STAT 420

The cutting speeds of four types of tools are being compared in an experiment. Five materials of varying degree of hardness are to be used as experimental blocks. Measurements of cutting time (in seconds) according to types of tool (Factor A) and Hardness of Material (Factor B) are given in the table below.

I Factor A levels, J Factor B levels

Factor A	Factor B 1	2	3	4	5	Factor A Means
1	12	2	8	1	7	$\bar{y}_{1\bullet} = 6$
2	20	14	17	12	17	$\bar{y}_{2\bullet} = 16$
3	13	7	13	8	14	$\bar{y}_{3\bullet} = 11$
4	11	5	10	3	6	$\bar{y}_{4\bullet} = 7$
Factor B Means	$\bar{y}_{\bullet 1} = 14$	$\bar{y}_{\bullet 2} = 7$	$\bar{y}_{\bullet 3} = 12$	$\bar{y}_{\bullet 4} = 6$	$\bar{y}_{\bullet 5} = 11$	$\bar{y}_{\bullet\bullet} = 10$

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\left(y_{ij} - \bar{y}_{\bullet\bullet}\right)^2 = J\sum_{i=1}^{I}\left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right)^2 + I\sum_{j=1}^{J}\left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)^2 + \sum_{i=1}^{I}\sum_{j=1}^{J}\left(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}\right)^2$$

SST	SSA	SSB	SSR
$IJ - 1$	$I - 1$	$J - 1$	$(I-1)(J-1)$

ANOVA table:

Source	**SS**	**DF**	**MS**	**F**
Factor A	310	3	103.3333	51.66667
Factor B	184	4	46	23
Residuals	24	12	2	
Total	518	19		

$$y_{ij} = \bar{y}_{\bullet\bullet} + \left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right) + \left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right) + \left(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}\right)$$

$$\begin{bmatrix} 12 & 2 & 8 & 1 & 7 \\ 20 & 14 & 17 & 12 & 17 \\ 13 & 7 & 13 & 8 & 14 \\ 11 & 5 & 10 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 \end{bmatrix}$$

$$+ \begin{bmatrix} -4 & -4 & -4 & -4 & -4 \\ 6 & 6 & 6 & 6 & 6 \\ 1 & 1 & 1 & 1 & 1 \\ -3 & -3 & -3 & -3 & -3 \end{bmatrix} + \begin{bmatrix} 4 & -3 & 2 & -4 & 1 \\ 4 & -3 & 2 & -4 & 1 \\ 4 & -3 & 2 & -4 & 1 \\ 4 & -3 & 2 & -4 & 1 \end{bmatrix}$$

$$+ \begin{bmatrix} 2 & -1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ -2 & -1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & -2 \end{bmatrix}$$

SST $\quad \sum_{i=1}^{I}\sum_{j=1}^{J}\left(y_{ij} - \bar{y}_{\bullet\bullet}\right)^2 \quad$ $IJ - 1$ d.f.

SSA $\quad \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right)^2 = J\sum_{i=1}^{I}\left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right)^2 \quad$ $I - 1$ d.f.

SSB $\quad \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)^2 = I\sum_{j=1}^{J}\left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)^2 \quad$ $J - 1$ d.f.

SSR $\quad \sum_{i=1}^{I}\sum_{j=1}^{J}\left(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}\right)^2 \quad$ $(I-1)(J-1)$ d.f.

$$\text{SSA} = 5 \times [\,(6-10)^2 + (16-10)^2 + (11-10)^2 + (7-10)^2\,]$$

$$= 5 \times [\,16 + 36 + 1 + 9\,] = 310.$$

$$\text{SSB} = 4 \times [\,(14-10)^2 + (7-10)^2 + (12-10)^2 + (6-10)^2 + (11-10)^2\,]$$

$$= 4 \times [\,16 + 9 + 4 + 16 + 1\,] = 184.$$

$$\text{SSR} = (12-6-14+10)^2 + (2-6-7+10)^2 + \ldots + (6-7-11+10)^2 = 24.$$

$$\text{SST} = (12-10)^2 + (2-10)^2 + \ldots + (6-10)^2 = 518.$$

```
> Time = c(12,2,8,1,7,20,14,17,12,17,13,7,13,8,14,11,5,10,3,6)
> A = c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4)
> B = c(1,2,3,4,5,1,2,3,4,5,1,2,3,4,5,1,2,3,4,5)
> results = glm(Time ~ factor(A) + factor(B))
> summary(aov(results))
            Df Sum Sq Mean Sq F value    Pr(>F)    
factor(A)    3 310.00  103.33  51.667 3.911e-07 ***
factor(B)    4 184.00   46.00  23.000 1.489e-05 ***
Residuals   12  24.00    2.00                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \qquad i = 1, 2, 3, 4, \quad j = 1, 2, 3, 4, 5,$$

ε_{ij} are independent $N(0, \sigma^2)$ random variables.

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0, \qquad \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 = 0,$$

$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

