

Package ‘correctedAUC’

July 22, 2025

Type Package

Title Correcting AUC for Measurement Error

Version 0.0.3

Date 2016-6-29

Author Bernard Rosner, Shelley Tworoger, Weiliang Qiu

Maintainer Weiliang Qiu <stwxq@channing.harvard.edu>

Depends R (>= 3.1.0), stats

Imports ICC, mnormt

Description Correcting area under ROC (AUC) for measurement error based on probit-shift model.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2016-06-29 23:22:20

Contents

AUCest.Reiser	2
AUCest.Rosner	3
genSimDataModelII	5
genSimDataModelIII	7
genSimDataReiser	8

Index	11
--------------	-----------

AUCest.Reiser	<i>Calculate AUC corrected for measurement error based on Reiser's (2000) method</i>
---------------	--

Description

Calculate AUC corrected for measurement error based on Reiser's (2000) method.

Usage

```
AUCest.Reiser(
  datFrame,
  sidVar = "subjID",
  obsVar = "y",
  grpVar = "grp",
  repVar = "myrep",
  alpha = 0.05)
```

Arguments

datFrame	a data frame with at least the following columns: y: numerical vector of observations; subjID: vector of subject ids; grp: group indicator: 1 means case; and 0 means control; myrep: integer vector indicating replication. should be consecutive positive integer starting from 1.
sidVar	character. variable name for subject id in the data frame datFrame.
obsVar	character. variable name for observations in the data frame datFrame.
grpVar	character. variable name for group indicator in the data frame datFrame.
repVar	character. variable name for replication indicator in the data frame datFrame.
alpha	confidence interval level $100(1 - \alpha)\%$

Value

A list of 4 elements

AUC.c	AUC corrected for measurement error based on Reiser's (2000) method.
sd.AUC.c	standard error of the estimated AUC corrected for measurement error based on Reiser's (2000) method.
AUC.c.low	lower bound of the AUC.c.
AUC.c.upp	upper bound of the AUC.c.

Author(s)

Bernard Rosner <stbar@channing.harvard.edu>, Shelley Tworoger <nhsst@channing.harvard.edu>, Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner B, Tworoger S, Qiu W (2015) Correcting AUC for Measurement Error. J Biom Biostat 6:270. doi:10.4172/2155-6180.1000270

Examples

```
set.seed(1234567)
tt=genSimDataReiser(
  nX = 100,
  nY = 100,
  sigma.X2 = 1,
  mu.X = 0.25,
  sigma.Y2 = 1,
  mu.Y = 0,
  sigma.epsilon2 = 0.5,
  sigma.eta2 = 0.5)

print(dim(tt$datFrame))
print(tt$datFrame[1:2,1:3])
print(tt$theta2)
print(tt$mu.true)
print(tt$AUC.true)

res = AUCest.Reiser(
  datFrame = tt$datFrame,
  sidVar = "subjID",
  obsVar = "y",
  grpVar = "grp",
  repVar = "myrep",
  alpha = 0.05)
print(res)
```

AUCest.Rosner

Calculate AUC.c for measurement error based on probit-shift model

Description

Calculate AUC.c for measurement error based on probit-shift model.

Usage

```
AUCest.Rosner(
  datFrame,
  sidVar = "subjID",
  obsVar = "y",
  grpVar = "grp",
  repVar = "myrep",
  alpha = 0.05)
```

Arguments

<code>datFrame</code>	a data frame with at least the following columns: <code>y</code> : numerical vector of observations; <code>subjID</code> : vector of subject ids; <code>grp</code> : group indicator: 1 means case; and 0 means control; <code>myrep</code> : integer vector indicating replication. should be consecutive positive integer starting from 1.
<code>sidVar</code>	character. variable name for subject id in the data frame <code>datFrame</code> .
<code>obsVar</code>	character. variable name for observations in the data frame <code>datFrame</code> .
<code>grpVar</code>	character. variable name for group indicator in the data frame <code>datFrame</code> .
<code>repVar</code>	character. variable name for replication indicator in the data frame <code>datFrame</code> .
<code>alpha</code>	confidence interval level $100(1 - \alpha)\%$

Value

A list of 9 elements:

<code>AUC.obs</code>	AUC estimated based on the Mann-Whitney statistic.
<code>AUC.c</code>	AUC corrected for measurement error based on the probit-shift model.
<code>ICC.x</code>	intra-class correlation for cases.
<code>ICC.y</code>	intra-class correlation for controls
<code>mu.mle</code>	maximum likelihood estimate of μ (i.e., the shift between the case distribution and the control distribution after probit transformation)
<code>AUC.obs.low</code>	lower bound of the <code>AUC.obs</code> .
<code>AUC.obs.upp</code>	upper bound of the <code>AUC.obs</code> .
<code>AUC.c.low</code>	lower bound of the <code>AUC.c</code> .
<code>AUC.c.upp</code>	upper bound of the <code>AUC.c</code> .

