Chapter 27

Randomized Block Designs

In addition to completely randomized designs discussed in previous chapters, a second kind of design is introduced and described in this chapter: *randomized block designs.*

27.1 Elements of Randomized Block Designs

A randomized block design is a restricted randomized design, in which experimental units are first organized into homogeneous blocks and then the treatments are assigned at random to these units within these blocks. The main advantage of this design is, if done properly, it provides more precise results. The main disadvantage is more assumptions are required (no interactions between treatments and blocks and constant variance from block to block) and there are fewer degrees of freedom than for a completely randomized design.

27.2 Model for Randomized Block Designs

SAS program: att9-27-2-prozac-RBD-treatmeans

The randomized block design model is

$$Y_{ij} = \mu_{..} + \rho_i + \tau_j + \varepsilon_{ij},$$

where ρ_i is the block effects, τ_j is the treatment effect and ε_{ij} is the error[1].

Exercise 27.1 (Treatment Plot: Prozac Yield)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study[2].

[1]Notice, the randomized block design is really just a two–factor ANOVA model where there is one case per treatment.

[2]This study is very similar to the prozac previous study done in chapter 21, only the data here is different from what it was before.

	block, blend →	1	2	3	4	5
treatment,	A	82	84	85	87	90
catalyst	B	88	90	92	95	97
	C	92	96	96	99	101
	D	94	97	100	102	103

From SAS, the five plots are more or less parallel, which indicates (*choose none, one or more*!)

1. the yield increases from catalyst A to catalyst D
2. there is no/little treatment–block interaction

which is "good" because then the analysis involves just the determining the influence of the treatments alone and not the influence of the treatments, blocks and treatment–block interaction.

27.3 Analysis of Variance and Tests

SAS program: att9-27-3-prozac-RBD-inference

The following ANOVA table is used to make inferences on this model,

Source	Degrees of Freedom, *df*	Sum Of Squares, *SS*	Mean Squares, *MS*
Blocks	$n-1$	$SSBL$	$MSBL = \frac{SSBL}{n-1}$
Treatments	$r-1$	$SSTR$	$MSTR = \frac{SSTR}{r-1}$
Error	$(n-1)(r-1)$	$SSBL.TR$	$MSBL.TR = \frac{SSBL.TR}{(n-1)(r-1)}$
Total	$nr-1$	$SSTO$	

Exercise 27.2 (Analysis of Variance and Tests: Prozac Yield)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5
treatment,	A	82	84	85	87	90
catalyst	B	88	90	92	95	97
	C	92	96	96	99	101
	D	94	97	100	102	103

Source	*df*	*SS*	*MS*
Blend (Blocks)	4	186	46.5
Catalyst	3	535	178.33
Error	12	6	0.50
Total	19	727	

1. *Test Treatment, Catalyst*
   Test if the catalyst effect is significant at $\alpha = 0.05$.

   (a) *Statement*
   The statement of the test is (check none, one or more):

   i. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a :\ \alpha_1 \neq \alpha_2, \alpha_1 = \alpha_3$.
   ii. $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus $H_a :\ \alpha_1 \neq \alpha_3, \alpha_1 \neq \alpha_2$.
   iii. $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus
   $H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3, 4$.
   iv. $H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \mu_{3\cdot} = \mu_{4\cdot}$ versus
   $H_a :$ at least one $\mu_{i\cdot} \neq \mu_{j\cdot},\ i, j = 1, 2, 3, 4$.

   (b) *Test*
   Since the test statistic is
   $F = \frac{MSTR}{MSBL.TR} = \frac{178.33}{0.5} = 356.67$,
   and the p–value, with $r - 1 = 4 - 1 = 3$ and $nr = 5(4) = 20$ degrees of freedom, is given by

   $$\text{p–value} = P(F \geq 356.67)$$

   which equals (circle one) **0.00** / **0.34** / **0.43**.
   (Use 2nd DISTR 9:Fcdf(356.67,E99,4,20).)
   The level of significance is 0.05.

   (c) *Conclusion*
   Since the p–value, 0.0, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average Prozac yields for the four catalysts are the same.

