Package 'Surrogate'

October 25, 2014

Type Package

Title Evaluation of surrogate endpoints in clinical trials

Version 0.1-3

Date 2014-10-24

Author Wim Van der Elst, Ariel Alonso & Geert Molenberghs

Maintainer Wim Van der Elst <Wim.vanderelst@gmail.com>

Description In a clinical trial, it frequently occurs that the most credible outcome to evaluate the effectiveness of a new therapy (the true endpoint) is difficult to measure. In such a situation, it can be an effective strategy to replace the true endpoint by a biomarker that is easier to measure and that allows for a prediction of the treatment effect on the true endpoint (a surrogate endpoint). The package 'Surrogate' allows for an evaluation of the appropriateness of a candidate surrogate endpoint based on the meta-analytic, information-theoretic, and causal-inference frameworks (for normally distributed endpoints). Part of this software has been developed using funding provided from the European Union's 7th Framework Programme for research, technological development and demonstration under Grant Agreement no 602552.

Depends MASS, nlme, msm, lme4

License GPL (>= 2)

BugReports Wim Van der Elst <Wim.vanderelst@gmail.com>

Repository CRAN

NeedsCompilation no

Date/Publication 2014-10-25 01:16:48

R topics documented:

ARMD *Data of the Age-Related Macular Degeneration Study*

Description

These are the data of a clinical trial involving patients suffering from age-related macular degeneration (ARMD), a condition that involves a progressive loss of vision. A total of 181 patients from 36 centers participated in the trial. Patients' visual acuity was assessed using standardized vision charts. There were two treatment conditions (placebo and interferon-α). The potential surrogate endpoint is the change in the visual acuity at 24 weeks (6 months) after starting treatment. The true endpoint is the change in the visual acuity at 52 weeks (1 year).

Usage

```
data(ARMD)
```

Format

A `data.frame` with 181 observations on 5 variables.

`Id` The Patient ID.

`Center` The center in which the patient was treated.

`Treat` The treatment indicator, coded as -1 = placebo and 1 = interferon-α.

`Diff24` The change in the visual acuity at 24 weeks after starting treatment. This endpoint is a potential surrogate for `Diff52`.

`Diff52` The change in the visual acuity at 52 weeks after starting treatment. This outcome serves as the true endpoint .

`BifixedContCont` *Fits a bivariate fixed-effects model to assess surrogacy in the meta-analytic multiple-trial setting (Continuous-continuous case)*

Description

The function `BifixedContCont` uses the bivariate fixed-effects approach to estimate trial- and individual-level surrogacy when the data of multiple clinical trials are available. The user can specify whether a (weighted or unweighted) full, semi-reduced, or reduced model should be fitted. See the **Details** section below.

Usage

```
BifixedContCont(Dataset, Surr, True, Treat, Trial.ID, Pat.ID, Model=c("Full"),
Weighted=TRUE, Min.Trial.Size=2, Alpha=.05)
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate endpoint values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.
`Treat`	The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.
`Trial.ID`	The name of the variable in `Dataset` that contains the trial ID to which the patient belongs.
`Pat.ID`	The name of the variable in `Dataset` that contains the patient's ID.
`Model`	The type of model that should be fitted, i.e., `Model=c("Full")`, `Model=c("Reduced")`, or `Model=c("SemiReduced")`. See the **Details** section below. Default `Model=c("Full")`.
`Weighted`	Logical. If `TRUE`, then a weighted regression analysis is conducted at stage 2 of the two-stage approach. If `FALSE`, then an unweighted regression analysis is conducted at stage 2 of the two-stage approach. See the **Details** section below. Default `TRUE`.
`Min.Trial.Size`	The minimum number of patients that a trial should contain in order to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.
`Alpha`	The α-level that is used to determine the confidence intervals around R^2_{trial}, R_{trial}, R^2_{indiv} and R_{indiv}. Default 0.05.

Details

When the full bivariate mixed-effects model is fitted to assess surrogacy in the meta-analytic framework (for details, Buyse & Molenberghs, 2000), computational issues often occur. In that situation, the use of simplified model-fitting strategies may be warranted (for details, see Burzykowski et al., 2005; Tibaldi et al., 2003).

The function `BifixedContCont` implements one such strategy, i.e., it uses a two-stage bivariate fixed-effects modelling approach to assess surrogacy. In the first stage of the analysis, a bivariate linear regression model is fitted. When a full or semi-reduced model is requested (by using the argument `Model=c("Full")` or `Model=c("SemiReduced")` in the function call), the following bivariate model is fitted:

$$S_{ij} = \mu_{Si} + \alpha_i Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij},$$

where S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_{Si} and μ_{Ti} are the fixed trial-specific intercepts for S and T, and α_i and β_i are the trial-specific treatment effects on S and T, respectively. When a reduced model is requested (by using the argument `Model=c("Reduced")` in the function call), the following bivariate model is fitted:

$$S_{ij} = \mu_S + \alpha_i Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_T + \beta_i Z_{ij} + \varepsilon_{Tij},$$

where μ_S and μ_T are the common intercepts for S and T (i.e., it is assumed that the intercepts for the surrogate and the true endpoints are identical in all trials). The other parameters are the same as defined above.

In the above models, the error terms ε_{Sij} and ε_{Tij} are assumed to be mean-zero normally distributed with variance-covariance matrix $\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{SS} & \\ \sigma_{ST} & \sigma_{TT} \end{pmatrix}.$$

Based on $\mathbf{\Sigma}$, individual-level surrogacy is quantified as:

$$R^2_{indiv} = \frac{\sigma^2_{ST}}{\sigma_{SS}\sigma_{TT}}.$$

Next, the second stage of the analysis is conducted. When a full model is requested by the user (by using the argument `Model=c("Full")` in the function call), the following model is fitted:

$$\widehat{\beta_i} = \lambda_0 + \lambda_1 \widehat{\mu_{Si}} + \lambda_2 \widehat{\alpha_i} + \varepsilon_i,$$

where the parameter estimates for β_i, μ_{Si}, and α_i are based on the full model that was fitted in stage 1.

When a reduced or semi-reduced model is requested by the user (by using the arguments `Model=c("Reduced")` or `Model=c("SemiReduced")` in the function call), the following model is fitted:

$$\widehat{\beta_i} = \lambda_0 + \lambda_1 \widehat{\alpha_i} + \varepsilon_i.$$

where the parameter estimates for β_i and α_i are based on the semi-reduced or reduced model that was fitted in stage 1.

When the argument `Weighted=FALSE` is used in the function call, the model that is fitted in stage 2 is an unweighted linear regression model. When a weighted model is requested (using the argument `Weighted=TRUE` in the function call), the information that is obtained in stage 1 is weighted according to the number of patients in a trial.

The classical coefficient of determination of the fitted stage 2 model provides an estimate of R^2_{trial}.

Value

An object of class `BifixedContCont` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Results.Stage.1`
The results of stage 1 of the two-stage model fitting approach: a `data.frame` that contains the trial-specific intercepts and treatment effects for the surrogate and the true endpoints (when a full or semi-reduced model is requested), or the trial-specific treatment effects for the surrogate and the true endpoints (when a reduced model is requested).

`Residuals.Stage.1`
A `data.frame` that contains the residuals for the surrogate and true endpoints that are obtained in stage 1 of the analysis (ε_{Sij} and ε_{Tij}).

`Results.Stage.2`
An object of class `lm` (linear model) that contains the parameter estimates of the regression model that is fitted in stage 2 of the analysis.

`Trial.R2` A `data.frame` that contains the trial-level coefficient of determination (R^2_{trial}), its standard error and confidence interval.

`Indiv.R2` A `data.frame` that contains the individual-level coefficient of determination (R^2_{indiv}), its standard error and confidence interval.

`Trial.R` A `data.frame` that contains the trial-level correlation coefficient (R_{trial}), its standard error and confidence interval.

`Indiv.R`	A `data.frame` that contains the individual-level correlation coefficient (R_{indiv}), its standard error and confidence interval.
`Cor.Endpoints`	A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.
`D.Equiv`	The variance-covariance matrix of the trial-specific intercept and treatment effects for the surrogate and true endpoints (when a full or semi-reduced model is fitted, i.e., when `Model=c("Full")` or `Model=c("SemiReduced")` is used in the function call), or the variance-covariance matrix of the trial-specific treatment effects for the surrogate and true endpoints (when a reduced model is fitted, i.e., when `Model=c("Reduced")` is used in the function call). The variance-covariance matrix `D.Equiv` is equivalent to the $\boldsymbol{D}$ matrix that would be obtained when a (full or reduced) bivariate mixed-effect approach is used; see function `BimixedContCont`).
`Sigma`	The 2 by 2 variance-covariance matrix of the residuals (ε_{Sij} and ε_{Tij}).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Burzykowski, T., Molenberghs, G., & Buyse, M. (2005). *The evaluation of surrogate endpoints*. New York: Springer-Verlag.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

Tibaldi, F., Abrahantes, J. C., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., et al., (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. *Journal of Statistical Computation and Simulation, 73,* 643-658.

See Also

`UnifixedContCont`, `UnimixedContCont`, `BimixedContCont`, `plot Meta-Analytic`

Examples

```
## Not run:
# Example 1, based on the ARMD data
data(ARMD)

# Fit a full bivariate fixed-effects model with weighting according to the
# number of patients in stage 2 of the two stage approach to assess surrogacy:
Sur <- BifixedContCont(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
Pat.ID=Id, Model="Full", Weighted=TRUE)

# Obtain a summary of the results
summary(Sur)

# Obtain a graphical representation of the trial- and individual-level surrogacy
```

```
plot(Sur)


# Example 2
# Conduct a surrogacy analysis based on a simulated dataset with 2000 patients,
# 100 trials, and Rindiv=Rtrial=.8
# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=100, R.Trial.Target=.8, R.Indiv.Target=.8,
Seed=123, Model="Reduced")

# Fit a reduced bivariate fixed-effects model with no weighting according to the
# number of patients in stage 2 of the two stage approach to assess surrogacy:
Sur2 <- BifixedContCont(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, , Model="Reduced", Weighted=FALSE)

# Show summary and plots of results:
summary(Sur2)
plot(Sur2, Weighted=FALSE)
## End(Not run)
```

`BimixedContCont` *Fits a bivariate mixed-effects model to assess surrogacy in the meta-analytic multiple-trial setting (Continuous-continuous case)*

Description

The function `BimixedContCont` uses the bivariate mixed-effects approach to estimate trial- and individual-level surrogacy when the data of multiple clinical trials are available. The user can specify whether a full or reduced model should be fitted. See the **Details** section below.

Usage

```
BimixedContCont(Dataset, Surr, True, Treat, Trial.ID, Pat.ID, Model=c("Full"),
Min.Trial.Size=2, Alpha=.05, ...)
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate endpoint values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.
`Treat`	The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.

`Trial.ID`	The name of the variable in `Dataset` that contains the trial ID to which the patient belongs.
`Pat.ID`	The name of the variable in `Dataset` that contains the patient's ID.
`Model`	The type of model that should be fitted, i.e., `Model=c("Full")` or `Model=c("Reduced")`. See the **Details** section below. Default `Model=c("Full")`.
`Min.Trial.Size`	The minimum number of patients that a trial should contain to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.
`Alpha`	The α-level that is used to determine the confidence intervals around R^2_{trial}, R_{trial}, R^2_{indiv} and R_{indiv}. Default 0.05.
`...`	Other arguments to be passed to the function `lmer` (of the R package `lme4`) that is used to fit the geralized linear mixed-effect models in the function `BimixedContCont`. For details, see the lme4 manual.

Details

The function `BimixedContCont` fits a bivariate mixed-effects model to assess surrogacy (for details, see Buyse et al., 2000). In particular, the following mixed-effects model is fitted:

$$S_{ij} = \mu_S + m_{Si} + (\alpha + a_i)Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_T + m_{Ti} + (\beta + b_i)Z_{ij} + \varepsilon_{Tij},$$

where i and j are the trial and subject indicators, S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_S and μ_T are the fixed intercepts for S and T, m_{Si} and m_{Ti} are the corresponding random intercepts, α and β are the fixed treatment effects for S and T, and a_i and b_i are the corresponding random treatment effects, respectively.

The vector of the random effects (i.e., m_{Si}, m_{Ti}, a_i and b_i) is assumed to be mean-zero normally distributed with variance-covariance matrix $\boldsymbol{D}$:

$$\boldsymbol{D} = \begin{pmatrix} d_{SS} & & & \\ d_{ST} & d_{TT} & & \\ d_{Sa} & d_{Ta} & d_{aa} & \\ d_{Sb} & d_{Tb} & d_{ab} & d_{bb} \end{pmatrix}.$$

The trial-level coefficient of determination (i.e., R^2_{trial}) is quantified as:

$$R^2_{trial} = \frac{\begin{pmatrix} d_{Sb} \\ d_{ab} \end{pmatrix}' \begin{pmatrix} d_{SS} & d_{Sa} \\ d_{Sa} & d_{aa} \end{pmatrix}^{-1} \begin{pmatrix} d_{Sb} \\ d_{ab} \end{pmatrix}}{d_{bb}}.$$

The error terms ε_{Sij} and ε_{Tij} are assumed to be mean-zero normally distributed with variance-covariance matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{SS} & \\ \sigma_{ST} & \sigma_{TT} \end{pmatrix}.$$

Based on $\mathbf{\Sigma}$, individual-level surrogacy is quantified as:

$$R^2_{indiv} = \frac{\sigma^2_{ST}}{\sigma_{SS}\sigma_{TT}}.$$

Note

When the full bivariate mixed-effects approach is used to assess surrogacy in the meta-analytic framework (for details, see Buyse & Molenberghs, 2000), computational issues often occur. Such problems mainly occur when the number of trials is low, the number of patients in the different trials is low, and/or when the trial-level heterogeneity is small (Burzykowski et al., 2000).

In that situation, the use of a simplified model-fitting strategy may be warranted (for details, see Burzykowski et al., 2000; Tibaldi et al., 2003).

