2011-01-01

Gamma and Generalized Gamma Distributions

Victor Hugo Jiménez Nava
University of Texas at El Paso, vhjimeneznava@miners.utep.edu

GAMMA AND GENERALIZED GAMMA DISTRIBUTIONS

VÍCTOR HUGO JIMÉNEZ NAVA

Department of Mathematical Science

APPROVED:

Panagis Moschopoulos, Ph.D., Chair

Ori Rosen, Ph.D.

Naijun Sha, Ph.D.

Dr. Max Shpak, Ph.D

Benjamín C. Flores, Ph.D.
Acting Dean of the Graduate School

to my

MOTHER and FATHER

with love

GAMMA AND GENERALIZED GAMMA DISTRIBUTIONS

by

VÍCTOR HUGO JIMÉNEZ NAVA

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Mathematical Science

THE UNIVERSITY OF TEXAS AT EL PASO

December 2011

Acknowledgements

I would like to express my deep-felt gratitude to my advisor, Dr. Panagis Moschopoulos of the Mathematical Science Department at The University of Texas at El Paso, for his advice, encouragement, guidance and constant support. I believe he has been more than patient. He has always tried to lead me in the right direction and all this time working with him I have learned a lot. He has always tried of providing clear explanations when I was confused and lost. But he was there every time dealing with my impertinences when working and giving me his time, in spite of anything else that was going on. I really appreciate all he did. I do not find words to thank him for all his patience.

I also wish to thank the other members of my committee, Dr. Naijun Sha, Dr. Ori Rosen, both of the Mathematical Science Department at The University of Texas at El Paso and Dr. Max Shpak of the Biological Science Department of the University of Texas at el Paso too. Their suggestions, comments and additional guidance were a big support to the completion of this work. I would like to thank as well to the Government of the State of Chihuahua Mexico for providing me with the financial support to study in this great university. They supported me with no complaints and always trying to make me a better citizen not only for the Chihuhua State but for all Mexico.

I would like to thank as well to my parents for all his support in order to complete this master degree successfully. During all this time they have given to me their help, cooperation and motivating me to continue in this task of my life. I really thank them for all what they have done.

Additionally, I want to thank The University of Texas at El Paso Mathematical Science Department professors and staff for all their hard work and dedication, providing me the means to complete my degree and prepare for a career as a statistician.

NOTE: This thesis was submitted to my Supervising Committee on the November 30, 2011.

Contents

	Page
Acknowledgements	iv
Table of Contents	v
List of Tables	vii
List of Figures	x
Chapter	
1 Introduction	1
2 The Gamma Distribution and its Basic Properties	3
2.1 History and Motivation	3
2.2 The Gamma Function	3
2.3 Getting the Gamma Density	5
2.4 Mean, Variance and Moment Generating Function	5
2.5 Particular Cases and Shapes	7
2.5.1 Standard Gamma	7
2.5.2 The Exponential	8
2.5.3 The Chi square distribution	9
2.5.4 Linear Combinations of Independent Chi-Square	10
3 Estimation of Parameters	12
3.1 Method of Moments	12
3.2 Maximum Likelihood Estimation	13
3.3 Applications	15
4 Distributions Related to the Gamma Distributions	19
4.1 Log Gamma Distribution	19
4.2 The Quotient of Two Independent Gamma	20
4.3 The Quotient $\frac{X_1}{X_1+X_2}$	20

4.4 Sum of k Independent Gamma . . . 21

4.4.1 Aprroximation for the Exact Distribution of the sum of k Independent Gamma . . . 23

Approximation with Two Moments . . . 23

Normal Approximation . . . 23

5 Generalized Gamma Distributions . . . 37

5.1 History . . . 37

5.2 Motivation and Getting the Generalized Gamma Density . . . 37

5.3 Basic Properties for the Generalized Gamma Distribution . . . 38

5.4 Particular Cases and Shapes . . . 39

6 Estimation of parameters . . . 44

6.1 Method of moments . . . 44

6.2 Maximum likelihood estimation . . . 46

7 Product and Ratio of two Generalized Gamma . . . 48

8 Hierarchical Models . . . 51

9 Random Numbers Generation for a Generalized Gamma Distribution . . . 55

9.1 Estimation of the sum of n Generalized Gamma Distributions . . . 57

References . . . 62

10 Appendix . . . 64

10.1 Code to Approximate the sum of k Independent Gamma Random Variables 64

10.2 Code to Approximate the Sum of k Independent Generalized Gamma Random Variables . . . 75

10.3 Bedfordshire Rainfall Statistics Table . . . 80

Curriculum Vitae . . . 83

List of Tables

3.1 Table of Observed and Expected Frequency 18

4.1 Approximation with two moments for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = 2, \alpha_2 = 1.5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 9,\ Var(Y) = 24.$ 27

4.2 Approximation with two moments for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = .7, \alpha_2 = .5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 3.1,\ Var(Y) = 8.3$ 28

4.3 Approximation with two moments for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 3, \alpha_4 = 2.5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 29.55,\ Var(Y) = 93.32$ 29

4.4 Approximation with two moments for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = .7, \alpha_2 = .3, \alpha_3 = .9, \alpha_4 = .5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 11.4,\ Var(Y) = 38.1$ 30

4.5 Approximation with two moments for the sum of the exact distribution of 10 gamma variables. $\alpha_1 = 2, \alpha_2 = 2.5, \alpha_3 = 3, \alpha_4 = 3.5, \alpha_5 = 4, \alpha_6 = 4.5, \alpha_7 = 5, \alpha_8 = 5.5, \alpha_9 = 6, \alpha_10 = 6.5\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4, \beta_6 = 4.5, \beta_7 = 5\beta_8 = 5.5, \beta_9 = 6, \beta_10 = 6.5,\ E(Y) = 201.25,\ Var(Y) = 1030.625.$ Both graphs are overlapping one to each other due to the good approximation 31

4.6 Normal approximation for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = 2, \alpha_2 = 1.5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 9,\ Var(Y) = 24.$ 32

4.7 Normal approximation for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = .7, \alpha_2 = .5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 3.1,\ Var(Y) = 8.3$ 33

4.8 Normal approximation for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 3, \alpha_4 = 2.5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 29.55,\ Var(Y) = 93.32$ 34

4.9 Normal approximation for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = .7, \alpha_2 = .3, \alpha_3 = .9, \alpha_4 = .5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4$, $E(Y) = 11.4$, $Var(Y) = 38.1$ Both graphs are overlapping one to each other due to the good approximation 35

4.10 Normal approximation for the sum of the exact distribution of 10 gamma variables. $\alpha_1 = 2, \alpha_2 = 2.5, \alpha_3 = 3, \alpha_4 = 3.5, \alpha_5 = 4, \alpha_6 = 4.5, \alpha_7 = 5, \alpha_8 = 5.5, \alpha_9 = 6, \alpha_10 = 6.5\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4, \beta_6 = 4.5, \beta_7 = 5\beta_8 = 5.5, \beta_9 = 6, \beta_{10} = 6.5$, $E(Y) = 201.25$, $Var(Y) = 1030.625$ Both graphs are overlapping one to each other due to the good approximation 36

9.1 Empirical distribtuion for a generalized gamma random variable. $\alpha = 2, \beta = 4$, $\delta = 2$, $E(X) = 5.31$, $Var(X) = 3.725$ 56

9.2 Normal approximation for the sum of the exact distribution of 2 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \beta_1 = 4, \beta_2 = 3, \delta_1 = 2, \delta_2 = 3$, $E(Y) = 10.33$, $Var(Y) = 4.3$ 58

9.3 Normal approximation for the sum of the exact distribution of 2 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \beta_1 = 4, \beta_2 = 3, \delta_1 = .7, \delta_2 = .8$, $E(Y) = 35.40$, $Var(Y) = 326.92$ 59

9.4 Normal approximation for the sum of the exact distribution of 5 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \alpha_3 = 10, \alpha_4 = 3, \alpha_5 = 78, \beta_1 = 4, \beta_2 = 3, \beta_3 = 81, \beta_4 = 37, \beta_5 = 125, \delta_1 = 2, \delta_2 = 3, \delta_3 = 4, \delta_4 = 5, \delta_5 = 6$, $E(Y) = 456.012$, $Var(Y) = 188.68$ 60

9.5 Normal approximation for the sum of the exact distribution of 10 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \alpha_3 = 10, \alpha_4 = 3, \alpha_5 = 7, \alpha_6 = .20, \alpha_7 = .9, \alpha_8 = 2.5, \alpha_9 = .3, \alpha_10 = .05\beta_1 = 4, \beta_2 = 3, \beta_3 = 5, \beta_4 = 6, \beta_5 = 2, \beta_6 = 500, \beta_7 = 18\beta_8 = 25, \beta_9 = 20, \beta_10 = 100.5, \delta_1 = 2, \delta_2 = 3, \delta_3 = 4, \delta_4 = 5, \delta_5 = 6, \delta_7 = 8, \delta_8 = 9, \delta_9 = 10, \delta_10 = 11$, $E(Y) = 405.1474$, $Var(Y) = 19932.65$. . 61

10.1 Rainfall Statistics in mm 80

10.2 Rainfall Statistics in mm (continuation) 81
10.3 Rainfall Statistics in mm (continuation) 82

List of Figures

2.1 Example of a standard gamma 8

2.2 Example of an exponential distribution 9

2.3 Example of a chi-square distribution, $\nu = 5$ 10

3.1 Relative Frequency Histogram of Bedfordshire Rainfall 16

3.2 $\widehat{f}(x) = \frac{1}{\Gamma(2.619)\cdot(20.99)^{(2.619)}} \cdot x^{2.617-1} e^{-x/20.9016}$ 17

5.1 Weibull, $h(x, 2, 2)$ 40

5.2 Half Normal, $g(x, 1)$ 41

5.3 Analyzing when δ increases 42

5.4 Analyzing when δ decreases 43

9.1 Empirical distribution for a Generalized Gamma 57

Chapter 1

Introduction

The purpose of this thesis is to review forms following the generalized gamma distribution. Such forms include the exponential, the standard gamma, the weibull and other distributions. These distributions are used in many fields, in particular life studies and they are very common in application of statistics. The generalized gamma distribution is given in a form that has been studied in the literature. We examine the estimation of parameters by the method of maximum likelihood and method of moments. Chapters from two to four concern to the gamma distribution. In chapter two the gamma distribution is briefly studied. How to obtain the density starting from the gamma function and obtaining its mean, variance and moment generating function. Particular cases and shapes are shown, those of the standard gamma, exponential and chi square distribution. For the chi square distribution we give the distribution for a linear combination of random variables each with chi-square distribution (Moschopoulos 1985). In chapter 3 we estimate the parameters for the gamma distribution by using the method of moments and the maximum likelihood estimation. We applied this estimation to show how data from real world can be used to fit a gamma distribution. In chapter four some distributions related to the gamma distribution are presented. We define the log gamma distribution as a random variable X that satisfies that $\log(X)$ follows a gamma distribution. The quotient for two independent gamma distributions is analyzed. The distribution for the sum of k independent gamma when the scale parameters are different is shown (Moschopoulos 1985). We also state two ways to approximate the exact distribution of this sum. One of them is with two moments and the other one is by means of a third moment normal approximation motivated with the ideas of (Jensen and Solomon 1972).

Chapters five to nine talk about the generalized gamma distribution. In chapter five the generalized gamma distribution is introduced. We obtain the density for the generalized gamma in the same way as we did for the gamma distribution, by means of a simple change of variable. Some basic properties for the generalized gamma distribution are stated. We obtain the gamma, weibull and half normal distribution as particular cases of the generalized gamma distribution. The shape of the generalized gamma when the parameter δ increases or decreases is obtained and analyzed and this leads to see that δ is a shape parameter. In chapter six we obtain the maximum likelihood estimators for the generalized gamma distribution. Chapter seven talks about the product and ratio of generalized gamma random variables. We start developing the density for the ratio of two random variables having a generalized gamma density. Then we give the distribution for the product of this kind of ratio (Coelho and Joao 2007). Chapter eight gives a new approach for generalized gamma distributions, those of the hierarchical models. We obtain a hierarchical model when $Y|\Lambda$ follows a Poisson distribution and Λ is a generalized gamma random variable. In chapter nine we obtain a way to generate random numbers for a generalized gamma distribution. We take a look at the empirical distribution of these kind of random variables. And finally we use the same normal approximation as in chapter 4 to estimate the distribution of the sum of k independent generalized gamma random variables.

Chapter 2

The Gamma Distribution and its Basic Properties

2.1 History and Motivation

The gamma distribution is known since the Laplace age is mentioned. The gamma distribution appears naturally in theory associated with normally distributed random variables, as the distribution of the sum of squares of independent standard normal variables. In applied work, gamma distributions give useful representations of many physical situations. They have been used to make realistic adjustments to exponential distributions in lifetimes problems. The model has an application in the theory of random processes in time, in particular in meteorological precipitation processes. Gamma distributions play a very important role when studying the formal theory of mathematical statistics.

2.2 The Gamma Function

The function

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy, \alpha > 0, \tag{2.1}$$

is known as the gamma function. The gamma function satisfies the following properties:

1. $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$
2. $\Gamma(n) = (n-1)!$, where n is a positive integer.
3. $\Gamma(1/2) = \sqrt{\pi}$.

To prove property 1 we have that

$$\Gamma(\alpha+1) = \int_0^\infty y^{(\alpha+1)-1} e^y dy = \int_0^\infty y^\alpha e^y dy = \alpha \int_0^\infty y^\alpha e^{-y} - y^\alpha e^{-y}\Big|_0^\infty,$$

using integration by parts. As the term $y^{\alpha-1}e^{-y}\big|_0^\infty \to 0$ as $y \to \infty$ we have that

$$\Gamma(\alpha+1) = \alpha \int_0^\infty y^{\alpha-1} e^{-y} = \alpha\Gamma(\alpha).$$

Property 2 follows immediately setting $\alpha = n$, where n is a positive integer and applying property 1 n times:

$$\begin{aligned}
\Gamma(n) &= (n-1) \times \Gamma(n-1) \\
&= (n-1) \times (n-2) \times \Gamma(n-2) \\
&= (n-1) \times (n-2) \times (n-3) \times \Gamma(n-3) \\
&\vdots \\
&= (n-1) \times (n-2) \times (n-3) \ldots 3 \times 2 \times 1 = (n-1)!
\end{aligned}$$

To prove property 3 we will use the fact that $\int_{-\infty}^\infty e^{-x^2} dx = \pi$. Now noting that e^{-x^2} is symmetric respect to zero we have

$$\begin{aligned}
\Gamma(1/2) &= \int_0^\infty e^{-x} x^{-1/2} dx. \\
&= \int_0^\infty e^{-u^2} du, \text{ setting } x = u^2, \text{ we obtain } du = x^{-1/2} dx \qquad (2.2) \\
&= \frac{\pi}{2}.
\end{aligned}$$

There are interesting alternative ways to define the gamma function (Continuous Univariate Distributions [1]). It can be defined in terms of a limit by

$$\Gamma(\alpha) = \lim_{n\to\infty} \frac{n^\alpha}{\alpha(1+\alpha)(1+\frac{\alpha}{2})\ldots(1+\frac{\alpha}{n})}. \qquad (2.3)$$

Note that this definition does not involve any integral at all. This definition is equivalent to

$$\frac{1}{\Gamma(\alpha)} = \alpha e^{\gamma\alpha} \cdot \prod_{i=1}^{n} [(1+\frac{\alpha}{n})e^{-\alpha/n}], \ \alpha > 0, n \text{ is an integer }, \qquad (2.4)$$

where γ is the Euler-Mascheroni constant given by

$$\gamma = \lim_{m\to\infty}[\sum_{n=1}^{m}\frac{1}{n} - \log m] \approx 0.5772156649\ldots. \tag{2.5}$$

The gamma function is studied in text of advanced calculus and arises often in applications.

2.3 Getting the Gamma Density

A function of the form

$$f(x) = cx^{\alpha-1}e^{-x/\beta},\ \ \alpha,\beta > 0, \tag{2.6}$$

defines a very special probability density. To find c we set

$$\int_0^\infty cx^{\alpha-1}e^{-x/\beta}dx = 1$$

Letting $y = x/\beta$ the integral above becomes

$$\int_0^\infty cx^{\alpha-1}e^{-x/\beta}dx = c\beta^\alpha\int_0^\infty y^{\alpha-1}e^{-y}dy = c\beta^\alpha\Gamma(\alpha) = 1,$$

from where we get

$$c = \frac{1}{\beta^\alpha\Gamma(\alpha)}.$$

Now with this value for c the gamma density in (2.6) becomes

$$f(x,\alpha,\beta) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta},\ \ \alpha,\beta > 0. \tag{2.7}$$

2.4 Mean, Variance and Moment Generating Function

We will use the notation $X \sim G(\alpha,\beta)$ to denote that a random variable X with density given by (2.7), follows a gamma distribution with parameters α,β.

