

# KJPipho\_PS\_8

Krista Pipho

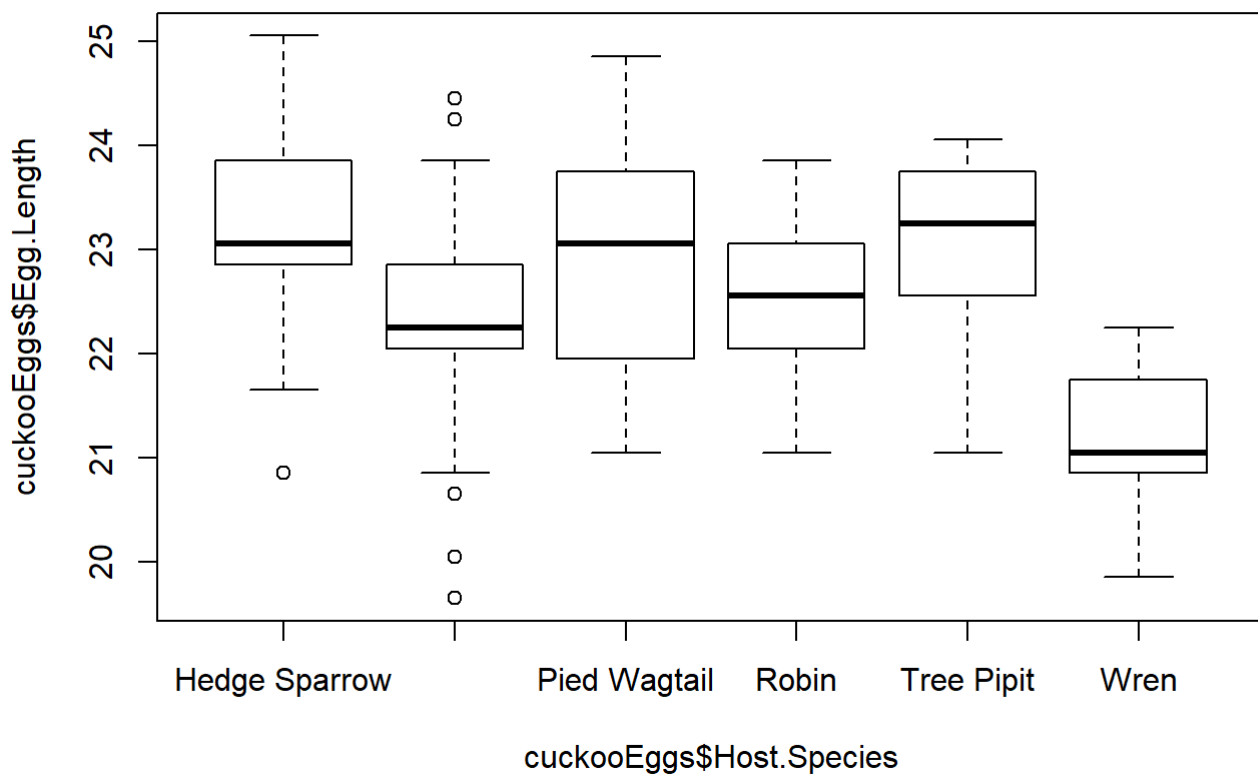
October 30, 2018

Question 1. Do cuckoo birds lay different size eggs in the nests of different bird species?

```
# First we load in the cuckooeggs data set and save it as cuckooEggs
cuckooEggs <- read.csv("C:\\Users\\xenon\\Desktop\\R Studio 2018\\cuckooeggs.csv")
# We print out the titles of the data variables
names(cuckooEggs)
```

