DigitalCommons@WayneState

Wayne State University

Wayne State University Dissertations

1-1-2010

Type Ii Robustness Of The Null Hypothesis Rho = 0 For Non-Normal Distributions

Stephanie Wren
Wayne State University

Follow this and additional works at: http://digitalcommons.wayne.edu/oa_dissertations

Part of the Statistics and Probability Commons

Recommended Citation

Wren, Stephanie, "Type Ii Robustness Of The Null Hypothesis Rho = 0 For Non-Normal Distributions" (2010). *Wayne State University Dissertations.* Paper 192.

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.

TYPE II ROBUSTNESS OF H_0: ρ=0 FOR NON-NORMAL DISTRIBUTIONS

by

STEPHANIE D. WREN

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF EDUCATION

2010

MAJOR: EDUCATION
EVALUATION AND RESEARCH

Approved by:

ADVISOR DATE

© COPYRIGHT BY

DR. STEPHANIE D. WREN

2010

All Rights Reserved

DEDICATION

I would like to dedicate my dissertation to my mother Vivian Wren. Your undying support and belief in me is what propelled me and motivated me through the moments in which I felt I could not go on. To my brother James Wren III, through rain, snow, sleet, or shine, I know you have got my back. And, I have got yours. To my niece Shelly Wren, I expect to see my name on the dedication page of your dissertation someday soon. And, finally, to my uncle Walter Scott, you do not have to wonder anymore because I am finally finished.

Dr. Stephanie D. Wren

ACKNOWLEDGEMENTS

I would like to acknowledge my committee members Dr. Shlomo Sawilowsky, Dr. Gail Fahoome, and Dr. Cheryl Somers. Dr. Sawilowsky I would thank you for walking me through this process and skillfully clarifying my missteps along the way. Dr. Fahoome I would like to thank you for taking the time to assist me with the Fortran code that I had to write because I do not think I could have done it without you. Dr. Somers I would like to thank you for not only agreeing to be a member of my committee without hesitation but also for giving me the chance to become a published author.

I would like to thank the College of Education at Wayne State University. I have had nothing short of a positive and rewarding experience as a student in the college of education. I would also like to acknowledge Paul Johnson in the graduate office in the College of Education at Wayne State University. I will always think of you whenever I hear the phrase "eleventh hour." And, lastly, I would like to acknowledge Dr. Donald Marcotte. Dr. Marcotte was my original advisor. Unfortunately, he became ill and passed away before I completed my dissertation. Under Dr. Marcotte's tutelage my interest in educational research truly blossomed. I will always appreciate and remember him for that.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

A power analysis can be performed in the planning phases of an experiment to determine if the sample size is sufficiently large to find a statistically significant result if it exists. Power analyses are an amalgamation of four components: (1) choice of one-tailed vs. two tailed tests (where appropriate), (2) the significance level (α), (3) the effect size, and (4) the sample size.

In the context of testing whether the Pearson product moment correlation coefficient, r, is statistically significant, the null hypothesis is H_0: $\rho = 0$. The alternative hypothesis may be either H_a: $\rho < 0$ or H_a: $\rho > 0$ which are unidirectional or H_a: $\rho \neq 0$ which is bidirectional. In research contexts where the direction may be postulated, the unidirectional test should be selected because it has more statistical power than a bi-directional test.

The significance level (α), the second component, is the a priori threshold of what constitutes a rare event. When the underlying assumptions of a statistical test are met, α also sets the likelihood of committing a Type I error; which is the probability of rejecting a true null hypothesis. This is commonly known as the false positive rate. Customary significance levels set for this threshold in the behavioral and social sciences are 1% and 5%. The larger α level is often selected when additional statistical power is required, whereas the lower α is chosen when the cost of a making a Type I error is greater.

According to Cohen (1988), an effect size is “the degree to which the phenomenon is present in the population.” (p. 9). Hence, r is a measure of effect.

Whether r is positive or negative, its squared magnitude is a measure of the relative strength of the linear relationship between an independent variable and a dependent variable. As the magnitude of r is hypothesized to be larger or smaller, power will increase or decrease respectively.

Sample size also has an effect on power. It is widely known that smaller samples are less accurate, reliable, and precise while larger samples are more accurate, reliable and precise. These differing characteristics can have an effect power. Additionally, a review of Cohen's (1988) correlation power tables will show that if r (effect size) is held constant, then smaller samples have lower power than larger samples. And, as the effect size, r, increases, the power of smaller samples increases, but the power of larger samples is greater. Thus, intangibles of sample sizes such as reliability, accuracy, and precision affect power just as the statistical properties of effect size as well.

As a descriptive statistic, there are statistical assumptions associated with r. Havlicek and Peterson (1976) noted these assumptions are (1) the scale of measurement is continuous for both variables, (2) a normal bivariate distribution exists, (3) each pair of scores is independent of the other scores, and (4) there is a rectilinear relationship between the variables

As an inferential statistic, violating these assumptions may have an adverse effect on the statistical properties of hypothesis tests on the correlation. For instance, Havlicek and Peterson (1976) performed a Monte Carlo simulation on the effects of violating the continuous scale of measurement assumption. Although there were slight deviations from the normal distribution, Havlicek and Peterson (1976) ruled those deviations to be marginal and concluded that the null hypothesis of $\rho=0$ was robust to violations of measurement scale.

Another Monte Carlo study addressed the issue of measurement error. Brooks, Kanyongo, Kyei-Blankson, and Gocmen (2002), found that increasing measurement error actually decreased statistical power while decreasing measurement error had the opposite effect. Several other researchers obtained mixed results when the normality statistical assumption was violated (Student, 1908, Soper, 1913, Fisher, 1915, Pearson, 1931, Pearson, 1932, and Rider, 1932).

Research Problem

The question remains how small samples drawn from non-normal populations affect the Type II error properties when testing the hypothesis H_a = some value other than 0. The literature is extensive regarding small samples and the null hypothesis $\rho=0$. Fisher, Soper, and Student have provided mathematically labor intensive formulas to replicate the frequency distribution of r when $\rho=0$ for normally distributed data (Student 1908, Soper 1913, and Fisher 1915). And, to tie together the work of Fisher, Soper, and Student, Florence Nightingale David (1938) completed the formulation of the ordinates to define the shape of the frequency distribution of r from a normally distributed population for samples as small as three to as large as 400. The tables which contained the ordinates also evaluated the frequency distribution of r for $\rho=0.0$ through $\rho=0.9$. Additionally, David (1938) also computed the probability integral for values of r ranging from -1.00 to +1.00 corresponding to their respective ordinates and ρ. The extensive tables created by David (1938) are the basis for the correlation power tables found in Cohen (1988). Given the research pertaining to the correlation coefficient, what remains to be investigated is the relationship between r and the Type II error rate. Blair and Lawson (1982) found that the correlation coefficient was not quite robust to the Type I error rate with respect to Bradley's (1978) liberal criterion. Sawilowsky and Hillman (1991) and Sawilowsky and Blair (1992) found that the independent samples t test for nonnormal datasets were robust with respect to Type I and Type II error rates. So, to clarify the research on r in small samples, an investigation of r in small samples from non-normal populations with $\rho \neq 0$ and its effect on Type II error properties will be studied here. More specifically, given a null hypothesis of $H_o{:}\rho=0$, what effect will altering a treatment of $H_a{:}\rho \neq 0$ by varying levels have on the number of false negatives of the correlation coefficient?

Significance of the Problem

This problem is significant given that many researchers use the correlation coefficient and power tables without regard to whether the parent population is normally distributed. Consequently, interpretation of the results may be based on samples that are not sufficient or power values that are underestimated or overestimated.

Limitations

The significance levels will be limited to 1% and 5%.

Chapter 2

Literature Review

Brooks, Kanyongo, Kyei-Blankson, and Gocmen (2002) addressed the issue of measurement error and its effect on r with a Monte Carlo simulation. They created normally distributed, yet low reliability data. The low reliability data was defined as the proportion or ratio of raw score variance to true score variance. They analyzed the effect of low reliability on, among other values, the correlation coefficient. They found that as reliability changes so does the raw score variance. In other words, reducing reliability resulted in increasing raw score variance. An increase in raw score variance decreased statistical power. They also found the converse to be true as well.

The test of the correlation appears to be robust to the violation of independence in very mild cases. Edgell and Noon (1984) conducted a Monte Carlo simulation on paired data in which both variables were dependent (one variable was the squared representation of the other) and the population correlation was zero. They found that if the correlation for the distribution of either variable was small and the probability of selecting the given distribution was large, then r was robust. Otherwise, r could be extremely sensitive to dependence.

Havlicek and Peterson (1976) studied the effect on r when the scale of measurement assumption was violated. Havlicek and Peterson (1976) performed a Monte Carlo simulation on the effects of violating the continuous scale of measurement statistical assumption. Nine groupings (in independent pairs) of data transformed into interval, ordinal or percentiles were created. Five thousand iterations of samples of size 5 and 15 were produced and evaluated to determine the proportion of rs that would exceed a given level of significance ranging from .5% to 5%. Although the distribution of rs deviated from the theoretical distribution, it was determined that the deviations were essentially negligible. Hence, hypothesis tests of Ho: r=0 appear to be robust to violations of scale of measurement.

Using a real data set, and Monte Carlo methods, Student (1908) studied the distribution of the sampled r when the population correlation ρ was zero or nonzero. His methodology entailed measuring the left middle finger of 3000 criminals. The measurements were stature and length. He put together 750 samples of four correlations (n=4) of stature and length. The population correlation for this grouping was $\rho=0.66$.

Next, to create independent samples whose population correlation would be zero, Student (1908) measured the stature of one group and correlated that with the lengths from a different group. This methodology resulted in 750 samples of size four as well. To create larger samples (n=8), Student (1908) combined two samples of size four by adding one sample to the tenth sample before it and after it. This procedure produced 750 samples of size eight for which $\rho=0$ or $\rho=0.66$.

Student (1908) used a frequency table to display the results. He found that in the independent case ($\rho=0$) more sample rs tended to group around zero than not. Yet, when there was a correlation in the parent population ($\rho=0.66$) the sampled correlations tended to be larger than ρ.

If the sample size was small, then the discrepancy tended to be large. For instance, in samples of size four with $\rho=0.66$ a majority of the sampled rs were in the range of 0.93 to 0.97. But, when the sample size was increased to eight, a majority of the sampled rs ranged from 0.78 to 0.82. So, conversely, as the sample size increased the error decreased as well.

Indeed, in the same study, Student (1908) conducted an examination on a larger sample and determined that when the frequency distribution of 100 samples of size 30 was analyzed, a majority of the correlations ranged from 0.68 to 0.70. This range of sample correlations was more reflective of the population correlation ($\rho=0.66$) than the correlations which corresponded to samples of size four and eight. So, the research of Student (1908) showed that there was a discrepancy between r and ρ not only when the samples were small, but also when the samples were small and ρ was large.

Soper (1913) continued the discussion regarding the error of the correlation coefficient by comparing his results to the results of Student (1908). In his study, Soper (1913) determined that if the sample size was small or the correlation coefficient was near its end points, then the frequency distribution of r was skewed. He also concluded that the value of r was more likely the modal value in the sample and not ρ or the mean value of r. Soper's (1913) results were extrapolated using a lengthy mathematical proof. However, he established the point that if n was small and ρ large, the modal value of r can be larger than ρ and possibly greater than one.

Additionally, when Soper (1913) compared the mean r to Student's (1908) mean r, in both cases, the error differences between mean r and ρ decreased as the sample size increased. Also, yet not surprising, as the sample size continued to increase mean r began to reflect ρ. Conversely, it can be concluded that descriptive estimates of ρ in small samples can be unstable.

In a more comprehensive study, Soper, Young, Cave, and Pearson (1917) also investigated the frequency distribution of r in small samples. Again using an elaborate mathematical proof, Soper et al. (1917) illustrated that in small samples (2-25) with large ρ (0.6 or 0.8) the frequency distribution was slow to approach the normal distribution. Although samples of 25 with $\rho=0$ more closely resembled the normal distribution, as ρ increased from 0.0 to 0.9 the frequency distribution tended to deviate quite severely from normality.

Up until this point, experimental research focused on investigating the correlation coefficient when the assumption of normality was not violated (Student, 1908; Soper, 1913; and; Fisher, 1915). Pearson (1931) undertook the task of investigating r when the assumption of normality was violated and $\rho=0$.

Regarding the consequences of violating the normality assumption, Pearson (1931) considered leptokurtic, slightly skewed, and very skewed distributions. Iterations of 250 to 395 were used to create sampling distributions of size 10 and 20. Although there were slight departures from the theoretical normal distribution (as measured by the standard deviation of r), the results appeared to indicate that r is essentially insensitive to the effects of non-normality.

As a continuation of his earlier work, Pearson (1932) investigated the effects on ρ if the population distribution for each variable differed (i. e. one variable is skewed and the other variable is leptokurtic). His methodology consisted of three series of xy variable pairings such that within each series at least one variable's distribution was symmetrical. Five hundred samples of size 10 and 20 each were drawn and their frequency distributions were developed. The chi square goodness of fit test was performed to compare the overall fit of the sampling distributions with the normal distribution. The results showed that the fit was "remarkably satisfactory."

Prior to Pearson's (1932) study of the effects of variable distribution on ρ, Baker (1930) investigated the effects of outliers on the correlation coefficient. Baker (1930) collected mortality rates from 50 large cities for several years and seasons for pneumonia and influenza. The years in the study represented the variables and the cities represented the bivariate data. The correlation coefficient was computed with extreme variables included and with extreme variables excluded. When the correlation coefficient was computed with the outliers included, Baker (1930) found that r was notably larger than the correlation computed with the outliers excluded.

Abdullah (1984) performed a similar study to Baker (1930). Abdullah (1984) investigated the effects of outliers on the robustness of the correlation coefficient. In the study, 100 observations were generated using a linear relationship in which y was a function of x which was normally distributed. The population correlation of the generated data was ρ=1. And, the sample correlation coefficient was r=0.984.

Abdullah (1984) then contaminated the data by replacing fifty data points with outliers. The contaminated data was generated using a linear relationship in which x was uniformly distributed and y was normally distributed. The proportion of contamination was in multiples of ten. From 0% contamination to 10% contamination, the correlation coefficient reduced from 0.984 to -0.070. As the amount of contamination (as measured by the outliers) increased, the greater the sample correlation coefficient was in error of the population correlation coefficient.

Srivastava and Lee (1984) measured contamination differently than Abdullah (1984). Srivastava and Lee (1984) defined a contaminated bivariate normal distribution as a "mixture of two bivariate normal distributions with zero mean but different covariances and mixing proportions 1-λ and λ, respectively." The intent of Srivastava and Lee (1984) was to study how well transformations of r performed when the population of r was defined as above and modeled using the probability density function $(1-\lambda)\varphi(x;0,I)+ \lambda\varphi(x;0,bI)$ where λ was between 0 and 1 inclusive, b, which represents the variance of the population, was at least 3, and $\rho=0$.

Srivastava and Lee (1984) selected samples of size six and ten. To approximate the sampling distribution of r, Student's t and Fisher's z approximations (among others) were used. Srivastava and Lee (1984) computed the proportion of rs, for each distribution, that violated nominal alpha of 5% and 10% based upon varying levels of contamination and variance. Srivastava and Lee (1984) found that Fisher's z and Student's t did not perform well under the given conditions. More specifically, as nominal alpha increased, the sample size increased, contamination increased, and the variance increased, so did the proportion of rs that violated alpha for Student's t and Fisher's z. So, although Student's t and Fisher's z may be close approximations of normal bivariate distributions of r, when the distribution is not normal, then Srivastava and Lee (1984) showed that Student's t and Fisher's z are not acceptable transformations of r.

At the behest of Pearson, Rider (1932) built upon the results of Pearson. Using a pseudo-random number generator, 1,000 samples of size five were obtained for rectangular, triangular, and normal distributions. Additionally, 500 samples of size ten were generated for the triangular and normal distributions. The rectangular distribution produced a population correlation of zero, the triangular distribution produced a population correlation of 0.5, and the normal distribution produced a population correlation of 0.9.

Using the Chi-squared goodness of fit test, Rider (1932) found that the observed frequency distribution for the rectangular population fit the normal theory frequency distribution quite well when $\rho=0$. The triangular distribution generated mixed results. The fit of the frequency distribution of samples of size five was acceptable; but, the fit of the frequency distribution of samples of size ten was weak. And, after performing Fisher's transformation on the normal distribution with $\rho=0.9$, Rider (1932) found that the fit of the distribution was weak also. So, when $\rho=0$, the fit of r was acceptable. When $\rho\neq0$, the fit of r was unacceptable as noted in the case described above. A detailed examination of Rider (1932) shows that a non-normal distribution with $\rho\neq0$ coupled with a small sample can produce a sampled r which may be in error.

The mixed results of the triangular distribution of Rider (1932) were contradicted by Kowalski (1972). Kowalski (1972) used a 'Fourier approach' to evaluating the correlation coefficient derived from a population that was not normal. One of his objectives was to test Pearson's claim that r was insensitive to non-normality.

Using a Fourier estimator, Kowalski (1972) evaluated the distribution of r by creating a non-normal population from a mixture of bivariate normal distributions. One hundred samples of size 30 were drawn from the indicated population with $\rho=0$. Kowalski (1972) found that r was sensitive to departures from normality which was in contrast to the results of Pearson.

Robustness studies using data gathered in true research settings should be considered a next step in investigating how r functions. Many of the distributions used to evaluate the robustness of r are the popular distributions used in most robustness studies (i.e. normal, exponential, or uniform distributions). Micceri (1989) decided to challenge the notion of normality. After collecting 440 psychometric and educational data sets from journal articles, statewide tests, national tests, college entrance exams, and various other research initiatives, Micceri (1989) evaluated each distribution to determine how closely each matched the Gaussian (normal) distribution. The normality characteristics measured were tail weight and symmetry/asymmetry. Although there were a few that were nearly Gaussian, none of the distributions inspected passed all of the tests for normality.

Micceri (1989) postulated that "60% of all distributions result directly from research and 33% percent from state, district, or university scoring programs." Hence, the use of theoretical distributions to test the robustness of r may not be as instructive as evaluating distributions which are used in practice.

Blair and Lawson (1982) expressed a similar concern regarding using the popular distributions for robustness studies. Their goal was to be able to provide a more rigorous test of the robustness of r when the population was not normal and when $\rho=0$. Blair and Lawson (1982) used a Bradley distribution which is a mixture of three normal probability distributions with varying means and standard deviations. The resulting population is 'L' shaped with a skew greater than three and kurtosis roughly 17.

A Monte Carlo simulation was employed. Samples of size 5, 30, 50, and 100 were repeatedly sampled 5,000 times. The correlation coefficient was computed each time and the proportion of nominal alpha (1%, 2%, 5%, and 10%) violations was recorded for each sample. Blair and Lawson (1982) found significant violations of nominal alpha for samples greater than size five. Nominal alpha in the lower tail was consistently deflated while nominal alpha in the upper tail was consistently inflated.

Sawilowsky and Blair (1992) also investigated the robustness of the t test with a twist which included the Type II error properties in populations that were not normally distributed. Ironically, the Sawilowsky and Blair (1992) study evaluated robustness and Type II error properties of the t test on the Micceri (1989) distributions. Those distributions are defined as the discrete mass at zero with gap, extreme asymmetry (growth), mass at zero, extreme asymmetry (decay), extreme bimodality, multimodality and lumpiness, digit preference, and smooth symmetric.

In their study, Monte Carlo methods were used on independent samples of sizes five through 15 (matched equally or unequally) for each of the eight Micceri (1989) distributions to determine the robustness of the t test to Type I error. The results showed that the divergence from nominal alpha of 5% or 1% were within the limits of Cochran's (1947) and Bradley's (1978) criterion.

Monte Carlo methods were used to test the robustness of the t test to Type II error rates. To do so, transformed variables were created by shifting the location of one of the variables by two-tenths, five-tenths, eight-tenths, or one and two-tenths of a standard deviation. Doing so produces the treatment effect needed to measure the Type II error rate. The results for this portion of the study showed that the Type II error rate was similar to the normal distribution. According to the authors, "…for shift alternatives, the rejection rate of the independent samples t test was maintained at a fairly consistent level, regardless of the population shape, sample size, and effect size."

A previous, yet quite similar study by Sawilowsky and Hillman (1991) produced similar results. In this study, the discrete mass at zero Micceri (1989) distribution was singled out as the data set to investigate given its likeness to psychometric or psychological measures. Regarding the Type II error rate specifically, Sawilowsky and Hillman concluded that, "…when confronted with nonnormal data sets such as this, psychology researchers need not make any modifications to Cohen's (1988) tables when making sample size determinations" (p. 3). In larger samples, the author's discovered that the results were in agreement with the normal distribution. In smaller samples, the loss of power was minimal to negligible at approximately 1%.

The Blair and Lawson (1982), Sawilowsky and Blair (1992), Sawilowsky and Hillman (1991), and the Micceri (1989) results makes an interesting statement about nonnormal populations, the correlation coefficient, and in particular Type II error rates. Blair and Lawson (1982) found that the correlation coefficient in small samples from non-normal populations affects the Type I error which by default will affect the Type II error rate. They determined that the Type I error in the lower tail of the distribution did not meet Bradley's (1978) liberal criterion for the 5% or 1% significance level. Micceri (1989) found that the practical distributions examined in his research did not mirror the normal distribution. But, Sawilowsky and Blair (1992) and Sawilowsky and Hillman (1991) showed that the independent samples t test was reasonably robust with respect to Type I and Type II error rates regarding the Micceri (1989) data sets. Hence, the goal of this study is to evaluate the relationship between the correlation coefficient from nonnormal populations and the effect it has on the Type II error rate. Doing so will similarly elucidate the agreement between the correlation coefficient and power as noted by Cohen (1988) as it pertains to the normal distribution.

Chapter 3

Methodology

A Monte Carlo simulation will be employed to evaluate the Type II error rate when the correlation coefficient is computed from small samples drawn from non-normal populations. Fortran Essential Lahey 90 version four along with its realpops 2.0 subroutine is the compiler and random number generator that will be used to perform the simulation. The following is a step by step explanation of the methodology that will be used for this investigation.

Step one will require extracting x and y scores of sample sizes 5, 15, 30, and 60 from the Normal Distribution, Chi square Distribution (1 DF and 2 DF), Laplace Distribution, and T Distribution (3 DF). Figures 1 – 4 are are author created illustrations of the distributions that will be used in the investigation. The illustrations were created using the drawing graphics in Micorsoft PowerPoint. Each of the graphics were converted to jpeg pictures and pasted into this document.

The sample sizes are within the range of those studied by Student (1908), Kowalski (1972), Rider (1932), Blair and Lawson (1982), and others. These scores will be uncorrelated and will be used to test the hypothesis that rho is 0 for each of the indicated distributions.

Step two will require computing the sample r by using the raw score formula for the linear correlation coefficient on the sample data extracted for each distribution. The null hypothesis that rho is 0 will be tested using the sample r computed in step two. The sample r computed in step two will be used in step three to calculate the t test statistic for bivariate data, $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$

In step four, the t test statistic calculated in step three will be compared to the t critical value for the corresponding sample sizes, degrees of freedom, and significance levels (5% and 1%). Steps one through four will be conducted 1 million times for each distribution, sample size, and significance level. The proportion of test statistics that exceed the critical values will be computed. This proportion will be used as a baseline to measure and compare the robustness of the t test for the normal distribution and nonnormal distributions. This is the typical type one error rate robustness study. It is expected that the results will reflect the significance levels of 5% and 1%. Next is an explanation of the investigation of the type two error rate for the correlation coefficient.

First, the Fleishman Method (Fleishman, 1978) will be used to correlate the data by 0.1, 0.5, 0.7, and 0.9 for each distribution and their corresponding skew and kurtosis. More specifically, Table 1 (Sawilowsky and Fahoome, 2003) details how variates extracted from the Normal Distribution will be correlated via a Fleishman Equation for each of the investigated distributions. The Fleishman Equation referred to in Table 1 (Sawilowsky and Fahoome, 2003) is the equation

$0 = \text{root}^2 * (B^2 + 6BD + 9D^2 + 2A^2\text{root}^2 + 6D^2\text{root}^4) - \rho.$

Next, the equation above will be solved within the written Fortran code to produce a positive solution root which will be used in the next set of equations:

$$1. \quad x = root * z_1 + z_2 * \sqrt{1 - root^2}$$

$$2. \quad y = root * z_1 + z_3 * \sqrt{1 - root^2}\ .$$

Equations 1 and 2 will be used to convert the variates extracted from the Normal Distribution into intermediate variates which reflect the distributions in Table 1 (Sawilowsky and Fahoome, 2003).

Subsequently, the following set of Fleishman Equations (Sawilowsky and Fahoome, 2003) will be used:

$$1. \quad X = A + BX_i + (-A)X_i^2 + DX_i^3$$

$$2. \quad Y = A + BY_i + (-A)Y_i^2 + DY_i^3\ .$$

Equations 3 and 4 will be used to convert the intermediate values found in equations 1 and 2 into values which are reflective of variates from their respective distributions with corresponding skew and kurtosis.

Table 1: Solutions To The Fleishman Equation For Selected Distributions

Distribution	Skew	Kurtosis	A	B	D
Normal Distribution	0	0	0	1	0
Chi Squared Distribution (1 DF)	2.8284	12	-0.5207	0.6146	0.02007
Chi Square Distribution (2 DF)	2	6	0.3137	0.8263	0.02271
Laplace Distribution	0	3	0	0.7828	0.0679
T Distribution (3 DF)	0	17	0	0.3938	0.1713

After which, the final set of variates, which will be correlated by 0.1, 0.5, 0.7, or 0.9, will be used to test the null hypothesis that rho is 0. In other words, by maintaining a null hypothesis of $\rho = 0$ and increasing the treatment effect (i.e. making ρ progressively stronger) this will elucidate Type II error properties of the t test and not Type I error properties.

