Behavioral Tolerance, Vigor, and resilience

David R. Clark, Jason Walsman, Rachael Kramp, and Jessica Stephenson

2024-03-25

Contents

1 Introduction

1.1 Overall summary

1.2 Quick experimental summary

1.3 Load in packages needed for analysis

```r
# Clear the working environment
rm(list = ls())
# Visualization
library(ggplot2)
library(visreg)
source("http://highstat.com/Books/BGS/GAMM/RCodeP2/HighstatLibV6.R")
# (generalized) Linear mixed modeling
library(lme4)
library(glmmTMB)
library(lmodel2)
# Statistcal analysis reporting and model validation
library(performance)
library(car)
library(lmtest)
library(DHARMa)
# Data wrangling
library(dplyr)
library(plyr)
library(tidyverse)
library(tidylog)
library(splancs)
```

1.4 Load in dataframe

```r
# Loading in data set
IndBehav <- read_csv("VIEBehavior_20210913_V2.csv", col_types = cols(fishID = col_character(),
    Infection = col_character(), InfDate = col_character(), BehavGroup = col_character(),
    dayofinf = col_character(), CountDay = col_character(), Behavdate = col_character(),
    TOD = col_character(), Duration = col_double(), Velocity = col_double(), Distance = col_double()))

# Filtering Data set so we can perform our analysis
IndBehav1 <- IndBehav %>%
    # Removing columns we dont need
select(-c(Blind.ID, InfLength, LateLength, dayofinf, Countintial, Notes, InfWeight))
```

```
## select: dropped 7 variables (Blind.ID, InfWeight, InfLength, LateLength, dayofinf, ...)
```

```r
# Quickly checking whether individuals have stress response from a new
# environment.

# Filter to only before measurements
IndBehav1_Before <- IndBehav1 %>%
    filter(TrialTime == "Before")
```

```
## filter: removed 829 rows (75%), 276 rows remaining
```

```r
# plot the velocity by time of day

ggplot(IndBehav1_Before, aes(x = TOD, y = Velocity, fill = Sex)) + geom_boxplot(aes(fill = Sex)) +
    geom_jitter() + theme_classic()
```
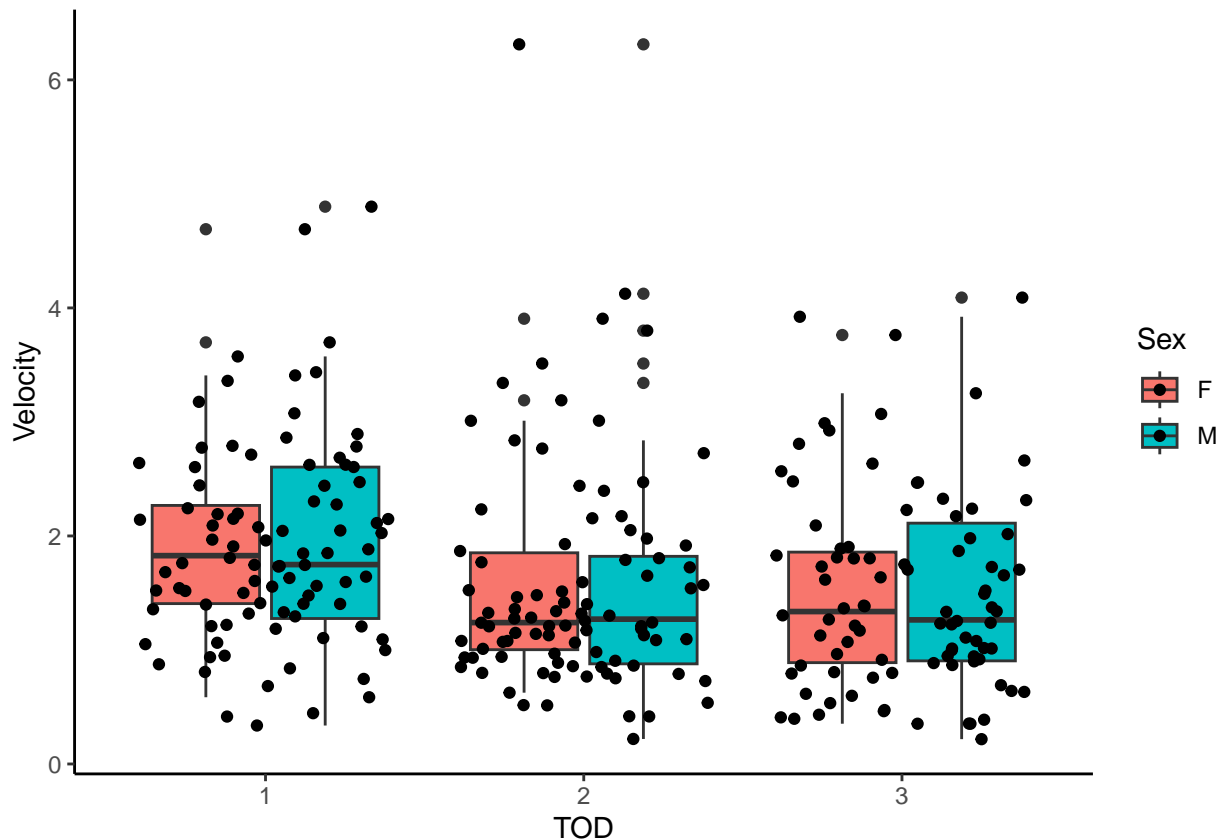


```r
# Quick Anova to verify that the difference is significantly different

anova(lm(Velocity ~ TOD * Sex, IndBehav1_Before))
```

```
## Analysis of Variance Table
##
## Response: Velocity
##            Df  Sum Sq Mean Sq F value   Pr(>F)
## TOD         2   9.570  4.7852  5.7966 0.003438 **
## Sex         1   0.010  0.0097  0.0117 0.913883
## TOD:Sex     2   0.148  0.0740  0.0896 0.914307
## Residuals 264 217.939  0.8255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# We do see a significant TOD effect so lets more specifically look at each
# individual
ggplot(IndBehav1_Before, aes(x = TOD, y = Velocity, color = fishID)) + geom_point() +
    geom_line(aes(group = fishID)) + theme_classic() + theme(legend.position = "none")
```
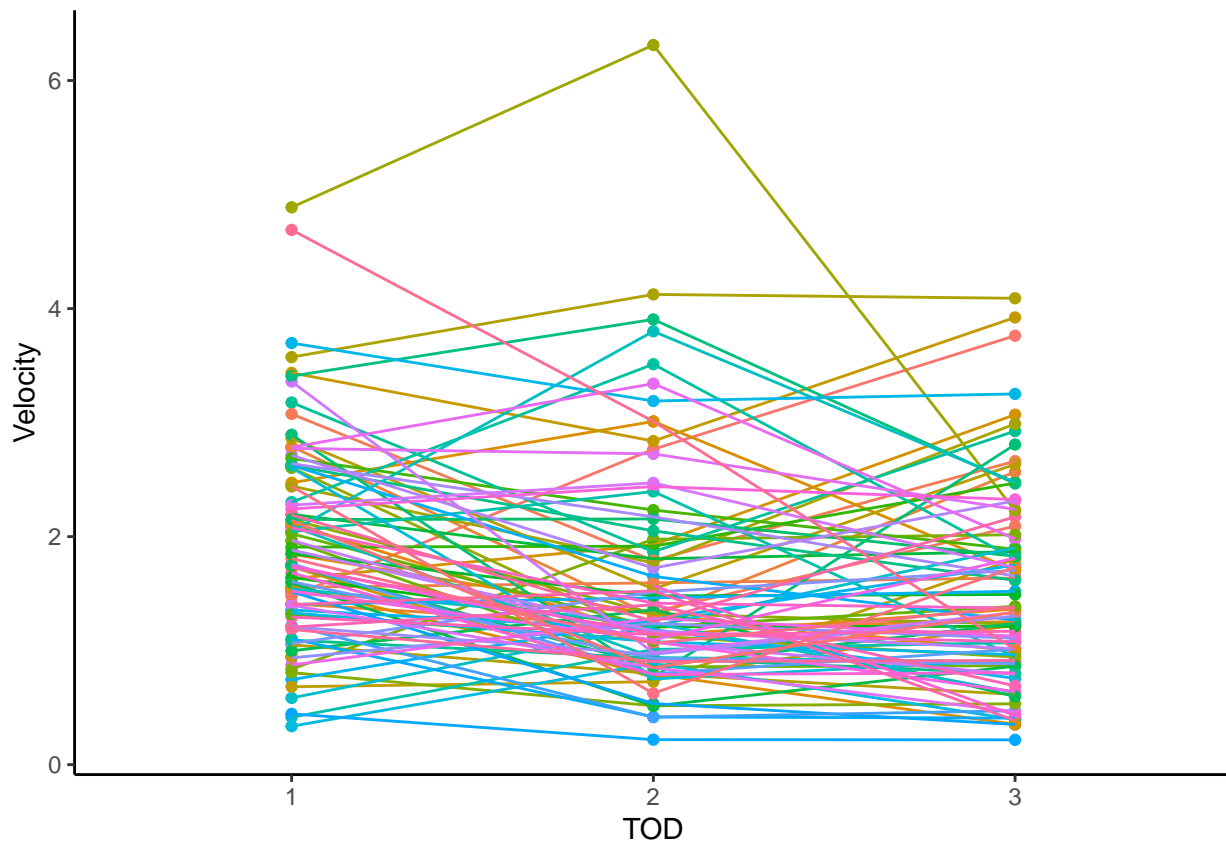
```
# Dropping the first video due to these TOD effects
IndBehav1 <- IndBehav1 %>%
    filter(TOD != "1" | TrialTime != "Before")
```

```
## filter: removed 93 rows (8%), 1,012 rows remaining
```

1.5 Quickly check for stress response when individauls are

1.6 Calculate metrics needed for further analyses

1.6.1 Calculating all body condition metrics

```
##### This code is just to examine how many fish cleared their infection
##### (13)#### IndBehavRec<- IndBehav1 %>% filter(Recov==1)%>% filter(Fish==1)


##### This code is just to examine how many fish died during their infection
##### (2)##### IndBehavDead<- IndBehav1 %>% filter(Died==1)%>% filter(Fish==1)

FemaleOnly <- IndBehav1 %>%
    filter(Sex == "F")
```

```
## filter: removed 495 rows (49%), 517 rows remaining
```

```r
MaleOnly <- IndBehav1 %>%
    filter(Sex == "M")
```

filter: removed 517 rows (51%), 495 rows remaining

```r
# Calculating SMI metrics Preinfection SMI#####

# PreinfectionSMI for females To calculate the SMI we take the OLS slope
lmodel2(log(as.numeric(PreWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = FemaleOnly)
```

RMA was not requested: it will not be computed.

No permutation test will be performed

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(PreWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = FemaleOnly)
##
## n = 517   r = 0.9103118   r-square = 0.8286675
## Parametric P-values:   2-tailed = 1.996223e-199   1-tailed = 9.981117e-200
## Angle between the two OLS regression lines = 4.099607 degrees
##
## Regression results
##   Method Intercept     Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -1.114762 0.4206982        22.81640                NA
## 2     MA -1.155627 0.4341311        23.46717                NA
## 3    SMA -1.240856 0.4621474        24.80390                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS      -1.165163       -1.064362  0.4041380   0.4372585
## 2     MA      -1.207946       -1.103957  0.4171464   0.4513293
## 3    SMA      -1.292137       -1.191380  0.4458838   0.4790043
##
## Eigenvalues: 0.00962744 0.0002518882
##
## H statistic used for computing C.I. of MA: 0.0002067558
```

```r
# Take the median length from summary
summary(FemaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.60   18.40   20.20   20.03   21.20   24.20
```

```r
# Adding PreSMI to Dataframe
FemaleOnly <- FemaleOnly %>%
    group_by(fishID) %>%
    mutate(PreSMI = PreWeight * ((20.2/PreLength)^0.4206982))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'PreSMI' (double) with 47 unique values and 0% NA
```

```
# PreinfectionSMI for males
lmodel2(log(as.numeric(PreWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = MaleOnly)
```

```
## RMA was not requested: it will not be computed.
```

```
## No permutation test will be performed
```

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(PreWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = MaleOnly)
##
## n = 495    r = 0.7784225    r-square = 0.6059415
## Parametric P-values:    2-tailed = 9.315674e-102    1-tailed = 4.657837e-102
## Angle between the two OLS regression lines = 6.76859 degrees
##
## Regression results
##   Method  Intercept     Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -0.4620565 0.1938185        10.96898                NA
## 2     MA -0.4750867 0.1984879        11.22660                NA
## 3    SMA -0.6160126 0.2489888        13.98170                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS     -0.5006629      -0.4234502  0.1799875   0.2076494
## 2     MA     -0.5147236      -0.4356638  0.1843606   0.2126918
## 3    SMA     -0.6556798      -0.5784876  0.2355417   0.2632036
##
## Eigenvalues: 0.004315198 9.775356e-05
##
## H statistic used for computing C.I. of MA: 0.000185703
```

```
# Take the median length from summary
summary(MaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13.70   14.50   15.30   15.32   16.00   18.50
```

```
# Adding PreSMI to Dataframe
MaleOnly <- MaleOnly %>%
    group_by(fishID) %>%
    mutate(PreSMI = PreWeight * ((15.3/PreLength)^0.1938185))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'PreSMI' (double) with 44 unique values and 0% NA
```

```
##### Late Infection SMI####

# Late infection SMI for females Take OLS slope from this model
lmodel2(log(as.numeric(LateWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = FemaleOnly)
```

```
## RMA was not requested: it will not be computed.
```

```
## No permutation test will be performed
```

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(LateWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = FemaleOnly)
##
## n = 473   r = -0.127178   r-square = 0.01617424
## Parametric P-values:   2-tailed = 0.005607781     1-tailed = 0.00280389
## Angle between the two OLS regression lines = 57.58662 degrees
##
## Regression results
##    Method  Intercept        Slope Angle (degrees) P-perm (1-tailed)
## 1     OLS    2.033718  -0.5976124        -30.86306                NA
## 2      MA 107.560934 -35.3034151        -88.37748                NA
## 3     SMA  14.504560  -4.6990243        -77.98610                NA
##
## Confidence intervals
##    Method 2.5%-Intercept 97.5%-Intercept  2.5%-Slope 97.5%-Slope
## 1     OLS      0.7499539        3.317482   -1.019621  -0.1756032
## 2      MA     63.1124483      365.209243 -120.038820 -20.6851920
## 3     SMA     13.2788937       15.845234   -5.139945  -4.2959269
##
## Eigenvalues: 0.1896467 0.008436892
##
## H statistic used for computing C.I. of MA: 0.0003994611
```

```
# Take the median length from summary
summary(FemaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.60   18.40   20.20   20.03   21.20   24.20
```

```
# Adding PreSMI to Dataframe
FemaleOnly <- FemaleOnly %>%
    group_by(fishID) %>%
    mutate(LateSMI = PreWeight * ((20.1/PreLength)^-0.5976124))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'LateSMI' (double) with 47 unique values and 0% NA
```

```r
# Late Infection SMI for males Take OLS slope from this model
lmodel2(log(as.numeric(LateWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = MaleOnly)
```

