## Package 'r2redux'

May 18, 2022

Type Package
Title R2 statistic
Version 1.0.3
Maintainer Moksedul Momiin <momin@cvasu.ac.bd></momin@cvasu.ac.bd>
<b>Description</b> R2 statistic for signficance test.
License `use_gpl3_license()`
Encoding UTF-8
<b>Roxygen</b> list(markdown = TRUE)
RoxygenNote 7.1.2.9000
R topics documented:
cc_trf
olkin_beta1_2
olkin_beta_inf
r2_diff
r2_enrich
r2_enrich_beta
r2_var
Index 1
cc_trf cc_trf function

## Description

This function transforms the predictive ability and its standard error between the observed scale and liability scale cc\_trf(R2,se, K,P), where R2 is the R squared value of the model for either observed scale or liability scale, se is the standard error for either observed scale or liability scale, K and P is the prevalence of cases in the population.

## Usage

```
cc\_trf(R2, se, K, P)
```

2 olkin\_beta1\_2

#### References

Lee, S. H., Goddard, M. E., Wray, N. R., and Visscher, P. M. A better coefficient of determination for genetic profile analysis. Genetic epidemiology, (2012). 36(3): p. 214-224.

#### **Examples**

```
To get the transformed predictive ability output=cc_trf(0.06, 0.002, 0.05, 0.05)

output$R21 (transformed R2 to liability scale, assuming 0.06 in observed scale)
0.2679337

output$sel (transformed se to liability scale, assuming 0.002 is observed scale)
0.008931123

output$R20 (transformed R2 to observed scale, assuming 0.06 in liability scale)
0.01343616

output$se0 (transformed to in observed scale, assuming 0.002 is liability scale)
0.000447872
```

olkin\_beta1\_2 *olkin\_beta1\_2 function* 

**Description** 

This function derives Information matrix for beta  $1^2$  and beta  $2^2$  where beta 1 and 2 are regression coefficients from a multiple regression model, i.e. y = x1 beta 1 + x2 beta 2 + e, where y, x1 and x2 are column-standardised, (i.e. in the context of correlation coefficients, see Olkin and Finn 1995).

#### Usage

```
olkin_beta1_2(omat, nv)
```

#### Arguments

omat 3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e.

omat=cor(dat) where dat is N by 3 matrix having variables in the order of cbind

(y,x1,x2)

nv sample size

#### References

Olkin, I. and J.D. Finn, Correlations redux. Psychological Bulletin, 1995. 118(1): p. 155.

olkin\_beta\_inf 3

#### **Examples**

To get information (variance-covariance) matrix of beta1 $^2$  and beta2 $^2$  where beta1 and 2 are regression coefficients from a multiple regression model.

```
dat=read.table("test_ukbb_thresholds_scaled") (see example files)
omat=cor(dat)[1:3,1:3]
omat
1.0000000 0.1958636 0.1970060
0.1958636 1.0000000 0.9981003
0.1970060 0.9981003 1.0000000
nv=length(dat$V1)
output=olkin_beta1_2(omat,nv)
output
output$info (2x2 information (variance-covariance) matrix)
0.04146276 0.08158261
0.08158261 0.16111124
output$var1 (variance of beta1^2)
0.04146276
output$var2 (variance of beta2^2)
0.1611112
output$var1_2 (variance of difference between beta1^2 and beta2^2)
0.03940878
```

olkin\_beta\_inf

olkin\_beta\_inf function

#### Description

This function derives Information matrix for beta1 and beta2 where beta1 and 2 are regression coefficients from a multiple regression model, i.e.  $y = x1 \cdot beta1 + x2 \cdot beta2 + e$ , where y, x1 and x2 are column-standardised (see Olkin and Finn 1995).

## Usage

```
olkin_beta_inf(omat, nv)
```

## Arguments

omat 3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e.

omat=cor(dat) where dat is N by 3 matrix having variables in the order of cbind

(y,x1,x2)

nv sample size

#### References

Olkin, I. and J.D. Finn, Correlations redux. Psychological Bulletin, 1995. 118(1): p. 155.