For example, a reduced bivariate-mixed effect model can be fitted instead of a full model (by using the `Model=c("Reduced")` argument in the function call). In the reduced model, the random-effects structure is simplified (i) by assuming that there is no heterogeneity in the random intercepts, or (ii) by assuming that the covariance between the random intercepts and random treatment effects is zero. Note that under this assumption, the computation of the trial-level coefficient of determination (i.e., R^2_{trial}) simplifies to:

$$R^2_{trial} = \frac{d^2_{ab}}{d_{aa}d_{bb}}.$$

Alternatively, the bivariate mixed-effects model may be abandonned and the user may fit a univariate fixed-effects model, a bivariate fixed-effects model, or a univariate mixed-effects model (for details, see Tibaldi et al., 2003). These models are implemented in the functions `UnifixedContCont`, `BifixedContCont`, and `UnimixedContCont`).

Value

An object of class `BimixedContCont` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Trial.Spec.Results`
A `data.frame` that contains the trial-specific intercepts and treatment effects on the surrogate and the true endpoints when a full model is requested (i.e., $\mu_S +$

m_{Si}, $\mu_T + m_{Ti}$, $\alpha + a_i$, and $\beta + b_i$), or the trial-specific treatment effects on the surrogate and the true endpoints when a reduced model is requested (i.e., $\alpha + a_i$, and $\beta + b_i$). Note that the results that are contained in `Trial.Spec.Results` are equivalent to the results in `Results.Stage.1` that are obtained when the functions `UnifixedContCont`, `UnimixedContCont`, or `BifixedContCont` are used.

`Residuals` A `data.frame` that contains the residuals for the surrogate and true endpoints (ε_{Sij} and ε_{Tij}).

`Fixed.Effect.Pars`
A `data.frame` that contains the fixed intercept and treatment effects for the surrogate and the true endpoints (i.e., μ_S, μ_T, α, and β).

`Random.Effect.Pars`
A `data.frame` that contains the random intercept and treatment effects for the surrogate and the true endpoints (i.e., m_{Si}, m_{Ti}, a_i, and b_i) when a full model is fitted (i.e., when `Model=c("Full")` is used in the function call), or that contains the random treatment effects for the surrogate and the true endpoints (i.e., a_i and b_i) when a reduced model is fitted (i.e., when `Model=c("Reduced")` is used in the function call).

`Trial.R2` A `data.frame` that contains the trial-level coefficient of determination (R^2_{trial}), its standard error and confidence interval.

`Indiv.R2` A `data.frame` that contains the individual-level coefficient of determination (R^2_{indiv}), its standard error and confidence interval.

`Trial.R` A `data.frame` that contains the trial-level correlation coefficient (R_{trial}), its standard error and confidence interval.

`Indiv.R` A `data.frame` that contains the individual-level correlation coefficient (R_{indiv}), its standard error and confidence interval.

`Cor.Endpoints` A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.

`D` The variance-covariance matrix of the random effects (the $\boldsymbol{D}$ matrix), i.e., a 4 by 4 variance-covariance matrix of the random intercept and treatment effects when a full model is fitted (i.e., when `Model=c("Full")` is used in the function call), or a 2 by 2 variance-covariance matrix of the random treatment effects when a reduced model is fitted (i.e., when `Model=c("Reduced")` is used in the function call).

`Sigma` The 2 by 2 variance-covariance matrix of the residuals (ε_{Sij} and ε_{Tij}).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Burzykowski, T., Molenberghs, G., & Buyse, M. (2005). *The evaluation of surrogate endpoints*. New York: Springer-Verlag.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

Tibaldi, F., Abrahantes, J. C., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., et al., (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. *Journal of Statistical Computation and Simulation, 73,* 643-658.

See Also

UnifixedContCont, BifixedContCont, UnimixedContCont, plot Meta-Analytic

Examples

```
## Not run:
# Example 1, based on the Schizo dataset (clinial trial in schizophrenic patients)
data(Schizo)

# When a reduced bivariate mixed-effect model is used to assess surrogacy,
# the conditioning number for the D matrix is very high:
Sur <- BimixedContCont(Dataset=Schizo, Surr=BPRS, True=PANSS, Treat=Treat, Model="Reduced",
Trial.ID=InvestId, Pat.ID=Id)

# Such problems often occur when the total number of patients, the total number
# of trials and/or the trial-level heterogeneity
# of the treatment effects is relatively small

# As an alternative approach to assess surrogacy, consider using the functions
# BifixedContCont, UnifixedContCont or UnimixedContCont in the meta-analytic framework,
# or use the information-theoretic approach


# Example 2
# Conduct a surrogacy analysis based on a simulated dataset with 2000 patients, 50 trials,
# and Rindiv=Rtrial=.8
# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=50, R.Trial.Target=.8, R.Indiv.Target=.8, Seed=123)

# Fit a full bivariate mixed-effects model to assess surrogacy:
Sur2 <- BimixedContCont(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, Model="Full")

# Show a summary and plots of results:
summary(Sur2)
plot(Sur2, Weighted=FALSE)
## End(Not run)
```

`CausalDiagramContCont` *Draws a causal diagram depicting the median correlations between the counterfactuals for a specified range of values of ICA or MICA*

Description

This function provides a diagram that depicts the medians of the correlations between the counterfactuals for a specified range of values of the individual causal association (ICA; ρ_Δ) or the meta-analytic individual causal association (MICA; ρ_M).

Usage

```
CausalDiagramContCont(x, Min=-1, Max=1, Cex.Letters=3, Cex.Corrs=2,
Lines.Rel.Width=TRUE, Col.Pos.Neg=TRUE)
```

Arguments

`x`	An object of class `ICA.ContCont` or `MICA.ContCont`. See `ICA.ContCont` or `MICA.ContCont`.
`Min`	The minimum values of (M)ICA that should be considered. Default=-1.
`Max`	The maximum values of (M)ICA that should be considered. Default=1.
`Cex.Letters`	The size of the symbols for the counterfactuals (S_0, S_1), T_0, T_1). Default=3.
`Cex.Corrs`	The size of the text depicting the median correlations between the counterfactuals. Default=2.
`Lines.Rel.Width`	Logical. When `Lines.Rel.Width=TRUE`, the widths of the lines that represent the correlations between the counterfactuals are relative to the size of the correlations (i.e., a smaller line is used for correlations closer to zero whereas a thicker line is used for (absolute) correlations closer to 1). When `Lines.Rel.Width=FALSE`, the width of all lines representing the correlations between the counterfactuals is identical. Default=`TRUE`.
`Col.Pos.Neg`	Logical. When `Col.Pos.Neg=TRUE`, the color of the lines that represent the correlations between the counterfactuals is red for negative correlations and black for positive ones. When `Col.Pos.Neg=FALSE`, all lines are in black. Default=`TRUE`.

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal inference and meta-analytic paradigms for the validation of surrogate markers.

Van der Elst, W., Alonso, A., & Molenberghs, G. (submitted). An exploration of the relationship between causal inference and meta-analytic measures of surrogacy.

See Also

ICA.ContCont, MICA.ContCont

Examples

```
## Not run:
# Generate the vector of ICA values when rho_T0S0=.91, rho_T1S1=.91, and when the
# grid of values {0, .1, ..., 1} is considered for the correlations
# between the counterfactuals:
SurICA <- ICA.ContCont(T0S0=.95, T1S1=.91, T0T1=seq(0, 1, by=.1), T0S1=seq(0, 1, by=.1),
T1S0=seq(0, 1, by=.1), S0S1=seq(0, 1, by=.1))

#obtain a plot of ICA

# Obtain a causal diagram that provides the medians of the
# correlations between the counterfactuals for the range
# of ICA values between .9 and 1 (i.e., which assumed
# correlations between the counterfactuals lead to a
# high ICA?)
CausalDiagramContCont(SurICA, Min=.9, Max=1)

# Same, for low values of ICA
CausalDiagramContCont(SurICA, Min=0, Max=.5)

## End(Not run)
```

`FixedContContIT` *Fits (univariate) fixed-effect models to assess surrogacy in the continuous-continuous case based on the Information-Theoretic framework*

Description

The function `FixedContContIT` uses the information-theoretic approach (Alonso & Molenberghs, 2007) to estimate trial- and individual-level surrogacy based on fixed-effect models when both S and T are continuous variables. The user can specify whether a (weighted or unweighted) full, semi-reduced, or reduced model should be fitted. See the **Details** section below.

Usage

```
FixedContContIT(Dataset, Surr, True, Treat, Trial.ID, Pat.ID,
Model=c("Full"), Weighted=TRUE, Min.Trial.Size=2,
Alpha=.05, Number.Bootstraps=500)
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate endpoint values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.

`Treat` The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.

`Trial.ID` The name of the variable in `Dataset` that contains the trial ID to which the patient belongs.

`Pat.ID` The name of the variable in `Dataset` that contains the patient's ID.

`Model` The type of model that should be fitted, i.e., `Model=c("Full")`, `Model=c("Reduced")`, or `Model=c("SemiReduced")`. See the **Details** section below. Default `Model=c("Full")`.

`Weighted` Logical. In practice it is often the case that different trials (or other clustering units) have different sample sizes. Univariate models are used to assess surrogacy in the information-theoretic approach, so it can be useful to adjust for heterogeneity in information content between the trial-specific contributions (particularly when trial-level surrogacy measures are of primary interest and when the heterogeneity in sample sizes is large). If `Weighted=TRUE`, weighted regression models are fitted. If `Weighted=FALSE`, unweighted regression analyses are conducted. See the **Details** section below. Default `TRUE`.

`Min.Trial.Size` The minimum number of patients that a trial should contain to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.

`Alpha` The α-level that is used to determine the confidence intervals around R_h^2 and R_{ht}^2. Default 0.05.

`Number.Bootstraps`
The standard error and confidence interval for R_h^2 is determined based on a bootstrap procedure. `Number.Bootstraps` specifies the number of bootstrap samples that are used. Default 500.

Details

Individual-level surrogacy

The following univariate generalised linear models are fitted:

$$g_T(E(T_{ij})) = \mu_{Ti} + \beta_i Z_{ij},$$

$$g_T(E(T_{ij}|S_{ij})) = \gamma_{0i} + \gamma_{1i} Z_{ij} + \gamma_{2i} S_{ij},$$

where i and j are the trial and subject indicators, g_T is an appropriate link function (i.e., an identity link when a continuous true endpoint is considered), S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, and Z_{ij} is the treatment indicator for subject j in trial i. μ_{Ti} and β_i are the trial-specific intercepts and treatment-effects on the true endpoint in trial i. γ_{0i} and γ_{1i} are the trial-specific intercepts and treatment-effects on the true endpoint in trial i after accounting for the effect of the surrogate endpoint.

The -2 log likelihood values of the previous models in each of the i trials (i.e., L_{1i} and L_{2i}, respectively) are subsequently used to compute individual-level surrogacy based on the so-called Variance Reduction Factor (VFR; for details, see Alonso & Molenberghs, 2007):

$$R_h^2 = 1 - \frac{1}{N}\sum_i exp\left(-\frac{L_{2i} - L_{1i}}{n_i}\right),$$

where N is the number of trials and n_i is the number of patients within trial i.

When it can be assumed (i) that the treatment-corrected association between the surrogate and the true endpoint is constant across trials, or (ii) when all data come from a single clinical trial (i.e., when $N = 1$), the previous expression simplifies to:

$$R_h^2 = 1 - exp\left(-\frac{L_2 - L_1}{N}\right).$$

Trial-level surrogacy

When a full or semi-reduced model is requested (by using the argument `Model=c("Full")` or `Model=c("SemiReduced")` in the function call), trial-level surrogacy is assessed by fitting the following univariate models:

$$S_{ij} = \mu_{Si} + \alpha_i Z_{ij} + \varepsilon_{Sij}, (1)$$

$$T_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij}, (1)$$

where i and j are the trial and subject indicators, S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_{Si} and μ_{Ti} are the fixed trial-specific intercepts for S and T, and α_i and β_i are the fixed trial-specific treatment effects on S and T, respectively. The error terms ε_{Sij} and ε_{Tij} are assumed to be independent.

When a reduced model is requested by the user (by using the argument `Model=c("Reduced")` in the function call), the following univariate models are fitted:

$$S_{ij} = \mu_S + \alpha_i Z_{ij} + \varepsilon_{Sij}, (2)$$

$$T_{ij} = \mu_T + \beta_i Z_{ij} + \varepsilon_{Tij}, (2)$$

where μ_S and μ_T are the common intercepts for S and T. The other parameters are the same as defined above, and ε_{Sij} and ε_{Tij} are again assumed to be independent.

When the user requested a full model approach (by using the argument `Model=c("Full")` in the function call, i.e., when models (1) were fitted), the following model is subsequently fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\mu_{Si}} + \lambda_2 \widehat{\alpha}_i + \varepsilon_i, (3)$$

where the parameter estimates for β_i, μ_{Si}, and α_i are based on models (1) (see above). When a weighted model is requested (using the argument `Weighted=TRUE` in the function call), model (3) is a weighted regression model (with weights based on the number of observations in trial i). The -2 log likelihood value of the (weighted or unweighted) model (3) (L_1) is subsequently compared to the -2 log likelihood value of an intercept-only model ($\widehat{\beta}_i = \lambda_3$; L_0), and R_{ht}^2 is computed based based on the Variance Reduction Factor (for details, see Alonso & Molenberghs, 2007):

$$R^2_{ht} = 1 - exp\left(-\frac{L_1 - L_0}{N}\right),$$

where N is the number of trials.

When a semi-reduced or reduced model is requested (by using the argument `Model=c("SemiReduced")` or `Model=c("Reduced")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\alpha}_i + \varepsilon_i,$$

where the parameter estimates for β_i and α_i are based on models (1) when a semi-reduced model is fitted or on models (2) when a reduced model is fitted. The -2 log likelihood value of this (weighted or unweighted) model (L_1) is subsequently compared to the -2 log likelihood value of an intercept-only model ($\widehat{\beta}_i = \lambda_3$; L_0), and R^2_{ht} is computed based on the reduction in the likelihood (as described above).

Value

An object of class `FixedContContIT` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Trial.Spec.Results` A `data.frame` that contains the trial-specific intercepts and treatment effects for the surrogate and the true endpoints (when a full or semi-reduced model is requested), or the trial-specific treatment effects for the surrogate and the true endpoints (when a reduced model is requested).

`R2ht` A `data.frame` that contains the trial-level surrogacy estimate and its confidence interval.

`R2h` A `data.frame` that contains the individual-level surrogacy estimate and its confidence interval.

`R2h.Single` A `data.frame` that contains the individual-level surrogacy estimate and its confidence interval under the assumption that the treatment-corrected association between the surrogate and the true endpoints is constant across trials or when all data come from a single clinical trial.

`Boot.CI`	A `data.frame` that contains the bootstrapped `R2h.Single` values.
`Cor.Endpoints`	A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.
`Residuals`	A `data.frame` that contains the residuals for the surrogate and true endpoints (ε_{Sij} and ε_{Tij}) that are obtained when models (1) or models (2) are fitted (see the **Details** section above).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A, & Molenberghs, G. (2007). Surrogate marker evaluation from an information theory perspective. *Biometrics, 63,* 180-186.

See Also

`MixedContContIT`, `plot Information-Theoretic`

Examples

