CONTRIBUTIONS TO DISTRIBUTION
THEORY WITH APPLICATIONS

A thesis submitted to the University of Manchester
for the degree of Master of Philosophy
in the Faculty of The Faculty of Science and Engineering

2020

By
Juncheng Hou
Department of Mathematics, School of Natural Sciences
Contents

Abstract 8

Declaration 10

Copyright 11

Acknowledgements 13

1 Introduction 14

1.1 Aims and motivation 14

1.2 Future work 15

2 On the generalized log-Moyal distribution 16

2.1 Introduction 16

2.2 Data application revisited 18

2.2.1 Analysis of the scaled Norwegian fire losses data set 19

2.2.2 Analysis of the Danish fire insurance losses data set 21
3 On an improved distribution better than the inverse Nagakami-$m$ distribution

3.1 Introduction

3.2 Data applications revisited

3.2.1 Analysis of the elevator data set

3.2.2 Analysis of the motor data set

3.2.3 Analysis of the machining center data set

3.2.4 Analysis of the electric power distribution station data set

3.3 Simulation study

4 The power series exponential power series distributions

4.1 Introduction

4.2 The PSEPS class of distributions

4.2.1 Density, survival and hazard rate functions

4.2.2 Quantiles and median

4.2.3 Moment generating function

4.2.4 Moments

4.2.5 Order statistics
4.3 Estimation and inference .................................................. 63
4.4 Simulation study ............................................................. 69
4.5 Illustrations using failure data sets ...................................... 72
   4.5.1 Generalized exponential geometric (Silva et al. (35)) and generalized exponential geometric (Bidram et al. (39)) distributions ........................................... 74
   4.5.2 Beta Fréchet (Barreto-Souza et al. (36)) and Weibull geometric (Barreto-Souza et al. (37)) distributions ............................................................. 76
   4.5.3 Gamma-exponentiated exponential (Ristic and Balakrishnan (38)) distribution ............................................................. 79
4.6 Conclusions ................................................................. 81
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Data analysis of the scaled Norwegian fire loss data set.</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>Data analysis for the Danish fire insurance losses data set.</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>Data analysis of the elevator data set.</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Data analysis of the motor data set.</td>
<td>37</td>
</tr>
<tr>
<td>3.3</td>
<td>Data analysis of the machining center data set.</td>
<td>39</td>
</tr>
<tr>
<td>3.4</td>
<td>Data analysis of the electric power distribution station data set.</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>Useful quantities for some power series distributions.</td>
<td>52</td>
</tr>
<tr>
<td>4.2</td>
<td>Fitted models, parameter estimates, log-likelihood values and $p$-values of the Kolmogorov-Smirnov statistic for the maximum stress data.</td>
<td>75</td>
</tr>
<tr>
<td>4.3</td>
<td>Fitted models, parameter estimates, log-likelihood values and $p$-values of the Kolmogorov-Smirnov statistic for the fiber data.</td>
<td>78</td>
</tr>
<tr>
<td>4.4</td>
<td>Fitted models, parameter estimates, log-likelihood values and $p$-values of the Kolmogorov-Smirnov statistic for the air conditioning data.</td>
<td>80</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Quantile plots of the fits of LogF, LogP and GlogM distributions on the log-log scale for the scaled 1990-1992 Norwegian data set. 21

2.2 Quantile plots of the fits of LogF, LogP and GlogM distributions on the log-log scale for the Danish fire insurance losses data set. 23

2.3 Empirical estimate of Value at Risk and estimates obtained from the fitted LogF, LogP and GlogM distributions. 24

2.4 Biases of the parameter estimates of the logF distribution. 26

2.5 Mean squared errors of the parameter estimates of the logF distribution. 27

2.6 Biases of the parameter estimates of the logP distribution. 28

2.7 Mean squared errors of the parameter estimates of the logP distribution. 29

3.1 The fitted reliability functions and the Kaplan-Meier estimator for the elevator data set. 36

3.2 The fitted reliability functions and the Kaplan-Meier estimator for the motor data set. 38
3.3 The fitted reliability functions and the Kaplan-Meier estimator for the machining center data set.

40

3.4 The fitted reliability functions and the Kaplan-Meier estimator for the electric power distribution station data set.

42

3.5 Biases of the parameter estimates of (3.2).

44

3.6 Mean squared errors of the parameter estimates of (3.2).

45

4.1 Possible shapes of (4.4) when $(M, N)$ has the (Poisson, Poisson), (Poisson, Geometric), (Poisson, Logarithmic), (Poisson, Binomial), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Geometric, Binomial), (Logarithmic, Poisson), (Logarithmic, Geometric), (Logarithmic, Logarithmic), (Logarithmic, Binomial), (Binomial, Poisson), (Binomial, Geometric), (Binomial, Logarithmic) and (Binomial, Binomial) distributions, arranged from top to bottom and left to right.

57

4.2 Possible shapes of (4.6) when $(M, N)$ has the (Poisson, Poisson), (Poisson, Geometric), (Poisson, Logarithmic), (Poisson, Binomial), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Geometric, Binomial), (Logarithmic, Poisson), (Logarithmic, Geometric), (Logarithmic, Logarithmic), (Logarithmic, Binomial), (Binomial, Poisson), (Binomial, Geometric), (Binomial, Logarithmic) and (Binomial, Binomial) distributions, arranged from top to bottom and left to right.

58

4.3 $\text{bias}_{\theta_1}(n)$ (top left), $\text{bias}_{\theta_2}(n)$ (top right) and $\text{bias}_{\beta}(n)$ (bottom left) versus $n = 10, 11, \ldots, 100$.

70

4.4 $\text{MSE}_{\theta_1}(n)$ (top left), $\text{MSE}_{\theta_2}(n)$ (top right) and $\text{MSE}_{\beta}(n)$ (bottom left) versus $n = 10, 11, \ldots, 100$.

71
Abstract

The whole thesis contains 4 chapters. Chapter 1 is the introductory chapter of my thesis and the main contributions are in Chapter 2 through to Chapter 4. The theme of these chapters is developing and reviewing statistical distributions.

Bhati and Ravi [Insurance: Mathematics and Economics, 79, 2018, 247-259] introduced a new heavy tailed distribution referred to as the generalized log-Moyal distribution. Chapter 2 points out that: i) this distribution is a particular case of many known distributions; ii) the two data sets considered can be fitted better by known distributions, with differences in AIC exceeding 60 and differences in BIC exceeding 50; iii) this distribution does not provide an adequate fit for one of the data sets while known distributions do.

A recent paper introduced a new distribution referred to as the inverse Nakagami-$m$ distribution. Chapter 3 points out that this distribution is a particular case of many known distributions. It also shows that four data sets can be fitted better by known distributions, with differences in information criteria exceeding 20. The better fits are justified also using three goodness of fit tests.

Chapter 4 introduces a new class of distributions motivated by systems having both series and parallel structures. Some mathematical properties of the new class (including the moment generating function, moments and order statistics) are derived. Estimation is addressed by the maximum likelihood method and the performance of the estimators assessed.
by a simulation study. An illustration using three failure data sets shows the usefulness of
the new class.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
4. Further information on the conditions under which disclosure, publication and commercialization of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectualproperty.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
Acknowledgements

I am greatly indebted to my supervisor, Dr Saralees Nadarajah, for his patience, encouragement, and professional instructions during my thesis writing. I feel grateful to all the teachers in the University of Manchester who once offered me valuable courses and advice during my study. Last but not least, I owe much to my friends and classmates for their valuable suggestions and critiques which helped in making the thesis a reality.
Chapter 1

Introduction

1.1 Aims and motivation

This thesis includes selected work of my contributions to statistical distribution theory in my MPhil study. The thesis is presented in the alternative format, each chapter presented in this thesis is a paper currently under review for a refereed journal. Chapters 2 and 4 are focused on some certain theoretical problems in the area of distribution theory. Chapters 2 and 3 is focused on application of distribution theory in a real world case.

The introduction part for each chapter gives specific academic background for certain problems we proposed. The corresponding conclusions are given at the end of each chapter. Thus we do not make a specific chapter for conclusion in this thesis.

The aims and motivation for the chapters in the thesis are as follows:

- the generalized log-Moyal and inverse Nakagami-m distributions have been proposed recently by Bhati and Ravi (2018) and in Louzada et al.(2018). Chapters 2 and 3 show that many known distributions in the literature contain these two distributions
as particular cases. These chapters also show that known distributions provide better fits than the two distributions.

- Chapter 4 introduces a class of distributions motivated by systems having series and parallel structures. The class is shown to provide better fits than five other distributions having at least the same number of parameters for three failure data sets.

1.2 Future work

We shall attempt to address some of the following in the future:

- multivariate generalizations of the generalized log-Moyal and inverse Nakagami-$m$ distributions, their statistical inference and data applications.
- multivariate generalizations of the class of distributions proposed in Chapter 5, their statistical inference and data applications.
Chapter 2

On the generalized log-Moyal distribution

2.1 Introduction

Bhati and Ravi (2018) introduced a novel heavy tailed distribution named the generalized log-Moyal (GlogM) distribution. Its probability density function is given by

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

(2.1)

for \( x > 0, \sigma > 0 \) and \( \mu > 0 \).

We would like to point out that there are many distributions in the literature that contain (2.1) as a particular case. For example, (2.1) is a particular case of the distribution due to Hoq et al. (1974) given by

\[
f(x) = \frac{\beta \theta^q}{\Gamma(q/\beta)} x^{-q-1} \exp \left[ -\left( \frac{\theta}{x} \right)^\beta \right]
\]
for \( x > 0, \ q > 0, \ \theta > 0 \) and \( \beta > 0 \); (2.1) is also a particular case of the distribution due to Lee and Gross (1991) given by

\[
\begin{align*}
    f(x) = \frac{m(ap)^{-bm}}{\Gamma(b)} x^{-bm-1} \exp \left[ -\left( \frac{1}{apx} \right)^m \right]
\end{align*}
\]

for \( x > 0, \ b > 0, \ m > 0, \ a > 0 \) and \( p > 0 \); (2.1) is also a particular case of the distribution due to Kalla et al. (2001) given by

\[
\begin{align*}
    f(x) = \frac{\beta b^\lambda}{n^\lambda \delta + 1, m} x^{-m-\beta-1} \left( 1 + \frac{b}{\delta x^\beta} \right)^{-\lambda} \exp \left( -\frac{1}{\delta x^\beta} \right)
\end{align*}
\]

for \( x > 0, \ \beta > 0, \ b > 0, \ m > 0, \ n > 0, \ \delta > 0 \) and \( \lambda \geq 0 \), where \( \Gamma_\lambda(\alpha, k) \) denotes a confluent hypergeometric function defined by

\[
\Gamma_\lambda(\alpha, k) = \int_0^\infty t^{\alpha-1} (t + k)^{-\lambda} \exp(-t) dt;
\]

(2.1) is also a particular case of the distribution due to Mead (2015) given by

\[
\begin{align*}
    f(x) = \frac{\beta \theta^{\alpha \beta}}{\Gamma_\lambda(\alpha, k)} x^{-\alpha \beta - 1} \left[ \left( \frac{\theta}{x} \right)^\beta + k \right]^{-\lambda} \exp \left( -\left( \frac{\theta}{x} \right)^\beta \right)
\end{align*}
\]

for \( x > 0, \ \alpha > 0, \ k > 0, \ \theta > 0, \ \beta > 0 \) and \( \lambda \geq 0 \); and so on.

In the rest of this chapter, we revisit the data application in Bhati and Ravi (2018) involving two data sets. We show that a distribution due to Nadarajah and Bakar (2014) provides better fits than the GlogM distribution. For the first data set, the difference in AIC values is over 60 and the difference in BIC values is over 50. For the second data set, the difference in AIC values is over 140 and the difference in BIC values is over 120. Goodness of fit tests show that the GlogM distribution does not provide an adequate fit to the second data set while a distribution due to Nadarajah and Bakar (2014) does.
The purpose of this chapter is not to study mathematical properties or physical interpretations relating to the distribution due to Nadarajah and Bakar (2014). Such details can be found in Nadarajah and Bakar (2014) and the many papers citing it.

