ACTUARIAL MODELING FOR INSURANCE CLAIM SEVERITY IN MOTOR COMPREHENSIVE POLICY USING INDUSTRIAL STATISTICAL DISTRIBUTIONS

BY

OYUGI MARGARET ACHIENG

Contact Details

Miss Margaret Achieng Oyugi
Graduate Trainee
Bosom Insurance Brokers Ltd, Kenya
P.O.BOX 74547-00200
NAIROBI,KENYA
EMAIL: margaretactuary@yahoo.com
MOB NO: +254 724 163 221

INTRODUCTION

General Insurance companies typically face two major problems when they want to use past or present claim amounts in forecasting future claim severity.

- Appropriate statistical distribution of claim
- Testing how well this statistical distribution fits their claims

OBJECTIVES

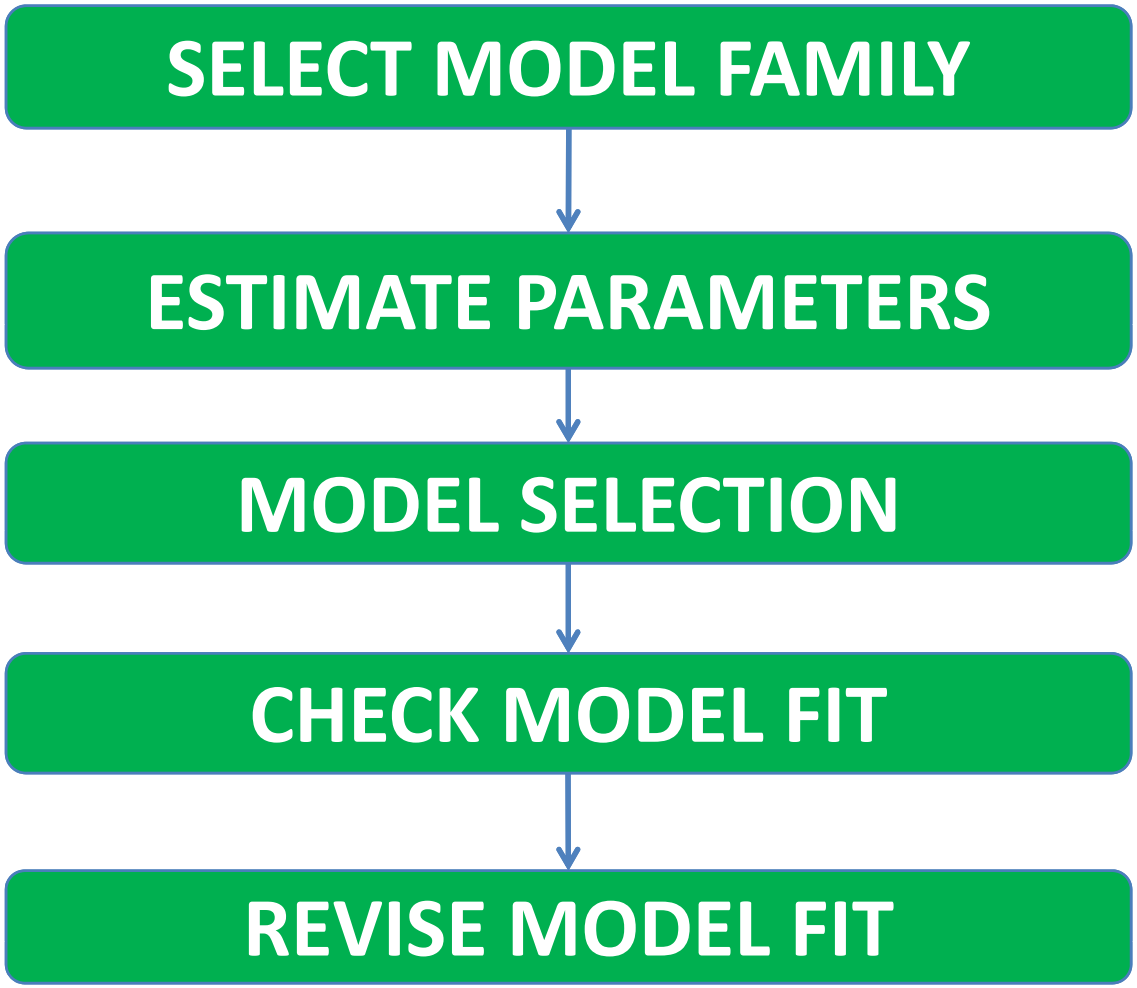
- Research aimed at:
  - Establishing one statistical distribution that could efficiently model the claim amounts
  - Testing goodness of fit as follows:
    - Mathematically (Akaike's Information Criterion - A.I.C)
    - Graphically (Quantile-Quantile Plots- Q-Q plots).

