International Journal of Computer Architecture and Mobility (ISSN 2319-9229)
Volume 7 -Issue 8, August 2019

# Newsvendor Problem and Forecasting based on Past Data

Vrashi Shrivastava and Manali Harinkhere

*Department of Computer Science Engineering*
*Shri Vaishnav Vidyapeeth Vishwavidyalaya*
*Gram Baroli,Sanwer Road,Indore 452002, India*

***Abstract – In operational analytics, Newsvendor Problem is an essential model used to define the problem of finding optimal solution for many industries. Forecasting methods are used to find this solution by using past demand data. The prediction made for inventory needs, demands for product, demands for services depends on various factors. The biggest factor is the type of data. For different types of data, the forecasting technique changes accordingly. Some techniques are designed to work with a specific kind of data, so if applied correctly, they give the best results possible but at the same time if applied on a wrong set of data, if fails to give usable results. In this paper, along with the Newsvendor problem, the different forecasting techniques depending on the different characteristics of data are discussed in detail.***

***Index Terms - Forecasting, Newsvendor problem, Trends and seasonality.***

## I. INTRODUCTION

In a world full of uncertainty, one of the main features is the uncertainty of future events. It plays a big role in many business decisions. A customer buys a mobile data plan before knowing it's actual data usage; government orders flu vaccines before the flu season starts for its' residents.These are some examples of Newsvendor problem in which lies the uncertainty. This uncertainty can be removed by predict the future demands which is called Forecasting. There are different types of forecasting methods to be used in different scenarios which are explained in later sections.

## III. NEWSVENDOR PROBLEM

The Newsvendor Problem is enormous and faced by almost every industry. It is a mathematical model used in Operational analytics to determine the optimal inventory solution. The basis of the problem lies in uncertainty about the future events. A retailer buying seasonal goods has to order the product from the supplier but needs to define the quantity of order before the actual demand is known, which can turn out to be profitable or loss depending on future events. The problem is named Newsvendor because the same situation is faced by a vendor selling magazines or newspaper as he has to buy the newspaper before he knows how many he will be able to sell and if some are left that will add to his loss. This problem arises in almost every industry and thus finding an optimal solution is essential. The solution is to predict the nearest possible value based on the past demand data of the product. This introduces us to the term, Forecasting, which means predicting the future. The characteristics of Newsvendor problem are:

- To make one decision (e.g. how much to buy?).
- There is an objective like maximize profits, minimize costs, improve market share.
- On occurrence of demand, cost and profit are realized.

  OBJECTIVE

## III. PROBABILITY DISTRIBUTION

While in the market, we cannot control the purchasing behaviour of customers. Hence, we cannot predict the exact demand in the future either. However, with the help of past data, we can make a very accurate prediction. The past data is generally modelled as the probability distribution to fulfil the requirement of the prediction method. To decide what future demand scenarios are possible, we need to estimate the likelihood of a scenario's realization.

**Discrete Probability Distributions** are described by distinct scenarios attached with probabilities of occurring. Two major factors that describe a probability distribution are - **Mean** (expected value) and **Standard Deviation** (variance from the mean). Though the mathematics of probability distribution is very vast, it's out of the scope of this paper since we only need the knowledge of a few aspects of it for forecasting methods described in this paper. The following statistics is also called the **Descriptive Statistics.**

**Mean for Discrete Probability Distribution** - the sum of products of scenario values (demand value) and their probabilities.

D= $p_1D_1 + p_2D_2 + p_3D_3... + p_nD_n$

D = i=1n$p_iD_i$

Mean reflects the demand value that we will get, on average, in the selling season.

**Standard Deviation** - it is the square root of the sum of the products of the probability and square of the difference between the mean and the scenario value(demand).

SD = p1 (D1-D)2 + p2 (D2-D)2 + p3 (D3-D)2

SD = i = 1npi (Di-D)2

The variance of actual random variables from the mean is called the standard deviation. In other words, how far away the actual values will be from the mean.

