UNIVERSITY OF NAIROBI

PRICING UNEMPLOYMENT INSURANCE USING BURR XII MIXTURE DISTRIBUTIONS AND C.A.P.M WITH APPLICATION TO USA DATA

Mutwiri Martin Kithinji

I56/75648/2014

August 1, 2016

A dissertation submitted to the graduate school for partial requirements for the award of master of science in actuarial science

Declaration

I do hereby declare that this work is based on a study I took, with reference to other people's work, which has been duly recognized/ acknowledged.

I certify that this is my original work which has not been presented in any other institution.

NAME	REGISTRATION NUMBER	SIGNATURE	DATE
Mutwiri Martin Kithinji	I56/75648/2014		

In the prescence of my supervisors:

Prof.R.O. Simwa _________________________ _________________________

Prof J.A.M. Otieno _________________________ _________________________

Abstract

The objective of this research is to consider varying unemployment duration in the pricing of unemployment insurance with application to USA data. The study assumes that unemployment duration follows Burr XII mixture distribution while the discount rate to use in the pricing of the scheme will be determined by fitting market data into the capital asset pricing model. The Burr XII mixture distribution has been used to model unemployment duration in order to allow for heterogeneity in the unemployment duration of the covered employees. The program will be administered by the government to cover her working citizenry so that in the event of an involuntarily job loss, one may receive unemployment benefits to help them pay their recurrence bills before they secure another job.The results yield a mean unemployment duration of approximately 16 weeks and premium contribution rate of 5.10% of the taxable wage base per month for a benefit of 45% of the taxable wage base per month, payable on weekly basis during spells of unemployment.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Insurance

Insurance is a financial service that involves the transfer of risk from one party, referred to as the insured, to another party, the insurer. The price paid by the insured for the security provided is referred to as the premium which may be a single payment or a series of payments. In return for the premiums, the insurer compensates the insured following the occurrence of the insured event. However, no payment is made if the insured event does not occur.
Through selling insurance policies to a large number of policyholders, the insurer spreads the risk to a large number of exposed individuals which by law of large numbers in statistics ensures that actual results converge to the expected value. The law allows the use of the expected value principle in determining the premiums to charge for a certain level of coverage.

1.2 Background

1.2.1 Unemployment Insurance

Unemployment insurance is a social welfare scheme that provides compensation for lost wages to workers during spells of involuntarily unemployment. The amount of benefits together with the level and method of contributions to the unemployment insurance fund depends on the design of the unemployment scheme. The unemployment insurance program is designed to only compensate employable individuals who are able and willing to work and who become unemployed through no fault of their own.
Research has shown that most unemployment insurance schemes are state run to benefit their citizens. However, there are some private insurers and labor unions that also provide unemployment covers. A Ghent system is a special case in which the unemployment benefits are distributed by labor unions or independent agencies.

The first unemployment insurance plans, supported by members' contributions, were adopted by some large trade unions in Switzerland in 1789. The idea spread across Europe and many such together with Ghent systems were established. Following popularization of the benefits brought about by the Unemployment Insurance programs, a movement began to develop national unemployment insurance. An attempt in 1895 was made to establish a compulsory unemployment insurance system in the Swiss canton of St. Gallen but it failed.

The first compulsory, national unemployment insurance system was adopted in Great Britain through the National Insurance Act of 1911. The act introduced a contributory scheme to cover the British workers against illness and unemployment. The contributions were a fixed amount from the workers, employers and taxpayers while the benefits, paid after one week of unemployment, were 7 shillings per week up to 15 weeks per year. The Unemployment Insurance Act 1920 ammended the National Insurance Act 1920 to create a dole system of payments for unemployed workers. The dole system covered over 11 million workers drawn from the entire civilian working population excluding civil servants, domestic workers, farm workers and rail workers and provided up to a maximum of 39 weeks of benefits payment.

In Germany, unemployment insurance was introduced in the year 1927 while in the United States, it originated from Wisconsin state in 1932 and its spread was catalyzed to all states by the federal government through enactment of the social security act in 1935. Most of the European countries later adopted unemployment insurance after the Second World War to promote equality, equity and public responsibility for those unable to avail themselves of the basic provisions for a good life.

Currently, most of the European, American and Asian countries have their own versions of Unemployment Insurance schemes. Only South Africa in Africa has unemployment insurance program, following promulgation of its first unemployment insurance act of 1966.

1.2.2 Unemployment Insurance in the United States of America

A. **Overview**

The Unemployment insurance scheme in the US is a federal-state partnership based upon federal law. The arrangement is anchored on a strong use of incentives to enhance efficiency. The Federal government ensures conformity and compliance of state programs through Federal Unemployment Tax Act (FUTA) and Social Security Act (SSA). States have enacted their own laws to regulate their individual schemes.

B. **Financing**

The program is entirely funded by employer contributions, both federal and state, although the states of Alaska, Pennyslvania and New Jersey levy unemployment contributions on employees to supplement employer contributions. Unemployment taxes to the federal government have been at a rate of 6% per annum of the first $\$7,000$ wage base per employee following the decline from 6.2% per annum in July 2011. Credit on contributions is available up to a maximum of 5.4% of FUTA taxable wages. However, the maximum tax discount is offered to employers who pay their respective state unemployment contribution in full, on time and on all the same as are subject to FUTA tax.

The Unemployment contributions to the federal government are used to pay for administrative costs incurred in running the Unemployment Insurance programs in all the states together with other associated programs, federal share of extended benefits and to pay for other third tier programs like loans to states with deficits in payment of benefits.

All states finance their Unemployment Insurance programs through contributions from subject employers on the wages of their covered employees. The contributions are deposited into the state's Unemployment Tax Fund (UTF) and are withdrawn by the state to pay the benefits or refunds on overpayment of contributions. Contrary to the federal contribution rate, most states use experience rate system to set the contribution rate for each employer. However, new employers are given a standard rate before their experience rate is determined. States sets their own tax base with some preferring to use the federal government's tax base.

C. **Eligibility**

An application brought to the State unemployment agency is reviewed to determine if the applicant qualifies to receive the benefits. To qualify, one must have worked for the base period or have earned the required wages as provided for in the State's labor laws and the cause of the unemployment must be out of control of the insured.

Upon commencement of the benefits, one must file weekly or biweekly claims and reports regarding any incomes from work or job offers refused as well as respond to any questions from the state labor office. Additionally, one must report to the Unemployment Insurance Claims office when required to do so.

D. **Waiting period**

Workers are required to file a claim with the Unemployment Insurance Agency of the state they worked for immediately they become unemployed. During claim, workers furnish the agency with the details of their immediate former employer to aid in authentication of the claim.According to the United States Department of Labor, it takes an average of two to three weeks after a claim is filed for one to receive the first benefits. However, some states take as low as one week to process a claim.

