αντικαθιστώντας τις συναρτήσεις  $F(x)$  και  $1 - F(x)$  με τις  $f(x)$  και  $a - x$  προκύπτει η διαφορική εξίσωση

$$\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{a - x}{c_1 + c_2x + c_3x^2}$$

που είναι η διαφορική εξίσωση που ικανοποιεί το σύστημα κατανομών του Pearson (1895).

Για  $g(x, y) = g(x)$  η διαφορική εξίσωση του Burr παίρνει τη μορφή

$$\frac{dF(x)}{dx} = F(x)(1 - F(x))g(x)$$

οπότε

$$\frac{dF(x)}{F(x)(1 - F(x))} = g(x)dx$$

και

$$\frac{dF(x)}{F(x)} + \frac{dF(x)}{1 - F(x)} = g(x)dx.$$

Με ολοκλήρωση παίρνουμε

$$\ln\left(\frac{1 - F(x)}{F(x)}\right) = -\int g(x)dx$$

οπότε

$$F(x) = \frac{1}{1 + \exp(-\int g(x)dx)}.$$

Από την παραπάνω σχέση προκύπτουν οι συναρτήσεις κατανομής του συστήματος των κατανομών του Burr ανάλογα με τη επιλογή που θα γίνει για τη μορφή της συνάρτησης  $g(x)$ . Στον Πίνακα 2.1 δίνονται οι συναρτήσεις κατανομής των 12 πιο γνωστών κατανομών του συστήματος κατανομών Burr και οι αντίστοιχες συναρτήσεις  $g(x)$  που τις παράγουν.

| Τύπος | $F(x)$                             | $g(x)$  | Τιμές του $x$                                |
|-------|------------------------------------|---|--|
| I     | $x$                                | $\frac{1}{x(1-x)}$  | (0,1)  |
| II    | $(1 + e^{-x})^{-k}$                | $\frac{ke^{-x}(e^{-x} + 1)^{k-1}}{(e^{-x} + 1)^k - 1}$  | $(-\infty, \infty)$                          |
| III   | $(1 + x^{-k})^{-r}$                | $\frac{kr}{x(1 - (x^{-k} + 1)^{-r})(x^k + 1)}$  | $(0, \infty)$                                |
| IV    | $[1 + (x^{-1}(c - x))^{1/c}]^{-k}$ | $\frac{k\left(\frac{c-x}{x}\right)^{(1/c)-1}}{x^2\left(1 - \left(\left(\frac{c-x}{x}\right)^{1/c} + 1\right)^{-k}\right)\left(1 + \left(\frac{c-x}{x}\right)^{1/c}\right)}$ | $(0, c)$                                     |
| V     | $(1 + ke^{-\tan(x)})^{-c}$         | $-\frac{kc\left(\left(ke^{-\tan(x)} + 1\right)^{-c} - 1\right)(\tan^2(x) + 1)}{k + e^{\tan(x)}}$  | $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ |
| VI    | $(1 + ke^{-c\sinh(x)})^{-r}$       | $\frac{ckr \cosh(x)e^{-\sinh(x)}}{\left(\left(ke^{-c\sinh(x)} + 1\right)^{-r} - 1\right)\left(ke^{-c\sinh(x)} + 1\right)}$  | $(-\infty, \infty)$                          |
| VII   | $2^{-r}[1 + \tanh(x)]^r$           | $\frac{r(\tanh(x) - 1)}{\left(\frac{\tanh(x)+1}{2}\right)^r - 1}$   | $(-\infty, \infty)$                          |
| VIII  | $(2\pi^{-1}\tan^{-1}(e^x))^{-r}$   | $\frac{re^x}{\tan^{-1}(e^x)(e^{2x} + 1)\left(1 - \left(\frac{2\tan^{-1}(e^x)}{\pi}\right)^r\right)}$  | $(-\infty, \infty)$                          |
| IX    | $1 - (e^x + 1)^{-r}$               | $\frac{re^x(e^x + 1)^{r-1}}{(e^x + 1)^r - 1}$   | $(-\infty, \infty)$                          |
| X     | $(1 - e^{-x^2})^k$                 | $\frac{2kxe^{-x^2}}{\left(1 - (1 - e^{-x^2})^k\right)(1 - e^{-x^2})}$   | $(0, \infty)$                                |

|            |                                   |   |        |
|------------|-----------------------------------|---|--------|
| <b>XI</b>  | $(x - (2\pi)^{-1}\sin(2\pi x))^k$ | $\frac{k(\cos(2\pi x) - 1)}{\left(x - \frac{\sin(2\pi x)}{2\pi}\right)\left(\left(x - \frac{\sin(2\pi x)}{2\pi}\right)^k - 1\right)}$ | (0,1)  |
| <b>XII</b> | $1 - (1 + x^c)^{-k}$              | $\frac{ckx^{c-1}(1 + x^c)^{k-1}}{(1 + x^c)^k - 1}$  | (0, ∞) |

Πίνακας 2. 1 - Δημοφιλείς κατανομές που ανήκουν στο σύστημα κατανομών του Burr

Για περισσότερες πληροφορίες ο αναγνώστης παραπέμπεται στο Κεφάλαιο 3 με τίτλο “Burr System of Frequency Distributions” του βιβλίου των Singh and Zhang (2020).

### 2.3 Η κατανομή Burr τύπου XII

Η κατανομή Burr εισήχθη για πρώτη φορά το 1942 από τον Irving Wingate Burr (δείτε επίσης Kumar (2017) και Hakim et al. (2021)). Ήταν αρχικά γνωστή ως κατανομή Burr τύπου XII, η οποία αποτελεί έναν από τους δώδεκα τύπους συνεχών κατανομών του συστήματος κατανομών του Burr. Η κατανομή Burr τύπου XII (ή κατανομή Singh–Maddala) είναι μια μονοκόρυφη κατανομή με μεγάλη ποικιλία σχημάτων. Χρησιμοποιείται για τη μοντελοποίηση πολλών φαινομένων, συμπεριλαμβανομένων των τιμών των καλλιεργειών, του εισοδήματος των νοικοκυριών, των τιμών δικαιωμάτων προαίρεσης, του κινδύνου ασφάλισης, και άλλων. Αν και υπάρχουν και άλλες κατανομές Burr, ο όρος "κατανομή Burr" αναφέρεται συνήθως στον τύπο XII.

Η κατανομή Burr τύπου XII με παραμέτρους  $c$  και  $k$  έχει συνάρτηση πυκνότητας που δίνεται από τον τύπο

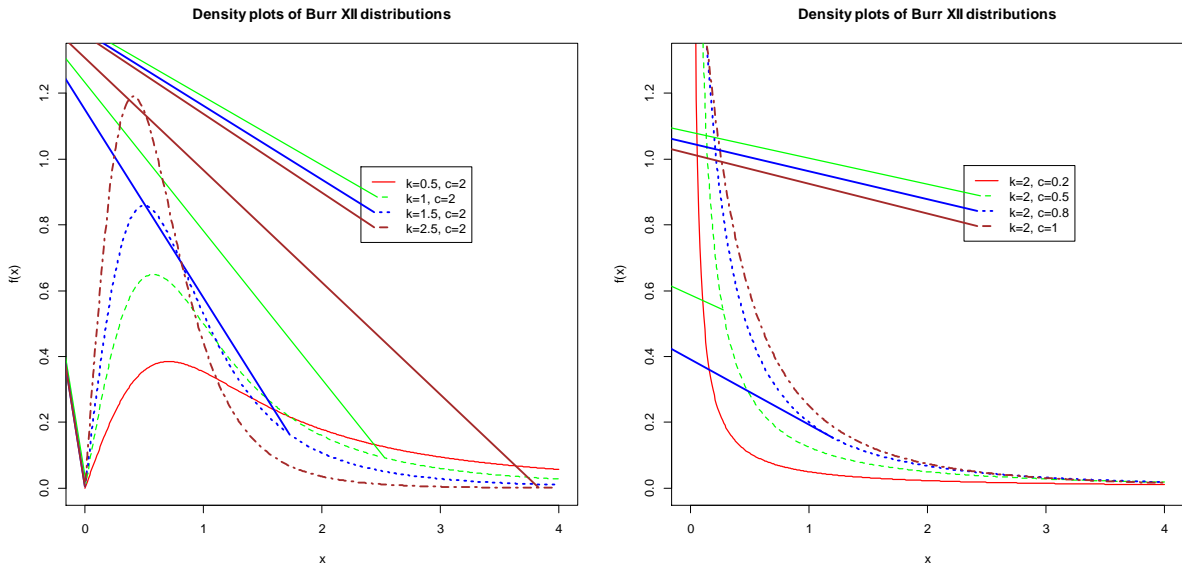
$$f(x; c, k) = ck \frac{x^{c-1}}{(1 + x^c)^{k+1}}, \quad x > 0$$

όπου  $c, k > 0$ . Για  $0 < c \leq 1$  η  $f(x; c, k)$  είναι φθίνουσα συνάρτηση, ενώ για  $c > 1$  είναι μονοκόρυφη στο σημείο

$$x_{\text{mod}} = \left(\frac{c-1}{ck+1}\right)^{1/c}.$$

Στο ακόλουθο σχήμα δίνονται συναρτήσεις πυκνότητας κατανομών Burr τύπου XII.





Σχήμα 2. 1 - Συναρτήσεις πυκνότητας διάφορων κατανομών Burr τύπου XII

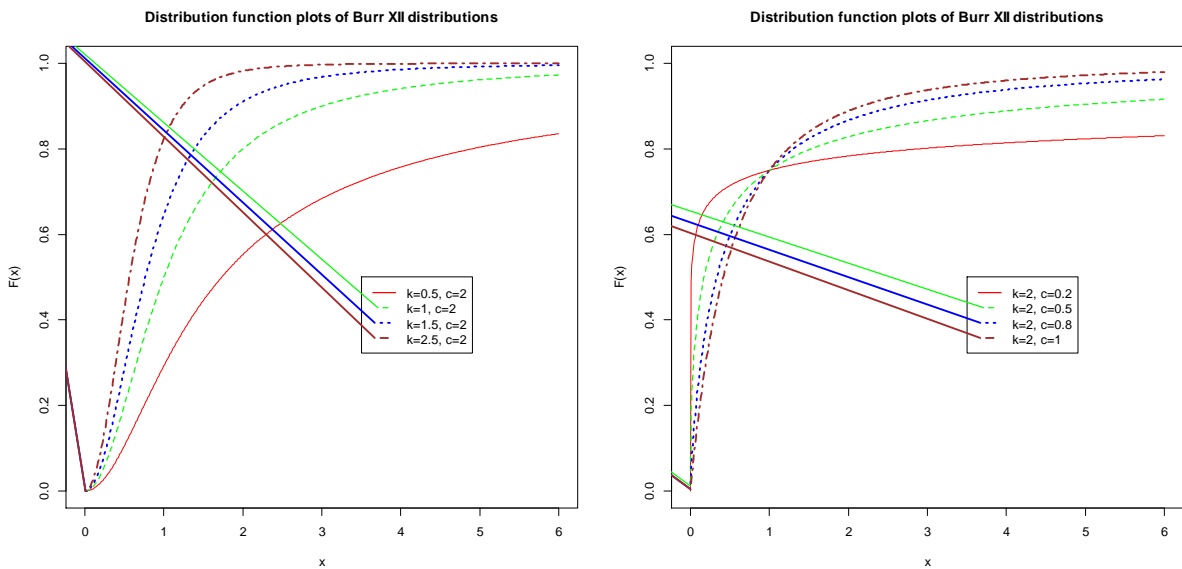
Η συνάρτηση κατανομής της κατανομής Burr τύπου XII δίνεται από τον τύπο

$$F(x) = 1 - (1 + x^c)^{-k}, \quad x > 0$$

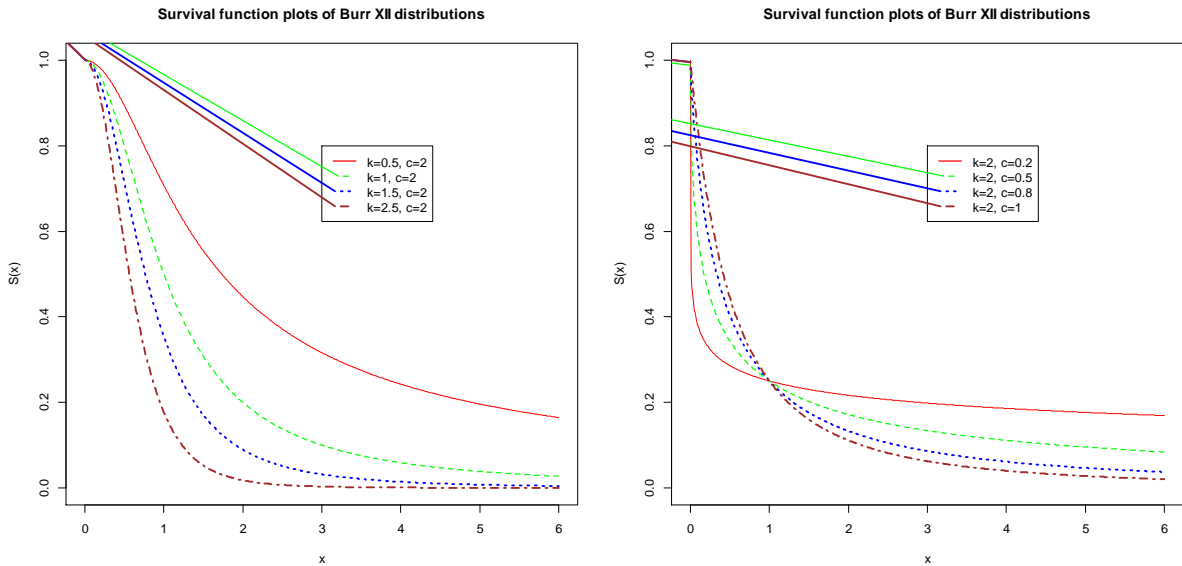
και η συνάρτηση επιβίωσης της από τον τύπο

$$S(x) = (1 + x^c)^{-k}, \quad x > 0.$$

Στα ακόλουθα σχήματα δίνονται γραφικές παραστάσεις συναρτήσεων κατανομής και συναρτήσεων επιβίωσης των κατανομών Burr του Σχήματος 2.1.



Σχήμα 2. 2 - Συναρτήσεις κατανομής διάφορων κατανομών Burr τύπου XII

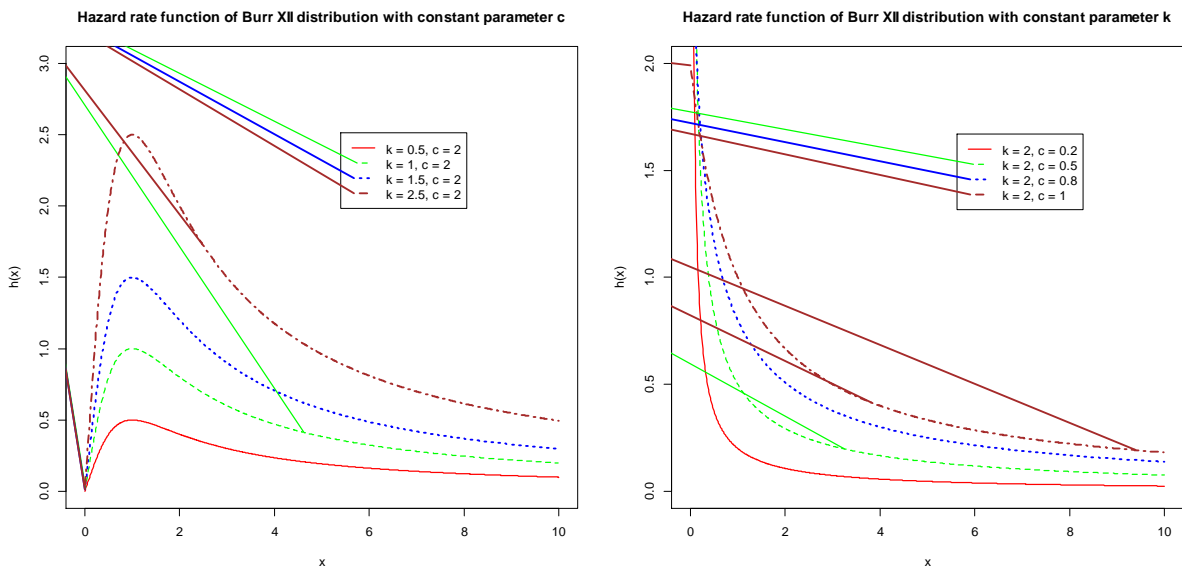


Σχήμα 2. 3 - Συναρτήσεις επιβίωσης διάφορων κατανομών Burr τύπου XII

Για τη συνάρτηση κινδύνου της κατανομής Burr τύπου XII έχουμε ότι

$$h(x) = \frac{f(x)}{S(x)} = \frac{kcx^{c-1}(1+x^c)^{-(k+1)}}{(1+x^c)^{-k}} = \frac{kcx^{c-1}}{1+x^c}, \quad x > 0.$$

Η συνάρτηση κινδύνου έχει σχήμα ανάποδης μπιανιέρας για  $c > 1$ , και είναι φθίνουσα για  $0 < c \leq 1$ .



Σχήμα 2. 4 - Συναρτήσεις κινδύνου διάφορων κατανομών Burr τύπου XII

## Ροπές

Για την  $m$ -οστή ροπή μια τυχαίας μεταβλητής  $X$  που έχει την κατανομή Burr τύπου XII, έχουμε

$$E(X^m) = \int_{-\infty}^{\infty} x^m f(x) dx = k \int_0^{\infty} x^m [(1+x^c)^{-1}]^{k+1} c x^{c-1} dx.$$

Θέτοντας  $u = (1+x^c)^{-1}$ , δηλαδή  $x = [(1-u)/u]^{1/c}$ , έχουμε ότι  $du = -cx^{c-1}(1+x^c)^{-2} dx = -cx^{c-1}u^2 dx$ . Συνεπώς,  $cx^{c-1} dx = -u^{-2} du$ . Είναι ξεκάθαρο ότι  $u \rightarrow 1$  όταν  $x \rightarrow 0$ , και  $u \rightarrow 0$  όταν  $x \rightarrow \infty$ . Έτσι

$$E(X^m) = k \int_1^0 \left[ \left( \frac{1-u}{u} \right)^{1/c} \right]^m u^{k+1} (-u^{-2}) du = k \int_0^1 u^{(k-m/c)-1} du.$$

Όταν  $k - m/c > 0$  και  $1 + m/c > 0$  (ή όταν  $m < kc$ ), το παραπάνω ολοκλήρωμα υπάρχει και είναι ίσο με  $B(k - m/c, 1 + m/c)$ , όπου  $B(\cdot; \cdot)$  είναι η συνάρτηση Βήτα. Επειδή  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  για οποιαδήποτε  $a, b > 0$ , όπου  $\Gamma(\cdot)$  είναι η συνάρτηση Γάμμα, έχουμε ότι

$$E(X^m) = kB\left(k - \frac{m}{c}, 1 + \frac{m}{c}\right) = \frac{\Gamma\left(k - \frac{m}{c}\right)\Gamma\left(1 + \frac{m}{c}\right)}{\Gamma(k)}, \quad m < kc.$$

Από την παραπάνω εξίσωση παίρνουμε ότι

$$\mu = E(X) = \frac{\Gamma\left(k - \frac{1}{c}\right)\Gamma\left(1 + \frac{1}{c}\right)}{\Gamma(k)}, \quad kc > 1.$$

Επίσης, για τη διακύμανση, αφού  $\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2$ , έχουμε ότι

$$\sigma^2 = \frac{\Gamma\left(k - \frac{2}{c}\right)\Gamma\left(1 + \frac{2}{c}\right)\Gamma(k) - \Gamma^2\left(k - \frac{1}{c}\right)\Gamma^2\left(1 + \frac{1}{c}\right)}{\Gamma^2(k)}, \quad kc > 2.$$

## Συντελεστής ασυμμετρίας και κέρτωσης

Ο συντελεστής λοξότητας

$$a_3 = \frac{E(X - \mu)^3}{\sigma^3}$$

και ο συντελεστής κέρτωσης

$$a_4 = \frac{E(X - \mu)^4}{\sigma^4}$$

της κατανομής Burr τύπου XII δίνονται από τους τύπους

$$a_3 = \frac{\Gamma^2(c)h_3 - 3\Gamma(c)h_2h_1 + 2h_1^3}{(\Gamma(c)h_2 - h_1^2)^{3/2}}$$

και

$$a_4 = \frac{\Gamma^3(c)h_4 - 4\Gamma^2(c)h_3h_1 + 6\Gamma(c)h_2h_1^2 - 3h_1^4}{(\Gamma(c)h_2 - h_1^2)^{3/2}}$$

όπου

$$h_j = \Gamma\left(\frac{j}{c} + 1\right)\Gamma\left(k - \frac{j}{c}\right), \quad j = 1, 2, 3, 4$$

και  $ck > 4$ .

