Original Article

International Journal of Educational Research and Technology
P-ISSN 0976-4089; E-ISSN 2277-1557
IJERT: Volume 4 [3]September 2013: 31 – 36

ISO 9001: 2008 Certified Organization
Website: www.soeagra.com/ijert/ijert.htm

Study on Interdependency of Secured Marks in Engineering and 10^{th} - 12^{th}

R. P. Rajoria[1], Shuchi Dave[2], Meenakshi Naval[3]
Poornima College of Engineering Jaipur (Raj.), India[1,2&3]
Email: rprajoria@poornima.org[1], shuchi.dave@poornima.org[2], meenakshinawal@poornima.org [3]

ABSTRACT
This paper is based on relationship between two continuous variable of average marks of class 10^{th} - 12^{th} and average marks of B-Tech (I to VIII Sem). This is a case study of B-Tech Students of batch 2011-12 who were admitted in 2008-09 at Poornima College of Engineering, Jaipur (Rajasthan) .To check this relation statistically significant Pearson Correlation has been introduced as a measure of Strength of a relationship between two variables. The data have been taken of different branches (Electronics Engineering, Electrical Engineering, Computer Science, Information Technology, and Mechanical Engineering) from the population to check the significance test. A null hypothesis is set and t-test is done to reject the hypothesis. On the basis of result analysis it can be said that dependency between two variables average marks of class 10^{th} - 12^{th} (PCM – Physics, Chemistry & Mathematics) and average marks of B-Tech (I to VIII Sem) Students is significant.
Keywords:- *Data of Electronics Engineering, Electrical Engineering, Computer Science, Information Technology, Mechanical Engineering, Degree of freedom, t-test, Correlation Coefficient, Null Hypothesis, p-level, t-table.*

Received 18/07/2013 Accepted 20/08/2013

INTRODUCTION

Statistics is used only to describe the central tendency and the variability among the scores in a single data set, or between levels of an independent variable (as in the independent group's t-Test).It is also used to assist in describing the relationship between different dependent Variables. Indeed, this is the one of the major uses of statistics: to determine whether one variable is statistically associated to another variable. A correlation coefficient shows a statistical relationship between two variables. Simply, a correlation exists when changes in one variable are statistically associated with systematic changes in another variable. Hence, a correlation is a type of bivariate relationship [3]. The strength of the relationship is indicated by the correlation coefficient: r but is actually measured by the coefficient of determination: r^2. The significance of the relationship is expressed in probability levels: p (e.g., significant at p =.05). This tells how unlikely a given correlation coefficient, r, will occur given no relationship in the population. The smaller the p-level, the more significant the relationship the larger the correlation, the stronger the relationship. The difference between what is observed and what is expected given the assumption of the population large enough to be significant -- to reject the assumption. The greater the difference -- the more the sample statistic deviates from the population parameter -- the more significant it is (http://janda.org).

Firstly a null hypothesis is set that there is no relationship between average marks of class 10^{th} - 12^{th} and average marks of B-Tech (I to VIII Sem) Students. The value of correlation coefficient 'r' & coefficient of determination 'r^2' was calculated. Then some scattered diagrams were plotted between X, Y & r^2. After that a straight line which indicates linear relation between X & Y is drawn. But in this case, some outlier points have been observed. So to check dependency of data, a t-test is done on all the samples. On the basis of t-test, the null-hypothesis is rejected. Which indicates type – I error in the data i.e. rejecting the null hypothesis when it is true (Gupta and Kapoor, 1999). In all the samples absolute value of 't' for one – tail test with 95% level of confidence, degree of freedom n-2, where n=Sample size of different branches has been observed . Then the errors has been calculated in the samples data which is probability of type- I error (α) =0.0005 which is less than the level of significance 0.05.

