# Beta diversity computed as Var(Y)

## Description

Compute estimates of total beta diversity as the total variance in a community data matrix **Y**, as well as derived SCBD and LCBD statistics, for 17 dissimilarity coefficients or the raw data table. Tests of significance of the LCDB indices are also produced.

## Usage

```
beta.div(Y,method="hellinger",sqrt.D=FALSE,samp=TRUE,nperm=999,save.D=FALSE)
```

## Arguments

Y        Community composition data. The file class can be either `data.frame` or `matrix`.

method        Name of one of the 17 dissimilarity coefficients, or "none". See Details. Default: method="hellinger".

sqrt.D        If `sqrt.D=TRUE`, the distances in matrix **D** are square-rooted before computation of SStotal, BDtotal and LCBD. This transformation may be useful for methods {"manhattan", "modmeanchardiff", "whittaker", "divergence", "canberra", "percentagedifference", "wishart"} since square-root transformation of the distances makes these **D** matrices Euclidean. Note 1 – Euclideanarity is useful for ordination by principal coordinate analysis; lack of this property does not affect the calculation of SStotal, BDtotal and LCBD. Note 2 – The logical value given to parameter `sqrt.D` has no incidence on calculations through methods {"euclidean", "profiles", "hellinger", "chord", "chisquare", "none"} since no **D** matrix is computed in those cases.

samp        If `samp=TRUE`, the abundance-based distances (ab.jaccard, ab.sorensen, ab.ochiai, ab.simpson) are computed for sample data. If `samp=FALSE`, they are computed for true population data.

nperm        Number of permutations for the tests of significance of LCBD indices.

save.D        If `save.D=TRUE`, the distance matrix will appear in the output list.

## Details

Calculations may be carried out in two ways, depending on the selected method.

• For untransformed or transformed raw data, the total sum of squares (SStotal) is first computed, then the total beta diversity (BDtotal), which is SStotal divided by ($n$ – 1), is calculated. This algorithm is used for methods {"euclidean", "profiles", "hellinger", "chord", "chisquare", "none"}. No dissimilarity matrix nor transformation of the data is computed when method is "euclidean" or "none" (no transformation). For methods "profiles", "hellinger", "chord" and "chisquare", the algorithm begins with computation of the same-name transformations of the

community data (Legendre and Gallagher 2001; Legendre and Legendre 2012, Section 7.7); SStotal and BDtotal are then computed on the transformed data.

• Calculations can also be conducted from a dissimilarity matrix. SStotal is computed by summing the squared dissimilarities in the lower triangular dissimilarity matrix and dividing by $n$; then, total beta diversity (BDtotal) is computed as above. Choices are: method = {"manhattan", "modmeanchardiff", "whittaker", "divergence", "canberra", "percentagedifference", "wishart", "kulczynski", "ab.jaccard", "ab.sorensen", "ab.ochiai", "ab.simpson"}. Equations for these dissimilarities are presented in Table 1 of Legendre and De Cáceres (submitted). See Chao et al. (2006) for details about the abundance-based coefficients (last 4 in the list).

*Species contributions to beta diversity* (SCBD indices for the species) can be computed for the untransformed or transformed raw data, but not for dissimilarity matrices. *Local contributions to beta diversity* (LCBD indices for the sites) can be computed in all cases: raw or transformed data, as well as dissimilarity matrices. See Legendre and De Cáceres (submitted) for details.

LCBD indices are tested for significance by random, independent permutations within the columns of table **Y**. This permutation method tests $H_0$ that the species are distributed at random, independently of one another, among the sites, while preserving the species abundance distributions found in the observed data. See Legendre and De Cáceres (submitted) for discussion.

## Value

Function beta.div returns a list containing the following results:

| | |
|---|---|
| SStotal_BDtotal | Total sum of squares and total beta diversity (= Var(**Y**)) of the data matrix. |
| SCBD | Vector of *Species contributions to beta diversity* (SCBD). |
| LCBD | Vector of *Local contributions to beta diversity* (LCBD) for the sites. |
| p.LCBD | P-values associated with LCBD indices. |
| method | Method selected. |
| note | Notes indicate whether the selected coefficient is Euclidean or not. |
| D | The distance matrix if save.D=TRUE. |

## References

Chao, A., R. L. Chazdon, R. K. Colwell and T. J. Shen. 2006. Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* 62: 361–371.

Legendre, P. and M. De Cáceres. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. (Submitted).

Legendre, P. and E. D. Gallagher, E.D. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129: 271–280.

Legendre, P. and Legendre, L. 2012. *Numerical Ecology. 3rd English edition*. Elsevier Science BV, Amsterdam.

## Author

Pierre Legendre, Département de sciences biologiques, Université de Montréal.
License: GPL-2.

## Example

### Example: mite data available in the vegan package
require(vegan)
data(mite)
res = beta.div(mite, "hellinger", nperm=999)

# Plot a map of the LCDB indices
# First, load the file of Cartesian coordinates of the 70 mite sampling sites
data(mite.xy)
plot(mite.xy, asp=1, type="n", xlab="x coordinates (m)", ylab="y coordinates (m)", main="Map of mite LCBD")
points(mite.xy, pch=21, col="white", bg="brown", cex=120*res$LCBD)

### Example using the log-transformed mite abundance data
mite.log = log1p(mite)
res = beta.div(mite.log, "none", nperm=999)

# Plot a map of the LCDB indices
# First, load the file of cartesian coordinates of the 70 mite sampling sites
data(mite.xy)
plot(mite.xy, asp=1, type="n", xlab="x coordinates (m)", ylab="y coordinates (m)", main="Map of mite LCBD")
points(mite.xy, pch=21, col="white", bg="brown", cex=120*res$LCBD)