DOCUMENT RESUME

ED 164 608 TM 008 175

AUTHOR Porter, Andrew C.
TITLE Precision for Simple Linear Models.
PUB DATE Mar 78
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Analysis of Covariance; *Analysis of Variance; Hypothesis Testing; Measurement Techniques; Predictive Measurement; Research Design; *Statistical Analysis; *Teaching Methods
IDENTIFIERS Error Analysis

ABSTRACT

A teaching aid appropriate for a beginning course on experimental design is presented. The aid is a numerical example which illustrates some of the theoretical interrelations among three competing design analysis strategies for estimating treatment effects in random assignment designs. The analysis strategies considered are analysis of variance (ANOVA) of the posttest, ANOVA of an index of response, and analysis of covariance (ANCOVA). The first relationship illustrated is that all three strategies estimate the same effects. Thus, choice of strategy may rest on precision and other factors not considered here, such as assumptions and robustness to violation of assumptions. The second relationship illustrated was precision; it was considered by comparison of mean square errors. Tables present a summary of the analysis strategies showing effects and error terms symbolically, as well as the numerical example itself. (Author/CP)

PRECISION FOR SIMPLE LINEAR MODELS*

Andrew C. Porter
Institute for Research on Teaching
Michigan State University

Introduction

The purpose of this paper is to present a numerical example which illustrates some of the theoretical interrelationships among several competing analysis strategies. The analysis strategies considered are ANOVA of the posttest, ANOVA of an index of response, and ANCOVA. The context for comparing the analysis strategies is any balanced design having random assignment of experimental units to levels of a fixed independent variable (T), a single random dependent variable (Y), and a single random covariable (X). The numerical example and the comparison of results from competing analysis strategies may facilitate students' understanding of more general relationships.

One of the topics typically considered in a beginning course on experimental design is methods for improving precision, where precision is defined as the standard error of a simple contrast. It is helpful to introduce these methods for random assignment experiments so that the more difficult problems of measuring change are not distracting. Thus, the student starts with the belief that effects estimated by ANOVA of the posttest are of interest and that they may be used to define the null hypotheses that motivated the study. A treatment effect for ANOVA of the posttest is defined

$$\alpha_{Y_{T\cdot}} = \mu_{Y_{T\cdot}} - \mu_{Y\cdot\cdot} ,$$

* Invited paper for Educational Statisticians Symposium, "Tips for Teaching Basic Statistical Concepts", AERA 1978.

where $\mu_{Y_{T.}}$ denotes the population mean of the T^{th} level of the independent variable on the posttest. The null hypothesis can then be stated

$$Ho: \sum_{T=1}^{t} \alpha_{Y_{T.}}^{2} = 0 ,$$

where t denotes the number of levels of T.

ANOVA of the posttest serves as a baseline against which other strategies may be judged. In judging the utility of a new strategy the student first must be convinced that the effects estimated by the new strategy are the ones of interest, i.e., the same as would be yielded by ANOVA of the posttest. Once the validity of the effects has been established, the student should consider precision. There are, of course, other criteria for selecting among design/analysis strategies, e.g., assumptions and robustness to violation of assumptions, but they are beyond the scope of this paper.

Alternative Strategies

ANOVA of index of response is a design/analysis strategy that might be considered in place of ANOVA of the posttest. An index of response is defined

$$Z_{TR} = Y_{TR} - KX_{TR} ,$$

where R denotes replication and K is any constant. The most popular form of index of response is when K equals one, i.e., a gain score. The design for ANOVA of index of response is identical to that for ANOVA of posttest except that information is also available on a covariable X. Once the index has been created it is used as the dependent variable in ANOVA. The effects estimated by ANOVA of Z are

$$\alpha_{Z_{T.}} = \alpha_{Y_{T.}} - K\alpha_{X_{T.}} ,$$

and are easily obtained by substitution. Since given random assignment all $\alpha_{X_{T.}}$ are zero, the student quickly sees that independent of choice of K, the strategy estimates the desired effects. Relevant to precision, the expected value of the mean square error for ANOVA of Z is

