Experimental Design 1: Blocking

Outline:

1. Randomization Principles
2. Randomized Complete Block Design (RCB)
3. Unbalanced RCB analysis (missing data)
4. Latin Square Designs (LS)
5. Thoughts on these designs

Examples:

1. Flax Example: Balanced RCB
2. Media Example: Unbalanced RCB
3. Latin Square Example
4. Graeco Latin Square Example

Some Perspective:

"The **experiment design** is the arrangement of experimental units used to control experimental error and at the same accommodate the **treatment design** in an experiment."
Robert Kuehl "Statistical Principles of Research Design and Experiment"

Recall (from CH6) that we considered two different approaches for comparing means for 2 treatments (H_0: $\mu_1=\mu_2$):

- Two-sample t-test (or Two-sample Wilcoxon test) when we have independent observations.
- Paired t-test (or Paired Wilcoxon test) when we have paired observations.

The appropriate analysis depends on the design.

Blocking extends the idea of pairing to cases where there are more than two treatments.

In this group of notes, we will talk about experimental designs when there is just a single factor (group of trts):
Completely Randomized Design (CRD)
Randomized Complete Block Design (RCB)
Latin Square Design (LS)

Keep in mind that research question (and treatment design) is really the same for this group of notes. We are interested in comparing means for a single factor.

So for a given treatment design, there can be different experimental designs.

1. Randomization Principles

Research goal is to compare 2 trts (A, B) $\Rightarrow H_0: \mu_A = \mu_B$

Unrestricted Randomization

A	B	A
A	B	B
B	A	B
A	B	A

Restricted (Paired) Randomization

A	A	B
B	B	A
B	A	A
A	B	B

Unrestricted randomization: 6 A's and 6 B's randomly assigned to plots, without restriction. Use two-sample t-test.

Paired randomization: within each of six pairs of plots, a coin flip decides which plot receives A and which plot receives B. Use paired t-test.

Two reasons for randomization:

1. Randomization assures unbiasedness (easy to see).

Randomization assures unbiasedness in the estimation of treatment differences, because each treatment has an equal chance of receiving the high-response and low-response experimental units.

2. Randomization justifies correct analysis (harder to see).

Randomization test:
Idea: If H0 (treatments have same effects) is true, then treatment assignment is just like assigning labels.
1. Compute the value of the test statistic (t-test or F-test) that you would have observed under all possible outcomes of random assignment of treatments to units.
2. Compute the randomization test p-value as the proportion of outcomes that have |t| or F larger than the actual observed value.

In our example:

1. If the randomization was done in a paired fashion, there are $2^6 = 64$ possible outcomes of randomization. (smallest pval?)

2. If the randomization was done without restriction ("Completely Randomized"), then there are
$$\frac{12!}{6!6!} = 924$$ possible outcomes to randomization.

Therefore: The randomization test p-value will depend on which type of randomization was done.

An important result about randomization test p-values: If the data are not too skewed, or have outliers:
a) The unpaired randomization test p-value is approximated by the two-sample t-test.
b) The paired randomization test p-value is approximated by the paired t-test.

Conclusions About Randomization:

1. If model assumptions hold (ex: normality), normal theory t-tests are acceptable. But randomization is generally done as a precaution against undetected deficiencies.

2. If model assumptions are in doubt and if you do randomize, you can use the randomization test. (As long as there are enough possible outcomes to the randomization.)

   We usually do the normal-theory t-tests, but can consider them as approximations to the randomization test.

   We now carry these ideas into more complicated situations.

2. Randomized Complete Block (RCB) Designs

We now consider a situation analogous to the paired randomization, but with four treatments (A, B, C, D).

Unrestricted Randomization

A	B	A
D	A	D
B	C	C
D	C	B

Restricted (blocked) Randomization

A	B	D
C	C	A
B	D	B
D	A	C

Unrestricted randomization (CRD): 3 reps of each treatment assigned to 12 plots, without restriction ⇒ One-way ANOVA analysis

Restricted Randomization (RCB): Plots grouped into 3 "blocks" of 4 plots. Treatments assigned randomly, with the restriction that each treatment occur exactly once in each block ⇒ RCB analysis

Idea of Blocking:

Assign experimental units (EU's) to blocks so that EU's in the same block are **as alike as possible**. Some blocks will have higher than average values; some blocks will have lower than average values. Comparison of treatment means will be more accurate, because we compare within block and each treatment has an equal number (=1) of observations in each block.

"Block what you can, randomize what you cannot."

Blocking is used to account for the effects of a few of the most important nuisance variables. Randomization is then used to reduce the contaminating effects of remaining nuisance variables.

Example: Greenhouse study on tomato seedlings.

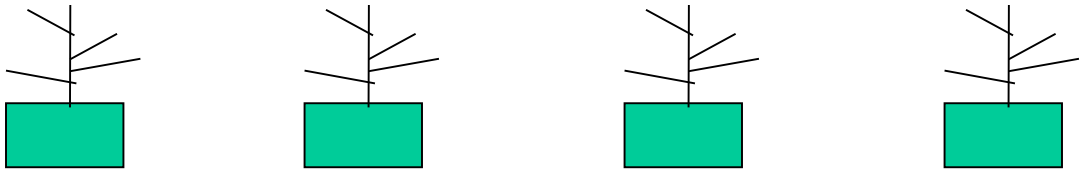
Objective: Compare 4 growing media (A, B, C, D)
using 3 reps (3 x 4=12 pots)

Method: Plant 12 seedlings in pots, three of each growing media.

