

# Data\_Analysis\_Project2

Parker Hicks

4/21/2021

## Data

```
tit_data <- read.csv("/Users/parkerhicks/Desktop/Biostats_data/tit_fitness_data.csv"), stringsAsFactors = FALSE
tit_dataOmit <- na.omit(subset(tit_data, select=c("Weight", "wing", "broodSizeWhenSampled")))
#tit_dataOmit$broodSizeWhenSampled <- tit_dataOmit$broodSizeWhenSampled / 10
```

## Hypothesis

Null: There is no difference in the brood size of wild great tits on the basis of weight. There is no difference in the brood size of wild great tits on the basis of wing size. There is no difference in the brood size of wild great tits on the basis of weight and wing size

Alternative: There is a difference in the brood size of wild great tits on the basis of weight. There is a difference in the brood size of wild great tits on the basis of wing size. There is a difference in the brood size of wild great tits on the basis of weight and wing size

## Linear Model

```
tit_glm <- glm(broodSizeWhenSampled ~ Weight*wing, family="poisson", data = tit_dataOmit)
summary(tit_glm)
```

```
##
## Call:
## glm(formula = broodSizeWhenSampled ~ Weight * wing, family = "poisson",
##      data = tit_dataOmit)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.87156  -0.19867  -0.04929   0.35734   1.20731
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.248544   2.016328  -0.619   0.536
## Weight       0.146801   0.120275   1.221   0.222
## wing        0.061178   0.041186   1.485   0.137
## Weight:wing -0.003377   0.002421  -1.395   0.163
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

```
## Null deviance: 42.072 on 162 degrees of freedom
## Residual deviance: 38.277 on 159 degrees of freedom
## AIC: 580.95
##
## Number of Fisher Scoring iterations: 4
```

### Confidence Intervals

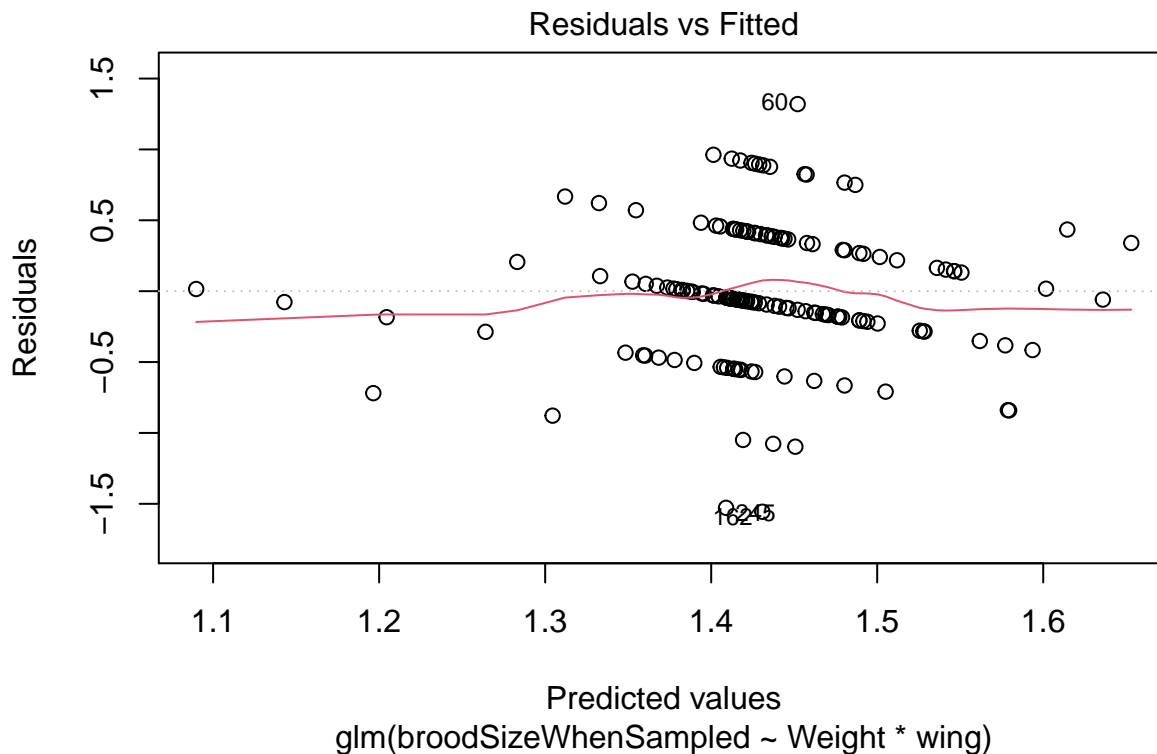
```
cov.tit_glm <- vcovHC(tit_glm, type="HC0")
std.err <- sqrt(diag(cov.tit_glm))
r.est <- cbind(Estimate= coef(tit_glm), "SE" = std.err,
"Pr(>|z|)" = 2 * pnorm(abs(coef(tit_glm)/std.err), lower.tail=FALSE),
Lower_CI = coef(tit_glm) - 1.96 * std.err,
Upper_CI = coef(tit_glm) + 1.96 * std.err)

