

Package ‘copas’

January 2, 2012

Title Statistical methods to model and adjust for bias in meta-analysis

Version 0.6-3

Author James Carpenter <James.Carpenter@lshtm.ac.uk>, Guido Schwarzer
<sc@imbi.uni-freiburg.de>

Maintainer Guido Schwarzer <sc@imbi.uni-freiburg.de>

Depends meta (>= 0.8)

Date 2009-11-26

Description Copas selection model.

License GPL (>= 2)

Repository CRAN

Date/Publication 2009-11-26 11:43:47

R topics documented:

copas	2
Crowther2003	7
plot.copas	8
print.copas	11
Index	14

 copas

Copas selection model analysis

Description

Perform a Copas selection model analysis for selection bias in meta-analysis.

The program takes an object of class `meta`, which is most easily created by an analysis using one of the functions `metabin`, `metacont` and `metagen` in the package `meta`, performs a 'Copas selection model analysis' and presents a graphical and tabular summary of the results. An object of class `copas` is created and this can be used to recreate the results table and graphs subsequently, without re-running the analysis, using the `print`, `summary` and `plot` function.

Usage

```
copas(x,
      gamma0.range=NULL, gamma1.range=NULL,
      ngrid=20, nlevels=10, levels=NULL,
      slope=NULL, left=NULL, rho.bound=0.9999,
      silent=TRUE, warn=options()$warn)
```

Arguments

- | | |
|---------------------------|--|
| <code>x</code> | An object of class <code>meta</code> , obtained from one of the functions <code>metabin</code> , <code>metacont</code> and <code>metagen</code> in the package <code>meta</code> . |
| <code>gamma0.range</code> | (Advanced users only) A numerical vector of length two specifying the range of <code>gamma0</code> values the program will explore.
The parameter <code>gamma0</code> is the constant in the probit selection model for study publication. Thus, the cumulative normal of <code>gamma0</code> is approximately the probability that a small study is published (in non-technical terms <code>gamma0</code> relates to the probability of publishing a small study, although its values are not restricted to the range $[0,1]$; larger values correspond to higher probabilities of publishing a small study). Most users will not need to specify a range for this parameter. When no argument is specified, the program uses an algorithm to determine a suitable range. This is based on the range of treatment effect standard errors in the meta-analysis, and is described in more detail below. |
| <code>gamma1.range</code> | (Advanced users only) A numerical vector of length two specifying the range of <code>gamma1</code> values the program will explore.
The parameter <code>gamma1</code> is the coefficient of study precision ($1/\text{standard error}$) in the probit selection model for study publication (in non-technical terms <code>gamma1</code> relates to the rate at which the probability of publishing a study increases as the standard error of the treatment effect it reports decreases; larger values correspond to higher probabilities of publishing a small study). Most users will not need to specify a range for this parameter. When no argument is specified, the program uses an algorithm to determine a suitable range. This is based on the range of treatment effect standard errors in the meta-analysis, and is described in more detail below. |

