General Letters in Mathematics (GLM)

Identifying factors influencing decision making using logistic regression

Obaid Mahmmood Mohsin Al Zawbaee [a*], Hassan Mstafa Tabra [b]

[a] *Prof. Dr., Cihan University Sulaimanya Camp Sulaimaniya, Kurdistan Region- Iraq*
[b] *Ass. Prof. Dr, Cihan University Sulaimanya Camp Sulaimaniya, Kurdistan Region- Iraq*
Emails: [a] obaed79@yahoo.com, [a] obed.muhsin@sulicihan.edu.krd, [b] hassan.tabra@sulicihan.edu.krd

Abstract:

The issue of decision-making is one of the important issues in modern management because of its impact on individuals, communities and countries, as most of the problems faced by individuals, communities and countries result from making incorrect decisions. The evolution of management science and the use of quantitative methods of treatment are also in a continuous growth as decisions are influenced by a variety of factors or variables. In this research, we considered the variables or factors related to the illusion of excellence and the level of ambition through the use of logistic regression technique to identify the variables that have a significant moral effect in the decision-making and are arranged according to their importance. A model by which to predict the extent to which the decision is affected by the environment. The result of the analysis demonstrated that out of ten factors that were identified and studied, three of them had a significant effect while the others had a non-significant effect, and the model achieves a correct classification of 81%.

1. Introduction

Decision making is considered a significant topic in modern management science. This is due to the fact that many things which go beyond the limits of the individuals clearly influence the policies of the societies and countries. This influence can be spotted in many political, economic, social and educational aspects. Consequently, decision-making will have a great significant since most of the problems faced by individuals which in turn faced by the entire societies, are the result of incorrect decision- making. A decision is considered a course chosen by the decider being the most appropriate available means to achieve a certain goal [5,21]. As a result of what modern management science witnessed in the development of analysis and treatment, additional cognitive sciences are involved in decision-making. In fact, the decisions in the age of cognition and globalization are taken by following the scientific methods relying on the computers and by the quantitative methods to have a precise decision [6]. The process of decision-making faces lots of problems influencing the levels of performance the most important of which those which result from the growth of the needs and desires in the inside and outside environment [3]. Hence, the environment of decision-making will be influenced by the existing factors. Within such environment, the decision made by the individual is influenced by many elements like, for example, the level of ambition, the illusion of preeminence and so on. Concerning the illusion of preeminence, it is one of the mental deviations which influences decision making. This kind of illusion is represented by the individual's inclination of having more supreme mentality and this leads to the deviation of the decision than what is rightly aimed at [15]. In this research, the factors concerned with the illusion of preeminence and the level of ambition are considered through employing the technique of logistic regression to identify the variables or factors which influence decision making. This is achieved

---

* Corresponding author

Email address: obaed79@yahoo.com, obed.muhsin@sulicihan.edu.krd (Obaid Al Zawbaee)

by determining ten variables then collecting their data through a questionnaire prepared for this purpose. Consequently, a model is built which can identify the variables with the moral influence on the studied phenomenon in addition to arranging those variables according to their importance [2].

It is well-known that the process of decision making is an integrated and a comprehensive process and out of such aspect the research problem emerged which can be represented by answering the following question:

- What are the variables of the significant influence on decision making motivated by preeminence illusion and ambition level?
- What is the arrangement of those variables concerning the importance of the capability of prediction?

The research value lies in identifying the organic relations between the environment and the process of decision making by the student and also in revealing the way of how the effects are arranged through determining the variables of the significant influence on the decision and then arranging them in accordance with the power of influence. This leads to arriving at suggestions which increase the precession of decision making in a way that develops his character in general. However, the aim of this research is building a statistical model which enables the prediction of decision making and its relation with the ambition level and preeminence illusion. In addition, the research relies on the hypothesis of the influence of all independent variables, i.e., all the logistic regression parameters for all independent variables are significant.

