International Journal of Information, Engineering & Technology

Volume 11, Issue 3, PP 28-51, ISSN: 2360-9194, December, 2020
Double Blind Peer Reviewed International Research Journal
garcjournalssubmit@gmail.com
Journal Series: Global Academic Research Consortium (garc)

Application of Correlation and Regressin Analysis on Student Performance in West African Examination Council (WAEC)

Modu Mohammed[1], Grema Modu Bako , Mjigumtum Joseph Madubu and Sule A. Askira

Department of Mathematics and Statistics, Ramat Polytchnic Maiduguri, Borno State, Nigeria

Abstract: *The main aim of this research is to assess an* Applications of Correlation and *Regression Analysis on Student Performance in West African Examination Council (WAEC) Base on the finding it is revealed that the result of the analysis of correlation coefficient in table 4.1,4.3, and 4.5 shows a negative and weak correlation and table 4.2 and 4.4 show positive and weak correlation and the Test at 5% and 1% level of significance of correlation coefficient of both tables shows there is no significant difference in performance between maths and physics, and the covariance method of estimation of regression parameters for β1 shows there is a significant difference in student performance and β0 shows there is no significant difference.it also reveals that the regression equation, In table 4.1, 4.3 and 4.5its indicates that the graph is moving downward and it is also shown that there is a slight decrease in student performance, and also that the regression equation in table 4.2 and 4.4 its indicates that the graph is moving upward and it is also shown that there is a slight increase in student performance. It's recommended that the government would employ qualified teachers; Students would dedicate themselves to hard work and dedication to learning and Provision of more laboratory equipment to enhanced student and learning process.*

Key words: *correlation, Regression and performance .*

Introduction

Statistical knowledge is an important skill in today's' technological societies. Statistics is widely practiced in science, economics, engineering, social sciences, health, sports, and many others Razak F. A *etal* (2017).

A Binary Logistic Regression model is used to determine the probability of the student performing rate founded on the stated factors. These factors play a significant role at 5% level of significance. Thus, a Logistic Regression model to forecast the Academic Performance will be an effective tool for the decision-making method Surendheran R (2017).

Galadanci (2017) Presented that there is a statistically significant and strong nonnegative relationship between students' theoretical and practical scores as a result of which the null hypotheses were rejected.

Ahono T. A *et al* (2018) investigate that the research reveals that there was a statistically significant, weak, and positive correlation (r=.142, n=396, p=.005) between self-efficacy of expectation and Mathematics achievement. The findings showed that self-efficacy of expectation predicted the achievement in mathematics among secondary school students

Izaak (2015) Investigate that there is a positive relationship between concentration in Physics and knowledge of Mathematics basic ideas with students' ability to solve physics problems.

A weighted ordinary least square hierarchical multiple regression method was employed to the achieved quantity and quality of computer usage, significant predictors of achievement were established, Jehanzeb R. C (2013).

A significant relationship between self-concept and the respondents' academic performance in Mathematics was found by Merson P eta (2020)

Pearson's correlation coefficients of four independent are correlated with student's academic performance although two are not. Though, using the regression analysis four variables is significant which include: Time appropriateness, people-friend connection, nature of Usage, and health addiction while Time duration and security/privacy problems are not significant, Sandra (2016).

The results obtained by the analyses conducted revealed that there were significant relationships between the students' academic achievement and student engagement as well as between their academic achievement and especially the dimensions of cognitive engagement, behavioral engagement, and sense of belonging, Selim G, Y. Y (2014).

A relationship between two quantitative variables usually involves a discussion of correlation and regression. When data is expressed in a standardized form, correlation and regression methods can be described very simply. The difference between fitting a line to points, and regression, is clarified by this simpler presentation. The use of n-1 in formulas for the standard deviation and the correlation coefficient is an unnecessary complication Weldon K. L. (2018).

Bayesian classification technique is used on student database to predict the student's division based on former year database and the study also shows that academic performances of the students are not always depending on their effort. It shows that other factors have got significant influence over students' performance Brijesh K. B. (2011).

A Binary Logistic Regression model is used to determine the probability of the student performing rate founded on the stated factors. These factors play a significant role at 5% level of significance. Thus, a Logistic Regression model to forecast the Academic Performance will be an effective tool for the decision-making method Surendheran R (2017).

Izaak (2015) Investigate that there is a positive relationship between concentration in Physics and knowledge of Mathematics basic ideas with students' ability to solve physics problems.

A weighted ordinary least square hierarchical multiple regression method was employed to the achieved quantity and quality of computer usage, significant predictors of achievement were establishing, Jehanzeb R. C (2013).

A significant relationship between self-concept and the respondents' academic performance in Mathematics was found by Merson P eta (2020)

Pearson's correlation coefficients of four independent are correlated with student's academic performance although two are not. Though, using the regression analysis four variables is significant which include: Time appropriateness, people-friend connection, nature of Usage, and health addiction while Time duration and security/privacy problems are not significant,

Sandra (2016).

The results obtained by the analyses conducted revealed that there were significant relationships between the students' academic achievement and student engagement as well as between their academic achievement and especially the dimensions of cognitive engagement, behavioral engagement, and sense of belonging, Selim G, Y. Y (2014).

A relationship between two quantitative variables usually involves a discussion of correlation and regression. When data is expressed in a standardized form, correlation and regression methods can be described very simply. The difference between fitting a line to points, and regression, is clarified by this simpler presentation. The use of n-1 in formulas for the standard deviation and the correlation coefficient is an unnecessary complication Weldon K. L. (2018).

