Experiments in rectangular areas:
restricted randomization or row-column designs?

R. A. Bailey

r.a.bailey@qmul.ac.uk

Lisboa, April 2012

Thanks to CAPES for support in Brasil
1. Statistician and scientist discuss the planned experiment ...

...

Do you  
▶ Go back to step 3, rerandomize and hope that the next field plan will be OK? (but maybe you will reject a large proportion of plans)
▶ Learn pertinent new information about constraints on the design, and so go back to step 1?
▶ Go back to step 2, and agree on a scheme of restricted randomization?
1. Statistician and scientist discuss the planned experiment …
2. … and agree on a design.
1. Statistician and scientist discuss the planned experiment …
2. … and agree on a design.
3. Statistician randomizes the design to produce a field plan.
1. Statistician and scientist discuss the planned experiment …
2. … and agree on a design.
3. Statistician randomizes the design to produce a field plan.
4. Scientist says “Oh, I can’t possibly do it that way because …”
1. Statistician and scientist discuss the planned experiment …
2. … and agree on a design.
3. Statistician randomizes the design to produce a field plan.
4. Scientist says “Oh, I can’t possibly do it that way because …”
1. Statistician and scientist discuss the planned experiment …

2. … and agree on a design.

3. Statistician randomizes the design to produce a field plan.

4. Scientist says “Oh, I can’t possibly do it that way because …”

Do you

► Go back to step 3, rerandomize and hope that the next field plan will be OK? (but maybe you will reject a large proportion of plans)
1. Statistician and scientist discuss the planned experiment …
2. … and agree on a design.
3. Statistician randomizes the design to produce a field plan.
4. Scientist says “Oh, I can’t possibly do it that way because …”

Do you
  ▶ Go back to step 3, rerandomize and hope that the next field plan will be OK? (but maybe you will reject a large proportion of plans)
  ▶ Learn pertinent new information about constraints on the design, and so go back to step 1?
1. Statistician and scientist discuss the planned experiment ...
2. … and agree on a design.
3. Statistician randomizes the design to produce a field plan.
4. Scientist says “Oh, I can’t possibly do it that way because …”

Do you
- Go back to step 3, rerandomize and hope that the next field plan will be OK? (but maybe you will reject a large proportion of plans)
- Learn pertinent new information about constraints on the design, and so go back to step 1?
- Go back to step 2, and agree on a scheme of restricted randomization?
The problem

An agricultural experiment to compare $n$ treatments. The experimental area has $r$ rows and $n$ columns.

Use a randomized complete-block design with rows as blocks. (In each row independently, choose one of the $n!$ orders with equal probability.)

What should we do if the randomization produces a plan with one treatment always at one side of the rectangle?
Federer (1955 book): guayule trees

<table>
<thead>
<tr>
<th>B</th>
<th>D</th>
<th>G</th>
<th>A</th>
<th>F</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>G</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>E</td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>G</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>F</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>E</td>
</tr>
</tbody>
</table>
Federer (1955 book): guayule trees

\[
\begin{array}{cccccc}
B & D & G & A & F & C & E \\
A & G & C & D & F & B & E \\
G & E & D & F & B & C & A \\
B & A & C & F & G & E & D \\
G & B & F & C & D & A & E \\
\end{array}
\]
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual

Solution: Simple-minded restricted randomization
Keep re-randomizing until you get a plan you like.
Analyse as usual.
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual

Solution: Simple-minded restricted randomization
Keep re-randomizing until you get a plan you like.
Analyse as usual.

Solution: Use a Latinized design, but analyse as usual
Deliberately construct a design in which
no treatment occurs more than once in any column.
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual

Solution: Simple-minded restricted randomization
Keep re-randomizing until you get a plan you like.
Analyse as usual.

Solution: Use a Latinized design, but analyse as usual
Deliberately construct a design in which
no treatment occurs more than once in any column.

Solution (following Yates):
Super-valid restricted randomization, with usual analysis
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual

Solution: Simple-minded restricted randomization
Keep re-randomizing until you get a plan you like.
Analyse as usual.

Solution: Use a Latinized design, but analyse as usual
Deliberately construct a design in which
no treatment occurs more than once in any column.

Solution (following Yates):
Super-valid restricted randomization, with usual analysis

Solution: Efficient row-column design,
with analysis allowing for rows and columns
Proposed courses of action

Solution (Fisher): Continue to randomize and analyse as usual

Solution: Simple-minded restricted randomization
Keep re-randomizing until you get a plan you like.
Analyse as usual.

Solution: Use a Latinized design, but analyse as usual
Deliberately construct a design in which
no treatment occurs more than once in any column.

Solution (following Yates):
Super-valid restricted randomization, with usual analysis

Solution: Efficient row-column design,
with analysis allowing for rows and columns

Solution: Use a carefully chosen Latinized design;
REML/ANOVA estimates of variance components
Continue to randomize and analyse as usual

- Simple to construct.
Continue to randomize and analyse as usual

- Simple to construct.
- Simple to randomize.
Continue to randomize and analyse as usual

- Simple to construct.
- Simple to randomize.
- Simple to analyse.
Continue to randomize and analyse as usual

- Simple to construct.
- Simple to randomize.
- Simple to analyse.
- Some treatment comparisons in some experiments will have a specially low or specially high variance, but the estimated variance is unbiased when averaged over all comparisons and all possible randomized plans.
Assumed model

$Y_\alpha$ is the response on plot $\alpha$.

$E(Y_\alpha) = \theta_i$ where $i$ is the treatment on $\alpha$.