FORMULATION OF THE PROBLEM

Data has been collected for different X & Y of sample size n=32 for Electrical, n=126 for Electronics, n=64 for Information & Technology, n=125 for Computer Science, n=61 for Mechanical engineering student. The average marks of 10th & 12th and B-Tech (I-VIII Sem) are shown in the table below

Table-1

Branch	No. of Students	Avg. marks of 10th & 12th	Avg. marks of B-Tech(I-VIII sem)
EE	32	69.847	64.183
EC	126	76.400	65.547
CS	125	77.084	68.351
ME	61	71.490	61.337
IT	64	74.118	65.389

Table 1: Average marks of 10-12 and B.Tech (I-VIII sem) of different branches

NULL HYPOTHESIS

The null hypothesis which is denoted by Ho. That means there is no relationship between average marks of students in 10th & 12th and average marks of students in B. Tech (I-VIII sem) or we can say that no relationship between X and Y in the population: ρ = 0.0. Under this common null hypothesis in correlation analysis: *r* = 0.0.
Symbolically we can say that

$$H_o: \mu = \mu_o$$

Where μ = mean marks of the students from all branches or population, μ_0 = mean marks of the students from specified branches. Null Hypothesis is the hypothesis of no difference. This is tested for possible rejection under the assumption that it is true [2].

PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENT (*r*)

The statistic most commonly used to measures the relationship between quantitative variables is the Pearson Produce-Moment Correlation Coefficient or more the Pearson correlation (r). The Pearson correlation is used to measure the strength and the direction of a linear relationship between two quantitative variables [3].
Formula to calculate correlation coefficient is given by equation (1)

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}} \quad (1)$$

Using equation (1), for Electrical Engineering when n=32, $\sum xy$=144985.3, $\sum x$=2235.1, $\sum y$=2053.86, $\sum x^2$=158777.0125, $\sum y^2$=133909.7 then r=.6490051 was calculated .
Similarly the value of r for other branches have been calculated and given in the table (2) below

Branch	r=coefficient of correlation
EE	0.6490051
EC	0.4404543
CS	0.5700877
ME	0.5119539
IT	0.54112754

Table 2: Coefficient of Correlation of different branches

MATHEMATICAL EQUATIONS

In order to find out the relationship between the variables equations of Straight lines from five different Sample size has been calculated by using the formulae

$$y = ax + b \quad (2)$$

$$\sum y = a\sum x + nb \quad (3)$$

$$\sum xy = a\sum x^2 + b\sum x \quad (4)$$

Where a = slope of the straight line, b = Intercept of the line on y axis, n = Sample size
For Electrical Engineering

$$y = 0.5745971x + 24.05289 \quad (5)$$

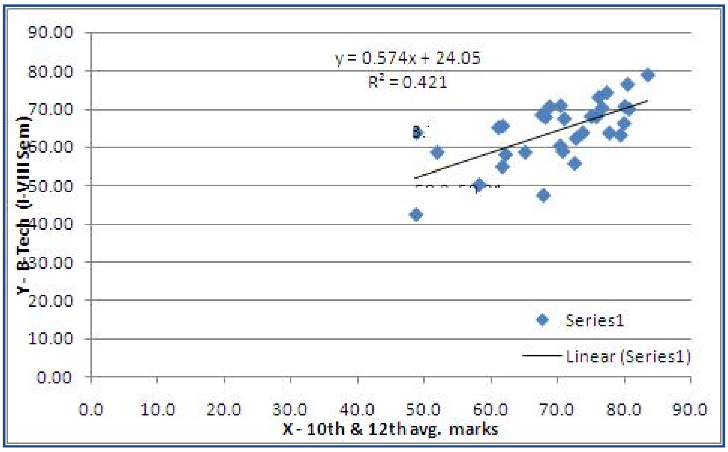

Fig (1)

For Electronics Engineering

$$y = 0.5895284x + 20.5072 \quad (6)$$

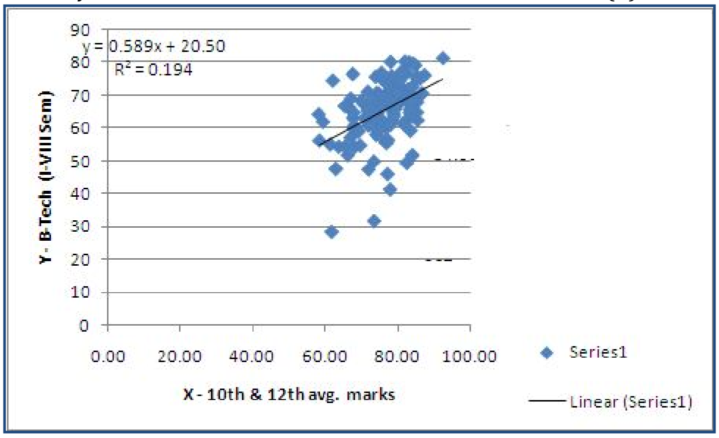

Fig (2)

