

eefAnalytics

February 03, 2021

Package: eefAnalytics

Version: 1.0.9

Date: 2021-2-03

Title: Robust Analytical Methods for Evaluating Educational Interventions using Randomised Controlled Trials Designs

Description: Analysing data from evaluations of educational interventions using randomised controlled trial designs. Various analytical tools to perform sensitivity analysis using different methods are supported (e.g. frequentist models with bootstrapping and permutations options, Bayesian models). The included functions can be used for simple randomised trials, cluster randomised trials and multisite trials. The methods can also be used more widely beyond education trials. Users can use these functions to evaluate other interventions designs using Frequentist and Bayesian multilevel models.

Author: Dimitris Vallis, Akansha Singh, Germaine Uwimpuhwe, Steve Higgins, ZhiMin Xiao, Ewoud De Troyer and Adetayo Kasim

Maintainer: Dimitris Vallis <dimitris.vallis@durham.ac.uk>

URL: <https://github.com/dimitris-90/eefAnalytics>

BugReports: <https://github.com/dimitris-90/eefAnalytics/issues>

STATA topics documented:

crtbayes.....	2
crtData.....	4
crtfreq.....	4
mstbayes.....	6
mstData.....	8
mstfreq.....	9
srtbayes.....	11
srtfreq.....	12

crtbayes: Bayesian analysis of cluster randomised education trials using non-informative Priors.

Description

crtbayes performs analysis of cluster randomised education trials using a multilevel model under a Bayesian setting, assuming non-informative¹ priors.

Usage

```
crtbayes varlist, int(intervention) ran(random) [thr(#) sepch diag noi save *]  
]
```

Arguments

Argument	Description
intervention()	A factor variable specifying the “intervention variable” as appearing in the formula and the data.
random()	A factor variable specifying the “clustering variable” as contained in the data.
threshold()	A scalar or vector of pre-specified threshold(s) for estimating Bayesian posterior probability such that the observed effect size is greater than or equal to the threshold(s).

¹ The Stata default priors are relatively non-informative for moderately scaled model parameters but may become informative for parameters with larger values (See [Stata manual](#)).

sepchains	Stores summary statistics for each chain.
diagnostics	Generates convergence diagnostic graphs for all chains.
noisily	Displays regression output for conditional and unconditional models.
save	Saves two datasets (mcmcUncCRT.dta, mcmcCondCRT.dta) containing the simulation output for the conditional and unconditional models.
*	Additional arguments to be passed to the command such as mcmcsz(#), burnin(#), rseed(#), nchains(#), and custom priors. Stata defaults of Bayesian mixed models apply.

Value

Stored output; a list consisting of

- **r(Beta)**: Estimates and credible intervals for variables specified in the model.
- **r(Uncond_ES#)**: Unconditional Hedges' g effect size and its 95% credible intervals for arm #.
- **r(Cond_ES#)**: Conditional Hedges' g effect size and its 95% credible intervals for arm #.
- **r(Cov)**: A matrix of variance decomposition into between cluster variance (Schools), within cluster variance (Pupils) and Total variance from conditional and unconditional model (model with only the intercept as a fixed effect). It also contains intra-cluster correlation (ICC).
- **r(SchEffects)**: A vector of the estimated deviation of each school from the intercept.
- **r(Cond_ProbES#)**: A matrix of Conditional Bayesian Posterior Probabilities such that the observed effect size is greater than or equal to a pre-specified threshold(s) for arm #.
- **r(Uncond_ProbES#)**: A matrix of Unconditional Bayesian Posterior Probabilities such that the observed effect size is greater than or equal to a pre-specified threshold(s) for arm #.
- **r(sepchains_#)**: Stores summary statistics for # number of chains separately.

Unconditional outputs within Uncond_ES, Cov and Uncond_ProbES are based on between cluster (if appropriate), within cluster and total variance from the unconditional model (model with only the intercept as a fixed effect).

