Package: regtomean (via r-universe)

September 11, 2024

Type Package

Title Regression Toward the Mean

Version 1.1

Imports formattable, effsize, mefa, plyr, plotrix, sjPlot, sjmisc, sjlabelled,

Description In repeated measures studies with extreme large or small values it is common that the subjects measurements on average are closer to the mean of the basic population. Interpreting possible changes in the mean in such situations can lead to biased results since the values were not randomly selected, they come from truncated sampling. This method allows to estimate the range of means where treatment effects are likely to occur when regression toward the mean is present. Ostermann, T., Willich, Stefan N. & Luedtke, Rainer. (2008). Regression toward the mean - a detection method for unknown population mean based on Mee and Chua's algorithm. BMC Medical Research Methodology.<doi:10.1186/1471-2288-8-52>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Depends R (>= 2.10)

NeedsCompilation no

Author Daniela Recchia [aut, cre], Thomas Ostermann [ctb]

Maintainer Daniela Recchia <daniela.rodriguesrecchia@uni-wh.de>

Date/Publication 2022-10-26 13:52:37 UTC

Repository https://infinitomagico.r-universe.dev

RemoteUrl https://github.com/cran/regtomean

RemoteRef HEAD

RemoteSha d538b6f968d0dc505969fe61ca13ed21c1cae712

Contents

---

cordata   *Correlation and Cohen's d effect sizes.*

Description

This function calculates the correlation for the data and Cohen's d effect sizes, both based on pooled and on treatment standard deviations.

Usage

```
cordata(Before,After,data)
```

Arguments

Before   a numeric vector giving the data values for the first (before) measure.

After   a numeric vector giving the data values for the second (after) measure.

data   an optional data frame containing the variables in the formula. By default the variables are taken from environment (formula).

Details

This function computes the correlation between both measures as also both effect sizes based on Cohen's d statistic.

The inputs must be numeric.

Value

Return a table containing the correlation, effect size pooled and effect size based on treatment.

Author(s)

Daniela R. Recchia, Thomas Ostermann.

References

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York:Academic Press.

See Also

cohen.d,cor

Examples

```
cordata("Before","After",data=language_test)
```

---

language_test   *Language Test in High School*

Description

A dataset with scores from 8 students who failed a high school test and could not get their diploma. They repeated the exam and got new scores.

Usage

```
data("language_test")
```

Format

A data frame with 8 observations on the following 9 variables.

Student a numeric vector

Before a numeric vector

After a numeric vector

'Total N' a numeric vector

Cross a numeric vector

'Pre-treatment Mean' a numeric vector

'Pre-treatment Std' a numeric vector

'Post-treatment Mean' a numeric vector

'Post-treatment Std' a numeric vector

Source

McClave, J.T; Dietrich, F.H.:"Statistics";New York, Dellen Publishing; 1988.

Examples

```
data(language_test)
## maybe str(language_test) ; plot(language_test) ...
```

---

meechua_eff.CI	*Calculates and plots treatment and regression effects as also its p-values.*

Description

This function calculates and plots treatment and regression effects of both before and after measures as also its p-values.

Usage

```
meechua_eff.CI(x,n,se.after)
```

Arguments

x	a data frame containing the results from `meechua_reg`. It is stored as `mod_coef`.
n	the original sample size (number of observations) from data.
se.after	the estimated standard error from `meechua_reg`. It is stored as `se.after`.

Details

After performing the `meechua_reg` the model coefficients `mod_coef` as also its global variable `se.after` are used as input in this function to estimate treatment and regression effects.

Value

Two plots are performed, the first "Treatment Effect and p-value" and the second "Confidence Intervals" for `mu`.

Author(s)

Daniela R. Recchia, Thomas Ostermann

References

Ostermann, T., Willich, Stefan N. & Luedtke, Rainer. (2008). Regression toward the mean - a detection method for unknown population mean based on Mee and Chua's algorithm. BMC Medical Research Methodology.

