Introduction to Statistics and Data Science using *eStat*

Chapter 9 Testing Hypothesis for Several Population Means

9.3 Analysis of Variance for Experiments of Two Factors

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

9.3 Analysis of Variance for Experiments of Two Factors

- If there are two factors (A and B) affecting the response variable,
  ⇨ two-way analysis of variances.
  ⇨ frequently used in experiments in engineering, medicine, agriculture.
- Response variable is observed at each combination of two factors.
  ⇨ repeat at least two experiments at each combination of two factors
- Two-way ANOVA tests:
  ⇨ test whether population means of each level of factor A are the same (main effect test of factor A)
  ⇨ test whether population means of each level of factor B are the same (main effect test of the factor B)
  ⇨ test whether effect of factor A is influenced by effect of factor B (interaction effect test).

9.3 Analysis of Variance for Experiments of Two Factors

[Example 9.3.1] Table 9.3.1 shows the yield of three repeated agricultural experiments for each combination of four fertilizer levels and three rise types to investigate yield of rice.

Fertilizer	Types of rice		
	1	2	3
1	64,66,70	72,81,64	74,51,65
2	65,63,58	57,43,52	47,58,67
3	59,68,65	66,71,59	58,45,42
4	58,50,49	57,61,53	53,59,38

1) **Find the average yield for each combination of fertilizers and rice types.**
2) **Using 『eStat』, draw a scatter plot with the rice types (1, 2 and 3) as X-axis and the yield as Y-axis. Separate the color of the scatter plot's dots by the type of fertilizer. Then, show the average of the combinations at each level on the scatter plot and connect them with lines for each type of fertilizer to observe.**
3) **Test main effects of fertilizers and rice types and test interaction effect of the two factors.**
4) **Using 『eStat』, check the results of the two-way analysis of variance.**

9.3 Analysis of Variance for Experiments of Two Factors

File: Ex931ANOVA2_YieldByRiceFertili

Analysis Var: 3: Yield

by Group: 1: Fertilizer

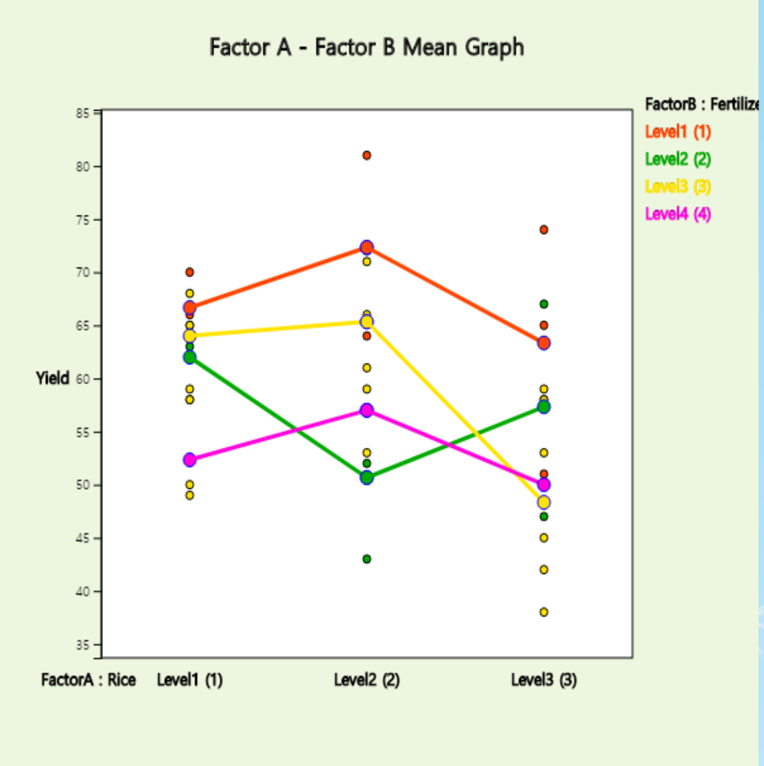
(Selected data: Raw Data) (Select up to two groups)

SelectedVar: V3 by V2,V1,

	Fertilizer	Rice	Yield	V4	V5
1	1	1	64		
2	1	1	66		
3	1	1	70		
4	1	2	72		
5	1	2	81		
6	1	2	64		
7	1	3	74		
8	1	3	51		
9	1	3	65		
10	2	1	65		
11	2	1	63		
12	2	1	58		
13	2	2	57		
14	2	2	43		
15	2	2	52		
16	2	3	47		
17	2	3	58		
18	2	3	67		
19	3	1	59		
20	3	1	68		
21	3	1	65		
22	3	2	66		
23	3	2	71		
24	3	2	59		
25	3	3	58		

Table 9.3.2 Average yield of rice by fertilizers and types of rice (unit kg)

