ST502: Final Project - Part 2

Apostolos Stamenos & Tyler Pollard

4/19/2022

In the previous section, we concluded that the t-test with the Satterthwaite approximation is more appropriate when there is no evidence that the population variances are equal. In order to understand the properties of both tests and to verify that our conclusion makes sense, we conducted a simulation study. A numerical simulation study can be used to evaluate the two testing procedures because unlike in a real data analysis, we know exactly what the true population parameters are. By generating enough datasets through a data-generating mechanism that we control, we can evaluate how the two testing procedures do when drawing inferences about the underlying populations.

(Tyler)Discuss design of (generated data sets) and mentioning how many combinations of the population characteristics with values considered which equals 135 unique data sets eqach containing 100 randomly generated samples for two samples

Table 1: **Varying Population Characteristics**

Population Characteristics	Values
True Variance	$\sigma_1^2 = 1, 4, 9$ and $\sigma_2^2 = 1$
Sample Size	$n_1 = 10, 30, 70$ and $n_2 = 10, 30, 70$
True Mean Difference	$\Delta = \mu_1 - \mu_2 = -5, -1, 0, 1, 5$

Figure 1: **Approximate alpha curves of the Pooled and Satterthwaite t-tests for different sample sizes and population variances**
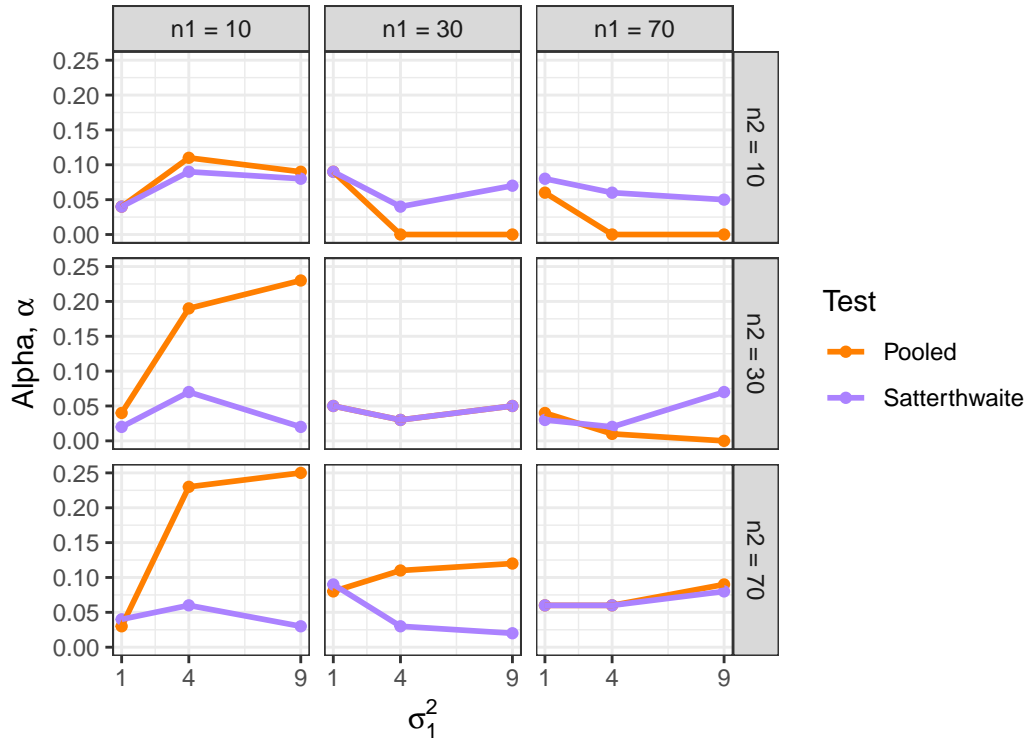
Regardless of n_1 and n_2, the power of both the pooled and Satterthwaite t-tests is greater for smaller values of σ_1^2. This result makes sense intuitively. When the variance of one population is small, it is easier for the hypothesis testing procedure to detect the true mean difference. When the sample sizes are the same (i.e., $n_1 = n_2$), the simulated powers of the pooled and Satterthwaite tests are approximately equal, regardless of σ_1^2. For $n_1 > n_2$, the Satterthwaite t-test tends to perform better, whereas for $n_1 < n_2$, the pooled variance t-test tends to perform better. As n_1 and n_2 increase, so does the power of both the pooled and Satterthwaite t-tests. This makes sense because for both tests, the standard error (i.e., the denominator of the test statistic) gets smaller and the degrees of freedom get larger for large sample sizes.

