

Analyzing Returns and Pattern of Financial Data Using Log-linear Modeling

Vivek Vijay*, Parmod K. Paul

Department of Mathematics, Indian Institute of Technology Jodhpur, India

Copyright ©2016 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Technical analysis is useful for forecasting the price movement through the analysis of historic data. This sort of movement has Turn of the year effect also and useful for short term prediction.

If the direction of price of two or more assets is same, it becomes necessary to analyze the returns also. We first use optimal band to predict the direction of price and create a contingency table of the data to analyze the pattern (movement) against returns. We use log-linear modeling for the analysis of the contingency table. We next include the volume of transactions as one more variable in the contingency table. The table consisting of three variables, Pattern, Returns and Volume is further analyzed by using log-linear modeling. We test various hypotheses of association for these variables by using Chi-square test for contingency tables.

Keywords Contingency Table, Log-linear Modeling, Optimal Band, Technical Analysis, Trading Band

1 Introduction

Stock market always attracts the investors to invest money according to their choice from which large profits can be earned. Fundamental driver behind maximizing profit is the strategy of buying and selling of the stocks. The buying and selling behaviour of investors is also affected by Turn of the year [Ritter 1988]. It is well documented that turn of the year, the average ratio of buying and selling, is more in first 9 days of January than mid January to mid December and last 9 days of December. Rozeff and Kinney (1976) also gave explanation about the January effect that the average of returns of stocks is higher in January month than in other months. There are number of articles available to discuss the Turn of the year effect. Jay R. Ritter (1988) proposed a theory based on the tax-loss-selling named, "parking-the proceeds" to explain the Turn of the Year effect on the NYSE daily returns from 17 Dec 1970 to 16 Dec 1985 using t-statistic. Barber and Odean (2008) tested the hypothesis based on attention grabbing stocks. These statistical tests confirm that the behaviour of individuals and institutions differ while buying and selling

the stocks.

There are several technical indicators proposed by researchers and financial experts for the prediction of pattern. Some of these indicators are Bollinger Band [Bollinger (2001)], Moving Average, Moving Average Convergence/Divergence, Relative Strength Index, Confidence Index, [Hoque and Gias (2009)] and Optimal Band [Vijay and Paul (2015)], to predict this buying and selling behaviour of stocks.

Most of these indicators are based on the past returns, their moments and/ or volume of transactions. For the short term investors/ traders, this analysis is important to make the decision of their investments. However, if the indicators exhibit similar pattern for two or more stocks, the decision is made on the basis of return and its association with pattern. We, here, classify the historic data as per their pattern by using optimal band [Vijay and Paul (2015)]. For each of the categories of pattern, we further divide the whole data into different categories of returns. If the interest lies in the classification of pattern then historic values of returns are used to predict the same but if one is interested in forecasting the returns then the historic value of pattern becomes more useful [Vijay and Paul (2015)]. Therefore, it becomes important to analyze the strength of dependence between the two variables, returns and pattern.

First, we use the historic data to see the buying and selling pattern by using the optimal band [Vijay and Paul (2015)]. The pattern data is then divided into three categories, namely, Sell(Y^S), Neutral(Y^N) and Buy(Y^B). This is further used to estimate the future category of returns, High, Moderate and Low. The whole data is then presented in the form of a 2-dimensional contingency table by using the variables, returns and pattern. Note that each of these variables has three categories. In technical analysis, one of the fundamental drivers is volume of transactions. We include the volume as third variable with its two categories, namely Up and Down. This division of volume is primarily based on the range of historic returns. This creates a 3-dimensional contingency table. A partial table is the cross-classification of two of these three variables for fixed level of the remaining one [Kateri (2014)]. Thus, there are two possible sets of partial tables corresponding to the variable volume, we test different hypotheses for these tables.

The hypotheses for 3-dimensional contingency table are

based upon:

1. Association between buying and high-return under up-volume and down-volume.
2. Relation between selling and high-return/ low-return under up-volume and down-volume.
3. Relation between neutral and all categories of returns under up and down volume.