Next, the linear correlation coefficient raw score formula will be used to compute the sample r on the final set of variates determined from equations 3 and 4 and for each of the indicated distributions. Continuing, the sample r computed from the raw score formula will be used to compute the t test statistic for bivariate data, $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$.

Then, the t test statistic will be compared to the t critical value for the corresponding sample sizes, degrees of freedom, and significance levels (5% and 1%). This process will be conducted 1 million times for each distribution, sample size, and significance level.

Finally, the proportion of test statistics that exceed the t critical values will be computed. These results will be compared to the baseline measurements previously explained. The expectation is that the percent of rejections for the null hypothesis will be more than the baseline measurement given that the null hypothesis is being made increasingly false. Upon completion of the simulation, the data gathered will be displayed in tables and charts.

Figure 1: Chi Square

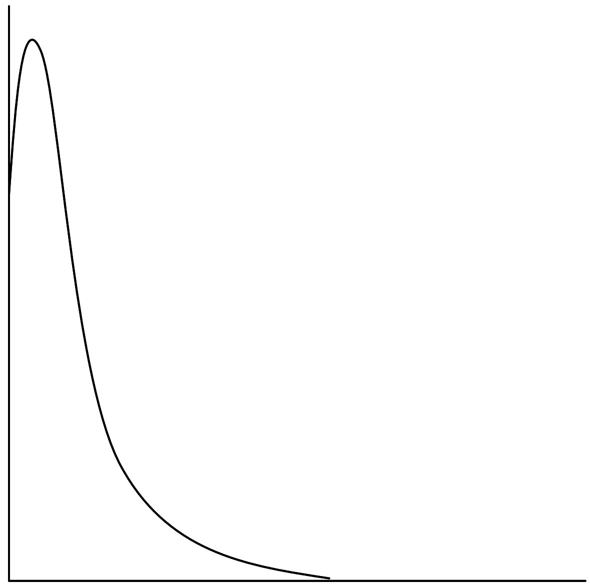


Figure 2: Laplace

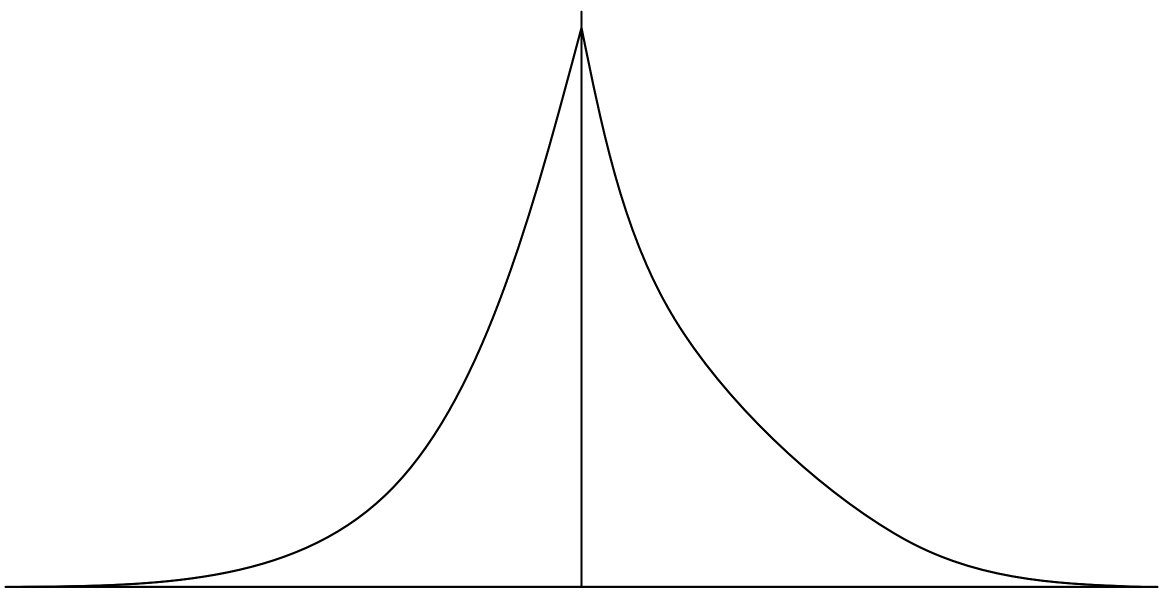


Figure 3: t Distribution

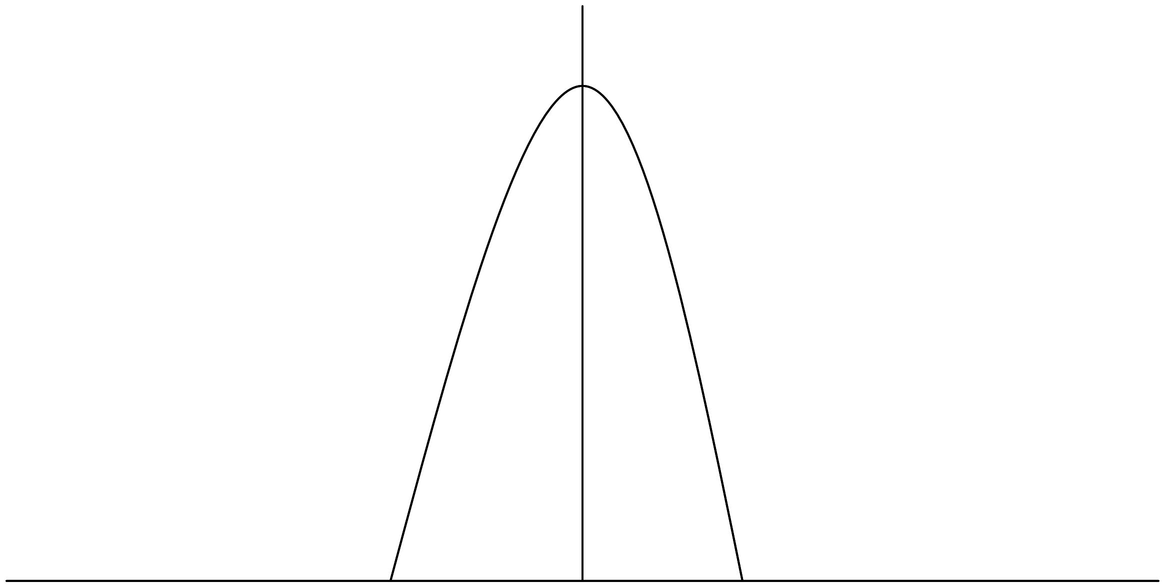
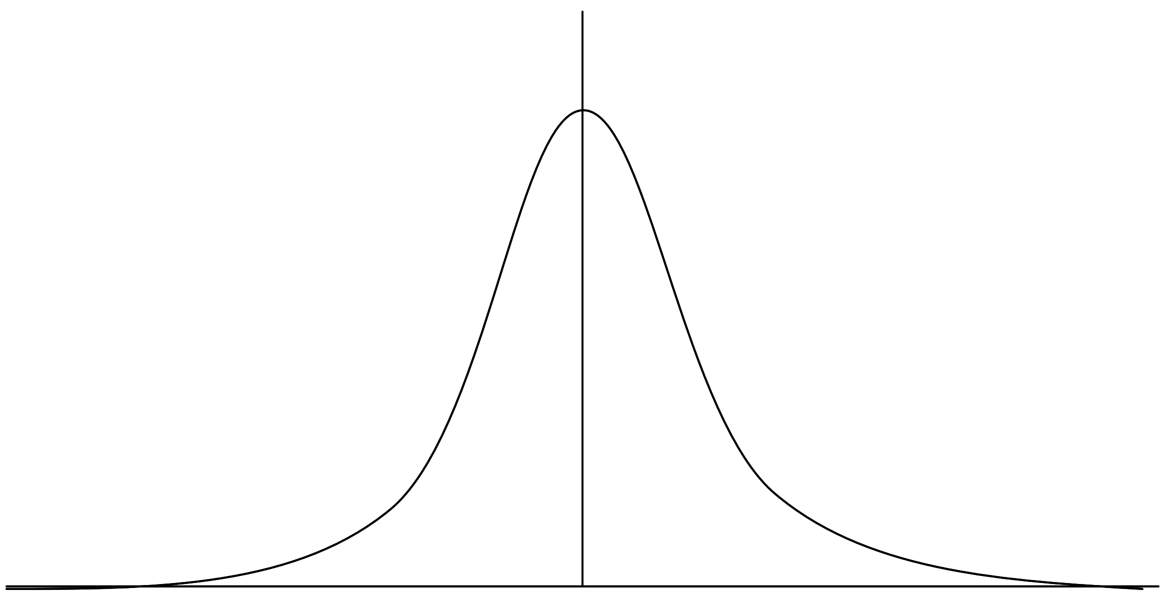


Figure 4: Normal Distribution



Chapter 4

Results

The results of the investigation are presented on the pages that follow. The results are presented in tables and charts. Tables 2 – 9 are the rejection rates for false null hypotheses corresponding to treatment effect, distribution, significance level, and sample size. The treatment effect for $\rho = 0.0, 0.1, 0.5, 0.7$, or 0.9 is located across the top row of each table. The distribution and sample size is located in the left column of each table. The title of each table indicates the significance level under investigation. The body of each table contains the rejection rates in decimal format for the upper tail and lower tail regions of each distribution. The values in red in each table are the sum of the upper tail and lower tail values for each treatment effect converted to a percentage. Summing and converting these values to percentages makes rejection rate interpretability evident.

Figures 5 – 50 are illustrations of each distribution corresponding to the red percentages from each of Tables 2 – 9. The vertical axis is the percent of rejections of false null hypotheses. The horizontal axes is the treatment effect for $\rho = 0.0, 0.1, 0.5, 0.7$, or 0.9. The title of each figure identifies the distribution, significance level, and sample size illustrated. The body of each figure is a line graph which represents the percentages of rejections of false null hypotheses.

Table 2: Type II Robustness at 5% Significance Level - Sample Size 5, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 5					
Normal Distribution					
upper tail	0.0250470	0.0341030	0.1270180	0.2687620	0.6720820
lower tail	0.0251560	0.0181060	0.0037570	0.0011580	0.0000830
	5.020300%	**5.220900%**	**13.077500%**	**26.992000%**	**67.216500%**
Chi Square Distribution (1 df)					
upper tail	0.0564260	0.0781430	0.2085880	0.3381290	0.6318800
lower tail	0.0089500	0.0072610	0.0027710	0.0011990	0.0001820
	6.537600%	**8.540400%**	**21.135900%**	**33.932800%**	**63.206200%**
Chi Square Distribution (2 df)					
upper tail	0.0399070	0.0594350	0.1932600	0.3403130	0.6772020
lower tail	0.0149470	0.0101480	0.0022610	0.0007490	0.0000720
	5.485400%	**6.958300%**	**19.552100%**	**34.106200%**	**67.727400%**
Laplace Distribution					
upper tail	0.0261990	0.0262810	0.0570940	0.1255350	0.4308880
lower tail	0.0261500	0.0244920	0.0105880	0.0037800	0.0003940
	5.234900%	**5.077300%**	**6.768200%**	**12.931500%**	**43.128200%**
T Distribution (3 df)					
upper tail	0.0259140	0.0300440	0.0655180	0.1389190	0.4242260
lower tail	0.0260070	0.0278530	0.0114020	0.0038790	0.0003920
	5.192100%	**5.789700%**	**7.692000%**	**14.279800%**	**42.461800%**

Table 3: Type II Robustness at 5% Significance Level - Sample Size 15, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 15					
Normal Distribution					
upper tail	0.0252240	0.0531030	0.5011600	0.8770330	0.9992850
lower tail	0.0250460	0.0107470	0.0000870	0.0000010	0.0000000
	5.027000%	**6.385000%**	**50.124700%**	**87.703400%**	**99.928500%**
Chi Square Distribution (1 df)					
upper tail	0.0508070	0.1000120	0.4768260	0.7744160	0.9910500
lower tail	0.0035600	0.0018330	0.0000750	0.0000080	0.0000000
	5.436700%	**10.184500%**	**47.690100%**	**77.442400%**	**99.105000%**
Chi Square Distribution (2 df)					
upper tail	0.0406750	0.0817130	0.5039500	0.8381700	0.9984680
lower tail	0.0100050	0.0037540	0.0000210	0.0000000	0.0000000
	5.068000%	**8.546700%**	**50.397100%**	**83.817000%**	**99.846800%**
Laplace Distribution					
upper tail	0.0259950	0.0278820	0.1411520	0.4670590	0.9784730
lower tail	0.0266130	0.0238540	0.0026850	0.0001310	0.0000000
	5.260800%	**5.173600%**	**14.383700%**	**46.719000%**	**97.847300%**
T Distribution (3 df)					
upper tail	0.0267770	0.0322590	0.1337660	0.3976870	0.9479080
lower tail	0.0269860	0.0282640	0.0040730	0.0002260	0.0000000
	5.376300%	**6.052300%**	**13.783900%**	**39.791300%**	**94.790800%**

Table 4: Type II Robustness at 5% Significance Level - Sample Size 30, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 30					
Normal Distribution					
upper tail	0.0251270	0.0751280	0.8280630	0.9952200	1.0000000
lower tail	0.0249180	0.0067590	0.0000030	0.0000000	0.0000000
	5.004500%	**8.188700%**	**82.806600%**	**99.522000%**	**100.000000%**
Chi Square Distribution (1 df)					
upper tail	0.0460630	0.1198440	0.7213500	0.9627520	0.9999790
lower tail	0.0039860	0.0012730	0.0000050	0.0000000	0.0000000
	5.004900%	**12.111700%**	**72.135500%**	**96.275200%**	**99.997900%**
Chi Square Distribution (2 df)					
upper tail	0.0378940	0.1011260	0.7829260	0.9874270	0.9999990
lower tail	0.0110010	0.0024110	0.0000010	0.0000000	0.0000000
	4.889500%	**10.353700%**	**78.292700%**	**98.742700%**	**99.999900%**
Laplace Distribution					
upper tail	0.0258300	0.0288200	0.2576570	0.7908980	0.9993400
lower tail	0.0258650	0.0229970	0.0006160	0.0000200	0.0000000
	5.169500%	**5.181700%**	**25.827300%**	**79.091800%**	**99.934000%**
T Distribution (3 df)					
upper tail	0.0264290	0.0321230	0.2122560	0.6706480	0.9993790
lower tail	0.0262680	0.0266610	0.0014770	0.0000130	0.0000000
	5.269700%	**5.878400%**	**21.373300%**	**67.066100%**	**99.937900%**

Table 5: Type II Robustness at 5% Significance Level - Sample Size 60, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 60					
Normal Distribution					
upper tail	0.0249430	0.1153850	0.9868140	0.9999980	1.0000000
lower tail	0.0254010	0.0033990	0.0000000	0.0000000	0.0000000
	5.034400%	**11.878400%**	**98.681400%**	**99.999800%**	**100.000000%**
Chi Square Distribution (1 df)					
upper tail	0.0422400	0.1553470	0.9346080	0.9938600	1.0000000
lower tail	0.0059280	0.0008760	0.0000000	0.0000000	0.0000000
	4.816800%	**15.622300%**	**93.460800%**	**99.386000%**	**100.000000%**
Chi Square Distribution (2 df)					
upper tail	0.0354980	0.1373880	0.9707040	0.9999670	1.0000000
lower tail	0.0135390	0.0014180	0.0000000	0.0000000	0.0000000
	4.903700%	**13.880600%**	**97.070400%**	**99.996700%**	**100.000000%**
Laplace Distribution					
upper tail	0.0257930	0.0300470	0.4727320	0.9788360	1.0000000
lower tail	0.0257400	0.0211880	0.0000750	0.0000000	0.0000000
	5.153300%	**5.123500%**	**47.280700%**	**97.883600%**	**100.000000%**
T Distribution (3 df)					
upper tail	0.0260270	0.0318350	0.3650470	0.9281790	1.0000000
lower tail	0.0260900	0.0240100	0.0003670	0.0000000	0.0000000
	5.211700%	**5.584500%**	**36.541400%**	**92.817900%**	**100.000000%**

Table 6: Type II Robustness at 1% Significance Level - Sample Size 5, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 5					
Normal Distribution					
upper tail	0.0048490	0.0071020	0.0299030	0.0744390	0.3089990
lower tail	0.0050910	0.0035110	0.0006640	0.0001990	0.0000180
	0.994000%	**1.061300%**	**3.056700%**	**7.463800%**	**30.901700%**
Chi Square Distribution (1 df)					
upper tail	0.0191180	0.0282330	0.0914230	0.1651120	0.3847360
lower tail	0.0015440	0.0012940	0.0004770	0.0002070	0.0000240
	2.066200%	**2.952700%**	**9.190000%**	**16.531900%**	**38.476000%**
Chi Square Distribution (2 df)					
upper tail	0.0101980	0.0167710	0.0704350	0.1441790	0.3879620
lower tail	0.0026740	0.0017970	0.0004220	0.0001530	0.0000150
	1.287200%	**1.856800%**	**7.085700%**	**14.433200%**	**38.797700%**
Laplace Distribution					
upper tail	0.0052740	0.0053170	0.0122650	0.0302820	0.1506480
lower tail	0.0052860	0.0048340	0.0019950	0.0007220	0.0000770
	1.056000%	**1.015100%**	**1.426000%**	**3.100400%**	**15.072500%**
T Distribution (3 df)					
upper tail	0.0053780	0.0069700	0.0170750	0.0415110	0.1773280
lower tail	0.0053490	0.0062990	0.0023560	0.0007560	0.0000940
	1.072700%	**1.326900%**	**1.943100%**	**4.226700%**	**17.742200%**

Table 7: Type II Robustness at 1% Significance Level - Sample Size 15, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 15					
Normal Distribution					
upper tail	0.0049830	0.0124230	0.2495410	0.6838760	0.9951750
lower tail	0.0048610	0.0018410	0.0000140	0.0000010	0.0000000
	0.984400%	**1.426400%**	**24.955500%**	**68.387700%**	**99.517500%**
Chi Square Distribution (1 df)					
upper tail	0.0215990	0.0479160	0.3132740	0.6126910	0.9676050
lower tail	0.0001700	0.0000920	0.0000040	0.0000020	0.0000000
	2.176900%	**4.800800%**	**31.327800%**	**61.269300%**	**96.760500%**
Chi Square Distribution (2 df)					
upper tail	0.0136230	0.0310830	0.3043050	0.6605370	0.9894410
lower tail	0.0008610	0.0003340	0.0000000	0.0000000	0.0000000
	1.448400%	**3.141700%**	**30.430500%**	**66.053700%**	**98.944100%**
Laplace Distribution					
upper tail	0.0058750	0.0061350	0.0441910	0.2282590	0.9115920
lower tail	0.0059130	0.0050820	0.0004160	0.0000100	0.0000000
	1.178800%	**1.121700%**	**4.460700%**	**22.826900%**	**91.159200%**
T Distribution (3 df)					
upper tail	0.0067060	0.0104600	0.0542500	0.2095350	0.8303450
lower tail	0.0067220	0.0089300	0.0010170	0.0000330	0.0000000
	1.342800%	**1.939000%**	**5.526700%**	**20.956800%**	**83.034500%**

Table 8: Type II Robustness at 1% Significance Level - Sample Size 30, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 30					
Normal Distribution					
upper tail	0.0049920	0.0192580	0.6181510	0.9757390	1.0000000
lower tail	0.0050110	0.0010360	0.0000000	0.0000000	0.0000000
	1.000300%	**2.029400%**	**61.815100%**	**97.573900%**	**100.000000%**
Chi Square Distribution (1 df)					
upper tail	0.0191230	0.0586690	0.5574230	0.9041380	0.9998830
lower tail	0.0001320	0.0000280	0.0000000	0.0000000	0.0000000
	1.925500%	**5.869700%**	**55.742300%**	**90.413800%**	**99.988300%**
Chi Square Distribution (2 df)					
upper tail	0.0125100	0.0399440	0.5953250	0.9505850	0.9999930
lower tail	0.0008740	0.0001560	0.0000000	0.0000000	0.0000000
	1.338400%	**4.010000%**	**59.532500%**	**95.058500%**	**99.999300%**
Laplace Distribution					
upper tail	0.0057360	0.0064000	0.1015300	0.5655700	0.9993810
lower tail	0.0057950	0.0047160	0.0000760	0.0000000	0.0000000
	1.153100%	**1.111600%**	**10.160600%**	**56.557000%**	**99.938100%**
T Distribution (3 df)					
upper tail	0.0068580	0.0106110	0.0946680	0.4420730	0.9938220
lower tail	0.0069110	0.0085830	0.0003660	0.0000020	0.0000000
	1.376900%	**1.919400%**	**9.503400%**	**44.207500%**	**99.382200%**

Table 9: Type II Robustness at 1% Significance Level - Sample Size 60, Normal and Nonnormal Distributions, and Rho

	0.0	0.1	0.5	0.7	0.9
Sample Size: 60					
Normal Distribution					
upper tail	0.0049720	0.0342150	0.9448120	0.9999630	1.0000000
lower tail	0.0049760	0.0004730	0.0000000	0.0000000	0.0000000
	0.994800%	**3.468800%**	**94.481200%**	**99.996300%**	**100.000000%**
Chi Square Distribution (1 df)					
upper tail	0.0164970	0.0762820	0.8497470	0.9969000	1.0000000
lower tail	0.0001910	0.0000260	0.0000000	0.0000000	0.0000000
	1.668800%	**7.630800%**	**84.974700%**	**99.690000%**	**100.000000%**
Chi Square Distribution (2 df)					
upper tail	0.0112650	0.0562040	0.9068780	0.9996710	1.0000000
lower tail	0.0011430	0.0000720	0.0000000	0.0000000	0.0000000
	1.240800%	**5.627600%**	**90.687800%**	**99.967100%**	**100.000000%**
Laplace Distribution					
upper tail	0.0055790	0.0066360	0.2443810	0.9187420	1.0000000
lower tail	0.0056430	0.0044160	0.0000060	0.0000000	0.0000000
	1.122200%	**1.105200%**	**24.438700%**	**91.874200%**	**100.000000%**
T Distribution (3 df)					
upper tail	0.0068220	0.0101950	0.1837760	0.7927040	0.9999960
lower tail	0.0069060	0.0076730	0.0000830	0.0000000	0.0000000
	1.372800%	**1.786800%**	**18.385900%**	**79.270400%**	**99.999600%**

Figure 5: Normal Distribution - Sample Size 5 at 5% Significance Level

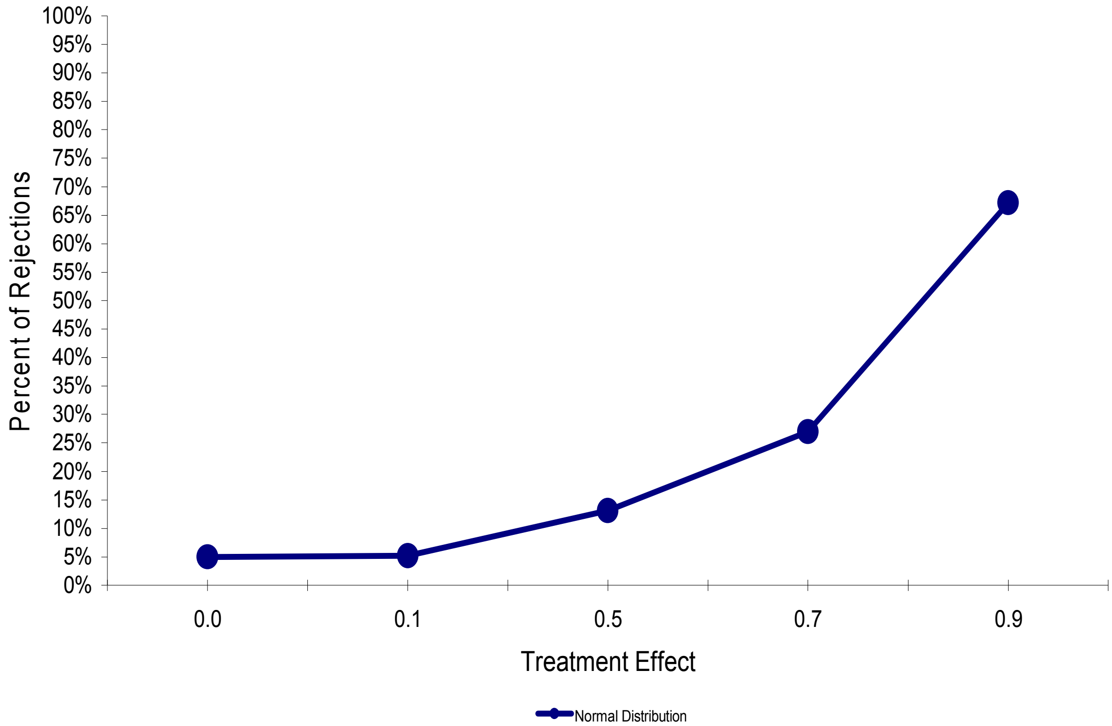


Figure 6: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 5 at 5% Significance Level

