Shiny Dashboard for Sample Size and Power Calculations

Matthew F. PartridgeF

4/25/2017

Table of Contents

[1. Introduction 1](#_Toc481010415)

[1.1 R and R Studio 2](#_Toc481010416)

[1.2 Packages 2](#_Toc481010417)

[2. Using the Application 2](#_Toc481010418)

[3. Application Tabs 3](#_Toc481010419)

[3.1 One Sample Mean 3](#_Toc481010420)

[3.2 One Sample Proportion 8](#_Toc481010421)

[3.3 Two Sample Means 12](#_Toc481010422)

[3.4 Two Sample Proportions 16](#_Toc481010423)

[3.5 Time to Event 20](#_Toc481010424)

[4. Further Improvements 25](#_Toc481010425)

[4.1 User Interface 25](#_Toc481010426)

[4.2 Server 26](#_Toc481010427)

[4.3 Alternative Hypothesis 27](#_Toc481010428)

[4.4 Time to Event 27](#_Toc481010429)

[4.5 New Pages 27](#_Toc481010430)

[5. References 27](#_Toc481010431)

# 1. Introduction

The purpose of this Shiny Dashboard is to create an updated version of the interactive Piface Java applet created by Russell V. Lenth. The ability to use interactive sliders and text boxes allows the user to see in real time how certain inputs, e.g. significance level or power, affect outputs, e.g. sample size or power. This will not only be a computational tool, but will also be educational to the user about co-dependencies of the inputs and outputs.

## 1.1 R and R Studio

This Shiny Dashboard is programmed using R version 3.4.0 and RStudio version 1.0.143. Four packages, besides the base packages included in R, are used for dashboard setup and statistical calculations.

## 1.2 Packages

### 1.2.1 [shiny](https://cran.r-project.org/web/packages/shiny/shiny.pdf)

The shiny package in R "Makes it incredibly easy to build interactive web applications with R."(RStudio, Shiny) It uses R functions to create HTML code for a web page. There are various different input, display, and settings options that can be customized all within the same shiny application.

### 1.2.2 [shinydashboard](https://cran.r-project.org/web/packages/shinydashboard/shinydashboard.pdf)

The shinydashboard package expands beyond the shiny package to add themes to the dashboard for a more attractive look.

### 1.2.3 [pwr](https://cran.rstudio.com/web/packages/pwr/pwr.pdf)

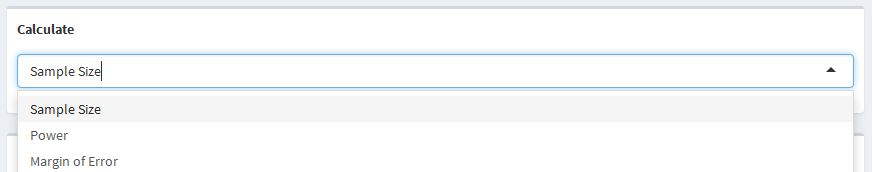
The pwr package calculates power and sample sizes for various different situations using the calculations of Cohen (Cohen, 1980) as a basis for the calculations. In this dashboard, the pwr package is used for the calculations of the One Sample Mean, One Sample Proportion, Two Sample Means, and Two Sample Proportions scenarios.

### 1.2.4 [gsDesign](https://cran.r-project.org/web/packages/gsDesign/gsDesign.pdf)

The gsDesign package is used for power and sample size calculations of time to event scenarios. Specifically, it incorporates factors targeting the time, recruitment, and censoring components that are more likely to occur in this scenario..

# 2. Using the Application

The first thing to choose is which scenario to use. After the scenario selection, what is being calculated is selected. All of the pages have a selection box located at the top of the dashboard.



The selectizeInput is used by all of the scenarios in order choose what is to be calculated. The example above is specifcally used for the one sample mean and one sample proportion scenarios while the other scenarios do not have the margin of error selection. The rest of the dashboard page will then update based on what is being calculated.

The setup of each of the pages is very similar with only the specific inputs being slightly different. The one sample mean, one sample proportion, two sample means, and two sample proportions pages are all set up with study information on the left-hand panel and calculation information as well as the calculation itself on the right-hand panel. The time to event page is slightly different in that the study information in on the left-hand panel, sample information is on the middle panel, and the calculation information and the calculation itself on the right-hand panel. For all of the pages, the calculation is outlined in green as seen below.



Within each of the pages, there are text, slider, and selector inputs. Each numeric variable has both a text box and a slider that can be manipulated. Updating either of them will have the same effect and will automatically be updated once the other is manipulated. The calculations are performed automatically after an input is manipulated. It should be noted that, although the calculated variable text box and slider can be manipulated, it will not cause any new calculations to be performed.

# 3. Application Tabs

## 3.1 One Sample Mean

### 3.1.1 Parameters

* Mean - The mean of the target sample
* Reference Mean - The mean of the reference population
* Standard Deviation - The standard deviation of the measure for the target sample
* Sample Size - The size of the target sample
* Significance Level - The probability of falsely rejecting a null hypothesis
* Power - The probability of correctly rejecting a null hypothesis
* Margin of Error - The distance in one direction of the confidence interval from the mean
* Confidence Interval - The interval in which the mean is likely to fall within

### 3.1.2 Use of Parameters

Sample size, power, and margin of error are all parameters that can be calculated, while sample size and power are both parameters that can be used as inputs for other calculations. All of the other parameters can be changed and manipulated resulting in different values for sample size, power, or margin of error.

### 3.1.3 Statistical Explanation

The one sample mean scenario looks to compare a sample mean and standard deviation to a reference population mean in a two-sided comparison. The null and alternative hypothesese are and respectively, where is the sample mean and is the reference population mean that the sample mean is being compared to. In order to compare the two means, the Student's -Test can be used:

where, and are defined as above, is the sample standard deviation, is the sample size, and is the t-value from the -distribution with degrees of fredom. The importance of the -test is that it standardizes the raw effect size, by dividing by the standard deviation of the population. Cohen defined this standardization, , as the Effect Size Index (ESI) (Cohen, 1988). Due to the fact that the sample size, , is unknown for the scenario of solving for the sample size, the -distribution can be estimated using the -distribution since as n gets larger. Thus, the -distribution will be used throughout this scenario.

**Power:** The power of a study is defined as the probability of correctly rejecting the null hypothesis, , when the alternative hypothesis, , is true. This can be expressed as the probability that the -score, , is more extreme than the standard normal value for a given significance level.

Where is the target mean, is the reference mean, is the true mean of the target population, denotes the Standard Normal quantile function, is the significance level, is the standard deviation of the target sample, is sample size of the target sample, and denotes the Standard Normal distribution function.

**Sample Size:** The sample size can be defined as the number of observations needed in order for the equation above to equate to a given power. With this idea in mind, the equation can be manipulated to solve for the samle size instead of the power. The sample size can be calculated by solving the following equation for a given power:

**Margin of Error:** The margin of error of an estimate is typically defined by the width of its confidence interval. A smaller margin of error implies more precision and vice versa. The margin of error is simply and the confidence interval for a value, is . Thus the confidence interval for a value, is