### Συμπεριφορά της δεξιάς ουράς

Η ουρά μιας κατανομής πιθανότητας (ή πιο σωστά η δεξιά ουρά) είναι το τμήμα της κατανομής που αντιστοιχεί σε μεγάλες τιμές της τυχαίας μεταβλητής. Οι μεγάλες τιμές έχουν μεγάλη επίδραση, και επομένως είναι σημαντικό να κατανοήσουμε τις ιδιότητες της ουράς μιας κατανομής. Μια τυχαία μεταβλητή που τείνει να αποδίδει μεγαλύτερη πιθανότητα σε μεγαλύτερες τιμές λέγεται ότι έχει βαριά ουρά.

Υπάρχουν διάφοροι τρόποι ταξινόμησης της βαρύτητας της ουράς μιας κατανομής. Ένας από αυτούς βασίζεται στην ύπαρξη όλων των ροπών της κατανομής. Μια κατανομή έχει ελαφριά ουρά εάν η  $m$ -οστή ροπή της υπάρχει για κάθε θετικό ακέραιο  $m$ . Διαφορετικά, θεωρούμε ότι έχει βαριά ουρά, εάν η  $m$ -οστή ροπή της υπάρχει για συγκεκριμένο θετικό ακέραιο  $m$  μόνο. Η  $m$ -οστή ροπή της κατανομής Burr τύπου XII υπάρχει μόνο για  $m < kc$ . Ως εκ τούτου, η κατανομή Burr τύπου XII έχει βαριά ουρά. Επίσης η συνάρτηση επιβίωσης της κατανομής Burr τύπου XII είναι  $S(x) = (1 + x^c)^{-k} \sim x^{-kc}$ , που δείχνει ότι η ουρά της κατανομής μειώνεται αλγεβρικά καθώς αυξάνει το  $x$  και, ως εκ τούτου, προκύπτει ότι η  $m$ -οστή ροπή της κατανομής Burr τύπου XII είναι άπειρη για  $m \geq kc$ .

Πληροφορίες σχετικά με την συμπεριφορά της ουράς της κατανομής Burr τύπου XII μπορούμε να πάρουμε και από την συνάρτηση κινδύνου. Μια κατανομή με φθίνουσα συνάρτηση κινδύνου έχει βαριά ουρά. Από το Σχήμα 2.4 προκύπτει ότι η συνάρτηση κινδύνου της κατανομής Burr τύπου XII μπορεί να είναι φθίνουσα στο διάστημα  $(0, \infty)$ , ή μπορεί να έχει σχήμα ανάποδης μπιανιέρας που μειώνεται σε μεγάλες τιμές αφού φτάσει πρώτα στην κορυφή της. Αυτό δείχνει ότι η κατανομή Burr τύπου XII είναι κατάλληλη

για τη μοντελοποίηση δεδομένων χρόνου επιβίωσης με βαριά ουρά των οποίων η συνάρτηση κινδύνου είναι είτε φθίνουσα, είτε έχει σχήμα ανάποδης μανιέρας.

### Εκτιμητές μέγιστης πιθανοφάνειας

Όταν μοντελοποιούμε δεδομένα, είναι σημαντικό να γνωρίζουμε τις τιμές των παραμέτρων του μοντέλου. Ωστόσο, στην πράξη, αυτές είναι άγνωστες. Επομένως, πρέπει να εκτιμήσουμε τις παραμέτρους με βάση δεδομένα που έχουμε στη διάθεσή μας. Έστω  $X_1, X_2, \dots, X_n$  ένα τυχαίο δείγμα μεγέθους  $n$ , από την Burr τύπου XII τυχαία μεταβλητή  $X$  με παραμέτρους  $c, k$ . Έστω ότι τα  $x_1, x_2, \dots, x_n$  είναι οι αντίστοιχες παρατηρηθείσες τιμές. Η συνάρτηση πιθανοφάνειας είναι

$$\begin{aligned} L(c, k; x_1, x_2, \dots, x_n) &= \prod_{i=1}^n f(x_i; c, k) = \prod_{i=1}^n ckx_i^{c-1}(1+x_i^c)^{-(k+1)} \\ &= c^n k^n \left( \prod_{i=1}^n x_i^{c-1} \right) \left[ \prod_{i=1}^n (1+x_i^c)^{-(k+1)} \right]. \end{aligned}$$

Οι εκτιμητές μέγιστης πιθανοφάνειας των παραμέτρων  $c$  και  $k$  μπορούν να βρεθούν μεγιστοποιώντας την συνάρτηση πιθανοφάνειας  $L$ . Έτσι μπορούν να προκύψουν από την επίλυση του συστήματος

$$\frac{\partial L(c, k; x_1, x_2, \dots, x_n)}{\partial c} = 0$$

και

$$\frac{\partial L(c, k; x_1, x_2, \dots, x_n)}{\partial k} = 0.$$

Ωστόσο γνωρίζουμε ότι η συνάρτηση πιθανοφάνειας  $L$  και ο λογάριθμός της  $\ln L$  μεγιστοποιούνται στις ίδιες τιμές των  $c$  και  $k$ . Συνεπώς επιλέγουμε να λύσουμε σύστημα

$$\frac{\partial \ln L(c, k; x_1, x_2, \dots, x_n)}{\partial c} = 0$$

και

$$\frac{\partial \ln L(c, k; x_1, x_2, \dots, x_n)}{\partial k} = 0$$

καθότι η επίλυση του είναι ευκολότερη. Έτσι

$$\begin{aligned}\ln L(c, k; x_1, x_2, \dots, x_n) &= \ln(k^n) + \ln(c^n) + \ln\left(\prod_{i=1}^n x_i^{c-1}\right) + \ln\left[\prod_{i=1}^n (1 + x_i^c)^{-(k+1)}\right] \\ &= n \ln(k) + n \ln(c) + (c-1) \sum_{i=1}^n \ln x_i - (k+1) \sum_{i=1}^n \ln(1 + x_i^c).\end{aligned}$$

Παραγωγίζοντας την συνάρτηση του λογαρίθμου της πιθανοφάνειας ως προς τη μια παράμετρο και στην συνέχεια ως προς την άλλη και θέτοντας τις παραγώγους ίσες με το μηδέν, βρίσκουμε τις μέγιστες τιμές των παραμέτρων, έστω  $\hat{c}$  και  $\hat{k}$ . Συνεπώς

$$0 = \frac{\partial \ln L(c, k; x_1, x_2, \dots, x_n)}{\partial k} = \frac{n}{\hat{k}} - \sum_{i=1}^n \ln(1 + x_i^{\hat{c}})$$

και

$$0 = \frac{\partial \ln L(c, k; x_1, x_2, \dots, x_n)}{\partial c} = \frac{n}{\hat{c}} + \sum_{i=1}^n \ln(x_i) - (\hat{k} + 1) \sum_{i=1}^n \frac{x_i^{\hat{c}}}{1 + x_i^{\hat{c}}}.$$

Έτσι

$$\hat{k} = \frac{n}{\sum_{i=1}^n \ln(1 + x_i^{\hat{c}})}$$

ενώ το  $\hat{c}$  προκύπτει από τη λύση της εξίσωσης

$$\frac{n}{\hat{c}} + \sum_{i=1}^n \ln(x_i) - \left[ \frac{n}{\sum_{i=1}^n \ln(1 + x_i^{\hat{c}})} + 1 \right] \sum_{i=1}^n \frac{x_i^{\hat{c}} \ln(x_i)}{1 + x_i^{\hat{c}}} = 0$$

(δείτε Rodriguez (2014) και Hakim et al (2021)).

### Επέκταση της Burr XII κατανομής

Όπως έχει ήδη αναφερθεί, ο Burr (1942) παρουσίασε για πρώτη φορά την κατανομή Burr ως οικογένεια δύο παραμέτρων. Ο Tadikamalla (1980) εισήγαγε μια περαιτέρω παράμετρο κλίμακας.

Η κατανομή Burr τύπου XII με παραμέτρους  $c$ ,  $k$  και  $a$  έχει συνάρτηση πυκνότητας που δίνεται από τον τύπο

$$f(x; c, k, a) = \frac{\frac{kc}{a} \left(\frac{x}{a}\right)^{c-1}}{\left(1 + \left(\frac{x}{a}\right)^c\right)^{k+1}}, \quad x > 0, \quad c > 0, \quad k > 0, \quad a > 0.$$

Η συνάρτηση κατανομής της κατανομής Burr τύπου XII με τρεις παραμέτρους δίνεται από τον τύπο

$$F(x; c, k, a) = 1 - \frac{1}{\left(1 + \left(\frac{x}{a}\right)^c\right)^k}.$$

Αντίστοιχα, η συνάρτηση επιβίωσης της κατανομής Burr τύπου XII με τρεις παραμέτρους δίνεται από τον τύπο

$$S(x; c, k, a) = \frac{1}{\left(1 + \left(\frac{x}{a}\right)^c\right)^k}.$$

Για τη συνάρτηση κινδύνου της κατανομής Burr τύπου XII με τρεις παραμέτρους έχουμε ότι

$$h(x; c, k, a) = \frac{\frac{kc}{a} \left(\frac{x}{a}\right)^{c-1}}{1 + \left(\frac{x}{a}\right)^c}.$$

## 2.4 Η κατανομή Burr τύπου II

Η κατανομή Burr τύπου II μελετήθηκε αρχικά σε μοντέλο δυαδικής επιλογής. Προκύπτει και ως περιθώρια συνάρτηση πυκνότητας από τη γενικευμένη διμεταβλητή λογιστική κατανομή του Gumbel.

Η συνάρτηση  $g(x)$  που γεννά την κατανομής Burr τύπου II με βάση τον Πίνακα 2.1 είναι η

$$g(x) = \frac{ke^{-x}(e^{-x} + 1)^{k-1}}{(e^{-x} + 1)^k - 1}, \quad k > 0, \quad x \in (-\infty, +\infty).$$

Έτσι

$$\int_{-\infty}^x g(x) dx = \int_{-\infty}^x \frac{ke^{-x}(e^{-x} + 1)^{k-1}}{(e^{-x} + 1)^k - 1} dx = -\ln((e^{-x} + 1)^k - 1)$$

και από τη σχέση

$$F(x) = \frac{1}{1 + \exp\left(-\int_{-\infty}^x g(x) dx\right)}$$

προκύπτει άμεσα ότι

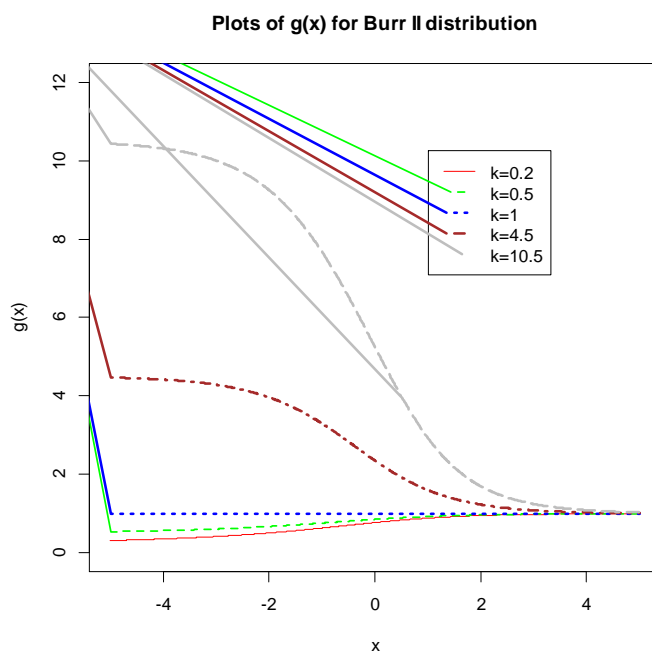
$$F(x) = (1 + e^{-x})^{-k}, \quad x \in (-\infty, +\infty).$$

Παραγωγίζοντας την παραπάνω σχέση, παίρνουμε την συνάρτηση πυκνότητας της κατανομής Burr II με παράμετρο  $k > 0$ , που δίνεται από τον τύπο

$$f(x) = \frac{ke^{-x}}{(1 + e^{-x})^{k+1}}, \quad x \in (-\infty, +\infty).$$

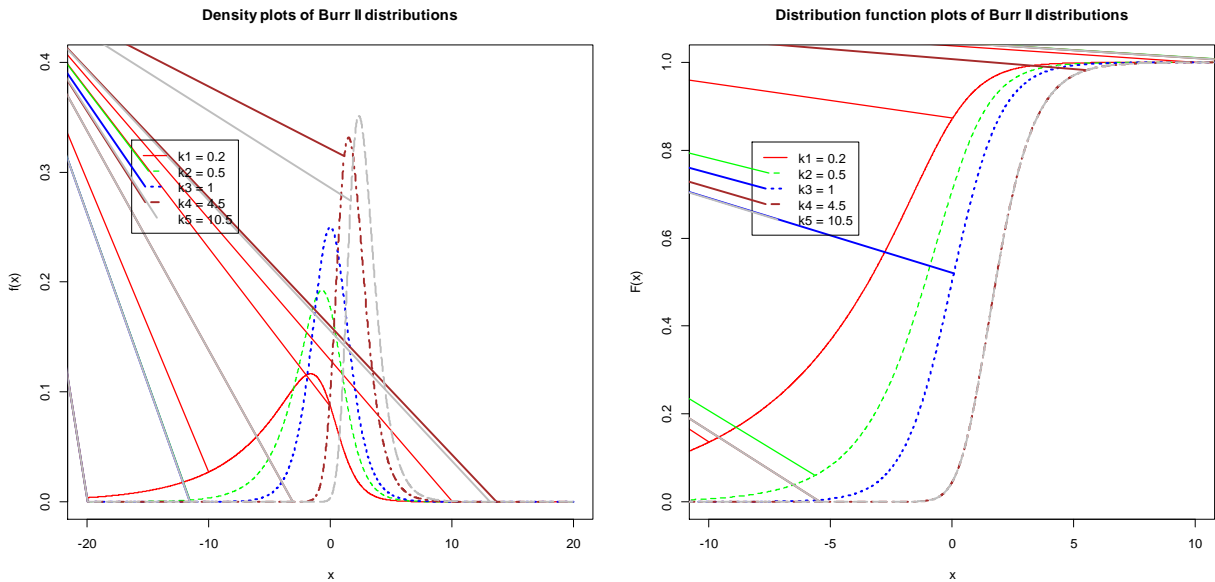
Η παραπάνω συνάρτηση πυκνότητας είναι μονοκόρυφη και γενικά είναι λοξή προς τα δεξιά, ωστόσο είναι συμμετρική όταν  $k = 1$ .

Παρακάτω δίνονται διάφορα σχήματα για τις συναρτήσεις  $g(x)$ ,  $f(x)$  και  $F(x)$ .



Σχήμα 2. 5 - Συναρτήσεις  $g(x)$  για την κατανομή Burr τύπου II





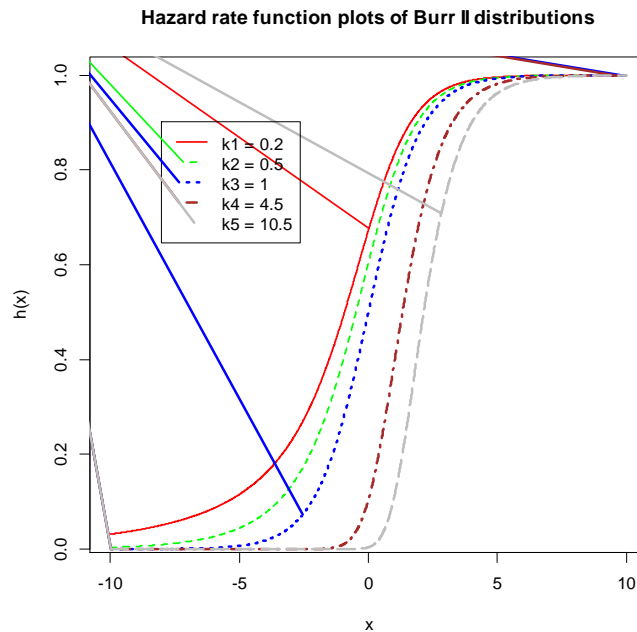
Σχήμα 2. 6 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου II

Η συνάρτηση επιβίωσης της κατανομής Burr τύπου II δίνεται από τον τύπο

$$S(x) = 1 - (1 + e^{-x})^{-k}, \quad x \in (-\infty, +\infty)$$

ενώ η συνάρτηση κινδύνου από τον τύπο

$$h(x) = \frac{f(x)}{S(x)} = \frac{ke^{-x}(1 + e^{-x})^{-(k+1)}}{1 - (1 + e^{-x})^{-k}}, \quad x \in (-\infty, +\infty).$$



Σχήμα 2. 7 - Συναρτήσεις κινδύνου για την κατανομή Burr τύπου II

## Διακριτή κατανομή Burr τύπου II

Ο Helmy (2018) μελέτησε μια διακριτή κατανομή Burr τύπου II. Θέτοντας  $\theta = e^{-k}$  έχουμε ότι η συνάρτηση επιβίωσης μιας τυχαίας μεταβλητής  $Y$  που έχει την κατανομή Burr τύπου II με παράμετρο  $k$  μπορεί να γραφτεί στη μορφή

$$S(y) = 1 - \theta^{\log(1+e^{-y})}$$

όπου  $\theta = e^{-k}$ . Ο Helmy (2018) ονόμασε την κατανομή της τυχαίας μεταβλητής  $X = [Y]$  ως διακριτή κατανομή Burr τύπου II με παράμετρο  $\theta$  ( $0 < \theta < 1$ ), την οποία θα συμβολίζουμε ως dBurrII ( $\theta$ ). Η συνάρτηση πιθανότητας  $p(x) = P(X = x)$  της  $X$  δίνεται από τον τύπο

$$p(x) = S(x) - S(x + 1) = \theta^{\log(1+e^{-(x+1)})} - \theta^{\log(1+e^{-x})}, \quad x = \dots, -2, -1, 0, 1, 2, \dots$$

ενώ η συνάρτηση κατανομής της από τον τύπο

$$F(x) = \theta^{\log(1+e^{-x})}, \quad x = \dots, -2, -1, 0, 1, 2.$$

Για την εκτίμηση της παραμέτρου  $\theta$  ο Helmy (2018) χρησιμοποίησε τη μέθοδο μέγιστης πιθανοφάνειας. Αν  $X_1, X_2, \dots, X_n$  ένα τυχαίο δείγμα μεγέθους  $n$  από την κατανομή dBurrII και

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

η συνάρτηση πιθανοφάνειας, τότε ο Helmy (2018) έδειξε ότι ο εκτιμητής μέγιστης πιθανοφάνειας του  $\theta$ , έστω  $\hat{\theta}$ , προκύπτει από τη λύση της εξίσωσης

$$\frac{d \log L}{d \theta} = \sum_{i=1}^n \left[ \frac{\log(1 + e^{-(x_i+1)})}{\hat{\theta}} + \frac{\varphi(x_i) \hat{\theta}^{\varphi(x_i)-1}}{\hat{\theta}^{\varphi(x_i)} - 1} \right] = 0$$

όπου

$$\varphi(x) = \log \left( \frac{1 + e^{-(x+1)}}{1 + e^{-x}} \right).$$

## 2.5 Η κατανομή Burr τύπου III

Από το σύστημα των κατανομών Burr, το μοντέλο της κατανομής Burr τύπου III χρησιμοποιείται ευρέως. Αναφέρεται και στη βιβλιογραφία ως κατανομή Dagum (Dagum (1977)) και χρησιμοποιείται σε μελέτες κατανομής εισοδήματος, μισθού και πλούτου. Στην αναλογιστική βιβλιογραφία, αναφέρεται με το όνομα αντίστροφη κατανομή Burr και στην μετεωρολογική επιστήμη αποτελεί ειδική περίπτωση της κατανομής κάπα (Mielke & Johnson (1974)). Αξίζει να σημειωθεί ότι η κατανομή Burr III εισήχθη στο σύστημα Burr για να μοντελοποιήσει δεδομένα διάρκειας ζωής ή δεδομένα επιβίωσης. Η χρησιμότητα και οι ιδιότητες της κατανομής Burr τύπου III συζητήθηκαν από τους Burr και Cislak (1968) & Johnson et al. (1995). Εάν μια τυχαία μεταβλητή  $X$  έχει κατανομή Burr τύπου XII, τότε η τυχαία μεταβλητή  $1/X$  έχει κατανομή Burr τύπου III. Για το λόγο αυτό, η κατανομή Burr τύπου III αναφέρεται συχνά ως η αντίστροφη κατανομή Burr.