Fertilizer (Factor B)	Types of Rice (Factor A) 1	2	3	Row Average
1	$\overline{y}_{11\cdot} = 66.7$	$\overline{y}_{12\cdot} = 72.3$	$\overline{y}_{13\cdot} = 63.3$	$\overline{y}_{1\cdot\cdot} = 67.4$
2	$\overline{y}_{21\cdot} = 62.0$	$\overline{y}_{22\cdot} = 50.7$	$\overline{y}_{23\cdot} = 57.3$	$\overline{y}_{2\cdot\cdot} = 56.7$
3	$\overline{y}_{31\cdot} = 64.0$	$\overline{y}_{32\cdot} = 65.3$	$\overline{y}_{33\cdot} = 48.3$	$\overline{y}_{3\cdot\cdot} = 59.2$
4	$\overline{y}_{41\cdot} = 52.3$	$\overline{y}_{42\cdot} = 57.0$	$\overline{y}_{43\cdot} = 50.0$	$\overline{y}_{4\cdot\cdot} = 53.1$
Column Average	$\overline{y}_{\cdot1\cdot} = 61.3$	$\overline{y}_{\cdot2\cdot} = 61.3$	$\overline{y}_{\cdot3\cdot} = 54.8$	$\overline{y}_{\cdot\cdot\cdot} = 59.1$



9.3 Analysis of Variance for Experiments of Two Factors

- **Testing main effect of rice types (factor A) :**
  H_0 : Average yields of the three rice types are the same.

  SSA = $12(61.3 - \bar{y}_{\dots})^2 + 12(61.3 - \bar{y}_{\dots})^2 + 12(54.8 - \bar{y}_{\dots})^2 = 342.39$
  ⇨ If SSA is close to zero, all sample means for rice are similar.
  ⇨ Mean squares of factor A : MSA = $\frac{SSA}{(3-1)}$

- **Testing main effect of fertilizer types (factor B) :**
  H_0 : Average yields of the four fertilizer types are the same.

  SSB=$9(67.4 - \bar{y}_{\dots})^2 + 9(56.7 - \bar{y}_{\dots})^2 + 9(59.2 - \bar{y}_{\dots})^2 + 9(53.1 - \bar{y}_{\dots})^2 = 1002.89$
  ⇨ If SSB is close to zero, all sample means for fertilizer are similar.
  ⇨ Mean squares of factor B : MSB = $\frac{SSB}{(4-1)}$

9.3 Analysis of Variance for Experiments of Two Factors

- **Testing interaction effect of rice types and fertilizers :**
  H_0 : There is no interaction effect between rice and fertilizer.

$$
\begin{aligned}
\text{SSAB} = {} & 3(66.7 - \bar{y}_{1..} - \bar{y}_{.1.} + \bar{y}_{...})^2 + 3(72.3 - \bar{y}_{1..} - \bar{y}_{.2.} + \bar{y}_{...})^2 + 3(63.3 - \bar{y}_{1..} - \bar{y}_{.3.} + \bar{y}_{...})^2 \\
& + 3(62.0 - \bar{y}_{2..} - \bar{y}_{.1.} + \bar{y}_{...})^2 + 3(50.7 - \bar{y}_{2..} - \bar{y}_{.2.} + \bar{y}_{...})^2 + 3(57.3 - \bar{y}_{2..} - \bar{y}_{.3.} + \bar{y}_{...})^2 \\
& + 3(64.0 - \bar{y}_{3..} - \bar{y}_{.1.} + \bar{y}_{...})^2 + 3(65.3 - \bar{y}_{3..} - \bar{y}_{.2.} + \bar{y}_{...})^2 + 3(48.3 - \bar{y}_{3..} - \bar{y}_{.3.} + \bar{y}_{...})^2 \\
& + 3(52.3 - \bar{y}_{4..} - \bar{y}_{.1.} + \bar{y}_{...})^2 + 3(57.0 - \bar{y}_{4..} - \bar{y}_{.2.} + \bar{y}_{...})^2 + 3(50.0 - \bar{y}_{4..} - \bar{y}_{.3.} + \bar{y}_{...})^2 \\
= {} & 588.94
\end{aligned}
$$

⇨ Mean squares of interaction AB : MSAB $= \dfrac{SSAB}{(3-1)(4-1)}$

- **Partition of Total Sum of Squares (SST)**

SST = SSA + SSB + SSAB + SSE
⇨ SSE = SST – (SSA + SSB + SSAB)

9.3 Analysis of Variance for Experiments of Two Factors

① Testing of the interaction effect on rice and fertilizer

$$F_0 = \frac{MSAB}{MSE} = \frac{\frac{\mathrm{SSAB}}{(3-1)(4-1)}}{\frac{\mathrm{SSE}}{24}} = 1.77$$

$$F_{6,24\,;\,0.05} = 2.51$$

Since $F_0 < F_{6,24\,;0.05}$, we conclude that there is no interaction. The interaction on rice and fertilizer in <Figure 9.3.1> is so small which is not statistically significant and it may due to other kind of random error. The calculated p-value of $F_0 = 1.77$ using 『eStat』 is 0.1488.

