A Time Series Analysis of the 1989 Space Shuttle Challenger Accident

Andrew Kiruluta

Contents


```
knitr::opts_chunk$set(echo = TRUE)
```

```
library(car)
library(dplyr)
library(Hmisc)
library(ggplot2)
library(mcprofile)
library(gridExtra)

df <- read.table(file = "challenger.csv", header = TRUE, sep = ",")
```

Exploratory Data Analysis

Introduction

The space shuttle Challenger explosion on takeoff at the Kennedy Space center in 1986 was a significant national traumatic event. It was later concluded that the culprit was the o-rings in the booster rocket sections. Given the data set from the Space Shuttle Challenger, we have been asked to analyze and test various models to find a good predictor of O-ring failure. We are then asked to choose a preferred model based on the analysis and use the same explanatory variables in a linear (rather then logistic) regression model. A very simple logistic regression model $\text{logit}(fail) = \log(\frac{fail}{1-fail}) = \beta_0 + \beta_1 Temp + \epsilon$ is determined to be the most explanatory and parsimonious given the data. After some analysis of the linear regression version of our chosen model, it was determined that logistic regression is more appropriate given that the example violates some of the basic conditions required for linear regression to be effective.

Exploratory Data Analysis

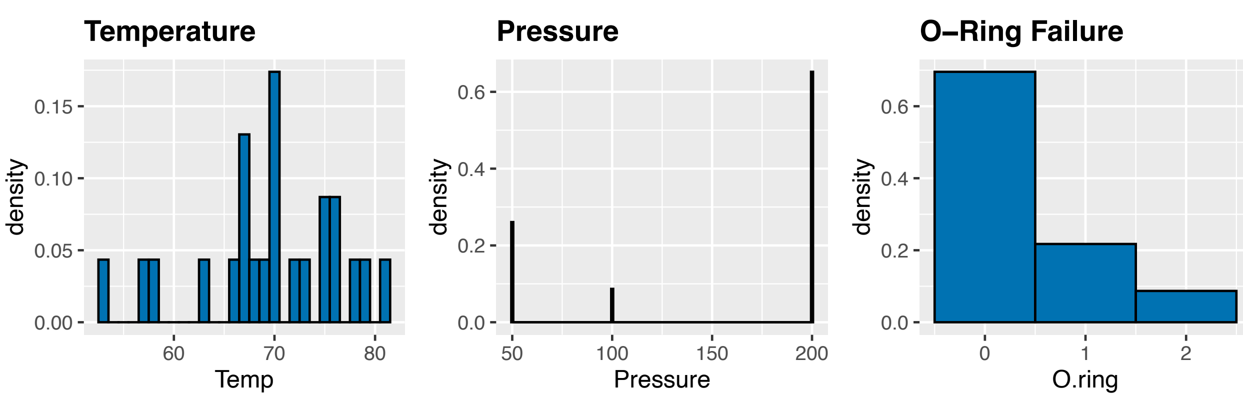
Univariate Analysis: A Look at the Explanatory variables Pressure and `Temp`, Response variable O.ring

First, we explore the features of our data set. We have 5 variables in the set. The Number variable represents the number of O-rings in the launch: which is invariantly 6 for all observations. The Flight variable is an identity property. Since it could describe the order in which the pre-accident launches occurred, it represents potentially confounding variable (e.g. relationship between Flight and Failures is positive, relationship between Flight and Temperature is positive leads to bias).

There are 23 observations of launches across a temperature range from 51F to 81F; only four temperatures were observed more than once. There are three pressure levels: 50, 100, and 200. We learned that putty alone can withstand pressure of 50 psi, thus actual pressure exerted on the O-ring were 0, 50, 150. 7 launches resulted in O-ring failure: 5 with 1 O-ring failure, and 2 with 2 O-ring failures for a total of 9 O-ring failures. There are no missing values in the data provided, and no evidence of incorrectly coded values (such as 999).

Pressure is an integer variable, but only takes 3 values, that is Pressure $\in \{50, 100, 200\}$. As a result, we'll convert this variable to a factor when creating the requisite models. The O.ring variable represents the number of O-rings that fail in any given launch. In some flights, we have more than one failing O-ring.

```
temp.plt <- ggplot(df, aes(x = Temp)) +
  geom_histogram(aes(y = ..density..), binwidth = 1, fill="#0072B2", colour="black") +
  ggtitle("Temperature") + theme(plot.title = element_text(lineheight=1, face="bold"))

pres.plt <- ggplot(df, aes(x = Pressure)) +
  geom_histogram(aes(y = ..density..), binwidth = 1, fill="#0072B2", colour="black") +
  ggtitle("Pressure") + theme(plot.title = element_text(lineheight=1, face="bold"))

oring.plt <- ggplot(df, aes(x = O.ring)) +
  geom_histogram(aes(y = ..density..), binwidth = 1, fill="#0072B2", colour="black") +
  ggtitle("O-Ring Failure") + theme(plot.title = element_text(lineheight=1, face="bold"))

grid.arrange(temp.plt, pres.plt, oring.plt, ncol=3)
```



A Visualization of individual variables reveals a few interesting features. First, the coldest temperature in our test launch data set is 53F: the majority of test flights were performed between 70F and

80F. Proportionally, most flights had no O-ring failures (16), of the remaining 7 flights, 5 had one failure, while 2 flights had two failures apiece. Finally, the majority of test launches used a nozzle pressure of 200 psi.

Basic Summary Data

Most temperatures in the data are within the range of 67F-75F, with a mean, median and mode of about 70F. Pressures used in a leak test performed prior to the launch are included in the data with a mode of 200 psi and a mean of 152 psi, and are heavily skewed towards 200psi.