Η συνάρτηση  $g(x)$  που γεννά την κατανομή Burr τύπου III με βάση τον Πίνακα 2.1 είναι η

$$g(x) = \frac{kr}{x(1 - (x^{-k} + 1)^{-r})(x^k + 1)}, \quad k, r > 0.$$

Συνεπώς

$$\int_0^x g(x) dx = \int_0^x \frac{kr}{x(1 - (x^{-k} + 1)^{-r})(x^k + 1)} dx = -\ln((1 + x^{-k})^r - 1).$$

Από τη γενική σχέση

$$F(x) = \frac{1}{1 + \exp(-\int_{-\infty}^x g(x))}$$

προκύπτει άμεσα ότι

$$F(x) = (1 + x^{-k})^{-r}, \quad x > 0.$$

Από την παραπάνω σχέση προκύπτει ότι η συνάρτηση επιβίωσης είναι

$$S(x) = 1 - F(x) = 1 - (1 + x^{-k})^{-r}, \quad x > 0.$$

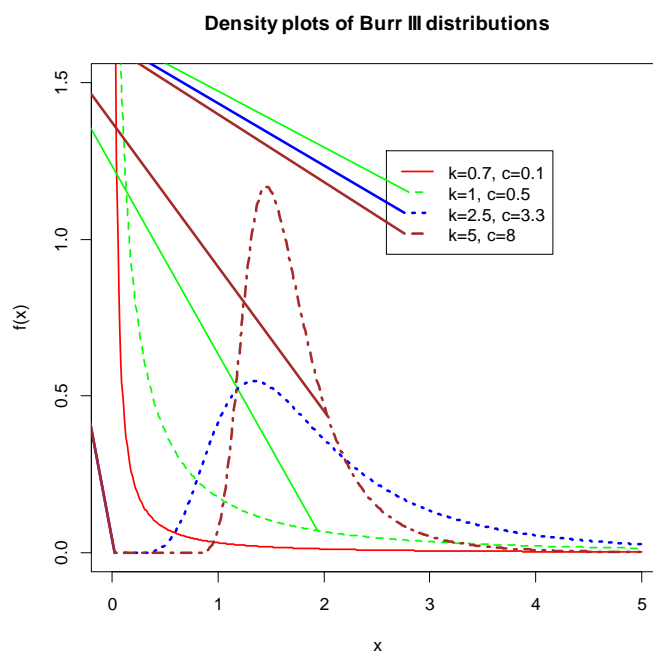
Παραγωγίζοντας τη συνάρτηση κατανομής  $F(x)$ , παίρνουμε την συνάρτηση πυκνότητας της κατανομής Burr τύπου III με παραμέτρους  $k, r > 0$ , που δίνεται από τον τύπο

$$f(x) = r k x^{-(k+1)} (1 + x^{-k})^{-(r+1)}, \quad x > 0.$$

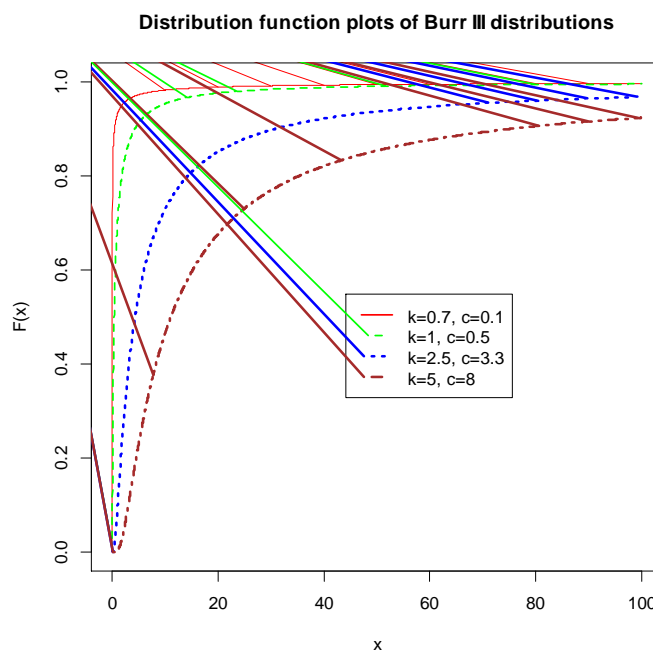
Η συνάρτηση κινδύνου δίνεται από τον τύπο

$$h(x) = \frac{f(x)}{S(x)} = \frac{rkx^{-(k+1)}(1+x^{-k})^{-(r+1)}}{1-(1+x^{-k})^{-r}}, x > 0.$$

Παρακάτω δίνονται διάφορα σχήματα για τις συναρτήσεις  $f(x)$  και  $F(x)$  της κατανομής Burr τύπου III.



Σχήμα 2. 8 - Συναρτήσεις πυκνότητας της κατανομής Burr τύπου III



Σχήμα 2. 9 - Συναρτήσεις κατανομής της κατανομής Burr τύπου III

## Εκτιμητές μέγιστης πιθανοφάνειας

Υποθέτουμε ένα δείγμα μεγέθους  $n$ ,  $(X_1, X_2, \dots, X_n)$ , από την κατανομή Burr τύπου III. Η συνάρτηση πιθανοφάνειας αυτού είναι

$$L(r, k) = \prod_{i=1}^n f(x_i; r, k) = (rk)^n \prod_{i=1}^n x_i^{-(k+1)} (1 + x_i^{-k})^{-(r+1)}.$$

Έτσι,

$$\ln L(r, k) = n \ln r + n \ln k - (k+1) \sum_{i=1}^n \ln x_i - (r+1) \sum_{i=1}^n \ln(1 + x_i^{-k}).$$

Παραγωγίζοντας την συνάρτηση του λογαρίθμου της συνάρτησης πιθανοφάνειας ως προς τις παραμέτρους  $r$  και  $k$  παίρνουμε το σύστημα

$$\frac{d \ln L(r, k)}{dr} = \frac{n}{r} - \sum_{i=1}^n \ln(1 + x_i^{-k}) = 0$$

$$\frac{d \ln L(r, k)}{dk} = \frac{n}{k} - \sum_{i=1}^n \ln x_i + (r+1) \sum_{i=1}^n \frac{x_i^{-k} \ln x_i}{1 + x_i^{-k}} = 0.$$

Έτσι καταλήγουμε στον εκτιμητή μέγιστης πιθανοφάνειας του  $r$  συναρτήσει του  $k$

$$\hat{r}(k) = \frac{n}{\sum_{i=1}^n \ln(1 + x_i^{-k})}.$$

Αντικαθιστώντας το  $r$  με το  $\hat{r}(k)$  στον τύπο του  $\ln L(r, k)$  παίρνουμε

$$\begin{aligned} g(k) &= \ln L(\hat{r}(k), k) \\ &= K + n \ln k - (k+1) \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln(1 + x_i^{-k}) \\ &\quad - n \ln \left[ \sum_{i=1}^n \ln(1 + x_i^{-k}) \right] \end{aligned}$$

όπου  $K$  είναι μια σταθερά. Επομένως, ο εκτιμητής μέγιστης πιθανοφάνειας του  $k$  (έστω  $\hat{k}_{MLE}$ ) μπορεί να επιτευχθεί με τη μεγιστοποίηση της συνάρτησης  $g(k)$  ως προς  $k$ . Μπορεί ακόμη να δειχθεί ότι το μέγιστο της συνάρτησης αυτής, μπορεί να ληφθεί ως λύση της ακόλουθης εξίσωσης

$$h(k) = k$$

όπου η  $h(k)$

$$h(k) = \left[ -\frac{1}{n} \sum_{i=1}^n \frac{\ln x_i}{1 + x_i^{-k}} - \frac{\sum_{i=1}^n \frac{x_i^{-k} \ln x_i}{1 + x_i^{-k}}}{\sum_{i=1}^n \ln(1 + x_i^{-k})} \right]^{-1}.$$

Μπορούμε να εφαρμόσουμε μια επαναληπτική διαδικασία για να βρούμε τη λύση  $\hat{k}_{MLE}$  της εξίσωσης  $h(k) = k$ . Όταν υπολογιστεί ο  $\hat{k}_{MLE}$ , μπορούμε να βρούμε και τον εκτιμητή μέγιστης πιθανοφάνειας του  $r$  (έστω  $\hat{r}_{MLE}$ ) ως  $r_{MLE} = \hat{r}(\hat{k}_{MLE})$ .

### Διακριτή κατανομή Burr τύπου III

Οι AL-Hunuti & AL-Dayian (2012) εισήγαγαν ένα διακριτό μοντέλο ανάλογο της κατανομής Burr τύπου III. Θέτοντας  $\theta = e^{-r}$  έχουμε ότι η συνάρτηση επιβίωσης μιας τυχαίας μεταβλητή  $Y$  που έχει την κατανομή Burr τύπου III με παραμέτρους  $k$  και  $r$  μπορεί να γραφτεί στη μορφή

$$S(y) = 1 - (1 + x^{-k})^{-r} = 1 - \exp(\log(1 + x^{-k})^{-r}) = 1 - \theta^{\log(1+x^{-k})}.$$

Οι AL-Hunuti & AL-Dayian (2012) ονόμασαν την κατανομή της τυχαίας μεταβλητής  $X = [Y]$  ως διακριτή κατανομή Burr τύπου III με παραμέτρους  $k$  και  $\theta$  ( $0 < \theta < 1$ ), την οποία θα συμβολίζουμε ως  $\text{dBurrIII}(k, \theta)$ . Η συνάρτηση πιθανότητας  $p(x) = P(X = x)$  της  $X$  δίνεται από τον τύπο

$$p(x) = S(x) - S(x + 1) = \theta^{\log(1+(x+1)^{-k})} - \theta^{\log(1+x^{-k})}, \quad x = 0, 1, 2, \dots$$

Η συνάρτηση κινδύνου  $h(x)$  της κατανομής  $\text{dBurrIII}(k, \theta)$  δίνεται από τον τύπο

$$h(x) = \frac{p(x)}{S(x)} = \frac{\theta^{\log(1+(x+1)^{-k})} - \theta^{\log(1+x^{-k})}}{1 - \theta^{\log(1+x^{-k})}}, \quad x = 0, 1, 2, \dots$$

Η ροπή  $r$ -τάξης της κατανομής  $\text{dBurrIII}(k, \theta)$  δίνεται από τον παρακάτω τύπο

$$\mu_r = \mu = \sum_{x=0}^{\infty} x^r f(x) = \sum_{x=0}^{\infty} x^r [S(x) - S(x + 1)]$$

από τον οποίο προκύπτει τόσο η μέση τιμή, όσο και η διακύμανση είναι

$$\mu_1 = \mu = \sum_{x=0}^{\infty} x \left[ \theta^{\log(1+(x+1)^{-k})} - \theta^{\log(1+x^{-k})} \right]$$

$$\sigma^2 = \mu_2 - \mu^2 = \sum_{x=0}^{\infty} x^2 \left[ \theta^{\log(1+(x+1)^{-k})} - \theta^{\log(1+x^{-k})} \right] - \left[ \sum_{x=0}^{\infty} x \left[ \theta^{\log(1+(x+1)^{-k})} - \theta^{\log(1+x^{-k})} \right] \right]^2.$$

Για περισσότερες ιδιότητες και εφαρμογές της κατανομής dBurr III( $k, \theta$ ) ο αναγνώστης παραπέμπεται στο άρθρο των AL-Huniti & AL-Dayian (2012).

### Επέκταση της Burr III κατανομής

Μέσω ερευνών διάφορων επιστημόνων, έχουν αναπτυχθεί αρκετές επεκτάσεις της κατανομής Burr III. Ενδεικτικά, παρατίθεται παρακάτω μια από τις πιο διαδεδομένες επεκτάσεις της κατανομής αυτής, με συνάρτηση κατανομής που δίνεται από τον τύπο

$$F(x; r, k, a) = \left( 1 + \left( \frac{x}{a} \right)^{-k} \right)^{-r},$$

όπου  $r > 0, k > 0, a > 0$ .

Δοθείσης της παραπάνω σχέσης, παραγωγίζοντας προκύπτει η συνάρτηση πυκνότητας πιθανότητας

$$f(x; r, k, a) = r a^k k x^{-k-1} \left( 1 + \left( \frac{x}{a} \right)^{-k} \right)^{-r-1}.$$

Τέλος, η συνάρτηση κινδύνου  $h(x)$  δίνεται από τον τύπο

$$h(x; r, k, a) = \frac{r a^k k x^{-k-1} \left( 1 + \left( \frac{x}{a} \right)^{-k} \right)^{-r-1}}{1 - \left( 1 + \left( \frac{x}{a} \right)^{-k} \right)^{-r}}.$$

Για περισσότερες λεπτομέρειες δείτε Παράγραφο 2.8.

## 2.6 Η κατανομή Burr τύπου X

Οι κατανομές Burr τύπου X με μία παράμετρο και Burr τύπου XII έχουν λάβει μεγάλη προσοχή στη βιβλιογραφία. Ειδικότερα η κατανομή Burr τύπου X έχει κάποιες αρκετά ενδιαφέρουσες ιδιότητες. Η συγκεκριμένη λοξή κατανομή μπορεί να χρησιμοποιηθεί αρκετά αποτελεσματικά στην ανάλυση δεδομένων χρόνου ζωής. Η κατανομή Burr X

έπαιξε επίσης σημαντικό ρόλο σε μελέτες αξιοπιστίας, στη μοντελοποίηση του χρόνου ζωής τυχαίων φαινομένων, στην υγεία, στη γεωργία και στη βιολογία. Αρκετές πτυχές της μονοπαραμετρικής ( $k = 1$ ) κατανομής Burr τύπου X, μελετήθηκαν από τους Meronci et al. (2016) (δείτε επίσης και την βιβλιογραφία που δίνεται).

Η συνάρτηση  $g(x)$  που γεννά την κατανομής Burr τύπου X, με βάση τον Πίνακα 2.1, είναι η

$$g(x) = \frac{2kxe^{-x^2}}{\left(1 - (1 - e^{-x^2})^k\right)(1 - e^{-x^2})}, \quad x, k > 0.$$

Ολοκληρώνοντας την  $g(x)$  παίρνουμε

$$\int_0^x g(x)dx = \int_0^x \frac{2kxe^{-x^2}}{\left(1 - (1 - e^{-x^2})^k\right)(1 - e^{-x^2})} dx = \ln\left(\frac{(1 - e^{-x^2})^k}{1 - (1 - e^{-x^2})^k}\right)$$

και από τη σχέση

$$F(x) = \frac{1}{1 + \exp\left(-\int_0^x g(x)\right)}$$

προκύπτει άμεσα ότι

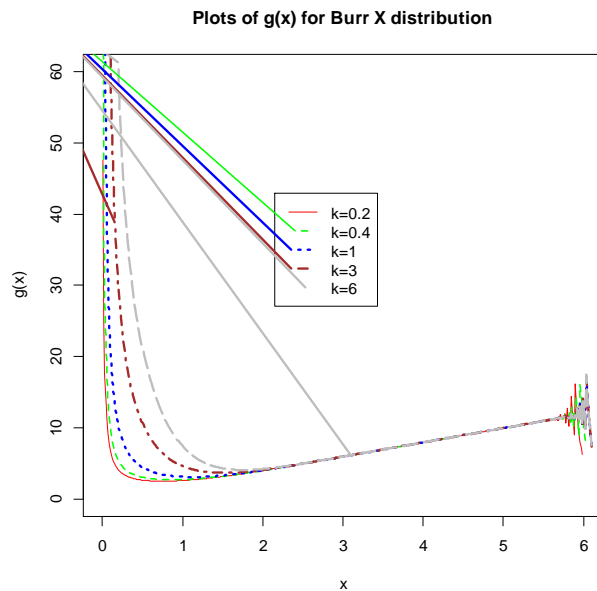
$$F(x) = \frac{1}{1 + \frac{1 - (1 - e^{-x^2})^k}{(1 - e^{-x^2})^k}} = (1 - e^{-x^2})^k, \quad x > 0.$$

Παραγωγίζοντας την παραπάνω σχέση, παίρνουμε την συνάρτηση πυκνότητας της κατανομής Burr X με παράμετρο  $k > 0$ , που δίνεται από τον τύπο

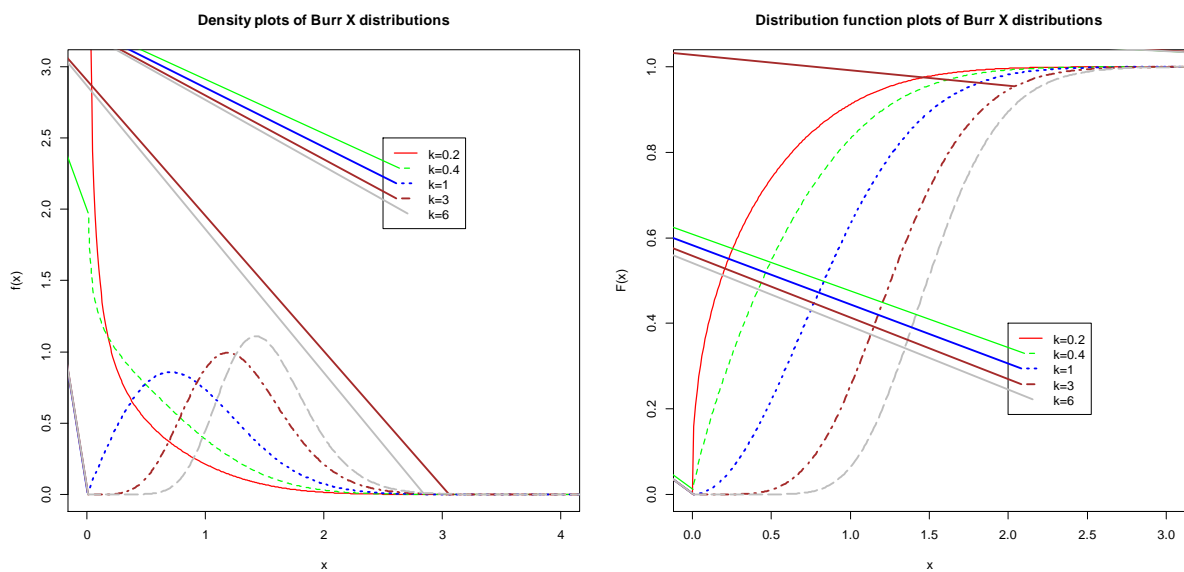
$$f(x) = 2kxe^{-x^2}(1 - e^{-x^2})^{k-1}, \quad x > 0.$$

Παρακάτω δίνονται διάφορα σχήματα για τις συναρτήσεις  $g(x)$ ,  $f(x)$  και  $F(x)$  της κατανομής Burr τύπου X.





