A SMOOTH POOLING PROCEDURE:

AN ALTERNATIVE TO A

SOMETIMES POOL PROCEDURE

IN THE ANALYSIS OF VARIANCE METHODOLOGY

BY

GEORGE D. McCLURE

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Statistics
University of Manitoba
Winnipeg, Manitoba

ISBN 0-315-71908-7

A SMOOTH POOLING PROCEDURE: AN ALTERNATIVE TO A SOMETIMES POOL PROCEDURE IN THE ANALYSIS OF VARIANCE METHODOLOGY

BY

GEORGE D. MCCLURE

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

ABSTRACT

Problems with the sometimes pooling procedure in the analysis of variance methodology motivate us to search for an alternative. The results of the sometimes pool procedure are discontinuous, arbitrary (due to the choice of the preliminary test size), and difficult to interpret. In this thesis we propose an alternative procedure which is continuous and is not arbitrary. We call this procedure the smooth pooling procedure.

Consider a randomized complete block experiment. Here, depending on the results of a preliminary test, the block and error sums of squares would sometimes be pooled in order to obtain an error mean square with more degrees of freedom. The proposed procedure utilizes an F statistic that has been constructed conditionally on the fact that the F statistic for blocks is less than or equal its observed value, and under the assumption that the block variance is 0. In this procedure the p-value that is used to assess treatment differences is the tail probability of this observed conditional F distribution.

The results produced by this procedure are intuitively attractive in that they pass smoothly from the result produced by the never pool procedure to the value produced by the always pool procedure (as the block p-value passes from 0 to 1). In addition, no arbitrary choice is involved in this procedure.

In this thesis the power function of this procedure is obtained for a specific example and compared to the power functions of the sometimes pool and classical never pool procedures. A computer program is provided that gives the results (p-values) which would result from this procedure in practice.

ACKNOWLEDGEMENTS

I am greatful to my advisor Dr. John F. Brewster for providing me with both an interesting problem and patient advice. The points raised by my committee members Dr. T. Berry, and Dr. S.W. Cheng were of help and are appreciated.

The financial support provided by a Natural Sciences and Engineering Research Council Scholarship was appreciated as was the support provided by the Department of Statistics through various forms of gainful employment.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER 1

INTRODUCTION AND SUMMARY

To pool or not pool sums of squares in the analysis of variance methodology is one of the classical conundrums of statistics. This issue has been studied for over forty years and remains for many unresolved (Bancroft and Han, 1977; Han, et al., 1988). As an illustration of the problem consider the following analysis of variance table.

Table 1.1 Oil content of Redwing flaxseed inoculated at different stages of growth with S. *linicola*, Winnipeg, 1947, in percent.

Source of variation	df	MS	E(MS)
treatment	5	6.33	$\sigma^2+3\sigma_t^2$
block	3	1.05	$\sigma^2+5\sigma_b^2$
error	15	1.31	σ^2

This ANOVA table is from an example of a randomized complete block design experiment used in Steel and Torrie (1980). No details about the nature of the blocks are given, but one might assume from the description of the treatments that we are dealing with a fixed or mixed effects model. However, since we are going to work with the random effects model in this thesis it suits our purposes to act as though this data resulted from the random effects model. Notice that while $E(MSB) \geq E(MSE)$ we observe $MSB < MSE$. The implication is that either the block variance, σ_b^2, is 0 or it is close to 0. The usual procedure in testing the main hypothesis of interest, $H_o:\sigma_t^2 = 0$, is to ignore the evidence about the lack of block variability and use the classical F test. This consists of constructing an F statistic as the ratio of MST to MSE which has an F distribution with (5,15) degrees of freedom under H_o. However, it seems wrong to ignore the evidence about the lack of block variability. If we knew that $\sigma_b^2 = 0$ then both MSB and MSE would be unbiased estimates of σ^2. If this were the case we would pool the information about σ^2 contained in MSB and MSE. This would allow us to construct a statistic with an F distribution with (5,18) degrees of freedom. This would result in a more sensitive and powerful test. The fact that the nature of the conclusion about the main factor of interest could be different if we knew there was no block variability is unsettling. The problem arises partly due to the nature of hypothesis testing --- in that even though we may accept $H_o:\sigma_b^2 = 0$ in the statistical sense, we do not accept it in the lay sense.

In an attempt to resolve this dilemma some researchers adopt a sometimes pool procedure (SPP). A sometimes pool procedure consists of first doing a preliminary test of the hypothesis $H_o:\sigma_b^2 = 0$. If this hypothesis is rejected, then we undertake the classical F test on the hypothesis about treatments (as described above). If we accept $H_o:\sigma_b^2 = 0$, however, then we act as though both MSB and MSE are estimates of the error variance σ^2 and pool these two estimates. The sometimes pool procedure is attractive in that it allows us to adjust the final test to compensate for the fact that σ_b^2 is near 0. In the literature on the SPP the classical procedure is referred to as the never pool procedure (NPP) (Paull, 1950).

There are three obvious drawbacks to the use of the SPP. Drawback number one is that it is arbitrary. This is due to the lack of an objective criteria on which to base the choice of the preliminary test size. Drawback number two is that the adjusted p-value for treatments is discontinuous when considered as a function of the p-value (F statistic) for blocks. The third drawback is that it is not clear how to interpret the results. These drawbacks will be illustrated in the following sections.

A new procedure which we call the smooth pooling procedure (SMTH) is proposed in this thesis. The new procedure is attractive in that it doesn't have either the arbitrary aspect or the discontinuity of the sometimes pool procedure, but does incorporate evidence about the block variability. The problem of the interpretation of the results does exist with this procedure however.

The smooth pooling procedure is motivated by work done by Brewster and Zidek (1974). Part of their work involved the estimation of σ^2 in the general linear model. In this problem they produced a smooth estimator which was admissible (within the class of scale equivariant estimators) and which dominated the usual inadmissible estimator. Our proposed procedure is based on an attempt to improve on the usual estimator of σ^2 that is used in the denominator of the test statistic, by incorporating information about the block variability.

In Section 1.1 we will use a numerical example to illustrate the various pooling procedures. Sections 1.2 and 1.3 deal with the pooling problem in a progressively more general manner.

1.1 AN EXAMPLE

ANOVA Table 1.2 will be used to illustrate how the various procedures are carried out in practice.

Table 1.2 A numerical ANOVA table used to illustrate the various pooling procedures in a randomized complete block experiment.

Source of variation	df	SS	MS
treatment	3	22.8	7.6
block	3	3.0	1.0
error	9	18.0	2.0

The assumed model for the data in Table 1.2 is $y_{ij} = \mu + TMT_i + BLK_j + \varepsilon_{ij}$ where

y_{ij} is the observation on the experimental unit which receives the ith treatment in the jth block,

μ is the overall mean,

TMT_i is the random effect of treatment i, $\sim N(0,\sigma_t^2)$,

BLK_j is the random effect of block j, $\sim N(0,\sigma_b^2)$,

ε_{ij} is a random error, $\sim N(0,\sigma^2)$,

and further TMT_i, BLK_j, and ε_{ij} are assumed to be independent. We want to test

$$H_o:\sigma_t^2 = 0 \qquad \text{vs.} \qquad H_a:\sigma_t^2 > 0.$$

1.1.1 THE NEVER POOL PROCEDURE (NPP)

This is the classical procedure. In this procedure we construct an F statistic as the ratio of the mean square for treatments to the mean square for error. Under the assumptions made about the model, and assuming that H_o is true, F has an F distribution with (3,9) degrees of freedom. If F is larger than $F(\alpha,3,9)$, where α is some chosen level of significance, then we reject the hypothesis $H_o:\sigma_t^2 = 0$ and conclude that the alternate hypothesis $H_a:\sigma_t^2 > 0$ is true. Otherwise we accept H_o. (The notation $F(\alpha,n_i,n_j)$ indicates the point on the F distribution with (n_i,n_j) degrees of freedom that is exceeded only $100\alpha\%$ of the time by a random variable with that distribution.)

If we use $\alpha = 0.05$ in testing $H_o:\sigma_t^2 = 0$ in the example in Table 1.2, then we accept H_o since the observed $F = 7.6/2 = 3.8 < 3.86 = F(0.05,3,9)$. The p-value of the observed F is 0.052.

1.1.2 THE SOMETIMES POOL PROCEDURE (SPP)

The sometimes pool procedure consists of two steps. First we pick a preliminary level of significance, α_p , and test the hypothesis

$$H_o:\sigma_b^2 = 0 \qquad \text{vs.} \qquad H_a:\sigma_b^2 > 0.$$

If this hypothesis is rejected then we proceed as under the NPP. That is, we construct the ratio of the mean square for treatments to the mean square for error, and compare this F statistic to $F(\alpha,3,9)$. If, however, $H_o:\sigma_b^2 = 0$ cannot be rejected then we proceed as follows. An adjusted estimate of the mean square for error, say AMSE, is calculated by adding the sums of squares for error and blocks and dividing by their combined degrees of freedom. Using this AMSE we calculate a pooled F statistic, F_p=MST/AMSE, and compare this to the F distribution with (3,(3+9)) degrees of freedom. The hypothesis $H_o:\sigma_t^2 = 0$ is rejected if F_p exceeds $F(\alpha,3,12)$. Otherwise we accept H_o.

In the example, we will use a preliminary test size of $\alpha_p = 0.25$ (a size often recommended in the literature) and a final test size of $\alpha = 0.05$. Since the result of the test for blocks is $F = 0.5 < 1.63 = F(0.25,3,9)$ we can't reject $H_o:\sigma_b^2 = 0$. In this case our adjusted estimate of the error variance is

$$\text{AMSE} = \frac{3+18}{3+9} = 1.75 .$$

Using this adjusted estimate of σ^2 to construct an F statistic we get $F_p = 4.34$ with (3,12) degrees of freedom. Using $\alpha = 0.05$ we reject $H_o:\sigma_t^2 = 0$ since $F_p = 4.34 > 3.49 = F(0.05,3,12)$. In this example the p-value of the observed F for blocks is 0.69. This means that given the observed ANOVA table any choice of the preliminary test size less than 0.69 results in a decision to accept $H_o:\sigma_b^2 = 0$, while any choice greater than 0.69 results in rejection of $H_o:\sigma_b^2 = 0$. Thus assuming a test of nominal size 0.05 of the hypothesis $H_o:\sigma_t^2 = 0$ we accept the hypothesis if our preliminary test size is greater than 0.69 and reject it if our preliminary test size is less than 0.69. Another way we can state this result is that any choice of $\alpha_p < 0.69$ results in an adjusted p-value for the main test of 0.027 while any choice of $\alpha_p > 0.69$ yields an adjusted p-value of 0.052. The fact that the final decision depends on this choice of the preliminary test size is disconcerting. This is an illustration of two of the drawbacks of using the SPP. First we have the lack of an objective rule on which to base the choice of the preliminary test size. This causes the result of the main test to be arbitrary. Secondly we have the problem of the interpretation of the adjusted p-value. The third drawback is illustrated in Section 1.2 (the drawback of the discontinuity in the adjusted p-value).

1.1.3 THE ALWAYS POOL PROCEDURE (APP)

For the sake of completeness it is noted that the literature also discusses a procedure known as the always pool procedure (APP). This consists of concluding on a subjective basis that a better model for the data is the completely randomized design model.

Analysis of the example as a completely randomized design results in an observed F of 4.34 with (3,12) degrees of freedom. Choice of $\alpha = 0.05$ for testing $H_o:\sigma_t^2 = 0$ in this case results in rejection of H_o since $F = 4.34 > 3.49 = F(0.05,3,12)$. An obvious problem with this procedure is that if there is substantial variance due to blocks and we pool the block and error mean squares we will have difficulty in rejecting $H_o:\sigma_t^2 = 0$ even when there is substantial variability due to treatments.

1.1.4 THE PROPOSED SMOOTH POOLING PROCEDURE (SMTH)

The proposed procedure consists of the following steps. First we obtain the distribution of the F statistic for treatments, conditional on the F statistic for blocks being less than or equal the observed F for blocks, and under the condition that the block variance is 0. Using this conditional F distribution we find the probability of observing a larger F statistic than that which we have observed. If this adjusted p-value is less than α then we reject

H_o:$\sigma^2_t = 0$. The required probability is found by numerical integration of the conditional F distribution over the appropriate region (i.e. from $F_{t:obs}$ to $+\infty$). This conditional F distribution has a density function given by

$$f^*(f_t; F_{b:obs}) = \frac{\int_0^{F_{b:obs}} f(f_t, f_b) df_b}{\int_0^{\infty} \int_0^{F_{b:obs}} f(f_t, f_b) df_b df_t} \quad ,$$

where
$$f(f_t, f_b) = \frac{f_t^{\frac{n_t}{2}-1} f_b^{\frac{n_b}{2}-1}}{(1 + \frac{n_b f_b}{n_e} + \frac{n_t f_t}{n_e})^{\frac{(n_t+n_b+n_e)}{2}}} k_1 \quad ,$$

with
$$k_1 = \left(\frac{n_t}{n_e}\right)^{\frac{n_t}{2}} \left(\frac{n_b}{n_e}\right)^{\frac{n_b}{2}} \frac{\Gamma(\frac{n_t+n_b+n_e}{2})}{\Gamma(\frac{n_t}{2})\Gamma(\frac{n_b}{2})\Gamma(\frac{n_e}{2})} \quad .$$

It can be shown that

$$\int_0^{F_{b:obs}} f(f_t, f_b) df_b = \frac{f_t^{\frac{n_t}{2}-1}}{(1 + \frac{n_t f_t}{n_e})^{\frac{n_e+n_t}{2}}} \left(\frac{n_t}{n_e}\right)^{\frac{n_t}{2}} \frac{\Gamma(\frac{n_e+n_t}{2})}{\Gamma(\frac{n_e}{2})\Gamma(\frac{n_t}{2})} k_2(f_t; F_{b:obs}) \quad ,$$

with $k_2(f_t; F_{b:obs}) = I[\frac{\frac{n_b F_{b:obs}}{n_e}}{(1 + \frac{n_b F_{b:obs}}{n_e} + \frac{n_t f_t}{n_e})}; \frac{n_b}{2}, \frac{n_e + n_t}{2}] \quad ,$

and $\int_0^{\infty} \int_0^{F_{b:obs}} f(f_t, f_b) df_b df_t = I[\frac{\frac{n_b F_{b:obs}}{n_e}}{(1 + \frac{n_b F_{b:obs}}{n_e})}; \frac{n_b}{2}, \frac{n_e}{2}] \quad .$

Here I(x;a,b) denotes the Beta cumulative distribution function with parameters a and b evaluated at x , i.e.,

$$I[x; a, b] = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} dy \quad .$$

Thus the adjusted p-value is obtained by numerically integrating the conditional F density, f^*, over the appropriate region rather than the integration of the unconditional F distribution over the same region. The density function of the unconditional F distribution is given by

$$f(f_t) = \frac{f_t^{\frac{n_t}{2}-1}}{(1+\frac{n_t f_t}{n_e})^{\frac{(n_t+n_e)}{2}}} \left(\frac{n_t}{n_e}\right)^{\frac{n_t}{2}} \frac{\Gamma(\frac{n_t+n_e}{2})}{\Gamma(\frac{n_t}{2})\Gamma(\frac{n_e}{2})} ,$$

and the integration of this function can be done by many of the standard computer packages or by table look up.

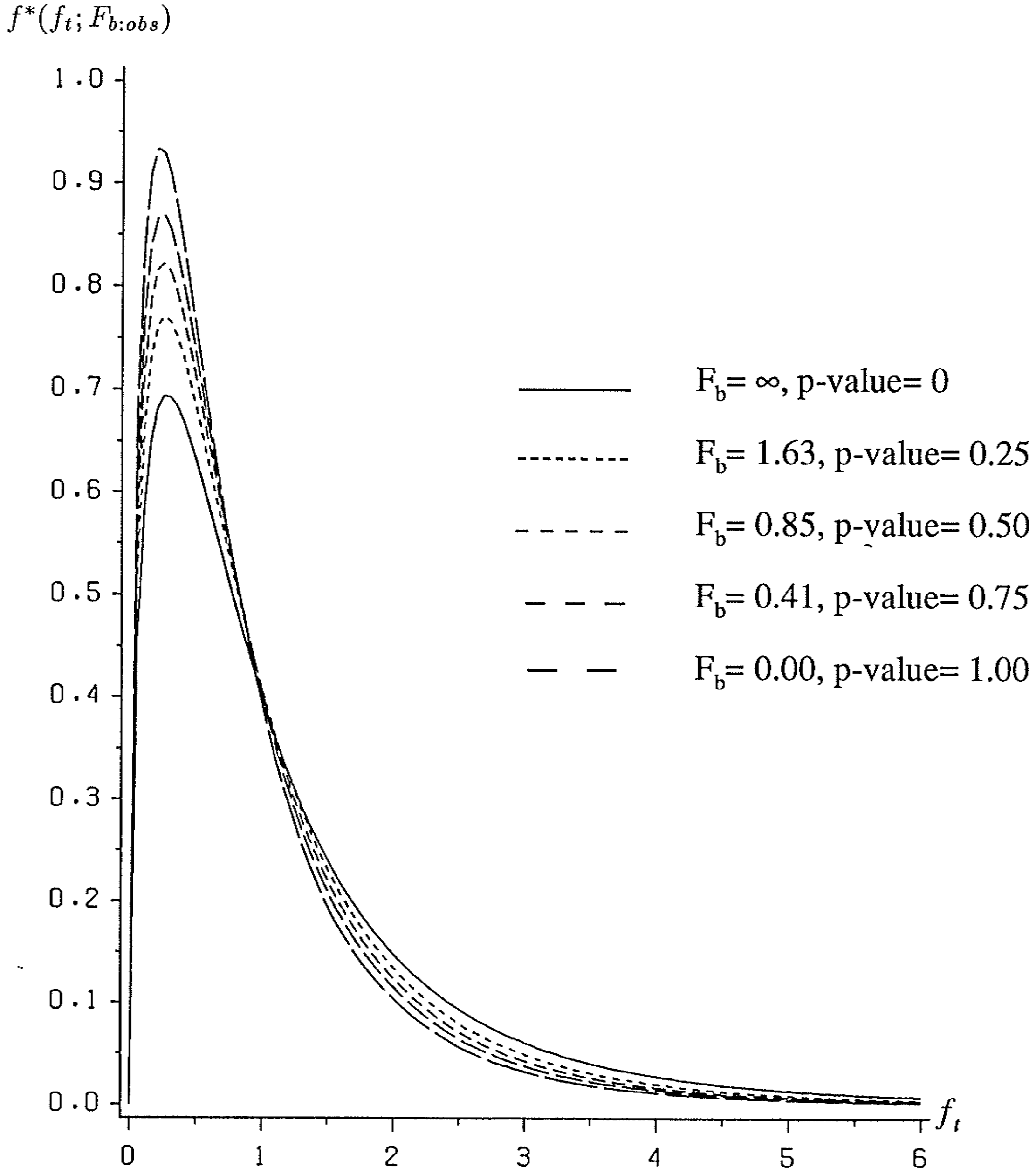
The conditional F distribution is illustrated in Figure 1.1.

In the given example the adjusted p-value is found to be 0.022 (see the Appendix). Use of $\alpha = 0.05$ as a criteria for quantifying rare events would indicate rejection of $H_o\!:\!\sigma_t^2 = 0$ in this example.

It can be shown that the conditional distribution approaches the classical never pool F distribution as the block p-value goes to 0, and also that conditional distribution approaches the always pool F distribution as the block p-value goes to 1. These facts are intuitively attractive.

As mentioned earlier the smooth pooling procedure is motivated by work on the estimation of σ^2 in the general linear model. The testing problem we are addressing is somewhat removed from this estimation problem. However, it is related. The methodology used to obtain the smooth dominating estimator of σ^2 was independent of the loss function that was chosen (Brewster and Zidek, 1974)! Indeed for a wide class of loss functions they showed that a smooth dominating estimator could be obtained from a specific conditional chi-squared distribution as opposed to the usual unconditional chi-squared

Figure 1.1 An example of a conditional F distribution --- for various block p-values (F statistics).



$n_t = 3, n_b = 3, n_e = 9$

distribution. This leads to the motivation of the specific conditional F distribution that is used to obtain the smooth testing procedure.

In summary we note that in the example used the adjusted p-values of the various procedures are ordered as NPP > SPP = APP > SMTH. In fact it can be shown that the smooth pooling procedure will always give the smallest adjusted p-value.

1.2 ANOTHER VIEW

As a further attempt to illustrate the issue at hand examine Table 1.3.

Table 1.3 A generalization of Table 1.2 in which the block mean square is allowed to vary.

Source	**df**	**MS**
treatment	3	7.6
block	3	MSB
error	9	2.0

This analysis of variance table is identical to Table 1.2 except that we allow the mean square for blocks to be variable in order to show how the results for the main test of interest change as the block mean square changes (i.e., as the block p-value or F statistic changes). For each of the procedures (the NPP, APP, SPP, and SMTH) Table 1.4 and Figures 1.2 to 1.5 illustrate what happens to the adjusted p-value for the test of $H_o:\sigma_t^2 = 0$ as the p-value for the test for blocks ranges from 0 to 1. Here, for each procedure, the adjusted p-value is the probability of observing a more extreme value of the statistic being used to make the test of $H_o:\sigma_t^2 = 0$ using the given procedure. We draw attention to this to emphasize that the various procedures have different sizes and that the interpretation of the results is not clear.

Table 1.4 and Figures 1.2 through 1.5 illustrate the following points about the various procedures.

1. The Never Pool Procedure ignores the information that the block variance is probably small or even nonexistent in the cases where the block p-value is large. If we have this information it seems wrong to ignore it. (See Figure 1.2 on page 13.)

Table 1.4 Adjusted p-values for the various pooling procedures as used with Table 1.3.

block p-value	adjusted p-values			
	NPP	APP	SPP(α_p=.25)	SMTH
0.9999	0.052	0.017	0.017	0.017
0.900	0.052	0.021	0.021	0.019
0.500	0.052	0.036	0.036	0.026
0.300	0.052	0.052	0.052	0.031
0.200	0.052	0.067	0.052	0.034
0.100	0.052	0.099	0.052	0.039
0.0001	0.052	0.667	0.052	0.052

Figure 1.2 Plot of the p-value for treatments versus the p-value for blocks, using the NPP to analyze the data in Table 1.3.

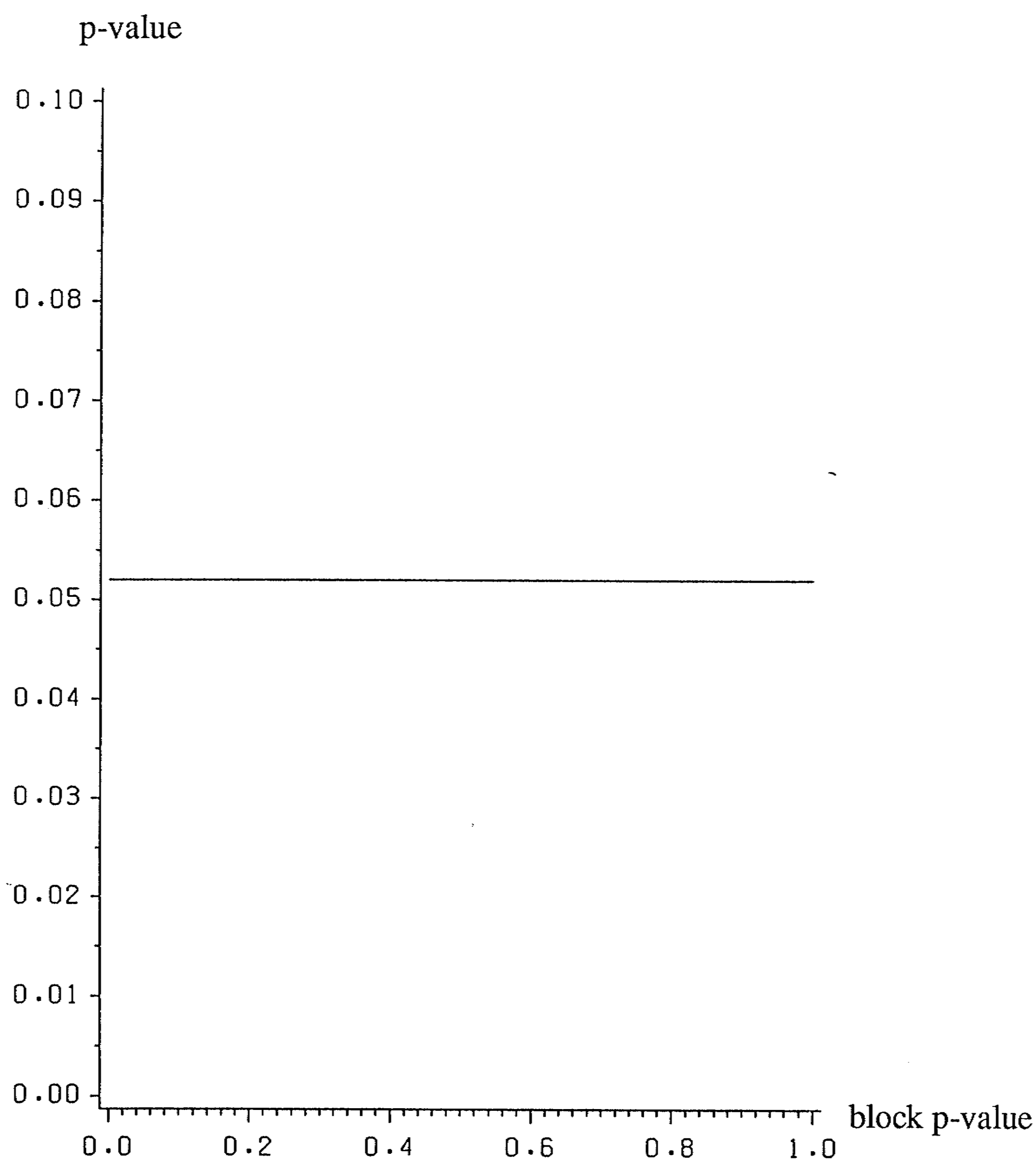


Figure 1.3 Plot of the adjusted p-value for treatments versus the p-value for blocks, using the APP to analyze the data in Table 1.3.

