CHAPTER 8. RANDOMIZED COMPLETE BLOCK DESIGN WITH AND WITHOUT SUBSAMPLES

The randomized complete block design (RCBD) is perhaps the most commonly encountered design that can be analyzed as a two-way AOV. In this design, a set of experimental units is grouped (blocked) in a way that minimizes the variability among the units within groups (blocks). The objective is to keep the experimental error within each block as well as possible. Each block contains a complete set of treatments, therefore differences among blocks are not due to treatments, and this variability can be estimated as a separate source of variation. The removal of an appreciable amount of this source of variation reduces experimental error and improves the ability of the experiment to detect smaller treatment differences. The greater the variability among blocks the more efficient the design becomes. In the absence of appreciable block differences the design is not as efficient as a completely randomized design (CRD). The CRD has more degrees of freedom for error and a smaller F value is required for significant difference among treatments. The paired sample experiment discussed in Chapter 6 is the simplest case of using the concept of blocking, where pairs are blocks.

8.1 Randomized Complete Block Design Without Subsamples

In animal studies, to achieve the uniformity within blocks, animals may be classified on the basis of age, weight, litter size, or other characteristics that will provide a basis for grouping for more uniformity within blocks. For plants in field trials, land is normally laid out in equal-sized blocks, each block being subdivided into as many equal-sized plots as there are treatments to be studied. In general, it is most efficient to have a single replicate of each treatment per block. There may be situations, however, when it is desirable to have more than one replicate per blocks.

Randomization

After experimental units have been grouped into blocks, treatments are assigned randomly within a block, and separate randomizations are made for each block.

To illustrate the randomization and the AOV for a RCBD, consider the layout of the field plots in Figure 8-1. The sugar beet root yield data shown in Figure 8-1 are the same as in Table 7-1 and Figure 7-1. This experiment was actually performed as a RCBD but was analyzed as a CRD in Chapter 7 to provide a basis for comparing the two designs.

The six treatments in each block were randomly assigned to the six plots by drawing random numbers from Appendix Table A-1 in the manner described in Chapter 7. Note in this case that there are only six random numbers (1 - 6) to be drawn for each block, e.g., for block 1 the random sequence was 3, 6, 5, 2, 1, and 4. Assigning treatments A-F to numbers 1-6 results in the block 1 treatment sequence.

Block 1	Block 2	Block 3	Block 4	Block 5
C 1 (40.9)	A 7 (33.4)	B 13 (37.4)	D 19 (40.1)	C 25 (39.8)
F 2 (40.6)	D 8 (41.7)	C 14 (39.5)	C 20 (38.6)	D 26 (40.0)
E 3 (39.7)	B 9 (37.5)	D 15 (39.4)	E 21 (38.7)	A 27 (33.9)
B 4 (38.8)	F 10 (41.0)	E 16 (39.2)	A 22 (32.2)	B 28 (38.4)
A 5 (31.3)	E 11 (40.6)	F 17 (41.5)	F 23 (41.1)	E 29 (41.9)
D 6 (40.9)	C 12 (39.2)	A 18 (29.2)	B 24 (35.8)	F 30 (39.8)

Figure 8-1. Field plots layout in a RCBD. Plots are numbered in the lower left. Treatments A-F are levels of nitrogen fertilizer from 0 - 250 lbs/acre in 50 lb increments. The number in parenthesis is the root yield per plot in tons/acre.

Analysis of Variance

To proceed with the AOV, the results shown in Figure 8-1 are organized by blocks and treatments in Table 8-1.

Table 8-1. Sugar beet root yield data (tons/acre).

A (0)	31.3	33.4	29.2	32.2	33.9	160.0	32.00
B (50)	38.8	37.5	37.4	35.8	38.4	187.9	37.58
C (100)	40.9	39.2	39.5	38.6	39.8	198.0	39.60
D (150)	40.9	41.7	39.4	40.1	40.0	202.1	40.42
E (200)	39.7	40.6	39.2	38.7	41.9	200.1	40.02
F (250)	40.6	41.0	41.5	41.1	39.8	204.0	40.80
Block ($Y_{.j}$) Total	232.2	233.4	226.2	226.5	233.8	$Y_{..}$=1152.1	
Block ($\bar{Y}_{.j}$) mean	38.70	38.90	37.70	37.75	38.97		$\bar{Y}_{..}$= 38.40

A generalized outline of the AOV for a RCBD is shown in Table 8-2. Our main concern in this design is still to test the equality of treatment means. However, now we can also test for a significant block effect.

Table 8-2. AOV for a RCBD.

Source	df	Sum of squares (SS)	Mean square (MS)	Observed F
Total	kr-1	TSSS		
Block	r-1	SSB	MSB	MSB/MSE
Treatment	k-1	SST	MST	MST/MSE
Exp. error	(k-1)(r-1)	SSE	MSE	

The AOV for the data in Table 8-1 is given in Table 8-3. Calculations for completing the table are shown below.

Table 8-3. Two-way AOV for the sugar beet yield data.

Source	df	SS	MS	F
Total	29	311.13		
Block	4	9.44	2.36	1.97
Treatment	5	277.69	55.54	46.28
Exp. error	20	24.00	1.20	

Step 1. Outline the AOV table and list the sources of variation and degrees of freedom to provide the entries for the first two columns of Table 8-3.

Step. 2. Correct factor (C)

$$C = Y^2../rk = (1152.1)^2/(5)(6) = 44244.48$$

where r is the number of blocks.

Step 3. Total sum of squares (TSS)

$$TSS = \Sigma\Sigma(Y_{ij} - \overline{Y}..)^2$$

$$= \Sigma\Sigma Y_{ij}^2 - C = 44555.61 - C = 311.13$$

Step 4. Block sum of squares (SSB) and mean square (MSB).

$$SSB = k\Sigma(\overline{Y}_{\cdot j} - \overline{Y}..)^2$$

$$= \Sigma\Sigma_{.j}^2/k - c = 44253.92 - C = 9.44$$

$$MSB = SSB/(r-1) = 9.44/4 = 2.36$$

Step 5. Treatment sum of squares (SST) and mean square (MST).

$$SST = r\Sigma(\overline{Y}_i. - \overline{Y}..)^2$$

$$= \Sigma Y_i^2./r - C = 44522.17 - C = 277.69$$

$$MST = SST/(k-1) = 277.69/5 = 55.54$$

Step 6. Error sum of squares (SEE) and mean square (MSE).

