Underwriting Factor for Determination of Health Insurance Premium Based on Generalized Linear Models (GLMS) Procedure

Riaman, Sukono
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia
riaman@unpad.ac.id; sukono@unpad.ac.id

Abdul Talib Bon
Department of Production and Operations, University Tun Hussein Onn Malaysia, Malaysia
talibon@gmail.com

ABSTRACT

The increasing cost of treatment in hospitals has made many people worry about the financial risks they face. One way to reduce this risk is through a health insurance program. Each customer must bind himself to an insurance company as a risk guarantor by paying money, which is called a premium. The amount of premium must be calculated correctly, so that it does not harm the customer or the insurance company. Therefore, there are several factors that need to be considered in determining the premium, such as age, occupation, medical history, lifestyle, and other factors determined by the company. These factors are called underwriting factors. In this study, Generalized Linear Models were used to determine the claims and the number of claims. Variables that significantly influence of claims are having a history of allergies, having a history of hereditary diseases, class of health facilities, and length of stay. While the variables that have a significant effect on many claims are occupation, smoking, having a history of allergies, and consuming alcohol. These two models are used to calculate the premium price from a customer using the compound model principle.

Keywords:
Health insurance, premiums, underwriting factors, GLMs

1. Introduction

During life, humans are always faced with various risks. Risk can occur anytime and anywhere, such as damage, accident, illness, or even death. This makes people worry about the welfare of their lives, considering the cost of necessities is increasing day by day. Therefore, to reduce financial risk, a person needs protection or something that can ensure the welfare of himself and his family. One way to do this is to take insurance (Sirait et al., 2020). Insurance is an agreement between the insured (customer) and the insurer (insurance) which aims to reduce the risk that may be experienced by the insured (Klafin et al., 2020b; Kalfin et al., 2021). There are several types of insurance products, including life insurance, general insurance, health insurance, reinsurance, and many more. One of the popular insurances in Indonesia is health insurance. According to Investopedia (2019), health insurance is a type of insurance coverage that pays for medical, surgical, or related health care costs as stated in the insurance policy.

Everyone who participates in insurance must pay a certain amount of money periodically to the insurer as a form of compensation for the transfer of risk that may be experienced by the insured at any time. The amount of premium paid is adjusted by taking into account various conditions/conditions found in the insured or what is called the underwriting factor. Insurance companies must calculate premiums appropriately, considering that premiums are one source of income from insurance companies. Many methods can be used to calculate premiums, one of which is Generalized Linear Models (GLMs).

GLMs are an extension of regression analysis where this statistical method is used to solve the problem of response variables that are no longer continuous but categorical by using certain logit connecting functions so that a model is obtained that is able to analyze the relationship between categorical responses and one or more explanatory variables. In this study, GLMs were used to obtain large claims model and multiple claims model. The premium calculation is done by using the compound model, where the net premium is obtained by multiplying the expectation value of the large claims model and the expected value of many claims.

Previously, research on the calculation of premiums with GLMs has been carried out by several researchers, such as David (2015) who discussed the calculation of premiums on vehicle insurance. In David's research (2015), it was concluded that premiums can be determined by considering the existing risk factors and the results of the research are useful representatives for the insurance company's business. Supriatna, et al (2017) discuss the prediction of premium rates on health insurance with many claims used only 0 and 1. In this study, a forward selection procedure was used to determine the covariates that affect the response variable, then a large claim model and many claims were formed. the basis for calculating premiums. Widianti (2018) discusses determining the net premium price for life insurance by taking into account the underwriting factor. In this study, it is necessary to do mortality modeling first, then a mortality table is formed to be able to calculate the amount of the premium.

In this study, the authors are interested in discussing the determination of inpatient health insurance premiums based on underwriting factors using GLMs. If in Suptiatna et al's research (2017) the regression procedure used is forward selection, then in this study the procedure used is backward elimination, the number of claims used is more than one, and the compound model is used for premium calculations.

