

REX: Insect Preference Trial Plots & Analyses

Moriah Young

September 17, 2021

```
## Loading required package: stats4

## Loading required package: lme4

## Loading required package: Matrix

##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##      lmer

## The following object is masked from 'package:stats':
##
##      step

## Loading required package: MASS

## Loading required package: survival

## Learn more about sjPlot with 'browseVignettes("sjPlot")'.

## Loading required package: carData

## Registered S3 methods overwritten by 'car':
##   method                               from
##   influence.merMod                     lme4
##   cooks.distance.influence.merMod     lme4
##   dfbeta.influence.merMod              lme4
##   dfbetas.influence.merMod             lme4

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.3    v purrr  0.3.4
## v tibble  3.0.6    v dplyr  1.0.4
## v tidyr   1.1.2    v stringr 1.4.0
## v readr   1.4.0    v forcats 0.5.1
```

```

## -- Conflicts ----- tidyverse_conflicts() --
## x tidyr::expand() masks Matrix::expand()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
## x tidyr::pack() masks Matrix::pack()
## x dplyr::recode() masks car::recode()
## x dplyr::select() masks MASS::select()
## x dplyr::slice() masks bbmle::slice()
## x purrr::some() masks car::some()
## x tidyr::unpack() masks Matrix::unpack()

##
## Attaching package: 'rstatix'

## The following object is masked from 'package:MASS':
##
## select

## The following object is masked from 'package:stats':
##
## filter

##
## Attaching package: 'janitor'

## The following object is masked from 'package:rstatix':
##
## make_clean_names

## The following objects are masked from 'package:stats':
##
## chisq.test, fisher.test

## [1] "animal"
## [2] "GHG"
## [3] "Irrigation"
## [4] "lookup_tables"
## [5] "microbes"
## [6] "REX Microbial sampling IDs complete.gsheel"
## [7] "REX_Data_Management.gdoc"
## [8] "REX_Data_Template.gsheel"
## [9] "REX_warmx_metadata.csv"
## [10] "REX_warmx_metadata.xlsx"
## [11] "soil"
## [12] "T7_ANPP"
## [13] "T7_plant_comp"
## [14] "T7_plant_phenology"
## [15] "T7_warmx_insect"
## [16] "T7_warmx_plant_traits"
## [17] "T7_warmx_VOC"
## [18] "Untitled form.gform"
## [19] "weather"

```

```

# Clean data
insects <- insects[-88, ] #this data point had a greater SLA_after than it's SLA_before even though it
# getting rid of NAs in data. NAs in data are from leaf
# samples collected for CN analyses and not used in the
# trials
insects1 <- na.omit(insects)
# delete date column
insects1$date <- NULL
# Change column names to lowercase
names(meta) <- tolower(names(meta))

```

Data Wrangling

```

# Take subplot average of sla before trial
sla_before <- insects1 %>% group_by(replicate, footprint_location,
  subplot_location) %>% dplyr::summarize(sla_before_mean = mean(sla_before,
    na.rm = TRUE))

```

'summarise()' has grouped output by 'replicate', 'footprint_location'. You can override using the '.

```

# Take subplot average of sla after trial
sla_after <- insects1 %>% group_by(replicate, footprint_location,
  subplot_location) %>% dplyr::summarize(sla_after_mean = mean(sla_after_1,
    na.rm = TRUE))

```

'summarise()' has grouped output by 'replicate', 'footprint_location'. You can override using the '.

```

insects2 <- left_join(sla_before, sla_after) # combine sla_before and sla_after into one dataframe

```

```

## Joining, by = c("replicate", "footprint_location", "subplot_location")

```

```

# calculate the proportion of leaves eaten
insects3 <- insects2 %>% group_by(replicate, footprint_location,
  subplot_location) %>% dplyr::summarize(prop_eaten = sla_after_mean/sla_before_mean)

```

'summarise()' has grouped output by 'replicate', 'footprint_location'. You can override using the '.

```

insects4 <- left_join(insects2, insects3)

```

```

## Joining, by = c("replicate", "footprint_location", "subplot_location")

```

```

insects5 <- left_join(meta, insects4)

```

```

## Joining, by = c("replicate", "footprint_location", "subplot_location")

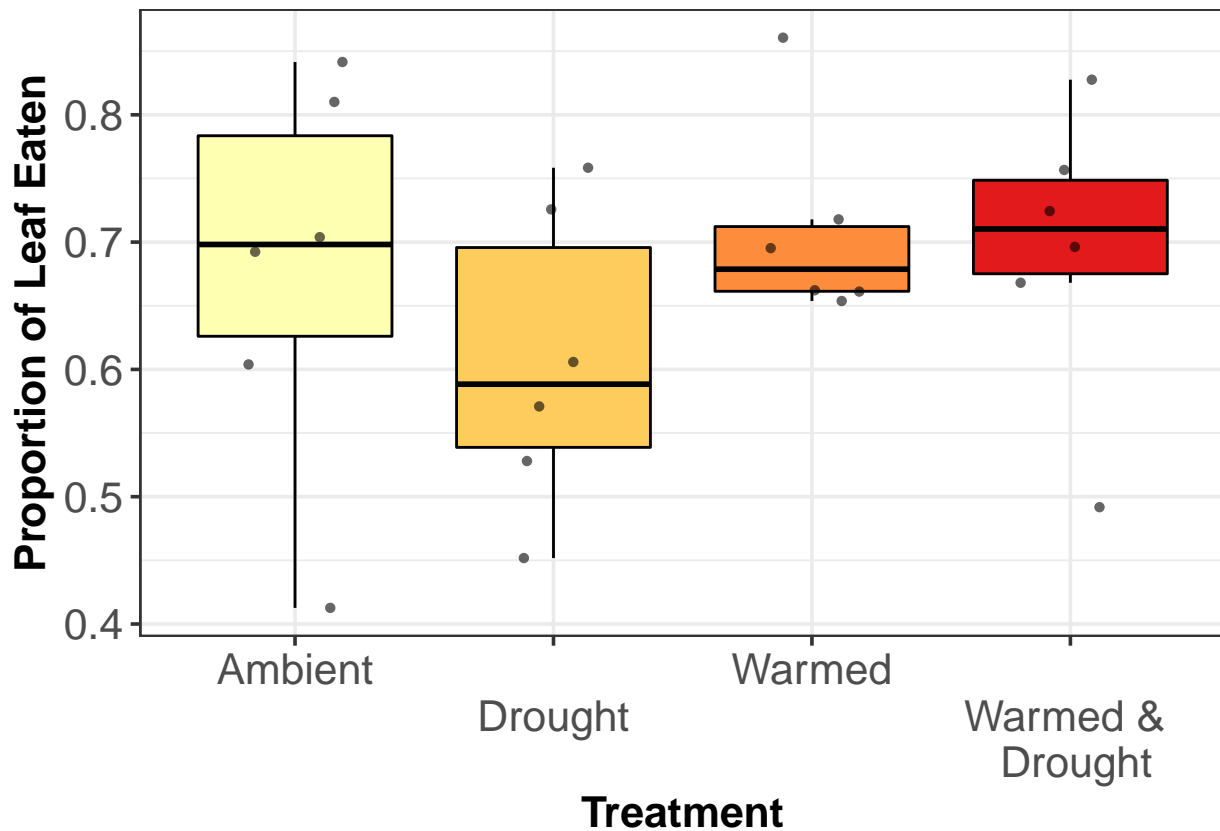
```

```

# get rid of NAs in data - irrigated control plots were not
# used in this experiment
insects6 <- na.omit(insects5)

