

# REX: Gall Analyses

Moriah Young, Kara Dobson

July 28, 2021

COLLABORATORS: Phoebe Zarnetske, Mark Hammond, Emily Parker, Kristin Wolford DATA INPUT:  
Clean & plot gall csv from the shared Google drive  
DATA OUTPUT: Code and Rmd are in the scripts folder in Github  
PROJECT: REX

```
# Clear all existing data  
rm(list=ls())
```

```
# Load packages  
library(bbmle)
```

```
## Loading required package: stats4
```

```
library(lmerTest)
```

```
## 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
```

```
library(fitdistrplus)
```

```
## Loading required package: MASS
```

```
## Loading required package: survival
```

```
library(sjPlot)
```

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

```
library(car)
```

```
## 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
```

```
library(emmeans)
```

```
library(tidyverse)
```

```
## -- 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()
```

```
library(ggpubr)
```

```
library(jtools) #summ
```

```
library(rstatix)
```

```
##
```

```
## Attaching package: 'rstatix'
```

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

```
##
```

```
##   select
```

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

```
##
```

```
##   filter
```

```
# Set working directory from .Renviron
dir <- Sys.getenv("ANALYSIS_DIR")
list.files(dir)
```

```
## [1] "animal"          "GHG"
## [3] "Irrigation"      "lookup_tables"
## [5] "microbes"        "REX_Data_Management.gdoc"
## [7] "REX_Data_Template.gsheet" "soil"
## [9] "T7_ANPP"         "T7_plant_comp"
## [11] "T7_plant_phenology" "T7_warmx_insect"
## [13] "T7_warmx_plant_traits" "T7_warmx_VOC"
## [15] "weather"
```

```
# Read in data
```

```
galls <- read.csv(file.path(dir, "T7_warmx_plant_traits/L1/T7_warmx_galls_L1.csv"))
```

```
# making it so that plant_height, gall_diameter, and gall_height are on the same row for the same plant
# so each unique plant will have just one row ompared to two in the raw data.
```

```
#galls2 <- galls %>%
#   select(-date) %>%
#   group_by(rep, footprint, treatment, plant_num, gall_present) %>%
#   summarise(across(everything(), na.omit)) %>%
#   distinct()
```

```
galls3<- galls %>% dplyr::select(rep, footprint, treatment, plant_num, gall_present, plant_height) %>%
  na.omit() #this is just plant height data
```

```
galls4 <- galls %>% dplyr::select(-plant_height) %>%
  na.omit() %>%
  full_join(galls3)
```

```
## Joining, by = c("rep", "footprint", "treatment", "plant_num", "gall_present")
```

```
galls5 <- galls4 %>% dplyr::select(-date) # get rid of date - we don't need this and some have NAs
```

```
# create a dataframe that only contains galled plants
```

```
gall_only <- galls5[!(galls5$gall_present == "no_gall"),]
gall_only <- gall_only[-131, ]
```

```
# Take subplot average of plant height
```

```
plant_height <- galls5 %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(plant_height = mean(plant_height, na.rm = TRUE))
```

```
## 'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the 'gr
```

```
# Take subplot average of gall height
```

```
gall_height <- galls5 %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(gall_height = mean(gall_height, na.rm = TRUE))
```

```
## 'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the 'gr
```

```
# Take subplot average of gall diameter
gall_diameter <- galls5 %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(gall_diameter = mean(gall_diameter, na.rm = TRUE))
```

## 'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the '.groups' argument.

```
mean_galls <- left_join(plant_height, gall_height)
```

```
## Joining, by = c("rep", "footprint", "treatment", "gall_present")
```

```
mean_galls <- left_join(mean_galls, gall_diameter)
```

```
## Joining, by = c("rep", "footprint", "treatment", "gall_present")
```

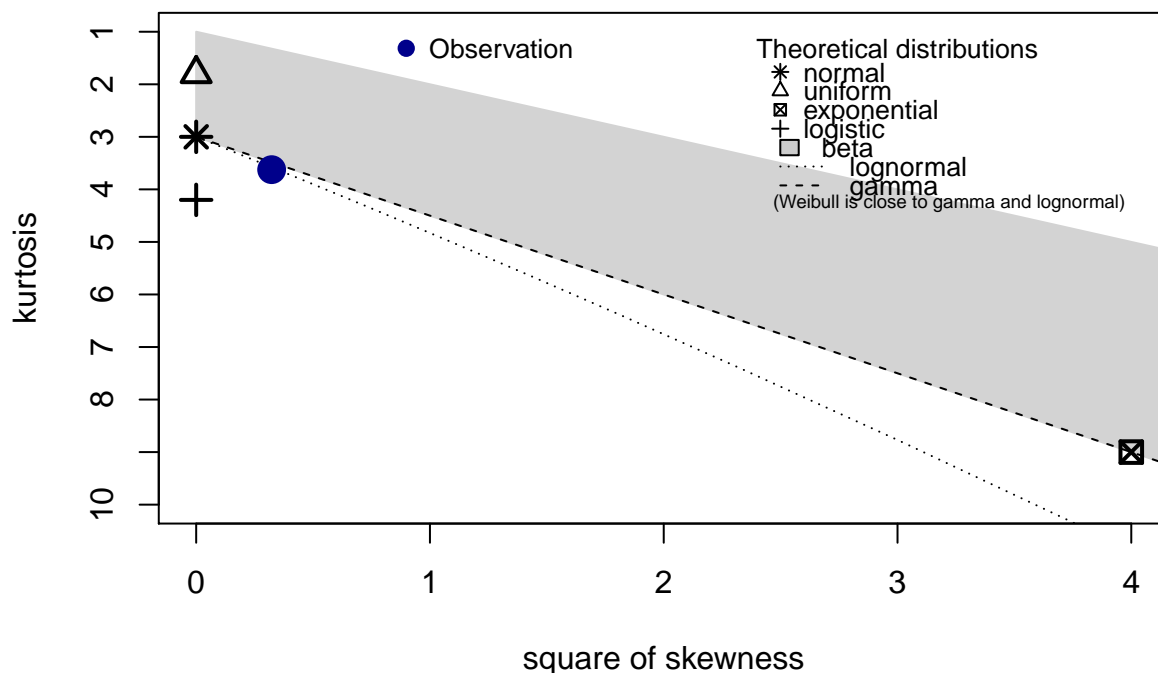
```
# Emily only wants to look at the effects of warming on goldenrod & galls so get rid of "drought" and "warmed_drought"
# gall_warmed <- galls2[!(galls2$treatment == "drought" | galls2$treatment == "warmed_drought"),]
```

Emily's questions: 1. Is there a relationship between plant height & likeliness of galling (i.e. are taller plants more likely to get a gall, or have multiple galls)? 2. Are plant height & area of gall related?

Looking at plant height first

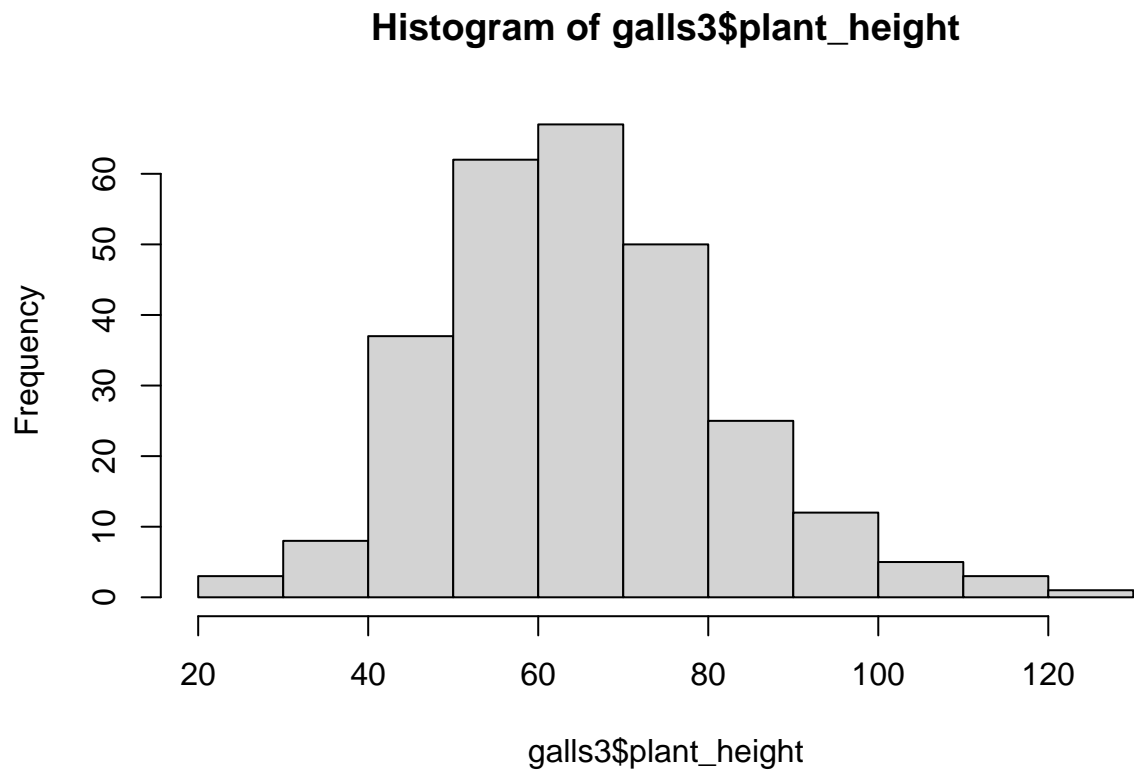
```
# Data exploration
descdist(galls3$plant_height, discrete = FALSE)
```

## Cullen and Frey graph



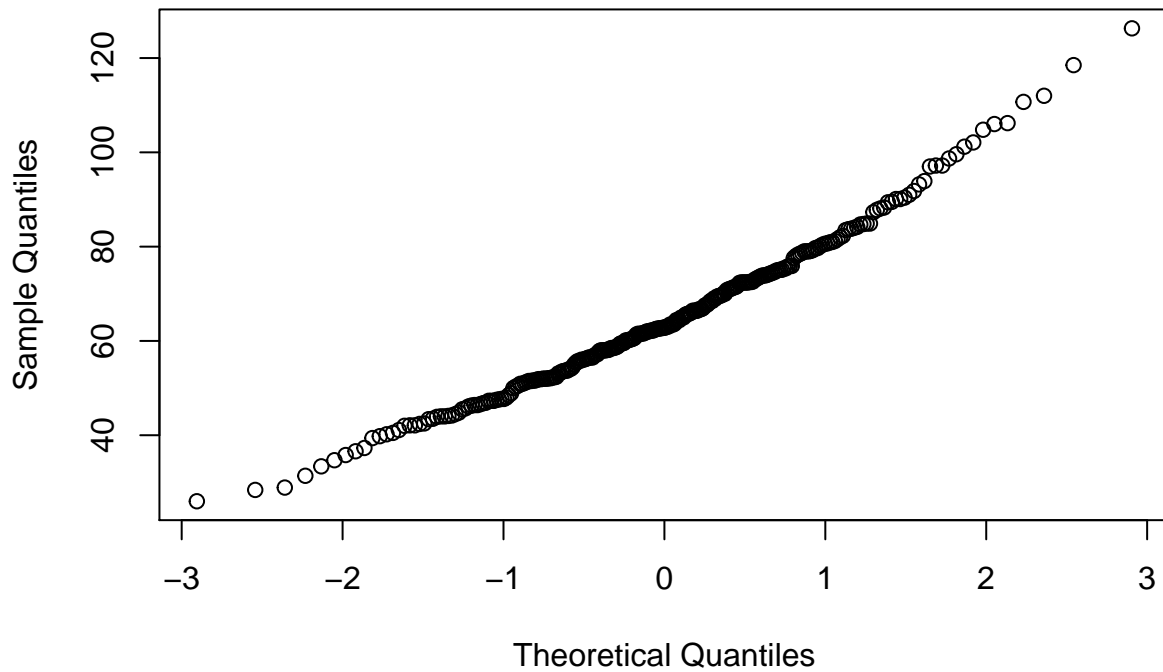
```
## summary statistics
## -----
## min: 26   max: 126.3
## median: 62.8
## mean: 64.92125
## estimated sd: 16.73332
## estimated skewness: 0.5681968
## estimated kurtosis: 3.624846
```

```
hist(galls3$plant_height)
```



```
qqnorm(galls3$plant_height)
```

## Normal Q-Q Plot

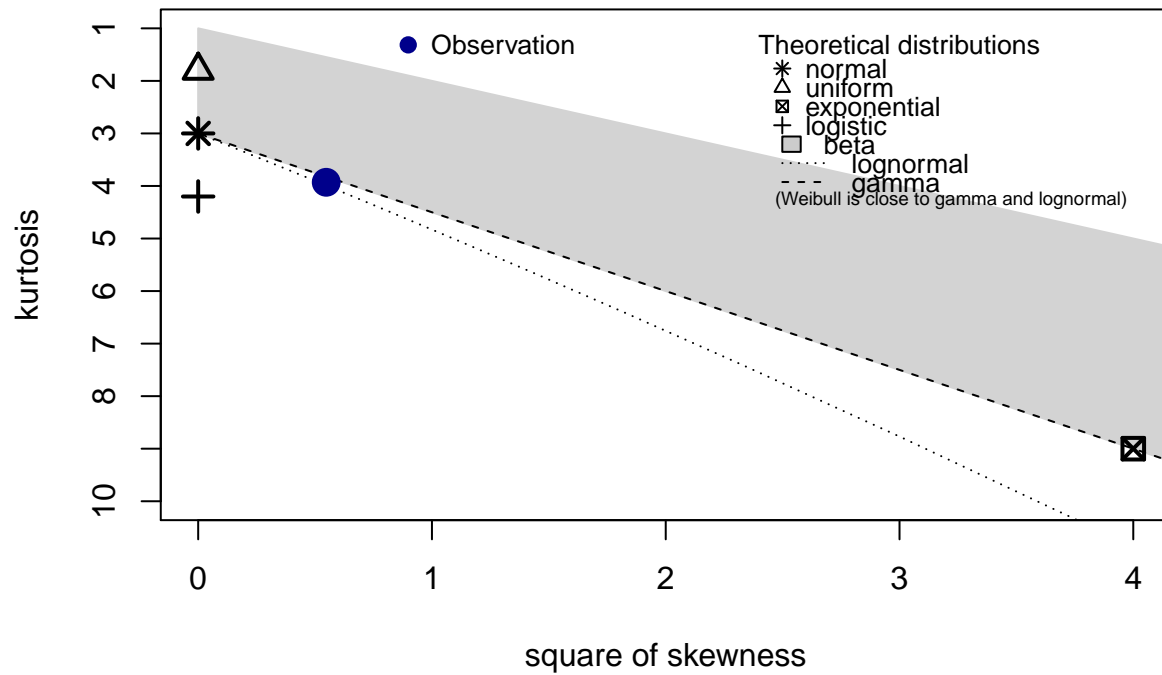


```
shapiro.test(galls3$plant_height) # not normally distributed
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  galls3$plant_height  
## W = 0.98138, p-value = 0.001245
```

```
# Let's look at gall vs no gall distributions  
no_gall_height <- galls3[(galls3$gall_present == "no_gall"),]  
just_gall_height <- galls3[(galls3$gall_present == "gall"),]  
  
# Data exploration for gall plants  
descdist(just_gall_height$plant_height, discrete = FALSE)
```

