

Investigation of the 1986 Space Shuttle Challenger Accident

W271 Group Lab 1

Michael Bollig | Amber Chen | Jeffrey Day

September 20, 2020

Contents

1. Introduction	2
1.1 Background	2
1.2 Exploratory Data Analysis (EDA)	2
2. Model Building Process	6
2.1 Assumption of Independence of O-rings	6
2.2 Logistic Regression Model	6
2.3 Likelihood Ratio Test (LRT)	8
2.4 Exclusion of Pressure Variable	8
3. Binary Logistic Regression Model	8
3.1 Model Estimation	8
3.2 O-ring Failure Predictions	9
3.3 Bootstrapped Confidence Intervals	12
3.4 Examination of Quadratic Term	13
4. Linear Regression Model	14
4.1 Model estimation	14
4.2 Classical Linear Model (CLM) Assumptions	16
4.3 Linear vs. Binary logistic regression	19
5. Implications of the Model	20

1. Introduction

1.1 Background

On January 28, 1986, the NASA space shuttle *Challenger* broke apart shortly after launching, due to a catastrophic O-ring failure attributed to an abnormally low temperature at launch of 31° F¹. In this Lab exercise, we follow, discuss, and reproduce elements of an investigative report on this event conducted by Dalal et al².

1.2 Exploratory Data Analysis (EDA)

Typically, when conducting an EDA the data must be summarized, however, there are actually few enough samples that displaying the entire data set is reasonable (Table 1).

```
knitr::kable(x = dat,
             format = "latex", longtable = T, escape = F,
             caption = "Full Challenger Data Set") %>%
  kable_styling(latex_options = c("striped", "hold_position", "repeat_header"),
              full_width = F)
```

Table 1: Full Challenger Data Set

Flight	Temp	Pressure	O.ring	Number
1	66	50	0	6
2	70	50	1	6
3	69	50	0	6
4	68	50	0	6
5	67	50	0	6
6	72	50	0	6
7	73	100	0	6
8	70	100	0	6
9	57	200	1	6
10	63	200	1	6
11	70	200	1	6
12	78	200	0	6
13	67	200	0	6
14	53	200	2	6
15	67	200	0	6
16	75	200	0	6
17	70	200	0	6
18	81	200	0	6
19	76	200	0	6
20	79	200	0	6
21	75	200	2	6
22	76	200	0	6
23	58	200	1	6

Additionally, we may use the *skim* function from the *skimr* package to gain a quick intuition about the variables included in this dataset.

¹https://www.nasa.gov/mission_pages/shuttle/shuttlemissions/archives/sts-51L.html

²Dalal, S., Fowlkes, E., & Hoadley, B. (1989). Risk Analysis of the Space Shuttle: Pre-Challenger Prediction of Failure. Journal of the American Statistical Association, 84(408), 945-957. doi:10.2307/2290069

```
# skimming without charts, because the charts are made of font-incompatible unicode characters
skim_without_charts(dat) %>% kable(format = "latex", booktabs = T,
  caption = "Summary of variables") %>%
  kable_styling(latex_options = c("striped", "scale_down", "hold_position", "repeat_header"),
    full_width = F)
```

Table 2: Summary of variables

skim_type	skim_variable	n_missing	complete_rate	numeric.mean	numeric.sd	numeric.p0	numeric.p25	numeric.p50	numeric.p75	numeric.p100
numeric	Flight	0	1	12.0000000	6.7823300	1	6.5	12	17.5	23
numeric	Temp	0	1	69.5652174	7.0570795	53	67.0	70	75.0	81
numeric	Pressure	0	1	152.1739130	68.2213324	50	75.0	200	200.0	200
numeric	O.ring	0	1	0.3913043	0.6563764	0	0.0	0	1.0	2
numeric	Number	0	1	6.0000000	0.0000000	6	6.0	6	6.0	6

After reviewing both the full data set (Table 1) and summarized variables (Table 2), the key takeaways are as follows:

- We will be working with a very limited ($n = 23$) data set
- No variable in the data set is missing any value
- All variables are stored as integers including the response variable
- There are no apparent extreme outlying data points (considering the min/max and quantiles of each variable)
- Of the 5 variables included in the data set, there are only 2 with explanatory power: *Temperature* and *Pressure*. *Flight* is an identifier, *Number* is static, and *O.ring* is the dependent variable.

The next step in our EDA is to examine the distributions of key variables that we will be using to build our model. First, we created a distribution of the temperature data available prior to launch, with the actual temperature at launch shown on the visualization (Fig. 1). Notably, the temperature at launch appears to be well below the rest of the data available at the time of launch.

```
# create plot
ggplot(data = dat, aes(x=Temp)) +
  # add histogram
  geom_histogram(colour = "black",
    aes(fill = "b"),
    alpha=0.6,
    bins=60) +
  # add extra data point to show launch date temperature
  geom_histogram(data = data.frame(Temp = 31),
    colour = "black",
    aes(fill = "r"),
    alpha=0.6,
    bins=60) +
  # pretty formatting
  scale_fill_manual(name="Group",
    values=c("r" = "red", "b"="blue"),
    labels=c("b"="Data available before launch",
      "r"="Temperature on launch day")) +
  labs(title = "Distribution of Temperature Data",
    x = expression(paste("Temperature (", degree, "F)")),
    y = "Number of Samples") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"))
```