A first sensible assumption is that the association between pattern and returns exhibits a linear trend. The linear trend is measured by Pearson's correlation coefficient, defined through their categories [Anderson (1996)]. For the purpose of testing of hypotheses, we use observed frequency and expected frequency to find the test statistic $Z(H)$ under hypothesis (H). It is known that $Z(H)$ follows Chi-square distribution with degree of freedom equal to the number of unconstrained log linear model parameters which are set to zero under H and, at a particular level of significance [Boulesteix (2006)]. All the hypotheses of independence or conditional independence can be equivalently represented in terms of interaction parameters of a log-linear model. Log linear modeling is a widely used method for the analysis of a contingency table. Parameters of log linear model describe the interaction/ association among two or more variables. One of the advantages of using log-linear model is that it goes beyond a single summary statistics and specify how the cell counts depend on the levels of categorical variables. They model the association and interaction pattern among categorical variables. These are appropriate when there is no clear distinction between response and explanatory variables, or there are more than two responses [Vellaisamy and Vijay (2007)]. If any hypothesis of independence is accepted then the interaction parameters can be assumed to be zero. If the hypotheses of independence is rejected, the values of these interaction parameters help in analyzing the influence of different categories of variables. The log-linear modelling, therefore, helps us identifying the level of a variable which has strong influence on another variable. Hence, this approach is not only useful for prediction of pattern but also deals with its association with other variables. We demonstrate the process of classification of the data in the form of a contingency table. Various hypotheses are tested by using χ^2 test of independence/ conditional independence. The association, if exists, is described by the parameters of log-linear model.

The structure of the paper is given below:

Section-2 deals with formation of contingency tables by using trading band approach. Section-3 presents, briefly, the log-linear modeling for 2 and 3-dimensional contingency tables. Various hypotheses of association are also represented in terms of interaction parameters. Analysis of contingency tables is shown in Section-4. Conclusion and future aspects are presented in Section-5.

2 Contingency table for Returns, Pattern and Volume of transactions

Consider the series X_1, X_2, \dots, X_n of returns of a stock. We define the process of construction of a contingency table for pattern and returns of the series. We first use optimal band [Vijay and Paul (2015)] to divide the data into three categories of pattern, namely, Sell, Neutral and Buy.

Once divided, the cardinality of each of these subsets of the time series data will represent the count of each category of pattern. We further divide, for each categories of pattern, these subsets into subsets corresponding to the returns, that is, High, Moderate and Low. We use the following algorithm to construct a 2-dimensional contingency table.

Step-1 Define

$$\begin{aligned}\alpha &= \text{Max}(X_1, X_2, \dots, X_n); \\ \delta &= \text{Min}(X_1, X_2, \dots, X_n); \\ \beta_i &= \text{Max}(X_i, X_{i+1}, \dots, X_{i+4}), 1 \leq i \leq n-4; \\ \gamma_i &= \text{Min}(X_i, X_{i+1}, \dots, X_{i+4}), 1 \leq i \leq n-4.\end{aligned}$$

Step-2 Define the linear function [Vijay and Paul (2015)].

$$f = a * \alpha + b * \bar{\beta} + c * \bar{\gamma} + d * \delta$$

$$\begin{aligned}\text{where } \bar{\beta} &= \text{mean}(\beta_1, \beta_2, \dots, \beta_{n-4}) \\ &= \frac{1}{n-4} \sum_{i=1}^{n-4} \beta_i, \\ \bar{\gamma} &= \text{mean}(\gamma_1, \gamma_2, \dots, \gamma_{n-4}) \\ &= \frac{1}{n-4} \sum_{i=1}^{n-4} \gamma_i\end{aligned}$$

The parameters a, b, c and d are obtained in the following step-3.

Step-3 We obtain the parameters a, b, c and d by solving the optimization problem.

$$\begin{aligned}\underbrace{\text{Max}}_{a, b, c, d} f(\alpha, \bar{\beta}, \bar{\gamma}, \delta) &= a * \alpha + b * \bar{\beta} + c * \bar{\gamma} + d * \delta \\ \text{s.t } f &> 0 \\ f &< (\alpha - \bar{\beta})/2 \\ a, b, c, d &\in R\end{aligned}$$

