ROBUSTNESS OF NORMAL THEORY INFERENCE WHEN RANDOM EFFECTS ARE NOT NORMALLY DISTRIBUTED

by

MUDITHA VIRANGIKA DEVAMITTA PERERA

B.Sc., University of Colombo, Sri Lanka, 2006

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2011

Approved by:

Major Professor
Paul Nelson (Ph.D.)

Abstract

The variance of a response in a one-way random effects model can be expressed as the sum of the variability among and within treatment levels. Conventional methods of statistical analysis for these models are based on the assumption of normality of both sources of variation. Since this assumption is not always satisfied and can be difficult to check, it is important to explore the performance of normal based inference when normality does not hold. This report uses simulation to explore and assess the robustness of the F-test for the presence of an among treatment variance component and the normal theory confidence interval for the intra-class correlation coefficient under several non-normal distributions. It was found that the power function of the F-test is robust for moderately heavy-tailed random error distributions. But, for very heavy tailed random error distributions, power is relatively low, even for a large number of treatments. Coverage rates of the confidence interval for the intra-class correlation coefficient are far from nominal for very heavy tailed, non-normal random effect distributions.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am grateful to my major professor Dr. Paul Nelson for his enormous guidance, support and encouragement throughout this report. I was very fortunate to be a student of him and learn from his vast experience and expertise. My special thanks goes to Dr. Leigh Murray and Dr. Gary Gadbury for being my MS report committee members and for providing me with valuable feedback on this report. I would like express my appreciation to Dr. James Neill and the faculty in Department of Statistics for their guidance and support during my graduate studies and to Ms. Pamela Schierer and Ms. Angie Ladner for all their help given to me. Also, thank you all my friends in the Department of Statistics for your support in many ways and my husband for always being beside me and encouraging me to make this a success.

Chapter 1 - Introduction

Random effects models are widely used in disciplines such as agriculture, health sciences, and production engineering. This report will study the robustness of normal theory inference in completely randomized, balanced, one-way random effects models.

Studies used to compare treatments may be broadly classified along the following lines: (1) Inference is desired for a set of treatments specified by the experimenter. (2) The treatments studied are sampled from a population of treatments. Statistical models used in these cases are, called fixed effects models in (1), random effects models in (2). The goal of inference from (2) is to characterize and estimate the magnitude of the variance components. See page 287 of Milliken and Johnson (2009) for a fuller discussion of these issues. Almost all standard methods of making statistical inference from (2) are based on assumptions of independence and normality. But, when analyzing real life data, departures from these assumptions are to be expected. Normality for random effects models is an assumption whose validity can be difficult to check and to which there are few alternatives. Properties of normal based inference from these models have been investigated, but relatively little work has been done on the consequences of departures from normality. Therefore, the robustness of statistical inference from random effects models when normality does not hold is an important issue.

1.1 Completely Randomized, One-way, Balanced Design

A statistical model with a completely randomized, one-way balanced design can be used to model the response of the j^{th} unit in the i^{th} treatment (Y_{ij}) as follows:

$$\mathbf{Y_{ij} = \mu + a_i + e_{ij}} \quad ; i = 1, 2, \ldots., t \ ; j = 1, 2, \ldots \qquad 1.1$$

(1.1)

μ : overall mean

a_i : random effect common to all responses in the i^{th} treatment

e_{ij} : random error associated with the j^{th} unit in the i^{th} treatment

σ_a^2 : among treatment variance component

σ_e^2 : within treatment variance component

$\underline{a}$: $\{a_i\}$ is independent of $\underline{e} = \{e_{ij}\}$; $\{a_i\}$ are iid$(0, \sigma_a^2)$ and $\{e_{ij}\}$ are iid$(0, \sigma_e^2)$,

where $(0, \sigma^2)$ denotes a distribution with location zero and positive scale parameter σ

The total variation of Y_{ij} is $\sigma_a^2 + \sigma_e^2$. The usual assumption of normality stipulates that the errors and treatment effects have normal distributions.

1.1.1 The Significance Test for Among Treatment Variance Component

Since the treatment effects here are a sample from a population of effects, differences among the means of the treatments actually observed are typically not of interest in situations like this. The variability among the treatment effects (σ_a^2) is typically the main focus in random effects models. If σ_a^2 is zero, then all of the variation in the responses is due to the random error. But, if σ_a^2 is greater than zero, some of the variability is due to differences among the treatments. The relevant hypotheses to be tested are given by

$$H_0: \sigma_a^2 = 0 \text{ vs } H_1: \sigma_a^2 > 0.$$

The analysis of variance table shown in Table1.1 summarizes the variability in the observations with entries for sums of squares, degrees of freedom, mean squares and the expected mean squares for the among treatments and the within treatments sources of variation. It is the same table used in a fixed effects analysis.

Table 1.1 Analysis of Variance for the one-way classification with expected mean squares for the random effects model

Source of Variation	Degrees of Freedom	Sums of Square	Mean Square	Expected Mean Square	F_0
Total	t*r-1	SSTotal			
Among Treatment	t-1	SSA	MSA	$r\sigma_a^2 + \sigma_e^2$	MSA/MSW
Within Treatment	t*(r-1)	SSW	MSW	σ_e^2	

Moment estimators of the variance components are expressed as solutions to the following equations:

$$MSA = r\sigma_a^2 + \sigma_e^2, \tag{1.2}$$

$$MSW = \sigma_e^2. \tag{1.3}$$

The solutions are

$$\hat{\sigma}_e^2 = MSW, \tag{1.4}$$

$$\hat{\sigma}_A^2 = \frac{(MSA - MSW)}{r}. \tag{1.5}$$

If $\hat{\sigma}_A^2 < 0$, we typically set $\hat{\sigma}_A^2 = 0$.

The statistic computed from the analysis of variance table to test for the presence of a treatment effect is

$$F_0 = \frac{MSA}{MSW}. \tag{1.6}$$

Under normality, if H_0: $\sigma_a^2 = 0$ holds, the test statistic F_0 has the F distribution with (t-1) and t*(r-1) degrees of freedom. The null hypothesis is rejected at the α level of significance if $F_0 > F_{\alpha,(t-1),(N-1)}$. Again, note that this F-test of H_0: $\sigma_a^2 = 0$ vs H_1: $\sigma_a^2 > 0$ is valid if both <u>a</u> and <u>e</u> are normally distributed.

1.1.2 Intra-class Correlation Coefficient

The intra-class correlation coefficient (ρ_I) has a long history of application in several different fields of research. In epidemiologic research, ρ_I is commonly used to measure the degree of familial resemblance with respect to biological or environmental characteristics. In psychology, it plays a fundamental role in assessing the reliability of raters, where observations may be collected on a sample of r judges, (Donner, 1986). A third area of application is in sensitivity analysis, where ρ_I may be used to measure the effectiveness of an experimental treatment (Bradley and Schumann, 1957).

In general terms, the Intra-class correlation coefficient (ρ_I) is a measure of the variability of observations within treatments relative to that among treatments. It is given by

$$\rho_I = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}\,. \tag{1.7}$$

ρ_I can be estimated by $\hat{\rho}_I = \dfrac{(MSA - MSW)}{\{MSA + (r-1)MSW\}}\,.$ (1.8)

An exact $(1-\alpha)*100\%$ normal theory confidence interval estimator (Searle, Casella & McCulloch, 1992) for ρ_I is given by

$$\frac{F_0 - F_u}{F_0 + (r-1)F_u} < \rho_I < \frac{F_0 - F_l}{F_0 + (r-1)F_l} \tag{1.9}$$

where $F_u = F_{1-\alpha/2,(t-1),(N-1)}$ and $F_l = F_{\alpha/2,(t-1),(N-1)}$

The confidence interval in (1.9) for the intra-class correlation (ρ_I) defined in 1.7 is valid if both $\underline{a}$ and $\underline{e}$ are normally distributed. Studying the behavior of this interval estimate for ρ_I when both or either $\underline{a}$ and $\underline{e}$ are not normally distributed is one of the important applied questions that motivated this report.

1.1.3 Illustration

Source: Example 5.1 of Kuehl (2000).

An experiment was conducted to assess the variability in the tensile strengths of bars made from high-temperature castings of an alloy taken from three randomly selected castings at the same facility. Each casting was broken down into individual bars. Destructive tensile strength measurements were obtained on a random sample of 10 bars from each of three castings. The tensile strengths for each of the 30 bars in pounds per square inch (psi) are given in Table 1.2.

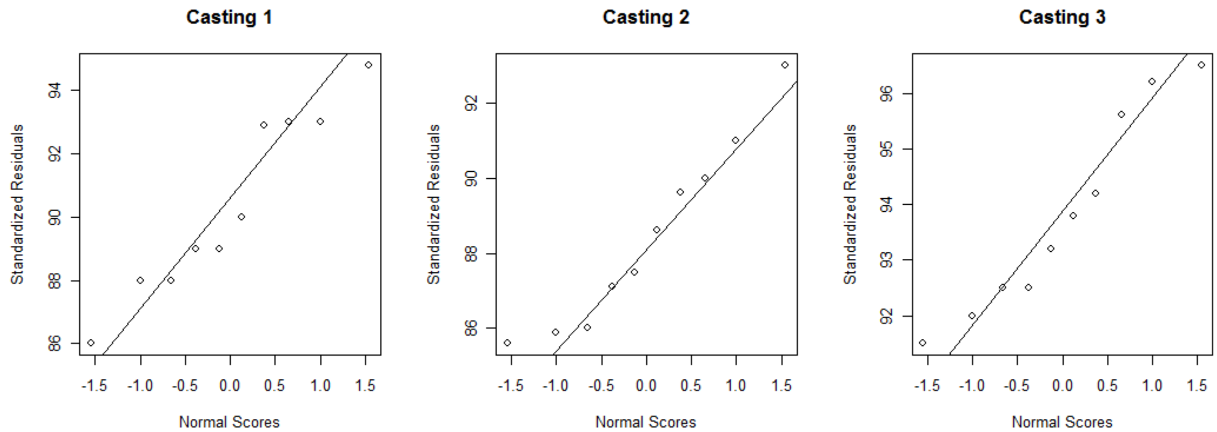
Table 1.2 Tensile strengths of bars from three separate castings of a high-temperature alloy

Castings		
1	**2**	**3**
88.0	85.9	94.2
88.0	88.6	91.5
94.8	90.0	92.0
90.0	87.1	96.5
93.0	85.6	95.6
89.0	86.0	93.8
86.0	91.0	92.5
92.9	89.6	93.2
89.0	93.0	96.2
93.0	87.5	92.5

Original source: G.J. Han and T.E. Rag Nathan (1988), which combined various other sources

The castings are treated as random effects here since they are randomly selected from a population of castings. The investigators were interested in studying the variation in tensile strength among castings produced by the facility. Also, the individual bars are a random sample of bars from a single casting and differences among them are also a random effect. Thus, there are two sources of variations present in this experiment, variation in castings and variation in bars within a casting.

Figure 1.1 Normal Probability Plots



The points on the normal probability plots of casting two and casting three closely follow a straight line. Therefore, these plots do not provide evidence of a departure from normality of tensile strengths from castings two and three. Although the points in the plot from casting one do not follow a straight line as closely as the other two plots, the departure from linearity does not appear to be big enough to raise questions about the assumption of normality.

If the among castings variance (σ_a^2) is zero, then all of the variation in tensile strengths in bars produced at this facility is due to differences among bars. But, if $\sigma_a^2 > 0$, some of this variability is due to differences among the castings. Since the casting effects here are a sample from a population of effects, the differences between the casting means are typically not of interest and as stated above, it is the variability among the casting effect (σ_a^2) that is the main interest in a situation like this. See page 151 of Kuehl (2000) for a further discussion of this issue. Table 1.3 gives the analysis of variance table for the tensile strength data in Table 1.1.

Table 1.3 The Analysis of Variance table for the tensile strength data under normality

Source of Variation	Degrees of freedom	Sums of Square	Mean Square	F	Pr > F
Total	**29**	**304.99**			
Among Treatment	**2**	**147.88**	**73.94**	**12.71**	**0.0001**
Within Treatment	**27**	**157.10**	**5.82**		

The moment estimate of the variance component for bars within castings is $\hat{\sigma}_e^2$ = MSW = 5.82 and the moment estimate of the variance component for among castings is $\hat{\sigma}_a^2$ = (MSA-MSW)/10 = 6.81. The estimated total variance of an observation on tensile strength is then given by $\hat{\sigma}_y^2 = \hat{\sigma}_a^2 + \hat{\sigma}_e^2$ = 12.63. The F_0 ratio to test H_0: $\sigma_a^2 = 0$ vs H_1: $\sigma_a^2 > 0$ is 12.71. Under the assumption of normality, the null hypothesis would be rejected with a p-value at most equal to 0.0001. Therefore, it could then be concluded that differences among castings contribute to the variation in the tensile strengths of bars. The estimate of intra-class correlation for castings of high-temperature alloys is

$\hat{\rho}_I$ = (73.94-5.82) / {73.94+ 9*(5.82)} = 0.54. With F_0 = 12.71, $F_u = F_{.025,2,27}$ = 4.242 and $F_l = F_{.975,2,27}$ = 0.025, the 95% normal theory confidence interval estimate for ρ_I is (0.17, 0.98).

In the past, most studies of robustness in this setting have dealt with the one-way fixed effects model. That is, taking {a_i, i = 1, 2,…., t} to be constants and defining $\sigma_a^2 = \sum_{i=1}^{t} \frac{(a_i - \bar{a})^2}{(t-1)}$, so that the question of normality pertains only to error terms {e_{ij}}.

Inkyung Jung and Pranab Kumar Sen (2008) conducted a simulation study to compare the behavior of different tests based on normality and a non-parametric test that they proposed for testing the significance of the among treatment variance component. A relatively small number of treatments and different distributions for the random effects were used in this study. It was found that the classical F-test is the most powerful for normal random errors in the homoscedastic case, regardless of the distribution of the random effect. But, it is not powerful under other distributions of the random errors. This study suggests that their proposed test is robust against non-normality of random errors in terms of the actual significance level and well behaved in terms of the power in the unbalanced case.

1.2 Aims and Objectives

The objective of my simulation study is to explore how the classical normal theory F-test for the treatment variance component in the random effects model (1.1) described above performs in terms of the actual significance level and the power of the test when the treatment components $\{a_i\}$ and/or the within treatment terms $\{e_{ij}\}$ are not normally distributed. Additionally, attention will be given to the coverage rates and median lengths of the exact normal theory confidence interval for the intra-class correlation coefficient. The impact of the number of treatments and the number of observations per treatment will also be considered.