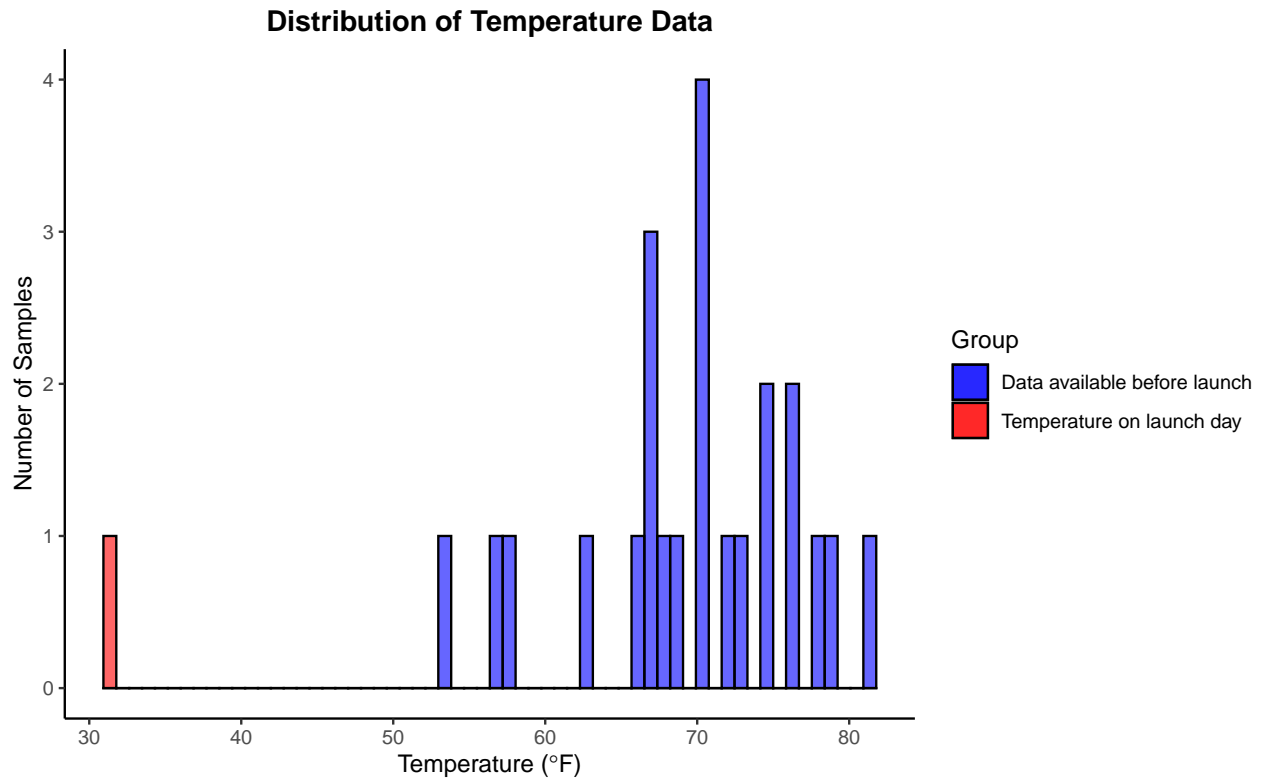


Figure 1: Temperature Range for Data Set

Knowing that the temperature at launch was abnormally low, we continue to investigate whether there is visual evidence of a relationship between temperature and the likelihood of O-ring failure. A bubble plot of temperature vs. O-ring failures shows that every flight launched below 67° F experienced at least 1 O-ring failure (Fig. 2). This suggests that the aforementioned relationship may indeed exist.

```
# create plot
ggplot(dat, aes(x=Temp, y=O.ring)) +
  # add points with count-based size
  geom_count(colour = "black", alpha=0.6, shape = 21, fill = "blue") +
  # scale and set breaks for point sizes
  scale_size_continuous(breaks = c(1,2), range=c(2,6)) +
  # pretty formatting
  labs(title = "Temperature vs. Number of O-rings With Thermal Distress",
       subtitle = "Size indicates number of overlapping events",
       x = expression(paste("Temperature (", degree, "F)")),
       y = "Number of O-rings With Thermal Distress") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))
```

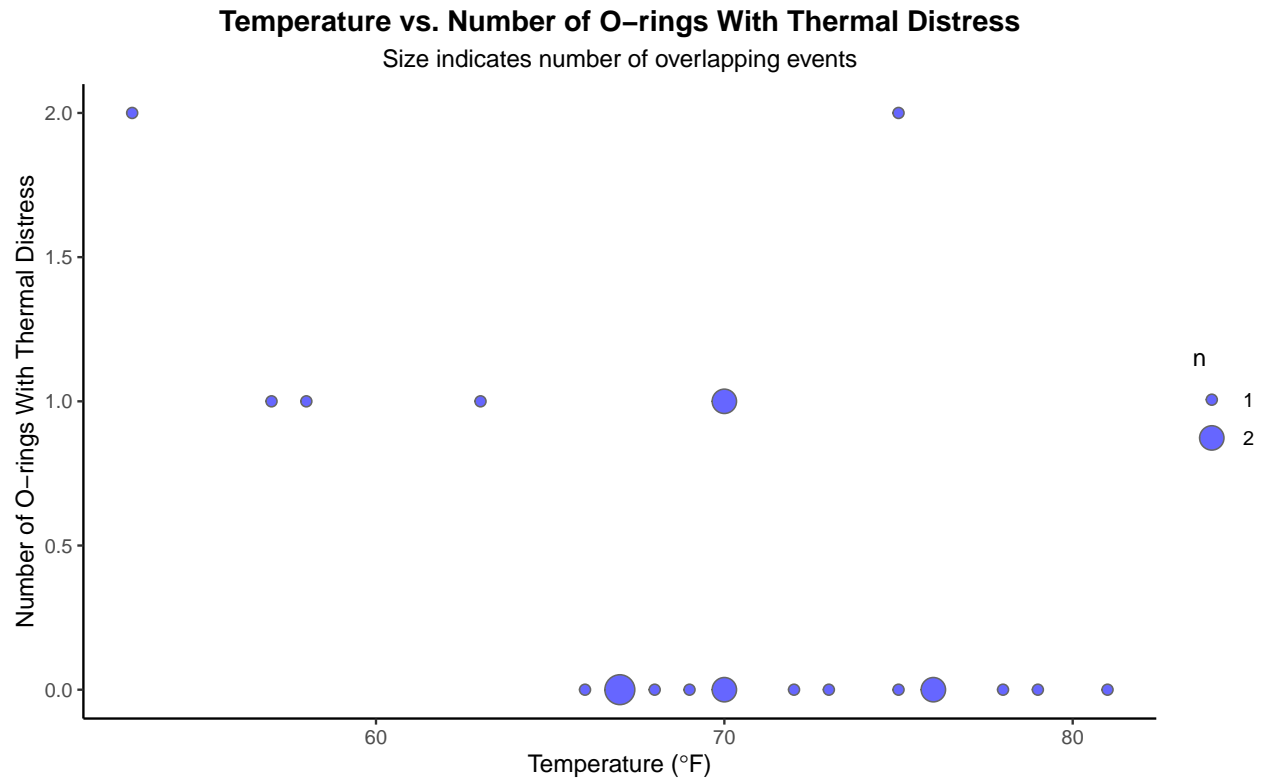


Figure 2: Temperature vs. O-ring Failures

Finally, we examine the only other explanatory variable available: *Pressure*. A box-and-whisker plot (Fig. 3) shows that only 1 O-ring failure occurred at a test pressure under 200 PSI. This indicates that pressure may have explanatory power in understanding whether an O-ring will fail.

```
# create plot
ggplot(dat, aes(x=as.factor(O.ring), y=Pressure)) +
  # add box plot
  geom_boxplot(outlier.shape = NA) +
  # add points, jittered due to overlap
  geom_jitter(width = 0.1, height = 0, colour = "black", alpha=0.6, shape = 21, size = 3, fill = "blue") +
  # pretty formatting
  labs(title = "Number of O-rings With Thermal Distress vs. Test Pressure",
       subtitle = "Data jittered to visualize overlapping points",
       x = "Number of O-rings With Thermal Distress",
       y = "Pressure (PSI)") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))
```


