

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**  
**Σχολή Χρηματοοικονομικής και Στατιστικής**



**Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ**  
**ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗ**  
**ΚΙΝΔΥΝΩΝ**

**Η ΚΑΤΑΝΟΜΗ WEIBULL ΩΣ ΜΟΝΤΕΛΟ**  
**ΖΗΜΙΟΚΑΤΑΝΟΜΗΣ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ**  
**ΕΠΙΣΤΗΜΗ**

**Ιωάννης Ράπτης**

*Διπλωματική Εργασία*

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

Πειραιάς

Σεπτέμβριος 2024

**University of Piraeus**  
**School of Finance and Statistics**



**Department of Statistics and Insurance Science**

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

**THE WEIBULL DISTRIBUTION AS A MODEL  
FOR LOSS DISTRIBUTIONS IN ACTUARIAL  
SCIENCE**

Ioannis Raptis

MSc Dissertation

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

Piraeus, Greece

September 2024



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

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής Πολίτης Κωνσταντίνος (Επιβλέπων)
- Αναπληρωτής Καθηγητής Μπούτσικας Μιχαήλ
- Επίκουρος Καθηγητής Τριανταφύλλου Ιωάννης

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



## Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Κωνσταντίνο Πολίτη, για την πολύτιμη συνδρομή του καθ' όλη την διάρκεια εκπόνησης της παρούσας Διπλωματικής Εργασίας.

Επιπλέον, θα ήθελα να απευθύνω ευχαριστίες στην οικογένεια και τον κύκλο των φίλων και συναδέλφων μου για την αμέριστη κατανόηση και υποστήριξη τους στο παρόν εγχείρημα.



## Περίληψη

Η κατανομή Weibull αποτέλεσε και εξακολουθεί να αποτελεί μια από τις κυριότερες κατανομές, αναφορικά με την επιστήμη της Στατιστικής και εφαρμοσμένης Στατιστικής σε παγκόσμια κλίμακα. Η πρώτη της επίσημη παρουσίαση στην παγκόσμια επιστημονική κοινότητα πραγματοποιήθηκε από τον Waloddi Weibull το 1951, παρόλα αυτά, οι απαρχές της μελέτης της φημολογούνται περί το 1920.

Συγκεκριμένα για την Αναλογιστική Επιστήμη, η κατανομή Weibull έχει καταλυτικό ρόλο στην ανάπτυξή της, αφού χρησιμοποιείται ως Ζημιοκατανομή στον κλάδο των Γενικών Ασφαλίσεων, καθώς επίσης και σε περιοχές όπως η Ανάλυση Επιβίωσης και η Θεωρία Κινδύνου, με εφαρμογές στην ένταση θνησιμότητας (force of mortality) και στους χρόνους ζωής (lifetimes) συνεχών τυχαίων μεταβλητών.

Η παρούσα εργασία θα εστιάσει στην κατανομή Weibull από την σκοπιά της Αναλογιστικής Επιστήμης, παρουσιάζοντας αναλυτικά τις βασικές ιδιότητές της, καθώς και θα γίνει μια απόπειρα να παρουσιαστούν ορισμένες εφαρμογές της μέσω αριθμητικών παραδειγμάτων, με την βοήθεια της προγραμματιστικής γλώσσας R.





## Abstract

The Weibull Distribution has always been one of the main Distributions, concerning the scientific field of Statistics and Applied Statistics, worldwide. The first official publication was performed by Waloddi Weibull in 1951, however, the origins of this distribution can be traced back to 1920.

In particular, and concerning Actuarial Science, the so-called Weibull Distribution maintains a pivotal role in its development, due to the fact that it is used as a Loss Distribution in General Insurance, as well as in the fields of Survival Analysis and Risk Theory, with applications such as in force of Mortality and lifetime data.

This thesis will focus on the Weibull distribution from the perspective of Actuarial Science, presenting its basic properties, as well as attempting to illustrate certain applications, via the use of R programming language.



## Περιεχόμενα

Ευχαριστίες .....	iii
Περίληψη.....	v
Abstract .....	vii
<b>Κεφάλαιο 1 – Η Κατανομή Weibull .....</b>	<b>5</b>
1.1 Εισαγωγή .....	5
1.2 Το μοντέλο του πιο αδύναμου κρίκου.....	6
1.3 Ο Κλασσικός Ορισμός της Κατανομής Weibull.....	8
1.4 Οι Εναλλακτικοί Ορισμοί Δύο και Τριών Παραμέτρων .....	12
1.5 Αθροιστική Συνάρτηση Κατανομής και Συνάρτηση Αξιοπιστίας Weibull..	13
1.6 Ένταση Κινδύνου και Αθροιστική Συνάρτηση Κινδύνου Weibull .....	18
1.7 Μέσος Υπολειπόμενος Χρόνος Ζωής (MRL) Weibull .....	20
1.8 Μέση Τιμή και Διακύμανση Weibull.....	24
<b>Κεφάλαιο 2 – Η Κατανομή Weibull ως Ζημιοκατανομή .....</b>	<b>26</b>
2.1 Εισαγωγή .....	26
2.2 Ροπογεννήτρια Κατανομής Weibull .....	27
2.3 Ουρά της Κατανομής Weibull .....	29
<b>Κεφάλαιο 3 – Βασικές Έννοιες στην Ανάλυση Επιβίωσης .....</b>	<b>32</b>
3.1 Εισαγωγή .....	32
3.2 Συνάρτηση Πυκνότητας Πιθανότητας Θανάτου .....	32
3.3 Συνάρτηση Κατανομής Διάρκειας Ζωής .....	33
3.4 Συνάρτηση Επιβίωσης.....	35
3.5 Ένταση Θνησιμότητας .....	37
3.6 Αθροιστική Συνάρτηση Κινδύνου .....	41
3.7 Μέσος Χρόνος ζωής και Μέση Υπολειπόμενη Ζωή.....	42
<b>Κεφάλαιο 4 – Εφαρμογές της Κατανομής Weibull .....</b>	<b>44</b>
4.1 Εισαγωγή .....	44
4.2 Εκτίμηση Παραμέτρων Κατανομής Weibull .....	44

4.3 Έλεγχος Καλής Προσαρμογής Κατανομής σε Πραγματικά Δεδομένα με βάση το Κριτήριο Πληροφορίας .....	53
4.4 Εκτίμηση Μέσου Υπολειπόμενου Χρόνου Ζωής Ενεργών Ασφαλιστηρίων Συμβολαίων Χαρτοφυλακίου.....	57
Σύνοψη - Συμπεράσματα .....	60
<b>ΠΑΡΑΡΤΗΜΑ.....</b>	<b>61</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ .....</b>	<b>72</b>

## Κατάλογος Σχημάτων

Σχήμα 1: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου θέσης $a$ ..9	
Σχήμα 2: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου κλίμακας $b$ .....	10
Σχήμα 3: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου σχήματος $c$ .....	11
Σχήμα 4: Η 3-Παραμετρική Αθροιστική Συνάρτηση Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $b$ ( $c$ και $a$ σταθερά).....	15
Σχήμα 5: Η Κανονικοποιημένη Αθροιστική Συνάρτηση Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $c$ .....	15
Σχήμα 6: Η 3-Παραμετρική Συνάρτηση Αξιοπιστίας της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $b$ ( $a$ και $c$ σταθερά) .....	17
Σχήμα 7: Η Κανονικοποιημένη Συνάρτηση Αξιοπιστίας της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $c$ .....	17
Σχήμα 8: Ένταση Κινδύνου της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $c$ .....	18
Σχήμα 9: Αθροιστική Συνάρτηση Κινδύνου της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου $c$ .....	20
Σχήμα 10: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με $0 < c < 1$ .....	22
Σχήμα 11: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με $c = 1$ .....	23
Σχήμα 12: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με $c > 1$ .....	23
Σχήμα 13: Γραφική Απεικόνιση της Μέσης Τιμής Weibull 2 Παραμέτρων για τις διάφορες τιμές της Παραμέτρου Σχήματος.....	25
Σχήμα 14: Γραφική Απεικόνιση της Διακύμανσης Weibull 2 Παραμέτρων για τις διάφορες τιμές της Παραμέτρου Σχήματος.....	25
Σχήμα 15: Προσεγγιστική Σχηματική Απεικόνιση Ροπογεννήτριας Κατανομής Weibull για διάφορες τιμές της παραμέτρου σχήματος $c > 1$ .....	29
Σχήμα 16: Διαγραμματική Απεικόνιση κατανομών με διαφορετικό τύπο ουράς, συγκριτικά με την ουρά της Εκθετικής Κατανομής .....	30
Σχήμα 17: Γραφική απεικόνιση ουράς Weibull, για τις διάφορες τιμές της παραμέτρου σχήματος $c$ .....	31
Σχήμα 18: Γραφική Αναπαράσταση της Συνάρτησης Κατανομής Διάρκειας Ζωής Weibull .....	35
Σχήμα 19: Θεωρητική Γραφική Αναπαράσταση Συνάρτησης Επιβίωσης.....	36
Σχήμα 20: Εμπειρική Γραφική Αναπαράσταση Συνάρτησης Επιβίωσης .....	37

<b>Σχήμα 21: Γραφική Αναπαράσταση Διαφορετικών Περιπτώσεων Έντασης Θνησιμότητας .....</b>	<b>39</b>
<b>Σχήμα 22: Γραφική Αναπαράσταση της Λεκανοειδούς Καμπύλης.....</b>	<b>40</b>
<b>Σχήμα 23: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας .....</b>	<b>49</b>
<b>Σχήμα 24: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου των Ροπών .....</b>	<b>50</b>
<b>Σχήμα 25: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Ελάχιστης Απόστασης .....</b>	<b>51</b>
<b>Σχήμα 26: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Ποσοστημορίων .....</b>	<b>52</b>
<b>Σχήμα 27: Γραφικές Αναπαραστάσεις Συνάρτησης Κατανομής των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας .....</b>	<b>56</b>
<b>Σχήμα 28: Q-Q Plots των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας .....</b>	<b>56</b>
<b>Σχήμα 29: P-P Plots των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας .....</b>	<b>57</b>
<b>Σχήμα 30: Γραφική Αναπαράσταση Μέσου Υπολειπόμενου Χρόνου Ζωής Χαρτοφυλακίου μέσω της κατανομής Weibull.....</b>	<b>58</b>





# Κεφάλαιο 1 – Η Κατανομή Weibull

## 1.1 Εισαγωγή

Σε αρκετές περιπτώσεις κατά την δημιουργία και ανάπτυξη μιας επιστημονικής μεθόδου, ενός θεωρήματος ή μιας διαδικασίας, η μέθοδος αυτή δεν έλαβε το όνομα του αρχικού δημιουργού της.

Ένα τέτοιο παράδειγμα αποτελεί και η γνωστή σε όλους μας **Κανονική Κατανομή** (Normal Distribution), επίσης γνωστή και ως **Κατανομή Gauss**, η οποία εμφανίζει τις απαρχές της περί τα μέσα του 18<sup>ου</sup> αιώνα από τον Abraham De Moivre. Στη συνέχεια, πραγματοποιείται μια δεύτερη αναφορά της κατανομής από τον Pierre Simon De Laplace στα τέλη του ίδιου αιώνα, περί το 1774, ως μια προσέγγιση για την Υπεργεωμετρική Κατανομή (Hypergeometric Distribution). Εν τέλει, ο Carl Friedrich Gauss θα ήταν αυτός που με δύο δημοσιεύσεις του, η μια το 1809, με τίτλο “Theoria Motus Corporum Coelestium” (μεταφρ. «Θεωρία Κινητικότητας των Ουράνιων Σωμάτων»), και η άλλη με ημερομηνία αναφοράς το 1816, με τίτλο “Bestimmung der Genauigkeit von Beobachtungen” (μεταφρ. «Προσδιορίζοντας την ακρίβεια των παρατηρήσεων»), περιέγραψε την Κανονική κατανομή ως «Μια μορφή ενός νόμου σφαλμάτων κατά τον υπολογισμό αστρονομικών αποστάσεων», συνδέοντας έτσι το όνομά του με μια από τις πλέον πιο γνωστές κατανομές παγκοσμίως.

Παρόμοια πορεία ακολούθησε και το ιστορικό πλαίσιο ανακάλυψης της γνωστής σε εμάς ως **Κατανομή Weibull**, με τις πρώτες επιστημονικές απόπειρες απεικόνισης της συγκεκριμένης κατανομής από τους Frechet (1927) και τους Rosin/Rammler/Sperling (1933). Η μελέτη που πραγματοποιήθηκε το 1939 από τον Σουηδό **Waloddi Weibull** στηρίχτηκε κυρίως σε 2 άρθρα και αφορούσαν αντοχές υλικών – στο πρώτο άρθρο με τίτλο “A Statistical Theory of the Strength of Material” (μεταφρ. «Μια Στατιστική Θεωρία για την Ισχύ των Υλικών»), παρουσιάζεται η Συνάρτηση Κατανομής της Weibull, βασιζόμενη σε εμπειρικές δοκιμές παρατηρήσεων σχετικά με την Κύρτωση, την Ελαστικότητα και την Στρεπτικότητα ράβδων κατασκευασμένων από γύψο και ψευδάργυρο. Στο δεύτερο άρθρο του με τίτλο “The Phenomenon of Rupture in Solids” (μεταφρ. «το Φαινόμενο της Ρήξης στα Στερεά») παρουσιάζονται οι Ροπές της Συνάρτησης Κατανομής, παράλληλα με την χρήση διαγραμματικών απεικονίσεων και πινάκων για την εκτίμηση των παραμέτρων.

Στα χρόνια που ακολούθησαν και ύστερα από τη λήξη του Δευτέρου Παγκοσμίου Πολέμου, από το 1949 και έπειτα, ο Weibull παρουσίασε την Κατανομή του στην διεθνή επιστημονική κοινότητα, με το σημαντικότερο δημοσίευμα να χρονολογείται περί το 1951.

Μέχρι και σήμερα, η κατανομή Weibull αποτελεί μια από τις σπουδαιότερες κατανομές μελέτης σε ένα ευρύ πεδίο εφαρμογών, από τις οποίες ξεχωρίζουν η Θεωρία Αξιοπιστίας Συστημάτων, καθώς και η Θεωρία Κινδύνου.

Στο συγκεκριμένο Κεφάλαιο και ως επί το πλείστον, αντλήθηκαν πληροφορίες από το Κεφάλαιο 2.2 του Συγγράμματος του Horst Rinne “The Weibull Distribution – A Handbook (2008)”

## 1.2 Το μοντέλο του πιο αδύναμου κρίκου

Πολλές φορές ταυτίζουμε μια στατιστική κατανομή με ένα αντίστοιχο φυσικό μοντέλο. Έτσι ενισχύουμε τόσο την επεξήγηση της κατανομής, όσο και την εμπειρική επιλογή μιας κατανομής, σε περίπτωση που έρθουμε αντιμέτωποι με ένα σύνολο δεδομένων, για τα οποία τυχόν κληθούμε να πραγματοποιήσουμε προσαρμογή με μια κατάλληλη κατανομή.

Στην περίπτωση της κατανομής Weibull, το κλασσικό φυσικό μοντέλο που την εκφράζει είναι αυτό που είναι άρρηκτα συνδεδεμένο με την προέλευσή της, ως κατανομή Ακραίων Τιμών.

Πιο συγκεκριμένα, το μοντέλο αυτό απαρτίζεται από  $n$  όμοιες μονάδες ή συστήματα, τα οποία είναι ενωμένα μεταξύ τους σε σειρά. Αυτό σημαίνει ότι, για να οδηγηθεί το σύστημα  $n$  σε αποτυχία, τότε θα πρέπει τουλάχιστον μια μονάδα να αποτύχει. Ισοδύναμα, το σύστημα  $n$  συνεχίζει να βρίσκεται σε λειτουργία, όταν όλες ανεξαιρέτως οι μονάδες που το απαρτίζουν βρίσκονται σε λειτουργία.

Το παραπάνω σύστημα σειριακής μορφής που παρουσιάστηκε, μπορεί να παρομοιαστεί στον φυσικό κόσμο με μια αλυσίδα, η οποία είναι τόσο ισχυρή όσο και ο πιο αδύναμος δεσμός της. Εναλλακτικά, μπορούμε να αναφέρουμε ότι μια αλυσίδα καταστρέφεται με το που καταστραφεί ο δεσμός μικρότερης διάρκειας ζωής που την απαρτίζει.

Έστω  $X_i (i = 1, \dots, n)$  οι τυχαίοι χρόνοι ζωής των προαναφερθέντων σειριακά συνδεδεμένων μονάδων και έστω ότι οι  $X_i$  είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, με συνάρτηση κατανομής  $F(x)$ . Τότε η  $Y_n$  υποδηλώνει τον συνολικό χρόνο ζωής του συστήματος με:

$$Y_n = \min_{1 \leq i \leq n} \{X_i\} \quad (1.2.1)$$

Η αθροιστική συνάρτηση κατανομής της  $Y_n$  σύμφωνα με τα όσα αναφέραμε μέχρι στιγμής προκύπτει ότι είναι:

$$F_{Y_n}(y) := \Pr(Y_n \leq y) = 1 - [1 - F(y)]^n \quad (1.2.2)$$

Για να μπορέσουμε να αποφύγουμε τον υπολογισμό δυνάμεων του  $[1 - F(y)]^n$  προχωράμε σε μια τεχνική μετασχηματισμού που είχε προτείνει ο Cramer και ουσιαστικά ορίζουμε μια νέα τυχαία μεταβλητή  $U_n$  με:

$$U_n := nF(Y_n) \quad (1.2.3)$$

Για κάθε  $u$  μέσα στο διάστημα  $[0, n]$  έχουμε ότι:

$$\Pr(U_n \leq u) = \Pr[nF(Y_n) \leq u] = \Pr[Y_n \leq F^{-1}(u/n)] \quad (1.2.4)$$

Και επομένως, αντικαθιστώντας στην σχέση (2.2.2) παραπάνω, έχουμε ότι:

$$G_n(u) = 1 - (1 - u/n)^n \quad (1.2.5)$$

Όσο το  $n \rightarrow \infty$ , η ακολουθία των τυχαίων μεταβλητών  $U_n$  συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή  $U$ , καθώς η ακολουθία των συναρτήσεων κατανομής  $G_n(u)$  συγκλίνει για κάθε  $u$ , στην συνάρτηση κατανομής:

$$G(u) := \lim_{n \rightarrow \infty} G_n(u) = 1 - e^{-u}, u \geq 0 \quad (1.2.6)$$

Η οποία έχει την αντίστοιχη συνάρτηση πυκνότητας πιθανότητας:

$$g(u) := \lim_{n \rightarrow \infty} g_n(u) = e^{-u}, u \geq 0 \quad (1.2.7)$$

Επομένως, γενικεύοντας τα συμπεράσματα που προκύπτουν από την σχέση (1.2.6), μπορούμε να πούμε ότι μια ακολουθία τυχαίων μεταβλητών  $Y_n$  συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή  $Y$ , με:

$$Y = F^{-1}(U/n) \quad (1.2.8)$$

Άρα, υπολογίζοντας την  $F^{-1}$ , μπορούμε να αποφανθούμε για την κατανομή της  $U$ .

Εφαρμόζοντας στην θέση της  $F(x)$  κατανομές όπως η **Ομοιόμορφη** στο διάστημα  $[a,b]$  και η **Εκθετική** με παράμετρο  $\lambda$ , μπορούμε να καταλήξουμε σε σύγκλιση της  $Y_n$  σε κατανομή Weibull.

Αυτό θα γίνει πιο σαφές στη συνέχεια, καθώς θα παρουσιαστεί και ο κλασσικός ορισμός της κατανομής Weibull.

### 1.3 Ο Κλασσικός Ορισμός της Κατανομής Weibull

Σύμφωνα με τον Rinne (2008), ο Weibull στην επίσημη παρουσίαση της κατανομής που κατέληξε να φέρει τιμητικά το όνομά του, βασίστηκε σε έναν τριπαραμετρικό ορισμό. Στην πορεία του χρόνου η ίδια κατανομή εκφράστηκε και με έναν εναλλακτικό τρόπο, μέσω ενός ορισμού ο οποίος εμπλέκει δύο παραμέτρους.

Για σκοπούς ιστορικής πιστότητας, θα ξεκινήσουμε εκφράζοντας την κατανομή Weibull μέσω του κλασσικού ορισμού της, αυτού των τριών παραμέτρων, και σε επόμενες παραγράφους θα δοθούν οι επιπρόσθετοι ορισμοί, παρουσιάζοντας παράλληλα την χρηστικότητα του καθενός στα επιμέρους επιστημονικά πεδία.

Έστω μια συνεχής τ.μ.  $X$ . Αυτή θα λέμε ότι ακολουθεί την κατανομή Weibull με παραμέτρους  $a$ ,  $b$  και  $c$ , όταν έχει την εξής συνάρτηση πυκνότητας πιθανότητας:

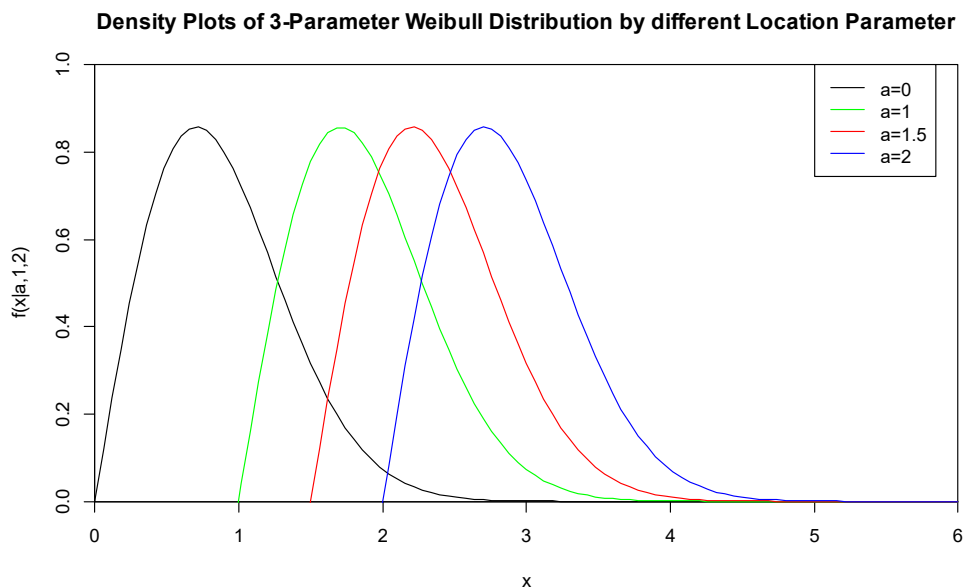
$$f_X(x|a, b, c) = \frac{c}{b} \left(\frac{x-a}{b}\right)^{c-1} \exp\left\{-\left(\frac{x-a}{b}\right)^c\right\}, x \geq a \quad (1.3.1)$$

Συμβολικά, γράφουμε ότι  $X \sim We(a,b,c)$

Αναλύοντας τις παραμέτρους, έχουμε ότι η  $a$  αποτελεί μια πραγματική μεταβλητή, η οποία συνήθως εκφράζει μια μονάδα του χρόνου, όπως λεπτά, ημέρες ή έτη. Συγκεκριμένα για την Ανάλυση Επιβίωσης, η μεταβλητή  $a$  καλείται και ως **ελάχιστη ή εγγυημένη** διάρκεια ζωής ενός ατόμου ηλικίας  $x$ .

Επομένως, παρατηρώντας ότι το  $a$  αποτελεί την ελάχιστη διάρκεια ζωής ενός ατόμου το οποίο έχει ήδη διανύσει ένα διάστημα ζωής  $x$  τουλάχιστον ίσο ή μεγαλύτερο του  $a$ , μπορούμε να περιορίσουμε το πεδίο ορισμού της  $a$  από ολόκληρο το  $\mathbb{R}$ , στο διάστημα  $[0, +\infty)$

Από πλευράς Στατιστικής, η παράμετρος  $a$  καλείται **Παράμετρος Θέσης (Location Parameter)**. Αυτό σημαίνει ότι, μια αλλαγή τιμής της παραμέτρου  $a$ , κρατώντας παράλληλα τις υπόλοιπες δύο μεταβλητές σταθερές, θα προκαλέσει μια παράλληλη μετατόπιση της γραφικής παράστασης της συνάρτησης πυκνότητας πιθανότητας κατά μήκος του άξονα των  $x$ . Για να μπορέσει να γίνει αυτό καλύτερα αντιληπτό, παρουσιάζουμε παρακάτω 4 διαφορετικές σ.π.π της κατανομής Weibull, οι οποίες διαφέρουν μεταξύ τους μόνο ως προς την παράμετρο θέσης  $a$ .