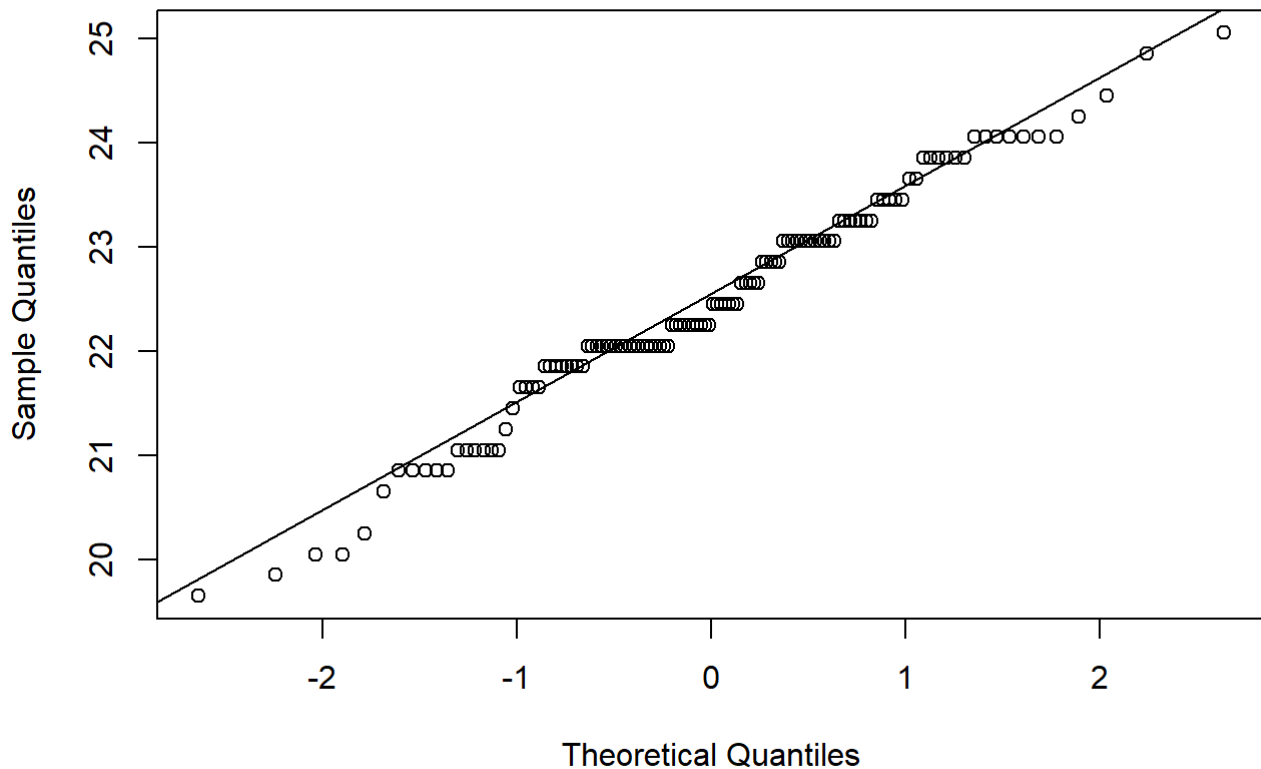
```
## [1] "Host.Species" "Egg.Length"
```

```
# Here we make a box plot of egg length by host species
plot(cuckooEggs$Egg.Length~cuckooEggs$Host.Species)
```



```
# Nothing looks crazy wrong with the box plot, we can check a qqnorm plot to better visualize whether the data is normal
qqnorm(cuckooEggs$Egg.Length)
qqline(cuckooEggs$Egg.Length)
```

