Lecture 15

Two-Way ANOVA and Multiple Measures

When we analyze an ANOVA problem, we are trying to understand how the different treatments change relative to the overall mean. We have two ways of thinking about this change: a fixed effect or a random effect. We will address this more next semester in the context of ANOVA as a linear model, but for now we will address the basic idea.

A fixed effect model says that an observation $X_{ij} = \mu + \alpha_i + \epsilon_{ij}$, where μ is the overall mean, α_i is the effect of treatment i, and ϵ_{ij} is the random error on this measurement. In this type of model, all the variability is contained in the error term.

A random effects model has the same basic structure $X_{ij} = \mu + A_i + \epsilon_{ij}$ but the capitalization on the treatment term suggests that it is also random, so there are two components of the model that have some random effect and what we measure is an estimate of the mean effect.

The difference between these two models is one of interpretation. It does not affect the computation at all. As we add another treatment variable, our model will expand, and there will be several possible ways of interpreting the model: we could have both effects fixed, both random, or one of each. We can also consider interaction terms where the two treatments enhance or suppress the effect of the treatments.

As we begin to introduce the computational side of two-way ANOVA, we are going to begin with the case where there is one observation for each combination of treatment levels. Then we'll extend the analysis to multiple measurements for each treatment level.

Our two-way model (without interactions) looks like this:

$$X_{ij} = \alpha_i + \beta_j + \epsilon_{ij}$$

Where $\mu_{ij} = \alpha_i + \beta_j$. This is a simple additive model and is consistent with our one-way ANOVA in that we can move any constant component of α and β (shift the means of these components to be 0), and express the model similarly to the way we did before as $X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$. The errors here are referred to as residuals which also have a mean of 0, but with a standard deviation of σ.

Our best estimates for these values is

$$\hat{\mu} = \bar{X}_{..}, \hat{\alpha}_i = \bar{X}_{i.} - \bar{X}_{..}, \hat{\beta}_j = \bar{X}_{.j} - \bar{X}_{..}$$

Recall that $\bar{X}_{..}$ is the grand mean, and the others with subscripts are the means over each treatment.

The calculations for the ANOVA model are similar to the one-way ANOVA, but we have an extra sum of squares (since we'll have one for each treatment variable), and when we calculate our ANOVA table, we'll end up with two F-statistics, which will analyze the effects of each treatment variable separately. So, we could get a case where neither variable is significant, both are significant, or where only one is significant. Our hypotheses will also be considered separately, with one set of null and alternative hypotheses for each treatment variable, which will be the same as for our one-way ANOVA: the means of each treatment level are the same, or at least one mean of one treatment level is different.

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{..})^2 \qquad df = IJ - 1$$

$$SSA = \sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{X}_{i.} - \bar{X}_{..})^2 = J\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2 \qquad df = I - 1$$

$$SSB = \sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{X}_{.j} - \bar{X}_{..})^2 = I\sum_{j=1}^{J}(\bar{X}_{.j} - \bar{X}_{..})^2 \qquad df = J - 1$$

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2 \qquad df = (I - 1)(J - 1)$$

The fundamental identity is

$$SST = SSA + SSB + SSE$$

Hypotheses	Test Statistic Value
H_{0A} versus H_{aA}	$f_A = \frac{MSA}{MSE}$
H_{0B} versus H_{aB}	$f_B = \frac{MSB}{MSE}$

Our ANOVA table will then have an extra row (one for each treatment variable).

SOURCE OF VARIATION	DF	SUM OF SQUARES	MEAN SQUARE	F
FACTOR A	$I - 1$	SSA	MSA	F_A
FACTOR B	$J - 1$	SSB	MSB	F_B
ERROR	$(I - 1)(J - 1)$	SSE	MSE	
TOTAL	$IJ - 1$	SST		

Recall that we are making assumptions about our data. We have assumed, among other things that, the errors are normally distributed with a mean of zero. We can calculate these from the fitted values of our model (this difference between our model and the observations is the error called a residual). We can test the assumptions of our model by constructing a normal probability plot, and a residual plot. A plot of the residuals vs. the fitted values should appear random. (We'll have more to say about residual plots when we look at regression in the next semester.)

When we have more than one observation per combination of treatments and levels, we can introduce the possibility of interaction effects. This introduces a third term into our model γ_{ij}.

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

When we build our ANOVA analysis, we will have an extra term for the interaction effect, and therefore an extra line in our ANOVA table, and an extra hypothesis to test for that interaction.

Procedurally, the interaction hypothesis is usually tested first. It's possible to end up with an interaction term but no main effect term, but this can sometimes be difficult to interpret.

Since we have more than one term per combinations of treatment, we'll need a third subscript to handle all the observations.

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

As before, the errors are still normally distributed and independent with a mean of zero and common standard deviation.

