Package ‘IDmeasurer’

May 9, 2019

Type    Package
Title   Assessment of Individual Identity in Animal Signals
Version 1.0.0
Maintainer Pavel Linhart <pavel.linhart@centrum.cz>
Description Provides tools for assessment and quantification of individual identity information in animal signals. This package accompanies a research article by Linhart et al. (2019) <doi:10.1101/546143>: ”Measuring individual identity information in animal signals: Overview and performance of available identity metrics”.
Depends R (>= 2.10)
License CC0
Encoding UTF-8
LazyData true
Imports MASS, infotheo, lme4
RoxygenNote 6.1.1
Suggests knitr, rmarkdown
VignetteBuilder knitr
NeedsCompilation no
Author Pavel Linhart [aut, cre]
Repository CRAN
Date/Publication 2019-05-09 15:10:10 UTC

R topics documented:

    ANmodulation .................................................. 2
    ANspec ............................................................. 3
    calcDistT ......................................................... 4
    calcDistW ......................................................... 5
    calcDS ............................................................ 6
    calcF .............................................................. 7
    calcHM ............................................................ 7
ANmodulation

Description

- **Species**: Little owl, *Athene noctua*
- **Number of individuals**: 33
- **Number of calls per individual**: 10
- **Number of acoustic variables**: 11
- **Individual identity**: HS=3.48

Calls of little owls were collected in Czech Republic and Hungary. Territorial calls of each male were recorded for three minutes after a short playback provocation (1 min) inside their territories from up to 50 m distance from the individuals. The recordings were made during comparable, favourable meteorological conditions (without strong wind or precipitation), from sunset until midnight between March and April of 2013 and 2014. This period covered the mating season. The period and the time of the day for recording were selected with regard to the peak in vocal activity of little owls both within a day and within a season. The original dataset included 54 males with more than 20 calls each (20-41 calls per individual, mean, SD = 26.9, 6.0) with good recording quality. The number of individuals and calls per individual was reduced to match parameters of the other datasets.
Eleven variables were measured to describe the modulation of the fundamental frequency within the call (shape of the call on spectrogram) and the duration of the call. Fundamental frequency was measured at ten measuring points spread equidistantly throughout the call duration. Variables were extracted in SASLab Pro by Avisoft.

**Usage**

`ANmodulation`

**Format**

A data frame with 330 rows and 12 variables:

- **id** factor, identity code of an individual emitting the call
- **f1, f3, f5, f7, f9, f11, f13, f15, f17, f19** fundamental frequency at the specific measuring point, in Hertz
- **dur** numeric, duration of the call, in seconds

**Source**


---

**Description**

- **Species**: Little owl, *Athene noctua*
- **Number of individuals**: 33
- **Number of calls per individual**: 10
- **Number of acoustic variables**: 7
- **Individual identity**: HS=4.68

Calls of little owls were collected in Czech Republic and Hungary. Territorial calls of each male were recorded for three minutes after a short playback provocation (1 min) inside their territories from up to 50 m distance from the individuals. The recordings were made during comparable, favourable meteorological conditions (without strong wind or precipitation), from sunset until midnight between March and April of 2013 and 2014. This period covered the mating season. The period and the time of the day for recording were selected with regard to the peak in vocal activity of little owls both within a day and within a season. The original dataset included 54 males with more than 20 calls each (20-41 calls per individual, mean, SD = 26.9, 6.0) with good recording quality. The number of individuals and calls per individual was reduced to match parameters of the
Variables were selected to measure basic parameters of call spectrum like the peak frequency, distribution of frequency amplitudes within spectrum, and range of the frequencies (minimum and maximum). Additionally, the duration of the call was measured. Variables were extracted in SASLab Pro by Avisoft.

Usage

A data frame with 330 rows and 8 variables:

- id: factor, identity code of an individual emitting the call
- dur: duration of the call, in seconds
- df: frequency of maximum amplitude within the spectrum - peak frequency, in Hertz
- minf, maxf: minimum and maximum frequency at -25dB relative to the call peak amplitude, in Hertz
- q25, q50, q75: frequencies at the three quartiles of amplitude distribution; frequencies below which lie 25, 50 and 75 percent of the energy of the call, respectively, in Hertz

Source


---

calcDistT

Calculate total distance in given dataset

Description

This function calculates centroid of all samples in a given dataset and sums distances between the centroid and each sample. Euclidean distances are used.