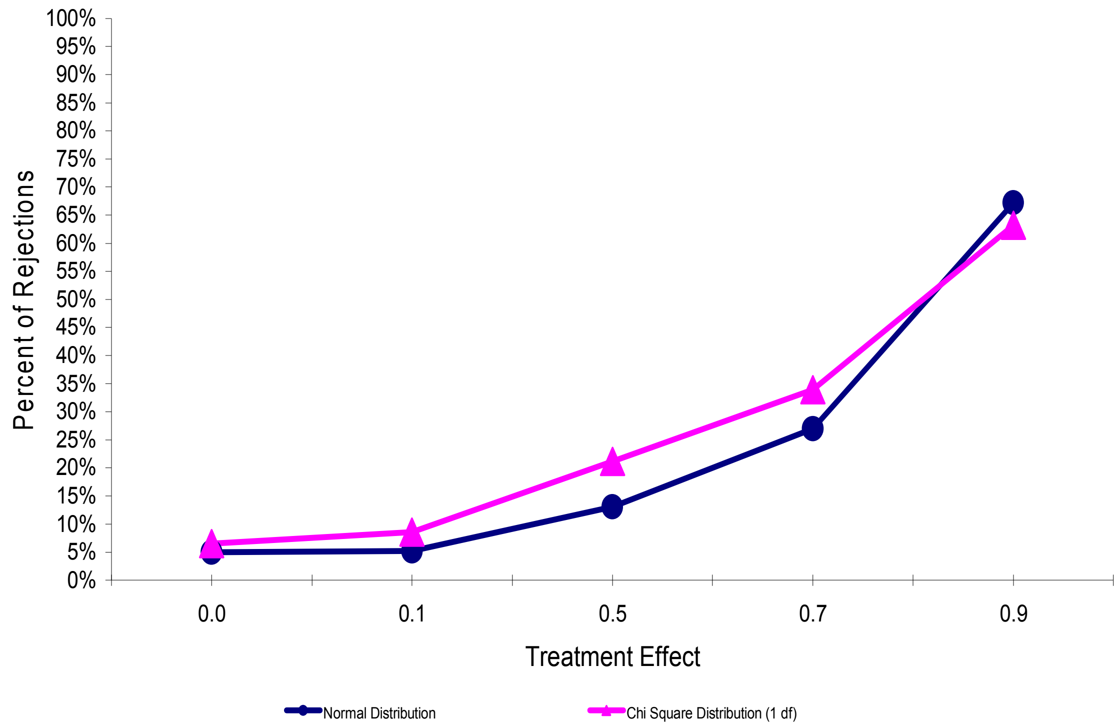


Figure 7: Normal Distribution Chi Square Distribution (2 DF) - Sample Size 5 at 5% Significance Level

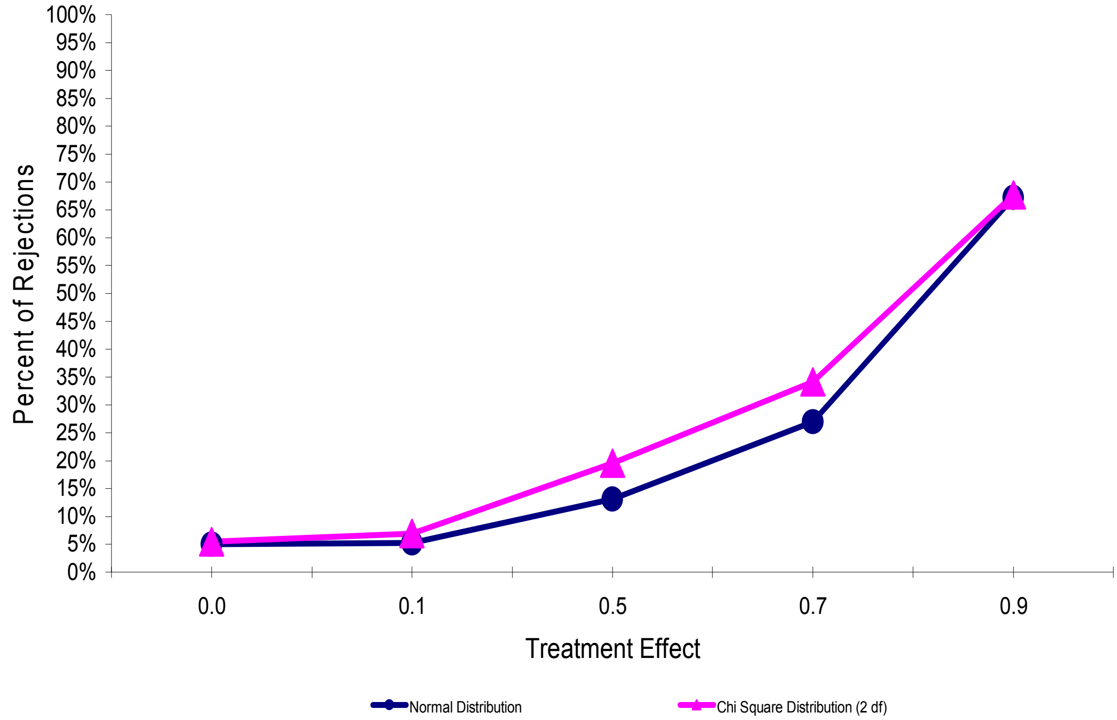


Figure 8: Normal Distribution and Laplace Distribution - Sample Size 5 at 5% Significance Level

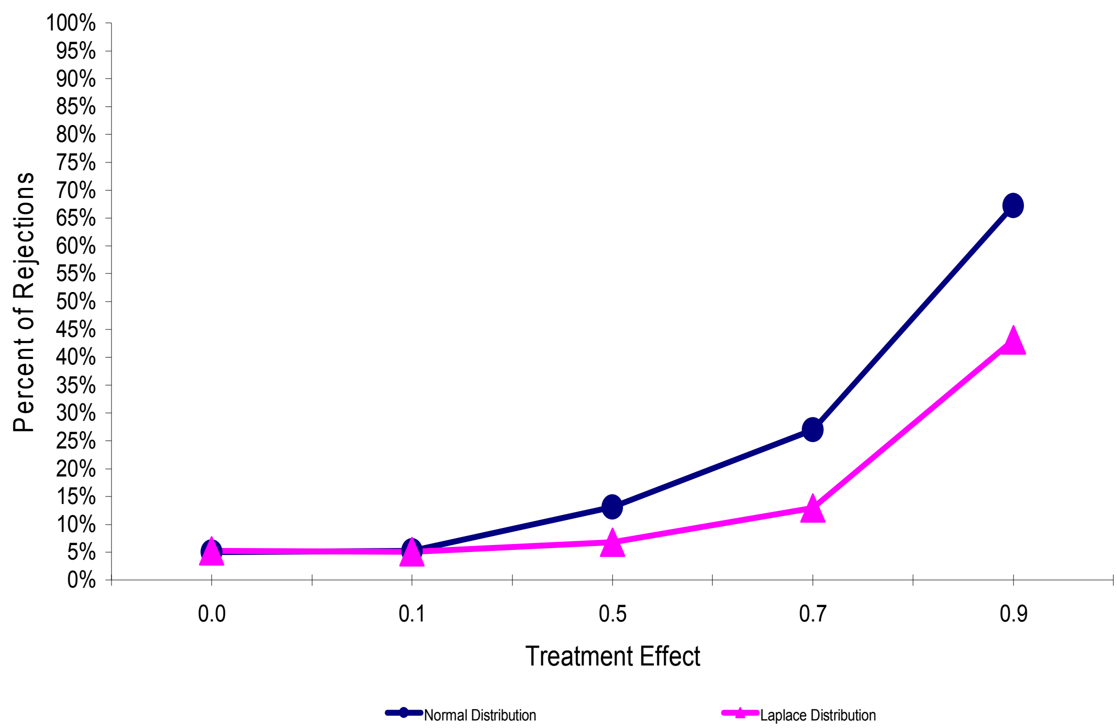


Figure 9: Normal Distribution and T Distribution (3 DF) - Sample Size 5 at 5% Significance Level

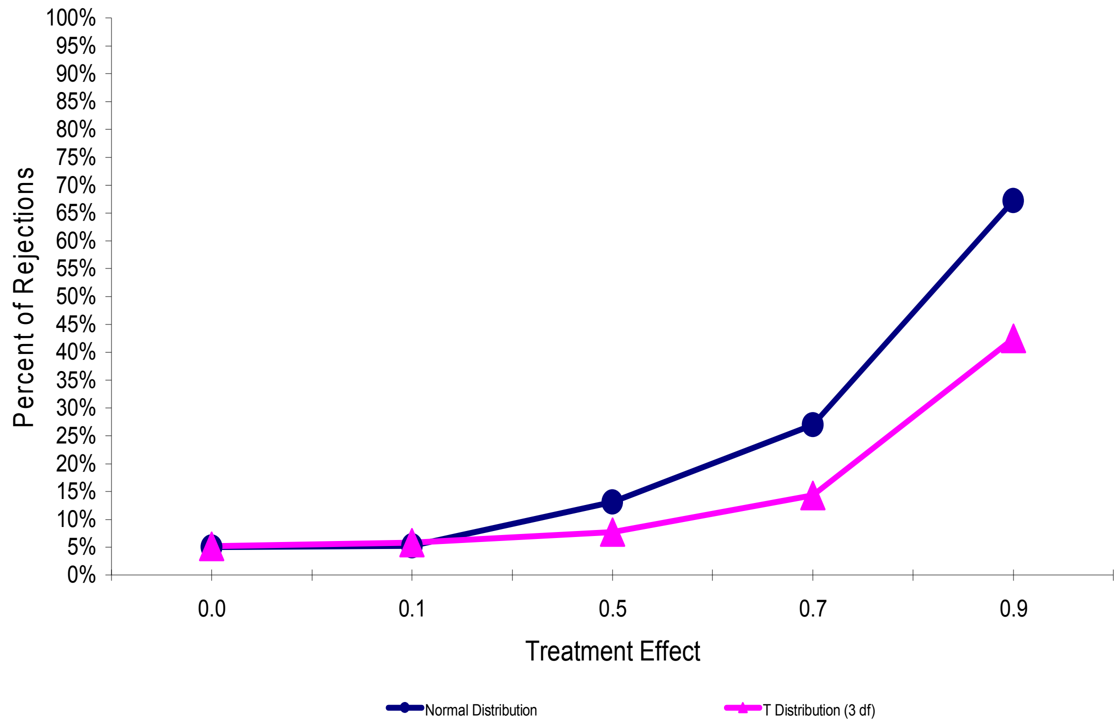


Figure 10: All Distributions - Sample Size 5 at 5% Significance Level

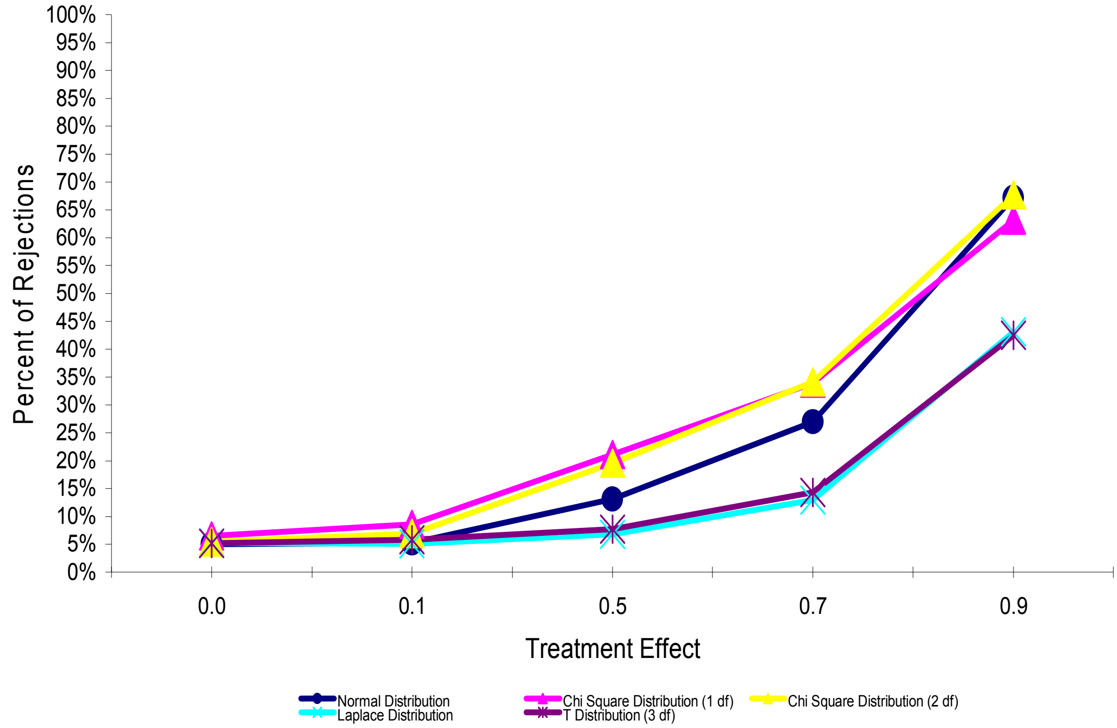


Figure 11: Normal Distribution - Sample Size 15 at 5% Significance Level

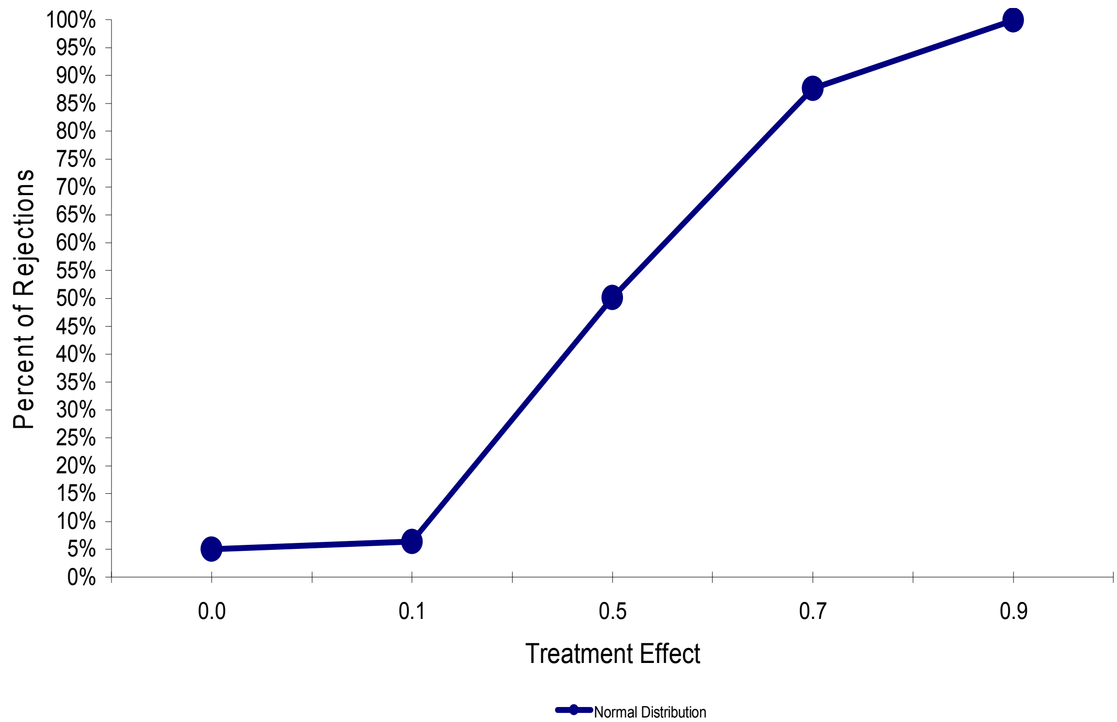


Figure 12: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 15 at 5% Significance Level

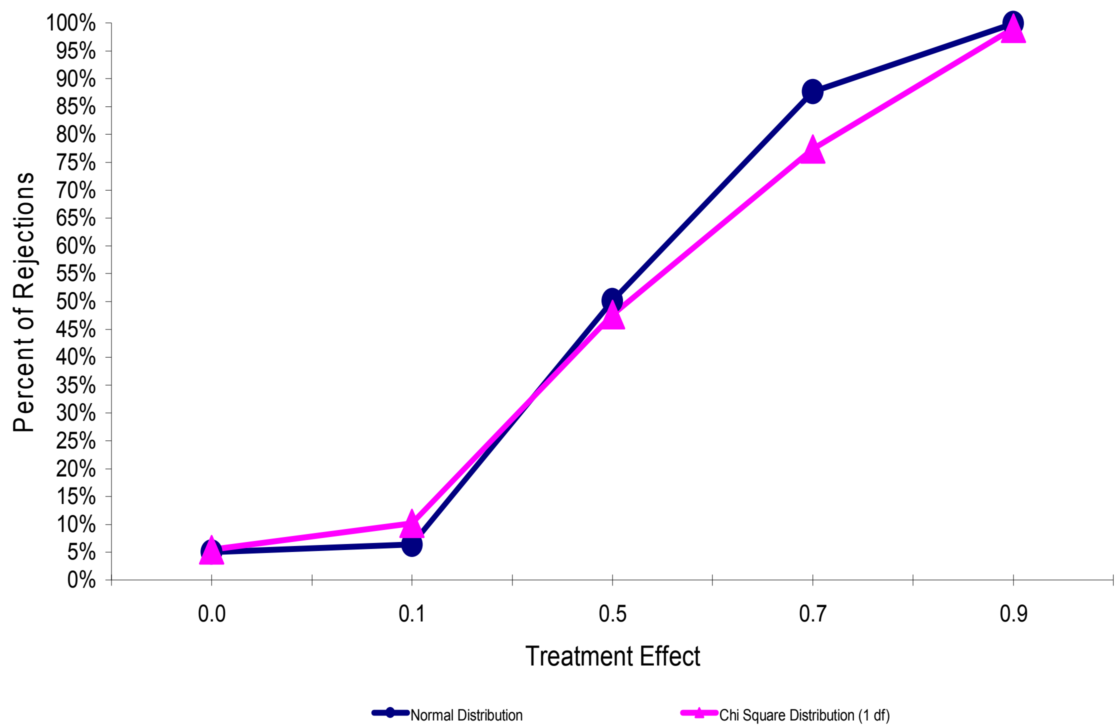


Figure 13: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 15 at 5% Significance Level

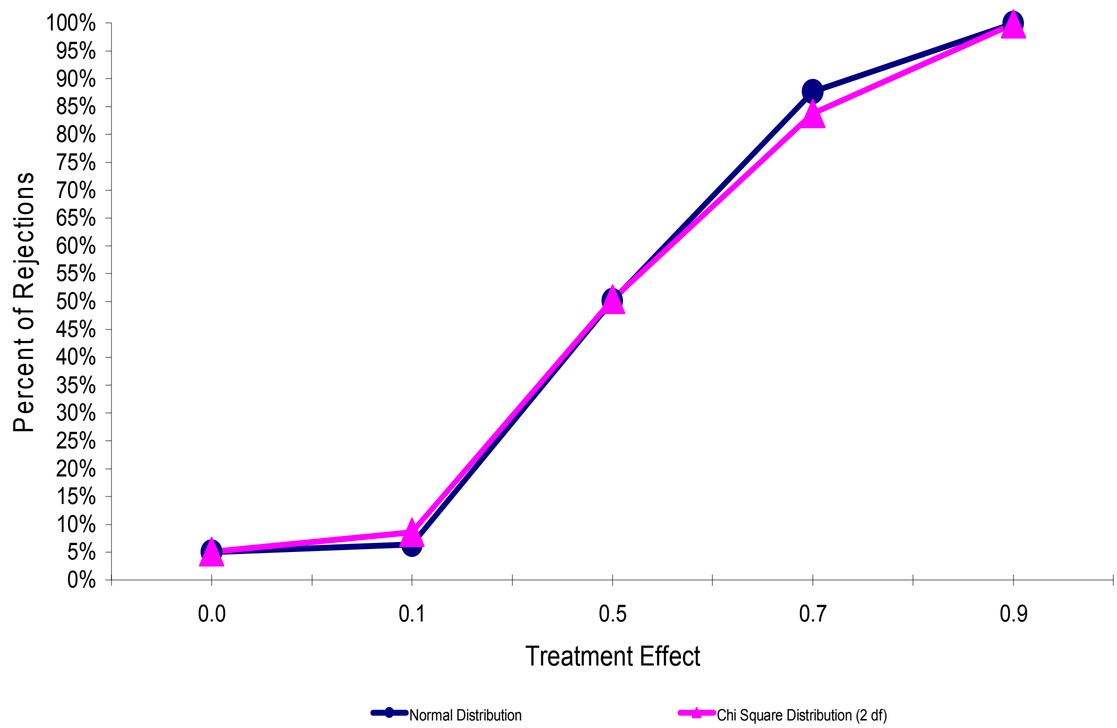


Figure 14: Normal Distribution and Laplace Distribution - Sample Size 15 at 5% Significance Level

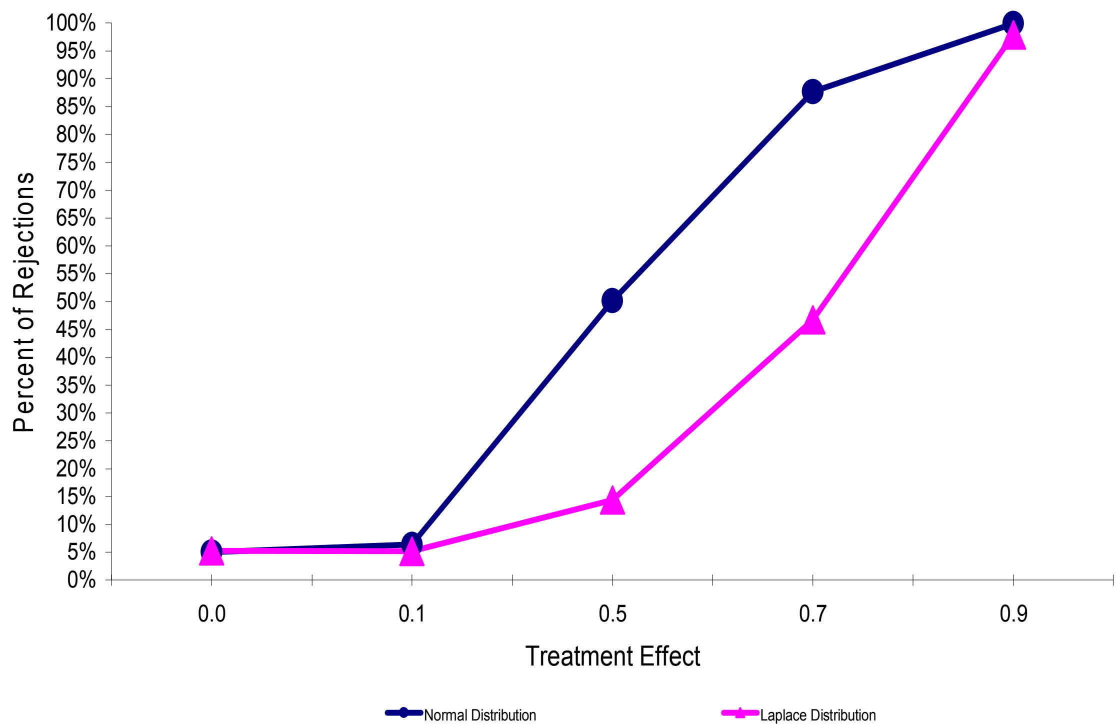


Figure 15: Normal Distribution and T Distribution (3 DF) - Sample Size 15 at 5% Significance Level