Bayesian classification technique is used on student database to predict the student's division based on former year database and the study also shows that academic performances of the students are not always depending on their effort. It shows that other factors have got significant influence over students' performance Brijesh K. B. (2011).

2. Methodology

CORRELATION COEFFICIENT

The correlation coefficient.

$$r = \text{Cov}\,\frac{\sum(x,y)}{\sqrt{x}\sqrt{y}}$$

This is the formula of correlation coefficient.

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

Let x be $x - x$ and $y - y$

$$r = \frac{\frac{1}{N}\sum(x-x)(y-y)}{\sqrt{\frac{1}{N}\sum(x-x)^2((\sum x-x)^2)\times\frac{1}{N}\sum(y-y)^2((\sum y-y)^2)}}$$

$$r = \frac{\frac{1}{N}\sum(xy+xy-xy+xy)}{\sqrt{\frac{1}{N}\sum(x-x)^2((\sum x-x)^2)\times\frac{1}{N}\sum(y-y)^2((y-y)^2)}}$$

$$r = \frac{\sum xy\,\frac{\sum xy}{N}}{\sqrt{\frac{1}{N}\sum(x-x)^2((\sum x-x)^2)\times\frac{1}{N}\sum(y-y)^2-((y-y)^2)}}$$

$$r = \frac{N\sum xy-\sum x\sum y}{\sqrt{[N\sum x^2-(\sum x)^2\times N\sum y^2-(\sum y)^2}}$$

OF HYPOTHESIS FOR CORRELATION COEFFICIENT

H_o: $\rho = 0$

H_i: $\rho \neq 0$

In this, r has a student's t-distribution given by tcal = $\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ with n-2 degree of freedom we reject H_o. If tcal > t$\frac{\alpha}{2}$, n-2 degree of freedom or otherwise accept.

Decision Critical

Obtain $t_{\alpha/2}$, n-2 d.f if $t_{cal} > t_{\alpha/2}$, n-2.

The population correlation coefficient is often estimated, hence to test hypothesis concerning ρ this is the stage we try to test the null H_o: ρ = 0 in this case, ρ student's t-distribution is given by:

$$tcal = \frac{r\sqrt{n-2}}{\sqrt{1+r^2}}$$

With n-2 degree of freedom, were r is the sample correlation, n is the number of observation. In this case, if t_{cal} is greater than t_{tab} then we reject H_o: ρ = 0. Which is the null hypothesis and if t_{tab} is greater than t_{cal} then we accept H_o: ρ ≠ 0. Which is alternative hypothesis where is given significant level, and t_{tab} value with n-2.

SIMPLE LINEAR REGRESSION MODEL

The simplest linear regression model is given by

$$y = \beta_o + \beta_1 x + e. \quad \text{Where:}$$

x = independent variable

β_o = population that gives the intercept

β_1 = population that gives the gradient

e = random error

3.0 DATA ANALYSIS

This sections is the data analysis and it's finding where mathematics as the independent variable (x), while the other science subject is the dependent variable (y)

DECISION RULE:

H_0: There is a relationship between the two subjects.

H_1: There is no relationship between the two subjects.

3.1 COMPUTATION OF CORRELATION COEFFICIENT ON STUDENT PERFORMANCE (r) BETWEEN MATHS(x) AND PHYSICS(y) 2016 TABLE 3.0 ORIGINAL DATA

S/N	X	Y	x^2	y^2	xy
1	50	39	2500	1521	1950
2	60	70	3600	4900	4200

3	40	50	1600	2500	2000
4	50	80	2500	6400	4000
5	60	79	3600	6241	4740
6	75	60	5625	3600	4500
7	65	73	4225	5329	4745
8	67	80	4489	6400	5360
9	39	50	1521	2500	1950
10	50	60	2500	3600	3000
11	60	50	3600	2500	3000
12	70	60	4900	3600	4200
13	50	40	2500	1600	2000
14	40	50	1600	2500	2000
15	49	70	2401	4900	3430
16	39	59	1521	3481	2301
17	40	70	1600	4900	2800
18	60	80	3600	6400	4800
19	40	50	1600	2500	2000
20	50	70	2500	4900	3500
21	59	60	3481	3600	3540
22	50	90	2500	8100	4500
23	39	73	1521	5329	2847
24	70	49	4900	2401	3430
25	60	50	3600	2500	3000
26	39	90	1521	8100	3510
27	40	80	1600	6400	3200
28	50	59	2500	3481	2950
29	60	60	3600	3600	3600
TOTAL	1521	1851	83205	123783	97053

n = 29

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$r = \frac{29(97053) - (1521)(1851)}{\sqrt{\left[29(83205) - (1521)^2\right]\left[29(123783) - (1851)^2\right]}}$$

$$= \frac{2814537 - 2815371}{\sqrt{(2412945 - 2313441)(3589707 - 3426201)}}$$

$$= \frac{-834}{\sqrt{(99504)(163506)}} = \frac{-834}{127551.9542} \Rightarrow r = -0.0065$$

REMARK

The relationship is negative and is also a weak correlation between maths and physics

TEST FOR SIGNIFICANCE OF CORRELATION COEFFICIENT AT α = 5% AND 1% LEVEL

Test statistics

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{-0.0065\sqrt{29-2}}{\sqrt{1-(-0.0065)^2}}$$

$$= \frac{-0.016(5.19615)}{0.99997} = -0.0338$$

Therefore t_{cal} = -0.0818

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 27 = 2.052

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.771

Remarks

The result shows that t_{tab} > t_{cal} that is 2.052>2.771, 2.771>-0.0338, 2.771>0.0338, therefore, H_0 is accepted and H_1 is rejected and concluded that there is no significant difference between the student performance in maths and physics in both 5% and 1% level of significance.