Usage

calcDistT(df)

Arguments

- df: A data frame with the first column indicating individual identity.

Value

Numeric. Total distance.
CalcDistW

See Also

Other calcHM support function: calcDistW, calcMeanVec

Examples

calcDistW(ANmodulation)

---

CalcDistW  
*Calculate average within individual distance*

Description

This function calculates centroid for each individual and sums distances of samples from centroid for that particular individual. When the within individual sum of distances is known for each individual, it calculates their mean. Euclidean distances are used.

Usage

calcDistW(df)

Arguments

df  
A data frame with the first column indicating individual identity.

Value

Numeric. Average within individual distance in dataset.

See Also

Other calcHM support function: calcDistT, calcMeanVec

Examples

calcDistW(ANmodulation)
Description

This function will take the specified data frame and will perform linear discrimination analysis with leave-one-out crossvalidation and equal priors for each individual (e.g., all priors will be set to 1/10 in case that the dataset contains 10 individuals). Variables are not modified in any way; scaling, centering, transformation of variables, or principal component analysis, etc., if required, need to be done on dataset before calling this function.


Usage

`calcDS(df)`

Arguments

df

A data frame with the first column noting individual identity of the sample.

Value

Proportion of samples correctly classified by LDA in df.

See Also

Other individual identity metrics: `calcF`, `calcHM`, `calcHSngroups`, `calcHSnpergroup`, `calcHSntot`, `calcHSvarcomp`, `calcHS`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweentot`, `calcPIC`

Examples

`calcDS(ANmodulation)`
calcF

**Calculate F-values for individual identity traits**

**Description**

This function calculates ANOVA F-values (type I sum of squares) for all identity traits in dataset along with its significance. Each trait is used as dependent and identity code as independent variable.


**Usage**

```r
calcF(df)
```

**Arguments**

- `df` A data frame with the first column indicating individual identity.

**Value**

A data frame with 11 rows and 3 columns (trait, f-value, and p-value).

**See Also**

Other individual identity metrics: `calcDS`, `calcHM`, `calcHSngroups`, `calcHSnpergroup`, `calcHSntot`, `calcHSvarcomp`, `calcHS`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweenot`, `calcPIC`

**Examples**

```r
calcF(ANmodulation)
```

---

calchM

**Calculate information capacity (HM)**

**Description**

This function calculates information capacity of a signal.

Usage

calchm(df)

Arguments

df A data frame with the first column indicating individual identity.

Value

Numeric value. Individual identity information capacity HM (in bits) in dataset.

See Also

Other individual identity metrics: \texttt{calcDS}, \texttt{calcF}, \texttt{calchsngroups}, \texttt{calchspergroup}, \texttt{calchsntot}, \texttt{calchsvarcomp}, \texttt{calchS}, \texttt{calcMI}, \texttt{calcPICbetweenmeans}, \texttt{calcPICbetweenot}, \texttt{calcPIC}

Examples

calchm(ANmodulation)

\begin{verbatim}

| calchS | Calculate Beecher's information statistic (HS, variant = HSnper-group) |

\end{verbatim}

Description

This function calculates Beecher’s information statistic (HS) for all variables within the dataset.


calchS (equivalent to \texttt{calchspergroup}) is the correct variant of the function calculating Beecher’s information statistic. The other variants use total sample size (\texttt{calchsntot}) or number of individuals in dataset (\texttt{calchsngroups}) instead of number of samples per individual to calculate HS. \texttt{calchsvarcomp} calculates HS from variance components of mixed models. HS values calculated by \texttt{calchsvarcomp} were found to be twice as large compared to HS calculated by standard approach.

Please note, \texttt{sumHS = TRUE} should be used in datasets where individuality traits are uncorrelated. If traits are correlated, Principal component analysis (PCA) should be applied and HS should be calculated on uncorrelated principal components instead of original trait variables.

Usage

calchS(df, sumHS = TRUE)
Arguments

- **df**: A data frame with the first column indicating individual identity.
- **sumHS**: `sumHS = TRUE` (default) will sum partial HS values of each trait variable; `sumHS = FALSE` provides partial HS values separately for each individuality trait in a dataset.