Step-4 We next define, for $1 \leq i \leq n-4$,

$$\text{Upper Band}[UB_1] = \beta_i + f(\alpha, \bar{\beta}, \bar{\gamma}, \delta);$$

$$\text{Upper Band}[UB_2] = \beta_i - f(\alpha, \bar{\beta}, \bar{\gamma}, \delta);$$

$$\text{Middle Layer}[ML] = f(\alpha, \bar{\beta}, \bar{\gamma}, \delta);$$

$$\text{Lower Band}[LB_1] = \gamma_i + f(\alpha, \bar{\beta}, \bar{\gamma}, \delta);$$

$$\text{Lower Band}[LB_2] = \gamma_i - f(\alpha, \bar{\beta}, \bar{\gamma}, \delta).$$

Let us now denote by Y^S , Y^N and Y^B , the subsets corresponding to the categories Sell, Neutral and Buy of pattern respectively. We have the following rule:

$$\begin{aligned}X_i \in Y^S &\text{ if } UB_1 \leq X_i < UB_2 \\ X_i \in Y^N &\text{ if } UB_2 \leq X_i < LB_1 \\ X_i \in Y^B &\text{ if } LB_1 \leq X_i \leq LB_2 \\ \text{for } 1 \leq i \leq n-4.\end{aligned}$$

The total cell count corresponding to the Sell, Neutral and Buy are given by the cardinality of the sets Y^S , Y^N and Y^B .

Next, we divide each of the subsets Y^S , Y^N and Y^B into High, Moderate and Low returns.

Step-1 Consider the set Y^B , and denote the maximum, minimum and average values of the set Y^B by Y_{max}^B , Y_{min}^B and Y_{Ave}^B respectively.

Define the intervals

$$\begin{aligned} I_H^B &= (Y_{Ave}^B + 0.3 * Y_{max}^B, Y_{max}^B) \\ I_M^B &= (Y_{Ave}^B + 0.5 * Y_{min}^B, Y_{Ave}^B + 0.3 * Y_{max}^B) \\ I_L^B &= (Y_{min}^B, Y_{Ave}^B + 0.5 * Y_{min}^B) \end{aligned}$$

Step-2 The classification is defined by the following rule:
Let $y \in Y^B$, then

$$\begin{aligned} y \in I_H^B &\Rightarrow y \in Y^{BH} \\ y \in I_M^B &\Rightarrow y \in Y^{BM} \\ y \in I_L^B &\Rightarrow y \in Y^{BL} \end{aligned}$$

Here, Y^{BH} , Y^{BM} and Y^{BL} are subsets of Y^B corresponding to the categories High, Moderate and Low of returns.

Similarly, we obtain the subsets $\{Y^{SH}, Y^{SM}, Y^{SL}\}$ corresponding to Y^S and $\{Y^{NH}, Y^{NM}, Y^{NL}\}$ corresponding to Y^N .

The 2-dimensional contingency table for the variables pattern and returns is formed by the counts given by cardinality of these subsets.

The table is represented above

Table 1. Frequency Table

Pattern	Returns		
	High	Moderate	Low
Sell	$ Y^{SH} $	$ Y^{SM} $	$ Y^{SL} $
Neutral	$ Y^{NH} $	$ Y^{NM} $	$ Y^{NL} $
Buy	$ Y^{BH} $	$ Y^{BM} $	$ Y^{BL} $

We next present a concrete example.

Example: We consider the Maruti Sazuki Co. daily returns data from 23 Nov 2007 to 23 Nov 2009. The total data points are $n = 478$.

Step-1 We obtain

$$\alpha = 0.1597, \bar{\beta} = 0.0348,$$

$$\bar{\gamma} = -0.0305, \delta = -0.0987,$$

where, $\bar{\beta} = \frac{1}{n-4} \sum_{i=1}^{n-4} \beta_i$ and $\bar{\gamma} = \frac{1}{n-4} \sum_{i=1}^{n-4} \gamma_i$.

Step-2 We now find a linear function f defined as

$$f = a * \alpha + b * \bar{\beta} + c * \bar{\gamma} + d * \delta.$$

The initial values of the parameters are chosen as

$$a = 0.5, b = 0.7, c = 1, d = 1.$$

The estimated values are [Vijay and Paul (2015)],

$$\hat{a} = 0.1458, \hat{b} = 0.5071,$$

$$\hat{c} = 0.4315, \hat{d} = 0.2084.$$

Note that, different initial values of parameters may give different estimates but the function's value remains unchanged.

