MULTINOMIAL LOGISTIC REGRESSION: USAGE AND APPLICATION IN RISK ANALYSIS

Anass BAYAGA
School of Initial Teacher Education (SITE), Faculty of Education,
University of Fort Hare, South Africa

E-mail: abayaga@ufh.ac.za

Abstract: *The objective of the article was to explore the usage of multinomial logistic regression (MLR) in risk analysis. In this regard, performing MLR on risk analysis data corrected for the non-linear nature of binary response and did address the violation of equal variance and normality assumptions. Additionally, use of maximum likelihood (-2log) estimation provided a means of working with binary response data. The relationship of independent and dependent variables was also addressed.*
The data used included a cohort of hundred risk analyst of a historically black South African University. In this analysis, the findings revealed that the probability of the model chi-square (17.142) was 0.005, less than the level of significance of 0.05 (i.e. p<0.05). Suggesting that there was a statistically significant relationship between the independent variable-risk planning (Rp) and the dependent variable-control mechanism (control mecs) (p<0.05). Also, there was a statistically significant relationship between key risks assigned (KSA) and time spent on risk mitigation. For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought institution spend too little time on Rp decreased by 74.7%. Moreover, the findings revealed that survey respondents who had less confidence in control mecs were less likely to be in the group of survey respondents who thought institution spent about the right amount of time on risk planning.

Key words: *Binary variable; Log odds ratio; Logistic regression model; (log) Likelihood ratio statistic; -2 Log Q; Wald statistic; Model fit; -2 Log L. Quantitative risk analysis*

1. Context of study

Modeling of risk processes such as risk awareness, risk identification, monitoring and reporting, planning and mitigation etc is among rather difficult subjects tackled by risk analyst especially in applying multinomial logistic regression in dynamic (social) setting. Invariably though, social science research (Yu, Lai & Wang, 2008; Fan & Xiao, 2006) problems somewhat call for analysis and prediction of a dichotomous[1] outcomes. Traditionally, such research outcomes were addressed by either ordinary least squares (OLS) regression or linear discriminant function analysis (Hosmer & Lemeshow, 2000). However, both techniques, as a result of their nature, depend on strict statistical assumptions, thus, normality of independent variables, linearity of relationships, multicollinearity among independent variables, equal dispersion matrices for discriminant analysis (Tabachnick et al., 2001). These assumptions which are not easily observed in a dynamic setting are part of multiple[2] regression.

Introduction of multinomial logistic regression was an alternative regression analysis to cater for conditions that do not necessarily obey the assumptions listed above with the exception of multicollinearity (Hosmer & Lemeshow, 2000). In the last decade, the technique, like other univariate and multivariate data analysis methods, started to find a prominent place in the medicine, engineering and the manufacturing industries. This development led researchers in risk analysis to build more accurate and useful statistical models by applying it in risk analysis (Liebenberg & Hoyt, 2003; Hosmer & Lemeshow, 2000). Meanwhile, despite being in use in general statistical analysis for many years, it has received rather little attention in the risk analysis literature compared to other regression applications regarding modelling of explanatory and response variable (Liebenberg & Hoyt, 2003;Hosmer & Lemeshow, 2000).

Recent studies (Crane, Gibbons, Jolley, Van Belle, 2007; Hedeker, 2003; Menard, 2002; Tabachnick, Fidell & Osterlind, 2001; Harrell, 2001; Hosmer & Lemeshow, 2000) have noted that modelling relationship between explanatory (predictor) and response variables is a fundamental activity encountered in risk analysis. The accounts of these studies suggest that simple linear regression is often used to investigate the relationship between a single predictor variable and a single response (dependent) variable. But, when there are several explanatory variables though, multinomial logistic regression is used.