The purpose of this study is to determine the variables that affect the number of claims and the amount of claims, estimate the model parameters, and determine the pure premium using the compound model.

2. Methods

Underwriting is the process of identifying and selecting risks, classifying risk levels, and making decisions regarding prospective insurance customers. The underwriting process is an effort so that prospective customers get justice in determining the amount of premiums. In insurance, the underwriting process is carried out by assessing risk based on age, occupation, medical history, lifestyle, and other factors determined by the insurance company. (Allianz Indonesia, 2020).

Meanwhile, Generalized Linear Models (GLMs) are an extension of the linear regression model with the assumption that the predictor has a linear effect but does not assume a certain distribution of the response variable and is used when the response variable is a member of the exponential family (Nelder and Wedderburn, 1972). GLMs aim to determine the effect of the explanatory variable on the response variable. The advantage of GLMs lies in the distribution of the response variables. In this method, the response variable does not have to be normally distributed, but a distribution that belongs to the exponential family, namely: binomial, poisson, negative binomial, normal, gamma, inverse gaussian.

For the response variable y, the GLMs are:

$$f(y) = \exp\left\{\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)\right\} \tag{1}$$

$$g(\mu) = x'\beta$$

where,

$f(y)$: probability function of response variable y with exponential family distribution

$g(\mu)$: link function, determines how the mean is related to the variable x.

According to McCullagh and Nelder (1989), GLMs consist of 3 components, namely:

1. Random components: components of Y that are independent of each other. The response variable is assumed to come from the exponential family.
2. Systematic component: covariates $x_1, x_2, \dots, x_p$ which gives the linear predictor η, where
$$\eta = \sum_{1}^{p} x_j\beta_j = \beta_0 + X_1\beta_1 + \cdots + X_p\beta_p$$
3. The link function g(.), describes the relationship of the linear predictor η with the mean μ. This relationship can be written as $\eta = g(\mu)$.

The general form of the probability function of the response variable Y in GLMs is given in equation (1), where θ is the canonical parameter and ϕ is the dispersion parameter. The probability function given in equation (1) is part of the exponential family.
Selecting the functions $b(\theta)$ and $c(y, \phi)$ determines the actual probability function such as binomial, normal, or gamma. Condition $b(\theta)$,

$$E(y) = \dot{b}(\theta)$$
$$Var(y) = \phi\, \ddot{b}(\theta)$$

where $\dot{b}(\theta)$ and $\ddot{b}(\theta)$ is the first and second derivative of $b(\theta)$.

$$\ddot{b}(\theta) = \frac{\partial \dot{b}(\theta)}{\partial \theta} = \frac{\partial \mu}{\partial \theta} = V(\mu)$$

so that equation (4) can be written as:

$$Var(y) = \phi\, V(\mu) \text{ atau } V(\mu) = \frac{Var(y)}{\phi}$$

The *link* function is a function that connects the linear predictor η with the expected value of the y response, namely μ. Each distribution has a special link function that has a statistical similarity β in the linear predictor $\eta = \sum x_j \beta_j$. The canonical link exists when $\theta = \eta$, where θ is the canonical parameter.
Modeling using GLMs will involve two types of variables, namely response variables and explanatory variables. Before performing the regression, the explanatory variables that are categorical need to be selected based on the base level. Base level is the basis of measurement in the regression that will compare the base level with other level categories in a variable. The base level is selected based on the standard of the largest number of observations (not sparse) at a certain level in a variable. Regression using categorical variables will be related to dummy variables. The dummy variable is a dummy variable whose number is $j - 1$ with j being the number of levels in the response variable.