Examples

Simple model:

- `crtbayes Posttest Prettest, int(Intervention) ran(School)`

Model using custom simulation options and all diagnostic options:

- `crtbayes Posttest Prettest, int(Intervention) ran(School) thr(0.1) mcmcsize(50000) burnin(50000) rseed(1234) nchains(4) sepch diag save`

Model using custom simulation options with three-arm intervention variable and custom priors:

- `crtbayes Posttest Prettest, int(Intervention2) mcmcsize(50000) burnin(50000) rseed(1234) nchains(4) prior({Posttest:_cons}, uniform(-50,50))`

crtData: Cluster Randomised Trial Data.

Description

A cluster randomised trial dataset containing 22 schools. The data contains a random sample of test data of pupils and not actual trial data.

Format

A data frame with 265 rows and 5 variables

Details

- Posttest: posttest scores
- Prettest: prettest scores
- Intervention: the indicator for intervention groups in a two-arm trial, coded as 1 for intervention group and 0 for control group.
- Intervention2: a simulated indicator for intervention groups in a three-arm trial.
- School: numeric school identifier

crtfreq: Analysis of Cluster Randomised Education Trials using Multilevel Model under a Frequentist Setting.

Description

`crtfreq` performs analysis of cluster randomised education trials using a multilevel model under a frequentist setting.

Usage

```
crtfreq varlist, int(intervention) ran(random) [np(#) nb(#) seed(#) noi show m
l *]
```

Arguments

Argument	Description
intervention()	A factor variable specifying the “intervention variable” as contained in the data.
random()	A factor variable specifying the “clustering variable” as contained in the data.
nperm()	A number of permutations required to generate a permuted p-value. Default is NULL.
nboot()	A number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if nperm or nboot are specified.
showprogress	Displays progress of permutations/bootstraps using dots.
ml	Fits model via maximum likelihood. Default is RMLE.
*	Additional options that allow the user to configure the maximum likelihood estimation such as <code>technique()</code> , <code>difficult</code> , <code>iterate()</code> . Stata defaults apply.

Value

Stored output; a list consisting of:

- **r(Beta)**: Estimates and confidence intervals for variables specified in the model.
- **r(Uncond_ES#)**: Unconditional Hedges’ g effect size and its 95% confidence intervals for arm #.
- **r(Cond_ES#)**: Conditional Hedges’ g effect size and its 95% confidence intervals for arm #. If nboot is specified, CIs are replaced with bootstrapped CIs.
- **r(Cov)**: A matrix of variance decomposition into between cluster variance, within cluster variance (Pupils) and Total variance from conditional and unconditional models (model with only the intercept as a fixed effect). It also contains intra-cluster correlation (ICC).
- **r(SchEffects)**: A vector of the estimated deviation of each school from the intercept and intervention slope.

Unconditional outputs within `Uncond_ES` and `Cov` are based on between cluster (if appropriate), within cluster and total variance from the unconditional model (model with only the intercept as a fixed effect)

- List of generated variables attached to existing dataset when nperm and/or nboot is specified (I# denotes number of interventions and C/Unc denotes Conditional and

Unconditional estimates) containing the bootstrapped/permutated effect sizes using residual variance (Within) indicated by suffix “_W” and total variance (Total) indicated by suffix “_T”:

- BootC_I#_W , BootC_I#_T , BootUnc_I#_W , BootUnc_I#_T
- PermC_I#_W , PermC_I#_T , PermUnc_I#_W , PermUnc_I#_T

Existing variables generated from previous use are replaced.

Examples

Simple model:

- `crtfreq Posttest Prettest, int(Intervention) ran(School)`

Model using permutations including factor parameters and additional maximization options:

- `crtfreq Posttest Prettest, int(Intervention) ran(School) nb(3000) technique(dfp) difficult`

Model using permutations and bootstraps with three-arm intervention variable and maximum likelihood estimation:

- `crtfreq Posttest Prettest, int(Intervention) ran(School) np(3000) nb(2000) ml show`

mstbayes: Bayesian analysis of Multisite Randomised Education Trials using non-informative Priors.