Chapter 2 - Simulation Procedure

R software was used to carry out the simulation study of the size and power of the F-test of $H_0:\sigma_a^2=0$ vs $H_1:\sigma_a^2>0$, the coverage rate and the median width of the confidence interval for the intra-class correlation coefficient ρ_I under a variety of conditions when the assumption that both the random effects $\{a_i\}$ and error terms $\{e_{ij}\}$ are normally distributed is not satisfied. Without loss of generality, the error variance component (σ_e^2) was taken to be one and the overall mean (μ) set equal zero throughout the entire simulation study. Hence, σ_a^2 represents the ratio of the among to the within variance components.

The independent variables of the factorial design used in my simulation experiment were: the number of treatments (t), the number of replications per treatment (r), among treatment variance component (σ_a^2) and the distributions of the random effects and error terms. The levels of the factors I used are given below. I relied on the random number generator used by R to provide an effective randomization of all 1640 treatment combinations in my study.

The values chosen for the parameters are listed below.

Table 2.1 Levels of Factors Used in Simulation Study

Parameter	**Values**
t	5, 10, 50, 100, 500
r	5, 10
σ_a	0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5
σ_e^2	1
μ	0

2.1 Densities Used for Both Random Effects

- Normal

$$f(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

μ : location parameter, σ : scale parameter.

- Logistic

$$f(x;\mu,\theta)=\frac{\exp\{-(x-\mu)/\theta\}}{\theta(1+\exp\{-(x-\mu)/\theta\})^2},$$

μ : location parameter, θ : scale parameter.

- Double Exponential

$$f(x;\mu,\beta)=\frac{1}{2\beta}\exp\left\{\frac{|x-\mu|}{\beta}\right\},$$

μ : location parameter, β : scale parameter.

- Cauchy

$$f(x;x_0,\gamma)=\frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right],$$

x_0: location parameter, γ : scale parameter.

The values for the scale parameters of the logistic, the double exponential and the Cauchy were selected such that these distributions have the same inter-quartile range as the corresponding normal distribution. The location parameters were all set to zero.

Table 2.2 Inter-Quartile Range for the Distributions in terms of the Scale Parameter

Distribution	**Inter-quartile range**
Normal	1.349σ
Logistic	2.197θ
Double Exponential	$2\ln(2)\beta$
Cauchy	2γ

Table 2.3 Values taken for the Scale Parameters in Order to Have Same Inter-quartile Range as the Corresponding Normal Distribution

Distribution	**Scale**
Logistic	$\theta = 1.349\sigma / 2.197$
Double Exponential	$\beta = 1.349\sigma / 2\ln(2)$
Cauchy	$\gamma = 1.349\sigma / 2$

2.1.1 Generating data from Double Exponential Distribution

The R software package does not have a direct function for generating random numbers from a double-exponential distribution. Instead, I implemented an algorithm based on the following facts.

A random variable X has a double-exponential (0, 1) distribution if its probability density function is

$$f(x) = \begin{cases} \frac{1}{2}e^{x}, & x < 0 \\ \frac{1}{2}e^{-x}, & x \geq 0 \end{cases}$$

The CDF of f(x)

For $x < 0$,

$$F(x) = \int_{-\infty}^{x} \frac{1}{2} e^{t} \, dt$$

$$= \frac{1}{2} [e^{t}]_{-\infty}^{x}$$

$$= \frac{1}{2} [e^{x} - e^{-\infty}]$$

$$= \frac{1}{2} e^{x} \; ; \; -\infty < x < 0$$

For $x \geq 0$,

$$F(x) = \int_{0}^{x} \frac{1}{2} e^{-t} \, dt + \frac{1}{2}$$

$$= \frac{1}{2} [-e^{-t}]_{0}^{x} + \frac{1}{2}$$

$$= \frac{1}{2} [-e^{-x} - (-e^{0})] + \frac{1}{2}$$

$$= \frac{1}{2} [1 - e^{-x}] + \frac{1}{2}$$

$$= 1 - \frac{1}{2} e^{x} \qquad ; 0 \leq x < \infty$$

Let the random variable U~Uniform(0,1), then

U=F(X) => X = F^{-1}(U), as given by

$$U = \frac{1}{2} e^{X} \; ; \; -\infty < X < 0$$

$$\therefore X = \ln(2U) \; ; 0 < U < \tfrac{1}{2} \tag{2.1}$$

$$U = 1 - \frac{1}{2} e^{-X} \qquad ; 0 \leq X < \infty$$

$$\therefore X = -\ln 2(1 - U) \qquad ; \tfrac{1}{2} \leq U < 1 \, . \tag{2.2}$$

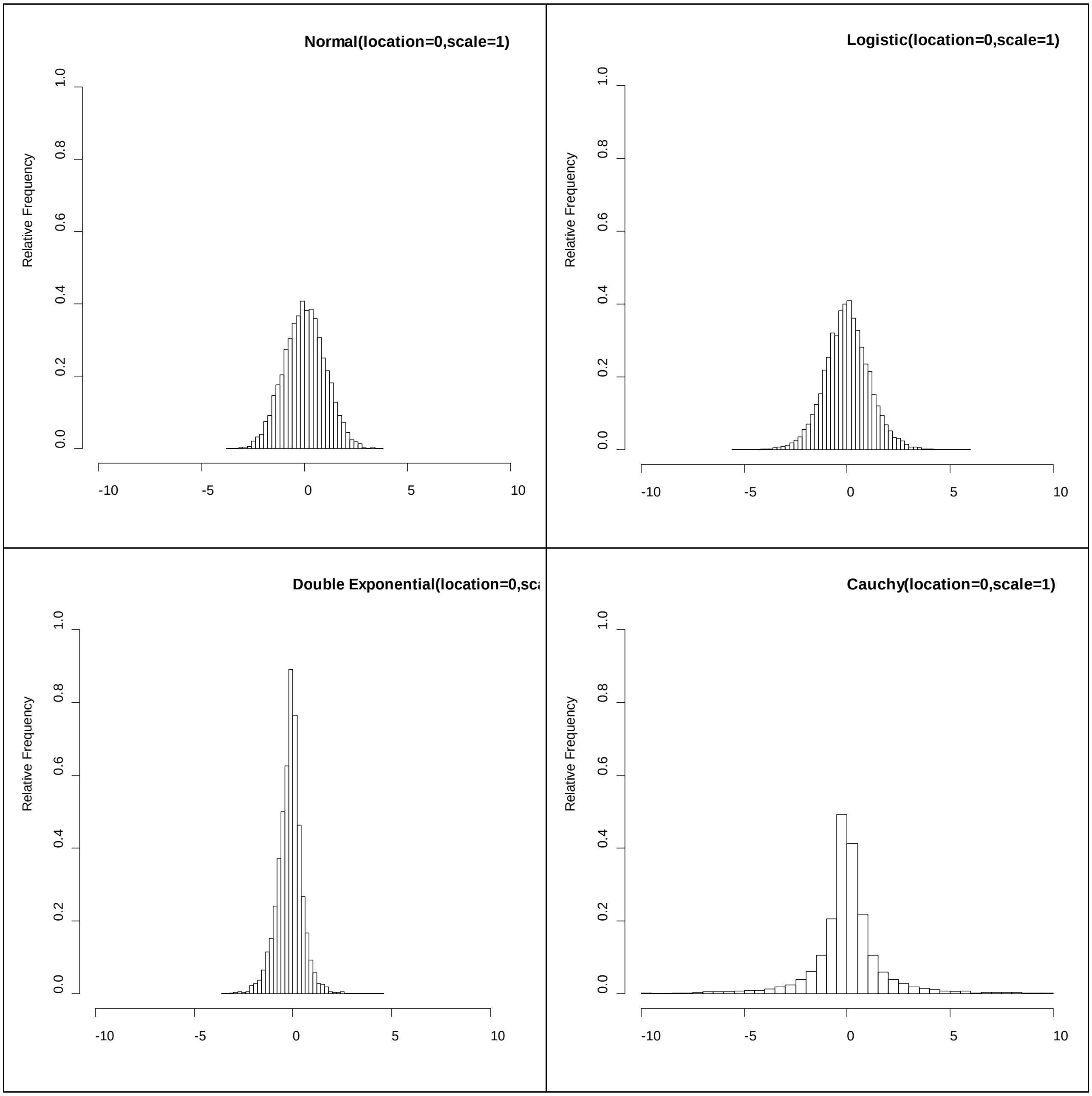
2.1.1.1 *Algorithm: Generate random numbers from double-exponential (location=0, scale) distribution*

- Generate random numbers $\{U_i\}$ from a uniform (0, 1)
- For $0 < U < \frac{1}{2}$ and, $\frac{1}{2} \leq U < 1$, use (1) and (2) respectively to obtain X
- Transform X to $\frac{1.349\sigma}{2\ln(2)}X$ so that it has the same inter-quartile range as the

  Normal (0, σ^2) distribution
- Transformed values ~double-exponential (location=0, scale=$\frac{1.349\sigma}{2\ln(2)}$)

To illustrate the procedures I used to generate data, ten thousand random numbers were generated from each normal, logistic, double exponential and Cauchy distributions and the corresponding histograms are shown Figure 2.1 below. Only data in the interval from -10 to 10 (9336 values) was used in constructing the histogram of Cauchy distribution to make for a less distorted comparison with the other histograms.

Figure 2.1 Histograms of 10,000 random numbers generated from Normal, Logistic, Double-Exponential and Cauchy distributions



Compared to the normal distribution, the logistic distribution has heavier tails. The double-exponential distribution is more peaked and has heavier tails than the normal distribution. Simulating the Cauchy distribution resulted in data with a minimum of -3028.698 and maximum of 3185.403. Like the normal, all three non-normal distributions are symmetric and all three have heavier tails than the normal. Data generated from the logistic, double-exponential and Cauchy distributions were used to explore the behavior of inferences based on the assumption of normality as described in Chapter 1.

One thousand independent sets of data were generated for each combination of parameter values and the distribution of the two random effects. Then, power of the normal theory *F* test, coverage rate and the median width for the confidence interval for the intra-class correlation coefficient were estimated for each combination and stored in a file.

2.2 Power Calculation

The number of times the *F*-test rejected H_0: $\sigma_a^2 = 0$ at the 5% nominal significance level out of 1000 data sets was recorded at each parameter combination. Then, the average number of rejections was calculated for each treatment combination and it was taken as the estimated power of the *F*-test for the case where σ_a^2 values are greater than zero and as the estimated actual size of the test for the case where σ_a^2 is equal to zero.

2.3 Coverage Rates and Widths of Confidence Intervals for the Intra-class Correlation Coefficient

An exact $(1-\alpha)*100\%$ normal theory confidence interval estimator for the Intra-class Correlation Coefficient (ρ_I) is given by

$$\frac{F_0 - F_u}{F_0 + (r-1)F_u} < \rho_I < \frac{F_0 - F_l}{F_0 + (r-1)F_l}, \qquad (2.3)$$

where F_o: Observed test statistic for the F-test,

F_u: $\mathrm{F}(\alpha, t-1, N-t)$,

F_l: $(1-\alpha/2, t-1, N-t)$.

2.3.1 Calculate the Coverage Rate

According to (1.8), there might be situations where ρ_I estimate is negative and according to (1.9) the lower limit of the confidence interval is negative or both lower and upper limits are negative. But by definition (1.7), ρ_I is nonnegative. One course of action suggested by (Searle, 1970, p.407) is to replace the negative ρ_I estimates by zero. Therefore, the following steps were taken when estimating the coverage rate for the confidence interval for the intra-class correlation coefficient.

- If the upper limit of the confidence interval is negative for a particular data set, then that data set was deleted from further analysis.
- Among the remaining data sets, if the lower limit of the confidence interval is negative, the lower limit was set to zero.
- Next, the number of times the population intra-class correlation coefficient falls within the nominal 95% confidence interval for the intra-class correlation for remaining number of data sets of each design point was recorded. Then, the average number of times the population intra-class correlation coefficient falls within the confidence interval was calculated for each design point and was considered as the estimated coverage rate.

2.3.2 Width

- Similarly, widths of the 95% confidence interval for the intra-class correlation were calculated for each data set remaining after deleting the data sets which results negative upper limit for the confidence interval and the average width was computed for each parameter combination. Since the distributions of the confidence interval widths were highly skewed in some cases, the median width instead of the mean width was used to summarize the widths obtained for each parameter-distribution setting.

Chapter 3 - Simulation Results

The results of the simulated samples of data for each of the parameter and distribution combinations for the two random effects are presented and discussed in this chapter. Results only five representative values of σ_A (0, 0.5, 1, 1.5, 2 and 2.5) are presented in the Tables in this chapter and in the Appendix A because the general pattern is visible from a portion of σ_A values considered in the study.

3.1 Size of F-test for the Significance of the Among Treatments Variance Component

The F-test was conducted using a nominal size of 0.05 for the simulated data for each design point. The estimated actual size of the F-test for different random effect distributions for different numbers of treatments and replications are shown in the Table 3.1. The standard errors of the estimates in Table 3.1 are bounded above by $\sqrt{\frac{0.07(1-0.07)}{1000}} = 0.0081$.

Table 3.1 Actual Size of the F-test for Different Distributions of Random Error (Nominal Size = 0.05)

Number of Treatments	Replicates per Treatment	Actual Size			
		eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy
5	5	0.055	0.061	0.036	0.014
5	10	0.049	0.048	0.049	0.018
10	5	0.045	0.041	0.034	0.018
10	10	0.039	0.054	0.048	0.021
50	5	0.062	0.048	0.043	0.017
50	10	0.055	0.051	0.05	0.02
100	5	0.054	0.044	0.048	0.016
100	10	0.052	0.042	0.058	0.017
500	5	0.058	0.059	0.043	0.022
500	10	0.053	0.063	0.047	0.011

As expected, when the random error is normally distributed the actual size of the F-test is close to the nominal value of 0.05 in all of the situations considered. An interesting fact to observe from this table is that when random error has a Cauchy distribution, the actual size of the test is very conservative. In the situations where random error has a logistic or double exponential distribution, the test has an acceptable size close to 0.05.

3.2 Power Analysis

Table 3.2 consists of the estimated power values for 5 and 100 treatments, and 5 and 10 replications per treatment, for different σ_A values ranging from 0.5 to 2.5. Table 3.2 shows that for a fixed number of treatments and replications per treatment, power increases with σ_A in all cases. It can be observed that double exponential errors give higher power compared to the other distributions of random errors irrespective of the treatment effect distribution. The *F*-test has a lower power on average for each σ_A value when the random error is distributed as Cauchy compared to other error distributions while the treatment effect is either normal, logistic or double exponential. This is consistent with the results in Table 3.1 where it was seen that the test was very conservative when the error terms had a Cauchy distribution. But, when both random effects are Cauchy, power is somewhat higher. According to Table 3.2 it can be seen that lowest powers here resulted with double exponential treatment random effects and Cauchy random errors. Also, this table shows that power of the test approaches one as σ_A increases, all other factors remaining fixed except for the case where random errors are Cauchy. It can be observed from Table 3.2, for 100 treatments, that except for the case where the random error is Cauchy, other combinations of distributions considered for both random effects give almost perfect power here. When the random errors are Cauchy, an interesting observation is that for Cauchy treatment random effects, the *F*-test performs better than for normal treatment effects.

