ASSESSING THE PREDICTIVE ACCURACY OF BAYESIAN METHODS

Milan Marček[1,2]

[1]Institute of Computer Science, Faculty of Philosophy and Science, The Silesian University Opava

[2]The Faculty of Management Science and Informatics, University of Žilina

Abstract

A recurring problem concerning marketing operatives at all levels is a forecast of the demand for new products. This article discusses a technique, Bayesian method, by which the forecast of such products can be calculated. This technique is used to examine the forecast accuracy of lightning current arristers during the year 2004.

***Key words**: Bayesian estimation, prior and posterior distribution, ex ante forecast, double exponential smoothing.*

1 Introduction

Much Econometric and statistical forecasting techniques are widely used in managing production systems. They have also found frequent application in a variety of other problem areas, including financial planning, investment analysis, distribution planning, marketing. For example, consider the two common economic theories of demand and production. Each one is based on elementary economic principles and, if enough historical data are available, provides a basis for a theoretical multivariate regression model. Usually the demand function for a good relates the quantity of a good sold to several explanation variables. The theoretical foundation for demand analysis is based on the modern theory consumer behaviour. In determining the demand the important roles play the price of the good and the consumer's real income. Other variables that should be included are the price of other goods (either substitutes or complements) in consumption, the size of the market, the expected future price of the good and tastes and preferences.

Nevertheless, it is likely that the price of the good, the income, the size of market and the price of substitutes play the most important role in determining demand. It is important to remember that the model is being developed for forecasting purposes. Therefore, it is extremely important to keep the model relatively simple in terms of the number of the explanatory variables. Remember that, for each period in an ex ante forecast of the dependent variable, it is also necessary to provide a forecast for each of the independent variables.

Most dependent variables may be related to time or may be explained by simple mathematical functions of time. Regression models are statistical techniques for modelling an investigating the relationship between two or more variables. As we mentioned above, regression analysis requires that enough historical data of dependent and independent variables are available. In many decision problems there are little or non useful historical information available at the time of decision making. In such situations Bayesian methods are often useful to determine the early forecast.

The next section contains a brief description of forecasting models based on Bayesian approach that are linear in the unknown parameters. The data, fitted model and prediction results of the Bayesian procedure are given in section 3. Assessing the predictive accuracy of models using Bayesian procedures and exponential smoothing methods is discussed in section 4. The final section contains a summary of our results.

2 Models

Retail and whole sales are important components of the business company. One of the most basic problems facing the marketing or product manager is how to predict sales of new products in situations where no historical data are available. Marketing decisions must be made in the context of insufficient information about processes that are dynamic, stochastic and downright difficult.

This article discusses a technique, Bayesian method in forecasting, which should prove helpful in modelling of such situations. The methodology of Bayesian method, i.e. Bayesian parameter estimation and prediction for general time series model is described in [5]. We will this methodology apply and further extend for general (causal) model.

The functional form of the causal regression model has the form

$$y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + \dots, b_k x_{kt} + u_t\,, \tag{1}$$

where $\{x_i\}_{i=1}^k$ represent a series of independent variables for i = 1, 2, ..., k and t = 1, 2, ..., n, $\{b_i\}_{i=1}^k$ are partial regression coefficients (the model parameters), $\{y_t\}_{t=1}^n$ is dependent variable, $\{u_t\}_{t=1}^n$ is random error term.

It is convenient to express the model (1) in matrix notation of the following form

$$\mathbf{y} = \mathbf{Xb} + \mathbf{u}\,, \tag{2}$$

where $\mathbf{y} = \{y_t\}_{t=1}^n$ is a column vector of the n dependent values, $\mathbf{b} = (b_0, b_1, ..., b_k)$ is a column vector of the k + 1 unknown parameters, $\mathbf{u} = \{u_t\}_{t=1}^n$ is an column vector of the errors, and $\mathbf{X}$ is a matrix of the independent values of the form

$$\mathbf{X} = \begin{pmatrix} 1, & x_{11}, & x_{21}, ..., x_{k1} \\ 1, & x_{21}, & x_{22}, ..., x_{k2} \\ & \dots & \\ 1, & x_{1n}, & x_{2n}, ..., x_{kn} \end{pmatrix}$$

As a practical example, consider a firm which mass-produces new systems of lightning current arresters. The marketing manager wishes to forecast monthly sales of these products of new unknown foreign markets. The functional form of this process relates the size of sales to two explanatory variables: $\{x_{1t}\}_{t=1}^n$ - the advertising expenditures, $\{x_{2t}\}_{t=1}^n$ - service and cost of repairs. No similar item has been sold in the past. Thus, we have the model

$$y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + u_t\,, \quad t = 1, 2, ..., n \tag{3}$$

The manager expects that the sales during first three months do not fall under the specified level. Using this information the last-squares model parameter estimates $\hat{\mathbf{b}} = (\hat{b}_0, \hat{b}_1, \hat{b}_2)$ can be calculated.

