

# Package ‘R2ROC’

July 6, 2023

**Title** AUC Statistics

**Version** 1.0.0

**Description** AUC statistic for significance test. Variance and covariance of AUC values used to assess the 95% CI and p-value of the AUC difference for both nested and non-nested model.

**License** GPL (>=3)

**URL** <https://github.com/mommy003/R2ROC>

**Encoding** UTF-8

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.1.2

**Depends** R (>= 2.10)

**LazyData** true

## R topics documented:

auc_diff . . . . .	2
auc_var . . . . .	4
dat1 . . . . .	5
dat2 . . . . .	6
olkin_auc1 . . . . .	6
olkin_auc12 . . . . .	7
olkin_auc12_1 . . . . .	7
olkin_auc12_13 . . . . .	8
olkin_auc12_3 . . . . .	8
olkin_auc12_34 . . . . .	9
olkin_auc1_2 . . . . .	9
<b>Index</b>	<b>10</b>

auc\_diff

*auc\_diff function***Description**

This function estimates  $\text{var}(\text{AUC}(y \sim x[,v1]) - \text{AUC}(y \sim x[,v2]))$  where AUC is the Area Under ROC curve 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**

```
auc_diff(dat, v1, v2, nv, kv)
```

**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
kv	Population prevalence

**Value**

This function will estimate significant difference between two PRS (either dependent or independent and joint or single). To get the test statistics for the difference between  $\text{AUC}(y \sim x[,v1])$  and  $\text{AUC}(y \sim x[,v2])$  (here we define  $\text{AUC1} = \text{AUC}(y \sim x[,v1])$  and  $\text{AUC2} = \text{AUC}(y \sim x[,v2])$ ). The outputs are listed as follows.

mean_diff	AUC differences between AUC1 and AUC2
var	Variances of AUC differences
upper_diff	Upper value of the differences
lower_diff	Upper value of the differences
p	Two tailed P-value for significant difference between AUC1 and AUC2
p_one_tail	One tailed P-value for significant difference
heller_p	P-value based on Heller's test for significant difference
heller_upper_diff	Upper limit of 95% CI for the difference basedon Heller's test
heller_lower_diff	Lower limit of 95% CI for the difference basedon Heller's test

**Examples**

```
#To get the test statistics for the difference between AUC(y=x[,v1])
#and AUC(y=x[,v2])
dat=dat1 #(this example embedded within the package)
nv=length(dat$V1)
kv=sum(dat$V1)/length(dat$V1)# pop. prevalence
#kv is 0.10 for the pre-adjusted phenotype
```

```
#please see dat2 embedded within the package)
v1=c(1)
v2=c(2)
output=auc_diff(dat,v1,v2,nv,kv)

#R2ROC output
#output$mean_diff (mean difference of AUC1 and AUC2)
#0.1756046

#output$var (variance of AUC difference)
#9.274356e-05

#output$upper_diff (upper limit of 95% CI for difference)
#0.1944801

#output$lower_diff (lower limit of 95% CI for difference)
#0.1567292

#output$p (two-tailed P-value for the differences is
#significantly different from zero)
#2.747031e-74

#output$p_one_tail (one-tailed P-value for the differences
#is significantly different from zero)
#1.373515e-74

#To get the test statistics for the difference between
#AUC(y=x[,v1]+x[,v2]) and AUC(y=x[,v2])
dat=dat1 #(this example embedded within the package)
nv=length(dat$V1)
kv=sum(dat$V1)/length(dat$V1)# pop. prevalence
#kv is 0.10 for the pre-adjusted phenotype
#please see dat2 embedded within the package)
v1=c(1,2)
v2=c(2)
output=auc_diff(dat,v1,v2,nv,kv)

#R2ROC output
#output$mean_diff (mean difference of AUC1 and AUC2)
#0.1793682

#output$var (variance of AUC difference)
#0.0001190366

#output$upper_diff (upper limit of 95% CI for difference)
#0.2007526

#output$lower_diff (lower limit of 95% CI for difference)
#0.1579839

#output$p (two-tailed P-value for the differences is
#significantly different from zero)
#9.87014e-61

#output$p_one_tail (one-tailed P-value for the differences
#is significantly different from zero)
```

```
#4.93507e-61

#output$heller_p (two-tailed P-value based on Hellers test
#for the differences is significantly different from zero)
#4.2085e-237

#output$heller_upper_diff (upper limit of 95% CI for
#difference based on Hellers test)
#0.2013899

#output$heller_lower_diff (lower limit of 95% CI for
#difference based on Hellers test)
#0.1586212
```

---

auc_var	<i>auc_var function</i>
---------	-------------------------

---

## Description

This function estimates  $\text{var}(\text{AUC}(y \sim x[,v1]))$  where AUC is the Area Under ROC curve 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

```
auc_var(dat, v1, nv, kv)
```

## 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
kv	Population prevalence

## Value

This function will test the null hypothesis for AUC. To get the test statistics for  $\text{AUC}(y \sim x[,v1])$ . The outputs are listed as follows.

auc	AUC
var	Variance of AUC
upper_auc	Upper limit of 95% CI for AUC
lower_auc	Lower limit of 95% CI for AUC