Table 3.2 Estimated Power for the F-test for Treatments= (5, 100) and Replications = (5, 10)

Number of Treatments	Replicates per Treatment	σ_A	Power															
			ai ~ Normal				ai ~ Logistic				ai ~ Double Exponential				ai ~ Cauchy			
			eij ~Normal	eij ~Logistic	*eij ~Double Exponential*	**eij ~Cauchy**	eij ~Normal	eij ~Logistic	*eij ~Double Exponential*	**eij ~Cauchy**	eij ~Normal	eij ~Logistic	*eij ~Double Exponential*	**eij ~Cauchy**	eij ~Normal	eij ~Logistic	*eij ~Double Exponential*	eij ~Cauchy
5	5	0.5	0.317	0.24	*0.739*	**0.039**	0.371	0.304	*0.755*	**0.053**	0.112	0.104	*0.312*	**0.031**	0.649	0.635	*0.652*	0.253
5	5	1	0.756	0.724	*0.96*	**0.136**	0.777	0.726	*0.962*	**0.199**	0.301	0.287	*0.621*	**0.065**	0.891	0.847	*0.851*	0.382
5	5	1.5	0.918	0.89	*0.989*	**0.26**	0.924	0.888	*0.994*	**0.291**	0.483	0.476	*0.78*	**0.106**	0.938	0.943	*0.95*	0.486
5	5	2	0.973	0.952	*0.996*	**0.396**	0.977	0.943	*0.998*	**0.403**	0.64	0.61	*0.862*	**0.156**	0.975	0.967	*0.972*	0.573
5	5	2.5	0.991	0.977	*0.997*	**0.453**	0.986	0.977	*0.998*	**0.481**	0.695	0.68	*0.903*	**0.202**	0.993	0.986	*0.979*	0.658
5	10	0.5	0.582	0.513	*0.882*	**0.062**	0.615	0.551	*0.906*	**0.07**	0.211	0.184	*0.485*	**0.027**	0.789	0.769	*0.986*	0.235
5	10	1	0.915	0.873	*0.994*	**0.177**	0.927	0.912	*0.99*	**0.217**	0.502	0.458	*0.756*	**0.053**	0.945	0.921	*0.995*	0.452
5	10	1.5	0.981	0.98	*0.999*	**0.29**	0.978	0.974	*0.999*	**0.308**	0.683	0.633	*0.874*	**0.095**	0.985	0.978	*0.998*	0.533
5	10	2	0.988	0.986	*1*	**0.389**	0.995	0.991	*0.999*	**0.427**	0.802	0.76	*0.911*	**0.139**	0.996	0.989	*0.998*	0.608
5	10	2.5	0.998	0.996	*0.999*	**0.467**	0.997	0.993	*1*	**0.508**	0.842	0.811	*0.949*	**0.22**	0.993	0.993	*0.999*	0.681
100	5	0.5	0.999	0.996	*1*	**0.021**	1	0.999	*1*	**0.027**	0.596	0.492	*0.996*	**0.017**	1	1	*0.755*	0.478
100	5	1	1	1	*1*	**0.043**	1	1	*1*	**0.045**	0.998	0.992	*1*	**0.016**	1	1	*0.938*	0.699
100	5	1.5	1	1	*1*	**0.102**	1	1	*1*	**0.124**	1	1	*1*	**0.036**	1	1	*0.978*	0.782
100	5	2	1	1	*1*	**0.19**	1	1	*1*	**0.206**	1	1	*1*	**0.039**	1	1	*0.988*	0.815
100	5	2.5	1	1	*1*	**0.271**	1	1	*1*	**0.325**	1	1	*1*	**0.075**	1	1	*0.997*	0.873
100	10	0.5	1	1	*1*	**0.022**	1	1	*1*	**0.023**	0.954	0.865	*1*	**0.015**	1	1	*0.996*	0.5
100	10	1	1	1	*1*	**0.052**	1	1	*1*	**0.057**	1	1	*1*	**0.027**	1	1	*0.999*	0.694
100	10	1.5	1	1	*1*	**0.132**	1	1	*1*	**0.145**	1	1	*1*	**0.028**	1	1	*0.997*	0.778
100	10	2	1	1	*1*	**0.2**	1	1	*1*	**0.216**	1	1	*1*	**0.05**	1	1	*0.999*	0.847
100	10	2.5	1	1	*1*	**0.283**	1	1	*1*	**0.34**	1	1	*1*	**0.093**	1	1	*1*	0.875

[1]

[1] Boldface numbers indicate low estimated power values in Cauchy random error cases.

Observing the power curves for all the combinations of distributions of the treatment random effect and the random error in Appendix A, it can be seen that power increases when the number of treatments increases, the other factors remaining fixed, except for the cases where the random error is Cauchy and random effect is either normal, logistic or double exponential. Figure 3.1 below shows the power curves for the *F*-test for differing numbers of treatments when the treatment random effect and the random error are both normal. It can be seen that the rate at which power approaches one, increases as the number of treatments increases. Figure 3.2, Figure 3.3 and Figure 3.4 show the power curves for the *F*-test when the random error is Cauchy and the treatment random effect is normal, logistic and double exponential for different combinations of treatments and replications per treatment. From Figures 3.2, 3.3 and 3.4 it can be observed that contrary to what would be desired, the powers of the *F*-test for different σ_A 's actually decrease when the number of treatments increases and the number of replicates is held fixed.

Figure 3.1 Power Curves for the F-test when Treatment Random effect and Random Error are both Normal for Fixed Number of Replications and Varying Number of Treatments

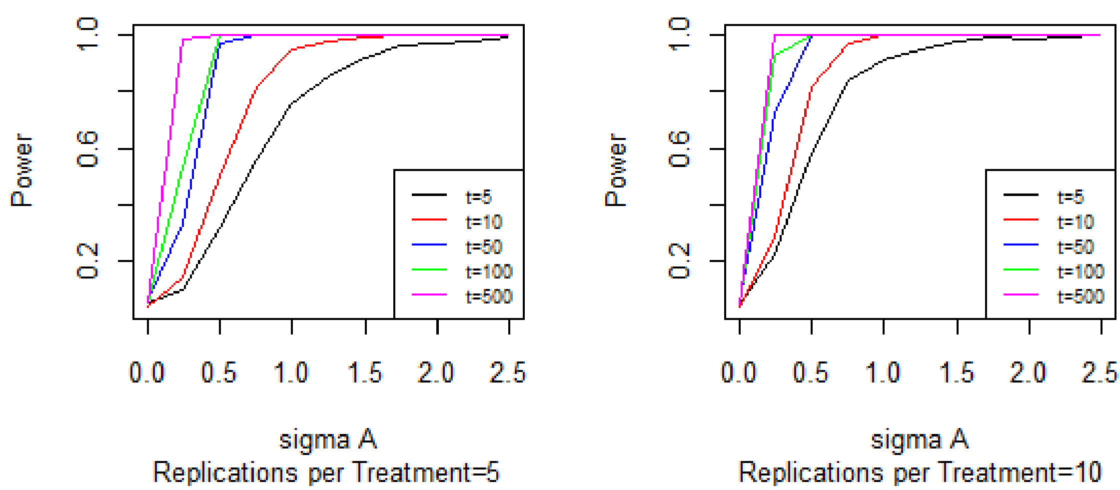


Figure 3.2 Power Curves for the F-test when Treatment Random Effect is Normal and Random Error is Cauchy Fixed Number of Replications and Varying Number of Treatments

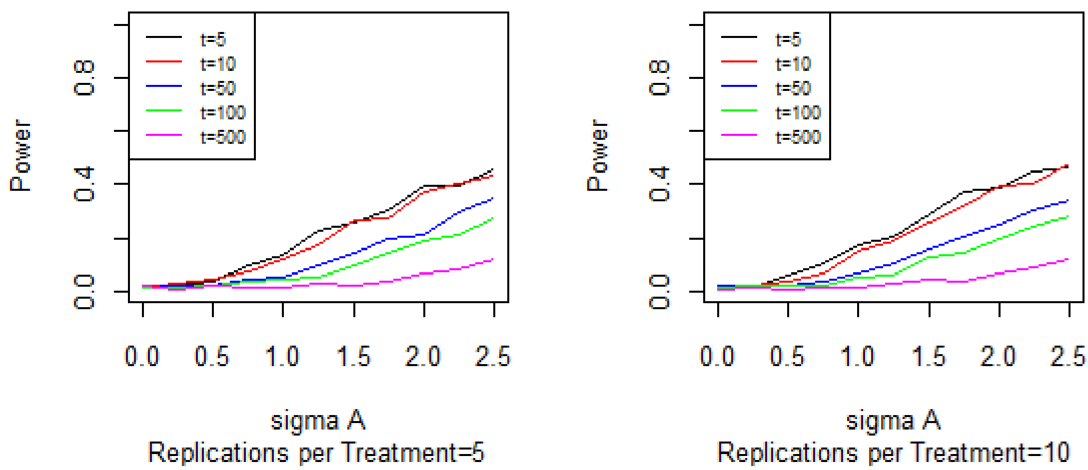


Figure 3.3 Power Curves for the F-test when Treatment Random Effect is Logistic and Random Error is Cauchy Fixed Number of Replications and Varying Number of Treatments

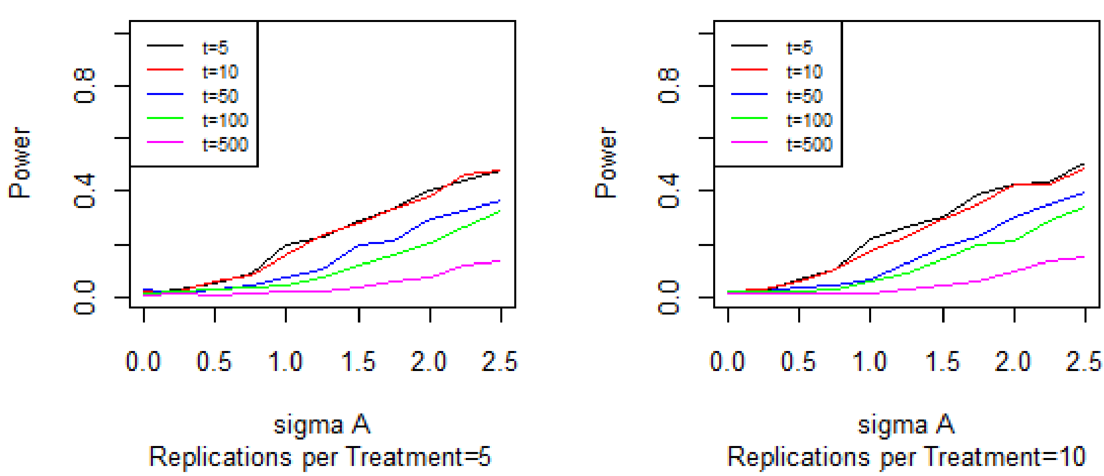


Figure 3.4 Power Curves for the F-test when Treatment Random Effect is Double Exponential and Random Error is Cauchy Fixed Number of Replications and Varying Number of Treatments

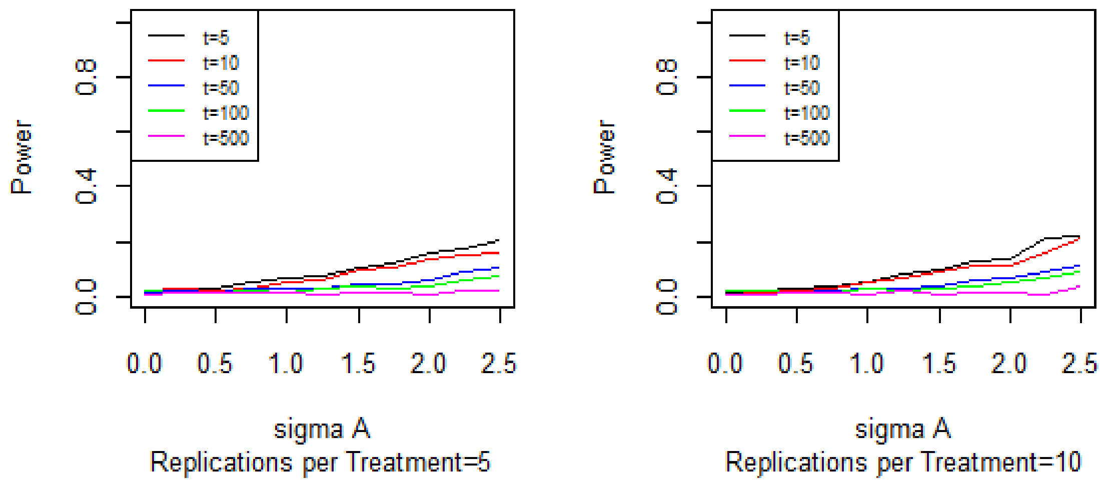


Figure 3.5 shows the power curves for the case where 5 and 10 number of replications while number of treatments is fixed when both random effects are normal. Figures 3.6, 3.7and 3.8 display the power curves for the case where 5 and 10 number of replications for fixed number of treatments when random error is Cauchy and treatment random effect is normal, logistic and double exponential respectively. Observation of Figure 3.5 indicates that when the number of replications increases the power also increases for a fixed number of treatments and power curves for the two replication sizes tend to overlap as the treatment size increases while power values approach to one. This pattern of the power curves is closely followed by all the other random effect combinations of the distributions except for the cases where random error is Cauchy and random effect is either normal, logistic or double exponential. When random error has a Cauchy distribution and random effect has any other distributions considered in this study, power values are approximately equal and very low for 5 replications and for 10 replications when the number of treatments and σ_A are fixed. This can be observed from Figures 3.6, 3.7 and 3.8.

