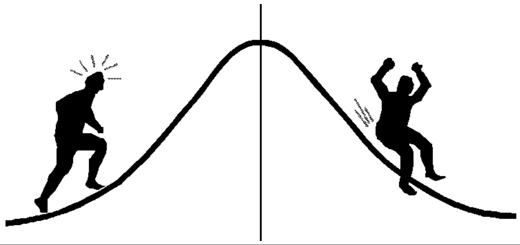


BIOMETRICS INFORMATION

Power Analysis Workshop

What is power? . . . Why should we use it?

Revised August 2001

by

Wendy Bergerud
Vera Sit
Biometrics Section
Forest Science Research Branch
B. C. Ministry of Forests

These workshop notes are based on the Biometrics Information Handbook #2 titled Power Analysis Handbook for the Design and Analysis of Forestry Trials by Amanda F. Linnell Nemec. This handbook is the basic text for this workshop.

British Columbia
Ministry of Forests
Research Program

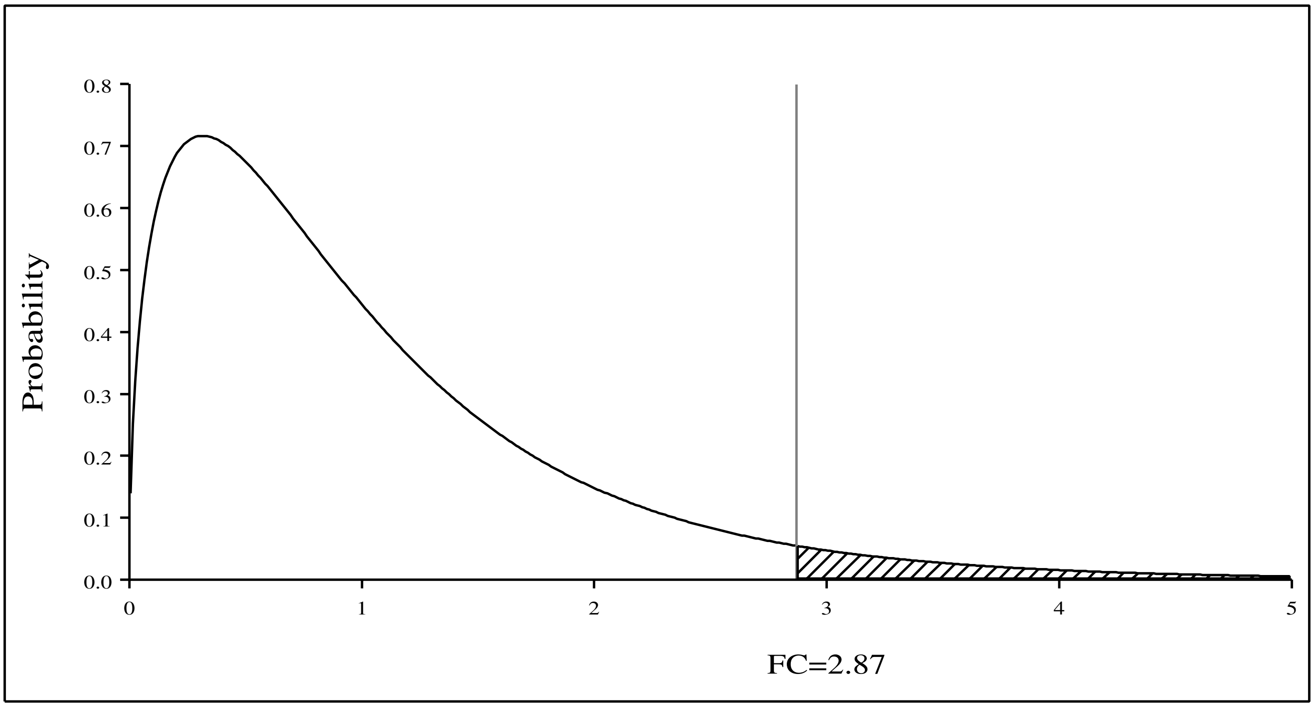
TABLE OF CONTENTS

APPENDICES

What is Power? . . . Why should we use it?

1 The Null Hypothesis and α

First, let's remind you of something that you are familiar with that is a lot like power. As an example, let's use a one-way experiment using ANOVA for data analysis. Suppose that we had **a** = 4 groups or treatments with a sample size of 10 t.u.'s per group. What does it mean when an F-value of 3.81 is observed? Since this F-value has 3, 36 df the probability value is 0.018. Recall that the null hypothesis associated with this ANOVA is that there are no group or treatment differences. If we were able to do this same experiment a 100,000 times, and each time the null hypothesis were true, we would observe the following frequency curve of observed F-values:



This is called the F-distribution and the critical F-value, FC, is chosen so that the area to the right of FC is equal to α. If $\alpha = 0.05$ then FC = 2.87 and if $\alpha = 0.01$ then FC = 4.38. The probability of a single observed F-value is the area under the curve to the right of the observed value. For F = 3.81 this probability is p = 0.018. This probability value implies that 18 out of 1,000 times this large or larger of an F-value would be observed even though the null hypothesis is true. In this case, we would argue that this large an F-value is unlikely if the null hypothesis were true. Thus we reject the null hypothesis. We might state in a report that the experiment presents strong evidence against the null hypothesis.

EXERCISE: Use SAS to obtain p-values and FC values.

```
/* Fprobs.sas */

data fprobs;
  input alpha @@;
  dfh  = 3;   dfe = 36;
  FC   = finv(1-alpha, dfh, dfe, 0);
  prob = 1 - probf(FC, dfh, dfe, 0);
cards;
0.20 0.10 0.05 0.025 0.01 0.005 0.001
;
proc print;
title 'Listing of Critical F-values for specific alpha values';
title2 'Check that the prob and alpha values are identical';
run;
```

OUTPUT:

```
            Listing of Critical F-values for specific alpha values           1
              Check that the prob and alpha values are identical

                OBS     ALPHA     DFH     DFE         FC         PROB

                 1      0.200      3       36     1.62781     0.200
                 2      0.100      3       36     2.24261     0.100
                 3      0.050      3       36     2.86627     0.050
                 4      0.025      3       36     3.50468     0.025
                 5      0.010      3       36     4.37710     0.010
                 6      0.005      3       36     5.06165     0.005
                 7      0.001      3       36     6.74361     0.001
```

2 <u>When the null hypothesis is not rejected</u>

Suppose, that instead of an F-value of 3.81, a value of 1.72 had been observed. The probability for this value is 0.18. In this case, there is little or no evidence against the null hypothesis. Surely an interesting question to ask is: "What were the chances of rejecting the null hypothesis?" or "How much ability did the experiment have to reject the null hypothesis if, in fact, it were not true?" These questions relate to the POWER of an experiment. *Power is the ability to reject the null hypothesis when it is not true.* The actual power depends on the specific alternate hypothesis under consideration.

3 <u>Power or 1-β</u>

A chosen α-value, or an observed probability value, p, is interpreted as the chance of rejecting the null hypothesis, H_o, when in fact it is true. This is also known as the Type I error rate. The confidence placed in not rejecting H_o is 1-α or 1-p. Similarly, if we assume that a <u>specific</u> alternate hypothesis, H_a, is true for an experiment, the probability of incorrectly accepting H_o is known as β and the confidence placed in rejecting it is 1-β. Observed values will be denoted as p_a and 1-p_a. When looked at from the H_o point of view, 1-β, is the power of the experiment or the ability of the experiment to detect a specific departure (H_a) from H_o. β is also known as the Type II Error. These concepts are conveniently summarized by the following table.

	Reality	Decision	Consequence	Conclusion
Null Hypothesis (H_o)	True	Retain H_o →	Ok	Treatments the same
	True	Reject H_o →	Type I Error (α)	Treatments believed different when same
	False	Retain H_o →	Type II Error (β)	Treatments believed same when different
	False	Reject H_o →	OK	Treatments different

This table is copied from A Statistics Primer for Foresters, by Susan Stafford in Journal of Forestry, March, 1985.

4 The non-centrality parameter, λ or nc

Departures from H_o can be summarized by a noncentrality value or parameter, denoted by nc. We will define nc[1] as

$$\lambda = nc = \frac{SS_{H_a}}{\sigma^2},$$

where σ^2 is the variance or square of the standard deviation. It is usually estimated by the MSE of an ANOVA[2]. SS_{H_a} is the sums of squares of the expected means for H_a, i.e.

$$SS_{H_a} = n \sum(\mu_i - \mu)^2 = n * SSM.$$

where n is the sample size for each mean, the means are denoted by μ_i, and the mean of the means (grand mean) is denoted by μ. $SSM = \sum(\mu_i - \mu)^2$. For H_o, nc = 0. Since nc is always non-negative, the larger it is, the "farther" H_a is from H_o.

As an example, suppose that the expected means for the 4-group ANOVA above were:

$$H_a: 10, 15, 20, 25 .$$

The grand mean for this set is $\mu = (10 + 15 + 20 + 25)/4 = 17.5$, and

$$SSM = [(10-17.5)^2 + (15-17.5)^2 + (20-17.5)^2 + (25-17.5)^2] =$$

---

1 There are different definitions of the non-centrality parameter, so check the definition carefully when reading other literature on power.

2 see Biometrics Information Pamphlet #25: ANOVA: The Within Sums of Squares as an Average Variance.

SSM = [56.25 + 6.25 + 6.25 + 56.25] = 125

hence SS_{H_a} = 10*SSM = 10 * 125 = 1250.

SAS can do this calculation directly using PROC GLM or ANOVA. This is particularly useful when the designs are more complicated. This will be covered in more detail later.

QUESTION: What is the null hypothesis for this experiment?

If the best available estimate of σ^2 is MSE = 100 then λ = nc = 1250/100 = 12.5. This nc-value can now be used to determine the power of this experiment to detect the group means given by H_a above. The following steps accomplish this.

i) Determine the critical F-value for specific value(s) of α.
   i.e. FC = FINV(1-α, df_d, df_h, 0).

ii) Determine the power given the nc, FC , and the df's.
   i.e. POWER = 1 - PROBF(FC, df_d, df_h, nc).

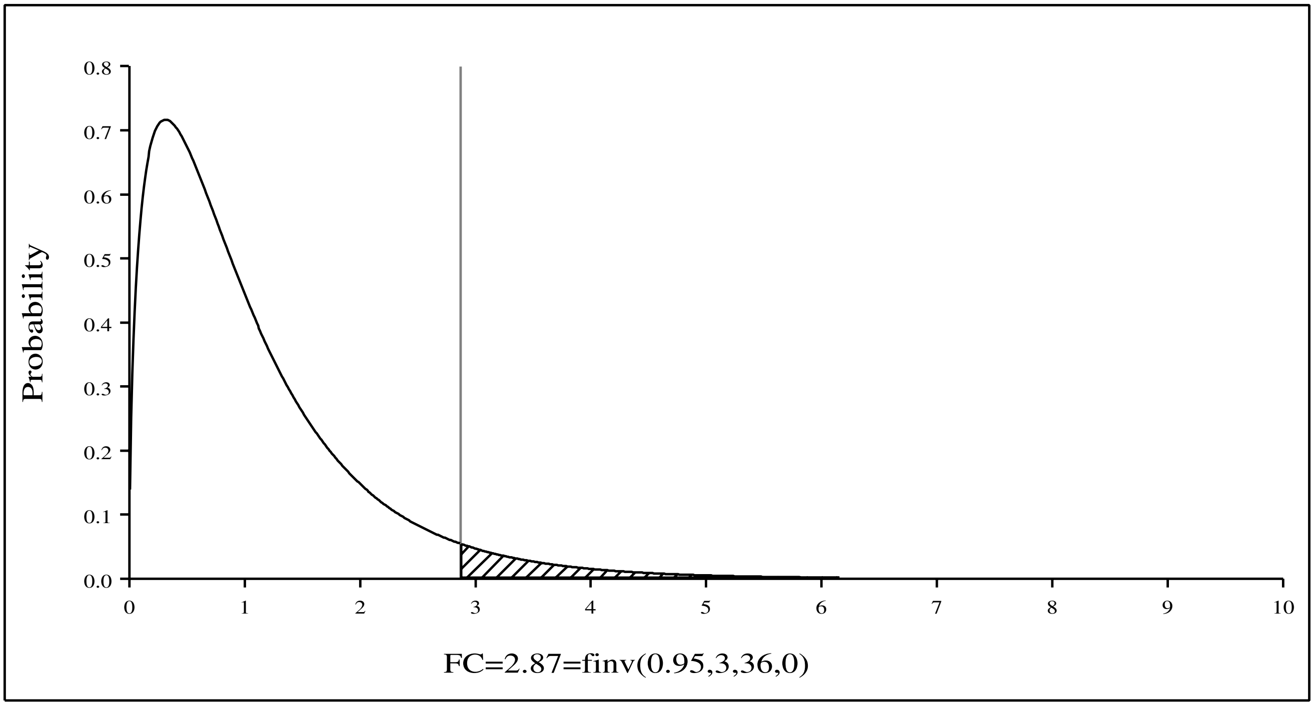
For an α = 0.05, the power for our example is 0.812. This means that H_o will be rejected 812 out of 1,000 times when this specific H_a is true. The corresponding SAS calculations are:

```
/* Power.sas */

data power;
  fc    = finv(1-0.05, 3, 36, 0);
  power = 1-probf(fc, 3, 36, 12.5);
run;
proc print noobs;
title1 'Example in Main Text';
title2 'Results of Power Calculation';
run;
```

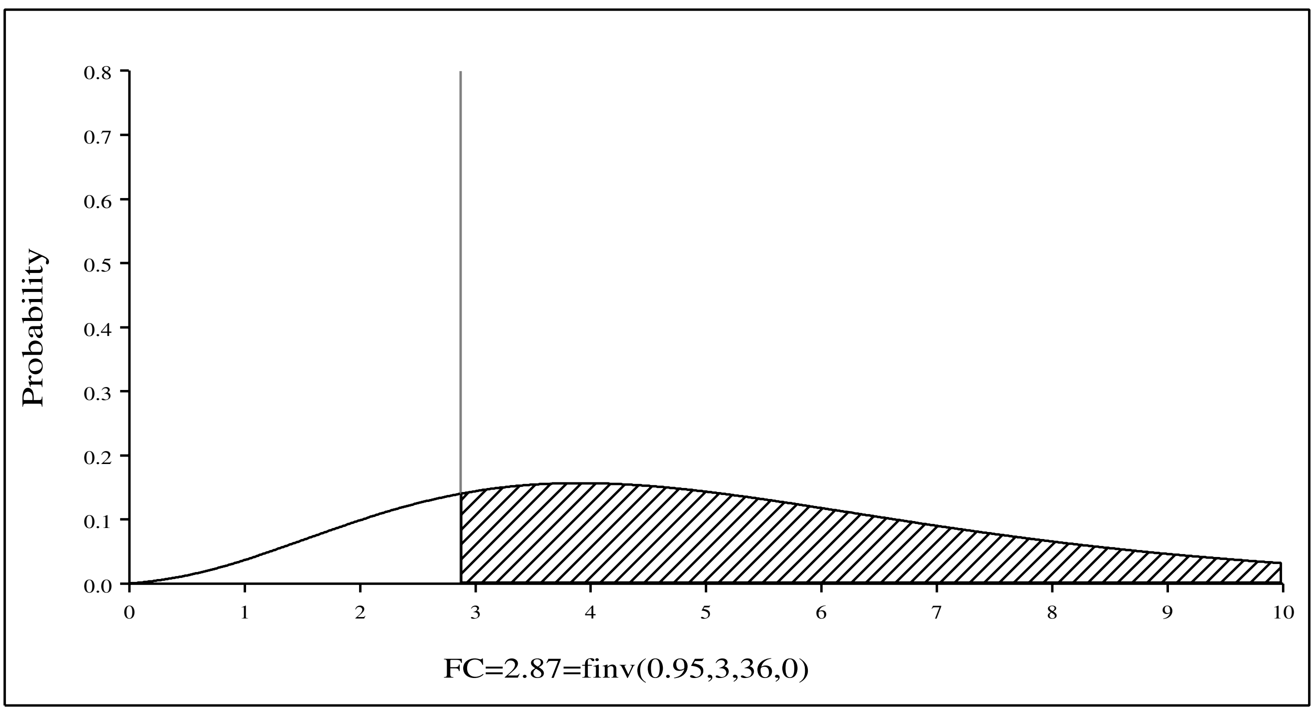
OUTPUT:

```
                             Example in Main Text                            1
                         Results of Power Calculation


                                FC        POWER

                             2.86627    0.81196
```

