Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 15(1) 2022: 24-35
http://ejournal.undip.ac.id/index.php/media_statistika

ESTIMATION OF IBNR AND RBNS RESERVES USING RDC METHOD AND GAMMA GENERALIZED LINEAR MODEL

Tiara Yulita[1], Adhitya Ronnie Effendie[2]
[1] Actuary Study Program, Sumatera Institute of Technology, Indonesia
[2] Mathematics Study Program, Gadjah Mada University, Indonesia

e-mail: tiara.yulita@at.itera.ac.id

DOI: 10.14710/medstat.15.1.24-35

Article Info:
Received: 24 March 2021
Accepted: 10 June 2022
Available Online: 27 July 2022

Keywords:
IBNR, RBNS, RDC, Gamma Generalized Linear Models, Reserves.

Abstract: Estimation of claims reserves is a very important role for insurance companies because the information will be used to assess the insurance company's ability to meet future claim payment obligations. In practice, claims reserves are divided into two Incurred but Not Reported (IBNR) and Reported but Not Settled (RBNS). Reserving by Detailed Conditioning (RDC) is one of the individual methods that can estimate claims reserves of both the IBNR and RBNS, which involves detailed condition so-called claim characteristics, and some information else so-called background variable. The result of estimating claims reserves using RDC with background variable is not stable because many combinate of calculation from each background variable. The purpose of this study is to overcome these problems, which we can combine RDC and Gamma Generalized Linear Model (GLM) as an effective method for estimating claims reserves. By using Bootstrapping Individual Claims Histories (BICH) method, the results show that estimation of claims reserves using RDC and Gamma GLM gives the fewest value of Mean Square Error of Prediction (MSEP) rather than RDC with Poisson GLM, RDC, and Chain Ladder. Where the smaller the value of the resulting MSEP estimate, the closer to the actual claim reserve value.

1. INTRODUCTION

In some types of non-life insurances, sometimes claims payment is done more than one time. It requires a long time counted from the occurrence period. These characteristics in insurance are known as long tailed business or third part liability (TPL) (Verrall et al, 2010). Because of the existence of the time-lag between reported claim and finalized claim shown a term called outstanding claims liability. To overcome the problem of the outstanding claims liability, a company has to have enough fund specified to pay the liability are called reserves. According to Pigeon et al (2014) insurance claims reserve are divided into IBNR and RBNS. Estimated claims reserve IBNR and RBNS have a very important role in an insurance company because it can lead to bankruptcy if the estimate is bad.

The method of estimating claims reserves can be divided into two parts, the aggregate method, and the individual method. Some aggregate methods used to estimate claims reserve are Chain Ladder, Bornhuetter Fergusson, Benktander Hovinen, and Cape-Cod. These methods are widely applied in practice because simple and give accurate results (Wilandari et al, 2021). According to Godecharle & Antonio (2015), if the data is aggregated, then a lot of important claim information will be omitted when estimating claims reserves. Therefore, it is necessary to estimate the claim reserve using an individual method that includes claim information. Rosenlund (2012) estimates claims reserves using the individual method called RDC. In RDC, the information of each individual claim of policyholders is included as a condition in claims reserves estimation.

Based on Effendie & Pebriawan (2017) for each policy and the number of them is large, RDC needs to use the claim information which would then be formed into the claim characteristics and background variable as a condition in the calculation. The background variable will result in estimates that are not stable, due to the combination of each background variable calculation. The purpose of this study is to overcome these problems, which will estimate the claim reserves for both IBNR and RBNS using the RDC combined with the Gamma GLM method to analyze the background variables that are included as requirements in the RDC. We also determine a more accurate and effective estimate of claim reserves from several methods, including Chain Ladder, RDC, RDC and Poisson GLM, and RDC and Gamma GLM using the BICH method where the smallest MSEP is chosen.

2. LITERATURE REVIEW

2.1. Claims Reserves

According to Wüthrich & Merz (2008), the history of a typical non-life insurance claim there are two different types of claims reserves for past exposures. IBNR (incurred but not reported) reserves, this claim is an event that occurred and was announced in the information media but has not been reported to the insurance company. And RBNS (reported but not settled) reserves, this reserves to state the time of payment or settlement. The reserves included claims in processing of being settled.

2.2. Individual Run-Off Triangle

The run-off triangle is used to predict the size of claims or the number of claims that will occur in the future and is often used in the long-tailed business insurance. The individual run-off triangle is a presentation of claims data for each individual policyholder such that the information corresponding to the individual claims can also be displayed (Drieskens et al, 2012). The general form of the individual run-off triangle can be seen in Table 1.

