The Disparity Between Exponentially and Lognormally Distributed Mean Severities in General Insurance Business

Gbenga Michael Ogungbenle, University of Jos
Email: gbengarising@gmail.com

Abstract: *The purpose of this study is to enable us to obtain mean losses of an insured risk by means of the operational behaviour of density function with deductible modifications and then compare the mean severities under exponentially and log-normally distributed arbitrary policy in a cost per loss and cost per payment circumstances. The mean losses model thus obtained for an arbitrary policy in general insurance under deductible coverage modifications is meant to reduce the number and magnitude of claims received. Furthermore, the mean losses are then used to compute premium numerically, based on the applied deductible. Rate relativity data on deductible was obtained through a non-life insurance agent operating in property insurance market in Lagos. The result show that despite log-normal severity distribution has a thicker tail than the exponential distribution, its cost per loss payment* $\langle Y_L \rangle$ *is correspondingly lower in value than the values of exponential mean loss that is* $\langle Y_L \rangle_{\log normal} < \langle Y_L \rangle_{\exp ential}$ *.While the cost per payment is uniformly constant throughout the entire domain of definition for the deductible under exponential distribution, the insurer experiences higher cost per payment than expected in the subinterval* $45 \le D \le 1$ *under lognormal regime. It is, therefore, recommended that the insurer is advised to apply deductible in this subdomain to disapprove nuisance claims and control the problem of moral hazard.*

Keywords: Cost per loss; Cost per payment; Deductible; Exponential distribution; Log-normal distribution

1. Introduction

In general insurance underwriting practice, underwriting line of business is often restricted by a deductible clause. The idea about a deductible being introduced into the policy contract and the resulting pricing architecture with respect to insurance premium is of fundamental significance in the underwriting sector. The correct computation of premium is important because inappropriate level of pricing may lead to a significant loss or that the underwriter may be edged out of business. Therefore, it generally behoves the underwriter to ascertain a technique of computing the premium in such a way to integrate deductible clause. The rationale behind introducing deductible is essentially to minimize the claim handling charges by eliminating cover for frequent

ISSN 2664-3413 (Print) 2664-3421 (Online)
Daffodil International University Journal of Business and Entrepreneurship, Vol. 13, No. 2, PP. 1-18, December, 2020

nuisance small claims and moreover put in place some measures in favour of scheme holder to prevent claims by means of a bounded level of sharing in the cost of claims. Generally speaking, the fact that a deductible will eventually force the scheme holder to receive a part indemnity in future gives an advantage to minimize the impact of loss so that the liability of both the insured and the insurer are minimized. For small losses, it is probable that the loss administrative charges will be more than the actual losses, consequently, the underwriter would expect and enforce that the scheme holder pays it. The insured may choose to seek high deductible value to arrive at a lower premium since lower pricing would be preferable by the scheme holder. Even though the indemnity value of the scheme holder may be minimized in future by the chosen level of deductible, it is certain that the policyholder's retention level is greater than zero and consequently, loss is avoided. In general insurance, the maximum accumulated amount of losses retained by the insured under deductible policy modifications is usually set as part of the terms of the policy conditions thus specifying the amount which the insured is responsible for, in accordance with the insured peril. When an insured event occurs, the value of the deductible is subsequently defrayed from the claim payment. The higher the deductible, the lower is the premium payable on the insurance policy thus establishing an inverse relationship between premium and deductible. In this work, we investigate the effect of payment per loss and payment per payment on two loss distributions.

2. Literature Review

Claims modelling is a core insurance issue because a good knowledge of loss distribution is the basis of underwriting decisions taken in insurance sector in respect of premiums, expected profit, reserve and re-insurance arrangement. Insurers usually keep record of data base bearing information on deductibles, policies and claims applicable for ratemaking purposes. We observe in Pacakova & Brebera (2015) and Zacaj et al. (2015) that the severity distributions used for risk theoretical analysis could be evaluated only after rigorous data processing because the generation of loss distribution arising from insurance data is very difficult. However, deductible statistics is quite limited since database containing information on deductible, policy and claim may be missing or unavailable. In Zacaj et al. (2015), Bakar et al. (2015), we also observe that claim generating processes is essentially tedious under social-economic conditions and consequently, the magnitude of claim could be obtained by the claim size management of an insurance contract. Following Raschke (2019), medium size claim could be subject of log-normal regime influenced by base distribution function $F_b\left(x\right)$ and base survival distribution $S_b\left(x\right)$ and resulting in a random risk. Insurance managers place strong emphasis on severities in relation to selecting the adequate probability model to analyze claims data. A clear knowledge of loss distribution in risk theory is therefore needed to summarize and model claims data. In actuarial discipline, the ability to understand and interpret claim data is a requirement to build a good claim model which allows us to make critical underwriting judgement in estimating premiums, expected profit, reserve and the knowledge of the distribution of insurance claims data can further be used to advise underwriters on reinsurance matters. In general insurance business, the objective of risk theory is to understand, quantify dimensions and analyze risk of severity. Tse (2009), observes that the actuarial risk theory is responsible for building models of pricing based on observations of the random variables for the size of claim out go $C(e) = \sum_{0}^{N(e)} Y_i I_{(N(e)>0)}$ and frequency of claims $N(e)$ where Y_i are random variables describing individual claims, the exposure e is expressed as $\frac{e}{\zeta} = s^{S}$, is the value of insurance under cover , ζ is the time horizon during which period the value under coverage has been exposed to severity risk and $I_{(N(e)>0)}$ is the indicator function. The main responsibility of a claim actuary in the insurance underwriting is to obtain appropriate values on the cash flows in the insurance industry. From the knowledge of the cash flows, actuarial models are constructed to describe and analyze cash flow process. General insurance including casualty and health insurance describes the most critically challenging sector for claims actuaries because it is driven by data where the cash outflow is the claim outgo. Analysis of severity claims based model such as lognormal and exponential distribution using appropriate density functions, form the basis of solving actuarial claim issues arising in general insurance. Afify et al. (2020) reports that claims managers seldom bother on occurrence of claim but are concerned with the random processes describing the severity value that the claims manager pays as indemnity rather than the particular events resulting in the claim process. Consequently, claims managers should have a good knowledge of loss distribution models which consists of total amount of claims payable by insurers over a defined period. Insurance industry is data driven probably in large amount which could be infrequent and consequently, it is necessary to identify suitable densities models having the characteristics of heavy tails such as log-normal distribution and high skewness such as exponential distribution. Afify et al. (2020) again report that Loss distributions are a description of risk exposure units e the degree of which is computed from risk indicator metrics which are functions of the model. Underwriting managers usually employ the risk metrics to assess the level at which the insurance firms are exposed to defined areas of risk arising from vagaries of underlying variables such as price of equity, interest rate and inflation. Tse (2009) posits, the upper tail of a severity size distribution in general business insurance can be modelled by log-normal distribution although, the upper tail may not necessarily be log-normally distributed. The models should be adequate to the extent of enabling decisions relating to solvency requirements, loading, premium analysis, technical reserves, expected profits forecast, profitability, reinsurance, and the influence of deductibles on severities. The size of claim at a particular time is

