Investigation of the 1989 Space Shuttle Challenger Accident

Srishti Mehra

The Space Shuttle Challenger disaster was a fatal accident in the United States space program that occurred on January 28, 1986, when the Space Shuttle Challenger (OV-099) broke apart 73 seconds into its flight, killing all seven crew members aboard; it was the first fatal accident involving an American spacecraft in flight. The mission carried the designation STS-51-L and was the tenth flight for the Challenger orbiter and twenty-fifth flight of the Space Shuttle fleet. The crew was scheduled to deploy a communications satellite and study Halley's Comet while they were in orbit. The spacecraft disintegrated over the Atlantic Ocean, off the coast of Cape Canaveral, Florida, at 11:39 a.m. EST (16:39 UTC).

The investigation of that failure by estimating probabilities of that failure in different conditions is what this report covers. There is a published report by Dalal et al published in the Journal of the American Statistical Association, Vol. 84, No. 408 (Dec., 1989), pp. 945- 957 which was referenced.

Loading the data

```
challenger_df<-read.csv("challenger.csv")
attach(challenger_df)
challenger_df$atleast_one_failure <- ifelse(O.ring>0,1,0)
challenger_df$pi<-O.ring/Number
```

Added two variables here that will be used later in the report: 1. atleast_one_failure - represents 1 for the flights that had at least one O.ring failure and 0 for the flights that had no failure. 2. pi - represents the ratio: $\frac{number of failed O.rings}{total number of O.rings}$

Part 1. EDA of the data set, including univariate, bivariate, and trivariate analysis, and examination of anomalies and missing values.

```
describe(challenger_df)
```

```
## challenger_df
##
##  7  Variables      23  Observations
## ---------------------------------------------------------------------------------
## Flight
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##       23        0       23        1       12        8      2.1      3.2
##      .25      .50      .75      .90      .95
##      6.5     12.0     17.5     20.8     21.9
##
## lowest :  1  2  3  4  5, highest: 19 20 21 22 23
## ---------------------------------------------------------------------------------
## Temp
```

```
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##       23        0       16    0.992    69.57    7.968     57.1     59.0
##      .25      .50      .75      .90      .95
##     67.0     70.0     75.0     77.6     78.9
##
## lowest : 53 57 58 63 66, highest: 75 76 78 79 81
##
## Value         53    57    58    63    66    67    68    69    70    72    73
## Frequency      1     1     1     1     1     3     1     1     4     1     1
## Proportion 0.043 0.043 0.043 0.043 0.043 0.130 0.043 0.043 0.174 0.043 0.043
##
## Value         75    76    78    79    81
## Frequency      2     2     1     1     1
## Proportion 0.087 0.087 0.043 0.043 0.043
## --------------------------------------------------------------------------------
## Pressure
##        n  missing distinct     Info     Mean      Gmd
##       23        0        3    0.706    152.2    67.59
##
## Value         50   100   200
## Frequency      6     2    15
## Proportion 0.261 0.087 0.652
## --------------------------------------------------------------------------------
## O.ring
##        n  missing distinct     Info     Mean      Gmd
##       23        0        3    0.654   0.3913   0.6087
##
## Value          0     1     2
## Frequency     16     5     2
## Proportion 0.696 0.217 0.087
## --------------------------------------------------------------------------------
## Number
##        n  missing distinct     Info     Mean      Gmd
##       23        0        1        0        6        0
##
## Value       6
## Frequency  23
## Proportion  1
## --------------------------------------------------------------------------------
## atleast_one_failure
##        n  missing distinct     Info      Sum     Mean      Gmd
##       23        0        2    0.636        7   0.3043   0.4427
##
## --------------------------------------------------------------------------------
## pi
##        n  missing distinct     Info     Mean      Gmd
##       23        0        3    0.654  0.06522   0.1014
##
## Value      0.0000000 0.1666667 0.3333333
## Frequency         16         5         2
## Proportion     0.696     0.217     0.087
## --------------------------------------------------------------------------------
```

```r
head(challenger_df)
```

```
##   Flight Temp Pressure O.ring Number atleast_one_failure        pi
## 1      1   66       50      0      6                   0 0.0000000
## 2      2   70       50      1      6                   1 0.1666667
## 3      3   69       50      0      6                   0 0.0000000
## 4      4   68       50      0      6                   0 0.0000000
## 5      5   67       50      0      6                   0 0.0000000
## 6      6   72       50      0      6                   0 0.0000000
```

Checking all unique values for variable 'Number'

```r
unique(Number)
```

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

This shows that the 'Number' variable which represents the number of O.rings in a flight is the same in all flights being analyzed.

Temperature

I know I do not have any missing values for Temperature from the output of the describe function. I will now see the distribution of the 'Temp' variable to get an idea of the range it's values form.

```r
paste("Minimum temperature at launch for sample flights: ", min(Temp))
```

```
## [1] "Minimum temperature at launch for sample flights:  53"
```

```r
paste("Maximum temperature at launch for sample flights: ", max(Temp))
```

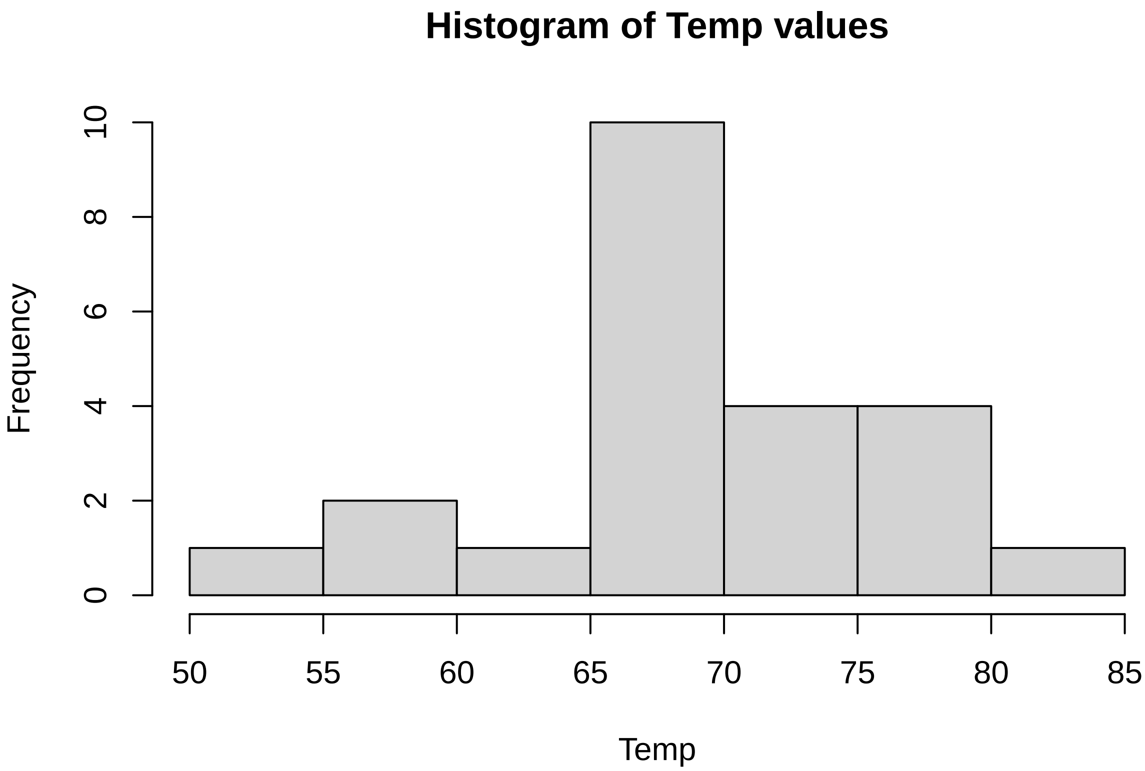
```
## [1] "Maximum temperature at launch for sample flights:  81"
```

```r
paste("Number of flights launched at <60F Temperature: ", length(which(Temp<60)))
```

```
## [1] "Number of flights launched at <60F Temperature:  3"
```

