# MA4605 Lecture 12 B - ANOVA table for $2^3$ Factorial Design

In the last class, we looked at the procedure for computing estimates for the main effects, the interaction effect and the sum of squares for the ANOVA table.

We will continue with a discussion of the ANOVA table itself.

As before – the key calculations in the ANOVA table are the mean square values and consequently the F test statistics. A full table for all factorial designs is presented below.

Starting on the left hand side, the list of main and interaction effects. For the case of three factors, there are seven effects in total, and also the residuals.

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | Expected Mean Squares | $F_0$ |
|---|---|---|---|---|---|
| $A$ | $SS_A$ | $a-1$ | $MS_A$ | $\sigma^2 + \frac{bcn\sum\tau_i^2}{a-1}$ | $\frac{MS_A}{MS_E}$ |
| $B$ | $SS_B$ | $b-1$ | $MS_B$ | $\sigma^2 + \frac{acn\sum\beta_j^2}{b-1}$ | $\frac{MS_B}{MS_E}$ |
| $C$ | $SS_C$ | $c-1$ | $MS_C$ | $\sigma^2 + \frac{abn\sum\gamma_k^2}{c-1}$ | $\frac{MS_C}{MS_E}$ |
| $AB$ | $SS_{AB}$ | $(a-1)(b-1)$ | $MS_{AB}$ | $\sigma^2 + \frac{cn\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$ | $\frac{MS_{AB}}{MS_E}$ |
| $AC$ | $SS_{AC}$ | $(a-1)(c-1)$ | $MS_{AC}$ | $\sigma^2 + \frac{bn\sum\sum(\tau\gamma)_{ik}^2}{(a-1)(c-1)}$ | $\frac{MS_{AC}}{MS_E}$ |
| $BC$ | $SS_{BC}$ | $(b-1)(c-1)$ | $MS_{BC}$ | $\sigma^2 + \frac{an\sum\sum(\beta\gamma)_{jk}^2}{(b-1)(c-1)}$ | $\frac{MS_{BC}}{MS_E}$ |
| $ABC$ | $SS_{ABC}$ | $(a-1)(b-1)(c-1)$ | $MS_{ABC}$ | $\sigma^2 + \frac{n\sum\sum\sum(\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$ | $\frac{MS_{ABC}}{MS_E}$ |
| Error | $SS_E$ | $abc(n-1)$ | $MS_E$ | $\sigma^2$ | |
| Total | $SS_T$ | $abcn-1$ | | | |

The sum of squares for each effect, and residuals is presented on the next column. As per last class, we have learned how to compute the Sum of Squares for effects from first principles. (The residual sum of squares will be provided directly in any exam question).

The summation of the sums of squares for effects and residuals yields the total sum of squares.

The next column presents the degrees of freedom for each corresponding effect.  For the main effects – the degrees of freedom is the number of levels (i.e. k-1). In the case of $2^3$ design, each factor has two level. Hence the corresponding degrees of freedom is simply 1 (i.e. 2-1) for each main effect.

For the interaction effects, the degrees of freedom are the products of the degrees of freedom for the underlying main effects. Again in the case of $2^3$ design, the degrees of freedom will be 1.

The Degrees of freedom for residuals is the difference between the sum of degrees of freedom and the total degrees of freedom ( total number of observations -1).

As with all other ANOVA procedures, the mean squares are the ratios of the sums of squares to the corresponding degrees of freedom (i.e. SSx/dfx). In the case of $2^3$ design, the degrees of freedom for effects are all 1 in value. Hence the sums of squares and mean square having the same value. For the case of mean square for residuals, the calculation proceeds in the conventional manner.

The F –value test statistics are compute as the ratio of each mean square for an effect and residual (e.g. $MS_A/MS_{resid}$). A demonstration will follow shortly.

The corresponding p-values are computed from the F distribution, with the degrees of freedom for the effect and residuals (e.g 1,8). However in this module calculation will not be required, as they will be presented as part of the output.

