Two-way ANOVA without interaction

Recall: Two-way ANOVA may be used to determine whether the effects of two factors (denoted by Factor A and Factor B hereafter) on a response are significant or not, and whether there is a significant interaction effect between the two factors.

- Model: the \( m \)-th observed response when Factor A is of level \( i \) and Factor B is of level \( j \) is

\[ X_{i,j,m} = \mu_{i,j} + \varepsilon_{i,j,m} \]

for \( 1 \leq i \leq k, 1 \leq j \leq b, 1 \leq m \leq \ell \), where \( \ell \) is the number of observations when Factor A is of level \( i \) and Factor B is of level \( j \) and \( \varepsilon_{i,j,m} \)'s are IID \( N(0, \sigma^2) \). The mean \( \mu_{i,j} \) is often expressed as

\[ \mu_{i,j} = \mu + \alpha_i + \beta_j + \gamma_{i,j}, \quad (1) \]

where

\[ \sum_{i=1}^{k} \alpha_i = 0 = \sum_{j=1}^{b} \beta_j, \]

\[ \sum_{j=1}^{b} \gamma_{i,j} = 0 \text{ for each } i \text{ and } \sum_{i=1}^{k} \gamma_{i,j} = 0 \text{ for each } j. \]

- The SSA (sum of squares due to Factor A) is

\[ \sum_{i=1}^{k} \sum_{j=1}^{b} \sum_{m=1}^{\ell} (\bar{X}_i - \bar{X}_G)^2 = b\ell \sum_{i=1}^{k} (\bar{X}_i - \bar{X}_G)^2. \]

- The SSB (sum of squares due to Factor B) is

\[ \sum_{i=1}^{k} \sum_{j=1}^{b} \sum_{m=1}^{\ell} (\bar{X}_j - \bar{X}_G)^2 = k\ell \sum_{j=1}^{b} (\bar{X}_j - \bar{X}_G)^2. \]

- The SSI (sum of squares due to interaction) is

\[ \ell \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{X}_{i,j} - \bar{X}_i - \bar{X}_j + \bar{X}_G)^2 \]

- The SSE (sum of squares due to error) is

\[ \sum_{i=1}^{k} \sum_{j=1}^{b} \sum_{m=1}^{\ell} (X_{i,j,m} - \hat{X}_{i,j})^2. \]
Note.

\[ \text{SS Total} = \text{SSA} + \text{SSB} + \text{SSI} + \text{SSE} \]  \hspace{1cm} (2)

- When \( \ell = 1 \), we assume that there is no interaction effect and use SSI as the new SSE. Let \( X_{i,j} = X_{i,j,1} = \bar{X}_{i,j} \), then (2) becomes

\[ \text{SS Total} = \text{SSA} + \text{SSB} + \sum_{i=1}^{k} \sum_{j=1}^{b} (X_{i,j} - \bar{X}_{i} - \bar{X}_{j} + \bar{X}_{G})^2, \]

SSE (assuming no interaction)

where

\[ \text{SSA} = \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{X}_{i} - \bar{X}_{G})^2 = b \sum_{i=1}^{k} (\bar{X}_{i} - \bar{X}_{G})^2 \]

and

\[ \text{SSB} = \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{X}_{j} - \bar{X}_{G})^2 = k \sum_{j=1}^{b} (\bar{X}_{j} - \bar{X}_{G})^2. \]

- Note that in the two-way ANOVA without interaction model,

degrees of freedom for SSE = \( n - 1 - (k - 1) - (b - 1) = (k - 1)(b - 1) \).

- Testing problems.
  - Testing whether there is no Factor A effect (different treatment levels of Factor A have the same effect on the mean of the response).
    \[ H_0 : \alpha_1 = \cdots = \alpha_k = 0. \]  \hspace{1cm} (3)
  - Test whether there is no Factor B effect (different treatment levels of Factor B have the same effect on the mean of the response).
    \[ H_0 : \beta_1 = \cdots = \beta_b = 0 \]  \hspace{1cm} (4)

- Two-way ANOVA example.
  - Response: weight loss.
  - Factor A: receiving different drugs.
  - Factor B: doing different exercises.

- For testing (3), the test statistic is

\[ F = \frac{\text{SSA}/(k - 1)}{\text{SSE}/((k - 1)(b - 1))}. \]

\[- F \sim F(k - 1, (k - 1)(b - 1)) \text{ under } H_0. \]
- The $F$ test rejects $H_0$ at level $\alpha$ if $F > f_{\alpha,k-1,(k-1)(b-1)}$.

- For testing (4), the test statistic is
  \[ F = \frac{SSB/(b - 1)}{SSE/((k - 1)(b - 1))}. \]

- $F \sim F(b - 1, (k - 1)(b - 1))$ under $H_0$.

- The $F$ test rejects $H_0$ at level $\alpha$ if $F > f_{\alpha,b-1,(k-1)(b-1)}$.