### 2.2 Data application revisited

Nadarajah and Bakar (2014) introduced composite lognormal distributions given by the probability density function

\[
f(x) = \begin{cases} 
\frac{\psi \left( \frac{\log x - \mu}{\sigma} \right)}{(1 + \phi) \sigma \Phi \left( \frac{\log \theta - \mu}{\sigma} \right)}, & \text{if } 0 < x \leq \theta, \\
\phi f_2(x) \left[ 1 - F_2(\theta) \right], & \text{if } \theta < x < \infty,
\end{cases}
\]

where $\theta$ is a cutoff point, $\psi(\cdot)$ denotes the standard normal probability density function, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, $f_2(x)$ denotes a valid probability density function and $F_2(x)$ denotes the corresponding cumulative distribution function. Furthermore,

\[
\phi = \frac{\psi \left( \frac{\log \theta - \mu}{\sigma} \right)}{\theta \sigma f_2(\theta) \Phi \left( \frac{\log \theta - \mu}{\sigma} \right)} \left[ 1 - F_2(\theta) \right].
\]

Different distributions can be obtained for (2.2) by specifying different forms for $f_2$. We obtain the lognormal-Pareto (LogP) distribution when $f_2$ is specified by

\[
f_2(x) = \frac{\alpha \beta^\alpha}{(x + \beta)^{\alpha+1}}
\]

for $x > 0$, $\alpha > 0$ and $\beta > 0$. We obtain the lognormal-Fréchet (LogF) distribution when $f_2$
is specified by

\[
    f_2(x) = \frac{\alpha \beta}{x^2} \left( \frac{\beta}{x} \right)^{\alpha - 1} \exp \left[ -\left( \frac{\beta}{x} \right)^{\alpha} \right]
\]

for \( x > 0, \alpha > 0 \) and \( \beta > 0 \).

We fitted the GlogM, LogP and LogF distributions to two actuarial data sets: the Norwegian fire losses and Danish fire insurance losses data sets. The Norwegian fire losses data set was scaled by dividing each observation by 1000. The method of maximum likelihood was used to fit distributions for both data sets. The R package \texttt{CompLognormal} (Nadarajah and Bakar, 2013) was used to fit the LogP and LogF distributions.

### 2.2.1 Analysis of the scaled Norwegian fire losses data set

The data are Norwegian fire claim data exceeding five hundred thousand Norwegian krones. The data were rounded to thousand Norwegian krones and then divided by 1000. The parameter estimates, log likelihood values, values of the Akaike information criterion (AIC) and values of the Bayesian information criterion (BIC) are given in Table 2.1. The LogF distribution has the smallest AIC and the smallest BIC in years 1990, 1991 and 1992. The differences with AIC and BIC values of the GlogM distribution are huge.
Table 2.1: Data analysis of the scaled Norwegian fire loss data set.

<table>
<thead>
<tr>
<th>Year</th>
<th>Model</th>
<th>GlogM</th>
<th>LogF</th>
<th>LogP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameters</td>
<td>$\left(\hat{\mu}, \hat{\sigma}\right)$</td>
<td>$\left(\hat{\sigma}, \hat{\mu}, \hat{\beta}, \hat{\alpha}\right)$</td>
<td>$\left(\hat{\sigma}, \hat{\mu}, \hat{\beta}, \hat{\alpha}\right)$</td>
</tr>
<tr>
<td>Estimates</td>
<td>(0.863, 0.319)</td>
<td>(-14.407, -0.693, -0.179, 0.530)</td>
<td>(-14.125, -0.693, 0.267, 0.856)</td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>-743.6</td>
<td>-721.7</td>
<td>-728.4</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>AIC</td>
<td>1491.2</td>
<td>1451.3</td>
<td>1464.8</td>
</tr>
<tr>
<td>BIC</td>
<td>1500.0</td>
<td>1469.1</td>
<td>1482.6</td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td>(0.840, 0.318)</td>
<td>(-12.908, -0.693, -0.258, 0.499)</td>
<td>(-16.952, -0.693, 0.219, 0.861)</td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>-724.3</td>
<td>-696.5</td>
<td>-697.6</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>AIC</td>
<td>1452.6</td>
<td>1401.0</td>
<td>1403.1</td>
</tr>
<tr>
<td>BIC</td>
<td>1461.5</td>
<td>1418.7</td>
<td>1420.9</td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td>(0.838, 0.334)</td>
<td>(-20.458, -0.693, -0.396, 0.382)</td>
<td>(-15.185, -0.693, -0.374, 0.551)</td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>-769.5</td>
<td>-735.2</td>
<td>-736.2</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>AIC</td>
<td>1543.1</td>
<td>1478.4</td>
<td>1480.5</td>
</tr>
<tr>
<td>BIC</td>
<td>1552.0</td>
<td>1496.1</td>
<td>1498.2</td>
<td></td>
</tr>
</tbody>
</table>
2.2.2 Analysis of the Danish fire insurance losses data set

The Danish fire insurance losses data set consists of 2492 losses in millions of Danish Kroner during the period from 1980 to 1990, arising from the fire claims in Copenhagen. The
parameter estimates, log likelihood values, values of AIC and values of the BIC are given in Table 2.2.2. The LogF distribution has the smallest AIC and the smallest BIC. The differences with AIC and BIC values of the GlogM distribution are once again huge.

Table 2.2: Data analysis for the Danish fire insurance losses data set.

<table>
<thead>
<tr>
<th>Models→</th>
<th>GlogM</th>
<th>LogF</th>
<th>LogP</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}$</td>
<td>1.312</td>
<td>0.003</td>
<td>0.135</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.413</td>
<td>0.447</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>-0.287</td>
<td>-1.012</td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.321</td>
<td>-1.718</td>
<td>-1.702</td>
</tr>
<tr>
<td>Lmax</td>
<td>-3933.0</td>
<td>-3859.3</td>
<td>-3860.5</td>
</tr>
<tr>
<td>AIC</td>
<td>7870.00</td>
<td>7726.6</td>
<td>7728.9</td>
</tr>
<tr>
<td>BIC</td>
<td>7872.8</td>
<td>7749.9</td>
<td>7752.2</td>
</tr>
</tbody>
</table>
Figure 2.2: Quantile plots of the fits of LogF, LogP and GlogM distributions on the log-log scale for the Danish fire insurance losses data set.
Figure 2.3: Empirical estimate of Value at Risk and estimates obtained from the fitted LogF, LogP and GlogM distributions.

The quantile plots of the fitted distributions on the log-log scale are presented in Figure 2.2. The plots indicate that the GlogM, LogF and LogP distributions give similar fits in the middle and lower parts. But both LogF and LogP distributions appear to provide better fits than the GlogM distribution in the upper tail.

Figure 2.3 plots estimates of Value at Risk (VaR) computed from the fitted LogF, LogP and GlogM distributions. Also plotted are the empirical estimates of VaR. We see that the estimates from the fitted LogF and LogP distributions are consistently closer to the empirical estimates than the estimates from the fitted GlogM distribution. This is further evidence that LogF and LogP distributions provide better fits.
2.3 Simulation study

In this section, we conduct a simulation study to assess the performance and accuracy of the maximum likelihood estimators of the better fitting logF and logP distributions. The following scheme was used:

1. simulate a sample of size \( n \) from the logF / logP distribution;
2. estimate \((\mu, \alpha, \beta, \sigma)\);
3. repeat steps 1 and 2 ten thousand times;
4. hence, estimate the biases and the mean squared errors for the four parameters;
5. repeat steps 1 to 4 for \( n = 20, 21, \ldots, 500 \).

The plots of the biases versus \( n \) for the logF distribution are shown in Figure 2.4. The plots of the mean squared errors versus \( n \) for the logF distribution are shown in Figure 2.5. The plots of the biases versus \( n \) for the logP distribution are shown in Figure 2.6. The plots of the mean squared errors versus \( n \) for the logP distribution are shown in Figure 2.7.
Figure 2.4: Biases of the parameter estimates of the logF distribution.
Figure 2.5: Mean squared errors of the parameter estimates of the logF distribution.
Figure 2.6: Biases of the parameter estimates of the logP distribution.
Figure 2.7: Mean squared errors of the parameter estimates of the logP distribution.

We can observe the following from the figures for both the logF and logP distributions: the biases can be positive or negative but approach zero as \( n \) approaches 500; the biases appear largest for \( \hat{\sigma} \) and smallest for \( \hat{\mu} \); the biases appear reasonably small at around \( n = 500 \); the mean squared errors gradually decrease with increasing \( n \); the mean squared errors appear largest for \( \hat{\sigma} \) and smallest for \( \hat{\mu} \); the mean squared errors appear reasonably small at around \( n = 500 \).

In the simulation scheme, we have taken the initial parameter values as the estimated values for the scaled Norwegian fire losses data set. The results were similar for a wide range
of other initial values including the estimated values for the other data set.

2.4 Conclusions

We have pointed out that the GlogM distribution is a particular case of many known distributions, some known as early as the 1970s. We have also suggested a distribution that can provide better fits to the Norwegian fire losses and Danish fire insurance losses data sets. The distribution suggested is a composite distribution and has four parameters, but there are free software for fitting it. So the distribution is easy to handle. Its fit gives a difference in AIC as high as 150 and a difference in BIC as high as 140.

We also tested the goodness of the fitted distributions by the Kolmogorov Smirnov, Carmer von Mises and Anderson-Darling tests. For the scaled Norwegian fire losses data set, all of the fitted distributions gave adequate fits at the 5 percent level of significance. For the Danish fire insurance losses data set, the $p$ values of the tests for the GlogM distribution were $0.002$, $0.002$ and $2.408 \times 10^{-7}$. The $p$ values of the tests for the LogP distribution were $0.297$, $0.242$ and $0.098$. Hence, for the Danish fire insurance losses data set, the GlogM distribution does not provide an adequate fit while the LogP distribution does. The $p$ values reported in Bhati and Ravi (2018) for both data sets do not appear correct.
Bibliography


Journal of Statistical Planning and Inference, 29, 137-143.


posite lognormal distributions. R Journal, 5, 97-103.

Chapter 3

On an improved distribution better than the inverse Nagakami-\(m\) distribution

3.1 Introduction

Louzada et al. (2018) introduced a novel distribution named the inverse Nakagami-\(m\) distribution. Its probability density function is given by

\[
f(x) = \frac{2}{\Gamma(\mu)} \left( \frac{\mu}{\Omega} \right)^\mu x^{-2\mu-1} \exp\left( -\frac{\mu}{\Omega x^2} \right)
\]

(3.1) for \(x > 0\), \(\mu > 0\) and \(\Omega > 0\).

We would like to point out that there are many distributions in the literature that contain (3.1) as a particular case. For example, (3.1) is a particular case of the distribution due to
CHAPTER 3. ON AN IMPROVED DISTRIBUTION BETTER THAN THE INVERSE NAGAKAMI-M

Hoq et al. (1974) given by

\[
f(x) = \frac{\beta \theta^q}{\Gamma(q/\beta)} x^{-q-1} \exp \left[ -\left( \frac{\theta}{x} \right)^\beta \right]
\]

(3.2)

for \( x > 0, \ q > 0, \ \theta > 0 \) and \( \beta > 0 \); (3.1) is also a particular case of the distribution due to

Lee and Gross (1991) given by

\[
f(x) = \frac{m(ap)^{-bm}}{\Gamma(b)} x^{-bm-1} \exp \left[ -\left( \frac{1}{apx} \right)^m \right]
\]

for \( x > 0, \ b > 0, \ m > 0, \ a > 0 \) and \( p > 0 \); (3.1) is also a particular case of the distribution
due to Kalla et al. (2001) given by

\[
f(x) = \frac{\beta \theta^\alpha}{\Gamma(\lambda)} x^{-\lambda-1} \left[ \left( \frac{\alpha}{x} \right) \theta + k \right]^{-\lambda} \exp \left[ -\left( \frac{\theta}{x} \right)^\beta \right]
\]

for \( x > 0, \ \alpha > 0, \ k > 0, \ \theta > 0, \ \beta > 0 \) and \( \lambda \geq 0 \); and so on.