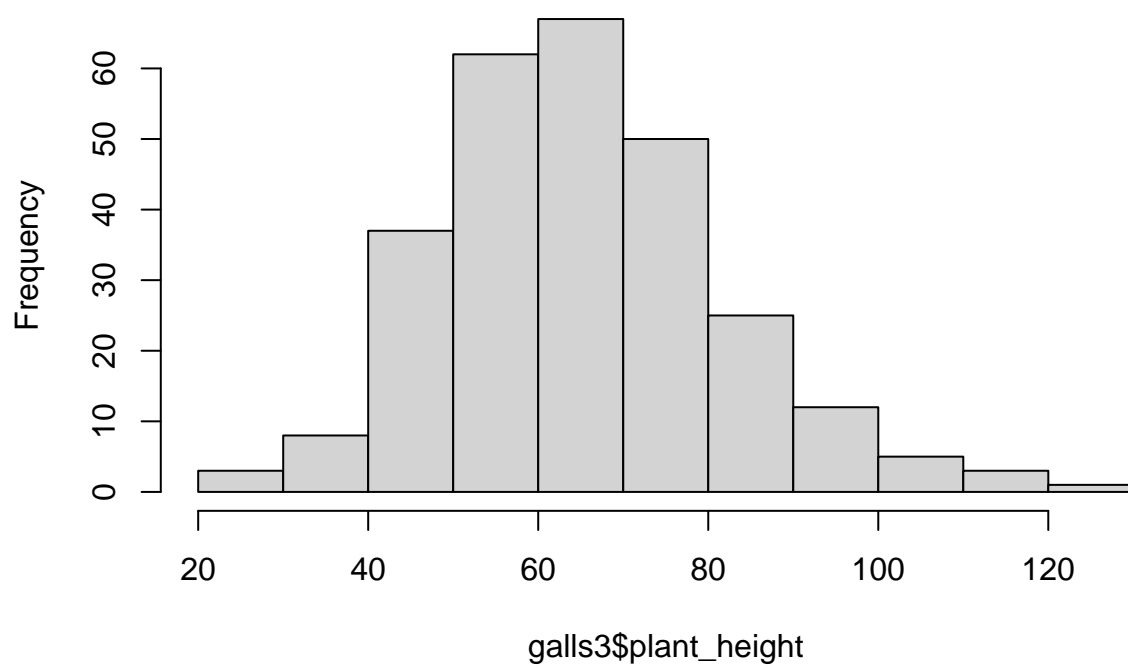
## Cullen and Frey graph



```
## summary statistics
## -----
## min: 37.3   max: 126.3
## median: 63.5
## mean: 66.32824
## estimated sd: 16.10932
## estimated skewness: 0.7400842
## estimated kurtosis: 3.929762
```

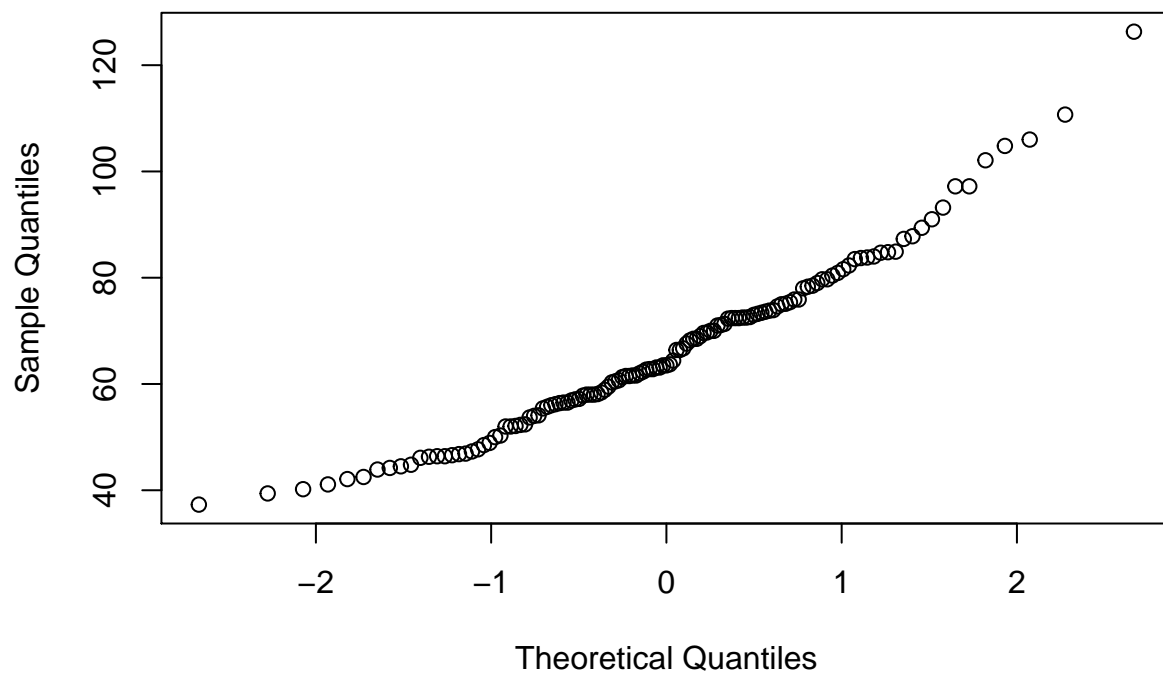
```
hist(galls3$plant_height)
```

**Histogram of galls3\$plant\_height**



```
qqnorm(just_gall_height$plant_height)
```

**Normal Q-Q Plot**



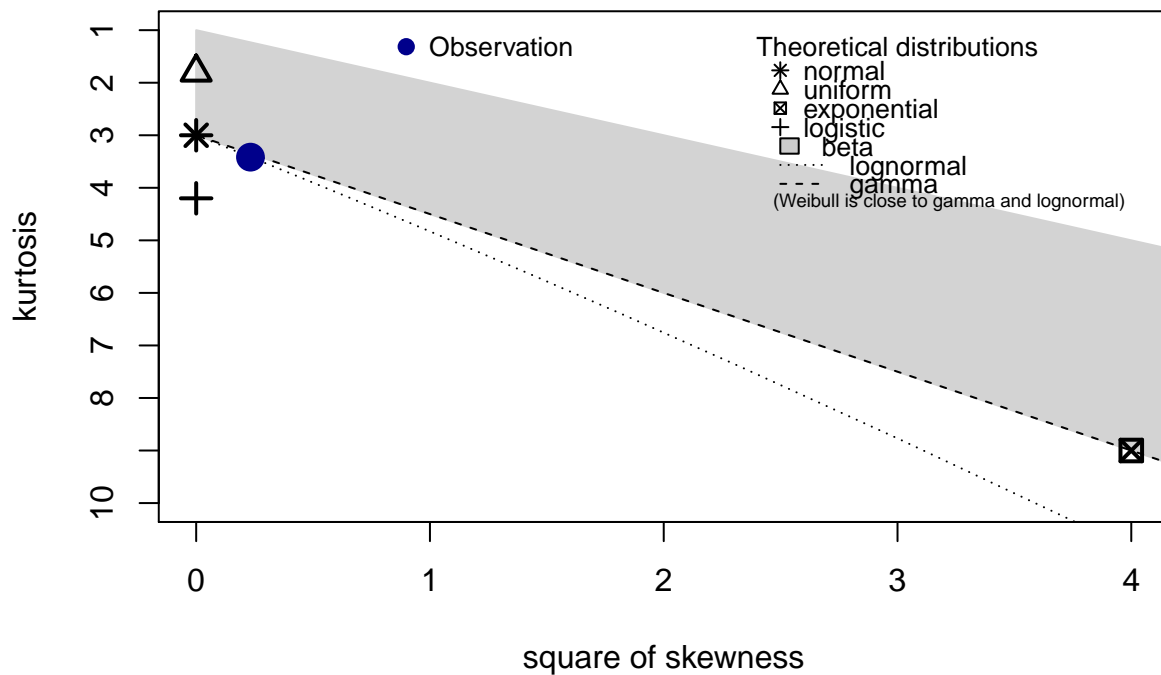


```
shapiro.test(just_gall_height$plant_height) # not normally distributed
```

```
##
## Shapiro-Wilk normality test
##
## data: just_gall_height$plant_height
## W = 0.96571, p-value = 0.002157
```

```
# Data exploration for non gall plants
descdist(no_gall_height$plant_height, discrete = FALSE)
```

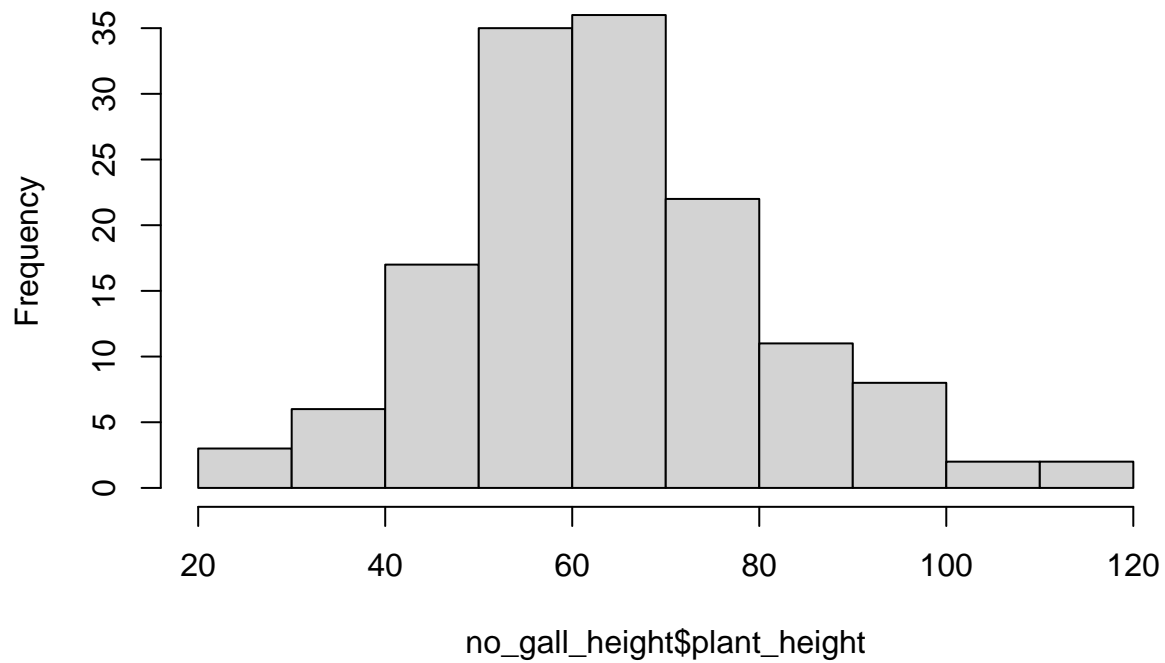
## Cullen and Frey graph



```
## summary statistics
## -----
## min: 26 max: 118.5
## median: 62.35
## mean: 63.62324
## estimated sd: 17.2438
## estimated skewness: 0.4819897
## estimated kurtosis: 3.416888
```

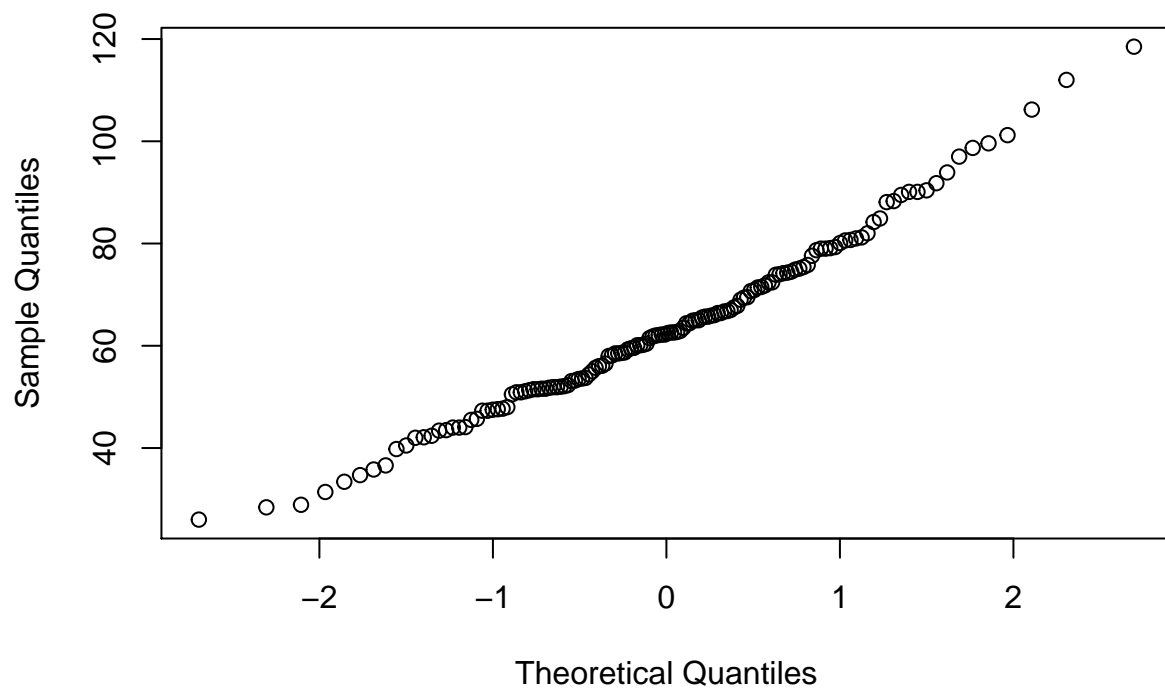
```
hist(no_gall_height$plant_height)
```

**Histogram of no\_gall\_height\$plant\_height**



```
qqnorm(no_gall_height$plant_height)
```

**Normal Q-Q Plot**

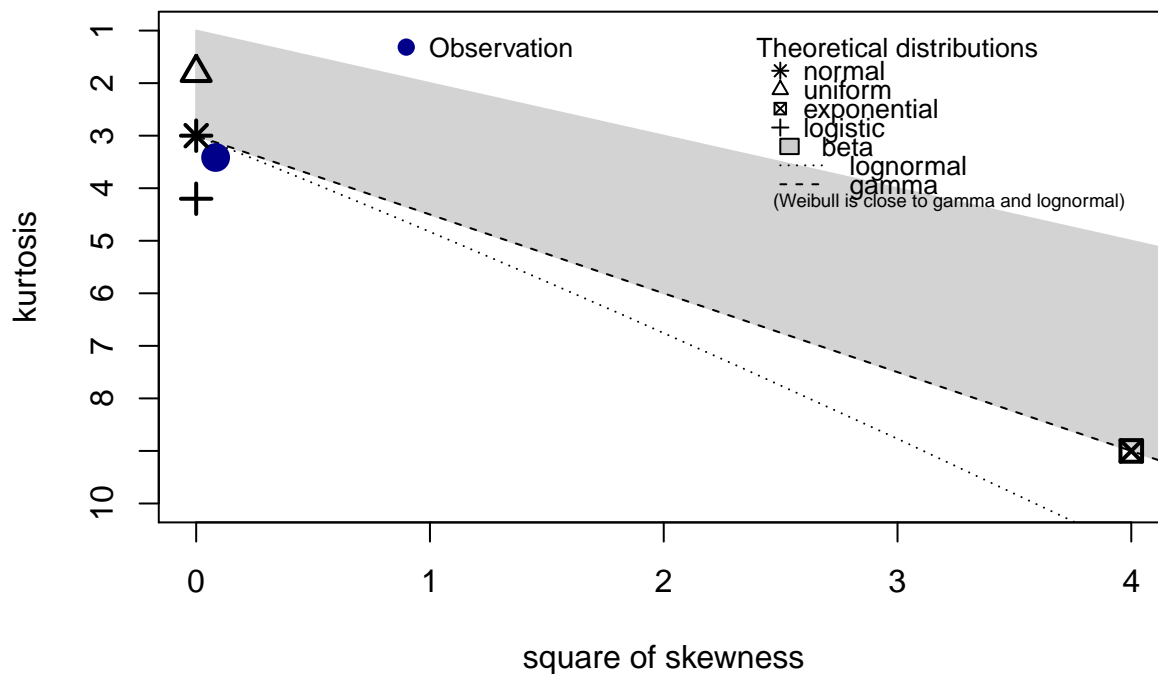


```
shapiro.test(no_gall_height$plant_height) # normally distributed
```

```
##
## Shapiro-Wilk normality test
##
## data: no_gall_height$plant_height
## W = 0.9835, p-value = 0.08565
```

```
# let's take the lognormal of plant height (gall and no gall)
galls3$lph <- log(galls3$plant_height)
descdist(galls3$lph, discrete = FALSE)
```