Value

For `sumHS = TRUE`: Numeric vector of two elements indicating: 1) HS summed over variables that significantly differ between individuals (in one-way Anova with individual as independent and a specific signal trait as dependent variable; or 2) HS summed over all variables in dataset.

For `sumHS = FALSE`: Data frame with three columns and number of rows equal to number of variables in dataset. First column includes names of traits considered for individuality. Second column includes significance test for each trait (from one-way ANOVA with individual identity as independent factor and trait as dependent variable). Third column includes values of HS for each variable trait.

See Also

Other individual identity metrics: `calcDS`, `calcF`, `calcHM`, `calcHSngroups`, `calcHSngroups`, `calcHSnpergroup`, `calcHSntot`, `calcHSvarcomp`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweentot`, `calcPIC`

Examples

```r
calcHS(ANmodulation)
temp <- calcPCA(ANmodulation)
calcHS(temp)
```

---

**calcHSngroups**: Calculate Beecher’s information statistic (HS, variant = HSngroups)

Description

This function calculates Beecher’s information statistic (HS) for all variables within the dataset.


`calcHS` (equivalent to `calcHSngroups`) is the correct variant of the function calculating Beecher’s information statistic. The other variants use total sample size (`calcHSntot`) or number of individuals in dataset (`calcHSngroups`) instead of number of samples per individual to calculate HS. `calcHSvarcomp` calculates HS from variance components of mixed models. HS values calculated by `calcHSvarcomp` were found to be twice as large compared to HS calculated by standard approach.
Please note, `sumHS = TRUE` should be used in datasets where individuality traits are uncorrelated. If traits are correlated, Principal component analysis (PCA) should be applied and HS should be calculated on uncorrelated principal components instead of original trait variables.

**Usage**

```
calcHSngroups(df, sumHS = T)
```

**Arguments**

- `df`: A data frame with the first column indicating individual identity.
- `sumHS`: `sumHS = TRUE` (default) will sum partial HS values of each trait variable; `sumHS = FALSE` provides partial HS values separately for each individuality trait in a dataset.

**Value**

For `sumHS = TRUE`: Numeric vector of two elements indicating indicating: 1) HS summed over variables that significantly differ between individuals (in one-way Anova with individual as independent and a specific signal trait as dependent variable; or 2) HS summed over all variables in dataset.

For `sumHS = FALSE`: Data frame with thre columns and number of rows equal to number of variables in dataset. First column includes names of traits considered for individuality. Second column includes significance test for each trait (from one-way ANOVA with individual identity as independent factor and trait as dependent variable). Third column includes values of HS for each variable trait.

**See Also**

Other individual identity metrics: `calcDS, calcF, calcHM, calcHSngroup, calcHSntot, calcHSvarcomp, calcHS, calcMI, calcPICbetweenmeans, calcPICbetweenot, calcPIC`

**Examples**

```
calcHSngroups(ANmodulation)
temp <- calcPCA(ANmodulation)
calcHSngroups(temp)
```

---

| `calcHSngroup` | Calculate Beecher’s information statistic (HS, variant = HSnper-group) |
Description

This function calculates Beecher’s information statistic (HS) for all variables within the dataset.


calchSnpergroup is the correct variant of the function calculating Beecher’s information statistic. The other variants use total sample size (calchSnntot) or number of individuals in dataset (calchSngrps) instead of number of samples per individual to calculate HS. calchSvarcomp calculates HS from variance components of mixed models. HS values calculated by calchSvarcomp were found to be twice as large compared to HS calculated by standard approach.

Please note, sumHS = TRUE should be used in datasets where individuality traits are uncorrelated. If traits are correlated, Principal component analysis (PCA) should be applied and HS should be calculated on uncorrelated principal components instead of original trait variables.

Usage

calchSnpergroup(df, sumHS = TRUE)

Arguments

df A data frame with the first column indicating individual identity.
sumHS sumHS = TRUE (default) will sum partial HS values of each trait variable; sumHS = FALSE provides partial HS values separately for each individuality trait in a dataset.

Value

For sumHS = TRUE: Numeric vector of two elements indicating: 1) HS summed over variables that significantly differ between individuals (in one-way Anova with individual as independent and a specific signal trait as dependent variable; or 2) HS summed over all variables in dataset.