I calculated those under 60F because the paper mentions that the probability of failure would have dropped from 0.13 to at least 0.02 if the launch was postponed till a temperature of 60F was observed. These are only 3 in our dataset out of the 23 being analyzed.

```r
hist(Temp, main = "Histogram of Temp values", xlab = "Temp")
```



The temperatures at launch of the flights being analyzed are mostly above 65F or number of flights analyzed launched at temperatures below 65 are also few. I later notice that the number of failures in O-rings are also lower in number for these set of flights, which is something to ponder over.

Pressure

I know I do not have any missing values for Pressure from the output of the describe function. I will now see the distribution of the 'Pressure' variable to get an idea of the range it's values form.

```
table(Pressure)
```

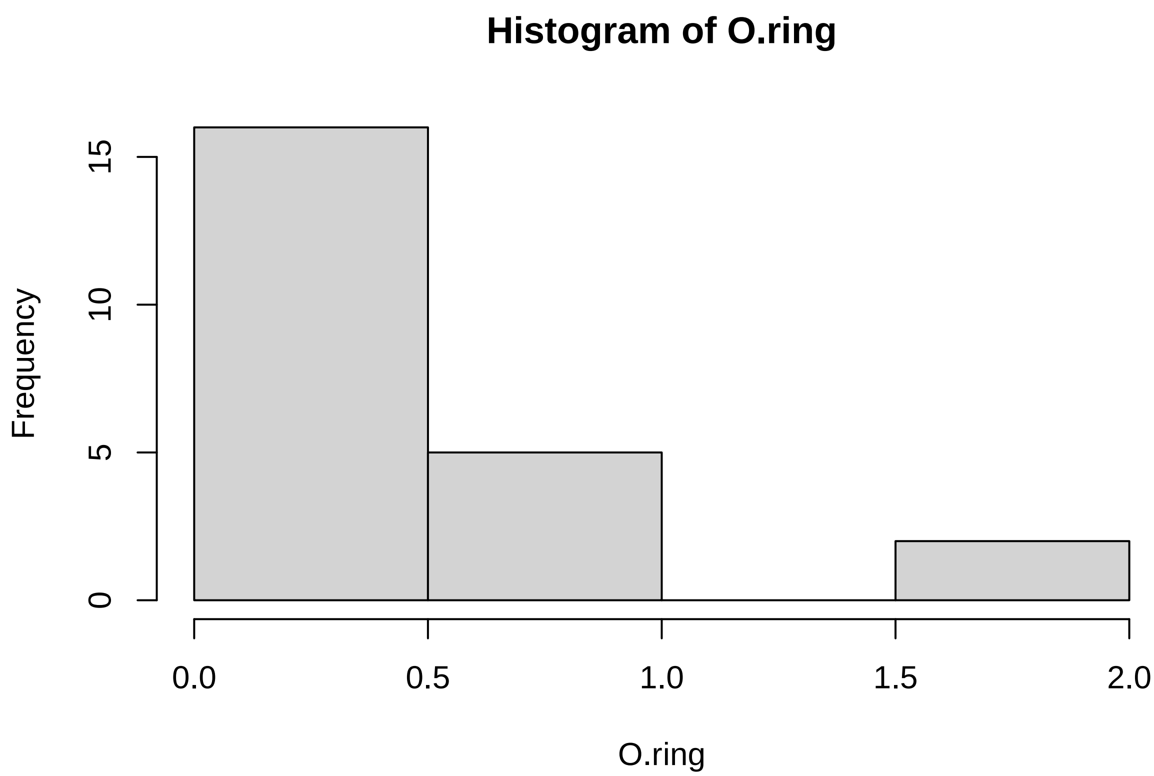
```
## Pressure
##  50 100 200
##   6   2  15
```

Most flights analyzed are launched at 200 but the distribution does not indicate clearly whether the low frequency of failures is related or not.

O-ring Failures

I know I do not have any missing values for O-ring failures from the output of the describe function. Let's also observe the distribution of the O-ring failures and see what percent of them fail.

```
hist(O.ring)
```

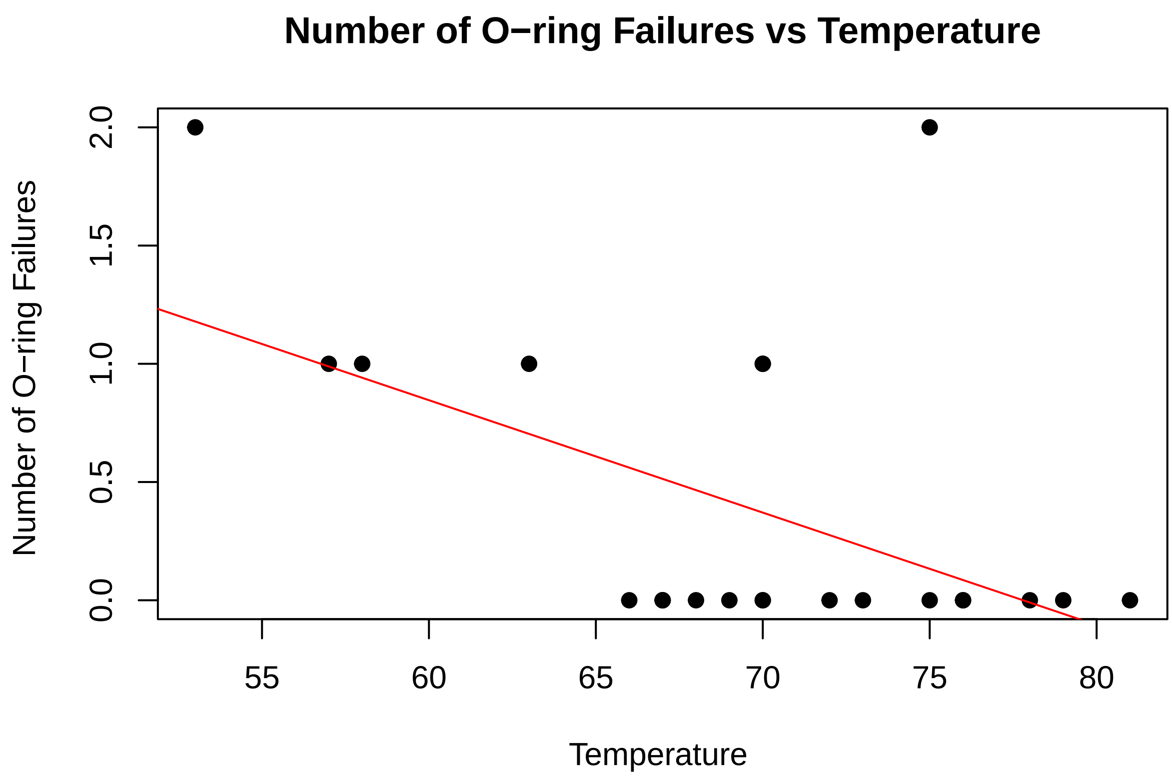


So, a maximum of 2 and minimum of 0 O-rings have failed out of 6 in the flights being examined with the cases of 0 failures being more than twice as high as at least one failure.

