The Overall Two-Factor Analysis

Chapter 11, Sections 4-7

ERSH 8310

Today's Class

- More two-factor ANOVA:
  - Linear statistical model.
  - Blocking factors.
  - Measuring effect size.
  - Determining sample size.

THE STATISTICAL MODEL

The Linear Model

- The model underlying the present analysis is:

$$Y_{ijk} = \mu_T + \alpha_j + \beta_k + (\alpha\beta)_{jk} + E_{ijk.}$$

- Where:
  - μ_T is the overall mean of the population.
  - $\alpha_j = \mu_j - \mu_T$ is the average treatment effect at level $a_{j.}$
  - $\beta_k = \mu_k - \mu_T$ is the average treatment effect at level $b_{k.}$
  - $(\alpha\beta)_{jk} = \mu_{jk} - \mu_j - \mu_k + \mu_T$ is the interaction effect at cell $a_j b_{k.}$
  - $E_{ijk} = Y_{ijk} - \mu_{jk}$ is the experimental error associated with each score.

Null Hypotheses

- The null hypothesis for the A main effect:

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$$

- The null hypothesis B main effect:

$$H_0: \beta_1 = \beta_2 = \ldots = \beta_b = 0$$

- The null hypothesis for the A ×B interaction:

$$H_0: (\alpha\beta)_{11} = (\alpha\beta)_{12} = \ldots = (\alpha\beta)_{ab} = 0$$

Expected Mean Squares

- The $E(MS_{S/AB}) = \sigma_{error}{}^2$ and other expected values of the mean squares are:

$$E(MS_A) = \sigma_{error}{}^2 + bn(\theta_A{}^2),$$

$$E(MS_B) = \sigma_{error}{}^2 + an(\theta_B{}^2),$$

$$E(MS_{A\times B}) = \sigma_{error}{}^2 + n(\theta_{A\times B}{}^2),$$

Fixed-Effects Model Expected Mean Squares

$$\theta_A^2 = \frac{\sum_{j=1}^{a} \alpha_j^2}{a - 1},$$

$$\theta_B^2 = \frac{\sum_{k=1}^{b} \beta_k^2}{b - 1},$$

$$\theta_{A \times B}^2 = \frac{\sum_{j=1}^{a} \sum_{k=1}^{b} (\alpha\beta)_{jk}^2}{(a - 1)(b - 1)}.$$

E(MS) Notes

- Note that E(MS)'s are the average values of the respective mean squares obtained with repeated sampling from a given set of population.

- Under the assumption that the null hypotheses are true, no effects are present.
  - Consequently, all θ^2 terms will be deleted from the expected mean squares.

- The ratios of the mean squares of the effect and the error will follow the respective F distributions with a set of two degrees of freedom (numerator and denominator).

The Linear Model in a 2 ×2 Table

- In a 2 ×2 design, the linear model is particularly simple because all the tests have one degree of freedom so are equivalent to contrasts on the means.

Violating the Assumptions

- The independence assumption is fundamental to the analysis.
- Normality and variance homogeneity concern with the distribution of scores within the treatment populations.
- For the test of normality, we may use Kolmogorov-Smirnov type tests (Conover, 1999).
- We can test heterogeneity with either the Levene test or the Brown-Forsythe test.
- Alternatively, we may use the F_{max} procedure.
  - If $F_{max} > 3$, then we may employ a more stringent significance level.

DESIGNS WITH A BLOCKING FACTOR

The Randomized-Blocks Design

- One may introduce a blocking factor to capture variability that is irrelevant to the effect of interest and thereby reduce the size of error term.
  - Perhaps to make groups of subjects more homogeneous within a block.
- In a randomized-blocks design, the blocking is part of the original study and controls the assignment of subjects to groups.

RBD Example

- Suppose that a researcher is investigating the effects of four sets of instructional material on how well college students learn a body of quantitative material - for example, say statistics.
- The simplest procedure is to obtain a sample of 60 students, say, and to randomly assign $n=15$ subjects to each of four groups to create a completely randomized single-factor experiment.
- Suppose the researcher realizes there is great variability in the student performance arising from differences in their quantitative skills before the study started.

More of the RBD Example

- The variability makes the $MS_{S/A}$ large and limits the power of the design to detect differences among the instruction conditions.
  - To increase power one can increase the sample size – but here cannot do that.
  - Another way to increase power is to decrease the variability of the scores.
- Creating a blocking factor will aid in decreasing the variability of the scores.

Table 11.10

Table 11.10: A comparison of a completely randomized single-factor design and a randomized-blocks design.

