One-way ANCOVA model with one fixed concomitant variable

The relevant model is

$$y_{ij} = \mu + \alpha_i + \gamma(x_{ij} - \bar{x}_{i0}) + e_{ij}, \quad i = 1(1)k, \; j = 1(1)n_i, \quad \sum_{i=1}^{k} n_i = n. \quad \text{----} (1)$$

Where y_{ij} denotes the j^{th} observation of interest variable corresponding to i^{th} class or level of the factor,

μ denotes the general effect,

α_i denotes the additional fixed effect due to i^{th} class, I

x_{ij} ~~e_{ij}~~ denotes the ~~random errors~~ j^{th} observation of the concomitant variable X corresponding to the i^{th} class,

γ denotes the regression coefficient associated to the variable X and

e_{ij} denotes the random errors independent to all other model components in (1).

The basic interest is to test $H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

against $H_1: \alpha_i \neq 0$ for atleast one i, $i = 1(1)k$.

Let us assume that $e_{ij} \overset{iid}{\sim} N(0, \sigma_e^2) \; \forall (i,j)$.

From (1) it is clear that $\sum_{i=1}^{k} n_i \alpha_i = 0$. Let us denote $x_{ij}' = (x_{ij} - \bar{x}_{i0})$.

Now, the least-square estimates of μ, α_i and γ are obtained by minimizing $E = \sum_i \sum_j (y_{ij} - \mu - \alpha_i - \gamma x_{ij}')^2$ w.r.t. μ, α_i and γ.

$$\frac{\partial E}{\partial \mu} = 0 \Rightarrow \sum_i \sum_j y_{ij} = n\mu + \sum_i n_i \alpha_i \Rightarrow \hat{\mu} = \bar{y}_{00}$$

$$\frac{\partial E}{\partial \alpha_i} = 0 \Rightarrow \sum_j y_{ij} = n_i \mu + n_i \alpha_i \Rightarrow \hat{\alpha}_i = \bar{y}_{i0} - \bar{y}_{00} \; \forall i.$$

$$\frac{\partial E}{\partial \gamma} = 0 \Rightarrow \sum_i \sum_j y_{ij} \cdot x_{ij}' = \mu \sum_i \sum_j x_{ij}' + \sum_i \alpha_i \sum_j x_{ij}' + \gamma \sum_{i,j} x_{ij}'^2$$

$$\Rightarrow \hat{\gamma} = \frac{\cancel{\sum_i \sum_j (y_{ij} - \bar{y}_{00})(x_{ij}} \sum_i \sum_j y_{ij} \cdot x_{ij}'}{\sum_i \sum_j x_{ij}'^2} \quad \text{....} (2)$$

$\hat{\gamma}$ can also be written as $\dfrac{\sum_i \sum_j (y_{ij} - \bar{y}_{i0})(x_{ij} - \bar{x}_{i0})}{\sum_i \sum_j (x_{ij} - \bar{x}_{i0})^2}$. $\left(\because \sum_j x_{ij}' = 0 \; \forall i\right)$

Now, the unrestricted residual SS is

$$S_1^{2\prime} = \underset{\substack{\mu, \alpha_i, \gamma \\ \forall i}}{\text{Min}} \sum_i \sum_j (y_{ij} - \mu - \alpha_i - \gamma x_{ij}')^2$$

$$= \sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\gamma} x_{ij}')^2$$

$$= \sum_i \sum_j \left[y_{ij} - \bar{y}_{00} - (\bar{y}_{i0} - \bar{y}_{00}) - \hat{\gamma}(x_{ij}' - \bar{x}_{i0}')\right]^2$$