For sumHS = FALSE: Data frame with three columns and number of rows equal to number of variables in dataset. First column includes names of traits considered for individuality. Second column includes significance test for each trait (from one-way ANOVA with individual identity as independent factor and trait as dependent variable). Third column includes values of HS for each variable trait.

See Also

calcPIC, calchS

Other individual identity metrics: calcDS, calcF, calchM, calchSngrps, calchSnntot, calchSvarcomp, calchS, calcMI, calcPICbetweenmeans, calcPICbetweenot, calcPIC
Examples

calchSnperegroup(ANmodulation)
temp <- calcPCA(ANmodulation)
calchSnperegroup(temp)

calchSntot

Calculate Beecher’s information statistic (HS, variant = HSntot)

Description

This function calculates Beecher’s information statistic (HS) for all variables within the dataset.


calchS (equivalent to calchSnperegroup) is the correct variant of the function calculating Beecher’s information statistic. The other variants use total sample size (calchSntot) or number of individuals in dataset (calchSngroups) instead of number of samples per individual to calculate HS. calchSvarcomp calculates HS from variance components of mixed models. HS values calculated by calchSvarcomp were found to be twice as large compared to HS calculated by standard approach.

Please note, sumhs = TRUE should be used in datasets where individuality traits are uncorrelated. If traits are correlated, Principal component analysis (PCA) should be applied and HS should be calculated on uncorrelated principal components instead of original trait variables.

Usage

calchSntot(df, sumhs = T)

Arguments

df A data frame with the first column indicating individual identity.
sumhs sumhs = TRUE (default) will sum partial HS values of each trait variable; sumhs = FALSE provides partial HS values separately for each individuality trait in a dataset.

Value

For sumhs = TRUE: Numeric vector of two elements indicating indicating: 1) HS summed over variables that significantly differ between individuals (in one-way Anova with individual as independent and a specific signal trait as dependent variable; or 2) HS summed over all variables in dataset.

For sumhs = FALSE: Data frame with thre columns and number of rows equal to number of variables in dataset. First column includes names of traits considered for individuality. Second column
Calculation of Beecher's information statistic (HS) for all variables within the dataset. Includes significance test for each trait (from one-way ANOVA with individual identity as independent factor and trait as dependent variable). Third column includes values of HS for each variable trait.

See Also

- `calcPIC`, `calcHS`
- Other individual identity metrics: `calcDS`, `calcF`, `calcHM`, `calcHSngroups`, `calcHSnpergroup`, `calcHSvarcomp`, `calcHS`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweenot`, `calcPIC`

Examples

```r
calcHSntot(ANmodulation)
temp <- calcPCA(ANmodulation)
calcHSntot(temp)
```

---

**Description**

This function calculates Beecher's information statistic (HS) for all variables within the dataset.


`calcHS` (equivalent to `calcHSnpergroup`) is the correct variant of the function calculating Beecher's information statistic. The other variants use total sample size (`calcHSntot`) or number of individuals in dataset (`calcHSngroups`) instead of number of samples per individual to calculate HS. `calcHSvarcomp` calculates HS from variance components of mixed models. HS values calculated by `calcHSvarcomp` were found to be twice as large compared to HS calculated by standard approach.

Please note, `sumHS = TRUE` should be used in datasets where individuality traits are uncorrelated. If traits are correlated, Principal component analysis (PCA) should be applied and HS should be calculated on uncorrelated principal components instead of original trait variables.

**Usage**

```r
calcHSvarcomp(df, sumHS = T)
```

**Arguments**

- `df` A data frame with the first column indicating individual identity.
- `sumHS` `sumHS = TRUE` (default) will sum partial HS values of each trait variable; `sumHS = FALSE` provides partial HS values separately for each individuality trait in a dataset.
Value

For `sumHS = TRUE`: Numeric vector of two elements indicating: 1) HS summed over variables that significantly differ between individuals (in one-way Anova with individual as independent and a specific signal trait as dependent variable); or 2) HS summed over all variables in dataset.

For `sumHS = FALSE`: Data frame with three columns and number of rows equal to number of variables in dataset. First column includes names of traits considered for individuality. Second column includes significance test for each trait (from one-way ANOVA with individual identity as independent factor and trait as dependent variable). Third column includes values of HS for each variable trait.

