Statistical Methods for Discrete Response, Time Series, and Panel Data (W271): Lab 1

Daniel Van Lunen,Aaron Olson

Introduction

A post-disaster investigation revealed the Challenger explosion was caused by O.ring failures. Before the launch, there was debate about whether to launch due to low temperatures. In this lab, we model the probability of O.ring failure using data from 23 launch tests to investigate the causes of failure and whether a launch at the 31 degree temperature was especially risky. Our findings indicate temperature has a significant impact on O.ring failure, while pressure does not. Our final logistic regression model is

$$log(\frac{P(O.ring\ Failure)}{P(O.ring\ doesn't\ fail)}) = 5.08498 - 0.11560 \times Temp$$

This yields an estimated 81.2% chance of O.ring failure with 90% bootstrapped confidence between 12.5% and 99.3% at the 31 temperature on the day of Challenger loss. This risk could have been avoided by waiting for a warmer day where we predict a much lower chance of failure. For example, at 72 degrees we predict the chance of failure between 1.0% and 6.9% degrees with 90% confidence

EDA

```
d <- read.csv('challenger.csv')
d %>% glimpse()
```

```
## Observations: 23
## Variables: 5
## $ Flight   <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16...
## $ Temp     <int> 66, 70, 69, 68, 67, 72, 73, 70, 57, 63, 70, 78, 67, 5...
## $ Pressure <int> 50, 50, 50, 50, 50, 50, 100, 100, 200, 200, 200, 200,...
## $ O.ring   <int> 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 2, 0, 0, 0, 0,...
## $ Number   <int> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6,...
```

```
summary(d)
```

```
##      Flight         Temp          Pressure         O.ring
##  Min.   : 1.0   Min.   :53.00   Min.   : 50.0   Min.   :0.0000
##  1st Qu.: 6.5   1st Qu.:67.00   1st Qu.: 75.0   1st Qu.:0.0000
##  Median :12.0   Median :70.00   Median :200.0   Median :0.0000
##  Mean   :12.0   Mean   :69.57   Mean   :152.2   Mean   :0.3913
##  3rd Qu.:17.5   3rd Qu.:75.00   3rd Qu.:200.0   3rd Qu.:1.0000
##  Max.   :23.0   Max.   :81.00   Max.   :200.0   Max.   :2.0000
##      Number
##  Min.   :6
```

```
##  1st Qu.:6
##  Median :6
##  Mean   :6
##  3rd Qu.:6
##  Max.   :6
```

```r
describe(d)
```

```
## d
##
##  5  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
##
## Value         53    57    58    63    66    67    68    69    70    72
## Frequency      1     1     1     1     1     3     1     1     4     1
## Proportion 0.043 0.043 0.043 0.043 0.043 0.130 0.043 0.043 0.174 0.043
##
## Value         73    75    76    78    79    81
## Frequency      1     2     2     1     1     1
## Proportion 0.043 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
## ---------------------------------------------------------------------------
```

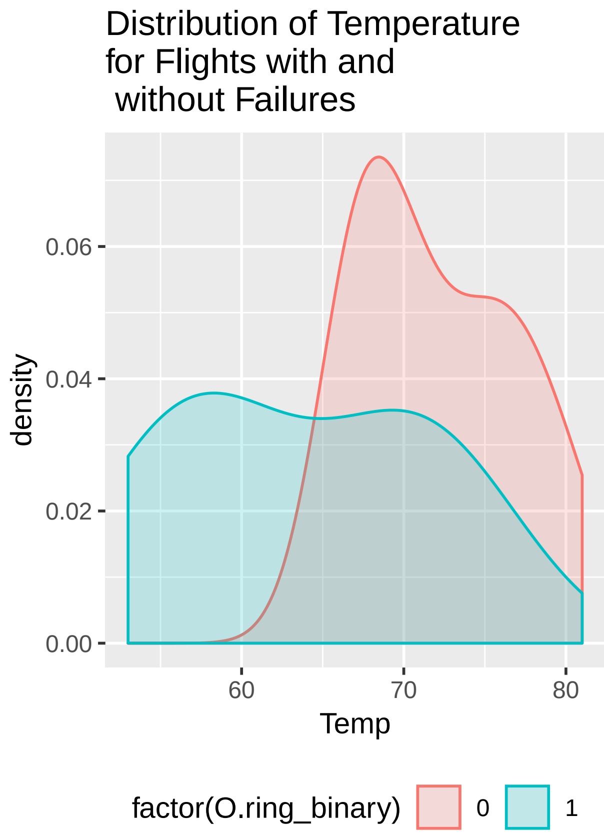
```
d$Pressure.factor <- factor(d$Pressure,levels = c(50,100,200))
d$O.ring_binary <- factor(ifelse(d$O.ring>0,1,0))
d$O.ring_factor <- factor(d$O.ring, levels = c(0,1,2))
```

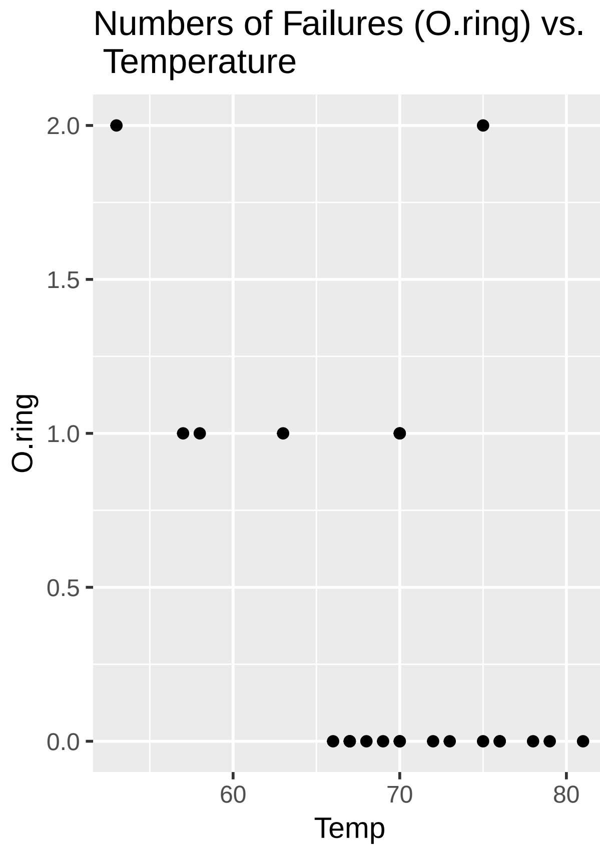
The dataset has 5 variables:

1. Flight appears to be an index for each flight and ranges from 1 to 23.
2. Temp is the Farenheit temperature on the day of the launch
3. Pressure is the psi level pressure used. It increases sequentially with the flight number as the paper mentioned they increased the pressure from 50 to 100 then 200 in their tests.
4. O.ring represents the number of O.ring failures (ranges between 0 and 2)
5. Number is always 6. This is the total number of O.rings.

