Package: granovaGG (via r-universe)

September 1, 2024

Version 1.4.1.9000

Date 2023-11-22

Title Graphical Analysis of Variance Using ggplot2

Description Create what we call Elemental Graphics for display of
anova results. The term elemental derives from the fact that
each function is aimed at construction of graphical displays
that afford direct visualizations of data with respect to the
fundamental questions that drive the particular anova methods.
This package represents a modification of the original granova
package; the key change is to use 'ggplot2', Hadley Wickham's
package based on Grammar of Graphics concepts (due to
Wilkinson). The main function is granovagg.1w() (a graphic for
one way ANOVA); two other functions (granovagg.ds() and
granovagg.contr()) are to construct graphics for dependent
sample analyses and contrast-based analyses respectively. (The
function granova.2w(), which entails dynamic displays of data,
is not currently part of 'granovaGG'.) The 'granovaGG'
functions are to display data for any number of groups,
regardless of their sizes (however, very large data sets or
numbers of groups can be problematic). For granovagg.1w() a
specialized approach is used to construct data-based contrast
vectors for which anova data are displayed. The result is that
the graphics use a straight line to facilitate clear
interpretations while being faithful to the standard effect
test in anova. The graphic results are complementary to
standard summary tables; indeed, numerical summary statistics
are provided as side effects of the graphic constructions.
granovagg.ds() and granovagg.contr() provide graphic displays
and numerical outputs for a dependent sample and contrast-based
analyses. The graphics based on these functions can be
especially helpful for learning how the respective methods work
to answer the basic question(s) that drive the analyses. This
means they can be particularly helpful for students and
non-statistician analysts. But these methods can be of
assistance for work-a-day applications of many kinds, as they

1

can help to identify outliers, clusters or patterns, as well as
highlight the role of non-linear transformations of data. In
the case of granovagg.1w() and granovagg.ds() several arguments
are provided to facilitate flexibility in the construction of
graphics that accommodate diverse features of data, according
to their corresponding display requirements. See the help files
for individual functions.

URL https://github.com/briandk/granovaGG

BugReports https://github.com/briandk/granovaGG/issues

License MIT + file LICENSE

Depends R (>= 2.14.0),

Imports dplyr, ggplot2 (>= 0.9.2), magrittr, RColorBrewer, tibble,
tidyr

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Collate 'granovaGG-package.R' 'granovagg.1w-helpers.R'
'theme-defaults.R' 'shared-functions.R' 'granovagg.1w.R'
'granovagg.contr.R' 'granovagg.ds.R'

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

RemoteUrl https://github.com/briandk/granovagg

RemoteRef HEAD

RemoteSha 7014d7472f863f68cedce97a4d3f232056581ad1

Contents

---

granovaGG-package   *Elemental Graphics for Analysis of Variance Using ggplot2*

Description

This collection of functions in granovaGG provides what we call elemental graphics for display of
anova results. The term elemental derives from the fact that each function is aimed at construction
of graphical displays that afford direct visualizations of data with respect to the fundamental questions that drive the particular anova methods. This package represents a modification of the original
granova package; the key change is to use ggplot2, Hadley Wickham's package based on Grammar
of Graphics concepts (due to Wilkinson). The main function is granovagg.1w (a graphic for one
way anova); two other functions (granovagg.ds and granovagg.contr) are to construct graphics for
dependent sample analyses and contrast-based analyses respectively. (The function granova.2w,
which entails dynamic displays of data, is not currently part of granovaGG.) The granovaGG functions are to display data for any number of groups, regardless of their sizes (however, very large
data sets or numbers of groups can be problematic). For granovagg.1w a specialized approach is
used to construct data-based contrast vectors for which anova data are displayed. The result is that
the graphics use a straight line to facilitate clear interpretations while being faithful to the standard
effect test in anova. The graphic results are complementary to standard summary tables; indeed, numerical summary statistics are provided as side effects of the graphic constructions. granovagg.ds
and granovagg.contr provide graphic displays and numerical outputs for a dependent sample and
contrast-based analyses. The graphics based on these functions can be especially helpful for learning how the respective methods work to answer the basic question(s) that drive the analyses. This
means they can be particularly helpful for students and non-statistician analysts. But these methods
can be of assistance for work-a-day applications of many kinds, as they can help to identify outliers,
clusters or patterns, as well as highlight the role of non-linear transformations of data. In the case of
granovagg.1w and granovagg.ds several arguments are provided to facilitate flexibility in the construction of graphics that accommodate diverse features of data, according to their corresponding
display requirements. See the help files for individual functions.

