Package 'GK2011'

July 21, 2025

Type Package

Title Gaines and Kuklinski (2011) Estimators for Hybrid Experiments

Version 0.1.3

Date 2016-05-26

Description Implementations of the treatment effect estimators for hybrid (self-selection) experiments, as developed by Brian J. Gaines and James H. Kuklinski, (2011), ``Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection," American Journal of Political Science 55(3): 724-736.

License GPL (>= 2)

URL https://github.com/leeper/GK2011

BugReports https://github.com/leeper/GK2011/issues

LazyData TRUE

Imports stats

Suggests testthat

RoxygenNote 5.0.1

NeedsCompilation no

Author Thomas J. Leeper [aut, cre]

Maintainer Thomas J. Leeper <thosjleeper@gmail.com>

Repository CRAN

Date/Publication 2016-05-27 10:59:07

Contents

GK2011-package *GK2011*

Description

Gaines and Kuklinski (2011) Estimators for Hybrid Experiments

Author(s)

Thomas J. Leeper <thosjleeper@gmail.com>

References

Brian J. Gaines and James H. Kuklinski, (2011), "Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection," *American Journal of Political Science* 55(3): 724-736.

See Also

estimate

ajps *Gaines and Kuklinski (2011) AJPS data*

Description

Subset of data from Gaines and Kuklinski (2011)

Usage

```
ajps
```

Format

tr The treatment indicator, where 1=treatment, 2=control, 3=chose treatment, 4=chose control.

therm.obama A "feeling thermometer" toward John McCain.

therm.mccain A "feeling thermometer" toward Barack Obama.

pid An indicator of party identification, where -1=Republican, 0=Independent, 1=Democrat.

Details

This dataset contains a subset of variables, extracted from the dataset used by Gaines and Kuklinski (2011).

Source

Brian J. Gaines and James H. Kuklinski, (2011), "Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection," *American Journal of Political Science* 55(3): 724-736.

See Also

estimate

Examples

```
data(ajps)

# replicate Gaines and Kuklinski (2011) Table 2
pmean <- function(x) sprintf("%0.1f", mean(x))
cbind(
  # Democrats
  aggregate(cbind(therm.mccain, therm.obama) ~ tr,
            data = ajps[ajps$pid == 1, ], FUN = pmean)[, 1:3],
  n_dem = aggregate(therm.obama ~ tr,
                    data = ajps[ajps$pid == 1, ], FUN = length)[, 2],
  # Republicans
  aggregate(cbind(therm.mccain, therm.obama) ~ tr,
            data = ajps[ajps$pid == -1, ], FUN = pmean)[, 2:3],
  n_rep = aggregate(therm.obama ~ tr,
                    data = ajps[ajps$pid == -1, ], FUN = length)[, 2]
)

# effects for McCain among Democrats
with(ajps[ajps$pid == 1, ], {
  estimate(rand = tr %in% 1:2, tr = tr %in% c(1,3), y = therm.mccain)
})
# effects for McCain among Republicans
with(ajps[ajps$pid == -1, ], {
  estimate(rand = tr %in% 1:2, tr = tr %in% c(1,3), y = therm.mccain)
})

# effects for Obama among Democrats
with(ajps[ajps$pid == 1, ], {
  estimate(rand = tr %in% 1:2, tr = tr %in% c(1,3), y = therm.obama)
})
# effects for Obama among Republicans
with(ajps[ajps$pid == -1, ], {
  estimate(rand = tr %in% 1:2, tr = tr %in% c(1,3), y = therm.obama)
})
```

Description

Estimators for Hybrid Experiments

Usage

```
estimate(rand, tr, y, iterations = 5000L)
```

Arguments

`rand`	An integer or logical vector specifying whether each observation is from the random (1) or self-selection (0) arm of the experiment.
`tr`	An integer or logical vector specifying whether each observation was treated (1) or control (0), regardless of the arm of the experiment.
`y`	A numeric vector specifying outcome values.
`iterations`	An integer specifying the number of bootstrap iterations used to estimate standard errors.

Details

The package provides R implementations of the treatment effect estimators for hybrid (self-selection) experiments, as developed by Gaines and Kuklinski (2011). These functions estimate local average treatment effects for unobserved population subgroups inclined and disinclined to be treated, as revealed by a three-condition (two-arm) experimental design. In the design, participants are randomly assigned to one of three conditions: 1) treatment (T), 2) control (C), or 3) self-selection (S) of treatment or control. The design enables the estimation of three treatment effects:

1. First, the sample average treatment effect is estimated from conditions (1) and (2) as:\ $\hat{t} = \bar{Y}_T - \bar{Y}_C$
2. The effect for those inclined to choose treatment is given by:\ $\hat{t}_s = \frac{\bar{Y}_S - \bar{Y}_C}{\hat{\alpha}}$ where $\hat{\alpha}$ is the observed proportion of individuals in group S that choose T (rather than C).
3. The effect for those disinclined to choose treatment (or, equivalently, inclined to choose control) is given by:\ $\hat{t}_n = \frac{\bar{Y}_T - \bar{Y}_S}{1-\hat{\alpha}}$

By definition, the sample average treatment effect is an average of the other two effects.

Value

A data.frame containing the following variables:

- `Effect`, a character vector of effect names ("t", "t_s", "t_n", "naive")
- `Estimate`, a numeric vector of effect estimates
- `SE`, a numeric vector of bootstrapped standard errors
- `t`, a t-statistic for the effect
- `p`, a two-tailed p-value

The return value will also carry an attribute "alpha", indicating the estimated proportion α.

Author(s)

Thomas J. Leeper <thosjleeper@gmail.com>

References

Brian J. Gaines and James H. Kuklinski, (2011), "Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection," *American Journal of Political Science* 55(3): 724-736.

See Also

ajps

Examples

```
# create fake data
set.seed(12345)
d <-
data.frame(rand = c(rep(1, 200), rep(0, 100)),
           tr = c(rep(0, 100), rep(1, 100), rep(0, 37), rep(1, 63)),
           y = c(rnorm(100), rnorm(100) + 1, rnorm(37), rnorm(63) + 3))

# estimate effects
estimate(rand = d$rand, tr = d$tr, y = d$y)
```

Index