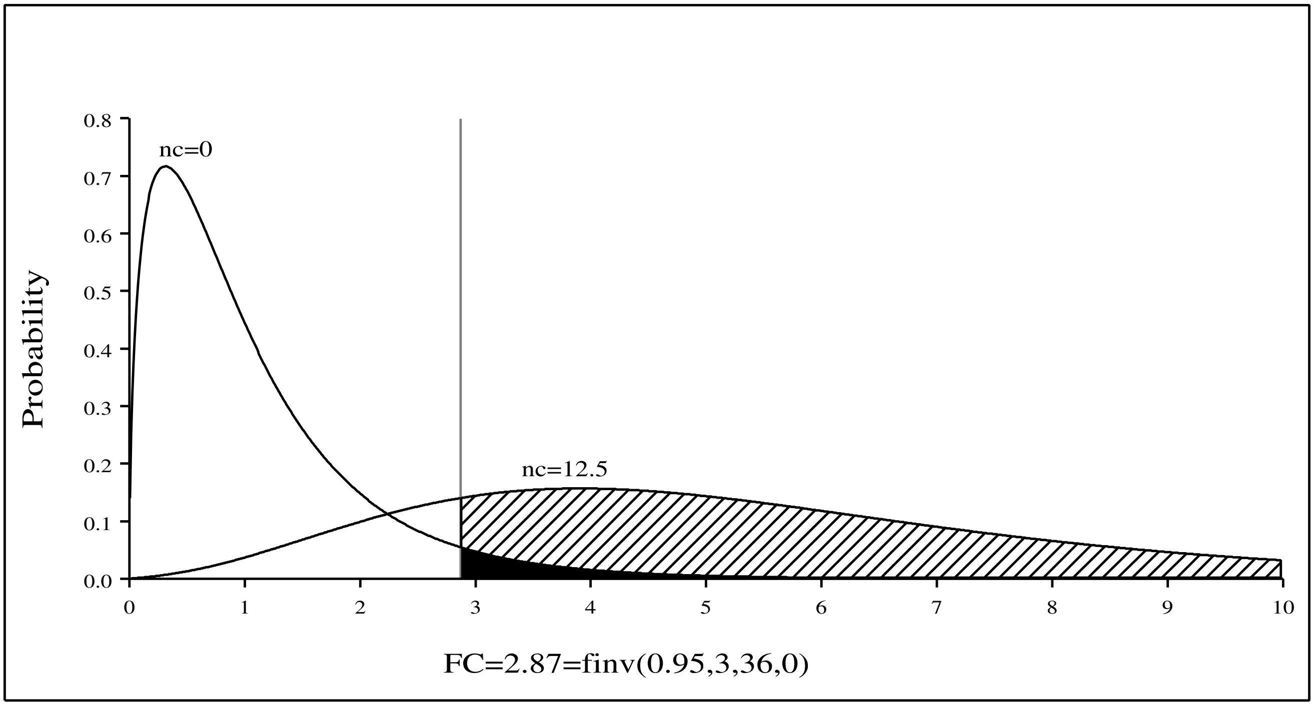
To see more clearly what is happening, let's look again at the F-distribution. If H_o is true for df = 3, 36, then, as before, the F-distribution is (note the different horizontal scale):



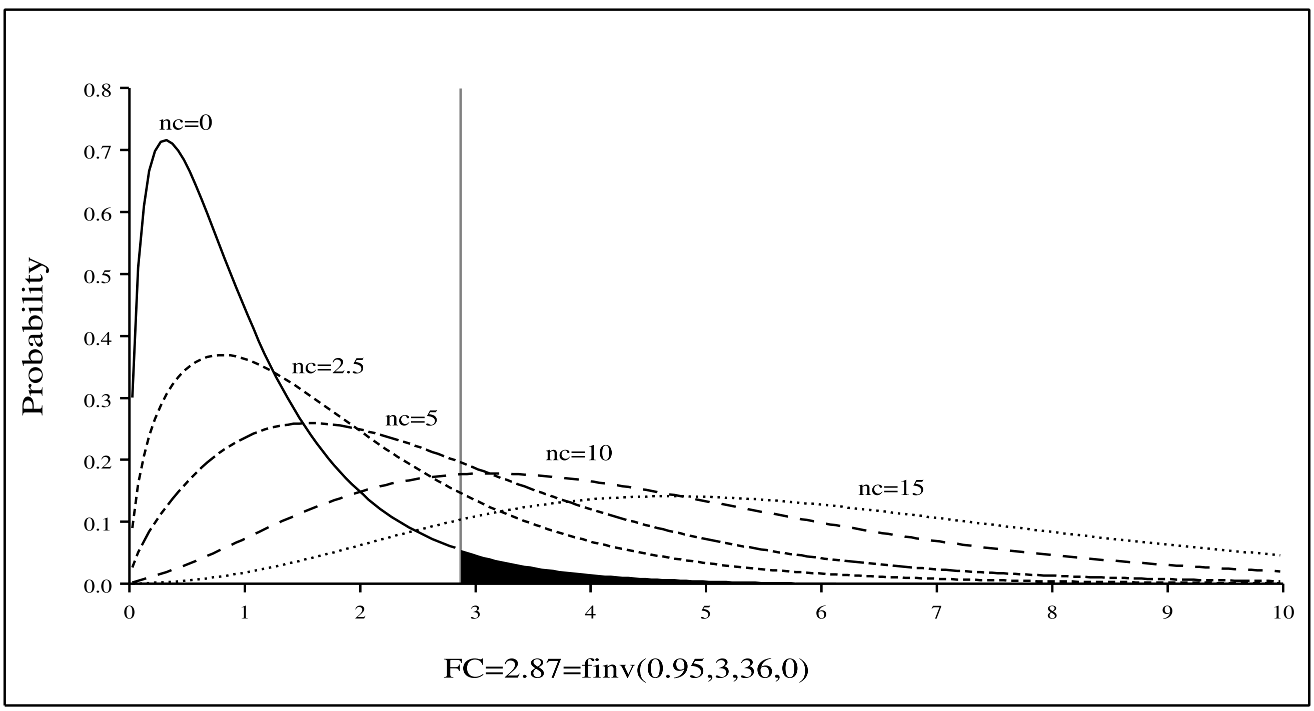
If H_a is true for nc = 12.5 then the corresponding non-central F-distribution is:



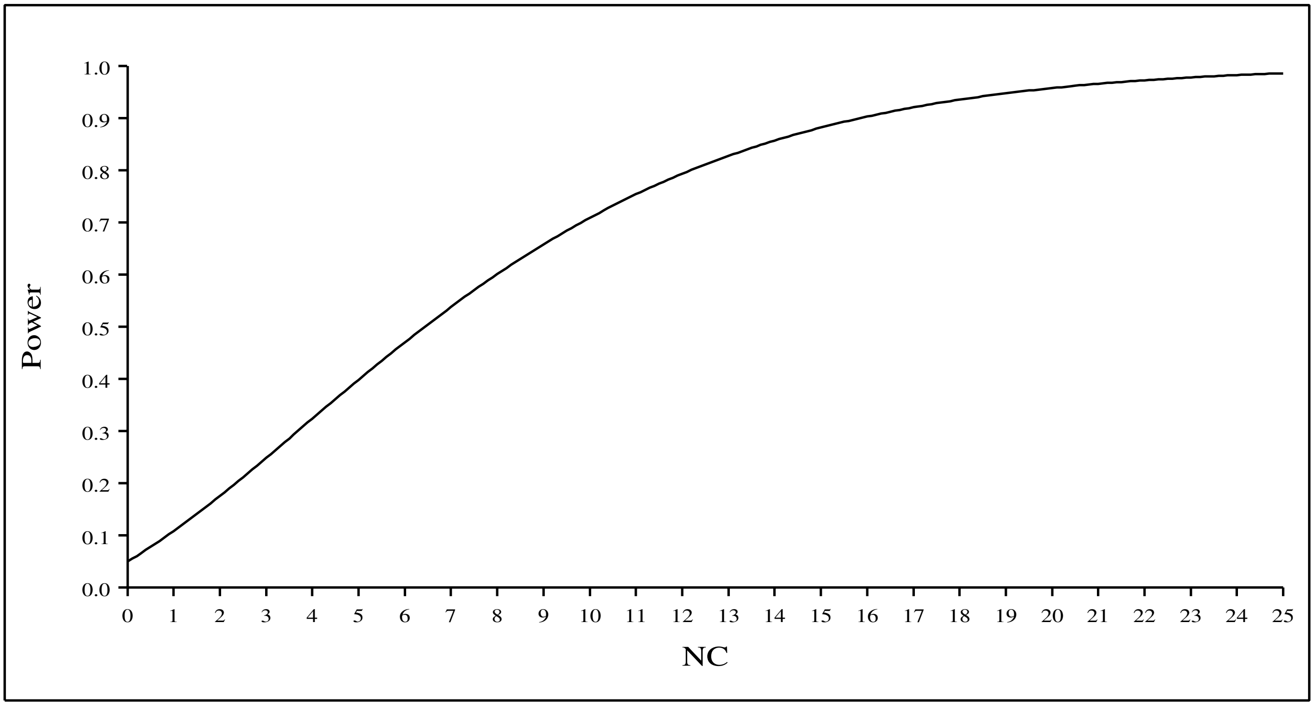
The area to the right of FC is now $1-\beta$, the probability of rejecting H_o given that H_a is true. If the 2 graphs are put together then the effect of nc is quite clear.



Let's expand this graph by looking at non-central F-distributions for a range of nc-values.

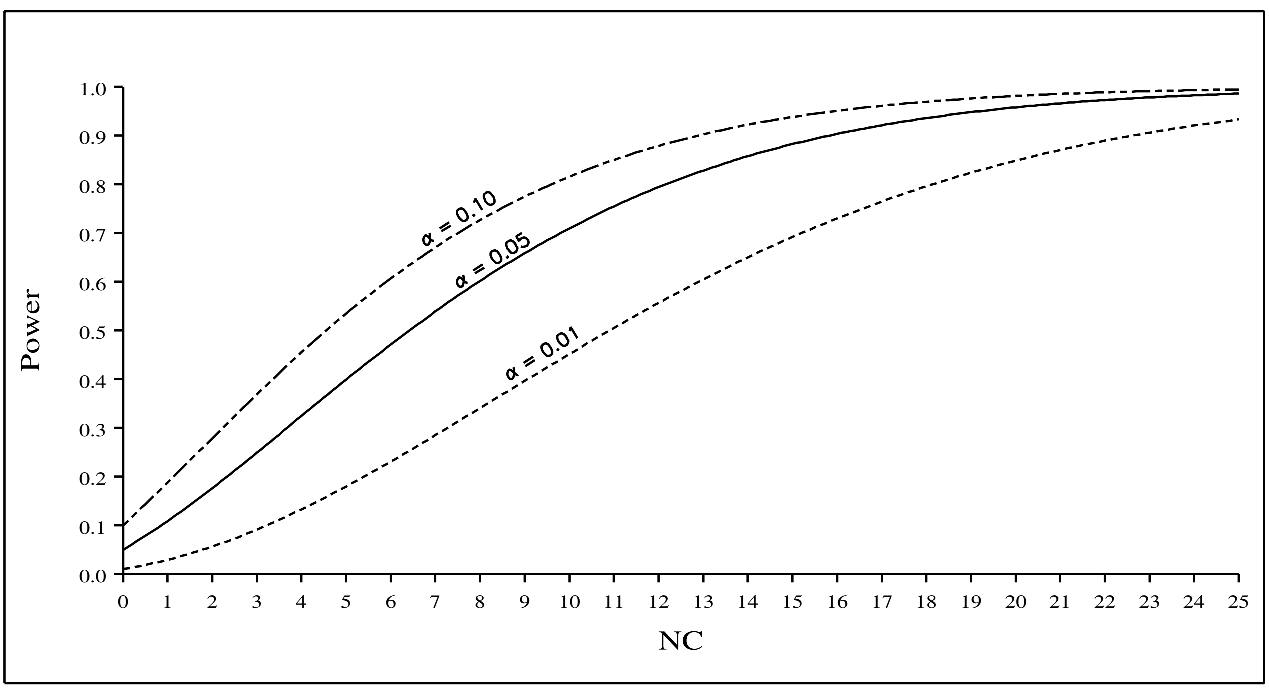


It is evident that the larger nc is, the greater is the power of the experiment. Another representation of the above graph clearly demonstrates this:

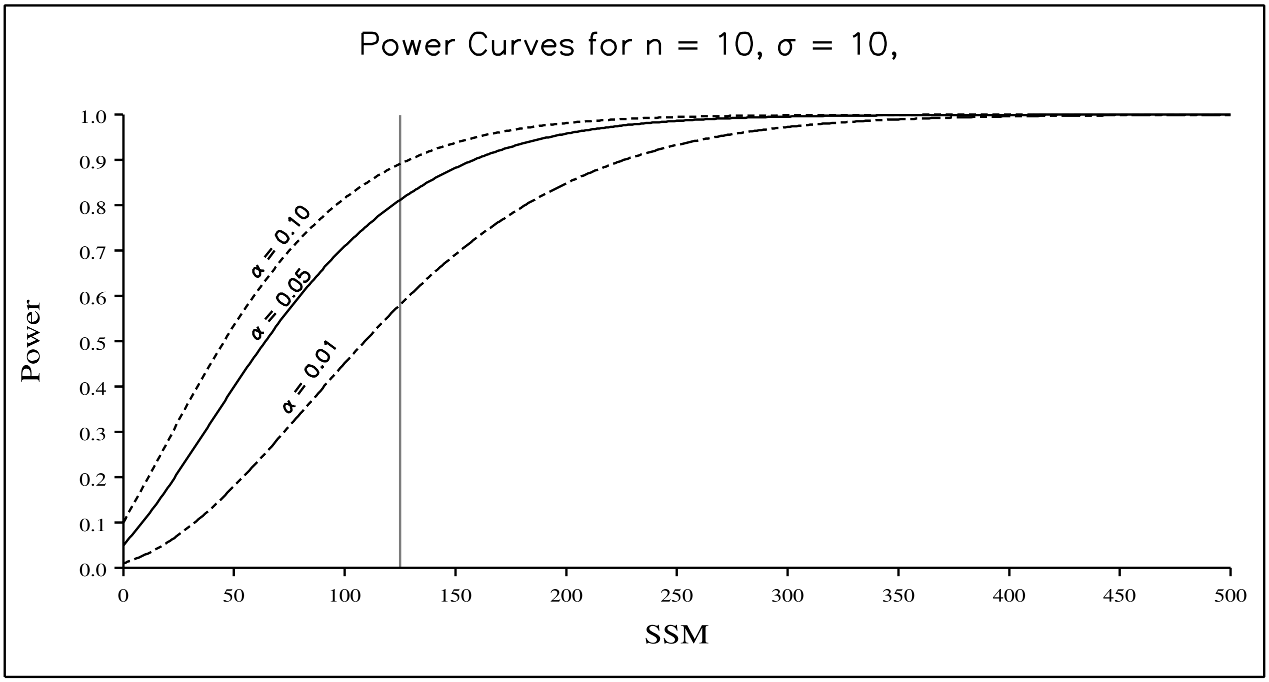


QUESTION: What do you think the effects of n, α and σ will have on power?