Details

Package:	granovaGG
Version:	1.0
License:	GPL (>= 2)

Author(s)

Brian A. Danielak <brian@briandk.com>
Robert M. Pruzek <RMPruzek@yahoo.com>

with contributions by:
William E. J. Doane <wil@DrDoane.com>
James E. Helmreich <James.Helmreich@Marist.edu>
Jason Bryer <jason@bryer.org>

References

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. New York: Springer.

Wilkinson, L. (1999). The Grammar of Graphics. Statistics and computing. New York: Springer.

See Also

granovagg.1w granovagg.ds granovagg.contr

---

anorexia   *Anorexia Data on Weight Change*

Description

The anorexia data frame has 72 rows and 3 columns. Weight change data for young female anorexia patients.

Format

A dataframe with 72 observations of three variables:

Treat Factor of three levels: "Cont" (control), "CBT" (Cognitive Behavioural treatment) and "FT" (family treatment).

Prewt Pretreatment weight of subject, in pounds.

Postwt Postreatment weight of subject, in pounds.

Source

Hand, D. J., Daly, F., McConway, K., Lunn, D. and Ostrowski, E. eds (1993) A Handbook of Small Data Sets. Chapman & Hall, Data set 285 (p. 229)

References

Venables, W. N. and Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth edition. Springer.

anorexia.sub	*Family Treatment Weight change data for young female anorexia patients (subset).*

Description

The MASS package includes the dataset anorexia, containing pre and post treatment weights for young female anorexia patients. This is a subset of those data, containing only those patients who received Family Treatment.

Format

A dataframe with 17 observations on the following 2 variables, no NAs.

Prewt Pretreatment weight of subject, in pounds.

Postwt Postreatment weight of subject, in pounds.

Source

Hand, D. J., Daly, F., McConway, K., Lunn, D. and Ostrowski, E. eds (1993) A Handbook of Small Data Sets. Chapman & Hall, Data set 285 (p. 229)

References

Venables, W. N. and Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth edition. Springer.

arousal	*Arousal in Rats*

Description

40 rats were given divided randomly into four groups and assigned to one of four treatments: placebo, drug A, drug B, or both drug A and drug B. Response is a standard measure of physiological arousal.

Format

A data frame with 40 observations, 10 in each of 4 columns the corresponding to placebo, drug A, drug B and both drug A and drug B; no NAs.

Placebo Rats receiving a placebo treatment.

Drug.A Rats receiving only drug A.

Drug.B Rats receiving only drug B.

Drug.A.B Rats receiving both drug A and drug B.

Source

Richard Lowry. Concepts & Applications of Inferential Statistics. Vassar College, Poughkeepsie, N.Y., 2010, http://faculty.vassar.edu/lowry/webtext.html

---

blood_lead	*Blood lead levels of lead workers' children matched with similar control children.*

Description

Children of parents who had worked in a factory where lead was used in making batteries were matched by age, exposure to traffic, and neighborhood with children whose parents did not work in lead-related industries. Whole blood was assessed for lead content yielding measurements in mg/dl

Format

A dataframe with 33 observations on the following 2 variables, no NAs.

Exposed Blood lead level of exposed child, mg/dl.

Control Blood lead level of exposed child, mg/dl.

Source

Morton, D., Saah, A., Silberg, S., Owens, W., Roberts, M., Saah, M. (1982). Lead absorption in children of employees in a lead related industry. American Journal of Epidemiology, 115:549-555.

References

See discussion in Section 2.5 of Enhancing Dependent Sample Analyses with Graphics, Journal of Statistics Education Volume 17, Number 1 (March 2009).

---

granovagg.1w	*Elemental Graphic Display for One-Way ANOVA*

Description

Graphic to display data for a one-way analysis of variance – that is for unstructured groups. Also to help understand how data play out in the context of the basic one-way model, how the F statistic is generated for the data at hand, etc. The graphic may be called 'elemental' or 'natural' because it is built upon the central question that drives one-way ANOVA (see details below).