EXCECUTIVE SUMMARY

Steps followed in the modeling process:

- Selecting a model family
- Estimating model parameters
- Specification of the criteria to choose one model from the family of distributions
- Check model fit
- Revise model fit if necessary

STEPS IN MODELLING



VARIABLE

- The variable modelled was claim amounts from First Assurance Company limited, motor comprehensive policy (June 2006-June 2007).

- This modeling process was aided by the **MATLAB** software

SELECT MODEL FAMILY

Therefore four statistical distributions were used, these included:

- Gamma
- Exponential
- Lognormal
- weibull

MAX. LIKELIHOOD ESTIMATOR

- Let X_i be the ith claim amount, where $1 \leq i \leq n$.

  n = number of claims in the data set

  L = likelihood function

  θ is the parameter

  f(x) is the probability distribution function of a specific distribution

MAX. LIKELIHOOD ESTIMATOR

- The likelihood function is given by:

  $L = \prod f(X)$..... (1)

  To get maximum likelihood, differentiate equation (1)

  $M.L.E = \delta L / \delta \theta$..... (2)

  Therefore to solve the value of the parameter, equate equation (2) to zero:

  $\delta L / \delta \theta = 0$..... (3)

SELECTING ONE DISTRIBUTION

- The criteria for choosing one distribution out of the four was based on the values of the estimated maximum likelihood estimator

- The larger the Maximum likelihood estimator, the better the model

CHECKING MODEL FIT

- Just because a distribution got the highest likelihood out of the four distributions is not sufficient evidence to show that its the right distribution for the claims data

- Therefore to check for the goodness of fit of the distribution to the claims data Q-Q (Quantile-Quantile) plots were used

THE Q-Q PLOTS

- First Q - stands for the quantiles of the sampled data set
- Second Q stands for the quantile of the distribution being checked
- The Q-Q plot is a plot of the target population quantile(y) against the respective sample quantile (x)

AKAIKE'S INFORMATION CRITERION (A.I.C)

A.I.C criterion is defined by:

A.I.C = -2 (Maximum log-likelihood) +

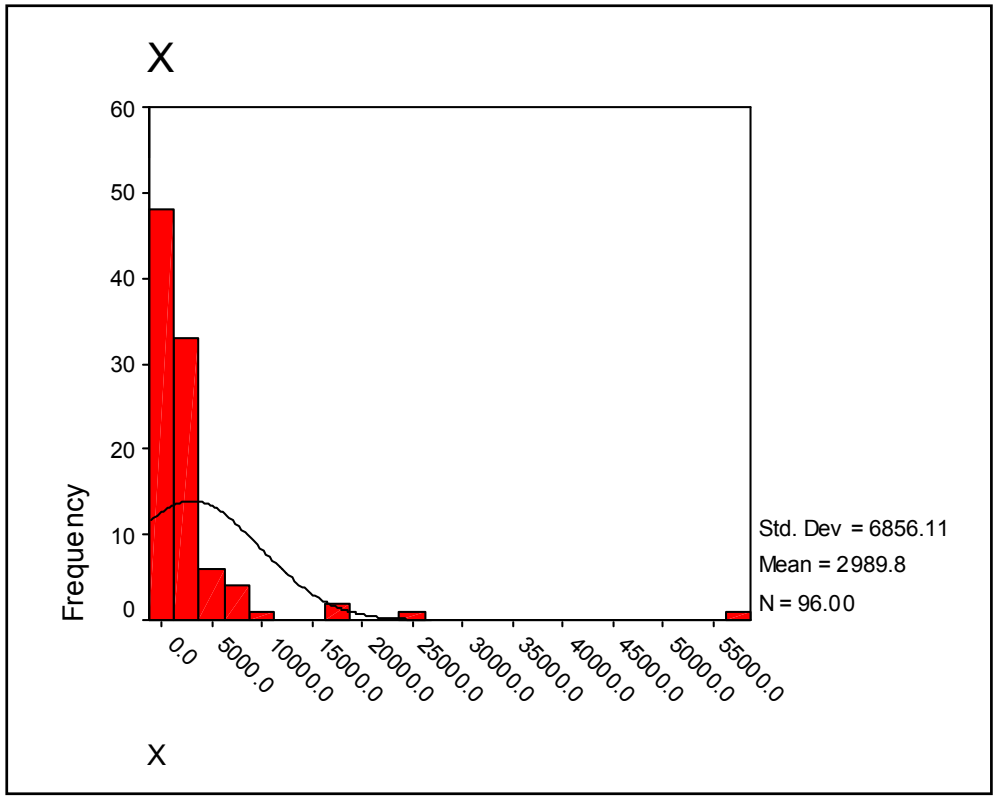
2 (No. of parameters estimated)