However, often the response (dependent variable) as some of the authors argued (Menard, 2002; Tabachnick et al., 2001; Harrell, 2001; Hosmer & Lemeshow, 2000) is not a numerical value. Instead, the response is simply a designation of one of two possible outcomes (a binary response); example, alive or dead, success or failure, yes or no. Data involving relationship between explanatory variables and binary responses proliferate in just about every discipline from engineering to the natural sciences, to medicine etc. Invariably though, what remains a matter of concern for many practitioners and theorists of risk analysis in University is the questions of how to model relationship between explanatory variables and a binary response variable (Liebenberg & Hoyt, 2003; Tabachnick et al., 2001; Harrell, 2001; Hosmer & Lemeshow, 2000).

Scholars (Hamilton, 2003; Hendrickx, 2000; McCullagh, 1980) have argued that the difficulty for both practitioners and theorists in modeling of risk processes steams from the social setting within which risk parameters are applied. Additionally, the authors suggested that little scholarly literature has delved into application of multinomial logistic regression in analysis of risks parameters particularly in a University context.

Whiles, there remain little studies being conducted in risk analysis with regards to MLR in the context of a University, many recent studies (Van Gelderen, Thurik & Bosma, 2006; Jalilvand & Switzer, 2005; McNeil, Frey & Embrechts, 2005; Mishra & El-Osta, 2002) have encouraged its usage due to its relevance to the field of risk analysis. Following the above, the paper explored the application of multinomial logistic regression via University-wide risk analysis. It does this by using concepts from simple and multiple linear regressions which are carried over to MLR. Additionally, ideas of maximum likelihood estimation are central to the modelling of the MLR data.

1.1. Multinomial Logistic Regression

The multinomial (polytomous) logistic regression model is a simple extension of the binomial logistic regression model. It is used when the dependent variable has more than two nominal or unordered categories, in which dummy coding[3] of independent variables is quite common. In using multinomial logistic regression in risk analysis, the *dependent* (response) variable is dummy coded into multiple 1/0 variables (cf. sections 3 for details). This means that there is a variable for all categories but one, so if there are M categories, there will be M-1 dummy variables. All but one category has its own dummy variable. Each category's dummy variable has a value of 1 for its category and a 0 for all others. One category, the reference category, does not need its own dummy variable, as it is uniquely identified by all the other variables being 0. With regards to the above, risk analyst using multinomial logistic regression can then estimate a separate binary logistic regression model for each of those dummy variables. The result is M-1 binary logistic regression models. The most significant factor to consider here is that each one tells the effect of the predictors of risk on the probability of success in that category, in comparison to the reference category. Noting though that each model has its own intercept and regression coefficients- the reason being that predictors of risk analysis processes could affect each category differently (cf. sections 2 & 3 for details).

1.2. Why *Multinomial* Logistic Regression instead of other Techniques?

Most of multivariate analysis techniques require the basic assumptions of normality and continuous data, involving independent and/or dependent variables as aforementioned. This necessity manifests itself also in the application of MLR data collection and measurement steps in risk analysis, but to varying degree. Thus, whereas much stronger interval and ratio scales provide a good basis for a more comprehensive multivariate analysis, commonly used risk measurement scales such as five-point likert, ordinal, and nominal scales are usually considered unsuitable for multivariate analysis techniques, due to various assumption as listed above. For this reason, multinomial logistic regression was used where the above assumptions tend to be violated. This is evident in one main way in MLR analysis. Thus, it has alternative data distribution assumptions, suggesting that it generates more appropriate and correct findings in terms of model fit and correctness of the analysis regardless of any assumption (cf. section 3.1 for details).

A multinomial logistic regression model is a form of regression where the outcome variable (risk factor-dependent variable) is binary or dichotomous and the independents are continuous variables, categorical variables, or both. The application of multinomial logistic regression in risk analysis arises when an analyst analyses relationships between a non-metric dependent variable and metric or dichotomous independent variables. It compares multiple groups of risk processes such as risk mitigation, planning, monitoring identification, through a combination of binary logistic regressions (cf. section 3). The comparisons are equivalent to the comparisons for a dummy-coded dependent variable, with the group with the highest numeric score used as the reference group. Additionally, it provides a set of coefficients for each of the two comparisons. Multinomial logistic regression exists to handle the case of dependents with more classes. This is referred to as the multivariate case.

