Pak. J. Statist.
2022 Vol. 38(3), 315-340

INFINITE MIXTURE OF RAYLEIGH-RAYLEIGH DISTRIBUTION AND ITS APPLICATION TO MOTOR INSURANCE CLAIMS

Uraiwan Jaroengeratikun[1], **Sukanda Dankunprasert**[2] and **Tosaporn Talangtam**[2§]

[1] Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand. Email: uraiwan.j@sci.kmutnb.ac.th

[2] Faculty of Science, Khon Kaen University, Khon Kaen, Thailand. Email: sukanda_d@kkumail.com and tosata@kku.ac.th

[§] Corresponding author Email: tosata@kku.ac.th

ABSTRACT

In this paper, the Rayleigh-Rayleigh (RR) distribution is presented. The RR distribution is a new model which is constructed based on infinite mixture distribution and some properties are described. The model is applied to the motor insurance claims of 3 voluntary plans and compared to the current models that are derived from exponential and some traditional models. The maximum likelihood estimation (MLE) is the parameter estimation. The various measurements of model fitting are the Kolmogorov–Smirnov test (K–S test), the Anderson–Darling test (AD test), Akaike information criterion (AIC) and Bayesian information criterion (BIC). We have found that the RR distribution is more suitable for the data than some current distributions such as Weibull, Gamma, Exponential and Rayleigh.

KEYWORDS

AIC, BIC, Exponential-Exponential (EE) distribution, Exponentiated Exponential-Exponential (EEE) distribution, Gamma-Exponential distribution (GED) distribution, Generalized Pareto distribution (GPD).

1. INTRODUCTION

The Rayleigh and Exponential distributions are usually implemented for model construction. They are easily applied for model fitting, since there is only one parameter for its model. However, many modelers always modify or create new distributions based on them to be heavy-tailed or better fit than the traditional models can handle.

The Rayleigh distribution is applicable to many areas of science, such as noise theory, lethality and radar return. It is a good model for approximation in engineering practice. The proof of Rayleigh distribution and its generalizations are explained by Beckmann (1964). In 1967, Archer described some properties of the Rayleigh distribution random variables and their sums and products. Karim *et al*. (2011) presented Rayleigh mixture distributions based on weight functions. These are a mixture of Rayleigh distribution with sampling of Chi-square, *t* and *F* distributions. The method of moment is use for the estimation of parameters. Muhammad (2014) created a two parameter generalization Inverse Rayleigh distribution. The data sets, both simulation and real data, are better fitted by the model than the modified inverse Rayleigh, inverse Rayleigh and inverse Exponential distributions. The real data consist of 72 exceedances for the years 1958–1984 which correspond to the exceedances of flood peaks (in m3/s) of the Wheaton River near Carcross in Yukon Territory, Canada. In 2017, Akarawak *et al*. introduced the Gamma-Rayleigh distribution that is a new member of the Gamma-X family of generalized distributions. It is applied to the two sets of survival data which are secondary data on the time needed for patients to recover from typhoid fever and the time to drop out from an insurance policy. The distribution mostly produced fits and was competitive among the Gamma, Rayleigh, Weibull and Lognormal distributions with maximum likelihood estimation (MLE).

Exponential distribution has a fundamental role in describing the area of reliability theory and is commonly used to model the waiting time for how long one is required for a successful or failed event to occur. The parameter of its distribution is sometimes called the failure rate or intensity of the exponential. It is suitable for small sized data. There are many new distributions which are constructed from Exponential distribution. For example, Kareema *et al*. (2013) introduced Exponential Pareto Distribution for which some proofs are described. Mahmoud (2014) constructed and proved the Exponentiated Inverted Weibull Distribution. This model is applied to aggregated loss for reinsurance premium pricing. Zahida Perveen *et al*. (2016) presented the size-biased double Weighted Exponential distribution and found that it fits the ball bearing data records better than the size-biased Rayleigh and size-biased Maxwell distributions using the Anderson-Darling (AD) and Cramer-von Mises tests. Dankunprasert (2017) presented Gamma-Exponential distribution (GED) and Exponential-Exponential distribution (EED) constructed from an infinite mixture model for severity claims of motor insurance. Recently, Dankunprasert *et al*. (2021) explained the derivative of inverse Pareto distribution (IPD) and GED which are derived from different constructions but formed by the same distribution. Some properties and tail behavior of IPD are presented. This is a good competitor with the generalized Pareto distribution (GPD) for the modeling of tail distribution, by using the Kolmogorov–Smirnov test (K-S test), Anderson–Darling test (AD test), Akaike information criterion (AIC) and Bayesian information criterion (BIC), for both simulation data and Danish fire data.

In this paper, we are interested in a construction of the models using infinite mixture distributions. This new model is presented by Rayleigh-Rayleigh (RR) distribution and compared with the model derivative from Exponential and Rayleigh distributions consisting of GED and EED. We have shown that GED can be derived from Exponential-Exponential distribution based on exponentiated distribution in an alternative method for model construction. Thus, the GED is abbreviated as EEE distribution in this paper. Some properties of RR distribution and EED or EE distribution are described. Some traditional distributions of Exponential (Exp), Rayleigh (R), Lognormal (LN), Gamma (Gam), Weibull (Wei) and generalized Pareto distribution (GPD) are compared for model fitting. The models are applied to 3 data sets of motor insurance claims from a voluntary plan. The various measurements of model fitting are the K-S test, AD test, AIC and BIC. Their results are analyzed for model selection.

2. MODEL

2.1 Construction of New Models

2.1.1 Infinite Mixture Models

Suppose a random variable X is a conditional distribution given by λ. We denote its probability density function (pdf) by $f(x|\lambda)$. The continuous weighting function is treated as a pdf for λ, say $g(\lambda)$. Accordingly, the joint pdf is $f(x|\lambda)g(\lambda)$ and the compound pdf can be thought of as the marginal (unconditional) pdf of X,

$$h(x) = \int f_X(x|\lambda)g_\Lambda(\lambda)d\lambda.$$

where $f_X(x|\lambda)$ is the pdf of X with parameter λ and $g_\Lambda(\lambda)$ is the pdf of Λ. The distribution function can be determined from

$$\begin{aligned} H(x) &= \int_{-\infty}^{x}\int f_X(y|\lambda)g_\Lambda(\lambda)d\lambda\, dy \\ &= \int\int_{-\infty}^{x} f_X(y|\lambda)g_\Lambda(\lambda)\, dy\, d\lambda \\ &= \int F_X(x|\lambda)g\,(\lambda)d\lambda. \end{aligned}$$

where $F_X(x|\lambda)$ is the cumulative density function (cdf) of X with parameter λ and $g_\Lambda(\lambda)$ is the pdf of Λ.