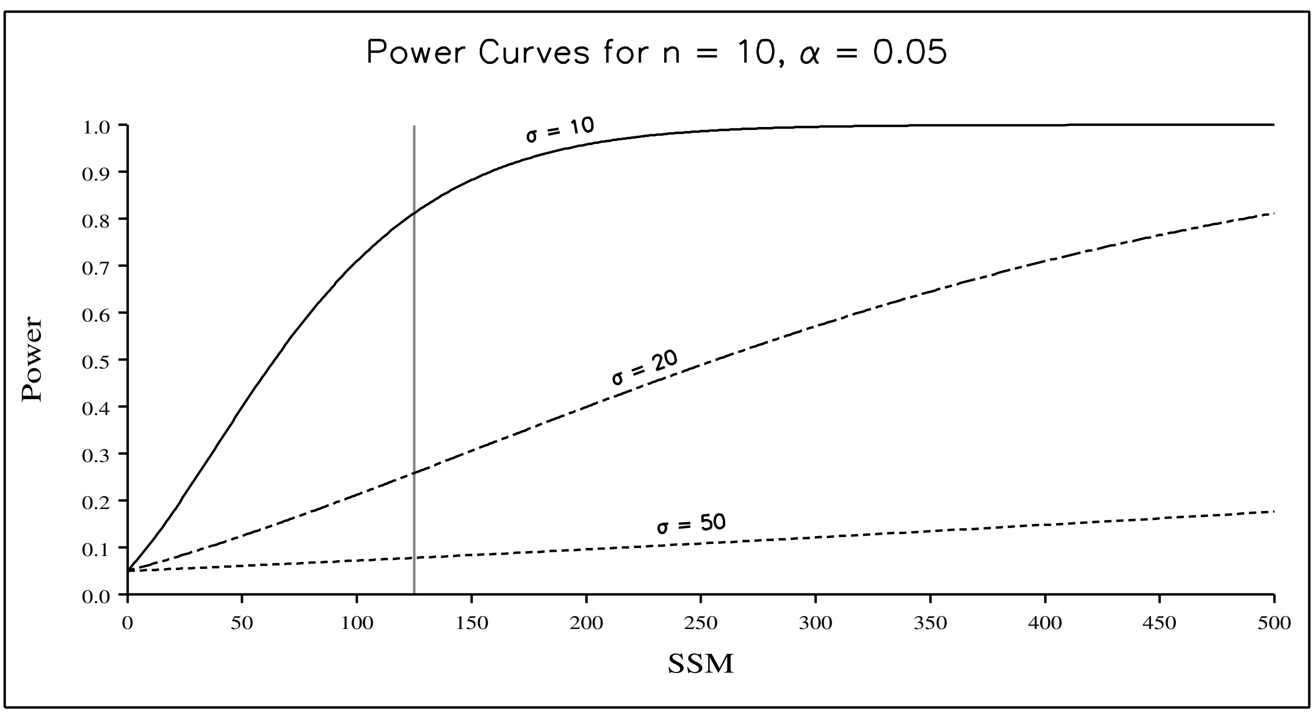
The following graph shows the affect of changing α.



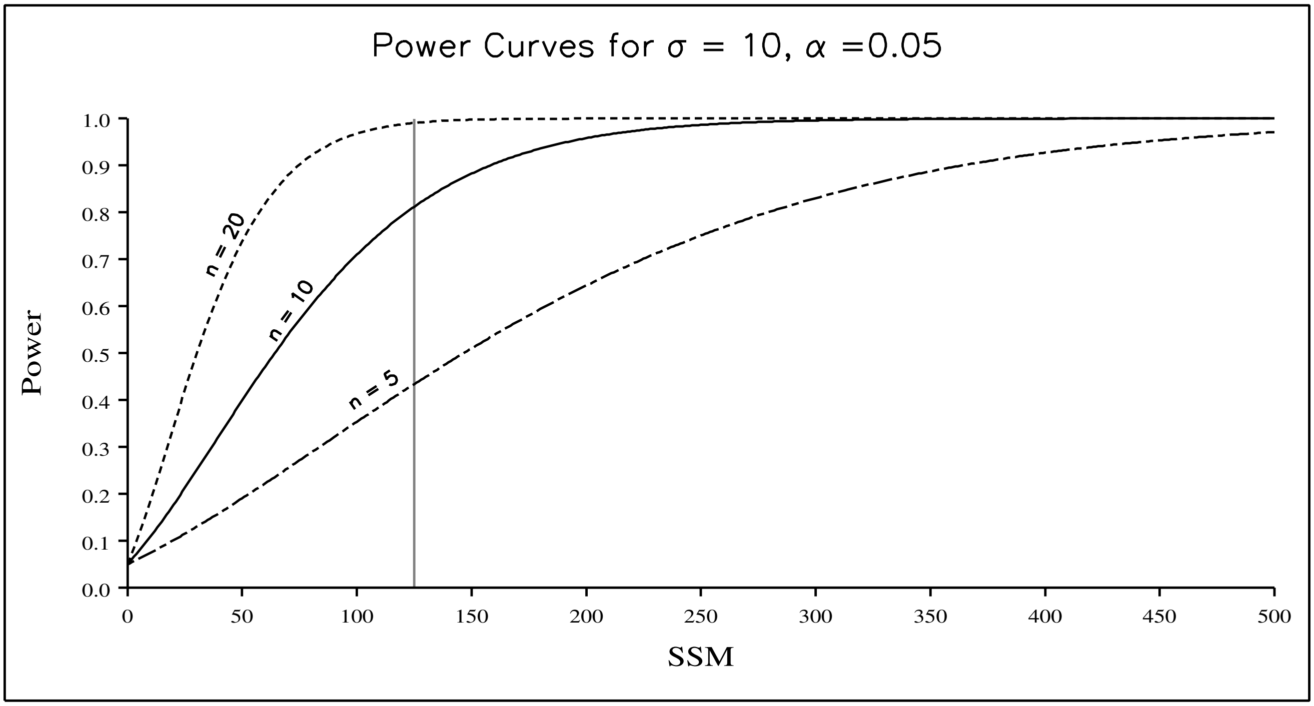
This next graph is the same as the previous except that the x-axis has been changed to SSM. Since n = 10 and σ = 10 for this example, SSM = 10 times the nc of the previous graph.



The following graphs shows the influence of σ on power.



This last graph shows the affect of sample size, n, on power.



QUESTION: Could other values of the expected means have the same nc-value?

We will expand on this topic in a later section.

5 An alternative to a specific list of expected means for H_a

Determining a specific list of expected means for H_a is bound to be difficult under most circumstances. A simpler method relies on the researcher determining the minimum and maximum expected values for the means. To do this, an effect size index, d, also known as the *standardized range of the treatment means* is used. It is calculated by:

$$d = \frac{\mu_{max} - \mu_{min}}{\sigma} .$$

For any specific value of d the range of possible nc values is:

for **a** even[3]: $$\frac{nd^2}{2} \leq nc \leq \frac{nd^2\mathbf{a}}{4}$$

for **a** odd: $$\frac{nd^2}{2} \leq nc \leq \frac{nd^2(\mathbf{a}^2-1)}{4\mathbf{a}}$$

---

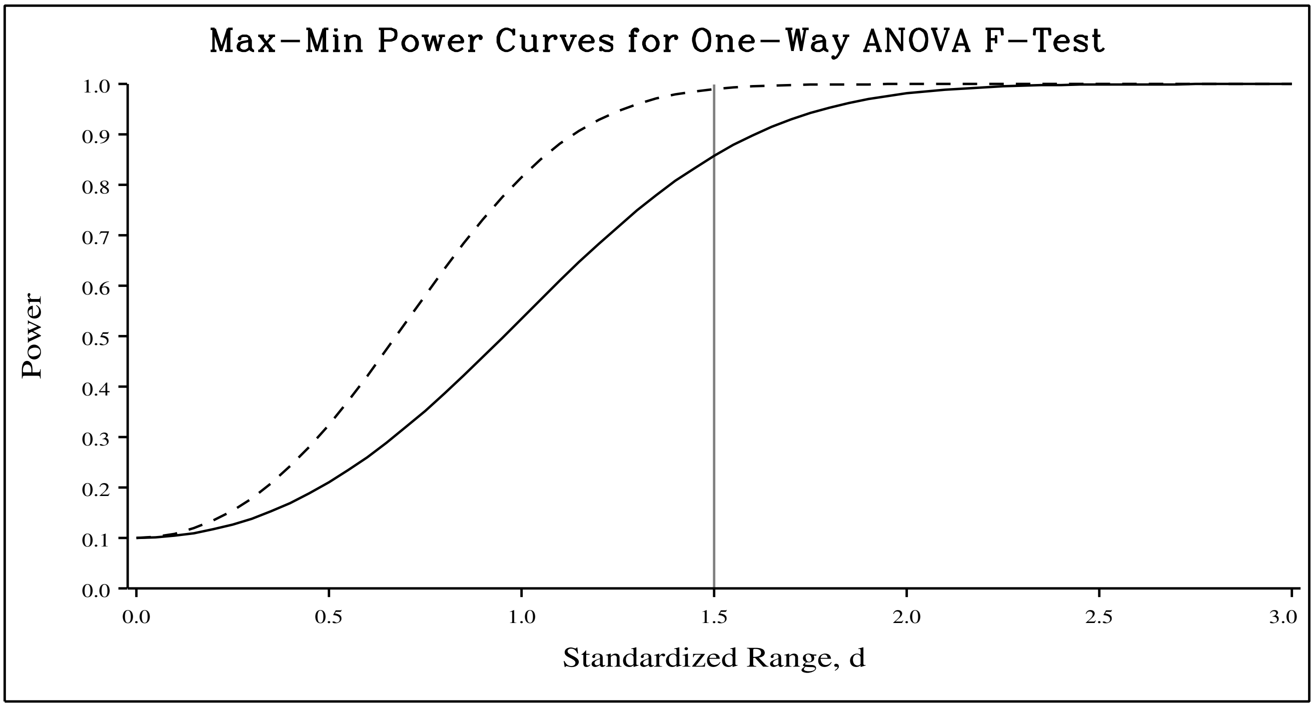
[3]For **a** = 2 (the two-sample t-test), $nc = \frac{nd^2}{2}$, with no range of values.

Power curves can then be determined for the minimum and maximum values of λ or nc given specific values of d, n, α and σ. For our example:

$$d = \frac{25 - 10}{10} = 1.50 \text{ , so that } d^2 = 2.25 \text{ .}$$

Hence $$11.25 \leq nc \leq 22.50 \text{ .}$$

A graph showing the power envelope for this situation is shown below. The vertical line is at d = 1.50.



6 The non-centrality parameter for two groups

When a = 2, that is, when the ANOVA has only 2 groups or treatments then a two-sample t-test can be used to test for differences as the two tests are equivalent[4]. In this case, the effect size index, d, is simply the difference between the two means divided by the standard deviation, σ. This effect size, d, is used directly in the calculation of the non-centrality parameter for the t-distribution[5]. The nc for the F-test no longer has a range of possible values since the minimum is $nd^2/2$ and the maximum is now $nd^2 2/4 = nd^2/2$. In fact, $nd^2/2$ is the square of the non-centrality parameter for the t-test given equal sample sizes for the two means (proof is an exercise for the reader).

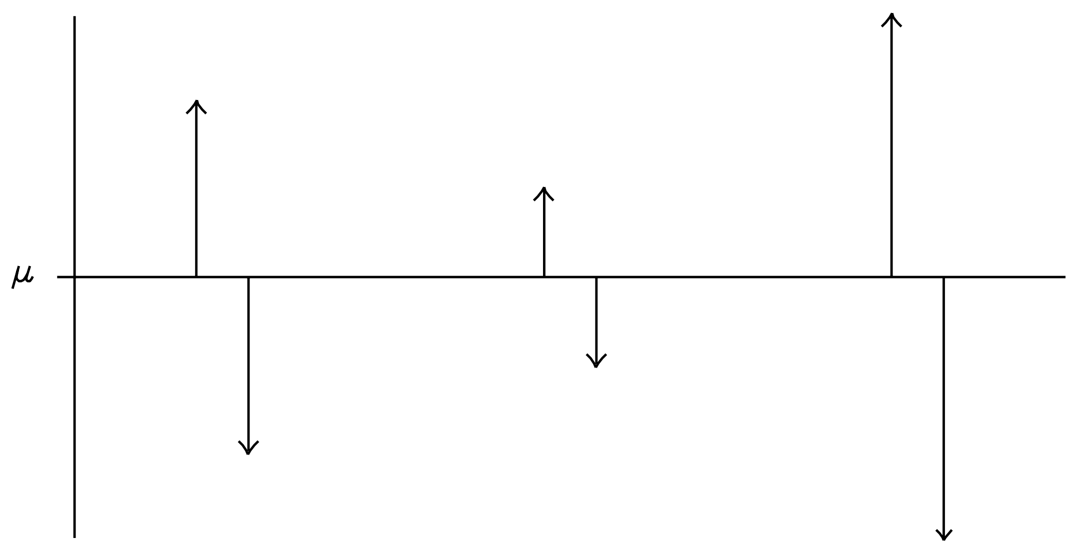
[4] See Biometrics Information Pamphlet # 27: When the t-test and F-test are equivalent.

[5] see page 8 and section 2.3.2 of the Power Analysis Handbook.

When a=2, each value of nc has only one possible pattern for the predicted means. Some calculations will make this clear. In this case, $SSH_a = n \sum(\mu_i - \mu)^2 = n [(\mu_1 - \mu)^2 + (\mu_2 - \mu)^2]$. If the differences between a mean and the grand mean (also known as residuals[6]) are denoted by r_i, then $SSH_a = n\sum(r_i)^2 = n[r_1^2 + r_2^2]$. Since $\mu = (\mu_1 + \mu_2)/2$ then it follows that:

$$r_1 = \mu_1 - (\mu_1 + \mu_2)/2 = (\mu_1 - \mu_2)/2 \quad \text{while} \quad r_2 = \mu_2 - (\mu_1 + \mu_2)/2 = (\mu_2 - \mu_1)/2$$

Hence, it follows that $r_1 = -r_2$ so that each mean is the same distance from the grand mean, but in different directions. Three such pairs of residuals are shown below.



Accordingly, $SSH_a = 2nr_1^2 = 2nr_2^2$ and the noncentrality parameter, $nc = 2nr_1^2/\sigma^2 = 2nr_2^2/\sigma^2$. The difference between the two means is $2r_1 = 2r_2$ and the effect size index is $d = 2r_1/\sigma = 2r_2/\sigma$.

The calculations above demonstrate that there is only one pattern of residuals possible for each nc value when there are just two groups in an experiment. This is not the case when there are three or more groups. Each nc value can be associated with many different patterns of the means. These patterns are more easily discussed using the differences of the means from the grand mean (residuals).

---

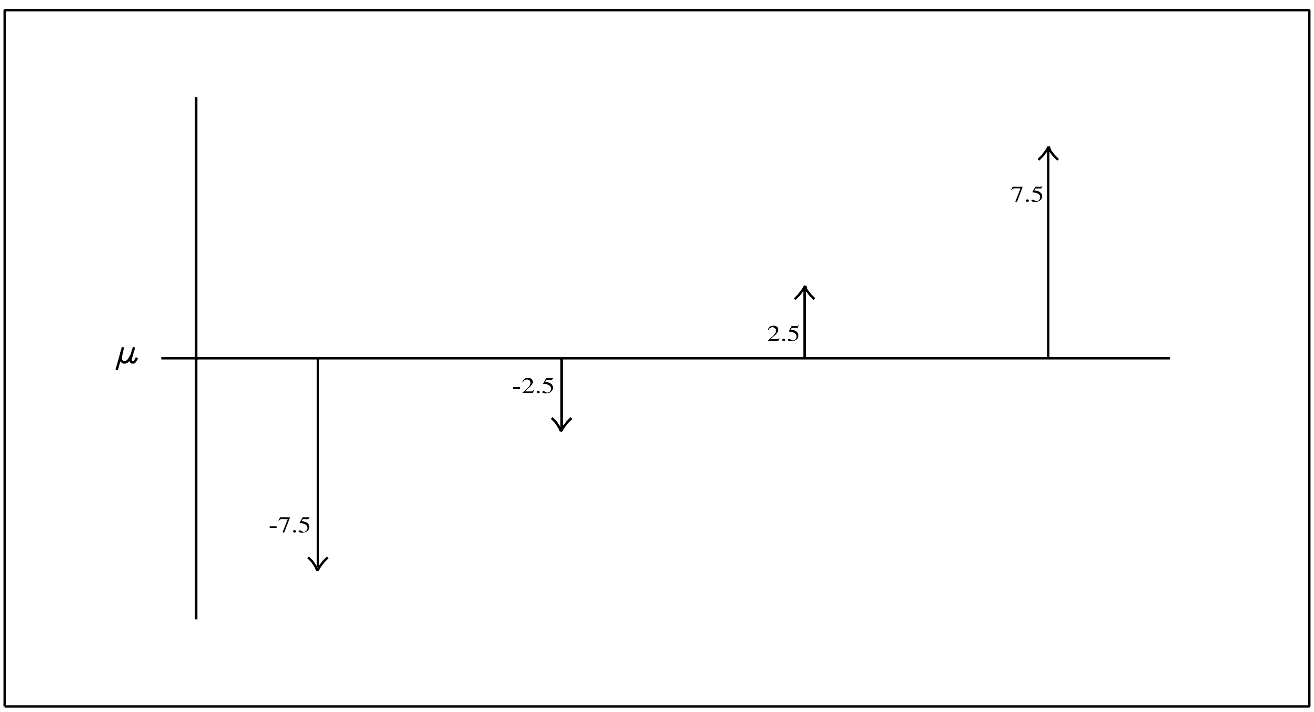
[6] see Biometrics Handbook #1 titled Pictures of Linear Models.

