Package ‘sizeMat’

April 14, 2020

Type Package
Title Estimate Size at Sexual Maturity
Version 1.1.1
Date 2020-04-14
Maintainer Josymar Torrejon-Magallanes <ejosymart@gmail.com>
Description Estimate morphometric and gonadal size at sexual maturity for organisms, usually fish and invertebrates. It includes methods for classification based on relative growth (using principal components analysis, hierarchical clustering, discriminant analysis), logistic regression (Frequentist or Bayes), parameters estimation and some basic plots.
Depends R (>= 3.5.0)
Imports MCMCpack, matrixStats, MASS, grDevices, graphics, stats, utils
Suggests rmarkdown, knitr
VignetteBuilder knitr
License GPL-2
LazyData TRUE
Encoding UTF-8
RoxygenNote 7.1.0
NeedsCompilation no
Author Josymar Torrejon-Magallanes [cre, aut]
Repository CRAN
Date/Publication 2020-04-14 13:40:02 UTC

R topics documented:

sizeMat-package .................................................. 2
classify_mature .................................................. 4
Crab data ......................................................... 5
Fish maturity ..................................................... 6
gonad_mature .................................................... 6
morph_mature .................................................... 8
Estimate morphometric and gonadal size at sexual maturity for organisms, usually fish and invertebrates. It includes methods for classification based on relative growth (principal components analysis, hierarchical clustering, discriminant analysis), logistic regression (frequentist or Bayes), parameters estimation and some basic plots. The size at sexual maturity is defined as the length at which a randomly chosen specimen has a 50% chance of being mature.

Details

Package: sizeMat
Type: Package
The Size at Morphometric and Gonad maturity are estimating using different functions (process).
1) The estimation of the Size at Morphometric Maturity involves two processes:
   1.1) A Principal Components Analysis is conducted with two allometric variables (x: independent variable, y: dependent variable) in log base, allowing to distinguish two groups that would represent juveniles and adult. The individuals are assigned to each group using a hierarchical classification procedure (hierarchical cluster). This method is based on establishing a predetermined number of groups (in this case, two) and assigning individuals to one of the groups according to their loads on the two axes of the PCA (Corgos & Freire, 2006). Using the results of the classification (PCA + cluster), a discriminant analysis (linear or quadratic) is carried out to obtain a discriminating function that permitted any individuals to be classified as a juvenile or an adult on the basis of the X and Y allometric variables.
   1.2) After classification, the logistic approach is used. The size at 50% maturity ($L_{50}$) is estimated as the length at which a randomly chosen specimen has a 50% chance of being mature (Somerton 1980, Roa et al. 1999, Corgos & Freire 2006). In the regression analysis, $X$ (i.e: carapace width) is considered the explanatory variable and the classification $CS$ (juveniles: 0, adults: 1) is considered the response variable (binomial).

The variables are fitted to a logistic function with the form:

$$P_{CS} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot X)}}$$

where:
$P_{CL}$ is the probability of an individual of being mature at a determinate $X$ length.

$\beta_0$ (intercept) and $\beta_1$ (slope) are parameters estimated.

The ($L_{50}$) is calculated as:

$$L_{50} = -\frac{\beta_0}{\beta_1}$$

Some basic plotting (classification, $\beta_0$, $\beta_1$ and $L_{50}$ histogram, and maturity ogive) are also provided.

2) The estimation of Size at Gonad Maturity use the logistic approach.

To estimate size at gonadal maturity, the database must contains the stage of sexual maturity and at least one allometric variable (i.e: total length, fork length, carapace width). The stage of sexual maturity is referred to the gonadal maturation stages (i.e: I, II, III, IV or 0, 1, etc).

So, in the regression analysis, the allometric variable (i.e: total length) is considered the explanatory variable and the stage of sexual maturity (immature: 0, mature: 1) is considered the response variable (binomial). The regression analysis is performed in the same way as the size at morphometric maturity.

**Author(s)**

Josymar Torrejon-Magallanes <ejosymart@gmail.com>

**References**


**Examples**

#See examples for functions morph_mature() and gonad_mature().
Classify mature

**Description**

Classify the individuals in two groups (0: juveniles, 1: adults) based on relative growth.

