

# Package: countgmifs (via r-universe)

August 25, 2024

**Title** Discrete Response Regression for High-Dimensional Data

**Version** 0.0.2

**Description** Provides a function for fitting Poisson and negative binomial regression models when the number of parameters exceeds the sample size, using the the generalized monotone incremental forward stagewise method.

**Depends** R (>= 3.5.0), MASS

**License** GPL (>=2)

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 6.0.1.9000

**Repository** <https://kelliejarcher.r-universe.dev>

**RemoteUrl** <https://github.com/kelliejarcher/countgmifs>

**RemoteRef** HEAD

**RemoteSha** b94d7caea39893f615005d861f0be7d9b66aa900

## Contents

|                              |   |
|------------------------------|---|
| countgmifs-package . . . . . | 2 |
| coef.countgmifs . . . . .    | 2 |
| countgmifs . . . . .         | 3 |
| plot.countgmifs . . . . .    | 4 |
| predict.countgmifs . . . . . | 5 |
| print.countgmifs . . . . .   | 6 |
| summary.countgmifs . . . . . | 6 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>8</b> |
|--------------|----------|

---

|                    |  |
|--------------------|--|
| countgmifs-package | <i>Discrete Response Regression for High-Dimensional Data: Discrete Response Generalized Monotone Incremental Forward Stagewise Regression</i> |
|--------------------|--|

---

### Description

This package provides a function that fits a Poisson or negative binomial model when the number of parameters exceeds the sample size, using the the generalized monotone incremental forward stagewise method.

### Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.

This package contains functions for fitting a penalized discrete response model (either negative binomial or Poisson) and extracting estimated coefficients, predictions, and plots. The model and methods can be used when the response to be predicted is discrete, and is particularly relevant when there are more covariates than observations.

### Author(s)

Kellie Archer [aut, cre] Kellie J. Archer <archer.43@osu.edu>

Maintainer: Kellie Archer <archer.43@osu.edu> Kellie J. Archer <archer.43@osu.edu>

### References

Makowski M., Archer K.J. (2015) Generalized monotone incremental forward stagewise method for modeling count data: application predicting micronuclei frequency. *Cancer Informatics*, 14(Suppl 2), 97–105.

---

|                 |                                    |
|-----------------|------------------------------------|
| coef.countgmifs | <i>Extract Model Coefficients.</i> |
|-----------------|------------------------------------|

---

### Description

A generic function which extracts the model coefficients from a fitted model object fit using countgmifs

### Usage

```
## S3 method for class 'countgmifs'
coef(object, model.select = "BIC", ...)
```

**Arguments**

|              |   |
|--------------|---|
| object       | an countgmifs fitted object.  |
| model.select | when x is specified any model along the solution path can be selected. The default is model.select="BIC" which calculates the predicted values using the coefficients from the model having the lowest BIC. Other options are model.select="AIC" or any numeric value from the solution path. |
| ...          | other arguments.  |

**See Also**

See Also [countgmifs](#), [predict.countgmifs](#), [summary.countgmifs](#), [plot.countgmifs](#)

---

|            |   |
|------------|---|
| countgmifs | <i>Discrete Response Generalized Monotone Incremental Forward Stagewise Regression.</i> |
|------------|---|

---

**Description**

This function can fit a Poisson or negative binomial model when the number of parameters exceeds the sample size, using the the generalized monotone incremental forward stagewise method.

**Usage**

```
countgmifs(formula, data, x = NULL, offset, subset, epsilon = 0.001,
  tol = 1e-05, scale = TRUE, verbose = FALSE, family = "nb", ...)
```

**Arguments**

|         |   |
|---------|---|
| formula | an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The left side of the formula is the ordinal outcome while the variables on the right side of the formula are the covariates that are not included in the penalization process. Note that if all variables in the model are to be penalized, an intercept only model formula should be specified. |
| data    | an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model.  |
| x       | an optional matrix of predictors that are to be penalized in the model fitting process.   |
| offset  | this can be used to specify an a priori known component to be included during fitting (e.g., denominator term). This should be NULL or a numeric vector of length equal to the number of cases.   |
| subset  | an optional vector specifying a subset of observations to be used in the fitting process.   |
| epsilon | small incremental amount used to update a coefficient at a given step.  |

|         |  |
|---------|--|
| tol     | the iterative process stops when the difference between successive log-likelihoods is less than this specified level of tolerance. |
| scale   | logical, if TRUE (default) the penalized predictors are centered and scaled.   |
| verbose | logical, if TRUE the step number is printed to the console (default is FALSE).   |
| family  | the type of count response model to be fit. Default is 'nb' for negative binomial; user can also specify 'poisson'.                |
| ...     | other arguments.   |

### See Also

See Also [coef.countgmifs](#), [summary.countgmifs](#), [predict.countgmifs](#), [plot.countgmifs](#)