7 The non-centrality parameter for three or more groups

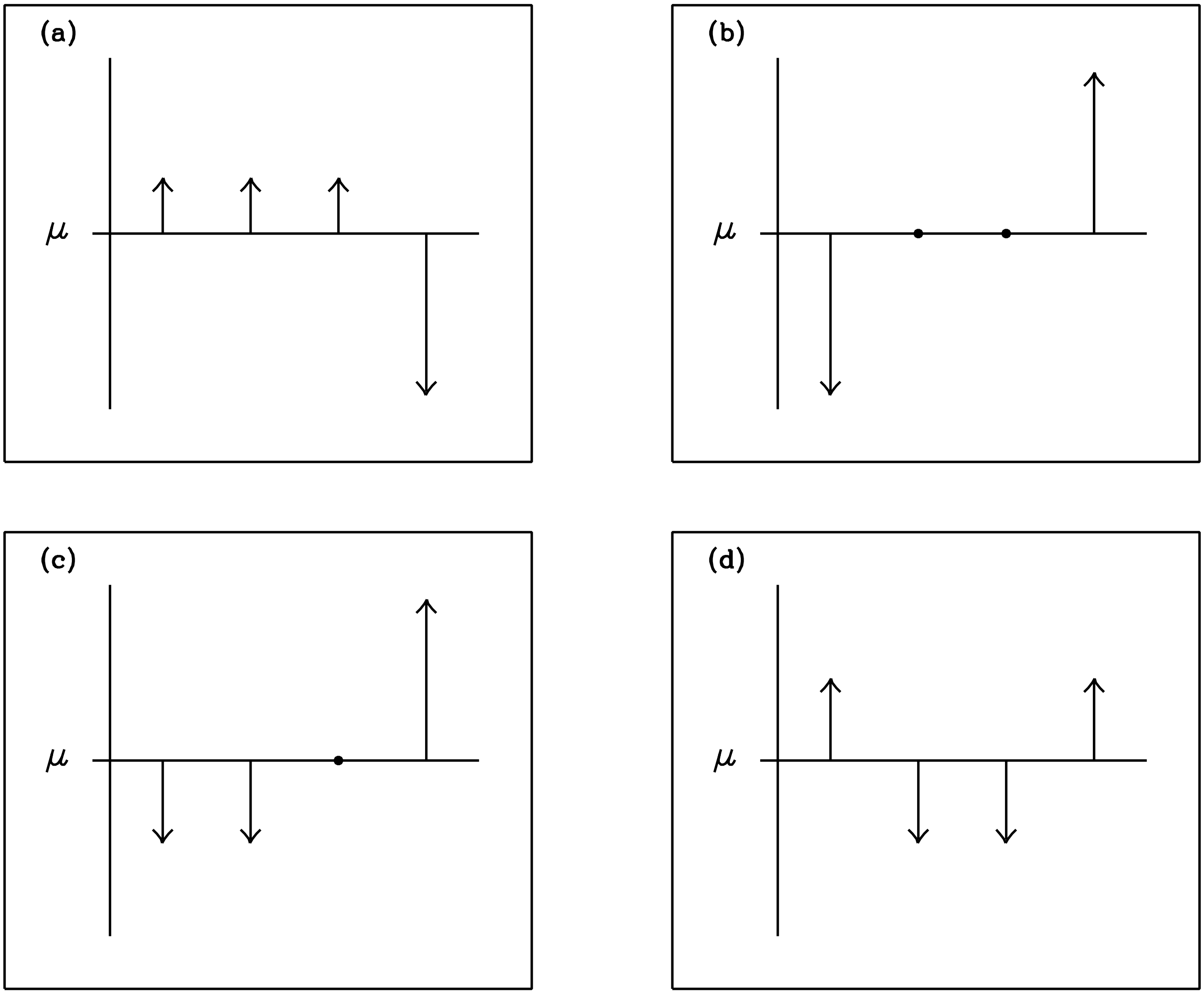
Each nc value can be associated with many different sets of means when there are three or more groups in an ANOVA. If nc, the sample size, n, and the standard deviation, σ, are kept fixed, then all that is left to vary is SSM = $\sum(\mu_i-\mu)^2 = \sum r_i^2$. This term describes the sum of the squared differences or residuals (SSM) between group means and the grand mean. Any specific pattern of group means will have the same SSM regardless of the grand mean. For example, the following sets of means all have an SSM of 125.

Means with SSM = 125:			
	10, 15, 20, 25	15, 20, 10, 25	380, 385, 390, 395
	1550, 1555, 1560, 1565	0, 5, 10, 15	-10, -5, 0, 5

The differences $r_i = (\mu_i - \mu)$ for all these sets are -7.5, -2.5, 2.5 and 7.5 so that the SSM = $(-7.5)^2 + (-2.5)^2 + (2.5)^2 + (7.5)^2 = 125$. Thus, we find that it is the residuals of the group means from the grand mean which contain the salient information about the pattern of means associated with a specific nc value. The above sets of means all have the same values for the residuals and, if the order of the means is irrelevant, have the same pattern. The only difference is the value of the grand mean, μ. The pattern associated with the above sets of means can be represented as:



Other patterns of four means with an SSM = 125 are:



If the first pattern (a) has three arrows of length k and one arrow of length 3k, then $k = 5\sqrt{5}/2\sqrt{3}$. Similarly, the second pattern (b) has two arrows of equal length, k, so that $k = 5\sqrt{5}/\sqrt{2}$. The third pattern (c) has two arrows of length k and one arrow of length 2k so that $k = 5\sqrt{5}/\sqrt{6}$. The fourth pattern has four arrows of equal length so that $k = 5\sqrt{5}/2$.

QUESTION: Can you come up with some other patterns that have the same SSM?

For the calculation of SS_{H_a}, the order of the means does not matter. The order will matter substantially, though, if we are interested in doing contrast tests and associated power or sample size calculations.

8 Power calculations for contrasts

Suppose that there were two contrasts of interest for the simple ANOVA example that we have been using, and that they are:

Group	Expected Means, μ_i	Contrast Coefficients Linear	Control vs others
1	10	-3	-3
2	15	-1	1
3	20	1	1
4	25	3	1

The SSM for the linear contrast is calculated by:

$$SSM_{linear} = \frac{[-3(10) - 1(15) + 1(20) + 3(25)]^2}{-3^2 + -1^2 + 1^2 + 3^2} = \frac{50^2}{20} = 125$$

and for the Control vs others:

$$SSM_{control} = \frac{[-3(10) + 1(15) + 1(20) + 1(25)]^2}{-3^2 + 1^2 + 1^2 + 1^2} = \frac{30^2}{12} = 75.00$$

The non-centrality parameters and corresponding power are then calculated as before for SSH_a = n*SSM.

9 General sample size/design steps:

1) Draw pictures or develop clear descriptions of what would happen if:
   i) there were no treatment effects -- the null hypothesis.
   ii) one or more possible ways for the treatments to have different affects -- the alternate hypotheses. Calculate SSH_a or d^2 for each scenario. These are zero for the null hypothesis.

2) Develop an estimate or a range of possible values for σ or σ^2 based on whatever information is available. (see Biometrics Information #25: ANOVA: The Within Sums of Squares as an Average Variance).

3) Choose acceptable levels of α and β (the Types I and II Error rates).

4) Determine FC for the desired sample size or for a range of sample sizes (using the SAS function FINV).

5) Determine the corresponding power for all combinations of scenarios and values of σ^2 and sample size (using the SAS function PROBF).

6) Decide on an acceptable sample size.

10 Power calculations using SAS

We will now put all the above bits together into complete SAS programs. These programs are useful when planning experiments. The first program uses a list of specific expected means while the second uses the range of expected means.

i) One-way ANOVA: using specific expected means for planning

```
/*  One-way ANOVA with specific expected means */

options linesize=78 pagesize=55 nodate;

data means;
  a = 4;  frq = 10;            *<====  Put in your own numbers for a;
  do atreat = 1 to a;          * (be careful if you change the value for frq);
    input y @@;
    output; end;
cards;
10 15 20 25
;*<====  Put in your own numbers;
;
/**  Obtaining the treatment and contrast SS's (SSH) using PROC GLM **/

proc glm data=means outstat=stat;
  class atreat;  freq frq;
  model y = atreat / ss3;      ***  Put in your own numbers for contrasts;
  contrast 'Control vs others' atreat -3  1  1  1;
  contrast 'Linear response  ' atreat -3 -1  1  3;
title 'One-way ANOVA - Example for text';
run;
proc print data=stat;
title2 'Listing of output data from the outstat=stat of PROC GLM';
run;

/** Calculating the SSM's (i.e. SSH for a sample size of one) **/
/**  and Calculating the POWER for a range of conditions      **/

data power (keep=source dfh dfe n ssm ssh nc fc alpha power);
  set stat; frq=10; retain cells;
  if _type_='ERROR' then cells = df/(frq-1);
  if _type_='SS3' or _type_='CONTRAST' then do;
      dfh = df; ssm = ss/frq; source = _source_;
 /*****************************************/
 /*****  Put your own values in here *****/
  do alpha = 0.01, 0.05, 0.10;
    do n = 2 to 30 by 2;
       do sigma = 10;
 /*****************************************/
```

```
       dfe = (n-1)*cells; ssh = n*ssm;
       nc = ssh/(sigma**2);
       fc = finv(1-alpha, dfh, dfe, 0);
      power = 1-probf(fc, dfh, dfe, nc);
      if power = . then power = 1.0;
      output;
  end; end; end; end;
run;

/** Printing and plotting out the results **/

proc print noobs; by source alpha notsorted;
  where alpha = 0.05;          *<=== Used to reduce output;
  var ssm ssh dfh dfe n fc nc power;
title2 'Results of Power Analysis';  run;
proc plot; by source notsorted;
  plot power*n = alpha / vref = 0.80, 0.90;
run;
proc sort data=power; by alpha n;  run;
proc plot; by alpha;
  where alpha = 0.05;          *<=== Used to reduce output;
  plot power*n = source / vref = 0.80, 0.90;
run;
```

The output from this program is:

```
                     One-way ANOVA - Example for text                      1

                    General Linear Models Procedure
                        Class Level Information

                      Class     Levels    Values

                      ATREAT         4    1 2 3 4


               Number of observations in data set = 40


                    One-way ANOVA - Example for text                      2

                    General Linear Models Procedure

Dependent Variable: Y
Frequency:          FRQ
                                    Sum of            Mean
Source                  DF         Squares          Square   F Value     Pr > F

Model                    3     1250.000000      416.666667  99999.99     0.0

Error                   36        0.000000        0.000000

Corrected Total         39     1250.000000

                  R-Square            C.V.        Root MSE               Y Mean

                  1.000000               0               0           17.5000000


Source                  DF     Type III SS     Mean Square   F Value     Pr > F

ATREAT                   3     1250.000000      416.666667  99999.99     0.0

Contrast                DF     Contrast SS     Mean Square   F Value     Pr > F

Control vs others        1      750.000000      750.000000  99999.99     0.0
Linear response          1     1250.000000     1250.000000  99999.99     0.0


                     One-way ANOVA - Example for text                      3
            Listing of output data from the outstat=stat of PROC GLM

  OBS   _NAME_   _SOURCE_             _TYPE_      DF     SS        F      PROB

    1     Y      ERROR                ERROR       36      0        .        .
    2     Y      ATREAT               SS3          3   1250   99999.99      0
    3     Y      Control vs others    CONTRAST     1    750   99999.99      0
    4     Y      Linear response      CONTRAST     1   1250   99999.99      0
```

Results of Power Analysis

------------------------- SOURCE=ATREAT ALPHA=0.05 -------------------------

SSM	SSH	DFH	DFE	N	FC	NC	POWER
125	250	3	4	2	6.59138	2.5	0.12266
125	500	3	12	4	3.49029	5.0	0.32934
125	750	3	20	6	3.09839	7.5	0.53147
125	1000	3	28	8	2.94669	10.0	0.69474
125	1250	3	36	10	2.86627	12.5	0.81196
125	1500	3	44	12	2.81647	15.0	0.88939
125	1750	3	52	14	2.78260	17.5	0.93742
125	2000	3	60	16	2.75808	20.0	0.96574
125	2250	3	68	18	2.73950	22.5	0.98178
125	2500	3	76	20	2.72494	25.0	0.99055
125	2750	3	84	22	2.71323	27.5	0.99521
125	3000	3	92	24	2.70359	30.0	0.99762
125	3250	3	100	26	2.69553	32.5	0.99884
125	3500	3	108	28	2.68869	35.0	0.99944
125	3750	3	116	30	2.68281	37.5	0.99974

-------------------- SOURCE=Control vs others ALPHA=0.05 ---------------------

SSM	SSH	DFH	DFE	N	FC	NC	POWER
75	150	1	4	2	7.70865	1.5	0.15879
75	300	1	12	4	4.74723	3.0	0.35742
75	450	1	20	6	4.35124	4.5	0.52357
75	600	1	28	8	4.19597	6.0	0.65723
75	750	1	36	10	4.11317	7.5	0.75958
75	900	1	44	12	4.06171	9.0	0.83496
75	1050	1	52	14	4.02663	10.5	0.88877
75	1200	1	60	16	4.00119	12.0	0.92622
75	1350	1	68	18	3.98190	13.5	0.95173
75	1500	1	76	20	3.96676	15.0	0.96881
75	1650	1	84	22	3.95457	16.5	0.98007
75	1800	1	92	24	3.94454	18.0	0.98739
75	1950	1	100	26	3.93614	19.5	0.99209
75	2100	1	108	28	3.92901	21.0	0.99508
75	2250	1	116	30	3.92288	22.5	0.99696

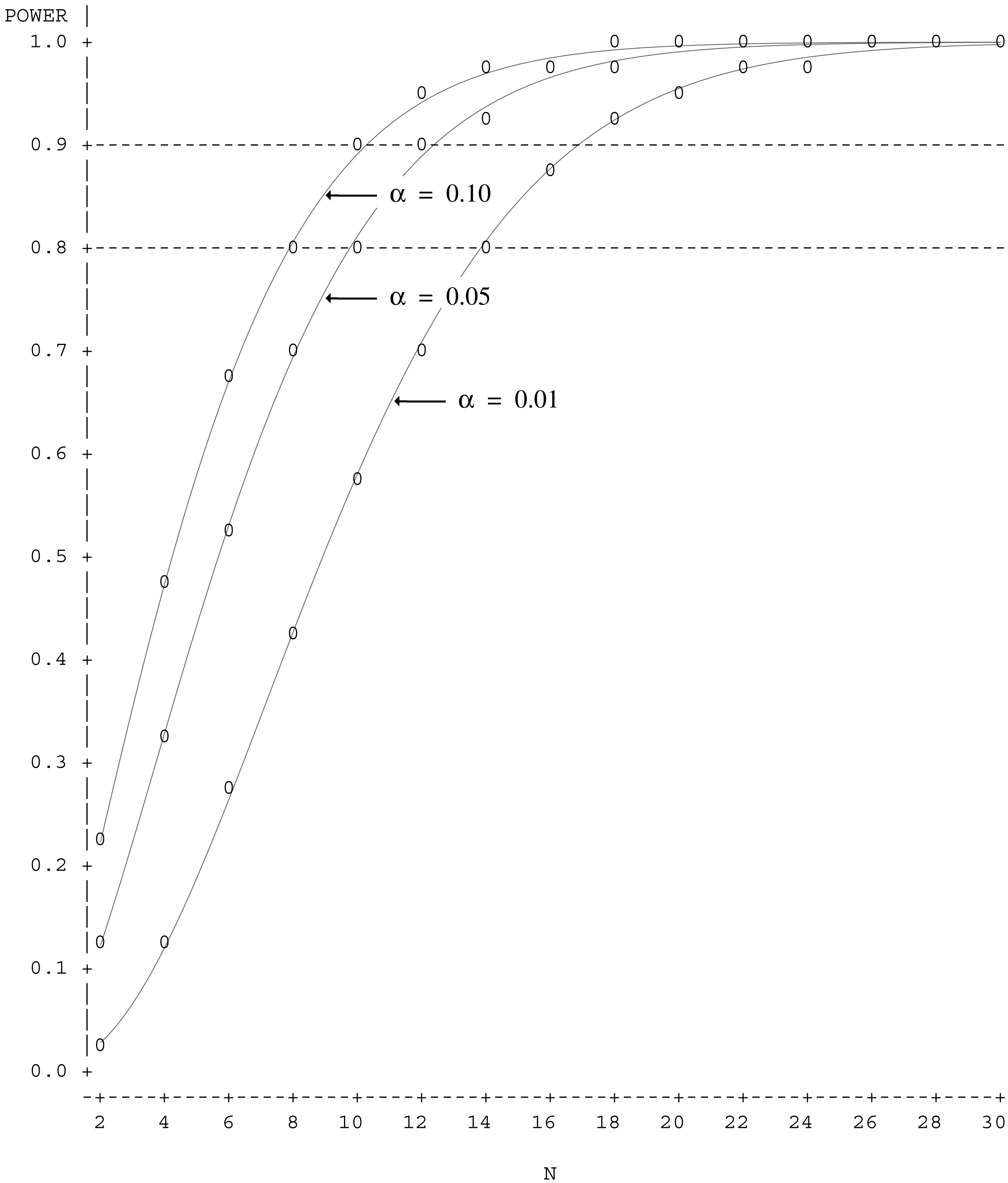
Results of Power Analysis

--------------------- SOURCE=Linear response ALPHA=0.05 ----------------------

SSM	SSH	DFH	DFE	N	FC	NC	POWER
125	250	1	4	2	7.70865	2.5	0.23134
125	500	1	12	4	4.74723	5.0	0.53819
125	750	1	20	6	4.35124	7.5	0.74063
125	1000	1	28	8	4.19597	10.0	0.86262
125	1250	1	36	10	4.11317	12.5	0.93053
125	1500	1	44	12	4.06171	15.0	0.96616
125	1750	1	52	14	4.02663	17.5	0.98402
125	2000	1	60	16	4.00119	20.0	0.99265
125	2250	1	68	18	3.98190	22.5	0.99669
125	2500	1	76	20	3.96676	25.0	0.99854
125	2750	1	84	22	3.95457	27.5	0.99937
125	3000	1	92	24	3.94454	30.0	0.99973
125	3250	1	100	26	3.93614	32.5	0.99989
125	3500	1	108	28	3.92901	35.0	0.99995
125	3750	1	116	30	3.92288	37.5	0.99998

