ST502: Final Project

Apostolos Stamenos & Tyler Pollard

4/19/2022

Part 1

Let $Y_{i1}, \ldots, Y_{in}$ be a simple random sample, where $i = 1$ denotes that the individual was selected from the population of nonsmokers and $i = 2$ denotes that the individual was selected from the population of smokers. For the samples from both population, we assume the parametric model $Y_{i1}, \ldots, Y_{in} \overset{\text{iid}}{\sim} N\left(\mu_i, \sigma_i^2\right)$.

Table 1: **Summary of tests**: .

Test	Point Estimate	SE	Test Statistic	p-value	Confidence Interval
Pooled Variance	0.yyy	0.0xx			
Satterthwaite	0.yyy	0.0xx			

Figure 1: **yadayadayada**

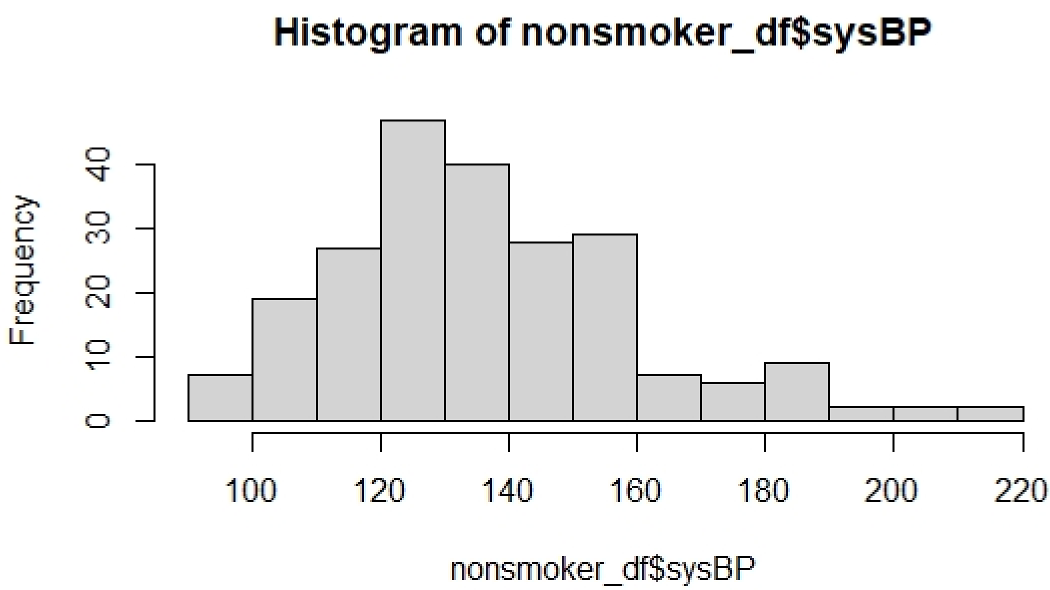


Figure 2: **yadayadayada**

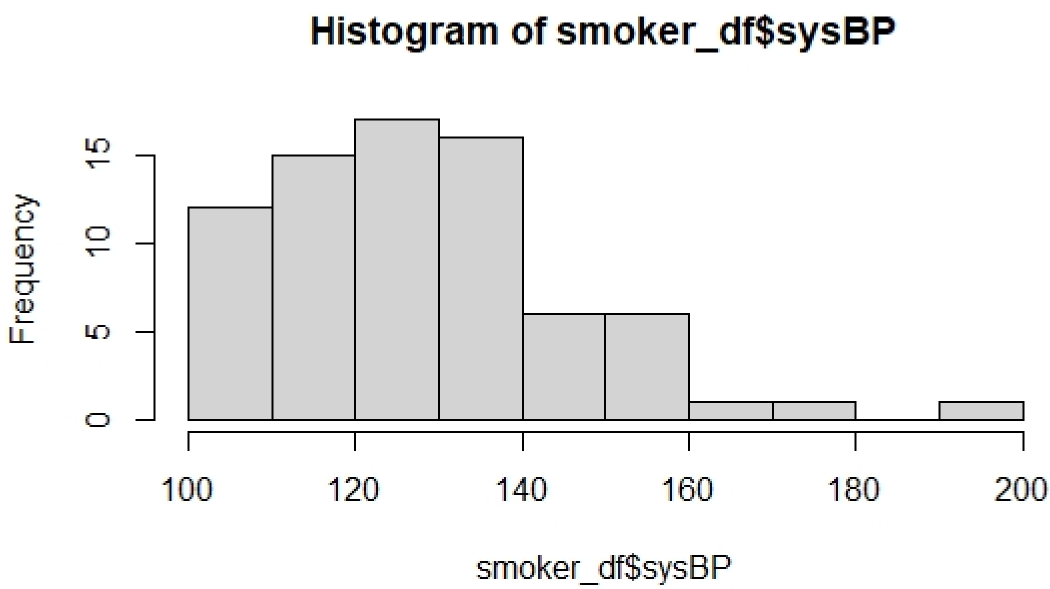


Figure 3: **yadayadayada**

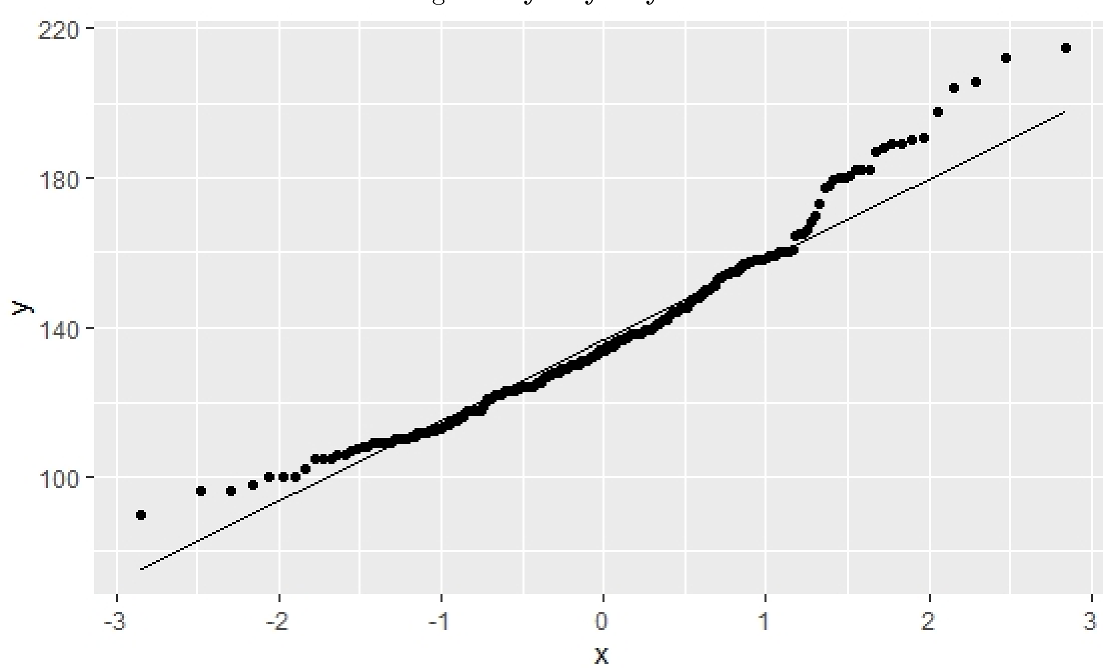
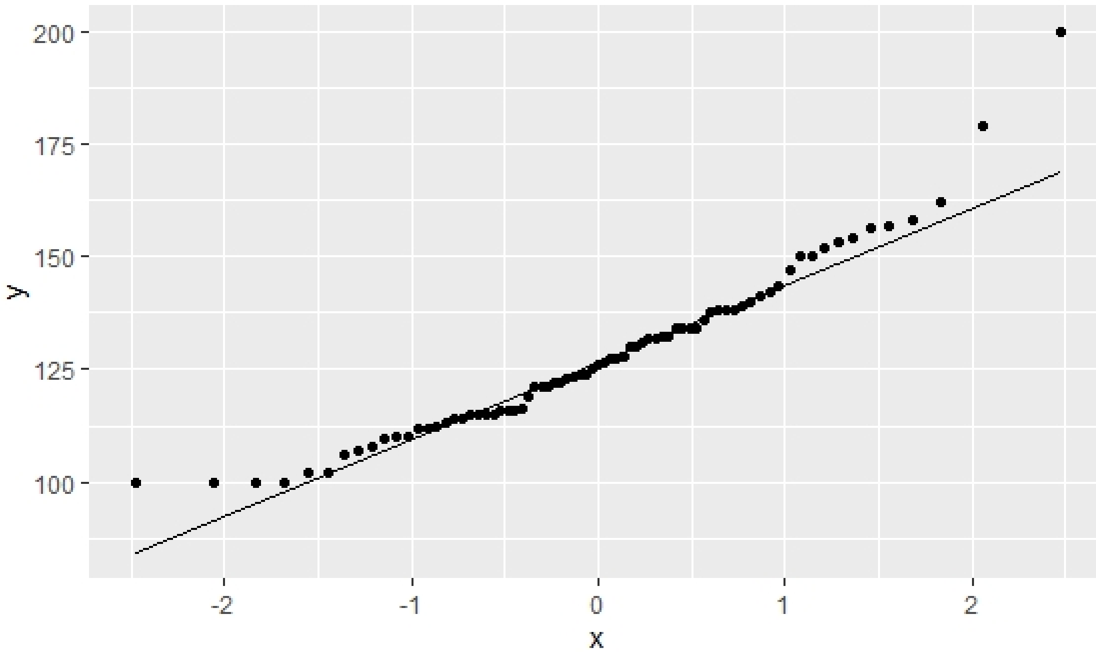
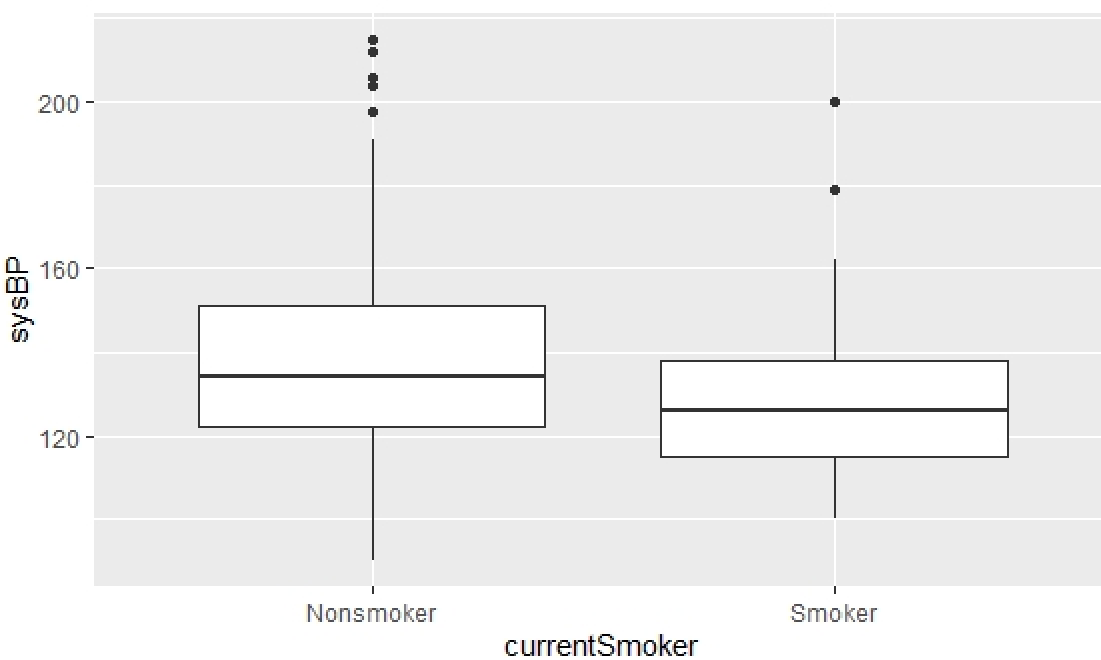