In the rest of this chapter, we revisit two of the data applications in Louzada et al. (2018)
and two other data applications. We show that (3.2) provides better fits than the inverse
Nakagami-m distribution. The better fits are assessed in terms of the following information
criteria: the Akaike information criterion (AIC) due to Akaike (1974); the consistent Akaike
information criterion (CAIC) due to Bozdogan (1987); the corrected Akaike information
criterion (AICc) due to Hurvich and Tsai (1989); the Hannan-Quinn criterion (HQIC) due
to Hannan and Quinn (1979). The better fits are also assessed in terms of $p$ values of the
Kolmogorov Smirnov, Cramer von Mises and Anderson-Darling tests.

For the first data set, the differences in information criteria values are over 20. For this
data, both (3.1) and (3.2) give adequate fits in terms of the goodness of fit tests. For the
second data set, the differences in information criteria values are over 10 and (3.2) gives the
only adequate fit in terms of the goodness of fit tests. For the third data set, the differences
in information criteria values are over 10. Moreover, (3.2) gives an adequate fit in terms of
all three goodness of fit tests while (3.1) gives an adequate fit in terms of only two of the
three goodness of fit tests. For the fourth data set, the differences in information criteria
values are over 15 and (3.2) gives the only adequate fit in terms of the goodness of fit tests.

The purpose of this chapter is not to study mathematical properties or physical interpre-
tations relating to the distribution in (3.2). Such details can be found in Hoq et al. (1974)
and the many papers citing it.

3.2 Data applications revisited

In this section, we revisit two of the data sets considered by Louzada et al. (2018) plus
two other data sets: data related to an agricultural machine’s elevator (Louzada et al.,
2018); data related to an agricultural machine’s motor (Louzada et al., 2018); data related
to machining centers (Dai et al., 2003); data related to electric power distribution stations
(Alwan et al., 2013). We fitted the distributions given by (3.1) and (3.2) to the data sets.
The method of maximum likelihood was used to fit distributions for both data sets. For
computational stability, all data sets were scaled by dividing each observation by 100. It is
common practice that data are scaled to small values before a distribution is fitted. This
is because the maximisation routines like optim in R work more efficiently with small data values. This is a well known fact. The scaling of data applies to fitting of any distribution, even a one parameter exponential distribution for example. The scaling of data has nothing to do with the number of parameters a distribution has.

### 3.2.1 Analysis of the elevator data set

The parameter estimates, AIC values, AICc values, HQIC values, CAIC values and $p$ values of the three goodness of fit tests are given in Table 3.2.1. (3.2) gives the smaller values of AIC, AICc, HQIC and CAIC. The differences with AIC, AICc, HQIC and CAIC values of (3.1) are highly significant. The $p$ values show that both (3.1) and (3.2) give adequate fits with respect to Cramer von Mises and Anderson-Darling tests.

<table>
<thead>
<tr>
<th>Model</th>
<th>(3.1)</th>
<th>(3.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters estimates</td>
<td>$(\hat{\mu}, \Omega)$</td>
<td>$(\hat{\beta}, \hat{\theta}, \hat{q})$</td>
</tr>
<tr>
<td></td>
<td>$(0.341, 3597.164)$</td>
<td>$(153.512, 0.010, 0.652)$</td>
</tr>
<tr>
<td>AIC</td>
<td>-151.5</td>
<td>-173.5</td>
</tr>
<tr>
<td>AICc</td>
<td>-151.3</td>
<td>-173.0</td>
</tr>
<tr>
<td>HQIC</td>
<td>-145.0</td>
<td>-171.2</td>
</tr>
<tr>
<td>CAIC</td>
<td>-145.5</td>
<td>-164.5</td>
</tr>
<tr>
<td>KS</td>
<td>0.018</td>
<td>0.012</td>
</tr>
<tr>
<td>CvM</td>
<td>0.114</td>
<td>0.110</td>
</tr>
<tr>
<td>AD</td>
<td>0.058</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Figure 3.1: The fitted reliability functions and the Kaplan-Meier estimator for the elevator data set.

Figure 3.1 compares the empirical reliability function with the fitted versions. The fitted values of (3.1) and (3.2) appear very close to each other.

### 3.2.2 Analysis of the motor data set

The parameter estimates, AIC values, AICc values, HQIC values, CAIC values and $p$ values of the three goodness of fit tests are given in Table 3.2.2. (3.2) once again gives the smaller values of AIC, AICc, HQIC and CAIC. The differences with AIC, AICc, HQIC and CAIC
values of (3.1) are once again highly significant. The $p$ values show that (3.2) gives an adequate fit with respect to Cramer von Mises and Anderson-Darling tests. None of the tests support an adequate fit of (3.1).

Table 3.2: Data analysis of the motor data set.

<table>
<thead>
<tr>
<th>Model→</th>
<th>(3.1)</th>
<th>(3.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>$(\hat{\mu}, \hat{\Omega})$</td>
<td>$(\hat{\beta}, \hat{\theta}, \hat{q})$</td>
</tr>
<tr>
<td>estimates</td>
<td>(0.455, 1902.077)</td>
<td>(0.688, 0.056, 1.098)</td>
</tr>
<tr>
<td>AIC</td>
<td>-159.7</td>
<td>-172.7</td>
</tr>
<tr>
<td>AICc</td>
<td>-159.5</td>
<td>-172.3</td>
</tr>
<tr>
<td>HQIC</td>
<td>-158.0</td>
<td>-170.1</td>
</tr>
<tr>
<td>CAIC</td>
<td>-153.3</td>
<td>-163.2</td>
</tr>
<tr>
<td>KS</td>
<td>0.0002</td>
<td>0.014</td>
</tr>
<tr>
<td>CvM</td>
<td>0.008</td>
<td>0.120</td>
</tr>
<tr>
<td>AD</td>
<td>$5.155 \times 10^{-5}$</td>
<td>0.055</td>
</tr>
</tbody>
</table>
Figure 3.2: The fitted reliability functions and the Kaplan-Meier estimator for the motor data set.

Figure 3.2 comparing the empirical reliability function with the fitted versions shows that the fitted values of (3.2) are a lot closer to the empirical reliability function.

3.2.3 Analysis of the machining center data set

The parameter estimates, AIC values, AICc values, HQIC values, CAIC values and $p$ values of the three goodness of fit tests are given in Table 3.2.3. (3.2) gives the smaller values of AIC, AICc, HQIC and CAIC. The differences with AIC, AICc, HQIC and CAIC values of
(3.1) are highly significant. The $p$ values show that (3.2) gives an adequate fit with respect to all the three tests. Only Cramer von Mises and Anderson-Darling tests support an adequate fit of (3.1). Their $p$ values are a lot smaller than those for (3.2).

Table 3.3: Data analysis of the machining center data set.

<table>
<thead>
<tr>
<th>Model →</th>
<th>(3.1)</th>
<th>(3.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters estimates</td>
<td>$\left(\hat{\mu}, \hat{\Omega}\right)$</td>
<td>$\left(\hat{\beta}, \hat{\theta}, \hat{q}\right)$</td>
</tr>
<tr>
<td>(0.266, 5.961)</td>
<td>(0.063, $1.980 \times 10^{37}$, 14.165)</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>95.2</td>
<td>82.0</td>
</tr>
<tr>
<td>AICc</td>
<td>95.9</td>
<td>83.5</td>
</tr>
<tr>
<td>HQIC</td>
<td>95.6</td>
<td>82.6</td>
</tr>
<tr>
<td>CAIC</td>
<td>99.2</td>
<td>88.0</td>
</tr>
<tr>
<td>KS</td>
<td>0.031</td>
<td>0.560</td>
</tr>
<tr>
<td>CvM</td>
<td>0.050</td>
<td>0.400</td>
</tr>
<tr>
<td>AD</td>
<td>0.059</td>
<td>0.447</td>
</tr>
</tbody>
</table>
Figure 3.3: The fitted reliability functions and the Kaplan-Meier estimator for the machining center data set.

Figure 3.3 compares the empirical reliability function with the fitted versions. The fitted values of (3.2) are a lot closer to the empirical reliability function.

### 3.2.4 Analysis of the electric power distribution station data set

The parameter estimates, AIC values, AICc values, HQIC values, CAIC values and \( p \) values of the three goodness of fit tests are given in Table 3.2.4. (3.2) gives the smaller values of AIC, AICc, HQIC and CAIC. The differences with AIC, AICc, HQIC and CAIC values of
(3.1) are highly significant. The $p$ values show that (3.2) gives an adequate fit with respect to all the three tests. None of the tests support an adequate fit of (3.1).

Table 3.4: Data analysis of the electric power distribution station data set.

<table>
<thead>
<tr>
<th>Model $\rightarrow$</th>
<th>(3.1)</th>
<th>(3.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters estimates</td>
<td>$(\hat{\mu}, \hat{\Omega})$</td>
<td>$(\hat{\beta}, \hat{\theta}, \hat{q})$</td>
</tr>
<tr>
<td>$(0.189, 483.082)$</td>
<td>$(0.060, 5.115 \times 10^{35}, 8.710)$</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>58.6</td>
<td>39.4</td>
</tr>
<tr>
<td>AICc</td>
<td>59.1</td>
<td>40.5</td>
</tr>
<tr>
<td>HQIC</td>
<td>59.4</td>
<td>40.6</td>
</tr>
<tr>
<td>CAIC</td>
<td>63.2</td>
<td>46.3</td>
</tr>
<tr>
<td>KS</td>
<td>0.021</td>
<td>0.370</td>
</tr>
<tr>
<td>CvM</td>
<td>0.016</td>
<td>0.494</td>
</tr>
<tr>
<td>AD</td>
<td>0.022</td>
<td>0.578</td>
</tr>
</tbody>
</table>
Figure 3.4: The fitted reliability functions and the Kaplan-Meier estimator for the electric power distribution station data set.

Figure 3.4 compares the empirical reliability function with the fitted versions. The fitted values of (3.2) are a lot closer to the empirical reliability function.

### 3.3 Simulation study

In this section, we conduct a simulation study to assess the performance and accuracy of the maximum likelihood estimators of the better fitting distribution (3.2). The following scheme
was used:

1. simulate a sample of size $n$ from (3.2);
2. estimate $(q, \theta, \beta)$;
3. repeat steps 1 and 2 ten thousand times;
4. hence, estimate the biases and the mean squared errors for the three parameters;
5. repeat steps 1 to 4 for $n = 20, 21, \ldots, 500$.

The plots of the biases versus $n$ are shown in Figure 3.5. The plots of the mean squared errors versus $n$ are shown in Figure 3.6.
Figure 3.5: Biases of the parameter estimates of (3.2).
We can observe the following from the figures: the biases can be positive or negative but approach zero as \( n \) approaches 500; the biases appear largest for \( \hat{\theta} \) and smallest for \( \hat{q} \); the biases appear reasonably small at around \( n = 500 \); the mean squared errors gradually decrease with increasing \( n \); the mean squared errors appear largest for \( \hat{\theta} \) and smallest for \( \hat{q} \); the mean squared errors appear reasonably small at around \( n = 500 \).

In the simulation scheme, we have taken the initial parameter values as the estimated values in Section 3.2.1. The results were similar for a wide range of other initial values including the estimated values for the other three data sets.
Bibliography


Chapter 4

The power series exponential power series distributions

4.1 Introduction

The exponential distribution is the simplest and oldest model for failure times. But it has a constant hazard rate function. Most real life systems do not exhibit constant hazard rates.