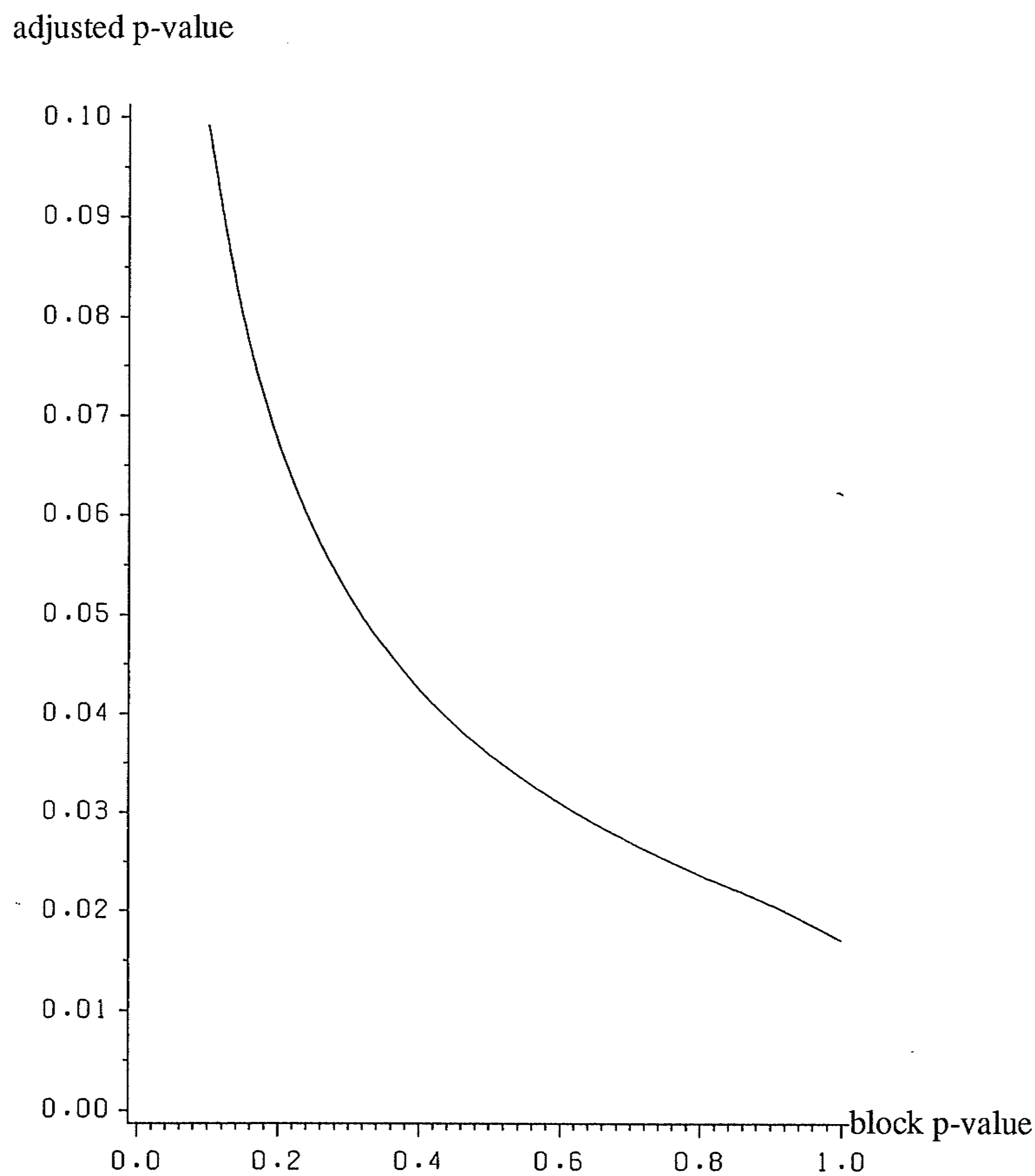


Figure 1.4 Plot of the adjusted p-value for treatments versus the p-value for blocks, using a SPP (α_p=.25) to analyze the data in Table 1.3.

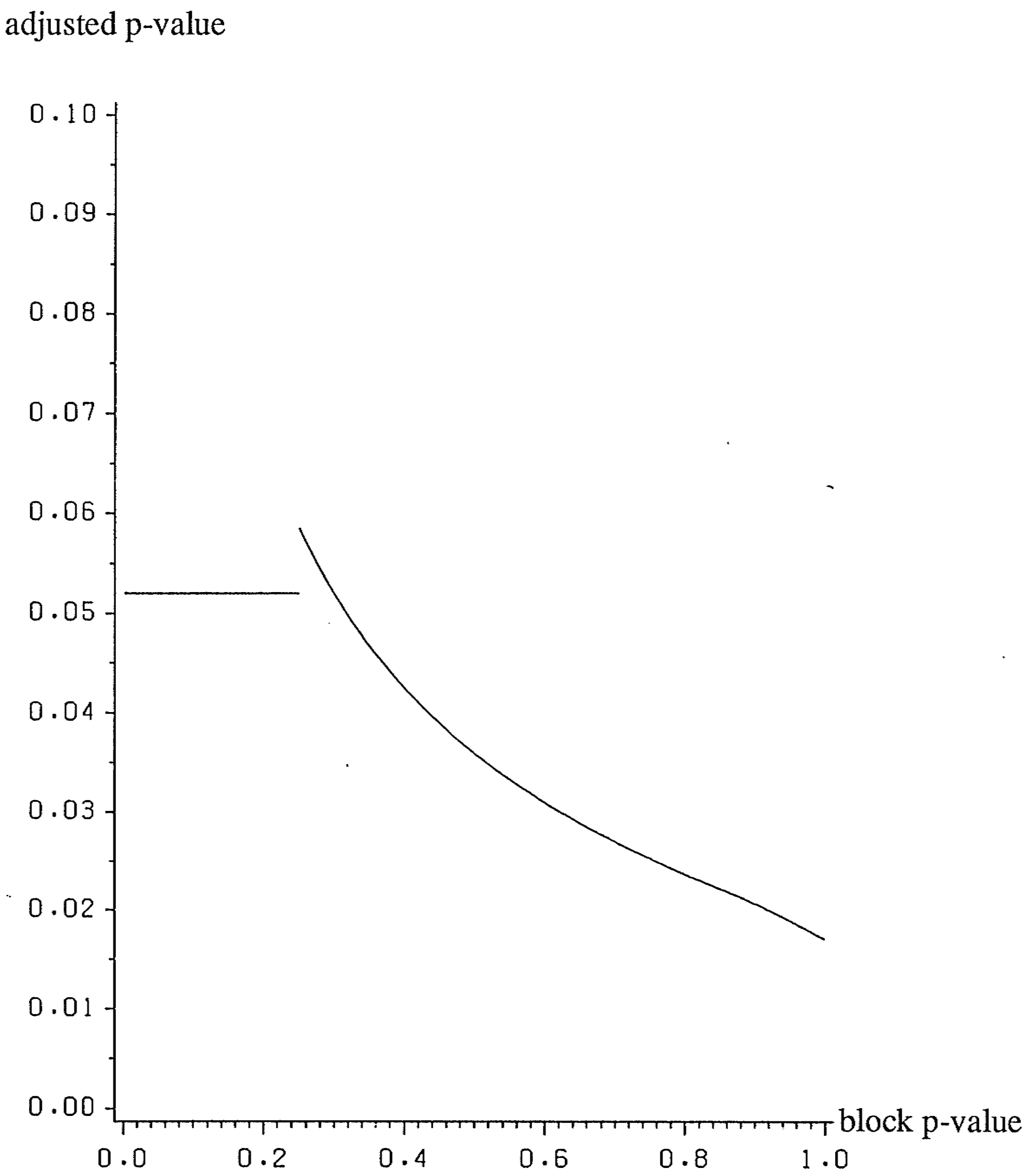
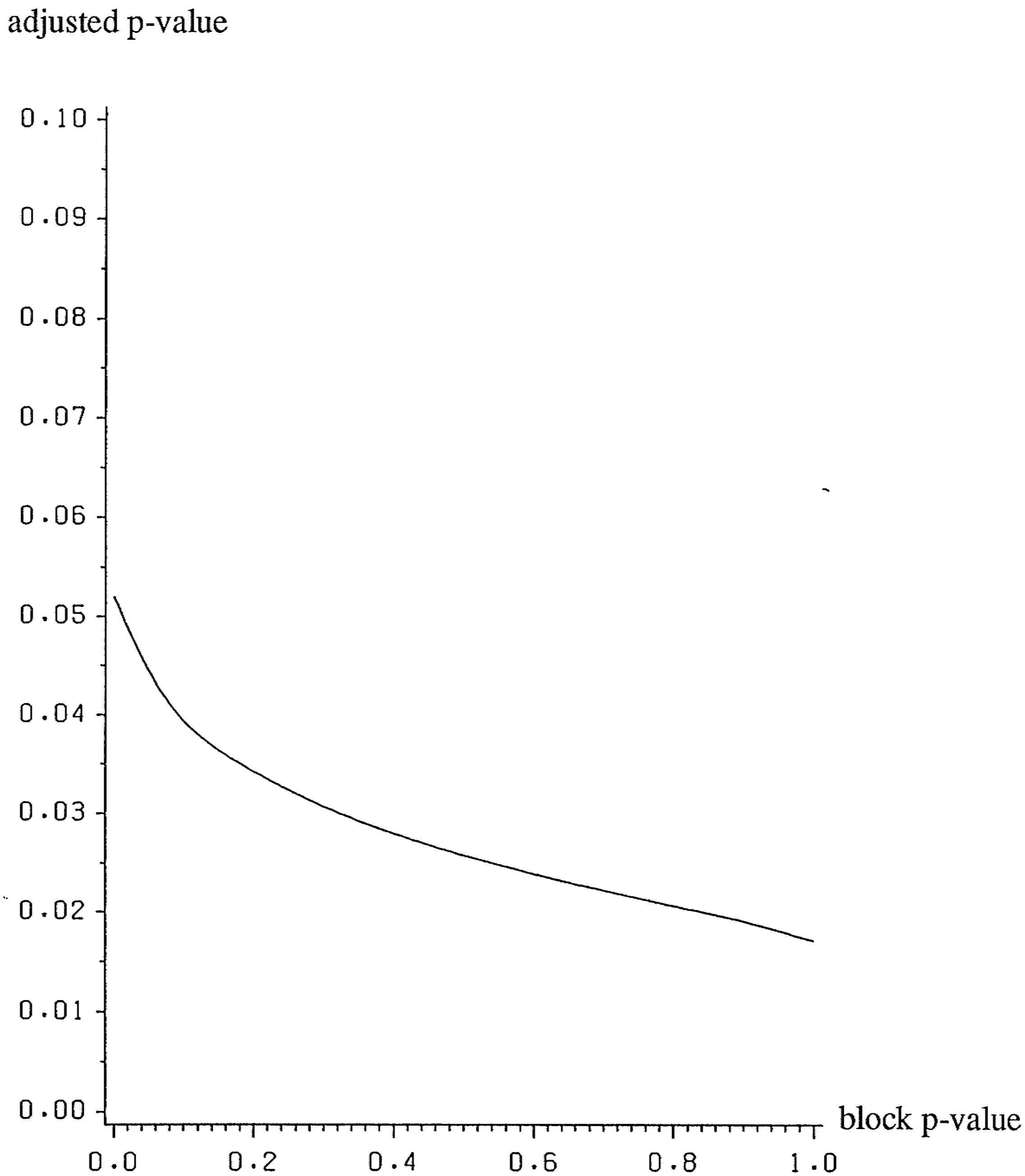


Figure 1.5 Plot of the adjusted p-value for treatments versus the p-value for blocks, using the SMTH to analyze the data in Table 1.3.



2. The Always Pool Procedure ignores the information provided by blocking. This seems wrong especially where there is evidence of significant variability due to blocks. (See Figure 1.3 on page 14.)

3. The Sometimes Pool Procedure results in a discontinuous function when we consider the adjusted p-value of the procedure as a function of the p-value for blocks. (See Figure 1.4 on page 15.) This fact alone is unsatisfactory and we also have the problem of choosing the significance level of the preliminary test, as well as the problem of the correct interpretation of the results.

4. The results from the Smooth Pooling Procedure are pleasing in that as the p-value for blocks ranges from 0 to 1 (from very significant to very insignificant) the adjusted p-value of the procedure passes smoothly from the value given by the Never Pool Procedure when the block p-value is 0 to the value given by the Always Pool Procedure when the block p-value is 1. (See Figure 1.5 on page 16.) Intuitively this seems like what we should be doing. In this procedure as well, however, there is the problem of the interpretation of the results.

1.3 THE GENERAL PROBLEM

We have been considering the pooling problem from the point of view that we are conducting a randomized complete block experiment. The pooling problem arises in other situations, however, so we will describe the problem from a more general point of view.

The problem we are faced with is as follows. We have three mean squares V_1, V_2, and V_3 --- called the auxiliary mean square (V_1), the classical denominator mean square (V_2), and the treatment mean square (V_3). The expected values of these mean squares are σ_1^2, σ_2^2, and σ_3^2, and their degrees of freedom are n_1, n_2, and n_3. We are interested in whether the treatments have introduced variability into the results of the experiment.

We wish to test the hypothesis $H_o:\sigma_3^2 = \sigma_2^2$ against the alternate hypothesis $H_a:\sigma_3^2 > \sigma_2^2$. This may be tested by constructing an F statistic as the ratio of V_3 to V_2. Since V_2 **always** has the same expectation as V_3 when H_o is true it is used as the denominator mean square in the classical NPP. Therefore we call it the classical denominator mean square. Under the usual assumptions this F statistic has an F distribution with (n_3,n_2) degrees of freedom when H_o is true.

It is possible that $\sigma_2^2 = \sigma_1^2$. If we knew this were the case then we would have conducted the experiment in a more efficient manner. However, given the manner in which we did conduct the experiment, how should we test for variability due to treatments? Since σ_1^2 may equal σ_2^2, V_1 may contain auxiliary information that could be used in estimating the quantity that is used as the denominator mean square for the main test of interest. We therefore call V_1 the auxiliary mean square.

Historically when faced with this problem researchers have proceeded in one of three ways --- by using the NPP, the APP, or the SPP. The never pool procedure uses the classical F statistic previously described. Philosophically its basis is that it is always a valid procedure regardless of whether $\sigma_2^2 = \sigma_1^2$

or not. The always pool procedure makes the assumption that indeed $\sigma_2^2 = \sigma_1^2$. On the basis of this assumption the information contained in V_2 and V_1 is pooled. The resulting pooled mean square is used for the denominator mean square in constructing an F_p statistic which is used to make the test of $H_o{:}\sigma_t^2 = 0$. Philosophically the basis of this procedure is that, in the researcher's opinion, σ_2^2 does equal σ_1^2 and therefore a more powerful test is achieved by this procedure. The sometimes pool procedure is conducted as follows. First a preliminary test of the hypothesis $H_o{:}\sigma_2^2 = \sigma_1^2$ is made at level α_p. If this hypothesis is rejected the F statistic of the never pool procedure is used to test the hypothesis of no variability due to treatments. If, however, $H_o{:}\sigma_2^2 = \sigma_1^2$ cannot be rejected, then the F_p statistic of the always pool procedure is used. The philosophical basis of this procedure is that it protects the researcher against ignoring the possibility that $\sigma_2^2 = \sigma_1^2$, and at the same time protects them from assuming that σ_2^2 and σ_1^2 are equal when in fact they may not be.

Fundamental differences in the results associated with different experimental designs provoke the need to look at two examples of the general situation. These differences in the results are due to the fact that in some models we find E(classical denominator MS) $\leq$ E(auxiliary MS), whereas in other models this ordering is reversed. We refer to these as Type A and Type B models, respectively.

As an illustration of the general problem, and using a Type A Model, consider Table 1.5. Table 1.5 is based on a randomized complete block experiment.

Table 1.5 An ANOVA table of a Type A Model; an experiment in which E(classical denominator MS) ≤ E(auxiliary MS).

Source of variation	df	MS	E(MS)
treatment	$n_3=n_t$	$V_3=V_t$	$\sigma_3^2=\sigma^2+n_t\sigma_t^2$
block	$n_1=n_b$	$V_1=V_b$	$\sigma_1^2=\sigma^2+n_b\sigma_b^2$
error	$n_2=n_e$	$V_2=V_e$	$\sigma_2^2=\sigma^2$

We are interested in testing the hypothesis $H_o:\sigma_t^2 = 0$ against the hypothesis $H_a:\sigma_t^2 > 0$. This is equivalent to testing the hypothesis $H_o:\sigma_3^2 = \sigma_2^2$ against the hypothesis $H_a:\sigma_3^2 > \sigma_2^2$. This can be tested using the classical F test. Under H_o, the statistic

$$F = \frac{V_3}{V_2} = \frac{V_t}{V_e}$$

has an F distribution with (n_t,n_e) degrees of freedom. If we routinely follow this procedure then this is what is called the never pool procedure.

If $\sigma_b^2 = 0$ then $\sigma_1^2 = \sigma^2$ and $\sigma_2^2 = \sigma^2$, so in this case both V_e and V_b are estimates of σ^2. If this is the case, then a more powerful test of the hypothesis $H_o:\sigma_t^2 = 0$ may be achieved by pooling the information contained in V_e and V_b and by constructing a pooled statistic F_p. If indeed $\sigma_b^2 = 0$, then F_p has the F distribution with (n_t,n_e+n_b) degrees of freedom. If this is routinely done then we have the always pool procedure.

The sometimes pool procedure consists of doing a preliminary test of the hypothesis $H_o:\sigma_b^2 = 0$, or equivalently $H_o:\sigma_2^2 = \sigma_1^2$. If this hypothesis is accepted, we use the F_p statistic of the always pool procedure to test the hypothesis of no variability due to treatments. If however we reject $H_o:\sigma_b^2 = 0$, then we use the F statistic of the never pool procedure.

As an example of the general problem, and using a Type B Model, consider Table 1.6. This is based on a replicated two-way random effects experiment.

Table 1.6 An ANOVA table of a Type B Model; an experiment in which E(classical denominator MS) $\geq$ E(auxiliary MS).

Source	df	MS	E(MS)
tmtA	$n_3 = n_t$	$V_3 = V_t$	$\sigma_3^2=\sigma^2+K\sigma_{AB}^2+JK\sigma_A^2$
tmtB	not of interest		
A•B	$n_2 = n_n$	$V_2 = V_n$	$\sigma_2^2=\sigma^2+K\sigma_{AB}^2$
error	$n_1 = n_e$	$V_1 = V_e$	$\sigma_1^2=\sigma^2$

In this case we are interested in testing the hypothesis $H_o:\sigma_A^2 = 0$ against the hypothesis $H_a:\sigma_A^2 > 0$, which again is equivalent to testing the hypothesis $H_o:\sigma_3^2 = \sigma_2^2$ against the hypothesis $H_a:\sigma_3^2 > \sigma_2^2$. In this case, if $\sigma_{AB}^2 = 0$, then both V_2 and V_1 are estimates of σ^2. In this case the various procedures are conducted in an analogous manner to the case of the randomized complete block experiment.

The nature of the results for the sometimes pool procedure are quite different depending on which model we are employing. This is due to the fact that the auxiliary mean square that is pooled (V_1) when pooling is undertaken has a smaller expectation than the classical denominator mean square (V_2) in Model Type B, while the situation is reversed in Model Type A. As we have indicated we are going to work primarily with the random effects model which is traditionally referred to as model II. A little reflection shows that we have the Type A and B situations under both model I and model II. Therefore we have models I-A, I-B, II-A, and II-B, as well as various mixed models. Chapter 2 of this thesis is devoted to a discussion of the nature and behaviour of the sometimes pool procedure under models II-A and II-B.

1.4 RESULTS AND THESIS OUTLINE

In this thesis the behaviour of the proposed Smooth Pooling Procedure is investigated, as used with a random model where the sizes of the expected mean squares are ordered as in the case of the randomized complete block experiment (i.e. E(classical denominator MS) $\leq$ E(auxiliary MS)). The Smooth Pooling Procedure is attractive and should be considered as an alternative to the classical NPP or the SPP, but further work is necessary before a stronger recommendation could be made. Chapter 2 of this thesis provides a detailed description of the nature and behaviour of sometimes pooling procedures. In Chapter 3 we outline our proposed smooth procedure and compare it to the NPP and the SPP. Chapter 4 provides a summary of the thesis together with suggestions for further research.

CHAPTER 2

THE SOMETIMES POOL PROCEDURE

In Chapter 2 we will describe the behaviour of the sometimes pool procedure. Two bibliographies on the subject of the SPP under various scenarios have been published in the statistical literature (Bancroft and Han, 1977; Han, et al., 1988).

As discussed in Section 1.3 there are a number of situations to be considered. We have the fixed and random effects models as well as the mixed ones. In addition, under both the fixed and random models we have situations where the classical denominator mean square that is used in testing the main hypothesis of interest has an expectation that is greater than the expectation of the auxiliary mean square which may be pooled, and also situations where this ordering in the size of the expected mean squares is reversed.

In Section 2.1 we describe the nature of the SPP used with a nested model with random treatments and nuisance factor, or from a replicated two-way random effects layout. This is an example of the random effects model where the expected value of the classical denominator mean square for the main test of interest is greater than the expected value of the auxiliary mean square. This situation has been studied extensively (Bozivich, et al., 1956; Paull, 1950).

In Section 2.2 the behaviour of the SPP used with a randomized complete block design with random treatments and blocks is described. This is an example of the random effects model where the expected value of the classical denominator mean square for the main hypothesis of interest is smaller than the expected value of the auxiliary mean square. The behaviour of the SPP in such problems does not appear to have been investigated.

It is difficult to illustrate the behaviour of the SPP in a nice fashion. What we are attempting to do is display or describe the operating characteristic of the SPP, and to compare it to the operating characteristic of an appropriate NPP. In order to simplify the problem, for fixed degrees of freedom, we are going to study a SPP with preliminary test size α_p, and nominal size 0.05. Now given this SPP with fixed degrees of freedom, nominal size 0.05, and a preliminary test size α_p, to what should we compare its operating characteristic? A couple of obvious candidates are the NPP of size 0.05 and a NPP of the same size as the SPP. There are difficulties with the use of either of these as a means of judging the behaviour of the SPP. On the one hand researchers who have studied the SPP in the past have recognized that it is not fair to compare the SPP to the NPP of size 0.05. On the other hand it has been pointed out that direct comparisons of size α tests based on tests which do and which do not incorporate a preliminary test of significance are perhaps not appropriate, so in this context it is unfair to compare the SPP of size α_T to a NPP of the same size (Cohen, 1974). The size of the SPP, α_T, is the supremum of the Type I error probabilities (i.e., the supremum of the probability of rejecting the main of hypothesis of interest taken over all values of the nuisance parameter and using the null value of the main parameter).

The problem arises because we are trying to compare the NPP procedure which has a constant probability of type I error over all values of the nuisance parameter with the SPP procedure which has a probability of type I error that varies depending on the value of the nuisance parameter.

Because neither of these comparison procedures is entirely satisfactory the literature has considered a third possibility. This third method consists of the following steps. First the nuisance parameter, θ_{21}, is assumed known. Using this value of the nuisance parameter the "size" α_θ, is determined (i.e., the probability of rejecting the main hypothesis of interest at its null value and at the assumed nuisance parameter value). In this third method of comparison a NPP of size α_θ is used as a basis of comparison for the SPP (Paull, 1950).

Which of these three bases of comparison should we use? Since none is satisfactory to all we will use all three. That is we will compare the SPP of size α_T to Never Pool Procedures of sizes 0.05, α_T, and α_θ, at various values of the nuisance parameter.

2.1 THE SOMETIMES POOL PROCEDURE WITH A TYPE B RANDOM EFFECTS MODEL

In Section 2.1 we discuss the sometimes pool procedure when used with a random effects model, where the expected mean squares are ordered as E(classical denominator MS) $\geq$ E(auxiliary MS). In order to see how this could arise consider Table 2.1 which is a reproduction of Table 1.6.

Table 2.1 arises from a replicated two-way random effects model. Using the standard notation the model is

$$y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} ,$$
$$i=1,..,I \qquad j=1,..,J \qquad k=1,..,K \ ,$$
$$A_i \sim \text{iid } N(0,\sigma_a^2) \ ,$$
$$B_j \sim \text{iid } N(0,\sigma_b^2) \ ,$$
$$AB_{ij} \sim \text{iid } N(0,\sigma_{ab}^2) \ ,$$
$$\varepsilon_{ijk} \sim \text{iid } N(0,\sigma^2) \ ,$$

where the model random variables are all mutually independent.

Table 2.1 An example of a Type B Random Effects Model arising from a replicated two-way experiment.

Source	**df**	**MS**	**E(MS)**
tmtA	$n_3=n_t$	$V_3=V_t$	$\sigma_3^2=\sigma^2+K\sigma_{AB}^2+JK\sigma_A^2$
tmtB	not of interest		
A • B	$n_2=n_n$	$V_2=V_n$	$\sigma_2^2=\sigma^2+K\sigma_{AB}^2$
error	$n_1=n_e$	$V_1=V_e$	$\sigma_1^2=\sigma^2$

Our main interest is in testing whether there is variation due to factor A. In addition there is some doubt as to whether the AB interaction is present and we have decided that we will utilize a SPP in order to accomplish our goal. In this situation then, V_t is our treatment mean square, V_n is the classical denominator mean square for the main test of interest, and V_e is the auxiliary mean square. In statistical terms we want to test the hypothesis $H_o:\sigma_A^2 = 0$ against the hypothesis $H_a:\sigma_A^2 > 0$ which is equivalent to testing the hypothesis

$H_o:\sigma_3^2 = \sigma_2^2$ against the hypothesis $H_a:\sigma_3^2 > \sigma_2^2$. The SPP can be stated in the following form.

1. Test H_o': $\sigma_2^2 = \sigma_1^2$ against the hypothesis H_a': $\sigma_2^2 > \sigma_1^2$ at level α_p.

2. If H_o' is rejected then use the statistic

$$F = \frac{V_t}{V_n} \sim F(n_t, n_n)$$

to test H_o at level α.

3. If H_o' is accepted then use the statistic

$$F_p = \frac{(n_e + n_n)V_t}{n_e V_e + n_n V_n} \sim F(n_t, n_e + n_n)$$

to make the test of H_o at level α.

We note that a similar type of model arises from a nested experiment. Table 2.2 illustrates this situation.

Table 2.2 An example of a Type B Random Effects Model arising from a nested experiment.

Source	df	MS	E(MS)
A	$n_3 = n_t$	$V_3 = V_t$	$\sigma_3^2 = \sigma^2 + K\sigma_{AB}^2 + JK\sigma_A^2$
B under A	$n_2 = n_n$	$V_2 = V_n$	$\sigma_2^2 = \sigma^2 + K\sigma_{AB}^2$
error	$n_1 = n_e$	$V_1 = V_e$	$\sigma_1^2 = \sigma^2$

To summarize we have an ANOVA table of the same form as Table 2.3 and we note that in this case $\sigma_3^2 \geq \sigma_2^2 \geq \sigma_1^2$.

Table 2.3 Generalized Type B Random Effects Model.

