**Predicting and Interpreting the Odds Ratios of Hospital Admission Types**

Joshua Arrabaca - 220029955

## Introduction

This summary is intended for professionals in the medical industry, and provides a non-technical overview on the analysis conducted on hospital admissions data. This data contains almost 1,500 records of admissions, and includes patient ethnicity, age, length of stay, and whether the patient died.

In particular, this summary seeks to answer: “Can the probability of admission type be predicted by previous admissions data?”. The data was analysed in a software called R, and analysis conducted to obtain the odds ratios of the different admission types: ‘Elective’, ‘Urgent’, ‘Emergency’, as well as how the probabilities change depending on length of stay and other factors.

## Methods

The admissions data contained several columns of categorical data (i.e. data that was not numeric) such as hospital, ethnicity, age bracket, and admission type. The data was imported into R, and the following columns were cast as factors (i.e. turned into categories so R could process them): died, whether the patient died or not; white, the patient’s ethnicity; age, the patient’s age range; age80, whether the patient was at least 80 years old; and admission, whether the patient’s entry was Elective, Emergency, or Urgent.

A multinomial model was fitted; this type of model can predict categories instead of usual numerical predictions. To assess the model, a statistical tool called ANOVA was used to check which of the variables mentioned above could be dropped to improve predictive accuracy.

Predictions were computed to interpret how probabilities and odds rations of admission types relate to each other. For these, two predictions were done: one where the patient lived, and the other where they died. For both, the ethnicity was set to ‘White’, and Length of Stay set to the median of 8 days.

Effect plots was also created. This type of plot shows how the admission predictions can change over the length of stay, ethnicity, age, and died.

## Results and Recommendations

The initial model was comprised of all the above variables. However, the ANOVA showed that the age variables were not useful in the model and were therefore dropped. The model was updated to have the ff:

- **Response Variable:** Admission Type
- **Predicting Variables:** Died, White, Length of Stay

The resulting predictions show that the odds ratio when a patient died was 9.93 for Elective admission compared to Emergency (i.e. it's 9.93 more likely that a patient is admitted Electively rather than via Emergency if the patient died). However, the odds of Elective admission to Urgent was only 3.78, while Emergency to Urgent was 1/ 0.381 (or Urgent to Emergency was 2.62).

If a patient lived, the odds ratios significantly change. Here, the odds ratio of Elective to Emergency admission becomes 23.16 (i.e., it's 23.16 more likely that a patient is admitted electively rather than via emergency if the patient lived). The odds the odds of elective admission to urgent becomes 5.78, and Emergency to Urgent was 1/0.25 (Urgent to Emergency was 4).

The effect plots showed that Elective admissions were around 80% when length of stay is 0, but gradually decreases as length of stay increases. Admissions are also predominantly White for Elective admissions, while Death has some effect overall on admission type (higher deaths for Urgent and Emergency).

Therefore, hospital administrators and staff should prepare more for Elective admissions over Urgent or Emergency admissions, and if patients are do admitted via Emergency, patients who live tend to have a longer length of stay compared to other admission types.

# Practical 4 Technical Notes

Joshua Arrabaca - 220029955

2023-11-24

A multinom model was fitted using the nnet library The model had `admissions` as the response, and `died`, `white`, `los`, `age`, and `age80` as the predictors. ANOVA was conducted on the model, and it showed that age has p-values of 0.74 and 1.00 respectively. These were dropped and an updated model fitted.

The new model only had `died`, `white` and `los` as the predictors, and these covariates had significant p-values (at least on the 0.01 level).

Predictions were obtained using the median of `los` (which is 8.0), and used to predict the odds ratios when patients `died='0'` and `died='1'`. For these two sets of predictions, white was set to `'1'`, and los to `median(dat$los)`.

Effects plot were also created using the `effects` library and `plot(effect(covariate, model))`.

To check the model assumption of independence, plots were created using:

```
‘plot(residuals(model)[,columnNumber])‘
```

where columnNumber referred to each of the column numbers for ‘Elective’, ‘Urgent’, and ‘Emergency’.

On checking the model assumptions, the linear relationship assumption is satisfied (when plotting the log odds of the outcomes against los). Independent observations also seems to hold true as residual plots show the values are fairly close to each other.

Refer to the tables, plots and code outputs in the succeeding pages.

