Practical7

09/10/18

Show all R commands and outputs

1) Returning to the dataset orange- can use all 5 trees here.

age=Orange$age

circ=Orange$circumference

tree=Orange$Tree

Look at the variable tree. It is a factor.

Notice the levels for tree are set so that the trees are ordered.

Using subsetting, look at the summary of each tree's circumference individually.

e.g. summary(circ[tree==1]) etc

What do you think the ordering on the levels of tree (i.e. 3 < 1 < 5 < 2 < 4) were based on?

2) We will make a single plot with the data from all 5 trees as follows:

par(mfrow=c(1,1),pch=8) #set plot window AND default plotting character

set the title to the tree identification by using ‘main= Orange Tree 1 etc’

set the x and ylab to 'Age in Days' and 'Circumference'.

plot(age[tree==1],circ[tree==1],col=1) #need to add in title and axis labels

points(age[tree==2],circ[tree==2],col=2,pch=9) #change default plotting character

At this point look at the graph- is anything unusual?

Why are there only 3 points for the second tree?

Recall that the tree factor was ordered by size-

Because tree2 is larger than tree 1 the axes, which were based on the data for the ‘plot’ command, do not extend far enough to include the data for tree 2.

We need to set the variable ylim right at the start so that there is enough room for the data for all 5 trees. We also should check whether we need limits for the x-axis (age).

max(age) or max(circ)

But min also could matter potentially- so best to use ‘summary’.

Check summary of both age and circ.

Now set the initial plot command as follows:

plot(age[tree==1],circ[tree==1],col=1,ylim=c(30,214)) #need to add in title and axis labels

3) We also could do this (try it):

plot(age[tree==1],circ[tree==1],col=1,ylim=c(min(circ),max(circ)) #need to add in title and axis labels

Now follow on with the plot for trees 2-5 (with plotting characters set below)

points(age[tree==2],circ[tree==2],col=2,pch=9) #change default plotting character

points(age[tree==3],circ[tree==3],col=3,type='l') #change default plotting character

points(age[tree==4],circ[tree==4],col=4,lty=2) #change default plotting character

points(age[tree==5],circ[tree==5],col=5,type='c') #change default plotting character

note that unless we explicitly change pch it will go to the default that was set with par.

type is ‘points’, ‘lines’, or ‘both’

lty is the line design- e.g. dots, dashes,solid