$$\sigma_Z^2 = \sigma_Y^2 (1 - \rho_{XY}^2) + a^2 \sigma_X^2 ,$$

where $a = K - \beta_{Y \cdot X}$, ρ_{XY} is the correlation of X and Y and $\beta_{Y \cdot X}$ is the slope of Y on X pooled within levels of T (Porter and Chibucos, 1973). While this expression makes it easy to see that setting $K = \beta_{Y \cdot X}$ yields minimum error variance, it is also true that ANOVA of Z will be more precise than ANOVA of Y whenever

$$0 < K/\beta_{Y \cdot X} < 2.$$

Yet a third design/analysis strategy for improving precision is the analysis of covariance. An ANCOVA effect is defined

$$\alpha'_{Y_{T.}} = \alpha_{Y_{T.}} - \beta_{Y \cdot X}\, \alpha_{X_{T.}} ,$$

where prime denotes adjusted effect. As was true for index of response, the second term of an ANCOVA effect disappears under random assignment and ANCOVA is seen to estimate the effects of interest. The expected value of the mean square error for ANCOVA is

$$\sigma_{Y \cdot X}^2 = \sigma_Y^2 (1 - \rho_{XY}^2) \left[1 + \frac{1}{\text{d.f.} - 2}\right];$$

where d.f. denotes degrees of freedom.

Teaching ANOVA of index of response prior to ANCOVA has some advantages. First, ANOVA of index of response is an intuitive procedure which is computationally straightforward. ANCOVA is computationally mysterious but can be thought of as essentially an index of response in which $\beta_{Y \cdot X}$ is estimated from the sample rather than known a priori.* The smaller the sample the poorer the estimate of $\beta_{Y \cdot X}$, and so the larger $[1 + \frac{1}{d.f.-2}]$ and thus error variance. Conversely, as sample size goes to infinity, the sample estimate of $\beta_{Y \cdot X}$ converges on the parameter, and the error variance reduces to its minimum. Thus, a second advantage of teaching ANOVA of index of response first is that, consistent with intuition, the limiting form of the error variance for ANCOVA is identical to the minimum error variance for index of response. A summary of the effects and variances of the three strategies considered is provided in Table 1.

A Numerical Example

Table 2 contains data that can be used to illustrate the interrelationships presented in Table 1. In the example, there are two treatments and two levels of a second independent variable, say sex (S), crossed with treatments. The design is balanced with five individuals per cell. The data in Table 2 have several properties which facilitate instruction. First, the observations on the posttest (Y) and the covariable (X) are all whole numbers of modest size. Second, the correlation between X and Y is exactly .8 within each cell. Third, the within cell variances of X and Y are equal so that the slope of Y on X is equal to the correlation between Y and X. Finally, the treatment level means on the covariable are exactly equal. While this is unlikely given random assignment, it is consistent with the long run expectation and will facilitate illustration of the Table 1 relationships.

* ANCOVA is not, however, computationally equivalent to ANOVA of $Z = Y - \hat{\beta}_{Y \cdot X}$ where $\hat{\beta}_{Y \cdot X}$ is the sample estimate of slope.

Ignoring the covariable information it is possible to analyze the posttest data using a two-way fixed effects ANOVA. The results of such an analysis are reported in Table 4. The F test for a sex by treatment interaction is zero providing no evidence for rejecting the null hypothesis. The F test for a treatment effect is 2 which does not exceed the tabled critical value of 4.49 for $\alpha = .05$. While the F test for a sex main effect was significant, it was probably never in doubt. From the point of view of precision, the treatment effects

$$\hat{\alpha}_{Y_{T.}} = \overline{Y}_{T.} - \overline{Y}_{..} = \pm .5$$

(Table 3) were not judged to be different from zero using an error variance of $MS_{I:TS} = 2.5$.