```
## Not run:
# Example 1
# Based on the ARMD data

data(ARMD)
# Assess surrogacy based on a full fixed-effect model
# in the information-theoretic framework:
Sur <- FixedContContIT(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
Pat.ID=Id, Model="Full", Number.Bootstraps=50)
# Obtain a summary of the results:
summary(Sur)

# Example 2
# Conduct an analysis based on a simulated dataset with 2000 patients, 100 trials,
# and Rindiv=Rtrial=.8

# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=100, R.Trial.Target=.8, R.Indiv.Target=.8,
             Seed=123, Model="Full")
# Assess surrogacy based on a full fixed-effect model
# in the information-theoretic framework:
Sur2 <- FixedContContIT(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, Model="Full", Number.Bootstraps=50)

# Show a summary of the results:
summary(Sur2)
## End(Not run)
```

`ICA.ContCont` *Assess surrogacy in the causal-inference single-trial setting (Individual Causal Association, ICA) in the Continuous-continuous case*

Description

The function `ICA.ContCont` quantifies surrogacy in the single-trial causal-inference framework. See **Details** below.

Usage

```
ICA.ContCont(T0S0, T1S1, T0T0=1, T1T1=1, S0S0=1, S1S1=1, T0T1=seq(-1, 1, by=.1),
T0S1=seq(-1, 1, by=.1), T1S0=seq(-1, 1, by=.1), S0S1=seq(-1, 1, by=.1))
```

Arguments

`T0S0`	A scalar or vector that specifies the correlation(s) between the surrogate and the true endpoint in the control treatment condition that should be considered in the computation of ρ_Δ.
`T1S1`	A scalar or vector that specifies the correlation(s) between the surrogate and the true endpoint in the experimental treatment condition that should be considered in the computation of ρ_Δ.
`T0T0`	A scalar that specifies the variance of the true endpoint in the control treatment condition that should be considered in the computation of ρ_Δ. Default 1.
`T1T1`	A scalar that specifies the variance of the true endpoint in the experimental treatment condition that should be considered in the computation of ρ_Δ. Default 1.
`S0S0`	A scalar that specifies the variance of the surrogate endpoint in the control treatment condition that should be considered in the computation of ρ_Δ. Default 1.
`S1S1`	A scalar that specifies the variance of the surrogate endpoint in the experimental treatment condition that should be considered in the computation of ρ_Δ. Default 1.
`T0T1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and T1 that should be considered in the computation of ρ_Δ. Default `seq(-1, 1, by=.1)`, i.e., the values $-1, -0.9, -0.8, \ldots, 1$.
`T0S1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and S1 that should be considered in the computation of ρ_Δ. Default `seq(-1, 1, by=.1)`.
`T1S0`	A scalar or vector that contains the correlation(s) between the counterfactuals T1 and S0 that should be considered in the computation of ρ_Δ. Default `seq(-1, 1, by=.1)`.
`S0S1`	A scalar or vector that contains the correlation(s) between the counterfactuals S0 and S1 that should be considered in the computation of ρ_Δ. Default `seq(-1, 1, by=.1)`.

Details

Based on the causal-inference framework, it is assumed that each subject *j* has four counterfactuals (or potential outcomes), i.e., T_{0j}, T_{1j}, S_{0j}, and S_{1j}. Let T_{0j} and T_{1j} denote the counterfactuals for the true endpoint (T) under the control ($Z = 0$) and the experimental ($Z = 1$) treatments of subject *j*, respectively. Similarly, S_{0j} and S_{1j} denote the corresponding counterfactuals for the surrogate endpoint (S) under the control and experimental treatments, respectively. The individual causal effects of Z on T and S for a given subject *j* are then defined as $\Delta_{T_j} = T_{1j} - T_{0j}$ and $\Delta_{S_j} = S_{1j} - S_{0j}$, respectively.

In the single-trial causal-inference framework, surrogacy can be quantified as the correlation between the individual causal effects of Z on S and T (for details, see Alonso et al., submitted):

$$\rho_{\Delta} = \rho(\Delta_{T_j}, \Delta_{S_j}) = \frac{\sqrt{\sigma_{S_0S_0}\sigma_{T_0T_0}}\rho_{S_0T_0} + \sqrt{\sigma_{S_1S_1}\sigma_{T_1T_1}}\rho_{S_1T_1} - \sqrt{\sigma_{S_0S_0}\sigma_{T_1T_1}}\rho_{S_0T_1} - \sqrt{\sigma_{S_1S_1}\sigma_{T_0T_0}}\rho_{S_1T_0}}{\sqrt{(\sigma_{T_0T_0} + \sigma_{T_1T_1} - 2\sqrt{\sigma_{T_0T_0}\sigma_{T_1T_1}}\rho_{T_0T_1})(\sigma_{S_0S_0} + \sigma_{S_1S_1} - 2\sqrt{\sigma_{S_0S_0}\sigma_{S_1S_1}}\rho_{S_0S_1})}},$$

where the correlations $\rho_{S_0T_1}$, $\rho_{S_1T_0}$, $\rho_{T_0T_1}$, and $\rho_{S_0S_1}$ are not estimable. It is thus warranted to conduct a sensitivity analysis (by considering vectors of possible values for the correlations between the counterfactuals – rather than point estimates).

When the user specifies a vector of values that should be considered for one or more of the counterfactual correlations in the above expression, the function `ICA.ContCont` constructs all possible matrices that can be formed as based on these values, identifies the matrices that are positive definite (i.e., valid correlation matrices), and computes ρ_{Δ} for each of these matrices. The obtained vector of ρ_{Δ} values can subsequently be used to examine (i) the impact of different assumptions regarding the correlations between the counterfactuals on the results (see also `plot Causal-Inference`), and (ii) the extent to which proponents of the causal-inference and meta-analytic frameworks will reach the same conclusion with respect to the appropriateness of the candidate surrogate at hand.

The function `ICA` also generates output that is useful to examine the plausibility of finding a good surrogate endpoint (see `GoodSurr` in the **Value** section below). For details, see Alonso et al. (submitted).

Notes

A single ρ_{Δ} value is obtained when all correlations in the function call are scalars.

Value

An object of class `ICA.ContCont` with components,

`Total.Num.Matrices`
: An object of class `numeric` that contains the total number of matrices that can be formed as based on the user-specified correlations in the function call.

`Pos.Def`
: A `data.frame` that contains the positive definite matrices that can be formed based on the user-specified correlations. These matrices are used to compute the vector of the ρ_{Δ} values.

`ICA`
: A scalar or vector that contains the individual causal association (ICA; ρ_{Δ}) value(s).

`GoodSurr`
: A `data.frame` that contains the ICA (ρ_{Δ}), σ_{Δ_T}, and δ.

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal-inference and meta-analytic paradigms for the validation of surrogate markers.

See Also

MICA.ContCont, Single.Trial.RE.AA, plot Causal-Inference

Examples

```
## Not run:
# Generate the vector of ICA.ContCont values when rho_T0S0=rho_T1S1=.95,
# sigma_T0T0=90, sigma_T1T1=100,sigma_ S0S0=10, sigma_S1S1=15, and
# the grid of values {-1, -.9, ..., 1} is considered for the correlations
# between the counterfactuals:
SurICA <- ICA.ContCont(T0S0=.95, T1S1=.95, T0T0=90, T1T1=100, S0S0=10, S1S1=15,
T0T1=seq(-1, 1, by=.1), T0S1=seq(-1, 1, by=.1), T1S0=seq(-1, 1, by=.1),
S0S1=seq(-1, 1, by=.1))

# Examine and plot the vector of generated ICA values:
summary(SurICA)
\dontrun{plot(SurICA)}

# Obtain the positive definite matrices than can be formed as based on the
# specified (vectors) of the correlations (these matrices are used to
# compute the ICA values)
SurICA$Pos.Def
## End(Not run)
```

LongToWide — *Reshapes a dataset from the 'long' format (i.e., multiple lines per patient) into the 'wide' format (i.e., one line per patient)*

Description

Reshapes a dataset that is in the 'long' format into the 'wide' format. The dataset should contain a single surrogate endpoint and a single true endpoint value per subject.

Usage

```
LongToWide(Dataset, OutcomeIndicator, IdIndicator, TreatIndicator, OutcomeValue)
```

Arguments

Dataset	A data.frame in the 'long' format that contains (at least) five columns, i.e., one that contains the subject ID, one that contains the trial ID, one that contains the endpoint indicator, one that contains the treatment indicator, and one that contains the endpoint values.
OutcomeIndicator	The name of the variable in Dataset that contains the indicator that distinguishes between the surrogate and true endpoints.
IdIndicator	The name of the variable in Dataset that contains the subject ID.
TreatIndicator	The name of the variable in Dataset that contains the treatment indicator. For the subsequent surrogacy analyses, the treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.
OutcomeValue	The name of the variable in Dataset that contains the endpoint values.

Value

A data.frame in the 'wide' format, i.e., a data.frame that contains one line per subject. Each line contains a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.

Author(s)

Wim Van der Elst, Ariel Alonso, and Geert Molenberghs

Examples

```
# Generate a dataset in the 'long' format that contains
# S and T values for 100 patients
Outcome <- rep(x=c(0, 1), times=100)
ID <- rep(seq(1:100), each=2)
Treat <- rep(seq(c(0,1)), each=100)
Outcomes <- as.numeric(matrix(rnorm(1*200, mean=100, sd=10),
                                    ncol=200))
Data <- data.frame(cbind(Outcome, ID, Treat, Outcomes))

# Reshapes the Data object
LongToWide(Dataset=Data, OutcomeIndicator=Outcome, IdIndicator=ID,
           TreatIndicator=Treat, OutcomeValue=Outcomes)
```

MICA.ContCont *Assess surrogacy in the causal-inference multiple-trial setting (Meta-analytic Individual Causal Association; MICA) in the continuous-continuous case*

Description

The function `MICA.ContCont` quantifies surrogacy in the multiple-trial causal-inference framework. See **Details** below.

Usage

```
MICA.ContCont(Trial.R, D.aa, D.bb, T0S0, T1S1, T0T0=1, T1T1=1, S0S0=1, S1S1=1,
T0T1=seq(-1, 1, by=.1), T0S1=seq(-1, 1, by=.1), T1S0=seq(-1, 1, by=.1),
S0S1=seq(-1, 1, by=.1))
```

Arguments

`Trial.R`	A scalar that specifies the trial-level correlation coefficient (i.e., R_{trial}) that should be used in the computation of ρ_M.
`D.aa`	A scalar that specifies the between-trial variance of the treatment effects on the surrogate endpoint (i.e., d_{aa}) that should be used in the computation of ρ_M.
`D.bb`	A scalar that specifies the between-trial variance of the treatment effects on the true endpoint (i.e., d_{bb}) that should be used in the computation of ρ_M.
`T0S0`	A scalar or vector that specifies the correlation(s) between the surrogate and the true endpoint in the control treatment condition that should be considered in the computation of ρ_M.
`T1S1`	A scalar or vector that specifies the correlation(s) between the surrogate and the true endpoint in the experimental treatment condition that should be considered in the computation of ρ_M.
`T0T0`	A scalar that specifies the variance of the true endpoint in the control treatment condition that should be considered in the computation of ρ_M. Default 1.
`T1T1`	A scalar that specifies the variance of the true endpoint in the experimental treatment condition that should be considered in the computation of ρ_M. Default 1.
`S0S0`	A scalar that specifies the variance of the surrogate endpoint in the control treatment condition that should be considered in the computation of ρ_M. Default 1.
`S1S1`	A scalar that specifies the variance of the surrogate endpoint in the experimental treatment condition that should be considered in the computation of ρ_M. Default 1.
`T0T1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and T1 that should be considered in the computation of ρ_M. Default `seq(-1, 1, by=.1)`, i.e., the values $-1, -0.9, -0.8, \ldots, 1$.
`T0S1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and S1 that should be considered in the computation of ρ_M. Default `seq(-1, 1, by=.1)`.
`T1S0`	A scalar or vector that contains the correlation(s) between the counterfactuals T1 and S0 that should be considered in the computation of ρ_M. Default `seq(-1, 1, by=.1)`.
`S0S1`	A scalar or vector that contains the correlation(s) between the counterfactuals S0 and S1 that should be considered in the computation of ρ_M. Default `seq(-1, 1, by=.1)`.

Details

Based on the causal-inference framework, it is assumed that each subject *j* in trial *i* has four counterfactuals (or potential outcomes), i.e., T_{0ij}, T_{1ij}, S_{0ij}, and S_{1ij}. Let T_{0ij} and T_{1ij} denote the counterfactuals for the true endpoint (T) under the control ($Z = 0$) and the experimental ($Z = 1$) treatments of subject *j* in trial *i*, respectively. Similarly, S_{0ij} and S_{1ij} denote the corresponding counterfactuals for the surrogate endpoint (S) under the control and experimental treatments of subject *j* in trial *i*, respectively. The individual causal effects of Z on T and S for a given subject *j* in trial *i* are then defined as $\Delta_{T_{ij}} = T_{1ij} - T_{0ij}$ and $\Delta_{S_{ij}} = S_{1ij} - S_{0ij}$, respectively.

In the multiple-trial causal-inference framework, surrogacy can be quantified as the correlation between the individual causal effects of Z on S and T (for details, see Alonso et al., submitted):

$$\rho_M = \rho(\Delta_{Tij},\, \Delta_{Sij}) = \frac{\sqrt{d_{bb}d_{aa}}R_{trial} + \sqrt{V(\varepsilon_{\Delta Tij})V(\varepsilon_{\Delta Sij})}\rho_{\Delta}}{\sqrt{V(\Delta_{Tij})V(\Delta_{Sij})}},$$

where

$$V(\varepsilon_{\Delta Tij}) = \sigma_{T_0T_0} + \sigma_{T_1T_1} - 2\sqrt{\sigma_{T_0T_0}\sigma_{T_1T_1}}\rho_{T_0T_1},$$

$$V(\varepsilon_{\Delta Sij}) = \sigma_{S_0S_0} + \sigma_{S_1S_1} - 2\sqrt{\sigma_{S_0S_0}\sigma_{S_1S_1}}\rho_{S_0S_1},$$

$$V(\Delta_{Tij}) = d_{bb} + \sigma_{T_0T_0} + \sigma_{T_1T_1} - 2\sqrt{\sigma_{T_0T_0}\sigma_{T_1T_1}}\rho_{T_0T_1},$$

$$V(\Delta_{Sij}) = d_{aa} + \sigma_{S_0S_0} + \sigma_{S_1S_1} - 2\sqrt{\sigma_{S_0S_0}\sigma_{S_1S_1}}\rho_{S_0S_1}.$$

The correlations between the counterfactuals (i.e., $\rho_{S_0T_1}$, $\rho_{S_1T_0}$, $\rho_{T_0T_1}$, and $\rho_{S_0S_1}$) are not identifiable from the data. It is thus warranted to conduct a sensitivity analysis (by considering vectors of possible values for the correlations between the counterfactuals – rather than point estimates).

When the user specifies a vector of values that should be considered for one or more of the correlations that are involved in the computation of ρ_M, the function `MICA.ContCont` constructs all possible matrices that can be formed as based on the specified values, identifies the matrices that are positive definite (i.e., valid correlation matrices), and computes ρ_M for each of these matrices. An examination of the vector of the obtained ρ_M values allows for a straightforward examination of the impact of different assumptions regarding the correlations between the counterfactuals on the results (see also `plot Causal-Inference`), and the extent to which proponents of the causal-inference and meta-analytic frameworks will reach the same conclusion with respect to the appropriateness of the candidate surrogate at hand.

Notes

A single ρ_M value is obtained when all correlations in the function call are scalars.

Value

An object of class `MICA.ContCont` with components,

`Total.Num.Matrices`
: An object of class `numeric` which contains the total number of matrices that can be formed as based on the user-specified correlations.

`Pos.Def`
: A `data.frame` that contains the positive definite matrices that can be formed based on the user-specified correlations. These matrices are used to compute the vector of the ρ_M values.

ICA	A scalar or vector of the ρ_Δ values.
MICA	A scalar or vector of the ρ_M values.

Warning

The theory that relates the causal-inference and the meta-analytic frameworks in the multiple-trial setting (as developped in Alonso et al., submitted) assumes that a reduced or semi-reduced modelling approach is used in the meta-analytic framework. Thus R_{trial}, d_{aa} and d_{bb} should be estimated based on a reduced model (i.e., using the `Model=c("Reduced")` argument in the functions `UnifixedContCont`, `UnimixedContCont`, `BifixedContCont`, or `BimixedContCont`) or based on a semi-reduced model (i.e., using the `Model=c("SemiReduced")` argument in the functions `UnifixedContCont`, `UnimixedContCont`, or `BifixedContCont`).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal-inference and meta-analytic paradigms for the validation of surrogate markers.

See Also

`ICA.ContCont`, `plot Causal-Inference`, `UnifixedContCont`, `UnimixedContCont`, `BifixedContCont`, `BimixedContCont`

Examples