The rth moment for the gamma distribution is given by

$$\mu'_r = E(X^r) = \frac{\beta^r\Gamma(\alpha+r)}{\Gamma(\alpha)},\ \ \alpha,\beta > 0, r \text{ a positive integer.} \tag{2.8}$$

The derivation for the rth moment comes from the definition

$$\begin{aligned}
E(X^r) &= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^{\alpha-1} \cdot x^r e^{-x/\beta} dx \\
&= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^{\alpha+r-1} e^{-x/\beta} dx \\
&= \frac{1}{\beta^\alpha \Gamma(\alpha)} \Gamma(\alpha+r)\beta^{\alpha+r} = \frac{\beta^r \Gamma(\alpha+r)}{\Gamma(\alpha)}.
\end{aligned}$$

Thus we have that the mean is equal to

$$\mu_1' = E(X) = \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\beta\alpha\gamma(\alpha)}{\gamma(\alpha)} = \alpha\beta. \tag{2.9}$$

The variance is given by

$$Var(X) = \mu_2' - (\mu_1')^2 = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} - \big(\frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)}\big)^2 = \alpha\beta^2. \tag{2.10}$$

The moment generating function for the gamma is given by

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} \cdot \frac{1}{\beta^\alpha\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx, \ -\infty < t < \infty. \tag{2.11}$$

In order to get this moment generating function we will try to express the right side above as a gamma density. We do the change of variable $y = x(1-\beta t)$ then:

$$\begin{aligned}
E(e^{tX}) &= \int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)} x^{\alpha-1} e^{-x(1-\beta t)/\beta} dx \\
&= \int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)} \frac{y^{\alpha-1}}{(1-\beta t)^{\alpha-1}} \cdot \frac{1}{1-\beta t} e^{-y} dy \\
&= \frac{1}{(1-\beta t)^{\alpha-1}} \cdot \frac{1}{1-\beta t} \cdot \frac{1}{\beta^\alpha\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{1}{(1-\beta t)^\alpha}.
\end{aligned}$$

One application for the moment generating function of a gamma distribution is that of the reprodcutive property. If $X_1 \sim G(\alpha_1, \beta)$ and $X_2 \sim G(\alpha_2, \beta)$ are independent with respective densities

$$f_1(x, \alpha_1, \beta) = \frac{1}{\beta^{\alpha_1}\Gamma(\alpha_1)} x^{\alpha_1-1} e^{-x/\beta}, \ \alpha_1, \beta > 0 \tag{2.12}$$

and

$$f_2(x, \alpha_2, \beta) = \frac{1}{\beta^{\alpha_2}\Gamma(\alpha_2)} x^{\alpha_2-1} e^{-x/\beta}, \ \alpha_1, \beta > 0, \tag{2.13}$$

then $X_1 + X_2 \sim G(\alpha_1 + \alpha_2, \beta)$. That is, the density for the sum of the random variables $X_1 + X_2$ is

$$f_{X_1+X_2}(x, \alpha_1 + \alpha_2, \beta) = \frac{1}{\beta^{\alpha_1+\alpha_2}\Gamma(\alpha_1 + \alpha_2)} x^{\alpha_1+\alpha_2-1} e^{-x/\beta}. \tag{2.14}$$

This easily seen using the properties for the moment generating function. Note that $M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$ because X_1 and X_2 are independent. So we have

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t) = (1 - \beta t)^{-\alpha_1} \cdot (1 - \beta t)^{-\alpha_2} = (1 - \beta t)^{-(\alpha_1+\alpha_2)}.$$

Here we use the fact that there exists a unique correspondence between moment generating function and distribution. This means that as $M_{X_1+X_2}(t)$ has the form of a moment generating function for a gamma with parameters β, $\alpha_1 + \alpha_2$ then $X_1 + X_2$ has a gamma distribution with parameters β and $\alpha_1 + \alpha_2$. We can even generalize this to k independent gamma's. Let $X_i, \dots X_k$ be k independent random variables each one following a gamma distribution with parameters (α_i, β), $i = 1, 2, \dots k$. Then the moment generating function for these k independent gamma is

$$M_{X_1 \cdot \ldots \cdot X_k}(t) = M_{X_1}(t) \cdot \ldots M_{X_k}(t) = (1 - \beta t)^{-\alpha_1} \cdot \ldots \cdot (1 - \beta t)^{-\alpha_k} = (1 - \beta t)^{-\sum_{i=1}^{k} \alpha_i}.$$

This is the moment generating function for a gamma with parameters $\sum_{i=1}^{k} \alpha_i$ and β. So by the same reasoning as in the two parameters case we conclude that the random variable $X_1 + \dots X_n$ has a gamma distribution with parameters $(\sum_{i=1}^{k} \alpha_i, \beta)$.

2.5 Particular Cases and Shapes

2.5.1 Standard Gamma

Several forms of the gamma density can be obtained by changing the parameters in (2.7). We will analyze these forms and study them for some values. One of them is obtained by setting $\beta = 1$ in (2.7). That is

$$f(x, \alpha, 1) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-x}, x > 0, \ \alpha > 0, \tag{2.15}$$

this form for the gamma is called the standard gamma. The integral

$$\Gamma_x(\alpha) = \int_0^x t^{\alpha-1} e^t dt$$

is called the incomplete gamma function.

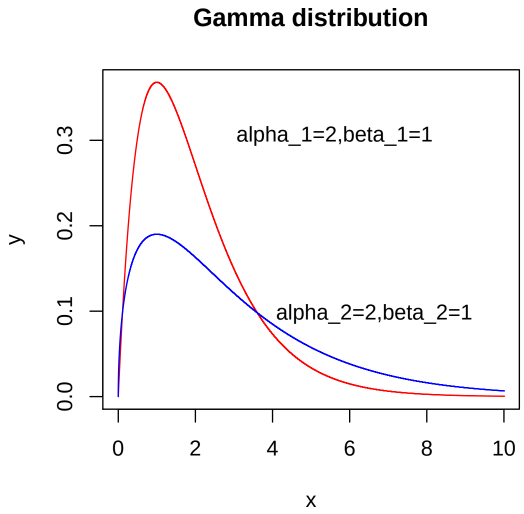


Figure 2.1: Example of a standard gamma

2.5.2 The Exponential

Another very important form of the gamma distribution is obtained by setting $\alpha = 1$ in (2.7). This is the exponential distribution.

$$f(x, 1, \beta) = e^{-x/\beta}, \; x > 0, \; \beta > 0. \tag{2.16}$$

This distribution arises often in practice. It can be used to model the waiting time between arrivals of a Poisson process (see [8] pp 203). According to (2.9) and (2.10) we can see that the the mean and variance for the exponential are equal to β and β^2 respectively.

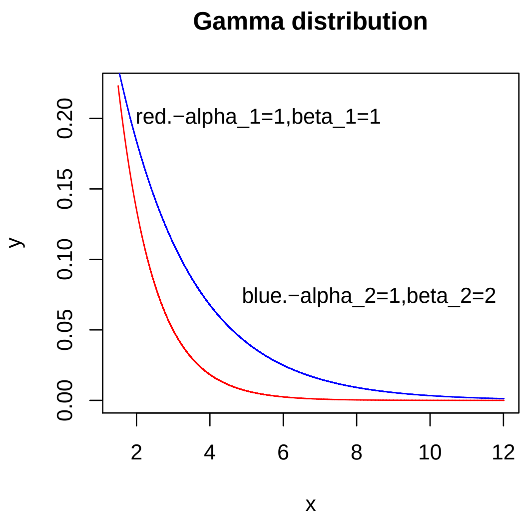


Figure 2.2: Example of an exponential distribution

2.5.3 The Chi square distribution

A chi-square distribution is perhaps the most important special case of a gamma distribution. This distribution is defined with $\alpha = \nu/2, \nu > 0$ and $\beta = 2$.

$$f(x, \nu/2, 2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} \;, x > 0, \nu > 0. \tag{2.17}$$

We then say that the density in (2.17) is a Chi-square with ν degrees of freedom. If a random variable X follows a chi-square distribution with ν degrees of freedom we will denote it as $X \sim \chi^2(\nu)$. This distribution plays a very important role in many applications. It also serves as a way to characterize the square of standard normal distributions. For example if $U_1, U_2 \ldots U_n$ are standard normal random variables $(N(0,1))$ then the distribution of $\sum_{i=1}^{n} U_i^2$ has a chi-square distribution with n degrees of freedom.

The mean and variance of the chi square $\chi^2(\nu)$ distribution are ν and 2ν respectively.

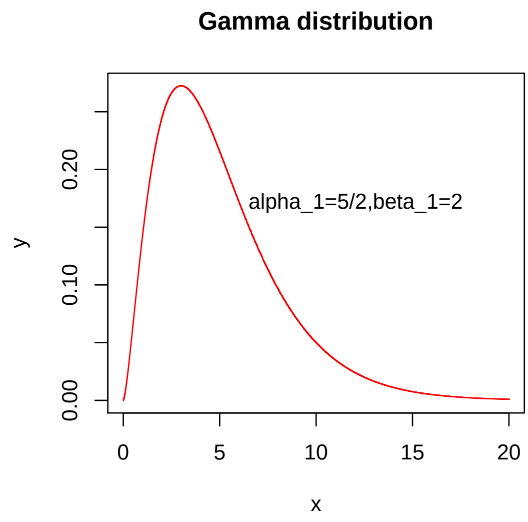


Figure 2.3: Example of a chi-square distribution, $\nu = 5$

2.5.4 Linear Combinations of Independent Chi-Square

We include here the distribution of a linear combination of independent chi-square ([10] Moschopoulos 1985), $X_i \sim \chi^2(n_i)$ defined by

$$Y = \sum_{i=1}^{p} c_i \chi^2(n_i), \;\; c_i > 0, \;\; n_i \text{ a positive integer for } i = 1, 2, 3 \ldots p \tag{2.18}$$

where $\chi^2(n_i)$, $i = 1 \ldots p$ is independent with n_i degrees of freedom. The moment generating function for the equation in (2.18) is obtained as

$$M(t) = \prod_{i=1}^{p} (1 - 2c_i t)^{-m_i}, m_i = n_i/2. \tag{2.19}$$

For $t < \min(1/2c_i)$ then

$$(1 - 2c_i t)^{-mi} = (c_1/c_i)^{mi} \sum_{r=0}^{\infty} \frac{(m_i)_r}{r!} (1 - c_1/c_i)^r (1 - 2c_i t)^{-(r+m_r)}. \tag{2.20}$$

So that with this information and multipliying the expression in (2.19) we get

$$M(t) = \big(\sum_{i=2}^{p} b_i \sum_{j=0}^{\infty} a_j (1 - 2c_1 t)^{-(s+j)} \big), s = \sum_{i=1}^{p} m_i, \tag{2.21}$$

with

$$b_i = (c_1/c_i)^{m_i} \text{ and } A(i,r) = (m_i)_r(1 - c_1/c_i)^r/r!,$$

besides the $a_j s$ satisfy the relation

$$\prod_{i=2}^{p}[\sum_{r=0}^{\infty} A(c_i, r)x^{-r}] = \sum_{j=0}^{\infty} a_j x^{-j}.$$

From this we can get the $a_j s$ recursively that is

$$a_j = A_j^{(p)}, A_j^{(i)} = \sum_{k=0}^{j} A_k^{i-1} A(c_i, j-k) \tag{2.22}$$

where $i = 3, 4, \ldots, p$ $j = 0, 1, 2\ldots$ and for $r = 0, 1, 2, \ldots$

$$A_r^{(2)} = A(C_2, r).$$

And now we invert (2.21) term by term. Since the density corresponding to the factor $(1 - 2c_1 t)^{-(s+j)}$ is the gamma density $g_j(y)$ where

$$g_j(y) = y^{s+j-1}e^{-y/2c_1}/(2c_1)^{s+j}\Gamma(s+j)$$

the density function $F(w) = P(Y \leq w)$ (the probability that the random variable Y is less or equal than w) is

$$F(w) = (\prod_{i=2}^{p} b_i) \sum_{j=0}^{\infty} a_j \int_0^{\infty} g_j(y)dy. \tag{2.23}$$

([10] Moschopoulos 1985).

Chapter 3

Estimation of Parameters

3.1 Method of Moments

We know from (2.8) that the r-th moment for $X \sim G(\alpha, \beta)$ is given by

$$\mu'_r = \frac{\beta^r \Gamma(\alpha + r)}{\Gamma(\alpha)},$$

these are the population moments for $r = 1, 2, 3 \ldots$.

If we consider random variables $X_1, X_2, \ldots X_n \sim G(\alpha, \beta)$, and if we observe a particular value from each, $x_1, x_2, \ldots x_n$, then the sample moments are given by $\frac{\sum_{i=1}^{n} x_i^r}{n}$. The method of moments consists on setting those population moments equal to the sample moments. Considering the cases $r = 1, 2$, we must solve the equations

$$\frac{\beta \Gamma(\alpha + 1)}{\Gamma(\alpha)} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3.1}$$

and

$$\frac{\beta^2 \Gamma(\alpha + 2)}{\Gamma(\alpha)} = \frac{\sum_{i=1}^{n} x_i^2}{n} \tag{3.2}$$

for β and α.

By using the properties for the gamma function (3.1) and (3.2) become

$$\alpha\beta = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{3.3}$$

$$\beta^2(\alpha + 1)\alpha = \frac{\sum_{i=1}^{n} x_i^2}{n}. \tag{3.4}$$

Solving (3.3) for β we obtain $\beta = \frac{\overline{x}}{\alpha}$, where $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$. Now if we substitute in (3.4) for

α we have

$$\begin{aligned}
&\beta^2(\alpha+1)\alpha = \frac{\sum_{i=1}^n x_i^2}{n}\\
&\Leftrightarrow \frac{\overline{x}^2}{\alpha^2}(\alpha+1)\alpha = \frac{\sum_{i=1}^n x_i^2}{n} (\text{ substituing } \beta = \frac{\overline{x}}{\alpha})\\
&\Leftrightarrow \frac{\overline{x}^2(\alpha+1)}{\alpha} = \frac{\sum_{i=1}^n x_i^2}{n}\\
&\Leftrightarrow \frac{(\alpha+1)}{\alpha} = \frac{\sum_{i=1}^n x_i^2}{\overline{x}^2 n}\\
&\Leftrightarrow 1+\frac{1}{\alpha} = \frac{\sum_{i=1}^n x_i^2}{\overline{x}^2 n}\\
&\Leftrightarrow \frac{1}{\alpha} = \frac{\sum x_i^2}{\overline{x}n} - 1 = \frac{\sum x_i^2 - \overline{x}^2 n}{\overline{x}^2 n},
\end{aligned}$$

from where finally we get

$$\widehat{\alpha} = \frac{\overline{x}^2 n}{\sum x_i^2 - \overline{x}^2 n}$$

and

$$\widehat{\beta} = \frac{\overline{x}}{\widehat{\alpha}}$$

3.2 Maximum Likelihood Estimation

Let $X_1, X_2, \ldots X_n \sim G(\alpha,\beta)$ for $i = 1, 2 \ldots n$. The likelihood function is given by

$$L(x_1,\ldots x_n,\alpha,\beta) = \prod_{i=1}^n f(x_i,\alpha,\beta) = \frac{1}{\Gamma(\alpha)^n \cdot \beta^{n\alpha}}(x_1,\ldots x_n)^{\alpha-1}e^{-\sum_{i=1}^n (x_i/\beta)}. \tag{3.5}$$

The maximum likelihood estimation method consists on finding the values for α and β such that (3.5) is a minimum.

Taking natural log at both sides of (3.5) and using the properties for logarithms we have

$$\ln(L(x_1\ldots x_n,\alpha,\beta)) = -n\ln(\Gamma(\alpha)) - n\alpha\ln(\beta) + (\alpha-1)\ln(x_1,\ldots x_n) - \Sigma_{i=1}^n (x_i/\beta). \tag{3.6}$$

The derivative respect to α in (3.6) set it equal to zero give the corresponding equation

$$\frac{\partial L}{\partial \alpha} = -\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} - n\ln(\beta) + \sum_{i=1}^n \ln(x_i) = 0.$$

This can be expressed as

$$\begin{aligned}\frac{\partial L}{\partial \alpha} &= -\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\Big(\ln(x_i) - \ln(\beta)\Big) = 0 \\ &= -\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\ln(x_i/\beta) = 0.\end{aligned} \tag{3.7}$$

Differentiating with respect to β

$$\frac{\partial L}{\partial \beta} = -\frac{-n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2}. \tag{3.8}$$

Setting (3.8) equal to zero the equation is

$$\begin{aligned}&-\frac{-n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2} = 0 \\ &\Leftrightarrow -n\alpha + \frac{\sum_{i=1}^{n} x_i}{\beta} = 0 \\ &\Leftrightarrow \frac{\sum_{i=1}^{n} x_i}{\beta} = n\alpha.\end{aligned}$$

From where we obtain

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} x_i}{n\alpha} = \frac{\overline{x}}{\alpha}. \tag{3.9}$$

Where we have used the notation $\widehat{\beta}$ to denote the maximum likelihood estimator for β.

Try to solve for α the equation (3.7) is quite complicated because of the function $\Gamma(\alpha)$. There is no closed way to solve for α in this equation. By adding the value for β in terms of α obtained in (3.9) to (3.7) we get

$$\begin{aligned}&-\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\ln(x_i/\beta) = -\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\ln(x_i\alpha/\overline{x}) \\ &= -n\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\big(\ln(x_i) + \ln(\alpha) - \ln(\overline{x})\big) \\ &= -n\psi(\alpha) + \sum_{i=1}^{n}\ln(x_i) + n\ln(\alpha) - n\ln(\overline{x}) \\ &= -\psi(\alpha) + \frac{\sum_{i=1}^{n}\ln(x_i)}{n} + \ln(\alpha) - \ln(\overline{x}) = 0 \\ &\Leftrightarrow \ln(\alpha) - \psi(\alpha) = \ln(\overline{x}) - \frac{\sum_{i=1}^{n}\ln(x_i)}{n},\end{aligned} \tag{3.10}$$

where $\psi(\alpha)$ is the digamma function defined as $\psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$. One approximation for the left side of (3.10) is

$$\ln(\alpha) - \psi(\alpha) \approx \frac{1}{\alpha}\left(\frac{1}{2} + \frac{1}{12\alpha + 2}\right).$$

So (3.10) becomes

$$\frac{1}{\alpha}\left(\frac{1}{2} + \frac{1}{12\alpha + 2}\right) \approx \ln(\overline{x}) - \frac{\sum_{i=1}^{n} \ln(x_i)}{n}.$$

Finally solving this equation for α we obtain the approximate maximum likelihood estimation for α

$$\widehat{\alpha} \approx \frac{3 - s + \sqrt{(s-3)^2 + 24s}}{12s} \tag{3.11}$$

where $s = \ln(\overline{x}) - \frac{1}{n}\sum_{i=1}^{n} \ln(x_i)$ ([5] S. C. Choi and R. Wette. (1969)).

3.3 Applications

Now we will use the previous estimators to fit a gamma density to a set of data.

Data set 1 The Bedfordshire county give the next information about the daily rainfall since january 1984 up to december 2010 by month (see Table 10.1). The histogram for this data is given in figure (3.1).