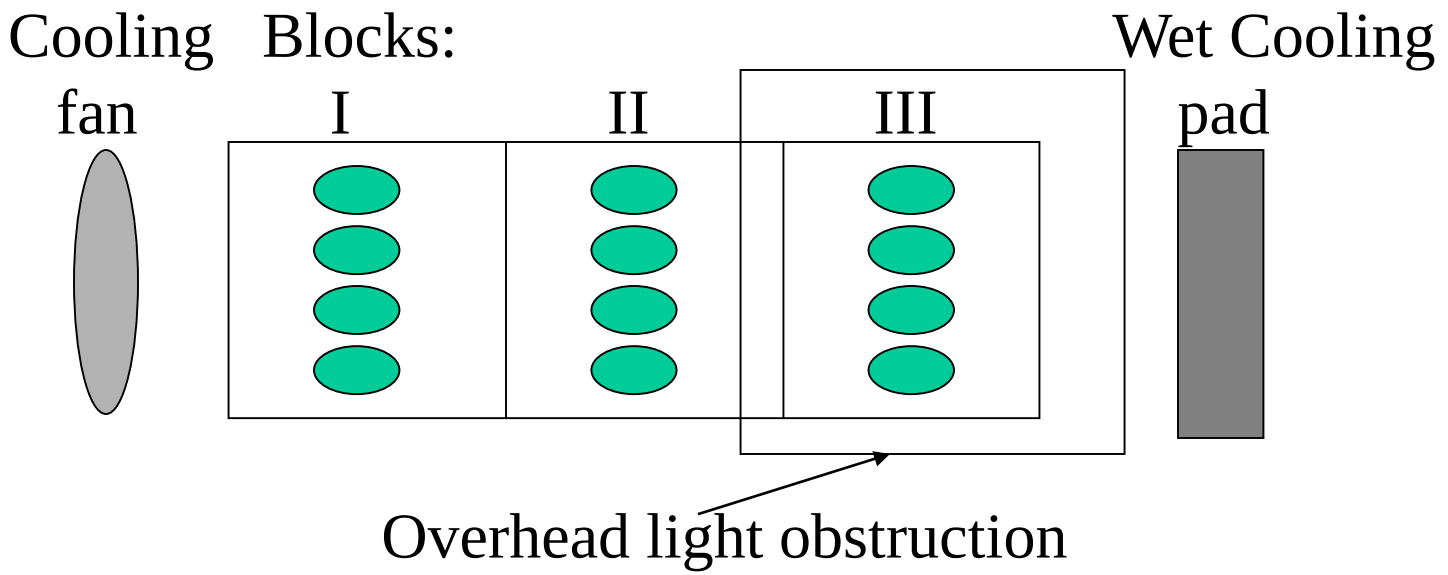
Blocking of seedlings: a) 4 largest seedlings = block 1
b) next 4 largest = block 2
c) 4 smallest = block 3

Randomly assign one of each media to each block.

Block 1:



Placement of pots in the greenhouse:



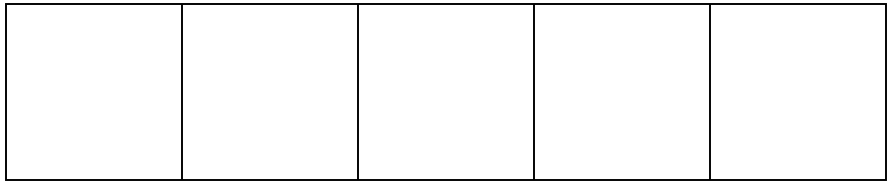
Lab analysis: analyze all samples from the same block in the same batch.

Summary: All conditions, except treatment, should be as alike as possible within a block. Differences should be, as much as possible, between blocks).

a) Seedling height (or health).
b) Environment in the greenhouse.
c) Analytical errors.

Other blocking examples:

1. Fertility in a field trial:

High  Low

2. Irrigation in a field trial (same picture as above, but with furrow irrigation flowing left to right)

3. Animal drug trials: block = litter.

4. Animal feeding trials: block = groups of animals with similar weights.

5. Subjective evaluations: block = rater or analyst.

6. Comparison of automobile fuels: block = driver.

7. If all samples cannot be run in a single day, can block by lab run.

Notes about randomized blocks:

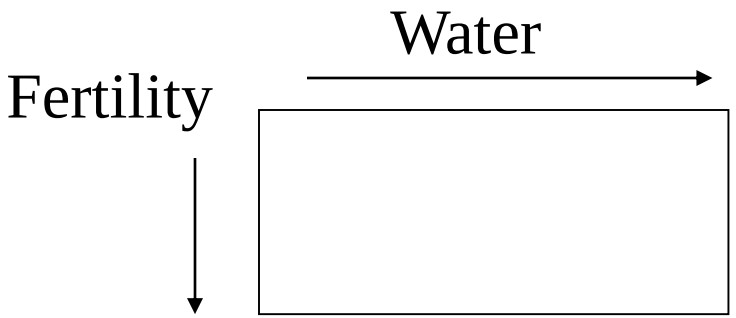
1. If it is randomized in blocks, it should be analyzed in blocks. Model selection usually not of interest.
2. The word "replication" sometimes is used to describe blocks and sometimes used to describe completely randomized reps. You have to determine whether the reps are blocks, or not, from the context.
3. An RCB analysis with two trts is the same as a paired t-test.