2. *Test block, blend; does the blocking work?*
   Test if the blend block is significant at $\alpha = 0.05$.

   (a) *Statement*
   The statement of the test is (check none, one or more):

   i. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a :\ \alpha_1 \neq \alpha_2, \alpha_1 = \alpha_3$.
   ii. $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus $H_a :\ \alpha_1 \neq \alpha_3, \alpha_1 \neq \alpha_2$.
   iii. $H_0 : \rho_i = 0$ versus $H_a :$ at least one $\rho_i \neq 0,\ i = 1, 2, 3, 4, 5$.
   iv. $H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \mu_{3\cdot} = \mu_{4\cdot}$ versus
   $H_a :$ at least one $\mu_{i\cdot} \neq \mu_{j\cdot},\ i, j = 1, 2, 3, 4$.

   (b) *Test*
   Since the test statistic is
   $F = \frac{MSBL}{MSBL.TR} = \frac{46.5}{0.5} = 93$,

and the p–value, with $n - 1 = 5 - 1 = 4$ and $nr = 5(4) = 20$ degrees of freedom, is given by

$$\text{p–value} = P(F \geq 93)$$

which equals (circle one) **0.00** / **0.34** / **0.43**.
The level of significance is 0.05.

(c) *Conclusion*
Since the p–value, 0.0, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average Prozac yields for the blends are the same; that is, the blend blocking variable is effective.

3. *Why block?*
True / **False**
It was felt (based on past studies possibly) that the prozac yield may be influenced by blend and so confound with the effect of the different catalysts on prozac yield. Blocking by blend reduced this possible confounding effect.

4. *Random versus fixed treatment*
The statement of the test for the treatment, when it is random, is (choose one)

(a) $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a :$ $\alpha_1 \neq \alpha_2, \alpha_1 = \alpha_3$.

(b) $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus $H_a :$ $\alpha_1 \neq \alpha_3, \alpha_1 \neq \alpha_2$.

(c) $H_0 :$ $\sigma_\tau^2 = 0$ versus $H_a :$ $\sigma_\tau^2 > 0$.

(d) $H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \mu_{3\cdot} = \mu_{4\cdot}$ versus
$H_a :$ at least one $\mu_{i\cdot} \neq \mu_{j\cdot}$, $i, j = 1, 2, 3, 4$.

Aside from the different test statements, the test for a random treatment proceeds in the same way as for the test of a fixed treatment. This is because, since there is only one observation per cell in a randomized block design, the error sums of squares is replaced by the interaction (between block and treatment) sums of squares and so the F test statistics must be the same in either case.

27.4 Evaluation of Appropriateness of Randomized Block Model

SAS program: att9-27-4-prozac-RBD-residuals

Exercise 27.3 (Evaluation of Appropriateness of Randomized Block Model) Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5	
treatment,	A	82	84	85	87	90	$\overline{Y}_{1\cdot} = 85.6$
catalyst	B	88	90	92	95	97	$\overline{Y}_{2\cdot} = 92.4$
	C	92	96	96	99	101	$\overline{Y}_{3\cdot} = 96.8$
	D	94	97	100	102	103	$\overline{Y}_{4\cdot} = 98.8$
		$\overline{Y}_{\cdot 1} = 89$	$\overline{Y}_{\cdot 2} = 91.75$	$\overline{Y}_{\cdot 3} = 93.25$	$\overline{Y}_{\cdot 4} = 95.75$	$\overline{Y}_{\cdot 5} = 97.75$	$\overline{Y}_{\cdot\cdot} = 93.5$

1. *Residuals versus fitted*
   From SAS, the residual plot indicates
   (choose one) **constant** / **nonconstant** variance.

2. *Normal probability plot of residuals*
   From SAS, the normal probability plot indicates
   (choose one) **normal** / **non–normal** residuals.

