DOE – Least square difference (LSD)

As a following up of the last blog on the discussion of the completely randomized design experiments, suppose the analysis of variance shows statistical evidence to reject the null hypothesis Ho of equal populations. The next question we should then ask is where the mean differences occur. Which is the culprit that makes such a difference?

Fisher's least square difference (LSD) procedure in making pairwise comparison of population means can be adopted to determine where such differences occur.

Consider the following experiment. In order to test whether the mean time needed to mix a batch of medicine in inert materials is the same for machines produced by three manufacturers, the ABC Chemical Co. obtained the following data on the time (in minutes) needed to mix these materials.

Manufacturer		
A	B	C
20	28	20
26	26	19
24	31	23
22	27	22

In here, $\bar{x}_A$ = 23; $\bar{x}_B$ = 28; $\bar{x}_C$ = 21.

We shall use these data to test whether the population mean times (in minutes) for mixing a batch of material differ for the three manufacturers. If they do, how do we find out which manufacturer stands out in this comparison?

By using Excel Statistical Tools pack for one-way ANOVA, we obtain the following outcomes:

ANOVA: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Manufacturer A	4	92	23	6.667
Manufacturer B	4	112	28	4.667
Manufacturer C	4	84	21	3.333

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	104	2	52	10.64	0.0043	4.26
Within Groups	44	9	4.89			
Total	148	11				

It is noted that the MSE = 4.89 and is obvious from the above ANOVA estimation that there are mean time differences amongst the three manufacturers as the F value of 10.64 is larger than the critical F value of 4.26 at degrees of freedom 2 and 9.

The LSD procedure tests the significant difference between two population means, usually the two closer means. The following table summarizes the LSD procedure:

FISCHER'S LSD PROCEDURE

$$H_o: \quad \mu_1 = \mu_2$$
$$H_1: \quad \mu_1 \neq \mu_2$$

TEST STATISTIC

$$t = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \qquad [1]$$

REJECTION RULE

P-value approach : Reject H_o if P-value $\leq \alpha$

Critical value approach : Reject H_o if $t \leq -t_{\alpha/2}$ or $t \geq t_{\alpha/2}$

As we would like to test if the two closer population means do differ, we can re-arrange the above test table as follows:

FISCHER'S LSD PROCEDURE BASED ON THE TEST STATISTIC $\bar{x}_i - \bar{x}_j$

$$H_o: \quad \mu_1 = \mu_2$$
$$H_1: \quad \mu_1 \neq \mu_2$$

TEST STATISTIC

$$\bar{x}_i - \bar{x}_j$$

REJECTION RULE AT A SIGNIFICANCE LEVEL α

Reject H_O if $|\bar{x}_i - \bar{x}_j| \geq LSD$

where

$$LSD = t_{\alpha/2}\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \qquad [2]$$

For the above example, the value of LSD at $\alpha/2=0.05$ is:

$$LSD = 2.262\sqrt{4.89\left(\frac{1}{4} + \frac{1}{4}\right)} = 3.54$$

Since $\bar{x}_A - \bar{x}_c = 2 < 3.54, and\ \bar{x}_B - \bar{x}_A = 5 > 3.54,$ it is clear therefore that there was a significant difference of means between the manufacturers A and B with 95% confidence.