ngrid	The program fits the Copas selection model over a grid defined by the range of values of gamma0 and gamma1 specified in the previous two arguments. This parameter fixes the square-root of the number of points in the grid.
nlevels	(Advanced users only). Fitting the Copas model over the grid specified by the previous three arguments results in a treatment estimate at every point in the grid. These can then be displayed on a contour plot where contours of treatment effect (z-axis) are shown by gamma0 (x-axis) and gamma1 (y-axis). This argument specifies the number of contour lines that will be drawn. Note (i) Calculations for the contour plot are performed by the function copas, so this argument has no effect in the plot function. (ii) If a large number of contour lines are desired, then you may wish to consider increasing the grid size (argument ngrid above). Leave this option unspecified if you are using the option levels below.
levels	A numerical vector of treatment values for which contour lines will be drawn. In more detail, fitting the Copas model over the grid specified by the arguments gamma0.range, gamma1.range and ngrid results in a treatment estimate at every point in the grid. These are then displayed on a contour plot where contours of treatment effect (z-axis) are shown by gamma0 (x-axis) and gamma1 (y-axis). This argument is a numerical vector which specifies the treatment effects for which contour lines will be drawn. It is usually not a good idea to set this argument for initial runs, as one does not know the range of treatment values that the contour plot will cover, and treatment values which do not correspond to values in the contour plot (defined by the range of gamma0 and gamma1) will not be plotted. Note (i) Calculations for the contour plot are performed by the function copas, so this argument has no effect in the plot function. (ii) Contours will not be drawn if a large number of contour lines are desired, then you may wish to consider increasing the grid size (argument ngrid above). Leave this option unspecified if you are using the option nlevels above.
slope	A numeric providing the slope of the line approximately orthogonal to contours in the contour plot. If the argument slope is NULL (default) the program seeks to estimate the slope of the contours in the region of the maximum, which are usually approximately parallel. Most users will leave the argument slope unspecified, at least for the first analysis of a data set, but in certain cases setting it manually can improve the results.
left	A logical indicating whether the cause of any selection bias is due to missing studies on the left or right of the funnel plot: left hand side if left=TRUE, right hand side if left=FALSE. This information is needed in order to be sure the test for presence of residual selection bias is calculated correctly. If not set, the linear regression test for funnel plot asymmetry (i.e., function metabias(..., meth="linreg")) is used to determine whether studies are missing on the left or right hand side. In the majority of cases this will work correctly.
rho.bound	(Advanced users only) A number giving the upper bound for the correlation parameter rho (see details below). This must be < 1, and usually > 0.95. The lower bound is calculated as -(the upper bound).

silent	A logical indicating whether information on progress in fitting the Copas selection model should be printed: <code>silent=TRUE</code> , do not print information (the default); <code>silent=FALSE</code> , print information.
warn	A number setting the handling of warning messages. It is not uncommon for numerical problems to be encountered during estimation over the grid of (<code>gamma0</code> , <code>gamma1</code>) values. Usually this does not indicate a serious problem. This option specifies what to do with warning messages. <code>warn=-1</code> : ignore all warnings; <code>warn=0</code> (the default): store warnings till function finishes; if there are less than 10, print them, otherwise print a message saying warning messages were generated; <code>warn=1</code> : print warnings as they occur; <code>warn=2</code> : stop the function when the first warning is generated. For further details see <code>help(options)</code> .

Details

Conduct a Copas selection model analysis to investigate, and attempt to correct for, selection/publication bias in a meta-analysis.

The Copas selection model consists of two models, which are fitted jointly. The first is the usual random effects meta-analysis model, and the second is a selection model, where study i is selected for publication if $Z > 0$, where

$$Z = \text{gamma0} + \text{gamma1}/(\text{SE}(i)) + \text{delta}(i)$$

The error $\text{delta}(i)$ is correlated with the error in the random effects meta-analysis, with correlation ρ . If $\rho=0$, the model corresponds to the usual random effects meta-analysis. As ρ moves from 0 to 1, studies with larger treatment estimates are more likely to be selected/published.

The software chooses a grid of `gamma0` and `gamma1` values, corresponding to a range of selection/publication probabilities for the study with the largest treatment effect standard error (often the smallest study). For each value in this grid, the treatment effect is estimated using the function `optim`. This information is used to produce the contour plot (top right panel of output from `plot.copas`).

Contours of constant treatment effect are usually locally parallel. The software estimates the slope of these contours, and combines this information with other parameter estimates from the model to explore (i) how the treatment estimate, and its standard error, change with increasing selection (bottom left panel, `plot.copas`) and (ii) how much selection needs to be accounted for before any remaining asymmetry in the funnel plot is likely to have occurred by chance (bottom right panel, `plot.copas`).

A table of results can be produced by the function `summary.copas`. A more detail output is provided by the function `print.copas`.

For a fuller description of the model, our implementation and specifically our approach to estimating the locally parallel contours, see Carpenter, Schwarzer, Rucker et. al (2008).