Recently, many generalizations of the exponential distribution have been proposed to accommodate monotonically increasing, monotonically decreasing, bathtub shaped and upside down bathtub shaped hazard rates. Most of these generalizations have been motivated by systems consisting of series components or systems consisting of parallel components. Some generalizations motivated by systems consisting of series components are the exponential-geometric (EG), exponential-Poisson (EP), exponential-logarithmic (EL), exponential power series (EPS), Weibull-power series (WPS), modified Weibull and extended Weibull power series (EWPS) distributions, see Adamidis and Loukas (1), Lai et al. (2), Kus (3), Tahmasbi and Rezaei (4), Chahkandi and Ganjali (5), Morais and Barreto-Souza (6), Almalki and
Yuan (7), Silva et al. (8) and Almalki and Nadarajah (9). Some generalizations motivated by systems consisting of parallel components are the exponentiated exponential-Poisson (EEP), complementary exponential-geometric (CEG), generalized exponential-power series (GEPS), Poisson-exponential (PE) and exponentiated modified Weibull extension distributions, see Barreto-Souza and Cribari-Neto (10), Cancho at al. (11), Louzada-Neto et al. (12), Mahmoudi and Jafari (13) and Sarhan and Apaloo (14). See also Singla et al. (15), Cordeiro et al. (16), Gomes et al. (17), Ortega et al. (18), Pinho et al. (19) and Tahir et al. (20).

The aim of this chapter is to introduce a distribution motivated by series and parallel units. Systems made up of series and parallel units arise in many areas of reliability, see Sharifi et al. (21), Nezhad (22) and Jang and Kim (23). Suppose a company has an optimum number of systems say $N$ functioning in series and independently, producing a certain product. As mentioned in Nadarajah et al. (24), this optimum number may be determined by such factors as economy, manpower and customer demand. So, $N$ can be regarded as a discrete random variable with support $\{1, 2, \ldots\}$. Suppose that each system is made of $M$ parallel units and that $M$ is also a discrete random variable with support $\{1, 2, \ldots\}$. Suppose also that the failure times of the units in the $i$th system say $Z_{i,1}, Z_{i,2}, \ldots, Z_{i,M}$ are independent and identical exponential random variables with rate parameter $\beta$. The failure time of the $i$th system for $i = 1, 2, \ldots, N$ is $Y_i = \max_{1 \leq j \leq M} Z_{i,j}$. The failure time of the system is $X = \min_{1 \leq i \leq N} Y_i$.

The distribution of $X$ is motivated by $N$ systems functioning in series and $M$ units working in parallel. We take $M$ and $N$ to be independent power series random variables. We shall refer to the distribution of $X$ as the Power Series Exponential Power Series (PSEPS) distribution. The power series distribution contains the geometric, Poisson, logarithmic and binomial distributions as particular cases. So, sixteen forms are possible for the distribution of $(M, N)$: (Poisson, Poisson), (Poisson, Geometric), (Poisson, Logarithmic), (Poisson,
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

Binomial), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Geometric, Binomial), (Logarithmic, Poisson), (Logarithmic, Geometric), (Logarithmic, Logarithmic), (Logarithmic, Binomial), (Binomial, Poisson), (Binomial, Geometric), (Binomial, Logarithmic) and (Binomial, Binomial).

The main motivations for the PSEPS distribution are that:

- the power series distribution contains all standard discrete distributions as particular cases;

- the power series distribution has received numerous recent theoretical developments - the class of distributions due to Chahkandi and Ganjali (5), the complementary exponential power series distribution due to Flores et al. (25), the compound class of extended Weibull power series distributions due to Silva et al. (8) and the class of distributions due to Bourguignon et al. (26) are all based on the power series distribution;

- the power series distribution has also received several recent applications — cure rate models for a cutaneous melanoma data due to Borges et al. (27), long-term survival models with latent activation due to Cancho et al. (28) and cure rate models for a cutaneous melanoma data due to Cancho et al. (29) are all based on the power series distribution;

- it is characterized by systems having series and parallel structures;

- it admits monotonically increasing, monotonically decreasing, bathtub shaped and upside down bathtub shaped hazard rates;

- it incorporates sixteen possible forms for the distribution of \((M, N)\);

- it provides better fits than five other distributions having at least the same number of parameters for three failure data sets.
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

The contents of the paper are organized as follows: In Section 4.2, we introduce the PSEPS class of distributions. The density, survival, hazard rate and moment generating functions as well as the moments, order statistics and quantiles are derived in Section 4.2.4. In Section 4.3, we investigate estimation by maximum likelihood. A simulation study is presented in Section 4.4 to assess finite sample performance of the maximum likelihood estimators. Real data illustrations of the PSEPS distribution are in Section 4.5. Finally, some concluding remarks are given in Section 4.6.

4.2 The PSEPS class of distributions

The new class of distributions can be defined as follows: Let $N$ be a truncated power series random variable specified by the probability mass function

$$\Pr(N = n) = \frac{a_n \theta_1^n}{C(\theta_1)}$$

for $n = 1, 2, \ldots$, where $a_n \geq 0$ depends only on $n$, $C(\theta_1) = \sum_{n=1}^{\infty} a_n \theta_1^n$ and $\theta_1 \in (0, s)$ ($s$ can be $+\infty$) is such that $C(\cdot)$ is finite. For example, the truncated binomial random variable has the probability mass function

$$\Pr(N = n) = \binom{k}{n} p^n (1-p)^{k-n}$$

$$\frac{1 - (1-p)^k}{(1-p)^{-k} - 1} = \frac{\binom{k}{n} \theta_1^n}{(1 + \theta_1)^{k} - 1}$$

for $n = 1, 2, \ldots$, where $\theta_1 = p/(1-p)$. Clearly, (4.2) takes the form (4.1) with $C(\theta_1) = (1 + \theta_1)^{k} - 1$ and $a_n = \binom{k}{n}$.

Let $C'(\cdot)$, $C''(\cdot)$ and $C'''(\cdot)$ denote the first, second and third derivatives of $C(\cdot)$. Table 4.1 lists some particular cases of the truncated power series distribution (geometric, truncated Poisson, logarithmic and truncated binomial). Power series distributions have been
Considered in Boehme and Powell (30) and Ostrovskaya (31). Detailed properties of power series distributions can be found in Noack (32).

\[
\begin{array}{cccccccc}
\text{Distribution} & a_n & C(\theta) & C'(\theta) & C''(\theta) & C'''(\theta) & C^{-1}(\theta) & s \\
\hline
\text{Poisson} & n!^{-1} & e^{\theta} - 1 & e^{\theta} & e^{\theta} & \log(\theta + 1) & \infty \\
\text{Geometric} & 1 & \theta(1 - \theta)^{-1} & (1 - \theta)^{-2} & 2(1 - \theta)^{-3} & 3(1 - \theta)^{-4} & \theta(1 + \theta)^{-1} & 1 \\
\text{Logarithmic} & 1 & -\log(1 - \theta) & (1 - \theta)^{-1} & (1 - \theta)^{-2} & 2(1 - \theta)^{-3} & 1 - e^{-\theta} & 1 \\
\text{Binomial} & \binom{k}{m} & (1 + \theta)^{k - 1} & \frac{k}{(1 + \theta)^{1 - k}} & \frac{k(k - 1)}{(1 + \theta)^{2 - k}} & \frac{k(k - 1)(k - 2)}{(1 + \theta)^{3 - k}} & (1 + \theta)^{\frac{1}{k} - 1} & \infty \\
\end{array}
\]

Table 4.1: Useful quantities for some power series distributions.

Assume that \( M \) is a power series random variable independent of \( N \) specified by the probability mass function

\[
\Pr(M = m) = \frac{b_m \theta_m^n}{D(\theta_2)}
\]

for \( m = 1, 2, \ldots \), where \( b_m \geq 0 \) depends only on \( m \), \( D(\theta_2) = \sum_{m=1}^{\infty} b_m \theta_m^n \) and \( \theta_2 \in (0, s) \) (\( s \) can be \( +\infty \)) is such that \( D(\cdot) \) is finite. Let \( D'(\cdot), D''(\cdot) \) and \( D'''(\cdot) \) denote the first, second and third derivatives of \( D(\cdot) \).

Given \( M \) and \( N \), suppose \( Z_{i,1}, Z_{i,2}, \ldots, Z_{i,M} \) are independent and identical exponential random variables with rate parameter \( \beta \). Let \( Y_i = \max_{1 \leq j \leq M} Z_{i,j} \). The conditional cumulative distribution function of \( Y_i | M = m \) is

\[
G_{Y_i|M=m}(x) = \left(1 - e^{-\beta x}\right)^m,
\]

the cumulative distribution function of a generalized exponential (GE) distribution with parameters \( m \) and \( \beta \) (Gupta and Kundu (33)). So, the unconditional cumulative distribution
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

function of $Y_i$ is

$$G_{Y_i}(x) = \sum_{m=1}^{\infty} \frac{b_m \theta_i^m}{D(\theta_2)} (1 - e^{-\beta x})^m = \frac{D(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)}.$$ 

Now, let $X = \min_{1 \leq i \leq N} Y_i$. Then the unconditional cumulative distribution function of $X$ is

$$F_X(x) = 1 - \Pr(X > x)$$

$$= 1 - \sum_{n=1}^{\infty} \frac{a_n \theta_1^n}{C(\theta_1)} [1 - G_{Y_i}(x)]^n$$

$$= 1 - \frac{C(\theta_1 (1 - G_{Y_i}(x)))}{C(\theta_1)}$$

$$= 1 - \frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x})) / D(\theta_2))}{C(\theta_1)}, \quad (4.3)$$

where $\beta > 0$ and $\theta_1, \theta_2 \in (0,s)$. We denote a random variable $X$ having this cumulative distribution function by $X \sim \text{PSEPS}(\theta_1, \beta, \theta_2)$. The particular cases of (4.3) for $\theta_2 \to 0$ and $\theta_1 \to 0$ are

$$F(x) = 1 - \frac{C(\theta_1 (1 - (1 - e^{-\beta x})^{c_1}))}{C(\theta_1)}$$

and

$$F(x) = 1 - \left[ 1 - \frac{D(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right]^{c_2},$$

respectively, where $c_1 = \min \{m \in \mathbb{N} : b_m > 0\}$ and $c_2 = \min \{n \in \mathbb{N} : a_n > 0\}$.

We see that $a_1 = 1!^{-1} = 1 > 0$ for the Poisson distribution, $a_1 = 1 > 0$ for the geometric distribution, $a_1 = 1^{-1} = 1 > 0$ for the logarithmic distribution and $a_1 = \binom{k}{1} = k > 0$ for the binomial distribution. So, $c_1 = 1$ for all of the distributions in Table 4.1.
4.2.1 Density, survival and hazard rate functions

The probability density function of $X$ corresponding to (4.3) is

$$f_X(x) = \theta_1 \theta_2 \beta e^{-\beta x} \frac{C' \left( \theta_1 - \theta_1 D \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right) / D \left( \theta_2 \right) \right) D' \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right)}{C \left( \theta_1 \right) D \left( \theta_2 \right)}$$

(4.4)

for $x > 0$. The corresponding survival and hazard rate functions are

$$F_X(x) = \frac{C \left( \theta_1 - \theta_1 D \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right) / D \left( \theta_2 \right) \right)}{C \left( \theta_1 \right)}$$

(4.5)

and

$$h_X(x) = \theta_1 \theta_2 \beta e^{-\beta x} \frac{C' \left( \theta_1 - \theta_1 D \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right) / D \left( \theta_2 \right) \right) D' \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right)}{D \left( \theta_2 \right) C \left( \theta_1 - \theta_1 D \left( \theta_2 \left( 1 - e^{-\beta x} \right) \right) / D \left( \theta_2 \right) \right)},$$

(4.6)

respectively.

**Theorem 4.2.1.** The limiting cumulative distribution function of $\text{PSEPS}(\theta_1, \beta, \theta_2)$ for $\theta_1 \rightarrow 0^+$ and $\theta_2 \rightarrow 0^+$ is $1 - \left[ 1 - \left( 1 - e^{-\beta x} \right)^{c_1} \right]^{c_2}$, where $c_1 = \min \{ m \in \mathbb{N} : b_m > 0 \}$ and $c_2 = \min \{ n \in \mathbb{N} : a_n > 0 \}$. This limiting cumulative distribution function is the cumulative distribution function of $\min_{1 \leq i \leq c_2} Z_i$, where $Z_1, Z_2, \ldots, Z_{c_2}$ are independent random variables specified by the common cumulative distribution function $\left( 1 - e^{-\beta x} \right)^{c_1}$. 

