韓國 軍事運營分析 學會誌
第13券, 第1號, 1987, 6.

Power Algorithms for Analysis of Variance Tests

Hur, Seong Pil

Abstract

Power algorithms for analysis of variance tests are presented. In experimental design of operational tests and evaluations the selection of design parameters so as to attain an experiment with desired power is a difficult and important problem.

An interactive computer program is presented which uses the power algorithms for ANOVA tests and creates graphical presentations which can be used to assist decision makers in statistical design. ANOVA tests and associated parameters (such as sample size, types and levels of treatments, and alpha-level) are examined.

1. Introduction

1-1 Description of the Problem

A statistical design is a plan according to which an experiment is patterned. It provides the basis upon which appropriate statistical tests and inferences can be made after the experiment has been performed. The selection of the experimental design to be used in a given situation is extremely important because it plays a predominant role in the effieiency of the experiment, the precision with which the objectives are met, and the total effort (and cost) expended upon the experiment.

The concern of this thesis is the development of power algorithms for analysis of variance tests. The experimenter wants a test to achieve the correct decision with as high probability as possible in practical operational tests and evaluations given certain conditions hold.

1-2 Scope of the Thesis

An interactive computer program is developed which uses power algorithms for ANOVA tests. The program can be used by decision makers who may not have deep knowledge of statistics or may

*) Korea Naval Academy

be unfamiliar with the details of the design of experiments. The output helps the decision maker design experiments appropriate for testing hypotheses with common statistical tests. The experiment designer can use the program to evaluate the effect upon power of varying underlying design parameters such as sample size.

Because the underlying statistical noncentral distributions used in computing power do not have tables, approximation methods are used for computing the noncentral CDFs. Approximation methods were developed which are efficient, requiring reduced computer CPU time.

1-3 Background

A hypothesis is a statement about the values of the parameters of a probability distribution. For example, suppose we think that the mean yield of a chemical process is more than 94.5 percent. Hypotheses to test this statement might be expressed formally as,

Ho: $\mu \leqslant 94.5$

Ha: $\mu > 94.5$

The statement Ho: $\mu \leqslant 94.5$ is called the null hypothesis, and Ha: $\mu > 94.5$ is called the alternative hypothesis. The value of the mean specified in the null hypothesis might be determined in one of two ways. It may result of some theory or model regarding the process under study, or it may be the result of contractual specifications.

To test a hypothesis we usually devise a procedure for taking a random sample, computing an appropriate test statistic, and then rejecting of failing to reject the null hypothesis Ho, depending on the outcome on the test statistic. Part of this procedure is specifying the set of values for the test statistic which lead to rejection of Ho. This set of values is called the critical region or rejection region for the test.

Two kinds of errors may be committed when testing hypotheses. If the null hypothesis is rejected when it is true, then a type I error has occurred. If the null hypothesis is not rejected when it is false, then a type II error has been made. The probabilities of these two errors are given special symbols:

α = P(type I error) = P(reject Ho | Ho is true)

β = P(type II error) = P(fail to reject Ho | Ho is false)

Generally, β is calculated using a noncentral distribution depending on a noncentrality parameter. The noncentrality parameter is usually a measure of the distance (in some sense) between the values of the parameter under the null and alternate hypotheses. Thus, type II error rates depend upon parameters of the distribution of the test statistic under the alternative hypothesis. Power is the probability the test would reject Ho when Ho is false, or $1-\beta$.

In addition, whatever the test procedure is, rejecting Ho when Ho is not true is usually something we want a test to achieve with as high a probability as possible. Therefore we want the power of the test, for a given value of non-centrality value, to be high. Power algorithms are used to assist in the "best" selection of parameters for a statistical design. We use approximation methods to compute the CDF's of the noncentral F, noncentral χ^2, and noncentral t distributions.

2. Noncentral Chi-square, F Distributions

2-1 Noncentral Chi-square Distribution

If W1, w2, ..., Wn are independently distributed as $N(\mu_i, 1)$, i=1,2, ..., n, then ΣWi^2 has a distribution known as a noncentral $\chi^2(n, \lambda)$ where n is the number of degrees of freedom and $\lambda = \Sigma\mu_i^2$ is the noncentrality parameter. When $\mu_1 = \mu_2 = \ldots = \mu_n = 0$, then $\lambda = 0$, and the noncentral $\chi^2(n, 0)$ reduces to the usual central $X^2(n)$ with n degrees of freedom. The cumulative distribution function of $\chi^2(n, \lambda)$ is,

$$F(X : n, \lambda) = Pr[\chi^2(n, \lambda) \leq X]$$

$$= e^{-\frac{1}{2}\lambda} \sum_{j=0}^{\infty} \frac{(\frac{1}{2}\lambda)^j}{j!} \frac{1}{2^{\frac{n}{2}+j}\Gamma(\frac{n}{2}+j)} \int_0^x y^{\frac{n}{2}+j-1} e^{-\frac{y}{2}} dy \; ; X > 0 \tag{2.1}$$

while $F(x; n, \lambda) = 0$ for $x \leq 0$

It is possible to express $F(x;n,\lambda)$ for $x>0$, in an easily remembered form as a weighted sum of central χ^2 CDF's with weights equal to the probabilities of a Poisson distribution with expected value $\lambda/2$. That is,

$$F(X; n, \lambda) = \sum_{j=0}^{\infty} \left[\frac{(\frac{1}{2}\lambda)^j}{j!} e^{\frac{1}{2}\lambda}\right] \cdot Pr[\chi^2_{(n+2j)} \leq X] \tag{2.2}$$

Thus a $\chi^2(n, \lambda)$ variable can be regarded as a mixture of central χ^2 variables. This interpretation is often useful in deriving the distribution of functions of noncentral χ^2 random variables.