3. *Tukey test for additivity*[3]

   (a) *Preliminary analysis*
   Since

$$\begin{aligned} SSBL.TR^* &= \frac{\left(\sum_i \sum_j (\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})(\overline{Y}_{\cdot j} - \overline{Y}_{\cdot\cdot}) Y_{ij}\right)^2}{\sum_i (\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})^2 \sum_j (\overline{Y}_{\cdot j} - \overline{Y}_{\cdot\cdot})^2} \\ &= \frac{[(85.6 - 93.5)(89 - 93.5)82 + (85.6 - 93.5)(91.75 - 93.5)84 + \cdots + (99.2 - 93.5)(97.75 - 93.5)103]^2}{[(85.6 - 93.5)^2 + \cdots + (99.2 - 93.5)^2][(89 - 93.5)^2 + \cdots + (97.75 - 93.5)^2]} \\ &= \frac{55.6}{(107)(46.5)} \\ &\approx 0.0111747 \end{aligned}$$

   and, using SAS for *SSTO*, *SSTR* and *SSBL*

$$\begin{aligned} SSRem^* &= SSTO - SSTR - SSBL - SSBL.TR^* \\ &= 727 - 535 - 186 - 0.011 \\ &= \end{aligned}$$

   (choose one) $\mathbf{-1.334}$ / $\mathbf{3.454}$ / $\mathbf{5.989}$

   (b) *Test for interaction*

      i. *Statement*
      The statement of the test is (check none, one or more):

      A. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a :$ $\alpha_1 \neq \alpha_2, \alpha_1 = \alpha_3$.

      B. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a :$ $\alpha_1 \neq \alpha_3, \alpha_1 \neq \alpha_2$.

      C. $H_0 : D = 0$ versus $H_a :$ $D \neq 0$

      D. $H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \mu_{3\cdot} = \mu_{4\cdot}$ versus
      $H_a :$ at least one $\mu_{i\cdot} \neq \mu_{j\cdot}$, $i, j = 1, 2, 3, 4$.

[3]We did this test for similar data (same factor names, but different data) previously in chapter 21.

ii. *Test*
the test statistic is

$$F^* = \frac{SSBL.TR^*}{1} \div \frac{SSRem^*}{nr-r-n} = \frac{0.0111747}{1} \div \frac{5.989}{(4)(5)-4-5} =$$

(choose one) **0.0105** / **0.0205** / **0.0305** and the critical value is and $F(1-\alpha; 1, nr-r-n) = F(1-0.05; 1, 11) = 4.844$

iii. *Conclusion*
Since the test statistic, 0.0205, is smaller than the critical value, 4.844, we (circle one) **accept** / **reject** the null hypothesis that *no* interactions are present (which seems to confirm our treatments plot above)

27.5 Analysis of Treatment Effects

SAS program: att9-27-5-prozac-RBD-effects

Detailed analysis of treatment effects, using, for example, Tukey, Scheffe and Bonferroni, is possible, just like it was for the completely randomized design.

Exercise 27.4 (Analysis of Treatment Effects: Prozac Yield)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5	
treatment,	A	82	84	85	87	90	$\overline{Y}_{1\cdot} = 85.6$
catalyst	B	88	90	92	95	97	$\overline{Y}_{2\cdot} = 92.4$
	C	92	96	96	99	101	$\overline{Y}_{3\cdot} = 96.8$
	D	94	97	100	102	103	$\overline{Y}_{4\cdot} = 98.8$
		$\overline{Y}_{\cdot 1} = 89$	$\overline{Y}_{\cdot 2} = 91.75$	$\overline{Y}_{\cdot 3} = 93.25$	$\overline{Y}_{\cdot 4} = 95.75$	$\overline{Y}_{\cdot 5} = 97.75$	$\overline{Y}_{\cdot\cdot} = 93.5$

Source	*df*	*SS*	*MS*
Blend (Blocks)	4	186	46.5
Catalyst	3	535	178.33
Error	12	6	0.50
Total	19	727	

Estimate $L_1 = \mu_{\cdot 1} - \mu_{\cdot 2}$ and $L_2 = \mu_{\cdot 2} - \mu_{\cdot 4}$ using the Bonferroni procedure with a 95 percent family confidence coefficient.