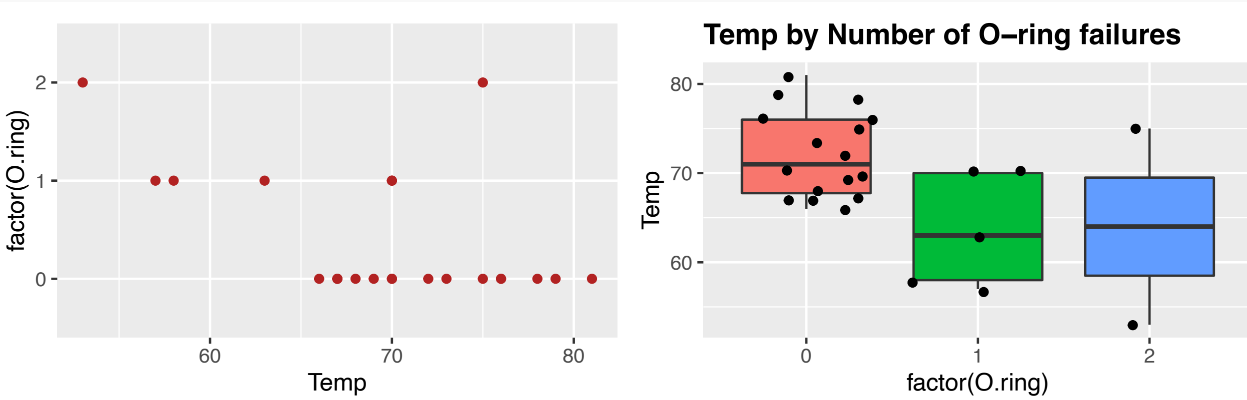
```
# Summary of the data series that provide some interesting summary data
summary(df[c("Temp", "Pressure", "O.ring")])
```

```
##       Temp          Pressure         O.ring
##  Min.   :53.00   Min.   : 50.0   Min.   :0.0000
##  1st Qu.:67.00   1st Qu.: 75.0   1st Qu.:0.0000
##  Median :70.00   Median :200.0   Median :0.0000
##  Mean   :69.57   Mean   :152.2   Mean   :0.3913
##  3rd Qu.:75.00   3rd Qu.:200.0   3rd Qu.:1.0000
##  Max.   :81.00   Max.   :200.0   Max.   :2.0000
```

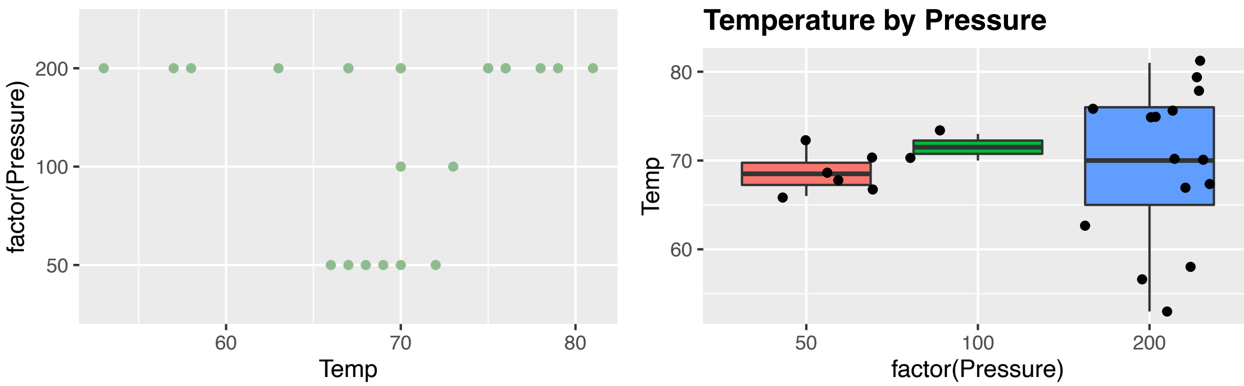
Bivariate Analyses: Relationships Between Time Series

There appears to be disproportionately more O-ring failures at lower temperatures than at higher temperatures. As such, all launches below 65F experienced at least 1 O-ring failure - it is also noted that visually there is only one documented case of failure above the mean/median value of 70F. There is no obvious visual interaction between the launch temperature and the PSI level used in the pre-launch pressure test. There is also no apparent relationship between pressure and O-ring failure based on visual inspection.

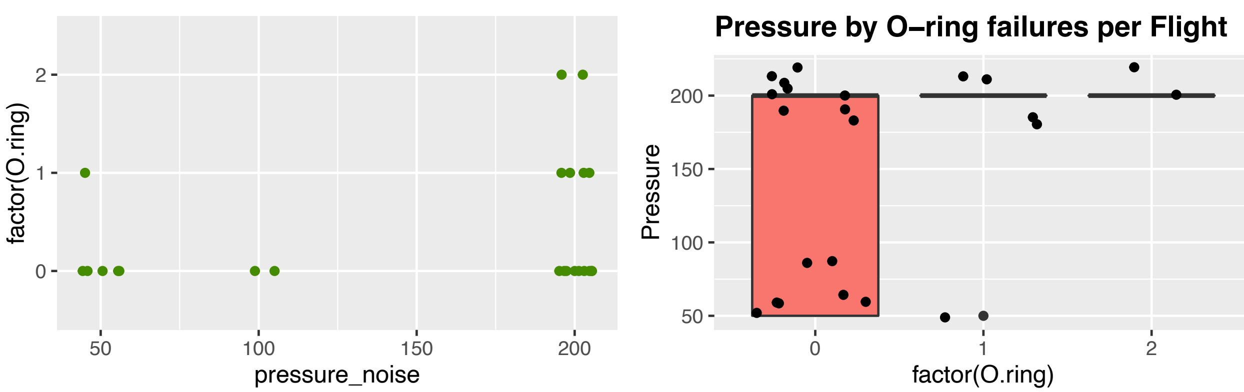
```
#Temp vs. O-ring Failures plots
otemp.plt <- ggplot(df, aes(Temp, factor(O.ring))) + geom_point(color="firebrick")
otemp.box <- ggplot(df, aes(factor(O.ring), Temp)) +
  geom_boxplot(aes(fill = factor(O.ring))) + geom_jitter() +
  guides(fill=FALSE) + ggtitle("Temp by Number of O-ring failures") +
  theme(plot.title = element_text(lineheight=1, face="bold"))
grid.arrange(otemp.plt, otemp.box, ncol=2)
```



```
#Temp vs. Pressure plots
tpres.plt <- ggplot(df, aes(Temp, factor(Pressure))) + geom_point(color="darkseagreen")
tpres.box <- ggplot(df, aes(factor(Pressure), Temp)) +
  geom_boxplot(aes(fill = factor(Pressure))) +
  geom_jitter() + guides(fill=FALSE) +  ggtitle("Temperature by Pressure") +
  theme(plot.title = element_text(lineheight=1, face="bold"))
grid.arrange(tpres.plt, tpres.box, ncol=2)
```




```
#Pressure vs. O-ring failure plots
noise <- runif(length(df$Pressure), min=-8, max = 8)
pressure_noise <- df$Pressure + noise
opres.plt <- ggplot(df, aes(pressure_noise, factor(O.ring))) + geom_point(color="chartreuse4")
opres.box <- ggplot(df, aes(factor(O.ring), Pressure)) +
  geom_boxplot(aes(fill = factor(O.ring))) + geom_jitter() + guides(fill=FALSE) +
  ggtitle("Pressure by O-ring failures per Flight") +
  theme(plot.title = element_text(lineheight=1, face="bold"))
grid.arrange(opres.plt, opres.box, ncol=2)
```