```
## Not run:
# Generate the vector of MICA values when R_trial=.8, rho_T0S0=rho_T1S1=.8,
# sigma_T0T0=90, sigma_T1T1=100,sigma_ S0S0=10, sigma_S1S1=15, D.aa=5, D.bb=10,
# and when the grid of values {-1, -0.9, ..., 1} is considered for the
# correlations between the counterfactuals:
SurMICA <- MICA.ContCont(Trial.R=.80, D.aa=5, D.bb=10, T0S0=.8, T1S1=.8,
T0T0=90, T1T1=100, S0S0=10, S1S1=15, T0T1=seq(-1, 1, by=.1),
T0S1=seq(-1, 1, by=.1), T1S0=seq(-1, 1, by=.1), S0S1=seq(-1, 1, by=.1))

# Examine and plot the vector of the generated MICA values:
summary(SurMICA)
plot(SurMICA)


# Same analysis, but now assume that D.aa=.5 and D.bb=.1:
SurMICA <- MICA.ContCont(Trial.R=.80, D.aa=.5, D.bb=.1, T0S0=.8, T1S1=.8,
T0T0=90, T1T1=100, S0S0=10, S1S1=15, T0T1=seq(-1, 1, by=.1),
T0S1=seq(-1, 1, by=.1), T1S0=seq(-1, 1, by=.1), S0S1=seq(-1, 1, by=.1))

# Examine and plot the vector of the generated MICA values:
summary(SurMICA)
```

```
plot(SurMICA)
## End(Not run)
```

`MinSurrContCont` *Examine the plausibility of finding a good surrogate endpoint in the Continuous-continuous case*

Description

The function `MinSurrContCont` examines the plausibility of finding a good surrogate endpoint in the continuous-continuous setting. For details, see Alonso et al. (submitted).

Usage

```
MinSurrContCont(T0T0, T1T1, Delta, T0T1=seq(from=0, to=1, by=.01))
```

Arguments

`T0T0`	A scalar that specifies the variance of the true endpoint in the control treatment condition.
`T1T1`	A scalar that specifies the variance of the true endpoint in the experimental treatment condition.
`Delta`	A scalar that specifies an upper bound for the prediction mean squared error when predicting the individual causal effect of the treatment on the true endpoint based on the individual causal effect of the treatment on the surrogate.
`T0T1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and T1 that should be considered in the computation of ρ^2_{min}. Default `seq(0, 1, by=.1)`, i.e., the values $0, 0.10, 0.20, \ldots, 1$.

Value

An object of class `MinSurrContCont` with components,

`T0T1`	A scalar or vector that contains the correlation(s) between the counterfactuals T0 and T1 that were considered (i.e., $\rho_{T_0T_1}$).
`Sigma.Delta.T`	A scalar or vector that contains the standard deviations of the individual causal treatment effects on the true endpoint as a function of $\rho_{T_0T_1}$.
`Rho2.Min`	A scalar or vector that contains the ρ^2_{min} values as a function of $\rho_{T_0T_1}$.

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal-inference and meta-analytic paradigms for the validation of surrogate markers.

See Also

ICA.ContCont, plot Causal-Inference, plot MinSurrContCont

Examples

```
# Assess the plausibility of finding a good surrogate when
# sigma_T0T0 = sigma_T1T1 = 8 and Delta = 1
MinSurr <- MinSurrContCont(T0T0 = 8, T1T1 = 8, Delta = 1)
summary(MinSurr)
## Not run: plot(MinSurr)
```

MixedContContIT — *Fits (univariate) mixed-effect models to assess surrogacy in the continuous-continuous case based on the Information-Theoretic framework*

Description

The function MixedContContIT uses the information-theoretic approach (Alonso & Molenberghs, 2007) to estimate trial- and individual-level surrogacy based on mixed-effect models when both S and T are continuous endpoints. The user can specify whether a (weighted or unweighted) full, semi-reduced, or reduced model should be fitted. See the **Details** section below.

Usage

```
MixedContContIT(Dataset, Surr, True, Treat, Trial.ID, Pat.ID,
Model=c("Full"), Weighted=TRUE, Min.Trial.Size=2,
Alpha=.05, Number.Bootstraps=500, ...)
```

Arguments

Dataset	A data.frame that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
Surr	The name of the variable in Dataset that contains the surrogate endpoint values.
True	The name of the variable in Dataset that contains the true endpoint values.
Treat	The name of the variable in Dataset that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.
Trial.ID	The name of the variable in Dataset that contains the trial ID to which the patient belongs.
Pat.ID	The name of the variable in Dataset that contains the patient's ID.
Model	The type of model that should be fitted, i.e., Model=c("Full"), Model=c("Reduced"), or Model=c("SemiReduced"). See the **Details** section below. Default Model=c("Full").

`Weighted` Logical. In practice it is often the case that different trials (or other clustering units) have different sample sizes. Univariate models are used to assess surrogacy in the information-theoretic approach, so it can be useful to adjust for heterogeneity in information content between the trial-specific contributions (particularly when trial-level surrogacy measures are of primary interest and when the heterogeneity in sample sizes is large). If `Weighted=TRUE`, weighted regression models are fitted. If `Weighted=FALSE`, unweighted regression analyses are conducted. See the **Details** section below. Default `TRUE`.

`Min.Trial.Size` The minimum number of patients that a trial should contain to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.

`Alpha` The α-level that is used to determine the confidence intervals around R_h^2 and R_{ht}^2. Default 0.05.

`Number.Bootstraps`
The standard error and confidence interval for R_h^2 is determined based on a bootstrap procedure. `Number.Bootstraps` specifies the number of bootstrap samples that are used. Default 500.

`...` Other arguments to be passed to the function `lmer` (of the R package `lme4`) that is used to fit the geralized linear mixed-effect models in the function `BimixedContCont`. For details, see the lme4 manual.

Details

Individual-level surrogacy

The following generalised linear mixed-effect models are fitted:

$$g_T(E(T_{ij})) = \mu_T + m_{Ti} + \beta Z_{ij} + b_i Z_{ij},$$

$$g_T(E(T_{ij}|S_{ij})) = \theta_0 + c_{Ti} + \theta_1 Z_{ij} + a_i Z_{ij} + \theta_{2i} S_{ij},$$

where i and j are the trial and subject indicators, g_T is an appropriate link function (i.e., an identity link when a continuous true endpoint is considered), S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, and Z_{ij} is the treatment indicator for subject j in trial i. μ_T and β are a fixed intercept and a fixed treatment-effect on the true endpoint, while m_{Ti} and b_i are the corresponding random effects. θ_0 and θ_1 are the fixed intercept and the fixed treatment effect on the true endpoint after accounting for the effect of the surrogate endpoint, and c_{Ti} and a_i are the corresponding random effects.

The -2 log likelihood values of the previous models (i.e., L_1 and L_2, respectively) are subsequently used to compute individual-level surrogacy (based on the so-called Variance Reduction Factor, VFR; for details, see Alonso & Molenberghs, 2007):

$$R_h^2 = 1 - exp\left(-\frac{L_2 - L_1}{N}\right),$$

where N is the number of trials.

Trial-level surrogacy

When a full or semi-reduced model is requested (by using the argument `Model=c("Full")` or `Model=c("SemiReduced")` in the function call), trial-level surrogacy is assessed by fitting the following mixed models:

$$S_{ij} = \mu_S + m_{Si} + (\alpha + a_i)Z_{ij} + \varepsilon_{Sij}, (1)$$

$$T_{ij} = \mu_T + m_{Ti} + (\beta + b_i)Z_{ij} + \varepsilon_{Tij}, (1)$$

where i and j are the trial and subject indicators, S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_S and μ_T are the fixed intercepts for S and T, m_{Si} and m_{Ti} are the corresponding random intercepts, α and β are the fixed treatment effects on S and T, and a_i and b_i are the corresponding random effects. The error terms ε_{Sij} and ε_{Tij} are assumed to be independent.

When a reduced model is requested by the user (by using the argument `Model=c("Reduced")` in the function call), the following univariate models are fitted:

$$S_{ij} = \mu_S + (\alpha + a_i)Z_{ij} + \varepsilon_{Sij}, (2)$$

$$T_{ij} = \mu_T + (\beta + b_i)Z_{ij} + \varepsilon_{Tij}, (2)$$

where μ_S and μ_T are the common intercepts for S and T. The other parameters are the same as defined above, and ε_{Sij} and ε_{Tij} are again assumed to be independent.

When the user requested that a full model approach is used (by using the argument `Model=c("Full")` in the function call, i.e., when models (1) were fitted), the following model is subsequently fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1\widehat{\mu_{Si}} + \lambda_2\widehat{\alpha}_i + \varepsilon_i, (3)$$

where the parameter estimates for β_i, μ_{Si}, and α_i are based on models (1) (see above). When a weighted model is requested (using the argument `Weighted=TRUE` in the function call), model (3) is a weighted regression model (with weights based on the number of observations in trial i). The -2 log likelihood value of the (weighted or unweighted) models (3) (L_1) is subsequently compared to the -2 log likelihood value of an intercept-only model ($\widehat{\beta}_i = \lambda_3$; L_0), and R^2_{ht} is computed based on the Variance Reduction Factor (VFR; for details, see Alonso & Molenberghs, 2007):

$$R^2_{ht} = 1 - exp\left(-\frac{L_1 - L_0}{N}\right),$$

where N is the number of trials.

When a semi-reduced or reduced model is requested (by using the argument `Model=c("SemiReduced")` or `Model=c("Reduced")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1\widehat{\alpha}_i + \varepsilon_i,$$

where the parameter estimates for β_i and α_i are based on models (2). The -2 log likelihood value of this (weighted or unweighted) model (L_1) is subsequently compared to the -2 log likelihood value of an intercept-only model ($\widehat{\beta}_i = \lambda_3$; L_0), and R^2_{ht} is computed based on the reduction in the likelihood (as described above).

Value

An object of class `MixedContContIT` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Trial.Spec.Results` A `data.frame` that contains the trial-specific intercepts and treatment effects for the surrogate and the true endpoints (when a full or semi-reduced model is requested), or the trial-specific treatment effects for the surrogate and the true endpoints (when a reduced model is requested).

`R2ht` A `data.frame` that contains the trial-level surrogacy estimate and its confidence interval.

`R2h` A `data.frame` that contains the individual-level surrogacy estimate and its confidence interval.

`Cor.Endpoints` A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.

`Residuals` A `data.frame` that contains the residuals for the surrogate and true endpoints (ε_{Sij} and ε_{Tij}) that are obtained when models (1) or models (2) are fitted (see the **Details** section above).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A, & Molenberghs, G. (2007). Surrogate marker evaluation from an information theory perspective. *Biometrics, 63,* 180-186.

See Also

`FixedContContIT`, `plot Information-Theoretic`

Examples

```
## Not run:
# Example 1
# Based on the ARMD data:
data(ARMD)
# Assess surrogacy based on a full mixed-effect model
# in the information-theoretic framework:
Sur <- MixedContContIT(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
Pat.ID=Id, Model="Full", Number.Bootstraps=50)
# Obtain a summary of the results:
summary(Sur)

# Example 2
# Conduct an analysis based on a simulated dataset with 2000 patients, 200 trials,
# and Rindiv=Rtrial=.8
# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=200, R.Trial.Target=.8, R.Indiv.Target=.8,
Seed=123, Model="Full")
# Assess surrogacy based on a full mixed-effect model
# in the information-theoretic framework:
Sur2 <- MixedContContIT(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, Model="Full", Number.Bootstraps=50)

# Show a summary of the results:
summary(Sur2)
## End(Not run)
```

`plot Causal-Inference` *Plots the (Meta-Analytic) Individual Causal Association*

Description

This function provides a plot that displays the frequencies, percentages, or cumulative percentages of the individual causal association (ICA; ρ_Δ) and/or the meta-analytic individual causal association (MICA; ρ_M) values. These figures are useful to examine the sensitivity of the obtained results with respect to the assumptions regarding the correlations between the counterfactuals (for details, see Alonso et al., submitted; Van der Elst et al., submitted). Optionally, it is also possible to obtain plots that are useful in the examination of the plausibility of finding a good surrogate endpoint when an object of class `ICA.ContCont` is considered.

Usage

```
## S3 method for class 'ICA.ContCont'
plot(x, Xlab.ICA, Main.ICA, Type="Percent",
Labels=FALSE, ICA=TRUE, Good.Surr=FALSE, Main.Good.Surr,
Par=par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)), col, ...)

