DEPARTMENT OF STATISTICS

The Ohio State University

OSU

COLUMBUS, OHIO

82   03   22   005

STRONGER TESTS USING HETEROSCEDASTIC ANOVA
WITH A SIMULATION OF MULTIPLE OBJECTIVE BUDGETING
CASE STUDY

by

Edward J. Dudewicz[1]

and

W. Thomas Lin[2]

Technical Report No. 250A
Department of Statistics
The Ohio State University
Columbus, Ohio 43210
March 1982

STRONGER TESTS USING HETEROSCEDASTIC ANOVA
WITH A SIMULATION OF MULTIPLE OBJECTIVE BUDGETING
CASE STUDY

The traditional F-tests assume that all the population variables under consideration have unknown means with equal unknown variances. The experimenter can control Type I error but not Type II error under the traditional F-test simulation experiment.[1]

The objective of this note is to show that the problem of controlling Type II error can be solved by using a new F-test, which also allows for unequal unknown population variances. The traditional analysis of variance (ANOVA) is based on the assumptions of normality, independence of the statistical errors, and equality of the variances of the errors. Studies of the robustness of the F-test have shown that the violation of normality has little effect on inferences about the means. However, the violation of independence or equality of variances can have a serious effect on inferences about the means, especially if the cell sample sizes are unequal (see, for example, Scheffé [1959] or Bishop [1976]). In practice, the assumption of equality of error variances seems to be often unjustified; in fact, even when the error variances are equal, the power of the F-test depends upon the unknown common variance, which renders it difficult to plan an experiment rationally. Recently, Bishop and Dudewicz [1978] [1981] developed new ANOVA procedures in the contexts of the one-way layout and higher-way layouts. Their procedures allow unequal and unknown population variances and give tests with level and power completely independent of the unknown variances.

The next section briefly outlines the two-way layout heteroscedastic-ANOVA (HANOVA) methodology. It is followed by case study

of Lin's [1978] multiple objective firm simulation for the two-way layout ANOVA procedure.[2] Finally, the conclusion and other possible business applications are outlined.

THE HANOVA PROCEDURE

Bishop and Dudewicz [1978] [1981] have developed heteroscedastic ANOVA (HANOVA) procedures in an r-way layout. For the purpose of this paper, only the two-way layout will be described.

The two-way layout under consideration is what is usually studied in ANOVA and is defined by

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk} \tag{1}$$

$$(i = 1, \ldots, I; \; j = 1, \ldots, J; \; k = 1, 2, \ldots, N),$$

where $\{X_{ijk}\}$ are the observed responses; it is assumed that the $\{X_{ijk}\}$ are independent and normally distributed with unknown mean $E(X_{ijk}) = \mu$ and unknown variance $Var(X_{ijk}) = \sigma_{ij}^2$; $\{e_{ijk}\}$ are assumed to be independent random variables with normal distributions with mean 0 and variance σ_{ij}^2, denoted $e_{ijk} \sim N(0, \sigma_{ij}^2)$, $0 < \sigma_{ij}^2 < \infty$; μ is the overall mean; α_i and β_j are the main effects of factors i and j; $\alpha\beta_{ij}$ is the interaction between i and j; and that

$$\sum_{i=1}^{I} \alpha_i = \sum_{j=1}^{J} \beta_j = \sum_{i=1}^{I} \alpha\beta_{ij} = \sum_{j=1}^{J} \alpha\beta_{ij} = 0.$$

The hypotheses of interest are

$$H_0: \alpha_i = 0 \text{ for all i} \tag{2}$$

$$H_1: \beta_j = 0 \text{ for all j} \tag{3}$$

$$H_2: \alpha\beta_{ij} = 0 \text{ for all i and j} \tag{4}$$

In this two-way layout there are I x J possible treatment combinations. Cell (i, j) refers to the combination of level i of the first factor and level j of the second factor. One seeks tests of the

above three hypotheses based on test statistics whose distributions are independent of the unknown variances.

Bishop and Dudewicz [1978] [1981] developed the following HANOVA procedure to test the above hypotheses based on test statistics whose distributions are independent of the unknown variances and both the level (i.e., probability of Type I error) and power (i.e., probability of Type II error) of the test are controllable:

1. Choose the Type I error level, Type II error level, and the tolerance of the difference between populations.

2. Compute the design constant $z > 0$ given the parameters specified in Step 1, and then in each cell (i,j) take an initial sample X_{ij1}, X_{ij2}, $X_{ij3} \cdots X_{ijn_o}$, of size n_o from each of k populations.[3]

3. Compute sample variance s_{ij}^2, the usual unbiased estimate of unknown population variances σ_{ij}^2, and define the final sample size:

$$N_{ij} = \max\{n_o + 1, \ \frac{s_{ij}^2}{z} + 1\}$$

where $[X]$ denotes the largest integer which is smaller than X.