2.1.2 Rayleigh-Rayleigh Distribution

Suppose a random variable X follows the Rayleigh distribution. Denote its cdf by $F(x|\sigma)$ where

$$F(x|\sigma) = 1 - \exp\left(-\frac{x^2\sigma^2}{2}\right);\sigma > 0, x \geq 0.$$

The Rayleigh distribution will be used as the mixing distribution. The pdf of the Rayleigh distribution is

$$g(\sigma) = \sigma t^2 \exp\left(-\frac{\sigma^2 t^2}{2}\right);\ t > 0, \sigma \geq 0.$$

The unconditional cdf of X is

$$\begin{aligned} F(x) &= \int_0^\infty \left[\left\{1 - \exp\left(-\frac{x^2\sigma^2}{2}\right)\right\}\cdot\left\{\sigma t^2 \exp\left(-\frac{\sigma^2 t^2}{2}\right)\right\}\right] d\sigma \\ &= \int_0^\infty \sigma t^2 \exp\left(-\frac{\sigma^2 t^2}{2}\right) d\sigma - \int_0^\infty \sigma t^2 \exp\left(-\frac{x^2\sigma^2}{2} - \frac{\sigma^2 t^2}{2}\right) d\sigma \\ &= \frac{x^2}{x^2 + t^2} \end{aligned}$$

Hence,

$$F(x) = \frac{x^2}{x^2 + t^2}; t > 0, x \geq 0.$$

The pdf is given by

$$f(x) = \frac{2xt^2}{(x^2 + t^2)^2}; t > 0, x \geq 0.$$

2.2 Exponentiated Exponential-Exponential Distribution

The Exponentiated Exponential-Exponential (EEE) distribution is the same distribution as the Gamma-Exponential distribution (GED) and the inverse Pareto distribution (IPD), but it was derived from a different method. Dankunprasert, S. (2017) constructed the GED for severity claims of motor insurance by using infinite mixture distribution. In this paper, the EEE distribution is derived from the Exponentiated model by using Exponential-Exponential (EE) distribution. The Exponentiated distribution is proposed by Gupta *et al.* (1998) for time failure data. Its model is specified as below.

Given a random variable X with the baseline distribution function of $F(x)$, the class of Exponentiated distribution is defined as

$$G_\alpha(x) = [F(x)]^\alpha.$$

where α is a positive real number. It has been called the Lehman alternative where α is a positive integer.

The EEE distribution is built from EE distribution by the Exponentiated model which is in the following form:

Suppose a random variable X follows an Exponential-Exponential distribution. Denote its cdf by $F(x|b)$ where

$$F(x|b) = 1 - \frac{b}{x + b}; b > 0, x \geq 0.$$

The cdf of the EEE distribution is easily described as

$$G_\alpha(x) = \left[\frac{x}{x + b}\right]^\alpha; \; b > 0, \alpha > 0, x \geq 0.$$

The pdf is given by

$$g_\alpha(x) = \frac{b\alpha x^{\alpha-1}}{(x + b)^{\alpha+1}}; \; b > 0, \alpha > 0, x \geq 0.$$

3. PROPERTIES

This section presents some properties of the distributions of Rayleigh-Rayleigh (RR), Exponential-Exponential (EE) and Exponentiated Exponential-Exponential (EEE) such as survival, hazard functions, value-at-risk (VaR), expected value and limited expected value. Some proof of the property for heavy tailed distributions are also explained.

3.1 Rayleigh-Rayleigh Distribution

Survival function

$$S(x) = \frac{t^2}{x^2 + t^2}.$$

Hazard function

$$h(x) = \frac{2x}{x^2 + t^2}.$$

VaR

$$\pi_p = \sqrt{\frac{t^2}{p^{-1} - 1}}.$$

Expected value

$$E[X] = \frac{\pi t}{2}.$$

Since $E[X^2]$ cannot be found because of upper limited of integrate, we will consider the limit of expectation in second order to be the variance. The limited loss random variable $X \wedge u$ is defined as

$$X \wedge u = \begin{cases} X & ; \quad X < u \\ u & ; \quad X \geq u \end{cases}$$

The kth limited expectation can be written as

$$E[(X \wedge u)^k] = \int_{-\infty}^{u} x^k f(x)dx + u^k[1 - F(u)]$$

Therefore, the second order of limited expectation of X on Rayleigh-Rayleigh distribution is in the form of

$$E[(X \wedge u)^2] = \int_{0}^{u} x^2 f(x)dx + u^2[1 - F(u)]$$

$$= \int_{0}^{u} x^2 \left(\frac{2xt^2}{(x^2 + t^2)^2}\right) dx + u^2 \left[1 - \frac{x^2}{x^2 + t^2}\right]$$

$$E[(X \wedge u)^2] = t^2 \left(\frac{(u^2 + t^2)\ln(x^2 + t^2) - u^2(2\ln(t) + 1) - 2t^2\ln(t) + u^2}{x^2 + t^2}\right).$$

Theorem 1:

Let X be a random variable. A heavy tailed distribution has a tail that is heavier than an Exponential distribution. Then RR has a heavy tail.

Proof:

Consider the survival function of RR, $S_{RR}(x)$, compared with survival function of Exp, $S_{Exp}(x)$. Then, the limit of the ratio will be the same, as can be seen by an application of L'Hôpital's rule:

$$\lim_{x\to\infty}\frac{S_{RR}(x)}{S_{Exp}(x)} = \lim_{x\to\infty}\frac{S'_{RR}(x)}{S'_{Exp}(x)} = \lim_{x\to\infty}\frac{-f_{RR}(x)}{-f_{Exp}(x)} = \lim_{x\to\infty}\frac{f_{RR}(x)}{f_{Exp}(x)}$$

$$= \lim_{x\to\infty}\frac{\frac{2xt^2}{(x^2+t^2)^2}}{\lambda\exp(-\lambda x)} = \lim_{x\to\infty}\frac{2xt^2\exp(\lambda x)}{\lambda(x^2+t^2)^2}.$$

Since exponential goes to infinity faster than polynomials, the limit is infinity. So, the RR has a heavier tail than the Exponential. Therefore, the RR is a heavy tailed distribution. □

3.2 Exponential-Exponential Distribution

Since this model has not been described with any references to its properties, we explain some of the properties in this section.

Survival function

$$S(x) = \frac{b}{x+b}.$$

Hazard function

$$h(x) = \frac{1}{x+b}.$$

VaR

$$\pi_p = b\left[\frac{p}{1-p}\right].$$

Expected value

$$E[X] = \int_{-\infty}^{\infty} x\cdot f(x)dx$$

$$= \int_{0}^{\infty} x\cdot\frac{b}{(x+b)^2}\,dx$$

$$= \int_{b}^{\infty} b\cdot\frac{(u-b)}{u^2}\,du \qquad ; u = x+b \text{ and } du = dx$$

$$= b\left[\lim_{a\to\infty}\ln a - \ln b - 1\right].$$

Since we cannot find the expected value because of the upper limit of integration, we will consider the limit of expectation for the first and second order to become the expected value and variance.

The first order of limited expectation of X

$$E[(X \wedge u)] = \int_{-\infty}^{u} xf(x)dx + u[1 - F(u)]$$

$$= \int_{0}^{u} x \cdot \frac{b}{(x+b)^2} dx + u\left[1 - \left(1 - \frac{b}{u+b}\right)\right]$$

$$= b \int_{b}^{u+b} \frac{u-b}{u^2} du + \frac{ub}{u+b}; where\ u = x + b\ and\ du = dx$$

$$= b \ln\left(\frac{u+b}{b}\right).$$

The second order of limited expectation of X

$$E[(X \wedge u)^2] = \int_{-\infty}^{u} x^2 f(x)dx + u^2[1 - F(u)]$$

$$= \int_{0}^{u} x^2 \cdot \frac{b}{(x+b)^2} dx + u^2\left[1 - \left(1 - \frac{b}{u+b}\right)\right]$$

$$= b \int_{b}^{u+b} \frac{u^2 - 2ub + b^2}{u^2} du + \frac{u^2 b}{u+b}; where\ u = x + b\ and\ du = dx$$

$$= b\left[\int_{b}^{u+b} 1\ du - \int_{b}^{u+b} \frac{2b}{u}\ du + \int_{b}^{u+b} \frac{b^2}{u^2}\ du\right] + \frac{u^2 b}{u+b}$$

$$= \frac{2b}{u+b}\left[u(u+b) - b(u+b)\, ln\left(\frac{u+b}{b}\right)\right].$$

Theorem 2:

Let X be a random variable. The probability distribution function is an Exponential-Exponential (EE) distribution such that

$$f(x) = \frac{b}{(x+b)^2}; \quad b > 0, x \geq 0.$$

Then EE has a heavy tail.