Description

mstbayes performs analysis of multisite randomised education trials using a multilevel model under a Bayesian setting assuming non-informative priors.

Usage

```
mstbayes varlist, int(intervention) ran(random) [thr(#) sepch diag noi save *]  
]
```

Arguments

Argument	Description
----------	-------------

intervention()	A factor variable specifying the “intervention variable” as contained in the data.
random()	A factor variable specifying the “clustering variable” as contained in the data.
threshold()	A scalar or vector of pre-specified threshold(s) for estimating Bayesian posterior probability such that the observed effect size is greater than or equal to the threshold(s).
sepchains	Stores summary statistics for each chain.
diagnostics	Generates convergence diagnostic graphs for all chains.
noisily	Displays regression output for conditional and unconditional models.
save	Saves two datasets (mcmcUncMST.dta, mcmcCondMST.dta) containing the simulation output for the conditional and unconditional models.
*	Additional Bayesian arguments to be passed to the command such as mcmcsize(#) burnin(#) rseed(#) nchains(#) and custom priors. Stata defaults of Bayesian mixed models apply.

Value

Stored output; a list consisting of

- **r(Beta)** : Estimates and credible intervals for variables specified in the model.
- **r(Uncond_ES#)** : Unconditional Hedges’ g effect size and its 95% credible intervals for arm #.
- **r(Cond_ES#)** : Conditional Hedges’ g effect size and its 95% credible intervals for arm #.
- **r(Cov)** : A list of variance decomposition into within cluster variance (Pupils) and Total variance from conditional and unconditional models (model with only the intercept as a fixed effect). It also contains intra-cluster correlation (ICC).
- **r(schCov)** : Variance decomposition into between cluster variance-covariance matrix (school by intervention).
- **r(UschCov)** : Variance decomposition for the Unconditional model into between cluster variance (School).
- **r(SchEffects)** : A matrix of the estimated deviation of each school from the intercept.
- **r(Uncond_ProbES#)** : A matrix of Unconditional Bayesian Posterior Probabilities such that the observed effect size is greater than or equal to a pre-specified threshold(s) for arm #.

- `r(Cond_ProbES#)` : A matrix of Conditional Bayesian Posterior Probabilities such that the observed effect size is greater than or equal to a pre-specified threshold(s) for arm #.
- `r(sepchains_#)` : Stores summary statistics for # number of chains separately.

Unconditional outputs within `Uncond_ES` , `Cov`, `UschCov` and `Uncond_ProbES`) are based on between cluster (if appropriate), within cluster and total variance from the unconditional model (model with only the intercept as a fixed effect).

Examples

Simple model:

- `mstbayes Posttest Prettest, int(Intervention) ran(School)`

Model using custom simulation options and all diagnostic options:

- `mstbayes Posttest Prettest, int(Intervention) ran(School) thr(0.1) mcmcsize(50000) burnin(50000) rseed(1234) nchains(4) sepch diag save`

Model using custom simulation options with three-arm intervention variable and custom priors:

- `mstbayes Posttest Prettest, int(Intervention2) ran(School) mcmcsize(50000) burnin(50000) rseed(1234) nchains(4) prior({Posttest:_cons}, uniform(-50,50))`

mstData: Multisite Trial Data.

Description

A multisite trial dataset containing 54 schools. This data contains a random sample of test data of pupils and not actual trial data.

Format

A data frame with 210 rows and 5 variables

Details

- `Posttest`: posttest scores
- `Prettest`: prettest scores
- `Intervention`: the indicator for the intervention groups in a two-arm trial, coded as 1 for intervention group and 0 for control group.
- `Intervention2`: a simulated indicator for intervention groups in a three-arm trial.