COVARIANCE METHOD OF ESTIMATION OF REGRESSION PARAMETERS

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 83205 - \frac{(1521)^2}{29}$$

$$= 83205 - 79773.82759 = 3431.1724$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 123783 - \frac{(1851)^2}{29}$$

$$= 123783 - 118144.8621 = 5638.1380$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 97053 - \frac{(1521)(1851)}{29}$$

$$= 97053 - 97081.75862 = -28.7586$$

Where $b_1 = \frac{\sum xy}{\sum x^2} = \frac{S_{xy}}{Sxx} = \frac{-28.75862}{3431.1724} = -0.0084$

$$\overline{X} = \frac{\sum x}{n} = \frac{1521}{29} = 52.4483$$

$$\bar{y} = \frac{\sum y}{n} = \frac{1851}{29} = 63.8276$$

$$b_0 = \bar{y} + b_1 x \bar{x} = 63.8276 + 0.44056 = 64.2682$$

$\hat{y} = b_0 + b_1 x = 64.2682 - 0.0084x$ Is the fitted regression line

THE VARIANCE OF ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$\sigma^2 = \sum e_i^2 = \frac{1}{n-2}\left(S_{yy} - b_1 S_{xy}\right)$$

$$= \frac{1}{27}\left(3431.1724 - (-0.0084)(-28.7586)\right)$$

$$\frac{1}{27}\left(3431.1724 - 0.241137\right) = 90.0345$$

$$Var(b_1) = \frac{\sigma^2 n}{S_{xx}} = \frac{90.0345}{3431.1724} = 0.0262$$

THE STANDARD ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$S.E(b_1) = \sqrt{Var(b_1)} = \sqrt{0.0262} = 0.162$$

$$Var(b_0) = \sigma^2\left(\frac{\bar{x}^2}{S_{xx}} + \frac{1}{n}\right)$$

$$= 90.0345\left(\frac{(52.4483)^2}{3431.1724} + \frac{1}{29}\right)$$

$$= 90.0345(0.8017 + 0.03448)$$

$$= 90.0345(0.8361) = 75.2853$$

STANDARD ERROR ABOUT β_0 IS GIVEN BY:

$$S.E(b_0) = \sqrt{\text{var}(b_0)} = \sqrt{75.2853} = 8.6767$$

TEST FOR REGRESSION PARAMETERS β_0 AND β_1

Test for hypothesis for β_0

$$t_{cal} = \frac{b_1 - \beta_{00}}{S.E(b_0)} = \frac{64.2682}{0.1620}$$

$$= 396.7173$$

Test statistic

$t_{tab} = t_{\frac{\alpha}{2},n-2} = 0.025,27 = 2.052$

$t_{tab} = t_{\frac{\alpha}{2},n-2} = 0.005.27 = 2.771$

REMARK:

The result shows that $t_{tab} < t_{cal}$ that is 2.052 < 396.7173 , 2.771<396.7173 ,therefore, H_0 is rejected and H_1 is accepted and concludes that there is no significant difference between the student performance in maths and physics in both 5% and 1% level of significance.

3.2 COMPUTATION OF CORRELATION COEFFICIENT (r) BETWEEN MATHS AND BIOLOGY 2016

TABLE 3.1 ORIGINAL DATA

S/N	X	Y	XY	X^2	Y^2
1	50	80	2500	6400	4000
2	60	40	3600	1600	2400
3	40	59	1600	3481	2360
4	50	89	2500	7921	4450
5	60	70	3600	4900	4200
6	75	39	5625	1521	2925
7	65	70	4225	4900	4550
8	67	45	4489	2025	3015
9	39	60	1521	3600	2340
10	50	70	2500	4900	3500
11	60	80	3600	6400	4800
12	70	73	4900	5329	5110
13	50	80	2500	6400	4000
14	40	40	1600	1600	1600
15	49	50	2401	2500	2450
16	39	60	1521	3600	2340
17	40	50	1600	2500	2000
18	60	70	3600	4900	4200
19	40	70	1600	4900	2800
20	50	50	2500	2500	2500
21	59	70	3481	4900	4130
22	50	50	2500	2500	2500
23	39	60	1521	3600	2340
24	70	70	4900	4900	4900
25	60	60	3600	3600	3600
26	39	70	1521	4900	2730

27	40	39	1600	1521	1560
28	50	59	2500	3481	2950
29	60	49	3600	2401	2940
TOTAL	1521	1772	83205	113680	93190

N = 29

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}}$$

$$r = \frac{29(93190) - (1521)(1772)}{\sqrt{\left[29(83205) - (1521)^2\right]\left[29(113680) - (1772)^2\right]}}$$

$$= \frac{2702510 - 2695212}{\sqrt{(2412945 - 2313441)(3296720 - 3139984)}}$$

$$= \frac{7298}{\sqrt{(99504)(156736)}} = \frac{7298}{124883.3814}, \Rightarrow r = 0.0584$$

REMARK

The correlation is positive but weak correlation

TEST FOR SIGNIFICANCE OF CORRELATION COEFFICIENT AT α = 5% AND α = 1% LEVEL

Test statistics

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.0584\sqrt{29-2}}{\sqrt{1-(0.0584)^2}}$$

$$= \frac{0.0584(3.196\grave{5})}{0.9983} = 0.3040$$

Therefore $t_{cal} = 0.3040$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 27 = 2.052

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.771

Remark

Since $t_{tab} > t_{cal}$ (2.052 > 0.3040, 2.771>0.3040), therefore (H_0) is accepted and H_1 is rejected and concluded that there is no significance difference between the student performance in maths and biology in both 5% and 1% level of significance.