E. **Unemployment Benefits**

States pay a benefit of between 40%–50% of average monthly earnings in the past one year before unemployment up to a state's maximum amount. Benefits are advanced on weekly basis up to a maximum of 26 weeks unless in the case of extended benefits during periods of high unemployment.

1.3 Problem Statement

Chuang and Yu, 2010, assumed the Weibull distribution to model the duration of unemployment on the basis that Weibull distribution is commonly assumed in most literature on unemployment duration. The Weibull distribution however, assumes homogeneity among the subject population. This might lead to erroneous results due to the heterogeneous nature of unemployment duration data. The study seeks to improve the model of Chuang and Yu, 2010 by incorporating the results of McDonald and Butler, 1987. However, we will mix the Weibull distribution with the gamma distribution, instead of the inverse generalized gamma, to allow the rate parameter of the Burr XII distribution not to depend on shape1 parameter of the same distribution. This will in turn take care of the heterogeneity in unemployment duration as well as provide a better fit to the unemployment duration data. We will also apply the model to price unemployment insurance by determining the appropriate premium contribution rate for the cover.

1.4 Justification of study

In a bid to empower their citizenry and ensure a positive economic growth rate, most governments have continually invested heavily in the education sector. For example in Kenya, the introduction of the free primary education in the year 2003 followed by the onset of subsidized secondary education together with the establishment of several day secondary schools five years later goes a long way in improving the education sector. Moreover, several university colleges and polytechnics have been chartered to become fully pledged universities. These together with the good public and private partnerships have ensured high literacy levels in the country and availability of skilled labor in the Kenyan economy.

However, the high turnover from colleges and universities to the job market, which is not matched with high investments to create more jobs, has increased the number of skilled unemployed people in the country. The situation gets worse for those who involuntarily lose their jobs and are looking for reemployment opportunities. They take longer to secure another job due to the crowding of the job market. The lack of any source of income during spells of unemployment to meet even routine bills may lead to frustration and eventually depression in worse case scenarios. It also leads to idle productive labor which is detrimental to the economy of the country.

Unemployment insurance seeks remedies this by providing benefits to pay for routine bills and challenge the government to encourage investments for job creation. This is as a result of reduced employer contributions to the scheme following reduction in the amount of benefit payments. This will in turn attract more investors into the economy due to the reduction in contribution burden.

Moreover, just like we have medical insurance to cover the risk of health-care or even any other type of insurance for that matter, why not have unemployment insurance to cover the risk of unemployment?

1.5 Objectives of the study

The main aim of the study is to price a nationalized unemployment insurance scheme. In order to achieve this, the study will seek

1. To determine the best statistical distribution for modeling unemployment spells.
2. To determine the appropriate interest rate to use in discounting future cash flows.
3. To come up with an appropriate pricing formula and apply it in calculating the premium contribution rate.

The next chapter outlines previous research in the area of unemployment insurance pricing as well as statistical and financial concepts used in this study.

Chapter 2. Literature review

Malinvaud, 1985 while reviewing literature on unemployment insurance, alludes that unemployment insurance is an exceptional insurance contract because, it is obligatory and is offered by the government in most countries. In such a set-up, unemployment insurance is considered as a social program whose main goal is to provide unemployment benefits to partially replace lost earnings for previously working individuals who become involuntarily unemployed and who are able, available and actively seeking for employment. The program specifications differ from country to country. None the less, a common factor in most countries is the way the contributions to the unemployment insurance fund are mobilized. Many of the unemployment insurance schemes charge a uniform percentage of the worker's pay earned between some minimum and maximum levels. However, this is not a fair premium since all employees are not exposed to the same level of unemployment risk. In his review he also cites problems of moral hazard, dis-utility and adverse effects while classifying the risk groups.

Beenstock, 1985 developed a model to solve the above problems. In his model he diversified the unemployment risk and assumed that the unemployment benefits are deterministic. According to the model, the unemployment insurance contract would automatically be enacted when a person starts working and the insured was required to pay premiums right from the onset of employment. They would then receive unemployment benefits in the event that they become involuntarily unemployed until they secure another job, if this occurs before the contract expires. To be able to determine the amount of premiums payable for the cover, Beenstock assumed that the insurer has identified various risk groups, just as is the case in car insurance, and considered each risk group separately. Since the benefits are deterministic, then equating the discounted value of the benefits to that of the premiums gives the amount of premiums payable.

Bronars, 1985 uses capital asset pricing model to determine the fair premiums in a theoretical model of a hypothetically regulated private market for unemployment insurance. Bronars improved on the existing work of Beenstock by undiversifying the unemployment risk and specifying an appropriate risk-adjusted interest rate for the unemployment insurance.

Further research courtesy of Blake and Beenstock, 1988 led to the development of a more generalized unemployment insurance model by allowing the unemployment probability to be stochastic. This model was however not successful since it failed to estimate the unemployment benefits according to the duration of unemployment.

The unfortunate failure of Blake and Beenstock, 1988 motivated Chuang and Yu, 2010 to extend the work of Bronars, 1985 by incorporating survival analysis with a more general form to estimate the unemployment duration and to calculate the fair premium rate for the unemployment insurance program. The study used data from the unemployment insurance program in Taiwan. In the development of the model, the Weibull distribution was used to estimate the average unemployment duration while the capital asset pricing model was used to determine the interest rate used to discount the benefits.

Bowers, 1980 probed issues surrounding unemployment duration ranging from methodological, measurement and results interpretation on existing statistics on unemployment duration and observed that a substantive amount of unemployment spells are of short durations although with some fluctuations especially during recessions. He used transition probabilities among the three states of employed, unemployed and not in the labor force to estimate the duration of unemployment. According to him, the short unemployment durations do not imply an active labor market so that in the event of a job loss, one is able to find his usual type of a job in a relatively short period. This is because a large portion of job changes occurs without any intervening spells of unemployment. More so, the ambiguity in labor force classification, particularly in differentiating between the unemployed and not in the labor force states, is problematic. This is because some of those who withdraw from the labor force experience a brief spell outside and soon reenter the labor force as unemployed again.

Salant, 1977 developed a model to sort between complete and incomplete unemployment durations. To accommodate both spells, he suggested a mixture model. The sorting model considered, that is Pareto, assumed a constant individual hazard rate which was allowed to vary among different individuals. The constant hazard rate was accounted for by exponential distribution while the variation was accounted for by the gamma. The resulting mixture that is, Pareto, yielded a decreasing hazard rate for the whole cohort of unemployed individuals.

McDonald and Butler, 1987 reviewed several generalized distributions that can be applied to unemployment duration with their relations established. Statistical tests on Salant's model (Pareto) and Burr XII revealed that the Burr XII distribution, a mixture of Weibull distribution and the inverse generalized gamma, was better than Pareto in estimating the spells of unemployment while allowing for heterogeneity in unemployment data. This might be attributed to the varying hazard function of the Weibull distribution.