ISSN 2664-3413 (Print) 2664-3421 (Online)
Daffodil International University Journal of Business and Entrepreneurship, Vol. 13, No. 2, PP. 1-18, December, 2020

of particular significance to the underwriting management of an insurance firm. The conditions under which claims data are gathered and future claims subsequently estimated is an enabling factor to estimate severity amounts in general insurance business as samples from definite but usually heavy tailed probability distributions. In view of Tse (2009), it is observed that as a probability based actuarial model for severity analysis, the probability of financial losses experienced by scheme holders and indemnified by the insurance firm should be clearly understood under the contract setting. Analytical actuarial distributions are applied to assess the cost to the extent that such distributions are positively skewed having high probability densities on the right tails. Since the distributions are specifically meant to analyze losses, they are tagged loss distribution models. Claims modeling remain the basis of information contents for underwriting firms to obtain estimate of premium, loadings, reserves, profits and capital required to ascertain overall profitability and to appraise the effect of deductible. Although, it is reasonable to fit probability distribution to claims data, analytical probability distributions are rather powerful techniques to employ in analyzing claims data and consequently, there is need to construct models which can be used to estimate the distribution of claims under exponentially and log-normally distributed actuarial data involving deductible clauses. For the purpose of this work, we are concerned with analysis of lognormal and exponential distributions of claims estimation for policies having deductible conditions. In general insurance business, claims data are usually skewed to the right tail and any distribution showing this type of behavior is sufficient for the analysis of severities (Sakthivel & Rajitha, 2017). The choice of these distributions is based on my experience and prior knowledge of claims data in the insurance underwriting as a professional underwriter. An important characteristic of a probability distribution to meet the requirements of a claims model is that it should be able to fit the data. It is assumed that there are no points of truncation in the data; the first moment of the distribution should at least exist. Claim modeling is therefore necessary because the construction of adequate interpretable loss models serves as the foundation of critical underwriting decisions taken in relation to premium and claims assessment to ascertain profitability.

3. Basic Derivations in General Insurance Business

Let $\zeta(Y)$ be the amount of an insurance loss for Y, S be the sum insured and V be the value as at the time of loss, the insurance amount , then if the insurance equals at least β of the value of the policy value as at the time of the loss, then we have

$$\zeta(Y) = \min\left(S; \frac{S \times Y}{\beta \times V}\right) \tag{1}$$

$$\zeta(Y) = \min\left(S; \frac{S \times Y}{\beta \times V}\right) \tag{2}$$

In line with the views of Tse (2009); Ogungbenle et al. (2020); Ogungbenle (2021), the cost per loss is given as $Y^L = (Y - D)_+$ (3)

$$E\left[(Y \wedge D)\right] = \int_0^D S_Y(y)\,dy \tag{4}$$

$$F_{Y^{L(y)}} = F_Y(y + D) \tag{5}$$

$$E\left(Y^L\right) = E\left[(Y - D)_+\right] = \int_D^\infty (y - D) f_Y(y)\,dy \tag{6}$$

$$E\left(Y^L\right) = E\left[(Y - D)_+\right] = \int_D^\infty \left(1 - F_Y(y)\right) dy \tag{7}$$

$$E\left[(Y - D)_+\right] = E(Y) - E\left[(Y \wedge D)\right] \tag{8}$$

Let the cost per payment $Y^P = (Y - D)_+ / Y > D$ (9)

$$F_{Y^{P(y)}} = \frac{F_Y(y + D) - F_Y(D)}{S_Y(D)} \tag{10}$$

$$S_{Y^{P(y)}} = \frac{1 - F_Y(y + D)}{1 - F_Y(D)}, \text{ if } Y^P = (Y - D)_+ / Y > D \tag{11}$$

$$\frac{E\left[(Y - D)_+\right]}{S_Y(D)} = \frac{E(Y) - E\left[(Y \wedge D)\right]}{S_Y(D)} \tag{12}$$

