Package: labstats (via r-universe)

November 16, 2026

Type Package

Title Data Sets for the Book ``Experimental Design for Laboratory Biologists"

Version 1.0

Date 2016-01-14

Author Stanley E. Lazic

Maintainer Stanley E. Lazic <stan.lazic@cantab.net>

Description Contains data sets to accompany the book ``Experimental Design for Laboratory Biologists: Maximising Information and Improving Reproducibility" (available mid 2016 from Cambridge University Press).

Depends R (>= 2.10)

License GPL-3

LazyData true

URL https://github.com/stanlazic/labstats

BugReports https://github.com/stanlazic/labstats/issues

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

Date/Publication 2016-12-04 21:51:27 UTC

RemoteUrl https://github.com/stanlazic/labstats

RemoteRef HEAD

RemoteSha e031bddfb246bfb53d52f7a70d25cf68a88d49ed

Contents

assay.window 2
block.covars 2
cellcount 3
CV 4
festing 4
fluoxetine 5

glycogen 5
goldstandard 6
hypertension 6
KH2004 7
locomotor 8
VPA 8

Index **10**

`assay.window` *Data to calculate an assay window*

Description

Simulated data for two variables to calculate an assay window between positive and negative control samples.

Format

A data frame with 40 rows and 3 variables:

A: Values for the first outcome variable.

B: Values for the second outcome variable.

condition: Indicates if a sample is a positive or negative control.

`block.covars` *Data to compare the use of blocking and covariate adjustment*

Description

Simulated data to illustrate the effects of blocking versus adjusting for a covariate.

Format

A data frame with 8 rows and 6 variables:

weight: A baseline measurement of body weight for eight rats. The rows of the data frame are sorted according to weight (lightest to heaviest).

block: Rats are grouped into four blocks based on body weight.

RBD: Treatment group when using a randomised block design (RBD).

CRD: Treatment group when using a completely randomised design (CRD).

y.RBD: Outcome variable under the RBD.

y.CRD: Outcome variable under the CRD.

Details

The experimental manipulation is a new diet versus a standard control diet and the outcome is the amount of food eaten on each diet. Since rats that weigh more at the beginning of the experiment are expected to eat more food, regardless of the diet, it would be beneficial to account for this source of variation. This can be done either through use of blocking or covariate adjustment and data for both designs are included. Note that only one design could be used in a real experiment but here we generate outcome values for two experiments using the same baseline body weight values.

For the randomised block design the eight rats are ranked according to baseline body weight and grouped into four blocks of two (the two lightest rats form the first block, the next two the second, and so on). Assignment to treatment group is done within blocks.

Examples

```
# Randomised block design
summary(aov(y.RBD ~ factor(block) + RBD, data=block.covars))

# Completely randomised design with weight as a covariate
summary(aov(y.CRD ~ weight + CRD, data=block.covars))
```

---

cellcount *Cell count data*

Description

Cell count data from a high-throughput screen with six 1536 well plates.

Format

A data frame with 9216 rows and 4 variables:

plate: Name of plate.

row: Row position of the well.

column: Column position of the well.

cell.count: Number of cells in the well.

CV *Data to estimate the coefficient of variation*

Description

Simulated data for two outcome variables.

Format

A data frame with 15 rows and 3 variables:

Sample: Sample identification number.

A: The first outcome variable.

B: The second outcome variable.

Details

The data contains five samples with three measurements made on each sample; these are "technical replicates" and do not contribute to the sample size. Which of the two outcome variables, A or B, would make a better primary outcome?

festing *Mouse liver enzyme randomised block design data set*

Description

Sixteen mice from four strains were used; half were assigned to the treated condition and half to the control. The experiment was split into two batches, which were run two months apart.

Format

A data frame with 16 rows and 4 variables:

strain: Four strains of mice were used: NIH, BALB/c, A/J, and 129/Ola.

treatment: Presence or absence of diallyl sulphide.

batch: The experiment was conducted in two batches.

value: Levels of glutathione-S-transferase (Gst).

Details

The purpose of the experiment was to test if diallyl sulphide (DS) affects the activity of the liver enzyme Gst. Four strains of mice were used and the experiment was conducted in two batches, where the housing conditions differed between batches. Both batch and strain are considered blocking variables.

References

Festing MFW (2014). Randomized block experimental designs can increase the power and reproducibility of laboratory animal experiments. *ILAR Journal* 55(3):472-476.

`fluoxetine` *Fluoxetine data set*

Description

Twenty rats were given fluoxetine and tested in the forced swim test (FST).

Format

A data frame with 20 rows and 2 variables:

dose: Dose of fluoxetine (0, 80, 160, or 240 mg/L).

time.immob: Total immobility time in the FST (seconds).

Details

Twenty rats were randomly assigned to a control condition or three doses of fluoxetine. Fluoxetine was administered in the drinking water. The amount of time the rats were immobile in the FST was measured and greater immobility time implies a more depressive phenotype.

