**BALANCED INCOMPLETE BLOCK DESIGNS (BIBD) & RELATED**

- As with CBD, \( b \) blocks, but now each with \( k \) units, \( k < t \)
- \( N = bk \)
- NOW all treatments cannot fit in any one block, so blocks are called "incomplete"
- Examples: \( b = 4, k = 3, t = 4 \)

<table>
<thead>
<tr>
<th>block</th>
<th>treatment</th>
<th>Usually a bad idea</th>
<th>block</th>
<th>treatment</th>
<th>Usually a better idea</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 2 3</td>
<td></td>
<td>1</td>
<td>1 2 3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1 2 3</td>
<td></td>
<td>2</td>
<td>1 2 4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1 2 3</td>
<td></td>
<td>3</td>
<td>1 3 4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4 4 4</td>
<td></td>
<td>4</td>
<td>2 3 4</td>
<td></td>
</tr>
</tbody>
</table>

Usually a bad idea

<table>
<thead>
<tr>
<th>block</th>
<th>treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 2 3</td>
</tr>
<tr>
<td>2</td>
<td>1 2 4</td>
</tr>
<tr>
<td>3</td>
<td>1 3 4</td>
</tr>
<tr>
<td>4</td>
<td>2 3 4</td>
</tr>
</tbody>
</table>

Usually a better idea
Rules for BALANCED Incomplete Block Designs (BIBD’s)

- **“First-order balance”**: Each treatment must appear the same number of times in the (entire) design:
  \[ r = \frac{N}{t} = \frac{bk}{t} \]

- **“Second-order balance”**: Each PAIR of treatments must appear together in the same number of blocks:
  \[ \lambda = \frac{r(k - 1)}{(t - 1)} \]
  - pick any specific treatment, say “1”
  - numerator is number of “non-1” assignments in blocks containing a “1”
  - denominator is number of “non-1” treatments
• BIBD’s don’t exist for every $t$, $b$, and $k < t$

• Necessary, but not Sufficient, Condition:
  
  $$r = \frac{bk}{t} \quad \text{and} \quad \lambda = \frac{r(k - 1)}{(t - 1)}$$

  must be integers

• Example: $t = 7$, $k = 5$
  
  - $r = \frac{bk}{t} = b[5/7] = \text{integer}$
  
  - $\lambda = [r][\frac{(k - 1)}{(t - 1)}] = b[5/7][4/6] = b[10/21] = \text{integer}$
  
  - smallest integer solution is $b = 21 \rightarrow \lambda = 10, r = 15$

  - EXIST? Yes, in this case use all possible combinations of 7 treatments in groups of size 5:

  $$\binom{7}{5} = 21$$
• Easy Case:
  - any $k < t$, one block for each unique subset of $k$ treatments
    $b = \binom{t}{k}$

    $r = b[k/t] = \frac{t!}{((t - k)!k!)[k/t]} = \binom{t - 1}{k - 1}$

    $\lambda = r[(k - 1)/(t - 1)] = \binom{t - 1}{k - 1} \frac{k - 1}{t - 1} = \binom{t - 2}{k - 2}$

  - the second design on the first slide is an example of this:
    $k = t - 1$, $b = t$, $r = t - 1$, $\lambda = t - 2$

• In General: combinatorics, tables
Model

- \( y_{i,j} = \alpha + \beta_i + \tau_j + \epsilon_{i,j} \) or \( \beta_i + \tau_j + \epsilon_{i,j} \)
  - as with CBD, but not all \((i, j)\) combinations are included

- \( y = X_1\beta + X_2\tau + \epsilon \)

- \([bk \times 1] = [bk \times b][b \times 1] + [bk \times t][t \times 1] + [bk \times 1]\)

Estimable Functions ... need to characterize \((I - H_1)X_2\)

