Package: factiv (via r-universe)

November 4, 2024

Title Instrumental Variables Estimation for 2^k Factorial Experiments

Version 0.1.0

Description Implements instrumental variable estimators for 2^K factorial experiments with noncompliance.

Depends R (>= 2.10)

Imports stats, generics, Formula

Suggests tibble, dplyr, rmarkdown, knitr

VignetteBuilder knitr

License GPL (>= 2)

URL https://github.com/mattblackwell/factiv

BugReports https://github.com/mattblackwell/factiv/issues

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Repository https://mattblackwell.r-universe.dev

RemoteUrl https://github.com/mattblackwell/factiv

RemoteRef HEAD

RemoteSha 6630518b843f5782686aa30374e6215e7e1351bc

Contents

compliance_profile *Complier covariate profiles*

Description

Calculates averages of covariates by compliance group in a 2^K factorial setting.

Usage

```
compliance_profile(formula, data, subset)
```

Arguments

`formula`	one-sided formula to indicate the treatment assignment, treatment uptake, and covariates. The right-hand side of the formula should have three components separated by the `|` symbol, with the first component containing the K binary treatment variables (treatment assignment), the second component containing the K binary instruments associated with each treatment variable (treatment uptake), and the third giving the covariates to be included in the profile. The order of the variables in the first two parts of the formula must match.
`data`	a data.frame on which to apply the `formula`.
`subset`	subset of the data to pass to estimation.

Value

A list with two objects:

`raw_table`	a data.frame whose rows represent the covariates and whose columns represent the different compliance groups. Each entry is the estimated mean of the covariate for that compliance group.
`std_table`	a data.frame similarly structured to raw_table but with the standardized difference between the compilance group means and the overall means in place of the raw means.

Author(s)

Matthew Blackwell

References

Matthew Blackwell and Nicole Pashley (2021) "Noncompliance in Factorial Experiments." Working paper.

Examples

```
data(newhaven)

cov_prof <- compliance_profile(~ inperson + phone | inperson_rand
  + phone_rand | age + maj_party + turnout_96, data = newhaven)

cov_prof
```

`iv_factorial` *IV Estimation of 2^K Factorial Design*

Description

Estimates principal stratum-specific effects and interactions in a 2^K factorial experiment

Usage

```
iv_factorial(formula, data, subset, method = "lm", level = 0.95)
```

Arguments

`formula`	formula specification of the factorial design with noncompliance. The right-hand side of the formula should have two components separated by the | symbol, with the first component containing the K binary treatment variables and the second component containing the K binary instruments associated with each treatment variable. The order of the variables in the formula must match.
`data`	A data.frame on which to apply the `formula`.
`subset`	subset of the data to pass to estimation.
`method`	character indiciating if the estimator should be `"lm"` using the least squares approach (default) or `"cmd"` to estimate via efficent minimum distance estimator.
`level`	the confidence level required.

Details

This function estimates treatment effects for 2^K factorial experiments in the face of noncompliance on all factors. A monotonicity assumption is assumed for both treatment-instrument pairs, along with treatment exclusion. See Blackwell (2017) for more details on those assumptions.

The procedure uses iterative generalized method of moments (GMM) to estimate both the proportions of each compliance class (also known as principal strata) and the average potential outcomes within those classes. It also provides estimates of several one-way, joint, and interactive treatment effects within these classes.

Under the above assumptions, the compliance classes are the product of the compliance classes for each treatment-instrument pair. For instance, `"cc"` is the class that would comply with both treatments, `"ca"` is the class that would comply with the first treatment and always take the second treatment, and `"cn"` is the class that would comply with the first treatment and never take the second treatment. Finally, note that treatment effects are only well-defined for compliance classes for which there is compliance on at least one treatment.

Value

A list of class `iv_factorial` that contains the following components:

- `rho` vector of estimated compliance class probabilities.
- `psi` vector of the estimated conditional mean of the outcome within the compliance classes.
- `vcov` estimated asymptotic variance matrix of the combined `rho` and `psi` parameters.
- `pcafe_est` vector of estimated main effects of each factor among perfect compliers.
- `pcafe_se` vector of estimated standard errors for the estimated effects in `tau`.
- `pcafe_cis` a matrix of confidence intervals for the PCAFE estimates.
- `level` the confidence level of the returned confience intervals.

Author(s)

Matt Blackwell

References

Matthew Blackwell (2017) Instrumental Variable Methods for Conditional Effects and Causal Interaction in Voter Mobilization Experiments, Journal of the American Statistical Association, 112:518, 590-599, doi:10.1080/01621459.2016.1246363

Matthew Blackwell and Nicole Pashley (2020) "Noncompliance in Factorial Experiments." Working paper.

Examples

```
data(newhaven)

out <- iv_factorial(turnout_98 ~ inperson + phone | inperson_rand
  + phone_rand, data = newhaven)

summary(out)
```

`iv_finite_factorial` *Finite-Sample IV Estimation of 2^K Factorial Design*

Description

Estimate main effect IV ratios for 2^K factorial experiments

Usage

```
iv_finite_factorial(formula, data, subset, level = 0.95)
```

Arguments

`formula`	formula specification of the factorial design with noncompliance. The right-hand side of the formula should have two components separated by the `|` symbol, with the first component containing the K binary treatment variables and the second component containing the K binary instruments associated with each treatment variable. The order of the variables in the formula must match.
`data`	a data.frame on which to apply the `formula`.
`subset`	subset of the data to pass to estimation.
`level`	the confidence level required.

