DISCUSSION 11

-Review on split-plot design :

1. When it is used? or what is the advantage of this type of designs?

   Split-plot designs are needed when the levels of some treatment factors are more difficult to change during the experiment than those of others. The designs have a nested blocking structure: split plots are nested within whole plots, which may be nested within blocks.

2. A model for split-plot design Randomized as a CRD:

$$Y_{ijk} = \mu + \alpha_i + \epsilon_{ij}$$
$$+ \beta_k + (\alpha\tau)_{ik} + \delta_{ijk}.$$
$$i = 1, 2, ..., a, j = 1, 2, ..., b, k = 1, 2, ..., c.$$

   where μ is the grand mean, α_i is the effect of factor A at level i; j is index for the replicates; τ_k the effect of factor B at level k, $(\alpha\tau)_{ik}$ is interaction effect.

   $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_W^2)$, error within the whole plot;

   $\delta_{ijk} \overset{iid}{\sim} N(0, \sigma_S^2)$, error within the subplot and random noise. the errors are independent. $i = 1, 2, ..., a, j = 1, 2, ..., b, k = 1, 2, ..., c.$

   Note the nested blocking structure: whole plots are nested within the blocks, and split-plots are nested within the whole plots.

3. How to analyze the data collected from this design?

ANOVA

Source	df	SS	MS	EMS
A	$a-1$	SSA	MSA	$\sigma_\delta^2 + c\sigma_\varepsilon^2 + \frac{bc\sum_i \alpha_i^2}{a-1}$
Whole Plot Error	$a(b-1)$	SSWPErr	MSWPErr	$\sigma_\delta^2 + c\sigma_\varepsilon^2$
C	$c-1$	SSC	MSC	$\sigma_\delta^2 + \frac{ab\sum_k \tau_k^2}{c-1}$
AC	$(a-1)(c-1)$	SSAC	MSAC	$\sigma_\delta^2 + \frac{b\sum_i\sum_k (\alpha\tau)_{ik}^2}{(a-1)(c-1)}$
SubPlot Error	$a(b-1)(c-1)$	SSSPErr	MSSPErr	σ_δ^2
Total	$abc-1$	SSTot		

The SubPlot Error is most easily calculated by subtraction. The WPError is equivalent to the SS for Whole Plots nested in A. In that sense, the Whole Plot Error reminds you of the sample or plot error in a design with subsampling. All other SS are calculated the usual way.

$\hat{\sigma}_S^2 = MSSPErr$

$\hat{\sigma}_W^2 = MSE^s = \frac{MSWPErrC - MSSPErr}{c}$, this is because there are b subplots in each whole-plot. the whole-plot error mean square is $c\sigma_W^2 + \sigma_S^2$

4. Tests/questions can be answered by the Anova table.

   Note for testing of equal effects of factor A, the whole-plot mean square error is used. It should also be used for main effect contrasts of factor A. (levels of factor A are assigned the whole plot.)

   For tests or main effect contrasts of factor B, or AB interaction contrasts, the split-plot mean square error is used.

-Review on some Model selection

1. Hierarchy rule, parsimony rule:

   Hierarchy rule: when you include interaction of two factors, say A, B, then you should want to keep A, B in the model. Help with the interpretation of the model.

   Parsimony rule: when two models perform similarly, we prefer the simple models to the complicated one.

2. Model comparison of two nested models(Lack-of-lack of test)

Source	Df	RSS	MRSS	F
reduced model	df1	RSS(reduced)	MRSS(reduced)	
full model	df2	RSS(full)	MRSS(full)	
Difference	df1-df2	RSS1-RSS2	MSS(diff)	MSS(diff)/MSS(full)

   If the F-test is not significant, this indicates that there is no significant difference between the two tested models. Based on **rule of parsimony**, we keep the simple(reduced) model instead of the full model. Otherwise, we keep the full model.

3. Model selection criteria(take into account both model fitting and parsimony)

   AIC, BIC, Adjusted R^2, Mallow's C_p

-Hand calculations in the example below:

Split Plot Arrangement

The split plot arrangement is specifically suited for a two or more factor experiment.

This arrangement can be used with the CRD, RCBD, and LS designs discussed in this course.

Features of this design are that plots are divided into whole plots and subplots.

