W271 Group Lab 1

Investigating the 1986 Space Shuttle Challenger Accident

Jonathan Phan, Priya Reddy, Spencer Zezulka

Contents

Abstract

The *Challenger* disaster of 1986 was caused by the failure of one of the shuttle's O-rings. This report reviews the validity of the hypothesis that O-ring failure is affected by temperature, and the accompanying analysis thereof from *Dalal et al., 1989*. An exploratory analysis of O-ring stress data derived from 23 preaccident shuttle launches is offered and accompanied by a probabilistic estimate of at least one of the 6 O-rings failing in the range of possible temperatures, as given by a logistic regression of failure on temperature. A comparison is made with a logistic model which also includes pressure, concluding that a model without the pressure term is more appropriate for predicting failure and further implying that temperature is the primary metric for assessing failure risk. A 90 percent confidence interval for the proportionate probability of failure of the 6 O-rings is assembled by means of a parametric bootstrap on the data, resulting in higher temperatures having lower upper limits than low temperatures. The report also juxtaposes the logistic model with an alternative linear model, and explains why the linear model cannot appropriately assess the temperature-failure relation. The report concludes that low temperature indeed significantly increases the likelihood of O-ring failure.

1 Introduction

1.1 Research question: Predict O-ring performance under the Challenger launch conditions and assess Challenger Report analysis on the effects of temperature on O-ring performance.

2 Data

2.1 Description

The data that we are working with is from is from the 23 pre-accident Challenger launches. The data collected includes information on the flight number, temperature (F), pressure (PSI), number of O-rings that failed, and the total number of O-rings on the flight. Each flight had 6 O-rings and thus could have had between 0 and 6 O-ring failures. To better understand the effect the feature of the data set (specifically temperature) had on the O-ring failures we are using a logistic regression model to estimate the probability that an O-ring will fail. In order to use this model we must assume the independence of each O-ring for each launch. The assumption of independence for each O-ring is necessary because logistic regression assumes that each trial (flight in this instance) is independent from each other. Logistic regression also uses the binomial distribution to model the probability success and/or failure. Some potential violations of this assumption of independence is that the failure of one O-ring may be due to another O-ring failing. There may also be other dependency problems, perhaps relating to the fact that each launch's O-rings may have gone through the same installation or manufacturing processes.

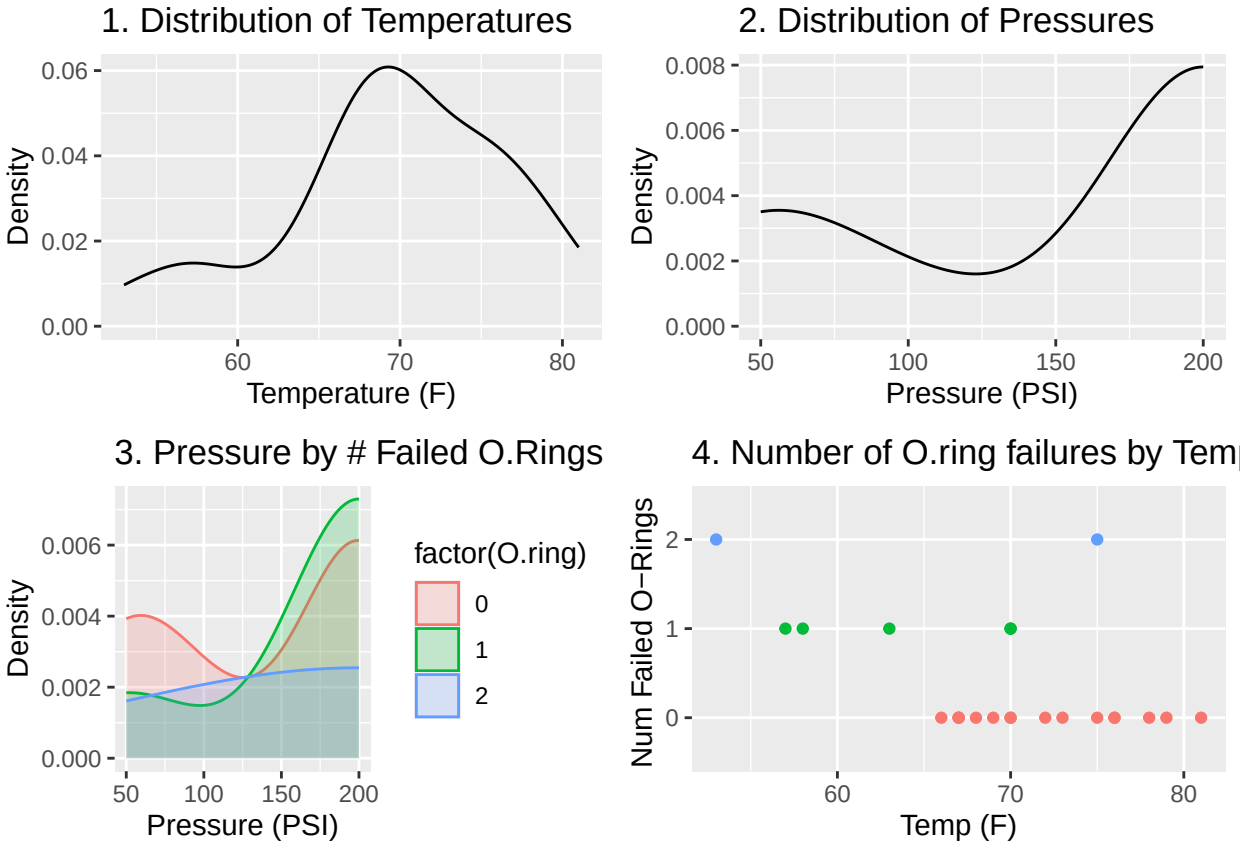
2.2 Key Features

Our variables of interest for this analysis are temperature and pressure.

Table 1: Table 1

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Temp	23	69.565	7.057	53	67	75	81
Pressure	23	152.174	68.221	50	75	200	200
O.ring	23	0.391	0.656	0	0	1	2

Looking at a Table 1, we can see that there are no missing values. We also see that for the given pre-accident launches the max number of O-ring failures was 2. There is no top or bottom coding.



Based on the plots above we can see in plot 1 that the distribution of the temperatures of the launches is roughly normal and in plot 2 that the distribution of the pressures is bi-modal. Plot 2 indicates that most launches occur at a higher pressure, but the plot 3, the distribution of pressures by number of failures, shows that O-ring failures occur both at ends of the pressure range. Further statistical testing later in this paper will show this lack of significance more clearly. Plot 4, the scatter plot of temperature vs the amount of failed O-rings shows that there are only 2 instances where 2 O-rings have failed in a given launch. The lack of more data points where 2 O-rings have failed may affect our later analysis. But overall, the fourth plot seems to show a negative relationship between temperature and number of O-ring failures, it appears that fewer O-ring failures occur at higher temperatures.

