

The impact of sample size re-estimation on the type I error rate in the analysis of a continuous end-point

by

SONGNIAN ZHAO

B.S., Tianjin University, 2010

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Major Professor
Dr. Christopher I. Vahl

Abstract

Sample size estimation is generally based on assumptions made during the planning stage of a clinical trial. Often, there is limited information available to estimate the initial sample size. This may result in a poor estimate. For instance, an insufficient sample size may not have the capability to produce statistically significant results, while an over-sized study will lead to a waste of resources or even ethical issues in that too many patients are exposed to potentially ineffective treatments. Therefore, an interim analysis in the middle of a trial may be worthwhile to assure that the significance level is at the nominal level and/or the power is adequate to detect a meaningful treatment difference. In this report, the impact of sample size re-estimation on the type I error rate for the continuous end-point in a clinical trial with two treatments is evaluated through a simulation study. Two sample size estimation methods are taken into consideration: blinded and partially unblinded. For the blinded method, all collected data for two groups are used to estimate the variance, while only data from the control group are used to re-estimate the sample size for the partially unblinded method. The simulation study is designed with different combinations of assumed variance, assumed difference in treatment means, and re-estimation methods. The end-point is assumed to follow normal distribution and the variance for both groups are assumed to be identical. In addition, equal sample size is required for each group. According to the simulation results, the type I error rates are preserved for all settings.

Table of Contents

List of Figures	iv
List of Tables	vi
Acknowledgements	vii
Chapter 1 - Introduction	1
1.1 Clinical trials and sample size	1
1.2 Hypothesis test	2
1.3 Type I error rate	3
1.4 Blinding	4
Chapter 2 - Methods	6
2.1 Sample size initialization	6
2.2 Sample size re-estimation	7
2.3 Two-sample t-test	7
Chapter 3 - Simulation Study	9
Chapter 4 - Simulation Results	12
Chapter 5 - Summary and Future Work	18
References	20
Appendix A - Histograms of Sample Size Distributions	21
Appendix B - Simulation SAS Code	45

List of Figures

Figure 4.1 Type I error rate for each setting	16
Figure 4.2 Type I error rate with different assumed variance for two methods	17
Figure 4.3 Type I error rate with assumed differences of treatment means for two methods	17
Figure A.1 Distribution of Sample Sizes for Setting 1 (Expected Variance: Pilot, method: control, $\epsilon^*:0.5$)	21
Figure A.2 Distribution of Sample Sizes for Setting 2 (Expected Variance: Pilot, method: control, $\epsilon^*:1$)	22
Figure A.3 Distribution of Sample Sizes for Setting 3 (Expected Variance: Pilot, method: control, $\epsilon^*:2$)	23
Figure A.4 Distribution of Sample Sizes for Setting 4 (Expected Variance: Pilot, method: pool, $\epsilon^*:0.5$).....	24
Figure A.5 Distribution of Sample Sizes for Setting 5 (Expected Variance: Pilot, method: pool, $\epsilon^*:1$).....	25
Figure A.6 Distribution of Sample Sizes for Setting 6 (Expected Variance: Pilot, method: pool, $\epsilon^*:2$).....	26
Figure A.7 Distribution of Sample Sizes for Setting 7 (Expected Variance: 1.5, method: control, $\epsilon^*:0.5$).....	27
Figure A.8 Distribution of Sample Sizes for Setting 8 (Expected Variance: 1.5, method: control, $\epsilon^*:1$)	28
Figure A.9 Distribution of Sample Sizes for Setting 9 (Expected Variance: 1.5, method: control, $\epsilon^*:2$).....	29
Figure A.10 Distribution of Sample Sizes for Setting 10 (Expected Variance: 1.5, method: pool, $\epsilon^*:0.5$).....	30
Figure A.11 Distribution of Sample Sizes for Setting 11 (Expected Variance: 1.5, method: pool, $\epsilon^*:1$).....	31
Figure A.12 Distribution of Sample Sizes for Setting 12 (Expected Variance: 1.5, method: pool, $\epsilon^*:2$).....	32
Figure A.13 Distribution of Sample Sizes for Setting 13 (Expected Variance: 3, method: control, $\epsilon^*:0.5$).....	33

Figure A.14 Distribution of Sample Sizes for Setting 14 (Expected Variance: 3, method: control, $\varepsilon^*:1$).....	34
Figure A.15 Distribution of Sample Sizes for Setting 15 (Expected Variance: 3, method: control, $\varepsilon^*:2$).....	35
Figure A.16 Distribution of Sample Sizes for Setting 16 (Expected Variance: 3, method: pool, $\varepsilon^*:0.5$).....	36
Figure A.17 Distribution of Sample Sizes for Setting 17 (Expected Variance: 3, method: pool, $\varepsilon^*:1$).....	37
Figure A.18 Distribution of Sample Sizes for Setting 18 (Expected Variance: 3, method: pool, $\varepsilon^*:2$).....	38
Figure A.19 Distribution of Sample Sizes for Setting 19 (Expected Variance: 4, method: control, $\varepsilon^*:0.5$).....	39
Figure A.20 Distribution of Sample Sizes for Setting 20 (Expected Variance: 4, method: control, $\varepsilon^*:1$).....	40
Figure A.21 Distribution of Sample Sizes for Setting 21 (Expected Variance: 4, method: control, $\varepsilon^*:2$).....	41
Figure A.22 Distribution of Sample Sizes for Setting 22 (Expected Variance: 4, method: pool, $\varepsilon^*:0.5$).....	42
Figure A.23 Distribution of Sample Sizes for Setting 23 (Expected Variance: 4, method: pool, $\varepsilon^*:1$).....	43
Figure A.24 Distribution of Sample Sizes for Setting 24 (Expected Variance: 4, method: pool, $\varepsilon^*:2$).....	44

List of Tables

Table 1.1 Type I error and type II error.	3
Table 3.1 Settings for the simulation study	11
Table 4.1 Sample size summary: average final sample size.....	13
Table 4.2 Type I error rate summary	15

Acknowledgements

I would like to take this opportunity to express my sincere appreciation to many people who advised and helped me to complete this report.

First, I would like to thank my advisor Dr. Christopher I. Vahl for constantly guiding me and giving me valuable advice on research and personal life. I would never have could finish my dissertation without his guidance. I would also like to thank my other committee members: Dr. Haiyan Wang and Dr. Perla Reyes.

I would like to express my gratitude to other faculty and staff members in my department for their help and support. Many thanks to Ms. Jo Ann Blackburn and Ms. Bonnie Messmer, for all the help in my school life. I also acknowledge my collegians in the department for their love and support during my graduate study.

Last but not the least, this report is dedicated to my family including my parents, Baozhong Zhao and Binmei Sun, my older brother, Funian Zhao, and my wife, Yan Kuang, for their endless love, support, understanding, and faith in me.

Chapter 1 - Introduction

1.1 Clinical trials and sample size

A clinical trial is a designed experiment involving human subjects (or animals) to evaluate the efficacy of two or more interventions. For a traditionally designed experiment, experimental units are identified before the treatments are randomized to them. However, the subjects are often recruited into the study over time in a clinical trial. In general, the larger the number of subjects involved in a trial, the more meaningful the statistical results are. Hence, the subject recruitment efforts and costs play a key role for the success of a clinical trial and the sample size estimation will be extremely essential because of the complexity of the recruitment.

Sample size estimation is generally based on assumptions during the planning stage of a clinical trial. Much of time, there are only very limited information available to calculate the initial sample size, which may result a poor estimation. For instance, an insufficient sample size may not have the capability to produce statistically significant results, while an over-sized study will lead to a waste of resources or even raise ethical issues. Therefore, an interim analysis in the middle of a trial may be worthwhile to assure that the significance level is at the nominal level and/or the power is adequate to detect a meaningful treatment difference.

To calculate the sample size, the type of endpoint (or response variable) should be identified. There are two main types typically encountered in clinical trials: binary endpoint and continuous endpoint. The type of endpoint determines the appropriate statistical analysis and the approach to calculate sample size. In this report, the sample size re-estimation with a continuous type of endpoint in clinical trials will be studied. The basic assumptions applied in this study include: 1. The endpoint follows normal distribution. 2. The variances for control group and test group are assumed to be identical. 3. Equal sample size is required for each group.

In this report, we extend the results of Cong (2016), who studied sample-size re-estimation for a binary endpoint, by examining the impacts of sample size re-estimation on the type I error rate for the continuous end-point through a simulation study. The simulation study is designed with different combinations of assumed variance, assumed difference in treatment means and re-estimation methods. The objective of this report is to determine whether the type I error rates are preserved for all combinational settings, including different blinding strategies. A secondary objective is to study how the sample sizes are changed correspondingly after the interim analysis.

1.2 Hypothesis test

The type of hypothesis tests adopted for the study depends on the purpose of the trial. Typically, there are three types of hypothesis tests: superiority, non-inferiority, and equivalence. In this work, only the test of superiority is taken into consideration, which is usually used to test if the investigational (or test) intervention (e.g. a new therapy) is superior to a control therapy. The formal hypothesis test is shown below:

$$\begin{array}{ll}
 H_0: \mu_T \leq \mu_C + \delta & \text{OR} & H_0: \mu_T - \mu_C \leq \delta \\
 H_A: \mu_T > \mu_C + \delta & & H_A: \mu_T - \mu_C > \delta.
 \end{array}$$

where $\delta > 0$ is called the superiority margin and is based on the clinical knowledge of the disease. The terms μ_T and μ_C represent the mean values of the response variables for the test and control group, respectively. Without loss of generality, we assume here that a larger mean response is better. If H_0 is rejected, the test intervention is concluded to be clinically superior to the control. For the sake of simplicity, a test of statistical superiority is considered in this study, which means $\delta = 0$. Then the original hypothesis test becomes:

$$H_0: \mu_T \leq \mu_C$$

OR

$$H_0: \mu_T - \mu_C \leq 0$$

$$H_A: \mu_T > \mu_C$$

$$H_A: \mu_T - \mu_C > 0.$$

1.3 Type I error rate

For sample size re-estimation, type I error rate and type II error rate (i.e. $1 - \text{power}$) are two crucial parameters. Both of them must be specified before the calculation of sample size. The type I error and type II error could be explained based on the hypothesis test in section 1.2.

Considering the hypothesis test with H_0 and H_A , there are four possible conclusions which may be drawn: H_0 is not rejected when it is true, H_0 is rejected when it is true, H_0 is not rejected when it is false, and H_0 is rejected when it is false. The result is shown in Table 1.1

Table 1.1 Type I error and type II error.

	H_0 is not rejected	H_0 is rejected
H_0 is true	$1 - \alpha$	Type I error rate = α
H_0 is false	Type II error rate = β	Power = $1 - \beta$

From the Table 1.1, the type I error rate is usually denoted by α , which indicates the probability that a true H_0 is incorrectly rejected. Similarly, the type II error rate is usually denoted by β , which indicates the probability that a false H_0 is not rejected. In addition, the power is defined as $1 - \beta$, referring to correctly rejecting a false H_0 . In this report, the goal is to ensure the trial's assumption on the variance is accurate by taking an interim look at the data while not inflating the type I error rate.

1.4 Blinding

For the study of clinical trials, one of the best methods to compare effects of test interventions is the randomized clinical trial (RCT) (Armitage, Berry and Matthews 2008). The most important goal of randomization is to avoid bias. However, a particular risk of bias in clinical trials is due to the subjectivity in assessment, specifically the subjective human behaviors involved in the experiments (Day and Altman 2000). To minimize this type of bias, another commonly used tool, named blinding, is introduced in this section.

Blinding refers to a masking of the treatment group assignments in a clinical trial, usually a RCT (Karanicolas 2010). There are four main blinding types, including: unblinded, single blinded, double blinded, and triple blinded. Unblinded indicates that both of subjects and the investigators know which intervention has been assigned. Single blinded means that the subject does not know which intervention has been assigned, while the investigator does. For double blinded, both of subjects and investigators do not recognize which intervention has been assigned. Sometimes, the triple blinded is mentioned in some specific cases, representing that all involved personnel including the statisticians do not know which intervention has been assigned. From the researcher's perspective, the blinding could be classified into three groups: completely unblinded, partially unblinded, and blinded (or no unblinded). The partially unblinded indicates that researchers only recognize the groups but do not know what treatment is assigned to each group, while the blinded means researchers have no information about the groups and treatment assignments (Gould 2001).

In literature, some researchers claim that unblinded methods may impact the type I error rate after sample size re-estimation compared to blinded method. For instance, a simulation study is conducted to show that the unblinded method adopted for sample size re-estimation may inflate

the type I error rate if the sample size is small (Wittes et al. 1999). When using a variance estimate with continuous end-point clinical trials for sample size estimation, the type I error rate may be inflated (Wittes and Brittain 1990; Birkett and Day 1994). Hence, two sample size estimation methods are taken into consideration in this study: partially unblinded and blinded. For the blinded method, all collected data for two groups are used to estimate the variance, while only data from the control group are used to re-estimate the sample size for the partially unblinded method.

Chapter 2 - Methods

2.1 Sample size initialization

Consider a clinical trial with two treatments (control and test). Assume the corresponding responses for both groups follow normal distributions with the same variance σ^2 . If there is a pilot study before the trial, the collected data could be used to estimate the variance; however, if not applicable, a point estimate of σ^2 is assumed at the planning stage, denoted by $(\sigma^2)^*$. To estimate the needed sample size to compare the treatment effects of two groups, assume that the mean difference is $\mu_T - \mu_C = \varepsilon^*$. In addition, initialization of sample size also depends on the desired type I error rate and the desired power. For instance, the type I error rate usually is chosen as 5%, i.e. $\alpha=0.05$; similarly, the power is assumed to be 80%, i.e. $1 - \beta = 0.8$. Considering the test of statistical superiority, the sample size can be calculated using the formula in Chow, Wang and Shao (2007):

$$n_1 = kn_2 \quad (2.1)$$

$$n_2 = \frac{(z_\alpha + z_\beta)^2 (\sigma^2)^* (1 + 1/k)}{(\varepsilon^* - \delta)^2} \quad (2.2)$$

where n_1 represents the sample size for control group, n_2 represents the sample size for test group, k is the ratio n_1/n_2 , and z_α and z_β are the upper α and β critical values from a standard normal distribution. Based on our assumption that there are equal sample sizes for each arm or treatment group, $k = 1$. For the test of statistical superiority, $\delta = 0$. Hence, the initial sample size can be calculated as:

$$n = n_1 = n_2 = \frac{2(z_\alpha + z_\beta)^2 (\sigma^2)^*}{(\varepsilon^*)^2} \quad (2.3)$$

$$N = 2n \quad (2.4)$$

Suppose that the assumed $(\sigma^2)^*$ is 3.0 and ε^* is equal to 1.0. With a 5% significance level and 80% power, the initial sample size is then $N = 2n = 74$.

2.2 Sample size re-estimation

Sample size re-estimation sometimes happens in the middle of the trial. For the sake of assuring accuracy and reliability, investigators evaluate the interim data to determine if changing the sample size is necessary. In this study, the interim analysis is conducted at the halfway of the trial, i.e. $n/2$ patients are completed for each group. In clinical trials, researchers may choose blinded or unblinded methods to conduct the trial based on the specific conditions. Typically, unblinding data may lead to a bias resulting that the type I error rate may be inflated in the end of trial, so both of blinded and unblinded are adopted in this simulation study so as to compare the differences between these two methods. As mentioned in section 1.4, two methods are involved into the sample size re-estimation: blinded method and partially unblinded method.

Based on the updated estimated variance, the re-estimated sample size N^* could be calculated using Eqs (2.3) and (2.4). If the original sample size N is smaller than N^* , N^* is chosen as the final sample size; otherwise, N is the final sample size. Hence, the final sample size N_{final} is given below:

$$N_{final} = \text{Max}(N, N^*) \quad (2.5)$$

2.3 Two-sample t-test

To compare the efficacy differences between two treatments, the statistical superiority test is conducted in this study. The hypothesis for this test is shown as follows:

$$H_0: \varepsilon \leq 0$$

$$H_A: \varepsilon > 0$$

where $\mu_T - \mu_C = \varepsilon$. Since the test of superiority is a right-tailed two-sample t-test, the test statistic is given below:

$$t = \frac{\bar{y}_2 - \bar{y}_1}{\hat{\sigma}_{pool} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2.6)$$

$$(\hat{\sigma}_{pool})^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (2.7)$$

where \bar{y}_1 and \bar{y}_2 are the sample means of two treatment groups, s_1^2 and s_2^2 are their sample variances, and n_1 and n_2 are their sample sizes. The region of rejection for the null hypothesis is under the condition that $t > t_{0.05, n_1+n_2-2}$, where $t_{0.05, n_1+n_2-2}$ represents the upper 0.05 critical value of student's t distribution with degrees of freedom equal to $n_1 + n_2 - 2$.

Chapter 3 - Simulation Study

For the sake of evaluating the type I error for the test of superiority after the interim analysis, a simulation study was designed with a variety of combinations of assumed variance, assumed difference in treatment means, and sample-size estimation methods. All settings are shown in Table 3.1. The simulation was conducted by SAS 9.4 platform (SAS Institute Inc.).

The response is assumed to follow a normal distribution with a true variance of 3.0. The assumed variance, $(\sigma^2)^*$, used in the initial sample size estimate were set to 1.5, 3.0, and 4.0, representing the cases of underestimation, exact estimation, and overestimation respectively. In addition, a pilot study was also taken into consideration, which meant a small sample size (we chose 10 in the simulation) was selected before the trial and the true variance was estimated based on the results of the pilot study. The assumed difference in treatment means, ε^* , between control group and test group were assumed to be 0.5, 1.0, and 2.0. Two sample size re-estimation methods in this simulation study were adopted: using only data from control group and using the pooled data from both of treatment groups without unblinding the treatment codes. Hence, there are total 24 settings in this study: $(3+1)$ assumed variances times 3 assumed differences in treatment means times 2 sample-size re-estimation methods. The complete SAS code is given in Appendix B.

For each setting, the simulation was implemented for 5000 replications with the process described below:

1. If applicable, generate the data for the pilot study (sample size is 10), and calculate $(\sigma^2)^*$, otherwise use an assumed variance $(\sigma^2)^*$.
2. Initialize the sample size n^* based on $(\sigma^2)^*$.
3. Generate the data for half the sample size n^* in each arm.

