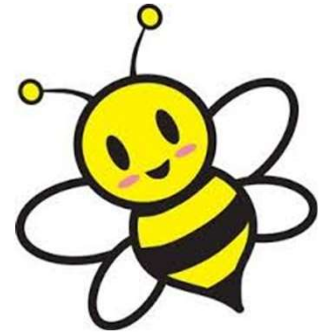
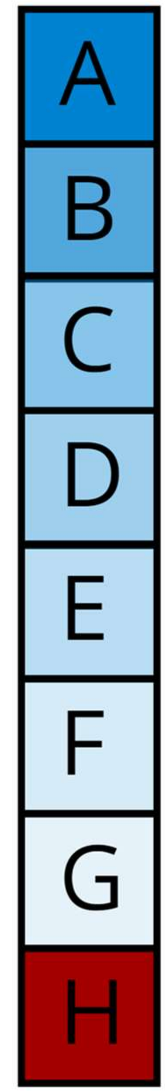


Do increasing concentrations of lime sulphur solution act as stronger honey bee repellents?

Ethan Scott, Daniel Girvitz

Background

1. 8 groups of 100 bees were released into 8 different chambers for two hours.
2. Each chamber, except one control, contained cells of dry comb with lime sulphur emulsion in sucrose solution.
3. Seven different concentrations of lime sulphur, labelled A - G, ranged from a concentration of 1/100 to 1/1,562,500, in successive factors of ⅕.
4. The one control, labelled H, contained no lime sulphur emulsion in the sucrose solution.
5. The decrease in volume of the solutions in the various cells was measured, in order to determine the potency of the lime sulphur emulsion as a honey bee repellent.



Dataset

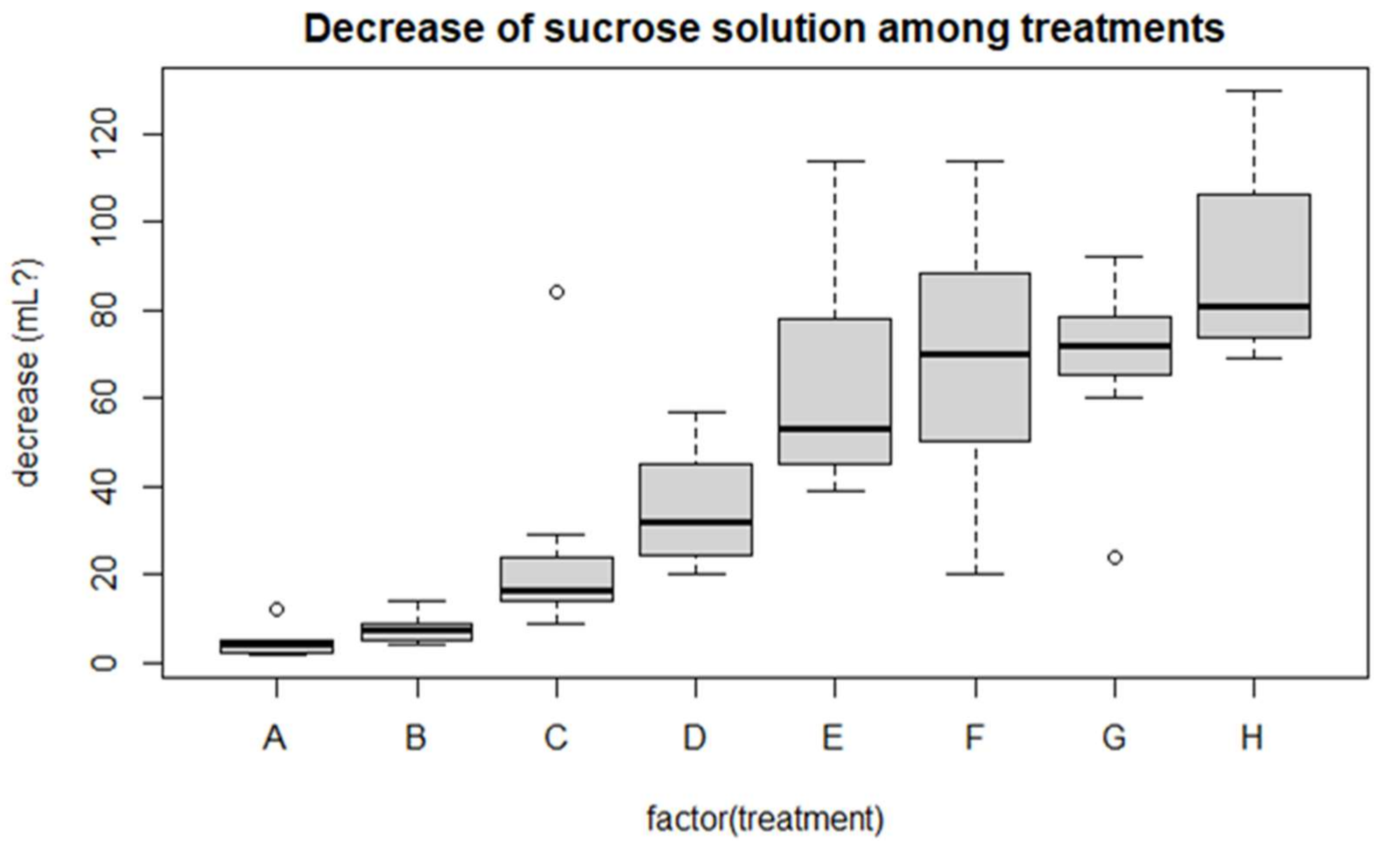
- An 8x8 Latin square design with coded treatments.
- Since we haven’t yet learned about Latin square design, we disregarded that information.
- We kept only two variables: “treatment”, which is a factor, and “decrease”, which is discrete numerical.
- It appears that decreases in solutions were calculated for 8 random cells of dry comb from each of the 8 groups
    - n=8 (instances in each treatment)
    - a=8 (treatments)
    - N=n*a=8*8=64 (total instances)