From the example, used in both this weeks Lab classes and Lectures, the ANOVA table is implemented in R and presented below:

```
> summary(Model2)
            Df Sum Sq Mean Sq F value  Pr(>F)
A            1    189     189   1.199 0.30540
B            1   3570    3570  22.640 0.00143 **
C            1    977     977   6.193 0.03761 *
A:B          1     86      86   0.543 0.48239
A:C          1    588     588   3.729 0.08956 .
B:C          1   1314    1314   8.333 0.02030 *
A:B:C        1    203     203   1.288 0.28932
Residuals    8   1262     158
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1
```

The p-values ( i.e. `Pr(>F)`) indicate the outcome of a test of significance for each effect.

The null hypothesis is the main effect is not significant (i.e. has no bearing on the response variable).

The rule of thumb of rejecting the null hypothesis only if the p-value is less than 0.01 is overly stringent. Instead, a significant result is indicated by the presence of asterisks. (If there is an asterisk, reject null hypothesis – Students may use this rule of thumb anywhere in the module.)

From the output, we see that main effects for factors B and C are significant (agitation and temperature).

Also significant is the interaction effect between the two factors. Of interest is the rather low p-value for the interaction effect for A and C, despite the fact that A is not significant. We do not accept the result as significant, but it is interesting that it got flagged by a ".". ( In an exam situation, students may disregard this, but in practical usage, it may warrant a re-appraisal of how the experiment was performed).

**<u>Regression Model</u>**

We can formulate our results, constructing a regression model that we can use to predict responses. As we have seen previously, the important factors are B, C and the interaction effect of B and C.

The regression equation would necessarily take the following form:

$$Y = \beta 0 + \beta 1X_B + \beta 2X_C + \beta 3X_{BC}$$

$X_B$, $X_C$ and $X_{BC}$ represent the factors B , C and interaction respectively.

The regression coefficients are estimated as one half of the corresponding effect estimates, with ***β0*** as the overall mean of all responses. (from last class: 44.4375)

| | Contrast | Effect | regression coefficient |
|---|---|---|---|
| B | -239 | -29.875 | -14.9375 |
| C | 125 | 15.625 | 7.8125 |
| BC | -145 | -18.125 | -9.0625 |

$$Y = 44.4375 - 14.9375X_B + 7.8125X_C - 9.0625X_{BC}$$

**<u>Interactions plot</u>**

Interactions plot creates a single interaction plot if two factors are entered, or a matrix of interaction plots if 3 to 9 factors are entered.

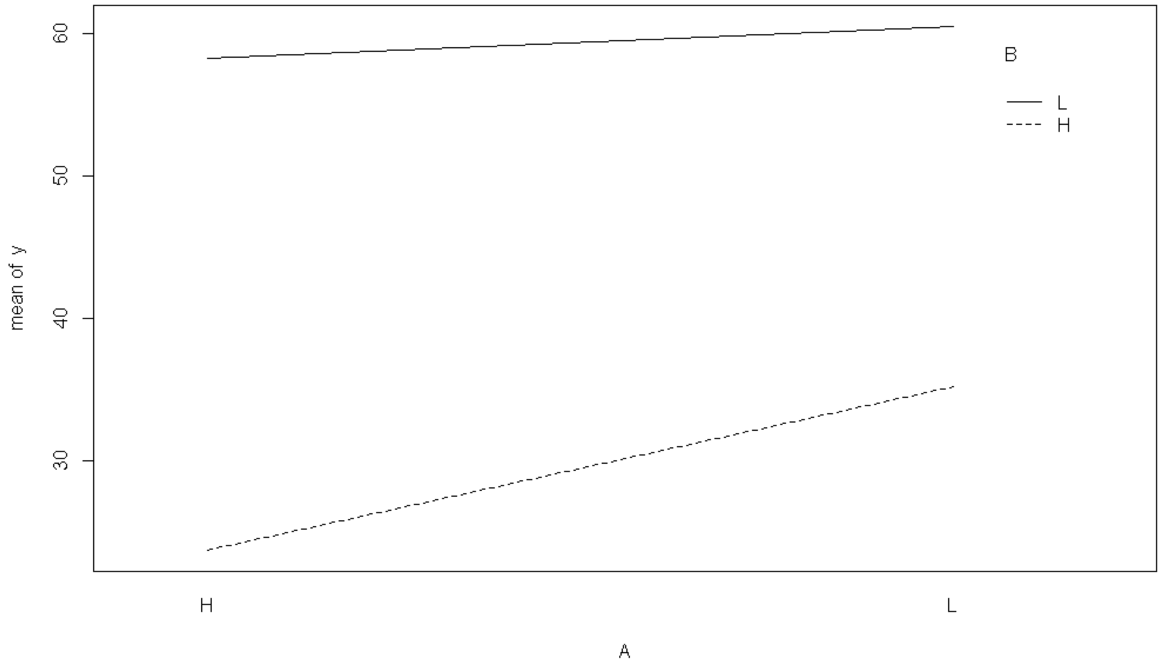
An interactions plot is a plot of means for each level of a factor with the level of a second factor held constant.

