INVESTIGATING POWER OF ANALYSIS OF COVARIANCE METHODS

James R. Schwenke

J. Mark Donovan

Applied Statistics in Agriculture

INVESTIGATING POWER OF ANALYSIS OF COVARIANCE METHODS

James R. Schwenke, J. Mark Donovan
Covance Periapproval Services, Inc.
One Radnor Corporate Center, Radnor, Pennsylvania 19087

Introduction

Analysis of covariance is a well-utilized statistical methodology. The procedure involves a series of statistical tests to first construct a most significant analysis model to characterize the effect of the covariate on response. Pairwise comparisons among treatments are then based on the finalized model.

For traditional Normal error assumptions, each step of the process is based on exact statistical tests. However, the series of statistical tests defines a conditional probability scheme with possible multiplicity issues. The question then becomes if the analysis of covariance methodology considered in entirety is able to maintain a nominal level of significance with good power.

Several procedures have been proposed in the literature suggesting different sequences of tests and understandings of analysis of covariance. This simulation study is being conducted to compare among a number of these analysis strategies. The initial goal was to investigate power of detecting treatment differences using the various analysis of covariance strategies. But, before power can be considered, the ability of the methodology to maintain a nominal level of significance must be investigated.

Simulation Strategy

What follows is a summary of a simulation conducted to investigate the ability of various analysis of covariance testing strategies to maintain a nominal (0.05) level of significance. Instead of reporting the traditional one simulation run of 1000 iterations, five runs of 1000 iterations are reported here. It is felt that this gives a better representation of variation among simulation runs and thus a better summary of the characteristics of the testing procedures.

For each simulation, a 0.05 level of significance was set. A $N(0, 5)$ error distribution was arbitrarily assumed. To investigate the level of significance, four treatment groups were assumed with zero treatment means (intercepts) and zero covariate effects (slopes). A sample size of 25 per treatment group was studied. For each set of simulations, the seed number was held constant to allow direct comparison among runs.

Several statistical tests with known properties were included to evaluate how well each simulation run adhered to the analysis assumptions and how well each characterized a 0.05 level of significance. The normality assumption was tested using the Shapiro-Wilk test and homogeneity of variances among treatment groups was tested using a pragmatic extension of Levene's test. In addition, the traditional F-tests for significant covariate slopes (H_0: $\beta_i = 0$, for all i), heterogeneous covariate slopes (H_0: $\beta_i = \beta_j$, i $\neq$ j), and homogeneous covariate slopes (H_0: $\beta = 0$) were considered, separately.

Analysis of Covariance Methods

Twelve analyses of covariance testing strategies were compared in this simulation to investigate the ability to maintain a nominal level of significance. Each method was given a name for easy reference. The primary analysis of covariance methods of interest are defined below.

COMPLETE:
  1) test significant covariate effect (H_0: $\beta_i = 0$, for   all i)

       a) if not significant, conduct pairwise comparisons among treatments assuming an analysis of variance model

       b) if significant, continue

   2) test heterogeneous covariate effects (H_0: $\beta_i = \beta_j$, $i \neq j$)

       a) if not significant, continue

       b) if significant, conduct pairwise comparisons among treatments at 25th, 50th, and 75th percentile of covariate effect

   3) test homogeneous covariate effect (H_0: $\beta = 0$)

       a) if not significant, conduct pairwise comparisons among treatments assuming an analysis of variance model

       b) if significant, conduct pairwise comparisons among treatments adjusted for common covariate effect

TRADITIONAL:

   1) test heterogeneous covariate effects (H_0: $\beta_i = \beta_j$, $i \neq j$)

       a) if not significant, continue

       b) if significant, conduct pairwise comparisons among treatments at 25th, 50th, and 75th percentile of covariate effect

   2) test homogeneous covariate effect (H_0: $\beta = 0$)

       a) if not significant, conduct pairwise comparisons among treatments assuming an analysis of variance model

       b) if significant, conduct pairwise comparisons among treatments adjusted for common covariate effect

COVARIATE:

   1) test heterogeneous covariate effects (H_0: $\beta_i = \beta_j$, $i \neq j$)

       a) if not significant, assume homogeneous covariate effect model and conduct pairwise comparisons among treatments adjust for common covariate effect

       b) if significant, conduct pairwise comparisons among treatments at 25th, 50th, and 75th percentile of covariate effect

