Chapter 12
Analysis of Covariance

Any scientific experiment is performed to know something that is unknown about a group of treatments and to test certain hypothesis about the corresponding treatment effect.

When variability of experimental units is small relative to the treatment differences and the experimenter do not wishes to use experimental design, then just take large number of observations on each treatment effect and compute its mean. The variation around mean can be made as small as desired by taking more observations.

When there is considerable variation among observations on the same treatment and it is not possible to take an unlimited number of observations, the techniques used for reducing the variation are

(i) use of proper experimental design and
(ii) use of concomitant variables.

The use of concomitant variables is accomplished through the technique of analysis of covariance. If both the techniques fail to control the experimental variability then the number of replications of different treatments (in other words, the number of experimental units) are needed to be increased to a point where adequate control of variability is attained.

Introduction to analysis of covariance model

In the linear model

$$Y = X_1\beta_1 + X_2\beta_2 + ... + X_p\beta_p + \varepsilon,$$

if the explanatory variables are quantitative variables as well as indicator variables, i.e., some of them are qualitative and some are quantitative, then the linear model is termed as analysis of covariance (ANCOVA) model.

Note that the indicator variables do not provide as much information as the quantitative variables. For example, the quantitative observations on age can be converted into indicator variable. Let an indictor variable be

$$D = \begin{cases} 1 \text{ if age} \geq 17 \text{ years} \\ 0 \text{ if age} < 17 \text{ years.} \end{cases}$$

Now the following quantitative values of age can be changed into indicator variables.

Ages (years)	Ages
14	0
15	0
16	0
17	1
20	1
21	1
22	1

In many real application, some variables may be quantitative and others may be qualitative. In such cases, ANCOVA provides a way out.

It helps is reducing the sum of squares due to error which in turn reflects the better model adequacy diagnostics.

See how does this work:

$$\begin{aligned}
&\text{In one way model}: Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, && \text{we have } TSS_1 = SSA_1 + SSE_1 \\
&\text{In two way model}: Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, && \text{we have } TSS_2 = SSA_2 + SSB_2 + SSE_2 \\
&\text{In three way model}: Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ik}, && \text{we have } TSS_3 = SSA_3 + SSB_3 + SS\gamma_3 + SSE_3
\end{aligned}$$

If we have a given data set, then ideally

$$\begin{aligned}
&TSS_1 = TSS_2 = TSS_3 \\
&SSA_1 = SSA_2 = SSA_3; \\
&SSB_2 = SSB_3
\end{aligned}$$

So $\text{SSE}_1 \geq SSE_2 \geq SSE_3$.

Note that in the construction of F-statistics, we use $\dfrac{SS(effects)/df}{SSE/df}$.

So F-statistic essentially depends on the $SSEs$.

Smaller $SSE \Rightarrow$ large $F \Rightarrow$ more chance of rejection.

Since SSA, SSB etc. here are based on dummy variables, so obviously if SSA, SSB, etc. are based on quantitative variables, they will provide more information. Such ideas are used in ANCOVA models and we construct the model by incorporating the quantitative explanatory variables in ANOVA models.

In another example, suppose our interest is to compare several different kinds of feed for their ability to put weight on animals. If we use ANOVA, then we use the final weights at the end of experiment. However, final weights of the animals depend upon the initial weight of the animals at the beginning of the experiment as well as upon the difference in feeds.

Use of ANCOVA models enables us to adjust or correct these initial differences.

ANCOVA is useful for improving the precision of an experiment. Suppose response Y is linearly related to covariate X (or **concomitant variable**). Suppose experimenter cannot control X but can observe it. ANCOVA involves adjusting Y for the effect of X. If such an adjustment is not made, then the X can inflate the error mean square and makes the true differences is Y due to treatment harder to detect.

If, for a given experimental material, the use of proper experimental design cannot control the experimental variation, the use of concomitant variables (which are related to experimental material) may be effective in reducing the variability.

Consider the one way classification model as

$$E(Y_{ij} = \beta_i \qquad i = 1,..., p, j = 1,..., N_i,$$
$$Var(Y_{ij}) = \sigma^2.$$

If usual analysis of variance for testing the hypothesis of **equality of treatment effects** shows a highly significant difference in the treatment effects due to some factors affecting the experiment, then consider the model which takes into account this effect

$$E(Y_{ij}) = \beta_i + \gamma t_{ij} \qquad i = 1,..., p, j = 1,..., N_i,$$
$$Var(Y_{ij}) = \sigma^2$$

where t_{ij} are the observations on concomitant variables (which are related to X_{ij}) and γ is the regression coefficient associated with t_{ij}. With this model, the variability of treatment effects can be considerably reduced.

