Index

Section A. Planning, Budgeting and Forecasting
Section A.2 – Forecasting techniques

Prelude

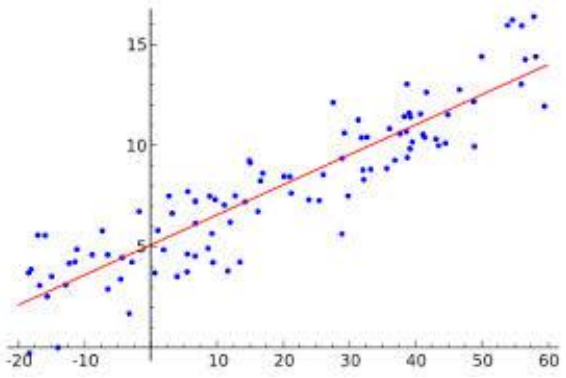
Planning involves forecasting key revenue and cost drivers. Forecasting can be done using qualitative (market research, anecdotes, opinions from experts, historical analogy) and quantitative methods. Here we are focusing on following quantitative methods of forecasting:

1) **Linear Regression Analysis**
   a) Single variable and
   b) Multi variable
2) **Learning Curve Analysis**
   a) Cumulative Average Time Learning Model and
   b) Incremental Unit Time Learning Model
3) **Moving Averages**
   a) Weighted Average and
   b) Exponential Smoothing
4) **Time Series Analysis**
5) **Expected Value Techniques and Sensitivity Analysis**

A. Understanding of a simple regression equation and the measures associated with it

What is a Linear Regression Analysis?

1. Statistical tool / model to establish relationship between one variable (called dependent variable, Y) with another (or a group of another) variable (called independent variable(s), Xi)
2. The relationship is then translated into a linear regression equation and used to predict / forecast the value of dependent variable (Y) given the values of independent variable(s)



1) **Types of Regression Analysis:**
   a) A simple regression analysis uses / involves only one independent variable
   b) A multiple regression analysis uses / involves multiple independent variables
2) **A simple linear regression line has an equation of the form Y = a + bX, where**
   a) X is the explanatory variable
   b) Y is the dependent variable
   c) b is the slope of the line and measures change in y w.r.t unit change in x

d) as is the intercept (the value of y when x = 0)

3) A linear regression is fitting a straight line to data and explaining the change in one variable through changes in other variables. This is based on following assumptions:

a) **Linearity:** Linear relationship between X and Y (Y varies directly with first power of X)

b) **Constant Process:** Process relating the variables is constant or stationary

c) **No auto correlation:** Dependent variable is not auto-correlated – this implies the errors measured by Y(actual) – Y(predicted) are normally distributed with zero mean and a constant standard deviation

d) **No multi-co linearity:** The independent variables are independent of each other. They are not correlated with each other

4) In real life, we hardly come across a situation where these assumptions are met. We nevertheless perform regression analysis. So, it's likely that analysis doesn't yield efficient results. **We therefore have various measures to test the efficiency of a regression analysis or model:**

a) **R Squared:** Also known as coefficient of determination; takes a value between 0 and 1; explains the extent to which changes in dependent variable can be explained by changes in dependent variables; a statistical measure of how close the data are to the fitted regression line;

i) 0 indicates that the model explains none of the variability of the response data around its mean.

ii) 1 indicates that the model explains all the variability of the response data around its mean.

b) **T value:** Measure of strength of relationship between the independent and dependent variable:

i) A value of 0 means no significant relationship between the two and hence the independent variable should be removed from the regression analysis

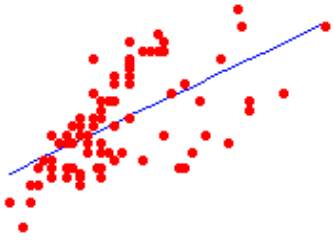
ii) should be more than 2 to indicate a strong relationship between the dependent and independent variables

c) **Standard Error (SE):** A measure of the accuracy of predictions; a measure of dispersion around the regression line

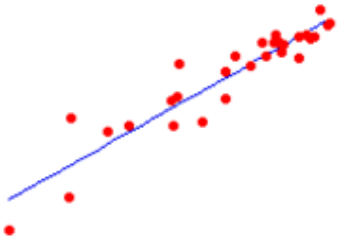
i) ~68% of observations should fall within ± 1 x SE

ii) ~95% of observations should be within ± 2 x SE

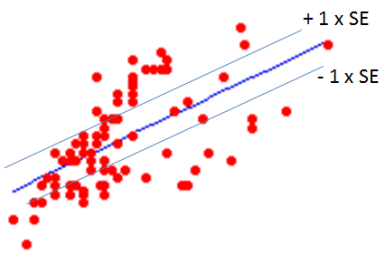
Regression Fit with Low R squared value



Regression Fit with High R squared Value



Graphical Representation of Standard Error



B. A multiple regression equation and when it is an appropriate tool to use for forecasting

Multiple Linear Regression Equation: $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