- \( H_1 = X_1(X'_1X_1)^{-1}X'_1 = \frac{1}{k}\text{diag}(J_{k\times k}...J_{k\times k}) \)
  - like CBD, but now \( k \neq t \)
• $H_1X_2 = \frac{1}{k}$

\[
\begin{bmatrix}
\text{# of times trt j appears in blk 1} \\
\text{# of times trt j appears in blk 2} \\
\vdots \\
\text{# of times trt j appears in blk b}
\end{bmatrix}
\]

$\rightarrow$ columns indexed by $j$ $\rightarrow$

$(k$ rows$)$

• Not the same as a CRD with $r$ units/trt group

• $(I - H_1)X_2 =$

\[
\begin{bmatrix}
\text{# of times trt j appears in this row of } X_2 - \\
\frac{1}{k} \text{ no. of times trt j appears in blk 1} \\
\vdots \\
\frac{1}{k} \text{ no. of times trt j appears in blk b}
\end{bmatrix}
\]
• e.g. for $k = 4$, $t = 6$, treatments 1 - 4 in first block:

$$X_{2|1} = \begin{bmatrix}
\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 \\
-\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 \\
-\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & 0 & 0 \\
-\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots 
\end{bmatrix}$$

• Note all rows sum to zero, so all linear combinations of rows add to zero, so only contrasts between treatments are estimable (again)
Reduced Normal Equations: (Need $X'_2|_1 X_{2|1} = X'_2 (I - H_1) X_2$)

- $X'_2 X_2 = rI$
- $(X'_1 X_1)^{-1} = k^{-1}I$
- $\{X'_1 X_2\}_{ij} = \mathcal{N}(i, j)$, an “incidence matrix”
  - number of times treatment $j$ appears in block $i$
  - always 0 or 1
- $\{(X'_1 X_2)'(X'_1 X_2)\}_{j,j'} = \sum_i \mathcal{N}(i, j) \mathcal{N}(i, j')$
  - number of blocks containing treatments $j$ AND $j'$
  - always $r$ for $j = j'$
  - always $\lambda$ for $j \neq j'$
  - $(r - \lambda)I + \lambda J$
- $(X'_2 X_1)(X'_1 X_1)^{-1}(X'_1 X_2) = \frac{r-\lambda}{k}I + \frac{\lambda}{k}J$
• Left side of R.N.E.:
  \[ [X'_2X_2 - X'_2H_1X_2]\hat{\tau} \]
  \[ [rI - (r - \lambda)/kI - \lambda/kJ]\hat{\tau} \]
  \[ \frac{\lambda}{k} [tI - J]\hat{\tau} = \mathcal{I}_{2|1}\hat{\tau} \]

• Right side of R.N.E.:
  \[ X'_2(I - X_1(X'_1X_1)^{-1}X'_1)y \]
  \[ X'_2y - \frac{1}{k}X'_2X_1X'_1y \]
  \[ T - k^{-1}(X'_2X_1)B \]
  * \(T = t\)-element vector of treatment totals
  * \(B = b\)-element vector of block totals
  * \(\{(X'_2X_1)B\}_j = \) sum of block totals containing trt \(j\)
• Multiply both Right and Left expressions above by $k/\lambda$
  - $[tI - J] \hat{\tau} = k/\lambda [T - k^{-1}(X_2'X_1)B]$
  - Now divide both sides of this by $t$
  - $j$th row: $\hat{\tau}_j - \bar{\hat{\tau}} = k/(\lambda t) \left[ y_{.j} - k^{-1} \sum_i N(i,j) y_i. \right]$
  - Second term in brackets is the sum of block averages for (only!) blocks that contain a unit with treatment $j$.
  - Many books use $Q_j$ for $[-]$, “adjusted treatment $j$ total”
  - If $c'\tau$ is estimable,
    \[ \widehat{c'\tau} = k/(\lambda t) \sum_j c_j Q_j \] (since $\bar{\hat{\tau}}$ drops out)
  - Note that $\frac{k}{\lambda t}$ “plays the role of” $\frac{1}{r}$ ... but they aren’t equal
Variance of Estimable Functions:

- $\hat{c}'\tau = \frac{k}{(\lambda t)}c'[X_2' - X_2'H_1]y$

- $Var[\hat{c}'\tau]$
  
  $= \frac{k^2\sigma^2}{(\lambda t)^2}c'[X_2' - X_2'H_1][X_2 - H_1X_2]c$
  
  $= \frac{k^2\sigma^2}{(\lambda t)^2}c'[X_2X_2' - X_2'H_1X_2]c$

  left-side matrix!

  $= \frac{k^2\sigma^2}{(\lambda t)^2}c'[^t\lambda t/k I - \lambda/k J]c$

  $J$ term drops out for estimable functions

  $= \frac{k\sigma^2}{(\lambda t)}[k\sigma^2/(\lambda t)]c'c$

  again, $\frac{k}{\lambda t}$ “plays the role of” $\frac{1}{r}$...
When is it worth using?

- \( \text{Var}_{BIBD}[c'\tau] = [k\sigma_{BIBD}^2/(\lambda t)]c'c \)
- \( \text{Var}_{CRD}[c'\tau] = [\sigma_{CRD}^2/(N/t)]c'c \)

- Use \( r = N/t \) and \( \lambda = (N/t)[(k-1)/(t-1)] \) ...

- \( \text{Var}_{BIBD}/\text{Var}_{CRD} = [k/(k-1)]/[t/(t-1)] \left[\frac{\sigma_{BIBD}^2}{\sigma_{CRD}^2}\right] \)

- "Multiplying ratio" is greater than one if \( k < t \)

- Compare to earlier results, e.g. \( \text{Var}_{LSD}/\text{Var}_{CRD} = \sigma_{LSD}^2/\sigma_{CRD}^2 \)
• Multiplying ratio increases as $k$ decreases relative to $t$, e.g.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$k$</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>1/1</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>72/70</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>54/50</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>36/30</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>18/10</td>
</tr>
</tbody>
</table>

• Expected squared CI length criterion becomes:

$$
\sqrt{\frac{k/(k-1)}{t/(t-1)}} \frac{\sigma_{BIBD}}{\sigma_{CRD}} < \frac{t(1 - \alpha/2; N - t)}{t(1 - \alpha/2; N - b - t + 1)}
$$
ANOVA: SS’s for blocks and treatments aren’t orthogonal ...

- \( \text{SST}(\text{blocks}^*) = \mathbf{y}' \mathbf{H}_1 \mathbf{y} = \sum_{i=1}^{b} k(\bar{y}_i - \bar{y}..)^2 \)
- \( \text{SST}(\text{blocks}^*, \text{treatments}) = \mathbf{y}' \mathbf{Hy} \)

\[
\text{SST}(\text{treatments}|\text{blocks}^*) = \text{SST}(\text{blocks}^*, \text{treatments}) - \text{SST}(\text{blocks}^*)
\]
\[
= \frac{k}{\lambda_t} \sum_j Q_j^2 \quad \text{(slide 10)}
\]
\[
= \frac{k}{\lambda_t} \sum_j [y_{.j} - \frac{1}{k} \sum_i \mathcal{N}(i,j)y_{i.}]^2
\]
\[
= \frac{(t-1)k}{(k-1)t} \sum_j r[\bar{y}_{.j} - \bar{y}[^j]]^2
\]

where \( \bar{y}[^j] \) is the average of all observations from blocks that include a unit assigned to treatment \( j \)

*: “blocks” includes the intercept here, which would ordinarily be removed first in “correction for the mean”
Similar result for Test Power