The estimation of the parameters by means of the maximum likelihood is given by (3.11)

$$\widehat{\alpha} \approx \frac{3 - s + \sqrt{(s-3)^2 + 24s}}{12s}$$

where s is given by $\ln(\overline{x}) - \frac{1}{n}\sum_{i=1}^{n} \ln(x_i)$ and

$$\widehat{\beta} = \frac{\overline{x}}{\alpha}.$$

The estimation of the parameters lead to $\widehat{\alpha} \approx 2.617$ and $\widehat{\beta} = 20.9016$. So that the estimated density function is given by

$$\widehat{f}(x) = \frac{1}{(20.9016)^{2.617}} \cdot \Gamma(2.617) \cdot x^{2.617} e^{-x/20.9016}$$

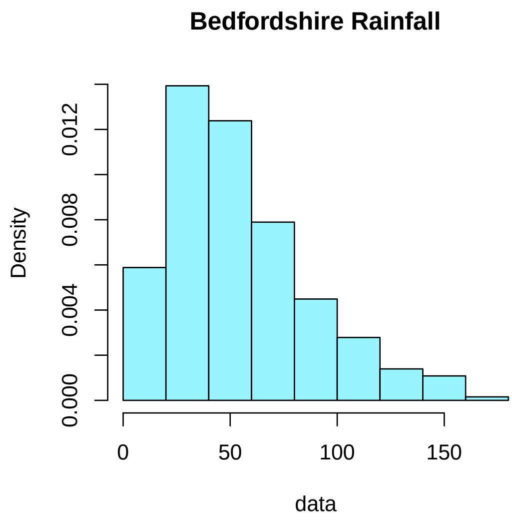


Figure 3.1: Relative Frequency Histogram of Bedfordshire Rainfall

At first glance it seems to be a very good fit to this data in figure (3.2), where we have the estimated density with the estimated parameters as above. We can be sure of this by doing a goodness of fit test. With the hypotheses

$$H_0 : \text{Data follows a gamma distribution.}$$

$$H_a : \text{Data doesn't follow a gamma distribution.}$$

The used statistic is

$$\chi^2 = \sum_{i=1}^{9} \frac{(f_i - e_i)^2}{e_i},$$

The data in table (10.1) was arranged into nine class intervals. The table shows 324 numbers representing rainfall in milimeters. We omit the value for august in 2003 because it is zero. Here f_i and e_i are the observed and expected frequency for each class respectively, for $i = 1, 2, \ldots, 9$. Under the hypothesis that H_0 is true the expected frequency is defined as $e_{i+1} = 323 \times \left(\widehat{f}(y_{i+1}) - \widehat{f}(y_i)\right)$, $i = 0, 2, \ldots 8$ and $y_0 = 0, y_1 = 20, y_2 = 40, \ldots y_9 = 180$. This information is summarized in table (3.3).

We reject H_0 if $\chi^2 \geq \chi^2_{m-t-1}$, otherwise we do not reject H_0. Here m is the number of

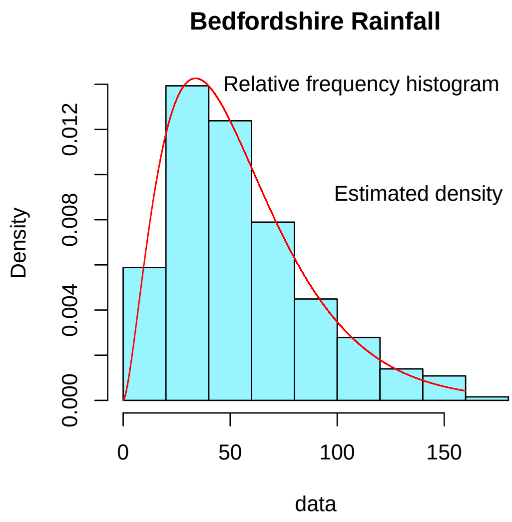


Figure 3.2: $\widehat{f}(x) = \frac{1}{\Gamma(2.619)\cdot(20.99)^{(2.619)}} \cdot x^{2.617-1}e^{-x/20.9016}$

terms in the summation and t is the number of parameters estimated. In this case $m = 9$ and $t = 2$. Hence we have that

$$\chi^2 = 3.0421 < \chi^2_6 = 12.5916,$$

so that we do not reject H_0 and conclude that this data does follow a gamma distribution.

Table 3.1: Table of Observed and Expected Frequency

Class Interval	Observed Frequency	Expected Frequency
0-20	38	38.76
20-40	90	88.43
40-60	80	79.40
60-80	51	53.18
80-100	29	30.88
100-120	18	16.49
120-140	9	8.32
140-160	7	4.04
160-180	1	1.903

Chapter 4

Distributions Related to the Gamma Distributions

4.1 Log Gamma Distribution

We say that a random variable X is log gamma distributed if $\ln(X)$ is gamma distributed. This means that if $Y = \ln(X)$ then the density for Y denoted by $f(y, \alpha, \beta)$ is given by

$$f_Y(y, \alpha, \beta) = \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} y^{\alpha-1} e^{-y/\beta}, \beta, \alpha, y > 0. \tag{4.1}$$

We will get the density for X. Noting that Y is a transformation for X the density for X is given by $f(x, \alpha, \beta) = f_Y(\ln(x)) \cdot \left|\frac{dy}{dx}\right|$. As we know that $\ln(x) = y$ and $\left|\frac{dy}{dx}\right| = \frac{1}{x}$ we obtain

$$\begin{aligned} f(x, \alpha, \beta) &= \frac{(\ln(x))^{\alpha-1}}{x\beta^\alpha\Gamma(\alpha)} \cdot e^{-\ln(x)/\beta} \\ &= \frac{(\ln(x))^{\alpha-1}}{x\beta^\alpha\Gamma(\alpha)} \cdot x^{-1/\beta} = \frac{(\ln(x))^{\alpha-1}}{\beta^\alpha\Gamma(\alpha)} \cdot x^{-(\beta+1)/\beta}, x > 1. \end{aligned}$$

Some properties ([6] P.C Consul and G. C. Jain) for the log gamma distribution are:

1. If Y is a log gamma with density $f(x, \alpha, \beta)$ then Y^n is also a log gamma distribution with density $f(x, \alpha/n, \beta)$.

2. If $Y_1, Y_2, \ldots Y_n$ are independent log gamma variables with densities $f(x_i, \alpha, \beta_i)$ for $i = 1, 2 \ldots,$ then $U = \Pi_{i=1}^n Y_i$ is also a log gamma density with density $f(u, \alpha, \sum_{i=1}^n \beta_i)$.

3. If $Y_1, Y_2, \ldots Y_n$ is a sequence of independent log gamma variables, where Y_i has density $f(x_i, \alpha, \beta_i)$ for all $i = 1, 2 \ldots,$ then the product $U = \Pi_{i=1}^n Y_i^n$ is also a log gamma variable with density $f(u, \alpha/n, \sum_{i=1}^n \beta_i)$.

4.2 The Quotient of Two Independent Gamma

Now if we consider the ratio X/Y, where $X \sim G(\alpha_1, \beta_1)$ and $Y \sim G(\alpha_2, \beta_2)$ and they are independent. Setting

$$u = X + Y, v = \frac{X}{Y},$$

then the jacobian becomes $|J| = \frac{-(1-v)^2}{u}$. So

$$\begin{aligned} f_{U,V}(u,v) =& \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u} \big(\frac{uv}{1+v}\big)^{\alpha_1 - 1} \big(\frac{u}{1+v}\big)^{\alpha_2 - 1} \frac{u}{1+v}^{2} \\ =& \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u} u^{\alpha_1+\alpha_2-1} v^{\alpha_2 - 1} (1+v)^{-(\alpha_1+\alpha_2)}. \end{aligned}$$

If we integrate this with respect to u we obtain

$$\begin{aligned} f_V(v) =& \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^\infty e^{-u} u^{\alpha_1+\alpha_2-1} v^{\alpha_2-1} (1+v)^{-(\alpha_1+\alpha_2)} du \\ =& \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} v^{\alpha_2-1} (1+v)^{-(\alpha_1+\alpha_2)} \int_0^\infty e^{-u} u^{\alpha_1+\alpha_2-1} du \\ =& \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} v^{\alpha_2-1} (1+v)^{-(\alpha_1+\alpha_2)} \Gamma(\alpha_1+\alpha_2), \end{aligned}$$

because $\displaystyle\int_0^\infty e^{-u} u^{\alpha_1+\alpha_2-1} du = \Gamma(\alpha_1+\alpha_2)$.

4.3 The Quotient $\dfrac{X_1}{X_1 + X_2}$

If X_1 and X_2 are independent random gamma variables with parameters (α_1, β) and (α_2, β) respectiveley, then $X_1/(X_1 + X_2)$ has a beta distribution with parameters α_1, α_2. We can deduce this from

$$\begin{aligned} f(x_1) &= \frac{1}{\beta_1^{\alpha_1} \cdot \Gamma(\alpha_1)} x_1^{\alpha_1 - 1} e^{-x/\beta}, \\ f(x_2) &= \frac{1}{\beta_2^{\alpha_2} \cdot \Gamma(\alpha_2)} x_1^{\alpha_2 - 1} e^{-x/\beta}, \end{aligned}$$

then as the random variables are independent we have

$$f(x_1, x_2) = \frac{1}{\Gamma(\alpha_1) \cdot \Gamma(\alpha_2)} \frac{1}{\beta^{\alpha_1}} \cdot \frac{1}{\beta^{\alpha_2}} e^{\frac{-(x_1+x_2)}{\beta}} x_1^{\alpha_1-1} x_2^{\alpha_2-1}. \tag{4.2}$$

Now we set

$$u = X_1 + X_2,\ v = X_1/(X_1 + X_2),$$

then

$$X_1 = uv,\ X_2 = u(1 - v).$$

With this the jacobian becomes $|J| = u$. So

$$\begin{aligned} f_{U.V}(u, v) =& \frac{u}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u/\beta} \cdot (uv)^{\alpha_1 - 1} u^{\alpha_2 - 1} (1 - v)^{\alpha_2 - 1} \\ =& \frac{u}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u/\beta} u^{\alpha_1 + \alpha_2 - 1} v^{\alpha_1 - 1} (1 - v)^{\alpha_2 - 1}. \end{aligned} \tag{4.3}$$

Therefore, integrating (4.3) respect to v the sum has distribution

$$f(u) = \frac{e^{-u/\beta} u^{\alpha_1 + \alpha_2 - 1}}{\Gamma(\alpha_1 + \alpha_2)},$$

which is a gamma distribution with parameters $\alpha_1 + \alpha_2$ and β. Now using (4.3) and integrating respect to u the ratio $v = X_1/(X_1 + X_2)$ has distribution

$$f_V(v) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1) \cdot \Gamma(\alpha_2)} v^{\alpha_1 - 1} \cdot (1 - v)^{\alpha_2 - 1},$$

which is a beta function with parameters α_1, α_2.

4.4 Sum of k Independent Gamma

We can say even more about the random variable $Y = \sum_{i=1}^{k} X_i$ where each variable has gamma distribution, that is $X_i \sim (\alpha_i, \beta_i)$, $i = 1, 2 \ldots n$ and they are independent ([9], Moschopoulos 1985) . We can obtain the exact distribution for Y. We know that each X_i has moment generating function given by

$$M(t) = (1 - \beta_i t)^{-\alpha_i},$$

then Y has moment generating function given by

$$M(t) = \prod_{i=1}^{k} (1 - \beta_i t)^{-\alpha_i}.$$

The idea is to express $M(t)$ as

$$M(t) = C(1 - \beta_i t)^{-\rho} \cdot \exp(\sum_{k=1}^{\infty} \gamma_k (1 - \beta_i t)^{-k}),$$

where $\exp(a) = e^a$ and

$$C = \prod (\beta_1/\beta_i)^{\alpha_i},$$

with $\beta_1 = \min(\beta_1, \beta_2 \dots \beta_n)$ (the minimum among all β's),

$$\gamma_k = \sum_{i=1}^{n} \alpha_i (1 - \beta_1/\beta_i)^k / k,$$

$$\rho = \sum_{i=1}^{n} \alpha_i > 0.$$

And now if we let

$$\exp(\sum_{k=1}^{\infty} \gamma_k (1 - \beta_i t)^{-k}) = \sum_{k=0}^{\infty} \delta_k (1 - \beta_1 t)^{-k}.$$

By differentiating with respect to $(1 - \beta t)^{-1}$, the coefficients δ_k can be obtained by means of the formula

$$\delta_{k+1} = \frac{1}{k+1} \sum_{i=1}^{k+1} i \gamma_i \delta_{k+1-i}.$$

This leads to the next theorem which gives an expression for Y ([9] Moschopoulos 1985).

Theorem 1. *If $X_i \sim (\alpha_1, \beta_i)$ and independently distributed for $i = 1, 2, \dots n$. Then the density for Y can be expressed as*

$$g(y) = C \sum_{k=0}^{\infty} \delta_k y^{\rho+k-1} e^{-y/\beta_1} / [\Gamma(\rho + k) \beta_1^{\rho+k}], y > 0 \tag{4.4}$$

and 0 elsewhere.

4.4.1 Aprroximation for the Exact Distribution of the sum of k Independent Gamma

Approximation with Two Moments

If $X_i \sim G(\alpha_i, \beta_i)$ for $i = 1 \ldots n$, and $Y = X_1 + \ldots + X_n$. Then we know that the density is given by (4.4)

$$g(y) = C \sum_{k=0}^{\infty} \delta_k y^{\rho+k-1} e^{-y/\beta_1} / [\Gamma(\rho + k) \beta_1^{\rho+k}], y > 0. \tag{4.5}$$

If we want to compute $G(y)$, the distribution for this sum, we need to evaluate this inifinite sum. But we can approximate it assuming that $Y \sim G(\alpha, \beta)$ and knowing two moments of this assumed distribution. If $Y \sim G(\alpha, \beta)$ then $EY = \alpha\beta$ and $Var(Y) = \alpha\beta^2$, from where we obtain

$$\beta = \frac{Var(Y)}{E(Y)},$$

and

$$\alpha = \frac{E(Y)^2}{Var(Y)},$$

but $E(Y) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots \alpha_n\beta_n$ and $Var(Y) = \alpha_1\beta_1^2 + \ldots + \alpha_n\beta_n^2$. So that with these values we obtain assuming that $Y \sim G(\alpha, \beta)$

$$\beta = \frac{\alpha_1\beta_1^2 + \ldots + \alpha_n\beta_n^2}{\alpha_1\beta_1 + \alpha_2\beta_2 + \ldots \alpha_n\beta_n}$$

and

$$\alpha = \frac{(\alpha_1\beta_1 + \alpha_2\beta_2 + \ldots \alpha_n\beta_n)^2}{(\alpha_1\beta_1^2 + \ldots + \alpha_n\beta_n^2)}.$$

If we use this values for α and β we can approximate (4.5) with $G(\alpha, \beta)$, where α and β are given as above.

Normal Approximation

If X_i, $i = 1 \ldots n$ are independent standard normal variables and considering the following linear combination ([18] Solomon and Jensen 1972)

$$Q_k(\mathbf{c}, \mathbf{a}) = \sum_{j=1}^{k} c_j(x_j^2 + a_j), c_j > 0, 1 \geq j \leq k, \tag{4.6}$$

then the sth cumulant of $Q_k(\mathbf{c}, \mathbf{a})$ is

$$\kappa_s = 2^{s-1}(s-1)! \sum_{j=1}^{k} c_j^s (1 + s a_j^2).$$

Under the condition that the parameters c_j, a_j are bounded, it can be shown that the distribution of Q_k tends to a gaussian distribution as $\theta_1 \to \infty$ where $\theta_1 = E(Q_k(\mathbf{c}, \mathbf{a}))$. So following this idea (Solomon and Jensen 1972) used the transformation $(Q_k/\theta_1)^h$ to accelerate the rate of convergence and thus to provide a Gaussian approximation for moderate values of θ_1. The moments of $(Q_k/\theta_1)^h$ expanded in powers of $(\theta_1)^{-1}$ were obtained. In particular the third central moment is

$$\begin{aligned}\mu_3(h) = & \frac{4h^2}{\theta_1^2}\big(2\phi_3 + 3(h-1)\phi_2^2\big) \\ & \frac{h^2(h-1)}{\theta_1^3}\big(72\phi_4 + 24(7h-10)\phi_2\phi_3\big) \\ & + 4(17h^2 - 55h + 44)\phi_2^2 + 0(\theta_1^{-4}),\end{aligned} \tag{4.7}$$

where $\phi_r = \theta_r/\theta_1, r = 2, 3$. A similar expression for γ_1, the skewness of $(Q_k/\theta_1)^h$ is

$$\begin{aligned}\gamma_1(h) =& \frac{4\big(2\phi_3 + 3(h-1)\phi_2^2\big)}{\theta_1^{1/2}(2\phi_2^{3/2})} \\ & \cdot (h-1)\big(72\phi_4 + 24(7h-10)\phi_2\phi_3 + \frac{4(17h^2 - 55h + 44)\phi_2^3\big)}{\theta_1^{5/2}} \\ & + 0(\theta_1^{-5/2}).\end{aligned} \tag{4.8}$$

Now h is chosen so that the leading term in γ_1 and μ_3 vanishes. This is

$$\frac{2\theta_3}{\theta_1} + 3(h-1)\big(\frac{\theta_2}{\theta_1}\big)^2 = 0. \tag{4.9}$$

This yields to

$$h = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}. \tag{4.10}$$

Now the distribution of the random variable $(Q_k/\theta_1)^h$ can be approximated by a Gaussian distribution with mean

$$\mu_1'(h) = 1 + \theta_2 h(h-1)/\theta_1^2, \tag{4.11}$$

and variance

$$Var_{(Q_k/\theta_1)^h}(h) = 2\theta_2 h^2/\theta_1. \tag{4.12}$$

In summary the variable $z = \theta_1\big((Q_k/\theta_1)^h - 1 - \theta_2 h(h-1)/\theta_1^2\big)/(2\theta_2 h^2)$ is approximately a standard normal variable.