References

Lazic SE (2008). Why we should use simpler models if the data allow this: relevance for ANOVA designs in experimental biology. *BMC Physiology* 8:16.

`glycogen` *Glycogen content in rat livers*

Description

An experiment measuring the glycogen content in rat livers that uses subsampling.

Format

A data frame with 36 rows and 4 variables:

Glycogen: Glycogen content.

Treatment: A numeric treatment group indicator (1, 2, 3).

Rat: Rat identification number. The numbers are not unique; each treatment has a rat numbered 1 and 2, but these not the same rats (rats are nested under treatment).

Liver: A numeric variable indicating the liver piece. Each liver was divided into three pieces.

Details

Six rats were randomised to three treatment conditions (two per condition). Their livers were divided into three pieces and two measurements were taken on each piece for 6 x 3 x 2 = 36 observations. The rats are the experimental units and there are two levels of subsampling.

References

Sokal RR and Rohlf FJ (1995). *Biometry*. WH Freeman and Co., New York, NY.

`goldstandard` *Data to compare a new method against a gold standard*

Description

Simulated data for one gold standard and four other outcome variables.

Format

A data frame with 20 rows and 5 variables:

goldstandard: Values for the gold standard outcome variable.

perfect.cor: An outcome variable that is identical to the gold standard.

noise: An outcome variable created by adding noise to the gold standard.

location: An outcome variable created by adding noise plus a shift in location.

scale: An outcome variable created by adding noise plus a shift in scale.

Details

The data contains a gold standard plus four other simulated variables that vary from the gold standard in different ways.

`hypertension` *Hypertension repeated measures data*

Description

Repeated measures data set from Casella (2008) testing the effect of high and low dietary calcium on blood pressure.

Format

A data frame with 30 rows and 4 variables:

Subject: Patient identification number.

Diet: High or low calcium condition.

Time: One of three time points (units unknown).

Y: Blood pressure, the outcome variable.

Details

Ten patients were randomly assigned to a high or low dietary calcium condition (five per group). Blood pressure measurements were taken at three time points.

References

Casella G (2008). *Statistical Design*. Springer, New York, NY.

KH2004 *Repeated measures data on rat muscles*

Description

Repeated measures data set from Kristensen and Hansen (2004) testing the effect of pinacidil on the force of muscle contraction.

Format

A data frame with 210 rows and 4 variables:

time: Observation number, from 0 to 14. Observations were taken every 30 seconds.

cond: Condition, either Placebo or Pinacidil.

rat: Rat identification number.

values: Force of muscle contraction, normalised to the first time point.

Details

Seven rats were euthanized and both soleus muscles (lower leg) were extracted from each rat. One muscle from each pair was randomised to the pinacidil condition and the other to the control condition. Fifteen measurements of the force of muscle contraction were taken every thirty seconds on each muscle.

References

Kristensen M, Hansen T (2004). Statistical analyses of repeated measures in physiological research: a tutorial. *Adv Physiol Educ.* 28(1-4):2-14.

locomotor *Locomotor activity rat data*

Description

Distance travelled by rats in the open field test for a drug and control condition.

Format

A data frame with 282 rows and 4 variables:

drug: Drug or control indicator.

animal: Rat identification number.

time: Time point of observation, in increments of 15 minutes.

dist: Distance travelled.

Details

Forty-seven rats were randomised to a control or drug condition. Rats were injected with either a drug or saline at time = 0. Locomotor activity in the open field was recorded for 90 minutes. The distance travelled in each 15 minute interval is reported.

VPA *Valproic acid (VPA) data set*

Description

A mouse split-plot experiment with two treatments.

Format

A data frame with 24 rows and 5 variables:

litter: Mice were from one of six litters.

sex: Male or Female.

group: VPA or saline (SAL).

drug: MPEP or saline (SAL).

activity: Locomotor activity.

Details

Six pregnant female mice were randomly assigned to receive an injection of valproic acid (n = 3) or saline (n = 3). The offspring of these mice (n = 24) were then randomly assigned to receive an injection of the glutamate receptor antagonist MPEP (n = 12) or saline (n = 12). There are two levels of randomisation: the pregnant females and their offspring.

This is a subset of the full data set from Mehta et al., which contained fourteen litters. The litters were selected to have an equal sample size across all conditions for illustrative purposes. The complete data can be found in the supplementary material of Lazic and Essioux (2013).

References

Mehta MV, Gandal MJ, Siegel SJ (2011). mGluR5-antagonist mediated reversal of elevated stereotyped, repetitive behaviors in the VPA model of autism. *PLoS ONE* 6(10):e26077.

Lazic SE, Essioux L (2013). Improving basic and translational science by accounting for litter-to-litter variation in animal models. *BMC Neuroscience* 14:37.

Index

assay.window, 2

block.covars, 2

cellcount, 3
CV, 4

festing, 4
fluoxetine, 5

glycogen, 5
goldstandard, 6

hypertension, 6

KH2004, 7

locomotor, 8

VPA, 8