$\text{Var}(Y_\alpha) = \sigma^2$ for all $\alpha$

$\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho\sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau\sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}$

with $0 \leq \rho \leq 1$ and $0 \leq \tau \leq 1$. 
Concurrence

$\lambda_{ij} = $ number of pairs of plots in the same column getting treatments $i$ and $j$.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>D</th>
<th>G</th>
<th>A</th>
<th>F</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>G</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>G</td>
<td>E</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>F</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>
Concurrence

\( \lambda_{ij} = \) number of pairs of plots in the same column getting treatments \( i \) and \( j \).

\[
\begin{array}{ccccccc}
B & D & G & A & F & C & E \\
A & G & C & D & F & B & E \\
G & E & D & F & B & C & A \\
B & A & C & F & G & E & D \\
G & B & F & C & D & A & E \\
\end{array}
\]

\( \lambda_{AD} = 0 + 1 + 0 + 1 + 0 + 0 + 1 = 3 \)
Concurrence

\( \lambda_{ij} = \) number of pairs of plots in the same column getting treatments \( i \) and \( j \).

\[ \begin{array}{cccccc}
B & D & G & A & F & C & E \\
A & G & C & D & F & B & E \\
G & E & D & F & B & C & A \\
B & A & C & F & G & E & D \\
G & B & F & C & D & A & E \\
\end{array} \]

\[ \lambda_{AD} = 0 + 1 + 0 + 1 + 0 + 0 + 0 + 1 = 3 \]
\[ \lambda_{AB} = 2 + 1 + 0 + 0 + 0 + 1 + 0 + 0 = 4 \]
Concurrence

\[ \lambda_{ij} = \text{number of pairs of plots in the same column getting treatments } i \text{ and } j. \]

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>D</th>
<th>G</th>
<th>A</th>
<th>F</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>G</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>G</td>
<td>E</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>F</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>

\[ \lambda_{AD} = 0 + 1 + 0 + 1 + 0 + 0 + 0 + 1 = 3 \]
\[ \lambda_{AB} = 2 + 1 + 0 + 0 + 0 + 1 + 0 + 0 = 4 \]
\[ \lambda_{AA} = 1 + 1 + 0 + 1 + 0 + 1 + 0 + 1 = 5 \]
Concurrence

\( \lambda_{ij} = \) number of pairs of plots in the same column getting treatments \( i \) and \( j \).

\[
\begin{array}{cccccc}
B & D & G & A & F & C \\
A & G & C & D & F & B \\
G & E & D & F & B & C \\
B & A & C & F & G & E \\
G & B & F & C & D & A \\
\end{array}
\]

\[
\lambda_{AD} = 0 + 1 + 0 + 1 + 0 + 0 + 1 = 3 \\
\lambda_{AB} = 2 + 1 + 0 + 0 + 0 + 1 + 0 = 4 \\
\lambda_{AA} = 1 + 1 + 0 + 1 + 0 + 1 + 1 = 5 \\
\lambda_{BB} = 4 + 1 + 0 + 0 + 1 + 1 + 0 = 7
\]
Pairwise variance

\[
\text{Var}(Y_\alpha) = \sigma^2 \quad \text{for all } \alpha
\]

\[
\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho \sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau \sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}
\]

\[
V_{ij} = \text{variance of the estimator of } \theta_i - \theta_j
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r - 2r \rho + (\lambda_{ii} - r) \tau + (\lambda_{jj} - r) \tau - 2\lambda_{ij} \tau \right]
\]
Pairwise variance

\[
\text{Var}(Y_\alpha) = \sigma^2 \quad \text{for all } \alpha
\]

\[
\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho \sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau \sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}
\]

\[
V_{ij} = \text{variance of the estimator of } \theta_i - \theta_j
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r - 2r \rho + (\lambda_{ii} - r) \tau + (\lambda_{jj} - r) \tau - 2\lambda_{ij}\tau \right]
\]

↑

same

plot
Pairwise variance

\[
\text{Var}(Y_\alpha) = \sigma^2 \quad \text{for all } \alpha
\]

\[
\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho \sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau \sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}
\]

\[
V_{ij} = \text{variance of the estimator of } \theta_i - \theta_j
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r - 2r \rho + (\lambda_{ii} - r) \tau + (\lambda_{jj} - r) \tau - 2\lambda_{ij} \tau \right]
\]

↑

same

\n
↑

same

\n
plot

row
Pairwise variance

\[
\text{Var}(Y_\alpha) = \sigma^2 \quad \text{for all } \alpha
\]

\[
\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho \sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau \sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}
\]

\[
V_{ij} = \text{variance of the estimator of } \theta_i - \theta_j
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r - 2r \rho + (\lambda_{ii} - r) \tau + (\lambda_{jj} - r) \tau - 2\lambda_{ij} \tau \right]
\]

↑

same plot

↓

same row

↑

same column
Pairwise variance

\[
\text{Var}(Y_\alpha) = \sigma^2 \quad \text{for all } \alpha
\]

\[
\text{Cov}(Y_\alpha, Y_\beta) = \begin{cases} 
\rho \sigma^2 & \text{if } \alpha \neq \beta \text{ in same row} \\
\tau \sigma^2 & \text{if } \alpha \neq \beta \text{ in same column} \\
0 & \text{if } \alpha \neq \beta \text{ otherwise}
\end{cases}
\]

\[
V_{ij} = \text{variance of the estimator of } \theta_i - \theta_j
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r - 2r \rho + (\lambda_{ii} - r) \tau + (\lambda_{jj} - r) \tau - 2\lambda_{ij} \tau \right]
\]

\[
= \frac{\sigma^2}{r^2} \left[ 2r(1 - \rho) + (\lambda_{ii} + \lambda_{jj} - 2\lambda_{ij} - 2r) \tau \right]
\]
Pairwise variance in the example

\[
\begin{array}{cccccc}
  B & D & G & A & F & C \\ 
  A & G & C & D & F & B \\ 
  G & E & D & F & B & C \\ 
  B & A & C & F & G & E \\ 
  G & B & F & C & D & A \\
\end{array}
\]

