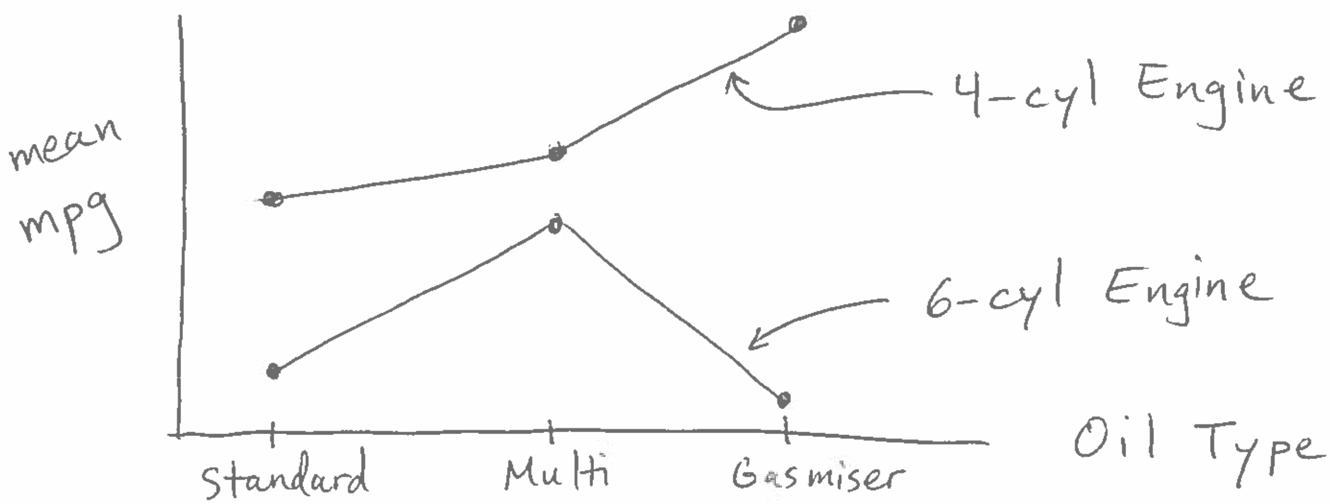
Interpreting a Significant Interaction

- **Generally done by examining Interaction Plots.** (Profile Plots)

Example (Gas mileage data): Could plot $\bar{Y}_{ij\cdot}$ against A, separately for each level of C.

- Could plot $\bar{Y}_{ij\cdot}$ against C, separately for each level of A.



Conclusions:

- 4-cylinder engines seem to get better gas mileage than 6-cylinder engines, but the effect of engine type is more pronounced for Gasmiser and Standard Oils than for Multi Oil.

- No oil type is uniformly the best — Gasmiser is best for 4-cyl engines, but Multi is best for 6-cyl engines.

Specific Comparisons

• If any of the F-tests reveal that the factor(s) have significant effects on the response, we can perform:

- **• Preplanned comparisons (contrasts)**
- **• Post-hoc multiple comparisons (Fisher LSD or Tukey)**

in order to determine which factor levels produce significantly different mean responses.

• This is straightforward when there is <u>no significant interaction</u> between factors.

• We may then treat each factor separately, and use contrasts or multiple comparisons to compare mean responses among the levels of each factor.

• Basically just like in previous chapter, except we do it for two factors separately.

Check SAS's ordering of levels from Class Level Info

Example: Suppose there were no engine × oil interaction. Let's compare cheap oil (standard) vs. expensive oils (others).

Contrast of interest: $L = \mu_{std} - \dfrac{\mu_{Gas} + \mu_{multi}}{2}$

$L = -\frac{1}{2}\mu_{gas} - \frac{1}{2}\mu_{multi} + \mu_{std}$

Test $H_0: L = 0$ vs. $H_a: L \neq 0$. $t^* = -2.78$, P-value $= .0105 < .05 = \alpha$. We reject H_0 and conclude a significant difference in mean mileage between standard ("cheap") oil and other oil types.

Not really valid in this analysis b/c we have interaction.

• **If we do have significant interaction (as we actually did in the gas mileage example), we must investigate contrasts about one factor <u>given a specific level of the other factor</u>.**

<u>Example 1</u>: Do the mean mileages of 4-cylinder and 6-cylinder engines differ significantly, when the oil type is "Gasmiser"?

$$E(Y_{4\text{-cyl},G}) = \mu + \alpha_{4\text{-cyl}} + \gamma_G + (\alpha\gamma)_{4\text{-cyl},G}$$

$$E(Y_{6\text{-cyl},G}) = \mu + \alpha_{6\text{-cyl}} + \gamma_G + (\alpha\gamma)_{6\text{-cyl},G}$$

Our contrast is $E(Y_{4\text{-cyl},G}) - E(Y_{6\text{-cyl},G})$

Relevant contrast:

$$L = \alpha_{4\text{-cyl}} - \alpha_{6\text{-cyl}} + (\alpha\gamma)_{4\text{-cyl},G} - (\alpha\gamma)_{6\text{-cyl},G}$$

We test: $H_0: L = 0$ vs. $H_a: L \neq 0$

<u>Example 2</u>: Do the mean mileages for the cheap oil ("standard") and the expensive oils differ significantly, when the engine is "4-cylinder"?

