ST4241 — Design and Analysis of Clinical Trials
Lecture 22: BIBD (cont.) and introduction to factorial designs

Chen Zehua

Department of Statistics & Applied Probability

8:00-10:00 am, Tuesday, November 1, 2016
Outline

Regression approach to BIBD

Application to interexaminer reliability study

Combination of BIBD and Latin square designs

Introduction to factorial designs
The regression model

The response $X$ in a BIBD is expressed in another linear model (with reparametrization) as:

$$X = \mu + \sum_{i=2}^{n} \gamma_i b_i + \sum_{j=2}^{g} \beta_j t_j + \epsilon,$$

where

$$b_i = \begin{cases} 
1, & \text{if block } i, \\
0, & \text{otherwise, } i = 2, \ldots, n;
\end{cases}$$

$$t_j = \begin{cases} 
1, & \text{if treatment } j, \\
0, & \text{otherwise, } j = 2, \ldots, g.
\end{cases}$$
Inference based on the regression model

- The hypothesis testing for treatment effects is equivalent to testing $H_0 : \beta_2 = \cdots = \beta_g = 0$.
- The hypothesis testing for block effects is equivalent to testing $H_0 : \gamma_2 = \cdots = \gamma_n = 0$.
- A contrast of the treatment effects is expressed in terms of the $\beta$ parameters. The multiple comparison on the treatment effects are done through the tests on the corresponding $\beta$ parameters.
The F-statistics for the treatment effect and block effect can be obtained in either one of the following two methods: (i) through the construction of the Wald statistic for testing the regression parameters; (ii) by fitting the model with different order of arguments in the formula specification of \( \text{lm} \).

(a) place the treatment factor as the last argument in the formula specification to get the ANOVA table with correct F-value for treatment effects;

(b) place the block factor as the last argument in the formula specification to get the ANOVA table with correct F-value for block effects;
Illustration by depression assessment data

- The ANOVA table with correct F-value for examiners:
  \[
  \text{lm.fit1} = \text{lm}(x \sim \text{patient} + \text{examiner})
  \]
  \[
  \text{anova(lm.fit1)}
  \]

- The ANOVA table with correct F-value for patients:
  \[
  \text{lm.fit2} = \text{lm}(x \sim \text{examiner} + \text{patient})
  \]
  \[
  \text{anova(lm.fit2)}
  \]

- Alternatively, the two $F$-ratios can be computed by
  \[
  b = \text{lm.fit2}$coef
  \]
  \[
  v = \text{vcov(lm.fit2)}
  \]
  \[
  b.e = b[2:6]
  \]
  \[
  v.e = v[2:6,2:6]
  \]
  \[
  F.e = t(b.e) \times \text{solve(v.e)} \times b.e/5
  \]
  \[
  b.p = b[7:15]
  \]
  \[
  v.p = v[7:15,7:15]
  \]
  \[
  F.p = t(b.p) \times \text{solve(v.p)} \times b.p/9
  \]
Reliability of measurement

- Measurement reliability concerns with the quality of data: whether or not the data obtained are reliable.

- A measurement on a characteristic of a patient can be expressed as

\[ X_i = S_i + \epsilon_i, \]

where \( S_i \) is the true value of the characteristic which follows a distribution with mean \( \mu \) and variance \( \sigma^2_s \), and \( \epsilon_i \) is a random measurement error distributed with mean zero and variance \( \sigma^2_{\epsilon} \).

- The reliability coefficient is defined as \( R = \frac{\sigma^2_s}{\sigma^2_X} \), where \( \sigma^2_X \) is the expectation of the sample variance of the \( X_i \)'s. In the above simple case, \( \sigma^2_X = \sigma^2_s + \sigma^2_{\epsilon} \).
Reliability of the measurements given by different examiners

- Interexaminer reliability arises when measurements of the subjects are taken by different examiners or raters. Suppose the measurement of each subject is taken by a randomly assigned examiner from $g$ examiners. The measurement on subject $i$ taken by examiner $j$ can be expressed as

$$X_{ij} = S_i + \alpha_j + \epsilon_{ij}.$$  

- In this case,

$$\sigma^2_X = \sigma^2_s + \frac{1}{g} \sum_{j=1}^{g} \alpha_j^2 + \sigma^2_\epsilon,$$

and the reliability coefficient is given by

$$R = \frac{\sigma^2_s}{\sigma^2_s + \frac{1}{g} \sum_{j=1}^{g} \alpha_j^2 + \sigma^2_\epsilon}.$$
Estimation of reliability coefficient

- The BIBD can be applied for the interexaminer reliability study to estimate the reliability coefficient. The example of depression assessment in lecture 21 is in fact an interexaminer reliability study.

