

6 Analysis

With the goal to figure out whether Properties (Area, Perimeter, Circularity and Aspect Ratio) of mitochondria are different by Locations (Proximal end, Middle and Distal end) in a single muscle fiber cell from a young mouse, we set our null hypothesis (H_0) to be that the mean Area, Perimeter, Circularity and Aspect Ratio of mitochondria are equal by Locations and alternative hypothesis (H_A) to be that the means are not all equal. Moreover, I also conducted pairwise comparison tests to find out the difference between Locations. In the overall and pairwise comparison hypothesis tests, MLE and Weighted Mean were used as the estimators of mean Area for the whole mitochondria and for each location; 2nd Order Taylor's Approximation Estimator and Weighted Mean were for Perimeter and Arithmetic Mean was for Circularity and Aspect Ratio for their independence to size of Area.

Overall Hypothesis Test:

$$\begin{cases} H_0 : \mu_{i_P} = \mu_{i_M} = \mu_{i_D}, i = \{\text{Area, Perimeter, Circularity, Aspect Ratio}\} \\ H_A : \text{At least one } \mu_{i_j} \neq \mu_{i_k}, j, k = \{P, M, D\} \end{cases}$$

Pairwise Comparison Test:

$$\begin{cases} H_0 : \mu_{i_j} = \mu_{i_k}, i = \{\text{Area, Perimeter, Circularity, Aspect Ratio}\}; j, k = \{P, M, D\} \\ H_A : \mu_{i_j} \neq \mu_{i_k} \end{cases}$$

Instead of choosing the standard ANOVA (F-test) and T-test as our overall and pairwise comparison test methods, we decided to use a Permutation method. The reasons are that the

estimator of population mean in ANOVA and T-test is sample mean but it is not appropriate to Area and Perimeter because of their size-biased samples. And for Circularity and Aspect Ratio, the data violated the normality assumption of ANOVA and T-test (Figure 3.3.3 and Figure 3.3.4). Hence, the statistics in our case for the overall test was $\sum_{i=\{P,M,D\}} (\hat{\mu}_i - \hat{\mu})^2$ and for the pairwise comparison test was $\hat{\mu}_i - \hat{\mu}_j$, where $i = \{P, M, D\}$. Their sampling distributions were constructed by assuming the null hypothesis is true, treating the observed data as population, finding out all the possible combinations of elements in groups and then calculating the statistics for every combination. However, it was not efficient for us to calculate all the combinations so we randomly drew large amount of the combinations from the complete set to obtain an approximate sampling distribution. Then, the approximate P-value was the probability for the statistics more extreme than the observed one. After finished all the hypothesis test, we used Bootstrap technique with Bonferroni's correction to obtain confidence interval of true difference of properties means by Locations.

Bonferroni's correction is a method to control Strong Familywise Error Rate in multi-comparison hypothesis tests and to have simultaneous confidence interval for the mean differences. The significance level for each test are defined as α/m , where m is the number of test in this multi-comparison test. So, the simultaneous confidence interval for the mean differences will be $(1 - \frac{\alpha}{m})\%$.