$$E(Y_{4\text{-cyl},S}) = \mu + \alpha_{4\text{-cyl}} + \gamma_S + (\alpha\gamma)_{4\text{-cyl},S}$$

$$E(Y_{4\text{-cyl},G}) = \mu + \alpha_{4\text{-cyl}} + \gamma_G + (\alpha\gamma)_{4\text{-cyl},G}$$

$$E(Y_{4\text{-cyl},M}) = \mu + \alpha_{4\text{-cyl}} + \gamma_M + (\alpha\gamma)_{4\text{-cyl},M}$$

Relevant contrast:

$$L = \gamma_S - \tfrac{1}{2}\gamma_G - \tfrac{1}{2}\gamma_M + (\alpha\gamma)_{4\text{-cyl},S} - \tfrac{1}{2}(\alpha\gamma)_{4\text{-cyl},G} - \tfrac{1}{2}(\alpha\gamma)_{4\text{-cyl},M}$$

We test:

$H_0: L = 0$ vs. $H_a: L \neq 0$

Conclusions based on computer output:

Example 1: $t^* = 6.74$, P-value < .0001. Reject H_0, conclude the 4-cylinder and 6-cylinder engines have different mean mileages when Oil Type is "Gasmiser".

Example 2: : $t^* = -2.54$, P-value = .018. Reject H_0: At $\alpha = .05$, conclude the standard oil has different mean mileage than other oils, when engine is 4-cylinder.

Post-Hoc Comparisons

- **If there is significant interaction, we test for significant differences in mean response for <u>each pair</u> of <u>factor level combinations</u>.**

We test: $H_0: E(Y_{i'j'k'}) = E(Y_{i''j''k''})$ for each $i' \neq i''$ or $j' \neq j''$

a series of null hypotheses

Could write: $H_0: \mu_{i'j'} = \mu_{i''j''}$ for <u>all</u> $i' \neq i''$ or $j' \neq j''$

- **Again, Fisher LSD procedure has P{Type I error} = α <u>for each comparison</u>.**

- **Tukey procedure has P{at least one Type I error} = α <u>for the entire set of comparisons</u>.**

- **For Tukey procedure, we conclude a difference in mean response is significant, at level α, if:**

$$\left|\bar{Y}_{i'j'\cdot} - \bar{Y}_{i''j''\cdot}\right| > q_\alpha(t, df)\sqrt{\frac{MSW}{n}} \quad \text{(balanced data)}$$

(for $i' \neq i'', j' \neq j''$) where $q_\alpha(t, df)$ given in Table A.7.

Here t = # of factor-level combinations (ac).

and df = # of within-cell d.f.

$t = (2)(3) = 6$

Example (Gas mileage data): error(within) $df = ac(n-1) = 24$

$n = 5$

Using $\alpha = .05$:

$q_{.05}(6, 24) = 4.37$ (Table A.7)

So $4.37\sqrt{\frac{1.084}{5}} = 2.035$. An example:

Difference in mean mileage between (4-cyl, multi) and (6-cyl, standard):

$|\bar{Y}_{4\text{-cyl, multi}} - \bar{Y}_{6\text{-cyl, standard}}| = |24.08 - 21.72| = 2.36 > 2.035$,

so Tukey procedure judges these population means to be different.

- Tukey procedure designed to compare all such pairs of cell population means (see SAS code/output)

Additional Considerations

- **What if we have no replication (i.e., $n = 1 \rightarrow$ one observation for each cell)?**

- **We then have no estimate of σ^2 (the variation among responses in the same cell).**

- **Solution: Assume there is no interaction. The interaction MS will then serve as an estimate of σ^2.**

- **If we do this, and interaction <u>does exist</u>, then our F-tests will be biased (conservative $\rightarrow$ less likely to reject H_0).**

Three or More Factors

• If we have three or more factors, we have the possibility of higher-order interactions.

Example: Factors A, B, and C:

3 sets of main effects (for A, B, C)

3 two-factor interactions ($A \times B$, $A \times C$, $B \times C$)

1 three-factor interaction ($A \times B \times C$)

• If the 3-way interaction is significant, this implies, for example, that the $A \times B$ interaction is not consistent across the levels of C.

• Having 3 or more factors means having lots of "cells".

• If resources are limited, the number of replicates could be small ($n = 1$? $n = 2$?)

• It may be better to assume higher-order interactions do not exist (often they are of no practical interest anyway).

• Thus we could devote more degrees of freedom to estimating σ^2.

• Analysis of three-factor studies can be done with software in a similar way.

Example: (Table 9.27 data, p. 515)

Response: Rice yield

Factors: Location (4 levels)
Variety (3 levels)
Nitrogen (4 levels)

- We have $n = 1$ observation for each factor level combination.

Analysis: When we included the 3-way interaction, we had no estimate of σ^2 (no MSW) and we could not do F-tests.

- Solution: Leave off 3-way interaction.

New analysis: Found significant Location × Variety interaction. No significant interaction involving Nitrogen.

- Main effects F-test about Nitrogen was significant → this shows the mean yield differs at the different levels of Nitrogen.
- Further analysis on Nitrogen factor: Nitrogen levels 60 and 150 have significantly different mean yields (150 level appears to have a higher mean).
- All other comparisons between pairs of nitrogen levels are not significant [Tukey procedure]