------------------------------ SOURCE=ATREAT ------------------------------

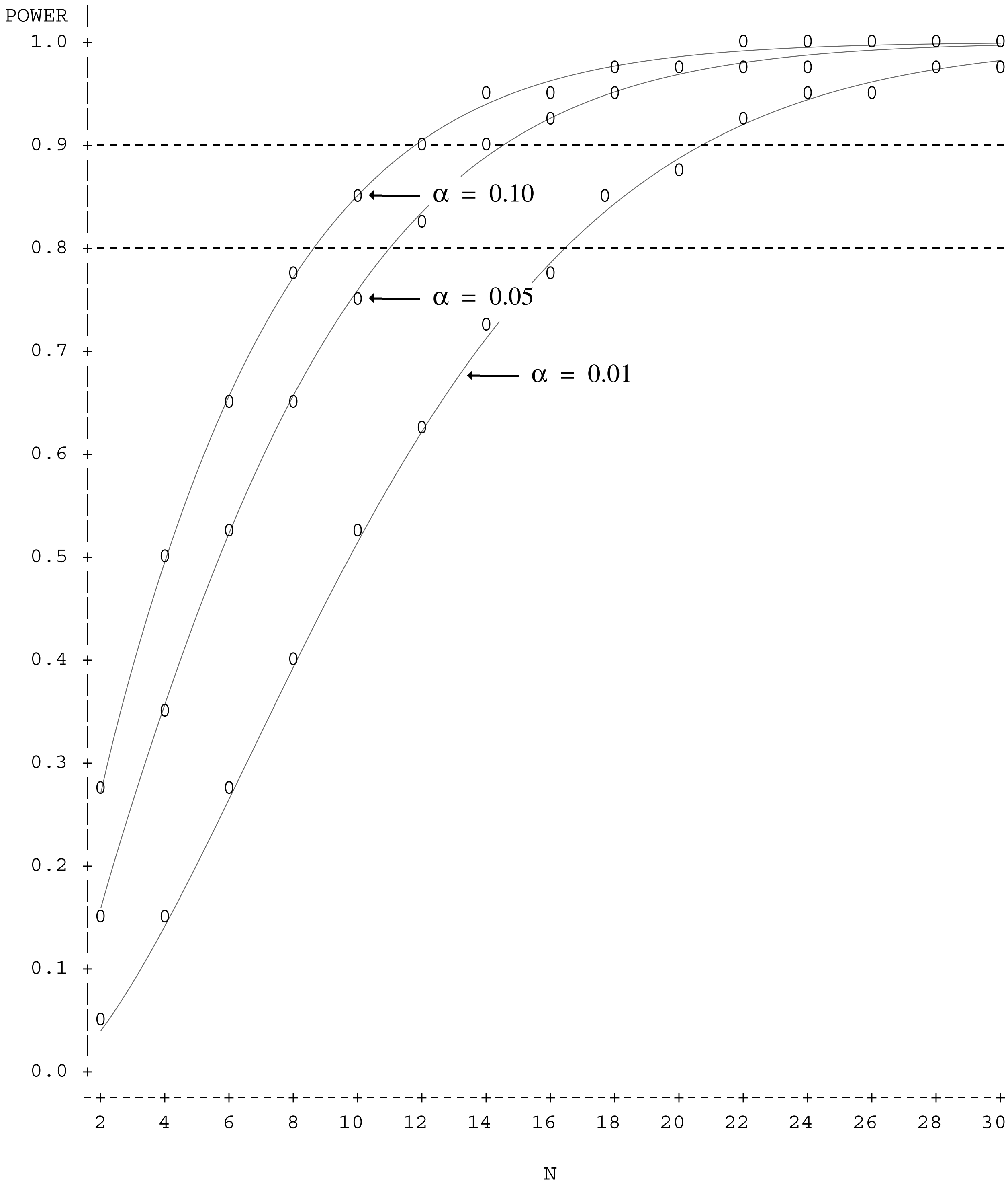
Plot of POWER*N. Symbol is value of ALPHA.



NOTE: 10 obs hidden.

------------------------- SOURCE=Control vs others -------------------------

Plot of POWER*N. Symbol is value of ALPHA.

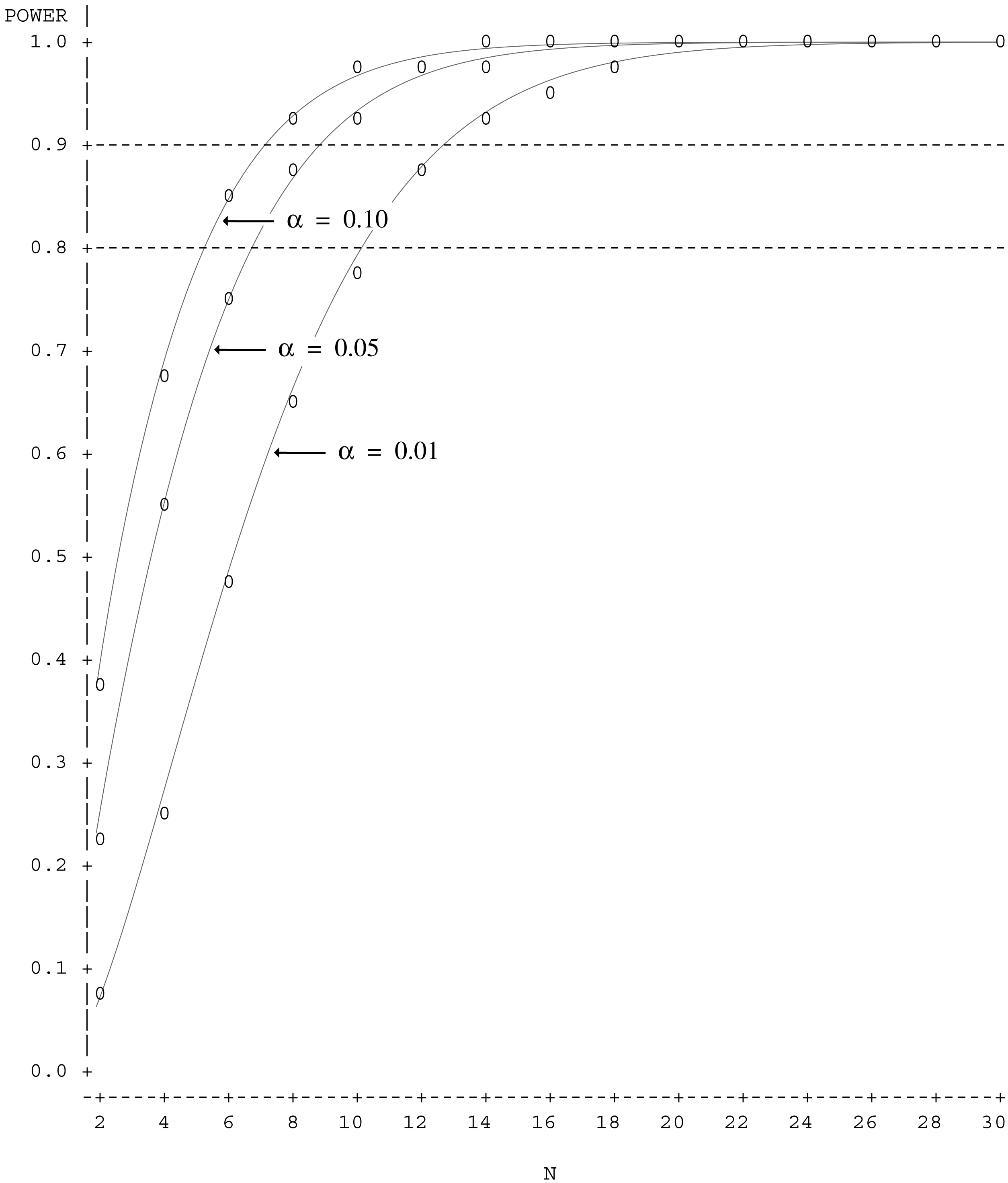


NOTE: 4 obs hidden.

Results of Power Analysis

-------------------------- SOURCE=Linear response --------------------------

Plot of POWER*N. Symbol is value of ALPHA.

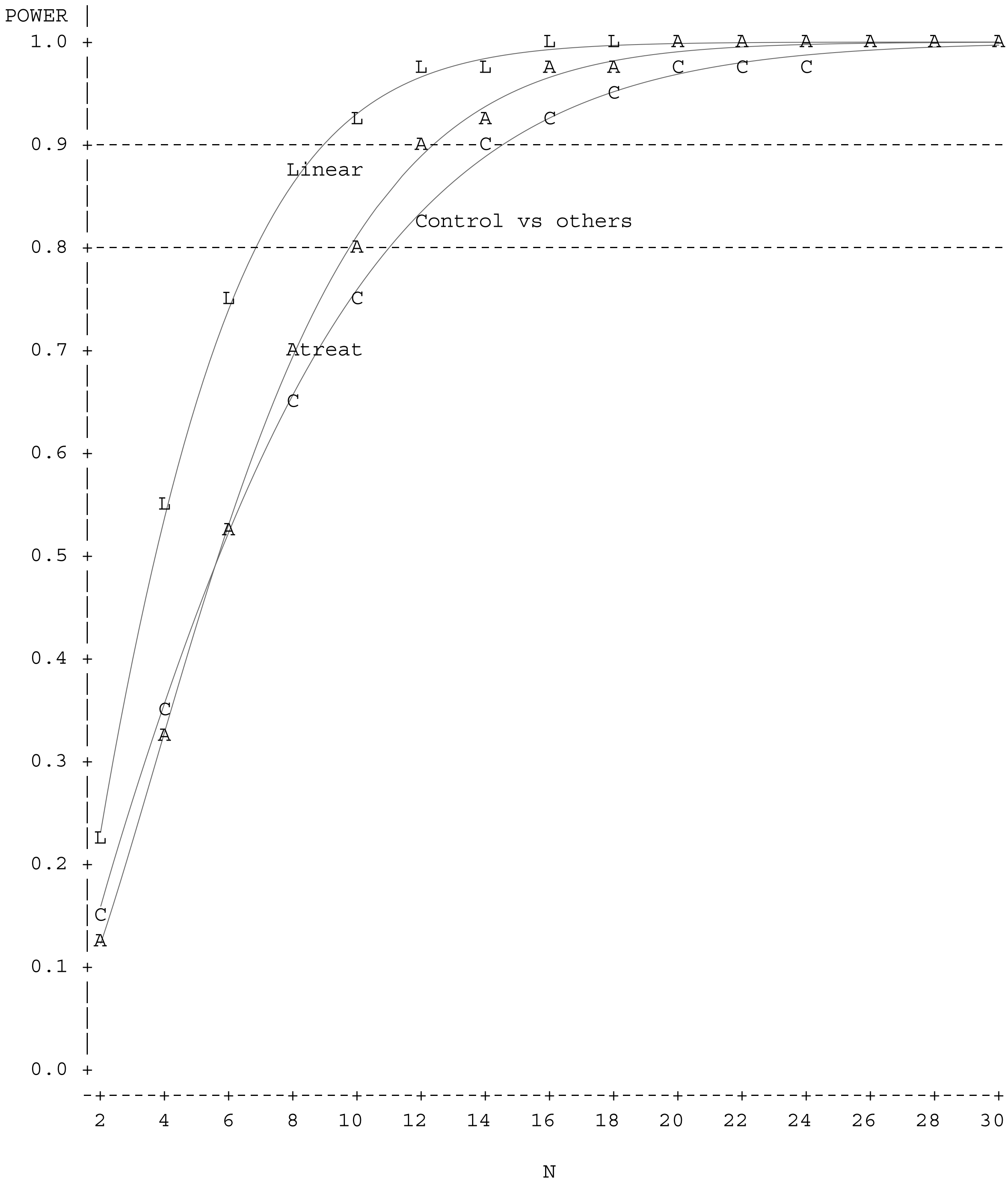


NOTE: 15 obs hidden.

Results of Power Analysis

-------------------------------- ALPHA=0.05 --------------------------------

Plot of POWER*N. Symbol is value of SOURCE.



