Model Based Statistics in Biology.
Part IV. The General Linear Model. Multiple Explanatory Variables.
Chapter 13.4 Fixed*Random Effects (Randomized Block)

ReCap. Part I (Chapters 1,2,3,4), Part II (Ch 5, 6, 7)
ReCap Part III (Ch 9, 10, 11)
ReCap Multiple Regression (Ch 12)
13.1 Fixed Effects ANOVA (no interactive effects)
13.2 Fixed Effects ANOVA (interactive effects)
13.3 Fixed*Random Effects (Paired t-test)
13.4 Fixed*Random Effects (Randomized Block)
13.5 Fixed*Random Effects (Repeated Measures)
13.6 Nested Random Effects (Hierarchical ANOVA)
13.7 Random within Fixed (Hierarchical ANOVA)
13.8 More Than Two Factors (to be written)

on chalk board

Tribolium growth data
Sokal and Rohlf Box 11.4
Ch13.xls

ReCap Part I (Chapters 1,2,3,4) Quantitative reasoning is based on models, including statistical analysis based on models.
ReCap Part II (Chapters 5,6,7)
Hypothesis testing uses the logic of the null hypothesis to declare a decision.
Estimation is concerned with the specific value of an unknown population parameter.
ReCap (Ch 9, 10,11) The General Linear Model with a single explanatory variable.
ReCap (Ch 12) GLM with more than one regression variable (multiple regression)
ReCap (Ch 13) GLM with more than one categorical variable (ANOVA).
Two fixed factors (Ch 13.1, Ch13.2)
One fixed and one random factor (Paired t-test)

Today: Special case of Two way ANOVA: Randomized Blocks.
One factor fixed by design, the other factor is random.

Wrap-up.
The randomized block design is analyzed with a general linear model consisting of two explanatory variables on a nominal scale.
One of these is fixed (two or more classes),
the other is random (two or more classes).
We are interested in the fixed effects, controlled for the random effects.
A special case of randomized blocks – the paired t-test.

Introduction. The randomized block design has a fixed and a random factor. The randomized block is an example of statistical control, in which the effects of one variable (the random factor) are removed in order to arrive at a better test for the fixed factor. This results in a more sensitive test because it removes some of the noise in the data before testing. The paired t-test is a special case of the randomized block, in which the fixed factor has just two categories.

Statistical control is used when manipulative control is not possible.
- Epidemiology. (manipulative control is unethical)
- Many field situations.
  - Manipulative control is impossible at large scales.
  - Manipulative control can be expensive, even at small scales.
  - Manipulative control can generate artefacts, and so a study with well designed statistical controls can be more informative.

Blocks	Gtype	Wt
1	1	0.958
2	1	0.971
3	1	0.927
4	1	0.971
1	2	0.986
2	2	1.051
3	2	0.891
4	2	1.010
1	3	0.925
2	3	0.952
3	3	0.829
4	3	0.955

1. Construct model

Data

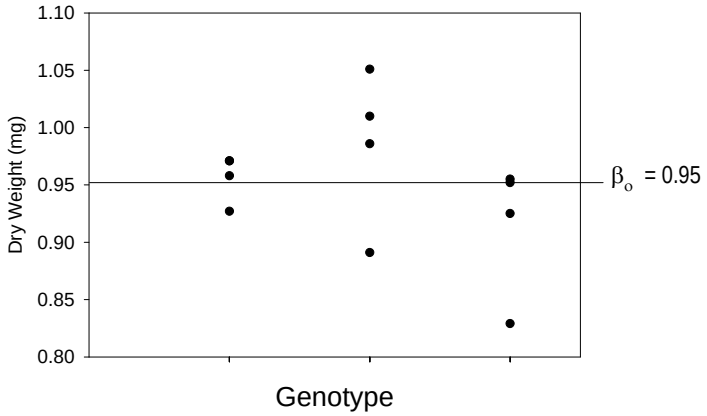
To illustrate this analysis, we use data from Box 11.4 in Sokal and Rohlf 1995, p 350. Dry weights (mg) of 3 genotypes of *Tribolium castaneum* in 4 experiments.

Verbal model.

Does weight of flour beetle *Tribolium* vary among genotype, after controlling for differences among experiments ?

Each experiment is a block.
The analysis will take out the effects of blocks. This will be a more sensitive test of whether weight varies among genotypes.

Graphical model.



Response variable is beetle mass
The first explanatory variables is genotype, a fixed factor under the control of the experimenter.
The second explanatory variable is block, a random factor beyond the control of the investigator.

1. Construct model

Formal model.

Symbol	Name	Fixed or random?	Units	Type of measurement scale
M	mass of beetles		mg	ratio
G	genotype (I II III)	Fixed		categorical (nominal)
B	block	Random		categorical (nominal)

Write GLM: $M = \beta_o + \beta_B \cdot B + \beta_G \cdot G + \beta_{GxB} \cdot G \cdot B$ + residual

S&R81 $Y_{ijk} = \mu + B_j + \alpha_i + (\alpha \cdot B)_{ij} + \varepsilon_{ij}$

Common practice is to use roman letters to distinguish random from fixed effects.
Genotype is treatment effect, so Greek letter is used.
Block is random effect so Roman letter is used.