- Hypothesis: $\tau_1 = \tau_2 = \ldots = \tau_t$
- Compare CRD with $r = N/t$, BIBD with same $N$
- For hypothetical $\tau$, CRD $Q(\tau) = r \sum_i (\tau_i - \bar{\tau})^2$
- For CRD, the distribution of the test statistic is:
  $$F'(t - 1, N - t, \frac{Q(\tau)}{\sigma^2_{CRD}})$$
- For BIBD, the distribution of the test statistic is:
  $$F'(t - 1, N - b - t + 1, \frac{t/(t-1)}{k/(k-1)} \frac{Q(\tau)}{\sigma^2_{BIBD}})$$
- i.e. noncentrality is decreased, and estimation variances increased by the same factor
YOUDEN “SQUARE” DESIGNS

• Recall that in LSD, \(\#(\text{rows}) = \#(\text{columns}) = \#(\text{treatments})\)

• Now consider row-column designs in which
  – ignoring rows leads to a CBD in columns
  – ignoring columns leads to a BIBD in rows

• For example:

  \[
  \begin{array}{ccc}
  1 & 2 & 3 \\
  2 & 4 & 1 \\
  3 & 1 & 4 \\
  4 & 3 & 2 \\
  \end{array}
  \]

• Generally: \(t\) treatments, \(t\) rows, \(k < t\) columns
- **Structure:**
  - Columns are *orthogonal to* Treatments (this matters)
  - Rows are *orthogonal to* Columns (this DOESN’T matter)
  - Rows are *not orthogonal to* Treatments (this matters)

- **$H_1$** is more complicated than it would be for a CBD or a BIBD, but
  - $H_1X_2$ is the same as for a BIBD with $t$ blocks of size $k$.
  - RNE’s are as with BIBD with row blocks (only)
  - Point estimates, variance formulae, and non-centrality parameters can be computed ignoring row blocks

- Can replicate Youden Designs in any of the 3 ways discussed with LSDs
PARTIALLY BALANCED INCOMPLETE BLOCK DESIGNS

- For two associate classes ...
- \(t\) treatments, \(b\) blocks, \(k < t\) units/block
- Every treatment is applied to \(r\) units overall (as in BIBD’s)
- Every pair of treatments are either:
  - *First associates*, applied together in \(\lambda_1\) blocks, or
  - *Second associates*, applied together in \(\lambda_2\) blocks
- Each treatment has the same number of 1st associates and the same number of 2nd associates
- For any two treatments that are \(i\)th associates, there are:
  - \(p_1^i\) other treatments that are 1st associates of each
  - \(p_2^i\) other treatments that are 2nd associates of each
  - \(q^i\) other treatments that are 1st associates of one and 2nd associates of the other
• For example:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

- $t = 6$, $b = 3$, $k = 4$
- $r = 2$
- Treatments can be divided into two groups: $(1,2,3)$ and $(4,5,6)$
- Any two treatments within a group are 1st associates, with $\lambda_1 = 1$
- Any two treatments in different groups are 2nd associates, with $\lambda_2 = 2$

• Analysis is somewhat messier than with BIBD’s, but there is far more flexibility in size for designs for given $t$ and $k$
BIBD’s: INTER- and INTRA-BLOCK INFORMATION

- Fixed-Effects analyses we’ve considered derive treatment information only from within-block comparisons (INTRA-block information)

- INTER-block comparisons are considered “sacrificed”; confounded with unknown block effects.

- Recall $\hat{c}' \tau = \frac{k}{\lambda t} \sum_j c_j [1 \times y_{.j} - \frac{1}{k} \sum_i N(i, j) y_i.]$
  - for every block containing treatment $j$, weights for each data value in the estimate of $\hat{c}' \tau$ are $\frac{k}{\lambda t} \times c_j$ times:
    
    $-\frac{1}{k}, \ -\frac{1}{k}, \ ..., \ 1 - \frac{1}{k}, \ -\frac{1}{k}, \ -\frac{1}{k}, ...$

  - for every block not containing treatment $j$, weights are all zero.