## Normal Q-Q Plot



```
# The data appears quite normal, let us formally check using the shapiro test
stestCE <- shapiro.test(cuckooEggs$Egg.Length)
stestCE
```

```
##
##  Shapiro-Wilk normality test
##
## data:  cuckooEggs$Egg.Length
## W = 0.98241, p-value = 0.1193
```

```
# The p value is not low enough to reject that the data is normal. Let us test weather the va
riance is equal
btestCE <- bartlett.test(cuckooEggs$Egg.Length~cuckooEggs$Host.Species)
btestCE
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  cuckooEggs$Egg.Length by cuckooEggs$Host.Species
## Bartlett's K-squared = 4.4794, df = 5, p-value = 0.4826
```

```
# The p value is not low enough to reject that the variances are equal. Since both the assump
tion of normality and the assumption of equal variance are met we can procede with the ANOVA
analysis
cuckooEggsAnova <- aov(cuckooEggs$Egg.Length~cuckooEggs$Host.Species)
cuckooEggsAnova
```

```
## Call:
##   aov(formula = cuckooEggs$Egg.Length ~ cuckooEggs$Host.Species)
##
## Terms:
##               cuckooEggs$Host.Species Residuals
## Sum of Squares              42.93965   94.24835
## Deg. of Freedom                5       114
##
## Residual standard error: 0.9092524
## Estimated effects may be unbalanced
```

```
summary(cuckooEggsAnova)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## cuckooEggs$Host.Species    5  42.94    8.588   10.39 3.15e-08 ***
## Residuals              114  94.25    0.827
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(cuckooEggsAnova)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = cuckooEggs$Egg.Length ~ cuckooEggs$Host.Species)
##
## $`cuckooEggs$Host.Species`
##              diff              lwr              upr      p adj
## Meadow Pipit-Hedge Sparrow -0.82253968 -1.629133605 -0.01594576 0.0428621
## Pied Wagtail-Hedge Sparrow -0.21809524 -1.197559436  0.76136896 0.9872190
## Robin-Hedge Sparrow        -0.54642857 -1.511003196  0.41814605 0.5726153
## Tree Pipit-Hedge Sparrow   -0.03142857 -1.010892769  0.94803563 0.9999990
## Wren-Hedge Sparrow         -1.99142857 -2.970892769 -1.01196437 0.0000006
## Pied Wagtail-Meadow Pipit   0.60444444 -0.181375330  1.39026422 0.2324603
## Robin-Meadow Pipit         0.27611111 -0.491069969  1.04329219 0.9021876
## Tree Pipit-Meadow Pipit     0.79111111  0.005291337  1.57693089 0.0474619
## Wren-Meadow Pipit          -1.16888889 -1.954708663 -0.38306911 0.0004861
## Robin-Pied Wagtail         -0.32833333 -1.275604766  0.61893810 0.9155004
## Tree Pipit-Pied Wagtail     0.18666667 -0.775762072  1.14909541 0.9932186
## Wren-Pied Wagtail          -1.77333333 -2.735762072 -0.81090459 0.0000070
## Tree Pipit-Robin           0.51500000 -0.432271433  1.46227143 0.6159630
## Wren-Robin                 -1.44500000 -2.392271433 -0.49772857 0.0003183
## Wren-Tree Pipit            -1.96000000 -2.922428738 -0.99757126 0.0000006
```

