

# Analysis of Variance and Design of Experiments

General Linear Hypothesis and Analysis of Variance

∴

Lecture 10

Multiple Comparison Tests



Shalabh

Department of Mathematics and Statistics  
Indian Institute of Technology Kanpur



Slides can be downloaded from <http://home.iitk.ac.in/~shalab/sp1>

## One-way classification with fixed effect linear models of full rank:

Let  $y_{ij} (j = 1, 2, \dots, n_i)$  be a random sample from the  $i^{\text{th}}$  normal population with mean  $\beta_i$  and variance  $\sigma^2, i = 1, 2, \dots, p$ , i.e.,

$$Y_{ij} \sim N(\beta_i, \sigma^2), j = 1, 2, \dots, n_i; i = 1, 2, \dots, p.$$

The random samples from different populations are assumed to be independent of each other.

These observations follow the set up of linear model

$$Y = X\beta + \varepsilon$$

The null hypothesis of interest is  $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = \beta$  (say)

and  $H_1 : \text{At least one } \beta_i \neq \beta_j (i \neq j)$  where  $\beta$  and  $\sigma^2$  are unknown.

## Case of rejection of $H_0$

If  $F > F_{1-\alpha}(p-1, n-p)$ , then  $H_0 : \beta_1 = \beta_2 = \dots = \beta_p$  is rejected.

This means that at least one  $\beta_i$  is different from others which is responsible for the rejection.

So the objective is to investigate and find out such  $\beta_i$  and divide the population into groups such that the means of populations within the groups are the same.

This can be done by pairwise testing of  $\beta$ 's.

Test  $H_0 : \beta_i = \beta_k$  ( $i \neq k$ ) against  $H_1 : \beta_i \neq \beta_k$  .

## Student-Newman-Keuls test:

The Student-Newman-Keuls test compares the range is compared with  $100\alpha\%$  points on critical Studentized range  $W_p$  given by

$$W_p = q_{\alpha, p, \gamma} \sqrt{\frac{s^2}{n}}.$$

The observed range  $R = \bar{y}_p^* - \bar{y}_1^*$  is now compared with  $W_p$ .

Let the effects  $\beta_1, \beta_2, \dots, \beta_p$  are denoted as  $\beta_1^*, \beta_2^*, \dots, \beta_p^*$  corresponding to  $\bar{y}_1^*, \bar{y}_2^*, \dots, \bar{y}_p^*$  respectively

## Student-Newman-Keuls test:

- If  $R < W_p$  then stop the process of comparison and conclude that  $\beta_1^* = \beta_2^* = \dots = \beta_p^*$
- if  $R > W_p$  then
  - i. divide the ranked means  $\bar{y}_1^*, \bar{y}_2^*, \dots, \bar{y}_p^*$  into two subgroups containing  $(\bar{y}_p^*, \bar{y}_{p-1}^*, \dots, \bar{y}_2^*)$  and  $(\bar{y}_{p-1}^*, \bar{y}_{p-2}^*, \dots, \bar{y}_1^*)$ .
  - ii. Compute the ranges  $R_1 = \bar{y}_p^* - \bar{y}_2^*$  and  $R_2 = \bar{y}_{p-1}^* - \bar{y}_1^*$ .

Then compare the ranges  $R_1$  and  $R_2$  with  $W_{p-1}$ .

Continue with this procedure until a group of  $i$  means is found whose range does not exceed  $W_i$ .

## **Student-Newman-Keuls test:**

**By this method, the difference between any two means under test is significant when the range of the observed means of each and every subgroup containing the two means under test is significant according to the Studentized critical range.**

## Duncan's multiple comparison test:

The test procedure in Duncan's multiple comparison test is the same as in the Student-Newman-Keuls test except the observed ranges are compared with Duncan's  $100\alpha\%$  critical range

$$D_p = q_{\alpha_p, p, \gamma}^* \sqrt{\frac{s^2}{n}}$$

where  $\alpha_p = 1 - (1 - \alpha)^{p-1}$ ,  $q_{\alpha_p, p, \gamma}^*$  denotes the upper  $100\alpha\%$  points of the Studentized range based on Duncan's range.

Tables for Duncan's range are available.

## **Duncan's multiple comparison test:**

**Duncan felt that this test is better than the Student-Newman-Keuls test for comparing the differences between any two ranked means.**

**Duncan regarded that the Student-Newman-Keuls method is too stringent in the sense that the true differences between the means will tend to be missed too often.**



## **Duncan's multiple comparison test:**

**Duncan notes that in testing the equality of a subset  $k, (2 \leq k \leq p)$  means through null hypothesis, we are in fact testing whether  $(p - 1)$  orthogonal contrasts between the  $\beta$ 's differ from zero or not.**

## **Duncan's multiple comparison test:**

If these contrasts were tested in separate independent experiments, each at level  $\alpha$ , the probability of incorrectly rejecting the null hypothesis would be  $\left[1 - (1 - \alpha)^{p-1}\right]$ . So Duncan proposed to use  $\left[1 - (1 - \alpha)^{p-1}\right]$  in place of  $\alpha$  in the Student-Newman-Keuls test.