Cummins, 1991 points out the problem posed on insurance pricing due to parallelism in research on the three major paradigms of insurance. These are statistical modeling, financial modeling and economics. Although few attempts have been made to integrate research in the three areas, the technicality and high specialization exhibited in each have posed a great challenge in the exercise. Cummins, 1991 made an attempt to integrate the three by looking at both statistical and financial models and how they are applied in insurance together with some of the errors made in application. Some of the statistical models, concepts and laws looked at include individual and collective risk models, central limit theorem, law of large numbers and the concept of homogeneity of risks. The financial models explored include the application of capital asset pricing model in determining underwriting rate of return, discrete time discounted cash flow models, option pricing models and sensitivity analysis of the assets and liabilities of the insurance firm. Although economic models were not considered in his integration, he notes that financial models consider insurance variables in an economic setup which in a way incorporates economic models.

Stephen, 1988 applied the capital asset pricing model to discount loss reserves for an insurer with an objective of determining the tax liability of the insurer. The problem however, as is in many applications of C.A.P.M, was how to determine the beta of the loss reserves. This is because listed insurers deal with more than one type of cover and no market data is available for single coverage's. By splitting the beta of equity into beta of assets and beta of liabilities and using the industry average of beta of assets, Fairley determined beta of liabilities to be -0.21. Upon apportioning the -0.21 into various components of the liabilities using industry reported averages, Stephen found the beta of loss reserves according to market data to be -0.24. This together with the short term US Treasury bill rate as the risk free rate and market rate of return proxied from a broad market index was used to estimate the risk adjusted rate of return.

Wang, 2002 in his paper suggested a universal way of pricing financial and insurance risks using a Wang Transform on either the asset prices or returns. This was to address the assumption of normality of returns in Capital Asset Pricing Model. A Wang transform involves finding the standard normal inverse of the cumulative distribution function of the returns and then getting the standard normal of this after adding the risk premium. This gives the cumulative distribution function of the risk adjusted returns from which we can determine the expected risk adjusted return.

Delignette-Muller and Dutang, 2014 motivated by feedback from users of their earlier packages, developed fitdistrpus package in r programming to provide functions for fitting data into various probability distributions. By comparing the various fits, a researcher is able to select the distribution of best fit for the data under study. The package is also able to handle estimation of parameters of the fits using various methods such as maximum likelihood estimation.

Dutang, Goulet, Pigeon, et al., 2008 developed Package 'actuar' which is a collection of actuarial functions in the fields of loss modeling, risk theory, credibility theory and tailed distributions. This is aimed at improving functionality of normal statistical packages in solving complex actuarial models in the r programming environment.

Chapter 3. Methodology

This chapter deals with the derivation of the various models applied in this study.

3.1 Method of Mixtures

Singular distributions assume among other things, homogeneity in the data. This assumption can be relaxed by randomizing one of the parameters of the singular distribution to form a mixture distribution. Mixture distributions allow for heterogeneity hence providing better fits to heterogeneous data than singular distributions.

Let H be a distribution function depending on a parameter θ and g the probability density function of θ. Then

$$F(u) = \int_{-\infty}^{\infty} H(u,\theta)g(\theta)d\theta \tag{3.1}$$

is a monotone function of X increasing from 0 to 1 and hence a distribution function. If H has a continuous density h, then the probability density f of u is given by

$$f(u) = \int_{-\infty}^{\infty} h(u,\theta)g(\theta)d\theta \tag{3.2}$$

In cases where θ changes discretely, then

$$f(u) = \sum_{n} h(u,\theta_n)p_n \tag{3.3}$$

where p_n, the probability of $\theta = \theta_n$ is such that $p_n \geq 0$ and $\sum_n p_n = 1$

In Equations (3.1), (3.2) and (3.3), the parameter θ is treated as a random variable and a new probability distribution is defined in the $U, \theta - plane$ which acts as our new sample space. Densities of the form (3.1) and distributions of the form (3.2) and (3.3) are generally referred to as mixtures. [see Feller, 1968]

3.1.1 Burr XII, Pareto and log-logistic mixed distributions

Suppose now, $\theta \sim gamma(a,c)$ and $h(u,\theta)$ is Weibull$(u;b,\theta)$. Then we have

$$\begin{aligned} f(u) &= \int_0^\infty \theta b u^{b-1} \exp^{-\theta u^b} \left(\frac{c^a}{\Gamma a} \theta^{a-1} \exp^{-c\theta} \right) d\theta \qquad (3.4) \\ &= \frac{c^a}{\Gamma a} b u^{b-1} \int_0^\infty \theta^a exp^{-\theta(c+u^b)} d\theta \\ &= \frac{c^a}{\Gamma a} b u^{b-1} \frac{\Gamma(a+1)}{(c+u^b)^{a+1}} \\ &= \frac{abc^a u^{b-1}}{(c+u^b)^{a+1}} \\ &= \frac{abu^{b-1}}{c\left(1+\frac{u^b}{c}\right)^{a+1}} \\ f(u) &= \frac{ab(\frac{u}{s})^b}{u\left[1+(\frac{u}{s})^b\right]^{a+1}} \qquad ; u > 0, \quad a,b,s > 0 \qquad (3.5) \end{aligned}$$

with $c = s^b$ and where a and b are shape parameters while s is the scale parameter. $f(u)$ is the Burr XII probability distribution function.

When $a = 1$, we have a special case of the Burr XII distribution called the log-logistic distribution. Similarly, when $b = 1$,we have a special case of the Burr XII distribution called the shifted Pareto distribution.

3.1.2 Burr XII specification of unemployment duration

In calculating the amount of benefits, one needs to estimate the duration of receipt of the claims. The best estimate of this would be the expected value of the unemployment duration. However, due to the incomplete spells of unemployment duration, it would be inappropriate to use a statistical average of the unemployment duration as the estimated duration of receipt of the claims.

In this regard, we will apply survival analysis on the unemployment duration to estimate the duration of receipt of the claims. Let $U \in (0,\infty)$ denote the duration of unemployment and $f(u) = P(U = u)$ denote the probability density function of U. The cumulative distribution function of U is thus expressed as

$$F(u) = Pr(U \geq u) = \int_0^u f(s) ds \qquad (3.6)$$

$F(u)$ is the probability that the unemployment duration is less than or equal to u. For an individual who has been unemployed for u periods, the conditional probability density that this individual will be reemployed back into the labor force at time u is defined by the hazard function $h(u)$ and can be written as:

$$h(u)du = Pr(u < U \geq u + du | U > u) = f(u)/s(u) \tag{3.7}$$

where $s(u)$ is the survival function defined as

$$s(u) = Pr(U > u) = \int_u^\infty f(s)ds \qquad ; s(0) = 1 \quad ; s(\infty) = 0 \tag{3.8}$$

We assume that unemployment duration follows Burr XII distribution when calculating the duration of benefit payment.