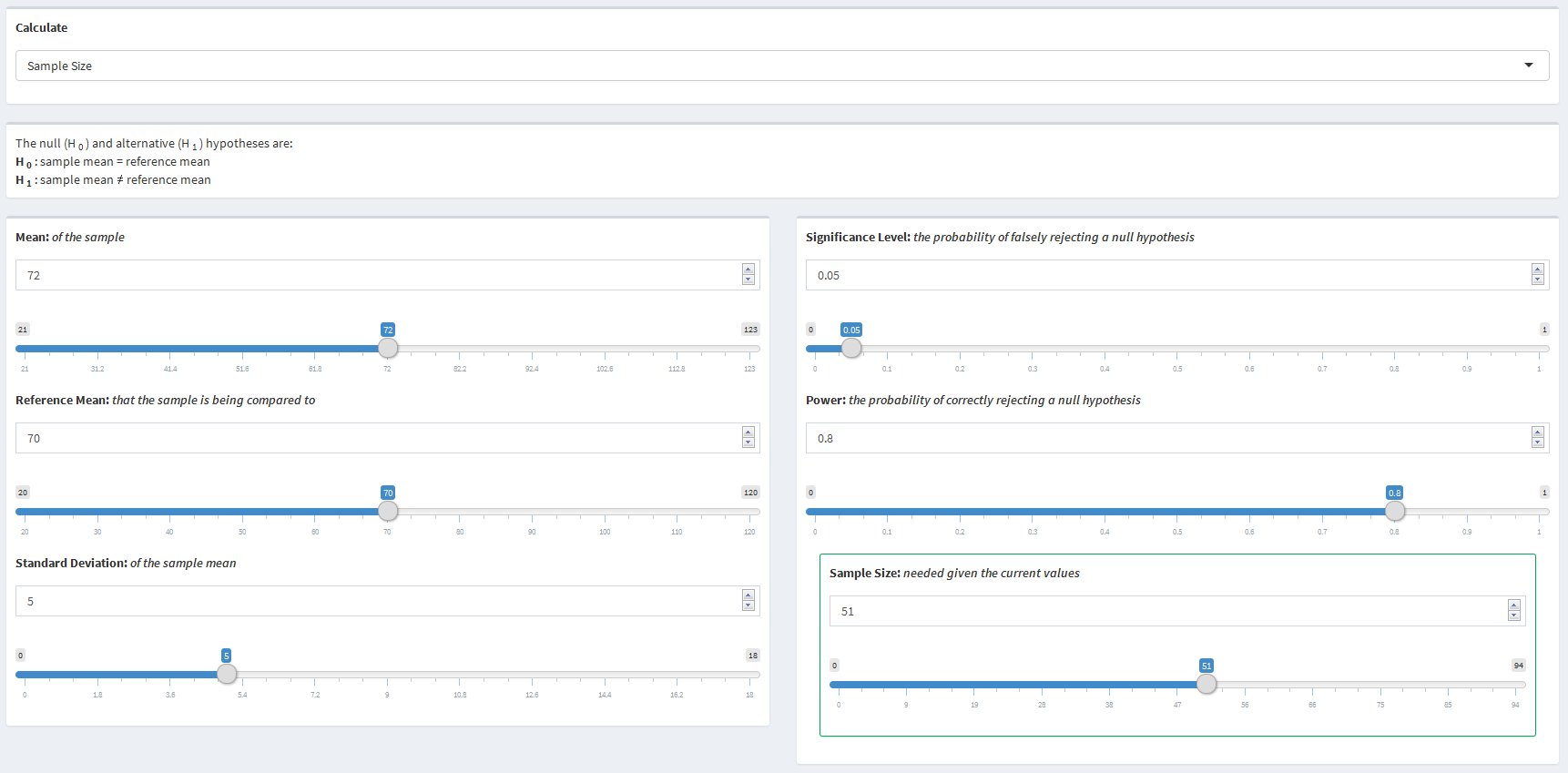
### 3.1.4 Example

**Calculating Sample Size:** A study is being performed to examine the heart rates of people after walking for 5 minutes. Heart rate is a measurement that can be summarized using the mean. Say it is known that the average resting heart rate is 70 beats per minute. The goal of the study is to determine if the heart rate of participants walking for 5 minutes is different than an average resting heart rate. The researches performed a pilot study and have data about the mean and standard deviation of the heart rate, say 72 and 5 respectively. They decide that they want their study to be based on a significance level of 0.05 with 80% power. The researches want to know how large of a sample they will need to determine participants walking for 5 minutes have a different heart rate than the resting heart rate of an average person.

The researches would use the following values to determine what they need:

* Mean = 72
* Reference Mean = 70
* Standard Deviation = 5
* Significance Level = 0.05
* Power = 0.8

With the values put in by the researchers, they would find that a sample of 51 participants would be required.

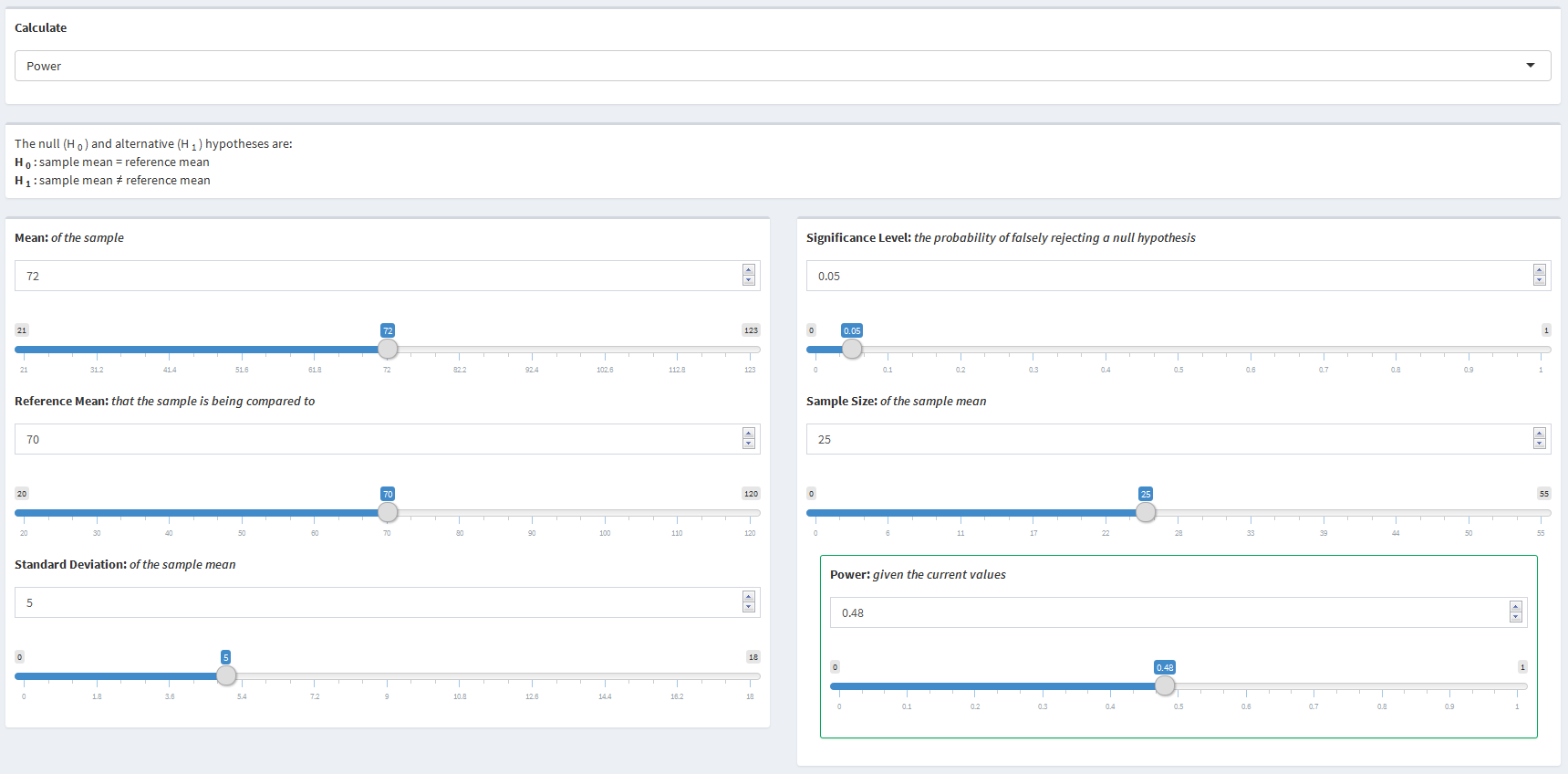


**Calculating Power:** Instead of determining how large of a sample the researches would need before they performed the study, they decided to collect the data now and later determine how much power they have. They were able to record heart rates after walking 5 minutes for 25 participants and found a mean of 72 and a standard deviation of 5. They still require a significance level of 0.05. The researchers want to know how much power their study has.

The researchers would use the following values to determine what they need:

* Mean = 72
* Reference Mean = 70
* Standard Deviation = 5
* Significance Level = 0.05
* Sample Size = 25

With the values put in by the researches, they would find that their study has 48% power.

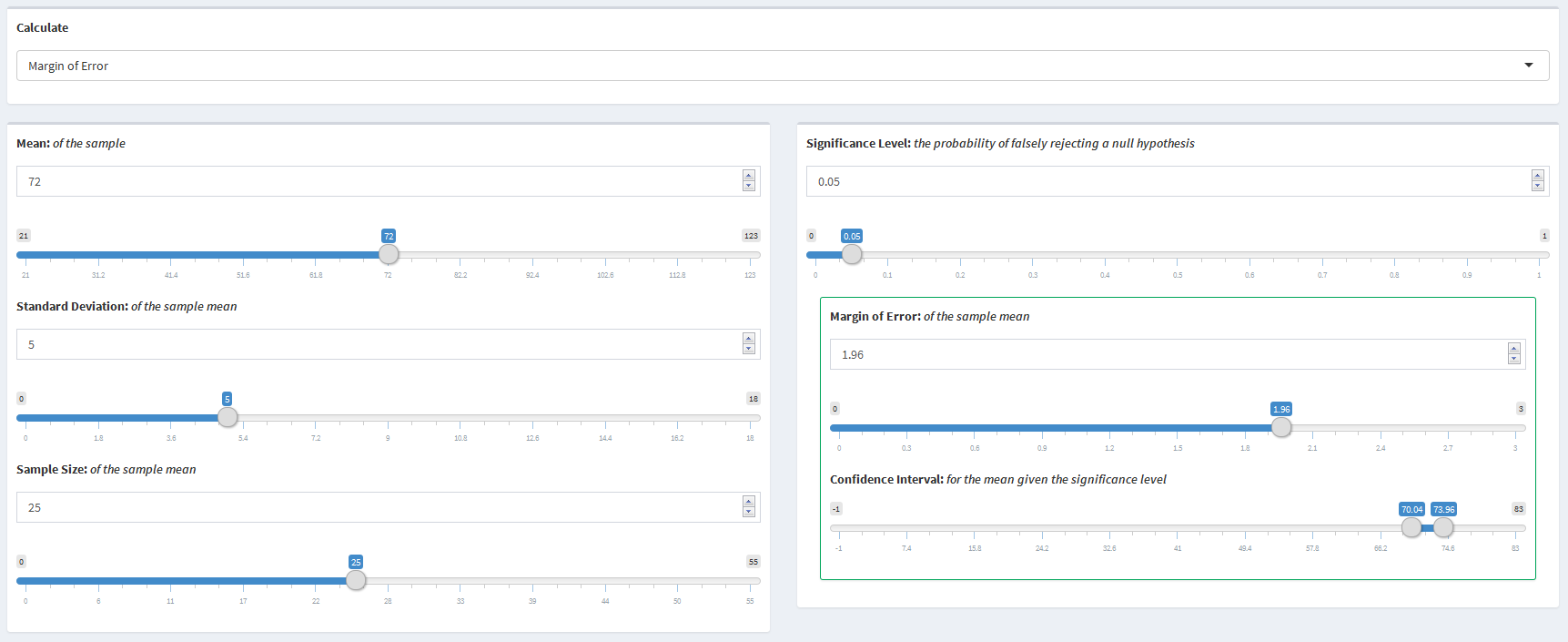


**Calculating Margin of Error:** The researchers are now trying to report their findings from the study and would like to present a 95% confidence interval for the mean heart rate after walking for 5 minutes. The researches want to know the margin of error of their measure.