Advantages of RCB designs:

1. Accounts for variability, thereby increasing accuracy of treatment comparisons.
2. Simple to implement and easy to analyze.
3. In principle, can accommodate any number of treatments.

Disadvantages and limitations of RCB design:

1. Hard to find homogeneous blocks, when there are many trts.
    a) variety trials
    b) animal siblings used as blocks
2. Often more than one source of variability cannot be accommodated:



3. To be efficient, treatment differences should be nearly the same in each block. Say variety A is drought tolerant, B is not. Then low-water blocks might have larger treatment differences than high-water blocks.
4. More complication caused by missing data than with CRD design.

Flax RCB Example: (Steele and Torrie) Oil content in Redwing flaxseed inoculated at five different growth stages with the organism that causes " pasmo" (plus control).

Six treatments (t=6) randomly assigned in four blocks (b=4).

y_{ij} = the response for the ith trt in the jth block.

$= \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad i = 1,...,t \quad j = 1,...,b$

α_i = the ith treatment effect $\quad \beta_j$ = the jth block effect

ε_{ij} = the random error : independent, normal

$\mathrm{E}(\varepsilon_{ij}) = 0, \quad \mathrm{Var}(\varepsilon_{ij}) = \sigma^2$

Parameter estimation by LS. Find $\hat{\mu}$, $\hat{\alpha}_i$, and $\hat{\beta}_j$ that minimize:

$$\text{SSResid} = \text{SSE} = \sum_{i=1}^{t}\sum_{j=1}^{b}\left(y_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \right)^2$$

Notation for sample means and sums:

We use the "." (dot) and "-" (bar) notation:

"." in place of a subscript means sum over that subscript.

"-" over a letter means to divide by the number of terms summed.

$\bar{y}_{i.}$ = the mean for the ith treatment

$\bar{y}_{.j}$ = the mean for the jth block

$\bar{y}_{..}$ = the overall mean

(The same formulas, without the "-" would be totals.)

Hypothesis tests are based on the ANOVA table, which is constructed from the following:

SSTotal = total sum of squares $= \sum_{i,j} (y_{ij} - \bar{y}_{..})^2$

SSTrt = sum of squares for treatments $= b\sum_{i=1}^{t} (\bar{y}_{i.} - \bar{y}_{..})^2$

SSBlock = sum of squares for blocks $= t\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2$

$$\text{SSResid} = \sum_{i=1}^{n}\sum_{j=1}^{b}\left(y_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \right)^2$$

$$= \text{SSTotal - SSTrt - SSBlock}$$

Notes:

1. SSTotal and SSTrt formulas same as one-way model.
2. SSTrt, SSBlock use same formula with i,j reversed.

ANOVA Table for RCB Analysis:

Source	df	MS = SS/df	F test statistic	p-value
Blocks	b-1	MSBlock	MSB/MSResid	p_b
Treatments	t-1	MSTrt	MST/MSResid	p_t
Resid	(b-1)(t-1)	MSResid		
Total	tb-1			

Note dfResid formula works when data is balanced (no missing values).
Another way to calculate is by subtraction:
dfResid = Total # observations – 1 – dfBlocks - dfTrts

Hypothesis tests for RCB Analysis:

1. Test of Trt effect (primary question)

$H0: \alpha_1 = \alpha_2 = ... = \alpha_t$ vs $HA:$ at least one α_i differs from the others

$$F = \frac{MSTrt}{MS\,\mathrm{Re}\,sid} \qquad \mathrm{df}_1 = \mathrm{dfTrt} = t - 1 \quad \mathrm{df}_2 = df\mathrm{Resid}$$

Reject if $F > F_{a,df1,df2}$ or compute p - value.

2. Test of Block effect (secondary question)

$H0: \beta_1 = \beta_2 = ... = \beta_b$ vs $HA:$ at least one β_j differs from the others

$$F = \frac{MSBlock}{MS\,\mathrm{Re}\,sid} \qquad \mathrm{df}_1 = \mathrm{dfBlock} = b - 1 \quad \mathrm{df}_2 = \mathrm{dfResid}$$

Reject if $F > F_{a,df1,df2}$ or compute p - value.

RCB Analysis using R:

```
Model1 <- lm(Y ~ block + trt, data = )
Anova(Model1, type = 3)
emmeans(Model1, pairwise ~ trt)
```

Notes:

1. The summary() output does not directly answer common research questions. Use `Anova()` from car and `emmeans()` from emmeans package instead.
2. When **balanced (no missing data):**
   - Type 1 tests from `anova()` are the same as Type 3 from `Anova()` for this analysis.
   - emmeans and simple means are the same.
3. If the data is **not balanced**, Type 3 tests from Anova() are recommended. emmeans should still be used.

Notes (continued)

4. Use `emmeans( , pairwise ~ trt)` to get Tukey adjusted pairwise comparisons.
5. In this analysis, we are treating blocks as fixed, later we will treat them as random. However, if we are interested in treatment differences, the results will be the same. More on this in Random2 notes.
6. In this group of notes, we assume that we have exactly one observation per block*treatment combination. If you have multiple replicates per block*treatment combination, an interaction term can be included in the model.