Value

An object of class `copas` with corresponding `print`, `summary`, `plot` function. The object is a list containing the following components:

TE	Vector of treatment effects plotted in treatment effect plot
seTE	Vector of standard error of TE

TE.random	Usual random effects estimate of treatment effect
seTE.random	Usual standard error of TE.random
left	Whether selection bias expected on left or right
rho.bound	Bound on rho
gamma0.range	Range of gamma0 (see help on copas arguments above)
gamma1.range	Range of gamma1 (see help on copas arguments above)
slope	Slope of line approximately orthogonal to contours in contour plot
regr	A list containing information on regression lines fitted to contours in contour plot
ngrid	Square root of grid size
nlevels	Number of contour lines
gamma0	Vector of gamma0 values at which model fitted (determined by gamma0.range and grid). x-axis values for contour plot
gamma1	vector of gamma1 values at which model fitted (determined by gamma1.range and grid). y-axis values for contour plot
TE.contour	Treatment values (ie z-axis values) used to draw contour plot.
x.slope	x coordinates for 'orthogonal line' in contour plot
y.slope	y coordinates for 'orthogonal line' in contour plot
TE.slope	Vector of treatment values plotted in treatment effect plot
seTE.slope	Standard error of TE.slope
rho.slope	Vector of estimated rho values corresponding to treatment estimates in TE.slope
tau.slope	Vector of estimated heterogeneity values corresponding to treatment estimates in TE.slope
loglik1	Vector of log-likelihood values corresponding to treatment estimates in TE.slope
conv1	Numerical vector indicating convergence status for each treatment estimate in TE.slope - see parameter convergence in function optim
message1	Character vector - translation of conv1
loglik2	Vector of log-likelihoods from fitting model to evaluate presence of residual selection bias
conv2	Numerical vector indicating convergence status for models to evaluate presence of residual selection bias - see parameter convergence in function optim
message2	Character vector - translation of conv2
publprob	Vector of probabilities of publishing the smallest study, used in x-axis of bottom two panels in function plot.copas
pval.rsb	P-values for tests on presence of residual selection bias, plotted in bottom right panel in plot.copas
N.unpubl	Approximate number of studies the model suggests remain unpublished
sm	Effect measure (OR - odds ratio, RR - risk ratio, RD - risk difference, AS - arcsin difference)
call	Call to copas function
x	Details of meta-analysis input into copas function

Author(s)

James Carpenter <James.Carpenter@lshtm.ac.uk>, Guido Schwarzer <sc@imbi.uni-freiburg.de>

References

Carpenter JR, Schwarzer G, Rucker G and Kunstler R (2008), Empirical evaluation shows the Copas selection model provides a useful summary in 80% of meta-analyses. *Revision submitted to Journal of Clinical Epidemiology*.

Copas J (1999), What works?: Selectivity models and meta-analysis. *Journal of the Royal Statistical Society, Series A*, **162**, 95–109.

Copas J, Shi JQ (2000), Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, **1**, 247–262.

Copas JB, Shi JQ (2001), A sensitivity analysis for publication bias in systematic reviews. *Statistical Methods in Medical Research*, **10**, 251–265.

See Also

[plot.copas](#), [summary.copas](#), [metabias](#), [metagen](#), [funnel](#)

Examples

```
# Basic example
#
# Load data
#
data(Fleiss93)
#
# Perform meta-analysis
# (Note event.e indicates events, n.e total in exposed arm;
#      event.c indicates events, n.c total in control arm)
#
meta1 <- metabin(event.e, n.e, event.c, n.c, data=Fleiss93, sm="OR")
summary(meta1)
#
# To perform a basic Copas-selection model analysis
#
cop1 <- copas(meta1)
plot(cop1)
summary(cop1)
#
# Interpretation:
#
# a. The initial meta-analysis shows the fixed and random effects pooled
# ORs differ; consistent with asymmetry in the funnel plot and
# possible selection bias. Both fixed effect and random effects model
# show a significant treatment effect in this dataset.
#
# b. Plotting the copas analysis shows
#
# (i) funnel plot: asymmetry indicates possible selection bias.
```

```

#
# (ii) contour plot treatment effect declines steadily as selection
#       increases (no selection, top right, log OR < -0.12; increasing
#       selection as move to left of plot, log OR rises to -0.03.
#
# (iii) Treatment effect plot suggests that even with no selection,
#       p-value for treatment effect is larger than 0.05 which is
#       different from the result of the usual random effects model
#       (see output of summary(cop1). This difference is due to the
#       use of different methods to estimate the between-study
#       variance: maximum-likelihood in Copas analysis compared to
#       method-of-moments in usual random effects model.
#       The p-value for treatment effect is increasing with
#       increasing selection.
#
# (iv) P-value for residual selection bias plot: this shows that even
#       with no selection bias, the p-value for residual selection bias
#       is non-significant at the 10% level. As expected, as selection
#       increases the p-value for residual selection bias increases too.

# Repeat the same example, setting all the arguments of the copas
# function:
#
cop2 <- copas(meta1,
              gamma0.range=c(-0.5,2.1), # range of gamma0 parameter
              gamma1.range=c(0, 0.08), # range of gamma1 parameter
              ngrid=20,                # specify a 20X20 grid (finer than default)
              levels=c(-0.13, -0.12, -0.1, -0.09, -0.07, -0.05, -0.03),# specify contour lines
              slope=0.2,               # specify slope of 'orthogonal' line in contour plot
              left=FALSE,              # as any selection bias due to missing studies on right
              rho.bound=0.998,         # constrain rho between [-0.998, 0.998]
              silent=FALSE,            # update user on progress
              warn=-1                   # suppress warning messages
              )
plot(cop2)
#
# Print table of results used to draw treatment effect plot:
#
summary(cop2)

