Constrained Forecasting of the Number of IBNR Claims

Louis G. Doray[*]

Abstract

This paper considers the problem of forecasting the number of claims incurred. After substracting the number of claims reported to date, the number of claims incurred but not reported (IBNR) can be forecasted. The basic model assumes that the number of claims per accident period follow an auto-regressive moving average time series process. Instead of assuming the data are available in the usual claim run-off triangle format, we assume that the only data available are the number of claims reported at the valuation date for each accident interval of an observation period. Box-Jenkins methods are used to forecast the ultimate number of claims incurred and to obtain approximate confidence intervals for the number of claims incurred. The forecast of the ultimate number of claims incurred is used to derive the IBNR forecast. We show how additional information on the number of claims reported by the valuation date can be incorporated in the model, when the process is auto-regressive.

Key words and phrases: *time series, Box-Jenkins, quadratic programming, truncated normal distribution.*

1

1   Introduction

In all lines of insurance, there is usually a delay between the occurence of the event giving rise to a claim and the time that the claim is first reported to the insurance company. This reporting delay, however, is more serious in certain lines of insurance than in others. For some lines, the reporting lag may be measured in days or weeks (as in the case of life insurance), while in others it may be measured in years or decades (as in the case of claims arising from environmental hazards such as asbestos). Reinsurers experience longer delays since they must wait for claims to exceed the retention level before being notified. Notwithstanding this reporting lag, the insurance company must estimate, at the end of each valuation year, the liability arising from the claims that have been incurred but not reported (IBNR).

Various statistical models have been proposed to estimate IBNR claims. For example, Van Eeghen (1981) and Taylor (1986) contain a survey of some of those models. Time series models have been successfully used to model past claims amounts and forecast future claims amounts. Lemaire (1982) used an autoregressive model where the amount paid in a certain accident and development year is a linear combination of the amount in the cell above it and the one to its left. He estimated the parameters by a least-squares method. Verrall (1989) considered a similar model but used maximum likelihood theory; his model was selected with the Akaike Information Criterion (AIC).

A common thread running through most research on IBNR is the assumption that the claims paid by the insurer can be grouped according to accident and development year, resulting in a triangular array of numbers. This is called the claim run-off triangle in the literature. In this paper however, we consider a subset of the traditional structure for the data set. We assume that the only data available are the number of claims reported at the valuation date for each accident month of an observation period.

The paper is organized as follows: Section 2 gives a review of the basics of Box-Jenkins time series analysis. Section 3 shows how to estimate the number of IBNR claims with an $ARMA(p,q)$ model. The additional information on the number of claims reported to-date is then incorporated into the model. By minimizing the sum of the squared forecast errors, subject to the ultimate number of claims incurred for a certain accident month being at least equal to the number of claims reported by the valuation date, the problem is transformed into a quadratic programming problem with linear inequality constraints. The forecasted number of claims incurred is calculated with a modified simplex algorithm. Revised and smaller confidence intervals for the ultimate number of claims incurred with respect to each

accident period can be obtained, using the truncated normal distribution. In section 4, we work out with actual data, an example of the application of the the methods developed in the previous section, to the estimation of incurred and IBNR claims, and to the derivation of approximate confidence intervals for these estimates. Section 5 discusses non-stationary time series. The conclusions follow in Section 6.

2 Forecasting Using Box-Jenkins Methods

A discrete-time stochastic process $\{Z_t, t = 1, \ldots, n\}$, where Z_t takes a a real valued at time t, is said to be weakly stationary (see Brockwell and Davis (1991)) if:

1. $E(|Z_t^2|) < \infty$, for $t = 1, \ldots, n$

2. $E(Z_t) = \mu$ is constant for $t = 1, \ldots, n$, and

3. $Cov(Z_r, Z_s) = Cov(Z_{r+t}, Z_{s+t})$, for $r, s, r + t, s + t = 1, \ldots, n$, i.e., the covariance structure depends only on the distance $|r - s|$.

