School of Mathematics and Statistics

University of St Andrews

Executive Summary

Investigation of changes in the probability of patients' admission to hospitals.

This report examines the findings of the research about how the probability of type of hospital admission changes when patients are admitted to hospitals based on personal choice (electively), as an urgent case (referred by General Practice), or as an emergency (by ambulance). For this purpose, factors such as age, race (whether white or non-white), length of stay and subsequent death are recorded in the admission data set to predict the different types of hospital admission. The target audience of this report is hospital administrators who are interested in analysis of changes in patients' admission to hospitals under certain situations. An equation is formulated to summarize how each variable recorded in the dataset affects the probability of hospital admission. This report uses a nominal multinomial model to fit the regression model as it accommodates dependent variables with multiple categories. Our model excludes hospital identification number as a variable from the model since it doesn't provide valuable information about the patients.

Preliminary findings suggest that the possibility of patients accessing the hospital both in emergency and urgent cases given that they died, with 95 percent confidence, is likely to increase by 2.33 and 1.52 times respectively compared with elective. Moreover, the possibility of white patients being admitted to hospital, in both emergency and urgent categories, tends to decrease by 24.25 % and 52.20% respectively with respect to elective (see table 1 in Appendix). Similarly, the model estimate indicates that there is a positive association between patients' length of stay and both urgent and emergency types. Our findings further reveal that the possibility of length of stay (in days) in hospital for white and non-white patients who either died or survived is likely to increase under emergency circumstances. That said, the length of stay for white and non-white patients referred by the local GP tends to increase only for the first 40 days, thereafter a negative trend is detected from the model (see Figure 1 in Appendix).

Even though this report used the best model, some unexplained patterns are still detected in the dataset. The proportions of our predictions appear to be overestimated for patients being admitted to hospital based on personal choice (elective) and overestimated for those patients who visited under emergency and urgent cases (see Figure 2 in Appendix). Furthermore, statistical assessment of explanatory variables points to the presence of strong dependencies between race and subsequent death. As a result the predicted estimates are unstable and have high errors. Additionally, the diagnosis of underlying assumptions of our model illustrates that samples do not follow a natural ordering and as a consequence observations appear to be dependent. However, model assessment of the sample data confirms that it represents the dataset satisfactorily. This report further investigated independence of our samples from irrelevant alternatives outcomes. Our findings conclude that additions of choices to the types of hospital admission available will not change the likelihood of choosing *elective, urgent and emergency* cases by an equal proportion, since patients are referred to hospital by general practitioners in urgent cases and transported to hospital in emergency circumstances.

This report therefore is unsure about the conclusion drawn. Also there are more explanatory variables not captured in the current dataset that could present greater predictive power. Examples include the gender, specific group age, health conditions of patients, geographical location of hospitals and time period which are not specified in the data. Moreover, the type of hospital admission appears to be an invalid measure of the true emergency of patients. If anything it is only a measure of the access routes to hospitals. Thus, it presents a misleading perception of patients' trends under emergency cases. Furthermore, the dataset provides no information about the actual urgency of admissions. Considerations towards confounding variables like these are essential to generate inferences for a wider population from our sample.

Appendix

Table 1: Coefficients of Predicted Model

	(Intercept)	Died	White	Length of Stay
Emergency	0.02	2.33	0.75	1.08
Urgent	0.26	1.52	0.47	1.04

Table 2: Differences in Predicted and Original counts

Elective	*Emergency*	*Urgent*
135.11	−56.58	−78.53

Figure 1: Prediction of Patients' Length of Stay. The solid line is elective, dashed is emergency and dotted is urgent.