$$= \cancel{\sum_i \sum_j (y_{ij} - \bar{y}_{00})^2 + \sum_i \sum_j (\bar{y}_{i0} - \bar{y}_{00})^2} + \hat{\gamma}^2 \sum_i \sum_j (x_{ij}' - \bar{x}_{i0}')^2$$
$$\cancel{- 2\sum_i (\bar{y}_{i0} - \bar{y}_{00}) \sum_j (y_{ij} - \bar{y}_{00})}$$
$$\cancel{+ 2\hat{\gamma} \sum_i (\bar{y}_{i0} - \bar{y}_{00}) \sum_j (x_{ij}' - \bar{x}_{i0}')}$$
$$- 2 \sum_i \sum_j \cancel{(y_{ij} - \bar{y}_{00})(x_{ij}' - \bar{x}_{i0}')}.$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_{i0})^2 + \hat{\gamma}^2 \sum_i \sum_j (x_{ij}' - \bar{x}_{i0}')^2$$
$$- 2\hat{\gamma} \sum_i \sum_j (y_{ij} - \bar{y}_{i0})(x_{ij}' - \bar{x}_{i0}').$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_{i0})^2 - \hat{\gamma} \sum_i \sum_j (y_{ij} - \bar{y}_{i0})(x_{ij}' - \bar{x}_{i0}').$$

[Using (2)]

d.f. of S_1^2 is. $(\overline{n-k} - 1) = n-k-1$.

[$\because$ The first part of S_1^2 consists n observations on Y with k restrictions of group means $\bar{y}_{i0}$, $i=1(1)k$. So, it has d.f $(n-k)$ and in second part there is one restriction due to estimate $\hat{\gamma}$]

Next, the restricted (i.e. under H_0) residual SS is,

$$S_2^2 = \underset{\substack{\mu, \alpha_i, \gamma \\ H_0}}{\text{Min}} \sum_i \sum_j (y_{ij} - \mu - \alpha_i - \gamma x_{ij}')^2$$

$$= \underset{\mu, \gamma}{\text{Min}} \sum_i \sum_j (y_{ij}' - \mu - \gamma x_{ij}')^2.$$

$$= \sum_i \sum_j (y_{ij} - \hat{\hat{\mu}} - \hat{\hat{\gamma}} x_{ij}')^2.$$

where $\hat{\hat{\mu}} = \bar{y}_{00}$, $\hat{\hat{\gamma}} = \dfrac{\sum_i \sum_j (y_{ij} - \bar{y}_{00}) x_{ij}}{\sum_i \sum_j x_{ij}^2}$ ....(3).

So, $S_2^2 = \sum_i \sum_j (y_{ij} - \bar{y}_{00})^2 + \hat{\gamma}^2 \sum_i \sum_j (x_{ij} - \bar{x}_{i0})^2$
$\quad - 2\hat{\gamma} \sum_i \sum_j (y_{ij} - \bar{y}_{00})(x_{ij} - \bar{x}_{i0})$.
$= \sum_i \sum_j (y_{ij} - \bar{y}_{00})^2 - \hat{\gamma} \sum_i \sum_j (y_{ij} - \bar{y}_{00})(x_{ij} - \bar{x}_{i0})$, using (3)

$\left[\because \hat{\gamma}\right.$ can be written as $\left.\dfrac{\sum_i \sum_j (y_{ij} - \bar{y}_{00})(x_{ij} - x_{i0})}{\sum_i \sum_j (x_{ij} - x_{i0})^2}\right]$

d.f. of S_2^2 is $(\overline{n-1} - 1) = n-2$

So, the adjusted SS due to between groups variation is $SSB = (S_2^2 - S_1^2)$ with d.f. $= \overline{n-2} - \overline{n-k-1} = k-1$. and

SSE, adjusted ss due to error, is S_1^2 with d.f. $= n-k-1$.

Under H_0, SSB and SSE both follow χ^2 distribution with respective d.f.s. and they are independent.

Hence, $F = \dfrac{SSB/k-1}{SSE/n-k-1} = \dfrac{MSB}{MSE} \sim F_{k-1, n-k-1}$.

If observed F, say F_0, is found to be greater than $F_{\alpha; k-1, n-k-1}$, we reject H_0 at level α, otherwise we do not reject H_0.

ANCOVA Table

Source of variation	ANCOVA: Adjusted model incorporating by one concomitant variable			
	d.f.	SS	MS	F_0
Between Classes	$k-1$	SSB	$MSB = \dfrac{SSB}{k-1}$	$F_0 = \dfrac{MSB}{MSE}$
Due to concomitant variable	-1			
Error	$n-k-1$	SSE	$MSE = \dfrac{SSE}{n-k-1}$	
Total.	$n-1$	$TSS = \sum_i \sum_j (y_{ij} - \bar{y}_{00})^2$		

P.T.O.

Associated ANCOVA Table.

