STAT 213: Two-Way ANOVA With Interaction

Colin Reimer Dawson

Last Revised April 19, 2016

In this worksheet you will work through a two-way ANOVA analysis by hand, to help you get a feel for what is going on. Then, you will do the same thing in R, to see how to do it, and also to check your hand calculations.

Consider the following small dataset, recording Y: the amount of force in Newtons (the response) that it takes to separate two pieces of plastic that have been glued together, with two categorical predictors: X_1: thickness of the material (thin, moderate, thick), and X_2: the type of glue used (wood vs. plastic). There are two cases at each combination of factors. The data is presented in a table form, rather than the usual one-row-per-case form.

Thickness/Glue	Plastic	Wood	**Mean**
Thin	$52, 64$	$72, 60$	62
Moderate	$67, 55$	$78, 68$	67
Thick	$86, 72$	$43, 51$	63
Mean	66	62	64

The two-way ANOVA model with interaction has the following form

$$Y = \mu + \alpha_j + \beta_k + \gamma_{j,k} + \varepsilon$$

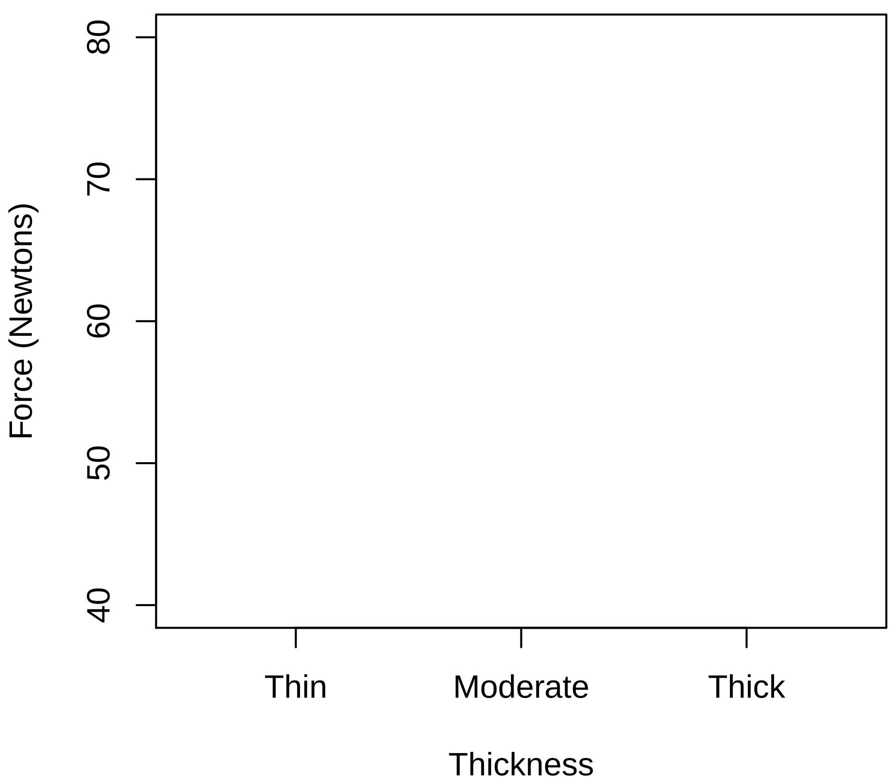
where j ranges over levels of the first factor (thickness) and k ranges over levels of the second factor (glue type), alphas reprent the deviations of the row means from the overall mean, betas represent the deviations of the column means from the overall mean, and gammas represent the deviation of the cell means from what would be predicted by the main effects model.

1. Estimate the various terms in the model:

$$
\begin{aligned}
\hat{\alpha}_{Thin} &= & \qquad\qquad \hat{\gamma}_{Thin,Plastic} &= \\
\hat{\alpha}_{Moderate} &= & \hat{\gamma}_{Thin,Wood} &= \\
\hat{\alpha}_{Thick} &= & \hat{\gamma}_{Moderate,Plastic} &= \\
&= & \hat{\gamma}_{Moderate,Wood} &= \\
\hat{\beta}_{Plastic} &= & \hat{\gamma}_{Thick,Plastic} &= \\
\hat{\beta}_{Wood} &= & \hat{\gamma}_{Thick,Wood} &=
\end{aligned}
$$

   Hint: to find the γs, first calculate the expected cell mean if only μ, αs and βs were in play, and compare to the observed cell mean.

2. Plot the cell means below. The x axis ranges over thicknesses; use a separate line for each glue type. Does there appear to be an interaction?



3. Calculate the sums of squares associated with each factor and with the interaction, using the following (note that here, $n_{jk} = 2$ for all j and k combinations):

$$SS_A = \sum_j \sum_k \sum_{i=1}^{n_{jk}} \hat{\alpha}_j^2 =$$

$$SS_B = \sum_j \sum_k \sum_{i=1}^{n_{jk}} \hat{\beta}_k^2 =$$

$$SS_{AB} = \sum_j \sum_k \sum_{i=1}^{n_{jk}} \hat{\gamma}_{jk}^2 =$$

$$SS_{Error} = SS_{Total} - SS_A - SS_B - SS_{AB} =$$

$$SS_{Total} = \sum_j \sum_k \sum_{i=1}^{n_{jk}} (y_{jki} - \bar{\bar{y}})^2 = 1573$$

Note that you can simplify the calculations a lot by using the fact that there are a lot of repeated values in the sums.

4. Fill in the ANOVA table below (use R for the P-values)

Source	df	SS	MS	F	P
Thickness					
Glue					
Interaction					
Residuals					
Total					

```r
## R code for computing P-value given an F-statistic
## Fill in numbers for F and the two df values
pf(F, DF1, DF2, lower.tail = FALSE)
```

5. Interpret the results.

6. Now let's see how to do all of that using R. The data is at

   $$\texttt{http://colinreimerdawson.com/data/glue.csv}.$$

   Read it in, and fit the two-way interaction model (use `aov()` in place of `lm()`).

7. Type the following to get the α, β and γ estimates (replace `MODEL` with the name of your model). Note that the levels of each factor will be sorted in alphabetical order; not the order they appear in the data. Do the results match what you computed by hand?

   ```
   model.tables(MODEL, type = "effects")
   ```

8. Use `plotModel()` (after loading the `mosaic` package) to get a plot of the means overlaid on the data. Does it match what you plotted by hand?

9. Use `summary()` to get the ANOVA table. Does it match what you computed by hand?