Usage

```
granovagg.1w(
  data,
  group = NULL,
  h.rng = 1,
  v.rng = 1,
  jj = NULL,
  dg = 2,
  resid = FALSE,
  print.squares = TRUE,
  xlab = "default_x_label",
  ylab = "default_y_label",
  main = "default_granova_title",
  plot.theme = "theme_granova_1w",
  ...
)
```

Arguments

`data`	Dataframe or vector. If a dataframe, the two or more columns are taken to be groups of equal size (whence `group` is NULL). If `data` is a vector, `group` must be a vector, perhaps a factor, that indicates groups (unequal group sizes allowed with this option).
`group`	Group indicator, generally a factor in case `data` is a vector.
`h.rng`	Numeric; controls the horizontal spread of groups, default = 1
`v.rng`	Numeric; controls the vertical spread of points, default = 1.
`jj`	Numeric; sets horiz. jittering level of points. `jj` gets passed as the `amount` parameter to [jitter](). When `jj` = NULL (the default behavior), the degree of jitter will take on a sensible value. In addition, if pairs of ordered means are close to one another and `jj` = NULL, the degree of jitter will default to the smallest difference between two adjacent contrasts.
`dg`	Numeric; sets number of decimal points in output display, default = 2
`resid`	Logical; displays marginal distribution of residuals (as a 'rug') on right side (wrt grand mean), default = FALSE.
`print.squares`	Logical; displays graphical squares for visualizing the F-statistic as a ratio of MS-between to MS-within
`xlab`	Character; horizontal axis label, can be supplied by user, default = `"default_x_label"`, which leads to a generic x-axis label ("Contrast coefficients based on group means").
`ylab`	Character; vertical axis label, can be supplied by user, default = `"default_y_label"`, which leads to a generic y-axis label ("Dependent variable (response)").
`main`	Character; main label, top of graphic; can be supplied by user, default = `"default_granova_title"`, which will print a generic title for graphic.
`plot.theme`	argument indicating a ggplot2 theme to apply to the graphic; defaults to a customized theme created for the one-way graphic
`...`	Optional arguments to/from other functions

Details

The one-way ANOVA graphic shows how the comparison of unstructured groups, viz. their means, entails a particular linear combination (L.C.) of the group means. In particular, we use the fact that the numerator of the one-way F statistic, the mean square between (MS.B), is a linear combination of the group means; each weight – one for each group – in the L.C. is (principally) a function of the difference between the group's mean and the grand mean, viz., (M~j~ - M..) where M~j~ denotes the jth group's mean, and M.. denotes the grand mean. The L.C. can be written as a sum of products of the form MS.B = Sum((1/df.B)(n_j (M_j - M..) M_j)) for j = 1...J. The denominator of the F-statistic, MS.W (mean square within), can be described as a 'scaling factor'. It is just the (weighted) average of the variances of the J groups (j = 1 ... J). (n~j~'s are group sizes.) The differences (M~j~ - M..) are themselves the 'effects' in the analysis. When the effects are plotted against the group means (the horizontal and vertical axes) a straight line necessarily ensues. Group means are plotted as triangles along this line. Once the means have been plotted, the data points (jittered) for the groups are displayed (vertical axis) with respect to the respective contrasts. Since the group means are just the fitted values in one-way ANOVA, and the deviations of the scores within groups are the residuals (subsetted by groups), the graphic can be seen as showing fitted vs. residual values for the line that shows the locus of ordered group means – from the smallest on the left) the the largest (on the right). If desired, the aggregate of all such residuals can be plotted (as a rug plot) on the right margin of the graphic centered on the grand mean (large green dot in 'middle'). The use of effects to locate groups this way yields what we term an 'elemental' graphic because it is based on the central question that drives one-way ANOVA.

Note that groups need not have the same size, nor do data need to reflect any particular distributional characteristics. Finally, the gray bars (one for each group) at the bottom of the graphic show the relative sizes of the group standard deviations with referene to the 'average' group s.d. (more precisely, the square root of the MS.W). This 'average' corresponds to the thin white line that runs horizontally across these bars.