4. Use the interim data to re-estimate the sample size n , and choose $\text{Max}(n, n^*)$ as the final sample size for each arm.
5. Generate the data for the rest of trial based on the final sample size.
6. Perform a two-sample t-test for the test of superiority and estimate the type I error rate by

$$\widehat{\text{Type I error rate}} = \frac{\text{the number of rejections of } H_0}{\text{total number of simulation runs}}$$

Table 3.1 Settings for the simulation study

Setting	$(\sigma^2)^*$	Method	ε^*
1	Pilot	Control Group	0.5
2	Pilot	Control Group	1
3	Pilot	Control Group	2
4	Pilot	Pooled	0.5
5	Pilot	Pooled	1
6	Pilot	Pooled	2
7	1.5	Control Group	0.5
8	1.5	Control Group	1
9	1.5	Control Group	2
10	1.5	Pooled	0.5
11	1.5	Pooled	1
12	1.5	Pooled	2
13	3	Control Group	0.5
14	3	Control Group	1
15	3	Control Group	2
16	3	Pooled	0.5
17	3	Pooled	1
18	3	Pooled	2
19	4	Control Group	0.5
20	4	Control Group	1
21	4	Control Group	2
22	4	Pooled	0.5
23	4	Pooled	1
24	4	Pooled	2

Chapter 4 - Simulation Results

In this chapter, the two main questions are discussed based on all simulation results: 1. How does each combination impact the final sample size if H_0 is true? 2. How does the sample size re-estimation impact the type I error rate?

To answer the first question, the distribution of final sample size in all 24 settings are given using histogram in Appendix A (from A.1 to A.24). A summary for the average final sample size of each setting is shown in Table 4.1. Based on Table 4.1 and Figure A.1 To Figure A.24, for each combination of assumed variance and assumed difference of treatment means, the results show that sample size re-estimation methods (partially unblinded and blinded) do not impact the final sample size if H_0 is true. For instance, the pairs of figure A.1 and A.4, A.2 and A.5, A.3 and A.6 have very similar sample size distribution. The same conclusion can be drawn if other pairs of settings with the same ε^* and σ^{2*} are compared. In addition, the average sample size for each method is almost the same. On the other hand, given the value of σ^{2*} and a re-estimation method, the final sample size will increase as the value of ε^* decreases which is as expected. The same conclusion could be obtained based on the Eq (2.3) in chapter 2. However, the value of ε^* does not have an impact on the difference between the original sample size and final sample size. For the σ^{2*} , the final sample size will increase as σ^{2*} increases based on the Eq. (2.3). The histograms in appendix A also validate this result. In addition, if the variance is underestimated, the final sample size will be larger than the original sample size for the most of cases; while if the variance is well estimated or overestimated, the final sample size will be very close to the original sample size. This result is reasonable since the calculation of sample size is positively correlated with the estimated variance. If the variance is underestimated, there will be a small sample size in the beginning of the trial. After the completion of half of sample size and the re-estimation of the

variance using the collected data, the updated estimated variance will be closer to the true value which leads to a larger sample size.

Table 4.1 Sample size summary: average final sample size

Assumed Treatment Differences ε^*	Assumed variance σ^{2*}	Method	
		Control	Pooled
0.5	Pilot	177.5	176.4
1	Pilot	45.2	44.2
2	Pilot	11.8	11.5
0.5	1.5	148.9	148.0
1	1.5	37.6	36.6
2	1.5	10.3	9.4
0.5	3	158.2	155.1
1	3	42.0	40.4
2	3	11.9	10.8
0.5	4	197.2	197.0
1	4	50.0	49.3
2	4	13.3	12.6

In summary, re-estimation methods do not impact the final sample size if H_0 is true. The assumed difference in treatment means impact the final sample size but do not contribute to the differences between original sample size and final sample size. The assumed variance impacts both of final sample size and the differences of sample sizes after interim analysis.

For the second question, the type I error rate is calculated for each setting using 5000 replications. The summary of type I error rate and the standard error is given in table 4.2. Figure 4.1 shows the result for each setting with three lines. The middle line is the nominal type I error rate of 0.05, and the upper and lower lines represent the margin of error for a (non-simultaneous)

95% confidence interval for the true type I error rate based on inference for one population proportion (i.e. $\pm 1.96\sqrt{0.05 \times 0.95 / 5000}$). All points fall between upper bound and lower bound but for one point which is smaller than the lower bound, meaning its type I error rate deflated. This one result is likely due to random chance since the confidence bounds were calculated for one proportion. Hence, the interim analysis does not appear to be inflating the type I error rate based on these simulation results. In addition, there are no obvious differences for the impact of re-estimation methods on the type I error rate according to the Figures 4.2 and 4.3.

Table 4.2 Type I error rate summary

$(\sigma^2)^*$	Method	ε^*	Type I error rate	Standard error
Pilot	Control	0.5	0.0508	0.0031
Pilot	Control	1	0.0486	0.0030
Pilot	Control	2	0.0489	0.0031
Pilot	Pooled	0.5	0.0508	0.0031
Pilot	Pooled	1	0.0476	0.0030
Pilot	Pooled	2	0.0467	0.0030
1.5	Control	0.5	0.0496	0.0031
1.5	Control	1	0.0510	0.0031
1.5	Control	2	0.0470	0.0030
1.5	Pooled	0.5	0.0474	0.0030
1.5	Pooled	1	0.0518	0.0031
1.5	Pooled	2	0.0434	0.0029
3	Control	0.5	0.0542	0.0032
3	Control	1	0.0508	0.0031
3	Control	2	0.0530	0.0032
3	Pooled	0.5	0.0514	0.0031
3	Pooled	1	0.0476	0.0030
3	Pooled	2	0.0500	0.0031
4	Control	0.5	0.0546	0.0032
4	Control	1	0.0508	0.0031
4	Control	2	0.0508	0.0031
4	Pooled	0.5	0.0544	0.0032
4	Pooled	1	0.0504	0.0031
4	Pooled	2	0.0478	0.0030

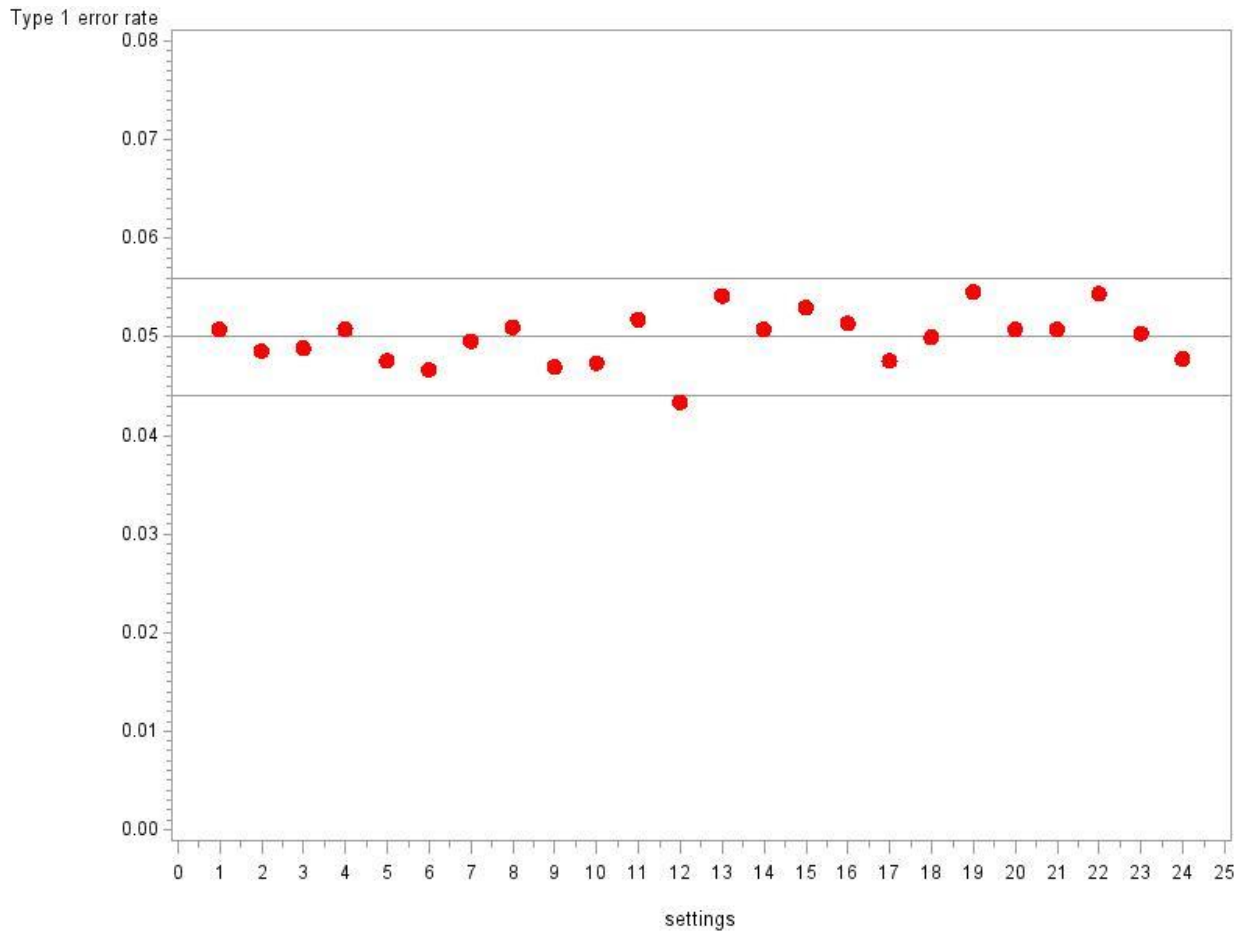


Figure 4.1 Type I error rate for each setting

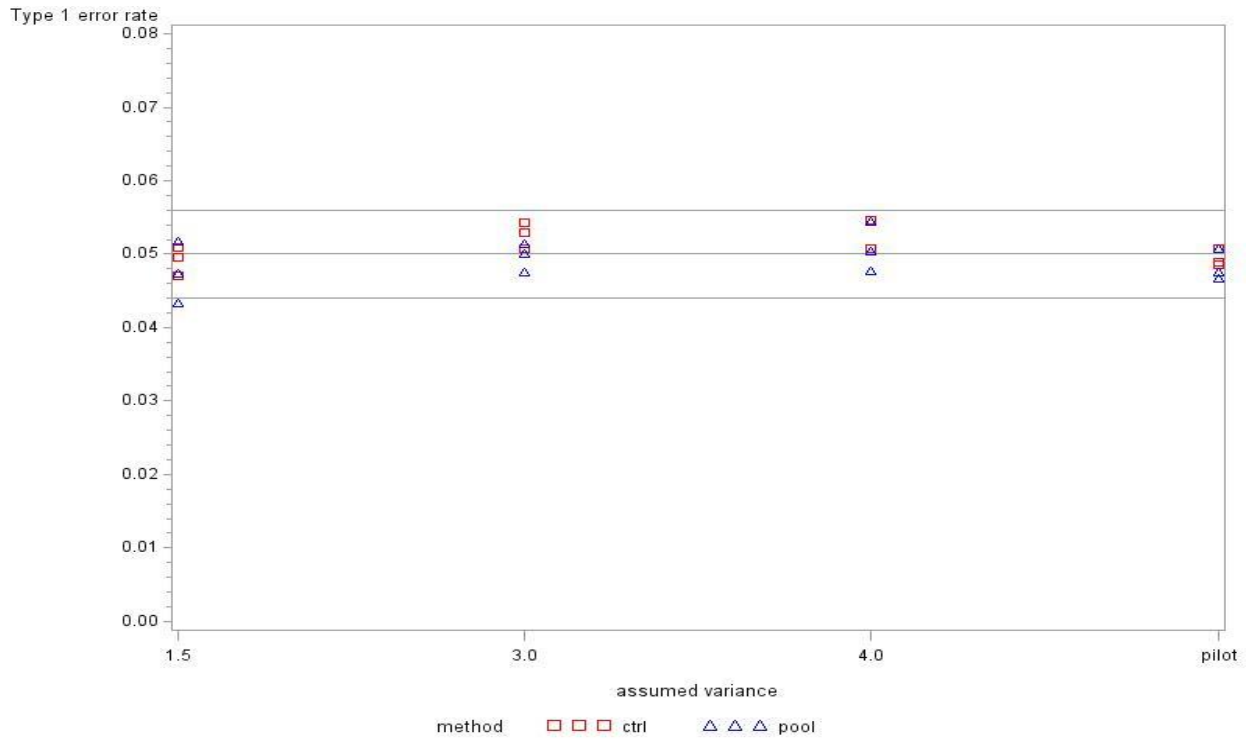


Figure 4.2 Type I error rate with different assumed variance for two methods

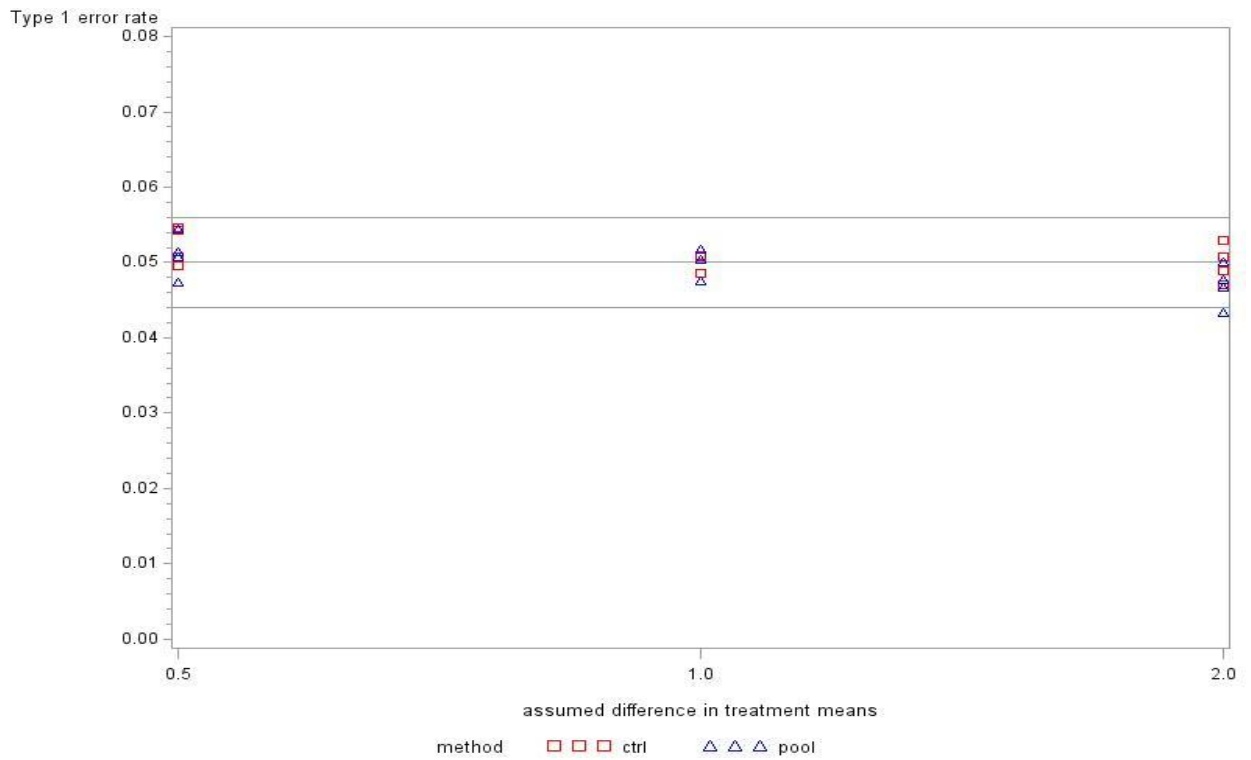


Figure 4.3 Type I error rate with assumed differences of treatment means for two methods

Chapter 5 - Summary and Future Work

In this study, the impacts of sample size re-estimation on the type I error rate in a clinical trial with continuous type of endpoints are investigated. This simulation study follows several assumptions for the endpoints and sample size. The distribution of endpoints is assumed to be normal with the same variance for the response variable in control and test groups. For each arm, the sample size is assumed to be equal. The sample size of the clinical trial for each group is calculated based on the assumed variance, assumed difference in treatment means, desired type I and type II error rates. To initialize the sample size, either a pilot study is given or the assumed variance is guessed based on prior experience. In addition, two blinding strategies are taken into consideration: partially unblinded and blinded. In the middle of the trial, the observed data are used to re-estimate the sample size. A test of superiority is conducted to evaluate the effects of two treatments. The overall type I error rate is estimated after completing all simulation replications for each combination of parameters and blinding methods.

The main conclusion of this study is drawn based on the simulation results with respect to how each setting impacts the final sample size and type I error rate. It turns out that the similar final sample size is obtained using the two blinding strategies when H_0 is true. It is reasonable since the identical mean and variance are assumed so that data from control group or from the pooled sample follows the same probability distribution. The impacts of assumed difference in treatment means and assumed variance to the final sample size are also discussed. Based on the Eq. (2.3), the sample size is smaller as the ϵ^* increases. But ϵ^* does not have any impact on the sample size re-estimation since this parameter is not updated. On the other hand, the assumed variance has a positive impact on the initial sample size. If the variance is underestimated, the final sample size will be larger than the initial sample size, while if the variance is closely or

overestimated, the final sample size will be similar to the initial sample size. Both of blinding strategies does not inflate the type I error rate in our simulation. In fact, the type I error rate is preserved very well for all simulation settings.

For the future work, there are several possible interesting extensions. First, the endpoint may not follow normal distribution, so we may investigate if the same conclusion will hold for endpoint with a non-normal distribution. Second, the identical variance is assumed for each treatment. The type I error rate might be violated after sample size re-estimation if the two treatment arms have different variances. Third, the halfway point of the clinical trial is chosen as the time when sample size is re-estimated. For different scenarios, there may be better choices for the selection of interim analysis time. Finally, only the type I error rate is studied in this work. The next step should involve the discussion of the test of power, which also plays an essential role in clinical trial research.