The method used to estimate the parameters in this study is the Maximum Likelihood Estimator (MLE) method. The first step that needs to be done in this method is to determine the likelihood function. Likelihood function is defined as Equation (2)

$$L(y; \theta, \phi) = \prod_{i=1}^{n} f(y_i; \theta, \phi). \tag{2}$$

To simplify calculations, Equation (2) is transformed into the form natural logarithm or called log-likelihood, defined as:

$$\ln\big(L(y; \theta, \phi)\big) = \ell(\theta, \phi) = \sum_{i=1}^{n} \ln f(y_i; \theta, \phi). \tag{3}$$

Then, by substituting Equation (1) into Equation (3), we get

$$\ell(\beta, \phi) = \sum_{i=1}^{n} \ln f(y_i; \beta, \phi) = \sum_{i=1}^{n} \left\{ \frac{y_i \theta_i - b(\theta_i)}{\phi} + c(y, \phi) \right\} \tag{4}$$

assuming the variables y_i are independent. Then reduce Equation (4) to β_j, the result is obtained as in Equation (5).

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ell}{\partial \theta_i} \frac{\partial \theta_i}{\partial \beta_j} \tag{5}$$

where

$$\frac{\partial \ell}{\partial \theta_i} = \frac{y_i - \dot{a}(\theta_i)}{\phi} = \frac{y_i - \mu_i}{\phi}, \qquad \frac{\partial \theta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \eta_i} x_{ij} \tag{6}$$

In this calculation, $\eta_i = x_i' \beta$ are the i components of x_j. Next, maximize the likelihood function by decreasing the log-likelihood function $\frac{\partial \ell}{\partial \beta_j} = 0$, so that Equation (7).

$$\sum_{i=1}^{n} \frac{\partial \theta_i}{\partial \eta_i} x_{ij} (y_i - \mu_i) = 0 \qquad \Leftrightarrow \qquad \mathrm{X}^{\mathrm{T}} \mathrm{D}(y - \mu) = 0 \tag{7}$$

Equation (7) is difficult to solve directly. Therefore, the solution can be done numerically using Newton-Raphson (Jong, 2008).

Kolmogorov Smirnov test. According to Tse, (2009), the Kolmogorov Smirnov test is a test used to see the fit of the model on continuous random variables. The hypothesis test used on Kolmogorov Smirnov is.
H_0 : data from distribution F^*
H_1 : data does not come from the F^*
The critical point of the D test statistic for the value $\alpha = 0,05$ is $\frac{1,36}{\sqrt{n}}$. The value of the test statistic is less than the critical point, then the selected model can be said to be close to the true distribution.

Chi-Square Test, according Tse (2009), *Chi-Square test is used for grouped data. Initial hypothesis test on Chi-Square is*
H_0 : data comes from a particular distribution
H_1 : data does not come from the distribution specified in H_0
To test the initial hypothesis used

$$\chi^2 = \sum_{j=1}^{k} \frac{(e_j - n_j)^2}{e_j} = \left(\sum_{j=1}^{k} \frac{n_j^2}{e_j} - n \right) - n \quad (8)$$

H_0 was rejected when $\chi^2 > \chi^2_{\alpha,k-1-p}$ where k is the number of classes and p is the number of parameters to be estimated.

Parameter Significance Test, in GLMs regression analysis is used to determine the covariates that have a significant effect on the response variable. In this study, the technique used was backward elimination. After determining the significant covariates, further testing is carried out. The parameter significance test that is often used is the Likelihood Ratio (LR) test and the Wald test.
Likelihood Ratio (LR), according to Jong (2008), the likelihood ratio is defined as $\lambda = \frac{\hat{L}}{\tilde{L}}$ where $\hat{L}$and $\tilde{L}$ are the likelihoods of the unrestricted and restricted models. LR test statistic is

$$2 \ln \lambda = 2(\hat{\ell} - \tilde{\ell}) \quad (9)$$

with
dengan
$\hat{\ell}$: likelihood function of the model in which all explanatory variables are involved (unrestricted model)
$\tilde{\ell}$: the likelihood function of a model with an explanatory variable of $p - q, p$ is the number of parameters in the model and q is the number of parameters considered 0 (restricted model)
The critical point of the test statistic is $\chi^2_{\alpha,q}$ so H_0 **is rejected if** $2 \ln \lambda$ is greater than the tipping point.