The model (3) is the classical linear regression equation. To generate forecasts of future observations using Bayes methodology we will assume that variance of the random component σ_u^2 is known. Then the vector $\hat{\mathbf{b}} = (\hat{b}_0, \hat{b}_1, \hat{b}_2)$ follows the multivariate normal prior distribution with mean $E(\mathbf{b}) = \overline{\mathbf{b}}' = (\overline{b}_0', \overline{b}_1', \overline{b}_2')$, the variance $var\,(\{b_i\}_{i=1}^k) = \sigma_{b_i'}^2$ and the covariance $cov(b_i', b_j')$, i.e.

$$N(\overline{\mathbf{b}}', \mathbf{V}') \equiv (2\pi)^{-(\frac{1}{2})} \left|\mathbf{V}'^{-1}\right|^{1/2} \exp\{-\tfrac{1}{2}[\mathbf{b} - \overline{\mathbf{b}}']^T \mathbf{V}'^{-1} [\mathbf{b} - \overline{\mathbf{b}}']\}\,, \tag{4}$$

where $\mathbf{V}'$ is the variance-covariance matrix of the prior distribution. This probability distribution measures our subjective information above $\mathbf{b}$.

After one period, or generally after T periods, we have observed $y_1, y_2, ..., y_T$. The problem is to modify the estimate $\overline{\mathbf{b}}'$ and the measure of uncertainty expressed by $var\,(\{b_i\}_{i=1}^k)$ in light of this information. This can be done using the normal posterior of the parameters with mean $\overline{\mathbf{b}}''$ and variance-covariance matrix $\mathbf{V}''$ as [5]

$$\overline{\mathbf{b}}'' = \mathbf{G}''^{-1}(\mathbf{G}'\overline{\mathbf{b}}' + \mathbf{X}^T\mathbf{X}\hat{\mathbf{b}}), \quad (5)$$

where $\mathbf{G}' = \sigma_u^2\mathbf{V}'$, $\mathbf{G}'' = \mathbf{G}' + \mathbf{X}^T\mathbf{X}$. From the expression (5) is seen that the parameters of the posterior can be determined from simple algebraic combination of the prior parameters and the results of last-squares analysis of the time series data.

3 The Data and Results

The data were taken from 10 consecutive months of lightning current arresters purchasing history (y_t). The data for each independent variable (x_{1t}, x_{2t}) are collected by firm (see [3]). Table 1 presents the monthly data for all variables (y_t - lightning current arristers purchasing, x_{1t} - the advertising expenditures, x_{2t} - service and cost of repairs).

TABLE 1: DEPENDENT, INDEPENDENT AND FORECAST VALUES WHEN BAYESIAN METHOD IS USED FOR A CAUSAL LINEAR MODEL (SEE TEXT FOR DETAILS)

T	1	2	3	4	5	6	7	8	9	10
y_t	4101	4132	4158	4188	4214	4193	4272	4281	4301	4317
x_{1t}	121	121	128	120	130	1125	138	140	142	140
x_{2t}	97	99	108	113	110	105	114	116	120	121
$\hat{y}_t$				4179.2	4212.1	4171.2	4265.8	4284.0	4308.3	4302.3

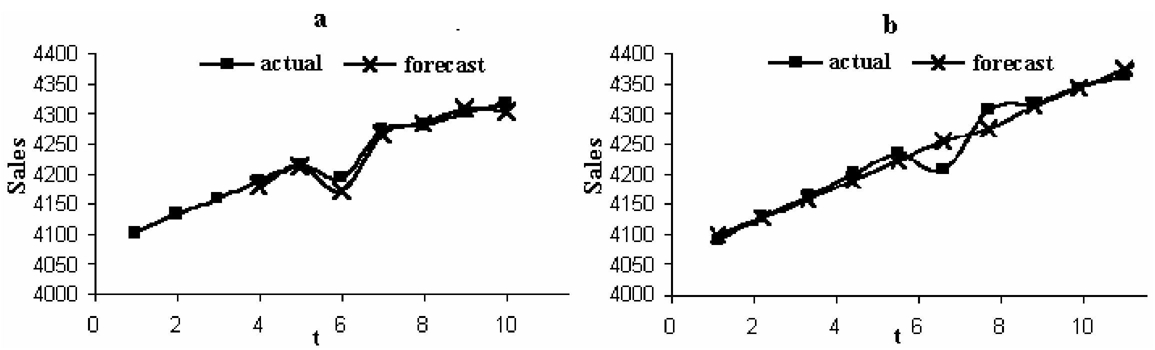

Figure: 1: Sales – actual and forecast values a) Bayesian method b) single exponential smoothing

The calculation of forecast values for a given example can be summarised as follows (detailed procedures can be found in [5]