References

- Armitage, P., Berry, G. and Matthews, J. N. S. (2008), *Statistical methods in medical research* John Wiley & Sons.
- Birkett, M. A. and Day, S. J. (1994), "Internal pilot studies for estimating sample size," *Statistics in medicine* , 13, 2455-2463.
- Chow, S., Wang, H. and Shao, J. (2007), *Sample size calculations in clinical research*. CRC press.
- Cong, D. (2016). The effect of sample size re-estimation on type I error rates when comparing two binomial proportions. MS Report, Kansas State University. <http://krex.k-state.edu/dspace/handle/2097/34504> (last accessed 04-07-2017)
- Day, S. J. and Altman, D. G. (2000), "Blinding in clinical trials and other studies," *Bmj* , 321, 504.
- Gould, A. L. (2001), "Sample size re- estimation: recent developments and practical considerations," *Statistics in medicine* , 20, 2625-2643.
- Karanicolas, P. J. (2010), "Practical tips for surgical research: blinding: who, what, when, why, how?" *Canadian journal of surgery* , 53, 345.
- SAS Institute Inc. , *SAS 9.4*, Cary, NC, USA.
- Wittes, J. and Brittain, E. (1990), "The role of internal pilot studies in increasing the efficiency of clinical trials," *Statistics in medicine* , 9, 65-72.
- Wittes, J., Schabenberger, O., Zucker, D., Brittain, E. and Proschan, M. (1999), "Internal pilot studies I: type I error rate of the naive t- test," *Statistics in medicine* , 18, 3481-3491.

Appendix A - Histograms of Sample Size Distributions

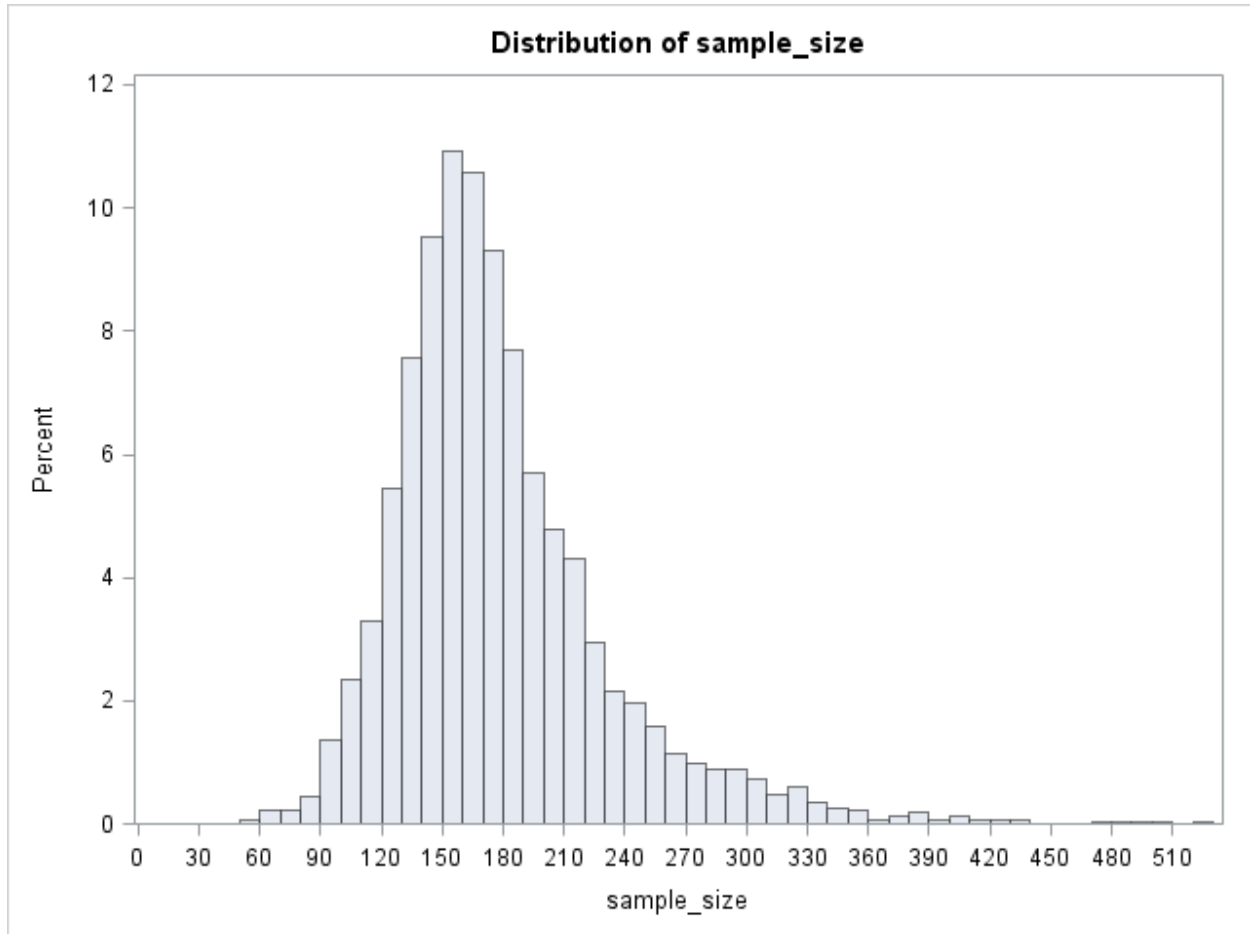


Figure A.1 Distribution of Sample Sizes for Setting 1 (Expected Variance: Pilot, method: control, $\epsilon^*:0.5$)

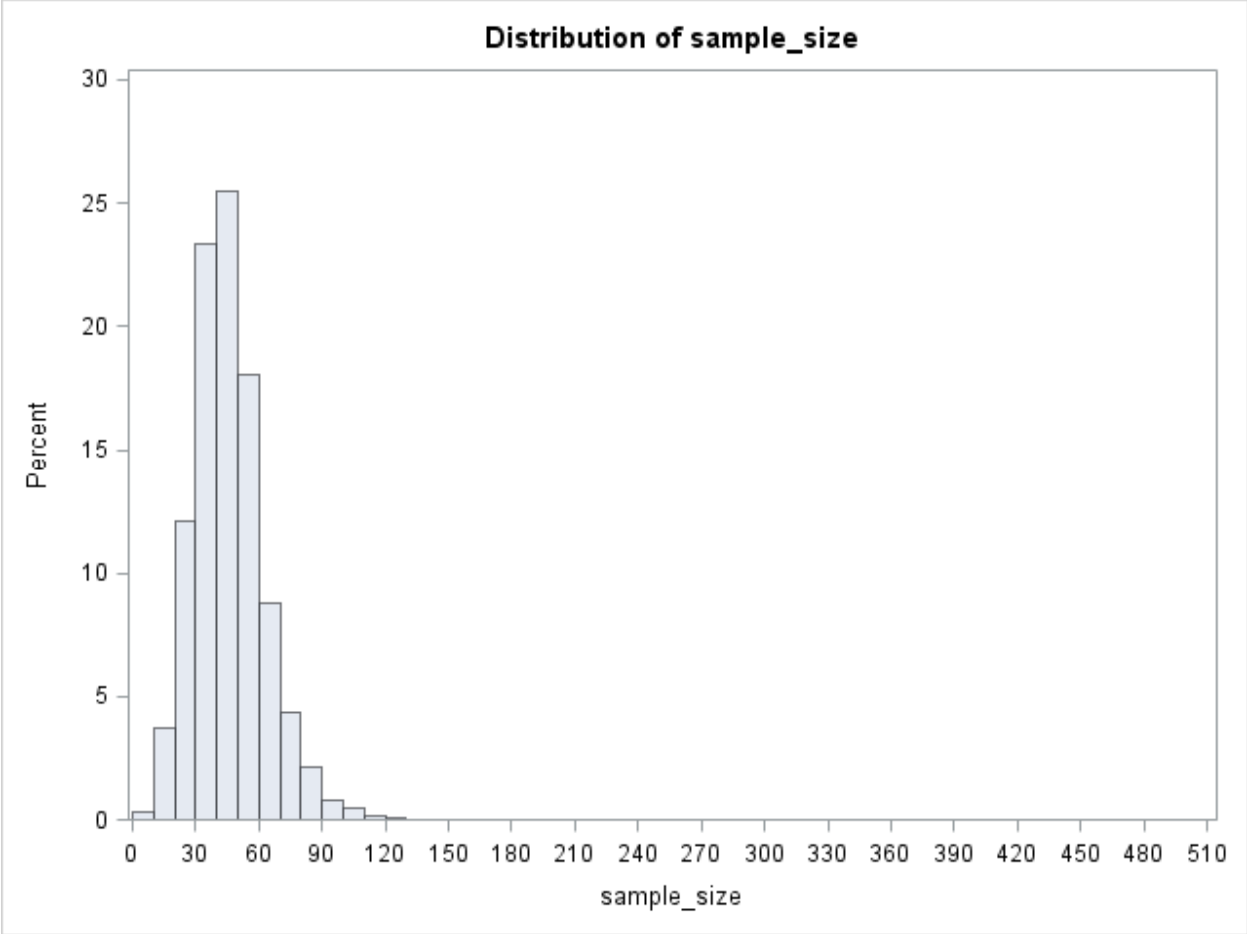


Figure A.2 Distribution of Sample Sizes for Setting 2 (Expected Variance: Pilot, method: control, $\epsilon^*:1$)

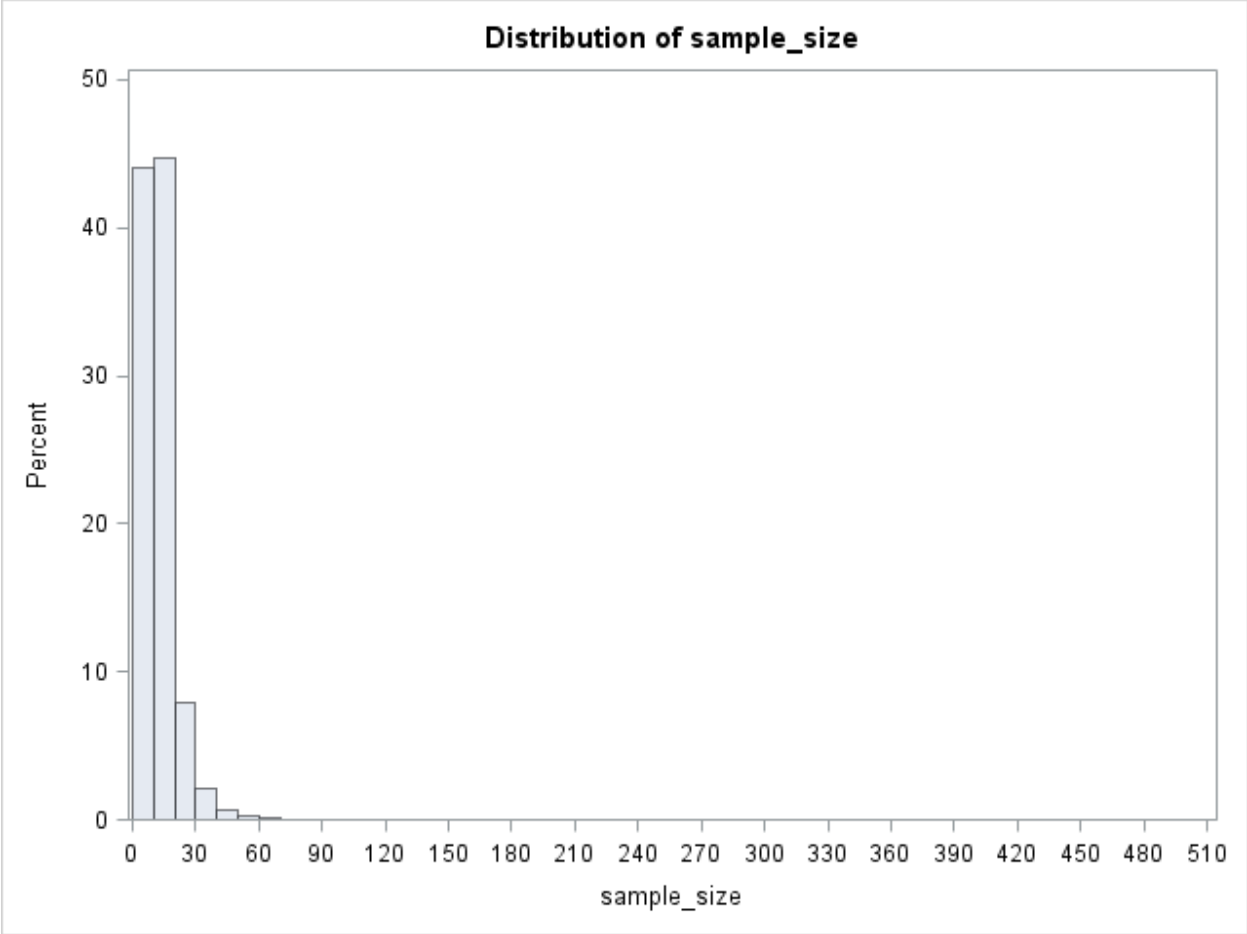


Figure A.3 Distribution of Sample Sizes for Setting 3 (Expected Variance: Pilot, method: control, $\epsilon^*:2$)

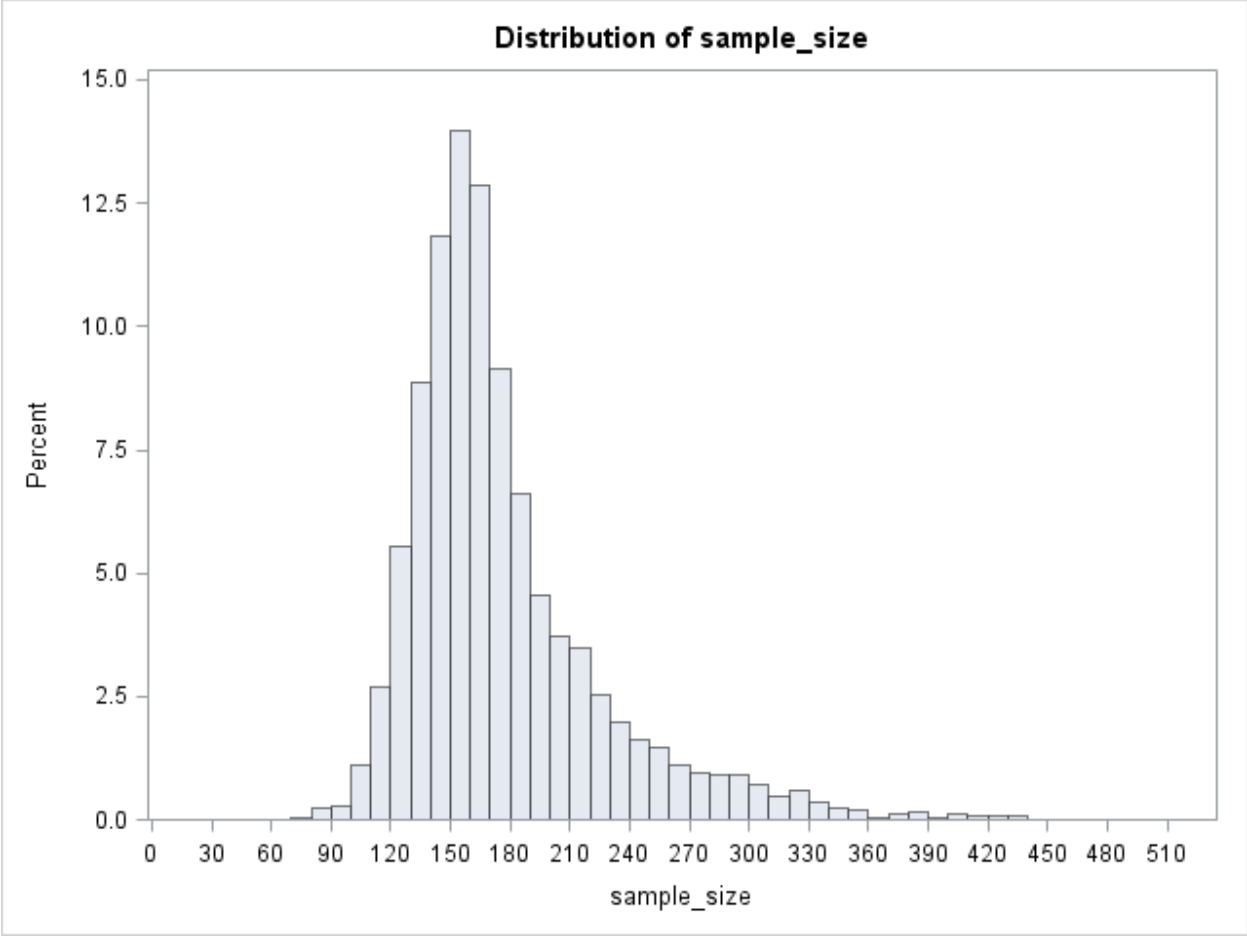


Figure A.4 Distribution of Sample Sizes for Setting 4 (Expected Variance: Pilot, method: pool, $\epsilon^*:0.5$)

