



STATISTICAL EVALUATION OF FORECASTING METHODS.

Sanjeev Kumar Khichi

Associate Professor, Community Medicine, SHKM Govt. Medical College, Nalhar, Nuh, Haryana

Dr.Nand Kishore Singh*

Associate Professor & Statistician, Community Medicine, SHKM Govt. Medical College Nalhar, Nuh, Haryana *Corresponding Author

Neha Nehra

Research Scholar in Sun Rise University.

Sarita Dahiya

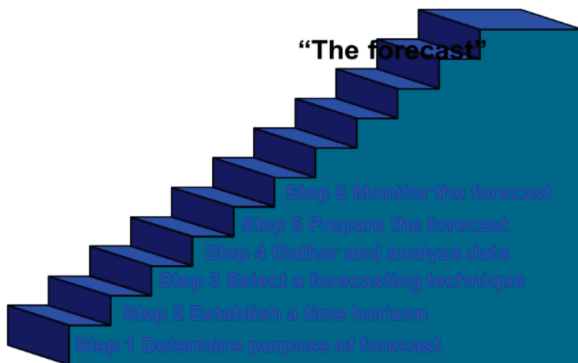
Assistant Professor, Psychology, Govt. College for Women, Rewari, Haryana

ABSTRACT

In this paper some systematic steps were used for forecasting. Generate forecasts for data with different patterns: level, trend, seasonality, and cyclical. Describe causal modeling using linear regression. Compute forecast accuracy. Explain how forecasting models should be selected. Decide what needs to be forecast. Level of detail, units of analysis & time horizon required. Evaluate and analyze appropriate data. Identify needed data. Select and test the forecasting model. Cost, ease of use & accuracy. Generate the forecast accuracy over time.

KEYWORDS :

Evaluation Steps



- Association between the predictor and the predicted
- Predictor variables - used to predict values of variable interest, sometimes called independent variables
- Predicted variable = Dependent variable
- Regression - technique for fitting a line to a set of points
- Linear regression is the most widely used form of regression
- The objective is to obtain an equation of a straight line that minimizes the sum of squared vertical deviations of data points from the line.
- $y = \alpha + bx$
- Where
- y = predicted (dependent) variable
- x = predictor (independent) variable
- b = slope of the line
- α = value of y when $x = 0$ (the height of line at the y intercept)

Given n data points, find the intercept a and the slope b to

Minimize the sum of squared errors =

Minimize the sum of deviations from the line =

$$\text{Minimize } \sum_{i=1}^n (y_i - a - bx_i)^2$$

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad a = \frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n x_i}{n}$$

- Time-ordered sequence of observations taken at regular intervals over a period of time
- Future values of the series can be estimated from past values.

Types of Variations in Time Series Data

- *Trend* - long-term movement in data
- *Seasonality* - short-term regular variations in data
- *Cycles* - wavelike variations of long-term
- *Irregular variations* - caused by unusual circumstances
- *Random variations* - caused by chance
- Forecast error: = Actual - Forecast = $A(t-1) - F(t-1)$

$$F_t = F_{t-1} + \alpha (A_{t-1} - F_{t-1})$$

Forecast today = Forecast yesterday + (alpha) * (Forecast error yesterday) Each new forecast is equal to the previous forecast plus a percentage of the previous error.

$$\text{Forecast error} = \text{Actual} - \text{Forecast}$$

$$MAD = \frac{\sum_{i=1}^n |A_i - F_i|}{n}$$

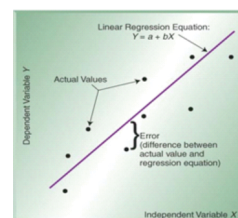
	Qualitative Methods	Quantitative Methods
1. Characteristics	Based on human judgment, opinions; subjective and nonmathematical.	Based on mathematics; quantitative in nature.
2. Strengths	Can incorporate latest changes in the environment and "inside information."	Consistent and objective; able to consider much information and data at one time.
3. Weaknesses	Can bias the forecast and reduce forecast accuracy.	Often quantifiable data are not available. Only as good as the data on which they are based.