For example, in any agricultural experimental, if the experimental units are plots of land then, t_{ij} can be measure of fertility characteristic of the j^{th} plot receiving i^{th} treatment and X_{ij} can be yield.

In another example, if experimental units are animals and suppose the objective is to compare the growth rates of groups of animals receiving different diets. Note that the observed differences in growth rates can be attributed to diet only if all the animals are similar in some observable characteristics like weight, age etc. which influence the growth rates.

In the absence of similarity, use t_{ij} which is the weight or age of j^{th} animal receiving i^{th} treatment.

If we consider the quadratic regression in t_{ij} then in

$$E(Y_{ij}) = \beta_i + \gamma_i t_{ij} + \gamma_2 t_{ij}^2, \quad i = 1,..., p, j = 1,..., n_i,$$
$$Var(Y_{ij}) = \sigma^2.$$

ANCOVA in this case is the same as ANCOVA with two concomitant variables t_{ij} and t_{ij}^2.

In two way classification with one observation per cell,

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j + \gamma t_{ij}, \quad i = 1,..., I, \ j = 1,..., J$$

or

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j + \gamma_i t_{ij} + \gamma_2 w_{ij}$$

with $\sum_i \alpha_i = 0, \ \sum_j \beta_j = 0$,

then (y_{ij}, t_{ij}) or (y_{ij}, t_{ij}, w_{ij}) are the observations in $(i, j)^{th}$ cell and t_{ij}, w_{ij} are the concomitment variables.

The concomitant variables can be fixed on random.

We consider the case of fixed concomitant variables only.

One way classification

Let $Y_{ij}\ (j=1...n_i,\ i=1...p)$ be a random sample of size n_i from i^{th} normal populations with mean

$$\mu_{ij} = E(Y_{ij}) = \beta_i + \gamma t_{ij}$$
$$Var(Y_{ij}) = \sigma^2$$

where β_i, γ and σ^2 are the unknown parameters, t_{ij} are known constants which are the observations on a concomitant variable.

The null hypothesis is

$$H_0 : \beta_1 = ... = \beta_p .$$

Let

$$\bar{y}_{io} = \frac{1}{n_i}\sum_j y_{ij};\ \ \bar{y}_{oj} = \frac{1}{p}\sum_i y_{ij},\ \ \bar{y}_{oo} = \frac{1}{n}\sum_i\sum_j y_{ij}$$
$$\bar{t}_{io} = \frac{1}{n_i}\sum_j t_{ij};\ \ \bar{t}_{oj} = \frac{1}{p}\sum_i t_{ij},\ \ \bar{t}_{oo} = \frac{1}{n}\sum_i\sum_j t_{ij}$$
$$n = \sum_i n_i .$$

Under the whole parametric space (π_Ω), use likelihood ratio test for which we obtain the $\hat{\beta}_i$'s and $\hat{\gamma}$ using the least squares principle or maximum likelihood estimation as follows:

Minimize $S = \sum_i\sum_j (y_{ij} - \mu_{ij})^2$

$$= \sum_i\sum_j (y_{ij} - \beta_i - \gamma t_{ij})^2$$

$$\frac{\partial S}{\partial \beta_i} = 0 \text{ for fixed } \gamma$$
$$\Rightarrow \beta_i = \bar{y}_{io} - \gamma \bar{t}_{io}$$

Put β_i in S and minimize the function by $\frac{\partial S}{\partial \gamma} = 0,$

i.e., minimize $\sum_i\sum_j \left[y_{ij} - \bar{y}_{io} - \gamma(t_{ij} - \bar{t}_{io}) \right]^2$ with respect to γ gives

$$\hat{\gamma} = \frac{\sum_i\sum_j (y_{ij} - \bar{y}_{io})(t_{ij} - \bar{t}_{io})}{\sum_i\sum_j (t_{ij} - \bar{t}_{io})^2}.$$