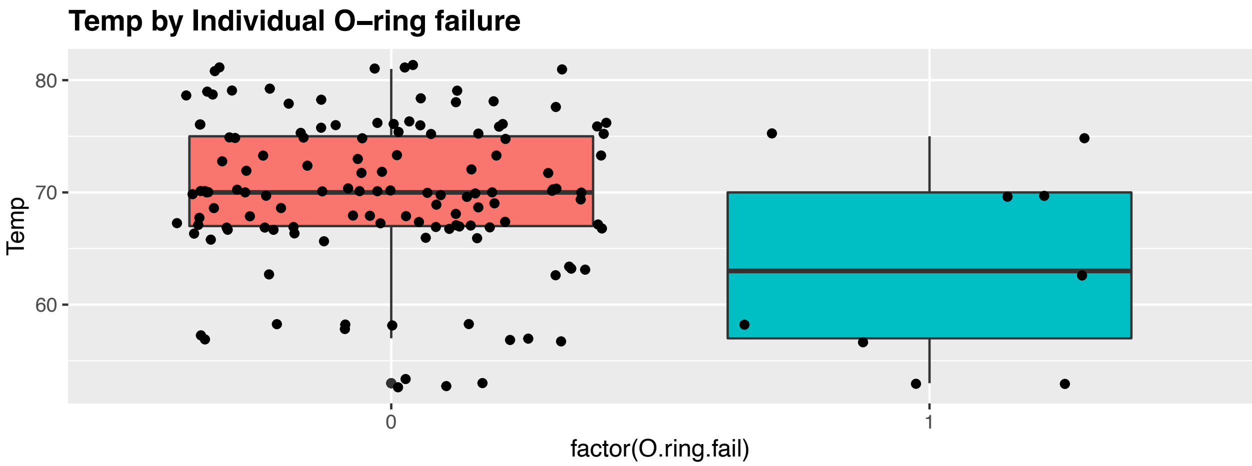


Considering each O-ring as an Independent Observation

A separate plot was created by treating every individual O-ring's behavior as a separate, independent event rather than accumulating failures into a single launch event. This new plot nicely displays that O-rings appear to fail more frequently as `Temp` declines below the 65F to 70F level.

```
#create a new dataframe that treats each O-ring independently
df2 <- data.frame(expand.grid(Flight = seq(1, 23, 1), O.ring.label = seq(1, 6, 1),
```

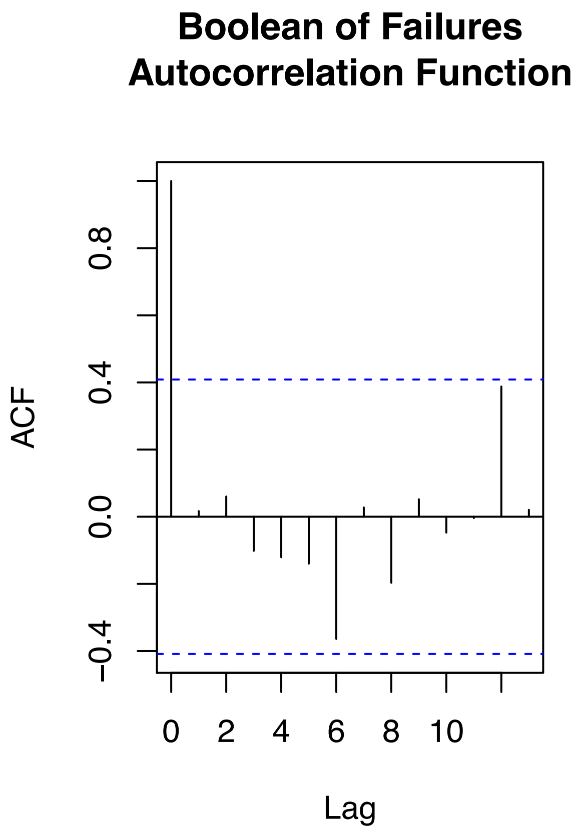
```
                                          Temp = 0, O.ring.fail = 0))
for(row in rownames(df2)){
  fl <- df2[row, ]$Flight
  df2[row, ]$Temp <- df[df$Flight == fl, ]$Temp # set Temp
  df2[row, ]$O.ring.fail <- ifelse(df2[row, ]$O.ring.label <=
                                  df[df$Flight == fl, ]$O.ring, 1, 0)  # set O.ring.fail
}
ggplot(df2, aes(factor(O.ring.fail), Temp)) +
  geom_boxplot(aes(fill = factor(O.ring.fail))) + geom_jitter() +
  guides(fill=FALSE) + ggtitle("Temp by Individual O-ring failure") +
  theme(plot.title = element_text(lineheight=1, face="bold"))
```