Now, motivated with this ideas, we will consider another approximation for the distribution of $Y = X_1+X_2+\ldots X_n$, with $X_i \sim G(\alpha_i, \beta_i), i = 1, 2 \ldots n$. We propose that $(Y/\theta_1)^h$ with h given as in (4.10) and $E(Y) = \theta_1$, can be approximated with a standard normal distribution with mean μ_h and variance σ_h^2 given as follows. We introduce the cumulant κ_n of the random variable Y defined as

$$\kappa_n = g^{(n)}(t) = \frac{d^n}{dt}g(t),$$

where

$$g(t) = \ln\big(E(e^{tY})\big) = -\sum_{n=1}^{\infty}\frac{1}{n}\big(1 - E(e^{tY})\big)^n$$
$$= \sum_{n=1}^{\infty}\frac{1}{n}\big(-\sum_{m=1}^{\infty}\mu'_m\frac{t^m}{m!}\big)^n$$
$$= \mu'_1 t + \big(\mu'_2 - \mu'_1\big)\frac{t^2}{2!} + \big(\mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3\big)\frac{t^3}{3!} + \ldots.$$

We have used the fact that $\ln(x) = -\sum_{n=1}^{\infty}\frac{(1-x)^n}{n}$

Hence we have that

$$\kappa_1 = g'(0) = \mu',$$

$$\kappa_2 = g''(0) = \mu'_2 + (\mu'_1)^2,$$

$$\kappa_3 = g^{(3)}(0) = \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3.$$

In general $\kappa_n = g^{(n)}(0)$. Now if we define

$$h = 1 - \kappa_1\kappa_3/(3\cdot\kappa_1^2),$$

$$\mu_h = 1 + \frac{h(h-1)\cdot\phi_2}{2\cdot\kappa_1} + \frac{h\cdot(h-1)\cdot(h-2)\cdot(4\cdot\phi_3 + 3\cdot(h-3)\cdot\phi_2^2)}{24\cdot\kappa_1^2}$$

and

$$\sigma_h^2 = \frac{\phi_2 h^2}{\kappa_1} + \frac{(h-1)\cdot h^2\big(2\phi_3 + (3h-5)\phi_2^2\big)}{2\kappa_1^2},$$

with

$$\phi_2 = \frac{\kappa_2}{\kappa_1},$$
$$\phi_3 = \frac{\kappa_3}{\kappa_1},$$

then we take

$$Z = \big(\frac{(Y/\theta_1)^h - \mu_h}{\sigma_h}\big) \tag{4.13}$$

to be approximately a standardized Gaussian variable. So that we can use this information to approximate the distribution of Y. This kind of approximation has been used in literature with other applications. Moschopoulos in 1983 ([11]) and Moschopoulos and Mudholkar in 1983 ([12]) used and actually improved the efficiency of these approximations. In the next pages tables with the normal and two moments approximation are shown. We consider here the distribution $P(Y \leq t)$ where $Y = X_1 + \ldots + X_n$.

Table 4.1: Approximation with two moments for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = 2, \alpha_2 = 1.5, \beta_1 = 3, \beta_2 = 2$, $E(Y) = 9$, $Var(Y) = 24$.

t	Exact	Approximate	Difference
5	.2122	.2137	-.0014
6	.3045	.3050	-.0005
7	.3989	.3985	.0003
8	.4899	.4888	.0011
9	.5740	.5724	.0015
10	.6490	.6472	.0017
11	.7143	.7125	.0017
12	.7698	.7682	.0015

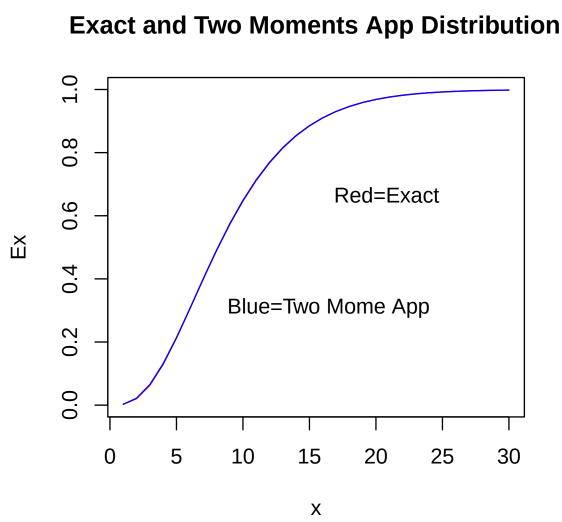


Table 4.2: Approximation with two moments for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = .7, \alpha_2 = .5, \beta_1 = 3, \beta_2 = 2$, $E(Y) = 3.1$, $Var(Y) = 8.3$

t	Exact	Approximate	Difference
1	.2404	.2442	-.0038
3	.6110	.6099	.0011
5	.8070	.8051	.0019
7	.9051	.9040	.0010
9	.9533	.9531	.0002
11	.9770	.9772	-.0001
13	.9887	.9889	-.0002
15	.9944	.9946	-.0002

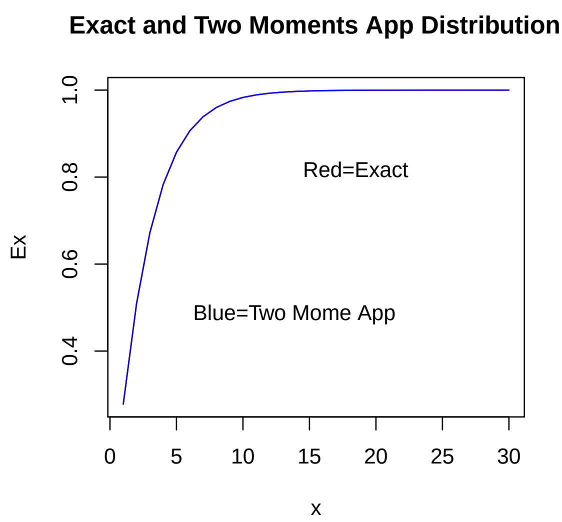


Table 4.3: Approximation with two moments for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 3, \alpha_4 = 2.5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 29.55,\ Var(Y) = 93.32$

t	Exact	Approximate	Difference
20	.1558	.1565	-.0007
23	.2661	.2661	-.0000
26	.3931	.3924	.0006
29	.5219	.5207	.0011
32	.6401	.6388	.0012
35	.7400	.7390	.0010
38	.8191	.8185	.0006
41	.8783	.8781	.0002
44	.9206	.9206	.0000

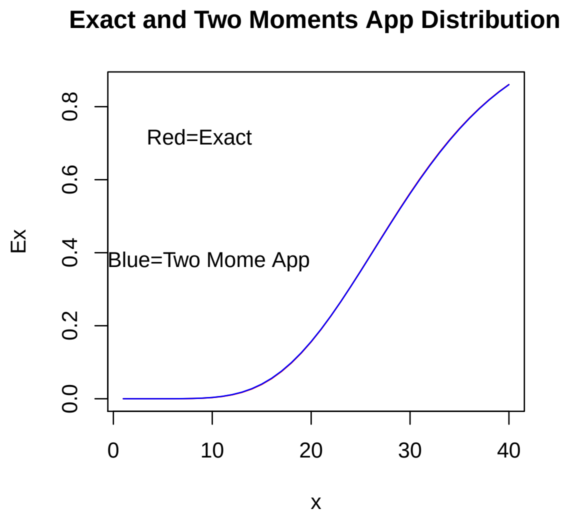


Table 4.4: Approximation with two moments for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = .7, \alpha_2 = .3, \alpha_3 = .9, \alpha_4 = .5, \alpha_5 = 1.2\beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 11.4,\ Var(Y) = 38.1$

t	Exact	Approximate	Difference
12	.6112	.6088	.0023
14	.7184	.7161	.0023
16	.8013	.7994	.0018
18	.8627	.8615	.0011
20	.9067	.9062	.0005
22	.9375	.9374 -	.0000
24	.9586	.9589 -	.0002
26	.9728	.9733 -	.0004
28	.9886	.9828 -	.0004

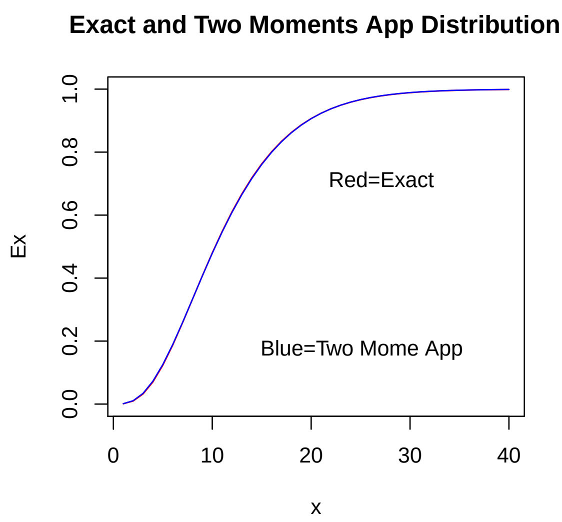


Table 4.5: Approximation with two moments for the sum of the exact distribution of 10 gamma variables. $\alpha_1 = 2, \alpha_2 = 2.5, \alpha_3 = 3, \alpha_4 = 3.5, \alpha_5 = 4, \alpha_6 = 4.5, \alpha_7 = 5, \alpha_8 = 5.5, \alpha_9 = 6, \alpha_10 = 6.5 \beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4, \beta_6 = 4.5, \beta_7 = 5 \beta_8 = 5.5, \beta_9 = 6, \beta_10 = 6.5$, $E(Y) = 201.25$, $Var(Y) = 1030.625$. Both graphs are overlapping one to each other due to the good approximation

t	Exact	Approximate	Difference
180	.2636	.2633	.003
200	.5068	.5056	.0011
220	.7330	.7322	.0007
240	.8828	.8829	-.0000
260	.9576	.9581	-.0004
280	.9868	.9875 -	.0006
300	.9958	.9968 -	.0009
320	.9980	.9993 -	.0012
340	.9984	.9998 -	.0013

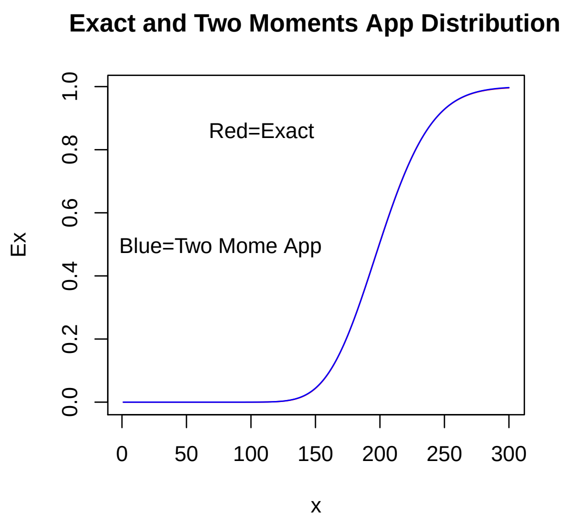


Table 4.6: Normal approximation for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = 2, \alpha_2 = 1.5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 9,\ Var(Y) = 24$.

t	Exact	Approximate	Difference
5	.2122	.2108	.0014
6	.3045	.3030	.0015
7	.3989	.3977	.0011
8	.4899	.4892	.0006
9	.5740	.5738	.0001
10	.6490	.6493	-.0003
11	.7143	.7149	-.0006
12	.7698	.7706	-.0008

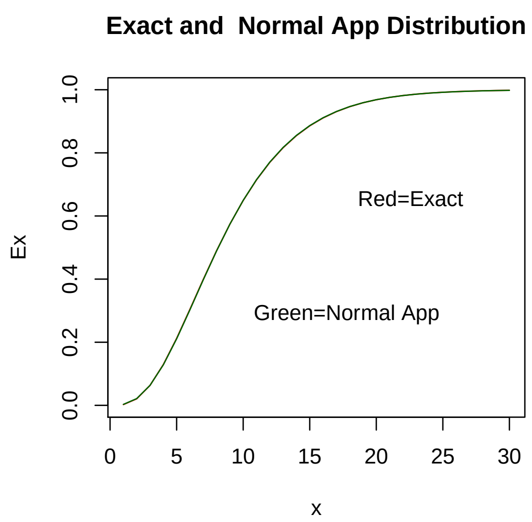


Table 4.7: Normal approximation for the sum of the exact distribution of 2 gamma variables. $\alpha_1 = .7, \alpha_2 = .5, \beta_1 = 3, \beta_2 = 2,\ E(Y) = 3.1,\ Var(Y) = 8.3$

t	Exact	Approximate	Difference
1	.2404	.2358	.0045
3	.6110	.6100	.0009
5	.8070	.8091	-.0021
7	.9051	.9069	-.0018
9	.9533	.9543	-.0009
11	.9770	.9773	-.0003
13	.9887	.9886	.0000
15	.9944	.9942	.0001

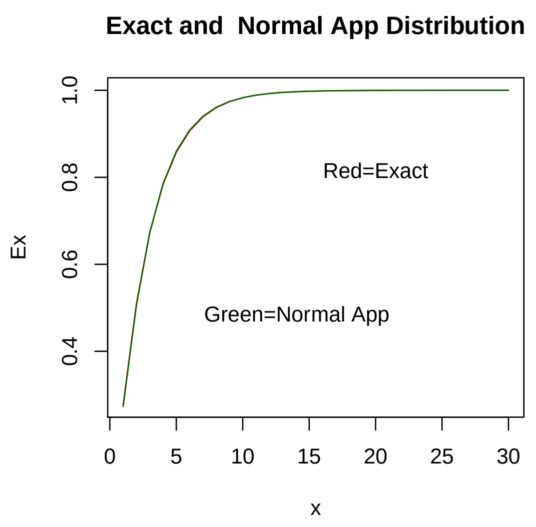


Table 4.8: Normal approximation for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 3, \alpha_4 = 2.5, \alpha_5 = 1.2 \beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4,\ E(Y) = 29.55,\ Var(Y) = 93.32$

t	Exact	Approximate	Difference
20	.1558	.1554	-.0003
23	.2661	.2656	-.0004
26	.3931	.3924	.0003
29	.5219	.5218	.0000
32	.6401	.6403	-.0001
35	.7400	.7404	-.0003
38	.8191	.8195	-.0003
41	.8783	.8786	-.0002
44	.9206	.9207	-.0001

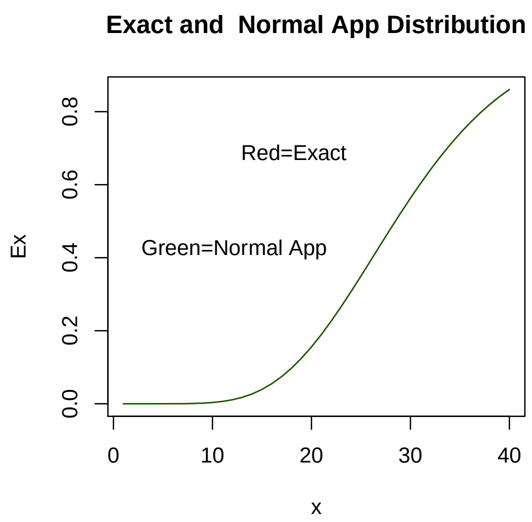


Table 4.9: Normal approximation for the sum of the exact distribution of 5 gamma variables. $\alpha_1 = .7, \alpha_2 = .3, \alpha_3 = .9, \alpha_4 = .5, \alpha_5 = 1.2 \beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4$, $E(Y) = 11.4$, $Var(Y) = 38.1$ Both graphs are overlapping one to each other due to the good approximation

t	Exact	Approximate	Difference
12	.6112	.6114	-.0001
14	.7184	.7191	-.0006
16	.8013	.8021	-.0008
18	.8627	.8634	-.0007
20	.9067	.9073	-.0005
22	.9375	.9378	-.0003
24	.9586	.9588 -	.0001
26	.9728	.9729 -	.0000
28	.9886	.9823	.0000

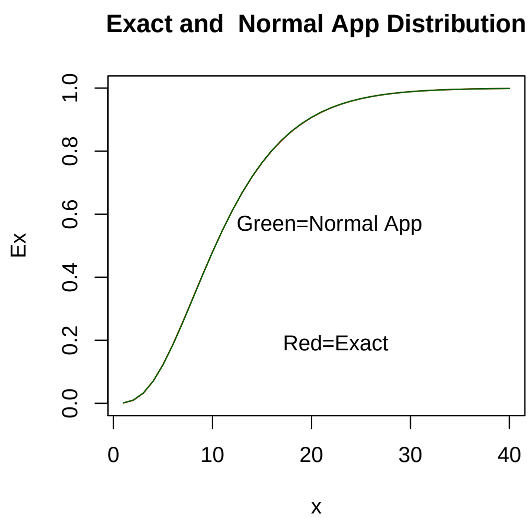
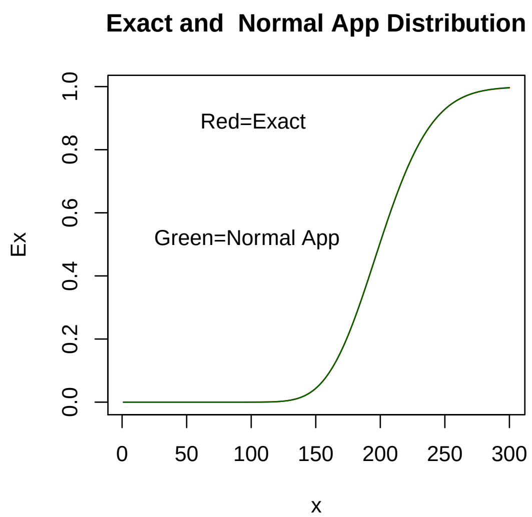


Table 4.10: Normal approximation for the sum of the exact distribution of 10 gamma variables. $\alpha_1 = 2, \alpha_2 = 2.5, \alpha_3 = 3, \alpha_4 = 3.5, \alpha_5 = 4, \alpha_6 = 4.5, \alpha_7 = 5, \alpha_8 = 5.5, \alpha_9 = 6, \alpha_10 = 6.5 \beta_1 = 2, \beta_2 = 2.5, \beta_3 = 3, \beta_4 = 3.5, \beta_5 = 4, \beta_6 = 4.5, \beta_7 = 5 \beta_8 = 5.5, \beta_9 = 6, \beta_{10} = 6.5$, $E(Y) = 201.25$, $Var(Y) = 1030.625$ Both graphs are overlapping one to each other due to the good approximation

t	Exact	Approximate	Difference
180	.2636	.2635	.0000
200	.5068	.5068	.0000
220	.7330	.7331	-.0000
240	.8828	.8829	-.0000
260	.9576	.9577	-.0001
280	.9868	.9871 -	.0003
300	.9958	.9966 -	.0007
320	.9980	.9992 -	.0011
340	.9984	.9998 -	.0013



Chapter 5

Generalized Gamma Distributions

5.1 History

Generalized gamma distributions were discussed since the beginning of the 20th century. It appeared when fitting such distribution to an observed distribution of income rates. There was not a good interest in this kind of distribution up to the middle of the 20th century. From that time on, the interest in the generalized gamma was increasing. Properties, applications, distributions which were related to a generalized gamma were showing up. Stacy (see [15] 1963) published a paper which developed the idea of a generalized gamma introducing basic properties. It was the first paper talking entirely about this generalized distribution.