COVARIANCE ESTIMATION OF REGRESSION PARAMETERS

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 83205 - \frac{(1521)^2}{29}$$

$$= 83205 - 79773.82759 = 3431.1724$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 113680 - \frac{(1772)^2}{29}$$

$$= 113680 - 108275.3103 = 5404.6897$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 93190 - \frac{(1521)(1772)}{29}$$

$$= 93190 - 92938.34483 = 251.6552$$

Where $b_1 = \frac{\sum xy}{\sum x^2} = \frac{S_{xy}}{Sxx} = \frac{251.6552}{3431.1724} = 0.0733$

$$\overline{X} = \frac{\sum x}{n} = \frac{1521}{29} = 52.4483$$

$$\overline{y} = \frac{\sum y}{n} = \frac{1772}{29} = 61.1034$$

$$b_0 = \overline{y} + b_1 \overline{x} = 61.1034 + (0.0733)(52.4483) = 61.1034 - 3.8445 = 57.2589$$

$$\hat{y} = b_0 + b_1 x_1 = 57.2589 + 0.0733 x_1 \text{ is the fitted regression line}$$

THE VARIANCE OF ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$\sigma^2 = \sum e_i^2 = \frac{1}{n-2}\left(S_{yy} - b_1 S_{xy}\right)$$

$$= \frac{1}{29-2}\left(3431.17724 - (0.0733)(251.6552)\right)$$

$$\frac{1}{27}(3412.7261) = 126.3973$$

$$Var(b_1) = \frac{\sigma^2 n}{S_{xx}} = \frac{126.3973}{3431.17724} = 0.0368$$

THE STANDARD ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$S.E(b_0) = \sqrt{Var(b_0)} = \sqrt{0.0368} = 0.1919$$

THE VARIANCE OF ERROR TERM ABOUT (β_1)

$$Var(b_0) = \sigma^2 \left(\frac{\overline{x}^{\,2}}{S_{xx}} + \frac{1}{n} \right)$$

$$=126.3973\left[\frac{(52.4483)}{3431.1724}+\frac{1}{29}\right]=126.3973\times 0.8362=105.6912$$

STANDARD ERROR TERM ABOUT β_1 IS GIVEN BY:

$$S.E(b_0)=\sqrt{\text{var}(b_0)}=\sqrt{105.6912}=10.2806$$

THE TEST FOR REGRESSION PARAMETERS β_0 AND β_1

$$t_{cal}=\frac{b_1-\beta_{00}}{S.E(b_0)}=\frac{57.2589}{0.1919}=298.3788$$

TEST STATISTIC

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 27 = 2.052

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.771

REMARK

The result shows that $t_{tab} < t_{cal}$ that is 2.052 < 298.37788, 2.771< 298.37788, therefore, H_0 is rejected and H_1 is accepted and conclude that there is significance difference between the student performance in Maths and Biology in both 5% and 1% level of significance.

TEST FOR HYPOTHESIS FOR β_1

$$t_{cal}=\frac{b_1-\beta_{10}}{S.E(b_1)}=\frac{0.0733}{10.2806}=0.0071$$

TEST STATISTIC

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 27 = 2.052

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.771

REMARK

The result shows that $t_{tab} > t_{cal}$ that is 2.052 > 0.0071, 2.771> 0.0071 therefore H_0 is accepted and H_1 is rejected and conclude that there is no significance difference between in Maths and Biology in both 5% and 1% level of significance.

4.3 COMPUTATION OF CORRELATION COEFFICIENT (r) BETWEEN MATHS AND CHEMISTRY 2017

TABLE 3.2 ORIGINAL DATA

S/N	X	Y	X^2	Y^2	XY
1	85	89	7225	7921	7565
2	83	39	6889	1521	3237

3	72	50	5184	2500	3600
4	50	65	2500	4225	3250
5	66	48	4356	2304	3168
6	80	43	6400	1849	3440
7	45	45	2025	2025	2025
8	58	58	3364	3364	3364
9	50	53	2500	2809	2650
10	40	40	1600	1600	1600
11	45	81	2025	6561	3645
12	50	73	2500	5329	3650
13	70	40	4900	1600	2800
14	39	88	1521	7744	3432
15	70	70	4900	4900	4900
16	66	39	4356	1521	2574
17	37	88	1369	7744	3256
18	60	80	3600	6400	4800
19	50	76	2500	5776	3800
20	60	60	3600	3600	3600
21	47	49	2209	2401	2303
22	70	90	4900	8100	6300
23	90	49	8100	2401	4410
24	49	42	2401	1764	2058
25	50	72	2500	5184	3600
26	70	80	4900	6400	5600
27	50	78	2500	6084	3900
28	80	39	6400	1521	3120
29	59	80	3481	6400	4720
30	73	66	5329	4356	4818
TOTAL	1814	1870	116034	125904	111185

