APPLICATION OF AN EMPIRICAL BAYES TECHNIQUE TO TELEPHONE TRAFFIC FORECASTING

R. N. PANDYA, W. R. ROBINSON
Bell-Northern Research, Ottawa, Canada

ABSTRACT

This paper discusses the feasibility of using an empirical Bayes technique for forecasting telephone-subscriber traffic. The specific technique considered is derived from the James-Stein estimator. After a brief discussion of current forecasting practice, the paper provides a description of the James-Stein estimator. Its application to forecasting subscriber traffic is explained and comparative results for a group of eight local crossbar switching offices are presented and discussed. The results indicate that the method is simple to implement, requires little additional data and provides improved forecasting accuracy.

1. INTRODUCTION

For efficient planning and provisioning of switching offices, a large telephone company like Bell Canada relies on accurate estimates of:

(a) future demand for telephone connections, and
(b) future pattern of telephone usage.

The estimates of these two variables are then combined to provide an estimate of future telephone traffic. Since the planning function is carried out at various organization levels (Switching office, District, Area, Region and Company), forecasts for the above variables are made at each of these levels. Further, the forecasts are needed one to three years in advance, the time it takes to plan and implement necessary additions.

The future demand for telephone connections (Main Stations Forecast) is generally estimated from relevant economic and business indicators by fitting suitable multiple regression models [1]. Even though the number of main stations served by a switching office may be fixed, the traffic load on the office will exhibit hourly, daily and seasonal variations. The quantification of 'telephone usage' is therefore based on the concept of Busy Season Busy Hour (BSBH) traffic load [2]. Under this concept, a single-parameter representation of telephone usage is provided by the Average Busy Season Busy Hour (ABS BH) traffic per main station (CCS/MS). The ABS BH CCS/MS measure of telephone usage is obtained by averaging approximately sixty data points.

At the switching office level, the CCS/MS forecasts are presently generated through a two-step procedure. An initial or base forecast is obtained by fitting a suitable trend line to the past CCS/MS data for the office and then projecting it into the future. This base forecast is called the Trend View. The final forecast, called the Traffic View, is obtained by applying an adjustment factor to the base forecast [3], [4]. The adjustment is based on informed judgement and reflects anticipated changes in the office characteristics.

Results on the sensitivity of capital investment to errors in CCS/MS forecasts and on the forecasting accuracy of the present methods were reported in [3] and [4]. The results indicate that the simple trend procedure does not track the movements in CCS/MS with sufficient accuracy, thus leading to relatively poor forecasts. This is specially noticeable in the case of crossbar and stored program controlled switching offices.

To obtain improved CCS/MS forecasts, the simple trend procedure needs to be replaced by a more suitable alternative. Some of the factors that should be considered in evaluating a suitable replacement are:

- forecasting accuracy,
- ease of implementation,
- changeover costs,
- personnel retraining requirements,
- additional data requirements, and
- suitability for top-down bottom-up forecasting.

The next section briefly describes the empirical Bayes method and its application to CCS/MS is outlined in Sections 3 and 4. Comparative results on CCS/MS forecasts for eight local crossbar switching offices are presented and discussed in Section 5. Conclusions and plans for future work are summarized in Section 6.

2. EMPIRICAL BAYES TECHNIQUE

Classical statistics provides inferential or estimation techniques based entirely on the sample information. The Bayesian approach introduces an additional ingredient - the prior distribution. The prior distribution embodies the a priori information that may be available about the parameters being estimated. If the performance of the Bayesian estimator is to be judged on the basis of a suitable 'loss' or 'risk' function, then we enter the realm of Bayesian Decision Theory and the resulting estimators are sometimes referred as Decision Functions.

In practice, an exact knowledge of the prior distribution is not always available, thus restricting the application
of Bayesian methods. However, in many instances, the defining parameters of the prior distribution (mean and variance) can be estimated from the available data. Bayesian procedures that utilize such 'estimated priors' are classified as empirical Bayes techniques [5]. The empirical Bayes technique considered for the present application is a form of the so called James-Stein estimator. The James-Stein estimator can be used advantageously when the following conditions apply:

(a) more than three parameters are to be estimated,
(b) the true values of the parameters to be estimated are relatively close together, and
(c) the loss function to be minimized is a composite loss function which is additive over the set of parameters being estimated.

The problem of estimating future CCS/MS values meets the above conditions in so far as:

(a) CCS/MS is to be estimated for all switching offices in the company,
(b) switching offices with similar characteristics can be grouped together, and
(c) the operating company would be interested in a procedure that reduces the sum of estimation errors over all the switching offices.

