NDK_PCR_PRFTest

Last Modified on 03/14/2016 11:39 am CDT

- C/C++
- .Net

```
int __stdcall NDK_PCR_PRFTest ( double **  X,
                                size_t     nXSize,
                                size_t     nXVars,
                                double *   Y,
                                size_t     nYSize,
                                double     intercept,
                                LPBYTE     mask1,
                                size_t     nMaskLen1,
                                LPBYTE     mask2,
                                size_t     nMaskLen2,
                                double     alpha,
                                WORD       nRetType,
                                double *   retVal
                              )
```

Returns an array of cells for the i-th principal component (or residuals).

Returns

status code of the operation

Return values

NDK_SUCCESS	Operation successful
NDK_FAILED	Operation unsuccessful. See **Macros** for full list.

Parameters

[in]	**X**	is the independent variables data matrix, such that each column represents one variable
[in]	**nXSize**	is the number of observations (i.e. rows) in X
[in]	**nXVars**	is the number of variables (i.e. columns) in X
[in]	**Y**	is the response or the dependent variable data array (one dimensional array)
[in]	**nYSize**	is the number of elements in Y
[in]	**intercept**	is the constant or the intercept value to fix (e.g. zero). If missing (NaN), an intercept will not be fixed and is computed normally

[in]	**mask1**	is the boolean array to select a subset of the input variables in X. If missing (i.e. NULL), all variables in X are included.
[in]	**nMaskLen1**	is the number of elements in mask1
[in]	**mask2**	is the boolean array to select a subset of the input variables in X. If missing (i.e. NULL), all variables in X are included.
[in]	**nMaskLen2**	is the number of elements in mask2
[in]	**alpha**	is the statistical significance of the test (i.e. alpha)
[in]	**nRetType**	is a switch to select the return output (1 = P-Value (default), 2 = Test Stats, 3 = Critical Value.)
[out]	**retVal**	is the calculated test statistics/

Remarks

1. The underlying model is described **here**.
2. Model 1 must be a sub-model of Model 2. In other words, all variables included in Model 1 must be included in Model 2.
3. The coefficient of determination (i.e. R^2) increases in value as we add variables to the regression model, but we often wish to test whether the improvement in R square by adding those variables is statistically significant.
4. To do so, we developed an inclusion/exclusion test for those variables. First, let's start with a regression model with K_1 variables: $$Y_t = \alpha + \beta_1 \times X_1 + \cdots + \beta_{K_1} \times X_{K_1}$$ Now, let's add few more variables $\left(X_{K_1+1} \cdots X_{K_2}\right):$ $$Y_t = \alpha + \beta_1 \times X_1 + \cdots + \beta_{K_1} \times X_{K_1} + \cdots + \beta_{K_1+1} \times X_{K_1+1} + \cdots + \beta_{K_2} \times X_{K_2}$$
5. The test of hypothesis is as follows: $$H_o : \beta_{K_1+1} = \beta_{K_1+2} = \cdots = beta_{K_2} = 0$$ $$H_1 : \exists \beta_{i} \neq 0, i \in \left[K_1+1 \cdots K_2\right]$$
6. Using the change in the coefficient of determination (i.e. R^2) as we added new variables, we can calculate the test statistics: $$\mathrm{f}=\frac{(R^2_{f}-R^2_{r})/(K_2-K_1)}{(1-R^2_f)/(N-K_2-1)}\sim \mathrm{F}_{K_2-K_1,N-K2-1}$$ Where:
   - R^2_f is the R^2 of the full model (with added variables).
   - R^2_r is the R^2 of the reduced model (without the added variables).
   - K_1 is the number of variables in the reduced model.
   - K_2 is the number of variables in the full model.
   - N is the number of observations in the sample data.
7. The sample data may include missing values.
8. Each column in the input matrix corresponds to a separate variable.
9. Each row in the input matrix corresponds to an observation.
10. Observations (i.e. row) with missing values in X or Y are removed.
11. The number of rows of the response variable (Y) must be equal to the number of rows of the explanatory variables (X).
12. The MLR_ANOVA function is available starting with version 1.60 APACHE.

Requirements

```
Header
SFSDK.H

Library
SFSDK.LIB

DLL
SFSDK.DLL
```

References

Hamilton, J .D.; Time Series Analysis , Princeton University Press (1994), ISBN 0-691-04289-6

Tsay, Ruey S.; Analysis of Financial Time Series John Wiley & SONS. (2005), ISBN 0-471-690740

See Also

[template("related")]