

Fitting the Relationship between Financial Variables and Stock Price through Fuzzy Regression Case study: Iran Khodro Company

**Seyed Heidar Mirfakhr- Al –Dini
Hassan Dehghan Dehavi,
Elham Zarezadeh**

Islamic Azad University
Science Research-Yazd Branch, Iran

**Hamed Armesh
Mahmoud Manafi
Saeed Zraezadehand**

Doctoral of Business Administration candidates
Multimedia University, Malaysia

Abstract

One of the best ways of investment is investing in stock exchange. For investing in stock, it is a crucial fact to be sure of the reasonable share price in the determined date and also predicting the future changes in stock price. Because some parts of the returns out of this type of investment comes from changes in the stock exchange. Thus, the increasing tendency is observed towards stock price prediction in the capital market. So far many researchers have tried to discover the relationship between the stock price and financial and non-financial variables by using the regression method. However, the fuzzy regression is not used thoroughly for finding this relationship. In the present study, the regression method based on the fuzzy sets theory has been used to fit the relationship between the financial variables and stock price of Iran Khodro Company (case study) from 1377 till 1386. The financial variables of this study are: Earning per Share (EPS), Dividends per Share (DPS) and Price to Earnings ratio (P/E). Eventually, the fuzzy linear regression model for examining the relationship between DPS, EPS and P/E variables and stock price of Iran Khordro Company has been presented. The empirical results of this research indicate that there is a positive and significant relationship between Earning per Share (EPS) and stock price of the company. However, there is a negative and significant relationship between Dividends per Share (DPS) and Price to Earnings ratio (P/E) of the said company.

Key words: Stock prices, Fuzzy regression, Dividends per Share, Earning per Share, Price to Earnings ratio.

Introduction

The common purpose of all investors is to seek profit and to increase their wealth in the capital markets. Nowadays, in the developed countries, most of the real and legal investments are carried out through stock markets. Indeed, stock market is one of the most important investment centers in these countries. Therefore, the stock listed companies are trying to augment shareholders wealth by increasing the financial activities and adopting the best distribution policies. The development of the stock exchange markets has a considerable role in country's economic growth. Consequently, some policies should be adopted to make the investors show more interest in investing in stock markets. In order to encourage investors to gain financial assets, some conditions should be provided through which the investors would be able to obtain the required knowledge for appropriate economic decision making. At the time, both the individual and company investors have a lot of concern regarding the investment in the stocks and these concerns are due to the extreme fluctuation of share prices. The extreme fluctuations of share prices are also affected by many factors and identifying these factors help to predict the stock price and determine the reasonable purchase price. Thus, the main objective of the researchers, investors and financial analysts is the knowledge of effective factors for stock price prediction.

Statement of the problem

The knowledge of effective factors on the stock price is very important while deciding to invest in stock exchange. With the increase in the amount of stock transactions (due to the implementation of article 44 of constitution) and activation of the stock exchange organization, the stock pricing gained great significance in Iran.

In this respect, many empirical studies have been conducted to evaluate the relationship between financial and non-financial variables and stock prices in market and also to predict stock prices with different mathematical and statistical methods. But so far, less of fuzzy linear regression has been used. The incentive of this research is to study the relationship between financial variables (EPS, DPS, P/E) and the stock prices with the fuzzy linear regression method.

Research variables calculation

This study examines the relationship between dependent variables (stock price) and independent variables (Dividends Per Share, Earning Per Share, Price to Earnings ratio) through fuzzy linear regression method. The data which has been used in this study is collected from Iran Khodro Company. In order to calculate EPS and DPS variables, the related data for a period of 3 month for 10 years (1377 to 1386) has been extracted through the annual report of general meeting of board of directors and also Rahavard Novin Cd , so that the total number of each variable is equal to 40. The only constrain is the DPS related data as there was only one DPS available for each year the data related to this variable are divided by 4.

The price to earnings ratio will be calculated using the data included in Iran Khodro's annual report of general meeting. For this purpose the after-tax data related to stock price average, common stock number and the earnings (loss) will be extracted from annual report of general meeting at the date of meeting, then the common stock numbers will be multiplied by the stock price average and the result is divided by the after-tax net earnings (losses). The dependent variable (stock price) can be obtained through the Rahavard Novin Cd. After receiving Iran Khodro's latest stock price at the end of each month, the data are transferred to the excel software and 3 month stock price average will be calculated for a 10 year period (1377 to 1386).

