The confidence intervals are calculated differently depending on the hypothesis for the value of the correlation coefficient. If /RHO is not used the confidence intervals are computed using the critical value F_{c2}, otherwise they are computed using the critical Z_{c2} and sigmaZ.

References

Zar, J.H., *Biostatistical Analysis*, 4th ed., 929 pp., Prentice Hall, Englewood Cliffs, New Jersey, 1999.

See Also

Chapter III-12, **Statistics** for a function and operation overview; **StatsCircularCorrelationTest**, **StatsMultiCorrelationTest**, and **StatsRankCorrelationTest**.

StatsLinearRegression

`StatsLinearRegression` `[`***`flags`***`] [`***`wave0, wave1`***`,…]`

The StatsLinearRegression operation performs regression analysis on the input wave(s). Output is to the W_StatsLinearRegression wave in the current data folder or optionally to a table. Additionally, the M_DunnettMCElevations, M_TukeyMCSlopes, and M_TukeyMCElevations waves may be created as specified.

Flags

/ALPH = *val* Sets the significance level (default *val*=0.05).

/B=*beta0* Tests the hypothesis that the slope *b*= *beta0* (default is 0). The results are expressed by the t-statistic, which can be compared with the tc value for the two-tailed test. Get the critical value for a one-tailed test using `StatsStudentCDF(1-alpha,N-2)`. It does not work with /MYVW.

/BCIW Computes two confidence interval waves for the high side and the low side of the confidence interval. The new waves are named with _CH and _CL suffixes respectively appended to the Y wave name and are created in the current data folder. For multiple runs a numeric suffix will also be appended to the names.

/BPIW[=*mAdditional*]

Computes prediction interval waves for the high side and the low side of the confidence interval on a single additional measurement (default). Use *mAdditional* to specify additional measurements. The new waves are named with _PH and _PL suffixes respectively appended to the Y wave name and are created in the current data folder. For multiple runs a numeric suffix will also be appended to the names.

/DET=*controlIndex* Performs Dunnett’s multicomparison test for the elevations. The test requires more than two Y waves for regression, the test for the slopes should not reject the equal slope hypothesis, and the test for the elevations should reject the equal elevation hypothesis. *controlIndex* is the zero-based index of the Y wave representing the control (X waves do not count in the index specification). The test compares the elevation of every Y wave with the specified control.

Output is to the M_DunnettMCElevations wave in the current data folder or optionally to a table. For every Y wave and control Y wave combination, the results include SE, q, q' (shown as qp), and the conclusion with 1 to accept the hypothesis of equal elevations or 0 to reject it. Use /TAIL to determine the critical value and the sense of the test. If you use /TUK you will also get the Tukey test for the set of elevations.

/MYVW={*xWave*, *yWave*}

Specifies that the input consists of multiple Y values for each X value. It ignores all other inputs and the results are appropriate only for multiple Y values at each X point.

yWave is a 2D wave of values arranged in columns. Use NaNs for padding where rows do not have the same number of entries as others. It will use the X scaling of *yWave* when *xWave* is null, `/MYVW={*,`*`yWave`*`}`.

It first tests the hypothesis (H_0) that the population regression is linear in an analysis of variance calculation. It generates results 1-7 (see Details) as well as: Among Groups SS, Among Groups DF, Within Groups SS, Within Groups DF, Deviations from Linearity SS, Deviations from Linearity DF, F statistic defined by the ratio of Deviation from Linearity MS to Within Groups MS, and the critical value Fc.

Next, it tests the hypothesis that the slope beta=0. If the original H_0 was accepted, the new F statistic=regressionMS/residualMS. Otherwise the with the critical F=regressionMS/WithinGroupsMS with a corresponding critical value. Finally, it reports the values of the coefficient of determination r2 and the standard error of the estimate S_{YX}.

/PAIR Specifies that the input waves are XY pairs, where each pair must be an X wave followed by a Y wave.

/Q No results printed in the history area.

/RTO Reflects the regression through the origin.

/T=*k* Displays results in a table. *k* specifies the table behavior when it is closed.

k=0: Normal with dialog (default).

k=1: Kills with no dialog.

k=2: Disables killing.

/TAIL=*tCode* Sets the sense of the test when applying Dunnett's test (see /DET). *tCode* is 1 or 2 for a one-tail critical value and 4 for a two-tail critical value.

