BST 224 HW1

BohaoZou

4/10/2020

Question 2:

(a)

In this project, the researchers used two different treatments (1. Receive Azathioprine alone (AZ, group 1) 2. Receive Azathioprine plus Methylprednisommne (AZ+MP, group 2)) for patients and we want to investigate (i) whether both treatments lower AFCR over 18 month period and (ii) whether treatment with AZ+MP results in different immune system response than does AZ alone, and, (iii) if so how it is different in terms of response over time.

At the first step, we should explore the data structure. In this data set, the **id** variable represents the identity of suffers. The **time** is the longitudinal variable and it indicates how long has passed after the initiation of treatment. The **AFCR** is the response variable. Low value of AFCR are evidence that immunity is improving and may have a better treatment effect. The **group, prior treatment indicator and age** are x-covariates. Those variables may affect the response variable AFCR in some situations.

X-Covariates Explore

We must notice that the immunity is different between different age and if received the prior treatment is also important factor which effects the immunity. So, we could not ignore those essential factors.

There are 150 suffers participated in this project. The youngest suffer is 32 years old and the oldest is 73 years old. There are 92 suffers who has received prior treatments and 58 suffers who did not receive the prior treatments. The distribution of age and prior treatments are displayed below.
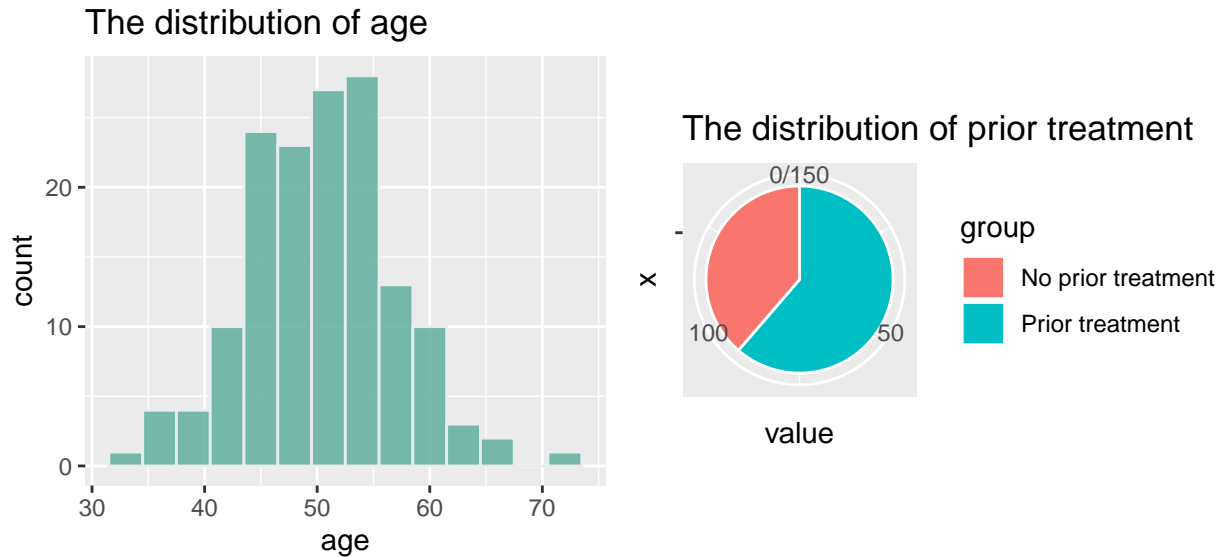

Figure 1 : The left side is the distribution of age variable. The right side is the distribution of Prior treatment indicator.

From the distribution, we can get that the age of most suffers are concentrated in the interval of 42 - 56 and the number of suffers who recieved prior treatment is more than the number of suffers who didn't receive the prior treatment.

Longitudinal variable explore

In the data set, the longitudinal variable time is an discrete variable. It contains 7 unique time points which are 0, 3, 6, 9, 12, 15 and 18 respectively. Those number indicate how many months have passed after the first

treatment. Because the longitudinal variable is an discrete variable. It is difficult to use points to display the trend of response variable.