Source	df	MS	E(MS)
treatments	$n_3=n_t$	$V_3=V_t$	σ_3^2
nuisance	$n_2=n_n$	$V_2=V_n$	σ_2^2
expt. error	$n_1=n_e$	$V_1=V_e$	σ_1^2

The source labelled nuisance results from a factor which possibly, but not necessarily, introduces variability into the experimental results. If we knew definitely that it was a source of variability then the classical never pool test in which $F = V_t/V_n$ is compared to $F(\alpha;n_t,n_n)$ would be used to ascertain whether there is variability due to treatments. If we were certain that the nuisance source did not introduce extra variability then we would use $F_p = (n_e+n_n)V_t/(n_eV_e+n_nV_n)$ to make the main test. We are going to use a sometimes pool procedure in which we make a preliminary test to see what conclusion we can make about the nuisance variability. If our preliminary test is significant, at some agreed upon level, then we will conclude that not pooling is the right action. If, on the other hand, our preliminary test is not significant, we will pool. That is, denoting the F statistic for the preliminary test by F_n and the F statistic for the main test by F_t, we will reject the main hypothesis of interest (that there is no variability due to treatments) if

(1) $F_n = V_n/V_e > F(\alpha_p;n_n,n_e)$ and $F_t = V_t/V_n > F(\alpha_1;n_t,n_n)$

or if

(2) $F_n = V_n/V_e \leq F(\alpha_p;n_n,n_e)$ and $F_t = V_t/V_n > \dfrac{F(\alpha_2;n_t,n_n+n_e)(n_n+n_eF_n)}{F_n(n_n+n_e)}$.

The conditions in (2) are equivalent to

$$F_n = V_n/V_e \leq F(\alpha_p;n_n,n_e) \text{ and } F_p = \frac{(n_e+n_n)V_t}{n_eV_e+n_nV_n} \geq F(\alpha_2;n_t,n_n+n_e).$$

Early work on the SPP for this model considered that eight parameters were involved and discussed the problem from this point of view (Bozivich, et al., 1956). These eight parameters are: n_e, the degrees of freedom for experimental error; n_n, the degrees of freedom for the nuisance parameter; n_t, the treatment degrees of freedom; α_p, the level of the preliminary test; α_1, and α_2, the nominal levels of the main test under the cases of rejecting and not rejecting the preliminary hypothesis; $\theta_{21} = \sigma_2^2/\sigma_1^2$, the ratio of the nuisance factor variance to the error variance (the nuisance parameter); and $\theta_{32} = \sigma_3^2/\sigma_2^2$, the ratio of the treatment factor variance to the nuisance factor variance (the main parameter). From our studies it would appear that the general behaviour of a SPP is similar irrespective of the relative magnitudes of the various degrees of freedom so we will with the exception of a few brief comments not treat them as parameters in our investigation. In addition, we will fix our interest on main tests with size near 0.05. Since tests with size in this vicinity can be achieved by choosing α_1 and α_2 both to be 0.05 (Bozivich, et al., 1956) we can thus eliminate two more parameters from our investigations. We are thus left to discuss the behaviour of the SPP in terms of the preliminary test size α_p --- which is under the control of the researcher, and the nuisance parameter θ_{21} --- which is an unknown quantity.

It has been shown (Paull, 1950) that there are two main cases to be considered, and these are referred to as Class A and Class B tests. The classification of a test depends on the chosen size of the preliminary test. For a given problem there is a value of the preliminary test size, referred to in the literature as the borderline value of the preliminary test size (denoted $\bar{\alpha}_p$), that separate Class A and B tests. Tests with $\alpha_p < \bar{\alpha}_p$ are the ones commonly encountered in practice and these are called Class A tests. Tests conducted at $\alpha_p = \bar{\alpha}_p$ are called borderline tests in the literature. Class B tests are tests with $\alpha_p > \bar{\alpha}_p$.

It can be shown that the borderline value of the preliminary test size, $\bar{\alpha}_p$, is the value of α_p which makes the SPP a continuous procedure as a function of the block p-value (although not smooth).

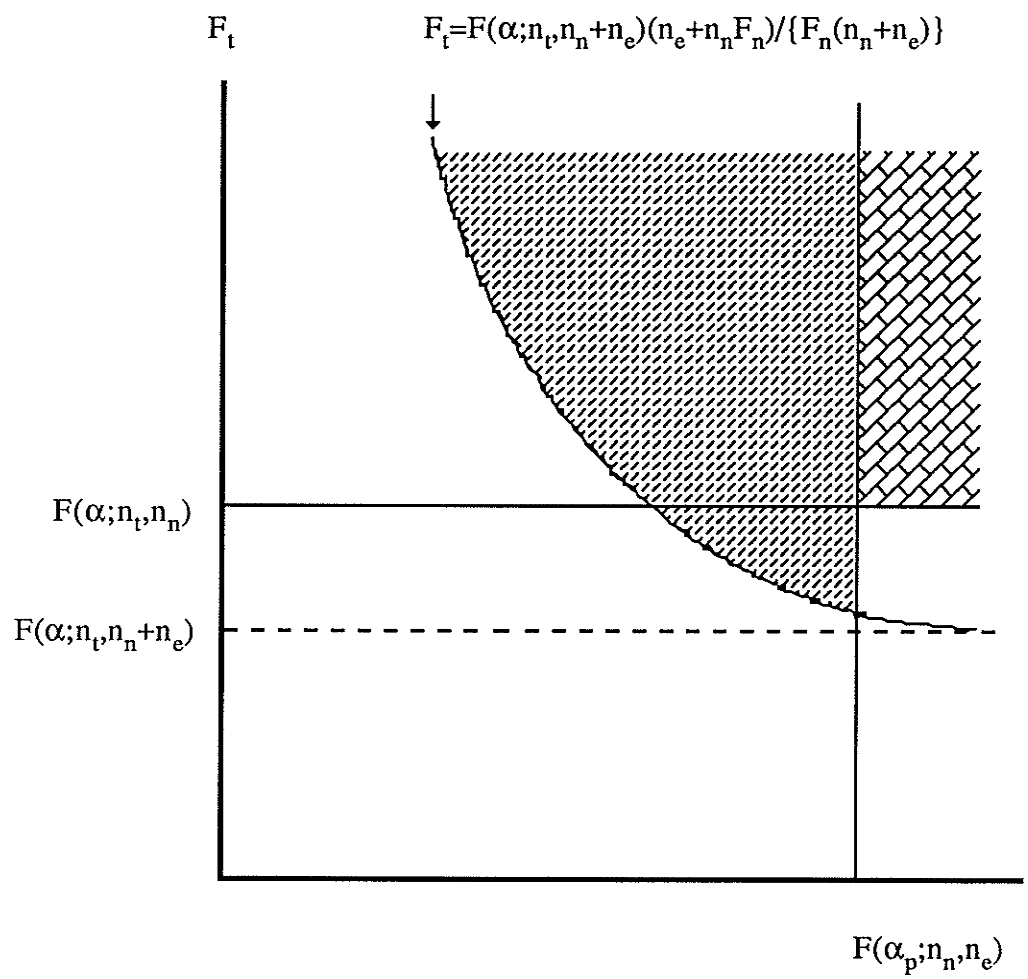
Figures 2.1, 2.2, and 2.3 are illustrations of the rejection regions of Class A, Class B, and borderline tests.

Based on extensive investigations over a wide range of parameter values (the eight parameters we have mentioned) the literature (Bozivich, et al., 1956) has made the following recommendations for most sets of degrees of freedom commonly encountered.

(1) If the research worker is reasonably sure that only small values of the nuisance parameter θ_{21} can be envisioned then use of the Class A procedure with $\alpha_p = 0.25$ is recommended.

(2) If the research worker is not willing to make the assumption that θ_{21} is small then the literature recommends the use of $\alpha_p = \bar{\alpha}_p$. In this case

Figure 2.1 Rejection Region of a Class A SPP ($\alpha_p < \bar{\alpha}_p$) in a Type B Random Effects Experiment.



1) the shaded area in total is the rejection region of the SPP

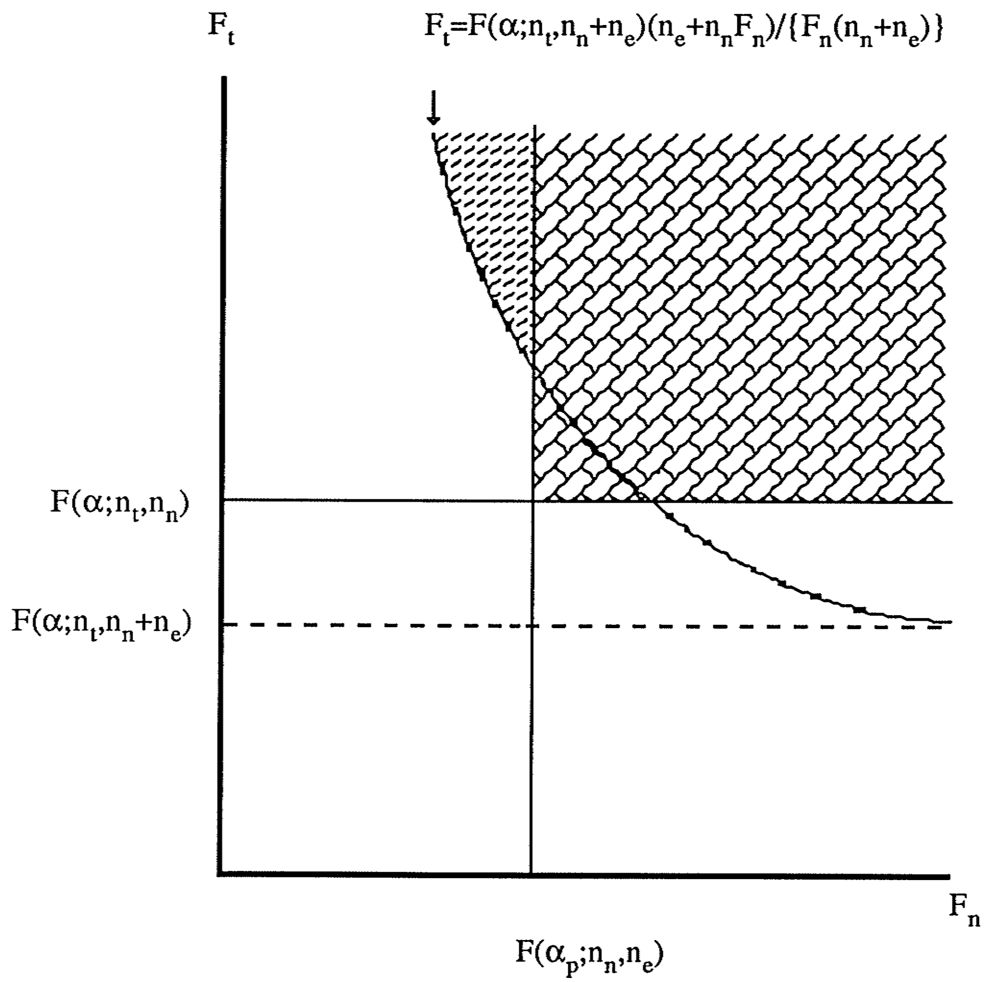
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the curved line is the portion of the SPP rejection region due to pooling

3) the shaded area to the right of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_n)$ is the portion of the SPP rejection region due to not pooling

4) the area above the solid horizontal at $F(\alpha;n_t,n_n)$ is the NPP rejection region

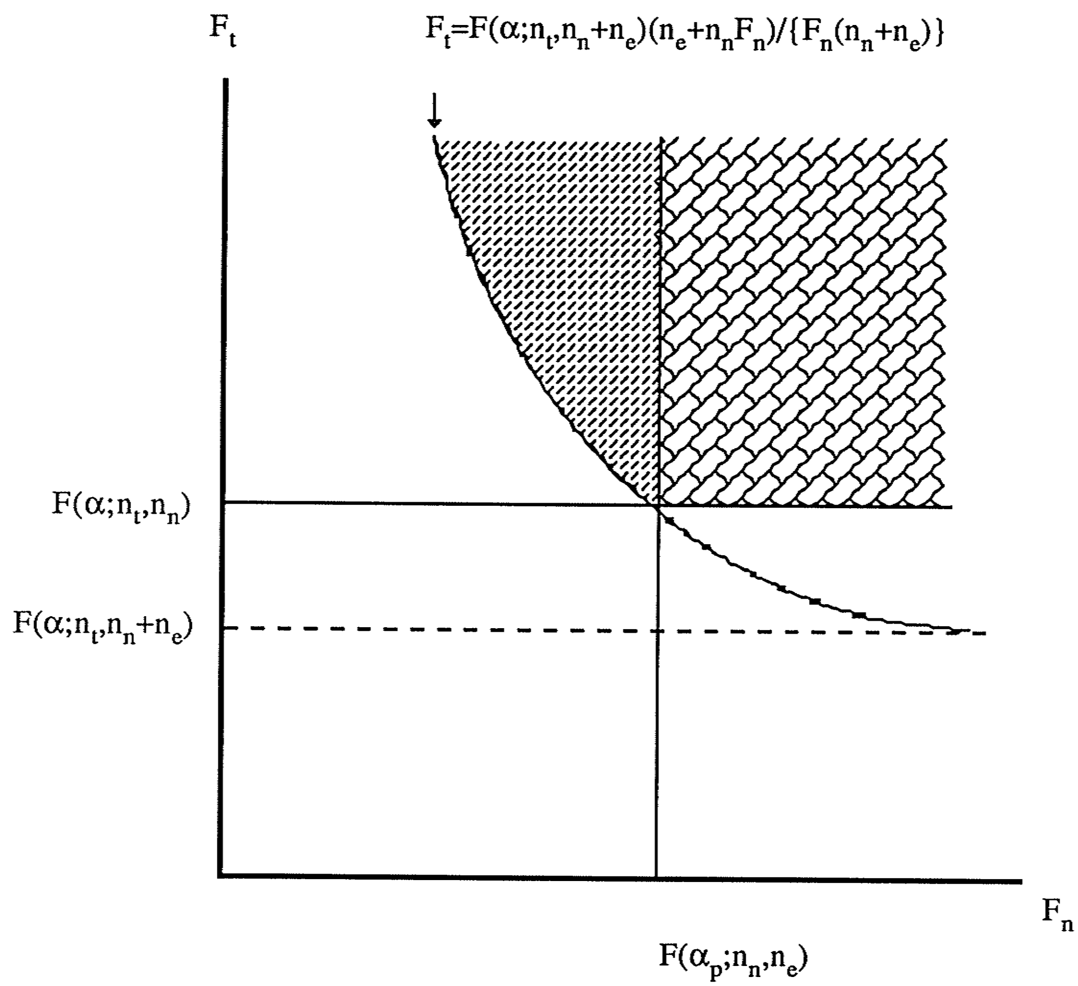
5) the area above the curved line is the APP rejection region in terms of F_t

Figure 2.2 Rejection Region of a Class B SPP ($\alpha_p > \bar{\alpha}_p$) in a Type B Random Effects Experiment.



1) the shaded area in total is the rejection region of the SPP
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the curved line is the portion of the SPP rejection region due to pooling
3) the shaded area to the right of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_n)$ is the portion of the SPP rejection region due to not pooling
4) the area above the solid horizontal at $F(\alpha;n_t,n_n)$ is the NPP rejection region
5) the area above the curved line is the APP rejection region in terms of F_t

Figure 2.3 Rejection Region of a Borderline SPP ($\alpha_p = \bar{\alpha}_p$) in a Type B Random Effects Experiment.



1) the shaded area in total is the rejection region of the SPP
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the curved line is the portion of the SPP rejection region due to pooling
3) the shaded area to the right of the vertical line at $F(\alpha_p;n_n,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_n)$ is the portion of the SPP rejection region due to not pooling
4) the area above the solid horizontal at $F(\alpha;n_t,n_n)$ is the NPP rejection region
5) the area above the curved line is the APP rejection region in terms of F_t

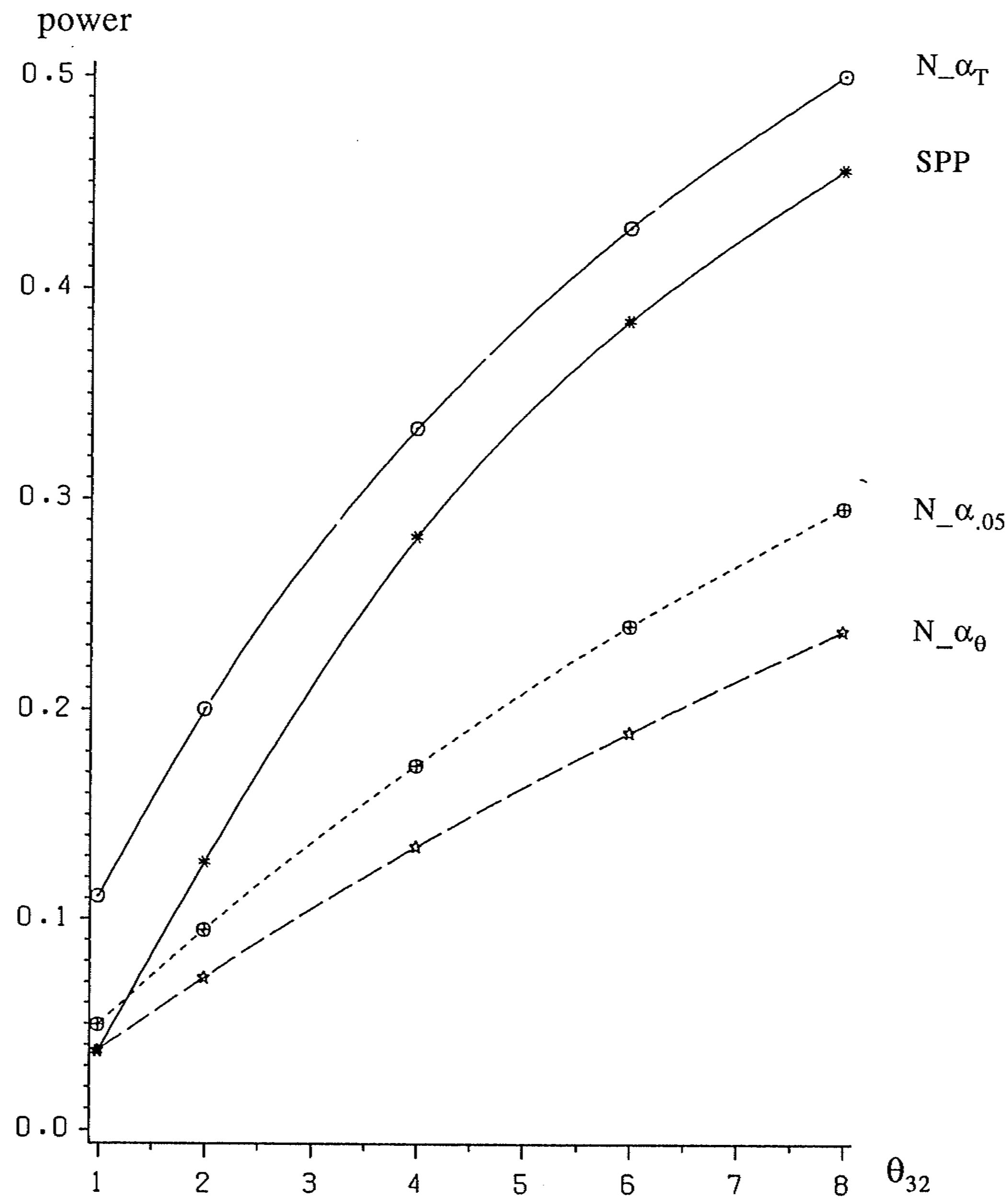
however the literature cautions that the research worker may wish to use the never pool procedure.

The reasons underlying these recommendations involve controlling the disturbance in the size of the procedure while optimizing the gain in power over the NPP for a broad range of degrees of freedom.

Figures 2.4 - 2.8 help us to visualize the general nature of Class A tests as well as their behaviour at the value of the preliminary test size recommended in the literature, if it is assumed that the nuisance parameter is small. In the following observations we denote the true size of the SPP by α_T, and the value of the "size" considered in the literature for a fixed nuisance parameter value by α_θ. Regardless of the "sizes" of two compared procedures we will say that a procedure is uniformly more (less) powerful at a fixed value of the nuisance parameter if it is more (less) powerful at all alternative values of the main parameter. In the figures throughout the rest of the thesis the labels N_α_T, N_$\alpha_{.05}$, N_α_θ, N_$\alpha_{T=.05}$, N_$\alpha_{T\theta}$, and SPP, denote, respectively, NPP's of sizes α_T, $\alpha_{.05}$, α_θ, $\alpha_{T=.05}$, $\alpha_{T\theta}$ and the SPP under discussion.

1. When $\theta_{21} = 1$ (i.e., there is no variability due to the nuisance factor) the power of the SPP at $\theta_{32} = 1$ is less than the nominal level of 0.05 (i.e., α_θ is less than 0.05). The SPP is less powerful than the NPP of size 0.05 when θ_{32} is very close to 1 but more powerful everywhere else. The SPP is uniformly more powerful than a NPP of size α_θ, but is uniformly less powerful than a NPP of size α_T. See Figure 2.4.

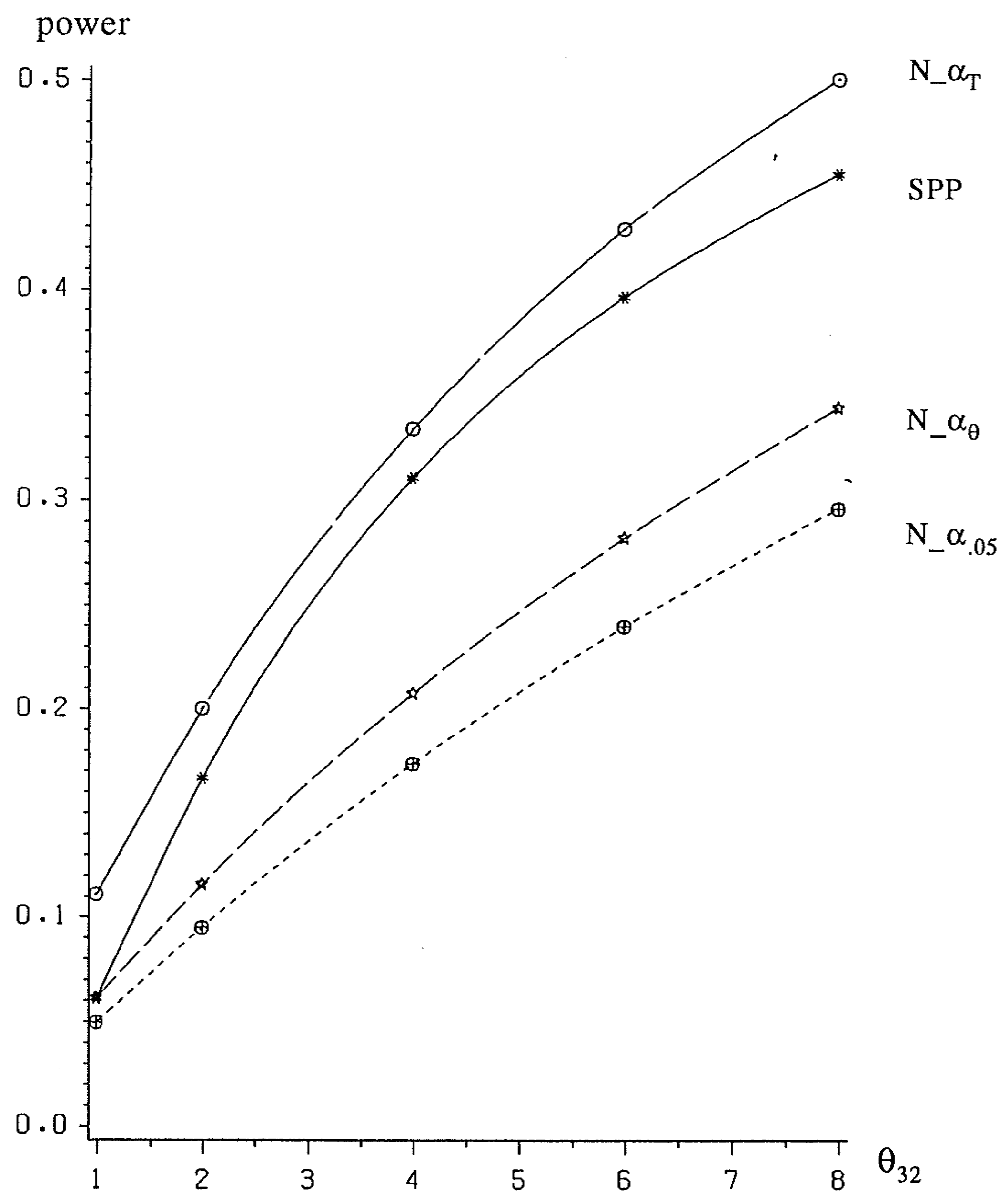
Figure 2.4 The Power Curve of a Class A SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.