4. Take $(N_{ij} - n_o)$ additional observations from cell (i,j) populations and calculate the weighting coefficients $a_{ij1}, \ \ldots, \ a_{ijN_{ij}}$ such that[4]

$$a_{ij1} = \cdots = a_{ijn_o} = \frac{1 - (N_{ij} - n_o) \, b_{ij}}{n_o}$$

$$a_{ijn_o+1} = \cdots = a_{ijN_{ij}} = b_{ij}$$

and $b_{ij} = \frac{1}{N_{ij}} \ [\ 1 + \frac{n_o \, (N_{ij} \, z - s_{ij}^2)}{\sqrt{(N_{ij} - n_o) \, s_{ij}^2}} \ ]$

-4-

5. Compute cell sample generalized means $\widetilde{\overline{X}}_{ij.} = \sum\limits_{k=1}^{N_{ij}} a_{ijk} X_{ijk}$

where $X_{ij1}, \ldots, X_{ijN_{ij}}$ is the final set of observations for cell

(i,j). Then compute group means for treatments i and j:

$$\widetilde{\overline{X}}_{i..} = \frac{1}{J} \sum\limits_{J=1}^{J} \widetilde{\overline{X}}_{ij.}$$

$$\widetilde{\overline{X}}_{.j.} = \frac{1}{I} \sum\limits_{i=1}^{I} \widetilde{\overline{X}}_{ij.}$$

and grand mean:

$$\widetilde{\overline{X}}... = \frac{1}{IJ} \sum\limits_{i=1}^{I} \sum\limits_{j=1}^{J} \widetilde{\overline{X}}_{ij.}$$

6. Test hypotheses (2),(3),(4) based on the following quadratic forms
    of independent, identically distributed Student's t variates:

$$\widetilde{F}_0 = J \sum\limits_{i=1}^{I} \frac{(\widetilde{\overline{X}}_{i..} - \widetilde{\overline{X}}...)^2}{z} \tag{5}$$

$$\widetilde{F}_1 = I \sum\limits_{j=1}^{J} \frac{(\widetilde{\overline{X}}_{.j.} - \widetilde{\overline{X}}...)^2}{z} \tag{6}$$

$$\widetilde{F}_2 = \sum\limits_{i=1}^{I} \sum\limits_{j=1}^{J} \frac{(\widetilde{\overline{X}}_{ij.} - \widetilde{\overline{X}}_{i..} - \widetilde{\overline{X}}_{.j.} + \widetilde{\overline{X}}...)^2}{z} \tag{7}$$

The test of H_0 proceeds by rejecting H_0 if and only if $\tilde{F}_0 \geq \tilde{F}_{0,n_o}^{\alpha}$ where $\tilde{F}_{0,n_o}^{\alpha}$ is the upper α^{th} percent point of the null distribution of $\tilde{F}_0$. Tests for the other hypotheses are performed in a similar manner, namely reject H_1 (or H_2) if and only if the corresponding statistic $\tilde{F}_1$ (or $\tilde{F}_2$) is greater than the upper α^{th} percent point $\tilde{F}_{1,n_o}^{\alpha}$ (or $\tilde{F}_{2,n_o}^{\alpha}$) of its respective null distribution.[5]

In general, this new procedure uses z to replace traditional F-test's σ^2/N. The power of the test is one of the inputs to determine the value of z. Therefore, the Type II error can be controlled. The new procedure also uses generalized cell means $\tilde{\bar{X}}_{ij}$ to replace the traditional F-test cell sample means. Bishop and Dudewicz [1981] have proved that $\{\tilde{\bar{X}}_{ij}\}$ are independent random variables and $\tilde{F}_0$, $\tilde{F}_1$, $\tilde{F}_2$ are independent of σ_{ij}^2. Therefore, the assumption of equal variances can be eliminated by using the new procedure. Bishop and Dudewicz [1981] also showed that the power of the test using $\hat{F}$ is much better than the power of the traditional F-test when the variances are unequal. Therefore, the new procedure is more efficient than the traditional procedure.