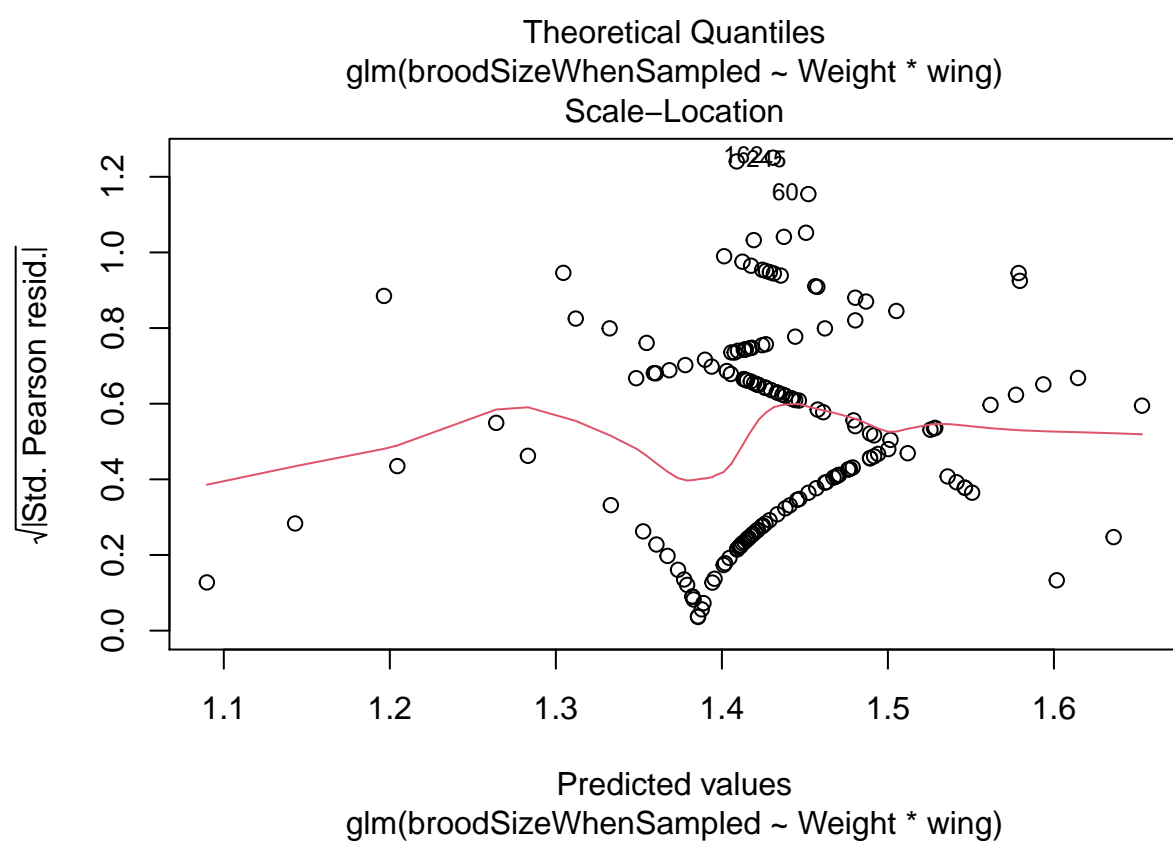
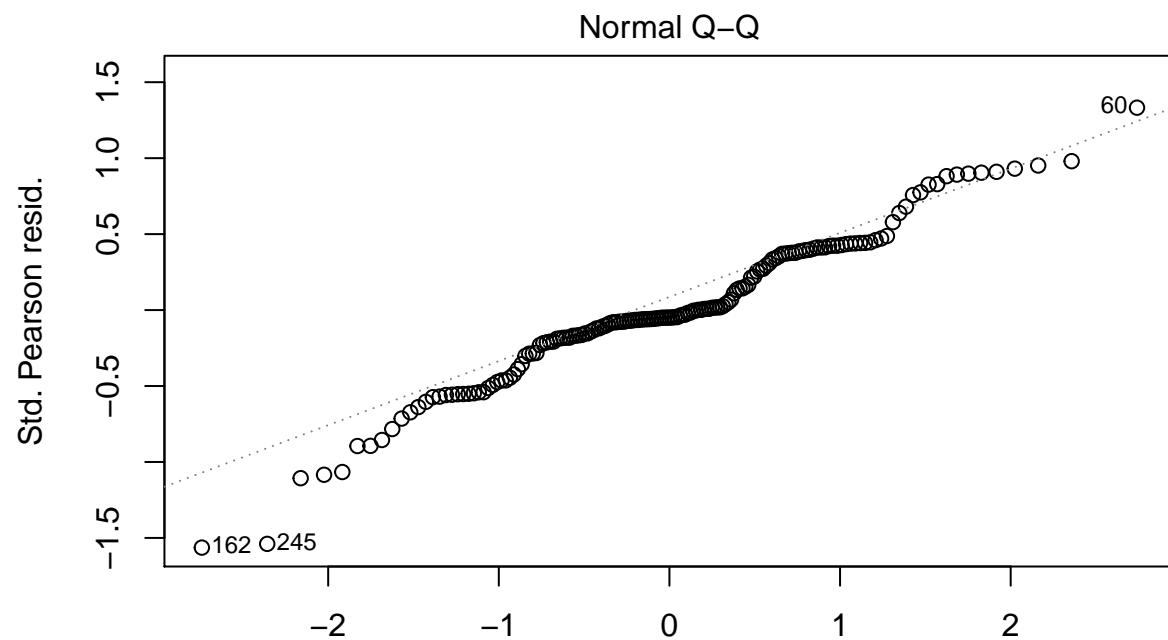
r.est
```

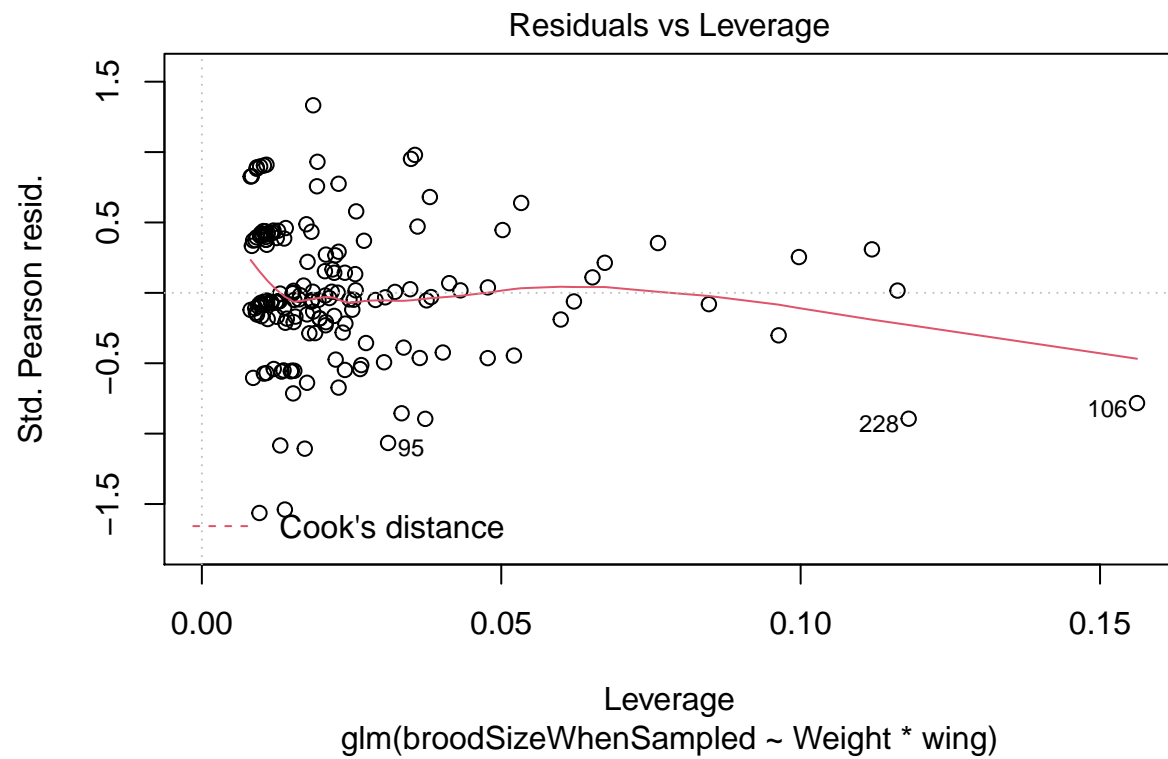
```
##           Estimate          SE    Pr(>|z|)    Lower_CI    Upper_CI
## (Intercept) -1.24854405 0.919561651 0.174539843 -3.050884887 0.553796785
## Weight      0.14680090 0.055743487 0.008450791 0.037543669 0.256058140
## wing        0.06117815 0.019014265 0.001293188 0.023910193 0.098446110
## Weight:wing -0.00337726 0.001137251 0.002981202 -0.005606273 -0.001148248
```

#Assumption Tests