**Continuous probability distribution** is the most common type of probability distribution. It is used when

1 A random variable model has a large number of scenarios on a small interval out of all the possible intervals of demand value.

2 The probability of any scenario being realized is very low.

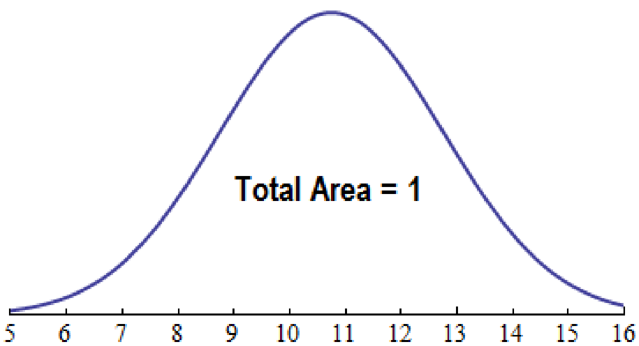


Fig 1.1 continuous probability distribution

Given the cardinal rule of probability, the area under the continuous probability distribution curve is 1.

The shaded area under the curve represents the probability of a random variable X's value is greater than $X_1$ and smaller than $X_2$.

The most popular example of continuous is Normal Distribution.

**Normal distribution** allows the underlying variable to take any value from negative infinity (−∞) to (+∞). One feature that makes is the most encountered distribution is that it is completely characterized by two factors −

1. Mean **μ**
2. Standard deviation ***σ***

For a normal distribution curve, the calculation of these two statistical values is fairly easy. That's how it can describe the distribution of relative future changes in the Stock Market values, FX rates, and analysing the Income Distribution In Economy.

-insert original normal distribution image-

Apart from the normal distribution, there are other popular continuous distribution curves that are widely used. For example - Exponential Probability Distribution and Beta Probability Distribution. But those are out of the scope of this paper since, for the methods discussed in this paper, only the knowledge of normal and continuous probability distribution is necessary.

**Predictive Statistics** - descriptive statistics however, can not be used for forecasting the future demand in some cases for some particular methods. But it can easily be done using the predictive statistics. One can calculate the predictive statistics using previously calculated descriptive statistics.

Mean for Prediction = Descriptive Sample Mean

The reason that one can use the Descriptive Mean for prediction as well is because it is an unbiased estimate of true demand and distribution. However, the standard deviation for the prediction has to be changed. Simply because it is insufficient data.

The predictive standard deviation for normally distributed data is calculated by -

Standard Deviation for prediction $(\sigma) = s + s / \sqrt{n}$

Where s denotes the descriptive standard deviation and n is the total number of data points.

## IV. Forecasting

Point forecasts are very rarely accurate because the demand could be a random variable. So, for a forecast to be good, it should not be a single number. It should be a range or distribution information like Mean and standard deviation. Aggregated forecasts are usually more accurate. However, the accuracy of forecasts decreases as we try to predict further in the future. One thing to pay attention to while making a prediction is that no known information should be excluded unless there is a very strong reason for it. There are more than one ways to predict the demand depending on the circumstances.

### A. *Subjective Forecasting*

It is the preliminary forecasting technique. It contributes a little more than just a wild guess. However, it is not based on historic data analytics. Few common examples of subjective forecasting methods are - Composite forecasting, customer survey, jury of executive opinion and delphi method are some of the methods used that can be called Subjective.

### B. *Forecasting with Past Data Objectively*

The past data can be used in liaison with some forecasting techniques like Causal models and Time Series Methods.

**Causal Models** can be explained through causal analysis(analysis of the cause).

For example - if there is a demand 'A' to be predicted. Assume that there are 'n' variables (causes) that are influencing the demand A. in a causal model for the same scenario, the demand A will be formulated as a theoretical function including all the causes 'n'.

Causal models however, are not easy to use, create or optimize as per use. They are generally very complex and need advance and intricate tools along with high expertise in the domain.