Example 1. Suppose that we are interested in the effects of taking three weight loss drugs $D_1$, $D_2$ and $D_3$ while doing two types of exercises: jogging or walking, at the same time. 30 participants are assigned to receive one of the three drugs and required to jog or walk for 40 minutes 3 times per week. Their weight losses (in kilograms) are given below.

<table>
<thead>
<tr>
<th>Jogging group</th>
<th>Walking group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug</td>
<td>$D_1$</td>
</tr>
<tr>
<td>weight loss</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Determine whether the following conclusions can be made at the 0.05 significance level.

(a) Different drugs have different effects on weight loss.

(b) Different exercises have different effects on weight loss.

- Solution.

  - Let Factor $A$ be the drug factor and Factor $B$ be the exercise factor, and compute SSA and SSB. The grand mean is
    \[ \bar{X}_G = \frac{5.4 + 4.6 + 6.7 + 5.1 + 4.3 + 4.5}{6} = \frac{30.6}{6} = 5.1. \]

  Let $\bar{X}_i$ be the sample mean for the $D_i$ group, then
    \[ \bar{X}_1 = \frac{5.4 + 5.1}{2} = 5.25, \]
    \[ \bar{X}_2 = \frac{4.6 + 4.3}{2} = 4.45 \]

  and
    \[ \bar{X}_3 = \frac{6.7 + 4.5}{2} = 5.6. \]

  \[
  SSA = 2 \left( (\bar{X}_1 - \bar{X}_G)^2 + (\bar{X}_2 - \bar{X}_G)^2 + (\bar{X}_3 - \bar{X}_G)^2 \right) 
  = 2 \left( (5.25 - 5.1)^2 + (4.45 - 5.1)^2 + (5.6 - 5.1)^2 \right) 
  = 1.39 
  \]
Let $\bar{X}_1$ and $\bar{X}_2$ be the sample means for the jogging group and the walking group respectively, then

$$\bar{X}_1 = \frac{5.4 + 4.6 + 6.7}{3} = \frac{16.7}{3}.$$  

and

$$\bar{X}_2 = \frac{5.1 + 4.3 + 4.5}{3} = \frac{13.9}{3}.$$  

$$SSB = 3 \sum_{j=1}^{2} (\bar{X}_j - \bar{X}_G)^2$$

$$= 3 \left( \left( \frac{16.7}{3} - 5.1 \right)^2 + \left( \frac{13.9}{3} - 5.1 \right)^2 \right)$$

$$= \frac{3.92}{3}.$$  

– Compute SSE.

$$SSE = SS\text{ total} - SSA - SSB$$

$$= \left( \sum_{i=1}^{3} \sum_{j=1}^{2} \bar{X}_{i,j}^2 \right) - 3 \cdot 2(\bar{X}_G)^2 - SSA - SSB$$

$$= (5.4)^2 + (4.6)^2 + (6.7)^2 + (5.1)^2 + (4.3)^2 + (4.5)^2 - 6(5.1)^2$$

$$-1.39 - \frac{3.92}{3}$$

$$= \frac{3.61}{3}$$  

– The $F$ statistic for testing the drug effect is

$$\frac{SSA/(3 - 1)}{SSE/((3 - 1)(2 - 1))} = \frac{1.39/2}{(3.61/3)/2} = 1.155125.$$  

From the table in Appendix B.4, $f_{0.05,2,2} = 19 > 1.155125$, so we conclude that different drugs have the same effect on weight loss at the 0.05 level.

– The $F$ statistic for testing the exercise effect is

$$\frac{SSB/(2 - 1)}{SSE/((3 - 1)(2 - 1))} = \frac{3.92/3}{(3.61/3)/2} = 2.171745.$$  

From the table in Appendix B.4, $f_{0.05,1,2} = 18.51 > 2.171745$, so we conclude that different exercises have the same effect on weight loss at the 0.05 level.

• R codes for two-way ANOVA without interaction
weight.loss <- c(5.4, 4.6, 6.7, 5.1, 4.3, 4.5)
drug <- as.factor(c(1,2,3,1,2,3))
exercise <- as.factor(c(1,1,1,2,2,2))
anova(lm(weight.loss ~ drug + exercise))

The R output is:

Analysis of Variance Table

Response: weight.loss
                      Df Sum Sq Mean Sq F value Pr(>F)
drug                  2 1.3900  0.69500  1.1551 0.4640
exercise              1 1.3067  1.30667  2.1717 0.2785
Residuals            2 1.2033  0.60167

• Suppose that we are given the above ANOVA table. Then we can
  answer the following questions.
  (a) Is there a significant drug effect on weight loss at 0.05 level?
  (b) Is there a significant exercise effect on weight loss at 0.05 level?
  (c) How many levels does the drug factor have?
  (d) How many levels does the exercise factor have?
  (e) What is the sample size?

Answers. (a) No. (b) No. (c) 3. (d) 2. (e) 6.