Good essay on econometrics

Profession, Student

Introduction

In this study, we are interested in the performance of boys and girls in mathematics subjects in high school. In this study, the data was obtained from the international mathematical Olympiad website. In this study the dependent variable is performance of boys and girls in mathematics the independent variable are gender, proficiency, and ethnicity. According to Jeffery M. Wooldridge,(2000) the performance in mathematics is influenced by gender, proficiency of the students in the subject and the ethnicity. According to murphy M. (2000) gender is the greatest factor that influences the performance of mathematics. The research will be Causal / explanatory research type, and inductive methodology approach will be used and hence the results will be highly reliable and empirical in nature. Both of quantitative and qualitative research methods required to be used in this research project in the following two phases

Literature review

Many researchers have devoted on the topic about the performance of mathematics subject in both male and female is high academic level. Payne and Isaacs (2005) advocated that lack of commitment by student in mathematics subject, gender, proficiency in the areas of mathematics and race of students are the contributing factors in performance in mathematics . As a result, it is believed that if all the students are provided with equal opportunity and under same learning environment, both male and female student would perform equally well in mathematics. Besides gender, race and proficiency, the altitude of students toward mathematic subject is

believed to have influence performance. Dintiman and Ward (2003) pointed out that proficiency is one of the three factors that contribute to performance in mathematics. For high performance in mathematics, many researchers have noted that the student with high performance have good proficiency of the mathematic subject even at their early age.

Problem statement

The aim of this study is to examine the relationship between the performance of the mathematics subject with, gender, proficiency in the mathematics subject, and the race of the student. The researcher also attempted to find out the best linear combination of variables explaining the variance for the dependent variable scores in mathematics subject as well as the independent variables; the gender, proficiency, and race. In this study, the performance of the students both male and female in mathematics subject was recorded and their proficiency in the subject traced and recorded. In this study, the dependent variable was the performance of the student in mathematics subjects for the last 10 years. We will also try to construct the linear regression model between the dependent variable, student performance in mathematic subject and the independent variables; the gender of the student, proficiency of the student in the mathematics subject, and the race of the student. The assumption of the regression model is that the data came from a normally distributed population. Thus, we will conduct the test to determine if the data is normally distributed and hence conduct a regression analysis of the dependent and independent variables. Further we will examine if the performance of the girls in dependent on the performance of boys. We decide to use two sample t-tests to determine if

the performance of girls is independent or dependent on the performance of boys.

Study design

Based on the study question, a survey questionnaire were formulated and used to survey and obtain the important data and based on Literature review, number of questions addressing the research problem and the research questions are framed. The survey format contains two sections; the first section is designed with the questionnaire to analyze the variables, while the second section is designed to gather information about the variables of interest in the research questions.

The survey questionnaire developed based on the extensive literature review into 4 factors, mainly the education data files. The education performance files were examined based on the formulated study questions to obtain the information on the variables of interest in the study. The results of the data obtained, was stored in a tabular table with headers as the variable names.

Data description

The data was obtained from the international mathematical Olympiad website. In this study the dependent variable is performance of boys and girls in mathematics the independent variable are gender, proficiency, and ethnicity. According to Jeffery M. Wooldridge,(2000) the performance in mathematics is influenced by gender, proficiency of the students in the subject and the ethnicity. The above secondary data will be cleaned to ensure that the outlier and any wrong information will not be used in the study. This will ensure that we will obtain a conclusion that is justifiable to

any road infrastructure projects.

The survey samples were selected through Non probability, convenient sampling method; this was to ensure the data that was used in the survey was not biased. The total sample size that was accepted to be a justifiable for modeling the data was 15 and had 4 variables that were to be modeled using the mentioned data.

Hypothesis

The following summary highlights the range of quantitative testing that will be applied over each of the 3 primary hypotheses.

- Hypothesis 1: There is no significant difference in the performance of mathematics subject and the gender either male or female.

A comparative differentiation between male and female performance in mathematics subject was conducted. The primary variables in this analysis include mathematic performance relative to gender of the students. The quantification of these datasets reflects two difference formulas, whereby performance will be assessed according to marks attained while gender will be assessed based of male or female

- Hypothesis 2: There is no significant difference in the performance mathematics subject and the proficiency of the student in the subject.

This particular hypothesis differentiates between mathematic performance of student and the proficiency of the student. The quantitative underpinnings for this model will replicate the dual segment analysis (performance and proficiency) referenced in the first hypothesis; however, will eliminate the control for race.

Hypothesis 3: There is no significant relationship between mathematics performance and the race of the student.

Leveraging the quantitative outputs from Hypotheses 1, this particular hypothesis involves a statistical comparison of the relationship between performance of the student in mathematics and the race of the student. Using the statistical foundations of regression analysis and correlation we will be able to test and draw conclusion about the hypothesis 1, 2, and 3. In order to identify statistical correlations associated with Hypotheses 1 and 2, this test will include differentiation between male and female datasets.

Checking of model assumptions

The model assumptions are that the data set is normally distributed.

We first test if the mathematics performance is normally distributed. In this case we conduct a normal probability test

The proficiency

A test of normality was carried out so as to determine if the data for mathematics performance were normally distributed so as to determine the type of analysis to be carried out. If the data was not normally distributed, then special methods could be necessary for its analysis (Gergen, 1997). In order to test for normality, the frequency distribution for percentage cost and time overrun for was plotted for six classification of projects. A further normality test carried out was the One Sample Kolmogorov-Smirnov Test.