**Usage**

```r
classify_mature(
  data,
  varNames = c("x", "y"),
  varSex = "sex",
  selectSex = NULL,
  method = "ld"
)
```

**Arguments**

- `data` data.frame with allometric variables and sex category (male, female). If sex category contains NA's, that row will be filtered.
- `varNames` the name of two allometric variables to be used for analysis.
- `varSex` the name of the variable containing sex information.
- `selectSex` sex category to be used for analysis. If `selectSex = NULL` all the individuals will be used in the analysis.
- `method` a character string indicating the discriminant analysis method, linear discriminant analysis "ld", quadratic discriminant analysis "qd". We suggest begin the analysis using the `method = "ld"`.

**Details**

Classify the individuals in two groups (juveniles = 0 and adult = 1).

A Principal Components Analysis was conducted with two allometric variables (x: independent variable, y: dependent variable) in log base, allowing to distinguish two groups that would represent juveniles and adult. The individuals are assigned to each group using a hierarchical classification procedure (hierarchical cluster with agglomeration method: "Ward.D" and the distance measure: "euclidean"). This method is based on establishing a predetermined number of groups (in this case, two) and assigning individuals to one of the groups according to their loads on the two axes of the PCA (Corgos & Freire, 2006).

Using the results of the classification (PCA + cluster), a discriminant analysis (linear or quadratic) is conducted to obtain a discriminating function that permitted any individuals to be classified as a juvenile or an adult on the basis of the X and Y allometric variables.

**Value**

A data.frame of class 'classify', with x (independent), y (dependent) and classification of maturity (juveniles = 0, adult = 1) variables.
Examples

```r
data(crabdata)

classify_data = classify_mature(crabdata, varNames = c("carapace_width", "chela_height"),
                         varSex = "sex_category", selectSex = NULL, method = "ld")

classify_data
```

Description

A database containing allometric measures and other attributes of 223 crabs of the specie *Chionectestanneri*. The variables carapace width and chela height were published in Somerton & Donaldson (1996).

Usage

```r
data(crabdata)
```

Format

A data frame with different variables:

- year
- month
- carapace_width (mm.)
- carapace_length (mm.)
- chela_height (mm.)
- chela_width (mm.)
- sex_category (m: male, f: female).

Source


References

Fish maturity

Fish gonadal maturity data

Description
A data set containing the total length and the stages of sexual maturity of 1733 fishes.

Usage
data(matFish)

Format
A data frame with two variables:
- `total_length` the total length in cm.
- `stage_mat` the stages of sexual maturity (gonadal maturation stages, I: immature, II-III-IV: mature)

---

gonad_mature

Estimate gonadal maturity

Description
Estimate size at gonad maturity.

Usage
gonad_mature(
data, 
varNames = c("allometric", "stage"), 
inmName = "inm", 
matName = "mad", 
method = "fq", 
niter = 999, 
seed = 70388 
)

Arguments
data data.frame with allometric variables and stage of sexual maturity (gonad maturation stages).

varNames a character string indicating the name of the allometric and the stage of sexual maturity variables to be used for analysis.
inmName a character string indicating the name or names of the immaturity stage.
matName a character string indicating the name or names of the maturity stage.
gonad_mature

method a character string indicating the method to be applied, "fq" frequentist GLM, or "bayes" Bayes GLM (MCMClogit function).
niter number of iterations (bootstrap resampling).
seed a single value, interpreted as an integer.

Details

Estimate the size at gonad maturity using a logistic regression with X variable and stages of sexual maturity (two categories: immature and mature).

The function requires a data.frame with the X (allometric variable) and the stage of sexual maturity (gonad maturation stage).

The argument varNames requires a character string indicating the name of one allometric and the stage of sexual maturity variable to be used for analysis (i.e varNames = c("total_length", "stage_mat")). So the argument varNames must contain two character strings only, the first is the allometric variable and the second is the stage of sexual maturity.

The arguments inmName and matName require a character string indicating the name of the stages of sexual maturity in the data.frame. The argument could contain one character string or could be a vector (i.e inmName = "I", matName = c("II", "III", "IV")).

