DOCUMENT RESUME

ED 061 282 TM 001 233

AUTHOR Pohlmann, John T.
TITLE The Effects of Pooling the Interaction and Within Components on the Alpha and Power for Main Effects Tests.
PUB DATE Apr 72
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Analysis of Variance; Computer Programs; Factor Analysis; *Hypothesis Testing; *Interaction; Mathematical Models; *Statistical Analysis; *Tests of Significance
IDENTIFIERS *Monte Carlo Method

ABSTRACT

The Monte Carlo method was used, and the factors considered were (1) level of main effects in the population; (2) level of interaction effects in the population; (3) alpha level used in determining whether to pool; and (4) number of degrees of freedom. The results indicated that when the ratio degrees of freedom (axb)/degrees of freedom (within) was large (1/4), pooling resulted in a disturbance in the actual alpha for the main effect test. The magnitude and nature of the disturbance was dependent on the alpha level employed in testing the interaction effects. The use of an alpha of .25 for the interaction effects resulted in a congruence between actual alpha and nominal alpha, and a slight increase in power. (Author/CK)

THE EFFECTS OF POOLING THE INTERACTION AND WITHIN COMPONENTS ON THE ALPHA AND POWER FOR MAIN EFFECTS TESTS

John T. Pohlmann

Southern Illinois University at Carbondale

Abstract

The Monte Carlo method was used, and the factors considered were (1) level of main effects in the population; (2) level of interaction effects in the population; (3) alpha level used in determining whether to pool; and (4) number of $d.f._{within}$. The results indicated that when the ratio $d.f._{axb}/d.f._{within}$ was large (1/4), pooling resulted in a disturbance in the actual alpha for the main effect test. The magnitude and nature of the disturbance was dependent upon the alpha level employed in testing the interaction effects. The use of an alpha of .25 for the interaction effects resulted in a congruence between actual alpha and nominal alpha, and a slight increase in power.

THE EFFECTS OF POOLING THE INTERACTION AND WITHIN COMPONENTS ON THE ALPHA AND POWER FOR MAIN EFFECTS TESTS

John T. Pohlmann

Southern Illinois University at Carbondale

Objectives of the Inquiry

The purpose of this study was to determine the effects of pooling the interaction component with the within component, on the actual alpha and power of the resulting main effects tests. In this study, only the two-way fixed effects ANOVA, with two levels of each main effect, was considered. The initial structural mode' including interaction, was

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk} \qquad (1)$$

where i = 1, 2; j = 1, 2; k = 1,$\cdots$ (2, 3, or 4); and ε_{ijk} as the error term. The revised structural model, excluding interaction, was

$$X_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} \qquad (2)$$

where i = 1, 2; j = 1, 2; k = 1,$\cdots$ (2, 3, or 4), as in model (1), but where ε_{ijk} is the pooled interaction and within terms. In terms of sample values, the pooled error term is given by

$$MS_{error} = (SS_{axb} + SS_{within}) / (d.f._{axb} + d.f._{within})$$

Decisions about what terms should be included in the model are based upon either theory or experience (previous research). In the absence of theory and prior knowledge, a

researcher must rely on the data available in his experiment as a basis for revising an initial model. Tests undertaken to revise a structural model, given the data, are called preliminary tests. For example, an insignificant preliminary test of the interaction term in model (1), at some specified alpha level, would result in pooling the interaction term with the within term to give model (2).

Previous research in the area of preliminary tests suggested that two schools of thought prevail. The first school of thought may be characterized as the "never pool" school. Proponents of this line of reasoning include Ostle (1963), and Scheffe (1959). Ostle (1963) reasoned that since pooling results in a disturbance in the actual alpha level for other tests in the experiment, the interaction component should never be pooled; no matter what the result of the F-test for interaction. Sheffe (1959) stated that since little is known about the operating characteristics of such procedures, researchers should try to avoid pooling by designing the experiment such that there will be a sufficient number of error degrees of freedom.