```
> qf(0.95,3,12)
[1] 3.490295
> qf(0.99,3,12)
[1] 5.952545
```

$F = 51.66667$

Reject H_0

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

```
> qf(0.95,4,12)
[1] 3.259167
> qf(0.99,4,12)
[1] 5.411951
```

$F = 23$

Reject H_0

Six children are tested for pulse rate before and after watching a violent movie with the following results.

Pulse Rate

Child	*Before*	*After*
1	102	112
2	96	108
3	89	94
4	104	112
5	90	102
6	85	96

(a) Using the paired t test, test for differences in the before and after mean pulse rates. Let $\gamma = 0.05$, and use a two-sided test.

(b) Test for differences in the pulse rate means employing a two-factor fixed effects model and the F test. Again, let $\gamma = 0.05$. Are the results consistent with those of part (a)?

(c) Square the observed t value and the corresponding critical point. Compare the results with the observed F ratio and the F critical point, respectively. They should (within rounding error) be the same. Comment on the importance of these results.

(d) Using the t and F distributions, calculate 95-percent confidence intervals for the difference in mean pulse rates. Comment on the results.

(e) Compare the values of MSR from (b) and s_D^2 from (a).

For (a) and (d):

Matched Pair Comparison:

Pair			Difference
1	X_1	Y_1	$D_1 = X_1 - Y_1$
2	X_2	Y_2	$D_2 = X_2 - Y_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	X_n	Y_n	$D_n = X_n - Y_n$

Assume that the differences $D_i = X_i - Y_i$ are a random sample from normal distribution with mean δ and standard deviation σ_D.

A confidence interval for δ is $\overline{D} \pm t_{\alpha/2} \cdot \frac{s_D}{\sqrt{n}}$. The number of degrees of freedom $= n - 1$.

To test the hypothesis $H_0 : \delta = \delta_0$, use the test statistic $t = \frac{\overline{D} - \delta_0}{s_D / \sqrt{n}}$. $n - 1$ degrees of freedom

a)

```
> ybefore = c(102,96,89,104,90,85)
> yafter = c(112,108,94,112,102,96)
> ydiff = yafter - ybefore
> ydiff
[1] 10 12  5  8 12 11
> dbar = mean(ydiff); dbar
[1] 9.666667
> vardiff = var(ydiff); vardiff
[1] 7.466667
> sdiff = sqrt(vardiff); sdiff
[1] 2.73252
> tteststat = dbar/(sdiff/sqrt(6))   # test statistic
> tteststat
[1] 8.665407
> tcrit = qt(0.975,5)   # critical value
> tcrit
[1] 2.570582
> 2*(1-pt(tteststat,5))   # p-value
[1] 0.0003383023
```

OR

```
> t.test(yafter,ybefore,paired=TRUE)

        Paired t-test

data:  yafter and ybefore
t = 8.6654, df = 5, p-value = 0.0003383
alternative hypothesis: true difference in means is not equal to
0
95 percent confidence interval:
  6.799063 12.534271
sample estimates:
mean of the differences
               9.666667
```

b), c)

```
> y = c(ybefore,yafter)
> y
 [1] 102  96  89 104  90  85 112 108  94 112 102  96
> when = c(rep(1,6),rep(2,6))
> when
 [1] 1 1 1 1 1 1 2 2 2 2 2 2
> child = c(rep(1:6,2))
> child
 [1] 1 2 3 4 5 6 1 2 3 4 5 6
>
> results = glm(y ~ factor(when) + factor(child))
> summary(aov(results))
              Df Sum Sq Mean Sq F value    Pr(>F)    
factor(when)   1 280.33 280.333  75.089 0.0003383 ***
factor(child)  5 582.67 116.533  31.214 0.0008918 ***
Residuals      5  18.67   3.733                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> tteststat^2
[1] 75.08929
>
> fcrit = qf(0.95,1,5); fcrit
[1] 6.607891
> tcrit^2
[1] 6.607891
```

Recall: One-way ANOVA is equivalent to two-sample *t* test with pooled variance if comparing two population means. One-way ANOVA can also be used to compare more than two population means.

Two-way ANOVA (without replications) is equivalent to paired *t* test if comparing two levels of one factor. Two-way ANOVA can also be used to compare more than two levels of one factor.

(For example, we can compare the pulse rate before, in the middle, immediately after, and 2 hours after the movie.)

d)

```
> dbar - tcrit*sdiff/sqrt(6)
[1] 6.799063
> dbar + tcrit*sdiff/sqrt(6)
[1] 12.53427
```

OR

```
> t.test(ydiff,mu=0,alt="two.sided",conf.level=0.95)