**Time Series Model** can also be called 'naive' method is nothing but the collection of data, values and and variable to be predicted. The purpose of this model is to identify and isolate the patterns associated with the past data and using those patterns, one can come up with descriptive statistics which can later help with the grid predictions about the future.

Here, the past data might have come caricature like trend, cycles/seasonality, randomness. The operations and the handling of data varies for each of the characteristics.

### C. *Forecasting for Stationary Series*

Stationary data is the data that shows no trend. So, it can be said that the future resembles the past roughly with the stationary data. So, past data can be efficiently used for forecasting. A stationary time series is of the form -

$D_t = \boldsymbol{\mu} + \varepsilon_t$

where $\boldsymbol{\mu}$ is a constant (sample average), $\varepsilon_t$ is random variable whose mean is zero and there is some standard deviation. The mean and standard deviation are not dependent on time.

There are two very common methods for forecasting with stationary series. Namely -

1. Moving Averages
2. Exponential Smoothing

Though the Exponential Smoothing method is out of the scope of this paper, the moving averages method is discussed in the next section

### D. *Moving Averages Method*

In simple words, it can be defined as the arithmetic average of 'n' most recent data points(observations). It is generally denoted as MA(n) where 'n' denotes the number of observations.

Example - if one wants to make a forecast for time period 't' in future using moving averages then -

$F_t = (D_{t-1} + D_{t-2} + D_{t-3} + \ldots + D_{t-n}) / n$

This is one step forecast. Meaning that this will find the next value/ data point after the existing data. In Moving Averages method, the one step forecast is the same as multi step forecast.

Example - if one wants to make a forecast for time period 't + 1' in future using moving averages, then -

$F_{t+1} = (D_t + D_{t-1} + D_{t-2} + \ldots + D_{t-(n-1)}) / n$

The method is called 'Moving Averages' because the chosen data points keep moving ahead as per the forecasting requirements and are always the most recent ones. As shown in the example, the new data point $D_t$ is added and $D_{t-n}$ is discarded.

<u>**What Kind of Data to be used with Moving Averages?**</u>

While choosing this method, the following things about the characteristics of data usage of the method should be kept in mind and should be corresponded with the requirement of forecasting- For last n data points,

1. All the data older than n data points is ignored completely and not used at all. Sometimes It could prove to be a waste of resources to have large quantities of data and not utilizing it to make better predictions.
2. All recent data points are weighted the same. Observation from two days ago holds the same value and is treated the same as last week's observation.
3. If one wants to give more weight to the more recent data and lesser as the data gets older, this method may not be the first choice for it. This however, can be achieved using the **Exponential Smoothing Technique** since it is based on this exact idea.

<u>**Advantages of Moving Averages method**</u> - there are the following advantages of this method.

1. It is easy to understand
2. It is easy to compute
3. It provides stable forecast

<u>**Disadvantages of Moving Averages method**</u> - along with all the advantages,there are some disadvantages of this method as well. They are as follows.

1. It lags behind trend if the data shows that behaviour
2. Choosing the number of data points to consider is a big question that affects the forecast immensely.
3. It is by no means a causal model. So it will not explain why the realization in the future behaviour comes out to be in a certain way, even though it gives a fairly good forecast.

## V. Forecasting with Trends and Seasonality

Sometimes it is clearly observable that the data has some characteristics. For example it may be continuously increasing or decreasing. This is called <u>**Trend**</u> in data. Generally there is a trend in data. If not a continuous trend, the data may also depict that the increase in value of a particular period of time and decrease in value for some other duration in time. This property of data is called <u>**Seasonality**</u>.

### A. *Forecasting When There is Trend in Data*

When the data shows a trend, there are two ways for forecasting as given following -

1. Using Moving Averages
2. Linear Regression

<u>**Moving Averages Method**</u> - if there is increasing or decreasing trend in data, forecasting using Moving Averages method actually lags behind the trend.