Hypothesis

H0: Equality of treatment means for the decrease in volume of the solutions in the various cells.

Ha: Inequality between at least two treatment means for the decrease in volume of the solutions in the various cells.

Boxplots

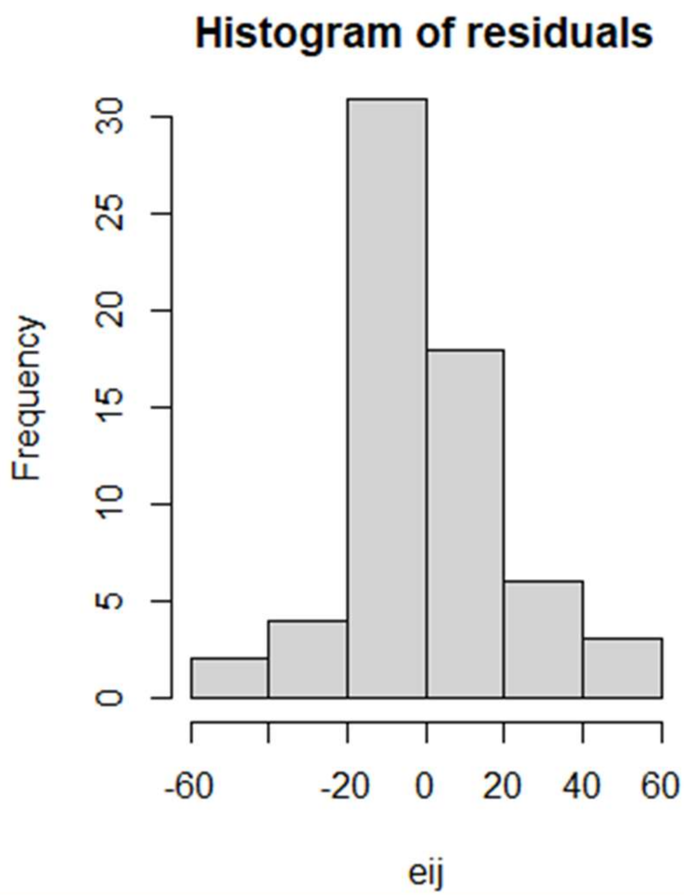


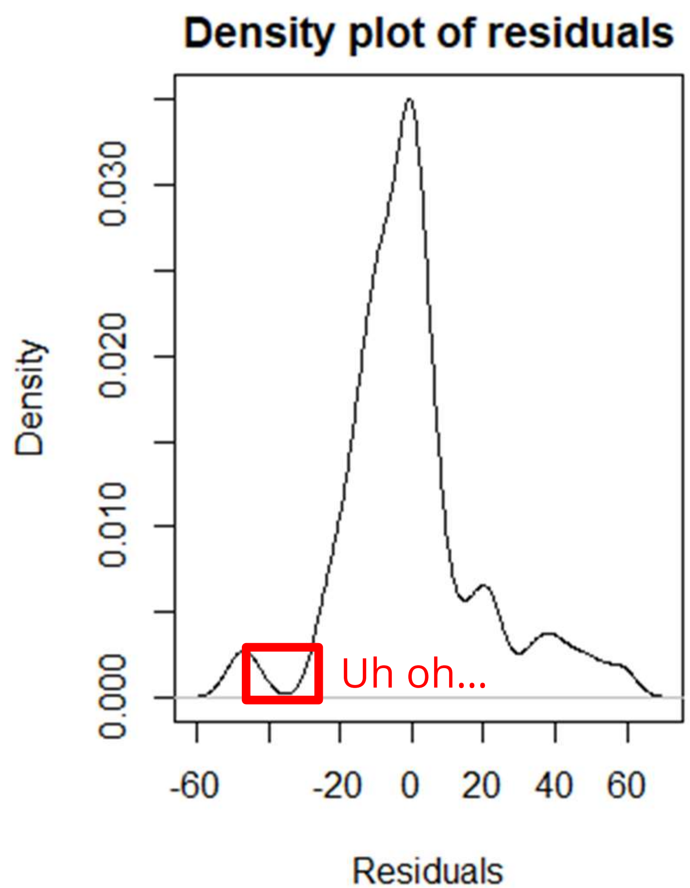
Decrease of sucrose solution among treatments
decrease (mL?)
0
20
40
60
80