I will now analyze the bi-variate relationships between O-ring failures and Temp and O-ring failures and Pressure

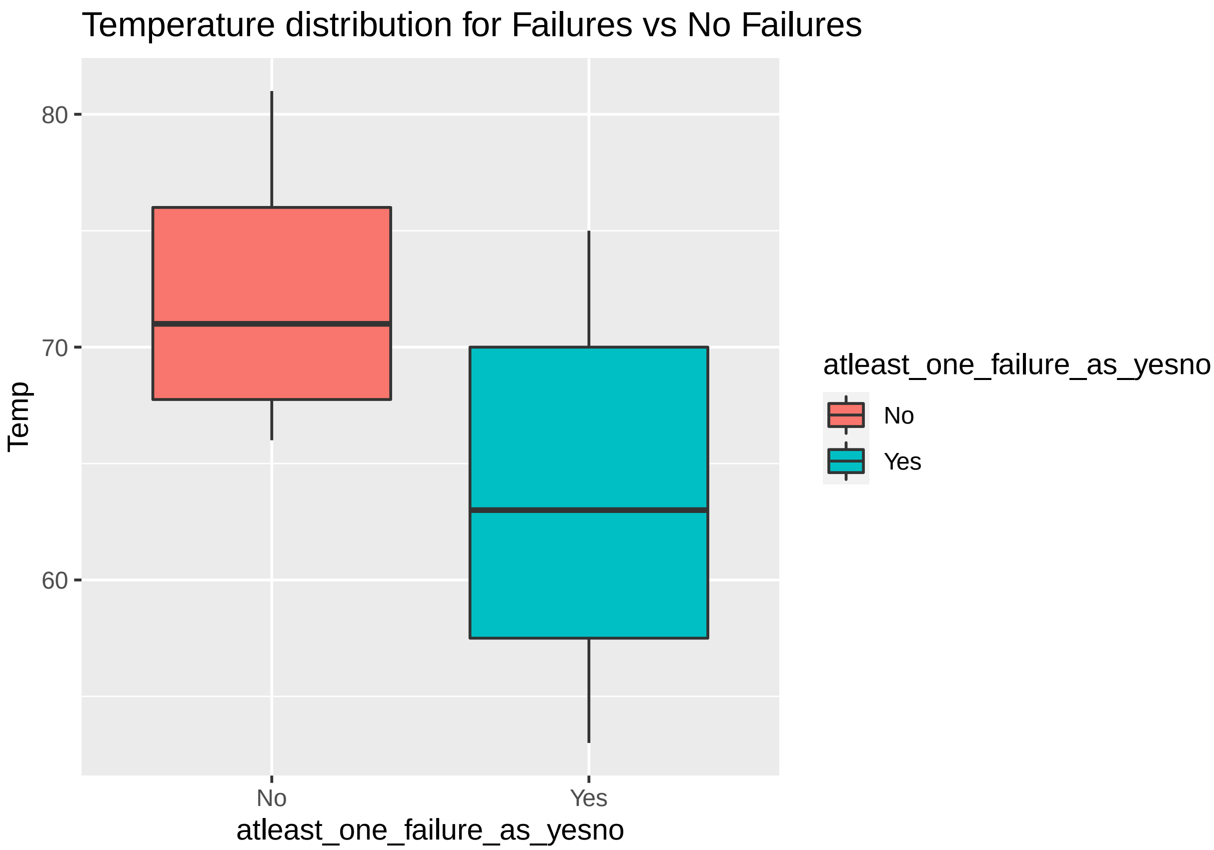
O-Ring Failures and Temp

```
plot(Temp, O.ring, main="Number of O-ring Failures vs Temperature",
   xlab="Temperature ", ylab="Number of O-ring Failures ", pch=19)
abline(lm(O.ring~Temp), col="red") # regression line (O.ring~Temp)
```



The relationship is not very linear but it seems like launches over 65 see no failures leaving out 2. Let's see what the box plots for temperature look for 0 failures v/s at least one failure.

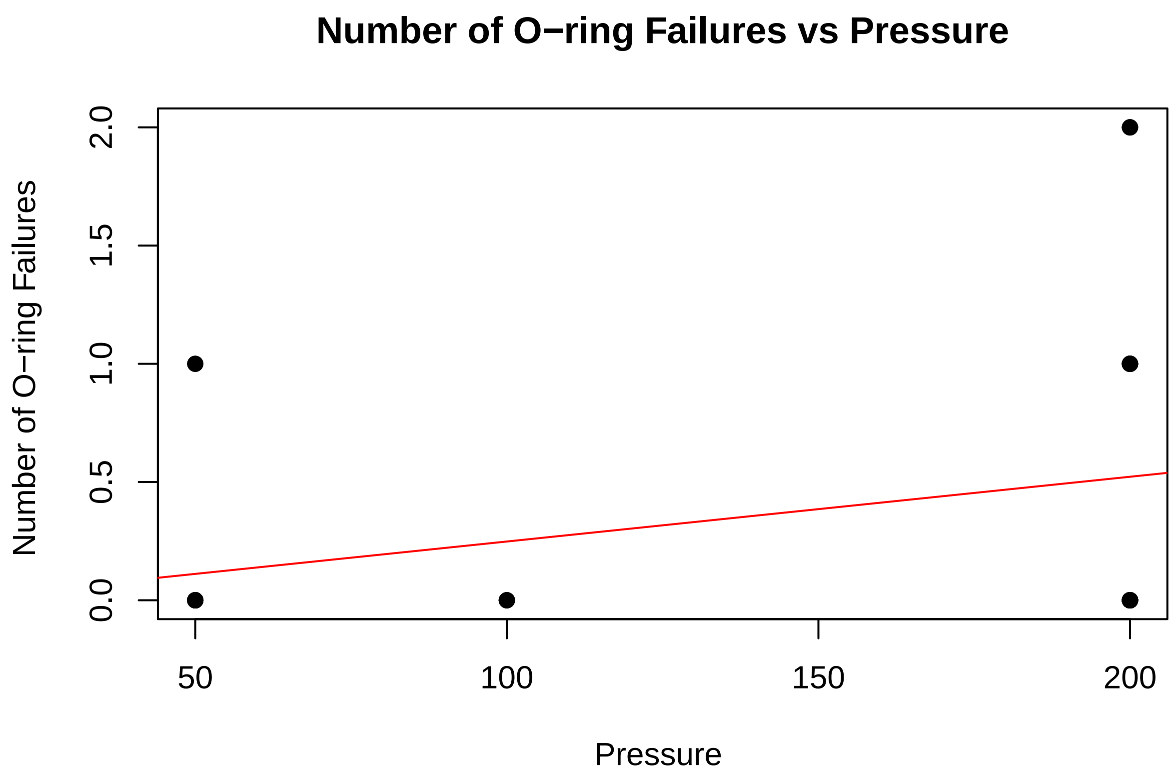
```
challenger_df$atleast_one_failure_as_yesno <- ifelse(O.ring>0,"Yes","No")
ggplot(challenger_df,mapping = aes(atleast_one_failure_as_yesno,Temp))+
  geom_boxplot(aes(fill = atleast_one_failure_as_yesno))+
  ggtitle("Temperature distribution for Failures vs No Failures")
```



There is considerable difference in the temperature distributions for flights that had an O-ring failure and those flights that had no O-ring failure.