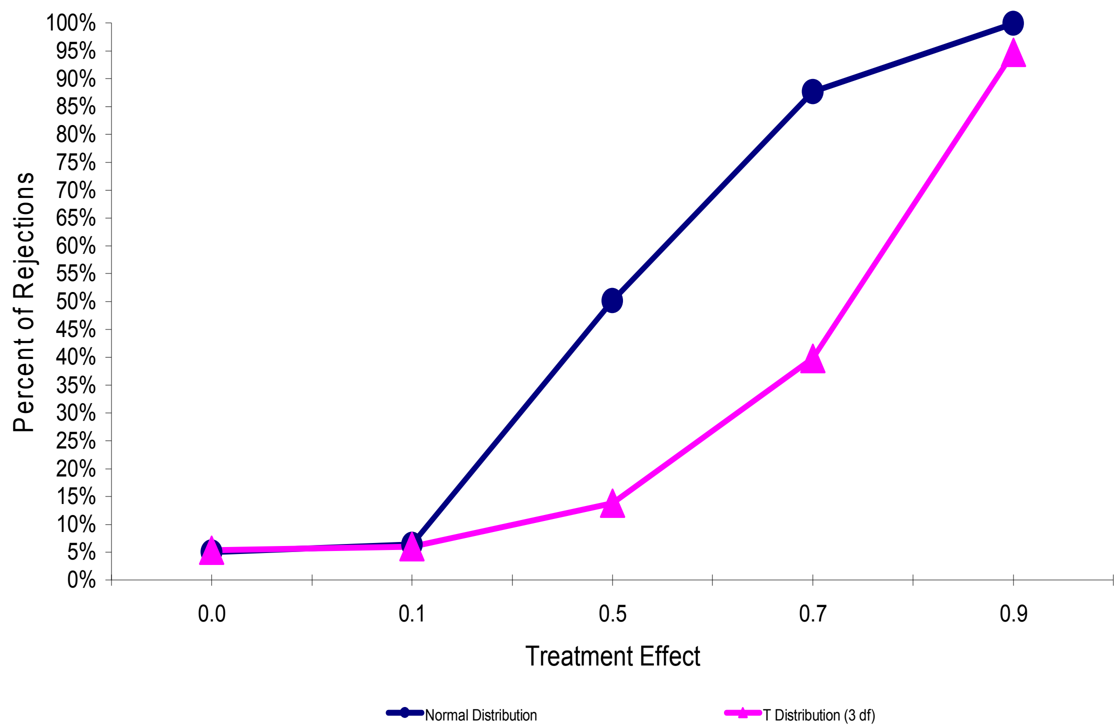


Figure 16: All Distributions - Sample Size 15 at 5% Significance Level

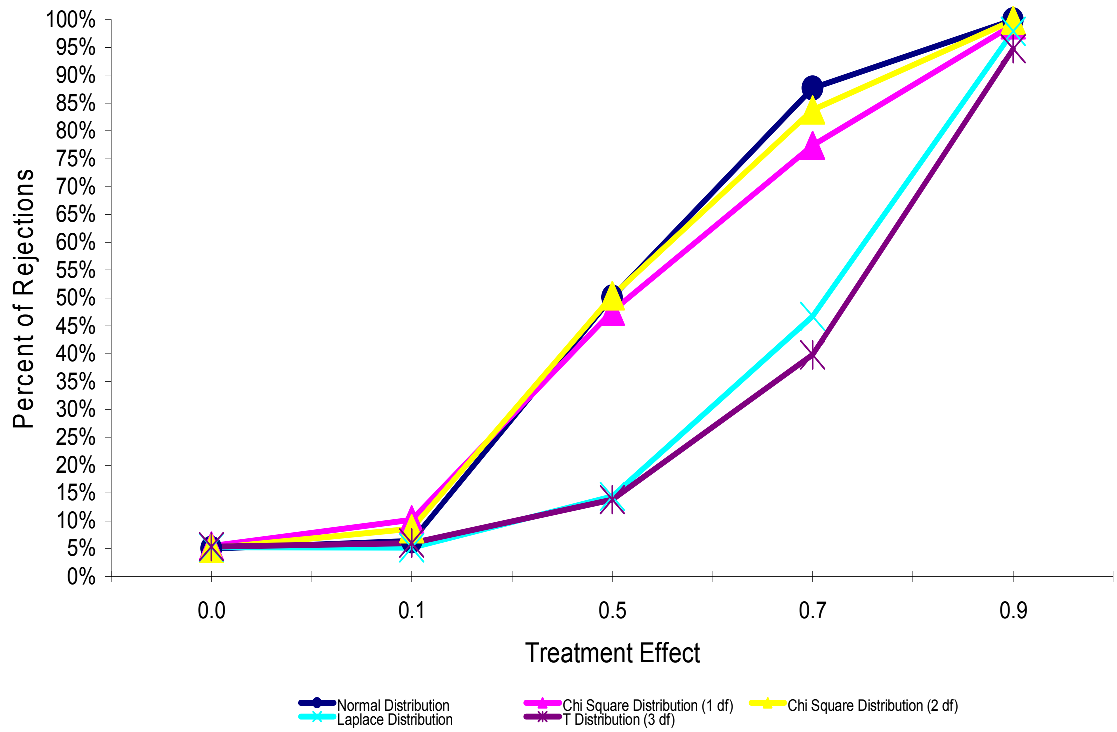


Figure 17: Normal Distribution - Sample Size 30 at 5% Significance Level

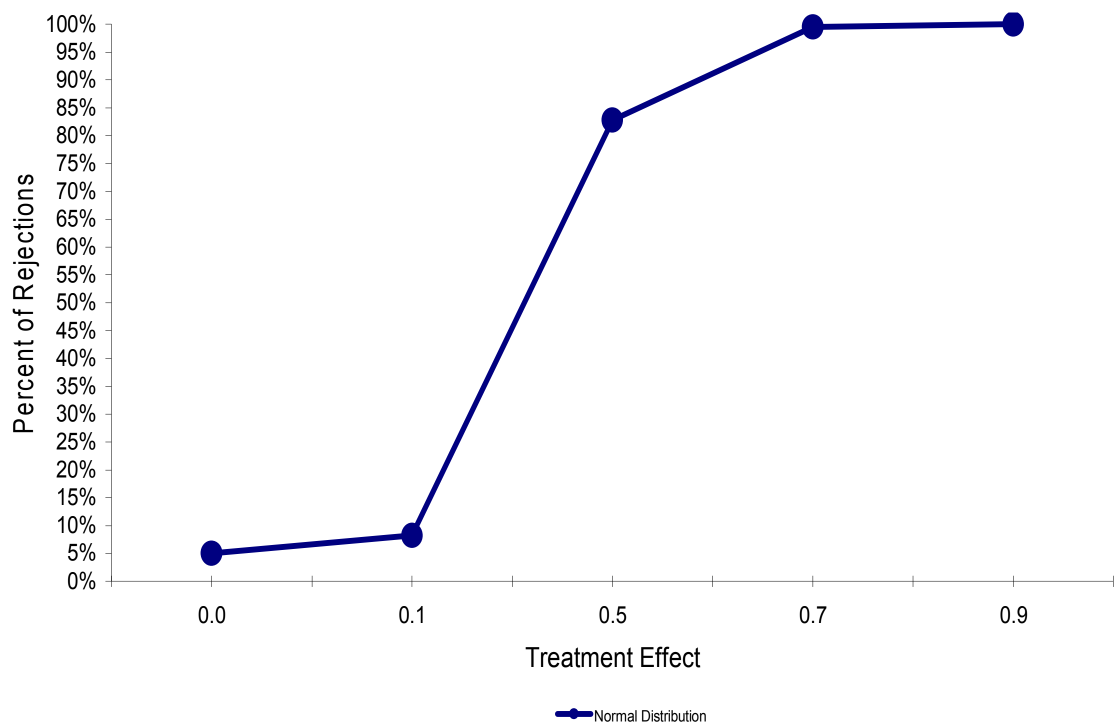


Figure 18: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 30 at 5% Significance Level

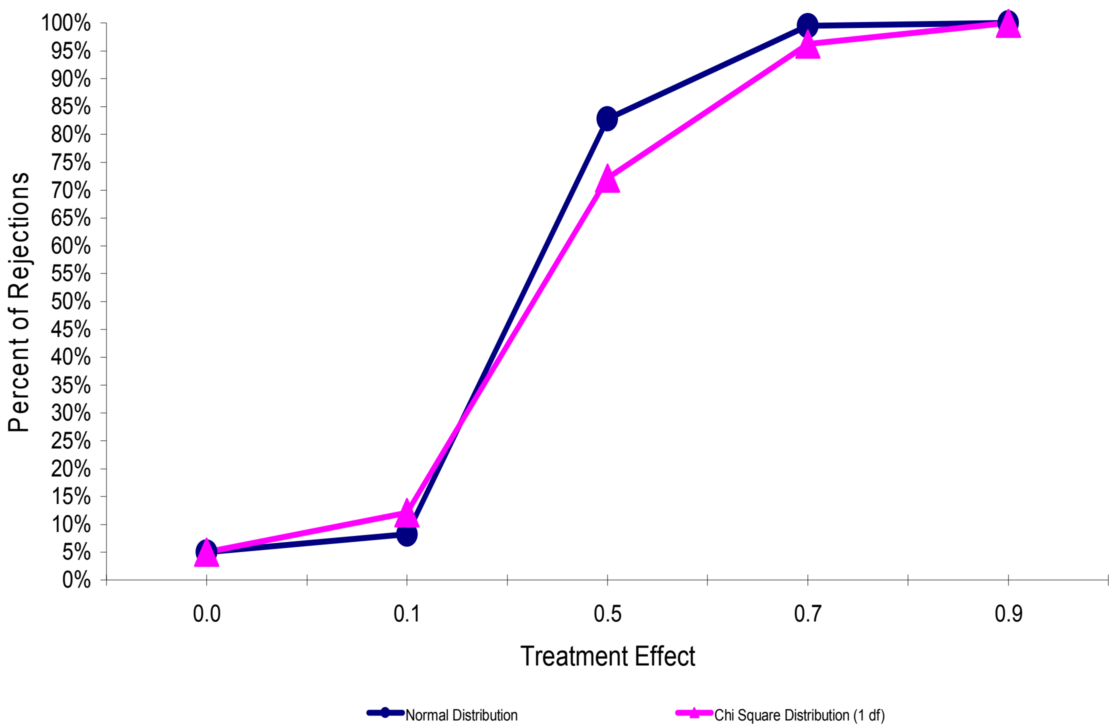


Figure 19: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 30 at 5% Significance Level

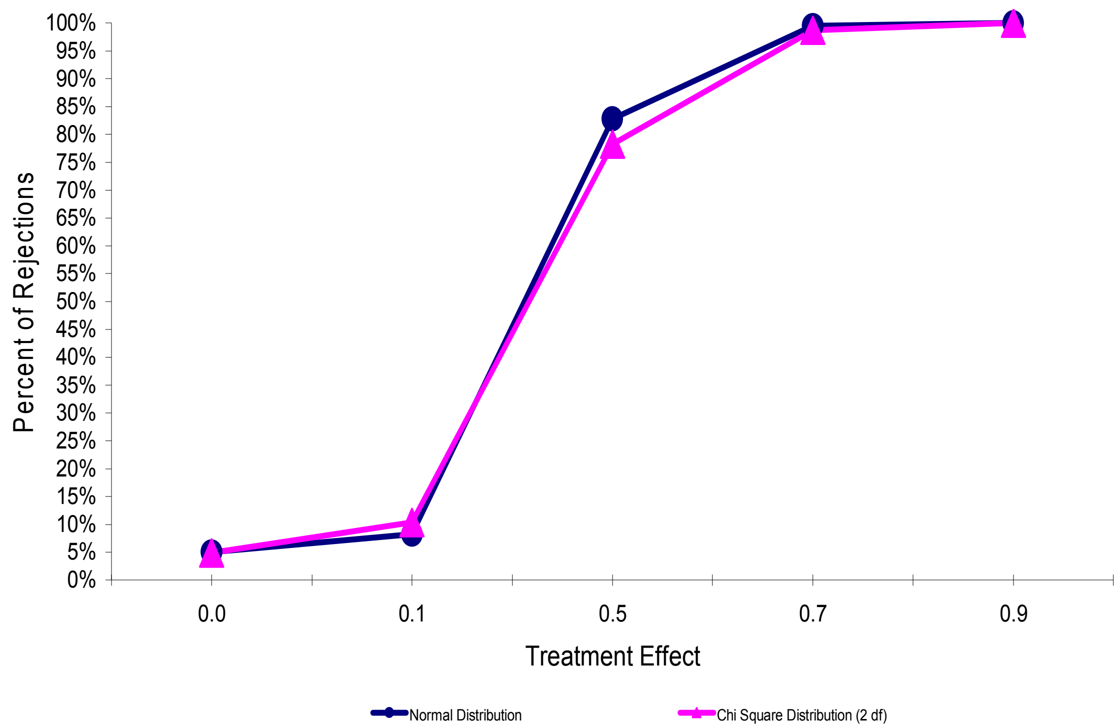


Figure 20: Normal Distribution and Laplace Distribution - Sample Size 30 at 5% Significance Level

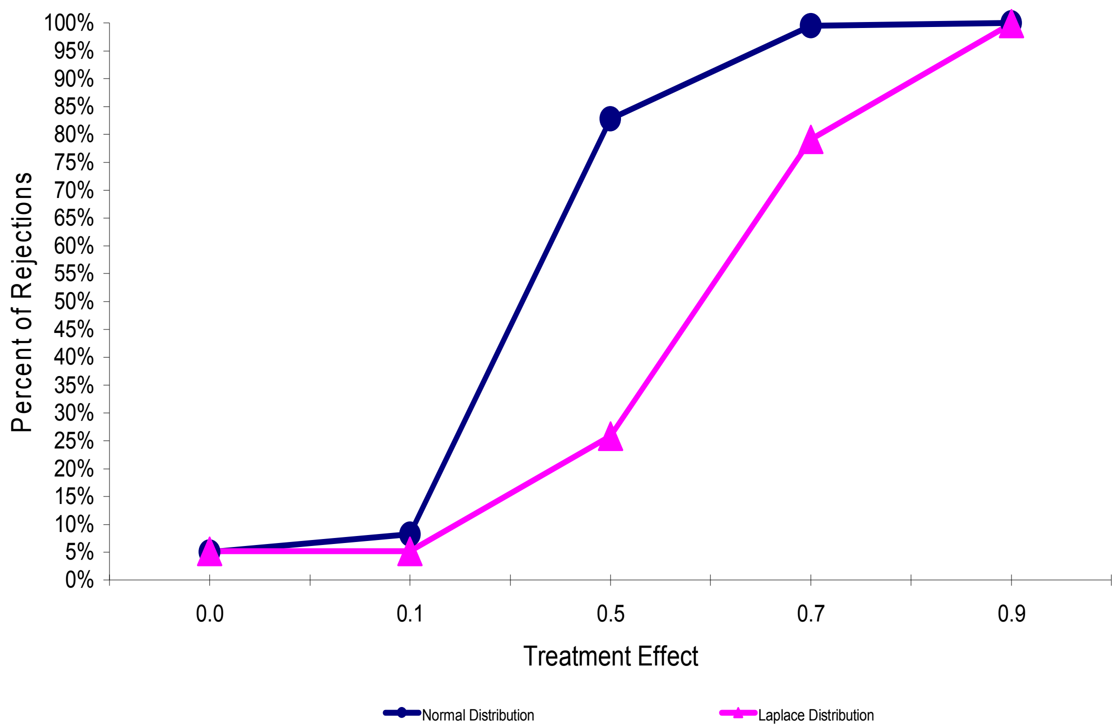


Figure 21: Normal Distribution and T Distribution (3 DF) - Sample Size 30 at 5% Significance Level

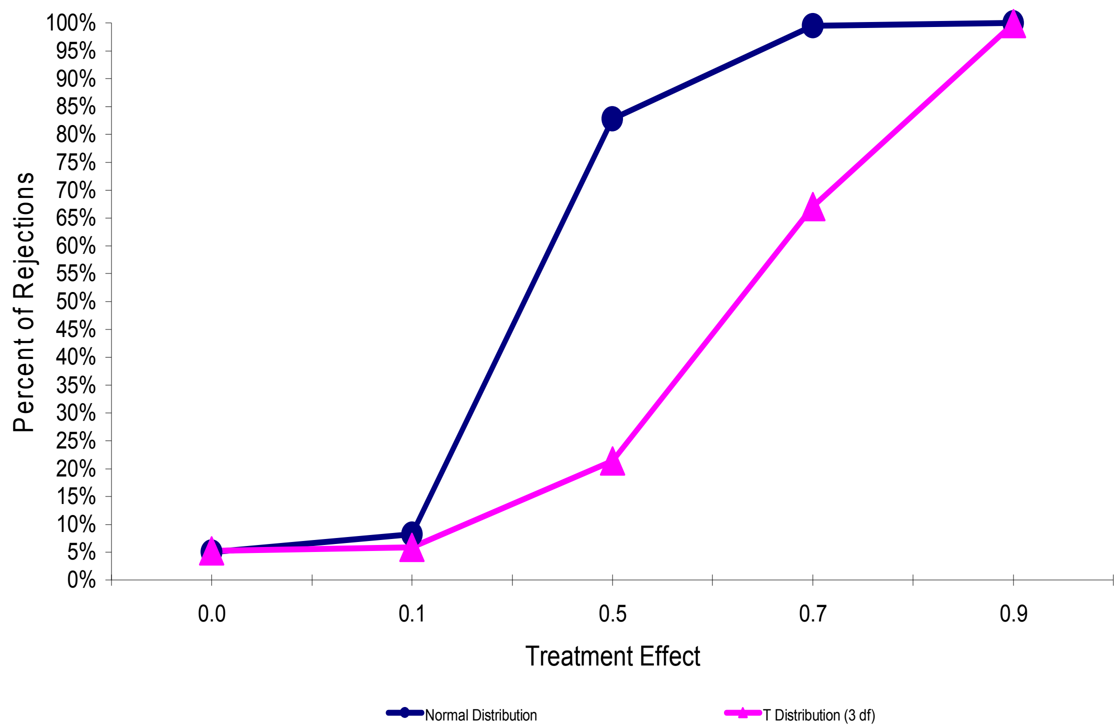


Figure 22: All Distributions - Sample Size 30 at 5% Significance Level

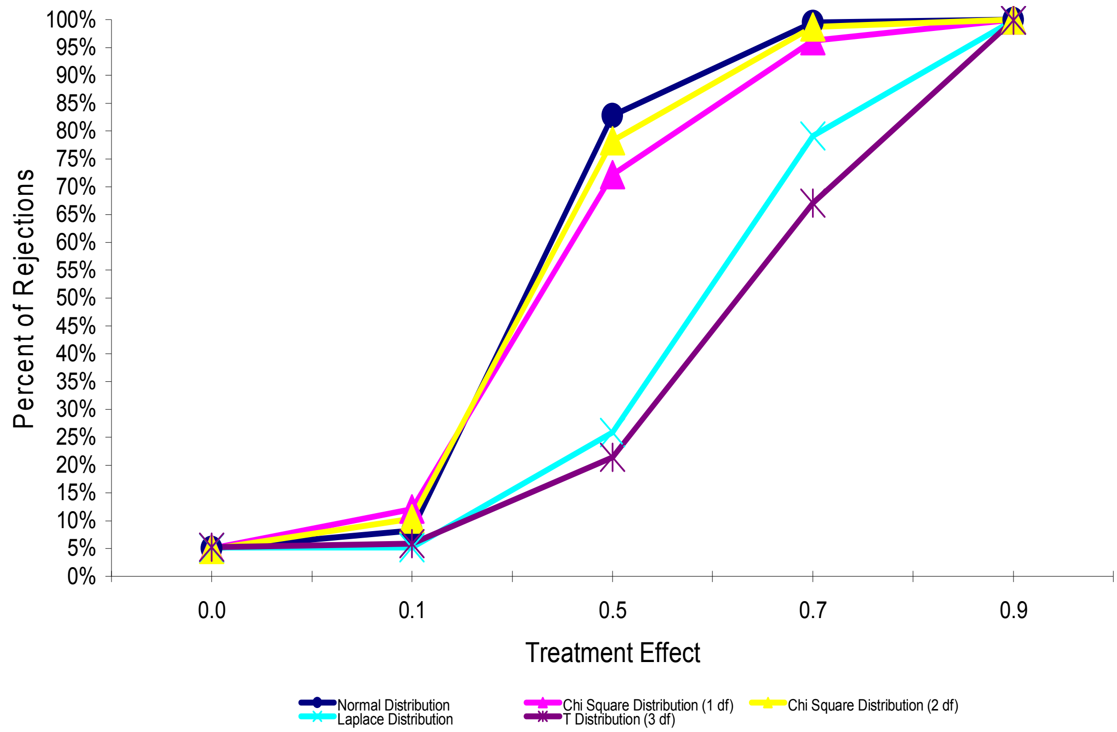


Figure 23: Normal Distribution - Sample Size 60 at 5% Significance Level

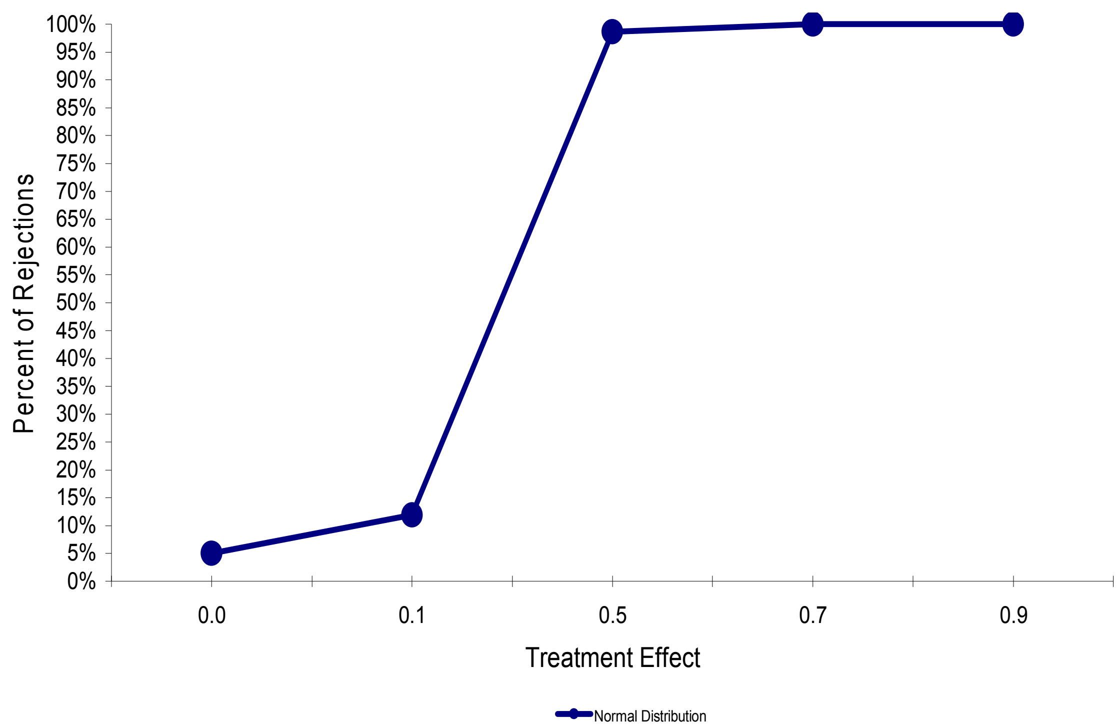


Figure 24: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 60 at 5% Significance Level

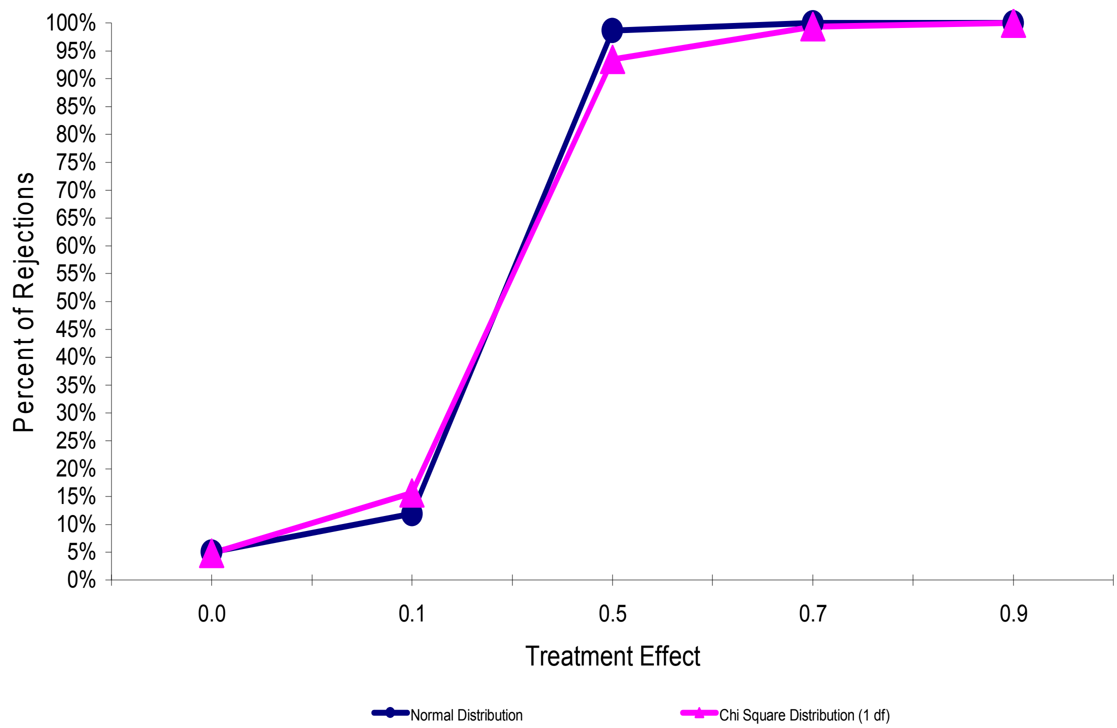


Figure 25: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 60 at 5% Significance Level

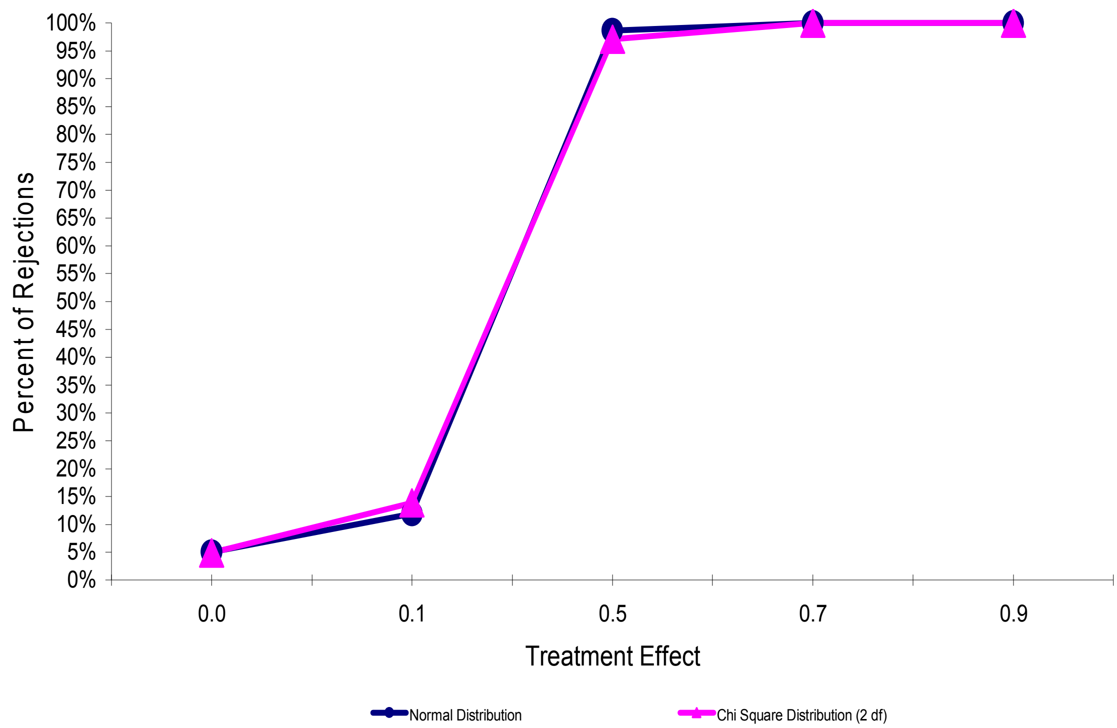


Figure 26: Normal Distribution and Laplace Distribution - Sample Size 60 at 5% Significance Level

