RANDOMIZED COMPLETE BLOCK DESIGN (CBD)

- As with CRD, t unstructured treatments
- b blocks, each of size t, one unit per treatment per block
- Pairs of units from the same block are thought to be "more alike" or more homogeneous, than pairs from different blocks
- $N = bt$ total units and observations

	treat. 1	treat. 2	...	treat. t
block 1	$y_{1,1}$	$y_{1,2}$	...	$y_{1,t}$
block 2	$y_{2,1}$	$y_{2,2}$	...	$y_{2,t}$
...	...	...	...	...
block b	$y_{b,1}$	$y_{b,2}$	...	$y_{b,t}$

- Ordinarily reflects model and assumptions for 2-way ANOVA

- Sometimes blocking is necessary, e.g. biology studies in which animals from the *same litter* are more alike than animals from *different litters*, but more than one litter is needed.
- Sometimes blocking is done by choice, e.g. industrial experiments where *lab-scale* properties of materials give more homogeneous results than *plant-scale*, but more runs are needed than can be produced in one lab "batch".
- Either way, the point is to COMPARE TREATMENTS across homogeneous units (to maximize experimental control).
- What's lacking here? True REPLICATION ... i.e. every two data values differ by at least one of block or treatment

- Suppose each block-and-treatment combine in a unique way to produce a general (unrestricted) mean:

$$y_{i,j} = \mu_{i,j} + \epsilon_{i,j},\ i = 1...b,\ j = 1...t$$

- No replication, one mean-parameter per observation, no ability to estimate σ:

source	df
Model|Mean	$bt - 1$
Residual	0
Mean Correction	1
Uncorrected Total	bt

- No formal inference is possible unless some additional *assumptions* are made ... d.f. "moved" from *model* to *residual*
- The usual model assumes additive structure between blocks and treatment effects (i.e. no interaction). Can check with, e.g.:
  - Tukey's "one degree of freedom test"
  - Simple graphics using $y_{i,j} - y_{i,j'} - y_{i',j} + y_{i',j'}$
    - "4 corners" contrasts from 2-way table
    - Should have $E = 0$, $Var = 4\sigma^2$ ...
    - $\binom{b}{2} \times \binom{t}{2}$ of these ... a very large number, and clearly not independent
    - Usually looking for atypical values with common i or j ...

Notes:

- The above remark about the "usual model assumes ... no interaction" is based on the blocks being treated as *fixed effects*.
- Alternatively, if blocks are regarded as *random effects*, and the block×treatment interaction (also a random effect) is included, the appropriate denominator SS is associated with this interaction.
- Note that the arithmetic is actually the same either way ... the $(b-1)(t-1)$ degree-of-freedom denominator can be interpretted as representing:
  - residual, in a model with fixed blocks and no interaction
  - (random) interaction, in a model with random blocks
- In more complex designs, the decision to call blocks "fixed" or "random" is actually more important than you may think. We will come back to this; for now, blocks are "fixed" unless otherwise stated.

Now consider models:

- $y_{i,j} = \beta_i + \tau_j + \epsilon_{i,j}$ or equivalently $\alpha + \beta_i + \tau_j + \epsilon_{i,j}$
- $i = 1...b, \quad j = 1...t$
- Ordering observations by block, then treatment within block

$$\mathbf{X} = \left(\begin{array}{c|cccc|c} \mathbf{1} & \mathbf{1} & & \ldots & & \mathbf{I} \\ \mathbf{1} & & \mathbf{1} & \ldots & & \mathbf{I} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{1} & & & \ldots & \mathbf{1} & \mathbf{I} \end{array}\right)$$

- This is the model matrix for the second model, drop first column for the first:

$$\mathbf{X} = \left(\begin{array}{cccc|c} \mathbf{1} & & \ldots & & \mathbf{I} \\ & \mathbf{1} & \ldots & & \mathbf{I} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ & & \ldots & \mathbf{1} & \mathbf{I} \end{array}\right)$$

- There is one linear dependency among columns with first model, or two with second
- Inferences of interest center on $\boldsymbol{\tau}$, regarding $(\alpha, \boldsymbol{\beta})$ as nuisance parameters ... Set up a partitioned model:
  - put $\boldsymbol{\tau}$-columns in $\mathbf{X}_2$
  - put $\boldsymbol{\beta}$-columns (and optionally the α-column) in $\mathbf{X}_1$