  - in either case, $\hat{c}' \tau = \sum_{i,j} w_{ij} y_{ij}$ where $\sum_j w_{ij} = 0$ for all $i$. 
• First, note that some of the INTRA-block info is not so obvious ... for example:

<table>
<thead>
<tr>
<th>block</th>
<th>treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 2 3</td>
</tr>
<tr>
<td>2</td>
<td>1 2 4</td>
</tr>
<tr>
<td>3</td>
<td>1 3 4</td>
</tr>
<tr>
<td>4</td>
<td>2 3 4</td>
</tr>
</tbody>
</table>

• Is there trt 1-vs-trt 2 INTRA-block information contributed from blocks 3 and 4?
  
  – Overall, $Var[\hat{\tau}_1 - \hat{\tau}_2] = c^T c \sigma^2 \frac{k}{\lambda t} = 2\sigma^2 \frac{k}{\lambda t} = \frac{3}{4} \sigma^2$
  
  – From just blocks 1 and 2,
  
  $Var[\hat{\tau}_1 - \hat{\tau}_2] = Var[\frac{1}{2}((y_{11} - y_{12}) + (y_{21} - y_{22}))] = \sigma^2$
  
  – So the answer is “Yes, but indirect”, e.g.

“Recovery” of INTER-block information

- For example design above, if \( y_{ij} = \beta_i + \tau_j + \epsilon_{ij} \),

\[
\begin{array}{c|ccc}
   +\text{block} & 1 & 2 & 3 \\
   \hline
   2 & \tau_4 - \tau_3 + 3\beta_2 - 3\beta_1 \\
   3 & \tau_4 - \tau_2 + 3\beta_3 - 3\beta_1 & \tau_3 - \tau_2 + 3\beta_3 - 3\beta_2 \\
   4 & \tau_4 - \tau_1 + 3\beta_4 - 3\beta_1 & \tau_3 - \tau_1 + 3\beta_4 - 3\beta_2 & \tau_2 - \tau_1 + 3\beta_4 - 3\beta_3 \\
\end{array}
\]

- Suppose block effects are assumed to be \textit{random}:
  \[
  \beta_i \sim \text{i.i.d. } \mathcal{N}(0, \delta^2) + \mu_\beta
  \]

- Then differences \textit{between} blocks also contain information about treatment differences
  (but ... some of these differences are correlated ... common \( \beta \)'s)
• Practical, Realistic Assumption?

• Sometimes not:
  – machinists ... perhaps some of those available were trained by same instructor, and that these are “systematically different” from others
  – litters of animals ... perhaps some of those available were bred/raised in same facility, ...

• Sometimes maybe:
  – autos actually selected randomly from a year’s production
  – fields actually selected randomly from Iowa farms
• $y_{ij} = \mu \beta + \tau_j + (\beta_i + \epsilon_{ij}) = \mu \beta + \tau_j + \epsilon_{ij}^*$
  
  - $E[\epsilon_{ij}^*] = 0$
  
  - $Var[\epsilon_{ij}^*] = \sigma^2 + \delta^2$
  
  - $Cov[\epsilon_{ij}^*, \epsilon_{i'j'}^*] = \begin{cases} 
    \delta^2, & i = i', j \neq j' \\
    0, & i \neq i', j \neq j' 
  \end{cases}$

• For block-ordered $y$:
  
  - $E[y] = \mu \beta I + X_2 \tau$
  
  - $Var[y] = \sigma^2 I + \delta^2 \text{diag}(J_k, J_k, ..., J_k) = \Sigma$

  - Full, general treatment would involve generalized least squares with unknown $\Sigma$
Neat Trick for BIBD

- Recall the INTRA-block estimator:
  \[
  \hat{c'}\tau = \frac{k}{(\lambda t)} \sum_j c_j [y_{.j} - k^{-1} \sum_i N(i, j)y_i].
  \]

