Fisher's Least Significant Difference (LSD) Method

We begin by describing Fisher's LSD method for constructing confidence intervals. The confidence interval for the difference $\mu_i - \mu_j$ is centered at the difference in sample means $\overline{X}_{i.} - \overline{X}_{j.}$. To determine how wide to make the confidence interval, it is necessary to estimate the standard deviation of $\overline{X}_{i.} - \overline{X}_{j.}$. Let J_i and J_j be the sample sizes at levels i and j, respectively. Since by assumption all observations are normally distributed with variance σ^2, it follows that $\overline{X}_{i.} - \overline{X}_{j.}$ is normally distributed with mean $\mu_i - \mu_j$ and variance $\sigma^2(1/J_i + 1/J_j)$. The variance σ^2 is estimated with MSE, for reasons explained previously in the discussion about confidence intervals for the treatment means (Section 9.1). Now the quantity

$$\frac{(\overline{X}_{i.} - \overline{X}_{j.}) - (\mu_i - \mu_j)}{\sqrt{\text{MSE}(1/J_i + 1/J_j)}}$$

has a Student's t distribution with $N - I$ degrees of freedom. (The value $N - I$ is the number of degrees of freedom used in computing MSE; see Equation 9.13.) The quantity $t_{N-I,\,\alpha/2}\sqrt{\text{MSE}(1/J_i + 1/J_j)}$ is called the least significant difference. This quantity forms the basis for confidence intervals and hypothesis tests.

Fisher's Least Significant Difference Method for Confidence Intervals and Hypothesis Tests

The Fisher's least significant difference confidence interval, at level $100(1 - \alpha)\%$, for the difference $\mu_i - \mu_j$ is

$$\overline{X}_{i.} - \overline{X}_{j.} \pm t_{N-I,\,\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)} \tag{9.32}$$

To test the null hypothesis $H_0 : \mu_i - \mu_j = 0$, the test statistic is

$$\frac{\overline{X}_{i.} - \overline{X}_{j.}}{\sqrt{\text{MSE}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}} \tag{9.33}$$

If H_0 is true, this statistic has a Student's t distribution with $N - I$ degrees of freedom. Specifically, if

$$|\overline{X}_{i.} - \overline{X}_{j.}| > t_{N-I,\,\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)} \tag{9.34}$$

then H_0 is rejected at level α.

The reason that the quantity $t_{N-I,\,\alpha/2}\sqrt{\text{MSE}(1/J_i + 1/J_j)}$ is called the least significant difference is that the null hypothesis of equal means is rejected at level α whenever the difference in sample means $|\overline{X}_{i.} - \overline{X}_{j.}|$ exceeds this value. When the design is balanced, with all sample sizes equal to J, the least significant difference is equal to $t_{N-I,\,\alpha/2}\sqrt{2\text{MSE}/J}$ for all pairs of means.

Example 9.9

In the weld experiment discussed in Section 9.1, hardness measurements were made for five welds from each of four fluxes A, B, C, and D. The sample mean hardness values were $\overline{X}_{A.} = 253.8$, $\overline{X}_{B.} = 263.2$, $\overline{X}_{C.} = 271.0$, and $\overline{X}_{D.} = 262.0$. The following output (from MINITAB) presents the ANOVA table.

```
One-way ANOVA: A, B, C, D

Source     DF        SS        MS        F        P
Factor      3    743.40   247.800     3.87    0.029
Error      16   1023.60    63.975
Total      19   1767.00

S = 7.998     R-Sq = 42.07%      R-Sq(adj) = 31.21%
```

Before the experiment was performed, the carbon contents of the fluxes were measured. Flux B had the lowest carbon content (2.67% by weight), and flux C had the highest (5.05% by weight). The experimenter is therefore particularly interested in comparing the hardnesses obtained with these two fluxes. Find a 95% confidence interval for the difference in mean hardness between welds produced with flux B and those produced with flux C. Can we conclude that the two means differ?

Solution

We use expression (9.32). The sample means are 271.0 for flux C and 263.2 for flux B. The preceding output gives the quantity MSE as 63.975. (This value was also computed in Example 9.3 in Section 9.1.) The sample sizes are both equal to 5. There are $I = 4$ levels and $N = 20$ observations in total. For a 95% confidence interval, we consult the t table to find the value $t_{16,\,.025} = 2.120$. The 95% confidence interval is therefore $271.0 - 263.2 \pm 2.120\sqrt{63.975(1/5 + 1/5)}$ or $(-2.92, 18.52)$.

To perform a test of the null hypothesis that the two treatment means are equal, we compute the value of the test statistic (expression 9.33) and obtain

$$\frac{271.0 - 263.2}{\sqrt{63.975(1/5 + 1/5)}} = 1.54$$

Consulting the t table with $N - I = 16$ degrees of freedom, we find that P is between $2(0.05) = 0.10$ and $2(0.10) = 0.20$ (note that this is a two-tailed test). We cannot conclude that the treatment means differ.

If it is desired to perform a fixed-level test at level $\alpha = 0.05$ as an alternative to computing the P-value, the critical t value is $t_{16,.025} = 2.120$. The left-hand side of the inequality (9.34) is $|271.0 - 263.2| = 7.8$. The right-hand side is $2.120\sqrt{63.975(1/5+1/5)} = 10.72$. Since 7.8 does not exceed 10.72, we do not reject H_0 at the 5% level.

The following output (from MINITAB) presents 95% Fisher LSD confidence intervals for each difference between treatment means in the weld experiment.

```
Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons

Simultaneous confidence level = 81.11%

A subtracted from:

    Lower  Center   Upper  -----+---------+---------+---------+---
B  −1.324   9.400  20.124              (--------*--------)
C   6.476  17.200  27.924                  (--------*--------)
D  −2.524   8.200  18.924             (--------*--------)
                           ------+---------+---------+---------+---
                            −12       0       12      24
B subtracted from:

     Lower  Center   Upper  ------+---------+---------+---------+---
C   −2.924   7.800  18.524              (--------*-------)
D  −11.924  −1.200   9.524      (--------*-------)
                            ------+---------+---------+---------+---
                             −12       0       12      24
C subtracted from:

     Lower  Center  Upper  ------+---------+---------+---------+---
D  −19.724  −9.000  1.724  (--------*------)
                           ------+---------+---------+---------+---
                            −12       0       12      24
```

The values labeled "Center" are the differences between pairs of treatment means. The quantities labeled "Lower" and "Upper" are the lower and upper bounds, respectively, of the confidence interval. Of particular note is the simultaneous confidence level of 81.11%. This indicates that although we are 95% confident that any given confidence interval contains its true difference in means, we are only 81.11% confident that *all* the confidence intervals contain their true differences.

In Example 9.9, a single test was performed on the difference between two specific means. What if we wanted to test every pair of means, to see which ones we could conclude to be different? It might seem reasonable to perform the LSD test on each pair. However, this is not appropriate, because when several tests are performed, the likelihood of rejecting a true null hypothesis increases. This is the multiple testing problem, which is discussed in some detail in Section 6.14. This problem is revealed