Relative Brain Mass

Kyle Emerson

5/10/2023

## R Markdown

This RMarkdown will contain start to finish analysis of Relative Brain Mass information compiled for Emerson Aim 3: Stress and the Microbiota Experiment. It is updated to be more streamlined and easier to knit the output and be placed into my data notebook. Note: This is updated for Manuscript Writing and all analysis was done in RStudio.

library(lme4)

## Warning: package 'lme4' was built under R version 4.1.3

## Loading required package: Matrix

## Warning: package 'Matrix' was built under R version 4.1.3

library(car)

## Warning: package 'car' was built under R version 4.1.3

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.1.3

library(readr)

## Warning: package 'readr' was built under R version 4.1.3

library(moments)

## Warning: package 'moments' was built under R version 4.1.3

library(psych)

## Warning: package 'psych' was built under R version 4.1.3

##   
## Attaching package: 'psych'

## The following object is masked from 'package:car':  
##   
## logit

library(pastecs)  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.1.3

##   
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':  
##   
## %+%, alpha

library(ggbiplot)

## Loading required package: plyr

## Warning: package 'plyr' was built under R version 4.1.3

## Loading required package: scales

## Warning: package 'scales' was built under R version 4.1.3

##   
## Attaching package: 'scales'

## The following objects are masked from 'package:psych':  
##   
## alpha, rescale

## The following object is masked from 'package:readr':  
##   
## col\_factor

## Loading required package: grid

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v tibble 3.1.7 v dplyr 1.0.9  
## v tidyr 1.2.0 v stringr 1.5.0  
## v purrr 0.3.4 v forcats 0.5.1

## Warning: package 'tibble' was built under R version 4.1.3

## Warning: package 'tidyr' was built under R version 4.1.3

## Warning: package 'dplyr' was built under R version 4.1.3

## Warning: package 'stringr' was built under R version 4.1.3

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x ggplot2::%+%() masks psych::%+%()  
## x scales::alpha() masks ggplot2::alpha(), psych::alpha()  
## x dplyr::arrange() masks plyr::arrange()  
## x scales::col\_factor() masks readr::col\_factor()  
## x purrr::compact() masks plyr::compact()  
## x dplyr::count() masks plyr::count()  
## x purrr::discard() masks scales::discard()  
## x tidyr::expand() masks Matrix::expand()  
## x tidyr::extract() masks pastecs::extract()  
## x dplyr::failwith() masks plyr::failwith()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::first() masks pastecs::first()  
## x dplyr::id() masks plyr::id()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::last() masks pastecs::last()  
## x dplyr::mutate() masks plyr::mutate()  
## x tidyr::pack() masks Matrix::pack()  
## x dplyr::recode() masks car::recode()  
## x dplyr::rename() masks plyr::rename()  
## x purrr::some() masks car::some()  
## x dplyr::summarise() masks plyr::summarise()  
## x dplyr::summarize() masks plyr::summarize()  
## x tidyr::unpack() masks Matrix::unpack()

library(modelbased)

## Warning: package 'modelbased' was built under R version 4.1.3

library(dplyr)  
library(emmeans)  
  
file.choose()

## [1] "C:\\R\\Aim 3 - Stress and the Microbiota\\Relative-Brain-Mass.docx"

df <- read.csv("C:\\Users\\kjeme\\OneDrive\\Desktop\\Woodley Lab\\Aim 3 - CORT & the Microbiota\\Aim 3 (Restarted)\\KE Data FIles\\Aim 3 - Relative Brain Mass\\Aim 3 Relative Brain Mass.csv")  
df$Water\_Trtmt = factor(df$Water\_Trtmt)  
df$Stressor = factor(df$Stressor)  
df$Trtmt\_Combo = factor(df$Trtmt\_Combo)  
df$Replicate = factor(df$Replicate)  
df$Gosner\_Stage = as.numeric(df$Gosner\_Stage)

## Warning: NAs introduced by coercion

shapiro.test(df$Brain\_Mass\_g)

##   
## Shapiro-Wilk normality test  
##   
## data: df$Brain\_Mass\_g  
## W = 0.98565, p-value = 0.289

Shapiro-Wilk test to ensure data is normal (> 0.05)

shapiro.test(df$Log\_Brain\_Mass)

##   
## Shapiro-Wilk normality test  
##   
## data: df$Log\_Brain\_Mass  
## W = 0.98675, p-value = 0.3529

Log transformed Brain Mass is normal as well.

leveneTest(df$Brain\_Mass\_g, df$Trtmt\_Combo, center = mean, na.rm = TRUE)

## Levene's Test for Homogeneity of Variance (center = mean: TRUE)  
## Df F value Pr(>F)  
## group 5 1.5036 0.1951  
## 104

Levenes test for homogeneity of variance. Data is homoscedastic.

leveneTest(df$Log\_Brain\_Mass, df$Trtmt\_Combo, center = mean, na.rm = TRUE)

## Levene's Test for Homogeneity of Variance (center = mean: TRUE)  
## Df F value Pr(>F)  
## group 5 0.7333 0.6001  
## 104