Thus $\hat{\beta}_i = \bar{y}_{io} - \hat{\gamma}\bar{t}_{io}$

$$\hat{\mu}_{ij} = \hat{\beta}_i + \hat{\gamma} t_{ij}.$$

Since $y_{ij} - \hat{\mu}_{ij} = y_{ij} - \hat{\beta}_i - \hat{\gamma} t_{ij}$
$= y_{ij} - \overline{y}_{io} - \hat{\gamma}(t_{ij} - \overline{t}_{io}),$

we have

$$\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2 = \sum \sum (y_{ij} - \overline{y}_{io})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \overline{y}_{io})(t_{ij} - \overline{t}_{io})\right]}{\sum_i \sum_j (t_{ij} - \overline{t}_{io})^2}.$$

Under $H_0 : \beta_1 = ... = \beta_p = \beta$ (say), consider $S_w = \sum_i \sum_j \left[y_{ij} - \beta - \gamma t_{ij} \right]^2$ and minimize S_w under sample space (π_w),

$$\frac{\partial S_w}{\partial \beta} = 0,$$
$$\frac{\partial S_w}{\partial \gamma} = 0$$
$$\Rightarrow \hat{\hat{\beta}} = \overline{y}_{oo} - \hat{\hat{\gamma}} \overline{t}_{oo}$$
$$\hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \overline{y}_{oo})(t_{ij} - \overline{t}_{oo})}{\sum_i \sum_j (t_{ij} - \overline{t}_{oo})^2}$$
$$\hat{\hat{\mu}}_{ij} = \hat{\hat{\beta}} + \hat{\hat{\gamma}} t_{ij}.$$

Hence

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j (y_{ij} - \overline{y}_{oo})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \overline{y}_{oo})(t_{ij} - \overline{t}_{oo})\right]^2}{\sum_i \sum_j (t_{ij} - \overline{t}_{oo})^2}$$

and

$$\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j \left[(\overline{y}_i - \overline{y}_{oo}) + \hat{\gamma}(t_{ij} - \overline{t}_{io}) - \hat{\hat{\gamma}}(t_{ij} - \overline{t}_{oo}) \right]^2.$$

The likelihood ratio test statistic in this case is given by

$$\lambda = \frac{\max_{w} L(\beta,\gamma,\sigma^2)}{\max_{\Omega} L(\beta,\gamma,\sigma^2)}$$

$$= \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}.$$

Now we use the following theorems:

Theorem 1: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follow a multivariate normal distribution $N(\mu, \Sigma)$ with mean vector μ and positive definite covariance matrix Σ. Then $Y'AY$ follows a noncentral chi-square distribution with p degrees of freedom and noncentrality parameter $\mu' A\mu$, i.e., $\chi^2(p, \mu' A\mu)$ if and only if ΣA is an idempotent matrix of rank p.

Theorem 2: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follows a multivariate normal distribution $N(\mu, \Sigma)$ with mean vector μ and positive definite covariance matrix Σ. Let $Y'A_1Y$ follows $\chi^2(p_1, \mu' A_1\mu)$ and $Y'A_2Y$ follows $\chi^2(p_2, \mu' A_2\mu)$. Then $Y'A_1Y$ and $Y'A_2Y$ are independently distributed if $A_1\Sigma A_2 = 0.$

Theorem 3: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follows a multivariate normal distribution $N(\mu, \sigma^2 I)$, then the maximum likelihood (or least squares) estimator $L'\hat{\beta}$ of estimable linear parametric function is independently distributed of $\hat{\sigma}^2; L\hat{\beta}$ follow $N\left[L'\beta, L'(X'X)^{-1}L\right]$ and $\frac{n\hat{\sigma}^2}{\sigma^2}$ follows $\chi^2(n-p)$ where $rank(X) = p.$

Using these theorems on the independence of quadratic forms and dividing the numerator and denominator by respective degrees of freedom, we have

$$F = \frac{n-p-1}{p-1} \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum \sum (y_{ij} - \hat{\mu}_{ij})^2} \sim F(p-1, n-p) \text{ under } H_0$$

So reject H_0 whenever $F \geq F_{1-\alpha}(p-1, n-p)$ at α level of significance.

