eefAnalytics

Package: eefAnalytics
Version: 1.1.0
Date: 2022-2-25

February 25, 2022
Title: Robust Analytical Methods for Evaluating Educational Interventions using Randomised Contr olled Trial Designs
Description: Analysing data from evaluations of educational interventions using randomised contr olled trial designs. Various analytical tools to perform sensitivity analysis using different methods are supported (e.g. frequentist models with bootstrapping and permutations options, Bayesian mode ls). 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, Adetayo Kasim and Jochen Einbeck.
Maintenance and Bug Reports: Dimitris Vallis <dimitris.vallis@durham.ac.uk>

STATA topics documented:

crtbayes: Bayesian analysis of cluster randomised education trials.

Description

Performs analysis of cluster randomised education trials using a multilevel model under a Bayesian setting, assuming non-informative[1] priors.

Usage

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

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).
sepchains	Stores summary statistics for each chain.
diagnostics	Generates convergence diagnostic graphs for all chains.
noisily	Displays regression output for conditional and unconditional models.

[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).

save	Saves two datasets (mcmcUncCRT.dta, mcmcCondCRT.dta) containing the simulation output for the conditional and unconditional models.
cond()	Wrapper for arguments to be passed only to the conditional specification.
uncond()	Wrapper for arguments to be passed only to the unconditional specification.
*	Additional arguments to be passed to both specifications 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 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

Model using custom simulation options and all diagnostic options:

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

Model using custom simulation options, custom priors for the constant in both conditional and unconditional models and custom prior for the Intervention variable:

- **crtbayes** Posttest Prettest**, int(**Intervention**) mcmcsize(**50000**) burnin(**5 0000**) rseed(**1234**) nchains(**4**) prior({**Posttest:_cons**}, uniform(**-50,50**)) cond(prior({**Posttest:2.Intervention**}, normal(**0,100**)))**

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 Modelling under a Frequentist Setting.

Description

Performs analysis of cluster randomised education trials using a multilevel model under a frequentist setting.

Usage

```
crtfreq varlist, int(intervention) ran(random) [np(#) nb(#) case(#) res perc
basic seed(#) noi nodot ml *]
```

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()	Number of permutations required to generate a permutated p-value. Default is NULL.
nboot()	Number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
case()	Case bootstrap, specifying inside the bracket which level should be re-sampled. Numlist can either be (1) for re-sampling with replacement at subject level, (2) at school level or (1,2) at school and then subject level. Default is (1).
residual	Non-parametric Residual bootstrap. Default is case bootstrap.
percentile	Percentile bootstrap confidence interval.
basic	Basic (Hall's) bootstrap confidence interval. Default is percentile.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if npermor nboot are specified.
nodot	Suppresses display of dots indicating progress of permutations/bootstraps. Default is one dot character displayed for every block of 10 replications.
ml	Fits model via maximum likelihood. Default is RMLE.
paste	Attaches bootstrapped or permutated effect sizes to existing dataset if nperm (PermC/Unc_I#_W/T) or nboot (BootC/Unc_I#_W/T) has been specified (I# denotes number of interventions, C/Unc denotes Conditional and Unconditional estimates and W/T within and total effect sizes). Existing variables generated from previous use are replaced.
*	Additional options that allow the user to configure the maximum likelihood estimation such as technique(), difficult, iterate(). 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.
- `r(UncondPv)`: A matrix containing the unconditional two-sided within and total effect size permutation p-values (available if nperm(#) has been selected).
- `r(CondPv)`: A matrix containing the conditional two-sided within and total effect size permutation p-values (available if nperm(#) has been selected).

Note: When using permutations/bootstraps consider changing the number of maximum iterations using the command iterate(). Number of supplementary permutations/bootstraps are automatically deployed when # number of permutated/bootstrapped models have failed to converge.

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)

Examples

Simple model:

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

Model using permutations including factor parameters, additional maximization options and attaching bootrstrapped effect sizes to existing dataset:

- `crtfreq Posttest Prettest, int(Intervention) ran(School) nb(3000) iterate(100) paste`

Model using case bootstraps resampling both levels with maximum likelihood estimation and additional maximization options:

- `crtfreq Posttest Prettest, int(Intervention) ran(School) nb(2000) ml case(1,2) iterate(100)`

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

Description

Performs analysis of multisite randomised education trials using a multilevelmodel under a Bayesian setting assuming non-informative priors.