Source of Variation	d.f.	Adjusted by one concommitant variable			
		d.f.	SS	MS	F
Between classes	$k-1$	$k-1$	$SSB = S_2^2 - S_1^2$	$MSB = \frac{SSB}{k-1}$	$F_0 = \frac{MSB}{MSE}$
Error	$n-k$	$n-k-1$	$S_1^2 = SSE$	$MSE = \frac{SSE}{n-k-1}$	
Total	$n-1$	$n-2$	S_2^2		
Concomitant Information		1			

Two-way ANCOVA model with one fixed concomitant variable and one observation per cell

OR, Analysis of Covariance for an RBD with one ~~observation~~ concomitant variable.

The relevant model is

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma(x_{ij} - \bar{x}_{io}) + e_{ij},$$
$$i = 1(1)p,\ j = 1(1)q. \quad \text{------} (1)$$

Where y_{ij} denotes the observation of the response variable corresponding to the i^{th} row and j^{th} column,

μ denotes the general fixed effect,

α_i denotes the additional fixed effect corresponding to the i^{th} level of factor A ~~row~~, with $\sum_i \alpha_i = 0$,

β_j denotes the additional fixed effect corresponding to the j^{th} level of factor B ~~column~~, with $\sum_j \beta_j = 0$,

x_{ij} denotes the (i,j)th observation of the concomitant variable under study,

γ denotes the regression coefficient associated to the concomitant variable X and

e_{ij} denotes the random error independent to all other model components in (1).

The basic interest is to test

(i) $H_{0A}: \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$
against H_{1A} : not satisfying H_{01} and

(ii) $H_{0B}: \beta_1 = \beta_2 = \cdots = \beta_q = 0$
against H_{02A} : not satisfying H_{02}.

To carry out the tests, let us assume that $e_{ij} \overset{iid}{\sim} N(0, \sigma_e^2)\ \forall\, i,j$.

Now, the least square estimates of μ, α_i, β_j and γ are obtained by minimizing, $E = \sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma x'_{ij})^2$ w.r.t. μ, α_i, β_j and γ for all i and j, where $x'_{ij} = x_{ij} - \bar{x}_{io}\ \forall\, i,j$.

$$\frac{\partial E}{\partial \mu} = 0 \Rightarrow \sum_i \sum_j y_{ij} = n\mu + q\sum_i \alpha_i + p\sum_j \beta_j + \gamma \sum_i \sum_j x'_{ij}$$
$$\Rightarrow \hat{\mu} = \bar{y}_{oo} \quad [\because \textstyle\sum \alpha_i = \sum_j \beta_j = 0]$$

$$\frac{\partial E}{\partial \alpha_i} = 0 \Rightarrow \sum_j y_{ij} = q\mu + q\alpha_i + \sum_j \beta_j + \gamma \sum_j x'_{ij}$$

$$\Rightarrow \hat{\alpha}_i = \bar{y}_{i0} - \bar{y}_{00}$$

$$\frac{\partial E}{\partial \beta_j} = 0 \Rightarrow \hat{\beta}_j = \bar{y}_{0j} - \bar{y}_{00} \quad \text{and}$$

$$\frac{\partial E}{\partial \gamma} = 0 \Rightarrow \sum_i\sum_j y_{ij}\, x_{ij}' = \mu \sum_i\sum_j x_{ij}' + \sum_i \alpha_i \sum_j x_{ij}' + \sum_i\sum_j \beta_j x_{ij}' + \gamma \sum_i\sum_j x_{ij}'^2$$

$$\Rightarrow \sum_i\sum_j (y_{ij} - \hat{\beta}_j)\, x_{ij}' = \gamma \sum_i\sum_j x_{ij}'^2$$

$$\Rightarrow \hat{\gamma} = \frac{\sum_i\sum_j (y_{ij} - \bar{y}_{0j})\, x_{ij}'}{\sum_i\sum_j x_{ij}'^2} = \frac{\sum_i\sum_j (y_{ij} - \bar{y}_{0j})(x_{ij} - \bar{x}_{i0})}{\sum_i\sum_j (x_{ij} - \bar{x}_{i0})^2}$$