4 r2\_diff

## **Examples**

To get information (variance-covariance) matrix of beta1 $^2$  and beta2 $^2$  where beta1 and 2 are regression coefficients from a multiple regression model.

```
dat=read.table("test_ukbb_thresholds_scaled") (see example files)
omat=cor(dat)[1:3,1:3]
omat
1.0000000 0.1958636 0.1970060
0.1958636 1.0000000 0.9981003
0.1970060 0.9981003 1.0000000
nv=length(dat$V1)
output=olkin_beta_inf(omat,nv)
output
\verb"output$info (2x2 information (variance-covariance) matrix")
0.2531406 -0.2526212
-0.2526212 0.2530269
output$var1 (variance of beta1^2)
0.2531406
output$var2 (variance of beta2^2)
0.2530269
output$var1_2 (variance of difference between beta1^2 and beta2^2)
1.01141
```

r2\_diff

r2\_diff function

#### Description

This function estimates  $var(R2(y\sim x[,v1]) - R2(y\sim x[,v2]))$  where R2 is the R squared value of the model, y is N by 1 matrix having the dependent variable, and x is N by M matrix having M explanatory variables. v1 or v2 indicates the ith column in the x matrix (v1 or v2 can be multiple values between 1 - M, see Arguments below)

#### Usage

```
r2_diff(dat, v1, v2, nv)
```

## **Arguments**

dat	N by $(M+1)$ matrix having variables in the order of $cbind(y,x)$
v1	This can be set as $v1=c(1)$ or $v1=c(1,2)$
v2	This can be set as $v2=c(2)$ , $v2=c(3)$ , $v2=c(1,3)$ or $v2=c(3,4)$
nv	sample size

 $r2_diff$  5

```
To get the test statistics for the difference between R2(y\sim x[,v1]) and
R2(y \sim x[,v2]). (here we define R2_1=R2(y \sim x[,v1])) and R2_2=R2(y \sim x[,v2])))
dat=read.table("test_ukbb_thresholds_scaled") (see example files)
nv=length(dat$V1)
v1=c(1)
v2=c(2)
output=r2_diff(dat,v1,v2,nv)
output
r2redux output
output$var1 (variance of R2_1)
0.0001437583
output$var2 (variance of R2_2)
0.0001452828
output$var_diff (variance of difference between R2_1 and R2_2)
5.678517e-07
output$r2_based_p (p-value for significant difference between R2_1 and R2_2)
0.5514562
output$mean_diff (differences between R2_1 and R2_2)
-0.0004488044
output$upper_diff (upper limit of 95% CI for the difference)
0.001028172
output$lower_diff (lower limit of 95% CI for the difference)
-0.001925781
To get the test statistics for the difference between R2(y \sim x[,v1]+x[,v2]) and
 \mbox{R2} \ (\mbox{$y$}\mbox{$x$} \mbox{$(y$}\mbox{$x$}\mbox{$(,v2]$}) \mbox{. (here } \mbox{$R2$}\mbox{$1$}\mbox{$=$R2} \mbox{$(y$}\mbox{$x$}\mbox{$(,v1]$}\mbox{$+$x$}\mbox{$(,v2]$}) \mbox{ and } \mbox{$R2$}\mbox{$=$R2} \mbox{$(y$}\mbox{$x$}\mbox{$(,v1]$}\mbox{$)$}) 
dat=read.table("test_ukbb_thresholds_scaled") (see example files)
nv=length(dat$V1)
v1=c(1,2)
v2=c(1)
output=r2_diff(dat,v1,v2,nv)
output
r2redux output
output$var1 (variance of R2_1)
0.0001475195
output$var2 (variance of R2_2)
0.0001437583
output$var_diff (variance of difference between R2_1 and R2_2)
2.321425e-06
```

6 r2\_enrich

```
output$r2_based_p (p-value for significant difference between R2_1 and R2_2) 0.4369177

output$mean_diff (differences between R2_1 and R2_2) 0.0006042383

output$upper_diff (upper limit of 95% CI for the difference) 0.004887989

output$lower_diff (lower limit of 95% CI for the difference) -0.0005574975
```

r2\_enrich

r2\_enrich

#### **Description**

This function estimates var(t1/(t1+t2)) where  $t1 = R2(y\sim x[,v1]+x[,v2]) - R2(y\sim x[,v1])$  and  $t2 = R2(y\sim x[,v1]+x[,v2]) - R2(y\sim x[,v2]))$  where R2 is the R squared value of the model, y is N by 1 matrix having the dependent variable, and x is N by M matrix having M explanatory variables. v1 or v2 indicates the ith column in the x matrix (v1 or v2 should be a single interger between 1 - M, see Arguments below)