Consequently, the mean excess loss $\varepsilon_Y(D)$

is given as

$$\varepsilon_Y(D) = \frac{\int_D^{\infty} (y-D) f_Y(y)\,dy}{S_Y(D)} \tag{13}$$

$$\varepsilon_Y(D) = \frac{\int_D^{\infty} S_Y(y)\,dy}{S_Y(D)} \tag{14}$$

$$\varepsilon_Y(D) S_Y(D) = E(Y) - E\left[(Y \wedge D)\right] \tag{15}$$

The loss elimination ratio LER is usually obtained as the fraction of the expected loss that the underwriter would not pay because of deductible D

$$LER = \frac{E(Y) - E\left[(Y-D)_+\right]}{E(Y)} = \frac{E\left[(Y \wedge D)\right]}{E(Y)} \tag{16}$$

For the exponential distribution

$$LER = 1 - e^{-\frac{D}{\theta}}, \tag{17}$$

The indicated deductible relativity $REL(D)$ is the ratio of the payment per loss with a deductible D to the payment per loss with the base deductible $\overline{D}$,

$$REL(D) = \frac{\int_D^{\infty} (y-D) f_Y(y)\,dy}{\int_{\overline{D}}^{\infty} (y-D) f_Y(y)\,dy} = \frac{E(Y) - \int_0^{D} S_Y(y)\,dy}{E(Y) - \int_0^{\overline{D}} S_Y(y)\,dy} \tag{18}$$

$$REL(D) = \frac{E\left[(Y-D)_+\right]}{E\left[(Y-\overline{D})_+\right]} = \frac{E(Y) - E\left[(Y \wedge D)\right]}{E(Y) - E\left[(Y \wedge \overline{D})\right]} \tag{19}$$

Consequently, the loss elimination ratio is otherwise obtained as

$$LER = \frac{E\left[(Y \wedge D)\right] - E\left[(Y \wedge \overline{D})\right]}{E(Y) - E\left[(Y \wedge \overline{D})\right]} \tag{20}$$

Let $F_Y(y)$, $S_Y(y)$, $f_Y(y)$ be the distribution, survival and probability density function respectively where Y constitutes some insurance risk and denote as $\eta(y)$, the payment function that corresponds to a deductible. Then if $E[\eta(y)]$ exists, the risk premium P_R is obtained as $P_R = E[\eta(y)]$. Where deductible is not applicable, then $\eta(y) = y$ and provided $E(y)$ exists, then $P_R = E(y)$.

Let $y = e^X$. The distribution of y is called lognormal distribution. Klugman et al. (2004) defines CDF as $F(s) = \Phi\left(\frac{(\log_e s - \mu)}{\sigma}\right) = \int_0^S \frac{1}{\sigma y\sqrt{2\pi}} e^{-\frac{\left[(\log_e y - \mu)^2\right]}{2\sigma^2}} dy$ (21)

where $s > 0$, $\sigma > 0$ and $s \in R$ real number and $\Phi\left(\frac{(\log_e s - \mu)}{\sigma}\right)$ is the standard normal distribution function. The lognormal distribution is applicable in modelling claims, costs and moreover its thick right tail property is well adapted to many loss decisions. An agreement between the scheme holder and the underwriter integrating a deductible amount ρ means that the underwriter compensates for only the part of the claim which exceeds the amount of the deductible ρ. Consequently, if the claim size falls below this amount, then the loss will be repudiated as it will not be covered by the policy.

Then $\eta(y)_\rho(y) = \max(0; y - \rho)$ and we have $P_R = E[\eta(y)]$ becomes

$$P_R(\rho) = P - E(y,\rho) = P_R(\rho) - \rho \times (S_Y(\rho)) \quad (22)$$

$$P - E(y,\rho) = P_R(\rho) - \rho \times (S_Y(\rho)) \quad (23)$$

$$-E(y,\rho) = -P + P_R(\rho) - \rho \times (S_Y(\rho)) \quad (24)$$

$$E(y,\rho) = P - P_R(\rho) + \rho \times (S_Y(\rho)) \quad (25)$$

But by definition, $E(Y,y)$ the expected limited value function is given by

$$E(Y,u) = \int_0^u sf(s)ds + uS_Y(u) \quad (26)$$

$$E(Y,\rho) = \int_0^\rho sf(s)ds + \rho S_Y(\rho) \quad (27)$$

$$\int_0^\rho sf(s)ds + \rho S_Y(\rho) = P - P_R(\rho) + \rho \times (S_Y(\rho)) \quad (28)$$

$$\int_0^\rho sf(s)ds = P - P_R(\rho) \quad (29)$$

3.1 Main Results: Theorem 1

$$P_R^{LOGNORMAL}(\rho) = P - E(y,\rho) = e^{\mu+\frac{\sigma^2}{2}}\left[1-\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)\right] - \rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\} \tag{30}$$

Proof

$$F(s) = \Phi\left(\frac{(\log_e s-\mu)}{\sigma}\right) = \int_0^S \frac{1}{\sigma y\sqrt{2\pi}} e^{-\frac{\left[(\log_e y-\mu)^2\right]}{2\sigma^2}} dy \tag{31}$$

The function then $\eta(y) = y = e^{\mu+\frac{\sigma^2}{2}}$ describes the condition where deductible does not apply

$$E(Y,\rho) = e^{\mu+\frac{\sigma^2}{2}}\left[1-\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)\right] - \rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\} \tag{32}$$

taking $P = e^{\mu+\frac{\sigma^2}{2}}$ and substituting in equation below we have

$$P_R(\rho) = P - E(y,\rho) \tag{33}$$

$$P_R(\rho) = E(y) - E(y,\rho) \tag{34}$$

$$E(Y^n) = e^{n\mu+\frac{n^2\sigma^2}{2}} \tag{35}$$

$$E\left[(Y\wedge y)^n\right] = e^{n\mu+\frac{n^2\sigma^2}{2}}\left[\Phi\left(\frac{\log_e x-\mu-n\sigma^2}{\sigma}\right)\right] + y^n\left\{1-F_Y(y)\right\} \tag{36}$$

$$E\left[(Y\wedge y)^n\right] = e^{n\mu+\frac{n^2\sigma^2}{2}}\left[\Phi\left(\frac{\log_e x-\mu-n\sigma^2}{\sigma}\right)\right] + y^n S_Y(y) \tag{37}$$