## Loading the Data

The data is loaded, and the following is the dataframe head.

```
##   provnum died white los age80 age admission
## 1   30001    0     1   4     0   4  Elective
## 2   30001    0     1   9     0   4  Elective
## 3   30001    1     1   3     1   7  Elective
## 4   30001    0     1   9     0   6  Elective
## 5   30001    1     1   1     1   7  Elective
## 6   30001    1     1   4     0   5  Elective
```

The following show the levels of admission

```
## [1] "Elective"  "Emergency" "Urgent"
```

## Model Fitting

This model was fitted with the following call:

```
## multinom(formula = admission ~ died + white + age + age80 + los,
##     data = dat, trace = FALSE)
```

## ANOVA for original model

ANOVA was run on this model. Note that age and age80 have non-significant p-values:

```
## Analysis of Deviance Table (Type II tests)
##
## Response: admission
##       LR Chisq Df Pr(>Chisq)
## died    20.076  2   4.37e-05 ***
## white   11.918  2   0.002582 **
## age     12.118 16   0.735818
## age80    0.000  2   1.000000
## los     80.773  2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Updating model

The model is refitted to remove the age covariates:

```
## multinom(formula = admission ~ died + white + los, data = dat,
##     trace = FALSE)
```

## ANOVA on Updated Model

ANOVa was run again on the updated model, and the resulting ANOVA shows signficant p-values for ‘died’, ‘white’, and ‘los’.

```
## Analysis of Deviance Table (Type II tests)
##
## Response: admission
##       LR Chisq Df Pr(>Chisq)
## died    19.519  2  5.774e-05 ***
## white   10.888  2   0.004323 **
## los     81.637  2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Predicted Probabilities in the Link Function

When predictions are made, the following probabilities are obtained.

```
##    Elective  Emergency    Urgent
## 1 0.7327583 0.07378399 0.1934577
## 2 0.8223260 0.03549518 0.1421789
```

## Odds Ratios for Patients Who Died

The above probabilities are then taken as ratios, the log of each ratio computed, then the result is exponentiated. The following are the odds ratios for ‘died=1’ for ‘Elective to Emergency’, ‘Elective to Urgent’, and ‘Emergency to Urgent’ respectively:

```
## [1] "9.93112856606233 3.78769184555286 0.381395912897206"
```

