# A Study of Exponential-Exponential and Exponentiated Inverted Weibull Distributions for Motor Insurance Claim

Sukanda Dankunprasert[1*] and Tosaporn Talangtam[2]
[1] Mathematics Department, Faculty of Science, Khon Kaen University, Khon Kaen, 40002, Thailand, sukanda_d@kkumail.com
[2] Mathematics Department, Faculty of Science, Khon Kaen University, Khon Kaen, 40002, Thailand, tosata@kku.ac.th

**Abstract**

In this paper, we study the modeling of claim severity and compare these models. The models are Exponential-Exponential (EE), Exponentiated Inverted Weibull (EIW) and Exponential (EXP) distributions which are appropriat for motor insurance claims. The claims data sets of individual claims and aggregate claims are considered. The maximum likelihood estimation (MLE) is used for parameters estimation and the goodness of fit test (GOF) is Kolmogorov-Smirnov test (K-S test). We conclude that the EE, EIW and EXP are appropriate for models for the data of individual claims, aggregated claims by day and aggregated claims by month, respectively.

***Keywords***: Maximum likelihood estimation (MLE), Fixed Point Iteration method, Aggregate Claims

## 1. Introduction

Losses of non-life insurance are referred to as claim or liabilities for which insurance companies are responsible for the insured or policyholder. Claim modeling is an important task for them to be able to estimate or forecast the behavior of claims which will occur in the future. Exponential distribution is helpful for insurance business because it has more properties for explain and use to claim modeling. The building claim models are developed by many authors relative to their data in hands. For example, Reference [6] constructed Exponential-Exponential distribution (EE) for severity claim of motor insurance. It used infinite mixture model that was constructed by Exponential and Exponential distribution. Reference [1] is a generalization to the Inverted Weibull distribution via adding a new shape parameter by exponentiation to distribution function, $F$, and the new distribution function, $F^{\theta}$.

Reference [3] contains some lessons for loss distributions and their modeling. Aggregate loss models, see chapter 9 of [3], explain two ways to build a model for the amount paid on all claims occurring in a fixed time period on a defined set of insurance contracts. The individual model is a natural construction for an insurance portfolio. The **individual risk model** represents the aggregate loss as

$$S = X_1 + X_2 + \cdots + X_n, \quad n = 0,1,2,\ldots \tag{1}$$

The collective model has been suggested as an approximation to the individual model. The **collective risk model** has the representation as

$$S = Y_1 + Y_2 + \cdots + Y_N, \quad N = 0,1,2,\ldots \tag{2}$$

where $Y_N$ are independent and identically distributed (i.i.d.) random variables and mutually independent random variables and are independent of $N$.

The real claims data can be recorded by individual and collective claims which are appropriate for each insurance policies. Reference [5] constructed a model and simulations in non-life insurance mainly in many cases that we cannot find a sufficient number of real data. A typical example is the collective risk, that we know only one value of the calendar year. Reference [4], Estimation of the operational risk capital under the Loss Distribution Approach requires evaluation of aggregate (compound) loss distributions which is one of the classic problems in risk theory.

Therefore, we consider Exponential, Exponentiated Inverted Weibull and Exponential-Exponential distributions and apply individual and aggregate claims.

## 2. Research Methodology

In this section, we present methodologies of modeling for Exponential, Exponentiated Inverted Weibull and Exponnetial-Exponential distributions. The maximum likelihood estimation (MLE) is used for parameters estimation and the goodness of fit test is Kolmogorov test (K-S test).

### *2.1 Parameters Estimation*

The method of maximum likelihood provides estimators which are usually quite satisfactory and frequently used in actuarial mathematics.

Let a vector $X = (x_1, x_2, \ldots, x_n)'$ be an independent observation. The amount $x_i$ is paid for the $i^{th}$ contract. The MLE provides the estimated parameters for fitting the distribution to the data set.