See Also

calcPIC, calcHS

Other individual identity metrics: calcDS, calcF, calcHM, calcHSngroups, calcHSngroups, calcHSntot, calcHS, calcMI, calcPICbetweenmeans, calcPICbetween, calcPIC

Examples

calcHSvarcomp(ANmodulation)
temp <- calcPCA(ANmodulation)
calcHSvarcomp(temp)

calcMeanVec

*Calculate the centroid of the individual identity traits*

Description

This function calculates centroid of the individual identity traits. Euclidean distances are used.

Usage

`calcMeanVec(df)`

Arguments

`df` A data frame with the individual identity traits without identity codes (NA will be produced for the column with identity code).

Value

Numeric vector with the centroid values for each trait.

See Also

Other calcHM support function: calcDistT, calcDistW
calcMI

Examples

#incorrect use (with identity codes, NA will be produced):
calcMeanVec(ANmodulation)

#correct use (with identity codes removed)
calcMeanVec(ANmodulation[-1])

calcMI

Calculate Mutual information (MI)

Description

This function calculates Mutual information of actual and predicted individual identities in the dataset. It uses Linear discrimination analysis (MASS::lda) to predict individual identity. Settings for LDA are identical to those used in calcDS function, i.e., LDA uses leave-one-out cross-validation and priors are equal for each individual in dataset.


Usage

calcMI(df)

Arguments

df A data frame with the first column indicating individual identity.

Value

Numeric value of mutual information (in bits).

See Also

calcPIC, calcHS

Other individual identity metrics: calcDS, calcF, calcHM, calcHSngroups, calcHSnpergroup, calcHSntot, calcHSvarcomp, calcHS, calcPICbetweenmeans, calcPICbetweentot, calcPIC

Examples

calcMI(ANmodulation)
### calcPCA

**Convert raw trait variables into principal components**

**Description**

This function subjects the trait variables from the original dataset to the Principal component analysis (PCA, *stats::prcomp*) and calculates principal components scores for each sample. All variables are centered by subtracting the variable mean from a particular value and scaled to the unit variance by dividing the value by the standard deviation of a trait (*stats::prcomp* parameters `center = T, scale = T`). Some functions like, for example, `calchS` require uncorrelated input variables to calculate individual identity information properly.

**Usage**

```r
calcPCA(df)
```

**Arguments**

- `df` A data frame with the first column indicating individual identity.

**Value**

- `df` A data frame with the same attributes like the `df`, but the original individuality traits are replaced by principal components.

**Examples**

```r
summary(ANmodulation)
temp <- calcPIC(ANmodulation)
summary(temp)
```

### calcPIC

**Calculates potential of identity coding (PIC, variant=PICbetweenot)**

**Description**

This function calculates Potential of individual coding for all variables within the dataset.


**Usage**

```r
calcPIC(df)
```
calcPICbetweenmeans

Arguments

- **df**: A data frame with the first column indicating individual identity.

Value

Numeric vector with PIC values for each variable in df.

See Also

Other individual identity metrics: `calcDS`, `calcF`, `calcHM`, `calcHSGroups`, `calcHSNpergroup`, `calcHSntot`, `calcHSvarcomp`, `calcHS`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweentot`

Examples

```r
calcPIC(ANmodulation)
```

---

**Description**

This function calculates Potential of individual coding for all variables within the dataset.


**Usage**

```r
calcPICbetweenmeans(df)
```

Arguments

- **df**: A data frame with the first column indicating individual identity.

Value

Numeric vector with PIC values for each variable in df.

See Also

Other individual identity metrics: `calcDS`, `calcF`, `calcHM`, `calcHSGroups`, `calcHSNpergroup`, `calcHSntot`, `calcHSvarcomp`, `calcHS`, `calcMI`, `calcPICbetweenmeans`, `calcPICbetweentot`, `calcPIC`
Examples

```r
calcPICbetweenmeans(ANmodulation)
```

---

`calcPICbetweenot` Calculates potential of identity coding (PIC, variant= PICbetweenot)

### Description

This function calculates Potential of individual coding for all variables within the dataset.


### Usage

```r
calcPICbetweenot(df)
```

### Arguments

- `df` A data frame with the first column indicating individual identity.