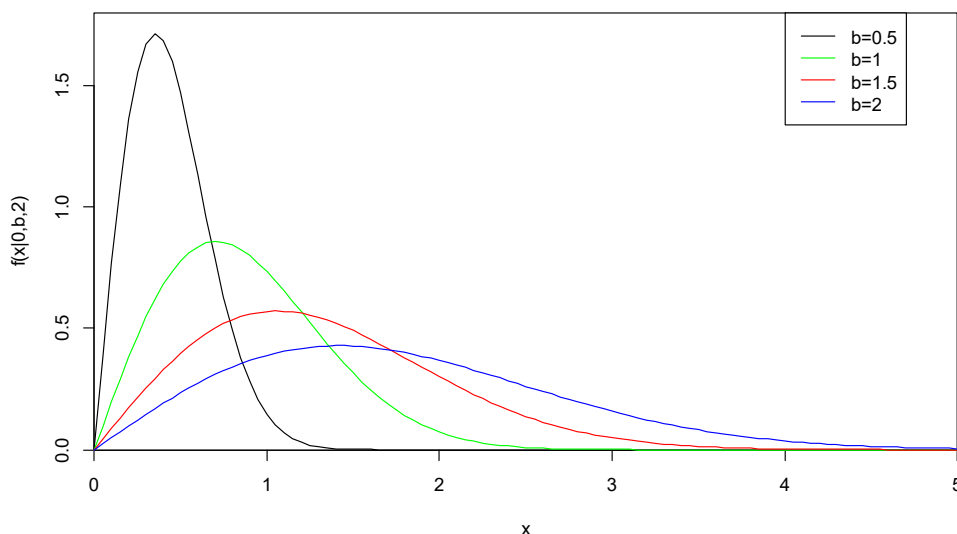
*Σχήμα 1: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου θέσης  $a$*

Προχωρώντας στην δεύτερη κατά σειρά παράμετρο  $b$ , και εκείνη με την σειρά της ορίζεται στο διάστημα  $[0, +\infty)$  και παριστάνει την ίδια μονάδα μέτρησης με την τ.μ  $x$ . Αν δηλαδή η  $x$  παριστάνει έτη, τότε και η  $b$  με την σειρά της εκφράζεται σε έτη κ.ο.κ.

Σύμφωνα με τον Rinne (2008), όταν η  $x$  εκφράζει την διάρκεια ζωής ενός ατόμου, τότε η  $b$  καλείται **Χαρακτηριστική Ζωή (Characteristic Life)**. Από την σκοπιά της Στατιστικής, η παράμετρος  $b$  είναι γνωστή και ως **Παράμετρος Κλίμακας (Scale Parameter)**, δηλαδή, μεταβάλλοντας την παράμετρο  $b$ , ενώ οι υπόλοιπες δύο παράμετροι της κατανομής παραμένουν σταθερές, τότε παρατηρείται μια αλλαγή στην κυρτότητα της καμπύλης πυκνότητας πιθανότητας. Συγκεκριμένα, μεγαλύτερες τιμές του  $b$  συνεπάγονται μείωση της κυρτότητας, ενώ μικρότερες τιμές προκαλούν το αντίστροφο αποτέλεσμα.

Μια γραφική αναπαράσταση των διαφορετικών τιμών της παραμέτρου  $b$  μπορεί να φανεί και στο Σχήμα 2.

Density Plots of 3-Parameter Weibull Distribution by different Scale Parameter



Σχήμα 2: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου κλίμακας  $b$

Από το Σχήμα 2 παρατηρείται επίσης ότι, λόγω της μεταβολής που προκαλείται για τις διάφορες τιμές της παραμέτρου  $b$ , αυτή επηρεάζει και την διασπορά της τ.μ.  $X$ . Μικρότερες τιμές του  $b$  οδηγούν σε μείωση της διασποράς της  $X$  και αντίστροφα.

Αξίζει να σημειωθεί ότι οι παράμετροι  $a$  και  $b$  καθορίζουν τον γραμμικό μετασχηματισμό του χρόνου ζωής  $x$  σε έναν αριθμό  $u$ , ως εξής:

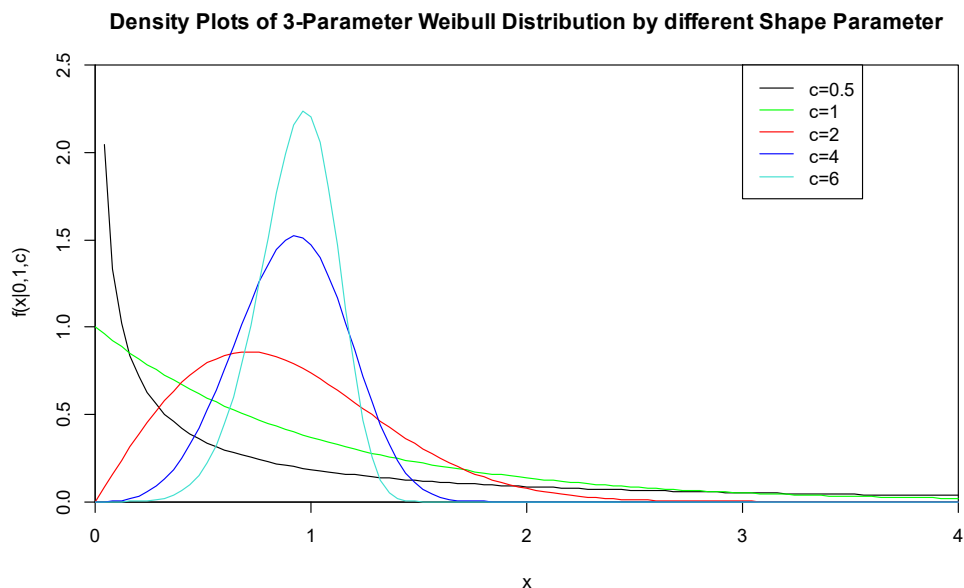
$$x \xrightarrow{a,b} \frac{x - a}{b} =: u \quad (1.3.2)$$

Συνεχίζοντας στην ερμηνεία της τρίτης παραμέτρου  $c$ , ως πεδίο ορισμού της λαμβάνει το διάστημα  $(0, +\infty)$ . Πολλές φορές αναφέρεται και ως **κλίση** της Weibull, καθώς αυτή προσδιορίζει την κλίση της αθροιστικής συνάρτησης κατανομής, ή αναφερόμενοι σε αναλογιστικές έννοιες, της αθροιστικής συνάρτησης κινδύνου.

Ερμηνεύοντας την παράμετρο  $c$  στατιστικά, εκείνη καλείται και ως **παράμετρος σχήματος (shape parameter)**. Ουσιαστικά, μεταβάλλοντας την  $c$  κρατώντας τις άλλες δύο παραμέτρους σταθερές, αυτό προκαλεί αλλαγή του σχήματος της γραφικής παράστασης της συνάρτησης πυκνότητας πιθανότητας. Πιο συγκεκριμένα, όταν  $c < 1$ , τότε επικρατεί η **εκθετική** συμπεριφορά στην σ.π., ενώ για  $c > 1$ , τότε έχουμε επικράτηση του

πολυωνυμικού χαρακτήρα της κατανομής (Rinne (2008)), κάτι το οποίο αποτυπώνεται αντίστοιχα και στο γράφημα της σ.π.π.

Μια ενδεικτική γραφική απεικόνιση των παραπάνω μπορεί να γίνει αντιληπτή και μέσω του Σχήματος 3 για τις διάφορες τιμές του  $c$ .



Σχήμα 3: Η Κατανομή Weibull για τις διάφορες τιμές της παραμέτρου σχήματος  $c$

Η παράμετρος  $c$  παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς από εκείνη εξαρτάται η μορφή της έντασης κινδύνου ή έντασης θνησιμότητας (βλ. Παράγραφο 1.6). Συγκεκριμένα, για τα διάφορα διαστήματα του χρόνου ζωής ενός πληθυσμού ατόμων, έχουμε τις παρακάτω επιρροές:

- Για  $c < 1$ , τότε η γραφική αναπαράσταση της σ.π.π. υποδεικνύει ότι η ένταση κινδύνου  $h(x)$  έχει φθίνουσα συμπεριφορά. Η συγκεκριμένη μελέτη θνησιμότητας είναι γνωστή και ως **Θνησιμότητα Νεογέννητων Ατόμων (Infant Mortality Rate)**. Με τον όρο αυτό, αναφερόμαστε στην υψηλότερη πιθανότητα θανάτου κατά την έναρξη του χρόνου ζωής του ατόμου ή της μονάδας
- Για  $c = 1$ , τότε αναφέρουμε ότι παρατηρείται μια **σταθερή** ένταση κινδύνου στον πληθυσμό, **ανεξάρτητη** του χρόνου
- Για  $c > 1$ , τότε αυτό ερμηνεύεται ως μια **αύξουσα** ένταση κινδύνου, το οποίο υποδεικνύει συνήθως έναν πληθυσμό αυξημένου γήρατος

Συνοψίζοντας όλα τα παραπάνω, το στοιχείο το οποίο διαχωρίζει την Κατανομή Weibull εκφρασμένη μέσω του κλασσικού ορισμού των τριών παραμέτρων, είναι η ύπαρξη της παραμέτρου θέσης. Επομένως, η χρήση της τριπαραμετρικής κατανομής Weibull ενδείκνυται σε περιπτώσεις κατά τις οποίες παρουσιάζονται διαστήματα μηδενικής θνησιμότητας σε έναν πληθυσμό.

Έστω ότι εμφανίζεται διάστημα μηδενικής θνησιμότητας  $t_0$ . Σε αυτή την κατάσταση, κρίνεται σκόπιμο να μετατοπιστεί η κλίμακα ηλικίας κατά  $t_0$ , έτσι ώστε να μπορέσει να εκφραστεί η συγκεκριμένη εγγυημένη διάρκεια ζωής του πληθυσμού μας. Αυτό επιτυγχάνεται, προσθέτοντας στην ηλικία κάθε ατόμου την τιμή  $t_0$ . Στην ουσία, αυτή η χρονική «διόρθωση» που πραγματοποιείται, θα ισούται με τον ελάχιστο χρόνο που απαιτείται για την εμφάνιση του πρώτου θανάτου στον πληθυσμό.

#### 1.4 Οι Εναλλακτικοί Ορισμοί Δύο και Τριών Παραμέτρων

Ορισμένοι εναλλακτικοί ορισμοί της σ.π.π της κατανομής Weibull περιλαμβάνουν την απαλοιφή ενός ή δύο παραμέτρων, προσδίδοντάς τους σταθερές τιμές με  $\mathbf{a=0}$ , είτε θέτοντας εναλλάξ  $\mathbf{b=1}$  ή  $\mathbf{c=1}$ . Αυτό έχει ως αποτέλεσμα την διατύπωση τριών διαφορετικών ορισμών για την σ.π.π:

$$f_X(x|0, b, c) = \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left\{-\left(\frac{x}{b}\right)^c\right\} \quad (1.4.1)$$

$$f_X(x|a, 1, c) = c(x - a)^{c-1} \exp\{-(x - a)^c\} \quad (1.4.2)$$

$$f_X(x|a, b, 1) = \frac{1}{b} \exp\left\{-\left(\frac{x - a}{b}\right)\right\} \quad (1.4.3)$$

Αξίζει επίσης να σημειωθεί ότι για την γραφική απεικόνιση των σ.π.π (1.4.1) έως (1.4.3), μπορεί κανείς να ανατρέξει στις αντίστοιχες γραφικές παραστάσεις των Σχημάτων 1 έως 3.

Η διπαραμετρική σχέση (1.4.1) αποτελεί την πιο ευρέως γνωστή μορφή της σ.π.π δύο παραμέτρων, και είναι γνωστή και ως **μορφή κλίμακας-σχήματος (scale-shape version)**

Προχωρώντας στις μονοπαραμετρικές αναπαραστάσεις της κατανομής Weibull, σημειώνουμε τις εξής σχέσεις για την σ.π.π:

$$f_X(x|0,1, c) = cx^{c-1} \exp\{-x^c\} \quad (1.4.4)$$



$$f_X(x|0, b, 1) = \frac{1}{b} \exp\left\{-\frac{x}{b}\right\} \quad (1.4.5)$$

$$f_X(x|a, 1, 1) = \exp\{-(x - a)\} \quad (1.4.6)$$

Παρατηρούμε ότι η σχέση (1.4.5) ουσιαστικά είναι η σ.π.π της **Εκθετικής Κατανομής**, ενώ η σχέση (1.4.6) αναπαριστά την **Μετατοπισμένη Εκθετική Κατανομή κατά  $a$  μονάδες**

Σε παρόμοιο τόνο με τις διπαραμετρικές αναπαραστάσεις, σημειώνουμε ότι η σχέση (1.4.4) αποτελεί την πιο ευρέως γνωστή μορφή της σ.π.π. της κατανομής Weibull μιας παραμέτρου, η οποία συχνά εμφανίζεται και με την ονομασία **ελαττωμένη ή κανονικοποιημένη σ.π.π Weibull (reduced or normalized Weibull density)**.

Γενικά, σε μια οικογένεια κατανομών η οποία περιέχει παραμέτρους θέσης και κλίμακας, η ελαττωμένη ή κανονικοποιημένη μορφή επιτυγχάνεται θέτοντας την παράμετρο θέσης ίση με το 0, καθώς επίσης και δίνοντας στην παράμετρο κλίμακας τιμή ίση με τη μονάδα.

Η ιδιότητα που κατέχει η ελαττωμένη ή κανονικοποιημένη μορφή μιας οικογένειας κατανομών, είναι ότι κάθε μέλος της οικογένειας αυτής μπορεί να εξαχθεί μέσω ενός **γραμμικού μετασχηματισμού** της κανονικοποιημένης μορφής. Έστω ότι η τ.μ.  $U$ , η οποία καλείται **ελαττωμένη μεταβλητή (reduced variable)**, με σ.π.π  $f_U(u)$ , τότε η τ.μ.  $X$ , με

$$X = a + bU$$

Θα έχει σ.π.π

$$f_X(x) = \begin{cases} \frac{1}{b} f_U\left(\frac{x-a}{b}\right), & \text{για } b > 0 \\ -\frac{1}{b} f_U\left(\frac{x-a}{b}\right), & \text{για } b < 0 \end{cases} \quad (1.4.7)$$

## 1.5 Αθροιστική Συνάρτηση Κατανομής και Συνάρτηση Αξιοπιστίας Weibull

Προχωρώντας περαιτέρω στην μελέτη της κατανομής Weibull, θα αναφερθούμε στην Αθροιστική Συνάρτηση Κατανομής και Συνάρτηση Αξιοπιστίας, βασιζόμενοι στην κανονικοποιημένη μορφή της κατανομής.

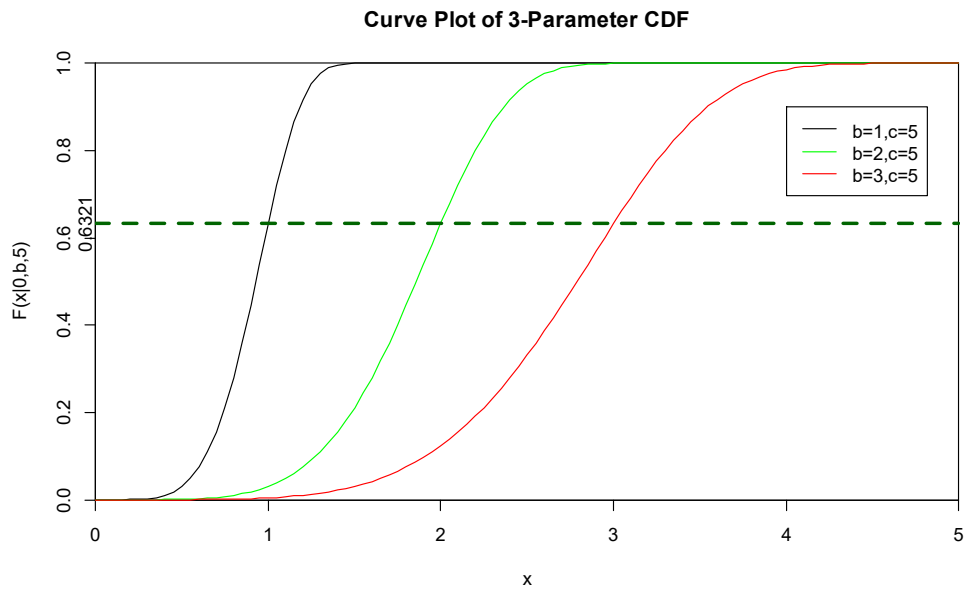
Συγκεκριμένα, ολοκληρώνοντας κατά μέλη την σχέση (1.4.4) έχουμε ότι:

$$F_X(x|c) = \Pr(X \leq x) = \int_0^x cu^{c-1} \exp(-u^c) du = 1 - \exp(-x^c) \quad (1.5.1)$$

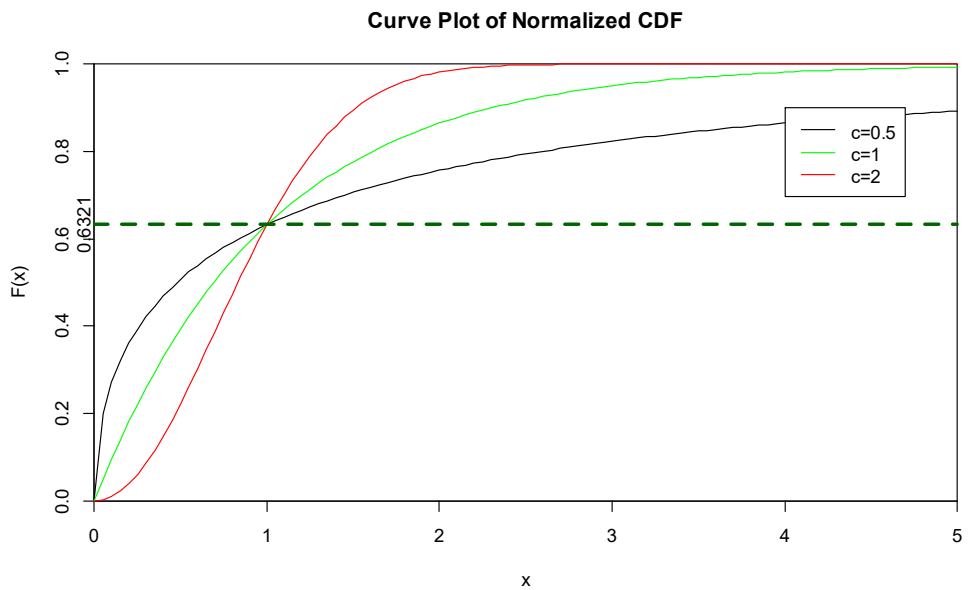
Ένεκα αναφοράς και στην γενική περίπτωση του 3-παραμετρικού ορισμού, και εργαζόμενοι με παρόμοιο τρόπο όπως και στην περίπτωση της κανονικοποιημένης μορφής, ολοκληρώνοντας κατά μέλη την 3-παραμετρική σ.π.π, ο τύπος της Αθροιστικής Συνάρτησης Κατανομής Weibull θα είναι ο εξής:

$$\begin{aligned} F_X(x|a, b, c) = \Pr(X \leq x) &= \int_0^x \frac{c}{b} \left(\frac{u-a}{b}\right)^{c-1} \exp\left\{-\left(\frac{u-a}{b}\right)^c\right\} du = \\ &\text{σ.π.π Εκθετικής Κατανομής με } \lambda=-1 \text{ και } x=\left(\frac{u-a}{b}\right)^c \\ &\overbrace{\int_0^x \left(-\exp\left\{-\left(\frac{u-a}{b}\right)^c\right\}\right)' du} = \\ &= 1 - \exp\left\{-\left(\frac{x-a}{b}\right)^c\right\} \end{aligned} \quad (1.5.2)$$

Στα παρακάτω σχήματα μπορούν να γίνουν εμφανείς οι γραφικές αναπαραστάσεις, τόσο της κανονικοποιημένης όσο και της 3-παραμετρικής Αθροιστικής Συνάρτησης Κατανομής Weibull, για διάφορες τιμές των παραμέτρων σε κάθε περίπτωση.



Σχήμα 4: Η 3-Παραμετρική Αθροιστική Συνάρτηση Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου  $b$  ( $c$  και  $a$  σταθερά)



Σχήμα 5: Η Κανονικοποιημένη Αθροιστική Συνάρτηση Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου  $c$

Παρατηρώντας τα 2 παραπάνω σχήματα, και ιδιαίτεως την γραφική απεικόνιση της Κανονικοποιημένης Weibull (Σχήμα 5), μπορούμε να εξάγουμε τα εξής δύο συμπεράσματα (Rinne (2008)):

1. Όλες οι διαφορετικές καμπύλες της κανονικοποιημένης Αθροιστικής Συνάρτησης Κατανομής (Κατανομής Ζωής), τέμνονται σε ένα κοινό σημείο με συντεταγμένες  $x = 1$  και  $F_X(x|c) = 1 - \exp\{-1\} \approx 0.6321$ . Όμοια για την 3-παραμετρική περίπτωση, προκύπτει ότι:  $F_X(a + b|a, b, c) = 1 - \exp\left\{-\left(\frac{a+b-a}{b}\right)^c\right\} = 1 - \exp\{-1\} \approx 0.6321$ .  
Επομένως, συμπεραίνουμε ότι στην Συνάρτηση Κατανομής Ζωής της Weibull, **υπάρχει πιθανότητα 63.21% περίπου, ένας χρόνος ζωής (lifetime) να είναι το πολύ b χρονικές μονάδες μεγαλύτερος από τον ελάχιστο χρόνο ζωής a**. Από την συγκεκριμένη παρατήρηση, μπορεί να γίνει αντιληπτός και ο λόγος για τον οποίο η παράμετρος b είναι επίσης γνωστή και ως «χαρακτηριστική ζωή», καθώς είναι χαρακτηριστικό όλων των Συναρτήσεων Κατανομής Ζωής της Weibull, να έχουν ένα κοινό (χαρακτηριστικό) σημείο, το σημείο **(a+b, 0.6321)**.
2. Οι συναρτήσεις κατανομής έχουν διαφορετική κυρτότητα η οποία προσδιορίζεται από την τιμή της παραμέτρου c. Συγκεκριμένα:
  - a. Αν  $0 < c \leq 1$ , τότε η F δεν παρουσιάζει κάποιο σημείο καμψής και είναι **κοίλη**
  - b. Αν  $c > 1$ , τότε η F παρουσιάζει σημείο καμψής, έστω  $x^*$ , και αυτή είναι **κυρτή** μέχρι και το σημείο όπου η πυκνότητα  $x \leq x^*$ , ενώ δεξιά του  $x^*$ , αυτή είναι **κοίλη**.