3 Analysis

3.1 Reproducing Previous Analysis

To evaluate the previous Challenger analysis, we first begin by constructing our logistic regression model using `Temp` and `Pressure` as our explanatory variables. Our initial model thus is in the form

$$logit(\hat{\pi}) = \beta_0 + \beta_1(temperature) + \beta_2(pressure)$$

We then conduct likelihood ratio tests, using the Anova command, to determine the significance of the inclusion of each of the explanatory variables.

Table 2: Table 2

	O-ring Failure	
	w/ Pressure	w/o Pressure
	(1)	(2)
Temperature (F)	-0.098^{**}	-0.116^{**}
	(0.045)	(0.047)
Pressure (PSI)	0.008	
	(0.008)	
Constant	2.520	5.085^{*}
	(3.487)	(3.052)
Observations	23	23
Log Likelihood	−15.053	−15.823
Akaike Inf. Crit.	36.106	35.647
Note:	*p<0.1; **p<0.05; ***p<0.01	

Our model with both explanatory variables was found to be:

$$logit(\hat{\pi}) = 2.52019 - 0.098297(temperature) + 0.008484(pressure)$$

Looking at table 2 we can see that only the inclusion of `Temp` appears to be significant (p<0.5). By contrast, the inclusion of `Pressure` is not significant at any level. This is in-line with what the authors found in their original analysis. The model coefficient for `Temp` shows us that for each increase in temperature by 1 degree, the odds that at least one 0-ring fails change by $e^{-0.098}$ or 0.9066 times. It might be more useful to interpret this to say that as the temperature decreases the probability of O-ring failure increases. `Pressure` was also found to be insignificant when we ran an Anova Ch-Squared test on our model. Even though it is insignificant, our model coefficients do shows that as pressure increases there is a marginal increase in chance of failure. In table 2 above note that when removing `Pressure` from our model, `Temp` becomes more significant, going from a p-value of 0.0285 to 0.014 and the AIC value of the model decreases.

Based on these results, the authors' decision to remove `Pressure` from their model is valid. We have chosen to remove `Pressure` as an explanatory variable because it's inclusion in our model is not significant. However, to get a clear understanding of the O-ring failure, furthering testing would be helpful. It is possible that the pressure could contribute to the erosion or the blow-by affecting O-rings. It is also possible that our data-set does not include conditions in which pressure was abnormal enough to affect the O-ring failure. Lastly, in their initial analysis the authors assume a linear relationship between pressure and the O-ring failure and that there is no interaction between `Temp` and `Pressure`. `Pressure` may also not be significant as a linear term, but may be significant after a transformation or after being included as an interaction term.

3.2 Confidence Intervals

Considering our analysis above validates the authors' decision to drop `Pressure` from their model. We move forward with creating the simplified logistic regression model of the form $logit(\hat{\pi}) = \beta_0 + \beta_1(temperature)$.

As seen in table 2, this simplified model is: $logit(\hat{\pi}) = 5.08498 - 0.11560(temperature)$. We can see that the coefficient for temperature has decreased in comparison to our previous model, so now for each increase in temperature by 1 degree, the odds that at least one 0 ring fails change by $e^{-0.11560}$ or 0.8908 times. In addition, we can see that the AIC value for the simplified model is smaller than the AIC for our more complex model, which further validates our decision to drop the `Pressure` term.

In order to check for any possible non-linear relationship between `Temp` and the O-ring failures, we also construct a quadratic model of the form: $logit(\hat{\pi}) = \beta_0 + \beta_1(temperature) + \beta_2(temperature^2)$.

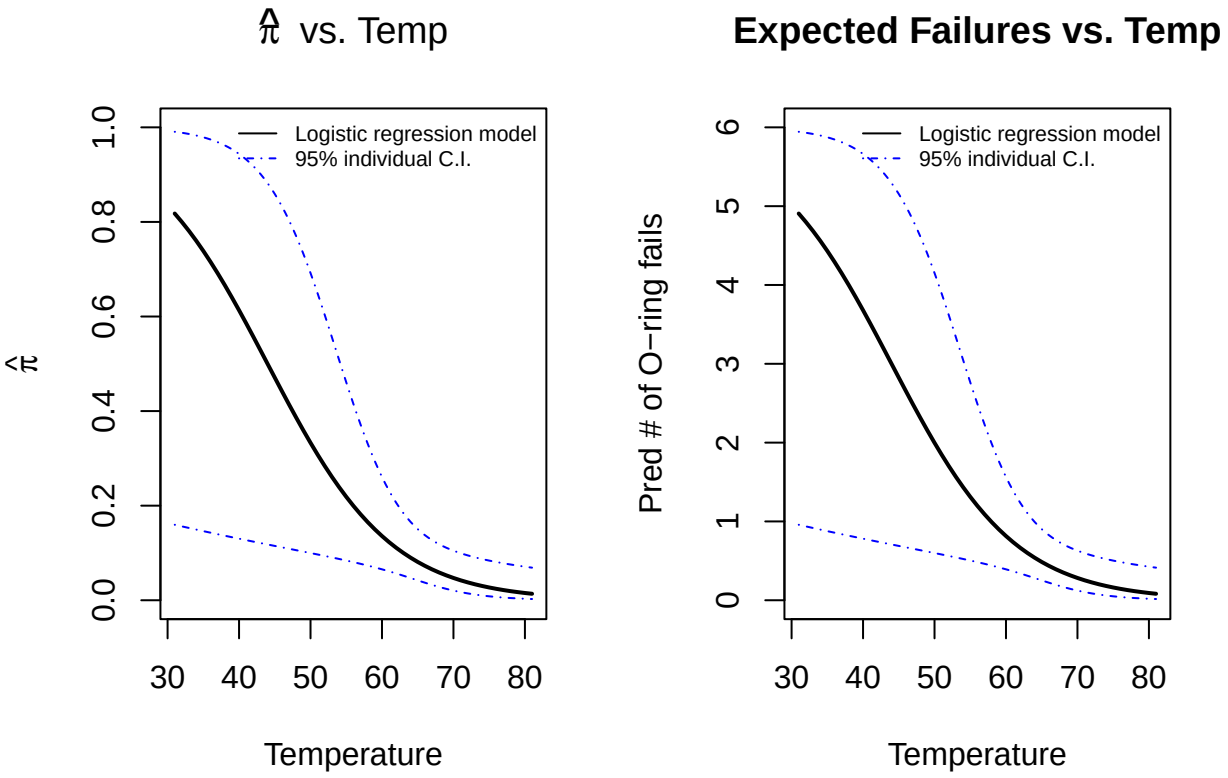
Table 3: Table 3

	O-ring Failure	
	Linear-Logistic	Quadratic-Logistic
	(1)	(2)
Temperature (F)	−0.116**	−0.651
	(0.047)	(0.741)
I(Temperature**2)		0.004
		(0.006)
Constant	5.085*	22.126
	(3.052)	(23.794)
Observations	23	23
Log Likelihood	−15.823	−15.576
Akaike Inf. Crit.	35.647	37.152
Note:		*p<0.1; **p<0.05; ***p<0.01