5.2 Motivation and Getting the Generalized Gamma Density

So far we have seen several properties for the gamma distribution. We will focus our attention on investigating the properties of the generalized gamma functions. If X is a random variable we say X has generalized gamma density if its pdf is given by

$$f(x) = cx^{\delta\alpha-1}e^{-(\frac{x}{\beta})^{\delta}}, x, \alpha, \beta, \delta > 0. \tag{5.1}$$

This distribution contains all previous densities we have just seen, such as gamma density and its densities derived from it. This is one form of the generalized gamma

that appears often in literature (see [1]), but there exist others forms that depend on the parameters. This is form is convenient for our purposes.

The constant c in (5.1) is the reciprocal of the integral

$$\int_0^\infty x^{\delta\alpha-1} e^{-(\frac{x}{\beta})^\delta} dx.$$

If we let $y = (x/\beta)^\delta$ we get $x^{\delta\alpha-1} = y^{\alpha-\frac{1}{\delta}} \cdot \beta^{\delta\alpha-1}$. And the differential dx is given by

$$dx = (\beta/\delta) \cdot y^{\frac{1-\delta}{\delta}}.$$

And then we get

$$\int_0^\infty cx^{\delta\alpha-1} e^{-(\frac{x}{\beta})^\delta} dx = \beta^{\delta\alpha} \cdot \frac{1}{\delta} \int_0^\infty cy^{\alpha-\frac{1}{\delta}+\frac{1}{\delta}-1} e^{-y} dy = \beta^{\delta\alpha} \cdot \frac{1}{\delta} \int_0^\infty cy^{\alpha-1} e^{-y} dy = 1.$$

But as

$$\int_0^\infty y^{\alpha-1} e^{-y} dy$$

has the form of a standard gamma we obtain

$$\int_0^\infty y^{\alpha-1} e^{-y} dy = \Gamma(\alpha).$$

Finally have

$$\beta^{\delta\alpha} \cdot \frac{1}{\delta} \int_0^\infty cy^{\alpha-1} e^{-y} dy = \beta^{\delta\alpha} \cdot \frac{1}{\delta} c\Gamma(\alpha) = 1.$$

From where we get

$$c = \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)}.$$

And thus the generalized gamma density becomes

$$f(x, \alpha, \beta, \delta) = \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)} \cdot x^{\delta\alpha-1} e^{-(\frac{x}{\beta})^\delta}. \tag{5.2}$$

5.3 Basic Properties for the Generalized Gamma Distribution

We will use the notation $X \sim GG(\alpha, \beta, \delta)$ to denote that a random variable X has a generalized gamma distribution with density given as in (5.2).

If $X \sim GG(\alpha, \beta, \delta), k > 0$ and m is a positive integer then

$$kX \sim GG(y, k\beta, \alpha, \delta), \tag{5.3}$$

$$X^m \sim GG(z, \beta^m, \alpha, \delta/m). \tag{5.4}$$

For example $(X/\beta)^\delta$, this random variable will follow a generalized gamma distribution

$$f(y, 1, \alpha, 1) = y^{\alpha-1}e^{-y}, y > 0,$$

which is the common gamma distribution.

If $F(x)$ is the cumulative distribution function for (5.2) given by

$$F(x) = \int_0^x f(x, \alpha, \beta, \delta)dx,$$

then

$$F(x) = \begin{cases} \Gamma_w(\alpha)/\Gamma(\alpha) & \text{if } \delta > 0 \\ 1 - \Gamma_w(\alpha)/\Gamma(\alpha) & \text{if } x < \delta \end{cases}$$

Where $w = (x/\beta)^\delta$ and Γ_w is the incomplete gamma function.

A necessary and sufficient condition for $|W|^{1/\delta}$ to have a $G(\alpha, \beta)$ distribution is that the pdf is of the form

$$p_W(w) = h(w)|w|^{\delta\alpha-1}e^{(-\frac{|w|^\delta}{\beta})},$$

with

$$h(w) + h(-w) = |\delta| \cdot \big(\beta^\alpha\Gamma(\alpha)\big)^{-1}, \text{for all } w.$$

([14] Roberts 1971).

5.4 Particular Cases and Shapes

If $\delta = 1$ in (5.2) we have the well known gamma distribution

$$f(x, \alpha, \beta) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1} \cdot e^{-x/\beta},$$

given by (2.7).

If $\alpha = 1$ (5.2) turns out the Weibull distribution, given by

$$h(x, \beta, \delta) = \frac{\delta x^{\delta-1}}{\beta^{\delta}} e^{-(\frac{x}{\beta})^{\delta}}. \tag{5.5}$$

We can also get from (5.2) the exponential distribution given by (2.16) just setting $\delta = 1$ in (5.1). This means that an exponential is a special case of a Weibull distribution. Figure (5.1) presents a graph for the Weibull when parameters are $\alpha = 1, \beta = 2, \delta = 2$.

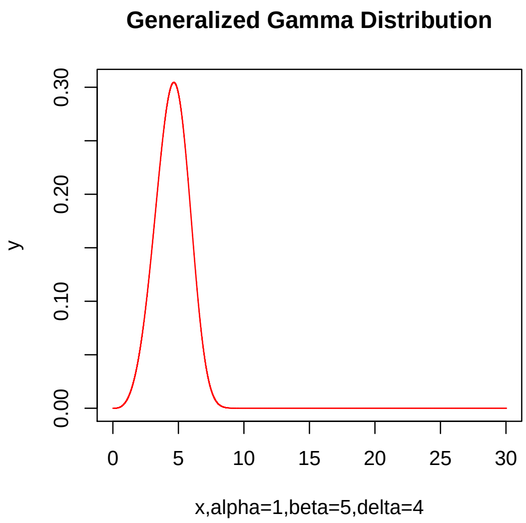


Figure 5.1: Weibull, $h(x, 2, 2)$

By setting $\alpha = 1/2, \delta = 2$ we get the *Half-Normal distribution* defined by

$$g(x, \beta) = \frac{2}{\beta\sqrt{\pi}} \cdot e^{-(\frac{x}{\beta})^2}, x > 0. \tag{5.6}$$

Where we have used the fact that $\Gamma(1/2) = \sqrt{\pi}$. Often this distribution appears in literature with the patameter $\beta = \sqrt{2}\sigma$, where σ is the standard deviation of a normal random variable. So (5.2) becomes

$$g(x, \sigma) = \frac{1}{\sigma} \cdot \sqrt{\frac{2}{\pi}} \cdot e^{-(\frac{x^2}{2\sigma^2})}.$$

Figure (5.2) shows a graph for the density of a half normal distribution

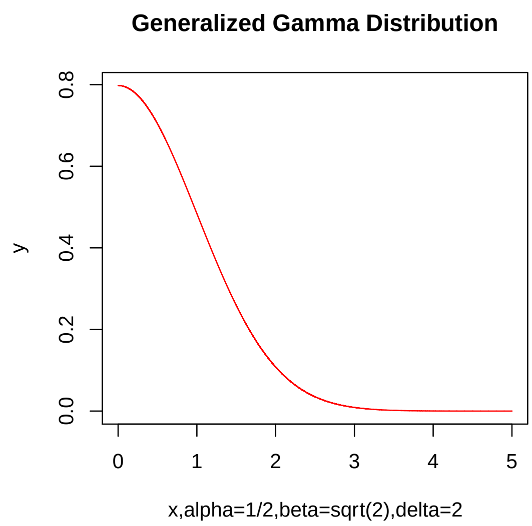


Figure 5.2: Half Normal, $g(x, 1)$

Is interesting to analyze the form of the generalized gamma when δ either increases or decreases. The generalized gamma distribution has two shape parameters α and δ. A plot with several graphs when δ is increasing is shown in figure (5.3). Note how the shape is becoming thinner due to the δ factor. When δ ranges between 0 and 1 the shape for the generalized gamma is skewed to the right. The skewness to the right becomes more notorious as $\delta \to 0$. Figure (5.4) shows several graphs analyzing this situation.

When δ decreases the graph of the generalized gamma is very skewed to the right.

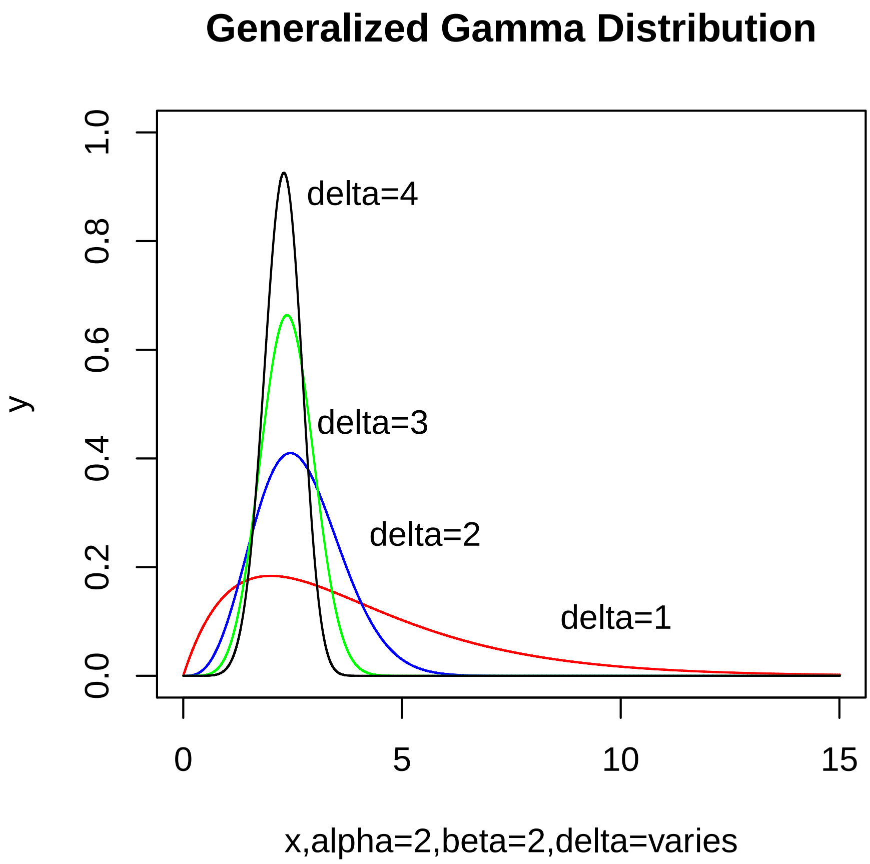


Figure 5.3: Analyzing when δ increases

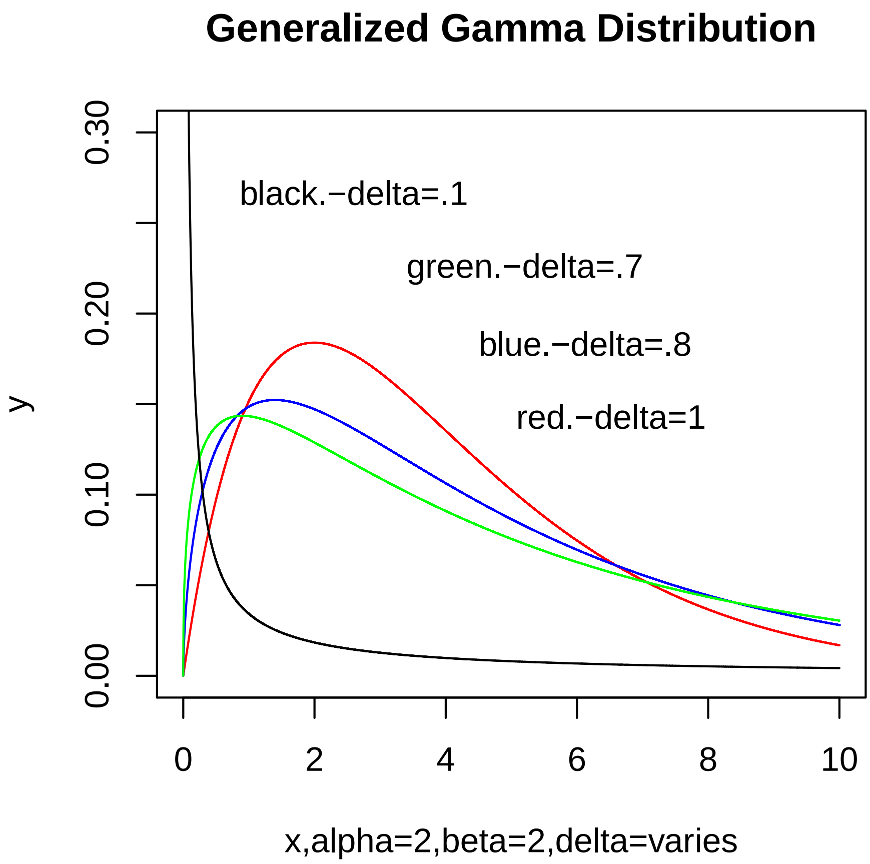


Figure 5.4: Analyzing when δ decreases

Chapter 6

Estimation of parameters

6.1 Method of moments

To find the moments of the generalized gamma we again use a change of variable for the equation

$$\int_0^\infty x^r f(x)dx = \int_0^\infty c x^{\delta\alpha + r - 1} e^{-(x/\beta)^\delta} dx, \tag{6.1}$$

and doing as before the change of variable $y = (x/\beta)^\delta$ we get

$$\frac{\delta}{\Gamma(\alpha)\beta^{\delta\alpha}\delta}\beta^{r+\delta\alpha-1}\beta\int_0^\infty y^{\alpha+\frac{r}{\delta}-1}e^{-y}dy.$$

So that the r^{th} moment is given by

$$\mu_r' = \frac{\beta^r\Gamma(\alpha + \frac{r}{\delta})}{\Gamma(\alpha)}. \tag{6.2}$$

Thus the equations are obtained by equating population and sample moments

$$\overline{x} = \frac{\beta\Gamma(\alpha + \frac{1}{\delta})}{\Gamma(\alpha)} \tag{6.3}$$

$$\sum_{i=1}^n \frac{x_i^2}{n} = \frac{\beta^2\Gamma(\alpha + \frac{2}{\delta})}{\Gamma(\alpha)} \tag{6.4}$$

$$\sum_{i=1}^n \frac{x_i^3}{n} = \frac{\beta^3\Gamma(\alpha + \frac{3}{\delta})}{\Gamma(\alpha)}. \tag{6.5}$$

Stacy ([16] 1965) suggested another method involving the method of moments. If we set

$$Z_i = \ln(X_i/\beta)^\delta = \delta(\ln(X_i) - \ln(\beta)), \tag{6.6}$$

where the X_is, $i = 1, \ldots n$, are independent generalized gamma random variables. And if we denote the central moment of any random variable, say Z by $\mu_k(Z)$, $k = 2, 3 \ldots$. So that (6.6) indicates that

$$\mu_k(Z_i) = \delta^k \mu_k(\ln(X_i)). \tag{6.7}$$

Applying (5.3) and (5.4) to $Z = \ln(X/\beta)^\delta$ we see that

$$Z = \ln(W) = \exp(\alpha z - \exp(z))/\Gamma(\alpha), \tag{6.8}$$

from where we get that the moment generating function of Z is

$$E(e^{\theta Z}) = \Gamma(\alpha + \theta)/\Gamma(\alpha). \tag{6.9}$$

It happens that any kth order derivative of $\Gamma(\alpha + \theta)$, evaluated at $\theta = 0$, is the same whether the derivative is taken with respect to θ or α. Hence the kth moment of Z is

$$E(Z^k) = \Gamma^{(k)}(\alpha)/\Gamma(\alpha), \tag{6.10}$$

where $\Gamma^{(k)}(\alpha)$ is the kth derivative of $\Gamma(\alpha)$ with respect to α. It follows from (6.6) and (6.7) that

$$\begin{gathered} \delta E(\ln(X) - \ln(\beta)) = \mu, \qquad (6.11) \\ \delta^2 \mu_2(\ln(X)) = \mu', \\ \delta^3 \mu_3(\ln(X)) = \mu''. \end{gathered}$$

In terms of the x_i, $y_i = \ln(x_i), i = 1, 2, \ldots n$. Solutions for this equations are

$$a_0 = \exp(\overline{y} - \mu_0/\delta_0) = \exp(\overline{y} - s_y \mu_0/(\mu_0')^{1/2}),$$

$$\delta_0 = (\mu_0')^{1/2}/s_y,$$

$$-|g_y| = \mu''/(\mu')^{3/2},$$

where β_0 is determined by the last equation , $\beta_0 = \mu(\beta_0)$ and $\mu_0' = \mu'(\beta_0)$. The sign of δ is according to whether g_y is less than or greater than zero respectively. Here $\overline{y}, s_y^2$ and g_y are respectively the sample mean, sample variance and sample skewness for the y_i.