- **Step 1 - Smoothing the level of the series**

$$S_t = \alpha A_t + (1 - \alpha)(S_{t-1} + T_{t-1})$$

- **Step 2 - Smoothing the trend**

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$



$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$a = \bar{Y} - b\bar{X}$$

$$Y = a + bX$$

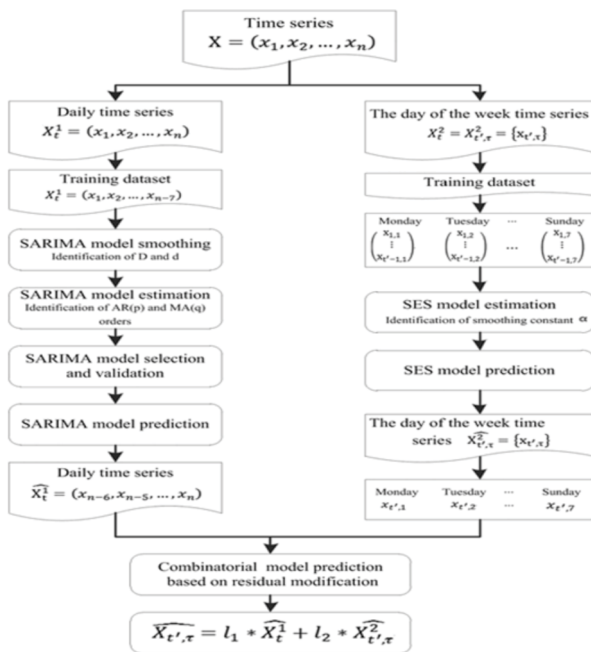
Forecast Software

- Spreadsheets
- Microsoft Excel, Quattro Pro, Lotus 1-2-3
- Limited statistical analysis of forecast data
- Statistical packages
- SPSS, SAS, NCSS, Minitab
- Forecasting plus statistical and graphics
- Specialty forecasting packages
- Forecast Master, Forecast Pro, Autobox, SCA

Guidelines for Selecting Software

- Does the package have the features you want?
- What platform is the package available for?
- How easy is the package to learn and use?
- Is it possible to implement new methods?
- Do you require interactive or repetitive forecasting?
- Do you have any large data sets?
- Is there local support and training available?
- Does the package give the right answers?

ARIMA and SES model for forecasting



DISCUSSION

Three basic principles of forecasting are: forecasts are rarely perfect, are more accurate for groups than individual items, and are more accurate in the shorter term than longer time horizons. The forecasting process involves five steps: decide what to forecast, evaluate and analyze appropriate data, select and test model, generate forecast, and monitor accuracy. Forecasting methods can be classified into two groups: qualitative and quantitative. Qualitative methods are based on the subjective opinion of the forecaster and quantitative methods are based on mathematical modeling.

Time series models are based on the assumption that all information needed is contained in the time series of data. Causal models assume that the variable being forecast is related to other variables in the environment. There are four basic patterns of data: level or horizontal, trend, seasonality, and cycles. In addition, data usually contain random

variation. Some forecast models used to forecast the level of a time series are: naïve, simple mean, simple moving average, weighted moving average, and exponential smoothing. Separate models are used to forecast trends and seasonality. A simple causal model is linear regression in which a straight-line relationship is modeled between the variable we are forecasting and another variable in then environment. The correlation is used to measure the strength of the linear relationship between these two variables. Three useful measures of forecast error are mean absolute deviation mean square error and tracking signal. There are four factors to consider when selecting a model: amount and type of data available, degree of accuracy required, length of forecast horizon, and pattern.

REFERENCES

1. Kahneman Daniel; Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk" (PDF). *Econometrica*. 47 (2): 263–291.
2. Jump up to: a b Kahneman, Daniel; Tversky, Amos (1977). "Intuitive prediction: Biases and corrective procedures" (PDF). *Decision Research Technical Report PTR-1042-77-6*. In Kahneman Daniel; Tversky, Amos (1982). "Intuitive prediction: Biases and corrective procedures". In Kahneman, Daniel; Slovic, Paul; Tversky, Amos (eds.). *Judgment Under Uncertainty: Heuristics and Biases*. pp.414
3. "Week 10: Reference Class Forecasting". *Conceptually*. Retrieved 20 April 2017.
4. "Outside view". *Less Wrong Wiki*. Retrieved 20 April 2017.
5. Flyvbjerg, Bent (2006). "From Nobel Prize to Project Management: Getting Risks Right". *Project Management Journal*. 37 (3): 5–15. arXiv:1302.3642. Bibcode:2013arXiv1302.3642F. SSRN 2238013.
6. "Council to borrow £231m for Edinburgh trams project". *BBC News*. BBC. 19 August 2011. Retrieved 19 August 2011.
7. AACE International, "Total Cost Management Framework, Section 7.3, Cost Estimating and Budgeting", 2011. p. 147
8. Marrow, E. and Yarossi, M., "Assessing Project Cost and Schedule Risk", 1990 AACE Transactions. pp H.6.1-7
9. Lovallo, D; Kahneman, D (2003). "Delusions of success. How optimism undermines executives' decisions". *Harvard Business Review*. 81 (7): 56–63. PMID 12858711.
10. Flyvbjerg, Bent (2008). "Public Planning of Mega-Projects: Overestimation of Demand and Underestimation of Costs". In Priemus, Hugo; Flyvbjerg, Bent; van Wee, Bert (eds.). *Decision-making on Mega-Projects*. doi:10.4337/9781848440173.00014. ISBN 9781848440173.
11. Flyvbjerg, Bent (2011). "Over Budget, Over Time, Over and Over Again: Managing Major Projects". In Morris, Peter W. G; Pinto, Jeffrey K; Söderlund, Jonas (eds.). *The Oxford Handbook of Project Management*.