From \( V_{BG} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{4}{5} \tau \right] \)
Pairwise variance in the example

<table>
<thead>
<tr>
<th>B</th>
<th>D</th>
<th>G</th>
<th>A</th>
<th>F</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>G</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>E</td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>D</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>G</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>F</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>E</td>
</tr>
</tbody>
</table>

From $V_{BG} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{4}{5} \tau \right]$ to $V_{EF} = \frac{2\sigma^2}{5} \left[ 1 - \rho + \tau \right]$
Pairwise variance in the example

\[
\begin{array}{ccccccc}
B & D & G & A & F & C & E \\
A & G & C & D & F & B & E \\
G & E & D & F & B & C & A \\
B & A & C & F & G & E & D \\
G & B & F & C & D & A & E
\end{array}
\]

From \( V_{BG} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{4}{5} \tau \right] \) to \( V_{EF} = \frac{2\sigma^2}{5} \left[ 1 - \rho + \tau \right] \)

with average \( V = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{1}{15} \tau \right] \).
Continue to randomize and analyse as usual: summary

- Simple to construct.
- Simple to randomize.
- Simple to analyse.
- Some treatment comparisons in some experiments will have a specially low or specially high variance, but the estimated variance is unbiased when averaged over all comparisons and all possible randomized plans.
Simple restricted randomization

Keep re-randomizing until you get a plan you like.
Analyse as usual.
Simple restricted randomization

Keep re-randomizing until you get a plan you like. Analyse as usual.

- Inefficient to produce plans: many will have to be rejected.
Simple restricted randomization

Keep re-randomizing until you get a plan you like. Analyse as usual.

- Inefficient to produce plans: many will have to be rejected. For the $5 \times 7$ rectangle, the proportion of plans with no repeat in any column is only 0.000006.
Simple restricted randomization

Keep re-randomizing until you get a plan you like. Analyse as usual.

- Inefficient to produce plans: many will have to be rejected. For the $5 \times 7$ rectangle, the proportion of plans with no repeat in any column is only 0.000006.

- The actual variance of treatment comparisons is lower, but the estimate of that variance is higher.
Calculations

\[ \lambda_{ij} = \text{number of pairs of plots in the same column getting treatments } i \text{ and } j \]

Note that \( \sum_{j=1}^{n} \lambda_{ij} = r^2 \) for each \( i \).

We know that \( V_{ij} = \frac{\sigma^2}{r^2} \left[ 2r(1 - \rho) + (\lambda_{ii} + \lambda_{jj} - 2\lambda_{ij} - 2r)\tau \right] \)
Calculations

\[ \lambda_{ij} = \text{number of pairs of plots in the same column getting treatments } i \text{ and } j \]

Note that \[ \sum_{j=1}^{n} \lambda_{ij} = r^2 \] for each \( i \).

We know that

\[ V_{ij} = \frac{\sigma^2}{r^2} \left[ 2r(1 - \rho) + (\lambda_{ii} + \lambda_{jj} - 2\lambda_{ij} - 2r)\tau \right] \]

Put \[ V = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} V_{ij} \] and put \[ D = \sum_{i=1}^{n} \lambda_{ii}. \]
Calculations

\( \lambda_{ij} = \) number of pairs of plots in the same column getting treatments \( i \) and \( j \)

Note that \( \sum_{j=1}^{n} \lambda_{ij} = r^2 \) for each \( i \).

We know that \( V_{ij} = \frac{\sigma^2}{r^2} [2r(1 - \rho) + (\lambda_{ii} + \lambda_{jj} - 2\lambda_{ij} - 2r)\tau] \)

Put \( V = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} V_{ij} \) and put \( D = \sum_{i=1}^{n} \lambda_{ii} \).

Calculations give \( V = \frac{2\sigma^2}{r^2} \left[ r(1 - \rho) + \left( \frac{D - r^2}{n-1} - r \right) \tau \right] \).
Spectral form of covariance matrix

\[ \text{Cov}(\mathbf{Y}) = \sigma^2[I + \rho(\mathbf{R} - I) + \tau(\mathbf{C} - I)] \]

where \( I \) is the identity matrix, 
\( \mathbf{R} \) is the matrix whose \((\alpha, \beta)\)-entry is equal to 1 if plots \( \alpha \) and \( \beta \) are in the same row and to 0 otherwise, 
\( \mathbf{C} \) is the similarly defined matrix for columns.
Spectral form of covariance matrix

\[ \text{Cov}(\mathbf{Y}) = \sigma^2 [\mathbf{I} + \rho (\mathbf{R} - \mathbf{I}) + \tau (\mathbf{C} - \mathbf{I})] \]

where \( \mathbf{I} \) is the identity matrix, \( \mathbf{R} \) is the matrix whose \((\alpha, \beta)\)-entry is equal to 1 if plots \( \alpha \) and \( \beta \) are in the same row and to 0 otherwise, \( \mathbf{C} \) is the similarly defined matrix for columns.