### Value

Numeric vector with PIC values for each variable in `df`.

### See Also

Other individual identity metrics: `calcDS, calcF, calcHM, calcHSgroups, calcHSnpergroup, calcHSntot, calcHSvarcomp, calcHS, calcMI, calcPICbetweenmeans, calcPIC`

### Examples

```r
calcPICbetweenot(ANmodulation)
```
Description

- **Species:** Corncrake, *Crex crex*
- **Number of individuals:** 33
- **Number of calls per individual:** 10
- **Number of acoustic variables:** 4
- **Individual identity:** HS=10.51


Corncrake calls were recorded at three sites in Poland and one in the Czech Republic. Recordings were made during the corncrake breeding season, from 8 to 30 July, in 2011 and in 2012. Males were recorded when calling spontaneously, in favourable conditions, at night (from 22.00 to 03.30, local time) from a distance of ca. 5-10 m. The original dataset comprised 104 males with 10 calls measured from each male.

Formants (second to fifth) were measured within the first syllable of the call. Formants were extracted by PRAAT.

Usage

CCformants

Format

A data frame with 330 rows and 5 variables:

- **id** factor, identity code of an individual emitting the call
- **s1f2, s1f3, s1f4, s1f5** formants 2 to 5, respectively, measured within the first syllable of the call, in Hertz

Source

Description

- **Species:** Corncrake, *Crex crex*
- **Number of individuals:** 33
- **Number of calls per individual:** 10
- **Number of acoustic variables:** 7
- **Individual identity:** HS=5.68
- **Reference:** Budka, M., & Osiejuk, T. S. (2013). Formant Frequencies are Acoustic Cues to Caller Discrimination and are a Weak Indicator of the Body Size of Corncrake Males. Ethology, 119, 960-969. doi:10.1111/eth.12141

Corncrake calls were recorded at three sites in Poland and one in the Czech Republic. Recordings were made during the corncrake breeding season, from 8 to 30 July, in 2011 and in 2012. Males were recorded when calling spontaneously, in favourable conditions, at night (from 22.00 to 03.30, local time) from a distance of ca. 5-10 m. The original dataset comprised 104 males with 10 calls measured from each male.

Seven variables were selected to measure duration of the first syllable of the call and its basic spectral parameters of each first syllable of the call like the peak frequency, distribution of frequency amplitudes within spectrum, and range of the frequencies (minimum and maximum). Additionally, the duration of the call was measured. Variables were extracted in SASLab Pro by Avisoft.

Usage

CCspec

Format

A data frame with 330 rows and 8 variables:

- **id** factor, identity code of an individual emitting the call
- **dur** duration of the call, in seconds
- **df** frequency of maximum amplitude within the spectrum - peak frequency, in Hertz
- **minf, maxf** minimum and maximum frequency at -25dB relative to the call peak amplitude, in Hertz
- **q25, q50, q75** frequencies at the three quartiles of amplitude distribution; frequencies below which lie 25, 50 and 75 percent of the energy of the call, respectively, in Hertz

Source

Budka, M., & Osiejuk, T. S. (2013). Formant Frequencies are Acoustic Cues to Caller Discrimination and are a Weak Indicator of the Body Size of Corncrake Males. Ethology, 119, 960-969. doi:10.1111/eth.12141
**Describe**

This function converts DS to HS. Because the model used to convert values is derived from `stats::loess` model, it cannot make predictions outside the range of the values used to construct the model. The model was tailored to the sampling frequently used by studies on individuality:
- number of individuals should be between 5 and 40 individuals
- number of observations per individual should be between 5 - 20
- DS value should be between 0 - 1

Consider increasing your sampling if number of individuals and/or number of observations are lower than 5. If number of individuals and observations exceeds the function limits, it might be acceptable to use the largest possible values allowed by the model to get the estimate - biases get smaller with larger sampling, so, if your sampling exceeds function limits a little bit, the estimate should be still quite precise.

**Usage**

`convertdstohs(nindivs, nobs, DS, se = FALSE)`

**Arguments**

- `nindivs` Number of individuals. Must be within 5-40 individuals.
- `nobs` Number of observations per individual. Must be within 5-20 observations.
- `DS` DS value to be converted into HS. Must be within 0-1.
- `se` should standard errors be computed (takes more time)?

