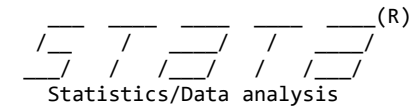

List of Nick's Toolbox Add-On Programs

Programs created in Stata Versions 11, 12, 13, 14, 15, 16
Nicholas J. Jackson, PhD MPH *04/01/2022*

Accelerated Longitudinal Designs

aldesign illustrates the impact of design parameters for an accelerated longitudinal design.

aldcost cost computation of ALDs.

aldsim simulation of ALDs.

Statistical Models

Cross-Sectional Analysis:

descdata Overall summary statistics and between group statistics (new program to replace *grpscompare*)

grpscompare Legacy Program for Overall summary statistics and between group statistics. Use *descdata* instead

glmreport Statistical GLM reporting program for main effects

nbregreport Statistical Negative Binomial Regression reporting program for main effects

regreport Statistical Linear Regression reporting program for main effects

logitreport Statistical Logistic Regression reporting program for main effects

polyfind Non-linear (Polynomial) regression finder. Identifies non-linear patterns in the data.

Longitudinal Analysis:

pairtest conducts paired ttest and wilcoxon sign rank on longitudinal designs

geereport Statistical GEE reporting program for main effects in longitudinal designs (Generalized Estimating Equations)

mixednbregreport Statistical Mixed Effects Negative Binomial Regression reporting program for main effects in longitudinal designs

mixedregreport Statistical Mixed Effects Linear Regression reporting program for main effects in longitudinal designs (Restricted Maximum Likelihood)

mixedlogitreport Statistical Mixed Effects Logistic Regression reporting program for main effects in longitudinal designs (Maximum Likelihood)

Survival Analysis:

survreport NEEDS HELPFILE Cox Regression analysis reporting program.

Correlation Analyses

corrreport Correlation table program

pcorrreport Partial Correlation table program

r2comp Bootstrapped R-Squared Comparison Program

pbcorr Point Biserial Correlation (X is Binary, Y continuous)

phicorr Phi Correlation Coefficient (X and Y are Binary)

Confounding (Mediation/Supression) & Moderation (Interaction)

>

Cross-Sectional Analysis:

confound Examines potential confounders in models. **NOTE:** See *estconf* or *bestconf* for more advanced confounding analysis.

estconf Estimates of Confounding. Conducts analysis of confounding through examining change in the direct effect.

bestconf Bootstrap Estimates of Confounding. Conducts analysis of confounding through examining bootstrap change in the direct effect (same as indirect effect in mediation).

mediation Conducts mediational (Positive confounding) analysis using Baron and Kenny steps, Sobel Test, and Bootstrap test of indirect effects. **NOTE:** See *estconf* or *bestconf* for additiona n analysis.

glmint Statistical exploration program for Two-Way interactions using a GLM

nbregint Statistical exploration program for Two-Way interactions using Negative Binomial Regression

regint Statistical exploration program for Two-Way interactions using Linear Regression

logitint Statistical exploration program for Two-Way interactions using Logistic Regression

Longitudinal Analysis:

geeint Statistical exploration program for Two-Way interactions in longitudinal designs (Generalized Estimating Equations)

mixedconf Conducts analysis of confounding for Mixed Effect and Multilevel Models.

mixednbregint Statistical exploration program for Two-Way interactions in longitudinal designs (Mixed Effects Negative Binomial Regression)

mixedregint Statistical exploration program for Two-Way interactions in longitudinal designs (Mixed Effects Linear Regression)

mixedlogitint Statistical exploration program for Two-Way interactions in longitudinal designs (Mixed Effects Logistic Regression)

Survival Analysis:

survint NEEDS HELPFILE Two-Way and Three-Way interactions in Cox PH Regression models.

Interaction Graphs

>

Cross-Sectional Analysis:

cinteract Two-Way Continuous X Continuous interaction exploration program for 2D Graphs in Stata

cinteract3d Two-Way Continuous X Continuous interaction extraction program for 3D Graphs in SigmaPlot

gginteract Two-Way Categorical X Categorical interaction exploration program for 2D Graphs in Stata

gcinteract Two-Way Categorical X Continuous interaction exploration program for 2D Graphs in Stata

Longitudinal Analysis:

cgeeinteract3d Generalized Estimating Equation Two-Way Continuous X Continuous interaction extraction program for 3D Graphs in SigmaPlot

Other Graphs

>

distdesc Summary statistics with graphical exploration of distributions

forest Forestplot graphs for Odds Ratios etc.

scatteradjust Used for creating scatter plots and linear predictions of 2 vars adjusted for covariates.

tabgraph Creates a frequency graph of an variable accross various other categorical vars (a visual cross tabs)

violin NEEDS HELPFILE Violin Plot.