Proof:

The hazard rate function for the EE is

$$h(x) = \frac{f(x)}{S(x)} = \frac{\frac{b}{(x+b)^2}}{\frac{b}{(x+b)}} = \frac{b}{(x+b)^2} \cdot \frac{(x+b)}{b} = \frac{1}{(x+b)}.$$

Since

$$h'(x) = -\frac{1}{(x+b)^2} < 0.$$

Using classification based on the hazard rate function, then $h(x)$ is decreasing. EE with decreasing hazard rate function has a heavy tail. $\square$

3.3 Exponentiated Exponential-Exponential Distribution

Dankunprasert *et al.* (2021) described some properties of GED or Exponentiated Exponential-Exponential (EEE) distribution without limited expectation. Thus, we present the limited expectation for the first and second order of EEE distribution in addition.

Survival function

$$S(x) = 1 - \left(\frac{x}{x+b}\right)^{\alpha}.$$

Hazard function

$$h(x) = \frac{b\alpha x^{\alpha-1}}{(x+b)[(x+b)^{\alpha} - x^{\alpha}]}.$$

VaR

$$\pi_p = \frac{b}{p^{-1/\alpha} - 1}.$$

The first order of limited expectation of X

$$\begin{aligned}E[(X \wedge u)] &= \int_{-\infty}^{u} x f(x) dx + u[1 - F(u)] \\ &= \int_{0}^{u} x \cdot \frac{\alpha b x^{\alpha-1}}{(x+b)^{\alpha+1}} dx + u\left[1 - \left(\frac{u}{u+b}\right)^{\alpha}\right] \\ &= \alpha b \sum_{n=0}^{-1-\alpha} \binom{-1-\alpha}{n} b^n \left[-\frac{u^{-n}}{n}\right] + u\left[1 - \left(\frac{u}{u+b}\right)^{\alpha}\right].\end{aligned}$$

The second order of limited expectation of X

$$\begin{aligned}E[(X \wedge u)^2] &= \int_{-\infty}^{u} x^2 f(x) dx + u^2[1 - F(u)] \\ &= \int_{0}^{u} x^2 \cdot \frac{\alpha b x^{\alpha-1}}{(x+b)^{\alpha+1}} dx + u^2\left[1 - \left(\frac{u}{u+b}\right)^{\alpha}\right] \\ &= \alpha b \int_{0}^{u} \sum_{n=0}^{-1-\alpha} \binom{-1-\alpha}{n} b^n x^{-n}\, dx + u^2\left[1 - \left(\frac{u}{u+b}\right)^{\alpha}\right]\end{aligned}$$

$$E[(X \wedge u)^2] = \alpha b \sum_{n=0}^{-1-\alpha} b^n \left[\frac{u^{1-n}}{n-1}\right] \binom{-1-\alpha}{n} + u^2\left[1 - \left(\frac{u}{u+b}\right)^{\alpha}\right].$$

4. MATERIAL AND METHODS

The following are the models, parameter estimation, measurements of model fitting and some descriptive statistics of data sets.

4.1 The Models

The loss distributions are selected for comparison which are composed of the distributions Rayleigh-Rayleigh (RR), Exponential-Exponential (EE), Exponentiated Exponential-Exponential (EEE), Exponential (Exp), Rayleigh (R), Lognormal (LN), Gamma (Gam), Weibull (Wei) and generalized Pareto distribution (GPD).

4.2 Parameter Estimation and Measurement of Model Fitting

4.2.1 Parameter Estimation

The method of maximum likelihood estimation (MLE) provides an estimator of RR by means of the following procedure:

The pdf of infinite mixture Rayleigh-Rayleigh distribution

$$f(x) = \frac{2xt^2}{(x^2+t^2)^2} \;\; ; \; t > 0, x \geq 0.$$

A random sample $x_1, x_2, ..., x_n$ are n independent observation on a random variable X.

The likelihood function can be written as follows:

$$L = \prod_{i=1}^{n} \frac{2x_i t^2}{({x_i}^2+t^2)^2}$$

and the natural log-likelihood function is in the form

$$\log L = \sum_{i=1}^{n} \log\left\{\frac{2x_i t^2}{({x_i}^2+t^2)^2}\right\}.$$

Taking the partial derivatives of the natural log-likelihood function with respect to the parameters is as follows:

$$\frac{d}{dt} log\, L(t) = \frac{2n}{t} - 4t\sum_{i=1}^{n} \frac{1}{(x_i^2+t^2)}.$$

We estimate $\hat{t}$ for t by $\frac{d}{dt}\log L(t) = 0$.

Thus,

$$\frac{2n}{t} - 4t\sum_{i=1}^{n} \frac{1}{(x_i^2+t^2)} = 0.$$

Solves the equations numerically for the estimated parameters by the fixed point iteration method.

4.2.2 Measurement of Model Fitting

There are four measurements for model fitting and their formulae are the following:

4.2.2.1 The Kolmogorov-Smirnov Test

The K-S test statistic is defined by

$$D = \sup_{x} |F_n(x) - F_X^*(x)|.$$

where $F_n(x)$ is empirical cdf of X with sample size n and $F_X^*(x)$ is the theoretical cumulative distribution of the distribution being tested.

4.2.2.2 Anderson-Darling Test (AD Test)

Anderson-Darling test

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\log F(x_i) + \log\{1 - F(x_{n-i+1})\}].$$

where F is the theoretical cumulative distribution of the distribution being tested and n is the number of data points, the number of observations, or equivalently, the sample size.

4.2.2.3 The Akaike Information Criterion (AIC)

$$AIC = 2k - 2\log\big(L(\theta)\big),$$

where k is the number of parameters estimated and $\log L(\theta)$ is the natural log-likelihood function.

4.2.2.4 Bayesian Information Criterion (BIC)

$$BIC = -2\log\big(L(\theta)\big) + k\log(n),$$

where $\log L(\theta)$ is the natural log-likelihood function and n is the number of observations.

5. APPLICATION

In this section, we apply the new model, RR and current models to the actual claim data for comparison.

5.1 Data

We consider the motor insurance claim according to 3 voluntary plans. The claim data payment of the year 2009 belongs to non-life insurance in Thailand which are named as Plan-A, Plan-B and Plan-C. The claim amount is the individual data in Thai Baht. The number of claim policies (n) of Plan-A, Plan-B and Plan-C are 42, 1,296 and 2,894, respectively. Some descriptive statistics are shown to be information about the dataset in Table 1.

We have seen that the observed claim amount of all plans are the right skewed distributions. The skewness of Plan A, B and C are 4.50, 10.67 and 32.60, respectively, which are illustrated by the histograms of claim amount with logarithmic scales as in the Figure 1.