2. Previous Studies

Mansoor [17] his study aimed at identifying the nature of the relation between the level ambition and decision making by the students. The model included 270 students from the departments of psychology and social sciences. The study concluded that there was a connecting relation with a statistical indication between ambition level and decision making by the sample. The study also concluded that there exist differences with significant indication in the mean average of the sample's performance on the scale of the ambition level.

Dhaifallah & obaid, [8] Also their study aimed at identifying the relation between decision making and the ambition level and diagnosing the capabilities of the students to make a decision. The study comprised 30 first year M.A. students from the College of Social and Humanitarian Sciences. The study concluded that decision making is connected with ambition level and also the existence of differences with statistical significant in making decision according to the specialized factor.

According to Alsmarat [4] his study aimed at diagnosing the influence of using the strategy of problem solving in teaching national education on developing the skills of decision making to a model of a students. The study demonstrated the preeminence of the empirical group in developing the skills concerned with making decision. It also demonstrated the preeminence of males over females and also there were no differences with statistical significant among the educational levels.

On the other hand, Ribanks [19] his study aimed at identifying the relation between mentality, personality characteristics and ambition levels. The model of the study comprised 1500 secondary and university students. The study concluded that there is an indicative and positive connection between the capability of mentality and achievement and some personality characteristics and ambition level. It also concluded that there exists a huge difference between the ambition level and both gender variable and specialty in favor of males.

Study by Kimm [12] aimed at analyzing the relation between participatory in decision making and job satisfaction of teachers there. The study comprised a model of 700 male and female teachers all over Korea. The study concluded that there exist differences with statistical indication between teachers' participation in making decision and job satisfaction because of the gender, experience, the size of the school and the subjects taught by the teacher. It was demonstrated that the real levels of the participation had a positive influence.

3. Logistic Regression

Achia, et al., [1] logistic regression is a statistical means which is used to examine and document the relation between the qualitative dependent variable and one or more variable of the independent variables irrespective of their kind whether quantitative or qualitative. The logistic regression is based on the basic assumption that the dependent variable (y) which we deal with is a binary variable which takes the values (1) with (P) probability and (0) with (1-P) probability and is:

$$E(y1x) = P(y = 1) = P \tag{1}$$

Thus, the value of the right side is between (1,0). The model is inapplicable from the point of view of regression [10,7]. To solve this problem, we insert a proper mathematical transformation on the dependent variable (y) where

$$0 \leq P \leq 1$$

Thus, the ratio ($\frac{P}{1-P}$) is a positive amount (0 ,∞) that is,

$$0 \leq \frac{P}{1-P} \leq \infty$$

Taking the natural logarithm, we get

$$-\infty \le \ln\left(\frac{P}{1-P}\right) \le \infty \tag{2}$$

Consequently, we can write the regression model in case of one independent variable:

$$\ln\left(\frac{P}{1-P}\right) = b_0 + \ b_1 \ x_1 \tag{3}$$

In case there is more than one independent variable, the model is

$$\ln\left(\frac{P}{1-P}\right) = b_0 + \sum_{i=1}^{n} b_j \ x_{ij} \tag{4}$$
$$j = 1, 2, ..., k$$
$$i = 1, 2, ..., n$$

Equation (4) above represents the logistic regression model and the transformation $\ln\left(\frac{P}{1-P}\right)$ is called (Logit Transformation) [11,9,14].