- School: numeric school identifier

mstfreq: Analysis of Multisite Randomised Education Trials using Multilevel Model under a Frequentist Setting.

Description

mstfreq performs analysis of multisite randomised education trials using a multilevel model under a frequentist setting.

Usage

```
mstfreq varlist, int(intervention) ran(random) [np(#) nb(#) seed(#) noi show *]
```

Arguments

Argument	Description
intervention()	A factor variable specifying the “intervention variable” as contained in the data.
random()	A factor variable specifying the “clustering variable” as contained in the data.
nperm()	A number of permutations required to generate a permuted p-value. Default is NULL.
nboot()	A number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if nperm or nboot are specified.
showprogress	Displays progress of permutations/bootstraps using dots.
ml	Fit model via maximum likelihood. Default is RMLE.
*	Additional options that allow the user to configure the maximum likelihood estimation such as <code>technique()</code> , <code>difficult</code> , <code>iterate()</code> ² . Stata defaults apply.

Value

Stored output; a list consisting of

- **r(Beta)**: Estimates and confidence intervals for variables specified in the model.

² Default iterations are set to 100.

- `r(Uncond_ES#)`: Unconditional Hedges' g effect size and its 95% confidence intervals for arm #.
- `r(Cond_ES#)`: Conditional Hedges' g effect size and its 95% confidence intervals for arm #. If `nboot` is specified, CIs are replaced with bootstrapped CIs.
- `r(Cov)`: A list of variance decomposition into within cluster variance (Pupils) and Total variance from conditional and unconditional models (model with only the intercept as a fixed effect). It also contains intra-cluster correlation (ICC).
- `r(schCov)`: Variance decomposition into between cluster variance-covariance matrix (school by intervention).
- `r(UschCov)`: Variance decomposition for the Unconditional model into between cluster variance (School).
- `r(SchEffects)`: A vector of the estimated deviation of each school from the intercept and intervention slope.
- List of generated variables attached to existing dataset when `nperm` and/or `nboot` is specified (I# denotes number of interventions and C/Unc denotes Conditional and Unconditional estimates) containing the bootstrapped/permutated effect sizes using residual variance (Within) indicated by suffix "_W" and total variance (Total) indicated by suffix "_T":
 - `BootC_I#_W`, `BootC_I#_T`, `BootUnc_I#_W`, `BootUnc_I#_T`
 - `PermC_I#_W`, `PermC_I#_T`, `PermUnc_I#_W`, `PermUnc_I#_T`

Note: # number of supplementary permutations/bootstraps are automatically employed when # number of permutated/bootstrapped models have failed to converge.

Existing variables generated from previous use are replaced.

Unconditional outputs within `Uncond_ES`, `Cov` and `UschCov` are based on between cluster (if appropriate), within cluster and total variance from the unconditional model (model with only the intercept as a fixed effect).

Examples

Simple model:

- `mstfreq Posttest Prettest, int(Intervention) ran(School)`

Model using permutations including factor parameters and additional maximization options:

- `mstfreq Posttest Prettest, int(Intervention) ran(School) nb(3000) technique(dfp) difficult`

Model using permutations and bootstraps with three-arm intervention variable and maximum likelihood estimation:

- `mstfreq Posttest Prettest, int(Intervention) ran(School) np(3000) nb(2000) ml show`

srtbayes: Analysis of Simple Randomised Education Trials using Bayesian Linear Regression Model with non-informative Priors.

Description

srtbayes performs analysis of educational trials under the assumption of independent errors among pupils using Bayesian framework. This can also be used with schools as fixed effect.