Table 1
Some Descriptive Statistics of Motor Insurance Claims

Item	Plan		
	A	**B**	**C**
Mean	7,405.02	17,662.45	14,638.17
Median	4,600	7,297	4,403
Skewness	4.50	10.67	32.60
Kurtosis	23	181	1,317
Standard deviation	11,349	41,332	90,606
Coefficient of variation	1.53	2.34	6.19

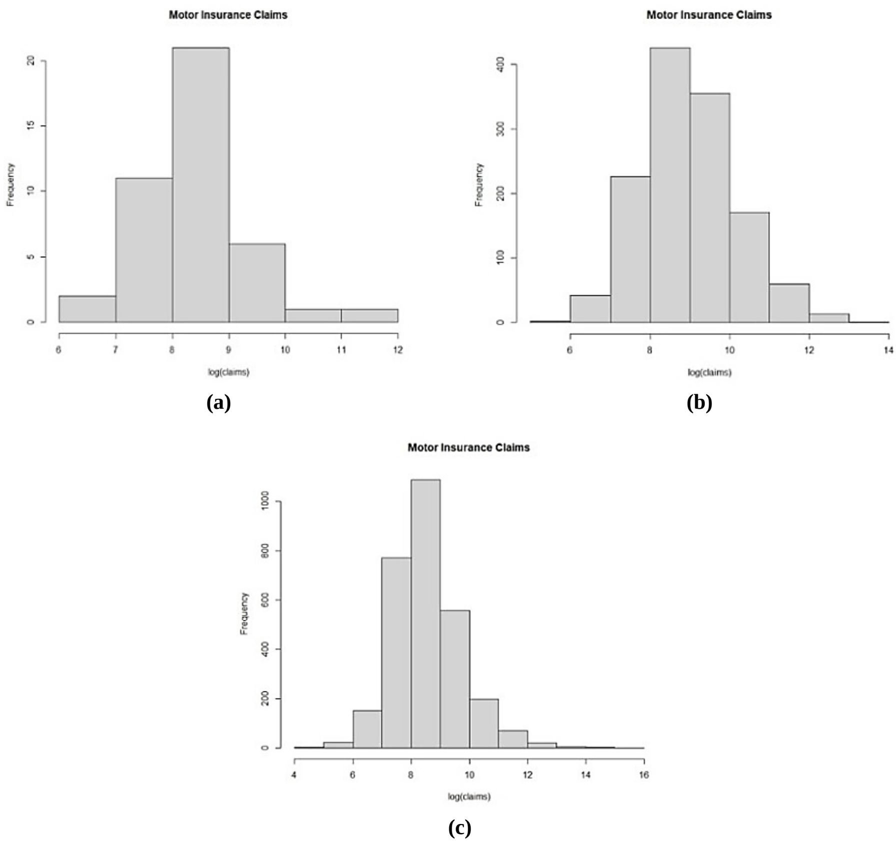


Figure 1: Histograms of log(claim) for (a) Plan-A (b) Plan-B and (c) Plan-C

5.2 Results

The 3 Tables show the estimated parameters, K-S test, AD test, AIC and BIC statistics. The results at a significant level $\alpha = 0.05$ of model fitting for each data set are as follows.

Table 2: Plan-A (n=42)

The claims data can be fitted by the distribution of RR, LN, EEE and Wei by the K-S and AD test. Based on the lowest value of the AIC and BIC, the RR distribution is the best fit for the data, following the distributions of LN, EEE, GPD, Exp, Wei, Gam, EE and R, respectively. The EE is the least suitable for the data.

Figure 2 shows the pdf of the distribution of RR, LN and EEE, respectively.

Figures 3 and 4 show the cdf of the distributions and P-P plots of the distributions of RR, LN and EEE versus the empirical distribution, respectively. They show that the fitted cdf for RR distribution is closer to the empirical cdf than the fits of the distributions of LN and EEE. The RR distribution is also good fit to the large claim amount for the dataset of Plan-A.

Table 3: Plan-B (n=1,296)

The claims data cannot be fitted by any models with the K-S or AD test. Based on the lowest value of the AIC and BIC, the LN distribution is the best fit for the data, followed by the distributions of EEE, GPD, RR, EE, Wei, Gam, Exp and R, respectively. The RR distribution fits to the data better than the EE distribution. The distributions of RR and EE are more suitable for the data than Wei, Gam, Exp and R.

Figure 5 shows the pdf of the distribution of LN, EEE and GPD, respectively. Figures 6 and 7 show the cdf of distributions and P-P plots of the distributions of LN, EEE and GPD versus the empirical distribution, respectively. Their figures demonstrate to confirm which the LN distribution is better fit than the distributions of EE and GPD.

Table 4: Plan-C (n=2,894)

The claims data cannot be fitted to any models with the K-S and AD tests. Based on the lowest value of the AIC and BIC, the EEE distribution is the best fit for the data, followed by the distributions of LN, RR, GPD, EE, Wei, Gam, Exp and R, respectively. The RR distribution fits to the data better than EE and some current distributions, except for EEE and LN. The EE distribution is more suitable for the data than Wei, Gam, Exp and R.

Figure 8 shows the pdf of the distribution of EEE, LN and RR, respectively. Figures 9 and 10 show the cdf of distributions and P-P plots of the distributions of EEE, LN and RR versus the empirical distribution, respectively. Similar to the dataset of Plan-B, the EEE distribution is the best fit and suitable for the Plan-C dataset.

Table 2
Motor Insurance Data of Plan-A (n=42)

Distributions	Estimated Parameters	Measurements of Model Fitting			
		K-S test	AD test	AIC	BIC
RR	$t = 4.3939 \times 10^3$	$D = 0.0961$	$AD = 0.2658$	817.8218	819.5595
		$p-value = 0.8329$	$p-value = 0.9612$		
EE	$b = 4.4661 \times 10^3$	$D = 0.2396$	$AD = 3.3038$	841.1190	842.8567
		$p-value = 0.01607$	$p-value = 0.0194$		
EEE	$b = 48.2180$	$D = 0.1440$	$AD = 1.2939$	826.6241	830.0994
	$\alpha = 66.1336$	$p-value = 0.3486$	$p-value = 0.2342$		
Exp	$\lambda = 1.1536 \times 10^{-4}$	$D = 0.2289$	$AD = 2.6899$	835.4222	837.1598
		$p-value = 0.0245$	$p-value = 0.0397$		
R	$\sigma = 9.5019 \times 10^3$	$D = 0.5476$	$AD = 29.7411$	917.2448	918.9825
		$p-value < 0.01$	$p-value < 0.01$		
LN	$\mu = 8.4227157$	$D = 0.1033$	$AD = 0.4244$	821.6733	825.1486
	$\sigma = 0.8981225$	$p-value = 0.7609$	$p-value = 0.8234$		
Gam	$a = 0.9212$	$D = 0.2191$	$AD = 2.5981$	838.1186	841.5939
	$r = 1.0627$ x 10^{-4}	$p-value = 0.0355$	$p-value = 0.0443$		
Wei	$a = 0.9716$	$D = 0.1791$	$AD = 1.7678$	836.3515	839.8268
	$b = 7.2844 \times 10^3$	$p-value = 0.1351$	$p-value = 0.1240$		
GPD	$b = 5.7185 \times 10^3$	$D = 0.2151$	$AD = 1.7659$	832.3463	835.8217
	$s = 0.2096$	$p-value = 0.0410$	$p-value = 0.1243$		