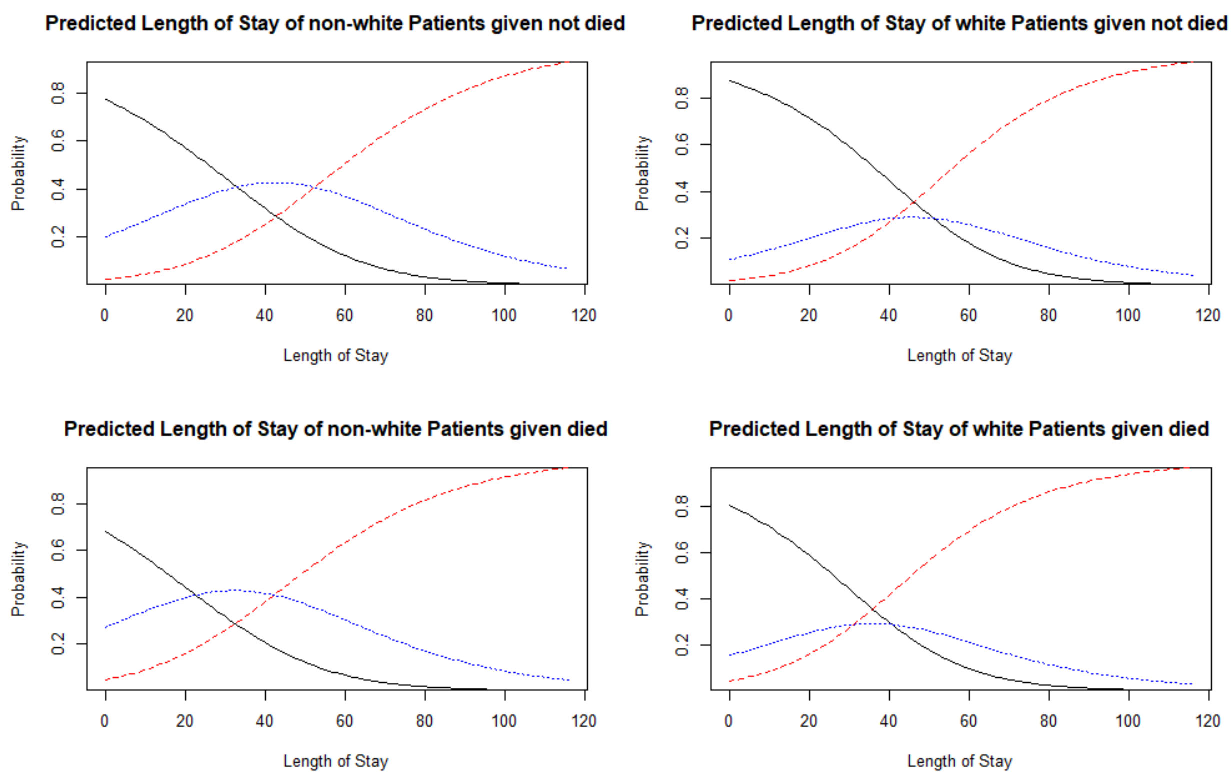
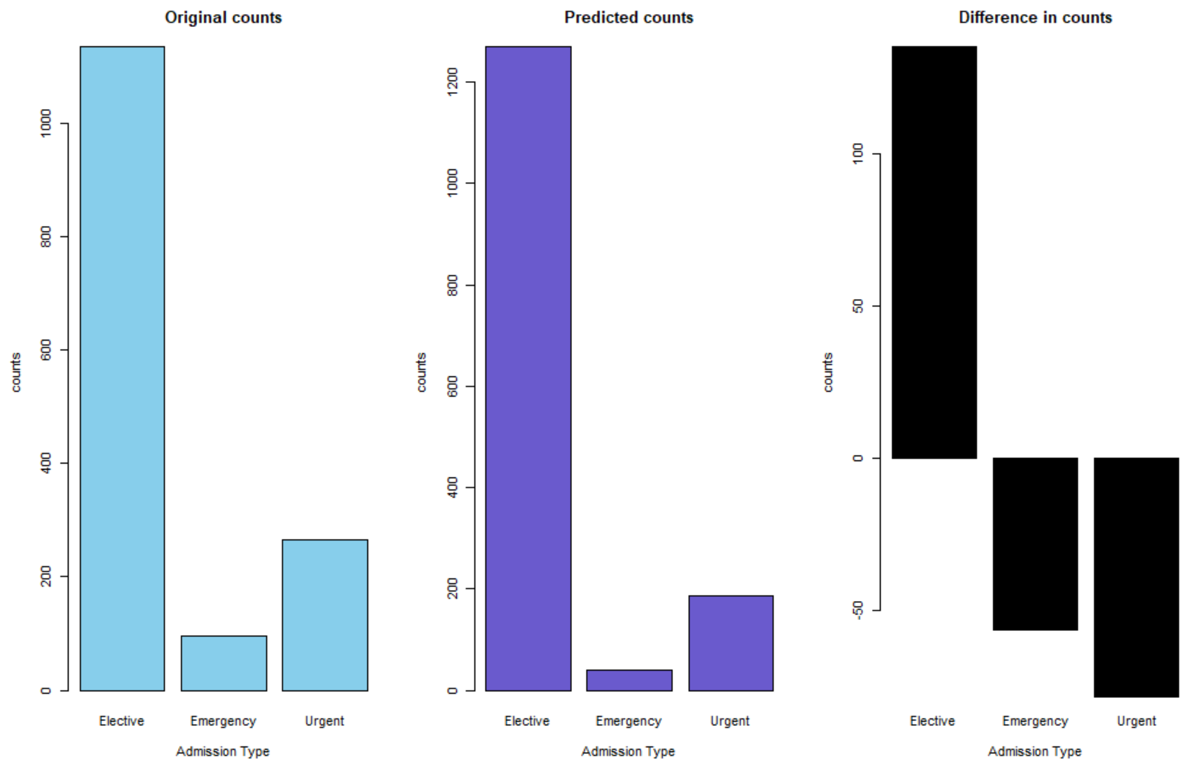