So \( \text{Cov}(\mathbf{Y}) = \xi_0 \mathbf{S}_0 + \xi_1 \mathbf{S}_1 + \xi_2 \mathbf{S}_2 + \xi_3 \mathbf{S}_3 \), where

\[
\begin{align*}
\xi_0 &= \sigma^2 (1 + (n - 1)\rho + (r - 1)\tau) \\
\xi_1 &= \sigma^2 (1 - \tau + (n - 1)\rho) \\
\xi_2 &= \sigma^2 (1 - \rho + (r - 1)\tau) \\
\xi_3 &= \sigma^2 (1 - \rho - \tau)
\end{align*}
\]

\( \mathbf{S}_1 \) = \( \frac{1}{n} \mathbf{R} - \frac{1}{rn} \mathbf{J} \) \( \mathbf{S}_2 \) = \( \frac{1}{r} \mathbf{C} - \frac{1}{rn} \mathbf{J} \) \( \mathbf{S}_0 \) = \( \frac{1}{rn} \mathbf{J} \) \( \mathbf{S}_3 \) = \( \mathbf{I} - \frac{1}{n} \mathbf{R} - \frac{1}{r} \mathbf{C} + \frac{1}{rn} \mathbf{J} \)
### Strata for analysis of variance

<table>
<thead>
<tr>
<th>Stratum</th>
<th>df</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1</td>
<td>$\xi_0$</td>
</tr>
<tr>
<td>Rows</td>
<td>$r - 1$</td>
<td>$\xi_1$</td>
</tr>
<tr>
<td>Columns</td>
<td>$n - 1$</td>
<td>$\xi_2$</td>
</tr>
<tr>
<td>Plots</td>
<td>$(r - 1)(n - 1)$</td>
<td>$\xi_3$</td>
</tr>
</tbody>
</table>

The expected sum of squares (SS) for treatments is:

$$E(\text{SS for treatments}) = Q + r(n - 1)\xi_2$$

where $Q$ is a positive-definite quadratic form in the treatment effects. Treatments are orthogonal to rows, so:

$$E(\text{SS for contrasts } \perp \text{to rows}) = Q + (n - 1)\xi_2 + (n - 1)(r - 1)\xi_3$$

Therefore:

$$E(\text{MS residual}) = E(\text{SS residual}) \left(1 + r - 1\right) = 1 \cdot \frac{r - 1}{r - 1} \left(\xi_3 - r\xi_2^2\right)$$
<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>$\xi_0$</td>
</tr>
<tr>
<td>rows</td>
<td>$r - 1$</td>
<td>$\xi_1$</td>
</tr>
<tr>
<td>columns</td>
<td>$n - 1$</td>
<td>$\xi_2$</td>
</tr>
<tr>
<td>plots</td>
<td>$(r - 1)(n - 1)$</td>
<td>$\xi_3$</td>
</tr>
</tbody>
</table>

}
### Strata for analysis of variance

<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>ξ₀</td>
</tr>
<tr>
<td>rows</td>
<td>r − 1</td>
<td>ξ₁</td>
</tr>
<tr>
<td>columns</td>
<td>n − 1</td>
<td>ξ₂</td>
</tr>
<tr>
<td>plots</td>
<td>(r − 1)(n − 1)</td>
<td>ξ₃</td>
</tr>
</tbody>
</table>

\[
E(\text{SS for anything}) = \text{SS}(E(\text{anything})) + \text{variance term}
\]
Strata for analysis of variance

<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>$\xi_0$</td>
</tr>
<tr>
<td>rows</td>
<td>$r - 1$</td>
<td>$\xi_1$</td>
</tr>
<tr>
<td>columns</td>
<td>$n - 1$</td>
<td>$\xi_2$</td>
</tr>
<tr>
<td>plots</td>
<td>$(r - 1)(n - 1)$</td>
<td>$\xi_3$</td>
</tr>
</tbody>
</table>

\[
E(\text{SS for anything}) = \text{SS}(E(\text{anything})) + \text{variance term}
\]

so \[E(\text{SS for treatments}) = Q + r(n - 1)V/2,\] where \(Q\) is a positive-definite quadratic form in the treatment effects.
Strata for analysis of variance

<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>$\zeta_0$</td>
</tr>
<tr>
<td>rows</td>
<td>$r - 1$</td>
<td>$\zeta_1$</td>
</tr>
<tr>
<td>columns</td>
<td>$n - 1$</td>
<td>$\zeta_2$</td>
</tr>
<tr>
<td>plots</td>
<td>$(r - 1)(n - 1)$</td>
<td>$\zeta_3$</td>
</tr>
</tbody>
</table>

\{ treatments \}

\{ residual \}

\[ E(\text{SS for anything}) = \text{SS}(E(\text{anything})) + \text{variance term} \]

so \[ E(\text{SS for treatments}) = Q + r(n - 1)\frac{V}{2}, \quad \text{where} \]

$Q$ is a positive-definite quadratic form in the treatment effects.

Treatments are orthogonal to rows, so (putting $y = \frac{\zeta_2}{\zeta_3}$)

\[ E(\text{SS for contrasts } \perp \text{ to rows}) = Q + (n - 1)\frac{\zeta_2}{\zeta_3} + (n - 1)(r - 1)\frac{\zeta_3}{\zeta_3} = \]

\[ \frac{15}{32} \]
Strata for analysis of variance

<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>$\bar{\xi}_0$</td>
</tr>
<tr>
<td>rows</td>
<td>$r - 1$</td>
<td>$\bar{\xi}_1$</td>
</tr>
<tr>
<td>columns</td>
<td>$n - 1$</td>
<td>$\bar{\xi}_2$</td>
</tr>
<tr>
<td>plots</td>
<td>$(r - 1)(n - 1)$</td>
<td>$\bar{\xi}_3$</td>
</tr>
</tbody>
</table>

$E(\text{SS for anything}) = \text{SS}(E(\text{anything})) + \text{variance term}$

so  
$E(\text{SS for treatments}) = Q + r(n - 1)V/2$,  
where  
$Q$ is a positive-definite quadratic form in the treatment effects. 