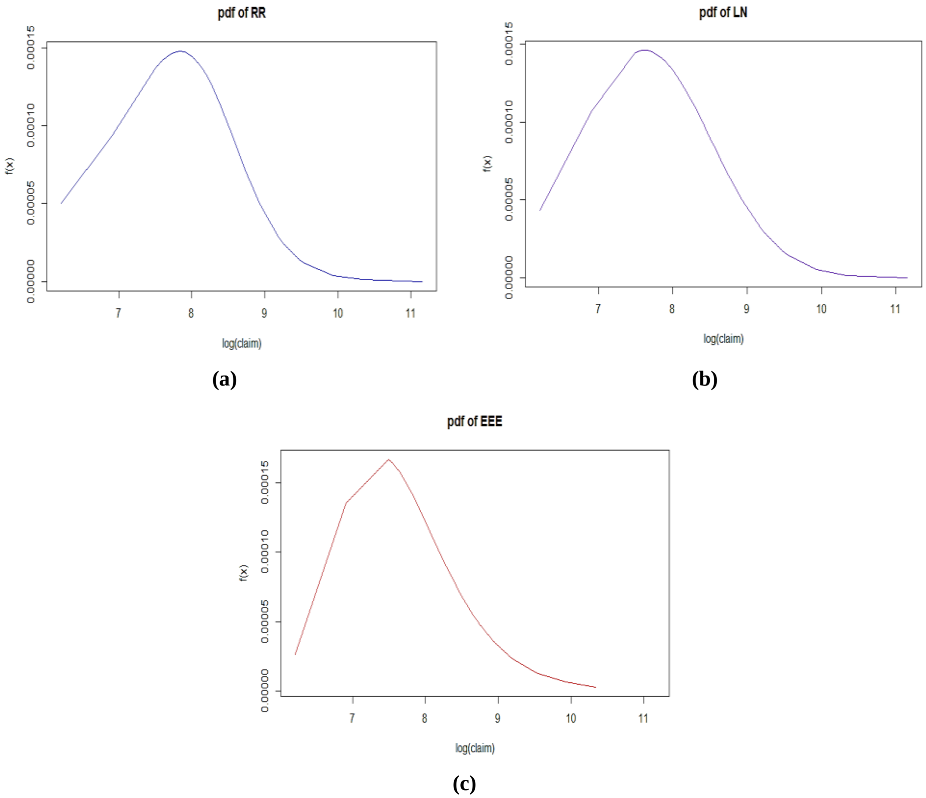
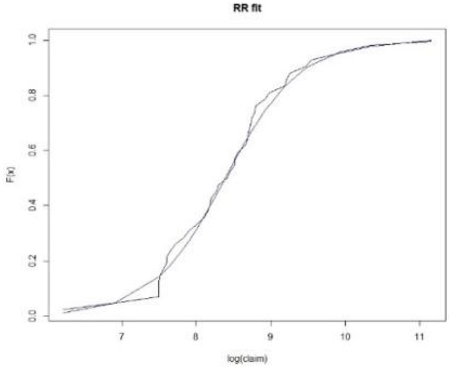
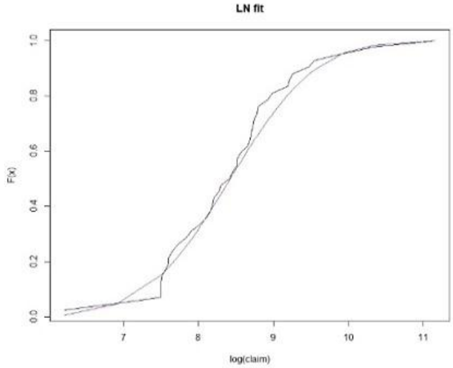


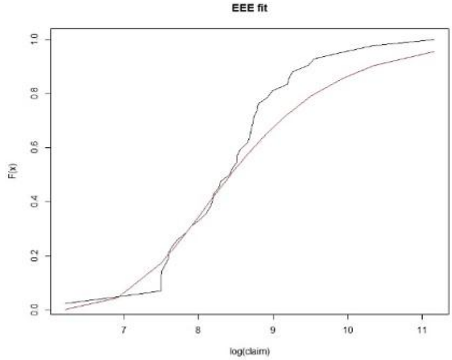
Figure 2: The pdf of Distributions (a) RR (b) LN and (c) EEE



(a)

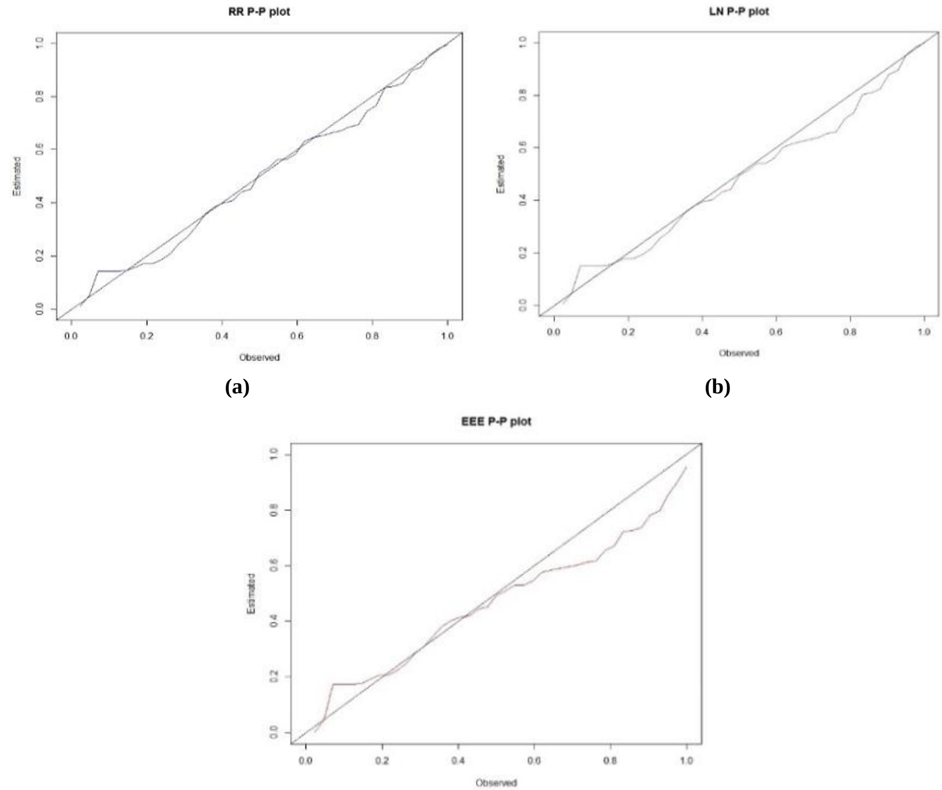


(b)



(c)