Step-3 We get

$$f(\alpha, \bar{\beta}, \bar{\gamma}, \delta) = 0.0072.$$

Step-4 We define the following bands:

$$\text{Upper Band}[UB_1] = \beta_i + 0.0072;$$

$$\text{Upper Band}[UB_2] = \beta_i - 0.0072;$$

$$\text{Middle Layer}[ML] = 0.0072;$$

$$\text{Lower Band}[LB_1] = \gamma_i + 0.0072;$$

$$\text{Lower Band}[LB_2] = \gamma_i - 0.0072.$$

This gives, the total numbers of data points in each of the categories in Y^S , Y^N , and Y^B as

$$|Y^S| = 136;$$

$$|Y^N| = 211;$$

$$|Y^B| = 131.$$

Now, we divide the data for each of the categories of pattern into the categories of returns (High, Moderate and Low). As an example, we use the following criteria for the set Y^B .

Step-1 We have

$$Y_{max}^B = 0.0208$$

$$Y_{min}^B = -0.0987$$

$$Y_{Ave}^B = -0.0265$$

Therefore,

$$I_H^B = (-0.002026, 0.0208)$$

$$I_M^B = (-0.0161, -0.002026)$$

$$I_L^B = (-0.0987, -0.0161)$$

Step-2 The total number of data point in categories Y^{BH} , Y^{BM} and Y^{BL} are given by

$$|Y^{BH}| = 54;$$

$$|Y^{BM}| = 71;$$

$$|Y^{BL}| = 06.$$

In the similar way, we divide the data corresponding to Y^S and Y^N to obtain the following 2-dimensional Table-2.

For short-term investors, another key factor is volume of the transactions. We next include the volume of transaction as third variable to construct a 3-dimensional contingency table. Each subset corresponding to the categories of Pattern and Returns is further divided into Up and Down categories of volume.

We define a constant $q \in (0,1)$ such that the data given in set, for example, Y^{BH} is divided into two categories of volume by the following relation

$$q * \max(vol) < Up Volume \leq \max(vol)$$

$$\min(vol) \leq Down Volume \leq q * \max(vol).$$

Here, q depends on the volume data. The two categories of data denoted by Y_U^{BH} and Y_D^{BH} are corresponding to Up and Down volume of transaction.

For each of the two categories of volume, we use the process given above to obtain the 3-dimensional contingency table. With value of $q = 0.4$, the range of volume of transaction is given by.

$$0.04592 < Up Volume \leq 0.1148$$

$$0 \leq Down Volume \leq 0.04592.$$

Using the above rule, we get the following 3-dimensional table.

Note that if the above 3-dimensional table is marginalized over third variable, that is, volume, we obtain the Table-2. We next give a brief description of log-linear modeling for analysis of contingency table. Also, we present a class of hypotheses for the log-linear models.

3 Log linear modeling

Log-linear modeling for 2-dimensional table: Log-linear model for a 2-dimensional table describes association between two categorical variables. A log-linear model expresses the cell counts depending on levels of the two categorical variables.

We now consider a 2-dimensional table of variables A and B. Let the categories of A and B be respectively $\{1, 2, \dots, I\}$ and $\{1, 2, \dots, J\}$. Assume that x_{ij} represents observed cell count of i th row and j th column of the table. Also, let $\mu_{ij} = E(x_{ij})$, where, $i = 1, \dots, I, j = 1, \dots, J$.

A saturated log-linear model for 2-dimensional contingency table is given by

$$\ln(\mu_{ij}) = \lambda_0 + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}, \quad (1)$$

where $i = 1, \dots, I$ and $j = 1, \dots, J$.

Here, λ_{ij}^{AB} is interaction effect of variables A and B, λ_i^A (λ_j^B) is main effect of A (B) and λ is overall effect. All these parameters satisfy the following constraints (Vellaisamy and Vijay (2007)).

$$\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = 0.$$

Table 2. 2-dim frequency table of 23 Nov 07 to 23 Nov 09

A: Pattern	B: Returns		
	High	Moderate	Low
Sell	09	66	61
Neutral	37	159	15
Buy	54	71	06

Table 3. 3-dim Frequency table of 23 Nov 07 to 23 Nov 09

A: Pattern	B: Returns	C: Volume	
		Up	Down
Sell	High	02	07
	Moderate	42	24
	Low	33	28
Neutral	High	17	20
	Moderate	78	81
	Low	05	10
Buy	High	28	26
	Moderate	44	27
	Low	04	02