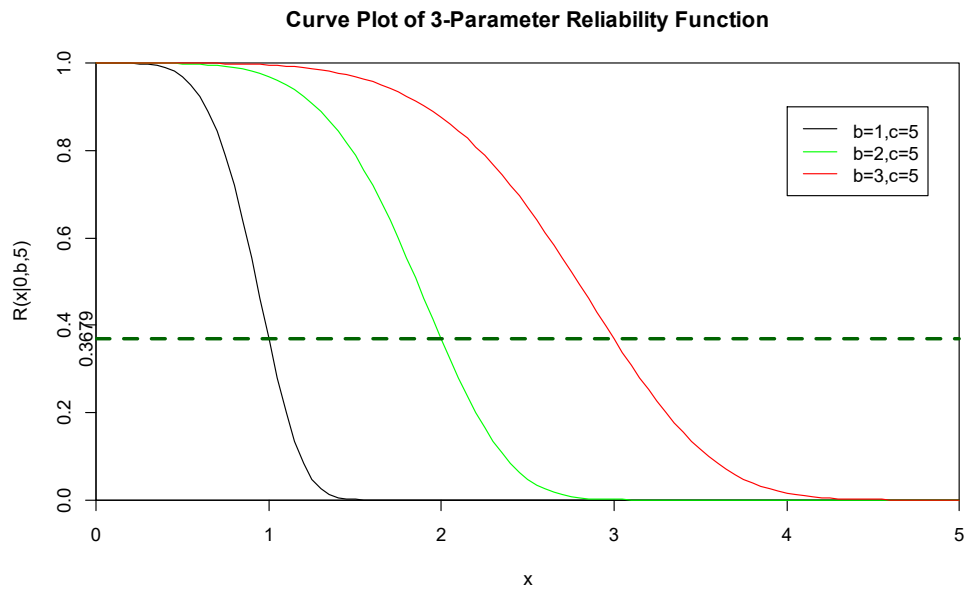
Η **Συνάρτηση Αξιοπιστίας R(x)** αποτελεί την συμπληρωματική συνάρτηση της F, με:

$$R(x|c) = 1 - F_X(x|c) = \exp\{-x^c\} \quad (1.5.3)$$

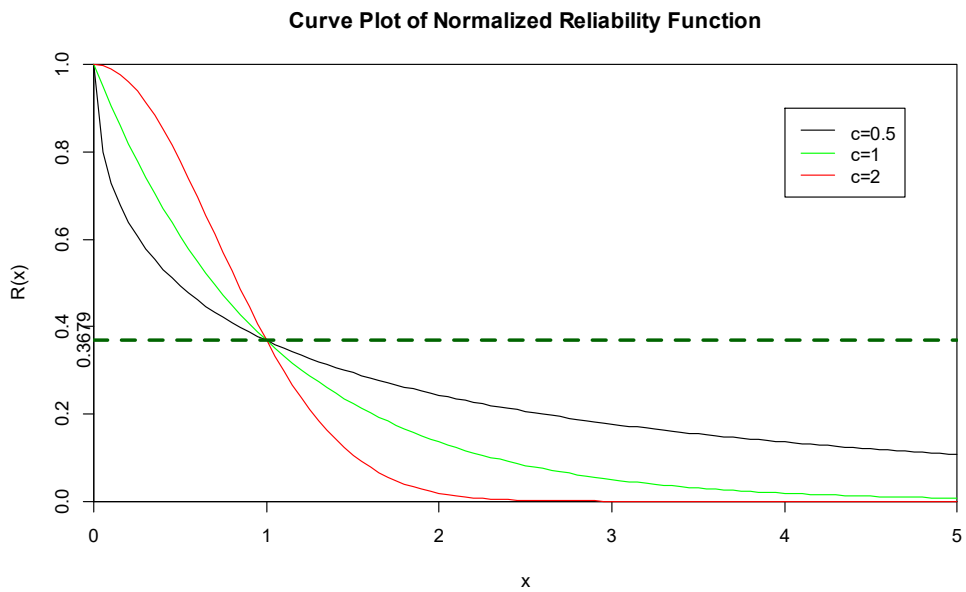
Σε αντιστοιχία με όσα ειπώθηκαν για την Συνάρτηση Κατανομής F(x), για την Συνάρτηση Αξιοπιστίας R(x) ισχύουν τα εξής:

1. Το χαρακτηριστικό σημείο έχει συντεταγμένες **(a+b, 0.3679)**
2. Η κυρτότητα της R(x) παρουσιάζει ακριβώς **αντίθετη** συμπεριφορά με αυτή της F(x)

Κατ' αντιστοιχία με την F(x) παρακάτω παρουσιάζονται οι γραφικές απεικονίσεις της R(x), τόσο μέσω της κανονικοποιημένης έκφρασης, όσο και μέσω του κλασσικού 3-παραμετρικού ορισμού.



Σχήμα 6: Η 3-Παραμετρική Συνάρτηση Αξιοπιστίας της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου  $b$  (α και  $c$  σταθερά)



Σχήμα 7: Η Κανονικοποιημένη Συνάρτηση Αξιοπιστίας της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου  $c$

## 1.6 Ένταση Κινδύνου και Αθροιστική Συνάρτηση Κινδύνου Weibull

Αναφορικά με την Ένταση Κινδύνου (ή Βαθμίδα Αποτυχίας) (Hazard Rate) ή Ένταση Θνησιμότητας (Force of Mortality) της Κατανομής Weibull, τότε για την κανονικοποιημένη μορφή, έχουμε την εξής σχέση:

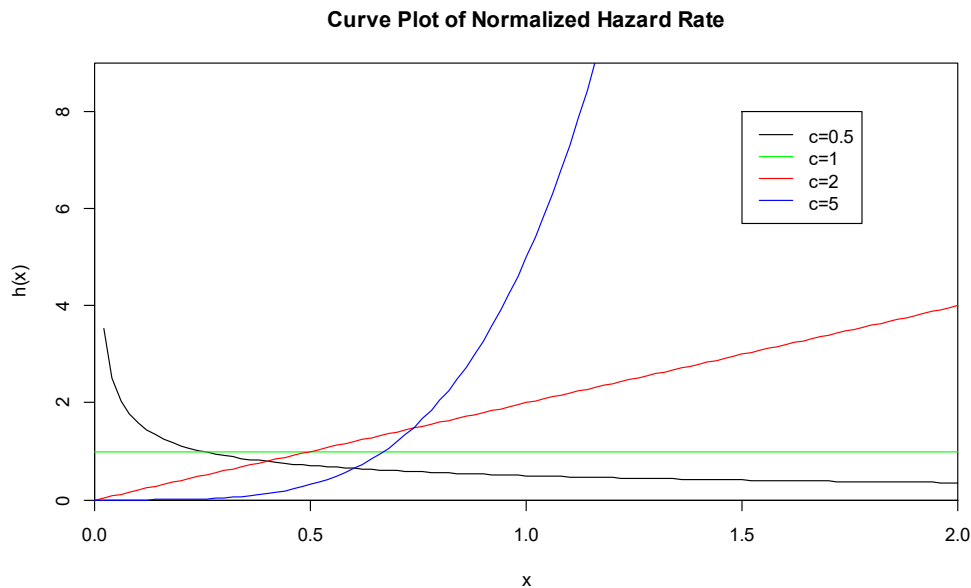
$$h_U(u|c) = \frac{f_U(u|c)}{R_U(u|c)} = \frac{cu^{c-1} \exp\{-u^c\}}{\exp\{-u^c\}} = cu^{c-1} \quad (1.6.1)$$

Ενώ στην περίπτωση του κλασσικού ορισμού τριών παραμέτρων θα έχουμε ότι:

$$h_X(x|a, b, c) = \frac{f_X(x|a, b, c)}{R_X(x|a, b, c)} = \frac{\frac{c}{b} \left(\frac{x-a}{b}\right)^{c-1} \exp\left\{-\left(\frac{x-a}{b}\right)^c\right\}}{\exp\left\{-\left(\frac{x-a}{b}\right)^c\right\}} = \frac{c}{b} \left(\frac{x-a}{b}\right)^{c-1} \quad (1.6.2)$$

Γενικά, καθοριστικό παράγοντα στην Ένταση Κινδύνου της κατανομής Weibull αποτελεί **η παράμετρος σχήματος c**, οπότε αρκεί να πραγματοποιηθεί μελέτη της συγκεκριμένης για να αποτυπωθεί η συμπεριφορά της Έντασης Κινδύνου.

Για τις διάφορες τιμές του c στην σχέση (1.6.1) έχουμε την παρακάτω σχηματική απεικόνιση της Έντασης Κινδύνου:



Σχήμα 8: Ένταση Κινδύνου της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου c

Από την παρατήρηση της γραφικής απεικόνισης, παράλληλα με την μελέτη των παραγώγων πρώτης και δεύτερης τάξης της Έντασης Κινδύνου, εξάγουμε τα ακόλουθα συμπεράσματα αναφορικά με την μελέτη γήρανσης ενός πληθυσμού, όταν ο χρόνος ζωής κάθε ατόμου ακολουθεί την κατανομή Weibull:

- $0 < c < 1$ , τότε έχουμε αρνητική γήρανση πληθυσμού
- $c = 1$ , τότε δεν υπάρχει γήρανση πληθυσμού (σταθερή γήρανση)
- $c > 1$ , τότε παρατηρείται θετική γήρανση πληθυσμού

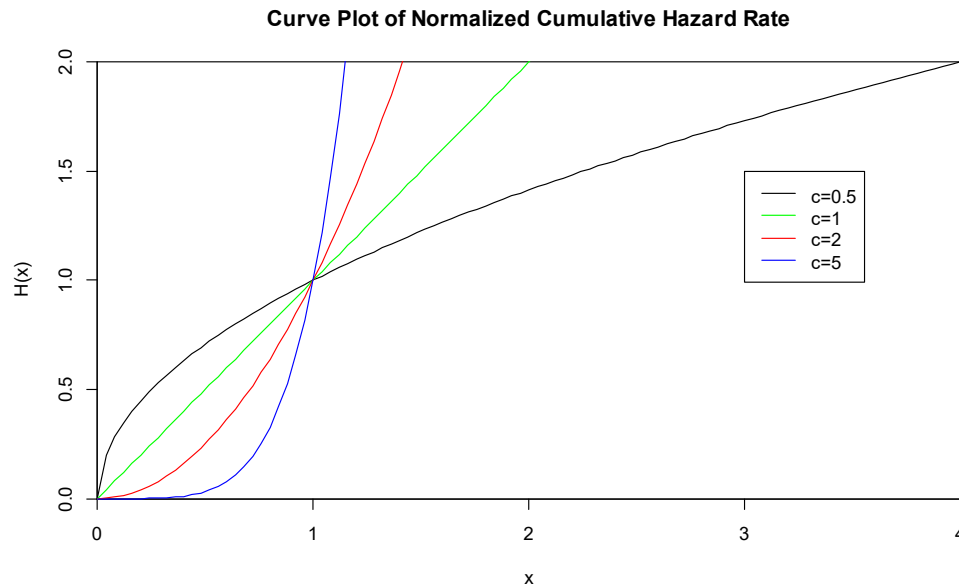
Μεταβαίνοντας στην Αθροιστική Συνάρτηση Κινδύνου (Cumulative Hazard Rate), έχουμε ότι αυτή για την κανονικοποιημένη μορφή της Weibull, ισούται με:

$$H_U(u|c) = \int_0^u h_U(v|c) dv = \int_0^u cv^{c-1} dv = u^c \quad (1.6.3)$$

Αντίστοιχα, για τον κλασσικό ορισμό της Weibull έχουμε ότι:

$$\begin{aligned} H_X(x|a, b, c) &= \int_0^x h_X(u|a, b, c) du = \int_0^x \frac{c}{b} \left(\frac{u-a}{b}\right)^{c-1} du \\ &= \int_0^x \overbrace{\left(\frac{u-a}{b}\right)^c}^{u \geq a} du \\ &= \left(\frac{x-a}{b}\right)^c \end{aligned} \quad (1.6.4)$$

Η γραφική αναπαράσταση της Αθροιστικής Συνάρτησης Κινδύνου απεικονίζεται στο Σχήμα 9:



Σχήμα 9: Αθροιστική Συνάρτηση Κινδύνου της Κατανομής Weibull για τις διάφορες τιμές της παραμέτρου  $c$

Από το Σχήμα προκύπτει ότι η αύξηση της Αθροιστικής Συνάρτησης Κινδύνου πραγματοποιείται:

- **Μονότονα, με μειούμενο ρυθμό** για  $0 < c < 1$
- **Γραμμικά**, για  $c = 1$
- **Μονότονα, με αυξανόμενο ρυθμό** για  $c > 1$

### 1.7 Μέσος Υπολειπόμενος Χρόνος Ζωής (MRL) Weibull

Αναφερόμενοι στον Μέσο Υπολειπόμενο Χρόνο Ζωής, ουσιαστικά αναζητούμε την αναμενόμενη τιμή της μεταβλητής  $Y$ , ορισμένης ως:

$$Y|x := X - x | X \geq x \quad (1.7.1)$$



Ουσιαστικά η μεταβλητή  $Y$  αποτελεί την **Υπολειπόμενη Ζωή (Residual Life)**, ενός ατόμου ή μιας μονάδας γραμμής παραγωγής, δοθέντος ότι ήδη έχει επιβιώσει για  $x$  μονάδες χρόνου.

Ως **Μέσος Υπολειπόμενος Χρόνος Ζωής (Mean Residual Life)** μιας τυχαίας μεταβλητής  $X$ , ορίζεται η εξής ποσότητα:

$$mrl(x) := E(X - x | X > x) \quad (1.7.2)$$

Στηριζόμενοι στα παραπάνω, και εκφράζοντας τον Μέσο Υπολειπόμενο Χρόνο Ζωής μέσω της Συνάρτησης Αξιοπιστίας  $R(x)$  για την κατανομή Weibull, θα έχουμε ότι (Rinne (2008)):

$$\begin{aligned} mrl(x) &= E(Y | X \geq x) = \frac{1}{R_X(x)} \int_x^{\infty} R_X(z) dz \\ &\stackrel{(1.5.3)}{\iff} mrl(x) = \frac{1}{\exp\{-x^c\}} \int_x^{\infty} \exp\{-z^c\} dz \end{aligned} \quad (1.7.3)$$

Μια ενδιαφέρουσα παρατήρηση του Μέσου Υπολειπόμενου Χρόνου Ζωής θα μπορούσε να είναι η συμπεριφορά που παρουσιάζει, για τις διάφορες τιμές της παραμέτρου κλίμακας  $c$ .

Συγκεκριμένα, η Μέση Υπολειπόμενη Διάρκεια Ζωής είναι μια **μονότονη** συνάρτηση, η οποία στην περίπτωση της κατανομής Weibull παίρνει τις ακόλουθες μορφές, ανάλογα με τις τιμές της παραμέτρου σχήματος  $c$  (Rinne (2008)):

- Για  $0 < c < 1$ , αυτή αποτελεί μια **αύξουσα** συνάρτηση
- Για  $c = 1$ , τότε έχουμε μια **σταθερή συνάρτηση στο 1**
- Για  $c > 1$ , η συνάρτηση είναι **φθίνουσα**

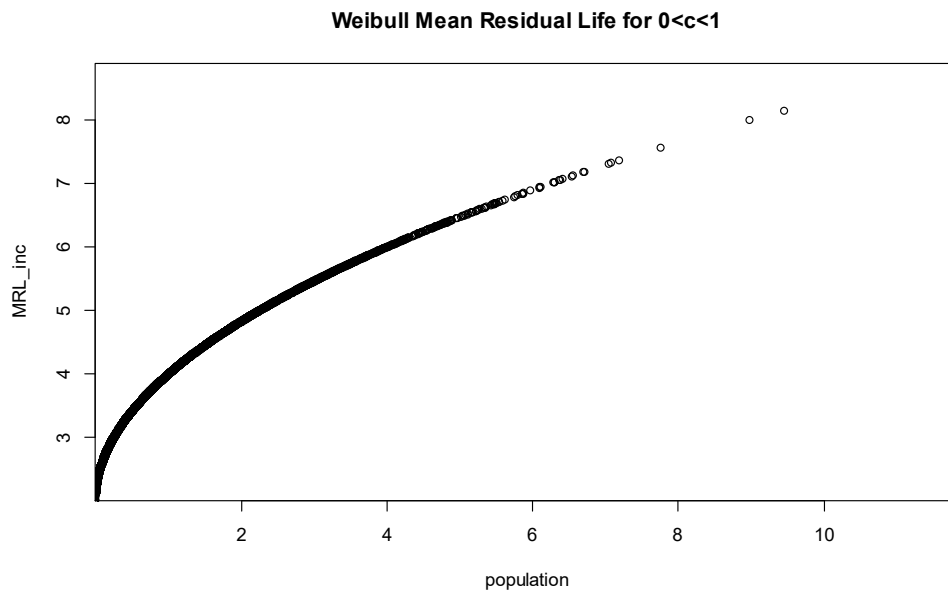
Επίσης, αξίζει να σημειωθεί ότι η αντίστροφη συμπεριφορά ως προς την μονοτονία, παρατηρείται για την βαθμίδα αποτυχίας  $h(x)$  ανά περίπτωση.

Στη συνέχεια παρουσιάζεται μια σχηματική απεικόνιση των όσων αναφέρθηκαν προηγουμένως, ορίζοντας τις εξής μεταβλητές για να καλύψουμε και τις τρεις παραπάνω περιπτώσεις σε ένα τυχαίο δείγμα 10.000 Παρατηρήσεων:

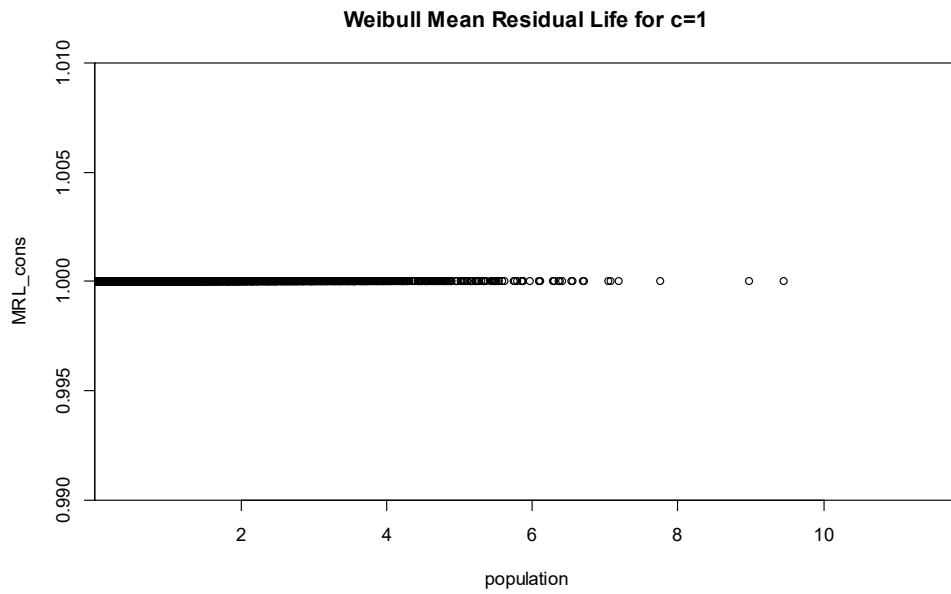
- i. MRL\_inc: Η μεταβλητή περιέχει τον υπολογισμό του MRL για τιμή παραμέτρου σχήματος  $c$  ίση με το 0.5
- ii. MRL\_cons: Η μεταβλητή περιέχει τον υπολογισμό του MRL για σταθερή τιμή παραμέτρου σχήματος  $c$  ίση με τη μονάδα
- iii. MRL\_dec: Η μεταβλητή περιέχει τον υπολογισμό του MRL για τιμή παραμέτρου σχήματος  $c$  ίση με το 2

Περισσότερες λεπτομέρειες, αναφορικά με την διαδικασία επιλογής του δείγματος και ορισμού των μεταβλητών, παρατίθενται στο Παράρτημα που βρίσκεται στο τέλος της Εργασίας.

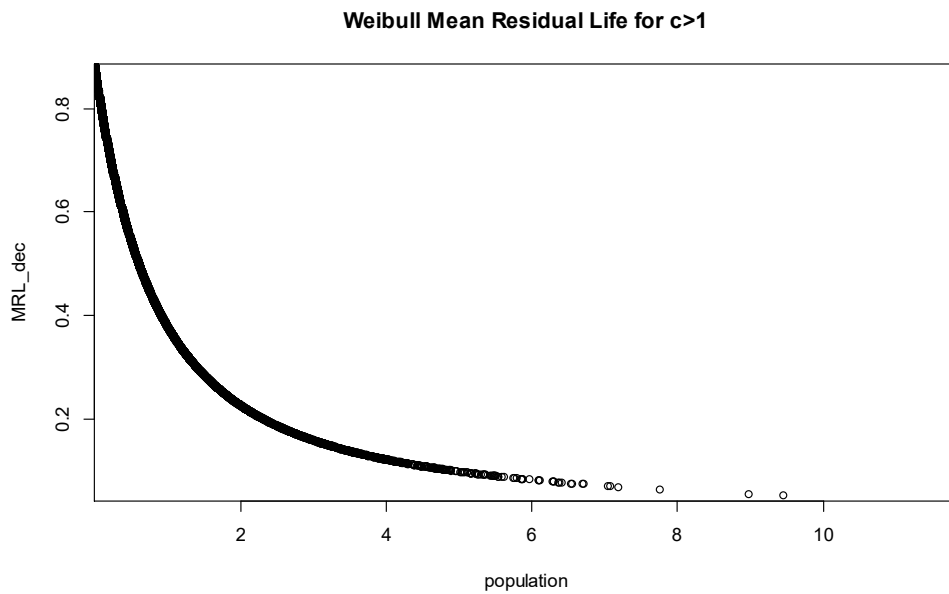
Προχωρώντας στον υπολογισμό των Μέσων Υπολειπόμενων Χρόνων Ζωής Weibull ανά περίπτωση, παρακάτω παρουσιάζονται σχηματικά τα αποτελέσματα για κάθε μία από τις εξεταζόμενες περιπτώσεις μας:



Σχήμα 10: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με  $0 < c < 1$



Σχήμα 11: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με  $c=1$



Σχήμα 12: Γραφική Αναπαράσταση του Μέσου Υπολειπόμενου Χρόνου Ζωής Weibull χαρτοφυλακίου με  $c>1$

## 1.8 Μέση Τιμή και Διακύμανση Weibull

Αναφορικά με την Μέση Τιμή και την Διακύμανση της κατανομής Weibull, αξίζει να σημειωθεί ότι αυτές προκύπτουν μέσω της συνάρτησης Γάμμα, και πιο συγκεκριμένα, μέσω της τιμής της συνάρτησης Γάμμα στα σημεία  $(1+1/c)$  και  $(1+2/c)$ .

Η Μέση Τιμή  $E(X)$  της Κατανομής Weibull για την περίπτωση των 2 Παραμέτρων, προκύπτει από την εξής σχέση (Rinne (2008)):

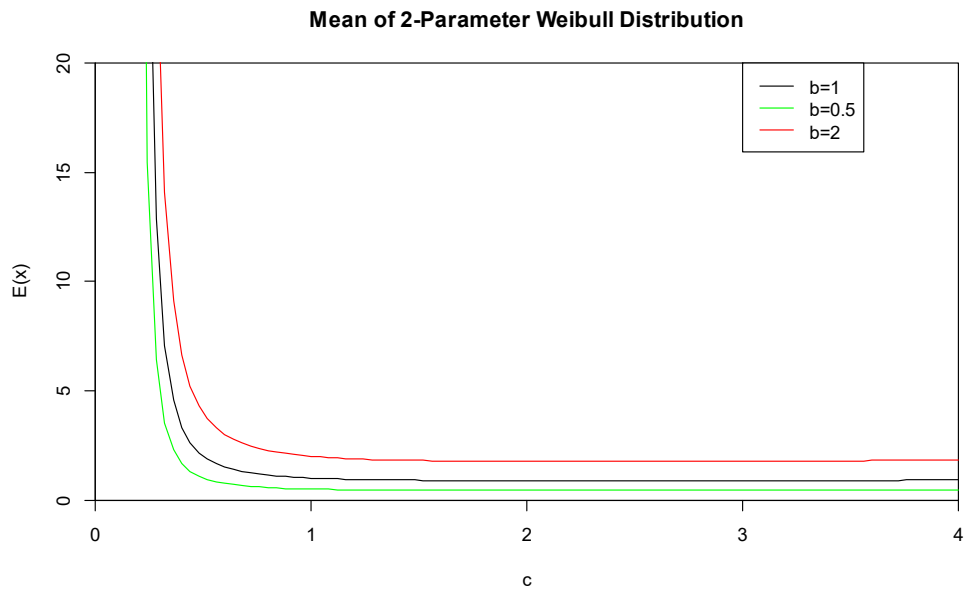
$$E(X) = b \Gamma \left( 1 + \frac{1}{c} \right) \quad (1.8.1)$$

Ενώ για την Διακύμανση  $Var(X)$  της Κατανομής Weibull στην διπαραμετρική περίπτωση, ισχύουν τα παρακάτω:

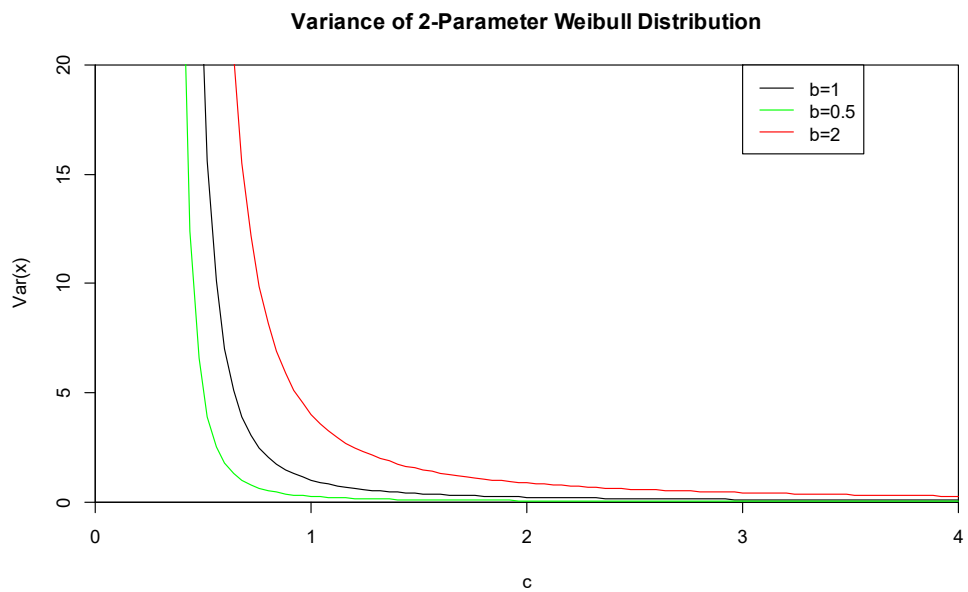
$$Var(X) = E(X^2) - E(X)^2 = b^2 \left[ \Gamma \left( 1 + \frac{2}{c} \right) - \left( \Gamma \left( 1 + \frac{1}{c} \right) \right)^2 \right] \quad (1.8.2)$$

Όπου  $b, c$  οι παράμετροι κλίμακας και σχήματος αντίστοιχα, και  $\Gamma \left( 1 + \frac{1}{c} \right), \Gamma \left( 1 + \frac{2}{c} \right)$  οι τιμές της συνάρτησης Γάμμα στα σημεία  $1 + \frac{1}{c}$  και  $1 + \frac{2}{c}$

Στα παρακάτω διαγράμματα, παρουσιάζονται γραφικά η Μέση Τιμή και η Διακύμανση της Διπαραμετρικής Weibull, για τις διάφορες τιμές της παραμέτρου  $b$ :



Σχήμα 13: Γραφική Απεικόνιση της Μέσης Τιμής Weibull 2 Παραμέτρων για τις διάφορες τιμές της Παραμέτρου Σχήματος



Σχήμα 14: Γραφική Απεικόνιση της Διακύμανσης Weibull 2 Παραμέτρων για τις διάφορες τιμές της Παραμέτρου Σχήματος

## Κεφάλαιο 2 – Η Κατανομή Weibull ως Ζημιοκατανομή

### 2.1 Εισαγωγή

Στο συγκεκριμένο Κεφάλαιο θα αναφερθούμε στο ρόλο που έχει η κατανομή Weibull ως Ζημιοκατανομή στην Αναλογιστική Επιστήμη.