As shown in Table 3, when we construct our model it appears that the `I(Temp^2)` term is statistically insignificant. We furthered confirmed this by conducting an Anova Chi-Squared test, which showed that the `I(Temp^2)` term has a p-value of 0.467 and is thus insignificant. In addition to these metrics we can also see that the AIC value of the non-quadratic model is smaller than that of the

quadratic model, indicating that the simpler model is better. As a result we can determine that the quadratic term is not needed in our model, and we can move forward in our analysis with the simplified temperature only model.

Using this simplified model we can construct two plots to visualize the relationship between $\hat{\pi}$ and `Temp`.



From the plots above we can see that the probability of O-ring failure increases as temperature decreases. However, we can also see that the confidence interval bands are much wider at lower temperatures than at higher temperatures. This shift occurs at around 60°F. This is likely because we had very few low temperature observations in our sample, 20 of the 23 launches in our sample occurred at temperatures greater than 60°F.

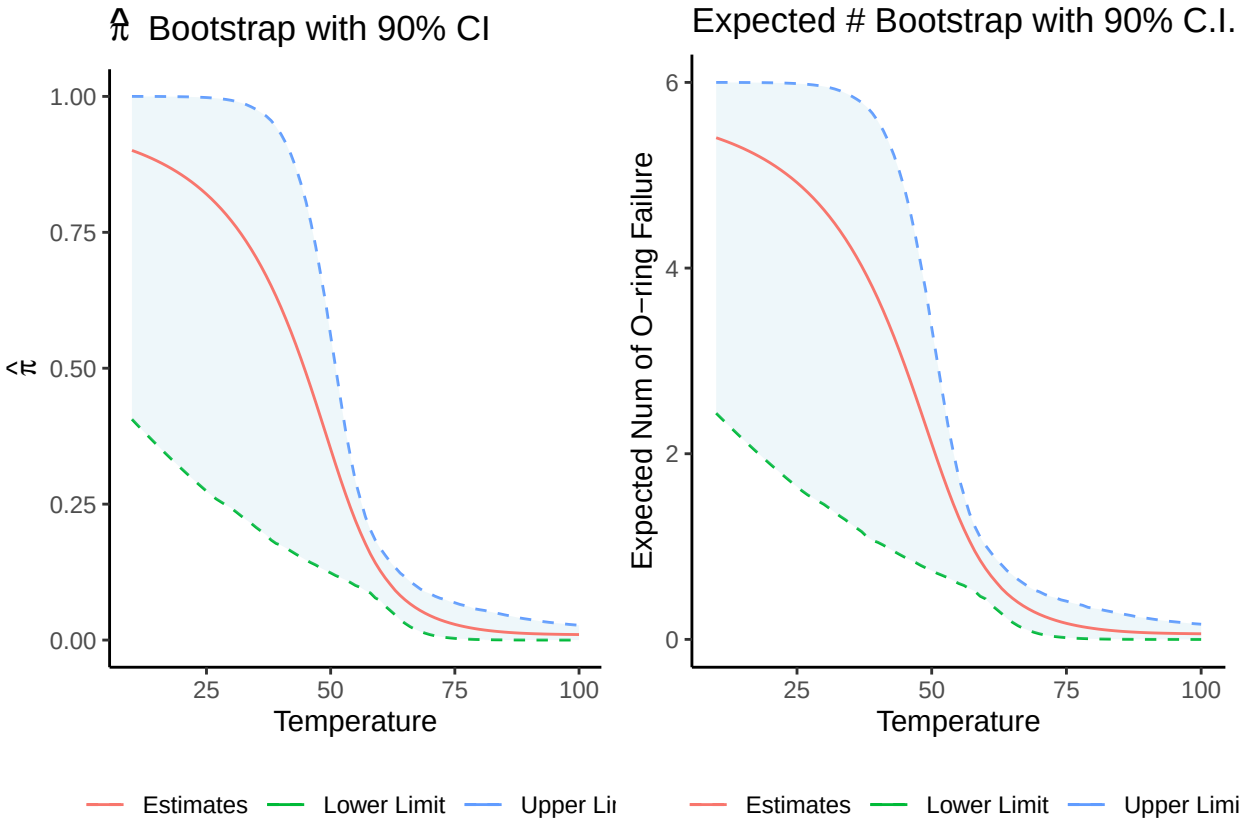
We know that the temperature was 31°F at launch for the Challenger in 1986. When we use our model to estimate the probability of an O-ring failure using this temperature,we find that there is a predicted 82% probability of O-ring failure, with a Wald 95% confidence interval of 15.96% to 99.06%. This is a very large CI band, which in part maybe due to the fact that the Wald CI does not work well with samples of less than 40, and our sample only has 23 observations. However, when we re-calculated the CI using the Profile Likelihood method to account for this we still found the confidence band to be quite large (14.18% to 99.05%).

As mentioned before this is likely due to the fact that our sample has very few data points for low temperatures. In addition, our sample has no data for temperatures lower than 53°F and we are estimating the log-likelihood of O-ring failure at 31°F. As a result, we are simply assuming that the logistic regression model we have calculated holds at this lower temperature, but it is possible that

our assumption of the linear relationship between `Temp` and the log-likelihood of O-ring failure is not sound. Even though we conducted a statistical test to rule out a quadratic relationship between `Temp` and pi_hat, it could be that within the range of our data-set any non-linear relationship is undetectable, but at 31°F (28° below our lowest observation) it is.

3.3 Bootstrap Confidence Intervals

As seen in the previous section, the confidence band for the lower temperature estimates is extremely wide. To better understand the variability in our predictions at lower temperatures we can use a bootstrap method. To create a parametric bootstrap and produce a resulting 90% confidence interval, we first generated 1000 samples (n=23) from the original data set with replacement. Then, for each of these 1000 samples, we generated a model of the form $logit(\hat{\pi}) = \beta_0 + \beta_1(temperature)$. Then for each of these 1000 models we predicted $\hat{\pi}$ at each integer temperature between 10° and 100° Fahrenheit. Using these predicted values we were able to compute the 0.05 and 0.95 quantile estimation at each temperature to construct our 90% confidence bands. For our estimation line we computed the mean estimation at each temperature.