The other school of thought may be characterized as the "sometimes pool" school. Proponents of this position, (Bennett & Franklin (1954); Bosovich, Bancroft & Hartley (1956); and Singh (1970)) reason that pooling is an acceptable procedure as long as the researcher is aware of the consequences of the pooling decision and can specify the

conditions under which pooling leads to tests where actual alpha and nominal alpha are in congruence, and which increase in power. Each of the proponents of the "sometimes pool" approach offer decision rules which a researcher may use in determining whether or not to pool. Bennett and Franklin (1954) state that a researcher should pool the interaction SS into error, if and only if the F-test for the interaction is not significant at the 5 percent level and the F-test is less than 2. Bosovich, Bancroft and Hartley (1956), in a study of the mixed effects model ANOVA, found that if the F-test for interaction is not significant at the .25 level, pooling results in a final F-test, when made at a nominal 5 percent level, which is very close to the 5 percent level. Singh (1970) found that the use of a preliminary alpha of .50 resulted in an err term that was least biased.

This study examined the effects of using various alpha levels in the preliminary test on the magnitude of the actual alpha and power of the final test of the main effects. Further, this study considered the effects of using these various preliminary alpha's when model (1) was correct ($E(MS_{axb}) > E(MS_{within})$, and when model (2) was correct ($E(MS_{axb}) = E(MS_{within})$.

Method

The design for this study was a 4X4X3X6 factorial design. The factors where (1) four levels of the magnitude of main effects in the population ($\phi^* = 0, 1, 2, 3$); (2) four levels

of the magnitude of the interaction effects in the population (ϕ = 0, 1, 2, 3); (3) three levels of degrees of freedom within (4, 8, 12), and (4) six alpha levels for the preliminary tests (.00 (always pool), .05, .10, .25, .50, and 1.00 (never pool)). In all cases there were two levels of each main effect and consequently $d.f._{axb}$ was always 1. Only a nominal alpha of .05 was used in testing the main effects.

Data Source

The Monte Carlo method was used in the conduct of this investigation. A computer program, written in FORTRAN IV, was developed by the author for the purpose of this study. Input for the program consisted of (1) the number of levels in factors A and B; (2) the number of replications (cell n); (3) a matrix of treatment effects for each treatment combination. These treatment effects were added to the random numbers as they were drawn from the random number generator. IBM subroutine GAUSS was used to generate normally distributed random numbers. Subroutine GAUSS sums 12 uniformly random values generated by the power residue method in subroutine RANDU. GAUSS then performs the appropriate linear transformations to give a normally distributed random number from a population with the parameters required by the user. The cell into which the random number was placed, determined which treatment effect was administered. The magnitude of the treatments effects was determined by the population values of the main effects and intereaction effects. Each combination of the population values of the main effects,

interaction effects and cell size was run for 2000 trials. The output from the program consisted of (1) the number of times, in 2000 trials, that the null hypothesis for the main effects was rejected at the .05 level, for each of the alpha levels used in the preliminary test. This output was converted into a proportion that approximated a limiting relative frequency that could be interpreted as a probability of rejecting the null hypothesis. When the null hypothesis was true with respect to the main effects, this value was interpreted as an actual alpha level, and when the null hypothesis was false with respect to the main effects, this value was interpreted as the power of the test.

Results

In Table 1 the results are presented in terms of the proportion of times in 2000 trials that the null hypothesis, for main effects, was rejected. The nominal alpha was .05 for all analyses.

The results are summarized below according to the relative magnitude of $d.f._{axb}$ to $d.f._{within}$, and the alpha level employed in the preliminary test. These results refer to the actual alpha and power associated with the main effects test under the various preliminary test conditions.

I. The Relative Magnitude of $d.f._{axb}$ to $d.f._{within}$

A. The largest disturbance in the magnitude of actual alpha and power occurred when the ratio $d.f._{axb}$ / $d.f._{within}$ was 1/4.

B. The least disturbance in the magnitude of actual alpha and power occurred when the ratio $d.f._{axb}$ / $d.f._{within}$ was 1/12.

II. The Alpha Level Employed in the Preliminary Test

A. The largest disturbance in actual alpha and power occurred under the alpha = .00 (always pool) condition. The use of this alpha level resulted in a consistent reduction in actual alpha and a loss of power, when interaction was present in the population.

B. When the ratio $d.f._{axb}$ / $d.f._{within}$ was 1/12, the use of alpha = .25 resulted in a congruence between actual alpha and nominal alpha, and a slight increase in power when the population interaction effect was small ($\phi < 1.0$). When the population interaction effect was large ($\phi > 2.0$) actual power and nominal power converged.