These parameters are estimated by using their maximum likelihood estimators given below.

$$\hat{\lambda}_0 = \frac{1}{IJ} \sum_i \sum_j \ln(\mu_{ij});$$

$$\hat{\lambda}_i = \frac{1}{J} \sum_j \ln(\mu_{ij}) - \hat{\lambda}_0;$$

$$\hat{\lambda}_j = \frac{1}{I} \sum_i \ln(\mu_{ij}) - \hat{\lambda}_0;$$

$$\begin{aligned} \hat{\lambda}_{ij} &= \ln(\mu_{ij}) - \frac{1}{I} \sum_i \ln(\mu_{ij}) - \frac{1}{J} \sum_j \ln(\mu_{ij}) \\ &+ \frac{1}{IJ} \sum_i \sum_j \ln(\mu_{ij}), \end{aligned}$$

To know more about log linear modeling and effect of levels of variables, see, Vijay (2011).

Log-linear modeling for 3-dimensional table: Log linear model for 3-dimensional contingency table is straight forward extension of (1), and is given by

$$\ln(\mu_{ijk}) = \lambda_0 + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}. \quad (2)$$

Maximum likelihood estimates of the parameters are also defined similarly. The above model is useful in explaining several interaction effects [Vellaisamy and Vijay (2007)], for example,

$$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = 0 \quad (\forall i, j, k) \Leftrightarrow (A \perp B \mid C),$$

that is, A and B are independent given C. Similarly,

$$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{ik}^{AC} = 0 \quad (\forall i, j, k) \Leftrightarrow (A \perp B, C).$$

All the hypotheses of independence/conditional independence for a 3-dimensional table are presented below-

$$H_1 : \lambda_{ijk}^{ABC} = \lambda_{ik}^{AC} = 0 \quad (C \perp A \mid B);$$

$$H_2 : \lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = 0 \quad (A \perp B \mid C);$$

$$H_3 : \lambda_{ijk}^{ABC} = \lambda_{jk}^{BC} = 0 \quad (B \perp C \mid A);$$

$$H_4 : \lambda_{ijk}^{ABC} = \lambda_{ik}^{AC} = \lambda_{jk}^{BC} = 0 \quad (C \perp A, B);$$

$$H_5 : \lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{jk}^{BC} = 0 \quad (B \perp C, A);$$

$$H_6 : \lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{ik}^{AC} = 0 \quad (A \perp B, C);$$

$$H_7 : \lambda_{ijk}^{ABC} = \lambda_{ik}^{AC} = \lambda_{jk}^{BC} = \lambda_{ij}^{AB} = 0 \quad (A \perp B \perp C).$$

4 Analysis of contingency table

We consider 2-dimensional and 3-dimensional contingency tables for analysis. Chi-square test statistic is used to test whether a set of log-linear parameters is zero or equivalently, to test the hypotheses of independence. If any of the hypothesis of independence or conditional independence is rejected, the parameters will be used to analyze the influence of the categories of variables.

4.1 Analysis of 2-dimensional table

The 2-dimensional table of pattern and returns is created for daily returns of Maruti Suzuki Co, using the process given in Section-3. We use the closing price of stocks of the company from 23 Nov 2007 to 23 Nov 2009 [Table-2]. We test the hypothesis of no association for this table by using

$$Z(H) = 2 * \sum_i \sum_j x_{ij} (\ln(x_{ij}) - \ln(\mu_{ij})).$$

The Z(H) value is 182.634 which is greater than $\chi_{0.95}^2$ with degree of freedom (4).

This leads to rejection of hypothesis of independence. That is, the variables A(Pattern) and B>Returns) are dependent on each other [Anderson (1996) pp 27-28].

We now obtain the maximum likelihood estimates of log linear parameters. The main effect parameters are given in Table-4 and 5. The interaction parameter is presented in Table-6.