```

```
# Subplot level graph png('insect_preference_treatments.png',
# units='in', width=7, height=6, res=300)
ggplot(insects6, aes(x = subplot_descriptions, y = prop_eaten,
  fill = subplot_descriptions)) + geom_boxplot(color = "black",
  outlier.shape = NA) + geom_jitter(shape = 16, position = position_jitterdodge(),
  alpha = 0.6) + # geom_errorbar(aes(ymin = lower, ymax =upper)) +
labs(x = "Treatment", y = "Proportion of Leaf Eaten") + scale_fill_brewer(palette = "YlOrRd") +
  scale_x_discrete(limits = c("ambient", "drought", "warmed",
    "warmed_drought"), labels = c(ambient = "Ambient", drought = "Drought",
    warmed = "Warmed", warmed_drought = "Warmed & \n Drought")),
  guide = guide_axis(n.dodge = 2)) + theme(legend.position = "none")
```



```
# dev.off()

# looking at the graph, there is more variability in the
# grasshoppers eating of ambient leaves and less so going
# from ambient to warmed + drought leaves.
```

```
# subset the leaves that had no herbivory - did certain
# treatments have more no herbivory on it's leaves?
no_herbivory <- subset(insects1, notes == "no_herbivory")
no_herbivory <- left_join(no_herbivory, meta)
```

```
## Joining, by = c("replicate", "footprint_location", "subplot_location")
```

```
summary(no_herbivory)
```

```
## replicate      footprint_location subplot_location plant_number
## Length:36      Min.      :1.000      Length:36      Length:36
## Class :character 1st Qu.:2.000      Class :character Class :character
## Mode :character Median :3.000      Mode :character Mode :character
##                  Mean  :3.194
##                  3rd Qu.:4.250
##                  Max.   :7.000
## sla_before      sla_after      sla_after_1      notes
## Min.      :1.340 Min.      :1.080 Min.      :1.340 Length:36
## 1st Qu.:2.390 1st Qu.:1.755 1st Qu.:2.390 Class :character
## Median :3.075 Median :2.610 Median :3.075 Mode :character
## Mean :3.255 Mean :2.801 Mean :3.255
## 3rd Qu.:3.857 3rd Qu.:3.458 3rd Qu.:3.857
## Max. :6.000 Max. :5.670 Max. :6.000
## id_number      treatment      footprint      subplot
## Min.      :185.0 Length:36      Length:36      Length:36
## 1st Qu.:238.5 Class :character Class :character Class :character
## Median :293.0 Mode :character Mode :character Mode :character
## Mean :274.1
## 3rd Qu.:321.0
## Max. :327.0
## subplot_descriptions unique_id
## Length:36      Length:36
## Class :character Class :character
## Mode :character Mode :character
##
##
##
```

```
tabyl(no_herbivory, subplot_descriptions) # pretty even among treatments
```

```
## subplot_descriptions n percent
## ambient 10 0.2777778
## drought 10 0.2777778
## warmed 8 0.2222222
## warmed_drought 8 0.2222222
```

```
# proportion of leaves eaten at individual level
```

```
indv_prop <- insects1 %>% group_by(replicate, footprint_location,
  subplot_location, plant_number) %>% dplyr::summarize(prop_eaten = sla_after_1/sla_before)
```

```
## 'summarise()' has grouped output by 'replicate', 'footprint_location', 'subplot_location'. You can o
```

```
#
```

```
indv_prop <- left_join(indv_prop, insects1)
```

```
## Joining, by = c("replicate", "footprint_location", "subplot_location", "plant_number")
```

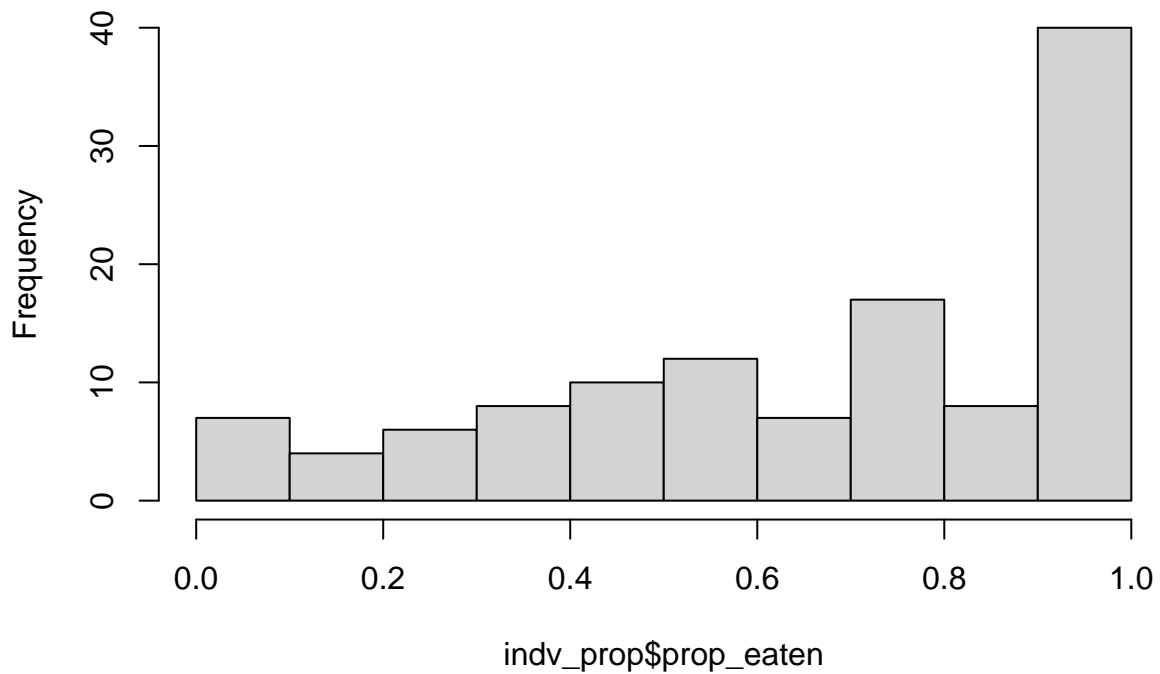
```
indv_prop <- left_join(meta, indv_prop)
```

```
## Joining, by = c("replicate", "footprint_location", "subplot_location")
```

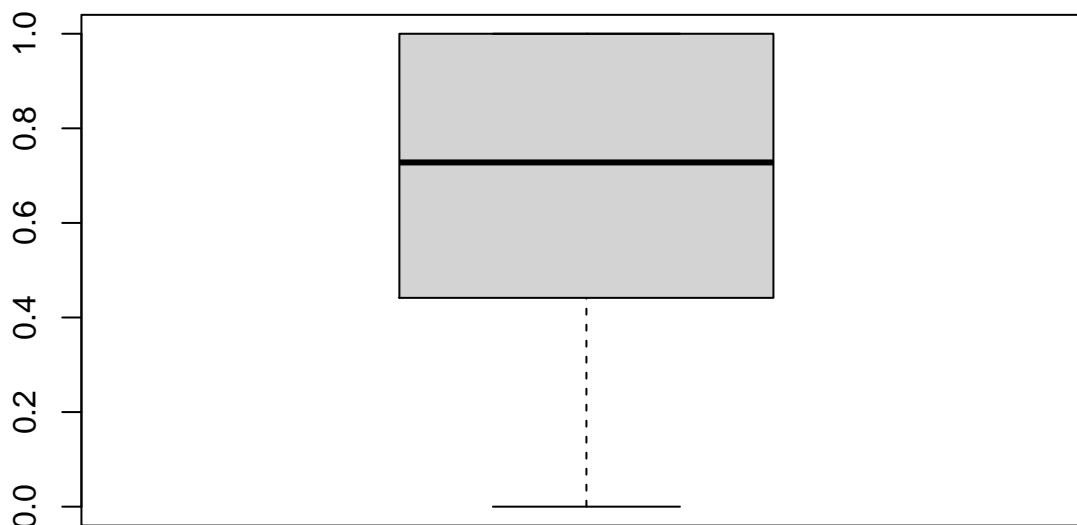
```
indv_prop <- na.omit(indv_prop)
```

```
hist(indv_prop$prop_eaten)
```