Figure 3: The cdf of Distributions (a) RR (b) LN and (c) EEE



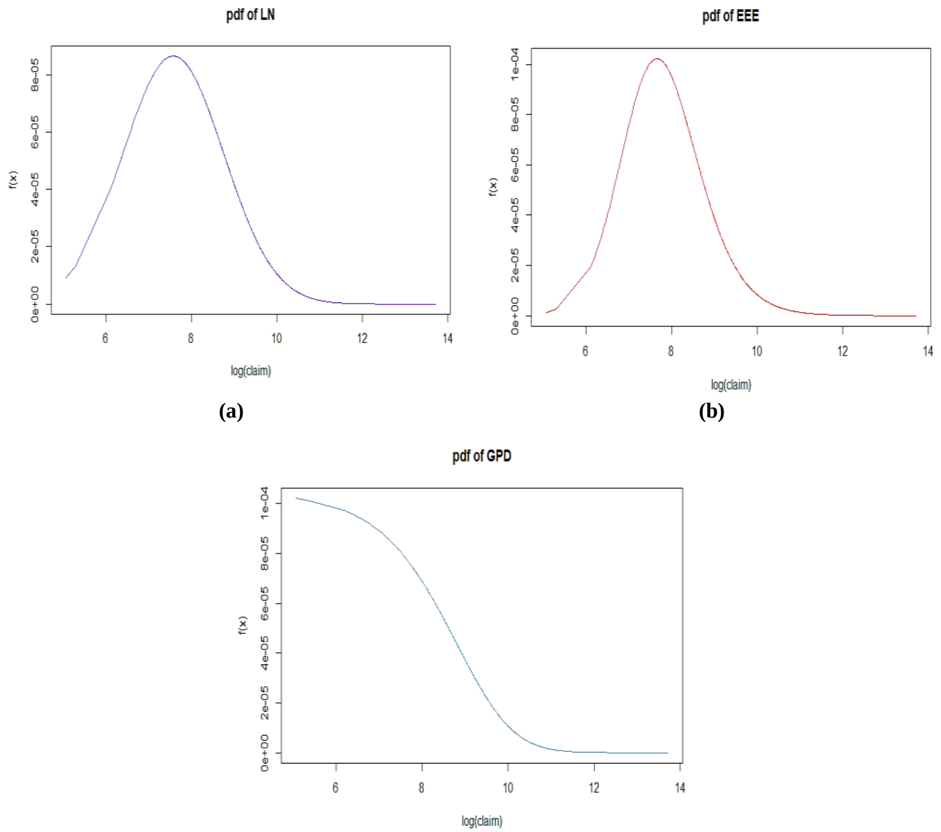
(a) **(b)**

(c)

Figure 4: The P-P Plot (a) RR (b) LN and (c) EEE

Table 3
Motor Insurance Data of Plan-B (n=1,296)

Distributions	Estimated Parameters	Measurements of Model Fitting			
		K-S test	AD test	AIC	BIC
RR	$t = 7.4389$ x 10^3	$D = 0.0875$	$AD = 33.4740$	27526.2502	27531.4173
		$p - value < 0.01$	$p - value < 0.01$		
EE	$b = 7.6240$ x 10^3	$D = 0.1215$	$AD = 37.0350$	27620.5172	27625.6842
		$p - value < 0.01$	$p - value < 0.01$		
EEE	$b = 1.1379$ x 10^3	$D = 0.0381$	$AD = 4.4783$	27364.9162	27375.2503
	$\alpha = 4.7260$	$p - value = 0.0462$	$p - value < 0.01$		
Exp	$\lambda = 5.6617$ x 10^{-5}	$D = 0.1961$	$AD > 38$	27941.6769	27946.8439
		$p - value < 0.01$	$p - value < 0.01$		
R	$\sigma = 3.1772$ x 10^4	$D = 0.6321$	$AD > 38$	33090.1921	33095.3591
		$p - value < 0.01$	$p - value < 0.01$		
LN	$\mu = 8.967189$	$D = 0.0466$	$AD = 3.3749$	27354.8190	27365.1530
	$\sigma = 1.180438$	$p - value < 0.01$	$p - value = 0.0177$		
Gam	$a = 0.7388$	$D = 0.1473$	$AD > 38$	27855.4500	27865.7900
	$r = 4.1826$ x 10^{-5}	$p - value < 0.01$	$p - value < 0.01$		
Wei	$a = 0.7752$	$D = 0.1160$	$AD = 33.0300$	27723.1490	27733.4831
	$b = 1.4425$ x 10^4	$p - value < 0.01$	$p - value < 0.01$		
GPD	$b = 9.5685$ x 10^3	$D = 0.0978$	$AD = 15.2370$	27486.1322	27496.4663
	$s = 0.4360$	$p - value < 0.01$	$p - value < 0.01$		



(a) **(b)** **(c)**

Figure 5: The pdf of Distributions (a) LN (b) EEE and (c) GPD

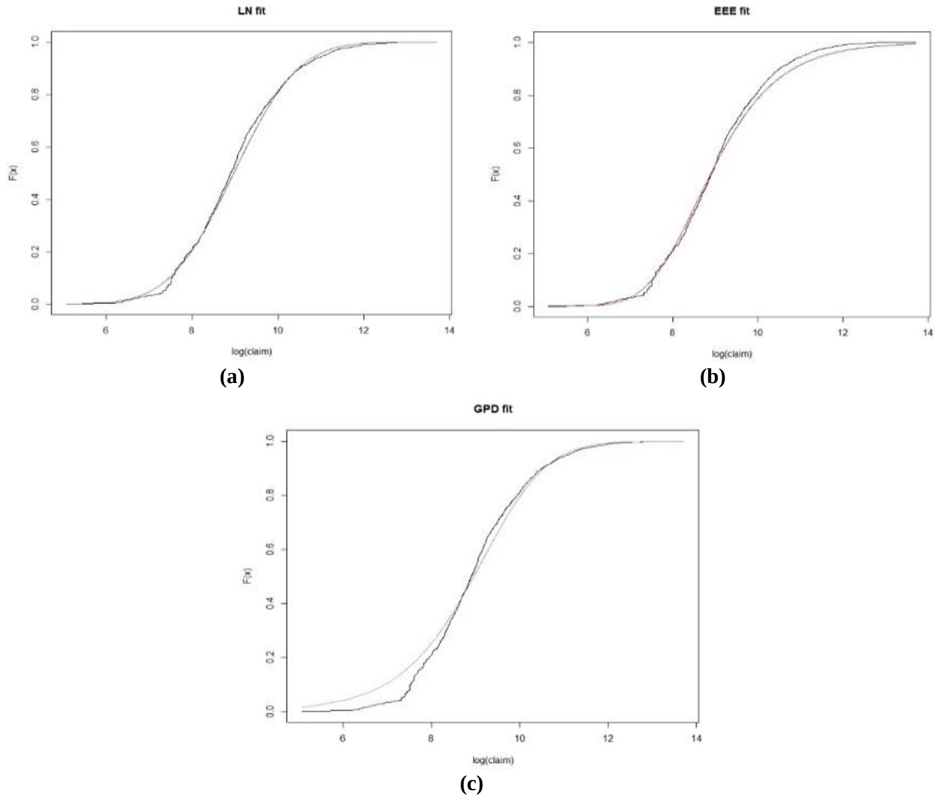
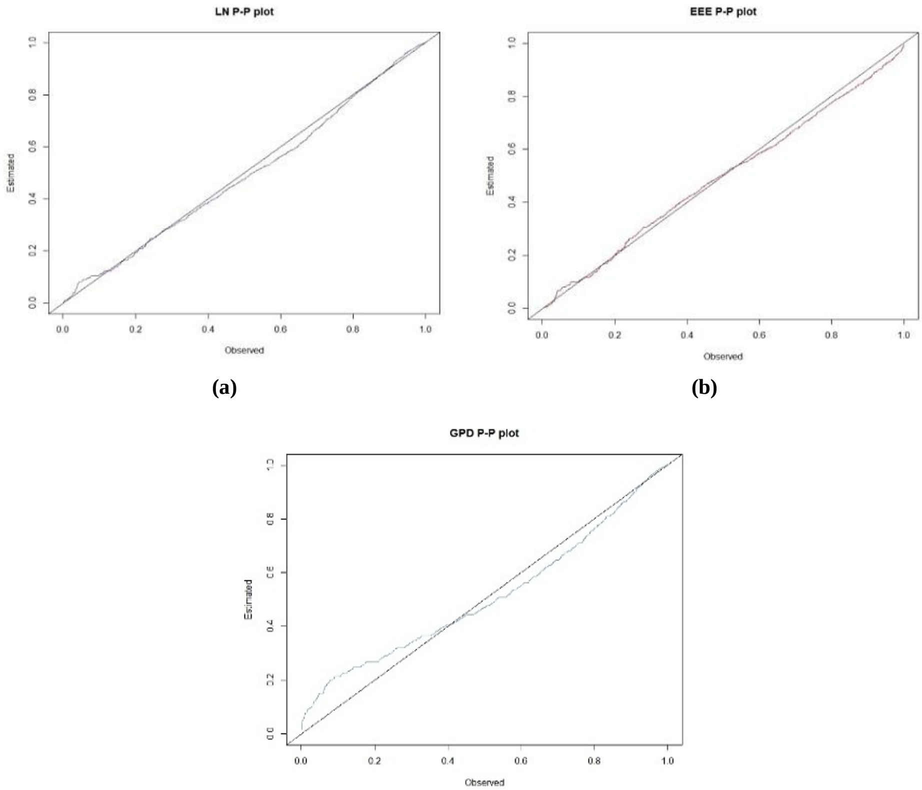