The logistic regression is a model which takes the shape of the logistic function as in figure (1) and is used in predicting the probability of a certain action to happen by documenting the data in a logistic curve shape [13].
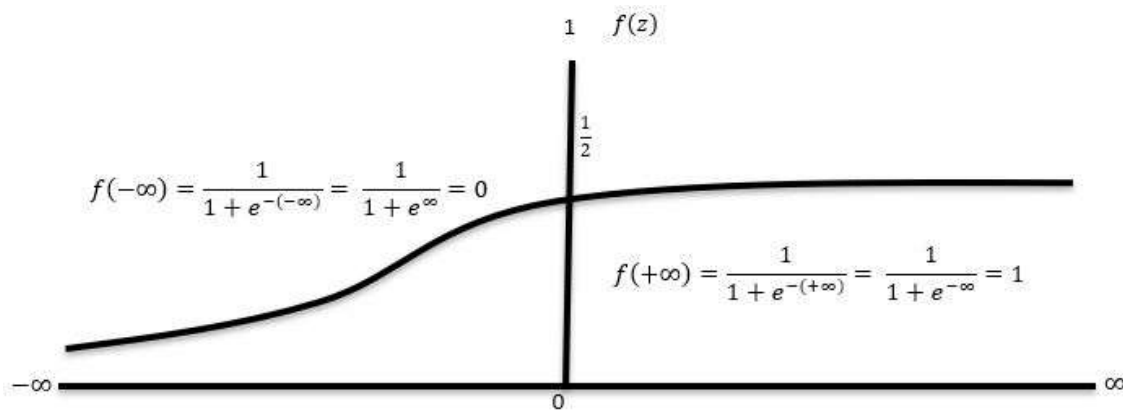
Figure (1): Logistic Function

The reliance of the regression in the above figure on the logistic curve which has the formula as in the equation

Range

$$0 \le f(z) \le 1$$

$$P = \frac{e^{a+bx}}{1 + e^{a+bx}} = \frac{1}{1 + e^{-(a+bx)}} \tag{5}$$

Instead of the straight-line equation $(y = b_0 + b_1 x + e)$ is the thing which made the topic of regression to be called logistic regression. From the characteristics of the logistic regression analysis is that it arranges the influence of the independent variables in a way which enables the researcher to conclude that a variable is considered more powerful than the other one.

In addition, the logistic regression does not require that the independent variables follow normal distribution and it also does not require that the relation between the independent variables and the dependent variable is to be a linear [18,20,16].

4. The Practical Side

The data used in the research are identified and determined by describing the variables of the research and determining sample size and the statistical analysis.

4.1 Sample and Data of the Research

The research sample is represented by 270 students from Cihan University-Slemani and the following adopted mechanism to choose the sample from population size (the total number of Cihan university students targeted in the research and did not include Law students and some new ones)

N= 830

$n=N/ [1+ N(d)^2]$
where:
n=sample size, N=population size, d = error = α =0.05
n= {830 / {1+830(0.05)^2} =270
The stratified Random sample is adopted and which takes into consideration the number of the students for each scientific department. The sample size is distributed as in the following:
Computer Science=25, English=35, Cultural= 16, Banking=20, Accounting=67, Accounting by IT=34, Business Administration=63, Translating=10
The class of the university students is chosen because they are qualified to make a decision but they lack experience. For this reason, we believe that preeminence illusion and ambition level and can be clearly seen in making decisions for such class. The questionnaire prepared for this purpose is adopted as clarified in appendix (1). The questionnaire form included the dependent variable (y) and ten independent variables and all of them are qualitative as in the following:
y is the dependent variable where
y= (1) the decision is influenced by environment y= (0) the decision is not influenced by environment
x_1 = (1) making decision takes long time from me, and it will take (0) if the answer is making decision does not take long time from me
X_2 = (1) when I have ready alternative decisions, and it will be = (0) when I do not have ready alternative decisions
X_3 = (1) if the respondent answer I rely on reasonable justifications in making decisions and = (0) when the respondent do not rely on reasonable justifications in making decisions
X_4 = (1) if the answer was, I discuss the decision to be made unanimously and = (0) if he answers I do not discuss the decision to be made unanimously
X_5 = (1) when the answer I change my decision after I made it and = (0) if the answer I do not change my decision after I made it
X_6 = (1) when the respondent believe that the personal effort diminishes the obstacles no matter how huge they are, and = (0) if he does not believe that the personal effort diminishes the obstacles no matter how huge they are
X_7 = (1) if the answer is life seems to me hopeless, and = (0) if the life does not seem to me hopeless
X_8 = (1) if the respondent considers the future very optimistic, and = (0) if he does not consider the future very optimistic
X_9 = (1) if the respondent hugely depends on others in solving his problems, and = (0) if don't depend on others in solving his problems
X_{10} = (1) I am much influenced by the criticism of others, and = (0) I am not influenced by the criticism of others.