Details

This function estimates the ratio of the effect of treatment assignment on the outcome to the effect of treatment assignment on treatment uptake in 2^K factorial experiments. The approach uses finite sample asymptotic inference to generate confidence intervals.

Value

A list of class `iv_finite_factorial` that contains the following components:

`tau`	a vector of estimated effect ratios for each factor.
`tau_cis`	a matrix of confidence intervals for each effect ratio. This matrix has 4 columns because it is possible to have disjoint confidence intervals in this method.
`tau_y`	a vector of the estimated effects on the outcome.
`v_tau_y`	the estimated sample variances of the effects on the outcome.
`tau_d`	a vector of the estimated effects on treatment uptake.
`v_tau_y`	the estimated sample variances of the effects on treatment uptake.
`level`	the confidence level of `tau_cis`.

Author(s)

Matt Blackwell

References

Matthew Blackwell and Nicole Pashley (2021) "Noncompliance in Factorial Experiments." Working paper.

Examples

```
data(newhaven)

out <- iv_finite_factorial(turnout_98 ~ inperson + phone | inperson_rand
  + phone_rand, data = newhaven)

out
```

```
joint <- iv_finite_factorial(turnout_98 ~ inperson + phone |
  inperson_rand + phone_rand, data = newhaven)

joint
```

---

newhaven *New Haven Voter Modelization Field Experiment*

Description

A data set that contains a subset of the voter mobilzation field experiment analyzed in Gerber and Green (2000).

Format

A data frame with 7865 observations and 6 variables:

ward ward of the residence.

turnout_98 indicator variable for voting in the 1998 general election.

inperson indicator variable for if the respondent received in-person canvassing.

phone indicator variable for if the respondent received phone canvassing.

inperson_rand indicator variable for if the respondent was randomized to receive in-person canvassing.

phone_rand indicator variable for if the respondent was randomized to receive phone canvassing.

age respondent age.

maj_party indicator variable for if the respondent is a registered Democrat or Republican.

turnout_96 indicator variable for whether or not the respondent voted in the 1996 general election.

Details

Data was cleaned and used in Bowers and Hansen (2009). It was a 2x2 factorial design with noncompliance on both factors. This is the subset of the subjects who lived in a single-person household and were not randomized to receive get-out-the-vote mailers. Blackwell (2017) analyzed the full data adjusting for noncompliance.

Source

https://doi.org/10.7910/DVN/Q6CID7

References

Gerber, Alan S., and Donald P. Green. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." The American Political Science Review 94, no. 3 (2000): 653. https://doi.org/10.2307/2585837.

Hansen, Ben B., and Jake Bowers. "Attributing Effects to a Cluster-Randomized Get-Out-the-Vote Campaign." Journal of the American Statistical Association 104, no. 487 (2009): 873–85. https://doi.org/10.1198/jasa.2009.ap06589.

Blackwell, Matthew. "Instrumental Variable Methods for Conditional Effects and Causal Interaction in Voter Mobilization Experiments." Journal of the American Statistical Association 112, no. 518 (2017): 590–99. https://doi.org/10.1080/01621459.2016.1246363.

`tidy.iv_factorial` *Tidy an iv_factorial object*

Description

Tidy summarizes information about the components of a model.

Usage

```
## S3 method for class 'iv_factorial'
tidy(x, conf.int = FALSE, conf.level = 0.95, ...)
```

Arguments

`x`	An `iv_factorial` object produced by a call to `iv_factorial()`
`conf.int`	Logical indicating whether or not to include a confidence interval in the tidied output. Defaults to `FALSE`
`conf.level`	The confidence level to use for the confidence interval if `conf.int = TRUE`. Must be strictly greater than 0 and less than 1. Defaults to 0.95, which corresponds to a 95 percent confidence interval.
`...`	Additional arguments. Not used. Needed to match generic signature only.

Value

A `tibble::tibble()` with columns:

`term`	The name of the effect term.
`estimand`	Which complier effect being estimated.
`estimate`	The estimated value of the effect.
`std.error`	The estimated standard error of the effect.
`conf.low`	Lower bound of the confidence interval for the estimate.
`conf.high`	Upper bound of the confidence interval for the estimate.

Author(s)

Matt Blackwell

`tidy.iv_finite_factorial` *Tidy an iv_finite_factorial object*

Description

Tidy summarizes information about the components of a model.

Usage

```
## S3 method for class 'iv_finite_factorial'
tidy(x, conf.level = 0.95, ...)
```

Arguments

`x`	An `iv_factorial` object produced by a call to `iv_finite_factorial()`
`conf.level`	The confidence level to use for the confidence interval. Must be strictly greater than 0 and less than 1. Defaults to 0.95, which corresponds to a 95 percent confidence interval.
`...`	Additional arguments. Not used. Needed to match generic signature only.

Value

A `tibble::tibble()` with columns:

`term`	The name of the effect term.
`estimand`	Which complier effect being estimated.
`estimate`	The estimated value of the effect.
`ci_1_lower`	Lower bound for the first interval of the Fieller confidence regresion for the estimate.
`ci_1_upper`	Upper bound for the first interval of the Fieller confidence regresion for tshe estimate.
`ci_2_lower`	Lower bound for the second interval of the Fieller confidence regresion for the estimate. Only non-`NA` when the confidence region is disjoint.
`ci_2_upper`	Upper bound for the second interval of the Fieller confidence regresion for the estimate. Only non-`NA` when the confidence region is disjoint.

Author(s)

Matt Blackwell

Index