

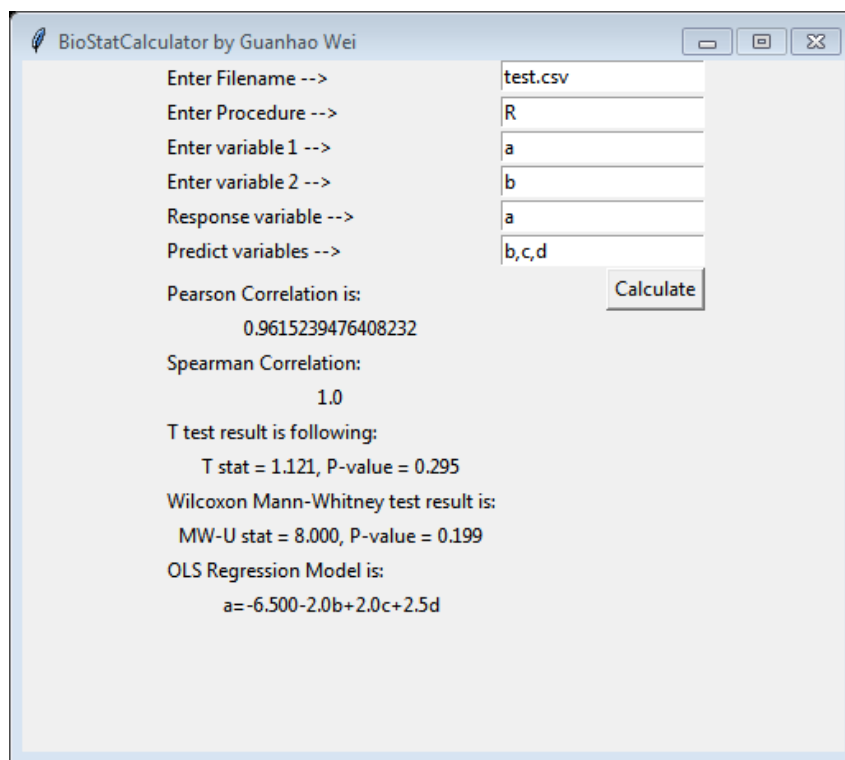
## GUI Biostatistics Inference Calculator

### Function Description:

This calculator works on several simple stat calculations that commonly used in biostat calculation. The user interface is as following, which has functions as following. We can read a given CSV file. Next, we can enter a procedure, which include Pearson's Correlation, Spearman's Correlation, Two sample T test, Two sample Wilcoxon Test, and Mutivariate Regression using ordinary Least square method.

In the procedure text box, we can enter 'P' or 'Pearson' to represent Pearson's Correlation, enter 'S' or 'Spearman' to represent Spearman's Correlation, enter 'T' or 'T test' to represent T sample t test, enter 'W' or 'Wilcoxon' to represent two sample Wilcoxon test and finally, enter 'R' or 'Regression' to represent Multivariate Linear Regression.

Next, we shall enter names of variable 1 and variable 2 that can be used to calculate correlations and tests. For regression analysis, we need to specify Response variable and Predict variable.



The screenshot shows a window titled "BioStatCalculator by Guanhao Wei". It contains several input fields and a "Calculate" button. The inputs are: "Enter Filename -->" with "test.csv", "Enter Procedure -->" with "R", "Enter variable 1 -->" with "a", "Enter variable 2 -->" with "b", "Response variable -->" with "a", and "Predict variables -->" with "b,c,d". Below the inputs, the results are displayed: "Pearson Correlation is: 0.9615239476408232", "Spearman Correlation: 1.0", "T test result is following: T stat = 1.121, P-value = 0.295", "Wilcoxon Mann-Whitney test result is: MW-U stat = 8.000, P-value = 0.199", and "OLS Regression Model is: a = -6.500 - 2.0b + 2.0c + 2.5d".

For Response variable, we can only enter one variable, and for predict variable, we can enter one or several variables, as long as they are in dataset, to do prediction, using comma as separation. Finally, there is a calculation button, this button depends on calculation procedure we chose.

Calculation results could be shown below the calculation button.

Pearson's Correlation is :

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

Spearman's Rank-Order Correlation is :

$$\rho_{\text{rg}_X, \text{rg}_Y} = \frac{\text{cov}(\text{rg}_X, \text{rg}_Y)}{\sigma_{\text{rg}_X} \sigma_{\text{rg}_Y}}$$

T test results contains test statistic and p values:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \sqrt{2/n}}$$

Wilcoxon test results contains test statistic and p values:

$$W = \sum_{i=1}^{N_r} [\text{sgn}(x_{2,i} - x_{1,i}) \cdot R_i]$$

Multivariate Linear Regression result is:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} = \left( \sum \mathbf{x}_i \mathbf{x}_i^\top \right)^{-1} \left( \sum \mathbf{x}_i y_i \right)$$

### Code description:

To generate a high efficient calculator GUI, we need to get help from existing Python module, which are tkinter, pandas, scipy, and statsmodels. So we need to install those packages first and import.

'tkinter' module can be used to generate GUI windows and frame.

'pandas' module can be used to read CSV file and construct dataset in dataframe. Specify each variable name and corresponding data.

'scipy' module mainly works on calculating correlations and tests.

'statsmodels' module works to do OLS regression and give output.

Then, our calculation could be separate into 3 steps.

First, input required information in the textbox. And then the code use getters to restore variable name and data.

Second, according to given calculation procedure specified, we can get do calculation using certain module and save output results.

Finally, select and rewrite saved results in a clear and simple way. Output all of the results in correct order and show on the windows as label type.

Also, if there is wrong input, for example, wrong path, wrong variable name, and wrong procedure name, then the error detection warning will show up. That could tell users they should double check their inputs under current calculation procedure.