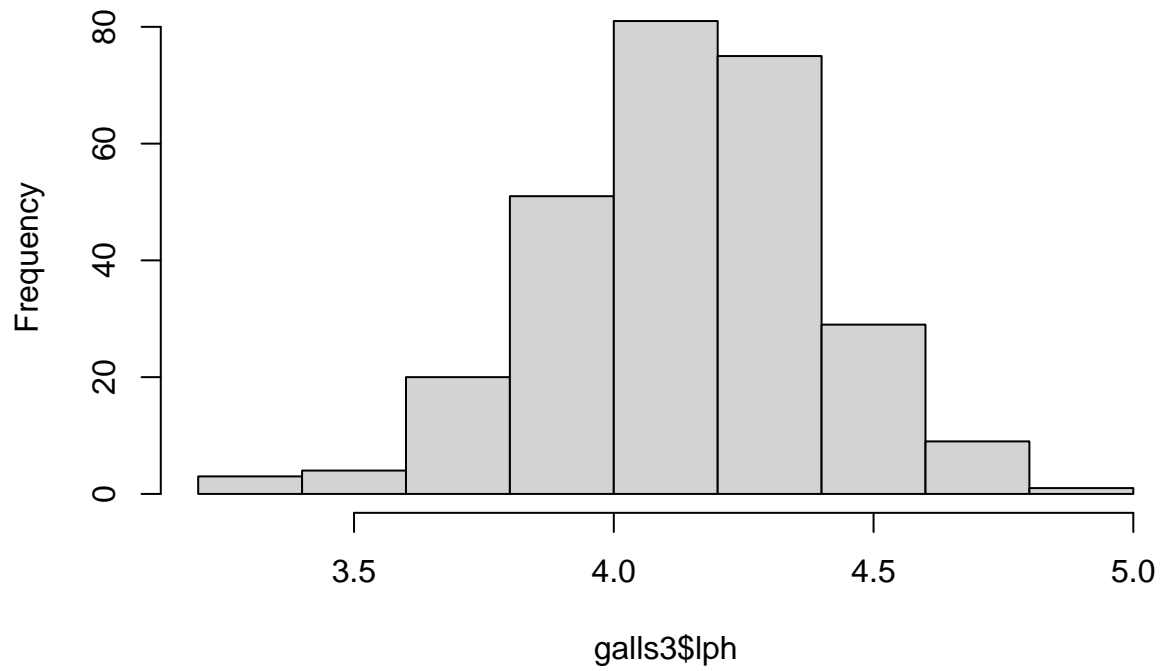
## Cullen and Frey graph



```
## summary statistics
## -----
## min: 3.258097 max: 4.83866
## median: 4.139955
## mean: 4.139809
## estimated sd: 0.2618149
## estimated skewness: -0.2884268
## estimated kurtosis: 3.415224
```

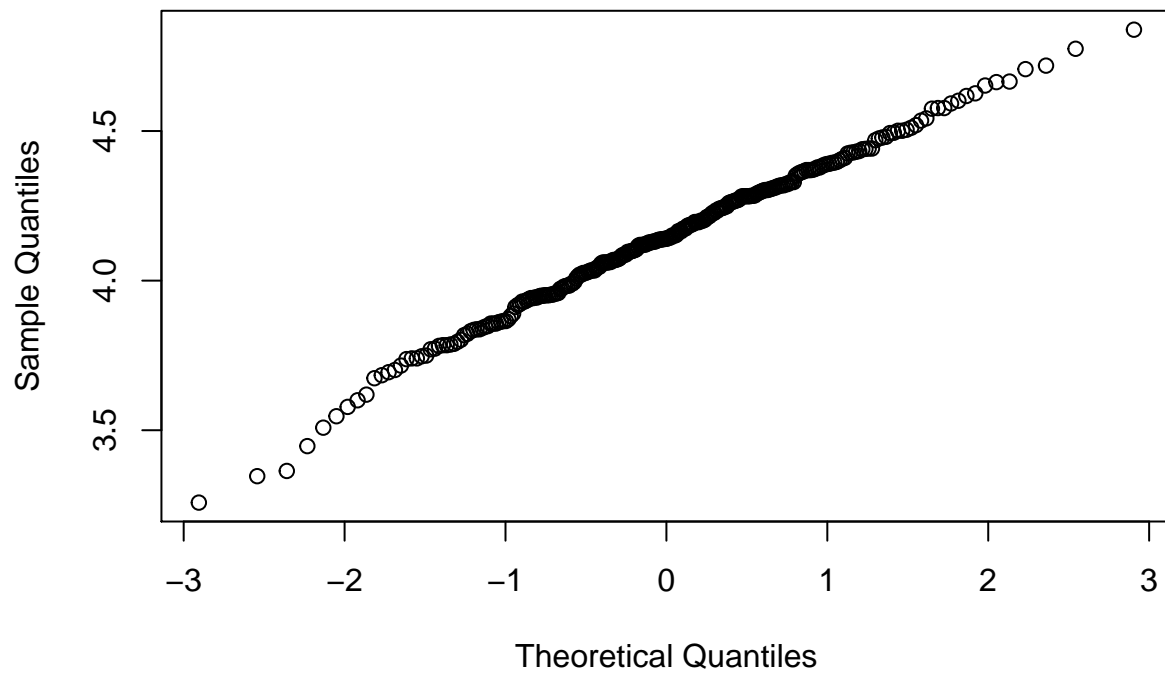
```
hist(galls3$lph)
```

### Histogram of galls3\$lph



```
qqnorm(galls3$lph)
```

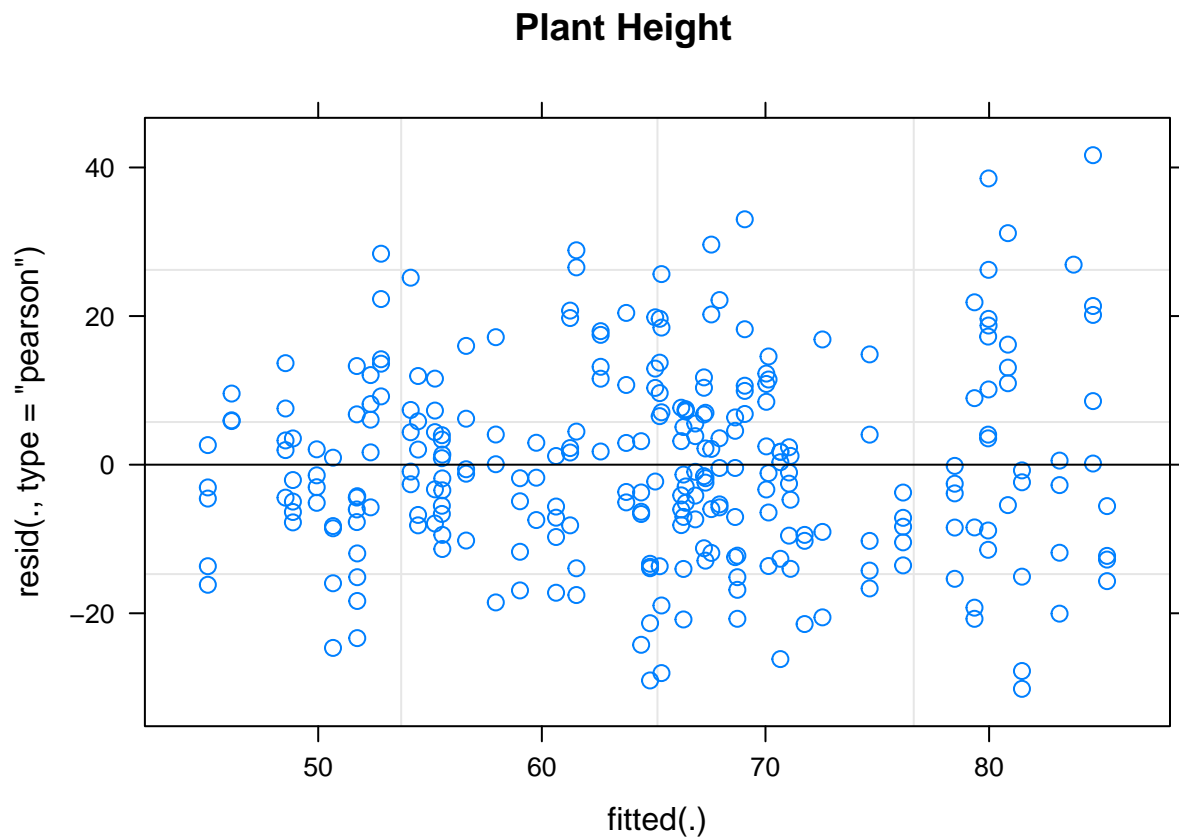
### Normal Q-Q Plot



```
shapiro.test(galls3$lph) # normally distributed!
```

```
##
## Shapiro-Wilk normality test
##
## data: galls3$lph
## W = 0.99305, p-value = 0.2349
```

```
# Assumption checking
m1 <- lmer(plant_height~ treatment + gall_present + (1|rep), data = galls3, REML=FALSE)
# Check Assumptions:
# (1) Linearity: if covariates are not categorical
# (2) Homogeneity: Need to Check by plotting residuals vs predicted values.
plot(m1, main = "Plant Height")
```



```
# Homogeneity of variance is ok here (increasing variance in resids is not increasing with fitted value)
# Check for homogeneity of variances (true if p>0.05). If the result is not significant, the assumption
leveneTest(residuals(m1) ~ galls3$treatment)
```

```
## 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  4  4.1255 0.002922 **
```

```
##          268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Assumption not met
```

```
leveneTest(residuals(m1) ~ galls3$gall_present)
```

```
## 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    1  4.4471 0.03588 *
##          271
## ---
```

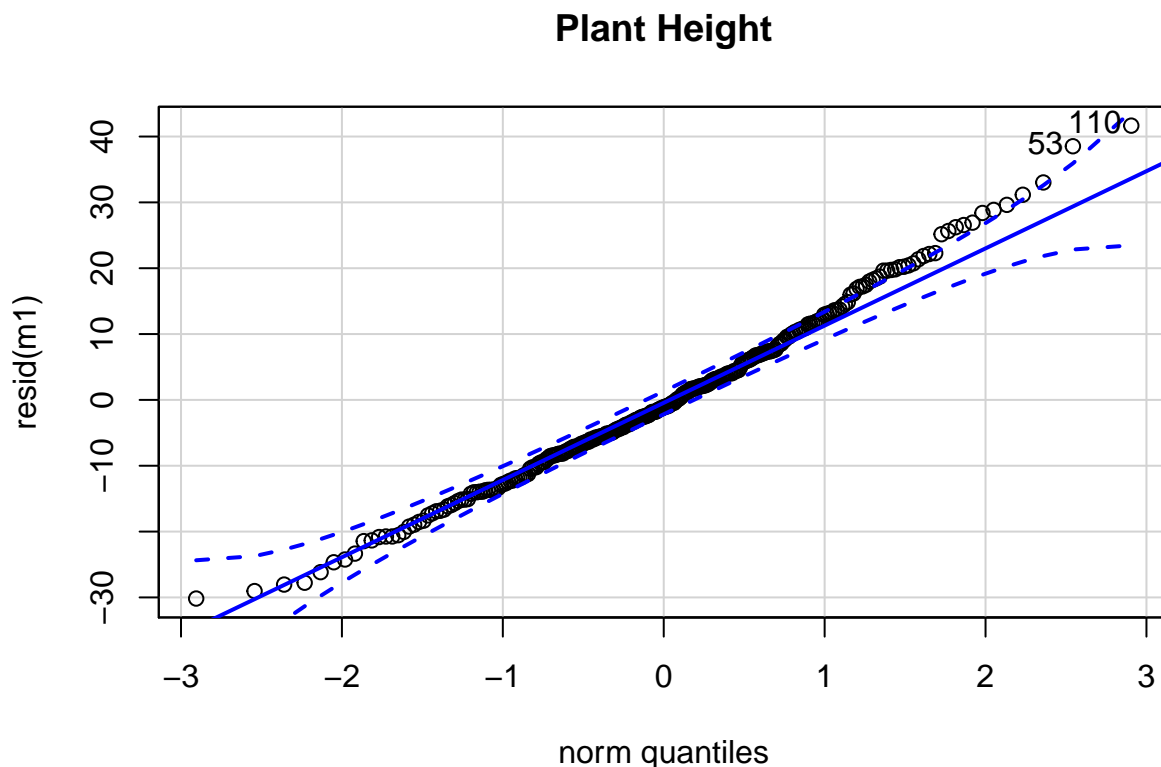
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Assumption not met
```

```
# (3) Normality of error term: need to check by histogram, QQplot of residuals, could do Kolmogorov-Smi
```

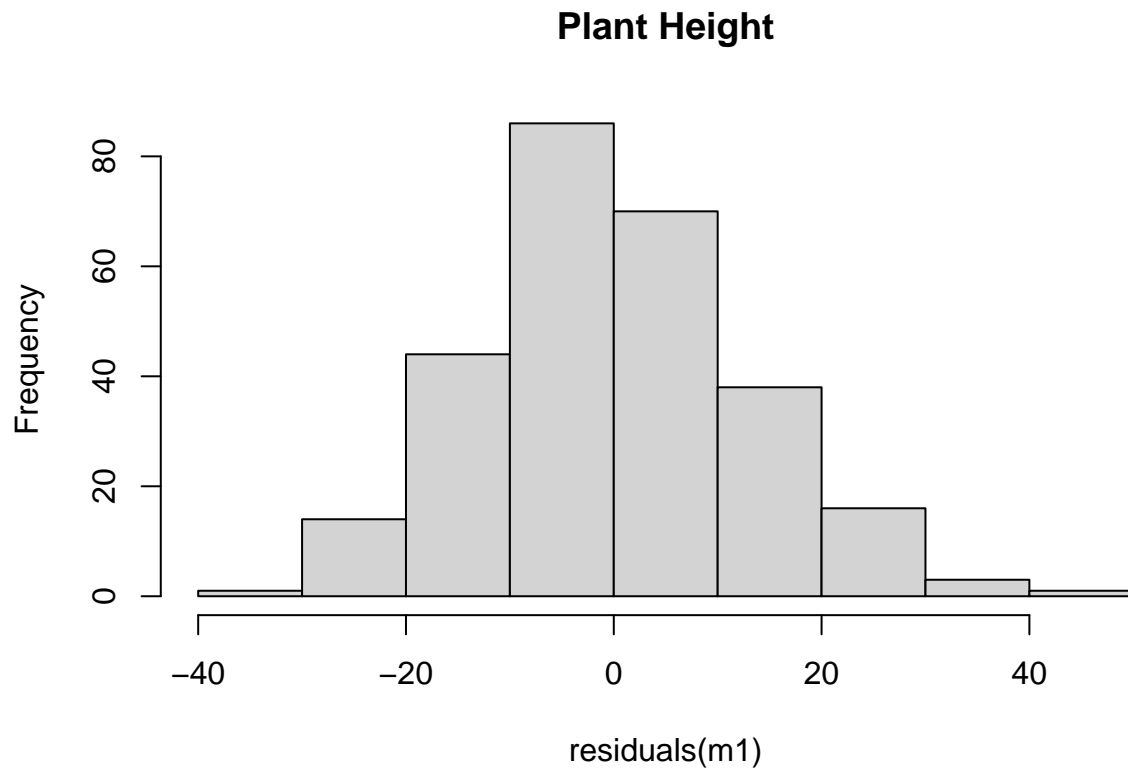
```
# Check for normal residuals
```

```
qqPlot(resid(m1), main = "Plant Height")
```



```
## [1] 110 53
```

```
hist(residuals(m1), main = "Plant Height")
```



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

```
##  
## Shapiro-Wilk normality test  
##  
## data: resid(m1)  
## W = 0.99091, p-value = 0.08861
```

```
# Model comparisons  
m2 <- lm(plant_height ~ treatment, data=galls3)  
m3 <- lm(plant_height ~ gall_present, data=galls3)  
m4 <- lmer(plant_height ~ treatment + (1|rep), data=galls3, REML=F)  
m5 <- lmer(plant_height ~ treatment * gall_present + (1|rep), data=galls3, REML=F)  
AICctab(m1, m2, m3, m4, m5, weights=T)
```

```
##      dAICc df weight  
## m5    0.0 12 0.686  
## m1    1.9  8 0.269  
## m4    5.4  7 0.045  
## m2   53.3  6 <0.001  
## m3  110.2  3 <0.001
```