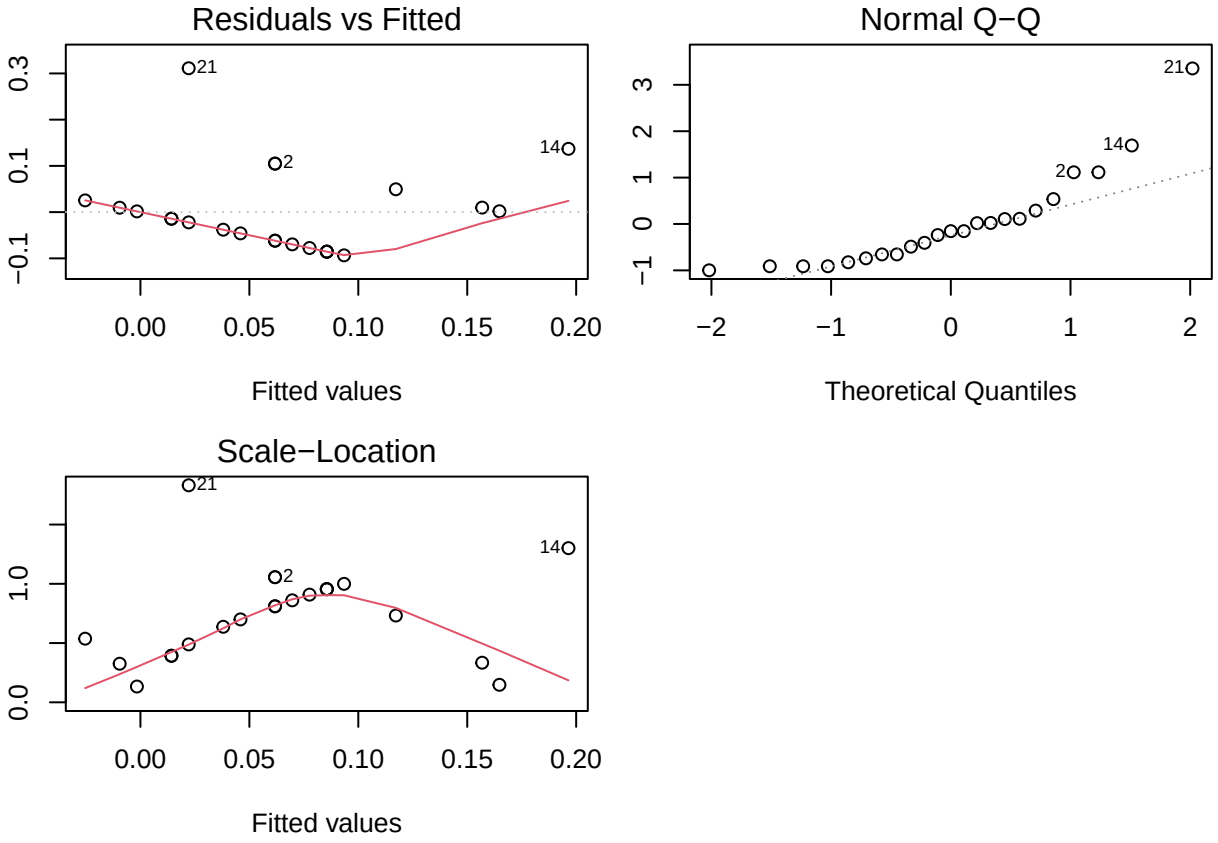
The plots above showcase the results of the bootstrap: it is clear that as temperature increases, the probability of having 6 of the O-ring failing decreases. It is also worth noting that the confidence interval is still much wider at lower temperatures (<65°F)than higher temperatures. As mentioned above, this is most likely due to the original data set having only a few observations at lower temperatures, leading to greater uncertainty for estimates at these temperatures.

3.4 Alternative Specification

As a final analysis of our model, we decided to compare it to a linear regression model with the same explanatory variables, this model takes the form $P(Oring.failures) = \beta_0 + \beta_1(temperature)$.

Table 4: Table 4

	O-ring Failure Linear
Temperature (F)	−0.008**
	(0.003)
Constant	0.616***
	(0.203)
Observations	23
R^2	0.261
Adjusted R^2	0.226
Residual Std. Error	0.236 (df = 21)
F Statistic	7.426** (df = 1; 21)
Note:	*p<0.1; **p<0.05; ***p<0.01



The model we generated takes the form $P(Oring.failures) = 0.616402 - 0.007923(temperature)$, as seen in table 4. According to this model, at a temperature of 31°F 37% of the O-rings would fail.

Following this interpretation we can see one immediate problem with this the assumption that the model of the relationship between `Temp` and the probability of O-ring failure is linear. The probability of failure is by definition bounded by [0,1], so a linear model, which is unbounded is not an appropriate model choice. We can see this since, with our calculated model, at temperatures greater than 77.97°F we get a predicted probability of less than 0, and at temperatures less than -48.415°F we get a predicted probability of failure of greater than 1.

In addition to this, when we evaluate the other assumptions needed to create a linear regression model through plotting the residuals, we can see that several other assumptions are violated. First, the samples used to create the model must be IID, this assumption is possibly violated since the O-rings are not necessary independent as each set of 6 O-rings on a launch are subjected to the same conditions and the failure of 1 O-ring might influence the state of the other O-rings. In addition when we plotted the residuals vs fitted values plot for the model we can see that the assumption of zero-conditional mean of residuals is violated, if the assumption were met we would see a linear relationship between the two, but from the plot we can see the non-linear relationship. When we look at the Scale-Location plot we can also see that the assumption of the homoskedasticity of the errors is also violated. The plotted residuals are not equally spread across the range of predictors. Lastly, from the Q-Q plot we can see that our model has a relatively normal distribution of its residuals, but that it is not as normal towards the upper quantiles, with a larger sample we may have been able to assume the Central Limit Theorem but with such a small sample we unable to do so with confidence. The only assumption that holds true is that of no perfect co-linearity. Since our model only includes one predictors, we know that there is no co-linearity.

We can also see that the model does not do a good job of explaining the variation in the data as seen by the low R^2 (0.2) in table 4. Based on these violations of the assumptions for creating a linear regression model, it appears that the the logistic regression model is a more appropriate model choice for this scenario.

4 Conclusions

Interpret the main result of your preferred model in terms of both odds and probability of failure. Summarize this result with respect to the question(s) being asked and key takeaways from the analysis.