The terms involved in λ can be simplified for computational convenience follows:

We can write

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2$$

$$= \sum_i \sum_j \left[y_{ij} - \hat{\hat{\beta}} - \hat{\hat{\gamma}} t_{ij} \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \overline{y}_{oo}) - \hat{\hat{\gamma}} (t_{ij} - \overline{t}_{oo}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \overline{y}_{oo}) - \hat{\hat{\gamma}} (t_{ij} - \overline{t}_{oo}) + \hat{\gamma}(t_{ij} - \overline{t}_{io}) - \hat{\gamma}(t_{ij} - \hat{t}_{io}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \overline{y}_{io}) - \hat{\gamma}(t_{ij} - \hat{t}_{io}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \overline{y}_{oo}) + \hat{\gamma}(t_{ij} - \overline{t}_{io}) - \hat{\hat{\gamma}}(t_{ij} - \overline{t}_{oo}) \right]^2$$

$$= \sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2 + \sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij}).$$

For computational convenience

$$\lambda = \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2} = \frac{\left(T_{yy} - \frac{T_{yt}^2}{T_{tt}} \right) - \left(E_{yy} - \frac{E_{yt}^2}{E_{tt}} \right)}{\left(E_{yy} - \frac{E_{yt}^2}{E_{yy}} \right)}$$

where

$$T_{yy} = \sum_i \sum_j (y_{ij} - \overline{y}_{oo})^2$$

$$T_{tt} = \sum_i \sum_j (t_{ij} - \overline{t}_{oo})^2$$

$$T_{yt} = \sum_i \sum_j (y_{ij} - \overline{y}_{oo})(t_{ij} - \overline{t}_{oo})$$

$$E_{yy} = \sum_i \sum_j (y_{ij} - \overline{y}_{io})^2$$

$$E_{tt} = \sum_i \sum_j (t_{ij} - \overline{t}_{io})^2$$

$$E_{yt} = \sum_i \sum_j (y_{ij} - \overline{y}_{io})(t_{ij} - \overline{t}_{io}).$$

Analysis of covariance table for one way classification is as follows:

Source of variation	Degrees of freedom	Sum of products			Adjusted sum of squares		F
		yy	yt	tt	Degress of feedom	Sum of squares	
Population	$p-1$	$P_{yy}(=T_{yy}-E_{yy})$	$P_{yt}(=T_{yt}-E_{yt})$	$P_{tt}(=T_{tt}-E_{tt})$	$p-1$	$q_1=q_0-q_2$	$\frac{n-p-1}{p-1}\frac{q_1}{q_2}$
Error	$n-p$	E_{yy}	E_{yt}	E_{tt}	$n-p-1$	$q_2=E_{yy}-\frac{E_{yt}^2}{E_{yy}}$	
Total	$n-1$	T_{yy}	T_{yt}	T_{tt}	$n-2$	$q_0=T_{yy}-\frac{T_{yt}^2}{T_{tt}}$	

If H_0 is rejected, employ multiple comprises methods to determine which of the contrasts in β_i are responsible for this.

For any estimable linear parametric contrast

$$\varphi=\sum_{i=1}^{p}C_i\beta_i \text{ with } \sum_{i=1}^{p}C_i=0,$$

$$\hat{\varphi}=\sum_{i=1}^{p}C_i\hat{\beta}_i=\sum_{i=1}^{p}C_i\overline{y}_i-\hat{\gamma}\sum_{i=1}^{p}C_{ii}\overline{t}_i$$

$$Var(\hat{\gamma})=\frac{\sigma^2}{\sum_i\sum_j(t_{ij}-\overline{t}_i)^2}$$

$$\Rightarrow Var(\hat{\varphi})=\sigma^2\left[\sum_i\frac{C_i^2}{n_i}+\frac{\left(\sum_i C_i\overline{t}_i\right)^2}{\sum_i\sum_j(t_{ij}-\overline{t}_i)^2}\right].$$

Two way classification (with one observations per cell)

Consider the case of two way classification with one observation per cell.

Let $y_{ij} \sim N(\mu_{ij}, \sigma^2)$ be independently distributed with

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + \gamma t_{ij}, \quad i = 1...I, j = 1...J$$
$$V(y_{ij}) = \sigma^2$$

where

μ : Grand mean

α_1 : Effect of i^{th} level of A satisfying $\sum_i^I \alpha_i = 0$

β_1: Effect of j^{th} level of B satisfying $\sum_i^J \beta_j = 0$

t_{ij} : observation (known) on concomitant variable.

The null hypothesis under consideration are

$$H_{0\alpha} : \alpha_1 = \alpha_2 = ... = \alpha_I = 0$$
$$H_{0\beta} : \beta_1 = \beta_2 = ... = \beta_J = 0$$

Dimension of whole parametric space $(\pi_\Omega) : I + J$

Dimension of sample space $(\pi_{w\alpha}) : J + 1$ under $H_{0\alpha}$

Dimension of sample space $(\pi_{w\beta}) : I + 1$ under $H_{0\beta}$

with respective alternative hypotheses as

$H_{1\alpha}$: At least one pair of α's is not equal

$H_{1\beta}$: At least one pair of β's is not equal.