LEAST-SQUARES ESTIMATION OF $\boldsymbol{\tau}$ in PARTITIONED MODEL

- $\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} \; = \; \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\boldsymbol{\tau} + \boldsymbol{\epsilon}$
- Here, $\boldsymbol{\beta} = (\alpha, \beta_1, \beta_2, ...\beta_b)$, or without α
- Full normal equations:

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\theta}} = \mathbf{X}'\mathbf{y}$$

- Reduced normal equations:

$$\mathbf{X}_2'(\mathbf{I} - \mathbf{H}_1)\mathbf{X}_2\hat{\boldsymbol{\tau}} = \mathbf{X}_2'(\mathbf{I} - \mathbf{H}_1)\mathbf{y}$$

- Similar to the development of $\mathbf{H}_A$ for a CRD:

$$\mathbf{H}_1 = \frac{1}{t}\begin{pmatrix} \mathbf{J}_{t\times t} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{t\times t} & ... & \mathbf{0} \\ ... & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{J}_{t\times t} \end{pmatrix} \quad (b \text{ diagonal blocks})$$

  since the blocking structure divides data into b groups of size t.

- So,

$$\mathbf{X}_2'\mathbf{H}_1 = \frac{1}{t}\begin{pmatrix} \mathbf{J}_{t\times t} & \mathbf{J}_{t\times t} & ... & \mathbf{J}_{t\times t} \end{pmatrix} = \frac{1}{t}\mathbf{J}_{t\times bt}$$

  because $\mathbf{X}_2' = (\mathbf{I}, \mathbf{I}, \mathbf{I}, ..., \mathbf{I})$, b times.

- What if α had been the ONLY nuisance parameter? (i.e. CRD)

$$\mathbf{X}_1 = \mathbf{1}_N \quad \mathbf{H}_1 = \tfrac{1}{N}\mathbf{J}_{N\times N}$$

- Then,

$$\mathbf{X}_2'\mathbf{H}_1 = \tfrac{1}{N}(b\mathbf{J}_{t\times N}) = \tfrac{1}{t}\mathbf{J}_{t\times bt}$$

- The SAME THING ... RNE's are the same for CBD's and CRD's, so $\widehat{\mathbf{c}'\boldsymbol{\tau}}$ is computed *as if there are no blocks*
- This is a result of orthogonality; each treatment-block combination appears the same number of times (once)

More about this:

- The RNE's are the same for a CBD and CRD (where $n_i = b$) because, apart from a rearrangement of the rows in the model matrix:
  - $\mathbf{X}_2$ is the same for each design (e.g. any given treatment receives the same number of units in each design), and
  - $\mathbf{H}_1\mathbf{X}_2$ is the same for each design
- The book says two designs satisfy "Condition E" for "*equivalent*" when they have this relationship.
- Other blocking arrangements are also "equivalent" to CRD's in this sense, e.g. "augmented" block designs ... see discussion in Chapter 4.

- So, for either CRD or CBD

$$\mathbf{X}_2'(\mathbf{I}-\mathbf{H}_1)\mathbf{X}_2\hat{\boldsymbol{\tau}} = \mathbf{X}_2'(\mathbf{I}-\mathbf{H}_1)\mathbf{y}$$

$$(\mathbf{X}_2'\mathbf{X}_2 - \tfrac{1}{t}\mathbf{J}_{t\times N}\mathbf{X}_2)\hat{\boldsymbol{\tau}} = \mathbf{X}_2'\mathbf{y} - \tfrac{1}{t}\mathbf{J}_{t\times N}\mathbf{y}$$

$$(b\mathbf{I} - \tfrac{b}{t}\mathbf{J}_{t\times t})\hat{\boldsymbol{\tau}} = \begin{pmatrix} y_{.1} \\ y_{.2} \\ \cdots \\ y_{.t} \end{pmatrix} - \tfrac{1}{t}\begin{pmatrix} y_{..} \\ y_{..} \\ \cdots \\ y_{..} \end{pmatrix}$$

$$\hat{\boldsymbol{\tau}} - \bar{\hat{\tau}}\mathbf{1} = \begin{pmatrix} \bar{y}_{.1} \\ \bar{y}_{.2} \\ \cdots \\ \bar{y}_{.t} \end{pmatrix} - \begin{pmatrix} \bar{y}_{..} \\ \bar{y}_{..} \\ \cdots \\ \bar{y}_{..} \end{pmatrix}$$

$$\hat{\tau}_j - \bar{\hat{\tau}} = \bar{y}_{.j} - \bar{y}_{..}$$

- Terms on left are non-unique because the same constant can be added or subtracted from each $\hat{\tau_j}$ in one solution to get another
- Can restrict the system (select a specific solution) by requiring:
  - $\bar{\hat{\tau}} = 0 \quad \rightarrow \quad \hat{\tau_j} = \bar{y}_{.j} - \bar{y}_{..}$
  - $\hat{\tau}_t = 0 \quad \rightarrow \quad \hat{\tau_j} = \bar{y}_{.j} - \bar{y}_{.t}$