O-Ring Failures and Pressure

```
plot(Pressure, O.ring, main="Number of O-ring Failures vs Pressure",
   xlab="Pressure ", ylab="Number of O-ring Failures ", pch=19)
abline(lm(O.ring~Pressure), col="red") # regression line (O.ring~Pressure)
```

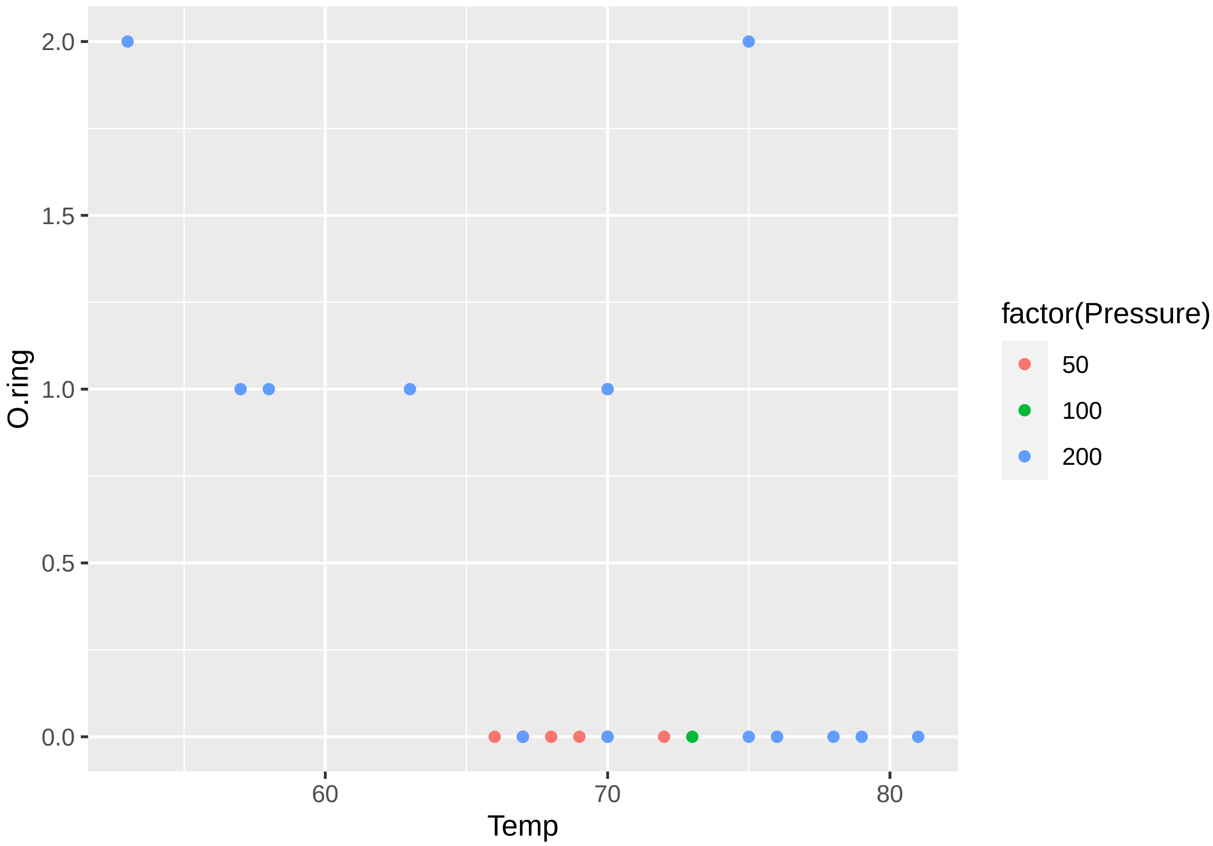


The relationship is not very linear here either but it seems like launches at 100 psi see no failure whereas other those at other pressures do. However, the number of different samples is low.

O-Ring Failures and Temp and Pressure

Finally let's observe what relationship O-ring failures have with temperature and pressure

```
ggplot(challenger_df,aes(Temp, O.ring,colour = factor(Pressure)))+geom_point()
```



The pressure variability seems to exist only in the flights that had no O-ring failures. All the ones with O-ring failures were launched at 200 psi, however, that does not say much about the Pressure variable's ability to inform failure of O-rings because there were no failures for flights launched at 200 psi as well.

So far I have understood that in our sample set of flights, there have been 0, 1, or 2 failure of O-rings (out of 6) in each flight. The launch temperature has ranged from 53F to 81F with lower number of sample flight launches being below 65F. The pressure at launch for sample flights was 50, 100, or 200 psi. I saw that the temperature distribution for launches with no failure in O-rings was visibly different from temperature distribution for launches with at least one O-ring failure and that there was no such difference for Pressure.

Part 2. Logistic Regerssion

a. The independence assumption is necessary because it allows us to calculate probability of number of O-ring failures in a flight from the failure probability of each without worrying about the correlation between each. It enables us to use maximum likelihood estimation for estimating parameters and reduces the number of parameters I will need to estimate by not having to estimate parameters for correlations between each variable with the other.

Potential Problems:

Binomial Model: It is described in the paper that an O-ring failure occurs due to distresses of two kind: erosion and blowby. Erosion is caused by excessive heat burning up the O-ring which is generally caused by a blowby and then causes blowby too. That excessive heat condition applies to all O-rings though so if I know 1 failure of O-ring occurred, I can get some information about another one occurring which is a problem in assuming independence between O-rings.

Binary Model: If I instead assume a binary model for at least one failure in the O-rings of a flight, I will need O-ring failures in one flight to be independent of those in other flights for all the 23 flights being analyzed. The potential problem with the independence assumption here is that the paper mentions parts of previous flights were recovered and used for future flights - if there was any erosion that was not till the extent of leading to a failure and that part was reused, that would inform us something about the probability of failure for the flight it was reused in.

b.

```
mod.glm.binom<-glm(pi ~ Temp+Pressure, family = binomial(link=logit),
                   data = challenger_df, weights=Number)
summary(mod.glm.binom)
```

Binomial Model:

```
##
## Call:
## glm(formula = pi ~ Temp + Pressure, family = binomial(link = logit),
##     data = challenger_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
```

The model in mod.glm.binom represents:

$$log\left(\frac{\pi}{1-\pi}\right) = 2.520195 - 0.098297\ Temp + 0.008484\ Pressure$$

where π is the probability of probability of failure per each O-ring.

The model summary shows us that the relationship between O.ring failure and the temperature at launch is statistically significant, the relationship between O.ring failure and pressure at launch, however, is not statistically significant.

Finding the standard deviation to assess right temperature increase to see the practical effect with (we could have seen the effect with a 1F increase but at 31F which was the temperature at the 1986 Challenger launch, 1 standard deviation of increase would have made much of a difference and 1F probably would not have).

```
sd(Temp)
```

```
## [1] 7.05708
```

According to this, the odds of seeing an O.ring in a flight launch change by $e^{7*(-0.098297)} = 0.502540523$ for every 7F (1 standard deviation) increase in temperature at launch. It can also be read as, odds of seeing an O.ring failure in a flight launch are 50% less for every 7F increase in temperature at launch.