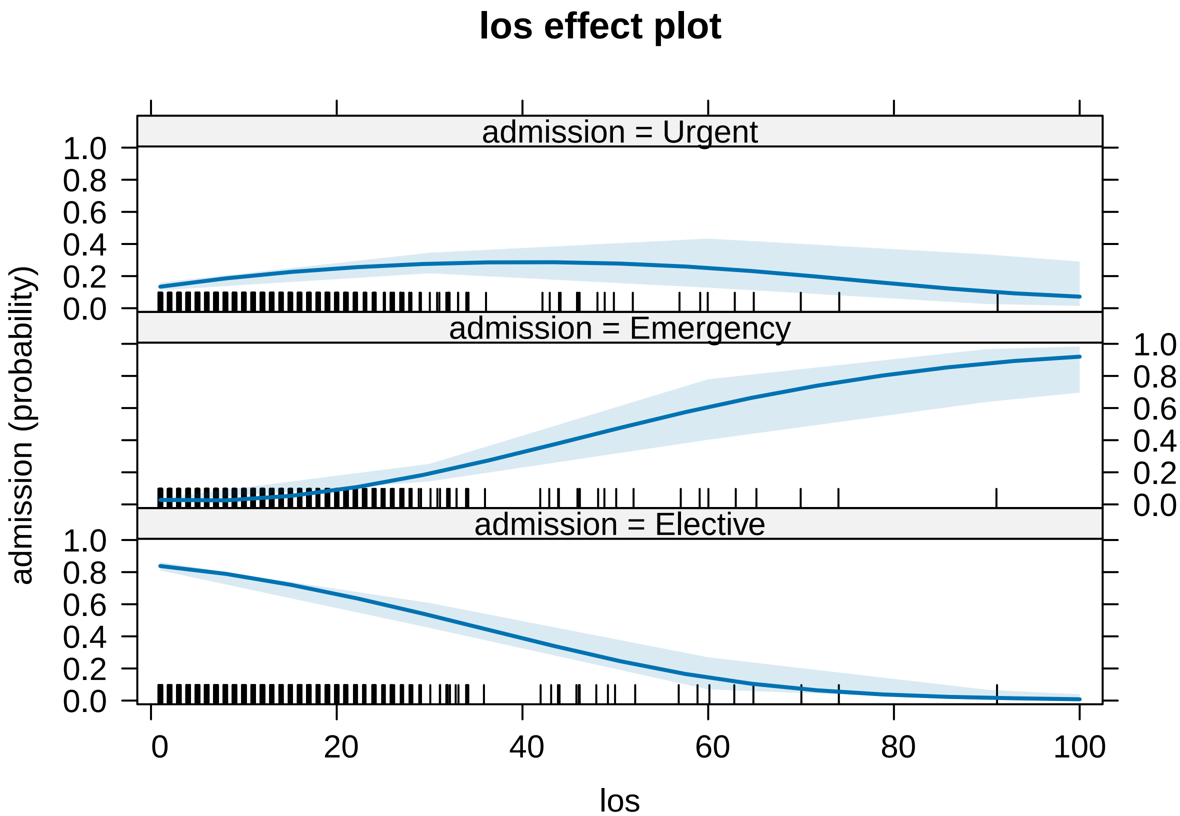
## Odds Ratios for Patients Who Lived

Here, the same ratios are computed, but this time for ‘Died=0’. These are the odds ratios for ‘Elective to