The researchers would use the following values to determine what they need:

* Mean = 72
* Standard Deviation = 5
* Significance Level = 0.05
* Sample Size = 25

With the values put in by the researches, they would find that their measure has a 1.96 margin of error resulting in a 95% confidence interval of 70.04 to 73.96.



## 3.2 One Sample Proportion

### 3.2.1 Parameters

* Proportion - The proportion affected in the target sample
* Reference Proportion - The proportion affected in the reference population
* Sample Size - The size of the target sample
* Significance Level - The probability of falsely rejecting a null hypothesis
* Power - The probability of correctly rejecting a null hypothesis
* Margin of Error - The distance in one direction of the confidence interval from the proportion
* Confidence Interval - The interval in which the proportion is likely to fall within

### 3.2.2 Use of Parameters

Sample size, power, and margin of error are all parameters that can be calculated, while sample size and power are both parameters that can be used as inputs for other calculations. All of the other parameters can be changed and manipulated resulting in different values for sample size, power, or margin of error.

### 3.2.3 Statistical Explanation

The one sample proportion scenario looks to compare a sample proportion to a reference population proportion in a two-sided comparison. The null and alternative hypothesese are and respectively, where is the target proportion affected, is the reference proportion affected that the target proportion is being compared to. In order to compare the two proportions, the Student's -Test can be used:

where, and are defined as above, is the true proportion of the target population affected, is the sample size, and is the t-value from the -distribution with degrees of fredom. Due to the fact that the sample size, , is unknown for the scenario of solving for the sample size, the -distribution can be estimated using the -distribution since as n gets larger. Thus, the -distribution will be used throughout this scenario.

**Power:** The power of a study is defined as the probability of correctly rejecting the null hypothesis, , when the alternative hypothesis, , is true. This can be expressed as the probability that the -score, , is more extreme than the standard normal value for a given significance level.

Where is the target proportion affected, is the reference proportion affected, is the true proportion of the target population affected, denotes the Standard Normal quantile function, is the significance level, is size of the target sample, and denotes the Standard Normal distribution function.

**Sample Size:** The sample size can be defined as the number of observations needed in order for the equation above to equate to a given power. With this idea in mind, the equation can be manipulated to solve for the samle size instead of the power. The sample size can be calculated by solving the following equation for a given power:

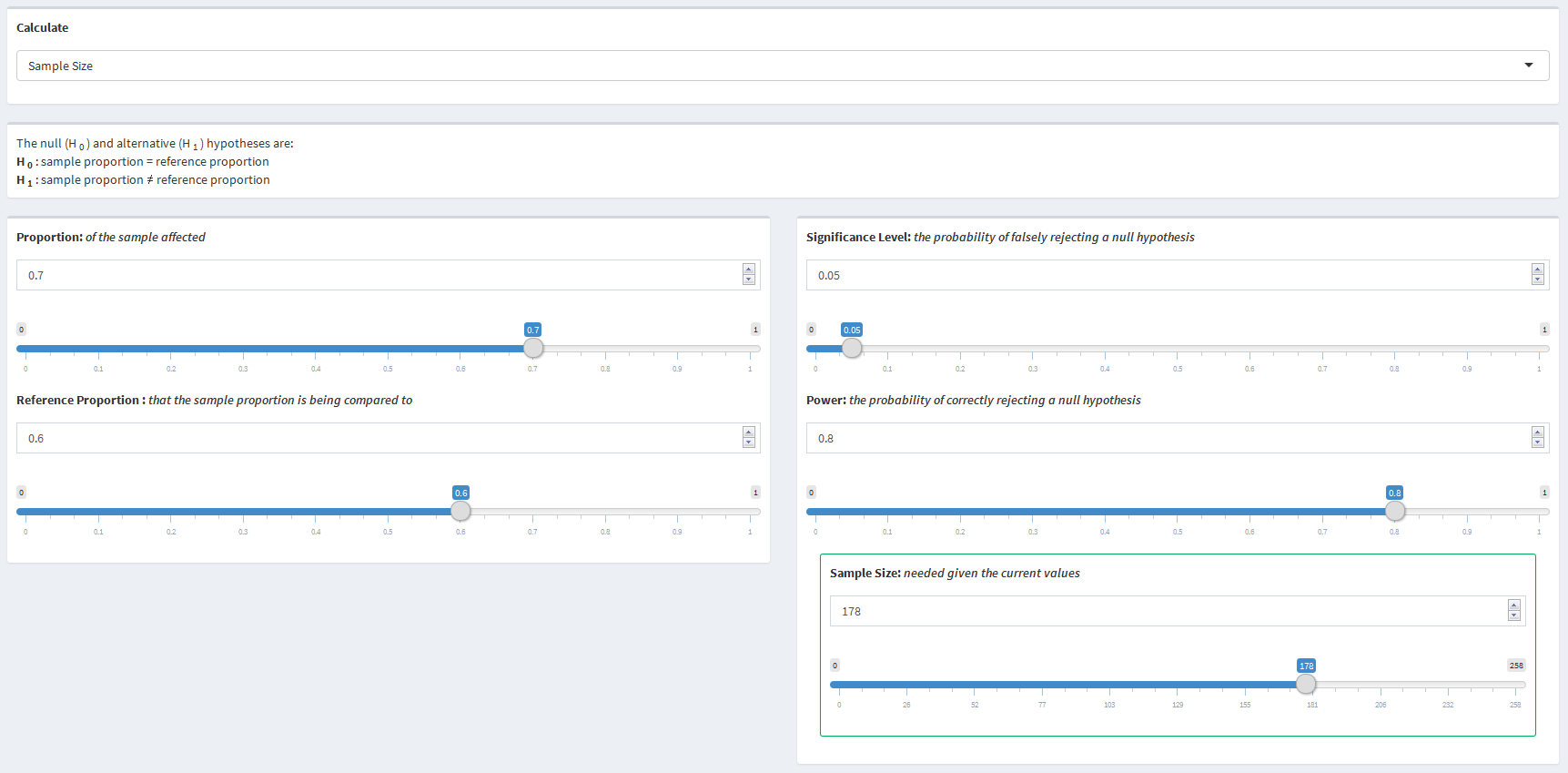
### 3.2.4 Example

**Calculating Sample Size:** A study is being performed to examine the effect of a new drug on treating headaches. Previous research has shown that the drug that is considered industry standard eases headaches in 60% of people. The goal of this study is to compare this new drug's effect on easing headaches to the industry standard. The researches have reason to believe that their new drug helps to ease headaches in 70% of people. They decide that they want their study to be based on a significance level of 0.05 with 80% power. The researches want to know how large of a sample they will need to determine the proportion affected by their new drug is different than the proportion of participants affected by the industry standard.

The researches would use the following values to determine what they need:

* Proportion: 0.7
* Reference Proportion: 0.6
* Significance Level = 0.05
* Power = 0.8

With the values put in by the researchers, they would find that a sample of 178 participants would be required.

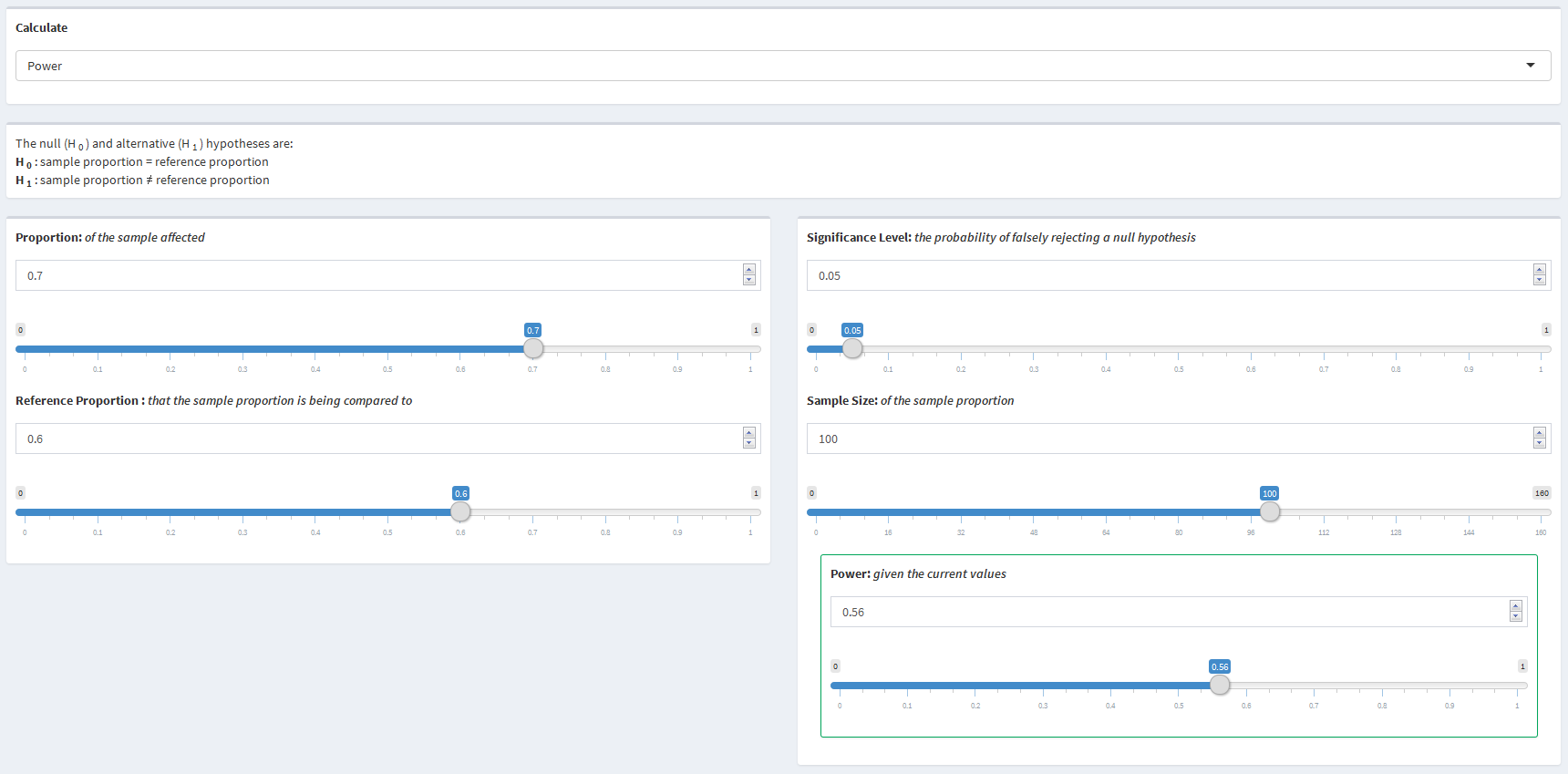


**Calculating Power:** Instead of determining how large of a sample the researches would need before they performed the study, they decided to collect the data now and determine how much power they would have later. They were able to collect data for 100 participants and found that 70% of them had their headaches eased. They still require a significance level of 0.05. The researches want to know how much power their study has.