Σε μια ασφαλιστική εταιρεία Γενικών Ασφαλίσεων, μια από τις βασικότερες προκλήσεις που ένας Αναλογιστής ή Διαχειριστής Κινδύνων καλείται να φέρει εις πέρας, είναι εκείνη της μοντελοποίησης της αβεβαιότητας που εμπλέκεται στην αναμενόμενη συχνότητα εμφάνισης **Ασφαλιστικών Απαιτήσεων** (Insurance Claims) καθώς και στην αναμενόμενη **Σφοδρότητα** (Severity) της κάθε Απαίτησης που θα εμφανιστεί. Απώτερος σκοπός των δύο αυτών εκτιμήσεων, είναι η αποτελεσματική Διαχείριση του Κινδύνου που πηγάζει από τα ασφαλιστικά προϊόντα, και η βέλτιστη αποτίμηση της κερδοφορίας αυτών.

Πιο συγκεκριμένα, η αβεβαιότητα η οποία προέρχεται από την συχνότητα εμφάνισης Απαιτήσεων μοντελοποιείται μέσω της χρήσης Διακριτών Κατανομών, ενώ η μοντελοποίηση της Σφοδρότητας των Απαιτήσεων προσεγγίζεται μέσω Συνεχών Κατανομών.

Ένας Αναλογιστής, λοιπόν, καλείται να κατανοήσει όσο το δυνατόν καλύτερα την μελλοντική έκβαση ενός χαρτοφυλακίου Ασφαλιστηρίων Συμβολαίων από Χρηματοοικονομική σκοπιά, μέσω του σχεδιασμού και της χρήσης ενός αποκαλούμενου **μοντέλου Ζημιών** και της χρήσης Τυχαίων Μεταβλητών.

Γενικά, τα μοντέλα Ζημιών αποτελούν αντικείμενα ζωτικής σημασίας για την Αναλογιστική Επιστήμη, ενώ είναι απαραίτητα για την εύρυθμη λειτουργία μιας Ασφαλιστικής Εταιρείας. Με την σειρά της, μια Ασφαλιστική Εταιρεία αποτελεί αναπόσπαστο μέρος της οικονομίας μιας χώρας, καθώς αποτελεί ένα πυλώνα σταθερότητας ως προς την αντιμετώπιση κινδύνων για μια μεγάλη μερίδα ατόμων.

Ένα κοινό χαρακτηριστικό που διέπει τα δεδομένα που αφορούν Γενικές Ασφαλίσεις, είναι η παρουσία **θετικής κυρτότητας**, γεγονός το οποίο καθιστά υποψήφιας προς χρήση κατανομές όπως η Λογαριθμοκανονική, η Pareto, η Γάμμα, η Burr, αλλά και την κατανομή Weibull για την μοντελοποίηση της σφοδρότητας των Απαιτήσεων στον τομέα των Γενικών Ασφαλίσεων.

Για το Κεφάλαιο αυτό θα παρατεθούν πληροφορίες οι οποίες στηρίζονται κυρίως στο Ακαδημαϊκό Άρθρο των Das και Nath (2019), καθώς επίσης και στις Πανεπιστημιακές Σημειώσεις του κ. Πολίτη (2023).

## 2.2 Ροπογεννήτρια Κατανομής Weibull

Ένα σημείο το οποίο παρουσιάζει ιδιαίτερο ενδιαφέρον στη μελέτη των Ζημιοκατανομών, είναι αυτό της Ροπογεννήτριας Συνάρτησης μιας κατανομής.

Έστω  $X$  μια τυχαία μεταβλητή. Τότε καλούμε ως **Ροπογεννήτρια Συνάρτηση** της τ.μ  $X$   $M_X$ , την συνάρτηση της πραγματικής μεταβλητής  $t$  με  $M_X: A \rightarrow \mathbb{R}$  και  $A \subseteq \mathbb{R}$ , η οποία για τη συνεχή περίπτωση εκφράζεται από την παρακάτω σχέση:

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \quad (2.2.1)$$

Αξίζει να σημειωθεί ότι αναγκαία (αλλά όχι ικανή) προϋπόθεση για την ύπαρξη της Ροπογεννήτριας Συνάρτησης, είναι η ύπαρξη όλων των ροπών  $\mu_k = E[X^k]$  της τ.μ  $X$

Επίσης ισχύει ότι:

$$M_X^{(k)}(t) = E[X^k e^{tX}] \Rightarrow M_X^{(k)}(0) = E[X^k] \quad (2.2.2)$$

όπου  $M_X^{(k)}(t)$  η παράγωγος  $k$  τάξης ως προς  $t$  της  $M_X(t)$

Προχωρώντας στην εύρεση της Ροπογεννήτριας Συνάρτησης της κατανομής Weibull, έχουμε ότι η σ.π.π μέσω του ορισμού δύο παραμέτρων είναι:

$$f_X(x|b, c) = \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left\{-\left(\frac{x}{b}\right)^c\right\}$$

και επίσης έχουμε ότι:

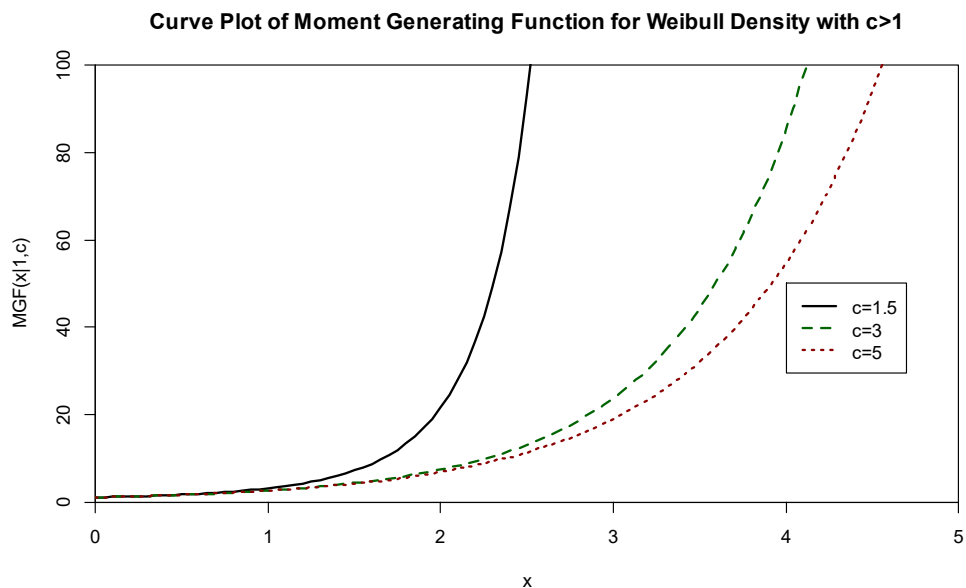
$$\begin{aligned}
 E[X^t] &= \int_0^{\infty} x^t \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left\{-\left(\frac{x}{b}\right)^c\right\} dx \\
 &\xrightarrow{u=\left(\frac{x}{b}\right)} \int_0^{\infty} e^{tbu} c u^{c-1} e^{-u^c} du \\
 &\xrightarrow{x=u^c} \int_0^{\infty} e^{btu^{\frac{1}{c}}} e^{-x} dx = \int_0^{\infty} \sum_{n=0}^{\infty} \frac{(bt)^n}{n!} x^{\frac{n}{c}} e^{-x} dx = \sum_0^{\infty} \frac{b^n t^n}{n!} \Gamma\left(\frac{n}{c} + 1\right), c \geq 1 \quad (2.2.3)
 \end{aligned}$$

Από την σχέση (2.2.3) μπορούμε να συμπεράνουμε ότι η Ροπογεννήτρια Συνάρτηση της κατανομής Weibull καθορίζεται από τις διάφορες τιμές της **παραμέτρου σχήματος c** ως εξής (Πολίτης (2023)):

- Για  $c > 1$ , η Ροπογεννήτρια υπάρχει, δηλαδή αυτή είναι πεπερασμένη εντός ενός διαστήματος της μορφής  $[0, d)$ , όμως ο αναλυτικός υπολογισμός της δεν είναι εφικτός
- Για  $c = 1$ , η Ροπογεννήτρια συνάρτηση ταυτίζεται με αυτήν της **Εκθετικής Κατανομής με Παράμετρο b**
- Για  $c < 1$ , η Ροπογεννήτρια συνάρτηση **δεν υπάρχει**

Παρακάτω θα επιχειρηθεί να δοθεί σχηματικά μια προσέγγιση του ανοικτού τύπου της Ροπογεννήτριας Συνάρτησης για την περίπτωση όπου  $c > 1$ , λαμβάνοντας υπόψιν τις πρώτες 150 τιμές του n στην σχέση (2.2.3).





Σχήμα 15: Προσεγγιστική Σχηματική Απεικόνιση Ροπογεννήτριας Κατανομής Weibull για διάφορες τιμές της παραμέτρου σχήματος  $c > 1$

### 2.3 Ουρά της Κατανομής Weibull

Μεταβαίνοντας σε ακόμα ένα σημαντικό σημείο μελέτης, αναφορικά με τις Ζημιοκατανομές, θα αναφερθούμε στην μελέτη της Ουράς μιας Κατανομής.

Ουσιαστικά με τον όρο «Ουρά» μιας κατανομής, αναφερόμαστε στην εμφάνιση «ακραίων» παρατηρήσεων σε ένα τυχαίο πείραμα ενός πληθυσμού. Οι παρατηρήσεις αυτές παρουσιάζουν ιδιαίτερα χαμηλό ρυθμό εμφάνισης συγκριτικά με την πλειονότητα των παρατηρήσεων ενός δείγματος, το οποίο τοποθετείται διαγραμματικά περίπου στο κεντρικό τμήμα μιας κατανομής (Bingham and Spradlin (2011)). Για αυτό πολλές φορές αναφέρονται και ως **Ακραίες Παρατηρήσεις**.

Έστω  $X$  μια τυχαία μεταβλητή και  $F(x)$  η αθροιστική συνάρτηση κατανομής της  $X$ . Τότε ως **ουρά κατανομής**  $\bar{F}(x)$  καλούμε το μέγεθος που εκφράζεται ως εξής:

$$\bar{F}(x) = 1 - F(x) \quad (2.3.1)$$

Ανάλογα με το αν αναφερόμαστε στην εμφάνιση είτε «πολύ μικρών», είτε «πολύ μεγάλων» παρατηρήσεων, τότε ανατρέχουμε στους όρους «αριστερή» ή «δεξιά» ουρά, αντίστοιχα.

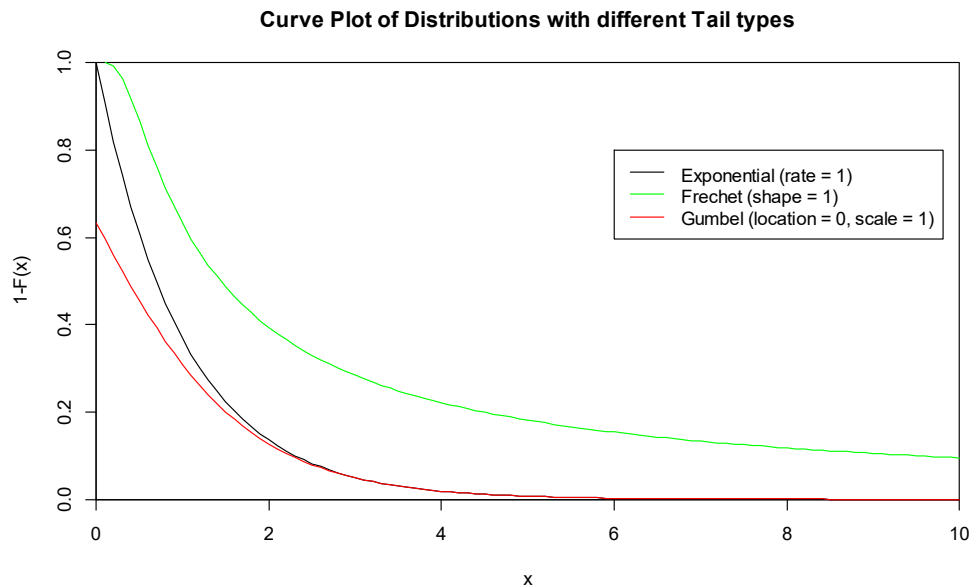
Στον Έλεγχο Υποθέσεων, ένας έλεγχος δεξιάς ουράς (right – tailed test) εξετάζει κατά πόσο ένα στατιστικό μέγεθος είναι σημαντικά μεγαλύτερο από μια αρχική υπόθεση. Ένας έλεγχος αριστερής ουράς (left – tailed test) υποδηλώνει τον ακριβώς αντίθετο έλεγχο.

Με τον όρο **βαριά ουρά** αναφερόμαστε σε μια κατανομή, η οποία έχει ουρά βαρύτερη από αυτήν οποιασδήποτε Εκθετικής Κατανομής (Bryson (1974)). Με άλλα λόγια, μια κατανομή η οποία έχει βαριά ουρά λέμε ότι συγκλίνει στο 0 με πιο αργό ρυθμό από ότι συγκλίνει η Εκθετική Κατανομή.

Διαγραμματικά, αυτό υποδηλώνει ότι θα υπάρχει μεγαλύτερη απόσταση μεταξύ της κατανομής και του οριζόντιου άξονα, συγκριτικά με αυτήν της Εκθετικής.

Στη συνέχεια, παραθέτουμε σχηματικά την απεικόνιση της ουράς δύο κατανομών, η μία αναπαριστά την κατανομή Fréchet, η οποία ανήκει στην οικογένεια κατανομών με βαριά ουρά, ενώ η άλλη την κατανομή Gumbel, η οποία κατηγοριοποιείται στις κατανομές με ελαφριά ουρά.

Επίσης, στο ίδιο γράφημα παρατίθεται και η ουρά της Εκθετικής Κατανομής, ως μέτρο σύγκρισης μεταξύ των δύο προαναφερθεισών κατανομών:



Σχήμα 16: Διαγραμματική Απεικόνιση κατανομών με διαφορετικό τύπο ουράς, συγκριτικά με την ουρά της Εκθετικής Κατανομής

Στην Αναλογιστική Επιστήμη, οι βαριές ουρές κατανομών χρησιμοποιούνται στην μελέτη εμφάνισης μιας πολύ μεγάλης Ασφαλιστικής Απαιτήσης, κατά πολύ μεγαλύτερη από τον μέσο όρο των Απαιτήσεων, σε ένα Χαρτοφυλάκιο Ασφαλιστηρίων Συμβολαίων.

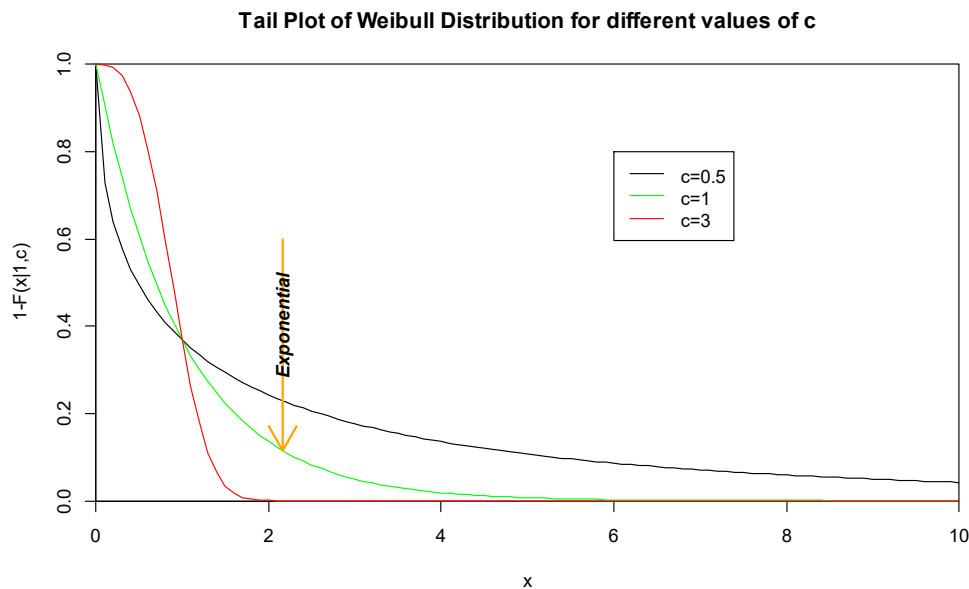
Προχωρώντας στην μελέτη ουράς της Κατανομής Weibull δύο παραμέτρων  $b, c$  για μια τ.μ  $X$ , αυτή μέσω της σχέσης (2.3.1) εκφράζεται ως εξής:

$$\bar{F}(x) = 1 - F(x) = e^{-bx^c}, x \geq 0 \quad (2.3.2)$$

Για τις διάφορες τιμές της παραμέτρου σχήματος  $c$  στην σχέση (2.3.2), λαμβάνουμε τις εξής συμπεριφορές (Πολίτης (2023)):

- Για  $c=1$ , τότε λαμβάνουμε την **Εκθετική Κατανομή**
- Για  $c<1$ , τότε η κατανομή έχει **βαριά ουρά**
- Για  $c>1$ , τότε η κατανομή έχει **ελαφριά ουρά** (πιο ελαφριά από την Εκθετική Κατανομή)

Στη συνέχεια αναπαριστούμε γραφικά τα παραπάνω συμπεράσματα της ουράς Weibull, για τις διάφορες τιμές του  $c$ :



Σχήμα 17: Γραφική απεικόνιση ουράς Weibull, για τις διάφορες τιμές της παραμέτρου σχήματος  $c$

## Κεφάλαιο 3 – Βασικές Έννοιες στην Ανάλυση Επιβίωσης

### 3.1 Εισαγωγή

Καθώς ο ρόλος της Κατανομής Weibull στην Αναλογιστική Επιστήμη είναι διττός, τόσο στο πλαίσιο των Γενικών Ασφαλίσεων όσο και σε αυτό των Ασφαλίσεων Ζωής, στο παρακάτω Κεφάλαιο κρίνεται σκόπιμο να παρατεθούν συνοπτικά ορισμένες βασικές έννοιες από την Ανάλυση Επιβίωσης, επεκτείνοντας τον χαρακτήρα της Weibull πέραν αυτού της Ζημιοκατανομής.

Στο συγκεκριμένο Κεφάλαιο και ως επί το πλείστον, αντλήθηκαν πληροφορίες από το Κεφάλαιο 2.1 του Συγγράμματος του Horst Rinne “The Weibull Distribution – A Handbook”, καθώς επίσης και από τις Διαφάνειες του μαθήματος «Ανάλυση Επιβίωσης» του κ. Ζήμερα (2020).

### 3.2 Συνάρτηση Πυκνότητας Πιθανότητας Θανάτου

Ένα από τα βασικά σημεία στην Ανάλυση Επιβίωσης, αναφορικά με την μελέτη της διάρκειας ζωής ενός ατόμου είναι ότι αυτή δεν μπορεί να υπολογιστεί με κάποια προκαθορισμένη και ντετερμινιστική μέθοδο. Αντίθετα, η **διάρκεια ζωής** ενός ατόμου συμβολίζεται με μια συνεχή, μη-αρνητική τυχαία μεταβλητή, έστω  $X$ .

Στο συγκεκριμένο σημείο θα προχωρήσουμε στον ορισμό των συναρτήσεων που προσδιορίζουν πλήρως την κατανομή που ακολουθεί η διάρκεια ζωής **ενός ατόμου** (τυχαία μεταβλητή  $X$ ) ή **περισσότερων ατόμων με κοινά χαρακτηριστικά** (ακολουθία τυχαίων μεταβλητών  $X_i$ ).

Έστω ότι έχουμε μια ομάδα ατόμων με χρόνους ζωής  $x_1, x_2, \dots, x_i$  τα οποία για κάθε  $i$  απεικονίζουν τους χρόνους επιβίωσης στην κάθε περίπτωση μέσα από το δείγμα που έχουμε συλλέξει, και έστω η συνεχής μεταβλητή  $X \in R^+$  ούτως ώστε κάθε  $x_i$  να είναι μια συγκεκριμένη τιμή που αυτή παίρνει, τότε:

Η συνάρτηση πυκνότητας πιθανότητας (ή πιο σύντομα σ.π.π)  $f(x)$  η οποία ικανοποιεί τις παρακάτω σχέσεις:

$$f(x) \geq 0 \quad \forall x \quad \text{και} \quad \int_{-\infty}^{+\infty} f(x) dx = 1 \quad (3.2.1)$$

καλείται **συνάρτηση πυκνότητας πιθανότητας αποτυχίας ή θανάτου (failure density)** και απεικονίζει την πιθανότητα θανάτου ενός ατόμου ηλικίας ( $x$ ).

Η αδέσμευτη πιθανότητα ενός νεογέννητου ατόμου να αποβιώσει στην ηλικία των  $x$  ετών δίνεται από την σχέση:

$$P(x - \Delta x/2 < X \leq x + \Delta x/2) \approx f(x) \Delta x \quad (3.2.2)$$

για πολύ μικρές τιμές του  $\Delta x$ .

Η αδέσμευτη πιθανότητα ενός νεογέννητου ατόμου να αποβιώσει μέχρι και το διάστημα που ορίζεται μεταξύ των ηλικιών  $x_l$  και  $x_u$  με  $x_l < x_u$ , δίνεται από την σχέση:

$$P(x_l < X < x_u) = \int_{x_l}^{x_u} f(x) dx \quad (3.2.3)$$

Επίσης, αναφορικά με την γραφική αναπαράσταση της  $f(x)$ , σημειώνεται ότι το διάγραμμα απεικόνισης της διάρκειας ζωής ονομάζεται **καμπύλη θανάτου**.