$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_K x_{iK} \quad (1)$$

Where π is the probability of an event, and x_1 through x_K are explanatory variables with respective coefficients β_1 through β_K . We first calculate the probability of O-ring failure for each flight test (Eq. 2), then express our first model in logit form (Eq. 3):

$$p.\text{failure} = \left(\frac{\text{O-rings with Thermal Distress}}{\text{Total O-rings on Flight}} \right) \quad (2)$$

$$\text{logit}(p.\text{failure}) = \beta_0 + \beta_1 \cdot \text{Temp} + \beta_2 \cdot \text{Pressure} \quad (3)$$

Our β coefficients are calculated using R's built-in *glm* function (Table 3).

```
# create variable indicating the probability of failure
dat$p.fail <- dat$O.ring / dat$Number

# fit model
modell1 <- glm(p.fail ~ Temp + Pressure,
              weights = Number,
              family = binomial,
              data = dat)

# summarize model
stargazer(modell1, type='latex', summary=F,
          dep.var.labels = c("Probability of O-ring Failures"),
          title = "Logistic Regression Model",
          header=F)
```

Table 3: Logistic Regression Model

	<i>Dependent variable:</i>
	Probability of O-ring Failures
Temp	−0.098** (0.045)
Pressure	0.008 (0.008)
Constant	2.520 (3.487)
Observations	23
Log Likelihood	−15.053
Akaike Inf. Crit.	36.106
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Plugging the coefficients back into Eq. 3, we arrive at our first model to estimate the probability of O-ring failure (Eq. 4):

$$\text{logit}(p.\hat{\text{failure}}) = 2.5202 + -0.0983 \cdot \text{Temp} + 0.0085 \cdot \text{Pressure} \quad (4)$$

2.3 Likelihood Ratio Test (LRT)

In our first model, the coefficient for *Pressure* was not found to be statistically significant in the model (Table 3). We use a LRT to judge the importance of each variable (Table 4):

```
# Likelihood Ratio Test
lrt <- Anova(model1, test="LR")
stargazer(lrt, type='latex', summary=F,
           title = "Likelihood Ratio Test (LRT)",
           header=F)
```

Table 4: Likelihood Ratio Test (LRT)

	LR Chisq	Df	Pr(>Chisq)
Temp	5.184	1	0.023
Pressure	1.541	1	0.215

- Note that the p-value for Temperature is significant at $p < 0.05$, while that for Pressure is not.

2.4 Exclusion of Pressure Variable

Based on the results of the model building process, there are two strong pieces of evidence that it is reasonable to exclude the *Pressure* variable from the model:

- Temperature is significant in the model ($p < 0.05$), while Pressure is not (Table 3)
- The LRT showed a low LR ($\Lambda = 1.541$) without significance for the Pressure variable against a Chi-squared distribution (Table 4)

While the evidence does suggest that for this data set *Pressure* does not add substantial value to the model, it is worth noting that temperature and pressure may not be independent in a real-world setting. Additionally, the lack of significance for *Pressure* may be partially attributed to the limited test sample set.

3. Binary Logistic Regression Model

3.1 Model Estimation

Considering the lack of significance of the Pressure variable, we simplify our model to only include temperature as an explanatory variable (Eq. 5).

$$\text{logit}(p.\text{failure}) = \beta_0 + \beta_1 \cdot \text{Temp} \quad (5)$$

Additionally, we consider a slight variation of this model, where we model the log-odds that *any* O-ring will fail (Eq. 6, Eq. 7):

$$\text{any.fail} = P(\text{O.ring} > 0) \quad (6)$$

$$\text{logit}(\text{any.fail}) = \beta_0 + \beta_1 \cdot \text{Temp} \quad (7)$$

Again, we calculate the coefficients for Eq. 5 and Eq. 7 in R (Table 5):


```

# create variable indicating the proportion of failures
dat$any.fail <- dat$O.ring > 0

# fit model
any_failure_model <- glm(any.fail ~ Temp,
                        family = binomial,
                        data = dat)

# fit model for expectation of number of failures
p_fail_model <- glm(p.fail ~ Temp,
                  weights = Number,
                  family = binomial,
                  data = dat)

# summarize models
stargazer(any_failure_model, p_fail_model, type='latex', summary=F,
          dep.var.labels = c("Probability of Any O-ring Failure", "Probability of O-ring Failure"),
          title = "Simplified Logistic Regression Model",
          header=F)

```

Table 5: Simplified Logistic Regression Model

	<i>Dependent variable:</i>	
	Probability of Any O-ring Failure	Probability of O-ring Failure
	(1)	(2)
Temp	−0.232** (0.108)	−0.116** (0.047)
Constant	15.043** (7.379)	5.085* (3.052)
Observations	23	23
Log Likelihood	−10.158	−15.823
Akaike Inf. Crit.	24.315	35.647

Note:

*p<0.1; **p<0.05; ***p<0.01

The coefficients for the two estimated models match those in the Dalal et al paper described as the “Binary Model” and the “Binomial Model”.

Note: a lower AIC for the Any-failure model shows a slightly better fit to the data points, however, it does not necessarily mean that this model will generalize better to new data. This is particularly important to consider when the data in question (31° F) is so far outside the range of the explanatory variable.