"Simple" vs Model based SE

General Form of SE = $\hat{\sigma}/\sqrt{n}$

"Simple" SE for group i = $s_i/\sqrt{n_i}$

Model based SE for group i = $\sqrt{MSResid}/\sqrt{n_i}$

The difference is that the simple SE allows standard deviation to be estimated separately for each group.
While the model based SE uses a pooled estimate (√MSResid). This makes sense because the model assumes equal variance.

Note that if sample sizes are equal, the model based SE will be the same for all groups. This is true for the Flax example.

Flax Example: simple summary statistics vs emmeans()

```
> SumStats
stage          n  mean     sd     SE
1seeding       4  35.1 0.990 0.495
2earlybl       4  34.3 2.23  1.120
3fullblm       4  34.0 0.638 0.319
4full100       4  36.7 0.258 0.129
5ripenin       4  36.0 0.915 0.457
6uninoc        4  37.0 0.585 0.293

> emmeans(Model1, pairwise ~ stage)
$emmeans
 stage    emmean         SE df lower.CL upper.CL
 1seeding 35.100 0.5732401 15 33.87817 36.32183
 2earlybl 34.300 0.5732401 15 33.07817 35.52183
 3fullblm 34.000 0.5732401 15 32.77817 35.22183
 4full100 36.700 0.5732401 15 35.47817 37.92183
 5ripenin 36.050 0.5732401 15 34.82817 37.27183
 6uninoc  37.025 0.5732401 15 35.80317 38.24683
```

Parameter Interpretation:

As we have seen, we can usually do an RCB analysis without working with the parameters directly. (In other words, we don't need to directly use the information from `summary()`).

But the same principles that applied to the ANCOVA can be applied here to understand:

1. An RCB analysis is just a regression on indicator variables. The model is over parameterized, some indicator variables are omitted.
2. The parameter estimates (Coef table) can be interpreted in terms of the mean responses.
3. In the summary output, "Intercept" is mean for 1st trt in 1st block (like a reference). Other coefficients are additive effects for other blocks/trts.

Review of Power:

- Power was covered in depth in STAT511.
- Power is the probability of rejecting H0 when HA is true.
- Loosely speaking, power is the probability of detecting a differences between means given there is a difference.
- Power calculations are done in advance of running a study. The power calculation depends on the planned analysis.
- Power calculations require conjectures about treatment differences and variability. These values are typically drawn from pilot data or previously published studies.
- Investigators typically aim for 80% or 90% power.
- Power increases as the sample size increases.
- Power increases as the difference between means increases.
- Power decreases as the variability increases.

Sample Size and Power in the RCB F-test:

As with all of the other sample size calculations, we need:

1. A conjecture of random variability, which in this case is the within-block variance (σ^2) Often this is taken from the MSResid of some ANOVA table in the literature.
2. Conjecture about the true alternative that we want to detect. Specific values for ($\mu_1, \mu_2, ..., \mu_t$).

Power is a function of the "noncentrality" parameter for the F-distribution, given the specific alternative:

$$\lambda = \frac{b\sum_{i=1}^{t}\left(\mu_i - \bar{\mu}_{\cdot}\right)^2}{\sigma^2}$$

The noncentrality parameter describes how far to the right the distribution of F is shifted when the alternative is true. Note that power increases as the number of blocks or the difference between the true means increases. Power increases as the variance decreases (i.e. if the EU's within a block are more homogeneous.)

Power in the RCB Example:

We conjecture that $\sigma=3$ and the treatment means are 4,5,6,7 ($t=4$). Then variance of treatment means is 1.67 (and standard deviation is 1.29).

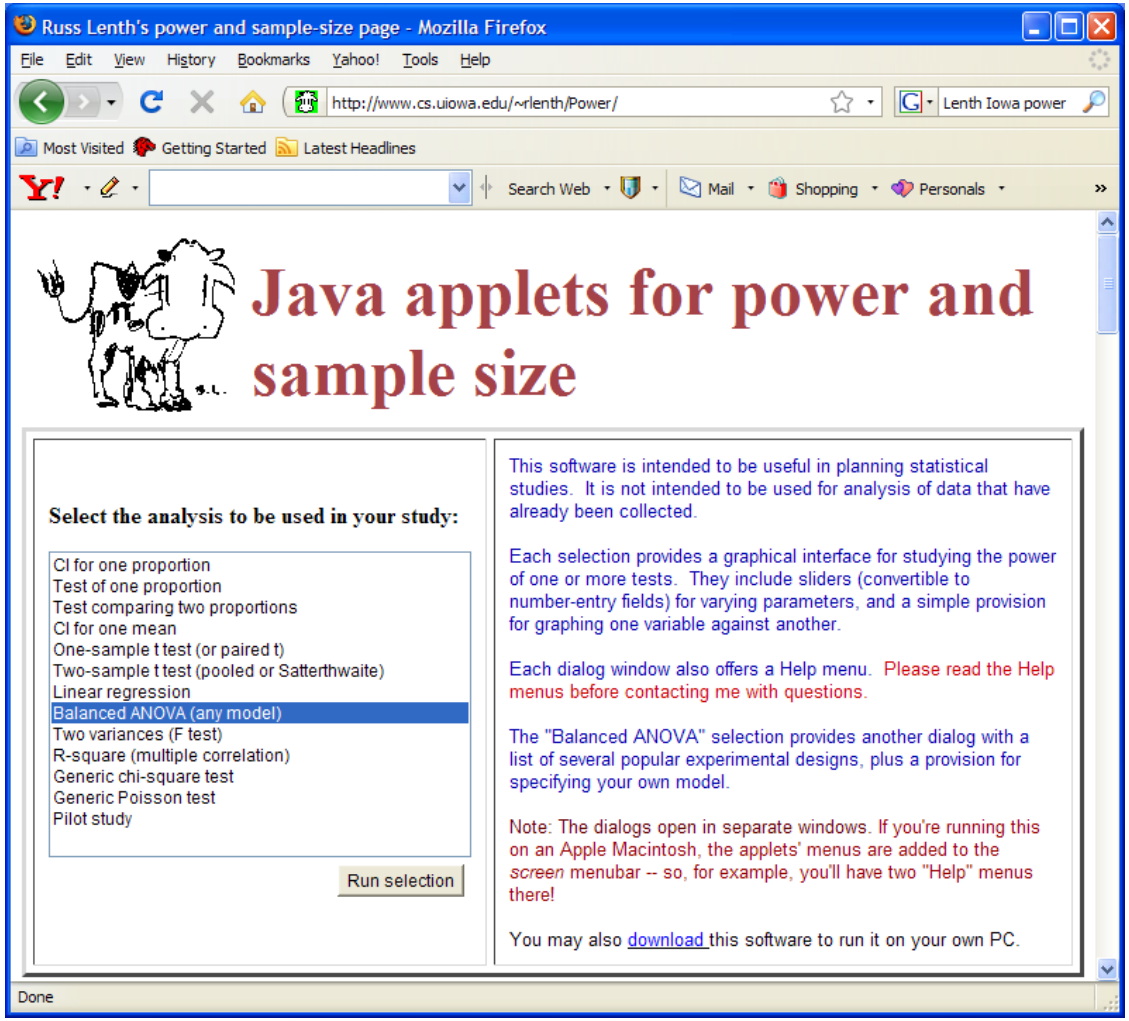
The investigators are interested in calculating power for $b=10$ blocks.