Conclusions and Implications

On the basis of these results, it was concluded that (1) when the ratio $d.f._{axb}$ / $d.f._{within}$ is relatively large (about 1/4), pooling becomes a questionable procedure due to the disturbance in Type I error rate; (2) when the ratio $d.f._{axb}$ / $d.f._{within}$ is relatively small (about 1/12) pooling, using a preliminary alpha of .25 results in a congruence between actual alpha and nominal alpha and a slight increase in the power of the main effects test. These results are

in keeping with those of Bosovich, Bancroft and Hartley (1956), in that a preliminary alpha of .25 resulted in main effects test that were superior to those when other preliminary alpha levels were used.

It is therefore suggested that a researcher, confronted with the decision of whether to pool or not, use the flow chart presented in Figure 1 to assist him in his decision making.

Table 1

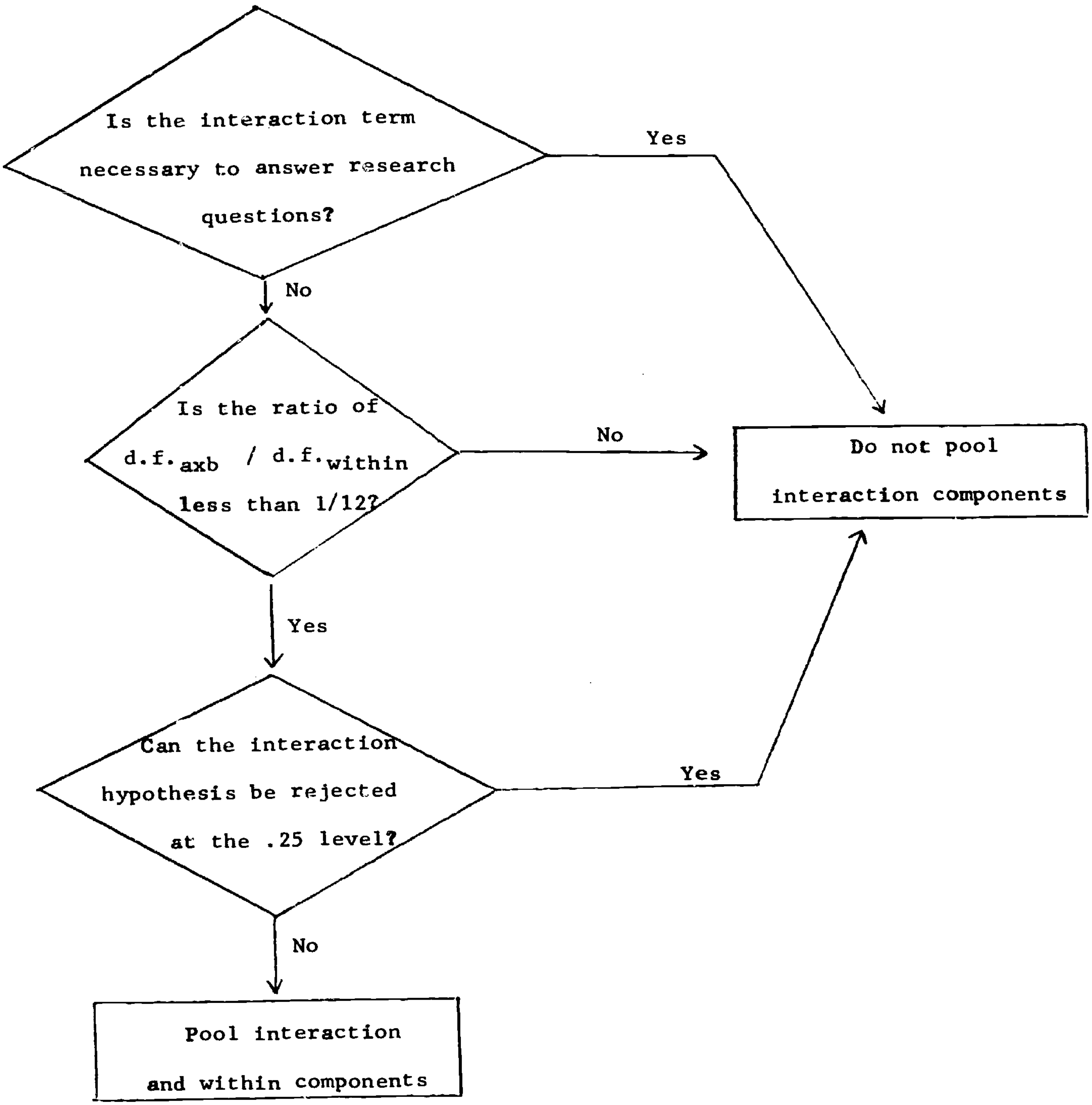
The Proportion of Rejected Null Hypotheses, Out of 2000 Trials, for the Main Effects Test, Using Six Alpha Levels for the Preliminary Test of the Interaction Effects ($d.f._{axb} = 1$)