```
# Models 5 and 1 fits the best  
summ(m5)
```

Observations	273
Dependent variable	plant_height
Type	Mixed effects linear regression

AIC	2204.95
BIC	2248.27
Pseudo-R <sup>2</sup> (fixed effects)	0.25
Pseudo-R <sup>2</sup> (total)	0.42

Fixed Effects					
	Est.	S.E.	t val.	d.f.	p
(Intercept)	59.61	3.79	15.74	17.29	0.00
treatmentdrought	2.84	3.49	0.82	267.10	0.42
treatmentirr_control	-3.11	3.58	-0.87	267.19	0.39
treatmentwarmed	14.85	3.65	4.07	267.10	0.00
treatmentwarmed_drought	21.80	3.55	6.15	267.09	0.00
gall_presentno_gall	3.16	3.46	0.91	267.10	0.36
treatmentdrought:gall_presentno_gall	-7.20	4.80	-1.50	267.06	0.13
treatmentirr_control:gall_presentno_gall	-6.44	5.03	-1.28	267.07	0.20
treatmentwarmed:gall_presentno_gall	-5.20	4.92	-1.06	267.06	0.29
treatmentwarmed_drought:gall_presentno_gall	-15.60	4.86	-3.21	267.04	0.00

p values calculated using Satterthwaite d.f.

Random Effects		
Group	Parameter	Std. Dev.
rep	(Intercept)	6.84
Residual		12.76

Grouping Variables		
Group	# groups	ICC
rep	6	0.22

`summ(m1)`

Observations	273
Dependent variable	plant_height
Type	Mixed effects linear regression

AIC	2207.49
BIC	2236.36
Pseudo-R <sup>2</sup> (fixed effects)	0.23
Pseudo-R <sup>2</sup> (total)	0.40



Fixed Effects					
	Est.	S.E.	t val.	d.f.	p
(Intercept)	63.41	3.43	18.47	11.32	0.00
treatmentdrought	-1.08	2.44	-0.44	267.05	0.66
treatmentirr_control	-6.68	2.58	-2.59	267.58	0.01
treatmentwarmed	12.03	2.49	4.82	267.05	0.00
treatmentwarmed_drought	13.52	2.47	5.47	267.08	0.00
gall_presentno_gall	-3.80	1.59	-2.40	267.34	0.02

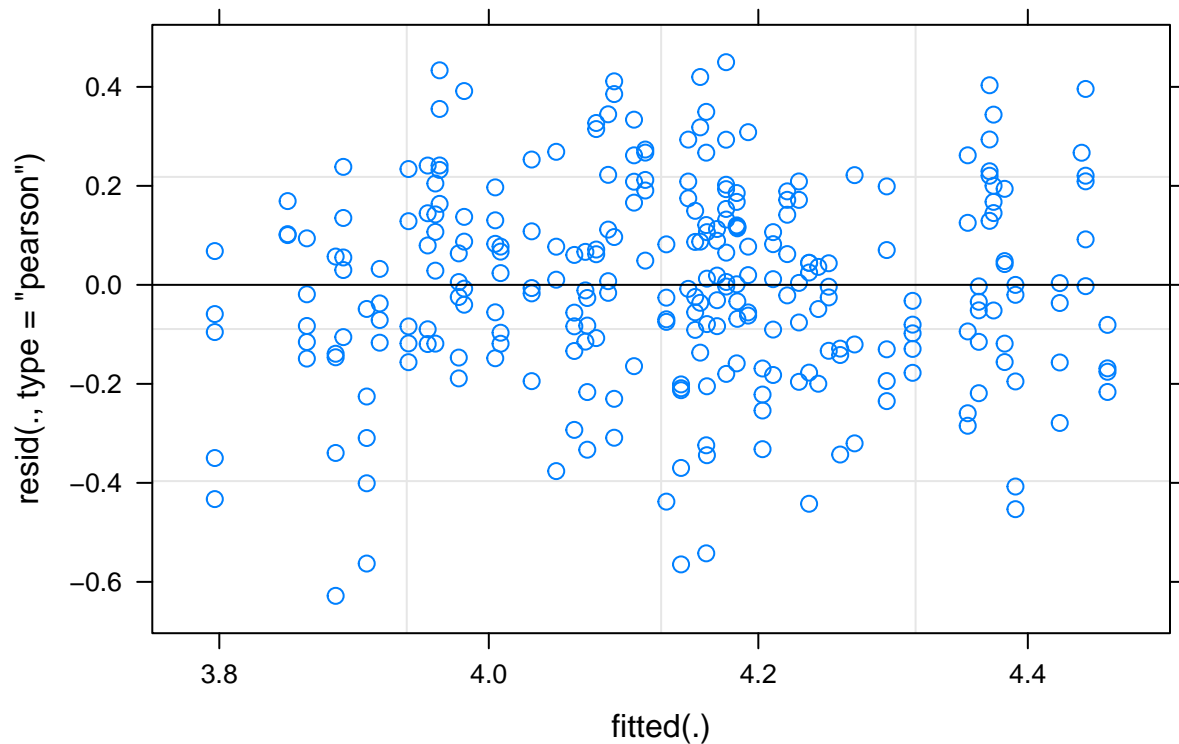
p values calculated using Satterthwaite d.f.

Random Effects		
Group	Parameter	Std. Dev.
rep	(Intercept)	6.91
Residual		13.01

Grouping Variables		
Group	# groups	ICC
rep	6	0.22

```
# checking assumptions with log of plant heights
m1l <- lmer(lph ~ treatment + gall_present + (1|rep), data = galls3, REML=FALSE)
# Check Assumptions:
# (1) Linearity: if covariates are not categorical
# (2) Homogeneity: Need to Check by plotting residuals vs predicted values.
plot(m1l, main = "Plant Height")
```

## Plant Height



```
# Homogeneity of variance is ok here (increasing variance in resids is not increasing with fitted value.
# Check for homogeneity of variances (true if p>0.05). If the result is not significant, the assumption
leveneTest(residuals(m1l) ~ galls3$treatment)
```

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

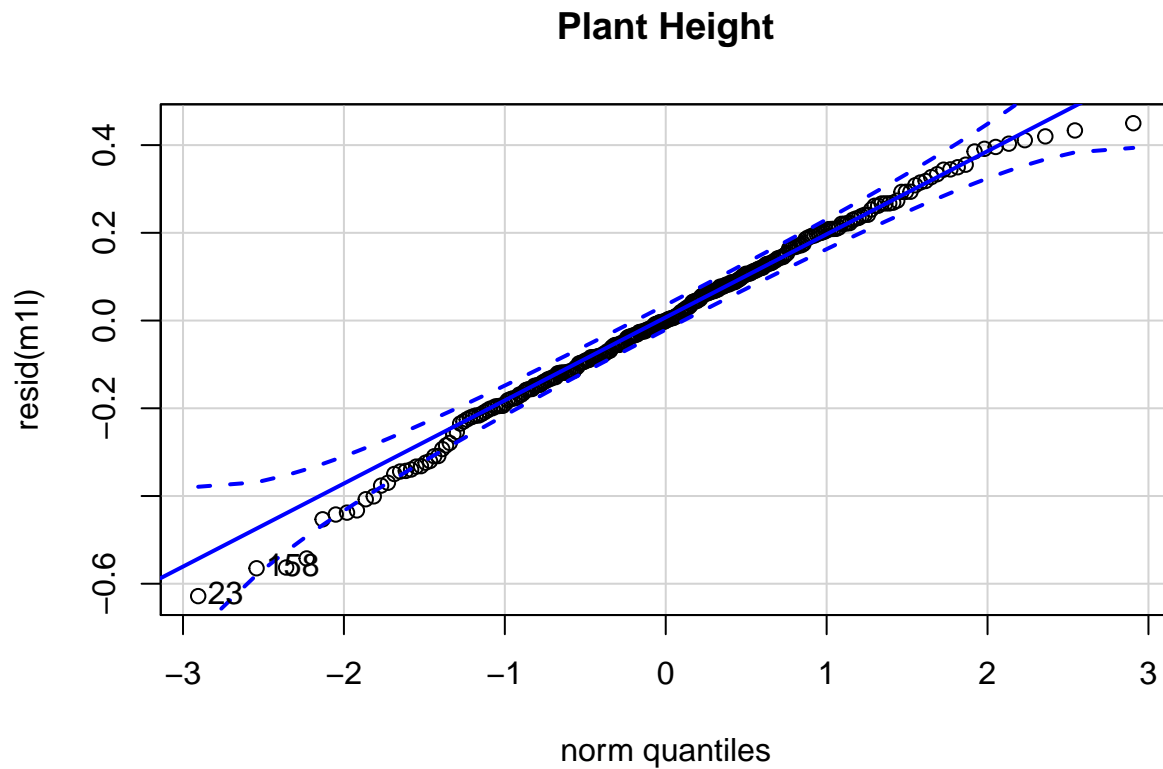
```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  4  1.1588 0.3294
##      268
```

```
# Assumption met
leveneTest(residuals(m1l) ~ galls3$gall_present)
```

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

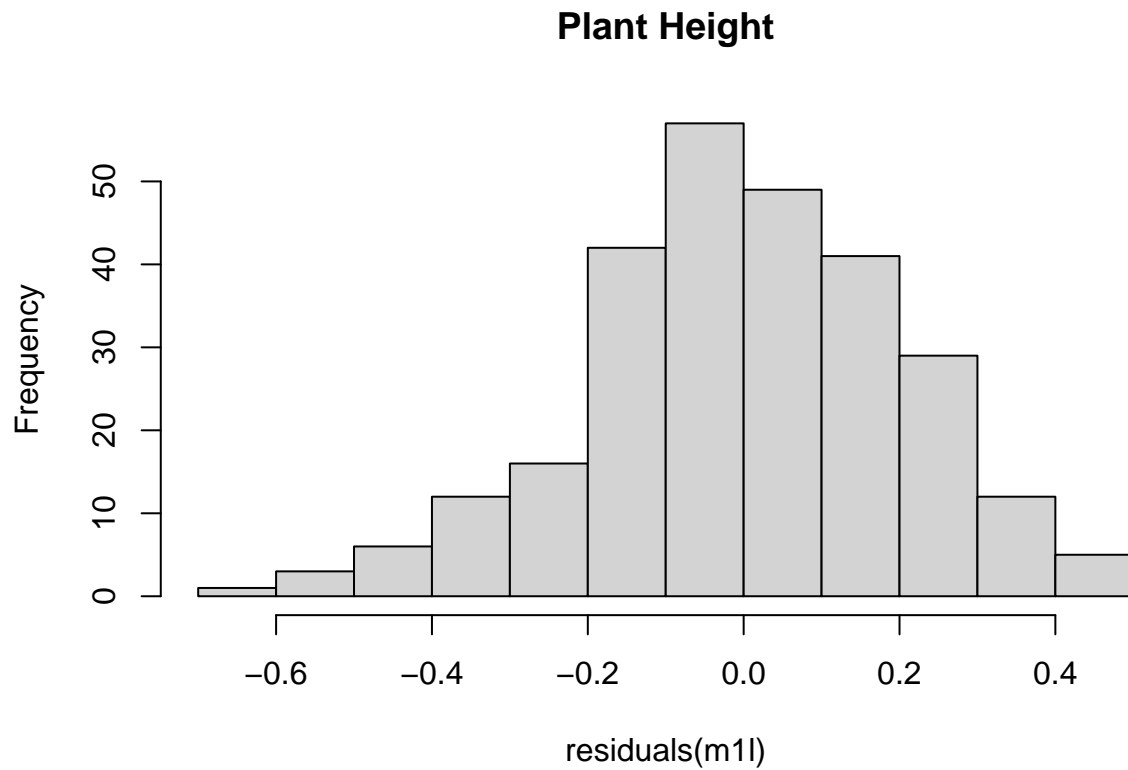
```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value  Pr(>F)
## group  1  6.883 0.009195 **
##      271
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Assumption not met
# (3) Normality of error term: need to check by histogram, QQplot of residuals, could do Kolmogorov-Smi
# Check for normal residuals
qqPlot(resid(m1l), main = "Plant Height")
```



```
## [1] 23 158
```

```
hist(residuals(m1l), main = "Plant Height")
```



```
shapiro.test(resid(m1l)) # Normal
```

```
##
## Shapiro-Wilk normality test
##
## data: resid(m1l)
## W = 0.99184, p-value = 0.1365
```

```
# Model comparisons with log plant height
m2l <- lm(lph ~ treatment, data=galls3)
m3l <- lm(lph ~ gall_present, data=galls3)
m4l <- lmer(lph ~ treatment + (1|rep), data=galls3, REML=F)
m5l <- lmer(lph ~ treatment * gall_present + (1|rep), data=galls3, REML=F)
AICctab(m1l, m2l, m3l, m4l, m5l, weights=T)
```

```
##      dAICc df weight
## m1l   0.0  8  0.629
## m5l   1.3 12  0.329
## m4l   5.4  7  0.042
## m2l  57.0  6 <0.001
## m3l 110.1  3 <0.001
```

```
# Model 1 fits the best but it makes sense that there is an interaction between treatment and gall pres
summ(m5l)
```

Observations	273
Dependent variable	lph
Type	Mixed effects linear regression

AIC	-64.43
BIC	-21.12
Pseudo-R <sup>2</sup> (fixed effects)	0.24
Pseudo-R <sup>2</sup> (total)	0.42

Fixed Effects					
	Est.	S.E.	t val.	d.f.	p
(Intercept)	4.07	0.06	67.52	16.53	0.00
treatmentdrought	0.04	0.05	0.66	267.10	0.51
treatmentirr_control	-0.06	0.06	-1.03	267.19	0.30
treatmentwarmed	0.23	0.06	3.98	267.10	0.00
treatmentwarmed_drought	0.31	0.06	5.55	267.09	0.00
gall_presentno_gall	0.03	0.05	0.62	267.10	0.54
treatmentdrought:gall_presentno_gall	-0.11	0.08	-1.44	267.06	0.15
treatmentirr_control:gall_presentno_gall	-0.10	0.08	-1.27	267.07	0.21
treatmentwarmed:gall_presentno_gall	-0.09	0.08	-1.16	267.06	0.25
treatmentwarmed_drought:gall_presentno_gall	-0.21	0.08	-2.72	267.04	0.01

p values calculated using Satterthwaite d.f.

Random Effects		
Group	Parameter	Std. Dev.
rep	(Intercept)	0.11
Residual		0.20

Grouping Variables		
Group	# groups	ICC
rep	6	0.23

```
summary(m5l)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: lph ~ treatment * gall_present + (1 | rep)
## Data: galls3
##
##      AIC      BIC   logLik deviance df.resid
##    -64.4    -21.1    44.2    -88.4      261
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.1310 -0.5953  0.0321  0.6944  2.1344
##
## Random effects:
```

```

## Groups Name Variance Std.Dev.
## rep (Intercept) 0.01222 0.1105
## Residual 0.03991 0.1998
## Number of obs: 273, groups: rep, 6
##
## Fixed effects:
##
## Estimate Std. Error df
## (Intercept) 4.07316 0.06032 16.53025
## treatmentdrought 0.03588 0.05458 267.10170
## treatmentirrr_control -0.05784 0.05609 267.19027
## treatmentwarmed 0.22772 0.05715 267.10144
## treatmentwarmed_drought 0.30833 0.05551 267.09315
## gall_presentno_gall 0.03340 0.05416 267.09664
## treatmentdrought:gall_presentno_gall -0.10835 0.07510 267.05854
## treatmentirrr_control:gall_presentno_gall -0.10003 0.07878 267.07002
## treatmentwarmed:gall_presentno_gall -0.08924 0.07699 267.06049
## treatmentwarmed_drought:gall_presentno_gall -0.20672 0.07606 267.04074
##
## t value Pr(>|t|)
## (Intercept) 67.523 < 2e-16 ***
## treatmentdrought 0.657 0.512
## treatmentirrr_control -1.031 0.303
## treatmentwarmed 3.984 8.73e-05 ***
## treatmentwarmed_drought 5.555 6.72e-08 ***
## gall_presentno_gall 0.617 0.538
## treatmentdrought:gall_presentno_gall -1.443 0.150
## treatmentirrr_control:gall_presentno_gall -1.270 0.205
## treatmentwarmed:gall_presentno_gall -1.159 0.247
## treatmentwarmed_drought:gall_presentno_gall -2.718 0.007 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
## (Intr) trtmntd trtmntr_ trtmntw trtmntw_ gll_p_ trtmntd:__
## trtmntdrght -0.487
## trtmntrr_cn -0.474 0.524
## tretmntwrmd -0.462 0.511 0.498
## trtmntwrmd_ -0.478 0.529 0.514 0.503
## gll_prsntn_ -0.491 0.542 0.528 0.515 0.533
## trtmntdr:__ 0.354 -0.727 -0.381 -0.371 -0.384 -0.721
## trtmntr:__ 0.336 -0.371 -0.708 -0.355 -0.365 -0.686 0.495
## trtmntwr:__ 0.343 -0.379 -0.370 -0.742 -0.374 -0.701 0.506
## trtmntw:__ 0.349 -0.386 -0.376 -0.367 -0.729 -0.712 0.513
## trtmntr:__ trtmntw:__
## trtmntdrght
## trtmntrr_cn
## tretmntwrmd
## trtmntwrmd_
## gll_prsntn_
## trtmntdr:__
## trtmntr:__
## trtmntwr:__ 0.483
## trtmntw:__ 0.488 0.500

```

