Ben Lovelace

Professor Michael

DASE 4570

5/12/2020

Final Project

![A screenshot of a cell phone

Description automatically generated]() For my final project, I decided to look at data from a study on the effects of ibuprofen on the physiology and survival of patients with sepsis. This study was conducted because previous studies that had confirmed the effects of Ibuprofen on sepsis patients used small sample sizes (less than 30). The sample used in this study was much larger and consisted of 455 patients from all over the United States and Canada. Before a linear mixed model is created to identify a relationship between treatment and temperature of the patient, the fields of the data should be explained. The first field of the dataset is the studyid, this is a unique id number assigned to each of the 455 patients of the study. The second field is the hour the temperature measurement was taken. The hours the measurements were taken are 0, 4, 16, 24, and 36. The third field is the temperature of the patient at the specific hour. The fourth field is the treatment, this field can have a value of 1 for ibuprofen treatment or 0 for placebo. The fifth field is the age of the patient. The sixth field is the race of the patient 0 is white, 1 is black and 2 is other. The sixth field is the Acute Physiology and Chronic Health Evaluation (APACHE II) score. The seventh and last field is the surv3 or survival status of the patient. 1 means the patient died within 15 days of the study, 2 means the patient died 15-30 days after the study and 3 means the patient survived past 30 days of the study. To get a better idea of the spread of the data, a summary of each column filtered by the treatment level can be seen below.

A close up of a map

Description automatically generatedThe mean for the temperature is 99.96 for placebo patients while the mean for ibuprofen patients is 99.00. This could be a potential relationship between the treatment and the temperature. To visualize this potential relationship, a scatter plot with the x axis being the hour and the y axis being the temperature. A linear regression line could be fitted in top of the scatter plot to see what the average temperature does over time. The plot can be seen below.

The plot shows that the average temperature of the patients goes down over time. However, one major assumption is violated in this simple linear model, that is the assumption of independent measures. Since every patient has multiple measures over the course of the study, they cannot be considered independent. In order to further understand this potential relationship properly, a linear mixed model can be created to identify how much of this change over time is due to the treatment level while controlling for the multiple measures per patient. The first model is going to be a random intercept with a formula as follows, temp ~ hour + treat + age + (1|studyid). This model includes fixed effects believed to impact the temperature of the subject the most. The random effects of this model account for the multiple measures per patient by controlling for the studyid. The summary of this model can be seen below.

![A close up of text on a black background

Description automatically generated]()

By looking at the coefficients of the fixed effects, it can be seen that the treatment has the greatest impact on temperature over time with a value of -0.96. This model can be improved on because random intercept models acknowledge the fact that every patient starts off with different temperatures but fails to acknowledge the fact that every patient won’t exhibit the same relationship between the fixed effects and the response variable. To fix this, a random effects model can be used. The new model looks like this temp ~ hour + treat + age + (hour|studyid). The only change needed was to add the effect of interest, hour, into the previous model. The output for this model can be seen below.

![A screenshot of a cell phone

Description automatically generated]()

Looking at the coefficients of the fixed effects, it can see that the impact of these variables on the temperature have increased. In order to compare the fitness level of each model, the AIC and BIC values can be compared. Since both the AIC and BIC values for the random effects model are lower, it is the more fit model. To visualize the random effects model, the random intercepts and slopes can be plotted. Below is a visualization that captures the relationship between the treatment and the temperature.

A picture containing text, map, table, truck

Description automatically generatedThe blue lines represent patients who received the ibuprofen while the red lined represent patients who received the placebo. The plot shows that overall the patients with the ibuprofen treatment had noticeably lower temperatures than the patients who had the placebo. This shows that there is indeed a relationship between the treatment and the temperature of the patient.