Author(s)

Bernard Rosner <stbar@channing.harvard.edu>, Shelley Tworoger <nhsst@channing.harvard.edu>, Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner B, Tworoger S, Qiu W (2015) Correcting AUC for Measurement Error. *J Biom Biostat* 6:270. doi:10.4172/2155-6180.1000270

Examples

```
set.seed(1234567)
tt=genSimDataModelIII(
  nX = 100,
  nY = 100,
  mu = 0.25,
  lambda = 0,
  sigma.X2 = 1,
  sigma.Y2 = 1,
```

```

sigma.e.X = 1,
sigma.e.Y = 1)

print(dim(tt$datFrame))
print(tt$datFrame[1:2,1:3])
print(tt$AUC.true)

res = AUCest.Rosner(
  datFrame = tt$datFrame,
  sidVar = "subjID",
  obsVar = "y",
  grpVar = "grp",
  repVar = "myrep",
  alpha = 0.05)
print(res)

```

genSimDataModelIII	<i>Generate one simulated data set based on Model II in Rosner et al's (2015) manuscript</i>
--------------------	--

Description

Generate one simulated data set based on Model II in Rosner et al's (2015) manuscript.

Usage

```

genSimDataModelIII(
  nX,
  nY,
  mu,
  lambda,
  sigma.X2,
  sigma.Y2,
  sigma.e.X,
  sigma.e.Y)

```

Arguments

nX	integer. number of cases.
nY	integer. number of controls.
mu	difference of means between the case distribution and control distribution.
lambda	mean for controls.
sigma.X2	variance of the true value for cases.
sigma.Y2	variance of the true value for controls.
sigma.e.X	variance of the random error term for cases.
sigma.e.Y	variance of the random error term for controls.

Details

The Model II in Rosner et al.'s (2005) manuscript:

$$X_{ik,obs} = X_{i,true} + \epsilon_{ik}, \log(X_{i,true}) \sim N(\lambda + \mu, \sigma_X^2), \epsilon_{ik} \sim N(0, \sigma_\epsilon^2), i = 1, \dots, n_X, k = 1, 2$$

$$Y_{jl,obs} = Y_{j,true} + \xi_{jl}, \log(Y_{j,true}) \sim N(\lambda, \sigma_Y^2), \xi_{jl} \sim N(0, \sigma_\eta^2), j = 1, \dots, n_Y, l = 1, 2$$

Value

A list of 2 elements:

datFrame	A data frame with 4 elements: y: observations; subjID: subject ID; grp: group indicator; myrep: replication indicator.
AUC.true	true AUC value

Author(s)

Bernard Rosner <stbar@channing.harvard.edu>, Shelley Tworoger <nhsst@channing.harvard.edu>, Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner B, Tworoger S, Qiu W (2015) Correcting AUC for Measurement Error. J Biom Biostat 6:270. doi:10.4172/2155-6180.1000270

Examples

```
set.seed(1234567)
tt=genSimDataModelII(
  nX = 100,
  nY = 100,
  mu = 0.25,
  lambda = 0,
  sigma.X2 = 1,
  sigma.Y2 = 1,
  sigma.e.X = 1,
  sigma.e.Y = 1)

print(dim(tt$datFrame))
print(tt$datFrame[1:2,1:3])
print(tt$AUC.true)
```

genSimDataModelIII	<i>Generate one simulated data set based on Model III in Rosner et al's (2015) manuscript</i>
--------------------	---

Description

Generate one simulated data set based on Model III in Rosner et al's (2015) manuscript.

Usage

```
genSimDataModelIII(
  nX,
  nY,
  mu,
  lambda,
  sigma.X2,
  sigma.Y2,
  sigma.e.X,
  sigma.e.Y)
```

Arguments

nX	integer. number of cases.
nY	integer. number of controls.
mu	difference of means between the case distribution and control distribution.
lambda	mean for controls.
sigma.X2	variance of the true value for cases.
sigma.Y2	variance of the true value for controls.
sigma.e.X	variance of the random error term for cases.
sigma.e.Y	variance of the random error term for controls.