Table 1. Individual Run-Off Triangle

Occurrence Period	Claim ID	Development Period				
		1	$\cdots$	j	$\cdots$	n
1	$c_{1;1}$	$Y_{c_{1;1;1}}$	$\cdots$	$Y_{c_{1;1;j}}$	$\cdots$	$Y_{c_{1;1;n}}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	$c_{2;1}$	$Y_{c_{2;1;1}}$	$\cdots$	$Y_{c_{2;1;j}}$	$\cdots$	$Y_{c_{2;1;n}}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$c_{n;n}$	$Y_{c_{n;1;1}}$	$\cdots$	$Y_{c_{n;1;j}}$	$\cdots$	$Y_{c_{n;1;n}}$

In most cases, data are set up in what we call a development triangle where the rows represent accident years and the columns development years. Suppose $Y_{c_{i,j}}$ with $i, j \in \{1,2,\ldots,n\}$, is a random variable of a claim that occurred at the i^{th} occurrence period and paid at the j^{th} period of development where in claim ID C. The $Y_{c_{i,j}}$ is the incremental claim that is observed when $i + j \le n + 1$ and $i + j > n + 1$ are not observed (Mutaqin et al, 2008).

2.3. Contraction of Data

2.3.1. Individual Claim Information

According to Kartikasari et al (2017) in RDC, the claim information of N claims with each claim k $(k = 1, 2, \dots, N)$ on continuous time are discretized, we denote in (1) as a set of individual claim information

$$\{i(k), W(k), F(k), Y(k, 1), \cdots, Y(k, F(k))\} \quad (1)$$

The notation of $i(k)$ denotes the occurrence period of claim k with $i(k) \in \{1,2, \dots, n\}$. $W(k)$ denotes reporting period of claim k. $F(k)$ denotes the closing period for claim k. $Y(k, j)$ denotes payment for claims k at development period $j \in \{1,2, \dots, F(k)\}$.

2.3.2. Claim Characteristics

The individual claim information as described by (1) cannot be directly used as a condition of claim reserve estimation. Therefore, we have to create a claim information representative called claim characteristics to simplify the calculation process. The characteristics of claim in the RDC are divided into three parts, including claim length $L(k)$, reporting delay period $W(k)$, and quantile interval group number of cumulative payment $Q_t(k)$. The steps for the formation of these three parts will be explained below.

Claim Length

The claim length is the duration from claim reported up to claim finalized. We denote the length of claim k $(k = 1, 2, \dots, N)$ by $L(k)$ and define it as follows

$$L(k) = F(k) - W(k) + 1 \quad (2)$$

Claim statuses can be obtained from $L(k)$ and $W(k)$ and can be defined as follows: Closed claims are claims that have been reported and have been settled, where $W(k) \le n - i(k) + 1$ and $L(k) \le n - i(k) - W(k) + 2$, or $W(k)$ and $L(k)$ is known; IBNR is claims that have occurred but has not been reported, where $W(k) > n - i(k) + 1$ and $L(k) = 0$ or $W(k)$ and $L(k)$ are unknown; RBNS is claims that have been reported but not resolved, where $W(k) \le n - i(k) + 1$ and $L(k) \le n - i(k) - W(k) + 2$, or $W(k)$ is known but $L(k)$ is not observed.

Reporting Delay Period

Claims that have the same reporting delay period may have a similar development pattern. Based on that reason, a reporting delay period is used as the one of claim characteristics. In this case, we set w_0 as a maximum limit of $W(k)$. A claim is considered to be reported to the insurance company if $W(k) \ge w_0$. Therefore, we define $\min(W(k), w_0)$ as a boundary of reporting delay of an individual claim.

Quantile Interval Group Number of Cumulative Payment

The claims will be divided into several quantile intervals based on information from the cumulative payments. The cumulative payments, $H(k, t)$, for claim k on the time $t \in \{1,2, \cdots, n - i(k) - W(k) + 2\}$ starting from the reporting period is defined by

$$H(k, t) = \sum_{h=1}^{t} Y(k, h + W(k) - 1) \quad (3)$$

for computing the quantile intervals, we have to sort in ascending order $H(k,t)$ with $L(k) > t$ in a set of $H(\cdot,t)$. We split $H(\cdot,t)$ into q_0 quantile intervals, with q_0 is fixed in advance. If we denote the boundaries of quantile as $\{h_{t,1}, h_{t,2}, \cdots, h_{t,q_0}\}$, we can create the quantile intervals of $H(\cdot,t)$ as $[0, h_{t,1}], [h_{t,1}, h_{t,2}], \cdots, [h_{t,q_0-1}, h_{t,q_0}]$. The interval number to which $H(k,t)$ belongs is indicated by $Q_t(k)$ with $L(k) > t$, $Q_t(k) \epsilon \{1,2,\cdots,q_0\}$ and for $t > 0$

$$Q_t(k) = \text{quantile interval number of } H(k,t) \tag{4}$$

2.3.3. Background Variable

Kroon (2014) explained that using these claim markers to simulate future claim development may prove to be problematic, since some combinations of these markers may result in very small clusters. The background variable is used as an assessment factor in determining the amount of the premium, so very influential and a requirement for calculating claims reserves (Ohlsson & Johansson, 2010). Background variable always follows one of the categories: policy holder, insured item, geographical area. The background variable is denoted by A_γ, $\gamma \in \{1,2,\cdots,n\}$ where $A_1 = \{x_1, x_2, \cdots, x_m\}$. The value $\{x_1, x_2, \cdots, x_m\}$ are represented by a unique code per segment.