There are no missing, top-coded, or bottom-coded values.

```
p1 <- ggplot(d, aes(Temp, fill = factor(O.ring_binary), colour = factor(O.ring_binary))) +
 geom_density(alpha=0.2) +
  ggtitle("Distribution of Temperature\nfor Flights with and\n without Failures") +
    theme(legend.position="bottom")
p2 <- ggplot(d, aes(Temp, O.ring)) +
 geom_point() +
  ggtitle("Numbers of Failures (O.ring) vs.\n Temperature")
grid.arrange(p1,p2,ncol=2)
```





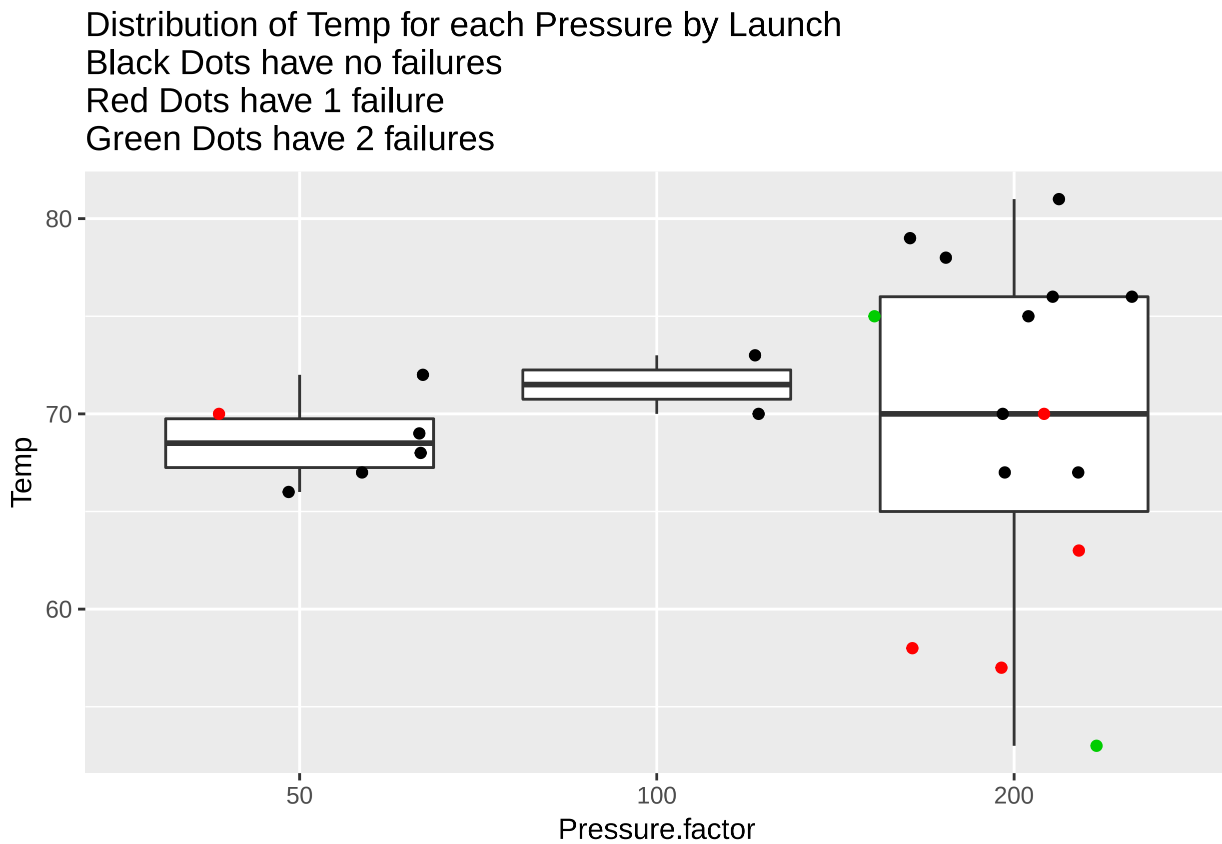
Temperature is skewed to the left with fewer observations for the colder temperatures we are most interested in. For launches with failures the temperature tends to be lower. There also appears to be one temperature that is far from the normal trend: 75 degrees with 2 failures. Given this seems completely possible due to randomness, we leave this point in the data. There is a sharp decline at approximately 65 degrees to nearly no failures above that temperature.

```
# how does the proportion of failures change with Pressure
d %>%
  group_by(Pressure) %>%
  summarise(pfailures=mean(O.ring/6))
```

```
## # A tibble: 3 x 2
##   Pressure pfailures
##      <int>     <dbl>
## 1       50    0.0278
## 2      100    0
## 3      200    0.0889
```

```
# how does temperature change with pressure
ggplot(d, aes(Pressure.factor, Temp)) +
  geom_boxplot() +
  geom_jitter(height = 0, colour=factor(d$O.ring+1)) +
  ggtitle(paste0("Distribution of Temp for each Pressure by Launch\n",
                 "Black Dots have no failures\n",
                 "Red Dots have 1 failure\n",
```

```
                 "Green Dots have 2 failures"))
```



Given only 3 values of pressure and only 2 data points at the value of 100, it is harder to judge the impact of pressure on O.ring failure. There are a higher proportion of failures at low pressures and high pressures so we might think this could be caused by failure to seal at low pressures and blow through at high pressures. However, it is possible it is also just explained by the temperature because the temperature on high and low pressure days tended to be lower. Thus, we will see if pressure still has an impact after accounting for temperature.

Modeling

4a Assumptions Underlying Logistic Model

It is necessary to assume that each O.ring is independent in the binomial model because if the probability that on O.ring fails is dependent on whether another O.ring fails, then the total number of failures will no longer follow a binomial distribution. The assumption is necessary so the authors can use each of the 6 O.rings in each of the 23 launches as a distinct datapoint in the logisitic regression. The problem is that if one O.ring fails, it likely increases the chance that additional O.rings on the same launch will fail because the conditions were "bad" enough to make one fail so they could make others fail as well.

Also detailed in the paper was the specifications for the field joint, whereby there was a primary and a backup/secondary o-ring. It is unlikely that the backup o-ring would be exposed to temperature or potential failure if the primary o-ring did not fail. This further supports an element of dependence amongst the o-ring variable. In this analysis however we consider each o-ring failure independently and assume one failure does not affect another.

