Study on Fruit Flies

Eric Chen, Valeria Martinez, Domenic Hoffman, Sidhart Sathya, Mark Combs

2023-11-01

Introduction

The birds and the bees... well the flies to be exact. Fruit flies are that pesky little animal that most people want out of their home. For researchers, they serve as a crucial part in research. Fruit flies (drosophila melanogaster) are one of the most extensively studied organisms in biological research. They are most commonly used in biomedial research and allow us to better understand the human condition.

Literature Review

In 1981, researchers Linda Partridge and Marion Farquhar from the University of Edinburgh decided to conduct a study on the effect of sexual activity of fruit flies on their lifespan. Since the cost of reproduction had already been determined to reduce longevity in females, the researchers sought to demonstrate the same effect for males (Partridge & Farquhar, 1981).

Due to its small size and short life cycle, it is easy to measure and compare the life spans for D. Melanogaster. Furthermore, D. Melanogaster shares multiple similarities with mammals with respect to aging. In the cellular and molecular levels, D. melanogaster and mammals both demonstrate deterioration of muscle and nerve tissue as well as accumulation of damaged DNA over time (Bhode & Tower, 2002). These similarities allow age and lifespan-related research on D. Melanogaster to be applied to mammalian and possibly human implications.

Methodology

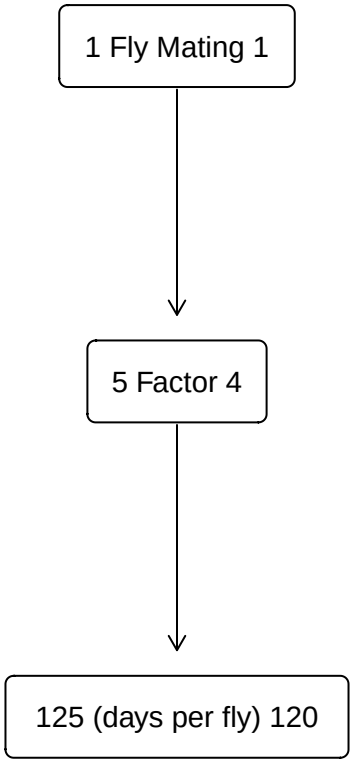
The data used in this report was from the study originally conducted by Partridge and Farquhar. The method of data collection hence is the same. The following is an excerpt from the original study:

"The flies used were an outbred stock collected in Dahomey in 1970. Sexual activity was manipulated by supplying individual males with receptive virgin females at a rate of one or eight virgins per day. The longevity of these males was recorded and compared with that of two control types. The first control consisted of two sets of individual males kept with newly inseminated females equal in number to the virgin females supplied to the experimental males. Newly inseminated females will not usually re-mate for at least 2 days 10' 11 thus they served as a control for any effect of competition with the male for food or space. The second control was a set of individual males kept with no females. There were 25 males in each of the experimental and control groups, and the groups were treated identically in respect of number of anaesthetizations (using C02) and provision of fresh food medium" (Partridge & Farquhar, 1981).

For analysis, we will use parametric shortcuts for testing the grand sample arithmetic mean (GSAM). In post-hoc analysis, we will consider all possible pairwise comparisons. Using Tukey's HSD to resolve the multiple comparison issues with an overall Type I Error risk of 0.01.

This report will utilize an Unusualness Threshold of 1%, meaning that if we observe an event that occurs no more than 1% of the time under the assumption of the null model, we will reject the null model.

Data Exploration



The above Hasse diagram illustrates a breakdown of what the data looks like as well as details on the degrees of freedom.