- When there is an increasing trend, the Moving Average prediction is usually below the demand.
- When there is decreasing trend shown by the data, the Moving Averages forecast is usually above the demand.

Note - when there is no trend in data, the moving averages method works better if as many data points are used for forecasting. But when there is a trend in data, the more data points one uses to forecast demand, the farther it is from the actual demand. Using more data points actually makes the prediction worse and more errorsome in case of trend.

<u>**Using Regression for Time Series Forecasting**</u> - the main idea in this method is to fit a line that has a slope to capture the trend in data. As discussed above, **Linear Regression** can be used in case of a trend is there in data. Using linear regression, forecast for a period t is calculated by -

Model: $D_t = a + bt$

Where t is the time period, b is the slope of the trend line, a is the intercept.

**'Best - Fit' Trend Line - Ordinary least square (OLS)**

It is a method that fits a trend line through data while minimizing the squared errors. How it is done is a straight line -

**$D_t$ = a + bt** is a fit through t = 1, 2, 3…., n data points and parameters 'a' and 'b' are chosen such that it ensures the minimum average squared distance of data from trend line. The new term introduced here '**Average Squared Distance**' is nothing but the Mean Squared Error.

## VI. Error in Forecasting

For any forecast, it can not be said that it is error free. There is always some noise in every forecast. The error in a forecast in period t is denoted as $e_t$. it is basically the difference between the forecast and the actual demand in period t.

$e_t = F_t - D_t$

The ways to measure errors are discussed following -

1. Mean absolute deviation (MAD) = $1 / n \sum |e_t|$ Through MAD, we find how much the forecast deviates from the actual demand, and its average is taken.
2. Mean Squared Error (MSE) = $1 / n \sum (e_t)2$ In MSE the error in the forecast is calculated, squared and its average is taken. It is generally used for fitting trend dataset.

Mean Absolute percentage error (MAPE) = $1 / n \sum |e_t / D_t| * 100$

Here the calculated error is divided by the actual demand and the percentage is taken. Then that average is calculated.

## VII. conclusion

There can be a very vast variety of data depending on different collection techniques. Utilizing that data to make informed and optimal decisions about the demand and supply is very important as it was discussed in the II section of this paper. To decide on a demand value, various forecasting techniques are discussed in this paper depending on the type of data. The descriptive and predictive statistics and the difference between them is stated in the III section. For data with no trend or seasonality, the Moving Averages method is gives very optimum results, at the same time, if the data has trend or seasonality, the method fails. the reason and best practices to forecast with trend or seasonality driven data is discussed in V section of this paper. For different data, the technique to draw relevant information from it also changes accordingly.

## References

[1] Zhenghai Zhang, Zhifang Jiang, Xiangxu Meng, Shenghui Cheng, Wei Sun, "Research on prediction method of api based on the enhanced moving average method," *DOI: 10.1109/ICSAI.2012.6223534*, 25 June 2012

[2] Zhongjun Tang," Mechanism for a New Demand Forecasting Paradigm 'Individual Demand Forecasting",Published in 2010 Third International Conference on Business Intelligence and Financial Engineering.

[3] Evan L. Porteus, The Newsvendor Problem," Part of the International Series in Operations Research & Management Science book series (ISOR, volume 115)

[4] Arthur V. Hill," THE NEWSVENDOR PROBLEM," unpublished.

[5] Dan Adelman, Dawn Barnes-Schuster, Don Eisenstein, "The Newsvendor Model" *Excerpted from "The Operations Quadrangle: Business Process Fundamentals"*

[6] Benzion, Uri & Cohen, Yuval & Peled, Ruth & Shavit, Tal. (2007). Decision-Making and the Newsvendor Problem – An Experimental Study. Ben-Gurion University of the Negev, Department of Economics, Working Papers. 59. 10.1057/palgrave.jors.2602470.

[7] Senthil Veeraraghavan, Demand Uncertainty, Forecasting, and the Newsvendor Problem, Operations Analytics, *University of Pennsylvania.*