4b Estimate Model

```r
m4b <- glm(formula = O.ring/Number ~ Temp + Pressure, data = d,
           family = binomial(link = "logit"), weights = Number)
summary(m4b)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp + Pressure, family = binomial(link = "logit"),
##     data = d, 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
```

We also estimate the model with a Pressure squared term because our EDA gives us the hypothesis that potentially middle-ground pressures are ok, but high and low are bad.

```r
m4b_psquared <- glm(formula = O.ring/Number ~ Temp + Pressure + I(Pressure^2),
                    data = d, family = binomial(link = "logit"), weights = Number)
summary(m4b_psquared)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp + Pressure + I(Pressure^2),
##     family = binomial(link = "logit"), data = d, weights = Number)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.04770  -0.67212  -0.56774  -0.00018   2.30791
```

```
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)     35.03744 6564.30727   0.005   0.9957
## Temp            -0.09558    0.04461  -2.143   0.0321 *
## Pressure        -0.80292  164.10766  -0.005   0.9961
## I(Pressure^2)    0.00324    0.65643   0.005   0.9961
## ---
## 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: 15.875  on 19  degrees of freedom
## AIC: 37.435
##
## Number of Fisher Scoring iterations: 17
```

And a model with a pressure and temperature interaction because it is possible that higher temperatures are worse at higher pressures because the rings are more malleable and prone to blow through.

```r
m4b_pt <- glm(formula = O.ring/Number ~ Temp + Pressure + Pressure:Temp,
              data = d, family = binomial(link = "logit"), weights = Number)
summary(m4b_pt)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp + Pressure + Pressure:Temp,
##     family = binomial(link = "logit"), data = d, weights = Number)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.0279  -0.6600  -0.5141  -0.1744   2.3724
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -21.765547  46.285607  -0.470    0.638
## Temp            0.252463   0.662989   0.381    0.703
## Pressure        0.130857   0.232643   0.562    0.574
## Temp:Pressure  -0.001769   0.003336  -0.530    0.596
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 24.230  on 22  degrees of freedom
## Residual deviance: 16.257  on 19  degrees of freedom
## AIC: 37.817
##
## Number of Fisher Scoring iterations: 6
```

And we estimate a model with just temperature for comparison.

```
m5a <- glm(formula = O.ring/Number ~ Temp, data = d,
           family = binomial(link = "logit"), weights = Number)
summary(m5a)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp, family = binomial(link = "logit"),
##     data = d, 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
```

4c LRTs for variable importance

```
# is temp or pressure signficant after accounting for the other?
Anova(m4b)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: O.ring/Number
##          LR Chisq Df Pr(>Chisq)
## Temp       5.1838  1     0.0228 *
## Pressure   1.5407  1     0.2145
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# is pressure and pressure^2 signficant after accounting for temp?
anova(m5a,m4b_psquared,test = "LR")
```

```
## Analysis of Deviance Table
##
```

```
## Model 1: O.ring/Number ~ Temp
## Model 2: O.ring/Number ~ Temp + Pressure + I(Pressure^2)
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     18.086
## 2        19     15.875  2   2.2112    0.331
```

```
# is there a significant interaction between temp and pressure
Anova(m4b_pt)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: O.ring/Number
##                LR Chisq Df Pr(>Chisq)
## Temp             5.1838  1     0.0228 *
## Pressure         1.5407  1     0.2145
## Temp:Pressure    0.2888  1     0.5910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The leave-one-out LR tests indicate that Temperature has a significant impact on the chance of an O.ring failure. On the other hand, after accounting for temperature, Pressure does not have a significant impact. Even when using a higher-order or interaction term, pressure does not have a significant impact after accounting for temperature.

4d Removing Pressure and Implications

The authors chose to remove pressure because they found it did not significantly reduce the residual deviance as the LRT above shows.[1] Removing this variable could potentially bias the coefficient on temperature. Temperature and pressure have a weak positive correlation and temperature has a negative correlation with failures. If pressure also has a negative impact on failures and is left out of the model, part of its impact will be absorbed as bias in the coefficient on temperature, making it seem like temperature is more important than it actually may be in causing a failure.

5a Estimate Model

```
m5a <- glm(formula = O.ring/Number ~ Temp, data = d,
           family = binomial(link = "logit"), weights = Number)
summary(m5a)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp, family = binomial(link = "logit"),
##     data = d, weights = Number)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.95227  -0.78299  -0.54117  -0.04379   2.65152
```

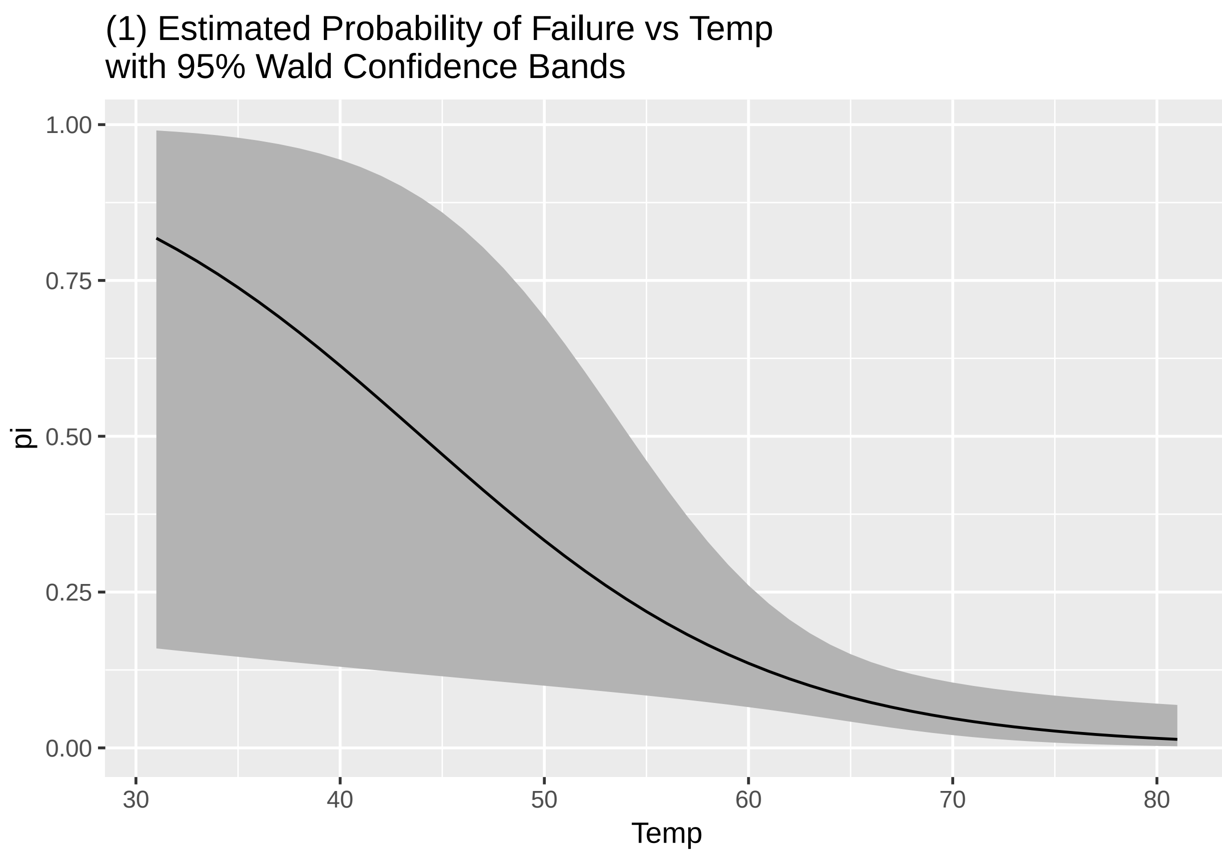
[1] they also performed bootstraped confidence intervals on the number of incidents holding pressure constant at 50/200 for different temperatures and found they overlapped greatly.