Research questions

The followings are the major research question which indicates the researcher's main motivation in conducting this research:

Is there a significance relationship between financial variables and Iran Khodro's stock price?

The selected financial variables are Price to Earnings ratio, Earning per Share and Dividends per Share.

3 minor questions are formed based on the main research quest ion:

- 1- Is there any significant relationship between Earning per Share and Iran khodro's stock price?
- 2- Is there any significance relationship between dividends per share and Iran khodro's stock price?
- 3- Is there any significance relationship between Price per Earnings ratio and Iran khodro's stock price?

Review of literature

The fuzzy set theory is presented since 1965. This theory provides a new mathematical model for analyzing the ambiguous concepts and variables and also systems which are based on approximate relationships. It should be noted that the aim of classic statistics theory is to study the patterns and uncertain statistical systems. The fuzzy set theory is applicable in the situation in which the dominant uncertainty on system and the desired patterns are of possibilistic and not of likelihood type. In a regression model, the error may be the result of the ambiguous and approximate relation between the variables rather than the uncertainty known as random error. In a regression model, you may face the ambiguous observations. In this case, the statistical regression model can be replaced by the fuzzy regression model.

The fuzzy regression , for the first time studied in 1982 by Tanaka et al., in fact they paid attention to the fuzzy coefficient regression (in other words: possibilistic regression). Their approach was followed by some researchers like Peters (1984), Luczynski & Matloka (1995) and Tanaka& Lee (1999) and expanded in different aspect. In 1982, Beger discussed the simplest fuzzy regression method with a different approach. In a regression model which is the result of common data, He predicted the value of dependent variable by putting the fuzzy values in the place of independent variable. Some application of this method were discussed and analyzed by Hshmati & Kandel in 1985. The distance regression analysis was described by Tanaka in 1998. In 1998, a method for the least fuzzy regression squares with the triangular number was presented by Diamond. Kandel et al. who extended this method to a non-triangular fuzzy numbers.

Triangular fuzzy numbers

Those types of fuzzy sets which are more popular in applicable problems and required to follow specific patterns in mathematical calculations are known as triangular fuzzy numbers.

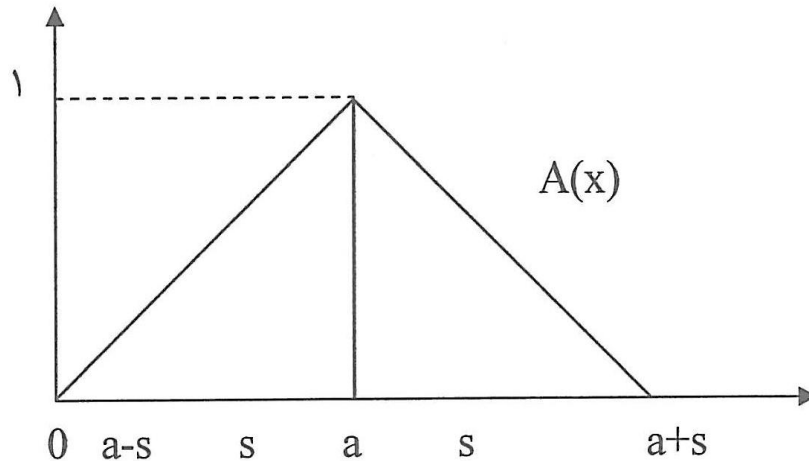
The triangular fuzzy numbers are often used in management decision making, business and finance and social sciences.

The triangular fuzzy numbers are shown like $\tilde{A} = (a, s^L, s^R)_T$.

a is the center value and S^L and S^R are left and right spreads of \tilde{A} respectively. If the left and the right spreads are equal ($S^R = S^L$), the triangular fuzzy number will be symmetric. In the case of inequality of these two values ($S^R \neq S^L$), the \tilde{A} fuzzy number will be asymmetric.

