Untitled

2025-02-21

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(tableone)
library(effsize)
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(car)
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
library(pscl)
```

```
## Classes and Methods for R originally developed in the
## Political Science Computational Laboratory
## Department of Political Science
```

```
## Stanford University (2002-2015),
## by and under the direction of Simon Jackman.
## hurdle and zeroinfl functions by Achim Zeileis.
```

```
library(ResourceSelection)
```

```
## ResourceSelection 0.3-6   2023-06-27
```

```
library(glmnet)
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-8
```

```
laryngoscope_data <- read.csv("C:/Users/dapah/Downloads/Laryngoscope.csv")
```

Data Exploration:

First, I'd like to look at the prescreening information collected. You randomized each participant into an intubation group. So lets verify the groups are as alike as possible. I'll start by showing the distribution of how many participants we have for each baseline attribute by intubation group. Then I'm just showing the averages for each attribute. aka Summary Statistics

```
# Count and proportion for categorical variables
laryngoscope_data %>%
  group_by(Randomization) %>%
  summarise(
    count = n(),
    Male = sum(gender == 1, na.rm = TRUE),
    Female = sum(gender == 0, na.rm = TRUE),
    ASA_1 = sum(asa == 1, na.rm = TRUE),
    ASA_2 = sum(asa == 2, na.rm = TRUE),
    ASA_3 = sum(asa == 3, na.rm = TRUE),
    ASA_4 = sum(asa == 4, na.rm = TRUE),
    Mallampati_1 = sum(Mallampati == 1, na.rm = TRUE),
    Mallampati_2 = sum(Mallampati == 2, na.rm = TRUE),
    Mallampati_3 = sum(Mallampati == 3, na.rm = TRUE),
    Mallampati_4 = sum(Mallampati == 4, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 12
##   Randomization count  Male Female ASA_1 ASA_2 ASA_3 ASA_4 Mallampati_1
##           <int> <int> <int>  <int> <int> <int> <int> <int>        <int>
## 1             0    49    10     39     0     7    40     2           14
## 2             1    50    11     39     0    15    32     3           21
## # i 3 more variables: Mallampati_2 <int>, Mallampati_3 <int>,
## #   Mallampati_4 <int>
```

```
laryngoscope_data %>%
  group_by(Randomization) %>%
  summarise(
    Age_Mean = mean(age, na.rm = TRUE),
    Age_SD = sd(age, na.rm = TRUE),
    BMI_Mean = mean(BMI, na.rm = TRUE),
    BMI_SD = sd(BMI, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 5
```

```
##   Randomization Age_Mean Age_SD BMI_Mean BMI_SD
##           <int>    <dbl>  <dbl>    <dbl>  <dbl>
## 1             0     48.5   14.1     42.5   5.91
## 2             1     50.3   12.2     41.4   4.44
```

Interpretation:

Everything appears to be in order. There aren't any participants participants with a Mallampati score of 4 in the control group. I'm not sure how this would impact our test. But we should be aware of this. Next we'll make sure everything appears similar.

Testing for Differences in Baseline Characteristics:

Continuing, I'm use statistical tests to compare each attribute by intubation group to ensure the study's internal validity I'm using a t-test to compare quantitative variables. For categorical/rating-scale variables I'm using a chi-square test if table cells are large or a fisher-test is any of the table's cells are small.

```r
# Checking Quantitative Vars
t.test(age ~ Randomization, data = laryngoscope_data)  # T-test (if normal)
```

```
##
##  Welch Two Sample t-test
##
## data:  age by Randomization
## t = -0.6835, df = 94.495, p-value = 0.496
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -7.066779  3.447188
## sample estimates:
## mean in group 0 mean in group 1
##         48.5102         50.3200
```

```r
t.test(BMI ~ Randomization, data = laryngoscope_data)
```

```
##
##  Welch Two Sample t-test
##
## data:  BMI by Randomization
## t = 1.0254, df = 89.042, p-value = 0.3079
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -1.018497  3.190801
## sample estimates:
## mean in group 0 mean in group 1
##        42.45469        41.36854
```

```r
# Checking Categorical Vars
table_asa <- table(laryngoscope_data$Randomization, laryngoscope_data$asa)
table_asa
```

```
##
##      2  3  4
##   0  7 40  2
##   1 15 32  3
```

```r
fisher.test(table_asa)
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  table_asa
## p-value = 0.1448
## alternative hypothesis: two.sided
```

```r
table_mall <- table(laryngoscope_data$Randomization, laryngoscope_data$Mallampati)
table_mall
```

```
##
##      1  2  3  4
##   0 14 21 13  0
##   1 21 18  7  4
```

```r
fisher.test(table_mall)
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  table_mall
## p-value = 0.05904
## alternative hypothesis: two.sided
```

```r
table_gen <- table(laryngoscope_data$Randomization, laryngoscope_data$gender)
table_gen
```

```
##
##      0  1
##   0 39 10
##   1 39 11
```

```r
chisq.test(table_gen)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_gen
## X-squared = 2.4044e-30, df = 1, p-value = 1
```