MULTIPLE OBJECTIVE BUDGETING SIMULATION CASE STUDY

Most of the business decisions faced by managers and accountants are multiple criteria or objective decision problems. Lin [1978] conducted a simulation study of a hypothetical firm with multiple objectives of profit and sales. He used the traditional F-tests to examine the effects of alternative planning models and accounting variance analysis techniques on the firm's profit and sales performance. The simulation results showed that profit and sales under multiple objective linear programming are higher than those under goal programming planning model. When comparing ex post accounting variance analysis with traditional variance analysis, the former resulted in higher sales, but there was no significant difference in profits. The major reason was that the observed sample means were very close; hence, the power of the test was very low. Since the variances were unknown, the traditional F-test was unable to set its sample size to control the power. We now apply the new HANOVA test.[6]

The main objective of the simulation experiment is to study the effects of combinations of production planning models and accounting variance analysis techniques on profit and sales performances. This study uses a 2 x 2 factorial design as shown in Figure 1.

Variance Analysis Factor

Planning	Mode 1	Mode 2
Factor	Mode 3	Mode 4

Figure 1, 2^2 Factorial Experiment

The four treatment combinations are:

Mode 1   Using goal programming planning model and traditional variance analysis technique (GP-TRAD)

Mode 2   Using goal programming planning model and ex post variance analysis technique (GP-EXPO)

Mode 3   Using multiple objective linear programming planning model and traditional variance analysis technique (MOLP-TRAD)

Mode 4   Using multiple objective linear programming planning model and ex post variance analysis technique (MOLP-EXPO)

The overall model is a production planning decision. The planning period is assumed to be a month. The total run length for the short-run production decisions is assumed to be two years or twenty-four periods. Sample size is determined by replicating runs of this total run length using different sets of pseudo-random numbers.

Following Bishop and Dudewicz's [1981] recommended approximation procedure for the two-way layout, the limiting distribution of $\tilde{F}_0$ is non-central chi-square with one degree of freedom and noncentrality parameter

$$\lambda_0 = \frac{2 \cdot \sum_{i=1}^{2} \alpha_i^2}{z}$$, denoted by $\chi_1^2(\lambda_0)$. Similarly, it can be shown that the limiting distribution of $\tilde{F}_1$ is noncentral chi-square with one degree of

freedom and noncentrality parameter $\lambda_1 = \frac{2 \cdot \sum_{j=1}^{2} \beta_j^2}{z}$, while the limiting distribution of $\tilde{F}_2$ is noncentral chi-square with one degree of freedom

and noncentrality parameter $\lambda_2 = \frac{\sum_{i=1}^{2} \sum_{j=1}^{2} (\alpha\beta_{ij})^2}{z}$.

For example, in this simulation the level of the test, i.e., probability of Type I error, was assumed to be 0.05, with power at least 0.95, i.e., Type II error probability = 0.05, and the tolerance of difference between modes for profit and sales were assumed to be $30 and $100, respectively.[7] From Haynam, Govindarajulu and Leone's [1970] tables of the cumulative noncentral chi-square distribution, the noncentrality parameters are 12.995 for both profit and sales. Then the standard errors of estimate z for profit and sales are 69.2574 and 769.5267 respectively.

Taking initial sample size of n_o = 30, the sample means and variances are shown in Table 1.[8]

- - - - - - - - - -
Insert Table 1 Here
- - - - - - - - - -

The final sample sizes and other simulation statistics are shown in Table 2.[9]

- - - - - - - - - -
Insert Table 2 Here
- - - - - - - - - -

From the central chi-square table, the critical value for $\chi^2_{1, .05}$ is 3.84. From the simulation outputs, the actual profit under goal programming is significantly different from the profit under multiple objective linear programming planning models, since $\tilde{F}_0$ = 3.99. Lin's [1978] study also showed a significant difference of profit between two models. The actual sales under goal programming is also significantly different from the sales under multiple objective linear programming planning models with $\tilde{F}_0$ = 75.40. This result is consistent with Lin's study. When comparing ex post variance analysis with traditional variance analysis, Lin [1978] found that the former results in statistically higher

means of actual sales per period, but that there is no significant difference on the means of actual profit per period. One major reason for nonsignificance in profit was that the power of the test in this study was low. In this study, the power of test is 95%, the simulated results ($\tilde{F}_1$ = 4.85 for profit and 5.54 for sales) show that there are statistically significant differences on both profit and sales performance between traditional and ex post accounting variance analysis techniques.

