

PCA Bootstrap Confidence Intervals Program

Steve Tonsor
Department of Biological Sciences
University of Pittsburgh
Pittsburgh PA 15260
Voice: 412-624-5491
Fax: 412-624-5863
tonsor@pitt.edu

Contact Steve Tonsor with any questions regarding the rather cumbersome output. **Please acknowledge me if you use this or any derivation of it.

This program bootstraps relationships among traits by randomly sampling the actual observations in a data set with replacement. For a data set with n observations, n random selections are made, each observation independently chosen from among the n actual observations. The correlation matrix and principle component eigenvectors and eigenvalues are produced by PROC PRINCOMP. This routine is repeated n times (the user sets this value), with the values output to a ASCII disk file (You can use this file in your own SAS program to make confidence intervals of any width you choose. However, you can also let the macro produce 95% confidence intervals.) BOOTPCA then uses proc univariate to produce bootstrapped-based 95% confidence intervals on the estimates of pc eigenvalues, trait correlations, and trait eigenvector coefficients, which are then printed to a SAS output window. Note that in the output window, for k traits, the traits will be named 'trait1' through 'traitk', and m treatments will be named simply 1 - m . The eigenvalue confidence intervals will be listed in m lines, one per treatment. The k eigenvalues will be listed as 'trait1' through 'traitk', although they are actually eigenvalue 1 through eigenvalue k .

CAUTION SHOULD BE EXERCISED IN INTERPRETING THE CONFIDENCE INTERVALS ON THE EIGENVECTOR COEFFICIENTS WHEN THE MAGNITUDE OF THE EIGENVALUE IS RELATIVELY SMALL, EVEN IF SIGNIFICANT. WHEN EIGENVALUES FOR TWO PCs ARE CLOSE IN VALUE, THEY CAN SWITCH IN MAGNITUDE, HENCE RANK, BETWEEN RESAMPLE ITERATIONS. THE PC IS ALSO PRONE TO REFLECTION. THE CONFIDENCE INTERVALS ON THE COEFFICIENTS CAN BECOME VERY LARGE AND PRONE TO TYPE II ERROR.

SEE Perez-Neto et al. 2003 Ecology 84(9):2347-2363 FOR A DISCUSSION OF THIS ISSUE.

Contact Steve Tonsor with any questions regarding the rather cumbersome output.
tonsor@pitt.edu
412-624-5491