SST	=	TSS - SSB - SST
	=	311.13 - 9.44 - 277.69 = 24.00
MSE	=	SSE/(k-1) (r-1) = 24/(5) (4) = 1.20

MSE represents the variability among experimental units that is not accounted for by any known source of variation. Thus the sum of squares for error is most easily obtained by subtracting the known sources of variation, i.e., blocks and treatments, from the total variation. To gain some insight into the nature of "experimental error" for this design, each observation can be expressed in the following form:

$$Y_{ij} = \mu + (\mu_i. - \mu) + (\mu._j - \mu) + \varepsilon_{ij}$$

where μ is the overall mean (estimated by $\overline{Y}.. = 38.4$), $\mu I. - \mu$ represents the i^{th} treatment effect (estimated by $\overline{Y}_i. - \overline{Y}..$), and $\mu._j - \mu$ represents the j^{th} block effect (estimated by $\overline{Y}._j - \overline{Y}..$). Thus the experimental error, ε_{ij}, is the difference between the observation, Y_{ij} and the effects of known sources of variation,

$$\varepsilon_{ij} = Y_{ij} - \mu - (\mu_i. - \mu) - (\mu._j - \mu)$$

$$= Y_{ij} - (\mu_i. + \mu_j - \mu)$$

To illustrate, we will calculate the estimated error component for plot 1 (which is treatment C in block 1),

$$\hat{\varepsilon}_{31} = Y_{31} - (\overline{Y}_3. + \overline{Y}._1 - \overline{Y}..)$$
$$= 40.9 - (39.6 + 38.7 - 38.4)$$
$$= 1.0$$

Performing this calculation for each plot of the experiment will yield the estimated errors. Squaring and summing these errors will result in the sum of squares for experimental error, i.e.,

$$SSE = \Sigma\Sigma(Y_{ij} - \overline{Y}_i. - \overline{Y}._j + \overline{Y}..)^2$$

Step 7. Calculate F values

For blocks: $F = MSB/MSE$

$= 2.36/1.20 = 1.97$ with 4 and 20 df.

For treatments: $F = MST/MSE$

$= 55.54/1.20$

$= 46.28$ with 5 and 20 df.

The F value for blocks is not significant at the 5% level (Appendix Table A-7), but the F value for treatment is highly significant ($P < 0.01$) and is considerably larger than the F value obtained when block effects are ignored in the AOV in Chapter 7.

8.2 Design Efficiency

In testing treatment differences, several alternative experimental designs may be used. However, the several designs that may be equally valid for testing treatment effects are rarely equally efficient. Efficiency may be defined in terms of the cost of experimentation, time to collect data, precision of the data obtained, etc. A commonly used index for comparing the efficiency of two different designs is the inverse ratio of the variance pr unit, i.e., the MSE's. Since different designs may have different degrees of freedom for error, a correction factor, suggested by Fisher, which multiplies the inverse ratio of variances will give a better measure of the relative efficiency (RE).

$$\text{RE (design A to design B)} = \frac{(df_A + 1)\,(df_B + 3)}{(df_B + 1)\,(df_A + 3)} \bullet \frac{MSE_B}{MSE_A}$$

where MSE_A is the mean square error obtained from design-A with degrees of freedom df_A, and MSE_B is the mean square of design-B with degrees of freedom df_B. If RE>1, design A is more efficient. If RE<1, the converse is true.

If a randomized complete block design (say, design-A) is used, one may want to estimate the relative efficiency compared with a completely randomized design (say, design-B).

This is possible by using the following equation to estimate the MSE of CRD (MSE_B) from the information obtained in the AOV of RCBD,

$$MSE_B = [SSB_A + SSE_A + (k-1)MSE_A] / (kr - 1)$$

where k is the treatment number and r is the number of blocks.

To illustrate, we will use the experiment shown in Figure 8-1 and the results presented in Table 8-3.

$$MSE_B = [9.44 + 24.00 + 5(1.20)] / 29 = 1.36$$

and

$$\text{RE (RCBD to CRD)} = \frac{(20+1)(24+3)}{(24+1)(20+3)} \bullet \frac{1.36}{1.20}(100)$$
$$= (0.986)\,(1.133)\,(100) = 111.8\%$$

Thus the estimated RE, 111.8%, implies the RCBD is slightly more efficient than the CRD for this experiment. It should be pointed out that although the F test for block effects is not significant at the 5% level, it is significant at the 13% level (calculated but not shown). In order to obtain as much information as 5 blocks for an RCBD, one needs 5(1.118) or about 6 replicates per treatment in a CRD.

8.3 RCBD with Subsamples

To illustrate a RCBD with subsamples, we will use the data for the sucrose content of sugar beet related to the N-fertilization presented in Table 7-4. As already pointed out, this experiment was actually done as an RCBD. The data are again presented in Table 8-4. Note that the only new features of the table are block totals and means. An outline of symbols for the AOV of a RCBD with subsamples is given in Table 8-5.

Table 8-4. % sucrose for two beet samples per plot from an RCBD.

Treatment (lb N/acre)	Sub-sample	Blocks 1	2	3	4	5	Total (Y_{i}..)	Mean ($\bar{Y}_i$..)
A(0)	1	16.5	16.4	15.7	16.6	16.0		
	2	16.4	15.8	15.3	16.1	16.8	161.60	16.16
	Y_{ij}.	32.9	32.2	31.0	32.7	32.8		
B(50)	1	16.0	14.4	15.5	15.6	16.4		
	2	16.6	13.9	16.6	16.2	16.2	157.40	15.74
	Y_{2j}.	32.6	28.3	32.1	31.8	32.6		
C(100)	1	15.1	15.0	15.9	16.1	15.0		
	2	15.6	14.3	16.2	15.2	14.5	152.90	15.29
	Y_{3j}.	30.7	29.3	32.1	31.3	29.5		
D(150)	1	15.6	14.7	15.6	15.4	15.6		
	2	15.5	15.2	15.5	14.6	15.2	152.90	15.29
	Y_{4j}.	31.1	29.9	31.1	30.0	30.8		
E(200)	1	13.5	14.2	14.5	15.4	14.9		
	2	14.3	13.3	15.1	15.1	13.3	143.60	14.36
	Y_{5j}.	27.8	27.5	29.6	30.5	28.2		
F(250)	1	14.2	12.5	15.1	14.0	14.3		
	2	13.0	12.6	14.3	14.8	14.6	139.40	13.94
	Y_{6j}.	27.2	25.1	29.4	28.8	28.9		
Block ($Y_{.j.}$) total		182.30	172.3	185.3	185.1	182.8	907.8 = Y...	
Block ($Y_{.j.}$) total Block ($\bar{Y}_{.j.}$)		15.19	14.36	15.44	15.43	15.23	15.13 ~ = ~ $\bar{Y}$	

Table 8-5. AOV of RCBD with n subsamples.