### 3.3 Συνάρτηση Κατανομής Διάρκειας Ζωής

Έστω  $X$  η συνεχής τ.μ., όπως αυτή ορίστηκε στην παράγραφο 3.2. Η (αθροιστική) συνάρτηση κατανομής  $F(x)$  με:

$$F(x) := P(X \leq x) = \int_0^x f(t) dt \quad (3.3.1)$$

Στην Ανάλυση Επιβίωσης καλείται **Συνάρτηση Κατανομής Διάρκειας Ζωής** (ή πιο σύντομα σ.κ).

Γνωρίζουμε ότι το γινόμενο  $f(x) dx = dF(x)$  εκφράζει την απειροστή πιθανότητα που, στην περίπτωση της Ανάλυσης Επιβίωσης, αυτό μεταφράζεται στο θάνατο του ατόμου

μέσα στο απειροστό διάστημα  $[x, x + \Delta x]$ . Επομένως, μπορούμε να αντιληφθούμε ότι η σχέση:

$$f(x) = \frac{dF(x)}{dx} \quad (3.3.2)$$

εκφράζει τον στιγμιαίο ρυθμό θανάτων.

Επίσης, η σχέση (3.2.2) μπορεί να γραφεί συναρτήσει της Συνάρτησης Κατανομής Διάρκειας Ζωής ως εξής:

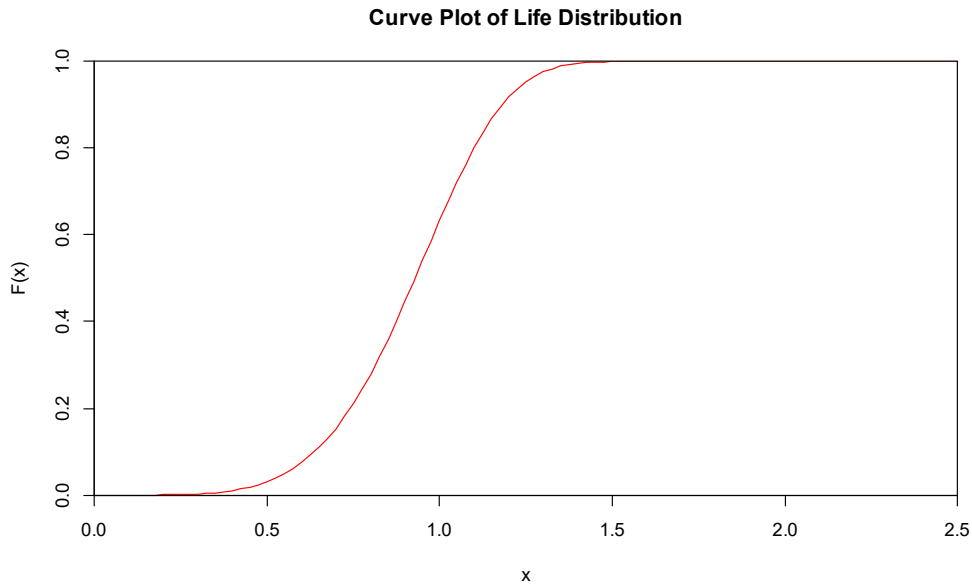
$$P(x - \Delta x/2 < X \leq x + \Delta x/2) \approx f(x) \Delta x = \lim_{\Delta x \rightarrow 0^+} (F(x + \Delta x/2) - F(x - \Delta x/2)) \quad (3.3.3)$$

Η  $F(x)$  ουσιαστικά εκφράζει την πιθανότητα ένα άτομο να αποβιώσει έως και την χρονική στιγμή  $x$ .

Επίσης, αξίζει να σημειωθεί ότι η σ.κ  $F(x)$  είναι μια αύξουσα συνάρτηση του  $x$ , η οποία ικανοποιεί τις δύο παρακάτω συνθήκες:

$$\lim_{x \rightarrow -\infty} F(x) = 0 \text{ και } \lim_{x \rightarrow +\infty} F(x) = 1 \quad (3.3.4)$$

Η συμπεριφορά της σ.κ. που αναλύεται στην σχέση (3.3.4) μπορεί να γίνει καλύτερα αντιληπτή μέσω και της γραφικής απεικόνισής της συνάρτησης κατανομής διάρκειας ζωής της κατανομής Weibull παρακάτω:



Σχήμα 18: Γραφική Αναπαράσταση της Συνάρτησης Κατανομής Διάρκειας Ζωής Weibull

### 3.4 Συνάρτηση Επιβίωσης

Έστω  $X$  η συνεχής τ.μ., όπως αυτή ορίστηκε στην παράγραφο 3.2. Η συνάρτηση  $S(x)$  η οποία εκφράζει την πιθανότητα για το συμπληρωματικό ενδεχόμενο της συνάρτησης κατανομής διάρκειας ζωής, καλείται **Συνάρτηση Επιβίωσης**. Δηλαδή, ορίζουμε:

$$S(x) := P(X > x) = 1 - P(X \leq x) = \int_x^{\infty} f(z) dz \quad (3.4.1)$$

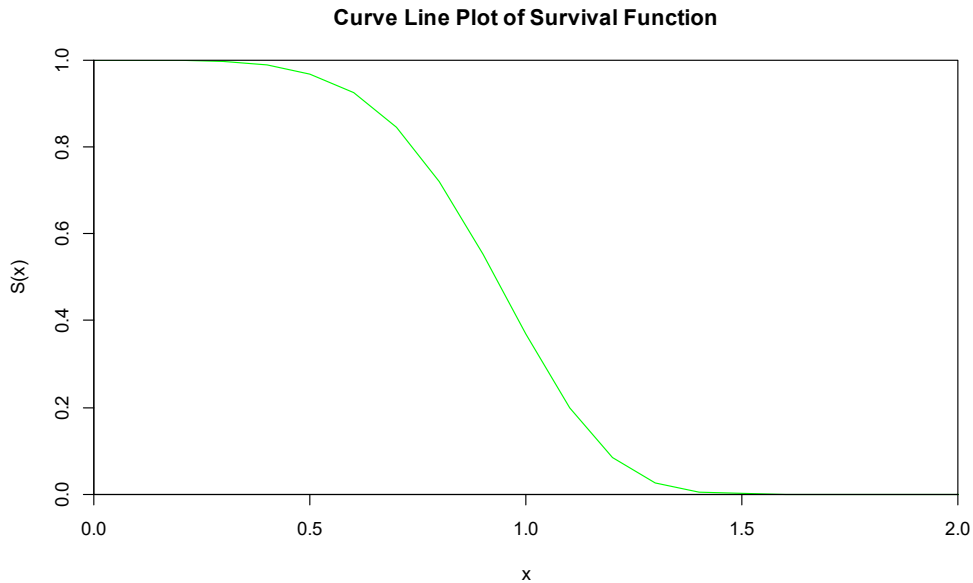
Σύμφωνα με τον ορισμό που δόθηκε από την σχέση (3.4.1), η  $S(x)$  εκφράζει την πιθανότητα το άτομο να επιβιώσει μέχρι και τη χρονική στιγμή  $x$ . Η συνάρτηση επιβίωσης αποτελεί μια φθίνουσα συνάρτηση του  $x$ , η οποία ικανοποιεί τις παρακάτω 2 συνθήκες:

$$\lim_{x \rightarrow -\infty} S(x) = 1 \text{ και } \lim_{x \rightarrow +\infty} S(x) = 0 \quad (3.4.2)$$

Παρατηρώντας την σχέση (3.4.2), συγκριτικά με την σχέση (3.3.4) της προηγούμενης παραγράφου, η οποία εξέφραζε τις αντίστοιχες συνθήκες που ικανοποιεί η σ.κ.  $F(x)$ , βλέπουμε ότι η συμπληρωματικότητα των 2 αυτών συναρτήσεων επεκτείνεται και στην μονοτονία τους.

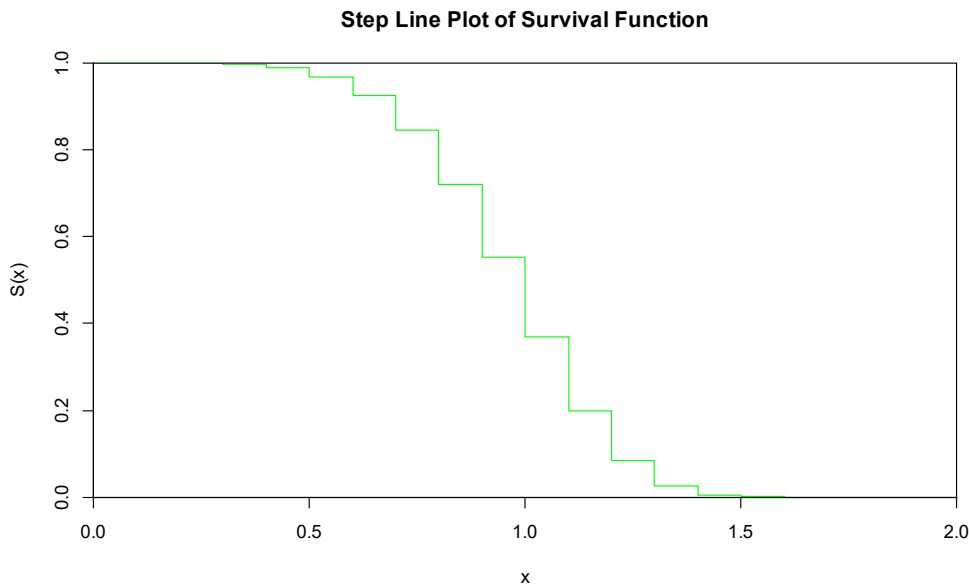
Αναφορικά με την γραφική παράσταση της συνάρτησης επιβίωσης, ο κ. Ζήμερας στις σημειώσεις της Ανάλυσης Επιβίωσης αναφέρει ότι αυτή αναπαρίσταται από μία **ομαλή (λεία) καμπύλη**. Αν και αυτό ισχύει θεωρητικά για συνεχείς κατανομές, στην πραγματικότητα σε εφαρμογές μέσα από την χρήση δεδομένων, η μορφή της αποτελεί μια **βηματική συνάρτηση (step function)**.

Η διαφορά στην μορφολογία των δύο προαναφερθέντων γραφικών απεικονίσεων μπορεί να διαπιστωθεί καλύτερα μέσω των παρακάτω σχημάτων, τα οποία απεικονίζουν την Συνάρτηση Επιβίωσης με δύο διαφορετικούς τρόπους, αυτόν της ομαλής (λείας) απεικόνισης των θεωρητικών κατανομών όπως η Weibull, και αυτόν της βηματικής απεικόνισης που ανταποκρίνεται σε ένα σύνολο δεδομένων.



Σχήμα 19: Θεωρητική Γραφική Αναπαράσταση Συνάρτησης Επιβίωσης





Σχήμα 20: Εμπειρική Γραφική Αναπαράσταση Συνάρτησης Επιβίωσης

Τέλος, σκόπιμο κρίνεται να αναφερθεί ότι, λόγω του ότι η χρονική περίοδος της μελέτης που πραγματοποιούμε είναι πάντοτε πεπερασμένη εντός ενός προκαθορισμένου χρονικού ορίζοντα, υπάρχει η πιθανότητα να υπάρχουν κίνδυνοι οι οποίοι ανταγωνίζονται μεταξύ τους, ως προς την εμφάνιση του γεγονότος θανάτου. Επομένως, είναι πιθανό ότι σε ορισμένα άτομα από έναν πληθυσμό που μελετάται μπορεί να μην συμβεί το αντίστοιχο μελετώμενο γεγονός.

### 3.5 Ένταση Θνησιμότητας

Έστω  $X$  μία συνεχής τ.μ., όπως αυτή ορίστηκε στην παράγραφο 3.2. Γενικά, ως **Ένταση Κινδύνου** ή **Βαθμίδα αποτυχίας (Hazard Rate)** με συμβολισμό  $h(x)$ , ορίζεται η:

$$h(x) := \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x \mid X > x)}{\Delta x} \quad (3.5.1)$$

η οποία εκφράζει την δεσμευμένη πιθανότητα θανάτου, ή εναλλακτικά τον στιγμιαίο ρυθμό αποβίωσης την χρονική στιγμή  $x$ , δοθέντος ότι το άτομο έχει επιβιώσει μέχρι την χρονική στιγμή  $x$ .

Στον Αναλογισμό, και συγκεκριμένα στην μελέτη των Ασφαλίσεων Ζωής, η Βαθμίδα Αποτυχίας καλείται ως Ένταση Θνησιμότητας (**Force of Mortality**) της τ.μ.  $X$  και συμβολίζεται με  $\mu(x)$ .

Επιπροσθέτως, η Ένταση Θνησιμότητας  $\mu(x)$  ικανοποιεί τις 2 εξής συνθήκες (Rinne (2008)):

$$\mu(x) \geq 0 \quad \forall x \quad \text{και} \quad \int_0^{\infty} \mu(x) dx = \infty \quad (3.5.2)$$

Επιπλέον, χρησιμοποιώντας τις σχέσεις που δόθηκαν για την σ.π.π. και την Συνάρτηση Επιβίωσης της συνεχούς τ.μ.  $X$ , μπορούμε να γράψουμε την σχέση (3.5.1) ως εξής:

$$\mu(x) = \frac{f(x)}{S(x)} \quad (3.5.3)$$

Αξίζει να αναφερθεί ότι, η ένταση θνησιμότητας ουσιαστικά αποτελεί ταυτόσημη έννοια με αυτήν της Έντασης Κινδύνου/Βαθμίδας Αποτυχίας, η οποία έχει ήδη οριστεί στην σχέση (1.6.1).

Οι διαφορετικές ονομασίες χρησιμοποιούνται ανάλογα με το πεδίο εφαρμογής του συγκεκριμένου μεγέθους κάθε φορά (αν μελετάμε την αποτυχία ενός συνόλου μονάδων παραγωγής ή την θνησιμότητα ενός πληθυσμού ατόμων)

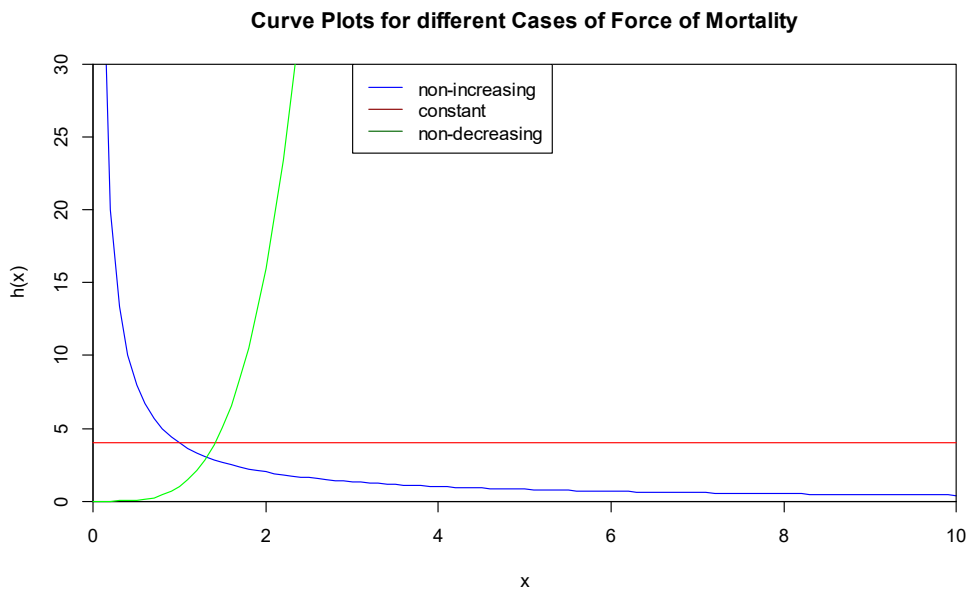
Η ένταση θνησιμότητας παρουσιάζει πολύ μεγάλο ενδιαφέρον στην Αναλογιστική Επιστήμη, καθώς χρησιμοποιείται στην μελέτη των **Πινάκων Επιβίωσης (lifetables)**.

Οι Πίνακες Επιβίωσης αποτελούν πολύτιμο εργαλείο των Αναλογιστών παγκοσμίως, με σκοπό την τιμολόγηση των ασφαλιστικών προϊόντων, καθώς και τον έλεγχο των αποθεμάτων που έχουν υπολογιστεί από τις τεχνικές προβλέψεις, για την πληρωμή των υποχρεώσεων (**liabilities**) που θα πρέπει να καταβάλει μια Ασφαλιστική Επιχείρηση Ασφαλίσεων Ζωής. Πιο συγκεκριμένα, οι Πίνακες Επιβίωσης χρησιμοποιούνται για να μπορέσουν να εκτιμηθούν τόσο το πλήθος, όσο και η χρονική στιγμή των μελλοντικών θανάτων των ασφαλισμένων στον χρόνο. Αυτό καθίσταται εφικτό μέσω της μελέτης ιστορικών δεδομένων (historical data), αναφορικά με την συχνότητα των θανάτων του παρελθόντος, και έτσι πραγματοποιούνται προβλέψεις για την εξέλιξη αυτών στα μελλοντικά χρονικά διαστήματα.

Επιπλέον, ενδιαφέρον παρουσιάζει και η μελέτη της παραγώγου 1<sup>ης</sup> τάξης της έντασης θνησιμότητας  $\mu'(x)$ . Σύμφωνα με τον Rinne (2008), το συγκεκριμένο μέγεθος παρουσιάζεται ως μέτρο **Γήρανσης Πληθυσμού**, και ισχύουν οι εξής 3 υποπεριπτώσεις για κάθε  $x$ :

$$\mu'(x) = \begin{cases} > 0, & \text{θετική γήρανση πληθυσμού (αύξουσα)} \\ = 0, & \text{μηδενική γήρανση πληθυσμού (σταθερή)} \\ < 0, & \text{αρνητική γήρανση πληθυσμού (φθίνουσα)} \end{cases} \quad (3.5.4)$$

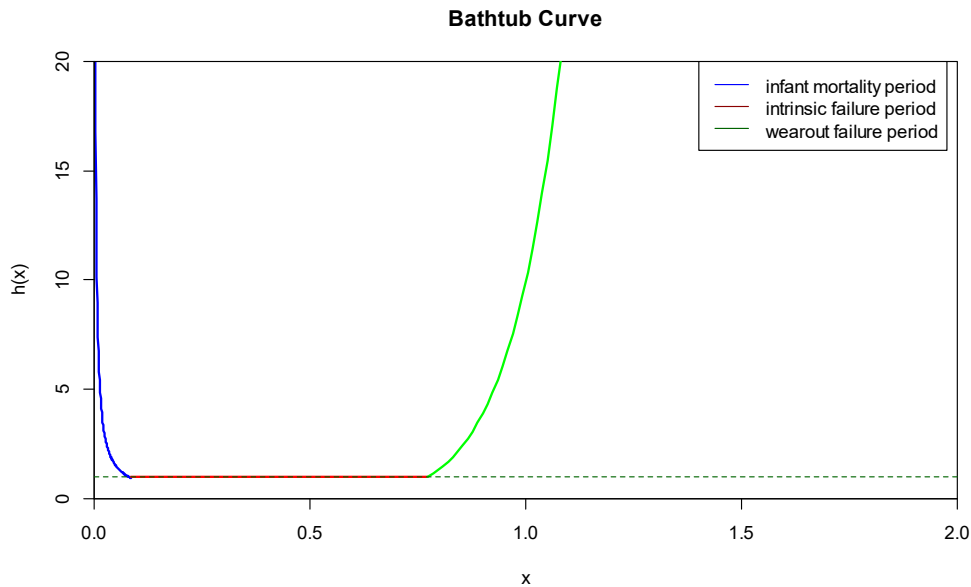
Οι 3 παραπάνω περιπτώσεις μπορούν να παρασταθούν γραφικά και ως εξής:



Σχήμα 21: Γραφική Αναπαράσταση Διαφορετικών Περιπτώσεων Έντασης Θνησιμότητας

Από το Σχήμα 21, μπορεί εύκολα να διαπιστώσει κανείς ότι, η ένωση των 3 διαφορετικών περιπτώσεων που απαρτίζουν την ένταση θνησιμότητας, σχηματίζει μια καμπύλη μορφολογίας “U”, η οποία απεικονίζει την ανθρώπινη θνησιμότητα σε ολόκληρη την

χρονική διάρκεια ζωής του ατόμου, από την γέννηση έως τον Θάνατο. Η συγκεκριμένη καμπύλη είναι γνωστή και ως **Λεκανοειδής Καμπύλη (Bathtub Curve)**.



Σχήμα 22: Γραφική Αναπαράσταση της Λεκανοειδούς Καμπύλης

Παρατηρώντας το Σχήμα 22, μπορούμε να διακρίνουμε τρεις διαφορετικές περιόδους (φάσεις) στη διάρκεια ζωής ενός ατόμου ή μιας μονάδας (Μπούτσικας (2003-8)):

- I. Η πρώτη περίοδος καλείται **βρεφική περίοδος (early life period/ infant mortality period)**. Σε αυτή τη χρονική περίοδο ένα άτομο ή μια μονάδα ξεκινάει με μεγάλη βαθμίδα αποτυχίας η οποία μειώνεται σταδιακά, καθώς όσο περνά ο χρόνος, η μονάδα λειτουργεί ομαλότερα. Η αρχικά αυξημένη βαθμίδα αποτυχίας μπορεί να αποδοθεί σε ενδεχόμενες αδυναμίες σχεδίασης ή κατασκευής αν μιλάμε για μονάδες παραγωγής, ή δυσμενών συνθηκών επιβίωσης (ασθένειες) αν αναφερόμαστε σε άτομα
- II. Αν το άτομο/η μονάδα περάσει το πρώτο στάδιο της βρεφικής περιόδου, τότε συνέχεια έχει η **χρήσιμη περίοδος (useful period) ζωής**. Σε αυτό το χρονικό διάστημα, η βαθμίδα αποτυχίας παραμένει περίπου σταθερή και σχετικά χαμηλή.
- III. Εάν το άτομο/η μονάδα κατορθώσει να περάσει και τις δύο προηγούμενες χρονικές περιόδους, τότε εισερχόμαστε στην λεγόμενη **περίοδο φθοράς (wear-out period)**. Εδώ, η βαθμίδα αποτυχίας αυξάνεται διαρκώς με την πάροδο του χρόνου. Οι θάνατοι/αποτυχίες εδώ συμβαίνουν λόγω γήρανσης του ατόμου ή φθοράς της μονάδας.

### 3.6 Αθροιστική Συνάρτηση Κινδύνου

Έστω  $X$  η συνεχής τ.μ., όπως αυτή ορίστηκε στην παράγραφο 3.2. Γενικά, ως **Αθροιστική Συνάρτηση Κινδύνου (Cumulative Hazard Rate)** του  $x$ , ορίζουμε την παρακάτω ποσότητα:

$$H(x) := \int_0^x h(z) dz \quad (3.6.1)$$

Αναφορικά με την Αθροιστική Συνάρτηση Κινδύνου, αξίζει να αναφερθεί ότι αυτή αποτελεί μια γνησίως αύξουσα συνάρτηση, η οποία ικανοποιεί τις παρακάτω συνθήκες (Ζήμερας (2020)):

$$H(0) = 0 \text{ και } \lim_{x \rightarrow \infty} H(x) = \infty \quad (3.6.2)$$

Επιπλέον, η  $H(x)$  συνδέεται με την συνάρτηση επιβίωσης μέσω της σχέσης:

$$H(x) = -\ln[S(x)] \quad (3.6.3)$$

### 3.7 Μέσος Χρόνος ζωής και Μέση Υπολειπόμενη Ζωή

Όπως είδαμε και στο Κεφάλαιο 1, ένα μέγεθος το οποίο παρουσιάζει ενδιαφέρον στην Αναλογιστική Επιστήμη και συγκεκριμένα στην Ανάλυση Επιβίωσης, αποτελεί και η μελέτη της μέσης (υπολειπόμενης) ζωής μιας ομάδας ατόμων με ομογενή χαρακτηριστικά.