Emergency’, ‘Elective to Urgent’, and ‘Emergency to Urgent’ respectively:

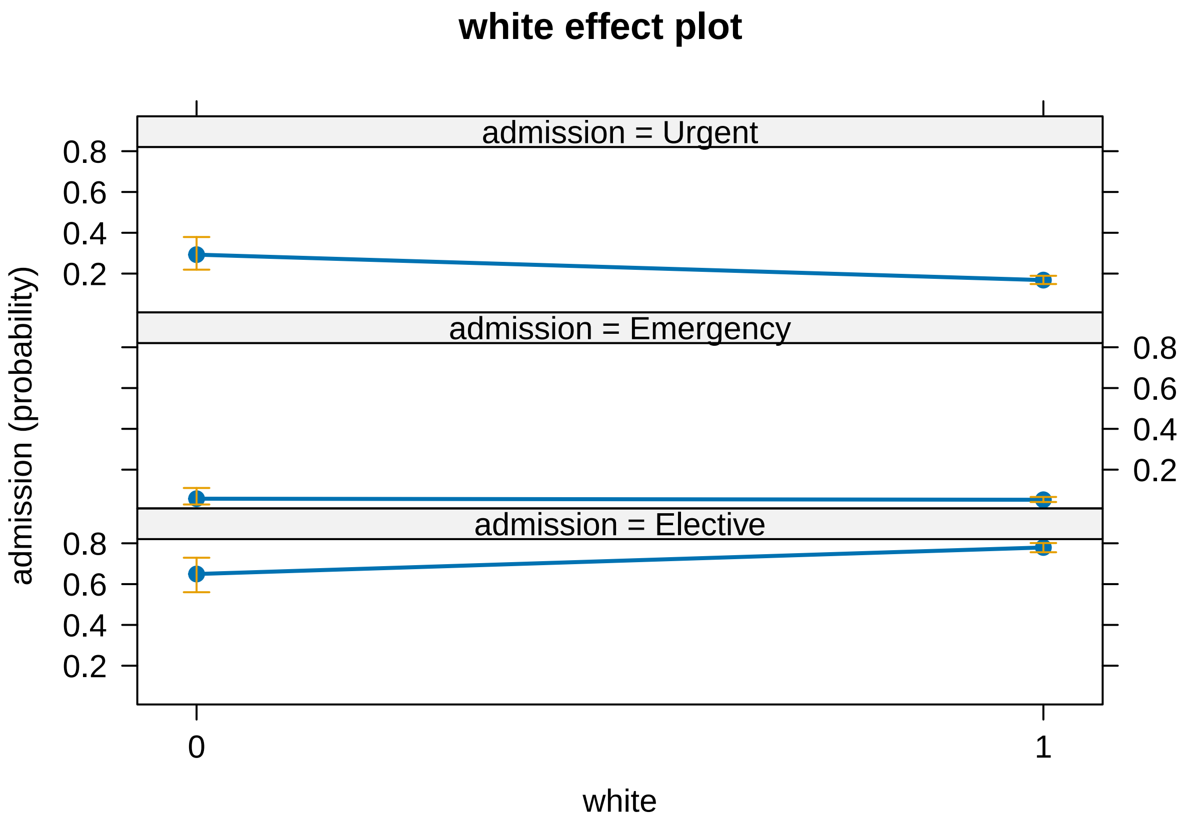
```
## [1] "23.1672567802009 5.7837426014532 0.249651594762659"
```