4.2 Results
Using the ready statistical program (SPSS), the data are analyzed after reviewing and tabulating them. The logistic regression model parameters are estimated relying on Min-2loglikelihood and Iterations. The optimum estimation of the model's parameters is obtained where:
-2Loglikelihood=230.839
In the fifth Iteration which we stopped at because the parameters change (B_0, B_1, ..., B_{10}) became zero. This is the best result that can be obtained for the parameters and that -2loglikelihood is in the minimum as clarified in table (1)

Table (1): Iteration History[a,b,c,d]

Iteration		-2 Log likelihood	Constant	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
									Coefficients				
Step 1	1	238.220	.564	-.145	.104	.465	.330	-.320	.282	.502	-.028	-.026	.233
	2	231.113	.446	-.228	.188	.688	.546	-.486	.439	.855	-.039	-.015	.361
	3	230.840	.402	-.249	.215	.732	.608	-.525	.475	.974	-.042	-.008	.390
	4	230.839	.400	-.250	.217	.734	.611	-.527	.477	.982	-.042	-.007	.391
	5	230.839	.400	-.250	.217	.734	.611	-.527	.477	.982	-.042	-.007	.391

a. Method: Enter
b. Constant is included in the model.
c. Initial -2 Log Likelihood: 252.332
d. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Table (2) shows the model parameters, column (B), and the Standard Error (S.E) for each parameter and (Wald Statistic), in addition to the significant parameters which we will explain later

Table (2): Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	x1	-.250	.345	.527	1	.468	.779	.396	1.530
	x2	.217	.363	.357	1	.550	1.242	.610	2.529
	x3	.734	.359	4.175	1	.041	2.083	1.030	4.211
	x4	.611	.363	2.832	1	.092	1.842	.904	3.752
	x5	-.527	.384	1.887	1	.170	.590	.278	1.252
	x6	.477	.345	1.906	1	.167	1.611	.819	3.171
	x7	.982	.478	4.211	1	.040	2.669	1.045	6.819
	x8	-.042	.387	.012	1	.913	.958	.449	2.046
	x9	-.007	.439	.000	1	.987	.993	.420	2.347
	x10	.391	.435	.809	1	.369	1.479	.630	3.468
	Constant	.400	.667	.360	1	.549	1.492		

a. Variable(s) entered on step 1: x1, x2, x3, x4, x5, x6, x7, x8, x9, x10.

To test the model's efficiency and the goodness of fit, we use here Log Likelihood Ratio which follows the distribution of (Chi-square -$\chi2$) according to the relation

$$\chi2 = 2\,[\text{Log L0-Log L1}]$$

Where:

L_1: the value of Log likelihood function which contains variable (i)

L_0: the value of Log likelihood function which contains variable (i-1)

Table (3): Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	21.493	10	.018
	Block	21.493	10	.018
	Model	21.493	10	.018

Hosmer and Lemeshow (1989) test was adopted to know the quality of model documentation. This test depends on calculating the statistics (χ^2) of the difference between the observed values and the expected ones as in table (4).

Table (4): Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.972	8	.761

We observe that they are not significant and this assures the quality of the model.