A regression analysis was used to investigate whether there exist a relationship between the variables. The research investigated the relationship between underlying variables such as average yearly cost overrun; average yearly time overrun, average yearly contract sum, average

yearly exchange rate, average yearly time overrun, average yearly inflation rate and average cost due to exchange rate fluctuation were investigated so as to be represented as a linear equation. According to Neumaier (2003), such models are very common especially in fields dealing with modelling for optimizing outputs by finding the best combination of all the input factors for such output variables.

The linear form of the mathematical formular used has the generalized simplified form of linear equation is as follows:

$Y = mX + C$

Where;

Y is the dependent (output) variable,

X is the input (independent) variable,

m is the coefficient of the input variable and,

c is the basic constant.

A correlation coefficient is a statistic that is used for testing whether there is any relationship between the variables being tested. It varies between -1 to +1. A correlation of +1 implies that there is a full positive relationship between two variables, hence an increase in one of the variable results to an increase in the value of the other. A correlation of 0 implies that there is totally no relationship between the variables being tested. Hence, the presence of one of the variables does not imply the presence of the other. A correlation of -1 implies that there is a full negative relationship between two variables; hence an increase in one of the variables results in a decrease in the value of the other. Any correlation near -1 implies a strong negative relationship and any near 0 from the negative implies a weak negative

relationship. Any correlation near 1 implies a strong positive relationship and any near 0 from the positive implies a weak positive relationship.

A correlation coefficient is a statistic that is used for testing whether there is any relationship between the variables being tested. It varies between -1 to +1. A correlation of +1 implies that there is a full positive relationship between two variables, hence an increase in one of the variable results to an increase in the value of the other. A correlation of 0 implies that there is totally no relationship between the variables being tested. Hence, the presence of one of the variables does not imply the presence of the other. A correlation of -1 implies that there is a full negative relationship between two variables; hence an increase in one of the variables results in a decrease in the value of the other. Any correlation near -1 implies a strong negative relationship and any near 0 from the negative implies a weak negative relationship. Any correlation near 1 implies a strong positive relationship and any near 0 from the positive implies a weak positive relationship.

ANOVA applied to normally (parametric) and non-normally distributed data

Analysis of Variance is not an individual test, but is actually a group of statistical models that are used together to divide the variance of a variable into its individual components, arising out of different influences on the variable. ANOVA is widely used in research for hypothesis testing with parametric data.

While the Student’s t-test is more suitable for parametric models, for non-parametric data the procedure usually adopted is to replace the values of

observations by their ranks, and then use the ranks instead of the observations.

Results of analysis

Test of the independence or dependence.

Two-Sample T-Test and CI: mathematic performance, Gender

Two-sample T for mathematic performance

Gender N Mean StDev SE Mean

1 82 517. 0 78. 2 8. 6

2 79 517. 9 74. 7 8. 4

Difference = mu (1) - mu (2)

Estimate for difference: -0. 9

95% CI for difference: (-24. 7, 22. 9)

T-Test of difference = 0 (vs not =): T-Value = -0. 07 P-Value = 0. 942 DF = 159

Both use Pooled StDev = 76. 4819

Regression analysis

Regression Analysis: mathematic perfo versus Gender, race, proficiency

The regression equation is

mathematic performance = 321 + 1. 76 Gender - 1. 81 race + 84. 0 proficiency

Predictor Coef SE Coef T P

Constant 321. 43 25. 59 12. 56 0. 000

Gender 1. 764 7. 167 0. 25 0. 806

race -1. 813 5. 064 -0. 36 0. 721

proficiency 83. 962 4. 970 16. 90 0. 000

S = 45. 4523 R-Sq = 65. 1% R-Sq(adj) = 64. 5%

Analysis of Variance

Source DF SS MS F P

Regression 3 605749 201916 97. 74 0. 000

Residual Error 157 324349 2066

Total 160 930098

Source DF Seq SS

Gender 1 31

race 1 16007

proficiency 1 589712

Analysis of the results

- Hypothesis 1: There is no significant difference in the performance of mathematics subject and the gender either male or female.

- Hypothesis 2: There is no significant difference in the performance mathematics subject and the proficiency of the student in the subject.

Hypothesis 3: There is no significant relationship between mathematics performance and the race of the student.

Conclusion

In above analysis of regression, it is clear indicated that there is a relationship between the performance of mathematics among student with proficiency of student in mathematic subject but there is no relationship

between the performance of student in math and the gender of the students or their race. From the two-sample t-test, the mean performance of girls in mathematics is 517. 9 while that of boys is 517. 0. The difference in the mean performance in mathematics between boys and girls is 0. 9. The p-value is 0. 942 which is greater than 0. 05 level of confidence level. We thus conclude that the performance of girls is independent of the performance of the boys in mathematics subjects.

Works Cited

Campbell, G, Stonehouse, G & Houston, B 2002, " Business Strategy," Butterworth-
Heinemann, Oxford, UK.

Farnham, D 1999, " Managing in a business context," CIPD Publishing, London, UK.

Hussey, DE 1998 " Strategic management: from theory to implementation," Butterworth-Heinemann, Oxford, UK.

Sadler, P & Craig, JC 2003. Strategic management. Kogan Page Publishers, London

Saloner, G et al. 2006. Strategic management. John Wiley, London, UK.

Sadler, P & Craig, JC 2011. TJ Hammond Transport Limited. Viewed on April 21,
2011