3.2 O-ring Failure Predictions

We will now use the two models in Table 5 to first estimate the probability of any O-ring failure, then estimate the number of O-rings expected to fail at 31° F. The logit function may be solved for π as follows:

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_K x_{iK})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_K x_{iK})} \quad (8)$$

Rewriting our Any-failure model (Eq. 7) in this form:

$$any.fail = \frac{\exp(\beta_0 + \beta_1 Temp)}{1 + \exp(\beta_0 + \beta_1 Temp)} \quad (9)$$

Finally, plugging in coefficients yields:

$$any.\hat{fail} = \frac{\exp(15.0429 + -0.2322 \cdot Temp)}{1 + \exp(15.0429 + -0.2322 \cdot Temp)} \quad (10)$$

By using this model, we are able to visualize the estimated probability of a flight experiencing at least one O-ring failure (Fig. 4)

```
# generate range of Temperatures
Temp <- seq(31, 81, 1)

# predict probability of at least 1 failure in linear form
pi <- predict(any_failure_model, list(Temp = Temp), type="link", se=T)

# calculate confidence intervals
alpha <- 0.05
pi$ci.upper = pi$fit + qnorm(1-alpha/2) * pi$se
pi$ci.lower = pi$fit + qnorm(alpha/2) * pi$se

# exponentiate out of logit form
pi$pi.hat <- exp(pi$fit)/(1+exp(pi$fit))
pi$pi.ci.upper = exp(pi$ci.upper)/(1+exp(pi$ci.upper))
pi$pi.ci.lower = exp(pi$ci.lower)/(1+exp(pi$ci.lower))

# plot
pi_df <- data.frame(Temp, pi)
ggplot(data = pi_df, aes(x=Temp, y=pi.hat)) +
  # add probability of any failure line
  geom_line() +
  # add confidence interval ribbon
  geom_ribbon(aes(ymin=pi.ci.lower,ymax=pi.ci.upper, fill = "b"), alpha=0.2) +
  # pretty formatting
  scale_fill_manual(name="",
                    values=("b" = "blue"),
                    labels=c("95% Wald CI")) +
  labs(title = "Temperature vs. Probability of Any O-ring Failure",
       subtitle = expression(hat(any.fail) == frac(e^{15.0429 - 0.2322 * Temp}, 1+e^{15.0429 - 0.2322
       x = expression(paste("Temperature (", degree, "F)")),
       y = expression(paste("Probability of Any O-ring Failure (", pi, ")")))) +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))
```

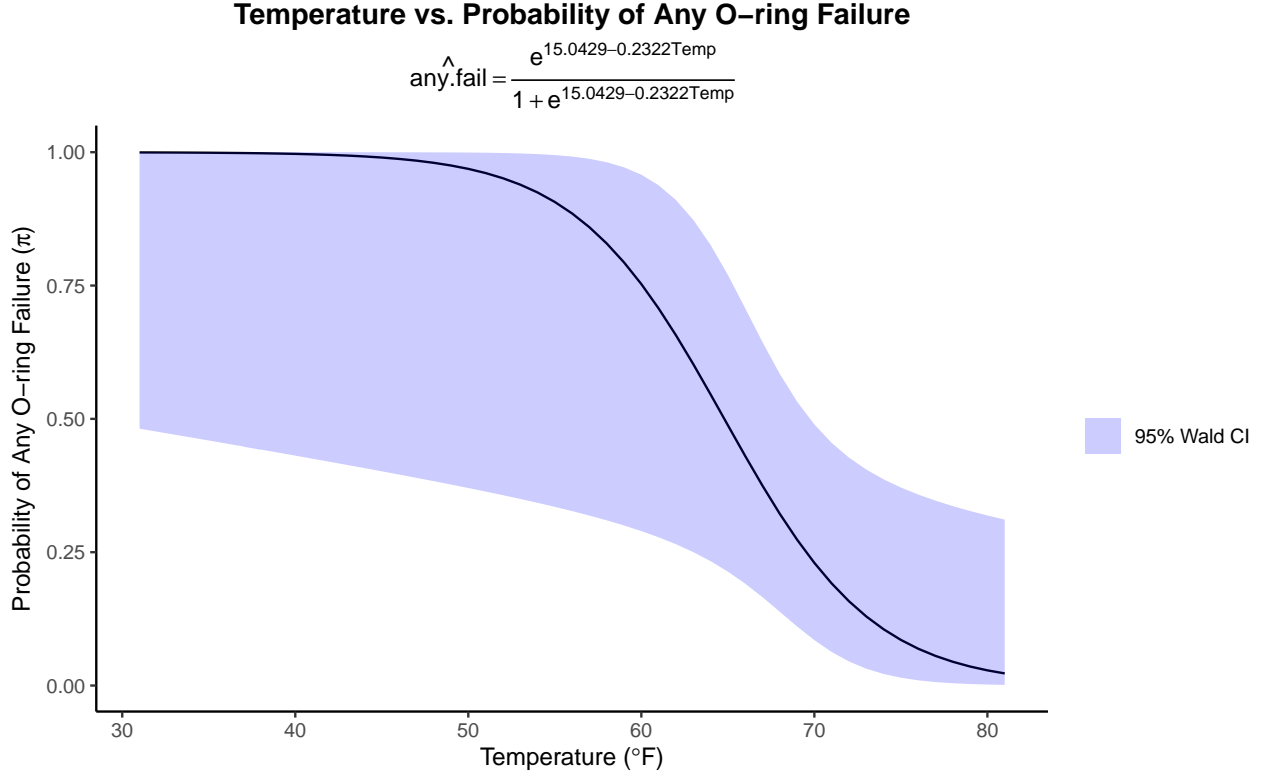


Figure 4: Probability of any O-ring failure

The probability of any O-ring failure occurring, given the model described by Eq. 10 is 0.9996, with 95% Wald Confidence interval [0.4816, 1]. Note that wider confidence interval bands indicate instability of the estimated response variables. The samples are concentrated in higher temperatures with no experiment conducted in an environment of lower than 50 degree temperature, resulting in unstable estimates. Hence, the CI bands are wider in the lower temperature area while the true response values provide narrow CI bands for the higher temperature area.

Finally, we use the *p.failure* model estimate the expected number of O-ring failures, plugging in coefficients from Table 5 as follows:

$$p.\hat{failure} = \frac{\exp(5.085 + -0.1156 \cdot Temp)}{1 + \exp(5.085 + -0.1156 \cdot Temp)} \quad (11)$$

```
# predict number of failures (multiplied by 6 because there are 6 O-rings per flight)
pred <- 6 * predict(p_fail_model, list(Temp = Temp), type="response")

# plot
ggplot(data = data.frame(Temp, pred),
       aes(x=Temp, y=pred)) +
  # add fit line
  geom_line() +
  # add existing data
  geom_count(data=dat, aes(x=Temp, y=O.ring), colour = "black", alpha=0.6, shape = 21, fill = "blue")
# pretty formatting
scale_size_continuous(name="Recorded O-ring failures", breaks = c(1,2), range=c(2,6)) +
```

```
labs(title = "Expected Number of O-rings With Thermal Distress",
     subtitle = expression(hat(p.failure) == frac(e^{5.085 - 0.1156 * Temp}, 1 + e^{5.085 - 0.1156 * Temp})),
     x = expression(paste("Temperature (", degree, "F)")),
     y = "Number of O-rings With Thermal Distress") +
theme_classic() +
theme(plot.title = element_text(hjust = 0.5, face="bold"),
      plot.subtitle = element_text(hjust = 0.5))
```