Σχήμα 2. 10 - Συναρτήσεις  $g(x)$  για την κατανομή Burr τύπου X



Σχήμα 2. 11 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου X

## Η διπαραμετρική Burr X κατανομή

Οι Surles & Padgett (2001) είναι αυτοί οι οποίοι τροποποίησαν την μονοπαραμετρική κατανομή Burr τύπου X ακολουθώντας τη μέθοδο των Mudholkar & Srivastava (1993) και την ονόμασαν διπαραμετρική κατανομή Burr X. Η διπαραμετρική κατανομή Burr X αποτελεί επέκταση της μονοπαραμετρικής κατανομής Burr X προσθέτοντας μια παράμετρο κλίμακας  $\eta$ . Περιλαμβάνει διάφορους τύπους κατανομών, όπως και την

μονοπαραμετρική Burr τύπου X όταν  $\eta = 1$ . Η αθροιστική συνάρτηση κατανομής διπαραμετρικής κατανομής Burr X δίνεται από τον τύπο

$$F(x; \eta, k) = (1 - e^{-(\eta x)^2})^k, \quad \eta > 0, \theta > 0.$$

Υπολογίζοντας την παράγωγο της παραπάνω συνάρτησης κατανομής, βρίσκουμε την συνάρτηση πυκνότητας πιθανότητας

$$f(x; \eta, k) = 2k\eta^2 x e^{-(\eta x)^2} (1 - e^{-(\eta x)^2})^{k-1}, \quad x > 0.$$

Η συνάρτηση κινδύνου της διπαραμετρικής κατανομής Burr τύπου X δίνεται από τον παρακάτω τύπο

$$h(x; \eta, k) = 2k\eta^2 x e^{-(\eta x)^2} [1 - e^{-(\eta x)^2}]^{k-1} \{1 - [1 - e^{-(\eta x)^2}]^k\}^{-1}.$$

Ακόμη η ροπή  $r$ -τάξης δίνεται παρακάτω.

$$E(X^r) = \frac{k}{\eta^r} \Gamma\left(\frac{r}{2} + 1\right) \sum_{j=0}^{k-1} \binom{k-1}{j} \frac{(-1)^j}{(j+1)^{\frac{r}{2}+1}}$$

(δείτε Merovci et al. (2016) και Surles & Padgett (2005)).

Από τη διπαραμετρική κατανομή Burr X μπορούν να προκύψουν οι κάτωθι κατανομές.

### I. Beta Burr τύπου X

Αν  $F(x)$  είναι η αθροιστική συνάρτηση κατανομής μιας οποιασδήποτε τυχαίας μεταβλητής  $X$ , τότε η αθροιστική συνάρτηση κατανομής της οικογένειας κατανομών Beta-G με παραμέτρους  $a > 0$  και  $b > 0$ , ορίστηκε από τους Eugene et al. (2002) ως εξής

$$G(x, a, b) = \frac{1}{B(a, b)} \int_0^{F(x)} t^{a-1} (1-t)^{b-1} dt.$$

Παρακάτω δίνεται και η αντίστοιχη συνάρτηση πιθανότητας.

$$g(x, a, b) = \frac{1}{B(a, b)} [F(x)]^{a-1} [1 - F(x)]^{b-1} f(x)$$

όπου  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  και  $f(x) = F'(x)$ .

Αν αντικαταστήσουμε τη συνάρτηση κατανομής  $F(x)$  και τη συνάρτηση πυκνότητας  $f(x)$  στον τύπο για τη συνάρτηση πυκνότητας  $g(x, a, b)$ , με τις αντίστοιχες συναρτήσεις της κατανομής της Burr τύπου X κατανομής προκύπτει η Beta – Burr τύπου X κατανομή, με συνάρτηση πυκνότητας πιθανότητας που δίνεται από τον τύπο

$$g(x, a, b, \eta, k) = \frac{2k\eta^2 x}{B(a, b)} \left[ [1 - e^{-(\eta x)^2}]^k \right]^{a-1} \left[ 1 - [1 - e^{-(\eta x)^2}]^k \right]^{b-1} \\ \times e^{-(\eta x)^2} [1 - e^{-(\eta x)^2}]^{k-1}, \quad y > 0$$

(δείτε, Ibrahim and Khaleel (2020)). Η συνάρτηση κατανομής της Beta-Burr τύπου X κατανομής είναι η

$$G(x, a, b, \eta, k) = \frac{1}{B(a, b)} \int_0^{[1 - e^{-(\eta x)^2}]^k} t^{a-1} (1 - t)^{b-1} dt$$

και η συνάρτηση κινδύνου είναι η

$$h(x, a, b, \eta, k) = \frac{g(x, a, b, \eta, k)}{1 - G(x, a, b, \eta, k)}.$$

## II. Gamma Burr τύπου X

Έστω  $F(x)$  είναι η αθροιστική συνάρτηση κατανομής μιας οποιασδήποτε τυχαίας μεταβλητής  $X$ . Από τους Ristic & Balakrishnan (2012) ορίστηκε μια γενικευμένη κλάση κατανομών Gamma, με συνάρτηση κατανομής που δίνεται από τον τύπο

$$G(x, v) = 1 - \frac{1}{\Gamma(v)} \int_0^{-\log[F(x)]} t^{v-1} e^{-t} dt, \quad x \in R, v > 0$$

και συνάρτηση πυκνότητας  $g(x, v) = G'(x, v)$  που δίνεται από τον τύπο

$$g(x, v) = \frac{f(x) [-\log[F(x)]]^{v-1}}{\Gamma(v)}, \quad x \in R, v > 0.$$

Αν αντικαταστήσουμε τη συνάρτηση κατανομής  $F(x)$  και τη συνάρτηση πυκνότητας  $f(x)$  στον τύπο για τη συνάρτηση πυκνότητας  $g(x, v)$ , με τις αντίστοιχες συναρτήσεις της κατανομής της Burr τύπου X κατανομής προκύπτει η Gamma - Burr τύπου X κατανομή, με συνάρτηση πυκνότητας πιθανότητας που δίνεται από τον τύπο

$$g(x, v, \eta, k) = \frac{2k\eta^2 x e^{-(\eta x)^2}}{\Gamma(v)} [1 - e^{-(\eta x)^2}]^{k-1} \left[ -k \log[1 - e^{-(\eta x)^2}] \right]^{v-1}.$$

(δείτε, Ibrahim and Khaleel (2020)). Η συνάρτηση κατανομής της είναι

$$G(x, v, \eta, k) = 1 - \frac{\gamma \left[ v, -k \log[1 - e^{-(\eta x)^2}] \right]}{\Gamma(v)}$$

όπου

$$\gamma[\delta, x] = \int_0^x t^{\delta-1} e^{-t} dt$$

είναι η μη πλήρης συνάρτηση Γάμμα.

### III. Weibull Burr τύπου X

Έστω  $F(x, \xi)$  και  $f(x, \xi)$ , αντίστοιχα, η αθροιστική συνάρτηση κατανομής και η συνάρτηση πυκνότητας πιθανότητας μιας οποιασδήποτε τυχαίας μεταβλητής  $X$  με διάνυσμα παραμέτρων  $\xi$ . Η συνάρτηση κατανομής της Weibull-G κατανομής με δύο επιπλέον παραμέτρους  $v$  και  $\omega$ , ( $v > 0, \omega > 0$ ) ορίστηκε από τους Bourguignon et al. (2014) ως εξής

$$G(x, v, \omega, \xi) = \int_0^{F(x, \xi)/(1-F(x, \xi))} v \omega x^{\omega-1} e^{-vx^\omega} = 1 - \exp\left(-v \left[ \frac{F(x, \xi)}{1 - F(x, \xi)} \right]^\omega\right).$$

Από την παραπάνω σχέση προκύπτει ότι η συνάρτηση πυκνότητας πιθανότητας της Weibull-G κατανομής είναι

$$g(x, v, \omega, \xi) = v \omega f(x, \xi) \frac{F(x, \xi)^{\omega-1}}{[1 - F(x, \xi)]^{\omega+1}} \exp\left(-v \left[ \frac{F(x, \xi)}{1 - F(x, \xi)} \right]^\omega\right).$$

Αν αντικαταστήσουμε τη συνάρτηση κατανομής  $F(x, \xi)$  και τη συνάρτηση πυκνότητας  $f(x, \xi)$  στον τύπο για τη συνάρτηση πυκνότητας  $g(x, v, \omega, \xi)$ , με τις αντίστοιχες συναρτήσεις της κατανομής της Burr τύπου X κατανομής προκύπτει η Weibull - Burr τύπου X κατανομή, με συνάρτηση πυκνότητας πιθανότητας που δίνεται από τον τύπο

$$g(x, v, \omega, \eta, k) = 2v\omega\eta^2 k x e^{-(\eta x)^2} \frac{[1 - e^{-(\eta x)^2}]^{k\omega-1}}{[1 - [1 - e^{-(\eta x)^2}]^k]^{\omega+1}} \times \exp\left(-v \frac{[1 - e^{-(\eta x)^2}]^{k\omega}}{[1 - [1 - e^{-(\eta x)^2}]^k]^\omega}\right)$$

(δείτε, Ibrahim and Khaleel (2020)). Η συνάρτηση κατανομής της είναι

$$G(x, v, \omega, \eta, k) = 1 - \exp\left(-v \frac{[1 - e^{-(\eta x)^2}]^{k\omega}}{[1 - [1 - e^{-(\eta x)^2}]^k]^\omega}\right).$$

## 2.7 Η Κατανομή Burr τύπου IX

Η συνάρτηση  $g(x)$  της Burr IX κατανομής δίνεται από τον παρακάτω τύπο.

$$g(x) = \frac{re^x(e^x + 1)^{r-1}}{(e^x + 1)^r - 1}, \quad x \in (-\infty, \infty), \quad r > 0.$$

Η συνάρτηση  $g(x)$  έχει τις ακόλουθες ενδιαφέρουσες ιδιότητες

$$g(x) > 0, \quad \lim_{x \rightarrow -\infty} g(x) = 1, \quad \lim_{x \rightarrow \infty} g(x) = r.$$

Ολοκληρώνοντας τη  $g(x)$  παίρνουμε

$$\int_{-\infty}^x g(x) dx = \int_{-\infty}^x \frac{re^x(e^x + 1)^{r-1}}{(e^x + 1)^r - 1} dx = \ln((e^x + 1)^r - 1).$$

Έτσι η συνάρτηση κατανομής  $F(x)$  της Burr IX κατανομής δίνεται από τον τύπο

$$\begin{aligned} F(x) &= \frac{1}{1 + \exp(-\int_{-\infty}^x g(x) dx)} = \frac{1}{1 + \exp(-\ln((e^x + 1)^r - 1))} \\ &= \frac{1}{1 + \frac{1}{(e^x + 1)^r - 1}} = 1 - \frac{1}{(e^x + 1)^r}. \end{aligned}$$

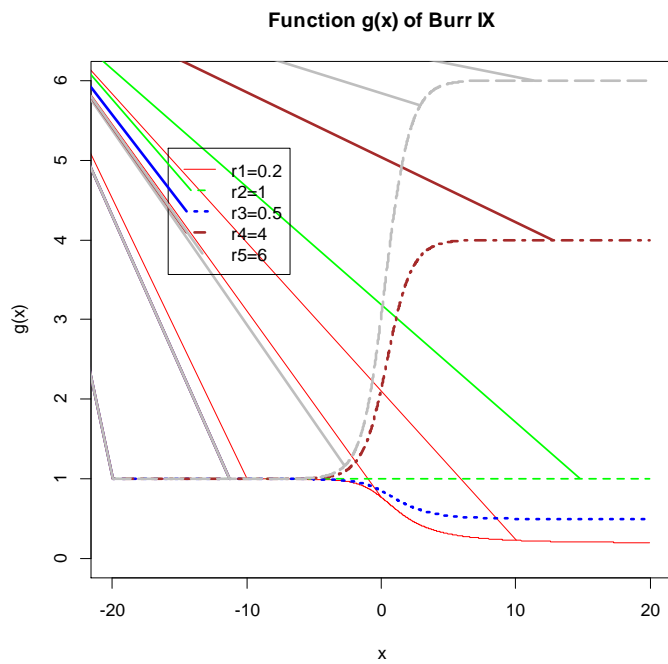
Συνεπώς

$$F(x) = 1 - \frac{1}{(e^x + 1)^r}, \quad x \in (-\infty, \infty), \quad r > 0.$$

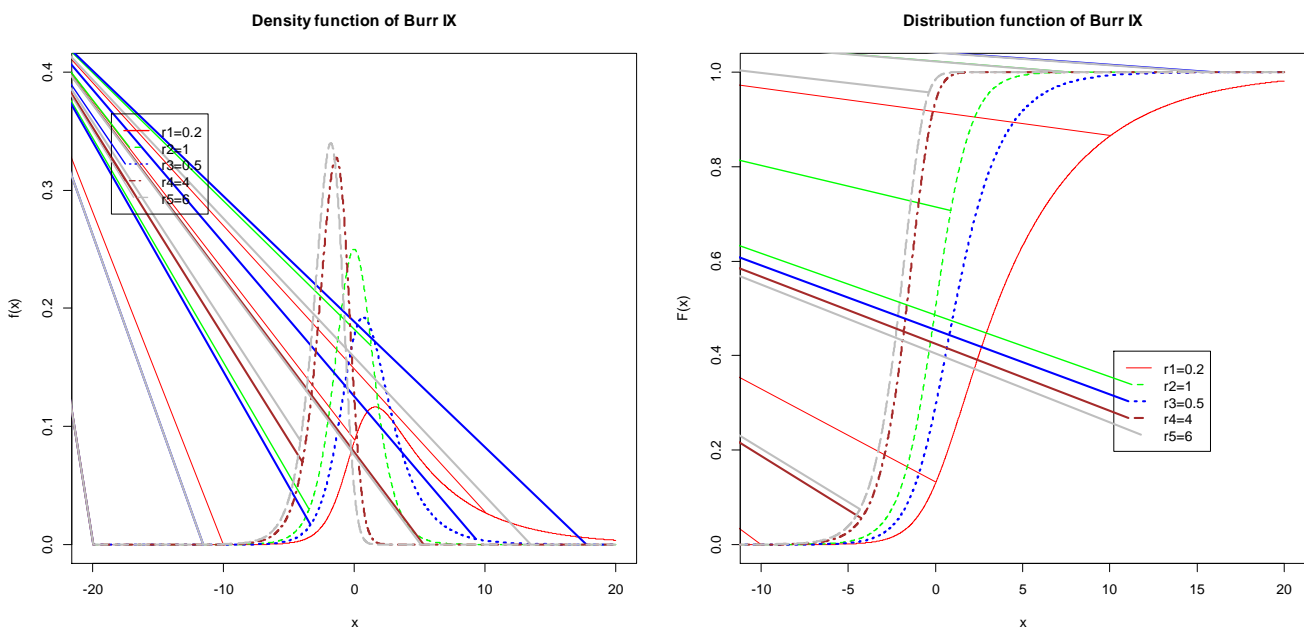
Παραγωγίζοντας την παραπάνω συνάρτηση κατανομής καταλήγουμε στη συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{re^x}{(e^x + 1)^{r+1}}, \quad x \in (-\infty, \infty), \quad r > 0.$$

Παρακάτω δίνονται διάφορα σχήματα για τις συναρτήσεις  $g(x)$ ,  $f(x)$  και  $F(x)$  της κατανομής Burr τύπου IX.



Σχήμα 2. 12 - Συναρτήσεις  $g(x)$  για την κατανομή Burr τύπου IX



Σχήμα 2. 13 - Συναρτήσεις πυκνότητας και κατανομής για την κατανομή Burr τύπου IX

### Επέκταση της Burr IX κατανομής

Σημειώνουμε ότι η κατανομή Burr IX που δόθηκε από τον Burr (1942) είχε συνάρτηση κατανομής

$$F(x; r, k) = 1 - \frac{2}{k((e^x + 1)^r - 1) + 2}, \quad x \in (-\infty, \infty), \quad r > 0, \quad k > 0.$$

Παραγωγίζοντας την παραπάνω συνάρτηση κατανομής καταλήγουμε στη συνάρτηση πυκνότητας πιθανότητας

$$f(x; r, k) = \frac{2rke^x(e^x + 1)^{r-1}}{(k((e^x + 1)^r - 1) + 2)^2}, \quad x \in (-\infty, \infty), \quad r > 0, \quad k > 0.$$

Θέτοντας  $k = 2$  προκύπτει η προαναφερθείσα κατανομή Burr IX με μια μόνο παράμετρο την  $r$ . Για περισσότερες λεπτομέρειες δείτε Singh and Zhang (2020).

## 2.8 Γνωστές κατανομές που προκύπτουν από την κατανομή

### Burr

Για πολλά χρόνια, χρησιμοποιώντας διαφορετικές τεχνικές, πολλοί ερευνητές αναπτύσσουν διάφορες τροποποιημένες μορφές της μητρικής κατανομής για να έχουν μεγαλύτερη ευελιξία στη μοντελοποίηση πραγματικών δεδομένων. Είναι χρήσιμο να αναφέρουμε μερικές από τις πιο γνωστές γενικεύσεις/τροποποιήσεις των κατανομών Burr.

#### Παραλλαγές της Burr III κατανομής

Από το σύστημα των κατανομών Burr, το μοντέλο κατανομής Burr III χρησιμοποιείται ευρέως. Οι Kham et al. (2021) εισήγαγαν την NMBIII κατανομή με συνάρτηση κατανομής

$$F(x; k, c, \lambda) = (1 + x^{-k}e^{-\lambda x})^{-c}$$

και συνάρτηση πυκνότητας πιθανότητας

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

Η αντίστοιχη συνάρτηση επιβίωσης, όπως και η συνάρτηση κινδύνου δίνονται παρακάτω από τους τύπους

$$S(x; k, c, \lambda) = 1 - (1 + x^{-k}e^{-\lambda x})^{-c}$$

και

$$h(x; k, c) = \frac{c \left( \frac{\lambda + k}{x} \right) (1 + x^{-k} e^{-\lambda x})^{-c-1}}{1 - (1 + x^{-k} e^{-\lambda x})^{-c}}.$$

Στον Πίνακα 2.2 παρουσιάζονται διάφορες ειδικές περιπτώσεις της NMBIII κατανομής (δείτε Kham et al. (2021)).

| Μοντέλο   | $\lambda$ | $c$ | $k$ | $F(x)$  |
|---|-----------|-----|-----|---|
| Burr III  | 0         | -   | -   | $(1 + x^{-k})^{-c}$                                 |
| Log logistic  | 0         | 1   | -   | $\frac{x^k}{1 + x^k}$                               |
| Modified loglogistic                                    | -         | 1   | -   | $\frac{x^k e^{-\lambda x}}{1 + x^k e^{-\lambda x}}$ |
| Logistic  | -         | 1   | 0   | $\frac{e^{-\lambda x}}{1 + e^{-\lambda x}}$         |
| Modified skew logistic                                  | -         | -   | 1   | $\frac{x e^{-\lambda x}}{1 + x e^{-\lambda x}}$     |
| Generalized logistic type I or Burr II or skew logistic | 1         | -   | 0   | $(1 + e^{-x})^{-c}$                                 |

Πίνακας 2. 2 - Ειδικές περιπτώσεις της NMBIII κατανομής

### Παραλλαγές της τριπαραμετρικής Burr XII κατανομής

Πολλοί συγγραφείς κατασκεύασαν γενικεύσεις της κατανομής BXII. Παρακάτω παρατίθεται η συνάρτηση κατανομής και η συνάρτηση πυκνότητας πιθανότητας της Burr XII κατανομής με τρεις παραμέτρους,

$$G_{a,b,c}(x) = 1 - \frac{1}{\left( \frac{x^a}{c^a} + 1 \right)^b}, \quad x \geq 0$$

και

$$g_{a,b,c}(x) = abc^{-a} \frac{x^{a-1}}{\left( \frac{x^a}{c^a} + 1 \right)^{b+1}}$$



όπου  $a > 0$  και  $b > 0$  είναι παράμετροι μορφής και το  $c > 0$  είναι παράμετρος κλίμακας (δείτε, π.χ., Gad et al (2019)). Για  $a = 1$  το μοντέλο BXII ανάγεται στο μοντέλο Lomax (Lx) ή Pareto τύπου II (PaII). Για  $b = 1$  το μοντέλο BXII ανάγεται στο μοντέλο log-logistic (LL). Όταν  $a = c = 1$ , το μοντέλο Burr XII ανάγεται στο μονοπαραμετρικό μοντέλο Lx ή μονοπαραμετρικό μοντέλο PaII. Για  $b = c = 1$ , το μοντέλο BXII ανάγεται στο μονοπαραμετρικό μοντέλο LL. Για  $c = 1$  το μοντέλο BXII ανάγεται στο μοντέλο Burr XII δύο παραμέτρων. Όταν  $b \rightarrow \infty$  το μοντέλο BXII ανάγεται στο διπαραμετρικό μοντέλο της κατανομής Weibull (W). Ακόμη, όταν  $c = 1$  και  $b \rightarrow \infty$  το μοντέλο BXII ανάγεται στην κατανομή Weibull (W) με μια παράμετρο.