The probability density function can, similarly, be expressed as a mixture of central χ^2 probability density functions

$$f(x) = C \sum_{j=0}^{\infty} \frac{\lambda^j \chi^{j-1}}{\Gamma(j+1)\, 2^j\, \Gamma(\frac{n}{2}+j)}, \text{ where } C = \frac{e^{-\lambda-\frac{x}{2}} X^{\frac{n}{2}}}{2^{\frac{n}{2}}} \tag{2.3}$$

The mean and variance of the distribution are

$E(x) = n + 2\lambda$ and $Var(x) = 2n + 8\lambda$ [Ref. 1:p. 130].

2.2 Noncentral F Distribution

If Y_1 and Y_2 are independent and Y_1 is $\chi_1^2(n_1, \lambda)$ and Y_2 is $\chi^2(n_2)$ then

$$V = (Y_1/n_1)/(Y_2/n_2) \tag{2.4}$$

is distributed as $F'(n_1, n_2, \lambda)$, the noncentral F distribution with n_1 and n_2 degrees of freedom and noncentrality parameter λ. Its density function is given by

$$f(v) = C \sum_{k=0}^{\infty} \frac{\lambda^k n_1^k \Gamma(\frac{n_1}{2} + \frac{n_2}{2} + k)\, v^{k-1}}{\Gamma(\frac{n_1}{2} + k)(n_2 + n_1 v)^k} \quad ; \; v \geq 0 \tag{2.5}$$

where

$$C = \frac{e^{-\lambda} n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} v^{\frac{n_1}{2}}}{\Gamma(\frac{n_2}{2})(n_2 + n_1 v)^{\frac{n_1+n_2}{2}}}$$

when $\lambda = 0$ this reduces reduces to the density function of the central $F(n_1, n_2)$ distribution. The mean and variance of the distribution are

$$E(v) = \frac{n_2}{n_2 - 2}\left(1 + \frac{2\lambda}{n_1}\right) \qquad \text{for } n > 2$$

and

$$Var(v) = \frac{2 n_2^2}{n_1^2 (n_2 - 2)} \left[\frac{(n_1 + 2\lambda)^2}{(n_2 - 2)(n_2 - 4)} + \frac{n_1 + 4\lambda}{n_2 - 4} \right] \qquad \text{for } n_2 > 4.$$

When $\lambda = 0$, these reduce to the mean and variance of the central $F(n_1, n_2)$ distribution. Derivation of (2.5) is shown in [Ref. 1:p. 189].

The cumulative distribution of V can be expressed in terms of an infinite series of multiples of incomplete beta functions, as follows:

$$\Pr[V \leq f_0] = \sum_{j=0}^{\infty} \left(\frac{[\frac{1}{2}\lambda_1]^j}{j!} e^{-\frac{1}{2}\lambda_1} \right) \cdot I_{\frac{n_1 f_0}{n_2 + n_1 f_0}}\left(\tfrac{1}{2} n_1 + j, \tfrac{n_2}{2}\right) \tag{2.6}$$

where $I_{\frac{n_1}{n_2 + n_1 f_0}}(\frac{1}{2} n_1 + j, \frac{n_2}{2})$ is the incomplete beta function.

In this thesis, we use Paulson's Approximation from Severo and Zelen [Ref. 2]. Their approximation is as follows:

$$\Pr(V \leq f_0) \sim \Phi(x)$$

where

$$X = \frac{(1 - \frac{2}{9n})(\frac{n_1 f_0}{n_1 + \lambda_1})^{\frac{1}{3}} - [1 - \frac{2}{9}(n_1 + 2\lambda_1)(n_1 + \lambda_1)^{-2}]}{[\frac{2}{9}(n_1 + 2\lambda_1)(n_1 + \lambda_1)^{-2} + \frac{2}{9} n_2^{-1} (\frac{n_1 f_0}{n_1 + \lambda_1})^{\frac{2}{3}}]^{\frac{1}{2}}} \text{, and} \tag{2.7}$$

Φ is the standard normal CDF.

The following table shows some values obtained using this approximation, together with exact values of Pr (V ≤ fo). It can be seen in Table 1 that this approximation gives about 2 decimal place accuracy. Thus, this approximation provides adequate acçuracy for computing the power of F-tests [Ref. 3:p. 202].

Table 1. CDF of a Noncentral F Distribution

n_1	n_2	λ	fo	approx	exact
3	10	4	3.708	0.7499	0.745
		4	6.552	0.9192	0.918
		16	3.708	0.2023	0.206
		16	6.552	0.190	0.517
3	20	4	3.098	0.7067	0.700
		4	4.938	0.8894	0.887
		16	3.098	0.1186	0.126
		16	4.938	0.3488	0.347
5	10	6	3.326	0.7337	0.731
		6	5.636	0.9143	0.914
		24	3.326	0.1553	0.158
		24	5.636	0.4629	0.461
5	20	6	2.711	0.6685	0.664
		6	4.103	0.8715	0.870
		24	2.711	0.0643	0.069
		24	4.1030	0.2437	0.245
8	10	9	3.072	0.7159	0.714
		9	5.057	0.9087	0.908
		36	3.072	0.1166	0.119
		36	5.057	0.4088	0.408
8	30	9	2.266	0.581	0.578
		9	3.173	0.8157	0.813
		36	2.266	0.0146	0.017
		36	3.173	0.0846	0.088

3. Description of the Power Program

3-1 Procedure Overview

The interactive program included in this thesis is written in FORTRAN 77. It was written for use on an IBM370 from IBM3278 terminal. It is an interactive program. The command 'THVS' is all that is required to start this program. The executive file name initializes the virtual machine environment and asks the user questions about the choice of program and compilation requirements. It then activates the selected program. The 'ANOVA' is interactive programs which will be discussed in detail in later chapters. The output from the programs are presented on the terminal screen.