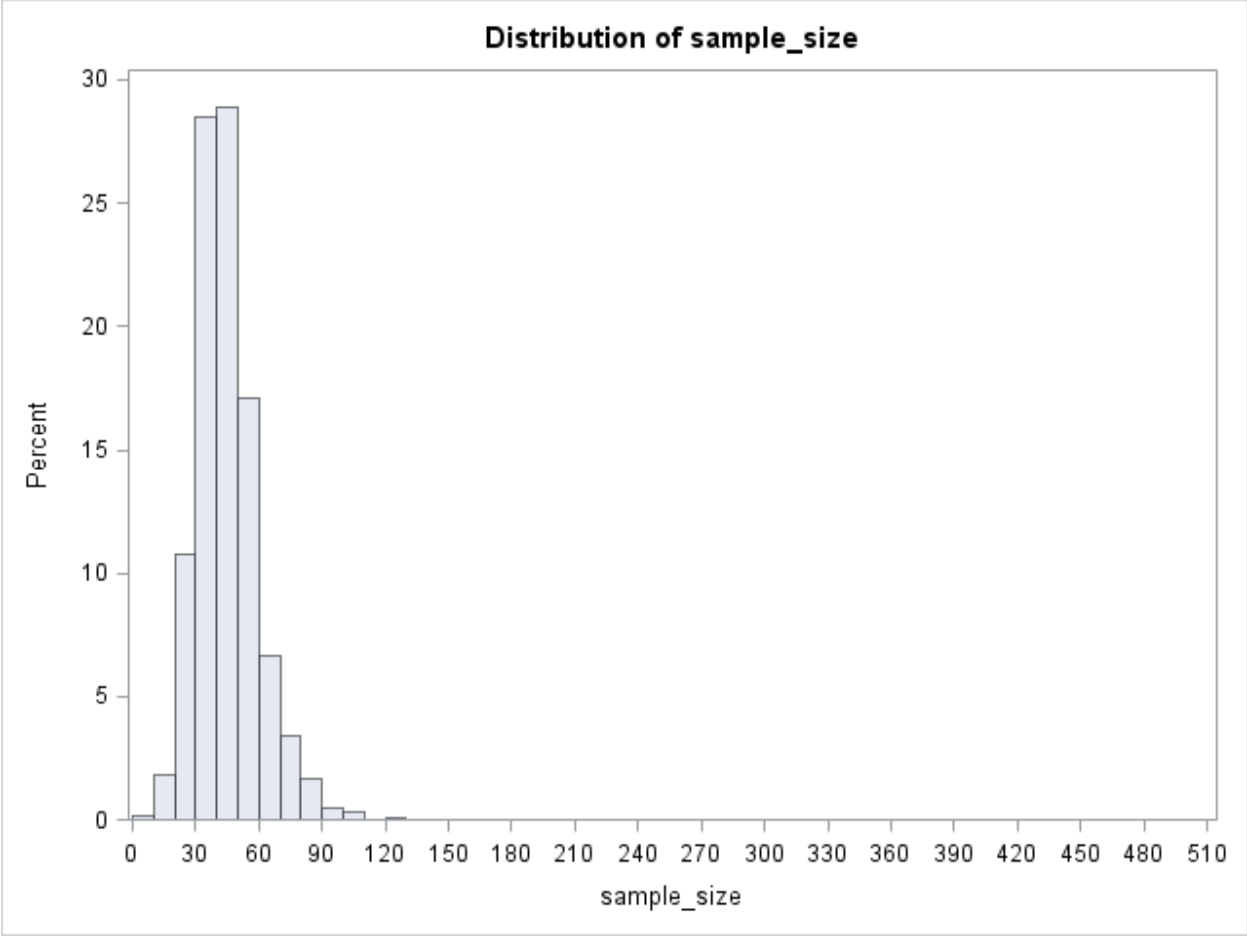


Figure A.5 Distribution of Sample Sizes for Setting 5 (Expected Variance: Pilot, method: pool, $\epsilon^*:1$)

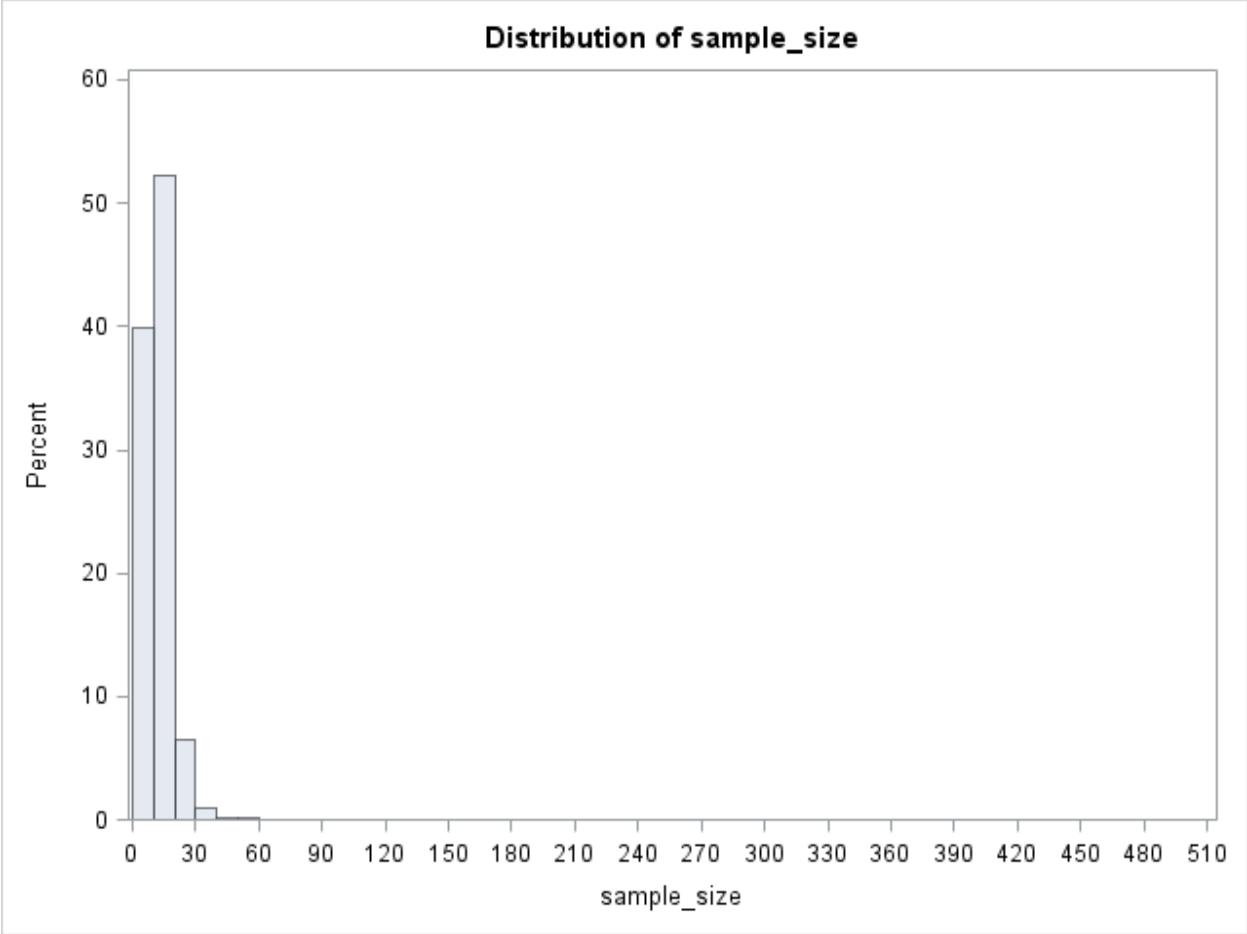


Figure A.6 Distribution of Sample Sizes for Setting 6 (Expected Variance: Pilot, method: pool, $\epsilon^*:2$)

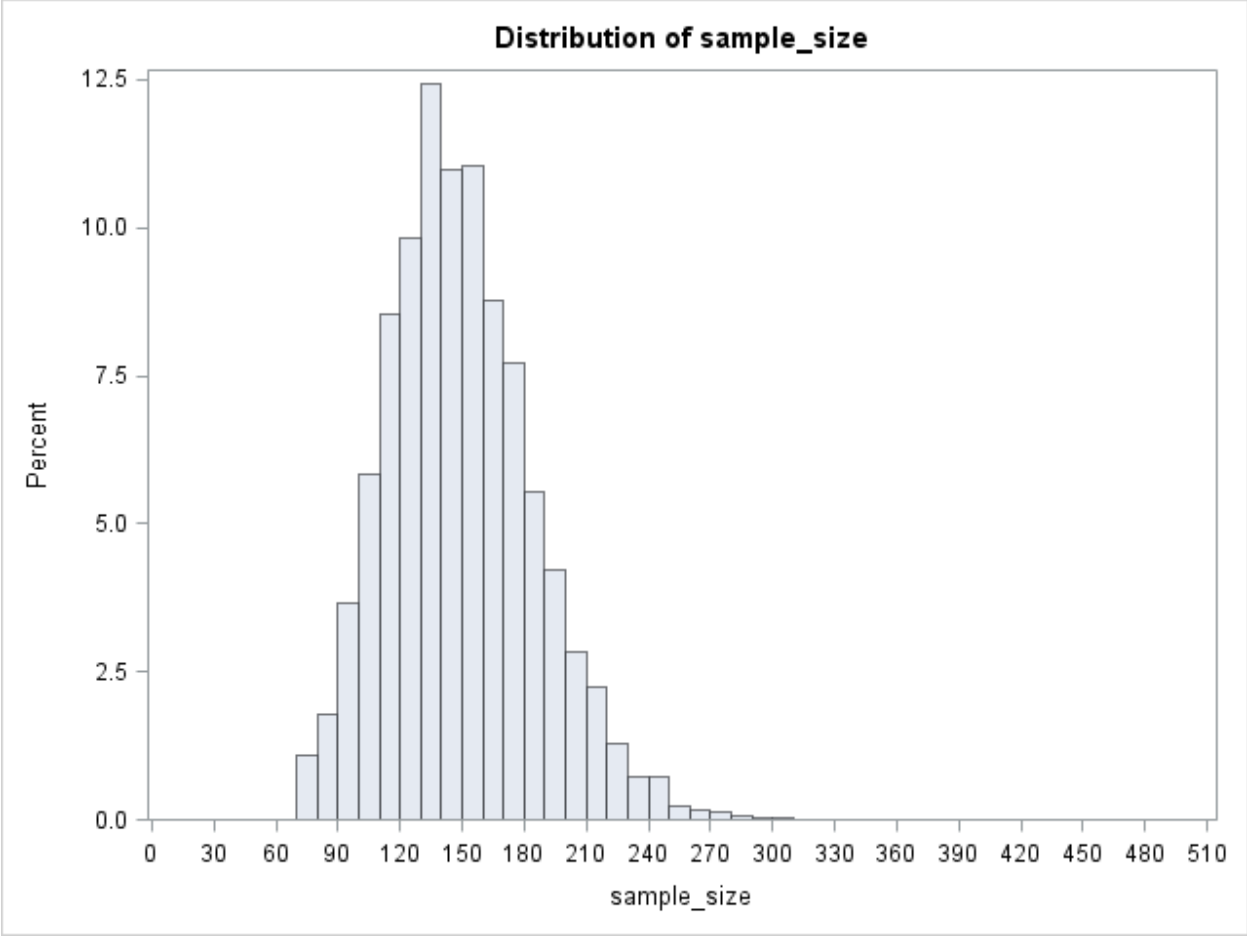


Figure A.7 Distribution of Sample Sizes for Setting 7 (Expected Variance: 1.5, method: control, $\epsilon^*:0.5$)

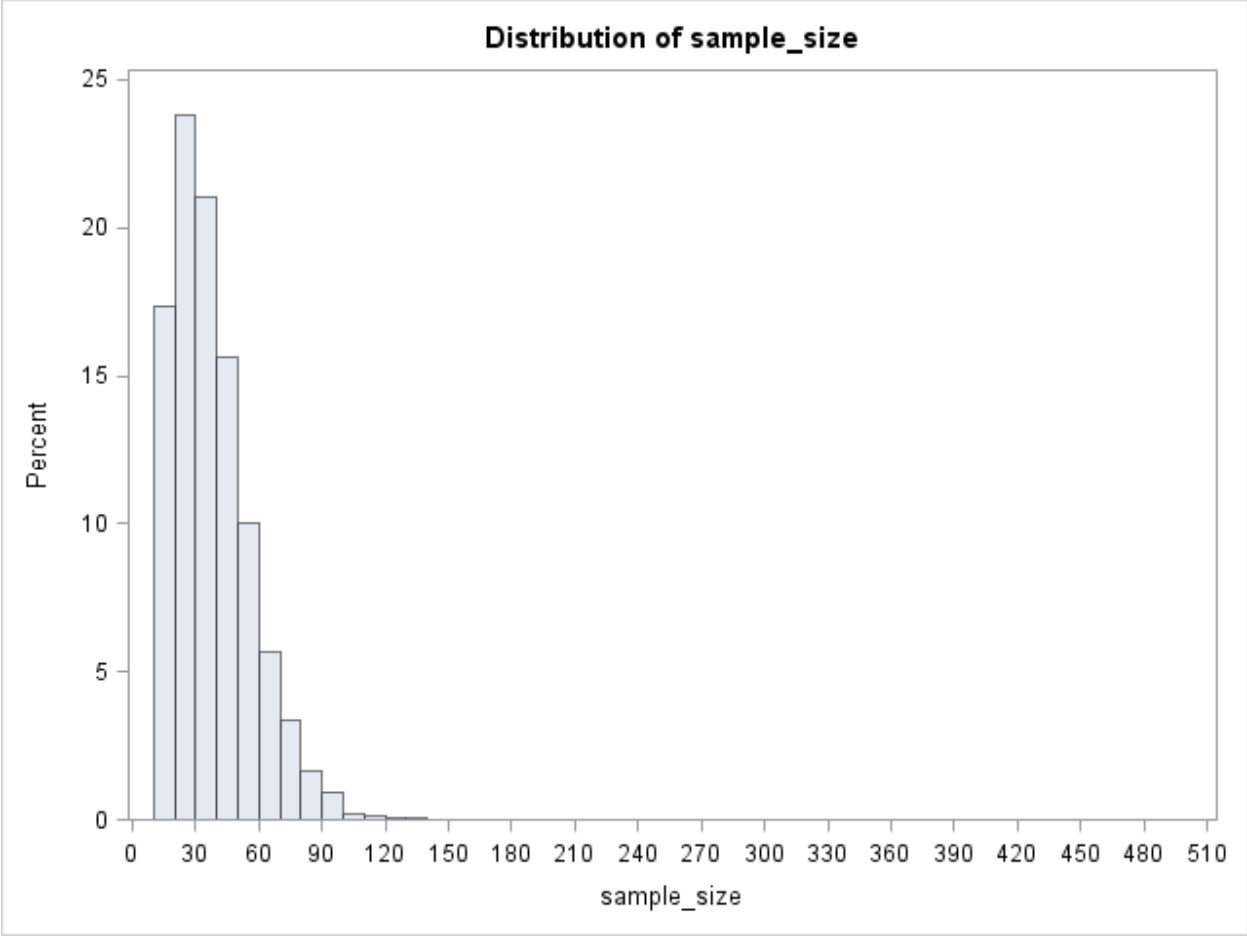


Figure A.8 Distribution of Sample Sizes for Setting 8 (Expected Variance: 1.5, method: control, $\epsilon^*:1$)

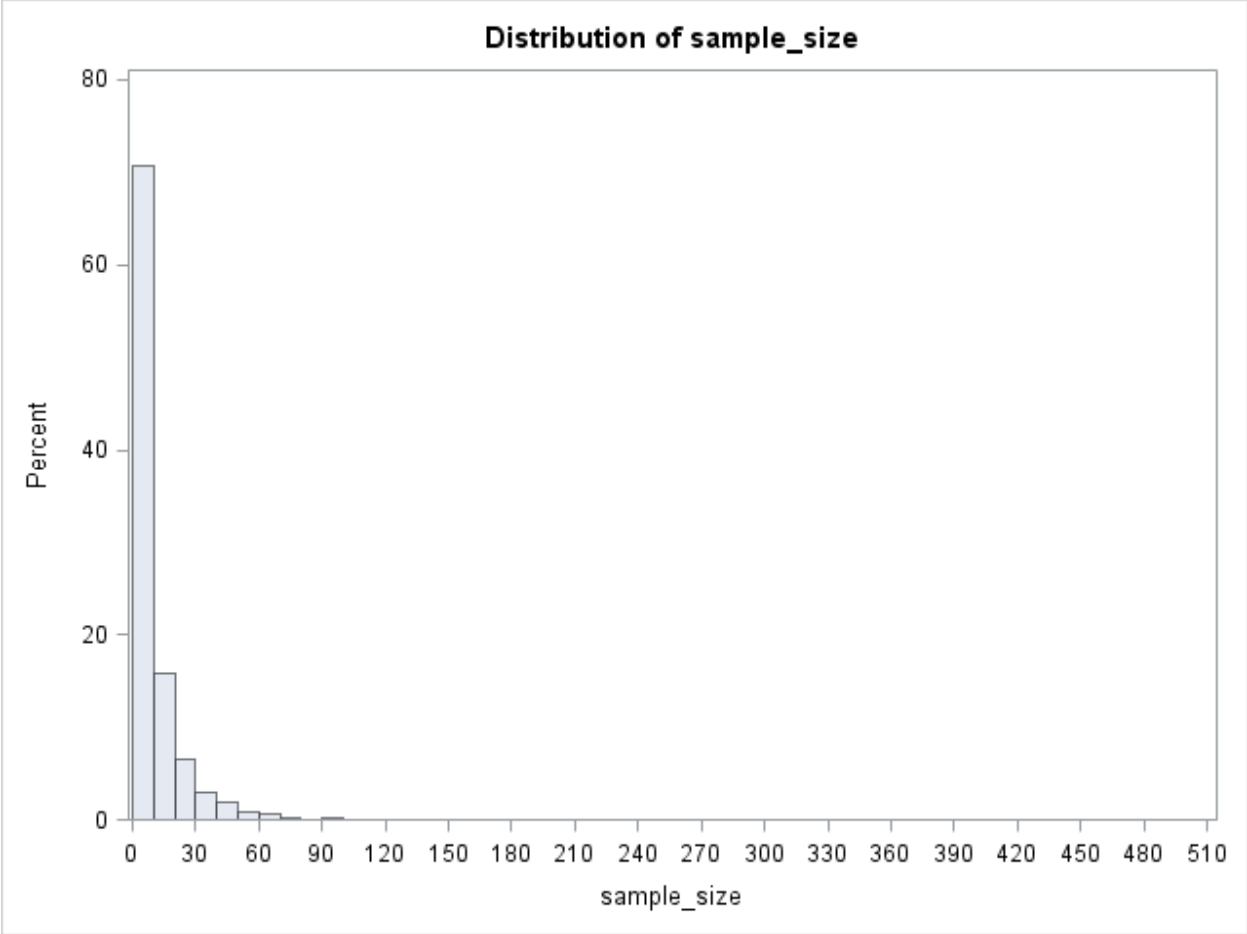


Figure A.9 Distribution of Sample Sizes for Setting 9 (Expected Variance: 1.5, method: control, $\epsilon^*:2$)