### Effects Plots

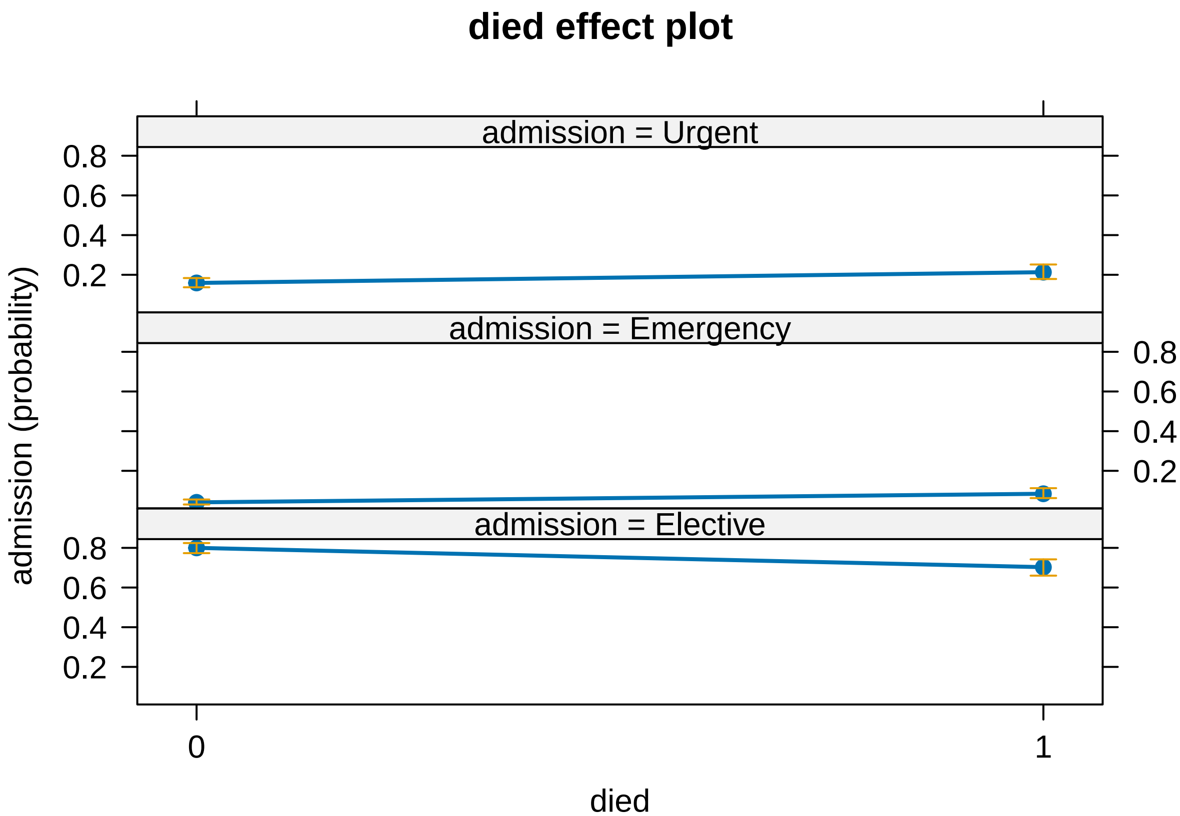
Plotting against `los` shows that Elections are at around 80% when los=0, but gradually decreases. Meanwhile, the probability of Emergency admission increases as length of stay increases. Urgent admissions only slightly increase to 20% before decreasing over the length of time.



There is a slight effect of ethinicity here; patients who identify as ‘white’ tend to be admitted via ‘Elective’. Non-white patients tend to be admitted as ‘Urgent’.