Wald test, Wald's test can only test the level of significance of one explanatory variable. The hypothesis used in this test is
H_o: $\beta_j = r$
H_1: $\beta_j \neq r$
for $j = 1, 2, 3, \ldots.$
The test statistic used in the Wald test is

$$\frac{(\widehat{\beta_j} - r)^2}{\phi \psi_j} \quad (10)$$

where ψ_j is the j^{th} element of the diagonal matrix $(X'WX)^{-1}$ and the value of r used to test is equal to 0. H_0 is rejected when the value of the test statistic is greater than $\chi^2_{\alpha,1}$ (Jong, 2008).

Diagnostic, the test to test the feasibility of the model is called the goodness of fit test. In GLMs, the deviance (Δ) is usually used. According to Jong (2008), deviance is defined as a measure of the distance between the fitted model and the saturated model. The hypothesis test used is
H_0: Model is matched
H_1: Model does not match
The test statistic of Δ is

$$\Delta \equiv 2(\breve{\ell} - \hat{\ell}) = 2 \sum_{i=1}^{n} \{ \frac{y_i(\breve{\theta}_i - \widehat{\theta}_i) - a(\breve{\theta}_i) + a(\widehat{\theta}_i)}{\phi} \quad (11)$$

where
$\breve{\ell}$: is the log-likelihood function of the saturated model
$\hat{\ell}$: is the log-likelihood function derived from the fitted model
The critical point of the deviance test is χ^2_{n-p}. If the deviance value is greater than the critical point then reject H_0.

Compound Model,
Let N be a random variable representing the number of claims in a certain period of time. Let X be a random variable that expresses the claim size. The total claims for the collective risk model are:

$$S = X_1 + X_2 + \cdots . + X_N, \qquad N = 0,1,2, \ldots$$

where $S = 0$ when $N = 0$.
According to (Klugman, 1998; Kalfin et al., 2020a), there are three basic assumptions that must be considered, namely:

1. Given that $N, X_1, X_2, \ldots, X_n$ is an independent random variable with identical distribution.
2. Given $N, X_1, X_2, \ldots, X_n$ is a random variable that does not depend on N.
3. The distribution of N does not depend on the values $X_1, X_2, \ldots$..

To get the premium formula based on the total expected value of the claim, assume that $E(X), E(N), Var(X)$, and $Var(N)$ exist. The formula that can be used to calculate pure risk premium is

$$E(S) = E(N)E(X) \tag{12}$$

3. Data Collection

The data used in this study is health insurance survey data. The data obtained consists of two parts, namely membership data and claim data. The membership data describes the characteristics of the participants which consist of: age, gender, marital status, occupation, smoker, drinking alcohol, and medical history. Meanwhile, claim data contains information about the insurance benefits received by participants, such as health facilities, diagnosis, length of treatment, claim size, and many claims.

4. Results and Discussion

The data used to determine the claim size model consists of the claim size variable as the response variable and 11 other explanatory variables. The explanatory variables consist of gender, age, marital status, occupation, smoking, alcohol, allergies, hereditary diseases, class of health facilities, and ICD. To determine the distribution that matches the big data claims, SAS software is used. Kolmogorov Smirnov test results show that the data is gamma distributed with the log link function.
Furthermore, the selection of covariates that affect the response variable is carried out using backward elimination. The initial step in the backward elimination method is to enter all explanatory variables into the model. Then the insignificant variables are removed one by one starting from the variable that has $p - value > \alpha$. Determination of covariates was carried out using SAS software with the Proc Genmod command. Based on the final results of the LR test, it is known that the variables that are significant to the size of the claim are history of allergies, hereditary diseases, class of health facilities, and length of stay.
After obtaining the best model, parameter estimation was carried out with MLE, the results are presented in Table 1.