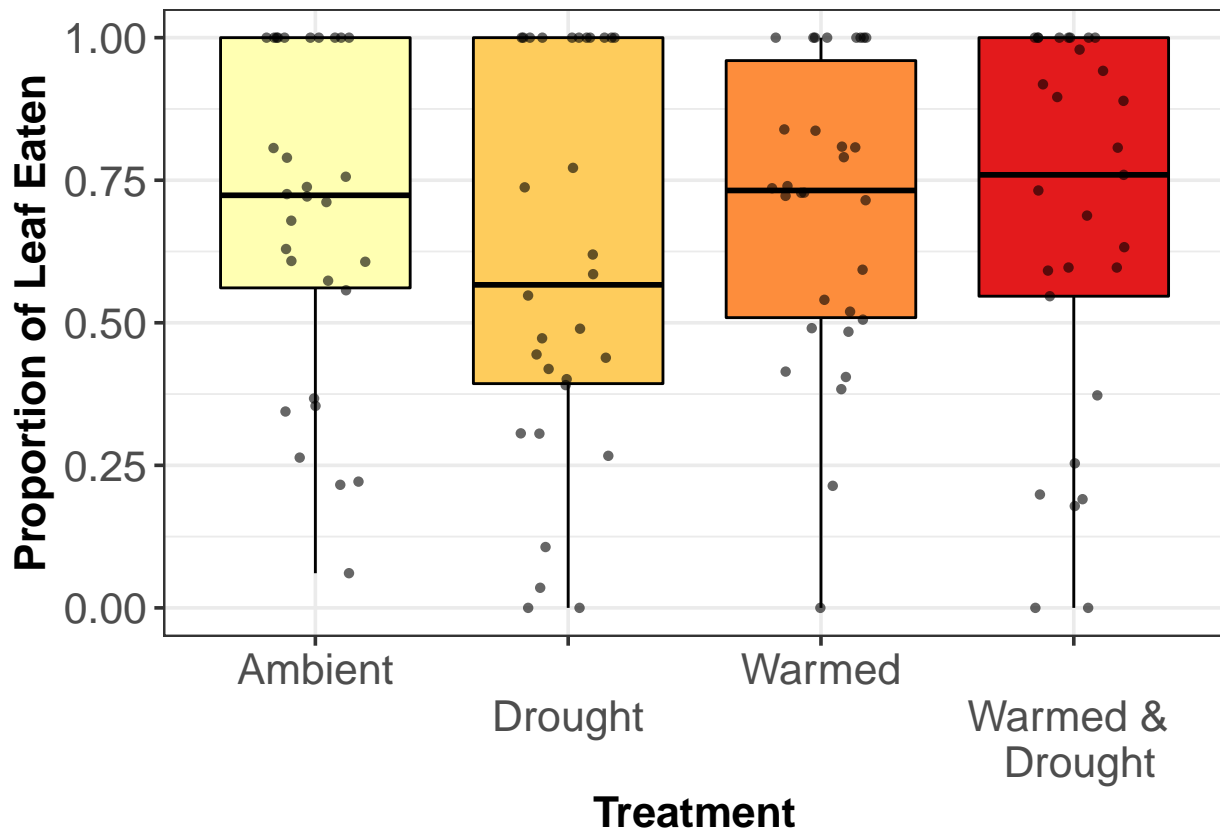
Histogram of indv_prop\$prop_eaten



```
boxplot(indv_prop$prop_eaten)
```



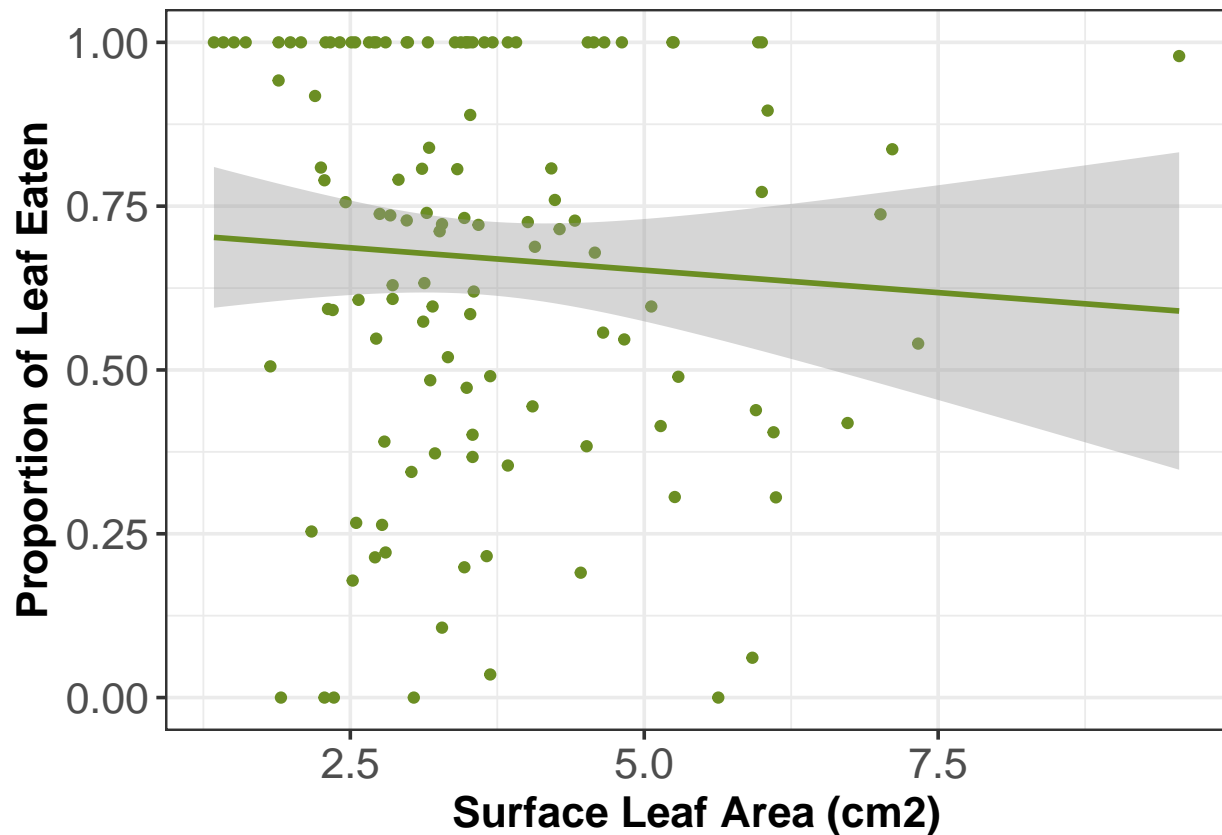
```
# png('indv_prop_treatments.png', units='in', width=7,
# height=6, res=300)
ggplot(indv_prop, aes(x = subplot_descriptions, y = prop_eaten,
  fill = subplot_descriptions)) + geom_boxplot(color = "black",
  outlier.shape = NA) + geom_jitter(shape = 16, position = position_jitterdodge(),
  alpha = 0.6) + # geom_errorbar(aes(ymin = lower, ymax =upper)) +
labs(x = "Treatment", y = "Proportion of Leaf Eaten") + scale_fill_brewer(palette = "YlOrRd") +
  scale_x_discrete(limits = c("ambient", "drought", "warmed",
    "warmed_drought"), labels = c(ambient = "Ambient", drought = "Drought",
    warmed = "Warmed", warmed_drought = "Warmed & \n Drought"),
  guide = guide_axis(n.dodge = 2)) + theme(legend.position = "none")
```



```
# dev.off()

# Regression
lm_indv_prop <- lm(prop_eaten ~ sla_before, data = indv_prop)
# png('regression_preference.png', units='in', width=5,
# height=5, res=300)
ggplot(indv_prop, aes(sla_before, prop_eaten)) + geom_point(color = "olivedrab") +
  geom_smooth(method = "lm", color = "olivedrab") + labs(x = "Surface Leaf Area (cm2)",
  y = "Proportion of Leaf Eaten")
```

```
## 'geom_smooth()' using formula 'y ~ x'
```