### Examples

```
set.seed(26)
n <- 50 # Sample size
p <- 500 # Number of covariates
intercept <- .5
#True parameter values for the 500 covariates
beta <- c(log(1.5), log(1.5), -log(1.5), -log(1.5), -log(1.5), rep(0,495))
alpha <- 0.5 # Intercept
x <- matrix(rnorm(n*p,0,1), nrow=n, ncol=p, byrow=TRUE) #Covariate values
colnames(x) <- paste("Var",1:p, sep="")
mu <- exp(intercept + crossprod(t(x),beta))
y <- rnbinom(n=n, size=1/alpha ,mu=mu) # Discrete response
data <- data.frame(y,x)
nb <- countgmifs(y ~ 1 , data=data, offset=NULL, x=x, epsilon=0.01, tol=0.001,
  scale=TRUE, verbose=FALSE)
coef.AIC <- coef(nb, model.select="AIC")
coef.AIC[coef.AIC!=0]
predict(nb, model.select="AIC")
plot(predict(nb, model.select="AIC"), y)
plot(nb)
```

---

plot.countgmifs

*Plot Solution Path for a Count GMIFS Fitted Model.*

---

### Description

This function plots either the coefficient path, the AIC, or the log-likelihood for a fitted countgmifs object.

### Usage

```
## S3 method for class 'countgmifs'
plot(x, type = "trace", xlab = NULL, ylab = NULL,
  main = NULL, ...)
```

**Arguments**

|      |   |
|------|---|
| x    | a countgmifs object.  |
| type | default is "trace" which plots the coefficient path for the fitted object. Also available are "AIC", "BIC", and "logLik". |
| xlab | a default x-axis label will be used which can be changed by specifying a user-defined x-axis label.                       |
| ylab | a default y-axis label will be used which can be changed by specifying a user-defined y-axis label.                       |
| main | a default main title will be used which can be changed by specifying a user-defined main title.                           |
| ...  | other arguments.  |

**See Also**

See Also [countgmifs](#), [coef.countgmifs](#), [summary.countgmifs](#), [predict.countgmifs](#)

---

predict.countgmifs      *Predict Outcome for Count GMIFS Fitted Model.*

---

**Description**

This function returns a numeric vector that is the predicted response from the countgmifs fitted object.

**Usage**

```
## S3 method for class 'countgmifs'
predict(object, neww = NULL, newdata, newx = NULL,
        model.select = "BIC", newoffset=NULL, ...)
```

**Arguments**

|              |   |
|--------------|---|
| object       | an ordinalgmifs fitted object.  |
| neww         | an optional formula that includes the unpenalized variables to use for predicting the response. If omitted, the training data are used.   |
| newdata      | an optional data.frame that minimally includes the unpenalized variables to use for predicting the response. If omitted, the training data are used.  |
| newx         | an optional matrix of penalized variables to use for predicting the response. If omitted, the training data are used.   |
| model.select | when x is specified any model along the solution path can be selected. The default is model.select="BIC" which calculates the predicted values using the coefficients from the model having the lowest BIC. Other options are model.select="AIC" or any numeric value from the solution path. |
| newoffset    | If an offset is used in the fit, then one must be supplied for making predictions.  |
| ...          | other arguments.  |

**See Also**

See Also [countgmifs](#), [coef.countgmifs](#), [summary.countgmifs](#), [plot.countgmifs](#)

---

`print.countgmifs`      *Print the Contents of a Count GMIFS Fitted Object.*

---

**Description**

This function prints the names of the list objects from an `countgmifs` fitted model

**Usage**

```
## S3 method for class 'countgmifs'
print(x, ...)
```

**Arguments**

`x`                    an `countgmifs` fitted object.  
`...`                other arguments.

**See Also**

See Also [countgmifs](#), [coef.countgmifs](#), [summary.countgmifs](#), [plot.countgmifs](#)

---

`summary.countgmifs`      *Summarize a Count GMIFS Object.*

---

**Description**

Prints the following items extracted from the fitted `countgmifs` object: the family used and model parameter estimates. For models that include `x`, the parameter estimates, AIC, BIC, and log-likelihood are printed for indicated `model.select` step or if `model.select` is not supplied the step at which the minimum BIC was observed.

**Usage**

```
## S3 method for class 'countgmifs'
summary(object, model.select = "BIC", ...)
```

**Arguments**

`object`                an `countgmifs` fitted object.  
`model.select`        when `x` is specified any model along the solution path can be selected. The default is `model.select="BIC"` which calculates the predicted values using the coefficients from the model having the lowest BIC. Other options are `model.select="AIC"` or any numeric value from the solution path.  
`...`                other arguments.

**See Also**

See Also [countgmifs](#), [coef.countgmifs](#), [predict.countgmifs](#), [plot.countgmifs](#)

# Index

## \* **methods**

- coef.countgmifs, 2
- countgmifs, 3
- plot.countgmifs, 4
- predict.countgmifs, 5
- print.countgmifs, 6
- summary.countgmifs, 6

## \* **models**

- countgmifs-package, 2

## \* **package**

- countgmifs-package, 2

## \* **regression**

- countgmifs, 3
- countgmifs-package, 2

coef.countgmifs, 2, 4-7

countgmifs, 3, 3, 5-7

countgmifs-package, 2

fitted.countgmifs (predict.countgmifs),  
5

plot.countgmifs, 3, 4, 4, 6, 7

predict.countgmifs, 3-5, 5, 7

print.countgmifs, 6

summary.countgmifs, 3-6, 6