The observations on X contained in Table 2 might have been used to form an index of response. Since gain scores have been so popular and because they have also been so heavily criticized (Cronbach and Furby, 1970) gain scores is a pedagogically good starting point. The results of an ANOVA of gain scores are reported in Table 5. Consistent with the conclusion from Table 1 that index of response tests the same hypothesis as ANOVA of the posttest, the mean square treatment is equal to five in both Tables 4 and 5. Similarly, using the means from Table 3 the treatment effects for gain scores are

$$\hat{\alpha}_{Z_{T.}} = \hat{\alpha}_{Y_{T.}} - \hat{\alpha}_{X_{T}}$$

$$= \pm .5 .$$

It must be remembered, however, that this identity of effects at the sample level is a function of exactly equal treatment means on X and should not be expected in practice.

Using gain scores the interaction F test remains zero but the treatment effect F test is 5 which is statistically significant at $\alpha = .05$. The significant F results from the error variance having been reduced from 2.5 to 1.0. The main point for the student is that these results could have been predicted from the relationships summarized in Table 1, i.e.,

$$\sigma_Z^2 = \sigma_Y^2 (1 - \rho_{XY}^2) + a^2 \sigma_Y^2$$

$$= 2.5 \ (1 - .8^2) + (1 - .8)^2 \ 2.5$$

$$= 1.0$$

Table 6 provides information on the relationship between $K/\beta_{Y \cdot X}$ and the mean square error. Consistent with the relationships given in Table 1, error variance is seen to be a minimum when $K/\beta_{Y \cdot X} = 1$. Further, the example illustrates that relative to improvement over ANOVA of the posttest the break-even points are $K/\beta_{Y \cdot X}$ equal 0 and 2.

As an aside, it is useful to point out to the student that the precision of gain scores for testing effects was quite good despite popular criticism about their low reliability. Explaining the paradox is beyond the scope of this paper but worth including in a course on experimental design (Porter, 1973). The beginning student might also be interested to know that on occasion a repeated measures design has been suggested as an alternative to the low reliability gain scores. While the algebra is tedious, the numerical example from Table 2 provides an easy illustration that the two procedures are identical. Using pretest as level one and posttest as level two of a repeated measures dimension crossed with T, S, and I; the ANOVA results are presented in Table 7. The mean squares in Table 7 are exactly half the comparable mean squares for gain scores given in Table 5 so that the factor cancels when forming the F tests.

Analysis of covariance is the final analysis strategy to be illustrated using the data from Table 2 and the results are presented in Table 8. Again, the mean square for treatments and treatments by sex interaction are as before, i.e., five and zero respectively. Similarly, ANCOVA adjusted treatment effects estimated using means from Table 3 are $\pm$ .5 as they were for each of the other design/analysis strategies. The mean square error for ANCOVA is seen to be .96 which could have been predicted from Table 1, i.e.,

$$\sigma^2_{Y.X} = \sigma^2_Y (1 - \rho^2_{XY}) \left[1 + \frac{1}{d.f.-2}\right]$$

$$= 2.5 (1 - .8^2) [1 + \frac{1}{14}]$$

$$\doteq .96$$

Again, these numerical results allow the student to verify predictions based on relationships summarized in Table 1. For example, the ANCOVA mean square error is not as small as the minimum mean square error for ANOVA of index of response. Further, not counting the loss of one degree of freedom, in the example ANCOVA is comparable to gain scores in precision. At this point the student may wish to refer to Cox (1957) where an index of precision is provided that reflects both mean square error and degrees of freedom.

In summary, the purpose of this paper has been to present a teaching aid that might be used in a beginning course on experimental design. The intent of the aid is to provide concrete illustrations of some of the interrelationships among three alternative design/analysis strategies for estimating treatment effects in random assignment designs. The first relationship to be illustrated is that all three strategies estimate the same effects. Thus, choice of strategy may rest on precision and other factors not considered here. The second relationship to be illustrated was precision and was considered by comparison of mean square errors.