- An unbiased estimate of $\sigma^2_\epsilon$ is given by
  \[ \hat{\sigma}^2_\epsilon = \text{RMS}. \]

- The unbiased estimate of $\nu^2 = \frac{1}{g} \sum_{j=1}^{g} \alpha_j^2$ is given by
  \[ \hat{\nu}^2 = \frac{g - 1}{r_{\text{EFF}}} \frac{\text{TMS(EB)} - \text{RMS}}{g}. \]

- An unbiased estimate of $\sigma^2_s$ is given by
  \[ \hat{\sigma}^2_s = \frac{1}{k_{\text{EFF}}} (\text{BMS(ET)} - \text{RMS}). \]
Estimation of reliability coefficient (cont.)

- The reliability coefficient $R$ is estimated by
  \[ R = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\nu}^2 + \hat{\sigma}_\epsilon^2}. \]

- Estimation of $R$ for depression assessment:
  \[
  \text{RMS} = 9.28, \quad \text{TMS(EB)} = 7.09, \quad \text{BMS(ET)} = 92.26.
  \]
  \[
  \text{EFF} = \frac{6 \times (3 - 1)}{3 \times (6 - 1)} = \frac{4}{5}; \quad \overline{\text{EFF}} = \frac{10 \times (5 - 1)}{5 \times (10 - 1)} = \frac{8}{9},
  \]
  \[
  \hat{\nu}^2 = \frac{(6 - 1)(7.09 - 9.26)}{6 \cdot 5 \cdot 4/5} = -0.4521,
  \]
  \[
  \hat{\sigma}_s^2 = \frac{1}{3 \cdot 8/9}(92.26 - 9.26) = 31.125,
  \]
  \[
  \hat{R} = \frac{31.125}{31.125 + (-0.4521) + 9.26} = 0.7794.
  \]
Latin square enlargement of BIBD

A BIBD scheme can be combined with Latin squares in the following way: each block is enlarged to a Latin square. Thus another factor of $k$ (the size of the blocks) levels can be controlled in addition to the block factor.

An example: The original BIBD scheme:

<table>
<thead>
<tr>
<th>Block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Example: enlarged scheme

The enlarged scheme:

<table>
<thead>
<tr>
<th>Block</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Block 1</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
</tr>
<tr>
<td>7</td>
<td>B</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
</tr>
<tr>
<td>9</td>
<td>A</td>
</tr>
<tr>
<td>10</td>
<td>B</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
</tr>
<tr>
<td>12</td>
<td>A</td>
</tr>
</tbody>
</table>
### Example: Data

The following table provides the data of a study with the enlarged scheme. A and B stands for two periods.

<table>
<thead>
<tr>
<th>Patient</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.0894 (A)</td>
<td>-1.3200 (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1.7577 (B)</td>
<td>-0.9817 (A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.0771 (B)</td>
<td></td>
<td>-1.7531 (A)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.9381 (A)</td>
<td>-1.6769 (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1.2044 (B)</td>
<td></td>
<td></td>
<td>-0.7795 (A)</td>
</tr>
<tr>
<td>6</td>
<td>-1.0395 (B)</td>
<td></td>
<td></td>
<td>-1.0426 (A)</td>
</tr>
<tr>
<td>7</td>
<td>-1.0991 (B)</td>
<td>-0.8092 (A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>-2.0245 (A)</td>
<td>-1.3374 (B)</td>
</tr>
<tr>
<td>9</td>
<td>-0.9846 (A)</td>
<td></td>
<td>-1.4712 (B)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>-1.1395 (B)</td>
<td>-1.6683 (A)</td>
</tr>
<tr>
<td>11</td>
<td>-0.8069 (A)</td>
<td></td>
<td></td>
<td>-1.1913 (B)</td>
</tr>
<tr>
<td>12</td>
<td>-0.7789 (A)</td>
<td></td>
<td></td>
<td>-1.1694 (B)</td>
</tr>
</tbody>
</table>
Example: computation