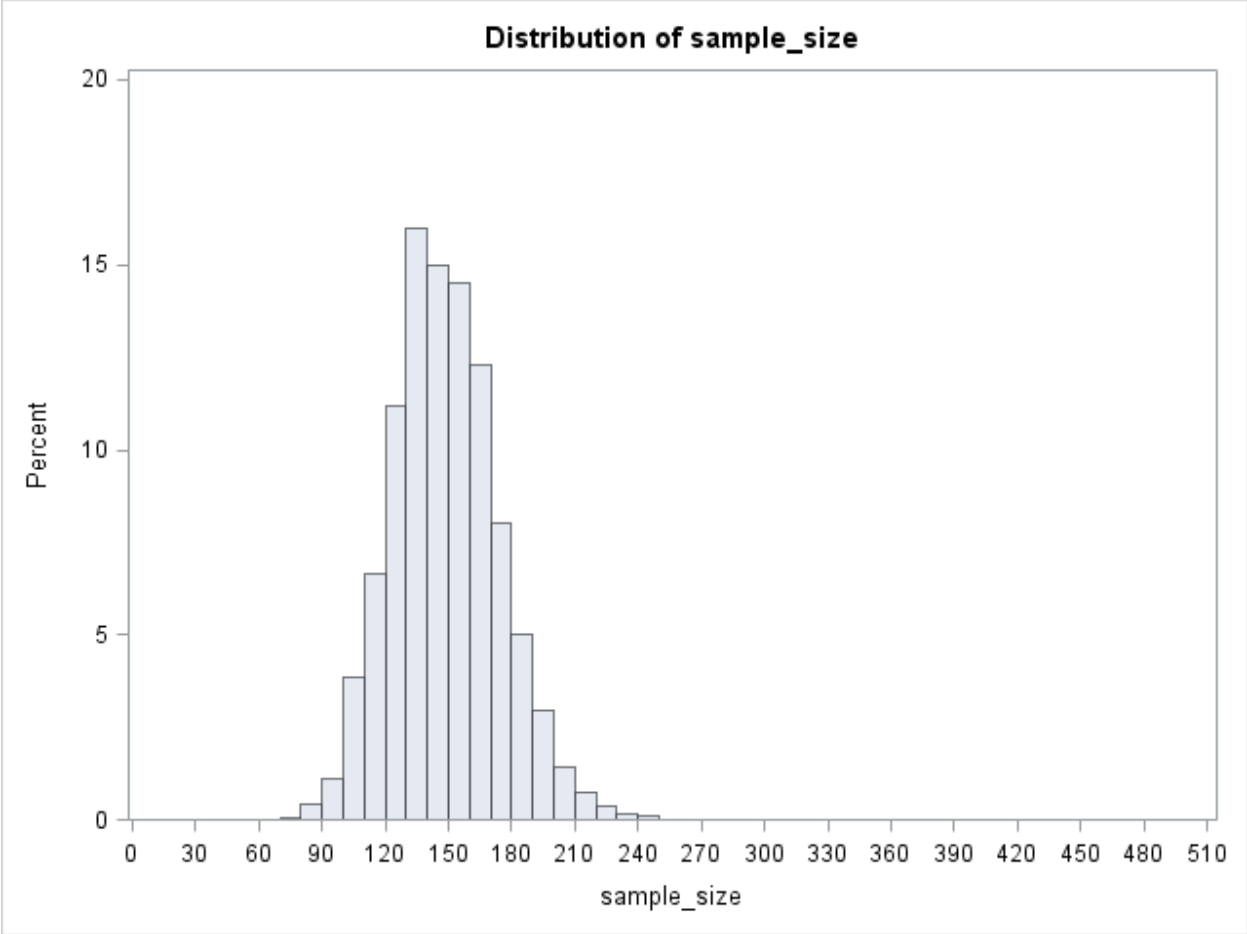


Figure A.10 Distribution of Sample Sizes for Setting 10 (Expected Variance: 1.5, method: pool, $\epsilon^*:0.5$)

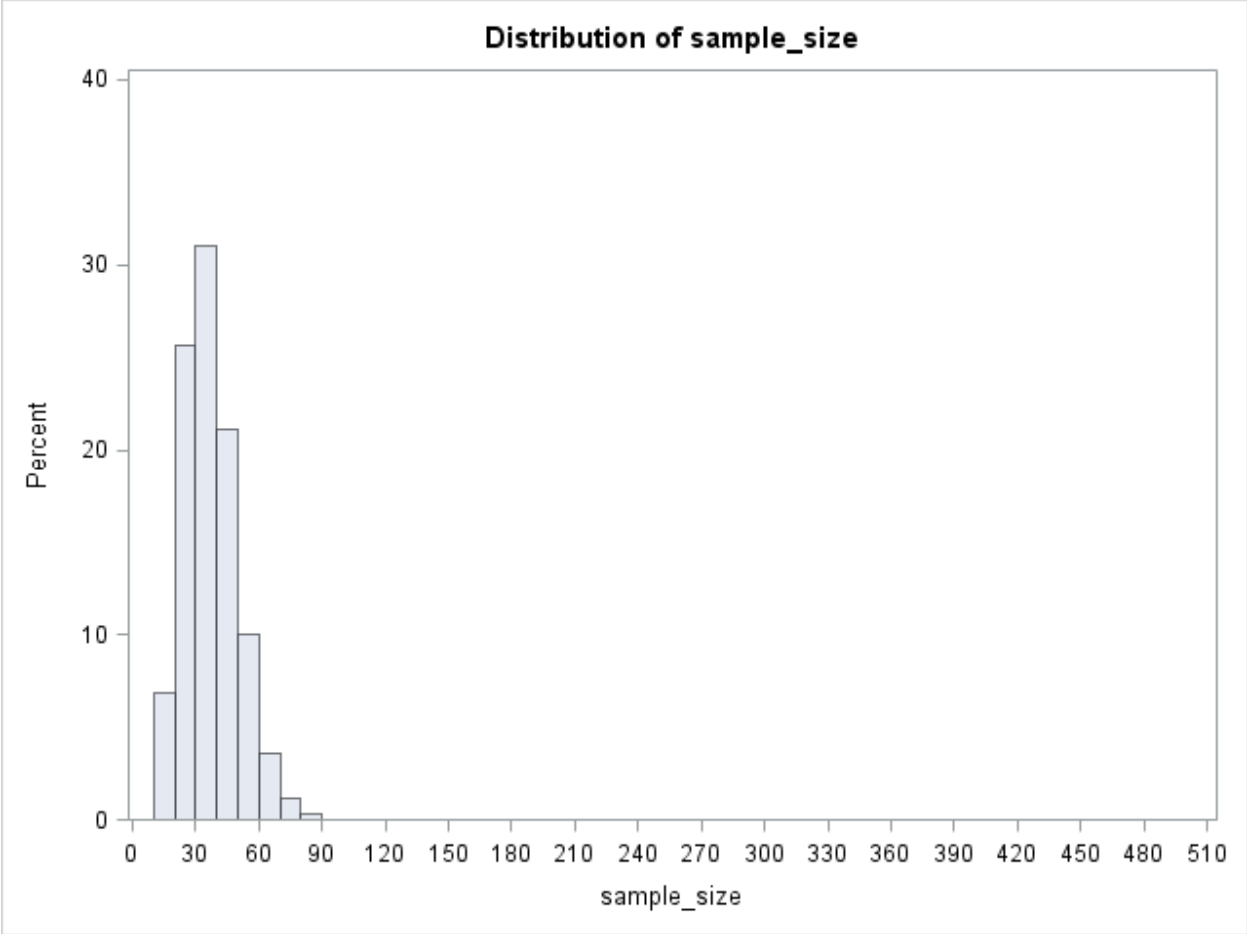


Figure A.11 Distribution of Sample Sizes for Setting 11 (Expected Variance: 1.5, method: pool, $\epsilon^*:1$)

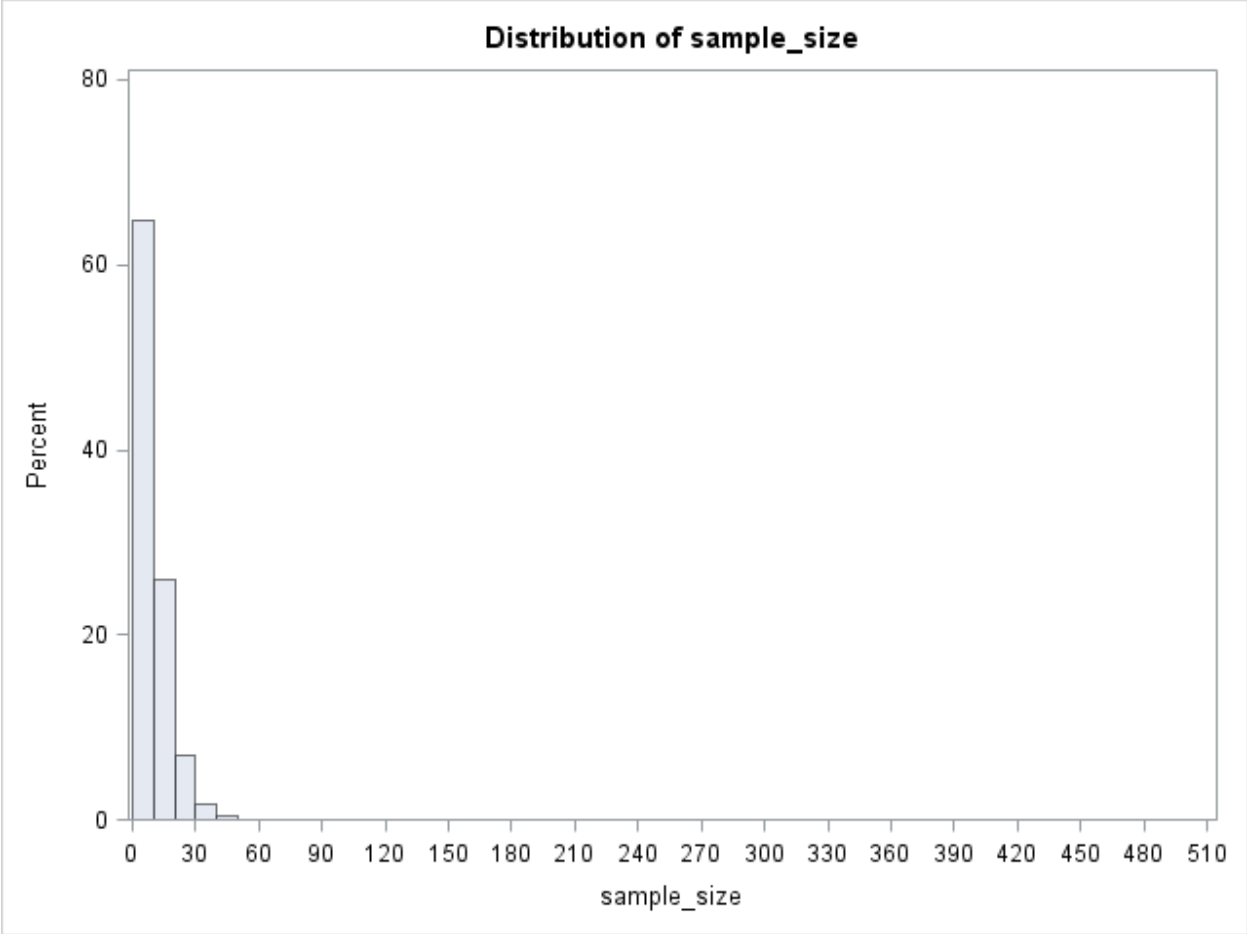


Figure A.12 Distribution of Sample Sizes for Setting 12 (Expected Variance: 1.5, method: pool, $\epsilon^*:2$)

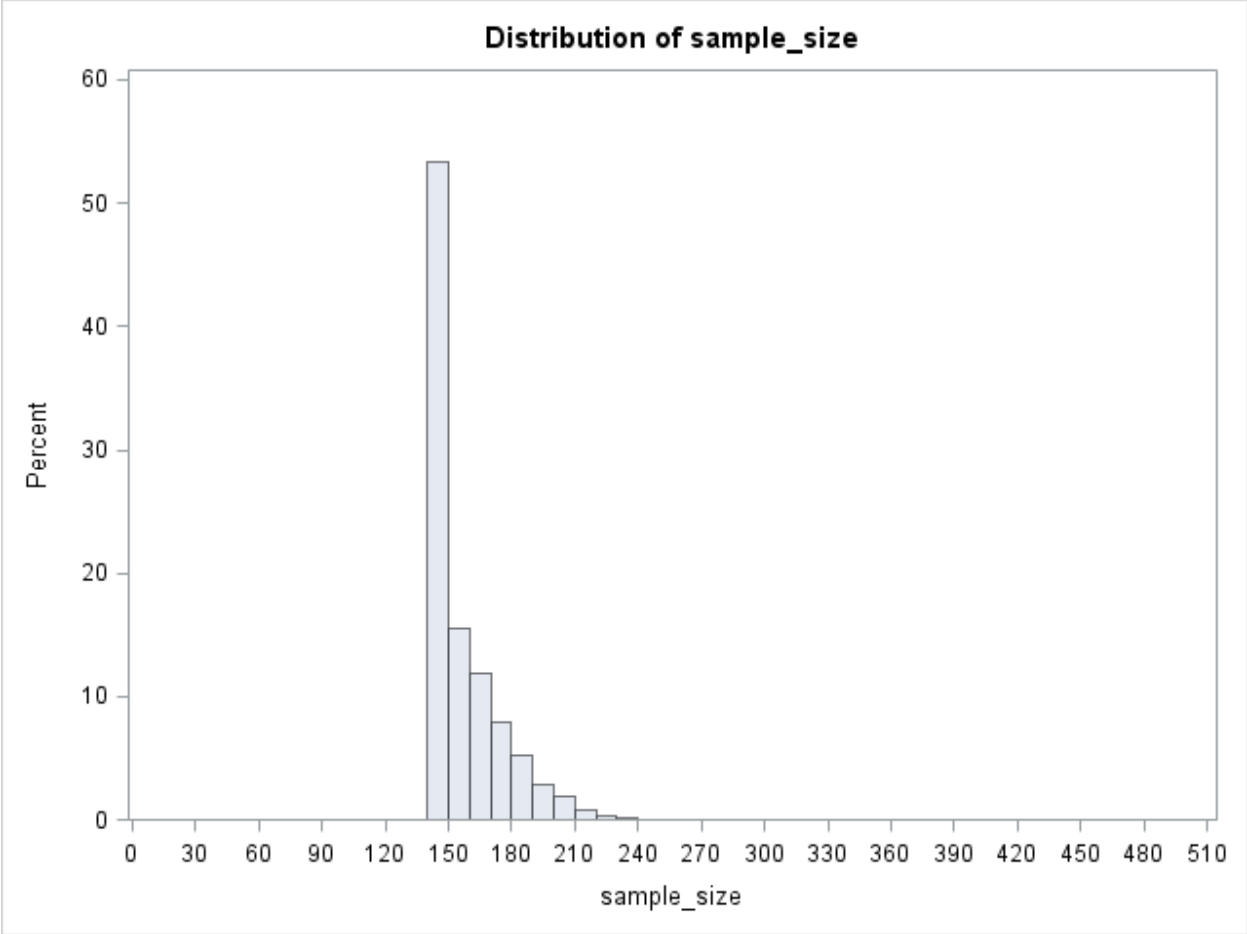


Figure A.13 Distribution of Sample Sizes for Setting 13 (Expected Variance: 3, method: control, $\epsilon^*:0.5$)

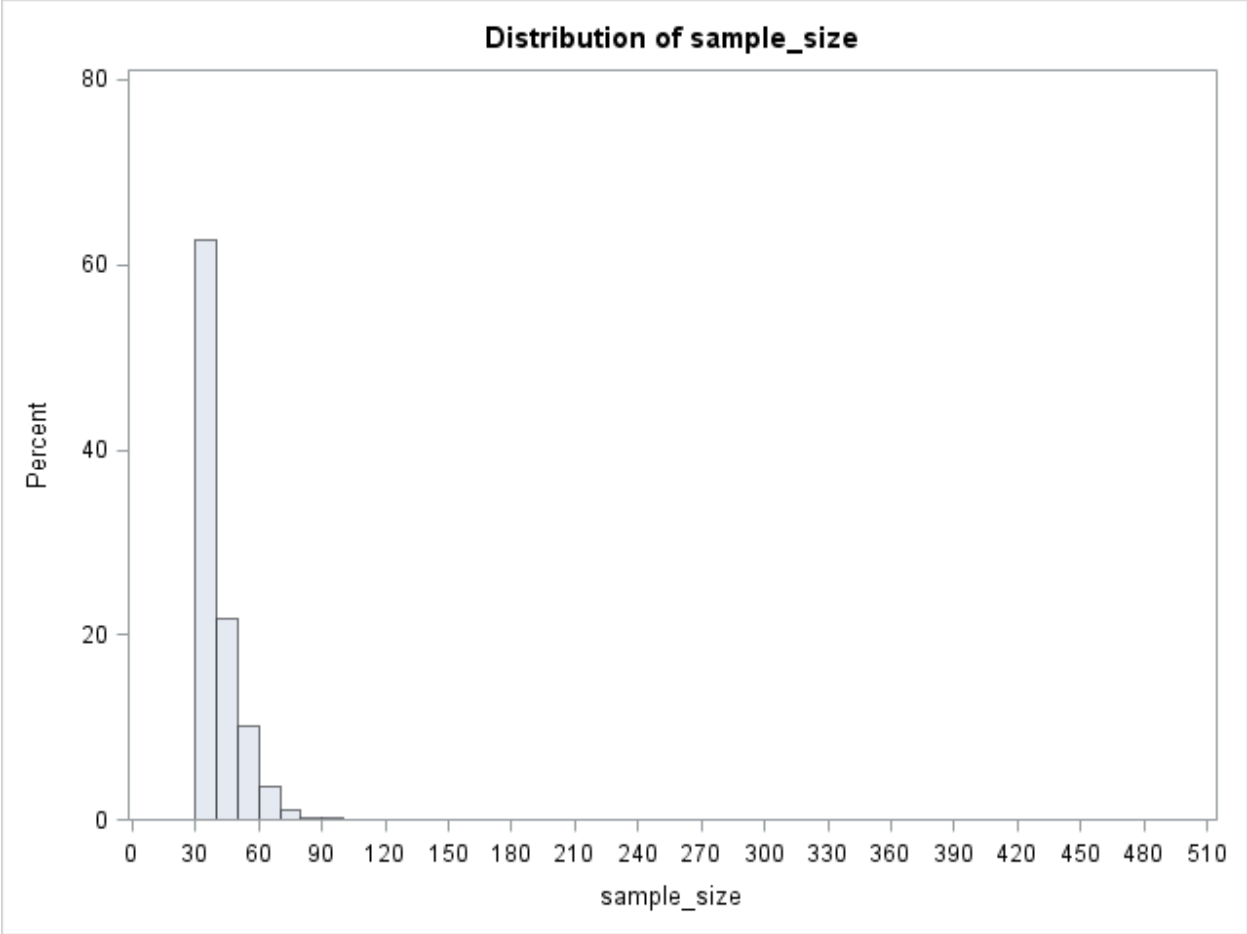


Figure A.14 Distribution of Sample Sizes for Setting 14 (Expected Variance: 3, method: control, $\epsilon^*:1$)