$\alpha_p = 0.25$

$n_t = 2 \;\; n_n = 2 \;\; n_e = 6$

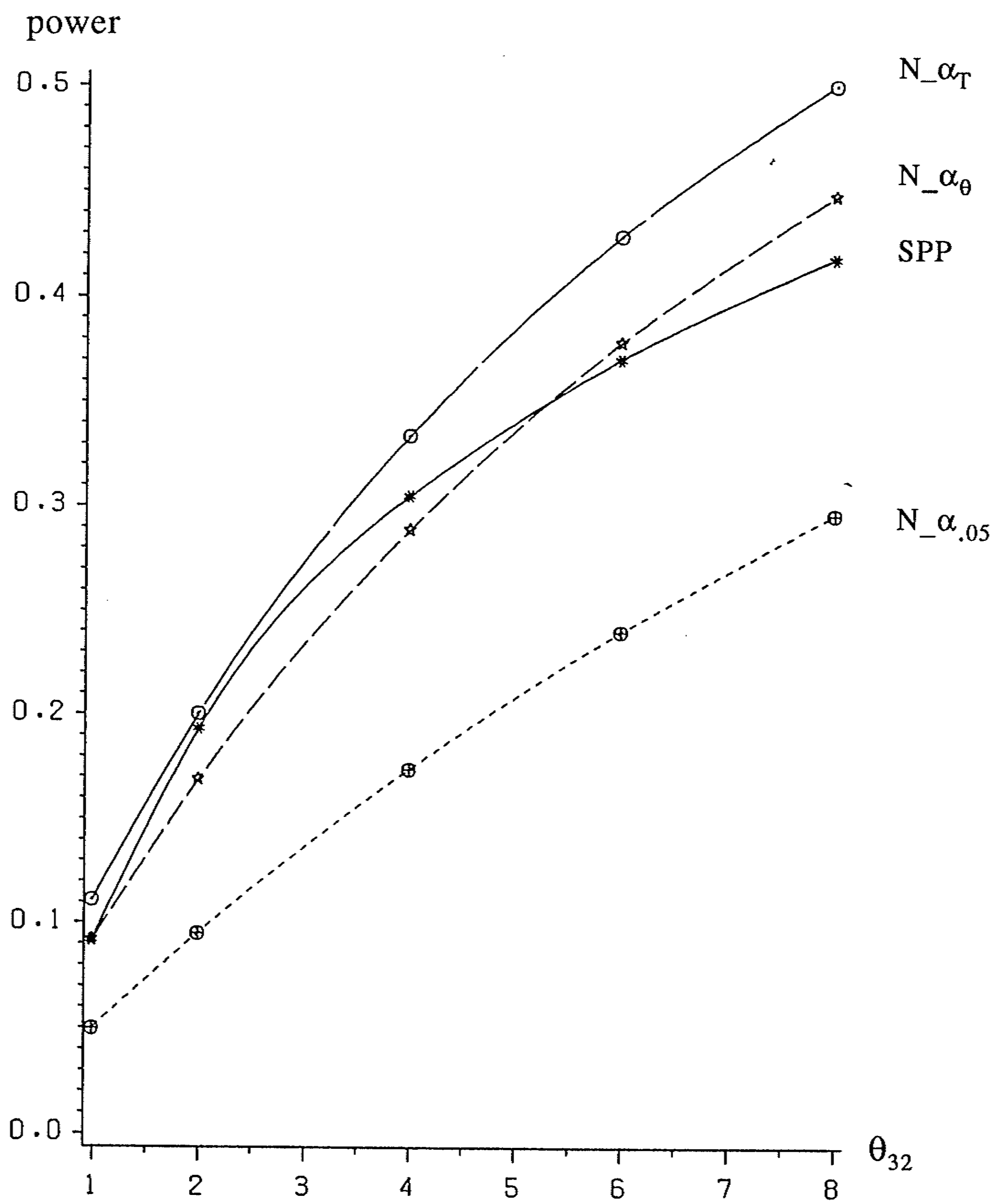
Figure 2.5 The Power Curve of a Class A SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 1.5$.



$\alpha_p = 0.25$

$n_t = 2 \;\; n_n = 2 \;\; n_e = 6$

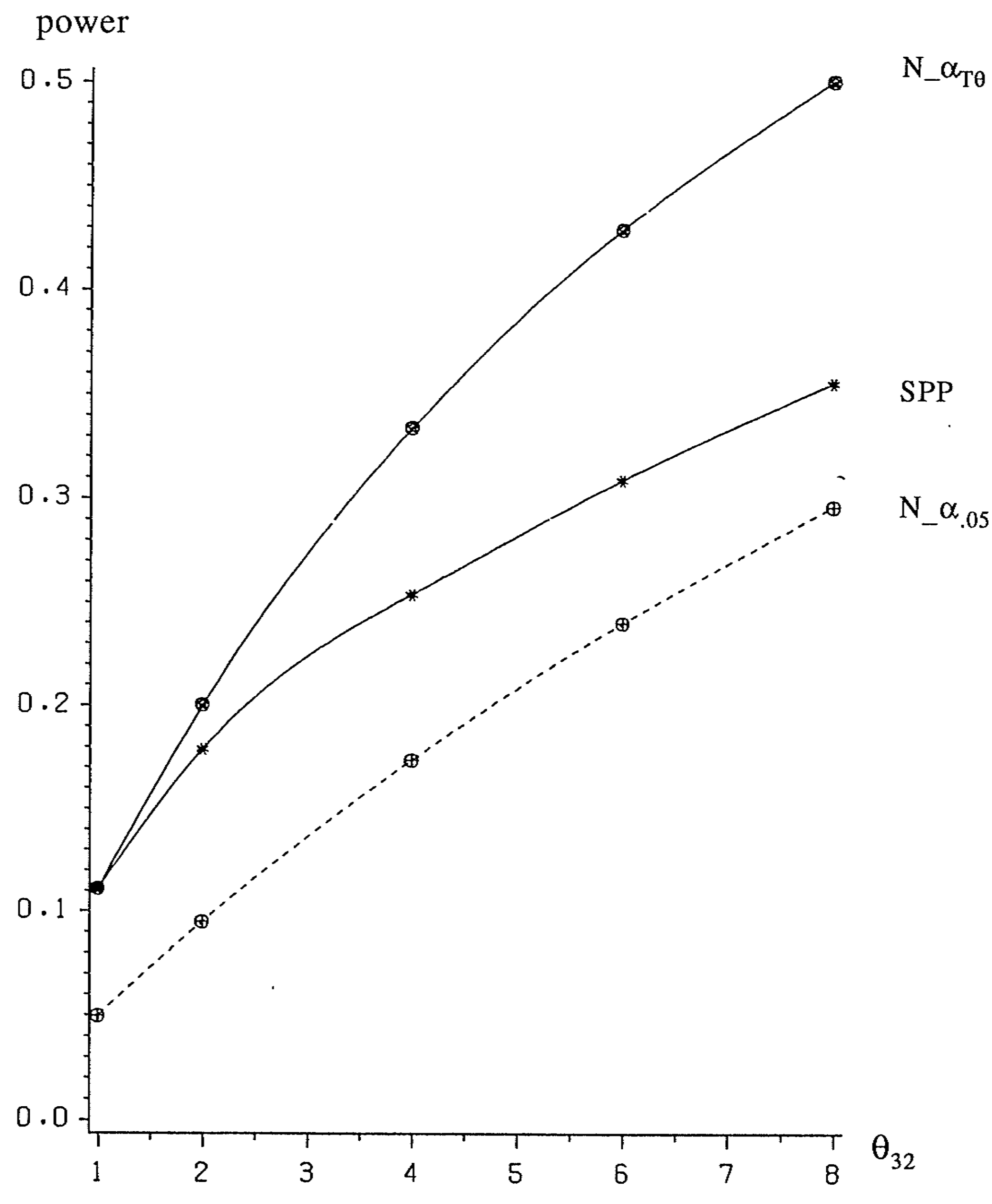
Figure 2.6 The Power Curve of a Class A SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 2.5$.



$\alpha_p = 0.25$

$n_t = 2$ $n_n = 2$ $n_e = 6$

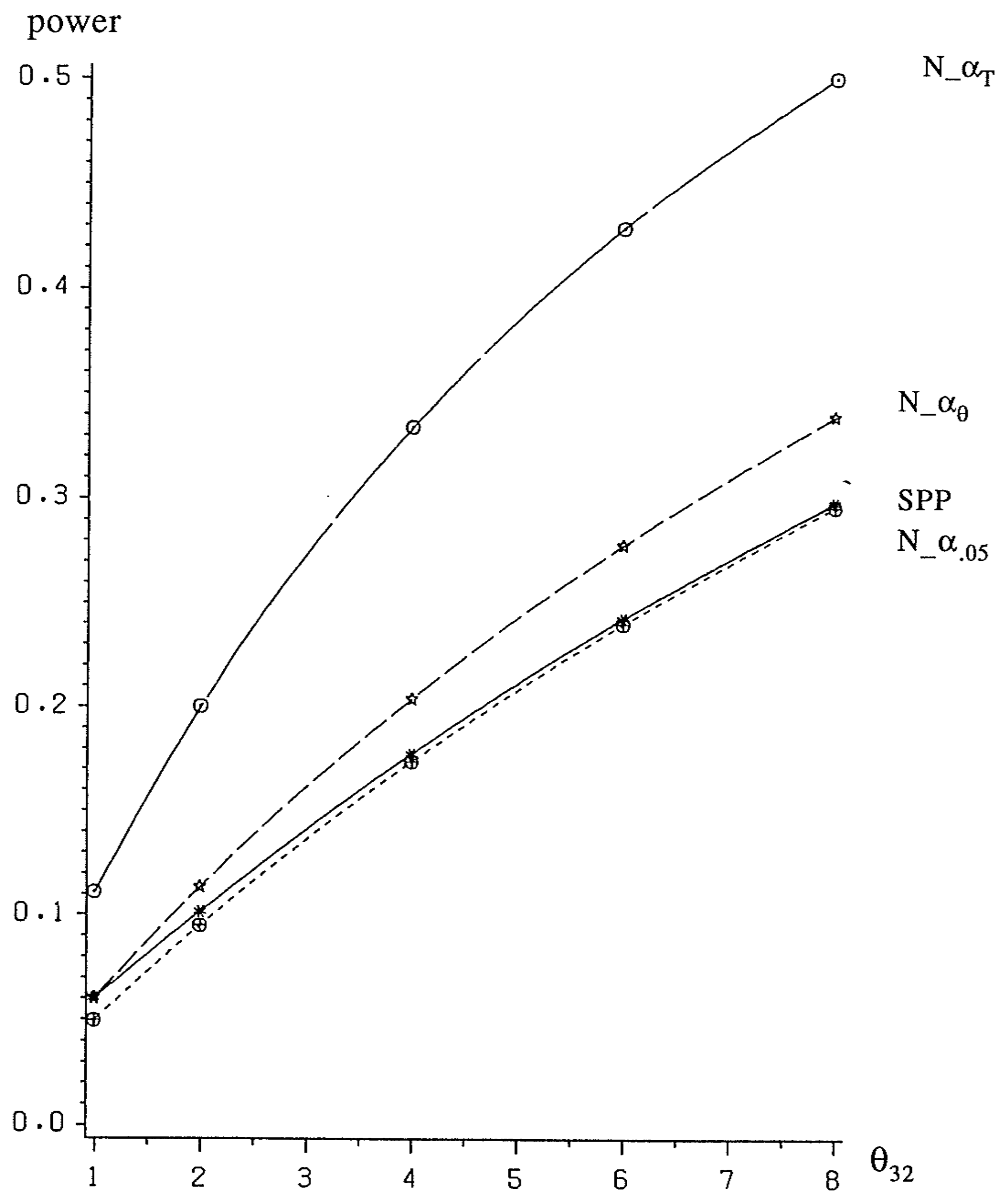
Figure 2.7 The Power Curve of a Class A SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 5.11$.



$\alpha_p = 0.25$

$n_t = 2 \; n_n = 2 \; n_e = 6$

Figure 2.8 The Power Curve of a Class A SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 40$.



$\alpha_p = 0.25$

$n_t = 2 \;\; n_n = 2 \;\; n_e = 6$

2. For small values of θ_{21} the "size", α_θ, is greater than the nominal level 0.05. Here the SPP is uniformly more powerful than both the adjusted NPP of size α_θ and the NPP of size 0.05. The SPP is uniformly less powerful than a NPP of size α_T. See Figure 2.5.

3. For intermediate values of θ_{21} the power functions of the SPP and the NPP of the same "size", α_θ, cross. That is, close to the null hypothesis value of the main test, $\theta_{32} = 1$, the SPP is more powerful than a NPP at level α_θ, but as θ_{32} increases from 1 these power functions cross and the NPP becomes more powerful. It is not surprising that, since α_θ is greater than 0.05, the SPP and the NPP at level α_θ, are uniformly more powerful than a NPP of size 0.05. The NPP of size 0.05, the SPP, and the NPP of size α_θ all are uniformly less powerful than a NPP of size α_T. See Figure 2.6.

4. For large values of θ_{21} the SPP is uniformly less powerful than the NPP of size α_θ. The "size", α_θ, is always greater than 0.05 although as θ_{21} approaches infinity α_θ approaches 0.05. The the SPP has "size" $\alpha_\theta > 0.05$ and it is uniformly more powerful than the NPP of size 0.05. In this region α_θ achieves its maximum value α_T. See Figures 2.7 and 2.8.

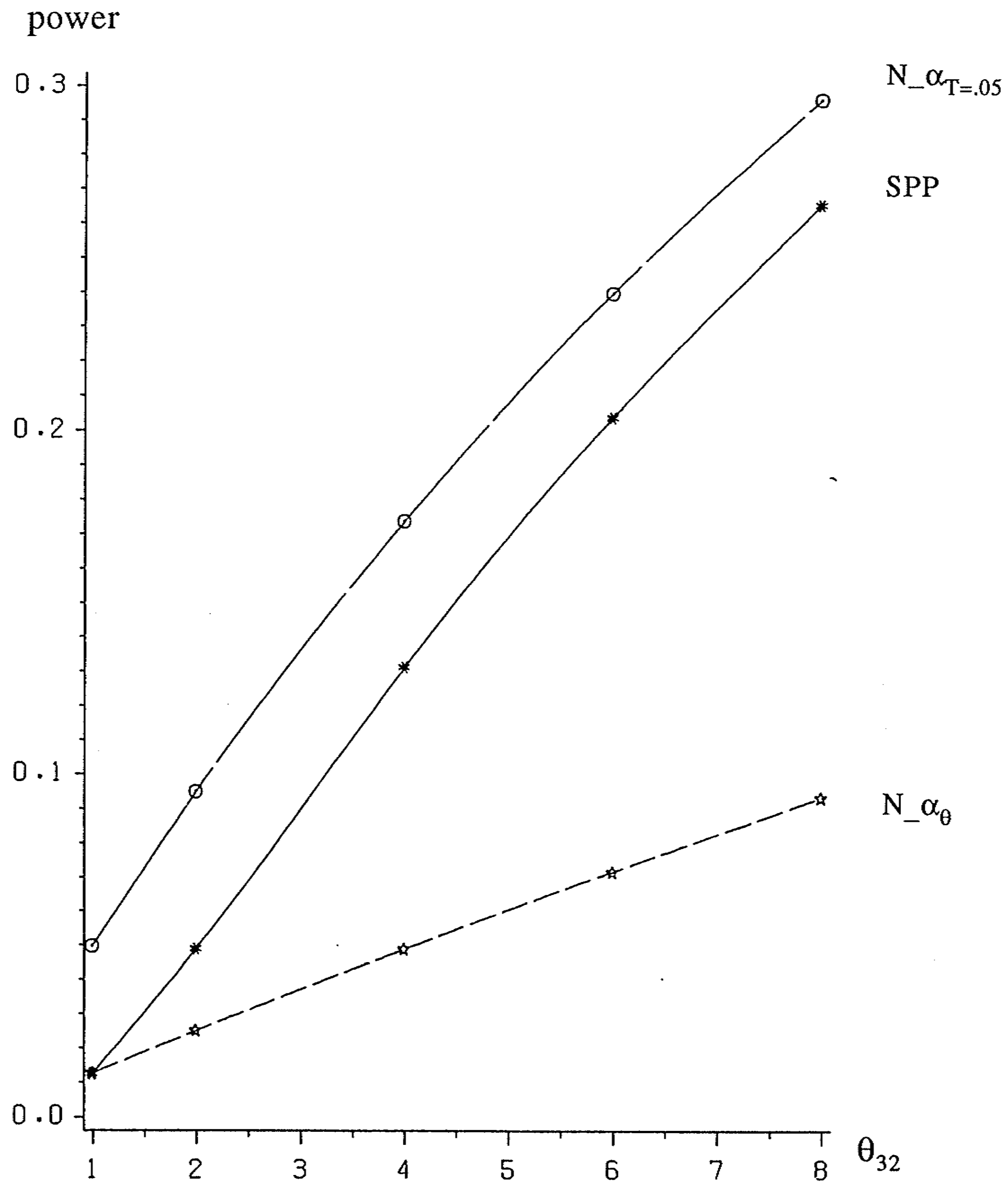
To summarize, while using the SPP recommended in the literature (under the assumption that the nuisance parameter is small) we observe the following:

- for small values of the nuisance parameter the SPP is uniformly more powerful than a NPP of size α_θ;

- for intermediate values of the nuisance parameter the SPP is more powerful than a NPP of size α_θ at small values of the main parameter and less powerful than a NPP of size α_θ at large values of the main parameter;
- for large values of the nuisance parameter the SPP is uniformly less powerful than a NPP of size α_θ;
- the SPP is more powerful than a NPP of size 0.05 except when both the main and nuisance parameter are very close to 1;
- the SPP is uniformly less powerful than a NPP of size α_T;
- the values of the nuisance parameter that constitute small, medium, and large are dependent on the degrees of freedom.

We now consider the procedure recommended when the research worker is unwilling to assume that the nuisance parameter is small (i.e., the borderline SPP). The size of the SPP, α_T, performed with a preliminary test size of $\bar{\alpha}_p$ is 0.05. Comparisons between the SPP for a given nuisance parameter value and a NPP of size $\alpha_T = 0.05$ are made using Figures 2.9 and 2.10. Using Figures 2.9 and 2.10 we see that a SPP is always more powerful than a NPP of size α_θ, but less powerful than a NPP of size $\alpha_T = 0.05$. We also see that, as the nuisance parameter increases in size, the power function of the SPP converges to the power function of the NPP of size 0.05.

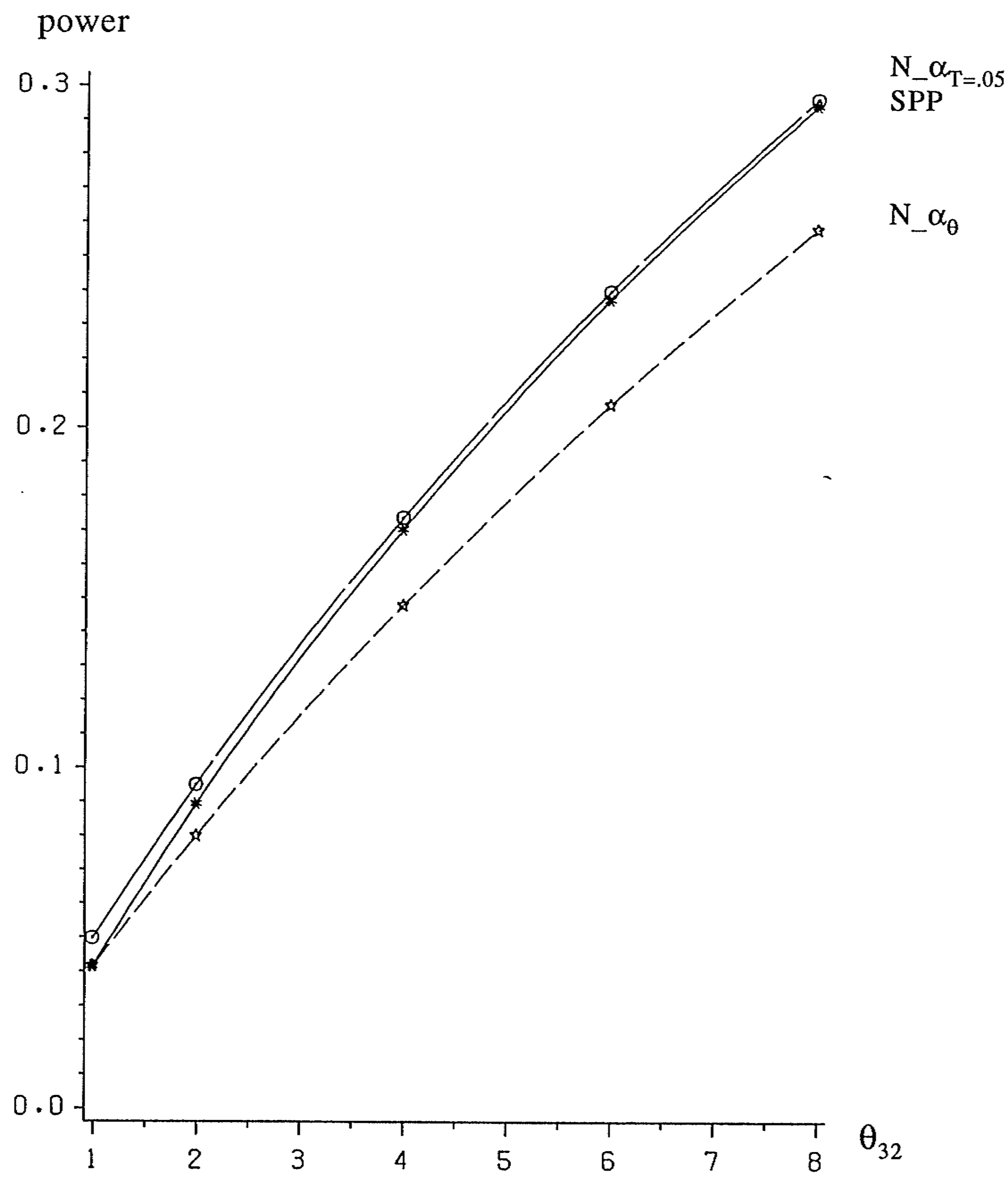
Figure 2.9 The Power Curve of a Borderline SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.



$\alpha_p = \bar{\alpha}_p$

$n_t = 2 \;\; n_n = 2 \;\; n_e = 6$

Figure 2.10 The Power Curve of a Borderline SPP in a Type B Random Effects Model when the Nuisance Parameter $\theta_{21} = 5$.



$\alpha_p = \bar{\alpha}_p$

$n_t = 2 \; n_n = 2 \; n_e = 6$

With Class B tests the "size" of the SPP, α_θ, is always less than 0.05 and greater than the corresponding α_θ value, which is achieved when the borderline value of the preliminary test size is used. The "size", α_θ, approaches $\alpha_T = 0.05$ as θ_{21} approaches infinity. The SPP is always more powerful than a NPP of size α_θ but less powerful than a NPP of size $\alpha_T = 0.05$. These differences are most pronounced for small values of θ_{21} and as θ_{21} approaches infinity the three power functions converge. Finally, the gain in power achieved by the SPP over the NPP of size α_θ is less than the gain in power achieved by the SPP over the NPP at the borderline value all other conditions remaining constant. A discussion of the use of the Class B procedure can be found in the literature (Paull, 1950).

2.2 THE SOMETIMES POOL PROCEDURE WITH A TYPE A RANDOM EFFECTS MODEL

In this section we illustrate the nature of a SPP when used with a random model where the expected mean squares are ordered such that the mean square that is pooled if pooling takes place (i.e., the auxiliary mean square) has an expectation that is greater than the classical denominator mean square (i.e., E(classical denominator mean square) $\leq$ E(auxiliary mean square)).

As an example of how this can occur consider Table 2.4 which is a reproduction of Table 1.5.

Table 2.4 results from a randomized complete block experiment in which the treatments and blocks are considered to be random effects. The assumed model for the data is $y_{ij} = \mu + TMT_i + BLK_j + \varepsilon_{ij}$ where

y_{ij}	is the observation on the experimental unit which receives the ith treatment in the jth block,
μ	is the overall mean,
TMT_i	is the random effect of treatment i, $\sim N(0,\sigma_t^2)$,
BLK_j	is the random effect of block j, $\sim N(0,\sigma_b^2)$,
ε_{ij}	is a random error, $\sim N(0,\sigma^2)$,

and further TMT_i, BLK_j, and ε_{ij} are assumed independent.

Table 2.4 An example of a Type A Random Effects Model arising from a randomized complete block experiment.

Source of variation	df	MS	E(MS)
treatment	$n_3=n_t$	$V_3=V_t$	$\sigma_3^2=\sigma^2+n_t\sigma_t^2$
block	$n_1=n_b$	$V_1=V_b$	$\sigma_1^2=\sigma^2+n_b\sigma_b^2$
error	$n_2=n_e$	$V_2=V_e$	$\sigma_2^2=\sigma^2$

Our main interest is in whether there is variability due to treatments and in order to consider the possibility that $\sigma_b^2 = 0$ we have decided to use a SPP.

In this situation, then, V_t is our treatment mean square, V_e is the classical denominator mean square for the main test of interest, and V_b is the auxiliary mean square. We want to test the hypothesis $H_o:\sigma_t^2 = 0$ against the hypothesis $H_a:\sigma_t^2 > 0$, which is equivalent to testing the hypothesis $H_o:\sigma_3^2 = \sigma_2^2$ against the hypothesis $H_a:\sigma_3^2 > \sigma_2^2$. The SPP consists of the following steps.

1. Test H_o': $\sigma_1^2 = \sigma_2^2$ against the hypothesis H_a': $\sigma_1^2 > \sigma_2^2$ at level α_p.

2. If H_o' is rejected then use the statistic

$$F = \frac{V_t}{V_e} \sim F(n_t, n_e)$$

to test H_o at nominal level α.

3. If H_o' is accepted then use the statistic

$$F_p = \frac{(n_e+n_b)V_t}{n_e V_e + n_b V_b} \sim F(n_t, n_e+n_b)$$

to make the test of H_o at nominal level α.

