

# Linnorm Package

March 5, 2016

**Type** Package

**Title** Linear model and normality based normalization (Linnorm) for RNA-seq expression dataset or large scale count data.

**Version** 1.0

**Date** 2015-11-27

**Maintainer** Ken Shun Hang Yip <shunyip@bu.edu>

**Description** This is a package for Linear model and normality based normalization (Linnorm) for RNA-seq or count data.

**License** MIT + file LICENSE

**Imports** Rcpp (>= 0.12.2),  
RcppArmadillo,  
MASS

**LinkingTo** Rcpp, RcppArmadillo, MASS

**RoxygenNote** 5.0.1

## R topics documented:

GammaSim . . . . .	1
Linnorm . . . . .	2
LogNormSim . . . . .	3
NBSim . . . . .	4
PoissonSim . . . . .	4
<b>Index</b>	<b>6</b>

---

GammaSim	<i>This function simulates a dataset with gamma distribution based on the input dataset.</i>
----------	--

---

## Description

This function simulates a dataset with gamma distribution based on the input dataset.

**Usage**

```
GammaSim(thisdata, NumRep = 3, NumDiff = 5000, NumFea = 20000,
  showinfo = FALSE, MaxLibSizeLog2FC = 0.5)
```

**Arguments**

thisdata	Matrix: The matrix or data frame that contains your dataset. Each row is a feature (ex. Gene) and each column is a replicate. Undefined values such as NA are not supported. It is assumed that all samples are replicates of the same sample.
NumRep	Integer: The number of replicates. This is half of the number of output samples. Defaults to 3.
NumDiff	Integer: The number of Differentially Changed Features. Defaults to 5000.
NumFea	Integer: The number of Total Features. Defaults to 20000.
showinfo	Logical: should we show data information on the console? Defaults to FALSE.
MaxLibSizeLog2FC	Double: The maximum library size difference from the mean that is allowed, in terms of log 2 fold change. Set to 0 to prevent program from generating library size differences. Defaults to 0.5.

**Examples**

```
simulateddata <- GammaSim(expMatrix)
```

---

Linnorm

*Linnorm Function*


---

**Description**

This function performs the Linear model and normality based normalization method (Linnorm).

**Usage**

```
Linnorm(datamatrix, showinfo = FALSE, method = "default",
  perturbation = 10, minZeroPortion = 2/3)
```

**Arguments**

datamatrix	The matrix or data frame that contains your dataset. Each row is a feature (ex. Gene) and each column is a sample. Undefined values such as NA are not supported.
method	"default" or "lambda" The program will output the transformed matrix if the method is "default". Otherwise, if the method is "lambda", the program will output a lambda value.

- perturbation** Integer  $\geq 2$ . This is the perturbation during the Iterated Local Search algorithm, when we are searching for the global minimum of the "deviation parameter" measure for lambda (Please refer to the article). The range of the areas searched locally will be the exponents of this perturbation, from closest to furthest from the local minima found by using the whole range. Increasing this value will make the program much faster, but less capable of finding the global minimum. Defaults to 10.
- minZeroPortion** double  $\geq 0$ ,  $\leq 1$  Rows without at least this portion of non-zero values will not be used in the calculation of normalizing parameter. Defaults to 2/3.

### Examples

```
normalizedExp <- Linnorm(expMatrix, showinfo = "TRUE")
```

---

LogNormSim	<i>This function simulates a dataset with log normal distribution based on the input dataset.</i>
------------	---

---

### Description

This function simulates a dataset with log normal distribution based on the input dataset.

### Usage

```
LogNormSim(thisdata, NumRep = 3, NumDiff = 5000, NumFea = 20000,
  showinfo = FALSE, MaxLibSizeLog2FC = 0.5)
```

### Arguments

- thisdata** Matrix: The matrix or data frame that contains your dataset. Each row is a feature (ex. Gene) and each column is a replicate. Undefined values such as NA are not supported. It is assumed that all samples are replicates of the same sample.
- NumRep** Integer: The number of replicates. This is half of the number of output samples. Defaults to 3.
- NumDiff** Integer: The number of Differentially Changed Features. Defaults to 5000.
- NumFea** Integer: The number of Total Features. Defaults to 20000.
- showinfo** Logical: should we show data information on the console? Defaults to FALSE.
- MaxLibSizeLog2FC** Double: The maximum library size difference from the mean that is allowed, in terms of log 2 fold change. Set to 0 to prevent program from generating library size differences. Defaults to 0.5.