The computation is similar to that for the original BIBD data, except an additional factor is added. One of the ANOVA table yielded from `anova(lm.fit)` is as follows:

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periods</td>
<td>1</td>
<td>0.1390</td>
<td>0.1390</td>
<td></td>
</tr>
<tr>
<td>Patient (IF)</td>
<td>11</td>
<td>1.1564</td>
<td>0.1051</td>
<td></td>
</tr>
<tr>
<td>Formulation(EP)</td>
<td>3</td>
<td>1.2799</td>
<td>0.4266</td>
<td>18.23</td>
</tr>
<tr>
<td>Res.</td>
<td>8</td>
<td>0.1870</td>
<td>0.0234</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>2.7623</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The design

- A $p$-factor ($p \geq 2$) factorial design is used to investigate the effects of $p$ factors including clinical factors and prognostic factors.
- In a $p$-factor factorial design, all combinations of the levels of the $p$ factors are considered. At each combination, more than one independent responses are obtained.
- With the factorial design, all the effects including the main effect of each factor, the two-way interaction between any two factors, the three-way interaction among any three factors, and so on, can be investigated.
The Analysis

- The data of a factorial design is analyzed through ANOVA.
- The procedure of ANOVA is (a) identify the sources of variations, (b) to compute the SS associated with each source, (c) to form proper $F$-ratios for testing the significance of the effects.
- The sources of variation in a factorial design could include all the main effects and all the interactions from two-factor to $p$-factor interactions. But in practice, only lower order interactions are of interest, higher order interactions can be ignored.
- The SS for various effects can be computed straightforwardly according to certain rules.
- A 3-factor factorial design is used for the illustration of the analysis.
3-factor factorial design

Suppose in factorial design with three factors: $A$, $B$ and $C$, the number of levels of the factors are $a$, $b$ and $c$ respectively, and there are $n$ responses at each combination of the levels.

The ANOVA table is

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$a-1$</td>
<td>SSA</td>
<td>MSA</td>
</tr>
<tr>
<td>$B$</td>
<td>$b-1$</td>
<td>SSB</td>
<td>MSB</td>
</tr>
<tr>
<td>$C$</td>
<td>$c-1$</td>
<td>SSC</td>
<td>MSC</td>
</tr>
<tr>
<td>$AB$</td>
<td>$(a-1)(b-1)$</td>
<td>SSAB</td>
<td>MSAB</td>
</tr>
<tr>
<td>$AC$</td>
<td>$(a-1)(c-1)$</td>
<td>SSAC</td>
<td>MSAC</td>
</tr>
<tr>
<td>$BC$</td>
<td>$(b-1)(c-1)$</td>
<td>SSBC</td>
<td>MSBC</td>
</tr>
<tr>
<td>$ABC$</td>
<td>$(a-1)(b-1)(c-1)$</td>
<td>SSABC</td>
<td>MSABC</td>
</tr>
<tr>
<td>Residual by subtraction</td>
<td>by subtraction</td>
<td>MSR</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$nabc - 1$</td>
<td>SST</td>
<td></td>
</tr>
</tbody>
</table>
The df for a main effect of a factor is the number of levels of that factor minus 1, the df of an interaction is the product of the df’s of the involved factors’ main effects. E.g, df for main effect of A, interactions AB and ABC are, respectively,

\[ a - 1, \quad (a - 1)(b - 1), \quad (a - 1)(b - 1)(c - 1). \]
Rules for computing SS

Let $X_{ijkl}$ denote the response of individual $l$ at levels $i, j$ and $k$ of $A, B$ and $C$ respectively.