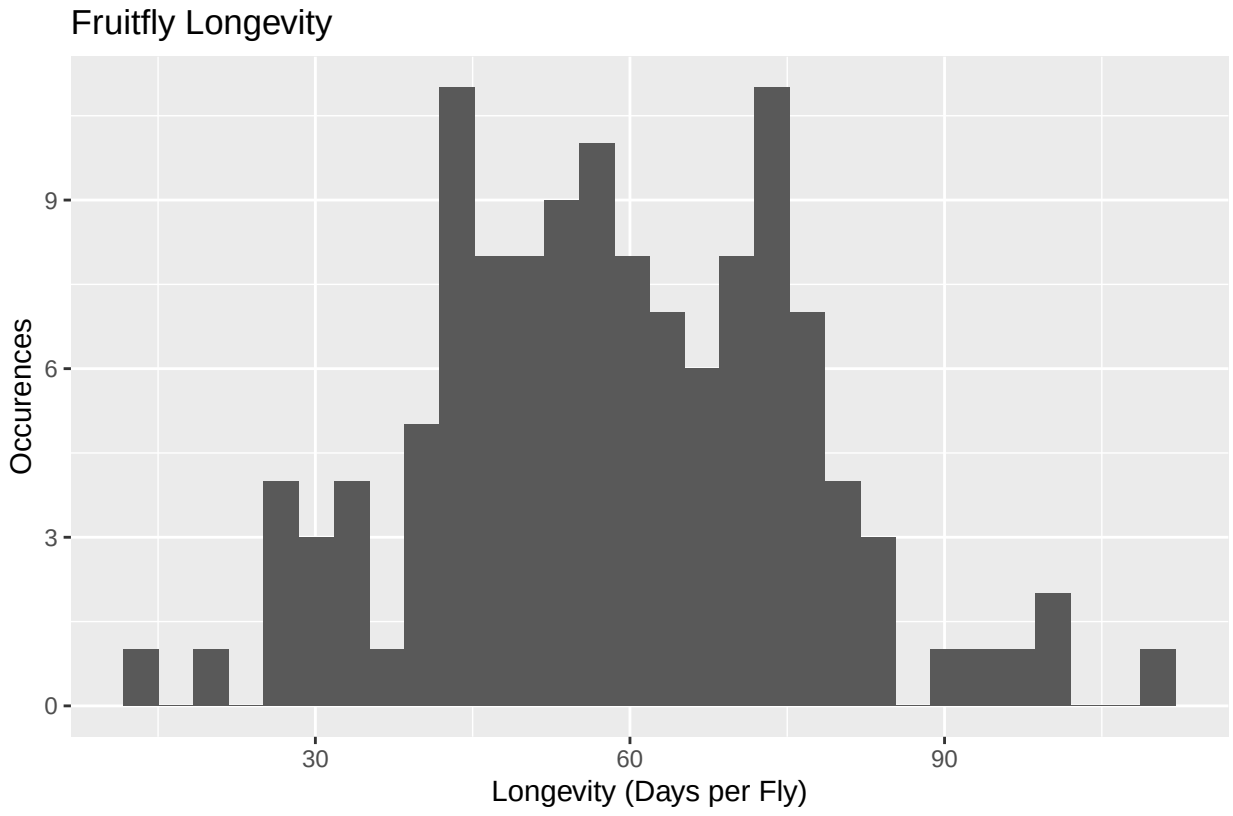
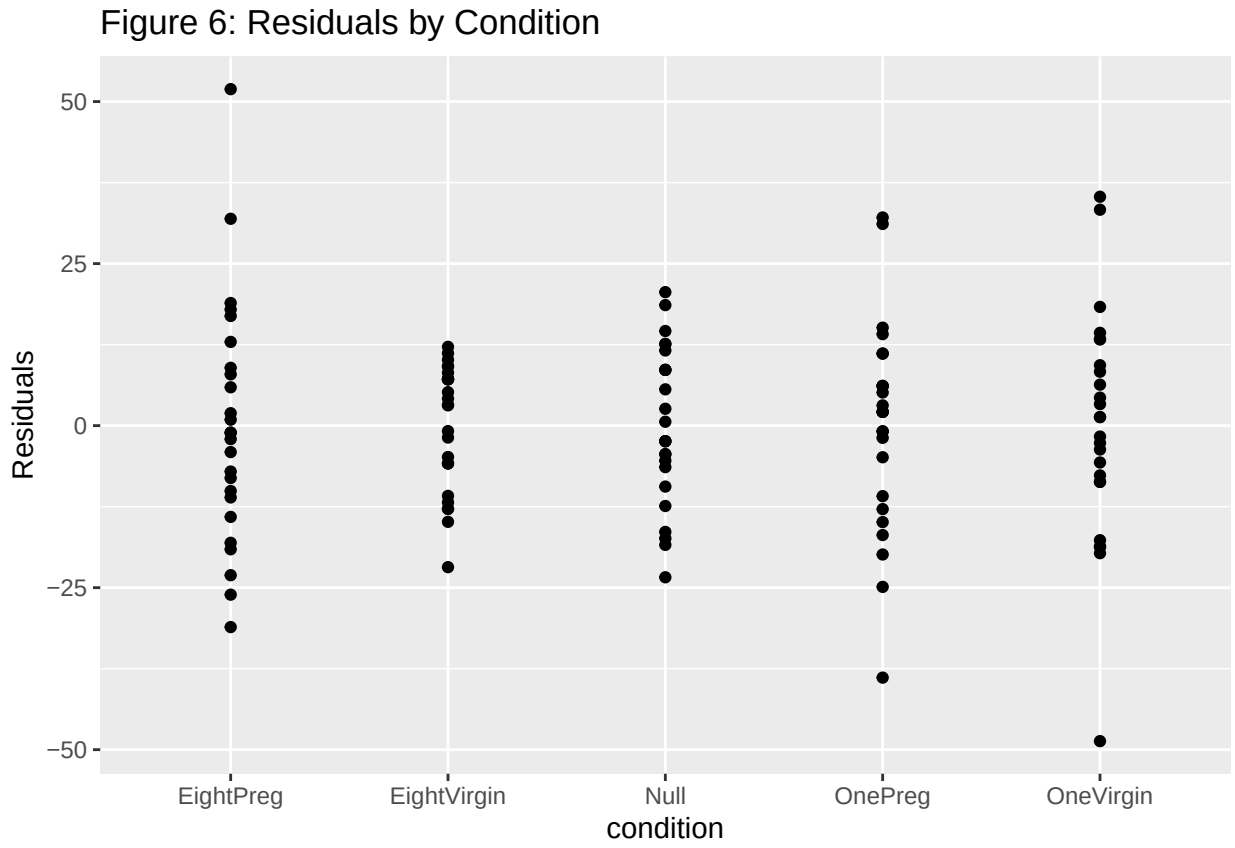