NOTE: 10 obs hidden.

ii) One-way ANOVA: using the effect size for planning

This program uses the minimum and maximum expected values for the means in the effect size index. The program is:

```
/* powrng.sas */

options linesize=78 pagesize=55 nodate;

/** Calculating minimum and maximum power for a range of conditions **/

data powrng;
 /****************************************/
 /*****  Put in your own values here *****/
  a = 4;  dfa = a - 1; max = 25; min = 10;
  do alpha = 0.01, 0.05, 0.10;
   do sigma = 10, 20, 50;
    do n = 2 to 30 by 2;
 /****************************************/
      dfe = a*(n-1);
      fc = finv(1-alpha, dfa, dfe, 0);
      d = (max - min) / sigma;
      nd2 = n*d*d;  r = mod(a,2);
      ************; ncmin = nd2/2;
      if r = 0 then ncmax = nd2*a/4;
               else ncmax = nd2*(a*a-1)/(4*a);
      powmin = 1-probf(fc, dfa, dfe, ncmin);
      powmax = 1-probf(fc, dfa, dfe, ncmax);
      if powmin = . then powmin = 1.0;
      if powmax = . then powmax = 1.0;
      output;
  end; end; end;
keep alpha a n dfe fc sigma d powmin powmax;

/** Printing and plotting out the results **/

proc print noobs; by alpha sigma;
  where alpha = 0.05;          *<==== Used to reduce output;
  var a n dfe fc d powmin powmax;
title 'Power envelope for One-way ANOVA with a levels';
run;

proc  plot; by alpha;
  plot powmax*n='+' powmin*n='*'
    / overlay vaxis = 0 to 1.0 by 0.1 vref=0.8, 0.9;
run;
```

The output from this program starts on the next page:

---------------------------- ALPHA=0.05 SIGMA=10 -----------------------------

A	N	FC	D	POWMIN	POWMAX
4	2	6.59138	1.5	0.11499	0.18582
4	4	3.49029	1.5	0.29895	0.55643
4	6	3.09839	1.5	0.48503	0.80722
4	8	2.94669	1.5	0.64367	0.92802
4	10	2.86627	1.5	0.76518	0.97596
4	12	2.81647	1.5	0.85143	0.99263
4	14	2.78260	1.5	0.90918	0.99789
4	16	2.75808	1.5	0.94611	0.99943
4	18	2.73950	1.5	0.96884	0.99985
4	20	2.72494	1.5	0.98239	0.99996
4	22	2.71323	1.5	0.99024	1.00000
4	24	2.70359	1.5	0.99469	1.00000
4	26	2.69553	1.5	0.99716	1.00000
4	28	2.68869	1.5	0.99850	1.00000
4	30	2.68281	1.5	0.99922	1.00000

---------------------------- ALPHA=0.05 SIGMA=20 -----------------------------

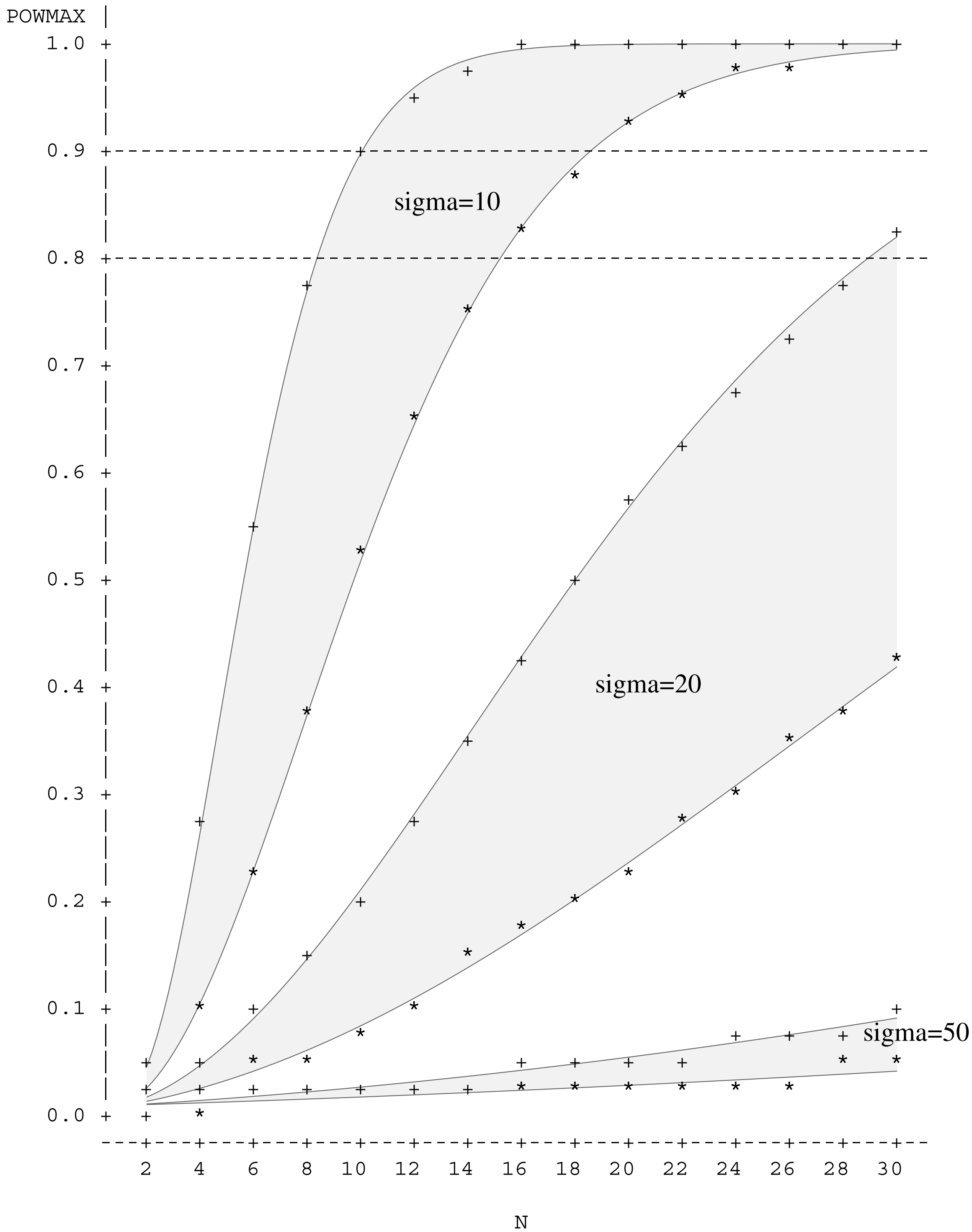
A	N	FC	D	POWMIN	POWMAX
4	2	6.59138	0.75	0.06543	0.08147
4	4	3.49029	0.75	0.10358	0.16485
4	6	3.09839	0.75	0.14513	0.25713
4	8	2.94669	0.75	0.18930	0.35202
4	10	2.86627	0.75	0.23529	0.44460
4	12	2.81647	0.75	0.28231	0.53136
4	14	2.78260	0.75	0.32964	0.61011
4	16	2.75808	0.75	0.37668	0.67975
4	18	2.73950	0.75	0.42287	0.73999
4	20	2.72494	0.75	0.46780	0.79112
4	22	2.71323	0.75	0.51110	0.83383
4	24	2.70359	0.75	0.55251	0.86898
4	26	2.69553	0.75	0.59181	0.89755
4	28	2.68869	0.75	0.62889	0.92050
4	30	2.68281	0.75	0.66364	0.93875

---------------------------- ALPHA=0.05 SIGMA=50 -----------------------------

A	N	FC	D	POWMIN	POWMAX
4	2	6.59138	0.3	0.05242	0.05487
4	4	3.49029	0.3	0.05793	0.06612
4	6	3.09839	0.3	0.06349	0.07772
4	8	2.94669	0.3	0.06915	0.08973
4	10	2.86627	0.3	0.07492	0.10214
4	12	2.81647	0.3	0.08080	0.11493
4	14	2.78260	0.3	0.08678	0.12807
4	16	2.75808	0.3	0.09288	0.14152
4	18	2.73950	0.3	0.09907	0.15526
4	20	2.72494	0.3	0.10536	0.16927
4	22	2.71323	0.3	0.11175	0.18350
4	24	2.70359	0.3	0.11822	0.19794
4	26	2.69553	0.3	0.12478	0.21255
4	28	2.68869	0.3	0.13143	0.22731
4	30	2.68281	0.3	0.13815	0.24219

------------------------------- ALPHA=0.01 -------------------------------

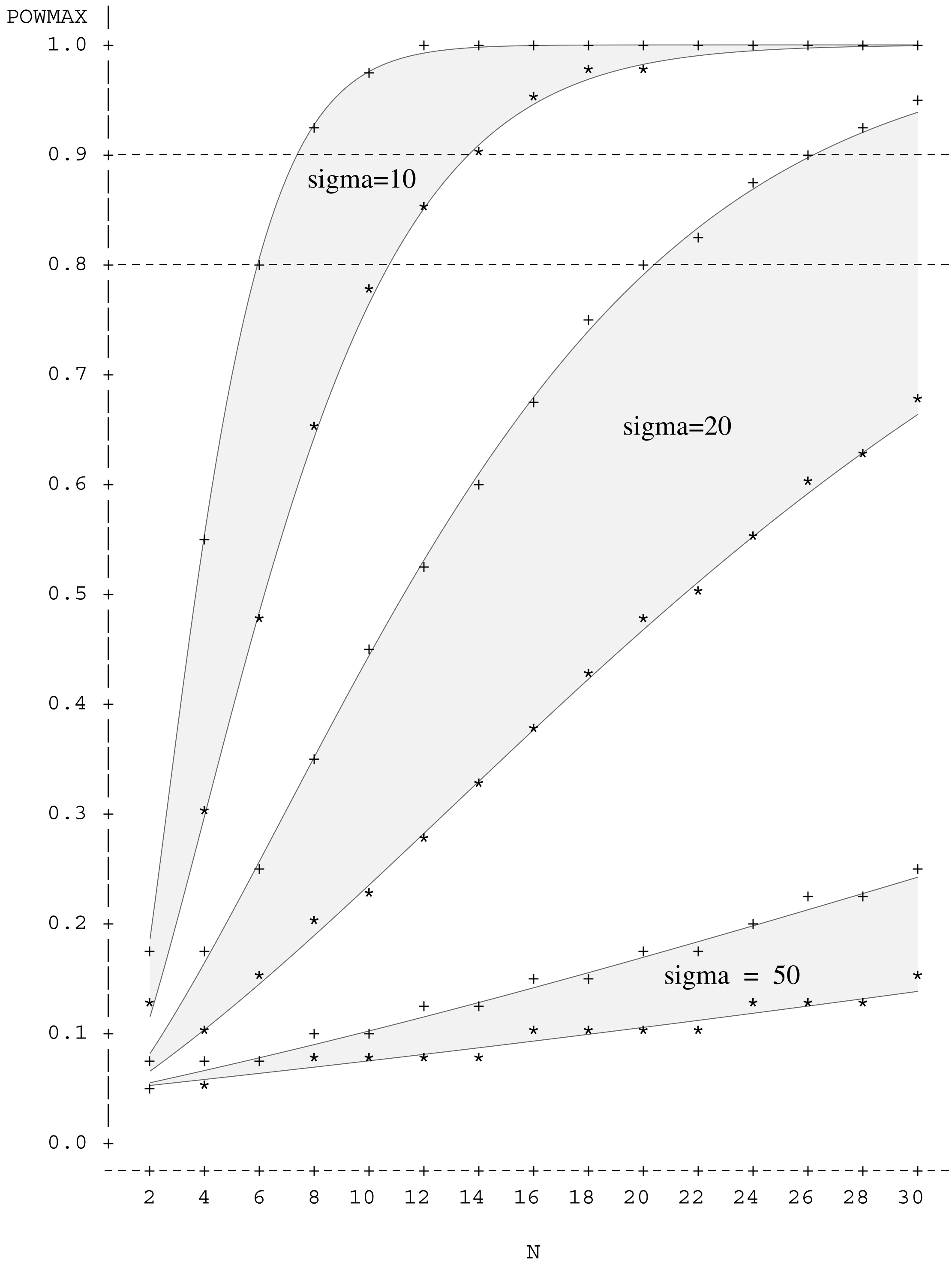
Plot of POWMAX*N. Symbol used is '+'.
Plot of POWMIN*N. Symbol used is '*'.



NOTE: 11 obs hidden.

------------------------------ ALPHA=0.05 ------------------------------

Plot of POWMAX*N. Symbol used is '+'.
Plot of POWMIN*N. Symbol used is '*'.

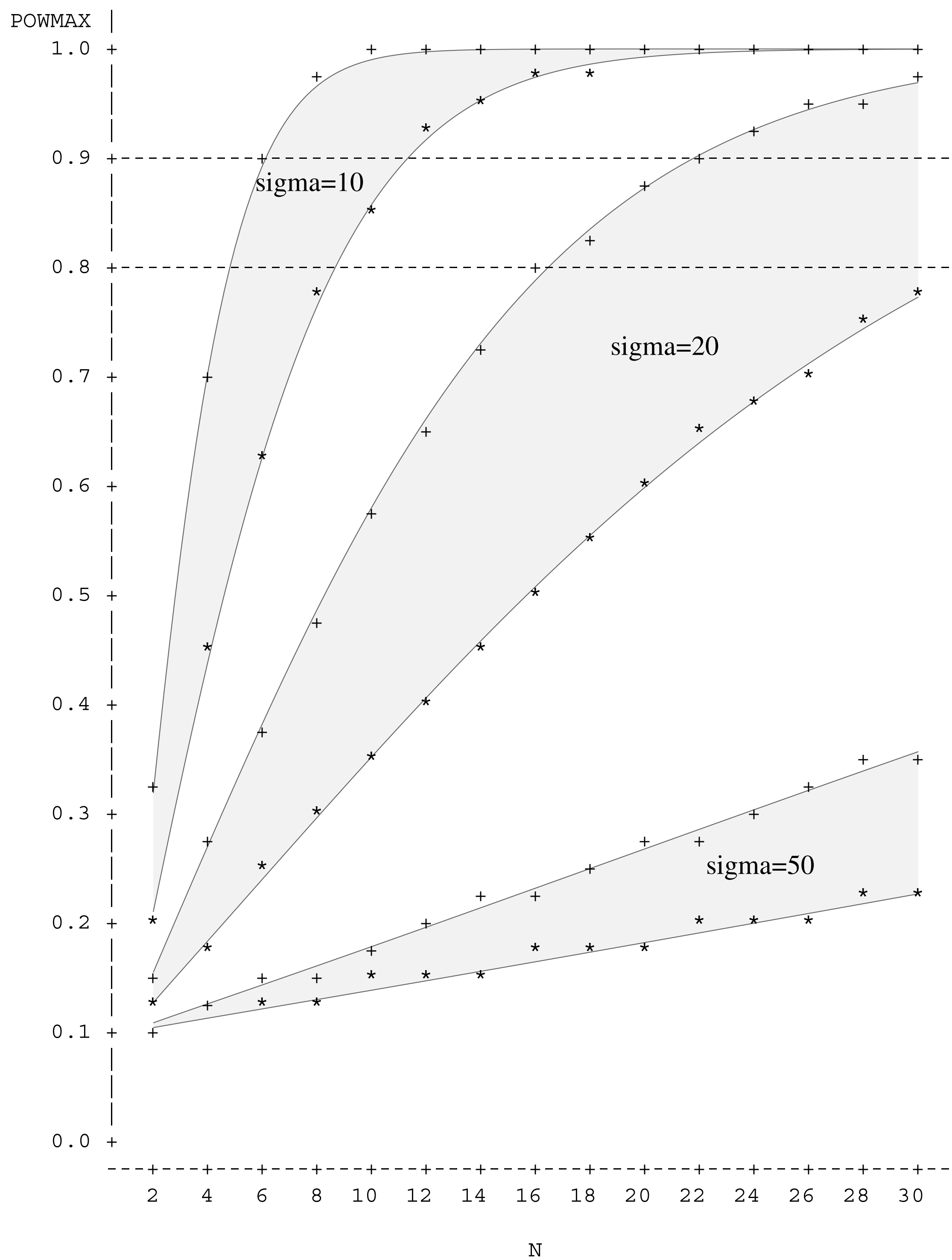


NOTE: 9 obs hidden.

Power envelope for One-way ANOVA with a levels 5

------------------------------- ALPHA=0.1 -------------------------------

Plot of POWMAX*N. Symbol used is '+'.
Plot of POWMIN*N. Symbol used is '*'.



NOTE: 9 obs hidden.

iii) Adjusting methods for the Randomized Block Design

Power analysis methods for the randomized block design are easily modified from the previous descriptions by simply using an estimate of the Mean Square for the Block x Treatment interaction as σ, instead of the MSE.

iv) Adjusting methods for two-way ANOVA: using specific expected means for planning

Calculate power for the interaction test first, then do the main effects.

