Chapter 7: One-factor Multi-sample Experiments

Introductory Statistics for Engineering Experimentation

Peter R. Nelson, Marie Coffin and Karen A.F. Copeland

Slides by Douglas Bates

Outline

Outline

Outline

Factors with more than two levels

- In the previous chapter we discussed inference for comparative experiments on two groups.
- In chapter 3 we showed how to model a continuous response as a function of one or more factors that could have multiple levels. We used the *R* function `aov` to estimate the cell means or the "effects" of the levels of each factor.
- The statistical assessment of whether or not the effects are significant is usually based on an *analysis of variance*; hence the name `aov` for the model-fitting function and the name `anova` for the extractor function that produces the analysis of variance table.
- The text book emphasizes a technique called the analysis of means (ANOM). This is not a widely-used technique.
- We will focus on the analysis of variance and another approach called `multiple comparisons` for follow-up analysis.

Overview of techniques

- In chapter 2 we used graphical methods, such as comparative dotplots and comparative density plots, to display a continuous response as it depends on levels of a factor. In section 4.4 we also discussed normal probability plots, which can be used as comparative plots for such data.
- In chapter 3 we discussed fitting models of the form

$$\mathcal{Y}_{ij} = \mu_i + \epsilon_{ij}, \quad i = 1, \ldots, I;\ j = 1, \ldots, n_i$$

or

$$\mathcal{Y}_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, \ldots, I;\ j = 1, \ldots, n_i$$

- These are the same model; just two different ways of writing it. The first is called the *cell means* form and the second is the *effects* form.
- We check for differences in the mean response in two stages: first we check if all the means could be equal and, if we reject this hypothesis, we check for which levels of the factor produce significantly different means.

R functions used in this chapter

- Preliminary plots are obtained with `dotplot`, `bwplot`, `densityplot` and `qqmath`, all in the comparative form. Model fits assume that the variances in the groups are more-or-less equal. Hence we check the plots for equal variances as well as equal means.
- We use `aov` to fit the model, `summary` or `anova` to obtain the analysis of variance table, and `model.tables` to obtain estimates of the cell means or the effects.
- If we reject the hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$ (or, equivalently, $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$) then we use `TukeyHSD` to perform multiple comparisons using Tukey's Honest Significant Difference method.
- We assess residual plots obtained with `plot(fm, which = 1)` and `plot(fm, which = 2)`

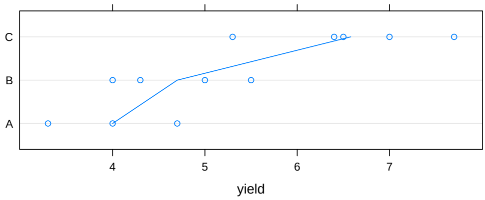
Section 7.1, Basic Inference

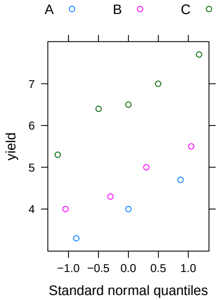
- We begin by plotting the data, preferably with comparative dotplots or comparative normal probability plots. (The text uses an error-bar chart in figure 7.1 but these are less informative than those mentioned above.)
- Group means and s^2, the mean square error, are evaluated by fitting an `aov` model and using `summary` or `anova`. The degrees of freedom for s^2, $n_1 + n_2 + \cdots + n_I - I$, is given in the table.
- You could use critical values from a T_ν distribution to calculate confidence intervals on the individual means (p. 248) but the practice is discouraged.

Examples 7.1.1 and 7.1.2

```
> reac <-
+     data.frame(yield = c(3.3, 4.0, 4.7,
+                          4.0, 5.0, 4.3, 5.5,
+                          5.3, 6.5, 6.4, 7.0, 7.7),
+                cat = factor(rep(LETTERS[1:3], c(3,4,5))))
> summary(fm1 <- aov(yield ~ cat, reac))
```

```
          Df  Sum Sq Mean Sq F value   Pr(>F)
cat        2 14.7012  7.3506  12.099 0.002813
Residuals  9  5.4680  0.6076
```





Section 7.5 Analysis of Variance

- The F ratio quoted in the analysis of variance table is the ratio of $MS_{\text{treatment}}$ to the mean square for error, MS_e. It is a "signal-to-noise" ratio based on the differences the differences between groups versus the differences within groups.
- Both the numerator and the denominator have degrees of freedom associated with them. In this case they are $I - 1$ (numerator) and $N - I$ (denominator) where N is the total number of observations $(N = n_1 + n_2 + \cdots + n_I)$
- The p-value is calculated from a theoretical distribution for this quantity, written F_{ν_1,ν_2}. The *R* functions for this distribution are `df`, `pf`, `qf` and `rf`.
- The hypothesis being tested is "are all the means the same?" versus "are there any differences?". Symbolically
  $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$, in the cell means form, or
  $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$ in the effects form.

Multiple comparisons

- If we reject H_0 in the analysis of variance the natural follow-up question is "so which group means are significantly different".
- It is tempting to use a series of t-tests to compare each pair of groups but doing so will inflate the probability of a false positive.
- There are several ways of controlling for this inflated false positive probability. We will use Tukey's Honest Significant Differences, `TukeyHSD`, which is preferred when the groups are of equal importance. (Other methods are used when we have, say, a control group that we wish to compare with each of several treatments.)

Example 7.1.1 cont'd

```
> TukeyHSD(fm1)

  Tukey multiple comparisons of means
    95% family-wise confidence level
Fit: aov(formula = yield ~ cat, data = reac)

$cat
    diff        lwr      upr     p adj
B-A 0.70 -0.9621403 2.362140 0.4954811
C-A 2.58  0.9906899 4.169310 0.0036432
C-B 1.88  0.4201254 3.339875 0.0144223
```