Our set of formulas then becomes

$$\text{SST} = \sum_i \sum_j \sum_k (X_{ijk} - \overline{X}_{...})^2 \qquad \text{df} = IJK - 1$$
$$\text{SSE} = \sum_i \sum_j \sum_k (X_{ijk} - \overline{X}_{ij.})^2 \qquad \text{df} = IJ(K - 1)$$
$$\text{SSA} = \sum_i \sum_j \sum_k (\overline{X}_{i..} - \overline{X}_{...})^2 \qquad \text{df} = I - 1$$
$$\text{SSB} = \sum_i \sum_j \sum_k (\overline{X}_{.j.} - \overline{X}_{...})^2 \qquad \text{df} = J - 1$$
$$\text{SSAB} = \sum_i \sum_j \sum_k (\overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...})^2 \qquad \text{df} = (I - 1)(J - 1)$$

With the SSAB term being the interaction effect sum of squares. And $SSA + SSB + SSAB + SSE = SST$

Hypotheses			Test Statistic Value
H_{0A}	versus	H_{aA}	$f_A = \dfrac{\text{MSA}}{\text{MSE}}$
H_{0B}	versus	H_{aB}	$f_B = \dfrac{\text{MSB}}{\text{MSE}}$
H_{0AB}	versus	H_{aAB}	$f_{AB} = \dfrac{\text{MSAB}}{\text{MSE}}$

Our ANOVA table further expands.

SOURCE OF VARIATION	DF	SUM OF SQUARES	MEAN SQUARE	F
FACTOR A	$I - 1$	SSA	MSA	F_A
FACTOR B	$J - 1$	SSB	MSB	F_B
FACTOR AB	$(I - 1)(J - 1)$	$SSAB$	$MSAB$	F_{AB}
ERROR	$IJ(K - 1)$	SSE	MSE	
TOTAL	$IJK - 1$	SST		

Our model can be fully a fixed effects model, a fully random effects model, or a mixed effects model. This doesn't change the computation of our models, only the interpretation.

We saw with one-way ANOVA, that if we determine that the null hypothesis for a treatment can be rejected, then we want to know which levels are different from which other levels for each treatment. We can extend Tukey's method to the two-way ANOVA. It will have different studentized range parameters for each confidence interval we construct, and as we saw with one-way ANOVA, we can graph the intervals in R rather than computing them by hand.

The general formulas are

$$w = Q \cdot (\text{estimated standard deviation of the sample means being compared})$$

$$= \begin{cases} Q_{\alpha,I,(I-1)(J-1)} \cdot \sqrt{\text{MSE}/J} & \text{for factor } A \text{ comparisons} \\ Q_{\alpha,J,(I-1)(J-1)} \cdot \sqrt{\text{MSE}/I} & \text{for factor } B \text{ comparisons} \end{cases}$$

These are the formulas for the single observation case, but these are readily extended to multiple observations. Tukey's method should only be applied when the null hypothesis is rejected.

We can look at the mtcars data in R and look at a two-way ANOVA analysis of mpg on both cylinders and gears.

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
cyl          2  824.8   412.4   38.00 1.41e-08 ***
gear         2    8.3     4.1    0.38    0.687    
Residuals   27  293.0    10.9                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

In this case, we can see that the P-value for cylinders is less than a significance level of 0.05, but the one for gear is 0.687, which is much too high. We can then say that we don't have enough evidence to think the number of gears affects miles per gallon, but we do have enough evidence to say that the number of cylinders in the engine does affect it. We plotted the graph of intervals from Tukey's method in the last lecture.

We can also have R check for interaction.

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
cyl          2  824.8   412.4  36.777 4.92e-08 ***
gear         2    8.3     4.1   0.368    0.696    
cyl:gear     3   23.9     8.0   0.710    0.555    
Residuals   24  269.1    11.2                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

This turns out not to be one of the weird cases where interaction survives but a main term does not. Here, we can't reject the null on the interaction and so our model does not change.

In the next lecture, we'll extend our ANOVA analysis to three factors. Conceivably, we can extend the model to N factors, but then the number of possible interactions can become quite large. The possible combinations of treatments and levels can easily explode, so we'll also discuss experimental design methods that will allow us to cover enough combinations of factors and levels to be able to perform an analysis. Three variables will be enough to get a look at how we might generalize our model, but we'll save the fully general model for factorial designs until next semester.

References:

1. https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineering_and_the_sciences.pdf
2. https://assets.openstax.org/oscms-prodcms/media/documents/IntroductoryStatistics-OP_i6tAI7e.pdf
3. http://www.statdistributions.com/f/
4. https://www.scribbr.com/statistics/anova-in-r/
5. https://r-graph-gallery.com/84-tukey-test.html