It can be shown that the survival and hazard functions corresponding to the Burr XII distribution are

$$s(u) = \left[1 + (\frac{u}{s})^b\right]^{-a} \tag{3.9}$$

$$h(u) = \frac{ab}{u\left[(\frac{s}{u})^b + 1\right]} \tag{3.10}$$

The duration of receipt of the unemployment benefits is

$$d = E(u) = \int_0^\infty uf(u)du \tag{3.11}$$

$$= \int_0^\infty \frac{ab(\frac{u}{s})^b}{u\left[1 + (\frac{u}{s})^b\right]^{a+1}} du$$

$$let \quad x = (\frac{u}{s})^b \Longrightarrow dx = \frac{b}{s}(\frac{u}{s})^{b-1}du$$

$$d = sc\int_0^\infty \frac{x^{\frac{1}{b}}}{(1+x)^{a+1}}dx$$

$$= asB(\frac{1}{b} + 1, a - \frac{1}{b}) \tag{3.12}$$

with the parameters determined from maximum likelihood estimation of the Burr XII distribution.

3.2 Maximum likelihood estimation of parameters

Given a random sample $X_1, X_2, \ldots, X_n$ from a distribution with one or more unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ restricted to a given sample parameter space Ω and probability density function $f(x_i; \theta_1, \theta_2, \ldots, \theta_m)$, then the joint probability density function

$$L(\theta_1, \theta_2, \ldots, \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_m) \tag{3.13}$$

is called the likelihood function which is evaluated at observed data points and is regarded as a function of only the θ_is. If $[g_1(x_1, x_2, \ldots, x_n), g_2(x_1, x_2, \ldots, x_n), \ldots, g_m(x_1, x_2, \ldots, x_n)]$ are the m ordered pairs that maximizes the likelihood function, then g_i s are called the maximum likelihood estimates of θ_is

Maximum likelihood estimation of parameters of Burr XII (u; a, b, s) distribution

Let $U_1, U_2, \ldots, U_n$ be a random sample of n independent and identically distributed Burr XII (u; a, b, s) random variables. The likelihood function is given by:

$$L(a, b, s) = \frac{a^n b^n}{s^{nb}} \big(\prod_{i=1}^{n} u_i\big)^{b-1} \left(\prod_{i=1}^{n} \left(1 + (\frac{u_i}{s})^b\right)\right) \qquad u > 0, i = 1, 2, \ldots, n \tag{3.14}$$

[see Okasha and Matter, 2015]

The log of the likelihood function is given by

$$\ln L(a, b, s) = n \ln a + n \ln b - nb \ln s + (b-1)\left(\sum_{i=1}^{n} \ln u_i\right) - (a+1)\left(\sum_{i=1}^{n} \ln \big(1 + \frac{u_i^b}{s^b}\big)\right) \tag{3.15}$$

Maximization of the log of the likelihood function yields

$$\frac{\partial \ln L(a, b, s)}{\partial a} = \frac{n}{a} - \sum_{i=1}^{n} \ln \left(1 + (\frac{u_i}{s})^b\right) = 0 \tag{3.16}$$

$$\frac{\partial \ln L(a, b, s)}{\partial b} = \frac{n}{b} - n \ln s + \sum_{i=1}^{n} \ln u_i - (a+1) \sum_{i=1}^{n} \left(\frac{\left(\frac{u_i}{s}\right)^b}{1 + \left(\frac{u_i}{s}\right)^b}\right) \ln \big(\frac{u_i}{s}\big) = 0 \tag{3.17}$$

$$\frac{\partial \ln L(a, b, s)}{\partial s} = \frac{-nb}{s} + \frac{b(a+1)}{s}\left(\sum_{i=1}^{n}\left(\frac{\left(\frac{u_i}{s}\right)^b}{1 + \left(\frac{u_i}{s}\right)^b}\right)\right) = 0 \tag{3.18}$$

Solving Equations (3.16), (3.17) and (3.18) together, we find that

$$\hat{a} = \frac{n}{\sum_{i=1}^{n} \ln \left(1 + \left(\frac{u_i}{\hat{s}}\right)^{\hat{b}}\right)} \tag{3.19}$$

and parameters s and b can be estimated using numerical methods such as Newton-Raphson. For Newton-Raphson, we make an initial guess, say b_0, to be the solution to Equation (3.17) (note that Equation (3.17) is differentiable) and improve the approximation with

$$b_1 = b_0 + \frac{f'(b_0)}{f''(b_0)} \tag{3.20}$$

The process is repeated iteratively using

$$b_{n+1} = b_n + \frac{f'(b_n)}{f''(b_n)} \tag{3.21}$$

until a sufficiently better value is reached according to the required level of accuracy. Newton-Raphson can also be applied in Equation (3.18) to estimate the value of parameter c.

3.3 Goodness of fit test statistics

Goodness of fit test statistics is used to test the fit of the data into various probability distributions to aid in the selection of the best fit distribution. These are methods that are used to infer whether a particular data set follows a specified statistical distribution or a class of distributions. Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed samples from an unknown distribution $G_n(x)$. If we wish to check whether at a specified level of significance this sample comes from a particular hypothesized distribution $G(x)$, we test the hypothesis:

H_o: Data follows a specified distribution

H_a: Data does not follow the specified distribution

This study applies three goodness of fit test statistics, that is, Kolmogorov Smirnov, Cramer-von Mises and Anderson-Darling tests.

Decisions on whether to reject the null hypothesis or not are based on either the p-value or the value of the test statistic. In all the above three tests, the null hypothesisis is rejected if the value of the test statistic is large. When comparing various probability fits, the distribution corresponding to the smallest value of the test statistic is the best fit in that case.

Kolmogorov Smirnov Test Statistic

This is a non-parametric two sided test statistic that compares the deviations of the empirical cdf of the data from the theoretical cdf. The highest value of the absolute deviations becomes the test statistic. That is

$$Teststatistic = sup|G_n(x) - G(x)| \tag{3.22}$$

where $G_n(x)$ and $G(x)$ are the empirical and theoretical cdfs respectively.

Cramer-von Mises Test Statistic and Anderson Darling Test statistic

Both Cramer-von mises test and Anderson darling are an extension of Kolmogorov test although the latter gives more weight to the tails than Cramer-von mises and Kolmogorov tests. The respective test statistics are given by

$$\text{Anderson Darling test statistic} = \int_{-\infty}^{\infty} \frac{[G_n(x) - G(x)]^2}{G(x)[1 - G(x)]} dG(x) \tag{3.23}$$

$$\text{Cramer-von test statistic} = \int_{-\infty}^{\infty} [G_n(x) - G(x)]^2 dG(x) \tag{3.24}$$

These two tests are more powerful than the Kolmogorov Smirnov test to certain deviations from the assumed distribution because they involve integration over the whole range of the data rather than considering a single case of the supremum.

3.3.1 Goodness of fit criterion

These are indices that supplement the goodness of fit statistics by aiding in choosing between competing models. The data analysis will incorporate both Aikake's and Bayesian information criterion.

