ONE WAY ANALYSIS OF VARIANCE

Let we suppose that 'N' observations y_{ij}; $(i=1,2,\cdots k, j=1,2,\cdots n_i)$ of a random variable Y are grouped, on some basis, into K-Classes of sizes $n_1, n_2, \cdots n_K$ respectively,

$$N = \sum_{i=1}^{K} n_i .$$

The linear model for analysis of one way classified data is

$$y_{ij} = \mu + \alpha_i + e_{ij} \; ; \; \begin{array}{l} i=1,2,\cdots k \\ j=1,2,\cdots n_i \end{array} \quad \cdots (*)$$

where y_{ij} = j^{th} observation in the i^{th} class

μ = general effect.

α_i = effect due to the i^{th} class.

e_{ij} = Errors $\sim N(0, \sigma_e^2)$ independently.

Also $\sum_{i=1}^{K} n_i \alpha_i = 0$

Estimation of the parameters:

μ and α_i are obtained by the method of least squares by minimising $E = \sum_i \sum_j e_{ij}^2$. The ~~estimates~~ ~~are~~ two normal equations are

$$\frac{\partial E}{\partial \mu} = 0 \quad \text{---(i)}$$

and

$$\frac{\partial E}{\partial \alpha_i} = 0 \quad \text{--(ii)}$$

(i) gives $\hat{\mu} = \bar{y}_{..}$

and (ii) gives $\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}$

where, $\bar{y}_{..} = \frac{1}{N} \sum_i \sum_j y_{ij}$ = grand mean

and $\bar{y}_{i.} = \frac{1}{n_i} \sum_j y_{ij}$ = mean of the i^{th} class

Partitioning Total Sum of Squares

Putting the value of $\hat{\mu}$, $\hat{\alpha}_i$ and $\hat{e}_{ij}$ in the model (*)

we get.

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})$$

Now,

$$(y_{ij} - \bar{y}_{..}) = (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})$$

Squaring both sides and taking summation over 'i' and 'j' we get.

$$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = \sum_i n_i (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2$$

[Product term vanishes due to sum of deviation from mean is zero]

$\therefore$ Total Sum of Squares (TSS) = Sum of Squares due to Class (SSC) + Sum of Squares due to Error (SSE)

Partitioning the degrees of freedom:

The total d.f. is also partitioned as

Total d.f = Class d.f + Error d.f

i.e. $(N-1) = (K-1) + (N-K)$

Null Hypothesis:

$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_K = 0$ i.e. there is no significant difference between the class means.

Alternative Hypothesis:

At least two of the class means are different from each other.

Mean Squares:

Mean Square due to class (MSC)

$$= \frac{SSC}{K-1}$$

Mean Square due to Error (MSE)

$$= \frac{SSE}{N-K}$$

Test statistic:

$$F = \frac{MSC}{MSE} \sim F_{K-1, N-K}$$ will give

is a test statistic for testing the null hypothesis. Ho will be rejected at α% level of significance if Calculated F value is greater than the tabulated value of F for $(K-1, N-K)$ d.f at α% level i.e. $F \geq F_{\alpha, K-1, N-K}$. otherwise Ho may be accepted.

ANOVA TABLE

Source of Variation	d.f	SS	M.S	F
Between Classes	$K-1$	SSC	MSC	$F = \frac{MSC}{MSE}$
Error	$N-K$	SSE	MSE	
Total	$N-1$	TSS	—	—