Table 1. Analysis of the estimated parameters of the large claims model

Parameter		*DF*	*Estimate*	*S.E*	*Wald Chi-Square*	*Pr > ChiSq*
Intercept		1	14.2293	0.0477	89111.7	< 0.0001
Alergi	1	1	0.1089	0.0432	6.34	0.0118
Penyakit	1	1	−0.1130	0.0430	6,91	0.0086
Kelas	1	1	0.4249	0.1171	13.16	0.0003
Kelas	2	1	−0.2710	0.0422	41.27	< 0.0001
Kelas	3	1	−0.3707	0.0817	20.59	< 0.0001
LM		1	0.1492	0.0064	598.49	< 0.0001

From Table 1 it can be seen that all variables are significant. So, the large claims model can be written as Equation (13).

$$\ln \hat{\mu} = 14.2293 + 0.1089\beta_{Ale1} - 0.1130\beta_{P1} + 0.4249\beta_{K1} - 0.2710\beta_{K2} - 0.3707\beta_{K3} + 0.1492\beta_{LM} \tag{13}$$

where
$\hat{\boldsymbol{\mu}}$: claim expectation value

x_{Ale1} : allergy variable
x_{p1} : hereditary disease variables
x_{K1} : VIP class variable
x_{K2} : class variable 2
x_{K3} : class variable 3
x_{LM} : variable length of stay

Furthermore, the model obtained was tested using the residual deviance test. Figure 1 is a residual deviance plot of the large claim model that has been obtained.

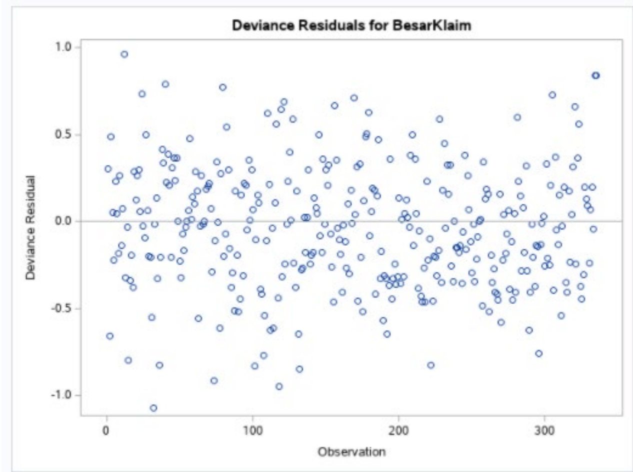

Fig 1. *Plot* deviance residual model for number of claim.

From Figure 1, it can be seen that the residual deviance value is in the interval [-1, 1], meaning that all observations contribute to the fit of the model. Deviance value obtained from the model is 329.5229. While the critical point is 372.2991. Deviance value is smaller than the critical point so it can be said that the model fits the data.

Claims Model

The data used to determine the multiple claims model consists of multiple claims as response variables and 8 explanatory variables. The explanatory variables consist of gender, age, marital status, occupation, smoking, alcohol, allergies, and inherited diseases. The mean value generated from the multiple claims data is 0.9571429 while the variance value is 2.0755219. The value of variance is greater than the mean value, so the multiple claims variable is assumed to have a negative binomial distribution. After the compatibility test, the values $\boldsymbol{p-value = 0,05530 > \alpha}$ dan $\boldsymbol{\chi^2 = 15,20376 < \chi^2_{0,05,8} = 15,5073}$, so it can be concluded that the multiple claims variable follows a negative binomial distribution with the log link function.

Furthermore, the selection of covariates that have a significant effect on the response variable is carried out using backward elimination. Based on the final results of the LR test, it is known that the variables that are significant to the claim are occupation, smoking, allergies, and alcohol. After obtaining the best model, parameter estimation was carried out using MLE, the results are presented in Table 2.