Figure 2: Fruitfly Longevity by Occurence

Now that we understand what we are looking at and what we hope to measure, we want to use a one-way ANOVA test.

One Way Assumptions

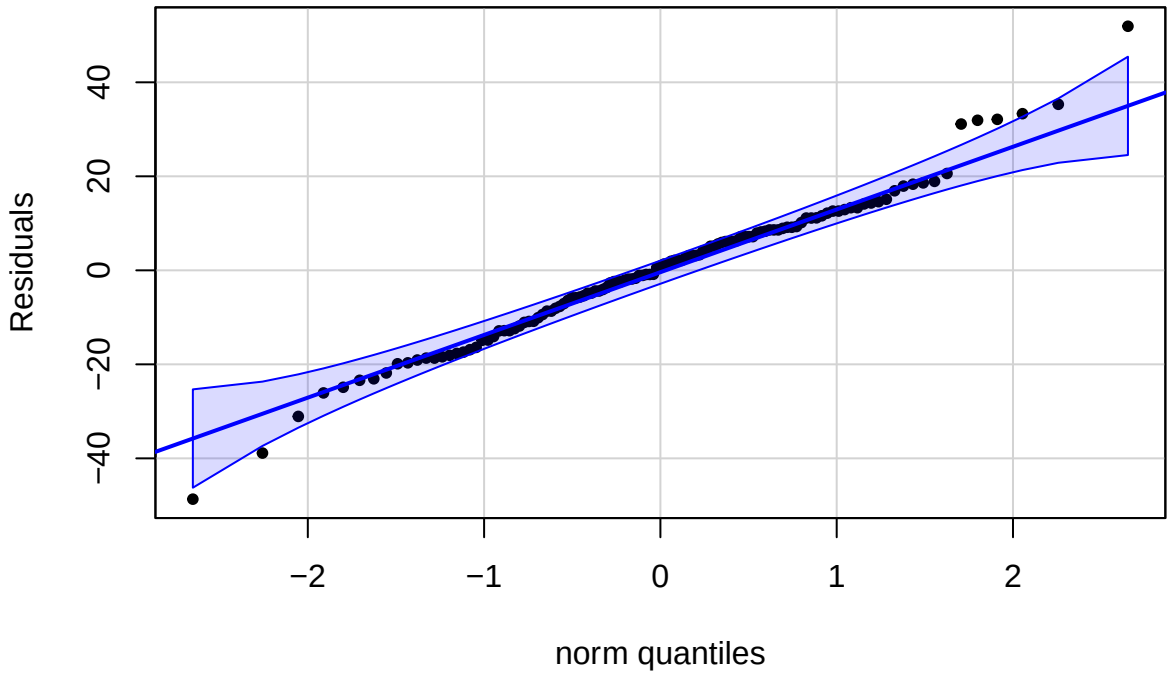
In this section, we will demonstrate the assumptions required for a one-way ANOVA test. ## Homoscedasticity