Interpretation of the variables and measure of efficiency of the regression model remain as per single linear regression equation (intercept, slope, R squared, T value, standard error)

When Multiple Regression is an appropriate tool to use for forecasting:

1. When a single independent variable is not able to explain the dependent variable in entirety. This implies the entire change in dependent variable is not explained by a single independent variable (R squared less than 1)
2. When the dependent variable is not expected to be impacted by seasonality, cyclicality or unexpected trend
3. When we know that an outcome is actually impacted by more than one input / variable

C. Calculate the result of a simple regression equation

Better explained by an example.

A marketing research firm has undertaken regression analysis on the historical sales data of Alpha Beta Corporation and has come up with following linkage between the monthly quantity sold (Q) and the sales price (P): Q = 25000 – 10 x P. Sonali Sundaram is preparing the sales budget for the company. What will be the budgeted annual revenue for the company if the sales price is budget at Rs. 50 / unit?

Solution:

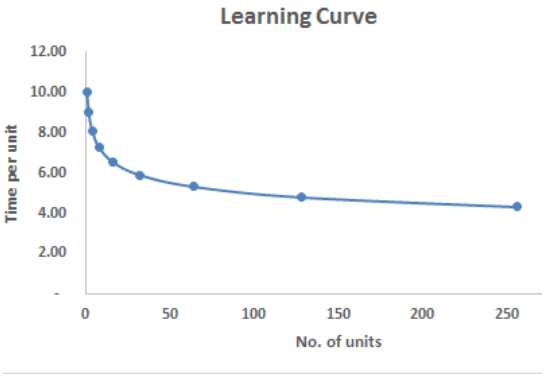
Monthly quantity as predicted by regression equation @ price of Rs. 50 / unit = 25000 – 10 x 50 = 24,500. Hence, Annual Quantity = 12 x 24500 = 2,94,000.

Budgeted Revenue = Sales Price x Annual Quantity = Rs. 1,47,00,000 = Rs. 1.47 Cr

D. An understanding of learning curve analysis

1. As an individual, group or organization completes a job, task or an activity; he /she / it tends to gain experience and learning. As a result when the same task is performed for the second time, the time and the cost involved should be lower than what they were previously.
2. This is the effect of learning and hence a plot of time/cost required per unit of production against cumulative units of production is called learning curve.
3. Learning curve analysis is thus a tool that helps us estimate the time / cost of production at different production levels.

- Learning curve analysis is thus a tool that helps us estimate the time / cost of production at different production levels.
- The graph on the right hand side depicts the time taken to complete the nth unit. For example, if time required to make 1st unit was 10 hours then time required to complete the 4th unit should be 8.10 hours as per 90% learning curve.
- The curve levels off after a certain point.



E. Calculate the results under a cumulative average-time learning model and under an incremental unit-time learning model

Two models used to capture different forms of learning are:

1. **Incremental unit-time learning model:** The incremental time needed to produce the last unit declines by a constant percentage each time the cumulative quantity of units produced doubles. Example: Calculate the time required to produce the nth unit at a cumulative production level of "n" assuming the learning rate of 90% and 10 hours requirement to make the first unit.

Cumulative Production Level (n)	Time required to produce nth unit (hours)	Formula
1	10.00	
2	9.00	= 10.00 x 90%
4	8.10	= 9.00 x 90%
8	7.29	= 8.10 x 90%
16	6.56	= 7.29 x 90%

2. **Cumulative average-time learning model:** The cumulative average time per unit declines by a constant percentage each time the cumulative quantity of units produced doubles. Example: Calculate the total time required to produced "n" units, assuming the learning rate of 90% and 10 hours requirement to make the first unit.

Cumulative Production Level (n)	Average Time required to produce "n" unit (hours)	Total time required to produce "n" units (hours)
1	10.00	10.00
2	9.00	18.00
4	8.10	32.40
8	7.29	58.32
16	6.56	104.98

Inferences from learning Curve Models:

1. Resource (Time or cost) required to make the next unit is lower than that required to make the previous unit

2. Rate of reduction is faster to begin with and slows down subsequently
3. The curve levels off ultimately implying learning and experience don't lower the quantum of resource required significantly after a point.

F. Understanding of Moving Averages, Weighted Moving Averages & Exponential Smoothing and Calculate Forecasts

- A time series data is a compilation of behaviour of a variable at a particular time interval (say daily, monthly, quarterly, hourly etc.) over a period of time
- An analysis of time series tells us many underlying behaviour of the variable but also exposes many noises (cyclicality, seasonality, irregularity etc. – explained in the subsequent LOS)
- Smoothing techniques are used to reduce irregularities (random fluctuations) in time series data. They provide a clearer view of the true underlying behaviour of the series. Smoothing is usually done to help us better see patterns, trends for example, in time series. One should generally smooth out the irregular roughness to see a clearer signal.

