Purpose:
Exposure to Risk (ETR) measurement of household insurance portfolio using Value- at- Risk (VaR) method

Keyword:
ETR, VaR, Negative Binomial, Weibull, Lognormal, Inverse Gaussian

Data set:
1. Theft.xls : Amounts (in $) of 120 theft claims made in a household insurance portfolio.

A. Fitting the data with various statistical distributions

The given data could potentially be fitted with the following distributions:

- Negative Binomial distribution

$$f(x|r,\beta) = \binom{x+r-1}{x}\left(\frac{1}{1+\beta}\right)^{r}\left(\frac{\beta}{1+\ \beta}\right)^{x} \quad x = 0,1,2\ \dots with\ r\ and\ \beta > 0$$

- Weibull distribution

$$f(x|\tau,\mu) = \left(\frac{\tau}{\mu}\right)\left(\frac{x}{\mu}\right)^{\tau-1} exp\left(-\left(\frac{x}{\mu}\right)^{\tau}\right) \quad with\ \mu > 0, \tau > 0\ and\ x > 0$$

- Lognormal distribution

$$f(x|\mu,\sigma) = \left(\frac{1}{x\sigma\sqrt{2\pi}}\right) exp\left(-\frac{1}{2}\left(\frac{logx-\mu}{\sigma}\right)^{2}\right) \quad with\ \mu\ \epsilon\ R, \sigma > 0\ and\ x > 0$$

- Inverse Gaussian distribution

$$f(x|\mu,\gamma) = \left(\frac{\gamma}{2\pi x^3}\right)^{\frac{1}{2}} exp\left(-\frac{\gamma(x-\mu)^2}{2\mu^2 x}\right) \quad with\ \ \gamma > 0\ , \mu > 0\ and\ x > 0$$

A. 1. Fitting the data using Method of Moments

Weibull, Lognormal and Inverse Gaussian distributions are fitted to the theft claims data using method of moments as followed:

1. Weibull distribution

$$E(X) = \lambda\Gamma\left(1+\frac{1}{\tau}\right)$$

$$Var(X) = \lambda^2\left[\Gamma\left(1+\frac{2}{\tau}\right) - \left(\Gamma\left(1+\frac{1}{\tau}\right)\right)^2\right]$$

$$E(X^2) = Var(X) + E(X)^2 = \lambda^2\Gamma\left(1+\frac{2}{\tau}\right)$$

$$\frac{E(X^2)}{E(X)^2} = \frac{\Gamma\left(1+\frac{1}{\tau}\right)}{\left[\Gamma\left(1+\frac{1}{\tau}\right)\right]^2} = \frac{\overline{X^2}}{(\bar{X})^2} \quad (1)$$

$$\frac{\overline{X^2}}{(\bar{X})^2} = \frac{19552940.1417}{2020.2917} = 4.7905$$

Using the Excel solver function on 1, we find $\hat{\tau}^{MoM} = 0.55389$.

$$\hat{\lambda}^{MoM} = \frac{E(X)}{\Gamma\left(1 + \frac{1}{\hat{\tau}^{MoM}}\right)} = \frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\hat{\tau}^{MoM}}\right)} = 1199.60565$$

Finding Method of Moment Estimates of Tao and Lambda						
		Tao_hat	0.55389	Lambda_hat	1199.60565	
Mean	2020.29167					
Mean Square (A)	19552940.1417	G(1+2t^-1)	13.5874			
Mean^2 (B)	4081578.4184	[G(1+t^-1)]^2	2.8363	A/B- Ratio (Objective)		
A/B	4.7905	Ratio	4.7905	0.0000000000	(Set to 0)	

2. Lognormal distribution

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right) = \bar{X}$$

$$E(X^2) = \exp(2\mu + 2\sigma^2) = \overline{X^2}$$

$$\mu + \frac{\sigma^2}{2} = \log(\bar{X})$$

$$2\mu + 2\sigma^2 = \log(\overline{X^2})$$

$$\widehat{\sigma^2}^{MoM} = \log\left(\overline{X^2}\right) - 2\log(\bar{X}) = 1.566642$$

$$\hat{\sigma}^{MoM} = \sqrt{\widehat{\sigma^2}^{MoM}} = 1.251656$$

$$\hat{\mu}^{MoM} = \log(\bar{X}) - \frac{\widehat{\sigma^2}^{MoM}}{2} = 6.8277$$

3. Inverse Gaussian distribution

$$E(X) = \hat{\mu}^{MoM} = \bar{X} = 2020.29167$$

$$Var(X) = \frac{(\hat{\mu}^{MoM})^3}{\hat{\lambda}^{MoM}} = \overline{X^2} - (\bar{X})^2$$

$$\hat{\lambda}^{MoM} = \frac{(\bar{X})^3}{\left(\overline{X^2} - (\bar{X})^2\right)} = 532.9834$$

A. 2. Fitting the data using Maximum Likelihood Estimator

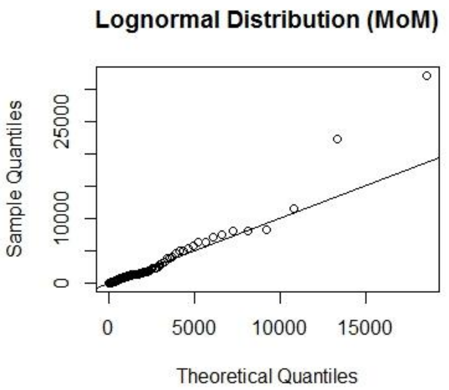
We then fit the claims data with 3 distributions specified in A using maximum likelihood estimator. Using R computations, the final outcome is as followed: [MLE.R]

