

How can I calculate the Bray-Curtis dissimilarity in R?

Authored by
stats writer

July 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate the Bray-Curtis dissimilarity in R?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=163689>

The Bray-Curtis dissimilarity is a measure of dissimilarity between two samples or groups, commonly used in ecology and biology. It calculates the proportion of dissimilarities in species composition between two samples, taking into account both the presence and absence of species. In R, the Bray-Curtis dissimilarity can be calculated using the "vegan" package, which provides a function called "vegdist". This function takes a matrix or data frame of species abundances as input and returns a matrix of Bray-Curtis dissimilarities. The resulting dissimilarity matrix can then be used for further analysis, such as clustering or ordination. This method allows for the quantitative comparison of species composition between samples and can provide valuable insights into community structure and diversity.

Calculate Bray-Curtis Dissimilarity in R

The is a way to measure the dissimilarity between two different sites.

It's often used in ecology and biology to quantify how different two sites are in terms of the species found in those sites.

It is calculated as:

$$BC_{ij} = 1 - (2 * C_{ij}) / (S_i + S_j)$$

where:

C_{ij} : The sum of the lesser values for the species found in each site.
 S_i : The total number of specimens counted at site i
 S_j : The total number of specimens counted at site j

The Bray-Curtis Dissimilarity always ranges between 0 and 1 where:

0 indicates that two sites have zero dissimilarity. In other words, they share the exact same number of each type of species. 1 indicates that two sites have complete dissimilarity. In other words, they share none of the same type of species.

For example, suppose a botanist goes out and counts the number of five different plant species (A, B, C, D, and E) in two different sites.

The following table summarizes the data she collected:

Count of Species					
	A	B	C	D	E
Site 1	4	0	2	7	8
Site 2	3	6	0	4	11

Using this data, she can calculate the Bray-Curtis dissimilarity as:

	Count of Species				
	A	B	C	D	E
Site 1	4	0	2	7	8
Site 2	3	6	0	4	11

$$C_{ij} = 3 + 0 + 0 + 4 + 8 = 15$$

$$S_i = 4 + 0 + 2 + 7 + 8 = 21$$

$$S_j = 3 + 6 + 0 + 4 + 11 = 24$$

Plugging these numbers into the Bray-Curtis dissimilarity formula, we get:

$$BC_{ij} = 1 - (2 * C_{ij}) / (S_i + S_j) = 1 - (2 * 15) / (21 + 24) = 0.33$$

The Bray-Curtis dissimilarity between these two sites is 0.33.

The following example shows how to calculate Bray-Curtis dissimilarity in R.

Example: Calculating Bray-Curtis Dissimilarity in R

First, let's create the following data frame in R to hold our data values:

```
#create data frame
df <- data.frame(A=c(4, 3),
```

```
B=c(0, 6),  
C=c(2, 0),  
D=c(7, 4),  
E=c(8, 11))
```

```
#view data frame  
df
```

```
A B C D E  
1 4 0 2 7 8  
2 3 6 0 4 11
```

We can use the following code to calculate the Bray-Curtis dissimilarity between the two rows of the data frame:

```
#calculate Bray-Curtis dissimilarity  
sum(apply(df, 2, function(x) abs(max(x)-min(x)))) /  
sum(rowSums(df))  
  
0.3333333
```

The Bray-Curtis dissimilarity turns out to be 0.33.

This matches the value that we calculated earlier by

hand.

Note: This formula will only work if each row in the data frame represents a distinct site.

Additional Resources

The following tutorials explain how to calculate other similarity metrics in R:

ARABPSYCHOLOGY.COM