Reciever Operating Curves and Model Discrimination Characteristics

>

nri Computes Net Reclassification Index (NRI) and Integrated Discrimination Improvement (IDI) between two predictive models

rocci NEEDS HELPFILE Creates percentile bootstrap confidence intervals for Sensitivity vs 1-Specificity ROC curve

rocop NEEDS HELPFILE (DONT USE OUT OF DATE-SEE **youden**) Creates standard ROC graph and computes Operating Point for the ROC

rocauc NEEDS HELPFILE Bootstrap Area Under the Curve (AUC) for ROC

specbias Computes Sensitivity, Specificity, LR+, LR- for predicted probabilities from a model. Also computes Spectrum Bias.

youden NEEDS HELPFILE Computes Youden's Index as the Operating Point

P Value Correction

>

benhoch Benjamini and Hochberg P Value correction

bholm Holm-Bonferonni P Value correction

bonfer Bonferonni P Value correction

roundp P Value rounding program

Goodness of Fit

>

semgof Goodness of fit statitics table for SEM postestimoation (RMSEA, TLI, CFI, etc.)

Concentration and Achievement Indicies

inequity Generate Concentration and Achievement Indicies with or without Covriate Adjustment

Data and Variable Manipulation

center Centers variables (subtracts mean)

charid Identifies the characters present in numeric variables for destring conversion

colprod Takes the product of values in a Column/Variable

coldup Finds duplicate observations between subjects

dotcounter Visually reports the progress of a loop

dropmiss Drops missing values

efcode Creates Effect Coded variables

logtran Log Transforms variables

lowercase Renames Variable Names to lowercase

medsplit Creates median split of a variable

moreobs Creates additional observations in the dataset

outliers Conducts basic outlier analysis -OUTDATED. See **outreport** instead

outreport Conducts outlier analysis using **outmad** , **outbox**, or **outsd** programs.

roundp P Value rounding program

spaceaft NEEDS HELPFILE Creates spacing to control output fortmating

quantnorm Conducts Quantile normalization for gene expression data.

rowdup Finds duplicate observations within subjects

standard Standardizes variables (subtracts mean, divides by standard deviation). Can also be used to set mean/sd for a variable.

strstrip Strips strings of values (advanced Substring/Destring Commands)

winsorize Winsorize variables

Wilcox Robust Methods Functions

>

bivar NEEDS HELPFILE Biweight Midvariance

bootse NEEDS HELPFILE Bootstrap Standard Errors

mad NEEDS HELPFILE Median Absolute Deviation (MAD)

medci NEEDS HELPFILE Confidence Interval and SE for Medians

onestep NEEDS HELPFILE One-Step M-Estimator

outbox Outlier Detection using Boxplot method. Also see **outreport**

outmad Outlier Detection using MAD-Median. Also see **outreport**

outsd Outlier Detection using Mean +/- x*SD rule. Also see **outreport**

pbvar NEEDS HELPFILE Percentage Bend Midvariance

twmean Trimmed and Winsorized Means

yuent Trimmed Means Two Sample T-test (Yuen's Method)

Data Extraction

exout Exportation of results to excel tables. Used to make Stata's built in Excel Exportation easier. Stata Version 12 Required.

resout extracts results from a model and places in a table

miresout extracts results from a multiply imputed model and places in a table

Data Simulation

>

choleskycorr Creates correlated random normal variables

coinflip Simulates a the filp of a coin

acesim Creates simulated data for Univariate ACE model

dieroll Simulates a die roll

distgen Generate Skewed/Heavy Tailed data based on G & H distribution (Contaminated Normal)

distmixed Generate Two-Component Mixture Distribution (Mixed Normal)

rtnormal Generate Truncated Random Normal Distribution

Behavioural Genetics

acemodel runs standard behavorial genetic model providing estimates of A, C, and E.

uniace Runs a series of nested ACE models for conducting a univariate analysis of Twin Data from the USC Twins Lab (a more specific version of the acemodel program above)

acesim Creates simulated data for Univariate ACE model

PsychoPhys Signal Analysis Programs

fftanalysis Fast Fourier Transform (FFT) for spectral EEG data analysis

psacompile Power Spectral Analysis Compilation Program for Spectral Analysis files created in Sandman. See Word Doc "UPENN PSA Compilation Instructions" for Details NOTE: Special program only to be used with UPENN PSA Data

siggen Sinusoid Signal Generation program.

Power/Sample Size Analysis

>

meanequiv NEEDS HELPFILE Power for tests of equivalence between two groups based on means using simulations.