Οι Cordeiro et al. (2018) όρισαν την οικογένεια κατανομών Burr XII-G (BXII-G) με συνάρτηση κατανομής

$$F_{\theta}(x) = 1 - \{[G_{\psi}(x) - 1]^{-\alpha} + 1\}^{-\beta}, \quad x \in \mathcal{R}$$

και συνάρτηση πυκνότητας πιθανότητας

$$f_{\theta}(x) = \alpha\beta \frac{g_{\psi}(x)G_{\psi}(x)^{\alpha-1}}{[1 - G_{\psi}(x)]^{\alpha+1}} \{[G_{\psi}(x) - 1]^{-\alpha} + 1\}^{-\beta-1}, \quad x \in \mathcal{R}$$

όπου  $g_{\psi}(x)$  είναι η βασική συνάρτηση πυκνότητας πιθανότητας και  $\theta = (\alpha, \beta, a, b, c)$ .

Το νέο μοντέλο BXII-BXII (Gad et al. (2019)) έχει συνάρτηση κατανομής

$$F_{\theta}(x) = 1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^{\alpha} + 1 \right\}^{-\beta}, \quad x \geq 0$$

και συνάρτηση πυκνότητας πιθανότητας

$$f_{\theta}(x) = \alpha\beta abc^{-a} x^{a-1} \left( \frac{x^a}{c^a} + 1 \right)^{b-1} \frac{\left[ - \left( \frac{x^a}{c^a} + 1 \right)^{-b} + 1 \right]^{\alpha-1}}{\left\{ \left[ \left( \frac{x^a}{c^a} + 1 \right)^b - 1 \right]^{\alpha} + 1 \right\}^{\beta+1}}, \quad x > 0.$$

Στον Πίνακα 2.3 παρουσιάζονται διάφορες ειδικές περιπτώσεις της BXII-BXII κατανομής (δείτε π.χ., Tadikamalla (1980) και Gad et al. (2019)).

| <b>N</b> | <b><math>\alpha</math></b> | <b><math>\beta</math></b> | <b><math>a</math></b> | <b><math>b</math></b> | <b><math>c</math></b> | <b>Κατανομή</b> | <b>F(x)</b>  |
|----------|----------------------------|---------------------------|-----------------------|-----------------------|-----------------------|-----------------|--|
| 1        | -                          | $\rightarrow\infty$       | -                     | -                     | -                     | W-BXII          | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-\beta}$ |
| 2        | -                          | $\rightarrow\infty$       | -                     | -                     | 1                     | W-BXII          | $1 - \{[(1 + x^a)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 3        | -                          | $\rightarrow\infty$       | 1                     | -                     | -                     | W-Lx            | $1 - \{[(1 + c^{-1}x)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 4        | -                          | $\rightarrow\infty$       | 1                     | -                     | 1                     | W-Lx            | $1 - \{[(1 + x)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 5        | -                          | $\rightarrow\infty$       | -                     | 1                     | -                     | W-LL            | $1 - \left[ \left( \frac{x^a}{c^a} \right)^\alpha + 1 \right]^{-\beta}$                            |
| 6        | -                          | $\rightarrow\infty$       | -                     | 1                     | 1                     | W-LL            | $1 - (x^{a\alpha} + 1)^{-\beta}$   |
| 7        | -                          | $\rightarrow\infty$       | -                     | $\rightarrow\infty$   | -                     | W-W             | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-\beta}$ |
| 8        | -                          | $\rightarrow\infty$       | -                     | $\rightarrow\infty$   | 1                     | W-W             | $1 - \{[(1 + x^a)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 9        | 1                          | -                         | -                     | -                     | -                     | Lx-BXII         | $1 - \left[ \left( 1 + \frac{x^a}{c^a} \right)^b \right]^{-\beta}$                                 |
| 10       | 1                          | -                         | -                     | $\rightarrow\infty$   | -                     | Lx-W            | $1 - \left[ \left( 1 + \frac{x^a}{c^a} \right)^b \right]^{-\beta}$                                 |
| 11       | 1                          | -                         | -                     | $\rightarrow\infty$   | 1                     | Lx-W            | $1 - [(1 + x^a)^b]^{-\beta}$   |
| 12       | 1                          | -                         | 1                     | -                     | -                     | Lx-Lx           | $1 - [(1 + c^{-1}x)^b]^{-\beta}$   |
| 13       | 1                          | -                         | 1                     | -                     | 1                     | Lx-Lx           | $1 - [(1 + x)^b]^{-\beta}$   |
| 14       | 1                          | -                         | -                     | 1                     | -                     | Lx-LL           | $1 - \left[ \left( 1 + \frac{x^a}{c^a} \right)^b \right]^{-1}$                                     |
| 15       | 1                          | -                         | -                     | 1                     | 1                     | Lx-LL           | $1 - [(1 + x^a)^b]^{-1}$   |
| 16       | 1                          | -                         | -                     | -                     | 1                     | Lx-BXII         | $1 - [(1 + x^a)^b]^{-\beta}$   |
| 17       | -                          | 1                         | -                     | -                     | -                     | LL-BXII         | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-1}$     |
| 18       | -                          | 1                         | -                     | 1                     | -                     | LL-LL           | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-1}$     |
| 19       | -                          | 1                         | -                     | 1                     | 1                     | LL-LL           | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-1}$     |
| 20       | -                          | 1                         | -                     | $\rightarrow\infty$   | -                     | LL-W            | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-1}$     |
| 21       | -                          | 1                         | -                     | $\rightarrow\infty$   | 1                     | LL-W            | $1 - \{[(1 + x^a)^b - 1]^\alpha + 1\}^{-1}$  |

|    |   |                      |   |                      |   |           |  |
|----|---|----------------------|---|----------------------|---|-----------|--|
| 22 | - | 1                    | 1 | -                    | - | LL-Lx     | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-1}$     |
| 23 | - | 1                    | 1 | -                    | 1 | LL-Lx     | $1 - \{[(1+x^a)^b - 1]^\alpha + 1\}^{-1}$  |
| 24 | - | 1                    | - | -                    | 1 | LL-BXII   | $1 - \{[(1+x^a)^b - 1]^\alpha + 1\}^{-1}$  |
| 25 | 1 | 1                    | - | -                    | - | BXII      | $1 - \left( 1 + \frac{x^a}{c^a} \right)^{-b}$  |
| 26 | 1 | 1                    | - | $\rightarrow \infty$ | - | W         | $1 - \left( 1 + \frac{x^a}{c^a} \right)^{-b}$  |
| 27 | 1 | 1                    | - | -                    | 1 | BXII      | $1 - (1+x^a)^{-b}$   |
| 28 | 1 | 1                    | - | $\rightarrow \infty$ | 1 | W         | $1 - (1+x^a)^{-b}$   |
| 29 | 1 | 1                    | 1 | -                    | - | Lx        | $1 - (1+c^{-1}x)^{-b}$   |
| 30 | 1 | 1                    | - | 1                    | - | LL        | $1 - \left( 1 + \frac{x^a}{c^a} \right)^{-1}$  |
| 31 | 1 | 1                    | 1 | -                    | 1 | Lx        | $1 - (1+x)^{-b}$   |
| 32 | 1 | 1                    | - | 1                    | 1 | LL        | $1 - (1+x^a)^{-1}$   |
| 33 | - | -                    | - | -                    | 1 | BXII-BXII | $1 - \{[(1+x^a)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 34 | - | -                    | - | $\rightarrow \infty$ | - | BXII-W    | $1 - \left\{ \left[ \left( 1 + \frac{x^a}{c^a} \right)^b - 1 \right]^\alpha + 1 \right\}^{-\beta}$ |
| 35 | - | -                    | - | $\rightarrow \infty$ | 1 | BXII-W    | $1 - \{[(1+x^a)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 36 | - | -                    | 1 | -                    | - | BXII-Lx   | $1 - \{[(1+c^{-1}x)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 37 | - | -                    | 1 | -                    | 1 | BXII-Lx   | $1 - \{[(1+x)^b - 1]^\alpha + 1\}^{-\beta}$  |
| 38 | - | -                    | - | 1                    | - | BXII-LL   | $1 - \left[ \left( \frac{x^a}{c^a} \right)^a + 1 \right]^{-\beta}$                                 |
| 39 | - | -                    | - | 1                    | 1 | BXII-LL   | $1 - (x^{aa} + 1)^{-\beta}$  |
| 40 | 1 | $\rightarrow \infty$ | 1 | $\rightarrow \infty$ | - | W-W       | $1 - (1+c^{-1}x)^{-b\beta}$  |

Πίνακας 2. 3 - Ειδικές περιπτώσεις της BXII-BXII κατανομής

# ΚΕΦΑΛΑΙΟ 3

## Εφαρμογή των κατανομών Burr σε πραγματικά δεδομένα

---

### 3.1 Εισαγωγή

Στην ενότητα αυτή θα προσαρμόσουμε κατανομές Burr σε τρία σύνολα δεδομένων και ταυτόχρονα θα τα συγκρίνουμε με άλλες γνωστές κατανομές. Η ανάλυσή μας θα βασιστεί στην γλώσσα προγραμματισμού R. Τα πιο σημαντικά εργαλεία για τη μελέτη της κατανομής Burr τα λαμβάνουμε από τα πακέτα `fitdistrplus` και `actuar` της R. Αρχικά θα εξεταστούν κάποια βασικά περιγραφικά μέτρα των δεδομένων ώστε να λάβουμε μια σαφή εικόνα του εκάστοτε δείγματος και εν συνεχεία θα προσαρμόσουμε στο κάθε δείγμα κάποιες κατανομές, μέσα στις οποίες συγκαταλέγεται και η Burr, με απώτερο σκοπό να παρατηρήσουμε ποια κατανομή εφαρμόζει καλύτερα στο κάθε ένα από τα σετ δεδομένων μας. Σε ότι αφορά τις συγκρίσεις με άλλες κατανομές θα γίνει χρήση του κριτηρίου Akaike.

### 3.2 Δεδομένα

Θα διεξάγουμε ανάλυση για τρία σετ δεδομένων. Το πρώτο σετ δεδομένων που θα χρησιμοποιηθεί ονομάζεται Danish και αποτελείται από ζημιές που προκλήθηκαν λόγω πυρκαγιάς στην Δανία. Τα δεδομένα αποτελούνται από 2167 ζημιές και αφορούν το χρονικό διάστημα 3 Ιανουαρίου 1980 έως 31 Δεκεμβρίου 1990. Αξίζει να σημειωθεί ότι τα δεδομένα καταγράφηκαν από την αντασφαλιστική εταιρεία Copenhagen Reinsurance και ότι η κάθε ζημιά αφορά μεμονωμένη απώλεια άνω του ενός εκατομμυρίου δανέζικων κορόνων (DKK). Στο σύνολο δεδομένων Danish έχει ήδη προσαρμοστεί η επίδραση του πληθωρισμού σε τιμές 1985. Τα δεδομένα αυτά βρίσκονται στο πακέτο `QRM` της R. Το δεύτερο σετ δεδομένων ονομάζεται Annual Rainfall Amount και έχει τιμές της ετήσιας βροχόπτωσης σε mm που καταγράφηκαν στο σταθμό μέτρησης U330058 (Akron-Canton WSOAP, Οχάιο, Η.Π.Α) από το 1953 έως το 2010. Τέλος, το τρίτο σετ δεδομένων ονομάζεται Maximum Daily Precipitation και έχει τιμές της μέγιστης ημερήσιας βροχόπτωσης σε mm ανά έτος από το 1948 έως το 2019 που καταγράφηκαν στο σταθμό

μέτρησης U330058 (Akron-Canton WSOAP, Οχάιο, Η.Π.Α). Το δεύτερο και το τρίτο σετ δεδομένων δίνονται στο Παράρτημα.

### 3.3 Ανάλυση δεδομένων

#### Δεδομένα Danish

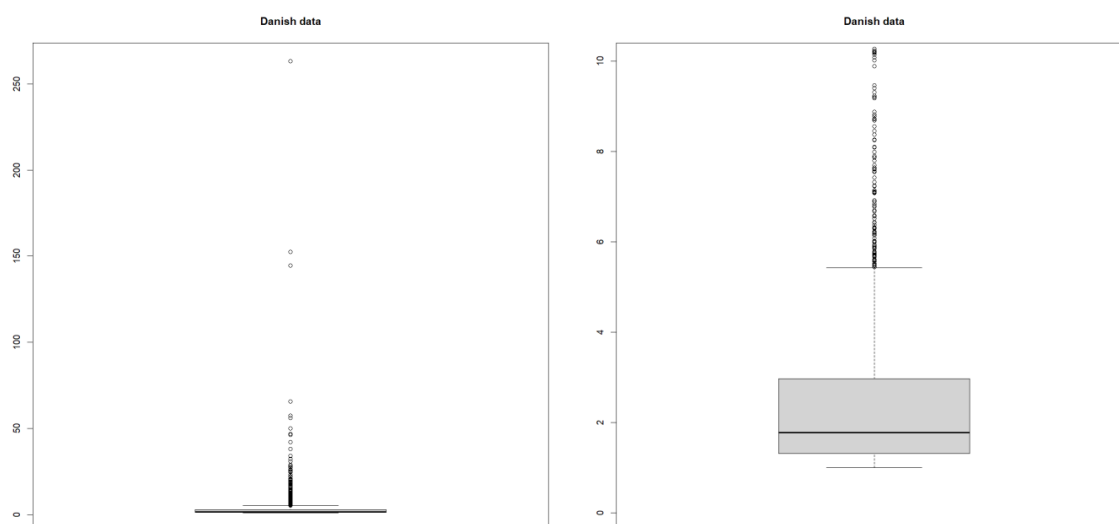
Αρχικά παρατίθενται περιγραφικά μέτρα των δεδομένων Danish.

| Περιγραφικά μέτρα           | Τιμή     |
|-----------------------------|----------|
| Μέση Τιμή                   | 3.385    |
| Διάμεσος                    | 1.778    |
| Τυπική Απόκλιση             | 8.507    |
| Λοξότητα                    | 18.74983 |
| Κύρτωση                     | 485.6461 |
| Ελάχιστο                    | 1        |
| Μέγιστο                     | 263.250  |
| 1 <sup>ο</sup> Τεταρτημόριο | 1.321    |
| 3 <sup>ο</sup> Τεταρτημόριο | 2.967    |
| 0.99 Ποσοστιαίο σημείο      | 26.043   |

Πίνακας 3. 1 - Περιγραφικά μέτρα δεδομένων Danish

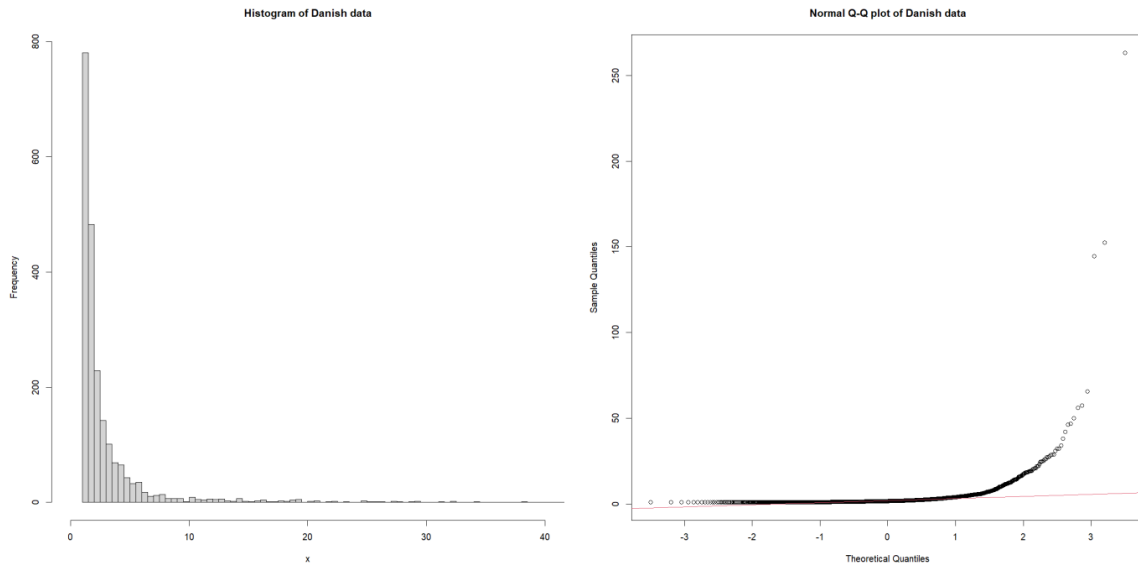
Από τον παραπάνω πίνακα είναι εύκολα αντιληπτό ότι πρόκειται για δεδομένα που έχουν έντονη θετική ασυμμετρία. Παρατηρώντας ακόμη το 99% ποσοστημόριο, το οποίο υποδηλώνει ότι το 99% του δείγματος βρίσκεται μέχρι την τιμή 26.043, αντιλαμβανόμαστε ότι το άνω 1% των παρατηρήσεων αφορά ακραίες τιμές που μπορεί να φτάσουν μέχρι την τιμή 263.250 (μέγιστη τιμή).

Το γεγονός ότι τα δεδομένα έχουν έντονη θετική ασυμμετρία φαίνεται από το θηκόγραμμα αυτών. Το θηκόγραμμα (BoxPlot) είναι μια εύχρηστη μέθοδος γραφικής παράστασης των υπό εξέταση δεδομένων, βάσει του οποίου γίνεται εύκολα αντιληπτή και η ύπαρξη ή μη ακραίων τιμών. Ενδεικτικά, παρατίθεται παρακάτω το θηκόγραμμα των δεδομένων Danish (πλήρες και τμήμα του), στο οποίο παρατηρούμε ότι υπάρχουν πάρα πολλές ακραίες τιμές.



Σχήμα 3. 1 - Θηκόγραμμα δεδομένων Danish

Ακολούθως δίνεται το ιστόγραμμα και το κανονικό διάγραμμα πιθανότητας των δεδομένων. Από το ιστόγραμμα παρατηρούνται πολλές ζημιές λιγότερο κοστοβόρες και λίγες ζημιές περισσότερο κοστοβόρες για την ασφαλιστική εταιρεία. Από το διάγραμμα Q-Q παρατηρείται ότι τα δεδομένα μας δεν προέρχονται από την κανονική κατανομή.



Σχήμα 3. 2 - Ιστόγραμμα και κανονικό Q-Q διάγραμμα δεδομένων Danish

Συνεχίζουμε βρίσκοντας τους εκτιμητές των παραμέτρων των διάφορων κατανομών που προσαρμόζουμε στα δεδομένα Danish και τα κριτήρια AIC και BIC. Για την εύρεση των εκτιμητών χρησιμοποιήθηκε η Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum-Likelihood Estimation), ενώ στις περιπτώσεις που δεν έδινε αποτελέσματα η συγκεκριμένη μέθοδος, χρησιμοποιήθηκε η μέθοδος Εκτίμησης Ελάχιστης Απόστασης (Maximum Goodness-of-Fit Estimation). Εφαρμόζοντας την μέθοδο Ελάχιστης Απόστασης, καταφέρνουμε να ελαχιστοποιήσουμε την «απόσταση» μεταξύ της θεωρητικής και της εμπειρικής συνάρτησης κατανομής των δεδομένων.