## S3 method for class 'MICA.ContCont'
```

```
plot(x, ICA=TRUE, MICA=TRUE, Type="Percent",
Labels=FALSE, Xlab.ICA, Main.ICA, Xlab.MICA, Main.MICA,
Par=par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)), col, ...)
```

Arguments

x	An object of class ICA.ContCont or MICA.ContCont. See ICA.ContCont or MICA.ContCont.
ICA	Logical. When ICA=TRUE, a plot of the ICA is provided. Default TRUE.
MICA	Logical. This argument only has effect when the plot() function is applied to an object of class MICA.ContCont. When MICA=TRUE, a plot of the MICA is provided. Default TRUE.
Type	The type of plot that is produced. When Type=Freq or Type=Percent, the Y-axis shows frequencies or percentages of ρ_Δ, ρ_M, and/or δ. When Type=CumPerc, the Y-axis shows cumulative percentages of ρ_Δ, ρ_M, and/or δ. Default "Percent".
Labels	Logical. When Labels=TRUE, the percentage of ρ_Δ, ρ_M, and/or δ values that are equal to or larger than the midpoint value of each of the bins are displayed (on top of each bin). Default FALSE.
Xlab.ICA	The legend of the X-axis of the ICA plot. Default "ρ_Δ".
Main.ICA	The title of the ICA plot. Default "ICA".
Xlab.MICA	The legend of the X-axis of the MICA plot. Default "ρ_M".
Main.MICA	The title of the MICA plot. Default "MICA".
Good.Surr	Logical. When Good.Surr=TRUE, a plot of δ is provided. This plot is useful in the context of examinating the plausibility of finding a good surrogate endpoint. Only applies when an object of class ICA.ContCont is considered. For details, see Alonso et al. (submitted). Default FALSE.
Main.Good.Surr	The title of the plot of δ. Only applies when an object of class ICA.ContCont is considered. For details, see Alonso et al. (submitted).
Par	Graphical parameters for the plot. Default par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)).
col	The color of the bins. Default col <- c(8).
...	Extra graphical parameters to be passed to hist().

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal inference and meta-analytic paradigms for the validation of surrogate markers.

Van der Elst, W., Alonso, A., & Molenberghs, G. (submitted). An exploration of the relationship between causal inference and meta-analytic measures of surrogacy.

See Also

ICA.ContCont, MICA.ContCont, plot MinSurrContCont

Examples

```
## Not run:
# Plot of ICA

# Generate the vector of ICA values when rho_T0S0=rho_T1S1=.95, and when the
# grid of values {0, .1, ..., 1} is considered for the correlations
# between the counterfactuals:
SurICA <- ICA.ContCont(T0S0=.95, T1S1=.95, T0T1=seq(0, 1, by=.1), T0S1=seq(0, 1, by=.1),
T1S0=seq(0, 1, by=.1), S0S1=seq(0, 1, by=.1))

# Plot the results:
plot(SurICA)

# Same plot but add the percentages of ICA values that are equal to or larger
# than the midpoint values of the bins
plot(SurICA, Labels=TRUE)

# Plot of both ICA and MICA

# Generate the vector of ICA and MICA values when R_trial=.8, rho_T0S0=rho_T1S1=.8,
# D.aa=5, D.bb=10, and when the grid of values {0, .1, ..., 1} is considered
# for the correlations between the counterfactuals:
SurMICA <- MICA.ContCont(Trial.R=.80, D.aa=5, D.bb=10, T0S0=.8, T1S1=.8,
T0T1=seq(0, 1, by=.1), T0S1=seq(0, 1, by=.1), T1S0=seq(0, 1, by=.1),
S0S1=seq(0, 1, by=.1))

# Plot the vector of generated ICA and MICA values
plot(SurMICA, ICA=TRUE, MICA=TRUE)
## End(Not run)
```

---

plot Information-Theoretic

Provides plots of trial- and individual-level surrogacy in the Information-Theoretic framework

---

Description

Produces plots that provide a graphical representation of trial- and/or individual-level surrogacy (R2_ht and R2_h) based on the Information-Theoretic approach of Alonso & Molenberghs (2007).

Usage

```
## S3 method for class 'FixedContContIT'
plot(x, Trial.Level=TRUE, Weighted=TRUE, Indiv.Level=TRUE,
Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial, Main.Trial, Main.Indiv,
```

```
Par=par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)), ...)

## S3 method for class 'MixedContContIT'
plot(x, Trial.Level=TRUE, Weighted=TRUE, Indiv.Level=TRUE,
Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial, Main.Trial, Main.Indiv,
Par=par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)), ...)
```

Arguments

x	An object of class MixedContContIT or FixedContContIT.
Trial.Level	Logical. If Trial.Level=TRUE, a plot of the trial-specific treatment effects on the true endpoint against the trial-specific treatment effect on the surrogate endpoints is provided (as a graphical representation of R_{ht}). Default TRUE.
Weighted	Logical. This argument only has effect when the user requests a trial-level surrogacy plot (i.e., when Trial.Level=TRUE in the function call). If Weighted=TRUE, the circles that depict the trial-specific treatment effects on the true endpoint against the surrogate endpoint are proportional to the number of patients in the trial. If Weighted=FALSE, all circles have the same size. Default TRUE.
Indiv.Level	Logical. If Indiv.Level=TRUE, a plot of the trial- and treatment-corrected residuals of the true and surrogate endpoints is provided. This plot provides a graphical representation of R_h. Default TRUE.
Xlab.Indiv	The legend of the X-axis of the plot that depicts individual-level surrogacy. Default "Residuals for the surrogate endpoint (ε_{Sij})".
Ylab.Indiv	The legend of the Y-axis of the plot that depicts individual-level surrogacy. Default "Residuals for the true endpoint (ε_{Tij})".
Xlab.Trial	The legend of the X-axis of the plot that depicts trial-level surrogacy. Default "Treatment effect on the surrogate endpoint (α_i)".
Ylab.Trial	The legend of the Y-axis of the plot that depicts trial-level surrogacy. Default "Treatment effect on the true endpoint (β_i)".
Main.Indiv	The title of the plot that depicts individual-level surrogacy. Default "Individual-level surrogacy".
Main.Trial	The title of the plot that depicts trial-level surrogacy. Default "Trial-level surrogacy".
Par	Graphical parameters for the plot. Default par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)).
...	Extra graphical parameters to be passed to plot().

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A, & Molenberghs, G. (2007). Surrogate marker evaluation from an information theory perspective. *Biometrics, 63,* 180-186.

See Also

MixedContContIT, FixedContContIT

Examples

```
## Load ARMD dataset
data(ARMD)

## Conduct a surrogacy analysis, using a weighted reduced univariate fixed effect model:
Sur <- MixedContContIT(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
Pat.ID=Id, Model=c("Full"))

## Request both trial- and individual-level surrogacy plots. In the trial-level plot,
## make the size of the circles proportional to the number of patients in a trial:
## Not run: plot(Sur, Trial.Level=TRUE, Weighted=TRUE, Indiv.Level=TRUE)

## Make a trial-level surrogacy plot using filled blue circles that
## are transparent (to make sure that the results of overlapping trials remain
## visible), and modify the title and the axes labels of the plot:
## Not run: plot(Sur, pch=16, col=rgb(.3, .2, 1, 0.3), Indiv.Level=FALSE, Trial.Level=TRUE,
Weighted=TRUE, Main.Trial=c("Trial-level surrogacy (ARMD dataset)"),
Xlab.Trial=c("Difference in vision after 6 months (Surrogate)"),
Ylab.Trial=c("Difference in vision after 12 months (True enpoint)"))
## End(Not run)

## Add the estimated R2_ht value in the previous plot at position (X=-2.2, Y=0)
## (the previous plot should not have been closed):
## Not run: R2ht <- format(round(as.numeric(Sur$R2ht[1]), 3))
text(x=-2.2, y=0, cex=1.4, labels=(bquote(paste("R"[ht]^{2}, "="~.(R2ht)))))
## End(Not run)

## Make an Individual-level surrogacy plot with red squares to depict individuals
## (rather than black circles):
## Not run: plot(Sur, pch=15, col="red", Indiv.Level=TRUE, Trial.Level=FALSE)

## Same plot as before, but now with smaller squares, a y-axis with range [-40; 40],
## and the estimated R2_h value in the title of the plot:
R2h <- format(round(as.numeric(Sur$R2h[1]), 3))
## Not run: plot(Sur, pch=15, col="red", Indiv.Level=TRUE, Trial.Level=FALSE, cex=.5,
ylim=c(-40, 40), Main.Indiv=bquote(paste("R"[h]^{2}, "="~.(R2h))))
## End(Not run)
```

`plot Meta-Analytic` *Provides plots of trial- and individual-level surrogacy in the meta-analytic framework*

Description

Produces plots that provide a graphical representation of trial- and/or individual-level surrogacy based on the meta-analytic approach of Buyse & Molenberghs (2000) in the single- and multiple-trial settings.

Usage

```
## S3 method for class 'BifixedContCont'
plot(x, Trial.Level=TRUE, Weighted=TRUE,
Indiv.Level=TRUE, Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial,
Main.Trial, Main.Indiv, Par=par(oma=c(0, 0, 0, 0),
mar=c(5.1, 4.1, 4.1, 2.1)), ...)

## S3 method for class 'BimixedContCont'
plot(x, Trial.Level=TRUE, Weighted=TRUE,
Indiv.Level=TRUE, Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial,
Main.Trial, Main.Indiv, Par=par(oma=c(0, 0, 0, 0),
mar=c(5.1, 4.1, 4.1, 2.1)), ...)

## S3 method for class 'UnifixedContCont'
plot(x, Trial.Level=TRUE, Weighted=TRUE,
Indiv.Level=TRUE, Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial,
Main.Trial, Main.Indiv, Par=par(oma=c(0, 0, 0, 0),
mar=c(5.1, 4.1, 4.1, 2.1)), ...)

## S3 method for class 'UnimixedContCont'
plot(x, Trial.Level=TRUE, Weighted=TRUE,
Indiv.Level=TRUE, Xlab.Indiv, Ylab.Indiv, Xlab.Trial, Ylab.Trial,
Main.Trial, Main.Indiv, Par=par(oma=c(0, 0, 0, 0),
mar=c(5.1, 4.1, 4.1, 2.1)), ...)
```

Arguments

`x`	An object of class `UnifixedContCont`, `BifixedContCont`, `UnimixedContCont`, `BimixedContCont`, or `Single.Trial.RE.AA`.
`Trial.Level`	Logical. If `Trial.Level=TRUE` and an object of class `UnifixedContCont`, `BifixedContCont`, `UnimixedContCont`, or `BimixedContCont` is considered, a plot of the trial-specific treatment effects on the true endpoint against the trial-specific treatment effect on the surrogate endpoints is provided (as a graphical representation of R_{trial}). If `Trial.Level=TRUE` and an object of class `Single.Trial.RE.AA` is considered, a plot of the treatment effect on the true endpoint against the treatment effect on the surrogate endpoint is provided, and a regression line that goes through the origin with slope RE is added to the plot (to depict the constant RE assumption, see `Single.Trial.RE.AA` for details). If `Trial.Level=FALSE`, this plot is not provided. Default `TRUE`.
`Weighted`	Logical. This argument only has effect when the user requests a trial-level surrogacy plot (i.e., when `Trial.Level=TRUE` in the function call) and when an

object of class UnifixedContCont, BifixedContCont, UnimixedContCont, or BimixedContCont is considered (not when an object of class Single.Trial.RE.AA is considered). If Weighted=TRUE, the circles that depict the trial-specific treatment effects on the true endpoint against the surrogate endpoint are proportional to the number of patients in the trial. If Weighted=FALSE, all circles have the same size. Default TRUE.

Indiv.Level — Logical. If Indiv.Level=TRUE, a plot of the trial- and treatment-corrected residuals of the true and surrogate endpoints is provided (when an object of class UnifixedContCont, BifixedContCont, UnimixedContCont, or BimixedContCont is considered), or a plot of the treatment-corrected residuals (when an object of class Single.Trial.RE.AA is considered). This plot provides a graphical representation of R_{indiv}. If Indiv.Level=FALSE, this plot is not provided. Default TRUE.

Xlab.Indiv — The legend of the X-axis of the plot that depicts individual-level surrogacy. Default "Residuals for the surrogate endpoint (ε_{Sij})" (without the i subscript when an object of class Single.Trial.RE.AA is considered).

Ylab.Indiv — The legend of the Y-axis of the plot that depicts individual-level surrogacy. Default "Residuals for the true endpoint (ε_{Tij})" (without the i subscript when an object of class Single.Trial.RE.AA is considered).

Xlab.Trial — The legend of the X-axis of the plot that depicts trial-level surrogacy. Default "Treatment effect on the surrogate endpoint (α_i)" (without the i subscript when an object of class Single.Trial.RE.AA is considered).

Ylab.Trial — The legend of the Y-axis of the plot that depicts trial-level surrogacy. Default "Treatment effect on the true endpoint (β_i)" (without the i subscript when an object of class Single.Trial.RE.AA is considered).

Main.Indiv — The title of the plot that depicts individual-level surrogacy. Default "Individual-level surrogacy" when an object of class UnifixedContCont, BifixedContCont, UnimixedContCont, or BimixedContCont is considered, and "Adjusted Association (rho_Z) when an object of class Single.Trial.RE.AA is considered.

Main.Trial — The title of the plot that depicts trial-level surrogacy. Default "Trial-level surrogacy" (when an object of class UnifixedContCont, BifixedContCont, UnimixedContCont, or BimixedContCont is considered) or "Relative Effect (RE)" (when an object of class Single.Trial.RE.AA is considered).

Par — Graphical parameters for the plot. Default par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)).

... — Extra graphical parameters to be passed to plot().

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

See Also

UnifixedContCont, BifixedContCont, UnifixedContCont, BimixedContCont, Single.Trial.RE.AA

Examples