Figure 6: The cdf of Distributions (a) LN (b) EEE and (c) GPD

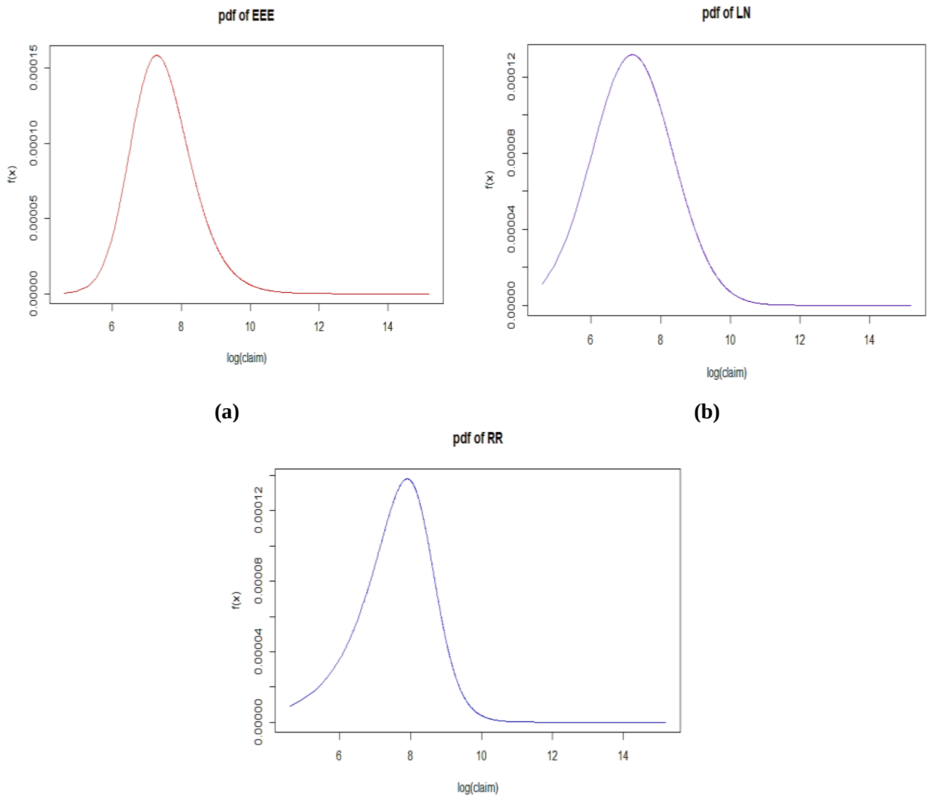


(a) **(b)** **(c)**

Figure 7: The P-P Plot (a) LN (b) EEE and (c) GPD

Table 4
Motor Insurance data of Plan-C (n=2,894)

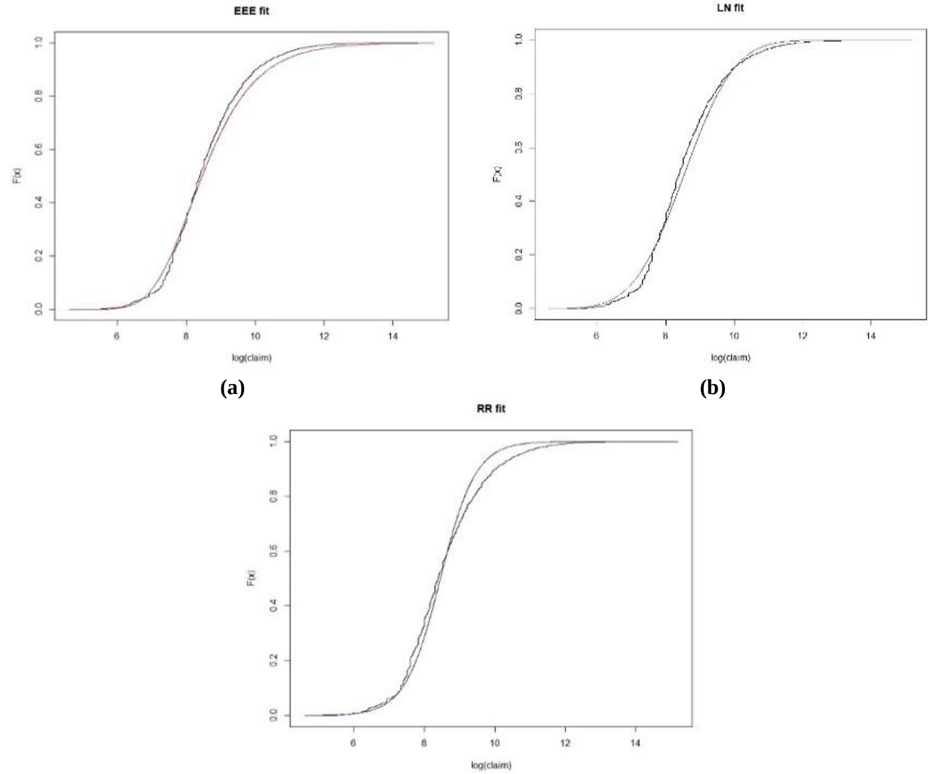
Distributions	Estimated Parameters	Measurements of Model Fitting			
		K-S test	AD test	AIC	BIC
RR	$t = 4.7023$ x 10^3	D = 0.0781	AD = 53.1680	58688.0303	58694.0007
		$p - value < 0.01$	$p - value < 0.01$		
EE	$b = 4.8967$ x 10^3	D = 0.1443	AD = 108.9600	59131.0029	59136.9733
		$p - value < 0.01$	$p - value < 0.01$		
EEE	$b = 5.1811$ x 10^2	D = 0.0494	AD = 16.9220	58409.3105	58421.2513
	$\alpha = 6.6460$	$p - value < 0.01$	$p - value < 0.01$		
Exp	$\lambda = 6.8315$ x 10^{-5}	D = 0.2837	$AD > 109$	61304.9500	61310.9200
		$p - value < 0.01$	$p - value < 0.01$		
R	$\sigma = 6.4888$ x 10^4	$D = 0.84204$	$AD > 109$	84614.3927	84620.3631
		$p - value < 0.01$	$p - value < 0.01$		
LN	$\mu = 8.5423$	D = 0.0677	AD = 21.7310	58542.1449	58554.0857
	$\sigma = 1.1648$	$p - value < 0.01$	$p - value < 0.01$		
Gam	$a = 0.5904$	D = 0.1957	$AD > 109$	60641.2200	60653.1600
	$r = 4.0330$ x 10^{-5}	$p - value < 0.01$	$p - value < 0.01$		
Wei	$a = 0.6922$	D = 0.1544	$AD > 109$	59826.8415	59838.7823
	$b = 9.4792$ x 10^3	$p - value < 0.01$	$p - value < 0.01$		
GPD	$b = 5.8894$ x 10^3	D = 0.1232	AD = 59.8190	58863.9813	58875.9221
	$s = 0.4887$	$p - value < 0.01$	$p - value < 0.01$		



