Twelfth International Conference
Financial and Actuarial Mathematics – FAM-2019

Classification of the IBNR Methods

Pavel Todorov Stoynov
Sofia University "St. Kliment Ohridski",
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Sofia, Bulgaria
Nikolay Kozlev,
Sofia University "St. Kliment Ohridski",
Sofia, Bulgaria

Abstract: The article considers a classification of the IBNR models in actuarial mathematics. .

Key words: insurance, actuarial mathematics, IBNR models.

Introduction

The IBNR models in insurance are a realistic way for determining the actuarial reserves taking into account the time. After the acceptance of the Solvency II directive in EU, they are widely used by the insurers acting at the territory of the union.

Apart from the abbreviation IBNR (Incurred but not Reported), other abbreviations are also used – RBNS (Reported but Not Settled), IBNFR (Incurred but not Fully Reported), IBNER (Incurred but Not Enough Reported), RBNFS (Reported but not Fully Settled). All they express the idea that there are time intervals between the occurrence of claims and their complete payment.

The basic IBNR model can be presented by a square table of random variables $X_{ij}, i = 1,2,\ldots,t,\ j = 1,2,\ldots t,$ where X_{ij} are the paid claims occurred in year i and paid after $j-1$ years, i. e. during the calendar year $i + j - 1,$ and $i = 1,2,\ldots,t$ is the year of occurrence of the actuarial event (year of origin, accident year), and $j = 1,2,\ldots t$ is the year of payment of the claims (development year). In this way, t^2 independent random variables are obtained, indexed by two indecesq among which $\frac{t(t+1)}{2}$ variables are observations of historical values till the calendar year. The observed values have indeces $(i, j): i + j - 1 \leq t.$ The explaining variables here are the row nimber i and the column number $j.$ The observed variable in each cell can be really paid claims, loss ratios, average values on all w_{ij} observations in the cell which play the role of natural weights or other variables.

Main point in this class of models is the so-called run-off triangle. It can be considered as a matrix $k \times k,$ in which the meaningful numbers are on and above the diagonal defined by the equation $k = i + j - 1.$ The numbers in the run-off triangle can be cumulative or incremental losses of the insurer for the corresponding period of occurrence and development.

IBNR models can be divided into two large groups - standard (classical) models and contemporary (stochastic) models.

The standard IBNR models can be divided into models using historical data and models using historical data and apriori estimates.

IBNR models can be also divided to models which uses historical data only for claims and models which use historical data for the calims and for the premiums collected.

From the combination of these two classifications we obtain four groups standard IBNR models which will be considered below.

Among the models with historical data only about claims is the classical Chain Ladder модел.

The classical Chain Ladder model uses the so-called chain-ladder factors defined by the mean chain ratios

$$f_k^{CL} = \frac{\sum_{i=1}^{n-k} S_{i(k+1)}}{\sum_{i=1}^{n-k} S_{ik}}, \quad k = 1,2,\ldots,n-1,$$

$$S_{ik} = \sum_{j=1}^{k} X_{ij},\ i = 1,2,\ldots,n; \quad k = 1,2,\ldots,(n-i+1).$$

From hereq the final Loss Development Factors – LDF are defined.

$$F_k^{CL} = \prod_{j=k}^{n-1} f_j^{CL},.\ k = 1,2,\ldots,(n-1); \quad F_n^{CL} = 1.$$

They represent mean ratio of the Ultimate Aggregate Paid Claims – UAPC – to the aggregated registered claims for every period after k years of development.

From the factors of development of the losses chain-ladder lag-factors can be determined:

$$p_i^{CL} = \frac{1}{F_{n-i+1}^{CL}},\ i = 1,2,\ldots,n.$$

They represent the mean share of the UAPC from period of occurrence $i,$ revealed til the day of the analysis in development period $n-i+1.$

The chain-ladder IBNR factors are also obtained

$$q_i^{CL} = 1 - p_i^{CL},\ i = 1,2,\ldots,n.$$

They are the mean share of UAPC from occurrence period $i,$ remained unknown till the day of the analysis in development period $n-i+1.$