Consider the estimation of parameters under the whole parametric space (π_Ω).

Find minimum value of $\sum_i \sum_j (y_{ij} - \mu_{ij})^2$ under π_Ω.

To do this, minimize

$$\sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma t_{ij})^2 .$$

For fixed γ, which gives on solving the least squares estimates (or the maximum likelihood estimates) of the respective parameters as

$$\mu = \bar{y}_{oo} - \gamma \bar{t}_o$$
$$\alpha_i = \bar{y}_i - \bar{y}_{oo} - \gamma(\bar{t}_{io} - \bar{t}_{oo}) \qquad (1)$$
$$\beta_j = \bar{y}_{oj} - \bar{y}_{oo} - \gamma(\bar{t}_{oj} - \bar{t}_{oo}).$$

Under these values of μ, α_i and β_{j}, the sum of squares $\sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma t_{ij})^2$ reduces to

$$\sum_i \sum_j \left[y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo} + \gamma(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}) \right]^2. \qquad (2)$$

Now minimization of (2) with respect to γ gives

$$\hat{\gamma} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})}{\sum_{i=1}^{I} \sum_{j=1}^{J} (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2}.$$

Using $\hat{\gamma}$, we get from (1)

$$\hat{\mu} = \bar{y}_{oo} - \hat{\gamma} \bar{t}_{oo}$$
$$\hat{\alpha}_i = (\bar{y}_{io} - \bar{y}_{oo}) - \hat{\gamma}(\bar{t}_{io} - \bar{t}_{oo})$$
$$\hat{\beta}_j = (\bar{y}_{oj} - \bar{y}_{oo}) - \hat{\gamma}(\bar{t}_{oj} - \bar{t}_{oo}).$$

Hence

$$\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2$$
$$= \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}) \right]}{\sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2}$$
$$= E_{yy} - \frac{E_{yt}^2}{E_{tt}}$$

where

$$E_{yy} = \sum_i \sum_j (y_{ij} - y_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2$$
$$E_{yt} = \sum \sum (y_{ij} - y_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})$$
$$E_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2.$$

Case (i) : Test of $H_{0\alpha}$

Minimize $\sum_i \sum_j (y_{ij} - \mu - \beta_j - \gamma t_{ij})^2$ with respect to μ, β_j and γ gives the least squares estimates) (or the maximum likelihood estimates) of respective parameters as

$$\Rightarrow \hat{\hat{\mu}} = \bar{y}_{oo} - \hat{\hat{\gamma}}\bar{t}_{oo}$$

$$\hat{\hat{\beta}}_j = \bar{y}_{oj} - \bar{y}_{oo} - \hat{\hat{\gamma}}(\bar{t}_{oj} - \bar{t}_{oo})$$

$$\hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{oj})(t_{ij} - \bar{t}_{oj})}{\sum_i \sum_j (t_{ij} - \bar{t}_{oj})^2} \qquad (3)$$

$$\hat{\hat{\mu}} = \hat{\hat{\mu}} + \hat{\hat{\beta}}_j + \hat{\hat{\gamma}} t_{ij}.$$

Substituting these estimates in (3) we get

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j (y_{ij} - \bar{y}_j)^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \bar{y}_{oj})(t_{ij} - \bar{t}_{oj})\right]}{\sum_i \sum_j (t_{ij} - \bar{t}_{oj})^2}$$

$$= E_{yy} + A_{yy} - \frac{\left[E_{yt} + A_{yt}\right]}{E_{tt} + A_{tt}}$$

where

$$A_{yy} = \sum_i J(\bar{y}_{io} - \bar{y}_{oo})^2$$

$$A_{tt} = \sum_i J(\bar{t}_{io} - \bar{t}_{oo})^2$$

$$A_{yt} = \sum_i J(\bar{y}_{io} - \bar{y}_{oo})(\bar{t}_{io} - \bar{t}_{oo})^2$$

$$E_{yy} = \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2$$

$$E_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2$$

$$E_{yt} = \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}).$$

Thus the likelihood ratio test statistic for testing $H_{0\alpha}$ is

$$\lambda_1 = \frac{\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 - \sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}.$$