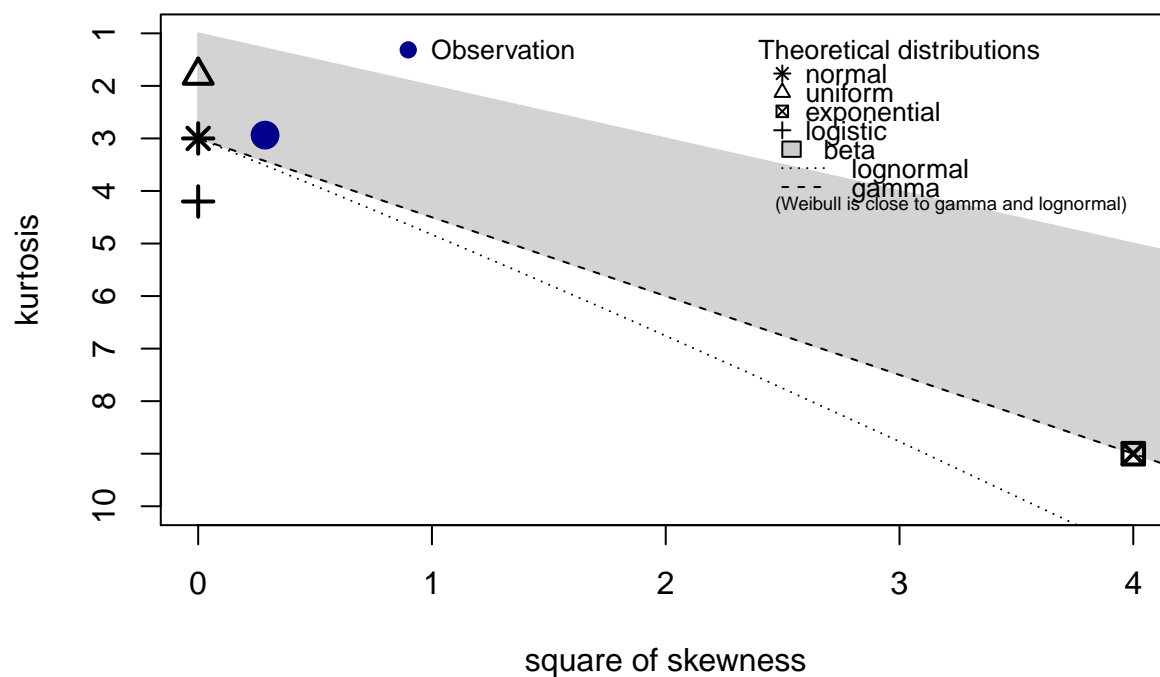
```
# dev.off()
summary(lm_indv_prop)
```

```
##
## Call:
## lm(formula = prop_eaten ~ sla_before, data = indv_prop)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69457 -0.20510  0.05407  0.31091  0.38906
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.72071    0.07849   9.182 1.83e-15 ***
## sla_before  -0.01369    0.02012  -0.680   0.498
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3091 on 117 degrees of freedom
## Multiple R-squared:  0.003938,    Adjusted R-squared:  -0.004575
## F-statistic: 0.4626 on 1 and 117 DF,  p-value: 0.4978
```

```
# Data Exploration and Analyses for proportion of leaf eaten
```

```
descdist(insects6$prop_eaten, discrete = FALSE)
```

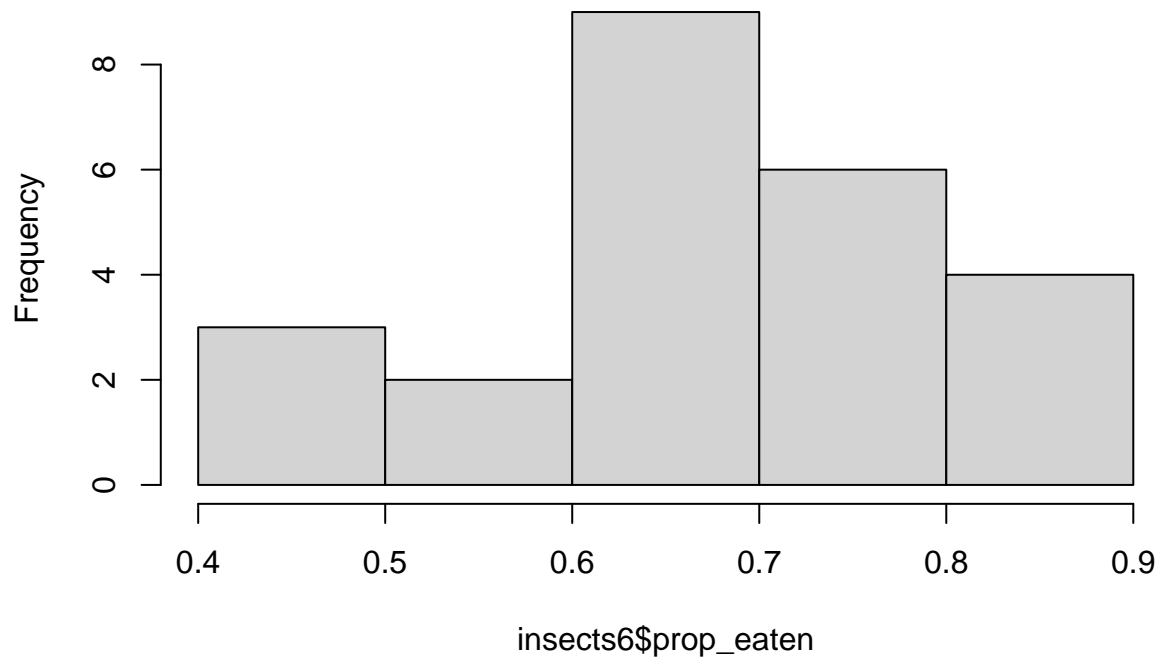

Cullen and Frey graph



```
## summary statistics
## -----
## min:  0.4126754  max:  0.8605882
## median:  0.6938382
## mean:  0.6716924
## estimated sd:  0.1181775
## estimated skewness:  -0.5348535
## estimated kurtosis:  2.937135
```

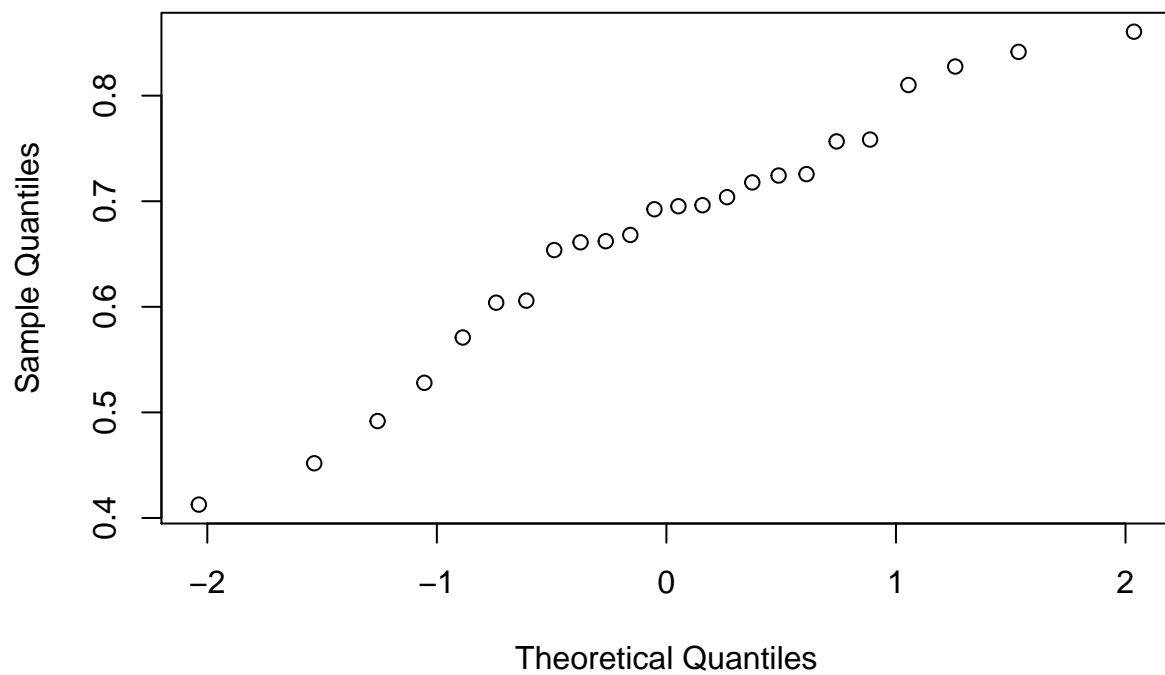
```
hist(insects6$prop_eaten)
```

Histogram of insects6\$prop_eaten



```
qqnorm(insects6$prop_eaten)
```

Normal Q-Q Plot



```
shapiro.test(insects6$prop_eaten) # p-value = 0.3238 so we can't reject the null hypothesis that the d
```