2.3.4. Segmentation

Segmentation is carried out based on different background variables for each claim. The basic idea is that the estimation of claims reserve using RDC is calculated per segment, where the numbers of segments depend on the number of unique codes from the initial background variable and all combinations of background variables that are used as segmentation references.

2.4. Gamma Generalized Linear Model

Gamma GLM is the right method to estimate the severity claims since Gamma distribution is one of the families of exponential distribution with a constant coefficient of variation. According to De Jong & Heller (2008), a gamma GLM is of the form

$$y \sim G(\mu, v) \text{ and } g(\mu_i) = log[E(Y_i)] \Leftrightarrow e^{log[\mu_i]} = e^{x_i^T \beta} \tag{5}$$

The canonical link for gamma distribution is the inverse function. Since parameters from a model with an inverse link are difficult to interpret, the log link is usually more useful. To estimate the parameter of Gamma distribution is used on claims reserve used likelihood and loglikelihood with probability in (5), and if the shape of the first derivative (Score Function) is not close-form then use Newton-Raphson iteration (Sunandi et al., 2022)

$$f(y_i) = \frac{1}{\Gamma(v)} \left(\frac{v y_i}{\mu_i}\right)^v e^{\left(-\frac{v y_i}{\mu_i}\right)} \frac{1}{y_i} \tag{6}$$

3. MATERIAL AND METHOD

3.1. Reserving by Detailed Conditioning

Rosenlund (2012) there are two steps to calculate claims reserves estimation using RDC. First, we compute an estimate of the probability distribution of claim length. Second, we compute an estimate of the expected value of the payment amount per future development period. The reserve can be obtained by multiplication the first and second one.

3.1.1 Estimation of Claim Length Probability

Rosenlund (2012) defines the estimation of the probability of claim length as

$$\hat{p}_\lambda(t,q,w) = \hat{r}_\lambda(t,q,w) \times \prod_{m=t+1}^{\lambda-1} \left(1 - \hat{r}_m(t,q,w)\right) \tag{7}$$

for $q \epsilon \{1,2,\dots,q_0\}$ and $w \epsilon \{1,2,\dots,w_0\}$ with the given $0 \le t \le n-1$ and $t+1 \le \lambda \le n$. Notation $\hat{p}_\lambda(t,q,w)$ express the probability that a claim has a claim length by λ. The notation of $\hat{r}_\lambda(t,q,w)$ is the estimated probability of finalized claim in the following period. The formula $\hat{r}_\lambda(t,q,w)$ is obtained from the following process

$$\begin{cases} \hat{r}_\lambda(t,q,w) = 1 & \lambda = n - w + 1 \\ \hat{r}_\lambda(t,q,w) = \dfrac{I_\lambda^F(t,q,w)}{J_\lambda(t,q,w)} & \lambda < n - w + 1 \end{cases} \tag{8}$$

where $I_\lambda^F(t,q,w)$ = number of claims closed given $L(k) = \lambda, Q_t = q, \min(W(k), w_0) = w$ and $J_\lambda(t,q,w)$ = number of claims reported given $L(k) \ge \lambda, Q_t = q, \min(W(k), w_0) = w$.

3.1.2 Estimation of Mean Payment

The estimation of mean payment is obtained via a combination of payments from opened claims and closed claims. Stage of observation estimation of mean payment is

1. For closed claims, the observations of claims are done in the period $h \le n - i - W + 2$ as payment with the given $0 \le t \le n-1, t+1 \le \lambda \le n$, and $t+1 \le h \le \lambda$. For all $q \epsilon \{1,2,\cdots,q_0\}$, $w \epsilon \{1,2,\cdots,w_0\}$, given $L(k) \le n - i - W + 2$, $L(k) = \lambda, Q_t = q, \min(W(k), w_0) = w$

   $I_{\lambda h}^F(t,q,w)$ is the number of closed claims given $L(k) = \lambda$

   $Y_{\lambda h}^F(t,q,w)$ is the number of closed claim payments given $L(k) \ge \lambda$.

2. For opened claims, the observations of claims are done in the period $t \le n-2$ with the given $0 \le t \le n-2, t+1 \le r \le n-1$, and $t+1 \le h \le r$ where r is opened claim period. For all $q \epsilon \{1,2,\cdots,q_0\}$, $w \epsilon \{1,2,\cdots,w_0\}$, given $r = n - i - W + 2, Q_t = q, \min(W(k), w_0) = w$.

   $I_r^o(t,q,w)$ is the number of opened claims given $L(k) = r$

   $Y_{rh}^o(t,q,w)$ is the number of opened claim payments given $L(k) > r$.