[Reference: Contributions to order statistics, Wiley 1962, Chapter 9 (Multiple decision and multiple comparisons, H.A. David, pages 147-148)].

## Case of unequal sample sizes:

When sample means are not based on the same number of observations, the procedures based on Studentized range, Student-Newman-Keuls test and Duncan's test are not applicable.

Kramer proposed that in Duncan's method, if a set of  $p$  means is to be tested for equality, then replace

$$q_{\alpha_p, p, \gamma}^* \frac{s}{\sqrt{n}} \text{ by } q_{\alpha_p, p, \gamma}^* s \sqrt{\frac{1}{2} \left( \frac{1}{n_U} + \frac{1}{n_L} \right)}$$

where  $n_U$  and  $n_L$  are the number of observations corresponding to the largest and smallest means in the data.

## Case of unequal sample sizes:

This procedure is only an approximate procedure but will tend to be conservative since means based on a small number of observations will tend to be overrepresented in the extreme groups of means.

Another option is to replace  $n$  by the harmonic mean of

$$n_1, n_2, \dots, n_p, \text{ i.e., } \frac{p}{\sum_{i=1}^p \left( \frac{1}{n_i} \right)}.$$

## The “Least Significant Difference” (LSD):

In the usual testing of  $H_0 : \beta_i = \beta_k$  against  $H_1 : \beta_i \neq \beta_k$ , the  $t$ -statistic

$$t = \frac{\bar{y}_{io} - \bar{y}_{ko}}{\sqrt{\widehat{Var}(\bar{y}_{io} - \bar{y}_{ko})}}$$

is used which follows a  $t$ -distribution, say with degrees of freedom ‘ $df$ ’. Thus  $H_0$  is rejected whenever

$$|t| > t_{df, 1 - \frac{\alpha}{2}}$$

and it is concluded that  $\beta_i$  and  $\beta_k$  are significantly different.

## The “Least Significant Difference” (LSD):

The inequality  $|t| > t_{df, 1-\frac{\alpha}{2}}$  can be equivalently written as

$$|\bar{y}_{io} - \bar{y}_{ko}| > t_{df, 1-\frac{\alpha}{2}} \sqrt{\widehat{Var}(\bar{y}_{io} - \bar{y}_{ko})}$$

If every pair of sample for which

$$|\bar{y}_{io} - \bar{y}_{ko}| \text{ exceeds } t_{df, 1-\frac{\alpha}{2}} \sqrt{\widehat{Var}(\bar{y}_{io} - \bar{y}_{ko})}$$

then this will indicate that the difference between  $\beta_i$  and  $\beta_k$  is significantly different. So according to this, the quantity

$$t_{df, 1-\frac{\alpha}{2}} \sqrt{\widehat{Var}(\bar{y}_{io} - \bar{y}_{ko})}$$

would be the least difference of  $\bar{y}_{io}$  and  $\bar{y}_{ko}$  for which it will be declared that the difference between  $\beta_i$  and  $\beta_k$  is significant.<sup>14</sup>

## The “Least Significant Difference” (LSD):

Based on this idea, we use the pooled variance of the two samples  $Var(\bar{y}_{io} - \bar{y}_{ko})$  as  $s^2$  and the Least Significant Difference (LSD) is defined as

$$LSD = t_{df, 1-\frac{\alpha}{2}} \sqrt{s^2 \left( \frac{1}{n_i} + \frac{1}{n_k} \right)}$$

If  $n_i = n_k = n$ , then

$$LSD = t_{df, 1-\frac{\alpha}{2}} \sqrt{\frac{2s^2}{n}}$$

Now all  $\frac{p(p-1)}{2}$  pairs of  $\bar{y}_{io}$  and  $\bar{y}_{ko}$ , ( $i \neq k = 1, 2, \dots, p$ ) are compared with *LSD*.

## The “Least Significant Difference” (LSD):

Use of *LSD* criterion may not lead to good results if it is used for comparisons suggested by the data (largest/smallest sample mean) or if all pairwise comparisons are done without correction of the test level.

If *LSD* is used for all the pairwise comparisons then these tests are not independent.

Such correction for test levels was incorporated in Duncan’s test.



