Answer on Question #78943 – Math – Statistics and Probability

Question

A group of 25 students took examinations in both pure mathematics and statistics. Their marks out of 150 in mathematics, x, and in statistics, y, were recorded and are summarized below.
Σ x = 1978, Σ x2 = 175840 Σ y = 2123 Σ y2 = 202257 Σ xy = 181572
i. Calculate Sxx, Syy and Sxy.
ii. Find the product moment correlation coefficient between the marks in pure Mathematics and Statistics.
iii. Starting your hypotheses clearly tests, at the 5% level of significance, whether or not there is evidence of a correlation.
iv. State an assumption needed for the test in part (iii) to be valid.

Solution

i. We have

$$S_{xx} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\Sigma\, x2 - n\bar{x}^2}{n-1} = \frac{175840 - \frac{1978^2}{25}}{24} = 805.86$$

$$S_{yy} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1} = \frac{\Sigma\, y2 - n\bar{y}^2}{n-1} = \frac{202257 - \frac{2123^2}{25}}{24} = 915.49$$

$$S_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{\Sigma\, xy - \frac{\Sigma x\, \Sigma y}{n}}{n-1} = \frac{181572 - \frac{1978 \cdot 2123}{25}}{24} = 566.68$$

ii. We have

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{566.68}{\sqrt{805.86 \cdot 915.49}} = 0.66$$

iii. The null-hypothesis is that there is no correlation between the marks in Math and in Statistics. The alternative hypothesis is that there is a correlation between these marks.
The test statistic is

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.66\sqrt{\frac{23}{1-0.66^2}} = 4.21$$

Then with F being t-distribution with n-2=23 degrees of freedom

P-value=2*(1-F(t))=2*(1-0.9998)=0.0003<0.05
We can reject the null-hypothesis that there is no correlation. So there is evidence of linear correlation.

iv. The observations should be independent of one another. Math marks variable should be approximately normally distributed. It should not contain any outliers.