When $n=1$, we have

$$E(Y) = e^{\mu+\frac{\sigma^2}{2}} \tag{38}$$

$$E\left[(Y\wedge y)\right] = e^{\mu+\frac{\sigma^2}{2}}\left[\Phi\left(\frac{\log_e x-\mu-\sigma^2}{\sigma}\right)\right] + y\left\{1-F_Y(y)\right\} \tag{39}$$

$$E\left[\left(Y \wedge y\right)\right] = e^{\mu+\frac{\sigma^2}{2}}\left[\Phi\left(\frac{\log_e x-\mu-\sigma^2}{\sigma}\right)\right]+yS_Y\left(y\right) \tag{40}$$

Consequently, as $y \to \rho$, we have

$$E\left[\left(Y \wedge \rho\right)\right] = e^{\mu+\frac{\sigma^2}{2}}\left[\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)\right]+\rho\left\{1-F_Y\left(\rho\right)\right\} \tag{41}$$

$$E\left[\left(Y \wedge y\right)\right] = e^{\mu+\frac{\sigma^2}{2}}\left[\Phi\left(\frac{\log_e x-\mu-\sigma^2}{\sigma}\right)\right]+yS_Y\left(y\right) \tag{42}$$

$$Z_1 = \frac{\log_e \rho-\mu-\sigma^2}{\sigma} \tag{43}$$

$$Z_2 = \frac{\log_e \rho-\mu}{\sigma} \tag{44}$$

then $F_Y\left(Z_1\right)=\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)$ and $F_Y\left(Z_2\right)=\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)$ (45)

$$E\left[\left(Y \wedge \rho\right)\right] = e^{\mu+\frac{\sigma^2}{2}}\left[\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)\right]+\rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\} \tag{46}$$

$$P_R\left(\rho\right)=e^{\mu+\frac{\sigma^2}{2}}-\left[e^{\mu+\frac{\sigma^2}{2}}\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)+\rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\}\right] \tag{47}$$

$$P_R\left(\rho\right)=e^{\mu+\frac{\sigma^2}{2}}-e^{\mu+\frac{\sigma^2}{2}}\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)-\rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\} \tag{48}$$

Therefore, the premium based on deductible ρ under the lognormal distribution frame work is given as follows

$$P_R\left(\rho\right)=e^{\mu+\frac{\sigma^2}{2}}\left[1-\Phi\left(\frac{\log_e \rho-\mu-\sigma^2}{\sigma}\right)\right]-\rho\left\{1-\Phi\left(\frac{\log_e \rho-\mu}{\sigma}\right)\right\} \tag{49}$$

ISSN 2664-3413 (Print) 2664-3421 (Online)
Daffodil International University Journal of Business and Entrepreneurship, Vol. 13, No. 2, PP. 1-18, December, 2020

3.2 Theorem 2

For the exponential distribution,

$$P_R^{EXPONENTIAL}\left(\rho\right)=\theta-\frac{\theta}{\Gamma\left(\frac{\rho}{\theta}\right)}\int_0^2 z^{\frac{\rho}{\theta}-1}e^{-z}dz-\rho e^{\frac{-\rho}{\theta}} \tag{50}$$

Proof

$$E\left(Y^n\right)=\theta^n\Gamma\left(n+1\right);n>-1 \tag{51}$$

if n is an integer, $E\left(Y^n\right)=\theta^n n!$ (52)

$$E\left[\left(Y\wedge y\right)^n\right]=\theta^n\Gamma\left(n+1\right)\Gamma\left(n+1;\frac{y}{\theta}\right)+y^n e^{\frac{-y}{\theta}},n>-1 \tag{53}$$

if $n>-1$ is an integer, $E\left[\left(Y\wedge y\right)^n\right]=\theta^n n!\Gamma\left(n+1;\frac{y}{\theta}\right)+y^n e^{\frac{-y}{\theta}},n>-1,$ (54)

If $n=1$

if n is an integer, $E\left(Y\right)=\theta$ (55)

if $n>-1$ is an integer, $E\left[\left(Y\wedge y\right)\right]=\theta\Gamma\left(2;\frac{y}{\theta}\right)+ye^{\frac{-y}{\theta}}$ (56)

$$P_R\left(\rho\right)=\theta-\left[\theta\Gamma\left(2;\frac{\rho}{\theta}\right)+\rho e^{\frac{-\rho}{\theta}}\right] \tag{57}$$

$$P_R\left(\rho\right)=\theta-\theta\Gamma\left(2;\frac{\rho}{\theta}\right)-\rho e^{\frac{-\rho}{\theta}} \tag{58}$$

Where $\Gamma\left(2;\frac{\rho}{\theta}\right)=\frac{1}{\Gamma\left(\frac{\rho}{\theta}\right)}\int_0^2 z^{\frac{\rho}{\theta}-1}e^{-z}dz$ (59)

is the incomplete gamma distribution function.

$$P_R(\rho)=\theta-\frac{\theta}{\Gamma\left(\frac{\rho}{\theta}\right)}\int_0^2 z^{\frac{\rho}{\theta}-1}e^{-z}dz-\rho e^{\frac{-\rho}{\theta}} \tag{60}$$