```
##### Multiple-trial setting

## Load ARMD dataset
data(ARMD)

## Conduct a surrogacy analysis, using a weighted reduced univariate fixed effect model:
Sur <- UnifixedContCont(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
Pat.ID=Id, Number.Bootstraps=100, Model=c("Reduced"), Weighted=TRUE)

## Request both trial- and individual-level surrogacy plots. In the trial-level plot,
## make the size of the circles proportional to the number of patients in a trial:
## Not run: plot(Sur, Trial.Level=TRUE, Weighted=TRUE, Indiv.Level=TRUE)

## Make a trial-level surrogacy plot using filled blue circles that
## are transparent (to make sure that the results of overlapping trials remain
## visible), and modify the title and the axes labels of the plot:
## Not run: plot(Sur, pch=16, col=rgb(.3, .2, 1, 0.3), Indiv.Level=FALSE, Trial.Level=TRUE,
Weighted=TRUE, Main.Trial=c("Trial-level surrogacy (ARMD dataset)"),
Xlab.Trial=c("Difference in vision after 6 months (Surrogate)"),
Ylab.Trial=c("Difference in vision after 12 months (True enpoint)"))
## End(Not run)

## Add the estimated R2_trial value in the previous plot at position (X=-7, Y=11)
## (the previous plot should not have been closed):
## Not run: R2trial <- format(round(as.numeric(Sur$Trial.R2[1]), 3))
text(x=-7, y=11, cex=1.4, labels=(bquote(paste("R"[trial]^{2}, "="~.(R2trial)))))
## End(Not run)

## Make an Individual-level surrogacy plot with red squares to depict individuals
## (rather than black circles):
## Not run: plot(Sur, pch=15, col="red", Indiv.Level=TRUE, Trial.Level=FALSE)

## Same plot as before, but now with smaller squares, a y-axis with range [-40; 40],
## and the estimated R2_indiv value in the title of the plot:
R2ind <- format(round(as.numeric(Sur$Indiv.R2[1]), 3))
## Not run: plot(Sur, pch=15, col="red", Indiv.Level=TRUE, Trial.Level=FALSE, cex=.5,
ylim=c(-40, 40), Main.Indiv=bquote(paste("R"[indiv]^{2}, "="~.(R2ind))))
## End(Not run)


##### Single-trial setting

## Conduct a surrogacy analysis in the single-trial meta-analytic setting:
SurSTS <- Single.Trial.RE.AA(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Pat.ID=Id)

# Request a plot of individual-level surrogacy and a plot that depicts the Relative effect
# and the constant RE assumption:
```

```
## Not run: plot(SurSTS, Trial.Level=TRUE, Indiv.Level=TRUE)
```

---

plot MinSurrContCont *Graphically illustrates the theoretical plausibility of finding a good surrogate endpoint in the continuous-continuous case*

Description

This function provides a plot that displays the frequencies, percentages, or cumulative percentages of ρ^2_{min} for a fixed value of δ (given the observed variances of the true endpoint in the control and experimental treatment conditions and a specified grid of values for the unidentified parameter $\rho_{T_0T_1}$; see MinSurrContCont). For details, see the online appendix of Alonso et al., submitted.

Usage

```
## S3 method for class 'MinSurrContCont'
plot(x, main, col, Type="Percent", Labels=FALSE,
Par=par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)), ...)
```

Arguments

x	An object of class MinSurrContCont. See MinSurrContCont.
main	The title of the plot.
col	The color of the bins.
Type	The type of plot that is produced. When Type=Freq or Type=Percent, the Y-axis shows frequencies or percentages of ρ^2_{min}. When Type=CumPerc, the Y-axis shows cumulative percentages of ρ^2_{min}. Default "Percent".
Labels	Logical. When Labels=TRUE, the percentage of ρ^2_{min} values that are equal to or larger than the midpoint value of each of the bins are displayed (on top of each bin). Only applies when Type=Freq or Type=Percent. Default FALSE.
Par	Graphical parameters for the plot. Default par(oma=c(0, 0, 0, 0), mar=c(5.1, 4.1, 4.1, 2.1)).
...	Extra graphical parameters to be passed to hist().

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal inference and meta-analytic paradigms for the validation of surrogate markers.

See Also

MinSurrContCont

Examples

```
# compute rho^2_min in the setting where the variances of T in the control
# and experimental treatments equal 100 and 120, delta is fixed at 50,
# and the grid G={0, .01, ..., 1} is considered for the counterfactual
# correlation rho_T0T1:

MinSurr <- MinSurrContCont(T0T0 = 100, T1T1 = 120, Delta = 50,
T0T1 = seq(0, 1, by = 0.01))

# Plot the results (use percentages on Y-axis)
## Not run: plot(MinSurr, Type="Percent")

# Same plot, but add the percentages of ICA values that are equal to or
# larger than the midpoint values of the bins
## Not run: plot(MinSurr, Labels=TRUE)
```

Pos.Def.Matrices *Generate 4 by 4 correlation matrices and flag the positive definite ones*

Description

Based on vectors (or scalars) for the six off-diagonal correlations of a 4 by 4 matrix, the function `Pos.Def.Matrices` constructs all possible matrices that can be formed by combining the specified values, computes the minimum eigenvalues for each of these matrices, and flags the positive definite ones (i.e., valid correlation matrices).

Usage

```
Pos.Def.Matrices(T0T1=seq(0, 1, by=.2), T0S0=seq(0, 1, by=.2), T0S1=seq(0, 1,
by=.2), T1S0=seq(0, 1, by=.2), T1S1=seq(0, 1, by=.2), S0S1=seq(0, 1, by=.2))
```

Arguments

`T0T1`	A vector or scalar that specifies the correlation(s) between T0 and T1 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`, i.e., the values $0, 0.20, \ldots, 1$.
`T0S0`	A vector or scalar that specifies the correlation(s) between T0 and S0 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`.
`T0S1`	A vector or scalar that specifies the correlation(s) between T0 and S1 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`.
`T1S0`	A vector or scalar that specifies the correlation(s) between T1 and S0 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`.
`T1S1`	A vector or scalar that specifies the correlation(s) between T1 and S1 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`.

S0S1 A vector or scalar that specifies the correlation(s) between S0 and S1 that should be considered to construct all possible 4 by 4 matrices. Default `seq(0, 1, by=.2)`.

Details

The generated object `Generated.Matrices` (of class `data.frame`) is placed in the workspace (for easy access).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

See Also

`Sim.Data.Counterfactuals`

Examples

```
## Generate all 4x4 matrices that can be formed using rho(T0,S0)=rho(T1,S1)=.5
## and the grid of values 0, .2, ..., 1 for the other off-diagonal correlations:
Pos.Def.Matrices(T0T1=seq(0, 1, by=.2), T0S0=.5, T0S1=seq(0, 1, by=.2),
T1S0=seq(0, 1, by=.2), T1S1=.5, S0S1=seq(0, 1, by=.2))

## Examine the first 10 rows of the the object Generated.Matrices:
Generated.Matrices[1:10,]

## Check how many of the generated matrices are positive definite
## (counts and percentages):
table(Generated.Matrices$Pos.Def.Status)
table(Generated.Matrices$Pos.Def.Status)/nrow(Generated.Matrices)

## Make an object PosDef which contains the positive definite matrices:
PosDef <- Generated.Matrices[Generated.Matrices$Pos.Def.Status==1,]

## Shows the 10 first matrices that are positive definite:
PosDef[1:10,]
```

Schizo *Data of five clinical trials in schizophrenia*

Description

These are the data of five clinical trial in schizophrenia. A total of 2128 patients were treated by 198 investiagators (psychiatrists). Patients' schizophrenic symptoms were measured using the PANSS, BPRS, and CGI. There were two treatment conditions (risperidone and control).

Usage

```
data(Schizo)
```

Format

A `data.frame` with 2128 observations on 198 variables.

`Id` The patient ID.

`InvestID` The ID of the investigator (psychiatrist) who treated the patient.

`Treat` The treatment indicator, coded as -1 = placebo and 1 = interferon-α.

`CGI` The change in the CGI score (= score at the start of the treatment - score at the end of the treatment).

`PANSS` The change in the PANSS score (= score at the start of the treatment - score at the end of the treatment).

`BPRS` The change in the PANSS score (= score at the start of the treatment - score at the end of the treatment).

`Sim.Data.Counterfactuals` *Simulate a dataset that contains counterfactuals*

Description

The function `Sim.Data.Counterfactuals` simulates a dataset that contains four (continuous) counterfactuals (i.e., potential outcomes) and a (binary) treatment indicator. The counterfactuals T_0 and T_1 denote the true endpoints of a patient under the control and the experimental treatments, respectively, and the counterfactuals S_0 and S_1 denote the surrogate endpoints of the patient under the control and the experimental treatments, respectively. The user can specify the number of patients, the desired mean values for the counterfactuals (i.e., $\boldsymbol{\mu}_c$), and the desired correlations between the counterfactuals (i.e., the off-diagonal values in the standardized $\boldsymbol{\Sigma}_c$ matrix). For details, see the papers of Alonso et al. (submitted) and Van der Elst et al. (submitted).

Usage

```
Sim.Data.Counterfactuals(N.Total=2000, mu_c=c(0, 0, 0, 0), T0S0=0, T1S1=0,
T0T1=0, T0S1=0, T1S0=0, S0S1=0, Seed=sample(1:1000, size=1))
```

Arguments

`N.Total`	The total number of patients in the simulated dataset. Default 2000.
`mu_c`	A vector that specifies the desired means for the counterfactuals S_0, S_1, T_0, and T_1, respectively. Default `c(0, 0, 0, 0)`.
`T0S0`	A scalar that specifies the desired correlation between the counterfactuals T0 and S0 that should be used in the generation of the data. Default 0.
`T1S1`	A scalar that specifies the desired correlation between the counterfactuals T1 and S1 that should be used in the generation of the data. Default 0.
`T0T1`	A scalar that specifies the desired correlation between the counterfactuals T0 and T1 that should be used in the generation of the data. Default 0.

T0S1	A scalar that specifies the desired correlation between the counterfactuals T0 and S1 that should be used in the generation of the data. Default 0.
T1S0	A scalar that specifies the desired correlation between the counterfactuals T1 and S0 that should be used in the generation of the data. Default 0.
S0S1	A scalar that specifies the desired correlation between the counterfactuals T0 and T1 that should be used in the generation of the data. Default 0.
Seed	A seed that is used to generate the dataset. Default `sample(x=1:1000, size=1)`, i.e., a random number between 1 and 1000.

Details

The generated object `Data.Counterfactuals` (of class `data.frame`) is placed in the workspace.

The specified values for T0S0, T1S1, T0T1, T0S1, T1S0, and S0S1 in the function call should form a matrix that is positive definite (i.e., they should form a valid correlation matrix). When the user specifies values that form a matrix that is not positive definite, an error message is given and the object `Data.Counterfactuals` is not generated. The function `Pos.Def.Matrices` can be used to examine beforehand whether a 4 by 4 matrix is positive definite.

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., & Burzykowski, T. (submitted). On the relationship between the causal inference and meta-analytic paradigms for the validation of surrogate markers.

Van der Elst, W., Alonso, A., & Molenberghs, G. (submitted). An exploration of the relationship between causal inference and meta-analytic measures of surrogacy.

See Also

`Sim.Data.MTS`, `Sim.Data.STS`

Examples

```
## Generate a dataset with 2000 patients, cor(S0,T0)=cor(S1,T1)=.5,
## cor(T0,T1)=cor(T0,S1)=cor(T1,S0)=cor(S0,S1)=0, with means
## 5, 9, 12, and 15 for S0, S1, T0, and T1, respectively:
Sim.Data.Counterfactuals(N=2000, T0S0=.5, T1S1=.5, T0T1=0, T0S1=0, T1S0=0, S0S1=0,
mu_c=c(5, 9, 12, 15), Seed=1)
```

`Sim.Data.MTS` *Simulates a dataset that can be used to assess surrogacy in the multiple-trial setting*

Description

The function `Sim.Data.MTS` simulates a dataset that contains the variables Treat, Trial.ID, Surr, True, and Pat.ID. The user can specify the number of patients and the number of trials that should be included in the simulated dataset, the desired R_{trial} and R_{indiv} values, the desired variability of the trial-specific treatment effects for the surrogate and the true endpoints (i.e., d_{aa} and d_{bb}, respectively), and the desired fixed-effect parameters of the intercepts and treatment effects for the surrogate and the true endpoints.

Usage