```
## RMA was not requested: it will not be computed.
```

```
## No permutation test will be performed
```

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(LateWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = MaleOnly)
##
## n = 418   r = -0.06095677   r-square = 0.003715728
## Parametric P-values:   2-tailed = 0.2136141   1-tailed = 0.106807
## Angle between the two OLS regression lines = 64.27059 degrees
##
## Regression results
##   Method  Intercept       Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS   1.492739   -0.472241       -25.27860                NA
## 2     MA 348.385697 -124.982883       -89.54158                NA
## 3    SMA  21.760991   -7.747145       -82.64494                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS     -0.5840109         3.56949  -1.217488    0.2730059
## 2     MA    135.2257943      -602.18348 216.205880  -48.4731822
## 3    SMA     19.7843318        23.93692  -8.528154   -7.0376608
##
## Eigenvalues: 0.2046989 0.00339751
##
## H statistic used for computing C.I. of MA: 0.0001594103
```

```r
# Take the median length from summary
summary(MaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13.70   14.50   15.30   15.32   16.00   18.50
```

```r
# Adding PreSMI to Dataframe
MaleOnly <- MaleOnly %>%
    group_by(fishID) %>%
    mutate(LateSMI = PreWeight * ((15.3/PreLength)^-0.472241))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'LateSMI' (double) with 44 unique values and 0% NA
```

```
#####

# Bringing the male and female dataframes together to create our overall
# dataframe
IndBehav2 <- rbind(FemaleOnly, MaleOnly)


# View(IndBehav2)
```

1.6.2 Checking that fish velocity is not different at different times of day.

We recorded the velocity of the fish at 3 time periods over the course of the day. Once in the monring between 0900-1100, once in the afternoon 1200-1400, 1500-1700. Therefore we wanted to confirm that there is no consistent Time of day effect before we collapse the behavioral measurements into means for each day.

```
##### First visually look at the TOD effect for behavior####
ggplot(IndBehav2, aes(TOD, Velocity)) + geom_boxplot() + geom_jitter() + xlab("Time of Day") +
    ylab("Velocity (cm/s)") + theme_classic()
```
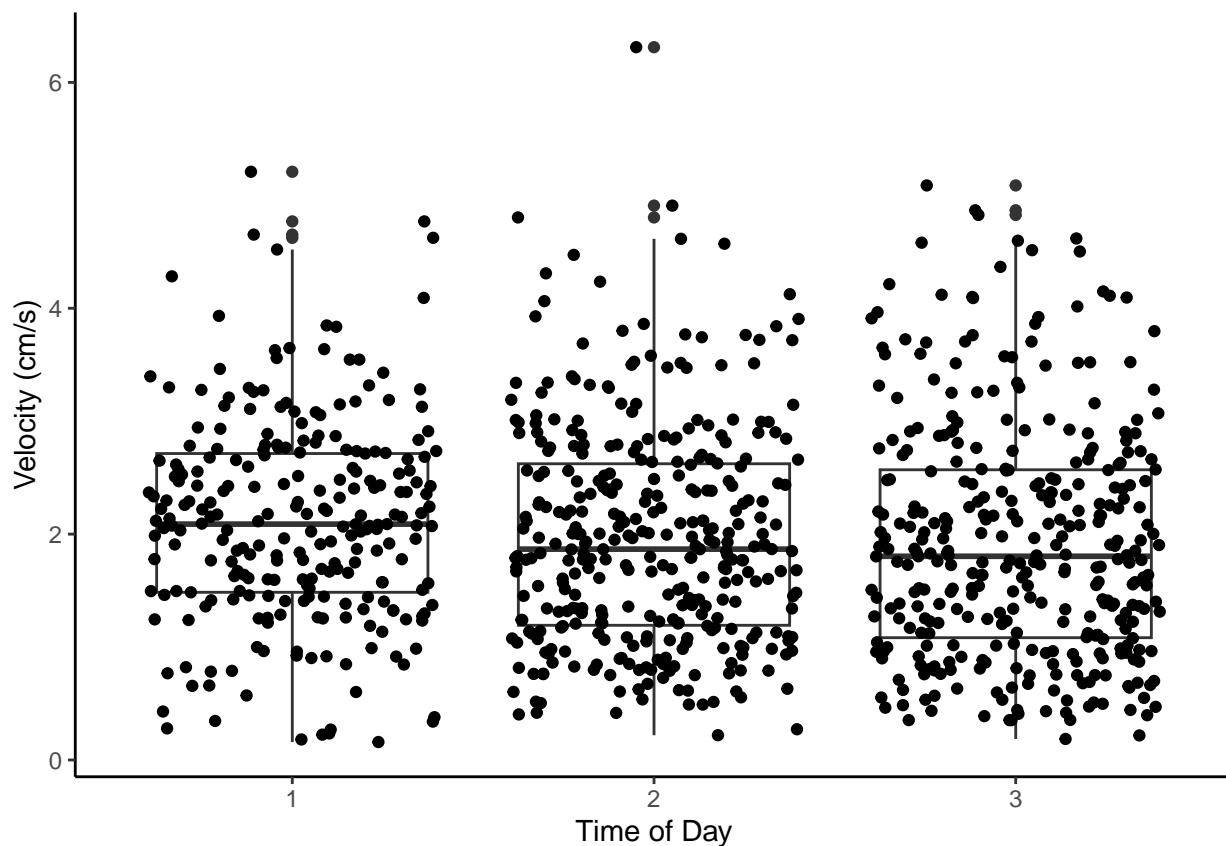


```
#### Verify that this is not statistically significant#####
anova(lm(Velocity ~ TOD, IndBehav2))
```

```
## Analysis of Variance Table
##
```

9

```
## Response: Velocity
##             Df Sum Sq Mean Sq F value  Pr(>F)
## TOD          2   5.55 2.77265  2.8353 0.05921 .
## Residuals  927 906.51 0.97789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##### Lets also confirm this does not change by sex####
ggplot(IndBehav2, aes(TOD, Velocity, fill = Sex)) + geom_boxplot() + geom_jitter() +
    xlab("Time of Day") + ylab("Velocity (cm/s)") + theme_classic()
```



```
##### Lets also confirm this does n0t change by Infection Status####
ggplot(IndBehav2, aes(TOD, Velocity, fill = Infection)) + geom_boxplot() + geom_jitter() +
    xlab("Time of Day") + ylab("Velocity (cm/s)") + theme_classic()
```
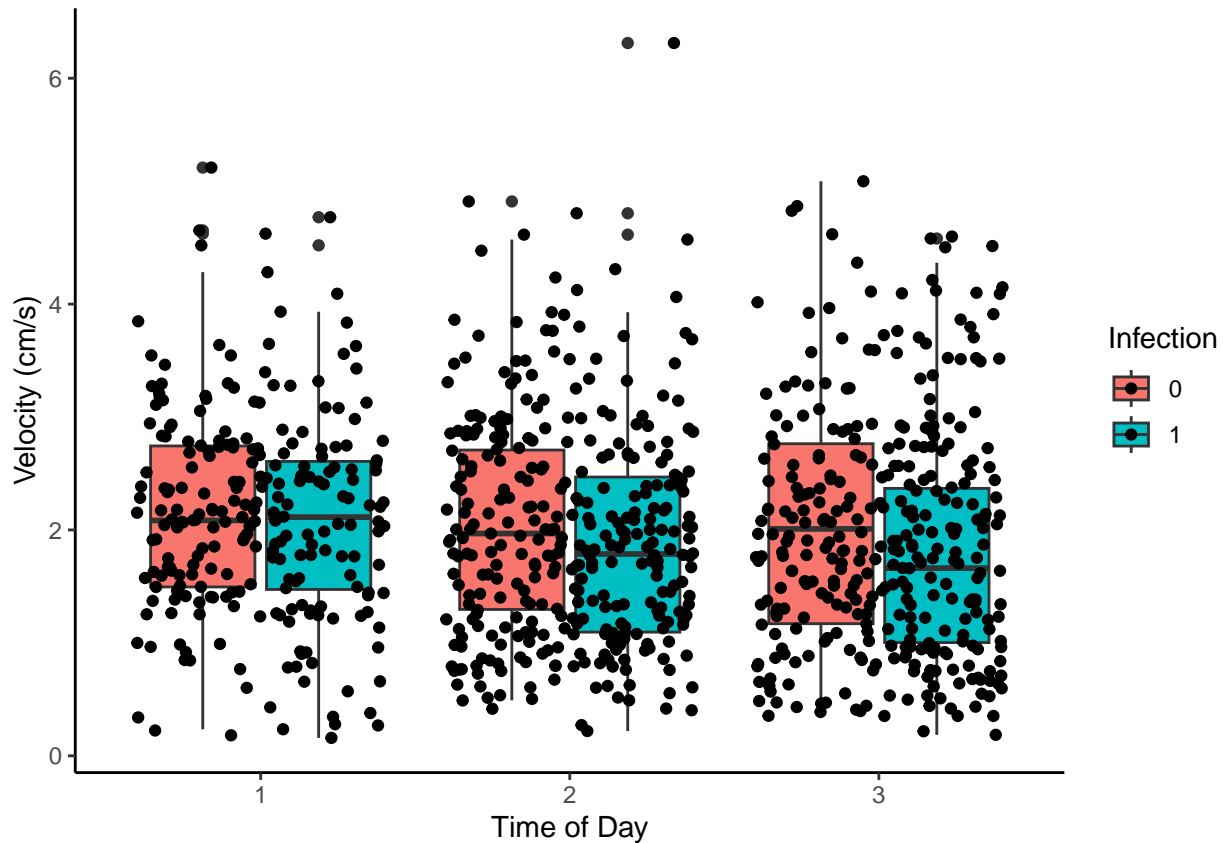
Visually there does not seem to be a distinguished pattern overall and an anova confirms it. It also does not seem to visually differ by sex or infection status. Therefore we can collapse the three different time points within a day to a mean velocity and variance of velocity across each time point, Before, Early, Late, and Later.
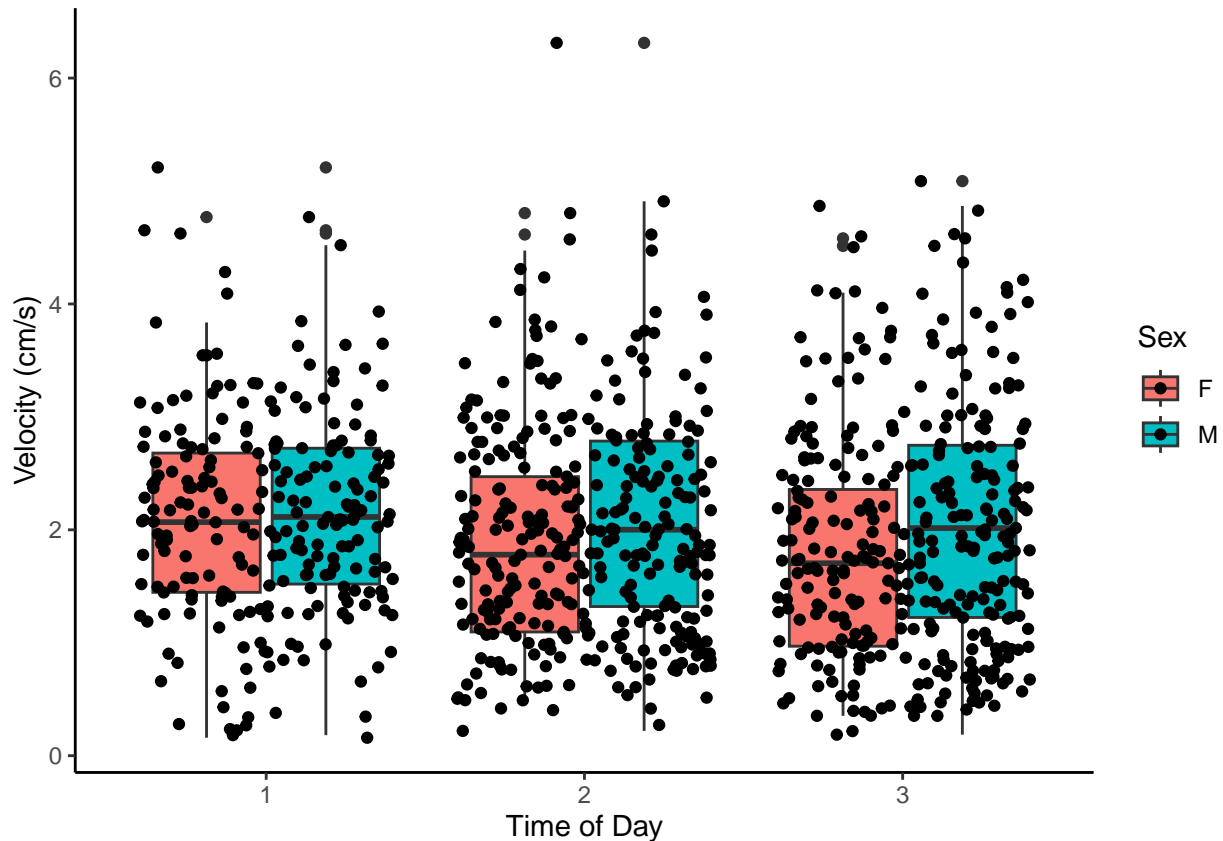
1.6.3 Calculating the behavioral tolerance metrics needed for analyses

We are calculating behavioral tolerance through two metrics. First is a measure of behavioral tolerance using linear mixed models using slope random effects for fishID. This allows us to extract each individuals change in activity with parasite burden and therefore allow us to calculate the tolerance (fishID slope) for each individual. This also allows us to use the random effect intercept for fishID as a measure of behavioral vigor. The second way we could calculate these metrics are a measure of change in behavior between two points of behavior (often referred to as point tolerance, CITATION FOR THIS). This is calculated in the second half of the code as the change in behavior between pre-infection and early, late, and later points of infection.

```
# Calulating the behavioral tolerance metrics were interested in for our
# analysis

# Random slope from random effects model Random effects model with random slope
# for each fish by worm burden prior to measuring its activity. This slope is
# the tolerance an individual has to changing its behavior as infection
# increases. The slope of the behavior is the individuals behavioral vigor, or
# its pre-infection behavior#####
BehavTolLM <- lmer(Velocity ~ (Wormbf | fishID), IndBehav2)
summary(BehavTolLM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Velocity ~ (Wormbf | fishID)
##    Data: IndBehav2
##
## REML criterion at convergence: 1253.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7065 -0.6145 -0.0122  0.5722  4.4823
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  fishID   (Intercept) 0.5552671 0.745162
##           Wormbf      0.0000108 0.003287 -0.40
##  Residual             0.5589316 0.747617
## Number of obs: 493, groups:  fishID, 48
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   1.9020     0.1075   17.69
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 2.74354 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```r
##### Extracting the intercept and slope from the linear mixed model####
betaTol <- data.frame(coef(BehavTolLM)$fishID)
##### Renaming the slope and intercept to align with our metrics#####
betaTol <- betaTol %>%
    rename(BehavTol = Wormbf, BehavVig = X.Intercept.)
```

```
## rename: renamed 2 variables (BehavTol, BehavVig)
```

```r
##### Creating a fishID column for combination with future dataframe####
betaTol <- cbind(fishID = rownames(betaTol), betaTol)

# Calculating the mean and variance in velocity per day, per fish

# We need to split the dataframes into the differnet time periods to calculate
# averages since mutate isnt splitting by TrialTime.

##### creating a dataframe to calculate the mean and variance of velocity for
##### before measurements.####

InBehavBefore <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Before") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining
```

```
## filter (grouped): removed 750 rows (81%), 180 rows remaining
```

```
## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)
```

```
## mutate (grouped): new variable 'AvgVel' (double) with 91 unique values and 0% NA
```

```
##                     new variable 'VarVel' (double) with 90 unique values and 1% NA
```

```
## distinct (grouped): removed 89 rows (49%), 91 rows remaining
```

```r
##### creating a dataframe to calculate the mean and variance of velocity for
##### early measurements.#####

InBehavEarly <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Early") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining
```

```
## filter (grouped): removed 670 rows (72%), 260 rows remaining
```

```
## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)
```

```
## mutate (grouped): new variable 'AvgVel' (double) with 87 unique values and 0% NA
```

```
##                     new variable 'VarVel' (double) with 87 unique values and 0% NA
```

```
## distinct (grouped): removed 173 rows (67%), 87 rows remaining
```

```r
##### creating a dataframe to calculate the mean and variance of velocity for
##### Late measurements. #####

InBehavLate <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Late") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining
```

```
## filter (grouped): removed 671 rows (72%), 259 rows remaining
```

```
## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 87 unique values and 0% NA

##                    new variable 'VarVel' (double) with 87 unique values and 0% NA

## distinct (grouped): removed 172 rows (66%), 87 rows remaining
```

```r
##### creating a dataframe to calculate the mean and variance of velocity for
##### Later measurements.####

InBehavLater <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Later") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining

## filter (grouped): removed 699 rows (75%), 231 rows remaining

## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 79 unique values and 0% NA

##                    new variable 'VarVel' (double) with 79 unique values and 0% NA

## distinct (grouped): removed 152 rows (66%), 79 rows remaining
```

```r
#### Adding these dataframes together to create one large one####

IndBehav3 <- rbind(InBehavBefore, InBehavEarly, InBehavLate, InBehavLater)


##### Pivoting wider so I can subtract velocity changes between periods to get
##### change in behavior####
IndBehav4 <- IndBehav3 %>%
    group_by(TrialTime, fishID, add = TRUE) %>%
    select(-Velocity) %>%
    pivot_wider(names_from = TrialTime, values_from = c(Wormbf, Ratebf, AvgVel, VarVel))
```

```
## group_by: 2 grouping variables (fishID, TrialTime)

## select: dropped one variable (Velocity)

## pivot_wider: reorganized (Wormbf, TrialTime, Ratebf, AvgVel, VarVel) into (Wormbf_Before, Wormbf_Earl
```

```
##### subset down to infected individuals only and calculate point tolerance
##### metrics for each timeframe. #####
IndBehav4Inf <- IndBehav4 %>%
    filter(Infection == 1) %>%
    mutate(EChBe = AvgVel_Early - AvgVel_Before, LChBe = AvgVel_Late - AvgVel_Before,
        LtrChBe = AvgVel_Later - AvgVel_Before, ERatebf = (Wormbf_Early - Wormbf_Before)/6,
        LRatebf = (Wormbf_Late - Wormbf_Early)/6, LtrRatebf = (Wormbf_Later - Wormbf_Late)/6) %>%
    pivot_longer(cols = starts_with("Wormbf"), names_to = "TrialTime", names_prefix = "Wormbf_",
        values_to = "Wormbf") %>%
    select(-c(AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later, Ratebf_Before,
        Ratebf_Early, Ratebf_Late, Ratebf_Later)) %>%
    filter(fishID != c(44, 118, 139, 94))
```

