teffects ipw — Inverse-probability weighting

Description

`teffects ipw` estimates the average treatment effect (ATE), the average treatment effect on the treated (ATET), and the potential-outcome means (POMs) from observational data by inverse-probability weighting (IPW). IPW estimators use estimated probability weights to correct for missing data on the potential outcomes. `teffects ipw` accepts a continuous, binary, count, fractional, or nonnegative outcome and allows a multivalued treatment.

See [CAUSAL] **teffects intro** or [CAUSAL] **teffects intro advanced** for more information about estimating treatment effects from observational data.

Quick start

ATE of binary `treat2` on `y` by IPW using a logistic model of `treat2` on `x` and `w`

```
teffects ipw (y) (treat2 x w)
```

Same as above, but estimate ATET

```
teffects ipw (y) (treat2 x w), atet
```

Same as above, but estimate potential-outcome means

```
teffects ipw (y) (treat2 x w), pomeans
```

ATE of `treat2` on `y` using heteroskedastic probit for `treat2` as a function of `x` and `w`

```
teffects ipw (y) (treat2 x w, hetprobit(x w))
```

ATE for treatment levels 2 and 3 of three-valued treatment `treat3`

```
teffects ipw (y) (treat3 x w)
```

Same as above, and specify that `treat3` = 3 is the control level

```
teffects ipw (y) (treat3 x w), control(3)
```

Same as above, specified using the label “MyControl” corresponding to `treat3` = 3

```
teffects ipw (y) (treat3 x w), control(MyControl)
```

Menu

Statistics > Causal inference/treatment effects > Continuous outcomes > Inverse-probability weighting (IPW)

Statistics > Causal inference/treatment effects > Binary outcomes > Inverse-probability weighting (IPW)

Statistics > Causal inference/treatment effects > Count outcomes > Inverse-probability weighting (IPW)

Statistics > Causal inference/treatment effects > Fractional outcomes > Inverse-probability weighting (IPW)

Statistics > Causal inference/treatment effects > Nonnegative outcomes > Inverse-probability weighting (IPW)

Syntax

`teffects ipw (`*ovar*`) (`*tvar tmvarlist* [, *tmodel* `noconstant`]`)` [*if*] [*in*] [*weight*]
[, *stat options*]

ovar is a binary, count, continuous, fractional, or nonnegative outcome of interest.

tvar must contain integer values representing the treatment levels.

tmvarlist specifies the variables that predict treatment assignment in the treatment model.

tmodel	Description
Model	
`logit`	logistic treatment model; the default
`probit`	probit treatment model
`hetprobit(`*varlist*`)`	heteroskedastic probit treatment model

tmodel specifies the model for the treatment variable.
For multivalued treatments, only `logit` is available and multinomial logit is used.

stat	Description
Stat	
`ate`	estimate average treatment effect in population; the default
`atet`	estimate average treatment effect on the treated
`pomeans`	estimate potential-outcome means

options	Description
SE/Robust	
`vce(`*vcetype*`)`	*vcetype* may be `robust`, `cluster` *clustvar*, `bootstrap`, or `jackknife`
Reporting	
`level(`#`)`	set confidence level; default is `level(95)`
`aequations`	display auxiliary-equation results
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
Advanced	
`pstolerance(`#`)`	set tolerance for overlap assumption
`osample(`*newvar*`)`	*newvar* identifies observations that violate the overlap assumption
`control(`# | *label*`)`	specify the level of *tvar* that is the control
`tlevel(`# | *label*`)`	specify the level of *tvar* that is the treatment
`coeflegend`	display legend instead of statistics

tmvarlist may contain factor variables; see [U] **11.4.3 Factor variables**.

bayesboot, bootstrap, by, collect, jackknife, and statsby are allowed; see [U] **11.1.10 Prefix commands**.

Weights are not allowed with the bootstrap prefix; see [R] **bootstrap**.

fweights, iweights, and pweights are allowed; see [U] **11.1.6 weight**.

coeflegend does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

noconstant; see [R] **Estimation options**.

Stat

stat is one of three statistics: ate, atet, or pomeans. ate is the default.

ate specifies that the average treatment effect be estimated.

atet specifies that the average treatment effect on the treated be estimated.

pomeans specifies that the potential-outcome means for each treatment level be estimated.

SE/Robust

vce(*vcetype*) specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (robust), that allow for intragroup correlation (cluster *clustvar*), and that use bootstrap or jackknife methods (bootstrap, jackknife); see [R] ***vce_option***.

Reporting

level(#); see [R] **Estimation options**.

aequations specifies that the results for the outcome-model or the treatment-model parameters be displayed. By default, the results for these auxiliary parameters are not displayed.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **Estimation options**.