Source	df	Sum of squares	Mean square	Observed F
Total (samples)	krn-1	TSS		
Exp. units	kr-1	SSU		
Blocks	r-1	SSB	MSB	MSB/MSE
Treatments	k-1	SST	MST	MST/MSE
Exp. error	(r-1)(k-1)	SSE	MSE	MSE/MSS
Sampling error	kr(n-1)	SSS	MSS	

The AOV for Table 8-4 is shown in Table 8-6 and the steps for completing this table are given below.

Table 8-6. AOV for the data of Table 8-4.

Source	df	SS	MS	F
Total (samples)	59	62.25		
Plots (exp. units)	29	55.71		
Blocks	4	9.53	2.38	4.25
Treatments	5	34.94	6.99	12.48
Exp. error	20	11.24	0.56	2.43
Sampling error	30	6.94	0.23	

Step 1: Degrees of freedom

Degrees of freedom are determined as shown in Table 8-5. For this experiment, treatments (k) = 6, blocks (r) = 5, and samples (n) = 2

Step 2: Correction factor

$$C = Y^2.../krn = 907.8^2/6(5)(2) = 13{,}735.01$$

Step 3: Total (samples) sum of squares.

$$TSS = \Sigma\Sigma\Sigma(Y_{ijh} - \overline{Y}...)^2$$

$$= \Sigma\Sigma\Sigma Y_{ijh}^2 - C$$

$$= 16.5^2 + 16.4^2 + ... + 14.6^2 - 13{,}735.01$$

$$= 62.25$$

Step 4: Sum of squares for plots (exp. units).

$$SSU = n\Sigma\Sigma(\overline{Y}_{ij}. - \overline{Y}...)^2$$

$$= \Sigma\Sigma Y^2_{ij}./n - C$$

$$= (32.9^2 + ... + 28.9^2)/2 - 13{,}735.01$$

$$= 55.71$$

Step 5: Sum of squares and mean square for blocks.

$$SSB = kn\Sigma(\overline{Y}._j. - \overline{Y}...)^2$$

$$= \Sigma Y^2._j./kn - C$$

$$= (182.30^2 + ... + 182.8^2)/6(2) - 13{,}735.01$$

$$= 9.53$$

$$MSB = SSB/(r-1) = 9.53/4 = 2.38$$

Step 6: Sum of squares and mean square for treatments.

$$SST = rn\ \Sigma\ (\overline{Y}_i.. - \overline{Y}...)^2$$

$$= \Sigma Y^2_i../rn - c$$

$$= (161.6^2 + \ldots + 139.4^2)/5\ (2) - 13{,}735.01$$

$$= 34.94$$

$$MST+ SST/(k-1) = 34.94/5 = 6.99$$

Step 7: Sum of squares and mean square for experimental error.

$$SSE = SSU - SSB - SST$$

$$= 55.71 - 9.53 - 34.94$$

$$= 11.24$$

$$MSE = SSE/(k-1)\ (r-1)$$

$$= 11.24/5(4)$$

$$= 0.56$$

Step 8: Sum of squares and mean square for sampling error.

$$SSS = TSS - SSU$$

$$= 62.65 - 55.71$$

$$= 6.94$$

$$MSS = SSS/kr(n-1)$$

$$= 6.94/6(5)(2-1)$$

$$= 0.23$$

Step 9: Calculate F values.

For testing the equality of block effects; F = MSB/MSE = 2.38/0.56 = 4.25, which is nearly significant at the 1% level ($F_{0.01,4,20} = 4.43$), indicating that not all block means are equal. Note that the block effect for sucrose content is considerably higher than the block effect for the root yield in Table 8-3, where F = 1.97. Thus, the relative efficiency of the RCBD versus CRD is higher for % sucrose than it was for root yield. As an exercise the reader may wish to calculate the RE.

For testing the equality of treatment effects; F = MST/MSE = 6.99/0.56 = 12.48, which exceeds the 1% tabular value ($F_{0.01,5,20} = 4.10$). Thus there are some significant treatment differences. For testing the significance of variability among experimental units treated alike over and above the variability due to sampling units, F = MSE/MSS = 0.56/0.23 = 2.43, which is greater than the 5% tabular value ($F_{0.05,20,30} = 1.93$). This indicates the existence of variability from plot to plot within treatments apart from sampling variability.

Note that MSE is used as the denominator in the F tests for blocks and treatments. This is because MSE contains all the random variation due to samples and experimental units. In fact MSE estimates $\sigma_s^2 + n\sigma_e^2$ where σ_s^2 is the variance component due to samples and σ_e^2 the component due to the experimental units. Both MSB and MST contain these random variations plus additional variation due to block or treatment effects, i.e

$$MST / MSE \sim (\sigma_s^2 \_ n\sigma_e^2 + \delta_t) / (\sigma_s^2 + n\sigma_e^2)$$

$$MSB / MSE \sim (\sigma_s^2 + n\sigma_e^2 + \delta_B) / (\sigma_s^2 + n\sigma_e^2)$$

where $\delta_g = rn\Sigma(\mu)^2 / (k-1)$ and

$$\delta_B = kn\Sigma(\mu._j - \mu)^2 / (r-1)$$

Since MSS only estimates σ_e^2 is a part of the random variation in the experiment, it should not be used as the divisor for testing block or treatment effects. For example:

$$MST / MSS \sim (\sigma_s^2 + n\sigma_e^2 + \delta_t) / \sigma_s^2$$

Now a significant σ_e^2 would result in a significant F test for treatment whether or not there is any real treatment effect. Therefore, replications of experimental units are essential to provide a valid estimate of experimental error for comparisons among treatments and should not be replaced by taking multiple samples from a single experimental unit per treatment.

8.4 The Relationship Between t and F Tests

In section 6.2, we discussed the use of paired samples to compare two treatments. Actually, this is the simplest RCBD and can be analyzed by AOV to yield the same statistical conclusion. To illustrate, the data in Table 6-2 are repeated in Table 8-7. Note that pairs of plots are now called blocks.

Table 8-7. Sugar beet root yield (tons/acre) paired plots (blocks).