Ορίζοντας τον **Μέσο Χρόνο Ζωής (Mean Lifetime)** μιας συνεχούς τ.μ., προχωρούμε στην διατύπωση της εξής σχέσης:

$$\mu = \int_0^{\infty} zf(z) dz \quad (3.7.1)$$

Μάλιστα αποδεικνύεται ότι η διακύμανση (variance) της τ.μ.  $X$ , μπορεί να συνδεθεί με την συνάρτηση επιβίωσης μέσα από την εξής σχέση:

$$\text{Var}(X) = 2 \int_0^{\infty} xS(x) dx - \left[ \int_0^{\infty} S(x) dx \right]^2 \quad (3.7.2)$$

Αναφερόμενοι στην σχέση (1.7.2) του Κεφαλαίου 1, για τον **Μέσο Υπολειπόμενο Χρόνο Ζωής (Mean Residual Life)** μιας τυχαίας μεταβλητής  $X$ , έχουμε ότι:

$$mrl(x) := E(X - x | X > x)$$

η οποία εκφράζει την αναμενόμενη (μέση) ζωή μιας παρατήρησης που έχει ήδη φτάσει στην χρονική στιγμή  $x$ , έχει επιβιώσει δηλαδή έως την χρονική στιγμή  $x$ . Παρατηρούμε ότι η ποσότητα  $mrl(x)$  αποτελεί μια δεσμευμένη μέση τιμή για τον υπολειπόμενο χρόνο επιβίωσης. Επίσης ισχύει ότι  $mrl(0) = \mu$ .

Συνεχίζοντας την ανάλυση της σχέσης (1.7.2) με την βοήθεια της συνάρτησης επιβίωσης  $S(x)$ , έχουμε ισοδύναμα ότι (Rinne (2008)):

$$mrl(x) = \frac{\int_0^{\infty} zf(z) dz}{S(x)} - x = \frac{1}{S(x)} \int_x^{\infty} S(z) dz, \text{ αν } \lim_{x \rightarrow \infty} x^2 f(x) = 0 \quad (3.7.3)$$

Επιπροσθέτως για την  $mrl(x)$ , αυτή αποδείχθηκε ότι ικανοποιεί τις εξής τρεις συνθήκες (Swartz 1973):

$$mrl(x) \geq 0 \text{ και } \frac{dmrl(x)}{dx} \geq -1 \text{ και } \int_0^{\infty} \frac{1}{mrl(x)} dx = \infty \quad (3.7.4)$$

Επίσης, μέσω της σχέσης (3.7.4), μπορεί να εξαχθεί το συμπέρασμα ότι ο μέσος χρόνος ζωής  $\mu$  ενός νεογέννητου ατόμου υπολογίζεται ως εξής:

$$mrl(0) = E(X) = \mu = \int_0^{\infty} S(z) dz \quad (3.7.5)$$

## Κεφάλαιο 4 – Εφαρμογές της Κατανομής Weibull

### 4.1 Εισαγωγή

Όπως ήδη αναφέρθηκε στα προηγούμενα Κεφάλαια, η Κατανομή Weibull αποτελεί μια κατανομή βαρύνουσας σημασίας, καθώς χρησιμοποιείται σε ένα ευρύτατο πεδίο εφαρμογών στην Αναλογιστική Επιστήμη και όχι μόνο.

Παρακάτω θα γίνει η απόπειρα να παρατεθούν ορισμένες ενδεικτικές εφαρμογές της Weibull, προσαρμόζοντας την κατανομή σε πραγματικά δεδομένα Γενικών Ασφαλίσεων.

Για την εκπόνηση των κατωτέρω εφαρμογών μέσω του προγραμματιστικού πακέτου R, έχουν ληφθεί υπ' όψιν, ως επί το πλείστον, οι Πανεπιστημιακές Σημειώσεις του μαθήματος «Γλώσσα Προγραμματισμού R με εφαρμογές στον Αναλογισμό» από τον κ. Αντζουλάκο (2023), καθώς επίσης έγινε χρήση του Συγγράμματος “Modern Actuarial Risk Theory Using R” - Kaas et al. (2009)

### 4.2 Εκτίμηση Παραμέτρων Κατανομής Weibull

Ένα από τα κυριότερα προβλήματα που καλούμαστε να φέρουμε εις πέρας για την μελέτη μιας κατανομής, είναι η εκτίμηση των παραμέτρων που την απαρτίζουν, με σκοπό την εύρεση εκείνων που επεξηγούν καλύτερα τα δεδομένα που διαχειριζόμαστε σε κάθε περίπτωση.

Για την Εκτίμηση των Παραμέτρων μιας κατανομής παρουσιάζουμε τέσσερις βασικές μεθοδολογίες:

#### 1) Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum Likelihood Estimation – MLE)

Έστω ότι έχουμε  $X = (X_1, X_2, \dots, X_n)$  ένα τυχαίο δείγμα ενός πληθυσμού με Συνάρτηση Πυκνότητας Πιθανότητας  $f(x; \theta)$ , όπου  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ .

Τότε, το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  το οποίο μεγιστοποιεί τη Συνάρτηση Πιθανοφάνειας:

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_k) \quad (4.2.1)$$



καλείται **Εκτιμητής Μέγιστης Πιθανοφάνειας (EMΠ)** του  $\theta$ .

## 2) Μέθοδος των Ροπών (Moment Matching Estimation – MME)

Όμοια με την προηγούμενη περίπτωση, έστω ότι έχουμε  $X = (X_1, X_2, \dots, X_n)$  ένα τυχαίο δείγμα ενός πληθυσμού με Συνάρτηση Πυκνότητας Πιθανότητας  $f(x; \theta)$ , όπου  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ .

Τότε, το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  το οποίο προκύπτει από την επίλυση του συστήματος  $\{\mu_1 = m_1, \mu_2 = m_2, \dots, \mu_k = m_k\}$  καλείται **Εκτιμητής του  $\theta$  με την μέθοδο των ροπών**. Στο παραπάνω σύστημα, με  $\mu_k$  συμβολίζονται οι **θεωρητικές ροπές** και με  $m_k$  οι **δειγματικές ροπές κ - τάξης**.

Συγκεκριμένα, για τον υπολογισμό των  $m_i, \mu_i$  με  $i=1, 2, \dots$  ισχύουν οι παρακάτω σχέσεις:

$$\mu_i = E(X^i) \text{ και } m_i = \frac{1}{n} \sum_{k=1}^n x_k^i \quad (4.2.2).$$

Στην περίπτωση όπου οι εκτιμώμενες παράμετροι είναι τουλάχιστον 2 (όπως συμβαίνει στους ορισμούς δύο και τριών παραμέτρων της κατανομής Weibull), μπορούμε να χρησιμοποιήσουμε τις παρακάτω σχέσεις:

$$\{\mu_1 = m_1, \mu'_2 = m'_2, \dots, \mu'_k = m'_k\}$$

όπου για  $i=2, 3, \dots$  γίνονται:

$$\mu'_i = E([X - E(X)]^i) \text{ και } m'_i = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^i \quad (4.2.3)$$

## 3) Εκτιμητής Ελάχιστης Απόστασης (Minimum Distance Estimation – MDE)

Μέσω της συγκεκριμένης μεθόδου, ελαχιστοποιείται μια επιλεγμένη απόσταση μεταξύ της θεωρητικής και της εμπειρικής συνάρτησης κατανομής.

Έστω ότι έχουμε  $X = (X_1, X_2, \dots, X_n)$  ένα τυχαίο δείγμα ενός πληθυσμού με Συνάρτηση Πυκνότητας Πιθανότητας  $f(x; \theta)$  και Συνάρτηση Κατανομής  $F(x; \theta)$ , όπου  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ . Τότε, η εκτίμηση των παραμέτρων της συνάρτησης κατανομής μπορεί να πραγματοποιηθεί με τρεις διαφορετικούς τρόπους:

- I. Μέσω της μεθόδου **Cramer-von Mises**, υπολογίζεται το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  που ελαχιστοποιεί την στατιστική συνάρτηση:

$$W_n^2 = n \int_{-\infty}^{\infty} [F(x; \theta) - F_n(x; \theta)]^2 dF(x) \quad (4.2.4)$$

II. Με την μέθοδο **Anderson-Darling**, υπολογίζεται το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  που ελαχιστοποιεί την στατιστική συνάρτηση:

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{[F(x; \theta) - F_n(x; \theta)]^2}{[F(x; \theta)][1 - F(x; \theta)]} dF(x) \quad (4.2.5)$$

III. Με χρήση της μεθόδου **Kolmogorov-Smirnov**, υπολογίζεται το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  που ελαχιστοποιεί την στατιστική συνάρτηση:

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x; \theta) - F(x; \theta)| \quad (4.2.6)$$

Στις παραπάνω περιπτώσεις, όπου  $F_n(x; \theta)$  η Εμπειρική Συνάρτηση Κατανομής.

Γενικά, αν  $(X_1, X_2, \dots, X_n)$  ανεξάρτητες και ισόνομες Τυχαίες Μεταβλητές με κοινή συνάρτηση κατανομής  $F(t)$ , τότε η Εμπειρική Συνάρτηση Κατανομής  $\widehat{F}_n(t)$  ορίζεται ως εξής:

$$\widehat{F}_n(t) = \frac{\text{πλήθος στοιχείων του δείγματος} \leq t}{n} = \frac{1}{n} \sum_{i=1}^n 1_{X_i \leq t} \quad (4.2.7)$$

όπου  $1_{X_i \leq t}$  η δείκτρια τυχαία μεταβλητή του ενδεχομένου  $\{X_i \leq t\}$

#### 4) Μέθοδος Ποσοστιαίων Σημείων (Quantile Matching)

Όμοια με τα προηγούμενα, έστω ότι έχουμε  $X = (X_1, X_2, \dots, X_n)$  ένα τυχαίο δείγμα ενός πληθυσμού με Συνάρτηση Πυκνότητας Πιθανότητας  $f(x; \theta)$  όπου  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ .

Τότε, το διάνυσμα  $\widehat{\theta}(x) = (\widehat{\theta}_1(x), \widehat{\theta}_2(x), \dots, \widehat{\theta}_k(x))$  που λειτουργεί ως εκτιμητής του διανύσματος  $\theta$ , υπολογίζεται εξισώνοντας θεωρητικά ποσοστιαία σημεία με τα αντίστοιχα εμπειρικά σημεία. Επομένως, θέλουμε να ισχύει η εξής σχέση:

$$F^{-1}(p_r; \theta) = Q_{n,p_r} \quad (4.2.8)$$

για  $r = 1, 2, \dots, k$  όπου  $Q_{n,p_r}$  είναι το  $p_r$  εμπειρικό ποσοστιαίο σημείο και  $F(x; \theta)$  είναι η θεωρητική Συνάρτηση Κατανομής του πληθυσμού.

## ΕΦΑΡΜΟΓΗ 4.1

Στην συνέχεια, θα παρουσιάσουμε μια εφαρμογή της Προσαρμογής της Κατανομής Weibull, βασιζόμενοι στις μεθόδους που περιεγράφηκαν προηγουμένως, χρησιμοποιώντας δεδομένα του πίνακα “dataOhlsson”, ο οποίος περιέχει τα κόστη αποζημιώσεων Δίκυκλων Οχημάτων, σε Ευρώ, της Σουηδικής πρώην Ασφαλιστικής Εταιρείας Wasa για το χρονικό διάστημα 1994-1998, με την βοήθεια του στατιστικού πακέτου R. Τα δεδομένα αποτελούν τμήμα του πακέτου “InsuranceData” στην R.

Πιο συγκεκριμένα, ο Πίνακας “dataOhlsson” αποτελείται από τις παρακάτω στήλες:

- **Agarald:** Η στήλη αυτή απεικονίζει την ηλικία του κάθε ασφαλισμένου, με τιμές από 0 έως 99 έτη
- **Kon:** Απεικονίζει μια κατηγορική μεταβλητή ανάλογα με την ηλικία του ασφαλισμένου, λαμβάνοντας τιμές “K” και “M”
- **Zon:** Αποτυπώνει την Γεωγραφική Ζώνη, με τιμές από 1 έως 7, οι οποίες αποτελούν διαμέριση όλων των συνοικιών της Σουηδίας
- **Mcklass:** Αποτελεί μεταβλητή που αναπαριστά την Κλάση του Οχήματος, συναρτήσει του EV Ratio ( $\text{Engine Power in kW} \times 100 / (\text{Vehicle weight in kg} + 75)$ ), στρογγυλοποιημένο στην πλησιέστερο ακέραιο. Τα 75 κιλά αναπαριστούν το μέσο βάρος του οδηγού. Τα EV Ratios χωρίζονται σε 7 κλάσεις
- **Fordald:** Αποτυπώνει την ηλικία του οχήματος, με εύρος από 0 έως και 99 έτη
- **Bonuskl:** Bonus class, με τιμές από το 1 έως το 7. Ένας νέος οδηγός ξεκινάει από την τιμή 1, και για κάθε έτος στο οποίο δεν διαμορφώνονται απαιτήσεις (claims), η bonus class αυξάνεται κατά 1. Ύστερα από την πρώτη εμφάνιση απαίτησης, η bonus class μειώνεται κατά 2. Ο οδηγός δεν μπορεί να επιστρέψει στην bonus class 7 με λιγότερα από 6 συνεχόμενα έτη χωρίς δημιουργία απαιτήσεων
- **Duration:** Ο αριθμός των ετών κατά τα οποία το ασφαλιστήριο συμβόλαιο βρίσκεται εν ισχύ.
- **Antskad:** Το πλήθος των απαιτήσεων ανά ασφαλιστήριο συμβόλαιο
- **Skadkost:** Το συνολικό κόστος των απαιτήσεων ανά ασφαλιστήριο συμβόλαιο

Θα προχωρήσουμε στην εκτίμηση των Παραμέτρων μέσω των μεθόδων που περιεγράφηκαν παραπάνω, για την κατανομή Weibull.

Με την βοήθεια των πακέτων “Fitdistrplus” και “Tidyverse”, γράφουμε τον παρακάτω κώδικα για να προσαρμόσουμε την κατανομή στο σύνολο των απαιτήσεων των δεδομένων μας με την μέθοδο της **Μέγιστης Πιθανοφάνειας** (στήλη “Skadkost”):

```
#Maximum Likelihood Estimation  
library(fitdistrplus)
```

```

library(tidyverse)
library(insuranceData)
data("dataOhlsson")
skadkost_nonzero<-filter(dataOhlsson,skadkost>0)
Wei.MLE<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mle")
summary(Wei.MLE)

```

Ο παραπάνω κώδικας μας δίνει τις εξής τιμές για τις παραμέτρους της κατανομής, όσων αφορά την προσαρμογή της Weibull στο σύνολο των απαιτήσεων των δεδομένων μας (στήλη “Skadkost”):

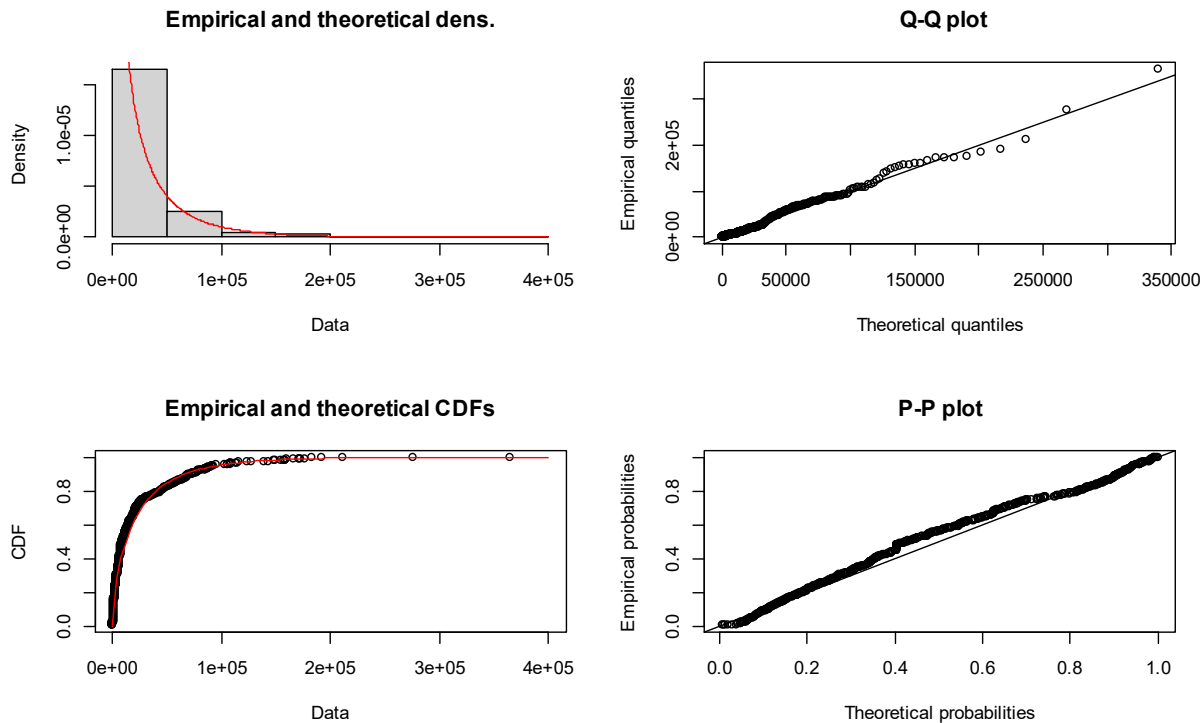
```

## Fitting of the distribution ' weibull ' by maximum likelihood
## Parameters :
##      estimate  Std. Error
## shape 7.026427e-01 2.048285e-02
## scale 2.043775e+04 1.089285e+03
## Loglikelihood: -7377.065  AIC: 14758.13  BIC: 14767.14
## Correlation matrix:
##      shape  scale
## shape 1.0000000 0.2708284
## scale 0.2708284 1.0000000

```

Διαγραμματικά, λαμβάνουμε την εξής πληροφορία για την εκτίμηση με την μέθοδο της Μέγιστης Πιθανοφάνειας:

plot(Wei.MLE)



Σχήμα 23: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας

Οι διάφορες μορφές γραφικής αναπαράστασης της Πυκνότητας των δεδομένων μας με βάση τις προσαρμοσμένες παραμέτρους (Συνάρτηση Πυκνότητας Πιθανότητας, Συνάρτηση Κατανομής, Q-Q Plot και P-P Plot) καταδεικνύουν μια καλή προσαρμογή πάνω στις θεωρητικές Καμπύλες και Ευθείες της Κατανομής, υποδηλώνοντας έτσι μια αρκετά καλή ερμηνεία των δεδομένων μας.

Προχωρώντας στην εκτίμηση των παραμέτρων με την **Μέθοδο των Ροπών**, έχουμε τον εξής κώδικα:

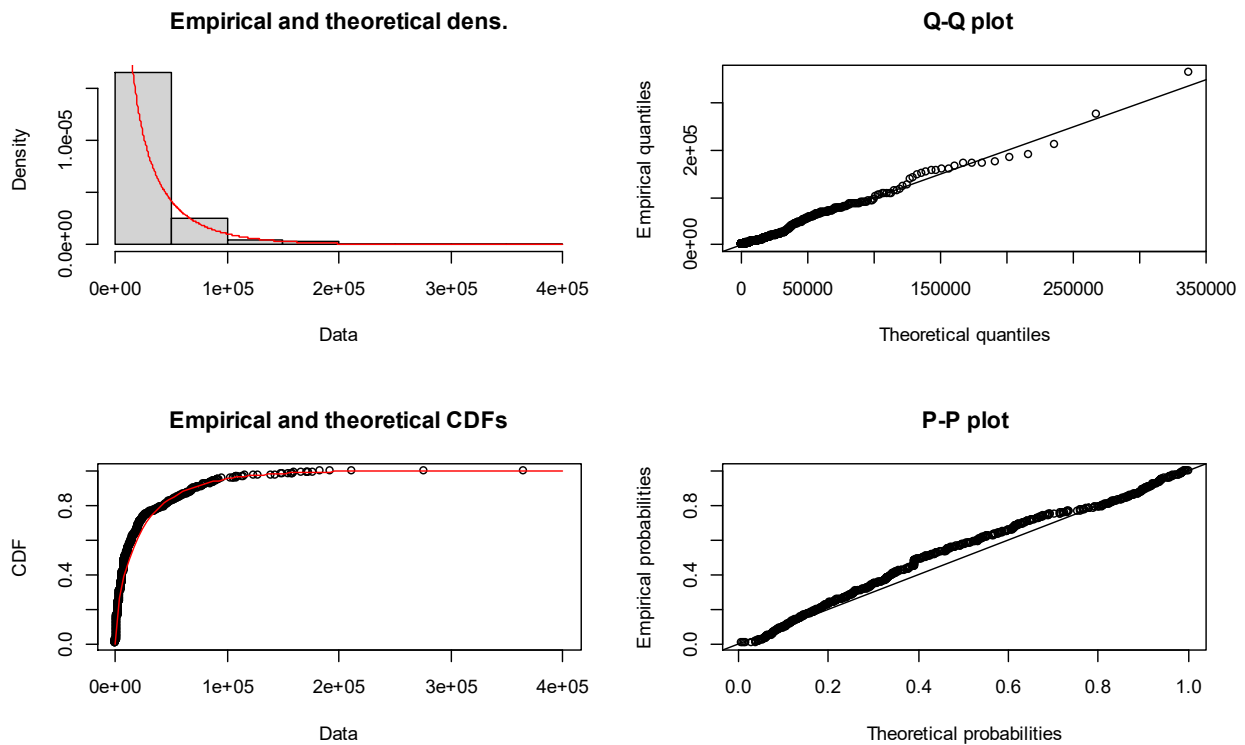
```
#Moment Matching Estimation  
data("dataOhlsson")  
skadkost_nonzero<-filter(dataOhlsson,skadkost>0)  
moment_function<-function(x,order) mean(x^order) #Function to calculate empirical moments - must be defined for Weibull  
Wei.MME<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mme",order=1:2,memp=moment_f
```

```
unction)
summary(Wei.MME)
```

Ο οποίος με την σειρά του μας υπολογίζει τις εκτιμώμενες παραμέτρους της Weibull ως εξής:

```
## Fitting of the distribution 'weibull' by matching moments
## Parameters :
## estimate
## shape 7.160188e-01
## scale 2.138821e+04
## Loglikelihood: -7377.969 AIC: 14759.94 BIC: 14768.95
```

Και γραφικά έχουμε ότι:



Σχήμα 24: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου των Ροπών

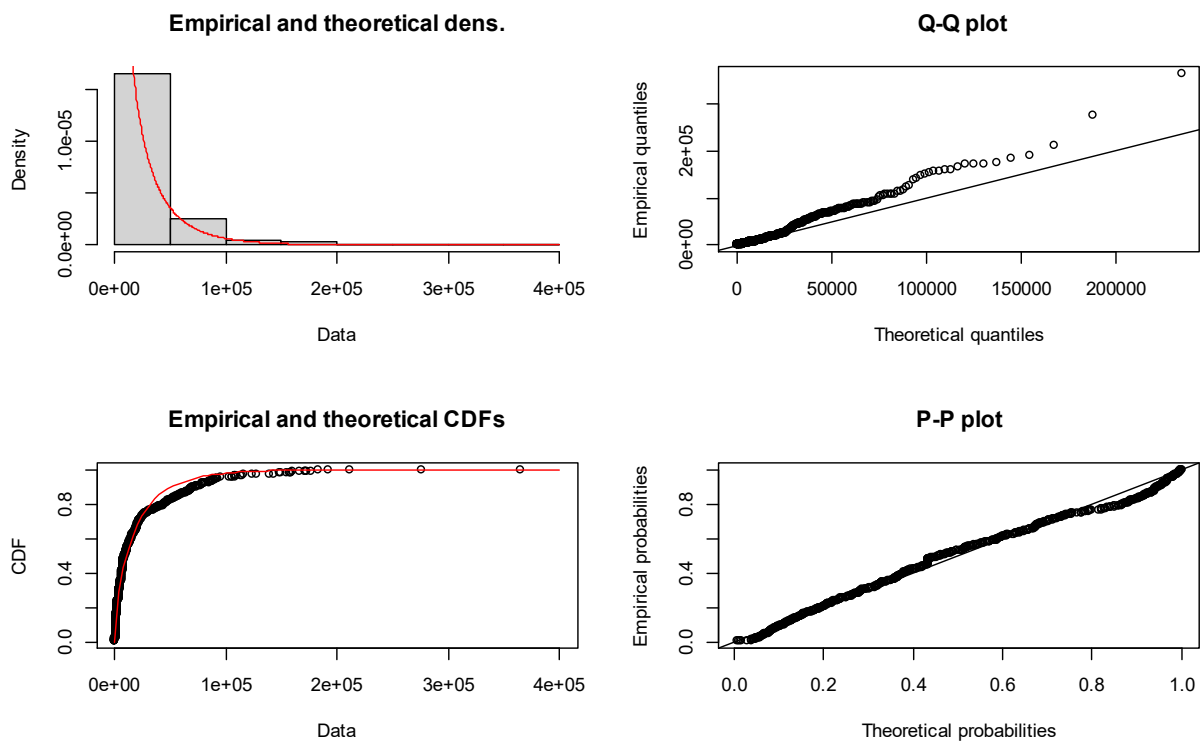
Στα διαγράμματα παρατηρείται όμοια συμπεριφορά στην προσαρμογή θεωρητικών και εμπειρικών δεδομένων με την μέθοδο της Μέγιστης Πιθανοφάνειας.