To know about the capability of the model which was built on classifying the observations, table (5) (classification table) clarifies this where the percentage of the correct classification: (overall percentage = 81%) to the two classification groups which it belongs to: (1+217)/269=81%. That is, the number of the observations classified correctly is 218 while there are 51 observations which were classified wrongly and the percentage is (19%):

Table (5): Classification Table[a]

			Predicted		
			y .00	1.00	Percentage Correct
Step 1	y	.00	1	47	2.1
		1.00	4	217	98.2
	Overall Percentage				81.0

a. The cut value is .500

From table (2) and with observing column (B) which represents the values of the model parameters, the model is:

Ln (p/(1-p)) = 0.4-0.25X_1+0.21X_2 +0.734X_3 +0.611X_4 -0.52X_5+0.477X_6 +0.982X_7 -0.042X_8 -0.007X_9 +0.391X_{10} (6)

The results in the same table point that the independent variables (X_7,X_3) have significant influence on decision making when ($\alpha = 0.05$) and that the variable (X_4) has a significant influence at ($\alpha = 0.10$).

Thus, the variables (X_7, X_3, X_4), are the variables which influence the dependent variable (y) and that the value of the parameters are:

B_7= 0.982 ------------EXP =2.669

B_3= 0.734 ---------------EXP =2.083

B_4= 0.611-----------------EXP =1.842

- Variable (X_7) occupied the first place in influencing the dependent variable (y).
- Variable (X_3) occupied the second place in influencing the dependent variable (y).
- Variable (X_4) occupied the third place in influencing the dependent variable (y).
- Concerning the rest of the variables, their influence on the dependent variable is not significant

And the adopted model is

$$\text{Ln}(p/(1-p)) = 0.982X_7 + 0.734X_3 + 0.611X_4 \qquad (7)$$

5. Conclusions and Recommendations

The most important conclusions and recommendations which the research arrived at are:

1. Some of the research hypothesis were achieved where the parameters (B_7, B_3, B_4), have significant values whereas the parameters (B_0, B_1, B_2, B_5, B_6, B_8, B_9, B_{10}) have no significant values.
2. The possibility of adopting the logistic regression through building a model by which it is possible to predict making decision.
3. The independent variables with the significant influence are:
   X_7 = life seems to me hopeless
   X_3 = I rely on reasonable justifications in making decisions
   X_4 = I discuss the decision to be made unanimously
4. The percentage of the correct classification for the proposed constructed model was (81%) and which shows whether the decision made is influenced by the environment or not.
5. The goodness of fit for the model was significant where ($\alpha = 0.10$) and the value (chi- square) $\chi^2 = 21.493$ at 10 degree of freedom.
6. The quality of documenting the model relying on testing (Hosmer & Lemeshow) and which in turn depends on the difference between the observed repetitions and the expected ones where the result was not significant which assures the quality of model documentation.
7. The expansion using the technique of logistic regression considering it an active method in this respect.

References:

[1] N. O. Achia, Anne Wangombe. Thomas and Nancy. Khadioli. *A Logistic Regression Model to Identify Key Determinants of Poverty Using Demographic and Health Survey Data*. European Journal of Social Sciences, Vol. 13(1) (2010), 38-45.

[2] A. E. Ahan, and R. Okafor. *Application of Logistic Regression Model to Graduating (CGPA of University Graduate-University of Lagos)*. Journal of Modern Mathematics and Statistics, 4(2) (2010), 58–62, https://doi.org/10.3923/jmmstat.2010.58.62

[3] M. AL-Nuaimi. and A. Muaayad, *Advanced Statistics in Decision Support focusing on Productive Business Organizations*. Amman. Al-Warraq Institution. (2007), 13-18

[4] J. D. Alsmarat. *The effect of using a problem-solving strategy in teaching the subject of national education in developing decision-making skills for tenth grade students in Jordan*, Al-Quds Open University Journal for Educational and Psychological Research and Studies. 1(2) (2013).

[5] A. K. M. AL-Sulah and H. M. Tabra. *Comprehensive in Business Economics*. Part 1. Juma Aden. (2014), 109.

