MS Module 11: Single-Factor ANOVA, Tukey's procedure – practice problems

(The attached PDF file has better formatting.)

Exercise 11.1: Tukey's honestly statistical differences

Five groups of 7 observations each have sample means of

$$\bar{x}_1 = 11.2,\ \bar{x}_2 = 12.1,\ \bar{x}_3 = 11.4,\ \bar{x}_4 = 11.1, \text{ and } \bar{x}_5 = 11.7$$

The total sum of squares (SST) is 24, and the treatment sum of squares (SSTr) is 12.

The groups have normal distributions with equal variances. We test the null hypothesis $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ at a 5% confidence level. If we reject the null hypothesis, we test which groups means differ significantly.

A. What is the error sum of squares SSE?
B. What is the mean square for the groups (treatment mean square)?
C. What is the mean squared error MSE?
D. What is the *F* value to test the null hypothesis?
E. What is the critical *F* value for α = 5%?
F. What is the critical *Q* value (studentized range distribution) for α = 5%?
G. What is *w* (the width of the difference) for Tukey's honestly statistical difference?
H. Which pairs of groups have different means by Tukey's analysis?

Part A: The exercise gives the total sum of squares and the treatment sum of squares; we derive the error sum of squares by subtraction:

$$\text{SSE} = \text{SST} - \text{SSTr} = 24 - 12 = 12.$$

Part B: The mean square is the sum of squares / degrees of freedom. The degrees of freedom for the groups is 5 – 1 = 4, so the mean square for the groups (the mean treatment sum of squares) is 12 / 4 = 3.

Part C: The degrees of freedom for the error is 5 × (7 – 1) = 30, so the MSE = 12 / 30 = 0.40.

Part D: The *F* value to test the null hypothesis is 3 / 0.40 = 7.50.

Part E: In the table of *F* values, $F_{0.05,4,30}$ = 2.69 (2.6896), so we reject the null hypothesis at a 5% significance level.

Part F: In the table of *Q* values, $Q_{0.05,5,30}$ = 4.10. The first degrees of freedom (second subscript of Q) is the number of groups (not the number of groups – 1).

Part G: The formula for *w* is (page 566 of second edition of the textbook or page 554 of the first edition):

$w = Q_{\alpha,\, i,\, i \times (j-1)} \times (\text{MSE} / j)^{½} = 4.10 \times (0.40 / 7)^{½} = 0.9801$

Part H: $\bar{x}_2 - \bar{x}_4$ = 12.1 – 11.1 = 1.00, which is greater than 0.9801, so groups 2 and 4 have different means. The other pair differences are all less than 0.9801, so they are not significant at a 5% level.

Question: How do we use the underlining procedure in the textbook?

Answer: We arrange the group means in ascending order:

$$\bar{x}_4 = 11.1,\ \bar{x}_1 = 11.2,\ \bar{x}_3 = 11.4,\ \bar{x}_5 = 11.7,\ \bar{x}_2 = 12.1$$

We draw a line under $\overline{x}_4 = 11.1$, $\overline{x}_1 = 11.2$, $\overline{x}_3 = 11.4$, $\overline{x}_5 = 11.7$, since the maximum difference is less than *w*. We do *not* continue this line under $\overline{x}_2 = 12.1$

We draw a second line under $\overline{x}_1 = 11.2$, $\overline{x}_3 = 11.4$, $\overline{x}_5 = 11.7$, $\overline{x}_2 = 12.1$, since the maximum difference is less than *w*.

The only two groups not connected by an underline are groups 2 and 4.

$$\overline{x}_1 = 11.2,\ \overline{x}_2 = 12.1,\ \overline{x}_3 = 11.4,\ \overline{x}_4 = 11.1, \text{ and } \overline{x}_5 = 11.7$$

Question: The term *honestly statistical difference* sounds unusual. A statistical difference is a number; it is not honest or dishonest.

Answer: Some statisticians in Tukey's day were computing the significance of each difference between groups Tukey reasoned that even all groups had the same mean and no group difference was significant at a 5% level, the likelihood that one group difference might appear significant depends on the number of groups.

Illustration: If an experiment has three groups with three groups differences (1 vs 2; 2 vs 3; 1 vs 3), and the *p* value for each difference is 5%, the likelihood that at least one difference would appear significant is more than 5%. If the differences were independent, the likelihood is $1 - (1 - 5\%)^3 = 14.2625\%$.

The group differences are not independent in these problems. Tukey developed a method for testing whether a group difference is significant given that we observe several interdependent group differences.