Our final model found that for every 1 degree increase in temperature the probability of O-ring failure decreases by 11%. Another way of looking at these results is that for a 10 degree decrease in the temperature, the estimated odds of failure proportionate to the number of O-rings changes by 3.177 times, with a 95% confidence that these odds change between 1.277 and 8.35 times for every 10 unit decrease in temperature.

These results confirm what the authors initially found in their analysis: at lower temperatures, there is a higher likelihood of an O-ring failure. Based on these results, it is likely that the low temperatures the day of the challenger launch contributed to the ultimately catastrophic O-ring failures.

Lab 1, Short Questions

Contents

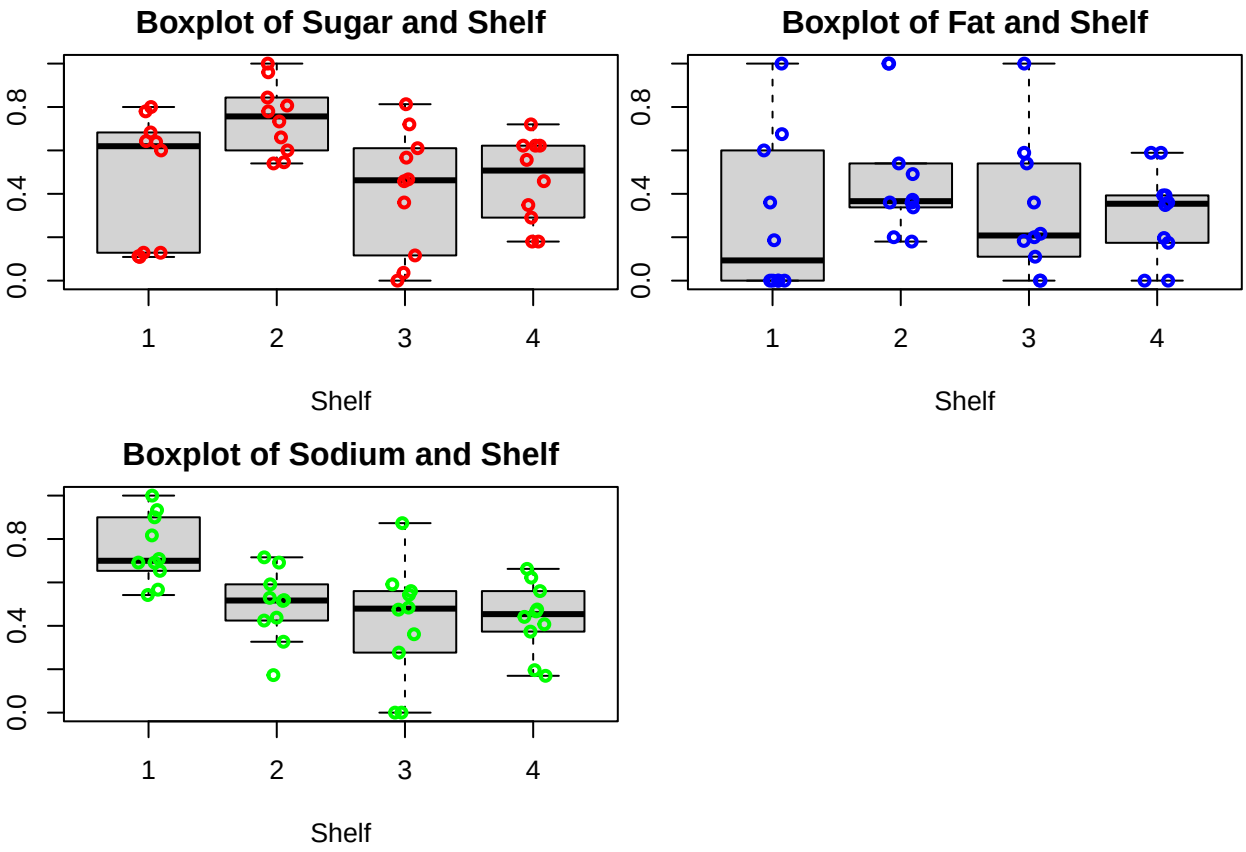
1 Strategic Placement of Products in Grocery Stores (5 points)

These questions are taken from Question 12 of chapter 3 of the textbook(Bilder and Loughin's "Analysis of Categorical Data with R.

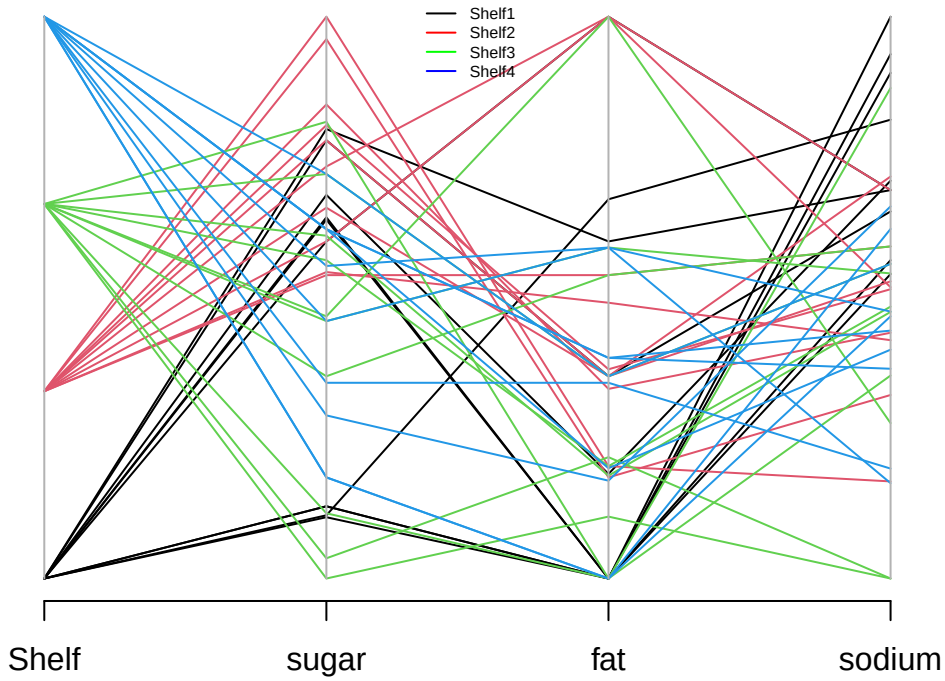
> *In order to maximize sales, items within grocery stores are strategically placed to draw customer attention. This exercise examines one type of item—breakfast cereal. Typically, in large grocery stores, boxes of cereal are placed on sets of shelves located on one side of the aisle. By placing particular boxes of cereals on specific shelves, grocery stores may better attract customers to them. To investigate this further, a random sample of size 10 was taken from each of four shelves at a Dillons grocery store in Manhattan, KS. These data are given in the* cereal_dillons.csv *file. The response variable is the shelf number, which is numbered from bottom (1) to top (4), and the explanatory variables are the sugar, fat, and sodium content of the cereals.*

1.1 Recode Data

(1 point) The explanatory variables need to be reformatted before proceeding further (sample code is provided in the textbook). First, divide each explanatory variable by its serving size to account for the different serving sizes among the cereals. Second, rescale each variable to be within 0 and 1. Construct side-by-side box plots with dot plots overlaid for each of the explanatory variables. Also, construct a parallel coordinates plot for the explanatory variables and the shelf number. Discuss whether possible content differences exist among the shelves.