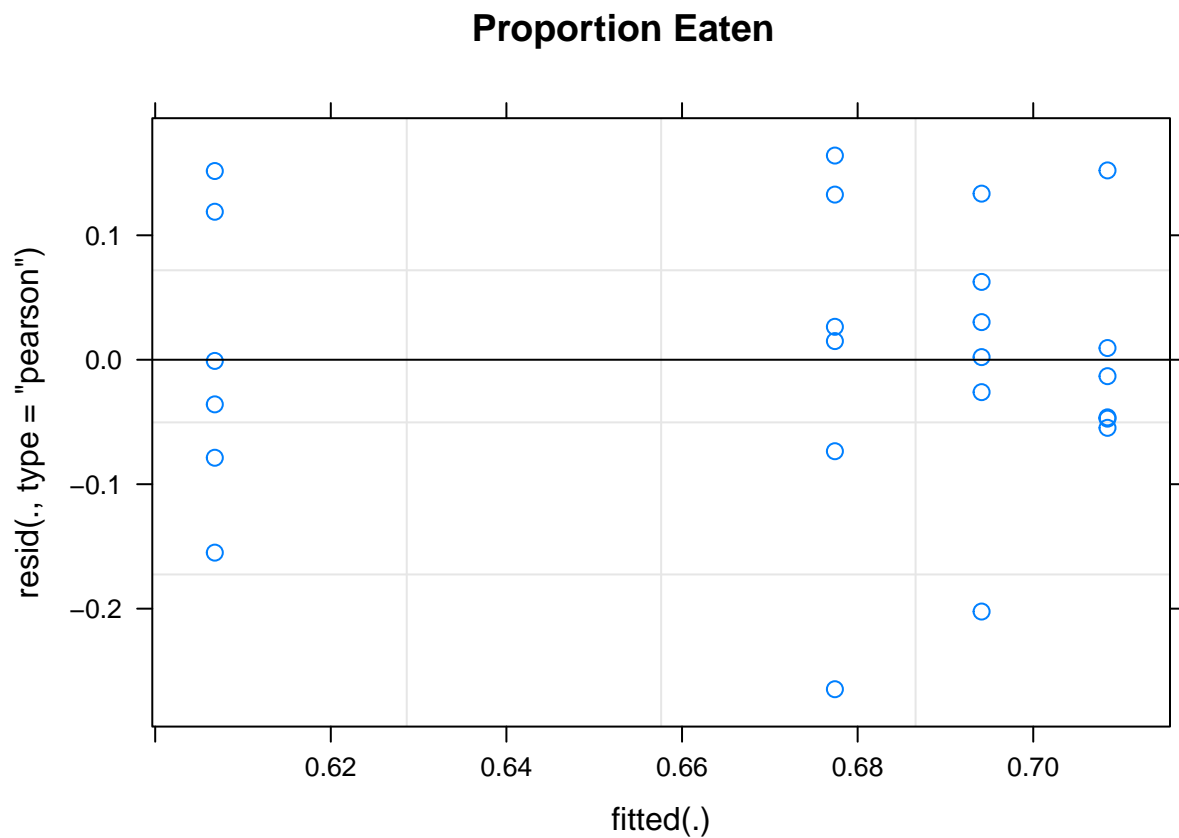
```
##  
## Shapiro-Wilk normality test  
##  
## data: insects6$prop_eaten  
## W = 0.9587, p-value = 0.4128
```

```
# aka normal distribution!
```

```
# Assumption checking  
m1 <- lmer(prop_eaten ~ subplot_descriptions + (1 | replicate),  
  data = insects6, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
# Check Assumptions: (1) Linearity: if covariates are not  
# categorical (2) Homogeneity: Need to Check by plotting  
# residuals vs predicted values.  
plot(m1, main = "Proportion Eaten")
```



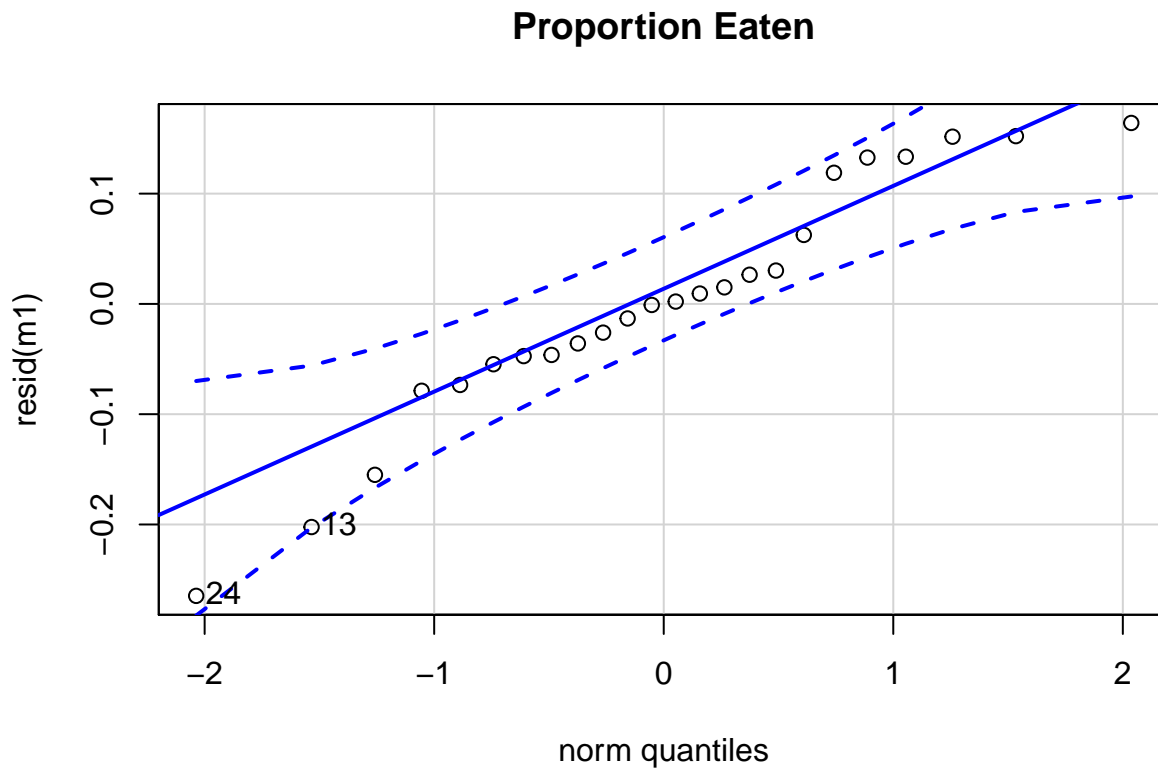
```
# Homogeneity of variance is ok here (increasing variance in  
# resids is not increasing with fitted values) Check for  
# homogeneity of variances (true if p>0.05). If the result is
```

```
# not significant, the assumption of equal variances
# (homoscedasticity) is met (no significant difference
# between the group variances).
leveneTest(residuals(m1) ~ insects6$subplot_descriptions)
```

```
## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  0.5634 0.6454
##      20
```