Techniques of Smoothing:

1. **Moving Averages:** Average of recent most set of data for the given fixed time frame. With every time period, last data point moves out and the recent most data moves in. Example: Three months moving average of Sales
2. **Weighted Moving Averages:** Similar to moving average method, but higher weights are assigned to recent data than the older data; assumes that recent observation will be a better predictor.
3. **Exponential Smoothing:** Forecast value is a weighted combination of the observed value at time t and the forecasted value at time t as shown by equation below:
   $F_{t+1} = \alpha \times D_t + (1 - \alpha) \times F_t$ where α is the smoothing constant
   Although the method is called a smoothing method, it's principally used for short run forecasting.

Sl. No.	Parameter	Moving Averages	Exponential Smoothing
1.	Effective	Effective in smoothing out sudden fluctuations in demand pattern in order to provide stable estimates Useful if forecast is assumed to stay fairly steady over time	Work well when the time series is stable without any significant trends, cyclicality or seasonality. Good for forecasting large number of dependent variables,
2.	Data Requirement	Requires maintaining extensive records of past data; requires historical data	Requires little record keeping of past data; minimal data requirement
3.	Weights	Equal or more weightage to recent data	Smoothing constant ranges from 0 to 1; subjectively chosen; Smaller constant gives more smoothing, larger constant gives less smoothing

Actual Sales of Alpha Beta Corporation in first 3 months of the years were as tabulated alongside. Sales Forecasts for March was Rs. 108 Cr. Calculate the sales forecasts for April based on moving averages, weighted moving averages (use weights of 20%, 30% and 50% for Jan, Feb and March data respectively) and Exponential Smoothing with constant of 0.92.

Sl. No.	Month	Sales (Rs. Cr)
1.	Jan	100
2.	Feb	106
3.	Mar	98

Solution:

Moving average forecast = (100 + 106 + 98) / 3 = 101.33

Weighted Moving average forecast = (20% x 100 + 30% x 106 + 50% x 98) / 3 = 100.80

Exponential Smoothing Forecast = 0.92 x 98 + (1 – 0.92) x 108 = 98.80

G. Understanding of Time Series Analysis including objectives and patterns

1. A time series data is a compilation of behaviour of a variable at a particular time interval (say daily, monthly, quarterly, hourly etc.) over a period of time.
2. An analysis of time series tells us many underlying behaviour of the variable but also exposes many noises / patterns:

Sl. No.	Pattern	Remarks
1.	Trends	1. Depicts a gradual shift to a higher (upward sloping) or lower (downward sloping) values 2. Due to change in population, technology, customer preference, pricing, competition etc. 3. Forecasts for dependent variable can be made based on trend line using regression analysis (time as independent variable) 4. Informed decision-making on expansion related decisions
2.	Cyclical	1. Repetitive pattern of data points lying above or below the trend line 2. Can last for more than a year 3. Result of macro-economic conditions 4. Improved decision-making in the light of macro-economic conditions
3.	Seasonal	1. Peak and Trough pattern during a year 2. Can occur due to seasonal change in preference and activities 3. Assists in inventory management
4.	Irregular Variations	1. Any outlier data points which don't form a trend, seasonality or cyclical pattern 2. Can occur due to non-recurring, extraordinary event 3. Such factors cannot be accounted for in the forecasts

H. List the benefits and shortcomings of regression analysis, learning curve analysis and time series analysis

Regression Analysis	Learning Curve Analysis	Time Series Analysis
Benefits		
1. Simple and most frequently used tool for forecasting 2. Can be used to trace the relationship and the strength of relationship with multiple variables at a time 3. Can handle large set of data as well to derive conclusions 4. Forces user to look at the relationship between the variable graphically and predict the trend intuitively	1. Important technique for predicting how long it will take to undertake future tasks. 2. Helps you take into account the impact of learning for resource planning, staffing, control and decision-making. 3. Emphasises the fact that initial resource requirements usually do not accurately represent future requirements.	1. A time series analysis clearly depicts all kinds of trends and noises and help managers take informed decision
Shortcomings		
1. Outliers can distort the results of regression analysis. Should be checked to ensure that any data recording error or an extraordinary event is not impacting the analysis 2. May not be a suitable tool if assumptions (linear relationship, stable process, no multi collinearity, no autocorrelation) don't hold true 3. Regression analysis should not be used for forecasting if independent variable lies outside the historical data set	1. Applies only for labor-intensive operations 2. Assumption of constant learning rate may not be valid and hence induce errors in forecasting 3. Productivity improvement can be due to factors other than learning so unreliable conclusions	1. Uses historical data for forecasting. Past patterns are expected to occur in future 2. Extraordinary or non-recurring events cannot be factored for forecasting