Thus, it is expected that multinomial logistic regression approach would do better when there is evidence of substantial departures from multivariate normality as is the case where there are some dichotomous or zero/one variables or where distributions are highly skewed or heavy-tailed especially in dynamic settings. In MLR however, hypotheses on significance of explanatory variables cannot be tested in quite the same way as say linear regression. Recall that in linear regression, where the response variables are normally distributed, one can use t- or F-test statistics for testing significance of explanatory variables. But in logistic regression, the response variables are Bernoulli distributed[4], meaning that a risk analyst has to use different test statistics, which exact distributions are unknown. In this paper though, the researcher would not go into any technical details about test statistics, but focus on interpreting the results of a MLR analysis (cf. methodology, results and discussion of results). To this effect, two different types of test statistics, the (log) likelihood ratio statistic (often referred to as the -2log or deviance) and the Wald statistic would be used. The model is written somewhat differently in some software (cf. SPSS & SAS) than usual mathematical approach. In some software (cf. SAS), the sign is a plus, suggesting that increases in predictor values leads to an increase of probability in the *lower*-numbered response categories. The converse is true for software such as SPSS with a minus sign between the intercept and all the regression coefficients (cf. section 3 for details).This is a convention ensuring that for positive beta coefficients, increases in predictor values leads to an increase of probability in the higher-numbered response categories. It is recommended risk analyst make sure they understand how the model is set up in any statistical package before interpreting results. In general, the likelihood statistic is superior to the Wald statistic (in the sense that it gives more reliable results), so the paper would mainly concentrate on the likelihood ratio statistic (the reason for considering the Wald statistic too, is that it is computationally easy and is given automatically in the output of most statistical computer packages.

Tabachnick et al. (2001) argued that multinomial logistic regression technique has number of major advantages as a summary to the discussion above: (1) it is more robust to violations of assumptions of multivariate normality and equal variance-covariance matrices across groups; and (2) it is similar to linear regression, but more easily interpretable diagnostic statistics. Further, advantages of the analysis that raise its popularity come from the following assumptions: (3) most importantly, MLR does not assume a linear relationship between the dependent and independent variables; (4) independent variables need not to be interval (5) MLR does not require that the independents be unbounded and lastly (6) normally distributed error terms are not assumed.

Widely use of MLR as a problem solving tool, particularly in the fields of medicine, psychology, mathematical finance and engineering are as a result of the above advantages listed. This listed relevance attracted the present author's attention in the case for University-wide risk analysis. Standing on the advantages, the purpose and extent of the usage of MLR in risk analysis research deserves a well-prepared review and application of the method (cf. sections 2 & 3). Such an application would provide valuable clues for future University-wide risk analyst about the scope of MLR, what would be expected from it, and how it would work in various risk problems. Following the above contestations and underlying advantages of MLR in risk analysis, the intent of this paper are as seen below (cf. 1.3).

1.3. Research Objectives

The aims of the paper are to:

a) Demonstrate the use of multinomial logistic regression in risk analysis

b) Evaluate the use of maximum likelihood methods to perform multinomial logistic regression in risk analysis.

c) Demonstrate how to assess the fit of a multinomial logistic regression model in risk analysis with the presence of independent variables.

d) Determine the significance of explanatory variables.

Following third research aim, the hypothesis posed is:

Ho = there is no difference between model without independent variables (WINV) and the model with independent variables (INV)

Ha = there is a difference between model without independent variables and the model with independent variables.