N = 30

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}}$$

$$r = \frac{30(111185) - (1814)(1870)}{\sqrt{\left[30(116034) - (1814)^2\right]\left[30(125904) - (1870)^2\right]}}$$

$$= \frac{3335550 - 3392180}{\sqrt{(3481020 - 3290596)(3777120 - 3496900)}}$$

$$= \frac{-56630}{\sqrt{(190424)(280220)}} = \frac{-56630}{230999.1629} \Rightarrow r = -0.2452$$

REMARK

The correlation is negative but week correlation

TEST FOR SIGNIFICANCE OF CORRELATION COEFFICIENT AT α = 5% AND 1% LEVEL

Test statistics

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{-0.2452\sqrt{30-2}}{\sqrt{1-(-0.2452)^2}} = \frac{-0.245(5.2915)}{0.9695} = -1.3235$$

Therefore t_{cal} = -1.3235

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

Since $t_{tab} > t_{cal}$ (2.048 > -1.3235 and 2.763>-1.3235), therefore H_0 is accepted and

H_1 reject and conclude that there is no significant difference between the student performance in maths and chemistry in both 5% and 1% level of significance.

COVARIANCE METHOD ESTIMATION OF REGRESSION PARAMETERS

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 116034 - \frac{(1814)^2}{30}$$

$$= 116034 - 109686.533 = 6347.4667$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 125904 - \frac{(1870)^2}{30}$$

$$= 125904 - 116563.3333 = 9340.6667$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 111185 - \frac{(1814)(1870)}{30} = 111185 - 113072.6667 = -1887.6667$$

Where $b_1 = \frac{\sum xy}{\sum x^2} = \frac{S_{xy}}{Sxx} = \frac{-1887.6667}{6347.4667} = -0.2974$

$$\overline{X} = \frac{\sum x}{n} = \frac{1814}{30} = 60.4667$$

$$\overline{y} = \frac{\sum y}{n} = \frac{1870}{30} = 62.3333$$

$$b_0 = \overline{y} - b_1\overline{x} = 62.3333 - (-0.2974)(60.4667) = 80.3161$$

$$\overline{y} = b_0 + b_1 x_1 = 80.3161 - 0.2974 x_1 \ \textit{fitted the regression line}$$

THE VARIANCE OF ERROR TERM ABOUT (β_1) IS GIVEN BY:

$$\sigma^2 = \sum e_i^2 = \frac{1}{n-2}\left(S_{yy} - b_1 S_{xy}\right)$$

$$= \frac{1}{30-2}\left(6347.4667 - \left(-0.2974\right)\left(-1887.6667\right)\right)$$

$$\frac{1}{28}\left(6347.4667 - 561.3921\right) = 206.6455$$

$$Var(b_1) = \frac{\sigma^2 n}{S_{xx}} = \frac{206.6455}{6347.4667} = 0.0326$$

THE STANDARD ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$S.E(b_1) = \sqrt{Var(b_0)} = \sqrt{0.0326} = 0.1804$$

THE VARIANCE OF ERROR TERM ABOUT (β_1)

$$Var(b_0) = \sigma^2\left(\frac{\overline{x}^{\,2}}{S_{xx}} + \frac{1}{n}\right)$$

$$= 206.6455\left(\frac{(60.4667)^2}{6347.4667} + \frac{1}{30}\right) = 206.6455(0.57600 + 0.0333) = 125.9160$$

STANDARD ERROR TERM ABOUT β_1 IS GIVEN BY:

$$S.E(b_0) = \sqrt{\text{var}(b_1)} = \sqrt{125.9160} = 11.2212$$

THE TEST FOR REGRESSION PARAMETERS FOR β_0 AND β_1

Test Statistics

$$\frac{b_1 - \beta_{00}}{S.E(b_0)} = \frac{80.3161}{0.1804} = 445.2112$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

REMARK:

The result obtained shows that $t_{tab} < t_{cal}$ that is 2.048<445.2112, 2.763<445.2112, therefore, H_0 is rejected and H_1 is accepted and concludes that there is a significant difference between the student performance in maths and chemistry in both 5% and 1% level of significance.

Test Statistics

$$\frac{b_1 - \beta_{00}}{S.E(b_0)} = \frac{-0.2974}{11.2212} = -0.0265$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

REMARK:

The result obtained shows that $t_{tab} > t_{cal}$ that is 2.048>-0.0265, 2.763>-0.0265, therefore, H_0 is accepted and H_1 is rejected and concludes that there is no significant difference between the student performance in maths and chemistry in both 5% and 1% level of significance.