<u>Example</u>

Whole plots are wheat varieties (a_0 to a_3) and subplots are rates of a herbicide (b_0 to b_2).

a_1	a_2	a_3	a_0
b_1	b_2	b_0	b_2
b_2	b_0	b_1	b_0
b_0	b_1	b_2	b_1

With a split plot arrangement, the precision for the measurement of the effects of the whole plot factor(s) are sacrificed to improve that of the subplot factor.

Measurement of the subplot factor and its interaction with the main-plot factor is more precise than that obtained with an RCBD with a factorial arrangement.

<u>Determining Which Factor to Use as the Whole and Subplot Factors</u>

With the split plot arrangement, plot size and precision of measurement of the effects are not the same for whole and subplot factors. Thus, assignment of a particular factor to either the whole or subplot is extremely important. To make a choice, the following guidelines are suggested:

1. **Degree of Precision:** for a greater deal of precision for factor B than factor A, assign factor B to the subplots and factor A to the whole plots.

2. **Relative Size of the Main Effects:** If the main effect of one factor (.e.g., factor A) is expected to be much larger and easier to detect than that of the other factor (e.g., factor B), factor A should be assigned to the whole plots and factor B to the subplots. This may increase the chances of detecting differences among levels of factor B.

3. **Management Practices:** Cultural practices required by a factor may dictate use of large plots. In such a case, such factors should be assigned to whole plots.

Randomization and Layout

The randomization procedure for split plots consists of two parts:

1. Randomly assign whole plot treatments to whole plots based on the experimental design used.

2. Randomly assign subplot treatments to subplots. The randomization procedure has no effect on assignment of subplot treatments to subplots.

Expected Mean Squares for the Split Plot Arrangement

The example to be given will be for an RCBD with factors A and B considered as random effects.

Source of variation	df	Expected mean square
Replicate	r-1	$\sigma^2 + b\sigma_\gamma^2 + ab\sigma_R^2$
A	a-1	$\sigma^2 + b\sigma_\gamma^2 + r\sigma_{AB}^2 + rb\sigma_A^2$
Error (a) = Rep x A	(r-1)(a-1)	$\sigma^2 + b\sigma_\gamma^2$
B	b-1	$\sigma^2 + r\sigma_{AB}^2 + ra\sigma_B^2$
AxB	(a-1)(b-1)	$\sigma^2 + r\sigma_{AB}^2$
Error (b) = Rep x B(A)	a(b-1)(r-1)	σ^2
Total	rab-1	

ANOVA of a Split Plot Arrangement

Data Set

Treatments		Blocks (j)			
A_i	B_k	1	2	3	$Y_{i.k}$
a_0	b_0	13.8	13.5	13.2	40.5
	b_1	15.5	15.0	15.2	45.7
	b_2	21.0	22.7	22.3	66.0
	b_3	18.9	18.3	19.6	56.8
Main plot total ($Y_{0j.}$)		69.2	69.5	70.3	209.0 = $Y_{0..}$
a_1	b_0	19.3	18.0	20.5	57.8
	b_1	22.2	24.2	25.4	71.8
	b_2	25.3	24.8	28.4	78.5
	b_3	25.9	26.7	27.6	80.2
Main plot total ($Y_{1j.}$)		92.7	93.7	101.9	288.3 = $Y_{1..}$
Rep total ($Y_{.j.}$)		161.9	163.2	172.2	497.3 = $Y_{...}$

Treatment Totals Table

	a_0	a_1	ΣB_k
b_0	40.5	57.8	98.3
b_1	45.7	71.8	117.5
b_2	66.0	78.5	144.5
b_3	56.8	80.2	137.0
ΣA_i	209.0	288.3	497.3

Step 1. Calculate Correction Factor:

$Y_{...}^2$ / (r x a x b) = 497.3^2 / (3 x 2 x 4) = **10,304.47**

Step 2. Calculate Total SS

ΣY_{ijk}^2 - CF = ($13.8^2 + 15.5^2 + 18.9^2 + \cdots + 27.6^2$) - CF = **516.2**

Step 3. Calculate Replicate SS

$\Sigma Y_{.j.}^2$ / axb - CF = ($161.9^2 + 163.2^2 + 172.2^2$) / (2x4) - CF = **7.87**