For Computer Science Engineering

$$y = 0.6087345x + 21.42778 \quad (7)$$

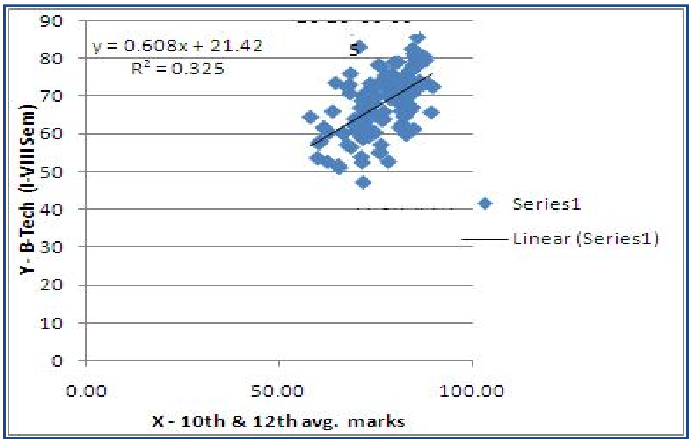

Fig (3)

For Mechanical Engineering

$$y = 0.5547793x + 21.67635 \quad (8)$$

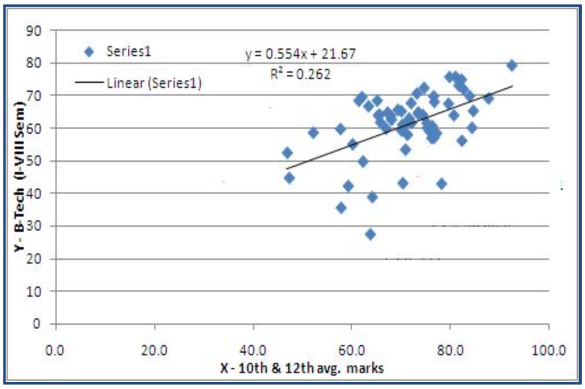

Fig (4)

For Information Technology Engineering

$$y = 0.6324496x + 18.51308 \quad (9)$$

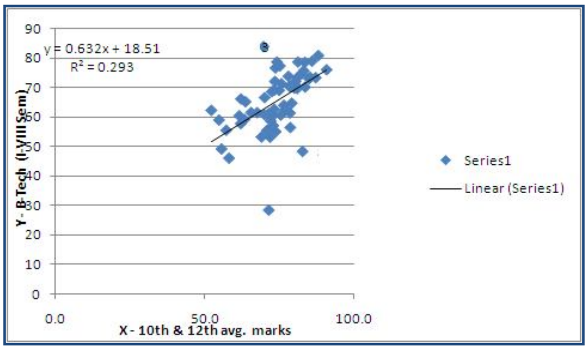

Fig (5)

In all above graphs a straight line has been fitted between the data X, Y and r^2
Where X is the avg. marks of 10th and 12th on the x axis, Y is the avg. marks of B-Tech (I to VIII Sem) on the y axis & r^2 is Coefficient of determination. The line shows relation between x & y.
Scatter graph plotted as above show a relationship between two variables is positive linear. A linear relationship means that the relationship between variables appears to occur in a straight line; that is, as values of one variable increase the values of a second variable increase; or the values of one variable decrease the values of another variable decrease. A positive linear relationship is observed when the values of both variables have a trend that occurs in the same direction. That is, as the values of one variable increase the values of the second variable also increase [3].