	Blocks					Treatment	
Treatment (lb N/acre)	1	2	3	4	5	Total $Y_{i\cdot}$	Mean $\bar{Y}_{i\cdot}$
50	38.8	37.6	37.4	35.8	38.4	188.0	37.6
100	40.9	39.2	39.5	38.6	39.8	198.0	39.6
Block total $Y_{\cdot j}$	79.7	76.8	7.69	74.4	78.2	$Y_{\cdot\cdot}$ = 386.0	
Block total $Y_{\cdot j}$	39.85	38.40	38.45	37.2	39.1	$\bar{Y}_{\cdot\cdot}$ = 38.6	
Block mean $\bar{Y}_{\cdot j}$							

Table 8-8. AOV of data in Table 8-7.

Source of variation	df	Sum of squares	Mean square	F
Total	9	18.26		
Block	4	7.67	1.92	12.80
Nitrogen	1	10.00	10.00	66.67
Exp. error	4	0.59	0.15	

$C = 386.0^2 / 10 = 14899.60$

$TSS = 38.82^2 + ... + 39.8^2 - C = 18.26$

$SSB = (79.7^2 + ... + 78.2^2)/2 - C = 7.67$

$SST = (188.0^2 + 198.02^2)/5 - C = 10.00$

$SSE = TSS - SSB - SST = 0.59$

Mean squares are obtained by dividing sum of squares by their respective degrees of freedom. For the comparison of the equality of treatment mean,

$$F = 10.00/0.15 = 66.67$$

Note that $\sqrt{66.67} = 8.16$ which approximately equals the t-value (t = 8.5) obtained in the section where data was analyzed as the paired t-test (section 6.2).

This result and that of section 7.3 demonstrate the fact that F and t tests are statistically equivalent in comparing two treatment means. An experiment involving two independent samples is the simplest CRD, and an experiment involving paired samples is the simplest RCBD.

8.5 Factorial Experiments

When two or more factors, each having 2 or more levels are investigated simultaneously in all possible combinations, the resulting treatments (the combinations) are said to be factorial. Factorial treatments can be studied in any appropriate experimental design. Examples of factorial experiments are the testing of 3 doses of a hormone on several breeds of animals, testing 2 or more varieties at different rates of N-fertilization, the study of bread quality as affected by the levels of protein in flour and baking temperatures.

If the effect of one factor is modified by the effect of another factor, they are said to interact. For instance, an interaction is present if one breed of cattle responds more to a hormone treatment at high concentrations but not at low concentrations than a second breed. For illustration, see Figure 8-2.

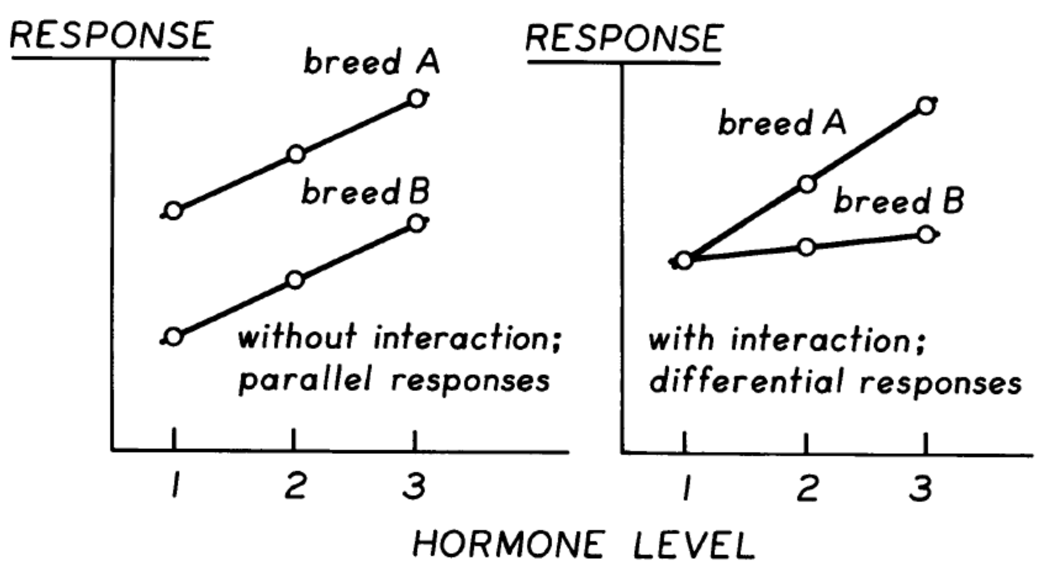


Figure 8.2. Two factors, with and without interaction.

Factorial treatments are usually designed to discover interactions which are often important in biological systems. When two factors interact, the effect of a factor varies depending upon the second factor. In this case, results from single factor experiments can be misleading. For example, consider a soil deficient in both nitrogen and phosphorus. A single factor experiment to determine the required amount of fertilizer nitrogen will only give the P_0

curve of Figure 8-3 which indicates a large quantity of nitrogen is needed for maximal yield. A factorial experiment with several levels of nitrogen and phosphorous will reveal the fact that less nitrogen is required in the presence of a higher level of phosphorous.

Even when there is no interaction, factorial experiments have the advantage of enlarging inferences about the main effects of each factor as each is tested over a wider range of conditions. Also, time and materials are saved in comparison to conducting single factor experiments. Consider testing two methods of cultivation on two wheat cultivars; the single factor approach would require twice the total number of experimental units for the same precision in evaluating the main effects as would be provided by a 2x2 factorial experiment.

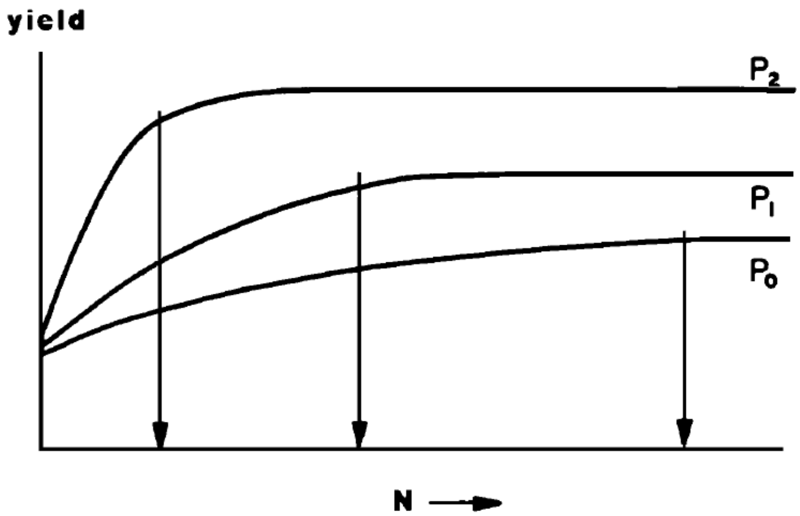


Figure 8-3. The interaction of nitrogen and phosphorus on crop yield.