100
120
A
B
C
D
E
F
G
H
factor(treatment)

ANOVA Assumptions

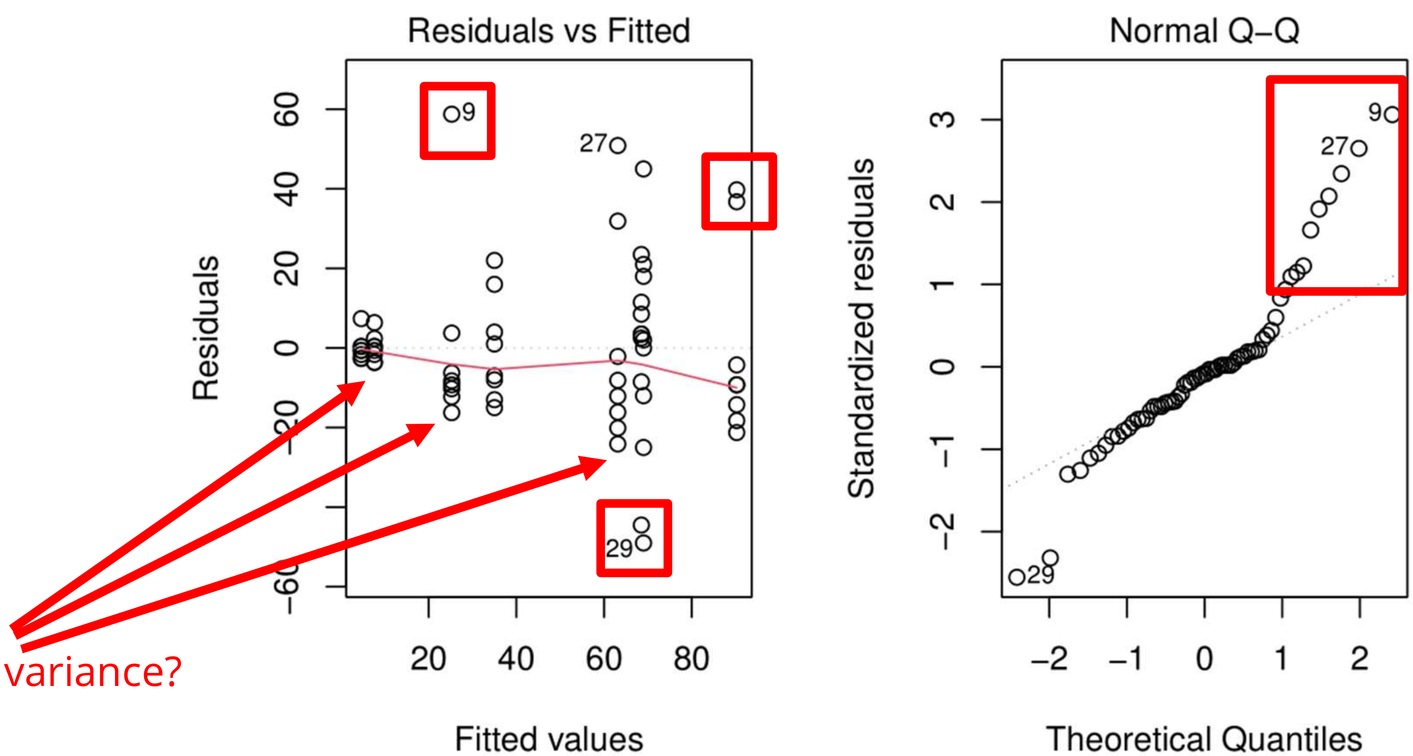
1. Data are **i.i.d normally distributed**
2. **Homogenity of variance** among treatments (in our case, decreases in solutions), which is called homoscedasticity, which is not to be pronounced…
3. Balanced design

Plots of residuals





Diagnostic plots



What’s up with this upward curvature?