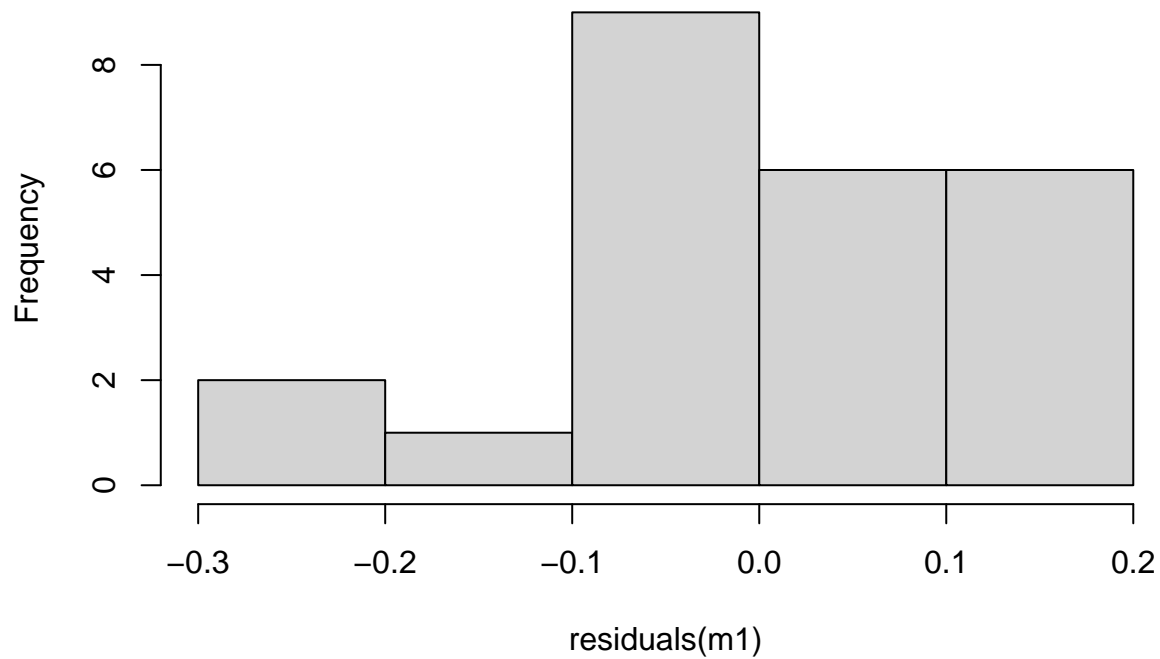
```
# Assumption met (3) Normality of error term: need to check
# by histogram, plot of residuals, could do
# Kolmogorov-Smirnov test. Check for normal residuals
qqPlot(resid(m1), main = "Proportion Eaten")
```



```
## 24 13
## 11 6
```

```
hist(residuals(m1), main = "Proportion Eaten")
```

Proportion Eaten



```
shapiro.test(resid(m1)) # Normal
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: resid(m1)  
## W = 0.94654, p-value = 0.2277
```

```
# Model comparisons  
m2 <- lm(prop_eaten ~ subplot_descriptions, data = insects6)  
m3 <- lmer(prop_eaten ~ subplot_descriptions + (1 | replicate),  
  data = insects6, REML = F)
```

```
## boundary (singular) fit: see ?isSingular
```

```
m4 <- lmer(prop_eaten ~ subplot_descriptions + (1 | replicate/footprint_location),  
  data = insects6, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AICcTab(m1, m2, m3, m4, weights = T)
```

```
##      dAICc df weight  
## m2 0.0    5  0.740  
## m1 3.6    6  0.122  
## m3 3.6    6  0.122  
## m4 7.7    7  0.016
```

```
# Model 2 fits the best
summary(m2)
```

```
##
## Call:
## lm(formula = prop_eaten ~ subplot_descriptions, data = insects6)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.264733 -0.049184  0.000584  0.076642  0.164070
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.67741    0.04870   13.909  9.6e-12 ***
## subplot_descriptionsdrought -0.07062    0.06887   -1.025    0.317
## subplot_descriptionswarmed  0.03105    0.06887    0.451    0.657
## subplot_descriptionswarmed_drought 0.01671    0.06887    0.243    0.811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1193 on 20 degrees of freedom
## Multiple R-squared:  0.1139, Adjusted R-squared:  -0.01897
## F-statistic: 0.8572 on 3 and 20 DF,  p-value: 0.4793
```

```
summ(m2)
```

Observations	24
Dependent variable	prop_eaten
Type	OLS linear regression

F(3,20)	0.86
R ²	0.11
Adj. R ²	-0.02

	Est.	S.E.	t val.	p
(Intercept)	0.68	0.05	13.91	0.00
subplot_descriptionsdrought	-0.07	0.07	-1.03	0.32
subplot_descriptionswarmed	0.03	0.07	0.45	0.66
subplot_descriptionswarmed_drought	0.02	0.07	0.24	0.81

Standard errors: OLS

```
# Post hoc test to compare different levels
emmeans(m2, list(pairwise ~ subplot_descriptions), adjust = "tukey")
```

```
## $'emmeans of subplot_descriptions'
## subplot_descriptions emmean SE df lower.CL upper.CL
## ambient              0.677 0.0487 20    0.576    0.779
## drought              0.607 0.0487 20    0.505    0.708
```

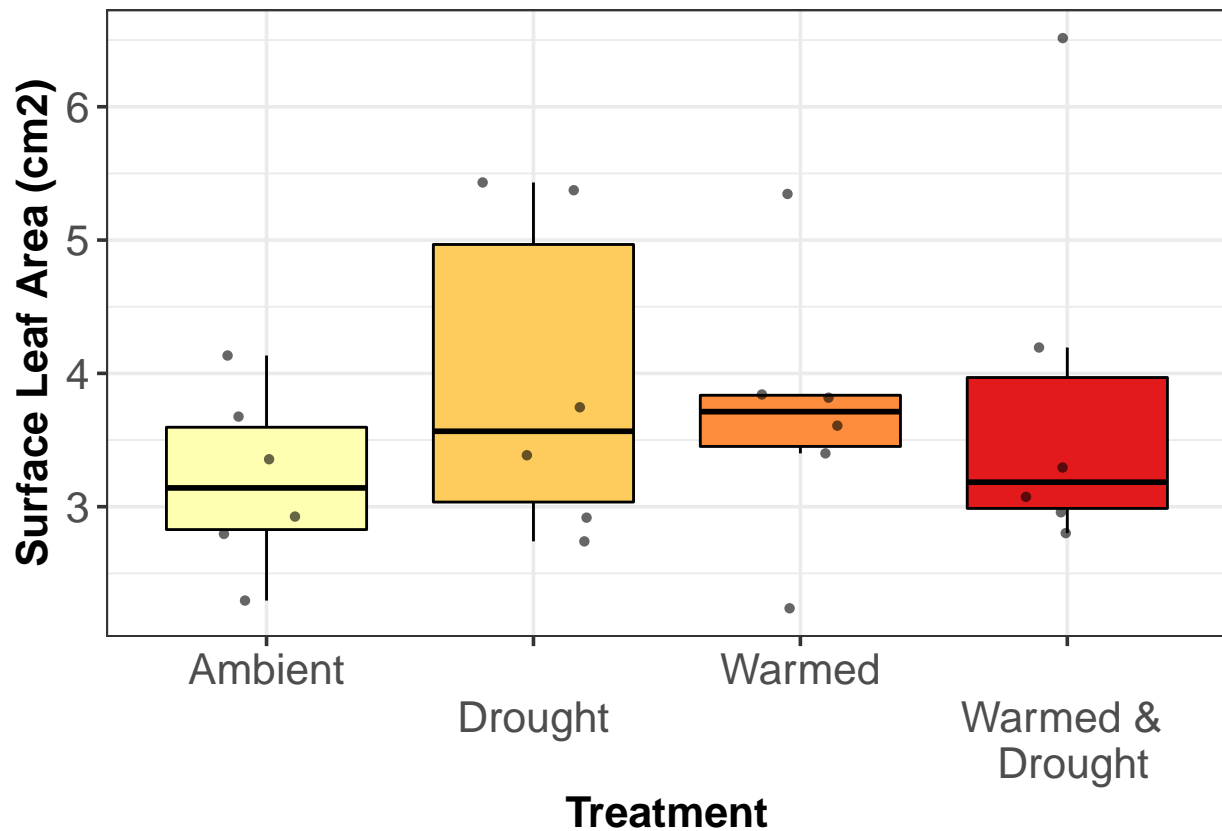
```
## warmed                0.708 0.0487 20    0.607    0.810
## warmed_drought        0.694 0.0487 20    0.593    0.796
##
## Confidence level used: 0.95
##
## $'pairwise differences of subplot_descriptions'
## 1                estimate      SE df t.ratio p.value
## ambient - drought      0.0706 0.0689 20   1.025  0.7368
## ambient - warmed      -0.0310 0.0689 20  -0.451  0.9687
## ambient - warmed_drought -0.0167 0.0689 20  -0.243  0.9948
## drought - warmed      -0.1017 0.0689 20  -1.476  0.4696
## drought - warmed_drought -0.0873 0.0689 20  -1.268  0.5928
## warmed - warmed_drought  0.0143 0.0689 20   0.208  0.9967
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

