**Assumptions**

The sample size has these common factors:

* The sampling method for each population is simple random sampling as the sample consists of n objects all of which are equally likely to occur.
* The samples are independent where the occurrence of one does not affect the probability of the next event.
* Each sample includes at least 10 successes and 10 failures as is evident in the table output (table(x) and table(y) in code).
* Each population is at least 20 times as big as its sample.

A double-blind test is helpful in reducing bias with testing. These simulations will create a Binomial distribution with discrete probabilities. With large enough n (sample size) and a fixed probability of success for each outcome, the Binomial distribution can be approximated by a Normal Distribution with an area under the curve that equals 1. Therefore, the two-proportioned z-test will suffice for this statistical analysis.

Let p1 = Probability of occurrences of 1’s in first sampling distribution = 0.02,

p2 = Probability of occurrences of 1’s in second sampling distribution = 0.05,

α = 0.05 95% is the most common confidence interval.

**Hypothesis**

H0: p1 = p2 Probability of occurrences of 1’s are the same in both sampling distributions.

HA: p1 ≠ p2 Probability of occurrences of 1’s are not the same in both sampling distributions.

**Data**

Table **x** is the probability of 1’s occurring at 2%. Table **y** is the probability of 1’s occurring at 5%.

x

0 1

4904 96

y

0 1

4784 216

2-sample test for equality of proportions with continuity correction

data: c(table(x)[[2]], table(y)[[2]]) out of c(5000, 5000)

X-squared = 87.7825, df = 1, p-value < 2.2e-16

alternative hypothesis: two.sided

95 percent confidence interval:

-0.04341409 -0.02818591

sample estimates:

prop 1 prop 2

0.0196 0.0554

As the p-value is < 2.2e-16 and <4.4e-16 for a two-sided interval test, it is still less than 0.05 and as a result, we reject the null hypothesis that the probabilities of the occurrences of 1’s in both sampling distributions are the same.

In the sample in the table mentioned above, the observed effect size for one sampling distribution is 96/500 and 216/5000 which is roughly 1.92% and 4.32% respectively. After conducting the two proportioned z-test, we find insufficient evidence to support the null hypothesis with 95% confidence. The following formula will show if the sample size of 5000 is large enough for the two samples.

Necessary Sample Size = (Zα/2+Zβ)^2 \* (p1(1-p1)+p2(1-p2)) / (p1-p2)^2

Standard Error = sqrt(p1(1-p1)/n)

Critical Value = 1.96 Critical Value in a Z-test is expressed as the Z-score

where Zα/2 is the critical value of the Normal distribution at α/2 (For a confidence level of 95%, α is 0.05 and the critical value is 1.96), Zβ is the critical value of the Normal distribution at β (For an assumed power of 80%, β is 0.2 and the critical value is 0.84) and p1 and p2 are the expected sample proportions of the two groups

Necessary Sample Size is calculated to be 584.52. As the sample sizes chosen for this simulation separately exceed 585 for both simulations, the sample size of 5000 is large, has a roughly normal approximation to the binomial distribution and the results are statistically significant to within 95% of the true proportion for the difference between the proportions. The assumed power takes an 20% chance of failing to detect a significant difference when it exists (a Type II error) within the suggested sample size of 585.

The power using our sample size which is 5000, is 1 which means the Type II error is very, very low and we have a nearly 0% chance of failing to detect a significant difference. This is directly affected by the confidence intervals as the higher the confidence interval, the smaller the region of acceptance and the higher the likelihood of rejecting the null hypothesis. The power of the test is also determined by values of the true parameters being tested. For example, an unbiased coin has 50% chance of producing a 1 which is far from the 2% and 5% probabilities of the biased coin that are tested. The larger the difference in effect size, the larger the power of the test. The extremely small probability of Type II error also affects Type I error by affecting the region of acceptance and presuming the probabilities are the same with 1-α level of significance within the sample.

**Input (in R, also attached as a separate file)**

x <- sample(c("1","0"), size=5000, replace=TRUE,prob=c(.02, .98)) table(x)

y <- sample(c("1","0"), size=5000, replace=TRUE,prob=c(.05, .95)) table(y)

prop.test(c(table(x)[[2]],table(y)[[2]]), c(5000,5000), alternative="two.sided", conf.level=.95) # test statistic

**Output**

x

0 1

4902 98

y

0 1

4723 277

2-sample test for equality of proportions with continuity correction

data: c(table(x)[[2]], table(y)[[2]]) out of c(5000, 5000)

X-squared = 87.7825, df = 1, p-value < 2.2e-16

alternative hypothesis: two.sided

95 percent confidence interval:

-0.04341409 -0.02818591

sample estimates:

prop 1 prop 2

0.0196 0.0554