Power can be calculated using Lenth's online power calculator.

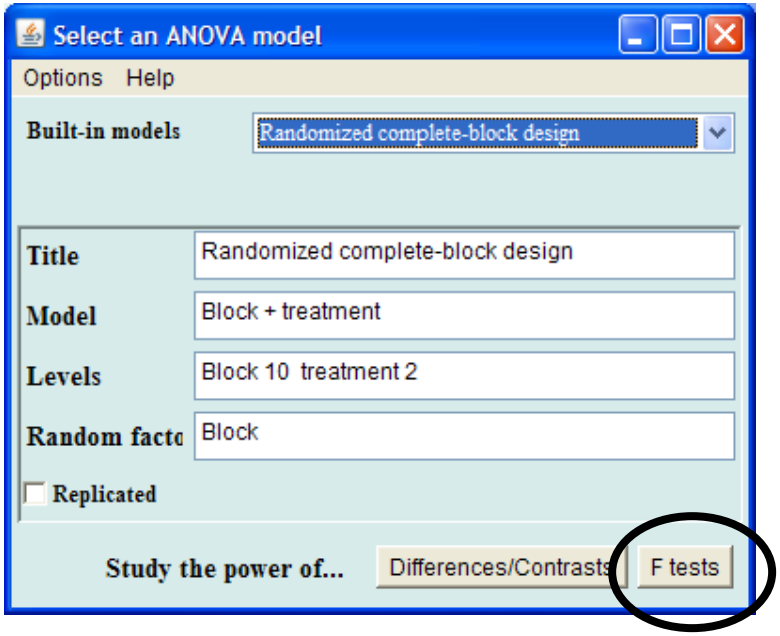
Based on the conjectures above, using $b=10$ blocks the power is 0.426.

Power Calculation for RCB using Lenth's Online Power Tool

http://www.cs.uiowa.edu/~rlenth/Power/



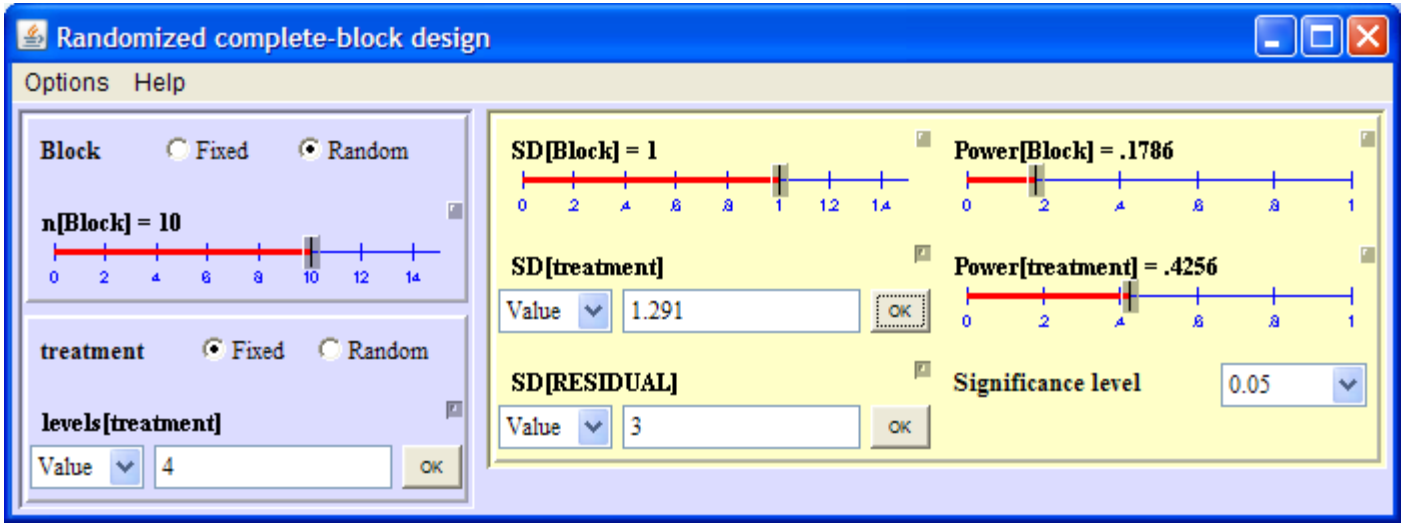
Select "Balanced ANOVA", then "Run Selection"