3.2 Instructions for Program Access

To start the program, make sure you have loaded (THVS, ANOVA) on your disk. In CMS (operating system mode), type 'THVS' and you will see:

Please provide the FILENAME for your VS FORTRAN program.

Now type the program name you want, for example 'ANOVA', the response is:

Do you need to compile your program? (y/n)

If you want to run, type 'Y' and your program will be loaded.

A detailed description of the ANOVA (Analysis of Variance Test) is given in Chapter IV.

Following the screen output from the selected test, the user will be asked some questions about the output. We will see the following:

Do you wish to BROWSE your output? (Y)

n

Print your output file? (Y)

n

Do you wish to XEDIT the program file? (Y/N)

n

Do you wish to run the program again? (Y)

n

Then return to CMS mode.

4. Power of the F-Test

4.1 Introduction

Under appropriate conditions, the best test for testing equality of several means is the analysis of variance. The analysis of variance has a wide application. It is one of the most useful techniques in the field of statistical inference. As in any hypothesis-testing situation, the power of the F test is of interest to the experimentor.

In this chapter we will discuss the power of F tests and provide an example. To give an overall evaluation of the power of F tests in the analysis of variance, we may use power curves. An important use of the power curve is to guide the experimentor in selecting the sample size (number of replicates) so that the design will be sufficiently sensitive to important potential differences in the treatments. We will consider power curves for one-way and multi-way analyses of variance(ANOVA's).

4.2 The One-way Classification Analysis of Variance

Suppose we have k different levels of a single factor that we wish to compare. The different levels of the factor are often called treatments. The observed responses from each of k treaments is random a sample on a random variable. The data would appear as in Figure 4-1.

Observation

Treatment					
	1	Y_{11}	Y_{12}	...	Y_{1n}
	2	Y_{21}	Y_{22}	...	Y_{2n}
	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.
	k	Y_{k1}	Y_{k2}	...	Y_{kn}

Figure 4-1. Typical Data for One-way Classification Analysis of Variance

We will find it useful to describe the observations by the linear statistical model:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \qquad (4.1)$$

where Y_{ij} is the $(ij)^{th}$ observation, μ is a parameter common to all treatments called the overall mean, ti is a parameter peculiar to the i^{th} treatment called the i^{th} treatment effect, and ε_{ij} is a random error component, assumed to be IID $N(0, \sigma^2)$.

The indices used are:

- i = the number of treatments, i = 1, 2, ..., k
- j = the sample size per treatment, j = 1, 2, ..., n

The objective in the ANOVA is to test appropriate hypotheses about the treatment effects. The variance σ^2 is assumed constant for all levels of the factor. This model is called the one-way classification analysis of variance because only one factor is investigated.

We are interested in testing the equality of the k treatment effects, so the appropriate hypothes are

$$H_o : \tau_1 = \tau_2 = ..., = \tau_k = \tau$$

$$H_a : \tau_i \neq \tau_j \text{ for some i, j}$$

That is, if the null hypothesis is true, then each observation is made up of the mean $\mu + \tau$ plus a realization of the random error ε_{ij}

The results of the ANOVA procedure is summarized in Table 3.

We may use the expected values of the mean squares to verify that Fo (in Table 3) is an appropriate test statistic for Ho. From the expected mean square we see that, in general, MSE is unbiased estimator of σ^2. Also, under the null hypothesis, MSt is an unbiased estimator of σ^2. However, if the null hypothesis is false, then the expected valued of MSt is greater than σ^2. The expected value of the mean square error is σ^2, and the expected value of the mean square between treatments is $\sigma^2 + (n\,\Sigma(t_i)^2)/(k-1)$.

Therefore, under the alternate hypothesis the expected value of the numerator of the test statistic (F_0) is greater than the expected value of the denominator and we would reject Ho with values of the test statistic which are too large. That is, we would reject Ho if

$$F_0 > F_{\alpha,(k-1),(N-k)}$$

where degrees of freedom 1 is K-1, degrees of freedom 2 is k(n-1) = N-k, and α is the type I error rate.

Table 3. The Analysis of Variance for a One-way Classification

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Between Treatments	SSt	k-1	MSt	MSt/MSE
Error (within te treatments)	SSE	k(n-1)	MSE	
Total	SST	kn-1		

The power of the test is:

$$1-\beta = P(Fo > Fa,(K-1),(N-k) \mid Ho \text{ is false}) \quad (4.2)$$

To evaluate the β in Equation 4.2 we need to know the distribution of the test statistic F_0 if the null hypothesis is false. It can be shown that, if Ho is false, the statistic F_0 has the noncentral F distribution with k-1 and k(n-1) degrees of freedom and noncentrality parameter λ, given by

$$\lambda = \frac{n\,\Sigma(\tau i)^2}{\sigma^2}$$

The noncentrality can be interpreted as the squared standardized distance between the origin and $(\tau_1, \tau_2, \cdots, \tau_k)$. The ratio $\Sigma(\tau i)^2/\sigma^2$ is called the squared standardized distance. If only an estimate of σ^2 is available, one may replace σ^2 with the estimate [Ref. 8:p. 34].