Interactions plots are useful for judging the presence of interaction, which means that the difference in the response at two levels of one factor depends upon the level of another factor. Parallel lines in an interactions plot indicate no interaction. The greater the departure of the lines from being parallel, the higher the degree of interaction. To use an interactions plot, data must be available from all combinations of levels

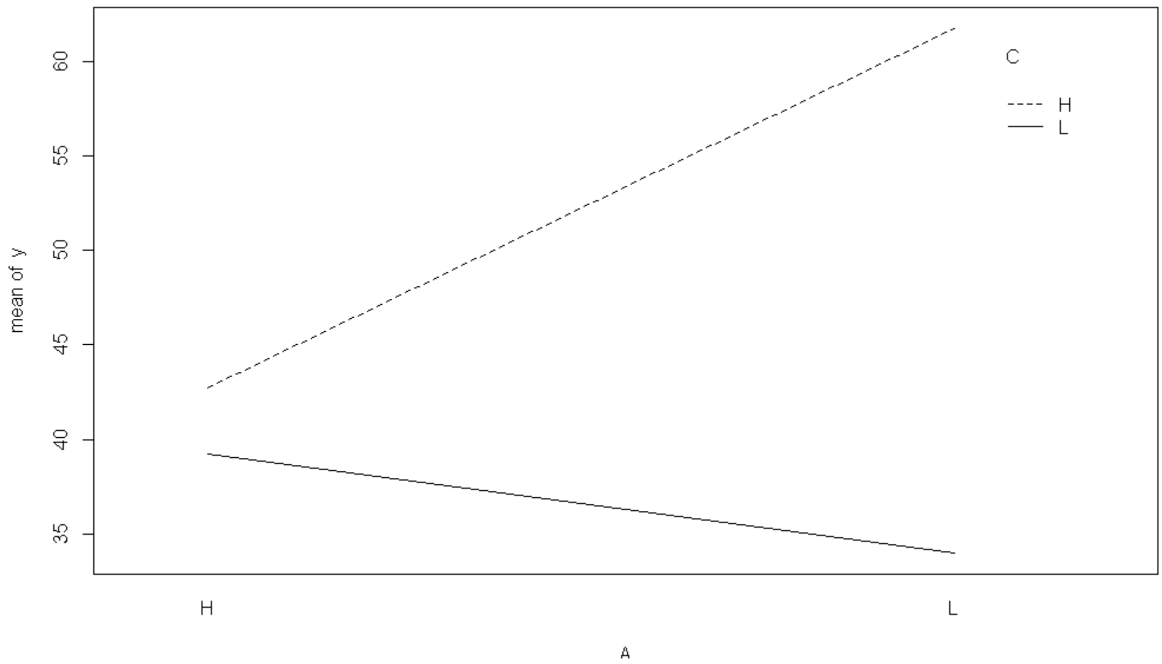
If there is crossing of lines in an interaction plot – there is strong evidence of interaction.

As with other procedures, when interpreting the plots – make sure there is a correspondence to the formal hypothesis test in the ANOVA table.

**AB interaction**




**AC interaction**



## **BC interaction**