MIXED MODELS:

   1) test significant covariate effect (H_0: $\beta_i = 0$, for   all i)

       a) if not significant, go to step #3

       b) if significant, continue

   2) test heterogeneous covariate effects (H_0: $\beta_i = \beta_j$, $i \neq j$)

       a) if not significant, continue

       b) if significant, conduct pairwise comparisons among treatments at 25th, 50th, and 75th percentile of covariate effect

   3) test homogeneous covariate effect (H_0: $\beta = 0$)

       a) if not significant, conduct pairwise comparisons among treatments assuming an analysis of variance model

       b) if significant, conduct pairwise comparisons among treatments adjusted for common covariate effect

Additional analysis of covariance testing strategies considered included the following.

HOMOGENEOUS:
1) test homogeneous covariate effect (H_0: $\beta = 0$)
    a) if not significant, conduct pairwise comparisons among treatments assuming an analysis of variance model
    b) if significant, conduct pairwise comparisons among treatments adjusted for common covariate effect

HETEROGENEOUS:
1) assume heterogeneous covariate effect model
    a) conduct pairwise comparisons among treatments at 25th, 50th, and 75th percentile of covariate effect

ANCOVA:
1) assume homogeneous covariate effect model
    a) conduct pairwise comparisons among treatments adjusted for common covariate effect

ANOVA:
1) assume analysis of variance model

For the COMPLETE, TRADITIONAL, COVARIATE, and MIXED MODELS strategies, two variations of each were investigated. For each of these strategies, if a heterogeneous covariate model is assumed, comparisons are conducted at the 25th, 50th, and 75th percentile of the covariate. Traditional testing strategies have pairwise treatment comparisons made without further protection for overall treatment differences at each percentile. This testing strategy is reported here as the "unprotected" results. The "protected" results incorporates a test for overall treatment differences at each percentile of the covariate. Pairwise treatment comparisons are considered only if the overall test for treatment differences is significant.

For each testing strategy, the Least Significant Difference (LSD) and Tukey's Honest Significant Difference (HSD) multiple comparison testing procedures are considered.

Results to Date

Initial tests for normality, homogeneity of variances, significant covariate effects (slopes), heterogeneous covariate effects, and homogeneous covariate effects reported in Table 1 indicate that each individual test maintains a 0.05 level of significance, as was expected. The empirical levels of significance based on 5000 iterations ranged from 0.0470 to 0.0528.

Results of the simulations using the LSD multiple comparison procedure reported in Table 2 indicate that none of the primary testing strategies adequately maintain the nominal level of significance. Of the eight primary testing strategies of interest, overall simulation results based on 5000 iterations ranged from 0.0618 to 0.0908.

Of the additional testing strategies considered, the analysis of variance (ANOVA) testing strategy did maintain the level of significance. This is to be expected because these simulations were conducted assuming no treatment or covariate effects. The HOMOGENEOUS and ANCOVA procedures assume a common covariate effect. Because the common covariate effect assumption utilizes only one degree of freedom, it is not surprising that the level of significance is maintained for these testing strategies.

Results of the simulations using Tukey's HSD multiple comparison procedure reported in Table 3 indicate an improvement in maintaining the nominal level of significance. Of the eight primary testing strategies of interest, overall simulation results based on 5000 iterations ranged from 0.0532 to 0.0618. The

COMPLETE and MIXED MODELS testing strategies did maintain the level of significance, with the MIXED MODELS procedure indicating a slightly higher level.

All of the additional testing strategies considered, the HOMOGENEOUS, HETEROGENEOUS, ANCOVA, and ANOVA testing strategies, were conservative.

Further Research

From the complete simulation results, it is evident that no one testing strategy, as currently defined, is able to maintain the nominal level of significance over a broad range of conditions. It is conjectured that because each procedure incorporates a series of tests, the conditional relationship of subsequent tests must be considered.

Further research is directed toward incorporating sequential testing strategies into the various analysis of covariance testing procedures.

The LSD procedure protects against the weak null hypothesis of no treatment effects (H_0: $\mu_1=\mu_2=\mu_3=\mu_4$). The weak null hypothesis is not appropriate if a heterogeneous covariate effect model is assumed. A heterogeneous covariate structure necessitates at least one treatment difference at some level of the covariate. This defines the strong null hypothesis which Tukey's HSD procedure protects against.