Value

Returns a plot object of class `ggplot`. The function also provides printed output including by-group statistical summaries and information about groups that might be overplotted (if applicable):

`group`	group names
`group means`	means for each group
`trimmed.mean`	20% trimmed group means
`contrast`	Contrasts (group main effects)
`variance`	variances
`standard.deviation`	
	standard deviations
`group.size`	group sizes
`overplotting information`	
	Information about groups that, due to their close means, may be overplotted

Author(s)

Brian A. Danielak <brian@briandk.com>
Robert M. Pruzek <RMPruzek@yahoo.com>

with contributions by:
William E. J. Doane <wil@drdoane.com>
James E. Helmreich <James.Helmreich@Marist.edu>
Jason Bryer <jason@bryer.org>

References

Fundamentals of Exploratory Analysis of Variance, Hoaglin D., Mosteller F. and Tukey J. eds., Wiley, 1991.

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. New York: Springer.

Wilkinson, L. (1999). The Grammar of Graphics. Statistics and computing. New York: Springer.

See Also

granovagg.contr, granovagg.ds, granovaGG

Examples

```
data(arousal)
#Drug A
granovagg.1w(arousal[,1:2], h.rng = 1.6, v.rng = 0.5)


###

data(anorexia)
wt.gain <- anorexia[, 3] - anorexia[, 2]
granovagg.1w(wt.gain, group = anorexia[, 1])


###

data(poison)
##Note violation of constant variance across groups in following graphic.
granovagg.1w(poison$SurvTime, group = poison$Group, ylab = "Survival Time")
##RateSurvTime = SurvTime^-1
granovagg.1w(poison$RateSurvTime, group = poison$Group,
ylab = "Survival Rate = Inverse of Survival Time")

##Nonparametric version: RateSurvTime ranked and rescaled
##to be comparable to RateSurvTime;
##note labels as well as residual (rug) plot below.
granovagg.1w(poison$RankRateSurvTime, group = poison$Group,
ylab = "Ranked and Centered Survival Rates",
main = "One-way ANOVA display, poison data (ignoring 2-way set-up)",
res = TRUE)


###

data(chickwts)
?chickwts # An explanation of the chickwts dataset
with(chickwts, granovagg.1w(weight, group = feed)) # Modeling weight as explained by feed type
```

---

granovagg.contr   *Elemental Graphic Display for Contrast Effect of ANOVA*

Description

Provides graphic displays that shows data and effects for a priori contrasts in ANOVA contexts; also corresponding numerical results.

Usage

```
granovagg.contr(
  data,
  contrasts,
  ylab = "default_y_label",
  plot.theme = "theme_granova_contr",
  jj = 1,
  ...
)
```

Arguments

data	Vector of scores for all equally sized groups, or a data.fame or matrix where each column represents a group.
contrasts	Matrix of column contrasts with dimensions (number of groups [G]) x (number of contrasts) [generally (G x G-1)].
ylab	Character; y axis label. Defaults to a generic granova title.
plot.theme	argument indicating a ggplot2 theme to apply to the graphic; defaults to a customized theme created for the contrast graphic
jj	Numeric; controls [jitter](#) and allows you to control the degree of jitter in the contrast plots. jj is divided by 100 and passed as the width parameter to [position_jitter](#).
...	Optional arguments to/from other functions.

Details

Function provides graphic displays of contrast effects for prespecified contrasts in ANOVA. Data points are displayed as relevant for each contrast based on comparing groups according to the positive and negative contrast coefficients for each contrast on the horizontal axis, against response values on the vertical axis. Data points corresponding to groups not being compared in any contrast (coefficients of zero) are ignored. For each contrast (generally as part of a 2 x 2 panel) a line segment is given that compares the (weighted) mean of the response variable for the negative coefficients versus the positive coefficients. Standardized contrasts are used, wherein the sum of (magnitudes) of negative coefficients is unity; and the same for positive coefficients. If a line is 'notably' different from horizontal (i.e. slope of zero), a 'notable' effect has been identified; however, the question of statistical significance generally depends on a sound context-based estimate of standard error for the corresponding effect. This means that while summary aov numerical results and

test statistics are presented (see below), the appropriateness of the default standard error generally requires the analyst's judgment. The response values are to be input in (a stacked) form, i.e. as a vector, for all cells (cf. arg. ylab). The matrix of contrast vectors contrasts must have G rows (the number of groups), and a number of columns equal to the number of prespecified contrasts, at most G-1. If the number of columns of contrasts is G-1, then the number per group, or cell size, is taken to be length(data)/G, where G = nrow(contrasts).