2x2 factorial experiment

	Cultivar		
Method	A	B	Ave
1	Y_{11}	Y_{12}	1
2	Y_{21}	Y_{22}	2
Ave	1	2	

Two single factor experiments

	Cultivar			Method	
rep	A	B	rep	1	2
1	C_{11}	C_{12}	1	M_{11}	M_{12}
2	C_{21}	C_{22}	2	M_{21}	M_{22}
Ave	1	2	Ave	1	2

Figure 8-4. A hypothetical 2x2 factorial experiment without replication gives the same precision with respect to the averages of cultivars and methods as obtained from two single factor experiments, each with two replications. Eight experimental units are used in the single factor experiments compared to four in the 2x2 factorial experiment.

To illustrate the analysis of a factorial experiment, we consider an experiment designed to study the effect of 5 nitrogen levels on winter wheat irrigated one or two times. The treatments of the experiment are the 10 combinations of 0, 80, 160, 240, and 320 lb N/acre times 1 or 2 irrigations. With 20 plots of a suitable size available for the experiment, the 10 treatments could be completely randomized with 2 replications each, or, as was actually done, the plots can be divided into 2 blocks and the treatments can be assigned randomly within each block. Thus we have an RCBD with 10 treatments and 2 blocks. The field plots along with wheat grain yield are shown in Figure 8-5.

Block I		Block II	
I_1N_0 31.7	I_2N_0 28.9	I_2N_{240} 72.9	I_1N_{80} 48.6
I_2N_{160} 68.5	I_1N_{240} 73.5	I_2N_{160} 66.7	I_2N_{320} 72.3
I_1N_{160} 56.8	I_2N_{80} 61.4	I_1N_{240} 59.4	I_1N_{320} 54.1
I_1N_{80} 50.0	I_2N_{320} 70.6	I_2N_0 40.0	I_2N_{80} 53.1
I_1N_{240} 57.3	I_1N_{320} 53.1	I_1N_{160} 60.6	N_1N_0 39.1
Block Totals			
$Y_{..1} = 561.8$		$Y_{..2} = 566.8$	

Figure 8-5. 10 treatments (5 levels of nitrogen X 2 levels of irrigation) arranged in an RCBD with grain yield data, 100 lbs/acre.

The treatment totals and means are organized in Table 8-9 to facilitate the AOV given in Table 8-10. Calculations are shown below.

$C = 1128.6^2/20 = 63686.90$

$TSS = 31.7^2 + ... + 39.1^2 - C = 66624.56 - C = 2937.66$

$SBB = (561.8^2 + 566.8^2)/10 - C = 63688.15 - C = 1.25$

$STT = (70.8^2 + ... + 142.9^2)/2 - C = 66547.98 - C = 2861.08$

$SSI = (510.7^2 + 617.9^2)/10 - C = 64261.49 - C = 574.59$

$SSN = (149.7^2 + ... + 250.1^2)/4 - C = 65850.02 - C = 2163.12$

`$SSIxN = SST - SSI - SSN$

$= 2861.08 - 574.59 - 2163.12 = 123.37$

$SSE = TSS - SSB - SST$

$= 2937.66 - 1.25 - 2861.08 = 75.33$

Table 8-9. Totals and means in parentheses for the treatments of Figure 8-5.

Irrigation (number)	Nitrogen levels (lb/acre) N_0	N_{80}	N_{160}	N_{240}	N_{320}	Irrigation Total ($Y_{i..}$)	Irrigation Mean ($\bar{Y}_{i..}$)
I_1	70.8 (35.40)	98.6 (49.30)	117.4 (58.70)	116.7 (58.35)	107.2 (53.6)	510.7	51.1
I_2	78.9 (39.45)	114.5 (57.25)	135.2 (67.60)	146.4 (73.20)	142.9 (71.45)	617.9	61.8
N-total	149.7	213.1	252.6	263.1	250.1	1128.6	
N-mean	37.4	53.3	63.2	65.8	62.5		56.4

Table 8-10. AOV of data in Table 8-9.

Source of variation	df	Sum of squares	Mean square	F
Total	19	2937.66		
Block	1	1.25	1.25	<1
Treatment	9	2861.08	317.90	37.98
Irrigation	1	574.59	574.59	68.64
N-level	4	2163.12	540.78	64.61
Interaction	4	123.37	30.84	3.68
Exp. error	9	75.33	8.37	

The design and analysis of this experiment leads to a discussion of several important features:

1. The observed F value for blocks is less than 1, resulting in a non-significant block effect. Blocking in this experiment did not improve the efficiency of the study. A CRD may have been the better choice for the experiment.

2. The new feature of this experiment involves treatments consisting of combinations of 2 factors. Note that the treatment sum of squares is partitioned into a main effect for each factor and their interaction. The degrees of freedom of these components also adds to the total degrees of freedom of treatments.

3. The number of replications for main effects are greater than the replications for the individual treatment. There are 10 replicates for each irrigation treatment and 4 replicates for each N-level. Thus if the interaction is not significant, this study is equivalent to 2 combined experiments -- one, an irrigation study with 10 replications per irrigation treatment and another, a N-fertilizer study with 4 replications per level of application of nitrogen.