- Burr XII

| <b>Εφαρμογή της μεθόδου Maximum Goodness-of-Fit Estimation στην Burr XII κατανομή</b> |           |
|---|-----------|
| Παράμετρος $k$  | 0.106015  |
| Παράμετρος $c$  | 13.100458 |
| Παράμετρος $a$  | 1.082401  |
| Loglikelihood   | -3433.696 |
| AIC   | 6873.392  |

|     |          |
|-----|----------|
| BIC | 6890.435 |
|-----|----------|

Πίνακας 3. 2 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr XII

- Burr III

| Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr III κατανομή |                           |
|---|---------------------------|
| Παράμετρος $a$  | $4.759638 \times 10^{-2}$ |
| Παράμετρος $k$  | 2.171037                  |
| Παράμετρος $r$  | $2.154767 \times 10^3$    |
| Loglikelihood   | -3588.364                 |
| AIC   | 7182.727                  |
| BIC   | 7199.771                  |

Πίνακας 3. 3 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III

- Burr IX

| Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr IX κατανομή |           |
|--|-----------|
| Παράμετρος $k$   | 2.4173992 |
| Παράμετρος $r$   | 0.2519601 |
| Loglikelihood  | -5219.956 |
| AIC  | 10443.91  |
| BIC  | 10455.27  |

Πίνακας 3. 4 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr IX

- Lognormal

| Εφαρμογή της μεθόδου Maximum Likelihood Estimation στη Lognormal κατανομή |           |
|---|-----------|
| Παράμετρος $\mu$  | 0.7869501 |
| Παράμετρος $\sigma$   | 0.7165545 |



|               |           |
|---------------|-----------|
| Loglikelihood | -4057.897 |
| AIC           | 8119.795  |
| BIC           | 8131.157  |

Πίνακας 3. 5 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal

- Pareto

| Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Pareto κατανομή |           |
|---|-----------|
| Παράμετρος $a$  | 5.369439  |
| Παράμετρος $s$  | 13.844821 |
| Loglikelihood   | -4622.833 |
| AIC   | 9249.666  |
| BIC   | 9261.029  |

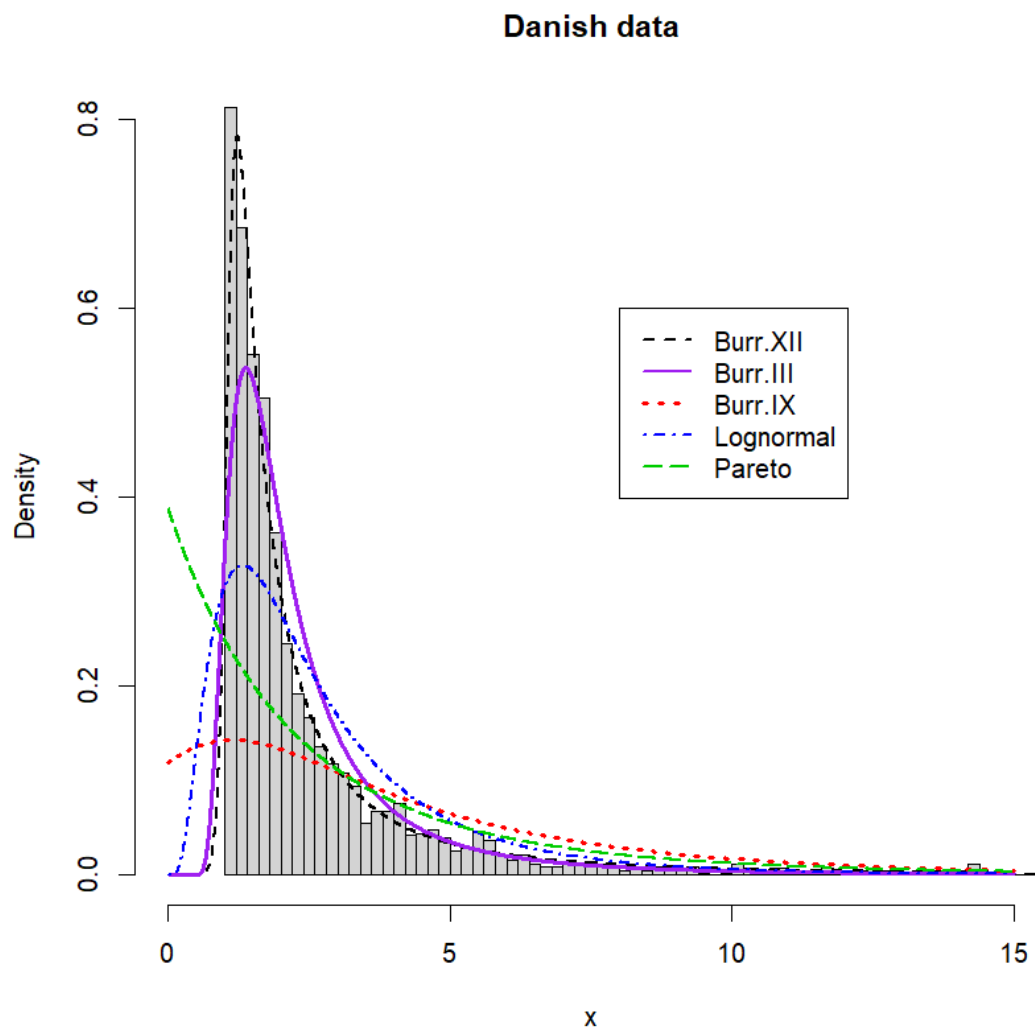
Πίνακας 3. 6 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto

Στη συνέχεια ερευνούμε κατά πόσο οι παραπάνω κατανομές προσαρμόζονται καλά στα δεδομένα Danish σύμφωνα με το κριτήριο AIC. Λαμβάνοντας υπόψη ότι η κατανομή με το μικρότερο AIC εφαρμόζει καλύτερα στα δεδομένα, καταλήγουμε στο συμπέρασμα ότι η Burr XII είναι η πιο κατάλληλη κατανομή για τα δεδομένα Danish συγκριτικά με τις υπόλοιπες.

| Goodness-of-fit criteria       | Burr XII        | Burr III | Burr IX  | Lognormal | Pareto   |
|--------------------------------|-----------------|----------|----------|-----------|----------|
| Akaike's Information Criterion | <b>6873.392</b> | 7182.727 | 10443.91 | 8119.795  | 9249.666 |
| Bayesian Information Criterion | 6890.435        | 7199.771 | 10455.27 | 8131.157  | 9261.029 |

Πίνακας 3. 7 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Danish

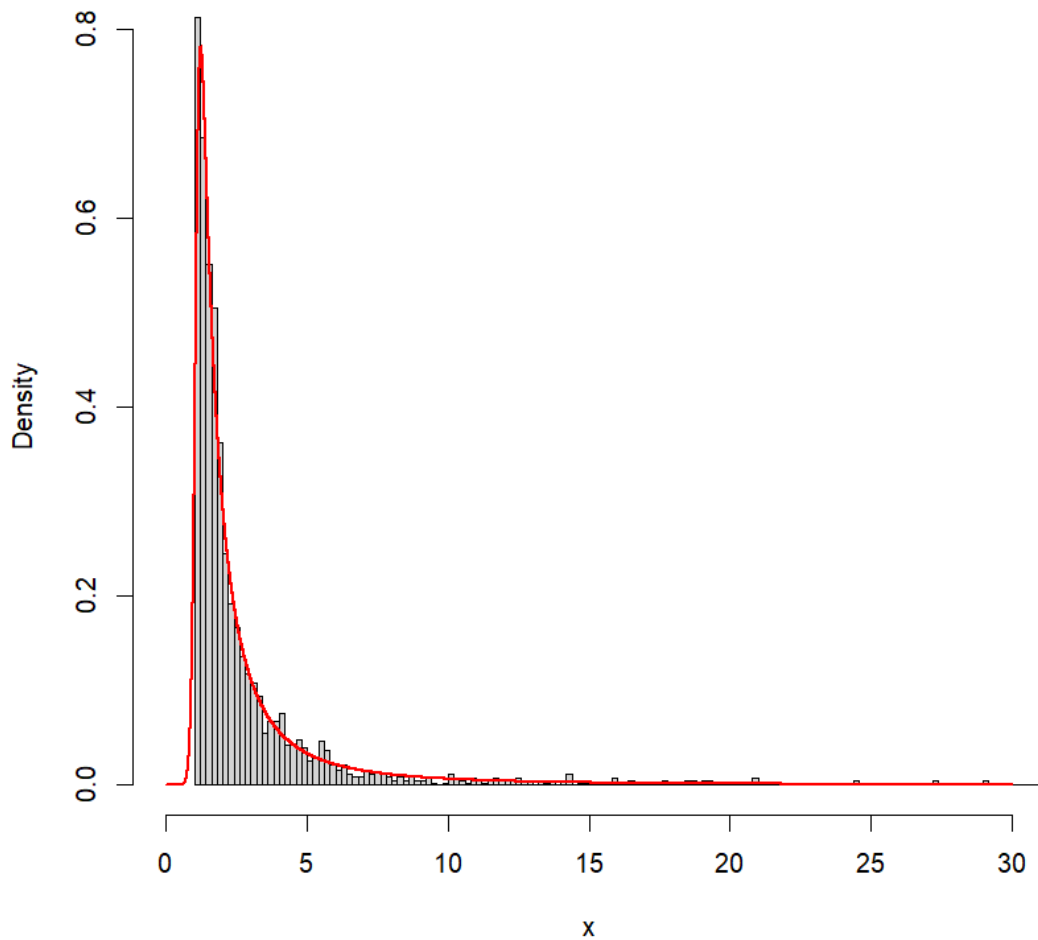
Για λόγους σύγκρισης, παρατίθεται παρακάτω ιστόγραμμα πυκνότητας των δεδομένων Danish πάνω στο οποίο προσαρμόζονται όλες οι κατανομές.



Σχήμα 3. 3 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Danish

Δεδομένου ότι η Burr XII εφαρμόζει καλύτερα στα δεδομένα Danish, παρακάτω παρατίθεται το ιστόγραμμα πυκνότητας των δεδομένων με μόνη προσαρμοσμένη κατανομή την κατανομή Burr XII.

### Fit of Burr.XII on Danish data



Σχήμα 3. 4 - Ιστόγραμμα πυκνότητας και προσαρμοσμένη κατανομή Burr XII στα δεδομένα Danish

### Δεδομένα Annual Rainfall Amount

Για τα δεδομένα Annual Rainfall Amount παρατίθενται παρακάτω τα πιο βασικά περιγραφικά τους μέτρα.

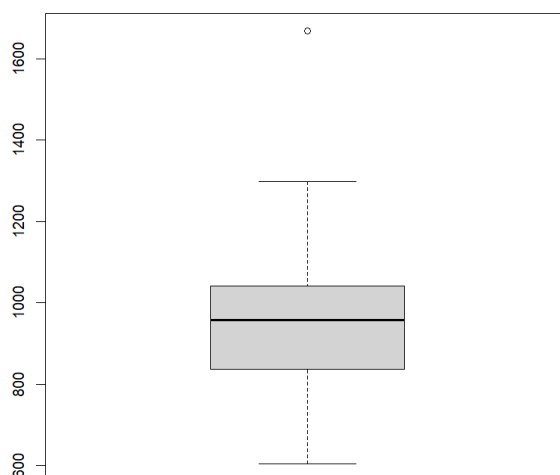
| Περιγραφικά μέτρα | Τιμή     |
|-------------------|----------|
| Μέση Τιμή         | 957.5    |
| Διάμεσος          | 957.8    |
| Τυπική Απόκλιση   | 174.8701 |

|                             |          |
|-----------------------------|----------|
| Λοξότητα                    | 1.007066 |
| Κύρτωση                     | 6.302904 |
| Ελάχιστο                    | 604.3    |
| Μέγιστο                     | 1668.8   |
| 1 <sup>ο</sup> Τεταρτημόριο | 838.8    |
| 3 <sup>ο</sup> Τεταρτημόριο | 1040.5   |
| 0.99 Ποσοστιαίο σημείο      | 1457.401 |

*Πίνακας 3. 8 - Περιγραφικά μέτρα δεδομένων Annual Rainfall Amount*

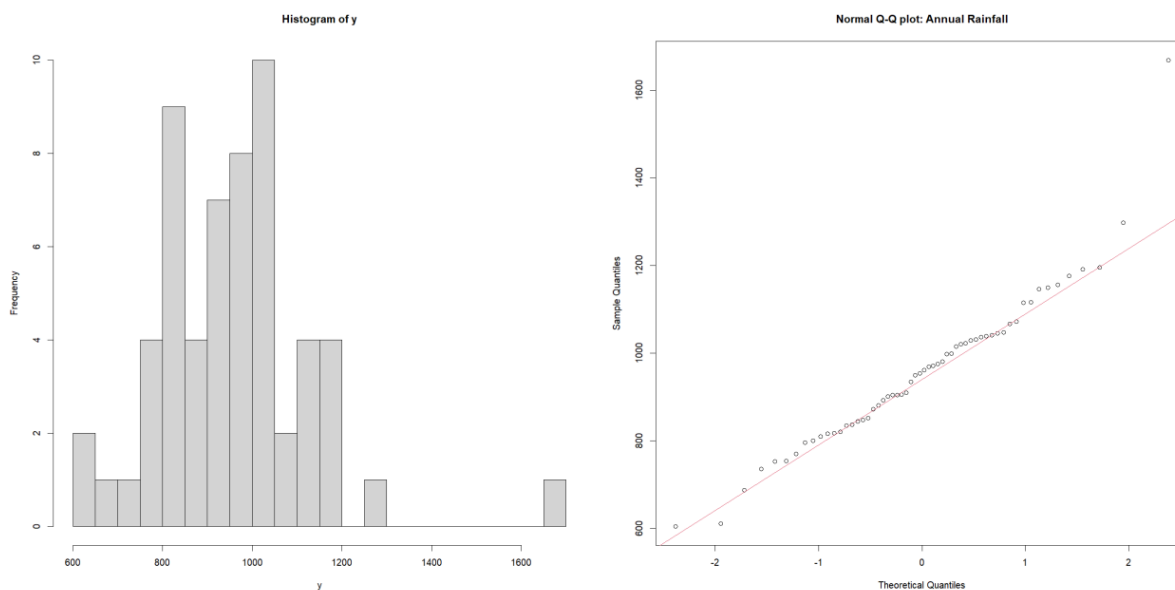
Από τον παραπάνω πίνακα είναι εύκολα αντιληπτό ότι πρόκειται για δεδομένα με ασθενή θετική ασυμμετρία. Παρατηρώντας ακόμη το 99% ποσοστημόριο, το οποίο υποδηλώνει ότι το 99% του δείγματος βρίσκεται μέχρι την τιμή 1457.401, αντιλαμβανόμαστε ότι το 1% των παρατηρήσεων αφορά ακραίες τιμές που μπορεί να φτάσουν μέχρι την τιμή 1668.8 (μέγιστη τιμή).

Το γεγονός ότι τα δεδομένα έχουν ασθενή θετική ασυμμετρία φαίνεται από το θηκόγραμμα αυτών. Ενδεικτικά, παρατίθεται παρακάτω το θηκόγραμμα των δεδομένων.



*Σχήμα 3. 5 - Θηκόγραμμα δεδομένων Annual Rainfall Amount*

Ακολούθως δίνεται το ιστόγραμμα και το κανονικό διάγραμμα πιθανότητας των δεδομένων.



Σχήμα 3. 6 - Ιστόγραμμα και κανονικό Q-Q διάγραμμα δεδομένων Annual Rainfall Amount

Συνεχίζουμε βρίσκοντας τους εκτιμητές των παραμέτρων των διάφορων κατανομών που προσαρμόζουμε στα δεδομένα Annual Rainfall Amount και τα κριτήρια AIC και BIC. Για την εύρεση των εκτιμητών χρησιμοποιήθηκε η Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum Likelihood Estimation).

- Burr XII

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr XII κατανομή</b> |                           |
|--|---------------------------|
| Παράμετρος $k$   | $3.252900 \times 10^{-4}$ |
| Παράμετρος $c$   | $6.914865 \times 10^3$    |
| Παράμετρος $a$   | $6.039220 \times 10^2$    |
| Loglikelihood  | -408.3097                 |
| AIC  | 822.6194                  |
| BIC  | 828.8007                  |

- Burr III

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr III κατανομή</b> |             |
|--|-------------|
| Παράμετρος $a$   | 197.382800  |
| Παράμετρος $k$   | 5.532835    |
| Παράμετρος $r$   | 3504.291065 |
| Loglikelihood  | -384.3576   |
| AIC  | 774.7152    |
| BIC  | 780.8965    |

Πίνακας 3. 10 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III

- Normal

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Κανονική κατανομή</b> |           |
|--|-----------|
| Παράμετρος $\mu$   | 957.5056  |
| Παράμετρος $\sigma$  | 173.3561  |
| Loglikelihood  | -381.3086 |
| AIC  | 766.6172  |
| BIC  | 770.7381  |

Πίνακας 3. 11 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κανονικής κατανομής

- Lognormal

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Lognormal κατανομή</b> |           |
|---|-----------|
| Παράμετρος $\mu$  | 6.8486681 |
| Παράμετρος $\sigma$   | 0.1762919 |

|               |           |
|---------------|-----------|
| Loglikelihood | -378.8556 |
| AIC           | 761.7112  |
| BIC           | 765.832   |

Πίνακας 3. 12 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal

- Pareto

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Pareto κατανομή</b> |            |
|--|------------|
| Παράμετρος $a$   | 1301855    |
| Παράμετρος $s$   | 1235841849 |
| Loglikelihood  | -456.1334  |
| AIC  | 916.2668   |
| BIC  | 920.3877   |

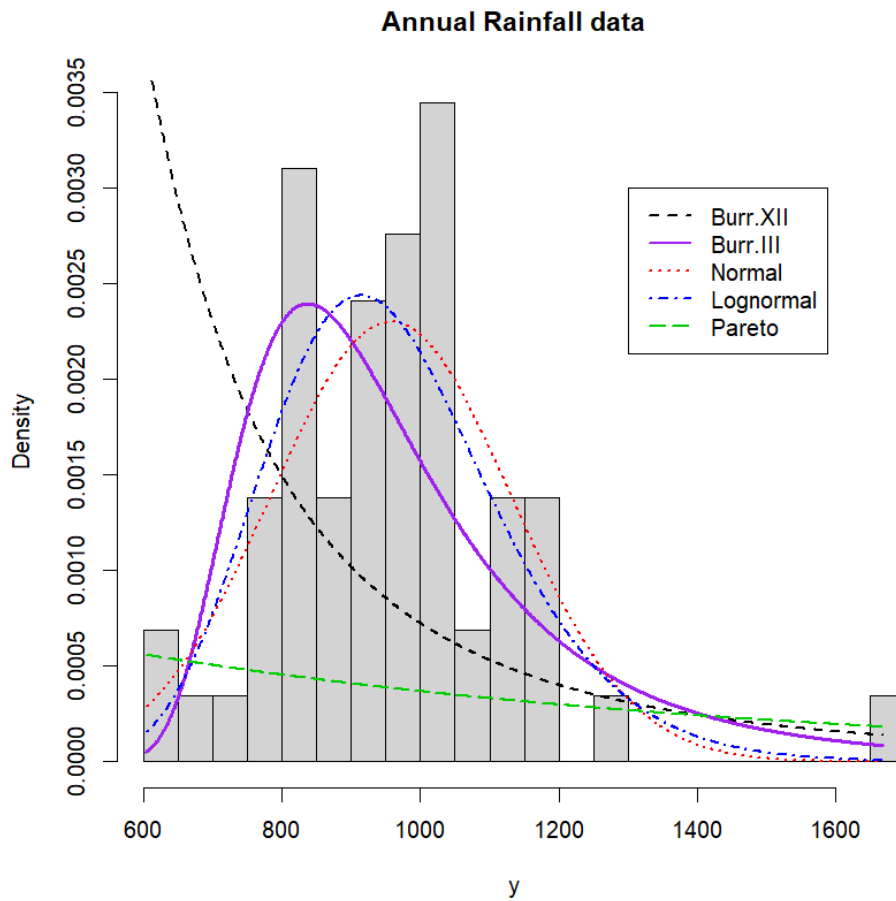
Πίνακας 3. 13 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto

Στη συνέχεια ερευνούμε κατά πόσο οι παραπάνω κατανομές προσαρμόζονται καλά στα δεδομένα Annual Rainfall Amount σύμφωνα με το κριτήριο AIC. Λαμβάνοντας υπόψη ότι η κατανομή με το μικρότερο AIC εφαρμόζει καλύτερα στα δεδομένα, καταλήγουμε στο συμπέρασμα ότι η Lognormal είναι η πιο κατάλληλη κατανομή για τα δεδομένα Annual Rainfall Amount συγκριτικά με τις υπόλοιπες.

| <b>Goodness-of-fit criteria</b> | <b>Burr XII</b> | <b>Burr III</b> | <b>Normal</b> | <b>Lognormal</b> | <b>Pareto</b> |
|---------------------------------|-----------------|-----------------|---------------|------------------|---------------|
| Akaike's Information Criterion  | 822.6194        | 774.7152        | 766.6172      | <b>761.7112</b>  | 916.2668      |
| Bayesian Information Criterion  | 828.8007        | 780.8965        | 770.7381      | 765.8320         | 920.3877      |

Πίνακας 3. 14 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Annual Rainfall Amount

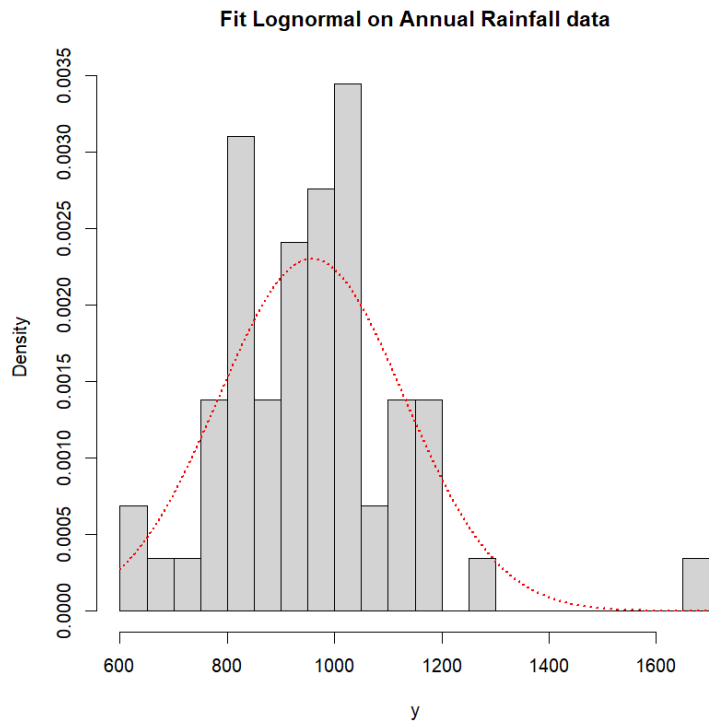
Για λόγους σύγκρισης, παρατίθεται παρακάτω ιστόγραμμα πυκνότητας των δεδομένων Annual Rainfall Amount πάνω στο οποίο προσαρμόζονται όλες οι κατανομές.