Maximization

maximize_options: iterate(#), [no]log, and from(*init_specs*); see [R] **Maximize**. These options are seldom used.

init_specs is one of

matname [, skip copy]

# [, # . . .] , copy

Advanced

pstolerance(#) specifies the tolerance used to check the overlap assumption. The default value is pstolerance(1e-5). teffects will exit with an error if an observation has an estimated propensity score smaller than that specified by pstolerance().

osample(*newvar*) specifies that indicator variable *newvar* be created to identify observations that violate the overlap assumption.

`control(#|`*label*`)` specifies the level of *tvar* that is the control. The default is the first treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. `control()` may not be specified with statistic `pomeans`. `control()` and `tlevel()` may not specify the same treatment level.

`tlevel(#|`*label*`)` specifies the level of *tvar* that is the treatment for the statistic `atet`. The default is the second treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. `tlevel()` may only be specified with statistic `atet`. `tlevel()` and `control()` may not specify the same treatment level.

The following option is available with `teffects ipw` but is not shown in the dialog box:

`coeflegend`; see [R] **Estimation options**.

Remarks and examples

Remarks are presented under the following headings:

Overview
Video example

Overview

IPW estimators use estimated probability weights to correct for the missing-data problem arising from the fact that each subject is observed in only one of the potential outcomes. IPW estimators use a two-step approach to estimating treatment effects:

1. They estimate the parameters of the treatment model and compute the estimated inverse-probability weights.
2. They use the estimated inverse-probability weights to compute weighted averages of the outcomes for each treatment level. The contrasts of these weighted averages provide the estimates of the ATEs. Using this weighting scheme corrects for the missing potential outcomes.

These steps produce consistent estimates of the effect parameters because the treatment is assumed to be independent of the potential outcomes after conditioning on the covariates. The overlap assumption ensures that predicted inverse-probability weights do not get too large. In fact, `teffects ipw` uses an estimation technique that implements both steps at once so that we do not need to correct the standard errors in the second step to reflect the uncertainty associated with the predicted treatment probabilities.

We will illustrate the use of `teffects ipw` by using data from a study of the effect of a mother's smoking status during pregnancy (`mbsmoke`) on infant birthweight (`bweight`) as reported by Cattaneo (2010). This dataset also contains information about each mother's age (`mage`), education level (`medu`), marital status (`mmarried`), whether the first prenatal exam occurred in the first trimester (`prenatal1`), and whether this baby was the mother's first birth (`fbaby`).

▷ Example 1: Estimating the ATE

We begin by using `teffects ipw` to estimate the average treatment effect of smoking on birthweight. We will use a probit model to predict treatment status, using `prenatal1`, `mmarried`, `mage`, the square of `mage`, and `fbaby` as explanatory variables:

```
. use https://www.stata-press.com/data/r19/cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)

. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu, probit)

Iteration 0:  EE criterion = 4.622e-21
Iteration 1:  EE criterion = 8.795e-26

Treatment-effects estimation                   Number of obs     =      4,642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: probit
──────────────┬────────────────────────────────────────────────────────────────
              │               Robust
      bweight │ Coefficient  std. err.      z    P>|z|     [95% conf. interval]
──────────────┼────────────────────────────────────────────────────────────────
ATE           │
      mbsmoke │
     (Smoker  │
          vs  │
  Nonsmoker)  │  -230.6886   25.81524    -8.94   0.000    -281.2856   -180.0917
──────────────┼────────────────────────────────────────────────────────────────
POmean        │
      mbsmoke │
   Nonsmoker  │   3403.463   9.571369   355.59   0.000     3384.703    3422.222
──────────────┴────────────────────────────────────────────────────────────────
```

The average birthweight if all mothers were to smoke would be 231 grams less than the average of 3,403 grams that would occur if none of the mothers had smoked.

◁

Sometimes, we are mainly concerned about those subjects that did in fact receive treatment, and we want to know how much the outcome changes as a result of treatment for that subpopulation. The ATET provides us with the answer. Moreover, the ATET can be estimated using weaker assumptions than are required to estimate the ATE; see [CAUSAL] **teffects intro advanced**.

▷ Example 2: Estimating the ATET

```
. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu, probit),
> atet

Iteration 0:  EE criterion = 4.633e-21
Iteration 1:  EE criterion = 7.494e-27

Treatment-effects estimation                   Number of obs     =      4,642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: probit
------------------------------------------------------------------------------
             |               Robust
     bweight | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
ATET         |
     mbsmoke |
    (Smoker  |
         vs  |
 Nonsmoker)  |  -225.1773   23.66458    -9.52   0.000     -271.559   -178.7955
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
  Nonsmoker  |   3362.837   14.20149   236.79   0.000     3335.003    3390.671
------------------------------------------------------------------------------
```

The average birthweight is 225 grams less when all the mothers who smoke do so than the average of 3,363 grams that would have occurred if none of these mothers had smoked.