In time series analysis, it is usually more convenient to use the zero-mean process Y_t defined as
$$Y_t = Z_t - \mu$$
In what follows, we will assume that the sequence of Y_t, for $t = 1, \ldots, n$, have a joint multivariate normal distibution. The observed time series will be represented by lower-case letters $\{z_1, \ldots, z_n\}$ and the centered observations by $\{y_1, \ldots, y_n\}$

2.1 Auto-Regressive and Moving Average Processes

There are three basic time series processes in the Box-Jenkins framework: auto-regressive, moving average and mixed auto-regressive moving average processes. Given that $\{a_t\}_{t=1}^{\infty}$ is a sequence of uncorrelated normal random variables with mean 0 and variance σ^2, then for $t = 1, 2, 3, \ldots$ these processes are briefly defined below:

- The *auto-regressive process* of order p, $AR(p)$, is represented as

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + a_t \tag{1}$$

where p is a positive integer and $\phi_1, \ldots, \phi_p$ are constants with $\phi_p \neq 0$.

- The moving average process of order q, $MA(q)$, is represented by

$$Y_t = a_t + \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q}. \tag{2}$$

  where q is a non-negative integer and $\theta_1, \ldots, \theta_q$ are constants with $\theta_q \neq 0$. Note that

$$Var[Y_t] = (1 + \theta_1^2 + \ldots + \theta_q^2)\sigma^2.$$

  Clearly, observations more than a distance of q steps apart are uncorrelated.

- A mixed auto-regressive moving average process, $ARMA(p,q)$, can be represented as

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} \\ &\quad + a_t + \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q}. \end{aligned} \tag{3}$$

2.2   Model Identification

This section gives a brief overview of the process of model identification using the Box-Jenkins method. The model selected can then be used to forecast the time series. Readers unfamiliar with time series analysis using this method should consult one of the many available references on the subject (e.g. Box and Jenkins (1976), Harvey (1981), Abraham and Ledolter (1983) or Pankratz (1983)).

The method consists of the following three steps:

1. Identification of the process generating the data, by looking at graphs of the sample autocorrelation function (a.c.f) and partial autocorrelation function (p.a.c.f.). The sample a.c.f. is the set of autocorrelations at lag k defined by

$$r_k = \frac{\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}.$$

   The partial autocorrelation at lag k is the correlation of the two residuals after regressing y_t and y_{t-k} on the intermediate observations $y_{t-k+1}, \ldots, y_{t-1}$.

2. Estimation of the parameters of the model; and

3. Verification tests to determine if the fit of the model is adequate.

For an $AR(p)$ process, the p.a.c.f. is zero after lag p, while the a.c.f. decays exponentially to zero. For the $MA(q)$ process, the a.c.f. cuts off after lag q, while the p.a.c.f. decays to zero.

After the estimation stage, the fit of the model can be checked through a test of goodness-of-fit. (See Brockwell and Davis (1991).) The Portmanteau statistic, also called the modified Box-Pierce statistic (Box and Pierce (1970)),

$$R = n \sum_{k=1}^{K} r_k^2(\hat{a}),$$

is calculated with K usually around 20. In this formula, $r_k(\hat{a})$ is the auto-correlation at lag k between the residuals,

$$\hat{a}_t = y_t - \hat{y}_t \tag{4}$$

$$\hat{a}_{t-k} = y_{t-k} - \hat{y}_{t-k}$$

and $\hat{y}_t$ is the value computed with the estimated value of the parameters. When an $ARMA(p, q)$ process is appropriate, R is distributed as a chi-squared random variable with $k - p - q$ degrees of freedom (χ_{k-p-q}^2); large values of the test statistic R indicate inadequacy of the model.

2.3 Forecasting

According to Anderson (1976), when the observed series of data is large, the estimation error in the parameters will not in general be serious. If we then assume that the model is known exactly for the past and that it will not change in the future, we can obtain forecasted values by minimizing the mean squared error of forecasts. Anderson (1976) shows that, for the ARMA process, the best l-step ahead forecast at time t, linear in a_t, is given by

$$\hat{y}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \ldots \tag{5}$$

where ψ_j, for $j = 1, 2, \ldots$, is the coefficient of x^j in the Taylor series expansion of

$$\psi(x) = \frac{1 + \theta_1 x + \ldots + \theta_q x^q}{1 - \phi_1 x - \ldots - \phi_p x^p}.$$

This forecast is unbiased and has minimum mean squared error. It therefore has minimum variance in the class of linear estimators.