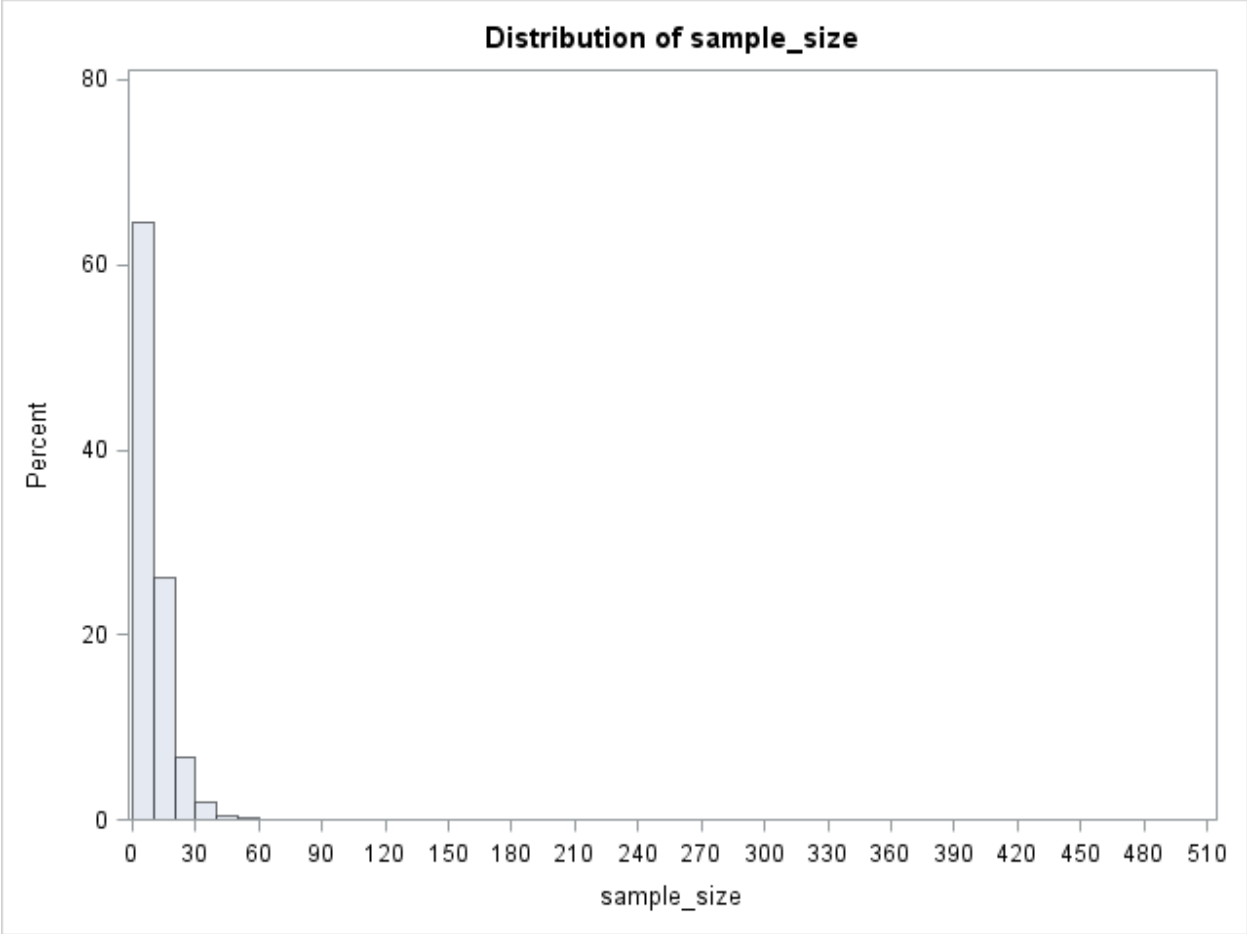


Figure A.15 Distribution of Sample Sizes for Setting 15 (Expected Variance: 3, method: control, $\epsilon^*:2$)

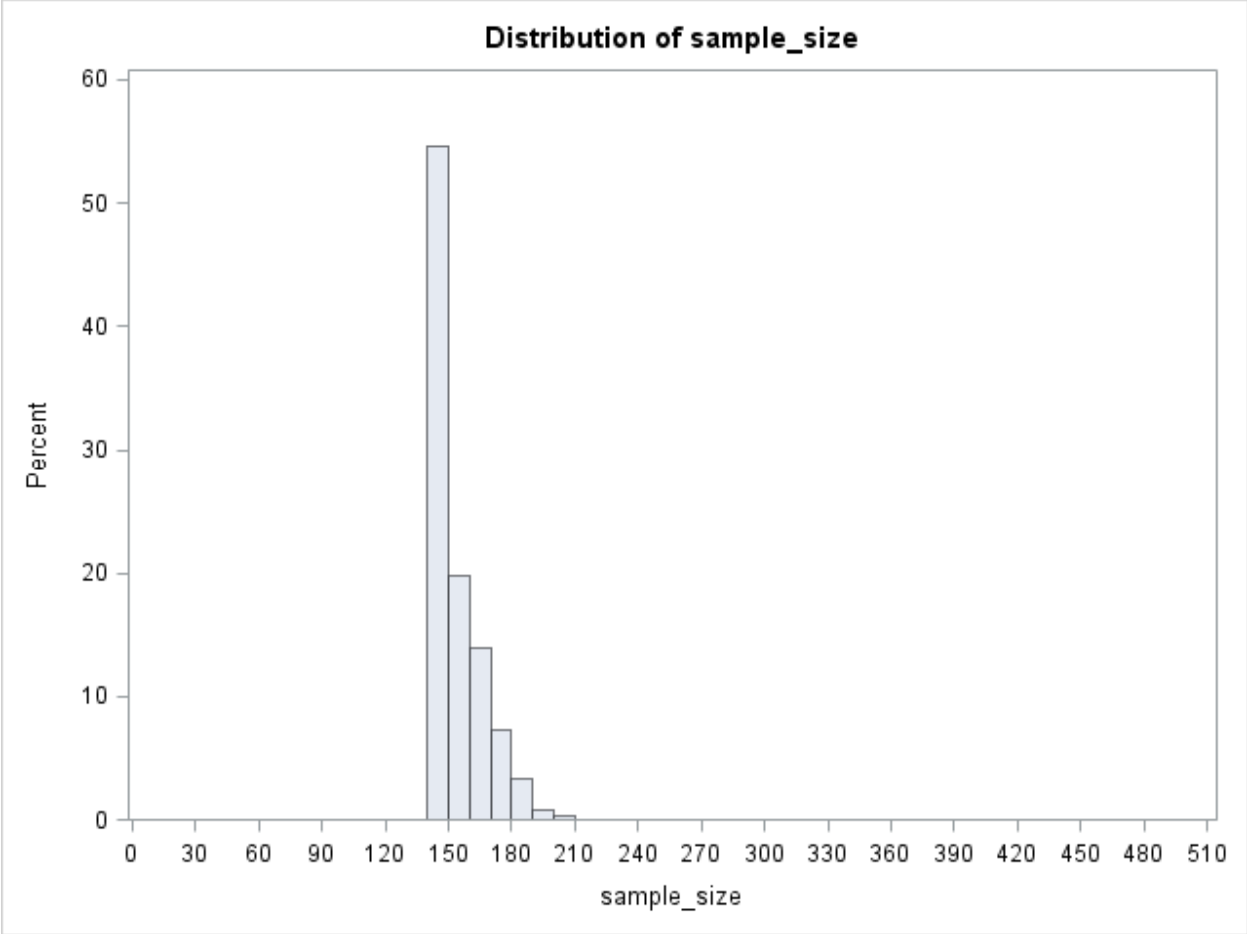


Figure A.16 Distribution of Sample Sizes for Setting 16 (Expected Variance: 3, method: pool, ϵ^* :0.5)

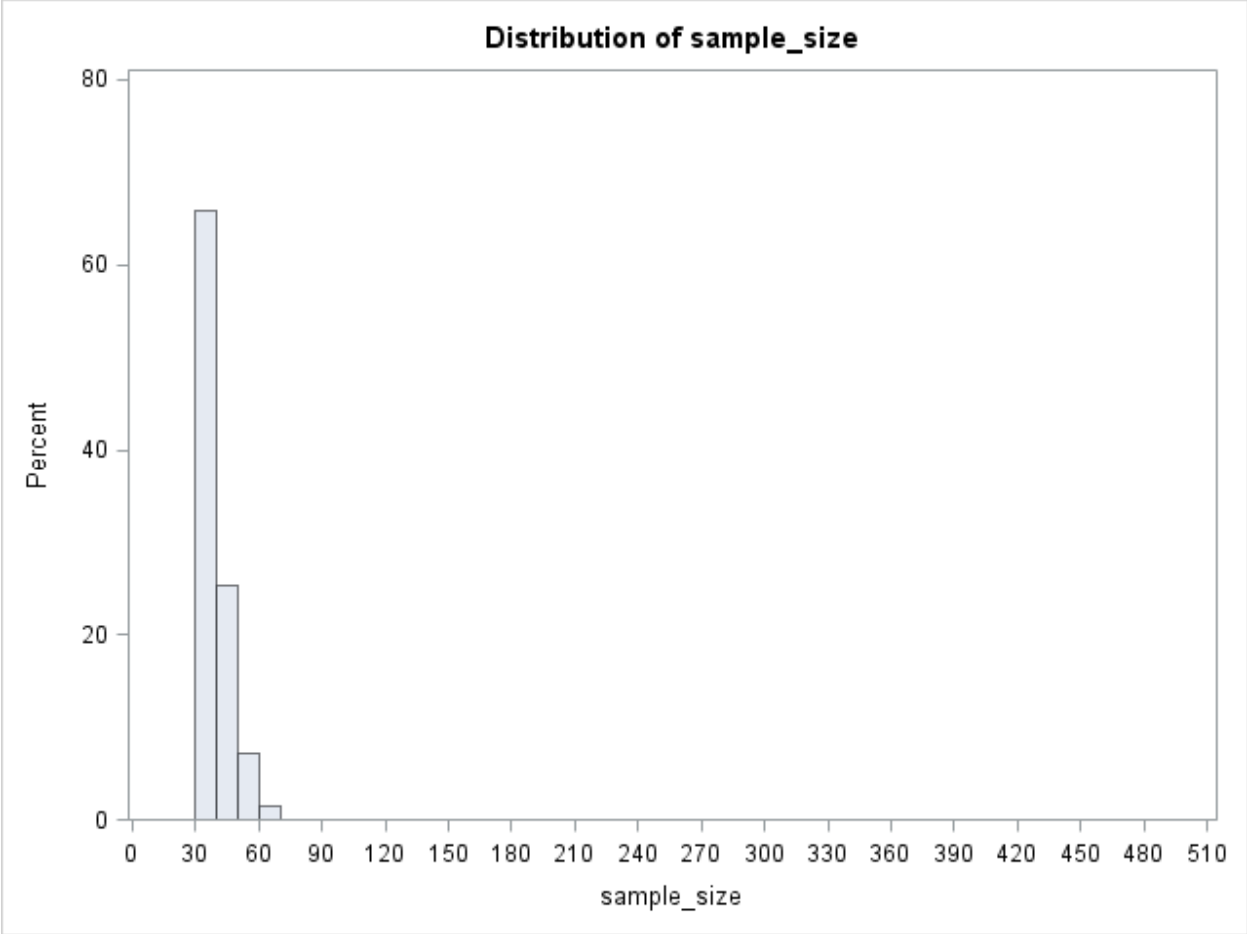


Figure A.17 Distribution of Sample Sizes for Setting 17 (Expected Variance: 3, method: pool, $\epsilon^*:1$)

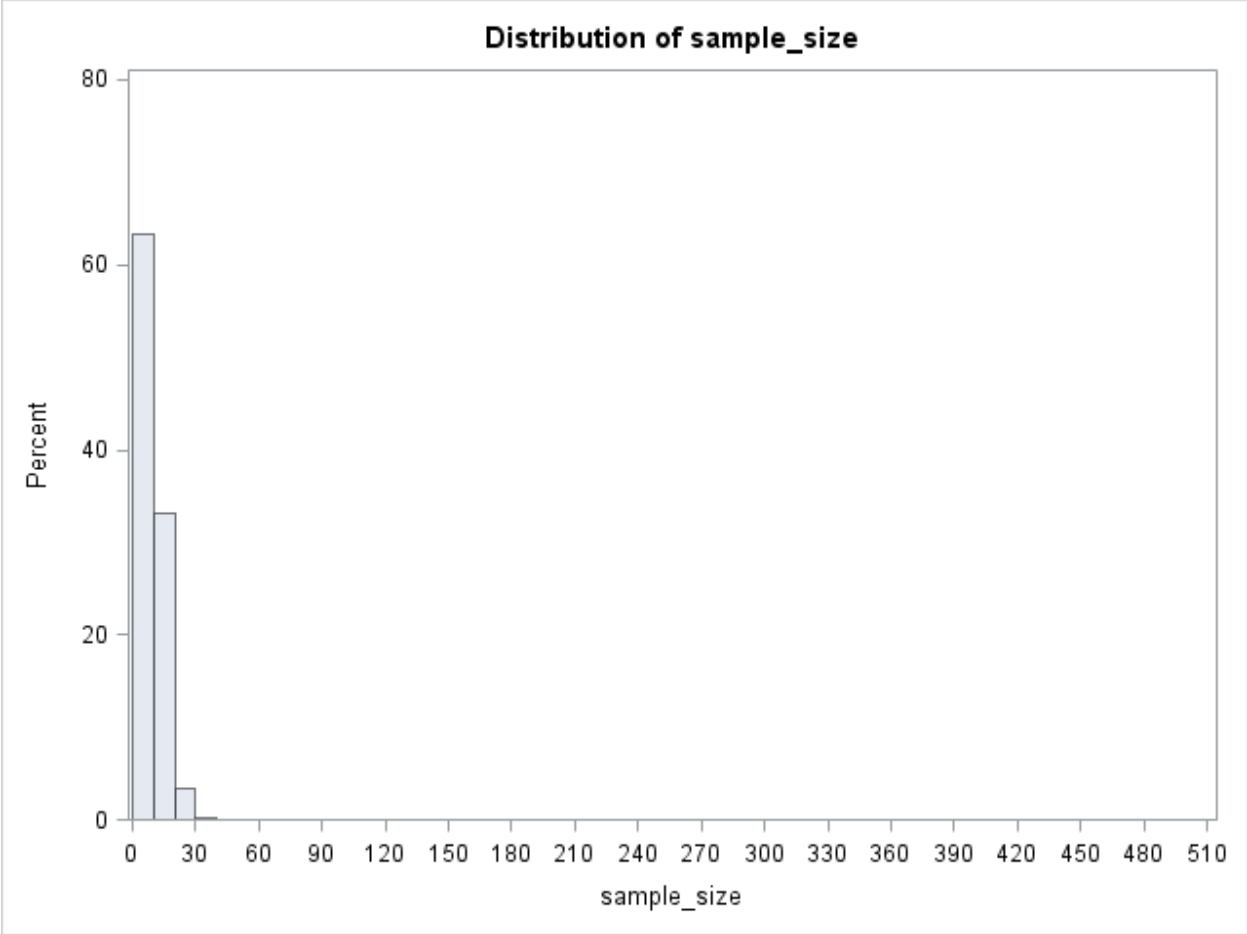


Figure A.18 Distribution of Sample Sizes for Setting 18 (Expected Variance: 3, method: pool, $\epsilon^*:2$)