(a) **(b)**

(c)

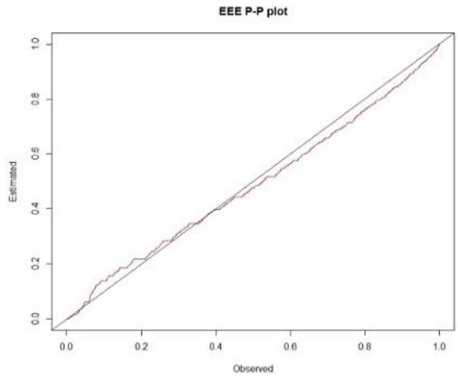
Figure 8: The pdf of Distributions (a) EEE (b) LN and (c) RR



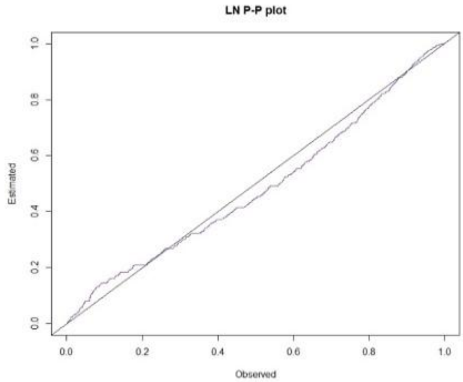
(a) (b)

(c)

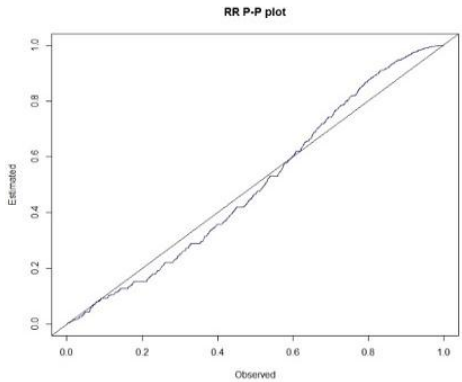
Figure 9: The cdf of Distributions (a) EEE (b) LN and (c) RR



(a)



(b)



(c)

Figure 10: The P-P Plot (a) EEE (b) LN and (c) RR

6. CONCLUSION

For all data, the Rayleigh-Rayleigh distribution is a better fit than the distributions of Exponential-Exponential and Rayleigh. The distributions of Rayleigh-Rayleigh and Exponential-Exponential are more suitable for the data than some traditional distributions such as Weibull, Gamma, Exponential and Rayleigh.

ACKNOWLEDGEMENTS

The author would like to thank the referees for their valuable comments.

REFERENCES

1. Akarawak, E.E.E., Adeleke, I.A. and Okafor, R.O. (2017). The Gamma-Rayleigh Distribution and Applications to Survival Data. *Nigerian Journal of Basic and Applied Science*, 25(2), 130-142.
2. Archer, C.O. (1967). *Some properties of Rayleigh Distributed Random Variables and of Their Sums and Products*. Systems Evaluation Division Technical Publications Division.
3. Barreto-Souza, W., Santos, H.S.A. and Cordeiro, M.G. (2010). The Beta Generalized Exponential Distribution. *Journal of Statistical Computation and Simulation*, 80, 159-172.
4. Bartoszewicz, J. (2009). On a representation of weighted distributions. *Statistics and Probability Letters*, 79(15), 1690-1694.
5. Bashir, S. and Naqvi, I.B. (2016). A new Weighted Exponential Distribution and its Applications on Waiting Time Data. *International Journal of Scientific and Research Publications*, 6(7), 698-702.
6. Beckmann, P. (1964). Rayleigh Distribution and Its Generalizations. *RADIO SCIENCE Journal of Research NBS/USNC-URSI*, 68(9), 927-932.
7. Dankunprasert, S. (2017). *The Modeling of Claim Severity with Infinite Mixture Models for Individual Data of Motor Insurance*. Master of Science Thesis in Applied Mathematics, Graduate School, Khon Kaen University.
8. Dankunprasert, S., Jaroengeratikun, U. and Talangtam, T. (2021). The Properties of Inverse Pareto Distribution and its Application to Extreme Events. *Thailand Statistician*, 19(1), 1-12.
9. Fatima, K. and Ahmad, S.P. (2017). Weighted Inverse Rayleigh Distribution. *International Journal of Statistics and Systems*, 12(1), 119-137.
10. Flaih, A., Elsalloukh, H., Mendi, E. and Milanova, M. (2012). The Exponentiated Inverted Weibull distribution. *Applied Mathematics and Information Sciences*, 6(2), 167-171.
11. Gupta, R.C., Gupta, P.L. and Gupta, R.D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics - Theory and Methods*, 27(4), 887-904.
12. Hogg, R.V., Craig, A. and McKean, J.W. (2005). *Introduction to mathematical statistics*. New Jersey: Prentice Hall.
13. Kareema, A.A. and Mohammad, A.B. (2013). Exponential Pareto Distribution. *Mathematical Theory and Modeling*, 3(5), 135-146.
14. Karim, R., Hossain, P., Begum, S. and Hossain, F. (2011). Rayleigh Mixture Distribution. *Journal of Applied Mathematics*. Article ID 238290, 17 pages.
15. Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2012). *Loss models: from data to decisions*. New York: John Wiley & Sons Inc.
16. Mahmoud, O.H. (2014). Construction Actuarial Model for Aggregate Loss under Exponentiated Inverted Weibull Distribution. *Applied Mathematical Sciences*, 8(162), 8085-8097.
17. Muhammad, A.H. (2016). Kumaraswamy Exponentiated Inverse Rayleigh Distribution. *Mathematical Theory and Modeling*, 6(3), 93-104.
18. Muhammad, S.K. (2014). Modified Inverse Rayleigh Distribution. *International Journal of Computer Applications*, 87(13), 28-33.

19. Oguntunde, P.E. Adejumo, A.O. Okagbue, H.I. and Rastogi, M.K. (2016). Statistical Properties and Applications of A New Lindley Exponential Distribution. *Gazi University Journal of Science*, 29(4), 831-838.
20. Perveen, Z., Ahmed, Z. and Ahmad, M. (2016). On Size-Biased Double Weighted Exponential Distribution (SDWED). *Scientific Research Publishing Inc. Open Journal of Statistics*, 2016(6), 917-930.
21. Zamani, H., Faroughi, P. and Ismail, N. (2014). Bivariate Poisson-weighted exponential distribution with applications. Proceedings of the *3rd International Conference on Mathematical Sciences AIP Conf. Proc*. 1602, 964-968. AIP Publishing LLC.