```
# Post hoc test to compare different levels
emmeans(m5l, list(pairwise ~ treatment), adjust = "tukey")
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## $'emmeans of treatment'
## treatment      emmean      SE    df lower.CL upper.CL
## ambient        4.09 0.0568 11.6      3.97      4.21
## drought        4.07 0.0563 11.1      3.95      4.20
## irr_control     3.98 0.0577 12.4      3.86      4.11
## warmed         4.27 0.0570 11.7      4.15      4.40
## warmed_drought 4.29 0.0566 11.4      4.17      4.42
##
## Results are averaged over the levels of: gall_present
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
##
## $'pairwise differences of treatment'
## 1 estimate      SE    df t.ratio p.value
## ambient - drought      0.0183 0.0382 276   0.479 0.9892
## ambient - irr_control   0.1079 0.0403 277   2.674 0.0605
## ambient - warmed     -0.1831 0.0392 276  -4.676 <.0001
## ambient - warmed_drought -0.2050 0.0387 276  -5.295 <.0001
## drought - irr_control   0.0896 0.0395 277   2.265 0.1597
## drought - warmed     -0.2014 0.0385 276  -5.235 <.0001
## drought - warmed_drought -0.2233 0.0379 276  -5.884 <.0001
## irr_control - warmed   -0.2910 0.0406 277  -7.167 <.0001
## irr_control - warmed_drought -0.3128 0.0401 277  -7.801 <.0001
## warmed - warmed_drought -0.0219 0.0390 276  -0.561 0.9804
##
## Results are averaged over the levels of: gall_present
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 5 estimates
```

```
emmeans(m5l, list(pairwise ~ treatment * gall_present), adjust = "tukey")
```

```
## $'emmeans of treatment, gall_present'
## treatment      gall_present emmean      SE    df lower.CL upper.CL
## ambient        gall          4.07 0.0642 19.8      3.94      4.21
## drought        gall          4.11 0.0624 17.4      3.98      4.24
## irr_control     gall          4.02 0.0637 19.1      3.88      4.15
## warmed         gall          4.30 0.0649 20.6      4.17      4.44
## warmed_drought gall          4.38 0.0633 18.5      4.25      4.51
## ambient        no_gall       4.11 0.0620 17.0      3.98      4.24
## drought        no_gall       4.03 0.0620 17.0      3.90      4.16
## irr_control     no_gall       3.95 0.0656 21.5      3.81      4.08
## warmed         no_gall       4.25 0.0620 17.0      4.11      4.38
## warmed_drought no_gall       4.21 0.0624 17.4      4.08      4.34
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
##
```

```

## $'pairwise differences of treatment, gall_present'
## 1 estimate SE df t.ratio
## ambient gall - drought gall -0.03588 0.0555 276 -0.646
## ambient gall - irr_control gall 0.05784 0.0571 276 1.014
## ambient gall - warmed gall -0.22772 0.0581 276 -3.916
## ambient gall - warmed_drought gall -0.30833 0.0565 276 -5.460
## ambient gall - ambient no_gall -0.03340 0.0551 276 -0.606
## ambient gall - drought no_gall 0.03907 0.0551 276 0.709
## ambient gall - irr_control no_gall 0.12448 0.0593 277 2.100
## ambient gall - warmed no_gall -0.17187 0.0551 276 -3.119
## ambient gall - warmed_drought no_gall -0.13500 0.0555 276 -2.431
## drought gall - irr_control gall 0.09372 0.0550 276 1.705
## drought gall - warmed gall -0.19185 0.0563 277 -3.409
## drought gall - warmed_drought gall -0.27245 0.0544 276 -5.011
## drought gall - ambient no_gall 0.00248 0.0529 276 0.047
## drought gall - drought no_gall 0.07495 0.0529 276 1.416
## drought gall - irr_control no_gall 0.16035 0.0571 277 2.808
## drought gall - warmed no_gall -0.13600 0.0529 276 -2.569
## drought gall - warmed_drought no_gall -0.09913 0.0534 276 -1.857
## irr_control gall - warmed gall -0.28557 0.0577 277 -4.945
## irr_control gall - warmed_drought gall -0.36617 0.0559 276 -6.545
## irr_control gall - ambient no_gall -0.09124 0.0545 276 -1.674
## irr_control gall - drought no_gall -0.01877 0.0545 276 -0.344
## irr_control gall - irr_control no_gall 0.06663 0.0583 277 1.142
## irr_control gall - warmed no_gall -0.22972 0.0545 276 -4.215
## irr_control gall - warmed_drought no_gall -0.19285 0.0550 276 -3.509
## warmed gall - warmed_drought gall -0.08060 0.0572 276 -1.410
## warmed gall - ambient no_gall 0.19433 0.0558 277 3.480
## warmed gall - drought no_gall 0.26680 0.0558 277 4.778
## warmed gall - irr_control no_gall 0.35220 0.0600 277 5.870
## warmed gall - warmed no_gall 0.05585 0.0558 277 1.000
## warmed gall - warmed_drought no_gall 0.09272 0.0563 277 1.648
## warmed_drought gall - ambient no_gall 0.27493 0.0540 276 5.096
## warmed_drought gall - drought no_gall 0.34740 0.0540 276 6.439
## warmed_drought gall - irr_control no_gall 0.43280 0.0581 277 7.446
## warmed_drought gall - warmed no_gall 0.13645 0.0540 276 2.529
## warmed_drought gall - warmed_drought no_gall 0.17332 0.0544 276 3.188
## ambient no_gall - drought no_gall 0.07247 0.0525 276 1.381
## ambient no_gall - irr_control no_gall 0.15787 0.0567 277 2.786
## ambient no_gall - warmed no_gall -0.13848 0.0525 276 -2.639
## ambient no_gall - warmed_drought no_gall -0.10161 0.0529 276 -1.920
## drought no_gall - irr_control no_gall 0.08540 0.0567 277 1.507
## drought no_gall - warmed no_gall -0.21095 0.0525 276 -4.020
## drought no_gall - warmed_drought no_gall -0.17408 0.0529 276 -3.289
## irr_control no_gall - warmed no_gall -0.29635 0.0567 277 -5.230
## irr_control no_gall - warmed_drought no_gall -0.25948 0.0571 277 -4.543
## warmed no_gall - warmed_drought no_gall 0.03687 0.0529 276 0.696
## p.value
## 0.9997
## 0.9913
## 0.0044
## <.0001
## 0.9998
## 0.9994

```



```
## 0.5289
## 0.0612
## 0.3116
## 0.7917
## 0.0255
## <.0001
## 1.0000
## 0.9215
## 0.1390
## 0.2374
## 0.6976
## 0.0001
## <.0001
## 0.8090
## 1.0000
## 0.9798
## 0.0014
## 0.0185
## 0.9233
## 0.0203
## 0.0001
## <.0001
## 0.9921
## 0.8232
## <.0001
## <.0001
## <.0001
## 0.2577
## 0.0503
## 0.9321
## 0.1463
## 0.2047
## 0.6557
## 0.8884
## 0.0030
## 0.0372
## <.0001
## 0.0003
## 0.9995
##
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 10 estimates
```

```
summ(m11)
```

Observations	273
Dependent variable	lph
Type	Mixed effects linear regression

```
# outlier <- galls3 %>%
#   identify_outliers(plant_height)
# outlier
# galls3_no_outliers <- galls[-c(53, 54, 103, 110), ]
```

AIC	-65.07
BIC	-36.20
Pseudo-R <sup>2</sup> (fixed effects)	0.22
Pseudo-R <sup>2</sup> (total)	0.40

Fixed Effects					
	Est.	S.E.	t val.	d.f.	p
(Intercept)	4.13	0.05	75.45	10.96	0.00
treatmentdrought	-0.02	0.04	-0.61	267.06	0.54
treatmentirr_control	-0.11	0.04	-2.81	267.56	0.01
treatmentwarmed	0.18	0.04	4.62	267.05	0.00
treatmentwarmed_drought	0.20	0.04	5.15	267.08	0.00
gall_presentno_gall	-0.07	0.02	-2.77	267.32	0.01

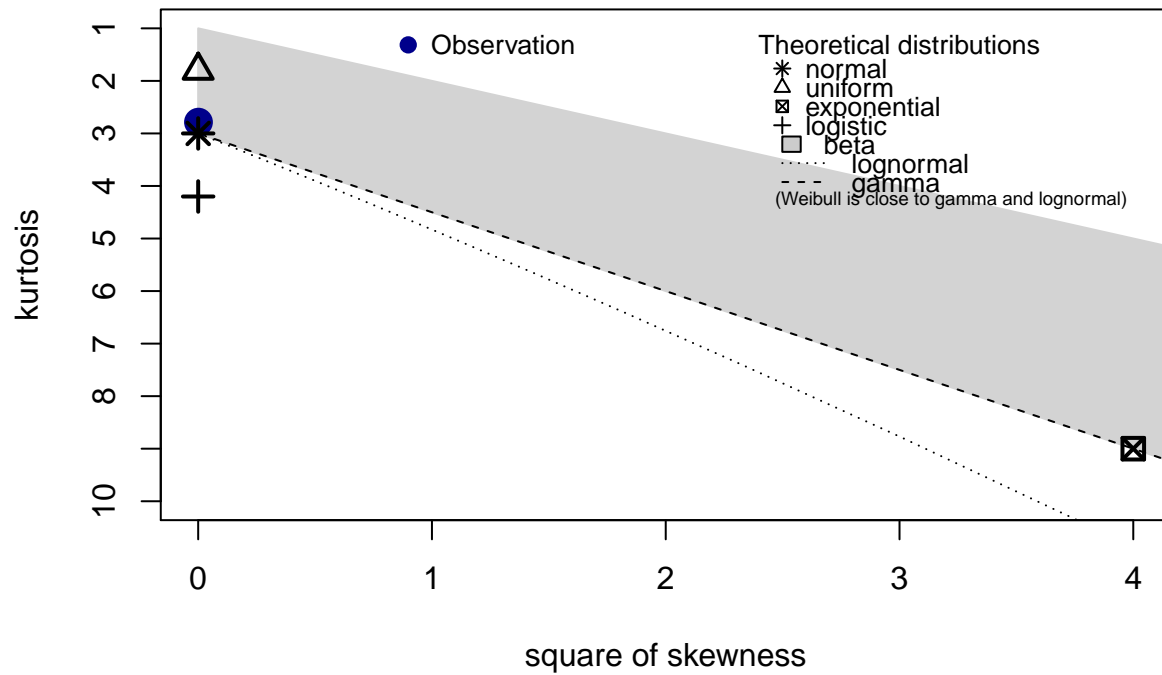
p values calculated using Satterthwaite d.f.

Random Effects		
Group	Parameter	Std. Dev.
rep	(Intercept)	0.11
Residual		0.20

Grouping Variables		
Group	# groups	ICC
rep	6	0.23

```
# taking the lognormal of plant height in the gall only dataframe
gall_only$lph <- log(gall_only$plant_height)
descdist(gall_only$lph, discrete = FALSE)
```

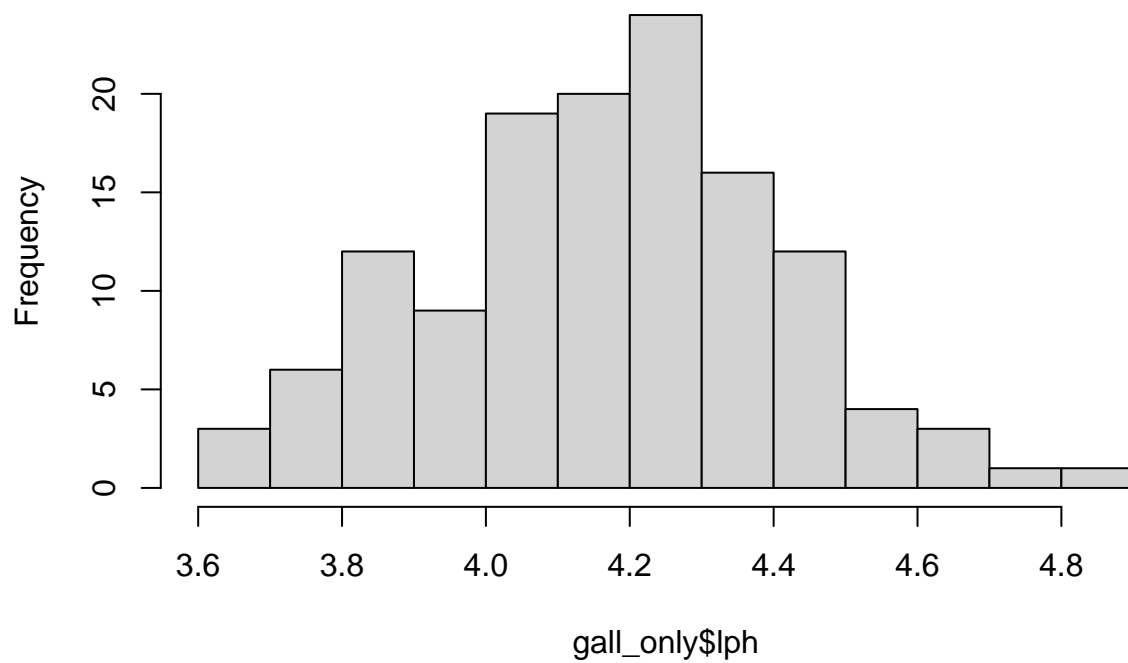
## Cullen and Frey graph



```
## summary statistics
## -----
## min: 3.618993   max: 4.83866
## median: 4.152612
## mean: 4.167847
## estimated sd: 0.2390164
## estimated skewness: 0.0357521
## estimated kurtosis: 2.785402
```

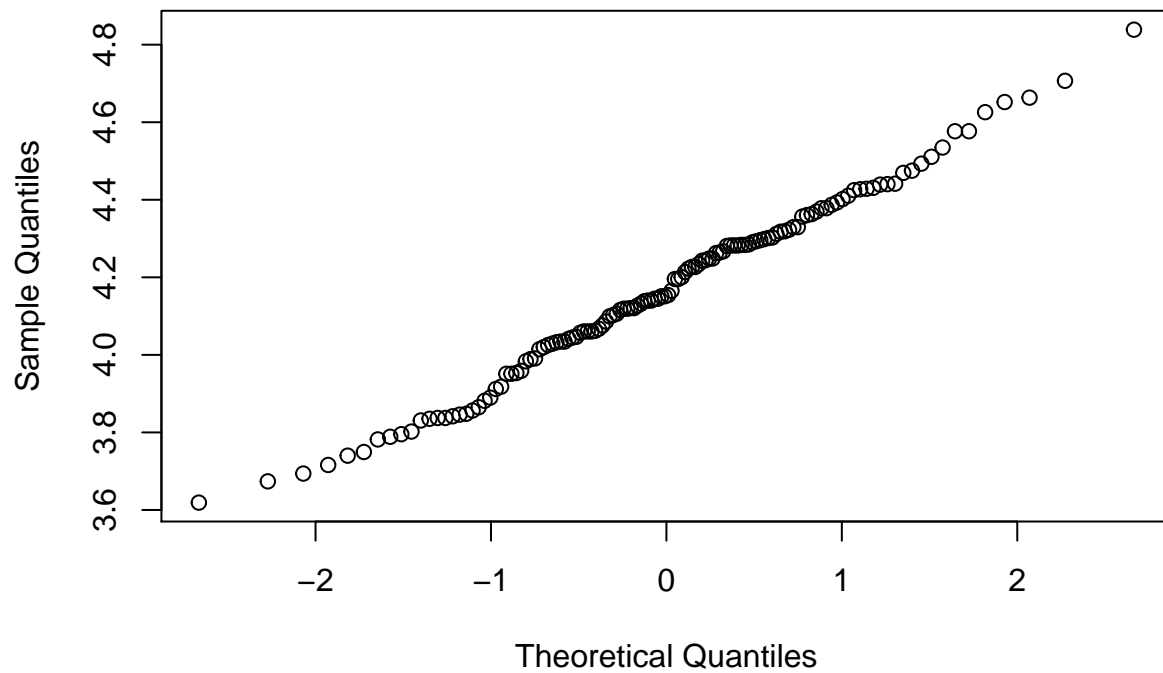
```
hist(gall_only$lph)
```

**Histogram of gall\_only\$lph**



```
qqnorm(gall_only$lph)
```

**Normal Q-Q Plot**



```
shapiro.test(gall_only$lph) # normally distributed!