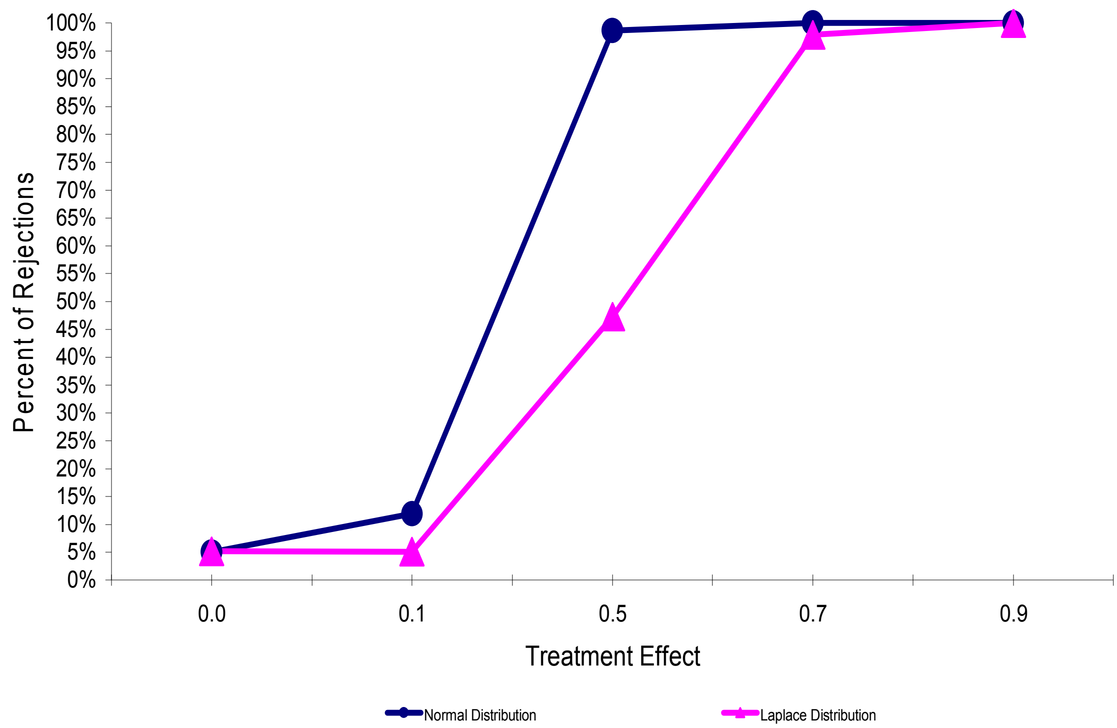


Figure 27: Normal Distribution and T Distribution (3 DF) - Sample Size 60 at 5% Significance Level

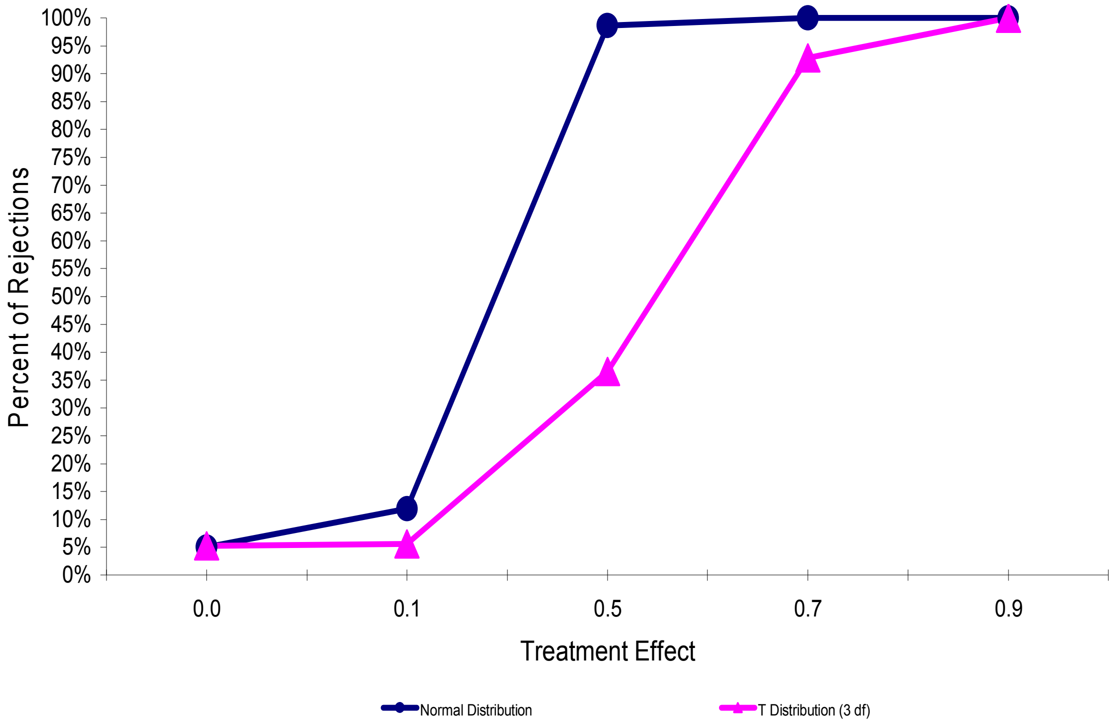


Figure 28: All Distributions - Sample Size 60 at 5% Significance Level

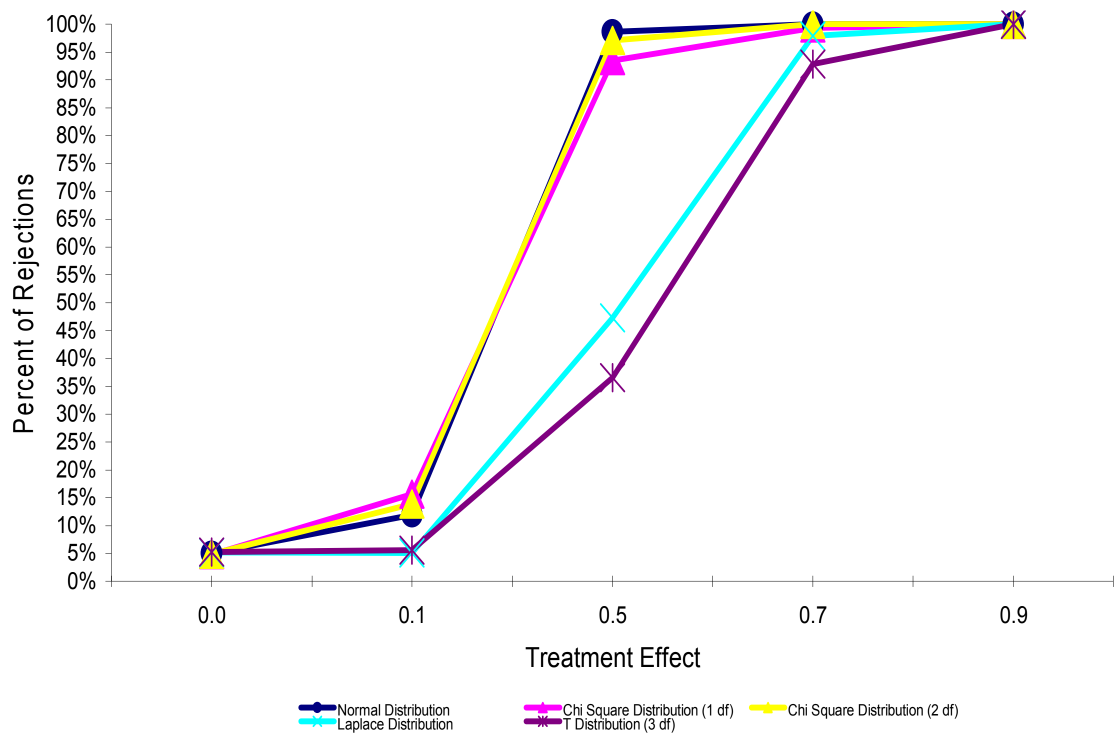


Figure 29: Normal Distribution - Sample Size 5 at 1% Significance Level

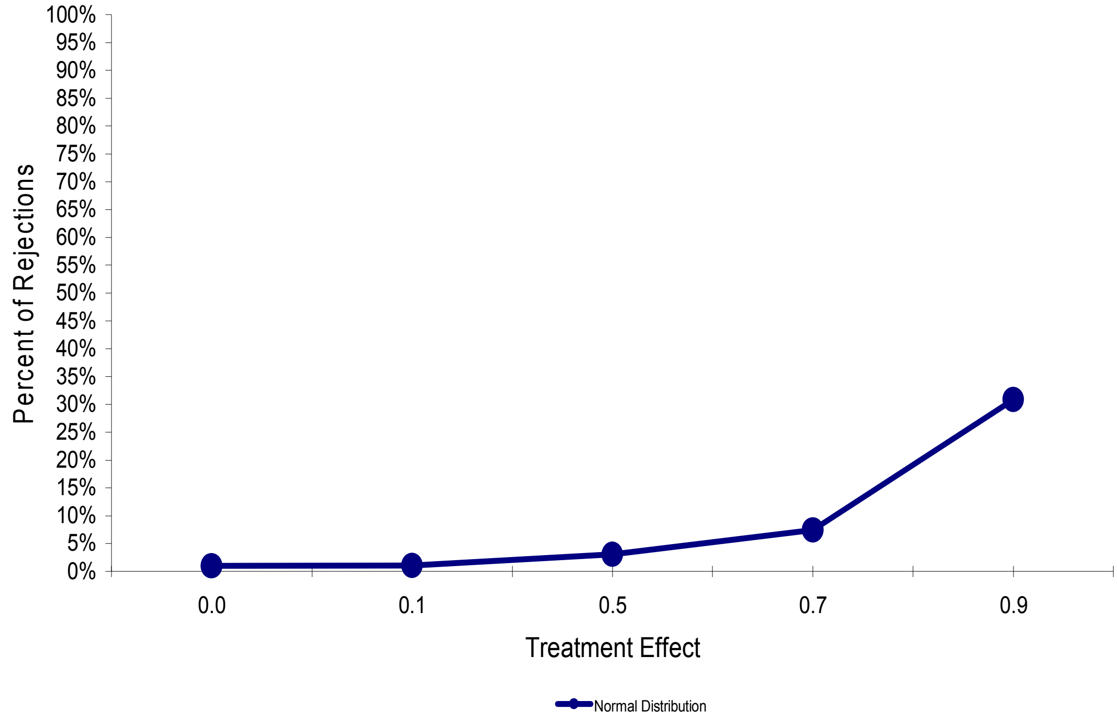


Figure 30: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 5 at 1% Significance Level

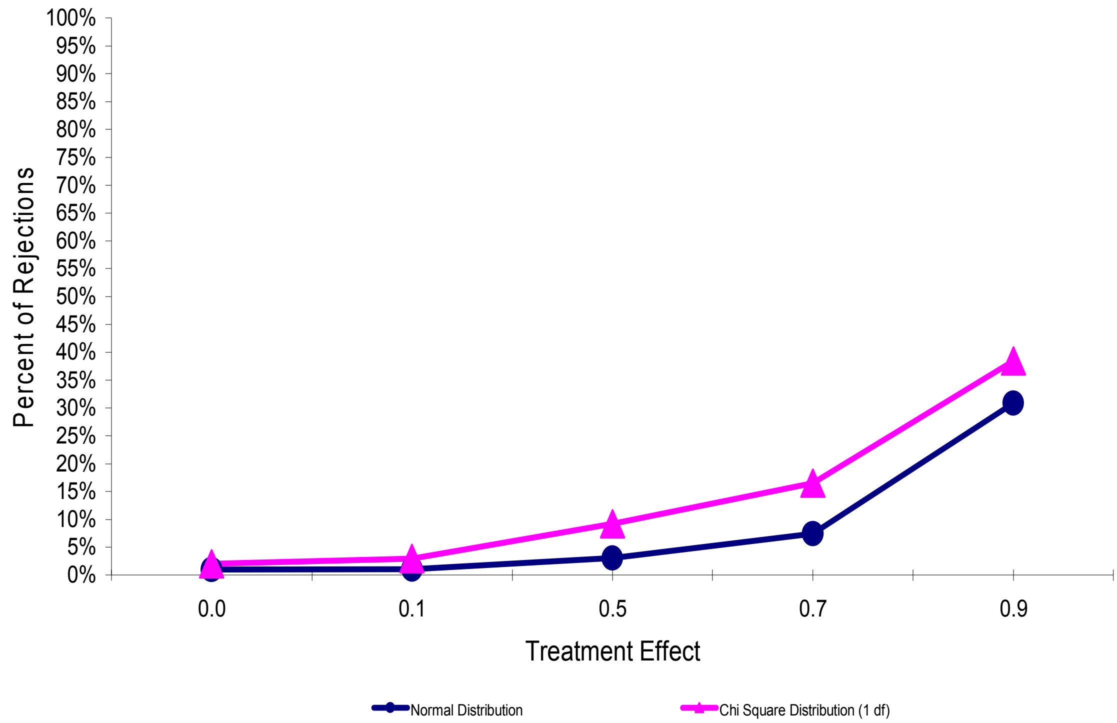


Figure 31: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 5 at 1% Significance Level

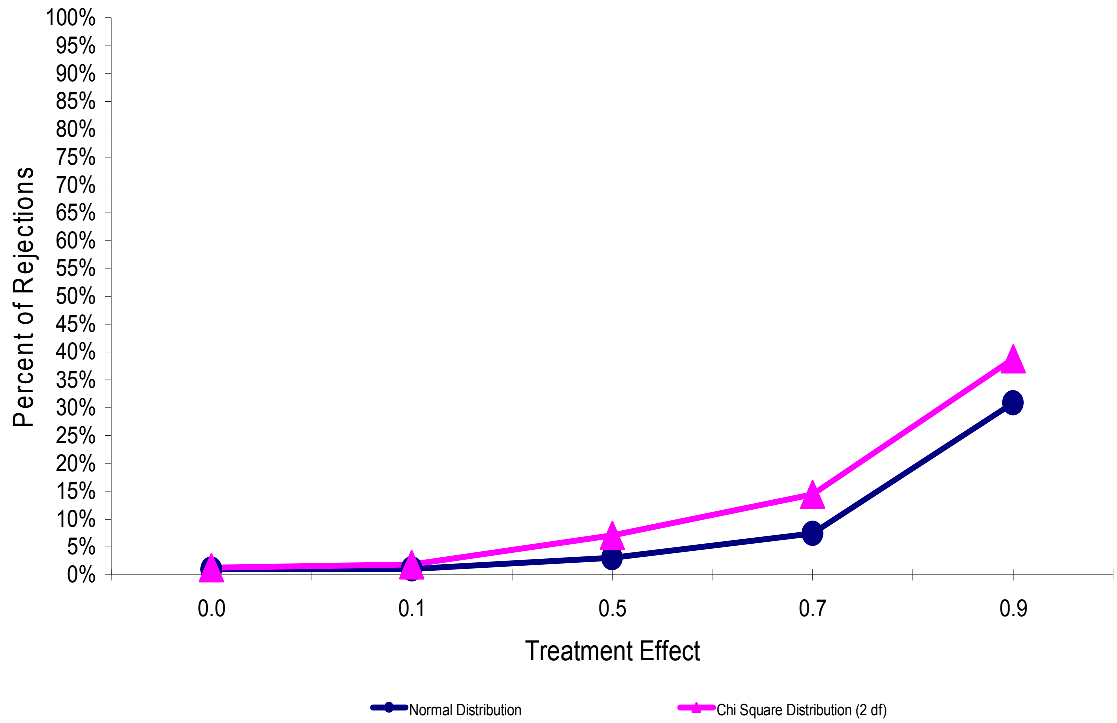


Figure 32: Normal Distribution and Laplace Distribution - Sample Size 5 at 1% Significance Level

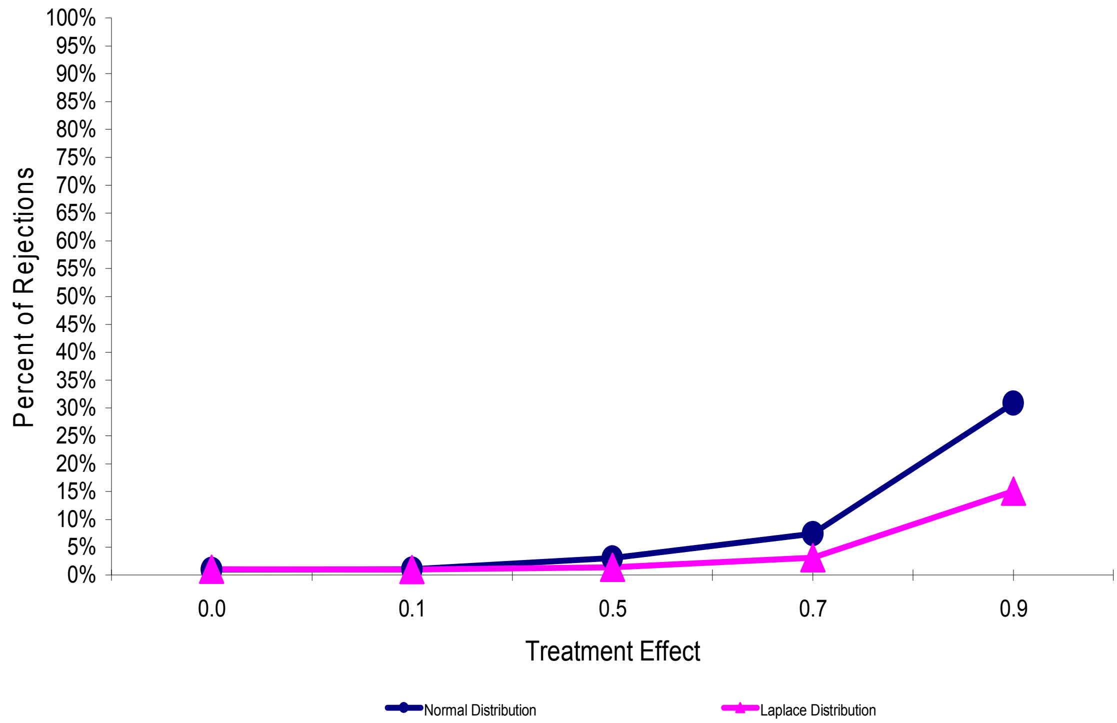


Figure 33: Normal Distribution and T Distribution (3 DF) - Sample Size 5 at 1% Significance Level

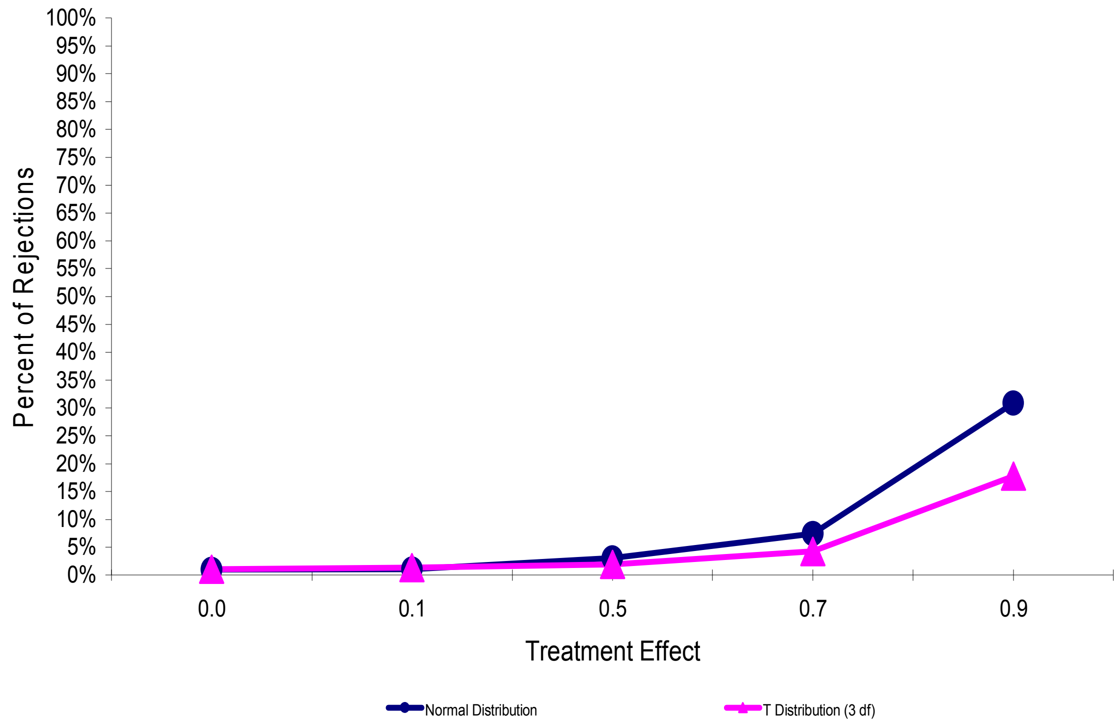


Figure 34: All Distributions - Sample Size 5 at 1% Significance Level

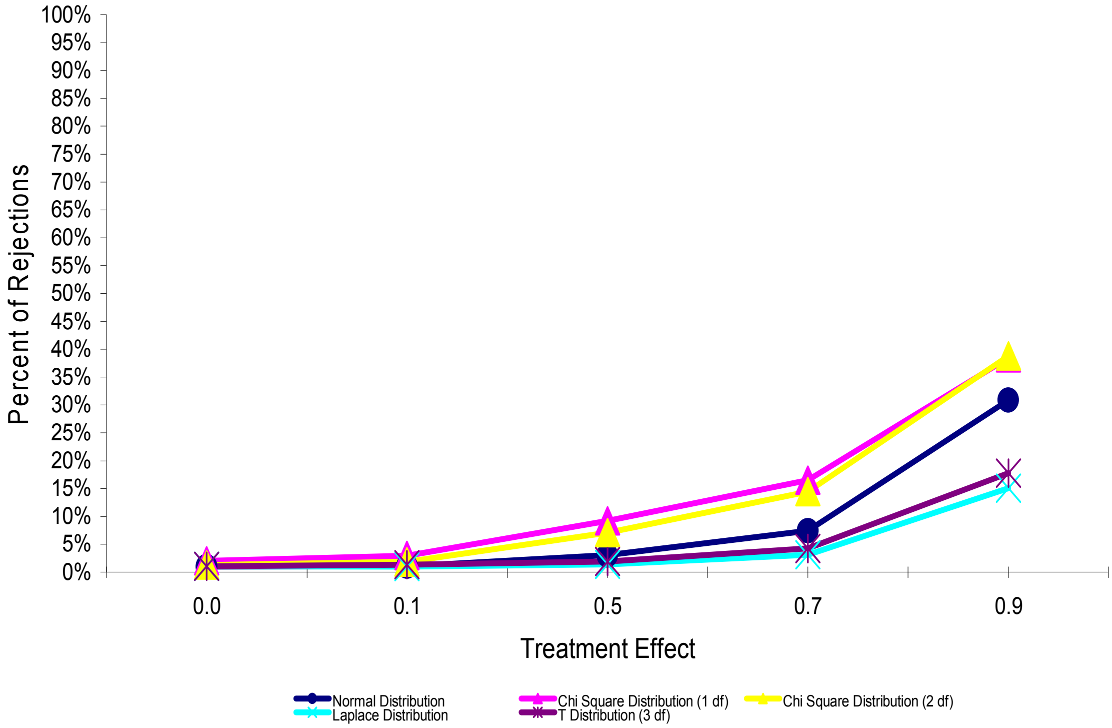


Figure 35: Normal Distribution - Sample Size 15 at 1% Significance Level

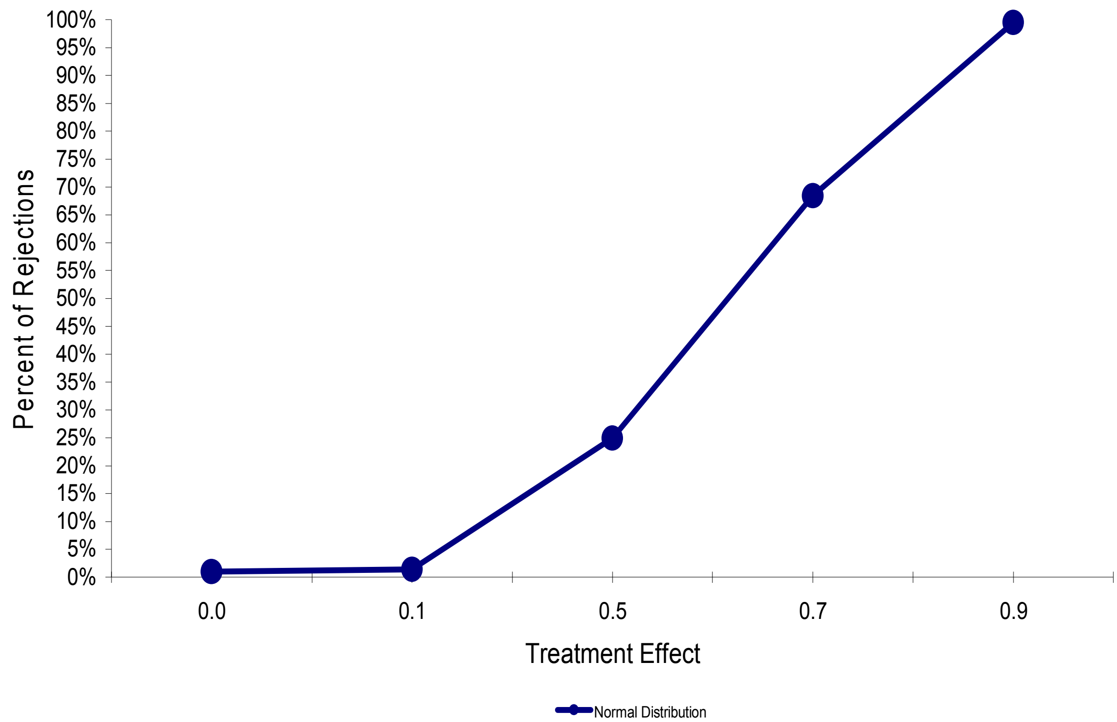


Figure 36: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 15 at 1% Significance Level

