Test Provider Report

Gev Chalikyan

October 2, 2025

1   Data Analysis

The results from 100 samples sent to two test providers were analyzed using several Python libraries. For clarity, data from test1 will be referred to as provided by Test Provider A and data from test2 as provided by Test Provider B.

The primary tools used were:

- **pandas**: Used for data processing and management.

- **matplotlib.pyplot**: Used for data visualization.

- **statsmodels.formula.api.ols**: Used for fitting the linear regression model.

An OLS regression was conducted to determine the relationship between the two providers, with the results from Test Provider A serving as the predictor variable for Test Provider B.

1.1   P-value

First, we will adopt the null hypothesis that the data from our two test providers have no relationship. Thus, disproving this hypothesis would mean that there must be a relationship present.

$$H_0 : \beta_1 = 0$$
$$H_a : \beta_1 \neq 0 \tag{1}$$

Observing our findings, we see the P-value calculated by our regression model has a value of $3.7213e - 41$. For the purposes of our analysis, we will consider our significance level to be represented by $\alpha = 0.05$.

To further our analysis, we compare our calculated P-value with our significance level:

$$3.7213e - 41 < 0.05 \tag{2}$$

As shown by inequality (2), our calculated P-value is well below our chosen significance level, therefore we reject our null hypothesis and conclude that there is a relationship between our two datasets.

1.2 R^2

To measure how well our model fits our data, we can analyze our calculated R^2 from our OLS regression:

$$R^2 = 0.8428 \tag{3}$$

This value tells us that 84.28% of the variability in our test results from Test Provider B can be explained by the test results from Test Provider A. This suggests a strong relationship between the two datasets, as we would expect.

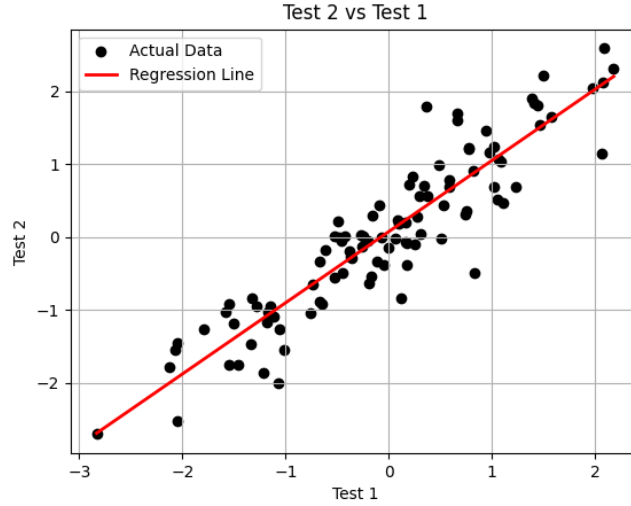## 1.3 Regression Visualized

Figure 1: Data from Test Provider B plotted against data from Test Provider A with the fitted regression line.

Our tests have calculated the estimated value for β_1 ($\hat{\beta}_1$) to be approximately 0.9759, indicating a nearly identical expected test result when providing a sample to either test provider.

Observing the line drawn in Figure 1, we can see that the regression line is very close to what we want to see if the tests are equivalent: $y = x$.

2

2 Conclusions

While the relationship of the sample data is not necessarily one-to-one, it is very close. This tells us that we can expect almost the same test results from either provider.

Therefore, based on our sample data, there is seemingly no difference between test results from either provider and we recommend testing samples with whichever provider is the most cost effective.