3.4 COMPUTATION OF CORRELATION COEFFICIENT (r) BETWEEN MATHS AND GEOGRAPHY 2017

TABLE 3.3 ORIGINAL DATA

S/N	X	Y	x^2	y^2	Xy
1	85	50	7225	2500	4250
2	83	60	6889	3600	4980
3	72	70	5184	4900	5040
4	50	45	2500	2025	2250
5	66	90	4356	8100	5940
6	80	60	6400	3600	4800
7	45	49	2025	2401	2205
8	58	50	3364	2500	2900
9	50	39	2500	1521	1950
10	40	63	1600	3969	2520
11	45	59	2025	3481	2655
12	50	39	2500	1521	1950
13	70	40	4900	1600	2800
14	39	39	1521	1521	1521
15	70	40	4900	1600	2800
16	66	48	4356	2304	3168
17	37	82	1369	6724	3034
18	60	39	3600	1521	2340

19	50	62	2500	3844	3100
20	60	58	3600	3364	3480
21	47	50	2209	2500	2350
22	70	66	4900	4356	4620
23	90	72	8100	5184	6480
24	49	55	2401	3025	2695
25	50	70	2500	4900	3500
26	70	59	4900	3481	4130
27	50	53	2500	2809	2650
28	80	49	6400	2401	3920
29	59	80	3481	6400	4720
30	73	90	5329	8100	6570
TOTAL	1814	1726	116034	105752	105318

N = 30

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$r = \frac{30(105318) - (1814)(1726)}{\sqrt{\left[30(116034) - (1814)^2\right]\left[30(105752) - (1726)^2\right]}}$$

$$= \frac{3156540 - 3130964}{\sqrt{(3481020 - 3290596)(3172560 - 2979676)}}$$

$$= \frac{25576}{\sqrt{(190424)(193484)}} = \frac{25576}{191947.9023}$$

$r = 01332$

REMARK

The correlation is positive but weak correlation

TEST FOR SIGNIFICANCE OF CORRELATION COEFFICIENT AT α = 5% AND 1% LEVEL

Test statistics

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.1332\sqrt{30-2}}{\sqrt{1-(0.1332)^2}} = \frac{0.1332(5.2915)}{\sqrt{1-0.0177}} = \frac{0.7048}{0.9911} = 0.7112$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab} > t_{cal}$ that is 2.48> 0.7112, 2.763>0.7112, therefore, H_0 is accepted and H_1 is rejected and conclude that there is no significance in maths and Geography in both 5% and 1% level of significance

COVARIANCE ESTIMATION OF REGRESSION PARAMETERS

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 11634 - \frac{(1814)^2}{30} = 11634 - 109686.5333 = 6347.4667$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 105752 - \frac{(1726)^2}{30} = 105752 - 99302.5333 = 6449.4667$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 105318 - \frac{(1814)(1726)}{30} = 105318 - 104365.4667 = 952.5333$$

Where $b_1 = \frac{\sum xy}{\sum x^2} = \frac{S_{xy}}{Sxx} = \frac{952.5333}{6347.46667} = 0.1501$

$$\overline{X} = \frac{\sum x}{n} = \frac{1814}{30} = 60.44467$$

$$\overline{y} = \frac{\sum y}{n} = \frac{1726}{30} = 57.5333$$

$$b_0 = \overline{y} - b_1 \overline{x} = 57.53333 - (0.1501)(60.46667) = 57.5333 - 9.0761 = 48.4572$$

$$\overline{y} = b_0 + b_1 x_1 = 48.4572 + 0.150 x_1 \ \textit{fitted the regression line}$$

THE VARIANCE OF ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$\sigma^2 = \sum e_i^2 = \frac{1}{n-2}\left(S_{yy} - b_1 S_{xy}\right)$$

$$= \frac{1}{30-2}\left(6347.4667 - (0.1501)(952.5333)\right) = \frac{1}{28}\left(6347.4667 - 112.9552\right) = 222.6611$$

$$Var(b_1) = \frac{\sigma^2 n}{S_{xx}} = \frac{222.6611}{6347.4667} = 0.0351$$

THE STANDARD ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$S.E(b_1) = \sqrt{Var(b_1)} = \sqrt{0.0351} = 0.1873$$

THE VARIANCE OF EROR TERM ABOUT (β_1)

$$Var(b_0) = \sigma^2 \left(\frac{\overline{x}^{\,2}}{S_{xx}} + \frac{1}{n} \right)$$

$$= 222.6611\left(\frac{(60.4667)^2}{6347.4667}+\frac{1}{30}\right)=222.6611(0.576+0.0333)=222.6611(0.60933)=135.6741$$

STANDARD ERROR ABOUT β_1 IS GIVEN BY:

$$S.E(b_0)=\sqrt{\text{var}(b_0)}=\sqrt{135.6741}=11.6479$$

THE TEST FOR REGRESSION PARAMETERS β_0 AND β_1

Test Statistics

$$t_{cal}=\frac{b_1-\beta_{00}}{S.E(b_0)}=\frac{48.4572}{0.1873}=258.7144$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab} < t_{cal}$ that is 2.048<258.7144, 2.763<258.7144, therefore, H_0 is rejected and H_1 is accepted and conclude that there is significance in maths and Geography in both 5% and 1% level of significance

Test Statistics

$$t_{cal}=\frac{b_1-\beta_{10}}{S.E(b_1)}=\frac{0.1501}{11.6479}=0.0129$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab} > t_{cal}$ that is 2.048>0.0129, 2.763>0.0129 therefore H_0 is accepted and H_1 is rejected and conclude that there is no significance in maths and Geography in both 5% and 1% level of significance