3. For outstanding claims in the previous period r, it will be complete in the next period $(L(k) = \lambda)$

$$I_{r\lambda}^o(t,q,w) = \frac{\hat{p}_\lambda(t,q,w)}{\hat{p}_{r+1}(t,q,w) + \cdots + \hat{p}_n(t,q,w)} I_r^o(t,q,w) \tag{9}$$

   $I_{r\lambda}^o(t,q,w)$ is the number of outstanding claims in the previous period will be complete in the next period given $\lambda = r+1, \cdots, n, Q_t = q, \min(W(k), w_0) = w$

$$Y_{r\lambda h}^o(t,q,w) = \beta_{rh}(t,q,w) \frac{Y_{\lambda h}^{r+1}(t,q,w)}{I_\lambda^{r+1}(t,q,w)} I_{r\lambda}^o(t,q,w) \tag{10}$$

   $Y_{r\lambda h}^o(t,q,w)$ is the number of outstanding claim payments in the previous period will be complete next period given $\lambda = r+1, \cdots, n, Q_t = q, \min(W(k), w_0) = w$

$$\beta_{rh}(t,q,w) = Y_{rh}^{o}(t,q,w)\left[\sum_{v=r+1}^{n} Y_{vh}^{r+1}(t,q,w)\frac{I_r^o(t,q,w)}{I_v^{r+1}(t,q,w)}\right]^{-1} \tag{11}$$

If $\beta_{rh}(t,q,w)$ will be undefined, the alternative way to calculate the value of

$$Y_{r\lambda h}^{o}(t,q,w) = \frac{\hat{p}_{\lambda}(t,q,w)}{\hat{p}_{r+1}(t,q,w)+\cdots+\hat{p}_{n}(t,q,w)}Y_{rh}^{o}(t,q,w) \tag{12}$$

4. The estimation of mean payment is obtained using the calculation of backward recursive. The backward recursive calculation on period $r=\lambda$ up to $r=h$ use the initial value

$$\begin{cases} I_{\lambda}^{(\lambda)}(t,q,w) = I_{\lambda}^{F}(t,q,w) \\ Y_{\lambda h}^{(\lambda)}(t,q,w) = Y_{\lambda h}^{F}(t,q,w) \end{cases} \tag{13}$$

The equation of recursive calculation for $r=\lambda-1,\lambda-2,\cdots,h$ can be expressed by

$$\begin{cases} I_{\lambda}^{(r)}(t,q,w) = I_{\lambda}^{(r+1)}(t,q,w) + I_{r\lambda}^{o}(t,q,w) \\ Y_{\lambda h}^{(r)}(t,q,w) = Y_{\lambda h}^{(r+1)}(t,q,w) + Y_{r\lambda h}^{o}(t,q,w) \end{cases} \tag{14}$$

5. After the recursive calculation up to the period of h is complete, we can obtain the final estimation of mean payment with $h=t+1,\cdots,n$ and $\lambda=h,\cdots,n$ that is expressed by

$$\hat{\mu}_{\lambda h}(t,q,w) = \frac{Y_{\lambda h}^{(h)}(t,q,w)}{I_{\lambda}^{(h)}(t,q,w)} \tag{15}$$

3.1.3 Estimation of Claim Reserve

The estimation of claim reserves is divided into two types, i.e. IBNR and RBNS. The total estimation of claims reserves in the occurrence period $i \in (1,2,\dots,n)$ can be expressed

$$\hat{R}_i = \hat{R}_i^I + \hat{R}_i^R \tag{16}$$

with $\hat{R}_i^I$ denotes IBNR claims reserves and $\hat{R}_i^R$ denotes RBNS claims reserves.

IBNR Reserve

IBNR reserves are claim reserves that have not yet been reported to the insurance company. Given $W(k)=w$, for each $W(k)=w\epsilon\{1,2,\dots,w_0\}$ the estimated IBNR reserves

$$\hat{R}(0,1,\min(w,w_0)\,) = \hat{p}_{\lambda}(0,1,\min(w,w_0)) \times \hat{\mu}_{\lambda h}(0,1,\min(w,w_0)) \tag{17}$$

The estimation of the number of claims per occurrences period $i \in (2,3,\dots,n)$ is done using the well-known performed using Overdispersed Poisson model by utilizing the data claims that have been reported to the company $(\hat{A}_{iw})$. The estimated IBNR claims reserves in the occurrence period $i \in (2,3,\dots,n)$ can be expressed by

$$\hat{R}_i^I = \sum_{w=n-i+2}^{n} \hat{A}_{iw} \times \hat{R}(0,1,\min(w,w_0)) \tag{18}$$

RBNS Reserve

RBNS reserve is a reserve of a claim that has been reported to the insurance company but that has not yet been finalized. We denote RBNS reserve at claim period i as following

$$\hat{R} = \hat{p}_{\lambda}(n-i-W+2,q,\min(w,w_0)\,) \times \hat{\mu}_{\lambda h}(n-i-W+2,q,\min(w,w_0)\,) \quad (19)$$

Suppose for the claim period i, $I^{o}(i,q,w)$ is the number of opened claims given $W(k)=w,q$. Then RBNS reserves for the claim period i is

$$\hat{R}_i^R=\sum_{w=1}^{n-i+1}\sum_{q=1}^{q_0} I^{o}(i,q,w)\times\hat{R}(n-i-W+2,q,\min(W(k),w_0)) \tag{20}$$