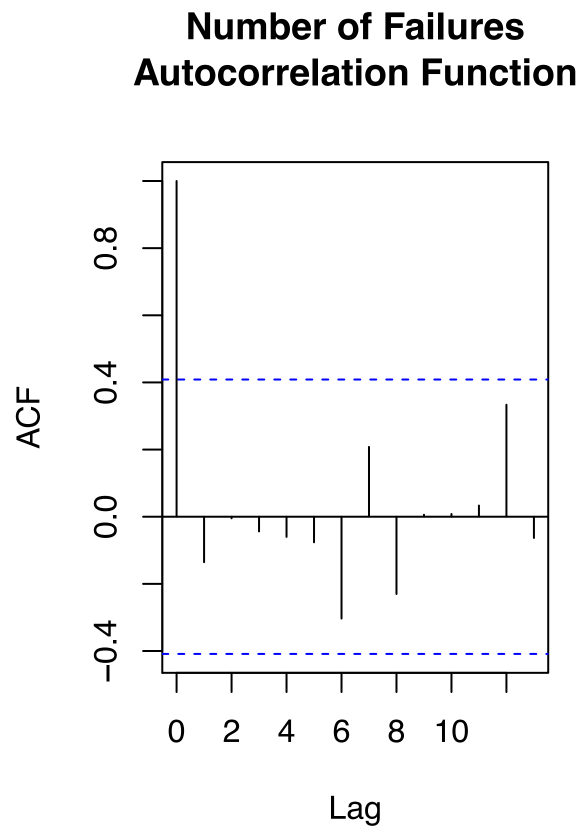


It is thus observed that of the six test launches during which an O-ring failure was observed, four were during the four coldest launches (Temperature by Failures plot). Proportionally, most failures also occur at a pressure of 200 psi, but it is difficult to immediately discern whether that is an individual effect of nozzle pressure or due to an overlap between the coldest launches and the nozzle pressure of 200 psi.

In this exploration, one would like to verify that O-ring failures are random and invariant to the order of testing. We will attempt to estimate the auto-correlation function of failure and O.ring testing order and observe if any amount of lag introduces significant correlation. As we can see in the correlograms, none of the spikes breach 5% significance blue line (except for lag=0 which is by design) and hence failures are random and invariant to the order of testing.

```
par(mfrow=c(1,2))
df$failure <- 1*(df$O.ring>0)
acf(df$failure, ylab="ACF", main=paste0("Boolean of Failures","\n","Autocorrelation Function"))
acf(df$O.ring, ylab="ACF", main=paste0("Number of Failures","\n","Autocorrelation Function"))
```





Why the assumption that probability of O-ring failure for each launch is independent necessary?

This issue was discussed in the EDA in part 1. The authors assume the probability of failure for each of the 6 O-rings is independent for each launch. This assumption is necessary in order to apply the binomial distribution to model the probability of failure. The binomial distribution assumes that the success/failure of each trial is independent, and in this case trials correspond to different O-rings in the same test. If binomial distribution assumptions do not hold, the logistic regression implying the odds of success/failure for each O-ring is invalid. Conceivably, the failure of one O-ring may contribute to some structural damage that causes other O-rings to fail, violating the independence assumption. On the other hand, the success of the primary O-ring may diminish the likelihood of failure of the second O-ring, if it does not experience the same conditions. There may also be omitted variables that influence O-ring quality or likelihood of failure, for example in conditions related to their production. These could also violate the independence assumption on a given launch or different launches.

Logistic regression model of probability of single O-ring failures based on linear relationship of temperature and pressure.

```
model1 <- glm(O.ring/Number ~ Temp + Pressure, data = df, family = binomial,
              weights=Number)
summary(model1)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp + Pressure, family = binomial,
##     data = df, weights = Number)
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.0361  -0.6434  -0.5308  -0.1625   2.3418
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.520195   3.486784   0.723   0.4698
## Temp        -0.098297   0.044890  -2.190   0.0285 *
## Pressure     0.008484   0.007677   1.105   0.2691
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 24.230  on 22  degrees of freedom
## Residual deviance: 16.546  on 20  degrees of freedom
## AIC: 36.106
##
## Number of Fisher Scoring iterations: 5
```

Likelihood ratio tests (LRTs) to judge the importance of the explanatory variables. Likelihood ratio tests were performed using this linear model as our alternative hypothesis and two reduced models setting the coefficients for temp and pressure respectively to zero, then conducting the ANOVA tests using the chi-squared distribution. We see that the inclusion of `Temp` in the model is significant, whereas the inclusion of `Pressure` is not even marginally significant.

```r
ha <- model1
h0 <- glm(O.ring/Number ~ Pressure, data = df, family = binomial, weights = Number)
anova(h0, ha, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: O.ring/Number ~ Pressure
## Model 2: O.ring/Number ~ Temp + Pressure
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     21.730
## 2        20     16.546  1   5.1838   0.0228 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
h0 <- glm(O.ring/Number ~ Temp, data = df, family = binomial, weights = Number)
anova(h0, ha, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: O.ring/Number ~ Temp
## Model 2: O.ring/Number ~ Temp + Pressure
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     18.086
```

```
## 2         20      16.546  1    1.5407   0.2145
```

Is Pressure Important? The lack of statistical significance of the pressure variable in the model above validates the authors' decision to remove Pressure from the model, however it is also reasonable to suggest that further testing may have still been warranted. The authors assume that the relationship between O-ring Failure and Pressure is linear, but some other transformation may be relevant. For example, a log transformation or a translation could be appropriate given the note in the paper that the putty covers pressure of 50 psi and thus it may be that only pressure in excess of 50 psi should be considered relevant to O-ring failure.