```r
# All of them look good except for our Mallampati variable
# There is no person in the control group with Mallampati score of 4
```

Interpretation:

Nothing was registered as significant by any of these tests except for that earlier note. That there appears to be some group differences due to the amount of participants with a Mallampati rating of 4. I wouldn't think that this would impact our study too much, but being aware of this is helpful for designing future studies.

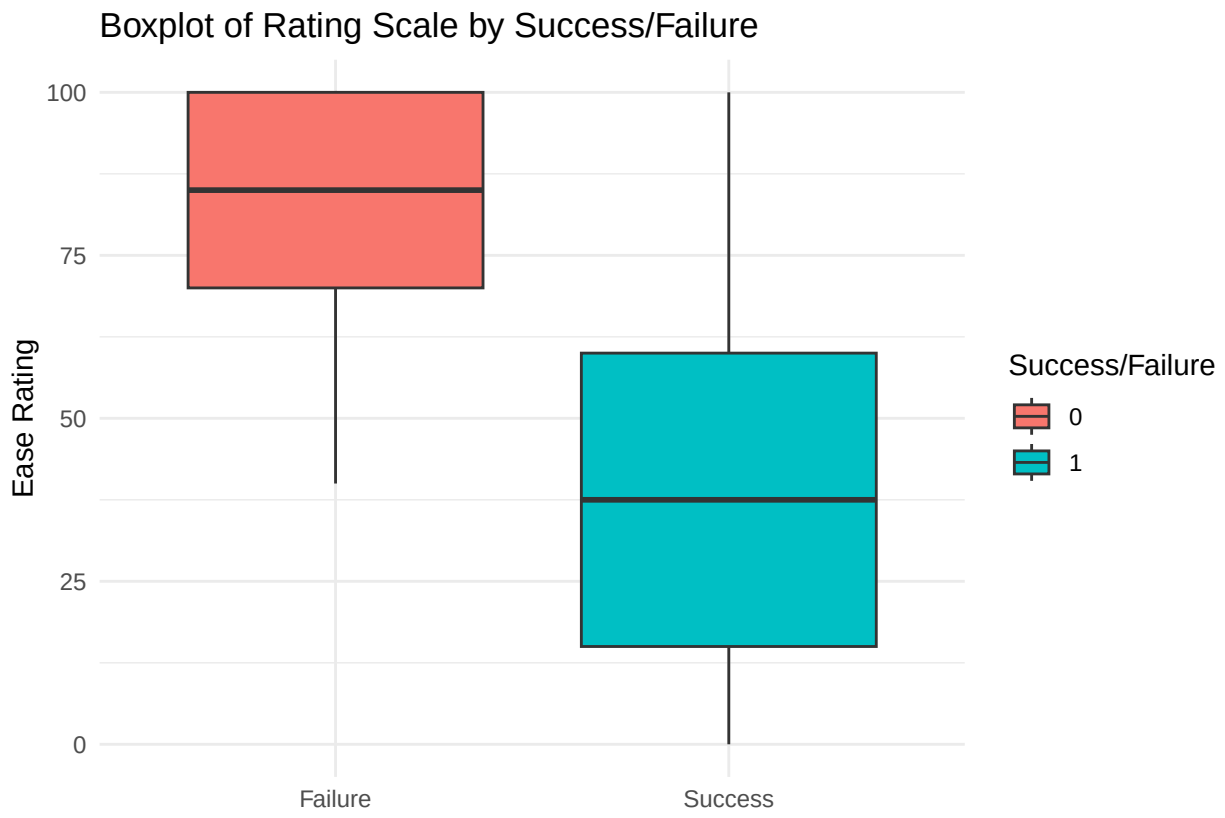
Understanding Doctors Ease-Rating

Here I want to see how many people succeed/ failed on the first attempt. I also want to see how this influences each doctors ease-rating.

```r
# Everyone who failed on the first attempt reported that the ease was difficult
# Is there a differnce here?


# Counting Amount of Successes on the first attempt #

laryngoscope_data %>%
  summarise(
    success = sum(attempt1_S_F == 1),
    failure = sum(attempt1_S_F == 0)
  )
```

```
##   success failure
## 1      88      11
```

```r
# Looking at relation between ease and succession of first attempt

laryngoscope_data %>%

  ggplot(aes(x = factor(attempt1_S_F), y = ease, fill = factor(attempt1_S_F))) +
  geom_boxplot() +
    scale_x_discrete(labels = c("0" = "Failure", "1" = "Success")) +  # Custom labels
  labs(
    title = "Boxplot of Rating Scale by Success/Failure",
    x = "",
    y = "Ease Rating",
    fill = "Success/Failure"
  ) +
  theme_minimal()
```



Interpretation: 11 people failed on the first attempts while the remaining 88 succeeded. Nearly all those who failed stated that the intubation was very difficult, which most who succeeded rated it comparatively easier. Later we'll test to see if this has relation to what randomization group the particiapants were in.

Dependence of Intubation Stlye on Success/Failure:

We'd like to use the fisher exact test to note any differences between between intubation groups and the success rate. So I'll create a table to conduct this test.

```
table_laryngoscope <- table(laryngoscope_data$Randomization, laryngoscope_data$attempt1_S_F)
dimnames(table_laryngoscope) <- list(
  Row_Names = c("Control", "Treatment"),
  Column_Names = c("no. Fail", "no. Succ")
)

# Testing dependence of Randomization with succ/fail

chisq.test(table_laryngoscope)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_laryngoscope
## X-squared = 0.36494, df = 1, p-value = 0.5458
```

```
fisher.test(table_laryngoscope)
```

```
##
```

```
##  Fisher's Exact Test for Count Data
##
## data:  table_laryngoscope
## p-value = 0.5246
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1098579 2.3436553
## sample estimates:
## odds ratio
##  0.5493335
```

```
table_laryngoscope
```

```
##            Column_Names
## Row_Names   no. Fail no. Succ
##   Control          4       45
##   Treatment        7       43
```

```
# Pval is large, fail to reject Ho
# intubation type does not affect success rate
```

Interpretation:

The fisher test reported a small p-value, so don't have much evidence to suggest differing successs rate by intubation type. In other words there is no relation between the amount of successful/failed intubations and the intubation type.

Comparing Time to Intubation:

Similarly, we want to compare intubation times for each group. For now we're only going to look at the first attempt. So, I will ignore all those who failed on the first attempt. To start, I'll test to see if the attempt times are normally distributed. If they are, I'll use a t-test to compare intubation times. If not, I'll use a wilcox test. One of our main questions is if the treatment will lead to faster intubation times, so lets test it out!