4.3 The Multi-way Classification Analysis of Variance

Many experiments require a study of the effects of two or more factors. It can be shown that, under certain conditions, factorial arrangements are the most efficient designs for this type of analysis.

One of the simplest factorial experiments involves only two factors or sets of treatments say factor A and factor B. Suppose there are a levels of factors A and b levels of factor B, and these are arranged in a 2-way factorial design; that is, each replication of the experiment contains all ab treatment combinations. Assume there are n replications of the experiment, and let Y_{ijk} represent the observation taken under the i^{th} level of factor A and the j^{th} level of factor B in the k^{th} replication.

The data can be summarized as shown in Figure 4.2. The order in which the abn observations are taken is selected at random.

Factor B

Factor A					
	1	Y_{111}, Y_{112} ,..., Y_{12n}	Y_{121}, Y_{122} ,..., Y_{12n}	...	Y_{1b1}, Y_{1b2} ,..., Y_{1bn}
	2	Y_{211}, Y_{212} ,..., Y_{21n}	Y_{221}, Y_{222} ,..., Y_{22n}	...	Y_{2b1}, Y_{2b2} ,..., Y_{2bn}
	. . .	...	...	. . .	...
	a	Y_{a11}, Y_{a12} ,..., Y_{a2n}	Y_{a21}, Y_{a22} ,..., Y_{a2n}	...	Y_{ab1}, Y_{ab2} ,..., Y_{abn}

Figure 4.2 Typical Data notation for a Two-way Classification.

The observations may be described by the linear model:

$$Y_{ijk} = \mu + \tau_i + \beta_j + \tau\beta_{ij} + \varepsilon_{ijk}, \tag{4.3}$$

where

- μ = the overall mean effect;
- τ_i = the true effect of i^{th} level of factor A, i = 1, 2, ..., a;

- β_j = the true effect of j^{th} level of factor B, j = 1, 2, ..., b;
- $\tau\beta_{ij}$ = the effect of the interaction between t_i and β_j; and
- ϵ_{ijk} = a random error component, assumed to be IID $N(0, \sigma^2)$.

Both factors are assumed to be fixed. It is usually assumed that the treatment effects are defined as deviations from the overall mean, so $\Sigma t_i = 0$ and $\Sigma\beta_j = 0$. Similarly, the interaction effects are fixed and usally defined so that $\Sigma(t\beta)_{ij} = 0$, where the summation is over either i or j. Since there are n replicates of the experiment, there are a total of abn observations.

We are interested in testing various hypothesis about the parameters in equation 4.3. An appropriate hypothesis testing procedure would be the analysis of variance. More spicifically, as we are considering two controllable sources of variation (A and B), the procedure is called the two-way classification analysis variance.

In order to test the hypothesis Ho; $t_i = 0$ for i=1, 2, ..., a (no row factor effects), Ho; $\beta_j = 0$ for j = 1, 2, ..., b (no column factoreffects), and Ho; $(t\beta)_{ij} = 0$ (no interaction effects), we can expresses the total sum of square as:

$$SST = SSA + SSB + SSAB + SSE. \tag{4.4}$$

Here, SSA is a sum of squares due to "rows" or factor A, SSB is a sum of squares due to "columns" or factor B, SSAB is a sum of squares due to the interaction between A and B, and SSE is a sum of squares due to error. The number of degrees of freedom associated with each sum of squares are shown in Figure 4.3.

If we assume ϵ_{ijk} are IID $N(0, \sigma^2)$ and apply Cochran's theorem (Theorem 3-1) under the null hypothesis of no effects, each sum of squares on the right-hand side of Table IV when divided divided by σ^2 is distributed as X^2 with the indicated number of degrees of freedom, and these statistics are independent.

Theorem 3-1 (Cochran). Let Z_i be IID N(0, 1) for i=1,2, ..., ν and suppose $\Sigma Z_i^2 = Q_1 + Q_2 + Q_3 + \cdots + Q_s$ where $s \leqslant \nu$, and the quadratic form Q_i has ν_i degrees of freedom (i=1,2, ..., s). Then the Q_1, Q_2, ..., Q_s are independent chi-square random variable with $\nu_1, \nu_2, \ldots, \nu s$ degrees of freedom, respectively, if only if

$$\nu = \nu_1 + \nu_2 + \ldots + \nu_s$$

Effect	Degrees of freedom
A	a-1
B	b-1
AB interaction	(a-1) (b-1)
Error	ab(n-1)
Total	abn-1

Figure 4.3 Table of degrees of freedom with sums of squares

Assuming that factors A and B are fixed, the expected values of the mean squares are:

$$E(MSA) = \sigma^2 + (bn \Sigma \tau_i^2)/(a-1) ; \quad (4.5)$$

$$E(MSB) = \sigma^2 + (bn \Sigma \beta i^2)/(b-1) ; \quad (4.6)$$

$$E(MSAB) = \sigma^2 + (n \Sigma \Sigma (\tau\beta)_{ij}^2)/(a-1)(b-1); \text{ and} \quad (4.7)$$

$$E(MSE) = \sigma^2 \quad (4.8)$$

Therefore, to test the hypothesis Ho: $t_i = 0$; i=1,2, ..., 1 (no rows factor effects), and Ho; $\beta_j = 0$; j = 1, 2, ..., b (no column factor effects), and Ho: $(t\beta)_{ij} = 0$ (no interaction effects), we would divide the corresponding mean square by mean square error. Under the null hypothesis of no effect, this ratio will follow an F distribution with appropriate numerator degrees of freedom and ab(n-1) denominator degrees of freedom, and the critical region will be located in the upper tail. The test procedure is usually summarized in an analysis of variance table, such as shown in Table IV.