The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i;\ \theta),\ \ i = 1, 2, \ldots, n,\ \ \theta \in \Omega.$$

The parameter $\theta$ is unknown where $f_X(x_i;\ \theta)$ is pdf of loss distribution of $X$ with parameter $\theta$.
We define the log-likelihood as

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f_X(x_i, \theta).$$

Then maximize $\ln L(\theta)$ to find the estimated parameters, that is

$$\frac{\partial}{\partial\theta}(\ln L(\theta)) = 0.$$

*2.2 The Models*

2.2.1 Exponential Distribution

We say that the random variable $X$ has an Exponential distribution (EXP) if its cumulative distribution function (cdf) takes the following form:

$$F(x) = 1 - \exp(-\frac{x}{\lambda})\ ; \lambda > 0, x > 0.$$

Therefore, the probability density function (pdf) is:

$$f(x) = \frac{1}{\lambda}\exp(-\frac{x}{\lambda})\ ; \lambda > 0, x > 0.$$

2.2.2 Exponentiated Inverted Weibull Distribution

We say that the random variable $X$ has a standard Exponentiated Inverted Weibull distribution (EIW), see [1], if its cumulative distribution function takes the following form:

$$F(x) = \left(\exp\left(-x^{-\beta}\right)\right)^{\theta}; \beta, \theta > 0, x > 0$$

which simplify the $\theta$-th power of the distribution function of the Exponentiated Inverted Weibull distribution. Here, $\beta$ and $\theta$ are the shape parameters. Therefore, the probability density function is:

$$f(x) = \theta\beta x^{-(\beta+1)}\left(\exp\left(-x^{-\beta}\right)\right)^{\theta}; \beta, \theta > 0, x > 0.$$

2.2.3 Exponential-Exponential Distribution

We say that the random variable $X$ has an Exponential-Exponential distribution (EE), see [6], if its cumulative distribution function (cdf) takes the following form:

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

Therefore, the probability density function (pdf) is:

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

*2.3 Parameters Estimation*

2.3.1 The pdf of Exponential distribution is

$$f_X(x) = \frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right).$$

The likelihood function is written as follows:

$$L(\beta) = \prod_{i=1}^{n}\frac{1}{\beta}exp\left(-\frac{x_i}{\beta}\right).$$

Then

$$\ln L(\beta) = \sum_{i=1}^{n}\ln\left\{\frac{1}{\beta}exp\left(-\frac{x_i}{\beta}\right)\right\}.$$

We estimate $\hat{\beta}$ for $\beta$ by $\frac{\partial}{\partial\mu} ln\, L(\beta) = 0$.

We obtain the estimated parameter $\beta$ as follows:

$$\beta = \frac{\sum_{i=1}^{n} x_i}{n}.$$

2.3.2 The pdf of Exponentiated Inverted Weibull distribution is

$$f(x) = \theta\beta x^{-(\beta+1)}\left(\exp\left(-x^{-\beta}\right)\right)^{\theta}; \beta, \theta > 0, x > 0.$$

The likelihood function can be written as follows:

$$L(\beta,\theta) = \prod_{i=1}^{n} \theta\beta x_i^{-(\beta+1)}\left(\exp\left(-x_i^{-\beta}\right)\right)^{\theta}$$

and the log-likelihood function is in the form

$$\ln L(\beta,\theta) = \sum_{i=1}^{n} \ln\left\{\theta\beta x_i^{-(\beta+1)}\left(\exp\left(-x_i^{-\beta}\right)\right)^{\theta}\right\}.$$

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

$$\frac{\partial}{\partial\beta} ln\, L(\beta,\theta) = \frac{n}{\sum_{i=1}^{n} \ln x_i - \theta\beta \sum_{i=1}^{n} x_i^{-\beta-1}}$$

and

$$\frac{\partial}{\partial\theta} ln\, L(\beta,\theta) = \frac{n}{\sum_{i=1}^{n} x_i^{-\beta}}.$$

We estimate $\hat{\beta}$ for $b$ by $\frac{\partial}{\partial\beta}\ln L(\beta,\theta) = 0$ and $\hat{\theta}$ for $\theta$ by $\frac{\partial}{\partial\theta}\ln L(\beta,\theta) = 0$.

Thus,

$$\frac{n}{\sum_{i=1}^{n} \ln x_i - \theta\beta \sum_{i=1}^{n} x_i^{-\beta-1}} = 0$$

and

$$\frac{n}{\sum_{i=1}^{n} x_i^{-\beta}} = 0.$$

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

2.3.3 The pdf of Exponential-Exponential distribution is

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

The likelihood function can be written as follows:

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

and the log-likelihood function is in the form

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

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

$$\frac{\partial}{\partial b} ln\, L(b) = \frac{n}{b} - 2\sum_{i=1}^{n} \frac{1}{(x_i+b)}.$$

We estimate $\hat{b}$ for $b$ by $\frac{\partial}{\partial b}\ln L(b) = 0$.