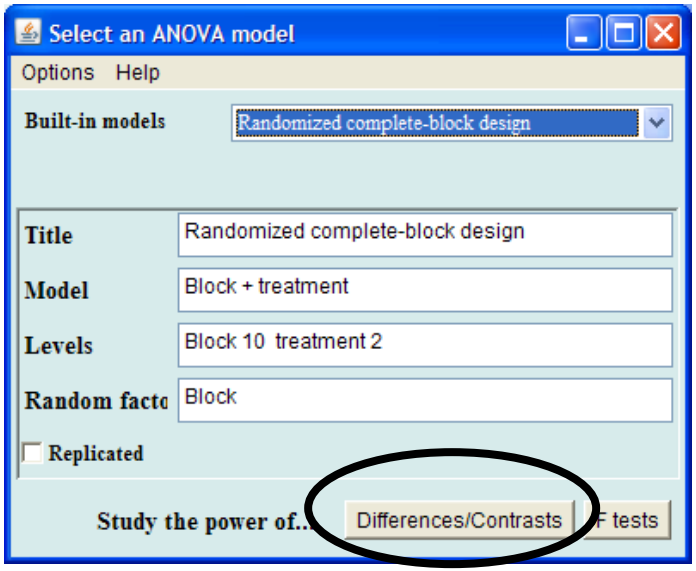


Select RCB from screen on the left, then "F tests"

Put inputs into screen below using sliders or little grey boxes to type numerical input. Std. dev. of treatment means is 1.291.

Power for treatments is 0.4256

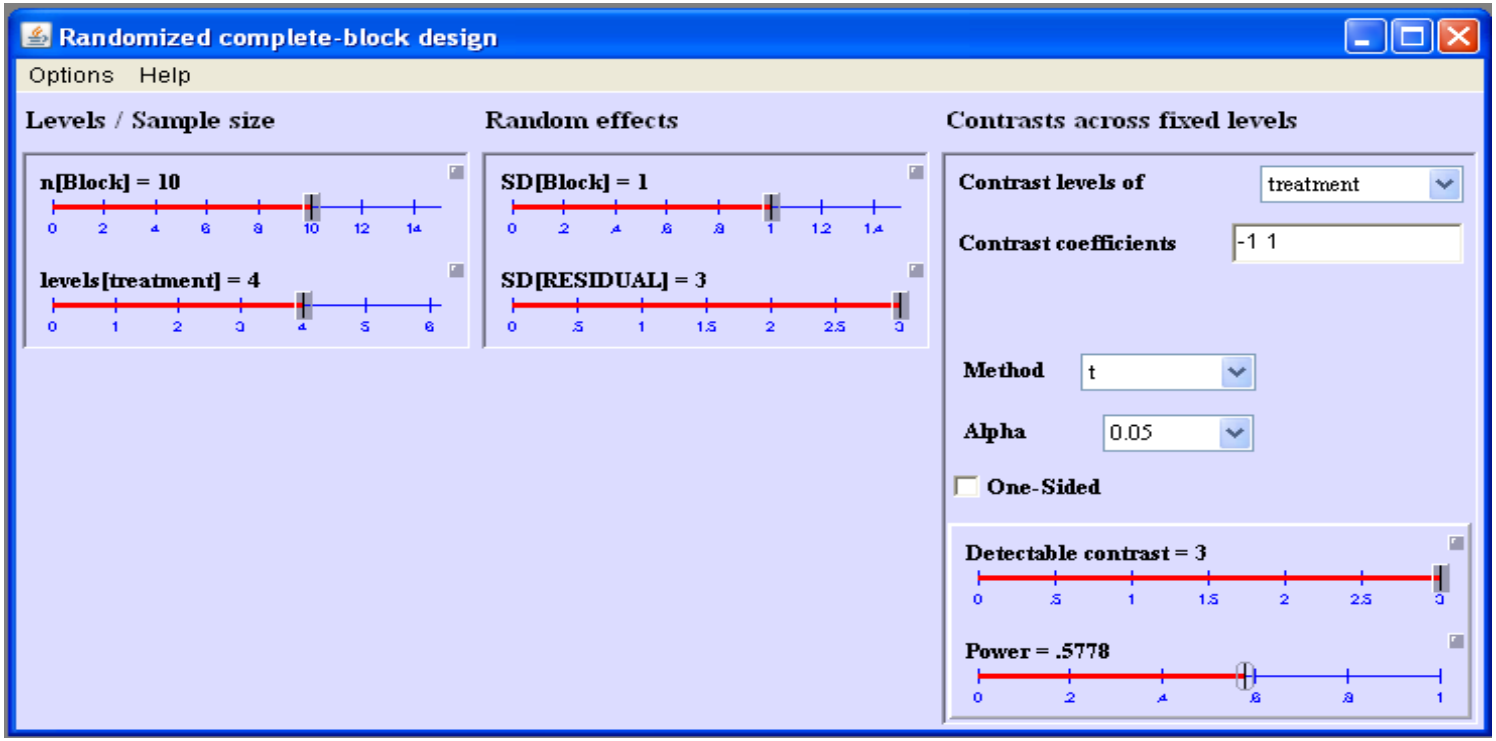




Now try using "Differences/Contrasts"