Treatments are orthogonal to rows, so (putting $y = \bar{\xi}_2 / \bar{\xi}_3$) 

$E(\text{SS for contrasts \perp to rows}) = Q + (n - 1)\bar{\xi}_2 + (n - 1)(r - 1)\bar{\xi}_3$

$= Q + (n - 1)(y + r - 1)\bar{\xi}_3$,  
so
Strata for analysis of variance

<table>
<thead>
<tr>
<th>stratum</th>
<th>df</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1</td>
<td>$\xi_0$</td>
</tr>
<tr>
<td>rows</td>
<td>$r-1$</td>
<td>$\xi_1$</td>
</tr>
<tr>
<td>columns</td>
<td>$n-1$</td>
<td>$\xi_2$</td>
</tr>
<tr>
<td>plots</td>
<td>$(r-1)(n-1)$</td>
<td>$\xi_3$</td>
</tr>
</tbody>
</table>

\[
E(\text{SS for anything}) = \text{SS}(E(\text{anything})) + \text{variance term}
\]

so \[E(\text{SS for treatments}) = Q + r(n-1)V/2, \quad \text{where} \]

$Q$ is a positive-definite quadratic form in the treatment effects.

Treatments are orthogonal to rows, so (putting $y = \frac{\xi_2}{\xi_3}$)

\[
E(\text{SS for contrasts } \perp \text{ to rows}) = Q + (n-1)\xi_2 + (n-1)(r-1)\xi_3 = Q + (n-1)(y + r - 1)\xi_3, \quad \text{so}
\]

\[
E(\text{MS residual}) = E \left( \frac{\text{SS residual}}{(n-1)(r-1)} \right) = \frac{1}{r-1} \left[ (y + r - 1)\xi_3 - \frac{rV}{2} \right].
\]
Overestimation of variance

The estimator of $V$ is $\hat{V} = \frac{2M}{r}$, where $M = $ MS residual.

We have shown that $E\left(\frac{2M}{r}\right) = \frac{2}{r-1} \left[ \frac{(y + r - 1)}{r} \xi_3 - \frac{V}{2} \right]$, so smaller $V \implies$ larger $\hat{V}$. 
Overestimation of variance

The estimator of $V$ is $\hat{V} = \frac{2M}{r}$, where $M = \text{MS residual}$.

We have shown that $E \left( \frac{2M}{r} \right) = \frac{2}{r - 1} \left[ \frac{(y + r - 1)}{r} \xi_3 - \frac{V}{2} \right]$, so smaller $V \implies$ larger $\hat{V}$.

Rewriting: $V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)(D - r^2)}{r^2(n - 1)} \right]$ with $D = \sum_{i=1}^{n} \lambda_{ii}$. 

Overestimation of variance

The estimator of $V$ is $\hat{V} = \frac{2M}{r}$, where $M = \text{MS residual}$.

We have shown that $E\left(\frac{2M}{r}\right) = \frac{2}{r-1} \left[ \frac{(y+r-1)}{r} \xi_3 - \frac{V}{2} \right]$, so smaller $V \implies$ larger $\hat{V}$.

Rewriting: $V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y-1)(D - r^2)}{r^2(n-1)} \right]$ with $D = \sum_{i=1}^{n} \lambda_{ii}$.

If simple restricted randomization forces no treatment to occur more than once per column, then $\lambda_{ii} = r$ for all $i$, so $D = rn$, so
Overestimation of variance

The estimator of $V$ is $\hat{V} = \frac{2M}{r}$, where $M = \text{MS residual}$. We have shown that $E\left(\frac{2M}{r}\right) = 2\left[\frac{(y + r - 1)}{r}\xi_3 - \frac{V}{2}\right]$, so smaller $V \implies$ larger $\hat{V}$.

Rewriting: $V = \frac{2\xi_3}{r} \left[1 + \frac{(y - 1)(D - r^2)}{r^2(n - 1)}\right]$ with $D = \sum_{i=1}^{n} \lambda_{ii}$.

If simple restricted randomization forces no treatment to occur more than once per column, then $\lambda_{ii} = r$ for all $i$, so $D = rn$, so

$$V = \frac{2\xi_3}{r} \left[1 + \frac{(y - 1)(n - r)}{r(n - 1)}\right]$$

and

$$\frac{2E(M)}{r} = \frac{2\xi_3}{r} \left[1 + \frac{(y - 1)n}{r(n - 1)}\right]$$

which over-estimates $V$ by $2(y - 1)\xi_3 / [r(n - 1)]$. 
Simple restricted randomization: summary

Keep re-randomizing until you get a plan with no treatment more than once in any column. Analyse as usual.

- Inefficient to produce plans: many will have to be rejected.
- Variance is overestimated:

\[
V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y-1)(n-r)}{r(n-1)} \right]
\]

and

\[
E(\hat{V}) = \frac{2E(M)}{r} = \frac{2\xi_3}{r} \left[ 1 + \frac{(y-1)n}{r(n-1)} \right]
\]
Simple restricted randomization: summary

Keep re-randomizing until you get a plan with no treatment more than once in any column. Analyse as usual.

- Inefficient to produce plans: many will have to be rejected.
- Variance is overestimated:

\[ V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)(n - r)}{r(n - 1)} \right] \]

and

\[ E(\hat{V}) = \frac{2E(M)}{r} = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)n}{r(n - 1)} \right] \]

- Genuine treatment differences may not be detected.
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column. Easy to do this directly, eg

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column. Easy to do this directly, eg

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column.
Easy to do this directly, eg

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column.
Easy to do this directly, eg

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column.
Easy to do this directly, eg

```
A  B  C  D  E  F  G
C  D  E  F  G  A  B
G  A  B  C  D  E  F
B  C  D  E  F  G  A
D  E  F  G  A  B  C
```

Randomize rows, columns, treatments.
Use a Latinized design, but analyse as usual

Deliberately construct a design in which no treatment occurs more than once in any column. Easy to do this directly, eg

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

Randomize rows, columns, treatments.