##
## Shapiro-Wilk normality test
##
## data:  gall_only$lph
## W = 0.99254, p-value = 0.7238

# Looking at plant height as a response variable and treatment as an explanatory variable.
w0 <- lm(lph ~ 1, data=gall_only)
w1 <- lm(lph ~ treatment, data=gall_only)
w2 <- lm(lph ~ gall_diameter, data=gall_only)
w3 <- lm(lph ~ gall_height, data=gall_only)
w4 <- lm(lph ~ treatment + (1|rep), data=gall_only)
w5 <- lm(lph ~ treatment + gall_diameter, data=gall_only)
w6 <- lm(lph ~ treatment + gall_height, data=gall_only)
w7 <- lm(lph ~ treatment * gall_diameter, data=gall_only)
w8 <- lm(lph ~ treatment * gall_height, data=gall_only)
w9 <- lm(lph ~ treatment + gall_diameter + (1|rep), data=gall_only)
w10 <- lm(lph ~ treatment + gall_height + (1|rep), data=gall_only)
w11 <- lm(lph ~ treatment * gall_diameter + (1|rep), data=gall_only)
w12 <- lm(lph ~ treatment * gall_height + (1|rep), data=gall_only)

AICctab(w0,w1,w2,w3,w4,w5,w6,w7,w8,w9,w10,w11,w12, weights=TRUE)
```

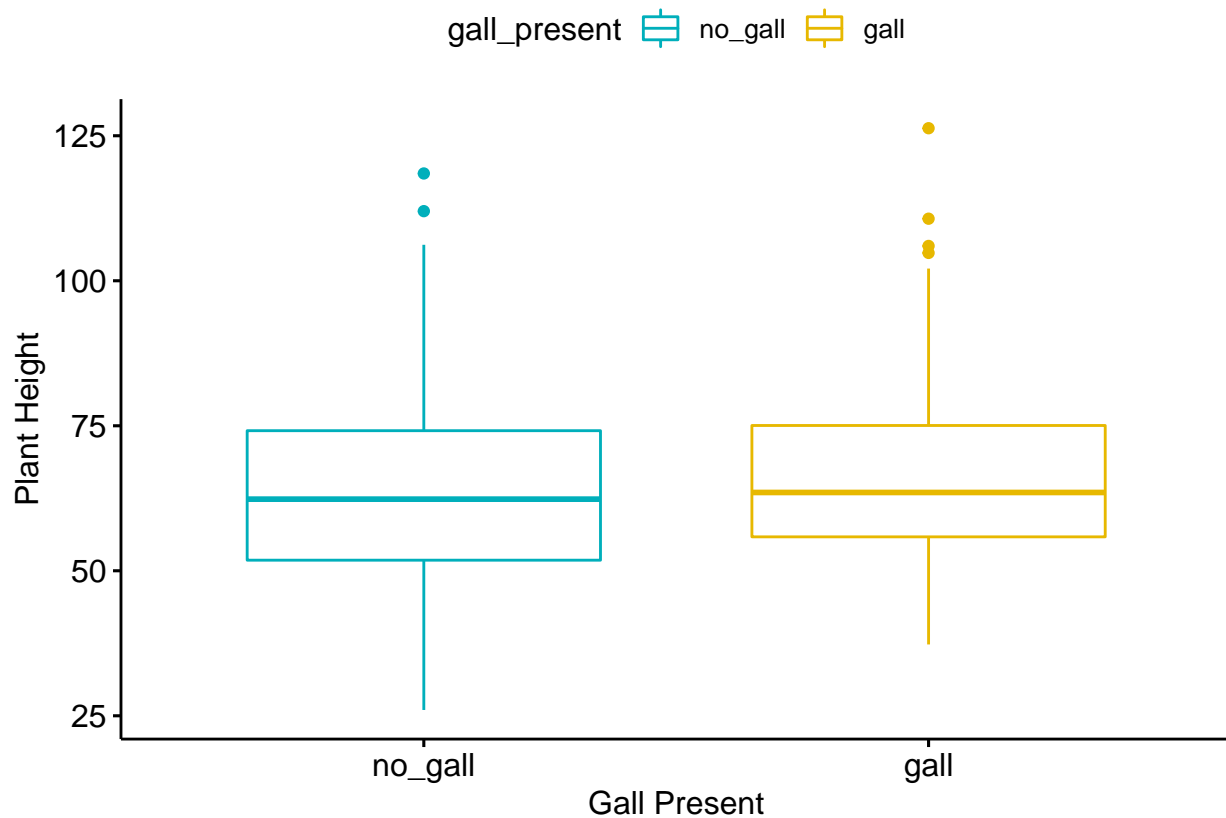
```
##      dAICc df weight
## w5    0.0  7  0.2281
## w9    0.0  7  0.2281
## w1    0.4  6  0.1838
## w4    0.4  6  0.1838
## w6    2.7  7  0.0602
## w10   2.7  7  0.0602
## w7    4.7 11  0.0216
## w11   4.7 11  0.0216
## w8    7.2 11  0.0063
## w12   7.2 11  0.0063
## w2   42.3  3 <0.001
## w0   46.4  2 <0.001
## w3   48.1  3 <0.001
```

Simple t-test of plant height between galled and non galled plants

```
group_by(galls3, gall_present) %>%
  summarise(count = n(),
            mean = mean(plant_height, na.rm = TRUE),
            sd = sd(plant_height), na.rm = TRUE)
```

```
## # A tibble: 2 x 5
##   gall_present count  mean    sd na.rm
## * <chr>      <int> <dbl> <dbl> <lgl>
## 1 gall           131  66.3  16.1 TRUE
## 2 no_gall        142  63.6  17.2 TRUE
```

```
ggboxplot(galls3, x = "gall_present", y = "plant_height",
          color = "gall_present", palette = c("#00AFBB", "#E7B800"),
          ylab = "Plant Height", xlab = "Gall Present")
```



```
# Shapiro-Wilk normality test for gall plant height
with(galls3, shapiro.test(plant_height[gall_present == "gall"])) # p = 0.002157 - not normally distributed
```

```
##
## Shapiro-Wilk normality test
##
## data: plant_height[gall_present == "gall"]
## W = 0.96571, p-value = 0.002157
```

```
# Shapiro-Wilk normality test for non galled plant height
with(galls3, shapiro.test(plant_height[gall_present == "no_gall"])) # 0.08565 - normally distributed
```

```
##
## Shapiro-Wilk normality test
##
## data: plant_height[gall_present == "no_gall"]
## W = 0.9835, p-value = 0.08565
```

```
# Because gall plant height does not pass the Shapiro-Wilk normality test, doing a non parametric two-s
# Wilcoxon rank test
res <- wilcox.test(plant_height ~ gall_present, data = galls3,
```

```

                                exact = FALSE)
res # p-value = 0.1872

```

```

##
## Wilcoxon rank sum test with continuity correction
##
## data: plant_height by gall_present
## W = 10161, p-value = 0.1872
## alternative hypothesis: true location shift is not equal to 0

```

*# p-value = 0.1872, which is greater than the significance level  $\alpha = 0.05$ , therefore we can conclude that there is no significant difference in plant height between galls and non-galls.*

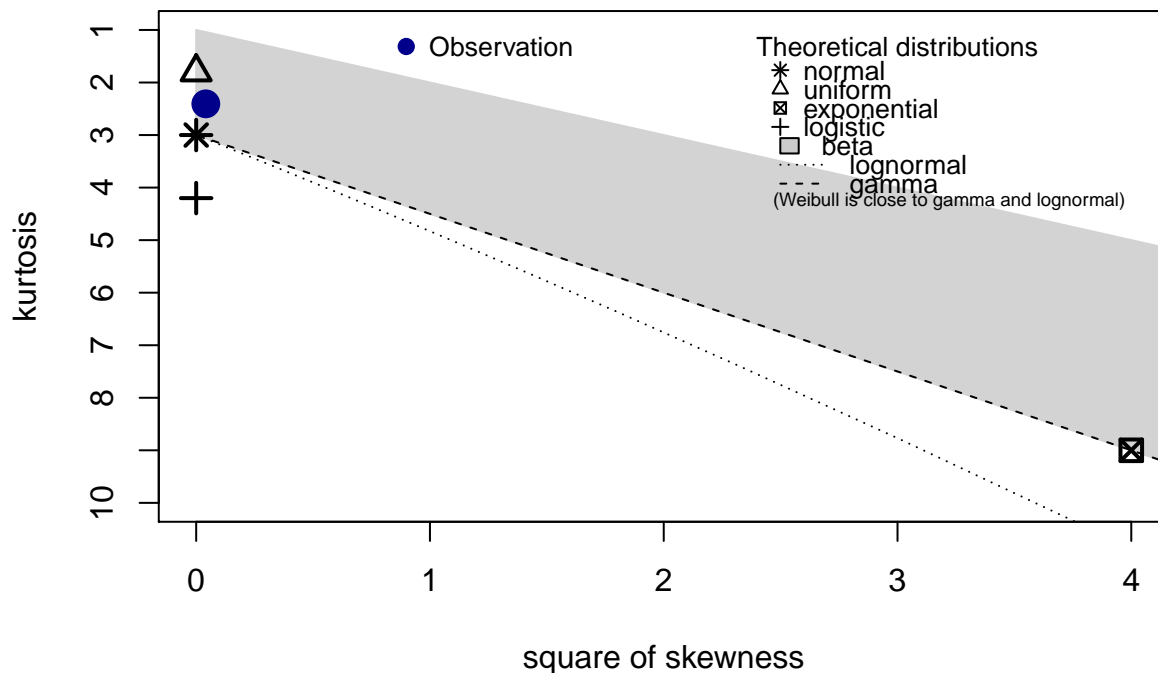
Gall diameter

```

# Data exploration
descdist(gall_only$gall_diameter, discrete = FALSE)

```

## Cullen and Frey graph



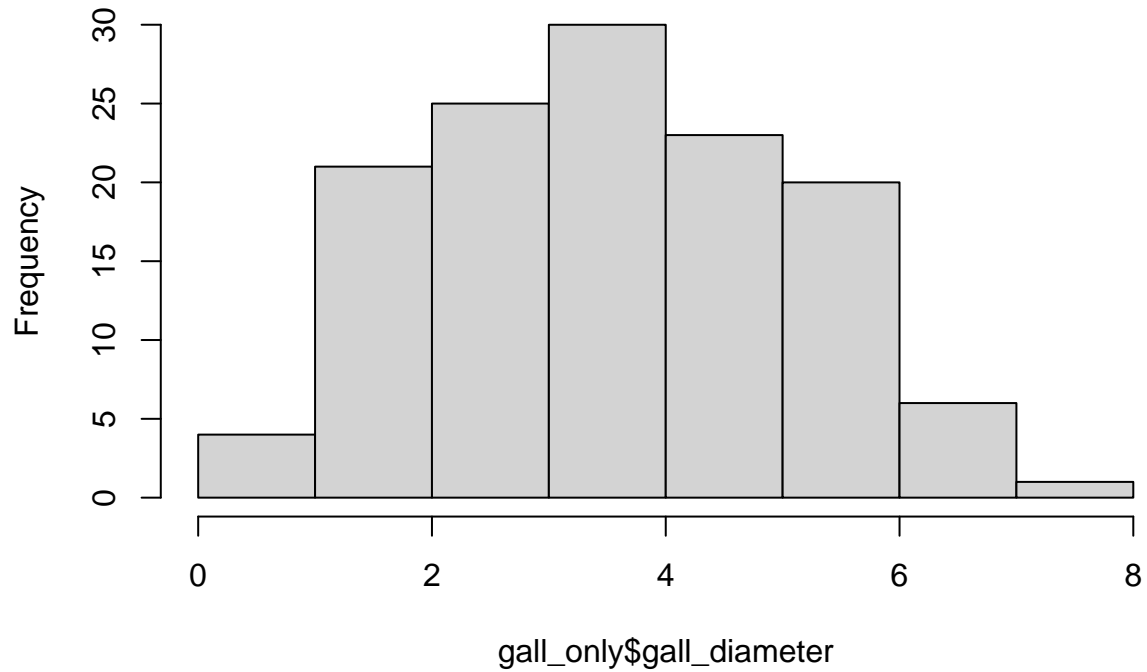
```

## summary statistics
## -----
## min: 0.9 max: 7.7
## median: 3.6
## mean: 3.633846
## estimated sd: 1.497572
## estimated skewness: 0.2029608
## estimated kurtosis: 2.406142

```

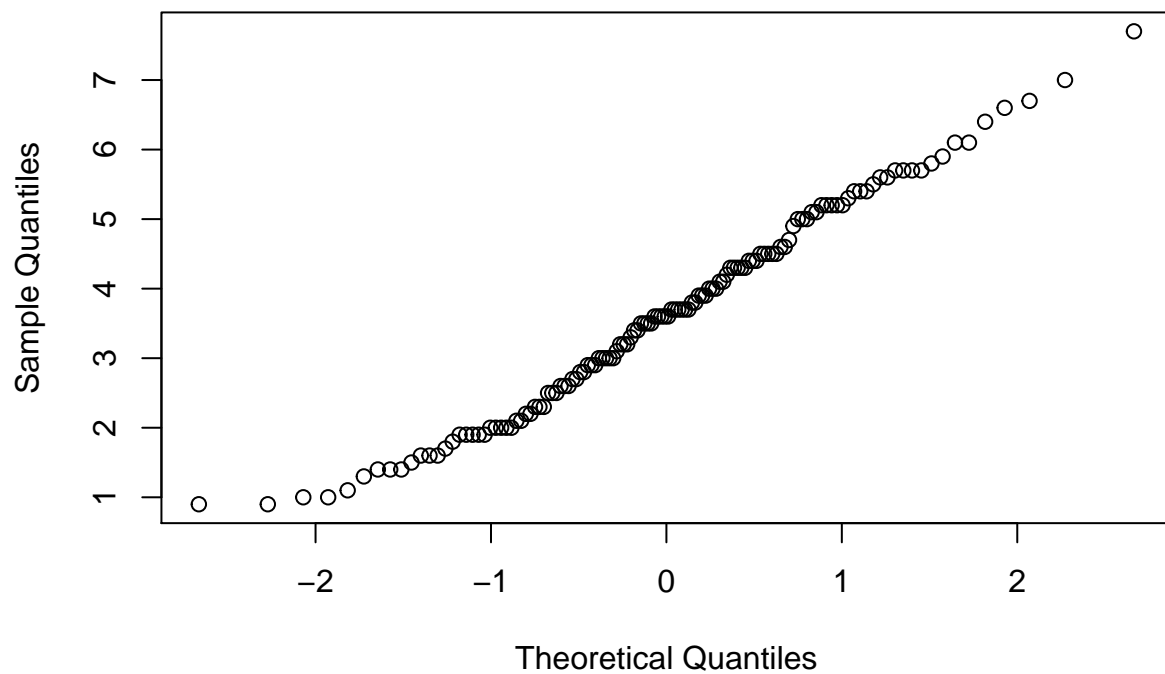
```
hist(gall_only$gall_diameter)
```

**Histogram of gall\_only\$gall\_diameter**



```
qqnorm(gall_only$gall_diameter)
```

**Normal Q-Q Plot**





```
shapiro.test(gall_only$gall_diameter)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  gall_only$gall_diameter  
## W = 0.98304, p-value = 0.1048
```