Figure 2: **Approximate power curves of the pooled and Satterthwaite t-tests for different samples sizes and population variances**
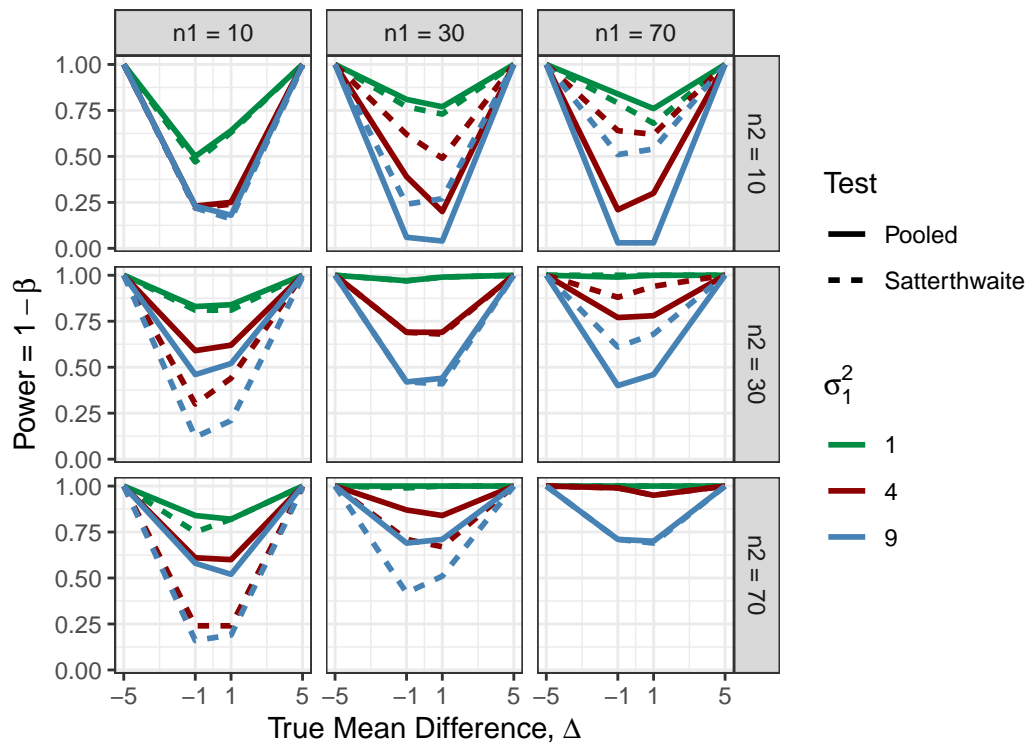
Results of simulations with the plots for both alpha and power (Apostolos focus on power Tyler on alpha)

Discuss conclusions based on plots and other theoretical arguments (Apostolos)

Need a conclusion paragraph

Appendix A: R Code

Team Member Contributions

These reports are the result of extensive collaboration facilitated by Zoom meetings and GitHub. Both group members took turns coding and writing the report. We had long discussions to figure out the best way to conduct the hypothesis tests and simulation study, and how to effectively present our findings.

Apostolos: performed the pooled variance t-test, wrote the code for many of the plots comparing smokers and nonsmokers, wrote the section assessing the normality assumption, wrote the section on testing the equality of variances, wrote a function to generate datasets for the simulation, wrote the section justifying the use of a simulation study, and wrote the section discussing the power of the tests.

Tyler: managed pull requests to the GitHub repository, performed the Satterthwaite t-test, wrote a function to calculate the probability of rejecting H_0 for the simulation, wrote code to calculate and process the simulated values of α and power, wrote the section discussing the design of the simulation study, and wrote the section discussing the Type I error rate of the tests.