

Perl script dist.pl - Features and Options

C. Oertel - M. Goeker

Features

- * Reads trees (topologies with or without branch lengths) in NEXUS format
- * Computes patristic (path-length) distances from each tree
- * Reads distance matrices in NEXUS and PHYLIP format
- * Computes mean distances, standard deviations and the inferences of the variance from multiple distance matrices (either entered directly or inferred from trees); makes use of weights for trees or distance matrices
- * Can cope with LINUX/UNIX, Mac, and DOS/Windows line breaks

Options

- 1 set all branch lengths to 1 (computes topological distances only). This is the default for trees without branch lengths.
- d compute patristic distance matrices only, no averages etc.
- a compute averages etc., expects distance matrices as input (as computed, e.g., with -d)
- f <n> start at the nth tree
- l <n> stop at the nth tree

Known issues

Trees containing negative branch lengths may cause erroneous output. Negative branch lengths may be set to zero using Rod Page's Treeview program, for instance (the tree file just needs to be saved again).

Input format 1

```
#NEXUS
begin trees;
  translate
    1 Cow,
    2 Pig,
    3 Camel,
    4 Dog;
  tree paup1 = (4:2.00, (1:3.00, (3:3.50, 2:5.00):5.50):0.00);
  tree paup2 = (4:2.00, (1:1.00, (3:3.10, 2:4.00):5.00):1.00);
  tree paup3 = (4:2.00, (1:2.00, (3:3.20, 2:6.00):6.00):2.00);
end;
```

Input format 2

PHYLIP distance matrix (squared or lower triangular). The taxon labels must not contain whitespace.

Output of patristic distances (= input format 3)

```
#NEXUS
[! Patristic distance matrices from treefile nexus.txt]
begin taxa;
```

```

dimensions ntax=4;
taxlabels
  cow
  pig
  camel
  dog
;
end;
[! paup1]
begin distances;
  format triangle=lower nodiagonal nolabels;
  matrix
    13.500000
    12.000000  8.500000
    5.000000  12.500000  11.000000
  ;
end;
[! paup2]
begin distances;
  format triangle=lower nodiagonal nolabels;
  matrix
    10.000000
    9.100000  7.100000
    4.000000  12.000000  11.100000
  ;
end;
[! paup3]
begin distances;
  format triangle=lower nodiagonal nolabels;
  matrix
    14.000000
    11.200000  9.200000
    6.000000  16.000000  13.200000
  ;
end;

```

Output of average distances etc.

```

#NEXUS
[! Mean pairwise distances from file nexus_distances.nex]
begin taxa;
  dimensions ntax=4;
  taxlabels
    cow
    pig
    camel
    dog
  ;
end;
begin distances [&N 3];
  format triangle=lower nodiagonal nolabels;
  matrix
    12.500000
    10.766667  8.266667
    5.000000  13.500000  11.766667
  ;
end;
begin inv_variances [&N 3];
  format triangle=lower nodiagonal nolabels;
  matrix
    0.315789

```

```
        0.668648    1.311953
        1.500000    0.315789    0.971922
    ;
end;
begin deviations [&N 3];
    format triangle=lower nodiagonal nolabels;
    matrix
        1.779513
        1.222929    0.873053
        0.816497    1.779513    1.014342
    ;
end;
```