Σχήμα 3. 7 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Annual Rainfall Amount

Δεδομένου ότι η Lognormal εφαρμόζει καλύτερα στα δεδομένα Annual Rainfall Amount, παρακάτω παρατίθεται το ιστόγραμμα πυκνότητας των δεδομένων με μόνη προσαρμοσμένη κατανομή την κατανομή Lognormal.





*Σχήμα 3. 8 - Ιστογράμμο πυκνότητας και προσαρμοσμένη κατανομή Lognormal στα δεδομένα Annual Rainfall Data*

### Δεδομένα Maximum Daily Precipitation

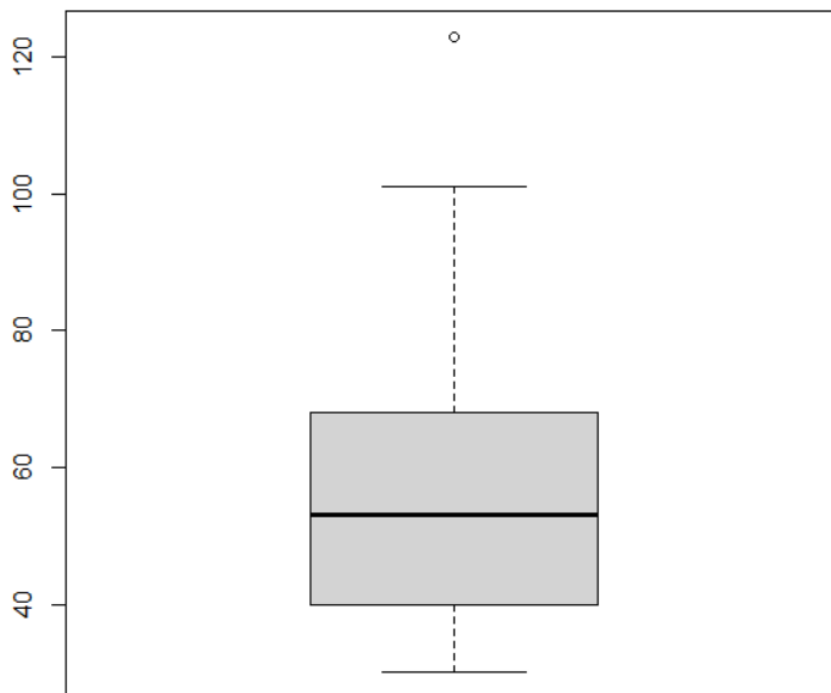
Για τα δεδομένα Maximum Daily Precipitation παρατίθενται παρακάτω τα πιο βασικά περιγραφικά τους μέτρα.

| <b>Περιγραφικά μέτρα</b> | <b>Τιμή</b> |
|--------------------------|-------------|
| Μέση Τιμή                | 56.24       |
| Διάμεσος                 | 53.10       |
| Τυπική Απόκλιση          | 18.61363    |
| Λοξότητα                 | 0.9937395   |
| Κύρτωση                  | 4.084921    |
| Ελάχιστο                 | 30.20       |
| Μέγιστο                  | 122.90      |

|                        |         |
|------------------------|---------|
| 1° Τεταρτημόριο        | 40.20   |
| 3° Τεταρτημόριο        | 67.67   |
| 0.99 Ποσοστιαίο σημείο | 107.422 |

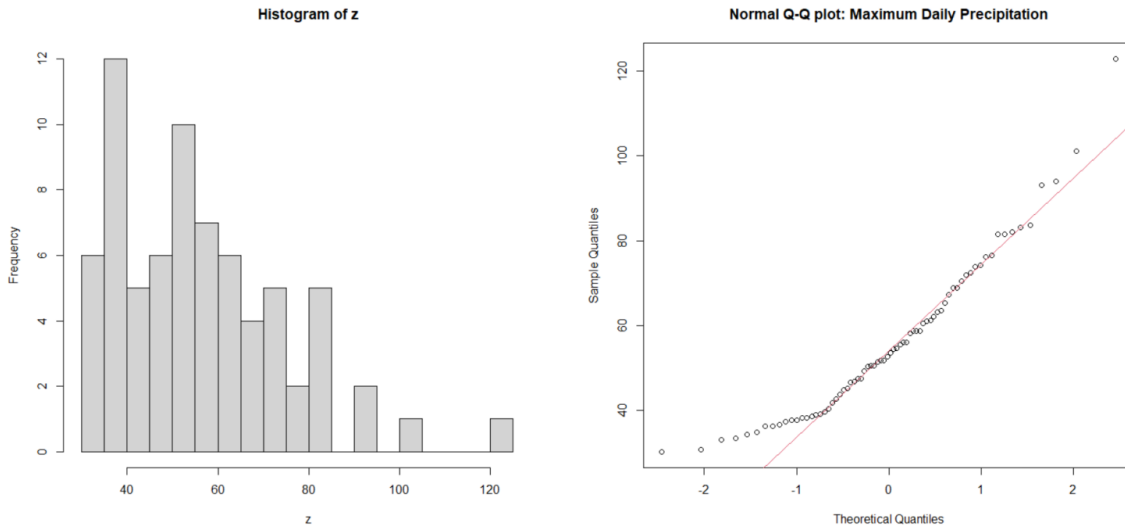
Πίνακας 3. 15 - Περιγραφικά μέτρα δεδομένων *Maximum Daily Precipitation*

Από τον παραπάνω πίνακα είναι εύκολα αντιληπτό ότι πρόκειται για δεδομένα που έχουν ασθενή θετική ασυμμετρία. Παρατηρώντας ακόμη το 99% ποσοστημόριο, το οποίο υποδηλώνει ότι το 99% του δείγματος βρίσκεται μέχρι την τιμή 107.422, ενώ το 1% των παρατηρήσεων αφορά ακραίες τιμές που μπορεί να φτάσουν μέχρι την τιμή 122.90 (μέγιστη τιμή).



Σχήμα 3. 9 - Θηκόγραμμα δεδομένων *Maximum Daily Precipitation*

Ακολούθως δίνεται το ιστόγραμμα και το κανονικό διάγραμμα πιθανότητας των δεδομένων.



Σχήμα 3. 10 - Ιστογράμμο και κανονικό Q-Q διάγραμμα δεδομένων Maximum Daily Precipitation

Συνεχίζουμε βρίσκοντας τους εκτιμητές των παραμέτρων των διάφορων κατανομών που προσαρμόζουμε στα δεδομένα Maximum Daily Precipitation και τα κριτήρια AIC και BIC. Για την εύρεση των εκτιμητών χρησιμοποιήθηκε η Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum Likelihood Estimation).

- Burr XII

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr XII κατανομή</b> |            |
|--|------------|
| Παράμετρος $k$   | 0.8456202  |
| Παράμετρος $c$   | 5.7653753  |
| Παράμετρος $a$   | 50.6582281 |
| Loglikelihood  | -307.0143  |
| AIC  | 620.0287   |
| BIC  | 626.8587   |

Πίνακας 3. 16 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr XII

- Burr III

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr III κατανομή</b> |           |
|--|-----------|
| Παράμετρος $a$   | 27.062792 |
| Παράμετρος $k$   | 3.849551  |
| Παράμετρος $r$   | 8.398621  |
| Loglikelihood  | -305.673  |
| AIC  | 617.346   |
| BIC  | 624.176   |

Πίνακας 3. 17 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr III

- Burr IX

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Burr IX κατανομή</b> |            |
|---|------------|
| Παράμετρος $k$  | 0.01132634 |
| Παράμετρος $r$  | 0.09498053 |
| Loglikelihood   | -311.1727  |
| AIC   | 626.3453   |
| BIC   | 630.8987   |

Πίνακας 3. 18 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Burr IX

- Lognormal

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Lognormal κατανομή</b> |           |
|---|-----------|
| Παράμετρος $\mu$  | 3.9795689 |
| Παράμετρος $\sigma$   | 0.3132986 |
| Loglikelihood   | -305.1294 |
| AIC   | 614.2589  |
| BIC   | 618.8122  |

Πίνακας 3. 19 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Lognormal

- Pareto

| <b>Εφαρμογή της μεθόδου Maximum Likelihood Estimation στην Pareto κατανομή</b> |           |
|--|-----------|
| Παράμετρος $a$   | 9424469   |
| Παράμετρος $s$   | 530567628 |
| Loglikelihood  | -362.139  |
| AIC  | 728.2779  |
| BIC  | 732.8312  |

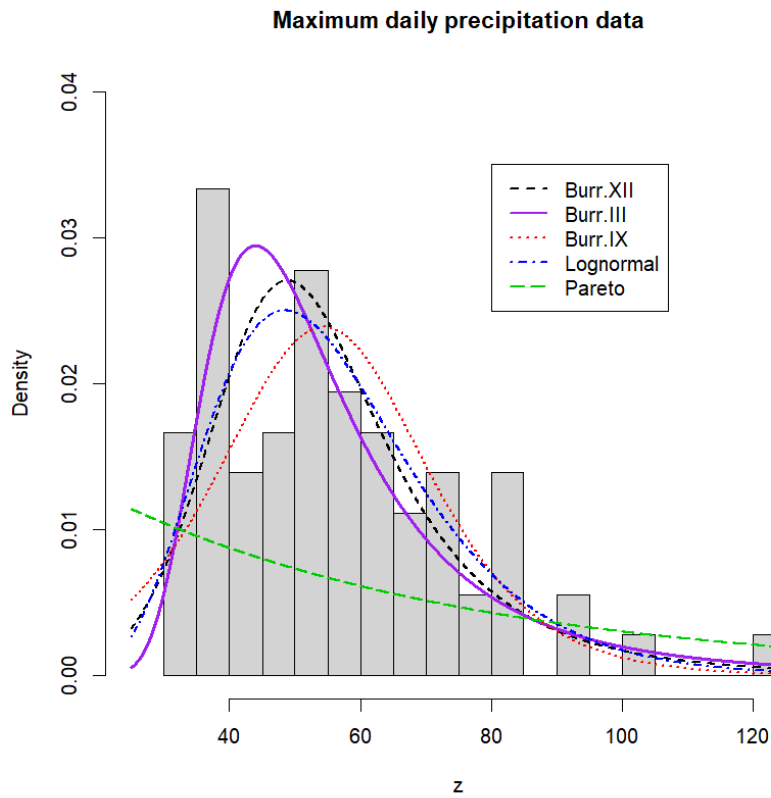
Πίνακας 3. 20 - Εκτιμήσεις παραμέτρων και κριτήρια AIC και BIC της κατανομής Pareto

Στη συνέχεια ερευνούμε κατά πόσο οι παραπάνω κατανομές προσαρμόζονται καλά στα δεδομένα Maximum Daily Precipitation σύμφωνα με το κριτήριο AIC. Λαμβάνοντας υπόψη ότι η κατανομή με το μικρότερο AIC εφαρμόζει καλύτερα στα δεδομένα, καταλήγουμε στο συμπέρασμα ότι η Lognormal είναι η πιο κατάλληλη κατανομή για τα δεδομένα Maximum Daily Precipitation συγκριτικά με τις υπόλοιπες.

| <b>Goodness-of-fit criteria</b> | <b>Burr XII</b> | <b>Burr III</b> | <b>Burr IX</b> | <b>Lognormal</b> | <b>Pareto</b> |
|---------------------------------|-----------------|-----------------|----------------|------------------|---------------|
| Akaike's Information Criterion  | 620.0287        | 617.346         | 626.3453       | <b>614.2589</b>  | 728.2779      |
| Bayesian Information Criterion  | 626.8587        | 624.176         | 630.8987       | 618.8122         | 732.8312      |

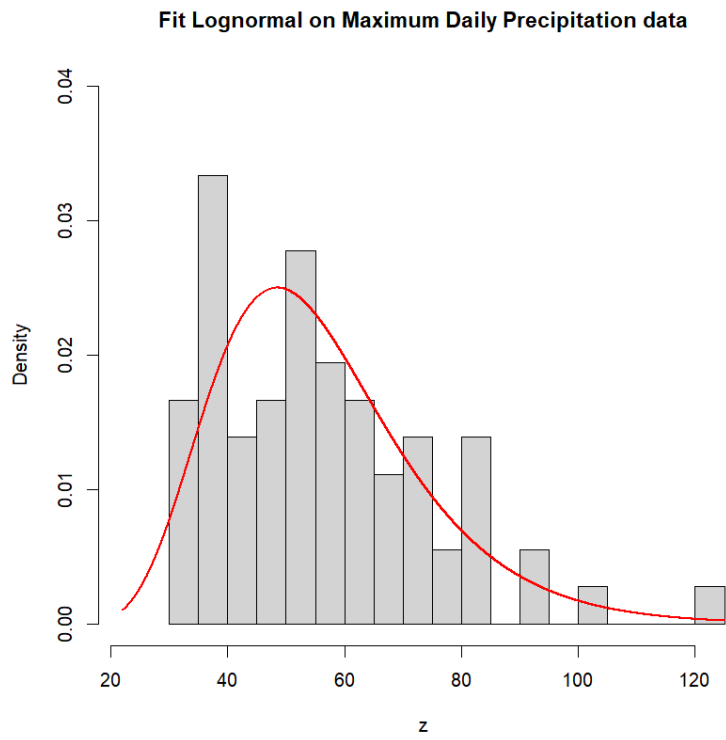
Πίνακας 3. 21 - Πίνακας κριτηρίων AIC και BIC των δεδομένων Maximum Daily Precipitation

Για λόγους σύγκρισης, παρατίθεται παρακάτω ιστόγραμμα πυκνότητας των δεδομένων Maximum Daily Precipitation πάνω στο οποίο προσαρμόζονται όλες οι κατανομές.



*Σχήμα 3. 11 - Οι πέντε προσαρμοσμένες κατανομές στα δεδομένα Maximum Daily Precipitation*

Δεδομένου ότι η Lognormal εφαρμόζει καλύτερα στα δεδομένα Maximum Daily Precipitation, παρακάτω παρατίθεται το ιστόγραμμα πυκνότητας των δεδομένων με μόνη προσαρμοσμένη κατανομή την κατανομή Lognormal.



*Σχήμα 3. 12 - Ιστόγραμμα πυκνότητας και προσαρμοσμένη κατανομή Lognormal στα δεδομένα Maximum Daily Precipitation*

### 3.4 Συμπεράσματα

Στόχος αυτού του κεφαλαίου ήταν η εφαρμογή της κατανομής Burr σε διάφορα σετ δεδομένων και η σύγκρισή της με την εφαρμογή άλλων κατανομών. Παρατηρήθηκε λοιπόν ότι, έχει αρκετά καλή εφαρμογή η κατανομή Burr και στα τρία τυχαία δείγματα. Πιο συγκεκριμένα, στα δεδομένα Danish η κατανομή Burr XII είχε την καλύτερη εφαρμογή συγκριτικά με τις υπόλοιπες κατανομές, ενώ στα άλλα δύο σετ δεδομένων μπορεί να μην είχε την καλύτερη εφαρμογή, ωστόσο προσέγγιζε σε έναν αρκετά ικανοποιητικό βαθμό τα δεδομένα μας. Άρα, εύκολα μπορούμε να συμπεράνουμε ότι είναι μια κατανομή με πολλές εναλλακτικές μορφές, όπου επιλέγοντας την κατάλληλη και προσαρμόζοντας την στα εκάστοτε δεδομένα, μπορούμε να εξάγουμε αποτελέσματα και συμπεράσματα συμβατά με τα εμπειρικά δεδομένα. Συνεπώς, η κατανομή Burr θα πρέπει να συμπεριλαμβάνεται στις κατανομές που χρησιμοποιεί ένας αναλογιστής ή διαχειριστής κινδύνων στις έρευνές του.