This was computed using a calculator

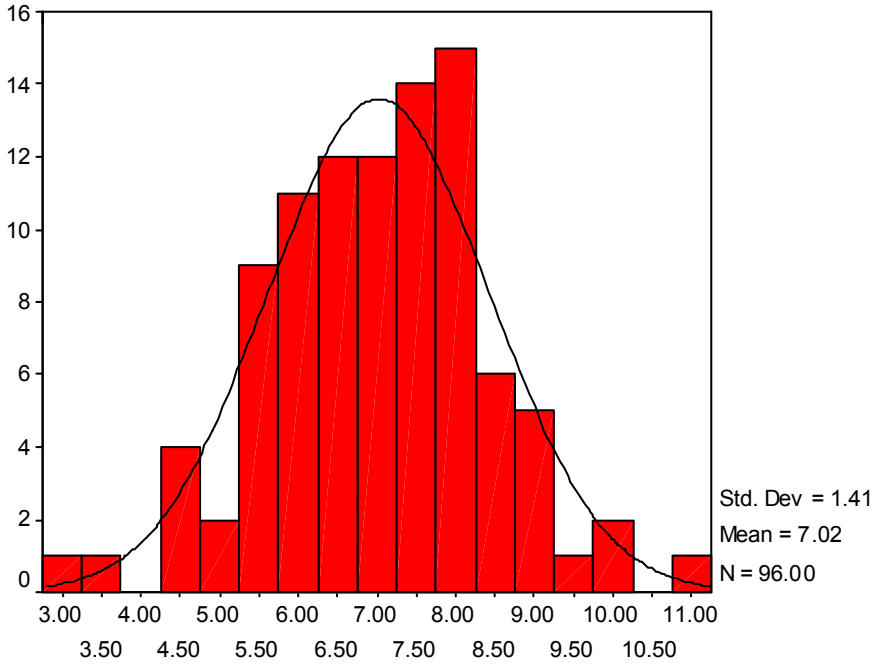
SUMMARY STATISTICS

N	Mean	Median	Mode	Skewness	Sum
125	148002.8	41978	11020	4.1	18500349

HISTOGRAMS OF CLAIMS



HISTOGRAM OF TRIPLE LOG



MAX LIKELIHOOD ESTIMATES

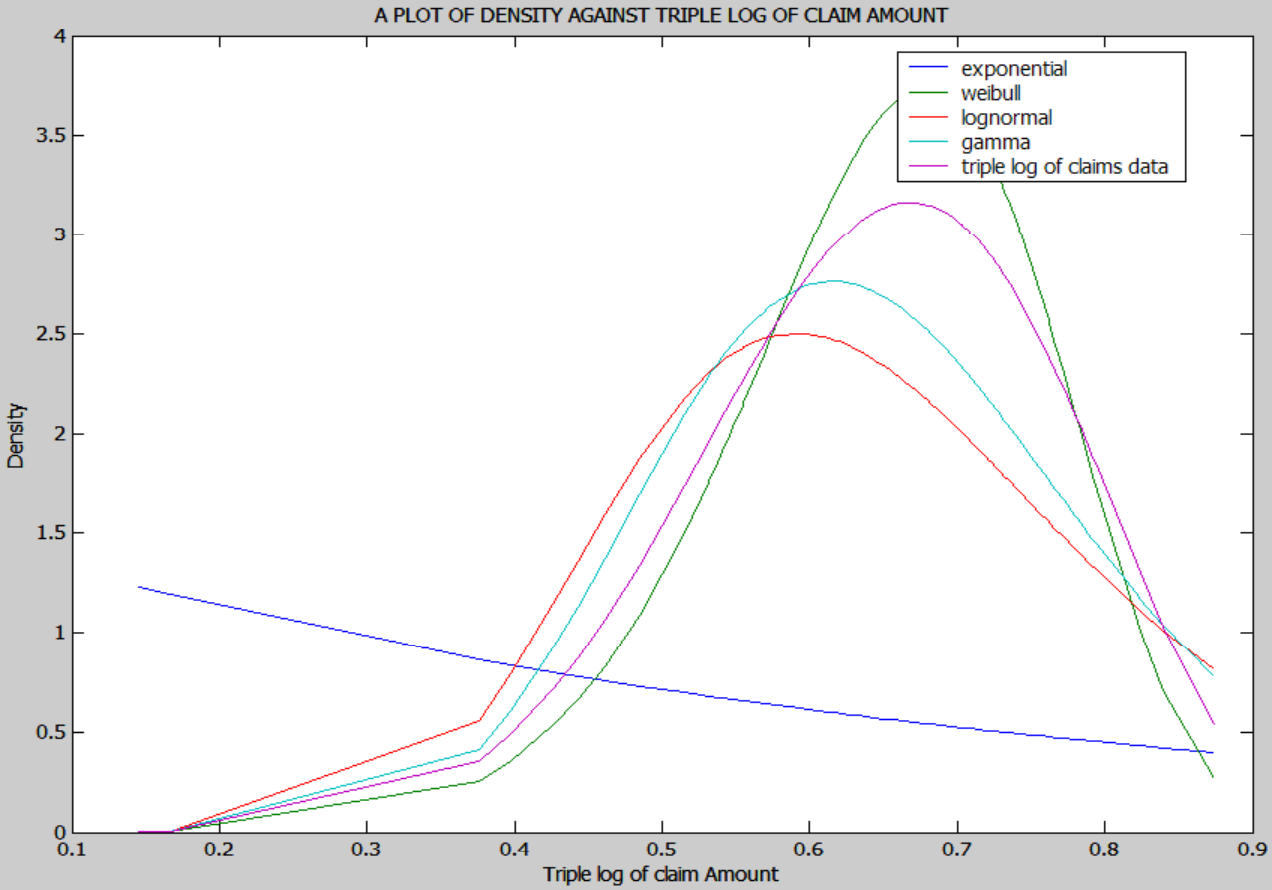
- The first step in fitting a model to a claims data is finding the parameter estimates of the particular statistical distribution.