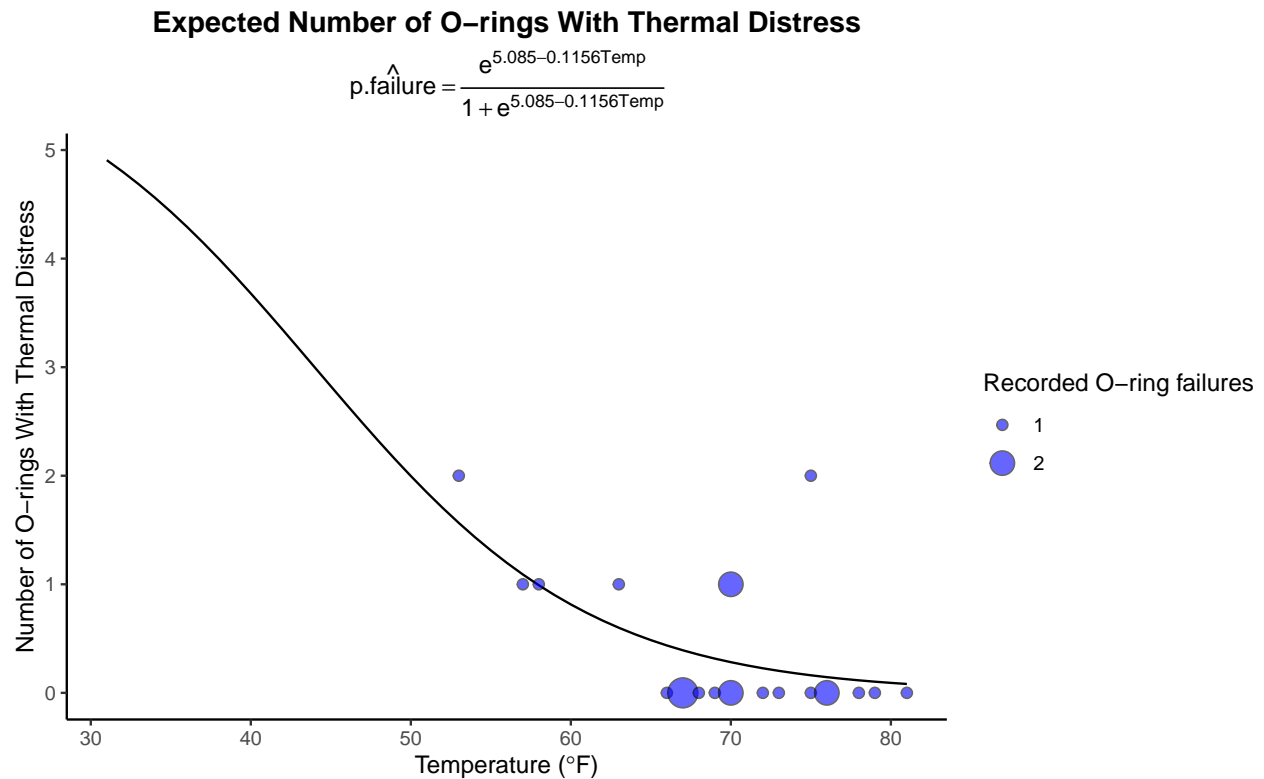


Figure 5: Estimated expected number of O-ring Failures

3.3 Bootstrapped Confidence Intervals

Rather than using Wald confidence intervals for the probability of any failure, we will use a parametric bootstrap to compute intervals for two temperatures of interest (31° F and 72° F):

1. Simulate 10,000 data sets ($n = 23$ for each) from the sample Temperature distribution
2. Use the modeled estimated probability at a specific temperature of interest (31° F or 72° F) to draw simulated O-ring failures
3. Estimate new models for each data set
4. Compute a new estimated probability at the same temperature for each model
5. Use the 0.05 and 0.95 quantiles from the distribution of probabilities as the 90% confidence interval limits

```
# define function for Monte Carlo simulation
single.sim <- function(temperature){
  # sample list of temperatures with replacement
  sample <- sample(dat$Temp, 23, replace=T)
```

```

# get pi at each sample point
sample.pi <- predict(any_failure_model, list(Temp = sample), type="response")

# draw response using pi for each point
sample.y <- rbinom(n = 23, size = 1, prob = sample.pi)

# estimate new model from sample
sample.model <- glm(sample.y ~ sample, family = binomial)

# compute specific pi from new sample model
predict(sample.model, list(sample = temperature), type="response")
}

# simulate 10,000 estimated pi at 31 degrees
sim.low <- replicate(10000, single.sim(31))
paste0("Parametric bootstrapped 90% Confidence interval at 31-degrees: [",
      round(quantile(sim.low, 0.05),4),
      ", ",
      round(quantile(sim.low, 0.95),4),
      "].")

```

[1] "Parametric bootstrapped 90% Confidence interval at 31-degrees: [0.9498, 1]."

```

# simulate 10,000 estimated pi at 72 degrees
sim.moderate <- replicate(10000, single.sim(72))
paste0("Parametric bootstrapped 90% Confidence interval at 72-degrees: [",
      round(quantile(sim.moderate, 0.05),4),
      ", ",
      round(quantile(sim.moderate, 0.95),4),
      "].")

```

[1] "Parametric bootstrapped 90% Confidence interval at 72-degrees: [0.0019, 0.3171]."

Both confidence intervals appear reasonable considering how they match the model used to generate them, with a smaller width of the band due to the added confidence of the Monte Carlo simulation.