Table 2. Analysis of the multi-claims model parameter estimation

Parameters		*DF*	*Estimate*	*S.E*	*Wald Chi-Square*	*Pr > ChiSq*
Intercept		1	−0.5116	0.1705	9.00	0.0027
Occupation	1	1	−0.3124	0.4758	0.43	0.5115
Occupation	2	1	0.7527	0.3371	4.99	0.0256
Occupation	3	1	0.4757	0.2564	3.44	0.0636
Occupation	4	1	0.2859	0.2120	1.82	0.1773
Occupation	5	1	0.5889	0.2176	7.33	0.0068
Smoking	1	1	−0.5200	0.2249	5.35	0.0207
Allergy	1	1	0.5302	0.1709	9.63	0.0019
Alcohol	1	1	1.8455	0.3982	21.48	< 0.0001

From Table 2, it is known that the work variables level 2 and 5, smoking, allergies, and alcohol which have a significant effect. Thus the multiple claims model can be written as Equation (14).

$$\ln \hat{\mu} = -0.5116 + 0.7527x_{Pk2} + 0.5889x_{Pk5} - 0.5200x_{M1} + 0.5302x_{Ale1} + 1.8455x_{Alk1} \quad (14)$$

where

$\hat{\mu}$: the expected value of many claims
x_{Pk2} : housewife work variable
x_{Pk5} : other job variables
x_{M1} : smoking variable
x_{Ale1} : allergy history variable
x_{Alk1} : alcohol consumption variable

Furthermore, the model obtained is tested using the deviance test. Deviance value obtained from the model is 344.7315. While the critical point value is 389.3136. Deviance value is smaller than the critical point so that it can be said that the model fits the data.

Premium Calculation using Compound Model

Calculation of the amount of health insurance premiums using case studies. Suppose an insurance customer is sick with the following criteria:

- 27 years old with unmarried status
- Treatment duration is 2 days
- Suffering from typhoid
- Work as a private employee
- The health facilities used are class 2
- Have a history of allergies and hereditary diseases
- Smoking but not consuming alcoholic beverages

Equation (13) is used to calculate the claim size, it is obtained

$$\ln \hat{\mu} = 14.2293 + 0.1089\beta_{Ale1} - 0.1130\beta_{P1} + 0.4249\beta_{K1} - 0.2710\beta_{K2} - 0.3707\beta_{K3} + 0.1492\beta_{LM}$$
$$\ln \hat{\mu} = 14.2293 + 0.1089(1) - 0.1130(1) + 0.4249(0) - 0.2710(1) - 0.3707(0) + 0.1492(2)$$
$$\ln \hat{\mu} = 14.2526$$
$$\hat{\mu} = 1{,}548{,}194.54$$

Untuk menghitung banyak klaim digunakan Persamaan (14), diperoleh

$$\ln \hat{\mu} = -0.5116 + 0.7527x_{Pk2} + 0.5889x_{Pk5} - 0.5200x_{M1} + 0.5302x_{Ale1} + 1.8455x_{Alk1}$$
$$\ln \hat{\mu} = -0.5116 + 0.7527(0) + 0.5889(0) - 0.5200(1) + 0.5302(1) + 1.8455(0)$$
$$\ln \hat{\mu} = -0.5014$$
$$\hat{\mu} = 0.6057$$

Sehingga didapat $E(X) = 1{,}548{,}194.54$ dan $E(N) = 0.6057$. Besar premi dapat dihitung menggunakan model *compound* mengacu pada Persamaan (12), seperti berikut ini:

$$E(S) = E(X)E(N)$$
$$= 1.548.194{,}54 \times 0{,}6057$$
$$= 937{,}741.44$$

5. Conclusion

Based on the results of the analysis and discussion, the variables that have a significant effect on the size of the claim are allergies, hereditary diseases, class of health facilities, and length of stay. While the variables that have a significant effect on many claims are the level 2 and 5 occupation variables, smoking, having a history of allergies, and consuming alcohol. The parameter estimation results for the large claims model are in Table 1 and the parameter estimates for the multiple claims model are in Table 2. The amount of premium that must be paid by customers who work as private employees, have a history of allergies and hereditary diseases, and smokers is Rp937,741.44.

