Subsequent Inferences for one-way ANOVA

- if the overall F test does not show significant differences among the groups, no further inferences are required
- if the overall test does show a significant difference, differences between particular means can be tested using

$$T = \frac{\bar{x}_{i.} - \bar{x}_{k.}}{\sqrt{MSE}\sqrt{\frac{1}{n_i} + \frac{1}{n_k}}}$$

- in these expressions, MSE is the estimate of σ^2, and the degrees of freedom are $N - a$, the same as for MSE
- however, adjustments must be made for *simultaneous inference i.e.* for the fact that several tests are being done
- the simplest adjustment is the *Bonferroni correction*, which reduces the significance level for each test so that the overall significance level is no larger than the desired level
- in a one-way ANOVA with a groups, there are $r = \binom{a}{2}$ natural comparisons between pairs of groups
- if you do r tests at level α, then the probability of rejecting at least one H_0 incorrectly could be as large as $r\alpha$

– for example for $r = 2$

$$P(reject\ at\ least\ one\ H_0) =$$

$$P(reject\ 1st) + P(reject\ 2nd)$$
$$-P(reject\ both) \leq 2\alpha$$

- to control the overall level, or *experimentwise error rate*, at α, each test should be done using $\alpha_* = \alpha/r$
- alternatively the P value should be multiplied by r
- similarly for r confidence intervals, use of α_* will give simultaneous confidence level $1 - \alpha$
- the confidence intervals for the difference in two means is

$$(\bar{x}_{i.} - \bar{x}_{k.} - t_{N-a}^{\alpha*/2}\sqrt{MSE}\sqrt{\frac{1}{n_i} + \frac{1}{n_k}},$$

$$\bar{x}_{i.} - \bar{x}_{k.} + t_{N-a}^{\alpha*/2}\sqrt{MSE}\sqrt{\frac{1}{n_i} + \frac{1}{n_k}})$$

Example: for the golf balls, the summary statistics are

i	$\bar{x}_i$	s_i^2	n_i
1	251.28	33.487	5
2	261.98	18.197	5
3	269.66	27.253	5

- the value for MSE is 26.312

- there are 3 possible pairwise comparisons between the groups
- the denominator of the test statistics is

$$\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_k}} = 5.1295(.6325) = 3.24$$

- the degrees of freedom are 12, and with $\alpha = .05$ and $\alpha_* = .05/3 = .0167$, the critical value of t is 2.7794
- this can be obtained using the commands

```
MTB > invcdf .00833;
SUBC> t 12.

Inverse Cumulative Distribution Function

Student's t distribution with 12 DF

P(X<=x)          x
    0.00833  -2.77969
```

- the test statistics are

$$t_{12} = \frac{251.18 - 261.98}{3.24} = -3.329$$

$$t_{13} = \frac{251.18 - 269.66}{3.24} = -5.697$$

and

$$t_{23} = \frac{261.18 - 269.66}{3.24} = -2.367$$

- the first two comparisons are significant at the .05 level but the third one is not
- confidence intervals for the differences in means are

$$-10.8 \pm 2.78(3.24) \quad or \quad (-19.81, -1.79)$$

$$-18.48 \pm 9.01 \quad or \quad (-27.49, -9.47)$$

and

$$-7.68 \pm 9.01 \quad or \quad (-16.69, 1.33)$$

Example: for the liver weights, the means in ascending order are

diet	B	C	A	D
n_i	8	6	7	8
mean	3.43	3.598	3.803	3.935

- the estimated standard deviation is $\sqrt{MSE} = .1899$
- there are 6 comparisons, so the appropriate table value for $\alpha = .05$ is $t_{.025/6,25} = 2.8649$, from MINITAB

- the pairwise differences in the means are

i/k	B	C	D
A	.373	.205	-.132
B		-.168	-.505
C		-.337	

- the absolute difference in means must exceed $t_{25}^{.025/6}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_k}}$, which
depends on the two sample sizes.

n_i/n_k	7	8
6	.3025	.2938
7		.2818
8		.2720

- using this table, we find that B and C, C and A and A and D are not statistically significant, the other 3 comparisons are significant

To carry out the multiple comparison procedure, you need to carry out all pairwise tests as described above, or construct all pairwise confidence intervals, with the appropriate adjustment of level, as follows.

In general, where $\alpha^* = \alpha/(\text{number of pairwise comparisons})$, the simultaneous confidence intervals for $\mu_i - \mu_j$ are given by

$$\bar{x}_{i.} - \bar{x}_{j.} \pm t_{\alpha^*/2, N-a} \sqrt{MSE} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

The anova table and summary statistics for the Diet data are as follows.

Source	DF	SS	MS	F	P
C2	3	1.1649	0.3883	10.84	0.000
Error	25	0.8954	0.0358		
Total	28	2.0603			

i	N	Mean	StDev	Diet
1	7	3.8029	0.2512	A
2	8	3.4300	0.1353	B
3	6	3.5983	0.1675	C
4	8	3.9363	0.1884	D

$\sqrt{MSE} = \sqrt{.0358} = .1899$ and there were 25 degrees of freedom for error.

The null hypothesis of equal treatment means was rejected at level $\alpha = .05$.

There are possible 6 pairwise comparisons, so $\alpha^* = \alpha/6 = .05/6 = .00833$, and $t_{\alpha^*/2,25} = t_{.00416,25} = 2.8649$, so

$$t_{\alpha^*/2,N-a}\sqrt{MSE} = 2.8649(.1899) = .544$$

i	j	n_i	n_j	$\bar{X}_{i.}$	$\bar{X}_{j.}$	CI for $\mu_i - \mu_j$	conclusion
1	2	7	8	3.8029	3.4300	(0.0913 , 0.6544)	$\mu_A \neq \mu_B$
1	3	7	6	3.8029	3.5983	(-0.0981 , 0.5073)	$\mu_A = \mu_C$
1	4	7	8	3.8029	3.9363	(-0.4149 , 0.1481)	$\mu_A = \mu_D$
2	3	8	6	3.4300	3.5983	(-0.4621 , 0.1255)	$\mu_B = \mu_C$
2	4	8	8	3.4300	3.9363	(-0.7783 ,-0.2343)	$\mu_B \neq \mu_D$
3	4	6	7	3.5983	3.9363	(-0.6407 ,-0.0353)	$\mu_C \neq \mu_D$

For example,

$$(3.8029 - 3.4300) - .544\sqrt{1/7 + 1/8} = .3729 - .2815 = .0914$$

One way in which the results are sometimes summarized is to:

- list the group means in ascending order
- label them according to their treatment
- starting with the smallest average, underline that average and any which are not significantly different from it
- repeat this with each subsequent average
- any two averages which are not underlined by the same line are significantly different