◁

We often express statistics as percentages to alleviate scaling issues and aid interpretation. In the present context, we may wish to express an ATE as a percentage of the untreated POM to gain a more intuitive measure of the effect of treatment.

▷ Example 3: Reporting the ATE as a percentage

Here we use the same model as in example 1, but we report the ATE as a percentage of the mean birthweight that would occur if no mothers smoke. First, we use `teffects ipw` to fit the model. We use the `coeflegend` option so that `teffects ipw` reports the names of the parameters. Then we use `nlcom` to obtain the statistic we want along with its delta-method-based standard error. We type

```
. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu, probit),
> coeflegend

Iteration 0:   EE criterion =  4.622e-21
Iteration 1:   EE criterion =  8.795e-26

Treatment-effects estimation                     Number of obs      =      4,642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: probit
--------------------------------------------------------------------------------
     bweight | Coefficient  Legend
-------------+------------------------------------------------------------------
ATE          |
     mbsmoke |
    (Smoker  |
         vs  |
 Nonsmoker)  |  -230.6886   _b[ATE:r1vs0.mbsmoke]
-------------+------------------------------------------------------------------
POmean       |
     mbsmoke |
  Nonsmoker  |   3403.463   _b[POmean:0.mbsmoke]
--------------------------------------------------------------------------------

. nlcom _b[ATE:r1vs0.mbsmoke] / _b[POmean:0.mbsmoke]

       _nl_1: _b[ATE:r1vs0.mbsmoke] / _b[POmean:0.mbsmoke]

--------------------------------------------------------------------------------
     bweight | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+------------------------------------------------------------------
       _nl_1 |  -.0677806   .0075169    -9.02   0.000    -.0825133   -.0530478
--------------------------------------------------------------------------------
```

The average birthweight falls by an estimated 6.8% when every mother smokes relative to the case when no mothers smoke. We also obtain a 95% confidence interval of a 5.3% to 8.3% reduction.

◁

Video example

Treatment effects: Inverse-probability weighting

Stored results

`teffects ipw` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(n`*j*`)`	number of observations for treatment level *j*
`e(N_clust)`	number of clusters
`e(k_eq)`	number of equations in `e(b)`
`e(k_levels)`	number of levels in treatment variable
`e(treated)`	level of treatment variable defined as treated
`e(control)`	level of treatment variable defined as control
`e(converged)`	`1` if converged, `0` otherwise

Macros

`e(cmd)`	`teffects`
`e(cmdline)`	command as typed
`e(depvar)`	name of outcome variable
`e(tvar)`	name of treatment variable
`e(subcmd)`	`ipw`
`e(tmodel)`	`logit`, `probit`, or `hetprobit`
`e(stat)`	statistic estimated, `ate`, `atet`, or `pomeans`
`e(wtype)`	weight type
`e(wexp)`	weight expression
`e(title)`	title in estimation output
`e(clustvar)`	name of cluster variable
`e(tlevels)`	levels of treatment variable
`e(vce)`	*vcetype* specified in `vce()`
`e(vcetype)`	title used to label Std. err.
`e(properties)`	`b V`
`e(estat_cmd)`	program used to implement `estat`
`e(predict)`	program used to implement `predict`
`e(marginsnotok)`	predictions disallowed by `margins`
`e(asbalanced)`	factor variables `fvset` as `asbalanced`
`e(asobserved)`	factor variables `fvset` as `asobserved`

Matrices

`e(b)`	coefficient vector
`e(V)`	variance–covariance matrix of the estimators

Functions

`e(sample)`	marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices

`r(table)`	matrix containing the coefficients with their standard errors, test statistics, p-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

`teffects ipw` implements a smooth treatment-effects estimator. All smooth treatment-effects estimators are documented in *Methods and formulas* of [CAUSAL] **teffects aipw**.

References

Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154. https://doi.org/10.1016/j.jeconom.2009.09.023.

Cerulli, G. 2014. treatrew: A user-written command for estimating average treatment effects by reweighting on the propensity score. *Stata Journal* 14: 541–561.

Drukker, D. M. 2014. Using gmm to solve two-step estimation problems. *The Stata Blog: Not Elsewhere Classified.* https://blog.stata.com/2014/12/08/using-gmm-to-solve-two-step-estimation-problems/.

Huber, C. 2015. Introduction to treatment effects in Stata: Part 1. *The Stata Blog: Not Elsewhere Classified.* https://blog.stata.com/2015/07/07/introduction-to-treatment-effects-in-stata-part-1/.

Also see

[CAUSAL] **teffects postestimation** — Postestimation tools for teffects

[CAUSAL] **teffects** — Treatment-effects estimation for observational data

[U] **20 Estimation and postestimation commands**

For suggested citations, see the FAQ on citing Stata documentation.