Same bias in estimator of variance as for simple restricted randomization, so there is a loss of power and genuine treatment differences may not be detected.
Super-valid restricted randomization

- Needs tables of designs.
Super-valid restricted randomization

- Needs tables of designs.
- Randomize rows, columns and treatments.
Super-valid restricted randomization

- Needs tables of designs.
- Randomize rows, columns and treatments.
- Analyse as usual.
Super-valid restricted randomization

- Needs tables of designs.
- Randomize rows, columns and treatments.
- Analyse as usual.
- **Same** average variance as in randomized complete-block design, but with **smaller range**.
Super-valid restricted randomization

- Needs tables of designs.
- Randomize rows, columns and treatments.
- Analyse as usual.
- **Same** average variance as in randomized complete-block design, but with **smaller range**.
- The estimator of variance is unbiased when averaged over all comparisons in this one experiment.
Condition for unbiased estimator of variance

We have

\[ V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)(D - r^2)}{r^2(n - 1)} \right] \]

with \( D = \sum_{i=1}^{n} \lambda_{ii} \),

and

\[ E(\hat{V}) = \frac{2}{r - 1} \left[ \frac{(y + r - 1)}{r} \xi_3 - \frac{V}{2} \right], \]
Condition for unbiased estimator of variance

We have

\[ V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)(D - r^2)}{r^2(n - 1)} \right] \]

with \( D = \sum_{i=1}^{n} \lambda_{ii} \),

and \( E(\hat{V}) = \frac{2}{r - 1} \left[ \frac{(y + r - 1)}{r} \xi_3 - \frac{V}{2} \right] \),

So \( V = E(\hat{V}) \iff D = r(r + n - 1). \)
We have \[ V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y-1)(D-r^2)}{r^2(n-1)} \right] \]

with \( D = \sum_{i=1}^{n} \lambda_{ii} \),

and \[ E(\hat{V}) = \frac{2}{r-1} \left[ \frac{(y+r-1)}{r} \xi_3 - \frac{V}{2} \right], \]

So \( V = E(\hat{V}) \iff D = r(r+n-1). \)

If treatment \( i \) occurs twice in any column then \( \lambda_{ii} \) increases by \( 2^2 - 2 \times 1^2 = 2 \).
Condition for unbiased estimator of variance

We have

\[ V = \frac{2\xi_3}{r} \left[ 1 + \frac{(y - 1)(D - r^2)}{r^2(n - 1)} \right] \]

with

\[ D = \sum_{i=1}^{n} \lambda_{ii}, \]

and

\[ E(\hat{V}) = \frac{2}{r - 1} \left[ \frac{(y + r - 1)}{r} \xi_3 - \frac{V}{2} \right], \]

So

\[ V = E(\hat{V}) \iff D = r(r + n - 1). \]

If treatment \( i \) occurs twice in any column then \( \lambda_{ii} \) increases by \( 2^2 - 2 \times 1^2 = 2 \).

If each pair of rows has one column with the same treatment but no treatment occurs more than twice in any column then

\[ D = r(r - 1) + rn = r(r + n - 1) \]

and so \( V = E(\hat{V}) \).
1. In every pair of rows, there is exactly one column in which the two treatments are the same.
A design from the tables

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>G</td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>G</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>B</td>
<td>G</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>D</td>
<td>G</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>A</td>
<td>F</td>
</tr>
</tbody>
</table>

1. In every pair of rows, there is exactly one column in which the two treatments are the same.
2. No treatment occurs more than twice in any column.
A design from the tables

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>G</td>
</tr>
<tr>
<td>A</td>
<td>G</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>B</td>
<td>G</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>A</td>
<td>F</td>
</tr>
</tbody>
</table>

1. In every pair of rows, there is exactly one column in which the two treatments are the same.
2. No treatment occurs more than twice in any column.
3. If $m_i =$ the number of columns in which treatment $i$ occurs twice, then $m_i - m_j \in \{-1, 0, 1\}$ for all other treatments $j$. 
1. In every pair of rows, there is exactly one column in which the two treatments are the same.

2. No treatment occurs more than twice in any column.

3. If \( m_i \) = the number of columns in which treatment \( i \) occurs twice, then \( m_i - m_j \in \{-1, 0, 1\} \) for all other treatments \( j \).

4. Subject to conditions (1)–(3), the spread of the variances of the estimators of simple treatment differences is as small as possible.
Pairwise variances in the example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>G</td>
</tr>
<tr>
<td>D</td>
<td>A</td>
<td>G</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>B</td>
<td>G</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>A</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

Minimum

\[ V_{AD} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{2}{5}\tau \right] \]

Maximum

\[ V_{AB} = \frac{2\sigma^2}{5} \left[ 1 - \rho + \frac{2}{5}\tau \right] \]

Average

\[ V = \frac{2\sigma^2}{5} (1 - \rho) \]
Pairwise variances in the example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>G</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>E</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>B</td>
<td>G</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>A</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

Minimum  \( V_{AD} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{2}{5} \tau \right] \cdots - \frac{4}{5} \tau \)

Maximum  \( V_{AB} = \frac{2\sigma^2}{5} \left[ 1 - \rho + \frac{2}{5} \tau \right] \cdots + \tau \)