2. Method

The study was undertaken in a historically black South African University in the greater Eastern Cape Province. The University has a high but relatively heterogonous population density of risk analyst in various committees mandated to risk manage the university. This means that the University is equipped with various (academic, finance, human resource, information systems/infrastructure) directors and or managers who are largely in the knowledge of risk analysis. Historically, the risk/quality unit in University is the overseer of risks and quality checks of the University, which is headed by a director. The population for this study was hundred (100) risk analysts in various committees which were stratified. Information was collected from respondents with the aid of a structured and validated interview schedule, consisting of closed ended questions, based on the objectives/questions of the study. Data analysis was with the aid of both descriptive and inferential analysis (cf. section 3). Independent variables for this study were variables of risk analysis, in this case grouped into two (i) institutional risk planning and (ii) institutional risk mitigation. The dependent variable was characteristics of risk variables, which was measured in terms of six elements, viz (i) institution embedded risk management into its planning and operational processes to a sufficient extent (ii) institutional policy documents deal with risk management issues; iinternal auditors conduct audits as part of statutory regulation (iii) institution has control mechanisms to mitigate risk (iv) responsibility for the oversight of individual key risks are assigned to appropriate managers (v) the institution's overall approach to risk management, as assessed for one-academic year is adequate for its strategic objectives (vi) the issues arising from audits are brought to the attention of the executive management team as appropriate. Since the independent variable was ordinal, the following cautions were considered, thus assumptions. Firstly, MLRA does not make any assumptions of normality, linearity, and homogeneity of variance for the independent variables (cf. sections 1.1 & 1.3). Secondly, because it does not impose these requirements, it is preferred to other class of analysis (e.g discriminant analysis) when the data does not satisfy these assumption

The overall test of relationship among the independent variables and groups defined by the dependent was based on the reduction in the likelihood values for a model which does not contain any independent variables and the model that contains the independent variables (cf. sections 3.1 & 3.2). This difference in likelihood followed a chi-square distribution x^2, and was referred to as the model chi-square (cf. section 3). The significance test for the final model chi-square was the researcher's statistical evidence of the presence of a relationship between dependent and independent variables (cf. section 3).

3. Results: Description of the data

This section describes results of the study. This included the description of the data. Firstly, consideration was given to overall test of relationship, this described the overall test of relationship. Secondly, strength of MLR relationship was tested, this was done to establish the strength of MLR relationship. Thirdly, evaluating for the usefulness of logistic model and relationship between the independent and dependent variables.

3.1. Overall test of Relationship

The first thing in MLR for any risk analyst is to describe the overall test of relationship, in this case a relationship between the dependent and independent variables (cf. section 2-method). The presence of a relationship between the dependent and combination of independent variables is based on the statistical significance of the final model chi-square in the table 1; termed model fitting information. In this analysis, the distribution reveals that the probability of the model chi-square (17.142) was 0.005, less than the level of significance of 0.05 (i.e. $p<0.05$). The null hypothesis that there was no difference between the model without independent variables and the model with independent variables was rejected (cf. section 1.4). As evidenced in table 3.1 this suggested that the existence of a relationship between the independent variables and the dependent variable was supported, hence accepting the alternate (Ha) hypothesis.

Table 3.1. Model fitting information

Model	-2logLikelihood	Chi-Square	df	Sig
Intercept Only	272.024	.	.	.
Final	225.334	17.142	6	.005

3.2. Strength of Multinomial Logistic Regression Relationship

Once the relationship is established, the next important thing to do is to establish the strength of multinomial logistic regression relationship. While, MLR does compute correlation measures to estimate the strength of the relationship (pseudo R square measures, such as Nagelkerke's R^2), these correlation measures do not really tell an analyst much about the accuracy or errors associated with the model. A more useful measure to assess the utility of a multinomial logistic regression model was classification accuracy, which compares predicted group membership based on the logistic model to the actual, known group membership, which is the value for the dependent variable (cf. section 3.2.1). To assess the strength of multinomial logistic regression relationship, however, the evaluation of the usefulness for logistic models was considered (cf. sections 3.2.1. & 3.2.2).

In this case, using Cox & Snell R Square and the Nagelkerke R square value, they provide an indication of the amount of variation in the dependent variable. These are described as pseudo R square. The distribution in table 3.2 below reveals that the values are 0.181 and 0.322 respectively, suggesting that between 18.1% percent and 32.2% percent of the variability is explained by this set of variables used in the model.