3.4 Examination of Quadratic Term

In every model estimated so far, temperature has been statistically significant. It is worth investigating whether the effect of temperature on probability changes as temperature changes. To do this, we add a quadratic term to our model (Eq. 12), calculate the coefficients (Table 6), and perform another LRT (Table 7).

$$\text{logit}(\text{any.fail}) = \beta_0 + \beta_1 \cdot \text{Temp} + \beta_2 \cdot \text{Temp}^2 \quad (12)$$

```

# fit quadratic model
quadratic_model <- glm(any.fail ~ Temp + I(Temp^2),
                      family = binomial,
                      data = dat)

# summarize models
stargazer(any_failure_model, quadratic_model, type='latex', summary=F,
          dep.var.labels = c("Any O-ring Failure", "Any O-ring Failure"),

```

```
title = "Comparison with Quadratic model",
header=F)
```

Table 6: Comparison with Quadratic model

	<i>Dependent variable:</i>	
	Any O-ring Failure	
	(1)	(2)
Temp	−0.232** (0.108)	−2.094 (2.185)
I(Temp^2)		0.014 (0.016)
Constant	15.043** (7.379)	78.483 (76.674)
Observations	23	23
Log Likelihood	−10.158	−9.694
Akaike Inf. Crit.	24.315	25.389
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

```
# anova
q_anova <- anova(any_failure_model, quadratic_model)
stargazer(q_anova, type='latex', summary=F,
           title = "Likelihood Ratio Test (LRT)",
           header=F)
```

Table 7: Likelihood Ratio Test (LRT)

	Resid. Df	Resid. Dev	Df	Deviance
1	21	20.315		
2	20	19.389	1	0.926

The result from the LRT for the quadratic term shows a deviance of 0.926 with a p-value of 0.336. Since p-value is greater than 0.05, there is no strong evidence that shows the quadratic term is significant. Therefore, no strong evidence suggests that a quadratic relationship exists between the probability of any O-ring failure and the temperature.

4. Linear Regression Model

4.1 Model estimation

As a thought experiment, we will now estimate a linear regression model (Eq. 13) using the same set of explanatory variables we used for the binary logistic regression model (only *Temperature*) to model the number of O-ring failures (*O.ring*).

$$O.ring = \beta_0 + \beta_1 \cdot Temp \quad (13)$$

```
# fit model
linear_model <- lm(O.ring ~ Temp, data = dat)

# summarize model
stargazer(linear_model, type='latex', summary=F,
  dep.var.labels = c("Probability of O-ring Failure"),
  title = "Linear Regression Model",
  header=F)
```

Table 8: Linear Regression Model

	<i>Dependent variable:</i>
	Probability of O-ring Failure
Temp	−0.048** (0.017)
Constant	3.698*** (1.220)
Observations	23
R ²	0.261
Adjusted R ²	0.226
Residual Std. Error	0.577 (df = 21)
F Statistic	7.426** (df = 1; 21)
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

The coefficients show a negative relationship between the explanatory variable and the dependent variable. This model estimates that for every 1° increase in temperature 0.048 fewer O-rings would fail. Alternatively, one fewer O-ring would fail for every approximate 21° increase in temperature. The F-statistic of 7.426 is significant with $p < 0.05$. The R^2 and adjusted R^2 are 0.2613 and 0.2261, respectively. The R^2 numbers indicate a fairly poor goodness-of-fit for the model, while the F-statistic indicates that the model is better than an intercept-only model. This model is visualized against the data in Fig. 5:

```
# generate range of Temperatures
Temp <- seq(31, 81, 1)

# predict probability of at least 1 failure in linear form
pi <- predict(linear_model, list(Temp = Temp))

# plot
pi_df <- data.frame(Temp, pi)
ggplot(dat, aes(x=Temp, y=O.ring)) +
  # add points with count-based size
  geom_count(colour = "black", alpha=0.6, shape = 21, fill = "blue") +
  # scale and set breaks for point sizes
  scale_size_continuous(breaks = c(1,2), range=c(2,6)) +
  # add probability of any failure line
  geom_line(data = pi_df, aes(x=Temp, y=pi)) +
  # pretty formatting
  labs(title = "Temperature vs. Number of O-ring failures",
    x = expression(paste("Temperature (", degree, "F)")),
    y = "Number of O-ring failures") +
```

```
theme_classic() +
theme(plot.title = element_text(hjust = 0.5, face="bold"),
      plot.subtitle = element_text(hjust = 0.5))
```

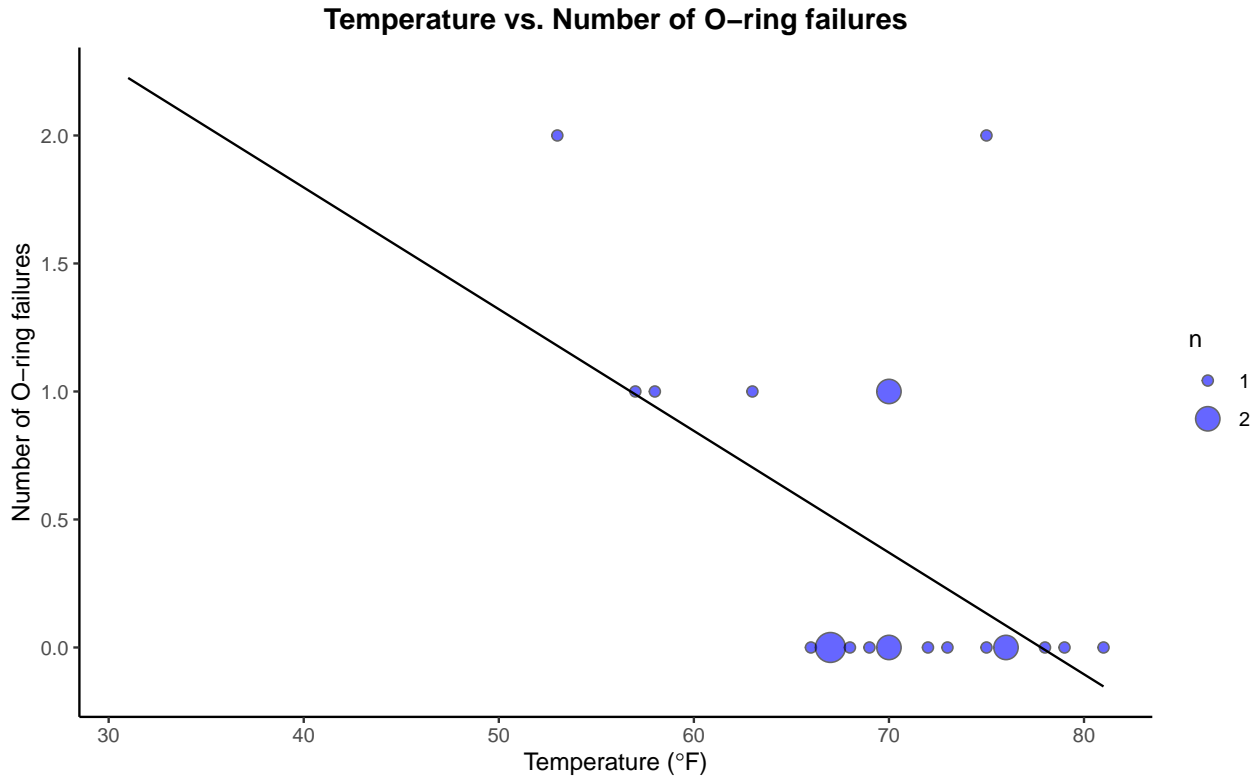


Figure 6: Linear Temperature vs. Number of O-ring failures

4.2 Classical Linear Model (CLM) Assumptions

Assumption 1: Linear in parameters

In this model, each of our β coefficients have linear relationships to the explanatory variables. Additionally, we have not constrained the error term in any way, so this assumption is met.