Overall, we can see that the variance in the data across features tends to be higher for shelves 1 and 3 than for their counterparts. Shelf 1 has the highest sodium mean sodium content and the lowest mean fat content. Shelves 2, 3, and 4 have similar mean sodium contents. Shelves 2 and 4 have a similar mean fat content. Shelves 1, 3, and 4 have similar mean sugar contents—shelf 2 has the highest by a considerable margin.



'Fill in: What do you observe in these parallel coordinates plots?' The parallel coordinates plots provide similar insights to above—it can be seen that the cereals from shelf 1 tend to have low sugar and fat contents as well as high sodium contents, while the cereals from shelf 2 tend to have high sugar and fat but mid sodium—the other two shelves are more evenly dispersed across all the nutrient profiles.

Fill in: Do content differences exist between the shelves?' It would appear from first glance at the EDA that there may be differences in contents based on the shelf of origin—it is possible that factors like average eye height by demographic might affect the distribution of how cereals are placed on the shelves, such as lower shelves containing more sugar to appeal to youngsters. However, a more rigorous analysis is required to substantiate such a claim.

1.2 Evaluate Ordinal vs. Categorical

(1 point) The response has values of $1, 2, 3$, and 4. Explain under what setting would it be desirable to take into account ordinality, and whether you think that this setting occurs here. Then estimate a suitable multinomial regression model with linear forms of the sugar, fat, and sodium variables. Perform LRTs to examine the importance of each explanatory variable. Show that there are no significant interactions among the explanatory variables (including an interaction among all three variables).

Ordinality should be taken into account when there is a natural ordering to the response—in this case, while central shelves may be more favorable given that they tend to be more level with average eye height, there is no consistent order between the shelves, so it doesn't make sense to include ordinality. We can not conclude that one shelf is necessarily "better" than another shelf. Although it is arguable that shelves 2 and 3 are

more preferable to the average human based on height, it is not wise to exclude children and taller people who's attention is possibly captured more by shelf 1 and 4.

```
## Call:
## multinom(formula = Shelf ~ sugar + fat + sodium, data = cereal2,
##     trace = FALSE)
##
## Coefficients:
##   (Intercept)      sugar        fat    sodium
## 2    6.900708   2.693071  4.0647092 -17.49373
## 3   21.680680 -12.216442 -0.5571273 -24.97850
## 4   21.288343 -11.393710 -0.8701180 -24.67385
##
## Std. Errors:
##   (Intercept)    sugar      fat   sodium
## 2    6.487408 5.051689 2.307250 7.097098
## 3    7.450885 4.887954 2.414963 8.080261
## 4    7.435125 4.871338 2.405710 8.062295
##
## Residual Deviance: 67.19028
## AIC: 91.19028
```

```
## Call:
## multinom(formula = Shelf ~ sugar + fat + sodium + sugar:fat +
##     fat:sodium + sugar:sodium + sugar:fat:sodium, data = cereal2,
##     maxit = 2000, trace = FALSE)
##
## Coefficients:
##   (Intercept)      sugar       fat     sodium   sugar:fat fat:sodium
## 2    -3.24345   3.659946  41.79546  -0.982032   42.640294  -46.61927
## 3    23.68888 -20.089119 102.84550 -26.479013 -106.626411 -168.01534
## 4    29.24182 -22.120274  50.25675 -31.093759   -2.334124 -101.04527
##   sugar:sodium sugar:fat:sodium
## 2    -5.566470        -74.27235
## 3     9.401781        187.85859
## 4    -1.150455         56.40446
##
## Std. Errors:
##   (Intercept)    sugar      fat   sodium sugar:fat fat:sodium sugar:sodium
## 2    27.66550 32.60086 133.3377 30.21397  181.9377   159.9864     36.71580
## 3    23.74289 26.13747 145.1512 25.46450  208.2268   194.3427     26.00302
## 4    23.69586 26.76547 144.3816 25.46274  208.8939   195.7386     30.40390
##   sugar:fat:sodium
## 2         211.5563
## 3         280.9464
## 4         290.8940
##
## Residual Deviance: 51.43707
## AIC: 99.43707
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: Shelf
##         LR Chisq Df Pr(>Chisq)
## sugar    22.7648  3  4.521e-05 ***
## fat       5.2836  3     0.1522
## sodium   26.6197  3  7.073e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: Shelf
##                   LR Chisq Df Pr(>Chisq)
## sugar               5.8375  3    0.11979
## fat                 6.2051  3    0.10205
## sodium              7.4655  3    0.05845 .
## sugar:fat           5.7756  3    0.12305
## fat:sodium          6.1020  3    0.10675
## sugar:sodium        3.6089  3    0.30691
## sugar:fat:sodium    4.6963  3    0.19544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

> 'Fill in: Write about what you learn as a result of these tests, using inline code evaluation.' As can be seen from the Type III Analysis of Deviance table, none of the interaction terms are statistically significant—however, we can see from the Type II Analysis of Deviance table that both sugar and sodium content are highly significant in predicting shelf placement. As a result we decide to use the model without the interaction terms.

1.3 Where do you think Apple Jacks will be placed?

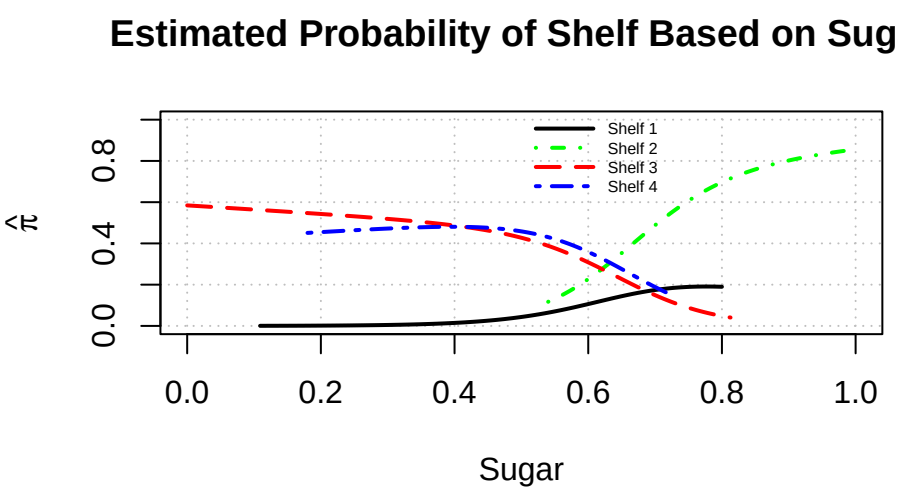
(1 point) Kellogg's Apple Jacks (http://www.applejacks.com) is a cereal marketed toward children. For a serving size of 28 grams, its sugar content is 12 grams, fat content is 0.5 grams, and sodium content is 130 milligrams. Estimate the shelf probabilities for Apple Jacks.