Figure 3.5 Power Curves for the F-test when both Random Effects are Normal for Fixed Number of Treatments and Varying Number of Replications

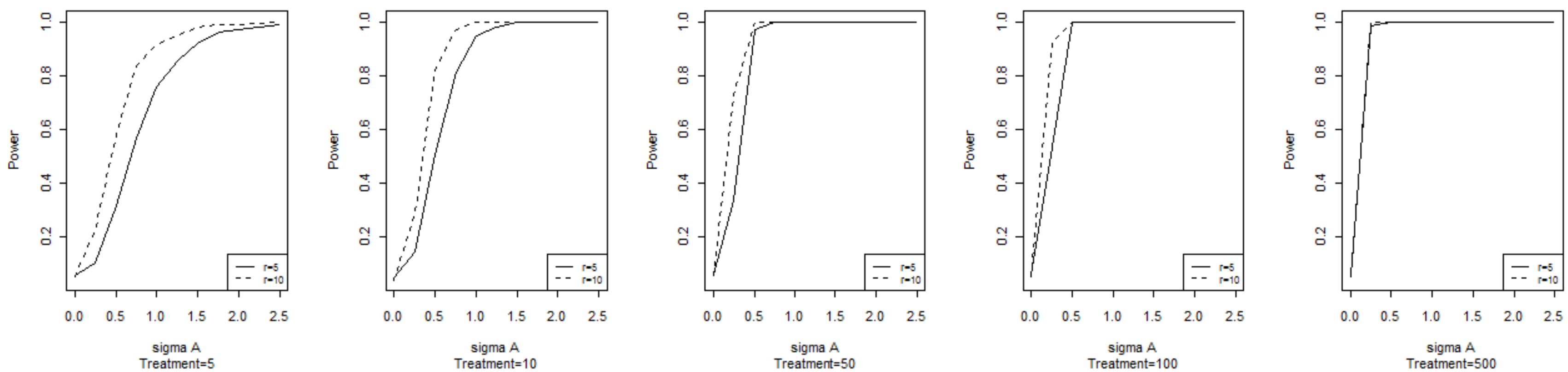


Figure 3.6 Power Curves for the F-test when both Treatment Random Effect is Normal and Random Error is Cauchy for Fixed Number of Treatments and Varying Number of Replications

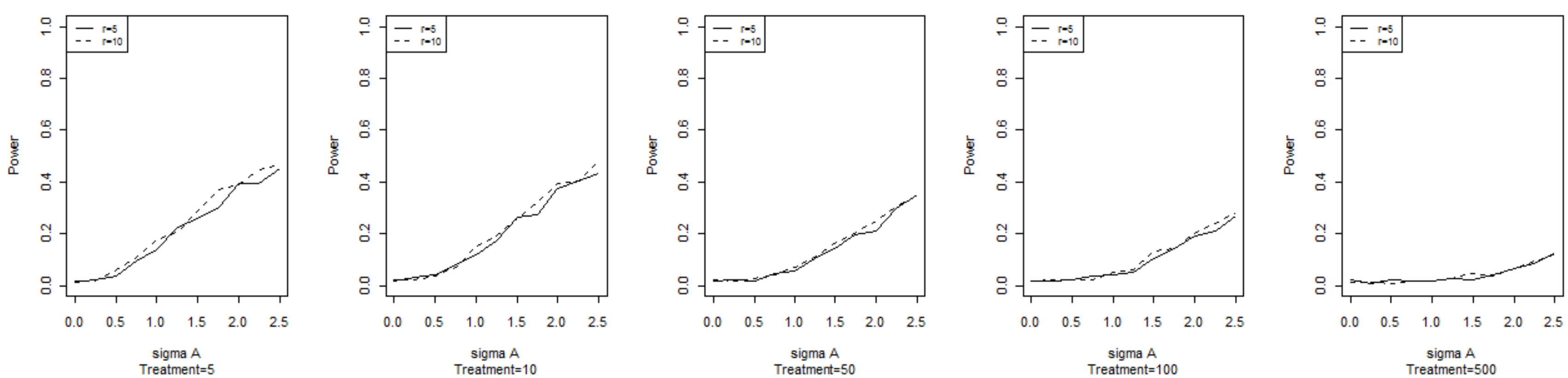


Figure 3.7 Power Curves for the F-test when both Treatment Random Effect is Logistic and Random Error is Cauchy for Fixed Number of Treatments and Varying Number of Replications

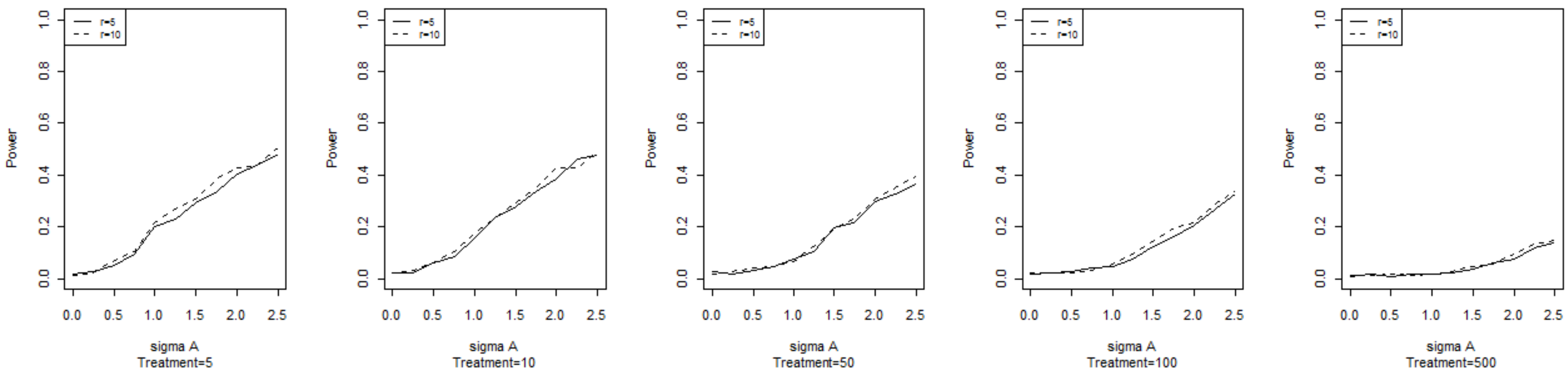
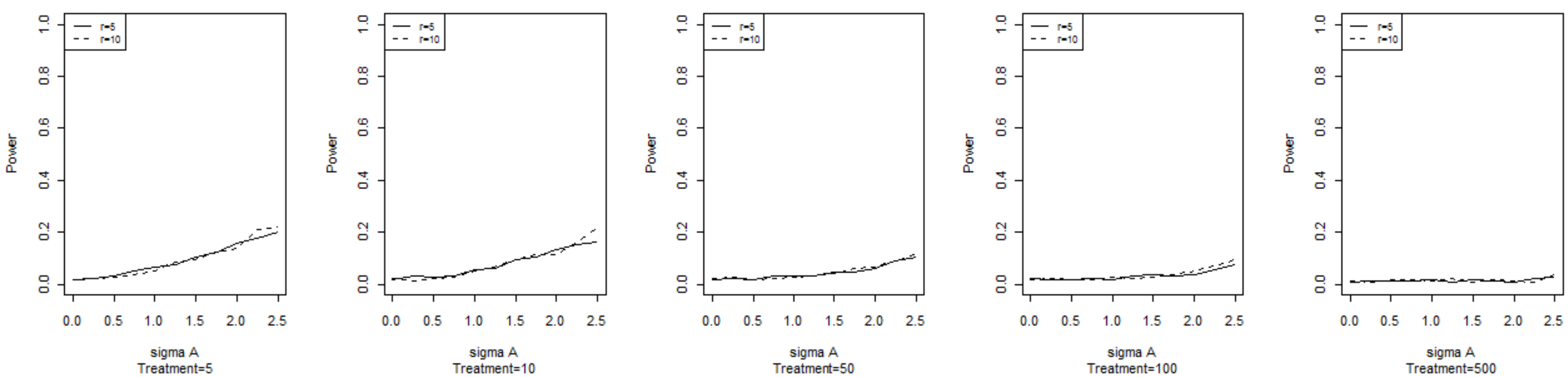


Figure 3.8 Power Curves for the F-test when both Treatment Random Effect is Double Exponential and Random Error is Cauchy for Fixed Number of Treatments and Varying Number of Replications



3.3 Intra-class Correlation Coefficient

A 95% normal theory confidence interval for the intra-class correlation coefficient (ρ_I) was computed for each combination of parameter values and the distributions of random effects. The estimated coverage rates and the median widths of the confidence interval for each design point were obtained. As mentioned in 2.3.1 and 2.3.2, coverage rates and median widths were estimated for the remaining data sets after deleting the datasets which gives negative estimates for both limits of the confidence interval. This deletion rate of data sets that resulted in negative limits for the confidence interval was low for every parameter and distribution combination. Out of the thousand data sets generated for each design point, not more than forty data sets have been deleted in each case due to the negative upper limit of the confidence interval. Appendix A.2 contains the number of data sets used for estimation of coverage rate and width for different design points.

3.3.1 Coverage Rate

The estimated coverage rates of the nominal 95% confidence interval for (ρ_I) closely achieves 0.95 when both random effects are normally distributed (Table A.2). ρ_I is equal to zero at σ_A equals zero and, as mentioned in Chapter 2, when the estimated lower limit of the confidence interval is a negative value, it is set to zero. Hence, the observed hundred percent coverage for the case $\sigma_A = 0$ correspond to cases where all the lower limits were negative. Since there was a decreasing pattern of the coverage rate when the number of treatments were increasing for all design points except when both random effects are normal, only 5 treatments to represent a small treatment size and 100 treatments to represent large treatment size are displayed in this chapter. The estimated coverage rates for the confidence interval when the number of treatments is equal to 5 and 100, and the replications per treatment are 5 and 10 is given in the Table A.2. As a whole, coverage rates have a decreasing pattern when σ_A increases. This is visible in Table A.2 where for fixed number of treatments and replications, coverage rate decreases as σ_A increases from 0.25 to 2.5. Observed coverage rates of approximately 0.95 are obtained when the treatment random effect is normal and the random error is normal or logistic for treatment size equals 5. Logistic treatment effects with normal or logistic errors gave acceptable coverage rates for small number of treatments. For Cauchy treatment effects, coverage is comparatively lower than the normal, logistic or double exponential errors. In the case where large number of treatments is considered, a coverage rate of approximately 0.95 resulted only when both random effects are normal. Also, when both random effects are logistic the coverage rate was around 0.90 on average for 100 treatments. When (i) ai ~ normal and eij ~ logistic (ii) ai ~ logistic and eij ~ normal (iii) both ai and eij ~ double exponential, the coverage rate is moderate for large treatment size. All the other distribution combinations give approximately zero coverage rate when number of treatments is one hundred. All of these observations about the coverage rates are clearly visible from the graphs shown in Figure 3.9 and Figure 3.10.

Figure 3.9 Estimated Coverage Rates for the 95% Confidence Interval for the Intra-class Correlation Coefficient when treatments=5, replications=5

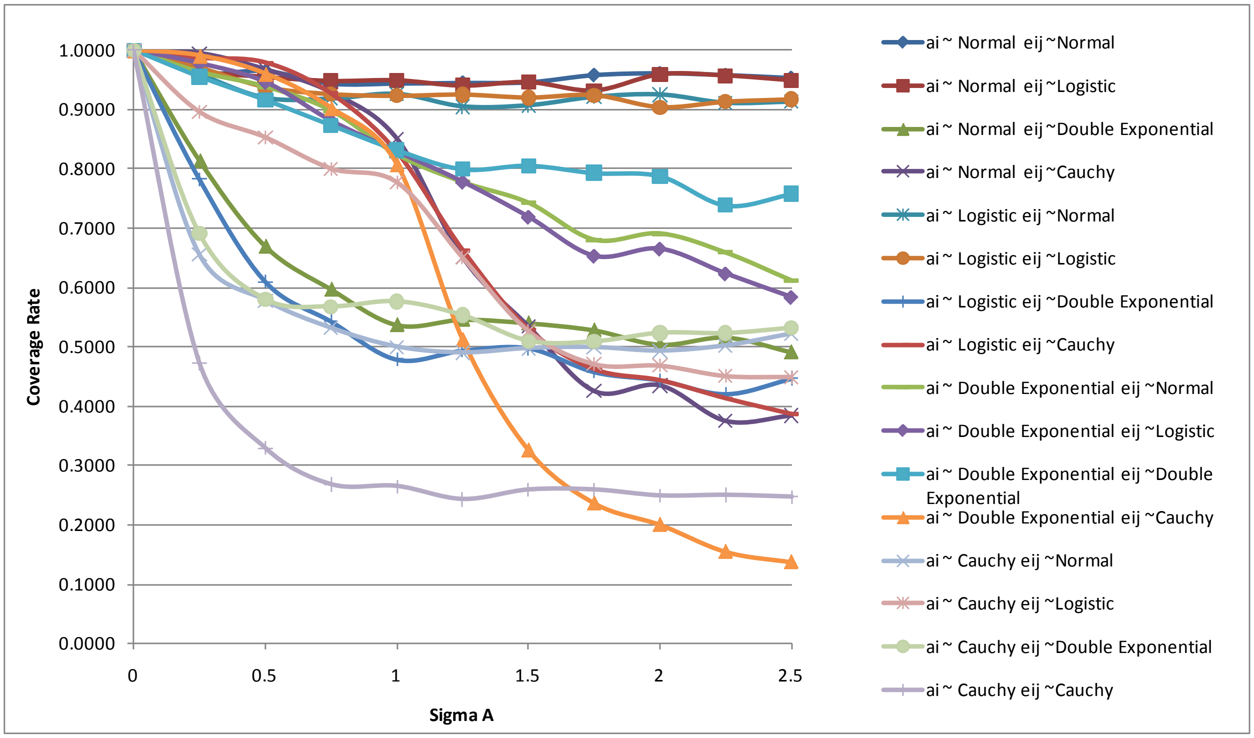
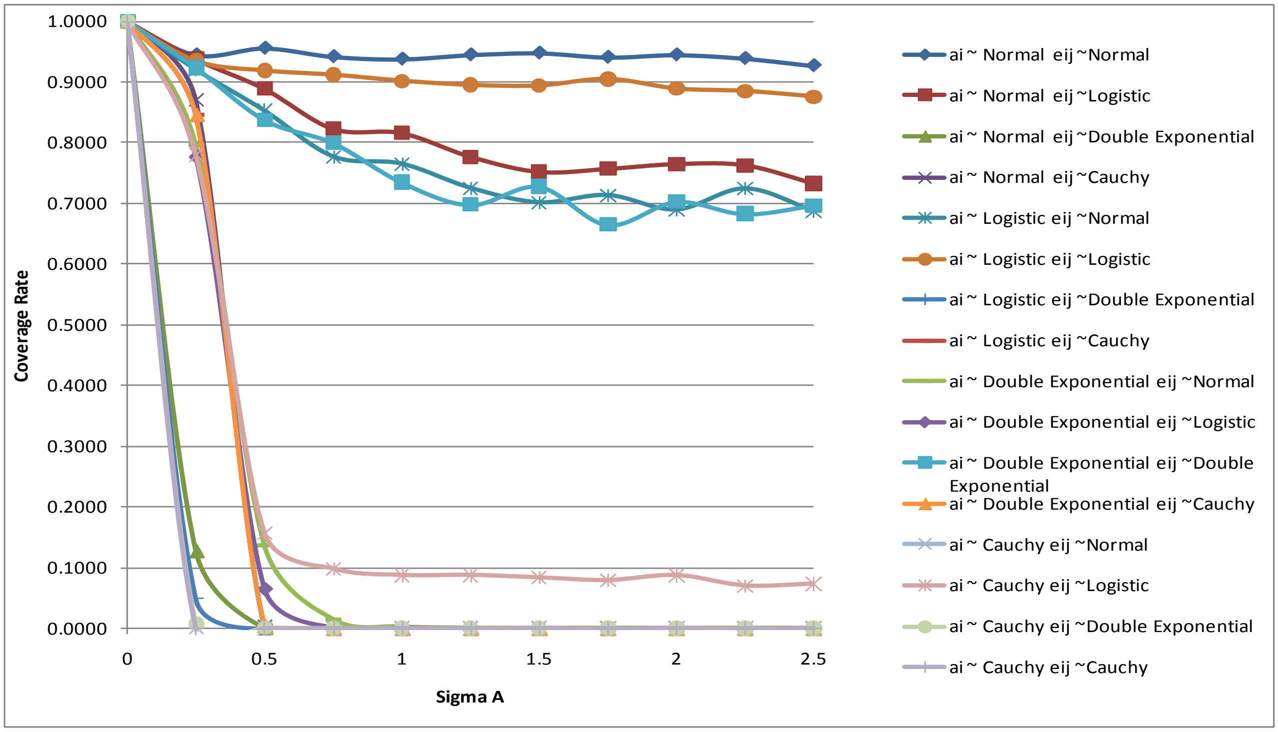