```

Description

Meta-analysis on phenobarbital prior to preterm birth for preventing neonatal periventricular haemorrhage

Usage

```
data(Crowther2003)
```

Format

A data frame with the following columns:

study study label

event.e number of periventricular haemorrhages in experimental group

total.e number of observations in experimental group

event.c number of periventricular haemorrhages in control group

total.c number of observations in control group

Source

Crowther CA, Henderson-Smart DJ (2003), Phenobarbital prior to preterm birth for preventing neonatal periventricular haemorrhage. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD000164. DOI: 10.1002/14651858.CD000164

Examples

```
data(Crowther2003)
metabin(event.e, total.e, event.c, total.c,
        data=Crowther2003,
        studlab=study)
```

plot.copas

Display results of Copas selection modelling

Description

Four plots (selectable by 'which') are currently available: (1) funnel plot, (2) contour plot, (3) treatment effect plot, (4) p-value for residual publication bias plot. By default, all plots are provided.

Usage

```
plot.copas(x,
           which=1:4,
           caption=c("Funnel plot", "Contour plot",
                    "Treatment effect plot",
                    "P-value for residual selection bias"),
           xlim.pp=NULL, level=0.95,
           orthogonal.line=TRUE, lines=FALSE,
           sign.rsb=0.1, warn=-1, ...)
```

Arguments

x	An object of class copas, generated by the copas function
which	Specify plots required: 1:4 produces all plots (default); 3 produces plot 3 etc; c(1,3) produces plots 1 and 3, and so on.
caption	Specify plot captions. Note that four captions must be specified even if fewer graphs are displayed (which is the case if the predefined captions are utilised). This must be considered if user-defined captions are provided. Captions corresponding to plots that are not displayed can be left empty. For example, if only plot 3 is selected, we might specify <code>caption=c("", "", "Plot 3", "")</code> .
xlim.pp	A vector of x-axis limits for plots 3 and 4, i.e. for the probability of publishing the study with largest standard deviation. E.g. to specify limits between 0.3 and 0.1 set <code>xlim.pp=c(0.3, 0.1)</code> .
level	The level used to calculate confidence intervals for plot 3 (treatment effect plot) (between 0 and 1).
orthogonal.line	A logical indicating whether the orthogonal line should be displayed in plot 2 (contour plot).
lines	(Diagnostic use only) A logical indicating whether regression lines should be plotted in contour plot. These regression lines attempt to summarise each contour of constant treatment effect by a straight line, prior to calculating the orthogonal line. Regression lines with a positive adjusted R^2 will be printed in green color, others will be printed in red color.
sign.rsb	The significance level for the test of residual selection bias (between 0 and 1).
warn	A number setting the handling of warning messages. It is not uncommon for numerical problems to be encountered during estimation over the grid of (γ_0 , γ_1) values. Usually this does not indicate a serious problem. This option specifies what to do with warning messages. <code>warn=-1</code> : ignore all warnings; <code>warn=0</code> (the default): store warnings till function finishes; if there are less than 10, print them, otherwise print a message saying warning messages were generated; <code>warn=1</code> : print warnings as they occur; <code>warn=2</code> : stop the function when the first warning is generated. For further details see <code>help(options)</code> .
...	other arguments to the function will be ignored (this option only included to conform with R standards)

Details

Takes an object created by the copas function and draws up to four plots to display the results of the Copas selection modelling.