The change in odds of seeing an O.ring failure with changes is pressure is not statistically significant.

```r
mod.glm.binary<-glm(factor(atleast_one_failure) ~ Temp+Pressure,
                    family = binomial(link=logit), data = challenger_df)
summary(mod.glm.binary)
```

Binary Model:

```
##
## Call:
## glm(formula = factor(atleast_one_failure) ~ Temp + Pressure,
##     family = binomial(link = logit), data = challenger_df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1993  -0.5778  -0.4247   0.3523   2.1449
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) 13.292360   7.663968   1.734   0.0828 .
## Temp        -0.228671   0.109988  -2.079   0.0376 *
## Pressure     0.010400   0.008979   1.158   0.2468
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 28.267  on 22  degrees of freedom
## Residual deviance: 18.782  on 20  degrees of freedom
## AIC: 24.782
##
## Number of Fisher Scoring iterations: 5
```

The model in mod.glm.binary represents:

$$log\left(\frac{\pi*}{1-\pi*}\right) = 13.292360 - 0.228671\ Temp + 0.010400\ Pressure$$

where $\pi*$ is the probability of failure of each O.ring.

The model summary also shows us that the relationship between at least one O.ring failure in a launch and the temperature at launch is statistically significant, but that with pressure is not.

According to this, the odds of seeing at least one failure of an O.ring in a flight launch change by $e^{7*(-0.228671)} = 0.201755845$ for every 7F (1 standard deviation) increase in temperature at launch. It can also be read as, odds of seeing at least one failure of an O.ring in a flight launch are 80% less for every 7F increase in temperature at launch.

The change in odds of seeing an O.ring failure with changes is pressure is not statistically significant.

c.

Binomial Model: Analyzing the importance of temperature as an explanatory variable:

```
mod.glm.binom.notemp<- glm(pi ~ Pressure, family = binomial(link=logit),
                           data = challenger_df, weights=Number)
anova(mod.glm.binom.notemp, mod.glm.binom, test='Chisq')
```

```
## Analysis of Deviance Table
##
## Model 1: pi ~ Pressure
## Model 2: pi ~ 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
```

I see that temperature is a statistically significant explanatory variable for assessing the probability of failure of an O-ring. It is evident from the p-value (assessed over a chi-squared distribution) for the anova test between the model that contains 'Temp' and the one that does not. Also from the large difference in residual deviance of the model that contains 'Temp' as compared to the one that does not.

Analyzing the importance of pressure as an explanatory variable:

```
mod.glm.binom.nopressure<- glm(pi ~ Temp, family = binomial(link=logit),
                               data = challenger_df, weights=Number)

anova(mod.glm.binom.nopressure, mod.glm.binom, test='Chisq')
```

```
## Analysis of Deviance Table
##
## Model 1: pi ~ Temp
## Model 2: pi ~ Temp + Pressure
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     18.086
## 2        20     16.546  1   1.5407   0.2145
```

I see that pressure is not a statistically significant explanatory variable for assessing the probability of failure of an O-ring. It is evident from the non-significant p-value p-value (assessed over a chi-squared distribution) for the anova test and not so large difference residual deviance of the model that contains 'Pressure' as compared to the one that does not.

Binary Model: Analyzing the importance of temperature as an explanatory variable:

```
mod.glm.binary.no.temp<- glm(factor(atleast_one_failure) ~ Pressure,
                             family = binomial(link=logit), data = challenger_df)
anova(mod.glm.binary.no.temp, mod.glm.binary, test='Chisq')
```

```
## Analysis of Deviance Table
##
## Model 1: factor(atleast_one_failure) ~ Pressure
## Model 2: factor(atleast_one_failure) ~ Temp + Pressure
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     26.536
## 2        20     18.782  1   7.7542 0.005359 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

I see that temperature is a statistically significant explanatory variable for assessing the probability of failure of at least one O-ring. It is evident from the p-value (here, <0.01 based on the chi-squared distribution) for the anova test and from the large difference in residual deviance of the model that contains 'Temp' as compared to the one that does not.

Analyzing the importance of pressure as an explanatory variable:

```
mod.glm.binary.no.pressure<- glm(factor(atleast_one_failure) ~ Temp,
                                  family = binomial(link=logit), data = challenger_df)
anova(mod.glm.binary.no.pressure, mod.glm.binary, test='Chisq')
```

```
## Analysis of Deviance Table
##
## Model 1: factor(atleast_one_failure) ~ Temp
## Model 2: factor(atleast_one_failure) ~ Temp + Pressure
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     20.315
## 2        20     18.782  1   1.5331   0.2156
```

I see that pressure is not a statistically significant explanatory variable for assessing the probability of failure of at least one O-ring. It is evident from the p-value and from the not-so-large difference residual deviance of the model that contains 'Pressure' as compared to the one that does not.

d. There is not much statistical significance to suggest that the variable 'Pressure' is important in assessing the probability of failure of an O.ring failure (or of at least one O-ring failure). I think the removal of pressure from the model is in line with the results obtained from hypothesis tests (LRT performed in part 2c and the Wald test values as seen in the summary of the models in part 2b). Addition of the variable does not give significant information on the probability of an O.ring failure. I also saw in the EDA in part 1 that there was not any relationship that could be derived between pressure and the flights that saw O.ring failures vs those that did not.

Part 3. Estimating probability of (with confidence intervals) failure at different temperatures

a. Estimating model

$$logit(\pi) = \beta_0 + \beta_1 Temp$$

```
mod.glm.binom.nopressure<- glm(pi ~ Temp, family = binomial(link=logit),
                                data = challenger_df, weights = Number)
summary(mod.glm.binom.nopressure)
```

```
##
## Call:
## glm(formula = pi ~ Temp, family = binomial(link = logit), data = challenger_df,
##     weights = Number)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.95227  -0.78299  -0.54117  -0.04379   2.65152
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.08498    3.05247   1.666   0.0957 .
## Temp        -0.11560    0.04702  -2.458   0.0140 *
## ---
```