$$\left[\because \sum_j x_{ij}' = 0\right] \quad \cdots\cdots (2)$$

Now, the unrestricted residual SS is

$$S_1^2 = \underset{\substack{\mu, \alpha_i, \beta_j, \gamma \\ \forall i, j}}{\text{Min}} \sum_i\sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma x_{ij}')^2$$

$$= \sum_i\sum_j (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma} x_{ij}')^2$$

$$= \sum_i\sum_j \left(y_{ij} - \bar{y}_{00} - (\bar{y}_{i0} - \bar{y}_{00}) - (\bar{y}_{0j} - \bar{y}_{00}) - \hat{\gamma} x_{ij}'\right)^2$$

$$= \sum_i\sum_j (y_{ij} - \bar{y}_{0j})^2 + \sum_i\sum_j (\bar{y}_{i0} - \bar{y}_{00})^2 + \hat{\gamma}^2 \sum_i\sum_j (x_{ij} - \bar{x}_{i0})^2$$

$$- 2\sum_i\sum_j (y_{ij} - \bar{y}_{0j})(\bar{y}_{i0} - \bar{y}_{00}) - 2\sum_i\sum_j (y_{ij} - \bar{y}_{0j})(x_{ij} - \bar{x}_{i0})\hat{\gamma}$$

$$+ 2\sum_i\sum_j (\bar{y}_{i0} - \bar{y}_{00})(x_{ij} - \bar{x}_{i0}).$$

$$= \sum_i\sum_j (y_{ij} - \bar{y}_{0j})^2 + q\sum_i (\bar{y}_{i0} - \bar{y}_{00})^2 - \hat{\gamma} \sum_i\sum_j (y_{ij} - \bar{y}_{0j})(x_{ij} - \bar{x}_{i0}).$$

$$- 2q\sum_i (\bar{y}_{i0} - \bar{y}_{00})^2 \quad \left[\because \sum_j (x_{ij} - \bar{x}_{i0}) = 0 \ \forall i. \text{ and using } (2)\right]$$

$$= \sum_i\sum_j (y_{ij} - \bar{y}_{0j})^2 - q\sum_i (\bar{y}_{i0} - \bar{y}_{00})^2 - \hat{\gamma} \sum_i\sum_j (y_{ij} - \bar{y}_{0j})(x_{ij} - \bar{x}_{i0}).$$

So, d.f. of S_1^2 is $(pq - q) - (p-1) - 1 = pq - p - q$

[$\because$ The first part consists pq observations on y with q restrictions of group means $\bar{y}_{0j}$, $j = 1(1)q$. So, the first part has d.f. $(pq - q)$. With similar logic, the second part has d.f. $(p-1)$ and in the third part, there is one new restriction only due to the estimate $\hat{\gamma}$.]

Next, the restricted (i.e. under H_{0A}) residual SS is

$$S_{2A}^2 = \underset{\substack{\mu, \alpha_i, \beta_j, \gamma \\ \forall i,j \\ H_{0A}}}{\mathrm{Min}} \sum_i \sum_j \left(y_{ij} - \mu - \alpha_i - \beta_j - \gamma x_{ij}' \right)^2$$

$$= \underset{\substack{\mu, \beta_j, \gamma \\ \forall j}}{\mathrm{Min}} \sum_i \sum_j \left(y_{ij} - \mu - \beta_j - \gamma x_{ij}' \right)^2.$$

$$= \sum_i \sum_j \left(y_{ij} - \bar{y}_{00} - (\bar{y}_{0j} - \bar{y}_{00}) - \hat{\gamma} x_{ij}' \right)^2.$$

$$= \sum_i \sum_j \left(y_{ij} - \bar{y}_{0j} \right)^2 - \hat{\hat{\gamma}} \sum_i \sum_j \left(y_{ij} - \bar{y}_{0j} \right) \left(x_{ij} - \bar{x}_{i0} \right)$$

$$\text{where } \hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{0j})(x_{ij} - \bar{x}_{i0})}{\sum_i \sum_j (x_{ij} - \bar{x}_{i0})^2}$$