Also passes for log transformed variable. Because our log transformed body mass values pass all assumptions (non-transformed body mass fails levenes test), I will move forward with our log transformed brain and body mass values.

anovaBrainMass <- aov(Log\_Brain\_Mass~Water\_Trtmt\*Stressor\*Log\_Body\_Mass, data = df)  
  
summary(anovaBrainMass, type = "III")

## Df Sum Sq Mean Sq F value Pr(>F)   
## Water\_Trtmt 1 0.0397 0.0397 26.838 1.18e-06 \*\*\*  
## Stressor 2 0.0106 0.0053 3.585 0.0314 \*   
## Log\_Body\_Mass 1 0.5100 0.5100 344.753 < 2e-16 \*\*\*  
## Water\_Trtmt:Stressor 2 0.0071 0.0035 2.390 0.0969 .   
## Water\_Trtmt:Log\_Body\_Mass 1 0.0020 0.0020 1.343 0.2492   
## Stressor:Log\_Body\_Mass 2 0.0025 0.0013 0.860 0.4262   
## Water\_Trtmt:Stressor:Log\_Body\_Mass 2 0.0074 0.0037 2.489 0.0882 .   
## Residuals 98 0.1450 0.0015   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## 34 observations deleted due to missingness

Note: type “III” makes no difference, since I believe that all of my numerical values are log transformed already.

This was our ANCOVA (univariate analysis) which takes body mass into account when developing our relative brain mass values. By confirming that there was no significant interaction between our treatments and body mass (> 0.05), we can say that the slopes of lines for our trait (i.e. brain mass) were paralell among each treatment group.

anovaBrainMass$residuals

## 1 4 5 6 7   
## 0.0221272577 0.0563432256 -0.0161864240 -0.0027087940 -0.0265992326   
## 9 10 11 12 13   
## -0.0312285989 -0.0081937214 -0.0354611610 -0.0081343737 -0.0118119279   
## 16 17 18 19 21   
## 0.0107136117 -0.0118533856 0.0027930247 -0.0302482737 0.0431783112   
## 22 23 25 27 28   
## 0.0332405924 0.0140298694 -0.0114129602 -0.0524715556 -0.0370926514   
## 29 30 31 33 34   
## -0.0680405008 -0.0272055438 0.0029909255 0.0034732401 0.0195594500   
## 35 36 37 39 40   
## -0.0388830150 0.0383566516 0.0022675933 0.0045564970 0.0349219436   
## 41 42 43 44 45   
## 0.0828382332 0.0357949365 0.0259521941 0.0227072752 0.0338601500   
## 46 49 51 52 53   
## -0.0721728634 -0.0242627680 0.0129774791 -0.0028105313 0.0345353063   
## 54 55 57 58 59   
## 0.0083416024 0.0659507474 -0.0079697581 0.0311279769 -0.0212148552   
## 61 63 64 65 66   
## 0.0001579597 0.0151743771 0.0097889270 0.0106001404 -0.0160422109   
## 67 68 70 71 72   
## 0.0207672849 -0.0476315291 -0.1146537072 0.0209965991 0.0041669594   
## 73 75 77 78 79   
## 0.0076033901 0.0019579155 -0.0003770332 -0.0721287454 -0.0055455358   
## 81 82 83 84 85   
## -0.0322274576 0.0275003062 -0.0370015983 0.0062240622 0.0113236423   
## 87 88 89 90 91   
## 0.0119567375 0.0270414295 -0.0250697600 0.0804448323 0.0196293864   
## 93 94 95 96 97   
## -0.0434626944 0.0027720729 0.0566529672 -0.0372939175 0.0204448942   
## 99 101 102 103 105   
## 0.0747594076 -0.0031513438 -0.0044284785 -0.0234662554 -0.0029904790   
## 106 109 111 112 114   
## 0.0364885283 -0.0447779173 -0.0620823501 -0.0625927638 0.0126288957   
## 115 116 117 118 119   
## 0.0689366505 0.0539255551 -0.0195177287 -0.0488807573 0.0273120851   
## 120 121 123 125 126   
## -0.0226079427 -0.0025673417 -0.0880211427 -0.0148878289 0.0582298767   
## 129 130 131 132 133   
## -0.0264825414 -0.0516317829 -0.0068645753 0.0005240291 -0.0005310513   
## 135 136 137 138 139   
## 0.0323651738 0.0490081061 -0.0098140331 0.0562629570 0.0295684874   
## 140 141 142 143 144   
## -0.0145348950 0.0345281037 0.0126157873 -0.0333029880 -0.0244643406

The Anova calculated our residuals from this statistical model that we will add to our EMMs. We will turn these into a separate data frame/vector in the global environment now, and then find our EMMs:

residuals <- anovaBrainMass$residuals  
  
model <-lm(Log\_Brain\_Mass~Water\_Trtmt\*Stressor\*Log\_Body\_Mass, data = df)  
  
means\_complex <- estimate\_means(model)

## We selected `at = c("Water\_Trtmt", "Stressor")`.

## NOTE: Results may be misleading due to involvement in interactions

means\_complex$Mean

## [1] -1.631404 -1.650008 -1.637535 -1.633611 -1.608334 -1.644068

Above are our EMMs for each treatment group that we will add to our residuals calculated above. These are located in a new data frame in the global environment. Each residual from each individual will be added to the EMM corresponding to what treatment group that tadpole was in. ~Same EMM as output from SPSS (Analyze -> GLM -> Univariate).