Step 4. Calculate A SS

$\Sigma Y_{i..}^2$ / rb - CF = ($209^2 + 288.3^2$)/(3x4) -CF = **262.02**

Step 5. Calculate Whole Plot SS

$\Sigma Y_{ij.}^2$ /b - CF = ($69.2^2 + 69.5^2 + \cdots + 101.9^2$) / 4 - CF = **274.92**

Step 6. Calculate Whole Plot Error SS = Error(a) SS

Whole Plot SS - A SS - Rep SS = **5.03**

Step 7. Calculate B SS

$\Sigma Y_{..k}^2$ / ra - CF = ($98.3^2 + 117.5^2 + 144.5^2 + 137.0^2$) / (3x2) - CF = **215.26**

Step 8. Calculate AxB SS

$\Sigma Y_{i.k}^2$/ r - CF - A SS - B SS = ($40.5^2 + 45.7^2 + \cdots + 80.2^2$)/3 - CF - A SS - B SS
= 18.7

Step 9. Calculate Error(b) SS = Total SS - Rep SS - A SS - Error(a) SS - B SS - AxB SS
= **7.24**

Step 10. Make ANOVA Table (Assuming A and B are fixed effects)

SOV	df	SS	MS	F
Replicate	2	7.87	3.935	6.53^{*}
A	1	262.02	262.02	104.183^{**}
Error(a)	2	5.03	2.515	
B	3	215.26	71.753	119.993^{**}
AxB	3	18.70	6.233	10.337^{**}
Error(b)	12	7.24	0.603	
Total	23	516.12		

<u>LSD's for Split Plot Arrangement</u>

1. To compare two whole plot means averaged over all subplot treatments (.e.g. a_0 vs. a_1)

$$t_{\alpha/2, err(a)df}\sqrt{\frac{2Error(a)MS}{rb}} = 4.303\sqrt{\frac{2(2.515)}{3x4}} = 2.79$$

2. To compare two subplot means average over all whole plot treatments (.e.g. b_0 vs. b_3)

$$t_{\alpha/2, err(b)df}\sqrt{\frac{2Error(b)MS}{ra}} = 2.179\sqrt{\frac{2(0.603)}{3x2}} = 0.98$$

3. To compare two subplot treatment means for the same whole plot treatment (.e.g. a_0b_0 vs. a_0b_3).

$$t_{\alpha/2, err(b)df}\sqrt{\frac{2Error(b)MS}{r}} = 2.179\sqrt{\frac{2(0.603)}{3}} = 1.38$$

4. To compare two whole plot means at the same or different sub plot treatments (.e.g. a_0b_0 vs. a_1b_0) or (.e.g. a_0b_0 vs. a_1b_3).

$$t'_{\alpha/2,ab}\sqrt{\frac{2[(b-1)Error(b)MS + Error(a)MS]}{rb}}$$

Where t'_{ab} is a weighted estimate of t that can be calculated using the following formula:

$$t'_{\alpha/2,ab} = \frac{(b-1)Error(b)MS * t_{\alpha/2,Err(b)df} + Error(a)MS * t_{\alpha/2,Err(a)df}}{(b-1)Error(b)MS + Error(a)MS}$$

NOTE: This formula is to calculate t, not the degrees of freedom of t as was done for other situations.

$$t'_{ab} = \frac{(4-1)(.603)(2.179) + (2.515)(4.303)}{(4-1)(.603) + 2.515} = \mathbf{3.414}$$

Therefore, the LSD $= 3.414\sqrt{\frac{2[(4-1)(0.603)+2.515]}{3x4}} = 2.90$

Table of Means for the example

	A levels		
B levels	a_0	a_1	Mean of B
b_0	13.5	19.3	16.4
b_1	15.2	23.9	19.6
b_2	22.0	26.2	24.1
b_3	18.9	26.7	22.8
Mean of A	17.4	24.0	

Table of Means for the example

LSD #4

B levels	A levels: a_0	a_1	Mean of B
b_0	13.5	19.3	16.4
b_1	15.2	23.9	19.6
b_2	22.0	26.2	24.1
b_3	18.9	26.7	22.8
Mean of A	17.4	24.0	

LSD #3

LSD #2

LSD #4

LSD #1