Adjusting with degrees of freedom and using the earlier result for the independence of two quadratic forms and their distribution

$$F_1 = \frac{(IJ-I-J)}{(I-1)}\left[\frac{\sum_i\sum_j (y_{ij}-\hat{\hat{\mu}}_{ij})^2 - \sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}{\sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}\right] \sim F(I-1, IJ-I-J) \text{ under } H_{o\alpha}.$$

So the decision rule is to reject $H_{o\alpha}$ whenever $F_1 > F_{1-\alpha}(I-1, IJ-I-J).$

Case b: Test of $H_{0\beta}$

Minimize $\sum_i\sum_j (y_{ij}-\mu-\alpha_i-\gamma t_{ij})^2$ with respect to μ, α_i and γ gives the least squares estimates (or maximum likelihood estimates) of respective parameters as

$$\begin{aligned}
\hat{\mu} &= \overline{y}_{oo} - \tilde{\gamma}\overline{t}_{oo} \\
\tilde{\alpha}_j &= \overline{y}_{io} - \overline{y}_{oo} - \tilde{\gamma}(\overline{t}_{io} - \overline{t}_{oo}) \\
\tilde{\gamma} &= \frac{\sum_i\sum_j (y_{ij}-\overline{y}_{io})(t_{ij}-\overline{t}_{io})}{\sum_i\sum_j (t_{ij}-\overline{t}_{io})^2} \qquad (4) \\
\tilde{\mu}_{ij} &= \tilde{\mu} + \tilde{\alpha}_i + \tilde{\gamma}_{ij}.
\end{aligned}$$

From (4), we get

$$\begin{aligned}
\sum_i\sum_j (y_{ij}-\tilde{\mu}_{ij})^2 &= \sum_i\sum_j (y_{ij}-\overline{y}_{io})^2 - \frac{\left[\sum_i\sum_j (y_{ij}-\overline{y}_{io})(t_{ij}-\overline{t}_{oj})\right]}{\sum_i\sum_j (t_{ij}-\overline{t}_{io})^2} \\
&= E_{yy} + B_{yy} - \frac{\left[E_{yt}+B_{yt}\right]^2}{B_{tt}}
\end{aligned}$$

$$B_{yy} = \sum_j I(\overline{y}_{oj} - \overline{y}_{oo})^2$$

where $B_{tt} = \sum_j I(\overline{t}_{oj} - \overline{t}_{oo})^2$

$$B_{yt} = \sum_j I(\overline{y}_{io} - \overline{y}_{oo})(\overline{t}_{oj} - \overline{t}_{oo})^2.$$

Thus the likelihood ratio test statistic for testing $H_{0\beta}$ is

$$F_2 = \frac{(IJ-I-J)}{(J-1)}\left[\frac{\sum_i\sum_j (y_{ij}-\tilde{\mu}_{ij})^2 - \sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}{\sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}\right] \sim F(J-1, IJ-I-J) \text{ under } H_{o\beta}.$$

So the decision rule is to reject $H_{0\beta}$ whenever $F_2 \geq F_{1-\alpha}(J-1, IJ-I-J).$

If $H_{o\alpha}$ is rejected, use multiple comparison methods to determine which of the contrasts α_i are responsible for this rejection. The same is true for $H_{o\beta}$.

The analysis of covariance table for two way classification is as follows:

Source of variation	Degrees of freedom	Sum of products yy	yt	tt			F
Between evels of A	$I-1$	A_{yy}	A_{yt}	A_{tt}	$I-1$	$q_0 = q_3 - q_2$	$F_1 = \frac{IJ-I-J}{I-1}\frac{q_0}{q_2}$
Between levels of B	$J-1$	B_{yy}	B_{yt}	B_{tt}	$J-1$	$q_1 = q_4 - q_2$	$F_2 = \frac{IJ-I-J}{J-1}\frac{q_1}{q_2}$
Error	$(I-1)(J-1)$	E_{yy}	E_{yt}	E_{tt}	$IJ-I-J$	$q_2 = E_{yy} - \frac{E_{yt}^2}{E_{tt}}$	
Total	$IJ-1$	T_{yy}	T_{yt}	T_{tt}	$IJ-2$		
Error + levels of A	$IJ-J$					$q_3 = (A_{yy}+E_{yy}) - \frac{(A_{yt}+E_{yt})^2}{A_{tt}+E_{tt}}$	
Error + levels of B	$IJ-I$					$q_4 = (B_{yy}+E_{yt}) - \frac{(B_{yt}+E_{yt})^2}{B_{tt}+E_{tt}}$	