Now, we are going to set up a second large data frame that includes residuals.

# Species\_ID <-df$Species\_ID  
# Log\_Brain\_Mass <- df$Log\_Brain\_Mass  
# Water\_Trtmt <- df$Water\_Trtmt  
# Stressor <-df$Stressor  
# Trtmt\_Combo <-df$Trtmt\_Combo  
# Replicate <- df$Replicate  
# df2 <- cbind(Species\_ID, Log\_Brain\_Mass)  
# df2 <- cbind(df2, Water\_Trtmt)  
# df2 <- cbind(df2, Stressor)  
# df2 <- cbind(df2, Trtmt\_Combo)  
# df2 <- cbind(df2, Replicate)  
# w = complete.cases(df2)  
# df2 = df2[w,]  
# df2 <- cbind(df2, residuals)  
# write.table(df2, file = "brainmasswithresiduals.csv", sep = ",")

If you are doing this for the first time, this will give you a new data frame CSV file containing your individuals and their residuals. Once you have your residuals correctly assigned to each tadpole, add them to the original master CSV file. Now that our new residuals (Residuals\_Brain\_Mass) have been added to the original master CSV file, we are going to Mass adjust those variables. To do so, I am going to add a new column with EMM next to each column. Then, I am going to add a third column representing our mass adjusted variable brain mass next to the EMM column I will add the Residuals to the EMM based on treatment group and have the resulting number be our MAV\_Brain\_Mass, in the original master CSV file.

MAV\_BrainMass\_glmm <- glmer(MAV\_Brain\_Mass~Water\_Trtmt\*Stressor + (1|Replicate), data = df, family = "gaussian")

## Warning in glmer(MAV\_Brain\_Mass ~ Water\_Trtmt \* Stressor + (1 | Replicate), :  
## calling glmer() with family=gaussian (identity link) as a shortcut to lmer() is  
## deprecated; please call lmer() directly

Anova(MAV\_BrainMass\_glmm)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: MAV\_Brain\_Mass  
## Chisq Df Pr(>Chisq)   
## Water\_Trtmt 3.9994 1 0.04552 \*  
## Stressor 1.7866 2 0.40930   
## Water\_Trtmt:Stressor 3.3750 2 0.18498   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Our anova shows a significant effect of water treatment on our relative brain mass.

emmeans(MAV\_BrainMass\_glmm, list(pairwise ~ Stressor), adjust = "tukey")

## NOTE: Results may be misleading due to involvement in interactions

## $`emmeans of Stressor`  
## Stressor emmean SE df lower.CL upper.CL  
## A -1.64 0.00779 18.2 -1.66 -1.62  
## B -1.64 0.00782 18.1 -1.65 -1.62  
## C -1.63 0.00768 17.1 -1.64 -1.61  
##   
## Results are averaged over the levels of: Water\_Trtmt   
## Degrees-of-freedom method: kenward-roger   
## Confidence level used: 0.95   
##   
## $`pairwise differences of Stressor`  
## 1 estimate SE df t.ratio p.value  
## A - B -0.00475 0.0110 18.2 -0.430 0.9036  
## A - C -0.01419 0.0109 17.6 -1.297 0.4152  
## B - C -0.00944 0.0110 17.6 -0.861 0.6711  
##   
## Results are averaged over the levels of: Water\_Trtmt   
## Degrees-of-freedom method: kenward-roger   
## P value adjustment: tukey method for comparing a family of 3 estimates

Non-significant.

Stress <- c("Vehicle Control", "Predator Cues", "CORT")  
#For labeling X axis  
  
ggplot(df, aes(x= factor(Stressor, level=c('B', 'A','C')), y = Relative\_Brain\_Mass, fill = Water\_Trtmt)) +  
 geom\_boxplot(width = 0.5, outlier.colour = "transparent", alpha =0.35) +  
 stat\_boxplot(geom = "errorbar", width = .5) +  
 theme\_classic() +  
 scale\_fill\_manual(values = c("seagreen3", "lightcyan"),  
 name = "Microbial Environment",  
 labels = c("Natural", "Autoclaved")) +  
 labs(x = "Stressor", y = "Relative Brain Mass (mg)") +  
 theme(aspect.ratio = 1) +  
 scale\_x\_discrete(labels = Stress)+  
 theme(axis.text = element\_text(face = "bold", size = 16)) +  
 theme(axis.title = element\_text(face = "bold", size = 20)) +  
 theme(legend.position = c(.83, .92)) +  
 theme(legend.title = element\_text(size = 14)) +  
 theme(legend.text = element\_text(size = 14))

## Warning: Removed 34 rows containing non-finite values (`stat\_boxplot()`).  
## Removed 34 rows containing non-finite values (`stat\_boxplot()`).