Forecast errors at various leads will however be correlated. The l-step ahead forecast error at time t is equal to

$$e_t(l) = Y_{t+l} - \hat{y}_t(l) = \sum_{j=0}^{l-1} \psi_j a_{t+l-j}.$$

with variance
$$Var[e_t(l)] = (1 + \psi_1^2 + \ldots + \psi_{l-1}^2)\sigma^2.$$

By estimating σ^2 by $\hat{\sigma}^2$ and ψ_j by $\hat{\psi}_j$, an approximate $100(1 - \alpha)\%$ confidence interval for the forecast is obtained:

$$\hat{y}_t(l) \pm [\hat{\sigma}^2 \sum_{j=0}^{l-1} \hat{\psi}_j^2]^{1/2} \Phi(\alpha/2), \tag{6}$$

where $\Phi(\alpha/2)$ is the $(100 - \alpha/2)$ percentile of the standard normal distribution.

3   Estimation of IBNR

3.1   The Basic Approach

Using the theory for $ARMA(p, q)$, we will show how it can be applied to estimate the number of IBNR claims. Let $\{z_t, t = 1, \ldots, n\}$ be the number of claims incurred during time periods $1, \ldots, n$. These time intervals could be months, quarters or years. We assume that the maximum delay between occurence and reporting of a claim is known.

To identify the model and estimate its parameters, we will use only the data of the n periods for which the number of claims is fully known. The process modelling the number of claims during each period can be identified by making graphs of the the sample a.c.f. and p.a.c.f.. The parameters of the model are then estimated by using one of the many time series software available. For an $ARMA(p, q)$ model, this would give the maximum likelihood (MLE) estimates $\hat{\theta}_1, \ldots, \hat{\theta}_q, \hat{\phi}_1, \ldots, \hat{\phi}_p$, their standard error, as well as the MLE for the process variance. A goodness-of-fit test to check the model adequacy is then performed.

Once the model has been validated, the forecasting of the number of claims incurred beyond time period n can be performed using equation (5). For example, the forecast at time n for periods $n + 1$ and $n + 2$ would be

$$\hat{y}_n(1) = \hat{\psi}_1 a_n + \hat{\psi}_2 a_{n-1} + \ldots$$
$$\hat{y}_n(2) = \hat{\psi}_2 a_n + \hat{\psi}_3 a_{n-1} + \ldots.$$

The standard error of the forecast is then calculated to get a confidence interval around the estimate. Let $r_{n+1}, r_{n+2}, \ldots$ be the number of claims reported at time $n + 1, n + 2, \ldots$. The number of IBNR claims predicted would then be $\hat{y}_n(1) - r_{n+1}, \hat{y}_n(2) - r_{n+2}, \ldots$.

Note that in this sub-section, we have not made any use of the partial number of claims reported for certain periods. This is done in the next sub-section.

3.2 Minimum Forecast Error With Constraints

To use the additional information on the partial number of claims reported for certain periods, we will again minimize the sum of the squared forecast errors, but now subject to the constraints that each forecast should be larger than or equal to the number of claims reported by the valuation date, i.e.

$$\text{Minimize} \quad \sum_{l=n+1} a_l^2, \quad \text{subject to } \hat{z}_l \geq r_l,$$

where r_l is the number of claims reported for period l.

This problem is a standard quadratic programming problem with linear inequality constraints. When we have an $AR(p)$ process, the problem can be rewritten in matrix form as

$$\begin{aligned}
\text{Minimize} \quad g(\hat{y}) \quad &= \quad \frac{1}{2}\hat{y}'Q\hat{y} \\
\text{subject to} \quad A\hat{y} \quad &\geq \quad b,
\end{aligned}$$

where Q, A are matrices, and b, y are vectors.

Writing the objective function in this form will make the matrix Q symmetric; it will be positive semi-definite since $g(\hat{y})$ is a sum of squares.

Hillier and Lieberman (1986) or Luenberger (1984) present algorithms to solve this type of problem when the matrix Q is positive semi-definite, using a modified simplex algorithm.

Note that the constraint of ultimate claim counts being no less than the number reported to date will not apply if complete salvage or subrogation recoveries are present and eliminate a previously reported claim; cumulative claim counts for a fixed accident period would then decline slightly at later evaluation dates. The method proposed here would not be applicable in this case.