Figure 3.10 Estimated Coverage Rates for the 95% Confidence Interval for the Intra-class Correlation Coefficient When Treatments=100, Replications=5



3.3.2 Width

Median width instead of mean width was used to summarize the widths of the confidence intervals because the distribution of width was highly skewed for some cases. Table A.3 shows estimated median widths when both the treatment random effect and the random error are normal. Also, it shows the relative median widths for different distribution combinations for two random effects. Relative median width is obtained by dividing the median width by the corresponding median width for the same values of t, r, and σ_A with both random effects being normally distributed. A small relative median width is only desirable for a parameter setting if the coverage rate is close to nominal. In many cases described below, small relative median widths actually corresponds to cases where the coverage rate is unacceptably low and therefore represent situations where the uncritical experimenter would incorrectly believe the interval gave precise information about σ_A.

Table A.3 shows that, as a whole, when the treatment random effect is normal or logistic and the random error is double exponential, the median width is narrower compared to the case where both random effects are normal for treatment size is 5 and treatment size is 10. For a Cauchy treatment random effect and either normal, logistic or double exponential random error relative median width is smaller compared to the situation when both normal random effects and this median width decrease with the increase of treatment size. But, when both random effects are Cauchy, the relative widths are larger than when only treatment random effect is Cauchy.

Figure 3.11 shows the relative median widths for different combinations of distributions for the two random effects changes with σ_A for number of treatments equals 5 and number of replications per treatment equals 5. There appear to be three groups of curves with regard to the variation of relative median width. When both a_i and e_{ij} are logistic or both a_i and e_{ij} double exponential, relative median width stays approximately closer to one which indicates that it is not much different from the median width when both a_i and e_{ij} are normal. An upward trend in the relative median width when σ_A increases approximately from 0.5 above can be seen in the cases where e_{ij} is Cauchy and a_i is any of the distributions considered here. In contrast, when e_{ij} double exponential and a_i is normal, logistic or Cauchy a decreasing pattern of the relative median width for increasing values of σ_A can be seen. Figure 3.12 shows the relative median widths for different combinations of distributions for the two random effects changes with σ_A for number of treatments equals 100 and number of replications per treatment equals 5. It appears that upward, downward and straight patterns visible in Figure 3.11 can be observed similarly in Figure 3.12. Relative median width is close to zero for σ_A values greater than 0.5 when a_i is Cauchy and e_{ij} is normal, logistic or double exponential.

Figure 3.11 Relative Median Widths for the 95% Confidence Interval for the Intra-class Correlation Coefficient when treatments=5 and replications per treatment=5

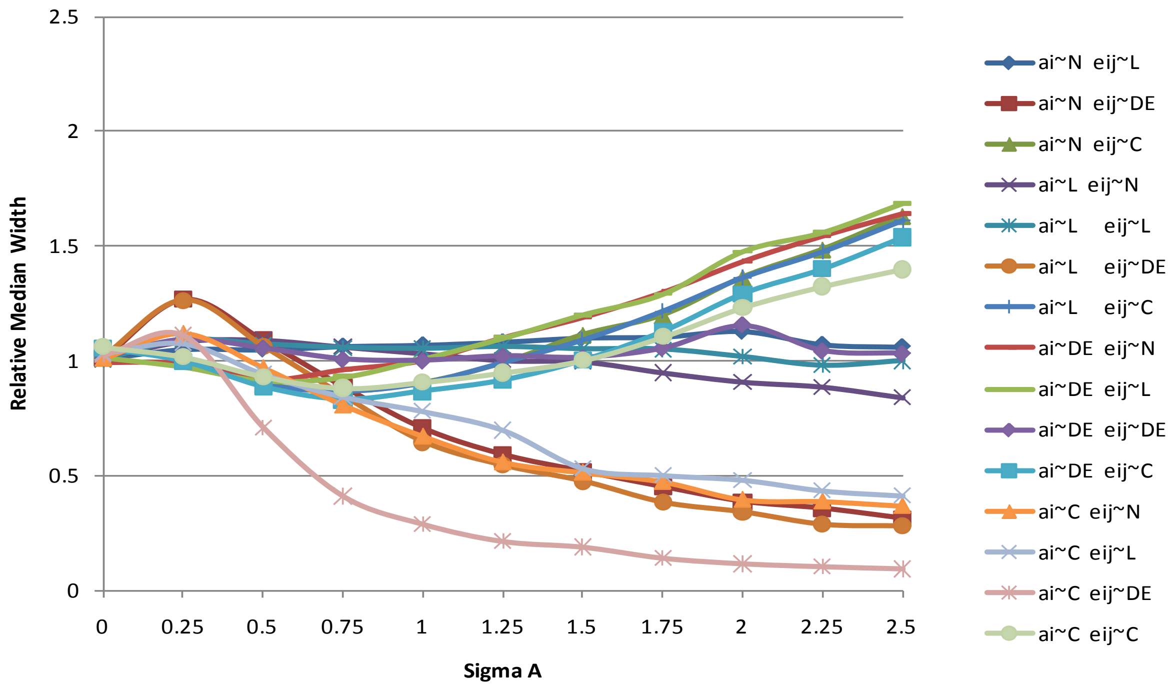
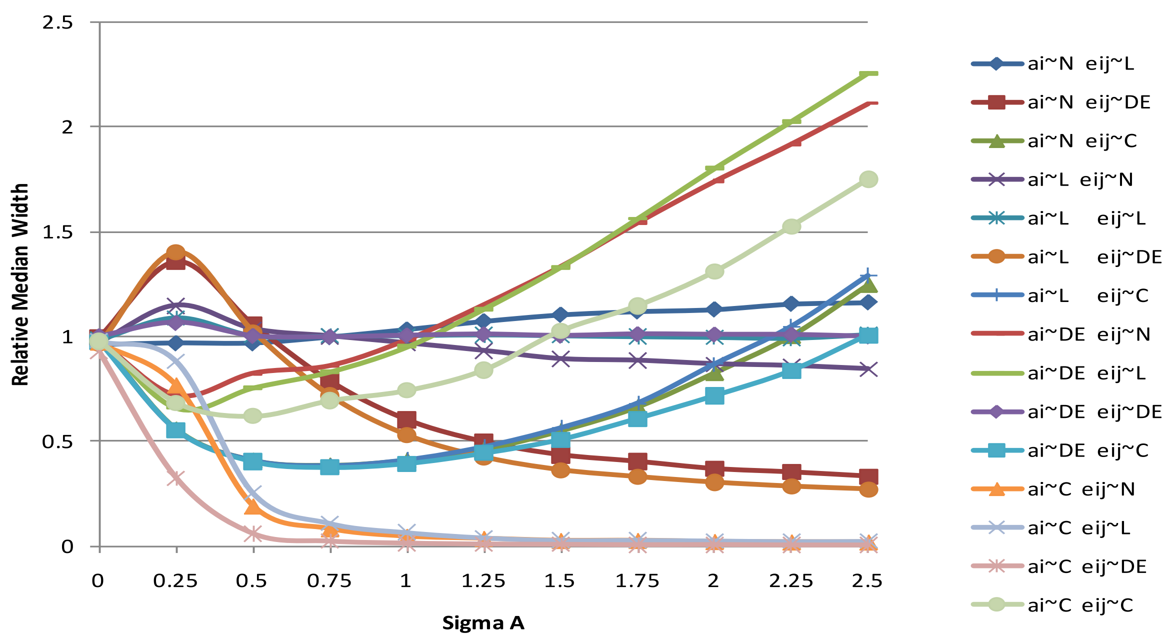


Figure 3.12 Relative Median Widths for the 95% Confidence Interval for the Intra-class Correlation Coefficient when treatments=100 and replications per treatment=5



Chapter 4 - Conclusions

The objectives of this report were to investigate the behavior of the actual size and power of the F-test to test for the presence of an among treatment variance component in a one-way, balanced, random effects model and assess the performance of the normal theory confidence interval for the intra-class correlation coefficient with respect to coverage rate and width, when one or both of the random effects are not normally distributed. A number of different data sets were simulated allowing, the treatment random effect and random error to have a variety of symmetric distributions, the logistic, double exponential, Cauchy and normal.

Analysis of simulated data sets showed that when random error is normal, logistic or double exponential, the actual size of the F-test was close to the nominal significance level. For Cauchy random errors, the F-test was very conservative in the sense of having type I error rates which were considerably less than the nominal value, $\alpha = 0.05$. An examination of the estimated power values of the normal theory F-test indicated that power is more sensitive to non-normal random errors than to non-normal treatment random effects. Cauchy random errors resulted in the lowest power values compared to the other distributions of random errors for each case where the treatment random effect is normal, logistic, double exponential or Cauchy. Except for Cauchy random errors, in all other cases, power approached one as the among variance component increases for the larger number of treatments used in this study.

The coverage rate of the normal theory confidence interval for the intra-class correlation coefficient depends on the distribution of the random effects and also on the treatment size. The distributions with heavier tails for the random effects resulted in low coverage rates. When the number of treatments is large, these coverage rates get even lower, sometimes approaching zero. Smaller widths for the confidence interval for the intra-class correlation coefficient were obtained when the assumption of normality was violated in one or both effects, especially if the random effects have heavy tailed distributions. Also, the coverage rates are low in these cases. Since a smaller width is desirable only if the actual coverage rate is close to its nominal value, non-normal random effects may mislead researchers into incorrectly believing that their results have led to a precise estimate of the intra-class correlation coefficient.

In conclusion, researchers should be cautious about the results of the F-test for the presence of the among treatment variance component in completely randomized one-way balanced random effects models when random errors are non-normal and have heavy tails since the power of the test can then be very low, even for a very large number of treatments. Also, the researcher should not rely on the nominal, normal theory coverage rate or the observed width of the confidence interval for the intra-class correlation coefficient when random effects are heavy tailed because narrow widths can occur with coverage rates much lower than nominal.

References

Jung, Inkyung and Kumar Sen, Pranab (2008). Robust testing for random effects in unbalanced heteroscdastic one-way models. Journal of Nonparametric Statistics, 20, 305-317.

Kuehl Robert O., (2000), Design of Experiments: Statistical Principles of Research Design and Analysis.

Searle, Shayle R. and Casella, George, and McCulloch, Charles E. (1992). Variance Components. New York: Wiley.

Searle, Shayle. R., (1970). Linear Models. New York: Wiley.

Milliken, George A. and Johnson, Dallas E., (2009). Analysis of Messy Data: Designed Experiments (Volume I).

Donner, Allan (1986). A review of inference procedures for the intra-class correlation coefficient in the one-way random effects models. International Statistics Review, 54, 67-82.

Bradley, R.A. and Schumann, D. E. W., (1957). The comparison of sensitivities of similar experiments: Applications. Biometrics, 13, 496-510.

Appendix A - Additional Plots and Tables

Figure A.1 Power Curves for the F-test when Treatment Random Effect is Normal and Random Error is Logistic

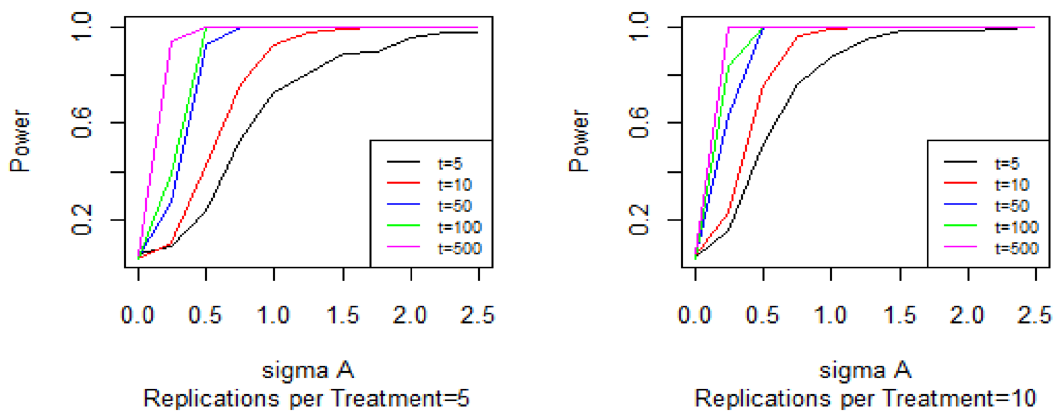


Figure A.2 Power Curves for the F-test when Treatment Random Effect is Normal and Random Error is Double Exponential

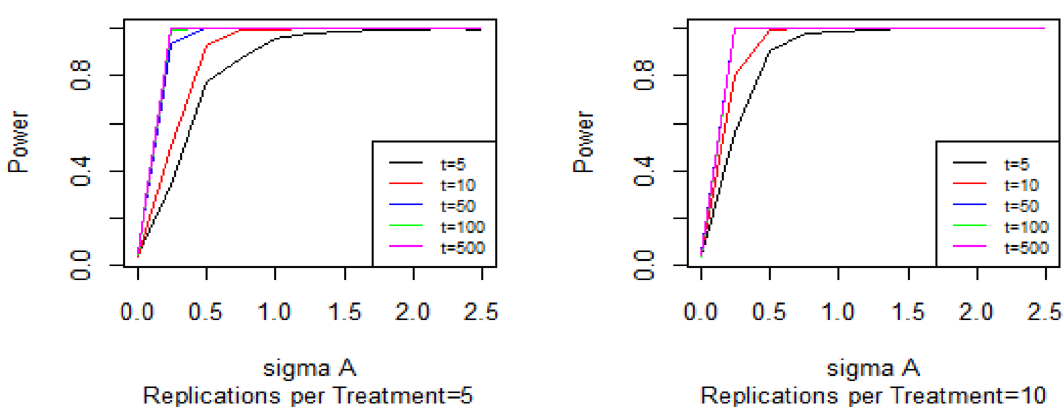


Figure A.3 Power Curves for the F-test when Treatment Random Effect is Logistic and Random Error is Normal