If the number of columns of contrasts is less than G-1 then the user must stipulate npg, the number in each group or cell. The function is designed for the case when all cell sizes are the same, and may be most helpful when the a priori contrasts are mutually orthogonal (e.g., in power of 2 designs, or their fractional counterparts; also when specific row or column comparisons, or their interactions (see the example below based on rat weight gain data)). It is not essential that contrasts be mutually orthogonal; but mutual linear independence is required. (When factor levels correspond to some underlying continuum a standard application might use con = contr.poly(G), for G the number of groups; consider also contr.helmert(G).) The final plot in each application shows the data for all groups or cells in the design, where groups are simply numbered from 1:G, for G the number of groups, on the horizontal axis, versus the response values on the vertical axis.

Value

a list of ggplot objects, one element per plot. That allows you to access any individual plot or plots, then modify them as you wish (with ggplot2 commands, for example).

The function also provides printed output:

Weighted Means	Table showing the (weighted) means for positive and negative coefficients for each (row) contrast, and for each row, the difference between these means, and the standardized effect size in the final column.
summary.lm	Summary results for a linear model analysis based on the R function lm (When effects are simple, as in an equal n's power of 2 design, mean differences will generally correspond to the linear regression coefficients as seen in the lm summary results.)
Contrasts	The contrast matrix you specified.

Author(s)

Brian A. Danielak <brian@briandk.com>
Robert M. Pruzek <RMPruzek@yahoo.com>

with contributions by:
William E. J. Doane <wil@drdoane.com>
James E. Helmreich <James.Helmreich@Marist.edu>
Jason Bryer <jason@bryer.org>

References

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. New York: Springer.

Wilkinson, L. (1999). The Grammar of Graphics. Statistics and computing. New York: Springer.

See Also

granovagg.1w, granovagg.ds, granovaGG

Examples

```
data(arousal)
contrasts22 <- data.frame( c(-.5,-.5,.5,.5),
  c(-.5,.5,-.5,.5), c(.5,-.5,-.5,.5) )
names(contrasts22) <- c("Drug.A", "Drug.B", "Drug.A.B")
granovagg.contr(arousal, contrasts = contrasts22)

data(rat)
dat6 <- matrix(c(1, 1, 1, -1, -1, -1, -1, 1, 0, -1, 1, 0, 1, 1, -2,
    1, 1, -2, -1, 1, 0, 1, -1, 0, 1, 1, -2, -1, -1, 2), ncol = 5)
granovagg.contr(rat[,1], contrasts = dat6, ylab = "Rat Weight Gain",
  xlab = c("Amount 1 vs. Amount 2", "Type 1 vs. Type 2",
  "Type 1 & 2 vs Type 3", "Interaction of Amount and Type 1 & 2",
  "Interaction of Amount and  Type (1, 2), 3"))
#Polynomial Contrasts
granovagg.contr(rat[,1],contrasts = contr.poly(6))

#based on random data
data.random <- rt(64, 5)
granovagg.contr(data.random, contrasts = contr.helmert(8),
  ylab = "Random Data")
```

---

granovagg.ds   *Elemental Graphic for Display of Dependent Sample Data*

Description

Plots dependent sample data beginning from a scatterplot for the X,Y pairs; proceeds to display difference scores as point projections; also X and Y means, as well as the mean of the difference scores.