Interaction Effects

			ϕ = 0						ϕ = 1.0						ϕ = 2.0						ϕ = 3.0					
			Preliminary Alpha						Preliminary Alpha						Preliminary Alpha						Preliminary Alpha					
		d.f. within	AP .00	.05	.10	.25	.50	NP 1.0	AP .00	.05	.10	.25	.50	NP 1.0	AP .00	.05	.10	.25	.50	NP 1.0	AP .00	.05	.10	.25	.50	NP 1.0
Main Effects	ϕ = 0	4	.05	.06	.06	.07	.06	.05	.03	.05	.06	.06	.06	.05	.01	.05	.05	.05	.05	.05	.00	.05	.05	.05	.05	.05
		8	.05	.05	.06	.06	.06	.05	.04	.05	.05	.06	.06	.05	.02	.05	.05	.05	.05	.05	.01	.05	.05	.05	.05	.05
		12	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.03	.05	.05	.05	.05	.05	.02	.05	.05	.05	.05	.05
	ϕ = 1.0	4	.21	.22	.23	.23	.22	.18	.14	.19	.20	.21	.20	.18	.04	.17	.18	.19	.18	.18	.01	.18	.18	.18	.18	.18
		8	.24	.25	.25	.25	.25	.23	.21	.23	.24	.24	.24	.23	.14	.21	.23	.23	.23	.23	.08	.23	.24	.24	.24	.24
		12	.26	.27	.27	.28	.27	.26	.23	.26	.26	.27	.26	.26	.18	.25	.26	.26	.26	.26	.12	.26	.27	.27	.27	.27
	ϕ = 2.0	4	.63	.64	.65	.66	.64	.58	.49	.57	.60	.62	.61	.58	.21	.53	.57	.58	.58	.58	.05	.56	.57	.58	.58	.57
		8	.72	.73	.73	.74	.73	.71	.69	.71	.72	.73	.73	.71	.58	.69	.70	.71	.71	.71	.41	.69	.71	.71	.71	.71
		12	.73	.75	.76	.76	.76	.74	.72	.74	.74	.75	.75	.74	.64	.72	.74	.74	.74	.74	.53	.73	.74	.74	.74	.74
	ϕ = 3.0	4	.92	.93	.93	.93	.92	.89	.84	.88	.90	.91	.90	.89	.56	.83	.88	.89	.89	.89	.22	.86	.88	.89	.89	.89
		8	.96	.96	.96	.96	.96	.96	.94	.94	.95	.95	.95	.94	.92	.95	.96	.96	.96	.96	.83	.94	.95	.95	.95	.95
		12	.97	.98	.98	.98	.98	.97	.97	.97	.98	.98	.98	.97	.96	.98	.98	.98	.98	.98	.91	.97	.97	.97	.97	.97

Figure 1

Flow Chart Outlining Pooling Decision Rules



Footnote

* ϕ is a non-centrality parameter for the F-distribution. In the central F distribution, $\phi = 0$. As the expected value of the F distribution increases, ϕ increases proportionally.

References

Bennett, C. A. and Franklin, N. L. Statistical Analysis in Chemistry and the Chemical Industry, New York: John Wiley, 1954.

Bozovich, H., Bancroft, T. A., and Hartley, H. O. Power of analysis of variance test procedures for certain incompletely specified models. Annals of Mathematical Statistics, 1956, 27, 1017-1043.

Ostle, B. Statistics in research, Ames Iowa: Iowa State University Press, 1963.

Sheffe, H. The analysis of variance, New York: John Wiley, 1959.

Singh, J. Pooling mean squares. Journal of the American Statistical Association, 1971, 66 (333), 82-85.