# ΠΑΡΑΡΤΗΜΑ

## Κεφάλαιο 2

### Κώδικας για τη διεξαγωγή των σχημάτων 2.1 έως 2.4

#### Διαγράμματα συνάρτησης πυκνότητας κατανομής Burr XII

##### Σχήμα 2.1 (α)

```
par(mfrow=c(1,2))
c1=2; k1=0.5
c2=2; k2=1
c3=2; k3=1.5
c4=2; k4=2.5
x <- seq(0,4,0.01)
y1 <- c1*k1*(x^(c1-1)) / ((1+x^c1)^(k1+1))
y2 <- c2*k2*(x^(c2-1)) / ((1+x^c2)^(k2+1))
y3 <- c3*k3*(x^(c3-1)) / ((1+x^c3)^(k3+1))
y4 <- c4*k4*(x^(c4-1)) / ((1+x^c4)^(k4+1))
plot(x,y1, xlim = c(0,4), type="l", ylim = c(0,1.3), col="red",
main="Density plot of Burr XII", ylab="f(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="brown", lty=4, lwd=3)
legend(locator(1),c("k=0.5, c=2","k=1, c=2","k=1.5, c=2","k=2.5,
c=2"),col=c("red","green","blue","brown"), lty=1:4, lwd=c(1,2,3,3))
```

##### Σχήμα 2.1 (β)

```
c1=0.2; k1=2
c2=0.5; k2=2
c3=0.8; k3=2
c4=1; k4=2
x <- seq(0,4,0.01)
y1 <- c1*k1*(x^(c1-1)) / ((1+x^c1)^(k1+1))
y2 <- c2*k2*(x^(c2-1)) / ((1+x^c2)^(k2+1))
y3 <- c3*k3*(x^(c3-1)) / ((1+x^c3)^(k3+1))
y4 <- c4*k4*(x^(c4-1)) / ((1+x^c4)^(k4+1))
plot(x,y1, xlim = c(0,4), type="l", ylim = c(0,1.3), col="red",
main="Density plot of Burr XII", ylab="f(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="brown", lty=4, lwd=3)
legend(locator(1),c("k=2, c=0.2","k=2, c=0.5","k=2, c=0.8","k=2,
c=1"),col=c("red","green","blue","brown"), lty=1:4, lwd=c(1,2,3,3))
```

#### Διαγράμματα συναρτήσεων κατανομής Burr XII

##### Σχήμα 2.2 (α)

```
par(mfrow=c(1,2))
c1=2; k1=0.5
c2=2; k2=1
c3=2; k3=1.5
c4=2; k4=2.5
x <- seq(0,6,0.01)
y1 <- 1-((1+x^c1)^(-k1))
y2 <- 1-((1+x^c2)^(-k2))
y3 <- 1-((1+x^c3)^(-k3))
y4 <- 1-((1+x^c3)^(-k4))
plot(x,y1, xlim = c(0,6), type="l", ylim = c(0,1), col="red",
main="Distribution function plots of Burr XII distributions",
xlab="x", ylab="F(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
```

```
lines(x,y4, col="grey", lty=4, lwd=4)
legend(locator(1),c("k=0.5, c=2","k=1, c=2","k=1.5, c=2","k=2.5,
c=2"),col=c("red","green","blue","grey"), lty=1:4, lwd=c(1,2,3,4))
```

### Σχήμα 2.2 (β)

```
c1=0.2; k1=2
c2=0.5; k2=2
c3=0.8; k3=2
c4=1; k4=2
x <- seq(0,6,0.01)
y1 <- 1-((1+x^c1)^(-k1))
y2 <- 1-((1+x^c2)^(-k2))
y3 <- 1-((1+x^c3)^(-k3))
y4 <- 1-((1+x^c3)^(-k4))
plot(x,y1, xlim = c(0,6), type="l", ylim = c(0,1), col="red",
main="Distribution function plots of Burr XII distributions",
xlab="x",ylab="F(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="grey", lty=4, lwd=4)
legend(locator(1),c("k=2, c=0.2","k=2, c=0.5","k=2, c=0.8","k=2,
c=0.8","k=2, c=1"),col=c("red","green","blue","grey"), lty=1:4,
lwd=c(1,2,3,4))
```

### Διαγράμματα συναρτήσεων επιβίωσης Burr XII

#### Σχήμα 2.3 (α)

```
par(mfrow=c(1,2))
c1=2; k1=0.5
c2=2; k2=1
c3=2; k3=1.5
c4=2; k4=2.5
x <- seq(0,6,0.01)
y1 <- ((1+x^c1)^(-k1))
y2 <- ((1+x^c2)^(-k2))
y3 <- ((1+x^c3)^(-k3))
y4 <- ((1+x^c4)^(-k4))
plot(x,y1, xlim = c(0,6), type="l", ylim = c(0,1), col="red",
main="Survival function plots of Burr XII distributions",xlab="x",
ylab="S(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="grey", lty=4, lwd=4)
legend(locator(1),c("k=0.5, c=2","k=1, c=2","k=1.5, c=2","k=2.5,
c=2"),col=c("red","green","blue","grey"), lty=1:4, lwd=c(1,2,3,4))
```

#### Σχήμα 2.3 (β)

```
c1=0.2; k1=2
c2=0.5; k2=2
c3=0.8; k3=2
c4=1; k4=2
x <- seq(0,6,0.01)
y1 <- ((1+x^c1)^(-k1))
y2 <- ((1+x^c2)^(-k2))
y3 <- ((1+x^c3)^(-k3))
y4 <- ((1+x^c4)^(-k4))
plot(x,y1, xlim = c(0,6), type="l", ylim = c(0,1), col="red",
main="Survival function plots of Burr XII distributions",
xlab="x",ylab="S(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="grey", lty=4, lwd=4)
legend(locator(1),c("k=2, c=0.2","k=2, c=0.5","k=2, c=0.8","k=2,
c=1"),col=c("red","green","blue","grey"), lty=1:4, lwd=c(1,2,3,4))
```

## Διαγράμματα συναρτήσεων κινδύνου Burr XII

### Σχήμα 2.4 (α)

```
par(mfrow=c(1,2))
c1=2; k1=0.5
c2=2; k2=1
c3=2; k3=1.5
c4=2; k4=2.5
x <- seq(0,6,0.01)
y1 <- (k1*c1*(x^(c1-1))*(1+x^c1)^-(k1+1))/(1+x^c1)^(-k1)
y2 <- (k2*c2*(x^(c2-1))*(1+x^c2)^-(k2+1))/(1+x^c2)^(-k2)
y3 <- (k3*c3*(x^(c3-1))*(1+x^c3)^-(k3+1))/(1+x^c3)^(-k3)
y4 <- (k4*c4*(x^(c4-1))*(1+x^c4)^-(k4+1))/(1+x^c4)^(-k4)
plot(x,y1, xlim = c(0,6), type="l", ylim = c(0,3), col="red",
main="Hazard Rate function of Burr XII with constant parameter c",
ylab="h(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="brown", lty=4, lwd=3)
legend(locator(1),c("k=0.5, c=2","k=1, c=2","k=1.5,
c=2","k=2.5,c=2"),col=c("red","green","blue","brown"), lty=1:4,
lwd=c(1,2,3,3))
```

### Σχήμα 2.4 (β)

```
c1=0.2; k1=2
c2=0.5; k2=2
c3=0.8; k3=2
c4=1; k4=2
x <- seq(0,10,0.01)
y1 <- (k1*c1*x^(c1-1))/(1+x^c1)
y2 <- (k2*c2*x^(c2-1))/(1+x^c2)
y3 <- (k3*c3*x^(c3-1))/(1+x^c3)
y4 <- (k4*c4*x^(c4-1))/(1+x^c4)
plot(x,y1, xlim = c(0,10), type="l", ylim = c(0,2), col="red",
main="Burr XII Hazard Rate function with constant parameter k",
ylab="f(x)")
lines(x,y2, col="green", lty=2, lwd=2)
lines(x,y3, col="blue", lty=3, lwd=3)
lines(x,y4, col="brown", lty=4, lwd=3)
legend(locator(1),c("k=2, c=0.2","k=2, c=0.5","k=2,
c=0.8","k=2,c=1"),col=c("red","green","blue","brown"), lty=1:4,
lwd=c(1,2,3,3))
```

## Κεφάλαιο 3

### Κώδικας δεδομένων Danish

```
library(QRM)
library(actuar)
library(fitdistrplus)
library(moments)
data(danish.df)
attach(danish.df)
names(danish.df)
x <- FIRE.LOSSES
summary(x)
sd(x) #Τυπική Απόκλιση
skewness(x) #Λοξότητα
kurtosis(x) #Κύρτωση
quantile(x,0.99) #99% Ποσοστημόριο (VaR0.99)

par(mfrow=c(1,2))
boxplot(x, main="Danish data") #Θηκόγραμμα
```

```

boxplot(x, main="Danish data", ylim=c(0, 10)) #Θηκόγραμμα

par(mfrow=c(1,2))
hist(x, nclass=400, xlim=c(0,40),main="Histogram of Danish data")
#Ιστόγραμμα
qqnorm(x, main="Normal Q-Q plot: Danish"); qqline(x, col = 2)
#Κανονικό διάγραμμα πιθανότητας

# Fitting Burr XII with method MGE
mystart2 <- list(shape1=1, shape2=1, scale=1)
fitBXII<- fitdist(x, "burr",method="mge",gof="CvM", start=mystart2)
# mge: Maximum Goodness-of-Fit Estimation
summary(fitBXII)

#Fitting Burr III with method MLE
dBIII<- function(x, k, r, a) {(k*r)/a * ( ( 1+(x/a)^(-k) )^(-r-1)
) * (x/a)^(-k-1) }
pBIII<- function(q, k, r, a) { 1+(q/a)^(-k) ^(-r)}
mystartBIII<- list(a=1, k=1, r=1)
fitBIII<- fitdist(x, "BIII",start=mystartBIII, lower = c(0,0,0))
summary(fitBIII)

#Fitting Burr IX with method MLE
dBIX<- function(x, k, r) ( 2*k*r*exp(x)*(1+exp(x))^(r-1) ) / (
( 2+k*( -1 + (1+exp(x))^r ) )^2 )
pBIX<- function(q, k, r) 1 - ( 2 / ( 2+k*( -1 + (1+exp(q))^r ) ) )
mystartBIX<- list(k=0.1, r=0.1)
fitBIX<- fitdist(x, "BIX", start=mystartBIX, lower = c(0,0))
summary(fitBIX)

# Fitting Lognormal with method MLE
fitln.MLE<- fitdist(x, "lnorm")
summary(fitln.MLE)

# Fitting Pareto with method MLE
fitP.MLE<- fitdist(x, "pareto")
summary(fitP.MLE)

# Statistics
gofstat(list(fitBXII,fitBIII,fitBIX,fitln.MGE, fitP.MGE),
fitnames=c("Burr XII","BurrIII","BurrIX","Lognormal", "Pareto"))
# Minimum AIC of Burr XII distribution

# Five pdfs in the same graph (prob=TRUE)
hist(x, nclass=1000, xlim=c(0,15), ylim=c(0,0.8), prob=TRUE)
t <- seq(0,15,0.0001)
y1 <-
dburr(t,shape1=fitBXII$estimate[1],shape2=fitBXII$estimate[2],scale=
fitBXII$estimate[3])
lines(t, y1, type="l", xlab="x", ylab="f(x)", col="black", lwd=2,
lty=2)
y2 <-
dBIII(t,a=fitBIII$estimate[1],k=fitBIII$estimate[2],r=fitBIII$estima
te[3])
lines(t, y2, type="l", xlab="x", ylab="f(x)", col="purple", lwd=3,
lty=1)
y3 <- dBIX(t,k=fitBIX$estimate[1],r=fitBIX$estimate[2])
lines(t, y3, type="l", xlab="x", ylab="f(x)", col="red", lwd=2,
lty=3)
y4 <- dlnorm(t, meanlog=fitln.MLE$estimate[1] ,
sdlog=fitln.MLE$estimate[2])
lines(t, y4, type="l", col="blue", lwd=2, lty=4)
y5 <- dpareto(t, shape=fitP.MLE$estimate[1] ,
scale=fitP.MLE$estimate[2])

```

```

lines(t, y5, type="l", col="green3", lwd=2, lty=5)
legend(8, 0.6, c("Burr.XII", "Burr.III", "Burr.IX", "Lognormal",
"Pareto"), lty=c(2,1,3,4,5), col=c("black", "purple", "red", "blue",
"green3"), lwd=c(2,2,2,2,2))

# Histogram for Burr.XII only
hist(x, nclass=1000, xlim=c(0,30), ylim=c(0,0.8), prob=TRUE,
main="Fit Burr.XII on Danish data")
t <- seq(0,30,0.0001)
y1 <-
dburr(t, shapel=fitBXII$estimate[1], shape2=fitBXII$estimate[2], scale=
fitBXII$estimate[3])
lines(t, y1, type="l", xlab="x", ylab="f(x)", col="red", lwd=2)

```

### Κεφάλαιο 3

#### Σετ δεδομένων Annual Rainfall Amount

| Year | Amount (mm) | Year | Amount (mm) |
|------|-------------|------|-------------|
| 1953 | 753.618     | 1982 | 880.872     |
| 1954 | 954.278     | 1983 | 975.106     |
| 1955 | 796.29      | 1984 | 892.302     |
| 1956 | 1195.324    | 1985 | 980.694     |
| 1957 | 816.356     | 1986 | 800.354     |
| 1958 | 949.452     | 1987 | 769.874     |
| 1959 | 1149.096    | 1988 | 843.788     |
| 1960 | 687.832     | 1989 | 998.728     |
| 1961 | 872.49      | 1990 | 1668.78     |
| 1962 | 736.092     | 1991 | 611.632     |
| 1963 | 604.266     | 1992 | 1146.556    |
| 1964 | 1036.828    | 1993 | 1047.242    |
| 1965 | 852.17      | 1994 | 1028.954    |
| 1966 | 752.856     | 1995 | 905.51      |
| 1967 | 809.752     | 1996 | 1191.768    |
| 1968 | 900.938     | 1997 | 820.166     |
| 1969 | 837.184     | 1998 | 1023.112    |
| 1970 | 971.042     | 1999 | 909.574     |
| 1971 | 817.88      | 2000 | 1155.7      |
| 1972 | 1114.552    | 2001 | 834.644     |
| 1973 | 934.466     | 2002 | 1031.748    |
| 1974 | 1072.134    | 2003 | 1297.94     |
| 1975 | 1041.146    | 2004 | 1176.02     |
| 1976 | 905.002     | 2005 | 1045.718    |
| 1977 | 1014.984    | 2006 | 1115.822    |
| 1978 | 847.344     | 2007 | 1038.606    |
| 1979 | 998.474     | 2008 | 1066.8      |
| 1980 | 969.264     | 2009 | 904.494     |
| 1981 | 1020.318    | 2010 | 961.39      |

### Κεφάλαιο 3

#### Σετ δεδομένων Maximum Daily Precipitation

| Year | Amount (mm) | Year | Amount (mm) | Year | Amount (mm) | Year | Amount(mm) |
|------|-------------|------|-------------|------|-------------|------|------------|
| 1948 | 61.2        | 1966 | 43.7        | 1984 | 34.3        | 2002 | 55.4       |
| 1949 | 36.3        | 1967 | 38.9        | 1985 | 54.4        | 2003 | 101.1      |
| 1950 | 44.7        | 1968 | 34.8        | 1986 | 30.7        | 2004 | 58.7       |
| 1951 | 38.6        | 1969 | 61          | 1987 | 50.5        | 2005 | 83.6       |
| 1952 | 49.3        | 1970 | 72.4        | 1988 | 94          | 2006 | 58.7       |
| 1953 | 62          | 1971 | 60.5        | 1989 | 36.6        | 2007 | 52.6       |
| 1954 | 65.3        | 1972 | 68.8        | 1990 | 83.1        | 2008 | 50.5       |
| 1955 | 41.7        | 1973 | 39.6        | 1991 | 33.3        | 2009 | 37.6       |
| 1956 | 67.3        | 1974 | 51.8        | 1992 | 56.1        | 2010 | 46.5       |
| 1957 | 68.8        | 1975 | 76.2        | 1993 | 42.7        | 2011 | 122.9      |
| 1958 | 73.9        | 1976 | 74.2        | 1994 | 93.2        | 2012 | 70.4       |
| 1959 | 71.9        | 1977 | 47.5        | 1995 | 51.8        | 2013 | 53.6       |
| 1960 | 38.1        | 1978 | 33          | 1996 | 47.5        | 2014 | 63.2       |
| 1961 | 81.5        | 1979 | 82          | 1997 | 50.3        | 2015 | 45.2       |
| 1962 | 39.1        | 1980 | 54.6        | 1998 | 51.3        | 2016 | 46.7       |
| 1963 | 30.2        | 1981 | 58.2        | 1999 | 58.7        | 2017 | 76.5       |
| 1964 | 81.5        | 1982 | 37.6        | 2000 | 56.1        | 2018 | 63.5       |
| 1965 | 38.1        | 1983 | 40.4        | 2001 | 36.3        | 2019 | 37.3       |

# ΒΙΒΛΙΟΓΡΑΦΙΑ

## Ελληνική

- Αντζουλάκος Δ., (2013), Σημειώσεις Μαθήματος Ανάλυση Επιβίωσης, ΠΜΣ Εφαρμοσμένη Στατιστική.
- Αντζουλάκος Δ., (2018), Σημειώσεις Μαθήματος Γλώσσα Προγραμματισμού R, Γλώσσα Προγραμματισμού R Με Εφαρμογές Στον Αναλογισμό, ΠΜΣ Αναλογιστική Επιστήμη & Διαχείριση Κινδύνων.
- Τζαβελάς, Γ., (2022), Σημειώσεις Μαθήματος Ζημιοκατανομές και Θεωρία Ακραίων Τιμών, ΠΜΣ Αναλογιστική Επιστήμη & Διαχείριση Κινδύνων.

## Ξένα

- Akaike, H., (1969), Fitting autoregressive models for prediction, *Ann. Inst. Statist. Math.*, 21, 243-247.
- Akaike, H., (1970), Statistical predictor identification, *Ann. Inst. Statist. Math.*, 22, 203-217.
- Akaike, H., (1974), A New Look at the Statistical Model Identification, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, 19(6), 716-723.
- AL-Huniti, A. A. and AL-Dayian, G. R., (2012), Discrete Burr Type III Distribution, *American Journal of Mathematics and Statistics*, 2(5), 145-152.
- Bourguignon, M., Silva, R. B. and Cordeiro, G. M., (2014), The Weibull-G family of probability Distributions, *Journal of Data Science*, 12(1), 53-68.
- Burr, I. W., (1942), Cumulative frequency functions, *Ann. Inst. Statist. Math.*, 13(2), 215-232.
- Burr, I. W. & Cislak, P. J., (1968), On a general system of distributions, I: its curveshape characteristics, II: the sample median, *Journal of the American Statistical Association*, 63(322), 627-635.
- Cordeiro, M., Yousof, H.M., Ramires, T.G. and Ortega, E.M.M, (2018), The Burr XII System of densities: properties, regression model and applications, *Journal of Statistical Computation and Simulation*, 88(3), 432-456.
- Dagum, C., (1977), A New model of personal income-distribution-specification and estimation, *Economie appliquée*, 30(3), 413-37.
- Efron, B. and Gong, G., (1983), A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation, *The American Statistician*, 37(1), 36-48.
- Eugene, N., Lee, C. and Famoye, F., (2002), Beta-normal distribution and its applications, *Comm in Stat-Theory & Methods*, 31(4), 497-512.

- Gad, A. M., Hamedani, G.G., Salehabadi, S. M. and Yousof, H. M., (2019), The Burr XII-Burr XII Distribution: Mathematical Properties And Characterizations, *Pak. J. Statist.*, 229-248.
- Hakim, A. R., Fithriani, I. and Novita, M., (2021), Properties of Burr distribution and its application to heavytailed survival time data, *J. Phys.: Conf. Ser.*, 1725, 012016, 1-10.
- Hannan, E.J and Quinn, B.G., (1979), The Determination of the Order of an Autogression, *J. R. Stat. Soc.*, 41(2), 190-195.
- Helmy, N., (2018), Discrete Burr Type II Distribution, *Scientific Journal of Faculty of Commerce-Al-Azhar University*, 20(1), 1-21.
- Ibrahim, N. A. and Khaleel, M. A., (2020), Generalizations of Burr Type X Distribution with Applications, *ASM Sc. J.*, 13, 1-8.
- Johnson, N. L., Kotz, S. & Balakrishnan, N., (1995), *Continuous Univariate Distributions*, Wiley Series in Probability and Statistics, 1(2).
- Khan, W., Jamal, F., Nasir, M. A., Abuzaid, H. A., Tahir, M. H. et al., (2021), New Modified Burr III distribution, properties and applications, *Mathematical and Computational Applications*, 26(4), 1-17.
- Kumar, D., (2017), The Burr Type XII Distribution With Some Statistical Properties, *Journal of Data Science*, 15(3), 509-534.
- Mallows, C.L., (1973), Some Comments on Cp, *Technometrics*, 15(4), 661-675.
- Merovci, F. et al., (2016), The beta Burr type X distribution properties with application, *SpringerPlus*, 5(697), 1-18.
- Mielke, Jr, P.W. and Johnson, E.S., (1974), Some generalized beta distributions of the second kind having desirable application features in hydrology and meteorology, *Water Resour. Res.*, 10 (2), 223-226.
- Mudholkar, G. S. and Srivastava, D. K., (1993), Exponentiated Weibull family for analysing bathtub failure rate data, *IEEE Trans. on Reliability*, 42(2), 299-302.
- Pandu, R., Tadikamalla, (1980), *A Look at the Burr and Related Distributions*, International Statistical Institute, 48(3), 2-3
- Ristic, M. M. and Balakrishnan, N., (2012), The Gamma-Exponentiated Exponential distribution, *J. Stat Comp & Sim*, 82(8), 1191-1206.
- Rodriguez, R. N., (2014), Burr Distributions, *Wiley StatsRef: Statistics Reference Online*, 1-7.
- Schwarz, G., (1978), ESTIMATING THE DIMENSION OF A MODEL, *The Annals of Statistics*, 6(2), 461-464.
- Singh, V. P. and Zhang, L., (2020), Systems of Frequency Distributions for Water and Environmental Engineering, *Cambridge University Press*, 40-62.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A., (2002), Bayesian measures of model complexity and fit., *J. R. Stat. Soc.*, 64(4), 583–639.



Surles, J. G. and Padgett, W. J., (2001), Inference for reliability and stress-strength for a scaled Burr type X distribution, *Lifetime Data Analysis*, 7(2), 187-200.

Surles, J. G. and Padgett, W. J., (2005), Some properties of a scaled Burr type X distribution, *Journal of Statistical Planning and Inference*, 128(1), 271–280.