Proof: Using $C(\theta_1) = \sum_{n=1}^{\infty} a_n \theta_1^n$ and $D(\theta_2) = \sum_{m=1}^{\infty} b_m \theta_2^m$, we have

$$
\lim_{\theta_2 \to 0^+} \frac{D(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} = \lim_{\theta_2 \to 0^+} \frac{\sum_{m=1}^{\infty} b_m \theta_2^m (1 - e^{-\beta x})^m}{\sum_{m=1}^{\infty} b_m \theta_2^m} = \lim_{\theta_2 \to 0^+} \frac{b_{c_1} (1 - e^{-\beta x})^{c_1} + \sum_{m=c_1+1}^{\infty} b_m \theta_2^{m-c_1} (1 - e^{-\beta x})^m}{b_{c_1} + \sum_{m=c_1+1}^{\infty} b_m \theta_2^{m-c_1}} = (1 - e^{-\beta x})^{c_1}
$$

and

$$
\lim_{\theta_1 \to 0^+} \lim_{\theta_2 \to 0^+} F(x) = 1 - \lim_{\theta_1 \to 0^+} \lim_{\theta_2 \to 0^+} \frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x})) / D(\theta_2))}{C(\theta_1)} = 1 - \lim_{\theta_1 \to 0^+} \frac{C(\theta_1 - \theta_1 (1 - e^{-\beta x})^{c_1})}{C(\theta_1)} \sum_{n=1}^{\infty} a_n \theta_1^n [1 - (1 - e^{-\beta x})^{c_1}]^n = 1 - \lim_{\theta_1 \to 0^+} \frac{a_{c_2} [1 - (1 - e^{-\beta x})^{c_1}]^{c_2} + \sum_{n=c_2+1}^{\infty} a_n \theta_1^{n-c_2} [1 - (1 - e^{-\beta x})^{c_1}]^n}{a_{c_2} + \sum_{n=c_2+1}^{\infty} a_n \theta_1^{n-c_2}} = 1 - [1 - (1 - e^{-\beta x})^{c_1}]^{c_2}.
$$

The proof is complete. □

Theorem 4.2.2. We have $f_X(x) \sim \theta_1 \theta_2 \beta e^{-\beta x} C'(0) D'(\theta_2)$ as $x \to \infty$ and $f_X(x) \sim \theta_1 \theta_2 \beta C'(\theta_1) D'(0)$ as $x \to 0$. So, the upper tail of the probability density function of $X$ decays exponentially while its lower tail approaches a constant.
Theorem 4.2.3. For the hazard rate function in (4.6), we have

$$\lim_{x \to 0} h(x) = \frac{\theta_1 \theta_2 \beta C'(\theta_1) D'(0)}{C' \left(\theta_1\right) D \left(\theta_2\right)}$$,

$$\lim_{x \to \infty} h(x) = 0$$.

So, both the initial and ultimate hazard rates are constants.

Proof: The proof for \(\lim_{x \to 0} h(x)\) follows immediately from Theorem 4.2.2. The proof for \(\lim_{x \to \infty} h(x)\) follows from (4.6) since

$$h(x) \sim \theta_1 \theta_2 \beta C'(0) D'(\theta_2) e^{-\beta x}$$ \hspace{1cm} (4.7)

as \(x \to \infty\). \(\square\)

Figure 4.1 shows possible shapes of the probability density function, (4.4), for all sixteen possible distributions of \((M, N)\). Figure 4.2 shows possible shapes of the hazard rate function, (4.6), for all sixteen possible distributions of \((M, N)\). We see that the probability density function can be monotonically decreasing or unimodal. The hazard rate function can be monotonically increasing, monotonically decreasing, bathtub shaped or upside down bathtub shaped.
Figure 4.1: Possible shapes of (4.4) when \((M,N)\) has the \((\text{Poisson, Poisson})\), \((\text{Poisson, Geometric})\), \((\text{Poisson, Logarithmic})\), \((\text{Poisson, Binomial})\), \((\text{Geometric, Poisson})\), \((\text{Geometric, Geometric})\), \((\text{Geometric, Logarithmic})\), \((\text{Geometric, Binomial})\), \((\text{Logarithmic, Poisson})\), \((\text{Logarithmic, Geometric})\), \((\text{Logarithmic, Logarithmic})\), \((\text{Logarithmic, Binomial})\), \((\text{Binomial, Poisson})\), \((\text{Binomial, Geometric})\), \((\text{Binomial, Logarithmic})\) and \((\text{Binomial, Binomial})\) distributions, arranged from top to bottom and left to right.
Figure 4.2: Possible shapes of (4.6) when \((M, N)\) has the (Poisson, Poisson), (Poisson, Geometric), (Poisson, Logarithmic), (Poisson, Binomial), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Geometric, Binomial), (Logarithmic, Poisson), (Logarithmic, Geometric), (Logarithmic, Logarithmic), (Logarithmic, Binomial), (Binomial, Poisson), (Binomial, Geometric), (Binomial, Logarithmic) and (Binomial, Binomial) distributions, arranged from top to bottom and left to right.
We now show that the survival function (4.5) can be expressed as an infinite linear combination of the survival function of the order statistics from the exponential power series distribution due to Chahkandi and Ganjali (5).

**Theorem 4.2.4.** Suppose $X_1, X_2, \ldots, X_n$ are independent and identical exponential power series random variables with their probability density function and cumulative distribution function specified by

$$h(x) = \frac{\theta_2^2 e^{-\beta x} D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)}$$

and

$$H(x) = \frac{D(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)},$$

respectively. Let $X^* = \min_{1 \leq i \leq n} X_i$. Then the survival function of $X \sim \text{PSEPS}(\theta_1, \beta, \theta_2)$ in (4.5) can be expressed as

$$F_X(x) = \sum_{n=1}^{\infty} \Pr(N = n) F_{X^*}(x).$$

**Proof:** First note that the survival function of $X^*$ can be expressed as

$$F_{X^*}(x) = \Pr \left( \min_{1 \leq i \leq n} X_i > x \right) = [\Pr(X_1 > x)]^n = [1 - H(x)]^n.$$

So,

$$\sum_{n=1}^{\infty} \Pr(N = n) F_{X^*}(x) = \sum_{n=1}^{\infty} \frac{a_n \theta_1^n}{C(\theta_1)} [1 - H(x)]^n = \frac{C'(\theta_1 [1 - H(x)])}{C'(\theta_1)},$$

which is equal to (4.5). The proof is complete. □
4.2.2 Quantiles and median

Quantiles are fundamental for estimation (for example, quantile estimators) and simulation. The $\xi$th quantile say $x_\xi$ of $X \sim \text{PSEPS}(\theta_1, \beta, \theta_2)$ is given by

$$x_\xi = G^{-1} \left\{ \frac{1}{\theta_2} D^{-1} \left( D (\theta_2) \left( 1 - \frac{C^{-1}(C(\theta_1)/(1 - \xi))}{\theta_1} \right) \right) \right\} ,$$

where $G^{-1}(y) = -\frac{1}{\beta} \log(1 - y)$, $0 < y < 1$, $C^{-1}(\cdot)$ denotes the inverse function of $C(\cdot)$, and $D^{-1}(\cdot)$ denotes the inverse function of $D(\cdot)$. In particular, the median is

$$x_{1/2} = G^{-1} \left\{ \frac{1}{\theta_2} D^{-1} \left( D (\theta_2) \left( 1 - \frac{C^{-1}(C(\theta_1)/2)}{\theta_1} \right) \right) \right\} .$$

We can use (4.8) to generate PSEPS variates.

4.2.3 Moment generating function

The moment generating function of $X \sim \text{PSEPS}(\theta_1, \beta, \theta_2)$ can be expressed as

$$M_X(t) = \sum_{n=1}^{\infty} \Pr(N = n) M_{X, \star}(t) ,$$
where

\[ M_X(t) = \int_0^\infty n e^{tx} h(x) [1 - H(x)]^{n-1} dx \]

\[ = - \int_0^\infty e^{tx} d((1 - H(x))^n) \]

\[ = \int_0^\infty d(e^{tx})(1 - H(x))^n \]

\[ = \int_0^\infty t e^{tx} (1 - H(x))^n dx \]

\[ = t \sum_{j=0}^n (-1)^j \binom{n}{j} I_H(j,t), \]

where

\[ I_H(j,t) = \int_0^\infty e^{tx} (H(x))^j dx. \]

### 4.2.4 Moments

Moment properties are fundamental for any distribution. For instance, the first four moments can be used to describe any data fairly well. Moments are also useful for estimation.

We can use \( M_X(t) \) to obtain \( E(X^k) \). But from direct calculations, we obtain

\[ E(X^k) = \int_0^\infty x^k f_X(x) dx \]

\[ = \sum_{n=1}^{\infty} n \Pr(N=n) \int_0^\infty x^k h(x) [1 - H(x)]^{n-1} dx \]

\[ = \sum_{n=1}^{\infty} n \Pr(N=n) \sum_{m=0}^{n-1} (-1)^m \binom{n-1}{m} \int_0^\infty x^k h(x) H^m(x) dx. \]
4.2.5 Order statistics

Order statistics have been used in a wide range of problems, including robust statistical estimation and detection of outliers, characterization of probability distributions and goodness-of-fit tests, entropy estimation, analysis of censored samples, reliability analysis, quality control and strength of materials.

Let $X_1, X_2, \ldots, X_n$ be independent and identical PSEPS($\theta_1, \beta, \theta_2$) random variables. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the corresponding order statistics. The probability density and cumulative distribution functions of the $i$th order statistic $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \frac{n! f_X(x)}{(n-i)!(i-1)!} \left[ 1 - \frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x}))/D(\theta_2))}{C(\theta_1)} \right]^{i-1}$$

and

$$F_{i:n}(x) = 1 - \frac{n!}{(n-i)!(i-1)!} \sum_{k=0}^{i-1} \frac{(-1)^k \binom{n-1}{k}}{n+k-i+1} \left[ \frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x}))/D(\theta_2))}{C(\theta_1)} \right]^{n+k-i+1},$$

respectively, for $x > 0$. An alternative expression for the latter is

$$F_{i:n}(x) = 1 - \frac{n!}{(n-i)!(i-1)!} \sum_{k=0}^{i-1} \frac{(-1)^{k \binom{n-1}{k}}}{k+i} \left[ 1 - \frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x}))/D(\theta_2))}{C(\theta_1)} \right]^{k+i}$$

for $x > 0$. 

"
4.3 Estimation and inference

Suppose \( x_1, x_2, \ldots, x_n \) is a random sample from the PSEPS(\( \theta_1, \beta, \theta_2 \)) distribution with all the parameters unknown. Here, we consider maximum likelihood estimation of the parameters.