```
# sub grouping to only those who completed the first attempt
lary <- laryngoscope_data[laryngoscope_data$attempt1_S_F == 1,]

shapiro.test(lary$attempt1_time)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  lary$attempt1_time
## W = 0.9442, p-value = 0.000871
```

```
# Normality test failed

# Testing to see if there is a differnce of intubation times based on intubation
# type
# Testing to see if the intubation times are different by group
wilcox.test(attempt1_time ~ Randomization, data = lary)
```

```
## Warning in wilcox.test.default(x = DATA[[1L]], y = DATA[[2L]], ...): cannot
## compute exact p-value with ties
```

```
##
```

```
##  Wilcoxon rank sum test with continuity correction
##
## data:  attempt1_time by Randomization
## W = 327.5, p-value = 9.211e-08
## alternative hypothesis: true location shift is not equal to 0
```

```
# Pvalue is very small, suggesting that there is a difference in intubation time
#
```

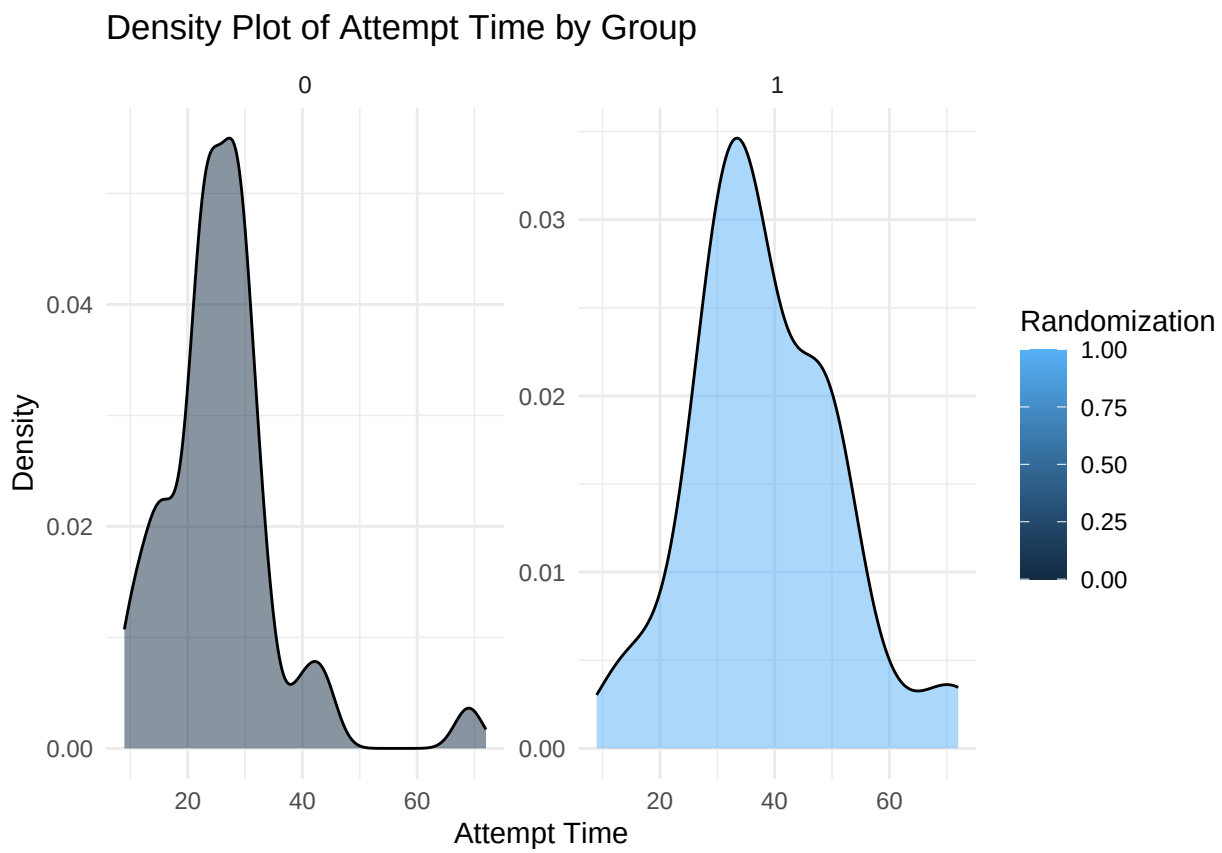
Interpretation:

The shapiro test reported a very low p-value, directing me towards the wilcox test (which handles non-normality). Likewise, the wilcox test reported a very low p-value. This suggests that there is a difference in the time to intubation based on which group the participants were in.

Visualizing Intubation Time

Lets get a better picture of what the distribution of the intubation times look like. I'll also display the means and standard deviations for each distribution. This gives us a numerical understanding of the shape of the distributions. Then, I'll verify my interpretation of the two distributions by computing an effect size using the pooled-sd. This will tell us which group had a faster intubation times.