3.2. Reserving by Detailed Conditioning and Gamma GLM

Based on steps from RDC, estimation of claims reserve using RDC and Gamma GLM has the same stages, only the background variables and their combinations need to be included. Below are the stages of RDC and Gamma GLM

1. The estimation of IBNR and RBNS per segment on a background variable using RDC (R1) has the following formula

$$^{x_m}\hat{R}_i\ (t,q,w)=\sum_{\lambda=t+1}^{n}\sum_{h=t+1}^{\lambda}{}^{x_m}\hat{p}_\lambda(t,q,w)\,{}^{x_m}\hat{\mu}_{\lambda h}(t,q,w) \tag{21}$$

If we get a background variable A_γ given $\gamma\in\{1,2,\cdots,n\}$ where $A_1=\{x_1,x_2,\cdots,x_m\}$, so we have $x_k\in A_1$ as a segment where $k\in\{1,2,\cdots,m\}$. The calculation $^{x_m}\hat{p}_\lambda(t,q,w)$ and $^{x_m}\hat{\mu}_{\lambda h}(t,q,w)$ are the same as the steps in the RDC, only they are carried out per segment on the background variable.

2. The estimation of IBNR and RBNS per segment on all combinations of background variables using RDC (R2) has the following formula

$$^{x_m y_n z_o}\hat{R}_i\ (t,q,w)=\sum_{\lambda=t+1}^{n}\sum_{h=t+1}^{\lambda}{}^{x_m y_n z_o}\hat{p}_\lambda(t,q,w)\,{}^{x_m y_n z_o}\hat{\mu}_{\lambda h}(t,q,w) \tag{22}$$

Suppose we have three background variables $A_1=\{x_1,x_2,\cdots,x_m\}$, $A_2=\{y_1,y_2,\cdots,y_n\}$, and $A_3=\{z_1,z_2,\cdots,z_o\}$. The calculation $^{x_m y_n z_o}\hat{p}_\lambda(t,q,w)$ and $^{x_m y_n z_o}\hat{\mu}_{\lambda h}(t,q,w)$ are the same as the steps in the RDC per segment on all combinations of background variables.

3. Smoothing of estimation RBNS claims reserve using Gamma GLM (R3).
Suppose we have three background variables and we assume that the background variable is independency polish, independency time, and homogeneity, so the multiplicative model is obtained as follows (Ohlsson & Johansson, 2010)

$$\mu_i=\gamma_0\gamma_{1,a}\gamma_{2,b}\gamma_{3,c} \tag{23}$$

Model (23) can be formed into a model of generalized linear models with circuit function log, which indicates Y_i Gamma distribution

$$\begin{aligned}&\beta_1=\log(\gamma_0),\beta_2=\log(\gamma_{1,2}),\beta_3=\log(\gamma_{2,2}),\beta_4=\log(\gamma_{2,3}),\\&\beta_5=\log(\gamma_{3,2}),\text{ So }\gamma_0=e^{\beta_1},\gamma_{1,2}=e^{\beta_2},\gamma_{2,2}=e^{\beta_3},\gamma_{2,3}=e^{\beta_4},\gamma_{3,2}=e^{\beta_5}\end{aligned} \tag{24}$$

If we build a dummy variable so we get $\log(\mu_i)=\sum_{j=1}^{5}x_{i,j}\beta_j$ given $i=1,2,\cdots,12$. To estimate the parameter of Gamma distribution used derivative from likelihood and loglikelihood function from formula (6) so we get $\frac{\partial l}{\partial\beta_j}=\sum_{i=1}^{12}\left\{-v+\frac{vy_i}{\mu_i}\right\}x_{i,j}\Leftrightarrow\sum_{i=1}^{12}vy_ix_{i,j}=\sum_{i=1}^{12}v\mu_ix_{i,j}$. The formula is not in closed form, so it can't provide a solution. This is caused by the relationship between the log-likelihood function equation derived against β_j. Thus, the estimated value of the Poisson model parameters must be calculated by maximizing the log-likelihood function numerically by the iteration technique known as Newton-Raphson method (McCulloch & Searle, 2004).

4. The estimation of RBNS claims reserve (R4).
   Calculation estimate of RBNS claims reserve for R4 using the below definition.

   Definition 1 (Rosenlund, 2012) *For each claim period and a variable background, which claims t the claim period i and background variables k, so*

$$R4(t) = R3(t) \times \frac{\sum_{u>1,k} R1(u)}{\sum_{u>i,k} R3(u)} \tag{25}$$

3.3. Calculation MSEP with Bootstrapping Individual Claim Histories (BICH)

One of the ways calculation errors in the study is the mean square error prediction (MSEP). The calculation of MSEP in estimation claims reserves in this paper is carried out using the Bootstrapping Individual Claim Histories (BICH). The bootstrap method BICH is given for estimating mean square prediction error and predictive distributions of non-life claim reserves under weak conditions. We assume that all claims are independent and that the historic claims are distributed as the object claims, possibly after inflation adjustment and segmentation on a background variable (Rosenlund, 2012).

