HW 11

In this assignment, several tests are conducted on comparing means of two populations with different combinations of actual mean μ and standard error σ. The first population has a fixed mean $\mu_1 = 0$ and a standard error $\sigma_1 = 1$. The second population can have a mean $\mu_2 = 0, 0.5, 1, 2$ and a standard error $\sigma_2 = 0.25, 1, 2, 5$. Sizes of two samples from two populations are 6 and 10 respectively, which are very small. Thus I expect that Welch's test gives a better result according to what I learned in class.

From the statistics in the output file, we can see that z-test always gives higher powers than pooled t-test and Welch's test do. In addition, z-test gives a lower power, 0.1531, to the parameter set $\mu_2 = 0.5, \sigma_2 = 5$, which means a different mean than μ_1, while giving a higher power, 0.1616, to the parameter set $\mu_2 = 0, \sigma_2 = 0.25$, which means a same mean. As a consequence, it may not be able to detect whether two populations have different means or actually have very different variances. Even worse situation can be seen in the results of pooled t-test. When second population has different means and the variances are $\sigma_2 = 2, 5$, the power are still low, even lower than the case that two populations have same mean and variance. In contrast, for variance $\sigma_2 = 0.25$, the power is too high even when two populations have the same mean. Consequently, we cannot distinguished a power implying different population means from a power revealing reversed information. Based on the findings, pooled t-test may be used only when two populations have the same variance. On the other hand, in Welch's test, the powers for populations with a same mean and the powers for those with different means can be easily distinguished. We can therefore adopt a threshold, such as 0.11, to determine whether two populations have the same mean or not.

For permutation tests, the results are basically the same for all repeating iterations. Thus we can see the obtained powers are either 0 or 1. Both permutation tests do not perform well on the cases which have standard error $\sigma_2 = 0.25$. The first permutation test has a false positive with parameter $\mu_2 = 0.5$ and $\sigma_2 = 5$, which is a very distant variance from $\sigma_1 = 1$. Otherwise, permutation tests can also provide good results if two populations do not have too different variances and means. The only drawback of these approaches is that there cannot be too many runs of permutation, or it will take a long time to finish.