The membership function of a triangular fuzzy number is as follows:

$$\tilde{A}(x) = \begin{cases} 1 - \frac{a-x}{s^L} & a - S^L \leq x \leq a \\ 1 - \frac{x-a}{s^R} & a \leq x \leq a + S^R \end{cases}$$



Fuzzy regression model

It is assumed that in a regression model with a fuzzy coefficient, the difference between the observed value of dependent variables and the model’s estimated values is caused by the ambiguity which exist in model’s structure. This ambiguity is applied to the fuzzy numbers which are considered as model’s coefficients. In the under study regression model, the model coefficients are indicated as fuzzy numbers but the variables and the related observations are the real numbers. Therefore, in this type of linear regression model the goal is based on all observations (Y_m, X_m) and... (Y_1, X_1), the fuzzy coefficients (numbers) $\tilde{A}_n, \dots, \tilde{A}_1, \tilde{A}_0$ are obtained in a way that the following would be the optimal model:

$$\tilde{Y} = F(x^0, A) = \tilde{A}_0 + \tilde{A}_1 x_1 + \dots + \tilde{A}_n x_n$$

Y is dependent variable or fuzzy output , $X = [x_1, x_2, \dots, x_n]$ is independent variable vector or input vector with real numbers and $A = \{A_0, A_1, \dots, A_n\}$ is a set of triangular fuzzy numbers.

If the \tilde{A}_i ’s, ($i=0, 1, \dots, n$) are the symmetric triangular fuzzy numbers and the x_i ’s are the real numbers, therefore based on the definition and the characteristics related to the fuzzy number’s total subtracting and scalar multiplication, the fuzzy output would be a symmetric triangular fuzzy number like $\hat{Y} = (F^C(x), F^S(x))$ in which $F^C(x)$ is center and $F^S(x)$ is \hat{Y} spread.

Determination of fuzzy coefficient

In a fuzzy regression, in order to estimate model’s parameters (fuzzy coefficient of the model) based on sets of observations some criteria should be taken into account.

In fact the goal of fuzzy regression method with a non-fuzzy data is to determine \tilde{A}_i , $i=1, \dots, n$ coefficient in the model in a such a way that:

Firstly, the fuzzy output \hat{Y} , for all values of ($j = 1, \dots, m$), Y_j at least include a large degree of membership as large as h . It means $\hat{Y} = (Y_j) \geq h$, $j = 1, \dots, m$.

This condition ensures that in the final model, the value membership Y_i , namely the i^{th} observed value of dependent variable of model at least should be the amount of h . In order to obtain the best fitting model, the value h is selected by the decision makers and it is the fitting degree of estimated linear model to a set of data. In addition, the parameter h can be defined as model’s validity and certainty level.

For fitting the fuzzy regression model of stock price and financial variables, the certainty or validity level is considered 0/51 ($h=0/5$) which is a common and reasonable certainty level.

Secondly, the ambiguity or fuzziness of model's output is the least possibility. This condition represents the fact that, in predicting the amount of dependent variables the ambiguity should be minimized.

The more is the spread of triangular fuzzy number, the more would be its ambiguity. Therefore, in all data, the total of **Y** fuzzy output spreads should be minimized to the lowest value.

There is an output from the model for each observation; therefore the total output ambiguity should be minimized. Means:

$$[1] \quad Z = 2ms_o + 2\sum_{i=1}^n \left[Si \sum_{j=1}^m X_{ji} \right]$$

In the above equation, **X_{ji}** represents the **jth** observation for the **ith** variable.

Constrains of the model:

$$(1-h)S_o + (1-h)\sum_{i=1}^x (S_i x_{ji}) - a_0^c - \sum_{i=1}^x (a_i^c x_{ji}) \geq -y_i \quad j = 1, 2, \dots, m$$

$$(1-h)S_o + (1-h)\sum_{i=1}^x (S_i x_{ji}) - a_0^c - \sum_{i=1}^x (a_i^c x_{ji}) \geq -y_i \quad j = 1, 2, \dots, m$$

In the above equation, **X_{ji}** represents the **jth** observation for the **ith** variable.

Be aware that there are two constrains for each observations. Therefore, the problem is finding the fuzzy coefficient of the model, equivalent to the minimization of **z** objective function with respect to the **2m** produced constrains by **m** observation.

Table 1 indicates the data related to the utilized variables in the process of model making.

The **y_i** represents Iran Khodro's stock price during the years 1377 till 1386. The **x_j**'s indicates independent variable in which **x_{j1}**, **x_{j2}**, **x_{j3}** are Earning per Share (EPS), Dividends per Share (DPS) and Price to Earnings ratio (P/E) respectively.