9.3 Analysis of Variance for Experiments of Two Factors

② Testing of the main effect on rice types (Factor A)

$$F_0 = \frac{MSA}{MSE} = \frac{\frac{SSA}{(3-1)}}{\frac{SSE}{24}} = 3.08$$

$$F_{2,24\,;\,0.05} = 3.40$$

Since $F_0 < F_{2,24\,;\,0.05}$, we cannot reject the null hypothesis that average yields of rice types are the same. There is not enough evidence statistically that average yields are different depending on rice types. The calculated p-value of $F_0 = 3.08$ using 『eStat』 is 0.0644.

9.3 Analysis of Variance for Experiments of Two Factors

③ Testing of the main effect on fertilizer types (Factor B)

$$F_0 = \frac{MSAB}{MSE} = \frac{\frac{\mathrm{SSB}}{(4-1)}}{\frac{\mathrm{SSE}}{24}} = 6.02$$

$$F_{3,24\,;\,0.05} = 3.01$$

Since $F_0 > F_{3,24\,;\,0.05}$, we reject the null hypothesis that average yields of fertilizer types are the same. There is enough evidence statistically that average yields are different depending on fertilizer types. Since there is no interaction effect by 1), we can conclude that fertilizer 1 produces more yields than other fertilizer. The calculated p-value of $F_0 = 6.02$ using 『eStat』 is 0.0033.

9.3 Analysis of Variance for Experiments of Two Factors

Factor	Sum of Squares	Degree of freedom	Mean Squares	F value	p-value
Rice Type	342.3889	2	171.1944	3.0815	0.0644
Feritlizer Type	1002.8889	3	334.2963	6.0173	0.0033
Interaction	588.9444	6	98.1574	1.7668	0.1488
Error	1333.3333	24	55.5556		
Total	3267.5556	35			

9.3 Analysis of Variance for Experiments of Two Factors

- **Statistical model of the two factor experiments:**

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}, \quad i = 1, 2, \dots, a; j = 1, 2, \dots, b; \text{ k=1,2,…,r}$$

μ: total mean

α_i: effect of i^{th} level of factor A

β_j: effect of j^{th} level of factor B

γ_{ij}: interaction effect of i^{th} level of factor A and j^{th} level of factor B

ϵ_{ijk} : error terms which are independent and follow $N(0, \sigma^2)$

9.3 Analysis of Variance for Experiments of Two Factors

- The total sum of squared distances from each observation to the total mean $\overline{Y}_{...}$ can be partitioned as the following sum of squares similar to the one way analysis of variance.

Total sum of squares: $SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(Y_{ijk} - \overline{Y}_{...})^2$: degree of freedom: $n-1$

Fator A sum of squares: $SSA = br\sum_{i=1}^{a}(\overline{Y}_{i..} - \overline{Y}_{...})^2$: degree of freedom: $a-1$

Factor B sum of squares: $SSB = ar\sum_{j=1}^{b}(\overline{Y}_{.j.} - \overline{Y}_{...})^2$: degree of freedom: $b-1$

Interaction sum of squares: $SSAB = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})^2$: degree of freedom: $(a-1)(b-1)$

Error sum of squares: $SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(Y_{ijk} - \overline{Y}_{ij.})^2$: degree of freedom: $n-ab$

9.3 Analysis of Variance for Experiments of Two Factors

Table 9.3.4 Two way analysis of variance table

Factor	Sum of Squares	Degree of Freedom	Mean Squares	F value
Factor A	SSA	a-1	MSA = SSA/(a-1)	F_1 = MSA/MSE
Factor B	SSB	b-1	MSB = SSB/(b-1)	F_2 = MSB/MSE
Interaction	SSAB	(a-1)(b-1)	MSAB = SSAB/((a-1)(b-1))	F_3 = MSAB/MSE
Error	SSE	n-ab	MSE = SSE/(n-ab)	
Total	SST	n-1		

9.3 Analysis of Variance for Experiments of Two Factors

1) F Test for the interaction effect:

$H_0 : \gamma_{ij} = 0,\ i = 1,2,\cdots,a;\ j = 1,2,\cdots,b$

If $F_3 = \mathrm{MSAB}/\mathrm{MSE} > F_{(a-1)(b-1),\, n-ab;\, \alpha}$, then reject H_0

2) F Test for the main effect of factor A:

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

If $F_1 = \mathrm{MSA}/\mathrm{MSE} > F_{(a-1),\, n-ab;\, \alpha}$, then reject H_0

3) F Test for the main effect of factor B:

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

If $F_2 = \mathrm{MSB}/\mathrm{MSE} > F_{b-1,\, n-ab;\, \alpha}$, then reject H_0

Thank you