filter (grouped): removed 39 rows (43%), 52 rows remaining

mutate (grouped): new variable 'EChBe' (double) with 49 unique values and 8% NA

new variable 'LChBe' (double) with 49 unique values and 8% NA

new variable 'LtrChBe' (double) with 42 unique values and 21% NA

new variable 'ERatebf' (double) with 23 unique values and 10% NA

new variable 'LRatebf' (double) with 42 unique values and 10% NA

new variable 'LtrRatebf' (double) with 37 unique values and 23% NA

pivot_longer: reorganized (Wormbf_Before, Wormbf_Early, Wormbf_Late, Wormbf_Later) into (TrialTime, W

select: dropped 8 variables (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later, AvgVel_Before,

filter (grouped): removed 4 rows (2%), 204 rows remaining

```
##### Make two dataframes with pivot longer for wormbf and velocity then merge
##### them together on fishID and TrialTime####
IndBehav4InfV <- IndBehav4 %>%
    filter(Infection == 1) %>%
    pivot_longer(cols = starts_with("AvgVel"), names_to = "TrialTime", names_prefix = "AvgVel_",
        values_to = "AvgVel") %>%
    select(fishID, AvgVel, TrialTime) %>%
    filter(fishID != c(44, 118, 139, 94))
```

filter (grouped): removed 39 rows (43%), 52 rows remaining

pivot_longer: reorganized (AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later) into (TrialTime, A

select: dropped 24 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

filter (grouped): removed 4 rows (2%), 204 rows remaining

```
##### Making the growth rate columns into one column for infected
##### individuals####
IndBehavGRInf <- IndBehav4 %>%
    filter(Infection == 1) %>%
    pivot_longer(cols = starts_with("Ratebf"), names_to = "TrialTime", names_prefix = "Ratebf_",
        values_to = "Ratebf") %>%
    select(fishID, TrialTime, Ratebf)
```

filter (grouped): removed 39 rows (43%), 52 rows remaining

pivot_longer: reorganized (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later) into (TrialTime,

select: dropped 24 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

```
##### Making the new growth rate columns into one column for infected
##### individuals#####
IndBehavNGRInf <- IndBehav4Inf %>%
    select(fishID, ERatebf, LRatebf, LtrRatebf) %>%
    rename(NRatebf_Early = ERatebf, NRatebf_Late = LRatebf, NRatebf_Later = LtrRatebf) %>%
    add_column(NRatebf_Before = 0) %>%
    pivot_longer(cols = starts_with("NRatebf"), names_to = "TrialTime", names_prefix = "NRatebf_",
        values_to = "NRatebf") %>%
    select(fishID, TrialTime, NRatebf) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

select: dropped 21 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

rename: renamed 3 variables (NRatebf_Early, NRatebf_Late, NRatebf_Later)

pivot_longer: reorganized (NRatebf_Early, NRatebf_Late, NRatebf_Later, NRatebf_Before) into (TrialTir

select: no changes

distinct (grouped): removed 608 rows (75%), 208 rows remaining

```
##### Making the Variance columns into one column for infected individuals#####
IndBehavVarInf <- IndBehav4Inf %>%
    select(-TrialTime) %>%
    pivot_longer(cols = starts_with("VarVel"), names_to = "TrialTime", names_prefix = "VarVel_",
        values_to = "VarVel") %>%
    select(fishID, TrialTime, VarVel) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

select: dropped one variable (TrialTime)

pivot_longer: reorganized (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later) into (TrialTime, V

select: dropped 19 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

distinct (grouped): removed 608 rows (75%), 208 rows remaining

```r
##### Merging the tolerance and vigor metrics ####
IndBehav4Inf1 <- merge(IndBehav4Inf, betaTol, by.x = "fishID", by.y = "fishID", .keep = all)

##### Merge the two dataframes together based on the fishID and TrialTime to
##### get all variables we want into one dataframe#####
IndBehav5Inf <- merge(IndBehav4InfV, IndBehav4Inf1, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf2 <- merge(IndBehav5Inf, IndBehavGRInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf3 <- merge(IndBehav5Inf2, IndBehavNGRInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf4 <- merge(IndBehav5Inf3, IndBehavVarInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
# Subset down to uninfected individuals and calculate their point tolerance
# metrics Creating a dataframe for calculating the tolerance metrics and then
# pivot longer by Wormbf####
IndBehav4Unf <- IndBehav4 %>%
    filter(Infection == 0) %>%
    mutate(EChBe = AvgVel_Early - AvgVel_Before, LChBe = AvgVel_Late - AvgVel_Before,
        LtrChBe = AvgVel_Later - AvgVel_Before, NRatebf_Early = (Wormbf_Early - Wormbf_Before)/5,
        NRatebf_Late = (Wormbf_Late - Wormbf_Early)/5, NRatebf_Later = (Wormbf_Later -
            Wormbf_Late)/5, ) %>%
    pivot_longer(cols = starts_with(c("Wormbf")), names_to = "TrialTime", names_prefix = c("Wormbf_"),
        values_to = "Wormbf") %>%
    select(-c(AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later, Ratebf_Before,
        Ratebf_Early, Ratebf_Late, Ratebf_Later))
```

filter (grouped): removed 52 rows (57%), 39 rows remaining

mutate (grouped): new variable 'EChBe' (double) with 39 unique values and 0% NA

new variable 'LChBe' (double) with 39 unique values and 0% NA

new variable 'LtrChBe' (double) with 39 unique values and 3% NA

new variable 'NRatebf_Early' (double) with one unique value and 100% NA

new variable 'NRatebf_Late' (double) with one unique value and 100% NA

new variable 'NRatebf_Later' (double) with one unique value and 100% NA

pivot_longer: reorganized (Wormbf_Before, Wormbf_Early, Wormbf_Late, Wormbf_Later) into (TrialTime, W

select: dropped 8 variables (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later, AvgVel_Before,

```r
##### Creating a second dataframe where we pivot longer for velocity####
IndBehav4UnfV <- IndBehav4 %>%
    filter(Infection == 0) %>%
    pivot_longer(cols = starts_with("AvgVel"), names_to = "TrialTime", names_prefix = "AvgVel_",
        values_to = "AvgVel") %>%
    select(fishID, TrialTime, AvgVel)
```

filter (grouped): removed 52 rows (57%), 39 rows remaining

pivot_longer: reorganized (AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later) into (TrialTime,

select: dropped 24 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

```r
##### Making the growth rate columns into one column for uninfected
##### individuals#####
IndBehavGRUnf <- IndBehav4 %>%
    filter(Infection == 0) %>%
    pivot_longer(cols = starts_with("Ratebf"), names_to = "TrialTime", names_prefix = "Ratebf_",
        values_to = "Ratebf") %>%
    select(fishID, TrialTime, Ratebf)
```

filter (grouped): removed 52 rows (57%), 39 rows remaining

pivot_longer: reorganized (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later) into (TrialTime,

select: dropped 24 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

```r
##### Making the new growth rate columns into one column for infected
##### individuals#####
IndBehavNGRUnf <- IndBehav4Unf %>%
    select(-TrialTime) %>%
    add_column(NRatebf_Before = 0) %>%
    pivot_longer(cols = starts_with("NRatebf"), names_to = "TrialTime", names_prefix = "NRatebf_",
        values_to = "NRatebf") %>%
    select(fishID, TrialTime, NRatebf) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

select: dropped one variable (TrialTime)

pivot_longer: reorganized (NRatebf_Early, NRatebf_Late, NRatebf_Later, NRatebf_Before) into (TrialTime

select: dropped 20 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

distinct (grouped): removed 468 rows (75%), 156 rows remaining

```r
##### Making the Variance columns into one column for infected individuals#####
IndBehavVarUnf <- IndBehav4Unf %>%
    select(-TrialTime) %>%
    pivot_longer(cols = starts_with("VarVel"), names_to = "TrialTime", names_prefix = "VarVel_",
        values_to = "VarVel") %>%
    select(fishID, TrialTime, VarVel) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

select: dropped one variable (TrialTime)

pivot_longer: reorganized (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later) into (TrialTime,

select: dropped 19 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

distinct (grouped): removed 468 rows (75%), 156 rows remaining

```
##### Merge the different datasets together#####
IndBehav5Unf <- merge(IndBehav4UnfV, IndBehav4Unf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf2 <- merge(IndBehav5Unf, IndBehavGRUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf3 <- merge(IndBehav5Unf2, IndBehavNGRUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf4 <- merge(IndBehav5Unf3, IndBehavVarUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)

# Creating the BehavTol and BehavVig columns for the uninfected and calculating
# the metric for uninfected individual so we can combine these dataframes
# together
IndBehav5Unf4 <- IndBehav5Unf4 %>%
    group_by(fishID) %>%
    add_column(BehavTol = 0, BehavVig = 0)
```

group_by: one grouping variable (fishID)

```
# Rbinding our infected and uninfected dataframes together into one overall
# dataframe
IndBehav5 <- rbind(IndBehav5Unf4, IndBehav5Inf4)

## Pivot longer for Change in behavior
IndBehav5ChBe <- IndBehav5 %>%
    select(-TrialTime) %>%
    rename(ChBehav_Early = EChBe, ChBehav_Late = LChBe, ChBehav_Later = LtrChBe) %>%
    add_column(ChBehav_Before = 0) %>%
    pivot_longer(cols = starts_with("ChBehav"), names_to = "TrialTime", names_prefix = "ChBehav_",
        values_to = "ChBehav") %>%
    select(fishID, TrialTime, ChBehav) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

select: dropped one variable (TrialTime)

rename: renamed 3 variables (ChBehav_Early, ChBehav_Late, ChBehav_Later)

pivot_longer: reorganized (ChBehav_Early, ChBehav_Late, ChBehav_Later, ChBehav_Before) into (TrialTir

select: dropped 29 variables (AvgVel, Sex, PreWeight, PreLength, Treatment, ...)

distinct (grouped): removed 1,092 rows (75%), 364 rows remaining

```
# Merge dataframes together for one large dataframe
IndBehav6 <- merge(IndBehav5, IndBehav5ChBe, by.x = c("fishID", "TrialTime"), by.y = c("fishID",
    "TrialTime"), all = TRUE)

# Removing some of the columns we dont want in the dataframe Create a point
# tolerance metric for each timepoint
IndBehav6 <- IndBehav6 %>%
    select(-c(ERatebf, LRatebf, LtrRatebf, NRatebf_Early, NRatebf_Late, NRatebf_Later,
```

```
      VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later)) %>%
   mutate(PTol = ChBehav/Wormbf) %>%
   mutate(PTol = replace(PTol, PTol == NaN, 0)) %>%
   mutate(PTol = replace(PTol, PTol == -Inf, 0)) %>%
   mutate(PTol = replace(PTol, PTol == Inf, 0)) %>%
   select(-c(EChBe, LChBe, LtrChBe)) %>%
   distinct(fishID, TrialTime, .keep_all = TRUE) %>%
   drop_na(Sex)
```

```
## select: dropped 10 variables (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later, NRatebf_Early,
```

```
## mutate: new variable 'PTol' (double) with 123 unique values and 63% NA
```

```
## mutate: no changes
```

```
## mutate: changed 8 values (2%) of 'PTol' (0 new NA)
```

```
## mutate: changed 7 values (2%) of 'PTol' (0 new NA)
```

```
## select: dropped 3 variables (EChBe, LChBe, LtrChBe)
```

```
## distinct: no rows removed
```

```
## drop_na: removed 17 rows (5%), 347 rows remaining
```

```
# Remove dataframes used for data wrangling and managing rm(c(IndBehav,
# IndBehav1,IndBehav2,IndBehav3,IndBehav3Inf1,IndBehav4,IndBehav4Inf,IndBehav4Inf1,IndBehav4InfV,IndBeh

# Setting some characters as factors
IndBehav6$TrialTime <- as.factor(IndBehav6$TrialTime)
IndBehav6$Sex <- as.factor(IndBehav6$Sex)
IndBehav6$Infection <- as.factor(IndBehav6$Infection)
IndBehav6$ContInf <- as.factor(IndBehav6$ContInf)
IndBehav6$Died <- as.factor(IndBehav6$Died)
IndBehav6$Treatment <- as.factor(IndBehav6$Treatment)