**Value**

If `se = FALSE`, Numeric value. DS for a specified number of individuals and number of observations per individual.
If `se = TRUE`, a list containing components fit, se, residual.scale, df. See `predict.loess` for more details.

**See Also**

Other metric conversion: `converthstods`

**Examples**

`convertdstohs(nindivs=10, nobs=10, DS=0.7)`
Description

This function converts HS to DS. Because the model used to convert values is derived from stats::loess model model, it cannot make predictions outside the range of the values used to construct the model. Our model was tailored to the values frequently used by studies on individuality:
- number of individuals should be between 5 and 40 individuals
- number of observations per individual should be between 5 - 20
- HS value should be between 0 - 12.9

Consider increasing your sampling if number of individuals and number of observations are lower than 5. If number of individuals and observations exceeds the function limits, it might be possible to use the largest possible values allowed by the model to get the estimate - biases get smaller with large sampling, so, if your sampling exceeds function limits a little bit, the estimate should be still quite precise.

Usage

converthstods(nindivs, nobs, HS, se = FALSE)

Arguments

nindivs    Number of individuals. Must be within 5-40 individuals.
nobs   Number of observations per individual. Must be within 5-20 observations.
HS    HS value to be converted into DS. Must be within 0-12.9 bits.
se    should standard errors be computed (takes more time)?

Value

If se = FALSE, Numeric value. DS for a specified number of individuals and number of observations per individual.
If se = TRUE, a list containing components fit, se, residual.scale, df. See predict.loess for more details.

See Also

Other metric conversion: convertdStoHS

Examples

converthstods(nindivs=10, nobs=10, HS=5)
GenerateMultivariate  Generate dataset with multiple individual identity traits

Description

This functions generates a dataset with desired parameters (number of individuals and number of observations per individual, number of variables and covariance between variables, individuality). Unlike for the function GenerateUnivariate, trait means are not customizable and are always set to 0.

Usage

GenerateMultivariate(nindivs, nobs, nvars, covar, individuality)

Arguments

nindivs  Indicates how many individuals should be in dataset
nobs  Indicates how many observations per individual should be in dataset
nvars  Indicates how many trait variables should be in dataset.
covar  Indicates covariance between variables in dataset. covar=0 for uncorrelated variables; covar=1 for fully correlated variables
individuality  Indicates the ratio of between to within individual variation in each trait variable.

Value

Data frame with the identity codes in the first column and trait variables in subsequent columns. Number of rows and columns depends on the parameters used to generate dataset.

See Also

Other Operations with datasets: GenerateUnivariate

Examples

id1 <- GenerateMultivariate(nindivs=10, nobs=10, nvars=2, covar=0, individuality=1)
GenerateUnivariate

Generate dataset with a single individual identity trait

Description

This function generates a dataset with desired parameters (number of individuals and number of observations per individual, mean of the parameter, individuality).

Usage

GenerateUnivariate(nindivs, nobs, betweenM, individuality)

Arguments

nindivs  Indicates how many individuals should be in dataset
nobs     Indicates how many observations per individual should be in dataset
betweenM Indicates the mean value of the trait.
individuality Indicates the ratio of between to within individual variation.

Value

Data frame with two columns. Identity codes are in the first column and the trait values are in the second column.

See Also

Other Operations with datasets: GenerateMultivariate

Examples

GenerateUnivariate(nindivs=10, nobs=10, betweenM=1000, individuality=2)
Description

- **Species**: Yellow-breasted boubou, *Laniarius atroflavus*
- **Number of individuals**: 33
- **Number of calls per individual**: 10
- **Number of acoustic variables**: 6
- **Individual identity**: HS=3.83
- **Reference**: Osiejuk, unpublished data

Male Yellow-breasted boubous were recorded in Bamenda region in Cameroon. Birds were recorded between 06.00 to 10.00 in the morning in 2016, typically, from the distance of 10 - 20 meters. The calls were recorded after short provocation with playback. Repertoire of males at the field site included three distinct call types and only the most common call typed labeled as “High wee woo” was used for this study. The original dataset comprised 33 individuals and 10 calls per individual.