The procedure for creating bootstrap T shadow data using claim data for finding the MSEP estimate is as follows: identifying claim data which is distributed as claim data T; taking a random sample of v with a return from the claim data so that the number of reported bootstrap shadow claim data is the same; calculating $R_i^{(v)}$as claim liability of T^i for $i = \{1,2, \dots, n\}$ with $R_i = \sum_{j=n-i+2}^{n} Y_{ij}$ and $\hat{R}_i^{(v)}$ as estimation of claim reserve for $i = \{1,2,, \dots n\}$; repeat step 2 and step 3 to B times from bootstrap claim data; calculating MSEP where according to Rosenlund (2012) defines a bootstrap MSEP estimation equation as

$$\tilde{\tau}_i = \sqrt{\frac{1}{B}\sum_{v=1}^{B}(R_i^{(v)} - \hat{R}_i^{(v)})}. \tag{26}$$

4. CASE STUDY AND RESULTS

4.1. Data Description

We have data of liability insurance claim from non-life insurance company, namely *Länsförsäkringar Alliance* in Sweden. It consisted of 1.710.629 data of liability insurance from opened claims and closed claims. The data starts from January 2011 to December 2012 with time units of study in months. The data claim consists of the information of claim identity (claim ID); occurred period; reporting period; period of claims closed; payment data (in Krona or Sweden currency); background variable A_1 is geographic area consist of : city (code 07), village (code 38); background variable A_2 is line of business consists of individual (code 02), company (code 05), professional (code 07); background variable A_3 is a type of damage object consisting of personal property (code HC), public property (code VF); background variable A_4 is damage : fire (code 010), accident (code 011), damage (code 130).

4.2. Result

For data, we describe reporting delay period, maximum, and minimum value of cumulative claim payments in Table 2. Computing programming using R program packages “glmmML” (Brostrom, 2020) and Rapp program (Rosenlund, 2021). Before starting calculates reserve estimation of liability insurance claim, we must determine the value of w_0 and q_0 based on descriptive statistics. Based on Table 2, we assume that a claim is reported late if $W(k) \geq 4$. Therefore, we set $w_0 = 4$ due to most of the reported claim until period $W(k) = 4$. For q_0 we set $q_0 = 10$ based on Table 2 where maximum value is 182.232 and minimum value is 0 so we can set claims payment is divided into ten groups based on cumulative payment also large amount of data.

Table 2. Reporting Delay, Maximum, and Minimum Cumulative Claim Payments of Liability Insurance Claim

W	1	2	3	4	5	6	7	8	9	10	11	12
Number of claims	121,393	22,206	21,970	21,518	0	0	0	0	0	0	0	0
%	64.9	11.9	11.7	11.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Max	64,407	128,814	103,770	121,488	151,860	182,232	118,874	78,128	0	0	0	0
Min	0	0	0	0	0	0	0	0	0	0	0	0

Before starting to calculate the reserve estimation of the liability insurance claim, we must determine the value of w_0 and q_0 based on descriptive statistics. Based on Table 2, we assume that a claim is reported late if $W(k) \geq 4$. Therefore, we set $w_0 = 4$ due to most of the reported claims until period $W(k) = 4$. For q_0 we set $q_0 = 10$ based on Table 2 where the maximum value is 182,232 and minimum value is 0 so we can set claims payment is divided into ten groups based on cumulative payment also a large amount of data.

In step one, we estimate IBNR and RBNS per segment on a background variable using RDC and we define the background variable used as the segment as geographic area. It is divided into two codes, code 07 and 38 so that it produces IBNR segments 07 and 38, also RBNS segments 07 and 38. Table 3 below is the summary estimation of IBNR and RBNS total per segment on a background variable in Krona Sweden (kr).

Table 3. Summary of Estimation IBNR & RBNS per Segment on a Background Variable (Area) Using RDC (R1) (In kr)

Occurred Period	IBNR Claims Reserve			RBNS Claims Reserve			IBNS Total Per Segment
	Segment 07	Segment 38	Total Per Segment	Segment 07	Segment 38	Total Per Segment	
201201	0	0	0	0	0	0	0
201202	0	0	0	45,954	108,262	154,216	154,216
201203	0	0	0	249,134	370,419	619,554	619,554
201204	0	0	0	591,187	869,258	1,460,446	1,460,446
201205	0	0	0	1,642,653	1,970,951	3,613,604	3,613,604
201206	0	0	0	3,274,764	3,486,887	6,761,651	6,761,651
201207	0	0	0	6,182,197	6,082,210	12,264,407	12,264,407
201208	0	0	0	9,518,421	9,455,757	18,974,178	18,974,178
201209	0	0	0	14,098,188	13,870,922	27,969,110	27,969,110
201210	4,153,782	7,384,768	11,538,550	16,554,340	13,666,640	30,220,980	41,759,530
201211	10,412,589	12,756,792	23,169,381	17,590,656	14,348,498	31,939,151	55,108,535
201212	16,395,554	16,835,711	33,231,265	19,530,753	15,237,598	34,768,351	67,999,616
Total			67,939,196			168,745,651	236,684,847

In step two, we estimate IBNR and RBNS per segment on all combinations of background variables using RDC. In claims data from *Länsförsäkringar Alliance* insurance we have 4 background variables, there are geographic area ($A_1 = \{7,38\}$), line of business ($A_2 = \{2,5,7\}$), type of damage object ($A_3 = \{HC, VF\}$), dan type of damage ($A_4 = \{10,11,130\}$), we get 36 combinations. Table 4 below is the summary estimation of IBNR and RBNS total per segment on all combinations of background variables.