The researchers would use the following values to determine what they need:

* Proportion = 0.7
* Reference Proportion = 0.6
* Significance Level = 0.05
* Sample Size = 100

With the values put in by the researches, they would find that their study has 56% power.

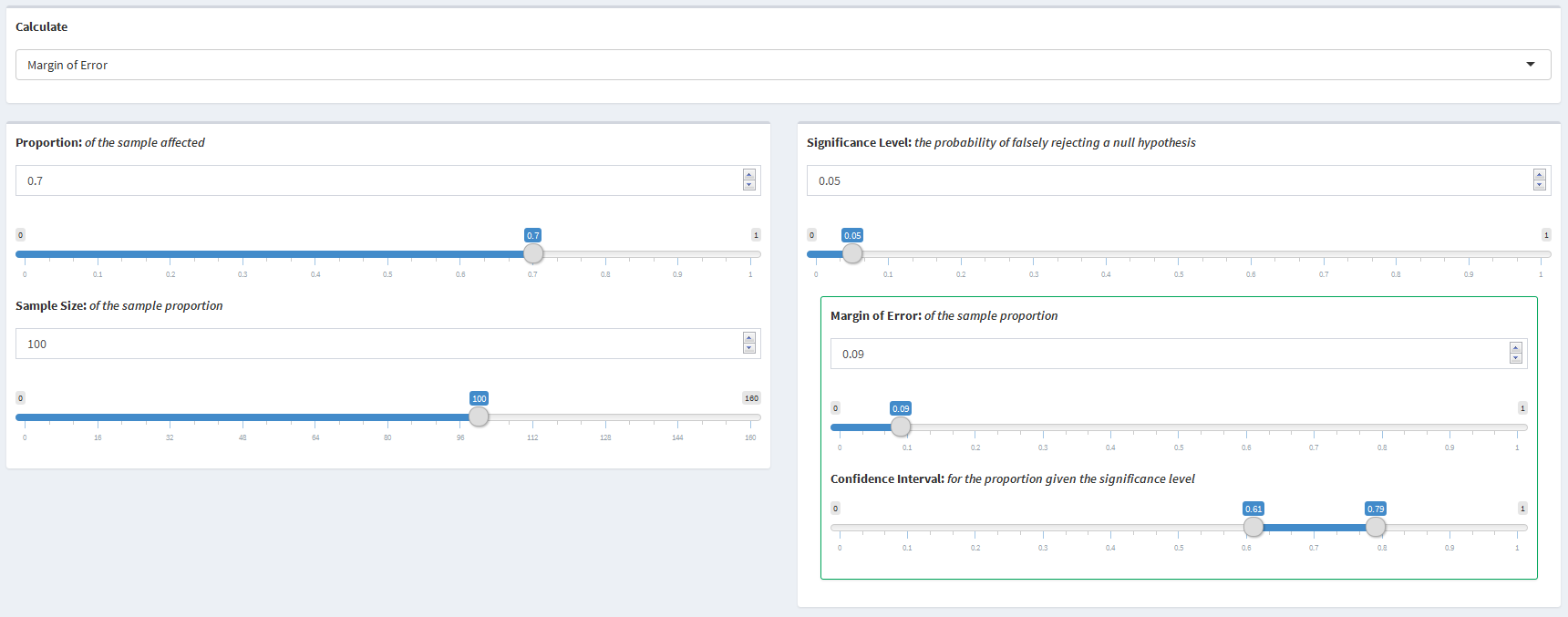


**Calculating Margin of Error:** The researchers are now trying to report their findings from the study and would like to present a 95% confidence interval for the proportion of participants affected by their new drug. The researches want to know the margin of error of their measure.

The researchers would use the following values to determine what they need:

* Proportion = 0.7
* Significance Level = 0.05
* Sample Size = 100

With the values put in by the researches, they would find that their measure has a 1.96 margin of error resulting in a 95% confidence interval of 70.04 to 73.96.



## 3.3 Two Sample Means

### 3.3.1 Parameters

* Mean One - The mean of the first sample
* Sample Size One - The size of the first sample
* Mean Two - The mean of the second sample
* Sample Size Two - The size of the second sample
* Standard Deviation - The standard deviation of the measure for the target sample
* Significance Level - The probability of falsely rejecting a null hypothesis
* Power - The probability of correctly rejecting a null hypothesis

### 3.3.2 Use of Parameters

Sample size and power are both parameters that can be calculated or used as input parameters for calculations. All of the other parameters can be changed and manipulated resulting in different values for sample size or power.

### 3.3.3 Statistical Explanation

The two sample mean scenario looks to compare a mean from one sample to an expected mean from another sample that are both part of a larger population in a two-sided comparison. It is assumed in this scenario that the standard deviation of both sample means are equal. The null and alternative hypotheses are and respectively, where is the mean from sample one, is the expected mean from sample two. In order to compare the two means, the Student's -Test can be used:

where is the sample mean from sample one, is the mean from sample two, and is the sample standard deviation. Since there are two samples in this scenario, represents the harmonic mean of the two sample size as defined by Cohen (1988) (), and is the -value from the -distribution with degrees of fredom. The importance of the -test is that it standardizes the raw effect size between the means by dividing by the standard deviation of the population. Cohen (1988) defined this standardization, , as the Effect Size Index (ESI). Due to the fact that the sample size, , is unknown for this situation, the -distribution will be estimated using the -distribution as as gets larger.

**Power:** The power of a study is the probability of correctly rejecting the null hypothesis, in this case, that . Given that the -distribution is being used to estimate the -distribution, this can be expressed as the probability that the -score, , is more extreme than the standard normal value for a given significance level.

Where is the mean for sample one, is the mean of sample two, is the true mean of population one, denotes the Standard Normal quantile function, is the significance level, is the standard deviation of the target sample, is the harmonic mean of the sample sizes, and denotes the Standard Normal distribution function.

**Sample Size:** The sample size can be defined as the number of observations needed in order for the equation above to equate to a given power. With this idea in mind, the equation can be manipulated to solve for the samle size instead of the power. The sample size can be calculated by solving the following equation for a given power:

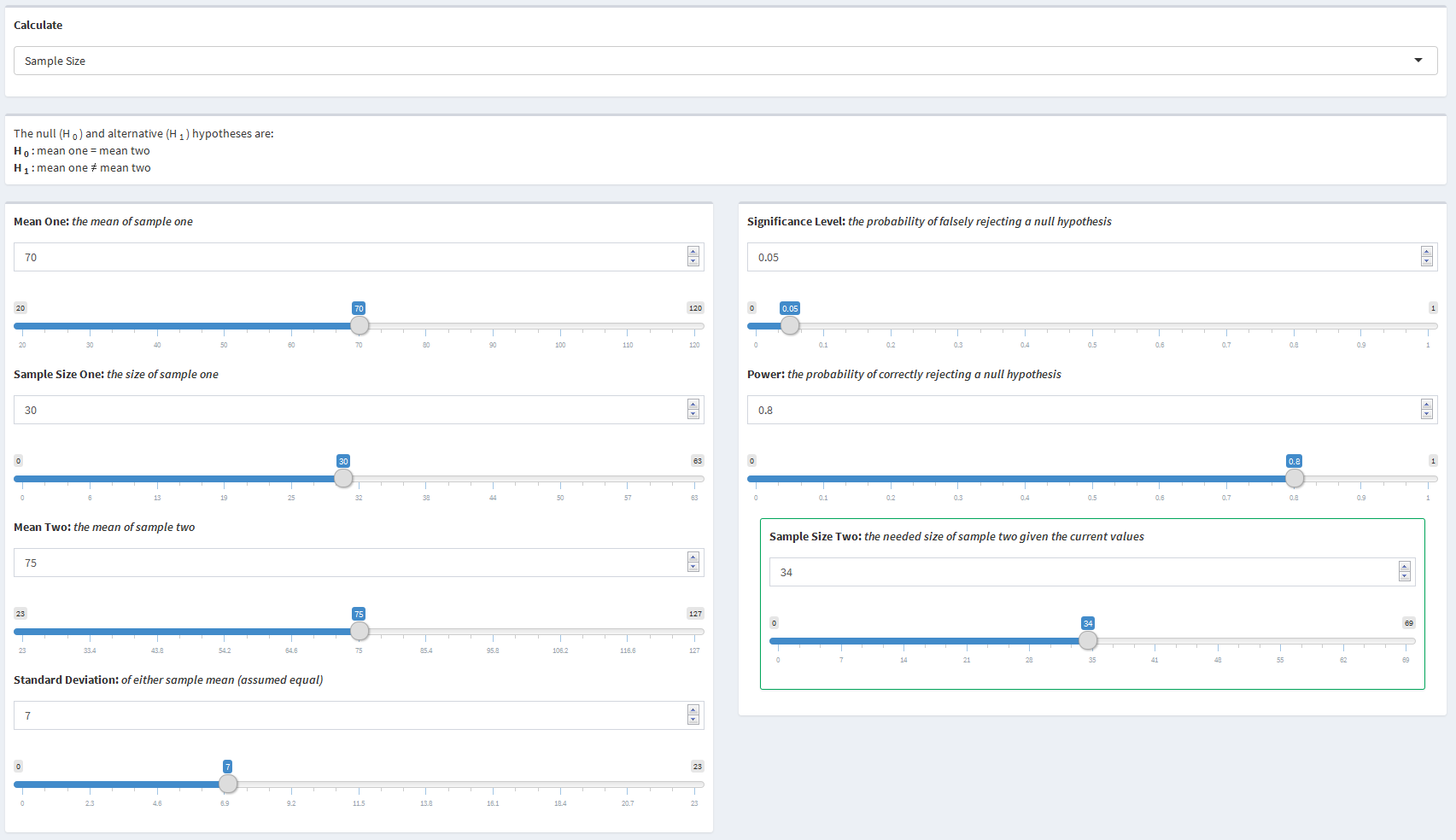
### 3.3.4 Example

**Calculating Sample Size:** A study is being performed to examine the resting hear rates of men and women. The researchers were able to recruit 30 men to particpate in the study and now need to determine how many women are needed in order to have sufficient power. They decide that they want their study to be based on a significance level of 0.05 with 80% power. They performed a pilot study that found the average male resting heart rate was 70 while the average female heart rate was 75. The standard deviation for the resting heart rate was found to be 7. The researches want to know how large of a sample of women is needed to be able to determine if there is a difference between the two genders.

The researches would use the following values to determine what they need:

* Mean One = 70
* Sample Size One = 30
* Mean Two = 75
* Standard Deviation = 7
* Significance Level = 0.05
* Power = 0.8

With the values put in by the researchers, they would find that a sample of 34 women would be required.

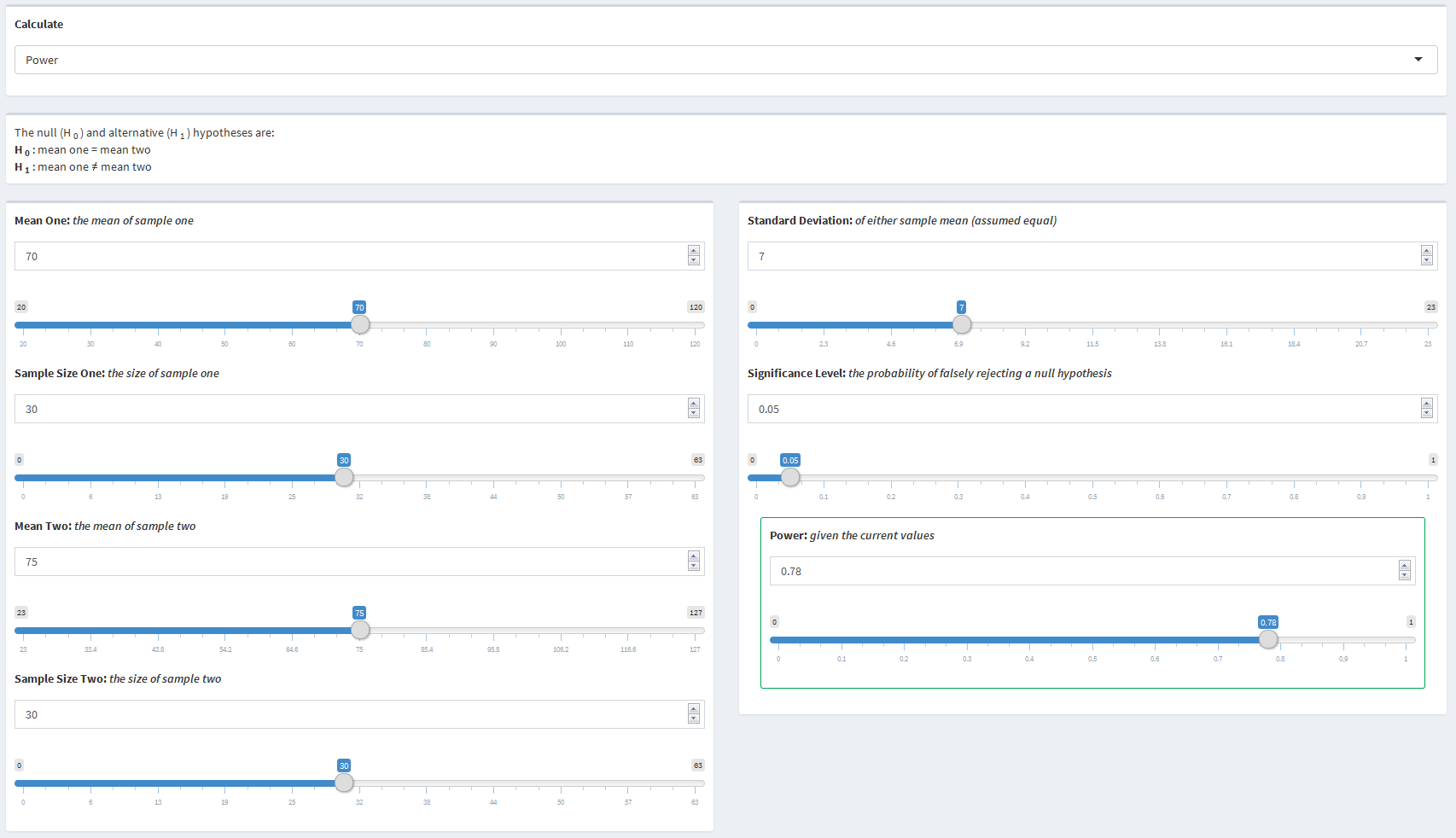


**Calculating Power:** The researchers decided, for ease, to recruit just as many women as they had men. Thus, the study obtained resting heart rate data for 30 men and 30 women. The resting heart rate measurement had a standard deviation of 7. They still require a significance level of 0.05. The researches want to know how much power their study has.

The researchers would use the following values to determine what they need:

* Mean One = 70
* Sample Size One = 30
* Mean Two = 75
* Sample Size Two = 30
* Standard Deviation = 7
* Significance Level = 0.05

With the values put in by the researches, they would find that their study has 78% power.



## 3.4 Two Sample Proportions

### 3.4.1 Parameters

* Proportion One - The proportion affected in sample one
* Sample Size One - The size of sample one
* Proportion Two - The proportion affected in sample two
* Sample Size Two - The size of sample two
* Significance Level - The probability of falsely rejecting a null hypothesis
* Power - The probability of correctly rejecting a null hypothesis

### 3.4.2 Use of Parameters

Sample size and power are both parameters that can be calculated or used as input parameters for calculations. All of the other parameters can be changed and manipulated resulting in different values for sample size or power.

### 3.4.3 Statistical Explanation

The two sample proportion scenario looks to compare the proportion affected from a target sample to the proportion affected from a reference sample. The null and alternative hypotheses are and respectively, where is the proportion affected from the target sample and is the proportion affected from the reference sample. A -test can be used to compare the two proportions:

where is the proportion affected from group , is the sample size from group , and is the -score from the -distribution.

**Power:** The power of a study is the probability of correctly rejecting the null hypothesis, . This can be expressed as the probability that the -score, , is more extreme than the standard normal value for a given significance level.

where is the proportion affected from the target sample, is the proportion affected from the reference sample, is the true proportion affected from the target sample, is the sample size from the target sample, is the sample size from the reference sample, is the significance level, denotes the Standard Normal quantile function, and denotes the Standard Normal distribution function.

**Sample Size:** The sample size can be defined as the number of observations needed in order for the equation above to equate to a given power. Since the the total sample size can be the sum of many different two sample combinations, one of the sample sizes must be selected in order to solved for the other sample size.The sample size can be calculated by solving the following equation for one of the sample sizes for a given power:

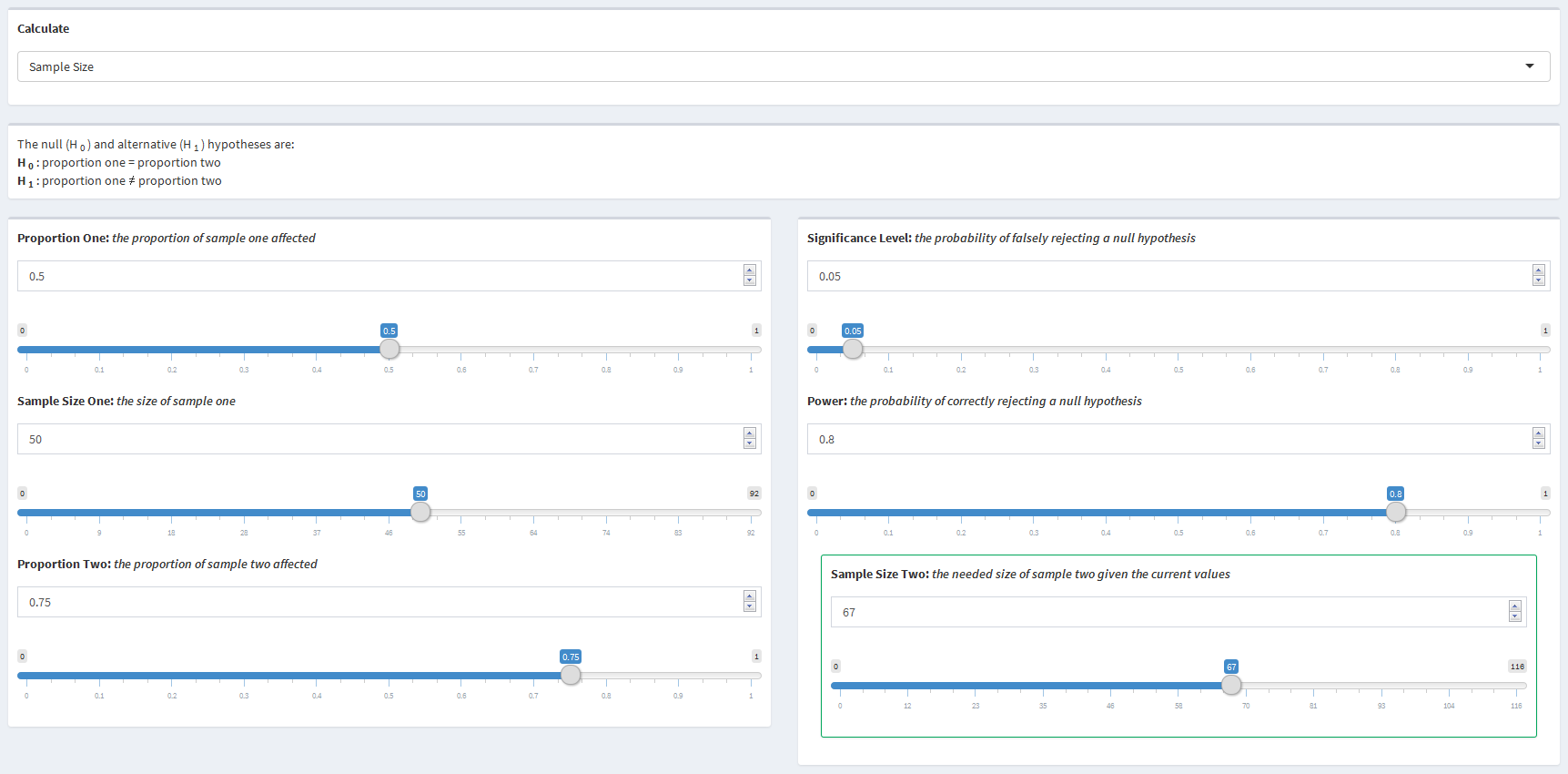
### 3.4.4 Example

**Calculating Sample Size:** A study is being performed to examine the effect of easing headaches between a new drug and the industry standard drug. The researchers have reason to believe that their new drug helps to ease headaches in 75% of people while the industry standard eases headaches in 50% of people. The researchers have recruited 100 participants to take the industry standard drug and are wondering how many participants they will need to recruit to take their new drug. They decide that they want their study to be based on a significance level of 0.05 with 80% power. The researchers want to know how large of a sample they will need to determine the proportion affected by their new drug is different than the proportion of participants affected by the industry standard.

The researches would use the following values to determine what they need:

* Proportion One: 0.5
* Sample Size One: 50
* Proportion Two: 0.75
* Significance Level = 0.05
* Power = 0.8

With the values put in by the researchers, they would find that a sample of 67 participants would be required.

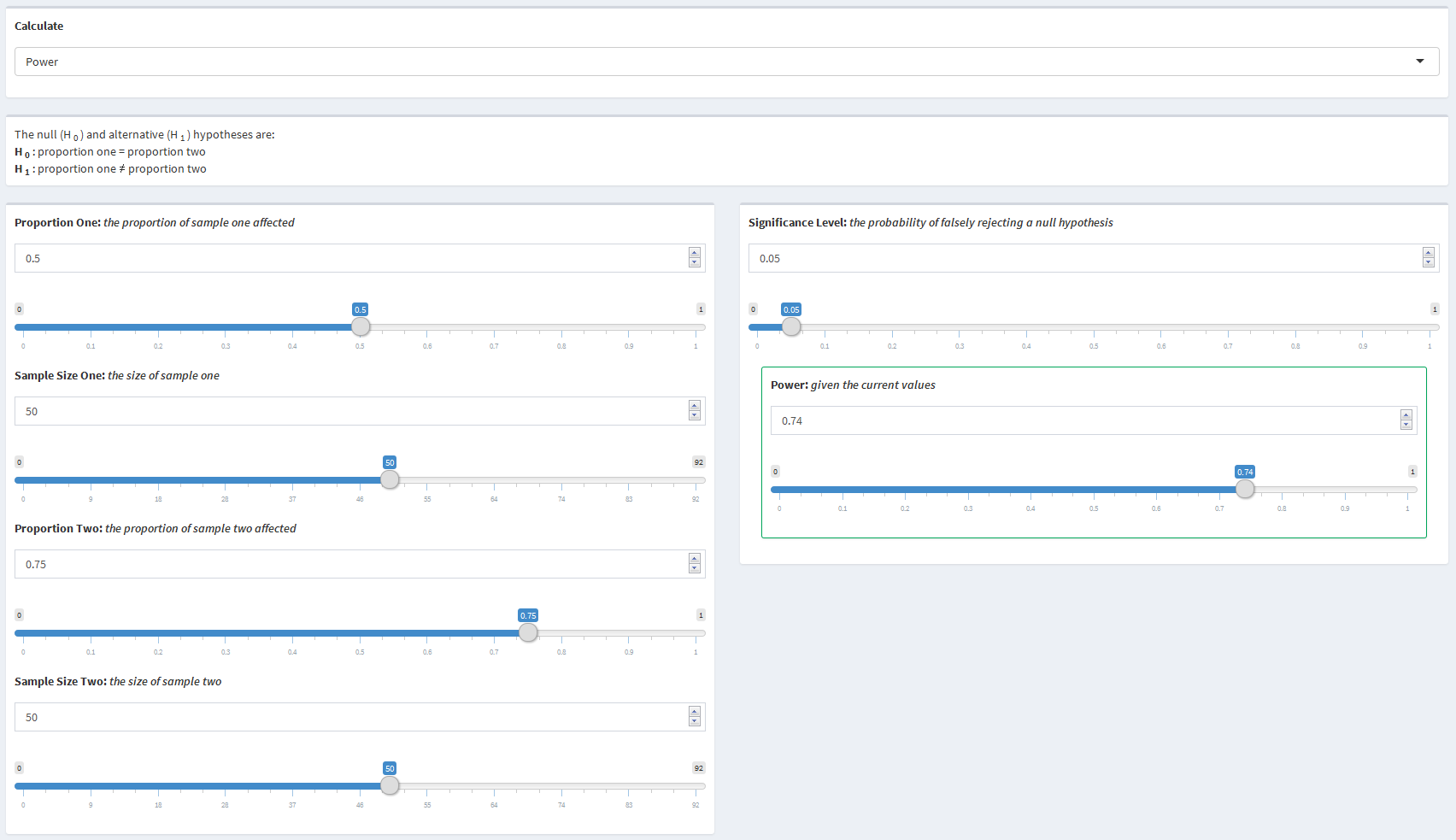


**Calculating Power:** The research decided, for ease, to recruit the same number of participants to take their drug as those who take the industry standard drug. This means they collected data on 50 participants taking each drug. They found that 50% of participants had their headache ease with the industry standard drug while 75% of participants had their headache ease with their new drug. They still require a significance level of 0.05. The researches want to know how much power their study has.

The researches would use the following values to determine what they need:

* Proportion One: 0.5
* Sample Size One: 50
* Proportion Two: 0.75
* Sample Size Two: 50
* Significance Level = 0.05

With the values put in by the researches, they would find that their study has 74% power.



## 3.5 Time to Event

### 3.5.1 Parameters

* Enrollment Schedule - The time at which participants are enrolled in the study
* Distribution of Enrollment - The distribution of how participants are enrolled
* Exponential Rate - The rate of growth or decay under the exponential enrollment assumption
* Study Duration - The duration of the study in units of time
* Enrollment Duration - The duration of enrollment in units of time
* Sample Allocation Ratio - The ratio of target sample to the reference sample
* Target Event Rate - The number of events per unit of time for the target sample
* Reference Event Rate - The number of events per unit of time for the reference sample
* Target Censoring Rate - The rate at which events will not be observed in the target sample
* Reference Censoring Rate - The rate at which events will not be observed in the reference sample
* Total Sample Size - The total sample size of both the target and reference samples added together
* Significance Level - The probability of falsely rejecting a null hypothesis
* Power - The probability of correctly rejecting a null hypothesis

### 3.5.2 Use of Parameters

Sample size and power are both parameters that can be calculated or used as input parameters for calculations. All of the other parameters can be changed and manipulated resulting in different values for sample size or power. Distribution of Enrollment becomes visible if either "All at Once" or "Throughout" are selected for Enrollment Schedule. Exponential Rate becomes visible if Distribution of Enrollment is visible and if "Exponential" has been selected. Enrollment Duration disappears when " Throughout" is selected for Distribution of Enrollment. The censor rate inputs should be set to 0 if the observed event rate is used for the event rate inputs. When a censor rate is greater than 0, it will change the inputted event rate to reflect what rate of events will actually be observed based on the censor rate.