When some indices of $X_{ijkl}$ are replaced by dots, it denotes the sum of the $X$’s over those indices, e.g.,

$$X_{i...} = \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} X_{ijkl}, \quad X_{ij...} = \sum_{k=1}^{c} \sum_{l=1}^{n} X_{ijkl}.$$  

These are called partial sums at level combinations of the factors whose indices are not replaced by dots.
Rules for computing SS (cont.)

- The uncorrected sum of squares (USS) for a subset of factors is computed as follows:
  
  (i) For each level combination of those factors, compute the partial sums, e.g., for USS of A and B, the partial sums are: \( X_{ij..}, \ i = 1, \ldots, a, j = 1, \ldots, b. \)

  (ii) Divide the square of each partial sum by the number of observations involved in that partial sum, e.g., \( \frac{X_{ij..}^2}{cn} \), then add them up to obtain the USS.

- Examples of USS:

\[
\begin{align*}
\text{USS}_A &= \sum_{i=1}^{a} \frac{X_{i..}^2}{bcn}, \\
\text{USS}_{AB} &= \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{X_{ij..}^2}{cn}, \\
\text{USS}_{ABC} &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \frac{X_{ijk..}^2}{n}.
\end{align*}
\]
Rules for computing SS (cont.)

- Procedure to compute sum of squares (SS):
  1. For each effect, expand its degrees of freedom formula.
  2. Replace each component of the expansion by their corresponding USS, and replace 1 by $\chi^2/(nabc)$.

- Examples:
  - For main effect of A, the df is $a - 1$,
    $$SS_A = USS_A - \frac{\chi^2}{nabc}.$$ 
  - For interaction between A and B, the expansion of df and SS are
    $$(a - 1)(b - 1) = ab - a - b + 1$$
    $$SS_{AB} = USS_{AB} - USS_A - USS_B + \frac{\chi^2}{nabc}.$$
Rules for computing SS (cont.)

- For interaction of A, B and C, the expansion of df and SS are
  \[(a - 1)(b - 1)(c - 1) = abc - ab - ac - bc + a + b + c - 1\]
  \[SS_{AB} = USS_{ABC} - USS_{AB} - USS_{AC} - USS_{BC}\]
  \[+USS_A + USS_B + USS_C - \frac{X^2}{nabc} .\]

F-ratios for testing the significance of the effects

- If all the effects are considered as fixed effects, the F-ratio for testing a particular effect is the MS of that effect divided by RMS, the residual mean sum of squares.
- If certain effects are random effects, the expected MS need to be derived, and the F-ratios are determined by the general rule discussed before.
The effect of a factor is measured by the contrasts of its level means. For example, for factor A, the contrasts are of the form:
\[ \sum_{i=1}^{a} c_i^{(a)} \bar{X}_i... \]
where \( \bar{X}_i... = X_i.../bcn \). In particular, if A has only two levels, the effect is measured by \( \bar{X}_2... - \bar{X}_1... \).

The two-way interaction of two factors, is measured by the contrasts of contrasts. For example, for interaction AB, the contrasts are of the form:
\[ \sum_{j=1}^{b} c_j^{(b)} \left[ \sum_{i=1}^{a} c_i^{(a)} \bar{X}_{ij}... \right] = \sum_{i=1}^{a} \sum_{j=1}^{b} c_i^{(a)} c_j^{(b)} \bar{X}_{ij}... \]
Contrasts of the above form are called two-way interaction contrasts.
Example: Interaction effects for factors with two levels

- When both factors have only two levels, the two-way interaction contrast becomes

\[(\bar{X}_{22..} - \bar{X}_{12..}) - (\bar{X}_{12..} - \bar{X}_{11..}) = \bar{X}_{22..} - \bar{X}_{12..} - \bar{X}_{12..} + \bar{X}_{11..}\]

- In general, an interaction of any order is measured by the contrasts of its immediate lower interaction contrasts. For example, when all factors have only two levels, the three-way interaction contrast is given by

\[(\bar{X}_{222.} - \bar{X}_{122.} - \bar{X}_{122.} + \bar{X}_{112.}) - (\bar{X}_{221.} - \bar{X}_{121.} - \bar{X}_{121.} + \bar{X}_{111.})\]