```
ggplot(lary, aes(x = attempt1_time, fill = Randomization)) +
  geom_density(alpha = 0.5) +
  facet_wrap(~ Randomization, scales = "free_y") +  # Separate plots side by side
  labs(title = "Density Plot of Attempt Time by Group",
       x = "Attempt Time",
       y = "Density") +
  theme_minimal()
```




```
# Shoutout STAT100
# Reminder: lary is a subgroup where we only show those who completed
#the first attempt
lary %>%
  group_by(Randomization) %>%
  summarise(
    Mean = mean(attempt1_time, na.rm = TRUE),
    SD = sd(attempt1_time, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 3
##   Randomization  Mean    SD
##           <int> <dbl> <dbl>
## 1             0  25.7  10.1
## 2             1  38.4  12.7
```

```
cohen.d(lary$attempt1_time ~ lary$Randomization)
```

```
## Warning in cohen.d.formula(lary$attempt1_time ~ lary$Randomization): Cohercing
## rhs of formula to factor
```

```
##
## Cohen's d
##
## d estimate: -1.105622 (large)
## 95 percent confidence interval:
##      lower      upper
## -1.5607826 -0.6504608
```

Interpretation:

Visually, the distributions' shape look roughly similar. But, the control group appears to have a mean that is much smaller than the treatment group. Cohen's D confirms this. We get an effect size of -1.106. This tells us that the control group reported much faster intubation times. Maybe we should implement the new intubation method. I think this is the most important result in this rmd.

Predicting Success/Failure and Intubation Time

To get a better idea how how all the baseline variables interplay with each other and out outcomes, we'll construct some models. First, we'll use a glm to predict S/F, because we have Bernoulli responses. Then we'll use a standard lm to predict intubation time.

```
# Checking what factors might predict our success in attempt1
# Using glm because we have binomial responses
laryngoscope_data <- laryngoscope_data[!is.na(laryngoscope_data$attempt1_S_F), ]

anova_model <- aov(attempt1_S_F ~ as.factor(Mallampati)+ as.factor(asa) + as.factor(gender), data = lary
summary(anova_model)
```

```
##                       Df Sum Sq Mean Sq F value   Pr(>F)
## as.factor(Mallampati)  3  0.346  0.1155   1.324 0.271479
## as.factor(asa)         2  0.054  0.0270   0.309 0.734728
## as.factor(gender)      1  1.428  1.4276  16.367 0.000109 ***
## Residuals             91  7.937  0.0872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 1 observation deleted due to missingness
```

```
logit_model <- glm(attempt1_S_F ~ age+ gender + BMI + as.factor(asa) + as.factor(Mallampati) + as.facto
                   data = laryngoscope_data, family = binomial)
summary(logit_model)
```

```
##
## Call:
## glm(formula = attempt1_S_F ~ age + gender + BMI + as.factor(asa) +
##     as.factor(Mallampati) + as.factor(Randomization), family = binomial,
##     data = laryngoscope_data)
##
## Coefficients:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                -3.57995    5.10261  -0.702   0.4829
## age                         0.03652    0.03614   1.011   0.3122
## gender                     -2.96975    0.92956  -3.195   0.0014 **
## BMI                         0.10934    0.10012   1.092   0.2748
## as.factor(asa)3            -0.30700    1.00910  -0.304   0.7610
## as.factor(asa)4            -0.23250    1.98799  -0.117   0.9069
## as.factor(Mallampati)2      1.67129    0.99179   1.685   0.0920 .
## as.factor(Mallampati)3      1.48439    1.14667   1.295   0.1955
## as.factor(Mallampati)4     -1.51380    2.02921  -0.746   0.4557
## as.factor(Randomization)1 -0.19795    0.86435  -0.229   0.8189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

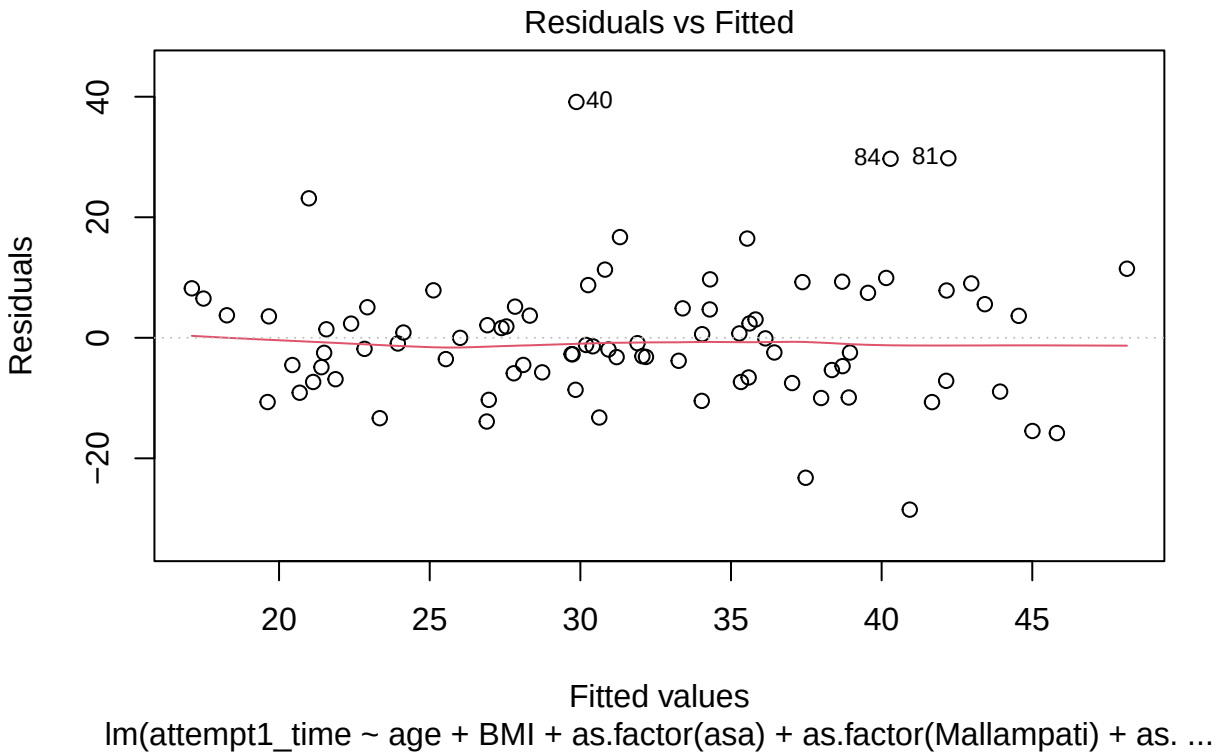
```
##     Null deviance: 68.350  on 95  degrees of freedom
## Residual deviance: 49.008  on 86  degrees of freedom
##   (3 observations deleted due to missingness)
## AIC: 69.008
##
## Number of Fisher Scoring iterations: 6
```

