International Journal Of

Recent Scientific Research

ISSN: 0976-3031

Volume: 7(11) November -2015

APPLICATION OF DUMMY VARIABLES IN MULTIPLE REGRESSION ANALYSIS

Abdullahi U. Usman, Hassan S. Abdulkadir and Kabiru Tukur

THE OFFICIAL PUBLICATION OF
INTERNATIONAL JOURNAL OF RECENT SCIENTIFIC RESEARCH (IJRSR)
http://www.recentscientific.com/ recentscientific@gmail.com

ISSN: 0976-3031

Available Online at http://www.recentscientific.com

International Journal of Recent Scientific Research
Vol. 6, Issue, 11, pp. 7440-7442, November, 2015

International Journal of Recent Scientific Research

RESEARCH ARTICLE

APPLICATION OF DUMMY VARIABLES IN MULTIPLE REGRESSION ANALYSIS

Abdullahi U. Usman, Hassan S. Abdulkadir and Kabiru Tukur

Jodhpur National University, Jodhpur, Rajasthan, India.

ARTICLE INFO

Article History:

Received 05th August, 2015
Received in revised form 08th September, 2015
Accepted 10th October, 2015
Published online 28st November, 2015

Key words

CGPA, ANOVA, fisher's test statistic, dummy variable, multiple regression

ABSTRACT

The target of any graduating student is to get a good and affordable placement after graduating from university. The purpose of this paper is to review the chance of getting successful placement. The paper tries to study the effect of three independent variables against a dependent variable i.e. cumulative grade point average (CGPA), types of job and engineering discipline against salary. The types of job under study are marketing, technical marketing, design and logistics. Electronics and communication, computer science and engineering, mechanical and electrical engineering are engineering discipline under consideration. The concepts of multiple regression analysis using Analysis of variance (ANOVA) are used for the placement's review opportunities in relation to salary structure.

INTRODUCTION

The concept of experimental design was broadly explained (montgomery, 2014). several researchers uses the concept of multiple regression analysis and ANOVA in their statistical research analysis. Multiple regression analysis are frequently used in different aspects of life. (oberkirchner *et al*, 2010) uses multiple regression to analyses and develop a model for the effect of essential material and process parameters to weight and moisture content of impregnated papers. (Bajpai, 2013) analyse university model using multiple regression and ANCOVA and found very essential. (Syla, 2013) study the significance of active-employment programs on employment levels using multiple regression. (Everarda *et al*, 2005) uses multiple regression result to study the importance of engaging student to graphical user interface in teaching statistical courses. (Oswald, 2012) shows how viewing multiple regression results through multiple lenses can give a better assessment to the researchers. (Kelley *et al*, 2003) shows that in multiple regression obtaining accurate parameter contributes more than having statistical significancy.

(Pazzani *et al*, 1981) shows how independent sign regression generate linear model that are almost accurate as multiple regression. (Ludlow, 2014) study suppressor variables and suppression effects in building regression model. (Moya-laraño *et al*, 2008) encourages ecological researchers to use partial regression in their studies. (Breheny *et al*, 2013) uses visreg package which is useful tool in visualizing the relationship between an explanatory variables that is estimated. Visreg construct convenient support in regression model.

Background

Jodhpur National University is a private university created under state government private university Act. The university is situated in the western part of Rajasthan, India. Thousands of students seek admissions in various disciplines every year. However, majority of their students seek employment opportunities within India as the unity in diversity atmosphere and abroad. The paper attempts to study the effects of CGPA, types of job and engineering discipline on the salary structure. The study was made in Jodhpur National University for the students offering engineering programs. Marketing, technical marketing, design and logistics are the four major trends as observed in India. Electronics and communication, computer science and engineering, mechanical and electrical engineering are the major engineering discipline chosen by the students. These jobs are offered mostly to the student with engineering background.

Corresponding author:* **Abdullahi U. Usman
Jodhpur National University, Jodhpur, Rajasthan, India.

