

# Package ‘OmicsWorkshop’

August 9, 2016

**Type** Package

**Title** Basic statistics, survival analysis and permutation testing

**Version** 1.0.1

**Date** 2015-09-15

**Author** Deepak Nag Ayyala, Shili Lin, Department of Statistics, The Ohio State University, Columbus, Ohio, USA.

**Maintainer** Deepak Nag Ayyala <ayyala.1@osu.edu>

**Description** Some basic functions useful for the workshop.

**Depends** R (>= 3.0.0)

**License** LGPL-3

**LazyLoad** YES

**Imports** stats, coin, survival, parallel, survMisc

## R topics documented:

OmicsWorkshop-package . . . . .	2
Attributes1 . . . . .	2
Attributes2 . . . . .	3
barplot_with_errors . . . . .	3
confint_plots . . . . .	4
Geneset1 . . . . .	4
Geneset2 . . . . .	5
multiple_hists . . . . .	5
perm_survival . . . . .	6
perm_t . . . . .	6
SurvivalData . . . . .	7
<b>Index</b>	<b>8</b>

OmicsWorkshop-package *Omics Workshop*

---

### Description

This package contains some functions and data sets for the Omics workshop

### Details

The functions in the package are useful for creating plots and performing permutation tests to be used in the Omics Workshop.

### Author(s)

Deepak Nag Ayyala, Shili Lin.

---

Attributes1

*Attributes for the data Geneset1*

---

### Description

The data set consists of medical attributes for the Geneset1 data set.

### Usage

```
data("Attributes1")
```

### Format

A data frame with 200 observations on the following 3 variables.

**First.Pathologic.Diagnosis.Biospecimen.Acquisition.Method.Type** First pathologic diagnosis biospecimen acquisition method

**Overall.Survival.Status** Overall survival status

**Patient.s.Vital.Status** Patient vital status

### Source

[cbioportal.org/study.do?cancer\\_study\\_id=brca\\_tcga](https://cbioportal.org/study.do?cancer_study_id=brca_tcga)

### Examples

```
data(Attributes1)
```

---

Attributes2

*Attributes for the data Geneset2*


---

**Description**

The data set consists of medical attributes for the Geneset2 data set.

**Usage**

```
data("Attributes1")
```

**Format**

A data frame with 1003 observations on the following 3 variables.

**First.Pathologic.Diagnosis.Biospecimen.Acquisition.Method.Type** First pathologic diagnosis biospecimen acquisition method

**Overall.Survival.Status** Overall survival status

**Patient.s.Vital.Status** Patient vital status

**Source**

[cbioportal.org/study.do?cancer\\_study\\_id=brca\\_tcga](https://cbioportal.org/study.do?cancer_study_id=brca_tcga)

**Examples**

```
data(Attributes2)
```

---

barplot\_with\_errors

*Bar and error plot*


---

**Description**

Construct the barplots whose height represents the mean and error bars which represent the standard deviation.

**Usage**

```
barplot_with_errors(X, use.col = TRUE, colors = 0)
```

**Arguments**

X A numeric matrix

use.col Logical variable. If TRUE, the columns will be treated as variables. If FALSE, rows will be treated as variables

colors A vector of length equal to number of rows of X to add colors to the bars.

**Author(s)**

Deepak Nag Ayyala

**Examples**

```
## Not run
# barplot_with_errors(matrix(rnorm(500*10, ncol = 10)), use.col = TRUE, colors = 1:10)
```

---

confint_plots	<i>Simultaneous confidence interval plot</i>
---------------	--

---

**Description**

Given N variables and a condition, plot the confidence intervals for all the variables.

**Usage**

```
confint_plots(X, groups)
```

**Arguments**

X	A numeric matrix whose columns correspond to variables and rows correspond to samples
groups	A indicator vector (numeric or character) which gives group membership of the samples. Must have only two factors.