Table 1
Summary of Analysis Strategies

Strategy	Effect	Error
ANOVA of Y	α_{Y_T}	σ_Y^2
ANOVA of Z	$\alpha_{Y_{T.}} - K\alpha_{X_{T.}}$	$\sigma_Y^2(1-\rho_{XY}^2) + a^2\sigma_X^2$
ANCOVA	$\alpha_{Y_{T.}} - \beta_{Y.X}\,\alpha_{X_{T.}}$	$\sigma_Y^2(1-\rho_{XY}^2)[1+\frac{1}{d.f.-2}]$

Table 2
Example Data

			X	Y	Y-X
T_1	S_1	I_1	8	9	1
		I_2	10	10	0
		I_3	6	8	2
		I_4	9	11	2
		I_5	7	7	0
	S_2	I_6	11	11	0
		I_7	14	14	0
		I_8	10	12	2
		I_9	13	15	2
		I_{10}	12	13	1
T_2	S_1	I_{11}	6	7	1
		I_{12}	8	8	0
		I_{13}	10	9	-1
		I_{14}	7	6	-1
		I_{15}	9	10	1
	S_2	I_{16}	12	12	0
		I_{17}	14	13	-1
		I_{18}	10	11	1
		I_{19}	13	14	1
		I_{20}	11	10	-1

T_1 and T_2: Treatments One and Two

S_1 and S_2: Two Sexes

I_1 - I_{20}: Individuals

X: Antecedent Information (covariable)

Y: Posttest or Outcome Variable

Y-X: Gain Score

Table 3

Means for The Data Presented in Table One

		S_1	S_2	
X	T_1	8	12	10
	T_2	8	12	10
		8	12	10

		S_1	S_2	
Y	T_1	9	13	11
	T_2	8	12	10
		8.5	12.5	10.5

		S_1	S_2	
Z	T_1	1	1	1
	T_2	0	0	0
		.5	.5	.5

Table 4

Analysis of Variance of the Posttest Data

Sources	d.f.	SS	MS	F*
T	1	5	5	2
S	1	80	80	32
TS	1	0	0	0
I:TS	16	40	2.5	

*F-Value Required For Statistical Significance at α = .05 with 1 and 16 degrees of freedom is 4.49.

Table 5

Analysis of Variance of the Gain Scores

Sources	d.f.	SS	MS	F*
T	1	5	5	5
S	1	0	0	0
TS	1	0	0	0
I:TS	16	16	1	

* Required F = 4.49 (d.f. 1, 16; α = .05)

Table 6

The Effect of Choice of K on the Precision of ANOVA of Index of Response

K	$K/\beta_{Y \cdot X}$	σ_Z^2
-.2	-.25	3.4
.0	.00	2.5
.6	.75	1.0
.8	1.00	.9
1.0	1.25	1.0
1.6	2.00	2.5
1.8	2.25	3.4

Table 7

Repeated Measures Analysis of Variance
Using Pretest and Posttest as Measures

Sources	d.f.	SS	MS	F**
T	1			
S	1			
I:TS	16			
M	1			
TM	1	2.5	2.5	5
SM	1	0	0	0
TSM	1	0	0	0
IM:TS	16	8	.5	

$F \geq 4.49$ required for significance at $\alpha = .05$, with 1 and 16 d.f.

Table 8

Analysis of Covariance of the Posttest Data Using X as the Covariate

Sources	d.f.	SS_X	SS_Y	SS_{XY}	SS_{Y}'	d.f.'	MS_{Y}'	F^*
T	1	0	5	0	5	1	5	5.21
S	1	80	80	80	1.07	1	1.07	1.11
TS	1	0	0	0	0	1	0	0
I:TS	16	40	40	32	14.4	15	.96	

*F must equal or exceed 4.54 to be statistically significant at $\alpha = .05$, for 1 and 15 d.f.

References

Cox, D. R. The use of a concomitant variable in selecting an experimental design. Biometrika, 1957, 44, 150-158.

Cronbach, L. J. & Furby, L. How we should measure "change" -- or should we? Psychological Bulletin, 1970, 74(1), Pp. 68-80.

Porter, A. C. Analysis strategies for some common evaluation paradigms. Paper presented at the meeting of the American Educational Research Association, 1973.

Porter, A. C. and Chibucos, T. R. Analysis issues in summative evaluation. Chapter which appeared in Evaluating Educational Programs and Products, Gary Borich (Ed.), Educational Technology Press, 1973.