I read chapter three from the MATLAB book. This chapter provided a review for linear algebra and statistics. The review on common distributions, probabilities and significance tests became very handy. I was conducting analysis on the TMR data for the past two weeks. Since the first and second data did not produce similar trend in the scatter plot, we tried to perform a chi-squared test to test the difference between the number of points in the uncued and cued groups, a matched pair t-test, and independent sample t-test to test the difference in the improvements of the two groups. Even though I just took the statistics class last semester, I forgot some of the equations we need and got a little bit confused on the one-sided or two-sided test’s p-value calculation. This chapter gave a good review of the major significance tests we used on the TMR data and cleared up my confusion. When I was taking a test and doing homework last semester, it seemed very easy to pick which test to use under what condition. However, in the lab dealing with real raw data, I am very slow on choosing the variables and the right statistic tests to exam further. For example, when we looked at the group of points with very high TMR effects, Dr. Reber suggested us to perform a chi-squared test, but I did not know chi-squared test could be used to test the categorical distribution. When we were dealing with the maximum improvements of each subject in cued and uncued conditions, I continued to use chi-squared test without realizing that the variables are uniquely matched and using a matched-pair t-test would be more appropriate for this case. I found that looking at the data in a real life setting gave very different perspectives than in a homework problem. We need to know the relationships between the variables, what should we calculate, how should we test the results and how should we present the data to best capture the possible patterns. Conducting the tests on real data set makes the statistical concepts more concrete and meaningful.

Chapter 3 of the *Theoretical Neuroscience* book moved the discussion to the decoding of the stimulus characteristics from the neuronal response and uses the decoding as a tool to assess the accuracy of the encoding from the computational models. This chapter has some overlap with the MATLAB book’s chapter. The probabilities theory and Bayesian theory were introduced in the MATLAB book and used as the foundations for a lot of adaptations in the decoding techniques in the modeling book. The chapter used the table of false alarm and hit rate in response to stimuli in different directions as a criterion to the performance of the decoding procedure. We used the same table to score the just noticeable difference task in the formation-formation project. One of my first job was to create the table using the data we have and found ways to produce the table on smaller segments of the data such as, specific changing dimension, specific changing range, etc. The difference in hit rate and false alarm rate could provide a reasonable evaluation for people’s correct judgment versus random guessing.

The *Theoretical Neuroscience* book contains very dense mathematics. The statistics concepts introduced in this book focuses on the theoretical aspects rather than the application like the MATLAB book or the analysis we did in the lab. Most of the theories used are brand new to me and the derivations are very challenging to follow, but I think it helps me to understand statistics and neuroscience better. I hope we have a class on this topic.

This week we got a new set on a paired special task. But the dataset has 60 variables and is not well formatted. I spent most of my time this week trying to clean up the data and figure out a solution to match the data of the paired special task to the one item task so that the results of the three data sets are comparable. I was not able to repeat the researcher’s previous statistically significant result. So, we ended up asking for the data in a cleaner format and it turns out the formatting was distorted due to compatibility issues as the researcher is currently in Germany. With the cleaner version, I managed to reproduce her previous results on Friday and repeated the analysis we did in our lab on the two other data sets. I did not find any patterns yet, but we will discuss the results on Wednesday lab meeting. In the previous analysis of the just noticeable difference task and the TMR studies, I found the results rather random, but Dr. Reber and other pot docs in the lab could spot the interesting behavior in the data and suggest different ways to further investigate the data. I think I definitely need a lot more training to improve the sensitivity and intuition about a given data set.