[Note that computation of the above S_{2A}^2 is same as that of S_1^2 in case of one-way ANCOVA model with one concomitant variable just replacing α_i by β_j in model]

d.f. of S_{2A}^2 is $(pq - q) - 1 = q(p-1) - 1$

Similarly, the restricted (due to H_{0B}) residual SS is

$$S_{2B}^2 = \underset{\substack{\mu, \alpha_i, \beta_j, \gamma \\ \forall i,j \\ H_{0B}}}{\mathrm{Min}} \sum_i \sum_j \left(y_{ij} - \mu - \alpha_i - \beta_j - \gamma x_{ij}' \right)^2$$

$$= \underset{\substack{\mu, \alpha_i, \gamma \\ \forall i}}{\mathrm{Min}} \sum_i \sum_j \left(y_{ij} - \mu - \alpha_i - \gamma x_{ij}' \right)^2$$

$$= \sum_i \sum_j \left(y_{ij} - \bar{y}_{i0} \right)^2 - \hat{\hat{\gamma}} \sum_i \sum_j \left(y_{ij} - \bar{y}_{i0} \right) \left(x_{ij} - \bar{x}_{i0} \right).$$

$$\text{where } \hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{i0})(x_{ij} - \bar{x}_{i0})}{\sum_i \sum_j (x_{ij} - \bar{x}_{i0})^2}$$

$\therefore$ d.f. of S_{2B}^2 is $p(q-1) - 1$.

So, the adjusted SS due to between levels of factor 'A' is, $SSA = (S_{2A}^2 - S_1^2)$ with

$$d.f. = pq-q-1-(pq-p-q) = p-1.$$

Adjusted SS due to error is, $SSE = S_1^2$ with

$$d.f = pq-p-q.$$

Under H_{0A}, SSA and SSE both follow χ^2 distn with respectives d.f.s. and they are independent.

Hence, $$F_A = \frac{SSA/(p-1)}{SSE/(pq-p-q)} = \frac{MSA}{MSE} \sim F_{\overline{p-1},\, \overline{pq-p-q}},$$

If observed F_A, say F_{A0}, is found to be greater than $F_{\alpha;\overline{p-1},\, \overline{pq-p-q}}$, we reject H_{0A} at level α, otherwise we do not reject H_{0A}.

Similarly, to carryout test for H_{0B}, we define.

$$F_B = \frac{SSB/(q-1)}{SSE/(pq-p-q)} = \frac{MSB}{MSE} \sim F_{\overline{q-1},\, \overline{pq-p-q}}$$

and set the rejection criteria for H_{0B} similarly as in case of H_{0A}, where $SSB = (S_{2B}^2 - S_1^2)$ with

$$d.f. = pq-p-1-(pq-p-q) = q-1.$$

i.e. if observed F_B, say F_{B0}, is greater than $F_{\alpha;\overline{q-1},\, \overline{pq-p-q}}$, we reject H_{0B} at level α, otherwise we do not reject H_{0B}.

Note: In the light of the analysis of covariance of an RBD model we do test the equality of treatment (here, factor 'A') effects only, not the equality of block (here, factor 'B') effects. So, relevent calculations for MSB and F_B are not be done accordingly. Now, we present the ANCOVA Table relevent for RBD model.

Associated ANCOVA Table for RBD.

Sources of Variation	d.f.	Adjusted by one concomitant variable			
		df.	SS	MS	F
Blocks (Factor B)	$q-1$	$q-1$			
Treatments (Factor A)	$p-1$	$p-1$	$SSA = S_{2A}^2 - S_1^2$	$MSA = \frac{SSA}{p-1}$	$F_{A} = \frac{MSA}{MSE}$
Error	$(p-1)(q-1)$	$pq-p-q$	$S_1^2 = SSE$	$MSE = \frac{SSE}{pq-p-q}$	
Total (Treatments + Error)	$pq-q$	$pq-q-1$	S_{2A}^2		
Total	$pq-1$	$pq-2$			

<u>Remark:</u> When we analyze the ANCOVA model with one fixed concomitant variable and one observation per cell, the associated ANCOVA Table will be same as the above table with addition to the following entries corresponding to the source of variation — 'Blocks (Factor B)':

SS: $SSB = S_{2B}^2 - S_1^2$,

MS: $MSB = \frac{SSB}{q-1}$,

F: $F_B = \frac{MSB}{MSE}$,

since unlike RBD model, in this case we have to carry out the testing of $H_{0B}: \beta_1 = \beta_2 = \cdots = \beta_q = 0$ also in addition the testing of $H_{0A}: \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$.