## Usage

```
r2_enrich(dat, v1, v2, nv, exp1)
```

#### **Arguments**

dat	N by $(M+1)$ matrix having variables in the order of $cbind(y,x)$
nv	sample size
exp1	The expectation of the ratio (e.g. # SNPs for the genomic region of interest / total # SNPs in genomic partitioning)
v1/v2	These can be set as $v1=1$ and $v2=2$ , $v1=2$ and $v2=1$ , $v1=3$ and $v2=2$ , or any combination as long as the value is between $1-M$

```
To get test statistics for the ratio, i.e. t1/(t1+t2). t1 = R2(y\sim x[,v1]+x[,v2]) - R2(y\sim x[,v1])) and t2 = R2(y\sim x[,v1]+x[,v2]) - R2(y\sim x[,v2])) (here we define R2_1=R2(y\sim x[,v1]), R2_2=R2(y\sim x[,v2]), R2_1=R2(y\sim x[,v1]+x[,v2]) dat=read.table("test_ukbb_enrichment_choles") (see example file) nv=length(dat\$V1) v1=c(1) v2=c(2) expected_ratio=0.04 (# SNPs for the regulatory/total # SNPs) output=r2_enrich(dat,v1,v2,nv,expected_ratio) output
```

r2\_enrich\_beta 7

```
r2redux output
output$var1 (variance of R2_1)
8.758455e-05
output$var2 (variance of R2_2)
7.36385e-05
output$var12 (variance of R2_12)
0.000102236
outputvar_diff1_2 (var of difference of R2(vx[,v1]) - R2(vx[,v2])))
6.074567e-05
outputvar_diff12_1 (var of difference of R2(y\sim x[,v1]+x[,v2]) - R2(y\sim x[,v1])))
1.184853e-05
\texttt{output\$var\_diff12\_2} \text{ (var of difference of R2}(y \sim x \texttt{[,v1]} + x \texttt{[,v2]}) \text{ - R2}(y \sim x \texttt{[,v2]})))
2.650564e-05
output\mbox{mean\_diff}12_1 (difference of R2(y \sim x[,v1] + x[,v2]) - R2(y \sim x[,v1])))
0.003048595
0.006845484
output$ratio (ratio = t1/(t1+t2))
0.6918768
output$ratio_var (variance of ratio, var(t1/(t1+t2))
0.1324076
output $enrich_p (p-value for testing the ratio significantly different
from the expectation (exp1))
0.07321821
output$upper_ratio (upper limit of 95% CI for the ratio)
1.405079
output$lower_ratio (lower limit of 95% CI for the ratio)
-0.02132515
```

r2\_enrich\_beta r2\_enrich\_beta

## Description

This function estimates var(t1/(t1+t2)) where  $t1 = beta1^2$  and  $t2 = beta2^2$ , and beta1 and 2 are regression coefficients from a multiple regression model, i.e.  $y = x1 \cdot beta1 + x2 \cdot beta2 + e$ , where y, x1 and x2 are column-standardised (see Olkin and Finn 1995). y is y by 1 matrix having the dependent variable, and y is y by 1 matrix having the ith explanatory variables. y is y by 1 matrix having the jth explanatory variables. y and y indicates the ith and jth column in the data (y1 or y2 should be a single interger between 1 - y4, see Arguments below). Note that y5 have y6 and y7 and y8 arguments below). Note that y8 have y9 and y9 and y9 and y9 and y9 have y9 and y9 have y

8 r2\_enrich\_beta

#### Usage

```
r2_enrich_beta(dat, v1, v2, nv, exp1)
```

#### **Arguments**

dat N by (M+1) matrix having variables in the order of cbind(y,x) nv sample size exp1 The expectation of the ratio (e.g. ratio of # SNPs in genomic partitioning) v1/v2 These can be set as v1=1 and v2=2, v1=2 and v2=1, v1=3 and v2=2, or any combination as long as the value is between 1 - M

#### References

Olkin, I. and J.D. Finn, Correlations redux. Psychological Bulletin, 1995. 118(1): p. 155.

```
To get the test statistic for the ratio which is significantly
different from the expectation.
var(t1/(t1+t2)), where t1 = beta1^2 and t2 = beta2^2.
betal and beta2 are regression coefficients from a multiple regression model,
i.e. y = x1 \cdot beta1 + x2 \cdot beta2 + e, where y, x1 and x2 are column-standardised
dat=read.table("test_ukbb_enrichment_choles") (see example file)
nv=length(dat$V1)
v1=c(1)
v2=c(2)
expected_ratio=0.04
output=r2_enrich_beta(dat,v1,v2,nv,expected_ratio)
output
r2redux output
output$var1 (variance of t1)
7.072931e-05
output$var2 (variance of t2)
3.161929e-05
output$var1_2 (difference between t1 and t2)
0.000162113
output$beta1_sq (t1)
0.01118301
output$beta2_sq (t2)
0.004980285
output$cov (covariance between t1 and t2)
-2.988221e-05
output$ratio (ratio = t1/(t1+t2))
0.6918768
output$ratio_var (variance of ratio, var(t1/(t1+t2))
```