The argument method requires a character string indicating which regression will be used for the test. If method = "fq" the logistic regression is based on GLM (frequentist), if method = "bayes" a sample from the posterior distribution of a logistic regression model using a random walk Metropolis algorithm is generated (see MCMClogit function).

The argument niter requires a number. For the GLM regression (method = "fq"), a non-parametric bootstrap method consists in generate B bootstrap samples, by resampling with replacement the original data. Then all statistics for each parameter can be calculated from each bootstrap sample (median and confidence intervals). For the method = "bayes", the argument niter is related to the number of Metropolis iterations for the sampler.

Value

An object of class 'gonadMat'.
model the summary statistics of the model.
A_boot the 'n iter' values of parameter A.
B_boot the 'n iter' values of parameter B.
L50 the 'n iter' values of parameter L50 (size at gonad maturity).
out a dataframe with the allometric variable "X", stage of sexual maturity, the fitted values for logistic regression and confidence intervals (95%). Also the summary statistics of the model is provided.

Examples

data(matFish)

gonad_mat = gonad_mature(matFish, varNames = c("total_length", "stage_mat"), inmName = "I", matName = c("II", "III", "IV"), method = "fq", niter = 50)
# 'niter' parameters:
gonad_mat$A_boot
gonad_mat$B_boot
gonad_mat$L50_boot
gonad_mat$out

morph_mature

## Estimate morphometric mature

### Description

Estimate size at morphometric maturity \( (L_{50}) \).

### Usage

```r
morph_mature(data, method = "fq", niter = 999, seed = 70388)
```

### Arguments

- **data**: an object of class 'classify' with the allometric variables \( (X, Y) \) and classification of maturity (juveniles = 0, adults = 1).
- **method**: a character string indicating the method to be applied, "fq" frequentist GLM, or "bayes" Bayes GLM (MCMClogit function).
- **niter**: number of iterations (bootstrap resampling).
- **seed**: a single value, interpreted as an integer.

### Details

Estimate the size at morphometric maturity using a logistic regression with \( X \) variable and maturity classification (two categories: juveniles and adults).

The function requires an object of class "classify" with the \( X, Y \) (allometric variables) and classification of maturity (juveniles = 0, adults = 1).

The argument method requires a character string indicating which regression will be used for the test. If method = "fq" the logistic regression is based on GLM (frequentist) and if method = "bayes" a sample from the posterior distribution of a logistic regression model using a random walk Metropolis algorithm is generated (see MCMClogit function).

The argument niter requires a number. For the GLM regression (method = "fq"), a non-parametric bootstrap method consists in generate B bootstrap samples, by resampling with replacement the original data. Then all statistics for each parameter can be calculated from each bootstrap sample (median and confidence intervals). For the method = "bayes", the argument 'niter' is related to the number of Metropolis iterations for the sampler.
Value

An object of class 'morphMat'.
model the summary statistics of the model.
A_boot the 'n iter' values of parameter A.
B_boot the 'n iter' values of parameter B.
L50 the 'n iter' values of parameter L50 (size at morphometric maturity).
out a dataframe with the allometric variables "X" and "Y", classification of maturity, the fitted values for logistic regression and confidence intervals (95%). Also the summary statistics of the model is provided.

Examples

data(crabdata)

classify_data = classify_mature(crabdata, varNames = c("carapace_width", "chela_height"),
varSex = "sex_category", selectSex = NULL, method = "ld")

my_mature = morph_mature(classify_data, method = "fq", niter = 50)

# 'niter' parameters:
my_mature$A_boot
my_mature$B_boot
my_mature$L50_boot
my_mature$out

\begin{tabular}{ll}
\textbf{nagelkerkeR2} & \textit{Nagelkerme method R-square} \\
\end{tabular}

Description

Estimate Nagelkerke’s R squared from the result of glm(). Evaluate the goodness of fit for logistic regression.

Usage

nagelkerkeR2(x)

Arguments

x An object of class 'glm'.