Table 4. The Analysis of Variance for a Two-way Classification

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_o
A treatments	SSA	a-1	MSA	MSA/MSE
B treatments	SSB	b-1	MSB	MSB/MSE
Interaction	SSAB	(a-1) (b-1)	MSAB	MSAB/MSE
Error	SSE	ab (n-1)	MSE	
Total	SST	abn - 1		

To the compute the power of tests in two-way ANOVAs. the procedure is the same as in the one-way case. Relationships among λ, the numerator degrees of freedom and the denominator degrees of freedom are shown in Table V.

In a similar way one can expand to the general multi-way ANOVA procedure [Ref. 8:p. 124].

4.4 The Algorithms and Flowchart

A program for computing a power table and power curve in general ANOVA's (fixed model) is shown in the Appendix F. The power of one-way ANOVA's are solved with these programs, using the following the sequence:

Step 1: Determine what variable to includesuch as sample size vs power, number of treatments vs power, α-level vs power, or noncentrality vs power.

Step 2: Determine the values of the maximum, minimum, and increment of the variable chosen.

Step 3: Given input data, compute the critical value, using the inverse of the central F-distribution

Step 4: Compute the power value in each case; that is, the CDF of noncentral F-distribution (Chapter 2).

Step 5: Print th power table or the power curve.

Table 5. Noncentrality Parameters for Power in a Two-way Anova

Factor	λ	Numerator (DOF)	Denominator (DOF)
A	$\frac{bn \sum \tau_j^2}{\sigma^2}$	a-1	ab (n-1)
B	$\frac{an \sum \beta_j^2}{\sigma^2}$	b-1	ab (n-1)
AB	$\frac{n \sum \sum \tau\beta_{ij}^2}{\sigma^2}$	(a-1) (b-1)	ab (n-1)

Multi-way ANOVAs have the same algorithms but multi-way ANOVA may include tests of interaction effects. The program considers only up to three-way interaction effects. We can explain how to compute degrees of freedom of error term in m-way ANVOVA's, assuming the balanced case, as follows: Total degrees of freedom is DOF (Total) = n Σ k(i) -1, where k(i) is the number of levels of the i^{th} factor [Ref. 1].

(1) If the model has only main effects without any interaction effects,

DOF 1 (Error) = DOF (Total) – Σ (k(i)-1).

(2) If the model has several factors and only 2-way interaction effects,

DOF 2 (Error) = DOF 1 (Error) – ΣΣ (k(i)-1) (k(j)-1).

(3) If the model has several factors and only 3-way interaction effects,

DOF 3 (Error) = DOF 1 (Error) – ΣΣΣ (k(i)-1) (k(j)-1) (k(k)-1).

(4) If the model has several factors and 2-way and 3-way interaction effects,

DOF 4 (Error) = DOF 3 (Error) – ΣΣ (k(i)-1) (k(j)-1).

If the model has more than 3-way interaction effects, then the user must modify the degrees of freedom of the error terms accordingly. A flowchart is shown in Figure 4.4

