



Case study: Am i forecasting correctly

Nilava Nandi

BA, (Economics with Econometrics), ICFAI University, Tripura, India

Abstract

The whole idea to present this case study is to simplify the concept of business decision making faced by small and medium enterprises. Often they rely on whims and fancies and comes out with wrong business decisions. To prevent such irrational approach, I presented a live case study. It's an attempt to serve as a guide to many beginners in the field of research and decision making by breaking bombastic statistical words into layman terms so that any professional can understand. Research Methodology is key for any business survival, hence a holistic study is given via case study in a simple and lucid language.

Keywords: forecasting, decision making, hypothesis, significance level

Introduction

Business enterprises especially the smaller one, let's say the young entrepreneurs are often busy with trends and forecasting. Forecasting is a very important decision making tools for "management professionals". A management professional engaged in decision making without taking into account current trends will risk his career as a management professional. The adverse effects of poor forecasting skills or amateurish modelling can have disastrous consequences in his revenue. Please note for a manager, revenue is more important than "net profit" concepts of accounting professionals. A manager forecasts with three objectives and all are inter-related i.e., revenue maximisation, sales maximisation and increasing customer satisfaction. His primary duty is to prevent over exploitation of the budget assigned to him and ensuring optimum resource utilisation. Hence, correct forecasting is essential. For all the new entrepreneurs, I am presenting in a layman language of correct manner of hypothesis.

Situation I: Ashish and his problem

So, Ashish is the manager of my Company handling sales division of our products. He did not want to stock up certain products which have lesser demand; hence he wanted to send the list of items in which the Company should only focus upon looking at last year sales trend. According to Ashish, the predicted frequencies of certain items should match with the last year analysis on the basis of colours. He felt 20% of the customers would choose colour A, 30% would choose colour B, 10% would choose C, 10% would choose colour D and 30% would choose colour E on the basis of last year sales analysis. Now the Company has 300 customer bases and it's difficult to interview each and every customer. So he randomly chose 150 customers and interviewed them. However, he found out following results.

Table 1: A

Products	Actual Observation	Predicted Observation
A	35	30
B	50	45
C	30	15
D	10	15
E	25	45

Source: Company Data

Looking at the above observation, Ashish is actually shocked that none of his predictions are matching with the actual observation as the study revealed that 150 people gave out different verdict completely. We all come across this dilemma. He had two options now, *firstly* to completely reject the study and begin the exhaustive study of 300 people which is a tedious task and expensive as well or he can follow the study conducted by him and accept the risk of product shortage or excess product which will ultimately result in loss of revenue. So, he was clueless and entered the CFO's room asking him to help him to arrive at correct figures. CFO being an accounting professional gave him the advice to study the whole population to avoid the risk. But the directors told him to present him within 4 days as he had already crossed his deadline. So, he also consulted me with the problem in the meantime. I detailed out the solution to him, we worked together like a team, and it was fun, exciting and was a learning experience for both of us.

Situation II: Solution Found in 4 Simple Steps

- I decided to explore the areas of "null and alternative hypothesis" here. Null hypothesis states that there are no significant differences between predicted and actual observation. However, alternative hypothesis says both the observations are different. The principle to reject null hypothesis is when calculated chi square is greater

than tabled value of chi square. The level of significance at which you can say with 95% confidence that the difference between the two observations are not due to chance alone is set at 0.05 which is the standard basis. Hence “Chi-Square Test” is used on these data.

Calculation of Chi Square

Table 2: B

1	2	3	4	5	6
Products	O	E	(O-E)	(O-E) ²	(O-E) ² ÷ E
A	35	30	5	25	0.83
B	50	45	5	25	0.56
C	30	15	15	225	15
D	10	15	-5	25	1.67
E	25	45	20	400	8.89
Total					26.95

Source: Company Data

- After finding the above data of 26.95, my next curiosity was to find *df* (degrees of freedom).
- $df = N-1 = 5-1 = 4$
- My next aim is to find out tabled value of Chi square. On observation it's found, that table value of degrees of freedom (*df*) which is 4(found above using $df = N-1$) at 0.05 level of significance is 9.49.
- Finally, it's now shown that calculated Chi square value is greater than tabled value of Chi square, hence we can reject null hypothesis. *This means that there exists significant differences between the data sets and that cannot be by chance alone.*

Case Solved

In this situation, rejection of null hypothesis implies that difference between predicted observation on the basis of last year's sales and actual observation from the current customer poll taken is not by chance, i.e., there is no chance variation in the sample which Ashish had taken. There is real difference between them. We just achieved a robust test on the given sets of data as goodness of fit. Hence, it would be advisable for Ashish not to ignore close his current poll while deciding upon the production of Company's items instead of relying upon previous trend.

Utility of Chi Square Test

Logic of Hypothesis Testing was first invented by Karl Pearson. It is also known as test of goodness-of-fit is an important contribution in modern theory of statistics. The main purpose of invention was to help the biologists, economists and psychologists. It is a non-parametric test for two main purposes:

- To test the hypothesis of no association between two or more groups or population or criteria (to check independence between two variables).
- To test how the observed distribution of data fits with the expected distribution of data, i.e., to test the goodness-of-fit.