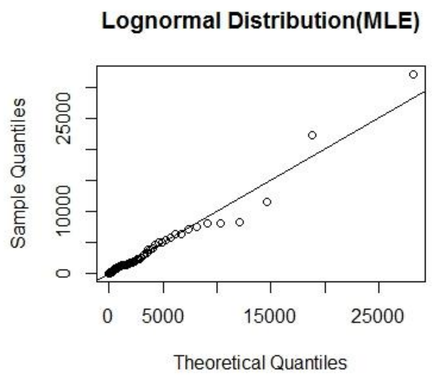
Distribution	Parameters	MLE	Log- likelihood Value
Weibull	τ	0.7155689	-1017.429
	λ	1560.387	
Lognormal	μ	6.624173	-1014.725
	σ	1.511246	
Inverse Gaussian	λ	163.9678	-1056.702
	μ	2020.035	

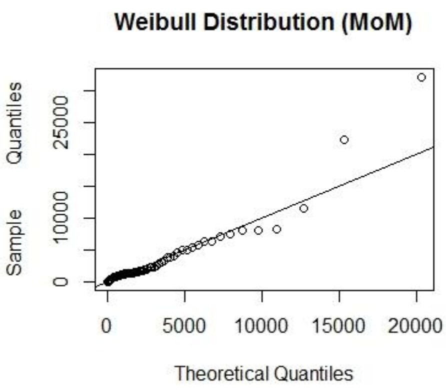
B. Fitness assessment

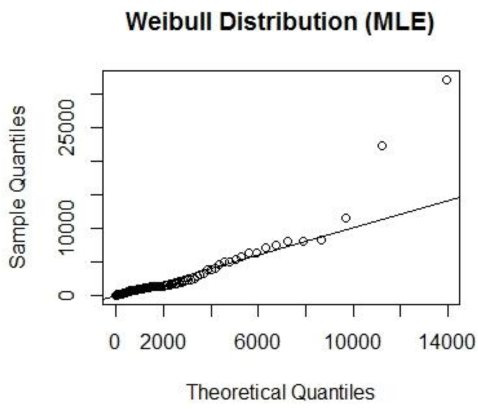
The accuracy of all fitted models of data is accessed using QQ plots

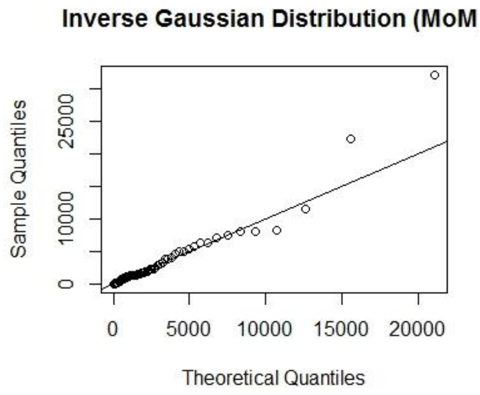
After finding the parameters associated with each distribution, QQ Plots are used to assess their accuracy of fit. [QQ Plot.R]

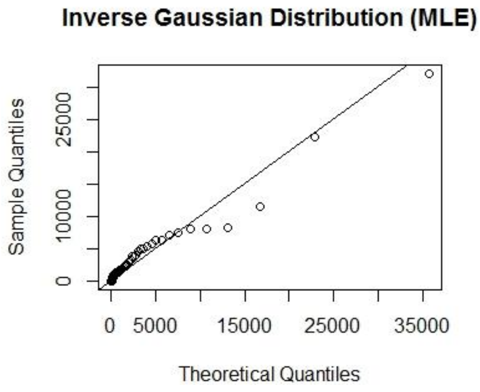












<u>C. Lognormal distribution provides the best fit</u>

Judging from the QQ plot and the log likelihood valus, lognormal distribution fitted the data best and it also gives the highest log likelihood value.

<u>D. VaR calculation</u>

The VaR values of these three loss models (fitted with method of moments and maximum likelihood method) are tabulated as followed: **[**VaR.R]

Distribution	p	VAR at (100p)%
Weibull	0.9	5407.385
(MoM)	0.95	8696.07
	0.99	18900.384
Lognormal	0.9	4590.578
(MoM)	0.95	7233.588
	0.99	16974.824
Inverse Gaussian	0.9	4899.667
(MoM)	0.95	8274.163
	0.99	19503.986
Weibull	0.9	5005.24
(MLE)	0.95	7230.032
	0.99	13185.943
Lognormal	0.9	5223.446
(MLE)	0.95	9044.856
	0.99	25332.842
Inverse Gaussian	0.9	4003.853
(MLE)	0.95	8822.737
	0.99	31885.545

Using the method of moment estimates, the Weibull model has the highest VaR at the p- level of 0.9 and 0.95, followed by the Inverse Gaussian and then the Lognormal. For the p- level of 0.99, Inverse Gaussian possess the highest VaR, then followed by Weibull and then Lognormal.

However, different results are produced if maximum likelihood is used instead. Using maximum likelihood estimates, the order of the VaR from the highest to the lowest at different levels is as followed:

90% VAR: Lognormal, Weibull, Inverse Gaussian
95% VAR: Lognormal, Inverse Gaussian, Weibull
99% VAR: Inverse Gaussian, Lognormal, Weibull

E. Weakness of VaR measurement

Despite the usefulness of VaR in calculating exposure to risk, it has some limitations which are outlined below:

1. Insensitivity beyond level p. p- level VaR does not tell us what happens beyond p.
2. VaR is not sub- additive. If a risk measure is sub- addictive, the addition of the risks of 2 portfolios considered separately should be greater than or equal to that of both portfolios considered together. The VaR measure fails to produce the properties of sub- additivity.
3. Fat- tailed distributions do not work well with VaR as they will lead to very low VaR numbers due to the underestimation of the probability of large moves.