Estimate the model with only ***`Temp`****.*

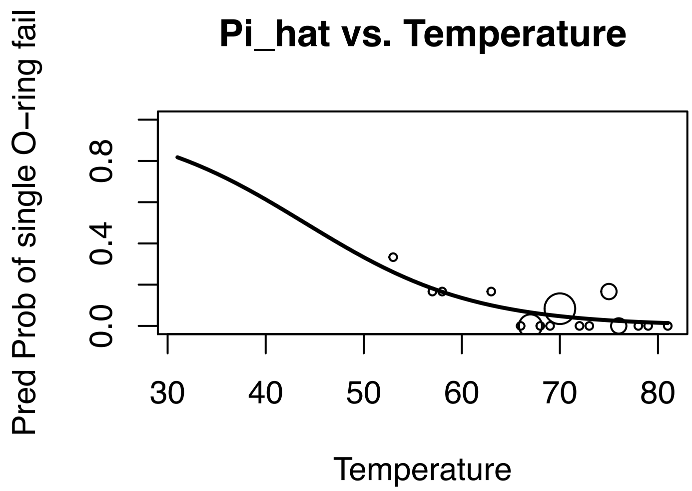
Simplified model: $logit(\pi) = \beta_0 + \beta_1 Temp$, where π is the probability of an O-ring failure. This model on `Temp` alone corresponds to the second `h0` model tested above and duplicated below for convenience. Using only a linear predictor on the `Temp` variable for the log-odds of yields an intercept of 5.085 and a coefficient for `Temp` of -0.116, which is significant at the 0.05 level.

```
model2 <- glm(O.ring/Number ~ Temp, data = df, family = binomial, weights = Number)
summary(model2)$coefficients
```

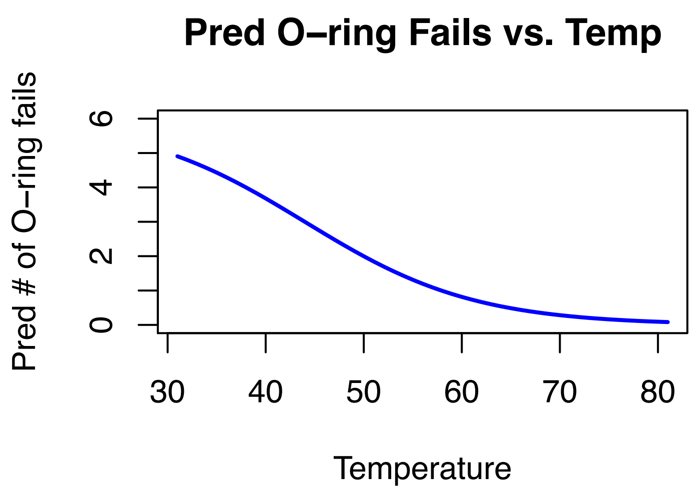
```
##                Estimate Std. Error   z value   Pr(>|z|)
## (Intercept)  5.0849772 3.05247412  1.665854 0.09574243
## Temp        -0.1156012 0.04702362 -2.458364 0.01395717
```

Q5b. ***Plot (1) $\hat{\pi}$ vs. `Temp` and (2) Expected number of failure vs. `Temp`.***

```
newdf <- data.frame(Temp = seq(from = 31, to = 81, by = 1)) #x-values to graph
## pi_hat vs. Temp
#calculate predicted values at each temp
lp.hat <- predict.glm(model2, newdata = newdf, type = "link", se.fit = TRUE)
lp.hat.mean <- lp.hat$fit
#calculate pi for each temp
pi.hat <- exp(lp.hat.mean) / (1 + exp(lp.hat.mean))

plot(newdf$Temp, pi.hat, ylim = range(c(0,1)),
     xlab = "Temperature", ylab = "Pred Prob of single O-ring fail",
     main = "Pi_hat vs. Temperature", type = 'l', col = 'black', lwd = 2)
#% failures for each temp
w <- aggregate(formula = O.ring/Number ~ Temp, data = df, FUN = sum)
# # of flights at each temp
n <- aggregate(formula = O.ring/Number ~ Temp, data = df, FUN = length)
symbols(x = w$Temp, y = (w$"O.ring/Number")/(n$"O.ring/Number"),
        circles = n$"O.ring/Number", inches = 0.08, add = TRUE)
```



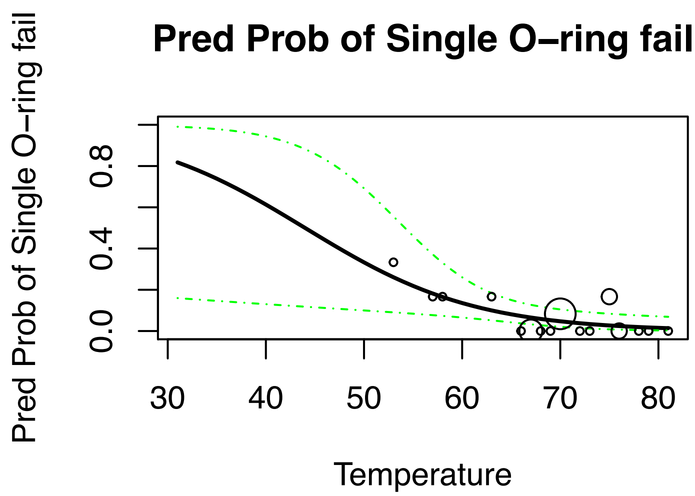

```
## expected number of failures vs. Temp
plot(newdf$Temp, pi.hat * 6, ylim = range(c(0,6)),
     xlab = "Temperature", ylab = "Pred # of O-ring fails",
     main = "Pred O-ring Fails vs. Temp", type = 'l', col = 'blue', lwd = 2)
```



Plot 95% Wald confidence interval bands. Why is the interval wider at lower temperatures? The bands are wider for lower temperature because there are very few observations in this region, which increases the standard error. The equation for a Wald confidence interval is: $\hat{\pi} - Z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{pi})}{n}} < \pi < \hat{\pi} + Z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$

```
#Create function to calculate CIs
ci.pi <- function(newdata, mod.fit.obj, alpha){
  linear.pred <- predict(object = mod.fit.obj, newdata = newdata, type = "link",
                         se = TRUE)
  #calculate linear CI from model
  CI.lin.pred.lower <- linear.pred$fit - qnorm(p = 1-alpha/2)*linear.pred$se
  CI.lin.pred.upper <- linear.pred$fit + qnorm(p = 1-alpha/2)*linear.pred$se
  #convert to pi
  CI.pi.lower <- exp(CI.lin.pred.lower) / (1 + exp(CI.lin.pred.lower))
  CI.pi.upper <- exp(CI.lin.pred.upper) / (1+ exp(CI.lin.pred.upper))
```

```
  list(lower = CI.pi.lower, upper = CI.pi.upper)
}

plot(newdf$Temp, pi.hat, ylim = range(c(0, 1)),
     xlab = "Temperature", ylab = "Pred Prob of Single O-ring fail",
     main= "Pred Prob of Single O-ring fail", type = 'l', col = 'black',
     lwd = 2)
curve(expr = ci.pi(newdata = data.frame(Temp = x), mod.fit.obj = model2,
                   alpha = 0.05)$lower, col = "green", lty = "dotdash", add = TRUE,
      xlim = c(31, 81))
curve(expr = ci.pi(newdata = data.frame(Temp = x), mod.fit.obj = model2,
                   alpha = 0.05)$upper, col = "green", lty = "dotdash", add = TRUE,
      xlim = c(31, 81))
symbols(x = w$Temp, y = (w$"O.ring/Number")/(n$"O.ring/Number"),
        circles = n$"O.ring/Number", inches = 0.08, add = TRUE)
```