/TUK Performs a Tukey-type test on multiple regressions on two or more Y waves. There are two possible Tukey-type tests: The first is performed if the hypothesis of equal slopes is rejected. It compares all combinations of two Y waves to identify if some of the waves have equal slopes. Output is to the M_TukeyMCSlopes wave in the current data folder or optionally to a table. For every Y wave pair, the results include the difference between slopes (absolute value), q, the critical value qc, and the conclusion set to 1 for accepting the equality of the pair of slopes or 0 for rejecting the hypothesis.

The second Tukey-type test is performed if all the slopes are the same but the elevations are not. The test (see /DET) compares all possible pairs of elevations to determine which satisfy the hypothesis of equality. Output is to the M_TukeyMCElevations wave in the current data folder.

/WSTR=*waveListString*

Specifies a string containing a semicolon-separated list of waves that contain sample data. Use *waveListString* instead of listing each wave after the flags.

/Z Ignores errors.

Details

Inputs may consist of Y waves or XY wave pairs. If X data are not used, the X values are inferred from the Y wave scaling. For multiple waves where only some have pairs, use the /PAIR flag and enter * in each place where the X values should be computed.

For each input StatsLinearRegression calculates:

1. Least squares regression line y=a+b*x.
2. Mean value of X: xBar.
3. Mean value of Y: yBar.
4. Sum of the squares (x_i-xBar)^2.
5. Sum of the squares (y_i-yBar)^2.
6. Sum of the product ($x_i y_i$-xyBar).
7. Standard error of the estimate $S_{YX}^2 = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}$.
8. F statistic for the hypothesis beta=0.
9. Critical F value Fc.
10. Coefficient of determination r2.
11. Standard error of the regression coefficient S_b.
12. t-statistic for the hypothesis beta=*beta0*, NaN if /B is not specified.
13. Critical value tc for the t-statistic above (used to calculate L1 and L2).
14. Lower confidence interval boundary (L1) for the regression coefficient.
15. Upper confidence interval boundary (L2) for the regression coefficient.

For two Y waves with the same slope, it computes a common slope (bc) and then tests the equality of the elevations (a). In both cases it computes a t-statistic and compares it with a critical value. If the elevations are also the same then it computes the common elevation (ac) and the pooled means of X and Y in (xp) and (yp).

For more than two Y waves it computes:

$$A_c = \sum_{j=1}^{W} A_j; \qquad A_j \equiv \sum x_i^2 = \sum_{i=0}^{n_j-1} X_i^2 - \frac{1}{n_j}\left(\sum_{i=0}^{n_j-1} X_i\right)^2$$

$$B_c = \sum_{j=1}^{W} B_j; \qquad B_j \equiv \sum xy = \sum_{i=0}^{n_j-1} XY - \frac{1}{n_j}\left(\sum_{i=0}^{n_j-1} X_i\right)\left(\sum_{i=0}^{n_j-1} Y_i\right)$$

$$C_c = \sum_{j=1}^{W} C_j; \qquad C_j \equiv \sum y^2 = \sum_{i=0}^{n_j-1} Y_i^2 - \frac{1}{n_j}\left(\sum_{i=0}^{n_j-1} Y_i\right)^2$$

$$SSp = \sum_{j=1}^{W} C_j - \frac{B_j^2}{A_j}$$

$$SSc = Cc - \frac{B_c^2}{A_c^2}$$

$$SSt = \sum_{j=1}^{W}\sum_{i=0}^{n_j} Y_{ji}^2 - \frac{1}{N}\left(\sum_{j=1}^{W}\sum_{i=0}^{n_j} Y_{ji}\right)^2 - \frac{\left(\sum_{j=1}^{W}\sum_{i=0}^{n_j} X_{ji}Y_{ji} - \frac{1}{N}\left(\sum_{j=1}^{W}\sum_{i=0}^{n_j} X_{ji}\right)\left(\sum_{j=1}^{W}\sum_{i=0}^{n_j} Y_{ji}\right)\right)^2}{\sum_{j=1}^{W}\sum_{i=0}^{n_j} X_{ji}^2 - \frac{1}{N}\left(\sum_{j=1}^{W}\sum_{i=0}^{n_j} X_{ji}\right)^2}$$