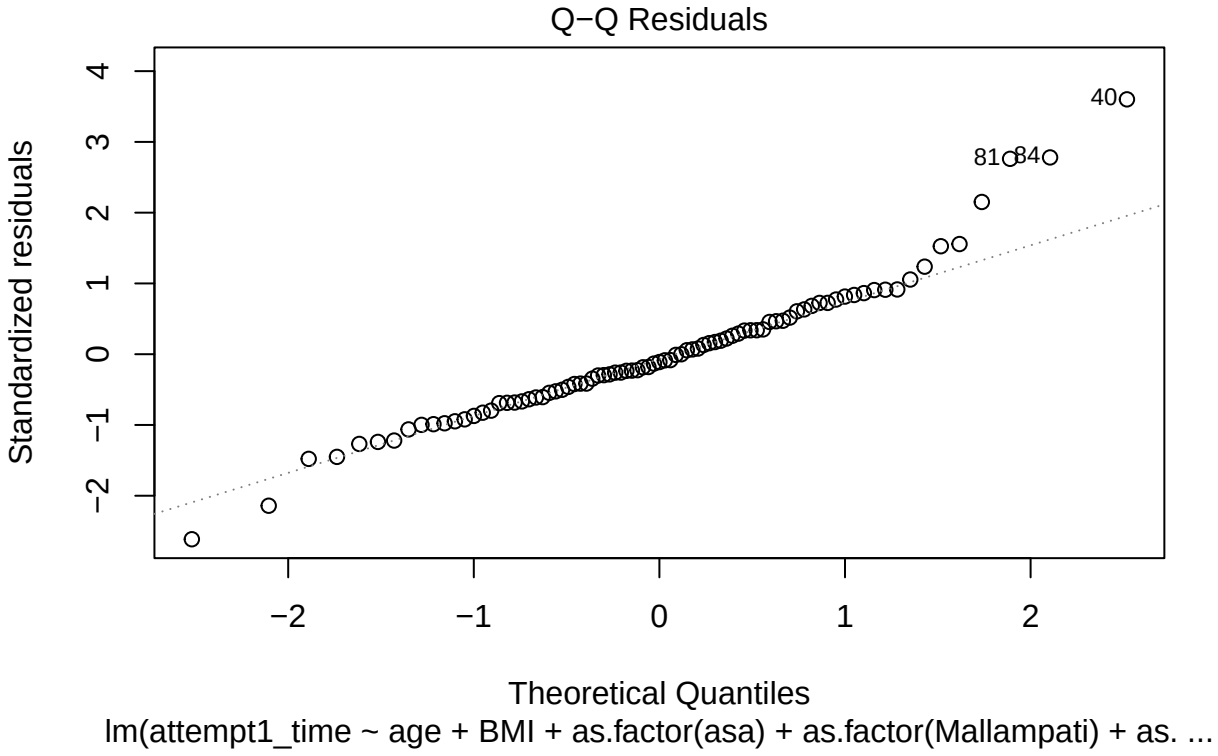
```
#plot(logit_model)
#pR2(logit_model)
#vif(logit_model)
#hoslem.test(laryngoscope_data$attempt1_S_F[1:length(fitted(logit_model))],
           # fitted(logit_model))

#AIC(logit_model)
#BIC(logit_model)
```