Table 4. Summary of Estimation Claims Reserve RBNS per Segment on All Combination of Background Variable Using RDC (R2) (In kr)

Occurred Period	Combination of Background Variable							Total
	7-2-HC-10	38-2-HC-10	7-5-HC-10	…	38-5-VF-130	7-7-VF-130	38-7-VF-130	
201201	0	0	0	…	0	0	0	0
201202	3,103	4,026	1,145	…	26	1,285	2,215	127,951
201203	19,787	24,059	12,119	…	3,955	25,710	5,029	557,802
201204	33,711	56,406	36,517	…	7,101	35,972	21,977	1,396,599
201205	161,421	126,650	106,003	…	20,453	89,972	49,735	3,541,319
201206	270,885	240,378	177,853	…	35,314	124,665	110,804	6,693,949
201207	421,536	367,482	502,169	…	101,042	286,520	225,198	12,198,806
201208	681,019	596,878	678,100	…	129,024	365,103	275,978	18,867,221
201209	1,079,443	947,212	1,004,489	…	131,292	396,117	534,513	27,891,519
201210	1,172,848	959,128	1,242,217	…	239,676	467,661	511,919	30,124,718
201211	1,264,348	950,247	1,222,012	…	222,709	554,404	470,761	31,807,693
201212	1,249,729	836,786	1,411,857	…	155,686	553,847	672,454	34,868,901

In step three, we smoothing estimation RBNS from step two used Gamma GLM with applied likelihood function or loglikelihood function and Newton-Raphson iteration. For step four, we calculated estimation RBNS from step three used (25). Then we can calculate the value of IBNR and RBNS. Step three, four, and summary we can see in Table 5.

Table 5. The Smoothing of Estimation RBNS Claims Reserve Using Gamma GLM (R3), Results of R4, and Summary of Estimation Claims Reserve Using RDC and Gamma GLM (In kr)

Occurred Period	RBNS R(3)	RBNS R(4)	Summary of Estimation Claims Reserve		
			IBNR	RBNS	Total
201201	0	0	0	0	0
201202	122,099	120,234	0	120,234	120,234
201203	552,563	544,121	0	544,121	544,121
201204	1,312,647	1,292,593	0	1,292,593	1,292,593
201205	3,449,132	3,396,440	0	3,396,440	3,396,440
201206	6,768,741	6,665,335	0	6,665,335	6,665,335
201207	12,105,639	11,920,702	0	11,920,702	11,920,702
201208	19,025,724	18,735,068	0	18,735,068	18,735,068
201209	27,916,547	27,490,066	0	27,490,066	27,490,066
201210	31,584,984	31,102,461	11,538,550	31,102,461	42,641,011
201211	32,345,010	31,850,876	23,169,381	31,850,876	55,020,257
201212	36,180,483	35,627,755	33,231,265	35,627,755	68,859,020
Total	171,363,568	168,745,651	67,939,196	168,745,651	236,684,847

For the same data, the following will be displayed the claim reserve estimation result using RDC with some combination of the characteristics claim. For the case of liability insurance claim data in this study, it has not been decided which method to use is the right one. Therefore, the use of the estimated MSEP is needed to be able to provide conclusions regarding the choice of a more accurate method for estimating the reserve for liability insurance claims in this study. The smaller the value of the resulting MSEP estimate, the closer the estimated claim reserve value is to the actual claim reserve value (Rosenlund, 2012).

Table 6. The Comparison of MSEP Estimation Value Obtained from Some Claims Reserve Estimation Methods (in kr)

No	Estimation Method Claim Reserve	Total Estimation Claim Reserve	Estimation of MSEP
1	RDC Gamma GLM with $w_0 = 1$ and $q_0 = 5$	236,900,708	1,759,125
2	RDC Gamma GLM with $w_0 = 1$ and $q_0 = 25$	236,222,627	1,689,349
3	RDC Gamma GLM with $w_0 = 2$ and $q_0 = 1$	237,355,075	1,749,433
4	RDC Gamma GLM with $w_0 = 2$ and $q_0 = 10$	237,224,820	1,524,234
5	RDC Gamma GLM with $w_0 = 3$ and $q_0 = 3$	238,047,793	1,481,452
6	RDC Gamma GLM with $w_0 = 3$ and $q_0 = 5$	237,292,180	1,496,902
7	RDC Gamma GLM with $w_0 = 4$ and $q_0 = 3$	237,670,059	1,473,476
8	RDC Gamma GLM with $w_0 = 4$ and $q_0 = 8$	236,679,618	1,412,413
9	RDC Gamma GLM with $w_0 = 4$ and $q_0 = 10$	236,684,847	1,390,087
10	RDC Poisson GLM with $w_0 = 4$ and $q_0 = 10$	236,015,678	1,409,944
11	RDC with $w_0 = 4$ and $q_0 = 10$	236,426,655	1,416,711
12	Chain Ladder	234,469,993	2,373,472

