

Package ‘leiv’

February 20, 2015

Version 2.0-7

Type Package

Title Bivariate Linear Errors-In-Variables Estimation

Date 2015-01-11

Maintainer David Leonard <davidsleonard@outlook.com>

Depends R (>= 2.9.0)

Imports methods, stats, graphics

Suggests grDevices

Description Estimate the slope and intercept of a bivariate linear relationship by calculating a posterior density that is invariant to interchange and scaling of the coordinates.

License GPL (>= 2)

URL <http://www.r-project.org>

Author David Leonard [aut, cre]

NeedsCompilation no

Repository CRAN

Date/Publication 2015-01-11 16:53:08

R topics documented:

leiv	2
leiv-internal	5
Index	6

Description

Generates a linear errors-in-variables object.

Usage

```
leiv(formula, data, subset, prior = NULL,
      n = NULL, cor = NULL, sdRatio = NULL, xMean = 0, yMean = 0,
      probIntCalc = FALSE, level = 0.95, subdivisions = 100,
      rel.tol = .Machine$double.eps^0.25, abs.tol = 0.1*rel.tol, ...)

## S4 method for signature 'leiv'
print(x, digits = max(3, getOption("digits") - 3), ...)
## S4 method for signature 'leiv,missing'
plot(x, plotType = "density", xlim = NULL, ylim = NULL,
      xlab = NULL, ylab = NULL, col = NULL, lwd = NULL, ...)
```

Arguments

formula	an optional object of class " <i>formula</i> " (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given in the 'Details' section of the documentation for <i>lm</i> . An intercept is always included and integrated out as a nuisance parameter: $y \sim x$, $y \sim \theta + x$, and $y \sim x - 1$ are equivalent. If not provided, the sufficient statistics <i>n</i> , <i>cor</i> , and <i>sdRatio</i> must be provided.
data	an optional data frame (or object coercible by <i>as.data.frame</i> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <i>environment(formula)</i> , typically the environment from which <i>leiv</i> is called.
subset	an optional vector specifying a subset of observations to be used in the fitting process.
prior	an optional object of class <i>leiv</i> to use as the prior density of the scale invariant slope; otherwise the rotationally invariant Cauchy density is used.
n	an optional sample size (if <i>formula</i> is missing).
cor, sdRatio	optional sample correlation $\text{cor}(x,y)$ and ratio $\text{sd}(y)/\text{sd}(x)$ (if <i>formula</i> is missing).
xMean, yMean	optional sample means $\text{mean}(x)$ and $\text{mean}(y)$ (if <i>formula</i> is missing).
probIntCalc	logical; if TRUE returns the shortest $(100*\text{level})\%$ probability intervals; if FALSE (the default) no probability intervals are returned.
level	the probability level requested (if <i>probIntCalc</i> = TRUE).
subdivisions	the maximum number of subintervals (see <i>integrate</i>).

rel.tol	the relative accuracy requested (see integrate).
abs.tol	the absolute accuracy requested (see integrate).
x	a leiv object.
digits	controls formatting of numeric objects.
plotType	specifies the type of plot; if plotType = "density" (the default) then the posterior density of the slope is plotted; if plotType = "scatter" then a scatter plot with the fitted line.
xlim, ylim	x limits c(x1, x2) and y limits c(y1, y2) of the plot.
xlab, ylab	labels for the x and y axes of the plot.
col, lwd	color and width of plotted lines.
...	additional argument(s) for generic methods.

Details

Use `leiv` to estimate the slope and intercept of a bivariate linear relationship when both variables are observed with error. The method is exact when the true values and the errors are normally distributed. The posterior density depends on the data only through the correlation coefficient and ratio of standard deviations; it is invariant to interchange and scaling of the coordinates.

Value

`leiv` returns an object of class "leiv" with the following components:

slope	the (posterior median) slope estimate.
intercept	the (maximum likelihood) intercept estimate.
slopeInt	the shortest (100*level)% probability interval of the slope.
interceptInt	the shortest (100*level)% probability interval of the intercept.
density	the posterior probability density function.
n	the number of (x,y) pairs.
cor	the sample correlation $\text{cor}(x, y)$.
sdRatio	the ratio $\text{sd}(y)/\text{sd}(x)$.
xMean	the sample mean $\text{mean}(x)$.
yMean	the sample mean $\text{mean}(y)$.
call	the matched call.
probIntCalc	the logical probability interval request.
level	the probability level of the probability interval.
x	the x data.
y	the y data.