6.2 Maximum likelihood estimation

The likelihood function for the generalized gamma is given by

$$L(x_1,\dots x_n,\alpha,\beta,\delta)=\prod_{i=1}^{n} f(x_i,\alpha,\beta,\delta)=\frac{\delta^n}{\Gamma(\alpha)^n\cdot\beta^{n\alpha\delta}}(x_1,\dots x_n)^{\delta\alpha-1}e^{-\sum_{i=1}^{n}(x_i/\beta)^{\delta}}. \quad (6.12)$$

Taking natural log we have

$$L(x_1\dots x_n,\alpha,\beta,\delta)=n\ln(\delta)-n\ln(\Gamma(\alpha))-n\delta\alpha\ln(\beta)+(\alpha\delta-1)\ln(x_1,\dots x_n)-\Sigma_{i=1}^{n}(x_i/\beta)^{\delta}. \quad (6.13)$$

Now taking derivative the corresponding equations are

$$\frac{\partial L}{\partial\alpha}=-\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)}-\delta n\ln(\beta)+\delta\sum_{i=1}^{n}\ln(x_i)=0.$$

This can be expressed as

$$\begin{aligned}\frac{\partial L}{\partial\alpha}&=-\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)}+\delta\ \sum_{i=1}^{n}\Big(\ln(x_i)-\ln(\beta)\Big)=0\\ &=-\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)}+\delta\sum_{i=1}^{n}\ln(x_i/\beta).\end{aligned} \quad (6.14)$$

Differentiating with respect to β

$$\frac{\partial L}{\partial\beta}=-\frac{-\delta n\alpha}{\beta}-\sum_{i=1}^{n}x_i^{\delta}(-\delta)\beta^{-\delta-1}=\frac{\delta}{\beta}[-n\alpha+\sum_{i=1}^{n}\ln(x_i/\beta)^{\delta}]=-n\alpha+\sum_{i=1}^{n}\ln(x_i/\beta)^{\delta}=0.$$

And from this last equation β turns out to be

$$\beta(\alpha,\delta)=\frac{(\sum_{i=1}^{n}(x_i)^{\delta})^{1/\delta}}{(n\alpha)^{1/\delta}} \quad (6.15)$$

it depends on α and from δ.

Differentiating with respect to δ

$$\frac{\partial L}{\partial\delta}=\frac{n}{\delta}-n\alpha\ln(\beta)+\alpha\sum_{i=1}^{n}\ln(x_i)-\sum_{i=1}^{n}\big((x_i/\beta)^{\delta})\cdot\ln(\frac{x_i}{\beta})\big).$$

Now considering $-n\ln(\beta)=-\sum_{i=1}^{n}\ln(\beta)$ and using the properties of the function ln we have

$$\frac{n}{\delta}+\alpha\sum_{i=1}^{n}\ln(\frac{x_i}{\beta})-\sum_{i=1}^{n}\big(\frac{x_i}{\beta}\big)^{\delta}\ln\big(\frac{x_i}{\beta}\big). \quad (6.16)$$

So putting the value of β (6.15) into (6.16) we have

$$\frac{n}{\delta} + \alpha \sum_{i=1}^{n} \ln\Big(\frac{x_i}{\frac{(\sum_{i=1}^{n}(x_i)^{\delta})^{1/\delta}}{(n\alpha)^{1/\delta}}}\Big) - \sum_{i=1}^{n} \Big(\frac{x_i}{\frac{(\sum_{i=1}^{n}(x_i)^{\delta})^{1/\delta}}{(n\alpha)^{1/\delta}}}\Big)^{\delta} \ln\Big(\frac{x_i}{\frac{(\sum_{i=1}^{n}(x_i)^{\delta})^{1/\delta}}{(n\alpha)^{1/\delta}}}\Big).$$

After rearranging terms we get an expression fot α in terms of δ :

$$\alpha(\delta) = \frac{1}{\delta \cdot \big[(\sum_{i=1}^{n} \ln(x_i)/n) - \big(\sum_{i=1}^{n} z_i^{\delta} \ln(x_i)\big)/\big(\sum_{i=1}^{n} x_i^{\delta}\big)\big]}. \tag{6.17}$$

If we substitute (6.17) and (6.15) into (6.14) we will have an equation in terms of δ only. This equation is

$$H(\delta) = -\psi(\alpha) + \delta \frac{\sum_{i=1}^{n} \ln(x_i)}{n} - \ln\big(\sum_{i=1}^{n} x_i^{\delta}\big) + \ln(n\alpha), \tag{6.18}$$

where $\psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ is the digamma function. And α is given by (6.17).

Thus if we are going to try to estimate the parameters for the generalized gamma distribution we need to solve (6.18) for δ. It is not always possible to find a solution for $H(\delta)$. In many cases it is not even possible to know if there exists a solution. Some authors say that the MLE may not exist unless $n > 400$. ([1] Continuous Univariate Distributions)

Chapter 7

Product and Ratio of two Generalized Gamma

The problem of obtaining an explicit expression, without involving any unsolved integrals, for both the probability density functon and the cumulative distribution of the product of two independent random variables with generalized gamma distributions is a challenging one. Expressions for the pdf of such a product had been obtained by some statisticians in terms of complicated functions. But these functions are not readily computable. They are usually computed in terms of the integrals that define them. And is even harder try to extend beyond the product of more of two variables. The computations get really complicated.

If

$$f(x_1) = \frac{\delta_1}{\Gamma(\alpha_1)\beta_1^{\delta_1\alpha_1}} x_1^{\delta_1\alpha_1-1} e^{-(\frac{x_1}{\beta_1})^{\delta_1}},$$

and

$$f(x_2) = \frac{\delta_2}{\Gamma(\alpha_2)\beta_2^{\delta_2\alpha_2}} x_2^{\delta_2\alpha_2-1} e^{-(\frac{x_2}{\beta_2})^{\delta_2}},$$

are two densities of two independent random variables X_1 and X_2 with generalized gamma distribution then the joint density is given by

$$f(x_1, x_2) = \frac{\delta_1}{\Gamma(\alpha_1)\beta_1^{\delta_1\alpha_1}} x_1^{\delta_1\alpha_1-1} e^{-(\frac{x_1}{\beta_1})^{\delta_1}} \cdot \frac{\delta_2}{\Gamma(\alpha_2)\beta_2^{\delta_2\alpha_2}} x_2^{\delta_2\alpha_2-1} e^{-(\frac{x_2}{\beta_2})^{\delta_2}}. \tag{7.1}$$

We want to find the distribution of $Y = X_1/X_2$. If we set $W = X_1$ then we have the variables

$$W = X_1, \text{ and } X_2 = \frac{W}{Y}.$$

The jacobian for these transformations is given by w/y^2. So the joint density of Y and W is

$$f_{Y,W}(w,y) = \frac{\delta_1}{\Gamma(\alpha_1)\beta_1^{\delta_1\alpha_1}} \cdot \frac{\delta_2}{\Gamma(\alpha_2)\beta_2^{\delta_2\alpha_2}} w^{\delta_1\alpha_1-1} e^{-(\frac{w}{\beta_1})^{\delta_1}} (w/y)^{\delta_2\alpha_2-1} e^{-(\frac{(w/y)}{\beta_2})^{\delta_2}} \cdot \frac{w}{y^2} =$$
$$c_1c_2 w^{\delta_1\alpha_1+\delta_2\alpha_2-2+1} \cdot (\frac{1}{y})^{\delta_2\alpha_2-1+2} \cdot e^{-(\frac{w}{\beta_1})^{\delta_1}} e^{-(w/(y\beta_2))^{\delta_2}} = \quad (7.2)$$
$$c_1c_2 w^{\delta_1\alpha_1+\delta_2\alpha_2-1} \cdot (\frac{1}{y})^{\delta_2\alpha_2+1} \cdot e^{-(\frac{w}{\beta_1})^{\delta_1}} e^{-(w/(y\beta_2))^{\delta_2}}.$$

Thus

$$f_Y(y) = \int_0^\infty c_1c_2 w^{\delta_1\alpha_1+\delta_2\alpha_2-1} \cdot (\frac{1}{y})^{\delta_2\alpha_2+1} \cdot e^{-(\frac{w}{\beta_1})^{\delta_1}} e^{-(w/(y\beta_2))^{\delta_2}} dw$$
$$c_1c_2(\frac{1}{y})^{\delta_2\alpha_2+1} \int_0^\infty w^{\delta_1\alpha_1+\delta_2\alpha_2-1} \cdot e^{-(\frac{w}{\beta_1})^{\delta_1}} e^{-(w/(y\beta_2))^{\delta_2}} dw = \quad (\text{assume } \delta = \delta_1 = \delta_2)$$
$$c_1c_2(\frac{1}{y})^{\delta_2\alpha_2+1} \int_0^\infty w^{\delta(\alpha_1+\alpha_2)-1} \cdot e^{-(\frac{w}{\beta_1})^{\delta}} e^{-(w/(y\beta_2))^{\delta}} dw =$$
$$c_1c_2(\frac{1}{y})^{\delta_2\alpha_2+1} \int_0^\infty w^{\delta(\alpha_1+\alpha_2)-1} \cdot e^{-w^\delta\left(\frac{1}{(\beta_1)^\delta}+\frac{1}{(y\beta_2)^\delta}\right)} dw =$$
$$c_1c_2(\frac{1}{y})^{\delta_2\alpha_2+1} \int_0^\infty w^{\delta(\alpha_1+\alpha_2)-1} \cdot e^{-w^\delta/(\beta^{1/\delta})^\delta} dw = \text{ where } \frac{1}{\beta} = \left(\frac{1}{(\beta_1)^\delta} + \frac{1}{(y\beta_2)^\delta}\right)$$
$$c_1c_2\frac{1}{y^{\delta\alpha_2+1}} \cdot \frac{\beta^{\alpha_1+\alpha_2}\Gamma(\alpha_1+\alpha_2)}{\delta} = c_1c_2\frac{1}{y^{\delta\alpha_2+1}} \frac{(\beta_1\beta_2 y)^{\delta(\alpha_1+\alpha_2)}}{((y\beta_2)^\delta + (\beta_1)^\delta)^{\alpha_1+\alpha_2}} \Gamma(\alpha_1+\alpha_2) =$$
$$c_1c_2 \cdot (\beta_1\beta_2)^{\delta(\alpha_1+\alpha_2)} \frac{\Gamma(\alpha_1+\alpha_2)}{\delta} \frac{y^{\delta\alpha_1+\delta\alpha_2-\delta\alpha_2-1}}{((y\beta_2)^\delta + (\beta_1)^\delta)^{\alpha_1+\alpha_2}} =$$
$$c_1c_2 \cdot (\beta_1\beta_2)^{\delta(\alpha_1+\alpha_2)} \frac{\Gamma(\alpha_1+\alpha_2)}{\delta} \frac{y^{\delta\alpha_1-1}}{((y\beta_2)^\delta + (\beta_1)^\delta)^{\alpha_1+\alpha_2}}$$
$$= c_1c_2 \cdot (\frac{\beta_1}{\beta_2})^{\delta(\alpha_1+\alpha_2)} \cdot \frac{\Gamma(\alpha_1+\alpha_2)}{\delta} \frac{y^{\delta\alpha_1-1}}{(y^\delta + (\frac{\beta_1}{\beta_2})^\delta)^{\alpha_1+\alpha_2}}.$$

We will say that a random variable which has this distribution has a GGR (*generalized gamma ratio*) with parameters $(\delta, k, \alpha_1, \alpha_2)$ where $k = \beta_1/\beta_2$.

Coelho and Mexia ([7] Coelho and Joao 2007) gave explicit and concise expressions for the pdf and cumulative distributions of

$$W = \prod_{j=1}^m X_j$$

where the X_j are independent random variables with all distinct shape parameters. If $X_j \sim GGR(k_j, \alpha_{1j}, \alpha_{2j}, \delta_j), j = 1 \dots m$. Where all the parameters are positive. Let

$$s_{1j} = \alpha_{ij} \text{ and } s_{2j} = \alpha_{2j},$$

then if $\delta_j s_{1j} \neq \delta_k s_{1k}$ and $\delta_j s_{2j} \neq \delta_k s_{2k}$, for all $j \neq k,\ j,k = 1,2 \dots n$, the density and cumulative distribution of the random variable

$$W = \prod_{j=1}^{m} X_j, \tag{7.3}$$

are given by

$$f_W(w) = \begin{cases} \lim_{n\to\infty} K_1 K_2 \sum_{j=1}^{m} \sum_{h=0}^{n} H_{2jh} d_{hj} (w/K^*)^{s^*_{1jh}} \frac{1}{w} & \text{if } 0 < w \leq K^* \\ \lim_{n\to\infty} K_1 K_2 \sum_{j=1}^{m} \sum_{h=0}^{n} H_{1jh} c_{hj} (w/K^*)^{-s^*_{2jh}} \frac{1}{w} & \text{if } w \geq K^*, \end{cases}$$

And

$$F_W(w) = \begin{cases} \lim_{n\to\infty} K_1 K_2 \sum_{j=1}^{m} \sum_{h=0}^{n} H_{2jh} \frac{d_{hj}}{s^*_{1jh}} (w/K^*)^{s^*_{1jh}} & \text{if } 0 < w \leq K^* \\ \lim_{n\to\infty} K_1 K_2 \sum_{j=1}^{m} \sum_{h=0}^{n} \left\{ H_{2jh} \frac{d_{hj}}{s^*_{1jh}} + H_{1jh} \frac{c_{hj}}{s^*_{2jh}} \left(1 - (\frac{w}{K^*})^{-s_{2jh}}\right) \right\} & \text{if } w \geq K^*, \end{cases},$$

where

$$K^* = \prod_{j=1}^{m} k_j^{-1/\delta_j},$$

$$K_1 = \prod_{j=1}^{m} \prod_{h=0}^{n} s^*_{1jh},\ K_2 = \prod_{j=1}^{m} \prod_{h=0}^{n} s^*_{2jh},$$

and, for $j = 1. \dots, m$ and $n = 0, 1, \dots,$

$$c_{hj} = \prod_{\eta=1}^{m} \prod_{\nu=}^{n} \frac{1}{s^*_{2\eta\nu} - s^*_{2jh}},\ c_{hj} = \prod_{\eta=1}^{m} \prod_{\nu=}^{n} \frac{1}{s^*_{1\eta\nu} - s^*_{1jh}},$$

and

$$H_{1jh} = \sum_{k=1}^{m} \sum_{l=0}^{n} \frac{d_{kl}}{s^*_{2jh} + s^*_{1kl}},\ H_{2jh} = \sum_{k=1}^{m} \sum_{l=0}^{n} \frac{c_{kl}}{s^*_{1jh} + s^*_{2kl}},$$

with

$$s^*_{1jh} = \delta^*_j (s_{1j} + h) \text{ and } s^*_{2jh} = \delta^*_j (s_{2j} + h),\ (j = 1, \dots m).$$

There are other considerations when the parameters are not the same. (See [7] Coelho and Joao 2007).

Chapter 8

Hierarchical Models

All of the cases seen thus far a random variable had a single distribution depending on parameters. Sometimes is more convenient to think of the things in a hierarchy. A well known hierarchical model is the following. An insect lays a large number of eggs , each surviving with probability p. On the average, how many eggs will survive? The "large number" of eggs laid is a random variable , often chosen to be Poisson (λ). Furthermore we assume that each eggs survival is independent, then we have Bernoulli trials. So if we set X = number of survivors and Y = number of Bernoulli trials we have

$$X|Y \sim \text{binomial}(Y, p)$$

$$Y \sim \text{Poisson}(\lambda).$$

We want to compute

$$\begin{aligned} P(X = x) &= \sum_{y=0}^{\infty} P(X = x, Y = y) \\ &= \sum_{y=0}^{\infty} P(X = x|Y = y)P(Y = y) \\ &= \sum_{y=0}^{\infty} [(yx)p^x(1-p)^{y-x}][\frac{e^{-\lambda}\lambda^y}{y!}]. \end{aligned}$$

Since $X|Y = y$ is binomial(y, p) and Y is Poisson(λ). If we now simplify this last expression

we have

$$
\begin{aligned}
P(X=x) &= \frac{(\lambda p e^{-\lambda})}{x!}\sum_{y=x}^{\infty}\frac{((1-p)\lambda)^{y-x}}{(y-x)!} \\
&= \frac{(\lambda p e^{-\lambda})}{x!}\sum_{y=x}^{\infty}\frac{((1-p)\lambda)^{t}}{(t)!} \\
&= \frac{(\lambda p e^{-\lambda})}{x!}e^{(1-p)}\lambda \\
&= \frac{(\lambda p e^{-\lambda})}{x!}e^{-\lambda p},
\end{aligned}
$$

so X $\sim$Poisson(λp) And from this is easily seen that

$$EX = \lambda p.$$

Now if X is a random variable with Generalized Gamma density in the form of (5.2) we will use the notation $X \sim GG(\alpha, \beta, \delta)$. We will also write $c(\alpha, \beta, \delta) = \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)}$ to refer to the constant in the generalized gamma density. Let's consider $Y|\Lambda \sim$ Poisson(Λ) and $\Lambda \sim$ Gg(α, β, δ). Then the distribution of Y is given by

$$f(y) = \int_0^\infty \frac{\lambda^y e^{-y}}{y!}\cdot c(\alpha,\beta,\delta)\cdot\lambda^{\alpha-1}\cdot e^{-(\frac{\lambda}{\beta})^\delta}d\lambda.$$

To evaluate this integral we wiil use the fact that $e^{-x} = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}x^k$. So for this case we have

$$
\begin{aligned}
f(y) &= \int_0^\infty \frac{\lambda^y e^{-\lambda}}{y!}\cdot c(\alpha,\beta,\delta)\cdot\lambda^{\alpha-1}\cdot\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}(\frac{\lambda}{\beta})^{\delta k}d\lambda \\
&= \frac{c(\alpha,\beta,\delta)}{y!}\int_0^\infty e^{-\lambda}\lambda^{y+\alpha-1}\cdot\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}(\frac{\lambda}{\beta})^{\delta k}d\lambda \\
&= \frac{c(\alpha,\beta,\delta)}{y!}\sum_{k=0}^{\infty}\int_0^\infty\frac{(-1)^k}{k!}(\frac{\lambda}{\beta})^{\delta k}\cdot\lambda^{y+\alpha-1}\cdot e^{-\lambda}d\lambda \\
&= \frac{c(\alpha,\beta,\delta)}{y!}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\beta^k}\int_0^\infty\lambda^{\delta k+y+\alpha-1}\cdot e^{-\lambda}d\lambda.
\end{aligned}
$$

Seeing the integral in the right hand side as $\Gamma(\delta k + y + \alpha)$ we get

$$f(y) = \frac{c(\alpha,\beta,\delta)}{y!}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\beta^k}\cdot\Gamma(\delta k+y+\alpha). \tag{8.1}$$

So we can get the distribution of Y just by changing parameters. For example if $\delta = \beta = 1$ we have that Λ is a random variable with standard gamma density, so

$$f(y) = \frac{c(\alpha,1,1)}{y!} \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \Gamma(k+y+\alpha) = \frac{1}{\Gamma(\alpha) y!} \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \Gamma(k+y+\alpha).$$

If $\alpha = \delta = 1$ we have an exponential distribution. So (8.1) becomes

$$\begin{aligned}
f(y) &= \frac{1}{\beta y!} \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{\beta^k k!} \Gamma(k+y+1) \\
&= \frac{1}{\beta y!} \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{\beta^k k!} \int_0^{\infty} \lambda^{k+y} e^{-\lambda} d\lambda \\
&= \frac{1}{\beta y!} \int_0^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{\beta^k k!} \cdot \lambda^k \lambda^y \cdot e^{-\lambda} d\lambda \\
&= \frac{1}{\beta y!} \int_0^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (\frac{\lambda}{\beta})^k \cdot \lambda^y \cdot e^{-\lambda} d\lambda \\
&= \frac{1}{\beta y!} \int_0^{\infty} e^{\frac{\lambda}{\beta}} \cdot e^{-\lambda} \cdot \lambda^y d\lambda = \frac{1}{\beta y!} \int_0^{\infty} e^{-\left(1+\frac{1}{\beta}\right)\lambda} \cdot \lambda^y d\lambda.
\end{aligned}$$

Now making the change of variable $u = (1 + \frac{1}{\beta})\lambda$ we arrive to

$$\begin{aligned}
f(y) &= \frac{1}{\beta \cdot y!} \int_0^{\infty} e^{-u} \cdot \frac{u^y}{(1+\frac{1}{\beta})^y} \cdot \frac{\beta}{\beta+1} du \\
&= \frac{1}{(1+\frac{1}{\beta})^y \cdot (\beta+1) \cdot y!} \int_0^{\infty} e^{-u} u^y du \\
&= \frac{\Gamma(y+1)}{(1+\frac{1}{\beta})^y \cdot (\beta+1) \cdot y!} = \frac{1}{(1+\frac{1}{\beta})^y \cdot (\beta+1)}.
\end{aligned}$$

Remember that for any two random variables we can relate them with the following formula

$$E(X) = E(E(X|Y)), \tag{8.2}$$

and

$$Var(X) = E(Var(X|Y)) + Var(E(X|Y)), \tag{8.3}$$

provided that the expectations exist. For the same case from above we have that

$$E(Y) = E(E(Y|\Lambda)) = E(\Lambda) = \frac{\beta \Gamma(\alpha + \frac{1}{\delta})}{\Gamma(\alpha)},$$

which is just the expected value given by (6.3).