Aikake's Information Criterion (AIC):

AIC balances between the complexity of the model and the statistical goodness of fit of the model by imposing a penalty for increasing the number of parameters in the model. It is defined as:

$$AIC = -2L(\hat{\theta}) + 2p \tag{3.25}$$

where $L(\hat{\theta})$ is the maximized log likelihood function and p is the number of parameters in the model.
The preferred model is the one corresponding to the lowest index.

Bayesian Information Criterion (BIC)

This is an improvement of the AIC in the sense that BIC factors in the size of the sample data in determining the amount of penalty to impose on a model due to increased number of parameters. It is defined as

$$BIC = -2L(\hat{\theta}) + 2p\ln(n) \tag{3.26}$$

where $L(\hat{\theta})$ and p are as defined above and n is the sample size of the data.
Just like in AIC, the preferred model is the one corresponding to the lowest index.

3.4 Pricing Framework

A fair premium is one that equates the expected present value of the benefits to that of the premium income according to the equivalence principle in insurance. To be able to do this, we need to determine the duration of benefits payment and that of receipt of the premiums. Another important factor to determine is the discount rate to use in discounting the future cash flows. We will for now assume that there are no miscellaneous incomes, expenses and taxes so that the present values of the premium income (P) and benefit payments (B) are as follows

$$PV(P) = \sum_{t=0}^{N} I_t(1+r_f)^{-t} \tag{3.27}$$

$$PV(B) = \sum_{t=0}^{N} B_t(1+E(r_b))^{-t} \tag{3.28}$$

where $t = 0, 1, \ldots, N$ indicates the time a cash flow is paid or received, P_t is the premium income at time t, B_t is the amount of benefits paid at time t, r_f is the risk free rate of return and r_b is the risk adjusted rate of return.

3.5 Capital Asset Pricing Model(C.A.P.M)

The capital asset pricing model is used to determine the theoretically appropriate required risk-adjusted rate of return. That is, r_b is such that

$$E(r_b) = r_f + (\overline{r}_m - r_f)\beta_u \tag{3.29}$$

where $\overline{r}_m$ is the expected return of the market portfolio and β_u, a measure of the correlation of unemployment rate and the market rate of return, is given by

$$\beta_u = \frac{cov(\text{unemployment, market})}{var(\text{market})} = \frac{cov(r_u, r_m)}{var(r_m)} \tag{3.30}$$

In this case r_u is the rate of unemployment.
The difference, $\overline{r}_m - r_f$, is referred to as the market premium. It shows the excess of the market return over the risk free rate. The market premium is directly proportional to the individual risk premium of the asset in question with beta being the constant of proportionality.
Both r_f and $\overline{r}_m$ are proxied from averages of historical returns of the risk free asset and the market respectively where the risk free rate is usually taken to be the return of a government security of the same term as the investment in consideration.

Chapter 4. Application and results

This chapter deals with fitting data into the models proposed in chapter three as well as determining the premium rate to charge for the scheme.

4.1 Data

The study uses secondary data from the United States of America retrieved from the Federal Reserve Bank of St. Louis and Robert Shriller's online data. It comprises of the mean monthly unemployment duration from January 1948 to February 2016 together with the number of first payments of unemployment insurance benefits and the corresponding covered jobs per month. This is as reported by the United States bureau of labor statistics but retrieved from Federal Reserve Bank of St. Louis. The unemployment duration in this case is the number of weeks a worker has been involuntarily unemployed. The other components of the data are the annual return on three-month treasury bills, retrieved from the Federal Reserve Bank of St. Louis; and Standards & Poors 500 Price Indices as reported by Robert Shiller in his online data.
Check References section for the data links.

4.2 Goodness of fit of the distributions on unemployment duration data

In order to determine the appropriate unemployment duration, we need to establish which distribution best fits the data.

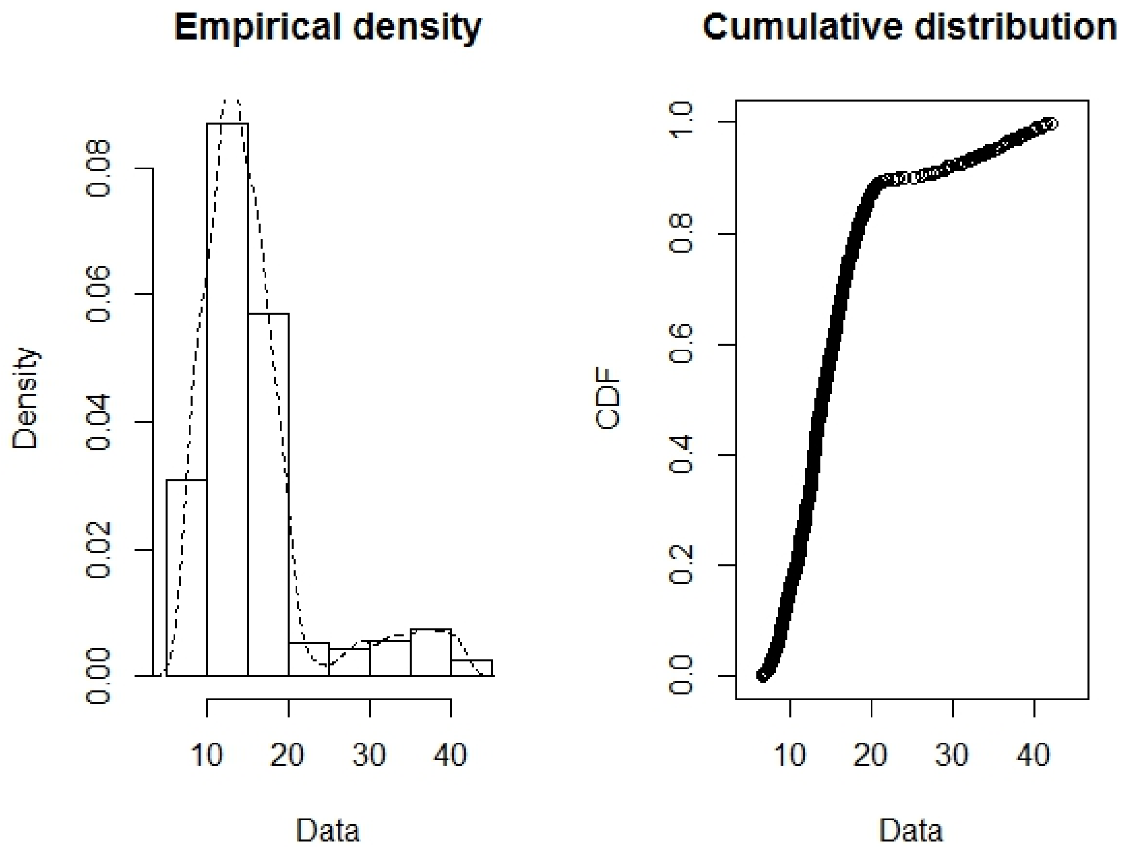

Figure 4.1: Empirical pdf and cdf

An empirical plot of both the density and distribution functions of the raw unemployment duration shown in Figure 4.1 indicates that the data follows one of the tailed distributions.