Figure 2: Differences in Predicted and Original counts



School of Mathematics and Statistics

University of St Andrews

Technical Summary

Investigation of changes in the probability of patients' admission to hospitals.

This report aims to investigate the findings of the research about how the probability of type of hospital admissions changes when patients attended hospitals based on their personal choice (electively), or under urgent (referred by GP) and emergency cases (by ambulance team). For this purpose, variables such as age groups, race (whether white or non-white), length of stay and subsequent death are examined to predict the different types of hospital admission (response variable). This report is intended for statistically literature administrators who are interested in analysis of changes in patients' admission to hospitals under certain situations. All preliminary analysis is performed with the help of R Programming Language. To fit the model, a multinomial model is used since it accommodates categorical response type with multiple levels. In this case, we are dealing with a nominal dataset as it appears to have no natural ordering. The 'multinomial logit model' computes log relative risk of two alternative scenarios (i.e. white and non-white), and models relative risk of each outcome with respect to a baseline (in this example elective is the reference level). The covariate '*provnum*' (which records hospital identification number) is also not included in our prediction as it doesn't present valuable information about the patients and types of hospital admission. In this report the Akaike information criterion (AIC) is utilized as an estimator of the relative quality of statistical models for a given dataset. The final model contains three variables including 'race' (white and non-white), 'length of stay', and 'subsequent death'.

The predicted model suggests that the likelihood of patients attending hospitals both in emergency and urgent cases given they died, with 95% confidence, is likely to increase by 2.33 and 1.52 times respectively compared with the baseline (see Table 1). Moreover, the possibility of white patients being admitted to hospital tends to decrease by 24.25 % and 52.20% for both emergency and urgent cases respectively with respect to baseline. This estimated result however is unreliable as its confidence intervals contain zero as a possible value. Also under emergency and urgent categories the probability of patients' length of stay tends to increase by 1.08 and 1.05 times compared with baseline respectively. Effect plots are used to show predicted probability of each covariates in our model (see Figure 2). Based on the estimated model, the probability of length of stay (measured in days) for white and non-white patients given they died or survived is likely to increase under emergency circumstances (see Figure 1). The length of stay in hospital for white and non-white patients under urgent cases, however, tends to increase only for the first 40 days. Thereafter a negative association is indicated by our predicted model. Model assessment of the estimated model using half-normal plots reveals that samples are normally distributed and well-fitted as they're close to a straight line (see Figure 3). Tests for serial correlation of residuals were completed using the 'Ljung–Box' test and 'ACF' plots. The result shows that residuals are autocorrelated, and thus independence assumption does not hold true. As regards 'Independent from irrelevant alternatives' assumption, this report argues that the relative likelihood of choosing elective, urgent and emergency will not change based on whether additional alternatives are present.