Details

The Model III in Rosner et al.'s (2005) manuscript:

$$X_{ik,obs} = X_{i,true} + \epsilon_{ik}, \log(X_{i,true}) \sim N(\lambda + \mu, \sigma_X^2), \log(\epsilon_{ik}) \sim N(0, \sigma_\epsilon^2), i = 1, \dots, n_X, k = 1, 2$$

$$Y_{jl,obs} = Y_{j,true} + \xi_{jl}, \log(Y_{j,true}) \sim N(\lambda, \sigma_Y^2), \log(\xi_{jl}) \sim N(0, \sigma_\eta^2), j = 1, \dots, n_Y, l = 1, 2$$

Value

A list of 2 elements:

datFrame	A data frame with 4 elements: y: observations; subjID: subject ID; grp: group indicator; myrep: replication indicator.
AUC.true	true AUC value

Author(s)

Bernard Rosner <stbar@channing.harvard.edu>, Shelley Tworoger <nhsst@channing.harvard.edu>, Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner B, Tworoger S, Qiu W (2015) Correcting AUC for Measurement Error. J Biom Biostat 6:270. doi:10.4172/2155-6180.1000270

Examples

```
set.seed(1234567)
tt=genSimDataModelIII(
  nX = 100,
  nY = 100,
  mu = 0.25,
  lambda = 0,
  sigma.X2 = 1,
  sigma.Y2 = 1,
  sigma.e.X = 1,
  sigma.e.Y = 1)

print(dim(tt$datFrame))
print(tt$datFrame[1:2,1:3])
print(tt$AUC.true)
```

genSimDataReiser

Generate one simulated data set based on Reiser's (2000) model

Description

Generate one simulated data set based on Reiser's (2000) model. The true AUC will also be calculated.

Usage

```
genSimDataReiser(
  nX = 100,
  nY = 100,
  sigma.X2 = 1,
  mu.X = 0.25,
  sigma.Y2 = 1,
  mu.Y = 0,
  sigma.epsilon2 = 0.5,
  sigma.eta2 = 0.5)
```

Arguments

nX	integer. number of cases.
nY	integer. number of controls.
sigma.X2	variance of the true value for cases.
mu.X	mean of the true value for cases.
sigma.Y2	variance of the true value for controls.
mu.Y	mean of the true value for controls.
sigma.epsilon2	variance of the random error term for cases.
sigma.eta2	variance of the random error term for controls.

Details

Reiser's (2000) measurement error model is:

$$X_{ik,obs} = X_{i,true} + \epsilon_{ik}, X_{i,true} \sim N(\mu_X, \sigma_X^2), \epsilon_{ik} \sim N(0, \sigma_\epsilon^2), i = 1, \dots, n_X, k = 1, 2$$

$$Y_{jl,obs} = Y_{j,true} + \xi_{jl}, Y_{j,true} \sim N(\mu_Y, \sigma_Y^2), \xi_{jl} \sim N(0, \sigma_\eta^2), j = 1, \dots, n_Y, l = 1, 2$$

Value

A list of 4 elements:

datFrame	A data frame with 4 elements: y: observations; subjID: subject ID; grp: group indicator; myrep: replication indicator.
theta2	$(\sigma_{\epsilon}^2 + \sigma_{\eta}^2) / (\sigma_X^2 + \sigma_Y^2)$
mu.true	$\mu_X - \mu_Y$
AUC.true	true AUC value

Author(s)

Bernard Rosner <stbar@channing.harvard.edu>, Shelley Tworoger <nhsst@channing.harvard.edu>, Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner B, Tworoger S, Qiu W (2015) Correcting AUC for Measurement Error. J Biom Biostat 6:270. doi:10.4172/2155-6180.1000270

Examples

```
set.seed(1234567)
tt=genSimDataReiser(
  nX = 100,
  nY = 100,
  sigma.X2 = 1,
  mu.X = 0.25,
  sigma.Y2 = 1,
```

```
mu.Y = 0,  
sigma.epsilon2 = 0.5,  
sigma.eta2 = 0.5)  
  
print(dim(tt$datFrame))  
print(tt$datFrame[1:2,1:3])  
print(tt$theta2)  
print(tt$mu.true)  
print(tt$AUC.true)
```

Index

* **method**

- AUCest.Reiser, [2](#)
- AUCest.Rosner, [3](#)
- genSimDataModelIII, [5](#)
- genSimDataModelIII, [7](#)
- genSimDataReiser, [8](#)

- AUCest.Reiser, [2](#)
- AUCest.Rosner, [3](#)

- genSimDataModelIII, [5](#)
- genSimDataModelIII, [7](#)
- genSimDataReiser, [8](#)