3.5 COMPUTATION OF CORRELATION COEFFICIENT (r) BETWEEN MATHS AND AGRIC 2018

TABLE 4.4 ORIGINAL DATA

S/N	X	Y	X^2	Y^2	XY
1	85	70	7225	4900	5950
2	75	50	5625	2500	3750
3	63	90	3969	8100	5670
4	73	100	5329	10000	7300

5	72	83	5184	6889	5976
6	62	88	3844	7744	5456
7	70	86	4900	7396	6020
8	81	87	6561	7569	7047
9	73	78	5329	6084	5694
10	60	89	3600	7921	5340
11	82	88	6724	7744	7216
12	90	77	8100	5929	6930
13	43	75	1849	5625	3225
14	50	48	2500	2304	2400
15	98	67	9604	4489	6566
16	89	78	7921	6084	6942
17	75	91	5625	8281	6825
18	81	67	6561	4489	5427
19	63	55	3969	3025	3465
20	70	58	4900	3364	4060
21	73	39	5329	1521	2847
22	49	78	2401	6084	3822
23	80	69	6400	4761	5520
24	90	68	8100	4624	6120
25	54	83	2916	6889	4482
26	62	90	3844	8100	5580
27	72	76	5184	5776	5472
28	49	46	2401	2116	2254
29	45	59	2025	3481	2655
30	55	78	3025	6084	4290
TOTAL	2084	2211	150944	169873	154301

N = 30

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$r = \frac{30(154301) - (2084)(2211)}{\sqrt{\left[30(150944) - (2084)^2\right]\left[30(169873) - (2211)^2\right]}} = \frac{4629030 - 4607724}{\sqrt{(4528320 - 4343056)(5096190 - 4888321)}}$$

$$= \frac{21306}{\sqrt{(185264)(207669)}} = \frac{21306}{196146.8573} \quad \Rightarrow r = 0.1086$$

REMARK

The correlation is positive but weak correlation

TEST FOR SIGNIFICANCE OF CORRELATION COEFFICIENT AT α = 5% AND 1% LEVEL

Test statistics

$$t_{cal} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.1086\sqrt{30-2}}{\sqrt{1-(0.1086)^2}} = \frac{0.57465}{0.9941} = 0.5781$$

Therefore $t_{cal} = 0.2931$

$t_{tab}, t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

$t_{tab}, t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab} > t_{cal}$ that is 2.048>0.5781, 2.763>0.5781, therefore, H_0 is accepted and H_1 is rejected and concluded that there is no significance in maths and Agric in both 5% and 1% level of significance

COVARIANCE METHOD ESTIMATION OF REGRESSION PARAMETERS

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 150944 - \frac{(2084)^2}{30} = 150944 - 144768.5333 = 6175.4667$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 169873 - \frac{(2211)^2}{30} = 169873 - 162950.7 = 6922.3$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 154301 - \frac{(2084)(2211)}{30} = 154301 - 153590.8 = 710.2$$

Where $b_1 = \frac{\sum xy}{\sum x^2} = \frac{S_{xy}}{Sxx} = \frac{710.2}{6175.4667} = 0.1150$

$$\overline{X} = \frac{\sum x}{n} = \frac{2084}{30} = 69.4667$$

$$\overline{y} = \frac{\sum y}{n} = \frac{2211}{30} = 73.7$$

$$b_0 = \overline{y} - b_1\overline{x} = 73.7 - (0.1150)(69.4667) = 73.7 - 7.98867 = 65.7113$$

$$\overline{y} = b_0 + b_1x_1 = 65.7113 + 0.115x_1 \ \textit{fitted the regression line}$$

THE VARIANCE OF ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$\sigma^2 = \sum e_i^2 = \frac{1}{n-2}\left(S_{yy} - b_1 S_{xy}\right)$$

$$= \frac{1}{30-2}\left(6175.4667 - (0.1150)(710.2)\right) = \frac{1}{28}\left(6175.4667 - 81.673\right) = 217.6354$$

$$Var(b_0) = \frac{\sigma^2 n}{S_{xx}} = \frac{217.6354}{6175.4667} = 0.03524$$

THE STANDARD ERROR TERM ABOUT (β_0) IS GIVEN BY:

$$S.E(b_1) = \sqrt{Var(b_0)} = \sqrt{0.03524} = 0.1877$$

$$Var(b_0) = \sigma^2\left(\frac{\bar{x}^2}{S_{xx}} + \frac{1}{n}\right)$$

$$= 217.6354\left(\frac{(69.4667)^2}{6175.4667} + \frac{1}{30}\right) = 217.6354(0.81472) = 177.31153$$

STANDARD ERROR TERM ABOUT β_1 IS GIVEN BY:

$$S.E(b_0) = \sqrt{\text{var}(b_1)} = \sqrt{177.31153} = 13.3158$$

THE TEST FOR REGRESSION PARAMETERS β_0 AND β_1

Test Statistics

$$t_{cal} = \frac{b_1 - \beta_{00}}{S.E(b_0)} = \frac{65.7113}{0.1877} = 350.0868$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab} < t_{cal}$ that is 2.048<350.0868, 2.763<350.0868, therefore, H_0 is rejected and H_1 is accepted and conclude that there is no significance in maths and Agric in both 5% and 1% level of significance

Test Statistics

$$t_{cal} = \frac{b_1 - \beta_{10}}{S.E(b_1)} = \frac{0.1150}{13.3158} = 0.0086$$

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.025,}$ 28 = 2.048

t_{tab}, $t_{\alpha/2}$. n-2 = $t_{0.005,}$ 27 = 2.763

Remarks

The result obtained shows that $t_{tab}>t_{cal}$ that is 2.048>0.0086, 2.763>0.0086, therefore, H_0 is accepted and H_1 is rejected and conclude that there is no significance in maths and Agric in both 5% and 1% level of significance.