```
##
## 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
```

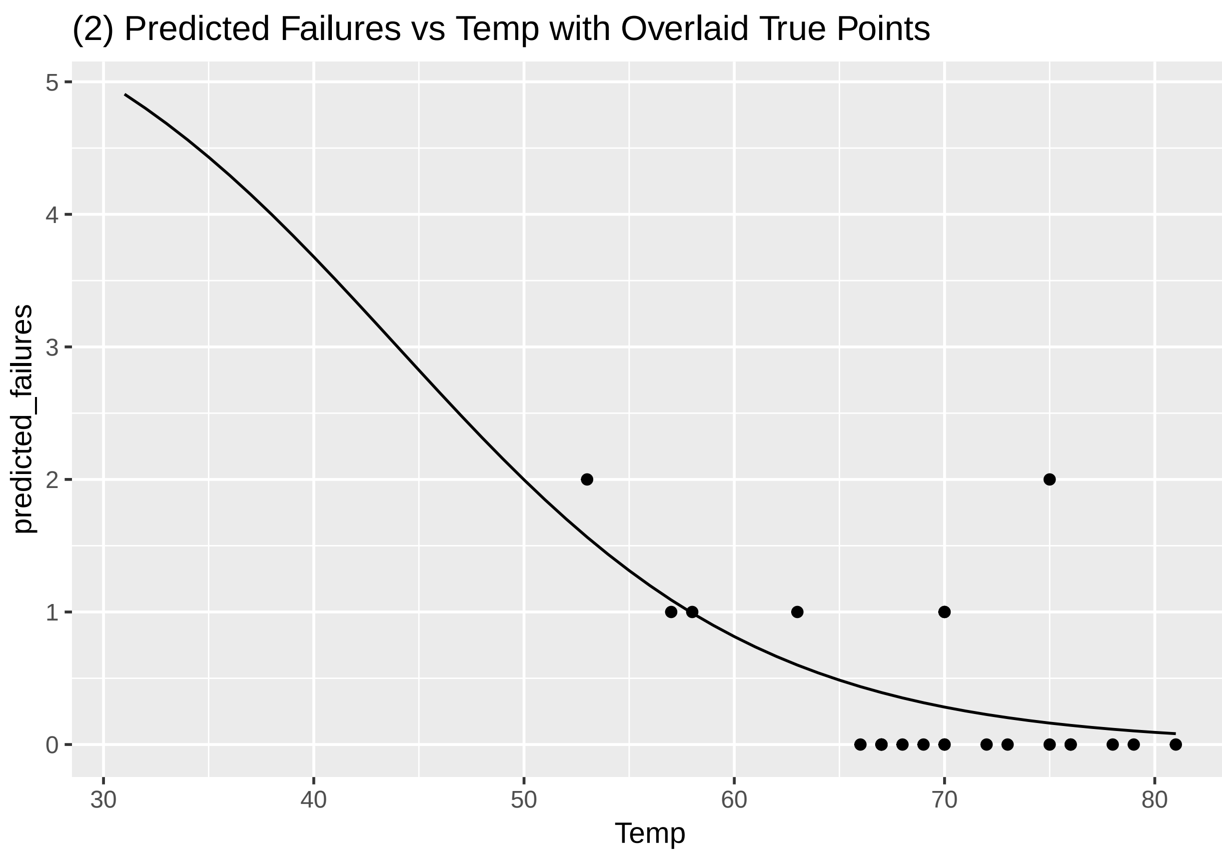
5b Pihat vs Temp and Expected failures vs Temp

Let's see how well the single covariate model lines up with the data by showing (1) the estimated probability of a failure vs. temperature and (2) the expected number of failures vs temperature in charts.

```r
alpha <- .05
predict.data <- data.frame(Temp=seq(31,81,by=1))
linear.pred <- predict(object = m5a,newdata = predict.data, type='link', se=T)
predict.data <- predict.data %>%
  mutate(
    z=linear.pred$fit,
    pi=exp(z)/(1+exp(z)),
    CI.upper.z= z + qnorm(1-alpha/2)*linear.pred$se.fit,
    CI.lower.z= z + qnorm(alpha/2)*linear.pred$se.fit,
    CI.upper.pi = exp(CI.upper.z)/(1+exp(CI.upper.z)),
    CI.lower.pi = exp(CI.lower.z)/(1+exp(CI.lower.z))
  )
predict.data %>%
  ggplot(aes(Temp)) +
  geom_ribbon(aes(ymin=CI.lower.pi,ymax=CI.upper.pi), fill="grey70") +
  geom_line(aes(y=pi)) +
  ggtitle(paste0("(1) Estimated Probability of Failure vs Temp\n",
                 "with 95% Wald Confidence Bands"))
```




```
data.frame(Temp <- seq(31,81,by=1)) %>%
  mutate(
    z=m5a$coefficients["(Intercept)"] + m5a$coefficients["Temp"]*Temp,
    predicted_failures=6*exp(z)/(1+exp(z))
    ) %>%
  ggplot(aes(Temp,predicted_failures)) + geom_line() +
  ggtitle("(2) Predicted Failures vs Temp with Overlaid True Points") +
  geom_point(data=d %>% mutate(predicted_failures=O.ring))
```



5c 95% Wald confidence bands

The confidence bands on *pi* are wider for lower temperatures because we do not have any data points at those low data points. The more we have to extrapolate, the less confident we will be (there is larger variance in predictions with fewer training points).

5d Confidence interval at 31 degrees and assumptions

```r
# Wald
predict.data %>% filter(Temp==31) %>% select(Temp,pi,CI.lower.pi,CI.upper.pi)
```

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