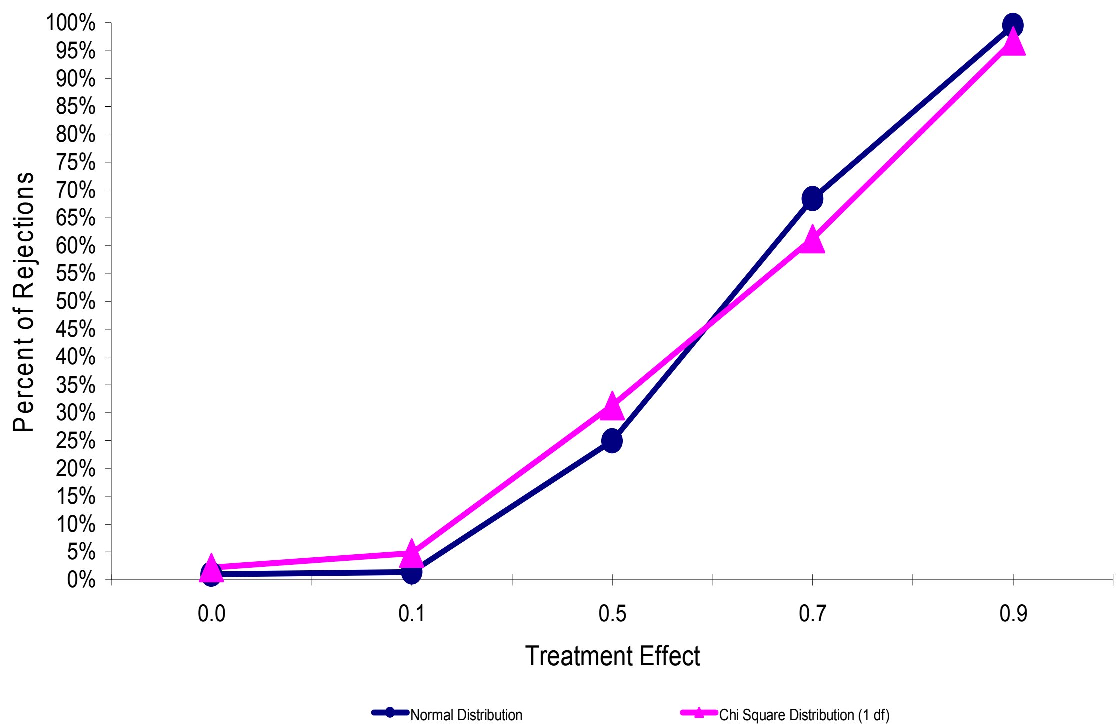


Figure 37: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 15 at 1% Significance Level

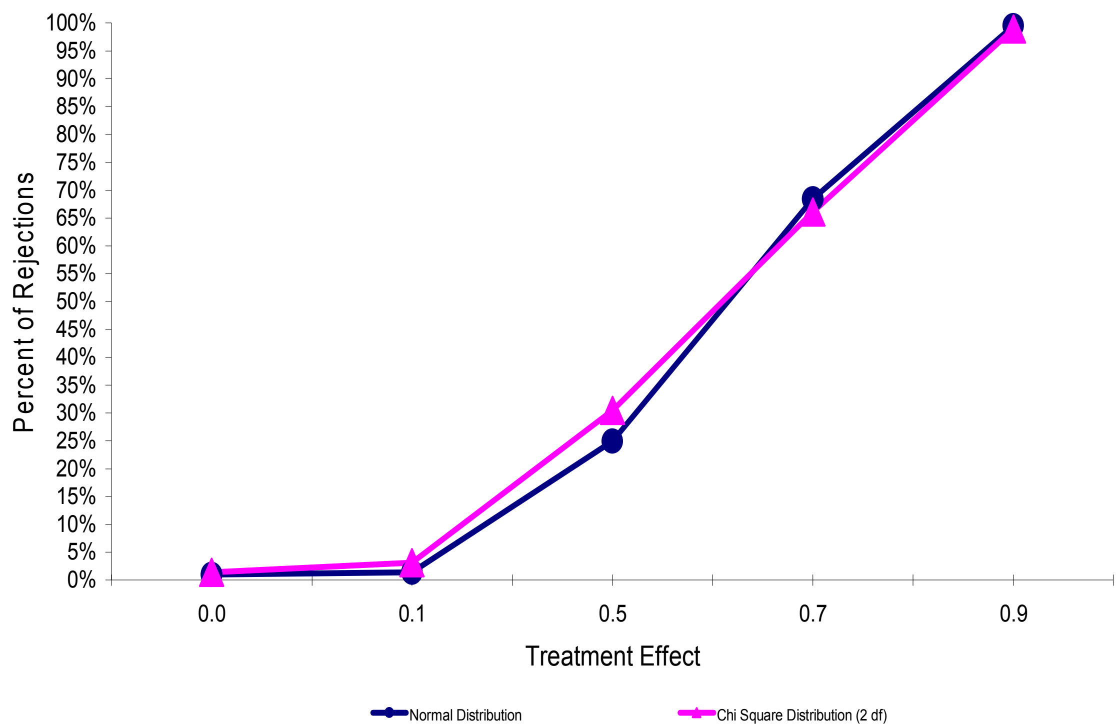


Figure 38: Normal Distribution and Laplace Distribution - Sample Size 15 at 1% Significance Level

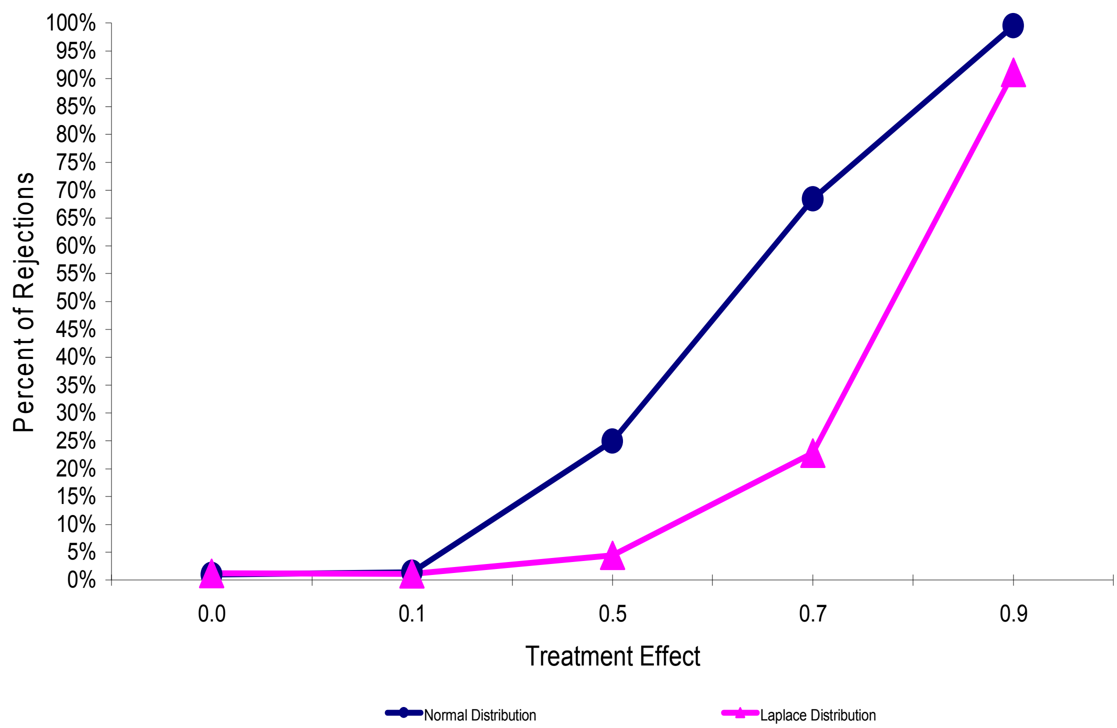


Figure 39: Normal Distribution and T Distribution (3 DF) - Sample Size 15 at 1% Significance Level

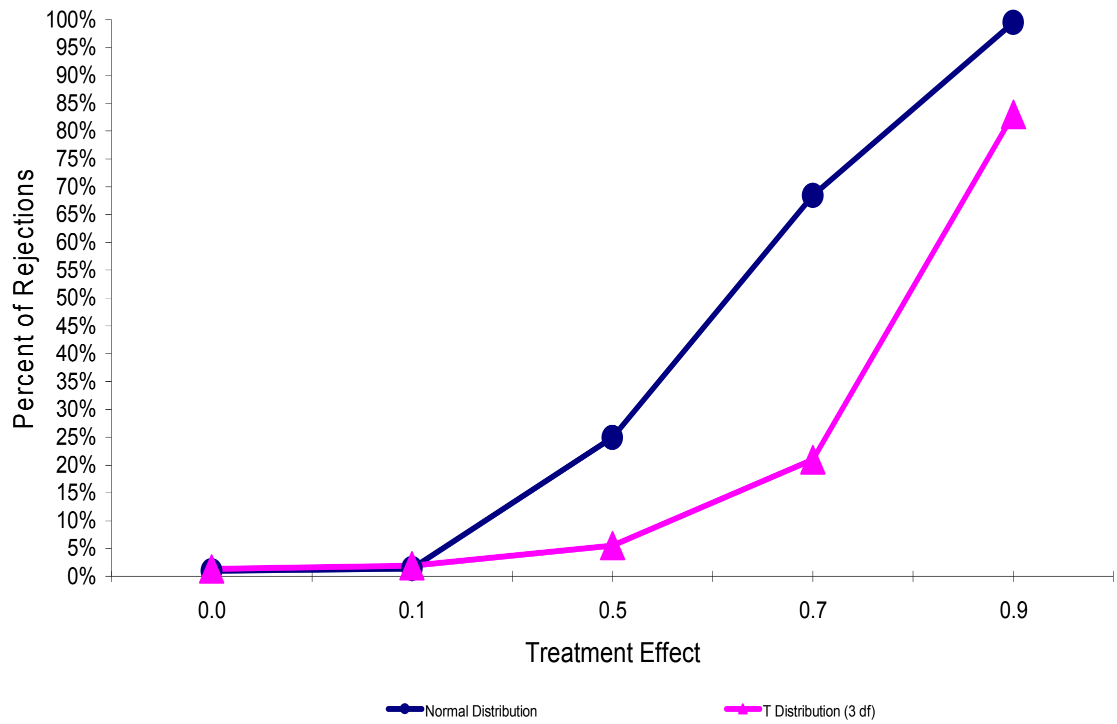


Figure 40: All Distributions - Sample Size 15 at 1% Significance Level

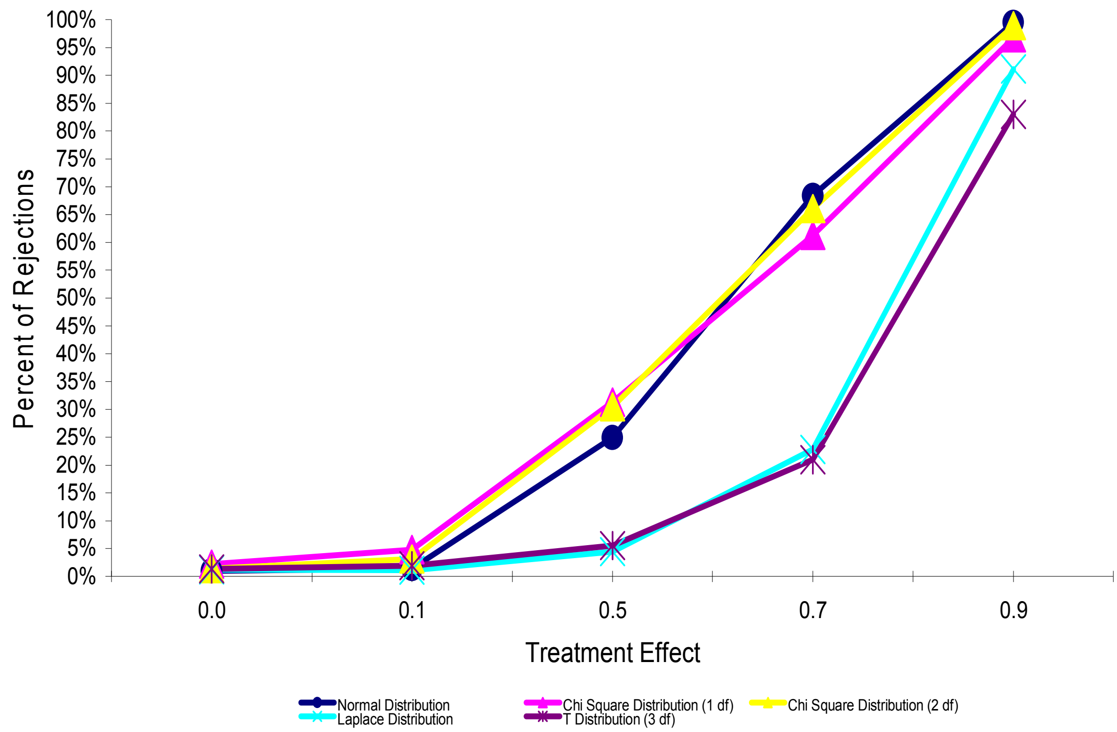


Figure 41: Normal Distribution - Sample Size 30 at 1% Significance Level

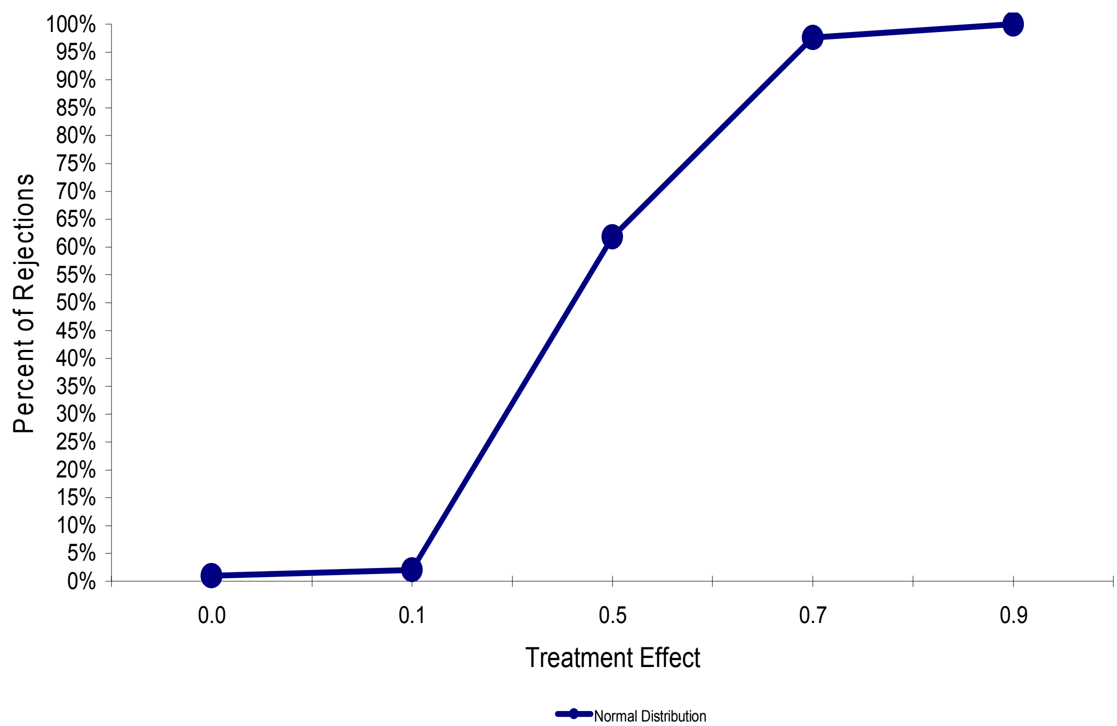


Figure 42: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 30 at 1% Significance Level

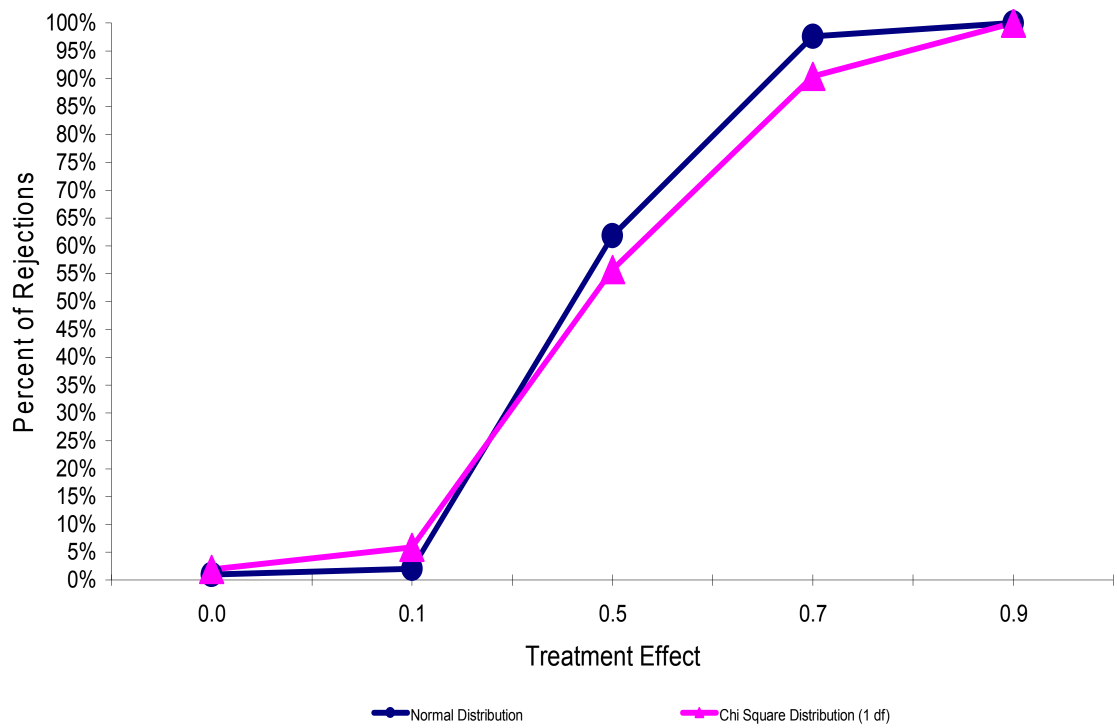


Figure 43: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 30 at 1% Significance Level

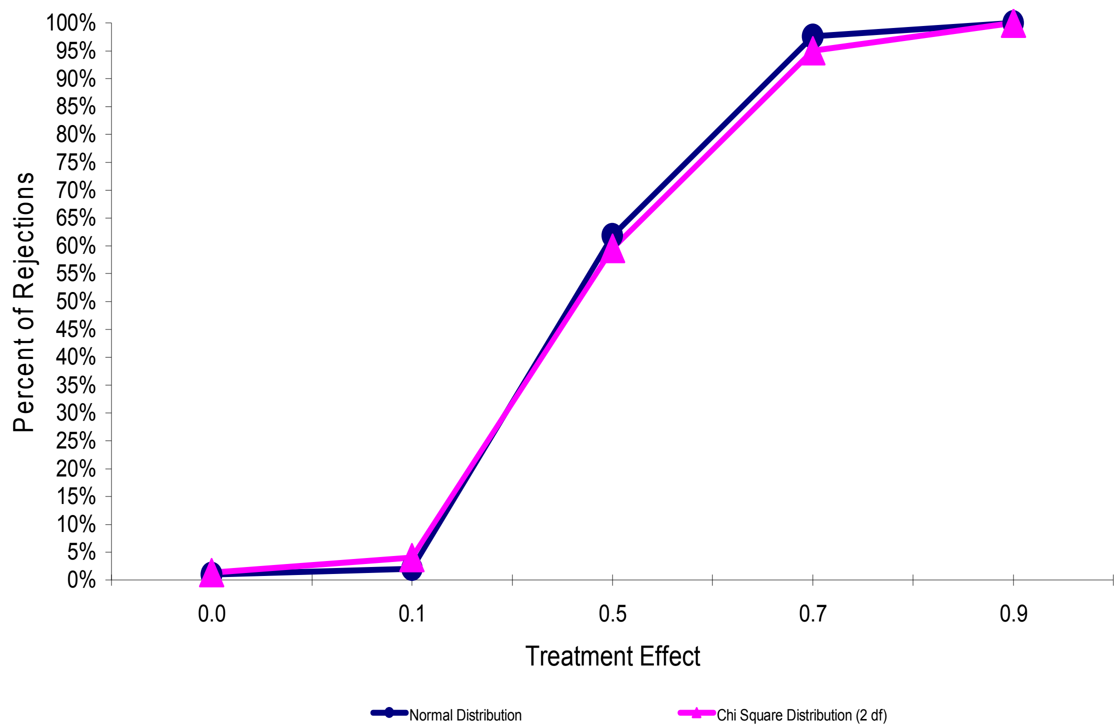


Figure 44: Normal Distribution and Laplace Distribution - Sample Size 30 at 1% Significance Level

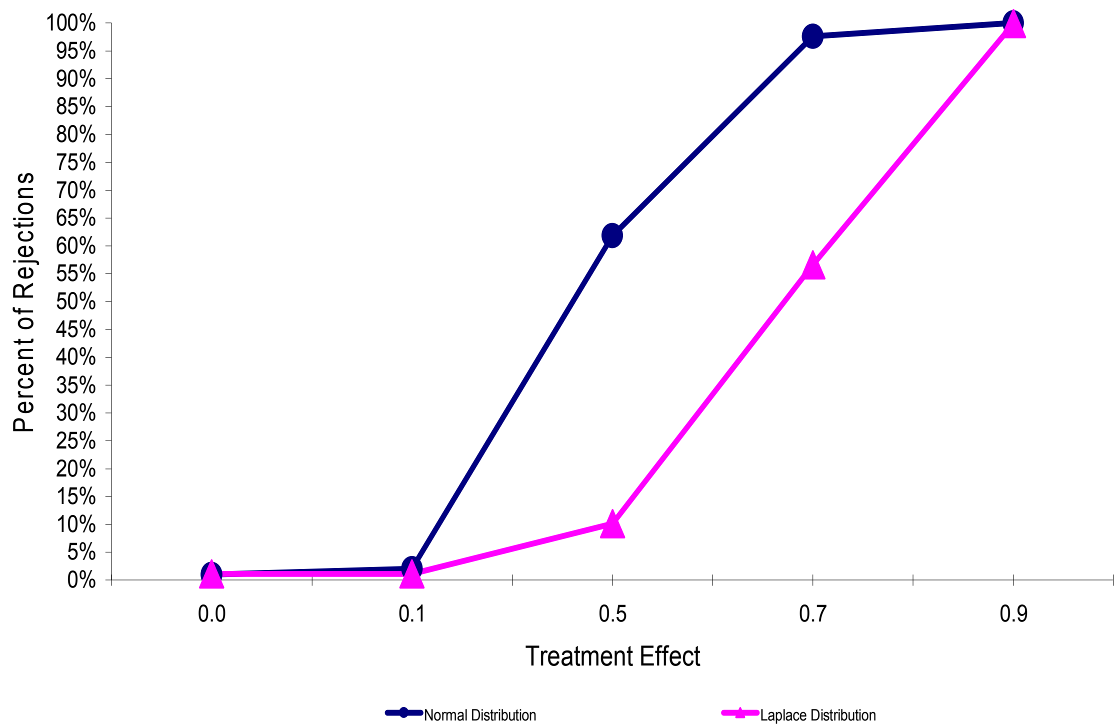


Figure 45: Normal Distribution and T Distribution - Sample Size 30 at 1% Significance Level

