Section 6.3 Randomized Complete Block Design (RCBD)

Examples of RCBD

- A chemist is studying the reaction rate of five chemical agents. Only five agents can be analyzed effectively per day. Since day-to-day differences may affect the reaction rate, each day is used as a block. **All five chemical agents are tested each day in independently random orders.**
- Blocking is to arrange experimental units in groups that are similar to each other.
- Blocking factor normally is not of primary interest to the investigator, mainly to remove nuisance effects.
- Blocking factor is an observational factor, related to the characteristic of the experimental unit, or the experimental settings. Not an experimental factor.

TABLE 7.3-1 **Compressive Strength of Concrete (100 pounds per square inch)**

Treatment	Batch 1	2	3	4	5	Treatment Mean
A	52	47	44	51	42	47.2
B	60	55	49	52	43	51.8
C	56	48	45	44	38	46.2
Batch mean	56	50	46	49	41	48.4

TABLE 7.3-2 **Data from a Randomized Complete Block Experiment**

Treatment	Block 1	2	$\cdots$	j	$\cdots$	b	Treatment Mean
1	Y_{11}	Y_{12}		$\cdots$		Y_{1b}	$\bar{Y}_{1\cdot}$
2	Y_{21}	Y_{22}		$\cdots$		Y_{2b}	$\bar{Y}_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
i	$\cdots$	$\cdots$	$\cdots$	Y_{ij}	$\cdots$	$\cdots$	$\bar{Y}_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
k	Y_{k1}	Y_{k2}		$\cdots$		Y_{kb}	$\bar{Y}_{k\cdot}$
Block mean	$\bar{Y}_{\cdot 1}$	$\bar{Y}_{\cdot 2}$	$\cdots$	$\bar{Y}_{\cdot j}$	$\cdots$	$\bar{Y}_{\cdot b}$	$\bar{Y}_{\cdot\cdot}$

RCBD with fixed factor and fixed block factor

- Model (with k treatments and b blocks)

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, i = 1, \ldots, k;\ j = 1, \ldots, b.$$

where iid $\varepsilon_{ij} \sim N(0, \sigma^2)$, $\sum_{i=1}^{k} \tau_i = \sum_{j=1}^{b} \beta_j = 0.$

- SS Decomposition: *SSTO = SSTR + SSB+SSE*

$$SSTO = \sum_{i=1}^{k} \sum_{j=1}^{b} (Y_{ij} - \overline{Y}_{..})^2, \overline{Y}_{..} = \frac{1}{kb} \sum_{i=1}^{k} \sum_{j=1}^{b} Y_{ij};$$

$$SSTR = \sum_{i=1}^{k} b(\overline{Y}_{i.} - \overline{Y}_{..})^2, \overline{Y}_{i.} = \frac{1}{b} \sum_{j=1}^{b} Y_{ij}, i = 1, \ldots, k;$$

$$SSB = \sum_{j=1}^{b} k(\overline{Y}_{.j} - \overline{Y}_{..})^2, \overline{Y}_{.j} = \frac{1}{k} \sum_{i=1}^{k} Y_{ij}, j = 1, \ldots, b;$$

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{b} (Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..})^2$$

ANOVA for RCBD

Source of Variation	SS	DF	MS	F
Treatment	SSTR	k-1	MSTR = SSTR/(k-1)	F_{TR} = MSTR / MSE
Block	SSB	b-1	MSB = SSB/(b-1)	F_B = MSB / MSE
Error	SSE	(k-1)(b-1)	MSE = SSE/((k-1)(b-1))	
Total	SSTO	kb-1		

Distribution of F-statistic in ANOVA

$$E(MSE)=\sigma^2, E(MSTR)=\sigma^2+\frac{b}{k-1}\sum_{i=1}^{k}\tau_i^{\,2},$$

$$E(MSB)=\sigma^2+\frac{k}{b-1}\sum_{j=1}^{b}\beta_j^{\,2}.$$

By Cochran's Theorem, SSTR, SSB, and SSE are mutually independent.

Under $H_0: \tau_i = 0$, i = 1,.., k, (H_1 : at least one $\tau_i \neq 0$)

$$F_{TR}=\frac{SSTR/(k-1)}{SSE/((k-1)(b-1))}\sim F(k-1,(k-1)(b-1)).$$

Under $H_0: \beta_j = 0$, j = 1,.., b, (H_1 : at least one $\beta_j \neq 0$)

$$F_{B}=\frac{SSB/(b-1)}{SSE/((k-1)(b-1))}\sim F(b-1,(k-1)(b-1)).$$

Parameter Estimation and Comparison Test

- ANOVA model

$$\text{LSE or MLE}: \quad \hat{\mu} = \bar{Y}_{..},$$

$$\hat{\tau}_i = \bar{Y}_{i.} - \bar{Y}_{..}, \hat{\beta}_j = \bar{Y}_{.j} - \bar{Y}_{..}, i = 1,\ldots.k,\ j = 1,\ldots.b.$$

$$\text{Reference Distribution for } \bar{Y}_{i.}, i = 1,\ldots,k:$$

$$\bar{Y}_{i.} \sim N\left(\mu + \tau_i, \frac{\sigma^2}{b}\right) \Rightarrow \frac{\bar{Y}_{i.} - (\mu + \tau_i)}{\sqrt{\frac{MSE}{b}}} \sim t((k-1)(b-1)), i = 1,\ldots,k$$