3.3 Confidence Intervals With Constraints

Since the errors in our ARMA model are assumed to be normal, the forecasted numbers of claims for each accident period will also have a normal distribution. But that forecasted number of claims must be greater than the number reported at the end of the observation period. The forecasted number of incurred claims will therefore have a normal distribution truncated from below. Johnson and Kotz (1970), and Patel and Read (1982) discuss the properties of the truncated normal distribution.

A random variable X has a truncated normal distribution, with lower

truncation point A, if its pdf is given by

$$f_X(x) = \sigma^{-1} Z\left(\frac{x-\mu}{\sigma}\right)\left[1 - \Phi\left(\frac{A-\mu}{\sigma}\right)\right]^{-1}, \quad x \geq A$$

where $Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$.

On pages 81-83 of Johnson and Kotz (1970), the following expressions for the expected value and the variance of X are derived.

$$E(X) \quad = \quad \mu + \frac{\sigma Z(\frac{A-\mu}{\sigma})}{1 - \Phi(\frac{A-\mu}{\sigma})} \tag{7}$$

$$Var(X) \quad = \quad \sigma^2 \left[1 + \frac{\left(\frac{A-\mu}{\sigma}\right) Z\left(\frac{A-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{A-\mu}{\sigma}\right)} - \left\{\frac{Z\left(\frac{A-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{A-\mu}{\sigma}\right)}\right\}^2\right] \tag{8}$$

The upper bound of the 95% confidence interval is obtained by solving for x the equation

$$0.95 = \frac{\left[\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi\left(\frac{A-\mu}{\sigma}\right)\right]}{1 - \Phi\left(\frac{A-\mu}{\sigma}\right)}. \tag{9}$$

4   An Example

4.1   The Data

Table 1 contains the number of third-party automobile liability claims reported by September 30, 1987 (the valuation date), to a propety/casualty insurance company, for each accident month of the observation period January 1980 to September 1987 and figure 1 is the graph of that time series. We assume that all the claims which occured during accident years 1980 to 1986 have been reported by the valuation date, and because of a reporting delay, the ultimate number of claims actually incurred for each month of accident year 1987 is at least as large as the number reported (next-to-last column of Table 1).

4.2   Model Determination

We will now analyze the data of Table 1. Let $\{z_t, t = 1, \ldots, 84\}$ represent the numbers of claims reported on September 30, 1987, for each month of the accident period 1980-1986, and let y_t be the centered observation. The graph of $\{z_t\}$ against time (see figure 1) shows no trend in the mean or

Table 1: Number of accidents reported by September 30, 1987

Month	Accident year								Mean
	1980	1981	1982	1983	1984	1985	1986	1987	(1980-86)
JAN	144	218	230	151	210	170	178	202	185.86
FEB	149	243	179	135	142	177	130	156	165.00
MAR	164	187	145	154	159	120	154	138	154.71
APR	124	189	143	144	132	102	134	153	138.29
MAY	196	244	169	189	167	156	213	198	190.57
JUN	208	230	169	206	180	195	201	178	198.43
JUL	226	266	153	198	186	186	201	127	202.29
AUG	190	226	161	206	157	184	203	142	189.57
SEP	234	229	173	176	185	167	219	93	197.57
OCT	260	265	154	220	192	167	205	-	209.00
NOV	234	179	189	208	197	167	193	-	195.29
DEC	257	201	153	197	153	260	162	-	197.57
Mean	198.83	223.08	168.17	182	171.67	170.92	182.75	-	

non-constant variance, indicating that the stationarity assumption is adequate for the data. Figure 2 contains the graphs of the sample a.c.f. and p.a.c.f.; we observe that the sample p.a.c.f. goes to zero after lag 1, suggesting the use of an $AR(1)$ process. Fitting that model to $\{y_1, \ldots, y_{84}\}$ with the ITSM (1991) software [1], we obtain the MLE of ϕ_1 and its estimated standard deviation (in brackets)

$$\hat{\phi}_1 = 0.5600628 \ (0.090391).$$

The model is therefore:

$$y_t = 0.5600628 y_{t-1} + a_t,$$

where a_t are independent $N(0, 885.562)$. An estimate of σ is thus $\hat{\sigma} = 29.76$.

Testing for randomness of the 20 residuals, we find a value of 13.6577 for the Portmanteau test statistic, which follows an asymptotic χ^2 distribu-

---

[1]The computer program PEST, contained in the ITSM software, uses a non-linear minimization procedure to search iteratively for the the value of the parameter ϕ_1 that maximizes the log-likelihood; the estimated value of σ^2 is then directly calculated.