### 3.5.3 Statistical Explanation

**Sample Size**: The time to event scenario looks to compare the event rates of two different samples in a two-sided comparison. Specifically, the hazard rates, and , are being compared by examining the risk ratio, of the two samples. Lachin (1981) derived a basic equation relating the sample size with the power when comparing the risk ratio:

.

where , , is the expected number of events given the event rate, and is defined by the distribution of entry into the study. It is often the case that studies will have recruitment periods that occur during the time of the study. One basic assumption is that entry into the study during the recruitment period will occur uniformly throughout that time. This can be reflected by letting , where is the event rate, T is the duration of the study, and R is the duration of recruitment. It may also be the case that entry into the study doesn't occur uniformly and instead occurs at an exponential rate, either convex or concave, during that time. In this case, let , where the exponential parameter. Entry will be convex, more entries earlier with less later on, if and entry will be concave, increasing entries as time goes on, if . Another common assumption made in these types of studies is that all of the possible events that could occur will be observed. That is, there is perfect information of all of the participants in the study. It is very possible that some events may not be able to be recorded or participants may leave the study before the official end of the trial among other possible reasons. These types of scenarios result in censoring of the data as the full information from that participant in the study is not known. Each group can have its own censoring rate, which represents the probability of being censored within each group. Then, the observed event rate () can be estimated by deflating the expected event rate () using the censoring rate () where: . The observed event rate can then be substituted into the sample size equation for each instance of its corresponding expected event rate.

The final sample size equation is then:

where and can be replaced with observed event rates if censoring is present and can changed depending on they type of entry into the study that is expected as outlined above.

**Power**: Calculations for the power of the study are slightly more involved than the calculations for sample size due to the way the R package being used is built (gsDesign Citation). Although calculations for sample size and power can be performed in terms of both the event rates or the number of expected events, the gsDesign package can only calculate power in terms of the number of observed events. Thus, conversions from the event rates to the number of expected events are performed and then the power is calculated. The first step is to calculate the required sample size in terms of the event rates, as shown above in the sample size calculations using any power

where all of the same parameter specifications that were outlined above can also be made. The expected number of events can then be calculated using this sample size estimation

where N is the estimated sample size from above, D represents the event, and is defined as found in the sample size calculations above. It should be noted that by using the sample size estimate, the expected number of events is dependent on the parameter and thus, the power. Dividing the expected number of events calculated above by the sample size estimate from above will result in the number of events per person, . The events per person estimate no longer uses the sample size estimate and is thus free from dependancy on the parameter. This is why any power could be used for the initial sample size estimation. In *The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions (1981)*, Schoenfeld derived a test statistics for comparing survival distributions in terms of the expected number of events. The statistic is asymptotically normal with unit variance. He also showed that using the assumpitons of the log-rank test, which is commonly used for comparing survival curves, the mean for the test statistic reduces to , where is the event rate for the group, N is the sample size, , and d is the combined probability of an event occurring. This can be further simplified by multiplying N and d to get the expected number of events, . The power can then be calculated in terms of the expected number of events following the equation outlined by Schoenfeld:

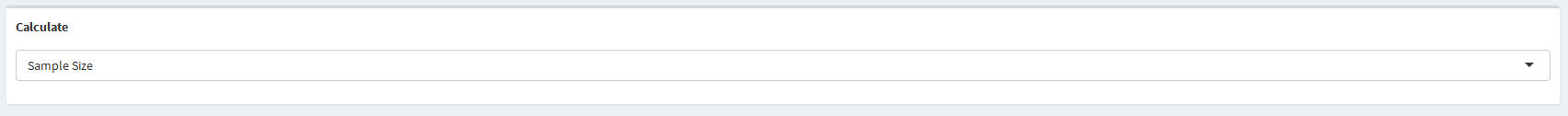
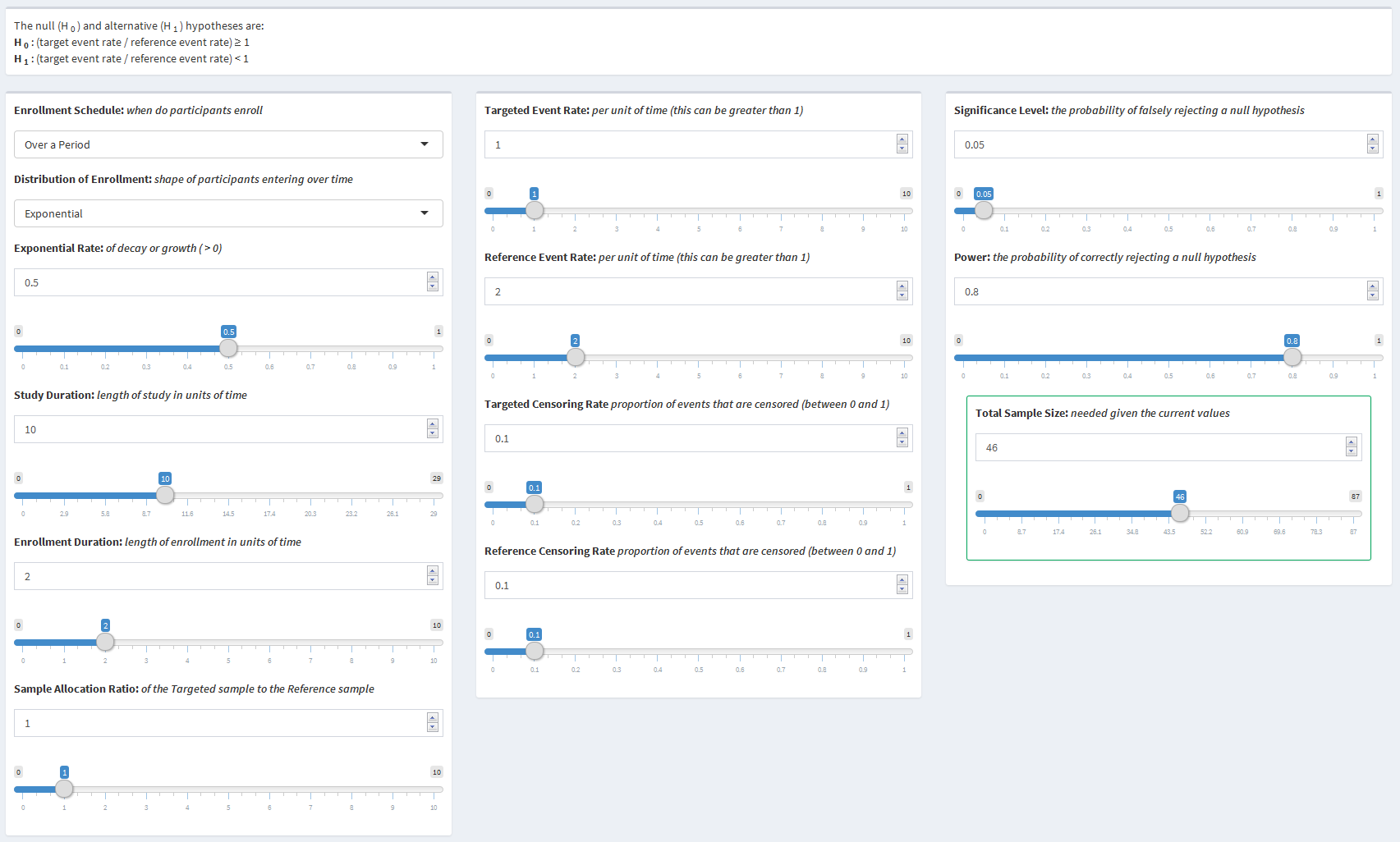
### 3.5.4 Example

**Calculating Sample Size:** A study is being performed to examine the effectiveness of a new drug for preventing breast cancer reccurrence. The goal of the study is to compare the inclusion of this new drug with general lifestyle guidlines to just following the lifestyle guidlines. The study is set to be 10 years long in order to allow for the possiblity of relapse with participants being recruited within the first 2 years of the study. It is expected that recruitment will start out relatively quickly, but then slow gradually. The researches will recruit the same number of participants for each of the two groups. Using general lifestyle guidlines, the researchers expect to see 2 reccurrences per yearhowever, the researches expect that using their new drug will reduce the number of reccurrences to only 1 per year. The researchers have various methods set up for following the participants over the course of the study, but they believe that both groups have a 10% chance of each recurrence not being observed or found by the researchers. The researchers decide that they want to have a significance level of 0.05 and 80% power. The researchers want to know how many participants will need to be recruited with a one-to-one allocation ratio between the two groups.

The researches would use the following values to determine what they need:

* Enrollment Schedule: Over a Period
* Distribution of Enrollment: Exponential
* Exponential Rate: 0.5
* Study Duration: 10 (years)
* Enrollment: 2 (years)
* Sample Allocation Ratio: 1
* Target Event Rate: 1 (recurrence per year)
* Target Censoring Rate: 0.1
* Reference Censoring Rate: 0.1
* Significance Level: 0.05
* Power: 0.8

With the values put in by the researches, they would find that they need to recruit a total of 58 participants, 29 in each arm.