Then we have

$$U_i^{CL} = \frac{S_{i(n-i+1)}}{p_i^{CL}}$$

for the final aggregate claims paid $U_i^{CL},\ i = 1,2,\ldots,n.$ For the incurred but not reported claims IBNR the following equation holds:

$$IBNR_i^{CL} = q_i^{CL} U_i^{CL},\ i = 1,2,\ldots,n.$$

The classical Chain Ladder model can be divided in the contxt of the multiplicative generalized regression model

$$X_{ij} \approx \alpha_i \beta_j \gamma_k.$$

Her X_{ij} are independent exponentially distributed random variables. The year of the occurrence of the claim $i,$ the year of payment j and the calendar year $k = i + j - 1$ are the explanatory independent variables and X_{ij} are the explained dependent variables. Then EX_{ij} is an exponent of the linear $\log\alpha_i + \log\beta_j + \log\gamma_k$, i. e. the generalized regression model has a logarithmic function of the regression function. The variables $\alpha_i, \beta_j, \gamma_k$ are dummy, indicating the row, the column and the diagonal. For them an estimate can be found by the method of maximal likelihood at different assumptions about the distribution of X_{ij}. Then the estimated values are

$$\hat{X}_{ij} = \hat{\alpha}_i \hat{\beta}_j \hat{\gamma}_k.$$

One common assumption about γ_k, when $k > t,$ is that $\gamma_k = \gamma^k$ for some real number $\gamma.$

If we suppose that $X_{ij} \in Po(\alpha_i\beta_j),\ \gamma_k \equiv 1$ and we look for such estimates $\widehat{\alpha}_i, \hat{\beta}_j,\ i, j = 1,\ldots,t,$ for which the summing by i and j of the products $\widehat{\alpha}_i\hat{\beta}_j,\ i + j - 1 \le t,$ is equal to the corresponding sums of the observed variables, i. e.

$$\sum_i \hat{\alpha}_i \hat{\beta}_j = \sum_i X_{ij} = C_j,$$

$$\sum_j \hat{\alpha}_i \hat{\beta}_j = \sum_j X_{ij} = R_i,$$

which means the method of Marginal Totals is used, the future losses of the insurer for $(i, j) : i + j - 1 > t$ can be estimated by $\hat{\alpha}_i \hat{\beta}_j,\ i + j - 1 > t.$ Since the estimates by the method of Marginal Totals in case of $w_{ij} = 1$ are given by the equations

$$\alpha_i = \frac{\sum_j X_{ij}}{\sum_j \beta_j}$$

$$\beta_j = \frac{\sum_i X_{ij}}{\sum_i \alpha_i},$$

the Chain Ladder method leads to solving this system of dependencies for $(i, j) : i + j - 1 \le t.$ One of the approaches to solvethe systemis the method of Verbeek (1972).

Wuth the Verbeek method (1972), as one of the parameters is not needed, an assumption is made that $\sum_{j=1}^{t} \beta_j = 1.$ Also

$$\sum_i X_{ij} = C_j,\ i + j \le t$$

$$\sum_j X_{ij} = R_i,\ i + j \le t.$$

From summing of the first row we obtain

$$\hat{\alpha}_1 (\hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_t) = R_1,$$

from where

$$\hat{\alpha}_1 = R_1.$$

After, from

$$\hat{\alpha}_1 \hat{\beta}_t = C_t,$$

is obtained

$$\hat{\beta}_t = \frac{C_t}{R_1}.$$

If we assume that for some $n < t$ the estimates $\hat{\beta}_{t-n+2}, \ldots, \hat{\beta}_t$ and $\hat{\alpha}_1, \ldots, \hat{\alpha}_{n-1}$ are found, the next estimates can be maid by using the equations

$$\hat{\alpha}_n (\hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_{t-n+1}) = R_n,$$

$$\hat{\alpha}_1 (\hat{\alpha}_1 + \cdots + \hat{\alpha}_n) \hat{\beta}_{t-n+1} = C_{t-n+1}.$$

As $(\hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_t) = 1,$ the first equation gives an estimate for $\hat{\alpha}_n$, and from the second equation we may determine $\hat{\beta}_{t-n+1}$. This procedure is repeating for $n = 2, \ldots, t.$

As examples of modification of Chain Ladder Double Chain Ladder (DCL) and Munich Chain Ladder (MCL) can be shown.

Double Chain Ladder – DCL – uses two run-off triangles – triangle of payments and triangle of the number of claims. The classical Chain Ladder method is used for both run-off triangles.

When calculating reserves in the context of IBNR models separate run-off triangles are used for the paid losses and for the incurred losses, i. e. Paid losses and assigned reserves. The method Munich Chain Ladder (MCL) combines paid and incurred losses considering and projecting their ratio.