Variables were selected to measure basic spectral parameters of each “high weewoo” call like the peak frequency, distribution of frequency amplitudes within spectrum, and range of the frequencies (minimum and maximum). Additionally, the duration of the call was measured. Variables were extracted in Raven Pro 1.5 by the Cornell Bioacoustic Research Program.

Usage

`LAhighweewoo`
Format

A data frame with 330 rows and 7 variables:

- **id** factor, identity code of an individual emitting the call
- **dur** duration of the call, in seconds
- **df** frequency of maximum amplitude within the spectrum - peak frequency, in Hertz
- **minf, maxf** minimum and maximum frequency at -25dB relative to the call peak amplitude, in Hertz
- **q25, q75** frequencies at the two quartiles of amplitude distribution; frequencies below which lie 25 and 75 percent of the energy of the call, respectively, in Hertz

Source

Osiejuk, unpublished data

SSgrunts

*Domestic pig, Sus scrofa domestica - piglet grunts*

Description

- **Species**: Domestic pig, *Sus scrofa domestica*
- **Number of individuals**: 33
- **Number of calls per individual**: 10
- **Number of acoustic variables**: 10
- **Individual identity**: HS=3.18

Piglet grunts were recorded in 2014 and 2015 at the research farm of Institute of Animal Science in Prague. To elicit the calls piglets were separated from the litter and sow and were recorded in social isolation in age of 25-30 days after birth from the distance of 1m. Piglets were recorded opportunistically during the day. The original dataset comprised 97 piglets coming from 13 different litters and 10 calls per individual piglet.

Variables were selected to be the most informative regarding the individual identity by the reference study. Variables were extracted using the LMA 2008 software for analysis of animal sounds by Kurt Hammerschmidt.

Usage

SSgrunts
SSgrunts

Format

A data frame with 330 rows and 8 variables:

- **id**: factor, identity code of an individual emitting the call
- **fp1amean**: mean amplitude of the 1st frequency peak, relative amplitude
- **q1mean**: mean frequency of the 1st DFreqA; equivalent to q25 in ANspec, CCspec, and LAhigh-weewoo, in Hertz
- **csmean**: mean correlation coefficient of successive time segments
- **tonalF0**: estimation of segments with detectable fundamental frequency, in percents
- **q3mean**: mean frequency of the 3nd DFreqA; equivalent to q75 in ANspec, CCspec, and LAhigh-weewoo, in Hertz
- **noise**: percentage of noise time segments, in percents
- **ranmean**: mean frequency range, in Hertz
- **q2mean**: mean frequency of the 2nd DFreqA; equivalent to q50 in ANspec, CCspec, and LAhigh-weewoo, in Hertz
- **df2pr**: percentage of time segments where 2nd DomFreqB detected, in percents
- **q2min**: min frequency of the 2nd DFreqA, in Hertz

Source

Index

*Topic **datasets**
  - ANmodulation, 2
  - ANspec, 3
  - CCformants, 19
  - CCSpec, 20
  - LAhighweewoo, 25
  - SSgrunts, 26

ANmodulation, 2
ANspec, 3
calcDistT, 4, 5, 14
calcDistW, 5, 5, 14
calcDS, 6, 7–11, 13–15, 17, 18
calcF, 6, 7, 8–11, 13–15, 17, 18
calchM, 6, 7, 7, 9–11, 13–15, 17, 18
calchS, 6–8, 8, 10, 11, 13–15, 17, 18
calchSnsgroups, 6–9, 9, 11, 13–15, 17, 18
calchSnpergroup, 6–10, 10, 13–15, 17, 18
calchSntot, 6–11, 12, 14, 15, 17, 18
calchSvarcomp, 6–11, 13, 13, 15, 17, 18
calcMeanVec, 5, 14
calcMI, 6–11, 13, 14, 15, 17, 18
calcPCA, 16
calcPIC, 6–11, 13–15, 16, 17, 18
calcPICbetweenmeans, 6–11, 13–15, 17, 17, 18
calcPICbetweenentot, 6–11, 13–15, 17, 18
CCformants, 19
CCspec, 20
convertDSetoHS, 21, 22
convertHStoDS, 21, 22
GenerateMultivariate, 23, 24
GenerateUnivariate, 23, 24
IDmeasurer, 24
IDmeasurer-package (IDmeasurer), 24
LAhighweewoo, 25
SSgrunts, 26