

Author(s)

M.Ruschhaupt

References

Choi et al, Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*, 2003, i84-i90.

dstar

Tools for Meta-analysis of gene expression data.

Description

A small number of meta-analysis functions for comparing two gene expression experiments are provided.

Usage

```
dstar(d, n)
getdF(data, categ)
sigmad(d, ng1, ng2)
```

Arguments

d	A vector of t-statistics, i.e. the output of <code>getdF</code> .
n	The number of t-statistics.
data	The data used to compute t-statistics, either a <i>matrix</i> or an <code>ExpressionSet</code> .
categ	A vector of 0's and 1's indicating group membership.
ng1	The number of samples in group 1.
ng2	The number of samples in group 2.

Details

The functions `getdF` compute t-test statistics for the input data and group membership (note that group membership must be indicated by a vector of 0's and 1's).

The function `dstar` computes an unbiased estimate of the t-test. The function `sigmad` computes the variance estimate of `dstar`.

Value

The different functions have different return values, but generally they are vectors of the requested quantities.

Author(s)

L. Lusa, R. Gray and R. Gentleman

References

Choi et al, Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*, 2003, i84-i90.

Examples

```
x = matrix(rnorm(1000), ncol=10)
ds = getdF(x, rep(c(0,1), c(5,5)))
dst = dstar(ds, ncol(x))
sgd = sigmad(ds, 5, 5)
```

`f.Q`*Compute Cochran's Q statistic*

Description

Compute Cochran's Q statistic for testing whether the a fixed effects or a random effects model will be appropriate.

Usage

```
f.Q(dadj, varadj)
```

Arguments

<code>dadj</code>	A matrix, each row is a gene, each column a study, of the estimated t-statistics.
<code>varadj</code>	A matrix, each row is a gene, each column a study, of the estimated, adjusted variances of the t-statistics.

Details

A straightforward computation of Cochran's Q statistic. If the null hypothesis that the data are well modeled by a fixed effects design is true then the estimate Q values will have approximately a chi-squared distribution with degrees of freedom equal to the number of studies minus one.

Value

A vector of length equal to the number of rows of `dadj` with the Q statistics.

Author(s)

L. Lusa and R. Gentleman

References

Choi et al, Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*, 2003, i84-i90.

See Also

[dstar](#), [sigmad](#)

Examples

```
##none now, this requires a pretty elaborate example
```

Nevins	<i>Intensity data for 46 Affymetrix slides with tissue samples of breast tumors</i>
--------	---

Description

Intensity data for 46 Affymetrix hu6800 slides with tissue samples of breast tumors. See vignette Nevins.pdf in the /scripts directory for details of the processing.

Usage

```
data(Nevins)
```

Format

Nevins is an ExpressionSet containing the data from 46 Affymetrix chips.

Source

```
http://data.cgt.duke.edu/west.php
```

References

Predicting the clinical status of human breast cancer by using gene expression profiles, West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA Jr, Marks JR, and Nevins JR. Proc Natl Acad Sci U S A 98(20):11462-7 (2001)

Examples

```
data(Nevins)
Nevins
```

tau2.DL	<i>estimating my and tau in a REM</i>
---------	---------------------------------------

Description

tau2.DL is an estimation of tau in a random effects model (REM) using Cochran's Q statistic.

Usage

```
tau2.DL(Q, num.studies, my.weights)
mu.tau2(my.d, my.vars.new)
var.tau2(my.vars.new)
```

Arguments

<code>Q</code>	A vector of Cochran's Q statistics.
<code>num.studies</code>	The number of studies used for the meta-analysis.
<code>my.weights</code>	A matrix with one column for each experiment containing the variances of the effects that should be combined.
<code>my.d</code>	A matrix, with one column for each experiment, containing the effects that should be combined.
<code>my.vars.new</code>	A matrix, with one column for each experiment, containing the variances of the effects that should be combined.

Author(s)

L. Lusa and R. Gentleman

References

Choi et al, Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*, 2003, i84-i90.

See Also

[dstar](#), [sigmad](#)

Examples

```
# please have a look at the vignette
```

zScores

Tools for Meta-analysis of gene expression data.

Description

A small number of meta-analysis functions for computing zScores for FEM and REM and computing FDR.