The variance for Y is given by

$$\begin{aligned} Var(Y) &= E(Var(Y|\Lambda)) + Var(E(Y|\Lambda)) \\ &= E(\Lambda) + Var(\Lambda) = E(\Lambda) + E(\Lambda^2) - (E(\Lambda))^2 \\ &= \frac{\beta\Gamma(\alpha+\frac{1}{\delta})}{\Gamma(\alpha)} + \frac{\beta^2\Gamma(\alpha+\frac{2}{\delta})}{\Gamma(\alpha)} + \frac{\beta^2\Gamma(\alpha+\frac{1}{\delta})^2}{\Gamma(\alpha)^2}. \end{aligned}$$

Chapter 9

Random Numbers Generation for a Generalized Gamma Distribution

If we set the change of variable $y = (x/\beta)^\delta$ in (5.2) a standard gamma distribution with parameter α is obtained.

$$\begin{aligned} f(x,\alpha,\beta,\delta)dx &= \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)} \cdot x^{\delta\alpha-1}e^{-(\frac{x}{\beta})^\delta} \\ &= \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)} \cdot (\beta y^{1/\delta})^{\delta\alpha-1}e^{-(\frac{\beta y^{1/\delta}}{\beta})^\delta} \cdot \frac{\beta y^{\frac{1-\delta}{\delta}}}{\delta} (\text{ because } dx = \frac{\beta y^{\frac{1-\delta}{\delta}}}{\delta}) \\ &= \frac{1}{\Gamma(\alpha)} \cdot y^{\alpha-1/\delta+1/\delta-1} \cdot e^{-y} \\ &= f(y) = \frac{1}{\Gamma(\alpha)} \cdot y^{\alpha-1}e^{-y}. \end{aligned}$$

So we can generate random numbers for y (which is easy with the adequate software) in this gamma density and then relate this numbers to the generalized gamma by means of the equation

$$x = \beta y^{1/\delta}. \tag{9.1}$$

This can work to estimate the cummulative distribution when $X \sim GG(\alpha,\beta,\delta)$. That is, for any $t > 0$

$$P(X \le t) = \int_0^t \frac{\delta}{\beta^{\delta\alpha}\Gamma(\alpha)} \cdot x^{\delta\alpha-1}e^{-(\frac{x}{\beta})^\delta}dx. \tag{9.2}$$

If we generate n random numbers $x_1, x_2, \ldots x_n$ from X, $t > 0$ then the empirical distribution is defined as

$$\widehat{F}_n(x) = \sum_{i=1}^n \mathbf{1}\{x_i \le t\}. \tag{9.3}$$

Here $\mathbf{1}\{x\}$ is the indicator function which is equal to 1 if $x \leq t$ and 0 otherwise. Consider for example $X \sim GG(2,4,2)$. Generating $n = 5000$ for a standard gamma with parameter $\alpha = 2$ and using (9.1) we obtain the next table

Table 9.1: Empirical distribtuion for a generalized gamma random variable. $\alpha = 2, \beta = 4$, $\delta = 2$, $E(X) = 5.31$, $Var(X) = 3.725$

t	Empirical distribution
1	.0024
2	.0264
3	.1070
4	.2648
5	.4644
6	.6586
7	.8054
8	.9030
9	.9592
10	.9840

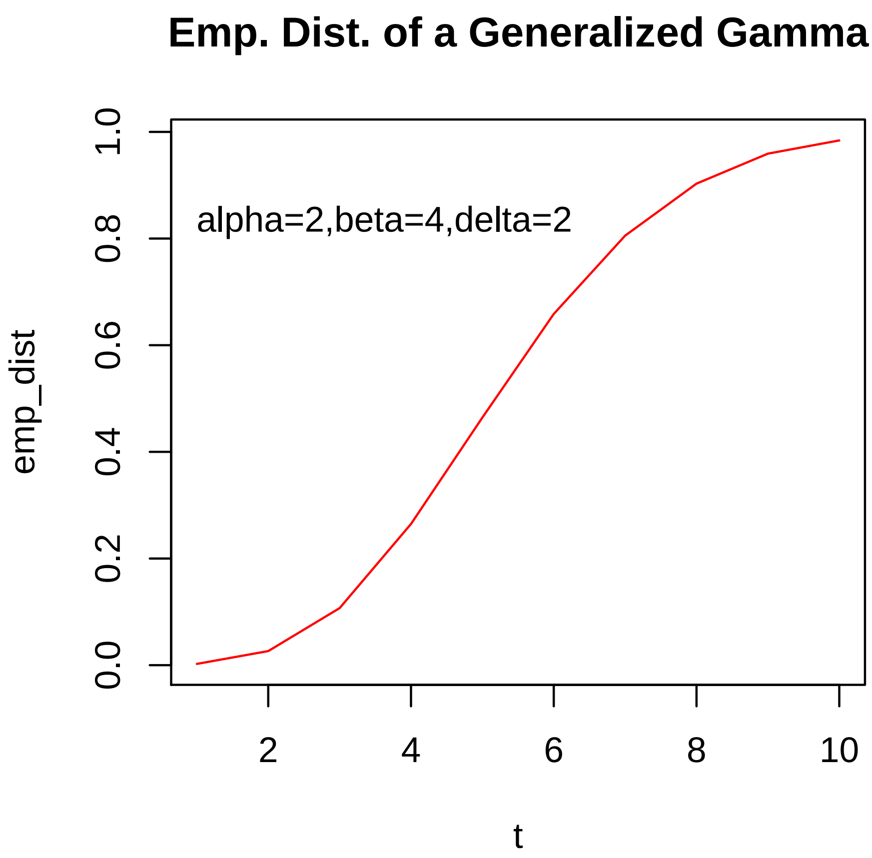


Figure 9.1: Empirical distribution for a Generalized Gamma

9.1 Estimation of the sum of n Generalized Gamma Distributions

As we have (9.1) to obtain the empirical distribution for a generalized gamma, we can obtain empirical distribution for $Y = X_1 + X_2 + \ldots + X_n$, where $X_i \sim GG(\alpha_i, \beta_i, \delta_i)$ for $i = 1, 2, \ldots n$. And we can use the normal approximation we used in (4.13) and see how it works here for a sum of generalized gamma distributions.

Table 9.2: Normal approximation for the sum of the exact distribution of 2 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \beta_1 = 4, \beta_2 = 3, \delta_1 = 2, \delta_2 = 3$, $E(Y) = 10.33$, $Var(Y) = 4.3$

t	Empirical	Normal Approximation	Difference
5	.0011	.0013	-.0002
7	.0436	.0429	.0005
9	.2736	.2708	.0028
11	.6446	.6442	.0004
13	.8927	.8957	-.0030
15	.9810	.9809	.0001
17	.9981	.9977	.0004
19	.9998	.9998	.0000

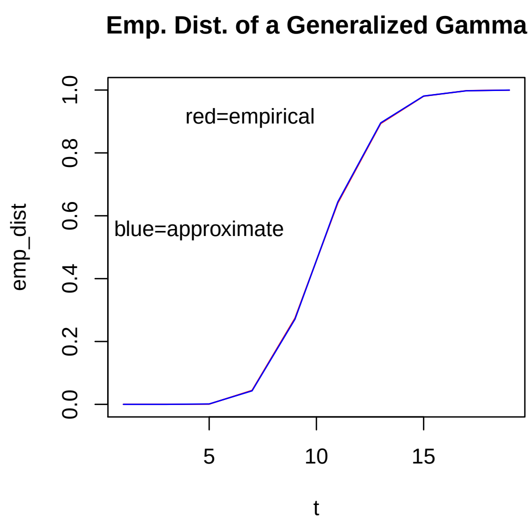


Table 9.3: Normal approximation for the sum of the exact distribution of 2 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \beta_1 = 4, \beta_2 = 3, \delta_1 = .7, \delta_2 = .8$, $E(Y) = 35.40$, $Var(Y) = 326.92$

t	Empirical	Normal Approximation	Difference
11	.0305	.0265	.0040
21	.2162	.2113	.0049
31	.4763	.4766	-.0003
41	.6914	.6933	-.0018
51	.8310	.8324	-.0014
61	.9120	.9119	.0001
71	.9537	.9768	-.0008
81	.9764	.9768	-.0004
91	.9883	.9882	.0002

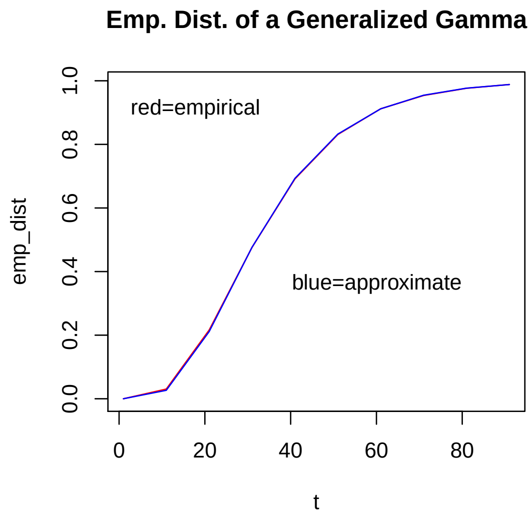


Table 9.4: Normal approximation for the sum of the exact distribution of 5 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \alpha_3 = 10, \alpha_4 = 3, \alpha_5 = 78, \beta_1 = 4, \beta_2 = 3, \beta_3 = 81, \beta_4 = 37, \beta_5 = 125, \delta_1 = 2, \delta_2 = 3, \delta_3 = 4, \delta_4 = 5, \delta_5 = 6$, $E(Y) = 456.012$, $Var(Y) = 188.68$

t	Empirical	Normal Approximation	Difference
410	.0006	.0006	.0000
420	.0054	.0051	.0003
430	.0313	.0308	.0006
440	.1233	.1224	.0009
450	.3268	.3279	-.0010
460	.6087	.6109	-.0021
470	.8446	.8459	-.0013
480	.9622	.9614	.0009
490	.9942	.9942	.0000

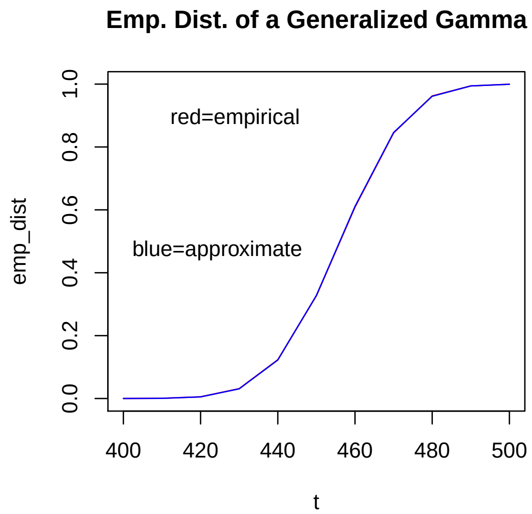
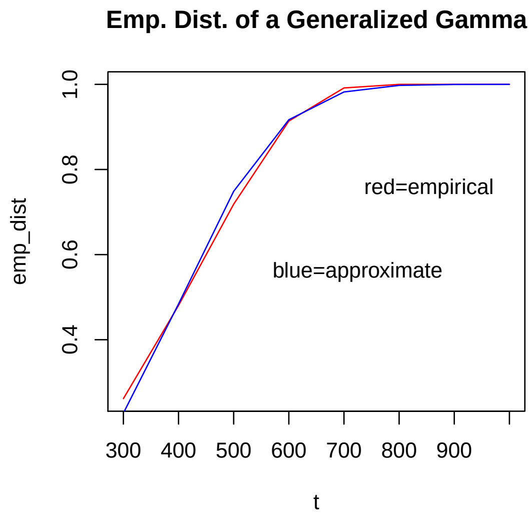


Table 9.5: Normal approximation for the sum of the exact distribution of 10 generalized gamma variables. $\alpha_1 = 2, \alpha_2 = 5, \alpha_3 = 10, \alpha_4 = 3, \alpha_5 = 7, \alpha_6 = .20, \alpha_7 = .9, \alpha_8 = 2.5, \alpha_9 = .3, \alpha_10 = .05 \beta_1 = 4, \beta_2 = 3, \beta_3 = 5, \beta_4 = 6, \beta_5 = 2, \beta_6 = 500, \beta_7 = 18 \beta_8 = 25, \beta_9 = 20, \beta_10 = 100.5, \delta_1 = 2, \delta_2 = 3, \delta_3 = 4, \delta_4 = 5, \delta_5 = 6, \delta_7 = 8, \delta_8 = 9, \delta_9 = 10, \delta_10 = 11$, $E(Y) = 405.1474$, $Var(Y) = 19932.65$.

t	Empirical	Normal Approximation	Difference
300	.2572	.2275	.0298
400	.4798	.4840	-.0042
500	.7174	.7487	-.0312
600	.9127	.9168	-.0040
700	.9915	.9821	.0094
800	.9998	.9976	.0022
900	1	.9998	.0002
1000	1	1	.0000



References

[1] Continuous Univariate Distributions, 1994, Norman L. Johnson, Samuel Kotz, N. Balakrishnan, Second edition, Vol. 1 Wiley Series in Probability and Mathematical Statistics.

[2] Amorosol, (1925). Ricerche intorno alla curva dei redditi. Ann. Mat. Pura Appl. Ser. 4 21 123-159

[3] D'Addario, R. (1932). Intorno alla curva dei redditi di Amoroso. Riv. Italiana Statist. Econ. Fnanza, anno 4, No. 1

[4] Van B. Parr,J.T. Webster, A method for discriminating between Failure Density Functions Used in Reliability Predictions,Technometrics, Vol 7. No. 1 (Feb. 1965), pp 1-10

[5] S. C. Choi and R. Wette, (1969) Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias, Technometrics, 11(4) 683690

[6] P.C Consul and G. C. Jain, Statistische Theorie,On the Log-Gamma Distribution and its Properties

[7] Carlos A. Coelho and Joao t. Mexia, (2007), On the Distribution of Product and Ratio of Independent Generalized Gamma-Ratio, The Indian Journal of Statistics (2003-), Vol. 69, No.2 pp 221-255.

[8] John E. Freunds, Mathematical Statistics with Applications, Seventh Edition, Pearson Prentice Hall (2004), Upper Saddle River, New Jersey 221-255.

[9] P.G. Moschopoulos, (1985) The Distribution of the sum of Independent Gamma Random Variables, Annals of the Institute of Statistical Mathematics, 37, Part A, 541-544

[10] P.G. Moschopoulos, W.B. Canada (1984) The distribution function of a linear combination of Chi-squares, Comps and Maths with applications, Vol 10. No. 4/5 pp. 383-386

[11] P.G. Moschopoulos,(1983) On a new Transformation to Normality, Communication Statistics-Theory and Methods, 12(16), pp. 1873-1878

[12] P.G. Moschopoulos, Govind, S. Mudholkar, (1983) A Likelihood Ratio Based Normal Approximation for the Non-Null Distribution of the Multiple Correlation Coefficient, Communication Statistics-Simulations and Computations, 12(3), pp 355-371

[13] Pham-Gia, Turkkan, N. (2002). Operations on the generalized-F variables and applications. Statistics, 36, 195-209.

[14] Roberts,C.D. (1971). On the distribution of random variables whose mth absolute power is gamma, Sankhya, Series A, 33, pp. 229-232

[15] E.W. Stacy, (1962) A generalization of the gamma distribution. The Annals of Mathematical Statistics, Vol. 33, No. 3, pp. 1187-1192

[16] E.W. Stacy and G. A. Mihram. (1965) Estimation for a generalized gamma distribution, Technometrics, Vol 7 No.3 pp.349-358

[17] Stacy E. W. Mihram G.A. (1965). Parameter estimation for a generalized gamma distribution, Technometrics, 7, pp. 349-358.

[18] D.R. Jensen and Herbert Solomon (1972) A Gaussian Approximation to the Distribution of a Definite Quadratic Form, JASA, Vol.67. No. 340, pp 898-902.

Chapter 10

Appendix

All the codes have been written in the software R.