- 100(1-α)% Confidence Interval of i-th treatment mean$(\mu + \tau_i)$:

$$\bar{Y}_{i.} \pm t_{\frac{\alpha}{2}}((k-1)(b-1)) \cdot \sqrt{\frac{MSE}{b}}$$

Follow-up Comparison and Test

- 100(1-α)% Confidence Interval of j-th block mean $(\mu + \beta_j)$

$$\bar{Y}_{.j} \pm t_{\frac{\alpha}{2}}\left((k-1)(b-1)\right) \cdot \sqrt{\frac{MSE}{k}}$$

- 100(1-α)% Confidence Interval of $(\mu_i - \mu_j)$:

$$\left(\bar{Y}_{i.} - \bar{Y}_{j.}\right) \pm t_{\frac{\alpha}{2}}\left((k-1)(b-1)\right) \cdot \sqrt{\frac{2MSE}{b}}$$

- 100(1-α)% Tukey's Simultaneous C.I. for $(\mu_i - \mu_j)$ any pair (i,j)

$$\left(\bar{Y}_{i.} - \bar{Y}_{j.}\right) \pm q\left(\alpha, k, (k-1)(b-1)\right) \cdot \sqrt{\frac{MSE}{b}}$$

Model Diagnosis

- Normality Check for iid errors -- histogram and QQ plot

$$\varepsilon_{ij} \sim N\left(0, \sigma^2\right), \sum_{i=1}^{k} \tau_i = \sum_{j=1}^{b} \beta_j = 0.$$

- Variances in different treatment groups and block groups are roughly the same

$$\bar{Y}_{i.} = \frac{1}{b}\sum_{j=1}^{b} Y_{ij} \sim N\left(\mu + \tau_i, \frac{\sigma^2}{b}\right),$$

$$\bar{Y}_{.j} = \frac{1}{k}\sum_{i=1}^{k} Y_{ij} \sim N\left(\mu + \beta_j, \frac{\sigma^2}{k}\right)$$

- Residual vs fitted value $\sum_{i=1}^{k}\sum_{j=1}^{b} e_{ij}\hat{Y}_{ij} = 0$

Example: (SAS code)

avgain is the mean average daily gain of all calves in a pen. Average daily gain is (final weight minus initial weight)/(number of days) on study.

```
data weightgain;
input farm pen diet $ avgain;
cards;
 1 1 C 2.19
 1 2 D 2.44
 1 3 B 3.02
...

6 21 A 1.73
 6 22 D 1.49
 6 23 B 1.96
6 24 C 1.33
 ;
```

```
/* Without Blcok Effect*/
proc glm data=weightgain;
class diet;
model avgain=diet;
lsmeans diet / pdiff; run;

/* with block effect */
 proc glm data=weightgain;
class farm diet;
model avgain=farm diet / p clm;
 lsmeans diet / pdiff cl adjust=tukey;
output out=OutData p=predict r=residual; run;

proc gplot data=Outdata;
plot predict*residual avgain*residual; run;
```

Section 6.4 Design with two blocks

- Latin Square: each of k treatments is assigned to each level of each of the two blocking factors in a special way

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \\ 3 & 5 & 4 & 2 & 1 \\ 4 & 1 & 5 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{bmatrix}.$$

ANOVA for Latin Square with two blocks

Source of Variation	SS	DF	MS	F
Treatment	SSTR	k-1	MSTR	F_{TR} = MSTR / MSE
Row	SS_{Row}	k-1	MS_{Row}	F_{Row} = MS_{Row} / MSE
Column	SS_{Col}	k-1	MS_{Col}	F_{Col} = MS_{Col} / MSE
Error	SSE	k^2-3k+2	MSE	
Total	SSTO	k^2-1		

Example 4-1

The yields (in bushels per 0.1 of an acre) of 4 corn hybrids.

		Column			Row Total	Treatment Total
	10-A	14-B	7-C	8-D	39	$T_A = 53$
Row	7-D	18-A	11-B	8-C	44	$T_B = 44$
	5-C	10-D	11-A	9-B	35	$T_C = 30$
	10-B	10-C	12-D	14-A	46	$T_D = 37$
Column Total	32	52	41	39		

H_0: $\tau_A = \tau_B = \tau_C = \tau_D = 0$

SSTR=72.5, SSE=10.5 , F_{TR}=MSTR/MSE=13.8

k-1=3, k^2-3k+2=16-12+2=6, $F_{0.01}(3, 6)$=9.78, p-value<0.01.