**Author(s)**

Deepak Nag Ayyala

**Examples**

```
## Not run
# confint_plots(matrix(rnorm(500*10, ncol = 10)), groups = rep(c(1,2), c(250, 250)))
```

---

Geneset1	<i>Data set to be used in the lecture.</i>
----------	--

---

**Description**

The data set normalized RNA-seq counts for 1000 genes

**Usage**

```
data("Geneset1")
```

**Format**

A data frame with 200 observations on 1000 genes

**Source**

[cbioportal.org/study.do?cancer\\_study\\_id=brca\\_tcga](https://cbioportal.org/study.do?cancer_study_id=brca_tcga)

**Examples**

```
data(Geneset1)
```

---

 Geneset2

*Data set to be used in the exercise session.*


---

**Description**

The data set normalized RNA-seq counts for 1000 genes

**Usage**

```
data("Geneset2")
```

**Format**

A data frame with 1003 observations on 1000 genes

**Source**

[cbioportal.org/study.do?cancer\\_study\\_id=brca\\_tcga](https://cbioportal.org/study.do?cancer_study_id=brca_tcga)

**Examples**

```
data(Geneset2)
```

---

multiple\_hists

*Plot multiple density histograms***Description**

Plot multiple density histograms overlapping within the range of the data.

**Usage**

```
multiple_hists(X, by.col = TRUE, colors = NULL, breaks = 10, ylim = c(0,1))
```

**Arguments**

X	A numeric matrix
by.col	Logical variable. If TRUE, columns will be used as the variables. If FALSE, rows will be used as variables.
colors	Vector of length equal to number of variables to be used for coloring the histograms
breaks	Number of breaks to be used when constructing the histogram
ylim	Y-axis limits to be used when plotting the histograms

**Author(s)**

Deepak Nag Ayyala

**Examples**

```
## NOT RUN
# multiple_hists(matrix(rnorm(50*2), ncol = 2), by.col = TRUE,
#   colors = c(2,4), breaks = 10, ylim = c(0,0.5))
```

---

perm\_survival                      *Permutation test for survival data*

---

**Description**

Performs a single replication of the permutation test for survival data

**Usage**

```
perm_survival(surv.object, group)
```

**Arguments**

surv.object      An object of type Surv obtained using the survival package

group            A vector whose length is equal to the number of samples which gives the group membership of the samples

**Value**

Log-rank test statistic obtained using a permutation of the group memberships

**Author(s)**

Deepak Nag Ayyala

**Examples**

```
## NOT RUN
#data(lung)
#perm_survival(surv.object = Surv(lung$time, lung$status),
#   group = lung$sex)
```

---

perm\_t                              *Permutation test statistic for the t-test*

---

**Description**

Performs a single replication of the permutation test using the two-sample t-test statistic.

**Usage**

```
perm_t(X, Y)
```

**Arguments**

X            A numeric vector  
Y            A numeric vector

**Value**

Two sample t-test statistic calculated under the assumption of unequal variances.

**Author(s)**

Deepak Nag Ayyala

**Examples**

```
perm_t(rnorm(20), rnorm(30))
```

---

SurvivalData	<i>Survival data recorded for the 1003 subjects participated in the TCGA provisional study</i>
--------------	--

---

**Description**

The data set normalized RNA-seq counts for 1000 genes

**Format**

A data frame with 1003 observations on the following 6 variables.

**PATIENT.ID** Patient ID

**Overall.Survival..Months.** Overall survival months

**Overall.Survival.Status** Overall survival status

**Patient.s.Vital.Status** Patient vital status

**Person.Gender** Gender

**Tissue.Source.Site** Tissue source site

**Source**

[cbioportal.org/study.do?cancer\\_study\\_id=brca\\_tcga](http://cbioportal.org/study.do?cancer_study_id=brca_tcga)

**Examples**

```
data(SurvivalData)
```

# Index

- \*Topic **\textasciitildekw2**
  - perm\_survival, 6
- \*Topic **barplot\_with\_\_errors**
  - barplot\_with\_errors, 3
- \*Topic **confint\_plots**
  - confint\_plots, 4
- \*Topic **datasets**
  - Attributes1, 2
  - Attributes2, 3
  - Geneset1, 4
  - Geneset2, 5
  - SurvivalData, 7
- \*Topic **multiple\_hists**
  - multiple\_hists, 5
- \*Topic **package**
  - OmicWorkshop-package, 2
- \*Topic **perm\_survival**
  - perm\_survival, 6
- \*Topic **perm\_t**
  - perm\_t, 6

Attributes1, 2  
Attributes2, 3

barplot\_with\_errors, 3

confint\_plots, 4

Geneset1, 4  
Geneset2, 5

multiple\_hists, 5

OmicWorkshop (OmicWorkshop-package), 2  
OmicWorkshop-package, 2

perm\_survival, 6  
perm\_t, 6

SurvivalData, 7