# #Exporting dataframe for saving
write.csv(IndBehav6, "IndividualBehaviors_20240416.csv")
```

Now we have to decided how best to calculate whether an individual has recovered. In the experiment only had 13 fish recover from infection totally, therefore we are thinking of another metric that indicates a fish is in recovery during infection.

1.7   Description of data, structure, and type

fishID: The individual ID of each fish used in the trial.
TrialTime: The point of infection where each behavior trial happened. Early - early infection (6 days), Late - late in infection (12 days), Later - later in infection (18 days)
AvgVel: The average velocity from three separate behavioral trials of each fish for each of the Trial Times.

(cm/s)

Sex: The sex of the individual. F - female, M - male

PreWeight: The weight of the individual prior to their fist behavior trial and pre-infection. ($grams$)

PreLength: The length of the individual prior to their first behavior trial and pre-infection (mm)

Treatment: What treatment the fish received prior to their first behavior and pre-infection. VIE - visible implant elastomer implant, UNTOUCHED - control individual, received no injection nor implant (mm)

Infection: Whether or not the individual was infected with Gyrodactylus turnbulli. 1 - infected, 0 - uninfected

LateWeight: The weight of the individual after their final behavior trial and after infection. ($grams$)

wormJump: The number of worms that jumped from the donor fish to the trial fish during manual infections.

AUC2: The area under the curve of infection over the total infection trajectory for each individual.

RecovPeriod: The time frame in which an individual started recovering from infection. This is calculated by looking at the growth rate of the worms between each count and see when the worm growth rate was decreasing overall.

ContInf: Whether or not the individual was controling infection. We qualified controling infection by having a negative growth rate post peak infection. 1 - Controlled, 0 - Uncontrolled

Died: Whether or not the individual Died from infection during the experimental trial. 1 - Died, 0 - did not died

PreSMI: The body condition of the individual prior to their first behavioral trial and pre infection. (mm/g)

LateSMI: The body condition of the individual after to their last behavioral trial and after infection. (mm/g)

Wormbf:The number of worms on the fish prior to each trial.

Ratebf:The rate of growth of worms on the fish prior to each trial from the worm count immediately prior to the trial.

NRatebf:The rate of growth of worms on the fish prior to each trial calculated as the change in worms between each time point. (For Example, From before to early).

VarVel:The variance in velocity from the separate behavioral trials during each time point. ($(cm/s)^2$)

BehavTol:The behavioral tolerance of each individual calculated as the slope of a random effect model where fishID is the random effect term.

BehavVig:The behavioral Vigor of each individual calculated as the intercept of a random effect model where fishID is the random effect term

ChBehav:The change in behavior between trial times. (cm/s)

PTol:Point tolerance metric calculated by dividing the change in behavior between time points by the worms on the fish before the trial ($cm/s/worms$)

1.8   Visualizing the relationships between variables in our dataset

```
pairs(~AvgVel + TrialTime + Sex + Treatment + Infection + LateWeight + AUC2 + Died +
    PreSMI + LateSMI + Wormbf + Ratebf + NRatebf + VarVel + BehavTol + BehavVig +
    ChBehav, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav6)
```
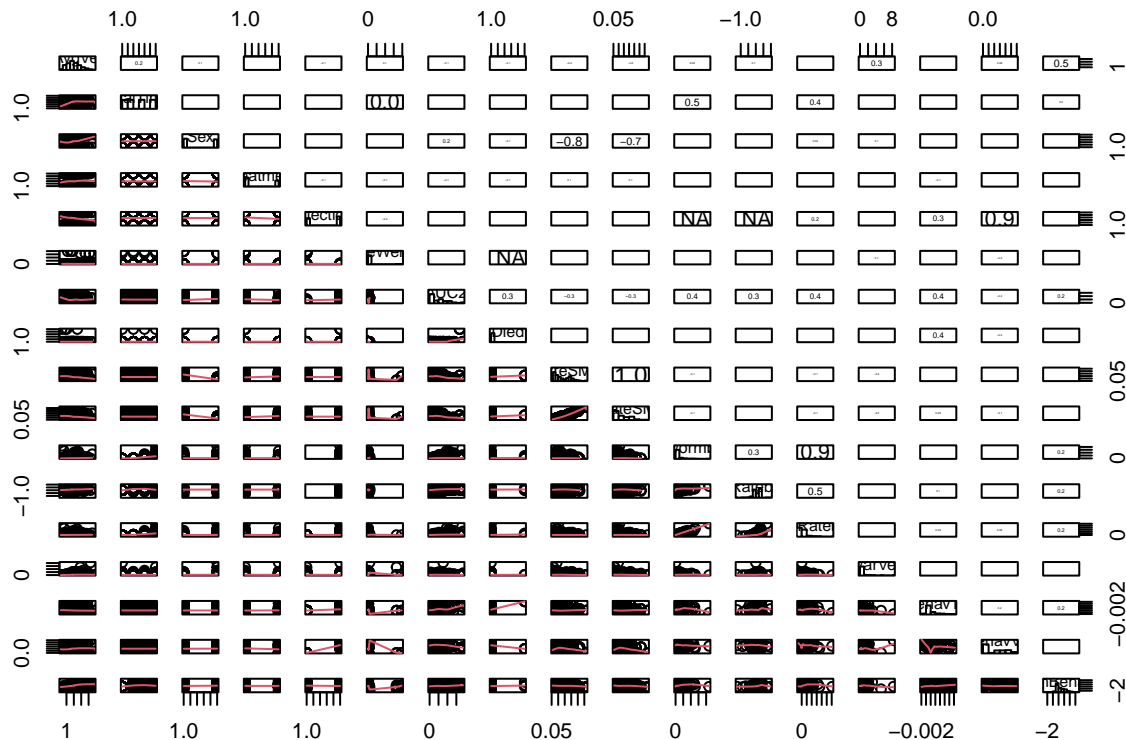
1.9 Visualize some patterns in the raw data

```r
# Load in dataset from above. This bit of code is meant to save time so people
# dont have to rerun the entire data parsing and calculating step above. Code
# will be saved for reproducibility.
IndBehav7 <- read_csv("IndividualBehaviors_20240501.csv")
```

```
## Rows: 347 Columns: 27
## -- Column specification ------------------------------------------------------
## Delimiter: ","
## chr  (3): TrialTime, Sex, Treatment
## dbl (23): fishID, AvgVel, AvgVelBef, VarvelBef, PreWeight, PreLength, Infect...
## lgl  (1): ContrPeriod
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# Setting some of the factors back to factors
IndBehav7$fishID <- as.factor(IndBehav7$fishID)
IndBehav7$TrialTime <- as.factor(IndBehav7$TrialTime)
IndBehav7$Sex <- as.factor(IndBehav7$Sex)
IndBehav7$Infection <- as.factor(IndBehav7$Infection)
IndBehav7$ContInf <- as.factor(IndBehav7$ContInf)
IndBehav7$Died <- as.factor(IndBehav7$Died)
IndBehav7$Treatment <- as.factor(IndBehav7$Treatment)
```

```r
# Calculating tissue tolerance for each individual.
```

```r
IndBehav7 <- IndBehav7 %>%
    mutate(ChSMI = LateSMI - PreSMI) %>%
    mutate(TisTol = ChSMI/Totworm)
```

```
## mutate: new variable 'ChSMI' (double) with 85 unique values and 0% NA
## mutate: new variable 'TisTol' (double) with 47 unique values and 46% NA
```

```r
# Subsetting down to female and males only to scale their SMI Males
IndBehavM <- IndBehav7 %>%
    filter(Sex == "M") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

```
## filter: removed 180 rows (52%), 167 rows remaining
## mutate: new variable 'ScPSMI' (double) with 41 unique values and 0% NA
##          new variable 'ScLSMI' (double) with 41 unique values and 0% NA
```

```r
# Females
IndBehavF <- IndBehav7 %>%
    filter(Sex == "F") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

```
## filter: removed 167 rows (48%), 180 rows remaining
## mutate: new variable 'ScPSMI' (double) with 45 unique values and 0% NA
##          new variable 'ScLSMI' (double) with 45 unique values and 0% NA
```

```r
# Combine the two separate dataframes together
IndBehav8 <- rbind(IndBehavF, IndBehavM)


# Filtering the data set for individuals who are not controlling for infection.
IndBehavCont <- IndBehav7 %>%
    filter(ContInf == "0" | ContInf == "1")
```

```
## filter: removed 152 rows (44%), 195 rows remaining
```

```r
# Filtering to only infected individuals for tolerance metrics
IndBehavI <- IndBehav7 %>%
    filter(Infection == "1")
```

```
## filter: removed 156 rows (45%), 191 rows remaining
```

```r
# Scaling some variables to make them biologically comparable and better for
# model fitting

IndBehav8 <- IndBehav8 %>%
    mutate(ScAvgVelBef = scale(AvgVelBef), ScVarvelBef = scale(VarvelBef), ScNRatebf = scale(NRatebf),
        ScBehavTol = scale(BehavTol), ScTotworm = scale(Totworm), ScAUC = scale(AUC2),
        ScBehavVig = scale(BehavVig), ScChSMI = scale(ChSMI), ScTisTol = scale(TisTol))
```

```
## mutate: new variable 'ScAvgVelBef' (double) with 87 unique values and 1% NA
##         new variable 'ScVarvelBef' (double) with 86 unique values and 2% NA
##         new variable 'ScNRatebf' (double) with 85 unique values and 37% NA
##         new variable 'ScBehavTol' (double) with 49 unique values and 0% NA
##         new variable 'ScTotworm' (double) with 47 unique values and 46% NA
##         new variable 'ScAUC' (double) with 49 unique values and 43% NA
##         new variable 'ScBehavVig' (double) with 87 unique values and 0% NA
##         new variable 'ScChSMI' (double) with 85 unique values and 0% NA
##         new variable 'ScTisTol' (double) with 47 unique values and 46% NA
```

```r
IndBehav8$fishID <- as.character(IndBehav8$fishID)
```

```r
unique(IndBehav8$fishID)
```

```
##  [1] "2"   "4"   "12"  "14"  "15"  "47"  "50"  "59"  "72"  "74"  "83"  "85"
## [13] "91"  "93"  "111" "131" "133" "135" "142" "144" "1"   "3"   "5"   "11"
## [25] "13"  "46"  "49"  "56"  "57"  "71"  "73"  "75"  "81"  "82"  "84"  "92"
## [37] "95"  "112" "113" "114" "115" "132" "134" "141" "145" "9"   "16"  "19"
## [49] "43"  "45"  "76"  "78"  "87"  "107" "109" "116" "117" "119" "120" "126"
## [61] "129" "138" "140" "146" "6"   "7"   "18"  "20"  "44"  "51"  "52"  "53"
## [73] "54"  "55"  "77"  "80"  "86"  "88"  "89"  "90"  "106" "108" "110" "128"
## [85] "130" "136" "148"
```

```r
summary(IndBehav8)
```

```
##     fishID           TrialTime      AvgVel          AvgVelBef
##  Length:347         Before:86   Min.   :0.2186   Min.   :0.2186
##  Class :character   Early :87   1st Qu.:1.2542   1st Qu.:0.9563
##  Mode  :character   Late  :87   Median :1.9052   Median :1.2278
##                     Later :87   Mean   :1.9654   Mean   :1.5268
##                                 3rd Qu.:2.5363   3rd Qu.:2.0114
##                                 Max.   :4.7609   Max.   :4.2697
##                                 NA's   :11       NA's   :3
##    VarvelBef        Sex        PreWeight        PreLength       Treatment
##  Min.   :0.000001   F:180   Min.   :0.0570   Min.   :13.70   UNTOUCHED:207
##  1st Qu.:0.016805   M:167   1st Qu.:0.0795   1st Qu.:15.40   VIE      :140
##  Median :0.119540           Median :0.1100   Median :17.10
##  Mean   :0.547397           Mean   :0.1343   Mean   :17.79
##  3rd Qu.:0.801353           3rd Qu.:0.1930   3rd Qu.:20.30
##  Max.   :4.269705           Max.   :0.3380   Max.   :24.20
##  NA's   :7
##  Infection   LateWeight         wormJump        Totworm          AUC2
##  0:156     Min.   : 0.0480   Min.   :0.000   Min.   :   3.0   Min.   :   4.0
##  1:191     1st Qu.: 0.0870   1st Qu.:2.000   1st Qu.: 133.0   1st Qu.: 268.5
##            Median : 0.1310   Median :2.000   Median : 333.0   Median : 759.5
##            Mean   : 0.5955   Mean   :1.909   Mean   : 560.4   Mean   :1097.7
##            3rd Qu.: 0.1720   3rd Qu.:2.000   3rd Qu.: 887.0   3rd Qu.:1584.5
##            Max.   :19.6000   Max.   :3.000   Max.   :1950.0   Max.   :3758.5
##            NA's   :31        NA's   :172     NA's   :160      NA's   :148
##    Died      ContrPeriod    ContInf      PreSMI           LateSMI
##  0  :199   Mode:logical   0  :116    Min.   :0.05737   Min.   :0.05611
##  1  :  4   NA's:347       1  : 79    1st Qu.:0.07962   1st Qu.:0.07852
##  NA's:144                 NA's:152   Median :0.11601   Median :0.10839
```

```
##                                         Mean   :0.13378   Mean   :0.13560
##                                         3rd Qu.:0.18909   3rd Qu.:0.19642
##                                         Max.   :0.31326   Max.   :0.37766
##
##       Wormbf            Ratebf             NRatebf            VarVel
##   Min.   :  0.00   Min.   :-1.24245   Min.   :-12.667   Min.   :0.000001
##   1st Qu.:  0.00   1st Qu.: 0.00000   1st Qu.:  0.000   1st Qu.:0.045519
##   Median :  9.00   Median : 0.07192   Median :  0.000   Median :0.164248
##   Mean   : 56.95   Mean   : 0.09886   Mean   :  5.174   Mean   :0.388202
##   3rd Qu.: 62.00   3rd Qu.: 0.26494   3rd Qu.:  3.708   3rd Qu.:0.442258
##   Max.   :876.00   Max.   : 0.97296   Max.   :121.167   Max.   :8.339405
##   NA's   :166      NA's   :168        NA's   :127       NA's   :13
##      BehavTol           BehavVig          ChBehav            PTol
##   Min.   :-0.0020206   Min.   :0.8218   Min.   :-2.5006   Min.   :-0.72749
##   1st Qu.: 0.0000000   1st Qu.:1.2964   1st Qu.: 0.0000   1st Qu.:-0.00146
##   Median : 0.0000000   Median :1.6804   Median : 0.1510   Median : 0.00343
##   Mean   : 0.0004484   Mean   :1.7828   Mean   : 0.4252   Mean   : 0.01994
##   3rd Qu.: 0.0004072   3rd Qu.:2.1921   3rd Qu.: 0.9175   3rd Qu.: 0.02145
##   Max.   : 0.0043354   Max.   :3.9298   Max.   : 3.4553   Max.   : 0.63456
##                                        NA's   :11        NA's   :213
##       ChSMI             TisTol            ScPSMI.V1
##   Min.   :-0.0197622   Min.   :-0.00647   Min.   :-1.6648685
##   1st Qu.:-0.0034791   1st Qu.:-0.00001   1st Qu.:-0.7178001
##   Median : 0.0003388   Median : 0.00000   Median :-0.0863437
##   Mean   : 0.0018169   Mean   : 0.00005   Mean   : 0.0000000
##   3rd Qu.: 0.0059333   3rd Qu.: 0.00001   3rd Qu.: 0.5545737
##   Max.   : 0.0643940   Max.   : 0.00468   Max.   : 3.0180687
##                        NA's   :160
##      ScLSMI.V1          ScAvgVelBef.V1     ScVarvelBef.V1     ScNRatebf.V1
##   Min.   :-1.545105   Min.   :-1.571727   Min.   :-0.654146   Min.   :-1.31585
##   1st Qu.:-0.745671   1st Qu.:-0.685482   1st Qu.:-0.634066   1st Qu.:-0.38162
##   Median :-0.174070   Median :-0.359270   Median :-0.511296   Median :-0.38162
##   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000
##   3rd Qu.: 0.586155   3rd Qu.: 0.582184   3rd Qu.: 0.303482   3rd Qu.:-0.10812
##   Max.   : 3.204945   Max.   : 3.295354   Max.   : 4.448220   Max.   : 8.55499
##                       NA's   :3           NA's   :7           NA's   :127
##     ScBehavTol.V1       ScTotworm.V1        ScAUC.V1          ScBehavVig.V1
##   Min.   :-1.9776696   Min.   :-1.06586   Min.   :-1.09754   Min.   :-1.576856
##   1st Qu.:-0.3591724   1st Qu.:-0.81727   1st Qu.:-0.83212   1st Qu.:-0.798058
##   Median :-0.3591724   Median :-0.43481   Median :-0.33940   Median :-0.168043
##   Mean   : 0.0000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000
##   3rd Qu.:-0.0330042   3rd Qu.: 0.62459   3rd Qu.: 0.48848   3rd Qu.: 0.671727
##   Max.   : 3.1134508   Max.   : 2.65734   Max.   : 2.67007   Max.   : 3.523079
##                        NA's   :160        NA's   :148
##      ScChSMI.V1         ScTisTol.V1
##   Min.   :-1.584879   Min.   :-5.19255
##   1st Qu.:-0.388963   1st Qu.:-0.04623
##   Median :-0.108562   Median :-0.03779
##   Mean   : 0.000000   Mean   : 0.00000
##   3rd Qu.: 0.302332   3rd Qu.:-0.02897
##   Max.   : 4.595987   Max.   : 3.68576
##                        NA's   :160
```

2 Calculating athe total area of a polygon based on the number of worms and the avergae velocity of each individual

This metric is used as a combination as the tolerance (change in behavior) and resistance (the change in parasites over the course of infeciton).

