Consulting Meeting

Friday Team

September 30,2022

1 Basic Information

- Client: Kwang Jin Choi
- Affiliation: Hospital Medicine

2 The Problem

There are around 5000 hospitalized covid patients who either receive or don't receive a medication for covid during their stay at the hospital. The goal is to see the effect of this medication on whether a patient requires ventilation or not. Some of the other covariates included in the study are age,sex,whether a patient has diabetes or not, some comorbidities, etc.

3 Method Used

Let $Y_i = 0/1$ denote whether the i-th patient requires ventilation, $M_i = 0/1$ denote whether that patient received medication or not and let X_i denote the other covariates.

3.1 Model 1

The model is that of a logistic regression.

$$Y_i \sim \text{Bernoulli}(p_i)$$
$$\text{logit}(p_i) = \alpha + \beta M_i + X_i^T \beta$$

3.2 Model 2

Now, two different types of ventilation are considered which are invasive ventilation and non invasive ventilation. He looked at the 2×2 contingency table of ventilation(Yes/No) vs medication(Yes/No) for both types of ventilation. The results of proportion of ventilations required in the two groups were contrary to what he was expecting. Then he fit the logistic regression model (same as Model 1) to both the groups.

4 Suggestions During the Meeting

- Since, the client is not sure how exactly a patient is assigned to receive one of the two different ventilations or whether and how the covariates or the information of a patient is related to what ventilation they are assigned to, it does not make sense to divide the patient population into two groups. This is basically just similar to doing the same analysis to two randomly split partitions of the data. So the contrary results in the contingency table could just be a result of that particular random split.

- **How to strengthen the results of the analysis?** We could test for the significance of the covariates included in the model using the LRT test(performed by anova in R) and include only the significant covariates. We could also do some sort of model selection to include less number of covariates, like perform backward elimination as there are quite a few covariates.