15C_aphid_cat_1

nbutool

02/03/2022

The purpose of this document is to display the results of the Power analysis for each mean temperature (15, 19, 23, 27) for the aphid data. For each temperature there will be 4 results total as follows.

1. Anova summary comparing 0 to 0.95 m<0, categorization based on whole sequence
2. Anova summary comparing 0 to 0.95 m>0, categorization based on whole sequence
3. Anova summary comparing 0 to 0.95 m<0, categorization based on first 1/4th of sequence
4. Anova summary comparing 0 to 0.95 m>0, categorization based on first 1/4th of sequence

In this document we'll focus on cat_1

##No. of acceptable samples per temperature

```r
#Power analysis for autocorrelation experiments - duckweed
#based on file named simpleanovathermal.Rmd

#Download the data from github repo and check import
datin <- read.csv("https://raw.githubusercontent.com/Cuddington-Lab/thermal-experiments/main/expdata_met
                  header=TRUE, stringsAsFactors = TRUE)
str(datin)
```

```
## 'data.frame':    136 obs. of  23 variables:
##  $ Experiment_Number: int  2 2 2 2 3 3 3 3 4 4 ...
##  $ Mean_Temp        : int  15 15 15 15 19 19 19 19 23 23 ...
##  $ Experiment_Start : Factor w/ 40 levels "","02-Aug-20",..: 34 34 34 34 2 2 2 2 13 13 ...
##  $ Experiment_End   : Factor w/ 40 levels "","02-Feb-22",..: 40 40 40 40 9 9 9 9 20 20 ...
##  $ Profile_name     : Factor w/ 119 levels "","Simplelong 15_095_06",..: 22 35 28 36 46 48 49 51 53 5
##  $ Autocorrelation  : num  0 0.9 0.6 0.9 0 0.6 0.95 0.95 0 0.6 ...
##  $ Incubator        : int  1 3 5 6 1 3 5 6 1 3 ...
##  $ Offspring_Plant1 : int  0 6 0 0 12 19 22 19 21 23 ...
##  $ Offspring_Plant2 : int  8 6 3 0 16 16 23 17 21 26 ...
##  $ Offspring_Plant3 : int  1 0 4 1 21 14 8 8 22 20 ...
##  $ Duckweed_Rep1    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Duckweed_Rep2    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Duckweed_Rep3    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ cat_1            : Factor w/ 4 levels "","N","N/A","P": 3 1 3 1 3 3 2 4 3 3 ...
##  $ cat_1_4          : Factor w/ 4 levels "","N","N/A","P": 3 1 3 1 3 3 4 2 3 3 ...
##  $ Program_mean     : num  NA NA NA NA NA ...
##  $ Obs_mean         : num  NA NA NA NA NA ...
##  $ Program_sd       : num  NA NA NA NA NA ...
##  $ Obs_sd           : num  NA NA NA NA NA ...
##  $ Program_ac       : num  NA NA NA NA NA ...
```

```
## $ Obs_ac            : num  NA NA NA NA NA ...
## $ Gaps              : Factor w/ 3 levels "","n","y": 2 2 2 2 1 2 2 2 2 2 ...
## $ Gap_size          : Factor w/ 8 levels "","35h","36h",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```r
#Exclude NAs and samples with standard deviations too different from set value of 2.5
#(code for duckweeds, needs to be adapted for aphids)
datin <- subset(datin, Offspring_Plant1 != "NA" & Obs_sd < 2.7 & Obs_sd > 2.2 & Gaps != "y")

#Create new treatment label and check
table(datin$Autocorrelation, datin$cat_1)
```

```
##
##           N N/A  P
##   0       0  0  22  0
##   0.95    0 23   0 23
```

```r
levels(datin$cat_1) = c("","m<0","", "m>0" )
datin$label<-paste0(datin$Autocorrelation, datin$cat_1)
table(datin$label)
```

```
##
##        0 0.95m<0 0.95m>0
##       22      23      23
```

```r
#Create new column including sum of fronds (sumFro)
datin$sumFro=datin$Offspring_Plant1+datin$Offspring_Plant2+datin$Offspring_Plant3
datin <- subset(datin, sumFro != "NA")
table(datin$Mean_Temp, datin$label)
```

```
##
##        0 0.95m<0 0.95m>0
##   15  6       5       6
##   19  0       1       1
##   23  3       5       5
##   27 13      12      11
```