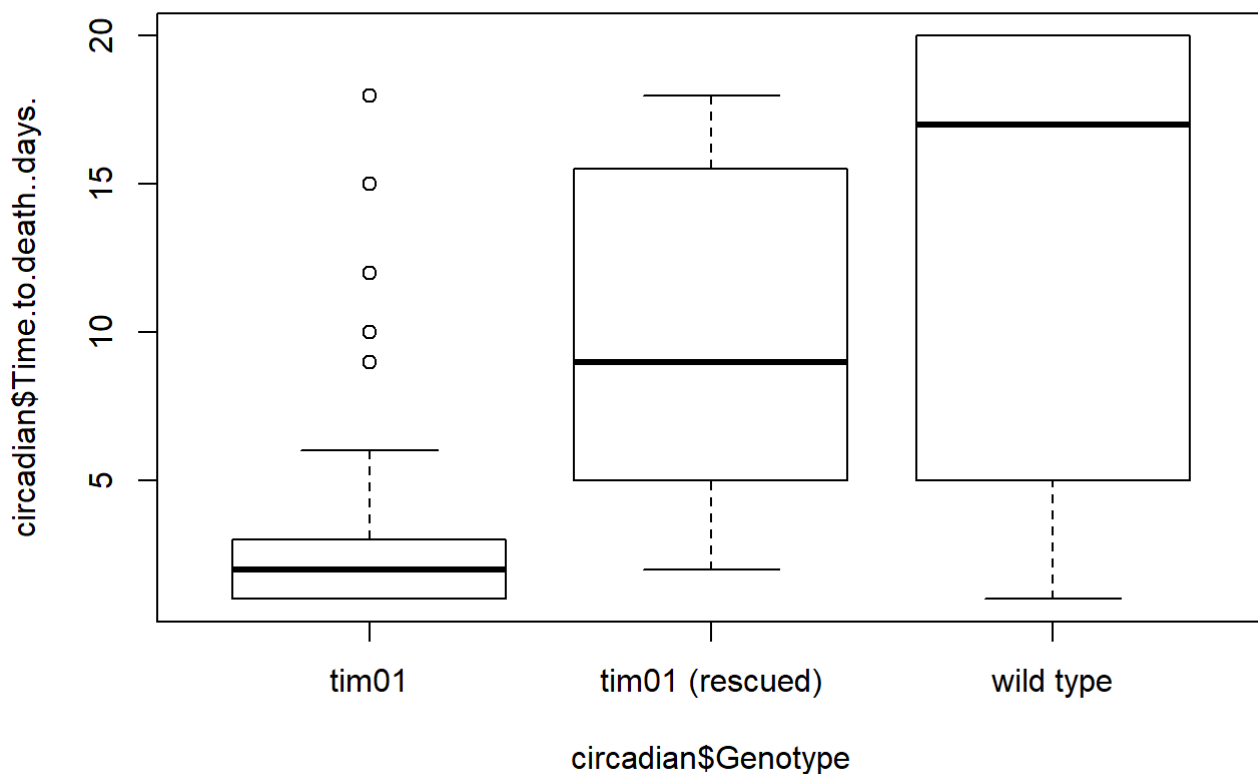
We can see statistically significant differences in egg length between the eggs laid in Meadow Pipit vs Hedge Sparrow nests, between Wren and Hedge Sparrow nests, between tree pipit and meadow pipit nests, between wren and meadow pipit nests, between wren and pied wagtail nests, between wren and robin nests, and between wren and tree pipit nests. This shows that there are many instances of egg size variance depending on host species.

Question 2. Can circadian disruptions affect robustness to infections?

```
# First we load in the circadian mutant health data set and save it as circadian
circadian<-read.csv("C:\\Users\\xenon\\Desktop\\R Studio 2018\\circadian mutant health.csv")
# Next we print out the variable names for the data set
names(circadian)
```

```
## [1] "Genotype"          "Time.to.death..days."
```