Among the models with curren data based only on claims and using apriori data, the model Bornhuetter-Ferguson (1972) and the model Cape Code can be cited.

At Cape Code model, the mean ratio of the losses (ratio of total losses to mean premiums) is used. More specifically, the ratio

$$L = \frac{\sum_{i=1}^{n} S_{i(n-i+1)}}{\sum_{i=1}^{n} p_i^{CL} P_i}$$

is defiend where P_i are the premiums collected during the occurrence period $i.$ Then we have

$$IBNR_i^{CC} = q_i^{CL} LP_i,\ i = 1,2,\ldots,n.$$

The total claims are correspondingly

$$U_i^{CC} = S_{i(n-i+1)} + IBNR_i^{CC} = p_i^{CL} U_i^{CL} + (1 - p_i^{CL}) LP,\ i = 1,2,\ldots,n.$$

At Bornhuetter-Ferguson model, instead of mean ratio of losses, an estimate of the loss ratio $L_i,\ i == 1,2,\ldots,n$ is used for every occurrence period $i.$ Then

$$IBNR_i^{BF} = q_i^{CL} L_i P_i,\ i = 1,2,\ldots,n.$$

For the aggregate claims we have

$$U_i^{BF} = S_{i(n-i+1)} + IBNR_i^{BF} = p_i^{CL} U_i^{CL} + (1 - p_i^{CL}) L_i P,\ i = 1,2,\ldots,n.$$

One of the weeknesses of Chain Ladder is that the prognoses about the reserves can be unstable. A change by $p\%$ of the sum of the number of the claims in a given row leads to a change with the same percentage in all predicted values of the numbers of the claims in this row. The method of Bornhuetter-Ferguson (1972) is a trial to stabilize the claims prognoses.

At this method, it is supposed that there are preliminary estimates for the total amount of the claims which will appear for a given initial period $i.$ More precisely, it is supposed that

$$E(X_{i1} + X_{i2} + \cdots + X_{it}) = M_i$$

where M_i are known and are also called schedule or budget ultimate losses. So, the estimates $\frac{M_i}{P_i}$ are given where P_i is the premium income for year $i.$

If we consider the case when the number of the insurance policies n_i for every occurrence year is known, and X_{ij} is the total number of claims and the loss ratio $\frac{M_i}{P_i} = const$ is constant, which means that

$$\frac{M_i}{P_i} = \frac{M_1}{P_1},$$

$$\frac{M_i}{n_i \overline{p}} = \frac{M_1}{n_1 \overline{p}},$$

wher $\overline{p}$ is the mean premium per policy, then

$$M_i = \frac{n_i}{n_1} M_1.$$

From other side,

$$E(X_{i1} + X_{i2} + \cdots + X_{it}) = M_i = \sum_j \hat{\alpha}_i \hat{\beta}_j \approx \hat{\alpha}_i$$

And it follows that $M_1 = \hat{\alpha}_1,$ from where

$$M_i = \frac{n_i}{n_1} \hat{\alpha}_1.$$

The last equation gives a preliminary estimate of the mean value of the total loss, corresponding to occurrence year $i.$

With the method Chain Ladder, the vectors $\hat{\alpha}$ are $\hat{\beta}$ constructed supposing that the sum of the coordinates of the vector β is equal to, thus minimizing the devience between the future data and the historic data, i. e. The upper left triangle, of matrix $\hat{\alpha}\hat{\beta}'.$ After that the part of the future, i. e. the lower right triangle of the matrix $\hat{\alpha}\hat{\beta}'$ is used for determining of the reserves. The sum by rows of the future part of the matrix $\vec{M}\hat{\beta}'$ is determined by the methods of Bornhuetter-Ferguson.

Bornhuetter-Ferguson method combines internal development year factors β_i with external a priori estimates of the loss ratio by occurrence years. At Chain Ladder method, the (development year factors, as well as the accident year factors are estimated internally. It can be shown (Verrrall, 2004) that both methods are two border cases of the Bayesian estimates at generalized linear regression models with loose prior distribution at the method Chain ladder and tight prior distribution at the method of Bornhuetter-Ferguson.