4. Note that the interaction was calculated by subtraction,

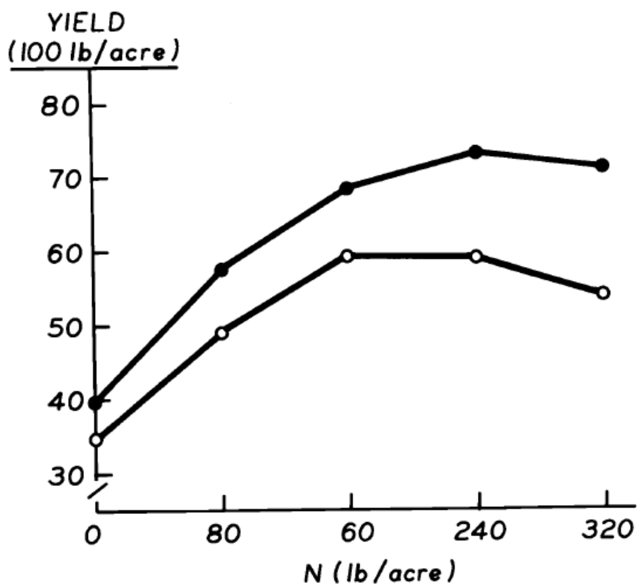
$$SSI \times N = SST - SSI - SSN$$

It can be calculated directly by the formula: r is the number of blocks (or the number of replications per treatment combination in a CRD).

$$SSIxN = r\Sigma\Sigma(\overline{Y}_{ij}.-\overline{Y}_{i}..-\overline{Y}._{j}.+\overline{Y}...)^2$$

5. The observed F for interaction (F = 3.68) is significant at the 5% level ($F_{0.05,4,9} = 3.63$). This indicates that the trend of the yield response for nitrogen-levels depends on the number of irrigations. It is equivalent to conclude that the difference of yields to irrigations depends on the applied nitrogen-level.

Figure 8-6 illustrates the nature of the interaction. It is obvious that 2-irrigations produced higher yields than 1-irrigation for all levels of N applications. However, the rates of yield increase or yield response curves over different levels of N application are different between the two irrigation schemes. For instance, the required N-level for maximum yield is greater for 2 than 1 irrigations.



SUMMARY

1. Blocking is a design technique that is used to remove the known variation among experimental units from the unexplainable random variation. Thus the error mean square can be reduced or the power of detecting treatment differences can be increased.

2. A randomized complete block design is a design which controls one criterion of heterogeneity besides the treatments, e.g., given kn experimental units these can be grouped into n blocks of k units each in such a way that the units are as uniform as possible within a block (experimental units may be classified on the basis of age, weight, general vigor, soil fertility, soil moisture, etc. Any one of these characteristics can be used as a criterion for blocking). The treatments are randomized within each block, the randomization being carried out separately for each block. With k treatments in blocks of size k, each treatment occurs once. This constitutes completeness.

3. The relative efficiency of design-A to design-B is estimated by:

   $$\mathrm{RE}\,(\mathrm{A\ to\ B}) = \frac{(\mathrm{df_A}+1)(\mathrm{df_B}+3)}{(\mathrm{df_B}+1)(\mathrm{df_A}+3)} \bullet \frac{\mathrm{MSE_B}}{\mathrm{MSE_A}}$$

   If RE > 1, design-A is more efficient, If RE < 1, the converse is true.

4. If the effect of one factor depends on the effect of another factor, they are said to interact. Factorial experiments are usually designed to discover interactions between factors which are different types of treatments.

   If there is more than one replication of all treatments in each block, the treatment by block interaction can also be separated from the experimental error term. The term treatment by block interaction means that the differences among treatments change between blocks. Sometimes, blocks may represent soil fertility or soil moisture; other times, they may be the animal's initial body weight, or age, etc. Thus, if a significant treatment by block interaction is found, the interpretation of differences among treatments must be refined within each block. A simple graphic presentation of treatment means for each block will usually clarify the type of interaction.

5. Degrees of freedom and sum of squares of the AOV for a RCBD: One experimental unit per k treatments and per r blocks.

	Without Subsampling		With S Subsamples	
Source	df	SS	df	SS
Total	kr - 1	$\Sigma\Sigma(Y_{ij} - \overline{Y}..)^2$	krs - 1	$\Sigma\Sigma\Sigma(Y_{ijh} - \overline{Y}...)^2$
Block	r - 1	$k\Sigma(\overline{Y}._j - \overline{Y}..)^2$	r - 1	$ks\Sigma(\overline{Y}._j. - \overline{Y}...)^2$
Treatment	k - 1	$r\Sigma(\overline{Y}_i. - \overline{Y}..)^2$	k - 1	$rs\Sigma(\overline{Y}_i.. - \overline{Y}...)^2$
Exp. error	(r - 1)(k - 1)	$\Sigma\Sigma(\overline{Y}_{ij} - \overline{Y}_i. - \overline{Y}._j + \overline{Y}..)^2$	(r - 1)(k - 1)	$s\Sigma\Sigma(\overline{Y}_{ij}. - \overline{Y}.. - \overline{Y}._j. + \overline{Y}...)^2$
Sampling error			kr (s - 1)	$\Sigma\Sigma\Sigma(Y_{ijh} - \overline{Y}_{ij}.)^2$

6. Degrees of freedom sand sum of squares of the AOV for a RCBD: n replications for each of k treatments and r blocks

Without Subsampling			With s Subsamples	
Total	krn - 1	$\Sigma\Sigma\Sigma(Y_{ijh} - \overline{Y}...)^2$	krns - 1	$\Sigma\Sigma\Sigma\Sigma(Y_{ijhn} - \overline{Y}....)^2$
Block	r - 1	$kn\Sigma(\overline{Y}._j. - \overline{Y}...)^2$	r - 1	$kns\Sigma(\overline{Y}._j.. - \overline{Y}....)^2$
Treatment	k - 1	$rn\Sigma(\overline{Y}_i.. - \overline{Y}...)^2$	k - 1	$rns\Sigma(\overline{Y}_i... - \overline{Y}....)^2$
Block x Treatment	(r - 1)(k - 1)	$n\Sigma\Sigma(\overline{Y}_{ij}. - \overline{Y}_i.. - \overline{Y}._j. + \overline{Y}...)^2$	(r - 1)(k - 1)	$ns\Sigma\Sigma(\overline{Y}_{ij}.. - \overline{Y}_i... - \overline{Y}._j.. + \overline{Y}....)^2$
Exp. error	kr (n - 1)	$\Sigma\Sigma\Sigma(Y_{ijh} - \overline{Y}_{ij}.)^2$	kr (n - 1)	$s\Sigma\Sigma\Sigma(\overline{Y}_{ijh}. - \overline{Y}_{ij}..)^2$
Sampling error			krn (s - 1)	$\Sigma\Sigma\Sigma\Sigma(Y_{ijhn} - \overline{Y}_{ijh}.)^2$

EXERCISES

1. Five judges score 4 products on a 10-point scale. The results are shown in the following table. Analyze the data and make an inference whether there is a significant difference between products or judges. (F=2.59 and 12.97)

Judge (or block)	Product (or treatment) A	B	C	D	Total
1	7	10	7	8	32
2	9	10	5	6	30
3	8	8	5	7	28
4	7	8	4	4	23
5	8	9	6	4	27
Total	39	45	27	29	140
Mean	7.8	9.0	5.4	5.8	

2. A plant breeder conducted an experiment to study the inheritance of some agronomic characters in a cross of safflower. Five generations (parents, F_1, F_2 and backcross) and 3 locations were used in the study. The average number of seeds per head from 5 plants are shown in the following table.