Table 2: Forecasted numbers of claims for 1987 using the Box-Jenkins method

Month	Forecast	$(MSE)^{.5}$	Lower 95% C.B.	Upper 95% C.B.	Reported
JAN	172.27	29.76	113.94	230.60	202
FEB	178.02	34.11	111.16	244.88	156
MAR	181.24	35.36	111.93	250.55	138
APR	183.05	35.75	112.98	253.12	153
MAY	184.06	35.87	113.75	254.37	198
JUN	184.63	35.90	114.27	254.99	178
JUL	184.94	35.92	114.54	255.34	127
AUG	185.12	35.92	114.72	255.52	142
SEP	185.22	35.92	114.82	255.62	93
OCT	185.27	35.92	114.87	255.67	-
NOV	185.31	35.92	114.91	255.71	-
DEC	185.32	35.92	114.92	255.72	-

tion with 19 degrees of freedom (the critical value at the 5% level is 30.1). The model therefore provides an adequate fit to the data. Figure 3 contains a residual plot; two residuals are outside the 95% confidence interval $[-1.96\hat{\sigma}, 1, 96\hat{\sigma}]$.

4.3   Forecasting the Number of Claims Incurred

Using the above $AR(1)$ model, the next twelve monthly forecasts for 1987, as given by equation (6), appear in Table 2, along with the square root of their mean square error and 95% confidence bounds (C.B.). In order to forecast, we need to assume that the average monthly number of accidents and the variance will stay the same in 1987 as in previous years.

The 95% confidence interval for the forecasted number of claims for an accident month in 1987 is quite wide and widens as the forecast is further in the future. It covers the actual number of claims reported at September 30, 1987, for all accident months, except September (for this accident month, only the claims with a reporting lag of 0 month can be included).

In the next subsection, we see how these confidence intervals can be narrowed using the number of accidents incurred and reported in 1987 (last

column of Table 2).

4.4   Minimum Forecast Error With Constraints

In Section 4.3, to forecast the number of claims for accident year 1987, we used only the number of claims that occured during accident years 1980 to 1986. We now use the additional information on the number of claims reported for accident year 1987 to get a better forecast of the number of claims incurred in 1987.

For the $AR(1)$ process of section 4.2, the quadratic programming problem discussed in section 3.2 becomes:

$$\text{Minimize} \quad \sum_{l=85}^{96} (\hat{y}_l - 0.5600628\hat{y}_{l-1})^2$$

subject to

$$\hat{y}_{84} = -23.34524, \quad \hat{y}_{91} \geq -58.34524,$$
$$\hat{y}_{85} \geq 16.65476, \quad \hat{y}_{92} \geq -43.34524,$$
$$\hat{y}_{86} \geq -29.34524, \quad \hat{y}_{93} \geq -92.34524,$$
$$\hat{y}_{87} \geq -47.34524, \quad \hat{y}_{94} \geq -185.34524,$$
$$\hat{y}_{88} \geq -32.34524, \quad \hat{y}_{95} \geq -185.34524,$$
$$\hat{y}_{89} \geq -12.65476, \quad \hat{y}_{96} \geq -185.34524.$$
$$\hat{y}_{90} \geq -7.34524,$$

The figures on the right of the inequality signs represent the number of claims reported on September 30, 1987, for accident months December 1986 to December 1987 minus the average monthly number of claims for accident years 1980 to 1986 (185.34524).

In the $AR(1)$ case, $Q = \{Q_{ij}\}$ is a 13×13 symmetric tridiagonal matrix where

$$Q_{i,j} = \begin{cases} 2\hat{\phi}_1^2 & \text{if } i = j = 1 \\ -2\hat{\phi}_1 & \text{if } i = j - 1, \text{ for } j = 2, 3, \ldots, 13 \\ -2\hat{\phi}_1 & \text{if } i = j + 1, \text{ for } j = 1, 2, \ldots, 12 \\ 2(1 + \hat{\phi}_1^2) & \text{if } i = j, \text{ for } j = 2, 3, \ldots, 12 \\ 2 & \text{if } i = j = 13, \text{ and} \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

$\hat{\phi}_1 = 0.560$ and A is the identity matrix.