The value of McFadden R-square of estimated model is 5.14% suggesting that overall the model is poorly fitted across the regression line. This report is therefore unsure about the conclusion drawn. Also there are more explanatory variables not captured in the data that can offer greater predictive power. Examples include the gender, health conditions of patients, locations of hospitals. Moreover, the type of hospital admission appears to be an invalid measure of the true emergency of patients. If anything it is only a measure of the access routes to hospitals, and presents a misleading perception of patients' trend under emergency cases. Furthermore, the dataset provides no information about the actual urgency of admissions. Considerations towards confounding variables like these are critical to generate inferences to a wider population from the sample.

Appendix

Table1: Coefficients of Predicted Model

	(Intercept)	Died	White	Length of Stay
Emergency	0.02	2.33	0.75	1.08
Urgent	0.26	1.52	0.47	1.04

*Table 3. P-values of the Predicted Model: significance level denoted by 0.001 '***'*

	(Intercept)	Died	White	Length of Stay
Emergency	0.000 ***	0.000 ***	0.465	$2.22e-16$ ***
Urgent	$5.5e-09$ ***	0.003 ***	0.000 ***	$1.683e-06$ ***

Table 2: Confidence intervals of the estimated Model

Emergency	2.5 %	97.5%
(Intercept)	-4.32	-2.72
Died	0.40	1.28
White	-1.02	0.47
Length of Stay	0.06	0.10

Urgent	2.5 %	97.5%
(Intercept)	-1.79	-0.89
Died	0.14	0.70
White	-1.16	-0.31
Length of Stay	0.02	0.05

Figure 3: Prediction of Patients' Length of Stay. The solid line is elective, dashed is emergency and dotted is urgent.