Table 4. Value of $\hat{\lambda}_i^A$ for 2-dimensional table

$\hat{\lambda}_i^A$	i=1	i=2	i=3
	-0.1317	0.1367	-0.0051

Table 5. Value of $\hat{\lambda}_j^B$ for 2-dimensional table

$\hat{\lambda}_j^B$	j=1	j=2	j=3
	0.1038	0.7999	-0.9037

Table 6. Value of $\hat{\lambda}_{ij}^{AB}$ for 2-dimensional table

$\hat{\lambda}_{ij}^{AB}$	j=1	j=2	j=3
i=1	-0.4550	0.1134	0.3417
i=2	0.0139	0.1169	0.1308
i=3	0.4411	-0.2303	-0.2109

Note that the parameter $\hat{\lambda}_{1j}^{AB}$ and $\hat{\lambda}_{2j}^{AB}$ attain the maximum value for j=3, that is, corresponding to Low returns. This implies that when the data exhibit selling pattern, there are more chances that the next return will be low in comparison to Medium and High. On the other hand, $\hat{\lambda}_{3j}^{AB}$ is maximum for j= 1 which shows that when the buying pattern is exhibited, there are more chances of the next return to be high. The main effect $\hat{\lambda}_i^A$ shows that the stock remain Neutral most of the time. $\hat{\lambda}_j^B$ shows that the returns of the stock is maintained at Moderate returns. Similarly, one can interpret the other values of interaction parameters.

4.2 Analysis of 3-dimensional table

We consider the returns for two different periods.

4.2.1 The analysis for returns for the period 23 Nov 2007 to 23 Nov 2009

We now divide the data given in above table as per the intensity of volume, that is, Up and Down. This creates a 3-dimensional Table-3 for volume of transactions (Up, Down), pattern and returns. We test the seven hypotheses, given in Section-3, by using the test statistic

$$Z(H) = 2 * \sum_i \sum_j \sum_k x_{ijk} (\ln(x_{ijk}) - \ln(\mu_{ijk})).$$

The following table presents the value of test statistic and standard chi-square value to test it at 95% level of significance.

Table 7. Hypothesis and Z(H) value table of 23 Nov07 to 23 Nov09

Hypo	Z(H) Value	Df	$\chi_{0.95}^2$
H_1	11.312	6	12.592
H_2	132.384	6	12.592
H_3	8.843	8	15.507
H_4	13.546	10	18.307
H_5	134.618	10	18.307
H_6	137.087	8	15.507
H_7	139.321	12	21.026

Clearly, H_1 , H_3 and H_4 cannot be rejected. This shows that the volume is independent of pattern and returns. Under H_1 , H_3 and H_4 the model contains the following non zero parameters: $\hat{\lambda}_{ij}^{AB}$, $\hat{\lambda}_i^A$, $\hat{\lambda}_j^B$ and $\hat{\lambda}_k^C$. The Maximum likelihood estimates for these parameters are shown in following tables.

Table 8. Value of $\hat{\lambda}_i^A$ for 3-dimensional table

$\hat{\lambda}_i^A$	i=1	i=2	i=3
	-0.0356	0.1170	-0.0815

Table 9. Value of $\hat{\lambda}_j^B$ for 3-dimensional table

$\hat{\lambda}_j^B$	j=1	j=2	j=3
	-0.1315	0.4250	-0.2935

Table 10. Value of $\hat{\lambda}_k^C$ for 3-dimensional table

$\hat{\lambda}_k^C$	k=1	k=2
	-0.004	0.004

Table 11. Value of $\hat{\lambda}_{ij}^{AB}$ for 3-dimensional table

$\hat{\lambda}_{ij}^{AB}$	j=1	j=2	j=3
i=1	-0.4813	-0.1092	0.5905
i=2	0.0587	0.1368	-0.1955
i=3	0.4226	-0.0276	-0.3950

Table-11 exhibits a pattern similar to that of Table-6. These estimated values of log-linear parameters can be interpreted similarly.

4.2.2 Data between Jan 2009- May 2009

Again, we take same company data but different period. We now create a similar 3-dimensional table for the period of Jan 2009 to May 2009. The contingency table is given below:

Table 12. 3-dim frequency table of Jan 09- May 09

A: Pattern	B: Returns	C: Volume	
		High	Low
Sell	High	3	4
	Moderate	7	18
	Low	7	6
Neutral	High	3	2
	Moderate	6	6
	Low	2	1
Buy	High	9	11
	Moderate	6	10
	Low	2	2

Table-12 of test statistic with respect to the hypotheses suggests that hypotheses H_1 , H_3 and H_4 cannot be rejected at 95% level of significance. The non-zero interaction parameters are estimated and given in the below tables:

Table 13. Hypotheses and Z(H) value table of Jan 09- Mar 09

Hypo	Z(H) Value	Df	$\chi_{0.95}^2$
H_1	2.417	6	12.592
H_2	16.391	6	12.592
H_3	3.763	8	15.507
H_4	5.431	10	18.307
H_5	19.405	10	18.307
H_6	18.060	8	15.507
H_7	21.074	12	21.026

Table 14. Value of $\hat{\lambda}_i^A$ for 3-dimensional table

$\hat{\lambda}_i^A$	i=1	i=2	i=3
	0.1609	0.0248	-0.1857

Table 15. Value of $\hat{\lambda}_j^B$ for 3-dimensional table

$\hat{\lambda}_j^B$	j=1	j=2	j=3
	0.0025	0.2662	-0.2687

Table 16. Value of $\hat{\lambda}_k^C$ for 3-dimensional table

$\hat{\lambda}_k^C$	k=1	k=2
	0.00005	-0.00005

Once again the tables of parameters exhibit similar pattern, that is, when the pattern is Sell, there are more chances of return being low. For this data we get similar results.

Table 17. Value of $\hat{\lambda}_{ij}^{AB}$ for 3-dimensional table

$\hat{\lambda}_{ij}^{AB}$	j=1	j=2	j=3
i=1	-0.2634	-0.0165	0.2799
i=2	-0.0527	0.0727	-0.0200
i=3	0.3161	-0.0562	-0.2599

5 Conclusion and future aspect

We analyze the relationships among pattern, returns and volume of transactions of stock market data. The data is presented in the form of contingency tables. These tables are analyzed by using log-linear modeling and the hypotheses of interactions are tested by using chi-square test statistic. We use these tables for further analysis to see the strength of relationship among the variables by using the maximum likelihood estimates of various parameters of interaction. The Maruti-Suzuki Co. stock data clearly shows that pattern and returns are independent of the volume of transactions. Also, the log-linear model parameters show that the influence of categories of each of these variables will not depend upon the categories of other variables uniformly. The selling pattern and low value of next day return have more correlation than the other categories. Also, this analysis is not affected by Turn of the year effect.

One can similarly include more variables to analyze the multi-dimensional contingency tables. The construction of categories can also be defined by using other technical indicators, such as relative strength index, principal volume oscillator etc. These tables may be further expanded to include more categories of each of the variables, for example, we have included two categories of volume and similarly three categories of other two variables.

Acknowledgements

Second author is thankful to the ministry of human resources and development (MHRD) for the fellowship grant. We thank the institute, IIT Jodhpur, for providing all the required facilities. We are also thankful to the referees for their valuable inputs to strengthen the paper.

REFERENCES

- [1] Anderson, E. B. (1996). *Introduction to the Statistical Analysis of Categorical Data*. Springer, New York.
- [2] Bollinger, J. (2002). *Bollinger on Bollinger Bands*. McGraw-Hill, ISBN 0-07-127368-3.
- [3] Boulesteix, A.L. (2006). Maximally Selected Chi-squared Statistics for Ordinal Variables. *Biometrical Journal* 48, 451 - 462.
- [4] Brad, M. B. and Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors. *The Review of Financial Studies*, 21(2), 785-818.
- [5] Hoque, M. S. and Gias, S. (2009). Understanding the Technical Market Indicators. Unpublished Article.

- [6] Kateri, M. (2014). *Contingency Table Analysis: Method and Implementation Using R*, Birkhäuser.
- [7] Slavkovic A. B. (2006). *Analysis of Discrete Data*. Lecture Notes, Pennsylvania State University
- [8] Vellaisamy, P. and Vijay, V. (2007). Some Collapsibility Results for N-Dimensional Contingency Tables. *Ann.Inst.Statist.Math*, 59, 557 - 576.
- [9] Vijay, V. (2011). Relationships Between Full and Layer Models With Applications to Level Merging. *Communications in Statistics-Theory and Methods*, 40(4), 745 - 761.
- [10] Vijay, V. and Paul, P. (2015). A New Trading Band for Prediction of Buy and Sell Signals and Forecasting of States. *International Journal of Applied Management Sciences and Engineering*, 2(2), 34-54.
- [11] Ritter, J. R. (1988). The Buying and Selling Behaviour of Individual Investors at the Turn of the Year. *The Journal of Finance*, 42(3), 701-717.