Title

paramed --   causal mediation analysis using parametric regression models

Syntax

paramed varname, avar(varname) mvar(varname) a0(real) a1(real) m(real) yreg(string)

mreg(string) [cvars(varlist) nointeraction casecontrol fulloutput

c(numlist) bootstrap reps(integer 1000) level(cilevel) seed(passthru)]

varname - this specifies the outcome variable.

avar(varname) - this specifies the treatment (exposure) variable.

mvar(varname) - this specifies the mediator variable.

a0(real) - this specifies the natural level of the treatment (exposure).

a1(real) - this specifies the alternative treatment (exposure) level.

m(real) - this specifies the level of mediator at which the controlled direct effect is

to be estimated. If there is no treatment (exposure)-mediator interaction the

controlled direct effect is the same at all levels of the mediator and so an

arbitary value can be chosen.

yreg(string) - this specifies the form of regression model to be fitted for the outcome

variable. This can be either linear, logistic, loglinear, Poisson or Negative

binomial.

mreg(string) - this specifies the form of regression model to be fitted for the

mediator variable. This can be either linear or logistic.

Description

paramed performs causal mediation analysis using parametric regression models. Two

models are estimated: a model for the mediator conditional on treatment (exposure) and

covariates (if specified), and a model for the outcome conditional on treatment

(exposure), the mediator and covariates (if specified). It extends statistical

mediation analysis (widely known as Baron and Kenny procedure) to allow for the

presence of treatment (exposure)-mediator interactions in the outcome regression model

using counterfactual definitions of direct and indirect effects.

paramed allows continuous, binary or count outcomes, and continuous or binary mediators, and requires the user to specify an appropriate form for the regression models.

paramed provides estimates of the controlled direct effect, the natural direct effect, the natural indirect effect and the total effect with standard errors and confidence intervals derived using the delta method by default, with a bootstrap option also available. See references for precise definitions of these effects.

Options

cvars(varlist) - this option specifies the list of covariates to be included in the analysis. Categorical variables need to be coded as a series of dummy variables before being entered as covariates.

nointeraction - this option specifies whether a treatment (exposure)-mediator interaction is not to be included in the models (the default assumes an interaction is present).

fulloutput - this option specifies the output mode, which can be either reduced or full. The reduced output is the default option (if this option is omitted). The results matrix contains the controlled direct effect, natural direct effect, natural indirect effect and total effect. When the full option is specified, both conditional effects and effects evaluated at the mean covariate levels are shown.

c(numlist) - this option is used when the output option is full. When the output mode is full, fixed values must be provided for the covariates at which conditional effects are computed (the number of values must correspond to the number of covariates).

casecontrol - this option is used for implementing mediation analysis when data arise from a case-control design, provided the outcome in the population is rare. If this

option is omitted, the data will not be treated as from a case-control design.

bootstrap - this specifies whether a bootstrap procedure should be performed to compute bias-corrected bootstrap confidence intervals. The bootstrap procedure will not be performed if this option is omitted.

reps(integer 1000) - this specifies the number of replications for bootstrap. The default is 1000.

level(cilevel) - this specifies the confidence level for bootstrap. If this option is omitted, the current default level of 95% will be used.

seed(passthru) - this specifies the seed for bootstrap. If this option is omitted, a random seed will be used and the results cannot be replicated.

Assumptions

Let C be the measured covariates included in cvars(varlist). To obtain valid estimates of the controlled direct effects requires two assumptions:

(1) There are no unmeasured treatment (exposure)-outcome confounders given C

(2) There are no unmeasured mediator-outcome confounders given C

To estimate natural direct and indirect effects we need the assumptions (1) and (2) and require need two additional assumptions:

(3) There are no unmeasured treatment (exposure)-mediator confounders given C

(4) There is no effect of treatment (exposure) that confounds the mediator-outcome relationship

Note that assumptions (1) and (3) are satisified by random allocation of the treatment variable. See references for further details.

Examples

Setup

. use paramed_example.dta

Continuous outcome, continuous mediator, a binary treatment coded 0 and 1, two covariates, no interaction between treatment and mediator, delta method standard errors

```
. paramed y_cont, avar(treat) mvar(m_cont) cvars(var1 var2) a0(0) a1(1) m(1)
    yreg(linear) mreg(linear) nointer
```

Continuous outcome, binary mediator, a binary treatment coded 0 and 1, two covariates, include an interaction between treatment and mediator, bootstrap standard errors with default bootstrap settings

```
. paramed y_cont, avar(treat) mvar(m_bin) cvars(var1 var2) a0(0) a1(1) m(1)
    yreg(linear) mreg(logistic) boot
```

Binary outcome, binary mediator, a binary treatment coded 0 and 1, no covariates, no interaction between treatment and mediator, bootstrap standard errors with 500 replications and fixing the seed to 1234

```
. paramed y_bin, avar(treat) mvar(m_bin) a0(0) a1(1) m(1) yreg(logistic)
    mreg(logistic) nointer boot reps(500) seed(1234)
```

Count outcome with a Poisson model, binary mediator, a binary treatment coded 0 and 1, two covariates, no interaction between treatment and mediator, bootstrap standard errors with 1000 replications and fixing the seed to 1234

```
. paramed y_poisson, avar(treat) mvar(m_bin) cvars(var1 var2) a0(0) a1(1) m(1)
    yreg(poisson) mreg(logistic) nointer boot seed(1234)
```

Continuous outcome, binary mediator, a binary treatment coded 0 and 1, two covariates, interaction between treatment and mediator, and request full output.

```
. paramed y_cont, avar(treat) mvar(m_bin) cvars(var1 var2) a0(0) a1(1) m(1)
    yreg(linear) mreg(logistic) c(10 6) full
```

Saved results

paramed saves the following results in e():

Matrices

  e(b)   matrix containing direct, indirect and total effect estimates

  e(V)   matrix containing variance of the effect estimates

Authors

Hanhua Liu, Richard Emsley and Graham Dunn
Centre for Biostatistics
Institute of Population Health
The University of Manchester

Tyler VanderWeele and Linda Valeri
Harvard School of Public Health
Harvard University

Email: XXX or XXX

Further reading

Emsley RA, Liu H, Dunn G, Valeri L, VanderWeele TJ. Paramed: a command to perform causal mediation analysis using parametric models 2015; in preparation.

Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS Macros. *Psychological Methods* 2013;**18**:137–50.

VanderWeele TJ and Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface - Special Issue on Mental Health and Social Behavioral Science* 2009;**2**:457–68.