Average  \( V = \frac{2\sigma^2}{5} (1 - \rho) \cdots - \frac{1}{15} \tau \)
Pairwise variances in the example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>G</td>
<td>F</td>
<td>B</td>
<td>C</td>
<td>E</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>B</td>
<td>G</td>
<td>F</td>
<td>C</td>
<td>A</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>A</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

Minimum: \( V_{AD} = \frac{2\sigma^2}{5} \left[ 1 - \rho - \frac{2}{5}\tau \right] \cdots - \frac{4}{5}\tau \)

Maximum: \( V_{AB} = \frac{2\sigma^2}{5} \left[ 1 - \rho + \frac{2}{5}\tau \right] \cdots + \tau \)

Average: \( V = \frac{2\sigma^2}{5} (1 - \rho) \cdots - \frac{1}{15}\tau \cdots - \frac{2}{3}\tau \)
Super-valid restricted randomization: summary

- Needs tables of designs.
- Randomize rows, columns and treatments.
- Analyse as usual.
- Same average variance as in randomized complete-block design, but with smaller range.
- The estimator of variance is unbiased when averaged over all comparisons in this one experiment.
Super-valid restricted randomization: summary

- Needs tables of designs.
- Randomize rows, columns and treatments.
- Analyse as usual.
- Same average variance as in randomized complete-block design, but with smaller range.
- The estimator of variance is unbiased when averaged over all comparisons in this one experiment.
- There is no separate estimate of $\rho$ or $\tau$ (or $y$), so treatments must be randomized and a single standard error given for all differences.
Efficient row-column designs

- Needs tables of designs.
Efficient row-column designs

- Needs tables of designs.
- Randomize rows and columns.

More complicated analysis (should be available in software).

Average variance may be less than, or more than, the average variance in randomized complete-block design, depending on the size of the correlations.

Unbiased estimator of the variance of every treatment contrast.
Efficient row-column designs

- Needs tables of designs.
- Randomize rows and columns.
- More complicated analysis (should be available in software).
Efficient row-column designs

- Needs tables of designs.
- Randomize rows and columns.
- More complicated analysis (should be available in software).
- Average variance may be less than, or more than, the average variance in randomized complete-block design, depending on the size of the correlations.
Efficient row-column designs

- Needs tables of designs.
- Randomize rows and columns.
- More complicated analysis (should be available in software).
- Average variance may be less than, or more than, the average variance in randomized complete-block design, depending on the size of the correlations.
- Unbiased estimator of the variance of every treatment contrast.
Given an incomplete-block design for \( n \) treatments in \( n \) blocks of size \( r \), define the number \( A (0 < A < 1) \), depending on the design, by

\[
A = \frac{2\sigma^2}{rV}
\]

if the analysis uses information orthogonal to blocks.
Columns form an Incomplete-block design (IBD)

Given an incomplete-block design for $n$ treatments in $n$ blocks of size $r$,
define the number $A$ ($0 < A < 1$), depending on the design, by

$$A = \frac{2\sigma^2}{rV}$$

if the analysis uses information orthogonal to blocks.

Choose the **optimal** IBD: the one with the largest value of $A$. 

Hall's Marriage Theorem $\implies$ the blocks of this IBD can be arranged as the columns of a row-column design so that each treatment occurs once in each row.

Randomize rows and columns.
Given an incomplete-block design design for \( n \) treatments in \( n \) blocks of size \( r \),
define the number \( A (0 < A < 1) \), depending on the design, by

\[
A = \frac{2\sigma^2}{rV}
\]

if the analysis uses information orthogonal to blocks.

Choose the **optimal** IBD: the one with the largest value of \( A \).

Hall’s Marriage Theorem \( \implies \) the blocks of this IBD can be arranged as the columns of a row-column design so that each treatment occurs once in each row.
Given an incomplete-block design for \( n \) treatments in \( n \) blocks of size \( r \), define the number \( A (0 < A < 1) \), depending on the design, by

\[
A = \frac{2\sigma^2}{rV}
\]

if the analysis uses information orthogonal to blocks.

Choose the optimal IBD: the one with the largest value of \( A \).

Hall’s Marriage Theorem \( \implies \) the blocks of this IBD can be arranged as the columns of a row-column design so that each treatment occurs once in each row.

Randomize rows and columns.
Analyse by fitting rows, columns and treatments.

$$E(\text{MS residual}) = \xi_3$$

$$V_{ij} = \frac{2\xi_3}{rA_{ij}}$$

where $A_{ij}$ is known from the design

$$V = \frac{2\xi_3}{rA} = \frac{2\sigma^2}{rA} (1 - \rho - \tau)$$
Example of a row-column design

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

\[ V_{AB} = 1.044 \times \frac{2}{5} \xi_3 \]
\[ V_{AC} = 1.089 \times \frac{2}{5} \xi_3 \]
\[ V_{AD} = 1.091 \times \frac{2}{5} \xi_3 \]
\[ V = 1.075 \times \frac{2}{5} \xi_3 \]
Example of a row-column design

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>F</td>
<td>G</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

\[
V_{AB} = 1.044 \times \frac{2}{5} \xi_3 \quad \text{normal method}
\]

\[
V_{AC} = 1.089 \times \frac{2}{5} \xi_3 \quad V = \frac{2}{5} \left( \frac{y + 4}{5} \right) \xi_3
\]

\[
V_{AD} = 1.091 \times \frac{2}{5} \xi_3 \quad \text{averaged over randomizations}
\]

\[
V = 1.075 \times \frac{2}{5} \xi_3 \quad \text{N.B. } y = \frac{\xi_2}{\xi_3} \geq 1
\]
Efficient row-column designs: summary