1. *Assumptions of the model - prior distribution.*

The manager supposes that the initial sales rate will be between 3300 – 4300 per month. But with the advertising program planed, the manager expects the sales rate to increase steadily and estimate the amount of increase from month to month at between 80 – 120. Analogously, with the service program planed, the amount of steadily increase of the sales rate from month to month is estimated at between 100 – 140. The assumption about disturbance term u_t is, $u_t \equiv N(0,\sigma_u^2)$ with σ_u^2 = 2500. The prior variances of the parameters are assumed to be one-sixth of the ranges, i.e. $\sigma^2_{b'_0}$ = (4300 – 3300)^2/6 = 27777.778. Analogously $\sigma^2_{b'_1}$ = (140 – 100)^2/6 = 44.44, $\sigma^2_{b'_2}$ = (120 – 80)^2/6 = 44.44 and $\mathrm{cov}(b'_0, b'_1)$ = $\mathrm{cov}(b'_1, b'_2)$ = $\mathrm{cov}(b'_0, b'_2)$ = 0. Therefore, the prior variance-covariance matrix is

$$\mathbf{V}' = \begin{pmatrix} \sigma^2_{b'_0} & & 0 \\ & \sigma^2_{b'_1} & \\ 0 & & \sigma^2_{b'_2} \end{pmatrix} = \begin{pmatrix} 27777.78 & & 0 \\ & 44.44 & \\ 0 & & 44.44 \end{pmatrix}$$

2. *Posterior parameters at the time T = 3*:

- The least-squares estimates are calculated from the first $T = 3$ observations as

$$\hat{\mathbf{b}}^T = (\hat{b}_0, \hat{b}_1, \hat{b}_2) = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

$$= (45559.43,\ 15.5,\ 16.214).$$

- Using the formula (5), the posterior parameters of $\mathbf{b}$, given $\hat{\mathbf{b}}$, are calculated as

$$\bar{\mathbf{b}}'' = E(\mathbf{b}'') = \mathbf{G}''^{-1}(\mathbf{G}'\bar{\mathbf{b}}' + \mathbf{X}^T\mathbf{X}\hat{\mathbf{b}})$$

$$= (3225.031,\ 3.926,\ 4.255)$$

- The forecast for period $t = T + \tau$, for any lead time $\tau \geq 1$, we simply take expectation at origin T of

$$y_{T+\tau} = \bar{b}'' + \bar{b}_1'' x_{1,T+\tau} + \bar{b}_2'' x_{2,T+\tau} + u_{T+\tau}.$$

For $\tau = 1$

$$E(y_{T+1}) = \hat{y}_{T+1} = \bar{b}'' + \bar{b}_1'' x_{1,T+1} + \bar{b}_2'' x_{2,T+1}$$

$$\hat{y}_4 = \bar{b}'' + \bar{b}_1'' x_{1,4} + \bar{b}_2'' x_{2,4} = 4179.2$$

- Forecast for each future observation may be easily computed as a new observation becomes available through the three procedures above. Generally, we denote the current period by T. The forecast $\hat{y}_{T+\tau}$ for future observations $y_{T+\tau}$, for any lead time $\tau \geq 1$ is generated successively by setting the new origin of time equal to $T + \tau - 1$ and computing the new forecast for period $T + \tau$. These forecasts are shown in Tab. 1 and illustrated in Fig. 1a.

4 Assessing the Predictive Accuracy of the Alternative Forecasting Model

In this section, we shall present the forecast results of the sales using double exponential smoothing approach [4]. The results of updating the forecasts of sales using double exponential smoothing are presented in Fig. 1b. The weighting factor α was equal to 0.10.

Fig. 1a illustrates a good quantitative model. This model closely follows all the actual data. The model based on the double exponential smoothing (Fig. 1b) fails to fit the actual swing at the period $t = 6$.

5 Conclusion

We have illustrated that the model based on Bayesian methodology can be useful tool for economic time series modelling and forecasting. The Bayesian technique is attractive for forecasting because it provides a framework within which the time series can be successful modelled by only little or no observations. Finally, Bayesian approach is also attractive from a strictly model building point of view because: the final model will, in general, be a parsimonious representation of the data, the appropriate noise structure is easily determined and has relatively simple procedures for revising the parameters.

The disadvantage of models based on Bayesian approach is that the variance of the disturbance term must be known. For unknown variance case see [6] Some difficulties arise by nonlinear and seasonal models see [1], [5].

Acknowledgement: This work was supported by Slovak grant foundation under the grant No. 1/2628/05 and from the Grant Agency of the Czech Republic under the grant No. 402/05/2768.

References

[1] CARTER, R., F. *A Bayesian Approach to Seasonal Style Goods Forecasting*. Unpublished M. S. thesis, Georgia Institute of Technology, Atlanta, 1971

[2] CRAVENS, D., W. *Strategic Marketing*. The Mc Graw-Hill Companies, Inc. 1997

[3] KUBALOVÁ. E.: *Forecasting of uncertain events by Byesian methods in marketing*. Diploma work, Žilina: ŽU FRI, June 2004. (in Slovak)

[4] MARČEK, D., MARČEK, M. *Time series analysis and forecasting with applications in economics*. Žilina: ES ŽU, 2001 (in Slovak). ISBN 80-7100-870-2.

[5] MONGOMERY, D., C., JOHSTON, L., A., GARDINER, J., S. *Forecasting and Time Series Analysis*. Mc Graw-Hill, Inc. 1990. ISBN 0-07-042858-1.

[6] RAIFFA, H., SCHLAIFER, R. *Applied Statistical Decision Theory*. Harward University Press, Cambridge, Mass., 1961.

Author:
Milan Marček:
e-mail: marcek@fria.utc.sk