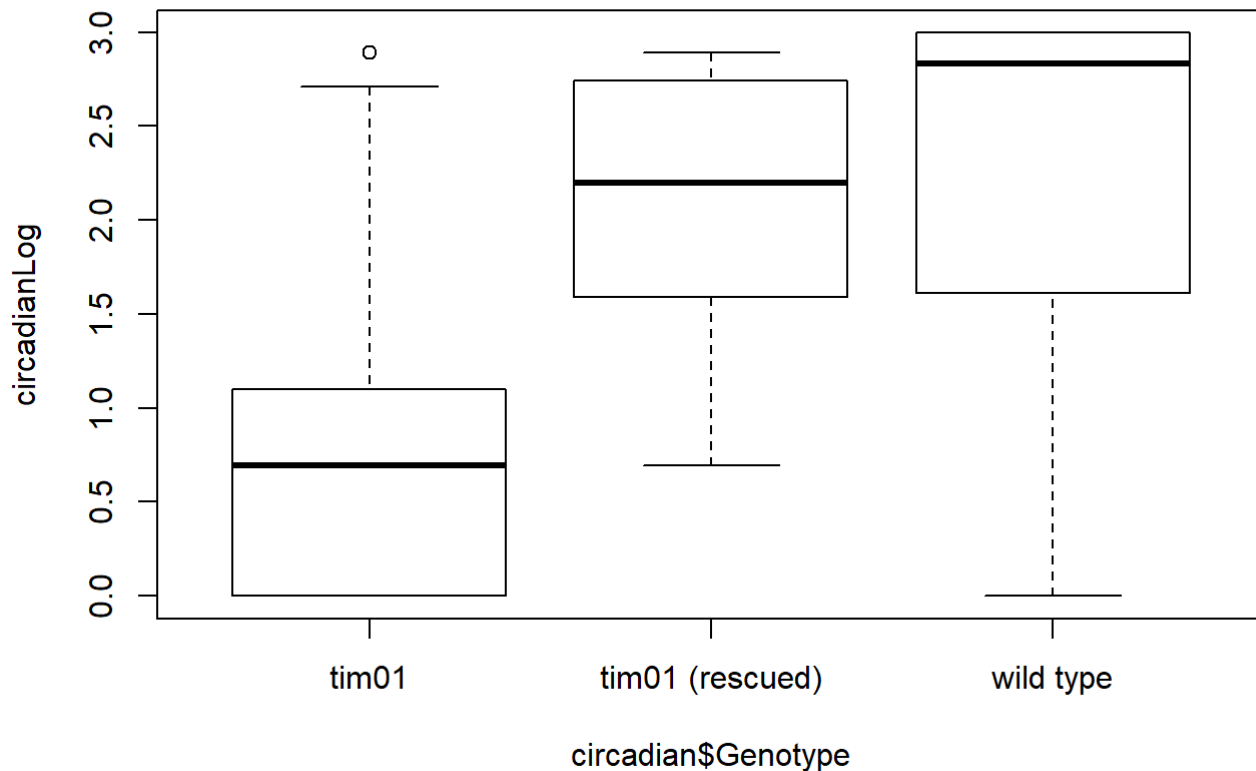
```
# Then we do an initial visualization via boxplot of survival time by mutation condition
plot(circadian$Time.to.death..days.~circadian$Genotype)
```



```
# There appear to be low levels of outlier and differences in mean, but Lets check if the data is normal
stestCIRC <- shapiro.test(circadian$Time.to.death..days.)
stestCIRC
```

```
##
## Shapiro-Wilk normality test
##
## data: circadian$Time.to.death..days.
## W = 0.84582, p-value = 1.072e-09
```

```
# The data is unfortunately not normal, so Lets log transform it and retest for normality
circadianLog <- log(circadian$Time.to.death..days.)
plot(circadianLog~circadian$Genotype)
```



```
shapiro.test(circadianLog)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  circadianLog
## W = 0.88623, p-value = 5.589e-08
```

```
# The data is still not normal. What if we transform with sqrt +.5 transformation and retest?
circadianRoot<-sqrt(circadian$Time.to.death..days.+0.5)
stestROOTCIRC <- shapiro.test(circadianRoot)
stestROOTCIRC
```

```
##
##  Shapiro-Wilk normality test
##
## data:  circadianRoot
## W = 0.87908, p-value = 2.62e-08
```

```
# The data is stilllll not normal. SO we will abandon ANOVA as an analysis option, and use the
kruskal wallis test instead.
kwTestCircadian <- kruskal.test(circadian$Time.to.death..days.~circadian$Genotype)
kwTestCircadian
```

```
##  
## Kruskal-Wallis rank sum test  
##  
## data: circadian$Time.to.death..days. by circadian$Genotype  
## Kruskal-Wallis chi-squared = 41.736, df = 2, p-value = 8.653e-10
```

```
summary(kwTestCircadian)
```

```
##           Length Class  Mode  
## statistic 1      -none- numeric  
## parameter 1      -none- numeric  
## p.value    1      -none- numeric  
## method     1      -none- character  
## data.name  1      -none- character
```

Using the Kruskal Wallis test we find with high certainty (Very low p value) that at least one of the groups has a significantly different mean. With a one way test we cannot tell which group or groups is different, only that at least one is. Regardless, we can now say that circadian rhythm health can directly exert an effect on survival of infections in fruit flies.