Table 3: Forecasted numbers of claims incurred for 1987 using quadratic programming

Month	Forecast	Month	Forecast
JAN	202.00	JUL	189.32
FEB	196.02	AUG	187.57
MAR	193.72	SEP	186.60
APR	194.31	OCT	186.05
MAY	198.00	NOV	185.74
JUN	192.44	DEC	185.57

Commercial software is available to solve quadratic programming problems of this type. Using the IMSL (1987) software [2], we obtained the forecasted values of Table 3.

The forecasts for accident months October, November and December 1987, are very close to those produced by the Box-Jenkins method, because the constraints for those months only specify that the number of accidents should be positive.

If the only information available for accident year 1987 was that the aggregate number of claims reported on September 30, 1987, totaled 1387 (without any information on the number of claims reported for each accident month), the constraints would become

$$\sum_{l=85}^{93} \hat{y}_i \geq -281.1716 \ \ (= 1387 - 9(185.34524)),$$
$$\hat{y}_l \geq -185.34524, \ \ l = 85, \ldots, 96,$$

since all z_i's need to be positive. The number of claims incurred for each accident month could then also be forecasted using quadratic programming.

Ordering among the number of claims to be forecasted for each accident month could also be accomodated; for example, the ordering

$$\hat{y}_i \geq \hat{y}_j \geq \hat{y}_k,$$

---

[2]The IMSL subroutine QPROG uses an efficient dual algorithm in quadratic programming for a positive definite matrix Q. It uses as a starting point the unconstrained minimum of the objective function and updates the solution with the Cholesky and QR factorizations.

Table 4: Forecasted numbers of claims for 1987 using the truncated normal distribution

Month	Mean	Std. Dev.	Lower 95% C.B.	Upper 95% C.B.
JAN	217.63	13.28	202.00	244.03
FEB	192.94	24.75	156.00	238.95
MAR	188.75	29.48	138.00	241.39
APR	195.58	27.29	153.00	245.64
MAY	222.11	19.22	198.00	259.74
JUN	209.19	22.86	178.00	252.87
JUL	189.06	32.16	127.00	244.97
AUG	193.00	29.81	142.00	246.30
SEP	185.75	35.22	114.82	255.62
OCT	185.27	35.92	114.87	255.67
NOV	185.31	35.92	114.91	255.71
DEC	185.32	35.92	114.92	255.72

can be transformed into the two linear inequalities

$$\hat{y}_i - \hat{y}_j \geq 0,$$

$$\hat{y}_j - \hat{y}_k \geq 0.$$

In the case of an AR process of order p, the matrix Q is still positive semi-definite, but becomes a band matrix.

4.5   Confidence Intervals With Constraints

The 95% confidence intervals for the forecasted number of claims incurred for each month of accident year 1987, which appear in Table 2, were quite wide. Using the techniques of Section 3.3 and the number of claims reported as of September 30, 1987, they can be narrowed.

Using formulas (7) and (8), we calculated the mean and standard deviation of the forecasted number of claims for each month of 1987. The results appear in Table 4. The upper bound of the 95% confidence interval was obtained by solving for x equation (9) and appears in the last column of Table 4.

Table 5: Estimated mean number of IBNR claims for 1987

Month	Forecast	Month	Forecast
JAN	15.63	JUL	62.06
FEB	36.94	AUG	51.00
MAR	50.75	SEP	92.75
APR	42.58	OCT	185.27
MAY	24.11	NOV	185.31
JUN	31.19	DEC	185.32

Comparing Table 2 and Table 4, we note that for the months of October, November and December, the truncation point at 0 has no effect on the mean, the standard deviation or the confidence interval, but for the other accident months, truncating from below reduces the standard deviation and the width of the confidence interval markedly, especially for the month of January.

Since the actuary is ultimately interested in the number of IBNR claims, the number of claims reported to date has only to be substracted from the estimated number of claims incurred to get the estimated number of IBNR claims; Table 5 contains the estimated mean number of IBNR claims.

In this analysis, we considered each accident month separately. We could consider the joint multivariate distribution of the forecasted number of claims incurred for each accident month and truncate each component. This gives rise to the truncated multinormal distribution. Tallis (1961) derived the mgf, mean and variance-covariance matrix of this distribution. However, the calculations require the evaluation of multivariate normal integrals, not a simple task.