Table 3.2. Pseudo R-Square

Step	Cox &Snell R^2	NagelKerke R^2
	.181	.322

3.2.1. Evaluating Usefulness for Logistic Models

The proportional by chance accuracy rate was computed by calculating the proportion of cases for each group based on the number of cases in each group in and then squaring and summing the proportion of cases in each group ($0.371^2 + 0.557^2 + 0.072^2 = 0.453$). The proportional by chance accuracy criteria however was 56.6% (1.25 x 45.3% = 56.6%). This warrants comparing accuracy rates. To characterise the model as useful, the study compared the overall percentage accuracy rate produced as 25% more than the proportional by chance accuracy. The classification accuracy rate was 60.5% (cf. table 3.3classification) which was greater than the proportional by chance accuracy criteria of 56.6%, suggesting that the model was useful.

Table 3.3. Classification

Observed	Predicted			
	1	2	3	% Correct
1	15	47	0	24
2	7	86	0	92.5
3	5	7	0	.0
Overall %	**16.2%**	**83.8%**	**0%**	**60.5**

3.2.2. Relationship of Independent and Dependent Variables

Once the above sections are clarified, it warrants a further scrutiny of the relationship of independent and dependent variables. There are two types of tests for individual independent variables (cf. 1.3 for details). The likelihood ratio test evaluates the overall relationship between an independent variable and dependent variables. While, the Wald test evaluates whether or not the independent variable is statistically significant in differentiating between two groups in each of embedded binary logistic comparisons (cf. table 3.4). Risk analyst need to be cautions though that if an independent variable has an overall relationship to the dependent variable, it does not necessarily suggest statistical significance. In fact, it might or might not be statistically significant in differentiating between pairs of groups defined by the dependent variable.

Table 3.4. Likelihood Ratio Tests[5]

Effect	-2log likelihood of Reduced Model	Chi-Square	df	Sig
Intercept	248.323	2.000	2	.010
i-Emb risk	248.625	2.150	2	.020
ii-Policy doc	260.395	3.423	2	.010
iii-Control mecs	265.195	8.200	2	.010

Following the argument above and referring to table 3.4, there is a statistically significant relationship between the independent variable risk planning (Rp) and the dependent variable ($0.010 < 0.05$). As well, the independent variable Rp is significant in distinguishing both category 1 of the dependent variable from category 3 of the dependent variable. ($0.027 < 0.05$) see table 3.5 for this case.

Table 3.5. Parameters Estimates

Risk planning [a]		B	S.E	Wald	df	Sig.	Exp(B)	95.0% C.I for EXP(B) Lower	Upper
1	Intercept	2.240	1.478	1.000	1	.001	.		.
	Emb risk	.019	.020	.906	1	.000	1.019	4.237	24.000
	Policy D.	.071	.106	.427	1	.011	1.073	8.696	302.304
	Control M	**-1.373**	**.620**	**4.913**	**1**	**.027**	**.253**	**.000**	
2	Intercept	3.639	2.456	2.195	1	.008	.	.245	1.000
	Emb risk	.003	.020	.017	1	.020	1.003	.655	1.200
	Policy D.	.002	.110	2.463	1	.117	1.188	.779	1.000
	Control M	.540	.401	4.392	1	.007	.191	.	.

The reference category is: 3

In addition, the independent variable Rp is significant in distinguishing category 2 (cf. table 3.5) of the dependent variable from category 3 of the dependent variable (0.007 < 0.05). What does this imply?

The above suggest that survey respondents who had less confidence in were less likely to be in the group of survey who thought the institution spent too much time on Rp (DV category 3). For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought the institution spent too little time on Rp decreased by 74.7%. (0.253 – 1.0 = -0.747).