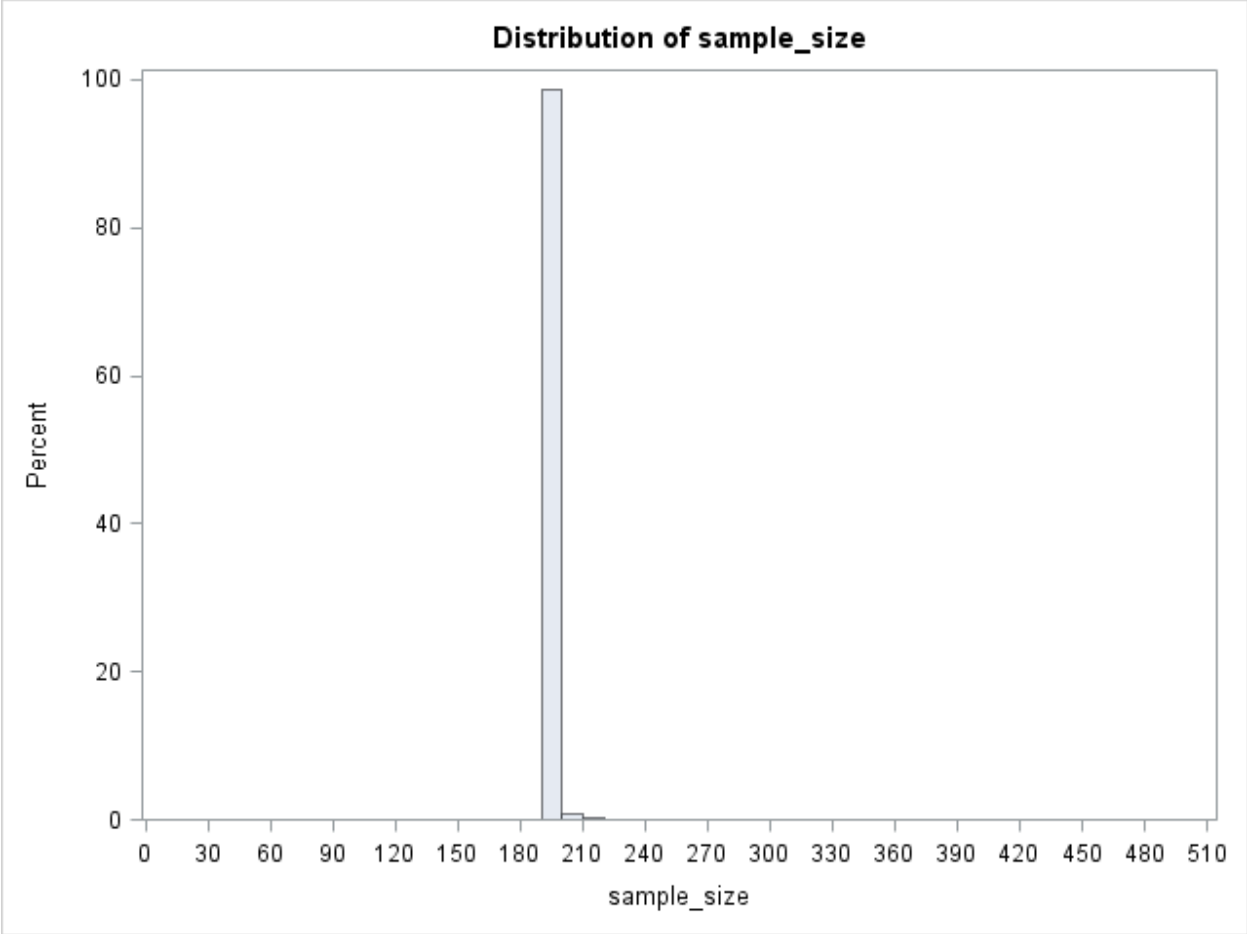


Figure A.19 Distribution of Sample Sizes for Setting 19 (Expected Variance: 4, method: control, $\epsilon^*:0.5$)

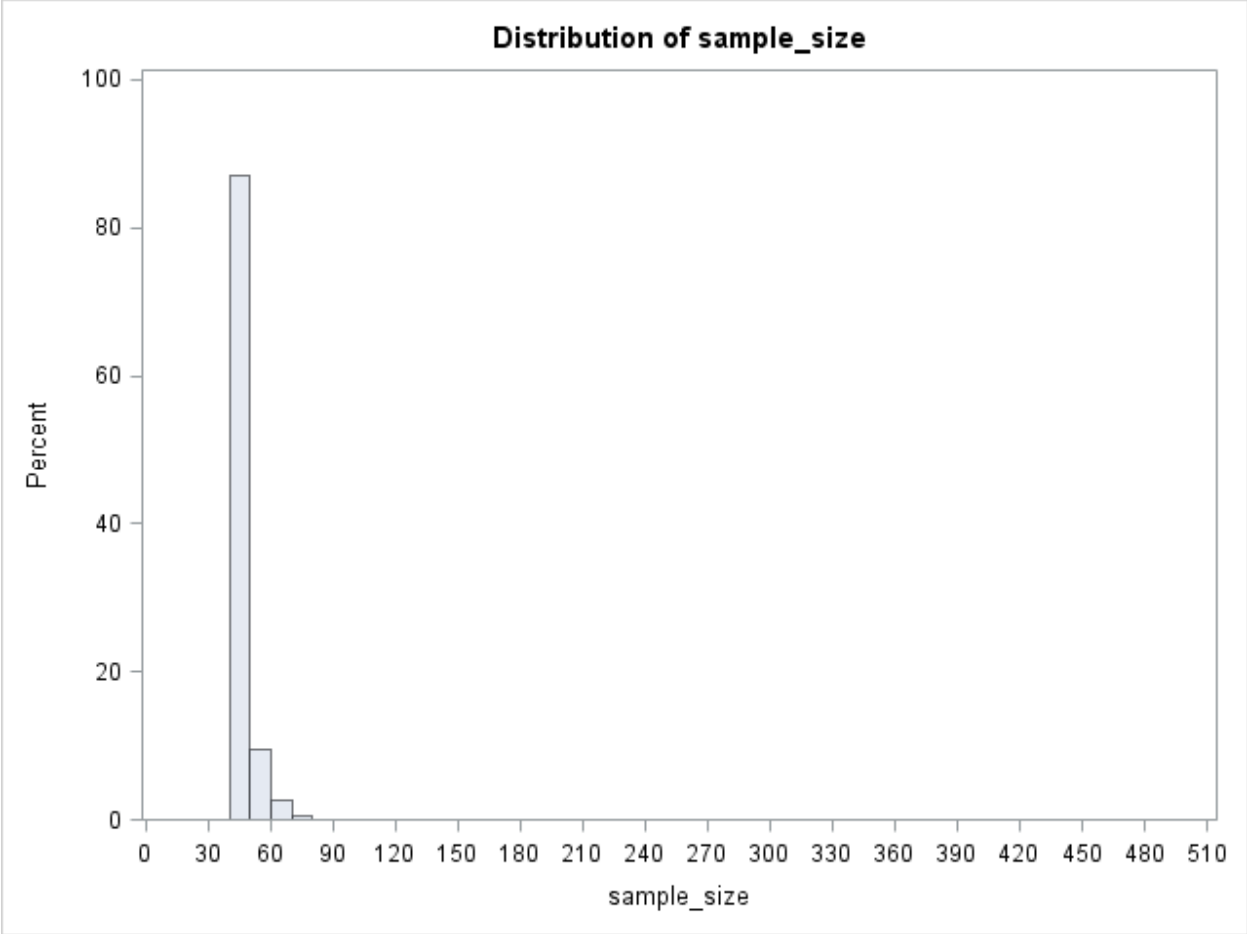


Figure A.20 Distribution of Sample Sizes for Setting 20 (Expected Variance: 4, method: control, $\epsilon^*:1$)

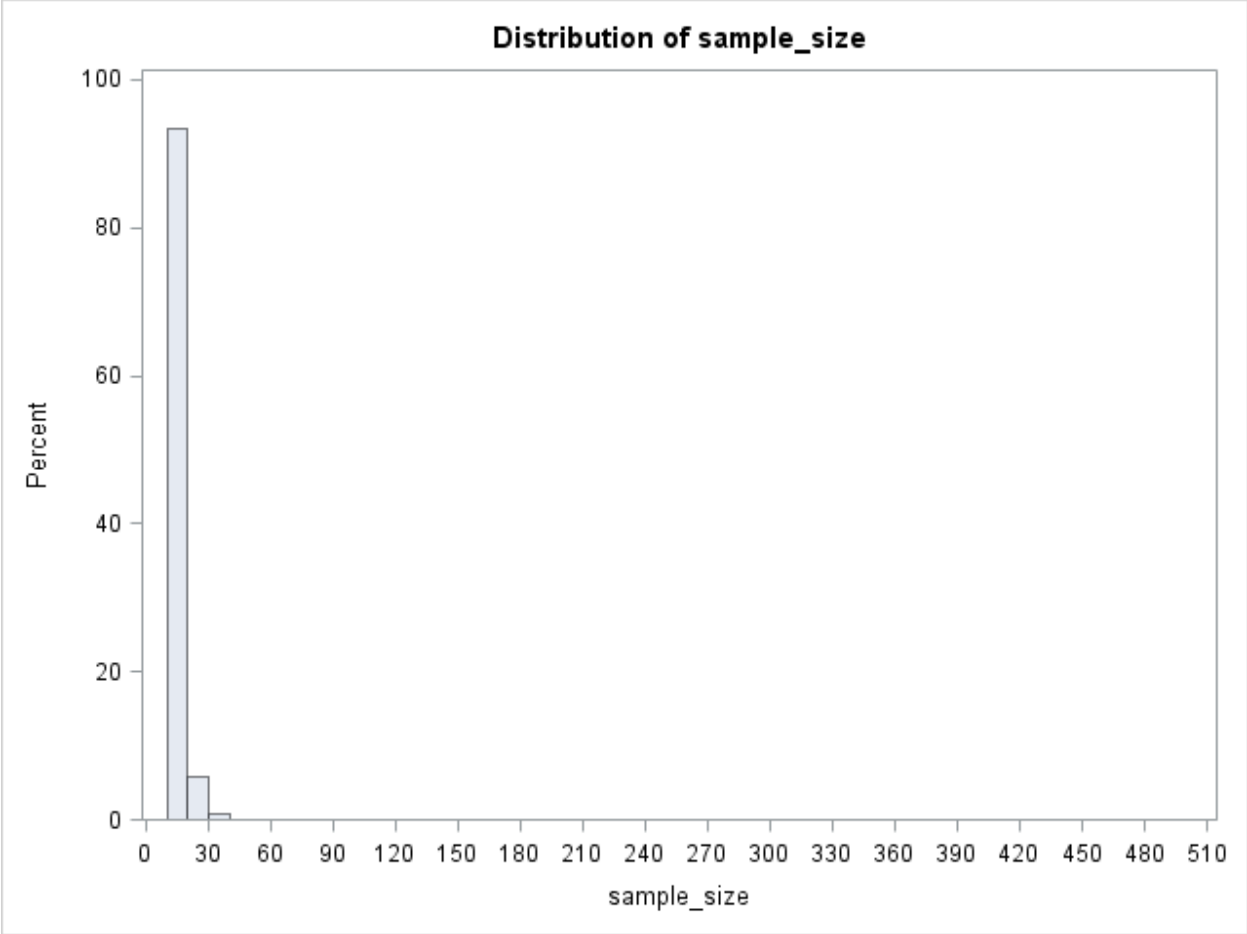


Figure A.21 Distribution of Sample Sizes for Setting 21 (Expected Variance: 4, method: control, $\epsilon^*:2$)

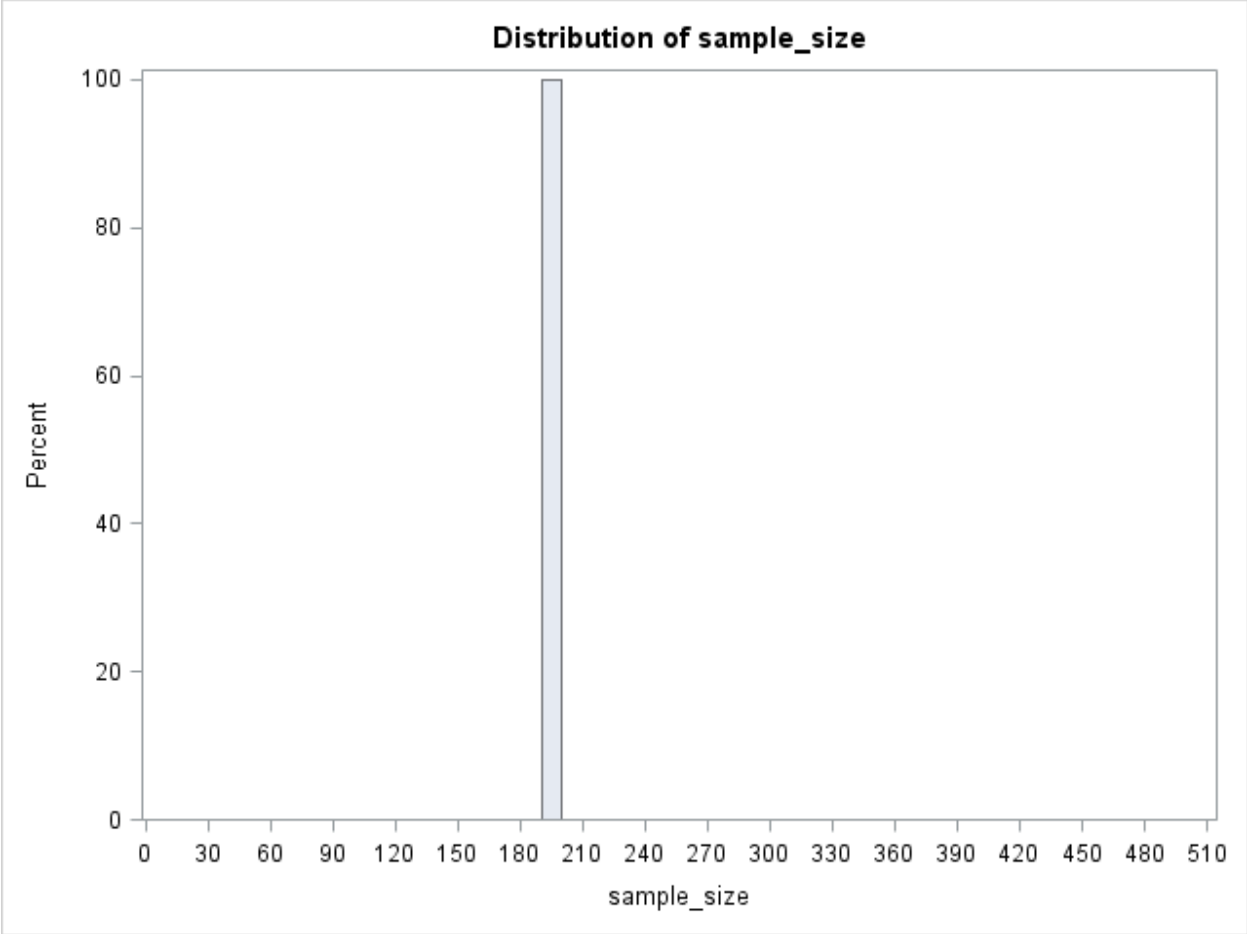


Figure A.22 Distribution of Sample Sizes for Setting 22 (Expected Variance: 4, method: pool, ϵ^* :0.5)

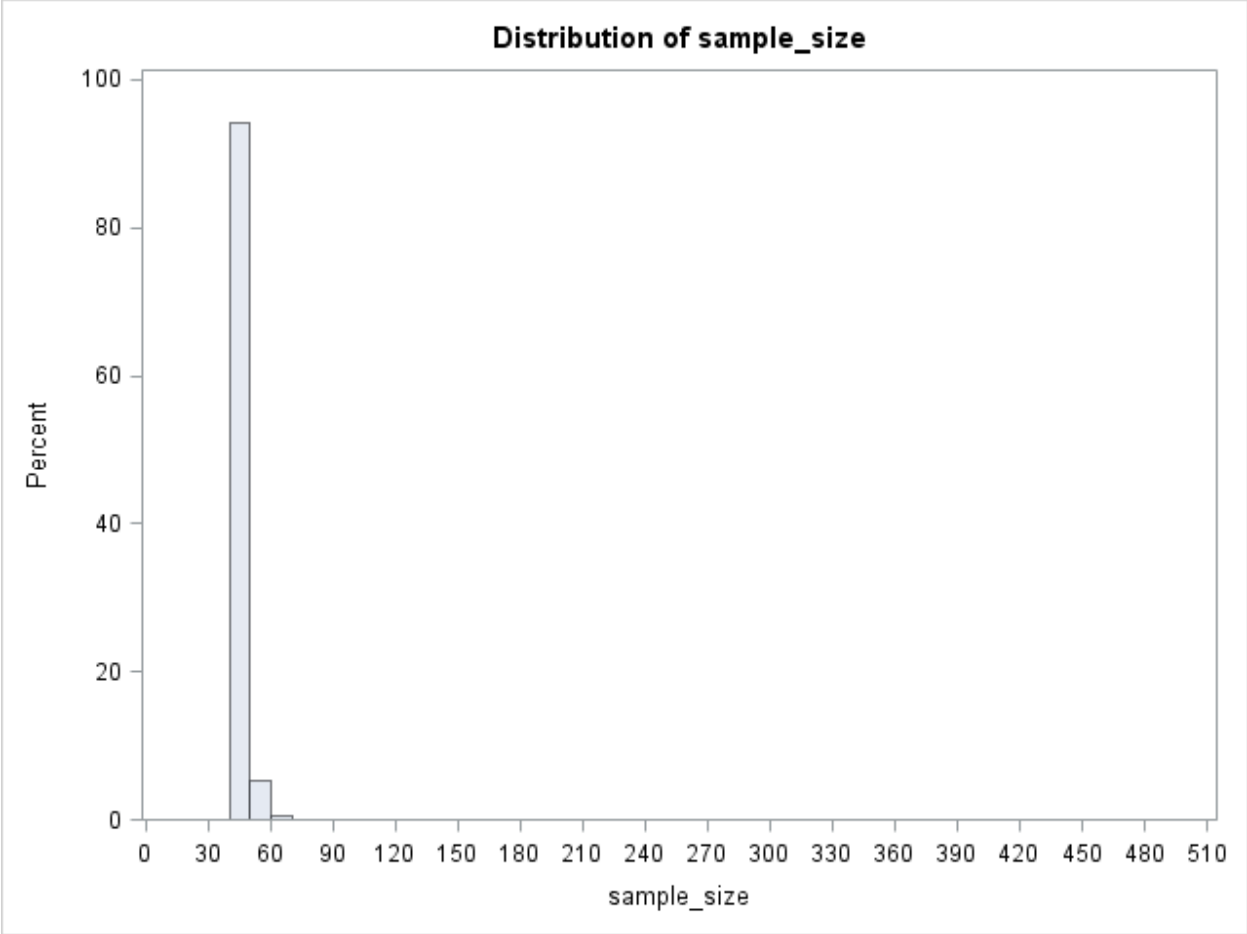


Figure A.23 Distribution of Sample Sizes for Setting 23 (Expected Variance: 4, method: pool, $\epsilon^*:1$)