15 deg C

```r
#Select mean temp of 15C
dat27 <- subset(datin, Mean_Temp == 15)

#Perform power analysis based on preliminary data
#Source: https://med.und.edu/daccota/_files/pdfs/berdc_resource_pdfs/sample_size_r_module.pdf
library(pwr)
```

Categorization for whole sequence

checking 0 to 0.95m<0

```
#Subset data to include only one 0.95 autocorrelation group and control group
#Obtain same number of samples for each group
datpower <- rbind(dat27[ sample(which (dat27$label == "0") ,5), ],
                  dat27[ sample(which (dat27$label == "0.95m<0") ,5), ])
```

Producing boxplot

```
#BoxPlot
tr=boxplot(sumFro~label, data=datpower,main=expression(paste("slope based on whole sequence – aphid: me
          xlab="autocorr treatment", ylab="summed nymphs",
          names = levels(as.factor(datpower$label)))
stripchart(sumFro~label, data=datpower,
          vertical = TRUE, method = "jitter",
          pch = 19, add = TRUE)
```

slope based on whole sequence – aphid: mean temperature 15°C
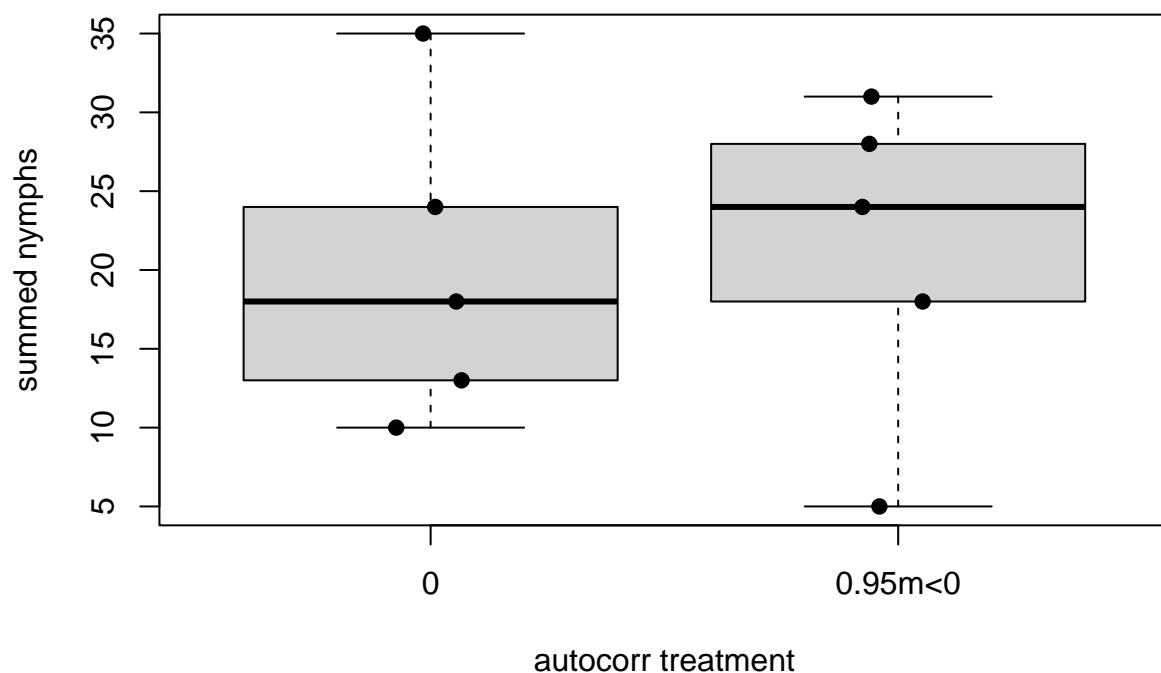
power analysis results

```
#Perform anova
anova <- aov(sumFro~label, data=datpower)
summary(anova)
```

```
##            Df Sum Sq Mean Sq F value Pr(>F)
## label       1    3.6     3.6   0.035  0.856
## Residuals   8  816.8   102.1
```

```
#Obtain effect size
library(effectsize)
#https://cran.r-project.org/web/packages/effectsize/vignettes/anovaES.html
effectsize <- eta_squared(anova, partial = FALSE)
summary(effectsize)
```

```
##    Parameter            Eta2              CI          CI_low      CI_high
## Length:1          Min.   :0.004388   Min.   :0.95   Min.   :0   Min.   :1
## Class :character  1st Qu.:0.004388   1st Qu.:0.95   1st Qu.:0   1st Qu.:1
## Mode  :character  Median :0.004388   Median :0.95   Median :0   Median :1
##                   Mean   :0.004388   Mean   :0.95   Mean   :0   Mean   :1
##                   3rd Qu.:0.004388   3rd Qu.:0.95   3rd Qu.:0   3rd Qu.:1
##                   Max.   :0.004388   Max.   :0.95   Max.   :0   Max.   :1
```
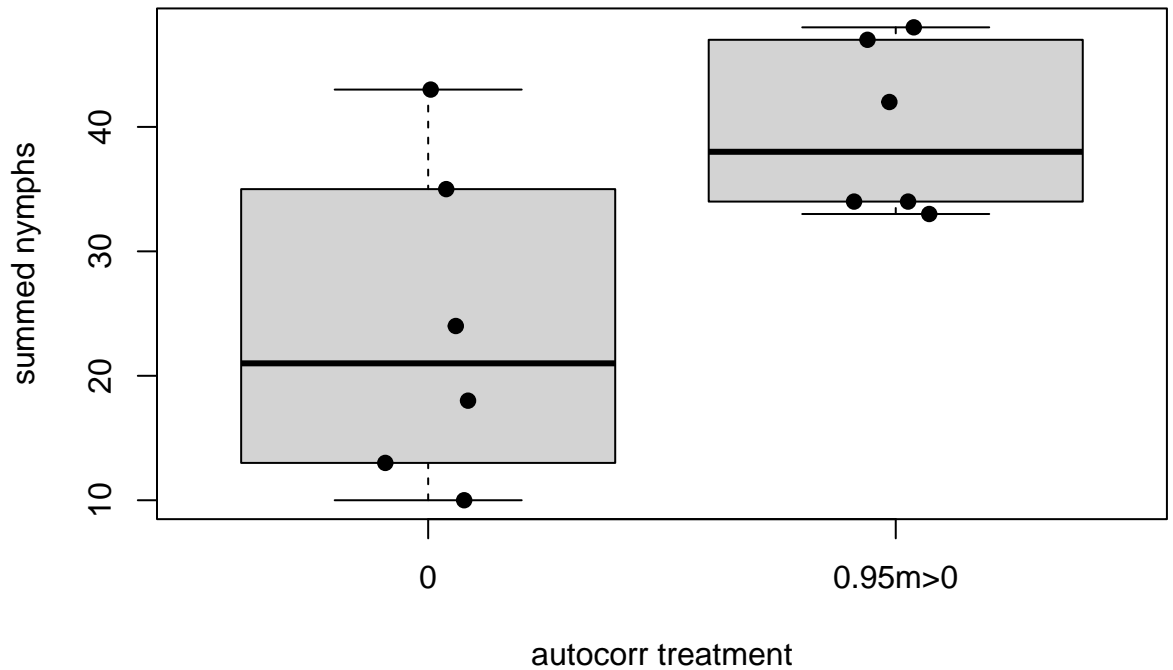
```
#Perform power test to obtain estimated "n" in each group based on effect size
pwr.anova.test(k = 2, f = effectsize$Eta2, sig.level = 0.05, power = 0.80)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##              k = 2
##              n = 203808.9
##              f = 0.004388103
##      sig.level = 0.05
##          power = 0.8
##
## NOTE: n is number in each group
```