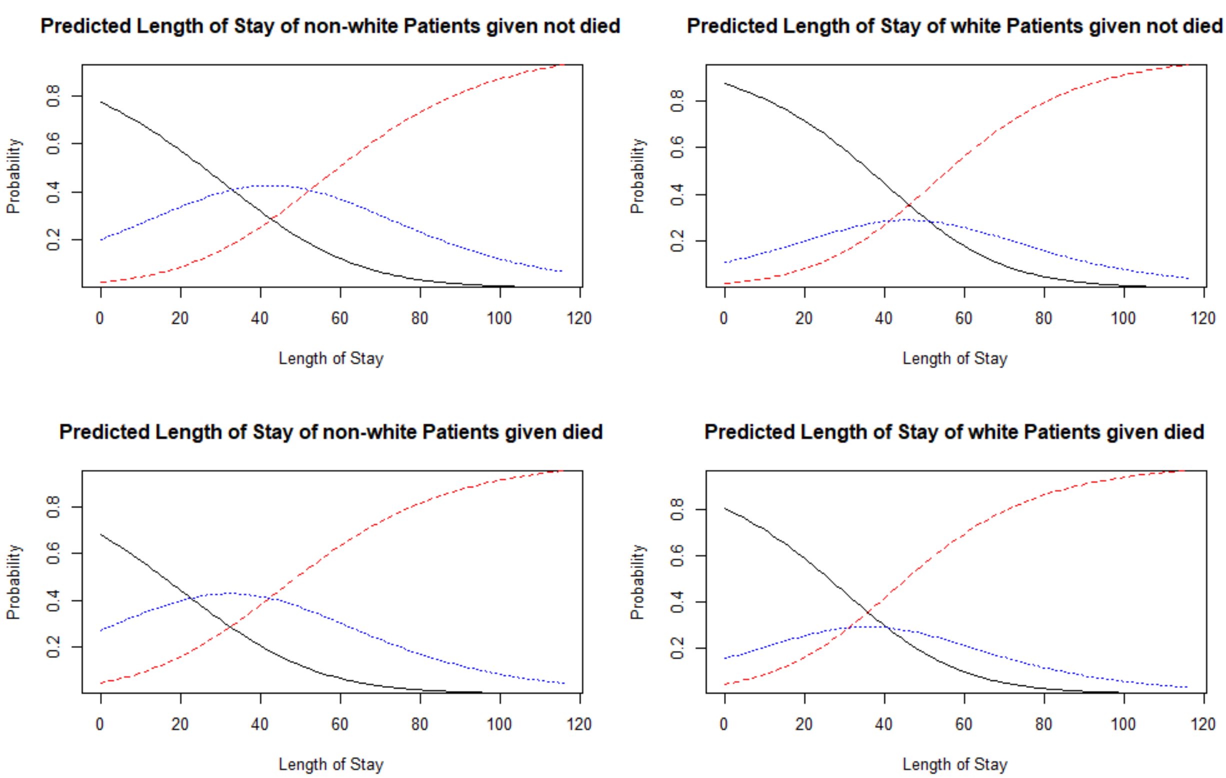


Figure 4: Effect Plots of Predicted Model

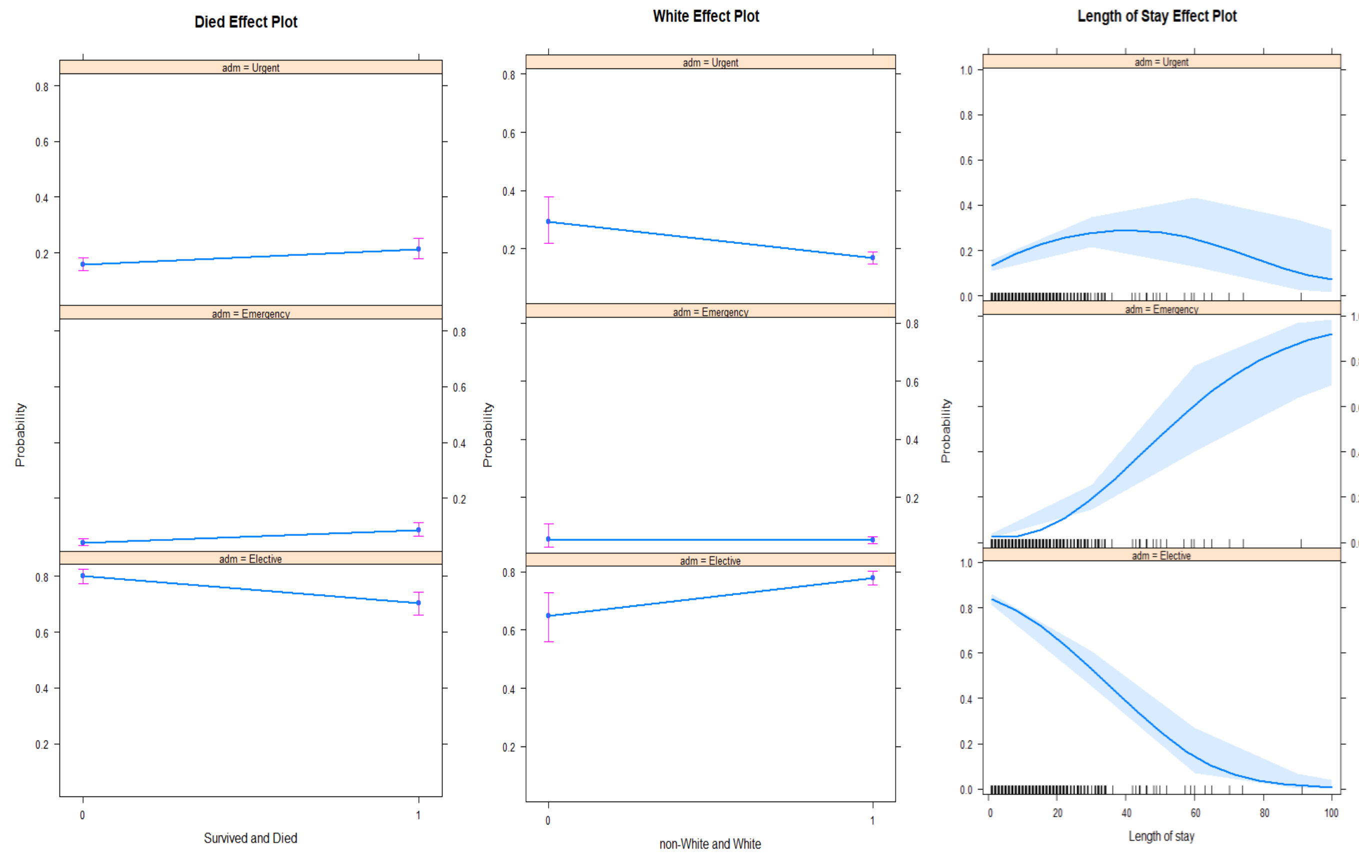