  ... this is really equivalent to selecting a specific g-inverse
- Estimable functions are $\sum c_j \tau_j$ such that $\sum c_j = 0$
- For any estimable function, $Var[\sum c_j \hat{\tau_j}] = \sigma^2 \sum c_j^2 \frac{1}{b}$ ... same formula for CRD or CBD

So, why is CBD better (at least sometimes)?

- Suppose ϵ represents primarily differences in physical units ...
- CRD:
  - Selecting a larger sample of units requires drawing from a less homogeneous population (with a larger variance)
  - $N = bt$ units drawn from a population with $Var[\epsilon] = \sigma^2_{CRD}$
- CBD:
  - Selecting smaller samples of units allows drawing from more homogeneous populations (with a smaller variance)
  - b groups of t units, each drawn from a (different) population with $Var[\epsilon] = \sigma^2_{CBD}$

- e.g. can perhaps choose between selecting:
  - 50 litters of animals, 6 animals from each litter
  - 300 animals from the general population
- or:
  - mix 5 batches of chemicals in lab on each of 20 days
  - pick 100 batches of chemicals from production facility

Expected squared length of CI for $\sum c_j \tau_j$:

- CI is: estimate $\pm\ t \times \sqrt{MSE \times \sum c_j^2/b}$
- Interval length is $2 \times t \times \sqrt{MSE \times \sum c_j^2/b}$
- Expected squared length (to avoid $E\sqrt{MSE}$) is:
  - CRD: $4t^2(1-\frac{\alpha}{2}, N-t)\sigma^2_{CRD}(\sum c_j^2/b)$
  - CBD: $4t^2(1-\frac{\alpha}{2}, N-b-t+1)\sigma^2_{CBD}(\sum c_j^2/b)$
- CBD quantity is smaller if:

  $$\sigma_{CBD}/\sigma_{CRD} < t(1-\tfrac{\alpha}{2}, N-t)/t(1-\tfrac{\alpha}{2}, N-b-t+1)$$

  and in this case, CI's are "tighter" on average than for CRD.
- t-ratio is only slightly < 1 unless t and b are small

Power of F-test for $\tau_1 = \tau_2 = ... = \tau_t$. Briefly,

- $E[$ SSE(Hyp_A) $] = \sigma^2_{CBD}(N - b - t + 1)$
- $E[$ SSE(Hyp_0) $] = (\boldsymbol{\beta}'\mathbf{X}_1' + \boldsymbol{\tau}'\mathbf{X}_2')(\mathbf{I} - \mathbf{H}_0)(\mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\boldsymbol{\tau}) + \sigma^2_{CBD}(N - b)$
  - d.f. for both SSE's are different than with CRD
  - $\mathbf{H}_0$ is different here than with CRD, *but*
  - $\mathbf{X}_1$ terms drop out for the same reason as before, *and*
  - $\mathbf{X}_2'\mathbf{H}_0$ is as with the CRD, *so*
- Quadratic form is still

  $$Q_{2|1}(\boldsymbol{\tau}) = \boldsymbol{\tau}'\mathbf{X}_2'(\mathbf{I} - \mathbf{H}_0)\mathbf{X}_2\boldsymbol{\tau} = \textstyle\sum_{j=1}^{t} b(\boldsymbol{\tau}_j - \bar{\tau}_.)^2$$

  as with CRD

- $[\text{SST}/(t-1)]/[\text{SSE}(Hyp_A)/(N-b-t+1)] \sim F'(t-1,\ N-b-t+1,\ Q_{2|1}(\boldsymbol{\tau})/\sigma^2_{CBD})$
- Compare to $F'(t-1,\ N-t,\ \ Q_{2|1}(\boldsymbol{\tau})/\sigma^2_{CRD})$ for CRD ... for candidate values of σ's and τ's
- In either case, the power is the probability of seeing a larger F than the critical value determined under the central F distribution ... calculations with SAS or R as discussed before for CRD
- Note that the critical value will be somewhat smaller for the CRD due to an increase in denominator d.f. ... Other things being equal, this would give the CRD slightly more power, but this is usually offset if σ_{CBD} is even slightly less than σ_{CRD}