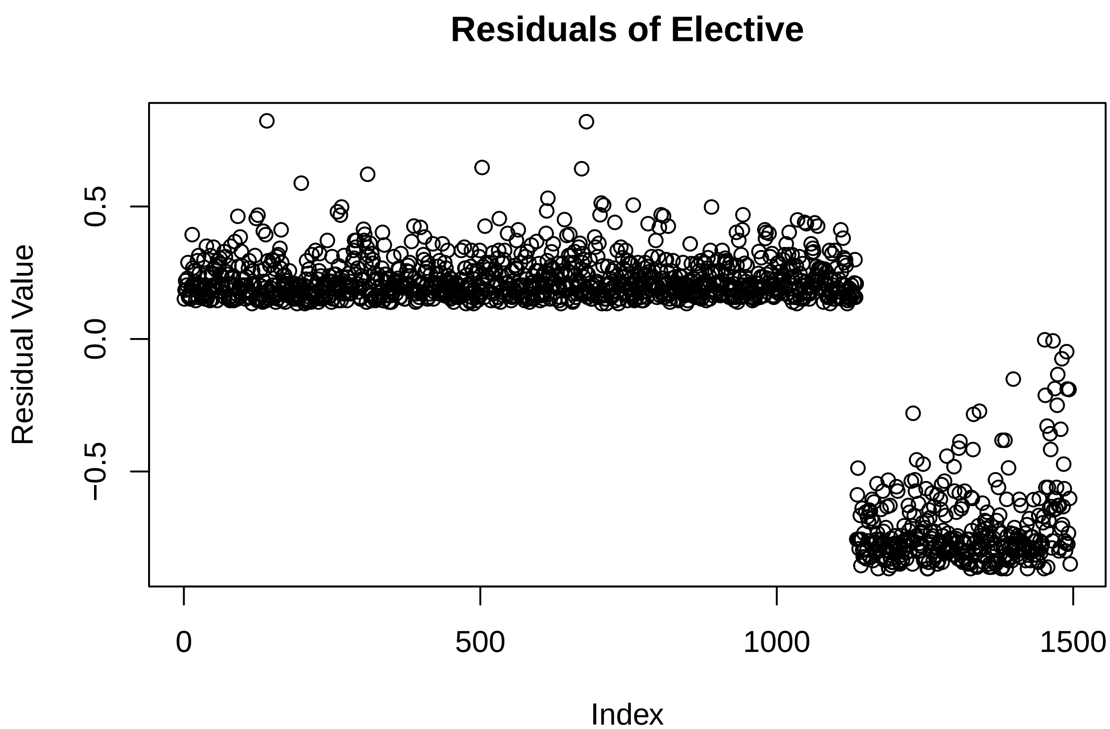


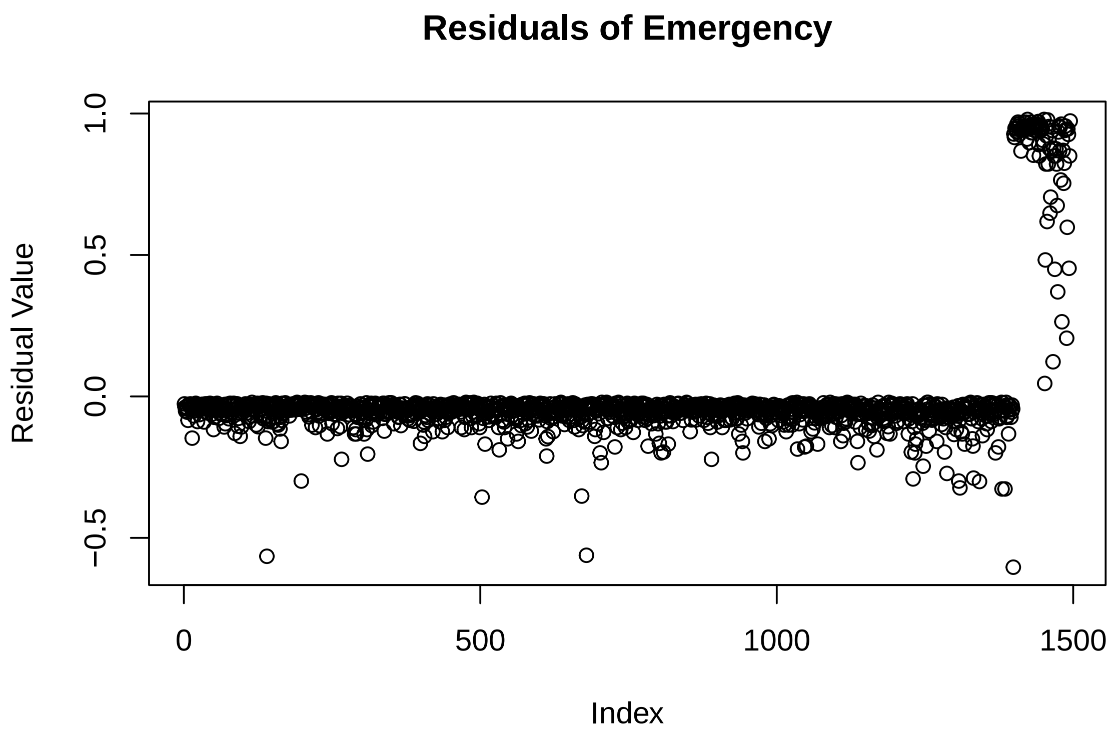
There is also a slight effect for `died`, where 'Urgent' and 'Emergency' admission have slightly higher incidence of patients with died=1, compared to 'Elective' admissions.