   Number of seeds/head of 5 generations of safflower.

Generation/location	1	2	3	Total
parent	33.63	30.25	26.32	90.20
parent	32.39	29.57	28.11	90.07
F_1	35.86	31.32	29.15	96.33
F_2	33.92	31.09	28.86	93.87
Backcross	38.03	35.13	31.52	104.68
Total	173.83	157.36	1243.96	475.15

   Can we conclude that there is a difference between generations or locations?
   (F=22.42 and 83.60)

3. In a randomized complete block experiment, four equally spaced levels of nitrogen were applied to a variety of barley. Blocking was based on the level of soil moisture. The yields (kg/plot) are,

	N-fertilizer			
Block	1	2	3	4
1	4.37	4.50	4.41	4.92
2	6.72	8.80	7.82	8.05
3	8.32	8.73	8.91	9.40
4	8.03	8.31	9.62	9.27

Construct the analysis of variance table. Estimate the means and their standard errors for treatments. Estimate the relative efficiency compared with a CRD.

(F=65.76 and 3.32) (RE=32.42 RCBD to CRD)

4. In a beef-feeding experiment, animals were randomly assigned to each of 4 diets with different percent of protein in 3 feedlots. The average body weight gains (kg/day) for each diet and feedlot are:

	Feedlot		
Diets	1	2	3
1	0.85	0.93	0.79
2	1.03	0.97	1.11
3	0.95	0.99	1.21
4	1.15	1.23	0.92

What is your conclusion on diet effects and feedlot differences? What is the 95% confidence interval of the difference between means of diet 1 and diet 2? (F=1.889 and <1)

5. Suppose that a drug company wishes to test the effects of five new compounds on the growth rate of white rats. It is possible that rats within the same litter may have similar response. Hence blocks are defined as litters. Twenty-five rats within each of 4 litters are chosen at random from a large group of rats, and 5 rats of each litter are placed in one pen to be given one of the five treatments. Body weight gains (g/day/animal) on the pen basis are given in the following table after 3 months of feeding the compound.

	Compound				
Litter	1	2	3	4	5
1	1.45	1.08	1.72	1.04	0.98
2	1.39	1.21	1.45	0.79	1.07
3	0.86	0.99	1.42	1.01	1.32
4	1.04	0.76	0.97	1.05	0.85

Does any difference exist among litters? Do all the compounds have the same effect on body weight gains? What is the 90% confidence interval between means of body weight gain of compounds 1 and 5? What is the 95% confidence interval between means of body weight gains of litter 1 and litter 4? Calculate the relative efficiency of this design compared with a CRD. (F=2.18 and 2.67; L=-0.132, U=0.392;

L=-0.0003 U=0.6403;
RE=1.156 RCBD to CRD)

6. Design an experiment in our own field of specialization to illustrate the technique of a two-way analysis of variance. Make up your own hypothetical data and show the analysis of variance results.

7. To compare the efficiencies of 5 kinds of desk calculators, A, B, C, D, and E, 5 operators were involved in the experiment in which one set of data was analyzed. The efficiency data (seconds) are presented in the following table.

	Calculator				
Operator	A	B	C	D	E
1	58	62	49	61	55
2	72	69	53	62	65
3	43	39	45	38	50
4	85	81	69	75	77
5	69	62	58	71	65

Test the hypothesis that the machines are equally efficient. Construct the analysis of variance table. (F=3.38 and 34.71)

8. Since the homogeneous condition of temperature and moisture in a growth chamber is questioned, 16 dishes of rice seeds were arranged in 4 rows and 4 columns to test the uniformity condition of the chamber. Each dish contained 100 seeds, the number of germinations are recorded as follows.

	Columns			
Rows	1	2	3	4
1	92	85	97	95
2	88	92	95	90
3	99	93	91	95
4	89	87	85	91

Is the environmental condition really homogeneous in the chamber?

(F=1.93 and 0.63)

9. An experiment station conducted a randomized complete block experiment in order to make comparisons among three varieties of barley. Six blocks were used. The yields (bushels per acre) are shown in the following table.

	Variety			
Block	A	B	C	Total
1	45	40	30	115
2	40	42	37	119
3	43	48	35	126
4	45	45	27	117
5	38	42	25	105
6	46	38	26	110
Total	257	255	180	692

Compute the analysis of variance and find 99% confidence intervals for each varietal mean. What is the relative efficiency of this design compared with a CRD?

(F=1.18 and 21.4)

10. Suppose that in an experiment there are 5 blocks and 4 treatments, and that there are 3 replications in each of the treatment-block cells. Suppose that the experiment yields the following results.

$\Sigma\Sigma\Sigma^2_{ijk} = 4110$, $C = Y^2_{...} / (4)(5)(3) = 720$

$\Sigma\Sigma^2_{ij.}/3 = 3150$

$\Sigma Y^2_{j..}/(5)(3) = 1080$

$\Sigma Y^2_{.j.} / (4)(3) = 990$

Construct an analysis of variance table, and test the following null hypotheses at $\alpha = 0.05$.

a) There are no treatment by block interaction effects. (F=6.25)
b) There are no differences in the treatment effects. (F=5.00)
c) There are no differences in the block effects. (F=2.81)

11. Suppose that for a nutrition study on beef cattle five diets are to be studied with Black Angus yearling steers. The measurement to be taken is the average daily gain during a suitable period of time. If 50 such steers are to be grouped into 5 blocks on the basis of initial body weights, and 2 steers within each block assigned at random to each of five specific diets which are of interest to the experimenter, what is the statistical model for the experiment, and the degrees of freedom for each source of variations.

12. Propose an experimental set-up of your planning with possible interaction between two factors, fully describe it, write out the analysis of variance table.

13. Given the following information from an experiment, complete the analysis of variance table, make all possible F tests, stating H_0 each time, and draw all appropriate conclusions:

Source of Variation	df	Sum of Squares
Total	134	9236.1
Treatments	8	640.5
Blocks	4	688.6
Interaction		
Error		3752.8

14. For various reasons, agricultural scientists are concerned with acid deposition from the atmosphere. Thus far no harmful effects of acid rain on domestic animals, crop yields, forest vegetation and soil composition have been found in the United States. Despite this, close monitoring and long-term measurements of acids and acid-forming substances in the rain are still important. Values of pH of precipitation in 1981 for a number of locations were recorded and shown in the following table.