An interesting discussion of the James-Stein estimator with simple examples is given in [6]. Theoretical foundations for the technique and its interpretation from a Bayesian viewpoint are provided in [7] and [8]. Its application to forecasting inventory demands is described in [9]. A very brief and formal description of the James-Stein estimator follows.

3. THE JAMES-STEIN ESTIMATOR

The James-Stein estimator can be used when several means are to be estimated and the loss function to be minimized is the sum of squared errors of estimation for the means. The James-Stein estimator Z for the mean \( \mu_i \) of the \( i \)th random variable \( X_i \) \((i = 1, 2, 3, \ldots, k; k > 3) \) has the general form:

\[
Z_i = (1-S_i) \bar{X} + S_i \bar{Z} \quad i = 1, 2, 3, \ldots, k
\]

where

\[
\bar{X}_i = \text{sample mean of random variable } X_i,
\]

\[
\bar{X} = \frac{1}{k} \sum_{j=1}^{k} X_j = \text{Group Mean},
\]

and \( S_i \) is the 'shrinking factor' associated with each mean to be estimated. The shrinking factor determines the degree of shrinking of the sample mean \( \bar{X}_i \) towards the Group Mean \( \bar{X} \) to provide the final James-Stein estimate \( Z_i \).

The expression that is generally used for \( S_i \) has the form

\[
S_i = \frac{(k-3) \sigma_i^2}{\sum_{j=1}^{k} (X_j - \bar{X})^2}
\]

where \( \sigma_i^2 \) is the sample variance of random variable \( X_i \).

The above form for the shrinking factor has the disadvantage that \( S_i \) can assume values that is greater than unity, thus moving \( Z_i \) away from the Group Mean \( \bar{X} \). It is suggested in (8) that \( S_i \) should be replaced by unity in such cases.

We have discovered that the following form for \( S_i \) yields more satisfactory results:

\[
S_i = \frac{1}{k} \frac{\sigma_i^2}{\sum_{j=1}^{k} (X_j - \bar{X})^2 + \sigma_i^2}
\]

The above expression not only ensures that \( S_i \) takes values between zero and one but it also has some intuitive appeal. If the first term in the denominator is interpreted as an estimate of the dispersion among the \( k \) population means, and the second term as the dispersion in the observed sample for the random variable, then \( S_i \) has the form:

\[
\text{Shrinking Factor} = \frac{\text{Sample Variance}}{\text{Group Variance} + \text{Sample Variance}}
\]

Thus \( Z_i \), the James-Stein estimate of \( \mu_i \) given by equation (2), will shrink more towards the Group Mean if the sample variance is relatively large (less faith in observed sample) or if the group variance is small (individual means are close together). On the other hand, if the sample variance is relatively small, we place more reliance on the observed sample and the shrinking towards the Group Mean is marginal.

From a Bayesian viewpoint, the group mean and the group variance represent the estimated priors. The James-Stein estimator thus belongs to the class of empirical Bayes estimators because it combines the sample information (sample mean and the sample variance) with the estimated priors (group mean and group variance) to provide the final estimates.

4. APPLICATION OF JAMES-STEIN METHOD TO CCS/MS FORECASTING

As mentioned in Section 1, the current practice for CCS/MS forecasting assumes that the past history of an office, as reflected in the trend line, has a significant influence on the future movements of CCS/MS. However, results of extensive data analysis carried out as part of this study as well as the results of multiple regression studies reported in [3] and [4] fail to support this hypothesis – especially for crossbar and stored program controlled offices. The results indicate that movement in CCS/MS for these offices is primarily dependent on the observed CCS/MS for the preceding year and one or more usage-related variables (e.g. percent business lines, percent toll messages). The proposed forecasting procedure based on the James-Stein technique accounts for the effect of the usage-related variables by:

(a) grouping together offices with 'similar' characteristics.
(b) defining a single CCS/MS figure which is representative of this group of offices (Group Average), and
(c) using a forecast which is a weighted sum of the observed CCS/MS and the Group Average.