Note

Numerical integration is used to normalize the posterior density. When the data is nearly linear, normalization using the default tolerance parameters may fail. Specifying `abs.tol = 1e-6` (or smaller) may help, but expect a longer run time. In general, `rel.tol` cannot be less than `max(50*.Machine$double.eps, 0.5e-28)` if `abs.tol <= 0`. In addition, when using a sharply peaked `leiv` object as a prior density, normalization may fail. In this case, an alternative is to first fit using the default Cauchy prior, then multiply by the appropriate ratio of prior densities and tackle the normalization outside of the `leiv` environment.

Author(s)

David Leonard

References

Leonard, David. (2011). "Estimating a Bivariate Linear Relationship." *Bayesian Analysis*, 6:727-754. DOI:10.1214/11-BA627.

Zellner, Arnold. (1971). *An Introduction to Bayesian Inference in Econometrics*, Chapter 5. John Wiley & Sons.

See Also

[lm](#) for formula syntax; [integrate](#) for control parameters.

Examples

```
## generate artificial data
set.seed(1123)
n <- 20
X <- rnorm(n, mean=5, sd=4) # true x
x <- X + rnorm(n, mean=0, sd=5) # observed x
Y <- 2 + X # true y
y <- Y + rnorm(n, mean=0, sd=3) # observed y

## fit with default options
fit <- leiv(y ~ x)
print(fit)
plot(fit) # density plot
dev.new()
plot(fit,plotType="scatter")

## calculate a density to use as an informative prior density of
## the scale invariant slope in a subsequent fit
fit0 <- leiv(n=10, cor=0.5, sdRatio=1.0)
print(fit0)

## refit the data using the informative prior density
fit1 <- leiv(y ~ x, prior=fit0, abs.tol=1e-6)
print(fit1)
```

leiv-internal *Probability Density Utilities*

Description

p50 calculates the median of the [leiv](#) posterior probability density. probInt calculates the shortest probability interval of the [leiv](#) posterior probability density for a given probability level.

Usage

```
p50(p, interval, subdivisions = 100,
    rel.tol = .Machine$double.eps^0.25,
    abs.tol = rel.tol)

probInt(p, interval, level, subdivisions = 100,
        rel.tol = .Machine$double.eps^0.25,
        abs.tol = rel.tol)
```

Arguments

p	a normalized probability density function.
interval	a vector containing the endpoints of the interval to be searched.
level	the probability level requested.
subdivisions	the maximum number of subintervals (see integrate).
rel.tol	the relative accuracy requested (see integrate).
abs.tol	the absolute accuracy requested (see integrate , optimize and uniroot).

Details

Internal functions for integrating the posterior density returned by the function [leiv](#). These functions are not meant to be called by the user.

Value

p50 returns a numeric scalar. probInt returns a 2-dimensional numeric vector of interval endpoints.

Note

p must accept a vector of inputs and produce a vector of function evaluations at those points. rel.tol cannot be less than $\max(50 \times .Machine$double.eps, 0.5e-28)$ if $abs.tol \leq 0$.

See Also

[leiv](#) for general information; [integrate](#) for control parameters.

Index

*Topic **models**

leiv, 2

*Topic **regression**

leiv, 2

as.data.frame, 2

formula, 2

integrate, 2–5

leiv, 2, 5

leiv-class (leiv), 2

leiv-internal, 5

leiv-package (leiv), 2

lm, 2, 4

numeric, 3

optimize, 5

p50 (leiv-internal), 5

plot, ANY, ANY-method (leiv), 2

plot, leiv, missing-method (leiv), 2

plot-methods (leiv), 2

print, ANY-method (leiv), 2

print, leiv-method (leiv), 2

print-methods (leiv), 2

probInt (leiv-internal), 5

uniroot, 5