As we check for homoscedasticity, we see there is an even distribution and no funneling. Therefore, we can assume homoscedasticity is satisified.

QQ Plot

Figure 7: QQ Plot



It appears that seven of the 125 observations were outside the 90% envelope. Since this is around 5.6% of total observations, we will act as if the Gaussian assumption is satisfied.

Independence of Observations

As no pattern is seen, we can assume independence of observations. Since all the assumptions have been met, we will proceed with a one-way ANOVA test.

Table 1: Figure 3: Modern ANOVA Table for Fruit Fly Data

Source	SS	df	MS	F	p-value	Eta Sq.	Omega Sq.	Epsilon Sq.
condition	11671.25	4	2917.8120	12.5321	0	0.2947	0.2696	0.2711
Residuals	27939.28	120	232.8273	NA	NA	NA	NA	NA

Post-Hoc – Pairwise Comparison

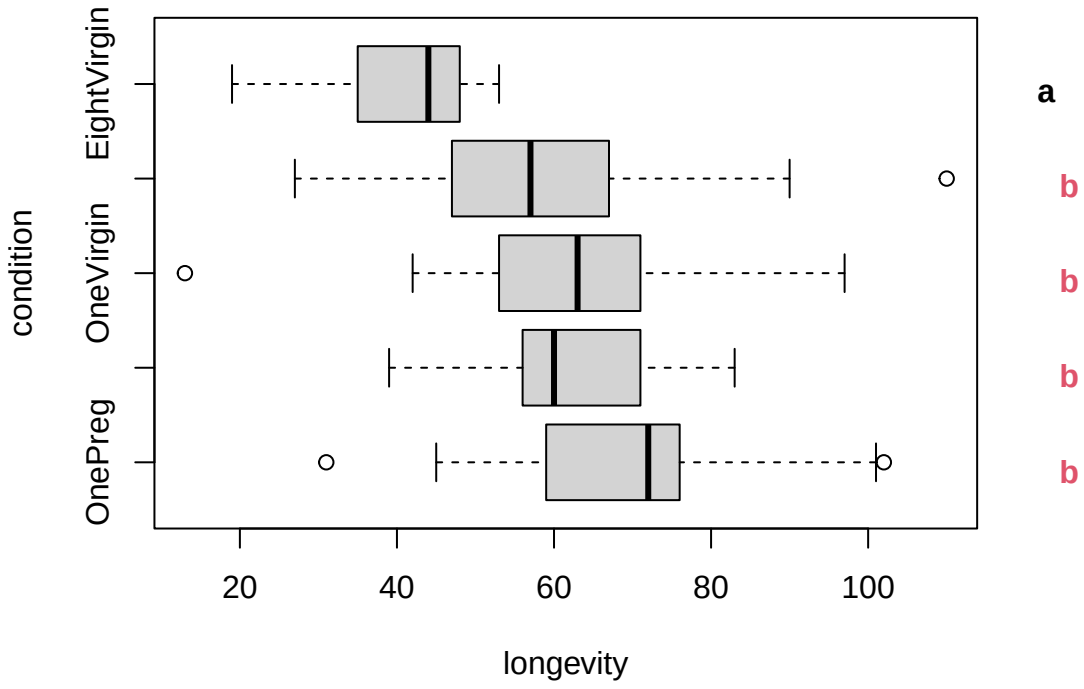
Table 2: Figure 4: Point Estimates from the Fruit Fly Study