Denoting the F statistic we use for the preliminary test by F_b, and the F statistic we use to test the main hypothesis by F_t, we will reject the main hypothesis of interest, $H_o:\sigma_t^2 = 0$, if

(1) $F_b = V_b/V_e > F(\alpha_p;n_b,n_e)$ and $F_t = V_t/V_e > F(\alpha;n_t,n_e)$

or if

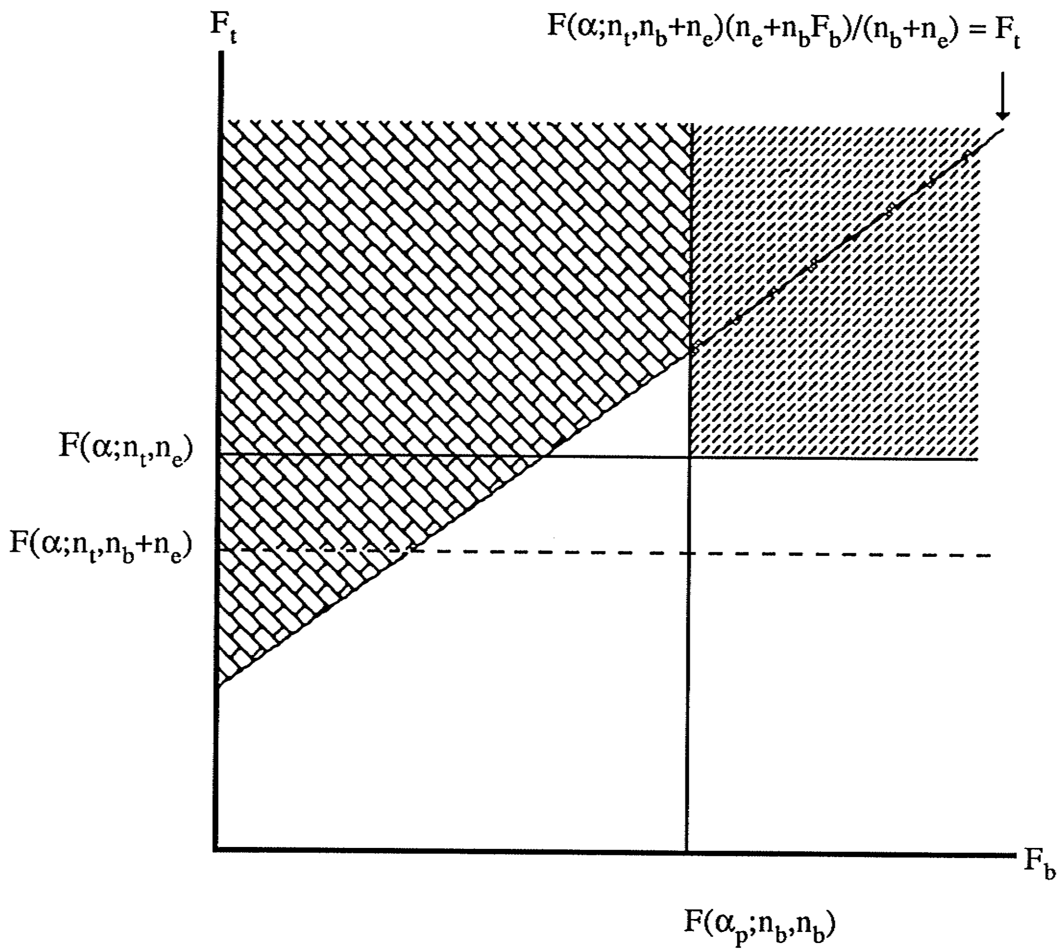
(2) $F_b = V_b/V_e \leq F(\alpha_p;n_b,n_e)$ and $F_t = V_t/V_e > \dfrac{F(\alpha;n_t,n_b+n_e)(n_e+n_bF_b)}{(n_b+n_e)}$,

which is equivalent to

$$F_b = V_b/V_e \leq F(\alpha_p;n_b,n_e) \text{ and } F_p = \frac{(n_e+n_b)V_t}{n_e V_e + n_b V_b} \geq F(\alpha;n_t,n_b+n_e).$$

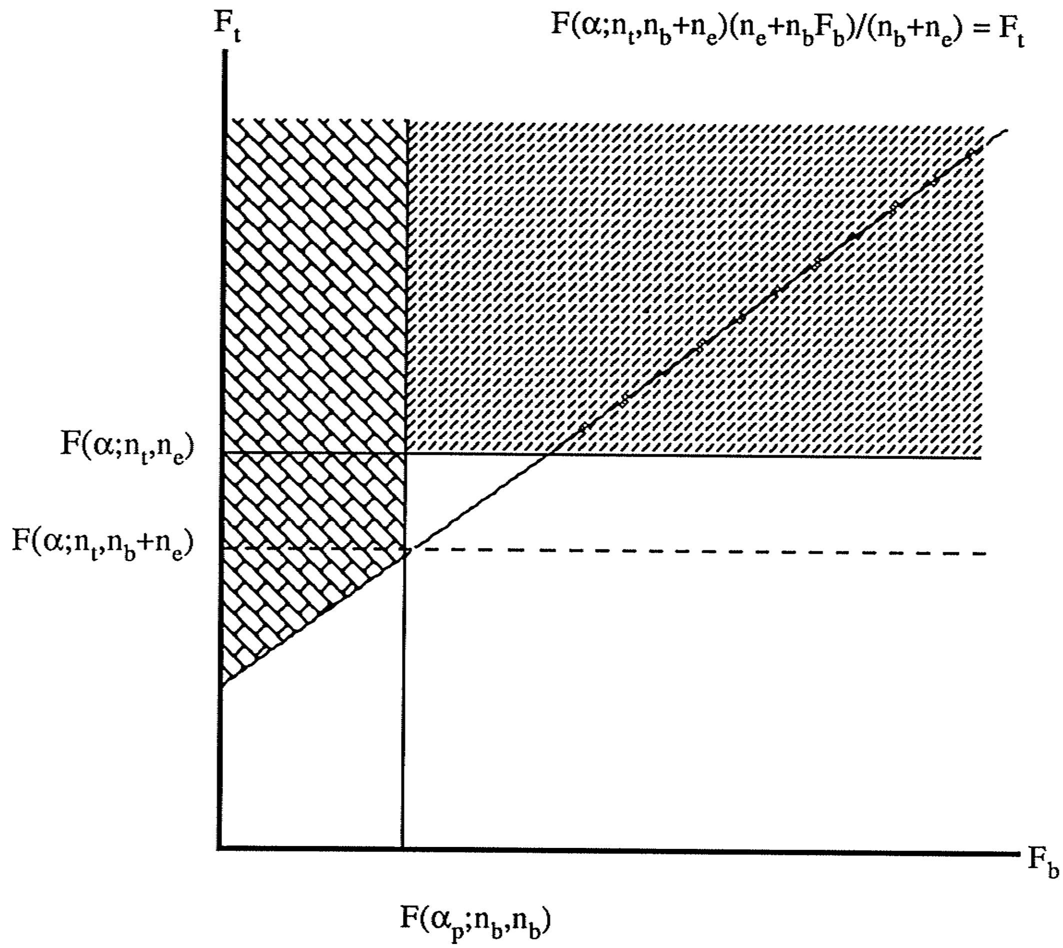
Here as well we have Class A tests, Class B tests, and the tests at the borderline value of the preliminary test size $\alpha_p = \bar{\alpha}_p$ that separates Class A tests from Class B tests. Figures 2.11 -2.13 illustrate the SPP in terms of its rejection region for Class A, B, and borderline tests.

Figure 2.11 Rejection Region of a Class A SPP ($\alpha_p < \bar{\alpha}_p$) in a Type A Random Effects Experiment.



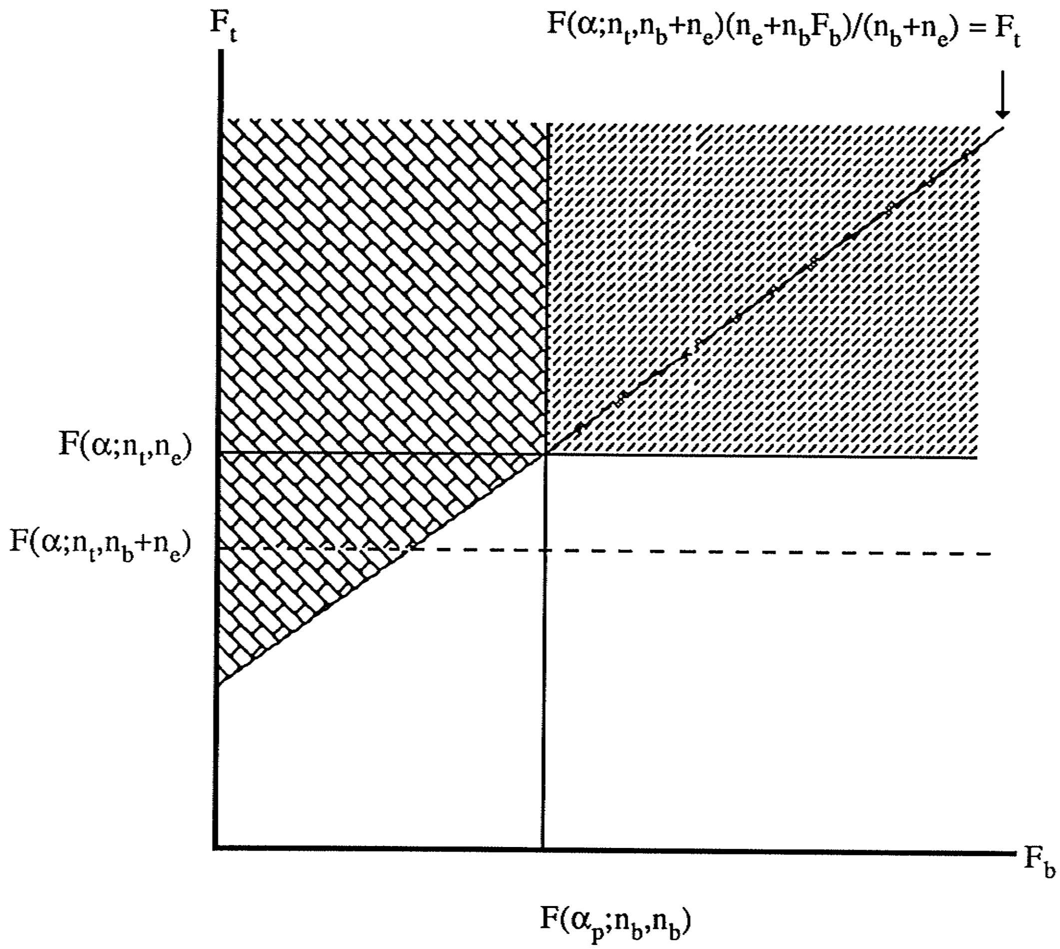
1) the shaded area in total is the rejection region of the SPP
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the diagonal line is the portion of the SPP rejection region due to pooling
3) the shaded area to the right of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_e)$ is the portion of the SPP rejection region due to not pooling
4) the area above the solid horizontal at $F(\alpha;n_t,n_e)$ is the NPP rejection region
5) the area above the diagonal line is the APP rejection region in terms of F_t

Figure 2.12 Rejection Region of a Class B SPP ($\alpha_p > \bar{\alpha}_p$) in a Type A Random Effects Experiment.



1) the shaded area in total is the rejection region of the SPP
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the diagonal line is the portion of the SPP rejection region due to pooling
3) the shaded area to the right of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_e)$ is the portion of the SPP rejection region due to not pooling
4) the area above the solid horizontal at $F(\alpha;n_t,n_e)$ is the NPP rejection region
5) the area above the diagonal line is the APP rejection region in terms of F_t

Figure 2.13 Rejection Region of a Borderline SPP ($\alpha_p = \bar{\alpha}_p$) in a Type A Random Effects Experiment.



1) the shaded area in total is the rejection region of the SPP
2) the shaded area to the left of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the diagonal line is the portion of the SPP rejection region due to pooling
3) the shaded area to the right of the vertical line at $F(\alpha_p;n_b,n_e)$ and above the solid horizontal line at $F(\alpha;n_t,n_e)$ is the portion of the SPP rejection region due to not pooling
4) the area above the solid horizontal at $F(\alpha;n_t,n_e)$ is the NPP rejection region
5) the area above the diagonal line is the APP rejection region in terms of F_t

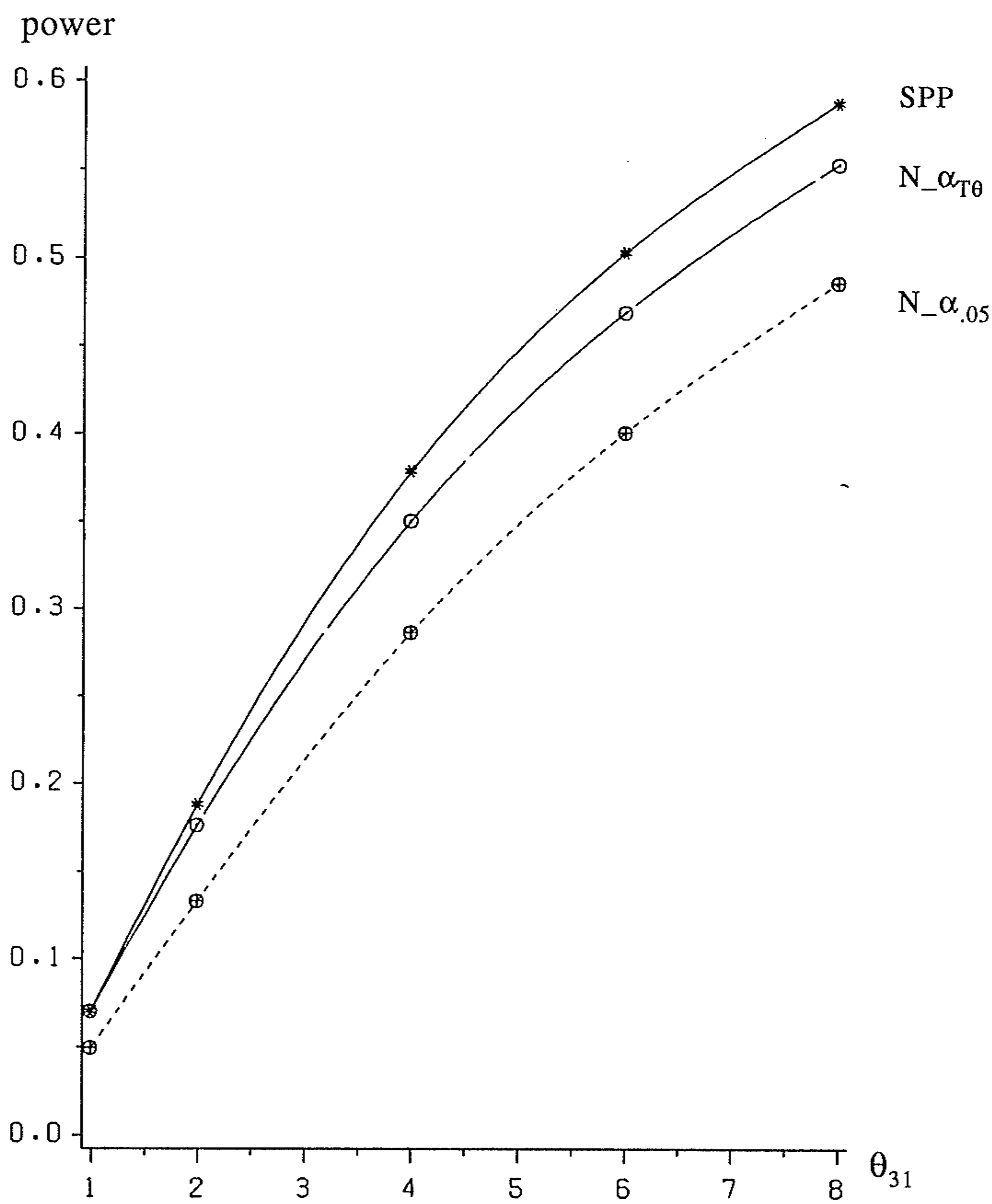
As we noted in the introduction to Chapter 2 the SPP does not appear to have been investigated with a random model where the expected classical denominator and auxiliary mean squares are ordered in this manner (i.e., a Type A random effects model). This being the case no recommendations have been made about the optimum value of the preliminary test size. In this case we would "recommend" the use of $\bar{\alpha}_p$ for reasons which will be clear after an examination of Figures 2.14 - 2.20. These figures and the notation correspond to those we used in Section 2.1 to illustrate the situation in the case of the Type B random effects model.

Figures 2.14 - 2.17 are plots of the probability of rejecting the main hypothesis of interest against the value of the main parameter of interest as the nuisance parameter ranges from small to large and for the borderline value of the preliminary test size. This is the value of the preliminary test size than seems optimum in some sense. Consider the following facts.

1. When $\theta_{21} = 1$ the SPP is uniformly more powerful than the NPP of size 0.05. It is also uniformly more powerful than the NPP of size α_θ which in this case is also the NPP of size α_T (since this is the value of the nuisance parameter at which the size of the SPP occurs). See Figure 2.14.

2. As θ_{21} departs from 1.0 the "size" of the SPP, α_θ, drops below α_T. At values of θ_{21} close to 1.0 the SPP is uniformly more powerful than the NPP of size 0.05 and the NPP of size α_θ. The SPP is also more powerful than the NPP of size α_T except at small values of the main parameter of interest. See Figure 2.15.

Figure 2.14 The Power Curve of a Borderline SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.



$\alpha_p = \bar{\alpha}_p$

$n_t = 2 \;\; n_n = 2 \;\; n_e = 4$

Figure 2.15 The Power Curve of a Borderline SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1.4$.

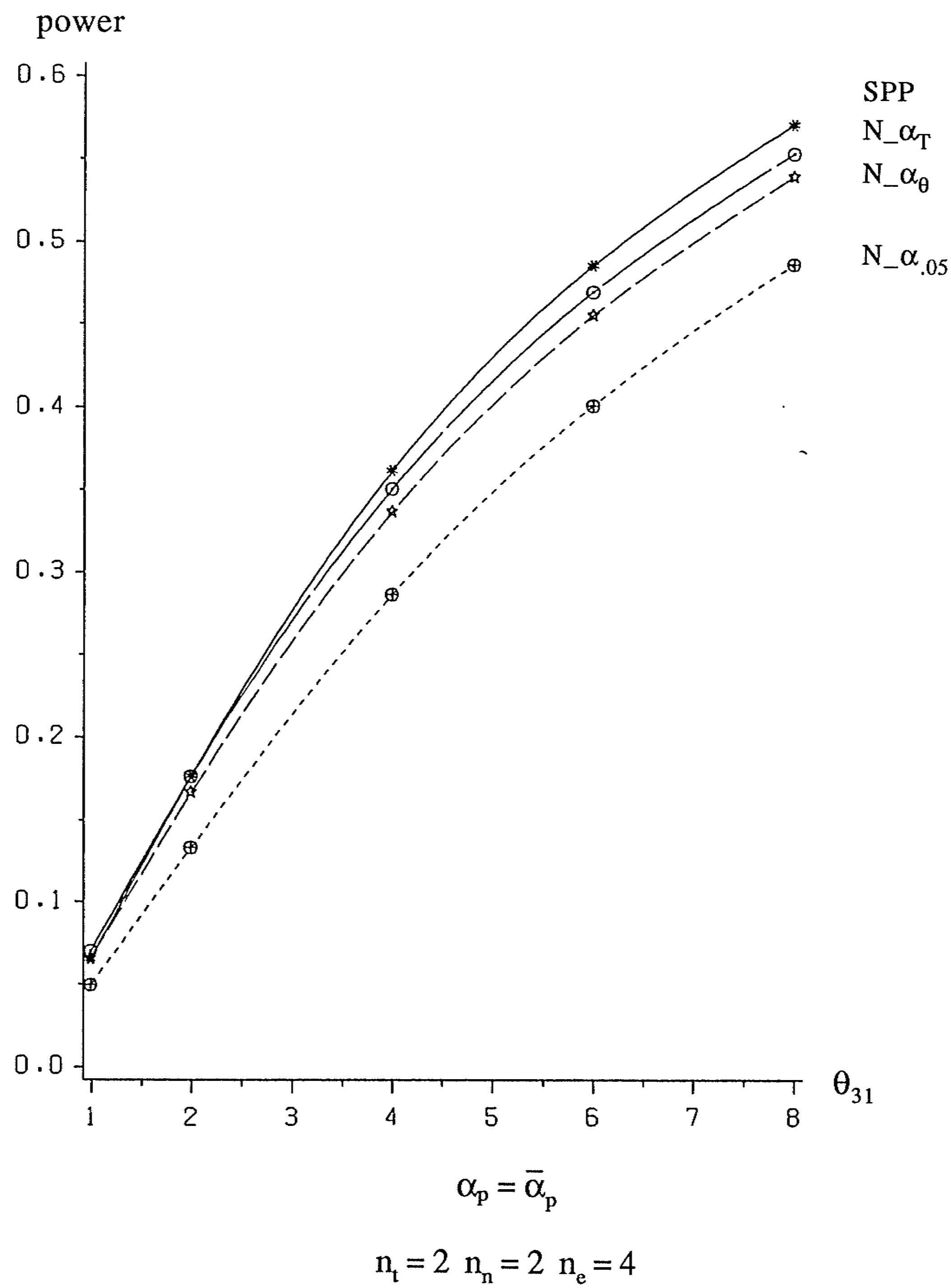


Figure 2.16 The Power Curve of a Borderline SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 2.0$.

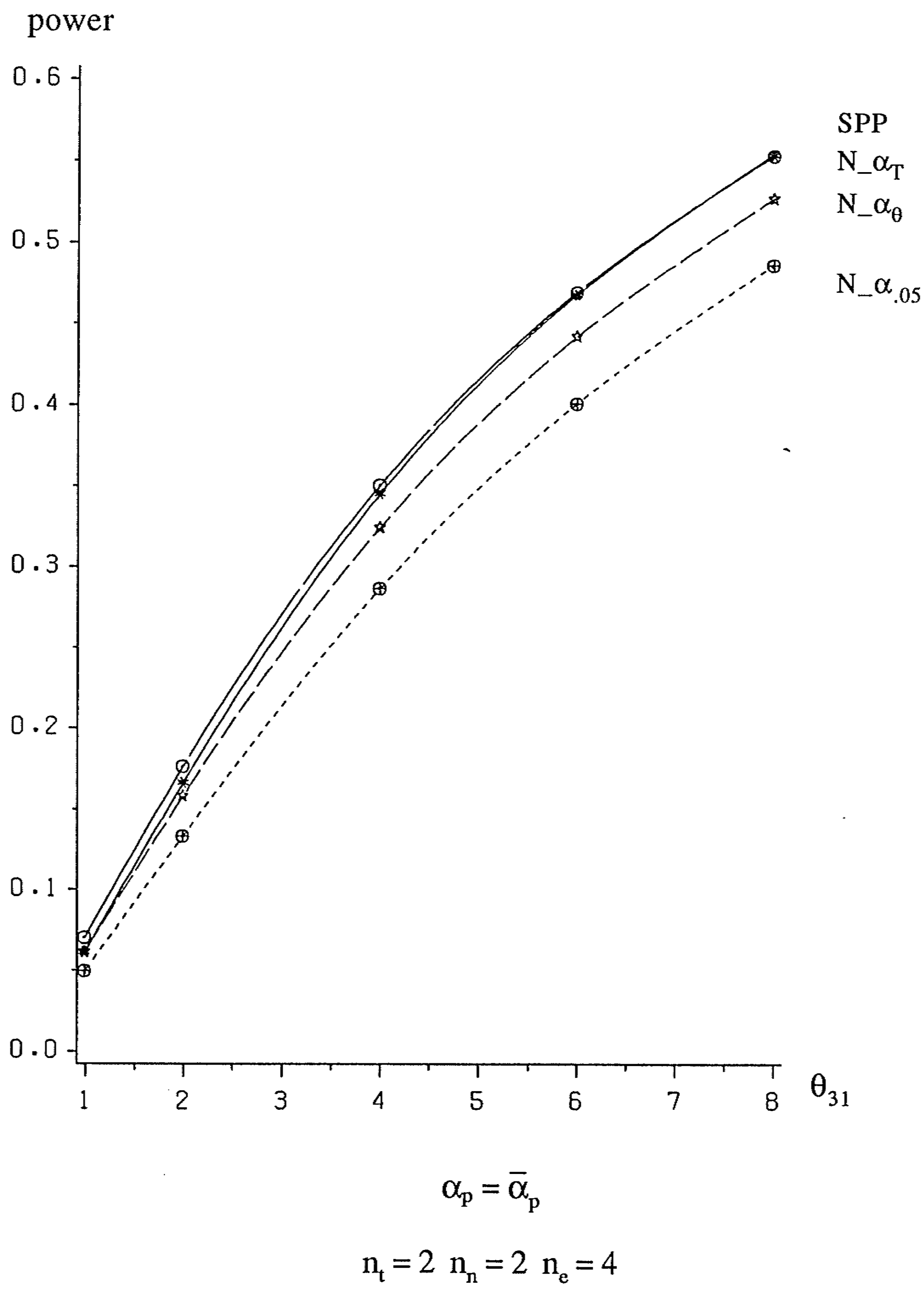
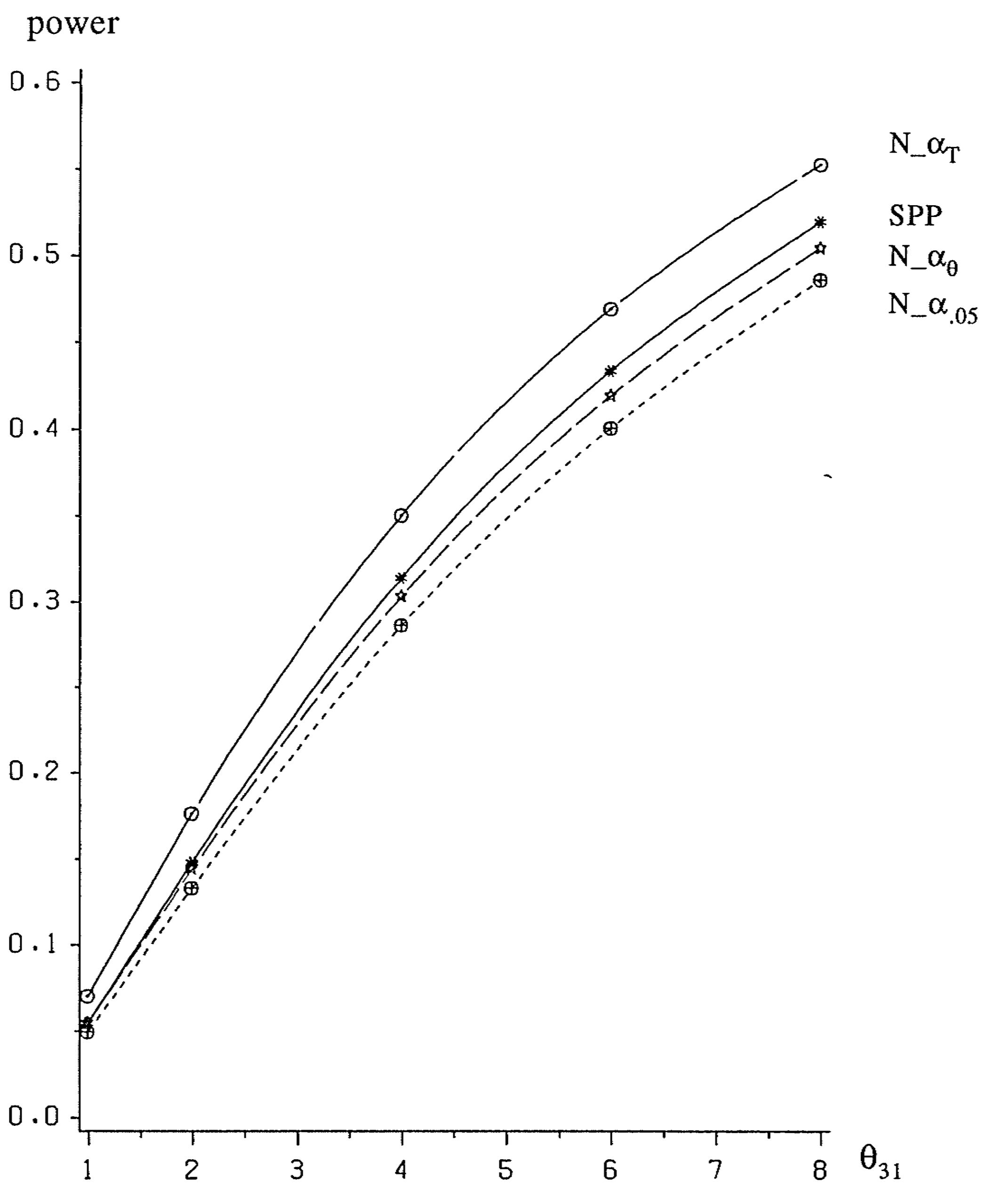


Figure 2.17 The Power Curve of a Borderline SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 5$.



