ANALYSIS OF COVARIANCE (ANOCOVA or ANCOVA)(*Page 1/3*)

Analysis of covariance is a combination of the Analysis of Variance (ANOVA) and regression analysis. Sometimes observations are taken on one or more independent variables besides the study variable from each experimental unit. The independent variable(s) is(are) called ancillary or concomitant variables. In such cases, ANOCOVA is applied to examine whether the variation in the study (dependent) variable (*y*) over the classes is due to the class effects or due to its dependence on the concomitant variable(s) (*x)*. The ANOCOVA controls the experimental error by taking into consideration the dependence of y on the x's.

Some examples where the technique of ANOCOVA can be used are as follows:

(i) The yield(*x*) of a crop may depend on the number(*y*) of plants per plot and we may consider the number(*y*) of plants as the concomitant variable and perform ANOCOVA.

(ii) In a study of the effect of drugs or diets on the growth of animals, the growth(y) may depend on the initial weight(x) of the animals and ANOCOVA may be performed.

ANALYSIS OF COVARIANCE OF ONE-WAY CLASSIFIED DATA:

Let us suppose that we have a set of observations (x_{ij}, y_{ij}), $i = 1, 2, \ldots\ldots, t; j = 1, 2, \ldots\ldots, r_i$ on the concomitant variable X and the dependent variable Y classified according to t classes.

The linear model is

$$y_{ij} = \mu_i + \beta(x_{ij} - x_{00}) + e_{ij} \qquad \ldots(1)$$

$i = 1, 2, \ldots\ldots, t; j = 1, 2, \ldots\ldots, r_i$ where e_{ij} are independently normal with zero means and variance σ^2 and $x_{oo} = \frac{\sum x_{ij}}{n}$ where $n = \sum_i^t r_i$ and β is the regression coefficient of *y* on *x.*

The least -square estimates of the parameters in model (1) are

$$\hat{\mu}_i = y_{i0} - \hat{\beta}(x_{i0} - x_{00})$$

$$\hat{\beta} = \frac{\sum_{i,j}(x_{ij} - x_{i0})(y_{ij} - y_{i0})}{\sum_{i,j}(x_{ij} - x_{io})^2} = \frac{E_{xy}}{E_{xx}} \text{ , say}$$

The unrestricted residual (error) *SS* for the above model is

$$SSE = \sum\sum_{ij}\left[y_{ij} - \hat{\mu}_i - \widehat{\beta}(x_{ij} - x_{00})\right]^2$$

$$= \sum\sum_{ij}\left[y_{ij} - y_{i0} - \widehat{\beta}(x_{ij} - x_{00})\right]^2 \quad (y_{i0} \text{ is the mean of } i^{th} \text{ class of Y})$$

$$= E_{yy} - \frac{E_{xy}^2}{E_{xx}} \qquad \ldots\ldots(2)$$

Where $E_{xx} = \sum_{ij}(x_{ij} - x_{00})^2$

$$E_{yy} = \sum_{ij}(y_{ij} - y_{00})^2$$

$E_{xy} = \sum_{ij}(x_{ij} - x_{i0})(y_{ij} - y_{i0})$ (*PTO*)

D.N. Adhikary.

Analysis of covariance
(Continued)

The unrestricted SSE has $df = n-t-1$

The null hypothesis to be tested is

H_0: μ_i are all equal i.e. the effects due to different classes are same when the regression of y on x is taken into consideration

The restricted error SS when H_0 is true is given by

$$(SSE)' = \text{Min} \sum_i \sum_j \{y_{ij} - \mu - \beta(x_{ij} - x_{00})\}^2 \text{ w.r.t. } \mu \text{ and } \beta$$

$$= \sum_{ij} \{y_{ij} - y_{00} - \hat{\beta}^*(x_{ij} - x_{00})\}^2 \text{ . This SS has } df = n-2$$

where $\hat{\mu} = y_{00} = \frac{1}{n}\sum_{ij} y_{ij}$ and $\hat{\beta}^* = \frac{\sum\sum(x_{ij} - x_{00})(y_{ij} - y_{00})}{\sum\sum(x_{ij} - x_{00})^2} = \frac{E'_{XY}}{E'_{XX}}$ (say)

are the estimates of μ and β under H_0.

We can show that

$$(SSE)' = E'_{YY} - \frac{E'^2_{XY}}{E'_{XX}}$$

where

$E'_{YY} = T_{YY} + E_{YY}$, $T_{YY} = \sum_i r_i(y_{i0} - y_{00})^2 = SSC$ for Y

$E'_{XX} = E_{XX} + T_{XX}$, $T_{XX} = \sum_i r_i(x_{i\cdot} - x_{00})^2 = SSC$ for X

and $E'_{XY} = T_{XY} + E_{XY}$, $T_{XY} = \sum_i (x_{i0} - x_{00})(y_{i0} - y_{00}) = SPC$ for X, Y

Appropriate test statistics for testing H_0 will be

$$F = \frac{\{(SSE)' - SSE\}/(t-1)}{SSE/(n-t-1)}$$

which follows the F distribution with $\nu_1 = t-1$ and $\nu_2 = n-t-1$ d.f. when H_0 is true.

The analysis of covariance table for one-way classified data with one concomitant variable will be as given below

(T.O)

ANALYSIS OF COVARIANCE (Continued)

ANOCOVA Table (one-way classified data)

Source of variations	df (degrees of freedom)	SS_{XX}	SP_{XY}	SS_{YY}	Estimate of β	Adjusted	
						SS_{YY}	df
Classes (Treatments)	$t-1$	T_{XX}	T_{XY}	T_{YY}			
Error	$n-t$	E_{XX}	E_{XY}	E_{YY}	$E_{XY}/E_{XX} = \hat{\beta}$	SSE	$n-t-1$
Total	$n-1$	E'_{XX}	E'_{XY}	E'_{YY}	$E'_{XY}/E'_{XX} = \hat{\beta}^*$	$(SSE)'$	$n-2$
Difference Total − Error		—				$(SSE)' - SSE$	$t-1$

— x —

Questions :

1. What is analysis of covariance? Discuss the analysis of a one-way classified data with a concomitant variable. (2015)
2. Write the linear model of one-way classified data when there is a concomitant variable. (2018)
3. Write an explanatory note on analysis of Covariance. (2018)