```
## 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: 18.086  on 21  degrees of freedom
## AIC: 35.647
##
## Number of Fisher Scoring iterations: 5
```

This model in mod.glm.binom.nopressure represents:

$$log\left(\frac{\pi}{1-\pi}\right) = 5.08498 - 0.11560\ Temp$$

where π is the probability of probability of failure per each O-ring.

The model summary shows us that the relationship between O.ring failure and the temperature at launch is statistically significant.

According to this model, the odds of seeing an O.ring in a flight launch change by $e^{7*(-0.11560)} = 0.445214095$ for every 7F (1 standard deviation) increase in temperature at launch. It can also be read as, odds of seeing an O.ring failure in a flight launch are 55% less for every 7F increase in temperature at launch. We can see that some dependence between Temperature and Pressure that must have been in play in the model that contained both reduced the effect we read temperature has on the probability of failure of an O.ring.

b and c. Plotting: (1) Probability of failure vs. Temp and (2) Expected number of failures vs. Temp.

Probability of failure:

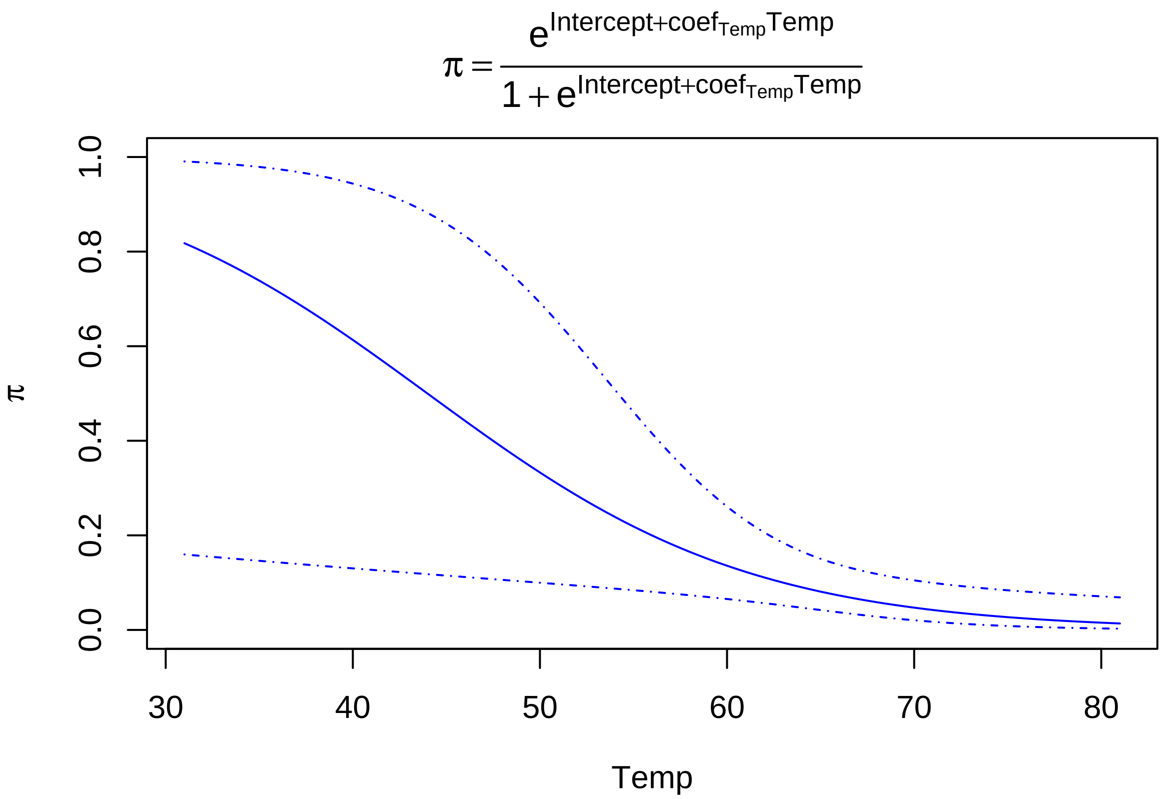
$$\pi = \frac{e^{5.08498-0.11560\ Temp}}{1+e^{5.08498-0.11560\ Temp}}$$

```
coef <- mod.glm.binom.nopressure$coefficients

curve(expr = exp(coef['(Intercept)'] + coef['Temp']*x)/(1+exp(coef['(Intercept)'] + coef['Temp']*x)),
    xlim = c(31,81),
    ylim = c(0,1),
    col = "blue",
    main = expression(pi == frac(e^{Intercept + coef[Temp]*Temp}, 1+e^{Intercept+coef[Temp]*Temp})),
    xlab =  expression(Temp), ylab = expression(pi))

ci.pi <- function (newdata, mod.fit.obj, alpha ) {
  linear.pred <- predict(object = mod.fit.obj, newdata=newdata,type="link", se=TRUE)
  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
  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)
}

curve (expr = ci.pi(newdata = data.frame(Temp = x),mod.fit.obj = mod.glm.binom.nopressure,
                    alpha = 0.05)$lower , col="blue", lty="dotdash", add=TRUE , xlim=c (31,81))
curve (expr = ci.pi(newdata = data.frame(Temp = x),mod.fit.obj = mod.glm.binom.nopressure,
                    alpha = 0.05)$upper , col="blue", lty="dotdash", add=TRUE , xlim=c (31,81))
```



The reason the interval bands are much wider for lower temperatures as compared to higher temperatures is because the amount of data available for lower temperatures is more sparse and therefore there is higher variance in the data points at lower temperatures. We saw in the EDA in Part 1, there are only 3 data points for temperatures under 60F (i.e. only 3 flights being analyzed were launched at temperatures under 60F). To achieve the same level of confidence with lesser data and more variance as with more data and lesser variance, the intervals need to wider.

Expected number of failures= Number * *pi*

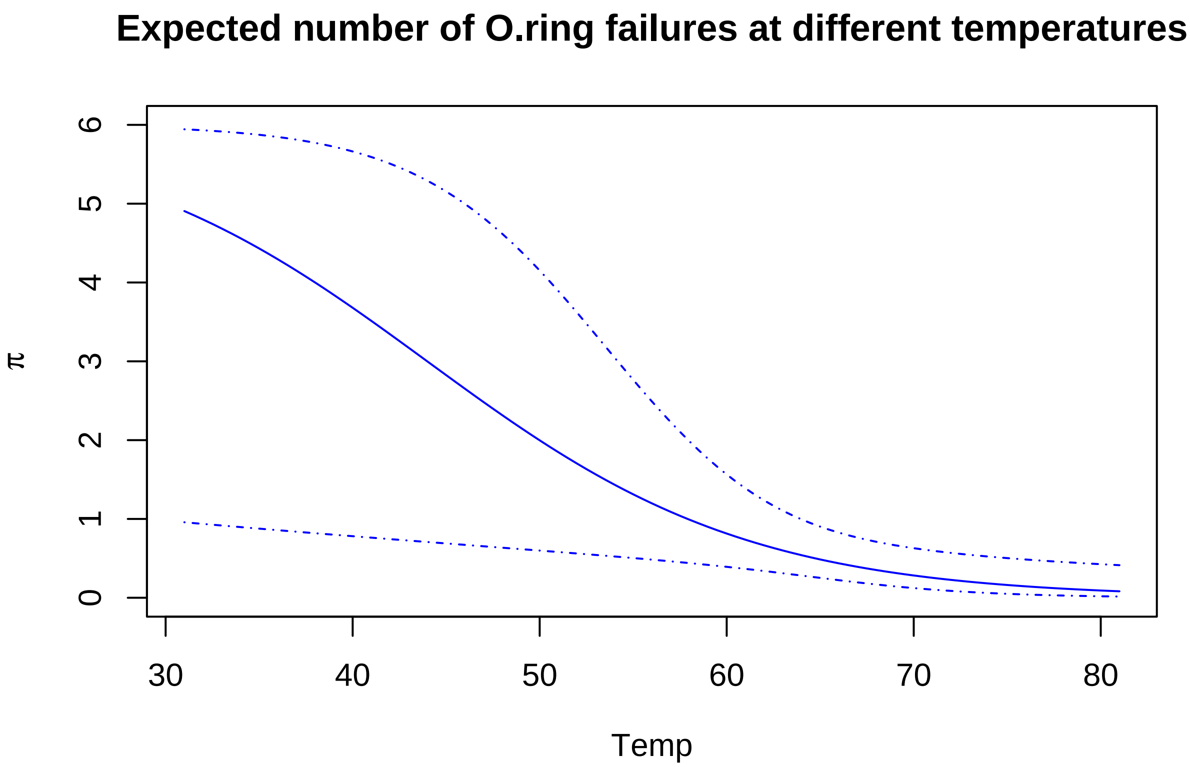
```
coef <- mod.glm.binom.nopressure$coefficients