Suppose for some comparison the difference between trt means is 4-1=3

Power for this contrast is 0.5778.



3. Unbalanced RCB (Missing Data)

Media Example: An experiment to compare the growth of snapdragons in various growing media. Seven growing media ($t = 7$, three blocks ($b = 3$), but with 2 missing observations. So there are a total of n=19 observations. Response variable is stem length. The model is the same as before.

Model: $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$ for non-missing i, j

Analysis is also the same as the balanced case**, just remember to use Anova(, type =3) for tests and emmeans() for estimates and pairwise comparisons.**

Predicted values: $\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$

	j = 1	2	3
i = 1	$\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_3$
2	$\hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_3$
3	$\hat{\mu} + \hat{\alpha}_3 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_3 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_3 + \hat{\beta}_3$
4	$\hat{\mu} + \hat{\alpha}_4 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_4 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_4 + \hat{\beta}_3$
5	$\hat{\mu} + \hat{\alpha}_5 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_5 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_5 + \hat{\beta}_3$
6	$\hat{\mu} + \hat{\alpha}_6 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_6 + \hat{\beta}_2$	$\hat{\mu} + \hat{\alpha}_6 + \hat{\beta}_3$
7	$\hat{\mu} + \hat{\alpha}_7 + \hat{\beta}_1$	$\hat{\mu} + \hat{\alpha}_7 + \hat{\beta}_2$	

emmeans for treatments are computed as row averages.

emmeans for blocks are computed as column averages.

Example (for Illustration): We compute predicted values for missing observations. Use the parameter estimate info (Coef table) to verify the predicted value.

Clarion (trt #1) in Block 1 is missing.

$$\hat{y}_{11} = \hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_1$$
$$= 34.443 + 0 + 0$$
$$= 34.443$$

Clinton (trt #2) in Block 2 is missing.

$$\hat{y}_{22} = \hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_2$$
$$= 34.443 - 2.055 - 1.509$$
$$= 30.879$$

Notes about unbalanced RCB analyses:

1. The emmean is called the "adjusted" mean (adjusted to as if each treatment had occurred in each block). The adjustment will be up if the missing value is in block with high values, and down if the missing value is in a block with low values.
2. The mean of the observed values for each treatment is called the "raw", "simple" or "unadjusted" mean.
3. Generally the emmeans and simple means will not be the same, except for treatments that have no missing values.
4. Differences between emmeans will be equal to the differences between estimated α's.
5. Differences between emmeans will have different SEs, depending on missing values.

Notes (continued)

6. The SEs will be larger for means of treatments having missing values.

7. In cases of missing values, the Type 1 tests from anova() and Type 3 tests from Anova() will NOT be the same. Use the Type 3 tests!

8. These methods are highly dependent on the assumption that the effect of each treatment will be the same in each block.

9. These methods assume that observations are missing due to reasons unrelated to the treatment (such as lost data, or botched lab analysis). If the observation is missing due to the effect of the treatment, this analysis will be misleading.

10. These methods work for small numbers of missing observations.

Type I vs Type 3 SS in RCB:

Model0: lm(Y ~ 1)
Model1A: lm(Y ~ trt)
Model1B: lm(Y ~ block)
Model2: lm(Y ~ trt + block)

Type 1 SS (Sequential): Add terms in order given in the model statement. Results depend on order of terms in the model!
SSTrt (ignoring block) = SSModel1A – SSModel0
SSBlock (adjusting for trt) = SSModel2 – SSModel1A

Type 3 SS (Marginal or Unique): Tests of each term as if added last. Results do NOT depend on order of terms in the model.
SSTrt (adjusting for block) = SSModel2 – SSModel1B
SSBlock (adjusting for trt) = SSModel2 – SSModel1A

4. The Latin Square (LS) Design

The Latin square design attempts to control for <u>two</u> sources of variability at the same time.

Example: 4x4 Latin Square

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

More fertility

More water

Latin Square Defining properties:

1. # Trts = # Rows = # Cols
2. Each treatment (letter) occurs exactly once in each row and once in each column.