Probability of an O-ring failure at 31F, comparison to the confidence interval, and necessary assumptions to apply the inference procedures At temperature of 31F, the model predicted that the probability of O-ring failure is 0.8178. The 95% Wald interval for π is $0.1596 < \pi < 0.9907$. Since we have only 23 data points, which is < 40, Wald CI generally does not work well. This is exacerbated by the fact that there is no data below 53F, so it is not even possible to estimate the local characteristics of the distribution at 31F. We therefore also check the profile likelihood ratio interval; the 95% interval for π is $0.1419 < \pi < 0.9905$. Despite small sample size, the profile likelihood ratio interval is not too far away from the Wald interval, thus we opt to report the profile likelihood ratio interval.

The key assumption being made is that there is a linear relationship between the temperature and the log-likelihood of O-ring failure. It is possible that either assumption is invalid, meaning that the logit is not the proper link-function for this relationship or there is a nonlinear relationship between temperature and the logit of the probability of O-ring failure. As the range of data we have for `Temp` is only 28 degrees (from 53F to 81F), 31F is 22 degrees lower than the minimum `Temp` we observe, which is almost as far away as the range of data we observe. A slightly non-linear relationship may not be as obvious with a range of 28 degree difference, but at 31F the deviance from linear relationship might be much more prominent.

```r
# Prob(failure) ~ temp = 31
model2.pred31 <- model2$coefficients[1] + model2$coefficients[2]*31

# Wald CI
predict.data<-data.frame(Temp=31)
pred31 <- predict(object = model2, newdata = predict.data, type = "link", se = TRUE)
pi.hat31 <- exp(pred31$fit) / (1 + exp(pred31$fit))
alpha <- 0.05
CI.pred31 <- pred31$fit + qnorm(p = c(alpha/2, 1-alpha/2))* pred31$se
CI.pi <- exp(CI.pred31)/(1 + exp(CI.pred31))
data.frame(predict.data, pi.hat31, lower = CI.pi[1], upper = CI.pi[2])
```

```
##   Temp  pi.hat31     lower     upper
## 1   31 0.8177744 0.1596025 0.9906582
```

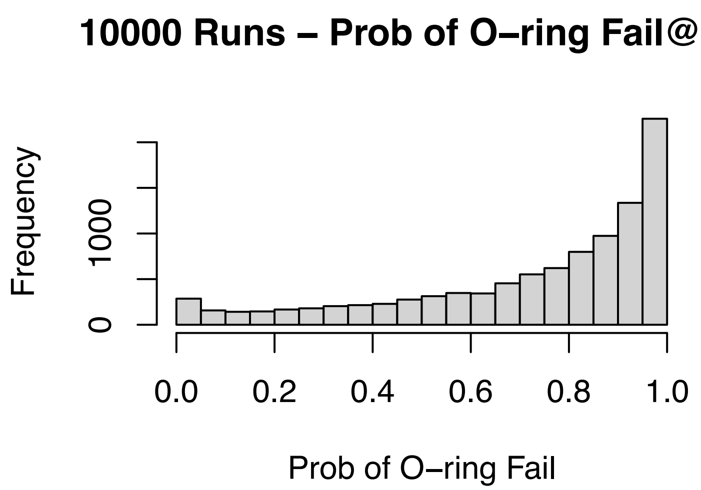
```r
# Profile Likelihood Ratio Interval
K <- matrix(data = c(1,31), nrow = 1, ncol = 2)
model2.combo <- mcprofile(object = model2, CM = K)
ci.logit.profile <- confint(object = model2.combo, level = 0.95)
exp(ci.logit.profile$confint)/(1 + exp(ci.logit.profile$confint))
```

```
##       lower     upper
## 1 0.1418508 0.9905217
```

A parametric bootstrap to compute the 90% c.i. at 31F and 72F. From secondary sources (NOAA: http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pdf/southeast_temp.shtml) on the nearest weather station, Orlando, the temperature is distributed approximately normal. We will thus sample with replacement as our bootstraping method. At temperature of 31F, the parametric bootstrapped 90% confidence interval for π is $0.1272 < \pi < 0.9936$ and at a temperature of 72F, the corresponding 90% confidence interval for π is $0.0101 < \pi < 0.0704$. It could be noted that the parametric bootstrapped CI is wider than the Wald CI and LRT intervals in Q5d despite a higher α. This demonstrates how liberal the Wald CI can be when data is sparse especially when no local samples are available.

```r
#suppress warnings
oldw <- getOption("warn")
options(warn = -1)

#define sigmoid function for computing values of pi
sigmoid = function(x) {
   1 / (1 + exp(-x))
}
#start with the parameter estimates from our model and our Temp data
beta0 = model2$coefficients[1]
beta1 = model2$coefficients[2]
x <- df$Temp
weights <- df$Number
#simulate new O.ring failure counts to estimate new model parameters
set.seed(23)
```