4. Materials and Methods

The actuarial modeling of claims amounts is significant because it helps in computing the estimated future claims used in insurance pricing. Usually, the modelling process is carried out through the use of appropriate continuous probability distributions whose expectations describe the severity value that scheme holders could claim. A random variable Y is said to be log-normally distributed if its probability density function is given by $g_Y(y)=\frac{1}{\sigma\sqrt{2\pi}}\frac{1}{z}e^{-\frac{\left[(\log_e z-\mu)^2\right]}{2\sigma^2}}$ (61)

Loss distribution function of lognormal distribution is given by-

$$G_Y(y)=\frac{1}{\sqrt{2\pi\sigma}}\int_D^y \frac{1}{z}e^{-\frac{\left[(\log_e z-\mu)^2\right]}{2\sigma^2}}dz \tag{62}$$

and $E(Y^r)=e^{\left(r\mu+\frac{1}{2}r^2\sigma^2\right)}$, $E(Y)=e^{\left(\mu+\frac{1}{2}\sigma^2\right)}$, $E(Y^2)=e^{\left(2\mu+2\sigma^2\right)}$ (63)

$$Var(y)=e^{\left(2\mu+2\sigma^2\right)}-\left[e^{\left(\mu+\frac{1}{2}\sigma^2\right)}\right]^2 \tag{64}$$

We observe that μ and σ^2 do not necessarily describe the mean and variance of y but are somewhat logarithmic in form. The expected loss under log-normally distributed risk with deductible conditions is defined by

$$\langle Y_L\rangle=\int_D^\infty (y-D)g_Y(y)dy=\int_D^\infty yg_Y(y)dy-\int_D^\infty Dg_Y(y)dy \tag{65}$$

$$\langle Y_L\rangle=\int_D^\infty yg_Y(y)dy-DS_Y(D) \tag{66}$$

where $\int_D^\infty yg_Y(y)dy=\int_D^\infty \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\left[(\log_e y-\mu)^2\right]}{2\sigma^2}}dy\cong e^{\left[\frac{\sigma^2}{2}+\mu\right]\left\{1-\Phi\left(\frac{\ln D-\mu-\sigma^2}{\sigma}\right)\right\}}$ (67)

$$S_Y(D)=\Pr\left[Z>\frac{\log_e y-\mu}{\sigma}\right] \tag{68}$$

From the definition of log-normal distribution, we observe that

$$g_Y(y) = \frac{d}{dy}\Phi\left(\frac{(\log_e y - \mu)}{\sigma}\right) = \frac{1}{\sigma} \times \frac{1}{y} \times \Phi'\left(\frac{(\log_e y - \mu)}{\sigma}\right) \tag{69}$$

$$\Phi(y) = \int_0^y \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds; \Phi\left(\frac{(\log_e y - \mu)}{\sigma}\right) = \int_0^{\frac{(\log_e y - \mu)}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds \tag{70}$$

$$\eta(s) = \int_{a(s)}^{b(s)} g(y,s)\,dy, \text{then, } \frac{d\eta(s)}{ds} = \int_{a(s)}^{b(s)} \frac{\partial}{\partial s} g(y,s)\,dy + g(b(s),s)\frac{\partial}{\partial s} b(s) - g(a(s),s)\frac{\partial}{\partial s} a(s) \tag{71}$$

applying equation the equation (71) to (70), we obtain

$$g_Y(y) = \left(e^{-\frac{\left[(\log_e y - \mu)^2\right]}{2\sigma^2}}\right)\left(\frac{\sqrt{2\pi}}{y \times \sigma \times 2\pi}\right) \tag{72}$$

$y > 0$, location μ and scale $\sigma > 0$

and therefore, $\int_{-\infty}^{y} g_S(s)ds = \frac{(\log_e y - \mu)}{\sigma}$ (73)

The import of the paper lies in highlighting techniques applicable in analyzing probability-based loss distributions as used in general insurance claims analysis. Schlesinger (1985); Thogersen (2016); Liu & Wang (2017); Woodard & Yi (2020) claim that under deductible conditions, the design of an adequate loss distribution which will model the severity of claims would enable insurance claims managers to have a good knowledge of claims data.

5. Mean Severity under Exponential and Log-normal Distributions

5.1 Data Analysis

In general insurance practice, *data on deductibles* is usually unavailable because they are claims which are only borne by individual scheme holder and because of the confidentiality of insurance data base. Instead of the raw deductible data, we obtained rate relativity on deductible through a non-life insurance *agent* operating in property insurance market in Lagos. In order to present logical arguments, we solve the following standard empirical illustration contained in literature of Tse (2009) by considering an insured risk Y with unit sum insured under insurance cover with specified deductibles D under exponential distributions $0.1 \le D \le 1, Y \sim EXP(\alpha), \alpha = 1$ and severity when log-normally distributed as $Y \sim LN(\mu, \sigma^2), assume, \mu = -\frac{1}{2}, \sigma^2 = 1$

5.2 Exponential Distribution

$$S_Y(D) = e^{-\alpha D}, g_Z(D) = \alpha e^{-\alpha D}, H_Y(y) = \frac{g_Y(D)}{S_Y(D)} \tag{74}$$

$$\langle Y_L \rangle = \langle (y - 0.15)_+ \rangle = \int_{0.15}^{\infty} e^{-y} dy = e^{-0.15} = 0.86071, \tag{75}$$

$$S_Y(D) = e^{-0.15} = 0.86071 \Rightarrow \frac{\langle Y_L \rangle}{S_Y(D)} = \langle Y_P \rangle_{\exp} = \int_0^{\infty} e^{-y} dy = 1 \tag{76}$$