Usage

```
granovagg.ds(
  data = NULL,
  revc = FALSE,
  main = "default_granova_title",
  xlab = NULL,
  ylab = NULL,
  conf.level = 0.95,
  plot.theme = "theme_granova_ds",
  northeast.padding = 0,
  southwest.padding = 0,
  ...
)
```

Arguments

`data`	is an n X 2 dataframe or matrix. First column defines X (intially for horzontal axis), the second defines Y.
`revc`	reverses X,Y specifications
`main`	optional main title (as character); can be supplied by user. The default value is `"default_granova_title"`, which will print a generic title for the graphic.
`xlab`	optional label (as character) for horizontal axis. If not defined, axis labels are taken from colnames of data.
`ylab`	optional label (as character) for vertical axis. If not defined, axis labels are taken from colnames of data.
`conf.level`	The confidence level at which to perform a dependent sample t-test. Defaults to `0.95` (95% Confidence)
`plot.theme`	argument indicating a ggplot2 theme to apply to the graphic; defaults to a customized theme created for the dependent sample graphic
`northeast.padding`	(numeric) extends axes toward lower left, effectively moving data points to the southwest. Defaults to zero padding.
`southwest.padding`	(numeric) extends axes toward upper right, effectively moving data points to the southwest. Defaults to zero padding. Making both southwest and northeast padding smaller moves points farther apart, while making both larger moves data points closer together.
`...`	Optional arguments to/from other functions

Details

Paired X and Y values are plotted as scatterplot. The identity reference line (for Y = X) is drawn. Parallel projections of data points to (a lower-left) line segment show how each point relates to its X-Y = D difference; semitransparent "shadow" points are used to display the distribution of difference scores, with thin grey lines leading from each raw datapoint to its shadow projection on the difference distribution. The range of that difference score distribution is drawn as a blue line beneath the shadow points and the mean difference is displayed as a heavy dashed purple line, parallel to the identity reference line. Means for X and Y are also plotted (as thin dashed vertical and horizontal lines), and rug plots are shown for the distributions of X (at the top of graphic) and Y (on the right side). The 95% confidence interval for the population mean difference is also shown graphically as a green band, perpendicular to the mean treatment effect line. Because all data points are plotted relative to the identity line, and summary results are shown graphically, clusters, data trends, outliers, and possible uses of transformations are readily seen, possibly to be accommodated.

In summary, the graphic shows all initial data points relative to the identity line, adds projections (to the 'north' and 'east') showing the marginal distributions of X and Y, as well as projections to the 'southwest' where the difference scores for each point are drawn. Means for all three distributions are shown using straight lines; the confidence interval for the population mean difference score is also shown. Summary statistics are printed as side effects of running the function for the dependent sample analysis.

Value

Returns a plot object of class `ggplot`.

Author(s)

Brian A. Danielak <brian@briandk.com>
Robert M. Pruzek <RMPruzek@yahoo.com>

with contributions by:
William E. J. Doane <wil@drdoane.com>
James E. Helmreich <James.Helmreich@Marist.edu>
Jason Bryer <jason@bryer.org>

References

Pruzek, R. M., & Helmreich, J. E. (2009). Enhancing Dependent Sample Analyses with Graphics. Journal of Statistics Education, 17(1), 21.

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. New York: Springer.

Wilkinson, L. (1999). The Grammar of Graphics. Statistics and computing. New York: Springer.

See Also

`granovagg.1w`, `granovagg.ds`, `granovaGG`

Examples

```
### Using granovagg.ds to examine trends or effects for repeated measures data.

# This example corresponds to case 1b in Pruzek and Helmreich (2009). In this
# graphic we're looking for the effect of Family Treatment on patients with anorexia.

data(anorexia.sub)

granovagg.ds(anorexia.sub,
             revc = TRUE,
             main = "Assessment Plot for weights to assess \
                     Family Therapy treatment for Anorexia Patients",
             xlab = "Weight after therapy (lbs.)",
             ylab = "Weight before therapy (lbs.)"
)

### Using granovagg.ds to compare two experimental treatments (with blocking)

# This example corresponds to case 2a in Pruzek and Helmreich (2009). For this
# data, we're comparing the effects of two different virus preparations on the
# number of lesions produced on a tobacco leaf.

data(tobacco)
granovagg.ds(tobacco[, c("prep1", "prep2")],
             main = "Local Lesions on Tobacco Leaves",
             xlab = "Virus Preparation 1",
```

```
                    ylab = "Virus Preparation 2"
)


### Using granovagg.ds to compare two experimental treatments (with blocking)

# This example corresponds to case 2a in Pruzek and Helmreich (2009). For this
# data, we're comparing the wear resistance of two different shoe sole
# materials, each randomly assigned to the feet of 10 boys.

data(shoes)
granovagg.ds(shoes,
             revc = TRUE,
             main = "Shoe Wear",
             xlab = "Sole Material B",
             ylab = "Sole Material A",
)


### Using granovagg.ds to compare matched individuals for two treatments

# This example corresponds to case 2b in Pruzek and Helmreich (2009). For this
# data, we're examining the level of lead (in mg/dl) present in the blood of
# children. Children of parents who had worked in a factory where lead was used
# in making batteries were matched by age, exposure to traffic, and neighborhood
# with children whose parents did not work in lead-related industries.

data(blood_lead)
granovagg.ds(blood_lead,
             sw = .1,
             main = "Dependent Sample Assessment Plot
             Blood Lead Levels of Matched Pairs of Children",
             xlab = "Exposed (mg/dl)",
             ylab = "Control (mg/dl)"
)
```