A SAS program to determine appropriate sample sizes is:

```
/*  Two-way ANOVA with specific expected means */

/* First do analysis with an interaction */
/* Example taken from Handbook #2 on Power Analysis */

options linesize=78 pagesize=55 nodate;

/**  Read in the expected means **/

data case1;
  a = 2;  b = 6; frq = 10;   *<==== Put in your own numbers for a and b;
  do species = 1 to a;       *  (be careful if you change the value for frq);
    do rate  = 0 to b-1;
    input y @@;
    output; end; end;
cards;
0.25 0.40  0.50 0.55  0.55 0.50
0.35 0.505 0.62 0.695 0.73 0.725
;
/**  Obtaining the treatment and contrast SS's (SSH's) using PROC GLM **/

proc glm data=case1 outstat=stat;
  class species rate;  freq frq;
     model y = species|rate / ss3;
     contrast 'Linear response'    rate -5 -3 -1  1  3  5;
     contrast 'Quadratic response' rate  5 -1 -4 -4 -1  5;
   title 'Two-way ANOVA Example -- with an interaction';
   run;
proc print data=stat noobs;
title2 'List of Sources and SS''s from PROC GLM';  run;
```

```
/** Calculating the SSM's and the POWER for each source **/
/**           for a variety of conditions              **/

data power1 (keep=source dfh dfe n ssm ssh nc fc alpha power);
  set stat; frq=10; retain cells;
  if _type_='ERROR' then cells = df/(frq-1);
  if _type_='SS3' or _type_='CONTRAST' then do;
    dfh = df; ssm = ss/frq; source = _source_;
 /*****************************************/
 /*****  Put your own values in here *****/
  do alpha = 0.01, 0.05, 0.10;
    do n = 2 to 30 by 2;
       do sigma = 0.4;
/*****************************************/
         dfe = (n-1)*cells; ssh = n*ssm;
         nc = ssh/(sigma**2);
         fc = finv(1-alpha, dfh, dfe, 0);
        power = 1-probf(fc, dfh, dfe, nc);
        if power = . then power = 1.0;
        output;
    end; end; end; end;
  run;

/** Printing and plotting out the results  **/

proc print noobs; by source alpha notsorted;
  where alpha = 0.05;
  var ssh dfh dfe n fc nc power;
title2 'Results of Power Analysis';
run;
proc sort data=power1; by alpha n;
run;
proc plot; by alpha; where alpha = 0.05;
  plot power*n = source / vref = 0.80, 0.90 vaxis = 0 to 1.0 by 0.1;
run;

/*******************************************/

/* Then do analysis without an interaction */

/**  Read in the expected means **/

data case2;
  a = 2;  b = 6; frq = 10;   *<==== Put in your own numbers for a and b;
  do species = 1 to a;       *  (be careful if you change the value for frq);
    do rate  = 0 to b-1;
    input y @@;
    output; end; end;
cards;
0.25 0.40  0.50 0.55  0.55 0.50
0.35 0.50  0.60 0.65  0.65 0.60
;
```

```
/**  Obtaining the treatment and contrast SS's (SSH's) using PROC GLM **/

proc glm data=case2 outstat=stat;
  class species rate;  freq frq;
  model y = species|rate / ss3;
  contrast 'Linear response'    rate -5 -3 -1  1  3  5;
  contrast 'Quadratic response' rate  5 -1 -4 -4 -1  5;
title 'Two-way ANOVA Example -- with no interaction';
run;

/** Calculating the SSM's and the POWER for each source **/
/**           for a variety of conditions               **/

data power2 (keep=source dfh dfe n ssm ssh nc fc alpha power);
  set stat; frq=10; retain cells;
  if _type_='ERROR' then cells = df/(frq-1);
  if _type_='SS3' or _type_='CONTRAST' then do;
      dfh = df; ssm = ss/frq; source = _source_;
 /****************************************/
 /*****  Put your own values in here *****/
  do alpha = 0.01, 0.05, 0.10;
    do n = 2 to 30 by 2;
       do sigma = 0.4;
/****************************************/
       dfe = (n-1)*cells; ssh = n*ssm;
       nc = ssh/(sigma**2);
       fc = finv(1-alpha, dfh, dfe, 0);
      power = 1-probf(fc, dfh, dfe, nc);
      if power = . then power = 1.0;
      output;
  end; end; end; end;
run;

/** Printing and plotting out the results  **/

proc print noobs; by source alpha notsorted;
  where alpha = 0.05;
  var ssh dfh dfe n fc nc power;
title2 'Results of Power Analysis';
run;
proc sort data=power2; by alpha n;
run;
proc plot; by alpha; where alpha = 0.05;
  plot power*n = source / vref = 0.80, 0.90 vaxis = 0 to 1.0 by 0.1;
run;
```

The output from this program is:

```
                  Two-way ANOVA Example -- with an interaction                1

                       General Linear Models Procedure
                           Class Level Information

                        Class    Levels    Values

                        SPECIES       2    1 2

                        RATE          6    0 1 2 3 4 5


                   Number of observations in data set = 120


                 Two-way ANOVA Example -- with an interaction                2

                       General Linear Models Procedure

Dependent Variable: Y
Frequency:          FRQ
                                   Sum of           Mean
Source                  DF        Squares         Square   F Value     Pr > F

Model                   11     2.44256250     0.22205114  99999.99     0.0

Error                  108     0.00000000     0.00000000

Corrected Total        119     2.44256250

                  R-Square           C.V.       Root MSE                Y Mean

                  1.000000              0              0            0.53125000


Source                  DF    Type III SS    Mean Square   F Value     Pr > F

SPECIES                  1     0.63802083     0.63802083  99999.99     0.0
RATE                     5     1.74518750     0.34903750  99999.99     0.0
SPECIES*RATE             5     0.05935417     0.01187083  99999.99     0.0

Contrast                DF    Contrast SS    Mean Square   F Value     Pr > F

Linear response          1     1.36718750     1.36718750  99999.99     0.0
Quadratic response       1     0.37800000     0.37800000  99999.99     0.0
```

Two-way ANOVA Example -- with an interaction 3

List of Sources and SS's from PROC GLM

NAME	_SOURCE_	_TYPE_	DF	SS	F	PROB
Y	ERROR	ERROR	108	0.00000	.	.
Y	SPECIES	SS3	1	0.63802	99999.99	0
Y	RATE	SS3	5	1.74519	99999.99	0
Y	SPECIES*RATE	SS3	5	0.05935	99999.99	0
Y	Linear response	CONTRAST	1	1.36719	99999.99	0
Y	Quadratic response	CONTRAST	1	0.37800	99999.99	0

Two-way ANOVA Example -- with an interaction 4

Results of Power Analysis

------------------------- SOURCE=SPECIES ALPHA=0.05 --------------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.12760	1	12	2	4.74723	0.7975	0.13053
0.25521	1	36	4	4.11317	1.5951	0.23324
0.38281	1	60	6	4.00119	2.3926	0.33098
0.51042	1	84	8	3.95457	3.1901	0.42307
0.63802	1	108	10	3.92901	3.9876	0.50769
0.76563	1	132	12	3.91288	4.7852	0.58381
0.89323	1	156	14	3.90176	5.5827	0.65109
1.02083	1	180	16	3.89364	6.3802	0.70968
1.14844	1	204	18	3.88745	7.1777	0.76007
1.27604	1	228	20	3.88257	7.9753	0.80294
1.40365	1	252	22	3.87862	8.7728	0.83906
1.53125	1	276	24	3.87537	9.5703	0.86925
1.65885	1	300	26	3.87264	10.3678	0.89428
1.78646	1	324	28	3.87032	11.1654	0.91490
1.91406	1	348	30	3.86832	11.9629	0.93178

--------------------------- SOURCE=RATE ALPHA=0.05 ---------------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.34904	5	12	2	3.10588	2.1815	0.12420
0.69807	5	36	4	2.47717	4.3630	0.27511
1.04711	5	60	6	2.36827	6.5445	0.43140
1.39615	5	84	8	2.32313	8.7259	0.57521
1.74519	5	108	10	2.29843	10.9074	0.69580
2.09422	5	132	12	2.28286	13.0889	0.79001
2.44326	5	156	14	2.27214	15.2704	0.85963
2.79230	5	180	16	2.26431	17.4519	0.90878
3.14134	5	204	18	2.25834	19.6334	0.94220
3.49037	5	228	20	2.25364	21.8148	0.96419
3.83941	5	252	22	2.24985	23.9963	0.97826
4.18845	5	276	24	2.24671	26.1778	0.98704
4.53749	5	300	26	2.24409	28.3593	0.99240
4.88652	5	324	28	2.24185	30.5408	0.99561
5.23556	5	348	30	2.23993	32.7223	0.99750

Results of Power Analysis

----------------------- SOURCE=SPECIES*RATE ALPHA=0.05 -----------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.01187	5	12	2	3.10588	0.07419	0.05215
0.02374	5	36	4	2.47717	0.14839	0.05559
0.03561	5	60	6	2.36827	0.22258	0.05897
0.04748	5	84	8	2.32313	0.29677	0.06238
0.05935	5	108	10	2.29843	0.37096	0.06585
0.07123	5	132	12	2.28286	0.44516	0.06938
0.08310	5	156	14	2.27214	0.51935	0.07297
0.09497	5	180	16	2.26431	0.59354	0.07662
0.10684	5	204	18	2.25834	0.66773	0.08033
0.11871	5	228	20	2.25364	0.74193	0.08410
0.13058	5	252	22	2.24985	0.81612	0.08792
0.14245	5	276	24	2.24671	0.89031	0.09181
0.15432	5	300	26	2.24409	0.96451	0.09575
0.16619	5	324	28	2.24185	1.03870	0.09974
0.17806	5	348	30	2.23993	1.11289	0.10379

--------------------- SOURCE=Linear response ALPHA=0.05 ----------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.27344	1	12	2	4.74723	1.7090	0.22568
0.54687	1	36	4	4.11317	3.4180	0.43631
0.82031	1	60	6	4.00119	5.1270	0.60565
1.09375	1	84	8	3.95457	6.8359	0.73389
1.36719	1	108	10	3.92901	8.5449	0.82565
1.64062	1	132	12	3.91288	10.2539	0.88854
1.91406	1	156	14	3.90176	11.9629	0.93022
2.18750	1	180	16	3.89364	13.6719	0.95708
2.46094	1	204	18	3.88745	15.3809	0.97401
2.73437	1	228	20	3.88257	17.0898	0.98448
3.00781	1	252	22	3.87862	18.7988	0.99084
3.28125	1	276	24	3.87537	20.5078	0.99466
3.55469	1	300	26	3.87264	22.2168	0.99691
3.82812	1	324	28	3.87032	23.9258	0.99823
4.10156	1	348	30	3.86832	25.6348	0.99900

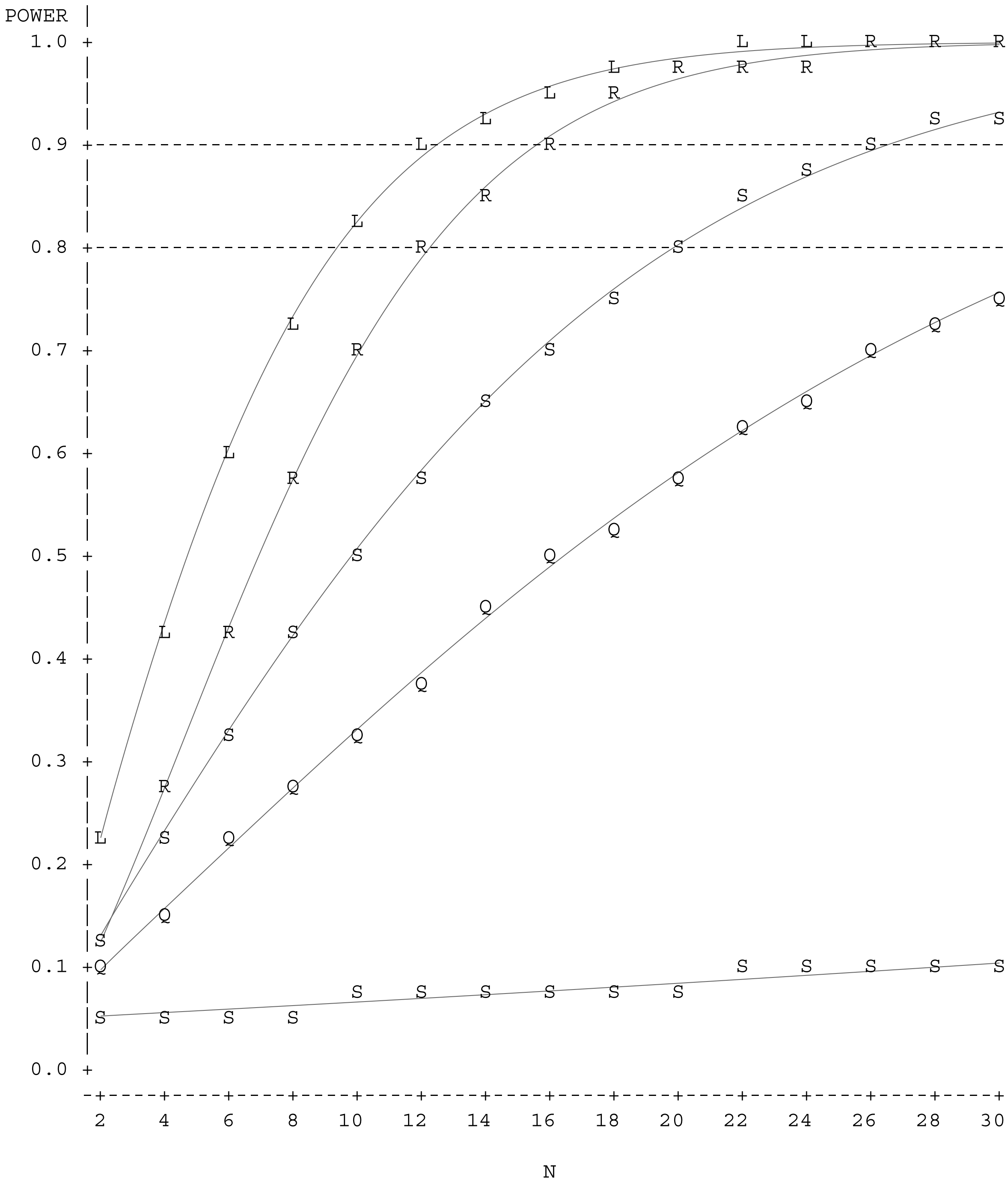
Results of Power Analysis

-------------------- SOURCE=Quadratic response ALPHA=0.05 --------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.0756	1	12	2	4.74723	0.4725	0.09716
0.1512	1	36	4	4.11317	0.9450	0.15721
0.2268	1	60	6	4.00119	1.4175	0.21611
0.3024	1	84	8	3.95457	1.8900	0.27441
0.3780	1	108	10	3.92901	2.3625	0.33145
0.4536	1	132	12	3.91288	2.8350	0.38664
0.5292	1	156	14	3.90176	3.3075	0.43948
0.6048	1	180	16	3.89364	3.7800	0.48963
0.6804	1	204	18	3.88745	4.2525	0.53687
0.7560	1	228	20	3.88257	4.7250	0.58107
0.8316	1	252	22	3.87862	5.1975	0.62217
0.9072	1	276	24	3.87537	5.6700	0.66017
0.9828	1	300	26	3.87264	6.1425	0.69515
1.0584	1	324	28	3.87032	6.6150	0.72720
1.1340	1	348	30	3.86832	7.0875	0.75645

------------------------------- ALPHA=0.05 -------------------------------

Plot of POWER*N. Symbol is value of SOURCE.



NOTE: 5 obs hidden.