1. $L_1 = \mu_{\cdot 1} - \mu_{\cdot 2}$.
From SAS and the ANOVA table above,
$\hat{L}_1 = \overline{Y}_{\cdot 1} - \overline{Y}_{\cdot 2} = 85.6 - 92.4 = -6.8$,

$s\{\hat{L}_i\} = \sqrt{MSBL.TR\left(\frac{1}{n}+\frac{1}{n}\right)} = \sqrt{\frac{2(0.5)}{5}} \approx 0.447$
$B = t(1-\frac{\alpha}{2g}; n_T - r) = t(1-\frac{0.05}{2(2)}; 12) = t(0.9875; 12) \approx 2.5600$
and so the CIs are
$-6.8 \pm 2.56(0.447) \approx$ (choose one)
$\mathbf{-8.04 \le L_1 \le -5.76}$ / $\mathbf{-7.94 \le L_1 \le -6.66}$ / $\mathbf{-7.94 \le L_1 \le -5.66}$

2. $L_2 = \mu_{.2} - \mu_{.4}$.
From SAS and the ANOVA table above,
$\hat{L}_2 = \overline{Y}_{.2} - \overline{Y}_{.4} = 92.4 - 98.8 = -6.4$,
$s\{\hat{L}_i\} = \sqrt{MSBL.TR\left(\frac{1}{n}+\frac{1}{n}\right)} = \sqrt{\frac{2(0.5)}{5}} \approx 0.447$
$B = t(1-\frac{\alpha}{2g}; n_T - r) = t(1-\frac{0.05}{2(2)}; 8) = t(0.9875; 12) \approx 2.56$
and so the CIs are
$-6.4 \pm 2.56(0.447) =$ (choose one)
$\mathbf{-8.94 \le L_2 \le -5.66}$ / $\mathbf{-7.544 \le L_2 \le -5.256}$ / $\mathbf{-7.94 \le L_1 \le -5.66}$

27.6 Factorial Treatments

The randomized block design model for a two–factor design is

$$Y_{ijk} = \mu_{..} + \rho_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk},$$

where ρ_i is the block effects, α_i is factor A, β_j is factor B, $(\alpha\beta)_{jk}$ is the interaction and ε_{ij} is the error[4].

Source	Degrees of Freedom, *df*	Sum Of Squares, *SS*	Mean Squares, *MS*
Blocks	$n-1$	$SSBL$	$MSBL = \frac{SSBL}{n-1}$
Treatments	$r-1$	$SSTR$	$MSTR = \frac{SSTR}{r-1}$
Factor A	$a-1$	SSA	$MSA = \frac{SSA}{a-1}$
Factor B	$b-1$	SSB	$MSB = \frac{SSB}{b-1}$
Interaction AB	$(a-1)(b-1)$	$SSAB$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$
Error	$(n-1)(r-1)$	$SSBL.TR$	$MSBL.TR = \frac{SSBL.TR}{(n-1)(r-1)}$
Total	$nr-1$	$SSTO$	

27.7 Planning Randomized Block Experiments

SAS program: att9-27-7-prozac-RBD-RBDvsCRD

An appropriate sample size, number of blocks and the power of a test are investigated for a randomized block design study.

Exercise 27.5 (Planning Randomized Block Experiments: Prozac Yield) Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

[4]The RBD design can be used in studies with a greater number of factors, such as a three factor design, for example.

	block, blend →	1	2	3	4	5	
treatment,	A	82	84	85	87	90	$\overline{Y}_{1\cdot} = 85.6$
catalyst	B	88	90	92	95	97	$\overline{Y}_{2\cdot} = 92.4$
	C	92	96	96	99	101	$\overline{Y}_{3\cdot} = 96.8$
	D	94	97	100	102	103	$\overline{Y}_{4\cdot} = 99.2$

Source	*df*	*SS*	*MS*
Blend (Blocks)	4	186	46.5
Catalyst	3	535	178.33
Error	12	6	0.50
Total	19	727	

1. *Determine power, $1-\beta$, given sample sizes ....*
   Determine the power of a test which tests if at least two of the four average prozac yields to the catalysts are different at $\alpha = 0.05$, where[5] $\mu_{\cdot 1} = 85$, $\mu_{\cdot 2} = 92$, $\mu_{\cdot 3} = 97$ and $\mu_{\cdot 4} = 99$, and $\sigma = 0.5$ and where $n_i = 5$, $i = 1, 2, 3, 4$.