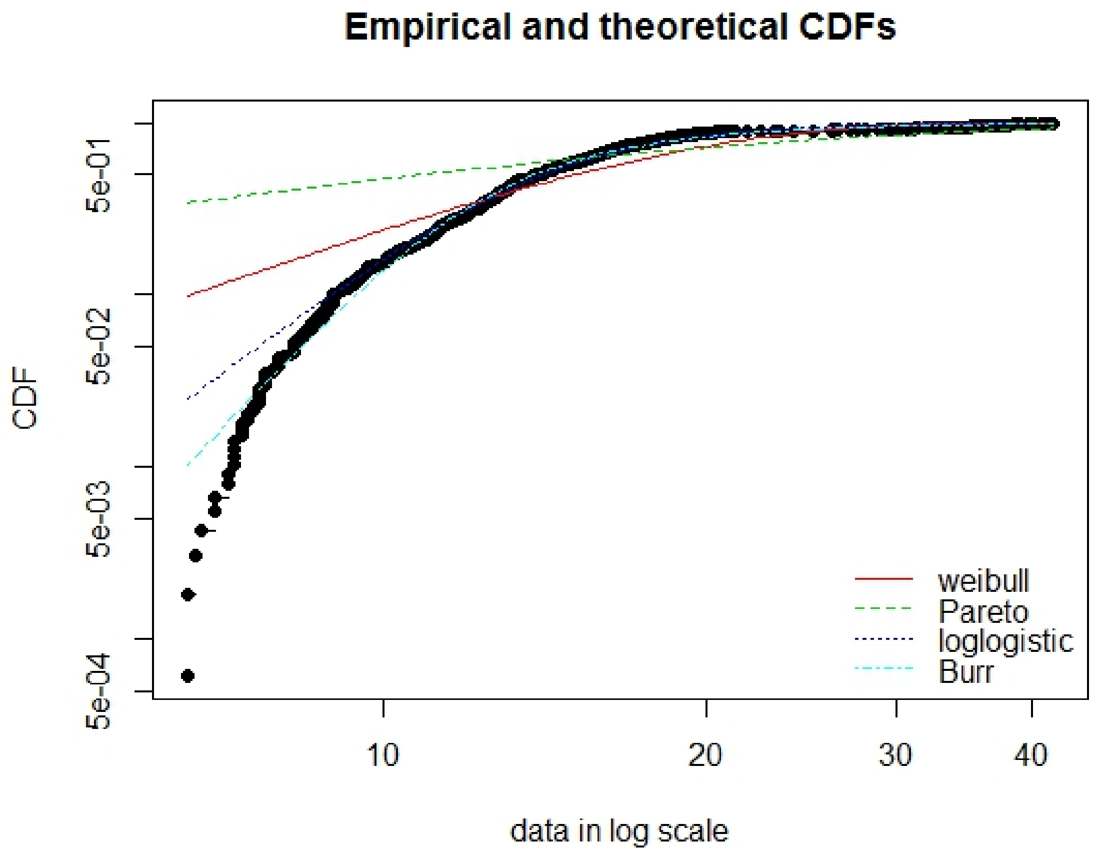

Figure 4.2: Empirical and theoretical cdfs

A plot of both empirical and theoretical cumulative distribution functions of the various chosen tailed distributions shown in Figure 4.2 indicates that both Burr distribution and loglogistic distribution can be used in modeling the duration of unemployment. The claim is further supported by comparing the quantile-quantile plots, Figure 5.1, and the probability plots, Figure 5.2, of these tailed distributions as shown in appendix2 .
This can be affirmed by observing the goodness of fit statistics of the tailed distributions below

Table 4.1: Goodness of fit statistics

	Weibull	loglogistic	Pareto	Burr
Kolmogorov-Smirnov statistic	0.1560211	0.05790557	0.3645346	0.04337078
Cramer-von Mises statistic	6.9333297	0.33442955	31.9466345	0.28938614
Anderson-Darling statistic	42.8112266	5.78689931	155.8114749	2.77486541

According to the results of the three statistical tests as outlined in Table 4.1, Burr emerges as the best candidate though competing with the loglogistic model.

Table 4.2: Goodness of fit criteria

	Weibull	loglogistic	Pareto	Burr
Akaike's Information Criterion	5436.162	5099.712	6158.474	5057.250
Bayesian Information Criterion	5445.581	5109.130	16167.893	5071.378

From the goodness of fit indices outlined in Table 4.2, Burr corresponds to the lowest indices under both AIC and BIC and is therefore the preferred model in estimating the duration of unemployment.

4.3 Assumptions of the model

1. No expenses.
2. The taxable wage base, S, will be constant.
3. The entry age into the labour force is 18 years and the retirement age is 65years in USA.
4. The risk-free rate is assumed to be the rate of return on investment of premiums.
5. A single unemployment spell for the insured.
6. Zero mortality during period of insurance coverage.

The above assumptions help us estimate the premium rate. They can however be relaxed as the model becomes more complex for better results.

4.4 Calculation of the premium rate

According to the equivalence principle in insurance, a fair premium is one that equates the expected present value of the benefits to that of the premium income. From section 3.1.2, evaluation of the duration of benefit receipt turns out to be the expected value of the

Burr distribution. Assuming that the policyholders pay premiums throughout their career life, including spells of unemployment, the mean present value of the premium income, MPV(P), is given by Equation (4.1). This will be achieved by using part of the benefits to pay the premium income. Therefore, for a state premium rate W of the monthly salary we have:

$$MPV(P) = W \times 12S \times \sum_{k=0}^{47}(1 + r_f)^{-k} \tag{4.1}$$

with S being the monthly taxable wage base of the insured.

We use the Benefit Event Valuation approach [see Bowers et al., 1997, Actuarial, 2014] to discount the contingent claims. The mean present value of a monthly benefit, MPV(B), of 45% of the taxable wage base per month, payable weekly during spells of unemployment is given by

$$MPV(B) = \sum_{k=0}^{2444}(1 + \overline{r}_b)^{\left(-\frac{k+m}{52}\right)} \times q_k^u \times \left\{\frac{(0.45 - W)}{4} \times S \times \sum_{t=0}^{d}(1 + \overline{r}_b)^{-\frac{t}{52}}\right\} \tag{4.2}$$

where k is the number of weeks since becoming unemployed involuntarily, m is the waiting period after applying for the unemployment benefits, q_k^u is the probability of a claim in week k, 52 is the number of weeks in a year and $\overline{r}_b$ is the expected risk-adjusted rate of return.

4.4.1 Parameter estimation

The parameters of the Burr distribution were estimated using Maximum likelihood estimation in R programming. The results are as outlined in table 4.3 below.