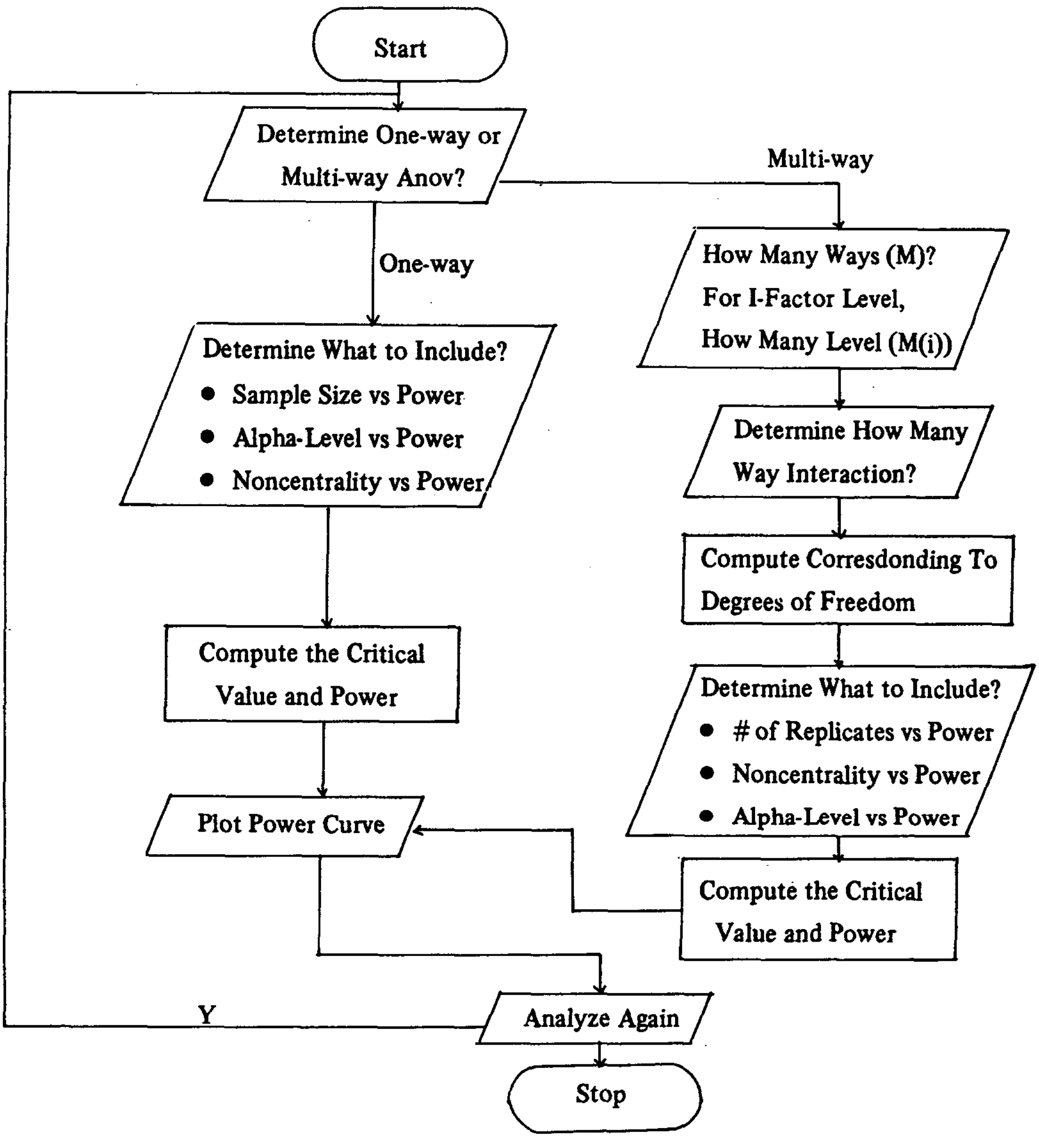


Figure 4.4. System Flowchart for F-Test Power

4.4 Examples of F-test

1. *Example of Onw-way ANOVA test (number of replicates vs power)*

   a. Scenario

A manufacturer suspects that the batches of raw materi..l urnished by his supplier differ significantly in calcium content. There are a large number of batches currently in the warehouse. Five of those are randomly selected for study. A chemist wants to know the appropriate sample size per each batch, in order to test

$$Ho : \tau_1 = \tau_2 = \cdots = \tau_k$$

against $$Ha : \tau i \neq \tau j \text{ for some } i,j$$

The chemist would like to know how many replicates to run if it is important to reject Ho with probability at least 0.9 when the standardized distance square ($\sum \tau i^2 / \sigma^2$) is 2 and $\alpha = .05$. Thus he would like to know what the power of the F-test is for a range of possible replicates. He decides to check replicates from 2 to 12 in increments of 1.

b. Inputs

Step 1: Select one-way ANOVA test (the number of replicates vs power).
Step 2: The number of treatment = 5.
Step 3: α - level = .05
Step 4: Standardized distance square = 2.
Step 5: Maximum replicates is 12, minimun replicates is 2, increment is 1.

c. Start Program

Do you want to analyze one way ANOVA (y/n)?
y
Do you want to plot n (# of observation per treatment vs power) (y/n)?
y
Number of k (# of treatment)?
?
5
Alpha-level?
?
.05
Standardized distance square value?
?
2.
Maximum n value (the maximum value on X-axis)?
?
12
Minimum n value (the minimum value on X-asix)?
(condition: N must be more than 2)
?
2

Increment n value?

?

