

# Simulate MRMC Data

There are two functions in this package that can simulate Multi-reader Multi-case (MRMC) data. `sim.NormalIG.Hierarchical` is the function to simulate the MRMC agreement data with no binary truth state. The `sim.gRoeMetz` is the function to simulate MRMC ROC data.

## 1. Using the new hierachical model to simulate the MRMC agreement data

The following is how to use the `sim.NormalIG.Hierachcial.R` to simulate MRMC data with no truth state, the MRMC agreement data

```
# configuration
nR = 5      #number of readers
nC = 100    #number of cases

config <- sim.NormalIG.Hierarchical.config(nR=nR, nC=nC, modalityID = c("testA","testB"))

# simulate MRMC study
set.seed(1, kind = "L'Ecuyer-CMRG")
dFrame.newH <- sim.NormalIG.Hierarchical(config)

# check the first and last few lines of the simulated dataframe
print(dFrame.newH[1:6, ])
```

```
##   caseID readerID modalityID      score
## 1  Case1  reader1      testA  0.6946134
## 2  Case2  reader1      testA  2.9965743
## 3  Case3  reader1      testA  1.7412831
## 4  Case4  reader1      testA -1.2549331
## 5  Case5  reader1      testA -0.2952628
## 6  Case6  reader1      testA  3.0468024
```

```
print(dFrame.newH[(nrow(dFrame.newH) - 5):nrow(dFrame.newH), ])
```

```
##      caseID readerID modalityID      score
## 995  Case95  reader5      testB -0.1549758
## 996  Case96  reader5      testB -1.2040039
## 997  Case97  reader5      testB  1.6544295
## 998  Case98  reader5      testB -1.2125349
## 999  Case99  reader5      testB -0.2017070
## 1000 Case100 reader5      testB  2.0605555
```

We simulated 5 readers and 100 cases for 2 modalities, so the total number of scores is  $5 \times 100 \times 2 = 1000$ .

## 2. Using sim.gRoeMetz in iMRMC package to simulate MRMC ROC data

The following is how to use the sim.gRoeMetz.R to simulate MRMC data with truth state, the MRMC ROC data

```
# configuration
nR = 5          #number of readers
nC.neg = 50     #number of positive cases
nC.pos = 50     #number of negative cases

config <- sim.gRoeMetz.config(nR = nR, nC.neg = nC.neg, nC.pos = nC.neg)

# simulate MRMC study
set.seed(1, kind = "L'Ecuyer-CMRG")
dFrame.gRM <- sim.gRoeMetz(config)

# check the first and last few lines of the simulated dataframe
print(dFrame.gRM[1:6, ])
```

```
##  readerID  caseID modalityID score
## 1    truth negCase1      truth    0
## 2    truth negCase2      truth    0
## 3    truth negCase3      truth    0
## 4    truth negCase4      truth    0
## 5    truth negCase5      truth    0
## 6    truth negCase6      truth    0
```

```
print(dFrame.gRM[(nrow(dFrame.gRM) - 5):nrow(dFrame.gRM), ])
```

```
##      readerID  caseID modalityID      score
## 1095  reader5 posCase45      testB  2.9342657
## 1096  reader5 posCase46      testB  1.0639768
## 1097  reader5 posCase47      testB  0.9885389
## 1098  reader5 posCase48      testB -0.5130517
## 1099  reader5 posCase49      testB  0.9254049
## 1100  reader5 posCase50      testB  2.4164279
```

The simulated data starts with truth state of each case, and followed by the reading scores from each of the readers. Since we simulate 50 positive cases and 50 negative cases, there are 100 lines for the truth and  $100 \times 5 \times 2 = 1000$  lines for the scores from 5 readers for 2 modalities.

To combine the truth data and the reader scores and change the data to a dataframe with 5 columns: readerID, caseID, modalityID, score, truth, we can use the function undoIMRMCdf in iMRMC package:

```
dFrame.gRM.2 <- undoIMRMCdf(dFrame.gRM)

# check the first and last few lines of the simulated dataframe
print(dFrame.gRM.2[1:6, ])
```

```
##      caseID readerID modalityID      score truth
## 1 negCase1  reader1      testA  0.36795249    0
## 2 negCase1  reader2      testB  0.05607467    0
```

```
## 3 negCase1 reader3 testB -0.46914121 0
## 4 negCase1 reader4 testA -1.38686722 0
## 5 negCase1 reader1 testB -0.49895246 0
## 6 negCase1 reader5 testA -1.64836669 0
```

```
head(dFrame.gRM.2)
```

```
##      caseID readerID modalityID      score truth
## 1 negCase1 reader1      testA  0.36795249      0
## 2 negCase1 reader2      testB  0.05607467      0
## 3 negCase1 reader3      testB -0.46914121      0
## 4 negCase1 reader4      testA -1.38686722      0
## 5 negCase1 reader1      testB -0.49895246      0
## 6 negCase1 reader5      testA -1.64836669      0
```

```
print(dFrame.gRM.2[(nrow(dFrame.gRM.2) - 5):nrow(dFrame.gRM.2), ])
```

```
##      caseID readerID modalityID      score truth
## 995 posCase9 reader5      testA  0.9233879      1
## 996 posCase9 reader4      testB  4.1614474      1
## 997 posCase9 reader4      testA  3.0874475      1
## 998 posCase9 reader2      testB  3.4945769      1
## 999 posCase9 reader1      testA  2.2259825      1
## 1000 posCase9 reader5      testB  1.3017587      1
```

```
tail(dFrame.gRM.2)
```

```
##      caseID readerID modalityID      score truth
## 995 posCase9 reader5      testA  0.9233879      1
## 996 posCase9 reader4      testB  4.1614474      1
## 997 posCase9 reader4      testA  3.0874475      1
## 998 posCase9 reader2      testB  3.4945769      1
## 999 posCase9 reader1      testA  2.2259825      1
## 1000 posCase9 reader5      testB  1.3017587      1
```