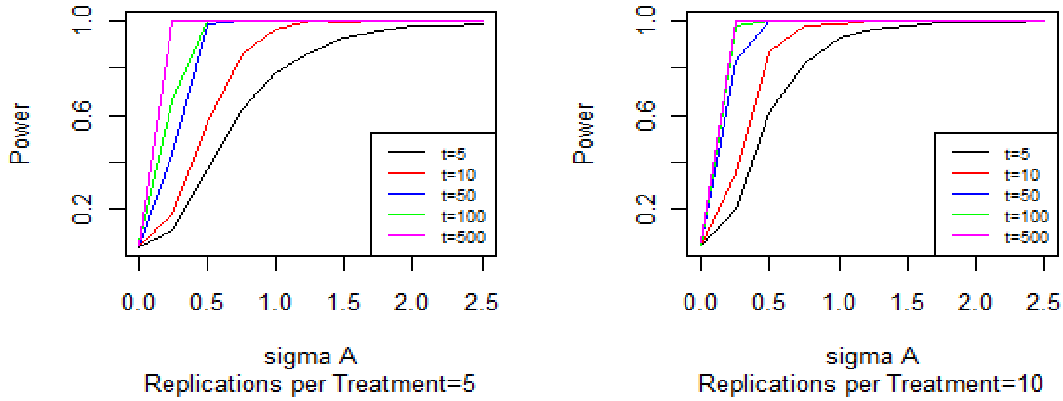


Figure A.4 Power Curves for the F-test when Treatment Random Effect is Logistic and Random Error is Logistic

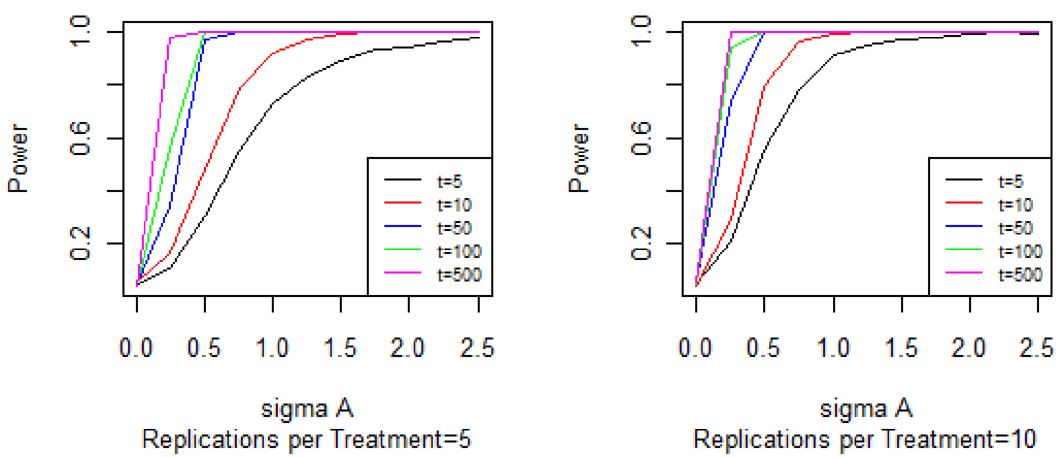


Figure A.5 Power Curves for the F-test when Treatment Random Effect is Logistic and Random Error is Double Exponential

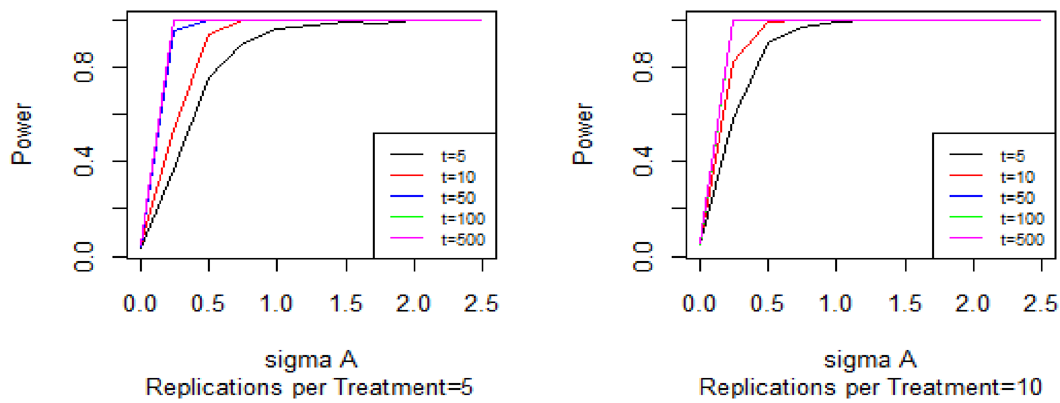


Figure A.6 Power Curves for the F-test when Treatment Random Effect is Double Exponential and Random Error is Normal

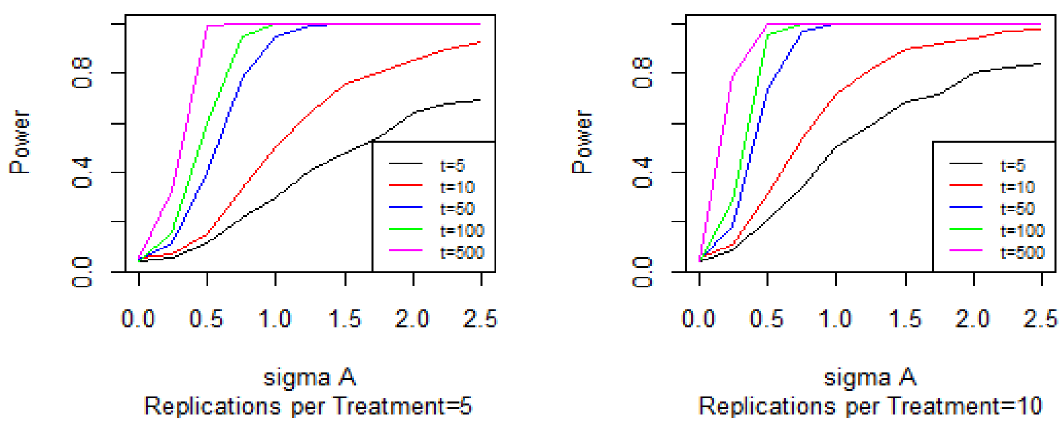


Figure A.7 Power Curves for the F-test when Treatment Random Effect is Double Exponential and Random Error is Logistic

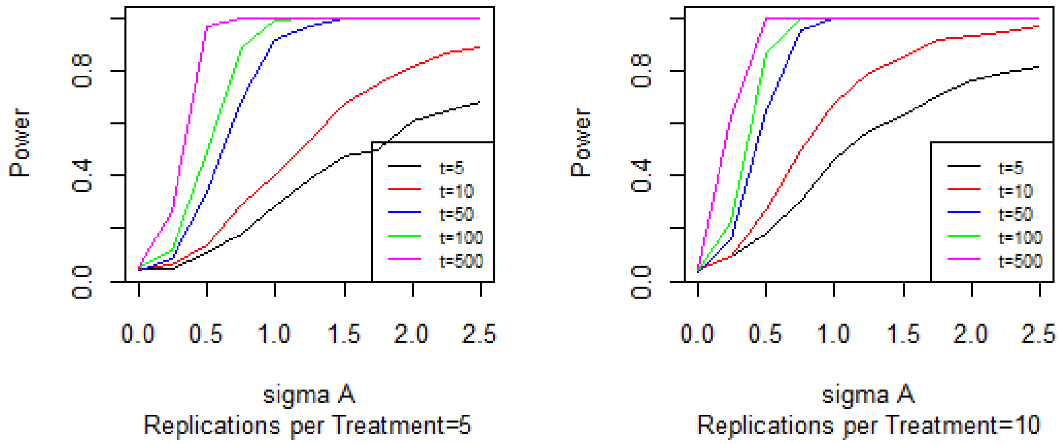


Figure A.8 Power Curves for the F-test when Treatment Random Effect is Double Exponential and Random Error is Double Exponential

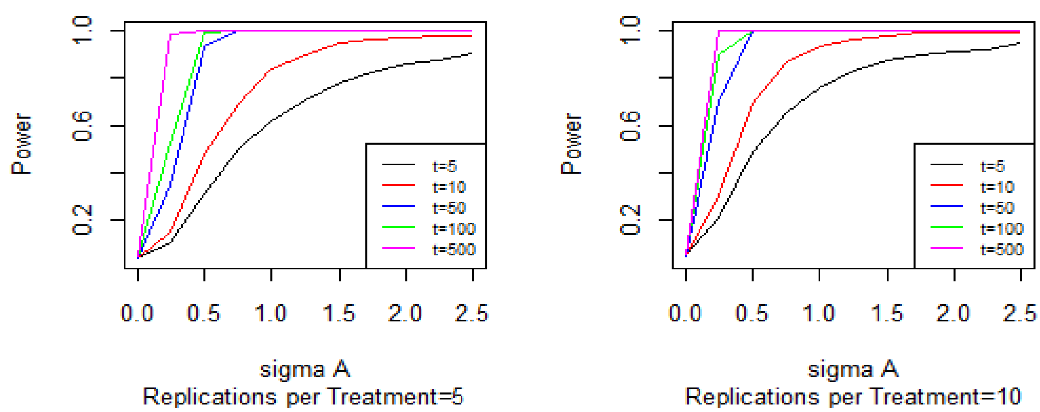


Figure A.9 Power Curves for the F-test when Treatment Random Effect is Cauchy and Random Error is Normal

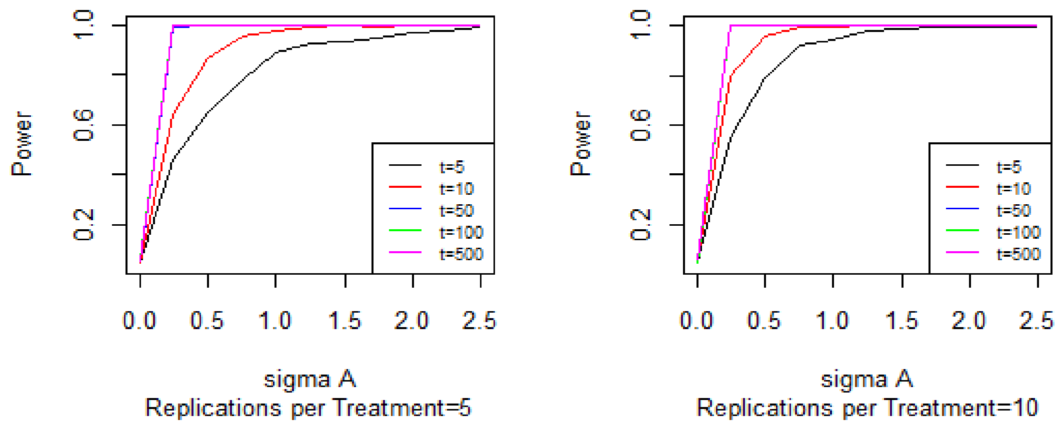


Figure A.10 Power Curves for the F-test when Treatment Random Effect is Cauchy and Random Error is Logistic

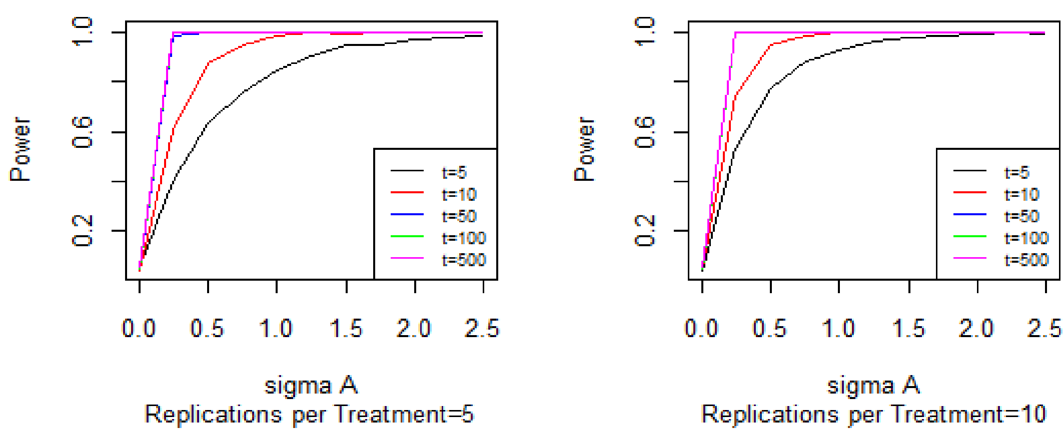


Figure A.11 Power Curves for the F-test when Treatment Random Effect is Cauchy and Random Error is Double Exponential

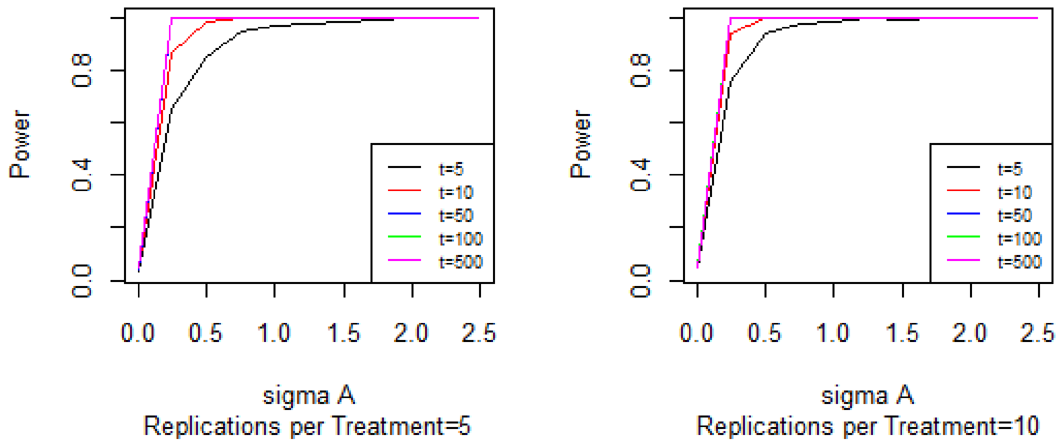


Figure A.12 Power Curves for the F-test when Treatment Random Effect is Cauchy and Random Error is Cauchy