   In this case, $\mu_{\cdot} = \frac{\sum n_i \mu_i}{n_T} = \frac{5(85)+5(92)+5(97)+5(99)}{20} = 93.25$,
   and so

$$\begin{aligned}\phi &= \frac{1}{\sigma}\sqrt{\frac{\sum n_i(\mu_i - \mu_{\cdot})^2}{r}} \\ &= \frac{1}{5}\sqrt{\frac{5(85-93.25)^2 + 5(92-93.25)^2 + 5(97-93.25)^2 + 5(99-93.25)^2}{4}} \\ &\approx 2.42 \approx 2.5\end{aligned}$$

   $\nu_1 = r - 1 = 4 - 1 = 3$
   $\nu_2 = n_T - r = 20 - 4 = 14$
   and so, using Table B.11 (1356), $1-\beta =$ (choose one)
   0.97 / 0.98 / 0.99

2. *One method for determining the number of blocks, using CI width.*
   What would be the required number of blocks, n, if it is desired to make all pairwise 95% (so $\alpha = 0.05$) CI widths ± 4, where $\sigma = 5$ and $r = 4$?

   This is an *iterative* procedure. Using block size $n = 10$ as a *starting* point,

$$\sigma\{\hat{L}\} = \sqrt{\sigma^2\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{(5)^2\left(\frac{1}{10} + \frac{1}{10}\right)} = 2.236$$

[5]Notice that the *guessed* means, $\mu_{\cdot 1} = 85$, $\mu_{\cdot 2} = 92$, $\mu_{\cdot 3} = 97$ and $\mu_{\cdot 4} = 99$ are *not* the same as the *observed* means, $\hat{Y}_{\cdot 1} = 85.6$, $\hat{Y}_{\cdot 2} = 92.4$, $\hat{Y}_{\cdot 3} = 96.8$ and $\hat{Y}_{\cdot 4} = 99.2$.

and so since $\alpha = 0.05$, $r = 4$, for Tukey (using Tables B.9 for q)

$$\begin{aligned} T &= \frac{1}{\sqrt{2}} q(1-\alpha; r, (n-1)(r-1)) \\ &= \frac{1}{\sqrt{2}} q(0.95; 4, (10-1)(4-1)) \\ &= \frac{1}{\sqrt{2}} q(0.95; 4, 27) \\ &\approx 3.85 \end{aligned}$$

In a similar way, trying $n = 10, 15, 20$,

contrast	$n = 10$	$n = 15$	$n = 20$
$\mu_i - \mu_j$	2.236	1.826	1.581
width of CI, $\pm T\sigma\{\hat{L}\}$	$\pm(3.85)(2.236) = \pm 8.61$	± 4.08	± 3.54

In other words, the width of the CI is closest to ± 4 when the number of blocks is $n =$ (choose one) **10** / **15** / **20**

3. *Another method for determining the number of blocks, using* Δ.
What is the number of blocks if the maximum difference between mean Prozac yields for the four catalysts is $\Delta = 10$ where $1 - \beta = 0.95$, $\sigma = 5$ and $\alpha = 0.05$?

   Since $r = 4$, $\alpha = 0.05$, $1 - \beta = 0.95$, $\alpha = 0.05$, $\sigma = 5$ and $\Delta = 10$,
   then $\frac{\Delta}{\sigma} = \frac{10}{5} = 2$
   and so, using Table B.12,
   $n =$ (choose one) **10** / **15** / **20**

4. *Comparing the two calculation of the number of blocks methods*
The two methods for determining an appropriate number of blocks in a randomized block design
(choose one) **always give** / **do not necessarily give** the same answer.