curve(expr = 6*exp(coef['(Intercept)'] + coef['Temp']*x)/(1+exp(coef['(Intercept)'] + coef['Temp']*x)),
    xlim = c(31,81),
    ylim = c(0,6),
    col = "blue",
    main = "Expected number of O.ring failures at different temperatures",
    xlab =  expression(Temp), ylab = expression(pi))

ci.pi.exp <- function (newdata, mod.fit.obj, alpha ) {
  linear.pred <- predict(object = mod.fit.obj, newdata=newdata,type="link", se=TRUE)
  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
  CI.pi.lower <- 6*exp(CI.lin.pred.lower)/(1+exp(CI.lin.pred.lower))
  CI.pi.upper <- 6*exp(CI.lin.pred.upper)/(1+exp(CI.lin.pred.upper))
  list (lower = CI.pi.lower, upper = CI.pi.upper)
}

curve (expr = ci.pi.exp(newdata = data.frame(Temp = x),mod.fit.obj = mod.glm.binom.nopressure,
                        alpha = 0.05)$lower , col="blue", lty="dotdash", add=TRUE , xlim=c (31,81))
curve (expr = ci.pi.exp(newdata = data.frame(Temp = x),mod.fit.obj = mod.glm.binom.nopressure,
```

```
                                   alpha = 0.05)$upper , col="blue", lty="dotdash", add=TRUE , xlim=c (31,81))
```



d. Estimate the probability (with confidence interval) of an O-ring failure at 31°, temperature at launch for the Challenger in 1986

```
paste("The probability of an O.ring failure at 31F is: ",
      round(exp(coef['(Intercept)'] + coef['Temp']*31)/(1+exp(coef['(Intercept)'] + coef['Temp']*31)),2
```

```
## [1] "The probability of an O.ring failure at 31F is:  0.82"
```

Wald Confidence Interval:

```
wi <- ci.pi(newdata = data.frame(Temp = 31),mod.fit.obj = mod.glm.binom.nopressure, alpha = 0.05)

paste("The Wald CI for probability of an O.ring failure at 31F: ",
      round(wi$lower,2), "to", round(wi$upper,2))
```

```
## [1] "The Wald CI for probability of an O.ring failure at 31F:  0.16 to 0.99"
```

Profile Likelihood Confidence Interval:

```
K <- matrix (data=c(1, 31) , nrow = 1 , ncol = 2)
linear.combo <- mcprofile(object=mod.glm.binom.nopressure, CM=K)
ci.logit.profile <- confint(object=linear.combo, level=0.95)

plci <- exp(ci.logit.profile$confint)/(1+exp(ci.logit.profile$confint))

paste("The Profile Likelihood CI for probability of an O.ring failure at 31F: ",
      round(plci$lower,2), "to",round(plci$upper,2))
```

```
## [1] "The Profile Likelihood CI for probability of an O.ring failure at 31F:  0.14 to 0.99"
```

Since the Profile Likelihood Ratio confidence interval bounds are very close to the Wald confidence interval bounds, we can use the Profile Likelihood Ratio.

For calculating a Wald Confidence Interval there is an assumption on normal approximation of the distribution which is generally true for larger sample sizes. Since we do not have a large sample size, the Profile Likelihood Ratio Confidence Interval is more statistically reliable.

e. Copmaring different types of confidence intervals

```
sample.size<-23
num.simulations<-1000
set.seed(0)
sample_temp<-replicate(num.simulations, sample(Temp, sample.size, replace=TRUE))

pi<-exp(coef['(Intercept)'] + coef['Temp']*sample_temp)/
  (1 + exp(coef['(Intercept)'] + coef['Temp']*sample_temp))

y<-rbinom(n=sample.size*num.simulations, size=6, prob = pi)
y.mat<-matrix(data=y, nrow = sample.size, ncol = num.simulations)

# Function to estimate model and obtain pi^'s
est_models<-function(y, x, t=31) {
  pi.est<-y/6
  mod.fit<-glm(formula = cbind(y,6-y) ~ x, family = binomial(link=logit))
  exp(coef(mod.fit)['(Intercept)'] + coef(mod.fit)['x']*t)/
    (1 + exp(coef(mod.fit)['(Intercept)'] + coef(mod.fit)['x']*t))
}

t31<-sapply(1:num.simulations, function(i) est_models(y.mat[,i], sample_temp[,i], t=31))
t72<-sapply(1:num.simulations, function(i) est_models(y.mat[,i], sample_temp[,i], t=72))


paste("The 90% confidence interval for Temperature at 31F is: ",
      round(quantile(t31, 0.05),2),"to",round(quantile(t31,0.95),2))
```

```
## [1] "The 90% confidence interval for Temperature at 31F is:  0.14 to 0.99"
```

```
paste("The 90% confidence interval for Temperature at 72F is: ",
      round(quantile(t72, 0.05),2),"to",round(quantile(t72,0.95),2))
```

```
## [1] "The 90% confidence interval for Temperature at 72F is:  0.01 to 0.07"
```

f. Evaluating importance of quadratic temperature term in model

```
mod.glm.binom.nopressure.quad.temp<- glm(pi ~ Temp + I(Temp^2), family = binomial(link=logit),
                                          data = challenger_df, weights=Number)
anova(mod.glm.binom.nopressure, mod.glm.binom.nopressure.quad.temp, test='Chisq')
```

```
## Analysis of Deviance Table
##
## Model 1: pi ~ Temp
## Model 2: pi ~ 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
```

The lack of statistical evidence from the p-value and the small difference in the residual deviance for the two models tells me that a quadratic term of Temperature is not needed for this model.

Part 4. Estimating linear regression and assessing assumptions needed for its validity

```
mod.lm <- lm(atleast_one_failure ~ Temp, data = challenger_df)
summary(mod.lm)
```