- Table 4.2 below gives the parameters of the four distributions Where:
  - μ was taken to be the value of the first parameter
  - δ was taken to be the second parameter value

PARAMETERS

Distributions	Log-Likelihood	A.I.C	μ	δ
Exponential	-109.6749	221.3500	0.8609	
Lognormal	182.0632	-360.1300	-0.1517	0.0628
Gamma	-6460.9000	12925.8000	255.8749	0.0034
Weibull	174.8366	-345.6700	0.8865	16.6898

PROBABILITY DENSITY FUNCTION



THE GOODNESS OF FIT TEST

- The central problem in analysis is which model to use for making inferences from the data- this is the model selection problem.
- The goodness of fit test was done both mathematically and affirmed graphically.
- This called for computation of the Akaike's Information Criterion (A.I.C) and plotting of the Q-Q plots

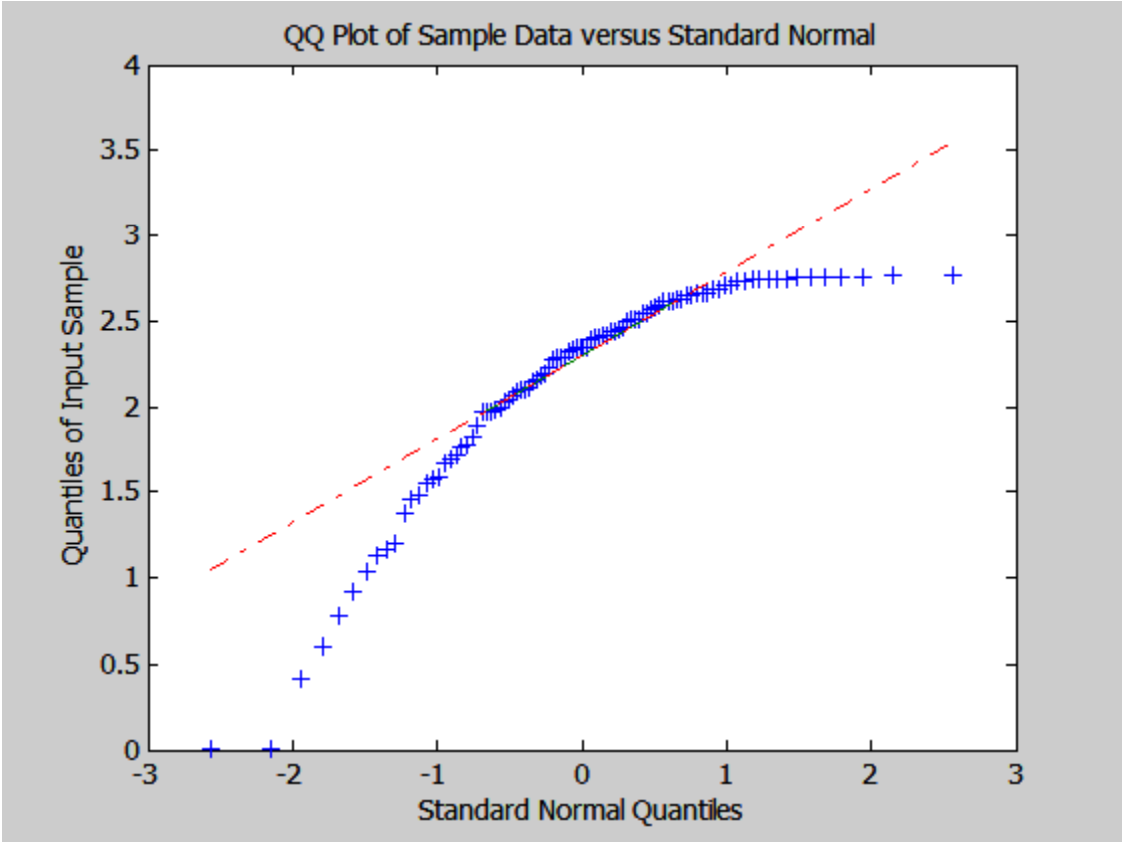
THE A.I.C INTERPRETATION

- In AIC one should select the model with the **smallest** value of AIC. With this regard the log-normal distribution had the smallest value of AIC.
- The log-normal distribution was estimated to be the closest to the unknown true distribution among the four candidate distributions considered