$\langle Y_p \rangle = \int_0^{\infty} g_Y(y) dy = 1$, hence, we can see that $\langle Y_L \rangle < \langle Y_p \rangle$ that is the cost per loss amount is less than the cost per payment amount. The loss eliminated (*LE*) and loss elimination ratio (*LER*) are as given below

$$LE(y) = \frac{1}{\alpha} - \langle Y_L \rangle = 1 - \langle Y_L \rangle, LER(y) = \frac{\langle Y \rangle - \langle Y_L \rangle}{\langle Y \rangle} \tag{77}$$

Table 1: Computed Values of D and LER for Exponentially Distributed Claim

DEDUCTIBLE DOMAIN $0.1 \leq D \leq 1$	COST PER LOSS $\langle Y_L \rangle$	LR $1-\langle Y_L \rangle$	LER	CHANGE IN LER
0.1	0.904837	0.095163	0.095163	0.0952
0.15	0.860708	0.139292	0.139292	0.0441
0.2	0.818731	0.181269	0.181269	0.042
0.25	0.778801	0.221199	0.221199	0.0399
0.3	0.740818	0.259182	0.259182	0.038
0.35	0.704688	0.295312	0.295312	0.0361
0.4	0.67032	0.32968	0.32968	0.0344
0.45	0.637628	0.362372	0.362372	0.0327
0.5	0.606531	0.393469	0.393469	0.0311
0.55	0.57695	0.42305	0.42305	0.0296
0.6	0.548812	0.451188	0.451188	0.0281
0.65	0.522046	0.477954	0.477954	0.0268
0.7	0.496585	0.503415	0.503415	0.0255
0.75	0.472367	0.527633	0.527633	0.0242
0.8	0.449329	0.550671	0.550671	0.023
0.85	0.427415	0.572585	0.572585	0.0219
0.9	0.40657	0.59343	0.59343	0.0208
0.95	0.386741	0.613259	0.613259	0.0198
1	0.367879	0.632121	0.632121	0.0189

Source: Author's Computation

5.3 Lognormal Distribution

$$\langle Y_L \rangle = \int_D^\infty (y-D) g_Y(y)\,dy = \int_D^\infty y g_Y(y)\,dy - \int_D^\infty D g_Y(y)\,dy =$$

$$\int_D^\infty y g_Y(y)\,dy - D S_Y(D) \tag{78}$$

$$\int_D^\infty y g_Y(y)\,dy = \int_D^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left[(\log_e y-\mu)^2\right]}{2\sigma^2}}\,dy \cong e^{\left[\frac{\sigma^2}{2}+\mu\right]\left\{1-\Phi\left(\frac{\ln D-\mu-\sigma^2}{\sigma}\right)\right\}} \tag{79}$$

$$S_Y(D) = \Pr\left[Z > \frac{\log_e y-\mu}{\sigma}\right] \tag{80}$$

TABLE 2: Computed Values of D and Log-normally Distributed Cost Per Loss

1	2	3	4	5	6	7	8	9
0.1	-2.302585093	-1.802585093	-1.802585093	-2.802585093	0.99746	0.00254	0.9642	0.90104
0.15	-1.897119985	-1.397119985	-1.397119985	-2.397119985	0.99176	0.00824	0.9188	0.85394
0.2	-1.609437912	-1.109437912	-1.109437912	-2.109437912	0.9825	0.0175	0.8661	0.80928
0.25	-1.386294361	-0.886294361	-0.886294361	-1.886294361	0.9703	0.0297	0.8122	0.76725
0.3	-1.203972804	-0.703972804	-0.703972804	-1.703972804	0.9558	0.0442	0.7592	0.72804
0.35	-1.049822124	-0.549822124	-0.549822124	-1.549822124	0.9393	0.0607	0.7085	0.691325
0.4	-0.916290732	-0.416290732	-0.416290732	-1.416290732	0.9215	0.0785	0.6613	0.65698
0.45	-0.798507696	-0.298507696	-0.298507696	-1.298507696	0.903	0.097	0.6172	0.62526
0.5	-0.693147181	-0.193147181	-0.193147181	-1.193147181	0.8836	0.1164	0.5765	0.59535
0.55	-0.597837001	-0.097837001	-0.097837001	-1.097837001	0.864	0.136	0.5391	0.567495
0.6	-0.510825624	-0.010825624	-0.010825624	-1.010825624	0.844	0.156	0.5044	0.54136
0.65	-0.430782916	0.069217084	0.069217084	-0.930782916	0.8241	0.1759	0.4723	0.517105
0.7	-0.356674944	0.143325056	0.143325056	-0.856674944	0.8042	0.1958	0.4431	0.49403
0.75	-0.287682072	0.212317928	0.212317928	-0.787682072	0.7847	0.2153	0.416	0.4727
0.8	-0.223143551	0.276856449	0.276856449	-0.723143551	0.7651	0.2349	0.3909	0.45238
0.85	-0.162518929	0.337481071	0.337481071	-0.662518929	0.7464	0.2536	0.3681	0.433515
0.9	-0.105360516	0.394639484	0.394639484	-0.605360516	0.7273	0.2727	0.3464	0.41554
0.95	-0.051293294	0.448706706	0.448706706	-0.551293294	0.7091	0.2909	0.3268	0.39864
1	0	0.5	0.5	-0.5	0.6915	0.3085	0.3085	0.383