	Estimate
Grand Mean	58.576
Null	-0.496
One Preg	-17.736
One Virgin	3.824
Eight Preg	11.304
Eight Virgin	3.104

Table 3: Figure 5: Pairwise Post Hoc Comparison via Tukey HSD

Pair	Difference	SE	DF	t	p-value
EightPreg - EightVirgin	17.24	4.316	120	3.995	0.001
EightPreg - Null	-4.32	4.316	120	-1.001	0.854
EightPreg - OnePreg	-11.80	4.316	120	-2.734	0.055
EightPreg - OneVirgin	-3.60	4.316	120	-0.834	0.919
EightVirgin - Null	-21.56	4.316	120	-4.996	0.000
EightVirgin - OnePreg	-29.04	4.316	120	-6.729	0.000
EightVirgin - OneVirgin	-20.84	4.316	120	-4.829	0.000
Null - OnePreg	-7.48	4.316	120	-1.733	0.418
Null - OneVirgin	0.72	4.316	120	0.167	1.000
OnePreg - OneVirgin	8.20	4.316	120	1.900	0.323

We can see all the comparisons of pairs to evaluate how they compare to the GSAM. Right away, we can tell that the greatest differences lie between the comparison of eight virgin fruit flies vs. one virgin fruit fly.



Results

Our research question was based on the abstract from the Partridge and Farquhar study, "How does sexual activity affect the longevity of the male fruit fly?"

The null hypothesis is: There is no cost of sexual activity of the male fruit fly that affects their longevity.

The alternative hypothesis is: There is a cost associated with sexual activity if the male fruit fly that affects their longevity.

Discussion

From the one-way ANOVA test, we will reject our null hypothesis in favor of the alternative, as the p-value acquired (0) is less than our unusualness threshold of 0.01.

While the original study did not go into as much depth as our analysis, we can continue under the assumption that the pairwise combinations that were statistically significant were: Eight pregnant vs. Eight virgin, Eight virgin vs. Null, Eight virgin vs. one pregnant, and Eight virgin vs. one virgin. For these pairs, we will reject our null hypothesis that there is no cost of sexual activity of the male fruit fly that affects their longevity. These pairs demonstrated a p-value lower than our unusualness threshold of 0.01.

However, for the eight pregnant vs. Null, eight pregnant vs. one pregnant, eight pregnant vs. one virgin, null vs. one pregnant, null vs. one virgin, and one pregnant vs. one virgin pairs, we would fail to reject our null hypothesis. These pairs demonstrated a p-value higher than our unusualness threshold of 0.01.

References

Bhole, D., & Tower, J. (2002). Fruit Flies, Drosophila. In D. J. Ekerdt (Ed.), Encyclopedia of Aging (Vol. 2, pp. 514-517). Macmillan Reference USA. https://link-gale-com.ezaccess.libraries.psu.edu/apps/doc/CX3402200152/GVRL?u=psucic&sid=bookmark-GVRL&xid=849d078d

Partridge, L., Farquhar, M. Sexual activity reduces lifespan of male fruit flies. Nature 294, 580–582 (1981). https://doi.org/10.1038/294580a0