```
# Visually inspecting the ploygons for the change in behavior as parasite load
# increases
ggplot(IndBehav7, aes(x = Wormbf, y = AvgVel, color = fishID)) + geom_point() + theme_classic() +
    theme(legend.position = "none") + geom_polygon(aes(fill = fishID, alpha = 1/10),
    rule = "evenodd")
```

```
## Warning: Removed 166 rows containing missing values or values outside the scale range
## ('geom_point()').
```
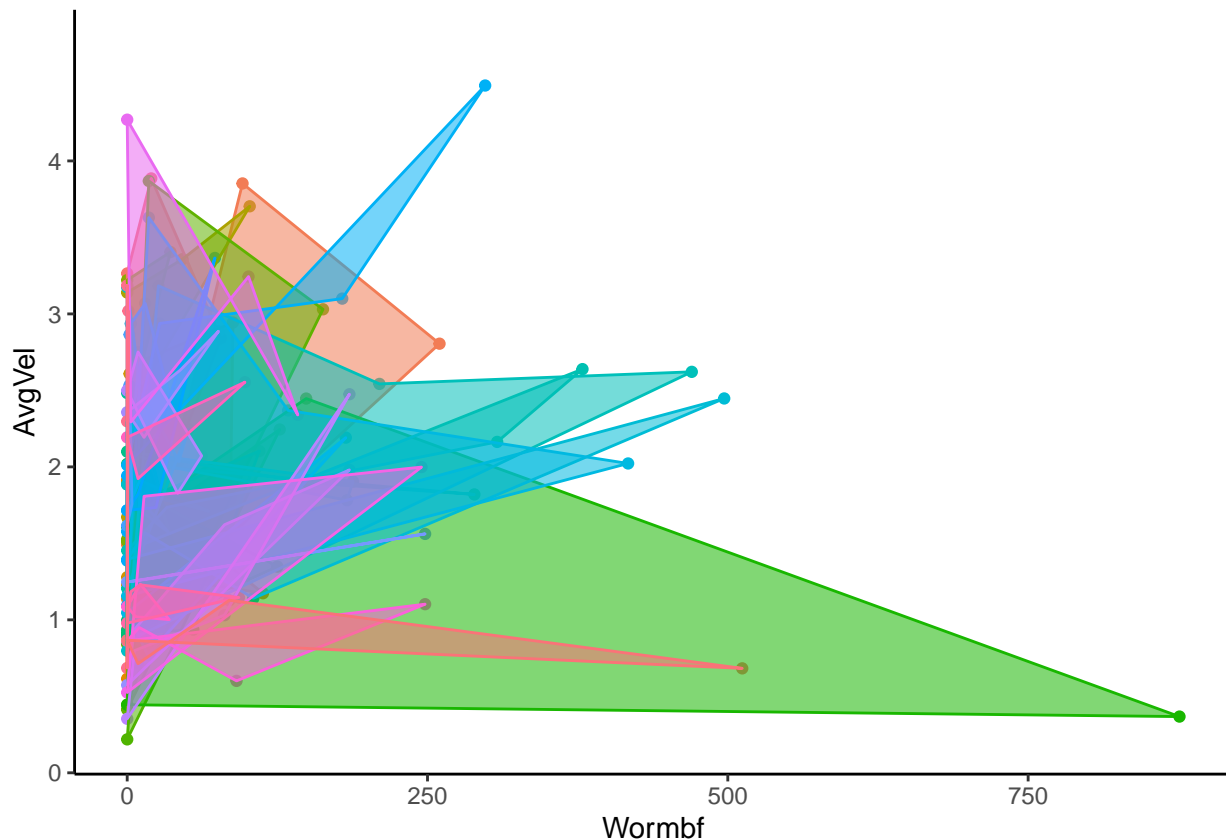


```
# Creating a dataframe that is missing the NAs for average velocity and Wormbf
# from fish who died/recovered during infection.
NonNABehav <- IndBehav8 %>%
    drop_na(c(AvgVel, Wormbf))
```

```
## drop_na: removed 166 rows (48%), 181 rows remaining
```

```
# Setting Levels as a character
Levels <- as.character(NonNABehav$fishID)
```

```r
# Subsetting down to only unique variables in fishID/Levels
Levels <- Levels %>%
    unique()



# Calculating the area of those polygons Creating empty dataframes and columsn
TotalAreas <- as.data.frame(Levels)
TotalAreas$PolyArea <- NA


# For Loop for calculating the area of the polygon of change in fish behavior
# with parasite ID
for (i in 1:length(TotalAreas$Levels)) {
    current_fish <- TotalAreas$Levels[i]

    Convex_hull <- NonNABehav %>%
        filter(fishID == current_fish) %>%
        select(AvgVel, Wormbf) %>%
        slice(chull(Wormbf, AvgVel))

    TotalAreas$PolyArea[i] <- as.matrix(Convex_hull) %>%
        areapl

    print(current_fish)

}
```

```
## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "1"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "3"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "5"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "11"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "13"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "46"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "49"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "56"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "57"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "71"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed 2 rows (50%), 2 rows remaining

## [1] "73"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "75"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "81"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "82"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining
```

```
## [1] "84"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "92"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed 2 rows (50%), 2 rows remaining

## [1] "95"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "112"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "113"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "114"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "115"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "132"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "134"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "141"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "145"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "6"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining
```

```
## [1] "7"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "18"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "20"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "51"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "52"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "53"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "54"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "55"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "77"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "80"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "86"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "88"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "89"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "90"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "106"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: removed one row (25%), 3 rows remaining

## [1] "108"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "110"

## filter: removed 177 rows (98%), 4 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed

## [1] "128"

## filter: removed 178 rows (98%), 3 rows remaining

## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)

## slice: no rows removed
```

```
## [1] "130"
```

```
## filter: removed 178 rows (98%), 3 rows remaining
```

```
## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)
```

```
## slice: no rows removed
```

```
## [1] "136"
```

```
## filter: removed 177 rows (98%), 4 rows remaining
```

```
## select: dropped 38 variables (fishID, TrialTime, AvgVelBef, VarvelBef, Sex, ...)
```

```
## slice: no rows removed
```

```
## [1] "148"
```

```r
# Renaming the Levels as fishID for merging
TotalAreas$fishID <- TotalAreas$Levels

# Removing the levels column
TotalAreas <- TotalAreas %>%
    select(-Levels)
```

```
## select: dropped one variable (Levels)
```

```r
# Merging the TotalAreas with our data frame
IndBehav8 <- merge(IndBehav8, TotalAreas, by.x = c("fishID"), by.y = c("fishID"),
    all = TRUE)
```

2.1 What hypotheses we want to test with these data and what data we can use to test them?

For uninfected and infected individuals only:

Do infected and uninfected individuals differ in their average and variation in velocity?

Is there sexual variation in the average and variation in velocity?

Do we see differences in the change in behavior and between infected

For infected individuals only:

Is there sexual variation in host behavioral tolerance and behavioral vigor?

Do host behavioral tolerance, tissue tolerance, and behavioral vigor correlate together and does this differ by sex?

Do hosts with lower behavioral tolerance have higher behavior resilience (i.e. how hosts behave once they have started to clear infection/have cleared infection)

Is there sexual variation in individuals behavioral resilience?

3 Do infected and uninfected individuals differ in their average and variation in velocity?

3.1 Visually inspection of the explanatory variables that will be used in the analyses

```
pairs(~AvgVel + VarVel + Sex + Infection + TrialTime, lower.panel = panel.smooth,
    diag.panel = panel.hist, upper.panel = panel.cor, data = IndBehav7)
```
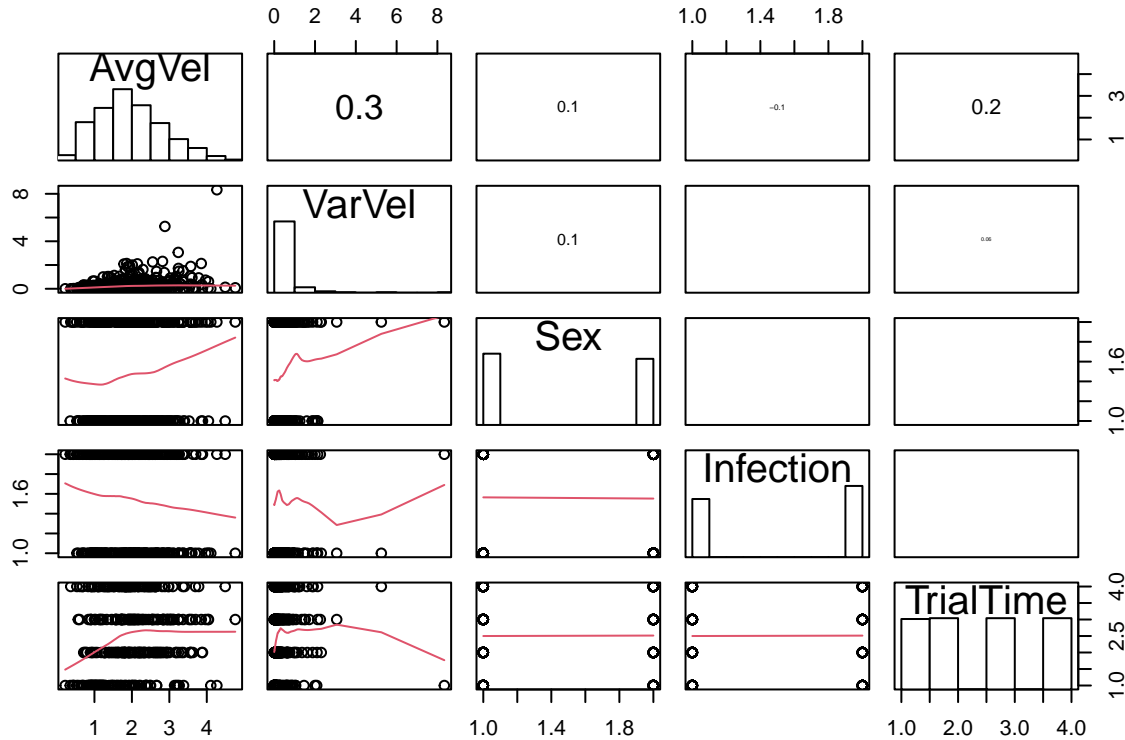

3.2 Does infection or sex variation impact the average velocity of individauls?

Individuals had 3 behavioral trials per time period of infection (i.e. 3 behavioral trials before infection) and therefore using preliminary analysis we showed that there is no difference due to time of day of these recordings so we averaged and quantified the variance of the velocities for that day to get an average velocity per trial time.

This analysis uses the average velocity for each individual at each trial point.

3.2.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how average velocity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic

- $AvgVel_{det} = \text{a} + b_1 \text{TrialTime} + b_2 \text{Infection} + b_3 \text{Sex} + a_i$

36

- Stochastic

    - $AvgVel \sim N(AvgVel_{det}, \sigma^2)$
    - $a_i \sim N(0, \sigma^2_{fishID})$

- Fixed

    - TrialTime
    - Infection
    - Sex

- Random

    - fishID

```r
# Fit a linear model for checking what explanatory factors are important for
# Average Velocity Note this is a linear mixed model because we have multiple
# measures per fish and therefore, need to account for non-independence between
# measures.
AvgVelLM <- lmer(AvgVel ~ TrialTime + Infection * Sex + ScPSMI + (1 | fishID), IndBehav8)



# Summary to see the relationship of the variables.
summary(AvgVelLM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: AvgVel ~ TrialTime + Infection * Sex + ScPSMI + (1 | fishID)
##    Data: IndBehav8
##
## REML criterion at convergence: 750.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4775 -0.5828 -0.1181  0.5605  3.2227
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  fishID   (Intercept) 0.2708   0.5204
##  Residual             0.3714   0.6095
## Number of obs: 336, groups:  fishID, 86
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      1.44475    0.14696   9.831
## TrialTimeEarly   0.60500    0.09294   6.510
## TrialTimeLate    0.74113    0.09294   7.974
## TrialTimeLater   0.38904    0.09586   4.059
## Infection1      -0.05602    0.18230  -0.307
## SexM             0.38490    0.19529   1.971
## ScPSMI          -0.15398    0.06698  -2.299
## Infection1:SexM -0.27055    0.26773  -1.011
##
## Correlation of Fixed Effects:
```

```
##             (Intr) TrlTmE TrilTmLt TrlTmLtr Infct1 SexM    ScPSMI
## TrialTmErly -0.316
## TrialTimeLt -0.316  0.500
## TrialTimLtr -0.316  0.485  0.485
## Infection1  -0.689  0.000  0.000     0.003
## SexM        -0.645  0.000  0.000     0.004    0.527
## ScPSMI      -0.081  0.000  0.000     0.000    0.118  0.139
## Infctn1:SxM  0.476  0.000  0.000     0.011   -0.695 -0.743 -0.197
```
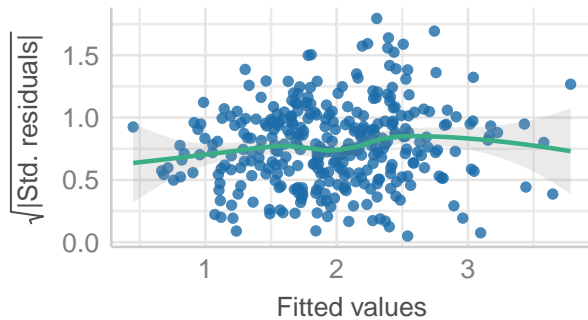
3.2.2 Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(AvgVelLM, check = c("qq", "normality", "homogeneity"))
```


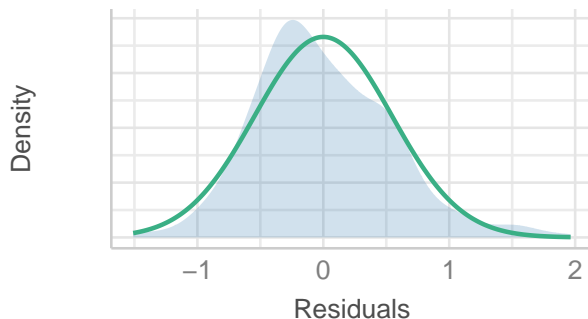**Homogeneity of Variance**
Reference line should be flat and horizontal