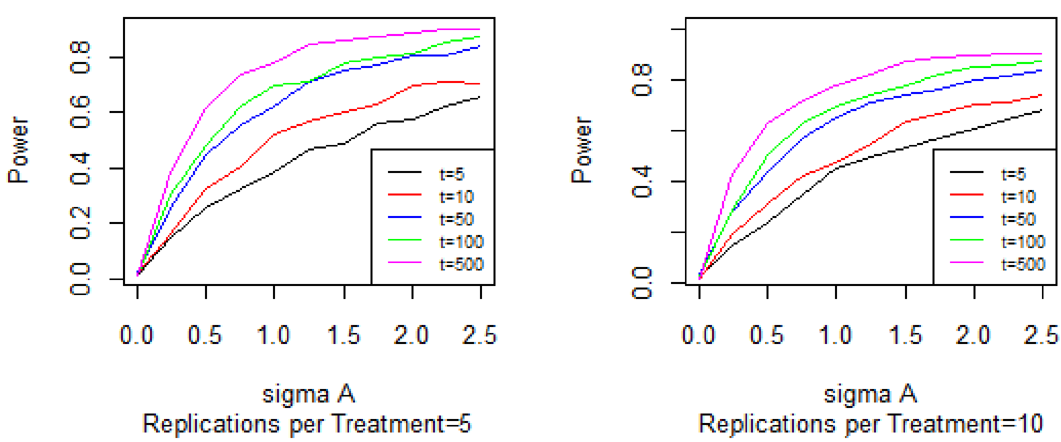


Table A.1 Estimated Coverage Rates for the 95% Confidence Interval for Intra-class Correlation Coefficient when Both Random Effects are Normal for Different Treatment Sizes

σ_A		ai~N eij~N								
	t=5, r=5	t=5, r=10	t=10, r=5	t=10, r=10	t=50,r=5	t=50,r=10	t=100, r=5	t=100,r=10	t=500,r=5	t=500,r=10
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.965	0.940	0.946	0.957	0.947	0.950	0.956	0.949	0.952	0.956
1	0.944	0.950	0.951	0.939	0.944	0.948	0.938	0.940	0.940	0.947
1.5	0.946	0.943	0.942	0.949	0.938	0.955	0.948	0.971	0.947	0.950
2	0.961	0.943	0.954	0.938	0.945	0.945	0.945	0.943	0.951	0.942
2.5	0.953	0.963	0.958	0.937	0.947	0.952	0.927	0.942	0.932	0.944

t= Number of treatments, r= Number of replicates per treatment

Table A.2 Estimated Coverage Rates for the 95% Confidence Interval for the Intra-class Correlation Coefficient for Treatments= (5, 100) and Replications = (5, 10)

Number of Treatments	Replicates per Treatment	σ_A	Coverage Rate															
			ai ~ Normal				ai ~ Logistic				ai ~ Double Exponential				ai ~ Cauchy			
			eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy
5	5	0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	5	0.5	**0.9648**	**0.9546**	0.6697	0.9677	**0.9194**	**0.9377**	0.61022	0.978788	0.9383	0.946083	0.9161	0.960406	0.5783	0.580321	0.329329	0.8530
5	5	1	**0.9440**	**0.9489**	0.5370	0.8508	**0.9260**	**0.9228**	0.479	0.827968	0.8239	0.828629	0.8317	0.807887	0.5000	0.576577	0.265	0.7771
5	5	1.5	**0.9460**	**0.9459**	0.5400	0.5343	**0.9069**	**0.9200**	0.498	0.529648	0.7427	0.718466	0.8056	0.326263	0.4975	0.51	0.259	0.5266
5	5	2	**0.9610**	**0.9590**	0.5040	0.4349	**0.9250**	**0.9040**	0.444	0.443662	0.6910	0.664659	0.7878	0.200605	0.4930	0.523524	0.249	0.4683
5	5	2.5	**0.9530**	**0.9490**	0.4910	0.3842	**0.9130**	**0.9170**	0.447	0.386774	0.6116	0.582915	0.7580	0.137652	0.5220	0.532	0.247	0.4493
5	10	0	**1.0000**	**1.0000**	1.0000	1.0000	**1.0000**	**1.0000**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	10	0.5	**0.9400**	**0.9468**	0.5910	0.9224	**0.9128**	**0.9319**	0.52	0.914141	0.8982	0.91001	0.8678	0.918367	0.5220	0.578736	0.255	0.8622
5	10	1	**0.9500**	**0.9430**	0.4950	0.4460	**0.9040**	**0.9390**	0.469	0.463243	0.7603	0.709127	0.8148	0.309476	0.4920	0.539	0.23	0.5792
5	10	1.5	**0.9430**	**0.9640**	0.4790	0.2679	**0.8870**	**0.9290**	0.418	0.291165	0.6580	0.6002	0.7698	0.084093	0.4760	0.513	0.22	0.3799
5	10	2	**0.9430**	**0.9340**	0.4990	0.2437	**0.8980**	**0.9320**	0.433	0.238716	0.6383	0.569709	0.7655	0.045226	0.4940	0.506	0.23	0.3825
5	10	2.5	**0.9630**	**0.9420**	0.4820	0.1996	**0.9020**	**0.9230**	0.411	0.219659	0.5926	0.553213	0.7630	0.046324	0.4910	0.488	0.227	0.3732
100	5	0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
100	5	0.5	**0.9560**	0.8890	0.0010	0.0030	0.8530	**0.9190**	0	0.004008	0.1360	0.065	0.8380	0.005107	0.0000	0.001	0	0.1567
100	5	1	**0.9380**	0.8160	0.0000	0.0000	0.7650	**0.9020**	0	0	0.0030	0.001	0.7350	0	0.0000	0	0	0.0874
100	5	1.5	**0.9480**	0.7520	0.0000	0.0000	0.7010	**0.8940**	0	0	0.0000	0	0.7280	0	0.0000	0	0	0.0834
100	5	2	**0.9450**	0.7650	0.0000	0.0000	0.6890	**0.8890**	0	0	0.0000	0	0.7030	0	0.0000	0	0	0.0873
100	5	2.5	**0.9270**	0.7330	0.0000	0.0000	0.6870	**0.8760**	0	0	0.0000	0	0.6970	0	0.0000	0	0	0.0731
100	10	0	**1.0000**	1.0000	1.0000	1.0000	1.0000	**1.0000**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
100	10	0.5	**0.9490**	0.8380	0.0000	0.0000	0.8010	**0.9240**	0	0	0.0080	0.002	0.7630	0	0.0000	0	0	0.0804
100	10	1	**0.9400**	0.7690	0.0000	0.0000	0.6850	**0.9100**	0	0	0.0000	0	0.7130	0	0.0000	0	0	0.0582
100	10	1.5	**0.9710**	0.7460	0.0000	0.0000	0.6570	**0.8850**	0	0	0.0010	0	0.6890	0	0.0000	0	0	0.0472
100	10	2	**0.9430**	0.7320	0.0000	0.0000	0.6720	**0.8700**	0	0	0.0010	0	0.7030	0	0.0000	0	0	0.0490
100	10	2.5	**0.9420**	0.7100	0.0000	0.0000	0.6660	**0.9020**	0	0	0.0000	0	0.6620	0	0.0000	0	0	0.0560

[2]

[2] Boldface numbers indicate estimated coverage rates close to nominal 0.95 coverage rate.

Table A.3 Relative Median Widths for the 95% Confidence Interval for the Intra-class Correlation Coefficient Number of treatments = (5,100) and Number of replications = (5 &10)

Number of Treatments	Replicates per Treatment	σ_A	Median Width	Relative Median Width														
				ai ~ Normal			ai ~ Logistic				ai ~ Double Exponential				ai ~ Cauchy			
			ai ~Normal eij~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy	eij ~Normal	eij ~Logistic	eij ~Double Exponential	eij ~Cauchy
5	5	0	**0.566**	1.009	1.010	1.056	1.012	1.006	1.019	1.053	0.996	1.016	1.015	1.053	1.012	1.034	1.025	1.060
5	5	0.5	**0.683**	1.048	1.089	0.899	1.091	1.072	1.062	0.898	0.919	0.915	1.055	0.885	0.970	0.942	0.709	0.932
5	5	1	**0.710**	1.067	0.705	0.904	1.033	1.053	0.647	0.903	1.004	1.010	1.000	0.867	0.671	0.780	0.288	0.908
5	5	1.5	**0.616**	1.097	0.518	1.113	0.997	1.051	0.476	1.091	1.192	1.202	1.014	1.002	0.514	0.531	0.188	1.003
5	5	2	**0.502**	1.125	0.386	1.363	0.910	1.017	0.342	1.362	1.434	1.476	1.155	1.293	0.395	0.481	0.116	1.232
5	5	2.5	**0.418**	1.059	0.312	1.631	0.841	1.001	0.281	1.611	1.642	1.684	1.032	1.539	0.368	0.413	0.094	1.398
5	10	0	**0.378**	1.072	1.057	1.116	0.988	1.011	1.041	1.102	0.994	1.002	1.030	1.103	0.988	1.049	1.074	1.117
5	10	0.5	**0.627**	1.079	1.098	0.688	1.111	1.089	1.066	0.697	0.817	0.802	1.022	0.677	0.950	0.982	0.733	0.771
5	10	1	**0.661**	1.064	0.700	0.718	1.030	1.042	0.685	0.731	0.991	0.953	0.997	0.661	0.688	0.769	0.291	0.824
5	10	1.5	**0.563**	1.092	0.500	0.980	0.986	1.065	0.438	0.980	1.209	1.194	1.041	0.791	0.506	0.590	0.162	0.937
5	10	2	**0.458**	1.116	0.419	1.272	0.920	1.029	0.358	1.305	1.473	1.484	1.116	1.006	0.438	0.480	0.125	1.151
5	10	2.5	**0.380**	1.058	0.335	1.603	0.827	0.977	0.292	1.621	1.657	1.721	1.054	1.287	0.381	0.382	0.097	1.393
100	5	0	**0.072**	0.961	0.998	0.978	0.972	0.987	0.965	0.977	0.980	0.995	1.004	0.973	0.972	0.980	0.926	0.978
100	5	0.5	**0.175**	0.968	1.055	0.408	1.035	1.004	1.021	0.408	0.827	0.758	1.005	0.404	0.193	0.254	0.060	0.620
100	5	1	**0.181**	1.033	0.603	0.409	0.968	1.005	0.530	0.411	0.987	0.952	1.007	0.395	0.049	0.067	0.014	0.742
100	5	1.5	**0.141**	1.102	0.436	0.548	0.893	1.005	0.364	0.564	1.336	1.332	1.006	0.509	0.029	0.029	0.008	1.027
100	5	2	**0.103**	1.127	0.370	0.825	0.869	0.997	0.305	0.868	1.742	1.802	1.011	0.717	0.024	0.027	0.007	1.308
100	5	2.5	**0.075**	1.161	0.334	1.244	0.845	1.010	0.273	1.291	2.113	2.257	1.003	1.007	0.022	0.025	0.005	1.749
100	10	0	**0.036**	0.954	0.970	0.967	0.931	0.966	0.940	0.965	0.947	0.948	0.967	0.968	0.931	0.938	0.931	0.970
100	10	0.5	**0.131**	0.936	1.225	0.271	1.061	0.998	1.206	0.271	0.673	0.628	0.990	0.269	0.276	0.296	0.090	0.468
100	10	1	**0.159**	1.015	0.640	0.232	0.978	1.006	0.567	0.233	0.847	0.800	1.002	0.224	0.060	0.069	0.017	0.518
100	10	1.5	**0.129**	1.087	0.450	0.313	0.903	1.009	0.382	0.313	1.234	1.203	1.015	0.278	0.028	0.040	0.008	0.713
100	10	2	**0.095**	1.141	0.380	0.454	0.883	1.009	0.315	0.461	1.656	1.682	1.022	0.388	0.023	0.032	0.007	1.025
100	10	2.5	**0.071**	1.157	0.343	0.661	0.852	1.007	0.280	0.724	2.037	2.150	1.024	0.539	0.018	0.025	0.005	1.436

[3]

[3] Boldface numbers indicate estimated median widths when the both random effects are normally distributed.

Table A.4 Number of data sets where upper bound of the 95% confidence interval is non-negative for all the combinations of parameters & distributions when treatments=5, replications per treatment = (5,10)

Number of Treatments	Replications per Treatment	σ_A	Number of data sets															
			ai~N eij~N	ai~N eij~L	ai~N eij~DE	ai~N eij~C	ai~L eij~N	ai~L eij~L	ai~L eij~DE	ai~L eij~C	ai~DE eij~N	ai~DE eij~L	ai~DE eij~DE	ai~DE eij~C	ai~C eij~N	ai~C eij~L	ai~C eij~DE	ai~C eij~C
5	5	0	973	979	984	984	975	974	974	988	975	967	990	984	975	979	982	988
5	5	0.5	994	991	998	990	993	995	998	990	988	983	989	985	996	996	999	993
5	5	1	1000	998	1000	992	1000	998	1000	994	994	992	998	989	1000	999	1000	996
5	5	1.5	1000	999	1000	992	999	1000	1000	995	995	991	998	990	999	1000	1000	995
5	5	2	1000	1000	1000	998	1000	1000	1000	994	1000	996	999	992	1000	999	1000	995
5	5	2.5	1000	1000	1000	997	1000	1000	1000	998	999	995	1000	988	1000	1000	1000	997
5	10	0	970	970	983	990	971	983	982	986	969	981	980	985	971	977	979	981
5	10	0.5	1000	996	1000	992	998	999	1000	990	982	989	991	980	998	997	1000	994
5	10	1	1000	1000	1000	991	1000	1000	1000	993	993	997	999	992	1000	1000	1000	998
5	10	1.5	1000	1000	1000	993	1000	1000	1000	996	997	998	999	987	1000	1000	1000	995
5	10	2	1000	1000	1000	997	1000	1000	1000	997	998	997	998	995	1000	1000	1000	996
5	10	2.5	1000	1000	1000	997	1000	1000	1000	997	999	996	1000	993	1000	1000	1000	994

N=normal, L=logistic, DE=double exponential, C= Cauchy

Table A.5 Number of data sets where upper bound of the 95% confidence interval is non-negative for all the combinations of parameters & distributions when treatments=10, replications per treatment = (5,10)