r2\_var 9

```
0.1324076
output$enrich_p (p-value for testing the ratio significantly different from
the expectation (expl))
0.07321821
output$upper_ratio (upper limit of 95% CI for the ratio)
1.405079
output$lower_ratio (lower limit of 95% CI for the ratio)
-0.02132515
output$enrich_p2 (p-value for testing the difference between t1/exp and t2/exp)
0.1997805
output$mean_diff (difference between t1/exp and t2/exp)
0.2743874
output$var_diff (variance of difference, t1/exp - t2/exp)
0.04579649
output$upper_diff (upper limit of 95% CI for the mean difference)
0.6938296
output$lower_diff (lower limit of 95% CI for the mean difference)
-0.1450549
```

r2\_var

r2\_var function

#### **Description**

This function estimates  $var(R2(y\sim x[,v1]))$  where R2 is the R squared value of the model, where R2 is the R squared value of the model, y is N by 1 matrix having the dependent variable, and x is N by M matrix having M explanatory variables. v1 indicates the ith column in the x matrix (v1 can be multiple values between 1 - M, see Arguments below)

#### Usage

```
r2_var(dat, v1, nv)
```

## Arguments

dat	N by $(M+1)$ matrix having variables in the order of $cbind(y,x)$
v1	This can be set as $v1=c(1)$ , $v1=c(1,2)$ or possibly with more values
nv	sample size

```
To get the test statistics for R2(y\sim x[,v1]) dat=read.table("test_ukbb_thresholds_scaled") (see example file) nv=length(dat$V1)
```

10 r2\_var

```
v1=c(1)
output=r2_var(dat, v1, nv)
r2redux output
output$var (variance of R2)
0.0001437583
output$r2_based_p (P-value under the null hypothesis, i.e. R2=0)
1.213645e-10
output$mean_r2 (R2)
0.03836254
output$upper_r2 (upper limit of 95% CI for R2)
0.06435214
output$lower_r2 (lower limit of 95% CI for R2)
0.01763347
To get the test statistic for R2(y \sim x[,v1]+x[,v2]+x[,v3])
dat=read.table("test_ukbb_thresholds_scaled") (see example file)
nv=length(dat$V1)
v1=c(1,2,3)
outout=r2_var(dat,v1,nv)
output
r2redux output
output$var (variance of R2)
0.0001499374
output$r2_based_p (R2 based P-value)
7.461267e-11
output$mean_r2 (R2)
0.03917668
output$upper_r2 (upper limit of 95% CI for R2)
0.06538839
output$lower_r2 (lower limit of 95% CI for R2)
0.01821657
```

# Index

*Topic <b>R2</b>	cc_trf,1
cc_trf,1	*Topic <b>of</b>
r2_diff,4	cc_trf,1
r2_enrich,6	olkin_beta1_2,2
r2_var,9	olkin_beta_inf,3
*Topic <b>Transformation</b>	$r2_{enrich, 6}$
cc_trf,1	r2_enrich_beta,7
*Topic <b>and</b>	*Topic <b>ratio</b>
cc_trf,1	r2_enrich, 6
*Topic <b>a</b>	r2_enrich_beta,7
r2_enrich_beta,7	*Topic <b>regression</b>
*Topic <b>beta^2</b>	r2_enrich_beta,7
r2_enrich_beta,7	*Topic <b>scale</b>
*Topic <b>between</b>	cc_trf, 1
cc_trf,1	*Topic <b>the</b>
r2_enrich,6	olkin_beta1_2, 2
r2_enrich_beta,7	olkin_beta_inf,3 *Topic variance
*Topic <b>context</b>	*Topic <b>variance</b> r2_diff, 4
olkin_beta1_2,2	r2_enrich, <b>6</b>
olkin_beta_inf, 3	r2_enrich_beta,7
*Topic correlation	r2_var, 9
olkin_beta1_2,2	12_var,
olkin_beta_inf,3	cc_trf,1
*Topic <b>from</b>	
r2_enrich_beta,7	olkin_beta1_2,2
*Topic information	olkin_beta_inf,3
olkin_beta1_2, 2	r2 diff 1
olkin_beta_inf,3	r2_diff, 4 r2_enrich, 6
r2_diff,4	r2_enrich_beta,7
r2_var, 9 *Topic <b>in</b>	r2_var, 9
olkin_beta1_2,2	12
olkin_beta_inf,3	
*Topic liability	
cc_trf, 1	
*Topic <b>matrix</b>	
olkin_beta1_2,2	
olkin_beta_inf,3	
r2_diff,4	
r2 var, 9	
*Topic <b>multiple</b>	
r2_enrich_beta, 7	
*Topic <b>observed</b>	