Biographies

Allianz Indonesia. 2020. Come on, Understand the Underwriting Process in Insurance. Allianz.co.id. https://www.allianz.co.id/explore/detail/yuk-memahami-proses-underwriting-dalam- asuransi/96581.

David, M. 2015. Auto Insurance Premium Calculation Using Generalized Linear Models. *Procedia Economics and Finance, 20.* https://doi.org/10.1016/S2212-5671(15)00059-3.

Futami, T. 1993. *Life Insurance Math* : *Matematika Asuransi Jiwa: Bagian 1. Terj. Dari Seimei Hoken Sugaku, Jokan (92 Revision)*. Incorporated Foundation Oriental Life Insurance Cultural Development Center.

Jong, P.de, Heller, and Gillian Z. 2008. *Generalized Linar Models for Insurance Data*. Cambridge University Press

Kagan, J. 2020. *What is Health Insurance?* Investopedia.

Kalfin, Sukono, Supian S, and Mamat, M. (2021). Mitigation and models for determining premiums for natural disaster insurance due to excessive rainfall. In *Journal of Physics: Conference Series*, 1722(1), p. 012058. IOP Publishing.

Kalfin, Sukono, Supian S, Mamat M, and Bon A T. (2020a) Model for Determining Natural Disaster Insurance Premiums in Indonesia Using the Black Scholes Method. *Proceedings of the International Conference on Industrial Engineering and Operations Management* Detroit Michigan, USA, pp. 2376–2385.

Kalfin,, Sukono,, Supian, S., Mamat, M., Bon, A.T. (2020b). A review mitigation efforts and natural disaster insurance schemes in indonesia. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Harare, Zimbabwe, pp. 1825–1835.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. 1998. *Loss Models: From Data To Decisions* (3rd ed.). Wiley

Latumaerissa, J. R. 2011. *Bank dan Lembaga Keuangan Lain*. Salemba Empat.

McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models* (Second). Chapman and Hall

Nelder, J. A., and Wedderburn, R. W. M. 1972. Generalized Linear Model. *Journal of the Royal Statistikal Society Series A (General), 135*(3).

Sirait, V., Hasriati, Putra, S., Kalfin, Sukono, and Bon A T. (2020). Estimating Expected Time for Recruitment of Human Resource Companies Health Insurance based on Lomax Distribution. *Proceedings of the International Conference on Industrial Engineering and Operations Management* Detroit Michigan, USA, pp. 2498– 2501.

Subagyo. 2002. *Bank dan Lembaga Keuangan Lain*. Sekolah Tinggi Ilmu Ekonomi.

Supriatna, A., Riaman, Sudradjat, dan Septiyani, T. 2017. Generalized Linear Models (GLM) untuk Data Asuransi dalam Menentukan Harga Premi. *Prosiding Seminar Nasional Pendidikan Matematika.*

Tse, Y-K. 2009. *Nonlife Actuarial Models Theory, Methods, and Evaluation*. Cambridge University Press.

Widianti, K. R. 2018. *Penentuan Harga Premi Bersih dengan Mempertimbangkan Faktor Underwriting Menggunakan Generalized Linear Models (GLM)*. UNPAD

Riaman is a teaching staff in the mathematics study program at the Mathematics Department of the University of Pajajaran since 1997. The mathematics degree was won from the Mathematics Department of the University of Padjadjaran in 1995, while the Masters of Actuarial Science was obtained from the Bandung Institute of Technology in 2000. The field of research currently being undertaken is the field of applied mathematics. More specifically are:

Sukono is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Currently as Chair of the Research Collaboration Community (RCC), the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.
fields of financial mathematics, actuarial mathematics, survival models, mathematical statistics, and reliability. Since 2001 until now he has joined as a member of IndoMS, and since 2017 joined IAENG.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He’s bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM and MIM.