        One Sample t-test

data:  ydiff
t = 8.6654, df = 5, p-value = 0.0003383
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  6.799063 12.534271
sample estimates:
mean of x
 9.666667
```

If the F tests reject the notion of no factor main effects, then it is of interest to compare the effects contributed by different levels of the factor. We can employ an extension of Scheffé's method of multiple comparisons, introduced in Theorem 4.1, for this purpose.

Let μ_i, $i = 1, 2, \ldots, I$, and μ_j, $j = 1, 2, \ldots, J$, denote the means for the ith level of factor 1 and the jth level of factor 2 [see equation (4.28)]. Since the I levels of factor 1 correspond to the rows and the J levels of factor 2 correspond to the columns of the two-way layout, it is convenient to refer to the μ_i's and μ_j's as the "row" and "column" means, respectively.

Theorem 4.2 Let

$$\psi_1 = \sum_{i=1}^{I} c_i\mu_i \qquad \text{and} \qquad \psi_2 = \sum_{j=1}^{J} c_j\mu_j$$

be contrasts in the row and column means, respectively. Then with $100 \times (1-\gamma)$-percent confidence, *all* contrasts in the means ψ_1 and ψ_2 are bracketed by the bounds,

$$\sum_{i=1}^{I} c_i\bar{Y}_i \pm \sqrt{F_\gamma(I-1, \nu)}\, s\, \sqrt{\frac{(I-1)}{JK}\sum_{i=1}^{I} c_i^2} \tag{4.37}$$

and

$$\sum_{j=1}^{J} c_j\bar{Y}_j \pm \sqrt{F_\gamma(J-1, \nu)}\, s\, \sqrt{\frac{(J-1)}{IK}\sum_{j=1}^{J} c_j^2} \tag{4.38}$$

respectively, where s^2 is the residual mean square from the appropriate ANOVA table, $F_\gamma(\nu_1, \nu_2)$ is the upper γ percentage point of an F distribution with ν_1 and ν_2 degrees of freedom, and

Replicated layout		*Unreplicated layout*	
$\bar{Y}_i = \bar{Y}_{i..}$,	$i = 1, 2, \ldots, I$	$\bar{Y}_i = \bar{Y}_{i.}$,	$i = 1, 2, \ldots, I$
$\bar{Y}_j = \bar{Y}_{.j.}$,	$j = 1, 2, \ldots, J$	$\bar{Y}_j = \bar{Y}_{.j}$,	$j = 1, 2, \ldots, J$
$\nu = IJ(K-1)$		$K = 1$	
		$\nu = (I-1)(J-1)$	

$\mu_{\text{after}} - \mu_{\text{before}}$ $\qquad$ $c_{\text{after}} = 1, \quad c_{\text{before}} = -1.$