```r
# LR profile
z <- mcprofile(object = m5a, CM = matrix(data = c(1, 31), nrow=1, ncol=2))
z.CI <- confint(object = z,level = .95)
(pi.CI.LR.profile <- exp(z.CI$confint)/(1+exp(z.CI$confint)))
```

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

The estimated probability of a failure at this temperature is about 81.8%. The LR profile confidence interval for the probability of failure of a given O.ring is between approximately 14.2% and 99.1%.

This is in rough agreement with the Wald confidence interval. The assumptions are 1. All O.ring failure probabilities are identical given temperature. 2. O.ring failures are independent of each other (i.e we trained on a random sample of O.ring tests). This is questionable because O.rings are clustered by launch (6 always happen in the same conditions). 3. The log-odds of an O.ring failure is linearly related to the Temperature. 4. After holding temperature constant there are no other variables that affect the probability of failure. This exogeneity assumption seems tenuous, but we do not have any other variables besides pressure to place in the regression. 5. The sample is large enough that the estimated linear combination of coefficients is normal (for Wald) or $-2log(LR) \sim \chi^2$ (for the profile interval). 6. The model has converged to the true values of the parameters. 7. There is no perfect colinearity of the independent variables (here this is met because we only have Temp as an indpendent variable).

5e Parametric Bootstrap

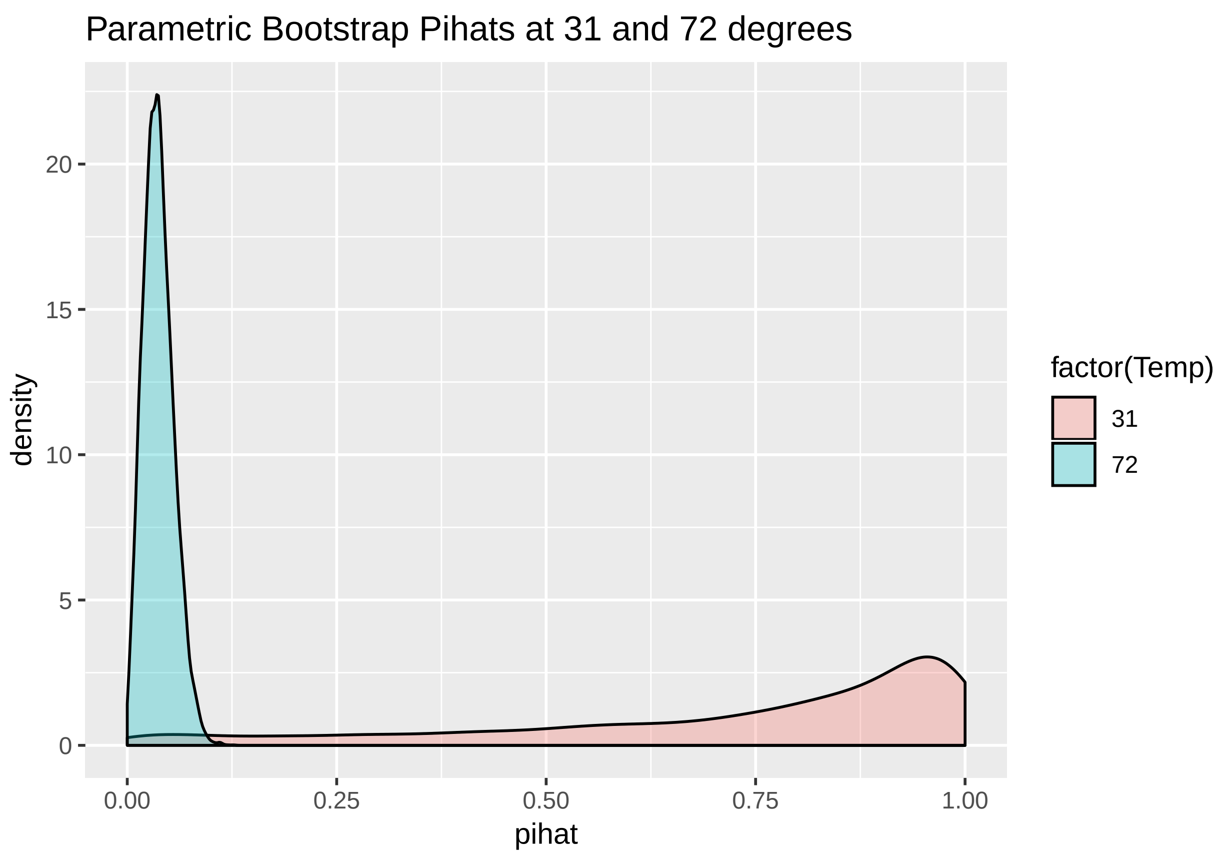
```
generate_pihats <- function(logisticmodel=m5a, data=d,
                            temps=data.frame(Temp=c(31,72))){
  # 1. Random sample of temperatures size nrow(d) from d with replacement
  bootstrap_df <- data.frame(Temp=sample(data$Temp, size = nrow(data), replace = T),
                             Number=6)
  # 2. Use model to estimate pi at each of those temperatures
  bootstrap_df$model.pi.hat <- predict(object = logisticmodel,
                                       newdata = bootstrap_df, type='response')
  # 3. Using those estimated pis, make random outcomes from binom(6, pi)
  bootstrap_df$O.ring <- rbinom(nrow(data), 6, bootstrap_df$model.pi.hat)
  # 4. Estimate a new logistic regression with those random outcomes and temperatures
  bootstrap_m <- glm(formula = O.ring/Number ~ Temp, data = bootstrap_df,
                     family = binomial(link = "logit"), weights = Number)
  # 5. Compute estimated pis at 31 and 72 degrees
  cbind(temps,pihat=predict(object = bootstrap_m, newdata = temps, type='response'))

}
# 6. Replicate many times
set.seed(42)
bootstrapped_pis <- do.call("rbind", replicate(10000, generate_pihats(),simplify = F))
# 7. Take 5th and 95th percentile of estimates for 31,72 as confidence interval
bootstrapped_pis %>%
  group_by(Temp) %>%
  summarise(CI.lower=quantile(pihat, probs=.05),
            CI.upper=quantile(pihat, probs=.95))
```

```
## # A tibble: 2 x 3
##    Temp CI.lower CI.upper
##   <dbl>    <dbl>    <dbl>
## 1    31   0.112    0.993
## 2    72   0.0106   0.0694
```