Table1- dependent and independent variables data

(j) Number	yj	Xj1	Xj2	Xj3	(j) Number	yj	Xj1	Xj2	Xj3
1	5095	473	250	3/1	21	3122	152	54	1
2	4525	473	250	3/1	22	8243	152	54	1
3	5040	473	250	1/3	23	8038	152	54	1
4	4803	473	250	1/3	24	5693	152	54	1
5	5580	511	250	4/9	25	6630	220	100	5/2
6	6213	511	250	4/9	26	5581	220	100	5/2
7	5482	511	250	4/9	27	4943	220	100	5/2
8	3909	511	250	4/9	28	4287	220	100	5/2
9	4486	292	200	7/8	29	4685	261	225	3/9
10	3955	292	200	7/8	30	3812	261	225	3/9
11	3518	292	200	7/8	31	3143	261	225	3/9
12	3545	292	200	7/8	32	3012	261	225	3/9
13	3358	196	130	5/3	33	3303	269	155	6/4
14	3308	196	130	5/3	34	3036	269	155	6/4
15	3190	196	130	5/3	35	9339	269	155	6/4
16	3123	196	130	5/3	36	2763	269	155	6/4
17	1113	165	138	14/8	37	2322	114	65	3/6
18	3213	165	138	14/8	38	1976	114	65	3/6
19	2698	165	138	14/8	39	1823	114	65	3/6
20	7630	165	138	14/8	40	1699	114	65	3/6

Based on equation 1, the objective function for 40 observations is as follows:

$$MinZ = 40S_o + S_1 \sum_{j=1}^{40} x_{j1} + S_2 \sum_{j=1}^{40} x_{j2} + S_3 \sum_{j=1}^{40} x_{j3}$$

Let the \tilde{A}_i 's be symmetric ($i=0, 1, \dots, n$), the constrains are as follows:

[2]

$$(1-h)S_o + (1-h)\sum_{i=1}^n (S_i X_{ji}) - a_o^c - \sum_{i=1}^n (a_i^c X_{ji}) \geq -y_i \quad j = 1, 2, \dots, n$$

$$(1-h)S_o + (1-h)\sum_{i=1}^n (S_i X_{ji}) - a_o^c + \sum_{i=1}^n (a_i^c X_{ji}) \geq y_i \quad j = 1, 2, \dots, n$$

X_{ji} represents the j^{th} observation for the i^{th} variable. There are two constrains for each observations.

The real numbers of the functional variables depends on the number of data sets variables.

In this study there are 80 constrains for 40 observations.

Let The z objective function (in symmetric state) be minimized with the use of linear programming method:

$$MinZ = 40S_o + 1061S_1 + 6268S_2 + 224S_3$$

The above mentioned equation should be minimized with respect to the following constrains by using the equation [2] and assuming $h=0/5$. Thus, There are 80 constrains:

$$1) 0/5S_o + 236/5S_1 + 125S_2 + 1/55S_3 + a_o^c + 473a_1^c + 250a_2^c + 3/1a_3^c \geq 5095$$

$$2) 0/5S_o + 236/5S_1 + 125S_2 + 1/55S_3 - a_o^c - 473a_1^c - 250a_2^c - 3/1a_3^c \geq -5095$$

⋮
⋮
⋮

$$79) 0/5S_o + 57S_1 + 32/5S_2 + 1/8S_3 + a_o^c + 114a_1^c + 65a_2^c + 3/6a_3^c \geq 1699$$

$$80) 0/5S_o + 57S_1 + 32/5S_2 + 1/8S_3 - a_o^c - 114a_1^c - 65a_2^c - 3/6a_3^c \geq -1699$$

By Considering the objective function and taking into account the 80 corresponding constrains, the objective function minimized with the use of linear programming method. The LINGO software is used for this purpose. It should be noted that in the linear programming problems, the problems variables are assumed non-negative. But in fuzzy regression model, the coefficient can be negative values.

Therefore, in the minimization programs of objective function, the variables of interest are considered free in sign.

Regression results

This part provides the fitting results of the relationship between Earning per Share, Dividends per Share and Price to Earnings ratio variables with regard to Iran Khodro's stock price. With the use of LINGO software and solving the above model, the following responds are obtained in table2.