```
##          1          2          3          4
## 0.05326849 0.47194264 0.20042742 0.27436145
```

> 'Fill this in: Where does your model predict apple jacks will be placed?' The model predicts that the most likely shelf for apple jack placement is shelf 2, with a probability of roughly 0.47.

1.4 Figure 3.3

(1 point) Construct a plot similar to Figure 3.3 where the estimated probability for a shelf is on the *y-axis* and the sugar content is on the *x-axis*. Use the mean overall fat and sodium content as the corresponding variable values in the model. Interpret the plot with respect to sugar content.



'Fill this in: What message does your plot give?' The plot shows that, as mentioned earlier in the EDA, the probability for high sugar content is greatest for shelf 2. It is interesting to see how shelf 3 and 4 follow each other closely. As sugar contents rise for these two shelves, their probability decreases. The realm of high sugar cereals is mainly absorbed by shelf 2.

1.5 Odds ratios

(1 point) Estimate odds ratios and calculate corresponding confidence intervals for each explanatory variable. Relate your interpretations back to the plots constructed for this exercise.

```
##     sugar       fat    sodium
## 0.2692078 0.2990292 0.2298359
```

Odds ratio for Shelf 2 vs Shelf 1

```
##        Odds-Ratio Inverse
## sugar        2.06    0.48
## fat          3.37    0.30
## sodium       0.02   55.74
```

Odds ratio for Shelf 3 vs Shelf 1

```
##        Odds-Ratio Inverse
## sugar        0.04   26.81
## fat          0.85    1.18
## sodium       0.00  311.36
```

Odds ratio for Shelf 4 vs Shelf 1

```
##        Odds-Ratio Inverse
## sugar        0.05   21.48
## fat          0.77    1.30
## sodium       0.00  290.31
```

CI for Shelf 2 parameter estimates

```
##         low    up
## sugar  0.14 29.68
## fat    0.87 13.04
## sodium 0.00  0.44
```

CI for Shelf 3 parameter estimates

```
##         low   up
## sugar  0.00 0.49
## fat    0.21 3.49
## sodium 0.00 0.12
```

CI for Shelf 4 parameter estimates