General Linear Models Procedure
Class Level Information

Class	Levels	Values
SPECIES	2	1 2
RATE	6	0 1 2 3 4 5

Number of observations in data set = 120

General Linear Models Procedure

Dependent Variable: Y
Frequency: FRQ

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	1.64166667	0.14924242	99999.99	0.0
Error	108	0.00000000	0.00000000		
Corrected Total	119	1.64166667			

R-Square	C.V.	Root MSE	Y Mean
1.000000	0	0	0.50833333

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SPECIES	1	0.30000000	0.30000000	99999.99	0.0
RATE	5	1.34166667	0.26833333	99999.99	0.0
SPECIES*RATE	5	0.00000000	0.00000000	99999.99	0.0

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Linear response	1	0.87500000	0.87500000	99999.99	0.0
Quadratic response	1	0.46666667	0.46666667	99999.99	0.0

------------------------- SOURCE=SPECIES ALPHA=0.05 -------------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.06	1	12	2	4.74723	0.375	0.08727
0.12	1	36	4	4.11317	0.750	0.13456
0.18	1	60	6	4.00119	1.125	0.18111
0.24	1	84	8	3.95457	1.500	0.22763
0.30	1	108	10	3.92901	1.875	0.27377
0.36	1	132	12	3.91288	2.250	0.31915
0.42	1	156	14	3.90176	2.625	0.36344
0.48	1	180	16	3.89364	3.000	0.40638
0.54	1	204	18	3.88745	3.375	0.44777
0.60	1	228	20	3.88257	3.750	0.48743
0.66	1	252	22	3.87862	4.125	0.52527
0.72	1	276	24	3.87537	4.500	0.56121
0.78	1	300	26	3.87264	4.875	0.59520
0.84	1	324	28	3.87032	5.250	0.62724
0.90	1	348	30	3.86832	5.625	0.65734

-------------------------- SOURCE=RATE ALPHA=0.05 --------------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.26833	5	12	2	3.10588	1.6771	0.10524
0.53667	5	36	4	2.47717	3.3542	0.21521
0.80500	5	60	6	2.36827	5.0312	0.33409
1.07333	5	84	8	2.32313	6.7083	0.45258
1.34167	5	108	10	2.29843	8.3854	0.56262
1.61000	5	132	12	2.28286	10.0625	0.65931
1.87833	5	156	14	2.27214	11.7396	0.74061
2.14667	5	180	16	2.26431	13.4167	0.80651
2.41500	5	204	18	2.25834	15.0937	0.85831
2.68333	5	228	20	2.25364	16.7708	0.89797
2.95167	5	252	22	2.24985	18.4479	0.92765
3.22000	5	276	24	2.24671	20.1250	0.94940
3.48833	5	300	26	2.24409	21.8021	0.96507
3.75667	5	324	28	2.24185	23.4792	0.97618
4.02500	5	348	30	2.23993	25.1562	0.98393

Results of Power Analysis

----------------------- SOURCE=SPECIES*RATE ALPHA=0.05 -----------------------

SSH	DFH	DFE	N	FC	NC	POWER
1.589E-31	5	12	2	3.10588	9.9314E-31	0.05
3.178E-31	5	36	4	2.47717	1.9863E-30	0.05
4.7671E-31	5	60	6	2.36827	2.9794E-30	0.05
6.3561E-31	5	84	8	2.32313	3.9726E-30	0.05
7.9451E-31	5	108	10	2.29843	4.9657E-30	0.05
9.5341E-31	5	132	12	2.28286	5.9588E-30	0.05
1.1123E-30	5	156	14	2.27214	6.952E-30	0.05
1.2712E-30	5	180	16	2.26431	7.9451E-30	0.05
1.4301E-30	5	204	18	2.25834	8.9382E-30	0.05
1.589E-30	5	228	20	2.25364	9.9314E-30	0.05
1.7479E-30	5	252	22	2.24985	1.0925E-29	0.05
1.9068E-30	5	276	24	2.24671	1.1918E-29	0.05
2.0657E-30	5	300	26	2.24409	1.2911E-29	0.05
2.2246E-30	5	324	28	2.24185	1.3904E-29	0.05
2.3835E-30	5	348	30	2.23993	1.4897E-29	0.05

--------------------- SOURCE=Linear response ALPHA=0.05 ----------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.175	1	12	2	4.74723	1.0937	0.16133
0.350	1	36	4	4.11317	2.1875	0.30176
0.525	1	60	6	4.00119	3.2812	0.42963
0.700	1	84	8	3.95457	4.3750	0.54293
0.875	1	108	10	3.92901	5.4687	0.63974
1.050	1	132	12	3.91288	6.5625	0.72009
1.225	1	156	14	3.90176	7.6562	0.78522
1.400	1	180	16	3.89364	8.7500	0.83701
1.575	1	204	18	3.88745	9.8437	0.87753
1.750	1	228	20	3.88257	10.9375	0.90878
1.925	1	252	22	3.87862	12.0312	0.93259
2.100	1	276	24	3.87537	13.1250	0.95055
2.275	1	300	26	3.87264	14.2187	0.96397
2.450	1	324	28	3.87032	15.3125	0.97390
2.625	1	348	30	3.86832	16.4062	0.98120

Results of Power Analysis

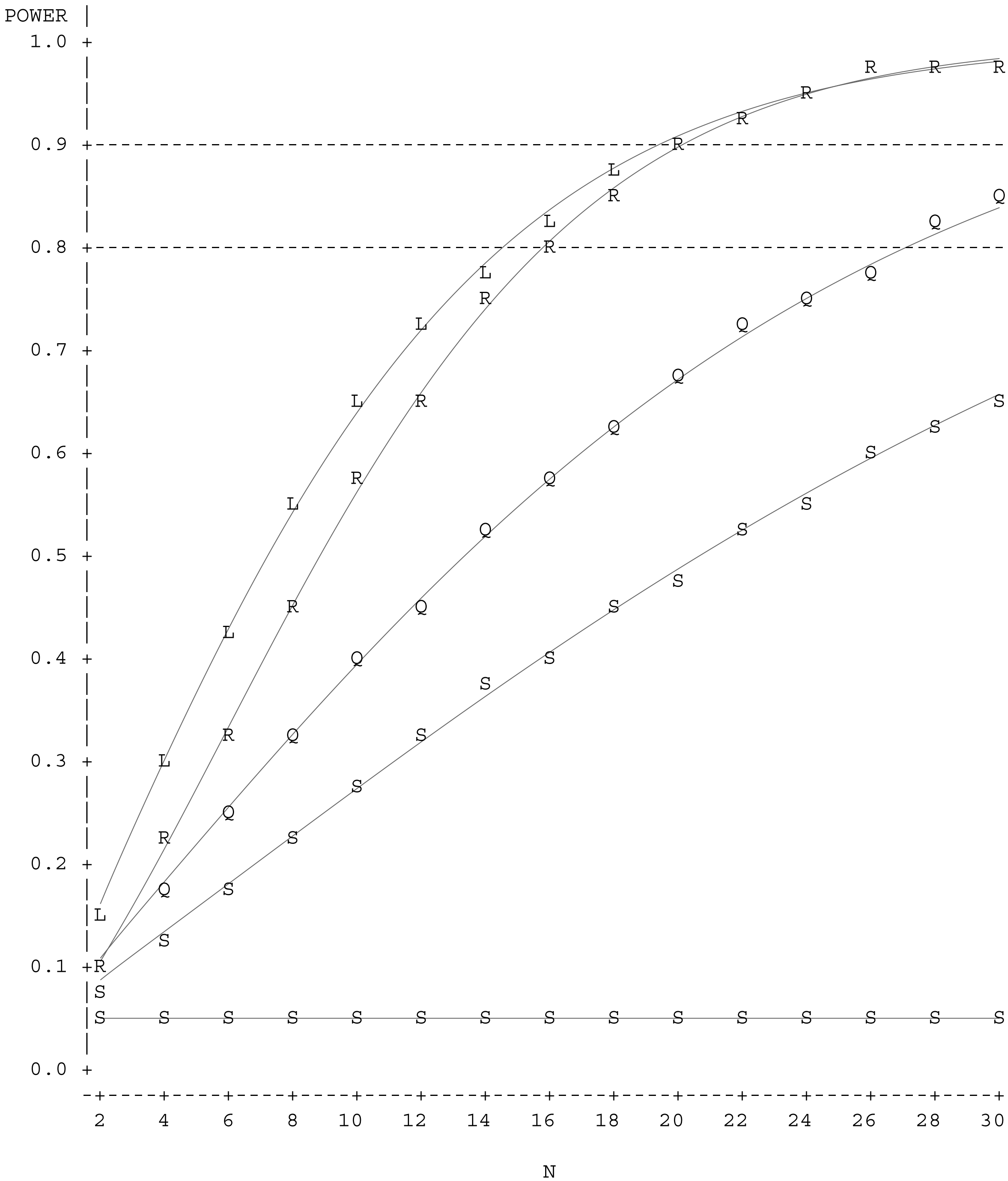
-------------------- SOURCE=Quadratic response ALPHA=0.05 --------------------

SSH	DFH	DFE	N	FC	NC	POWER
0.09333	1	12	2	4.74723	0.58333	0.10847
0.18667	1	36	4	4.11317	1.16667	0.18311
0.28000	1	60	6	4.00119	1.75000	0.25575
0.37333	1	84	8	3.95457	2.33333	0.32665
0.46667	1	108	10	3.92901	2.91667	0.39474
0.56000	1	132	12	3.91288	3.50000	0.45915
0.65333	1	156	14	3.90176	4.08333	0.51930
0.74667	1	180	16	3.89364	4.66667	0.57487
0.84000	1	204	18	3.88745	5.25000	0.62573
0.93333	1	228	20	3.88257	5.83333	0.67188
1.02667	1	252	22	3.87862	6.41667	0.71345
1.12000	1	276	24	3.87537	7.00000	0.75066
1.21333	1	300	26	3.87264	7.58333	0.78377
1.30667	1	324	28	3.87032	8.16667	0.81308
1.40000	1	348	30	3.86832	8.75000	0.83889

Results of Power Analysis

-------------------------------- ALPHA=0.05 --------------------------------

Plot of POWER*N. Symbol is value of SOURCE.



NOTE: 7 obs hidden.

Appendix 1: <u>Summary of Tests and SAS functions</u>

1) Comparing one mean, μ, with a standard μ_o, with σ known: using the z-distribution

	Null Hypothesis	Alternate Hypothesis	Effect Size & NonCentrality parameter	Decision Rule	Calculation of Critical Value	Calculation of Power
a)	$\mu \geq \mu_o$	$\mu = \mu_a < \mu_o$	$d = \frac{\mu_a - \mu_o}{\sigma}$	$z < zc$	$P(z<zc \mid H_o) = p_o=\alpha$ `SAS: zc=PROBIT(α)`	$P(z<zc \mid H_a) = 1-p_a$ `PROBNORM(zc-nc)`
b)	$\mu \leq \mu_o$	$\mu = \mu_a > \mu_o$		$z > zc$	$P(z>zc \mid H_o) = p_o=\alpha$ `SAS: zc=PROBIT(1-α)`	$P(z>zc \mid H_a) = 1-p_a$ `1-PROBNORM(zc-nc)`

c)	$\mu = \mu_o$	$\mu = \mu_a \neq \mu_o$	$nc = \delta = d\sqrt{n}$	$\lvert z \rvert < zc$	$P(\lvert z\rvert<zc \mid H_o) = p_o=\alpha$
					`SAS: zc=PROBIT(1-α/2)`
						$P(abs(z)>zc \mid H_a) = 1-p_a$
						`PROBNORM(zc-nc)`
						`+ 1 - PROBNORM(nc-zc)`

2) Comparing one mean, μ, with a standard μ_o, with σ unknown: using the t-distribution

	Null Hypothesis	Alternate Hypothesis	Effect Size d	Decision Rule	Calculation of Critical Value	Calculation of Power
a)	$\mu \geq \mu_o$	$\mu = \mu_a < \mu_o$	$d = \frac{\mu_a - \mu_o}{\sigma}$	$t < tc$	$P(t<tc \mid H_o) = p_o=\alpha$ `SAS: tc=TINV(α,df)`	$P(t<tc \mid H_a) = 1-p_a$ `PROBT(tc,df,nc)`
b)	$\mu \leq \mu_o$	$\mu = \mu_a > \mu_o$		$t > tc$	$P(t>tc \mid H_o) = p_o=\alpha$ `SAS: tc=TINV(1-α,df)`	$P(t>tc \mid H_a) = 1-p_a$ `1-PROBT(tc,df,nc)`

c)	$\mu = \mu_o$	$\mu = \mu_a \neq \mu_o$	$nc = \delta = d\sqrt{n}$	$\lvert t \rvert < tc$	$P(\lvert t\rvert<tc \mid H_o) = p_o=\alpha$
					`SAS: tc=TINV(1-α/2,df)`
						$P(abs(t)>tc \mid H_a) = 1-p_a$
						`PROBT(-tc,df,abs(nc))`
						`+ 1 - PROBT(tc,df,abs(nc))`

Note: σ is estimated by the observed standard deviation s .

3) Comparing 2 means, μ_1 and μ_2, σ unknown: using the 2-sample t-test

	Null Hypothesis	Alternate Hypothesis	Effect Size d	Decision Rule	Calculation of Critical Value	Calculation of Power
a)	$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$	$d = \frac{\mu_1 - \mu_2}{\sigma}$	$t < tc$	$P(t<tc \mid H_o) = p_o=\alpha$ `SAS: tc=TINV(α,df)`	$P(t<tc \mid H_a) = 1-p_a$ `PROBT(tc,df,nc)`
b)	$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	$nc = \delta =$	$t > tc$	$P(t>tc \mid H_o) = p_o=\alpha$ `SAS: tc=TINV(1-α,df)`	$P(t>tc \mid H_a) = 1-p_a$ `1-PROBT(tc,df,nc)`

c)	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$= d\sqrt{\frac{n_1 n_2}{2(n_1+n_2)}}$	$\lvert t \rvert < tc$	$P(\lvert t \rvert<tc \mid H_o) = p_o=\alpha$
					`SAS: tc=TINV(1-α/2,df)`
						$P(\lvert t \rvert>tc \mid H_a) = 1-p_a$
						`PROBT(-tc,df,nc) +`
						`1 - PROBT(tc,df,nc)`

Note: σ is estimated by the pooled standard deviation calculated by

$$S_p = \sqrt{\frac{(n-1)S_1^2 + (n-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{S_1^2 + S_2^2}{n}}$$

if $n_1 = n_2 = n$. S_1^2 and S_2^2 are the squared standard deviations for each mean.

4) Comparing **a** means, σ unknown: using ANOVA

Null Hypothesis	Alternate Hypothesis	Effect Size d	Decision Rule	Calculation of Critical Value	Calculation of Power
$\mu_1=\mu_2=...=\mu_a$	$\sum c_i\mu_i \neq 0$ where $\sum c_i = 0$	$nc = \lambda =$ $= \frac{n(\sum \mu_i - \mu)^2}{\sigma^2}$	$F > fc$	$P(F>fc \mid H_o) = p_o=\alpha$ `SAS: fc=FINV(1-α,df_h,df_e,0)`	$P(F>fc \mid H_a) = 1-p_a$ `1-PROBF(fc,df_h,df_e,nc)`

Note: σ^2 is estimated by the observed MSE or MSW from the ANOVA (see BI #25).

Appendix 2: NOTATION

H_o	Null hypothesis
H_a	Alternate hypothesis
α	Type I error or the significance level
$1-\alpha$	Confidence level
p	Observed confidence level, usually compared to α for determination of significance.
1-p	Observed confidence level
β	Type II error
$1-\beta$	Power
p_a	Observed type II error rate
$1-p_a$	Observed confidence level for rejecting the null hypothesis
μ	general symbol for the expected value of a mean, used here to represent the grand mean
a	number of groups or treatments for the one-way ANOVA situation.
μ_i	the expected value for an individual group or treatment mean, e.g., i = 1 to a.
μ_{max}	the maximum expected value for a group of means
μ_{min}	the minimum expected value for a group of means
r_i	the difference between a treatment mean and grand mean, i.e. $\mu_i - \mu$.
σ	standard deviation
s	sample (observed) standard deviation
n	sample size
d	effect size (for 2 groups) or effect size index (for 3 or more groups), $d = (\mu_{max} - \sigma_{min})/\sigma$
df_h	hypothesis or numerator degrees of freedom, e.g. $df_h = a - 1$.
df_e	error or denominator degrees of freedom, e.g. $df_e = a(n-1)$
SSM	the sums of squares of the expected means, $\sum(\mu_i - \mu)^2 = \sum r_i^2$, using an equal sample size of one for each mean
SSH_a	alternate hypothesis sums of squares; $SSH_a = n * SSM$ for the equal sample size situation (denoted by SSH on the printouts).
nc	also known as λ; the non-centrality parameter for the F-distribution, $nc = SSH_a/\sigma^2$.
FC	critical F-value for a specific α, df_h, and df_e assuming H_o true (i.e. nc = 0)

Appendix 3: ADDITIONAL READING

Cohen, J. 1977. Statistical power analysis for the behavioral sciences, New York: Academic Press.

Forbes, L.S. 1990. A note on statistical power. The Auk 107:438-453.

Goldstein, R. 1989. Power and sample size via MS/PC-DOS computers. The American Statistician, 43: 253-260.

Gray, J.S. 1990. Statistics and the precautionary principle. Marine Pollution Bulletin 21: 174-176.

Kupper, L.L. and K.B. Hafner. 1989. How appropriate are popular sample size formulas? The American Statistician 43: 101-105

* Markel, M.D., 1991. The Power of a statistical test: What does insignificance mean? Veterinary Surgery 20: 209-214

O'Brien, R.G., 1986. Power analysis for linear models. Proceedings of the Eleventh Annual SAS Users Group Conference, Cary, NC: SAS Institute.

O'Brien, R.G., 1986. Using the SAS system to perform analyses for log-linear models. Proceedings of the Eleventh Annual SAS Users Group Conference, Cary, NC: SAS Institute.

Pentico, D.W. 1981. On the determination and use of optimal samples sizes for estimating the difference in means. The American Statistician, 35: 40-42.

* Peterman, R.M. 1990. The Importance of reporting statistical power: The Forest decline and acidic deposition example. Ecology 71: 2024-2027.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Can. J. Fisheries and Aquatic Sciences, 47: 2-15.

Whysong, G.L. and W.W. Brady. 1987. Frequency sampling and Type II errors. J. Range Mgmt. 40: 472-474.

Two papers of interest but which I have not yet had a chance to read are:

Matloff, N.S., 1991. Statistical hypothesis testing - problems and alternatives. Environ. Entomology 20: 1246-1250.

Young, L.J. and J.H. Young, 1991. Alternative view of statistical hypothesis testing. Environ. Entomology 20: 1241-1245.

* Particularly useful reading.