See Also

[meechua_reg](#)

Examples

```
# First perform replicate_data and meechua_reg
replicate_data(50,60,"Before","After",data=language_test)
mee_chua_sort <- mee_chua[with(mee_chua,order(mu)),]
meechua_reg(mee_chua_sort)

# Model coeeficients (mod_coef) and se.after are stored in the environment
# as a result from the function meechua_reg
meechua_eff.CI(mod_coef,8,se.after)
```

---

meechua_plot   *Plot models from* meechua_reg

Description

This functions plots all 4 diagnostics plots for each linear regression model: "Residuals vs Fitted", "Normal Q-Q", "Scale-Location" and "Residuals vs Leverage".

Usage

```
meechua_plot(x)
```

Arguments

x   List containing the estimated linear models frommeechua_reg. It is stored as
                     models.

Details

For each model from models 4 diagnostic plots are performed. For the first model the numbers 1 to 4 should be given, for the second model numbers from to 8 to 12, and so on.

Value

Diagnostics plots for the set of models from meechua_reg.

Author(s)

Daniela R. Recchia, Thomas Ostermann.

References

Ostermann, T., Willich, Stefan N. & Luedtke, Rainer. (2008). Regression toward the mean - a detection method for unknown population mean based on Mee and Chua's algorithm. BMC Medical Research Methodology.

See Also

plot.lm,meechua_reg

Examples

```
# models are an output from meechua_reg
replicate_data(50,60,"Before","After",data=language_test)
mee_chua_sort <- mee_chua[with(mee_chua,order(mu)),]
meechua_reg(mee_chua_sort)

#models are the output from meechua_reg saved in the environment after running the function
meechua_plot(models)
```

---

meechua_reg   *Fit linear models on the (replication) data.*

Description

This function fit linear models for a subset of data frames.

Usage

```
meechua_reg(x)
```

Arguments

x   Data to be used in the regression.

Details

The data used for the regression must be sorted by `mu`.

A set of linear `models` will be estimated and model coefficients are saved and stored in `mod_coef`.

The estimated standard errror for the `after` measure is also stored in `se_after` to be used further in other functions.

Value

A table containing the estimations for each mu. Global variables `models`, `mod_coef`, `se_after` are stored for further analysis. The models are saved in an obejct called `mee_chua`, which is not automatically printted but is saved in the environment.

Author(s)

Daniela R. Recchia, Thomas Ostermann.

References

Ostermann, T., Willich, Stefan N. & Luedtke, Rainer. (2008). Regression toward the mean - a detection method for unknown population mean based on Mee and Chua's algorithm. BMC Medical Research Methodology.

See Also

lm,dlply

Examples

```
## get the values ##
mee_chua<-replicate_data(50,60,"Before","After",data=language_test)
## sort mu ##
mee_chua_sort <- mee_chua[with(mee_chua,order(mu)),]

meechua_reg(mee_chua_sort)
```

---

replicate_data	*Replicates before and after values 100 times.*

Description

This function replicates 100 times the before and after values giving a start and end reference.

Usage

```
replicate_data(start,end,by=NULL,Before,After,data)
```

Arguments

start	a start value for mu.
end	an end value for mu.
by	an increment for the sequence, if not informed (default) the increment is caluculated as the difference between end-start.
Before	a numeric vector giving the data values for the first (before) measure.
After	a numeric vector giving the data values for the second (after) measure.
data	an optional data frame containing the before and after variables in the formula. By default the variables are taken from environment (formula)

.

Details

In order to overcome the limitation of Mee and Chua's test regarding the population mean mu a replication of the data is performed.

After replicating the data the unknown population mean mu is systematically estimated over a range of values. Further estimations will be based on this new dataset.

Value

Return a data frame we could call mee_chua containing the values for mu, before and after.

Author(s)

Daniela R. Recchia, Thomas Ostermann.

References

Ostermann, T., Willich, Stefan N. & Luedtke, Rainer. (2008). Regression toward the mean - a detection method for unknown population mean based on Mee and Chua's algorithm. BMC Medical Research Methodology.

Galton, F. (1886). Regression towards mediocrity in heriditary stature. Journal of the Anthropological Institute (I5: 246-263).

See Also

rep

Examples

```
replicate_data(50,60,"Before","After",data=language_test)
```

Index