Source: Author's Computation

$$COLUMN\ 1 = 0.1 \le D \le 1; COLUMN\ 2 = \log_e D; COLUMN3\ = \log_e D - \mu;$$

$$COLUMN4\ = \frac{(\log_e D - \mu)}{\sigma} \tag{81}$$

$$COLUMN\ 5 = \frac{(\log_e D - \mu)}{\sigma} - \sigma; COLUMN6\ = \int_D^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left[(\log_e y - \mu)^2\right]}{2\sigma^2}} dy \tag{82}$$

$$COLUMN7\ = \Phi\left(\frac{(\log_e D - \mu)}{\sigma} - \sigma\right); COLUMN8\ = \Pr\left(Z > \frac{(\log_e D - \mu)}{\sigma}\right) \tag{83}$$

$$COLUMN9\ = \langle (Y_L) \rangle; \langle Y_P \rangle_{\log-normal} = \frac{\langle Y_L \rangle}{S_Y(D)} \tag{84}$$

From the definition of exponentially distributed loss, we can see that

$$\langle Y_L \rangle = \langle (y - 0.15)_+ \rangle = \int_{0.15}^{\infty} e^{-y} dy = e^{-0.15} = 0.86071 \tag{85}$$

In order to see the trend of the change in the loss eliminated in the domain for D over which it has been defined, it was revealed that $\langle Y_L \rangle < \langle Y_p \rangle$, that is the cost per loss amount is less than the cost per payment amount.

6 . Discussion of Results

In table 1 above, the ratio of the loss eliminated is roughly directly proportional to the deductibles. Using the last column of the table, and beyond a certain level, it seems high deductibles would not provide a reasonable fraction of the eliminated loss due to the underwriting firm. The considerations for deductible would be quite different depending on policy requirements, terms and on the risk preference of the scheme holder. The underwriter may choose to apply Table 1 and use it as a guide to confirm if the deductible proposed say by the scheme holder offers reasonable fraction of the losses eliminated to the underwriter. We observe that the cost per payment $|\langle Y_p \rangle| = 1$ is uniform throughout the domain of definition irrespective of the value of the deductible. However, Table 2 is a bit complex to compute due to the nature of lognormal distribution but shows a systematic process leading to the computation of cost per loss payment $\langle Y_L \rangle$ as shown in column 9. The rate relativity $0.1 \le D \le 1$ in column 1 forms the basis of computing the cost per loss and because its values are within 0 and 1, the logarithmic values are less than zero except at $D = 1$. A simple calculation of cost per payment $\langle Y_p \rangle_{\log normal}$ from Table 2 reveals a progressive increase from 0.934495 to 1.241491. While $|\langle Y_p \rangle| < 1, for, 0.1 \le D < 0.4$ $|\langle Y_p \rangle| > 1, for, 0.45 \le D < 1$ and consequently the insurer

In table 1 above, the ratio of the loss eliminated is roughly directly proportional to the deductibles. Using the last column of the table, and beyond a certain level, it seems high deductibles would not provide a reasonable fraction of the eliminated loss due to the underwriting firm. The considerations for deductible would be quite different depending on policy requirements, terms and on the risk preference of the scheme holder. The underwriter may choose to apply Table 1 and use it as a guide to confirm if the deductible proposed say by the scheme holder offers reasonable fraction of the losses eliminated to the underwriter. We observe that the cost per payment $\left|\left\langle Y_p \right\rangle\right| = 1$ is uniform throughout the domain of definition irrespective of the value of the deductible. However, Table 2 is a bit complex to compute due to the nature of lognormal distribution but shows a systematic process leading to the computation of cost per loss payment $\left\langle Y_L \right\rangle$ as shown in column 9. The rate relativity $0.1 \le D \le 1$ of D in column 1 forms the basis of computing the cost per loss and because its values are within 0 and 1, the logarithmic values are less than zero except at D $D = 1$. A simple calculation of cost per payment $\left\langle Y_p \right\rangle_{\log normal}$ from Table 2 reveals a progressive increase from 0.934495 to 1.241491. While $\left|\left\langle Y_p \right\rangle\right| < 1, for, 0.1 \le D < 0.4$ $\left|\left\langle Y_p \right\rangle\right| > 1, for, 0.45 \le D < 1$ and consequently the insurer experiences higher cost per payment than expected. The *underwriter is therefore advised to apply deductible in the second subdomain* to reduce the number and magnitude of nuisance claims advised and discourage moral hazards. Furthermore, the insurance firm should be cautious as high deductible is not financially ethical to the insured because the insured would bear a higher percentage of the loss burden arising from the insured peril. The insurer may impose high deductibles because of the underwriting cost saving conditions and loss control motivation mechanisms. Irrespective of the deductible type, the intent is to make the scheme holder very cautious since he would pay a proportion of the total loss. Provided that a large number of losses is lower than the deductible, the administrative costs incurred by the underwriter to offset liabilities and maintain solvency will *decrease*, hence the premium payable by the policy holder will decline compared to full coverage conditions. Underwriters working in distinct lines of business consider multiple views of severities in analyzing short-tailed lines of business. Comparing mean values in Tables 1 and 2 above, it is observed that despite lognormal severity distribution has a thicker tail than the exponential distribution, its mean loss $\left\langle Y_L \right\rangle$ column is observed to be correspondingly lower in value than the values of exponential mean loss, that is, $\left\langle Y_L \right\rangle_{\log - normal} < \left\langle Y_L \right\rangle_{\exp onential}$. We should be reminded that the accuracy of the results depends very much on rate relativity collected for the analysis.