$\alpha_p = \bar{\alpha}_p$

$n_t = 2 \ n_n = 2 \ n_e = 4$

3. As θ_{21} moves further away from 1.0 the SPP remains uniformly more powerful than both the NPP of size 0.05 and the NPP of size α_θ. As θ_{21} moves further from 1.0 the value of the main parameter of interest at which the SPP becomes more powerful than a NPP of size α_T also departs further from 1.0. See Figure 2.16.

4. As the nuisance parameter becomes large the SPP remains uniformly more powerful than the NPP of size α_θ, which in turn remains uniformly more powerful than a NPP of size 0.05. These procedures converge to the NPP of size 0.05 as the nuisance parameter approaches infinity. For these large values of the nuisance parameter the NPP of size α_T is the uniformly most powerful of all the considered procedures. See Figure 2.17.

Figures 2.18, 2.19, and 2.20 illustrate why Class A tests are not recommended for the random model where the values of the E(auxiliary MS) and E(classical denominator MS) are ordered as in the case of a randomized complete block experiment.

1. At $\theta_{21} = 1$ the SPP is the uniformly most powerful of any of the considered procedures. Consider Figure 2.18.

2. As θ_{21} increases from 1 the relationship quickly changes. At small values of θ_{21} the SPP remains more powerful than a NPP of size α_θ over the range of the main parameter of interest. Here the SPP drops below the power of the NPP of size 0.05 and the NPP of size α_T for

Figure 2.18 The Power Curve of a Class A SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.

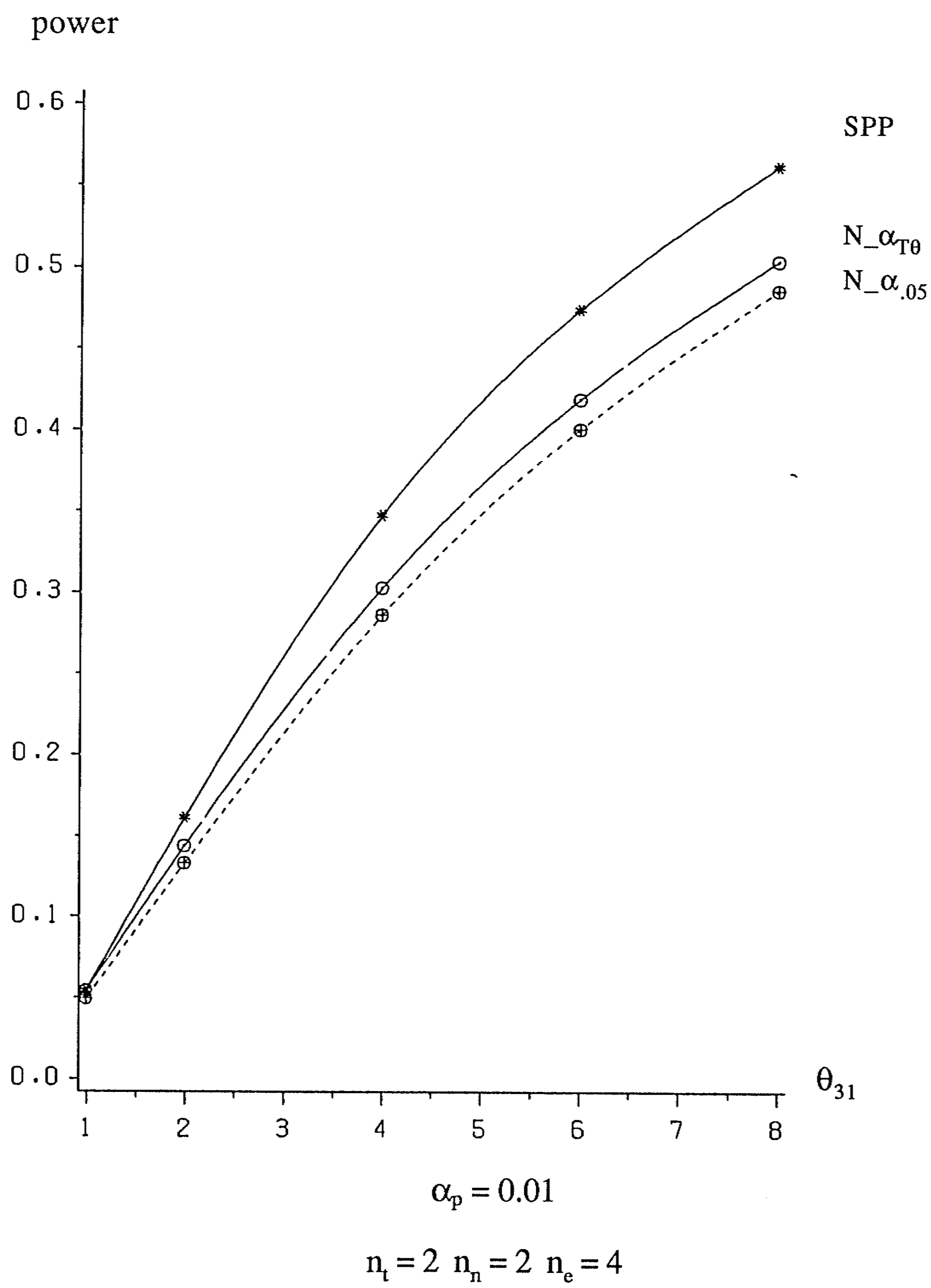


Figure 2.19 The Power Curve of a Class A SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 2.5$.

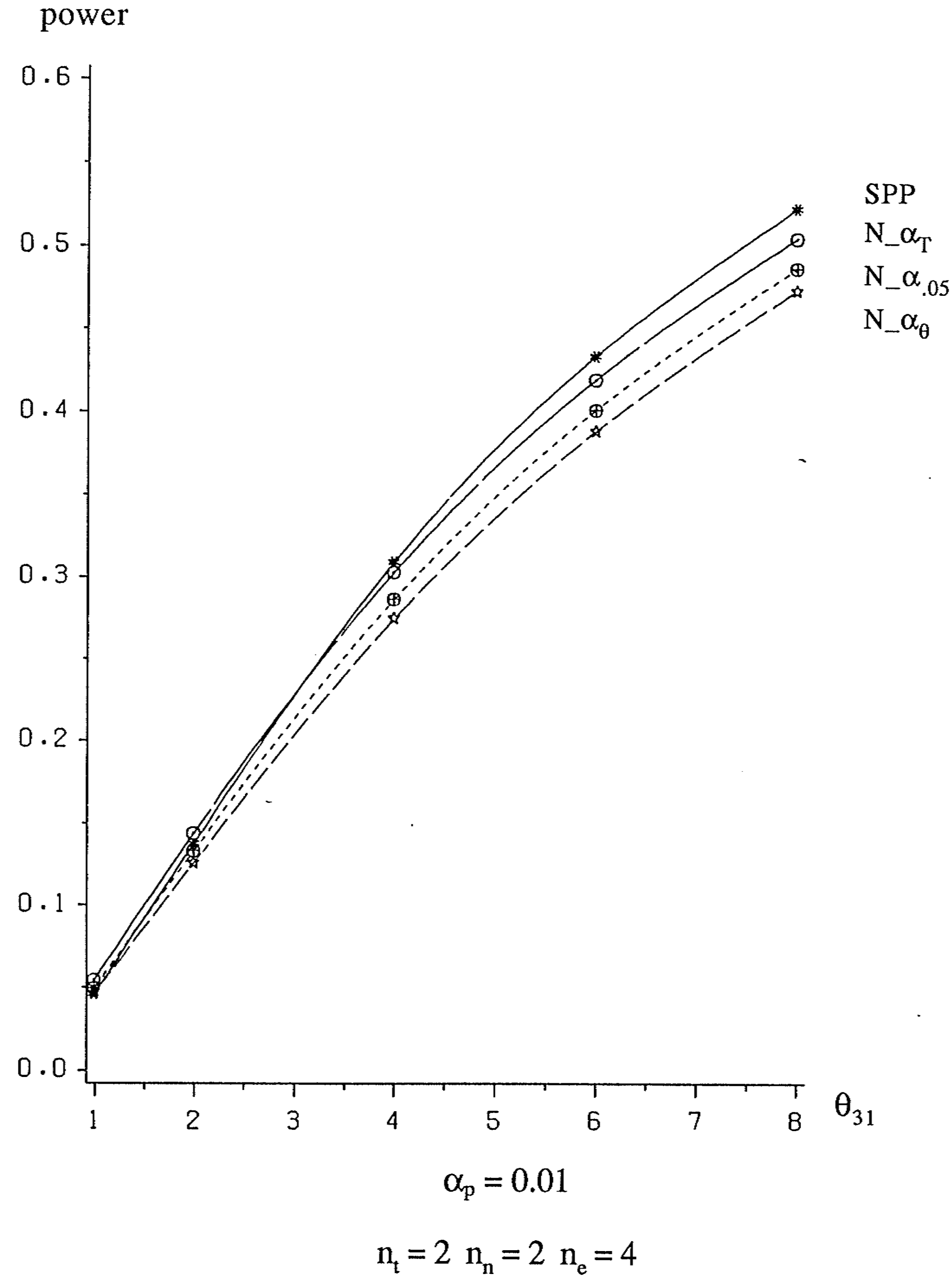
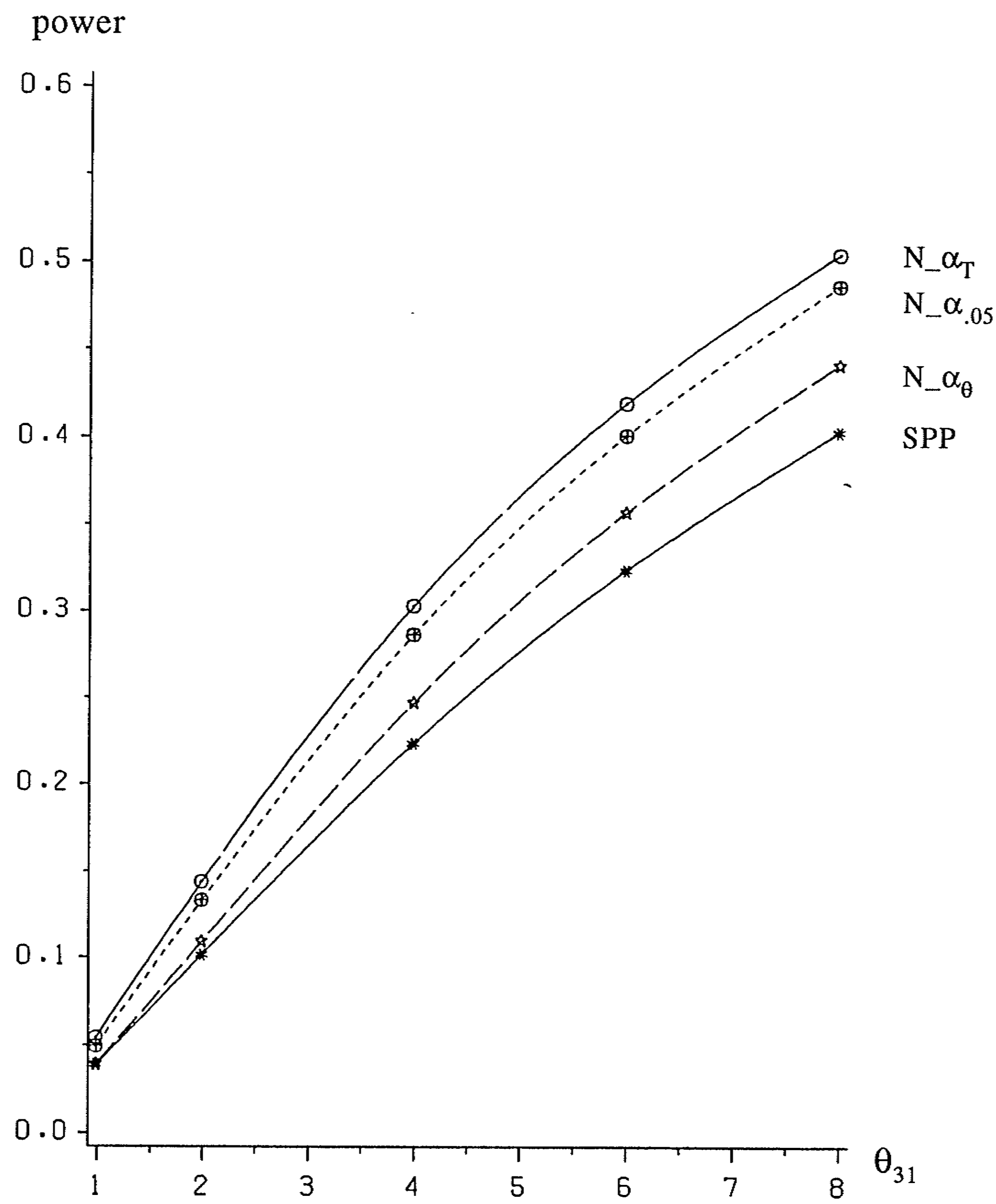


Figure 2.20 The Power Curve of a Class A SPP in a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 7$.



$\alpha_p = 0.01$

$n_t = 2$ $n_n = 2$ $n_e = 4$

small values of θ_{31} and becomes more powerful than these procedures as θ_{31} increases in size. See Figure 2.19.

3. As θ_{21} increases in size the SPP quickly becomes less powerful than the NPP of size 0.05, or the NPP of size α_θ. The SPP is uniformly less powerful than the NPP of size α_T. See Figure 2.20.

A Class B SPP behaves like a borderline SPP but the gains in power where gains are achieved are smaller in size.

Having considered the preceding facts about the SPP as utilized with Class A, Class B, and borderline values of the preliminary test size we can now explain our use of the words "optimum in some sense" to describe the borderline SPP. The borderline SPP ensures a gain in power over both the NPP of size α_θ, and the NPP of size 0.05. In addition it gives a gain in power over a NPP of size α_T for small values of the nuisance parameter. The same may be said of Class B tests but the borderline SPP gives greater gains in power where gains are achieved and since the disturbances in size are not appreciably different there is no advantage to Class B tests. Class A tests never perform in a manner superior to borderline tests and, as the nuisance parameter increases in size, their performance is noticeably inferior. These considerations are the basis for our description of the borderline SPP as optimum.

That is, the borderline SPP is sometimes better than all of the considered never pool procedures (this happens when θ_{21} is small), and is never the worst of the compared procedures; whereas, with the Class A SPP, when the nuisance parameter gets large the SPP is the least powerful of all.

CHAPTER 3

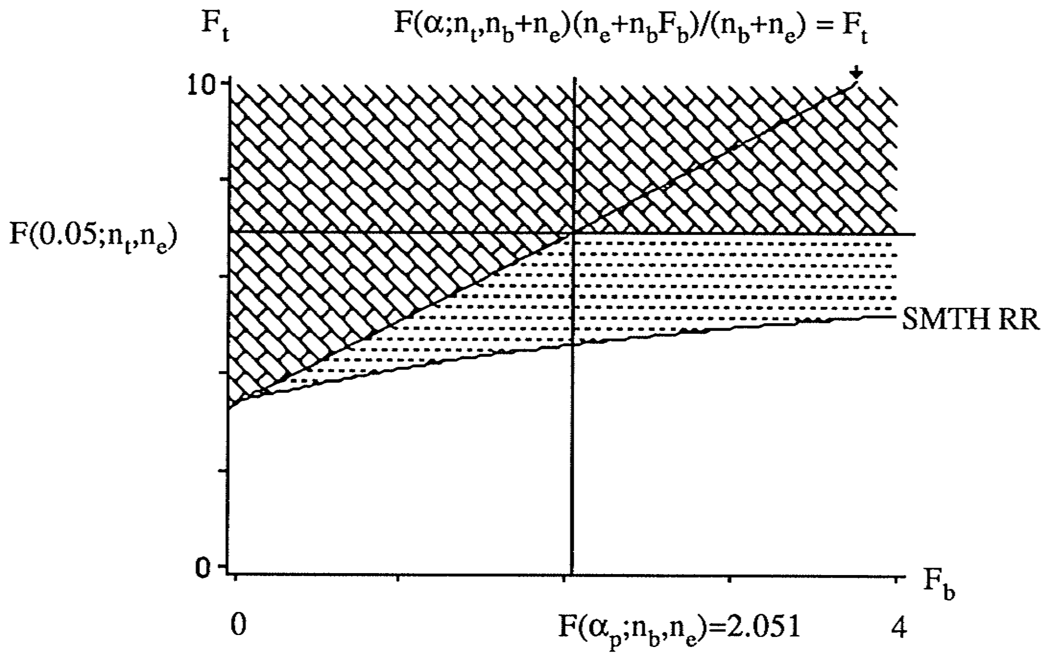
THE SMOOTH POOLING PROCEDURE

In Chapter 3 we illustrate the behaviour of our proposed smooth pooling procedure (SMTH) as used with a random effects model where the ordering of the expected mean squares is as in the case of the randomized complete block design.

3.1 THE SMOOTH POOLING PROCEDURE WITH A TYPE A RANDOM EFFECTS MODEL

In section 3.1 we will compare the results of the smooth pooling procedure (SMTH) with the results of a borderline sometimes pool procedure (SPP) and with the results of the never pool procedure (NPP). As can be seen from Figure 3.1 the SMTH (of nominal size 0.05) rejection region contains the rejection region of the borderline SPP (of nominal size 0.05), which in turn contains the rejection region of the NPP (of size 0.05). To reiterate a point made in Chapter 2, there are problems in comparing procedures when the results of some procedures depend on the value of a nuisance parameter while the results of some other procedures do not. In our case the results of the SMTH and SPP depend on the value of the nuisance parameter, θ_{21}, while the results of the NPP do not. We will make our comparison on the following basis. For fixed degrees of freedom we will determine the size

Figure 3.1 Rejection Regions of the NPP of size 0.05, and an SPP, and SMTH both of nominal size 0.05.



$n_b=2,\ n_t=2,\ n_e=4,\ \alpha_P = \bar{\alpha}_p=0.24380,\ \alpha=0.05$

1) The NPP rejection region of size 0.05 is the region above the horizontal line at $F(0.05;n_t,n_e)$.

2) The SPP rejection region of nominal size 0.05 is the region above the horizontal line at $F(0.05;n_t,n_e)$ and to the right of the vertical line at $F(\alpha_p;n_b,n_e)$ plus the region above the diagonal line $F_t=F(\alpha;n_t,n_b+n_e)(n_e+n_bF_b)/(n_b+n_e)$ and to the left of the vertical line at $F(\alpha_p;n_b,n_e)$. These are the portions due to not pooling and pooling respectively.

3) The SMTH rejection region of nominal size 0.05 is the rejection region of the SPP plus the region below the boundary of the SPP and above the curved line labelled SMTH RR. We note that the horizontal line $F(0.05;n_t,n_e)$ is an asymptote of the curve SMTH RR.

of the smooth pooling procedure. We will adjust our SPP to have the same size, α_T, as this smooth procedure and compare the power functions of these two procedures over a range of parameter values. We will also include in our comparison a NPP adjusted to the same size α_T. Figure 3.2 shows the rejection regions of the various procedures where the SPP and NPP have been adjusted to have the same size as the SMTH.

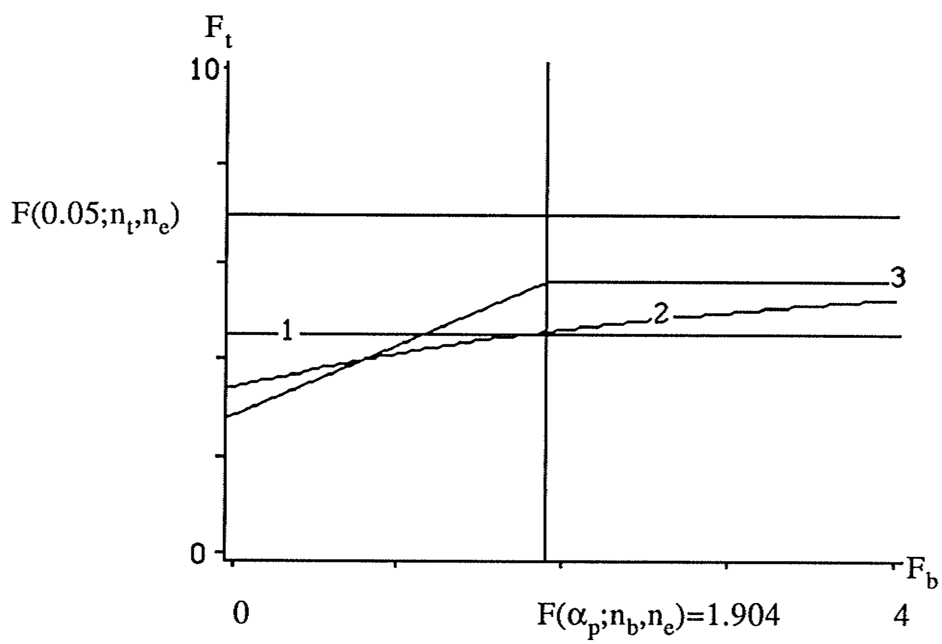
Figures 3.3 - 3.7 illustrate the behaviour of these procedures as the nuisance parameter takes on a range of values. We make the following observations based on these figures.

1) At values of θ_{21} close to 1 the SMTH and the SPP are about equal in power when the main parameter (θ_{31}) is small in value. As the main parameter becomes large the SPP is slightly more powerful than the SMTH. Both the SMTH and the SPP are always more powerful than the NPP of size α_T. See Figure 3.3.

2) As θ_{21} departs from 1 for all but small values of the main parameter the SMTH is the most powerful, followed by the SPP, which is in turn more powerful than the NPP of size α_T. When the main parameter is small the NPP of size α_T is slightly more powerful than the other procedures. See Figure 3.4.

3) As θ_{21} becomes larger the value of the main parameter at which the SMTH and the SPP become more powerful than the NPP of size α_T also becomes larger. See Figures 3.4 and 3.5.

Figure 3.2 Rejection Regions of the SMTH of size α_T (nominal size 0.05), and an SPP, and NPP both adjusted to have size α_T.



$n_b=2$, $n_t=2$, $n_e=4$, $\alpha_P = \bar{\alpha}_p=0.26245$, $\alpha=0.068957$

1) The NPP rejection region of size α_T is the region above the horizontal line labelled 1.

2) The SPP rejection region of size α_T is the region above the solid curve formed by the diagonal and horizontal line sections and labelled 3.

3) The SMTH rejection region of size α_T (nominal size 0.05) is the region above the curved line labelled 2. The horizontal line at $F(0.05;n_t,n_e)$ is an asymptote of this curve.

Figure 3.3 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of size α_T, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.

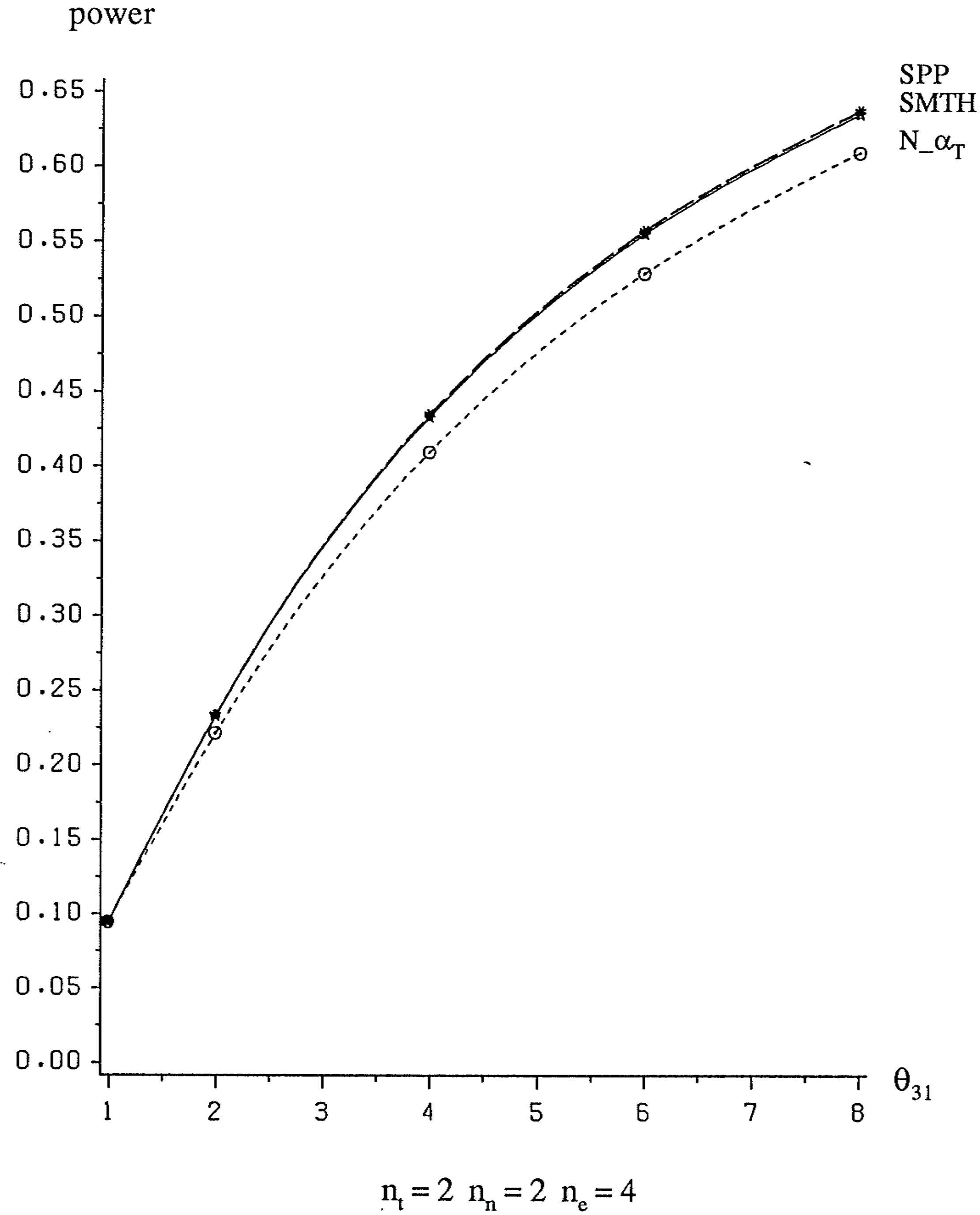


Figure 3.4 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of size α_T, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1.3$.