5. CONCLUSION

We use BICH to compare RDC with Chain Ladder and RDC is found overall best in simulated cases. From Table 6 we get an estimation of MSEP use BICH method that RDC with $w_0 = 4$ and $q_0 = 10$ is fewer value rather than chain ladder. It shows that RDC more accurate than Chain Ladder. Also in Table 6, where RDC with Poisson GLM has MSEP estimation is fewer value rather than RDC, and Chain Ladder. But if we have extra requirements claim information on RDC will result in estimates that are not stable, due to the combination of each background variable calculation. To overcome these problems, it can be used the method of Gamma GLM to analyze the background variables that were included as a condition in the RDC. In this data, we have many background variables so we must combine them. In Table 6 we get that RDC and Gamma GLM with characteristics claims of $w_0 = 4$ and $q_0 = 10$ is 1.390.087 kr. It gives a fewer value of MSEP estimation than RDC with Poisson GLM, RDC, and also Chain Ladder. We get estimation claims reserve using RDC based on Gamma GLM gives more effective and accurate results than some other method for estimation of claims reserve on liability insurance data and stabilize calculations from many combinate background variable.

REFERENCES

Brostrom, G. (2020). glmmML: Generalized Linear Models With Clustering. *Http://Cran. r-Project. Org/Web/Packages/GlmmML/Index. Html*.

De Jong, P., & Heller, G. Z. (2008). Generalized Linear Models for Insurance Data. *Cambridge Books*.

Drieskens, D., Henry, M., Walhin, J. F., & Wielandts, J. (2012). Stochastic Projection for Large Individual Losses. In *Scandinavian Actuarial Journal* (Issue 1). https://doi.org/10.1080/03461231003759708

Effendie, A., & Pebriawan, R. (2017). Estimation of IBNR and RBNS Reserve by Detailed Conditioning Method. *Far East Journal of Mathematical Sciences (FJMS)*, *101*, 2785–2801. https://doi.org/10.17654/MS101122785

Godecharle, E., & Antonio, K. (2015). Reserving by Conditioning on Markers of Individual Claims: A Case Study Using Historical Simulation. *North American Actuarial Journal*, *19*(4), 273–288. https://doi.org/10.1080/10920277.2015.1046607

Kartikasari, M. D., Effendie, A. R., & Wilandari, Y. (2017). Reserving by Detailed Conditioning on Individual Claim. *AIP Conference Proceedings*, *1827*(March 2017). https://doi.org/10.1063/1.4979428

Kroon, R. (2014). *Individual Reserving by Detailed Conditioning - A Parametric Approach*.

McCulloch, C. E., & Searle, S. R. (2004). *Generalized, Linear, and Mixed Models*. John Wiley & Sons.

Mutaqin, A. K., Tampubolon, D. R., & Darwis, S. (2008). Run-off Triangle Data dan Permasalahannya. *Statistika*, *8*(1), 55–59.

Ohlsson, E., & Johansson, B. (2010). *Non-Life Insurance Pricing with Generalized Linear Models*.

Pigeon, M., Antonio, K., & Denuit, M. (2014). Individual Loss Reserving using Paid-Incurred Data. *Insurance: Mathematics and Economics*, *58*(1), 121–131. https://doi.org/10.1016/j.insmatheco.2014.06.012

Rosenlund, S. (2012). Bootstrapping Individual Claim Histories. *ASTIN Bulletin: The Journal of the IAA*, *42*(1), 291–324. https://doi.org/10.2143/AST.42.1.2160744

Rosenlund, S. (2021). Rapp: The Free Actuarial Language. *Https://Www.Stigrosenlund.Se/Rapp.Htm*.

Sunandi, E., Notodiputro, K. A., & Sartono, B. (2022). a Study of Generalized Linear Mixed Model for Count Data Using Hierarchical Bayes Method. *Media Statistika*, *14*(2), 194–205. https://doi.org/10.14710/medstat.14.2.194-205

Verrall, R., Nielsen, J. P., & Jessen, A. H. (2010). Prediction of RBNS and IBNR Claims Using Claim Amounts and Claim Counts. *ASTIN Bulletin*, *40*(02), 871–887. https://doi.org/10.2143/AST.40.2.2061139

Wilandari, Y., Gunardi, & Effendie, A. R. (2021). Estimasi Cadangan Klaim Menggunakan Metode Deterministik dan Stokastik. *Jurnal Statistika*, *9*(1).

Wüthrich, M. V, & Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance* (Vol. 435). John Wiley & Sons.