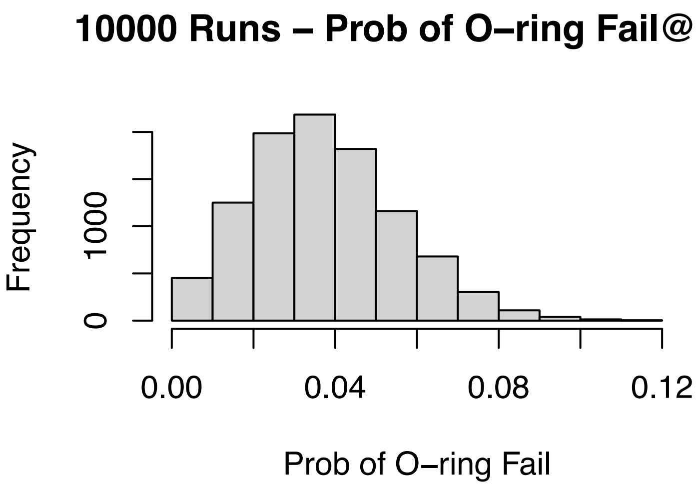
```r
sim <- function(){
  #Sample temp data with replacement (bootstrap)
  x.sample <- sample(x, 23, replace = TRUE)
  #Calculate pi
  pi <- sigmoid(beta0 + beta1*x.sample)
  #simulate new O.ring failure counts as binomial random variable with n=6
  #trials and p=pi probability of success
  y <- rbinom(n = length(x.sample), size = 6, prob = pi)

  #fit a new regression model on the simulated O.ring failure counts
  mod.fit <- glm(y/weights ~ x.sample, family = binomial, weights = weights)
  beta0.star = mod.fit$coefficients[1]
  beta1.star = mod.fit$coefficients[2]

  #use new model to compute predicted probability of O.ring failure at Temp = 31
  #and 72 degrees
  pi_star.31degrees <- sigmoid(beta0.star + beta1.star*31)
  pi_star.72degrees <- sigmoid(beta0.star + beta1.star*72)
  return(c(pi_star.31degrees,pi_star.72degrees))
}

#run simulation 10000 times
n=10000
sim_vals <- replicate(n,sim())
#plot distribution of computed pi values and return the 90% conf interval for
#Temp = 31 degrees
hist(sim_vals[1,], freq = T, xlab = "Prob of O-ring Fail",
     main = "10000 Runs - Prob of O-ring Fail@31F")
```




```r
quantile(sim_vals[1,],c(0.05,0.95))
```

```
##        5%       95%
## 0.1191875 0.9930378
```

```r
#plot distribution of computed pi values and return the 90% conf interval for
#Temp = 72 degrees
hist(sim_vals[2,], freq = T, xlab = "Prob of O-ring Fail",
     main = "10000 Runs - Prob of O-ring Fail@72F")
```




```r
quantile(sim_vals[2,],c(0.05,0.95))
```

```
##          5%        95%
## 0.01057184 0.06955134
```

```r
#restore old warning level
options(warn = oldw)
```

Is a quadratic term needed for `Temp`? This claim is tested by including a quadratic term on `Temp` and running a LRT using the chi-squared distribution to determine if its inclusion is statistically significant. The quadratic term's addition to the model is not statistically significant, suggesting either it shouldn't be included or some other variable transformations or terms should be conducted/tested first.

```r
model3 <- glm(O.ring/Number ~ Temp + I(Temp^2), data = df, family = binomial,
              weights = Number)
summary(model3)$coefficients
```

```
##                Estimate  Std. Error    z value  Pr(>|z|)
## (Intercept) 22.1261481 23.79442571  0.9298879 0.3524291
## Temp        -0.6508851  0.74075627 -0.8786765 0.3795767
## I(Temp^2)    0.0041405  0.00569214  0.7274066 0.4669769
```

```r
ha <- model3
anova(h0, ha, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: O.ring/Number ~ Temp
## Model 2: O.ring/Number ~ Temp + I(Temp^2)
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

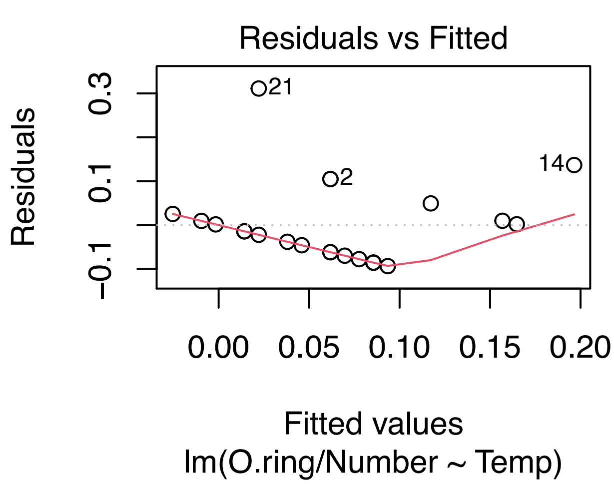
```
## 1        21      18.086
## 2        20      17.592  1    0.4947    0.4818
```

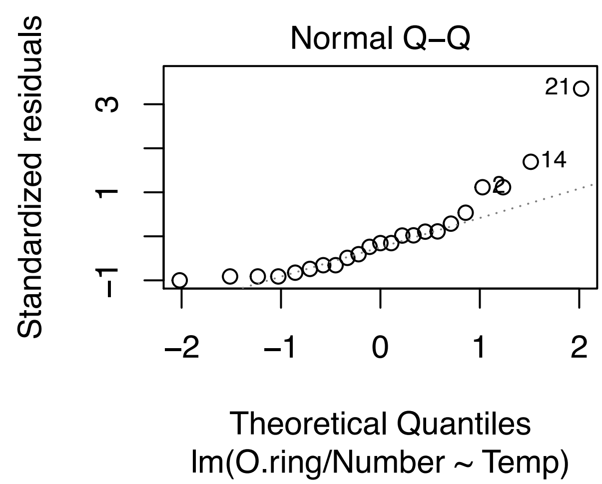
A Linear Regression Model

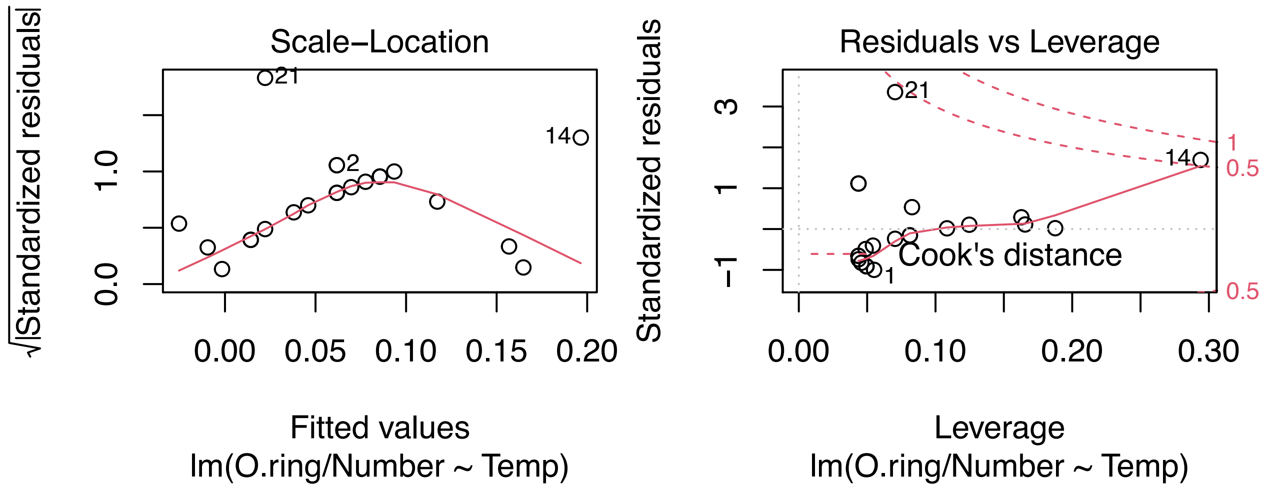
```
lin.model <- lm(O.ring/Number ~ Temp, data = df, weights = Number)
summary(lin.model)
```