Multiple Regression

The statistical techniques of extending linear regression so as to consider two or more independent variables are known as multiple regression analysis. Multiple linear regression takes the following form

$$Y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon_i \quad (1)$$

Where β_o is the intercepts, $\beta_1,\beta_2,\cdots\beta_n$ are regression coefficients

Parameter Estimation

The method of least square is typically used to estimate the regression coefficients in a multiple linear model. The method of least square chooses the β's in the equation (1) so that the sum of squares of errors ϵ_i is minimized. To solve the equations we may use matrix.

$$\boldsymbol{Y = X\beta + \varepsilon}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_o \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

,and $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

in general, y is an $(n \times 1)$ vector of the observations X is an $(n \times p)$ matrix of the levels of the independent variables,β is a $(p \times 1)$ vector of the regression coefficients, and ϵ is an $(n \times 1)$ vector of random errors.

Dummy Variables

These variables usually indicate the presence or absence of the "quality "or an attribute, such as male or female. They are essentially nominal scale variables. One way we could quantify such attributes is by contributing artificial variables that takes on values 1 or 0. 1 indicating the presence of that attribute and 0 indicating the absence of that attribute variables that assumed such 0 and 1 values are called dummy variables. Such variables are thus essentially a device to classify data into mutually exclusives categories such as presence and absence. Dummy variables can be incorporated in regression models just as easily as quantitative variables.

Analysis Of Variance

It involves classifying and cross-classifying data and then testing if the means of a specified classification differ significantly. The rationale behind ANOVA is that the amount of variation in a set of data can be attributed to chance and specified causes.

ANOVA allows for investigating any number of factors which are hypothesized to influence the dependent variable. The assumption of equal variances also implies that for all practical purposes, the means also comes from same population since any normal population is defined two parameters mean and variance. Thus the variance within samples is only because of random effect whereas the difference between samples is due to specified factor whose effect we are trying to study.

$$F = \frac{\textit{Estimate of population based on between sample variance}}{\textit{Estimate of population based on within sample variance}}$$

Table I Table Of Anova

Source of variation	Sum of squares	Degrees of freedom	Mean square	F-ratio
Treatments	$SS_{Treatments}$	$a-1$	$\frac{SS_{Treatments}}{a-1}$	$\frac{MS_{Treatments}}{MS_E}$
Blocks	SS_{Blocks}	$b-1$	$\frac{SS_{Blocks}}{b-1}$	
Error	SS_E	$(a-1)(b-1)$	$\frac{SS_E}{(a-1)(b-1)}$	
Total	SS_T	$N-1$		

Statistical Analysis

Assumptions

i. All samples are drawn from normally distributed population
ii. All samples are drawn independently of each other
iii. Within each sample, the observations are sampled randomly and independently of each other
iv. Factor effects are additive
v. All populations have a common variance

Objective

The objective here is to determine the impact of CGPA, types of jobs and engineering discipline in the salary structure of graduating student. And to determine whether can be fitted linearly. The hypothesis declared below;

H_o: There is no significant effect of CGPA, types of job and engineering discipline on the salary of graduating students.

H_1: CGPA, types of job and engineering discipline has significant effect on salary of graduating student.

Data Collection

The data is collected from office of placement Jodhpur National University offering engineering programs over a period of four years. The types of jobs and engineering discipline are categorical explanatory variables. CGPA and salary structures are quantitative explanatory variables. Four samples are collected each of size 20. The data was presented

on spreadsheets. It has been analyzed using MINITAB and SPSS by the concept of multiple regression analysis

Analysis

Table 2 Summary observation Taken Into Consideration Was 80. Below Table Summarizes All The Data

Variable	No. of obs	No. of obs with missing	mean	StDev	Minimum	Maximum
SALARY(1000)	80	0	4.234	4.234	11.000	27.000
CGPA(%)	80	0	13.13	13.13	50.00	92.00
Types of job	80	0				
discipline	80	0				

Table 3 Analysis Of Variance

Source	DF	SS	MS	F	P
Regression	7	1294.87	184.98	110.06	0.000
Residual Error	72	121.02	1.68		
Total	79	1415.89			

Source	DF	Seq SS
CGPA(%)	1	0.26
job_design	1	489.29
job_logistics	1	159.65
job_marketing	1	627.81
discipline_CSE	1	3.90
discipline_ECE	1	10.41
discipline_EE	1	3.56

Remark

fisher's F-test is an appropriate test statistic used.Table2 above shows that the probability corresponding to p-value is 0.000.it means that risk in assuming the null hypothesis is wrong is quite low. Therefore, we can conclude that with 95% confidence that the variables CGPA, types of job and discipline do bring a significant amount of information.