```
> MSR = sum(results$residuals^2)/5; MSR
[1] 3.733333
>
> mean(yafter)-mean(ybefore) - sqrt(qf(0.95,1,5))*sqrt(MSR)*sqrt((1/6)*2)
[1] 6.799063
> mean(yafter)-mean(ybefore) + sqrt(qf(0.95,1,5))*sqrt(MSR)*sqrt((1/6)*2)
[1] 12.53427
```

Two intervals are the same.

e) $\quad \mathbf{2} \times \text{MSR} = s_{\text{D}}^{2}.$

STAT 420

A two-factor analysis of variance experiment was performed with $I = 3$, $J = 2$, and $K = 4$ (a 3×2 factorial experiment with 4 replicates).

Factor A	Factor B 1	Factor B 2
1	23 18 17 20	20 16 15 19
2	26 23 20 27	30 24 29 27
3	23 21 24 16	27 19 21 23

a) Test at the 5% significance level to determine if factors A and B interact.

b) Test at the 5% significance level to determine if differences exist among the levels of factor A.

c) Test at the 5% significance level to determine if differences exist among the levels of factor B.

$$y_{ijk} = \bar{y}_{\bullet\bullet\bullet} + \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) + \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) + \left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right) + \left(y_{ijk} - \bar{y}_{ij\bullet}\right)$$

Two factors are said to **interact** if the difference between levels (treatment) of one factor depends on the level of the other factor.

(some combinations of levels of factors A and B result in higher responses and some result in lower)

Factors that do not interact are called additive.

$$y_{ijk} = \bar{y}_{\bullet\bullet\bullet} + \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}\right) + \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}\right)$$

$$+ \left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right) + \left(y_{ijk} - \bar{y}_{ij\bullet}\right)$$

$$\begin{bmatrix} 23 & 20 \\ 18 & 16 \\ 17 & 15 \\ 20 & 19 \\ \\ 26 & 30 \\ 23 & 24 \\ 20 & 29 \\ 27 & 27 \\ \\ 23 & 27 \\ 21 & 19 \\ 24 & 21 \\ 16 & 23 \end{bmatrix} = \begin{bmatrix} 22 & 22 \\ 22 & 22 \\ 22 & 22 \\ 22 & 22 \\ \\ 22 & 22 \\ 22 & 22 \\ 22 & 22 \\ 22 & 22 \\ \\ 22 & 22 \\ 22 & 22 \\ 22 & 22 \\ 22 & 22 \end{bmatrix} + \begin{bmatrix} -3.5 & -3.5 \\ -3.5 & -3.5 \\ -3.5 & -3.5 \\ -3.5 & -3.5 \\ \\ 3.75 & 3.75 \\ 3.75 & 3.75 \\ 3.75 & 3.75 \\ 3.75 & 3.75 \\ \\ -0.25 & -0.25 \\ -0.25 & -0.25 \\ -0.25 & -0.25 \\ -0.25 & -0.25 \end{bmatrix} + \begin{bmatrix} -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix}$$

$$+ \begin{bmatrix} 1.5 & -1.5 \\ 1.5 & -1.5 \\ 1.5 & -1.5 \\ 1.5 & -1.5 \\ \\ -1.25 & 1.25 \\ -1.25 & 1.25 \\ -1.25 & 1.25 \\ -1.25 & 1.25 \\ \\ -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \end{bmatrix} + \begin{bmatrix} 3.5 & 2.5 \\ -1.5 & -1.5 \\ -2.5 & -2.5 \\ 0.5 & 1.5 \\ \\ 2 & 2.5 \\ -1 & -3.5 \\ -4 & 1.5 \\ 3 & -0.5 \\ \\ 2 & 4.5 \\ 0 & -3.5 \\ 3 & -1.5 \\ -5 & 0.5 \end{bmatrix}$$

SST $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(y_{ijk}-\bar{y}_{\bullet\bullet\bullet}\right)^2$ $IJK-1$ d.f.

SSA $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)^2 = JK\sum_{i=1}^{I}\left(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)^2$ $I-1$ d.f.

SSB $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)^2 = IK\sum_{j=1}^{J}\left(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet}\right)^2$ $J-1$ d.f.

SSAB $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)^2$ $(I-1)(J-1)$ d.f.

$\displaystyle = K\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet}\right)^2$ $= IJ-I-J+1$ d.f.

SSR $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(y_{ijk}-\bar{y}_{ij\bullet}\right)^2$ $IJ(K-1)$ d.f.

ANOVA table:

Source	**SS**	**DF**	**MS**	**F**
Factor A	211	2	105.5	11.72222222
Factor B	6	1	6	0.666666667
Interaction	31	2	15.5	1.722222222
Residuals	162	18	9	
Total	410	23		

$\bar{y}_{11\bullet} = 19.5$ $\bar{y}_{21\bullet} = 24$ $\bar{y}_{31\bullet} = 21$

$\bar{y}_{12\bullet} = 17.5$ $\bar{y}_{22\bullet} = 27.5$ $\bar{y}_{32\bullet} = 22.5$

$\bar{y}_{1\bullet\bullet} = 18.5$ $\bar{y}_{2\bullet\bullet} = 25.75$ $\bar{y}_{3\bullet\bullet} = 21.75$

$\bar{y}_{\bullet 1\bullet} = 21.5$ $\bar{y}_{\bullet 2\bullet} = 22.5$ $\bar{y}_{\bullet\bullet\bullet} = 22$

$$\text{SSA} = JK \sum_{i=1}^{I} \left(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet} \right)^2 = 2 \times 4 \times [\,(18.5 - 22)^2 + (25.75 - 22)^2 + (21.75 - 22)^2\,]$$

$$= 8 \times [\,12.25 + 14.0625 + 0.0625\,] = 211.$$

$$\text{SSB} = IK \sum_{j=1}^{J} \left(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet} \right)^2 = 3 \times 4 \times [\,(21.5 - 22)^2 + (22.5 - 22)^2\,]$$

$$= 12 \times [\,0.25 + 0.25\,] = 6.$$

$$\text{SSAB} = K \sum_{i=1}^{I} \sum_{j=1}^{J} \left(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet} \right)^2$$

$$= 4 \times [\,(19.5 - 18.5 - 21.5 + 22)^2 + (17.5 - 18.5 - 22.5 + 22)^2$$

$$+ (24 - 25.75 - 21.5 + 22)^2 + (27.5 - 25.75 - 22.5 + 22)^2$$

$$+ (21 - 21.75 - 21.5 + 22)^2 + (22.5 - 21.75 - 22.5 + 22)^2\,] = 31.$$

$$\text{SSR} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(y_{ijk} - \bar{y}_{ij\bullet} \right)^2 = (23 - 19.5)^2 + (18 - 19.5)^2 + \ldots + (23 - 22.5)^2 = 162.$$

$$\text{SST} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(y_{ijk} - \bar{y}_{\bullet\bullet\bullet} \right)^2 = (23 - 22)^2 + (18 - 22)^2 + \ldots + (23 - 22)^2 = 410.$$