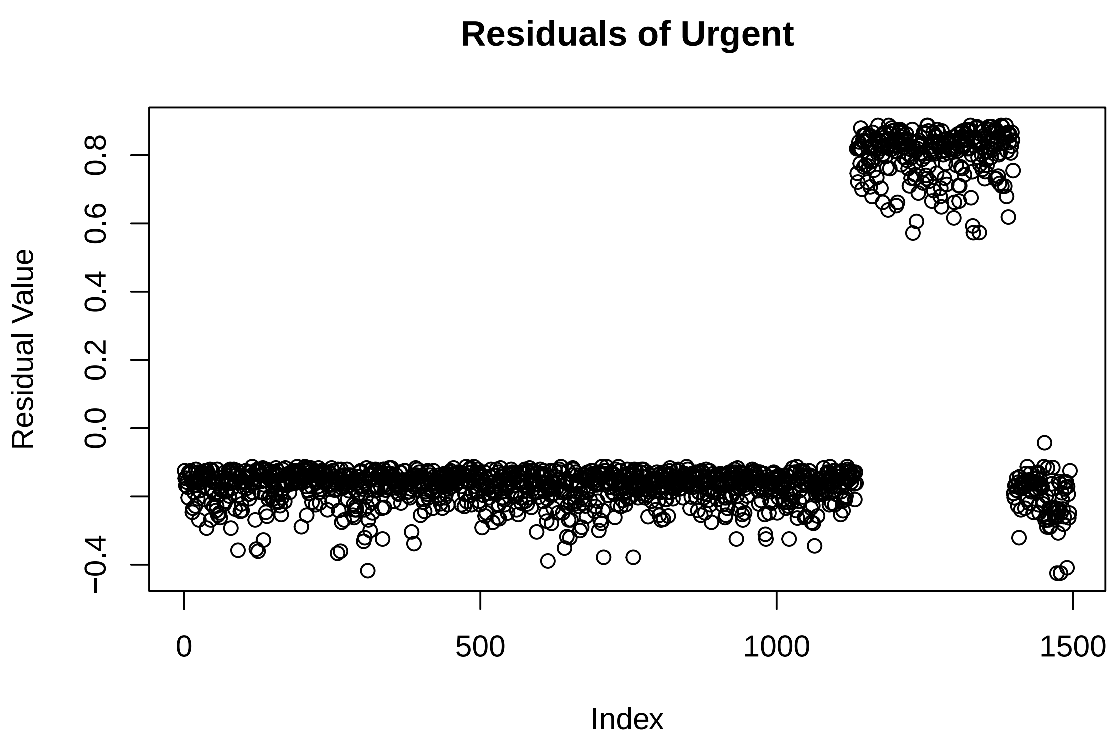


## Model Assumptions

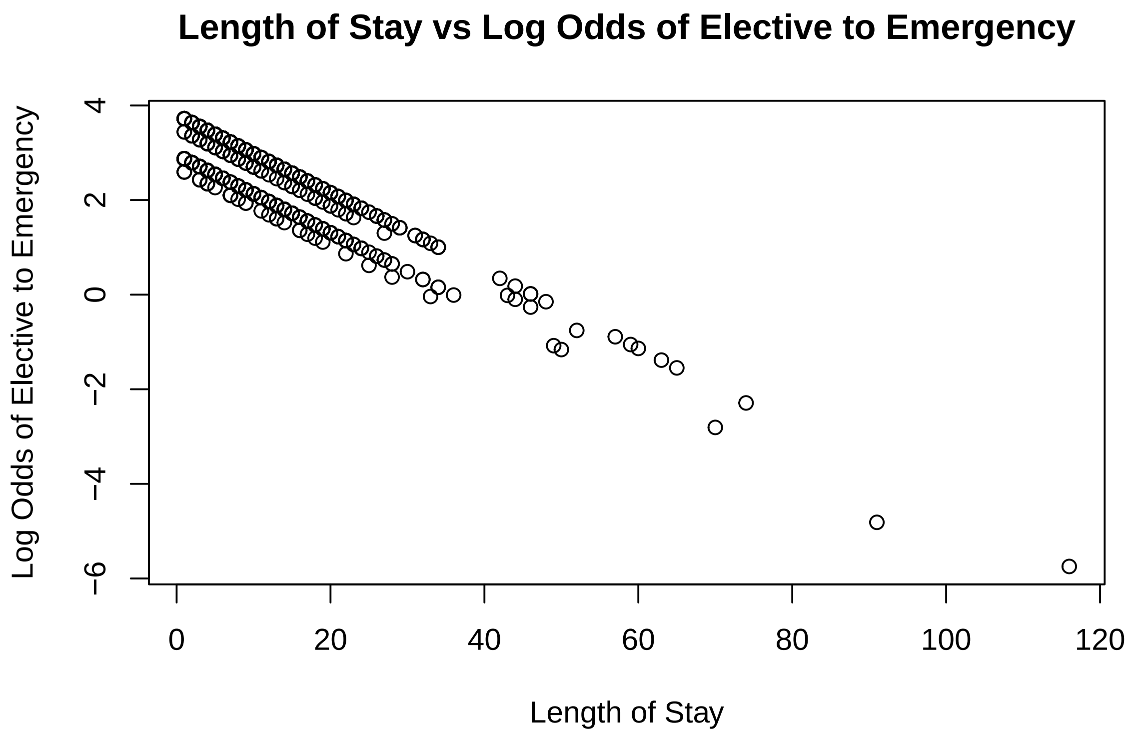
- **Independent observations -** This model assumes that each admission did not affect other admissions. In the below plots, most of the residuals are grouped fairly close, with a few residuals stretching out.