1

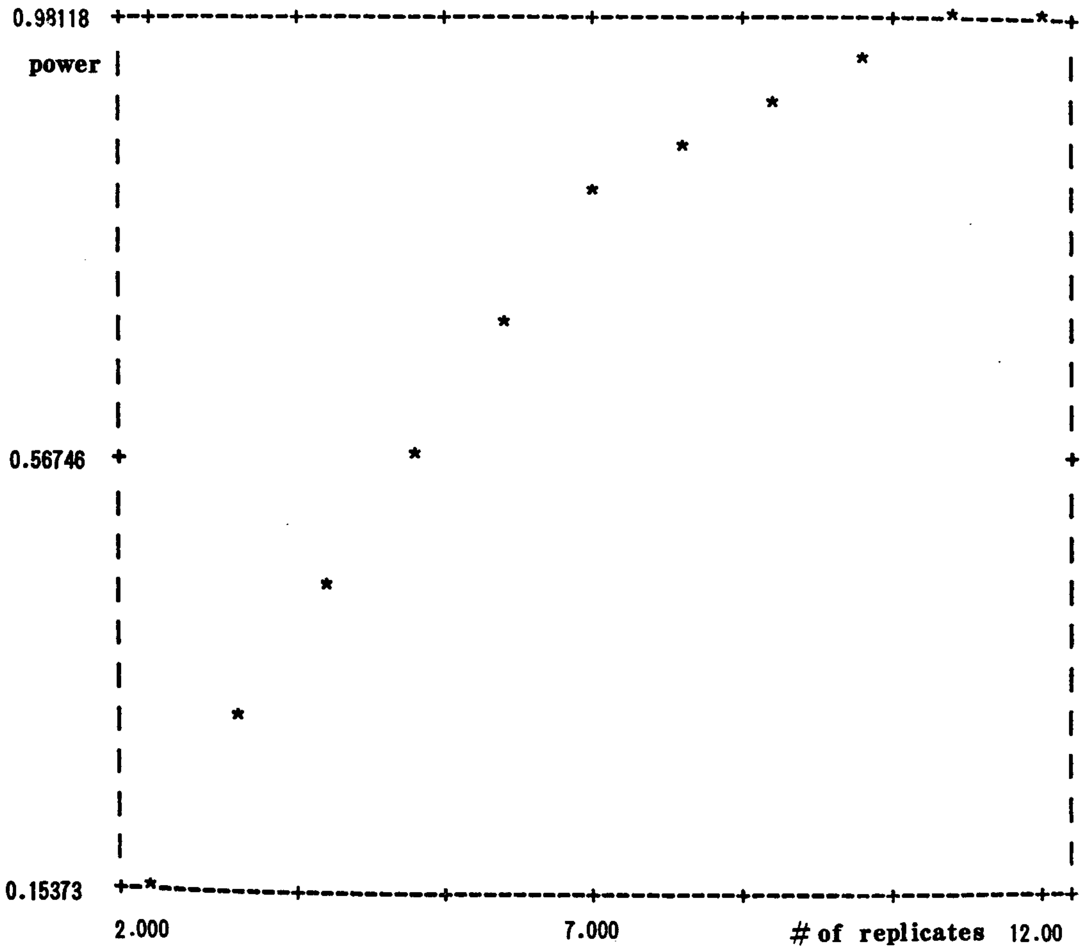
d. Output

The screen output (Table 8) is as follows:

Table 8. Output of the One-way Anov Example (Replicates vs Power)

dof 2	dof 2	α	F-inverse	λ	power
4	5	0.05	5.29087	4.00	0.15373
4	10	0.05	3.52496	6.00	0.30024
4	15	0.05	3.10397	8.00	0.44935
4	20	0.05	2.91676	10.00	0.58677
4	25	0.05	2.81108	12.00	0.70332
4	30	0.05	2.74323	14.00	0.79562
4	35	0.05	2.69600	16.00	0.86449
4	40	0.05	2.66122	18.00	0.91328
4	45	0.05	2.63455	20.00	0.94631
4	50	0.05	2.61345	22.00	0.96776
4	55	0.05	2.59634	24.00	0.98118

# of replicate	power
NN = 2	power = 0.15373
NN = 3	power = 0.30024
NN = 4	power = 0.44935
NN = 5	power = 0.58677
NN = 6	power = 0.70332
NN = 7	power = 0.79562
NN = 8	power = 0.86449
NN = 9	power = 0.91328
NN = 10	power = 0.94631
NN = 11	power = 0.96776
NN = 12	power = 0.98118



2. *Example of one-way ANOVA test (noncentrality vs power)*

a. Scenario

Five brands of batteries are under study. It is suspected that the lie (in weeks) of the five brands is different. Five batteries of each brand are tested. A manufacturer want to know that power as a function of the standardized distance square.˙ That is, the experimenter would like to know what the power of the F-test is for a range of possible standardized distances square. He decides to check standardized distances square from 1 to 5 in .2 increment, where n=5 and $\alpha = .05$.

b. Inputs

Step 1: Select one-way ANOVA test (the standardized distance square vs power)

Step 2: Number of treatment = 5

Step 3: Number of replicates = 5

Step 4: α - level = .05

Step 5: Maximum standardized distance square is 5 and minimum standardized distance square is 1 and the increment is .2.

c. Start Program

Start program:

Do you want to analyze one-way ANOVA (y/n)?

y

Do you want to plot n (# of observation per treatment) vs power (y/n)?

n

Do you want to plot alpha-level vs power (y/n)?

n

Do you want to plot noncentrality vs power (y/n)?

y

# of observations per treatment (n =) ?

?

5

# of treatment (k =)?

?

5

Alpha-level ?

?

.05

Maximum standardized distance square value range?

?

5

Minimum standardized distance square value range?

?

1

Increment standardized distance square value range?

?

.2

d. Output

The screen output (Table 9) is as follows:

Table 9. Output of the One-way Anova Example (noncentrality vs power)

dof 1	dof 2	α	F-inverse	λ	power
4	20	0.05	2.91676	5.00	0.30382
4	20	0.05	2.91676	6.00	0.36314
4	20	0.05	2.91676	7.00	0.42203
4	20	0.05	2.91676	8.00	0.47946
4	20	0.05	2.91676	9.00	0.53461
4	20	0.05	2.91676	10.00	0.58677
4	20	0.05	2.91676	11.00	0.63550
4	20	0.05	2.91676	12.00	0.68050
4	20	0.05	2.91676	13.00	0.72163
4	20	0.05	2.91676	14.00	0.75885
4	20	0.05	2.91676	15.00	0.79223
4	20	0.05	2.91676	16.00	0.82192
4	20	0.05	2.91676	17.00	0.84812
4	20	0.05	2.91676	18.00	0.87108
4	20	0.05	2.91676	19.00	0.89107
4	20	0.05	2.91676	20.00	0.90835
4	20	0.05	2.91676	21.00	0.92321
4	20	0.05	2.91676	22.00	0.93591
4	20	0.05	2.91676	23.00	0.94672
4	20	0.05	2.91676	24.00	0.95586
4	20	0.05	2.91676	25.00	0.96356

noncentrality	power
NN = 5.0	power = 0.30382
NN = 6 0	power = 0.36314
NN = 7.0	power = 0.42203
NN = 8.0	power = 0.47946
NN = 9.0	power = 0.53461
NN = 10.0	power = 0.58677