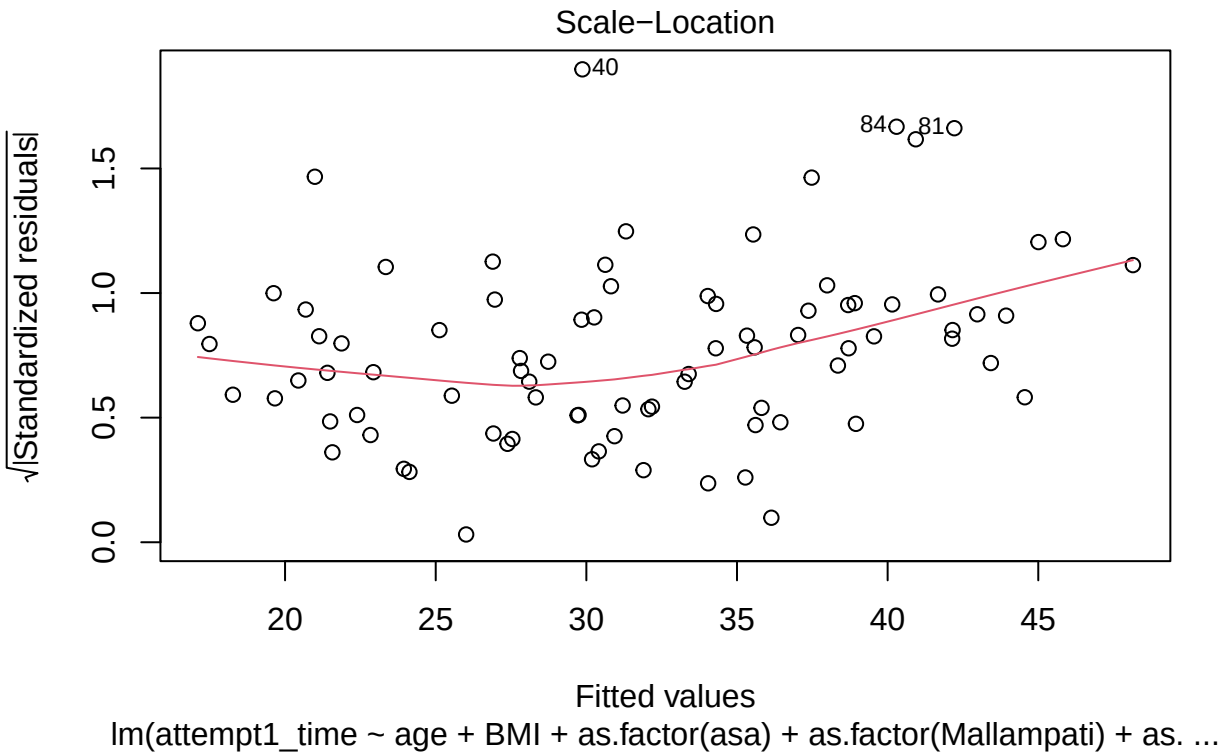
```
lm_mod <- lm(attempt1_time ~ age + BMI + as.factor(asa) + as.factor(Mallampati) + as.factor(Randomizati
summary(lm_mod)
```

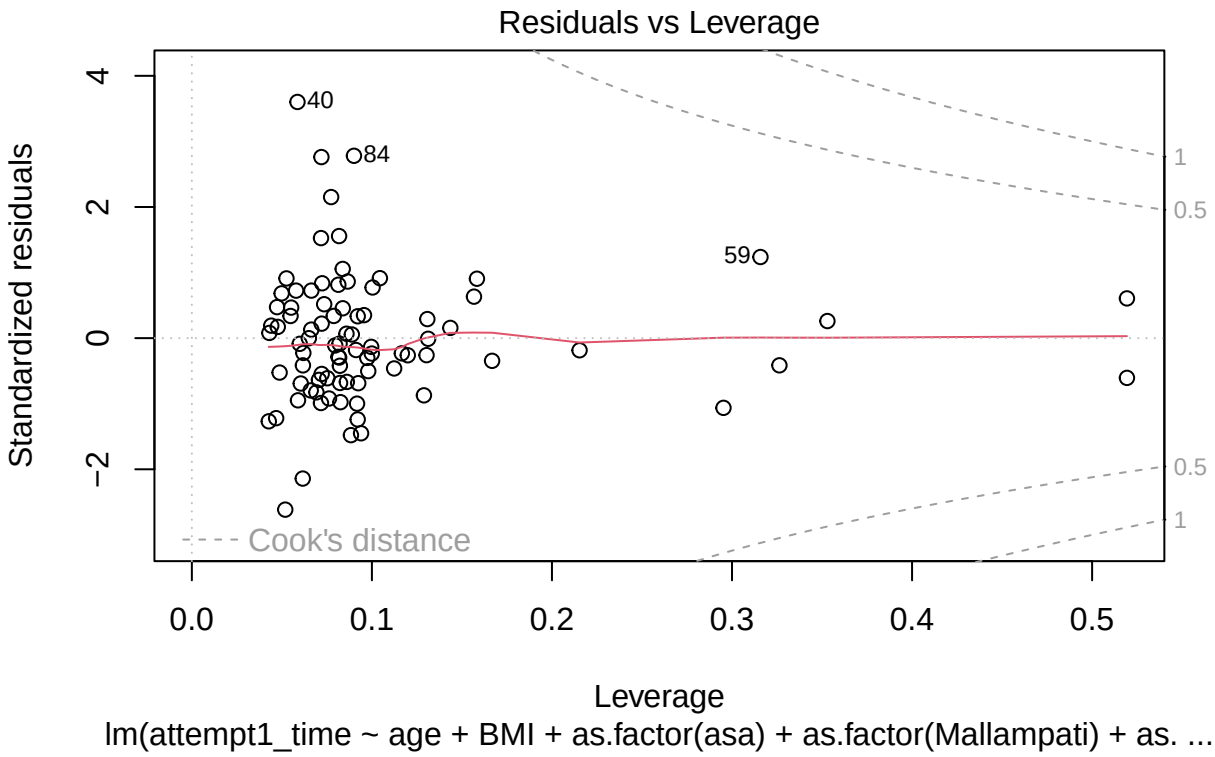
```
##
## Call:
## lm(formula = attempt1_time ~ age + BMI + as.factor(asa) + as.factor(Mallampati) +
##     as.factor(Randomization), data = lary)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -28.514  -6.586  -1.190   5.069  39.131
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  6.67154   13.26442   0.503  0.61644
## age                          0.27320    0.10153   2.691  0.00876 **
## BMI                          0.03453    0.26255   0.132  0.89572
## as.factor(asa)3              4.83764    3.06581   1.578  0.11873
## as.factor(asa)4              5.94905    6.21634   0.957  0.34160
## as.factor(Mallampati)2      -0.34105    2.87628  -0.119  0.90593
## as.factor(Mallampati)3       0.72791    3.54937   0.205  0.83806
## as.factor(Mallampati)4      -6.09344    8.88693  -0.686  0.49501
## as.factor(Randomization)1 13.56235    2.58684   5.243 1.38e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.2 on 76 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.3437, Adjusted R-squared:  0.2746
## F-statistic: 4.975 on 8 and 76 DF,  p-value: 5.641e-05
```

```
# using a linear model to predict attempt1_time
# I think a survival analysis might be more apt
# But I don't really know how to use them

plot(lm_mod)
```