```
# Looking at SLA (leaf traits) at subplot level
sla_before_trt <- left_join(meta, sla_before)
```

```
## Joining, by = c("replicate", "footprint_location", "subplot_location")
```

```
sla_before_trt <- na.omit(sla_before_trt)

# png('soca_SLA_treatments.png', units='in', width=7,
# height=6, res=300)
ggplot(sla_before_trt, aes(x = subplot_descriptions, y = sla_before_mean,
  fill = subplot_descriptions)) + geom_boxplot(color = "black",
  outlier.shape = NA) + geom_jitter(shape = 16, position = position_jitterdodge(),
  alpha = 0.6) + # geom_errorbar(aes(ymin = lower, ymax = upper)) +
labs(x = "Treatment", y = "Surface Leaf Area (cm2)") + scale_fill_brewer(palette = "YlOrRd") +
  scale_x_discrete(limits = c("ambient", "drought", "warmed",
    "warmed_drought"), labels = c(ambient = "Ambient", drought = "Drought",
    warmed = "Warmed", warmed_drought = "Warmed & \n Drought"),
  guide = guide_axis(n.dodge = 2)) + theme(legend.position = "none")
```



```
# dev.off()
```

```
# not much of a difference between treatments in SLA
```

```
# Regression
```

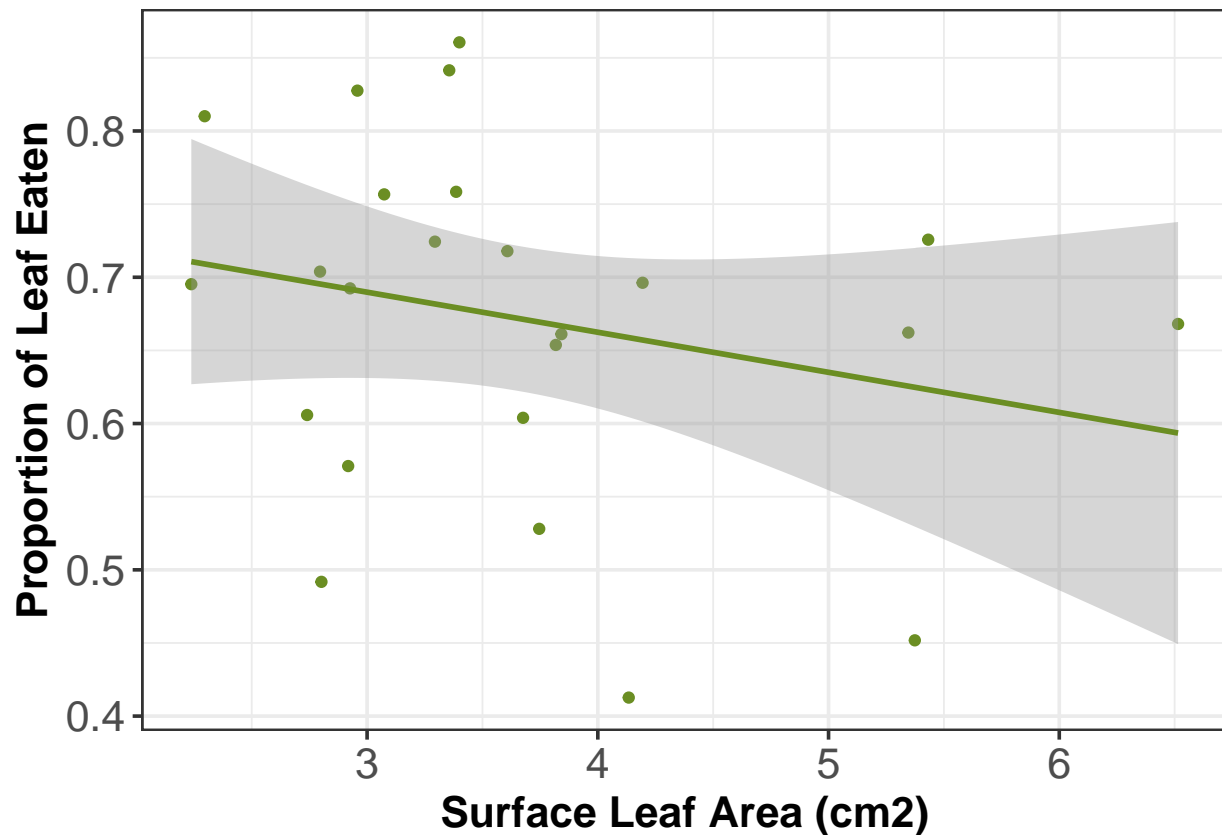
```
lm_plot_prop <- lm(prop_eaten ~ sla_before_mean, data = insects6)
```

```
# png('regression_preference.png', units='in', width=5,
```

```
# height=5, res=300)
```

```
ggplot(insects6, aes(sla_before_mean, prop_eaten)) + geom_point(color = "olivedrab") +  
  geom_smooth(method = "lm", color = "olivedrab") + labs(x = "Surface Leaf Area (cm2)",  
  y = "Proportion of Leaf Eaten")
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