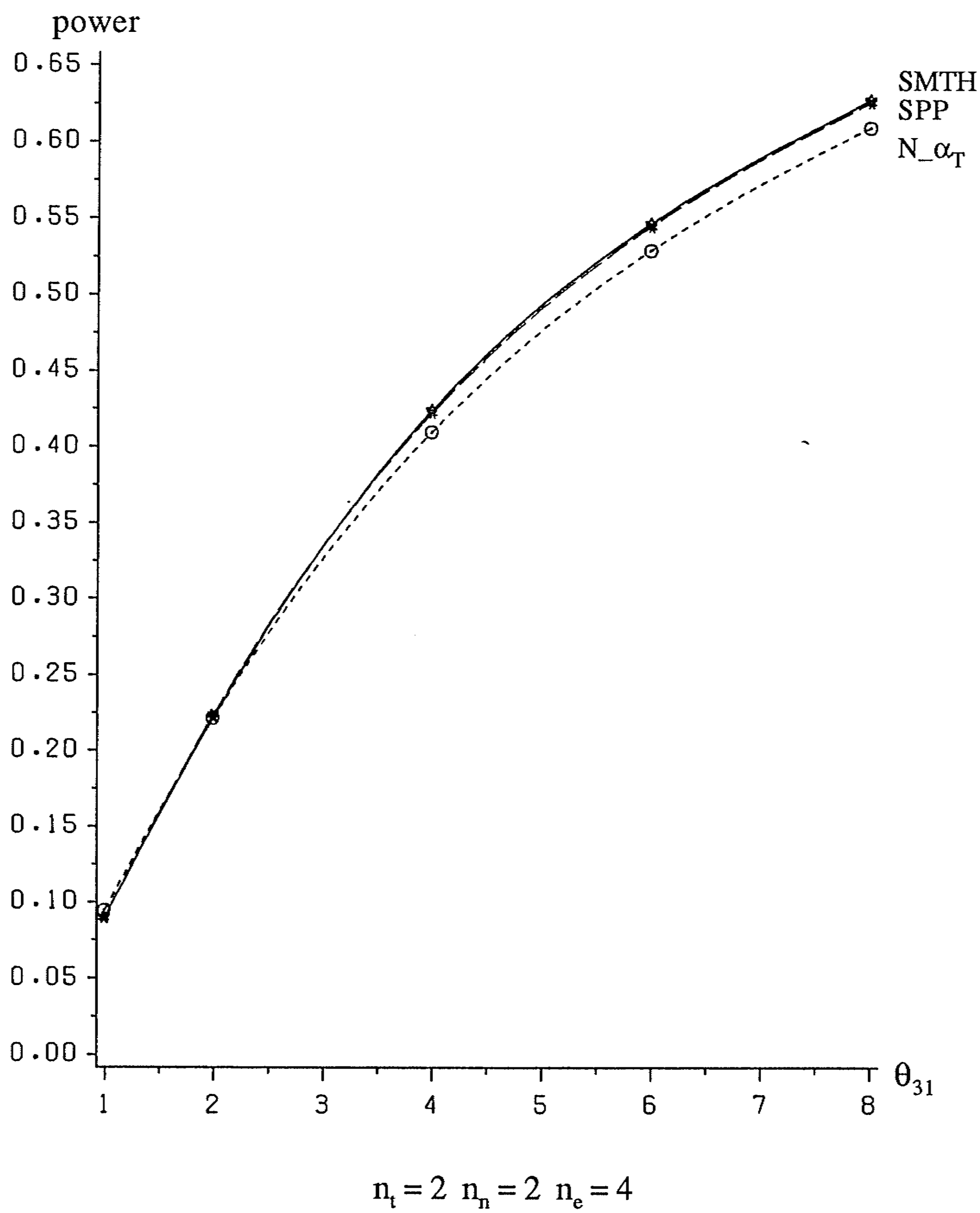


Figure 3.5 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of size α_T, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1.7$.

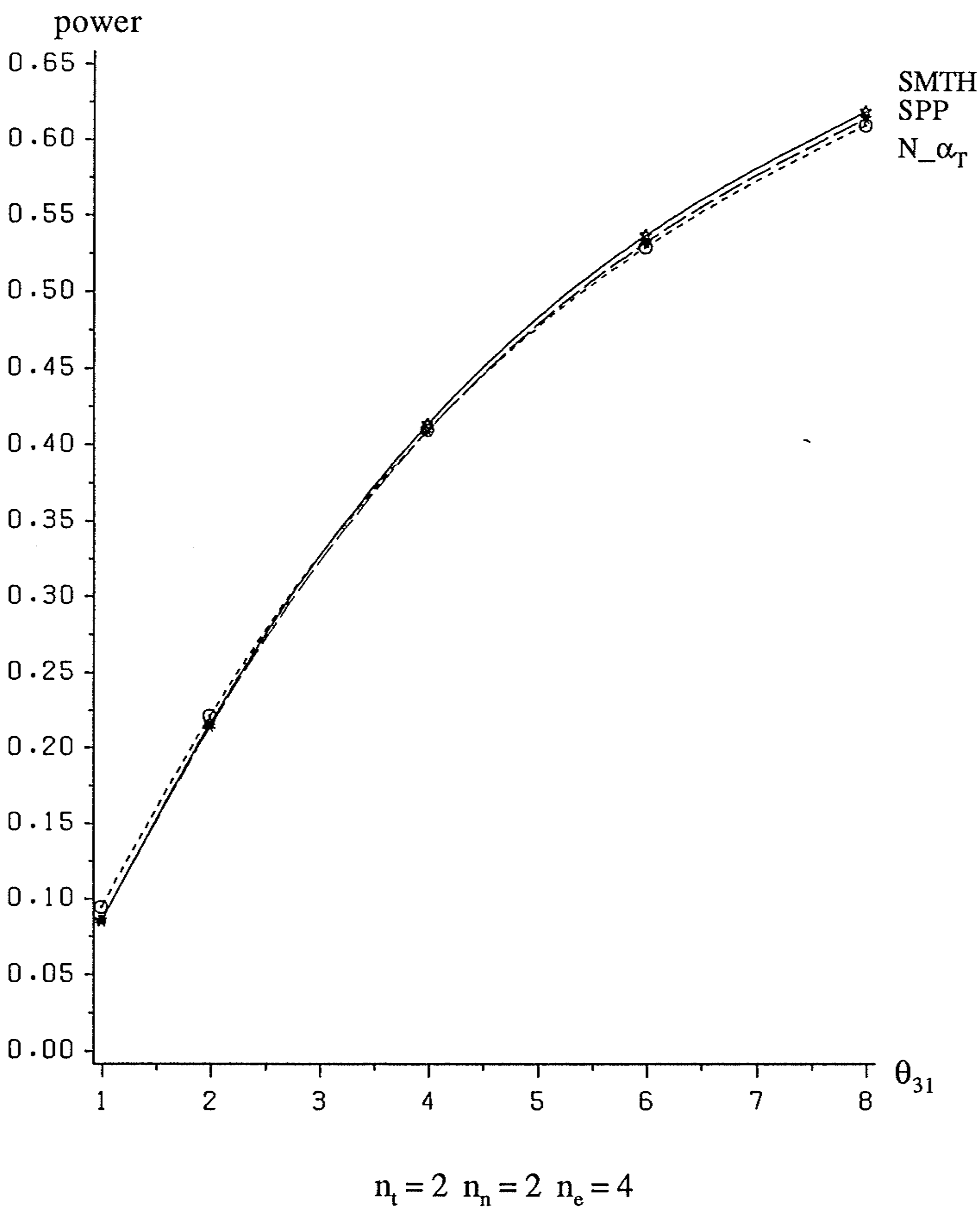
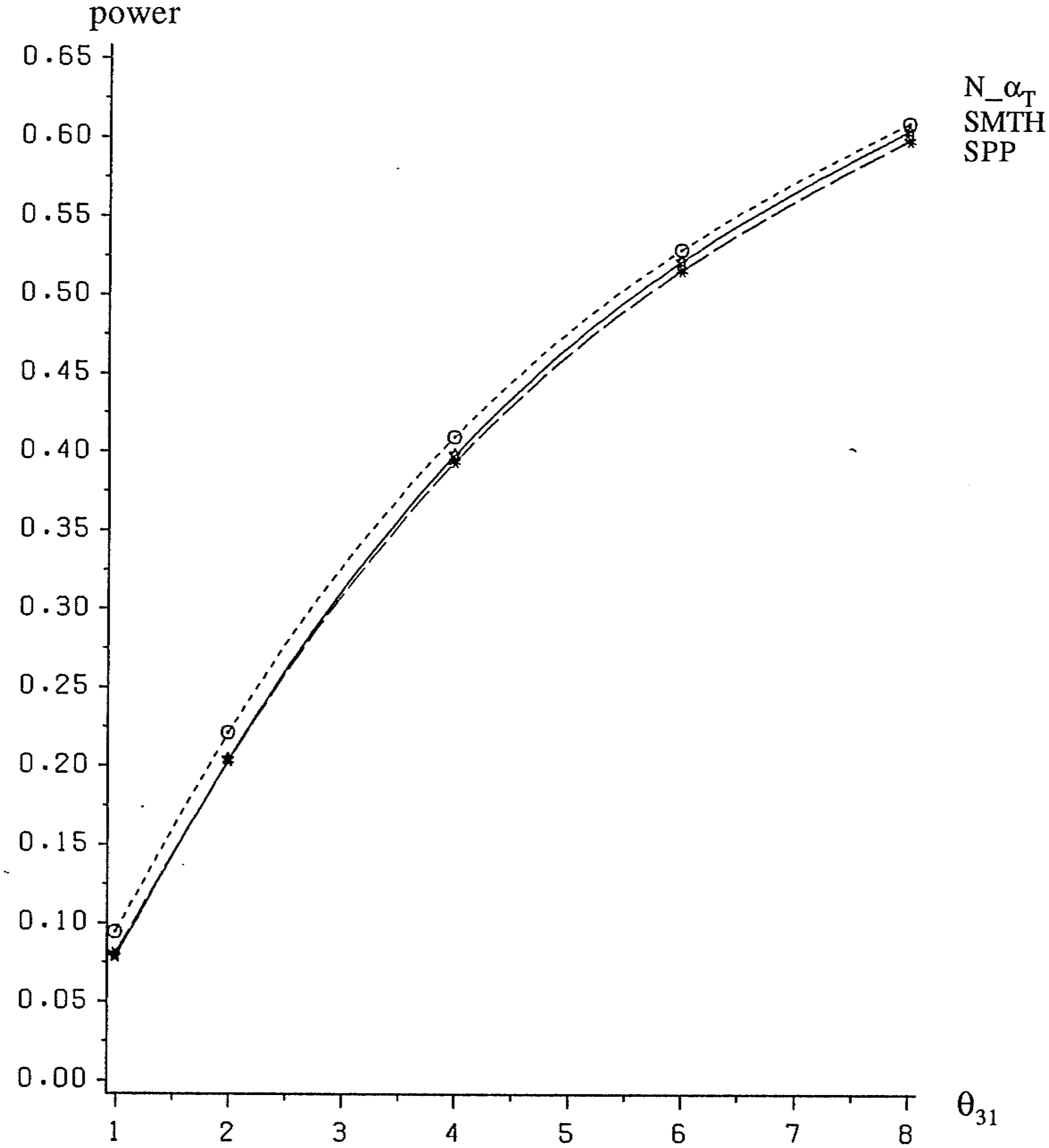
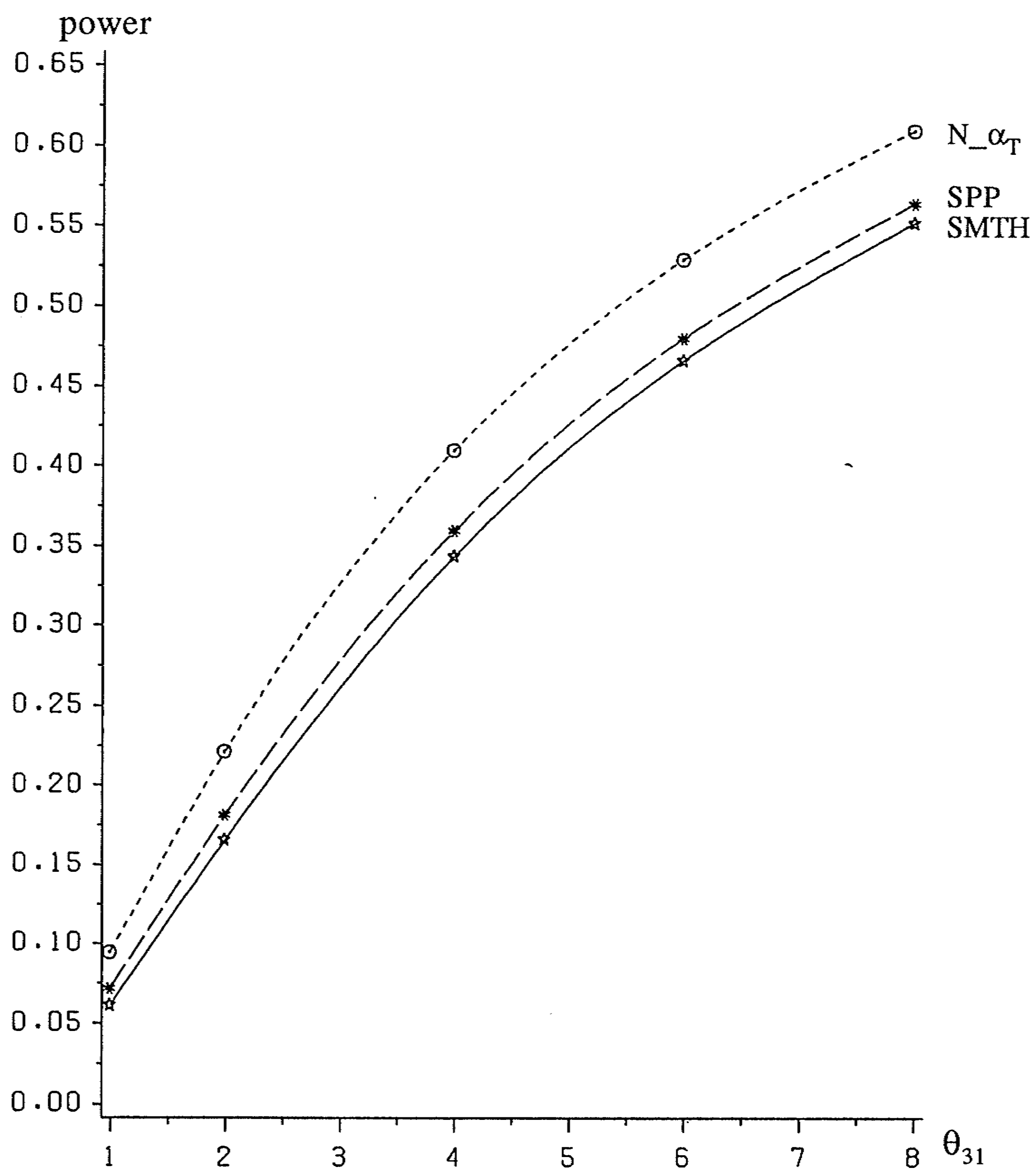


Figure 3.6 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of size α_T, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 2.5$.



$n_t = 2$ $n_n = 2$ $n_e = 4$

Figure 3.7 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of size α_T, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 10$.

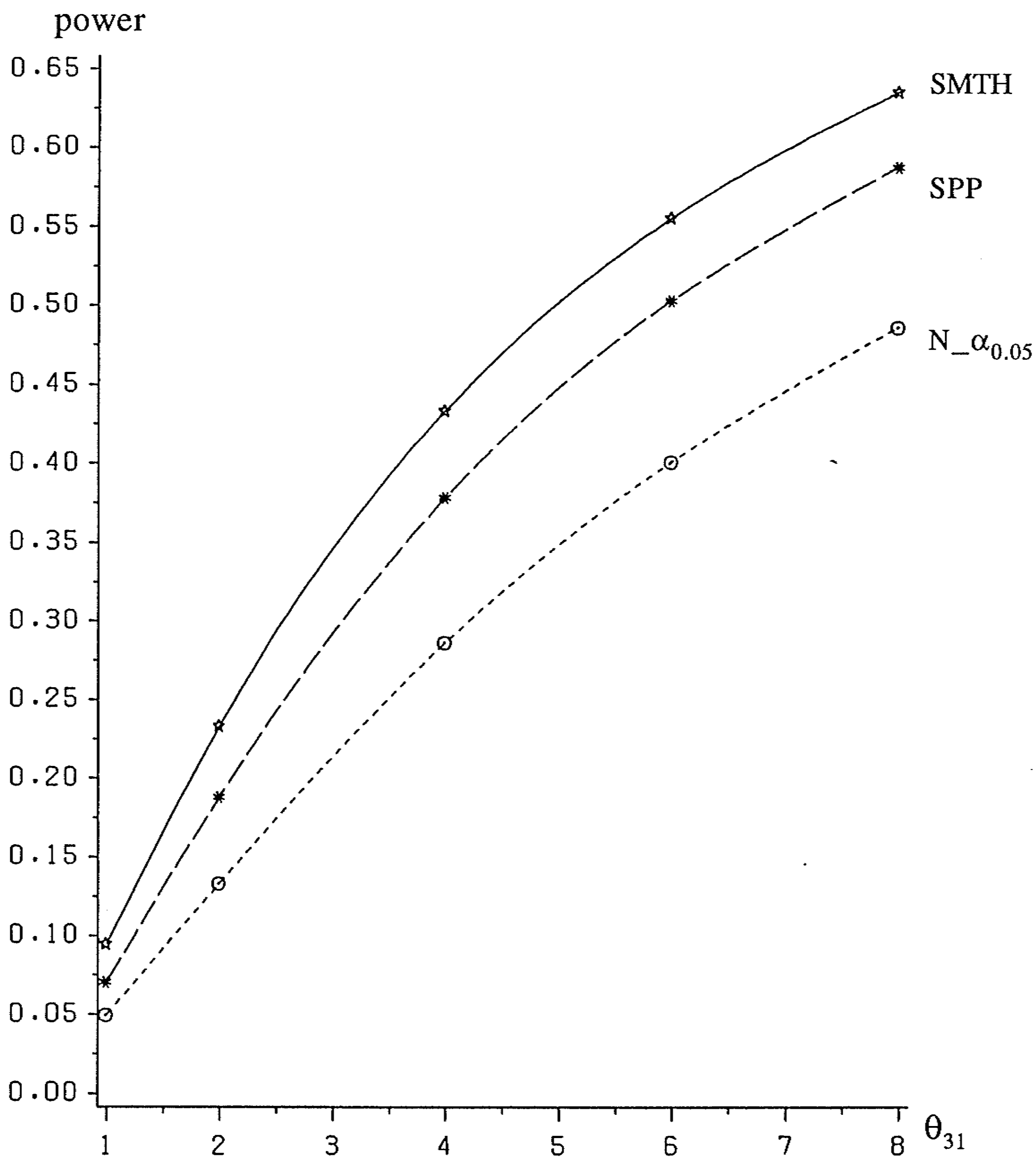


$n_t = 2 \;\; n_n = 2 \;\; n_e = 4$

4) When θ_{21} gets big enough no matter how big the main parameter gets then the NPP of size α_T is always the most powerful. Also at some intermediate value of θ_{21} the SPP overtakes the SMTH, but note that by the time this occurs the NPP of size α_T is already the most powerful. See Figures 3.6 and 3.7.

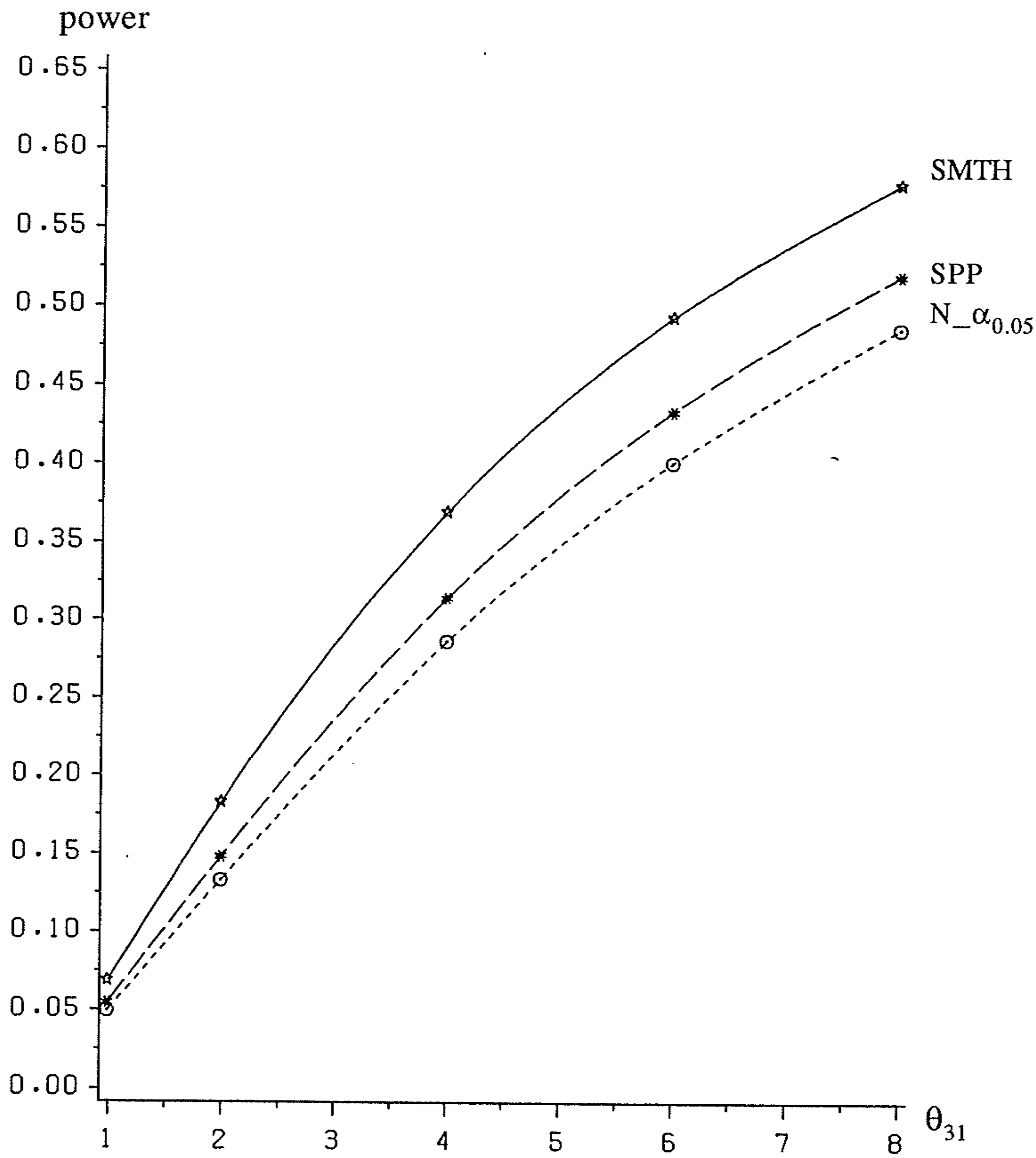
Figures 3.8 - 3.10 illustrate the results of the unadjusted procedures. Consideration of these figures shows that, while the smooth pooling procedure always has the highest probability of rejecting the hypothesis of interest in the region where the main hypothesis is false (a desirable feature), the same may be said when the main hypothesis is true (an undesirable feature).

Figure 3.8 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of nominal size 0.05, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 1$.



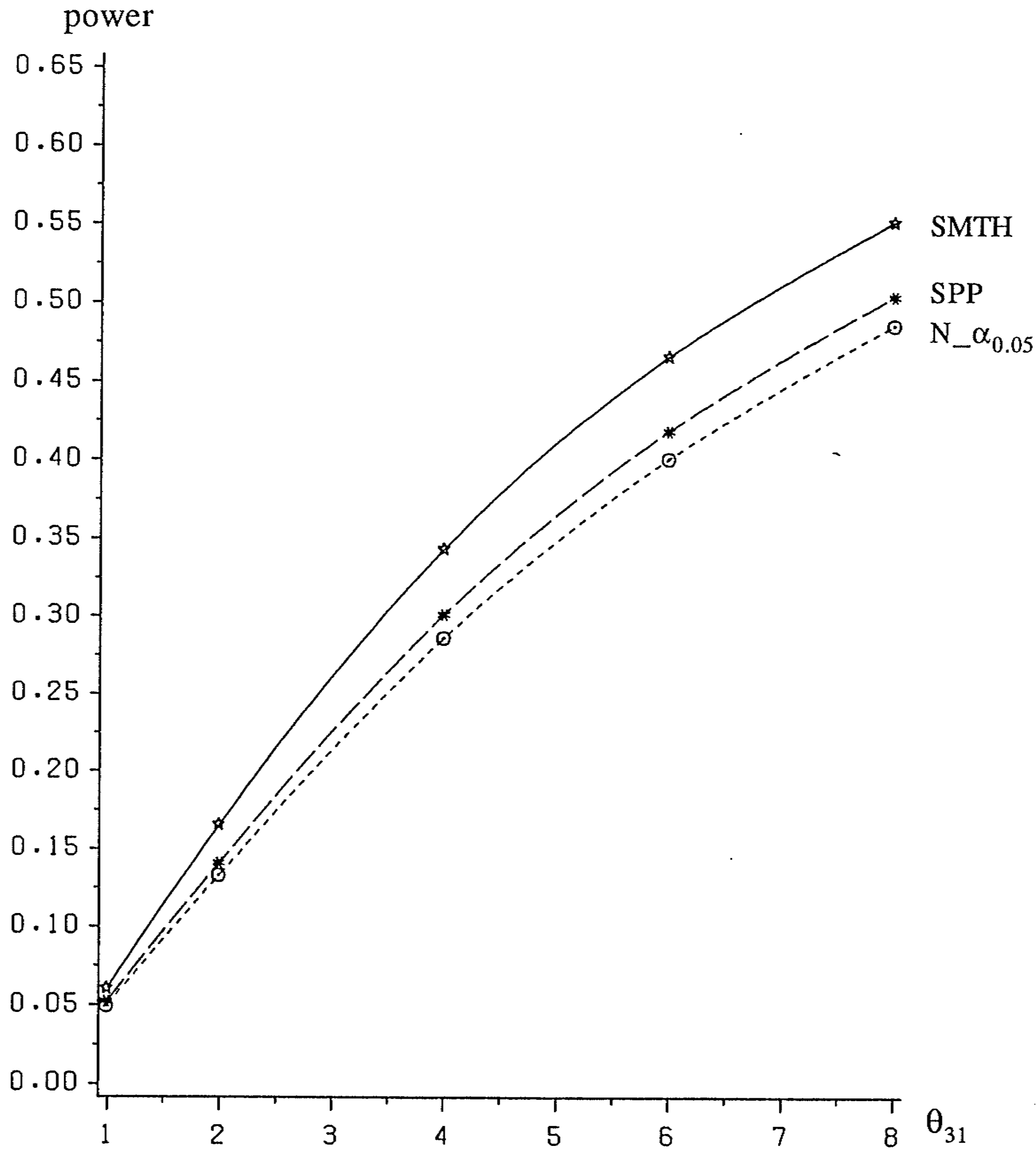
$n_t = 2 \ n_n = 2 \ n_e = 4$

Figure 3.9 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of nominal size 0.05, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 5$.