Thus,

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

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

*2.4 Goodness of fit test*

The goodness of fit (GOF) tests measure the compatibility of a random sample with a theoretical probability distribution function. We use the Kolmogorov-Smirnov Test (K-S test) to decide if a sample comes from a hypothesized continuous distribution. It based on the empirical cumulative distribution function (ECDF). The empirical cdf is denoted by

$$\mathrm{F_n(x)} = \frac{1}{\mathrm{n}}[\text{Number of observations} \leq \mathrm{x}].$$

The K-S test statistic is defined by

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

where $F_X^*(x)$ is the theoretical cumulative distribution of the distribution being tested.

## 3. Research Results and Discussion

We consider claims data set of Type-1, all coverages of voluntary motor insurance, for model fitting. There are 51,878 observations for a non-life insurance company in Thailand in 2009. There are 2,900,000 Baht, 131 Baht and 16,499 Baht for maximum, minimum and mean of claims data set, respectively.

Table 1 shows the statistical test value for models fitting of the individual and aggregated claims. At a significant level $\alpha = 0.01$, we found that the models cannot be fitted for individual and aggregated claims. For individual data, we found that the EE is the better fit to the individual data with D-value of 0.1259. The following are EXP with D-value 0.1999 and EIW with D-value 0 .5022, respectively. The EIW is the better fit to the aggregated claims data by day with 0.6888. The following are EXP with 0.7 and EE with 0.9060, respectively. For the aggregated claims data by month, the EXP is the best fit model with D-value 0.4389, following EE with D-value 0.4687 and EIW with D-value 0.6296, respectively.

Table 1: The fitting for Individual and Aggregated claims

| Data | Distribution | | | | | |
|---|---|---|---|---|---|---|
| | EXP | | EIW | | EE | |
| | Parameters | D-value | Parameters | D-value | Parameters | D-value |
| Individual Claims | | | | | | |
| | $\hat{\lambda} = 1.6498 \times 10^4$ | 0.1999 | $\hat{\theta} = 2.6976$<br>$\hat{\beta} = 0.1125$ | 0 .5022 | $\hat{b} = 7.0646 \times 10^3$ | 0.1259 |
| Time Period of Aggregated claims | | | | | | |
| Day | $\hat{\lambda} = 2.3449 \times 10^6$ | 0.7 | $\hat{\theta} = 2.7175$<br>$\hat{\beta} = 0.0685$ | 0.6888 | $\hat{b} = 2.2155 \times 10^6$ | 0.9060 |
| month | $\hat{\lambda} = 7.1141 \times 10^7$ | 0.4389 | $\hat{\theta} = 2.7182$<br>$\hat{\beta} = 0.0553$ | 0.6296 | $\hat{b} = 7.0722 \times 10^7$ | 0.4687 |

## 4. Conclusion

We conclude that the EE, EIW and EXP are appropriate models for the data of individual claims, aggregated claims by day and aggregated claims by month, respectively.

## 5. Acknowledgements

This research is supported by Science Achievement Scholarship of Thailand.

## 6. References

[1] Flaih A, Elsalloukh H, Mendi E, Milanova M, **The Exponentiated Inverted Weibull distribution**, Applied Mathematics and Information Sciences, 6, No. 2 (2012), 167-171.

[2] Hogg RV, Craig A and McKean JW. **Introduction to Mathematical Statistical**. 6th ed. New Jersey: Prentice Hall; 2005.

[3] Klugman SA, Panjer HH, Willmot GE. **Loss Models: From Data to Decisions**. 3rd ed. New York: John Wiley & Sons; 2008.

[4] Pavel V. Shevchenko, **Calculation of aggregate loss distributions**, The Journal of Operational Risk, 5, No. 2 (2010), 3-40.

[5] Viera P. **Modelling and Simulation in Non-life Insurance**: Proceedings of the 5th international conference on Applied mathematics, simulation, modelling; 2011 July 14 – 16; Corfu Island, Greece. USA: Wisconsin: Stevens Point, 2011.

[6] 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.