checking 0 to 0.95m>0

slope based on whole sequence – aphid: mean temperature 15°C

producing boxplot

power analysis results

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## label        1   752.1   752.1   7.015 0.0244 *
## Residuals   10  1072.2   107.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##   Parameter              Eta2              CI            CI_low
## Length:1          Min.   :0.4123   Min.   :0.95   Min.   :0.04207
## Class :character  1st Qu.:0.4123   1st Qu.:0.95   1st Qu.:0.04207
## Mode  :character  Median :0.4123   Median :0.95   Median :0.04207
##                   Mean   :0.4123   Mean   :0.95   Mean   :0.04207
##                   3rd Qu.:0.4123   3rd Qu.:0.95   3rd Qu.:0.04207
##                   Max.   :0.4123   Max.   :0.95   Max.   :0.04207
##    CI_high
## Min.   :1
## 1st Qu.:1
## Median :1
## Mean   :1
## 3rd Qu.:1
## Max.   :1


##
##      Balanced one-way analysis of variance power calculation
##
```

```
##                   k = 2
##                   n = 24.08857
##                   f = 0.4122699
##           sig.level = 0.05
##               power = 0.8
##
## NOTE: n is number in each group
```

Note the code is not shown for the second iteration comparing 0 to 0.95 m>0 since it's the same as the other code, just the sample size is changed.