The log-likelihood function of the parameters is

\[
\log L (\theta_1, \beta, \theta_2) = n \log (\theta_1 \theta_2 \beta) - \beta \sum_{i=1}^{n} x_i - n \log C (\theta_1) - n \log D (\theta_2) + \sum_{i=1}^{n} \log C' (\theta_1 - \theta_1 D (\theta_2 (1 - e^{-\beta x_i})) / D (\theta_2)) + \sum_{i=1}^{n} \log D' (\theta_2 (1 - e^{-\beta x_i})) \]  \( (4.9) \)

The normal equations are

\[
\frac{\partial \log L}{\partial \theta_1} = \frac{n}{\theta_1} - \frac{n C' (\theta_1)}{C (\theta_1)} + \sum_{i=1}^{n} \frac{C'' (\theta_1 - \theta_1 D (\theta_2 (1 - e^{-\beta x_i})) / D (\theta_2)) \left\{ 1 - D (\theta_2 (1 - e^{-\beta x_i})) \right\}}{D (\theta_2)} = 0,
\]

\[
\frac{\partial \log L}{\partial \theta_2} = \frac{n}{\theta_2} - \frac{n D' (\theta_2)}{D (\theta_2)} - \frac{\theta_1 D' (\theta_2)}{D (\theta_2)} \sum_{i=1}^{n} \frac{C'' (\theta_1 - \theta_1 D (\theta_2 (1 - e^{-\beta x_i})) / D (\theta_2)) D' (\theta_2 (1 - e^{-\beta x_i})) (1 - e^{-\beta x_i})}{D (\theta_2)} + \frac{\theta_1 D' (\theta_2)}{D^2 (\theta_2)} \sum_{i=1}^{n} \frac{C'' (\theta_1 - \theta_1 D (\theta_2 (1 - e^{-\beta x_i})) / D (\theta_2)) D (\theta_2 (1 - e^{-\beta x_i}))}{D (\theta_2) D (\theta_2)} + \sum_{i=1}^{n} \frac{D'' (\theta_2 (1 - e^{-\beta x_i}))}{D' (\theta_2 (1 - e^{-\beta x_i}))} (1 - e^{-\beta x_i}) = 0
\]
and

\[
\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} x_i - \frac{\theta_1 \theta_2}{D(\theta_2)} \sum_{i=1}^{n} C''(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x_i}))) / D(\theta_2) D'(\theta_2 (1 - e^{-\beta x_i})) x_i e^{-\beta x_i} \\
\theta_2 \sum_{i=1}^{n} \frac{D''(\theta_2 (1 - e^{-\beta x_i}))}{D'(\theta_2 (1 - e^{-\beta x_i}))} x_i e^{-\beta x_i} = 0.
\]

The maximum likelihood estimates of \((\theta_1, \theta_2, \beta)\) say \((\hat{\theta}_1, \hat{\theta}_2, \hat{\beta})\) are the simultaneous solutions of the equations \(\partial \log L / \partial \theta_1 = 0\), \(\partial \log L / \partial \theta_2 = 0\) and \(\partial \log L / \partial \beta = 0\).

Alternatively, the maximum likelihood estimates can be obtained by numerical maximization of (4.9). There are well established routines for numerical maximization like \texttt{nlm}\ or \texttt{optim}\ in the R statistical package (R Development Core Team (34)). Our numerical calculations showed that the surface of (4.9) was smooth for given smooth functions \(C(\cdot)\) and \(D(\cdot)\). The routines were able to locate the maximum of the likelihood surface for a wide range of smooth functions and for a wide range of starting values. However, to ease computations it is useful to have reasonable starting values. These can be obtained, for example, by the method of moments. For \(r = 1, 2, 3\), let \(m_r = n^{-1} \sum_{i=1}^{n} x_i^r\) denote the first three sample moments. Equating these moments with the theoretical versions given in Section 4.2.4, we have \(m_r = E(X^r)\) for \(r = 1, 2, 3\). These equations can be solved simultaneously to obtain the moments estimates.

In the simulation and real data sections later on, the maximum likelihood estimates were obtained by numerical maximization of (4.9). The \texttt{optim}\ routine in R was used for numerical maximization. The starting values of the parameters for numerical maximization were taken to correspond to all combinations of

- \(\theta_1 = 0.01, 0.02, \ldots, 10\), \(\theta_2 = 0.01, 0.02, \ldots, 10\) and \(\beta = 0.01, 0.02, \ldots, 10\) if the distribution of \((M, N)\) is \((\text{Poisson, Poisson})\), \((\text{Poisson, Binomial})\), \((\text{Binomial, Poisson})\) or
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

(Binomial, Binomial);

- \( \theta_1 = 0.01, 0.02, \ldots, 10 \), \( \theta_2 = 0.01, 0.02, \ldots, 0.99 \) and \( \beta = 0.01, 0.02, \ldots, 10 \) if the distribution of \((M, N)\) is (Poisson, Geometric), (Binomial, Geometric), (Poisson, Logarithmic) or (Binomial, Logarithmic);

- \( \theta_1 = 0.01, 0.02, \ldots, 0.99 \), \( \theta_2 = 0.01, 0.02, \ldots, 10 \) and \( \beta = 0.01, 0.02, \ldots, 10 \) if the distribution of \((M, N)\) is (Geometric, Poisson), (Geometric, Binomial), (Logarithmic, Poisson) or (Logarithmic, Binomial);

- \( \theta_1 = 0.01, 0.02, \ldots, 0.99 \), \( \theta_2 = 0.01, 0.02, \ldots, 0.99 \) and \( \beta = 0.01, 0.02, \ldots, 10 \) if the distribution of \((M, N)\) is (Geometric, Geometric), (Geometric, Logarithmic), (Logarithmic, Geometric) or (Logarithmic, Logarithmic).

The \texttt{optim} routine converged all the time and converged to a unique maximum all the time.

For interval estimation of \((\theta_1, \theta_2, \beta)\) and tests of hypothesis, one requires the Fisher information matrix. We can express the observed Fisher information matrix of \((\theta_1, \theta_2, \beta)\) as

\[
J = - \begin{pmatrix}
J_{1,1} & J_{1,2} & J_{1,3} \\
J_{1,2} & J_{2,2} & J_{2,3} \\
J_{1,3} & J_{2,3} & J_{3,3}
\end{pmatrix},
\]

where

\[
J_{1,1} = - \frac{n}{\theta_1^2} - n \left[ \frac{C''(\theta_1)}{C(\theta_1)} \right]^2 + n \left[ \frac{C'(\theta_1)}{C(\theta_1)} \right]^2 \\
\quad + \sum_{i=1}^n \left[ \frac{C''(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x_i})))}{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x_i}))) / D(\theta_2)} \right] \left\{ 1 - \frac{D(\theta_2 (1 - e^{-\beta x_i}))}{D(\theta_2)} \right\}^2 \\
\quad - \sum_{i=1}^n \left[ \frac{C''(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x_i})))}{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta x_i}))) / D(\theta_2)} \right] \left\{ 1 - \frac{D(\theta_2 (1 - e^{-\beta x_i}))}{D(\theta_2)} \right\}^2.
\]
\[ J_{1,2} = -\frac{\theta_1}{D(\theta_2)} \sum_{i=1}^{n} C''' \left( \theta_1 - \theta_i D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) / D(\theta_2) \right) \left\{ 1 - \frac{D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right)}{D(\theta_2)} \right\} D' \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \left( 1 - e^{-\beta x_i} \right) \]

\[ + \frac{\theta_1}{D^2(\theta_2)} \sum_{i=1}^{n} C'' \left( \theta_1 - \theta_i D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) / D(\theta_2) \right) \left\{ 1 - \frac{D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right)}{D(\theta_2)} \right\} D' \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \left( 1 - e^{-\beta x_i} \right) \]

\[ + \frac{\theta_1 D' \left( \theta_2 \right)}{D^2(\theta_2)} \sum_{i=1}^{n} C'' \left( \theta_1 - \theta_i D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) / D(\theta_2) \right) \left\{ 1 - \frac{D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right)}{D(\theta_2)} \right\} D' \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \left( 1 - e^{-\beta x_i} \right) \]

\[ - \frac{1}{D(\theta_2)} \sum_{i=1}^{n} C'' \left( \theta_1 - \theta_i D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) / D(\theta_2) \right) \left\{ 1 - \frac{D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right)}{D(\theta_2)} \right\} D' \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \left( 1 - e^{-\beta x_i} \right) \]

\[ + \frac{D' \left( \theta_2 \right)}{D^2(\theta_2)} \sum_{i=1}^{n} C'' \left( \theta_1 - \theta_i D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) / D(\theta_2) \right) \left\{ 1 - \frac{D \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right)}{D(\theta_2)} \right\} D' \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \left( 1 - e^{-\beta x_i} \right). \]
\[ J_{2,2} = -\frac{n}{\theta_2^2} - n \frac{D''(\theta_2)}{D(\theta_2)} + n \sum_{i=1}^{\infty} \frac{D''(\theta_2)}{D(\theta_2)} \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^2(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^2(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^2(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^3(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^3(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]

\[ + \frac{\theta_1^2}{D^3(\theta_2)} \sum_{i=1}^{\infty} \frac{C_{n+2}'''}{C_n'''} \left( \frac{\theta_1 - \theta_2 D'(\theta_2 (1 - e^{-\beta x}))}{D(\theta_2)} \right)^2 \left( D'(\theta_2 (1 - e^{-\beta x})) \right)^2 \left( 1 - e^{-\beta x} \right)^2 \]
\[ J_{2,3} = \frac{\theta_1^2 \theta_2}{D^2(\theta_2)} \sum_{i=1}^{n} \left[ \left( \theta_1 - \theta_1 \frac{D'}{D(\theta_2)} \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \right)^T \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \right] - \frac{\theta_1^2 \theta_2}{D(\theta_2)} \sum_{i=1}^{n} \left[ \left( \theta_1 - \theta_1 \frac{D'}{D(\theta_2)} \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \right)^T \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) \right] \left( \theta_2 \left( 1 - e^{-\beta x_i} \right) \right) x_i e^{-\beta x_i} \]

For large \( n \), the distribution of \( \sqrt{n} \left( \hat{\theta}_1 - \theta_1, \hat{\theta}_2 - \theta_2, \hat{\beta} - \beta \right) \) approximates to a trivariate normal distribution with zero means and variance-covariance matrix \( J^{-1} \). The properties of \( \left( \hat{\theta}_1, \hat{\theta}_2, \hat{\beta} \right) \) can be derived based on this normal approximation.
4.4 Simulation study

Here, we assess the performance of the maximum likelihood estimates \( \hat{\theta}_1, \hat{\theta}_2, \hat{\beta} \) with respect to sample size \( n \). The assessment is based on a simulation study:

1. generate ten thousand samples of size \( n \) from (4.4). The inversion method was used to generate samples, i.e., variates of the PSEPS distribution were generated using

\[
\frac{C(\theta_1 - \theta_1 D(\theta_2 (1 - e^{-\beta X}))/D(\theta_2))}{C(\theta_1)} = 1 - U,
\]

where \( U \sim U(0, 1) \) is a uniform variate on the unit interval;

2. compute the maximum likelihood estimates for the ten thousand samples, say \( \hat{\theta}_{1,i}, \hat{\theta}_{2,i}, \hat{\beta}_i \) for \( i = 1, 2, \ldots, 10000 \);

3. compute the biases and mean squared errors given by

\[
\text{bias}_f(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{f}_i - f)
\]

and

\[
\text{MSE}_f(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{f}_i - f)^2
\]

for \( f = \theta_1, \theta_2, \beta \).

We repeated these steps for \( n = 10, 11, \ldots, 100 \) with \( \theta_1 = 1, \theta_2 = 1, \beta = 1, C(\theta) = e^\theta - 1 \) and \( D(\theta) = e^\theta - 1 \), so computing \( \text{bias}_{\theta_1}(n), \text{bias}_{\theta_2}(n), \text{bias}_{\beta}(n) \) and \( \text{MSE}_{\theta_1}(n), \text{MSE}_{\theta_2}(n), \text{MSE}_{\beta}(n) \) for \( n = 10, 11, \ldots, 100 \).
Figure 4.3: bias_{\theta_1}(n) (top left), bias_{\theta_2}(n) (top right) and bias_{\beta}(n) (bottom left) versus \( n = 10, 11, \ldots, 100 \).
Figures 4.3 and 4.4 show how the three biases and the three mean squared errors vary with respect to $n$. The broken lines in Figure 4.3 correspond to the biases being zero. The following observations can be made:

- the biases for each parameter decrease to zero as $n \to \infty$;
- the mean squared errors for each parameter decrease to zero as $n \to \infty$;
- the biases for each parameter are generally positive;
- the biases appear largest for the parameter, $\beta$;
- the biases appear smallest for the parameter, $\theta_2$;
- the mean squared errors appear largest for the parameter, $\beta$;
the mean squared errors appear smallest for the parameter, $\theta_2$;

- the biases and mean squared errors for each parameter appear reasonably small for all $n \geq 40$.

We have presented results for only one choice for $\theta_1$, $\theta_2$, $\beta$, $C(\theta)$ and $D(\theta)$, namely that $\theta_1 = 1$, $\theta_2 = 1$, $\beta = 1$, $C(\theta) = e^\theta - 1$ and $D(\theta) = e^\theta - 1$. But the results were similar for a wide range of other choices. In particular, the biases always approached zero as $n$ increased, the mean squared errors always decreased to zero as $n$ increased, and the biases and mean squared errors for each parameter always appeared reasonably small for all $n \geq 40$.