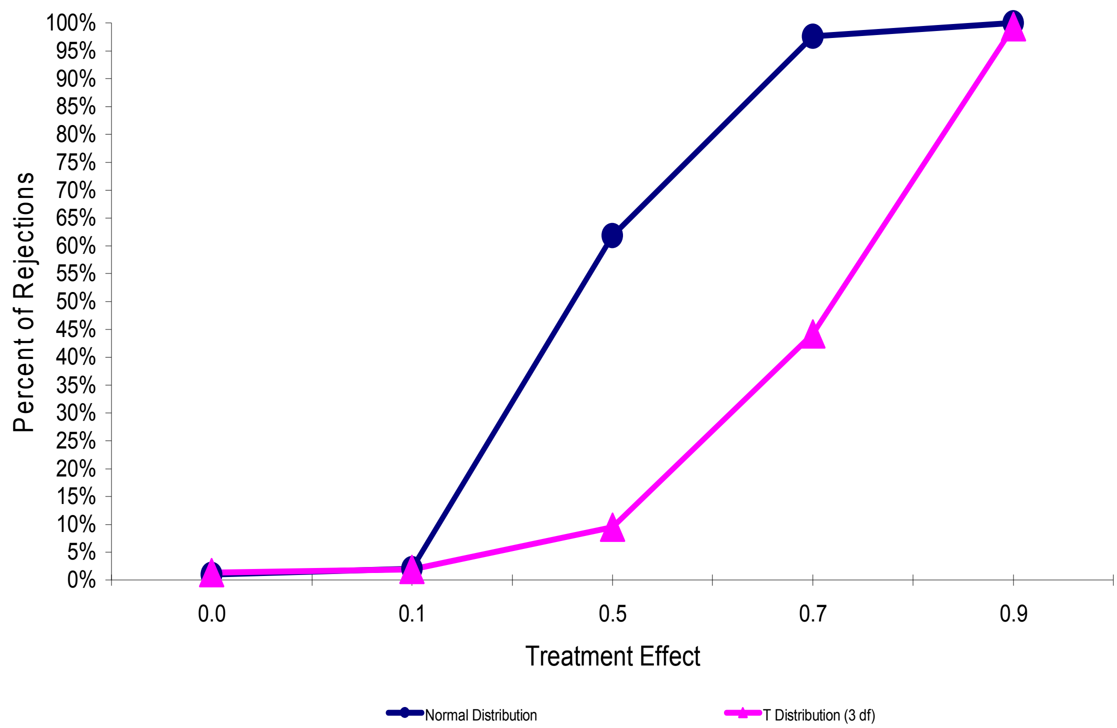


Figure 46: All Distributions - Sample Size 30 at 1% Significance Level

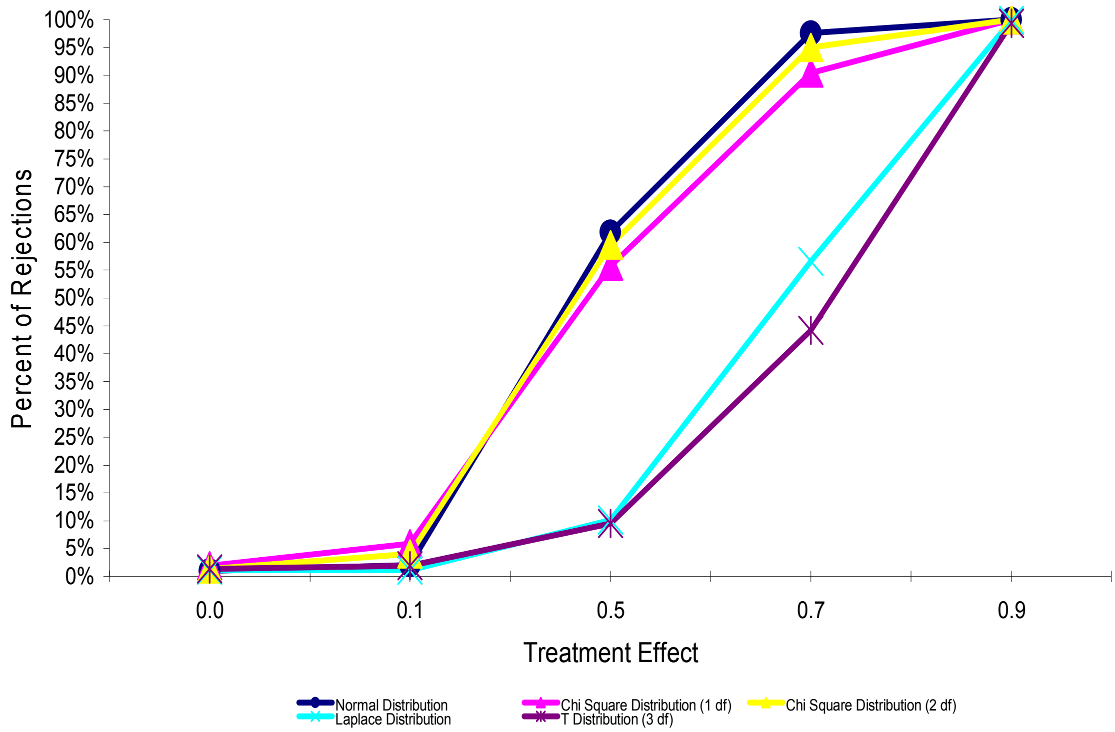


Figure 47: Normal Distribution - Sample Size 60 at 1% Significance Level

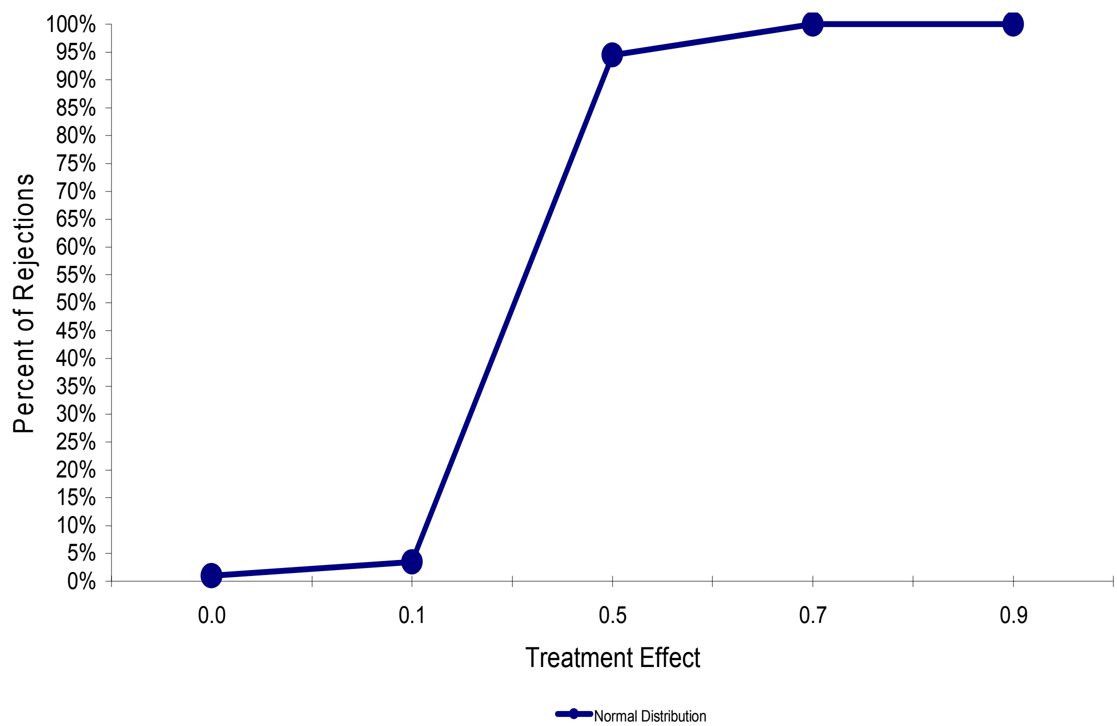


Figure 48: Normal Distribution and Chi Square Distribution (1 DF) - Sample Size 60 at 1% Significance Level

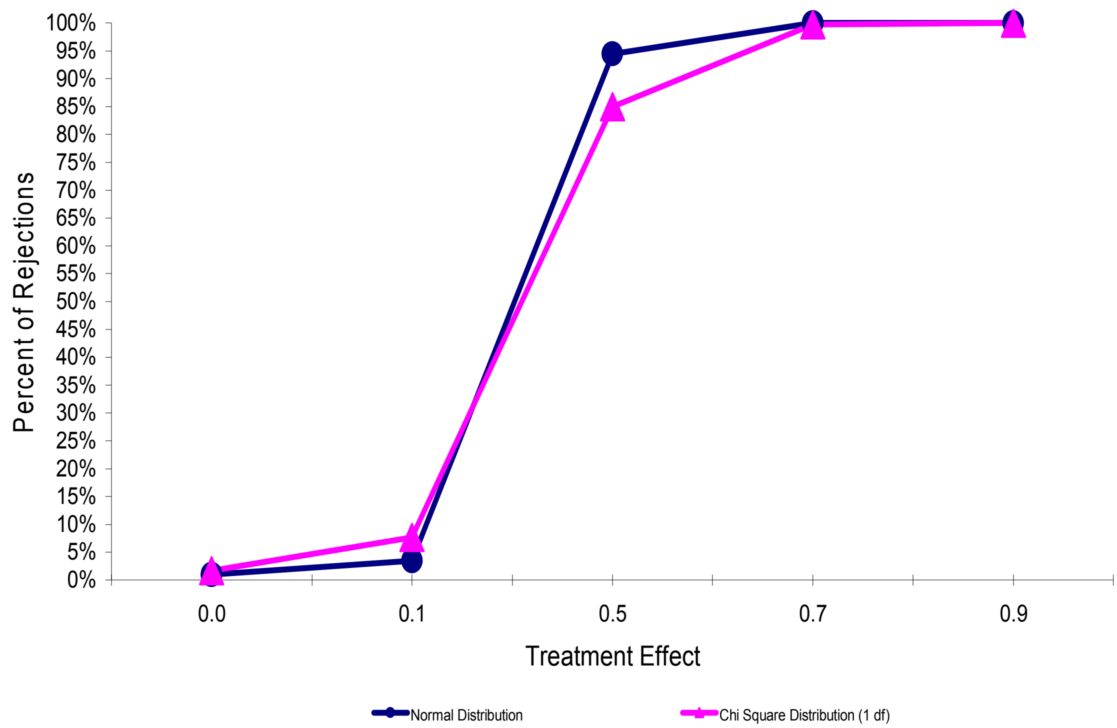


Figure 49: Normal Distribution and Chi Square Distribution (2 DF) - Sample Size 60 at 1% Significance Level

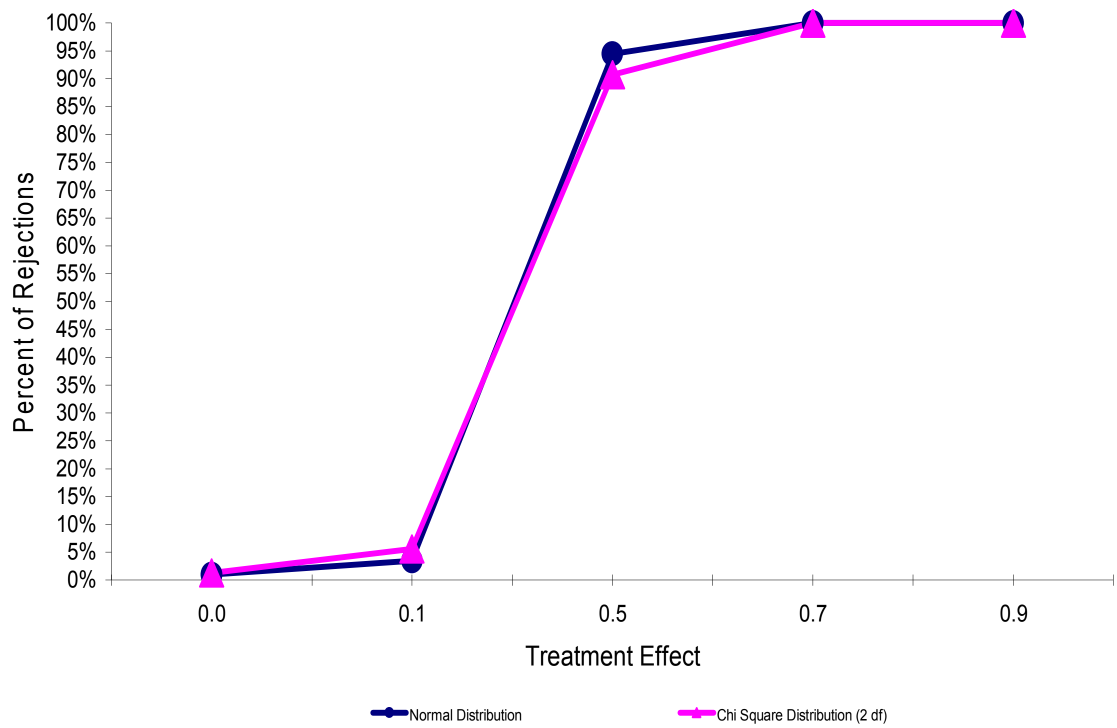


Figure 50: Normal Distribution and Laplace Distribution - Sample Size 60 at 1% Significance Level

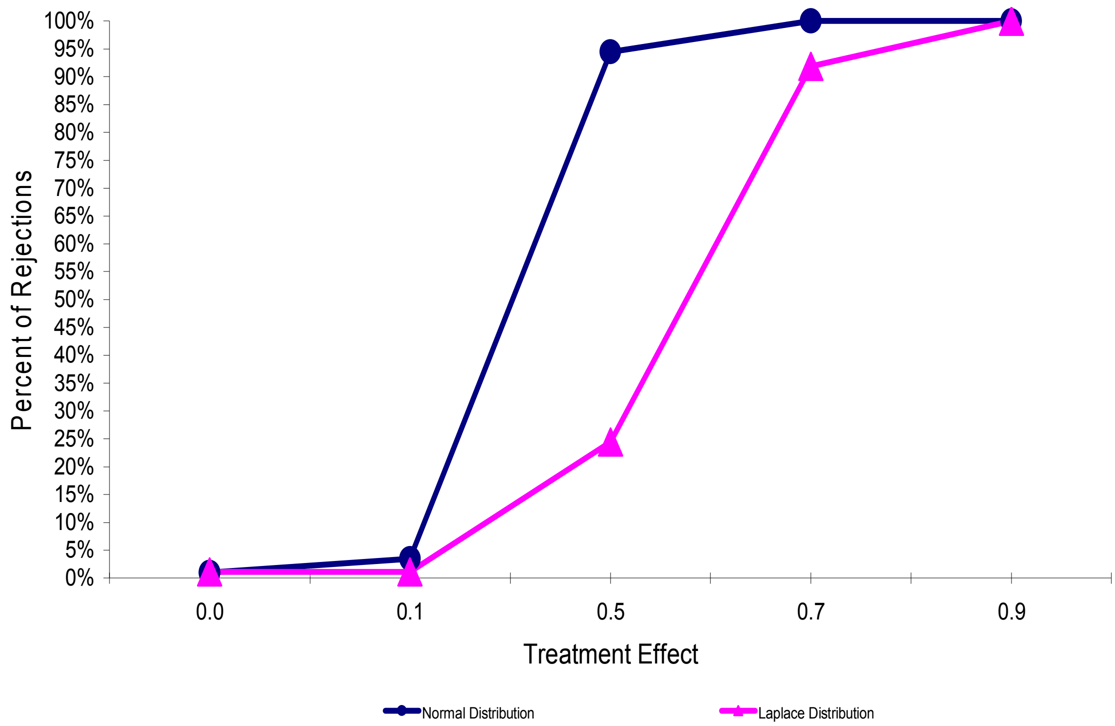


Figure 51: Normal Distribution and T Distribution (3 DF) - Sample Size 60 at 1% Significance Level

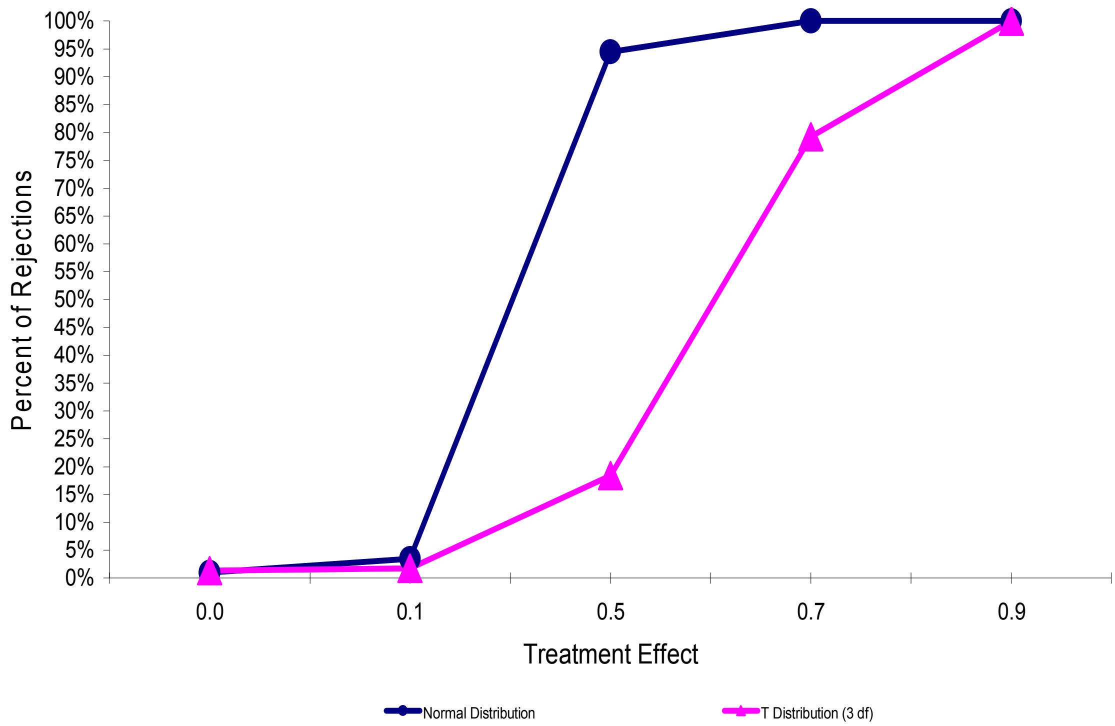
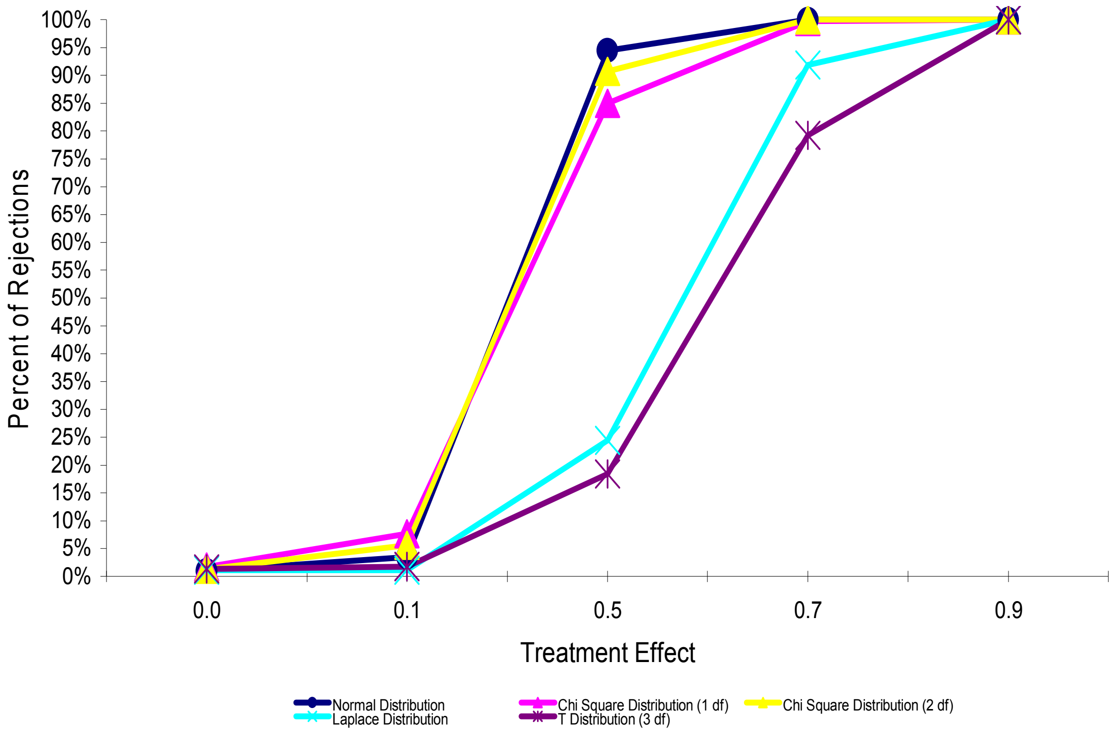


Figure 52: All Distributions - Sample Size 60 at 1% Significance Level



Chapter 5

Conclusion

Figures 5 – 9 and Table 2 illustrate the relationship between distribution and the 5% significance level for bivariate samples of size 5. Figure 5 shows the treatment effect on the Normal Distribution as rho is increased from 0.0 to 0.9. At 0.0 the results are consistent with that of the 5% significance level. As rho increases to 0.1, the percent of rejections are similar to the percent of rejections when rho is 0.0. Significantly noticeable difference are apparent beyond $\rho=0.2$. When $\rho=0.9$ the percent of rejections of a false null hypothesis is 67.2% for the Normal Distribution when the sample size is 5.

When the Chi Square Distribution (1 DF) is compared to the Normal Distribution (Figure 6), the percent of rejections of a false null hypothesis beyond $\rho=0$ is slightly greater with the largest difference when $\rho=0.5$. At $\rho=0.5$, the difference between the Normal Distribution and the Chi Square Distribution (1 DF) is approximately 8 percentage points.

The percent of rejections for a false null hypothesis for the Chi Square Distribution (2 DF) compared to the Normal Distribution (Figure 7) is similar to Figure 6 with the Chi Square Distribution (2 DF) rejecting at a greater rate when rho is larger than 0.1. In Figure 7, the largest difference occurs when rho is 0.7 (approximately 8 percentage points).

Unlike the Chi Square Distributions, the Laplace Distribution has a lower rate of rejection when compared to the Normal Distribution provided that rho is larger than 0.1.

The percent of rejection steadily increases as rho increases beyond 0.1. When rho reaches 0.9 the percent of rejections of a false null hypothesis is 67.2% for the Normal Distribution and 43.1% for the Laplace Distribution nearly a 24 percentage point difference in rejections.

Similarly, Figure 9 shows that the trend in the percent of rejections for the T Distribution (3 DF) compared to the Normal Distribution reflects that of the Laplace Distribution. When ρ=0.9, the Normal Distribution rejects 67.2% of false null hypotheses while the T Distribution (3 DF) rejects 42.4 % of false null hypotheses for a difference of approximately 25 percentage points.

In summary, for sample size n = 5 at the 0.05 nominal alpha level, Figure 10 illustrates that although the Chi Square Distributions may reject false null hypothesis at a higher rate when rho is larger than 0.1 by the time rho reaches 0.9 the rate of rejection for the Normal Distribution and the two Chi Square Distributions (1 and 2 DF) are basically the same. This is not the case for the Laplace and T Distributions (3 DF). The rate of rejection of false null hypotheses for the Normal Distribution is more than 1.5 times that of the T Distribution (3 DF) and the Laplace Distribution.

Figures 11 – 16 and Table 3 illustrate the percent of rejection of false null hypotheses for sample of size 15 at the 5% significance level. Figure 11 shows that the percent of rejections increase almost 8-fold when rho is between 0.1 and 0.5. And, the percent of rejections is relatively steady with the maximum reached when ρ=0.9 at this value the rate of rejections is 99.9%.

Unlike the results for the Chi Square Distribution (1 DF) for sample size 5, Figure 12 shows a variable trend in rejections for the Chi Square Distribution (1 DF) compared with the Normal Distribution for sample size 15. From ρ=0.0 to ρ=0.2, the percent of rejections for the Chi Square Distribution (1 DF) is greater than the Normal Distribution. Beyond ρ=0.2 the percent of rejections for false null hypotheses for the Chi Square Distribution (1 DF) is less than the Normal Distribution. However, when rho reaches 0.9 there is less than a 1 percentage point difference in the rate of rejections for the Normal Distribution compared to the Chi Square Distribution (1 DF).

The trend in rejections for the Chi Square Distribution (2 DF) is quite similar to the Normal Distribution. The largest difference noted is when ρ=0.7. At this treatment effect the Normal Distribution rejects at a rate of 87.7% while the Chi Square Distribution (2 DF) rejects at a rate of 83.8% for a difference of approximately 4 percentage points.

The Laplace Distribution is similar to the Normal Distribution when rho is between 0.0 and 0.1 and when rho is 0.9. However, between 0.1 and 0.9 there is a large difference in the percent of rejections of false null hypotheses. For instance, when rho is 0.5 and 0.7, the percent of rejections for the Laplace Distribution is 14.3% and 46.7%, respectively. These correspond to differences of 35.8 percentage points and 41.0 percentage points when compared to the Normal Distribution. Figure 15 shows that the trend in the results for the T Distribution (3 DF) is similar to that of the Laplace Distribution with the largest difference in rejections, when compared to the Normal Distribution, being 48.0 percentage points.

In summary, for samples of size 15 at the 5% significance level, Figure 16 illustrates that the rate of rejections of false null hypotheses increases for all distribution.

However, the rate of rejection for the Chi Squared Distributions (1 and 2 DF) is beginning to more closely reflect the Normal Distribution. Also, the Laplace Distribution and the T Distribution (3 DF) are quite similar to the Normal Distribution when rho is less than or equal to 0.1 or when rho is 0.9. In between those two markers, there is considerable differences in the rate of rejection of false null hypotheses for the Laplace Distribution and T Distribution (3 DF) when compared to the Normal Distribution.

Figures 17 – 22 and Table 4 detail the rate of rejections for samples of size 30 at the 5% significance level. Beyond ρ=0.1, the rate of rejection for false null hypotheses increases more than 10-fold for the Normal Distribution as indicated by Figure 17 and Table 4. The percent of rejections begin to level off when rho is greater than or equal to 0.7 achieving maximum power when ρ=0.9.

Figure 18 illustrates that the rate of rejections for the Chi Square Distribution (1 DF) begins to echo the rate of rejections for the Normal Distribution as rho increases beyond 0.5. Upon viewing Figure 19, the same can be said for the Chi Square Distribution (2 DF). However, the percent of rejections for the Chi Square Distribution with 2 degrees of freedom is almost a mirror image of the Normal Distribution.