Also, an assessment of table 3.6 revealed that there is a statistically significant relationship between key risks assigned (KSA) and the dependent variable, spending on risk mitigation (RisMit) (cf. table 3.2). Moreover, key risks assigned plays a statistically significant role in differentiating the too little group from the too much (reference) group (0.007 < 0.5). However, key risks assigned does not differentiate the about right group from the too much (reference) group (0.51 > 0.5). Survey respondents who were managers (code 1 for key risks assigned) were less likely to be in the group of survey respondents who thought institution spent too little time on RisMit (DV category 1), rather than the group of survey respondents who thought institution spent too much time on RisMit (DV category 3). Survey respondents who were managers were 88.5% less likely (0.115 – 1.0 = -0.885) to be in the group of survey respondents who thought institution spent too little time on RisMit.

Table 3.6. Parameter Estimates[6]

RisMit		B	S.E	Wald	df	Sig.	Exp(B)	95.0% C.I for EXP(B) Lower	Upper
Too Little	Intercept	8.434	2.233	14.261	1	.000	.	1.214	1.442
	audits kn	-.023	.017	1.756	1	.005	.677		
	APPRM	-.066	.102	.414	1	.000	.631		
	audits E	-.575	.251	5.234	1	.021	.563		
	KSA=1	**-2.167**	**.805**	**7.242**	**1**	**.007**	**.115**	**.119**	**1.910**
	KSA=2	0[b]	.	.	0	.		.143	1.303
About Right	Intercept	4.485	2.255	3.955	1	.0004		.245	1.102
	audits kn	-.001	.018	.003	1	.000	.999	.655	1.268
	APPRM	.011	.104	.011	1	.002	1.011	.779	1.300
	audits E	-.397	.257	2.375	1	.003	.673		
	KSA=1	**-1.606**	**.824**	**3.800**	**1**	**.003**	**.201**		
	KSA=2	0[b]							

4. Discussion of findings

The variables emb risk, policy documents, control mecs were useful predictors for distinguishing between groups on risk planning. These predictors differentiate survey respondents who thought institution spent too little time on Rp from survey respondents who thought institution spend too much time on Rp. It also differentiated survey respondents who thought institution spent about the right amount of time on Rp from survey respondents who thought institution spent too much time on Rp.

Among this set of predictors, confidence in control mecs was helpful in distinguishing among the groups defined by responses about spending on Rp. Survey respondents who had less confidence in control mecs were less likely to be in the group of survey respondents who thought institution spent too little time on Rp, rather than the group of survey respondents who thought institution spent too much time on Rp. For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought we spend too little time on Rp decreased by 74.7%. Survey respondents who had less confidence in control mecs were less likely to be in the group of survey respondents who thought institution spent about the right amount of time on Rp. For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought institution spent about the right amount of time on Rp decreased by 88.5%.

Implication MLR for risk analyst: Just as with ordinary least squares regression analyst need some means of determining the significance of the estimates of the model parameters (cf. section 3). The analyses also need a means of assessing the fit, or lack of fit, of the logistic model (Hedeker, 2003; Menard, 2002; Tabachnick et al., 2001). The deviance is twice the log-likelihood ratio statistic. The deviance for a logistic model can be likened to the residual sum of squares in ordinary least squares regression for the linear model. The smaller the deviance the better the fit of the logistic model. A small value for the deviance is an indication that there is a significant fit for the logistic model and some other model may be more appropriate. Asymptotically, the deviance has a $\chi 2$ distribution, therefore, to perform tests of hypotheses regarding the fit of the model the deviance is compared to the percentiles of a $\chi 2$ distribution.

Numerical problems: The maximum likelihood method used to calculate multinomial logistic regression is an iterative fitting process that attempts to cycle through repetitions to find an answer. Sometimes, the method will break down and not be able to converge or find an answer. Sometimes the method will produce wildly improbable results, reporting that a one-unit change in an independent variable increases the odds of the modeled event by hundreds of thousands or millions. These implausible results can be produced by multicollinearity, categories of predictors having no cases or zero cells, and complete separation whereby the two groups are perfectly separated by the scores on one or more independent variables.