Number of Treatments	Replications per Treatment	σ_A	Number of data sets															
			ai~N eij~N	ai~N eij~L	ai~N eij~DE	ai~N eij~C	ai~L eij~N	ai~L eij~L	ai~L eij~DE	ai~L eij~C	ai~DE eij~N	ai~DE eij~L	ai~DE eij~DE	ai~DE eij~C	ai~C eij~N	ai~C eij~L	ai~C eij~DE	ai~C eij~C
10	5	0	975	975	963	987	973	984	982	983	973	974	974	985	973	973	978	984
10	5	0.5	998	998	1000	982	998	999	1000	982	994	990	999	985	1000	1000	1000	990
10	5	1	1000	1000	1000	989	1000	999	1000	994	995	997	1000	988	1000	1000	1000	992
10	5	1.5	1000	1000	1000	997	1000	1000	1000	993	997	997	1000	990	1000	1000	1000	994
10	5	2	1000	1000	1000	997	1000	1000	1000	999	999	999	1000	990	1000	1000	1000	996
10	5	2.5	1000	1000	1000	998	1000	1000	1000	997	999	1000	1000	992	1000	1000	1000	997
10	10	0	977	969	975	980	972	975	974	982	978	981	983	978	972	971	983	992
10	10	0.5	999	999	1000	984	1000	999	1000	985	992	995	998	987	1000	1000	1000	992
10	10	1	1000	1000	1000	995	1000	1000	1000	989	999	997	1000	989	1000	1000	1000	993
10	10	1.5	1000	1000	1000	993	1000	1000	1000	992	1000	999	1000	993	1000	1000	1000	998
10	10	2	1000	1000	1000	997	1000	1000	1000	993	1000	1000	1000	993	1000	1000	1000	996
10	10	2.5	1000	1000	1000	992	1000	1000	1000	998	1000	1000	1000	989	1000	1000	1000	999

N=normal, L=logistic, DE=double exponential, C= Cauchy

Table A.6 Number of data sets where upper bound of the 95% confidence Interval is non-negative for all the combinations of parameters & distributions when treatments=50, replications per treatment = (5,10)

Number of Treatments	Replications per Treatment	σ_A	Number of data sets															
			ai~N eij~N	ai~N eij~L	ai~N eij~DE	ai~N eij~C	ai~L eij~N	ai~L eij~L	ai~L eij~DE	ai~L eij~C	ai~DE eij~N	ai~DE eij~L	ai~DE eij~DE	ai~DE eij~C	ai~C eij~N	ai~C eij~L	ai~C eij~DE	ai~C eij~C
50	5	0	962	970	978	980	971	979	974	983	982	977	973	985	971	959	958	989
50	5	0.5	1000	1000	1000	983	1000	1000	1000	984	998	1000	1000	984	1000	1000	1000	993
50	5	1	1000	1000	1000	989	1000	1000	1000	995	999	1000	1000	985	1000	1000	1000	998
50	5	1.5	1000	1000	1000	990	1000	1000	1000	996	1000	1000	1000	990	1000	1000	1000	999
50	5	2	1000	1000	1000	993	1000	1000	1000	997	1000	1000	1000	988	1000	1000	1000	998
50	5	2.5	1000	1000	1000	997	1000	1000	1000	997	1000	1000	1000	992	1000	1000	1000	996
50	10	0	985	980	968	979	971	970	979	990	976	976	979	991	971	982	980	987
50	10	0.5	1000	1000	1000	985	1000	1000	1000	983	1000	1000	1000	982	1000	1000	1000	993
50	10	1	1000	1000	1000	993	1000	1000	1000	992	1000	1000	1000	990	1000	1000	1000	996
50	10	1.5	1000	1000	1000	992	1000	1000	1000	996	1000	1000	1000	988	1000	1000	1000	1000
50	10	2	1000	1000	1000	991	1000	1000	1000	994	1000	1000	1000	987	1000	1000	1000	996
50	10	2.5	1000	1000	1000	998	1000	1000	1000	996	1000	1000	1000	993	1000	1000	1000	995

N=normal, L=logistic, DE=double exponential, C= Cauchy

Table A.7 Number of data sets where upper bound of the 95% confidence interval is non-negative for all the combinations of parameters & distributions when treatments=100, replications per treatment = (5,10)

Number of Treatments	Replications per Treatment	σ_A	Number of data sets															
			ai~N eij~N	ai~N eij~L	ai~N eij~DE	ai~N eij~C	ai~L eij~N	ai~L eij~L	ai~L eij~DE	ai~L eij~C	ai~DE eij~N	ai~DE eij~L	ai~DE eij~DE	ai~DE eij~C	ai~C eij~N	ai~C eij~L	ai~C eij~DE	ai~C eij~C
100	5	0	982	976	971	987	976	960	972	984	975	975	962	978	976	971	975	983
100	5	0.5	1000	1000	1000	987	1000	1000	1000	998	1000	1000	1000	979	1000	1000	1000	989
100	5	1	1000	1000	1000	987	1000	1000	1000	986	1000	1000	1000	989	1000	1000	1000	995
100	5	1.5	1000	1000	1000	990	1000	1000	1000	993	1000	1000	1000	987	1000	1000	1000	995
100	5	2	1000	1000	1000	994	1000	1000	1000	994	1000	1000	1000	989	1000	1000	1000	997
100	5	2.5	1000	1000	1000	992	1000	1000	1000	993	1000	1000	1000	995	1000	1000	1000	999
100	10	0	975	973	965	981	977	981	971	985	981	977	975	982	977	981	972	990
100	10	0.5	1000	1000	1000	984	1000	1000	1000	987	1000	1000	1000	986	1000	1000	1000	995
100	10	1	1000	1000	1000	986	1000	1000	1000	988	1000	1000	1000	987	1000	1000	1000	997
100	10	1.5	1000	1000	1000	994	1000	1000	1000	989	1000	1000	1000	982	1000	1000	1000	996
100	10	2	1000	1000	1000	998	1000	1000	1000	994	1000	1000	1000	991	1000	1000	1000	999
100	10	2.5	1000	1000	1000	995	1000	1000	1000	993	1000	1000	1000	991	1000	1000	1000	1000

N=normal, L=logistic, DE=double exponential, C= Cauchy

Table A.8 Number of data sets where upper bound of the 95% confidence interval is non-negative for all the combinations of parameters & distributions when treatments=500, replications per treatment = (5,10)

Number of Treatments	Replications per Treatment	σ_A	Number of data sets															
			ai~N eij~N	ai~N eij~L	ai~N eij~DE	ai~N eij~C	ai~L eij~N	ai~L eij~L	ai~L eij~DE	ai~L eij~C	ai~DE eij~N	ai~DE eij~L	ai~DE eij~DE	ai~DE eij~C	ai~C eij~N	ai~C eij~L	ai~C eij~DE	ai~C eij~C
500	5	0	970	966	970	989	963	972	973	993	975	974	970	995	963	971	968	986
500	5	0.5	1000	1000	1000	993	1000	1000	1000	993	1000	1000	1000	986	1000	1000	1000	998
500	5	1	1000	1000	1000	992	1000	1000	1000	991	1000	1000	1000	991	1000	1000	1000	1000
500	5	1.5	1000	1000	1000	991	1000	1000	1000	992	1000	1000	1000	987	1000	1000	1000	998
500	5	2	1000	1000	1000	991	1000	1000	1000	995	1000	1000	1000	992	1000	1000	1000	1000
500	5	2.5	1000	1000	1000	993	1000	1000	1000	996	1000	1000	1000	991	1000	1000	1000	999
500	10	0	969	973	972	989	975	964	975	991	965	982	969	989	975	975	977	990
500	10	0.5	1000	1000	1000	994	1000	1000	1000	993	1000	1000	1000	992	1000	1000	1000	996
500	10	1	1000	1000	1000	989	1000	1000	1000	990	1000	1000	1000	995	1000	1000	1000	997
500	10	1.5	1000	1000	1000	988	1000	1000	1000	991	1000	1000	1000	991	1000	1000	1000	1000
500	10	2	1000	1000	1000	993	1000	1000	1000	990	1000	1000	1000	991	1000	1000	1000	999
500	10	2.5	1000	1000	1000	994	1000	1000	1000	995	1000	1000	1000	996	1000	1000	1000	998

N=normal, L=logistic, DE=double exponential, C= Cauchy

Appendix B - R Code

##ai~Normal, eij~Normal##

```
set.seed(32111) # nn
T <- c(5,10,50,100,500) # number of treatment groups
R <- c(5,10) # number of replications per trt group
sd.Err <- 1 # the error stand deviation
SD.TRT <- c(0,0.25,0.50,0.75,1.00,1.25,1.50,1.75,2.00,2.25,2.5)
out1=NULL
out11=NULL
out2=NULL
out3=NULL
out4=NULL
out6=NULL
out7=NULL
out8=NULL
outd=NULL

for (t in T){
a <- numeric(t)
for (r in R){
 Y <- matrix(0, nrow=r, ncol=t)
for (sd.trt in SD.TRT) {
for (N in 1:1000) {
   for (i in 1:t) {
    ai] <- rnorm(1,0, sd.trt)
     for (j in 1:r) {
```

```
      Y[j,i] <- a[i]+rnorm(1,0,sd.Err)
     }
    }
x <- as.vector(Y)
trt <- as.factor(rep(1:t, each = r))

X <- t(Y)
X.bar <- apply(X, 1, mean) #trt sample means
A <- sum(X^2)
B <- sum(apply(X, 1, sum)^2)/r
C <- (sum(X))^2/(r * t)
dfT <- t-1
dfE <- t*(r-1)
crit <- 0.05 # critical value
MS.trt <- (B - C)/(t - 1)
MS.Error <- (A - B)/(t * (r - 1))

est.sigmaAsq=(MS.trt-MS.Error)/r
if (est.sigmaAsq<0) neg=1 else neg=0   #negative sigmaA-squared
estimates

#----------power-------------------------------
F.ratio <- MS.trt/ MS.Error
Falpha=qf(.95,dfT,dfE)
if (F.ratio>=Falpha) rej=1 else rej=0    #number of rejections

#---------Intra-class correlation--------------
ICC=(sd.trt)^2/((sd.trt)^2 + (sd.Err)^2)
```

```
est.ICC=(MS.trt-MS.Error)/(MS.trt+(r-1)*MS.Error)
FL=qf(0.05/2,t-1,t*r-1)
FU=qf(1-0.05/2,t-1,t*r-1)
if (F.ratio<FL)delete=1 else delete=0

if (delete==0){
UL=(F.ratio-FL)/(F.ratio+(r-1)*FL)
if (F.ratio<FU) LL=0 else LL=(F.ratio-FU)/(F.ratio+(r-1)*FU)
if (F.ratio<FU) negLL=1 else negLL=0
if (LL<ICC & UL>ICC) cover=1 else cover=0  #coverage
width=UL-LL
mid.CI=(LL+UL)/2
chk=abs(ICC-mid.CI)/(UL-LL)
if (mid.CI>=ICC) above=1 else above=0
} else{
UL=NA
LL=NA
negLL=NA
cover=NA
width=NA
mid.CI=NA
chk=NA
above=NA
}

out1=rbind(out1,c(t,r,sigmaA=sd.trt,pop.ICC=ICC,LL=LL,UL=UL
,width=width))
```

```
out2=rbind(out2,c(trt=t,rep=r,sigmaA=sd.trt,power=rej))
out3=rbind(out3,c(trt=t,rep=r,sigmaA=sd.trt,delete=delete
,ICC=ICC,coverage=cover,above.ICC=above,width=width))
}}}}

#------counting deleting datasets----------
outd=rep(0,1)
l=1:1000  # N=1000
for(m in 1:110) {         #t*r*sigma=110
#test=out3[l+(m-1)*1000,4] #N=1000
outd=rbind(outd,1000-sum(out3[l+(m-1)*1000,4]))
}
outd=outd[-1,]
out33=na.omit(out3)   #omit NA's

#------getting coverage rate, avg.width of the test------
out7=rep(0,10)
for(k in 1:110) {         #t*r*sigma=110
if (k==1){
test1=out33[1:outd[1],]
} else {
test1=out33[(sum(outd[1:k-1])+1):(sum(outd[1:k])),]
}
z1=apply(test1,2,mean)
out7=rbind(out7,z1)
}
out7=out7[-1,]
```

```
out7
#-------------------medain-------------------
out8=rep(0,4)
for(k in 1:110) {         #t*r*sigma=110
if (k==1){
test2=out33[1:outd[1],7:10]
} else {
test2=out33[(sum(outd[1:k-1])+1):(sum(outd[1:k])),7:10]
}
z2=apply(test2,2,median)
out8=rbind(out8,z2)
}
out8=out8[-1,]
out8

#----------------power-------------------
out6=rep(0,4)
l=1:1000 # N=1000
for(m in 1:110) {         #t*r*sigma=110
test=out2[l+(m-1)*1000,]  #N=1000
z=apply(test,2,mean)
out6=rbind(out6,z)
}
out6=out6[-1,]
out6

#-----------power curves------------------
```

```
par(mfrow=c(2,2))

nn=read.csv("F:\\MS\\report\\simulation\\4_4\\new\\data2nn.csv",
header=TRUE)

r=5

plot(nn$sigmaA,nn$power,type="n",xlab="sigma
A",ylab="Power",sub="Replications per Treatment=5")

lines(nn$sigmaA[1:11],nn$power[1:11],col="black") #t=5

lines(nn$sigmaA[23:33],nn$power[23:33],col="red")#t=10

lines(nn$sigmaA[45:55],nn$power[45:55],col="blue")#t=50

lines(nn$sigmaA[67:77],nn$power[67:77],col="green")#t=100

lines(nn$sigmaA[89:99],nn$power[89:99],col="magenta")#t=500

legend("bottomright",c("t=5","t=10","t=50","t=100","t=500"),lty=
c(1,1,1,1,1),col=c("black","red","blue","green","magenta"),cex=.
75)


#r=10

plot(nn$sigmaA,nn$power,type="n",xlab="sigma
A",ylab="Power",sub="Replications per Treatment=10")

lines(nn$sigmaA[22:22],nn$power[22:22],col="black") #t=5,r=5

lines(nn$sigmaA[34:44],nn$power[34:44],col="red") #t=10

lines(nn$sigmaA[56:66],nn$power[56:66],col="blue")#t=50

lines(nn$sigmaA[78:88],nn$power[78:88],col="green")#t=100

lines(nn$sigmaA[100:110],nn$power[100:110],col="magenta")#t=500

legend("bottomright",c("t=5","t=10","t=50","t=100","t=500")

,lty=c(1,1,1,1,1),col=c("black","red","blue","green","magenta"),
cex=.75)


#-----------------------------------------------------------

write.csv(out1,file="F:\MS\report\simulation\\data1nn.csv")
```

```
write.csv(out6,file="F:\MS\report\simulation\\data2nn.csv")
write.csv(out7,file="F:\MS\report\simulation\\data3nn.csv")
write.csv(out8,file="F:\MS\report\simulation\\data4nn.csv")
write.csv(outd,file="F:\MS\report\simulation\\datadnn.csv")
```