Usage

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

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.
cond()	Wrapper for arguments to be passed only to the conditional specification.
uncond()	Wrapper for arguments to be passed only to the unconditional specification.
*	Additional arguments to be passed to both specifications 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

Model using custom simulation options and all diagnostic options:

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

Model using custom simulation options, custom priors for the constant in both conditional and unconditional models and custom priors for each arm of the intervention variable:

- **mstbayes** Posttest Prettest**, int(**Intervention2**) mcmcsize(**50000**) burnin(**50000**) rseed(**1234**) nchains(**4**) prior({**Posttest:_cons**}, uniform(**-50,50**)) cond(prior({**Posttest:1.Intervention2**}, normal(**1.5,10**)) prior({**Posttest:2.Intervention2**}, normal(**0.5,5**)))**

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

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

Usage

```
mstfreq varlist, int(intervention) ran(random) [np(#) nb(#) case(#) res perc
basic seed(#) noi nodot *
```

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()	Number of permutations required to generate a permutated p-value. Default is NULL.
nboot()	Number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
case()	Case bootstrap, specifying inside the bracket which level should be re-sampled. Numlist can either be (1) for re-sampling with replacement at subject level, (2) at school level or (1,2) at school and then subject level. Default is (1).
residual	Non-parametric Residual bootstrap. Default is case bootstrap.
percentile	Percentile bootstrap confidence interval.
basic	Basic (Hall's) bootstrap confidence interval. Default is percentile.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if npermor nboot are specified.
nodot	Suppresses display of dots indicating progress of permutations/bootstraps. Default is one dot character displayed for every block of 10 replications.
ml	Fit model via maximum likelihood. Default is RMLE.

paste	Attaches bootstrapped or permutated effect sizes to existing dataset if nperm (PermC/Unc_I#_W/T) or nboot (BootC/Unc_I#_W/T) has been specified (I# denotes number of interventions, C/Unc denotes Conditional and Unconditional estimates and W/T within and total effect sizes). Existing variables generated from previous use are replaced.
*	Additional options that allow the user to configure the maximum likelihood estimation such as technique(),difficult, iterate(). 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 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 interceptand intervention slope.
- `r(UncondPv)`: A matrix containing the unconditional two-sided within and total effect size permutation p-values (available if nperm(#) has been selected).
- `r(CondPv)`: A matrix containing the conditional two-sided within and total effect size permutation p-values (available if nperm(#) has been selected).

Note: When using permutations/bootstraps consider changing the number of maximum iterations using the command iterate(). Number of supplementary permutations/bootstraps are automatically deployed when # number of permutated/bootstrapped models have failed to converge.

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, additional maximization options and attaching bootstrapped effect sizes on existing dataset:

- **mstfreq** Posttest Prettest, **int(**Intervention**) ran(**School**) nb(**3000**) iterate(**100**) paste**

Model using permutations with three-arm intervention variable and maximum likelihood estimation while hiding permutations progress:

- **mstfreq** Posttest Prettest, **int(**Intervention**) ran(**School**) np(**1000**) ml nodot iterate(**100**)**

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

Description

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 c() unc() *]
```

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.
noisily	Displays regression output for conditional and unconditional models.
save	Saves two datasets (mcmcUnc.dta, mcmcCond.dta) containing the simulation output for the conditional and unconditional models.
cond()	Wrapper for arguments to be passed only to the conditional specification.
uncond()	Wrapper for arguments to be passed only to the unconditional

specification.

*	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 forarm #.
- `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 ProbESare based on residual variancefrom the unconditional model (model with only the intercept as a fixed effect).

Examples

Model using custom simulation options and all diagnostic options:

```
srtbayes Posttest Prettest, int(Intervention) thr(0.1) mcmcsize(50000) b
urnin(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

Performs analysis of educational trials under the assumption of independent errorsamong pupils. This can also be used with schools as fixed effect.

Usage

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

Arguments

Argument	Description
intervention()	A factor variable specifying the "intervention variable" as contained in the data.
nperm()	Number of permutations required to generate a permutated p-value. Default is NULL.
nboot()	Number of bootstraps required to generate bootstrap confidence intervals. Default is NULL.
percentile	Percentile bootstrap confidence interval.
basic	Basic (Hall's) bootstrap confidence interval. Default is percentile.
seed()	Set seed. Default is 1020252.
noisily	Allow user to observe convergence of the conditional models if npermor nboot are specified.
nodot	Suppresses display of dots indicating progress of permutations/bootstraps. Default is one dot character displayed for every block of 10 replications.
paste	Attaches bootstrapped or permutated effect sizes to existing dataset if nperm (PermC/Unc_I#) or nboot (BootC/Unc_I#) has been specified (I# denotes number of interventions and C/Unc denotes Conditional and Unconditional estimates). Existing variables generated from previous use are replaced.

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.
- `r(Pv)`: Conditional and unconditional two-sided within and total effect size permutation p-values (available if nperm(#) has been selected).

Unconditional outputs within Uncod_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 bootstraps with three-arm intervention variable:

- `srtfreq Posttest Prettest, int(Intervention2) nb(2000)`