In general, Bishop and Dudewicz's [1978] [1981] HANOVA procedures eliminate the necessity of the assumption of equal variances. In addition, the experimenter receives the added benefits of a controllable power function and allocation of additional sample size where it is important. The new simulation results show that multiple objective linear programming results in higher profit and sales than those under goal programming planning models. This is consistent with the theory that the former is an optimizing model while the latter is a satisficing model. When comparing ex post variance analysis with traditional accounting variance analysis, the former results in higher amounts of both profit and sales. This result is consistent with the theory that the ex post variance analysis uses more relevant feedback information if the information is costless.

CONCLUSION

The assumption of equal variances in ANOVA is often in doubt, especially when most business and accounting variables do not come from equal variance populations. Previous accounting or business simulation studies such as Demski's [1967] [1970] implementation effects, Sundem's [1974] evaluating capital budgeting models, Onsi's [1975] simulation of organizational slack, Fellingham, Mock, Vasarhelyi's [1976] information choice, and Magee's [1976] analysis of alternative cost variance models have assumed equal variances and encountered inability to control Type II error. Lin's [1978] simulation of multiple objective budgeting models failed to control Type II error. Unfortunately, this means that a failure to assert a model effect in a simulation of accounting or business systems may be due not to the nonexistence of such effects, but rather to a Type II error.

This paper introduces a new methodology which eliminates the necessity of such an assumption of equal variances. In addition, the experimenter or manager receives the added benefits of a controllable power function and allocation of the additional sample size "where it counts" (i.e., where variability is high and may obscure means). An extension of Lin's [1978] multiple objective budgeting simulation has been illustrated as a case study. The new HANOVA procedures can be applied to many other business simulation experiments. For example, the different inventory costing methods can be considered together with different depreciation methods to test the effects of different factors on profit performance. Other applications include different audit procedures in variable sampling, different job shop scheduling schemes, and different advertising strategies.

-11-

1. Type I error is defined as the error of rejecting the null hypothesis when it is, in fact, true. Type II error is defined as the error of accepting the null hypothesis when it is actually false. Experiments in situations where variances are unknown and probably unequal are called "heteroscedastic experiments." Kleijnen and Naylor [1969] observed that "Only in rare cases can the assumption of a common known variance be expected to hold with computer simulation experiments with models of business and economic systems" (pp. 609-610).

2. The reason to replicate Lin's multiple objective budgeting simulation model is that it is a significant improvement over previous accounting simulations in terms of external validity. It is a more sophisticated model with an improved analytical technique.

3. One may view z as playing the role of σ^2/N, standard error of the mean, in the traditional r-way layout when the errors have equal variance σ^2, σ^2 is known, and N observations are taken in each cell. Bishop and Dudewicz [1978] developed tables for choosing z values in the one-way layout. For the two-way layout z values are related to noncentrality parameters in noncentral chi-square distributions.

4. While other choices of the $\{a_{ij}\}$ are possible, the above are conjectured by Dudewicz and Dalal [1975] to be robust against nonnormality of the errors.

6. The cutoff points $(\tilde{F}_{0, n_o}^{\alpha})$, $(\tilde{F}_{1, n_o}^{\alpha})$, and $(\tilde{F}_{2, n_o}^{\alpha})$ are not tabulated in the literature yet, and thus an approximation is called for. It has been shown by Bishop and Dudewicz [1981] that the limiting distribution of $\tilde{F}_0$ is noncentral chi-square with I-1 degrees of freedom and non-

centrality parameter $\Delta_0 = J \sum_{i=1}^{2} \alpha_i^2 \ / \ z$, denoted by $\chi_{I-1}^2 (\Delta_0)$.

Note that the closeness of the approximation depends in no way upon the σ_{ij}^2. That is, $\tilde{F}_0$, $\tilde{F}_1$, and $\tilde{F}_2$ are independent of unknown population variances.

6. See Lin [1978] for the description of the overall model in terms of planning, operations, and performance evaluation processes. Lin [1980] also showed examples of ex post accounting variance analysis under both goal programming and multiple objective linear programming planning models.

7. That is, $\sum_{i=1}^{2} \alpha_i^2$ and $\sum_{j=1}^{2} \beta_j^2$ are each assumed to be \$30 for profit, and each \$100 for sales.

8. The initial sample size should be large enough to allow Central Limit Theorem-type effects to take hold, and large enough to keep the multiplier z, used to set the second stage sample sizes, relatively small. A number of around 20 seems desirable, while a number less than 10 is not recommended.