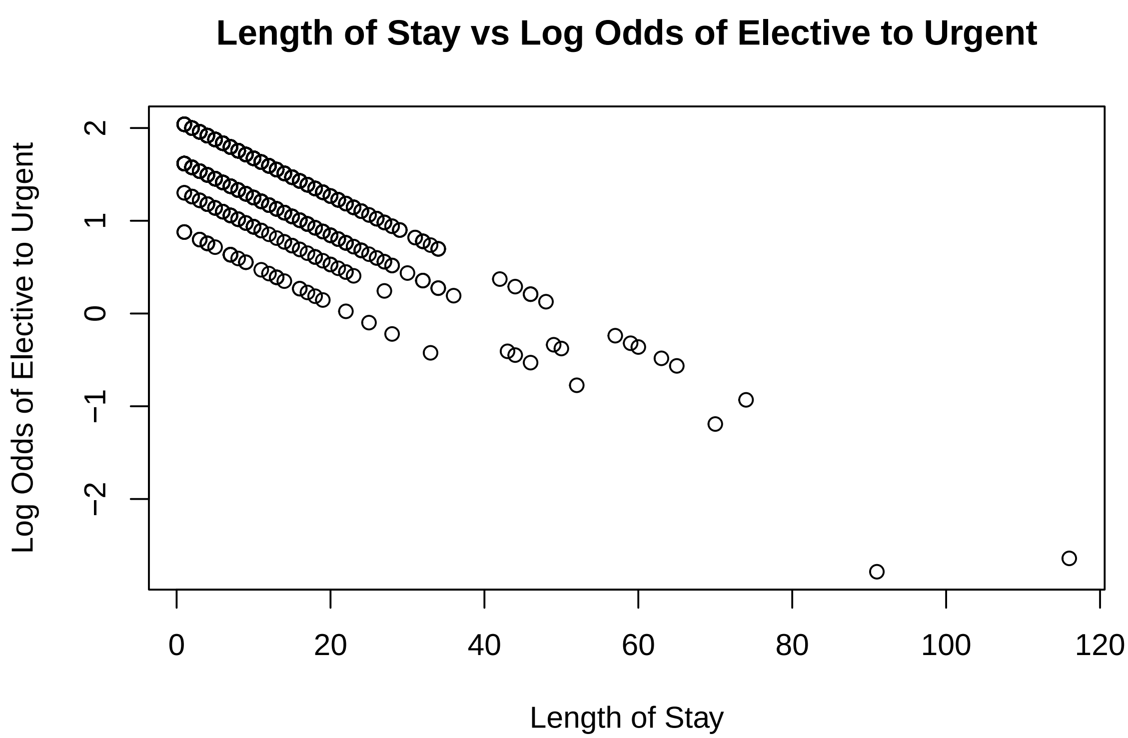


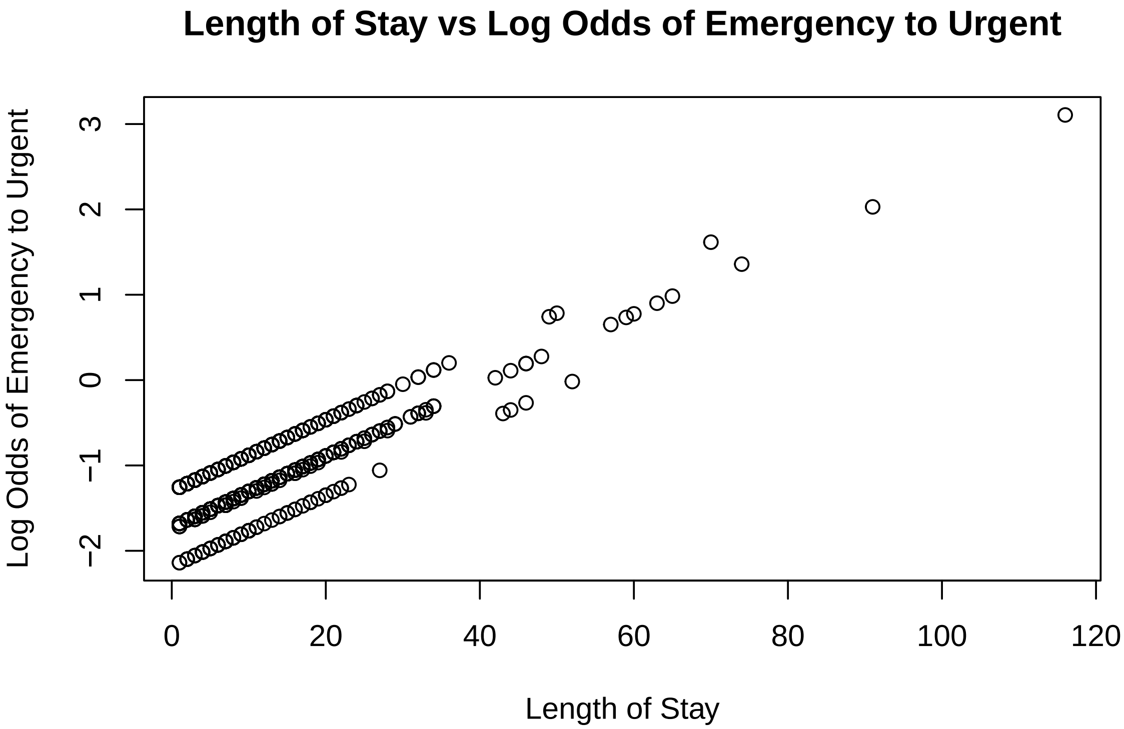

Residuals of Emergency
Residual Value
Index
1.0
0.5
0.0
−0.5
0
500
1000
1500


Residuals of Urgent
Residual Value
Index
0.8
0.6
0.4
0.2
0.0
−0.4
0
500
1000
1500

- **Linear relationship with covariates, on log odds scale -** When plotting Length of Stay against the Log odds of the different admissions, a linear relationship can be seen in the plots.






- **Independence from Irrelevant Alternatives (IIA)** - As with the above, it is assumed that admission odds do not affect other admission options