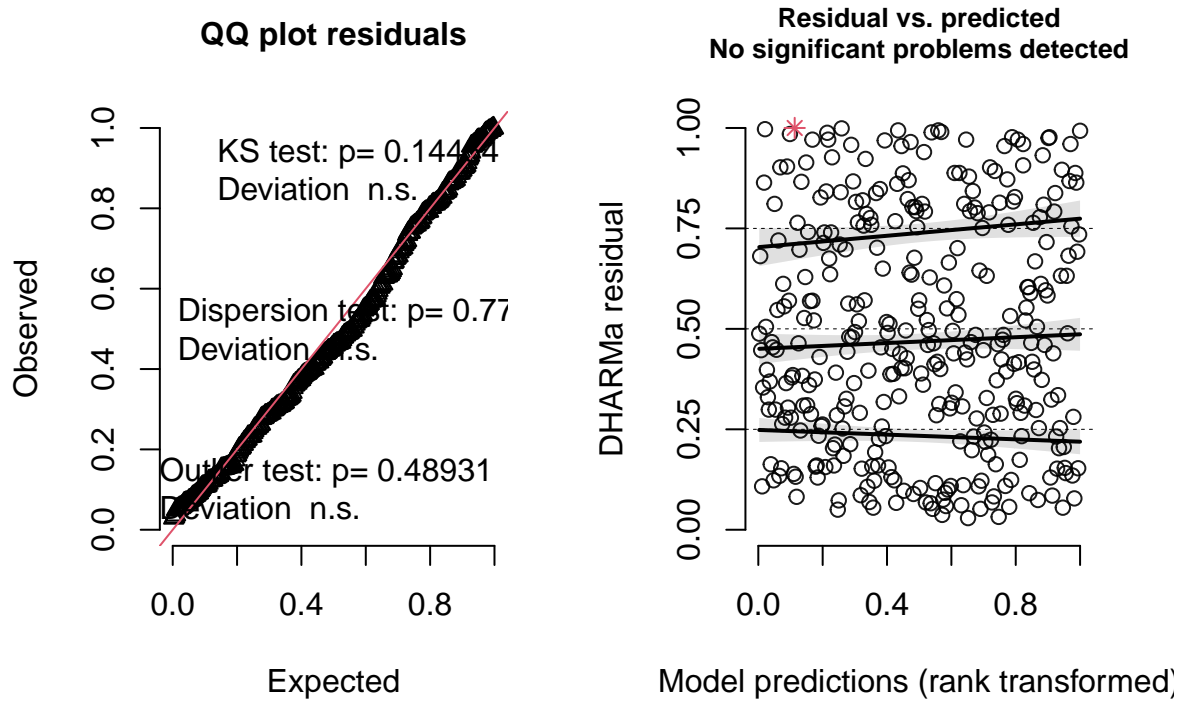**Normality of Residuals**
Dots should fall along the line
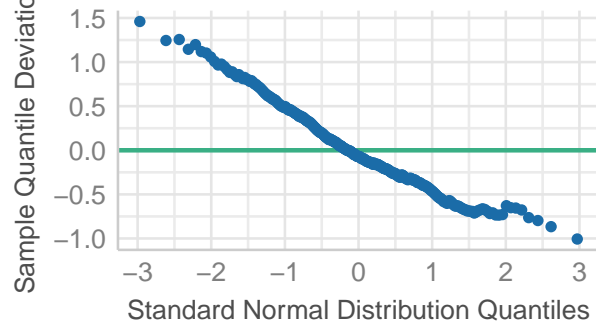
Normality of Residuals
Distribution should be close to the normal curve

```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_AvgVelLM <- simulateResiduals(AvgVelLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_AvgVelLM)
```

DHARMa residual

QQ plot residuals

KS test: p= 0.14444
Deviation n.s.

Dispersion test: p= 0.77
Deviation n.s.

Outlier test: p= 0.48931
Deviation n.s.

Observed

Expected

Residual vs. predicted
No significant problems detected

DHARMa residual

Model predictions (rank transformed)

```
# Testing for dispersion
testDispersion(sim_residuals_AvgVelLM)
```

DHARMa nonparametric dispersion test via sd of
residuals fitted vs. simulated

Frequency

Simulated values, red line = fitted model. p−value (two.sided) = 0.774
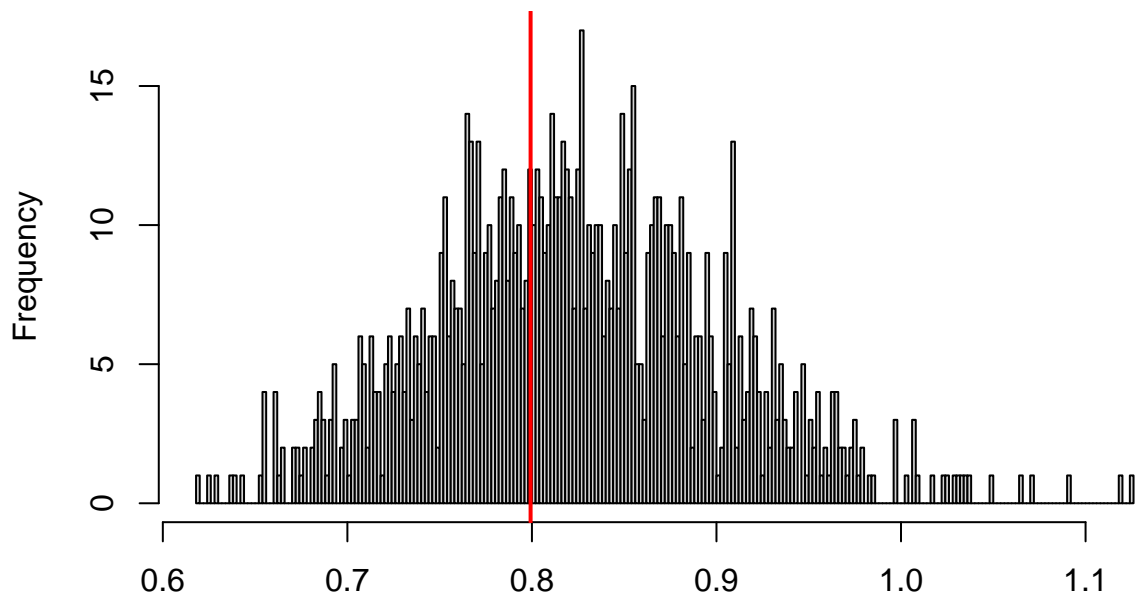
```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.9708, p-value = 0.774
## alternative hypothesis: two.sided
```

```
# All model validation looks good.
```

3.2.3   Testing the significance of factors in our model using a Kenward-Rodgers F test

```
# F test to test for signficance of slope of variables
Anova(AvgVelLM, test = "F", type = 3)
```

```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: AvgVel
##                     F Df  Df.res    Pr(>F)
## (Intercept)   96.6500  1 109.867 < 2.2e-16 ***
## TrialTime     24.2118  3 248.027 8.918e-14 ***
## Infection      0.0944  1  80.087   0.75941
## Sex            3.8843  1  80.131   0.05219 .
## ScPSMI         5.2849  1  81.168   0.02408 *
## Infection:Sex  1.0211  1  80.888   0.31526
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2.4   Visualize the important explanatory factors for average velocity

```
# TrialTimeGraph
AvgVelbyTT = visreg(AvgVelLM, scale = "response", "TrialTime", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Average Velocity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22)) + annotate("text", x = 0.25,
    y = 5, label = "p = 5.11e-10", size = 9, fontface = "italic")
```

```
## Conditions used in construction of plot
## Infection: 1
## Sex: F
## ScPSMI: -0.08630767
## fishID: 1
```

```
# Print the graph
print(AvgVelbyTT)
```
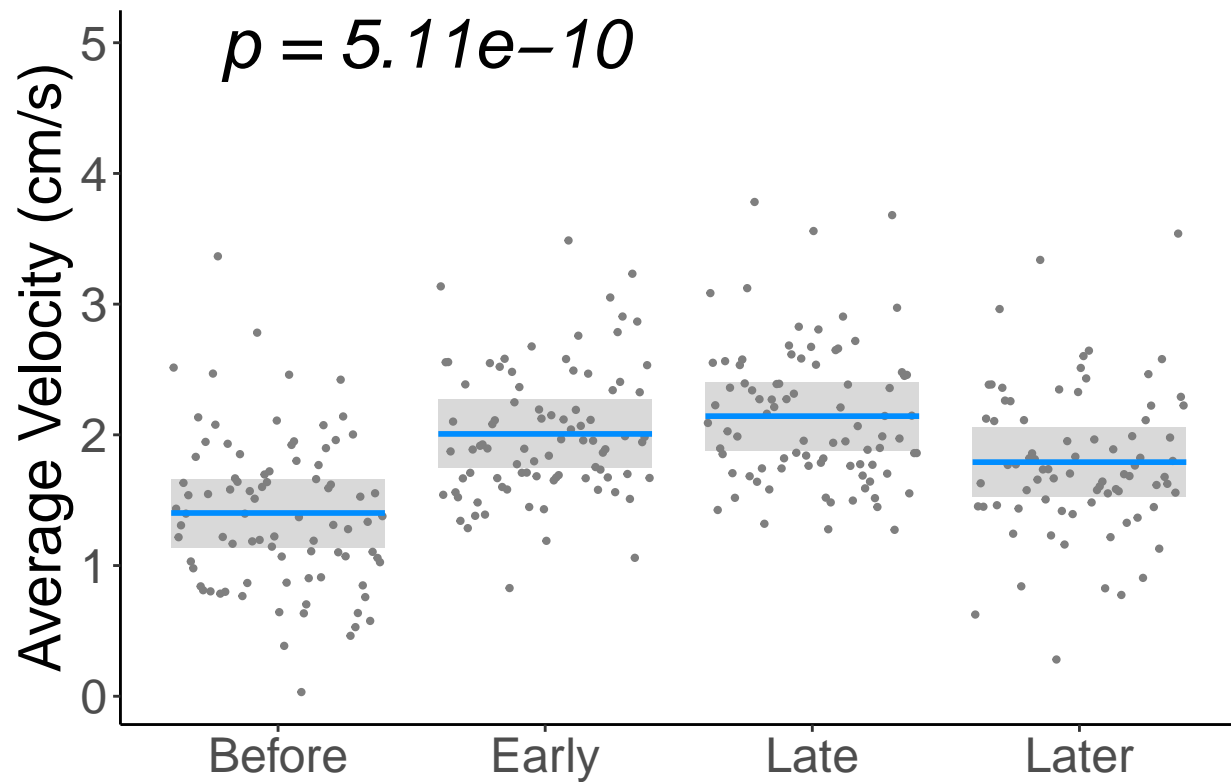
```
# TrialTimeGraph
AvgVelbySex = visreg(AvgVelLM, scale = "response", "Sex", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Average Velocity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22)) + annotate("text", x = 0.25,
    y = 5, label = "p = 0.027", size = 9, fontface = "italic")


## Conditions used in construction of plot
## TrialTime: Before
## Infection: 1
## ScPSMI: -0.08630767
## fishID: 1

# Print the graph
print(AvgVelbySex)
```
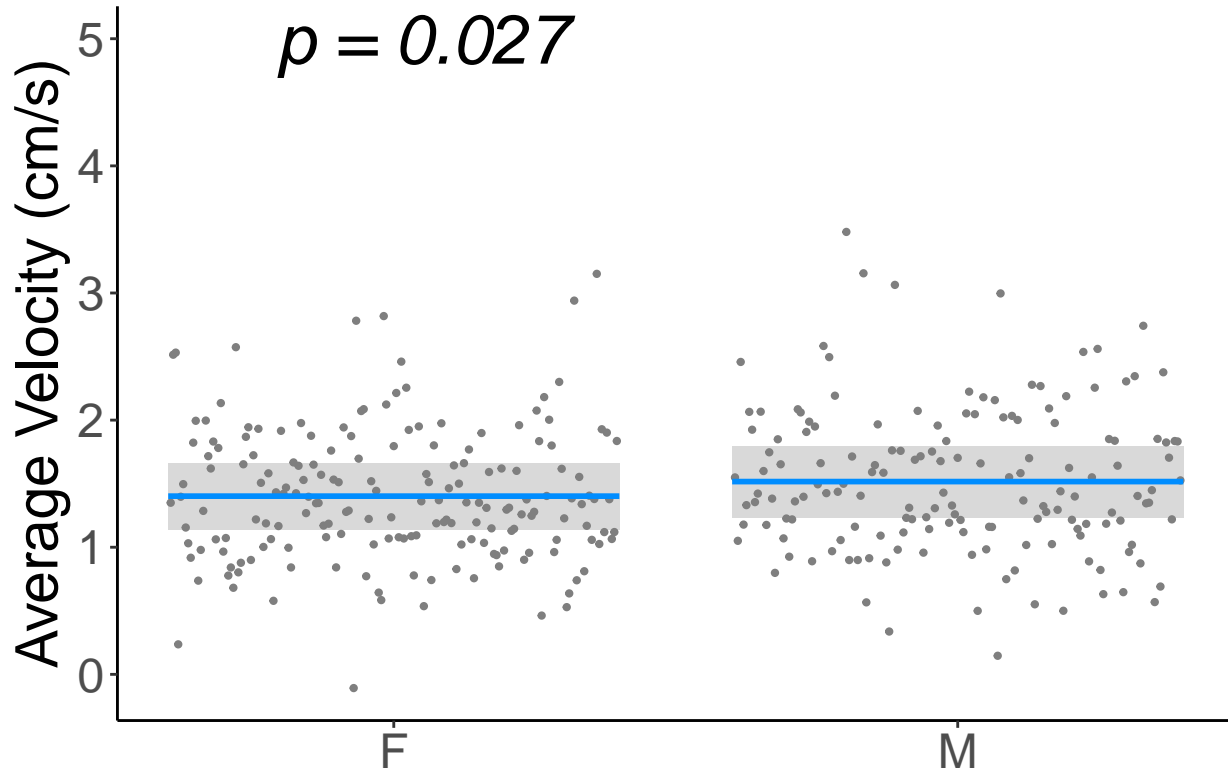
3.3 Does infection or sex variation impact the variance in velocity of individauls?

This analysis uses the variance of velocity for each individual at each trial point.

3.3.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how variance in velocity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic

- $VarVel_{det} = \text{a} + b_1\text{TrialTime} + b_2\text{Infection} + b_3\text{Sex} + a_i$

- Stochastic

  - VarVel $\sim N(VarVel_{det}, \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{fishID})$

- Fixed

  - TrialTime
  - Infection
  - Sex

- Random

– fishID

```
# Fit a linear model for checking what explanatory factors are important for
# Variance in Velocity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.
VarVelLM <- lmer(log(VarVel) ~ Sex * Infection + ScPSMI + TrialTime + (1 | fishID),
    IndBehav8)