Completely randomized (unblocked) design

Instructions (Factor A)

a_1	a_2	a_3	a_4
$n = 15$	$n = 15$	$n = 15$	$n = 15$

Source	df
A	$a - 1 = 3$
S/A	$a(n-1) = 56$
Total	$an - 1 = 59$

Randomized-blocks design

Instructions (Factor A)

Blocks	a_1	a_2	a_3	a_4
b_1	$n = 5$	$n = 5$	$n = 5$	$n = 5$
b_2	$n = 5$	$n = 5$	$n = 5$	$n = 5$
b_3	$n = 5$	$n = 5$	$n = 5$	$n = 5$

Source	df
A	$a - 1 = 3$
B	$b - 1 = 2$
$A \times B$	$(a-1)(b-1) = 6$
S/AB	$ab(n-1) = 48$
Total	$abn - 1 = 59$

Post-Hoc Blocking

- In a post-hoc design, the second factor is created after the data are collected.
- Use the analysis of covariance (see Chapter 15) when the potential blocking information is available as a numerical quantity.

MEASURING EFFECT SIZE

Complete Omega Squared

- A general formula for the complete omega squared is:

$$\omega_{\text{effect}}^2 = \frac{\sigma_{\text{effect}}^2}{\sigma_{\text{total}}^2}, \qquad (26)$$

- Where:

$$\sigma_{\text{total}}^2 = \sigma_A^2 + \sigma_B^2 + \sigma_{A \times B}^2 + \sigma_{S/AB}^2. \qquad (27)$$

- Recall, σ^2 comes from the MS for each term.
- We use also the F values to obtain the omega squared (see Equation 11.20).

Obtaining Effect Size Estimates

- We may obtain estimates of the above terms using:

$$\hat{\omega}^2_{\text{effect}} = \frac{\hat{\sigma}^2_{\text{effect}}}{\hat{\sigma}^2_{\text{total}}},$$

- Where:

$$\hat{\sigma}^2_{S/AB} = MS_{S/AB}.$$

More Formulae

$$\hat{\sigma}^2_{total} = \hat{\sigma}^2_{A} + \hat{\sigma}^2_{B} + \hat{\sigma}^2_{A\times B} + \hat{\sigma}^2_{S/AB} ,$$

$$\hat{\sigma}^2_{A} = \frac{df_A(MS_A - MS_{S/AB})}{abn} , \qquad (32)$$

$$\hat{\sigma}^2_{B} = \frac{df_B(MS_B - MS_{S/AB})}{abn} , \qquad (33)$$

$$\hat{\sigma}^2_{A\times B} = \frac{df_{A\times B}(MS_{A\times B} - MS_{S/AB})}{abn} , \qquad (34)$$

Partial Omega Squared

- The partial omega squared is defined as:

$$\omega_{effect}^2 = \frac{\sigma_{effect}^2}{\sigma_{effect}^2 + \sigma_{error}^2}.$$

- The estimate is given by

$$\hat{\omega}_{effect}^2 = \frac{\hat{\sigma}_{effect}^2}{\hat{\sigma}_{effect}^2 + \hat{\sigma}_{error}^2}.$$

- The partial omega squared relates the treatment component to the sum of only two (i.e., treatment and error) components.

Descriptive Effect Measures

- The squared correlation ratio R^2_{effect} can be obtained using:

$$R^2_{effect} = \frac{SS_{effect}}{SS_{total}}$$

$$R^2_{(effect)} = \frac{SS_{effect}}{SS_{effect} + SS_{S/AB}} .$$

More Descriptive Measures

- Another descriptive measure is the standard difference between means.
- For example,

$$d_{a1b1,a2b1} = \frac{\bar{Y}_{11} - \bar{Y}_{21}}{(SS_{11} + SS_{21})/(df_{11} + df_{21})}.$$

DETERMINING SAMPLE SIZE

Determining Sample Size from Population Effects

- Using the Sample Size Table with a Main Effect: Use Table 8.12.
  - Note that the effect with a partial measure (e.g., $\omega_{(A)}^2$) should be employed.
  - The sample size can be found for all the subjects at a level of whichever factor you use, not for those in one cell (e.g., b_n instead of n).

Power Charts

- Using the Power Charts:
  - If you need values of power or effect size other than those in Table 8.1, you can use the power charts in Appendix A.7. The total sample size N (e.g., abn) is

$$N = \frac{\phi^2 \, (df_{effect} + 1) \, \sigma_{error}^2}{\sigma_{effect}^2},$$

- where ϕ is the noncentrality parameter from Appendix A.7.

Using Estimated Treatment Effects

- The easiest way to transfer information from one study to another is to use the estimates of partial omega squared $[\hat{}(\omega)]_{(\text{effect})}^2$.
- We can use Table 8.1 if we are only interested in the main effects.
- We can use Equation 11.25 and the power charts if we need a value other than those in the table.

Estimating Power

- The power of an existing experiment can be obtained using:

$$\phi_{effect} = \sqrt{\frac{N\,\sigma_{effect}^2}{(df_{effect} + 1)\,\sigma_{effect}^2}}$$

and the power chart of Appendix A.7.

Final Thought

- Today we took more concepts we learned in the first 8 chapters and applied them to the two-way ANOVA.

- The process of using analysis of variance in multi-factor studies follows the process of using a one-way ANOVA.
- We will become more familiar with such concepts as we continue through the book.

Next Class

- Midterm handed out.
- Midterm discussion.