The Analysis of Covariance: An Alternative Case

By Moo Ki Bai*

The analysis of covariance is a very useful technique when one analyzes data with different groups or classes, or with different time periods. It can be used in testing differences in intercepts, in slopes of any economic function, and overall homogeneity between classes.

Of the text book treatments of the analysis, Johnston[1] presents neat analysis of covariance table and test formulas. His 'complete' table and the corresponding test formulas, however, can hardly be exclusive ones. In his test of differences in intercepts, slopes, i.e., coefficients of the independent variables, are assumed to be constant for all classes. But in the test of differences in slopes between classes, intercepts are allowed to be changed. Although such tests can be found very useful, researchers in various empirical studies, for example, demand function or production function studies, may want to test under different conditions. It is not unthinkable for one to seek a test of differential intercepts between classes with different slopes or a test of differential slopes with constant intercepts for all casses. The aim of this brief note is to make an alternative analysis of covariance table and the corresponding test formulas along this line.

All the relevant notations and equation formulations in the book will be used here. In order to make the conditions of intercepts and slopes the other way around, one interim model must be reformulated. It is an intermediate model which is an inbetween of the most and the least restricted models.

It can be postulated as

(1) $y = V\beta + u$[2]

Here y is an (nx1) column vector made up of p sub-vectors.

V is an $(n \times (pk-p+1))$ matrix and β a $((pk-p+1)\times 1)$ vector of coefficients. Specifically V is

---

* The writer is a full-time instructor in economics, College of Commerce, Seoul National University.

(1) J. Johnston, *Econometric Methods*, 2nd ed. New York, McGraw-Hill Book Co., 1972.

(2) Italic type was used throughout this note instead of Gothic type.

$$V=\begin{pmatrix} \left.\begin{matrix}1\\1\end{matrix}\right\} & X_1 & 0 & 0\cdots\cdots & 0 \\ \vdots & 0 & X_2 & 0\cdots\cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1\} & 0 & 0 & 0\cdots\cdots & X_p \end{pmatrix}$$

where

$$X_i=\begin{pmatrix} X_{i21} & \cdots\cdots\cdots\cdots & X_ik_1 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ X_{i2m} & \cdots\cdots\cdots\cdots & X_ik_m \end{pmatrix} \qquad i=1,\cdots\cdots,p$$

The first column of V matrix is units for an intercept.

If we indicate the result of applying least-squares to (1) as

(2) $\qquad y=V\hat{\hat{\beta}}+e^*$

where $\hat{\hat{\beta}}=(V'V)^{-1}V'y$ as usual and e^* indicates the vector of least-squares residuals, we then have

(3) $\qquad y'y=\hat{\hat{\beta}}'V'V\hat{\hat{\beta}}+e^{*\prime}e^*+2\hat{\hat{\beta}}'V'e^*$

or more simply

(4) $\qquad y'y=\hat{\hat{\beta}}'V'y+e^{*\prime}e^*.$

Therefore the sum of squares of residuals is

(5) $\qquad e^{*\prime}e^*=y'y-\hat{\hat{\beta}}'V'y$

with $(mp-pk+p-1)$ degree of freedom. The reduction in the residual sum of squares in moving from the most restricted model (6)

(6) $\qquad y=X\beta+u$

to (1) is thus

(7) $\qquad s's-e^{*\prime}e^*=\hat{\hat{\beta}}'V'y-\hat{\beta}'V'y=S^*_1$

with $(pk-k-p+1)$ degree of freedom. This incremental (or the reduction of the residual sum of squares) is due to the differential slope vectors with constant intercept.

By the same token, the reduction in the residual sum of spuares in moving from (1) to the model which allows both intercepts and slope vectors to change is

(8) $\qquad e^{*\prime}e^*-r'r=b'Z'y-\hat{\hat{\beta}}V'y=S^*_3$

with $(p-1)$ degree of freedom.

Table 1 and 2 show Johnston's analysis of covariance table and the corresponding F-test formulas, whereas Table 3 and 4 summarize the results of this note. $F_{(1)}$* in Table 4 is the test formula of differential intercepts by allowing slope vectors to change and $F_{(2)}$* is that of differential slope vectors with constant intercept. Naturally enough, $F^*_{(3)}$ in Table 4 and $F_{(3)}$ in Table 2 are identical except S_1+S_3 in Table 2 and $S^*_1+S^*_3$ in Table 4. However, S_1+S_3 must be equal to $S^*_1+S^*_3$ because both $F^*_{(3)}$ and $F_{(3)}$ are testing the exactly identical thing.

Table 1 **Johnston's Complete Analysis of Covariance Table**

Source	Sum of Squares	df	Mean Square
Z	Residual: $r'r=y'y-b'Z'y=S_4$	$p(m-k)$	$S_4/p(m-k)$
	Incremental (Differential slope vectors) $e'e-r'r=b'Z'y-\hat{\hat{\alpha}}'D'y-\hat{\hat{\beta}}'X'y=S_3$	$pk-p-k+1$	$S_3/(pk-p-k+1)$
X and D	Residual: $e'e=y'y-\hat{\hat{\alpha}}'D'y-\hat{\hat{\beta}}'X'y=S_2=S_3+S_4$	$mp-p-k+1$	$S_2/(mp-p-k+1)$
	Incremental (Differential intercepts) $s's-e'e=\hat{\hat{\alpha}}'D'y+\hat{\hat{\beta}}'X'y-\hat{\beta}'X'y=S_1$	$p-1$	$S_1/(p-1)$
X	Residual: $s's=y'y-\hat{\beta}'X'y$	$mp-k$	

Table 2 **Johnston's Test Formulas**

Test of	Formula
Differential intercepts	$F_{(1)}=\dfrac{S_1/(p-1)}{S_2/(mp-p-k+1)}$
Differential slope vectors	$F_{(2)}=\dfrac{S_3/(pk-p-k+1)}{S_4/p(m-k)}$
Overall homogeneity	$F_{(3)}=\dfrac{(S_1+S_3)/k(p-1)}{S_4/p(m-k)}$

Table 3　　**The Alternative Analysis of Covariance Table**

Source	Sum of Squares	df	Mean Square
Z	Residual:		
	$r'r = y'y - b'Z'y = S_4$	$p(m-k)$	$S_4/p(m-k)$
	Incremental (Differential intercepts)		
	$e^{*\prime}e^* - r'r = b'Z'y - \hat{\hat{\beta}}'V'y = S^*_3$	$p-1$	$S^*_3/(p-1)$
V	Residual:		
	$e^{*\prime}e^* = y'y - \hat{\hat{\beta}}'V'y = S^*_2 = S^*_3 + S_4$	$mp-pk+p-1$	$S^*_2/(mp-pk+p-1)$
	Incremental (Differential slope vectors)		
	$s's - e^{*\prime}e^* = \hat{\hat{\beta}}'V'y - \hat{\beta}'X'y = S^*_1$	$pk-k-p+1$	$S^*_1/(pk-k-p+1)$
X	Residual:		
	$s's = y'y - \hat{\beta}'X'y$	$mp-k$	

Table 4　　**The Alternative Test Formulas**

Test of	Formula
Differential intercepts	$F^*_{(1)} = \dfrac{S^*_3/(p-1)}{S_4/p(m-k)}$
Differential slope vectors	$F^*_{(2)} = \dfrac{S^*_1/(pk-p-k+1)}{S^*_2/(mp-pk+p-1)}$
Overall homogeneity	$F^*_{(3)} = \dfrac{(S^*_1+S^*_3)/k(p-1)}{S_4/p(m-k)}$