### 4.5 Illustrations using failure data sets

Here, we show that the PSEPS distribution performs better than five of the recently introduced distributions: the three-parameter generalized exponential geometric distribution due to Silva et al. (35), the four-parameter beta Fréchet distribution due to Barreto-Souza et al. (36), the three-parameter Weibull geometric distribution due to Barreto-Souza et al. (37), the three-parameter gamma-exponentiated exponential distribution due to Ristic and Balakrishnan (38), and the three-parameter generalized exponential geometric distribution due to Bidram et al. (39).

Each of these five distributions was proposed by showing that it performed better than several of the known distributions in the literature for at least one real data set. We show that the PSEPS distribution provides better fits than these five distributions for the respective data sets used in Silva et al. (35), Barreto-Souza et al. (36) (37), Ristic and Balakrishnan (38) and Bidram et al. (39).

Each distribution was fitted by the method of maximum likelihood. The PSEPS distribution was fitted by taking $(M, N)$ to follow the (Poisson, Poisson), (Poisson, Geometric),
(Poisson, Logarithmic), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Logarithmic, Poisson), (Logarithmic, Geometric), and (Logarithmic, Logarithmic) distributions. Each of these choices for \((M, N)\) allows for three parameters. We have not taken \((M, N)\) to follow (Poisson, Binomial), (Geometric, Binomial), (Logarithmic, Binomial), (Binomial, Poisson), (Binomial, Geometric), (Binomial, Logarithmic) or (Binomial, Binomial). Each of these choices for \((M, N)\) would allow for at least four parameters. We want to illustrate how the simplest of the PSEPS distribution compares to other known distributions.

The parameter estimates, associated standard errors, log-likelihood values and \(p\)-values of the Kolmogorov-Smirnov statistic for goodness of fit of the fitted distributions are shown in Tables 4.2, 4.3 and 4.4. The standard errors were computed by inverting the observed information matrix, see Section 4.3. All of the fitted distributions but the beta Fréchet distribution have the same number of parameters. So, the distribution giving the best fit can be chosen as the one giving the largest log-likelihood value or the one giving the largest \(p\)-value.

The data sets considered in Sections 4.5.1 to 4.5.3 are on fatigue life of 6061-T6 aluminum, breaking stress of carbon fibers and intervals, in service hours, between failures of an air-conditioning equipment. In these data sets, the 6061-T6 aluminum, carbon fiber and air-conditioning equipment can each be regarded as a “system”. According to Ross (40), any system can be represented either as a series arrangement of parallel structures or as a parallel arrangement of series structures. Hence, the 6061-T6 aluminum, carbon fiber and air-conditioning equipment can each be regarded as a series arrangement of parallel structures. It is reasonable to suppose that the number of units in the system, \(N\), and the number of sub-units in each unit, \(M\), are random variables because of factors like type of fiber, weight of fiber, length of fiber, type of aluminum, weight of aluminum, type of air-conditioning equipment, weight of air-conditioning equipment, etc. For simplicity, we
suppose \( M \) and \( N \) are independent random variables.

4.5.1 Generalized exponential geometric (Silva et al. (35)) and generalized exponential geometric (Bidram et al. (39)) distributions

The generalized exponential geometric (GEG1) (Silva et al. (35)) and generalized exponential geometric (GEG2) (Bidram et al. (39)) distributions are specified by the probability density functions

\[
f(x) = a b (1 - p) e^{-b x} \frac{(1 - e^{-b x})^{a-1}}{(1 - p e^{-b x})^{a+1}}
\]

and

\[
f(x) = a b (1 - p) e^{-b x} \frac{(1 - e^{-b x})^{a-1}}{[1 - p (1 - e^{-b x})^a]^2}
\]

respectively, for \( x > 0, \ a > 0, \ b > 0 \) and \( 0 < p < 1 \). These distributions were proposed by showing that they gave better fits than several known distributions for a data set on the fatigue life of 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at eighteen cycles per second. The data set consists of one hundred and one observations with maximum stress per cycle 31,000 psi:

CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

151, 151, 152, 155, 156, 157, 157, 157, 158, 159, 162, 163, 163, 164, 166, 166, 168, 170, 174, 201, 212

The data set was given by Birnbaum and Saunders (41).

Table 4.2: Fitted models, parameter estimates, log-likelihood values and p-values of the Kolmogorov-Smirnov statistic for the maximum stress data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter estimates (see)</th>
<th>$-\log L$</th>
<th>K-S p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEG1</td>
<td>$\hat{a} = 1.110(2.112)$, $\hat{b} = 1.493 \times 10^{-3}$ (5.243 \times 10^{-3}), $\hat{p} = 5.604 \times 10^{-1}$ (9.320 \times 10^{-2})</td>
<td>522.333</td>
<td>0.65</td>
</tr>
<tr>
<td>GEG2</td>
<td>$\hat{a} = 1.009(3.287)$, $\hat{b} = 1.476 \times 10^{-1}$ (5.640 \times 10^{-1}), $\hat{p} = 9.999 \times 10^{-1}$ (5.433 \times 10^{-1})</td>
<td>766.651</td>
<td>0.04</td>
</tr>
<tr>
<td>Poisson-Poisson</td>
<td>$\hat{\theta}_1 = 3.703 \times 10^{-7}$ (2.396 \times 10^{-6}), $\hat{\theta}_2 = 9.610(1.184)$, $\hat{\lambda} = 2.292 \times 10^{-2}$ (8.323 \times 10^{-1})</td>
<td>513.690</td>
<td>0.72</td>
</tr>
<tr>
<td>Poisson-Geometric</td>
<td>$\hat{\theta}_1 = 7.174 \times 10^{-7}$ (1.394 \times 10^{-6}), $\hat{\theta}_2 = 9.602 \times 10^{-1}$ (1.240 \times 10^{-2}), $\hat{\lambda} = 2.451 \times 10^{-2}$ (3.020 \times 10^{-3})</td>
<td>517.868</td>
<td>0.71</td>
</tr>
<tr>
<td>Poisson-Logarithmic</td>
<td>$\hat{\theta}_1 = 8.136 \times 10^{-7}$ (8.565 \times 10^{-6}), $\hat{\theta}_2 = 8.871 \times 10^{-1}$ (3.488 \times 10^{-2}), $\hat{\lambda} = 1.062 \times 10^{-2}$ (7.143 \times 10^{-1})</td>
<td>581.194</td>
<td>0.25</td>
</tr>
<tr>
<td>Geometric-Poisson</td>
<td>$\hat{\theta}_1 = 2.022 \times 10^{-6}$ (1.568 \times 10^{-1}), $\hat{\theta}_2 = 1.081(4.334)$, $\hat{\lambda} = 1.013 \times 10^{-2}$ (2.395 \times 10^{-1})</td>
<td>579.805</td>
<td>0.31</td>
</tr>
<tr>
<td>Geometric-Geometric</td>
<td>$\hat{\theta}_1 = 3.504 \times 10^{-7}$ (1.649 \times 10^{-1}), $\hat{\theta}_2 = 5.012 \times 10^{-1}$ (6.396 \times 10^{-1}), $\hat{\lambda} = 1.287 \times 10^{-2}$ (9.758 \times 10^{-4})</td>
<td>577.665</td>
<td>0.33</td>
</tr>
<tr>
<td>Geometric-Logarithmic</td>
<td>$\hat{\theta}_1 = 3.039 \times 10^{-9}$ (1.716 \times 10^{-3}), $\hat{\theta}_2 = 2.232 \times 10^{-1}$ (5.035 \times 10^{-1}), $\hat{\lambda} = 9.859 \times 10^{-3}$ (1.123 \times 10^{-3})</td>
<td>594.340</td>
<td>0.11</td>
</tr>
<tr>
<td>Logarithmic-Poisson</td>
<td>$\hat{\theta}_1 = 1.748 \times 10^{-10}$ (8.845 \times 10^{-8}), $\hat{\theta}_2 = 6.703(8.326 \times 10^{-1})$, $\hat{\lambda} = 2.164 \times 10^{-2}$ (1.300 \times 10^{-3})</td>
<td>529.053</td>
<td>0.61</td>
</tr>
<tr>
<td>Logarithmic-Geometric</td>
<td>$\hat{\theta}_1 = 3.561 \times 10^{-9}$ (2.635 \times 10^{-1}), $\hat{\theta}_2 = 7.313 \times 10^{-1}$ (5.833 \times 10^{-2}), $\hat{\lambda} = 1.380 \times 10^{-2}$ (9.882 \times 10^{-3})</td>
<td>560.452</td>
<td>0.44</td>
</tr>
<tr>
<td>Logarithmic-Logarithmic</td>
<td>$\hat{\theta}_1 = 1.668 \times 10^{-7}$ (2.388 \times 10^{-1}), $\hat{\theta}_2 = 2.848 \times 10^{-1}$ (4.591 \times 10^{-2}), $\hat{\lambda} = 9.699 \times 10^{-3}$ (7.830 \times 10^{-4})</td>
<td>592.455</td>
<td>0.19</td>
</tr>
</tbody>
</table>

We see from Table 4.2 that the PSEPS distribution with \((M, N)\) specified by the (Poisson, Poisson) and (Poisson, Geometric) distributions provides better fits than the GEG1 distribution. The PSEPS distribution with \((M, N)\) specified by the (Poisson, Poisson), (Poisson,
Geometric), (Poisson, Logarithmic), (Geometric, Poisson), (Geometric, Geometric), (Geometric, Logarithmic), (Logarithmic, Poisson), (Logarithmic, Geometric), and (Logarithmic, Logarithmic) distributions provides better fits than the GEG2 distribution. All of the fitted distributions but the GEG2 distribution provide adequate fits at the five percent level of significance.

The distribution giving the largest log-likelihood value and the largest $p$-value is the PSEPS distribution with $(M, N)$ specified by the (Poisson, Poisson) distribution. The parameter estimates of this distribution suggest that the aluminum coupon has a mean number of 1.0 units working in series and that each unit has a mean number of 9.6 sub-units working in parallel and that each sub-unit has an average failure time of 43.6.

4.5.2 Beta Fréchet (Barreto-Souza et al. (36)) and Weibull geometric (Barreto-Souza et al. (37)) distributions

The beta Fréchet (Barreto-Souza et al. (36)) and Weibull geometric (Barreto-Souza et al. (37)) distributions are specified by the probability density functions

$$f(x) = \frac{\lambda^b e^{-a(\beta/x)^b}}{x^{a+1} B(a, b)}$$

and

$$f(x) = \frac{ab^a (1 - p)x^{a-1}e^{-(bx)^a}}{[1 - pe^{-(bx)^a}]^2},$$

respectively, for $x > 0$, $a > 0$, $b > 0$, $\beta > 0$, $\lambda > 0$ and $x > 0$, $a > 0$, $b > 0$, $0 < p < 1$, respectively. These distributions were proposed by showing that they gave better fits than several known distributions for the following data on breaking stress of one hundred carbon fibers (in GPa):
3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11,
4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90,
3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22,
3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56,
3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83, 1.92,
1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59,
3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00, 1.22, 1.12, 1.71,
2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38,
1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.80,
1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65