Assumption 2: Random sampling

The data includes available results from all experiment flights except one flight that the motors were lost at sea. Each experiment was conducted using possibly the same rocket, and if not, the same design of the rocket. Since sampling procedure changed (pressure was adjusted partway through), and not necessarily from the same design or in equivalent circumstances, we cannot conclude that the data points are independently and identically distributed.

Assumption 3: No perfect collinearity

Since the model has only one independent variable, *Temperature*, there is no violation of multi-collinearity assumption.

Assumption 4: Zero conditional mean

Examining the following residuals vs fitted plot (Fig. 6), a zero conditional mean would be an approximately flat line on the 0-residual. This model differs greatly from that. Since the mean of residuals appears endogenous, we can conclude that the model does not satisfy the assumption of zero conditional mean. It is likely there are factors correlated with either our explanatory or dependent variable that are not controlled for in this model.

```
# convert to data frame
model_df <- fortify(linear_model)

# residuals vs. fitted
ggplot(data = model_df, aes(x = .fitted, y = .resid)) +
  geom_point(shape = 21,
             size = 3,
             colour = "black",
             fill = "grey",
             alpha=0.3) +
  geom_smooth(se=F, aes(y = .stdresid), alpha = 0.5, size = 0.5, method = "loess", span = 5, formula = y ~ x) +
  labs(title = "Residuals vs. Fitted",
       x = "Fitted values",
       y = "Residuals") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))
```

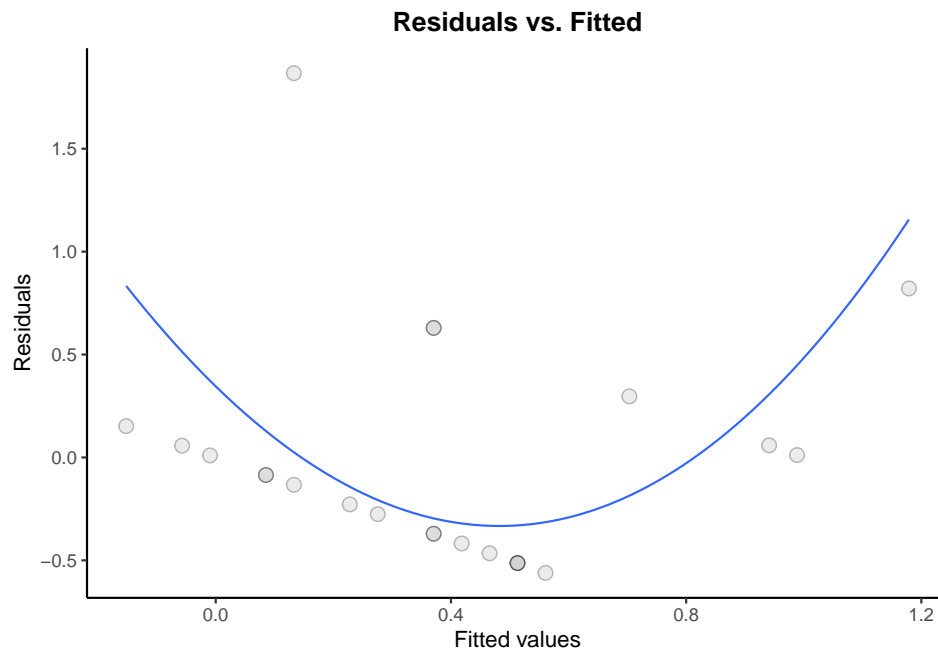


Figure 7: Residuals vs. Fitted

Assumption 5: Constant variance in the error term (Homoscedasticity)

A scale location plot (Fig. 8), shows a trend line sloping up from the left to the middle and then slides down to the right. This indicates the variance is not constant.

```
# fitted vs. sqrt of std. resid
ggplot(data = model_df, aes(x = .fitted, y = sqrt(abs(.stdresid)))) +
  geom_point(shape = 21,
             size = 3,
             colour = "black",
             fill = "grey",
             alpha=0.3) +
  geom_smooth(se=F, alpha = 0.5, size = 0.5, method = "loess", formula = y ~ x) +
  labs(title = "Scale-Location",
       x = "Fitted values",
       y = expression(sqrt("Standardized Residuals"))) +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))
```

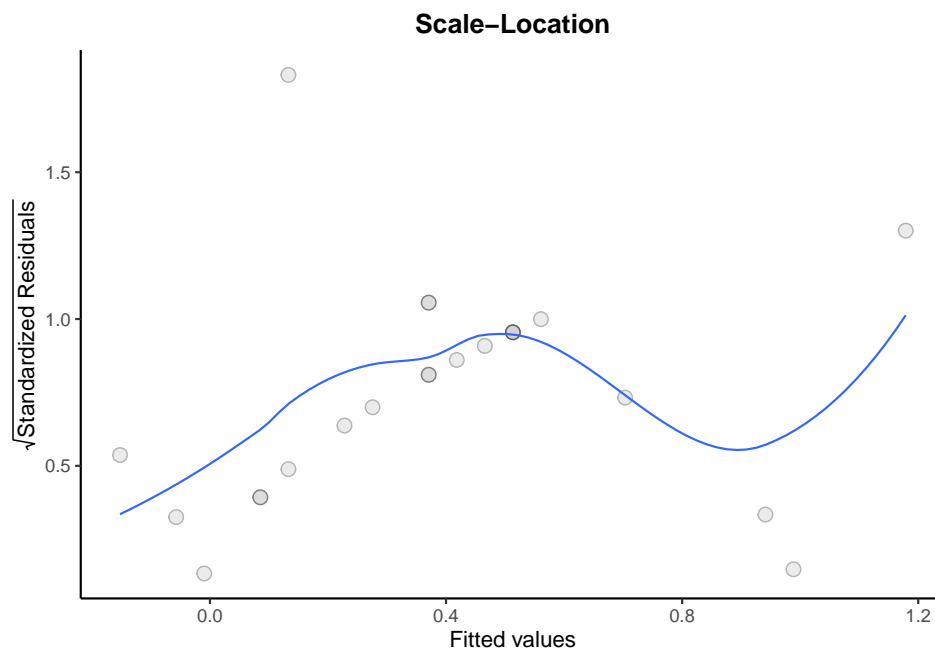


Figure 8: Scale-Location

Additionally, the results of a Breusch-Pagan test yield a BP value of 0.04 with a p-value of 0.83. Due to the non-significant p-value we are unable to reject the null hypothesis of homoscedasticity.