Latin Square model: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$

$\alpha_i = \text{the treatment effect}, \quad \beta_j = \text{the row effect}, \quad \gamma_k = \text{the column effect}$

$\varepsilon_{ijk} = \text{the random error: independent, normal}, \mathrm{E}(\varepsilon_{ijk}) = 0, \quad \mathrm{Var}(\varepsilon_{ijk}) = \sigma^2$

Least squares estimates:

Parameter estimates minimize: $\text{SSResid} = \sum_{\text{i,j,k observed}} \left(y_{ijk} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k) \right)^2$

Latin Square ANOVA Table:

Source	df	SS
Trt	$t - 1$	*SStrt*
Rows	$t - 1$	*SSrows*
Cols	$t - 1$	*SScols*
Resid	$(t-1)(t-2)$	*SStot-SStrt-SSrows-SScols*
Total	$t^2 - 1$	*SStot*

Latin Square Example:

A 4 by 4 LS is used to compare the mean yields from 4 varieties of wheat. Results:

F_{trt}=58.03 with df_{trt}=t-1=3 df_{Resid}=6 p<0.0001

Was the LS an efficient design for this problem? (i.e. Did the decrease in error variance achieved by the two-directional blocking gain more than was lost by reducing the dfResid by 2(t-1)?)
A formula from O&L calculates "relative efficiency" of the LS design (relative to the CRD) on the same experimental units. In practice, I usually just look at the F-tests for rows and columns. If significant, or nearly significant, then the LS was probably better than the CRD.

In the example: controlling for column variation was definitely worthwhile (F_{col}=5.00, p-value=0.0452) ; controlling for row variation probably was not (F_{row}=1.44,p-value=0.3219).

Notes about Latin Square Designs:

1. It is like blocking in two directions at once. If you ignore columns, it looks like a RCB using rows as blocks. If you ignore rows, it looks like a RCB using columns as blocks.
2. Only possible when: #rows = #columns = #treatments = t
3. Rows and columns do not have to be spatial:

Ex1:

Cars	1	2	3	4	Drivers
1					
2					
3					
4					

Letters=gas mix

Ex2:

Cow	1	2	3	4
				Period
1				
2				
3				
4				

Letters=feed trt

4. It is an “incomplete” design:
   (t rows) x (t cols) x (t trts)= t^3 possible combinations
   We observe only t^2 of these (one trt for each row/col combination)

5. It is a “balanced” design:
   a) each row occurs once with each col
   b) each trt occurs once in each row
   c) each trt occurs once in each col

6. Balance implies:
   Type 1 SS = Type 3 SS
   simple means = emmeans

7. Some care is needed when randomizing LS designs. See: *Experimental Design* by Cochran and Cox.

8. LS designs are very common with t = 4 to 7. Larger LS designs (t > 7) are seldom used because the additive row-column variability structure is rare over large areas or time spans.

9. Missing data is handled in the same fashion as missing data in the RCB. A small proportion of missing value is tolerable. The methods are the same as for the RCB. Just fit the model to the non-missing data and use Type 3 tests.

Graeco-Latin Squares (blocking in three directions)**:**

Assembly times (Y) compared for 4 methods:
Operators = Col (1, 2, 3, 4)
Order of assembly = Row (1, 2, 3, 4)
Location = Greek letter (α, β, δ, γ)
Method = Trt (A, B, C, D)

	1	2	3	4
1	Cβ	Bγ	Dδ	Aα
2	Bα	Cδ	Aγ	Dβ
3	Aδ	Dα	Bβ	Cγ
4	Dγ	Aβ	Cα	Bδ

Source	df
Operator	3
Order	3
Location	3
Method	3
Resid	3
Total	15

In this example the blocking variables (row, column and Greek letter) do not explain much variation, but reduce the dfResid to 3!!

Treatment effects are significant if row, column and Greek letter are eliminated from the analysis, but this might be a controversial choice. (You see why this design is not used very often.)

"Trade-off" between MSResid and dfResid

Recall that the LSD just gives the (unadjusted) Margin of Error (ME) for a pairwise comparison (ex: $\mu_1 - \mu_2$).

$$LSD_{0.05} = t_{\alpha/2, df\ \mathrm{Resid}} \sqrt{\frac{2MS\ \mathrm{Resid}}{n}}$$

Smaller LSD corresponds to higher power.

As we add blocking (or other) terms to the model, we:

1. Reduce dfResid (decreasing power).
2. Tend to reduce MSResid (increasing power).

There is no "cut-off" for dfResid, but I start to get concerned when dfResid ≤ 10.

Graeco-Latin Square Example:

Model1: dfResid = 3, $t_{\alpha/2}$= 3.18

MSResid = 27.5/3 = 9.17

LSD = 6.81

Model2: dfResid = 12, $t_{\alpha/2}$= 2.18

MSResid = 54.5/12 = 4.54

LSD = 3.28

Note: In this case MSResid is actually lower for the simpler model (Model2), but there is a tendency for MSResid to be lower for more complex models.

5. Thoughts on these designs

$$CRD: y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

$$RCB: y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

$$LS: y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

The key to deciding between these three designs is to model what you think the major sources of response variability would be <u>without treatments:</u>

$$CRD: y_{ij} = \mu + \varepsilon_{ij}$$

$$RCB: y_{ij} = \mu + \beta_j + \varepsilon_{ij}$$

$$LS: y_{ijk} = \mu + \beta_j + \gamma_k + \varepsilon_{ijk}$$

Suppose the following grids show the background response for a field (without applying trts). What design is appropriate for each scenario?

4	4	4	4
4	4	4	4
4	4	4	4
4	4	4	4

4	4	4	4
6	6	6	6
7	7	7	7
3	3	3	3

4	5	3	2
5	6	4	3
6	7	5	4
4	5	3	2

5	5	4	4
5	5	4	4
3	3	6	6
3	3	6	6

4	5	3	4
4	5	3	4
4	5	3	4
4	5	3	4

5	6	7	8
6	5	6	7
7	6	5	6
8	7	6	5