	California		Iowa		Virginia	
	Rural	Urban	Rural	Urban	Rural	Urban
	5.6	5.0	5.5	4.5	4.5	3.9
	5.8	4.9	5.2	4.9	3.9	3.8
	5.9	5.7	4.8	5.3	4.0	4.2
	5.4	5.9	5.0	5.2	4.6	4.6
	6.1	5.5	4.7	4.8	4.3	4.4
Total	28.8	27.0	25.2	24.7	21.3	20.9
Mean	5.76	5.40	5.04	4.94	4.26	4.18

Are there significant differences among the states, the type of locations (rural or urban) and/or interactions between states and types of locations. (F=41.52; 2.17; 0.545)

15. Microcomputers of four manufacturing companies with comparable power and compatible operating systems were compared for their efficiencies in speed in running simulation programs. These four types of machines were tested by four simulation programs and two locations; one in Davis and one in Berkeley. The following data were obtained:

16.

Lab.	Machine	1	2	3	4	Total
Davis	HP	6.8	7.1	8.9	13.3	36.1
	Tandy	5.2	7.4	9.7	11.7	34.0
	IBM	7.1	9.9	12.4	11.9	41.3
	AT&T	4.6	10.6	13.1	13.6	41.9
Berkeley	HP	5.2	6.8	9.1	9.9	31.0
	Tandy	6.2	7.1	10.7	12.2	36.5
	IBM	6.3	10.2	11.4	15.9	43.8
	AT&T	10.9	10.9	13.5	15.2	50.5
Total		52.3	70.0	88.8	103.7	314.8

a) Consider the programs as blocks, and locations and machines as treatment which should be compared. Is there any significant difference among treatments?

(F=6.723)

b) Is there a significant interaction between the locations and the machines?

(F=2.705)

16. The following are the yields in kilograms per plot that resulted when four treatment combinations of nitrogen and phosphate were applied to a grain variety in a randomized complete block experiment. Blocking was based on soil fertility, and each block contained 8 plots.

	Nitrogen Levels							
Block	N_0P_0		N_0P_1		N_1P_0		N_1P_1	
1	4.37	4.31	6.50	5.54	4.41	4.38	3.92	4.86
2	6.72	6.54	8.80	8.75	7.82	7.93	8.05	7.76
3	5.56	5.93	7.79	8.43	8.25	8.99	6.37	7.14

Perform the analysis of variance on the data and test all appropriate hypotheses.

(F=30.68, 0.11, 20.85, 71.07, 138.05, 5.56, 8.09, 2.44, 6.14)

17. The following data are field weights in pounds of corn for 18-hill plots. The treatments were different methods of application of a fertilizer: (1) Check (no fertilizer), (2) 300 lbs. per acre plower under, (3) 300 lbs. per acre broadcast

a) Construct the analysis of variance table. (F=25.42, 1.44, 1.37)

b) Construct the 95% confidence interval of the difference between means of treatment 2 and treatment 1. (L=5.82, U=11.88)

c) Construct the 95% confidence interval of the difference between means of treatment 3 and treatment 1. (L=4.44, U=10.50)

d) What is your interpretation of the results?

	Block					
Treatment	1		2		3	
1	45.1	46.6	51.2	49.3	52.4	44.2
2	56.7	57.3	54.6	55.0	60.1	58.2
3	53.3	55.0	54.7	58.2	55.2	57.2

18. In making bacterial counts on meat pies, samples are homogenized with a diluent and, after blending, are usually allowed to sit for a few minutes to permit foam to subside. In a study to determine whether the time after homogenization has any effect on the bacterial count, each of the 5 samples were subdivided into 8 subsamples and were randomly assigned to 4 times with 2 counts per time assay. The following data are the log counts.

a) Construct the analysis of variance table. (F=1.72, 0.23, 2.28)

b) Is there any reason to believe time after blending has an effect on count?
c) Perform a t-test to compare the means of one minute and 8 minute counts. (t=2.14, df=12)

	Time After Blending in Minutes			
Sample	1	2	3	4
1	3.73	3.72	3.59	2.78
	3.65	3.59	3.52	3.00
2	3.34	3.78	3.38	3.24
	3.20	3.10	3.36	3.18
3	3.76	3.28	3.14	3.11
	3.56	2.91	3.44	3.38
4	3.65	3.51	3.04	3.32
	3.35	3.36	2.95	3.69
5	3.72	3.58	3.42	3.19
	3.20	3.07	3.25	3.19

19. Young growth and old growth leaves of 4 varieties were chosen for ascorbic acid content comparisons. Three determinations for ascorbic acid were made on each leaf. Construct the analysis of variance table showing the partitioning of the degrees of freedom and sum of squares. (F=14.36, 8.66 <1, 30.24)

	Variety

Growth	Leaf	Det.	1	2	3	4
Young	1	1	9.23	12.73	10.78	11.15
		2	9.10	11.67	10.90	10.96
		3	9.18	12.58	10.69	10.73
	2	1	8.69	11.79	12.73	9.45
		2	8.95	12.03	12.53	9.73
		3	9.06	11.87	12.71	9.22
Old	1	1	8.31	10.75	9.60	10.10
		2	8.09	10.67	9.75	10.15
		3	8.13	10.12	9.72	10.23
	2	1	7.38	11.33	10.77	10.06
		2	7.25	11.46	10.21	9.78
		3	7.29	11.74	10.36	9.69

20. In a feeding trial, 2 breeds each of 4 horses were randomly assigned to each of 12 lots. Three rations containing different percentages of corn were randomly assigned to the lots so that each ration was used in 4 lots. Body weight gained over a 26-month period are given below.

	RATION					
	15% corn		30% corn		45% corn	
Breed	Lot 1	Lot 2	Lot 3	Lot 4	Lot 5	Lot 6
A	20	19	23	29	32	31
	18	23	28	22	29	25
	22	25	25	27	30	28
	24	26	30	25	28	33
	Lot 7	Lot 8	Lot 9	Lot 10	Lot 11	Lot 12
B	23	23	29	32	39	37
	21	28	33	37	38	35
	19	21	35	31	35	41
	25	18	30	34	37	44

Construct the analysis of variance table. Is there any significant difference between rations or breeds? (F=126.37, 71.63, 18.19, <1)