Assuming that there are N switching offices in the group (N≥3) and that \( \bar{C}_i(T) \) represents the observed ABBH CCS/MS for office \( i \) in year \( T \), the proposed forecasting equation has the form:

\[
\hat{C}_i(T+1) = (1-S_i) \bar{C}_i(T) + S_i \bar{C}(T); \quad i = 1, 2, \ldots, N \quad (5)
\]

where

\[
\bar{C}_i(T) = \frac{1}{M} \sum_{j=1}^{N} C_j(T); \quad \text{where}
\]

\[
M = \text{Total Number of Business Days in the Busy Season in Year } T.
\]

\[
C_j(T) = \text{ith observation on ABBH CCS/MS for office } j.
\]

\[
\bar{C}(T) = \frac{1}{N} \sum_{i=1}^{N} \bar{C}_i(T)
\]

= Group Average

\[
S_i = \frac{1}{N-3} \sum_{k=1}^{N} (C_k(T) - \bar{C}(T))^2 + \sigma_i^2
\]

= Shrinking Factor for Office \( i \)

\[
\sigma_i^2 = \frac{1}{M-1} \sum_{j=1}^{M} (C_j(T) - \bar{C}_j(T))^2 \quad (7)
\]

= Sample Variance of the ABBH CCS/MS data

where \( \sigma_i^2 \) is the sample variance of the ABBH CCS/MS data, i.e.,

\[
\sigma_i^2 = \frac{1}{M-1} \sum_{j=1}^{M} (C_j(T) - \bar{C}_j(T))^2
\]

Projections for more than one year can be obtained by using a recursion on (5).

5. RESULTS

The performance of the James-Stein method was initially tested on a group of eight local, crossbar switching offices. Except for the fact that they are all local crossbar switching centres, the offices in the group are quite diverse. For example, the size of the offices varies from 1,717 to 29,590 lines, the usage per main station varies from 1.99 to 2.67 CCS, the percentage of residential lines varies from 71% to 94% and their locations vary from city core to rural.

For this group of offices, forecasting performance for the proposed James-Stein method was compared with the Trend View, the Traffic View (Adjusted Trend View) and the Random Walk model (forecast = latest observation). The basis of comparison is the total RMSE (Root Mean Squared Error) over the group of eight offices and the performance is compared for one, two and three year forecasts. The results are given in Table 1.

The Traffic View represents the projections made by the traffic personnel at Bell Canada using the Trends View as the base or initial forecast. The Traffic View provides a significant improvement over the Trends View. This improvement reflects the additional information and insight contributed by the traffic personnel associated with these switching centres. Whichever mathematical model is used to generate the base forecasts, the final forecast should take into account the informed judgement of the field personnel.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELATIVE FORECASTING PERFORMANCE OF TREND VIEW, TRAFFIC VIEW, RANDOM WALK MODEL AND JAMES-STEIN METHOD FOR EIGHT LOCAL CROSSBAR OFFICES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FORECASTING METHOD</th>
<th>ONE-YEAR-AHEAD FORECASTS</th>
<th>TWO-YEAR-AHEAD FORECASTS</th>
<th>THREE-YEAR-AHEAD FORECASTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORECASTS</td>
<td>77 IN 77 IN 76</td>
<td>77 IN 77 IN 76</td>
<td>77 IN 77 IN 75</td>
</tr>
<tr>
<td>TREND VIEW</td>
<td>.369</td>
<td>.502</td>
<td>.672</td>
</tr>
<tr>
<td>TRAFFIC VIEW</td>
<td>.402</td>
<td>.510</td>
<td>.703</td>
</tr>
<tr>
<td>RANDOM WALK</td>
<td>.272</td>
<td>.385</td>
<td>.544</td>
</tr>
<tr>
<td>JAMES-STEIN</td>
<td>.225</td>
<td>.333</td>
<td>.378</td>
</tr>
</tbody>
</table>

6. CONCLUSIONS

- Results on the tested sample indicate that the proposed James-Stein method can lead to significant reductions in overall forecasting errors.
- The procedure is relatively simple and no additional data is required.
- Though the procedure should be applied to a group of similar offices, initial tests indicate that the method is quite robust with regard to changes in groupings.
- The Bayesian structure of the forecasting technique can provide a conceptual basis for reconciling top-down and bottom-up forecasts. For example, the James-Stein procedure could be used on all offices in an area with the Group Average being suitably modified to reflect the CCS/MS viewed from the areal level.

Work is currently underway to test the procedure on a larger sample of switching offices and compare its performance with multiple regression methods. Work is also planned towards evaluating the impact of group boundaries on the performance of the proposed procedure and towards evolving a top-down bottom-up approach to CCS/MS forecasting. Additional results will be available for presentation at the Ninth International Teletraffic Congress.

7. REFERENCES

[2] Blair, N.D., 'Forecasting Telephone Traffic in the
Bell System', Eighth International Teletraffic Congress, Melbourne, Australia, November 1976.