Table2- fuzzy regression set of responds

The minimized value of objective function	center of Fuzzy number				Spread of fuzzy number			
Min Z	a_o^c	a_1^c	a_2^c	a_3^c	S0	S1	S2	S3
272342/1	8/7	-32/97	-138/54	9536/09	10/45	-76/34	-72/94	16408/32

Source: researcher's finding

Based on the obtained results of table2-4, the coefficients of fuzzy model are stated as follows:

$$\tilde{A}_o (9536/09, 1640832)_T, \tilde{A}_1 (8/70, 1045)_T,$$

$$\tilde{A}_2 (-32/97, -76/34)_T, \tilde{A}_3 (-138/54, -72/94)_T$$

Thus, the optimal fuzzy model is as follows:

$$P = (9536/09, 1640632)_T + (8/70, 1045)_T \text{ EPS} - (32/97, 76/34)_T \text{ DPS} - (138/54, 72/94)_T \text{ P/E}$$

The fitting results of the relationship between financial variables and stock price by fuzzy linear regression method is summarized in table3.

Table 3- fitting results of the relationship between financial variable and stock price by fuzzy regression method

	Coefficients of fuzzy regression model (fuzzy number spreads, fuzzy number center)			
$P = \tilde{A}_0 + \tilde{A}_1(EPS) + \tilde{A}_2(DPS) + \tilde{A}_3(P/E)$	\tilde{A}_0	\tilde{A}_1	\tilde{A}_2	\tilde{A}_3
	(9536/09, 1648/32)	(8/70, 10/45)	(-32/97, -76/34)	(-138/54, -72/94)

Source: researcher’s finding

As it can be observed in table3, the calculated parameters in fuzzy regression model are fuzzy numbers which include the central value, the degree of ambiguity and their possible uncertainty.

In other words, based on the above fitted model(with regard to $h=0.5$ certainty level),for \tilde{A}_0 the symmetrical triangular fuzzy number like (9536/09 and 1648/32) means that it is not possible for the \tilde{A}_0 value to be less than -6872/23 and more than 25944/41. For EPS, the symmetrical triangular number like (8/70 and 10/45) means, the EPS value can to be less than -1/75 and more than 19/15. For DPS, the symmetrical triangular number like (-32/97 and -76/34) means the DPS cannot be less than 43/37 and more than 109/31. For P/E variable, the symmetrical triangular fuzzy number like (-138/54 and -72/94) shows the impossibility of P/E value to be less than -65/6 and more than -211/48.

Data finalization and decision making

One of the most important steps in application of fuzzy sets theory is non-fuzzification stage. At this stage the data and the results obtained from applying the fuzzy method turns into ordinary number. This is called finalization (non- fuzzy). Summarizing the existing data into a numeric value is more like a conversion from indefinite concept to a definite one. A fuzzy set includes a certain amount of ambiguous data and also a numerical value which contains an absolute data. Therefore, converting the ambiguous data into the exact and accurate amount of data indicates that the ambiguity can be removed by summarizing the main source of data.

In other words, with non-fuzzification action the ambiguous data sources are compacted into unambiguous data source. (Alhoghli pour omran, 1383) In this study, for finalizing the data, the high and low limit method has been used which is in accordance with the following formula:

$$\frac{a + 2b + c}{4}$$

The finalization result is represented in table 4.

Table4- fuzzy regression coefficients finalized value

P/E	DPS	EPS	variables
-138/5	-32/9	8/7	finalized value

Source: researcher’s finding

According to the fuzzy regression finalized coefficients which are presented in table 4, the regression model equation would be as follows:

$$Y = 9536 + 8/7X_1 - 32/9X_2 - 138/5X_3 + \varepsilon$$

The presented regression equation can be categorized as follows:

The intercept of equation is 953.

The algebraic expression $8/7x_1$ denotes that as per one unit increase in Earnings per Share, the stock price increase to the level of 8/7 unit.

The algebraic expression $-32/9 x_2$ denotes that se per one unit increase in dividends per share, the stock price decreased into the level of 32/9 unit.

The algebraic expression $-138/5 x_3$ denotes that as per one unit increase in price to earnings ratio, the stock price increased to the level of 138/5 unit.

Conclusion

Many accounting problems comprises the variables and relationships which are difficult to measure accurately (if not possible). On the other hand, the contingency judgments are not accurate.

Moreover the costs and benefits in many accounting problems are not accurately calculable. For example, the accurate measurement of the costs and interest of internal accounting control is not possible. In study of the system with average and inaccurate variables, considering the Fuzzy structure is necessary for system's modeling. In the study of the relationship between financial variables (PIE,DPS,EPS) and stock price, this structure can be considered as a Fuzzy linear function in which the parameters are determined by the Fuzzy sets. The present study has utilized the linear regression with Fuzzy coefficient for modeling the relationship between financial variables (P/E, DPS, and EPS) and stock price of Iran Khodro. The result of fitting the Fuzzy regression model suggests that there is a relationship between Earning per Share, Dividends per Share and Price to Earnings variables and stock price of Iran Khodro Company.