```
# density plots to visualize
bootstrapped_pis %>%
```

```r
  ggplot(aes(x=pihat,fill=factor(Temp))) +
  geom_density(alpha=.3) +
  ggtitle('Parametric Bootstrap Pihats at 31 and 72 degrees')
```

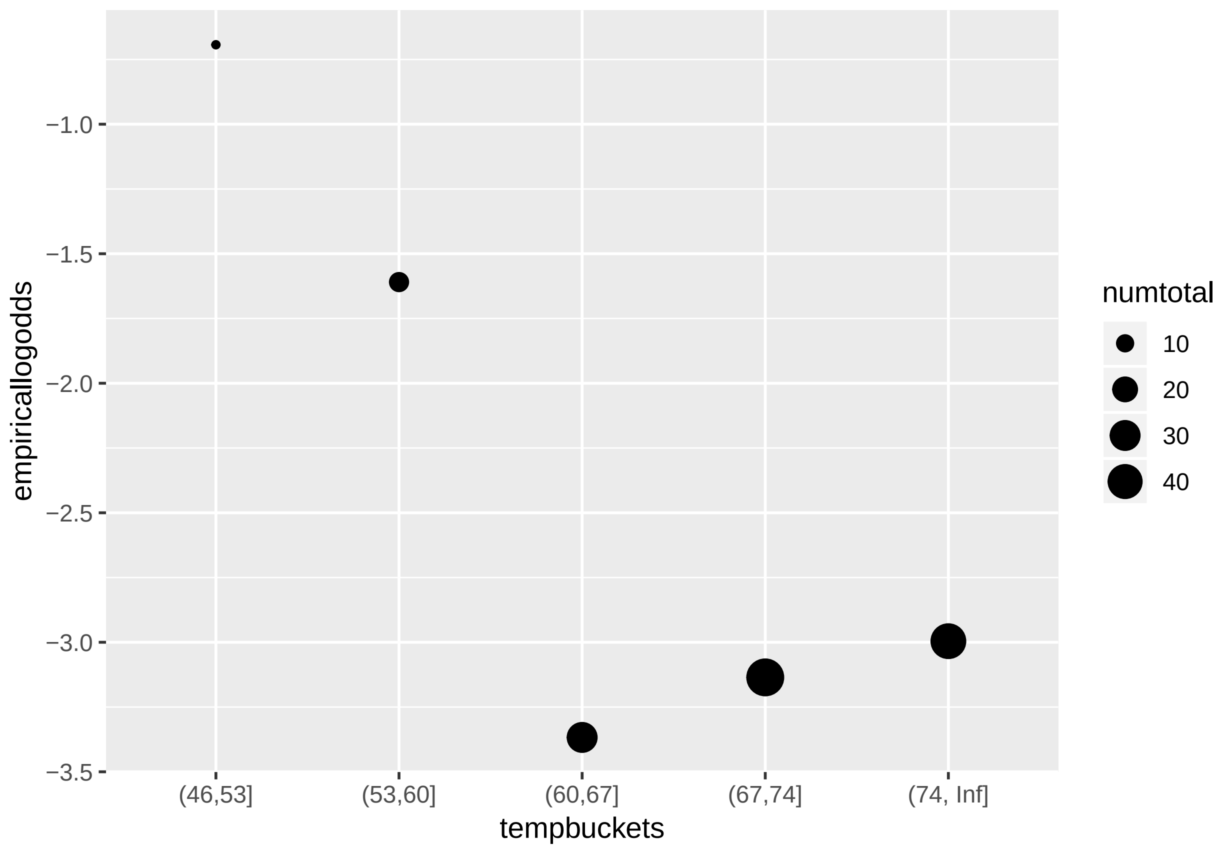


Our parametric bootstraps indicate that we estimate the probability of failure for a given O.ring to bet between about 12.5% and 99.3% at 31 degrees and 1.0% and 6.9% at 72 degrees. Again the interval is much wider for lower temperatures where we have fewer data points.

Note in our bootstrap sometimes we get complete separation, but this just makes our estimates 0 or 1 for the predicted probabilities which drop out after using the confidence interval.

5f is a quadratic term necessary?

```r
# empirical proportion of failures for buckets of temperatures
d$tempbuckets <- cut(d$Temp, breaks=c(-Inf,46,53,60,67,74,Inf))
d %>% group_by(tempbuckets) %>%
  summarise(empiricallogodds=log(mean(O.ring/6)/(1-mean(O.ring/6))),
            numtotal=sum(Number)) %>%
  ggplot(aes(x=tempbuckets,y=empiricallogodds,size=numtotal)) +
  geom_point()
```



Plotting the empircal log odds for different buckets reveals that above a certain temperature, failures seem quite rare. There does not seem to be enough data to confirm if the relationship is quadratic from a plot like this. Let's instead use a more general case LRT to see if the term on the squared temperature is significant.

```
# estimate new model with quadratic term
m5f <- glm(formula = O.ring/Number ~ Temp + I(Temp^2), data = d, family = binomial(link = "log
Anova(m5f, test.statistic = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: O.ring/Number
##           LR Chisq Df Pr(>Chisq)
## Temp       0.71878  1     0.3965
## I(Temp^2)  0.49470  1     0.4818
```

Interpret final model

After seeing that pressure, a higher order pressure term, a temperature-pressure interaction term, and a higher order temperature term all are not close to signficant, we arrive at the same final model as the authors of the paper.

```
summary(m5a)
```

```
##
## Call:
## glm(formula = O.ring/Number ~ Temp, family = binomial(link = "logit"),
##     data = d, 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
```

```
exp(-5*m5a$coefficients["Temp"]) # OR multiplier for 5 degree decrease
```

```
##    Temp
## 1.78248
```

Here for every 5 degree decrease in temperature, the odds of a failure for a given O.ring increase by 78.2%. The estimated probability of failure can be seen above in the chart for 5b. Noteably, the higher temperatures have a lower probability of failure even when accounting for the confidence bands.

What about a Linear Probability Model?

Estimate the model

```
failures <- c()
temps <- c()
for (i in 1:nrow(d)){
  failures <- c(failures,rep(1,d$O.ring[i]),rep(0,6-d$O.ring[i]))
  temps <- c(temps,rep(d$Temp[i],6))
}
d_LPM <- data.frame(failures=failures,Temp=temps)
LPM <- lm(failures~Temp,data = d_LPM)
summary(LPM)
```