Usage

```
zScores(esets, classes, useREM=TRUE, CombineExp=1:length(esets))
zScorePermuted(esets, classes, useREM=TRUE, CombineExp=1:length(esets))
zScoreFDR(esets, classes, useREM=TRUE, nperm=1000, CombineExp=1:length(esets))
multExpFDR(theScores, thePermScores, type="pos")
```

Arguments

<code>esets</code>	A list of ExpressionSets, one expression set per experiment. All experiments must have the same variables(genes).
<code>classes</code>	A list of class memberships, one per experiment. Each list can only contain 2 levels.
<code>useREM</code>	A logical value indicating whether or not to use a REM, TRUE, or a FEM, FALSE, for combining the z scores.

<code>theScores</code>	A vector of scores (e.g. t-statistics or z scores)
<code>thePermScores</code>	A vector of permuted scores (e.g. t-statistics or z scores)
<code>type</code>	"pos", "neg" or "two.sided"
<code>nperm</code>	number of permutations to calculate the FDR
<code>CombineExp</code>	vector of integer- which experiments should be combined-default:all experiments

Details

The function `zScores` implements the approach of Choi et al. for for a set of `ExpressionSets`. The function `zScorePermuted` applies `zScore` to a single permutation of the class labels. The function `zScoreFDR` computes a FDR for each gene, both for each single experiment and for the combined experiment. The FDR is calculated as described in Choi et al. Up to now ties in the zscores are not taken into account in the calculation. The function might produce incorrect results in that case. The function also computes `zScores`, both for the combines experiment and for each single experiment.

Value

A matrix with one row for each probe(set) and the following columns:

<code>zSco_Ex_</code>	For each single experiment the standardized mean difference, <code>Effect_Ex_</code> , divided by the estimated standard deviation, the square root of the <code>EffectVar_Ex_</code> column.
<code>MUvals</code>	The combined standardized mean difference (using a FEM or REM)
<code>MUsds</code>	The standard deviation of the <code>MUvals</code> .
<code>zSco</code>	The z statistic - the <code>MUvals</code> divided by their standard deviations, <code>MUsds</code> .
<code>Qvals</code>	Cochran's Q statistic for each gene.
<code>df</code>	The degree of freedom for the Chi-square distribution. This is equal to the number of combined experiments minus one.
<code>Qpvalues</code>	The probability that a Chi-square random variable, with <code>df</code> degrees of freedom) has a higher value than the value from the Q statistic.
<code>Chisq</code>	The probability that a Chi-square random variate (with 1 degree of freedom) has a higher value than the value of $zSco^2$.
<code>Effect_Ex_</code>	The standardized mean difference for each single experiment.
<code>EffectVar_Ex_</code>	The variance of the standardized mean difference for each single experiment.

Note that the three column names that end in an underscore are replicated, once for each experiment that is being analyzed.

Author(s)

M. Ruschhaupt

References

Choi et al, Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*, 2003, i84-i90.

Examples

```
data(Nevins)

##Splitting
thestatus <- Nevins$ER.status
group1 <- which(thestatus=="pos")
group2 <- which(thestatus=="neg")
rrr <- c(sample(group1, floor(length(group1)/2)),
         sample(group2, ceiling(length(group2)/2)))
Split1 <- Nevins[,rrr]
Split2 <- Nevins[,-rrr]

#obtain classes
Split1.ER <- as.numeric(Split1$ER.status) - 1
Split2.ER <- as.numeric(Split2$ER.status) - 1

esets <- list(Split1, Split2)
classes <- list(Split1.ER, Split2.ER)
theScores <- zScores(esets, classes, useREM=FALSE)
theScores[1:2,]
```

Index

*Topic **datasets**

Nevins, 4

*Topic **htest**

dstar, 2

f.Q, 3

tau2.DL, 4

*Topic **manip**

zScores, 5

CountPlot, 1

dstar, 2, 3, 5

f.Q, 3

getdF(dstar), 2

*getdF(ExpressionSet, numeric-method
(dstar)*, 2

*getdF(matrix, numeric-method
(dstar)*, 2

IDRplot(CountPlot), 1

mu.tau2(tau2.DL), 4

multExpFDR(zScores), 5

Nevins, 4

sigmad, 3, 5

sigmad(dstar), 2

tau2.DL, 4

var.tau2(tau2.DL), 4

zScoreFDR(zScores), 5

zScorePermuted(zScores), 5

zScores, 5