

# Package ‘ZLAvian’

May 15, 2024

**Type** Package

**Title** Zipf's Law of Abbreviation in Animal Vocalisations

**Version** 0.2.0

**Description** Assesses evidence for Zipf's Law of Abbreviation in animal vocalisation using IDs, note class and note duration. The package also provides a web plot function for visualisation.

**License** CC BY-SA 4.0

**URL** <https://github.com/CDr-er/ZLAvian>

**BugReports** <https://github.com/CDr-er/ZLAvian/issues>

**Maintainer** CD Durrant <cbdurrant@gmail.com>

**Depends** R(>= 4.3.0)

**Imports** lme4 (>= 1.1-34), performance (>= 0.10.4), doParallel (>= 1.0.17)

**Suggests** rmarkdown, knitr

**VignetteBuilder** knitr

**Encoding** UTF-8

**Language** en-GB

**LazyData** true

**RoxygenNote** 7.3.1

**NeedsCompilation** no

**Author** CD Durrant [com, cre, cph],  
Tucker Gilman [aut, cph]

**Repository** CRAN

**Date/Publication** 2024-05-15 11:00:03 UTC

## R topics documented:

Java.sparrow.notes . . . . .	2
plotZLA . . . . .	3
testZLA . . . . .	4

**Index**

7

---

Java.sparrow.notes      *Java sparrow note duration and frequency.*

---

**Description**

This dataset reports the durations and frequencies of use for 22,970 notes in 676 undirected songs produced by 73 Java sparrows (*Padda oryzivora*). Songs were recorded by Masayo Soma and colleagues and notes were assigned to types by Rebecca Lewis (Lewis et al. 2021; Lewis et al. 2023).

**Usage**

```
data("Java.sparrow.notes")
```

**Format**

A list with 22,970 observations of the following 3 variables.

duration duration, in ms, of each note in the dataset

note note type to which each note was assigned

ID identifier of the individual bird that produced each note

**Source**

<[https://figshare.manchester.ac.uk/articles/dataset/Code\\_and\\_Data\\_from\\_Lewis\\_et\\_al\\_2023\\_Animal\\_Behavior/22550149](https://figshare.manchester.ac.uk/articles/dataset/Code_and_Data_from_Lewis_et_al_2023_Animal_Behavior/22550149)>

**References**

Lewis, R.N., Soma, M., de Kort, S.R., Gilman, R.T. (2021) Like father like son: cultural and genetic contributions to song inheritance in an Estrildid finch. *Frontiers in Psychology* 12, 654198. (doi.org/10.3389/fpsyg.2021.654198)

Lewis, R.N., Kwong, A., Soma, M., de Kort, S.R., Gilman, R.T. (2023) Inheritance of temporal song features in Java sparrows. *Animal Behaviour* 206, 61-74. (doi.org/10.1016/j.anbehav.2023.09.012)

**Examples**

```
head(Java.sparrow.notes)
```

**Description**

Produces a web plot. Requires data from the testZLA function.

**Usage**

```
plotZLA(  
  inputObject,  
  title = "",  
  ylab = "duration (s)",  
  x.scale = "log",  
  y.base = 2  
)
```

**Arguments**

inputObject	output from the function testZLA
title	a title to be printed over the web plot
ylab	a label for the y-axis
x.scale	takes values "log" or "linear." Sets scaling for the x-axis.
y.base	takes values 2 or 10. Controls tick mark positions on the y-axis. When set to 2, tick marks indicate that values differ by a factor of 2. When set to 10, tick marks indicate that values differ by a factor of 10. If data was not log transformed in the analysis being illustrated, then the y-axis is linear and this argument is ignored.

**Details**

In the figure produced by plotZLA, each point represents a note or phrase type in the population repertoire. Note types are joined by a line if at least one individual produces both note types. The weight of the line is proportional to the number of individuals that produce both note types. The color of the lines indicates whether there is a positive (blue) or negative (red) concordance between the duration and frequency of use of the note types. Negative concordances are consistent with Zipf's law of abbreviation. Shades between blue and red indicate that the concordance is positive in some individuals and negative in others. For example, this can happen if some individuals use the note types more frequently than others, such that the rank order of frequency of use varies among individuals. Grey crosses centered on each point show the longest and shortest durations of the note type (vertical) and the highest and lowest frequencies of use (horizontal) in the population.

**Value**

Produces a web plot illustrating concordance between note duration and frequency of use within individuals. Requires output from the testZLA function.