Value

Rsquare Nagelkerke’s R squared.
Examples

```r
set.seed(7388)
n <- 300
x <- rnorm(n)
a <- 1
b <- -2
p <- 1/(1+exp(a+b*x))
y <- factor(ifelse(runif(n) < p, 1, 0), levels = 0:1)
mod1 <- glm(y ~ x, family=binomial)
nagelkerkeR2(mod1)
```

---

**plot.classify**  
*Plot method for classify class*

**Description**

Plot method for classify class

**Usage**

```r
## S3 method for class 'classify'
plot(
x,
xlab = "X",
ylab = "Y",
col = c(1, 2),
pch = c(4, 5),
cex = c(1, 1),
lty_lines = c(1, 1),
lwd_lines = c(1, 1),
legendPlot = TRUE,
cex_label = 0.8,
...)
```

**Arguments**

- `x` an object of class 'classify' with the allometric variables ("X", "Y") and classification of maturity (juveniles = 0, adults = 1).
- `xlab` a title for the x axis.
- `ylab` a title for the y axis.
- `col` the colors for juveniles and adults group.
- `pch` the character indicating the type of plotting.
- `cex` character expansion in the regression.
- `lty_lines` the line type in the regression.
`lwd_lines` the line width in the regression.
`legendPlot` legend in the plot (FALSE or TRUE).
`cex_label` size of the legendPlot
...
Additional arguments to the plot method.

**Examples**

data(crabdata)

classify_data = classify_mature(crabdata, varNames = c("carapace_width", "chela_height"),
varSex = "sex_category", selectSex = NULL, method = "ld")

## Showing different plots
plot(classify_data, xlab = "X")
plot(classify_data, xlab = "X", ylab = "Y", col = c(1, 2), pch = c(4, 5), cex = c(1, 3))
plot(classify_data, xlab = "Carapace width (mm.)", ylab = "Y", col = c(1, 2),
pch = c(4, 5), cex = c(1, 3), lwd_lines = c(1, 3))
plot(classify_data, xlab = "Carapace width (mm.)", ylab = "Y", col = c(1, 2),
pch = c(4, 5), cex = c(1, 3), lwd_lines = c(1, 3), main = "Classification")

---

**plot.gonadMat**  
*Plot method for gonadMat class (size at gonad maturity)*

**Description**

Plot method for gonadMat class (size at gonad maturity)

**Usage**

```r
## S3 method for class 'gonadMat'
plot(
x,  
xlab = "X",  
ylab = "Proportion mature",  
col = c("blue", "red"),  
lwd = 2,  
lty = 2,  
vline_hist = "black",  
lwd_hist = 2,  
lty_hist = 2,  
onlyOgive = FALSE,  
...  
)
```
### Arguments

- **x**
  - Object of class 'gonadMat' with the mature parameters and a data.frame with the X and stage of sexual maturity variables. Also the fitted values for the logistic regression and confidence intervals (95%).

- **xlab**
  - A title for the x axis.

- **ylab**
  - A title for the y axis.

- **col**
  - Color for the logistic curve and for the L50% size at gonad maturity.

- **lwd**
  - Line with for drawing fitted values and confidence intervals.

- **lty**
  - Line type line type for drawing fitted values and confidence intervals.

- **vline_hist**
  - Color of the vertical lines in the histogram. The lines represent the median and the confidence intervals.

- **lwd_hist**
  - Line with for the vertical line in the histogram.

- **lty_hist**
  - Line type for the vertical line in the histogram.

- **onlyOgive**
  - Plot only the ogive.

- **...**
  - Additional arguments to the plot method.

### Examples

```r
data(matFish)

gonad_mat = gonad_mature(matFish, varNames = c("total_length", "stage_mat"), inmName = "I", matName = c("II", "III", "IV"), method = "fq", niter = 50)

plot(gonad_mat, xlab = "Total length (cm.)", ylab = "Proportion mature", col = c("blue", "red"))
```

---

### plot.morphMat

*Plot method for morphMat class (size at morphometric maturity)*

#### Description

Plot method for morphMat class (size at morphometric maturity)