Ενδεικτικά, παρατίθεται και ο κώδικας για τις μεθόδους της **Ελάχιστης Απόστασης (Με την διαδικασία Anderson-Darling)** και των **Ποσοστιαίων Σημείων**, με τα αντίστοιχα αριθμητικά αποτελέσματα για την κάθε μέθοδο.

```
#Minimum Distance Estimation
data("dataOhlsson")
skadkost_nonzero<-filter(dataOhlsson,skadkost>0)
Wei.MGE<-fitdist(skadkost_nonzero$skadkost,"weibull",method="mge",gof="ADL")
summary(Wei.MGE)

## Fitting of the distribution ' weibull ' by maximum goodness-of-fit
## Parameters :
##      estimate
## shape 7.530278e-01
## scale 1.705808e+04
## Loglikelihood: -7387.632  AIC: 14779.26  BIC: 14788.28

plot(Wei.MGE)
```



Σχήμα 25: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Ελάχιστης Απόστασης

```
#Matching Quantile Estimation
data("dataOhlsson")
```

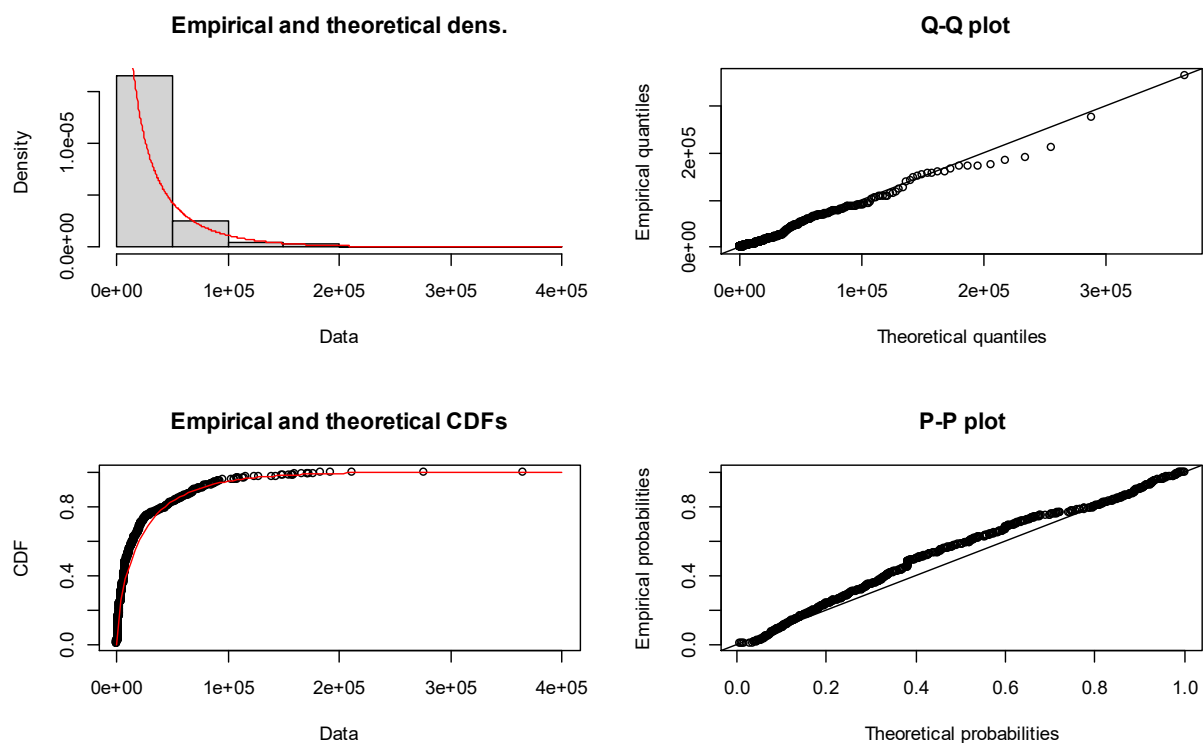
```

skadkost_nonzero<-filter(dataOhlsson,skadkost>0)
Wei.QME<-fitdist(skadkost_nonzero$skadkost,"weibull",method="qme",probs=c(8/10,9/10)) #Function to calculate empirical quantiles - must be defined for Weibull
summary(Wei.QME)

## Fitting of the distribution ' weibull ' by matching quantiles
## Parameters :
##      estimate
## shape 7.099382e-01
## scale 2.254835e+04
## Loglikelihood: -7379.589  AIC: 14763.18  BIC: 14772.19

plot(Wei.QME)

```



Σχήμα 26: Γραφικές Αναπαραστάσεις Πυκνότητας των δεδομένων με προσαρμοσμένες τις Παραμέτρους της Κατανομής Weibull μέσω της Μεθόδου Ποσοστημορίων

Παρατηρώντας τα εξαγόμενα αριθμητικά και σχηματικά αποτελέσματα από τις 4 διαφορετικές μεθοδολογίες, συμπεραίνουμε ότι η καταλληλότερη μεθοδολογία εκτίμησης των παραμέτρων, έτσι ώστε αυτές να είναι βέλτιστα προσαρμοσμένες στα δεδομένα μας, είναι αυτή της **Μέγιστης Πιθανοφάνειας**.



Βασικό κριτήριο στο συμπέρασμα αυτό αποτέλεσαν μεταξύ άλλων, και τα κριτήρια πληροφορίας AIC και BIC τα οποία θα αναλύσουμε εκτενώς στην επόμενη εφαρμογή της Παραγράφου 4.3.

#### 4.3 Έλεγχος Καλής Προσαρμογής Κατανομής σε Πραγματικά Δεδομένα με βάση το Κριτήριο Πληροφορίας

Εστιάζοντας στην κατάλληλη επιλογή ενός στατιστικού μοντέλου με σκοπό την βέλτιστη προσαρμογή στα δεδομένα μας, αξίζει να αναφερθεί ότι αυτό συνήθως επιλέγεται με βάση μια **τιμή κριτηρίου πληροφορίας**. Από τις ποικίλες μεθόδους που εφαρμόζονται για την εξαγωγή της τιμής αυτής, παραθέτουμε τις δύο σημαντικότερες (Burnham και Anderson (2004)):

##### A. Κριτήριο Πληροφορίας Akaike (AIC)

Το κριτήριο πληροφορίας Akaike εισήχθηκε για πρώτη φορά από τον Akaike περί το 1974 και φέρει τιμητικά το όνομα του Στατιστικού Hirotugu Akaike. Το συγκεκριμένο κριτήριο λαμβάνει χώρα σε πληθώρα προβλημάτων γραμμικής παλινδρόμησης, σε αναλύσεις χρόνων επιβίωσης και γενικότερα σε συγκρίσεις μοντέλων σε επίπεδο παραμέτρων.

Ο τύπος υπολογισμού του AIC έχει την παρακάτω μορφή:

$$AIC = -2\log L(\hat{\theta}) + 2 \dim(\theta) \quad (4.3.1)$$

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

Από τον έλεγχο καλής προσαρμογής για τα διάφορα μοντέλα, αυτό με την μικρότερη τιμή του AIC προσεγγίζει περισσότερο τα εμπειρικά δεδομένα, και επομένως παρουσιάζει την βέλτιστη προσαρμογή και επεξηγηματικότητα.

##### B. Μπεϋζιανό Κριτήριο Πληροφορίας (BIC)

Το Μπεϋζιανό κριτήριο πληροφορίας ή κριτήριο του Schwarz, τιμής ένεκεν για την ανάπτυξή του από τον Schwarz περί το 1978, αποτελεί ένα παράγωγο της Μπεϋζιανής Στατιστικής.

Για τον τύπο υπολογισμού του BIC, έχουμε το εξής:

$$BIC = -2\log L(\hat{\theta}) + \log n * \dim(\theta) \quad (4.3.2)$$

όπου  $L(\hat{\theta})$  είναι η μέγιστη τιμή πιθανοφάνειας,  $n$  το μέγεθος του δείγματος και  $\dim(\theta)$  η διάσταση του διανύσματος των παραμέτρων  $\theta$ .

Για τον έλεγχο καλής προσαρμογής των μοντέλων, και όμοια με το κριτήριο AIC, το μοντέλο με την μικρότερη τιμή του BIC προσεγγίζει περισσότερο τα εμπειρικά δεδομένα, και επομένως παρουσιάζει την βέλτιστη προσαρμογή και επεξηγηματικότητα.

Στην εφαρμογή που ακολουθεί, χρησιμοποιούμε τις προαναφερθείσες μεθόδους για να μπορέσουμε να αποφανθούμε για την βέλτιστη κατανομή μεταξύ τριών κατανομών, συμπεριλαμβανομένης και της Weibull.

## **ΕΦΑΡΜΟΓΗ 4.2**

Χρησιμοποιώντας τα ίδια δεδομένα με αυτά που χρησιμοποιήθηκαν στην Εφαρμογή της Παραγράφου 4.2, δηλαδή αυτά του πίνακα “dataOhlsson”, ο οποίος περιέχει τα κόστη αποζημιώσεων σε Ευρώ της Σουηδικής Ασφαλιστικής Εταιρείας Wasa για το χρονικό διάστημα 1994-1998, θα προχωρήσουμε στην αναζήτηση της βέλτιστης κατανομής αναφορικά με τα δεδομένα μας.

Θα εξετάσουμε τις εξής κατανομές:

- Ομοιόμορφη
- Κανονική
- Weibull

Χρησιμοποιώντας τη μέθοδο Μέγιστης Πιθανοφάνειας για την εκτίμηση των παραμέτρων, και με την βοήθεια του πακέτου “Fitdistrplus” εξάγουμε την παρακάτω πληροφορία προσαρμόζοντας καθεμία από τις παραπάνω κατανομές στο σύνολο των παρατηρήσεων που αφορούν τις απαιτήσεις των ασφαλιστήριων συμβολαίων:

i) Κανονική Κατανομή

	Εκτίμηση	Τυπικό Σφάλμα
Μέση Τιμή	25435.55	1482.910
Διακύμανση	38510.64	1048.576

Λογαριθμική Πιθανοφάνεια	-8025.011
--------------------------	-----------

ii) Ομοιόμορφη Κατανομή

	Εκτίμηση	Τυπικό Σφάλμα
Μέση Τιμή	16	-
Διακύμανση	365347	-

Λογαριθμική Πιθανοφάνεια	-8581.735
--------------------------	-----------

iii) Κατανομή Weibull

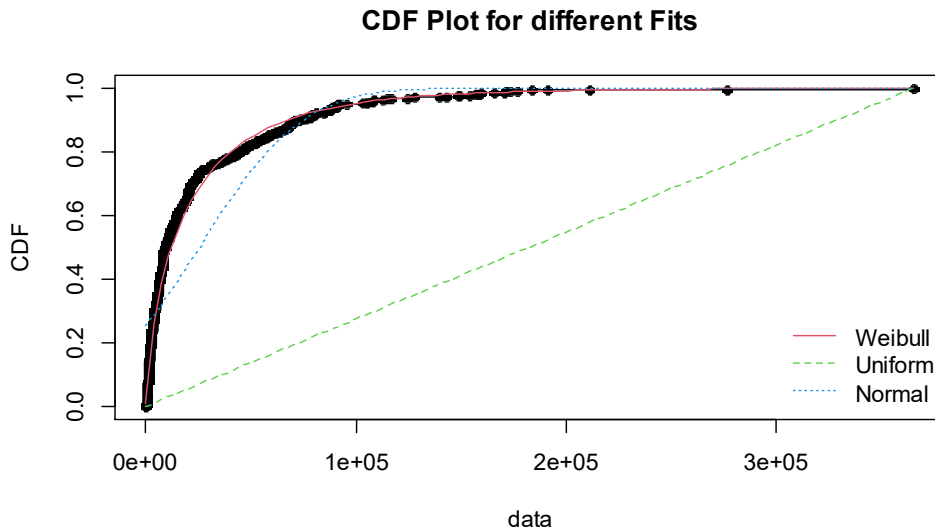
	Εκτίμηση	Τυπικό Σφάλμα
Σχήμα	7.026427e-01	2.048285e-02
Κλίμακα	2.043775e+04	1.089285e+03

Λογαριθμική Πιθανοφάνεια	-7377.065
--------------------------	-----------

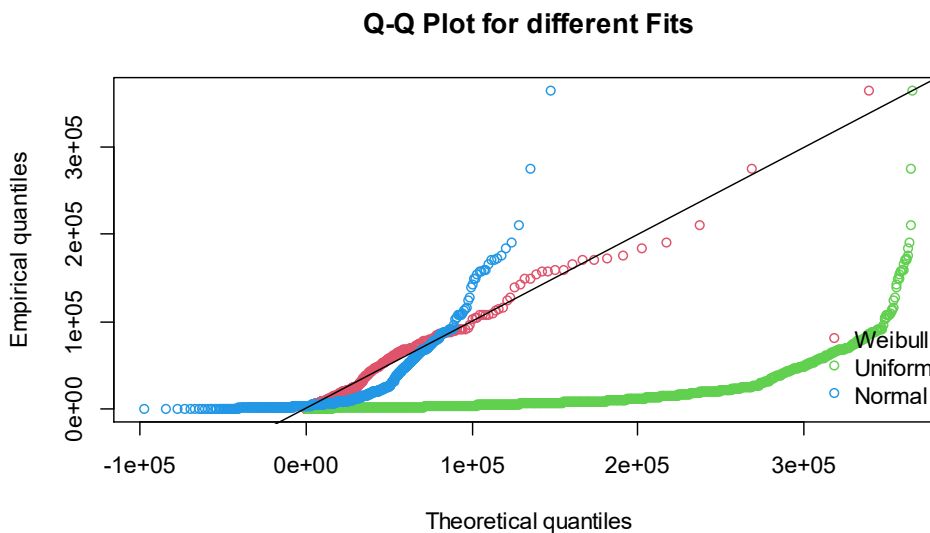
Ακόμη, προχωρώντας στις τιμές των κριτηρίων πληροφορίας AIC και BIC, αυτές παρουσιάζονται στον παρακάτω πίνακα:

	Weibull	Uniform	Normal
Akaike's Information Criterion (AIC)	14758.13	17167.47	16054.02
Bayesian Information Criterion (BIC)	14767.14	17176.48	16063.04

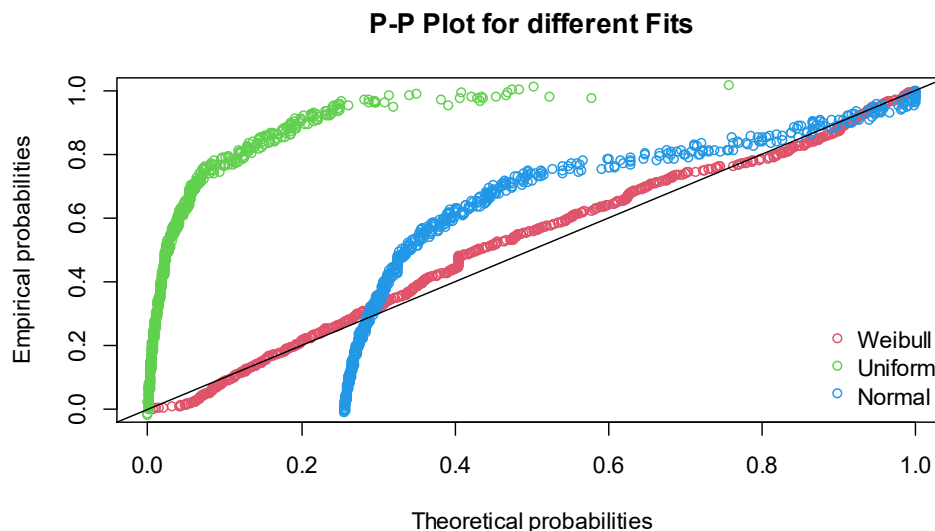
Επιπλέον, στα επόμενα σχήματα παρατίθενται και οι διαγραμματικές συγκρίσεις μεταξύ των 3 κατανομών και την προσαρμογή τους στα δεδομένα μας:



Σχήμα 27: Γραφικές Αναπαραστάσεις Συνάρτησης Κατανομής των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας



Σχήμα 28: Q-Q Plots των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας



Σχήμα 29: P-P Plots των εξεταζόμενων κατανομών συγκριτικά με τα πραγματικά δεδομένα, μέσω της Μεθόδου Μέγιστης Πιθανοφάνειας

Συμπερασματικά, παρατηρώντας τόσο τις τιμές των δεικτών AIC και BIC, καθώς και τις γραφικές απεικονίσεις των κατανομών, αποφαινόμεστε ότι η βέλτιστη κατανομή από τις εξεταζόμενες είναι η **κατανομή Weibull**.

#### 4.4 Εκτίμηση Μέσου Υπολειπόμενου Χρόνου Ζωής Ενεργών Ασφαλιστηρίων Συμβολαίων Χαρτοφυλακίου

Μια ακόμη εφαρμογή που θα αποπειραθούμε να παρουσιάσουμε στην συγκεκριμένη παράγραφο, αφορά την εκτίμηση του Μέσου Υπολειπόμενου Χρόνου Ζωής (MRL) ενός χαρτοφυλακίου ασφαλιστηρίων συμβολαίων, με βάση την χρονική διάρκεια την οποία τα συμβόλαια είναι ήδη ενεργά.

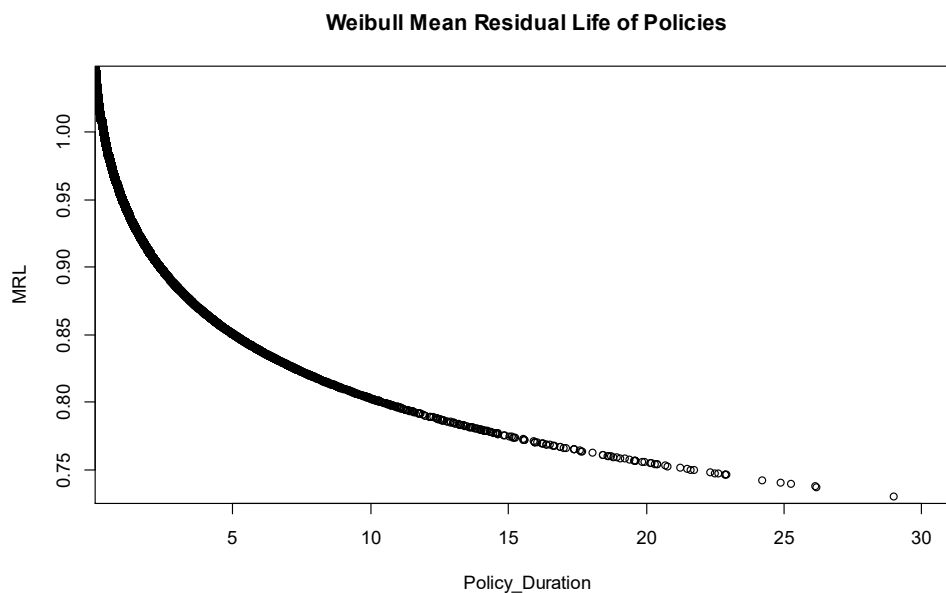
#### **ΕΦΑΡΜΟΓΗ 4.3**

Πιο συγκεκριμένα, θα επιχειρηθεί η εκτίμηση του Μέσου Υπολειπόμενου Χρόνου Ζωής των Ασφαλιστηρίων Συμβολαίων που περιέχονται στο σετ δεδομένων που χρησιμοποιήθηκε και στις προηγούμενες δύο παραγράφους του Κεφαλαίου 4, αυτή τη φορά στηριζόμενοι στην πληροφορία της χρονικής διάρκειας (σε έτη) ισχύος των ήδη ενεργών συμβολαίων. Η πληροφορία αυτή περιέχεται στην στήλη “Duration” του πίνακα “dataOhlsson”.

Αρχικά, και χρησιμοποιώντας το πακέτο “Fitdistrplus”, προχωρούμε στην βέλτιστη εκτίμηση των παραμέτρων της κατανομής Weibull για τα δεδομένα μας, η οποία παρουσιάζει τα ακόλουθα αποτελέσματα:

	Εκτίμηση	Τυπικό Σφάλμα
Σχήμα	1.092825	0.002969500
Κλίμακα	1.085742	0.004205318

Στη συνέχεια, χρησιμοποιώντας την εντολή “residlife”, η οποία περιέχεται στο Στατιστικό Πακέτο “reslife” της R, και η οποία υπολογίζει, μεταξύ άλλων, το MRL για ένα πλήθος κατανομών που προσδιορίζονται από τον χρήστη, προχωράμε στην εκτίμηση του MRL, χρησιμοποιώντας την κατανομή Weibull και ορίζοντας ως παραμέτρους Σχήματος και Κλίμακας τις παραπάνω, ο Μέσος Υπολειπόμενος Χρόνος Ζωής των δεδομένων μας λαμβάνει την ακόλουθη σχηματική απεικόνιση:



Σχήμα 30: Γραφική Αναπαράσταση Μέσου Υπολειπόμενου Χρόνου Ζωής Χαρτοφυλακίου μέσω της κατανομής Weibull

Τέλος, παρακάτω δίνονται και ορισμένες επιπλέον πληροφορίες που προκύπτουν από τον υπολογισμό του MRL, μέσω της εντολής “summary” του πακέτου της R:

```
summary(MRL)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.7256	0.9493	0.9551	0.9604	0.9821	1.0487

Από τα παραπάνω συμπεραίνουμε ότι, λόγω της τιμής της παραμέτρου σχήματος  $c$ , το MRL της στήλης “Duration” (διάρκεια ισχύος εξεταζόμενων ασφαλιστηρίων συμβολαίων) παρουσιάζει φθίνουσα συμπεριφορά.

## Σύνοψη - Συμπεράσματα

Η επιλογή της καταλληλότερης κατανομής, ικανής να επεξηγήσει τα δεδομένα που έχουμε στην διάθεσή μας κάθε φορά, αποτελεί απόφαση βαρύνουσας σημασίας στον Αναλογισμό, αλλά και την επιστήμη της Στατιστικής και των Πιθανοτήτων γενικότερα.

Ως πρόταση στην παραπάνω πρόκληση, η κατανομή Weibull παρουσιάζει μια ευρύτατη εφαρμογή σε ποικίλους τομείς όπως αυτών του Αναλογισμού, της Διαχείρισης Κινδύνων και της Θεωρίας Ακραίων Τιμών.

Στην συγκεκριμένη Διπλωματική Εργασία, και παράλληλα με την θεωρητική προσέγγιση, παρουσιάστηκαν ορισμένες εφαρμογές της Weibull, ούτως ώστε να κατανοηθεί περαιτέρω η χρησιμότητά της ως Ζημιοκατανομή στον κλάδο των Ασφαλίσεων.

Από τις Εφαρμογές που έλαβαν χώρα στο Κεφάλαιο 4, παρουσιάστηκαν διάφορες μεθοδολογίες εκτίμησης των παραμέτρων της κατανομής, ενώ πραγματοποιήθηκε και έλεγχος καλής προσαρμογής με βάση τις εκτιμώμενες τιμές των παραμέτρων στα δεδομένα μας, διαπιστώνοντας μια καλή εφαρμογή.

Επίσης, παρατηρήθηκε ότι η κατανομή Weibull αποτελεί μια αρκετά καλή κατανομή για την ανάλυση του Χαρτοφυλακίου που χρησιμοποιήσαμε ως δεδομένα έναντι άλλων κατανομών, σύμφωνα με τις τιμές κριτηρίου πληροφορίας AIC και BIC.

Παράλληλα, η κατανομή Weibull χρησιμοποιήθηκε και για την εκτίμηση της μέσης υπολειπόμενης διάρκειας ζωής του χαρτοφυλακίου μας, με συμπεριφορά η οποία είχε άμεση συσχέτιση με την τιμή της παραμέτρου σχήματος  $c$ .

Επομένως, η κατανομή Weibull θα πρέπει να συμπεριλαμβάνεται ως διαθέσιμη επιλογή μεταξύ των υποψήφιων κατανομών που ένας Αναλογιστής ή Διαχειριστής Κινδύνων δύναται να χρησιμοποιήσει στις αναλύσεις του.



## ΠΑΡΑΡΤΗΜΑ

Για τα αριθμητικά και γραφικά αποτελέσματα που παρουσιάστηκαν στα προηγούμενα Κεφάλαια της Διπλωματικής Εργασίας, χρησιμοποιήθηκε η γλώσσα προγραμματισμού R.

Παρακάτω δίνονται ανά Κεφάλαιο και με την σειρά που τα αποτελέσματα παρουσιάστηκαν, οι εντολές που χρησιμοποιήθηκαν:

### Κεφάλαιο 1