By applying the method Chain Ladder, different actuaries would achieve similar results. By applying the method Bornhuetter-Ferguson different results are possible about the prior estimate of the total sum of claims $S.$ This sum can be manipulated – idf the aim is to receive estimate of the reserves $R,$ then it is possible to give an estimate of the total sum of claims $S = \dfrac{R}{q}.$

The method Cape Code (CC) can be considered as a modification of the method Bornhuetter-Ferguson. With this method, we assume that the premiums or other measures of the volume of the business are known for historical accident years and the final loss ratios are identical for all claims occurenc years. This method is sometimes called also Standart-Buhlman method.

The models Optimal Bornhuetter-Ferguson and Optimal Cape Code are modifications of Bornhuetter-Ferguson and Cape Code.

Among the models with current data based on claims and premiums without using prior data the method Individual Loss Ratiocan be considered.

Among the models with current data based on claims and premiums and using prior data the models Collective Loss Ratio and Credibility Models of total Loss Ratio – Benktander Credibility Loss Ratio, Neuhaus Credibility Loss Ratio, Optimal Credibility Loss Ratio can be mentioned.

The model of Gunar Benktander (1976) can be considered as a mixed model based on the models Chain Ladder and Bornhuetter-Ferguson. At this model, the reserves are estimated by the formula

$$\begin{aligned} R_k^{GB} &= p_k R_k^{CL} + (1-p_k) R_k^{BF} = \\ &= \frac{p_k q_k S_k}{p_k} + q_k R_k^{BF} = \\ &= q_k (S_k + R_k^{BF}) = q_k U_k^{BF}, \end{aligned}$$

where $p_k = p_{ik}$ are the parts of the total claim which occurred during period i, and were paid during period $k,\ q_k = 1 - p_k,\ S_k = S_{ik} = \sum_{j=1}^{k} X_{ij}$ are the total claims paid till period $k,\ X_{ij}$ are the incremental claims for period of occurrence i and development period $j,\ R_k^{BF}$ are the reserves determined by the method of Bornhuetter-Ferguson for the period $k,\ U_k^{BF} = S_k + R_k^{BF}$ are the posterior estimates of the total claims by the method of Bornhuetter-Ferguson. While at the method of Bornhuetter-Ferguson the reserves are determined from the equation $R_k^{BF} = q_k U,$ where U are the prior estimates for the total claim and at the method Chain Ladder the reserves are determined from $R_k^{CL} = U_k^{CL} - S_k = \dfrac{S_k}{p_k} - S_k = q_k U_k^{CL},$ where U are the prior estimate for the total sum of claims, at the model of Gunar Benktander the process of Bornhuetter-Ferguson continues iteratively and with the development of the claims, i. e. with increase of the index of the development year $j,$ the weight p_k of the estimate of the total claims maid by the method Chain Ladder also inceases. The following relations hold:

$$R_k^{BF} = q_k U = q_k U_0,$$

$$U_1 = U_k^{BF} = U_k^{GB1} = S_k + q_k U_0 = (1 - q_k) U_k^{CL} + q_k U_0,$$

$$R_1 = R_k^{GB1} = q_k U_1 = q_k U_k^{BF} = (1-q_k)R_k^{CL} + q_k R_k^{BF},$$
$$U_2 = U_k^{GB2} = (1-q_k^2)U_k^{CL} + q_k^2 U_0,$$
$$\ldots$$
$$U_n = U_k^{BFn} = (1-q_k^n)U_k^{CL} + q_k^n U_0,$$
$$R_n = R_k^{GBn} = (1-q_k^n)R_k^{CL} + q_k^n R_k^{BF},$$
$$U_{n+1} = U_k^{BF(n+1)} = (1-q_k^{n+1})U_k^{CL} + q_k^{n+1} U_0,$$
$$\ldots$$
$$U_\infty = U_k^{CL},$$
$$R_\infty = R_k^{CL}.$$

Thus the reserves determined by the method Gunar Benktander are obtained through mixed model (credibility mixture) of the reserves, determined by Chain Ladder and the reserves determined by Bornhuetter-Ferguson. The optimal mix of the two models is defined by the conditions

$$R_k^c = cR_k^{CL} + (1-c)R_k^{BF},$$
$$c \in [0,1],$$
$$\min E(R_k^c - ER_k^c)^2 \to c^*.$$

Thus, about the reserves corresponding to the optimal model R_k^c it can be shown that they are always better than R_k^{CL} and R_k^{BF}, ant the estimate of reserves R_k^{GB1} is better than R_k^{CL} и R_k^{BF} in the most of the cases.