5   Non-Stationary Time Series

The theory presented thus far assumed that the time series was stationary, i.e. the mean of the process and the variance of the errors were constant over time. The stationarity assumption for the number of claims is not usually valid for a new line of business or a line subject to rapid growth; in such a case, it would be preferable to model claim frequency instead of counts. The assumption can be verified by plotting the observations against time. Other situations which vary from the stationary conditions can sometimes be accomodated in the Box-Jenkins method.

When the mean of the process increases linearly over time, differencing the original series will produce a new series which has a constant mean, and the theory developed previously can be applied to it. If an insurer experiences a constant growth in business, reflected in an exponentially increasing number of accidents, a logarithmic transformation of the data, followed by a differencing of the series will make it stationary. If the original data in the time series have a standard deviation which is proportional to its level, a logarithmic transformation will also make it stationary. When the inverse retransformation is used to compare with the original time series, care must be taken, because the forecasts will be biased. Pankratz and Dudley (1987) show how to correct for this bias.

It is conceivable that seasonality effects could affect certain lines of insurance. For example, in automobile insurance, the number of claims for damages to cars could increase during the winter months, due to bad weather conditions. These seasonal models can also be incorporated in the general framework of an ARMA process, by differencing corresponding monthly numbers, in successive years.

To get confidence intervals for the estimates of the claims numbers, the assumption of normality of the errors was essential. Without this assumption, for a weakly stationary time series (see section 2), we can obtain least-squares estimates of the parameters and best linear predictors for future values. However, we can not get confidence intervals based on the normal distribution nor on the truncated distribution.

6   Conclusions

In this paper, we have shown how to analyze the number of claims incurred in past accident periods to forecast the number for future periods. Additional information could also be taken into account to get better estimates. From these, the number of IBNR claims could be forecasted.

If the information available on the claims numbers or the claims amounts could be put into the usual claim run-off triangle format, a more traditional method of analysis like the chain-ladder method could be employed.

References

Abraham, B. and Ledolter, J. (1983). *Statistical Methods for Forecasting.* Wiley, New York.

Anderson, O.D. (1976). *Time Series Analysis and Forecasting: The Box-Jenkins Approach.* Butterworth & Co., London.

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, Oakland.

Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association*, **65**, 1509-1526.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods.* Springer-Verlag, New York.

Harvey, A. (1981). *Time Series Models.* Philip Allan Publishers Limited, Oxford.

Hillier, F.S. and Lieberman, G.J. (1986). *Introduction to Operations Research.* Holden-Day Inc., Oakland, California.

IMSL (1987). *Users Manual, Math/Library Fortran Subroutines for Mathematical Applications.* Version 1.0, Houston.

ITSM (1991). An Interactive Time Series Modelling Package for the PC. Brockwell, P.J. and Davis, R.A., Springer-Verlag.

Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions, Vol. 1* , Houghton Mifflin Company, Boston.

Johnson, N.L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* John Wiley and Sons, Inc., New York.

Lemaire, J. (1982). Claims provisions in liability insurance. *Journal of Forecasting*, **1**, 303-318.

Luenberger, D.G. (1984). *Linear and Nonlinear Programming.* Addison-Wesley, Reading, Massachusetts.

Pankratz, A. (1983). *Forecasting With Univariate Box-Jenkins Models: Concepts and Cases.* John Wiley and Sons, New York.

Pankratz, A. and Dudley, V. (1987). Forecasts of Power-transformed Series. *Journal of Forecasting*, **6**, 239-248.

Patel, J.K. and Read, C.B. (1982). *Handbook of the Normal Distribution.* Marcel Dekker, Inc., New York.

Tallis, G.M. (1961). The moment generating function of the truncated multi-normal distribution. *J. Roy. Statist. Soc. Ser. B*, **23**, 223-229.

Taylor, G.C. (1986). *Claim reserving in non-life insurance.* North-Holland, Amsterdam.

Van Eeghen, J. (1981). *Loss reserving methods.* Volume No. 1 in the series Surveys of actuarial studies, published by Nationale-Nederlanden N.V., Rotterdam.

Verrall, R.J. (1989). Modelling claims runoff triangles with two-dimensional time series. *Scandinavian Actuarial Journal*, 129-138.