Figure 20 and Table 4 demonstrates that the rate of rejections for the Laplace Distribution improves for sample size 30. However, for rho between 0.1 and 0.9, the rate of rejecting a false null hypothesis is considerably lower than the Normal Distribution. When ρ=0.5, the rate of rejecting a false null hypothesis for the Laplace Distribution is 57.0 percentage points less than the Normal Distribution. When rho increases to 0.7, the difference in the rate of rejection decreases to approximately 20 percentage points compared to the Normal Distribution.

Figure 21 shows that the fit of the T Distribution with 3 degrees of freedom is less compatible with the Normal Distribution than the Laplace Distribution although the trend in rejections of the T Distribution (3 DF) is similar to the Laplace Distribution. In other words, when rho is between 0.1 and 0.9, there is considerable differences in rejecting false null hypotheses for the T Distribution (3 DF) when compared to the Normal Distribution as was the case with the Laplace Distribution. But, when ρ=0.5, 21.3% of false null hypotheses are rejected for the T Distribution (3DF) compared to 25.8% for the Laplace Distribution and 82.8% for the Normal Distribution. This produces a 61.5 percentage point difference between the Normal Distribution and the T Distribution (3 DF).

Overall, for samples of size 30 at the 5% significance level, Figure 22 shows that percent of rejections have increase for all distributions. However, the Chi Square Distributions are becoming almost identical to the Normal Distribution's rate of rejections while the Laplace Distribution and T Distribution (3 DF) continue to reject at a much lower rate.

Figures 23 – 28 and Table 5 show the rejection rates for samples of size 60 at the 5% significance level. At ρ=0.5, Figure 23 illustrates that the Normal Distribution reaches almost maximum power with 98.6% of false null hypotheses rejected.

The percent of rejections for the Chi Square Distribution (1 DF) is very similar to the Normal Distribution with a trivial difference being noted when ρ=0.5. For this treatment effect, the rate of rejection for the Normal Distribution is 98.6% but 93.4% for the Chi Square Distribution (1 DF). This produces a small difference of 5.2 percentage points.

Figure 25 demonstrates that the Chi Square Distribution (2 DF) can essentially be considered a replica of the Normal Distribution. It reaches its maximum power at ρ=0.5 just as Chi Square with 1 degree of freedom and the Normal Distribution. However, at ρ=0.5, there is only a 1.6 percentage difference in rejections of false null hypotheses when compared to the Normal Distribution as opposed to 5.2 percentage points for the Chi Square Distribution (1 DF).

Table 5 and Figure 26 shows that the Laplace Distribution almost achieves maximum power when ρ=0.7. At this value the percent of rejections is 97.8%. And, the gap in rejections between a treatment effect of 0.1 and 0.9 has also decreased. For instance, with a sample of size 30, the Laplace Distribution was not comparable to the Normal Distribution until rho was equal to 0.9. At this value, the Laplace Distribution had a rejection rate of 99.9% and the Normal Distribution had a rejection rate of 100% . But, with samples of size 60, the Laplace Distribution begins to compare to the Normal Distribution when rho approaches 0.7. At this value, the Laplace Distribution has a rejection rate of 97.8% and the Normal Distribution has a rejection rate of 99.9%.

Figure 27 reveals that the gap in rejections for the T Distribution (3 DF) has decreased as well. The T Distribution (3 DF) also begins to compare to the Normal Distribution when ρ=0.7. At this value, the percent of rejections for the T Distribution (3 DF) is 92.8% compared to 99.9% rejections for the Normal Distributions.

In summary, for samples of size 60 at the 5% significance level, Figure 28 reveals that the rate of rejections increase for all of the distributions. The Chi Square Distributions are basically identical to the Normal Distribution and the Laplace Distribution and T Distribution (3 DF) are approaching the Normal Distribution.

However, appropriate power is more challenging for these two distributions when rho is between 0.1 and 0.7.

Figures 29 – 34 and Table 6 illustrate the rate of rejection of false null hypotheses at the 1% significance levels for samples of size 15. It can be readily seen that, for the Normal Distribution, the rate of rejection is lower for the 1% significance level at sample size 5 than for the 5% significance level. When the treatment effect, rho, is 0.9, the rate of rejections at the 1% significance level for sample size 5 is approximately 37 percentage points less than at the 5% significance level.

Figure 30 shows that the Chi Square Distribution (1 DF) has a higher rate of rejection when compared to the Normal Distribution with the greatest difference noted when the treatment effect is 0.7. At this value, the rate of false rejections for the Chi Square Distribution (1 DF) is 16.5% while the Normal Distribution is 7.4%. The Chi Square Distribution with 2 degrees of freedom, Figure 31, is slightly worse than the Chi Square Distribution with 1 degree of freedom with the same treatment effect. When ρ = 0.7, the Chi Square Distribution with 2 degrees of freedom has a false null hypothesis rejection rate of 14.4% compared to 16.5% and 7.5% for the Chi Square Distribution (1 DF) and Normal Distribution, respectively.

The rejection rates for the Laplace Distribution and the T Distribution (3 DF), Table 6, are quite similar to each other at the 1% significance level given a sample of size five. Figures 32 and 33 show that the rejection rate for the Normal Distribution begins to exceed the Laplace Distribution and the T Distribution (3 DF) as the when rho exceeds 0.5. When $\rho = 0.9$, the rate of rejections for the Normal Distribution is at least twice that of the Laplace Distribution and T Distribution.

In summary, Figure 34 shows that the Normal Distribution is between essentially mirror images of the two Chi Square Distributions (1 and 2 DF) and the Laplace Distribution and T Distribution (3 DF), respectively. Power is a significant challenge for all five distributions for a sample of size 5 at the 1% level of significance. Table 6 reveals that the maximum power at sample size 5 with 1% level of significance is 38.8% rejections of false null hypotheses for the Chi Square Distribution with 2 degrees of freedom when $\rho = 0.9$. This rejection rate is the largest of all the distributions investigated. However, evaluated at the same sample size, this rejection rate is less than the rejection rates for all of the distributions investigated at the5% significance level by at least 3.6 percentage points up to 28.9 percentage points when $\rho = 0.9$.

Figures 35 – 40 and Table 7 illustrate the rate of rejections of false null hypotheses for all distributions given a sample of size 15 and a significance level of 1%. The Normal Distribution makes a significant improvement in power when $\rho = 0.5$ by more than 8 times as much increasing from 3.5% rejections to 25.0% rejections. The Normal Distribution achieves approximately maximum power of 99.5% rejections when $\rho = 0.9$.

The rate of rejection for the Chi Square Distribution (1 DF) is slightly better than the Normal Distribution when rho is between 0.0 and 0.6. However, the rate of rejection dips below the Normal Distribution when rho grows larger than 0.6. At $\rho = 0.9$, the rate of rejection for false null hypotheses for the Chi Square Distribution (1 DF) is slightly lower than the Normal Distribution with a rate of 96.8% of rejections for the Chi Square Distribution (1 DF) versus 99.5% for the Normal Distribution. Figure 37 shows that the trend of rejections for the Chi Square Distribution with 2 degree of freedom is almost the same as the Chi Square Distribution with 1 degree of freedom. However, Table 7 shows that, in terms of maximum power, when $\rho = 0.9$, the Chi Square Distribution with 2 degrees of freedom is slightly better than the Chi Square Distribution with 1 degree of freedom, 98.9% rejections versus 96.8% rejections respectively. But, in comparison to the Normal Distribution, the rate of rejections for the Chi Square Distribution (2 DF) is marginally less than the rate of rejections for the Normal Distribution, 99.5% versus 98.9%, respectively.

At sample size 15, the rate of rejections for the Laplace Distribution improves greatly compared to the rate of rejection at sample size 5 (Table 7). However, its rate of rejection is still less than the Normal Distribution. Figure 38 shows a noteworthy gap in the rate of rejections for the Laplace Distribution when rho is between 0.1 and 0.9 compared to the Normal Distribution. The largest difference is when $\rho = 0.7$. At this value, the rate of rejections for the Normal Distribution is 68.4%. But, for the Laplace Distribution, it is 22.8%. This is a difference of 45.6 percentage points.

The trend in results of the t Distribution (3 DF) is similar to that of the Laplace Distribution (Figure 39). However, the same noteworthy gap is larger for the T Distribution (3 DF) than for the Laplace Distribution. When rho is between 0.1 and 0.9, the difference in rejections of false null hypothesis for the T Distribution (3 DF) in comparison to the Normal Distribution is as large as 47.4% (Table 7). This difference is 1.8 percentage points larger than the Laplace Distribution over the same range of treatment effect.

Figure 40 shows that, overall, as the sample size has increased from 5 to 15, the rate of rejection of false null hypotheses for the Normal Distribution has increased over the other investigated distributions. This is particularly the case as the treatment effect gets stronger. However, even with an increased sample size, power for the Laplace Distribution and the t Distribution (3 DF) is a challenge when rho is between 0.1 and 0.9.

Figures 41 – 46 and Table 8 illustrate the rate of rejection of false null hypotheses for samples of size 30 at the 1% significance level. Figure 41 shows that the Normal Distribution begins to achieve maximum power at a treatment effect of $\rho = 0.7$. At this value, the rate of rejection of false null hypotheses is 97.6%. At $\rho = 0.9$, the rate of rejection of false null hypotheses is 100%.

The rate of rejections for the Chi Square Distribution (1 DF) closely resembles that of the Normal Distribution (Figure 42). However, the rate of rejections of false null hypotheses favors the Normal Distribution comparatively. The Chi Square Distribution with 2 degrees of freedom, Figure 43, is an even closer fit than the Chi Square Distribution with 1 degree of freedom. The largest disagreement in rates of rejection of the Normal Distribution compared to the Chi Square Distribution with 2 degrees of freedom is 2.5 percentage points when $\rho = 0.7$ (Table 8).

The noteworthy gap in rejection rates persists for the Laplace Distribution compared to the Normal Distribution as shown in Figure 44. Although the rate of rejections has increased with an increase in sample, power is still a challenge for the Laplace Distribution when rho is between 0.1 and 0.9. The same observation can be made for the T Distribution with 3 degrees of freedom (Figure 45). Although the T

Distribution approaches maximum power with a treatment effect of 0.9, the largest difference in rejection rates when compared to the Normal Distribution is 53.4 percentage points when $\rho = 0.7$ (Table 8).

Overall, for samples of size 30 at the 1% significance level, the rates of rejection of false null hypotheses increases for all of the investigated distributions beyond a treatment effect of 0.1. The rejection rates are larger among the Chi Square Distributions (1 and 2 DF) with the Laplace Distribution and T Distribution (3 DF) lagging behind substantially until the treatment effect reaches 0.9.

Figures 47 – 52 and Table 9 detail the rates of rejection of false null hypotheses for samples of size 60 at the 1% significance level. The Normal Distribution begins to approach maximum power at $\rho = 0.5$. At this value, the proportion of rejections of false null hypotheses is 94.5% and increasing slightly to maximum power when $\rho = 0.9$.

With an increase in sample size, the power of the Chi Square Distributions (1 and 2 DF) begins to become more substantial with a treatment effect of 0.5. At this value, the rate of rejection for each Chi Square Distribution is 85.0% and 90.7%, respectively. Beyond a treatment effect of 0.5, each distribution increases steadily to maximum power and become mirror images of the Normal Distribution when $\rho = 0.7$ (Figures 48 and 49).

Figure 50 shows that the noteworthy gap referenced previously with respect to the Laplace Distribution has decreased from a treatment effect between 0.1 and 0.9 to a treatment effect between 0.1 and 0.7. Within this range, Table 9 shows that the largest difference in rates o f rejection of false null hypotheses between the Normal Distribution and the Laplace Distribution is 70.1 percentage points which occurs when $\rho = 0.5$. This difference is even larger for the T Distribution (3 DF). When $\rho = 0.5$, the rates of rejection for the Normal Distribution and T Distribution (3 DF) is 94.5% and 18.4% (Table 9). This produces a difference of 76.1 percentage points. Although this difference gets less as the treatment effect increases, power is a challenge for the T Distribution (3 DF) when the treatment effect is moderate.

With an increase in sample size, Figure 52 shows that the rejection rates of each of the non-normal distributions approaches the rate of the Normal Distribution. It is clear that the Chi Square Distributions (1 and 2 DF) are more reflective of the Normal Distribution while the Laplace Distribution and T Distribution (3 DF) do not begin to compare to the Normal Distribution until the treatment effect reaches 0.7.

The chi square distribution is a distribution created directly from the Normal Distribution. By repeatedly collecting samples of various sizes from the Normal Distribution and computing sample variances for each sample size, the chi square distribution is created. This observation can explain the close approximation in rejection rates or power of the Chi Square Distributions (1 and 2 DF) to the Normal Distribution. The differences in rejection rates of the Laplace Distribution and T Distribution (3 DF) in comparison to the Normal Distribution remains unexplained.

Table 10: Type II Robustness and Bradley's Conservative Criterion - Significance Level at 5%

	5%				
Sample Size: 5	**0.0**	**0.1**	**0.5**	**0.7**	**0.9**
Normal Distribution	5.020300%	5.220900%	13.077500%	26.992000%	67.216500%
Chi Square Distribution (1 df)	6.537600%	8.540400%	21.135900%	33.932800%	63.206200%
Chi Square Distribution (2 df)	5.485400%	6.958300%	19.552100%	34.106200%	67.727400%
Laplace Distribution	5.234900%	5.077300%	**6.768200%**	**12.931500%**	**43.128200%**
T Distribution (3 df)	5.192100%	5.789700%	**7.692000%**	**14.279800%**	**42.461800%**
Sample Size: 15					
Normal Distribution	5.027000%	6.385000%	50.124700%	87.703400%	99.928500%
Chi Square Distribution (1 df)	5.436700%	10.184500%	47.690100%	**77.442400%**	99.105000%
Chi Square Distribution (2 df)	5.068000%	8.546700%	50.397100%	83.817000%	99.846800%
Laplace Distribution	5.260800%	**5.173600%**	**14.383700%**	**46.719000%**	97.847300%
T Distribution (3 df)	5.376300%	6.052300%	**13.783900%**	**39.791300%**	94.790800%
Sample Size: 30					
Normal Distribution	5.004500%	8.188700%	82.806600%	99.522000%	100.000000%
Chi Square Distribution (1 df)	5.004900%	12.111700%	**72.135500%**	96.275200%	99.997900%
Chi Square Distribution (2 df)	4.889500%	10.353700%	78.292700%	98.742700%	99.999900%
Laplace Distribution	5.169500%	**5.181700%**	**25.827300%**	**79.091800%**	99.934000%
T Distribution (3 df)	5.269700%	**5.878400%**	**21.373300%**	**67.066100%**	99.937900%
Sample Size: 60					
Normal Distribution	5.034400%	11.878400%	98.681400%	99.999800%	100.000000%
Chi Square Distribution (1 df)	4.816800%	15.622300%	93.460800%	99.386000%	100.000000%
Chi Square Distribution (2 df)	4.903700%	13.880600%	97.070400%	99.996700%	100.000000%
Laplace Distribution	5.153300%	**5.123500%**	**47.280700%**	97.883600%	100.000000%
T Distribution (3 df)	5.211700%	**5.584500%**	**36.541400%**	92.817900%	100.000000%

Table 11: Type II Robustness and Bradley's Conservative Criterion - Significance Level at 1%

	1%				
Sample Size: 5	**0.0**	**0.1**	**0.5**	**0.7**	**0.9**
Normal Distribution	0.994000%	1.061300%	3.056700%	7.463800%	30.901700%
Chi Square Distribution (1 df)	2.066200%	2.952700%	9.190000%	16.531900%	38.476000%
Chi Square Distribution (2 df)	1.287200%	1.856800%	7.085700%	14.433200%	38.797700%
Laplace Distribution	1.056000%	1.015100%	**1.426000%**	**3.100400%**	**15.072500%**
T Distribution (3 df)	1.072700%	1.326900%	**1.943100%**	**4.226700%**	**17.742200%**
Sample Size: 15					
Normal Distribution	0.984400%	1.426400%	24.955500%	68.387700%	99.517500%
Chi Square Distribution (1 df)	2.176900%	4.800800%	31.327800%	**61.269300%**	96.760500%
Chi Square Distribution (2 df)	1.448400%	3.141700%	30.430500%	66.053700%	98.944100%
Laplace Distribution	1.178800%	**1.121700%**	**4.460700%**	**22.826900%**	91.159200%
T Distribution (3 df)	1.342800%	1.939000%	**5.526700%**	**20.956800%**	**83.034500%**
Sample Size: 30					
Normal Distribution	1.000300%	2.029400%	61.815100%	97.573900%	100.000000%
Chi Square Distribution (1 df)	1.925500%	5.869700%	55.742300%	90.413800%	99.988300%
Chi Square Distribution (2 df)	1.338400%	4.010000%	59.532500%	95.058500%	99.999300%
Laplace Distribution	1.153100%	**1.111600%**	**10.160600%**	**56.557000%**	99.938100%
T Distribution (3 df)	1.376900%	1.919400%	**9.503400%**	**44.207500%**	99.382200%
Sample Size: 60					
Normal Distribution	0.994800%	3.468800%	94.481200%	99.996300%	100.000000%
Chi Square Distribution (1 df)	1.668800%	7.630800%	**84.974700%**	99.690000%	100.000000%
Chi Square Distribution (2 df)	1.240800%	5.627600%	90.687800%	99.967100%	100.000000%
Laplace Distribution	1.122200%	**1.105200%**	**24.438700%**	91.874200%	100.000000%
T Distribution (3 df)	1.372800%	**1.786800%**	**18.385900%**	**79.270400%**	99.999600%

Given the demonstrated varying degrees of Type II robustness for ρ, the lower end of Bradley's (1978) conservative criterion ($0.9\alpha \le \gamma \le 1.1\alpha$) for Type I robustness should be considered as a guide for determining acceptable levels of rejection rates in comparison to the Normal Distribution. In other words, Table 10 and Table 11 show the relationship between the lower end of Bradley's (1978) conservative criterion ($0.9\alpha \le \gamma \le 1.1\alpha$) and Type II robustness. The left column of each table contains the sample sizes and distributions investigated in the study. The top two rows contain the significance level and treatment effect. The body of each table contains the rejection rates for each distribution and its corresponding sample size, treatment effect, and significance level. The rejection rates in red in the body of the tables are those values that would be less than 90% of the rejection rates established by the Normal Distribution. For instance, in Table 10 it can be seen quite clearly that when the treatment effect is 0.5, the rejection rates of the Laplace Distribution and T Distribution are less than 90% of the rejection rates of the Normal Distribution for all sample sizes. This would be a violation of Bradley's conservative criterion and hence considered unacceptable power levels. The same pattern can be seen in Table 11.

In Tables 10 and 11, when using Bradley's (1978) conservative criterion, it becomes obvious when sample size becomes an advantage and a disadvantage. When the treatment effect is small, say 0.1, increasing the sample size does not mitigate the failure of the Laplace Distribution and T Distribution to achieve 90% of the rejection rates of the Normal Distribution. If anything, increasing the sample size only increases the discrepancy in rejection rates of the Laplace Distribution and T Distribution in comparison to the Normal Distribution whether the significance level is 5% or 1%. So, increasing the sample size may actually be a disadvantage for small treatment effects. This pattern persists through moderate treatment effects (0.5) and moderately high treatment effects (0.7). However, sample size becomes an advantage when the treatment effect is moderately high (0.7) and the sample size is larger than 30. At this point, the rejection rates for the Laplace Distribution and T Distribution are at least 90% of the rejection rates for the Normal Distribution. And, increasing the sample size does provide the advantage of mitigating the discrepancies in rejection rates between the Laplace Distribution and T Distribution in comparison to the Normal Distribution particularly when the treatment effect is large (0.9).

So, for low treatment effects through moderately high treatment effects, small sample sizes can have a deleterious effect on the Type II robustness of ρ. Researchers and practitioners alike must be aware of and diminish this effect by either increasing the sample size beyond 30 for population correlations that are larger than 0.7 or by realizing the disadvantage of conducting a study when the population correlation is low to moderate.

Finally, the following limitations and recommendations are suggested for this study. First, the treatment effects measured were 0.1, 0.5, 0.7, and 0.9. To better understand the point at which the treatment begins to have an effect on the rejection rates of the null hypothesis, it is recommended that the intermediate treatment values should be analyzed as well as those in this study. Second, the customary significance levels (5% and 1%) were used in this study. To give the researcher more latitude in choosing a significance level for their study, it is recommended that other significance levels such as 0.5% and 10% should be evaluated also. Third, Bradley (1978), Blair and Lawson

(1982), and Micceri (1989), have noted the prevalence of researchers to use theoretical and/or mathematical distributions in empirical research even though these distributions do not occur frequently in social science or educational research. This study can be improved upon by replicating it using "real" datasets such as the Micceri (1989) datasets. Doing so will provide the practitioner with greater insight into the Type II robustness of r from datasets which are more common in educational and social science research settings. And, lastly, Soper (1913), Fisher (1915), and Student (1908) demonstrated that the frequency distribution of r can become skewed and even distorted as the magnitude of the population correlation approaches +1. To address this occurrence, this study should be replicated using the Fisher's Z Transformation. A study of this type will illuminate the robustness of r under conditions which are meant to generate a normal distribution for the frequency distribution of r.

REFERENCES

Abdullah, M. B. (1990). On a robust correlation coefficient. *The Statistician,* 39(4), 455-460.

Baker, G. A. (1930). The significance of the product moment coefficient of correlation with special reference to the character of the marginal distribution. *Journal of the American Statistical Association,* 25(172), 387-396.

Blair, C. R & Lawson, S. B. (1982). Another look at the robustness of the product-moment correlation coefficient to population non-normality. *Florida Journal of Educational Research,* 24, 11-15.

Bradley, J. V. (1978). “Robustness?” *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.

Brooks, G. P., Kanyongo, G. Y., Kyei-Blankson, L., & Gocmen, G. (2002, April). *Power reliability for correlation and ANOVA.* Paper presented at the Annual Meeting of the American Education Research Association, New Orleans, Louisiana.

Cochran, W. G. (1947). Some consequences when the assumptions for the analysis Of variance are not satisfied. *Biometrics*, 10, 417-451.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

David, F. N. (1938). *Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples.* Cambridge, England: The University Press.

Edgell, S. E. & Noon, S. M. (1984). Effect of violation of normality on the t-test of the correlation coefficient. *Psychological Bulletin*, 95(3), 576-583.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrik*a, 10(4), 507-521.

Fleishman, I. A. (1978). A method for simulating non-normal distributions. *Psychometrika, 43(4),* 521-532.

Havlicek, L.L. & Peterson, N. L. (1976). Robustness of the pearson correlation against violations of assumptions. *Perceptual and Motor Skills*, 43, 1319-1334.

Headrick, T, C. & Sawilowsky, S. S. (1999). Simulating correlated multivariate nonnormal distributions: Extending the fleishman power method. *Psychometrika* 64(1), 25-35.

Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product moment correlation coefficient. *Applied Statistics*, 21(1), 1-12.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin,* 105(1), 156-166.

Pearson, E. S. (1931). The test of significance for the correlation coefficient. *Journal of the American Statistical Association*, 26(174), 128-134.

Pearson, E. S. (1932). The test of significance for the correlation coefficient: Some Further Results. *Journal of the American Statistical Association*, 27(180), 424-426.

Rider, P. R. (1932). On the distribution of the correlation coefficient in small samples. *Biometrika*, 24(3/4), 382-403.

Sawilowsky, S. S. & Blair, C. R. (1992). A more realistic look at the robustness and type II error properties of the t test to departures from population normality. *Psychological Bulletin*, 111(2), 352-360.

Sawilowsky, S. S. & Fahoome, G. F. (2003). *Statistics through monte carlo simulation with fortran.* Oak Park, Michigan: JMASM, Inc.

Sawilowsky, S. S. & Hillman, S. B. (1991, August). *Sample size tables , "t" test, and a prevalent psychometric distribution*. Paper presented at the Annual meeting of the American Psychological Association, San Francisco, California.

Soper, H. E. (1913). On the probable error of the correlation coefficient to a second approximation. *Biometrika*, 9(1/2), 91-115.

Soper, H. E., Young, A. W., Cave, B. M., Lee, A., Pearson, K. (1917). On the distribution of the correlation coefficient in small samples. Appendix II to the papers of "student" and r. a. fisher. Biometrika, 11(4), 328 - 413.

Srivastava, M. S. & Lee, G. C. (1984). On the distribution of the correlation coefficient when sampling from a mixture of two bivariate normal densities: robustness of outliers. *The Canadian Journal of Statistics,* 12(2), 119-133.

Student (1908). Probable error of a correlation coefficient. *Biometrika*, 6(2/3), 302-310.

ABSTRACT

TYPE II ROBUSTNESS OF H_0: $\rho=0$ FOR NON-NORMAL DISTRIBUTIONS

by

STEPHANIE D. WREN

December 2010

Advisor: Dr. Shlomo Sawilowsky

Major: Education Evaluation and Research

Degree: Doctor of Education

This dissertation was an investigation of Type II robustness as it relates to sample size, non-normality, significance level, and treatment effect on the Pearson Product Moment Correlation Coefficient.

AUTOBIOGRAPHICAL STATEMENT

I have always had an interest in research. I originally wanted to do automotive research. However, once I began teaching in Detroit Public Schools and because education was always a passionate issue to me, I decided to shift my career path to educational research instead. And, I am glad I did. This career path has taken me from being a student to a published author to chairwoman of the Michigan Association for Institutional Research. I feel there is no limit to what I can accomplish as an educational researcher.

Dr. Stephanie D. Wren