The argument which specifies the plots to be drawn; plot numbers below will be produced by setting `which=1`, etc.

Plot 1: Funnel plot of studies in meta-analysis. Vertical grey line is usual random effects estimate (DerSimonian-Laird method); vertical broken line is fixed effects estimate.

Plot 2: Plot of contours of treatment effect (estimated by the Copas model) as the selection probability varies (the selection probability is a function of γ_0 and γ_1 - see `help(copas)` or the reference below).

Plot 3: Assuming the contours of treatment effect in Plot 2 are locally parallel, the results can be summarised in terms of the probability of publishing the study with the largest standard error. This plot displays the results of doing this, showing how the estimated treatment effect (and $100 \times \text{level}\%$ confidence interval) vary as the probability of publishing the study with the largest standard error decreases.

The three horizontal grey lines are the usual random effects treatment estimate (centre) \pm the $100 \times \text{level}\%$ confidence interval (upper/lower grey lines).

Plot 4: For any degree of selection (i.e. probability of the study with largest SE being published), we can calculate a p-value for the hypothesis that no further selection remains unexplained in the data. These plot displays these p-values against the probability that the study with the largest SE is published.

Under the copas selection model, probabilities of the smallest study being published which correspond to p-values for residual selection bias that are larger than 0.1 are more plausible. The corresponding treatment effect in plot 3 is thus the most plausible under the copas selection model.

Note

In the current version, fine control of the graphics parameters for the individual panels is not possible. However, all the data used to create the plots can be extracted manually from the object created by the copas function (see attributes list for copas) and used to create tailor-made plots.

Author(s)

James Carpenter <James.Carpenter@lshtm.ac.uk>, Guido Schwarzer <sc@imbi.uni-freiburg.de>

References

Carpenter JR, Schwarzer G, Rucker G and Kunstler R (2008), Empirical evaluation shows the Copas selection model provides a useful summary in 80% of meta-analyses. *Revision submitted to Journal of Clinical Epidemiology*.

See Also

[copas](#), [summary.copas](#), [metabias](#), [metagen](#)

Examples

```
# Simple example:
#
# Load data
#
data(Fleiss93)
#
# Perform meta-analysis (outcome measure is OR = odds ratio)
#
meta1 <- metabin(event.e, n.e, event.c, n.c,
                 data=Fleiss93, sm="OR")
#
# Perform copas analysis
#
cop1 <- copas(meta1)
```

```

#
# Plot results
#
plot(cop1)
#
# Only do plots 1 and 2 (without orthogonal line)
#
plot(cop1, which=1:2, orth=FALSE)
#
# Another example showing use of arguments
# Note the use of \n to create a new line in the caption
#
plot(cop1,
      which=3,
      caption=c("", "",
                "Variation in estimated treatment\n effect with selection",
                ""),
      xlim.pp=c(1,0.5))

```

print.copas

Print and summary method for Copas selection model

Description

Print and summary method for objects of class copas.

Usage

```

## S3 method for class 'copas'
print(x, sign.rsb=0.1, ...)

## S3 method for class 'copas'
summary(object, level=0.95, sign.rsb=0.1, ...)

## S3 method for class 'summary.copas'
print(x,
      digits = max(3, .Options$digits - 3),
      ...)

```

Arguments

x	An object of class copas or summary.copas.
object	An object of class copas.
level	The level used to calculate confidence intervals (between 0 and 1).
digits	Minimal number of significant digits, see print.default.
sign.rsb	The significance level for the test of residual selection bias (between 0 and 1).
...	other arguments to the function will be ignored (this option included only to conform with R standards)

Details

The `summary.copas` function prints a summary of a Copas analysis, performed using the function `copas`. It complements the graphical summary of the results, generated using `plot.copas`.