Assumption 6: Normal distribution of error terms

The histogram of residuals (Fig. 9, left) shows the residuals do not seem to be a normal distribution. The Normal Q-Q plot (Fig. 9, right) shows that the residuals are positively skewed.

```
# residuals hist
p1 <- ggplot(data = model_df, aes(x = .resid)) +
  geom_histogram(stat = 'bin', bins = 20, alpha = 0.3,
               fill = 'blue',
               color = 'black') +
  labs(title = "Histogram of Residuals",
       x = "Residuals",
```

```

    y = "Frequency") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))

# Q-Q plot
p2 <- ggplot(data = model_df, aes(sample = .stdresid)) +
  geom_qq(shape = 21,
          size = 3,
          colour = "black",
          fill = "grey",
          alpha=0.3) +
  geom_qq_line(alpha = 0.5, size = 0.5, color = "blue") +
  labs(title = "Q-Q Plot",
       x = "Theoretical Quantiles",
       y = "Standardized Residuals") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5))

grid.arrange(p1, p2, ncol = 2)

```

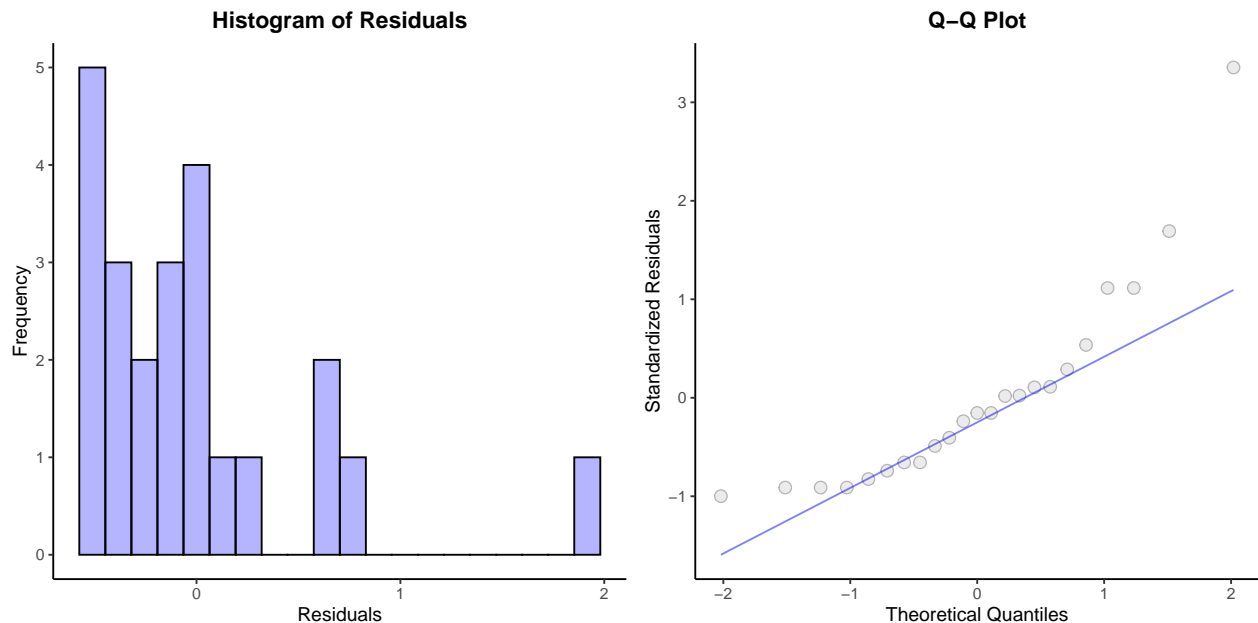


Figure 9: Normality of Errors

Finally, running a Shapiro-Wilk test on the distribution of our errors, $p = 0.001$. The significant p -value indicates we would reject the null hypothesis that the errors are normally distributed.

4.3 Linear vs. Binary logistic regression

We opt to reject the linear regression model and use the binary logistic regression due to the following reasons:

- Our purpose is to understand the probability of any O-ring failure on a rocket for a presence of a temperature effect, so the binary logistic regression is more appropriate to use than the linear regression as the dependent variable for the question is binary in nature (whether there is any O-ring failure in a flight).
- The dependent variable is not continuous, so a linear model is not appropriate overall
- The violations of assumption #2, #4, and #6 suggest that the relationship between the response and the explanatory variables may not be best represented in a linear model.

5. Implications of the Model

In the end, the strongest model candidate to be used was the simplified binary logistic regression model estimating whether or not any O-ring would fail (Eq. 14).

$$\text{logit}(\hat{any_fail}) = 15.0429 - 0.2322 \cdot Temp \quad (14)$$

Considering the consequences of even a single O-ring failure, this model provides valuable insight on the most critical question of the probability a catastrophic flight failure. In a logistic regression model, a c -unit increase in an explanatory variable corresponds to a change in the odds of the dependent variable of $e^{(c \cdot \beta)}$, where β is the coefficient for the explanatory variable in question.

$$\text{Odds ratio (OR) due to a } c\text{-unit change in Temp} = e^{(-0.2322 \cdot c)} \quad (15)$$

Applying Eq. 15, a 1° F decrease in temperature increases the odds of an O-ring failure occurring by 26.13% (Eq. 16).

$$OR = e^{(-0.2322 \cdot (-1))} = 1.2613 \quad (16)$$

The estimated probability of at least one O-ring failure at 31° F was estimated at 99.96%, with a Wald confidence interval of 48.16% to 100%. Due to the wide confidence interval as a result of limited data points, an alternative bootstrapped confidence interval of 95.49% to 100% was also calculated. The dramatically lower estimated probability of failure at higher temperatures reiterates that temperature has a significant impact in the performance of the primary O-rings. While a single O-ring failure does not necessarily imply a catastrophic launch failure will occur, in complex missions like launching a space shuttle any option to remove unnecessary risk should be taken into account. In Cape Canaveral, average temperatures in January reach into the 60s and 70s³, a temperature range with significantly lower risk of O-ring failure than the 31° day in which the Challenger launched. Our analysis indicates that waiting for a warmer launch day may have significantly decreased the probability of an O-ring failure.

³<https://weatherspark.com/m/18782/1/Average-Weather-in-January-in-Cape-Canaveral-Florida-United-States>