```
> Y <- c(23,18,17,20,26,23,20,27,23,21,24,16,
+         20,16,15,19,30,24,29,27,27,19,21,23)
> A <- c(1,1,1,1,2,2,2,2,3,3,3,3,1,1,1,1,2,2,2,2,3,3,3,3)
> B <- c(1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2,2,2)
```

```
> results2 <- glm(Y ~ factor(A) * factor(B))
> summary(aov(results2))
                    Df Sum Sq Mean Sq F value    Pr(>F)    
factor(A)            2  211.0   105.5 11.7222 0.0005499 ***
factor(B)            1    6.0     6.0  0.6667 0.4248908    
factor(A):factor(B)  2   31.0    15.5  1.7222 0.2068328    
Residuals           18  162.0     9.0                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, \qquad i = 1, 2, 3, \quad j = 1, 2, \quad k = 1, 2, 3, 4.$$

ε_{ijk} are independent $N(0, \sigma^2)$ random variables,

$$\alpha_1 + \alpha_2 + \alpha_3 = 0, \qquad \beta_1 + \beta_2 = 0,$$

$$(\alpha\beta)_{1j} + (\alpha\beta)_{2j} + (\alpha\beta)_{3j} = 0, \qquad j = 1, 2,$$

$$(\alpha\beta)_{i1} + (\alpha\beta)_{i2} = 0, \qquad i = 1, 2, 3.$$

a) $H_0: (\alpha\beta)_{11} = (\alpha\beta)_{12} = (\alpha\beta)_{21} = (\alpha\beta)_{22} = (\alpha\beta)_{31} = (\alpha\beta)_{32} = 0$

```
> qf(0.95,2,18)
[1] 3.554557
```

F = 1.7222. **Do NOT Reject H_0** Interaction $A \times B$ is NOT significant.

b) $H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0$

```
> qf(0.95,2,18)
[1] 3.554557
```

F = 11.7222. **Reject H_0** Factor A IS significant.

c) $H_0: \beta_1 = \beta_2 = 0$

```
> qf(0.95,1,18)
[1] 4.413873
```

F = 0.6667. **Do NOT Reject H_0** Factor B is NOT significant.

Fitting an additive model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}, \qquad i = 1, 2, 3, \quad j = 1, 2, \quad k = 1, 2, 3, 4.$$

ε_{ijk} are independent $N(0, \sigma^2)$ random variables,

$$\alpha_1 + \alpha_2 + \alpha_3 = 0, \qquad\qquad \beta_1 + \beta_2 = 0.$$

```
> results3 <- glm(Y ~ factor(A) + factor(B))
> summary(aov(results3))
                    Df  Sum Sq Mean Sq F value     Pr(>F)
factor(A)     2     211  105.50   10.933 0.000619 ***
factor(B)     1       6    6.00    0.622 0.439641
Residuals    20     193    9.65
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

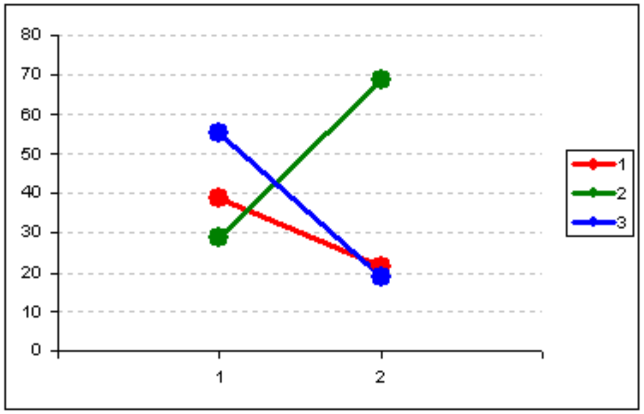
= =

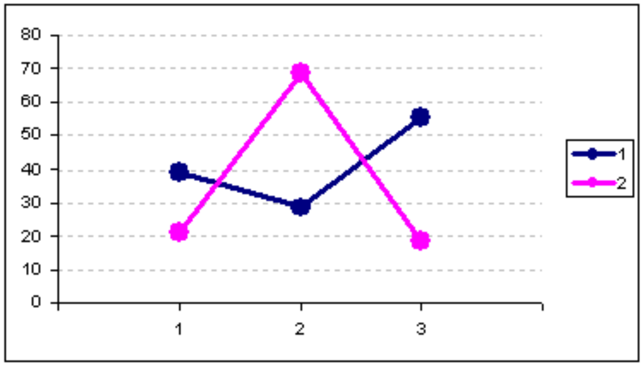
Recall Examples for 10/30/2012 (part 1):