```
Sim.Data.MTS(N.Total=2000, N.Trial=50, R.Trial.Target=.8, R.Indiv.Target=.8,
Fixed.Effects=c(0, 0, 0, 0), D.aa=10, D.bb=10, Seed=sample(1:1000, size=1),
Model=c("Full"))
```

Arguments

`N.Total`	The total number of patients in the simulated dataset. Default 2000.
`N.Trial`	The number of trials. Default 50.
`R.Trial.Target`	The desired R_{trial} value in the sumilated dataset. Default 0.80
`R.Indiv.Target`	The desired R_{indiv} value in the simulated dataset. Default 0.80.
`Fixed.Effects`	A vector that specifies the desired fixed-effect intercept for the surrogate, fixed-effect intercept for the true endpoint, fixed treatment effect for the surrogate, and fixed treatment effect for the true endpoint, respectively. Default `c(0, 0, 0, 0)`.
`D.aa`	The desired variability of the trial-specific treatment effects on the surrogate endpoint. Default 10.
`D.bb`	The desired variability of the trial-specific treatment effects on the true endpoint. Default 10.
`Model`	The type of model that will be fitted on the data when surrogacy is assessed, i.e., a full, semireduced, or reduced model (for details, see `UnifixedContCont`, `UnimixedContCont`, `BifixedContCont`, `BimixedContCont`).
`Seed`	The seed that is used to generate the dataset. Default `sample(x=1:1000, size=1)`, i.e., a random number between 1 and 1000.

Details

The generated object `Data.Observed.MTS` (of class `data.frame`) is placed in the workspace (for easy access).

The number of patients per trial in the simulated dataset is identical in each trial, and equals the requested total number of patients divided by the requested number of trials (`=N.Total/N.Trial`).

If this is not a whole number, a warning is given and the number of patients per trial is automatically rounded up to the nearest whole number. See **Examples** below.

Treatment allocation is balanced when the number of patients per trial is an odd number. If this is not the case, treatment allocation is balanced up to one patient (the remaining patient is randomly allocated to the exprimental or the control treatment groups in each of the trials).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

See Also

UnifixedContCont, BifixedContCont, UnimixedContCont, BimixedContCont, Sim.Data.STS

Examples

```
# Simulate a dataset with 2000 patients, 50 trials, Rindiv=Rtrial=.8, D.aa=10,
# D.bb=50, and fixed effect values 1, 2, 30, and 90:
Sim.Data.MTS(N.Total=2000, N.Trial=50, R.Trial.Target=.8, R.Indiv.Target=.8, D.aa=10,
D.bb=50, Fixed.Effects=c(1, 2, 30, 90), Seed=1)

# Sample output, the first 10 rows of Data.Observed.MTS:
Data.Observed.MTS[1:10,]

# Note: When the following code is used to generate a dataset:
Sim.Data.MTS(N.Total=2000, N.Trial=99, R.Trial.Target=.5, R.Indiv.Target=.8,
D.aa=10, D.bb=50, Fixed.Effects=c(1, 2, 30, 90), Seed=1)

# R gives the following warning:

# > NOTE: The number of patients per trial requested in the function call
# > equals 20.20202 (=N.Total/N.Trial), which is not a whole number.
# > To obtain a dataset where the number of patients per trial is balanced for
# > all trials, the number of patients per trial was rounded to 21 to generate
# > the dataset. Data.Observed.MTS thus contains a total of 2079 patients rather
# > than the requested 2000 in the function call.
```

Sim.Data.STS *Simulates a dataset that can be used to assess surrogacy in the single-trial setting*

Description

The function Sim.Data.STS simulates a dataset that contains the variables Treat, Surr, True, and Pat.ID. The user can specify the total number of patients, the desired R_{indiv} value (also referred to as the adjusted association (γ) in the single-trial meta-analytic setting), and the desired means of the surrogate and the true endpoints in the experimental and control treatment groups.

Usage

```
Sim.Data.STS(N.Total=2000, R.Indiv.Target=.8, Means=c(0, 0, 0, 0), Seed=
sample(1:1000, size=1))
```

Arguments

N.Total	The total number of patients in the simulated dataset. Default 2000.
R.Indiv.Target	The desired R_{indiv} (or γ) value in the simulated dataset. Default 0.80.
Means	A vector that specifies the desired mean for the surrogate in the control treatment group, mean for the surrogate in the experimental treatment group, mean for the true endpoint in the control treatment group, and mean for the true endpoint in the experimental treatment group, respectively. Default c(0, 0, 0, 0).
Seed	The seed that is used to generate the dataset. Default sample(x=1:1000, size=1), i.e., a random number between 1 and 1000.

Details

The generated object Data.Observed.STS (of class data.frame) is placed in the workspace (for easy access).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

See Also

Sim.Data.MTS, Single.Trial.RE.AA

Examples

```
# Simulate a dataset:
Sim.Data.STS(N.Total=2000, R.Indiv.Target=.8, Means=c(1, 5, 20, 37), Seed=1)
```

Single.Trial.RE.AA *Conducts a surrogacy analysis based on the single-trial meta-analytic framework*

Description

The function Single.Trial.RE.AA conducts a surrogacy analysis based on the single-trial meta-analytic framework of Buyse & Molenberghs (1998). See **Details** below.

Usage

```
Single.Trial.RE.AA(Dataset, Surr, True, Treat, Pat.ID, Alpha=.05,
Number.Bootstraps=500, Seed=sample(1:1000, size=1))
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, and a patient ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.
`Treat`	The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.
`Pat.ID`	The name of the variable in `Dataset` that contains the patient's ID.
`Alpha`	The α-level that is used to determine the confidence intervals around Alpha (which is a parameter estimate of a model where the surrogate is regressed on the treatment indicator, see **Details** below), Beta, RE, and γ. Default 0.05.
`Number.Bootstraps`	The number of bootstrap samples that are used to obtain the bootstrapp-based confidence intervals for RE and the adjusted association (γ). Default 500.
`Seed`	The seed that is used to generate the bootstrap samples. Default `sample(x=1:1000, size=1)`, i.e., a random number between 1 and 1000.

Details

The Relative Effect (RE) and the adjusted association (γ) are based on the following bivariate regression model (when the surrogate and the true endpoints are continuous variables):

$$S_j = \mu_S + \alpha Z_j + \varepsilon_{Sj},$$

$$T_j = \mu_T + \beta Z_j + \varepsilon_{Tj},$$

where the error terms have a joint zero-mean normal distribution with variance-covariance matrix:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{SS} & \\ \sigma_{ST} & \sigma_{TT} \end{pmatrix},$$

and where j is the subject indicator, S_j and T_j are the surrogate and true endpoint values of patient j, and Z_j is the treatment indicator for patient j.

The parameter estimates of the fitted regression model and the variance-covariance matrix of the residuals are used to compute RE and the adjusted association (γ), respectively:

$$RE = \frac{\beta}{\alpha},$$

$$\gamma = \frac{\sigma_{ST}}{\sqrt{\sigma_{SS}\sigma_{TT}}}.$$

Note

The single-trial meta-analytic framework is hampered by a number of issues (Burzykowski et al., 2005). For example, a key motivation to validate a surrogate endpoint is to be able to predict the effect of Z on T as based on the effect of Z on S in a new clinical trial where T is not (yet) observed. The RE allows for such a prediction, but this requires the assumption that the relation between α and β can be described by a linear regression model that goes through the origin. In other words, it has to be assumed that the RE remains constant across clinical trials. The constant RE assumption is unverifiable in a single-trial setting, but a way out of this problem is to combine the information of multiple clinical trials and generalize the RE concept to a multiple-trial setting (as is done in the multiple-trial meta-analytic approach, see `UnifixedContCont`, `BifixedContCont`, `UnimixedContCont`, and `BimixedContCont`).

Value

An object of class `Single.Trial.RE.AA` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Alpha` An object of class `data.frame` that contains the parameter estimate for α, its standard error, and its confidence interval. Note that `Alpha` is not to be confused with the `Alpha` argument in the function call, which specifies the α-level of the confidence intervals of the parameters.

`Beta` An object of class `data.frame` that contains the parameter estimate for β, its standard error, and its confidence interval.

`RE.Delta` An object of class `data.frame` that contains the estimated RE, its standard error, and its confidence interval (based on the Delta method).

`RE.Fieller` An object of class `data.frame` that contains the estimated RE, its standard error, and its confidence interval (based on Fieller's theorem).

`RE.Boot` An object of class `data.frame` that contains the estimated RE, its standard error, and its confidence interval (based on bootstrapping). Note that the occurence of outliers in the sample of bootstrapped RE values may lead to standard errors and/or confidence intervals that are not trustworthy. Such problems mainly occur when the parameter estimate for α is close to 0 (taking its standard error into account). To detect possible outliers, studentized deleted residuals are computed (by fitting an intercept-only model with the bootstrapped RE values as the outcome variable). Bootstrapped RE values with an absolute studentized residual larger than $t(1-\alpha/2n; n-2)$ are marked as outliers (where n = the number of bootstrapped RE values; Kutner et al., 2005). A warning is given when outliers are found, and the position of the outlier(s) in the bootstrap sample is identified. Inspection of the vector of bootstrapped RE values (see `RE.Boot.Samples` below) is recommended in this situation, and/or the use of the confidence intervals that are based on the Delta method or Fieller's theorem (rather than the bootstrap-based confidence interval).

`AA` An object of class `data.frame` that contains the adjusted association (i.e., γ), its standard error, and its confidence interval (based on the Fisher-Z transformation procedure).

`AA.Boot` An object of class `data.frame` that contains the adjusted association (i.e., γ), its standard error, and its confidence interval (based on a bootstrap procedure).

`RE.Boot.Samples`
: A vector that contains the RE values that were generated during the bootstrap procedure.

`AA.Boot.Samples`
: A vector that contains the adjusted association (i.e., γ) values that were generated during the bootstrap procedure.

`Cor.Endpoints`
: A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0T1}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.

`Residuals`
: A `data.frame` that contains the residuals for the surrogate and true endpoints that are obtained when the surrogate and the true endpoint are regressed on the treatment indicator.

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Burzykowski, T., Molenberghs, G., & Buyse, M. (2005). *The evaluation of surrogate endpoints*. New York: Springer-Verlag.

Buyse, M., & Molenberghs, G. (1998). The validation of surrogate endpoints in randomized experiments. *Biometrics, 54,* 1014-1029.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models (5th ed.)*. New York: McGraw Hill.

See Also

`UnifixedContCont`, `BifixedContCont`, `UnimixedContCont`, `BimixedContCont`, `ICA.ContCont`

Examples

```
## Not run:
# Example 1, based on the ARMD data:
data(ARMD)

# Assess surrogacy based on the single-trial meta-analytic approach:
Sur <- Single.Trial.RE.AA(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Pat.ID=Id)

# Obtain a summary and plot of the results
summary(Sur)
plot(Sur)


# Example 2
# Conduct an analysis based on a simulated dataset with 2000 patients
# and Rindiv=.8
```

```
# Simulate the data:
Sim.Data.STS(N.Total=2000, R.Indiv.Target=.8, Seed=123)

# Assess surrogacy:
Sur2 <- Single.Trial.RE.AA(Dataset=Data.Observed.STS, Surr=Surr, True=True, Treat=Treat,
Pat.ID=Pat.ID)

# Show a summary and plots of results
summary(Sur2)
plot(Sur2)
## End(Not run)
```

`UnifixedContCont` *Fits univariate fixed-effect models to assess surrogacy in the meta-analytic multiple-trial setting (continuous-continuous case)*

Description

The function `UnifixedContCont` uses the univariate fixed-effects approach to estimate trial- and individual-level surrogacy when the data of multiple clinical trials are available. The user can specify whether a (weighted or unweighted) full, semi-reduced, or reduced model should be fitted. See the **Details** section below.

Usage

```
UnifixedContCont(Dataset, Surr, True, Treat, Trial.ID, Pat.ID, Model=c("Full"),
Weighted=TRUE, Min.Trial.Size=2, Alpha=.05, Number.Bootstraps=500)
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate endpoint values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.
`Treat`	The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.
`Trial.ID`	The name of the variable in `Dataset` that contains the trial ID to which the patient belongs.
`Pat.ID`	The name of the variable in `Dataset` that contains the patient's ID.
`Model`	The type of model that should be fitted, i.e., `Model=c("Full")`, `Model=c("Reduced")`, or `Model=c("SemiReduced")`. See the **Details** section below. Default `Model=c("Full")`.

`Weighted` Logical. If `TRUE`, then a weighted regression analysis is conducted at stage 2 of the two-stage approach. If `FALSE`, then an unweighted regression analysis is conducted at stage 2 of the two-stage approach. See the **Details** section below. Default `TRUE`.

`Min.Trial.Size` The minimum number of patients that a trial should contain to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.

`Alpha` The α-level that is used to determine the confidence intervals around R^2_{trial}, R_{trial}, R^2_{indiv}, and R_{indiv}. Default 0.05.

`Number.Bootstraps`
The standard errors and confidence intervals for R^2_{indiv} and R_{indiv} are determined as based on a bootstrap procedure. `Number.Bootstraps` specifies the number of bootstrap samples that are used. Default 500.

Details

When the full bivariate mixed-effects model is fitted to assess surrogacy in the meta-analytic framework (for details, Buyse & Molenberghs, 2000), computational issues often occur. In that situation, the use of simplified model-fitting strategies may be warranted (for details, see see Burzykowski et al., 2005; Tibaldi et al., 2003).

The function `UnifixedContCont` implements one such strategy, i.e., it uses a two-stage univariate fixed-effects modelling approach to assess surrogacy. In the first stage of the analysis, two univariate linear regression models are fitted to the data of each of the i trials. When a full or semi-reduced model is requested (by using the argument `Model=c("Full")` or `Model=c("SemiReduced")` in the function call), the following univariate models are fitted:

$$S_{ij} = \mu_{Si} + \alpha_i Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij},$$

where i and j are the trial and subject indicators, S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_{Si} and μ_{Ti} are the fixed trial-specific intercepts for S and T, and α_i and β_i are the fixed trial-specific treatment effects on S and T, respectively. The error terms ε_{Sij} and ε_{Tij} are assumed to be independent.

When a reduced model is requested by the user (by using the argument `Model=c("Reduced")` in the function call), the following univariate models are fitted:

$$S_{ij} = \mu_S + \alpha_i Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_T + \beta_i Z_{ij} + \varepsilon_{Tij},$$

where μ_S and μ_T are the common intercepts for S and T (i.e., it is assumed that the intercepts for the surrogate and the true endpoints are identical in each of the trials). The other parameters are the same as defined above, and ε_{Sij} and ε_{Tij} are again assumed to be independent.

An estimate of R^2_{indiv} is provided by $r(\varepsilon_{Sij}, \varepsilon_{Tij})^2$.

Next, the second stage of the analysis is conducted. When a full model is requested (by using the argument `Model=c("Full")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\mu_{Si}} + \lambda_2 \widehat{\alpha}_i + \varepsilon_i,$$

where the parameter estimates for β_i, μ_{Si}, and α_i are based on the full models that were fitted in stage 1.

When a semi-reduced or reduced model is requested (by using the argument `Model=c("SemiReduced")` or `Model=c("Reduced")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\alpha}_i + \varepsilon_i.$$

where the parameter estimates for β_i and α_i are based on the semi-reduced or reduced models that were fitted in stage 1.

When the argument `Weighted=FALSE` is used in the function call, the model that is fitted in stage 2 is an unweighted linear regression model. When a weighted model is requested (using the argument `Weighted=TRUE` in the function call), the information that is obtained in stage 1 is weighted according to the number of patients in a trial.

The classical coefficient of determination of the fitted stage 2 model provides an estimate of R^2_{trial}.

Value

An object of class `UnifixedContCont` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Results.Stage.1`
The results of stage 1 of the two-stage model fitting approach: a `data.frame` that contains the trial-specific intercepts and treatment effects for the surrogate and the true endpoints (when a full or semi-reduced model is requested), or the trial-specific treatment effects for the surrogate and the true endpoints (when a reduced model is requested).

`Residuals.Stage.1`
A `data.frame` that contains the residuals for the surrogate and true endpoints that are obtained in stage 1 of the analysis (ε_{Sij} and ε_{Tij}).

`Results.Stage.2`
An object of class `lm` (linear model) that contains the parameter estimates of the regression model that is fitted in stage 2 of the analysis.

`Trial.R2` A `data.frame` that contains the trial-level coefficient of determination (R^2_{trial}), its standard error and confidence interval.

`Indiv.R2` A `data.frame` that contains the individual-level coefficient of determination (R^2_{indiv}), its standard error and confidence interval.

`Trial.R` A `data.frame` that contains the trial-level correlation coefficient (R_{trial}), its standard error and confidence interval.

`Indiv.R` A `data.frame` that contains the individual-level correlation coefficient (R_{indiv}), its standard error and confidence interval.

`Cor.Endpoints` A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.

`D.Equiv` The variance-covariance matrix of the trial-specific intercept and treatment effects for the surrogate and true endpoints (when a full or semi-reduced model is fitted, i.e., when `Model=c("Full")` or `Model=c("SemiReduced")` is used in the function call), or the variance-covariance matrix of the trial-specific treatment effects for the surrogate and true endpoints (when a reduced model is fitted, i.e., when `Model=c("Reduced")` is used in the function call). The variance-covariance matrix `D.Equiv` is equivalent to the $\boldsymbol{D}$ matrix that would be obtained when a (full or reduced) bivariate mixed-effect approach is used; see function `BimixedContCont`).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Burzykowski, T., Molenberghs, G., & Buyse, M. (2005). *The evaluation of surrogate endpoints*. New York: Springer-Verlag.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

Tibaldi, F., Abrahantes, J. C., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., et al., (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. *Journal of Statistical Computation and Simulation, 73,* 643-658.

See Also

`UnimixedContCont`, `BifixedContCont`, `BimixedContCont`, `plot Meta-Analytic`

Examples