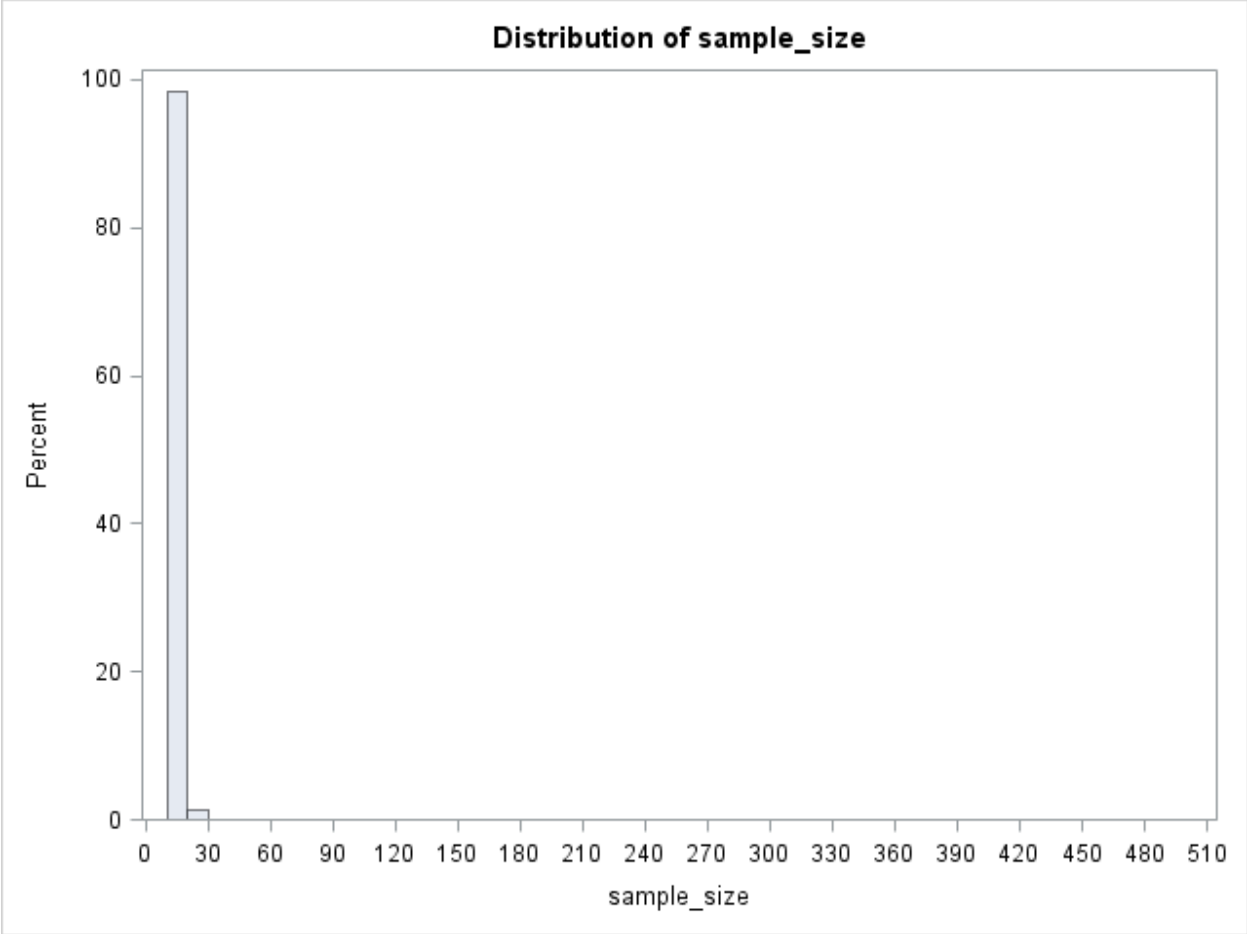


Figure A.24 Distribution of Sample Sizes for Setting 24 (Expected Variance: 4, method: pool, $\epsilon^*:2$)

Appendix B - Simulation SAS Code

Below is the SAS code used in the simulation studies.

```
libname master 'H:master report debug';

options nonotes nosource nosource2 errors=0;

proc sql;
  create table master.sim_result
    (method char(4),
     type char(5),
     mean num,
     true_variance num,
     guess_variance num,
     epslon num,
     t_test num,
     reject num,
     init_size num,
     real_size num);

  create table master.sim_summary
    (method char(4),
     type char(5),
     true_variance num,
     guess_variance num,
     epslon num,
     type1_error num);
run;

%let n_preexam = 10;

%macro sim(true_mu,true_var=,epslon=,index=,pilot=,guess_var=,method=);

/*pilot data*/
/*%if &pilot='pilot' %then
%do;*/
data pretest(drop=i);
  if &pilot='pilot' then
  do;
    call streaminit(10904+&index);
    do i = 1 to &n_preexam;
      temp = rand("Normal", &true_mu, sqrt(&true_var));
      output;
    end;
  end;
run;

proc MEANS data=pretest NOPRINT;
output out=pre_result std=std_star;
run;
/*%end;*/
```

```

data init;
if &pilot='pilot' then
    do;
        set pre_result;
    end;
else
    do;
        std_star = sqrt(&guess_var);
    end;

alpha = 0.05;
beta = 0.2;
delta = 0;
z_alpha = quantile("Normal", alpha);
z_beta = quantile("Normal", beta);
n_star = floor((z_alpha + z_beta)**2*std_star**2*2/(&epsilon - delta)**2);
call symput('init_size', n_star);
n1 = floor(n_star/2);
n2 = n1; /* test group*/
if &pilot='pilot' then
    do;
        call symput('var_star', 0);
    end;
else
    do;
        call symput('var_star', std_star**2);
    end;

run;

data pre_half_c(keep=result);
set init;
call streaminit(20905+&index);
do i = 1 to n1;
    result = rand("Normal", &>true_mu, sqrt(&>true_var));
    output;
end;
run;

data pre_half_t(keep=result);
set init;
call streaminit(30906+&index);
do i = 1 to n2;
    result = rand("Normal", &>true_mu, sqrt(&>true_var));
    output;
end;
run;

/*method pool*/

data pool;
if &method='pool' then
    do;
        set pre_half_c pre_half_t;
    end;
else

```

```

do;
    set pre_half_c;
end;
run;

proc MEANS data=pool NOPRINT;
output out=post_result mean=mu std=std;
run;

data newdata;
set init;
set post_result;
n_temp = floor((z_alpha + z_beta)**2*std**2/(&epsilon - delta)**2);
n = max(n_star, n_temp);
call symput('real_size', n);
n_total = 2*n;
n_rest = n - n1;
run;

data posttrial_c(keep=result);
set newdata;
call streaminit(40907+&index);
do i = 1 to n_rest;
    result = rand("Normal", &true_mu, sqrt(&true_var));
    output;
end;
run;

data posttrial_t(keep=result);
set newdata;
call streaminit(50908+&index);
do i = 1 to n_rest;
    result = rand("Normal", &true_mu, sqrt(&true_var));
    output;
end;
run;

data group_c;
set pre_half_c posttrial_c;
run;

proc MEANS data=group_c NOPRINT;
output out=stat_c mean=mu_c std=std_c;
run;

data group_t;
set pre_half_t posttrial_t;
run;

proc MEANS data=group_t NOPRINT;
output out=stat_t mean=mu_t std=std_t;
run;

data ttest_result;
set stat_c;
set stat_t;
set newdata(keep=n);

```

```

sigma2 = (std_c**2+std_t**2)/2;
t_test = (mu_c-mu_t)/sqrt(2*sigma2/n);
t_star = TINV(0.95,2*n-2);

if t_test > t_star then
    do;
        reject = 1;
    end;
else
    do;
        reject = 0;
    end;
call symput('reject', reject);
call symput('t_test', t_test);
run;

proc SQL;
insert into master.sim_result
values(&method, &pilot, &>true_mu, &>true_var, &var_star, &epsilon, &t_test,
&reject, &init_size, &real_size);
QUIT;

%mend sim;

%macro runit(run_true_mu=,run_true_var=,run_epsilon=, run_pilot=,
run_guess_var=, run_method=, rep=);

    %do k = 1 %to &rep;

        %sim(true_mu=&run_true_mu,true_var=&run_true_var,epsilon=&run_epsilon,ind
ex=&k, pilot=&run_pilot, guess_var=&run_guess_var, method=&run_method);
    %end;

%mend runit;

%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=1.5, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=1.5, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=3, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=3, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=4, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='fixed',
run_guess_var=4, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=1.5, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=1.5, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=3, run_method='pool', rep=5000);

```

```

%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=3, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=4, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='fixed',
run_guess_var=4, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=1.5, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=1.5, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=3, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=3, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=4, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='fixed',
run_guess_var=4, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='pilot',
run_guess_var=0, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=1, run_pilot='pilot',
run_guess_var=0, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='pilot',
run_guess_var=0, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=0.5, run_pilot='pilot',
run_guess_var=0, run_method='ctrl', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='pilot',
run_guess_var=0, run_method='pool', rep=5000);
%runit(run_true_mu=0,run_true_var=3,run_epsilon=2, run_pilot='pilot',
run_guess_var=0, run_method='ctrl', rep=5000);

/*
proc univariate data=master.sample_size noprint;
    histogram init_size real_size;
run;
*/

proc sql;

    insert into master.sim_summary
    select 'ctrl', 'pilot', 3, 0, 1, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='pilot' AND epsilon=1;
    insert into master.sim_summary
    select 'ctrl', 'pilot', 3, 0, 0.5, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='pilot' AND epsilon=0.5;
    insert into master.sim_summary
    select 'ctrl', 'pilot', 3, 0, 2, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='pilot' AND epsilon=2;
    insert into master.sim_summary
    select 'pool', 'pilot', 3, 0, 1, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='pilot' AND epsilon=1;
    insert into master.sim_summary
    select 'pool', 'pilot', 3, 0, 0.5, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='pilot' AND epsilon=0.5;
    insert into master.sim_summary
    select 'pool', 'pilot', 3, 0, 2, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='pilot' AND epsilon=2;

```



```

insert into master.sim_summary
select 'ctrl', 'fixed', 3, 1.5, 1, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=1 AND
guess_variance=1.5;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 1.5, 0.5, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=0.5 AND
guess_variance=1.5;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 1.5, 2, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=2 AND
guess_variance=1.5;
insert into master.sim_summary
select 'pool', 'fixed', 3, 1.5, 1, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=1 AND
guess_variance=1.5;
insert into master.sim_summary
select 'pool', 'fixed', 3, 1.5, 0.5, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=0.5 AND
guess_variance=1.5;
insert into master.sim_summary
select 'pool', 'fixed', 3, 1.5, 2, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=2 AND
guess_variance=1.5;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 3, 1, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=1 AND
guess_variance=3;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 3, 0.5, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=0.5 AND
guess_variance=3;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 3, 2, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=2 AND
guess_variance=3;
insert into master.sim_summary
select 'pool', 'fixed', 3, 3, 1, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=1 AND
guess_variance=3;
insert into master.sim_summary
select 'pool', 'fixed', 3, 3, 0.5, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=0.5 AND
guess_variance=3;
insert into master.sim_summary
select 'pool', 'fixed', 3, 3, 2, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=2 AND
guess_variance=3;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 4, 1, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=1 AND
guess_variance=4;
insert into master.sim_summary
select 'ctrl', 'fixed', 3, 4, 0.5, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=0.5 AND
guess_variance=4;
insert into master.sim_summary

```

```

        select 'ctrl', 'fixed', 3, 4, 2, sum(reject)/count(*) from
master.sim_result where method='ctrl' AND type='fixed' AND epslon=2 AND
guess_variance=4;
        insert into master.sim_summary
        select 'pool', 'fixed', 3, 4, 1, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=1 AND
guess_variance=4;
        insert into master.sim_summary
        select 'pool', 'fixed', 3, 4, 0.5, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=0.5 AND
guess_variance=4;
        insert into master.sim_summary
        select 'pool', 'fixed', 3, 4, 2, sum(reject)/count(*) from
master.sim_result where method='pool' AND type='fixed' AND epslon=2 AND
guess_variance=4;

```

```
quit;
```

```

goptions reset=all gsfname=grafout gsfmode=replace device=jpeg;
/*goptions reset=all cback=white border htitle=12pt htext=10pt;*/

```

```

data try1(drop=guess_variance);
/*
length function color $8;
retain xsys ysys '2' when 'a';
*/
set TMP2.sim_summary;
row = _n_;
if guess_variance = 0.0 then char_variance="pilot";
else if guess_variance = 1.5 then char_variance='1.5';
else if guess_variance = 3.0 then char_variance='3.0';
else if guess_variance = 4.0 then char_variance='4.0';

if epslon = 0.5 then char_epslon="0.5";
else if epslon = 1.0 then char_epslon='1.0';
else if epslon = 2.0 then char_epslon='2.0';

```

```
run;
```

```

proc contents data=try1;
run;

```

```

proc print data=try1;
run;

```

```

proc gplot data=try1;
plot type1_error * row / haxis=axis1 vaxis=axis2
vref=0.044 0.050 0.056
name="summary";
symbol1 interpol=none value=dot color=red h=1.2;
axis1 label=("settings")
order=(0 to 25 by 1);
axis2 label=("Type 1 error rate")
order=(0.00 to 0.08 by 0.01);

```

```
run;
```

```

proc gplot data=try1;
  plot type1_error * char_variance = method/ haxis=axis1 vaxis=axis2
        vref=0.044 0.050 0.056 name="variance";
  symbol1 value=square color=red h=0.8;
  symbol2 value=triangle color=blue h=0.8;
  axis1 label=("expected variance") ;
  axis2 label=("Type 1 error rate")
        order=(0.00 to 0.08 by 0.01);
run;

```

```

proc gplot data=try1;
  plot type1_error * char_epsilon = method/ haxis=axis1 vaxis=axis2
        vref=0.044 0.050 0.056 name="epsilon";
  symbol1 value=square color=red h=0.8;
  symbol2 value=triangle color=blue h=0.8;
  axis1 label=("expected efficacy difference") ;
  axis2 label=("Type 1 error rate")
        order=(0.00 to 0.08 by 0.01);
run;
quit;

```

```

filename grafout clear;

```

```

proc sql;
create table sample_size_1 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='ctrl' and old.epsilon=1);

```

```

proc univariate data=sample_size_1;
histogram sample_size/ name='sample size 1';
run;

```

```

proc sql;
create table sample_size_2 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='ctrl' and old.epsilon=0.5);

```

```

proc univariate data=sample_size_2;
histogram sample_size/ name='sample size 2';
run;

```

```

proc sql;
create table sample_size_3 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='ctrl' and old.epsilon=2);

```

```

proc univariate data=sample_size_3;

```

```

histogram sample_size/ name='sample size 3';
run;

proc sql;
create table sample_size_4 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='pool' and old.epsilon=1);

proc univariate data=sample_size_4;
histogram sample_size/ name='sample size 4';
run;

proc sql;
create table sample_size_5 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='pool' and old.epsilon=0.5);

proc univariate data=sample_size_5;
histogram sample_size/ name='sample size 5';
run;

proc sql;
create table sample_size_6 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=0 and old.method='pool' and old.epsilon=2);

proc univariate data=sample_size_6;
histogram sample_size/ name='sample size 6';
run;

proc sql;
create table sample_size_7 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='ctrl' and old.epsilon=1);

proc univariate data=sample_size_7;
histogram sample_size/ name='sample size 7';
run;

proc sql;
create table sample_size_8 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='ctrl' and old.epsilon=0.5);

proc univariate data=sample_size_8;
histogram sample_size/ name='sample size 8';
run;

```

```

proc sql;
create table sample_size_9 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='ctrl' and old.epsilon=2);

proc univariate data=sample_size_9;
histogram sample_size/ name='sample size 9';
run;

proc sql;
create table sample_size_10 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='pool' and old.epsilon=1);

proc univariate data=sample_size_10;
histogram sample_size/ name='sample size 10';
run;

proc sql;
create table sample_size_11 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='pool' and old.epsilon=0.5);

proc univariate data=sample_size_11;
histogram sample_size/ name='sample size 11';
run;

proc sql;
create table sample_size_12 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=1.5 and old.method='pool' and old.epsilon=2);

proc univariate data=sample_size_12;
histogram sample_size/ name='sample size 12';
run;

proc sql;
create table sample_size_13 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='ctrl' and old.epsilon=1);

proc univariate data=sample_size_13;
histogram sample_size/ name='sample size 13';
run;

proc sql;

```

```
create table sample_size_14 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='ctrl' and old.epsilon=0.5);
```

```
proc univariate data=sample_size_14;
histogram sample_size/ name='sample size 14';
run;
```

```
proc sql;
create table sample_size_15 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='ctrl' and old.epsilon=2);
```

```
proc univariate data=sample_size_15;
histogram sample_size/ name='sample size 15';
run;
```

```
proc sql;
create table sample_size_16 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='pool' and old.epsilon=1);
```

```
proc univariate data=sample_size_16;
histogram sample_size/ name='sample size 16';
run;
```

```
proc sql;
create table sample_size_17 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='pool' and old.epsilon=0.5);
```

```
proc univariate data=sample_size_17;
histogram sample_size/ name='sample size 17';
run;
```

```
proc sql;
create table sample_size_18 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=3 and old.method='pool' and old.epsilon=2);
```

```
proc univariate data=sample_size_18;
histogram sample_size/ name='sample size 18';
run;
```

```
proc sql;
create table sample_size_19 as
(select init_size as init, real_size as sample_size
```

```

from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='ctrl' and old.epsilon=1);

proc univariate data=sample_size_19;
histogram sample_size/ name='sample size 19';
run;

proc sql;
create table sample_size_20 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='ctrl' and old.epsilon=0.5);

proc univariate data=sample_size_20;
histogram sample_size/ name='sample size 20';
run;

proc sql;
create table sample_size_21 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='ctrl' and old.epsilon=2);

proc univariate data=sample_size_21;
histogram sample_size/ name='sample size 21';
run;

proc sql;
create table sample_size_22 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='pool' and old.epsilon=1);

proc univariate data=sample_size_22;
histogram sample_size/ name='sample size 22';
run;

proc sql;
create table sample_size_23 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='pool' and old.epsilon=0.5);

proc univariate data=sample_size_23;
histogram sample_size/ name='sample size 23';
run;

proc sql;
create table sample_size_24 as
(select init_size as init, real_size as sample_size
from Tmp2.Sim_result as old
where old.guess_variance=4 and old.method='pool' and old.epsilon=2);

```

```
proc univariate data=sample_size_24;  
  histogram sample_size/ name='sample size 24';  
run;
```