[6] N. Ayyub. *Theory of Administrative Decisions*. Tarbin Press. Damascus, (1989), 41

[7] Ş. Büyüköztürk, *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum (9. bs)*. Ankara: PegemA Yayıncılık, (2009)

[8] M. Dhaifallah and N. Obaid. *Decision-making and its relationship to the level of ambition among students of the first year of Master*: University of the Martyr Hama Akhdar in the Valley, Algeria, Master Thesis, (2015).

[9] A. Dutta, G. Bandopadhyay and S. Sengupta *Prediction of Stock Performance in the Indian Stock Market Using Logistic Regression*, 7 (1) (2012).

[10] I.S. Fagoyinbo, I.A. Ajibode and Y.O.A. Olaniran. *The Application of Logistic Regression Analysis to the Cummulative Grade Point Average of Graduating Students: A Case Study of Students' of Applied Science*, Federal Polytechnic, Ilaro-Developing Country Studies, 4(23) (2014), 26-30.

[11] D. W. Hosmer, and S. Lemeshow, *Applied Logistic Regression*. Wiley, N. Y., (1989).

[12] Funny. Kimm. *The relationship between decision Making patticipation in jobs satisfaction among wore an school teachers*, University of Iowa, Korea, (2001).

[13] J.E. King. *Running A Best-Subsets Logistic Regression: An Alternative to Stepwise Methods*, Educational and Psychological Measurement. 63(3) (2003), 392-403, https://doi.org/10.1177/0013164403063003003

[14] D. Kökdemir. *Belirsizlik durumlarında karar verme ve problem çözme.Yayım¬lanmamıs doktora tezi*, Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara, (2003).

[15] Justin. Kruger and David. Dunning. *Unskilled and Unaware of it, How Difficulties in Recording one's own incompetence lead to inflated self-Assessment*. Journal of Personality and social psychology. 77(6) (1999), 1121- 1134, https://doi.org/10.1037/0022-3514.77.6.1121

[16] H. Li, J. Sun and J. Wu. *Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods*. Expert Systems with Applications. 37(8) (2010), 5895- 5904, https://doi.org/10.1016/j.eswa.2010.02.016

[17] J. M. Mansoor. *Level of ambition and its relationship to decision-making among a sample of first year students at Damascus University*. Al-Baath University Journal. 39(7) (2017).

[18] W. Osbome and Jason. *Logits and tigers and bears, oh my! A brief look at the simple math logistic regression and how it can improve dissemination of results*. Practical Assesment, Research & Evaluation,17(11) (2012), 1-10.

[19] K. L. Ribanks. *The Relationship between mental ability and personal traits and the level of ambition*. psychology abstract. 88 (3) (2004), 626-628.

[20] D. Sahn and D. Stifel. *Exploring Alternative measure of welfare in Absence of Expenditure Data*. Review of income and wealth, 49(4) (2003): 463-489, https://doi.org/10.1111/j.0034-6586.2003.00100.x

[21] J. Todman and P. Dugard. *Approaching multivariate analysis: An introduction for psychology*. New York: Taylor & Francis Group, (2007).

Appendix (1)
Questionnaire Form

Dear Student

We would like to inform you that the purpose of this questionnaire is a scientific research. Consequently, we would like you to answer all the questions with yes or no. we made sure that all the questions are clear and simple and only need little time. Your answers will be used to obtain the correct scientific results.

Thanks in advance

Is your decision influenced by the environment?	Yes ☐	No ☐
Making a decision does not take a long time from me	Yes ☐	No ☐
I usually have ready alternative decisions	Yes ☐	No ☐
I rely on reasonable justifications in making decisions	Yes ☐	No ☐
I discuss the decision to be made unanimously	Yes ☐	No ☐
I change my decision after making it	Yes ☐	No ☐

I believe that the personal effort diminishes the obstacles no matter how huge they are
Yes   No

life seems to me hopeless	Yes ☐	No ☐
I consider the future very optimistic	Yes ☐	No ☐
I hugely depend on others in solving my problems	Yes ☐	No ☐
I am much influenced by the criticism of others	Yes ☐	No ☐

The researchers