If an independent variable has an overall relationship to the dependent variable, it might or might not be statistically significant in differentiating between pairs of groups defined by the dependent variable. The interpretation for an independent variable focuses on its ability to distinguish between pairs of groups and the contribution which it makes to change the odds of being in one dependent variable group rather than the other. Analyst should not interpret the significance of an independent variable's role in distinguishing between pairs of groups unless the independent variable also has an overall relationship to the dependent variable in the likelihood ratio test. The interpretation of an independent variable's role in differentiating dependent variable groups is the same as the researchers used in binary logistic regression. The difference in multinomial logistic regression is that analyst could have multiple interpretations for an independent variable in relation to different pairs of groups.

5. Conclusion

With reference to the research objectives, firstly, consideration was given to situations where response variables are binary random variables, taking the values 1 and 0, for 'success' and 'failure', risk analysis processes respectively. The parameters in the model were estimated using the method of maximum likelihood (cf. research objectives a-d). Odds ratios for the response variables were calculated from the parameters of the fitted model. In order to test hypotheses in logistic regression, the study used the likelihood ratio test and the Wald test. In order to evaluating usefulness for logistic models, the benchmark that was used to characterise the MLR model as useful was a 25% improvement over the rate of accuracy achievable by chance alone. This was referred to as by chance accuracy. The estimate of by chance accuracy that we will use is the proportional by chance accuracy rate, computed by summing the squared percentage of cases in each group. The odds of the response variable being success, for given values of the explanatory variables, are the ratio between the probability that the response is a success and the probability that the response is failure, given the values of the explanatory variables (cf. section 3). The odds ratio compares the odds of the response variable being success for two different sets of values of the explanatory variables.

Secondly, the finding also revealed that the probability of the model chi-square (17.142) was 0.005, less than the level of significance of 0.05 (i.e. $p<0.05$). Suggesting a statistically significant relationship between the independent variable risk planning (Rp) and the dependent variable ($0.010 < 0.05$). Also, an assessment of table 3.6 revealed that there is a statistically significant relationship between key risks assigned (KSA) and the dependent variable, spending on RisMit. Survey respondents who were managers (code 1 for key risks assigned) were less likely to be in the group of survey respondents who thought we spend too little time on RisMit, rather than the group of survey respondents who thought we spend too much time on RisMit. For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought institution spent too little time on Rp decreased by 74.7%. Survey respondents who had less confidence in control mecs were less likely to be in the group of survey respondents who thought we spend about the right amount of time on Rp. Implying that both risk planning and mitigation are integral component of risk analysis. There play important role in variables such as control mechanism policy formulation and audit. Hence, it is recommended that attention be given to such components of risk analysis in effort to quality assure an institution.

References

1. Crane, P.K. and Gibbons, L.E. and Jolley, L. and Van Belle, G. **Differential item functioning analysis with ordinal logistic regression techniques**. Medical care, 44; 2006, pp. 115-132
2. Crouhy, M., Galai, D. and Mark, R. **The essentials of risk management.** New York:McGraw-Hill, 2006
3. Falavigna, G. **New contents and perspectives in the risk analysis of enterprises,** International Journal of Business Performance Management. 8, 2008a, pp. 136-173
4. Falavigna, G. **A rating model simulation for risk analysis,** International Journal of Business Performance Management, 10, 2008b, pp. 269-29
5. Fan, J.X. and Xiao, J.J. **Cross-cultural differences in risk tolerance: a comparison between Chinese and Americans.** Journal of Personal Finance, 5, 2006, pp. 54-75
6. Hamilton, L.C. **Statistics with Stata,** Duxbury, Publisher Unknown, 2003
7. Harrell, F.E. **Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis** .UK, Springer, 2001
8. Hedeker, D. **A mixed-effects multinomial logistic regression model**. UK, John Wiley & Sons, Ltd, 2003
9. Hendrickx, J. **Special restrictions in multinomial logistic**, Stata Technical Bulletin regression, 56, 2000, pp. 18-26.