Table 4.3: Estimated parameters

Parameter	Estimate	Standard Error
Shape 1(a)	0.4955088	0.050208998
Shape 2 (b)	6.6921700	0.390146800
rate (1/s)	0.0853068	0.002159988

The mean unemployment duration (d) is equal to the expected value of the Burr distribution

$$d = 15.7598$$

r_f and $\overline{r}_m$ are average returns of treasury bills and the market respectively where individual entries of r_m are calculated as the annual rates of change of the $S\&P500$ index together with the associated dividends. From analysis we have:

$$\begin{aligned} r_f &= 4.21\% \\ \overline{r}_m &= 12.15\% \\ cov(r_m, r_u) &= -0.00013 \\ var(r_m) &= 0.025757 \end{aligned}$$

Replacing for the values of covariance of the market return and unemployment rate; and the variance of the market return in Equation (3.30) the Beta of unemployment becomes -0.01187. Values of $r_f, \overline{r}_m$ and the beta of unemployment yields a risk-adjusted rate of return $\overline{r}_b = 4.11\%$ using Equation (3.29). We assume that the waiting period $m = 2$weeks which is the average of what USA states take to process unemployment benefits. q_k^u is also estimated from the data and is assumed to be constant. It is the average proportion of successful claims (first payments) to the total number of covered jobs at the time of claim. From analysis, $q_k^u = 0.007071728$, for all $k = 1, 2, \ldots, 2444$ weeks

Equating the right sides of Equations (4.1) and (4.2) and replacing for the values of the variables outlined in this section yields a premium rate $W = 5.10\%$ of the taxable wage base per month.

Chapter 5. Conclusion and recommendation

5.1 Discussion and conclusions

From Equation (3.5) the scale parameter of the Weibull distribution was randomized by allowing it to follow the two parameter gamma distribution thereby accommodating heterogeneity in the data. This has led to a better fit of the unemployment data as confirmed from analysis; even Weibull provided a better fit than Pareto due to its varying hazard rate.
Observe that from Equation 3.10, $h(u)$ is a decreasing function of duration of unemployment u. This is an analogy to the normal expectation of decreasing probability of reemployment as the spell of unemployment lengthens.[see Salant, 1977]
As reported in section 4.4.1, both the beta and covariance of market return and unemployment rate are negative. This is expected because, when the market is doing well, that is, market rate of return is high; we would expect the rate of unemployment to be decreasing due to increased investments, hence a negative beta. This is a good strategy that States can adopt to control the amount of benefit payouts. Apart from reducing the number of retrenchment of workers, States can encourage new investments which in turn create more opportunities not only for new labor force but also for the retrenched workers. Reducing unemployment rate will also reduce the premium rate for the workers in the particular state. The premium rate W = 5.10% of the taxable wage base per month is within range of what most states demand under normal circumstances.

5.2 Recommendation

According to the study, States should adjust their premium contribution rates to approximately 5.10% of the taxable wage base to ensure solvency of their Unemployment Compensation Trust Fund.

From analysis, mixtures provide better fits to heterogeneous data than singular distributions. Specifically, the Burr XII distribution is recommended for modeling unemployment duration.

5.3 Limitations of study

The assumption of zero expenses and zero mortality among the covered employees is not the case in reality. The later increases the mean present value of the benefits income leading to a higher premium rate while the former reduces the mean present value of the outgo's from the Unemployment Compensation Trust Fund resulting to a lower premium rate. However, the assumption of zero expenses is in line with US Unemployment insurance program since administrative expenses for the program are paid for by the federal government.

The model assumed USA economy and therefore some adjustments should be made before applying it in another economy. Some of the factors that might require adjustment include retirement age and age of entry into the labour market among others.

The study was limited to only Burr XII distribution in modeling unemployment duration. Further research should be carried out to consider other mixture distributions that can also be used to model unemployment duration.

Acknowledgements

I thank African Mathematics Millennium Science Initiative (AMMSI) for the partial scholarship towards my masters and P. Weke, the Director School of Mathematics; University of Nairobi for recommending me for the award. Sincere regards also to my supervisors for the immense guidance during this research. Finally, I thank those who read and made several comments about the initial draft of this project.

References

Actuarial, Education Company (2014). *Pensions and other Benefits Specialist Technical (UK: Institute and Faculty of Actuaries)*. 14th ed. Institute and Faculty of Actuaries.

Beenstock, Michael (1985). “Competitive unemployment insurance pricing”. In: *Geneva Papers on Risk and Insurance*, pp. 23–31.

Blake, David and Michael Beenstock (1988). “The stochastic analysis of competitive unemployment insurance premiums”. In: *European Economic Review* 32.1, pp. 7–25.

Bowers, Newton L et al. (1997). *Actuarial Mathematics*. Illinois: The Society of Actuaries.

Bowers, Norman (1980). “Probing the issues of unemployment duration”. In: *Monthly Labor Review* 103.7, pp. 23–32.

Bronars, Stephen G. (1985). “Fair Pricing of Unemployment Insurance Premiums”. In: *Journal of Business*, pp. 27–47.

Chuang, Hwei-Lin and Min-Teh Yu (2010). “Pricing Unemployment Insurance–An Unemployment-Duration-Adjusted Approach”. In: *Astin Bulletin* 40.02, pp. 519–545.

Cummins, J. David (1991). “Statistical and financial models of insurance pricing and the insurance firm”. In: *The Journal of Risk and Insurance* 58.2, pp. 261–302.

Cummins, J David, Mary A Weiss, et al. (1991). “The structure, conduct, and regulation of the property-liability insurance industry”. In: *The financial condition and regulation of insurance companies*, pp. 117–154.

Delignette-Muller, Marie Laure and Christophe Dutang (2014). “fitdistrplus: An R Package for Fitting Distributions”. In: *J. Stat. Softw* 64, pp. 1–34.

Dutang, Christophe, Vincent Goulet, Mathieu Pigeon, et al. (2008). “actuar: An R package for actuarial science”. In: *Journal of Statistical Software* 25.7, pp. 1–37.

Feller, William (1968). *An Introduction to Probability Theory and Its Applications. Volume I*. John Wiley & Sons London-New York-Sydney-Toronto.

Malinvaud, Edmond (1985). “Unemployment insurance”. In: *Geneva Papers on Risk and Insurance*, pp. 6–22.

McDonald, James B. and Richard J. Butler (1987). “Some generalized mixture distributions with an application to unemployment duration”. In: *The Review of Economics and Statistics*, pp. 232–240.

Okasha, Mahmoud K and Mariam Y Matter (2015). “On the three-parameter Burr type XII distribution and its application to heavy tailed lifetime data”. In: *Journal: Journal of advances in Mathematics*.

Salant, Stephen W (1977). “Search theory and duration data: a theory of sorts”. In: *The Quarterly Journal of Economics*, pp. 39–57.

Stephen, P. D’Arc (1988). “Use of the CAPM to discount property-liability loss reserves”. In: *Journal of Risk and Insurance*, pp. 481–491.

Wang, Shaun S (2002). “A universal framework for pricing financial and insurance risks”. In: *Astin Bulletin* 32.02, pp. 213–234.