T-TEST TO TEST SIGNIFICANCE

In order to analyze further, test of significance is conducted. The simplest formula for computing the appropriate t value to test significance of a correlation coefficient is as follows:

$$t = r\sqrt{\frac{n-2}{1-r^2}} \quad (10)$$

Where (n-2) = degree of freedom, r = coefficient of correlation.
Calculated t value for all above mentioned samples by using equation (10) is given in table (3)

Table-3

Branch	df=degree of freedom	r^2	t value
EE	30	0.421	4.671641
EC	124	0.194	5.463162
CS	123	0.325	7.695597
ME	59	0.262	4.576658
IT	62	0.293	5.067407

Significant value of t distribution

α(1 tail)	0.05	0.025	0.01	0.005	0.025	0.001	0.0005
df							
30	1.69	2.004	2.45	2.75	3.02	3.38	3.64
59	1.67	2.00	2.39	2.66	2.91	3.23	3.46
62	1.66	1.99	2.38	2.65	2.91	3.22	3.45
123	1.65	1.97	2.35	2.61	2.85	3.15	3.37
124	1.65	1.97	2.35	2.61	2.85	3.15	3.37

Now the absolute value of t (from table 2.5) for one tail test with 95% level of confidence, degree of freedom n-2 is observed.

Table-4

Branch	df=degree of freedom	t value
EE	30	1.69
EC	124	1.6572
CS	123	1.6573
ME	59	1.6711
IT	62	1.6698

As a comparison of table (3) and table (4) ,we get that there is a significant difference between calculated value of t and absolute value. On the basis of this result we can say that the null hypothesis is rejected and it shows that there is a dependency between average marks of students in 10th & 12th class and average marks of students in B. Tech (I-VIII Sem). This relationship is strong.

T-graph of different sample

For EE

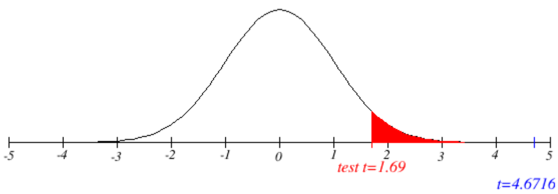

Fig (6)

For EC

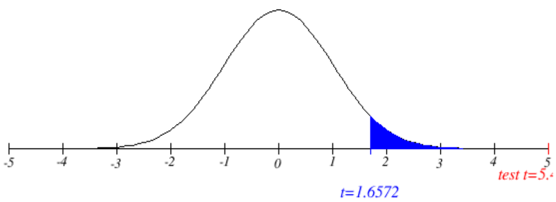

Fig (7)

Similarly different graph for other branches can also be drawn.

ERRORS

In the data of five samples sets we have type I error. It is represented by Greek letter alpha (α). In choosing a level of probability for a test, how much risk is there in committing a Type I error—rejecting the null hypothesis when it is, in fact, true. For this reason, the area in the region of rejection is sometimes called the alpha level because it represents the likelihood of committing a Type I error [Gupta and Kapoor, 1999]. This is the error of rejecting Ho when Ho is true that is α

α= probability of type – I error

=probability of rejecting Ho when Ho is true.

=0.0005 (using critical t table)

P-value > 0.0005= α for (1 – tail test) for all t values tabulated in table-3 with different degree of freedom.
As the p-value (α=.0005) is less than the level of significance (α* =.05) i.e. α < α* then we can say that relation is statistically significant. And we can also say that there is a dependency between average marks of students in 10^{th} & 12^{th} and average marks of students in B. Tech. this relationship is strong. So, as when average marks of 10^{th} & 12^{th} increases then average marks of B. Tech also increases.

CONCLUSION

In the Summary, the coefficient of correlation shows that the relationship between average marks of 10^{th} & 12^{th} & B-Tech (I-VIII Sem) is significantly strong. So we can conclude that Null Hypothesis is rejected & there is a dependency between average marks secured in Engineering Examination i.e. B.Tech (I-VIII Sem.) & 10^{th} - 12^{th} (PCM).

REFERENCES

1. http://janda.org/c10/lectures/topics06/L24-significanceR.htm.
2. Fundamental of Mathematical Statistics, by S.C Gupta & V.K. Kapoor, Sulatnchand & Sons, New-Delhi, 10th Revised Edition 1999.
3. humanattention.org/textbook/chapterpdfs/ch%2016%20correlation.pdf
4. *www.unc.edu/~jamisonm/fivesteps.htm*

Citation of This Article
R. P. Rajoria, Shuchi Dave, Meenakshi Naval. Study on Interdependency of Secured Marks in Engineering and 10^{th} - 12^{th} . Int. J. Educat. Res. Technol. Vol 4 [3] September 2013 : 31-36