```
# Normally distributed
```

```
# Assumption checking
```

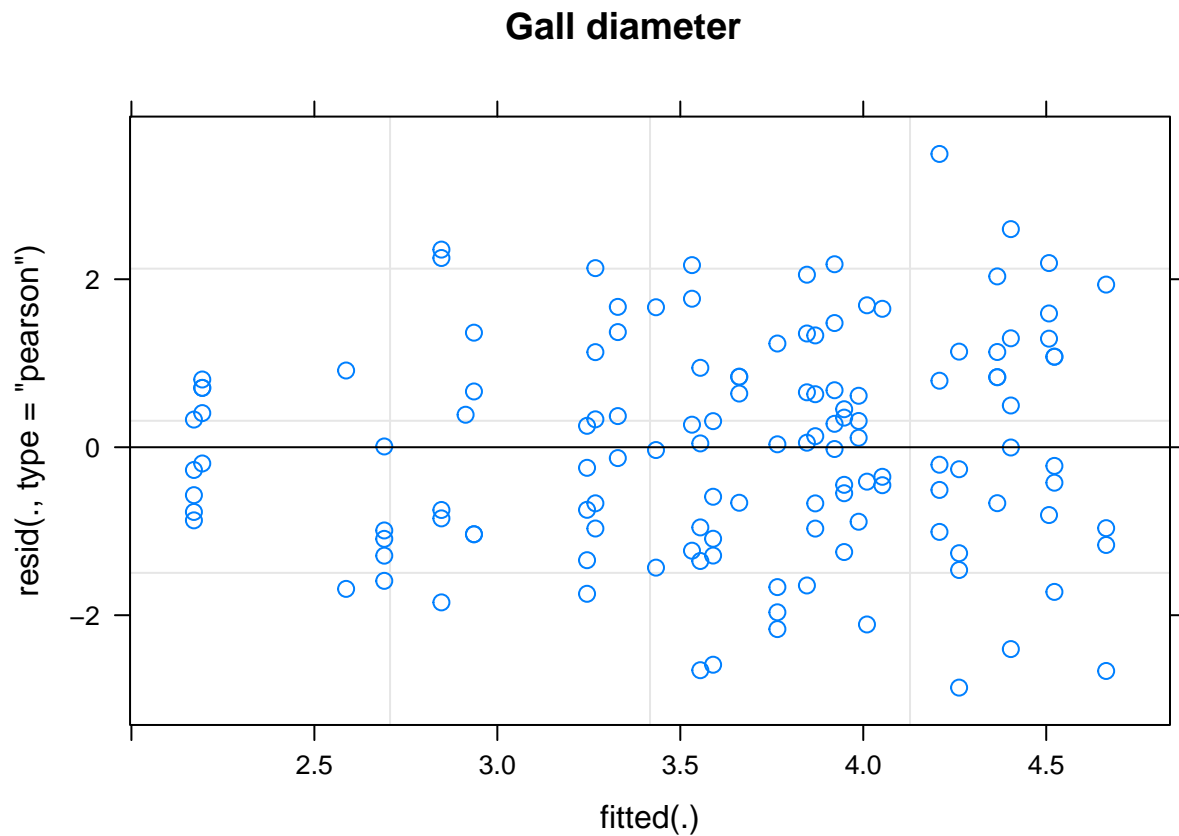
```
g.m1 <- lmer(gall_diameter ~ treatment + (1|rep), data = gall_only, REML=FALSE)
```

```
# Check Assumptions:
```

```
# (1) Linearity: if covariates are not categorical
```

```
# (2) Homogeneity: Need to Check by plotting residuals vs predicted values.
```

```
plot(g.m1, main = "Gall diameter")
```



```
# Homogeneity of variance is ok here (increasing variance in resids is not increasing with fitted value.  
# Check for homogeneity of variances (true if p>0.05). If the result is not significant, the assumption  
leveneTest(residuals(g.m1) ~ gall_only$treatment)
```

```
## 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  4  0.9338 0.4467
##      125
```

*# Assumption met*

```
leveneTest(residuals(g.m1) ~ as.factor(gall_only$rep))
```

```
## Levene's Test for Homogeneity of Variance (center = median)
```

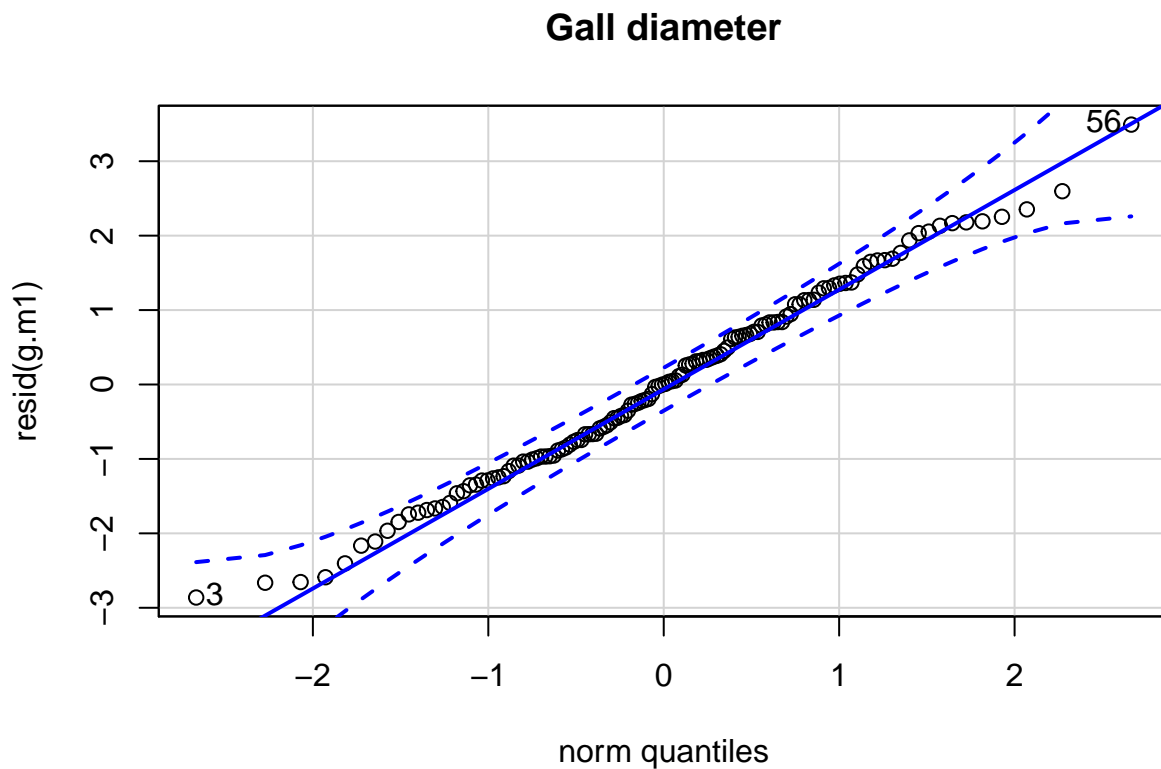
```
##      Df F value Pr(>F)
## group  5  0.4682 0.7994
##      124
```

*# Assumption met*

*# (3) Normality of error term: need to check by histogram, QQplot of residuals, could do Kolmogorov-Smirnov*

*# Check for normal residuals*

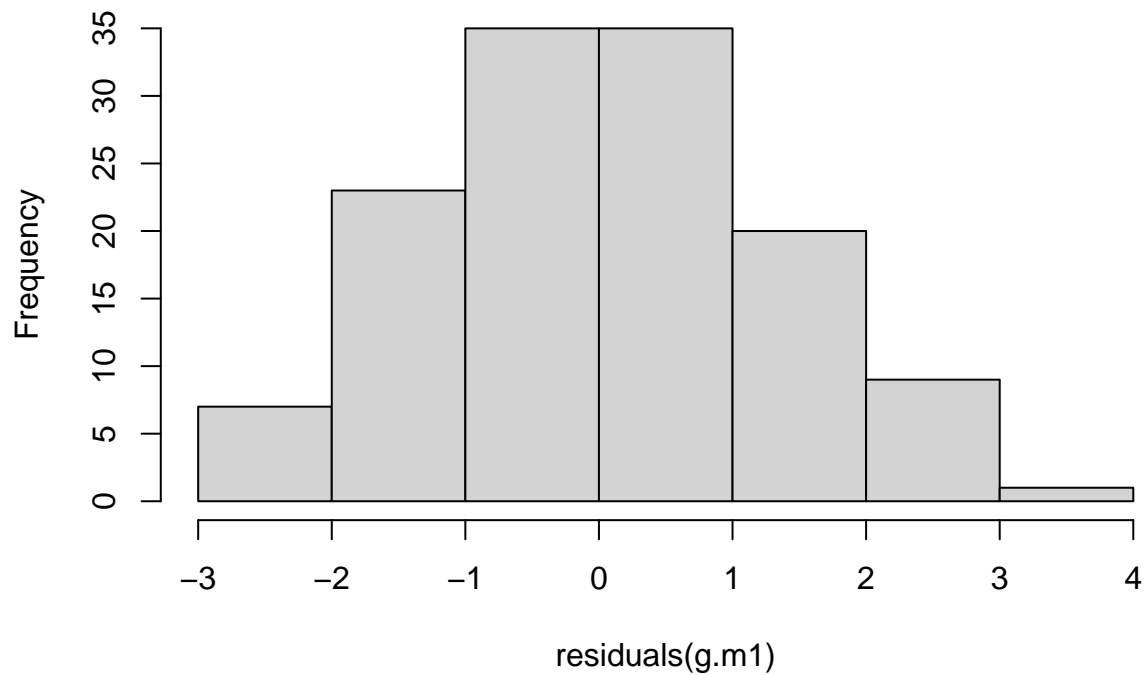
```
qqPlot(resid(g.m1), main = "Gall diameter")
```



```
## [1] 56 3
```

```
hist(residuals(g.m1), main = "Gall diameter")
```

## Gall diameter



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

```
##
## Shapiro-Wilk normality test
##
## data: resid(g.m1)
## W = 0.99217, p-value = 0.6857
```

```
# Model comparisons
g.m2 <- lm(gall_diameter ~ treatment, data=gall_only)
AICctab(g.m1, g.m2, weights=T)
```

```
##      dAICc df weight
## g.m1  0.0  7  1
## g.m2 16.1  6 <0.001
```

```
# Model 1 fits the best
summary(g.m1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: gall_diameter ~ treatment + (1 | rep)
## Data: gall_only
##
##      AIC      BIC    logLik deviance df.resid
##    463.7    483.8   -224.9    449.7     123
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.19051 -0.74005  0.00218  0.64212  2.67255
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   rep      (Intercept) 0.4427   0.6653
##   Residual              1.7070   1.3065
## Number of obs: 130, groups: rep, 6
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)      3.8026     0.3776   15.7128  10.070 2.96e-08 ***
## treatmentdrought      0.1562     0.3574  124.2091   0.437   0.663
## treatmentirr_control  -0.1043     0.3719  124.5496  -0.280   0.780
## treatmentwarmed      -0.5202     0.3743  124.2345  -1.390   0.167
## treatmentwarmed_drought -0.4974     0.3634  124.1837  -1.369   0.174
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) trtmntd trtmntr_ trtmntw
## trtmntdrght -0.510
## trtmntrr_cn -0.492  0.519
## tretmntwrmd -0.482  0.509  0.490
## trtmntwrmd_ -0.500  0.530  0.509  0.502
```

```
# Post hoc test to compare different levels
emmeans(g.ml, list(pairwise ~ treatment), adjust = "tukey")
```

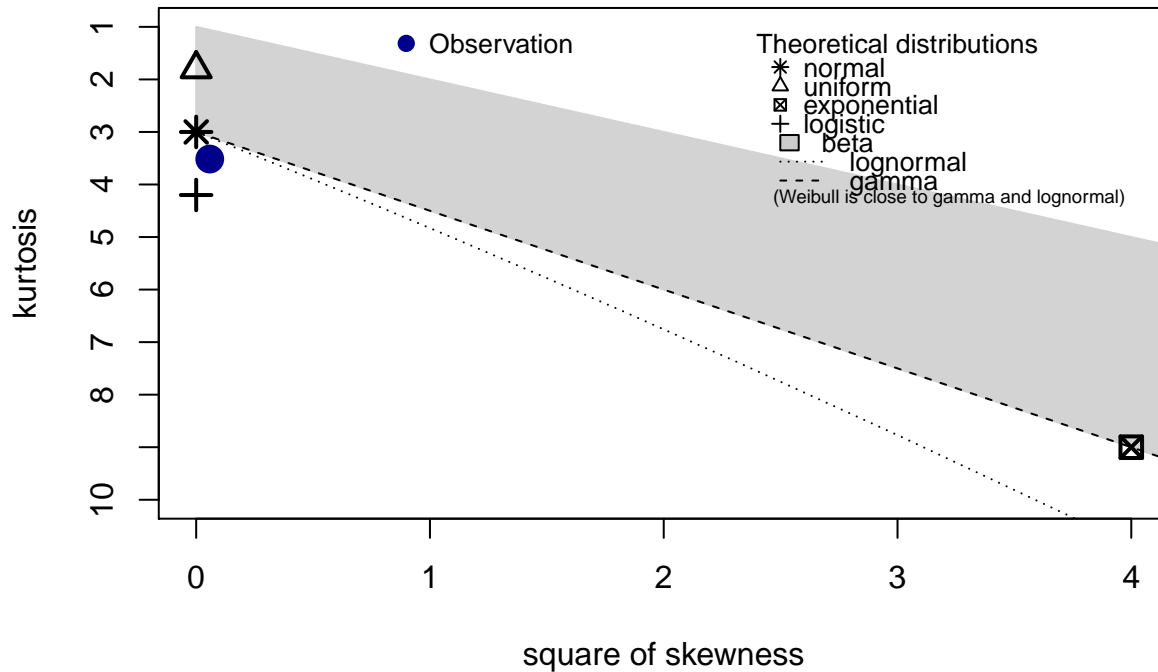
```
## $'emmeans of treatment'
##   treatment      emmean      SE    df lower.CL upper.CL
##   ambient      3.80 0.403 18.9      2.96      4.65
##   drought      3.96 0.390 16.5      3.13      4.78
##   irr_control   3.70 0.403 18.9      2.85      4.54
##   warmed       3.28 0.408 19.8      2.43      4.13
##   warmed_drought 3.31 0.396 17.6      2.47      4.14
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
##
## $'pairwise differences of treatment'
##   1              estimate      SE    df t.ratio p.value
##   ambient - drought      -0.1562 0.363 128  -0.430  0.9928
##   ambient - irr_control    0.1043 0.378 129   0.276  0.9987
##   ambient - warmed        0.5202 0.381 128   1.367  0.6499
##   ambient - warmed_drought 0.4974 0.370 128   1.346  0.6630
##   drought - irr_control    0.2604 0.364 128   0.715  0.9526
##   drought - warmed        0.6764 0.369 129   1.831  0.3603
##   drought - warmed_drought 0.6536 0.355 128   1.839  0.3560
##   irr_control - warmed     0.4159 0.384 129   1.084  0.8144
##   irr_control - warmed_drought 0.3931 0.371 129   1.060  0.8265
##   warmed - warmed_drought -0.0228 0.375 129  -0.061  1.0000
##
```

```
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 5 estimates
```

Gall height

```
# Data exploration
descdist(gall_only$gall_height, discrete = FALSE)
```

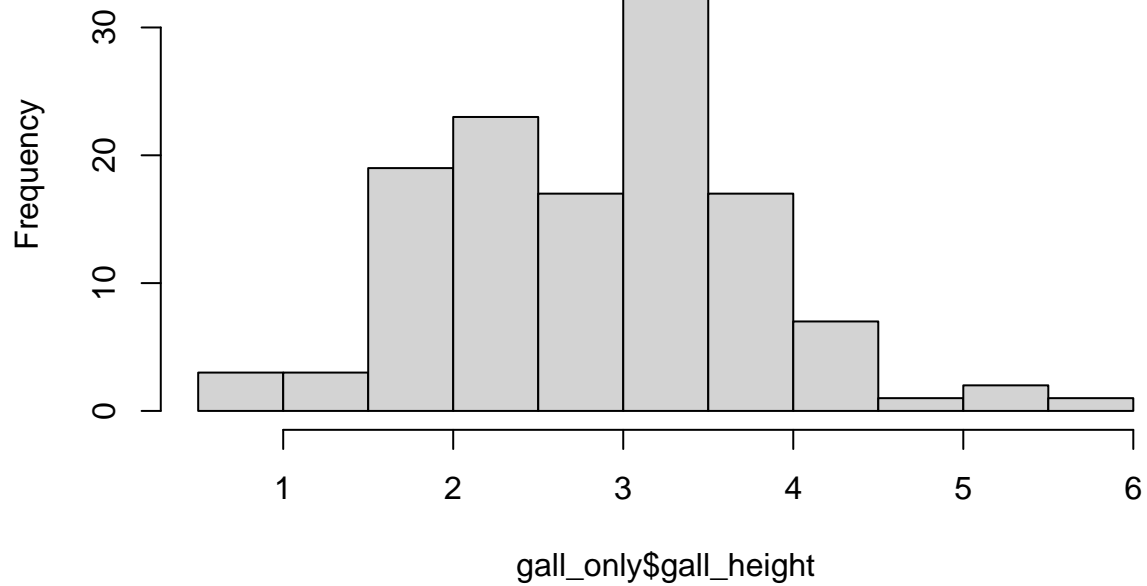
## Cullen and Frey graph