There is a positive relationship between (EPS) and stock price. Thus, with the required information regarding the earning per share which is one of the remarkable criteria in evaluation of companies' financial performance, one can predict the stock price. In other words, earning per share changes can be used for predicting the stock price. There is a negative relationship between Dividends per Share (DPS) and Iran Khodro's stock price. There is also a negative relationship between Price to Earnings ratio and stock price. It is predicted that the more the ratio amount decreases, the more the stock price increases.

References

1. Alhoghli Pouremran.N.1383. Fuzzy regression with coefficient. Master thesis, Tehran Tarbiat Moallem University *between Stock Price and EPS: Evidence Based on Taiwan Panel Data*, Economics Bulletin, Vol.3.,No.30,PP.1-12
2. Chang,H.L., Chen, Y.S., SU, C.W., Chang, Y.W.,2008, *The Relationship between Stock Price and EPS: Evidence Based on Taiwan Panel Data*, Economics Bulletin, Vol.3.,No.30,PP.1-12
3. Chang,P.T., 1994,*Fuzzy Regression Analysis*, Ph.D, Kansas State University.
4. Chang, P.Ch. and Chen, H., 2008, *A TSK Type Fuzzy Rule Based System for Stock Price Prediction*, Expert Systems with Applications, Vol.34,PP.135-144.
5. Chen, S.P. and Dang,J.F.,2008, *A Variable Spread Fuzzy Linear Regression Model with Higher Explanatory Power and Forecasting Accuracy*, Information Sciences , Vol.178, PP.3973-3988.
6. Cheng,Ch.B.,1998, *Nonparametric Fuzzy Regression and Its Applications*, Ph.D,Kansas State University.
7. Choi, S.H., Dong,K.H., 2004, *Note on Fuzzy Regression Model*, In: Proc.of the 7th Iranian Statistical Conference, Allameh- Tabatabaie University, PP.51-55.
8. Klir, G.J., Bo Yoan, 2002, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, New Jersey.
9. Majidi, S; Taheri, S. M; Alamatsaz .M, h, 1380. Linear regression with Fuzzy coefficient. Iran Sixth Statistic Conference Proceedings. Tarbiat Modares University.
10. Moskowitz, H. and Kim, K.,1993, *On assessing the H value in Fuzzy Linear Regression*, Fuzzy sets and systems, vol.58,PP.303-327.
11. Sanchez, J. and Gomez,A.T., 2003, *Applications of Fuzzy Regression In Actuarial Analysis*, Journal of Risk and Insurance, Vol.70,No.4, PP.665-699.
12. Taheri, S.M.,2003, *Trends in Fuzzy Statistics*, Austrian Journal of Statistics, Vol.32,No.3,PP.239-257.
13. Taheri, S, M, 1378. Introduction to Fuzzy sets theory. Mashahd, Jahad-e- Daneshgahi publishers, Boston.
14. Tanaka, H., Uejima, S.and Asai, K. 1982, *Linear Regression Analysis with Fuzzy Model*, IEEE Trans., Systems Man cybernet., Vol.12,PP.903-907.
15. Tanaka,H., Hayashi, Watada, 1989, *Possibilistic Linear Regression Analysis for Fuzzy Data*, Euro.J.Oper. Res., Vol.40,PP.389-396.
16. Wang, H.F. and Tsaur,R.C., 2000, *Insight of a Fuzzy Regression Model*, Fuzzy sets and Systems, Vol.12,PP.355-369.
17. Yen, K.K., Ghoshray,S. and Roig, G., 1999, *Linear Regression Model Using Triangular Fuzzy Number*, Fuzzy sets and systems , Vol.106, PP.167-177.
18. Zadeh,L.A., 1965, *Fuzzy sets, Information and control*, vol 18,PP.338-353.
19. Zebda, A., 1991, *The Problem of Ambiguity and vagueness in Accounting* , Behavioral Research in Accounting, Vol.3, PP.117-145.
20. Zimmermann, H.J., 1996, *Fuzzy set Theory and its Applications*, Kluwer Academic publishers, Boston.