4.0 SUMMARY, CONCLUSION AND RECOMMENDATIONS

4.1 SUMMARY

This project is summarized in five chapters, chapter one is the introduction, historical background of government girl's secondary school, aims and objectives, method of data collection, the significance of the study, scope, and limitation of the study, and definition of terms and concept used. Chapter two includes a literature review. Chapter three Methodology. Chapter four data analysis and chapter five Summary, Conclusion, and Recommendations.

4.2 CONCLUSION

From table 4.1 the result of the analysis of correlation coefficient r = -0.0065 which shows a negative and is also a weak correlation between maths and physics, Table 4.2 shows that r = 0.0584 which shows a positive and is also a weak correlation between maths and biology, Table 4.3 reveals that r = -0.2452 which shows a negative and is also a weak correlation between maths and Chemistry, From table 4.4 r = 0.1332 which shows a positive and is also a weak correlation between maths and Chemistry, From table 4.5 r = 0.10038 which shows a positive and is also a weak correlation between maths and agric.

Test at 5% and 1% level of significance of correlation coefficient of both tables shows there is no significant difference in performance between maths and physics, and the covariance method of estimation of regression parameters for β1 shows there is a significant difference in student performance and β0 shows there is no significant difference.

In table 4.1the regression equation it reveals that a = 64.2682 and b = -0.008, its indicates that the graph is moving downward and it is also shown that there is a slight decrease in student performance, and table 4.2 show the regression equation that a = 57.2589 and b = 0.0733, its indicates that the graph is moving upward and it is also shown that there is a slight increase in student performance, table 4.3 it reveals that the regression equation a = 80.3161 and b = -0.2974, its indicates that the graph is moving downward and it is also shown that there is a decrease in student performance, Table 4.4 found that the regression equation a = 48.4572 and b = 0.1501, its indicates that the graph is moving upward and it is also shown that there is a increase in student performance, Table 4.5 From the result of the regression equation, it reveals that a = 80.3161 and b = -0.2974, its indicates that the graph is moving downward and it is also shown that there is a decrease in student performance. Therefore in general the performance of the students is fluacting over the period of years. However the correlations indicate that student's performance shows independency. The government would employ qualified teachers. Students would dedicate themselves to hard work and dedication to learning. Provision of more laboratory equipment to enhanced student and learning process

Reference

Ahono T. A , Jairo P. A, and Calleb O .G (2018) "*Influence of Efficacy Expectation on Mathematics Achievement among Students in Secondary Schools in Kenya"* International Journal of Education and Research. Vol. 6 No. Vol. 6 No. 8

Brijesh K. B. and Saurabh P. (2011) *"Data Mining: A prediction for performance improvement using classification"* IJCSIS International Journal of Computer Science and Information Security, Vol. 9, No. 4,

Galadanci, B. S., and Mukhtar, M.I. (2017) *"A Correlational Study of Students 'Theoretical And Practical Examinations Scores In Computer Applications Courses In Bayero University Kano" Science World Journal Vol 12(No 2) www.scienceworldjournal.org ISSN 1597-6343*

Izaak H. W (2015) "*The Correlation Study of Interest at Physics and Knowledge of Mathematics Basic Concepts towards the Ability to Solve Physics Problems of 7th Grade Students at Junior High School in Ambon Maluku Province, Indonesia"*. Hindawi Publishing Corporation Education Research International Volume 2015, Article ID 396750, 6 pages http://dx.doi.org/10.1155/396750

Jehanzeb R. C and Bo Z.(2013) " *Quantity and quality of computer use and academic achievement: Evidence from a large-scale international test program"* international Journal of Education and Development using Information and Communication Technology (IJEDICT), Vol. 9, Issue 2, pp. 95-106

merson P, Antonietta G, Jonathan O. E, Aahron D, Regina S, and Redjie A (2020) *" Factors Affecting Mathematics Performance of Junior High School Students"* INTERNATIONAL ELECTRONIC JOURNAL OF MATHEMATICS EDUCATION e-ISSN: 1306-3030. Vol. 15, No. 1, em0556 https://doi.org/10.29333/iejme/5938

Razak F. A., Baharun N., Deraman N. A. and Ismail N. R. P (2017) " *ASSESSING STUDENTS' ABILITIES IN INTERPRETING THE CORRELATION AND REGRESSION ANALYSIS".* Journal of Fundamental and Applied Sciences ISSN 1112-9867

Sandra O.M and Ismail N, (2016) "*the impact of social media on students' academic performance- a case of malaysia tertiary institution"* International Journal of Education, Learning and Training Vol. 1 (No.1), ISSN: 2289-669 , DOI: 10.24924/ijelt/v1.iss1/14.21

Selim G, Y. Y (2014) "*the relationships between student engagement and their academic achievement*" international Journal on New Trends in Education and Their Implications Volume: 5 Issue: 4 Article: 19 ISSN 1309-6249.

Surendheran R, and Monisha R, 2017 *"Application of Logistic Regression model to determine the Academic Performance of MBA Students of Department of Management Studies"*, INTERNATIONAL JOURNAL OF MANAGEMENT & BUSINESS STUDIES VOL. 7, ISSUE 2, SSN : 2230- 9519 (Online) | ISSN : 2231-2463 (Print)

Weldon K. L. (2018) "*A Simplified Introduction to Correlation and Regression'*, Simon Fraser University, Burnaby, BC, Canada V5A 1S6