```
> Time = c(12,2,8,1,7,20,14,17,12,17,13,7,13,8,14,11,5,10,3,6)
> A = c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4)
> B = c(1,2,3,4,5,1,2,3,4,5,1,2,3,4,5,1,2,3,4,5)
> results = glm(Time ~ factor(A) + factor(B))
> summary(aov(results))
            Df  Sum Sq Mean Sq F value    Pr(>F)
factor(A)     3     310   103.3    51.67 3.91e-07 ***
factor(B)     4     184    46.0    23.00 1.49e-05 ***
Residuals    12      24     2.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
>
>
>
> results4 = glm(Time ~ factor(A) * factor(B))
> summary(aov(results4))
                    Df  Sum Sq Mean Sq
factor(A)            3     310   103.3
factor(B)            4     184    46.0
factor(A):factor(B) 12      24     2.0
```

STAT 420

Factor A	Factor B 1	Factor B 2	Factor B 3
1	56 23 52 28 35	43 25 16 27 32	47 43 52 61 74
2	16 14 18 27 31	58 62 68 72 83	15 14 22 16 27

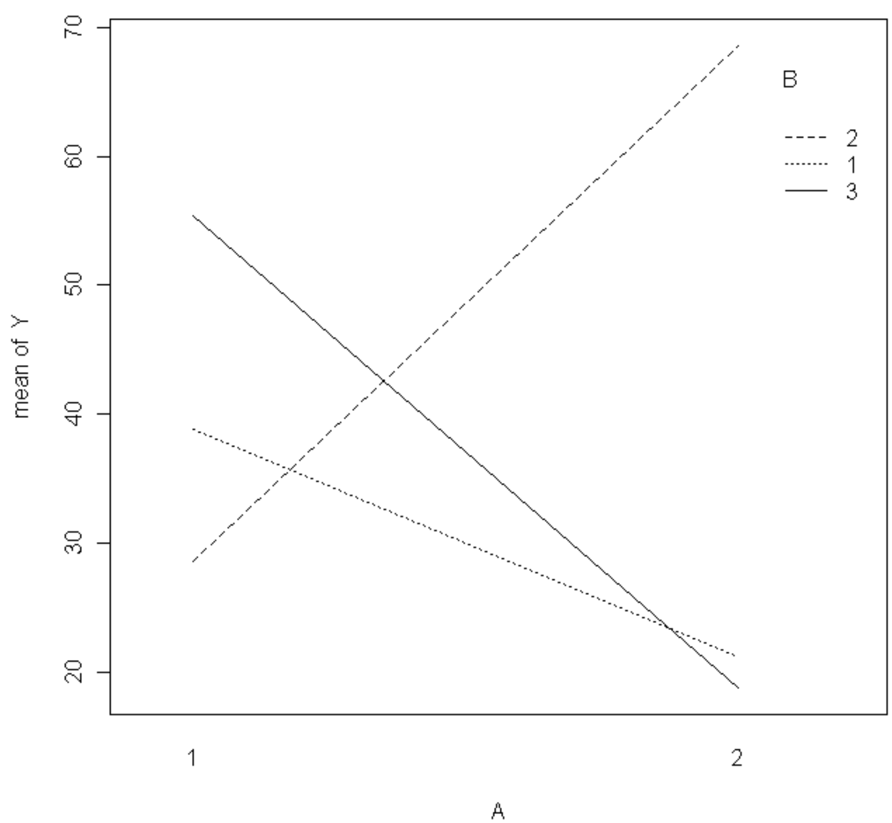