```
##         low   up
## sugar  0.00 0.61
## fat    0.19 3.16
## sodium 0.00 0.13
```

> 'Fill this in: What do you learn about each of these variables?' If we decrease the amount of sodium by a factor of 0.23, the odds of being placed on shelf 2 as opposed to shelf 1 increase by a factor of 1/0.02, or 50 times. The odds of high sodium cereals, as confirmed by the boxplot EDA, being on a shelf other than shelf 1 or 2 approaches 0. All else equal, the odds of a cereal being placed on shelf 2 over shelf 1 increase by a factor of 2.06 for a 0.27 increase in sugar. All else equal, the odds of a cereal being placed on shelf 2 over shelf 1 increase by a factor of 3.3 for a 0.30 increase in fat. This is also consistent with our earlier EDA.

2 Alcohol, self-esteem and negative relationship interactions (5 points)

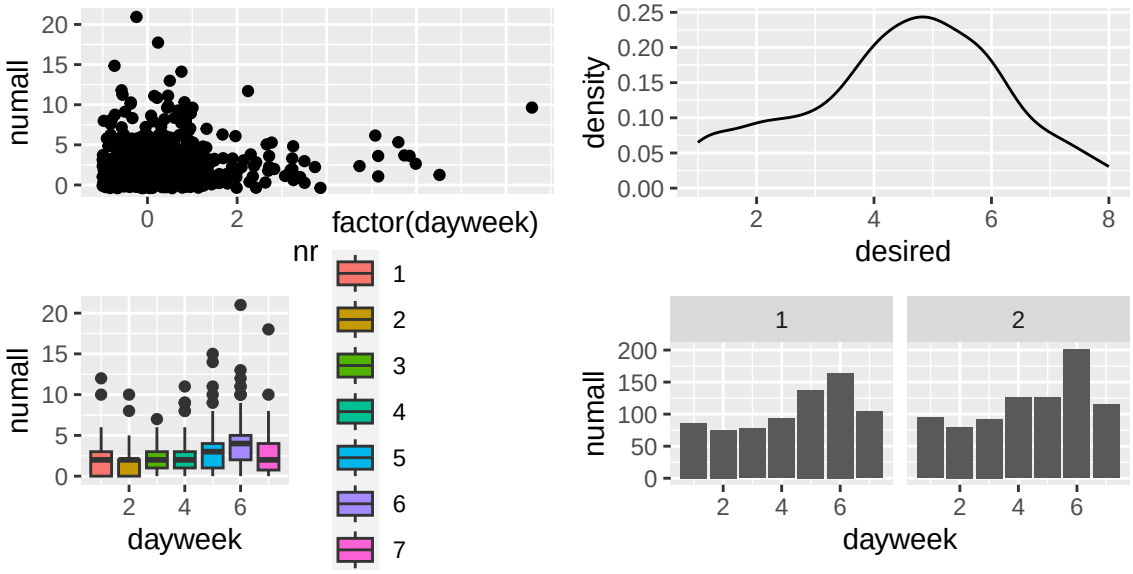
Read the example **'Alcohol Consumption'** in chapter 4.2.2 of the textbook(Bilder and Loughin's "Analysis of Categorical Data with R). This is based on a study in which moderate-to-heavy drinkers (defined as at least 12 alcoholic drinks/week for women, 15 for men) were recruited to keep a daily record of each drink that they consumed over a 30-day study period. Participants also completed a variety of rating scales covering daily events in their lives and items related to self-esteem. The data are given in the *DeHartSimplified.csv* data set. Questions 24-26 of chapter 3 of the textbook also relate to this data set and give definitions of its variables: the number of drinks consumed (`numall`), positive romantic-relationship events (`prel`), negative romantic-relationship events (`nrel`), age (`age`), trait (long-term) self-esteem (`rosn`), state (short-term) self-esteem (`state`).

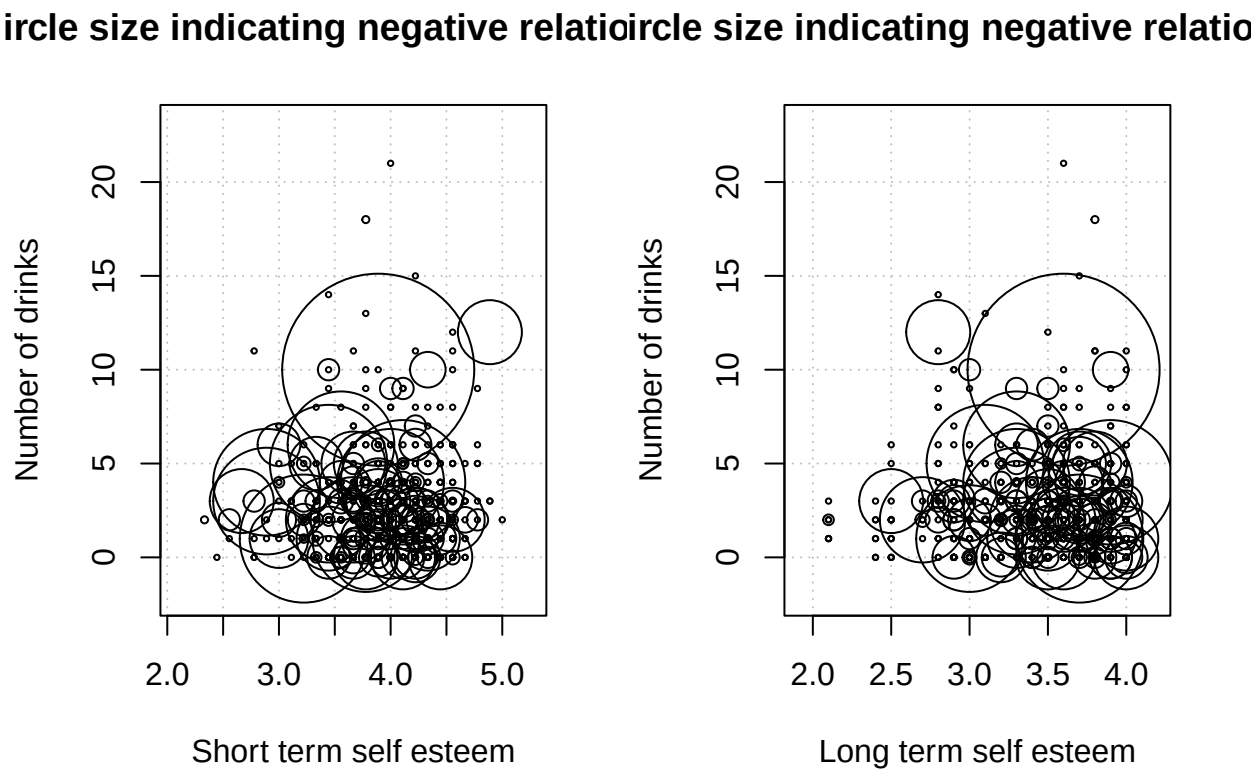
The researchers stated the following hypothesis:

> *We hypothesized that negative interactions with romantic partners would be associated with alcohol consumption (and an increased desire to drink). We predicted that people with low trait self-esteem would drink more on days they experienced more negative relationship interactions compared with days during which they experienced fewer negative relationship interactions. The relation between drinking and negative relationship interactions should not be evident for individuals with high trait self-esteem.*

2.1 EDA

(2 points) Conduct a thorough EDA of the data set, giving special attention to the relationships relevant to the researchers' hypotheses. Address the reasons for limiting the study to observations from only one day.





'Fill this in: What do you learn?' As can be seen from the pairplots and the boxplots,

the day of the week is correlated significantly with only the number of drinks and the desire to drink. This suggests that we should run our analysis across a given day to eliminate error based on variations in the day of the week. We can also see that the number of drinks is correlated significantly with positive relationship events, but not with negative relationship events, as hypothesized.

2.2 Hypothesis One

(2 points) The researchers hypothesize that negative interactions with romantic partners would be associated with alcohol consumption and an increased desire to drink. Using appropriate models, evaluate the evidence that negative relationship interactions are associated with higher alcohol consumption and an increased desire to drink.

```
##
## Call:
## glm(formula = numall ~ nrel, family = poisson(link = "log"),
##     data = saturday)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.8337  -1.3211  -0.5305   0.4733   5.9597
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.39003    0.05715  24.320   <2e-16 ***
## nrel         0.04971    0.05076   0.979    0.328
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 250.34  on 88  degrees of freedom
## Residual deviance: 249.43  on 87  degrees of freedom
## AIC: 508.83
##
## Number of Fisher Scoring iterations: 5
```

```
##
## Call:
## lm(formula = desired ~ nrel, data = saturday)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8467 -0.8453  0.1533  1.1547  3.1547
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.845267   0.184642  26.241   <2e-16 ***
## nrel        0.002914   0.178607   0.016    0.987
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.604 on 87 degrees of freedom
## Multiple R-squared:  3.059e-06,  Adjusted R-squared:  -0.01149
## F-statistic: 0.0002662 on 1 and 87 DF,  p-value: 0.987
```

'Fill this in: What do you learn?' From this table, we can see that the association between negative relationship events and either the number of drinks or desire to drink is not statistically significant.

2.3 Hypothesis Two

(1 point) The researchers hypothesize that the relation between drinking and negative relationship interactions should not be evident for individuals with high trait self-esteem. Conduct an analysis to address this hypothesis.

```
##
## Call:
## glm(formula = numall ~ nrel + esteem + nrel:esteem, family = poisson(link = "log"),
##     data = saturday)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.8711  -1.3742  -0.3801   0.6371   5.8848
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.416230   0.067001  21.137   <2e-16 ***
## nrel         0.050510   0.073759   0.685    0.493
## esteem      -0.105886   0.132011  -0.802    0.422
## nrel:esteem  0.006995   0.102600   0.068    0.946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 250.34  on 88  degrees of freedom
## Residual deviance: 248.70  on 85  degrees of freedom
## AIC: 512.1
##
## Number of Fisher Scoring iterations: 5
```

'Fill this in: What do you learn?' We can see that, indeed, this hypothesis holds — for individuals with high self esteem, there is no statistically significant evidence suggesting a relationship between negative relationship events and total drinks for individuals with high self esteem.