7. Conclusion

We have established a clear relationship between severity and deductible which enables us to use the extremely well known densities to deal with average loss functions. The idea of severity-deductible chains of proper actuarial risk function plays a central role in the analysis of insurance coverage. Therefore, given an insured risk of magnitude z, an insurance cover will pay a sum which is dependent on z and which function we could present as $h(z)$. The expected payment per loss will then be $\int_0^\infty h(z)dF_Z(z)$ while the number of losses assumed by a risk within the policy anniversary will be given by $m(e)$, where e is the exposure such that the aggregate loss payment is given by $\sum_{k=1}^{m(e)} h(z_j)$. The results arrived at on insurance claims size has meaningful application in general business insurance underwriting operations. A high precision estimation of severity of insurance claims and well-matched cash in-flow and cash out-flows is an advantage for the underwriter to meet up with solvency conditions as claims are advised and indemnified.

Specifically, the paper focuses on comparing the mean severities under exponentially and log-normally distributed arbitrary policy in a cost per loss and cost per payment events with deductible clauses. When a reliable empirical data on deductible is not available, it is not unreasonable to apply a theoretical loss distribution such as log-normal and exponential. In presenting the models as insurance statement of mean severity problem, the choice of the known severity distribution was done on the grounds of established analytical characterization of the distributions such as moments and value of distribution function in order to ease out calculation of severity charges which, by extension, may furthermore permit us to obtain expected value risk premium. The underlying assumption results in the observation that in computing distribution of severities under deductible theoretic regime, it is necessary to estimate the severity theoretic distribution by use of lognormal and exponential model. From the actuarial point of view, lognormal and exponential distribution permit us to modify the moment $\langle Y_L \rangle$ of the approximating distribution to fall in line with loss theoretic distribution as required. Bearing that it is necessary for underwriters to compute adequate premium, it is just as sufficient to compute the mean severities analytically or numerically. These results assist in obtaining the appropriate premium for the insured. The underwriter should match the premium income with the claims outgo while focusing on profitability but in the event of a parallel between the two parameters, with the premium received not adequate to cover the severity, the underwriter is faced with the probability of loss. An underwriter's operating income relies on how far it has understood the peril being insured and how it can bring down the expenses relating to claims management. The amount of premium an insurance firm charges for granting cover is a critical part of underwriting system. The premium as a function of losses, loss adjustment expenses and underwriting profit need to be adequate to cover the expected claims and must consider the probability that the underwriter may have to access its solvency requirements. The premium chargeable for insurance coverage depends therefore on the deductible level chosen that is obtained from the insurance coverage modifications. The paper recommends that the distribution models adopted will encourage underwriters to embrace risk management regulations through allocation of adequate capital and honouring liability obligations. Furthermore, constant re-training of technical staff on actuarial models of insurance losses when deductibles are integrated need to be facilitated and must be well informed on current market methodologies to solve issues relating to risk management changes and technologies. The technical knowledge enables the insurer to make informed financial decisions to ensure significant profitability. We need to mention that the insurance system software on insurance loss platform for incorporating core risk indicators together with claims reported such as cause of loss, premium and deductible could be spelt out based on the type of the business environment. This is meant to improve the efficiency of risk management process through actuarial modelling techniques.

References

1. Afify, A.Z., Gemeay, A.M. & Ibrahim, N.A. (2020). The heavy tailed exponential distribution risk measures estimation and application to actuarial data. *Mathematics,* 8, 1276, 1-28
2. Bakar, A.S., Hamzah, N.A., Maghsoudi, M. & Nadarajah, S. (2015). Modelling loss data using composite models. *Insurance: Mathematics and Economics*. 61(2015), 146-154.
3. Klugman, S.A., Panjer, H.H. & Willmot, G.E. (2004). *Loss models from data to decisions* (2nd ed) John Wiley & Inc. USA. 626-669
4. Liu, H. & Wang, R. (2017). Collective risk models with dependence uncertainty. *ASTIN Bulletin: The Journal of The IAA,* 47(2), 361–389.
5. Ogungbenle, M. G., Adeyele, J.S. & Mesioye, A.E. (2020). Modeling moments of insurance claim size under dirac-delta function. *Daffodil International University Journal of Science and Technology,* 15(1), 43-51
6. Ogungbenle, M.G. (2021). A theoretical Cauchy-Bunyakovsky-Schwartz inequality estimation of upper bound for the expected claim per Loss payment in non-life insurance business. *Journal of New Frontiers in Mathematics and Statistics, 2*(2), 1-19.
7. Pacakova, V. & Brebera, D. (2015). Loss distributions and simulations in general insurance and re- insurance. *International Journal of Mathematics and Computers in Simulation. 9,* 159-167.
8. Raschke, M. (2019). *Alternative modelling and inference methods for claim size distributions,* 1-24.
9. Sakthivel, K.M. & Rajitha, C.S. (2017). Artificial intelligence for estimation of future claim frequency in non-life insurance. *Global Journal of Pure and Applied Mathematics* 13(6), 1701-1710.
10. Schlesinger, H. (1985). Choosing a deductible for insurance contracts: Best or worst insurance policy?", *Journal of Risk and Insurance. 52*(3), 522-527.
11. Thogersen, J. (2016). Optimal premium as a function of the deductible: *Customer analysis and portfolio characteristics. Risks.* 4, 1-19.
12. Tse, Y.K. (2009). *Non-life actuarial models, Cambridge University Press,* U. K, 66-75.
13. Woodard, J.D. & Yi, J. (2020). Estimation of insurance deductible demand under endogenous premium rates". *Journal of Risk & Insurance, The American Risk and Insurance Association* 87(2), 477-500.
14. Zacaj, O., Dhamo, E. & Shehu, S. (2015). Modelling individual claims for motor third party liability of insurance companies in Albania. *International Journal of innovation in science and mathematics. 3*(3), 174-178.