5. *Efficiency: comparing randomized block design to completely randomized design*
Use the efficiency measure to determine how effective the blocking variable is when comparing the currently analyzed randomized block design to the previously analyzed completely randomized design.

   Since
   $df_1 = 16$ for the experimental error in base design (completely randomized design), 16 in this case
   $df_2 = 12$ for the experimental error in assessed design (randomized block

design);

$$\begin{aligned}\hat{E}' &= \frac{(df_2+1)(df_1+3)}{(df_2+3)(df_1+1)} \times \frac{(n-1)MSBL + n(r-1)MSBL.TR}{(nr-1)MSBL.TR} \\ &= \frac{(12+1)(16+3)}{(12+3)(16+1)} \times \frac{(5-1)(46.5)+5(4-1)(0.5)}{(5(4)-1)(0.5)} =\end{aligned}$$

(choose one) **10.54** / **15.53** / **19.72**
We need about 20 times as many replications per treatment with a completely randomized design to achieve the same variance of any estimated contrast as is obtained with blocking by blend. In other words, blocking by blend is very effective in this case.

27.8 Regression Approach to Randomized Block Designs

SAS program: att9-27-8-prozac-RBD-regression

Exercise 27.6 (Regression Approach to Randomized Block Designs)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5
treatment,	A	82	84	85	87	90
catalyst	B	88	90	92	95	97
	C	92	96	96	99	101
	D	94	97	100	102	103

1. *(Full) regression model, with treatments*
From SAS,

$$\begin{aligned}Y_{ij} &= \mu_{..} + \rho_1 X_{ij1} + \rho_2 X_{ij2} + \rho_3 X_{ij3} + \rho_4 X_{ij4} \\ &\quad + \tau_1 X_{ij5} + \tau_2 X_{ij6} + \tau_3 X_{ij7} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.5 - 4.5 X_{ij1} - 1.75 X_{ij2} - 0.25 X_{ij3} + 2.25 X_{ij4} \\ &\quad - 7.9 X_{ij5} + \hat{\tau}_2 X_{ij6} + 3.3 X_{ij7}\end{aligned}$$

where $\hat{\tau}_2 =$ (choose one) **−1.1** / **1.5** / **2.2**
and where

$$X_{ij1} = \begin{cases} 1, & \text{if case from block 1} \\ -1, & \text{if case from block 5} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij2} = \begin{cases} 1, & \text{if case from block 2} \\ -1, & \text{if case from block 5} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij3} = \begin{cases} 1, & \text{if case from block 3} \\ -1, & \text{if case from block 5} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij4} = \begin{cases} 1, & \text{if case from block 4} \\ -1, & \text{if case from block 5} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$X_{ij5} = \begin{cases} 1, & \text{if case from treatment 1} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

$$X_{ij6} = \begin{cases} 1, & \text{if case from treatment 2} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

$$X_{ij7} = \begin{cases} 1, & \text{if case from treatment 3} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

2. *(Reduced) regression model, without treatments*
From SAS,

$$\begin{aligned} Y_{ij} &= \mu_{..} + \rho_1 X_{ij1} + \rho_2 X_{ij2} + \rho_3 X_{ij3} + \rho_4 X_{ij4} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.5 + \hat{\rho}_1 X_{ij1} - 1.75 X_{ij2} - 0.25 X_{ij3} + 2.25 X_{ij4} \end{aligned}$$

where $\hat{\rho}_1 =$ (choose one) $\mathbf{-4.5}$ / $\mathbf{1.5}$ / $\mathbf{2.2}$

3. *Test of treatment effects, using regression approach*

   (a) *Statement*
   The statement of the test is (check none, one or more):

   i. $H_0 : \tau_i = 0$ versus $H_a :$ at least one $\tau_i \neq 0$, $i = 1, 2, 3, 4$.
   ii. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ versus $H_a : \alpha_1 \neq \alpha_3, \alpha_1 \neq \alpha_2$.
   iii. $H_0 : D = 0$ versus $H_a : D \neq 0$
   iv. $H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.} = \mu_{4.}$ versus
   $H_a :$ at least one $\mu_{i.} \neq \mu_{j.}$, $i, j = 1, 2, 3, 4$.