Unequal variance?

Shapiro-Wilk test

```{r}
shapiro.test((eData$decrease))
```

```
	Shapiro-Wilk normality test

data:  (eData$decrease)
W = 0.91892, p-value = 0.0004483
```

Yikes...

Modified Levene’s Test

```{r}
library(car)
# Levene's test with one independent variable
leveneTest(decrease ~ factor(treatment), data = eData)

```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  7  1.5132 0.1817
      56
```

FTR equality of variance

Ethan begins here.

Single-factor ANOVA

```
```{r}
one.way=aov(decrease~factor(treatment),data=eData)
summary(one.way)
```
```

```
                  Df Sum Sq Mean Sq F value  Pr(>F)    
factor(treatment)  7  56160    8023   19.06 9.5e-13 ***
Residuals         56  23570     421                    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

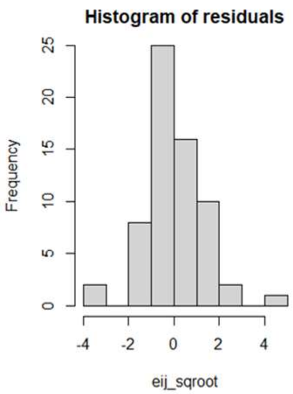
FTR equality of treatment means

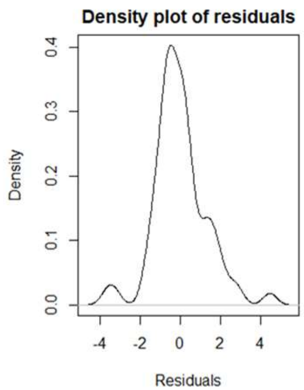
Square root transformation applied to ANOVA

```{r}
one.way.sqroot=aov(sqrt(decrease)~factor(treatment),data=eData)
summary(one.way.sqroot)
```

```
                  Df Sum Sq Mean Sq F value Pr(>F)
factor(treatment)  7  417.7   59.68   30.95 <2e-16 ***
Residuals         56  108.0    1.93
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

FTR





Shapiro-Wilk's test for normality

```{r}
shapiro.test(sqrt(eData$decrease))
```

```
        Shapiro-Wilk normality test

data:  sqrt(eData$decrease)
W = 0.94271, p-value = 0.005088
```

Reject

Tests for homogenity of variance (homoscedastiticity)

```
bartlett.test(sqrt(decrease) ~ factor(treatment), data = eData)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  sqrt(decrease) by factor(treatment)
## Bartlett's K-squared = 15.529, df = 7, p-value = 0.02979
```

```
leveneTest(sqrt(decrease) ~ factor(treatment), data = eData)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  7   0.763 0.6205
##       56
```

FTR

Fisher LSD Test for difference of means

```
N=64
n=8
a=8
MSE=8023
Fisher.LSD=qt(0.05, N-a, lower.tail=F)*sqrt(MSE*2/n)
Fisher.LSD
comparisons=c(
  means[1]-means[2],
  means[1]-means[3],
  means[1]-means[4],
  means[1]-means[5],
  means[1]-means[6],
  means[1]-means[7],
  means[1]-means[8],#Positive
  means[2]-means[3],
  means[2]-means[4],
  means[2]-means[5],
  means[2]-means[6],
  means[2]-means[7],
  means[2]-means[8],#positive
  means[3]-means[4],
  means[3]-means[5],
  means[3]-means[6],
  means[3]-means[7],
  means[3]-means[8],
  means[4]-means[5],
  means[4]-means[6],
  means[4]-means[7],
  means[4]-means[8],
  means[5]-means[6],
  means[5]-means[7],
  means[5]-means[8],
  means[6]-means[7],
  means[6]-means[8],
  means[7]-means[8])
abs(comparisons)-Fisher.LSD
```

Values from the difference of means for treatment groups A to H and B to H have positive values with the Fisher LSD test

