1 Hypothesis Testing

Since we have unequal sample sizes, the factor effect component sum of squares are no longer orthogonal. Therefore, we would use the general linear F-test instead for the testing parts. The basic idea is to compare SSE under the full model with SSE under the reduced model, and we want to test whether specific components could be drop out of the full model. The details are shown as below:

(1) Here the F-test statistic is: $F^* = \frac{\frac{SSE(R)-SSE(F)}{df_R-df_F}}{\frac{SSE(F)}{df_F}}$, where SSE(R) is SSE under the reduced model, df_R is the degree of freedom for the reduced model, SSE(F) is SSE under the full model, and df_F is the degree of freedom for the full model.

(2) F^* follows the F distribution, $F_{(df_R-df_F),df_F}$, under the null hypothesis (H_0).

(3) We would reject H_0 at level α if $F^* > F(1-\alpha;(df_R-df_F),df_F)$, or if the p-value $< \alpha$.

1.1 Test for Interaction Effects

First, we want to test whether or not interaction effects are present. This would assess whether the effect of the factor class size differs across the stratum.

$H_0 : all\ (\tau\beta)_{ij} = 0$

H_a: not all $(\tau\beta)_{ij}'s$ equal zero

Here the full model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$.

And the reduced model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + \epsilon_{ijk}$.

If we reject H_0 at level α, we conclude that there are interaction effects.

1.2 Test for Factor Main Effects

1.2.1 Class Size

Then, we want to test whether or not class size effects are present:

H_0: $\tau_1 = \tau_2 = \tau_3 = 0$

H_a: not all $\tau_i's$ equal zero.

(1) If there are interaction effects, then

- The full model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$.
- And the reduced model is: $Y_{ijk} = \mu_{..} + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$.

(2) If there are no interaction effects, then

- The full model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + \epsilon_{ijk}$.
- And the reduced model is: $Y_{ijk} = \mu_{..} + \beta_j + \epsilon_{ijk}$.

If we reject H_0 at level α, we conclude that the effects of class size are present.

1.2.2 School

Although the class size effects are of our primary interests, we also want to test whether or not school effects are present:

H_0: $\beta_1 = \beta_2 = ... = \beta_{76} = 0$

H_a: not all $\beta_i's$ equal zero.

(1) If there are interaction effects, then

- The full model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$.
- And the reduced model is: $Y_{ijk} = \mu_{..} + \tau_i + (\tau\beta)_{ij} + \epsilon_{ijk}$.

(2) If there are no interaction effects, then

- The full model is: $Y_{ijk} = \mu_{..} + \tau_i + \beta_j + \epsilon_{ijk}$.
- And the reduced model is: $Y_{ijk} = \mu_{..} + \tau_i + \epsilon_{ijk}$.

If we reject H_0 at level α, we conclude the effects of school are present.

1.3 Analysis of Class Size Effects

Because we are interested in the difference in the class size effects, we would do pairwise comparisons among the three class sizes. The Tukey's procedure will be used, and this procedure is conservative result when sample sizes are unequal.

First, we define the differnce between two factor level means $D_{ii'} = \mu_{i.} - \mu_{i'.}$. The point estimate for $D_{ii'}$ is $\hat{D}_{ii'} = \bar{Y_{i..}} - \bar{Y_{i'..}}$. Since $\bar{Y_{i..}}$ and $\bar{Y_{i'..}}$ are independent, the variance of $\hat{D}_{ii'}$ is $\sigma^2\{\hat{D}_{ii'}\} = \frac{\sigma^2}{b^2}\sum_j(\frac{1}{n_{ij}} + \frac{1}{n_{i'j}})$. And the estimated variance of $\hat{D}_{ii'}$ is $s^2\{\hat{D}_{ii'}\} = \frac{MSE}{b^2}\sum_j(\frac{1}{n_{ij}} + \frac{1}{n_{i'j}})$.

Then, we do simultaneous testing:

$H_0 : D_{ii'} = 0$

$H_a : D_{ii'} \neq 0$

If we control the family-wise confidence coefficient at level 1-α, the confidence interval for $D_{ii'}$ is of the form:

$\hat{D}_{ii'} \pm T \times s(\hat{D}_{ii'})$, where $T = \frac{1}{\sqrt{2}}q(1-\alpha; a, n_T - ab)$

We would check whether or not zero is contained in each interval. If zero is contained, we conclude H_0; otherwise, we conclude H_a.

2 Testing Result

We use significance level 0.05 for all the following tests.

2.1 Test for Interaction Effects

The results of F-test for interaction effects is shown in Table 1.

Table 1: Test for Interaction Effects

Model	Degree of Fredom	SSE	F^*	P-value
Full	114	34612		
Reduced	260	81345	1.0543	0.3855

Since p-value = 0.3855, we can not reject $H_0 : all\ (\tau\beta)_{ij} = 0$ at level of significance level 0.05. We conclude that there is no interaction between these two factors.

As a result, we would revised the full model by excluding the interaction effects for the following tests. Also, we use this new full model for the main analysis.

2.1.1 Class Size

The results of F-test for class type main effects is shown in the Table 2.

Table 2: Test for Factor Main Effects

Model	Degree of Fredom	SSE	F^*	P-value
Full	334	221371		
Reduced	336	232391	8.3137	0.0002995

Since p-value = 0.0002995, we reject H_0: $\tau_1 = \tau_2 = \tau_3 = 0$ at level of significance level 0.05. We conclude that there are class type main effects.

2.1.2 School

The results of F-test for school effects are shown in Table 3.

Table 3: Test for Factor Main Effects

Model	Degree of Fredom	SSE	F^*	P-value
Full	334	221371		
Reduced	335	230604	13.931	0.0002228

Since p-value = 0.0002228, we reject H_0: $\beta_1 = \beta_2 = ... = \beta_{76} = 0$ at level of significance level 0.05. We conclude that there are school main effects.

2.2 Analysis of Class Size Effects

From Table 4, we could see that all the confidence intervals do no contain zero. So we conclude that, at family-wise level $\alpha = 0,05$, $\mu_{1.}$ and $\mu_{2.}$, $\mu_{2.}$ and $\mu_{3.}$, $\mu_{1.}$ and $\mu_{3.}$ are different. Moreover, the small classes outperformed both regular classes and regular classes with aides.

Table 4: Pairwise comparisons of factor level means

	$\hat{D}_{ii'}$	Lower bound	Upper bound	p-value
D_{12}	0.984	-7.303	9.271	0.958
D_{32}	12.297	4.435	20.159	0.001
D_{31}	11.313	3.152	19.474	0.003