### Examples

```
simulateddata <- LogNormSim(expMatrix)
```

---

NBSim	<i>This function simulates a dataset with negative binomial distribution based on the input dataset.</i>
-------	--

---

### Description

This function simulates a dataset with negative binomial distribution based on the input dataset.

### Usage

```
NBSim(thisdata, NumRep = 3, NumDiff = 5000, NumFea = 20000,
      showinfo = FALSE, MaxLibSizeLog2FC = 0.5)
```

### Arguments

thisdata	Matrix: The matrix or data frame that contains your dataset. Each row is a feature (ex. Gene) and each column is a replicate. Undefined values such as NA are not supported. It is assumed that all samples are replicates of the same sample.
NumRep	Integer: The number of replicates. This is half of the number of output samples. Defaults to 3.
NumDiff	Integer: The number of Differentially Changed Features. Defaults to 5000.
NumFea	Integer: The number of Total Features. Defaults to 20000.
showinfo	Logical: should we show data information on the console? Defaults to FALSE.
MaxLibSizeLog2FC	Double: The maximum library size difference from the mean that is allowed, in terms of log 2 fold change. Set to 0 to prevent program from generating library size differences. Defaults to 0.5.

### Examples

```
simulateddata <- NBSim(expMatrix)
```

---

PoissonSim	<i>This function simulates a dataset with poisson distribution based on the input dataset.</i>
------------	--

---

### Description

This function simulates a dataset with poisson distribution based on the input dataset.

### Usage

```
PoissonSim(thisdata, NumRep = 3, NumDiff = 5000, NumFea = 20000,
          showinfo = FALSE, MaxLibSizeLog2FC = 0.5)
```

**Arguments**

thisdata	Matrix: The matrix or data frame that contains your dataset. Each row is a feature (ex. Gene) and each column is a replicate. Undefined values such as NA are not supported. It is assumed that all samples are replicates of the same sample.
NumRep	Integer: The number of replicates. This is half of the number of output samples. Defaults to 3.
NumDiff	Integer: The number of Differentially Changed Features. Defaults to 5000.
NumFea	Integer: The number of Total Features. Defaults to 20000.
showinfo	Logical: should we show data information on the console? Defaults to FALSE.
MaxLibSizeLog2FC	Double: The maximum library size difference from the mean that is allowed, in terms of log 2 fold change. Set to 0 to prevent program from generating library size differences. Defaults to 0.5.

**Examples**

```
simulateddata <- PoissonSim(expMatrix)
```

# Index

\*Topic **Binomial**

NBSim, [4](#)

\*Topic **CPM**

Linnorm, [2](#)

\*Topic **Count**

GammaSim, [1](#)

Linnorm, [2](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **Expression**

GammaSim, [1](#)

Linnorm, [2](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **FPKM**

Linnorm, [2](#)

\*Topic **Gamma**

GammaSim, [1](#)

\*Topic **Linnorm**

Linnorm, [2](#)

\*Topic **Log**

LogNormSim, [3](#)

\*Topic **Negative**

NBSim, [4](#)

\*Topic **Normal**

LogNormSim, [3](#)

\*Topic **Parametric**

Linnorm, [2](#)

\*Topic **Poisson**

PoissonSim, [4](#)

\*Topic **RNA-seq**

GammaSim, [1](#)

Linnorm, [2](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **RPKM**

Linnorm, [2](#)

\*Topic **Raw**

GammaSim, [1](#)

Linnorm, [2](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **Simulation**

GammaSim, [1](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **TPM**

Linnorm, [2](#)

\*Topic **distribution**

GammaSim, [1](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)

\*Topic **normalization**

Linnorm, [2](#)

\*Topic **transformation**

Linnorm, [2](#)

GammaSim, [1](#)

Linnorm, [2](#)

LogNormSim, [3](#)

NBSim, [4](#)

PoissonSim, [4](#)