```
         A          A          A          A          A          A          A          B
-71.904915 -54.279915 -44.529915 -16.404915 -10.529915 -11.029915  10.720085 -57.279915
         B          B          B          B          B          C          C          C
-47.529915 -19.404915 -13.529915 -14.029915   7.720085 -65.154915 -37.029915 -31.154915
         C          C          D          D          D          D          E          E
-31.654915  -9.904915 -46.779915 -40.904915 -41.404915 -19.654915 -69.029915 -69.529915
         E          F          F          G
-47.779915 -74.404915 -53.654915 -53.154915
```

Tukey's Test for difference of means

Tukey's Test gives the same results as the Fisher LSD test

```
Tukey = q_Alpha*sqrt(MSE/n)
Tukey
abs(comparisons)-Tukey
```

```
          A          A          A          A          A          A          A          B
-63.503224 -45.878224 -36.128224  -8.003224  -2.128224  -2.628224  19.121776 -48.878224
          B          B          B          B          B          C          C          C
-39.128224 -11.003224  -5.128224  -5.628224  16.121776 -56.753224 -28.628224 -22.753224
          C          C          D          D          D          D          E          E
-23.253224  -1.503224 -38.378224 -32.503224 -33.003224 -11.253224 -60.628224 -61.128224
          E          F          F          G
-39.378224 -66.003224 -45.253224 -44.753224
```

Confidence Intervals: Comparing Pairs of Treatment Means

```
     (means[1]-means[2])+c(-1,1)*Tukey [1] -69.50322  63.50322
     (means[1]-means[3])+c(-1,1)*Tukey [1] -87.12822  45.87822
     (means[1]-means[4])+c(-1,1)*Tukey [1] -96.87822  36.12822
     (means[1]-means[5])+c(-1,1)*Tukey [1] -125.003224    8.003224
     (means[1]-means[6])+c(-1,1)*Tukey [1] -130.878224    2.128224
     (means[1]-means[7])+c(-1,1)*Tukey [1] -130.378224    2.628224
A-H  (means[1]-means[8])+c(-1,1)*Tukey [1] -152.12822  -19.12178
     (means[2]-means[3])+c(-1,1)*Tukey [1] -84.12822  48.87822
     (means[2]-means[4])+c(-1,1)*Tukey [1] -93.87822  39.12822
     (means[2]-means[5])+c(-1,1)*Tukey [1] -122.00322   11.00322
     (means[2]-means[6])+c(-1,1)*Tukey [1] -127.878224    5.128224
     (means[2]-means[7])+c(-1,1)*Tukey [1] -127.378224    5.628224
B-H  (means[2]-means[8])+c(-1,1)*Tukey [1] -149.12822  -16.12178
     (means[3]-means[4])+c(-1,1)*Tukey [1] -76.25322  56.75322
     (means[3]-means[5])+c(-1,1)*Tukey [1] -104.37822   28.62822
     (means[3]-means[6])+c(-1,1)*Tukey [1] -110.25322   22.75322
     (means[3]-means[7])+c(-1,1)*Tukey [1] -109.75322   23.25322
     (means[3]-means[8])+c(-1,1)*Tukey [1] -131.503224    1.503224
     (means[4]-means[5])+c(-1,1)*Tukey [1] -94.62822  38.37822
     (means[4]-means[6])+c(-1,1)*Tukey [1] -100.50322   32.50322
     (means[4]-means[7])+c(-1,1)*Tukey [1] -100.00322   33.00322
     (means[4]-means[8])+c(-1,1)*Tukey [1] -121.75322   11.25322
     (means[5]-means[6])+c(-1,1)*Tukey [1] -72.37822  60.62822
     (means[5]-means[7])+c(-1,1)*Tukey [1] -71.87822  61.12822
     (means[5]-means[8])+c(-1,1)*Tukey [1] -93.62822  39.37822
     (means[6]-means[7])+c(-1,1)*Tukey [1] -66.00322  67.00322
     (means[6]-means[8])+c(-1,1)*Tukey [1] -87.75322  45.25322
     (means[7]-means[8])+c(-1,1)*Tukey [1] -88.25322  44.75322
```

Conclusion

1. Accept Ha in favour of H0 (yay)
2. Concentration of lime sulphur used to repel honey bees will be determined by the cost of the solution, required extent of power to repel, possible environmental impacts (contamination, effect on other species, etc.)

Thank you!

Additional argumentation

Although homogenity of variance (homoscedasticity) appears to be present, we have little evidence to argue for normality.

However odd it may be, non-normality works in our favour: It makes sense that A and B must be outliers, which implies strong concentrations of lime sulphur solution work as honey bee repellents. Treatments C through H should see a variation, because lesser concentrations wouldn't work as effectively. Furthermore, with 8 random cells chosen it is a natural consequence some have a greater decrease in sucrose solution than others - the honey bee's choice of cell is random.