## Examples

```
#To get the AUC for AUC(y=x[,v1])

dat=dat1 #(this example embedded within the package)
nv=length(dat$V1)
kv=sum(dat$V1)/length(dat$V1) # pop. prevalence
#kv is 0.10 for the pre-adjusted phenotype
#please see dat2 embedded within the package)
v1=c(1)
output=auc_var(dat,v1,nv,kv)

#R2ROC output
#output$auc (AUC)
#0.7390354

#output$var (variance of AUC)
#7.193337e-05

#output$upper_auc (upper limit of 95% CI for AUC)
#0.7556589

#output$lower_auc (lower limit of 95% CI for AUC)
#0.7224119
```

---

dat1

*Raw phenotypes and 2 sets of discovery PGSs*


---

## Description

A dataset containing raw phenotypes and multiple PGSs estimated two independent discovery population

## Usage

```
dat1
```

## Format

A data frame with 10000 rows and 3 variables:

**V1** Phenotype, raw case-control data

**V2** PGS1, for discovery population 1

**V3** PGS2, for discovery population 2

---

dat2	<i>Pre-adjusted phenotypes and 2 sets of discovery PGSs</i>
------	-------------------------------------------------------------

---

### Description

A dataset containing preadjusted phenotypes and multiple PGSs estimated two independent discovery population

### Usage

```
dat2
```

### Format

A data frame with 10000 rows and 3 variables:

**V1** Phenotype, preadjusted case-control data

**V2** PGS1, for discovery population 1

**V3** PGS2, for discovery population 2

---

olkin_auc1	<i>olkin_auc1 function</i>
------------	----------------------------

---

### Description

olkin\_auc1 function

### Usage

```
olkin_auc1(omat, nv, kv)
```

### 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
kv	Population prevalence

### Value

This function will be used as source code

---

olkin_auc12	<i>olkin_auc12 function</i>
-------------	-----------------------------

---

**Description**

olkin\_auc12 function

**Usage**

```
olkin_auc12(omat, nv, kv)
```

**Arguments**

omat	3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e. $\text{omat} = \text{cor}(\text{dat})$ where dat is N by 3 matrix having variables in the order of cbind(y,x1,x2)
nv	Sample size
kv	Population prevalence

**Value**

This function will be used as source code

---

olkin_auc12_1	<i>olkin_auc12_1 function</i>
---------------	-------------------------------

---

**Description**

olkin\_auc12\_1 function

**Usage**

```
olkin_auc12_1(omat, nv, kv)
```

**Arguments**

omat	3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e. $\text{omat} = \text{cor}(\text{dat})$ where dat is N by 3 matrix having variables in the order of cbind(y,x1,x2)
nv	Sample size
kv	Population prevalence

**Value**

This function will be used as source code

---

olkin\_auc12\_13      *olkin\_auc12\_13 function*


---

**Description**

olkin\_auc12\_13 function

**Usage**

```
olkin_auc12_13(omat, nv, kv)
```

**Arguments**

omat	3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e. $\text{omat} = \text{cor}(\text{dat})$ where dat is N by 3 matrix having variables in the order of cbind(y,x1,x2)
nv	Sample size
kv	Population prevalence

**Value**

This function will be used as source code

---

olkin\_auc12\_3      *olkin\_auc12\_3 function*


---

**Description**

olkin\_auc12\_3 function

**Usage**

```
olkin_auc12_3(omat, nv, kv)
```

**Arguments**

omat	3 by 3 matrix having the correlation coefficients between y, x1 and x2, i.e. $\text{omat} = \text{cor}(\text{dat})$ where dat is N by 3 matrix having variables in the order of cbind(y,x1,x2)
nv	Sample size
kv	Population prevalence

**Value**

This function will be used as source code



---

olkin\_auc12\_34      *olkin\_auc12\_34 function*


---

**Description**

olkin\_auc12\_34 function

**Usage**

olkin\_auc12\_34(omat, nv, kv)

**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
kv	Population prevalence

**Value**

This function will be used as source code

---

olkin\_auc1\_2      *olkin\_auc1\_2 function*


---

**Description**

olkin\_auc1\_2 function

**Usage**

olkin\_auc1\_2(omat, nv, kv)

**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
kv	Population prevalence

**Value**

This function will be used as source code

# Index

## \*Topic **R2**

auc\_diff, [2](#)

auc\_var, [4](#)

## \*Topic **datasets**

dat1, [5](#)

dat2, [6](#)

## \*Topic **information**

auc\_diff, [2](#)

auc\_var, [4](#)

## \*Topic **matrix**

auc\_diff, [2](#)

auc\_var, [4](#)

## \*Topic **source**

olkin\_auc1, [6](#)

olkin\_auc12, [7](#)

olkin\_auc12\_1, [7](#)

olkin\_auc12\_13, [8](#)

olkin\_auc12\_3, [8](#)

olkin\_auc12\_34, [9](#)

olkin\_auc1\_2, [9](#)

## \*Topic **variance**

auc\_diff, [2](#)

auc\_var, [4](#)

auc\_diff, [2](#)

auc\_var, [4](#)

dat1, [5](#)

dat2, [6](#)

olkin\_auc1, [6](#)

olkin\_auc12, [7](#)

olkin\_auc12\_1, [7](#)

olkin\_auc12\_13, [8](#)

olkin\_auc12\_3, [8](#)

olkin\_auc12\_34, [9](#)

olkin\_auc1\_2, [9](#)