9. In Lin's [1978] study, the sample size of the profit variable for all four modes was 126. The power of that test was around 0.80.

REFERENCES

Bishop, T. A. "Heteroscedastic ANOVA, MANOA, and Multiple-Comparisons," Unpublished Doctoral Dissertation, The Ohio State University, 1976.

Bishop, T. A. and E. J. Dudewicz, "Exact Analysis of Variance with Unequal Variances: Test Procedures and Tables," Technometrics (November 1978), pp. 419-430.

Bishop, T. A. and E. J. Dudewicz, "Heteroscedastic ANOVA," Sankhya, 43, 1981, Series B, pp. 40-57.

Demski, J. S., "The Decision Implementation Interface: Effects of Alternative Performance Measurement Models," The Accounting Review (January 1970), pp. 76-87.

Demski, J. S., "Implementation Effects of Alternative Performance Measurement Models in a Multivariable Context," The Acccounting Review (April 1971), pp. 268-278.

Dudewicz, E. J. and S. R. Dalal, "Allocation of Observations in Ranking and Selection with Unequal Variances," Sankhya, 37, 1975, Series B, pp. 28-78.

Fellingham, J. C., T. J. Mock and M. A. Vasarhelyi, "Simulation of Information Choice," Decision Sciences (April 1976), pp. 219-234.

Haynam, G. E., Z. Govindarajulu, and F. C. Leone, "Tables of the Cumulative Non-Central Chi-Square Distribution," in H. L. Harter and D. B. Owen (eds.), Selected Tables in Mathematical Statistics, Vol. I, Markham Publishing Company, Chicago, 1970, pp. 1-78.

Kleijnen, J. P. C. and T. H. Naylor, "The Use of Multiple Ranking Procedures to Analyze Simulations of Business and Economic Systems," Proceedings of the American Statistical Association, Business and Economic Statistical Section, 1969, pp. 605-615.

Lin, W. T., "Multiple Objective Budgeting Models: A Simulation," The Accounting Review (January 1978), pp. 61-76.

Lin, W. T., "An Accounting Control System Structured on Multiple Objective Planning Models," OMEGA, Vol. 8, No. 3, 1980, pp. 375-382.

Magee, R. P., "A Simulation Analysis of Alternative Cost Variance Investigation Models," The Accounting Review (July 1976), pp. 529-544.

Onsi, M., "Simulation of the Economic Factors Affecting Organizational Slack: A Factorial Design," Decision Sciences (January 1975), pp. 78-91.

Scheffé, H., The Analysis of Variance, John Wiley & Sons, Inc., New York, 1959.

Sundem, G. L., "Evaluation Simplified Capital Budgeting Models Using a Time - State Preference Metric," The Accounting Review (April 1974), pp. 306-320.

Table 1

INITIAL SAMPLE MEANS AND VARIANCES

$$(n_o = 30)$$

Variables		Mode 1 (GP-TRAD)	Mode 2 (GP-EXPO)	Mode 3 (MOLP-TRAD)	Mode 4 (MOLP-EXPO)
Actual Profit	Mean	762	792	812	812
Per Period:	Variance	12309	11794	10078	10096
Actual Sales	Mean	6220	6364	6539	6537
Per Period:	Variance	38682	33591	39816	37998

Table 2

SIMULATION STATISTICS

Variables and Modes	Final Sample Size	Weighting Coefficient a_{ij}	Weighting Coefficient b_{ij}	Generalized Cell Mean $\tilde{\overline{X}}_{ij.}$	Other Statistics
Profit:					
Mode 1: GP-TRAD	178	.0052	.0057	762	$\tilde{\overline{X}}... = 788$
Mode 2: GP-EXPO	171	.0051	.0060	796	$\tilde{F}_0 = 3.99$
Mode 3: MOLP-TRAD	146	.0067	.0069	794	$\tilde{F}_1 = 4.85$
Mode 4: MOLP-EXPO	146	.0063	.0070	797	$\tilde{F}_2 = 3.37$
Sales:					
Mode 1: GP-TRAD	51	.0177	.0244	6204	$\tilde{\overline{X}}... = 6388$
Mode 2: GP-EXPO	44	.0213	.0257	6330	$\tilde{F}_0 = 75.40$
Mode 3: MOLP-TRAD	52	.0181	.0208	6506	$\tilde{F}_1 = 5.54$
Mode 4: MOLP-EXPO	50	.0182	.0227	6510	$\tilde{F}2 = 4.84$