10.1 Code to Approximate the sum of k Independent Gamma Random Variables

```
#Number of terms
n=10

#Number of generation for the random numbers
n2=1000

#Setting the parameters
alpha=c(1:n)
beta=c(1:n)

alpha[1]=2
alpha[2]=2.5
alpha[3]=3
alpha[4]=3.5
alpha[5]=4
alpha[6]=4.5
```

```
alpha[7]=5
alpha[8]=5.5
alpha[9]=6
alpha[10]=6.5


beta[1]=2
beta[2]=2.5
beta[3]=3
beta[4]=3.5
beta[5]=4
beta[6]=4.5
beta[7]=5
beta[8]=5.5
beta[9]=6
beta[10]=6.5


#Normal approx

k1=sum(alpha*beta)
k2=sum(alpha*beta^2)
k3=sum(alpha*beta^3)
phi_2=k2/k1
phi_3=k3/k1
phi_4=k4/k1
```

```
h=1-k1*k3/(3*k2^2)

mu_h=1+h*(h-1)*phi_2/(2*k1)+h*(h-1)*(h-2)*(4*phi_3+3*(h-3)*phi_2^2)/(24*k1^2)
sigma_h=phi_2*h^2/(k1)+(h-1)*h^(2)*(2*phi_3+(3*h-5)*phi_2^2)/(2*k1^2)

#Estimation with random numbers
Gen=matrix(0,n,n2)
for(i in 1:n)
{
Gen[i,]=rgamma(n2,alpha[i],1/beta[i])
}

#Estimation of alpha
alpha_est=c(1:n)
s_est=c(1:n)
for(i in 1:n)
{
s_est[i]=log(mean(Gen[i,])) -sum(log(Gen[i,]))/n2
alpha_est[i]=(3-s_est[i]+sqrt((s_est[i]-3)^2
+24*s_est[i]))/(12*s_est[i])
}


#Estimation of the parameter "beta"
beta_est=c(1:n)
for(i in 1:n)
{
beta_est[i]=mean(Gen[i,])/(alpha_est[i])
```

```
}

beta_est
alpha_est

#Exact distribution
#Number of terms we will use in the infinite sum
k=100

#Length of the subintervals the interval is divided in
c=.1
#Right extreme of the interval
lim_r=300


#Setting the minimum between beta's
beta_min=min(beta)

#Setting the coefficient rho
rho=sum(alpha)


#Coefficients C and Lower gamma
beta_min_ov_beta=(beta_min^{rho})/(prod(beta^{alpha}))
C=prod(beta_min_ov_beta)
G=1-beta_min/beta
lower_gamma_matrix=matrix(1,n,(k+1))
for(j in 1:n)
```

```
{
for(i in 1:(k+1))
{
lower_gamma_matrix[j,i]=(alpha[j]*(G[j])^{i})/i
}
}
lower_gamma=colSums(lower_gamma_matrix)

#Setting the delta's
delta=c(1:(k+1))
delta[1]=lower_gamma[1]
for(j in 2:(k+1))
{
b=c(1:(j-1))
for(v in 1:(j-1)){b[v]=v*lower_gamma[v]*delta[(j-v)]}
d=sum(b)+j*lower_gamma[j]
delta[j]=(1/j)*(d)
}

#Domain
y=seq(0,lim_r,by=c)
#Density
density_sum_gamma=matrix(1,(k+1),(1/(c/lim_r))+1)
for(u in 1:(k+1))
{
density_sum_gamma[u,]=((gamma(rho+u)*(beta_min)^{rho+u})^{-1})
*delta[u]*y^{rho+u-1}*exp(-y/beta_min)
}
```

```
E=C*colSums(density_sum_gamma)+C*y^{rho-1}*exp(-y/beta_min)
*((gamma(rho)*(beta_min)^{rho})^{-1})


#Approximate
expect_z=sum(alpha*beta)
var_z=sum(alpha*beta^{2})
beta_z=var_z/expect_z
alpha_z=expect_z^(2)/var_z
App=1/(beta_z^{alpha_z}*gamma(alpha_z))*y^{alpha_z-1}*exp(-y/beta_z)


#Graphs for the distribution
#Exact
pgamma_vector_ex=c(1:lim_r)
for(j in 1:lim_r)
{
A=c(1:(k+1))
for(i in 1:(k+1))
{
A[i]=delta[i]*pgamma(j,rho+i,1/beta_min)
}
distribution_1=c(A,pgamma(j,rho,1/beta_min))
pgamma_vector_ex[j]=C*sum(distribution_1)
}
domain=c(1:lim_r)
Ex=pgamma_vector_ex
```

```
#Approximate
pgamma_vector_app=c(1:lim_r)
for(j in 1:lim_r)
{
pgamma_vector_app[j]=pgamma(j,alpha_z,1/beta_z,)
}
App=pgamma_vector_app


#Estimated
pgamma_vector_est=c(1:lim_r)
for(j in 1:lim_r)
{
pgamma_vector_est[j]=pgamma(j,alpha_z_est,1/beta_z_est)
}
Est=pgamma_vector_est


#Estimated_h
k1=sum(alpha*beta)
k2=sum(alpha*beta^2)
k3=2*sum(alpha*beta^3)
phi_2=k2/k1
phi_3=k3/k1
phi_4=k4/k1


h=1-k1*k3/(3*k2^2)
```

```
#h=1/3
mu_h=1+h*(h-1)*phi_2/(2*k1)+h*(h-1)*(h-2)*(4*phi_3+3*(h-3)*phi_2^2)/(24*k1^2)
sigma_h=phi_2*h^2/(k1)+(h-1)*h^(2)*(2*phi_3+(3*h-5)*phi_2^2)/(2*k1^2)
pgamma_vector_est_h=c(1:lim_r)
for(j in 1:lim_r)
{
z_h=((j/k1)^{h} -mu_h)/(sqrt(sigma_h))
pgamma_vector_est_h[j]=pnorm(z_h)
}
Est_h=pgamma_vector_est_h


par(mfrow=c(1,2))
plot(domain,Ex,type="l",col="red",xlab="x",main="Exact distribution")
#plot(domain,App,type="l",col="blue",xlab="x",main="Two moments app")
plot(domain,Est_h,type="l",col="brown",xlab="x",main="Normal Approx")
#plot(domain,Ex_est,type="l",col="black")
#plot(domain,Est,type="l",col="green")

#Distribution several times
#Approximate
a_1=180
b_1=340
c_1=20
vector_t=seq(a_1,b_1,by=c_1)
Approx_1=c(1:((b_1-a_1)/c_1 +1))
for(i in 1:((b_1-a_1)/c_1+(1)))
{
```

```
Approx_1[i]=pgamma(vector_t[i]/beta_z,alpha_z)
}
#Exact_dist=C*sum(distribution)

#Exact
Exact_1=c(1:((b_1-a_1)/c_1 +1))
for(j in 1:((b_1-a_1)/c_1+(1)))
{
A=c(1:(k+1))
for(i in 1:(k+1))
{
A[i]=delta[i]*pgamma(vector_t[j],rho+i,1/beta_min,)
}
distribution_2=c(A,pgamma(vector_t[j],rho,1/beta_min))
Exact_1[j]=C*sum(distribution_2)
}


#Estimated exact
Exact_1_est=c(1:((b_1-a_1)/c_1 +1))
for(j in 1:((b_1-a_1)/c_1+(1)))
{
A_est=c(1:(k+1))
for(i in 1:(k+1))
{
A_est[i]=delta_est[i]*pgamma(vector_t[j],rho_est+i,1/beta_min_est)
}
distribution_2_est=c(A_est,pgamma(vector_t[j],rho_est,1/beta_min_est,))
```

```
Exact_1_est[j]=C_est*sum(distribution_2_est)
}

#Estimated with two moments
Estimated_1=c(1:((b_1-a_1)/c_1 +1))
for(i in 1:((b_1-a_1)/c_1+(1)))
{
Estimated_1[i]=pgamma(vector_t[i],alpha_z_est,1/beta_z_est)
}

#Estimated with h
Normal_App_1=c(1:((b_1-a_1)/c_1 +1))
z_h=c(1:((b_1-a_1)/c_1 +1))
for(i in 1:((b_1-a_1)/c_1+(1)))
{
z_h[i]=((vector_t[i]/k1)^{h} -mu_h)/(sqrt(sigma_h))
Normal_App_1[i]=pnorm(z_h[i])
}
Approx=round(Approx_1,digits=7)
Exact=round(Exact_1,digits=7)
Estimated=round(Estimated_1,digits=7)
Est_ex=round(Exact_1_est,digits=7)
Normal_App=round(Normal_App_1,digits=7)
Diff_1=Exact_1-Approx_1
Diff_E_A=round(Diff_1,digits=7)
Diff_10_3_1=Diff*1000
Diff_10_3=round(Diff_10_3_1)
Diff_2=Exact-Estimated
```

```
Diff_E_Est=round(Diff_2,digits=7)
Diff_3=Exact-Est_ex
Diff_E_EE=round(Diff_3,digits=7)
Diff_4=Exact_1-Normal_App_1
Diff_E_Normal=round(Diff_4,digits=7)

#Printing out results
#Table_results=cbind(vector_t,Exact,Approx,Est_ex,Estimated,Normal_App,
Diff_E_A,Diff_E_EE,Diff_E_Est,Diff_E_Normal)
#Table_results=cbind(vector_t,Exact,Approx,Diff_E_A)
#Table_results=cbind(vector_t,Exact,Est_ex,Diff_E_EE)
Table_results=cbind(vector_t,Exact,Normal_App,Diff_E_Normal)
#Table_results=cbind(vector_t,Exact_1,Approx_1,
Estimated_h_1,Diff_1,Diff_4)
Table=data.frame(Table_results)
#write.table(Table,"C:/Users/Hugo/Desktop/Hugo/Thesis/
Simulation_Independent_Sum_of_Gamma/ex_norm_tabla_1.txt",sep=",")
```

10.2 Code to Approximate the Sum of k Independent Generalized Gamma Random Variables

```
#Sum of Generalized Gamma
#Empirical distribution
#Number of terms
n=5

#Number of random data generated
n1=50000

#Standard gamma numbers
y_ran_data=matrix(0,n,n1)

#Parameters
alpha=c(1:n)
beta=c(1:n)
delta=c(1:n)

alpha[1]=2
alpha[2]=5
alpha[3]=10
alpha[4]=3
alpha[5]=78
#alpha[6]=.2
#alpha[7]=.9
#alpha[8]=2.5
#alpha[9]=.3
```

```
#alpha[10]=.05

beta[1]=4
beta[2]=3
beta[3]=81
beta[4]=37
beta[5]=125
#beta[6]=500
#beta[7]=18
#beta[8]=25
#beta[9]=20
#beta[10]=100

#m=2
delta[1]=.5
delta[2]=.6
delta[3]=.3
delta[4]=.8
delta[5]=.2
#delta[6]=m+5
#delta[7]=m+6
#delta[8]=m+7
#delta[9]=m+8
#delta[10]=m+9

#Gen gamma numbers
x_vectors=matrix(0,n,n1)
for(i in 1:n)
```

```
{
y_ran_data[i,]=rgamma(n1,alpha[i])
x_vectors[i,]=beta[i]*(y_ran_data[i,]^{1/delta[i]})
}

sum_gen_gamma=colSums(x_vectors)
a_1=300
b_1=1000
c_1=100
vector_t=seq(a_1,b_1,by=c_1)
emp_dist=c(1:((b_1-a_1)/c_1 +1))
for(i in 1:((b_1-a_1)/c_1 +1))
{
cont=0
for(j in 1:n1)
{
if(sum_gen_gamma[j]<=vector_t[i]){cont=cont+1}
}
emp_dist[i]=cont/n1
}

#Cumulants
mean=sum(beta*(gamma(alpha+(1/delta)))/gamma(alpha))
var=sum(beta^{2}*(gamma(alpha+(2/delta)))/gamma(alpha)-
(beta*(gamma(alpha+(1/delta)))/gamma(alpha))^2)

k1=mean
k2=var
```

```
k3=sum(beta^{3}*gamma(alpha+(3/delta))/(gamma(alpha))-3*beta^{2}
*(gamma(alpha+(2/delta)))/gamma(alpha)
*(beta*(gamma(alpha+(1/delta)))/gamma(alpha))+2*(beta*(gamma(alpha+(1/delta)))
/gamma(alpha))^(3))
phi_2=k2/k1
phi_3=k3/k1
phi_4=k4/k1


h=1-k1*k3/(3*k2^2)

mu_h=1+h*(h-1)*phi_2/(2*k1)+h*(h-1)*(h-2)*(4*phi_3+3*(h-3)*phi_2^2)/(24*k1^2)
sigma_h=phi_2*h^2/(k1)+(h-1)*h^(2)*(2*phi_3+(3*h-5)*phi_2^2)/(2*k1^2)
Normal_App_sum_gg=c(1:((b_1-a_1)/c_1 +1))
z_h=c(1:((b_1-a_1)/c_1 +1))
for(i in 1:((b_1-a_1)/c_1+(1)))
{
z_h[i]=((vector_t[i]/k1)^{h} -mu_h)/(sqrt(sigma_h))
Normal_App_sum_gg[i]=pnorm(z_h[i])
}

diff=emp_dist-Normal_App_sum_gg

results=cbind(vector_t,emp_dist,round(Normal_App_sum_gg,4),round(diff,4))
#results=cbind(vector_t,emp_dist)
Results=data.frame(results)
#write.table(Results,"C:/Users/Hugo/Desktop/Hugo/Thesis
/Simulation_Independent_Sum_of_Gamma/sum_gen_gamm_7.txt",sep=",")
```

Results

```
plot(vector_t,emp_dist,type="l",col="red",xlab="t",main
="Emp. Dist. of a Generalized Gamma")
lines(vector_t,Normal_App_sum_gg,col="blue")
text(locator(1),"red=empirical")
text(locator(1),"blue=approximate")
```

10.3 Bedfordshire Rainfall Statistics Table

Source:http://www.zen40267.zen.co.uk/rainfall/rainfall.html. Table made by Andrew Leaper

Table 10.1: Rainfall Statistics in mm

Month	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999
January	46.3	36.8	70.6	668.5	19.6	19	69.4	32.8	37.3	57.9	26.3	67.1
February	82	60.1	22.6	66.8	28.4	43.5	30.5	19.3	71.2	69.7	41.3	16
March	33	22.2	83.8	32.7	40.7	55.1	33.4	19.4	37.4	81.6	43.7	38.1
April	15.9	30.1	34.3	3.7	43.4	30.1	77.9	19.4	24.8	93.7	123	43.1
May	51.1	31.2	110	146	84.5	22.1	48.7	43.1	62.4	58.7	83.8	49.6
June	43.7	55.4	70.7	136	8.9	57.6	29.1	43.7	33.4	22.8	24.9	86.3
July	22.9	121	77.6	110	57.5	43	63.2	67.6	80	89.4	55.8	18.1
August	173	65.7	91.4	31.5	73.2	42.8	125	0	75.9	58.7	55.6	64.9
September	58.5	9.3	48.6	34.3	62.9	66.3	12.7	22.7	21.3	67.3	48.2	56.9
October	44.6	42.4	56.6	93	101	55.9	109	36.5	102	93.4	153	53.1
November	36.5	95.8	80.6	50.3	90.9	37.8	55.8	88.1	102	28.6	103	14.1
December	22.2	78.5	23.9	32.9	58.1	29.7	21.4	59	128	39.4	98.1	82.9

Table 10.2: Rainfall Statistics in mm (continuation)

Month	1998	1997	1996	1995	1994	1993	1992	1991	1990	1989	1987
January	56.8	15.7	39.6	86.9	67.8	70.8	62.1	51.3	41.9	37.7	26.3
February	4.3	27.8	55.5	69.8	38.3	5.8	13.2	15.6	85.4	36.6	90.3
March	53.5	20.4	26.5	53.1	32.5	15	33	61.4	18	18	60.9
April	143.5	15	28.4	27.8	61.7	89.6	46.9	24.3	31.8	111.7	123
May	28.1	112.7	17.9	29.4	79.2	54.4	115.2	37	1	16.6	83.8
June	137.1	99	17	25.8	23.5	76.3	48.4	94.5	30.6	26.8	24.9
July	8.6	21.2	17	47	16.2	52.9	96.5	37.4	29.2	76.7	55.8
August	33.1	59.9	43.2	2.7	33.1	50.7	107.1	53.4	26.3	32.8	55.6
September	104.3	15.9	52.4	106.6	109.1	68	143.7	102.5	22.9	35.1	48.2
October	63.5	62.1	18.6	33.2	54.8	122.4	74.9	15.6	55.3	36.5	153
November	109.9	63.5	47.8	41.8	45.2	46.65	143.3	66.6	34.2	46.2	103
December	55.3	72.1	39.4	104.2	72.7	99.5	54.7	15.7	54.8	120.3	98.1

Table 10.3: Rainfall Statistics in mm (continuation)

Month	1988	1987	1986	1985	1984
January	97.7	32.5	54.5	42	51.3
February	25.7	17	35	26.8	38.5
March	56.8	60.55	43.1	28.9	39.5
April	15.8	56	75.5	22.5	20.5
May	65.2	47.7	71	50.3	57.5
June	42.1	132	22.5	98	69
July	145	52	41.5	24.2	22
August	13.8	89	85.5	65	46.5
September	71.5	33	25.5	15.5	79
October	59.3	155	43	24	37.5
November	13.8	69.4	79.5	39.5	87.5
December	42.8	13.66	30.5	101	43.5

Curriculum Vitae

Víctor Hugo Jiménez Nava was born on May 8, 1984. The second son of Jorge Sabino Jiménez and María Elva Nava Morales, he graduated from Juarez Autonomous University, Juárez, Chihuahua México, in the spring of 2009. He obtained a bachelor in mathematics. He has attended research summer programs in Mexico. He went to Colima, Colima at the Colima Autonomous University in 2007 and Guanajuato, Guanajuato, at the Research Center in Mathematics in 2008. He entered University of Texas at El Paso in the fall of 2009. While pursuing a master's degree in statistics he worked as a Teaching Assistant, and as instructor for two mathematical courses directed to freshman students. He submitted this work on december 2011 as a part of the requirements to complete his master in statistics at the University of Texas at El Paso.