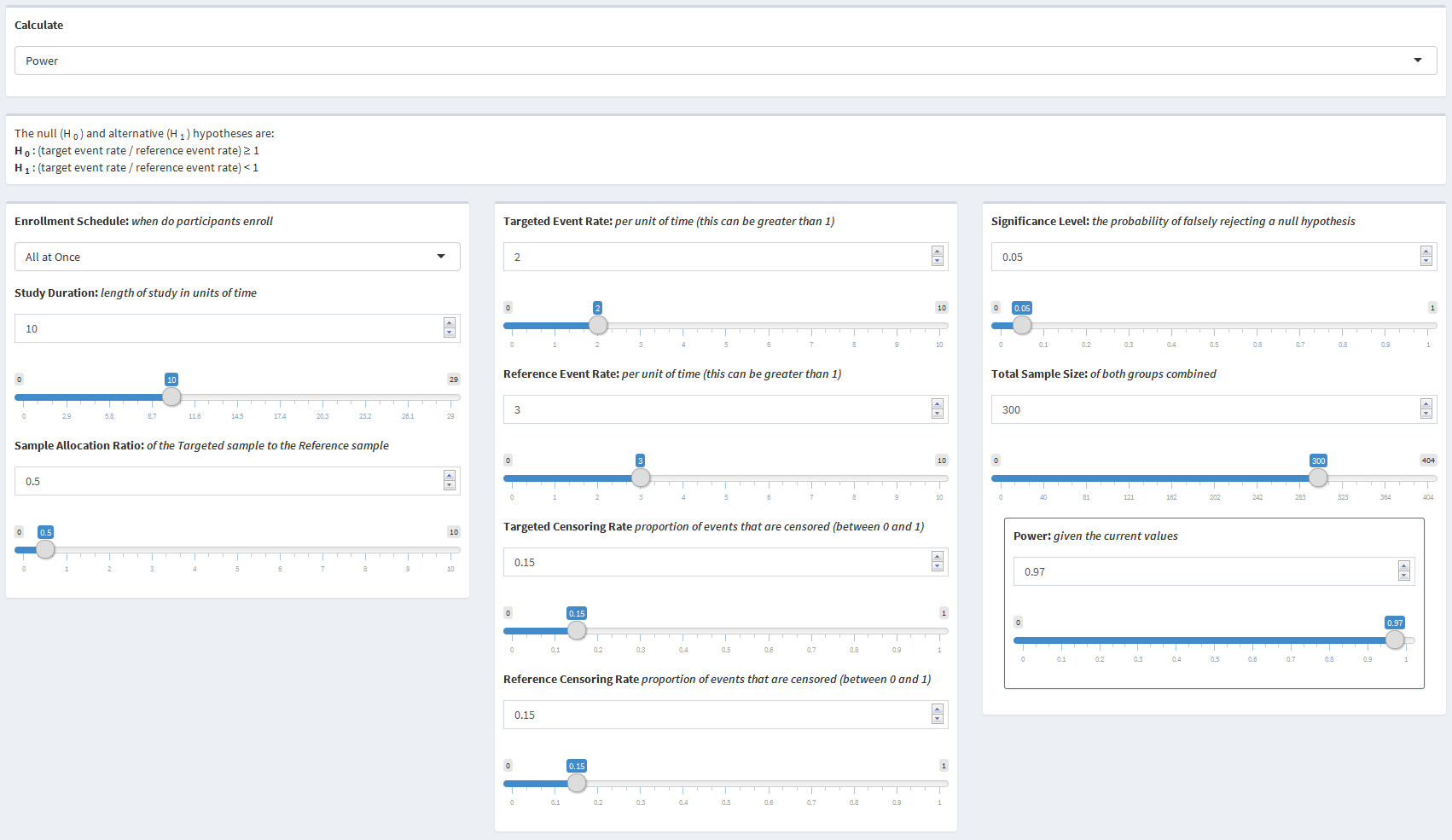
 

**Calculating Power**: A study was performed where the researchers were examining the effect of a new drug’s ability to help prevent cancer recurrence. The two arms of the study were general lifestyle guidelines with and without the new drug with 200 and 100 participants in each arm respectively. The study lasted 10 years with recruitment having been completed by the beginning of the study. The researchers found that group without the drug experienced recurrences at a rate of 3 per year, while the group taking the new drug experienced recurrences at a rate of 2 per year. It was estimated that both of the samples had a censor rate of 15%. They want to maintain a 0.05 significance level. The researchers want to know how much power their study had.

The researches would use the following values to determine what they need:

* Enrollment Schedule: All at Once
* Study Duration: 10 (years)
* Sample Allocation Ratio: 0.5
* Target Event Rate: 2 (recurrences per year)
* Reference Event Rate: 3 (recurrences per year)
* Target Censor Rate: 0.15
* Reference Censor Rate: 0.15
* Significance Level: 0.05
* Total Sample Size: 300

With the values put in by the researches, they would find that their study has 94% power.



# 4. Further Improvements

Like most things, this dashboard is not perfect and is not in its final state. Throughout working on this dashboard, there were many areas that were noted for possible upgrades or improvements. Below are just some of the possible areas for further improvement.

## 4.1 User Interface

### 4.1.1 Layout

The most important aspect of the user interface is its layout. The goal of this dashboard was to have the output visible while changing any of the inputs. This allows the user to to see the direct effect the input has on the output. That being said, it is also important for the inputs and outputs to be grouped in such a way that is intuitive. For the most part, the current layout does the best it can to maximize both of these aspects within the limitations of the environment. Further development can be performed to optimize the layout.

### 4.1.2 Calculation Selector

The original goal during the development of the dashboard was to have all of the inputs and outputs visible at the same time and to be interconnected. For example, the user would be able to change the power to have the sample size update, and then change the sample size to have the power update, all withouth having to select a specific variable to calculate. However, limitations of reactive values in Shiny meant that a fully interconnected page resulted in infinite loops of inputs updating. Further research into reactive values might result in achieving a fully interconnected page.

### 4.1.3 Slider Units

One of the nice features about Shiny Dashboards are the interactive sliders. Dynamic bounds and intervals appropriate for the value being presented can be added to maximize the viusal benefits of a slider. There are currently semi-dynamic bounds on most of the sliders, however, more appropriate bounds can be added, especially for numbers near zero.

### 4.1.4 Informative Dialog

One of the next steps for this dashboard is shaping it into an educational tool. While many users of this application will have an in depth knowledge of the statistical concepts used, many other users will have little to no understanding of the concepts. That being said, it is important that everyone using the dashboard understands how to use it and what the output means. Further improvements in this area will include an introductory page that explains how to use the dashboard similar to section 2 of this document, a box at the top of each page explaining each of the scenarios, and a box at the bottom of each page summarizing and explaining the results.

## 4.2 Server

### 4.2.1 Functions

Most of the code on the server side of the dashboard is unique to each page, and thus, can not be simplified by using a function across all of the pages. It is possible that the portion in the server that updates all of the inputs can be simplified to a be a function that, instead of updating only the inputs for the specific page and what is being calculated updates all of the inputs across the entire dashboard. This would save numerous lines of repeating code, but may result in less efficient computations.

### 4.2.2 Consolidation

As mentioned above, most of the server code is unique throughout and does not allow for much improvement. The updating of inputs has been identified as a possilbe area that can be improved. Besides upgrading it to a function, the number of times each page uses the updating code could be consolidated from two times to just once. Another code section that can be consolidated is in the solving of each calculation. Currently, the code defines all of the input variables into temporary variables first, then uses the temporary variables in the function call to calculate the output. The input variables can be used directly in the function call to consolidate a considerable amount of code.

## 4.3 Alternative Hypothesis

For simplicity, each page currently has a static alternative hypothesis (two-sided for one mean, one proportion, two means, and two proportions and one-sided for time to event). The functions used by the dashboard, however, have the capability to perform both one-sided and two-sided calculations for each scenario. Adding this functionality would allow for non-inferiority, superiority, and equivalence scenarios.

## 4.4 Time to Event

As time to event scenarios are generally more complex than the rest of the scenarios, it follows that there are more parameters that can be used for more unique situations. One parameter that can change is whether the risk difference or the risk ratio should be used as the comparison. The power and sample size equations change slightly depending on which comparison is being used. Although both of the comparisons are examining the same parameters, they are interpreted slightly differently and having the ability to choose instead of just calculating the risk ratio is definitely an upgrade.

## 4.5 New Pages

Although this dashboard covers most of the common scenarios that will be seen in practice, there are still numerous other scenarios that can be incorperated into the dashboard. Two of the possible additions could be paired versions of the two means and two proportions pages where each sample is compared with itself. Two other possible additions are versions of mean and proportion comparisons where more than two, k, means or proportions are being compared. These additional four scenarios would help to cover as many of the possible scenarios that would be seen in practice.

# 5. References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creating Shiny reactive variable that indicates which widget was last modified. (2015, July 6). *r - Creating Shiny reactive variable that indicates which widget was last modified - Stack Overflow. stack overflow*. <http://stackoverflow.com/questions/31250587>. Accessed 19 April 2017

HyLown Consulting LLC. *Overview of Power and Sample Size .com Calculators*. Power and Sample Size Calculators. HyLown Consulting LLC. <http://powerandsamplesize.com/Calculators/>. Accessed 19 April 2017

Lachin, J. M., & Foulkes, M. A. (1986). *Evaluation of Sample Size and Power for Analyses of Survival with Allowance for Nonuniform Patient Entry, Losses to Follow-Up, Noncompliance, and Stratification*. Biometrics, 42(3), 507-519. <doi:10.2307/2531201>

Lenth, R. V. (2006, August 14). Java applets for power and sample size. *Russ Lenth's power and sample-size page*. University of Iowa. https://homepage.divms.uiowa.edu/~rlenth/Power/. Accessed 26 April 2017

RStudio. *Shiny*. Shiny. RStudio. <http://shiny.rstudio.com/>. Accessed 19 April 2017

RStudio. *Reactivity: An overview*. Shiny - Reactivity: An overview. RStudio. <http://shiny.rstudio.com/articles/reactivity-overview.html>. Accessed 19 April 2017

RStudio. *Shiny Dashboard*. Shiny Dashboard. RStudio. <https://rstudio.github.io/shinydashboard/index.html>. Accessed 19 April 2017

RStudio. *R Markdown*. R Markdown. RStudio. <http://rmarkdown.rstudio.com/index.html>. Accessed 19 April 2017

Schoenfeld, D. A. (1981). *The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions*. Biometrika, 68(1), 316-319. <doi:10.2307/2335833>

Schoenfeld, D. A. (1983). *Sample-Size Formula for the Proportional-Hazards Regression Model*. Biometrics, 39(2), 499-503. <doi:10.2307/2531021>