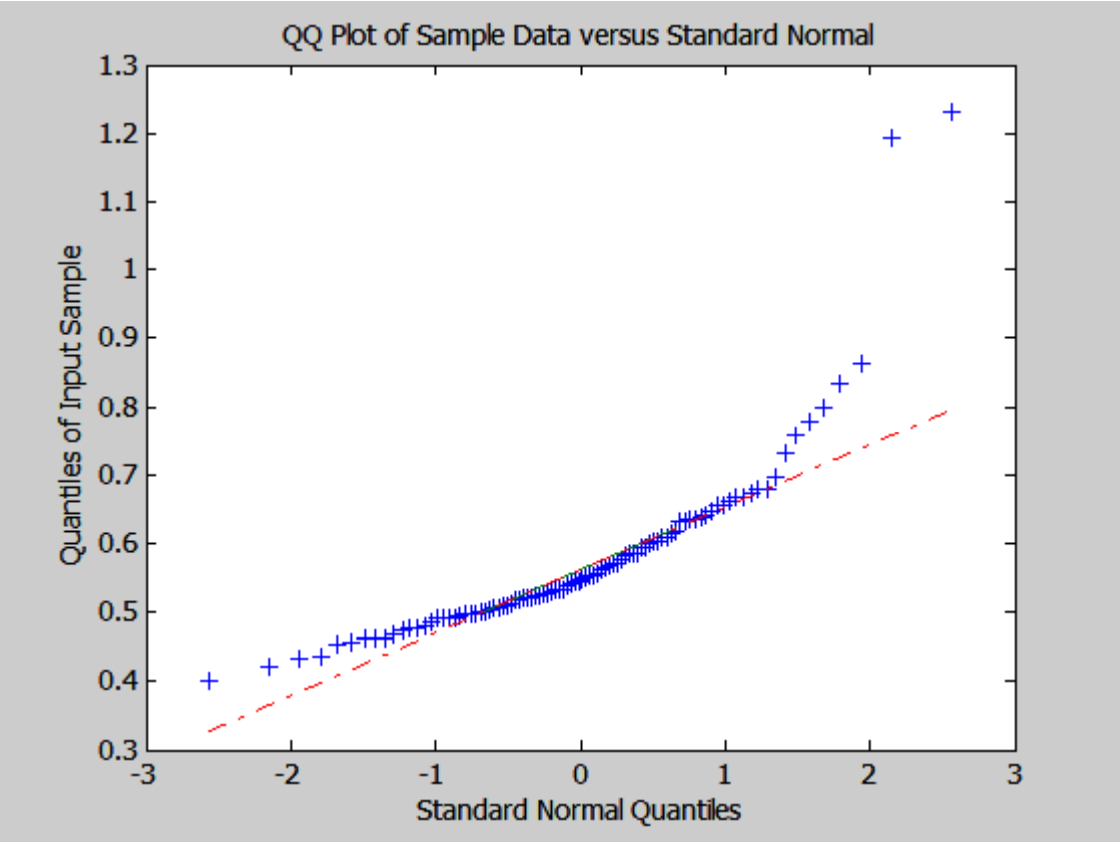
THE Q-Q PLOTS

- The Q-Q plots were constructed, and the selection criterion was based on the hypothesis below:
- **H0**: The statistical distribution provides the correct statistical model for the claims data
- **H1**: The statistical distribution does not provide the correct statistical model for the Claims data