#### Usage

```r
## S3 method for class 'morphMat'
plot(
  x,
  xlab = "X",
  ylab = "Proportion mature",
  col = c("blue", "red"),
  lwd = 2,
  lty = 2,
  vline_hist = "black",
  lwd_hist = 2,
)```

Arguments

- **x**: object of class 'morphMat' with the mature parameters and a data.frame with the allometric variables ('X', 'Y') and classification of maturity. Also the fitted values for the logistic regression and confidence intervals (95%).

- **xlab**: a title for the x axis.

- **ylab**: a title for the y axis.

- **col**: color for the logistic curve and for the L50% size at morphometric maturity.

- **lwd**: line with for drawing fitted values and confidence intervals.

- **lty**: line type line type for drawing fitted values and confidence intervals

- **vline_hist**: color of the vertical lines in the histogram. The lines represent the the median and the confidence intervals.

- **lwd_hist**: line with for the vertical line in the histogram.

- **lty_hist**: line type for the vertical line in the histogram.

- **onlyOgive**: plot only the ogive.

- **...**: Additional arguments to the plot method.

Examples

```r
data(crabdata)

classify_data = classify_mature(crabdata, varNames = c("carapace_width", "chela_height"), varSex = "sex_category", selectSex = NULL, method = "ld")

my_mature = morph_mature(classify_data, method = "fq", niter = 50)

plot(my_mature, xlab = "Carapace width (mm.)", ylab = "Proportion mature", col = c("blue", "red"))
```

print.classify

Print method for classify class

Description

Print method for classify class

Usage

```r
## S3 method for class 'classify'
print(x, ...)
```
Arguments

\textbf{x} \hspace{1cm} \text{an object of class 'classify' with the allometric variables ("X", "Y") and classification of maturity (juveniles = 0, adults = 1).}

\ldots \hspace{1cm} \text{Additional arguments to the print method.}

Value

The number of juveniles and adults. Also shows the regression analysis for juveniles and adults and an ANCOVA analysis to compare slopes.

Examples

data(crabdata)

\texttt{classify\_data = classify\_mature(crabdata, varNames = c("carapace\_width", "chela\_height"), varSex = "sex\_category", selectSex = NULL, method = "ld")}

print(classify\_data)

\begin{verbatim}
print.gonadMat                  Print method for gonadMat class (size at gonad maturity)
\end{verbatim}

Description

Print method for gonadMat class (size at gonad maturity)

Usage

\#

\texttt{S3 method for class 'gonadMat'

\texttt{print(x, ...)}}

Arguments

\textbf{x} \hspace{1cm} \text{object of class 'gonadMat' with the parameters of the logistic regression and a data.frame with the X and stage of sexual maturity. variables. Also the fitted values for the logistic regression and confidence intervals (95%).}

\ldots \hspace{1cm} \text{Additional arguments to the print method.}

Value

The median of the size at gonad maturity estimation, parameters and the Nagelkerke's R squared.
print.morphMat

Examples

data(matFish)

gonad_mat = gonad_mature(matFish, varNames = c("total_length", "stage_mat"), innName = "I", matName = c("II", "III", "IV"), method = "fq", niter = 50)

print(gonad_mat)

print.morphMat

Print method for morphMat class (size at morphometric maturity)

Description

Print method for morphMat class (size at morphometric maturity)

Usage

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

Arguments

x object of class 'morphMat' with the parameters of the logistic regression and a data.frame with the allometric variables ("X", "Y") and classification of maturity. Also the fitted values for the logistic regression and confidence intervals (95%).

... Additional arguments to the print method.

Value

The median of the size at morphometric maturity estimation, parameters and the Nagelkerke’s R square.

Examples

data(crabdata)

classify_data = classify_mature(crabdata, varNames = c("carapace_width", "chela_height"), varSex = "sex_category", selectSex = NULL, method = "ld")

my_mature = morph_mature(classify_data, method = "fq", niter = 50)

print(my_mature)
Index

classify_mature, 4
Crab data, 5
crabdata (Crab data), 5

Fish maturity, 6
gonad_mature, 6

matFish (Fish maturity), 6
morph_mature, 8

nagelkerkeR2, 9

plotclassify, 10
plot.gonadMat, 11
plot.morphMat, 12
print.classify, 13
print.gonadMat, 14
print.morphMat, 15

sizeMat (sizeMat-package), 2
sizeMat-package, 2