Residuals vs Fitted
Residuals
Fitted values
lm(attempt1_time ~ age + BMI + as.factor(asa) + as.factor(Mallampati) + as. ...
40
84
81
−20
0
20
40
20
25
30
35
40
45


Q−Q Residuals
Standardized residuals
Theoretical Quantiles
lm(attempt1_time ~ age + BMI + as.factor(asa) + as.factor(Mallampati) + as. ...
40
81
84
−2
−1
0
1
2
3
4


Scale−Location
40
84
81
√|Standardized residuals|
0.0
0.5
1.0
1.5
20
25
30
35
40
45
Fitted values
lm(attempt1_time ~ age + BMI + as.factor(asa) + as.factor(Mallampati) + as. ...




```
#shapiro.test(residuals(lm_mod))
#boxcox(lm_mod)
#plot(cooks.distance(lm_mod))

# Shapiro test suggests that residuals aren't normal
# but lambda = 0, is within the 95% CI of our boxcox

# gender is the only significant predictor
# I did check for interactions and found no significant ones
```

Interpretation:

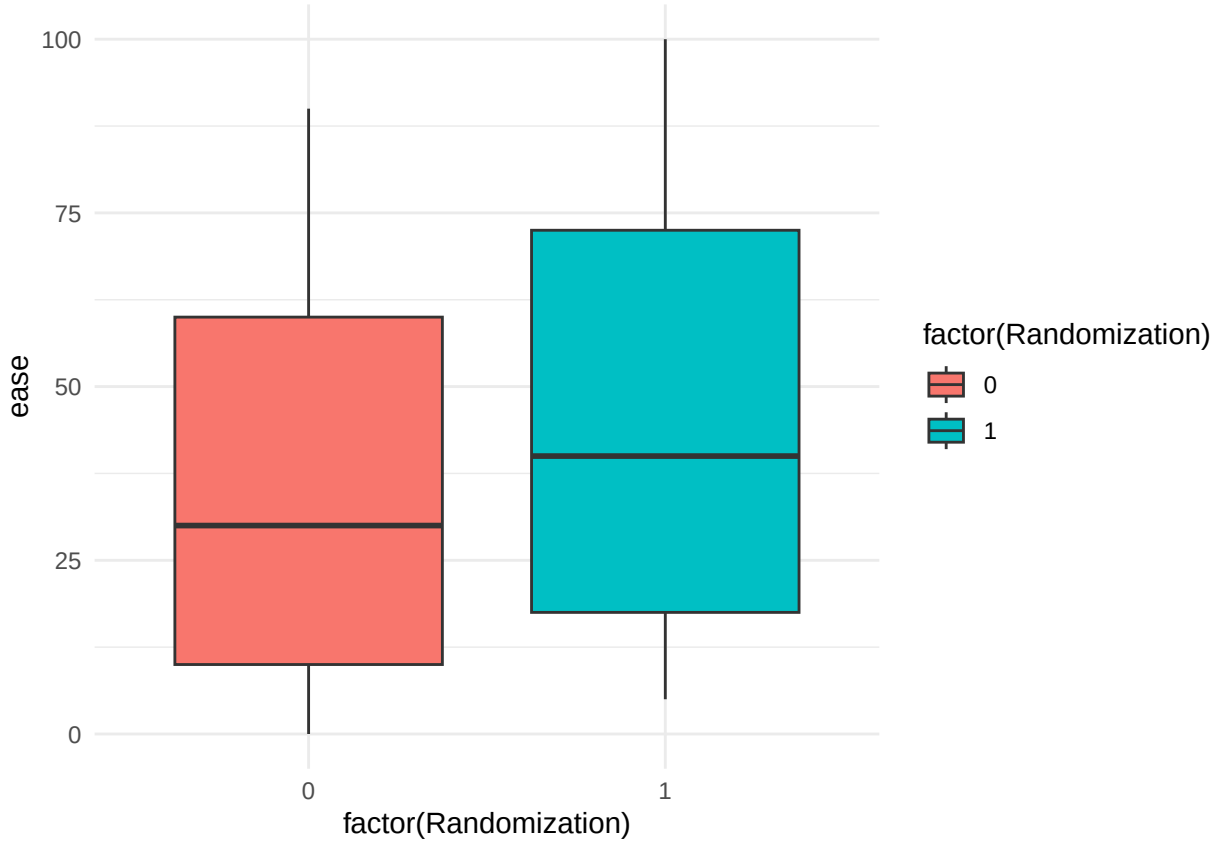
Most of the predictors for each response variable seem insignificant to prediction Gender seems to be the only significant predictor. There are no interactions either I'm not sure how useful this information is. Probably not very :)