```
##
```

```
## Call:
## lm(formula = failures ~ Temp, data = d_LPM)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.19647 -0.08554 -0.06177 -0.01423  0.97784
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.616402   0.209118   2.948  0.00377 **
## Temp        -0.007923   0.002991  -2.649  0.00904 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2425 on 136 degrees of freedom
## Multiple R-squared:  0.04905,    Adjusted R-squared:  0.04206
## F-statistic: 7.016 on 1 and 136 DF,  p-value: 0.009036
```

This linear probability model estimates that the probability of an O.ring failure decreases linearly with temperature. Every 5 degree decrease in temperature would increase the estimated probability of failure by 0.79% (absolute not relative change). Let's also check the linear model's assumptions.

Assumptions

Linear model, Random Sample, No Perfect Collinearity

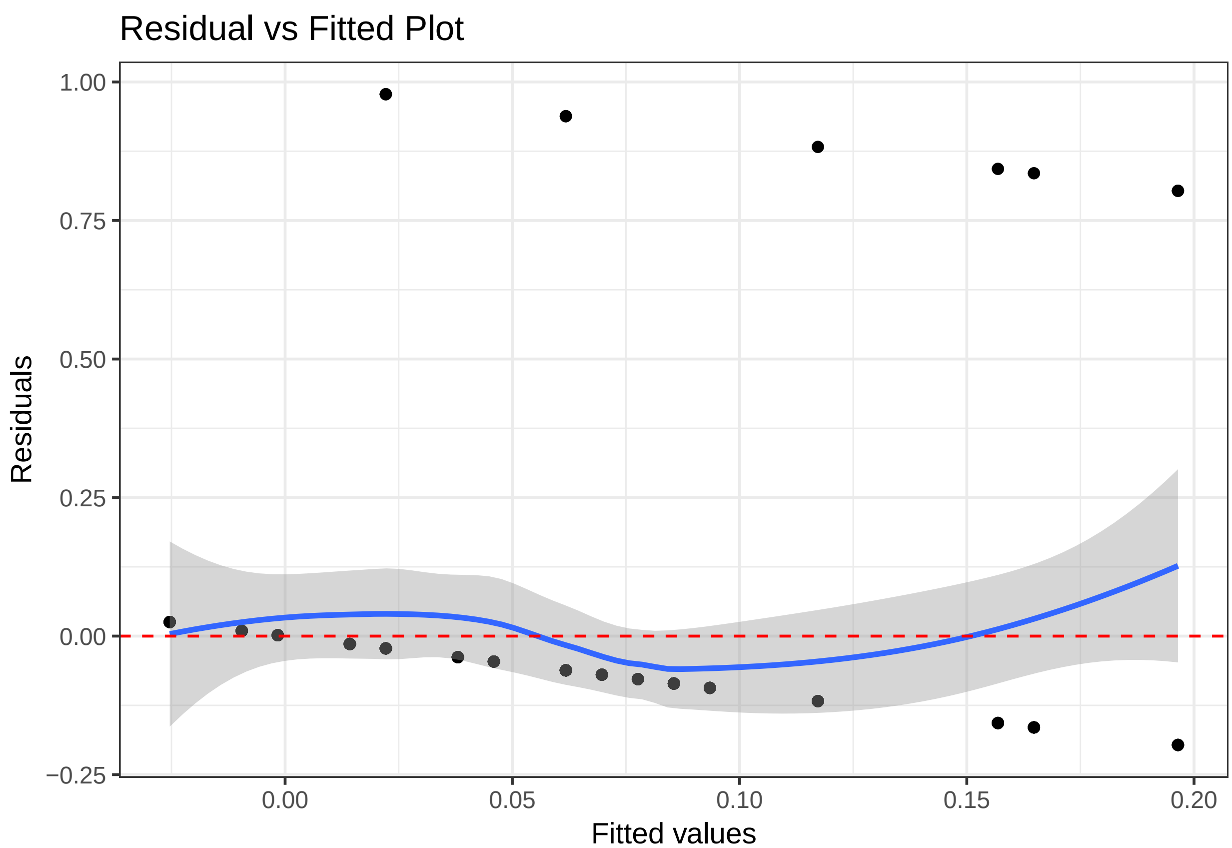
The first assumption is that the outcome is a linear function of the predictors. Though we are not placing any conditions yet on the error term, so this assumption is met, we can already see that a linear model may not be right because under this model, at temperatures above about 77.8 degrees, we predict a negative probability of failure, which does not make sense (especially given our data has data points here).

The linear model assumes a random sample, which is likely incorrect here for the same reasons mentioned in the logistic regression that outcomes on a given launch are likely correlated.

The no perfect collinearity assumption is met.

Zero conditional mean

```
ggplot(LPM, aes(.fitted, .resid))+geom_point() +
    stat_smooth(method="loess")+geom_hline(yintercept=0, col="red", linetype="dashed")+
    xlab("Fitted values")+ylab("Residuals") +
    ggtitle("Residual vs Fitted Plot")+theme_bw()
```

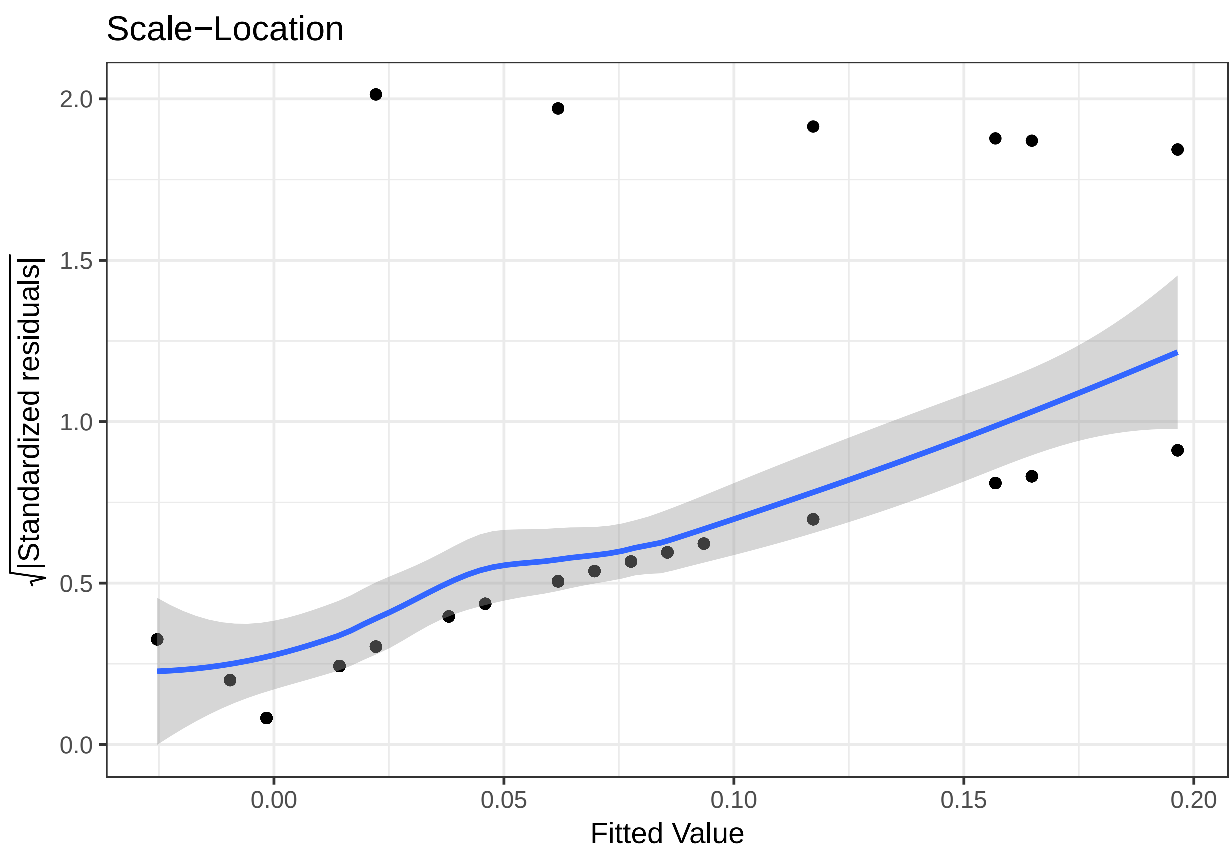


The residuals vs fitted value plot has confidence bands that contain zero throughout. However, there is an obvious, linearly decreasing pattern for two sets of points in the plot indicating there is some association between the error term and the fitted values. Therefore the temperature coefficient is likely biased.

Homoskedasticity

To get standard errors on the coefficients, we need a constant error variance.