$n_t = 2$ $n_n = 2$ $n_e = 4$

Figure 3.10 Power Comparisons of the Smooth, Sometimes, and Never Pool Procedures of nominal size 0.05, used with a Type A Random Effects Model when the Nuisance Parameter $\theta_{21} = 10$.



$n_t = 2 \; n_n = 2 \; n_e = 4$

CHAPTER 4

CONCLUSION

In Chapter 4 we conclude the thesis. Section 4.1 makes a recommendation as to which of the procedures should be used in practice. In Section 4.2 we discuss the potential for further research.

4.1 WHICH PROCEDURE SHOULD BE USED?

Intuitively it seems that all information arising from an experiment should be utilized. Classical never pool procedures that are used in the analysis of variance methodology ignore potential information, and historically, sometimes pooling procedures have been used in an attempt to incorporate more fully the information provided by experimental results. The smooth pooling procedure proposed in this thesis seems to be an improvement over the sometimes pool procedure and therefore a procedure that should be considered. The reasoning behind this statement is as follows. The results that are produced by the proposed procedure pass smoothly from those of the never pool procedure to those of the always pool procedure (as the block p-value passes from 0 to 1) whereas the sometimes pool procedure produces either the result of the never pool procedure or the result of the always pool procedure. Secondly, the SPP is arbitrary, due to the choice of the preliminary test size, while the smooth procedure is not. Finally it appears from our limited examination of these procedures that the range of nuisance parameter values where the proposed procedure is more powerful than the sometimes pool procedure is more extensive than the range of values where the sometimes pool procedure is more powerful. The proposed smooth pooling procedure is also more powerful than the classical never pool procedure for a substantial range of nuisance parameter values.

On the one hand, the proposed procedure is attractive because:

1) it is smooth;
2) it is not arbitrary;
3) it has good power properties over a wide range of parameter values.

On the other hand, the interpretation of the adjusted p-value of the smooth procedure is not straightforward.

4.2 DIRECTIONS FOR FUTURE RESEARCH

The question of "Why this procedure?" is a question of primary importance that needs to be addressed. Is it possible, for example, that this proposed smooth pooling procedure is a Bayes procedure? That is, is there a Bayesian interpretation or view of the problem which would yield the smooth pooling procedure p-value as the posterior probability that the null hypothesis $H_o:\sigma_t^2 = 0$ is true?

Once the philosophical questions have been resolved there are a number of issues of a technical nature which need to be addressed.

There is the question of how the proposed smooth pooling procedure would work in the case where the expected mean squares are ordered as in the case of the nested model (i.e. E(classical denominator MS) $\geq$ E(auxiliary MS)).

The results for the fixed models need to be developed and the results for all models (fixed, random, with both orderings of expected mean squares) should be investigated for a range of parameter values.

Finally we should consider how this type of procedure would be conducted in other situations. In a two factor experiment with interaction, for example, how would the smooth pooling procedure work if our interest was in factor A and we decided that we would smoothly incorporate both the interaction and treatment B information? Should we use the distribution of the F statistic for A, conditional on the F statistic for B and the F statistic for AB being less than or equal their observed values (and under the assumption of no B or AB variability)? Or should the conditioning be done in some ordered fashion?

These are only examples of the problems that need to be addressed before the smooth pooling procedure can be adopted.

APPENDIX

The following WATFIV program runs on the University of Manitoba mainframe. For user supplied degrees of freedom for treatments, blocks, and error and observed F statistics for treatments and blocks the p-value of the smooth procedure is calculated. In addition an estimate of the amount this estimate of the p-value deviates from the true unknown p-value is given. The last two lines of the appendix give the results calculated by the program for the example used in section 1.1.

```
C
C  CONDF IS A REAL VALUED FUNCTION OF THE VARIABLE FT
C  THAT CALCULATES THE CONDITIONAL F DENSITY FUNCTION
C
C  VARIABLE DICTIONARY FOR FUNCTION CONDF
C
C  FT  - VARIABLE SUPPLIED BY MAIN PROGRAM.
C  FB  - OBSERVED F STATISTIC FOR BLOCKS(SUPPLIED BY USER)
C  NT  - DEGREES OF FREEDOM FOR TREATMENTS(SUPPLIED BY
         USER).
C  NB  - DEGREES OF FREEDOM FOR BLOCKS(SUPPLIED BY USER).
C  NE  - DEGREES OF FREEDOM FOR ERROR(SUPPLIED BY USER).
C
C   A,B,C,B1,C1,K,K1,T1,T2,T3,T4,X2,X3,IER1,IER2 ARE
C   INTERMEDIATE VALUES CALCULATED IN THE PROGRAM.
C
      REAL FUNCTION CONDF(FT)
      REAL FT,FB,NE,NB,NT,A,B,C,B1,C1,K,K1,T1
      INTEGER IER1,IER2
      REAL T2,T3,T4,X2,X3
      NE=9.0
      NB=3.0
      NT=3.0
      A=NE/2
      B=NB/2
      C=NT/2
      B1=A+C
```

```
      C1=C-1
      FB=0.5
      T4=NB*FB/NE
      X2=T4/(1+T4)
      X3=T4/(1+T4+NT*FT/NE)
      CALL MDBETA(X2,B,A,T2,IER1)
      CALL MDBETA(X3,B,B1,T3,IER2)
      K1=GAMMA(B1)/GAMMA(A)/GAMMA(C)
      K1=K1*(NT/NE)**C
      T1=K1*FT**C1/(1+NT*FT/NE)**B1*T3/T2
      CONDF=T1
      RETURN
      END
C
C  MAIN PROGRAM CALCULATES THE P-VALUE
C  OF THE SMOOTH PROCEDURE
C
C  VARIABLE DICTIONARY FOR THE MAIN PROGRAM
C
C  DCADRE IS AN IMSL NUMERICAL INTEGRATION PACKAGE.
C  AJ,BJ ARE THE LOWER AND UPPER LIMITS OF INTEGRATION
C  OVER WHICH THE CONDITIONAL F DISTRIBUTION IS
C  INTEGRATED.  AJ IS SET CLOSE TO 0, WHILE BJ IS SET TO
C  THE OBSERVED F FOR TREATMENTS.  AERR, RERR ARE
C  CONTROLS ON THE  MAGNITUDE OF THE ABSOLUTE AND
C  RELATIVE ERRORS  ALLOWED IN THE NUMERICAL
C  INTEGRATION(SET BY USER).  ERROR IS AN ESTIMATE OF THE
C  ACTUAL ERROR IN THE NUMERICAL INTEGRATION
C  (CALCULATED BY PROGRAM).  IER IS A CODE CALCULATED
C  BY THE PROGRAM THAT  INDICATES POSSIBLE PROBLEMS IN
C  THE NUMERICAL INTEGRATION.  FB IS THE OBSERVED F FOR
C  BLOCKS(SUPPLIEDBY USER).  PVAL IS THE CALCULATED
C  P-VALUE FOR THE SMOOTH  PROCEDURE.
C
      INTEGER IER
      REAL DCADRE,AJ,BJ,AERR,RERR,ERROR,FB,PVAL
      EXTERNAL CONDF
      FB=0.5
      AJ=0.0000001
      BJ=3.80
      AERR=0.000001
      RERR=0.000001
      PVAL=1-DCADRE(CONDF,AJ,BJ,AERR,RERR,ERROR,IER)
```

```
PRINT,' '
 PRINT,'THE P-VALUE IS',PVAL,' .'
PRINT,' '
PRINT,'AN ESTIMATE OF THE ERROR IN THE P-VALUE
        IS',ERROR,' .'
STOP
END
```

```
THE P-VALUE IS      0.0223179 .

AN ESTIMATE OF THE ERROR IN THE P-VALUE IS  0.0000009 .
```

BIBLIOGRAPHY

This bibliography resulted from a comprehensive literature search which was conducted on the subject matter before the thesis was undertaken. Although only a few of the articles are referred to in the thesis the extensive list illustrates the amount of interest in this area and also serves as an aid for other researchers.

1. Agarwal, B. L. and Gupta, V. P. (1981) "Admissibility of Synthesised Variance Ratio Preliminary Procedures in a Mixed Model," Sankhya, Series B, 43, 212-221.

2. Agarwal, B. L. and Gupta, V. P. (1981) "*Admissibility of Three Satterthwaite F-statistics in One Preliminary Test Procedure in a Mixed Model*," Journal of the Indian Society of Agricultural Statistics, 33, 81-88.

3. Agarwal, B. L. and Gupta, V. P. (1983) "*Approximate Power of Test Procedures Using Two Preliminary Tests in the Mixed Model*," Journal of the Indian Society of Agricultural Statistics, 35, 56-68.

4. Ali, M. A. and Srivastava, S. R. (1981) "*On Certain Approximations of Power of a Test Procedure Using Two Preliminary Tests in a Mixed Model*," Bulletin of Mathematical Statistics, 19(3), 103-113.

5. Ali, M. A. and Srivastava, S. R. (1983) "*On Power Function of a Sometimes Pool Test Procedure in a Mixed Model-I: A Theoretical Investigation*," Journal of the Indian Society of Agricultural Statistics, 35, 80-90.

6. Arnold, Barry C. (1970) "*Hypothesis Testing Incorporating a Preliminary Test of Significance*," Journal of the American Statistical Association, 65, 1590-1596.

7. Bancroft, T. A. (1944) "*On Biases in Estimation due to the Use of Preliminary Tests of Significance*," Annals of Mathematical Statistics, 15, 190-200.

8. Bancroft, T. A. (1964) "*Analysis and Inference for Incompletely Specified Models Involving the Use of Preliminary Tests of Significance*," Biometrics, 20, 427-442.

9. Bancroft, T. A. (1975) "*Testimation, Testipredicting, and Testitesting as Aids in Using Snedecor and Cochran's "Statistical Methods"*," Biometrics, 31, 319-324.

10. Bancroft, T. A. and Han, Chein-Pai (1968) "*On Pooling Means When Variance is Unknown*," Journal of the American Statistical Association, 63, 1333-1342.

11. Bancroft, T. A. and Han, Chein-Pai (1976) "*On Pooling Means in Multivariate Normal Distributions*," Essays in Probability and Statistics, 1, 353-362.

12. Bancroft, T. A. and Han, Chein-Pai (1977) "*Inference based on Conditional Specification:a Note and a Bibliography*," International Statistical Review, 45, 117-127.

13. Bancroft, T. A. and Han, Chein-Pai (1978) "*Estimating Regression Coefficients Under Conditional Specification*," Communications in Statistics, Part A--Theory and Methods, 7, 47-56.

14. Bancroft, T. A. and Han, Chein-Pai (1983) "*A Note on Pooling Variances*," Journal of the American Statistical Association, 78, 981-983.

15. Bancroft, T. A. and Han, Chein-Pai (1986) "*Pooling Data*," Encyclopedia of Statistical Sciences, 7, 88-93.

16. Bass, Frank M. and Wittink, Dick L. (1978) "*Pooling Issues and Methods of Regression Analysis*," Journal of Marketing Research, 15, 277-279.

17. Basu, Debabrata (1977) "*On the Elimination of Nuisance Parameters*," Journal of the American Statistical Association, 72, 355-366.

18. Bennett, B. M. (1952) "*Estimation of Means on the Basis of Preliminary Tests of Significance*," Annals of the Institute of Statistical Mathematics, 4, 41-43.

19. Bennett, B.M. (1956) "*On the Use of Preliminary Tests in Certain Statistical Procedures*," Annals of the Institute of Statistical Mathematics, 8, 45-47.

20. Bhattacharjee, G. P. and Alan, S. S. (1982) "*Testing Equality of Variance After a Preliminary Test of Significance for Correlation Coeficient*," Journal of Statistical Planning and Inference, 6, 17-28.

21. Bock, M. E., Yancey, T. A. and Judge, G. G. (1973) "*The Statistical Consequences of Preliminary Test Estimators in Regression*," Journal of the American Statistical Association, 68, 109-116.

22. Box, G. E. P. and Tiao, G. C. (1968) "*Bayesian Estimation of Means in the Random Effects Model*," Journal of the American Statistical Society, 63, 174-181.

23. Bozivich, Helen, Bancroft, T. A. and Hartley, H. O. (1956) "*Power of Analysis of Variance Test Procedures for Certain Incompletely Specified Models*," Annals of Mathematical Statistics, 27, 1017-1043.

24. Brewster, J. F. and Zidek, J. V. (1974) "*Improving on equivariant estimators*," Annals of Statistics, 2, 21-38.

25. Brook, R. J. (1976) "*On the Use of a Regret Function to Set Significance Points in Prior Tests of Estimation*," Journal of the American Statistical Association, 71, 126-131.

26. Brown, L. (1967) "*The Conditional Level of the t Test*," Annals of Mathematical Statistics, 38, 1068-1071.

27. Buehler, R. J. and Feddersen, A. P. (1963) "*Note on a Conditional Property of Student's t*," Annals of Mathematical Statistics, 34, 1098-1100.

28. Cohen, Arthur (1974) "*To Pool or Not to Pool in Hypothesis Testing*," Journal of the American Statistical Association, 69, 721-725.

29. Da Silva, A. G. and Han, Chein-Pai (1984) "*Pooling means in a multivariate normal population*," Estadistica, 36, 63-74.

30. Der, William and Han, Chein-Pai (1979) "*Selection of a Preliminary Test Estimator in Linear Regression*," ASA Proceedings of Business and Economic Statistics Section, 441-446.

31. Dickey, J. M. (1967) "*A Bayesian Hypothesis-Decission Procedure*," Annals of Mathematical Statistics, 19, 367-369.

32. Dykstra, Richard L? (1981) "*An Isotonic Regression Algorithm*," Journal of Statistical Planning and Inference, 5, 355-363.

33. Faller, Alan J. (1981) "*An Average Correlation Coeficient*," Journal of Applied Meteorology, 20, 203-205.

34. Genest, Christian (1984) "*Pooling Operators with the marginalization property.*," Canadian Journal of Statistics, 12, 153-163.

35. Ghosh, Malay and Dey, Dipak K. (1986) "*On the inadmissibility of-preliminary test estimators when the loss involves a complexity cost*," Annals of the Institute of Statistical Mathematics, 38, 419-427.

36. Gupta, V. P. and Srivastava, S. R. (1965) "*Estimation after Preliminary Testing in ANOVA Model-I*," Biometrics, 21, 752-758.

37. Gupta, V. P. and Srivastava, S. R. (1969) "*Bias and Mean Square of an Estimation Procedure after Two Preliminary Tests of Significance in ANOVA Model-I*," Sankhya, Series A, 31, 319-332.

38. Gurland, J. and McCullough, R. S. (1962) "*Testing Equality of Means after a Preliminary Test of Equality of Variance*," Biometrika, 49, 403-417.

39. Han, Chein-Pai (1978) "*Nonnegative and Preliminary Test Estimators of Variance Components*," Journal of the American Statistical Association, 73, 855-858.

40. Han, Chien-Pai (1980) "*1980 On Estimating the Variance of the Sample Mean in Random Models Based on Conditional Specification,*" Communications in Statistics, Part A--Theory and Methods, A9, 1417-1425.

41. Han, Chein-Pai (1983) "*A Study of the Robustness of Preliminary Tests in Conditionally Specified Regression Models,*" ASA Proceedings of Business and Economic Statistics Section, 140-143.

42. Han, Chein-Pai, Rao, C. V. and Ravichandran, J. (1988) "*Inference Based on Conditional Specification:A Second Bibliography,*" Communications in Statistics, Part A--Theory and Methods, A17, 1945-1964.

43. Hill, R. Carter, Judge, George G. and Fomby, Thomas B. (1978) "*On Testing the Adequacy of a Regression Model,*" Technometrics, 20, 491-494.

44. Hirano, Katuomi (1978) "*A Note on the Level of Significance of the Preliminary Test in Pooling Variances,*" Journal of the Japanese Statistical Society, 8, 71-76.

45. Hirano, Katuomi (1978) "*On Level of Significance of the Preliminary Test in Pooling Means,*" Annals of the Institute of Statistical Mathematics, 30, 1-8.

46. Holms, Arthur G. (1980) "*Chain Pooling Model Selection for Two Level Fixed Effects Factorial Experiments*," Communications in Statistics, Part B--Simulation and Computation, B9, 51-71.

47. Huntsberger, D. V. (1955) "*A Generalization of a Preliminary Testing Procedure for Pooling Data*," Annals of Mathematical Statistics, 26, 734-743.

48. Inada, Koichi (1984) "*A Minimax Regret Estimator of a Normal Mean After a Preliminary Test*," Annals of the Institute of Statistical Mathematics, 36, 207-215.

49. Jain, R. C. (1978) "*Bias and Mean Square Error of a Variance Estimator After Three Preliminary Tests of Significance*," Indian Journal of Pure and Applied Mathematics, 9, 1-16.

50. Johnson, J. P., Bancroft, T. A. and Han, Chein-Pai (1977) "*A Pooling Methodology for Regressions in Prediction*," Biometrics, 33, 57-68.

51. Johri, S. K. and Gupta, V. P. (1981) "*Preliminary Test Estimation in Three-way Layout with Random Effects Model*," Journal of the Indian Statistical Association, 19, 51-62.

52. Kale, B. K. and Bancroft, T. A. (1967) "*Inference for Some Incompletely Specified Models Involving Normal Approximations to Discrete Data*," Biometrics, 1, 335-348.

53. Kasagi, Fumiyoshi (1978) "*A note on Bias Reduction of a Preliminary Test Estimator*," Memoirs of the Faculty of Science, Series A, Math, Kyushu University, 32, 319-323.

54. Kennedy, W. J. and Bancroft, T. A. (1971) "*Model Building for Prediction in Regression Based on Repeated Significance Tests*," Annals of Mathematical Statistics, 42, 1273-1284.

55. Klotz, J. H., Milton, R. C. and Zacks, S. (1969) "*Mean Square Efficiency of Estimators of Variance Components*," Journal of the American Statistical Association, 64, 1383-1402.

56. Lambert, Annick, Saleh, A. K. Md. Ehsanes and Sen, Pranab Kumar (1985) "*On Least Squares Estimation of Intercepts After a Preliminary Tesst on Parallelism of Regression Lines*," Communications in Statistics, Part A--Theory and Methods, A14, 793-807.

57. Lambert, Annick, Saleh, A. K. Md. Ehsanes and Sen, Pranab Kumar (1985) "*Test of Homogeneity of Intercepts After a Preliminary Test on Parallelism of Regression Lines*," Communications in Statistics, Part A--Theory and Methods, A14, 2243-2259.

58. Larson, H. J. and Bancroft, T. A. (1963) "*Biases in Prediction by Regression for Certain Incompletely Specified Models*," Biometrika, 50, 391-402.

59. Leamer, E. E. and Chamberlain, C. (1976) "*A Bayesian Interpretation of Pre-Testing*," Journal of the Royal Statistical Society, Series B, 38, 85-94.

60. Leonard, Tom (1977) "*A Bayesian Approach to Some Multinomial Estimation and Pretesting Problems*," Journal of the American Statistical Association, 72, 869-874.

61. Leonard, Tom and Ord, Keith (1976) "*An Investigation of the F Test Procedure as an Estimation Short-Cut*," Journal of the Royal Statistical Society, Series B, 38, 95-98.

62. Lin, H. T. (1978) "*Some Alternatives to Preliminary Tests in Problems of Pooling Data*," Journal of the Chinese Statistical Association, 16, 5989-6001.

63. Mead, Ronald, Bancroft, T. A. and Han, Chein-Pai (1975) "*Power of Analysis of Variance Test Procedures for Incompletely Specified Fixed Models*," Annals of Statistics, 3, 797-808.

64. Meeden, Glen and Arnold, Barry C. (1979) "*The Admissibility of a Preliminary Test Estimator When the Loss Incorporates a Complexity Cost*," Journal of the American Statistical Association, 74, 872-874.

65. Mosteller, Fredrick (1948) "*On Pooling Data*," Journal of the American Statistical Association, 43, 231-242.

66. Murphy, Kevin J. and Hofler, Richard A. (1983) "*The Pooling Problem:A Proposed Solution with an Illustration*," ASA Proceedings of Business and Economic Statistics Section, 260-264.

67. Nagata, Yasushi and Araki, Takaharu (1985) "*Practical Pooling Method in fixed ANOVA model, size and power trade off*," Reports of Statistical Application Research, 32(2), 1-9.

68. Nagata, Yasushi and Arki, Takaharu (1985) "*Interval Estimation of a Cell Mean in a Fixed ANOVA Model Under a Preliminary Test*," Reports of Statistical Application Research, 32(1), 1-10.

69. Ohta, Hiroshi (1981) "*A Procedure for Pooling Data by the Analysis of Means*," Journal of Quality Technology, 13, 115-119.

70. Ohtani, Kazuhiro (1983) "*Preliminary Test Predictor in the Linear Regression Model Including a Proxy Variable*," Journal of Japan Statistical Society, 13, 11-19.

71. Ohtani, Kazuhiro and Toyoda, Toshihisa (1978) "*Minimax Regret Critical Values for a Preliminary Test in Pooling Variances*," Journal of the Japan Statistical Society, 8, 15-20.

72. Ohtani, Kazuhiro and Toyoda, Toshihisa (1979) "*Optimal Critical Balues for a Preliminary Test in Pooling Means*," The Australian Journal of Statistics, 21, 157-161.

73. Ohtani, Kazuhiro and Toyoda, Toshihisa (1980) "*Estimation of Regression Coeficients After a Preliminary Test for Homoscedasticity*," Journal of Econometrics, 12, 151-159.

74. Olshen, Richard A. (1973) "*The Conditional Level of the F-Test*," Journal of the American Statistical Association, 68, 692-698.

75. Patel, H. I., Roy, T. and Huque, M. F. (1984) "*Some Preliminary Tests in Repeated Measures Designs*," SAS Users Group, 9, 748-753.

76. Paull, A. E. (1950) "*On a Preliminary Test for Pooling Mean Squares in the Analysis of Variance*," Annals of Mathematical Statistics, 21, 539-550.

77. Pratt, John W. (1976) "*A Discussion of the Question:For What Use Are Tests of Hypothesis and Tests of Significance?*," Communications in Statistics, Part A--Theory and Methods, A1, 779-787.

78. Raghunandanan, K. (1978) "*On Pooling Data to Estimate Variance*," Sankhya, Series B, 40, 94-104.

79. Rao, C. V. and Saxena, P. (1979) "*Bounds of Size of a Test Procedure Based on Two Preliminary Tests of Significance in a Mixed Model*," Sankhya, Series B, 40, 83-93.

80. Rao, C. V. and Saxena, P. (1981) "*On Approximation of Power of a Test Procedure Based on Preliminary Tests of Significance*," Communications in Statistics, Part A--Theory and Methods, A10, 1305-1321.

81. Saleh, A. K. M. E. and Sen, P. K. (1978) "*Nonparametric Estimation of Location Parameter After a Preliminary Test on Regression*," Annals of Statistics, 6, 154-168.

82. Saleh, A. K. Md. E. and Sen, P. K. (1983) "*Asymptootic properties of tests of hypothesis following a preliminary test*," Statistics & Decissions, 1, 455-478.

83. Sanathanan, Lalitha (1976) "*Priliminary Testing and the Minimax Regret Criterion in Regression Analysis*," ASA Proceedings of Business and Economic Statistics Section, 575-578.

84. Saxena, K. P. (1977) "*A Note on Relative Efficiency of Estimators Based on Preliminary Tests in Linear Regression Models*," Trabajos de Estadistica y de Investigacion Operativa, 28, 109-117.

85. Saxena, K. P. (1978) "*On Power of Sometimes Pool Tests and the Bayes Rule for Intercepts in Linear Regression Models Assuming Hetrogeneity of Variances*," Estadistica, 29, 158-166.

86. Sen, Pranab Kumar (1979) "*Asymptotic Properties of Maximum Likelihood Estimators Based on Conditional Specification*," Annals of Statistics, 7, 1019-1033.

87. Shukla, N. D. (1979) "*Estimation of a Regression Coeficient After Two Preliminary Tests of Significance*," Metrika, 26, 183-194.

88. Singh, J. (1971) "*Pooling Mean Squares*," Journal of the American Statistical Association, 66, 82-85.

89. Singh, J. and Srivastava, S. R. (1975) "*Asymptootic Power and Efficiency of a Sometimes Pool Test Procedure in a Component of Variance Model*," Progress of Mathematics, 9, 17-21.

90. Singh, L. and Saxena, P. (1980) "*On Power of a Sometimes Pool Test in a Split Plot Design*," Sankhya, Series B, 42, 153-164.

91. Singh, Umesh and Gupta, V. P. (1976) "*Bias and Mean Square Error of and Estimation Procedure After a Preliminary Test of Significance for a Three-way Layout in ANOVA Model-II*," Sankhya, Series B, 38, 164-178.

92. Srivastava, S. R. (1972) "*Pooling Mean Squares in ANOVA Model-II*," Journal of the American Statistical Association, 67, 676-679.

93. Srivastava, S. R. and Bancroft, T. A. (1967) "*Inferences Concerning a Population Coeficient from One or Possibly Two Samples Subsequent to a Preliminary Test*," Journal of the Royal Statistical Society, Series B, 29, 282,291.

94. Steel, Robert G.D. and Torrie, James H. (1980) "*Principles and Procedures of Statistics* " (2nd. ed.), eds. C. Napier and J.W. Maisel, New York: McGraw-Hill, pp. 199.

95. Stein, Charles (1964) "*Inadmissibility of the Usual Estimator for the Variance of a Normal Distribution with Unknown Mean*," Annals of the Institute of Statistical Mathematics, 16, 155-160.

96. Tailor, J. K. and Saxena, K. P. (1974) "*Approximate Formulae for Power Components of Sometimes Pool Test Procedure Based on Mixed Model*," Sankhya, Series B, 36, 267.

97. Thompson, W. A. Jr. (1962) "*The Problem of Negative Estimates of Variance Components*," Annals of Mathematical Statistics, 33, 273-289.

98. Vassar, Thomas B. II and Carter, Walter H. Jr. (1978) "*Minimizing Average Mean Squared Error:A Comparison of Preliminary Test Estimation to Other Procedures*," Journal of Statistical Computation and Simulation, 7, 55-70.

99. Wang, Ying Y. (1967) "*A Comparison of Several Variance Component Estimators*," Biometrika, 54, 301-305.

100. Wolde-Tsadik, Girma and Afifi, A. A. (1980) "*A Comparisom of the "Sometimes Pool", "Sometimes Switch", and "Never Pool" Procedures in the Two-way ANOVA Random Effects Model*," Technometrics, 22, 367-373.