Looking at Ease

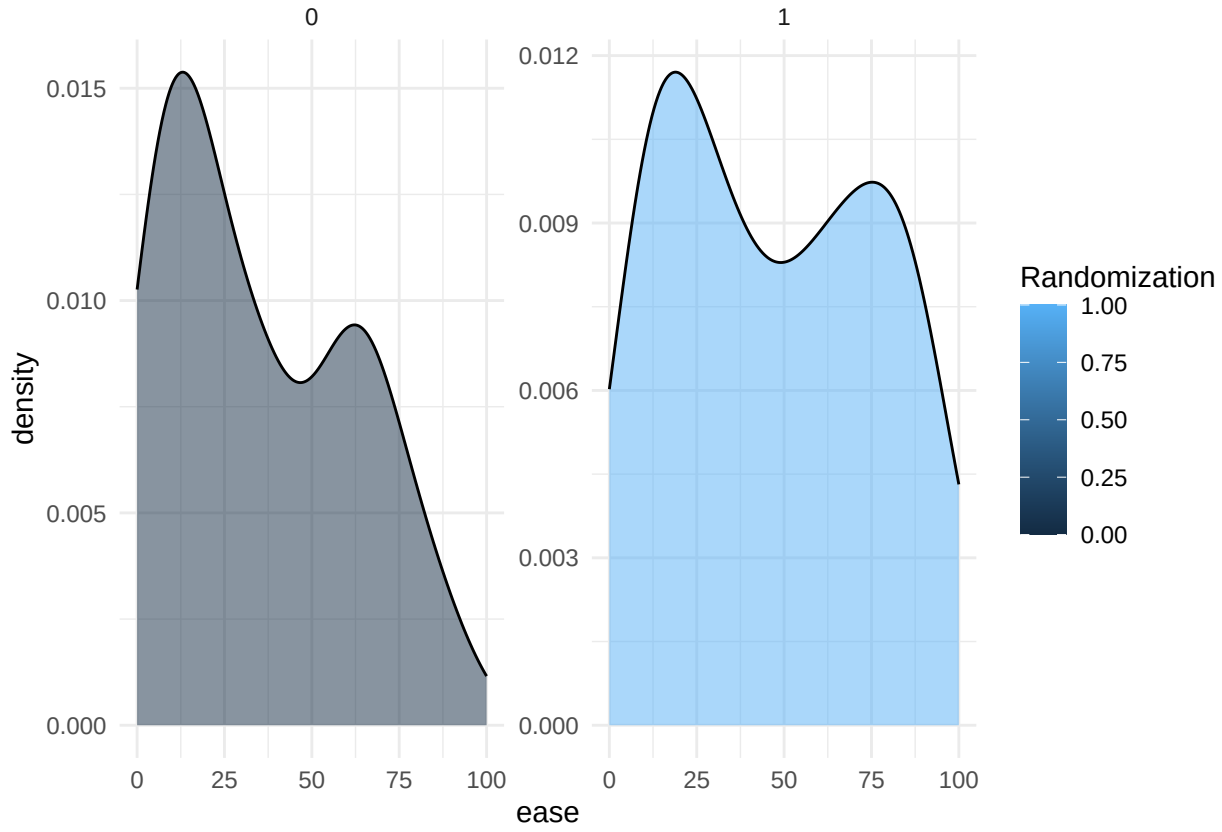
How did the doctors feel about either method of intubation? I'll calculate another effect size to determine any differences

```
lary %>%

  ggplot(aes(x = factor(Randomization), y = ease, fill = factor(Randomization))) +
  geom_boxplot() +
    #scale_x_discrete(labels = c("0" = "Failure", "1" = "Success")) +  # Custom labels
```

```
  theme_minimal()
```




```
ggplot(lary, aes(x = ease, fill = Randomization)) +
  geom_density(alpha = 0.5) +
  facet_wrap(~ Randomization, scales = "free_y") +  # Separate plots side by side
  theme_minimal()
```




```
lary %>%
  group_by(Randomization) %>%
  summarise(
    Mean = mean(ease, na.rm = TRUE),
    SD = sd(ease, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 3
##   Randomization  Mean    SD
##           <int> <dbl> <dbl>
## 1             0  34.8  26.6
## 2             1  47.0  29.5
```

```
cohen.d(lary$ease ~ lary$Randomization)
```

```
## Warning in cohen.d.formula(lary$ease ~ lary$Randomization): Cohercing rhs of
## formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.4334619 (small)
## 95 percent confidence interval:
##        lower        upper
## -0.862347333 -0.004576517
```

Interpretation:

The reported effect size was 0.433, which indicates a slight difference in ease rating by group. Those using the new intubation reported that this method was harder than the control method. Though the cohen's d CI has a lower bound of 0.005 which would indicate hardly any difference

Conclusion:

If there's no difference in the success rate, but doctors took longer and reported more difficulty when using the new intubation method, whats the point in changing what is standard practice. Based on this analysis I'd suggest using the traditional intubation method.

Redisgn Study:

To improve the internal validity of this study, I would recommend that somehow adjust for each doctor's experience level. Also it's important that each doctor have at least some training in both intubation methods. There are a couple ways to ensure this 1st method: When doctors first become elligible to conduct intubations in their training, they should be exposed to both methods 2nd method: Alternitivly, we can block doctors by experience level, and provide a training period where some doctors get to learn how the Pentax AWS works. - In both cases the doctors should be randomly assigned to each intubation method, as they were in the initial study - If we want to randomly assign the patients instead, we need to make sure the doctors have been exposed to the Pentax AWS