```
#Plot different Weibull Density Distribution Functions, while only changing Location Parameter
f<-function(x,a=0,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
g<-function(x,a=1,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
h<-function(x,a=1.5,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
t<-function(x,a=2,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
curve(f,from = 0,to = 6.0,col="black",main="Density Plots of 3-Parameter Weibull Distribution by different Location Parameter",
      xlab = "x",ylab = "f(x|a,1,2)",xaxs="i",yaxs="i",ylim=c(0,1.0))
curve(g,from = 0,to = 6.0,col="green",add = TRUE)
curve(h,from = 0,to = 6.0,col="red",add = TRUE)
curve(t,from = 0,to = 6.0,col="blue",add = TRUE)
legend(5,1,legend=c("a=0","a=1","a=1.5","a=2"),col=c("black","green","red","blue"),lty = c(1,1,1,1))

#Plot different Weibull Density Distribution Functions, while only changing Scale Parameter
f<-function(x,a=0,b=0.5,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
g<-function(x,a=0,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
h<-function(x,a=0,b=1.5,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
t<-function(x,a=0,b=2,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
curve(f,from = 0,to = 5.0,col="black",main="Density Plots of 3-Parameter Weibull Distribution by different Scale Parameter",
      xlab = "x",ylab = "f(x|0,b,2)",xaxs="i",yaxs="i",ylim=c(0,1.8),xlim = c(0.0,5.0))
curve(g,from = 0,to = 5.0,col="green",add = TRUE)
curve(h,from = 0,to = 5.0,col="red",add = TRUE)
curve(t,from = 0,to = 5.0,col="blue",add = TRUE)
legend(4,1.8,legend=c("b=0.5","b=1","b=1.5","b=2"),col=c("black","green","red","blue"),lty = c(1,1,1,1))
```

*#Plot different Weibull Density Distribution Functions, while only changing Shape Parameter*

```
f<-function(x,a=0,b=1,c=0.5)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
g<-function(x,a=0,b=1,c=1)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
h<-function(x,a=0,b=1,c=2)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
t<-function(x,a=0,b=1,c=4)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
u<-function(x,a=0,b=1,c=6)(c/b*((x-a)/b)^(c-1)*exp(-((x-a)/b)^c))
curve(f,from = 0,to = 4.0,col="black",main="Density Plots of 3-Parameter Weibull Distribution by different Shape Parameter",
      xlab = "x",ylab = "f(x|0,1,c)",xaxs="i",yaxs="i",ylim=c(0,2.5),xlim = c(0.0,4.0))
curve(g,from = 0,to = 4.0,col="green",add = TRUE)
curve(h,from = 0,to = 4.0,col="red",add = TRUE)
curve(t,from = 0,to = 4.0,col="blue",add = TRUE)
curve(u,from = 0,to = 4.0,col="turquoise",add = TRUE)
legend(3,2.5,legend=c("c=0.5","c=1","c=2","c=4","c=6"),col=c("black","green","red","blue","turquoise"),lty = c(1,1,1,1,1))
```

*#Plot a Curve from the Weibull Life Distribution with 3-parameter Definition (a constant) as well as the Normalized Definition*

```
F<-function(x,a=0,b=1,c=5)(1-exp(-((x-a)/b)^c))
G<-function(x,a=0,b=2,c=5)(1-exp(-((x-a)/b)^c))
H<-function(x,a=0,b=3,c=5)(1-exp(-((x-a)/b)^c))
curve(F,from = 0.0,to = 5.0,col="black",main="Curve Plot of 3-Parameter CDF",
      xlab = "x",ylab = "F(x|0,b,5)",xaxs="i",yaxs="i",ylim=c(0,1.0))
curve(G,from = 0.0,to = 5.0,col="green",add = TRUE)
curve(H,from = 0.0,to = 5.0,col="red",add = TRUE)
abline(h=0.6321,col="darkgreen",lty=2,lwd=3)
legend(4,0.9,legend=c("b=1,c=5","b=2,c=5","b=3,c=5"),col=c("black","green","red"),lty = c(1,1,1))
mtext(0.6321,side = 2,at=0.6321)
```

```
F<-function(x,c=0.5)(1-exp(-x^c))
G<-function(x,c=1)(1-exp(-x^c))
H<-function(x,c=2)(1-exp(-x^c))
curve(F,from = 0,to = 5.0,col="black",main="Curve Plot of Normalized CDF",
      xlab = "x",ylab = "F(x)",xaxs="i",yaxs="i",ylim=c(0,1.0))
curve(G,from = 0.0,to = 5.0,col="green",add = TRUE)
curve(H,from = 0.0,to = 5.0,col="red",add = TRUE)
abline(h=0.6321,col="darkgreen",lty=2,lwd=3)
legend(4,0.9,legend=c("c=0.5","c=1","c=2"),col=c("black","green","red"),lty = c(1,1,1))
mtext(0.6321,side = 2,at=0.6321)
```

*#Plot a Curve from the Weibull Reliability Distribution with 3-parameter Definition (a constant) as well as the Normalized Definition*

```
F<-function(x,a=0,b=1,c=5)(1-exp(-((x-a)/b)^c))
R1<-function(x) (1-F(x))
G<-function(x,a=0,b=2,c=5)(1-exp(-((x-a)/b)^c))
R2<-function(x) (1-G(x))
H<-function(x,a=0,b=3,c=5)(1-exp(-((x-a)/b)^c))
R3<-function(x) (1-H(x))
curve(R1,from = 0,to = 5.0,col="black",main="Curve Plot of 3-Parameter Reliability Function",
      xlab = "x",ylab = "R(x|0,b,5)",xaxs="i",yaxs="i",ylim=c(0,1.0))
curve(R2,from = 0,to = 5.0,col="green",add = TRUE)
curve(R3,from = 0,to = 5.0,col="red",add = TRUE)
abline(h=0.3679,col="darkgreen",lty=2,lwd=3)
legend(4,0.9,legend=c("b=1,c=5","b=2,c=5","b=3,c=5"),col=c("black","green","red"),lty = c(1,1,1))
mtext(0.3679,side = 2,at=0.3679)
```

```
F<-function(x,c=0.5)(1-exp(-x^c))
R1<-function(x) (1-F(x))
G<-function(x,c=1)(1-exp(-x^c))
R2<-function(x) (1-G(x))
H<-function(x,c=2)(1-exp(-x^c))
R3<-function(x) (1-H(x))
curve(R1,from = 0,to = 5.0,col="black",main="Curve Plot of Normalized Reliability Function",
      xlab = "x",ylab = "R(x)",xaxs="i",yaxs="i",ylim=c(0,1.0))
curve(R2,from = 0,to = 5.0,col="green",add = TRUE)
curve(R3,from = 0,to = 5.0,col="red",add = TRUE)
abline(h=0.3679,col="darkgreen",lty=2,lwd=3)
legend(4,0.9,legend=c("c=0.5","c=1","c=2"),col=c("black","green","red"),lty = c(1,1,1))
mtext(0.3679,side = 2,at=0.3679)
```

*#Plot a Curve from the Weibull Normalized Hazard Rate for different values of c*

```
h1<-function(x,c=0.5)(c*x^(c-1))
h2<-function(x,c=1)(c*x^(c-1))
h3<-function(x,c=2)(c*x^(c-1))
h4<-function(x,c=5)(c*x^(c-1))
curve(h1,from = 0,to = 2.0,col="black",main="Curve Plot of Normalized Hazard Rate",
      xlab = "x",ylab = "h(x)",xaxs="i",yaxs="i",ylim=c(0,9.0))
curve(h2,from = 0,to = 2.0,col="green",add = TRUE)
```

```

curve(h3,from = 0,to = 2.0,col="red",add = TRUE)
curve(h4,from = 0,to = 2.0,col="blue",add = TRUE)
legend(1.5,8.0,legend=c("c=0.5","c=1","c=2","c=5"),col=c("black","green","red",
,"blue"),lty = c(1,1,1))

```

*#Plot a Curve from the Weibull Normalized Cumulative Hazard Rate for different values of c*

```

H1<-function(x,c=0.5)(x^(c))
H2<-function(x,c=1)(x^(c))
H3<-function(x,c=2)(x^(c))
H4<-function(x,c=5)(x^(c))
curve(H1,from = 0,to = 4.0,col="black",main="Curve Plot of Normalized Cumulative Hazard Rate",
      xlab = "x",ylab = "H(x)",xaxs="i",yaxs="i",ylim = c(0,2))
curve(H2,from = 0,to = 4.0,col="green",add = TRUE)
curve(H3,from = 0,to = 4.0,col="red",add = TRUE)
curve(H4,from = 0,to = 4.0,col="blue",add = TRUE)
legend(3,1.5,legend=c("c=0.5","c=1","c=2","c=5"),col=c("black","green","red",
"blue"),lty = c(1,1,1))

```

*#Plot a Curve of the Weibull Mean for the 2-Parameter Definition*

```

E1<-function(c,b=1)(b*gamma(1+1/c))
E2<-function(c,b=0.5)(b*gamma(1+1/c))
E3<-function(c,b=2)(b*gamma(1+1/c))
curve(E1,from = 0,to = 4.0,col="black",main="Mean of 2-Parameter Weibull Distribution",
      xlab = "c",ylab = "E(x)",xaxs="i",yaxs="i",ylim = c(0,20))
curve(E2,from = 0,to = 4.0,col="green",add = TRUE)
curve(E3,from = 0,to = 4.0,col="red",add = TRUE)
legend(3,20,legend=c("b=1","b=0.5","b=2"),col=c("black","green","red"),lty =
c(1,1,1))

```

*#Plot a Curve of the Weibull Variance for the 2-Parameter Definition*

```

Var1<-function(c,b=1)(b^2*((gamma(1+2/c)-gamma(1+1/c)^2)))
Var2<-function(c,b=0.5)(b^2*((gamma(1+2/c)-gamma(1+1/c)^2)))
Var3<-function(c,b=2)(b^2*((gamma(1+2/c)-gamma(1+1/c)^2)))
curve(Var1,from = 0,to = 4.0,col="black",main="Variance of 2-Parameter Weibull Distribution",
      xlab = "c",ylab = "Var(x)",xaxs="i",yaxs="i",ylim = c(0,20))
curve(Var2,from = 0,to = 4.0,col="green",add = TRUE)
curve(Var3,from = 0,to = 4.0,col="red",add = TRUE)
legend(3,20,legend=c("b=1","b=0.5","b=2"),col=c("black","green","red"),lty =
c(1,1,1))

```

```

#Plot the Mean Residual Life of Weibull Distribution, for various values of c
library(reslife)

population<-rweibull(10000,1,1)
#Get Mean Residual Life Values with  $0 < C < 1$ 
MRL_inc<-residlife(population,distribution="weibull",parameters = c(shape=0.5
, scale=1),type = "mean")
#Get Mean Residual Life Values with  $C=1$ 
MRL_cons<-residlife(population,distribution="weibull",parameters = c(shape=1,
scale=1),type = "mean")
#Get Mean Residual Life Values with  $C > 1$ 
MRL_dec<-residlife(population,distribution="weibull",parameters = c(shape=2,s
cale=1),type = "mean")

#Plot the Graphs
plot(population,MRL_inc,xaxs="i",yaxs="i",main = "Weibull Mean Residual Life
for  $0 < c < 1$ ")

plot(population,MRL_cons,xaxs="i",yaxs="i",main = "Weibull Mean Residual Life
for  $c=1$ ")

plot(population,MRL_dec,xaxs="i",yaxs="i",main = "Weibull Mean Residual Life
for  $c > 1$ ")

```

## Κεφάλαιο 2

```

#Plot the Moment Generating Function of Weibull Distribution, for  $c > 1$ 
#MGF Function
MGF1<- function(x,b=1,c=1.5,n) {
  suma<-0
  for(n in 0:150) {
suma<-suma+(((x^n)*(b^n)/factorial(n))*gamma(1+n/c))
  }
return(suma)
}
MGF2<- function(x,b=1,c=3,n) {
  suma<-0
  for(n in 0:150) {
  suma<-suma+(((x^n)*(b^n)/factorial(n))*gamma(1+n/c))
  }
return(suma)
}
MGF3<- function(x,b=1,c=5,n) {
  suma<-0
  for(n in 0:150) {
  suma<-suma+(((x^n)*(b^n)/factorial(n))*gamma(1+n/c))
  }
}

```

```

    return(suma)
}
curve(MGF1(x),from = 0,ylim = c(0,100),xlim = c(0,5),xaxs = "i", yaxs = "i",
xlab = "x", ylab = "MGF(x|1,c)",
    main = "Curve Plot of Moment Generating Function for Weibull Density with
h c>1",col="black",lty = 1,lwd = 2)
curve(MGF2(x),from = 0,lty = 2,lwd = 2,col = "darkgreen", add = TRUE)
curve(MGF3(x),from = 0,lty = 3,lwd = 2,col = "darkred", add = TRUE)
legend(4,50,legend=c("c=1.5","c=3","c=5"),col=c("black","darkgreen","darkred"
),lty = c(1,2,3),lwd=c(2,2,2))

# Plot Distributions with different tail types
library(ordinal)

F<-function(x)(1-pexp(x, rate = 1)) #Exponential
G<-function(x,a=1)(1-exp((-x)^(-a))) #Frechet
H<-function(x)(1-pgumbel(x,location = 0, scale = 1)) #Gumbel
curve(F,from = 0,to = 10,col="black",main="Curve Plot of Distributions with d
ifferent Tail types",
    xlab = "x",ylab = "1-F(x)",xaxs="i",yaxs="i",ylim = c(0,1),xlim = c(0,1
0))
curve(G,from = 0,to = 10,col="green",lty = 1, add = TRUE)
curve(H,from = 0, to = 10,col="red",lty = 1, add = TRUE)
legend(6,0.8,legend=c("Exponential (rate = 1)","Frechet (shape = 1)","Gumbel
(location = 0, scale = 1)"),col=c("black","green","red"),lty = c(1,1,1),lwd=c
(1,1,1))

# Plot Tail Function of Weibull Distribution, for different values of shape p
arameter c
F1<-function(x)(1-pweibull(x,scale = 1,shape = 0.5))
F2<-function(x)(1-pweibull(x,scale = 1,shape = 1))
F3<-function(x)(1-pweibull(x,scale = 1,shape = 3))
curve(F1,from = 0,to = 10,col="black",main="Tail Plot of Weibull Distribution
for different values of c",
    xlab = "x",ylab = "1-F(x|1,c)",xaxs="i",yaxs="i",ylim = c(0,1),xlim = c
(0,10))
curve(F2,from = 0,to = 10,col="green",lty = 1, add = TRUE)
curve(F3,from = 0, to = 10,col="red",lty = 1, add = TRUE)
legend(6,0.8,legend=c("c=0.5","c=1","c=3"),col=c("black","green","red"),lty =
c(1,1,1),lwd=c(1,1,1))
arrows(x0 = 2.162046, y0 = 0.6, x1 = 2.162046, y1 = 0.1144649,col = "orange",
lwd = 2,lty = 1)
text(x = 2.162046, y = 0.4, label = "Exponential", srt = 90, font = 4 )
locator()

```

### Κεφάλαιο 3

```
library(fitdistrplus)

library(actuar)

library(tidyverse)

library(weibullness)

library(insuranceData)

#Plot the Failure Density of Weibull Distribution
f<-function(x,b=1,c=7) ((c/b)*(x/b)^(c-1)*exp(-(x/b)^c))
curve(f,from=0,to=3.5,xlab = "x",ylab = "f(x)",xaxs="i",yaxs="i")

#Plot a Curve from the Weibull Life Distribution
F<-function(x)(pweibull(x, scale = 1,shape = 5))
curve(F,from = 0,to = 2.5,col="red",main="Curve Plot of Life Distribution",
      xlab = "x",ylab = "F(x)",xaxs="i",yaxs="i",ylim=c(0,1.0))

#Plot a Step Line from the Weibull Life Distribution
curve(F,type="s",from = 0,to = 2.5,col="red",main="Step Line Plot of Life Dis
tribution",
      xlab = "x",ylab = "F(x)",xaxs="i",yaxs="i",ylim=c(0,1.0))

#Plot a Curve from the Weibull Survival Function
S<-function(x)(1-pweibull(x, scale = 1,shape = 5))
curve(S,from = 0,to = 10,col="green",main="Curve Line Plot of Survival Functi
on",
      xlab = "x",ylab = "S(x)",xaxs="i",yaxs="i",xlim=c(0,2.0))

#Plot a Step Line from the Weibull Survival Function
curve(S,type="s",from = 0,to = 10,col="green",main="Step Line Plot of Surviva
l Function",
      xlab = "x",ylab = "S(x)",xaxs="i",yaxs="i",xlim=c(0,2.0))
```

```

#Plot all Possible Curves Resulting from Force of Mortality Behaviour - Aging
Analysis
h1<-function(x) (4/x)
h2<-function(x) (0*x+4)
h3<-function(x) (x^4)
curve(h1,from = 0, to = 10,col="blue",main="Curve Plots for different Cases o
f Force of Mortality",
      xlab="x",ylab = "h(x)",xaxs="i",yaxs="i",xlim = c(0,10),ylim=c(0,30))
curve(h2,from = 0, to = 10,col="red",lty=1,add = TRUE)
curve(h3,from = 0, to = 10,col="green",lty=1,add = TRUE)
legend(3,30,legend = c("non-increasing","constant","non-decreasing"),
      col = c("blue","dark red","dark green"),lty = c(1,1,1))

#Plot Bathtub Curve of Weibull Distribution
h1<-function(x,c=0.1) (c*x^(c-1))
h2<-function(x,c=1) (c*x^(c-1))
h3<-function(x,c=10) (c*x^(c-1))
curve(h1,from=0,to=0.08435841,col="blue",main="Bathtub Curve",
      xlab="x",ylab = "h(x)",xaxs="i",yaxs="i",xlim = c(0,2),ylim=c(0,20),lwd
=2)
curve(h2,from=0.08435841,to= 0.7729101,col="red",lty=1,lwd=2,add = TRUE)
curve(h3,from=0.7729101,col="green",lty=1,lwd=2,add = TRUE)
legend(1.4,20,legend = c("infant mortality period","intrinsic failure period"
,"wearout failure period"),
      col = c("blue","dark red","dark green"),lty = c(1,1,1))
abline(h=1,col="darkgreen",lty=2,lwd=1)
mtext(1,side = 2,at=1)
locator()

```

## Κεφάλαιο 4

```

#Application in §4.2 - Assess the Best Estimate Parameters of Weibull Distrib
ution for Claims Data of
#Swedish insurance company Wasa, concerning Motorcycle Insurance

#Maximum Likelihood Estimation
library(fitdistrplus)
library(tidyverse)
library(insuranceData)
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
Wei.MLE<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mle")
summary(Wei.MLE)

plot(Wei.MLE)

```



```

#Moment Matching Estimation
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
moment_function<-function(x,order) mean(x^order) #Function to calculate empirical moments - must be defined for Weibull
Wei.MME<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mme",order=1:2,memp=moment_function)
summary(Wei.MME)

plot(Wei.MME)

#Minimum Distance Estimation
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
Wei.MGE<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mge",gof="ADL")
summary(Wei.MGE)

plot(Wei.MGE)

#Matching Quantile Estimation
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
Wei.QME<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "qme",probs=c(8/10,9/10)) #Function to calculate empirical quantiles - must be defined for Weibull
summary(Wei.QME)

plot(Wei.QME)

#Application in §4.3 - Assess the Goodness-of-fit amongst a variety of distributions for Claims Data of
#Swedish insurance company WASA, concerning Motorcycle Insurance, based on AIC and BIC

#Maximum Likelihood Estimation for Weibull Distribution
library(fitdistrplus)
library(tidyverse)
library(insuranceData)
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)

```

```

Wei.MLE<-fitdist(skadkost_nonzero$skadkost,"weibull",method = "mle")
summary(Wei.MLE)

#Maximum Likelihood Estimation for Uniform Distribution
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
Unif.MLE<-fitdist(skadkost_nonzero$skadkost,"unif",method = "mle")
summary(Unif.MLE)

#Maximum Likelihood Estimation for Normal Distribution
data("data0hlsson")
skadkost_nonzero<-filter(data0hlsson,skadkost>0)
Norm.MLE<-fitdist(skadkost_nonzero$skadkost,"norm",method = "mle")
summary(Norm.MLE)

#Compare above mentioned fits
cdfcomp(list(Wei.MLE,Unif.MLE,Norm.MLE),legendtext = c("Weibull","Uniform","Normal"),main = "CDF Plot for different Fits")

denscomp(list(Wei.MLE,Unif.MLE,Norm.MLE),legendtext = c("Weibull","Uniform","Normal"),main = "Density Plot for different Fits")

qqcomp(list(Wei.MLE,Unif.MLE,Norm.MLE),legendtext = c("Weibull","Uniform","Normal"),main = "Q-Q Plot for different Fits")

ppcomp(list(Wei.MLE,Unif.MLE,Norm.MLE),legendtext = c("Weibull","Uniform","Normal"),main = "P-P Plot for different Fits")

gofstat(list(Wei.MLE,Unif.MLE,Norm.MLE),fitnames = c("Weibull","Uniform","Normal"))

#Application in §4.4 - Calculate the Mean Residual Life of WASA Policies, based on their Duration
library(fitdistrplus)
library(tidyverse)
library(insuranceData)
library(reslife)
Policy_duration_nonzero<-filter(data0hlsson,duration>0)
Policy_Duration<-Policy_duration_nonzero$duration
#Find Best Estimate Parameters for Dataset, using Maximum Likelihood Estimation for Weibull Distribution

```

```
Policy_duration.MLE<-fitdist(Policy_Duration,"weibull",method = "mle")
summary(Policy_duration.MLE)

#Get Mean Residual Life Values per Insurance Policy
MRL<-residlife(Policy_Duration,distribution="weibull",parameters = c(shape=1.
092825,scale=1.085742),type = "mean")
summary(MRL)

#Plot the Graphs
plot(Policy_Duration,MRL,xaxs="i",yaxs="i",main = "Weibull Mean Residual Life
of Policies")
```

## ΒΙΒΛΙΟΓΡΑΦΙΑ

### Ελληνική

1. Αντζουλάκος, Δ. (2023). *Σημειώσεις Μαθήματος: «Γλώσσα Προγραμματισμού R με εφαρμογές στον Αναλογισμό»*, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς
2. Ζήμερας, Σ. (2020). *Διαφάνειες Μαθήματος: «Ανάλυση Επιβίωσης»*, Τμήμα Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών Μαθηματικών, Πανεπιστήμιο Αιγαίου
3. Μπούτσικας, Μ. (2003-8). *Σημειώσεις Μαθήματος: «Θεωρία Αξιοπιστίας»*, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς
4. Πολίτης, Κ. (2023). *Σημειώσεις Μαθήματος: «Θεωρία Κινδύνου II»*, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς

### Ξένη

1. Bingham, A., Spradlin, D. (2011). *The Long Tail of Expertise*. Pearson Education.
2. Bryson, M. (1974). *Heavy Tailed Distributions: Properties and Tests*. *Technometrics* 16(1)
3. Burnham, K. P., Anderson, D. R. (2004). "Multimodel inference: understanding AIC and BIC in Model Selection", *Sociological Methods & Research*, 33
4. Ihaka, R. (2003). *Notes of Course: "Statistics 120 – Information Visualization"*, Department of Statistics, University of Oakland
5. Jagriti, D., Dilip C. N. (2019). "Weibull Distribution as an Actuarial Risk Model: Computation of its probability of ultimate ruin and the moments of the time to ruin, deficit at ruin and surplus prior to ruin", *Journal of Data Science*, 17(1)
6. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2009). "Modern Actuarial Risk Theory Using R", *Second Edition*, Springer
7. Rinne, H. (2008). "The Weibull Distribution – A Handbook", Justus – Liebig University, Giessen, Germany
8. Swartz, G.B. (1973). *The mean residual lifetime function*; *IEEE Transactions on Reliability* 22, 108–109