- For random blocks, we can summarize INTER-block information using block totals:
  - Let \( \{z\}_i = y_{i.} \), i.e. the vector of block totals
  - \( z_{b\times1} = k\mu_\beta 1_{b\times1} + U_{b\times t}\tau_{t\times1} + \epsilon^*_{b\times1} \)
  - Elements of \( \epsilon^* \) are i.i.d. w/ variance \( k^2\delta^2 + k\sigma^2 \)
  - \( U = X'_1 X_2 \)!
– \( \mathbf{U} = \mathbf{X}'_1 \mathbf{X}_2 \) (again)

– \( \mathbf{U} \) has:
  * \( k \) 1’s in each row
  * \( r \) 1’s in each column
  * inner product of any two columns = \( \lambda \)

– One implication is that the sum of all columns in \( \mathbf{U} \) is \( k \mathbf{1} \)
  * \( (\mathbf{1}|\mathbf{U}) \) is of rank \( t \) (rather than \( t + 1 \))
  * only linear contrasts in \( \tau \)'s are estimable ...

– RNE’s: \( \mathbf{U}'(\mathbf{I} - \frac{1}{b} \mathbf{J}) \mathbf{U} \hat{\mathbf{\tau}} = \mathbf{U}'(\mathbf{I} - \frac{1}{b} \mathbf{J}) \mathbf{z} \)

– \( \hat{\mathbf{c}}' \mathbf{\tau} = k/(r - \lambda) \sum_j c_j [k^{-1} \sum_i \mathbf{N}(i, j)y_i - r \bar{y}..] \)

– \( \text{Var}[\hat{\mathbf{c}}' \mathbf{\tau}] = \left[(k^2 \delta^2 + k\sigma^2)/(r - \lambda)\right] \mathbf{c}' \mathbf{c} \)
Bottom Line:

• \( \sum_j \hat{c}_j \tau_j \) and \( \sum_j \tilde{c}_j \tau_j \) are independent statistics

• The weighted l.s. estimator (BLUE, given \( \Sigma \)) is a weighted average of the two, with weights inversely proportional to their respective variances.

\[
\sum_j \hat{c}_j \tau_j = \frac{w_1}{w_1 + w_2} \sum_j \hat{c}_j \tau_j + \frac{w_2}{w_1 + w_2} \sum_j \tilde{c}_j \tau_j
\]

• \( w_1 = \left[ k^2 \delta^2 + k \sigma^2 \right] / (r - \lambda) \) (generally larger)

• \( w_2 = \left[ k \sigma^2 \right] / (t \lambda) \)

• \( E[MSE] = \begin{bmatrix}
\sigma^2 & \text{INTRA: fit of individual data} \\
\delta^2 + k \sigma^2 & \text{INTER: fit of block-total data}
\end{bmatrix} \)

• Since, we ordinarily don’t know the weights, substitute:

\[ \hat{w}_1 = \text{MSE}_{\text{INTER}} / (r - \lambda) \]

\[ \hat{w}_2 = k \text{MSE}_{\text{INTRA}} / (t \lambda) \]
• If we knew $\sigma$ and $\delta$, we could write:

$$\text{Var} [\hat{c}'\tau] = \left( \frac{w_1}{w_1 + w_2} \right)^2 \text{Var} [\hat{c}'\tau] + \left( \frac{w_2}{w_1 + w_2} \right)^2 \text{Var} [\hat{c}'\tau]$$

$$= c'c \left[ \left( \frac{w_1}{w_1 + w_2} \right)^2 w_2 + \left( \frac{w_2}{w_1 + w_2} \right)^2 w_1 \right]$$

$$= c'c \frac{w_1 w_2}{w_1 + w_2}$$

• Again, can substitute estimates for $\sigma^2$ and $k^2 \delta^2 + k \sigma^2$, although this does not necessarily lead to an estimator that is better than $\hat{c}'\tau$. 