Figure 5: half-norm Plots of estimated models

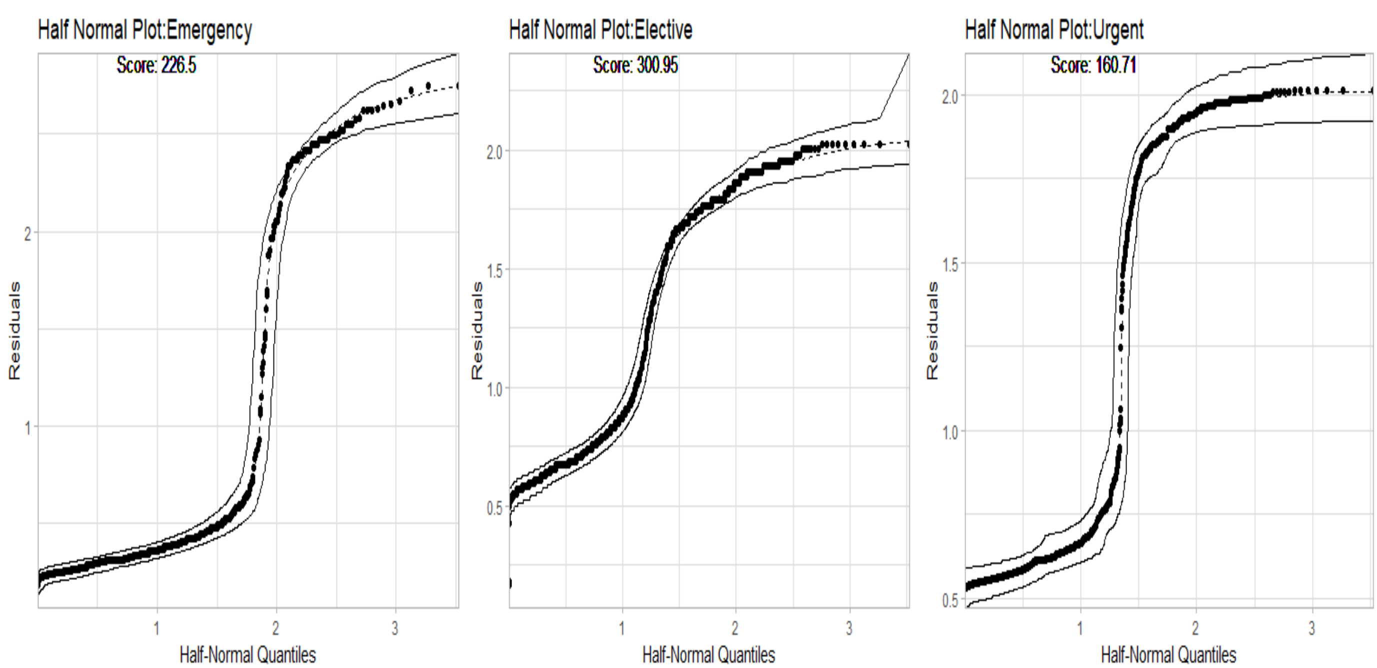


Figure 4: Bar Plots of explanatory variables provided in the dataset