```
## Not run:
# Example 1, based on the ARMD data
data(ARMD)

# Fit a full univariate fixed-effects model with weighting according to the
# number of patients in stage 2 of the two stage approach to assess surrogacy:
Sur <- UnifixedContCont(Dataset=ARMD, Surr=Diff24, True=Diff52, Treat=Treat, Trial.ID=Center,
```

```
Pat.ID=Id, Model="Full", Weighted=TRUE)

# Obtain a summary and plot of the results
summary(Sur)
plot(Sur)

# Example 2
# Conduct an analysis based on a simulated dataset with 2000 patients, 100 trials,
# and Rindiv=Rtrial=.8
# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=100, R.Trial.Target=.8, R.Indiv.Target=.8,
Seed=123, Model="Reduced")

# Fit a reduced univariate fixed-effects model without weighting to assess
# surrogacy:
Sur2 <- UnifixedContCont(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, Model="Reduced", Weighted=FALSE)

# Show a summary and plots of results:
summary(Sur2)
plot(Sur2, Weighted=FALSE)
## End(Not run)
```

`UnimixedContCont` *Fits univariate mixed-effect models to assess surrogacy in the meta-analytic multiple-trial setting (continuous-continuous case)*

Description

The function `UnimixedContCont` uses the univariate mixed-effects approach to estimate trial- and individual-level surrogacy when the data of multiple clinical trials are available. The user can specify whether a (weighted or unweighted) full, semi-reduced, or reduced model should be fitted. See the **Details** section below.

Usage

```
UnimixedContCont(Dataset, Surr, True, Treat, Trial.ID, Pat.ID, Model=c("Full"),
Weighted=TRUE, Min.Trial.Size=2, Alpha=.05, Number.Bootstraps=500,
...)
```

Arguments

`Dataset`	A `data.frame` that should consist of one line per patient. Each line contains (at least) a surrogate value, a true endpoint value, a treatment indicator, a patient ID, and a trial ID.
`Surr`	The name of the variable in `Dataset` that contains the surrogate endpoint values.
`True`	The name of the variable in `Dataset` that contains the true endpoint values.

`Treat` The name of the variable in `Dataset` that contains the treatment indicators. The treatment indicator should either be coded as 1 for the experimental group and -1 for the control group, or as 1 for the experimental group and 0 for the control group. The $-1/1$ coding is recommended.

`Trial.ID` The name of the variable in `Dataset` that contains the trial ID to which the patient belongs.

`Pat.ID` The name of the variable in `Dataset` that contains the patient's ID.

`Model` The type of model that should be fitted, i.e., `Model=c("Full")`, `Model=c("Reduced")`, or `Model=c("SemiReduced")`. See the **Details** section below. Default `Model=c("Full")`.

`Weighted` Logical. If `TRUE`, then a weighted regression analysis is conducted at stage 2 of the two-stage approach. If `FALSE`, then an unweighted regression analysis is conducted at stage 2 of the two-stage approach. See the **Details** section below. Default `TRUE`.

`Min.Trial.Size` The minimum number of patients that a trial should contain to be included in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded from the analysis. Default 2.

`Alpha` The α-level that is used to determine the confidence intervals around R^2_{trial}, R_{trial}, R^2_{indiv}, and R_{indiv}. Default 0.05.

`Number.Bootstraps`
The confidence intervals for R^2_{indiv} and R_{indiv} are determined as based on a bootstrap procedure. `Number.Bootstraps` specifies the number of bootstrap samples that are to be used. Default 500.

`...` Other arguments to be passed to the function `lmer` (of the R package `lme4`) that is used to fit the geralized linear mixed-effect models in the function `BimixedContCont`. For details, see the lme4 manual.

Details

When the full bivariate mixed-effects model is fitted to assess surrogacy in the meta-analytic framework (for details, Buyse & Molenberghs, 2000), computational issues often occur. In that situation, the use of simplified model-fitting strategies may be warranted (for details, see Burzykowski et al., 2005; Tibaldi et al., 2003).

The function `UnimixedContCont` implements one such strategy, i.e., it uses a two-stage univariate mixed-effects modelling approach to assess surrogacy. In the first stage of the analysis, two univariate mixed-effects models are fitted to the data. When a full or semi-reduced model is requested (by using the argument `Model=c("Full")` or `Model=c("SemiReduced")` in the function call), the following univariate models are fitted:

$$S_{ij} = \mu_S + m_{Si} + (\alpha + a_i)Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_T + m_{Ti} + (\beta + b_i)Z_{ij} + \varepsilon_{Tij},$$

where i and j are the trial and subject indicators, S_{ij} and T_{ij} are the surrogate and true endpoint values of subject j in trial i, Z_{ij} is the treatment indicator for subject j in trial i, μ_S and μ_T are the fixed intercepts for S and T, m_{Si} and m_{Ti} are the corresponding random intercepts, α and β are the fixed treatment effects for S and T, and a_i and b_i are the corresponding random treatment effects, respectively. The error terms ε_{Sij} and ε_{Tij} are assumed to be independent.

When a reduced model is requested (by using the argument `Model=c("Reduced")` in the function call), the following two univariate models are fitted:

$$S_{ij} = \mu_S + (\alpha + a_i)Z_{ij} + \varepsilon_{Sij},$$

$$T_{ij} = \mu_T + (\beta + b_i)Z_{ij} + \varepsilon_{Tij},$$

where μ_S and μ_T are the common intercepts for S and T (i.e., it is assumed that the intercepts for the surrogate and the true endpoints are identical in each of the trials). The other parameters are the same as defined above, and ε_{Sij} and ε_{Tij} are again assumed to be independent.

An estimate of R^2_{indiv} is computed as $r(\varepsilon_{Sij}, \varepsilon_{Tij})^2$.

Next, the second stage of the analysis is conducted. When a full model is requested by the user (by using the argument `Model=c("Full")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\mu_{Si}} + \lambda_2 \widehat{\alpha}_i + \varepsilon_i,$$

where the parameter estimates for β_i, μ_{Si}, and α_i are based on the models that were fitted in stage 1, i.e., $\beta_i = \beta + b_i$, $\mu_{Si} = \mu_S + m_{Si}$, and $\alpha_i = \alpha + a_i$.

When a reduced or semi-reduced model is requested by the user (by using the arguments `Model=c("SemiReduced")` or `Model=c("Reduced")` in the function call), the following model is fitted:

$$\widehat{\beta}_i = \lambda_0 + \lambda_1 \widehat{\alpha}_i + \varepsilon_i,$$

where the parameters are the same as defined above.

When the argument `Weighted=FALSE` is used in the function call, the model that is fitted in stage 2 is an unweighted linear regression model. When a weighted model is requested (using the argument `Weighted=TRUE` in the function call), the information that is obtained in stage 1 is weighted according to the number of patients in a trial.

The classical coefficient of determination of the fitted stage 2 model provides an estimate of R^2_{trial}.

Value

An object of class `UnimixedContCont` with components,

`Data.Analyze` Prior to conducting the surrogacy analysis, data of patients who have a missing value for the surrogate and/or the true endpoint are excluded. In addition, the data of trials (i) in which only one type of the treatment was administered, and (ii) in which either the surrogate or the true endpoint was a constant (i.e., all patients within a trial had the same surrogate and/or true endpoint value) are excluded. In addition, the user can specify the minimum number of patients that a trial should contain in order to include the trial in the analysis. If the number of patients in a trial is smaller than the value specified by `Min.Trial.Size`, the data of the trial are excluded. `Data.Analyze` is the dataset on which the surrogacy analysis was conducted.

`Obs.Per.Trial` A `data.frame` that contains the total number of patients per trial and the number of patients who were administered the control treatment and the experimental treatment in each of the trials (in `Data.Analyze`).

`Results.Stage.1`
: The results of stage 1 of the two-stage model fitting approach: a `data.frame` that contains the trial-specific intercepts and treatment effects for the surrogate and the true endpoints (when a full or semi-reduced model is requested), or the trial-specific treatment effects for the surrogate and the true endpoints (when a reduced model is requested).

`Residuals.Stage.1`
: A `data.frame` that contains the residuals for the surrogate and true endpoints that are obtained in stage 1 of the analysis (ε_{Sij} and ε_{Tij}).

`Fixed.Effect.Pars`
: A `data.frame` that contains the fixed intercept and treatment effects for S and T (i.e., μ_S, μ_T, α, and β) when a full, semi-reduced, or reduced model is fitted in stage 1.

`Random.Effect.Pars`
: A `data.frame` that contains the random intercept and treatment effects for S and T (i.e., m_{Si}, m_{Ti}, a_i and b_i) when a full or semi-reduced model is fitted in stage 1, or that contains the random treatment effects for S and T (i.e., a_i, and b_i) when a reduced model is fitted in stage 1.

`Results.Stage.2`
: An object of class `lm` (linear model) that contains the parameter estimates of the regression model that is fitted in stage 2 of the analysis.

`Trial.R2`
: A `data.frame` that contains the trial-level coefficient of determination (R^2_{trial}), its standard error and confidence interval.

`Indiv.R2`
: A `data.frame` that contains the individual-level coefficient of determination (R^2_{indiv}), its standard error and confidence interval.

`Trial.R`
: A `data.frame` that contains the trial-level correlation coefficient (R_{trial}), its standard error and confidence interval.

`Indiv.R`
: A `data.frame` that contains the individual-level correlation coefficient (R_{indiv}), its standard error and confidence interval.

`Cor.Endpoints`
: A `data.frame` that contains the correlations between the surrogate and the true endpoint in the control treatment group (i.e., ρ_{T0S0}) and in the experimental treatment group (i.e., ρ_{T1S1}), their standard errors and their confidence intervals.

`D.Equiv`
: The variance-covariance matrix of the trial-specific intercept and treatment effects for the surrogate and true endpoints (when a full or semi-reduced model is fitted, i.e., when `Model=c("Full")` or `Model=c("SemiReduced")` is used in the function call), or the variance-covariance matrix of the trial-specific treatment effects for the surrogate and true endpoints (when a reduced model is fitted, i.e., when `Model=c("Reduced")` is used in the function call). The variance-covariance matrix `D.Equiv` is equivalent to the $\boldsymbol{D}$ matrix that would be obtained when a (full or reduced) bivariate mixed-effects approach is used; see function `BimixedContCont`).

Author(s)

Wim Van der Elst, Ariel Alonso, & Geert Molenberghs

References

Burzykowski, T., Molenberghs, G., & Buyse, M. (2005). *The evaluation of surrogate endpoints*. New York: Springer-Verlag.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., & Geys, H. (2000). The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics, 1,* 49-67.

Tibaldi, F., Abrahantes, J. C., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., et al., (2003). Simplified hierarchical linear models for the evaluation of surrogate endpoints. *Journal of Statistical Computation and Simulation, 73,* 643-658.

See Also

UnifixedContCont, BifixedContCont, BimixedContCont, plot Meta-Analytic

Examples

```
## Not run:
# Conduct an analysis based on a simulated dataset with 2000 patients, 100 trials,
# and Rindiv=Rtrial=.8
# Simulate the data:
Sim.Data.MTS(N.Total=2000, N.Trial=100, R.Trial.Target=.8, R.Indiv.Target=.8,
Seed=123, Model="Reduced")

# Fit a reduced univariate mixed-effects model without weighting to assess surrogacy:
Sur <- UnimixedContCont(Dataset=Data.Observed.MTS, Surr=Surr, True=True, Treat=Treat,
Trial.ID=Trial.ID, Pat.ID=Pat.ID, Model="Reduced", Weighted=FALSE)

# Show a summary and plots of the results:
summary(Sur)
plot(Sur, Weighted=FALSE)
## End(Not run)
```

Index