I. Expected Value of Random Variables

1. For sake of simplicity, a random variable is a variable that can take any possible value. There always exists a finite probability for every possible outcome it can show.
2. For such a probabilistic situation, a random variable is quantified by its expected value given by sum total of its probability weighted values:
3. $X = p_1x_1 + p_2x_2 + p_3x_3 + \ldots\ldots p_nx_n$
4. Where X is the expected value of the variable x; x_1, x_2,…x_n are "n" different outcomes / value it can take and p_1, p_2,…p_n are the respective probabilities of these outcomes / values

Example:

Sales of Alpha Beta Corporation can take any of the three values depending upon the state of economy. Calculate the expected value of sales forecasts?

Economic Condition	Sales Estimate (Rs. Cr)	Probability
Good	5,000	0.15
Bad	2,000	0.15
Average	4,000	0.70

Expected value of sales forecasts = 5,000 x 0.15 + 2,000 x 0.15 + 4,000 x 0.70 = Rs. 3,850

J. Benefits and Shortcomings of Expected Value Techniques

Benefits

1. A method to forecast a variable even if it has too much of uncertainties associated with it
2. Aids in decision making – to do or not do
3. Considers all the possible states / outcomes / before decision making

Key Issues:

1. Estimation and probabilities assigned under different conditions can be subjective
2. Decision cannot be made in case of unreliable estimates
3. Expected value method assumes the decision maker is risk neutral. Will not be suitable for risk-taker or risk-averse decision makers

K. Use probability values to estimate future cash flows

Alpha Beta Corporation is planning to introduce a new product in its portfolio. The cost of introducing the product is Rs. 50 Cr. The product once introduced can fetch different level of revenues depending upon the state of economy next year. The gross margin is roughly 50% and fixed operating costs will be Rs. 10 Cr per annum. What is the expected break even period for this product?

Sl. No.	State	Probability	Revenue (Rs. Cr)
1.	Highly buoyant	10%	50
2.	Buoyant	20%	40
3.	Most likely	30%	30
4.	Pessimistic	40%	20

Solution: All financials in Rs. Cr

Sl. No.	State	Probability	Revenue	Variable Cost	Fixed Costs	Total expenses	Cash Flow
1.	Highly buoyant	10%	50	25	10	35	15
2.	Buoyant	20%	40	20	10	30	10
3.	Most likely	30%	30	15	10	25	5
4.	Pessimistic	40%	20	10	10	20	0

Expected cash flow per annum (Rs. Cr) = 15 x 10% + 10 x 20% + 5 x 30% + 0 x 40% = 5

Expected payback period = 50 / 5 = 10 years

L. Uses of sensitivity analysis

An output or a dependent variable is typically dependent upon several independent / input variables. Sensitivity Analysis tells us:

1. How sensitive the outcome / output is to a particular input?
2. The outcomes corresponding to different values of inputs?
3. Identify the variable(s) which significantly or weakly impact(s) the output / outcome

M. Perform a sensitivity analysis with different values for the probabilities of the states of nature and / or the payoffs

Example:

Alpha Beta Corporation has assigned different probabilities to the sales estimates under three macroeconomic scenarios. However two other estimates of probabilities are also available as depicted below. Find out the sensitivity of sales w.r.t economic conditions.

Economic Condition	Sales Estimate (Rs. Cr)	Expected Probabilities	Alternate Probabilities 1	Alternate Probabilities 2
Good	5,000	0.15	0.65	0.10
Bad	2,000	0.15	0.10	0.65
Average	4,000	0.70	0.25	0.25

- Expected sales under expected probabilities will be (Rs. Cr) = 5,000 x 0.15 + 2,000 x 0.15 + 4,000 x 0.70 = Rs. 3,850
- Expected sales under alternate probabilities 1 will be (Rs. Cr) = 5,000 x 0.65 + 2,000 x 0.10 + 4,000 x 0.25 = Rs. 4,450
- Expected sales under alternate probabilities 2 will be (Rs. Cr) = 5,000 x 0.10 + 2,000 x 0.65 + 4,000 x 0.25 = Rs. 2,800
- The significant variations in the sales forecast under three scenarios shows that it is highly sensitive to the economic conditions of the country.

N. Benefits & Shortcomings of Sensitivity Analysis

Benefits:

1. Allows a manager to do a what if / scenario analysis
2. Helps managers identify the most crucial / critical / sensitive variables in decision making and thus guides them to be very realistic and cautious in forecasting / projecting them

Shortcomings

1. Independent variables may be correlated and may not impact dependent variable individually but mutually result in substantial different outcome.
2. Independent variables may be correlated, hence all of them need to be changed if forecasts are to be accurate. A manager must understand this otherwise forecasts will be grossly inaccurate.