The data is due to Nichols and Padgett (42).
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter estimates (aux)</th>
<th>$-\log L$</th>
<th>K-S p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta Fréchet</td>
<td>$\hat{a} = 35.088(40.132)$, $\hat{b} = 194.676(201.019)$, $\hat{\theta} = 72.300(132.071)$, $\hat{\beta} = 1.862 \times 10^{-1} \left(9.023 \times 10^{-1}\right)$</td>
<td>149.026</td>
<td>0.71</td>
</tr>
<tr>
<td>Weibull geometric</td>
<td>$\hat{a} = 1.148 \left(1.119 \times 10^{-1}\right)$, $\hat{\beta} = 1.909 \times 10^{-1} \left(6.446 \times 10^{-2}\right)$, $\hat{\theta} = 2.557 \times 10^{-6} \left(2.331 \times 10^{-1}\right)$</td>
<td>185.198</td>
<td>0.59</td>
</tr>
<tr>
<td>Poisson-Poisson</td>
<td>$\hat{\theta}_1 = 6.355 \times 10^{-6}(4.175)$, $\hat{\theta}_2 = 12.215(2.371)$, $\hat{\beta} = 1.150 \left(5.799 \times 10^{-1}\right)$</td>
<td>144.652</td>
<td>0.73</td>
</tr>
<tr>
<td>Poisson-Logarithmic</td>
<td>$\hat{\theta}_1 = 7.864 \times 10^{-4}(2.309)$, $\hat{\theta}_2 = 9.868 \times 10^{-1} \left(1.492 \times 10^{-2}\right)$, $\hat{\beta} = 1.676 \left(1.513 \times 10^{-1}\right)$</td>
<td>142.123</td>
<td>0.78</td>
</tr>
<tr>
<td>Geometric-Logarithmic</td>
<td>$\hat{\theta}_1 = 1.502 \times 10^{-6} \left(3.459 \times 10^{-5}\right)$, $\hat{\theta}_2 = 8.863 \times 10^{-1} \left(1.370 \times 10^{-1}\right)$, $\hat{\beta} = 6.538 \times 10^{-1} \left(1.776 \times 10^{-1}\right)$</td>
<td>179.020</td>
<td>0.66</td>
</tr>
<tr>
<td>Logarithmic-Poisson</td>
<td>$\hat{\theta}_1 = 4.329 \times 10^{-9} \left(1.374 \times 10^{-3}\right)$, $\hat{\theta}_2 = 1.063 \left(5.285 \times 10^{-1}\right)$, $\hat{\beta} = 8.109 \times 10^{-1} \left(1.039 \times 10^{-1}\right)$</td>
<td>202.226</td>
<td>0.31</td>
</tr>
<tr>
<td>Geometric-Logarithmic</td>
<td>$\hat{\theta}_1 = 6.054 \times 10^{-8} \left(1.297 \times 10^{-1}\right)$, $\hat{\theta}_2 = 2.775 \times 10^{-1} \left(8.443 \times 10^{-1}\right)$, $\hat{\beta} = 8.071 \times 10^{-1} \left(7.020 \times 10^{-2}\right)$</td>
<td>212.418</td>
<td>0.23</td>
</tr>
<tr>
<td>Logarithmic-Geometric</td>
<td>$\hat{\theta}_1 = 2.272 \times 10^{-7} \left(1.258 \times 10^{-1}\right)$, $\hat{\theta}_2 = 1.876 \times 10^{-1} \left(7.099 \times 10^{-1}\right)$, $\hat{\beta} = 7.941 \times 10^{-1} \left(8.266 \times 10^{-2}\right)$</td>
<td>224.390</td>
<td>0.11</td>
</tr>
<tr>
<td>Logarithmic-Logarithmic</td>
<td>$\hat{\theta}_1 = 8.394 \times 10^{-8} \left(2.222 \times 10^{-1}\right)$, $\hat{\theta}_2 = 1.114961 \left(6.378 \times 10^{-1}\right)$, $\hat{\beta} = 6.857 \times 10^{-1} \left(1.007 \times 10^{-1}\right)$</td>
<td>189.879</td>
<td>0.55</td>
</tr>
<tr>
<td>Logarithmic-Geometric</td>
<td>$\hat{\theta}_1 = 1.091 \times 10^{-6} \left(2.339 \times 10^{-1}\right)$, $\hat{\theta}_2 = 5.594 \times 10^{-1} \left(5.922 \times 10^{-1}\right)$, $\hat{\beta} = 7.034 \times 10^{-1} \left(3.248 \times 10^{-2}\right)$</td>
<td>180.988</td>
<td>0.65</td>
</tr>
<tr>
<td>Logarithmic-Geometric</td>
<td>$\hat{\theta}_1 = 1.632 \times 10^{-8} \left(2.533 \times 10^{-1}\right)$, $\hat{\theta}_2 = 8.226 \times 10^{-1} \left(1.222 \times 10^{-1}\right)$, $\hat{\beta} = 5.547 \times 10^{-1} \left(1.145 \times 10^{-1}\right)$</td>
<td>182.878</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Table 4.3: Fitted models, parameter estimates, log-likelihood values and $p$-values of the Kolmogorov-Smirnov statistic for the fiber data.

We see from Table 4.3 that the PSEPS distribution with $(M, N)$ specified by the (Poisson, Poisson) and (Poisson, Geometric) distributions provides better fits than the beta Fréchet distribution. The PSEPS distribution with $(M, N)$ specified by the (Poisson, Poisson), (Poisson, Geometric), (Poisson, Logarithmic), (Logarithmic, Geometric), and (Logarithmic, Logarithmic) distributions provides better fits than the Weibull geometric distribution. The beta Fréchet distribution has one more parameter than the PSEPS distribution. Nonetheless
all of the fitted distributions provide adequate fits at the five percent level of significance.

The distribution giving the largest log-likelihood value and the largest $p$-value is the PSEPS distribution with $(M, N)$ specified by the (Poisson, Geometric) distribution. The parameter estimates of this distribution suggest that the fiber has a mean number of 1.00 units working in series and that each unit has a mean number of 1.01 sub-units working in parallel and that each sub-unit has an average failure time of 0.60.

4.5.3 Gamma-exponentiated exponential (Ristic and Balakrishnan (38)) distribution

The gamma-exponentiated exponential (GEE) distribution (Ristic and Balakrishnan (38)) is specified by the probability density function

$$f(x) = \frac{\lambda a^b}{\Gamma(b)} e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{a-1} \left[-\log \left(1 - e^{-\lambda x}\right)\right]^{b-1}$$

for $x > 0$, $a > 0$, $b > 0$ and $\lambda > 0$. This distribution was proposed by showing that it gave better fits than several known distributions for the following data set on the intervals, in service hours, between failures of the air-conditioning equipment in a Boeing 720 aircraft: 50, 44, 102, 72, 22, 39, 3, 15, 197, 188, 79, 88, 46, 5, 5, 36, 22, 139, 210, 97, 30, 23, 13, 14. The data was reported in Proschan (43).
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter estimates (aux)</th>
<th>– log L</th>
<th>K-S p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEE</td>
<td>( \hat{a} = 2.047(3.025) ), ( \hat{b} = 5.207(16.493) ), ( \hat{\lambda} = 2.261 \times 10^{-4}(1.055) )</td>
<td>13.155</td>
<td>0.06</td>
</tr>
<tr>
<td>Poisson-Poisson</td>
<td>( \hat{\beta}_1 = 1.301 \times 10^{-5} (6.348 \times 10^{-5}) ), ( \hat{\beta}_2 = 1.967 \times 10^{-6} (8.406 \times 10^{-7}) ), ( \hat{\lambda} = 1.510 (6.088 \times 10^{-1}) )</td>
<td>13.331</td>
<td>0.04</td>
</tr>
<tr>
<td>Poisson-Geometric</td>
<td>( \hat{\beta}_1 = 1.185 \times 10^{-5} (5.528) ), ( \hat{\beta}_2 = 3.749 \times 10^{-6} (2.768) ), ( \hat{\lambda} = 1.515 (6.118 \times 10^{-1}) )</td>
<td>13.331</td>
<td>0.04</td>
</tr>
<tr>
<td>Poisson-Logarithmic</td>
<td>( \hat{\beta}_1 = 1.391 \times 10^{-5} (1.944 \times 10^{-5}) ), ( \hat{\beta}_2 = 1.346 \times 10^{-7} (7.222 \times 10^{-8}) ), ( \hat{\lambda} = 1.503 (6.076 \times 10^{-1}) )</td>
<td>13.331</td>
<td>0.04</td>
</tr>
<tr>
<td>Geometric-Poisson</td>
<td>( \hat{\beta}_1 = 6.513 \times 10^{-2} (3.155) ), ( \hat{\beta}_2 = 1.141 \times 10^{-4} (6.613) ), ( \hat{\lambda} = 1.508 (6.092 \times 10^{-1}) )</td>
<td>13.331</td>
<td>0.04</td>
</tr>
<tr>
<td>Geometric-Geometric</td>
<td>( \hat{\beta}_1 = 1.702 \times 10^{-1} (13.653) ), ( \hat{\beta}_2 = 1.125 \times 10^{-1} (14.604) ), ( \hat{\lambda} = 1.508 (6.092 \times 10^{-1}) )</td>
<td>13.331</td>
<td>0.04</td>
</tr>
<tr>
<td>Geometric-Logarithmic</td>
<td>( \hat{\beta}_1 = 7.567 \times 10^{-1} (4.399 \times 10^{-1}) ), ( \hat{\beta}_2 = 9.999 \times 10^{-1} (4.991 \times 10^{-1}) ), ( \hat{\lambda} = 5.473 (6.311) )</td>
<td>12.482</td>
<td>0.11</td>
</tr>
<tr>
<td>Logarithmic-Poisson</td>
<td>( \hat{\beta}_1 = 5.948 \times 10^{-1} (6.305) ), ( \hat{\beta}_2 = 7.704 \times 10^{-1} (15.830) ), ( \hat{\lambda} = 1.511 (5.609 \times 10^{-1}) )</td>
<td>13.330</td>
<td>0.04</td>
</tr>
<tr>
<td>Logarithmic-Geometric</td>
<td>( \hat{\beta}_1 = 8.180 \times 10^{-1} (9.349 \times 10^{-1}) ), ( \hat{\beta}_2 = 5.986 \times 10^{-1} (1.358) ), ( \hat{\lambda} = 1.630 (9.410 \times 10^{-1}) )</td>
<td>13.318</td>
<td>0.08</td>
</tr>
<tr>
<td>Logarithmic-Logarithmic</td>
<td>( \hat{\beta}_1 = 9.194 \times 10^{-1} (4.663 \times 10^{-1}) ), ( \hat{\beta}_2 = 9.999 \times 10^{-1} (7.445 \times 10^{-1}) ), ( \hat{\lambda} = 5.499 (4.433) )</td>
<td>12.710</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 4.4: Fitted models, parameter estimates, log-likelihood values and p-values of the Kolmogorov-Smirnov statistic for the air conditioning data.

We see from Table 4.4 that the PSEPS distribution with \((M, N)\) specified by the (Geometric, Logarithmic) and (Logarithmic, Logarithmic) distributions provides better fits than the GEE distribution. Only the GEE distribution and the PSEPS distributions with \((M, N)\) specified by the (Geometric, Logarithmic), (Logarithmic, Geometric) and (Logarithmic, Logarithmic) distributions provide adequate fits at the five percent level of significance. These results however should be treated conservatively because Section 4.4 showed that the maximum likelihood estimators do not perform well for \(n < 40\).

The distribution giving the largest log-likelihood value and the largest \(p\)-value is the
CHAPTER 4. THE POWER SERIES EXPONENTIAL POWER SERIES DISTRIBUTIONS

PSEPS distribution with $(M, N)$ specified by the (Geometric, Logarithmic) distribution. The parameter estimates of this distribution suggest that the air-conditioning equipment in the Boeing 720 aircraft has a mean number of 144.6 units working in series and that each unit has a mean number of 1.3 sub-units working in parallel and that each sub-unit has an average failure time of 0.18.

4.6 Conclusions

We have introduced a new distribution referred to as the power series exponential power series distribution. We have derived mathematical expressions for its quantile, moment generating function, moments and order statistic properties. We have discussed maximum likelihood estimation of its parameters and derived the associated observed information matrix.

We have assessed the finite sample performance of the maximum likelihood estimators by a simulation study. This study shows that the estimators perform reasonably well for all sample sizes greater than forty.

We have compared the fit of the power series exponential power series distribution with five other distributions using three failure data sets. Four of the distributions have three parameters each. The other has four parameters. The new distribution was shown to give better fits than all five distributions for all three failure data sets.

A future work is to propose bivariate, multivariate, complex variate and matric variate generalizations of the power series exponential power series distribution.
Bibliography