- Needs tables of designs.
- Randomize rows and columns.
- More complicated analysis (should be available in software).
- Average variance may be less than, or more than, the average variance in randomized complete-block design, depending on the size of the correlations.
- Unbiased estimator of the variance of every treatment contrast.
Efficient row-column designs: summary

- Needs tables of designs.
- Randomize rows and columns.
- More complicated analysis (should be available in software).
- Average variance may be less than, or more than, the average variance in randomized complete-block design, depending on the size of the correlations.
- Unbiased estimator of the variance of every treatment contrast.
- There is no need to randomize treatments; the most important differences can be given the lowest variance.
Comparing super-valid restricted randomization and efficient row-column designs

\[
y = \frac{\xi_2}{\xi_3} = \frac{\text{columns stratum variance}}{\text{plots stratum variance}} \geq 1 \quad \text{(we believe)}
\]
Comparing super-valid restricted randomization and efficient row-column designs

\[ y = \frac{\tau_2}{\tau_3} = \frac{\text{columns stratum variance}}{\text{plots stratum variance}} \geq 1 \quad \text{(we believe)} \]

The best row-column design is more efficient than super-valid restricted randomization if and only if \( y \) exceeds the following value.
Comparing super-valid restricted randomization and efficient row-column designs

\[ y = \frac{\xi_2}{\xi_3} \geq 1 \quad \text{(we believe)} \]

The best row-column design is more efficient than super-valid restricted randomization if and only if \( y \) exceeds the following value.

<table>
<thead>
<tr>
<th>( r )</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.68</td>
<td>1.83</td>
<td>1.86</td>
<td>2.02</td>
<td>2.13</td>
<td>2.27</td>
</tr>
<tr>
<td>4</td>
<td>1.26</td>
<td>1.47</td>
<td>1.57</td>
<td>1.71</td>
<td>1.80</td>
<td>1.86</td>
</tr>
<tr>
<td>5</td>
<td>1.21</td>
<td>1.37</td>
<td>1.48</td>
<td>1.58</td>
<td>1.64</td>
<td></td>
</tr>
</tbody>
</table>
Use a carefully chosen Latinized design with REML/ANOVA estimates of variance components

Choose a design with the $\lambda_{ij}$ as equal as possible. Randomize rows and columns. Estimate treatment differences from the usual randomized-complete-blocks analysis.

$E(\text{MS residual from complete-block analysis}) = \xi_3 + \xi_2 - \xi_3 r (n - 1)$

$E(\text{MS residual from row-column analysis}) = \xi_3 Hence unbiased estimators of $\xi_2$ and $\xi_3$. But this estimator of $V=\frac{2r}{\xi_3 + (n-r)(\xi_2 - \xi_3)}$ does not have a $\chi^2$ distribution, so how do we do hypothesis tests? Also, there are so few effective df for $\xi_2$ that these estimates have very poor precision.
Use a carefully chosen Latinized design with REML/ANOVA estimates of variance components

Choose a design with the $\lambda_{ij}$ as equal as possible. Randomize rows and columns. Estimate treatment differences from the usual randomized-complete-blocks analysis.

$$E(\text{MS residual from complete-block analysis}) = \xi_3 + \frac{\xi_2 - \xi_3}{r(n - 1)}$$

$$E(\text{MS residual from row-column analysis}) = \xi_3$$
Use a carefully chosen Latinized design with REML/ANOVA estimates of variance components

Choose a design with the $\lambda_{ij}$ as equal as possible. Randomize rows and columns.
Estimate treatment differences from the usual randomized-complete-blocks analysis.

$$E(\text{MS residual from complete-block analysis}) = \xi_3 + \frac{\xi_2 - \xi_3}{r(n-1)}$$

$$E(\text{MS residual from row-column analysis}) = \xi_3$$

Hence unbiased estimators of $\xi_2$ and $\xi_3$ and of

$$V = \frac{2}{r} \left[ \xi_3 + \frac{(n-r)(\xi_2 - \xi_3)}{r(n-1)} \right].$$
Use a carefully chosen Latinized design with REML/ANOVA estimates of variance components

Choose a design with the $\lambda_{ij}$ as equal as possible. Randomize rows and columns. Estimate treatment differences from the usual randomized-complete-blocks analysis.

\[
E(\text{MS residual from complete-block analysis}) = \zeta_3 + \frac{\bar{\zeta}_2 - \bar{\zeta}_3}{r(n-1)}
\]

\[
E(\text{MS residual from row-column analysis}) = \zeta_3
\]

Hence unbiased estimators of $\zeta_2$ and $\zeta_3$ and of

\[
V = \frac{2}{r} \left[ \zeta_3 + \frac{(n-r)(\bar{\zeta}_2 - \bar{\zeta}_3)}{r(n-1)} \right].
\]

But this estimator of $V$ does not have a $\chi^2$ distribution, so how do we do hypothesis tests? Also, there are so few effective df for $\zeta_2$ that these estimates have very poor precision.
Use an efficient row-column design and analyse with combination of information.

Choose an optimal row-column design as before.
Use an efficient row-column design and analyse with combination of information

Choose an optimal row-column design as before.

Analyse using
- ANOVA with combination of information across strata, or
- REML, or
- mixed-model software.
Choose an optimal row-column design as before.

Analyse using

- ANOVA with combination of information across strata, or
- REML, or
- mixed-model software.

These should all be equivalent, but different implementations can give different results.
Choose an optimal row-column design as before.

Analyse using

- ANOVA with combination of information across strata, or
- REML, or
- mixed-model software.

These should all be equivalent, but different implementations can give different results.

As before, there are so few effective df for $\zeta_2$ that these estimates have very poor precision, and so are not recommended.