(b) *Test*
The test statistic is

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{541 - 6}{15 - 12} \div \frac{6}{12} =$$

(choose one) **44.6** / **356.7** / **446.7**
and the critical value is
and $F(1-\alpha; 1, nr - r - n) = F(1 - 0.05; 1, 5(4) - 4 - 5) = 4.844$

(c) *Conclusion*
Since the test statistic, 356.7, is larger than the critical value, 4.844, we (circle one) **accept** / **reject** the null hypothesis that there is *no* treatment effect (which, of course, confirms the test of the treatments above).

27.9 Covariance Analysis for Randomized Block Designs

SAS program: att9-27-9-prozac-RBD-covariance

Covariance analysis can sometimes be used to reduce the error variance even further than it already has been by using blocks.

Exercise 27.7 (Covariance Analysis for Randomized Block Designs)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5	
treatment,	A	82	84	85	87	90	$\overline{Y}_{1\cdot} = 85.6$
catalyst	B	88	90	92	95	97	$\overline{Y}_{2\cdot} = 92.4$
	C	92	96	96	99	101	$\overline{Y}_{3\cdot} = 96.8$
	D	94	97	100	102	103	$\overline{Y}_{4\cdot} = 98.8$

In addition, consider the following concomitant variable, *time–rank of testing*

	block, blend →	1	2	3	4	5	
treatment,	A	1	2	17	7	6	$\overline{Y}_{1\cdot} = 85.6$
catalyst	B	15	5	16	12	10	$\overline{Y}_{2\cdot} = 92.4$
	C	13	20	8	11	14	$\overline{Y}_{3\cdot} = 96.8$
	D	9	19	3	4	18	$\overline{Y}_{4\cdot} = 98.8$

For example, catalyst A with blend 2 was the 2nd batch of rice liquor tested; catalyst C with blend 4 was the 11th batch of rice liquor tested.

1. *(Full) covariance model, with treatments (and concomitant variable)*
From SAS,

$$\begin{aligned} Y_{ij} &= \mu_{..} + \rho_1 I_{ij1} + \rho_2 I_{ij2} + \rho_3 I_{ij3} + \rho_4 I_{ij4} \\ &\quad + \tau_1 I_{ij5} + \tau_2 I_{ij6} + \tau_3 I_{ij7} + \gamma x_{ij} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.5 - 4.5281 X_{ij1} - 1.7219 X_{ij2} - 0.23595 X_{ij3} + 2.19375 X_{ij4} \\ &\quad - 8.0096 X_{ij5} + \hat{\tau}_2 X_{ij6} + 3.3759 X_{ij7} - 0.0281 x_{ij} \end{aligned}$$

where $\hat{\tau}_2 =$ (choose one) $\mathbf{-1.1}$ / $\mathbf{-1.069}$ / $\mathbf{2.2}$
and where

$$I_{ij1} = \begin{cases} 1, & \text{if case from block 1} \\ -1, & \text{if case from block 5} \\ 0, & \text{otherwise,} \end{cases}$$

and X_{ij2}, X_{ij3}, X_{ij4} are defined in a similar way, and

$$I_{ij5} = \begin{cases} 1, & \text{if case from treatment 1} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

$$I_{ij6} = \begin{cases} 1, & \text{if case from treatment 2} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

$$I_{ij7} = \begin{cases} 1, & \text{if case from treatment 3} \\ -1, & \text{if case from treatment 4} \\ 0, & \text{otherwise.} \end{cases}$$

and $x_{ij} = X_{ij} - X_{...}$

2. *(Reduced) regression model, without treatments (but with concomitant variable)*
From SAS,

$$\begin{aligned} Y_{ij} &= \mu_{..} + \rho_1 X_{ij1} + \rho_2 X_{ij2} + \rho_3 X_{ij3} + \rho_4 X_{ij4} + \gamma x_{ij} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.5 + \hat{\rho}_1 X_{ij1} - 2.03685 X_{ij2} - 0.39342 X_{ij3} + 2.82369 X_{ij4} + 0.28658 x_{ij} \end{aligned}$$

where $\hat{\rho}_1 =$ (choose one) $\mathbf{-4.213}$ / $\mathbf{-4.345}$ / $\mathbf{-4.567}$