# Summary to see the relationship of the variables. #Summary to see the
# relagaussian()tionship of the variables.
summary(VarVelLM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log(VarVel) ~ Sex * Infection + ScPSMI + TrialTime + (1 | fishID)
##    Data: IndBehav8
##
## REML criterion at convergence: 1450.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3850 -0.4833  0.1835  0.5924  2.7039
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  fishID    (Intercept) 0.1622   0.4028
##  Residual              4.3440   2.0842
## Number of obs: 334, groups:  fishID, 86
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      -3.8597     0.3191 -12.095
## SexM              0.7653     0.3625   2.111
## Infection1        0.2850     0.3397   0.839
## ScPSMI           -0.2070     0.1252  -1.653
## TrialTimeEarly    1.7040     0.3198   5.328
## TrialTimeLate     1.6879     0.3198   5.278
## TrialTimeLater    1.7019     0.3284   5.182
## SexM:Infection1  -0.8013     0.5004  -1.601
##
## Correlation of Fixed Effects:
##            (Intr) SexM   Infct1 ScPSMI TrlTmE TrilTmLt TrlTmLtr
## SexM       -0.551
## Infection1 -0.578  0.523
## ScPSMI     -0.068  0.138  0.120
## TrialTmErly -0.504  0.000 -0.010 -0.002
## TrialTimeLt -0.504  0.000 -0.010 -0.002  0.506
## TrialTimLtr -0.504  0.006 -0.005 -0.002  0.493  0.493
## SxM:Infctn1  0.395 -0.737 -0.693 -0.196  0.007  0.007    0.024
```
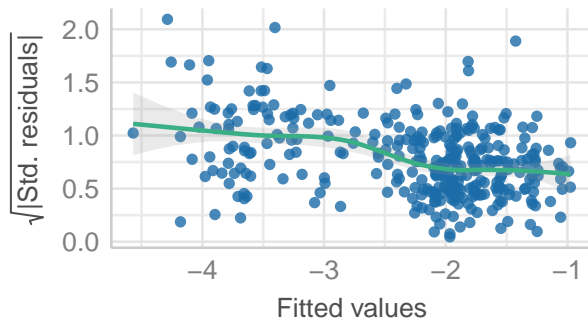
3.3.2 Validate that the model fits well and there are no problems

```r
# Using the check_model function from the performance package to check the
# model validation

check_model(VarVelLM, check = c("qq", "normality", "homogeneity"))
```
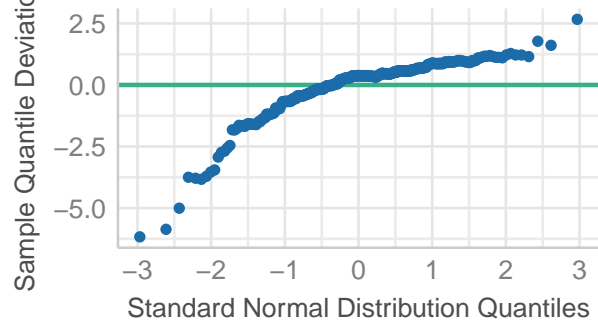

Homogeneity of Variance
Reference line should be flat and horizontal

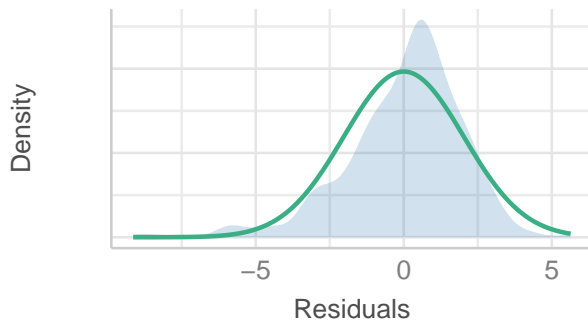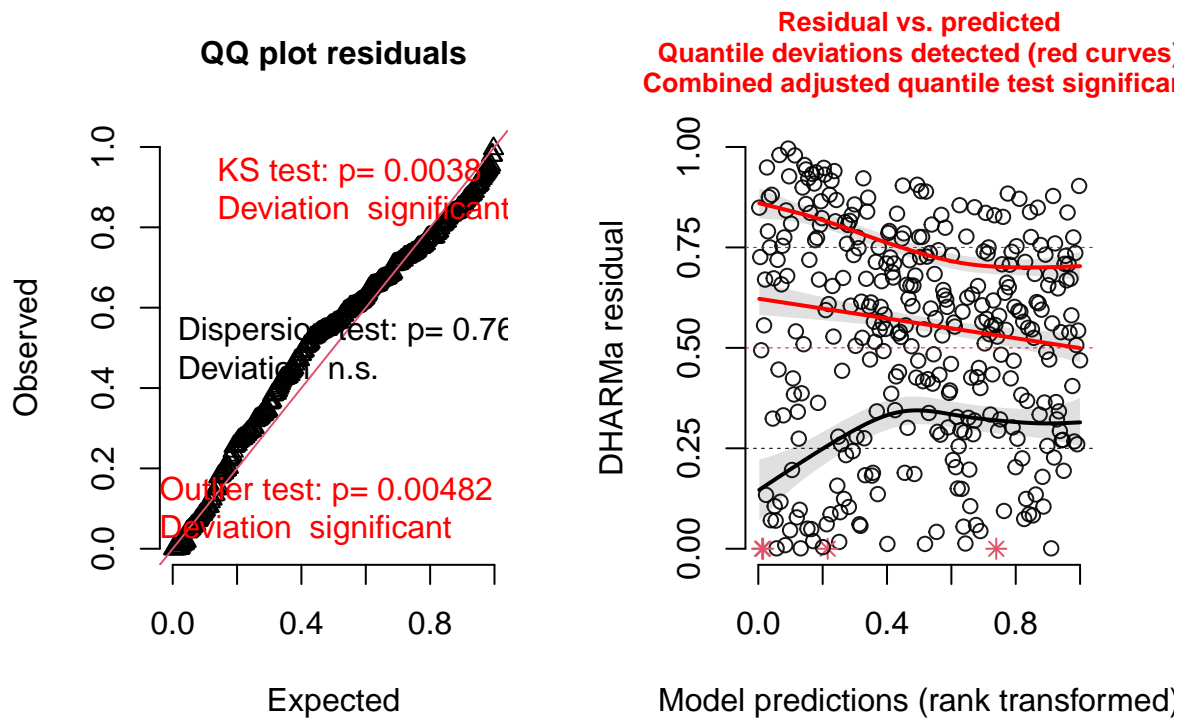Normality of Residuals
Dots should fall along the line

Normality of Residuals
Distribution should be close to the normal curve

```r
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_VarVelLM <- simulateResiduals(VarVelLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_VarVelLM)
```

DHARMa residual

QQ plot residuals

KS test: p= 0.0038
Deviation significant

Dispersion test: p= 0.76
Deviation n.s.

Outlier test: p= 0.00482
Deviation significant

Observed

Expected

**Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant**

DHARMa residual

Model predictions (rank transformed)

```
# Testing for dispersion
testDispersion(sim_residuals_VarVelLM)
```

DHARMa nonparametric dispersion test via sd of
residuals fitted vs. simulated
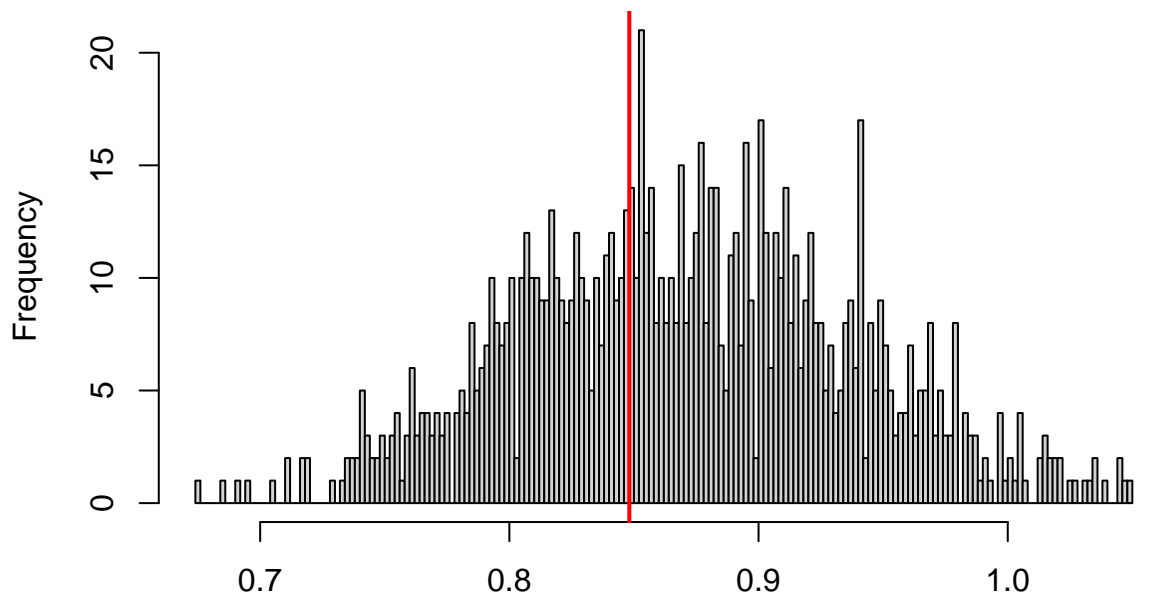
Frequency

Simulated values, red line = fitted model. p–value (two.sided) = 0.768

```
## 
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
## 
## data:  simulationOutput
## dispersion = 0.97468, p-value = 0.768
## alternative hypothesis: two.sided
```

```
# There are some problems with this model validation. It doesnt look model
# breaking but definitely should look at other model error structures to
# resolve the issues.
```

3.3.3 Testing the significance of factors in our model using a Kenward-Rodgers F test

```
# F test to test for signficance of slope of variables
Anova(VarVelLM, test = "F", type = 3)
```

```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
## 
## Response: log(VarVel)
##                       F Df  Df.res      Pr(>F)    
## (Intercept)   146.2746  1 180.904 < 2.2e-16 ***
## Sex             4.4567  1  77.804   0.03797 *  
## Infection       0.7035  1  78.772   0.40416    
## ScPSMI          2.7325  1  80.497   0.10222    
## TrialTime      13.8831  3 248.622 2.127e-08 ***
## Sex:Infection   2.5631  1  80.261   0.11332    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.3.4 Visualize the important explanatory factors for variance of velocity

```
# VariancebySex
VarVelbySex <- visreg(VarVelLM, scale = "response", "Sex", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Variance in Velocity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22)) + annotate("text", x = 0.25,
    y = 5, label = "p = 0.016", size = 9, fontface = "italic")
```

```
## Conditions used in construction of plot
## Infection: 1
## ScPSMI: -0.08630767
## TrialTime: Early
## fishID: 1
```

```
# Print the graph
print(VarVelbySex)
```
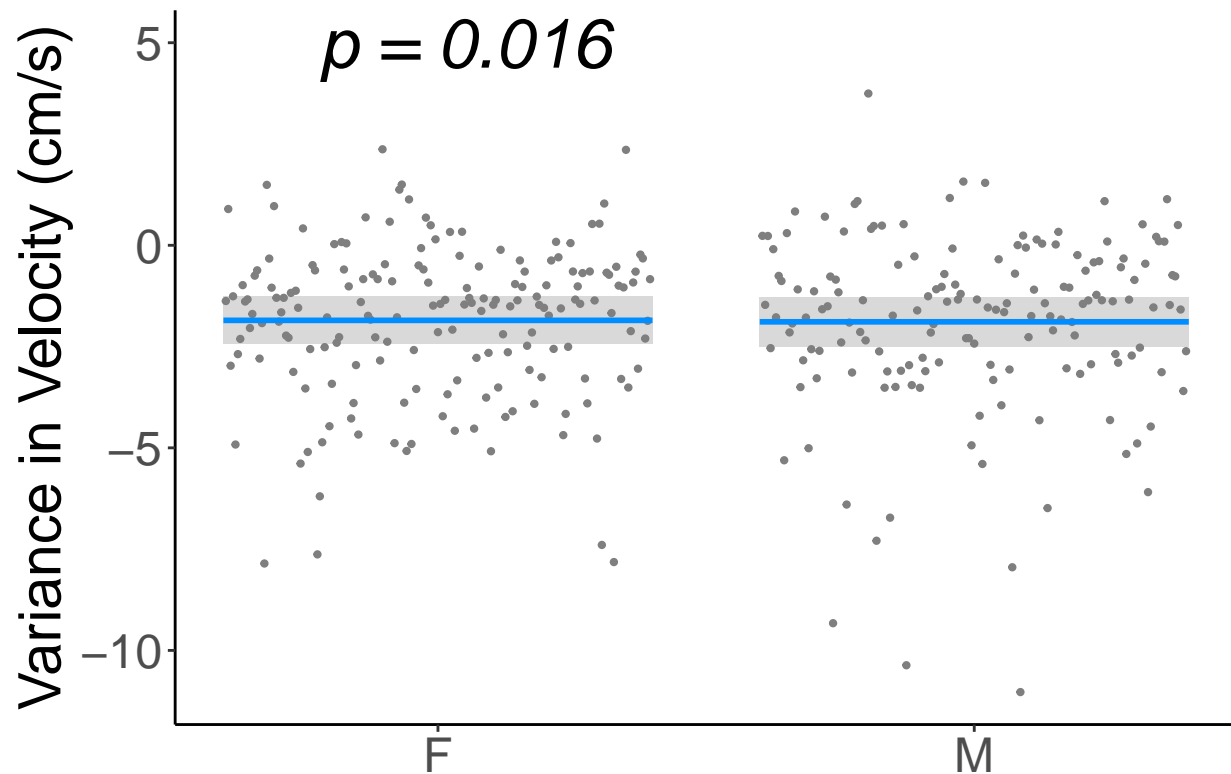
```
# #Variance by sex and split by infection Infnames<-c('Uninfected','Infected')
# names(Infnames)<-c('0','1') VarVelbySexnInf<-visreg(VarVelLM,
# scale='response', 'Sex', 'Infection', partial=T, gg=TRUE)+ theme_classic()+
# theme(legend.position='none')+ ylab('Variance in Velocity (cm/s)')+ xlab('
# ')+ theme(text=element_text(size=22))+ annotate('text',x=0.25, y=5, label =
# 'p = 0.018', size=9, fontface = 'italic')+ facet_wrap(~Infection, labeller =
# labeller(Infection = Infnames)) #Print the graph print(VarVelbySexnInf)
```

4 Do we see differences in change in behavior over time based on infection status and sexual variation?

4.1 Visually inspection of the explanatory variables that will be used in the analyses

```
pairs(~ChBehav + Sex + Infection + TrialTime + PreSMI + AUC2, lower.panel = panel.smooth,
    diag.panel = panel.hist, upper.panel = panel.cor, data = IndBehav7)
```
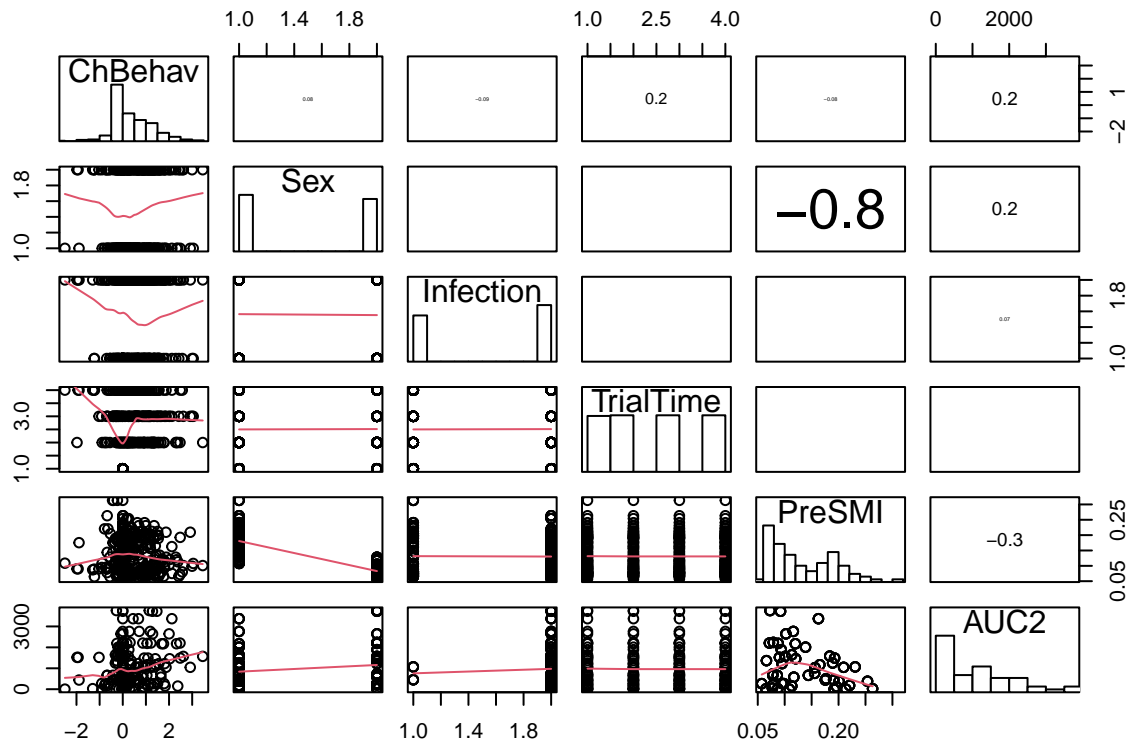
4.1.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how Change in velocity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic

- $ChVel_{det} = \text{a} + b_1\text{TrialTime} + b_2\text{Infection} + b_3\text{Sex} + a_i$

- Stochastic

  - ChVel ~ $N(ChVel_{det}, \sigma^2)$
  - a_i ~ $N(0, \sigma^2_{fishID})$

- Fixed

  - TrialTime
  - Infection
  - Sex

- Random

  - fishID

```
# Fit a linear model for checking what explanatory factors are important for
# Variance in Velocity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.
ChVelLM <- lmer(ChBehav ~ TrialTime + Infection * Sex + ScChSMI + (1 | fishID), IndBehav8)
```

```
# Summary to see the relationship of the variables.
summary(ChVelLM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ChBehav ~ TrialTime + Infection * Sex + ScChSMI + (1 | fishID)
##     Data: IndBehav8
##
## REML criterion at convergence: 740.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1435 -0.5500 -0.1007  0.5788  2.9930
##
## Random effects:
##  Groups    Name         Variance Std.Dev.
##  fishID    (Intercept) 0.2258   0.4752
##  Residual              0.3724   0.6103
## Number of obs: 336, groups:  fishID, 86
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       0.03070    0.14033   0.219
## TrialTimeEarly    0.60500    0.09307   6.501
## TrialTimeLate     0.74113    0.09307   7.964
## TrialTimeLater    0.37033    0.09596   3.859
## Infection1       -0.16842    0.17242  -0.977
## SexM              0.10002    0.18376   0.544
## ScChSMI          -0.04578    0.06203  -0.738
## Infection1:SexM   0.05362    0.24924   0.215
##
## Correlation of Fixed Effects:
##            (Intr) TrlTmE TrilTmLt TrlTmLtr Infct1 SexM   ScCSMI
## TrialTmErly -0.332
## TrialTimeLt -0.332  0.500
## TrialTimLtr -0.332  0.485  0.485
## Infection1  -0.688  0.000  0.000    0.004
## SexM        -0.645  0.000  0.000    0.005    0.534
## ScChSMI     -0.154  0.000  0.000    0.002    0.181  0.168
## Infctn1:SxM  0.479  0.000  0.000    0.011   -0.698 -0.744 -0.162
```
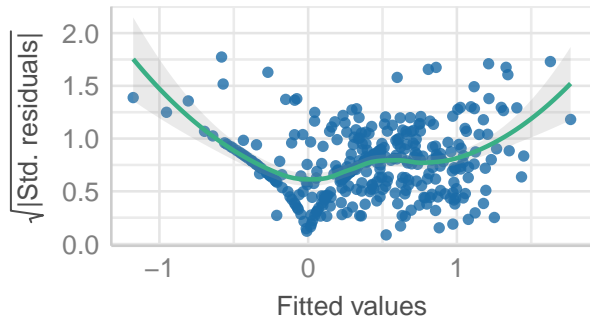
4.1.2 Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(ChVelLM, check = c("qq", "normality", "homogeneity"))
```
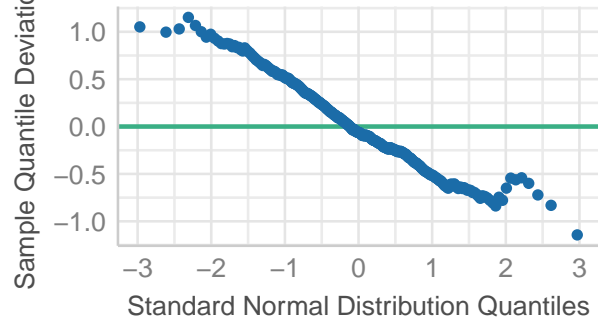
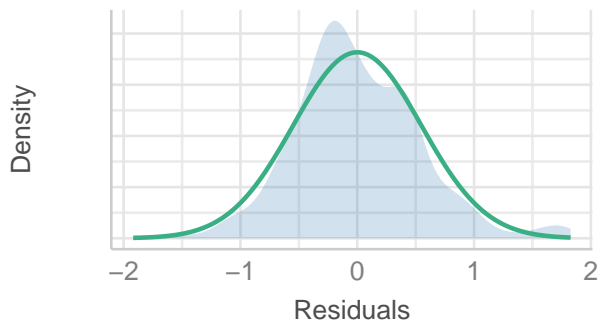## Homogeneity of Variance
Reference line should be flat and horizontal

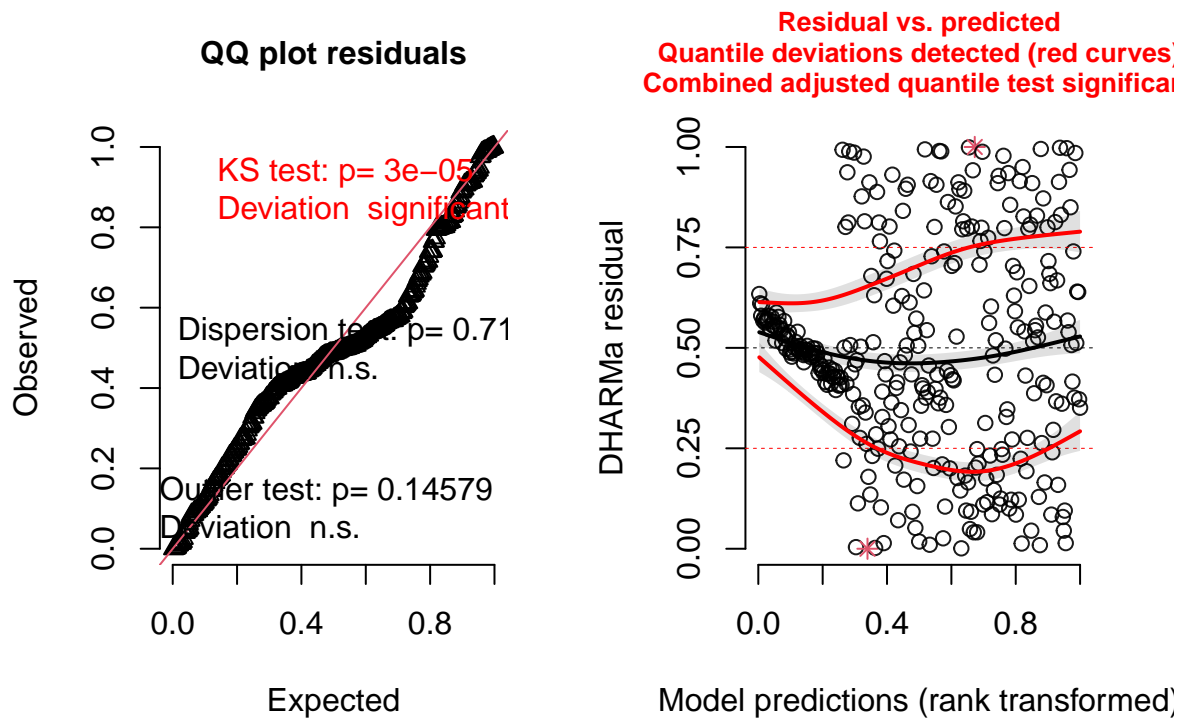## Normality of Residuals
Dots should fall along the line

Normality of Residuals
Distribution should be close to the normal curve


```r
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_ChVelLM <- simulateResiduals(ChVelLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_ChVelLM)
```

DHARMa residual

QQ plot residuals

KS test: p= 3e−05
Deviation significant

Dispersion test: p= 0.71
Deviation n.s.

Outlier test: p= 0.14579
Deviation n.s.

Observed / Expected

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant

DHARMa residual / Model predictions (rank transformed)

```r
# Testing for dispersion
testDispersion(sim_residuals_ChVelLM)
```

DHARMa nonparametric dispersion test via sd of residuals fitted vs. simulated
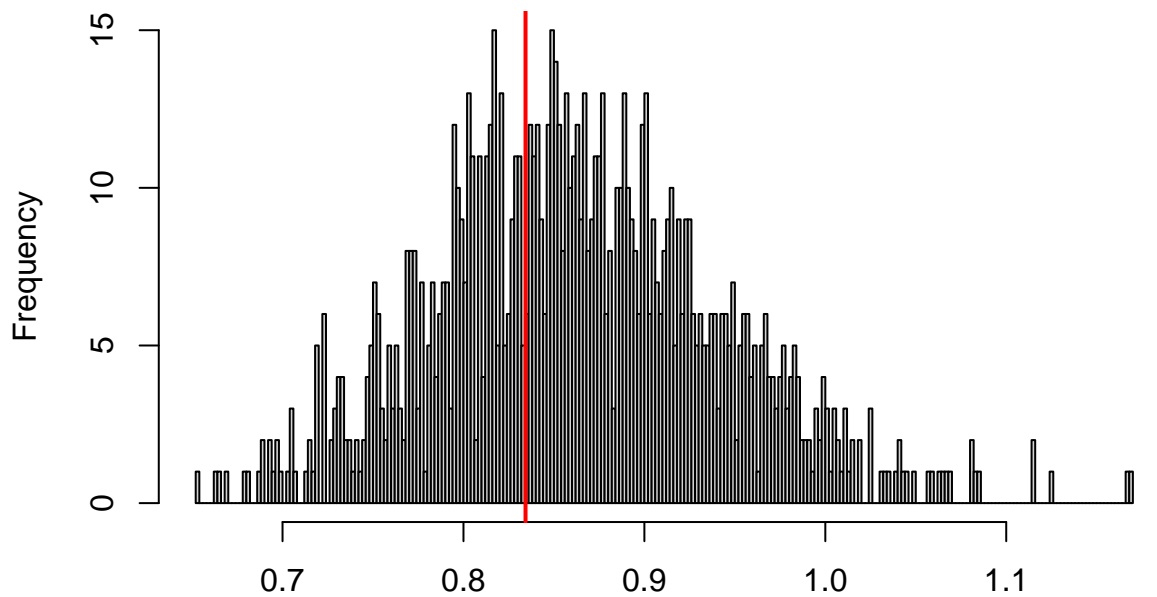
Frequency

Simulated values, red line = fitted model. p−value (two.sided) = 0.718

```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.96583, p-value = 0.718
## alternative hypothesis: two.sided
```

```
# There are some problems with this model validation. It doesnt look model
# breaking but definitely should look at other model error structures to
# resolve the issues.
```

4.1.3 Testing the significance of factors in our model using a Kenward-Rodgers F test

```
# F test to test for signficance of slope of variables
Anova(ChVelLM, test = "F", type = 3)
```

```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: ChBehav
##                     F Df  Df.res    Pr(>F)
## (Intercept)    0.0479  1 113.402    0.8272
## TrialTime     24.2837  3 248.160 8.197e-14 ***
## Infection      0.9542  1  79.952    0.3316
## Sex            0.2962  1  80.031    0.5878
## ScChSMI        0.5447  1  79.957    0.4627
## Infection:Sex  0.0463  1  80.875    0.8302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.1.4 Visualize the important explanatory factors for change in velocity

```
# TrialTimeGraph
ChVelbyTT = visreg(ChVelLM, scale = "response", "TrialTime", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Change in Velocity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22)) + annotate("text", x = 0.25,
    y = 5, label = "p = 4.12e-10", size = 9, fontface = "italic")
```

```
## Conditions used in construction of plot
## Infection: 1
## Sex: F
## ScChSMI: -0.1085622
## fishID: 1
```

```
# Print the graph
print(ChVelbyTT)
```

$p = 4.12e{-}10$
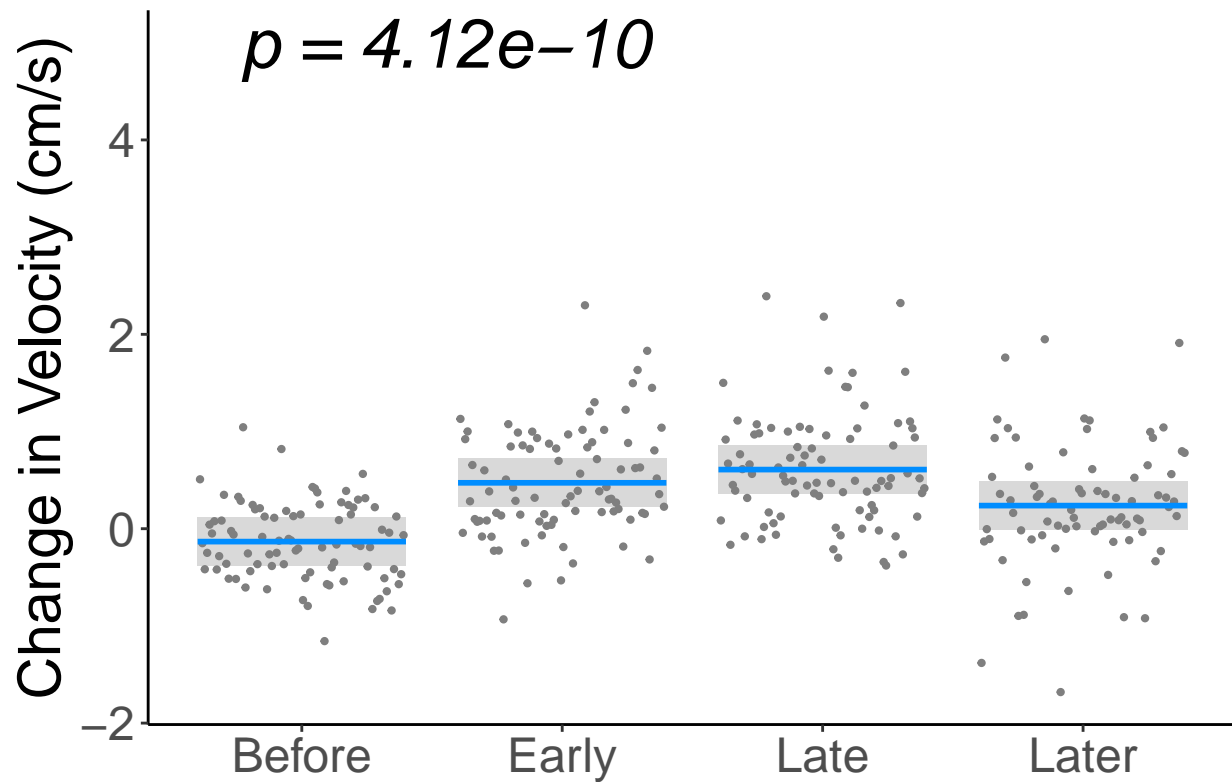
5 Is there sexual variation in host behavioral tolerance and behavioral vigor?

5.1 Visually inspection of the explanatory variables that will be used in the analyses