Specifically it prints a table where the:

first column corresponds to the x-axis in plots 3 & 4 from `plot.copas`;

second column corresponds to the treatment effect displayed in plot 3 from `plot.copas`;

third and fourth columns give the confidence intervals for this treatment effect,

fifth column gives the p-value for an overall treatment effect,

sixth column gives the p-value for residual publication bias (the y-axis of plot 4 from `plot.copas` (see `help(plot.copas)` under plot 4 for a further explanation of this p-value))

seventh column gives an approximate estimate of the number of studies the model suggests remain unpublished if the probability of publishing the study with the largest SE is as in column 1.

Below this is displayed the results of the Copas analysis for the smallest degree of selection for which the p-value for evidence of residual selection bias exceeds `sign.rsb` (default: 0.1). This is simply extracted from the corresponding row in the table above.

Lastly, the usual random effects estimate (based on the DerSimonian-Laird method) and 95% confidence interval is printed.

The function `print.copas` prints the summary information above together with the following information:

Range of `gamma0` values used (see `help(copas)`);

Range of `gamma1` values used (see `help(copas)`);

Largest SE of all studies in meta-analysis;

Range of probability publishing trial with largest SE;

The next table gives details relating to the summary of the contour plot. Specifically, it gives details from fitting a straight line to each treatment-contour in the contour plot. Column 1 (headed `level`) shows the treatment-contours; column 2 (`nobs`) shows the number of observations used by the contour plot command within the `copas` function to plot this contour line; column 3 (`adj.r.square`) shows the adjusted r-square from fitting a straight line to this contour; columns 4 & 5 show the slope and its standard error from fitting a straight line to this contour.

Value

A list is returned by the function `summary.copas` with the following elements:

<code>slope</code>	Results for points on orthogonal line (a list with elements <code>TE</code> , <code>seTE</code> , <code>lower</code> , <code>upper</code> , <code>z</code> , <code>p</code> , <code>level</code>).
<code>publprob</code>	Vector of probabilities of publishing the smallest study.
<code>pval.rsb</code>	P-values for tests on presence of residual selection bias
<code>N.unpubl</code>	Approximate number of studies the model suggests remain unpublished
<code>adjust</code>	Result of Copas selection model adjusted for selection bias (a list with elements <code>TE</code> , <code>seTE</code> , <code>lower</code> , <code>upper</code> , <code>z</code> , <code>p</code> , <code>level</code>).
<code>sign.rsb</code>	The significance level for the test of residual selection bias.

pval.rsb.adj	P-value for test on presence of residual selection bias for adjusted effect given in adjust.
N.unpubl.adj	Approximate number of studies the model suggests remain unpublished for adjusted effect given in adjust
random	Results for usual random effects model (a list with elements TE, seTE, lower, upper, z, p, level).
sm	A character string indicating underlying summary measure.
ci.lab	Label for confidence interval.

Author(s)

James Carpenter <James.Carpenter@lshtm.ac.uk>, Guido Schwarzer <sc@imbi.uni-freiburg.de>

See Also

[copas](#), [plot.copas](#) [metabias](#), [metagen](#)

Examples

```
# Load data
#
data(Fleiss93)
#
# Perform meta analysis, effect measure is odds ratio (OR)
#
meta1 <- metabin(event.e, n.e, event.c, n.c,
                 data=Fleiss93, sm="OR")
#
# Perform Copas analysis
#
cop1 <- copas(meta1)
summary(cop1, level=0.95)
```

Index

*Topic **datasets**

Crowther2003, [7](#)

*Topic **hplot**

plot.copas, [8](#)

*Topic **htest**

copas, [2](#)

*Topic **print**

print.copas, [11](#)

copas, [2](#), [10](#), [13](#)

Crowther2003, [7](#)

funnel, [6](#)

metabias, [6](#), [10](#), [13](#)

metagen, [6](#), [10](#), [13](#)

plot.copas, [6](#), [8](#), [13](#)

print.copas, [11](#)

print.summary.copas (print.copas), [11](#)

summary.copas, [6](#), [10](#)

summary.copas (print.copas), [11](#)