Code Appendix

```
# Setting Document Options
knitr::opts_chunk$set(
  echo = FALSE,
  warning = FALSE,
  message = FALSE,
  fig.align = "center"
)

# Add additional packages by name to the following list
packages <- c("tidyverse", "knitr",
"kableExtra", "car", "psych", "parameters","Rgraphviz","hasseDiagram","multcompView")
lapply(
  X = packages,
  FUN = library,
  character.only = TRUE
)

# Loading Helper Files and Setting Global Options
source("https://raw.github.com/neilhatfield/STAT461/master/rScripts/ANOVATools.R")
options("contrasts" = c("contr.sum", "contr.poly"))

# Read In CSV file
fruitFly <- read.csv("fruitflies.csv")
# Convert Condition to Factor
fruitFly$condition <- as.factor(fruitFly$condition)
#AOV for FF data
flyModel <- aov(
  formula = longevity ~ condition,
  data = fruitFly,
  na.action = "na.omit"
  )
modelLabels <- c("1 Fly Mating 1", "5 Factor 4", "125 (days per fly) 120")
modelMatrix <- matrix(
  data = c(FALSE, FALSE, FALSE, TRUE, FALSE, FALSE, TRUE, TRUE, FALSE),
  nrow = 3,
  ncol = 3,
  byrow = FALSE
)
hasseDiagram::hasse(
 data = modelMatrix,
 labels = modelLabels
)
ggplot(data = fruitFly, aes(x = longevity)) + geom_histogram() + ggtitle("Fruitfly Longevity") + xlab("Longevi

ggplot(data = fruitFly, aes(x = condition, y = flyModel$residuals)) + ylab("Residuals") +geom_point() +ggtitle
car::qqPlot(
x = flyModel$residuals,
distribution = "norm",
envelope = 0.90,
id = FALSE,
pch = 20,
ylab = "Residuals",
main = 'Figure 7: QQ Plot'
)
ggplot(
data = data.frame(
residuals = flyModel$residuals,
```

```
index = 1:length(flyModel$residuals)
),
mapping = aes(x = index, y = residuals)
) +
geom_point(size = 1.5) +
geom_line() +
theme_bw() +
geom_hline(
yintercept = 0,
linetype = "dashed",
color = "red"
) +
xlab("Measurement Order") +
ylab("Residuals")+ggtitle("Figure 8: Residuals by Measurement Order")

# Modern Table
parameters::model_parameters(
model = flyModel,
effectsize_type = c("eta", "omega", "epsilon") # Effect sizes
) %>%
knitr::kable(
digits = 4,
col.names = c(
"Source", "SS", "df", "MS", "F", "p-value",
"Eta Sq.", "Omega Sq.", "Epsilon Sq."),
caption = "Figure 3: Modern ANOVA Table for Fruit Fly Data",
booktabs = TRUE,
align = c("l", rep("c", 8))
) %>%
kableExtra::kable_styling(
font_size = 10,
latex_options = c("HOLD_position"))

pointEst <- dummy.coef(flyModel)
# pointEst # Look at the output of pointEst so you know in what order the estimates appear
pointEst <- unlist(pointEst)
names(pointEst) <- c("Grand Mean", "Null","One Preg","One Virgin","Eight Preg", "Eight Virgin")
data.frame("Estimate" = pointEst) %>%
knitr::kable(
digits = 4,
caption = "Figure 4: Point Estimates from the Fruit Fly Study",
booktabs = TRUE,
align = "c"
) %>%
kableExtra::kable_styling(
font_size = 15,
latex_options = c("HOLD_position")
)
# Convert the TukeyHSD results to a data frame
flyPairs <- emmeans::emmeans(
object = flyModel, # Your aov/lm object
specs = pairwise ~ condition, # Creates all pairs of the levels of the factor listed
adjust = "tukey", # How you want to control the error rate
level = 0.99 # 1 - Your overall Type I Error Rate
)
## Make a professional looking table
knitr::kable(
x = flyPairs$contrasts, # Grab the appropriate sub-object
digits = 3,
```

```
caption = "Figure 5: Pairwise Post Hoc Comparison via Tukey HSD",
col.names = c("Pair", "Difference", "SE", "DF", "t", "p-value"),
align = "lccccc",
booktabs = TRUE
) %>%
kableExtra::kable_styling(
bootstrap_options = c("condensed", "boardered"),
font_size = 12,
latex_options = c("HOLD_position")
)

multcompBoxplot(
  formula = longevity ~ condition,
  data = fruitFly,
  compFn = "TukeyHSD",
  plotList = list(
    boxplot = list(fig = c(0, 0.85, 0, 1)),
    multcompLetters = list(
      fig = c(0.8, 0.9, 0.1, 0.9),
      fontsize = 12,
      fontface = "bold"
    )
  )
)
```