```
##
## Call:
## lm(formula = O.ring/Number ~ Temp, data = df, weights = Number)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.22894 -0.16102 -0.03486  0.04311  0.76223
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.616402   0.203252   3.033  0.00633 **
## Temp        -0.007923   0.002907  -2.725  0.01268 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2357 on 21 degrees of freedom
## Multiple R-squared:  0.2613, Adjusted R-squared:  0.2261
## F-statistic: 7.426 on 1 and 21 DF,  p-value: 0.01268
```

```
plot(lin.model)
```








```r
#Temp below which we predict O-ring failure > 1
(1-lin.model$coefficients[1])/lin.model$coefficients[2]
```

```
## (Intercept)
##   -48.41402
```

```r
#Temp above which we predict O-ring failure < 0
(-lin.model$coefficients[1])/lin.model$coefficients[2]
```

```
## (Intercept)
##    77.79633
```

There are 6 model assumptions for the linear model. 1) We assume the true model is linear, which here is clearly invalid since it implies that for small enough or large enough temperatures the probability of O-ring failure will be outside of [0,1], which violates the laws of probability. 2) We assume samples are IID, but as discussed in 4a), the samples are not independent, in particular we have 6 samples per flight which all undergo roughly the same conditions outside of Temp/Pressure. 3) We assume there is no perfect collinearity between explanatory variables, which is not violated here as we only have a single explanatory variable. 4) We assume zero-conditional mean of residuals and exogeneity. The former appears violated in the residuals vs fitted plot above, as we expect negative residuals for intermediate temperatures, although it is difficult to tell with so few datapoints. Exogeneity holds as long as we assume a strictly associative relationship, however we are assuming that certain temperatures might directly cause failure, thus the model is implicitly causal. As a result, we must be aware of the possibility of omitted variable bias in our model, which may require some subject matter expertise to identify and test new explanatory variables outside of the scope of our current dataset. 5) We assume homoskedasticity of errors, which appears violated in the standardized residuals vs fitted plot, although again this is difficult to determine with such a small sample. 6) We assume errors are normally distributed, which appears to hold somewhat for our model looking at the normall q-q plot above, however we note that the sample is a bit too small to leverage the CLT, so the slight anormality is potentially a violation.

Given that certain linear model assumptions are explicitly violated in the example, most notably the assumption that the predicted failure for an O-ring should be bounded by [0,1], will not hold for temperatures below -48.4 degrees or above 77.8 degrees. Because of this, we would choose the binary logistic model over the linear model.

Overall Summary:

After eliminating other potential covariates like order of launch, pressure, shifts to pressure, interaction terms, square terms and log terms, we tested one more thing. Using a visual cue from the original `Temp` vs. failure chart in the EDA, we replaced the continuous variable `Temp` with a binary variable `bin.Temp` with values less than 65F set equal to 1. Our coefficient for the binary variable had an estimate of 1.9792, and while it had a relatively large standard error, was still highly significant. Exp(1.9792) = 7.237, which means that if the temperature is below the 65F threshold, odds of failure is 7.237:1 (or 6.237 times more likely) to occur as it would if the temperature is above the 65F threshold. This is a nice tidy answer, is reflective of our observations in EDA and has a great p-value but it reeks of p-hacking, and it would not be robust to further declining temperatures.

As a result it is ultimately best to go back to the basic O.ring ~ `Temp` single factor logistic regression model. That model is not as dramatic in terms of statistical significance but is still around a 95% confidence level and feels less forced. It implies that with every one degree increase in temperature the probability of an O-ring failure decreases 11% from what it was, and vice versa.

```
df$bin.Temp = df$Temp<65
model4 <- glm(O.ring/Number ~ bin.Temp, data = df, family = binomial,
              weights = Number)
summary(model4)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ bin.Temp, family = binomial, data = df,
##     weights = Number)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6547  -0.6547  -0.6547  -0.2582   2.4591
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.3142     0.5090  -6.511 7.46e-11 ***
## bin.TempTRUE   1.9792     0.7153   2.767  0.00566 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 24.230  on 22  degrees of freedom
## Residual deviance: 16.911  on 21  degrees of freedom
## AIC: 34.471
##
## Number of Fisher Scoring iterations: 5
```

```
exp(model4$coefficients[2])
```

```
## bin.TempTRUE
##     7.236842
```