10. Hosmer, D.W. and Lemeshow, S. **Applied logistic regression.** US, Wiley-Interscience, 2000
11. Jaafari, A. ***Management of risks, uncertainties and opportunities on projects: time for a fundamental* shift, *International***. Journal of *Project Management*. 19, 2000,pp 89-101
12. Jalilvand, A. and Switzer, J. **Evolution of Risk Management as an Integrated Field of Study: Evidence from International Data**. Place of publication unknown, publisher unknown, 2005
13. Kara, S. and Kayis, B. **Competing on capabilities-an analysis of supply chain flexibility in Australian manufacturing industry,** International Journal of Risk Assessment and Management,7, 2007, pp 79-99
14. Liebenberg, A. and Hoyt, R. **The determinants of enterprise risk management: evidence from the appointment of chief risk officers.** Risk Management and Insurance Review. 6, 2003, 37–52
15. McCullagh, P. **Methodological.** Journal of the Royal Statistical Society. Series B, 42, 1980, 109-142
16. McNeil, A. J., Frey, R. and Embrechts, P. **Quantitative Risk Management: Concepts, Techniques, and Tools**. Princeton: Princeton University Press, 2005
17. Menard, S.W. **Applied logistic regression analysis**. Uk: Sage Pubns, 2002
18. Mishra, A. and El-Osta, H. **Risk Management Through Enterprise Diversification: A Farm-Level Analysis.** CA: AAEA annual meeting, Long Beach, 2002
19. Models, B. L., Costs, T., Ventures, J. and Note, A.M. **Statistical models for strategic management**, US, Springer, 1997
20. Tabachnick, B.G. and Fidell, L.S. and Osterlind, S.J. **Using multivariate statistics.** US, Allyn and Bacon Boston, 2001
21. Van Gelderen, M., Thurik, R. and Bosma, N. **Success and risk factors in the pre-startup phase**.Small Business Economics*,* 26, 2006, pp319-335
22. Yu, L., Lai, K.K. and Wang, S.Y. **An evolutionary programming based knowledge ensemble model for business risk identification**, Soft computing applications in business, 2008, pp 57-72

---

[1] (Whether an outcome would succeed or fail, whether a respondent should be classified as a user or nonuser, whether a respondent is prone to engage in risky attitude or not, male/female, yes/no, user/nonuser, satisfied/unsatisfied, etc.)

[2] Note that logistic regression is a variation of mutilple regression, which does not necessarily obey the assumptions of the latter.

[3] Dummy coding provides one way of using categorical predictor variables in various kinds of estimation models. Dummy coding uses only ones and zeros to convey all of the necessary information on group membership.

[4]Bernoulli distribution, is a discrete probability distribution, which takes value 1 with success probability p and value 0 with failure probability $q = 1 - p$. It is a good model for any random experiment with two possible outcomes, for example, yes/no answer (of a respondent in an opinion poll), died/survived (in a drug trial) etc. A random variable x has a Bernoulli distribution with parameter $0 < p < 1$ if P(A) is the probability of outcome A. The parameter p is often called the "probability of success". For example, a single toss of a coin has a Bernoulli distribution with p=0.5 (where 0 = "head" and 1 = "tail").

[5] (i) institution embedden risk management into its planning and operational processes to a sufficient extent (emb risk) (ii) institutional policy documents deal with risk management issues; internal auditors conduct audits as part of statutory regulation (policy documents) (iii) institution has control mechanisms to mitigate risk (control mecs)

[6] RisMit: (iv) responsibility for the oversight of individual key risks are assigned to appropriate managers (key risks assigned-KSA) (v) the institution's overall approach to risk management, as assessed for one-academic year is adequate for its strategic objectives; (approach to risk management strategic objectives-APPRMSOBJ) (vi) the issues arising from audits are brought to the attention of the executive management team as appropriate (audits know to EMT)