**Author(s)**

CD Durrant and R. Tucker Gilman (2024)

**References**

Davis, M. K. and Chen, G. (2007) Graphic Kendall's tau. *Computational Statistics & Data Analysis*, 51(5), 2373-2378. (doi.org/10.2307/2346786)

Gilman, R. T., Durrant, C. D., Malpas, L., and Lewis, R. N. (2023) Does Zipf's law of abbreviation shape birdsong? *bioRxiv* (doi.org/10.1101/2023.12.06.569773)

**Examples**

```
#Test for evidence of ZLA in the songs of 73 Java sparrows
#and produce a web plot to illustrate the result. In testZLA,
#"null" is set to the minimum value to make the example run
#quickly, and thus the p-values reported are likely to be
#imprecise. This will not affect the web plot, because plotZLA
#does not rely on accurate p-values or null distributions.

ZLA.object=testZLA(Java.sparrow.notes, null = 99)

plotZLA(ZLA.object, ylab = "duration (ms)", x.scale = "linear")
```

---

testZLA

*Assess evidence for Zipf's law of abbreviation*

---

**Description**

Assesses evidence for Zipf's Law of Abbreviation in a population where samples from the population repertoire can be assigned to individuals.

**Usage**

```
testZLA(
  data,
  minimum = 1,
  null = 999,
  est = "mean",
  cores = 2,
  transform = "log"
)
```

**Arguments**

**data** a dataframe containing columns "note" (factor/character; identifies the note/phrase type of each token), "duration" (numeric; describes the duration of each token), and "ID" (factor; identifies the individual that produced each token). Other columns in the dataframe are ignored.

minimum	the minimum number of times a note type must appear in the data set to be included in the analysis. Must be a positive integer.
null	the number of permutations used to estimate the null distribution. Must be a positive integer 99 or greater.
est	takes values "mixed," "mean," or "median." If est = "mixed" then the expected logged duration for each note type in the population is computed as the intercept from an intercept-only mixed effects model, fit using the lmer() function of lme4 (Bates et al. 2015), that includes a random effect of individual ID. This computes a weighted mean across birds, and accords greater weights to birds that produce the note type more frequently. If est = "mean" then the expected logged duration for each note type in the population is computed as the mean of the means for the individual birds, with each individual bird weighted equally. This is faster than the "mixed" method. If est = "median" then the expected logged duration for each note type within birds is taken to be the median logged duration of the note type when produced by that bird, and the expected logged duration for each note type in the population is taken to be the median of the medians for the birds that produced that note type. The expected durations for note types are used in the permutation algorithm.
cores	divides (parallelizes) computation of the null distribution among cores. Cores must be an integer between 1 and the number of cores available on the users machine, inclusive.
transform	takes values "log" or "none." Indicates how duration data should be transformed prior to analysis. Gilman and colleagues (2023) argue that log transformation may often be appropriate for duration data, but some other measures might be better analysed as raw (untransformed) values.

### Value

a list with components:

overview	reports summary statistics for the dataset. These are "total notes," the total number of notes in the dataset; "total individuals," the number of individual birds in the dataset; "total note classes," the total number of note types represented in the dataset; "notes per individual," the average number of notes produced by individuals in the dataset; and "classes per individual," the average number of note classes that each individual produces.
shannon	the Shannon diversity of note types in the full dataset, and the mean Shannon diversity of note types produced by individuals in the dataset.
stats	a matrix reporting "individual level tau," the weighted mean concordance between note duration and frequency of use of notes within individuals, computed over all individuals in the dataset, with the concordance within each individual weighted by its inverse variance; "population level tau," the concordance between note duration and frequency of use in the full dataset, independent of the individuals that produced each note; and the p-values associated with each measure of concordance.
unweighted	the "individual level tau" and associated p-value, computed when concordances within individuals are not weighted by their inverse variances.

plotObject	data available to the function plotZLA and used to produce web plots illustrating individual level concordances.
thresholds	a matrix reporting the strength of concordance that would be needed to find significant ( $\alpha = 0.05$ ) evidence consistent with or contrary to ZLA in the study population. These are the 90 percent confidence intervals for the null distributions of tau (ie, individual level concordance). Values are reported for both the inverse-variance weighted and unweighted versions of tau.

### Author(s)

CD Durrant and R. Tucker Gilman (2024)

### References

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. (doi.org/10.18637/jss.v067.i01)

Gilman, R. T., Durrant, C. D., Malpas, L., and Lewis, R. N. (2023) Does Zipf's law of abbreviation shape birdsong? *bioRxiv* (doi.org/10.1101/2023.12.06.569773)

Lewis, R. N., Kwong, A., Soma, M., de Kort, S. R., Gilman, R. T. (2023) Java sparrow song conforms to Mezerath's law but not to Zipf's law of abbreviation. *bioRxiv* (doi.org/10.1101/2023.12.13.571437)

### Examples

```
#Test for evidence of ZLA in the songs of 73 Java sparrows.  
#Most parameters are set to their default values, but  
#"null" is set to the minimum value to make the example run  
#quickly. Thus, the taus reported will be accurate, but the  
#p-values will be imprecise.
```

```
testZLA(Java.sparrow.notes, null = 99)
```

# Index

## \* **datasets**

Java.sparrow.notes, [2](#)

Java.sparrow.notes, [2](#)

plotZLA, [3](#)

testZLA, [4](#)