---

poison	*Poison data from Biological Experiment*

Description

Survial times of animals in a 3 x 4 factorial experiment involving poisons (3 levels) and various treatments (four levels), as described in Chapter 8 of Box, Hunter and Hunter.

Format

This data frame was originally poison.data from the package BHH2, but as presented here has added columns; no NAs.

Poison Factor with three levels I, II, and III.

Treatment Factor with four levels, A, B, C, and D.

Group Factor with 12 levels, 1:12.

SurvTime Numeric; survival time.

RateSurvTime Numeric; inverse of SurvTime

RankRateSurvTime Numeric; RateSurvTime scores have been converted to ranks, and then rescaled to have the same median as and a spread comparable to RateSurvTime

Source

Box, G. E. P. and D. R. Cox, An Analysis of Transformations (with discussion), Journal of the Royal Statistical Society, Series B, Vol. 26, No. 2, pp. 211 - 254.

References

Box G. E. P, Hunter, J. S. and Hunter, W. C. (2005). Statistics for Experimenters II. New York: Wiley.

---

rat   *Weight gains of rats fed different diets*

Description

60 rats were fed varying diets to see which produced the greatest weight gain. Two diet factors were protein type: beef, pork, chicken and protein level: high and low.

Format

A data frame with 60 observations on the following 3 variables, no NAs.

Weight.Gain Weight gain (grams) of rats fed the diets.

Diet.Amount Amount of protein in diet: 1 = High, 2 = Low.

Diet.Type Type of protein in diet: 1 = Beef, 2 = Pork, 3 = Cereal.

Source

Fundamentals of Exploratory Analysis of Variance, Hoaglin D., Mosteller F. and Tukey J. eds., Wiley, 1991, p. 100; originally from Statistical Methods, 7th ed, Snedecor G. and Cochran W. (1980), Iowa State Press.

---

shoes   *Shoe wear data of Box, Hunter and Hunter*

Description

A list of two vectors, giving the wear of shoes of materials A and B for one foot each of ten boys.

Source

G. E. P. Box, W. G. Hunter and J. S. Hunter (1978) *Statistics for Experimenters*. Wiley, p. 100

References

Venables, W. N. and Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth edition. Springer.

---

tobacco   *Virus Preparation on Tobacco Leaves*

Description

This data is taken from Snedecor and Cochran (1980) and corresponds to a true matched pairs experiment. The data originally came from Youden and Beale in 1934 who "wished to find out if two preparations of a virus would produce different effects on tobacco plants. Half a leaf of a tobacco plant was rubbed with cheesecloth soaked in one preparation of the virus extract, and the second half was rubbed similarly with the second extract." (Page 86, Snedecor and Cochran, 1980) Each of the 8 points in the figure corresponds to the numbers of lesions on the two halves of one leaf with sides that had been treated differently.

Format

A dataframe with 8 observations on the following 2 variables, no NAs

prep1 Virus Preparation 1
prep2 Virus Preparation 2

Source

Youden, W. J., Beale, H. P. (1934). A statistical study of the local lesion method for estimating tobacco mosaic virus. In Contributions from Boyce Thompson Institute 6, page 437.

References

Snedecor, W., Cochran, W. (1980). Statistical methods. Iowa State University Press, Ames Iowa, seventh edition.

Index