```
## summary statistics
## -----
## min: 0.5   max: 5.9
## median: 3.05
## mean: 2.906154
## estimated sd: 0.9137837
## estimated skewness: 0.2409495
## estimated kurtosis: 3.5158
```

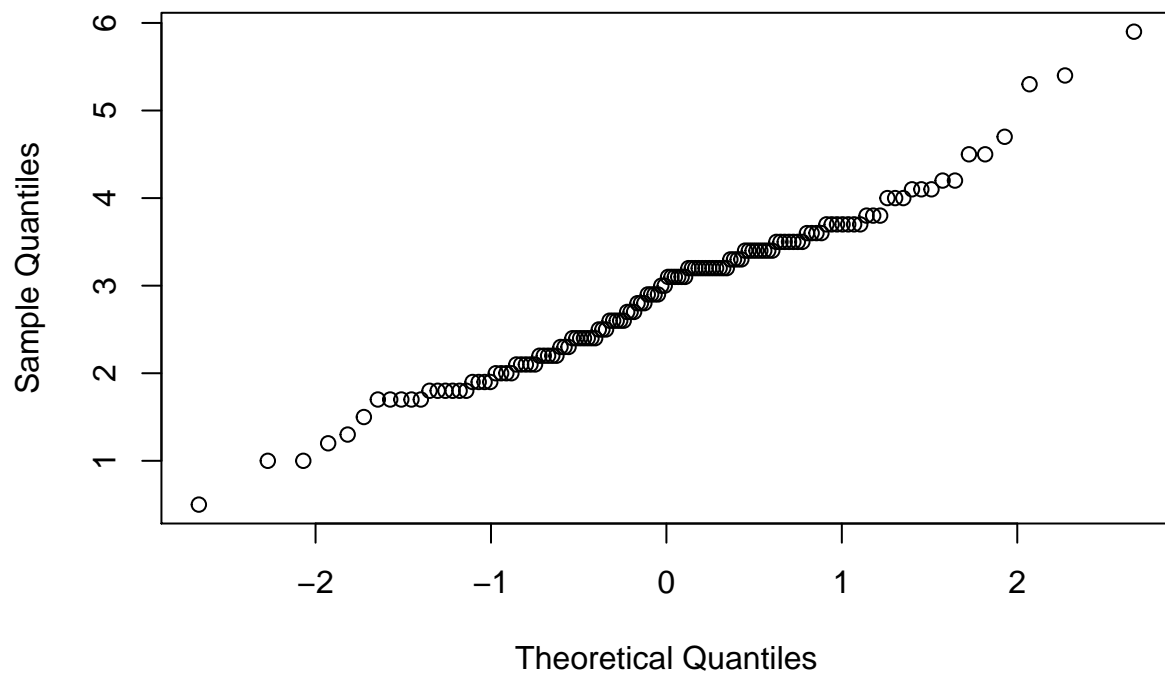
```
hist(gall_only$gall_height)
```

**Histogram of gall\_only\$gall\_height**



```
qqnorm(gall_only$gall_height)
```

**Normal Q-Q Plot**



```
shapiro.test(gall_only$gall_height)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  gall_only$gall_height
## W = 0.98255, p-value = 0.09358
```

```
# Normally distributed
```

```
# Assumption checking
```

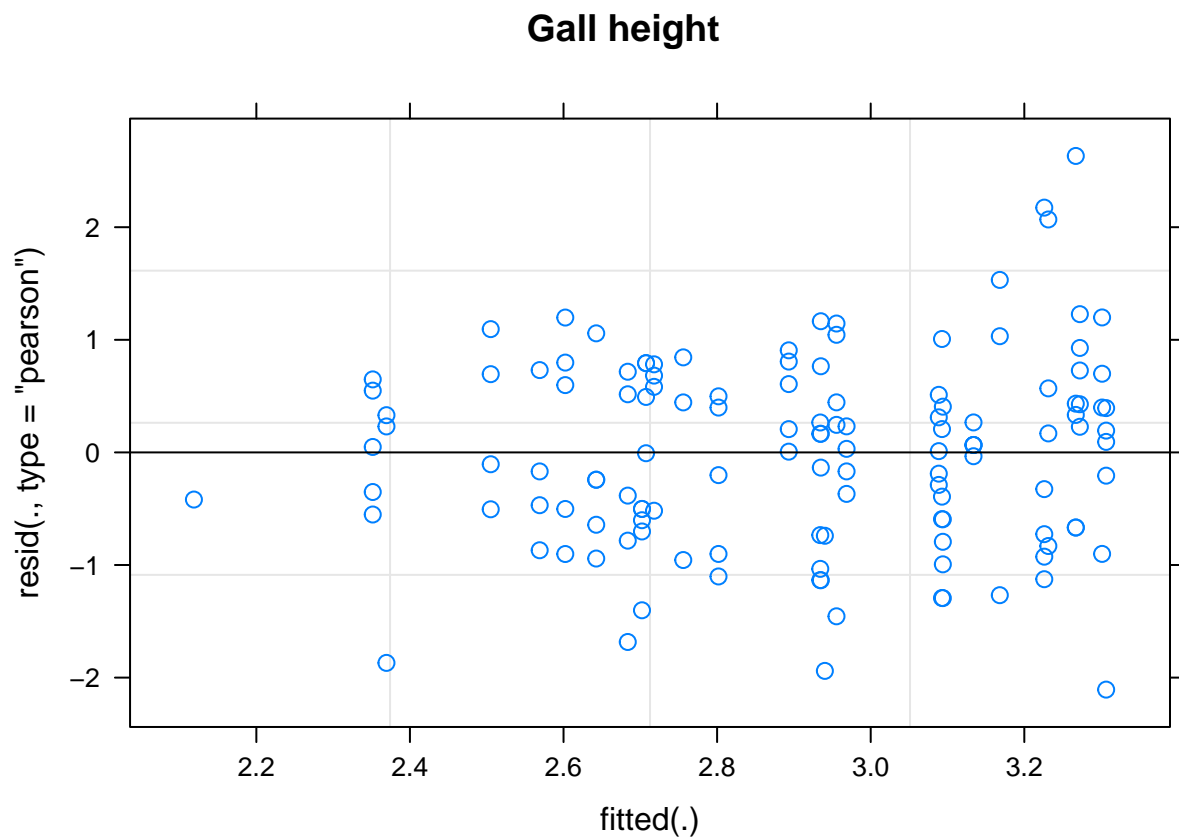
```
h.m1 <- lmer(gall_height ~ treatment + (1|rep), data = gall_only, REML=FALSE)
```

```
# Check Assumptions:
```

```
# (1) Linearity: if covariates are not categorical
```

```
# (2) Homogeneity: Need to Check by plotting residuals vs predicted values.
```

```
plot(h.m1, main = "Gall height")
```



```
# Homogeneity of variance is ok here (increasing variance in resids is not increasing with fitted value)
```

```
# Check for homogeneity of variances (true if p>0.05). If the result is not significant, the assumption
```

```
leveneTest(residuals(h.m1) ~ gall_only$treatment)
```

```
## 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  4  0.3904 0.8152
##      125
```

*# Assumption met*

```
leveneTest(residuals(h.m1) ~ as.factor(gall_only$rep))
```

```
## Levene's Test for Homogeneity of Variance (center = median)
```

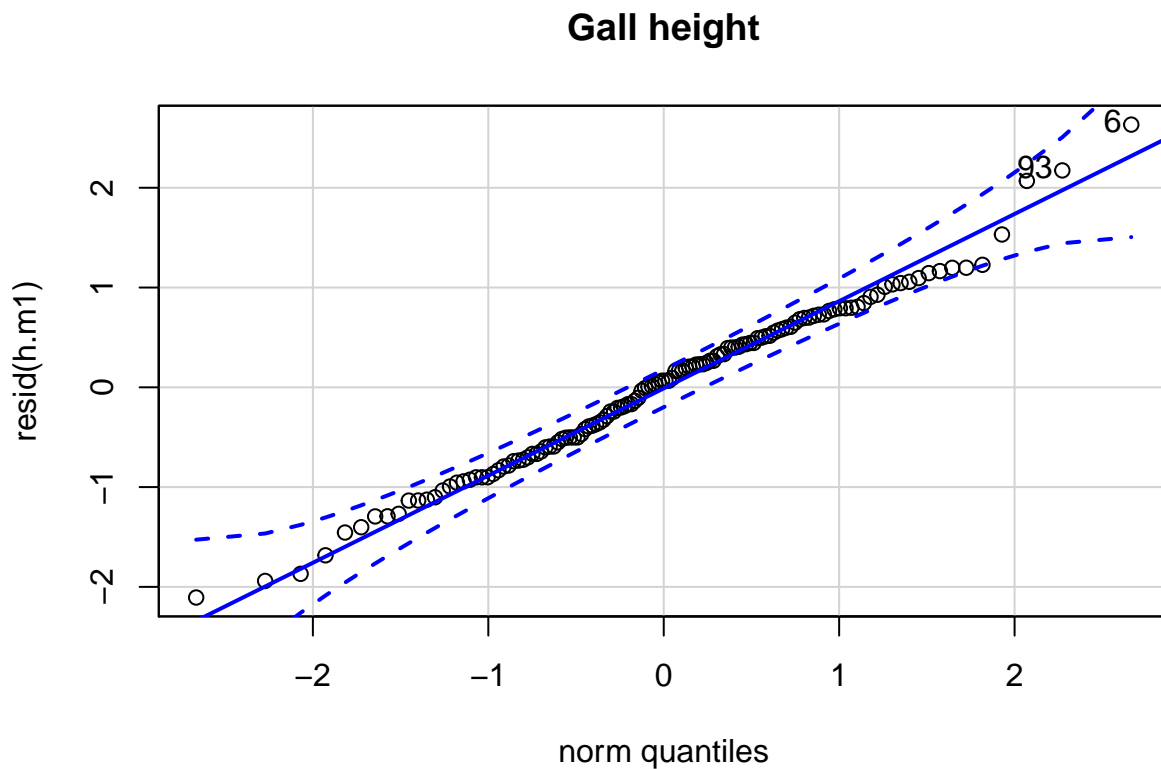
```
##      Df F value Pr(>F)
## group  5  0.251 0.9386
##      124
```

*# Assumption met*

*# (3) Normality of error term: need to check by histogram, QQplot of residuals, could do Kolmogorov-Smirnov*

*# Check for normal residuals*

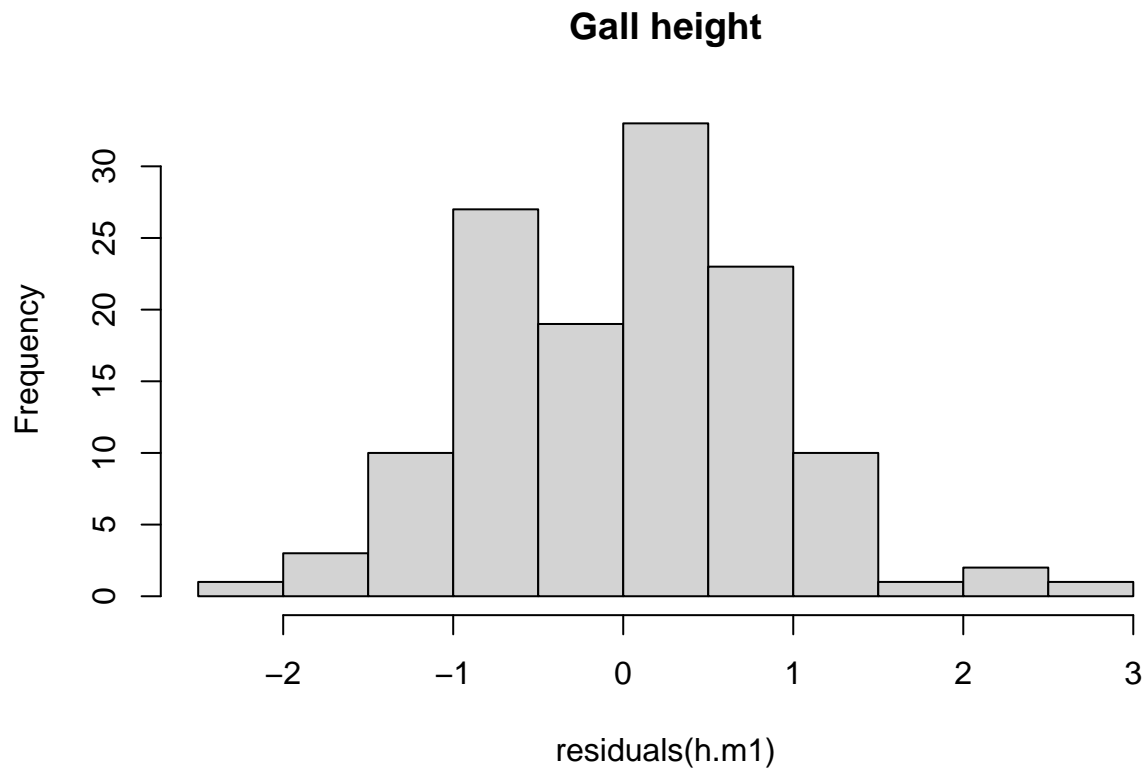
```
qqPlot(resid(h.m1), main = "Gall height")
```



```
## [1] 6 93
```

```
hist(residuals(h.m1), main = "Gall height")
```





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

```
##
## Shapiro-Wilk normality test
##
## data: resid(h.m1)
## W = 0.98993, p-value = 0.4677
```

```
# Model comparisons
h.m2 <- lm(gall_height ~ treatment, data=gall_only)
AICctab(h.m1, h.m2, weights=T)
```

```
##      dAICc df weight
## h.m1 0.0   7  0.71
## h.m2 1.8   6  0.29
```

```
# Model 1 fits the best
summary(h.m1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: gall_height ~ treatment + (1 | rep)
## Data: gall_only
##
##      AIC      BIC    logLik deviance df.resid
##    347.8    367.9   -166.9    333.8     123
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.47329 -0.70457  0.07763  0.67954  3.09126
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   rep      (Intercept) 0.06733  0.2595
##   Residual              0.72544  0.8517
## Number of obs: 130, groups: rep, 6
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)      2.7540     0.2010  28.9045  13.698  3.6e-14 ***
## treatmentdrought      0.3327     0.2329 123.9098   1.429   0.156
## treatmentirr_control    0.3380     0.2421 124.7086   1.396   0.165
## treatmentwarmed     -0.2507     0.2438 123.9484  -1.028   0.306
## treatmentwarmed_drought  0.1995     0.2368 123.8599   0.843   0.401
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) trtmntd trtmntr_ trtmntw
## trtmntdrght -0.624
## trtmntrr_cn -0.601  0.519
## tretmntwrmd -0.591  0.510  0.491
## trtmntwrmd_ -0.612  0.529  0.509  0.502
```

```
# Post hoc test to compare different levels
emmeans(h.m1, list(pairwise ~ treatment), adjust = "tukey")
```

```
## $'emmeans of treatment'
##   treatment      emmean    SE    df lower.CL upper.CL
##   ambient      2.75 0.212 38.0     2.32     3.18
##   drought      3.09 0.202 31.8     2.68     3.50
##   irr_control   3.09 0.212 37.9     2.66     3.52
##   warmed       2.50 0.216 39.8     2.07     2.94
##   warmed_drought 2.95 0.207 34.7     2.53     3.37
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
##
## $'pairwise differences of treatment'
##   1              estimate    SE    df t.ratio p.value
##   ambient - drought      -0.33273 0.237 128  -1.405  0.6256
##   ambient - irr_control   -0.33801 0.247 129  -1.371  0.6473
##   ambient - warmed        0.25071 0.248 128   1.011  0.8499
##   ambient - warmed_drought -0.19953 0.241 128  -0.829  0.9213
##   drought - irr_control   -0.00528 0.237 129  -0.022  1.0000
##   drought - warmed        0.58345 0.241 129   2.425  0.1152
##   drought - warmed_drought 0.13321 0.232 128   0.575  0.9785
##   irr_control - warmed     0.58873 0.250 130   2.356  0.1343
##   irr_control - warmed_drought 0.13848 0.242 129   0.573  0.9788
##   warmed - warmed_drought -0.45024 0.244 129  -1.844  0.3529
##
```

```
## Degrees-of-freedom method: kenward-roger  
## P value adjustment: tukey method for comparing a family of 5 estimates
```