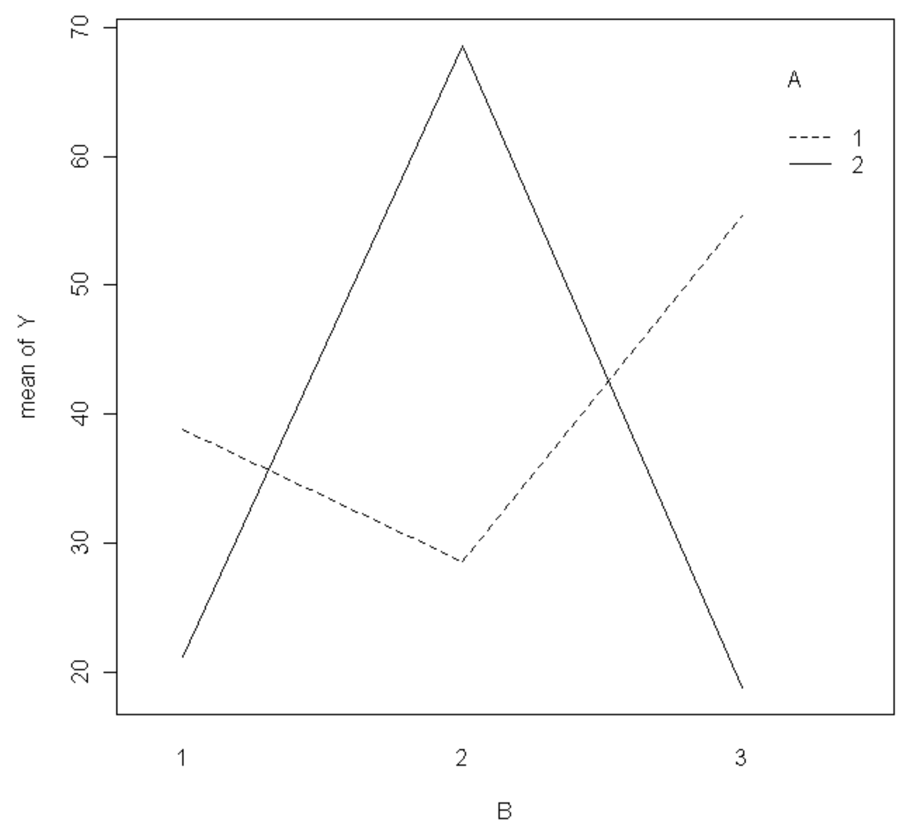
Means:

Factor A	Factor B 1	Factor B 2	Factor B 3	
1	38.8	28.6	55.4	40.93333
2	21.2	68.6	18.8	36.2
	30	48.6	37.1	38.56667

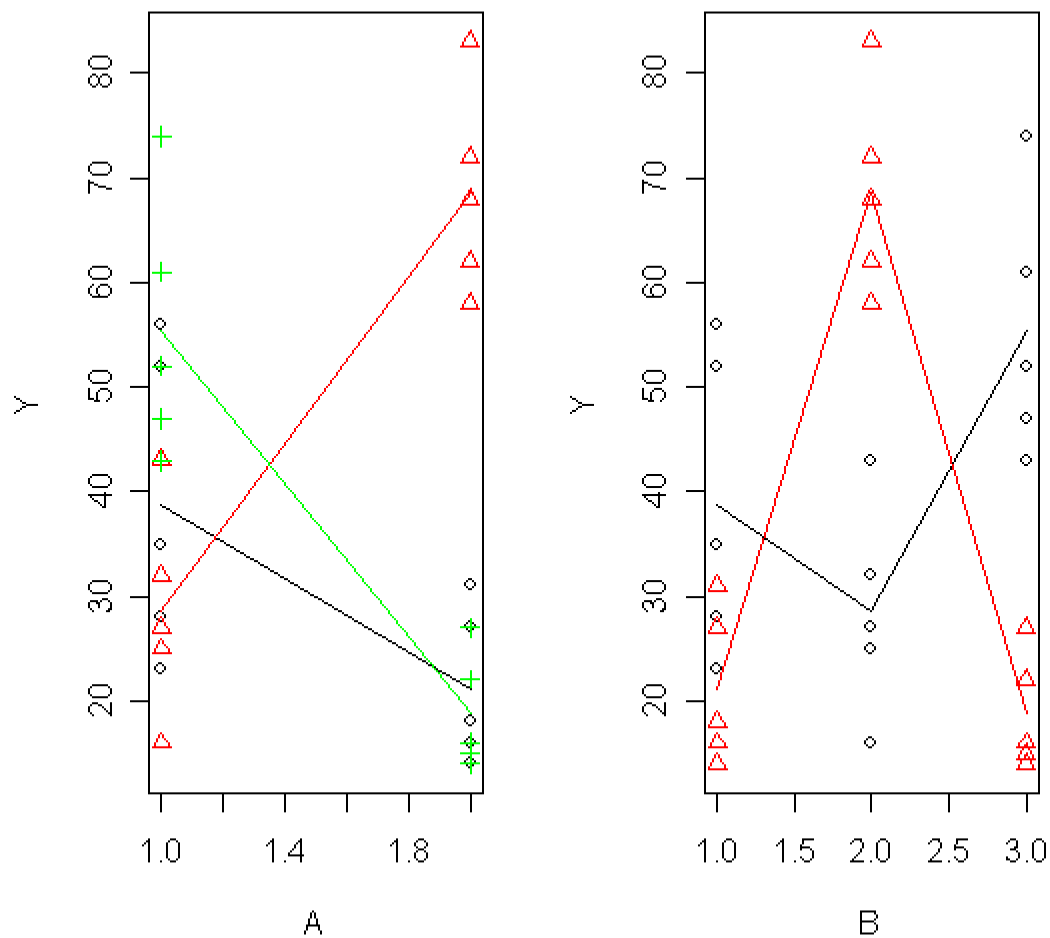
```
> Y = c(56, 23, 52, 28, 35, 43, 25, 16, 27, 32, 47, 43, 52, 61, 74,
+       16, 14, 18, 27, 31, 58, 62, 68, 72, 83, 15, 14, 22, 16, 27)
> A = c( 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)
> B = c( 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3)
> results = aov(glm(Y ~ factor(A) * factor(B)))
> summary(results)
                    Df Sum Sq Mean Sq F value    Pr(>F)
factor(A)            1  168.0   168.0  1.5667  0.222743
factor(B)            2 1762.1   881.0  8.2148  0.001915 **
factor(A):factor(B)  2 7955.3  3977.6 37.0875 4.555e-08 ***
Residuals           24 2574.0   107.2
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> interaction.plot(A,B,Y)
```




