Another Extra Topic! – Two-Factor ANOVA

Basics

What is it?

Essentially, Two-Factor ANOVA is multiple regression with two categorical explanatory variables (or factors). We use this method for initial model assessment, testing for interaction and/or individual factor effects. After an appropriate (general) model is found, we can answer specific questions about factor effects using a multiple regression approach.

Variables and Notation

Y: the response
A: factor 1 with 'a' treatment levels
B: factor 2 with 'b' treatment levels

Data: We will only consider a balanced design where we have the same number of observations (n') in each of the ab treatment combinations. ($n = abn'$)

Assumptions

We assume that, for each of the ab treatment combinations, the responses are independent and normally distributed with constant variability.

Special Models Defined – Additive vs. Non-Additive

1. *The Non-Additive (Saturated) Model*: In the non-additive model, the effects of one factor are possibly different at all levels of the other factor → there is interaction. Essentially, the separate lines model in multiple regression.

$$\mu(Y \mid A, B) = \beta_0 + A + B + AB$$

2. *The Additive Model*: In an additive model, the effects of one factor are all the same at all levels of the other factor → there is NO interaction. Essentially, the parallel lines model in multiple regression.

$$\mu(Y \mid A, B) = \beta_0 + A + B$$

Aside: Writing Models: Short Form vs. Long Form

When writing the model in regression form, the notation can get slightly bulky with a large number of indicator variables and/or interaction terms. An abbreviation for specifying a categorical explanatory variable is to list the categorical name in uppercase letters to represent the entire set of indicator variables used to model it.

Other Models (in Short Form)

3. $\mu(Y \mid A, B) = \beta_0 + A + AB$ (one factor and interaction)
4. $\mu(Y \mid A, B) = \beta_0 + B + AB$ (one factor and interaction)
5. $\mu(Y \mid A, B) = \mu(Y \mid A) = \beta_0 + A$ (one-way ANOVA for factor 1)

6. $\mu(Y \mid A, B) = \mu(Y \mid B) = \beta_0 + B$ (one-way ANOVA for factor 2)
7. $\mu(Y \mid A, B) = \beta_0 = \mu_Y$ (one-mean model)

Note: Models 3 and 4 rarely make sense. For example, model 3 states that factor 2 has no main effect while simultaneously saying that the factor 2 effects change for different levels of factor 1.

Method of Analysis

i. *Preliminary Analysis*: Graphical analysis to check assumptions. If serious violations exist, consider a transformation.
ii. *Two-Way ANOVA*: Assess additivity. An additive model (no interaction) is preferred to facilitate answering questions about the main factor effects. If the additive model is appropriate, carry out tests for main factor effects.
iii. *Multiple Regression Analysis*: When a sufficient model is found, we can use a multiple regression approach to answer specific questions about model effects.

***F*-tests in Non-Additive Table**

F-test #1: An overall *F*-test for any differences in the means of the *ab* group combinations of factor *A* and factor *B*:

H_0: $\mu(Y \mid A, B) = \beta_0 = \mu_Y$ (one-mean model)
H_a: $\mu(Y \mid A, B) = \beta_0 + A + B + AB$ (non-additive model)
$F_{1,0} \sim F_{ab\text{-}1,\, ab(n' - 1)}$

F-test #4: An *F*-test for interaction. This is the *F*-test for additivity.

H_0: $\mu(Y \mid A, B) = \beta_0 + A + B$ (additive model)
H_a: $\mu(Y \mid A, B) = \beta_0 + A + B + AB$ (non-additive model)
$F_{4,0} \sim F_{(a-1)(b-1),\, ab(n'-1)}$

We will generally test for interaction first. If it is found that there is little evidence of interaction, we will fit the additive model, and then test for the main factor effects.

***F*-tests in Additive Table**

F-test #1: An overall *F*-test for any differences in the means of the *ab* group combinations of factor *A* and factor *B*:

H_0: $\mu(Y \mid A, B) = \beta_0 = \mu_Y$ (one-mean model)
H_a: $\mu(Y \mid A, B) = \beta_0 + A + B$ (additive model)
$F_{1,0} \sim F_{a+b-2,\, abn' - a - b + 1}$

F-test #2: An *F*-test for factor *A* effects in the presence of factor *B*.

H_0: $\mu(Y \mid A, B) = \beta_0 + B$
H_a: $\mu(Y \mid A, B) = \beta_0 + A + B$ (additive model)
$F_{2,0} \sim F_{a-1,\, abn' - a - b + 1}$

F-test #3: An *F*-test for factor *B* effects in the presence of factor *A*.

H_0: $\mu(Y \mid A, B) = \beta_0 + A$
H_a: $\mu(Y \mid A, B) = \beta_0 + A + B$ (additive model)
$F_{3,0} \sim F_{b-1,\, abn' - a - b + 1}$

Two-Way ANOVA Table for Non-Additive Model (with interaction)
(Suppose A = Factor 1, B = Factor 2)

Source	SS	df	MS	F
Corrected Model	SSR(Extra)	$ab - 1$	$MS(\text{Extra}) = \dfrac{SSR(\text{Extra})}{ab-1}$	$F_{1,0} = \dfrac{MS(\text{Extra})}{MSE}$
(between)				
A	SSA	$a - 1$	$MSA = \dfrac{SSA}{a-1}$	$F_{2,0} = \dfrac{MSA}{MSE}$
B	SSB	$b - 1$	$MSB = \dfrac{SSB}{b-1}$	$F_{3,0} = \dfrac{MSB}{MSE}$
AB	$SSAB$	$(a - 1)(b - 1)$	$MSAB = \dfrac{SSAB}{(a-1)(b-1)}$	$F_{4,0} = \dfrac{MSAB}{MSE}$
Error (within) *Non-Additive*	SSE	$ab(n' - 1)$	$MSE = \dfrac{SSE}{ab(n'-1)}$	
Corrected Total	***TSS***	***abn′* – 1**		

Two-Way ANOVA Table for Additive Model
(Suppose A = Factor 1, B = Factor 2)

Source	SS	df	MS	F
Corrected Model	SSR(Extra)	$a + b - 2$	$MS(\text{Extra}) = \dfrac{SSR(\text{Extra})}{a+b-2}$	$F_{1,0} = \dfrac{MS(\text{Extra})}{MSE}$
(between)				
A	SSA	$a - 1$	$MSA = \dfrac{SSA}{a-1}$	$F_{2,0} = \dfrac{MSA}{MSE}$
B	SSB	$b - 1$	$MSB = \dfrac{SSB}{b-1}$	$F_{3,0} = \dfrac{MSB}{MSE}$
Error (within) *Additive*	SSE	$abn' - a - b + 1$	$MSE = \dfrac{SSE}{abn' - a - b + 1}$	
Corrected Total	***TSS***	***abn′* – 1**		