In this simulation study, either the LSD or HSD multiple comparison procedure was used to make all pairwise treatment comparisons, regardless of the covariate model assumed. From the results of this simulation, it is suggested that the LSD procedure be used to conduct pairwise comparisons if the homogeneous analysis of covariance model or the analysis of variance model is assumed. Tukey's HSD procedure should be used if the heterogeneous analysis of covariance model is assumed. Thus, the results of this simulation study indicate bounds on the level of significance.

Once the nominal level of significance can be reasonably maintained, a power study will be conducted considering a broad range of conditions.

Table 1

Analysis of Covariance -- Simulation Study Results
Probability of Detecting Residual Assumptions and Model Specification

Number of Iterations = 1000, Number of Observations = 25
Error Variance: N[0, 5], Level of Significance = 0.05
Covariate Distribution: N[5, 5]
Treatment Means: 0 0 0 0
Treatment Slopes: 0 0 0 0

Residual Analysis		
Simulation Run	Normality	Homogeneity of Variances
1	0.041	0.044
2	0.041	0.068
3	0.054	0.043
4	0.046	0.054
5	0.053	0.055
Combined	0.0470	0.0528

Probability of Detecting Covariate Effect			
Simulation Run	Significant Slopes	Heterogeneous Slopes	Homogeneous Slopes
1	0.041	0.039	0.048
2	0.049	0.053	0.048
3	0.055	0.049	0.047
4	0.041	0.043	0.057
5	0.053	0.062	0.047
Combined	0.0478	0.0492	0.0494

Table 2

Probability of Detecting Differential Treatment Effects
Using Fisher LSD Multiple Comparison Technique

Simulation Run	Complete Protected	Complete Unprotected	Traditional Protected	Traditional Unprotected
1	0.064	0.072	0.068	0.086
2	0.063	0.079	0.071	0.097
3	0.065	0.078	0.070	0.089
4	0.050	0.064	0.058	0.075
5	0.067	0.083	0.078	0.102
Combined	0.0618	0.0752	0.0690	0.0898

Simulation Run	Covariate Protected	Covariate Unprotected	Mixed Models Protected	Mixed Models Unprotected
1	0.068	0.086	0.064	0.072
2	0.075	0.101	0.063	0.079
3	0.070	0.089	0.067	0.080
4	0.059	0.076	0.052	0.066
5	0.078	0.102	0.068	0.084
Combined	0.0700	0.0908	0.0628	0.0762

Simulation Run	Homogeneous Covariate	Heterogeneous Covariate	ANCOVA	ANOVA
1	0.052	0.105	0.051	0.052
2	0.051	0.112	0.055	0.050
3	0.050	0.102	0.050	0.048
4	0.043	0.088	0.044	0.040
5	0.048	0.111	0.048	0.046
Combined	0.0488	0.1036	0.0496	0.0472

Table 3

Probability of Detecting Differential Treatment Effects
Using Tukey HSD Multiple Comparison Technique

Simulation Run	Complete Protected	Complete Unprotected	Traditional Protected	Traditional Unprotected
1	0.057	0.058	0.061	0.064
2	0.056	0.058	0.064	0.067
3	0.055	0.056	0.060	0.062
4	0.045	0.048	0.050	0.053
5	0.053	0.054	0.062	0.063
Combined	0.0532	0.0548	0.0594	0.0618

Simulation Run	Covariate Protected	Covariate Unprotected	Mixed Models Protected	Mixed Models Unprotected
1	0.060	0.063	0.057	0.058
2	0.066	0.069	0.056	0.058
3	0.061	0.063	0.057	0.058
4	0.048	0.051	0.046	0.049
5	0.062	0.063	0.054	0.055
Combined	0.0594	0.0618	0.0540	0.0556

Simulation Run	Homogeneous Covariate	Heterogeneous Covariate	ANCOVA	ANOVA
1	0.047	0.093	0.045	0.047
2	0.045	0.095	0.047	0.045
3	0.042	0.084	0.043	0.040
4	0.036	0.076	0.034	0.035
5	0.038	0.089	0.038	0.036
Combined	0.0416	0.0874	0.0414	0.0406