3. *Test of Treatment Effect.*
$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus
$H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3, 4.$

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{489.08 - 5.595}{14 - 11} \div \frac{5.595}{11} = 316.85$$

and so p–value is $P(F > 316.85; 3, 11) \approx$ (choose one) $\mathbf{0}$ / $\mathbf{0.01}$ / $\mathbf{0.07}$
since p–value $= 0 < \alpha = 0.05$
reject null; that is, mean Prozac yields are different for different catalysts

4. *Reduction in Error Variance?*
   for the ANOVA model, $MSE = 0.500$
   for the ANCOVA model,
   $MSE =$ (choose one) **0.0345** / **0.1488** / **0.5087**
   The error variance was (choose one) **increased** / **reduced** by introducing the time–rank concomitant variable. In other words, the ANCOVA model is a better one than the ANOVA model.

27.10 Nonparametric Rank F Test

Not covered.

27.11 Missing Observations

SAS program: att9-27-11-prozac-RBD-missing

The regression approach can be used to deal with missing observations.

Exercise 27.8 (Missing Observations: Prozac Yield)
Consider the effect of different catalysts on the Prozac yields from 10 milliliters of rice liquor. Use five different rice liquor blends as blocks for this study.

	block, blend →	1	2	3	4	5
treatment,	A	82	missing	85	87	90
catalyst	B	88	90	92	95	97
	C	92	96	96	missing	101
	D	94	97	100	102	103

Notice that a couple of the observations are missing in the table above and, more than this, notice that we do *not* have information on the concomitant variable and so will *not* be performing an ANCOVA analysis.

1. *(Full) regression model, with treatments, but missing observations*
   From SAS,

$$\begin{aligned} Y_{ij} &= \mu_{..} + \rho_1 X_{ij1} + \rho_2 X_{ij2} + \rho_3 X_{ij3} + \rho_4 X_{ij4} \\ &\quad + \tau_1 X_{ij5} + \tau_2 X_{ij6} + \tau_3 X_{ij7} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.4923 - 4.4923 X_{ij1} - 1.807 X_{ij2} - 0.2422 X_{ij3} + 2.283 X_{ij4} \\ &\quad - 7.944 X_{ij5} + \hat{\tau}_2 X_{ij6} + 3.328 X_{ij7} \end{aligned}$$

   where $\hat{\tau}_2 =$ (choose one) **−1.09** / **1.05** / **2.02**

2. *(Reduced) regression model, without treatments, and missing observations*
From SAS,

$$\begin{aligned} Y_{ij} &= \mu_{..} + \rho_1 X_{ij1} + \rho_2 X_{ij2} + \rho_3 X_{ij3} + \rho_4 X_{ij4} + \varepsilon_{ij} \\ \hat{Y}_{ij} &= 93.8 + \hat{\rho}_1 X_{ij1} + 0.533 X_{ij2} - 0.55 X_{ij3} + 0.866 X_{ij4} \end{aligned}$$

where $\hat{\rho}_1 =$ (choose one) $\mathbf{-4.8}$ / $\mathbf{-4.5}$ / $\mathbf{-2.2}$

3. *Test of Treatment Effect.*
$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ versus
$H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3, 4.$
The test statistic is

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{446.833 - 5.95594}{13 - 10} \div \frac{5.95594}{10} = 246.74$$

and so p–value is $P(F > 246.74; 3, 10) \approx$ (choose one) **0** / **0.01** / **0.07**
since p–value $= 0 < \alpha = 0.05$
reject null; that is, mean Prozac yields are different for different catalysts

4. *Effect of Missing Observations?*
The missing observations (choose one) **changed** / **did not change** the results of the test on the treatment effects.

27.12 Random Block Effects

Not covered.

27.13 Generalized Randomized Block Designs

Not covered.

27.14 Use of More Than One Blocking Variable

Not covered.