```
ggplot(LPM, aes(.fitted, sqrt(abs(.stdresid))))+geom_point(na.rm=TRUE)+
    stat_smooth(method="loess", na.rm = TRUE)+xlab("Fitted Value")+
    ylab(expression(sqrt("|Standardized residuals|")))+
    ggtitle("Scale-Location")+theme_bw()
```

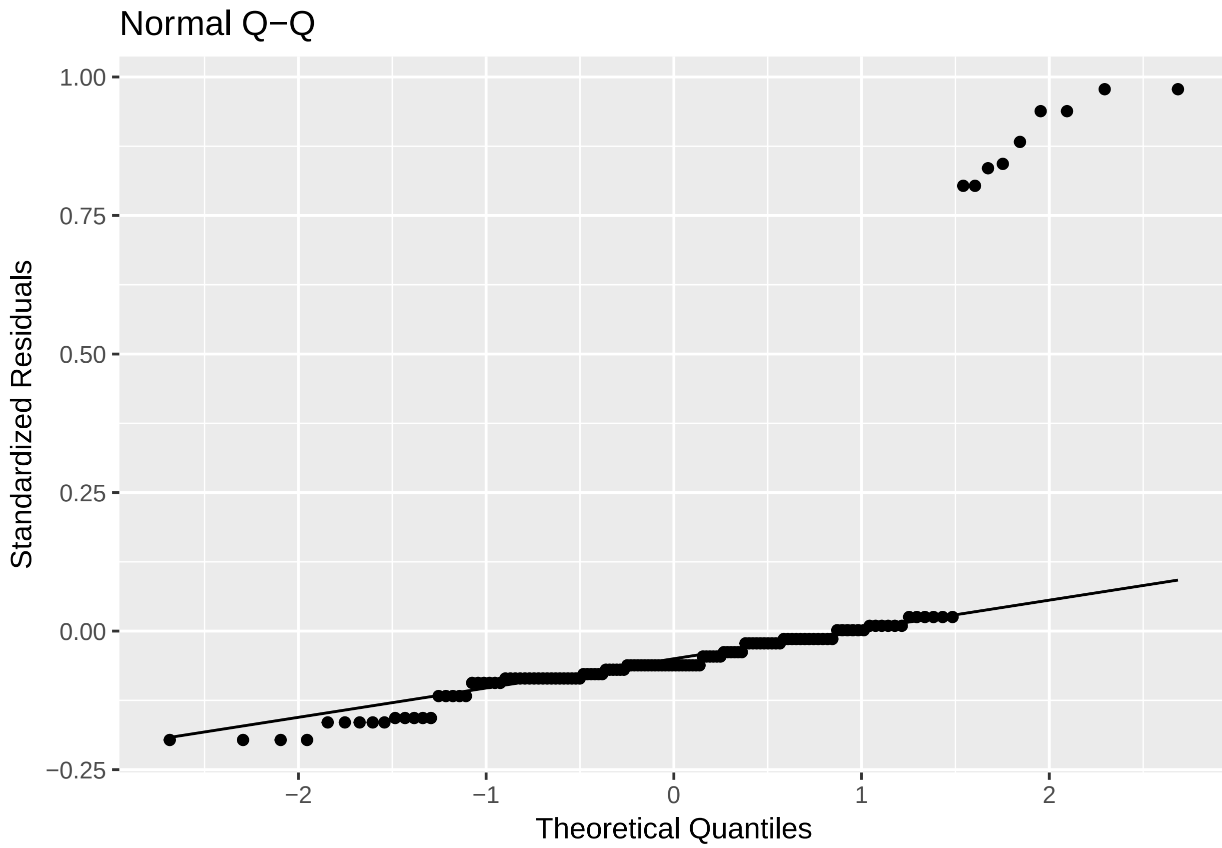


The Scale-Location plot reveals that the variance of the error term is clearly not constant: there is an upward trend in the variance of the residuals as the fitted values increase. Though we could use heteroskedastic-robust errors, the failure of the other assumptions makes the LPM a rough choice.

Normally Distributed Error

To conduct hypothesis t-tests for the LPM, we need normally distributed errors.

```
df <- data.frame(residuals=LPM$residuals)
ggplot(df, aes(sample=residuals)) +
    stat_qq() + stat_qq_line() +
    xlab("Theoretical Quantiles") + ylab("Standardized Residuals") +
    ggtitle("Normal Q-Q")
```



The residuals deviate significantly from the normal distribution.

Linear or Logistic?

A logistic model is much better because it keeps the estimated probabilities within $(0, 1)$ and doesn't require as many assumptions that are clearly failing with the linear model.

Conclusion

Decision makers at NASA would have done well to have statistically modelled the probability of O.ring failure before the Challenger launch. In this lab, we found that there was a strong relationship between O.ring failure and temperature. Our find model was

$$log(\frac{P(O.ring\ Failure)}{P(O.ring\ doesn't\ fail)}) = 5.08498 - 0.11560 \times Temp$$

This model indicates that the Challenger launch should have been delayed due to the high estimated probability of failure of the O.rings (between 12.5% and 99.3% chance of failure at the 31 temperature on chosen launch day with 90% confidence). This high estimated probability of failure could have convinced decision makers to wait until the temperature was higher. For example, until the temperature was 72 degrees where the predicted chance of failure is between 1.0% and 6.9% with 90% confidence. This could have potentially averted disaster.