Figure 4: **yadayadayada**



Both the histograms and the normal QQ plots indicate that the data are skewed to the right. The boxplots are fairly symmetrical if the outliers are excluded. We decided to keep the outliers in the analysis. By the Central Limit Theorem, even if the two datasets are not completely normal, their sample means are asymptotically normally distributed. Since the number of smokers and the number of nonsmokers are sufficiently large, the use of t-tests and confidence intervals are justified by the Central Limit Theorem.

Figure 5: **yadayadayada**



The boxplots indicate that the distribution of systolic blood pressure for nonsmokers is more spread out than the distribution of systolic blood pressure for smokers. Based on the boxplots, there is no indication that the true population variances are equal. To ensure robustness to the assumption of equal variances, we concluded that the t-test with the

Satterthwaite approximation is preferred.

Part 2

Table 2: **Table of blahblahblahblah**: For most of the scenarios, the power of the xxx test is higher than that for the xxx test.

True Parameters	Pooled	Satterthwaite
$\mu_1 - \mu_2 =, n_1 =, n_2 =, \sigma_1^2 =, \sigma_2^2 =$	0.yyy	0.0xx
$\mu_1 - \mu_2 =, n_1 =, n_2 =, \sigma_1^2 =, \sigma_2^2 =$	0.yyy	0.0xx