Assumptions

- The data must be randomly drawn from a population. So, Ashish here randomly selected 150 sample out of 300 customer base.
- Sample size should be sufficiently large. Smaller sample can lead to Type II error in hypothesis.

- There is no expected cut-off for sample size, however minimum sample size should be within 20-50, in our case we had taken 150 customers as sample.
- Variables under consideration must be mutually exclusive. Each variable must be counted once in particular category and not be repeated in other category. In our sample, those who chose A will not choose B. Hence, variables are mutually exclusive.

Different Scenario and its Applicability (Hypothetical)

For example, in medical field there are two therapies that are provided to certain group of patients. Few received spinal manipulative therapy (SMT) for 6 weeks and few received intermittent motorized traction (IMT). Suppose 190 patients received IMT and 200 received SMT are randomly selected taken as sample out of its patient list. Out of 190 cases in IMT, 95 said Yes they improved and 95 said No. However, in SMT, 45 said Yes and 155 said No towards recovery survey. The medical team would like to know whether there is an association between improvement and type of treatment received by the patients. This is illustrated in following table.

Table 3: C

Improved			
	Yes	No	Total
IMT	95	95	190
SMT	45	155	200
Total	140	240	390

Source: SGVUDE

In such cases, Expected values are calculated by applying following formula by the researcher:

Expected Values = Row total x Column Total / Grand Total
 = 190 x 140 / 390 = 68

Calculation of Chi Square

Table 4: D

Observation (O)	Expected (E)	(O-E)	(O-E) ²	(O-E) ² ÷ E
95	68	27	729	10.72
95	190-68 = 122	-27	729	5.98
45	72	-27	729	10.13
155	200-72 = 128	27	729	5.70
Total				32.53

Null Hypothesis is assumed that there exists no “significant differences” between treatment and recovery between IMT and SMT. We know that if calculated Chi Square is greater than Tabulated Chi Square, we reject Null Hypothesis. To find tabulated, we need degrees of freedom. In such cases, its calculated in following manner.

Degrees of Freedom or $df = (No. \text{ Of Rows} - 1) \times (No. \text{ Of Columns} - 1) = (2-1) \times (2-1) = 1$

If we take significance level of 0.05, then for $df = 1$, tabled chi square value will be 3.841. Thus, calculated value of 32.53 is greater than tabled value of 3.841, hence we can reject null hypothesis. Thus we can conclude that proportion of individual who received IMT and their recovery is significantly different to that of SMT. Thus, we can see how medical research team can also use hypothesis testing.

Uses of Chi Square in Business Organisations

Managers do need information daily on market and trends to

improve the customer satisfaction rate and at the same time keeping a check at inventory. Here, Ashish attempted the same thing. He basically did a market research on his own but he lacked a basic skill of substantiating his point when the confusion arose. Such market information is highly valuable to any organisation depending upon following factors:

- Ability and willingness to act on the information which in our case Directors was interested.
- Accuracy of the information which we substantiated mathematically and senior management is happy.
- The cost of information in terms of money and time which we reduced significantly by using statistical approach.
- And variability in information. There was a significant difference between observed and expected data which Ashish soon realised. Had Ashish gone along with his own logic, it would have created massive loss to the Company.
- Company's willingness to move by market trend instead of reliance on past trend.

Steps in Research

- The first step is to define the problem in the research conducted. Where is the confusion? Why there is confusion?
- Next step should be to find out whether my data is correct or not. Cross checking with secondary data is helpful here.
- Open attitude is needed while conducting research. This means that you need to be aware that you cannot study entire population, hence how data can be validated.
- Awareness of statistical tools that can substantiate the research is highly essential. When you cannot study entire population, you need to first understand whether it's parametric or non parametric data. While for parametric greater than 30, we would have applied different hypothesis testing technique completely.
- Test of significance is highly essential. While Chi Square can analyse on two means, ANOVA can analyse on multiple means. Awareness of significance testing hence lies of utmost importance.
- Market Research Report (MRR) should exclusively state the methodology applied for such analysis.

References

1. Abraham Sheena, Kulkarni Sachin. *Business Research Methodology*. Pune: Talentedge Education Ventures Pvt. Ltd.
2. Aggrawal B. *Basic Statistics*. New Delhi: New Age International Publisher, 2009.
3. Firth M. *Forecasting Methods in Business and Management*. London: Edward Arnold, 1977.
4. Jarett A. *Forecasting for Business Decisions*. London: Basil Blackwell, 1987.
5. McQuarrie, Edward, F. *The Market Research Toolbox: A Concise Guide for Beginners*.
6. *QuickMBA.com*. (n.d.). Retrieved, 2020. from <http://www.quickmba.com/marketing/research/>
7. Siegel S. *Non Parametric Statistics for Behavioural Sciences*. Tokyo: McGraw Hill Huga Konsa Ltd, 1956.
8. Singhal DR, Rana R. *Journal of the Practice of Cardiovascular Sciences*, 2015. Retrieved June 20, 2020, from <http://www.j-pcs.org/article.asp?issn=2395->

5144;year=2015;volume=1;issue=1;spage=69;epage=71;aulast=Rana