```
plot(tit_glm)
```



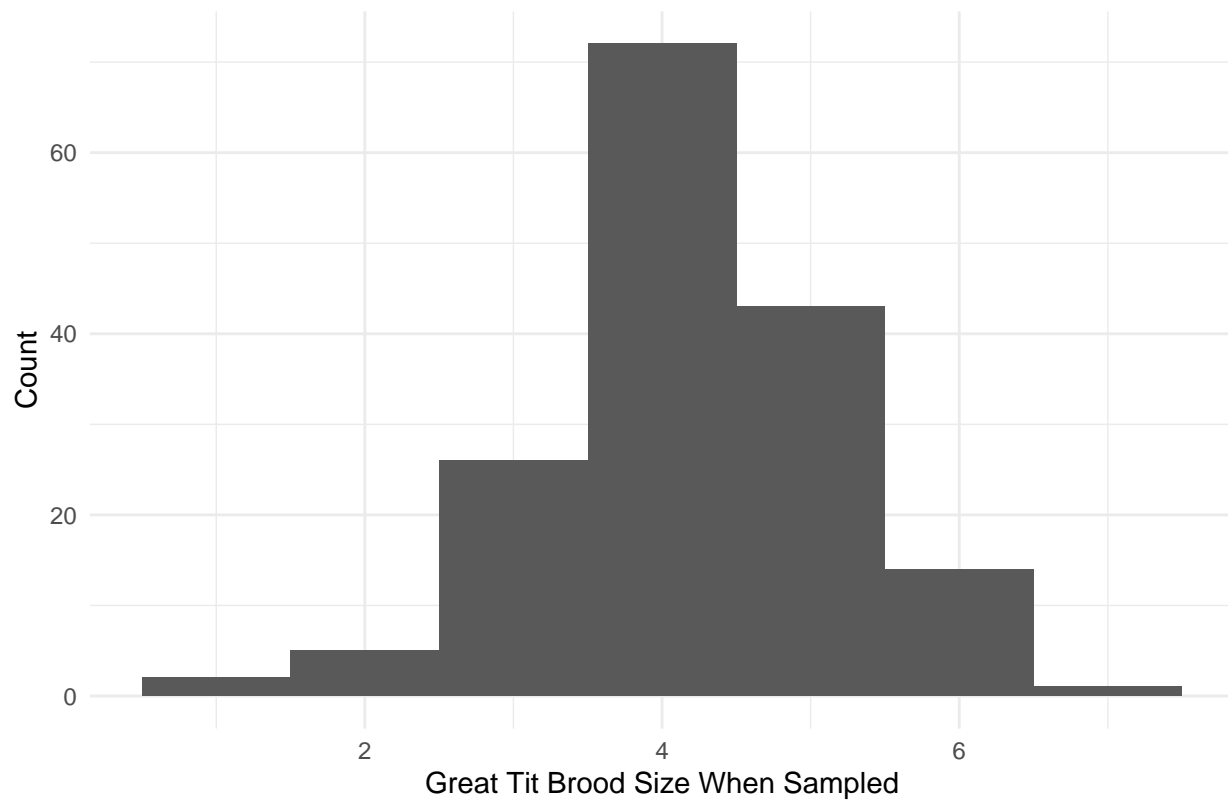




## Graph

```
ggplot(tit_data0mit, aes(x= broodSizeWhenSampled)) +
  geom_histogram(binwidth = 1) +
  theme_minimal() +
  xlab("Great Tit Brood Size When Sampled") +
  ylab("Count") +
  ggtitle("Distribution of Brood Sizes Among Great Tits ")
```

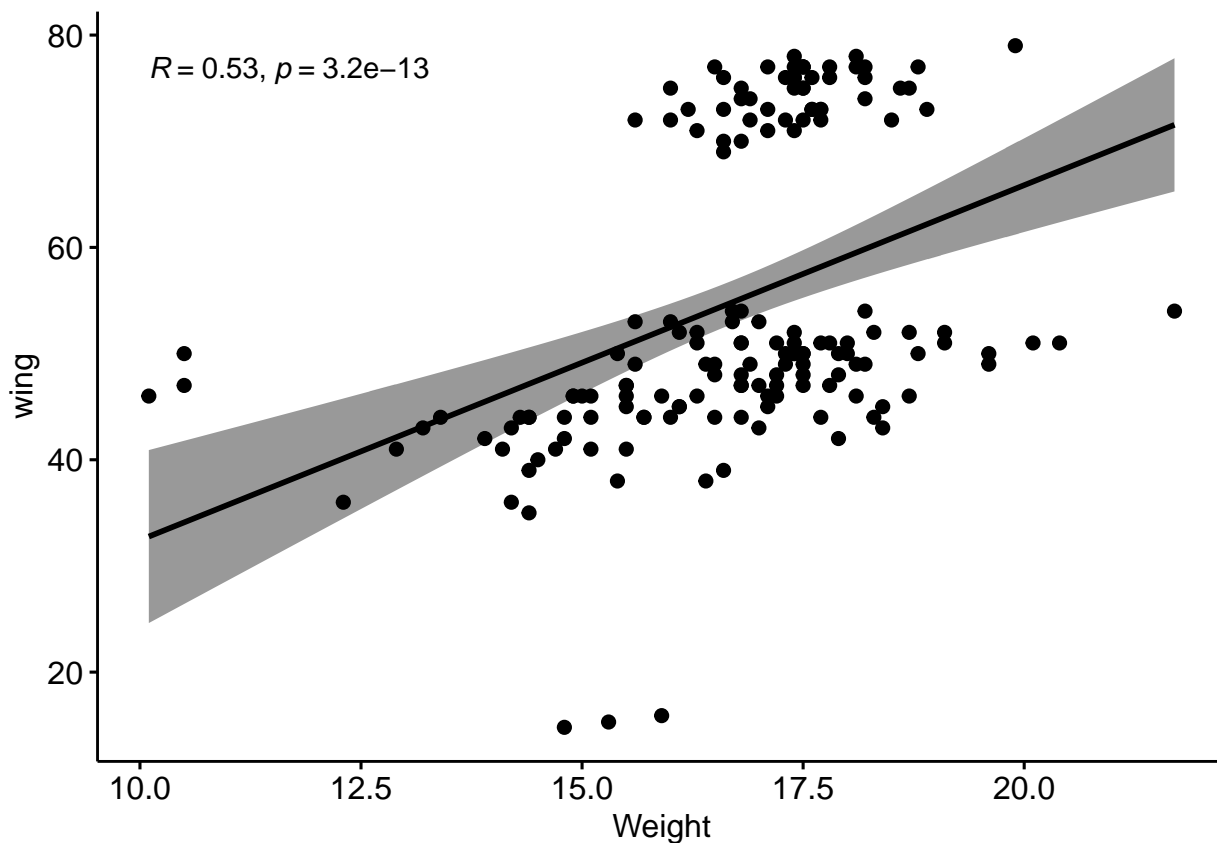
Distribution of Brood Sizes Among Great Tits



Brood size ~ weight

```
ggscatter(tit_data0mit, x="Weight", y="wing", add = "reg.line",  
          conf.int = TRUE, cor.coef = TRUE, cor.method = "spearman")
```

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



### Permutation Independence Test

```
independence_test(broodSizeWhenSampled ~ Weight*wing,
  data = tit_dataOmit)
```

```
##
## Asymptotic General Independence Test
##
## data: broodSizeWhenSampled by Weight, wing
## maxT = 2.2954, p-value = 0.04146
## alternative hypothesis: two.sided
```

### Brood size ~ weight \* wing size

```
ggplot(tit_dataOmit, aes(y=wing, x=Weight, color=factor(broodSizeWhenSampled)))+geom_point()+stat_smooth(m
  theme_minimal()
```

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