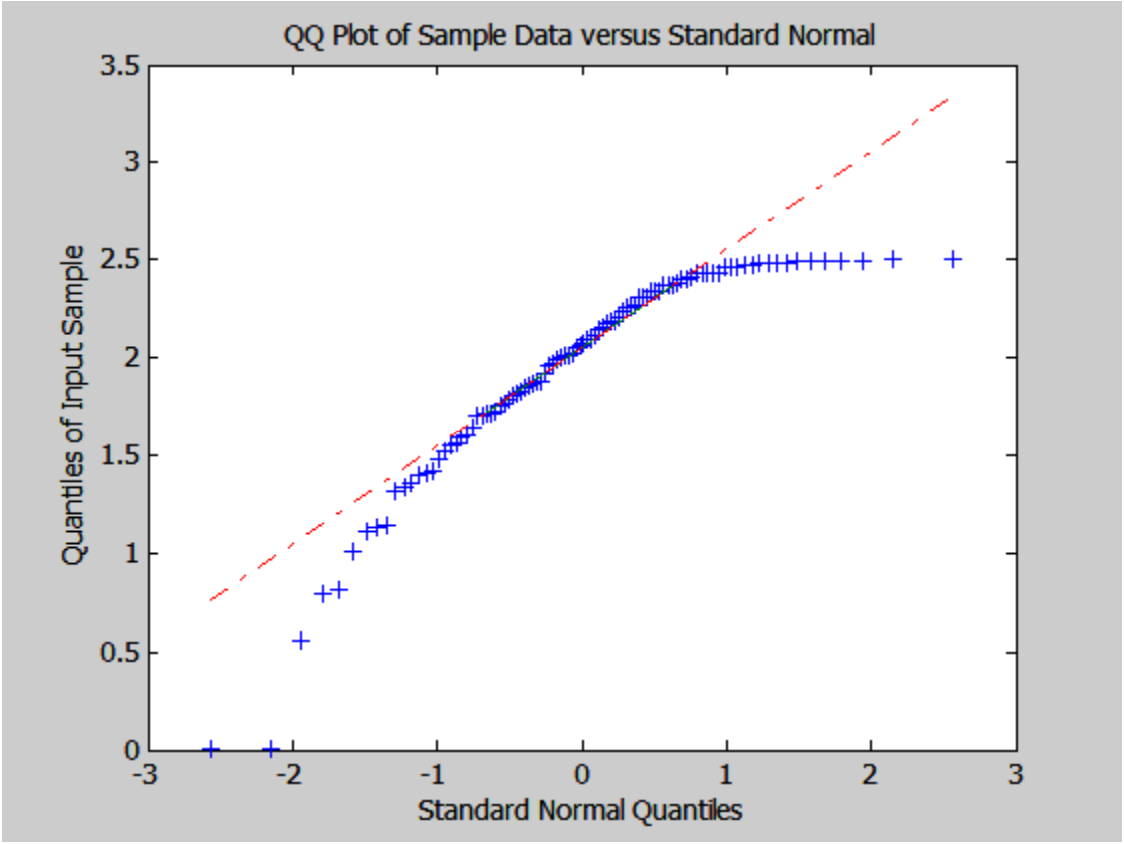
Q-Q PLOT OF GAMMA DISTRBN



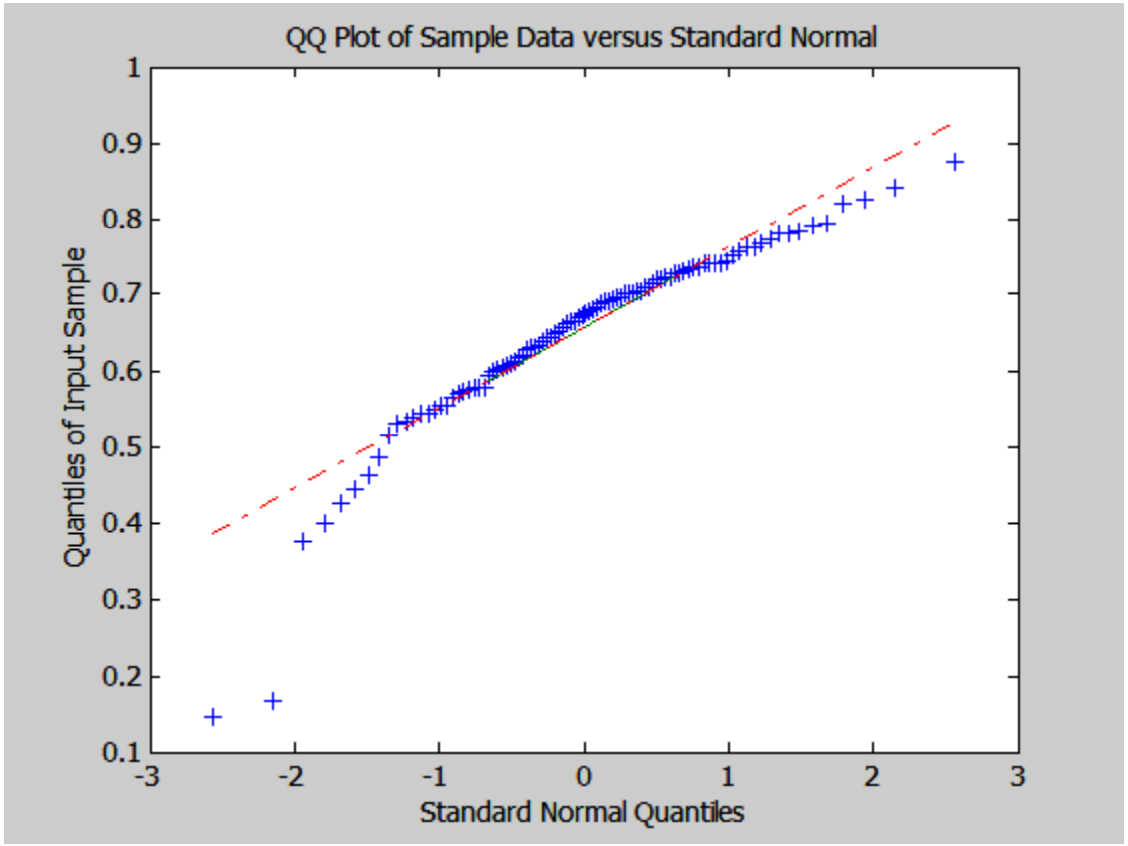
Q-Q PLOT OF EXPONENTIAL DISTRBN



Q-Q PLOT OF WEIBULL DISTRBN



Q-Q PLOT OF LOG-NORMAL DISTRBN



CONCLUSION OF THE Q-Q PLOTS

- The null hypothesis was not rejected for the lognormal distribution.
- Therefore a conclusion was made that at 99% level of confidence, the log-normal distribution provides the correct statistical model for the claims data.

CONCLUSION

- However the results of this study are dependent on a number of factors outside the modeling process.
- This means that the insurance company has to acknowledge these factors before using the results of this study in making future inferences. These factors include:

FACTORS TO CONSIDERED

I. Future claim amounts to be experienced by the insurance company may depend on an increase in **recklessness of drivers**

II. **Change of road traffic policies** may reduce future claims amounts. This maybe as a result of more traffic police-men being employed along the roads

FACTORS TO CONSIDER

III. **Increase in awareness** of the general public on the advantages of comprehensive insurance may lead to more claims amounts as more people would be moving from Third party (TPO) Insurance to Comprehensive insurance.

IV. **Increase of the number of motor vehicle** may alter the expected future claim amounts

LIMITATIONS

- However In adopting the actuarial modeling process displayed in this paper, the following are some of the likely limitation accompanying the results:

1. An assumption was made on the **absence of zero claim amounts** which in reality exists in comprehensive insurance policy.

LIMITATIONS

2. The choice of the **statistical distributions to be used was limited by the software** used in this research.

3. This research assumed that **all the claim amounts experiences between June 2006 and June 2007 were all reported**.

THE END

QUESTIONS??