```
##
## Call:
## lm(formula = atleast_one_failure ~ Temp, data = challenger_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.43762 -0.30679 -0.06381  0.17452  0.89881
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.90476    0.84208   3.450  0.00240 **
## Temp        -0.03738    0.01205  -3.103  0.00538 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3987 on 21 degrees of freedom
## Multiple R-squared:  0.3144, Adjusted R-squared:  0.2818
## F-statistic:  9.63 on 1 and 21 DF,  p-value: 0.005383
```

To be able to use linear regression to model a situation, we need to satisfy some assumptions. Below are my evaluations for each of those assumptions for this case:

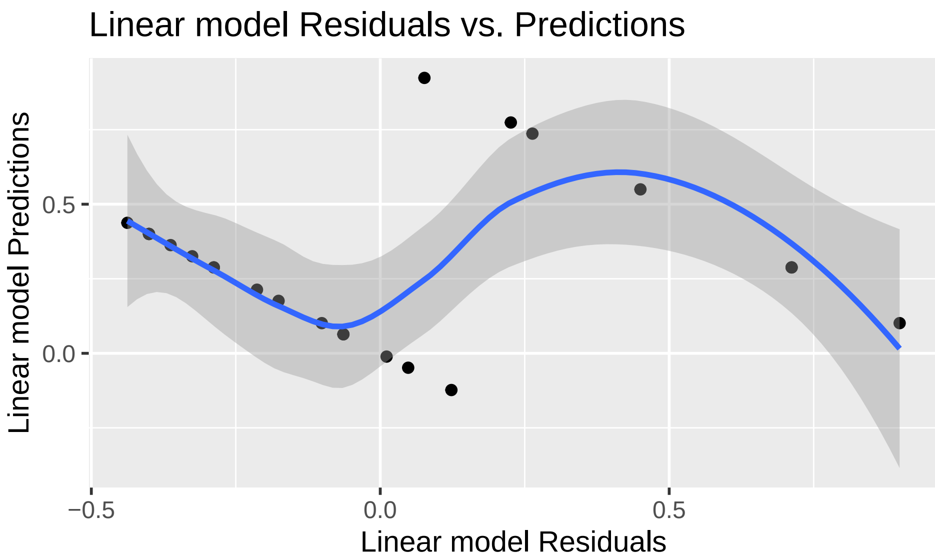
1. Independent and Identically distributed trials -

a. Identically distributed - For this assumption to be satisfied, all flights must be launched the same way, which in the practical scenario (like it was for the 23 flights in our dataset) is questionable. Therefore, I cannot say that this assumption is satisfied.
b. There is noted in the paper that parts of older launches have been reused for later ones so O.rings that were slightly deteriorated but did not fail earlier, if used for other flights, could inform us about the probability of failure of at least one O.ring for that flight. Thus, there are problems this assumption.

2. Linear Conditional Expectation - to evaluate this assumption, we will build a residuals vs predictions plot, and judge it by occular test. Since the plot below does not remain close to 0, we can assess that Linear Conditional Expectation assumption is not satisfied.

```
challenger_df %>% ggplot(aes(x=resid(mod.lm), y=predict(mod.lm)))+
  geom_point() + stat_smooth(se=TRUE) +
  labs(title = "Linear model Residuals vs. Predictions",
       x = "Linear model Residuals",
       y = "Linear model Predictions")
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



3. No Perfect Collinearity - Since we are using only one variable, this assumption can be satisfied because there is no other variable to evaluate correlation with.

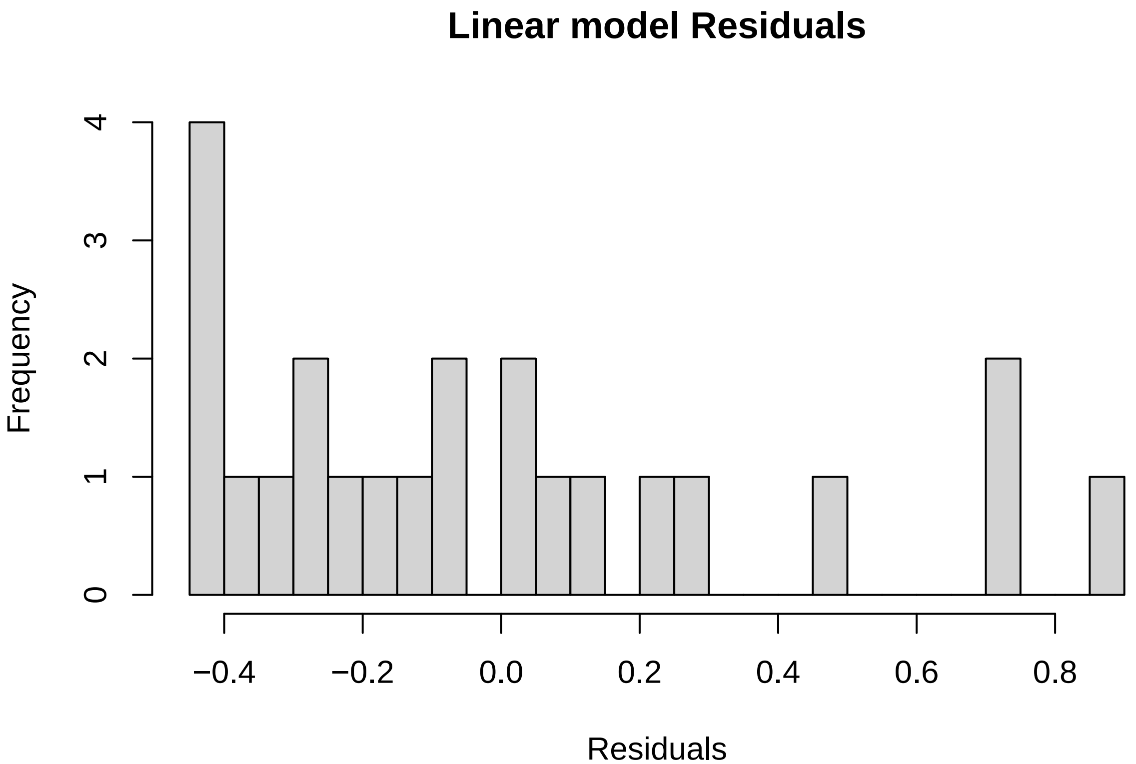
4. Homoskedastic Conditional Errors - We will use the Breush Pagan test to test the hypothesis that the variance is constant. Since the p-value of below is not significant to reject the null hypothesis, this assumption is satisfied.

```
lmtest::bptest(mod.lm)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  mod.lm
## BP = 0.0032704, df = 1, p-value = 0.9544
```

5. Normality of Errors - We will use a qqplot to see if the model's residuals have a normal distribution. By an occular judgement, we can see that the residuals are not normally distributed and therefore, this assumption is not met.

```
hist(mod.lm$residuals, breaks=23, main = "Linear model Residuals", xlab = "Residuals")
```



Not all assumptions for the Linear Regression model are met. The Binary Logistic Regression model needs fewer assumptions to be satisfied (Independence of Trials and No Multicollinearity). While the Independence of trials has problems, the no multicollinearity assumption is satisfied. Additionally, linear regression does not always bound the predicted outputs in the bounds that make sense (for example, linear regression model predicted probabilities can go under 0 and over 1). For these reasons, I would use Binary Logistic Regression to model whether there is at least one failure in the O.rings of a flight or not.

Part 5. Conclusion - Interpretation of model results

There are two variables that can be modeled and we have done so with two models: 1. The Binomial Logistic Regression which models the probability of O.ring failures, 2. The Binary Logistic Regression which models the probability of failure of at least one O.ring from all in the flight.

While, assumptions for the models have fewer problems for the Binary Logistic Regression than the Binomial Logistic Regression, I think the Binomial Logistic Regression is more useful since it can tell us how many O.rings are expected to fail and that is more granular detail for decision making of launch or no launch. We use only the 'Temp' variable as explanatory variable in the models. The temperature at flight launch is both statistically and practically significant.

The odds of seeing an O.ring failure in a flight launch are 55% less for every 7F increase in temperature at launch. At the temperature of the 1986 Challenger launch, 31F, the probability of failure of an O.ring was 82% which meant that ~5 rings out of the 6 were expected to fail in that launch. The odds of failure were

$$e^{5.08498-0.11560*31} = 4.487703$$

which indicates that the probability of an O.ring failing was 4.5 times the probability of it not failing. If statistical analysis augmented their decision, they would have probably chosen to not launch that morning and instead wait for a higher temperature. It could have possibly saved a launch that was highly probable to fail at that temperature.