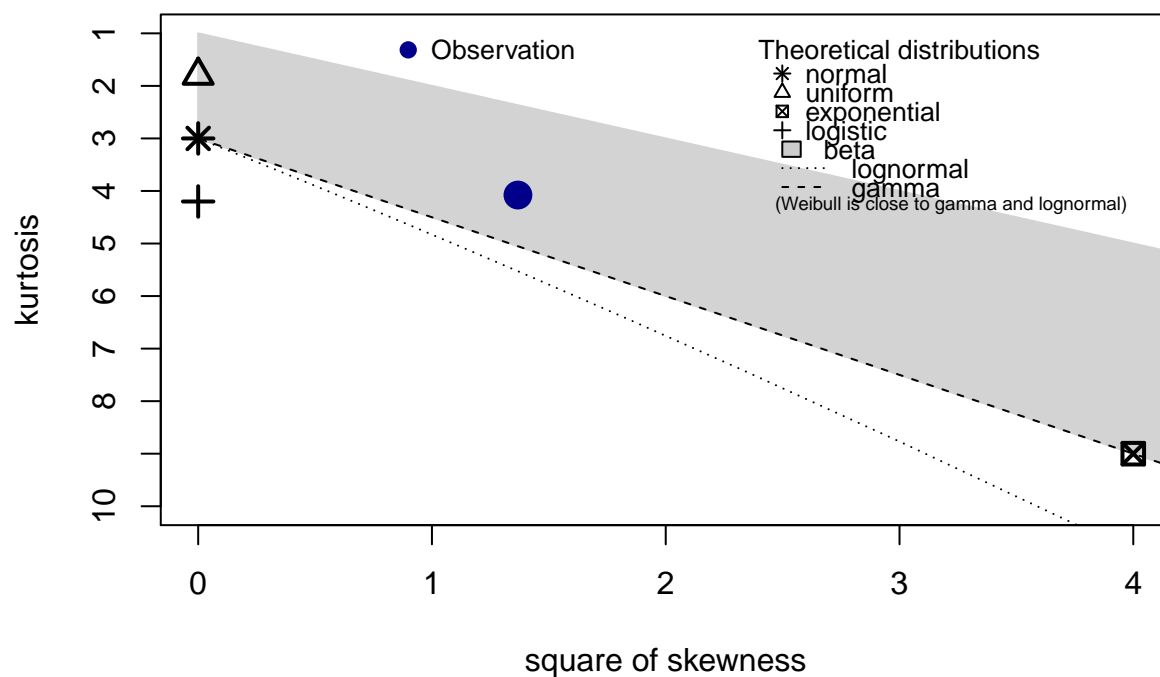
```
# dev.off()
summary(lm_plot_prop)
```

```
##
## Call:
## lm(formula = prop_eaten ~ sla_before_mean, data = insects6)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24607 -0.07330  0.02255  0.07571  0.18174
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.77198    0.08716   8.857 1.05e-08 ***
## sla_before_mean -0.02739    0.02289  -1.196   0.244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1171 on 22 degrees of freedom
## Multiple R-squared:  0.06109,    Adjusted R-squared:  0.01842
## F-statistic: 1.432 on 1 and 22 DF,  p-value: 0.2443
```

```
# Data Exploration and Analyses for SLA
```

```
descdist(sla_before_trt$sla_before_mean, discrete = FALSE)
```

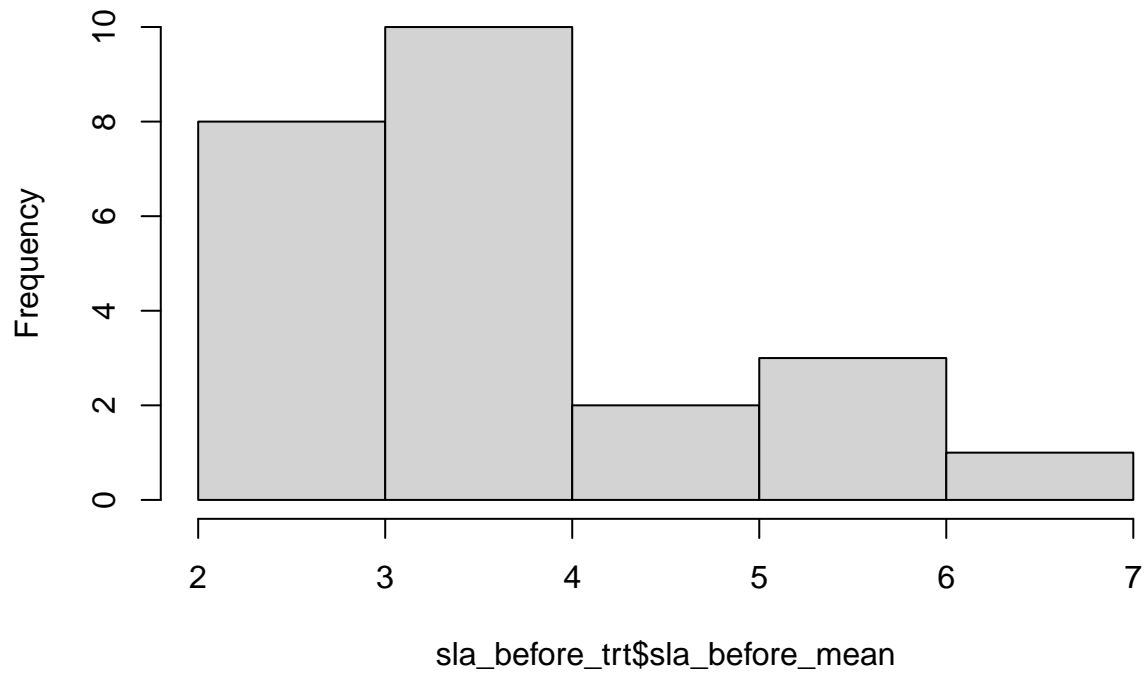
Cullen and Frey graph



```
## summary statistics
## -----
## min: 2.238 max: 6.515
## median: 3.393
## mean: 3.661208
## estimated sd: 1.066401
## estimated skewness: 1.16953
## estimated kurtosis: 4.077587
```

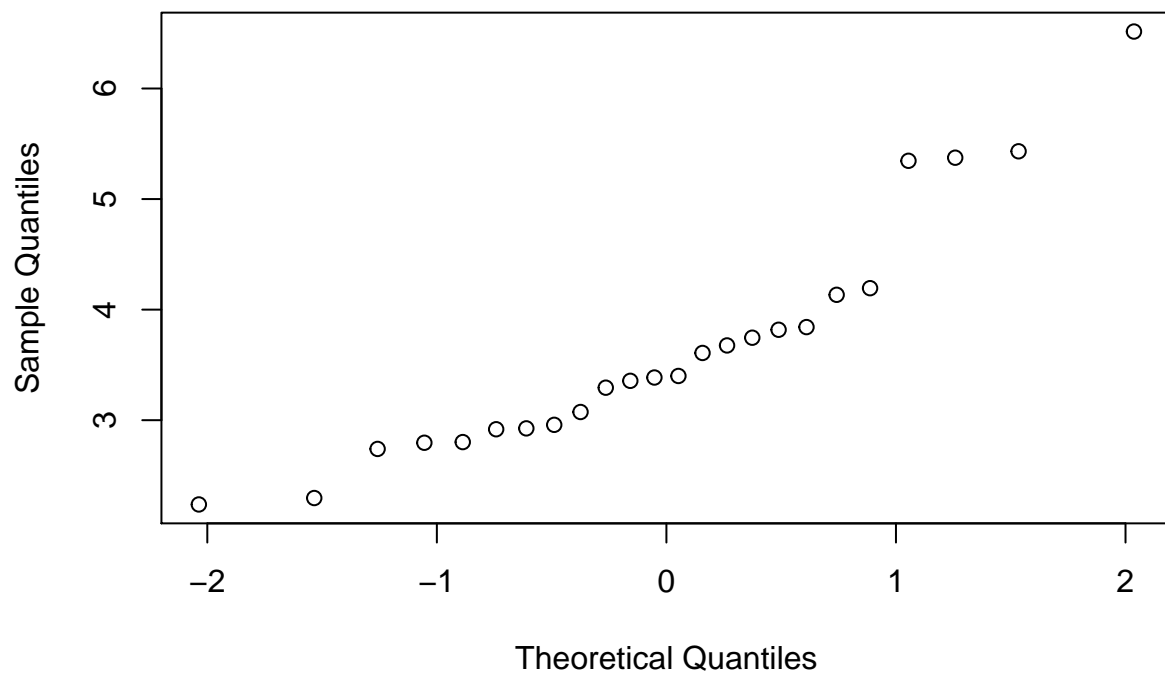
```
hist(sla_before_trt$sla_before_mean) # right skewed
```

Histogram of sla_before_trt\$sla_before_mean



```
qqnorm(sla_before_trt$sla_before_mean)
```

Normal Q-Q Plot



```
shapiro.test(sla_before_trt$sla_before_mean)  # p-value = 0.01431 so we reject the null hypothesis that
```

```
##
```

```
##  Shapiro-Wilk normality test
```

```
##
```

```
## data:  sla_before_trt$sla_before_mean
```

```
## W = 0.89155, p-value = 0.01431
```