Usage

```
srtbayes varlist, int(intervention) [thr(#) sepch diag noi save *]
```

Arguments

Argument	Description
intervention()	A factor variable specifying the “intervention variable” as contained in the data.
threshold()	A scalar or vector of pre-specified threshold(s) for estimating Bayesian posterior probability such that the observed effect size is greater than or equal to the threshold(s).
sepchains	Stores summary statistics for each chain.
diagnostics	Generates convergence diagnostic graphs for all chains.
save	Saves two datasets (mcmcUnc.dta, mcmcCond.dta) containing the simulation output for the conditional and unconditional models.
noisily	Displays regression output for conditional and unconditional models.
*	Additional Bayesian arguments to be passed to the command such as mcmcsample(#) burnin(#) rseed(#) nchains(#) and custom priors. Stata defaults of Bayesian mixed models apply.

Value

Stored output; a list consisting of

- `r(Beta)`: Estimates and credible intervals for variables specified in the model.
- `r(Uncond_ES#)`: Unconditional Hedges’ g effect size and its 95% credible intervals for arm #.
- `r(Cond_ES#)`: Conditional Hedges’ g effect size and its 95% credible intervals for arm #.

- `r(Sigma2)`: Conditional and Unconditional residual variance.
- `r(ProbES#)`: A matrix of Conditional and Unconditional Bayesian Posterior Probabilities such that the observed effect size is greater than or equal to a pre-specified threshold(s) for arm #.
- `r(sepchains_#)`: Stores summary statistics for # number of chains separately.

Unconditional outputs within `Uncond_ES`, `Sigma2` and `ProbES` are based on residual variance from the unconditional model (model with only the intercept as a fixed effect).

Examples

Simple model:

- `srtbayes Posttest Prettest, int(Intervention)`

Model using custom simulation options and all diagnostic options:

- `srtbayes Posttest Prettest, int(Intervention) thr(0.1) mcmcsiz
e(50000) burnin(50000) rseed(1234) nchains(4) sepch diag save`

Model using custom simulation options with three-arm intervention variable, a vector of thresholds and custom priors:

- `srtbayes Posttest Prettest, int(Intervention2) thr(0.3 0.4 0.5) mcmcsiz
e(50000) burnin(50000) rseed(1234) nchains(4) normalprior(10) igammapri
or(1 2)`

srtfreq: Analysis of Simple Randomised Education Trial using Linear Regression Model.

Description

`srtfreq` performs analysis of educational trials under the assumption of independent errors among pupils. This can also be used with schools as fixed effect.

Usage

```
srtfreq varlist, int(intervention) [np(#) nb(#) seed(#) noi show *]
```

Arguments

Argument	Description
<code>intervention()</code>	A factor variable specifying the “intervention variable” as contained in the data.

nperm()	A number of permutations required to generate a permuted p-value. Default is NULL.
nboot()	A number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if nperm or nboot are specified.
showprogress	Displays progress of permutations/bootstraps using dots.

Value

Stored output; a list consisting of

- **r(Beta)**: Estimates and confidence intervals for variables specified in the model.
- **r(Uncond_ES#)**: Unconditional Hedges' g effect size and its 95% confidence intervals.
- **r(Cond_ES#)**: Conditional Hedges' g effect size and its 95% confidence intervals. If nboot is specified, CIs are replaced with bootstrapped CIs.
- **r(Sigma2)**: Conditional and Unconditional residual variance.
- List of generated variables attached to existing dataset when nperm and/or nboot is specified (I# denotes number of interventions and C/Unc denotes Conditional and Unconditional estimates):
 - PermC_I# PermUnc_I#
 - BootC_I# BootUnc_I#

Existing variables generated from previous use are replaced.

Unconditional outputs within Uncond_ES and sigma2 are based on residual variance from the unconditional model (model with only the intercept as a fixed effect)

Examples

Simple model:

- **srtfreq** Posttest Prettest, **int**(Intervention)

Model using permutations including schools as fixed effects:

- **srtfreq** Posttest Prettest i.School, **int**(Intervention) **np**(3000)

Model using permutations and bootstraps with three-arm intervention variable:

- **srtfreq** Posttest Prettest, **int**(Intervention2) **np**(3000) **nb**(2000)