```
> interaction.plot(B,A,Y)
```



```
> # Cell means
> tmp = NULL
> for(a in 1:2)
+ {
+  for(b in 1:3)
+  {
+   tmp = c(tmp,mean(Y[A==a & B==b]))
+  }
+ }
> means = matrix(tmp,nrow=2,byrow=TRUE)
> tmp
[1] 38.8 28.6 55.4 21.2 68.6 18.8
> means
     [,1] [,2] [,3]
[1,] 38.8 28.6 55.4
[2,] 21.2 68.6 18.8
>
> # for 2 plots on the same page
> par(mfrow=c(1,2))
>
> plot(Y~A,col=B,pch=B)
> for(b in 1:3)
+ {
+  lines(1:2,means[,b],col=b)
+ }
>
> plot(Y~B,col=A,pch=A)
> for(a in 1:2)
+ {
+  lines(1:3,means[a,],col=a)
+ }
```



STAT 420

Consider a two-factor analysis of variance experiment was performed with $I = 2$, $J = 3$, and $K = 2$ $\left(\text{a } 2 \times 3 \text{ factorial experiment with 2 replicates}\right)$:

	B1	B2	B3
A1	Y_{111} Y_{112}	Y_{121} Y_{122}	Y_{131} Y_{132}
A2	Y_{211} Y_{212}	Y_{221} Y_{222}	Y_{231} Y_{232}

A1 – base category $\quad$ $\mathbf{v}_2$ – indicator of A2

(In general, will need $I - 1$ dummy variables for I levels of factor A.)

B1 – base category $\quad$ $\mathbf{w}_2$ – indicator of B2 $\quad$ $\mathbf{w}_3$ – indicator of B3

(In general, will need $J - 1$ dummy variables for J levels of factor B.)

Then will need $(I - 1) \times (J - 1)$ interaction terms $\mathbf{v}_i \, \mathbf{w}_j$.

	β_0	β_1	β_2	β_3	β_4	β_5
Y_{111} Y_{112}	1 1	0 0	0 0	0 0	0 0	0 0
Y_{121} Y_{122}	1 1	0 0	1 1	0 0	0 0	0 0
Y_{131} Y_{132}	1 1	0 0	0 0	1 1	0 0	0 0
Y_{211} Y_{212}	1 1	1 1	0 0	0 0	0 0	0 0
Y_{221} Y_{222}	1 1	1 1	1 1	0 0	1 1	0 0
Y_{231} Y_{232}	1 1	1 1	0 0	1 1	0 0	1 1
	$\mathbf{1}$	$\mathbf{v}_2$	$\mathbf{w}_2$	$\mathbf{w}_3$	$\mathbf{v}_2\,\mathbf{w}_2$	$\mathbf{v}_2\,\mathbf{w}_3$

$$\mathbf{Y} = \beta_0 \mathbf{1} + \beta_1 \mathbf{v}_2 + \beta_2 \mathbf{w}_2 + \beta_3 \mathbf{w}_3 + \beta_4 \mathbf{v}_2 \mathbf{w}_2 + \beta_5 \mathbf{v}_2 \mathbf{w}_3 + \boldsymbol{\varepsilon}.$$

11	β_0
12	$\beta_0 + \beta_2$
13	$\beta_0 + \beta_3$
21	$\beta_0 + \beta_1$
22	$\beta_0 + \beta_1 + \beta_2 + \beta_4$
23	$\beta_0 + \beta_1 + \beta_3 + \beta_5$

Interaction: $H_0: \beta_4 = \beta_5 = 0.$

(In general, $(I-1) \times (J-1)$ parameters.)

Factor A: $H_0: \beta_1 = 0.$

(In general, $I-1$ parameters.)

Factor B: $H_0: \beta_2 = \beta_3 = 0.$

(In general, $J-1$ parameters.)

Residuals DF $n - p$

$$= IJK - [(I-1) + (J-1) + (I-1)\times(J-1) + 1)]$$

$$= IJK - IJ = IJ(K-1).$$