Table 4 Table Of Coefficient

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18.412	2.387		7.714	.000
	CGPA	.029	.033	.091	.885	.379
	TYPES OF JOB	-1.666	.434	-.396	-3.837	.000
	CAREER	-.968	.353	-.282	-2.745	.008

a. Dependent Variable: salary

Predictor	Coef	SE Coef	T	P	VIF
Constant	16.665	1.274	13.08	0.000	
CGPA(%)	-0.00216	0.01180	-0.18	0.855	1.127
job_design	-1.8119	0.4771	-3.80	0.000	2.682
job_logistics	0.1841	0.5308	0.35	0.730	2.338
job_marketing	9.7530	0.5835	16.71	0.000	2.066
discipline_CSE	-0.9473	0.8198	-1.16	0.252	7.277
discipline_ECE	0.0314	0.8745	0.04	0.971	7.986
discipline_EE	-1.2510	0.8598	-1.45	0.150	7.016

Remark

from the above table CGPA, job_logistics and discipline_ECE has the least impact in the salary of graduating students.

CONCLUSION

The analysis above it shows that the impact of types of job contributed more on the salary structure in the university model when compared with other variables i.e. CGPA and engineering discipline. Job_marketing contributed more in the types of job when compared with other types of jobs offered in the university

References

1. Bajpai, P. (2013). Multiple Regression Analysis Using ANCOVA in University Model, *3*(5), 336–340. http://doi.org/10.7763/IJAPM.2013.V3.232
2. Breheny, P., & Burchett, W. (2013). Visualization of Regression Models Using visreg, 1–15.
3. D_&_A_Of_Exp.pdf. (n.d.).
4. Everarda, G., & Wei, C. (2005). COMMUNICATION OF MULTIPLE REGRESSION, 1–8.
5. Kelley, K., & Maxwell, S. E. (2003). Sample Size for Multiple Regression : Obtaining Regression Coefficients That Are Accurate, Not Simply Significant, *8*(3), 305–321. http://doi.org/ 10.1037/ 1082 -989X.8.3.305
6. Ludlow, L. (2014). Suppressor Variables : The Difference between " Is " versus " Acting As ," *22*(2), 1–28.
7. Moya-laraño, J., & Corcobado, G. (2008). Plotting partial correlation and regression in ecological studies, 35–46.
8. Oswald, F. L. (2012). Interpreting Multiple Linear Regression : A Guidebook of Variable Importance, *17*(9).
9. Pazzani, M. J., & Bay, S. D. (1981). The Independent Sign Bias : Gaining Insight from Multiple Linear Regression.
10. Quality Control of Impregnated Papers with a Multiple Regression Model. (2010), 5431.
11. Syla, S. (2013). Application of Multiple Linear Regression Analysis of Employment through ALMP, *3*(12), 252–258.
12. Draper, N.R.,& Smith, H (1998). Applied Rergression Analysis. Third Edition, Wiley, New York.
13. Montgomery, C.D. (2013). Design and Analysis of Experiment. Eight Edition, Wiley. New
14. York
15. Beri, G.C., (2005).Business Statistics. Second Edition. McGraw-Hill.

How to cite this article:
Lyrin K Vincent, Lekshmi B S and Basil K Jeemon., Performance Analysis Of Circular Patch Antenna Using Different Hybrid Substrates. *International Journal of Recent Scientific Research Vol. 6, Issue, 11, pp. 7440-7442, November, 2015*

****

ISSN 0976-3031