Data links

US. Bureau of Labor Statistics, Average (Mean) Duration of Unemployment [UEMPMEAN], retrieved from FRED, Federal Reserve Bank of St. Louis `https://research.stlouisfed.org/fred2/series/UEMPMEAN`, June 16, 2016.

Standards and Poors data: `www.econ.yale.edu/~shiller/data/ie_data.xls`

Board of Governors of the Federal Reserve System (US), 3-Month Treasury Constant Maturity Rate [DGS3MO], retrieved from FRED, Federal Reserve Bank of St. Louis `https://research.stlouisfed.org/fred2/series/DGS3MO`, June 16, 2016.

US. Employment and Training Administration, Initial Claims [ICNSA], retrieved from FRED, Federal Reserve Bank of St. Louis `https://research.stlouisfed.org/fred2/series/ICNSA`, June 16, 2016.

US. Employment and Training Administration, Insured Unemployment Rate [IURSA], retrieved from FRED, Federal Reserve Bank of St. Louis `https://research.stlouisfed.org/fred2/series/IURSA`, June 17, 2016.

Appendix

Appendix1 : R codes used in analysis

```
ud<- read.table("D:/study packs/project/Unemployment duration not seasonally adju
attach(ud)
plotdist(ud$VALUE, histo = TRUE, demp = TRUE)
descdist(ud$VALUE, boot = 1000)
ud1 <-ud$VALUE
ud.W <- fitdist(ud1, "weibull")
ud.P <- fitdist(ud1, "pareto", start = list(shape = 1, scale = 500))
ud.ll <- fitdist(ud1, "llogis", start = list(shape = 1, scale = 500))
ud.B <- fitdist(ud1, "burr", start = list(shape1 = 0.1, shape2 = 0.1, rate = 0.1)
cdfcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE, ylogscale = TRUE, legend
qqcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE, ylogscale = TRUE, legendt
ppcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE, ylogscale = TRUE, legendt
summary(ud.B)
gofstat(list(ud.W, ud.P, ud.ll, ud.B), fitnames = c("weibull", "Pareto", "llogis"

#PARAMETERS AND EXPECTED VALUE
a<-0.4955088; b<-6.6921700;
r<-0.0853068
d<-a*(1/r)*(factorial(1/b)*factorial(a-1-(1/b)))/factorial(a)
```

Appendix2 : QQ an PP plot

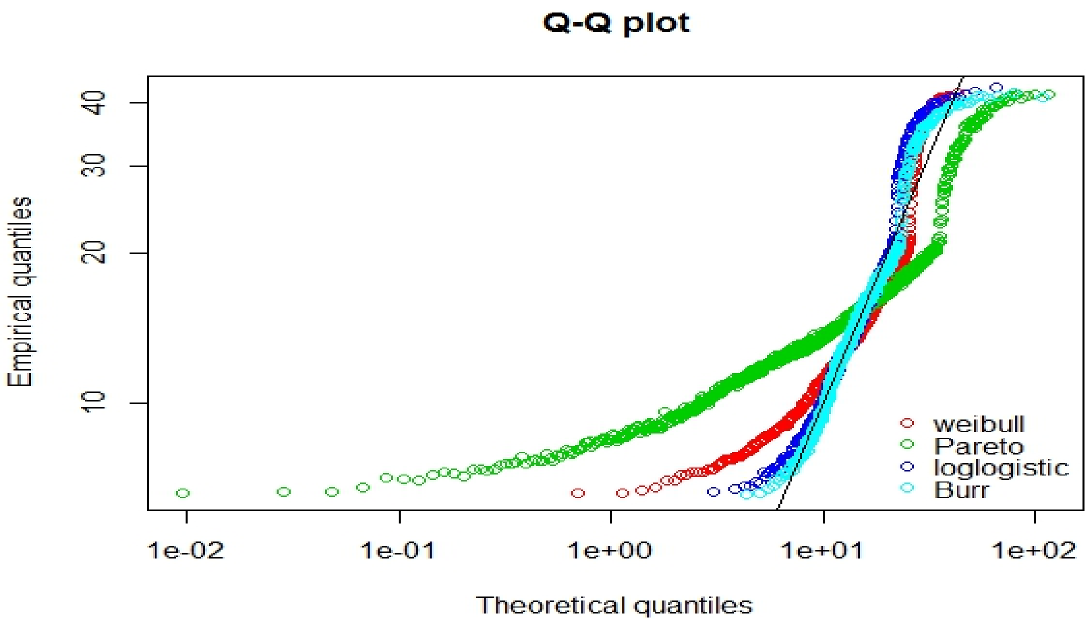

Figure 5.1: QQ-plot

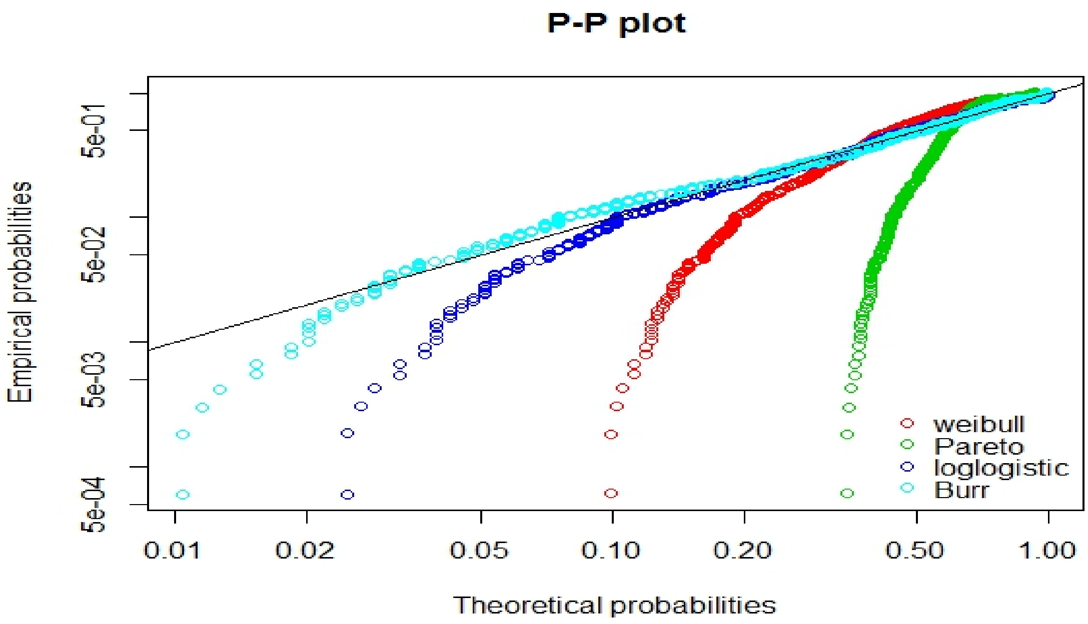

Figure 5.2: PP-plot