## Tukey's "Honestly significant Difference" (HSD)

In this procedure, the Studentized rank values  $q_{\alpha, n, \gamma}$  are used in place of  $t$ -quantiles and the standard error of the difference of pooled mean is used in place of standard error of mean in common critical difference for testing  $H_0 : \beta_i = \beta_k$  against  $H_1 : \beta_i \neq \beta_k$ .

Tukey's Honestly Significant Difference is computed as

$$HSD = q_{1-\frac{\alpha}{2}, p, \gamma} \sqrt{\frac{MS_{error}}{n}}$$

assuming all samples are of the same size  $n$ .

## Tukey's "Honestly significant Difference" (*HSD*)

All  $\frac{p(p-1)}{2}$  pairs  $|\bar{y}_{io} - \bar{y}_{ko}|$  are compared with *HSD*.

If  $|\bar{y}_{io} - \bar{y}_{ko}| > HSD$  then  $\beta_i$  and  $\beta_k$  are significantly different.

## **Multiple comparison based on confidence intervals:**

**We notice that all the multiple comparison test procedure discussed up to now are based on the testing of hypothesis. There is one-to-one relationship between the testing of hypothesis and the confidence interval estimation. So the confidence interval can also be used for such comparisons.**

## Example

Before discussing these procedures, let us consider the following example which illustrates the relationship between the testing of hypothesis and confidence intervals.

**Example:** Consider the test of hypothesis for

$$H_0 : \beta_i = \beta_j \quad (i \neq j = 1, 2, \dots, p)$$

or  $H_0 : \beta_i - \beta_j = 0$

or  $H_0 : \text{contrast} = 0$

or  $H_0 : L = 0.$

The test statistic for  $H_0 : \beta_i = \beta_j$  is

$$t = \frac{(\hat{\beta}_i - \hat{\beta}_j) - (\beta_i - \beta_j)}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_i - \hat{\beta}_j)}} = \frac{\hat{L} - L}{\sqrt{\widehat{\text{Var}}(\hat{L})}}$$

## Example

where  $\hat{\beta}$  denotes the maximum likelihood (or least-squares) estimator of  $\beta$  and  $t$  follows a  $t$ -distribution with  $df$  degrees of freedom.

This statistic, in fact, can be extended to any linear contrast,

say e.g.,  $L = \beta_1 + \beta_2 - \beta_3 - \beta_4$ ,  $\hat{L} = \hat{\beta}_1 + \hat{\beta}_2 - \hat{\beta}_3 - \hat{\beta}_4$ .

The decision rule is

reject  $H_0 : L = 0$  against  $H_1 : L \neq 0$

if

$$|\hat{L}| > t_{df} \sqrt{\widehat{Var}(\hat{L})}.$$

## Example

The **100**  $(1 - \alpha)$  % confidence interval of  $L$  is obtained as

$$P \left[ -t_{df} \leq \frac{\hat{L} - L}{\sqrt{\widehat{\text{Var}}(\hat{L})}} \leq t_{df} \right] = 1 - \alpha$$

or

$$P \left[ \hat{L} - t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})} \leq L \leq \hat{L} + t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})} \right] = 1 - \alpha$$

so that the **100** $(1 - \alpha)$ % confidence interval of  $L$  is

$$\left[ \hat{L} - t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})}, \hat{L} + t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})} \right]$$

and

$$\hat{L} - t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})} \leq L \leq \hat{L} + t_{df} \sqrt{\widehat{\text{Var}}(\hat{L})}.$$

## Example

If this interval includes  $L = 0$  between lower and upper confidence limits, then  $H_0 : L = 0$  is accepted. Our objective is to know if the confidence interval contains zero or not.

Suppose for some given data the confidence intervals for  $\beta_1 - \beta_2$  and  $\beta_1 - \beta_3$  are obtained as  $-3 \leq \beta_1 - \beta_2 \leq 2$  and  $2 \leq \beta_1 - \beta_3 \leq 4$ .

## Example

Thus we find that the interval of  $\beta_1 - \beta_2$  includes zero which implies that  $H_0 : \beta_1 - \beta_2 = 0$  is accepted. Thus  $\beta_1 = \beta_2$ .

On the other hand interval of  $\beta_1 - \beta_3$  does not include zero and so  $H_0 : \beta_1 - \beta_3 = 0$  is not accepted. Thus  $\beta_1 \neq \beta_3$ .

If the interval of  $\beta_1 - \beta_3$  is  $-1 \leq \beta_1 - \beta_3 \leq 1$  then  $H_0 : \beta_1 = \beta_3$  is accepted.

If both  $H_0 : \beta_1 = \beta_2$  and  $H_0 : \beta_1 = \beta_3$ , we can conclude that  $\beta_1 = \beta_2 = \beta_3$ .