NN = 11.0 power = 0.63550

NN = 12.0 power = 0.68050

NN = 13.0 power = 0.72163

NN = 14.0 power = 0.75885

NN = 15.0 power = 0.79223

NN = 16.0 power = 0.82192

NN = 17.0 power = 0.84812

NN = 18.0 power = 0.87108

NN = 19.0 power = 0.89107

NN = 20.0 power = 0.90835

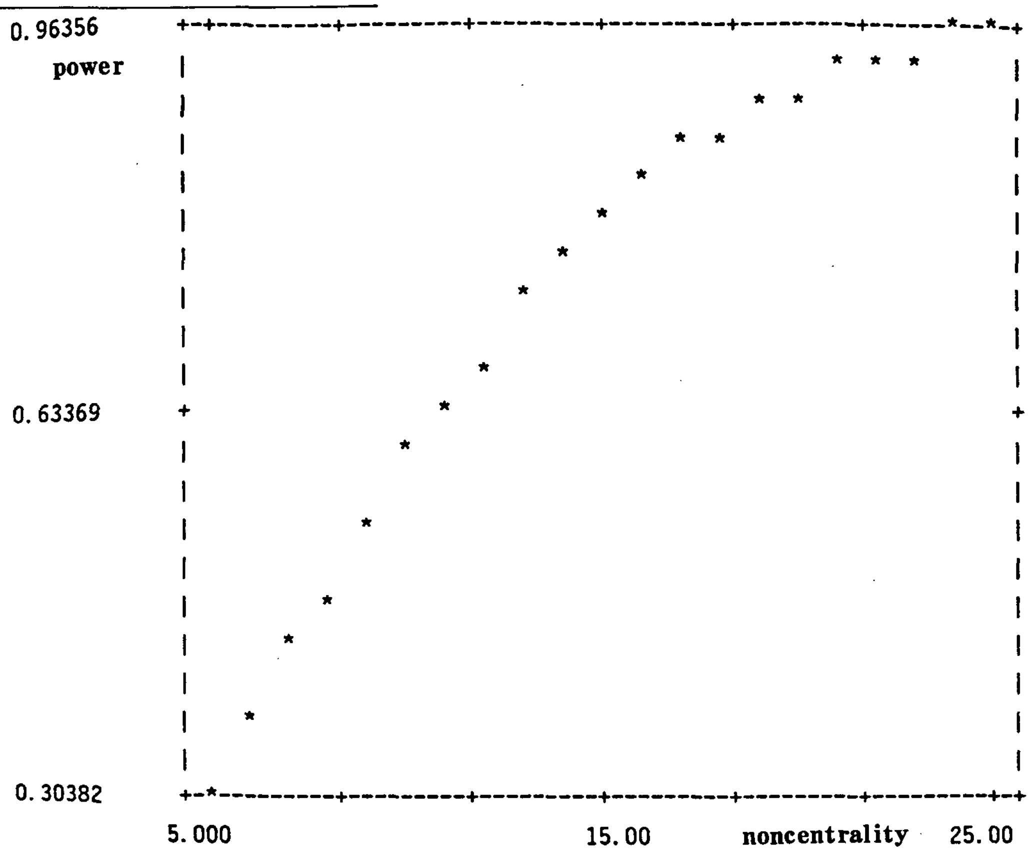
NN = 21.0 power = 0.92321

NN = 22.0 power = 0.93591

NN = 23.0 power = 0.94672

NN = 24.0 power = 0.95586

NN = 25.0 power = 0.96356



x-scale: "-" = 0.250E + 00 units

y-scale: "I" = 0.110E -01 units

5. Summary and Conclusions

Power considerations are useful in the design and assessment of statistical tests. The calculation of power usually involves noncentral distributions for which tables of probability are not available. A search was made for algorithms approximating the noncentral t, F and X^2 distributions. Algorithms giving sufficient accuracy and making efficient case of computer resources have been implemented in this thesis. Listings of the FORTRAN code for these implementations are included.

An interactive program to compute and display power curves for several t-test and F-test situations has been developed. This program is user friendly and is described in this thesis; a listing of the program is provided. This program should be useful to researchers, experiment designers and statisticians.

References

1. Searles, S.R, *"Linear Models,"* New York, John Wiley & Sons, Inc., 1971.
2. Johnson, N.L., Kotz, S. *"Continuous Univariate Distribution," Vol. 3* New York, Houghton Miflin, 1970.
3. Severo, N. and Zelen, M., *"Normal Approximation to the Chi-square and Noncentral F Probability function,"* Biometrika 52, 1960.
4. Patnaik, P.B., *"The Noncentral x^2 and F Distributions and Their Applications,"* Biometrika 36, 1949.
5. Bowker and Lieberman, *"Engineering Statistics,"* Englewood, Cliffs, N.J., Prentice-Hall. Inc., 1960.
6. Owen, D.B., *"The Power of Student's t-Test,"* Journal of the American Statistical Association, Vol. 60, 1965.
7. Abramowitz, M., an Stegun, I.A., *"Handbook of Mathematical Functions,"* Dover Publications, Inc., New York, 1965.
8. Douglas C. Montgomery, *"Design and Analysis Experiments,"* John Wiley and Sons, Inc., 1976.
9. Johnson, N.L. and Welch, B.L., *"Applications of the Noncentral t-Distribution,"* Biometrika 31, 1943.
10. Charles B. Kreitzberg and Ben Shneiderman, *"FORTRAN Programming (A Spiral Approach),"* Harcourt Brace Jovanovich, Inc., 1982.