In randomized block design, the interaction is assumed to be absent. We assume that response pattern will be similar across the blocks via control of experimental conditions.

```
Write full model on board,
cross out the interaction term.
```

full GLM: $M = \beta_o + \beta_G \cdot G + \beta_B \cdot B + \beta_{GxB} \cdot G \cdot B$ + residual

revised GLM $M = \beta_o + \beta_G \cdot G + \beta_B \cdot B$ + *residual*

2. Execute analysis.

Place data in model format:

Column labelled *M*, with response variable mass
Column labelled X_B with explanatory variable, X_B = 1,2,3, or 4
These are labels (categories), not numbers on ratio scale.
Column labelled X_G with explanatory variable, X_G = bb, +b, ++

Code model statement in statistical package according to the GLM

$$M = \beta_o + \beta_B \cdot X_B + \beta_G \cdot X_G + \varepsilon$$

```
MTB> ANOVA 'M' = 'XB'  'XG'
MTB> GLM   'M' = 'XB'  'XG'
SUBC> fits c4;
SUBC> res c5.
```

2. Execute analysis.

The grand mean.

$\hat{\beta}_0 = 12^{-1}\,\Sigma M = 12^{-1} \cdot 11.426 = 0.95217$ mg

Wt	grand mean	Gtype Effect	Block Effect	Fits	Res
0.958	0.952	0.00458	0.004	0.9609	-0.0029
0.971	0.952	0.00458	0.039	0.9959	-0.0249
0.927	0.952	0.00458	-0.070	0.8869	0.0401
0.971	0.952	0.00458	0.027	0.9833	-0.0123
0.986	0.952	0.03233	0.004	0.9887	-0.0027
1.051	0.952	0.03233	0.039	1.0237	0.0273
0.891	0.952	0.03233	-0.070	0.9147	-0.0237
1.010	0.952	0.03233	0.027	1.0110	-0.0010
0.925	0.952	-0.03692	0.004	0.9194	0.0056
0.952	0.952	-0.03692	0.039	0.9544	-0.0024
0.829	0.952	-0.03692	-0.070	0.8454	-0.0164
0.955	0.952	-0.03692	0.027	0.9418	0.0133

The fitted values are computed from the genotype and block means.

mean($M_{G=++}$) = 0.95675 mg

$\hat{\beta}_{G=++}$ = (0.95675 – 0.95217)
= 0.00458 mg

etc.

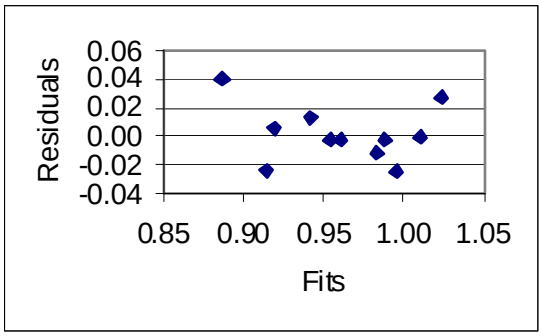
The table shows calculations in spreadsheet format.

3. Evaluate the model.

Plot residuals versus fits.

a. No line fitted in model,
so skip this evaluation.

b. No systematic change in residuals with increase in fitted values (*i.e.* no cones) so residual homogeneous, no need to revise error structure of model.



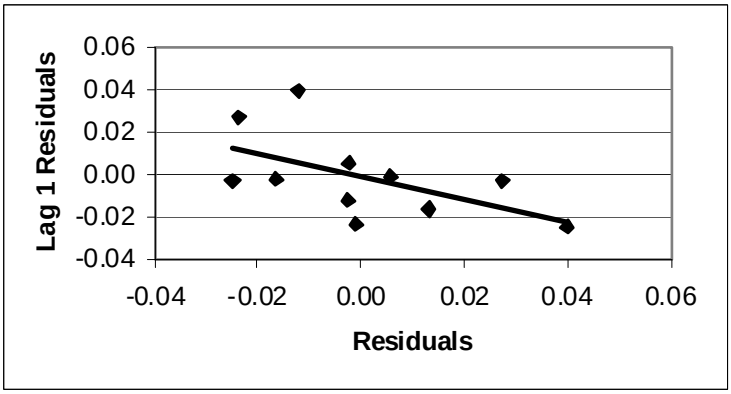
c. If n small, evaluate assumptions for computing p-values.

n = 12 so violations will distort p-values or confidence limits.

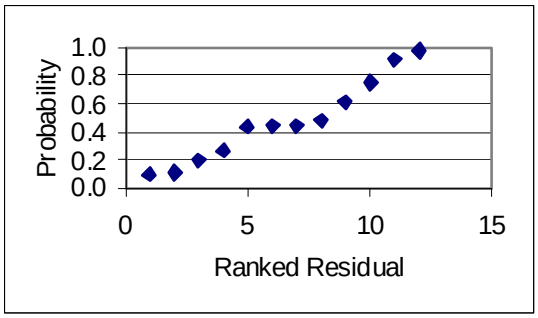
<u>Homogeneous</u>? Yes

<u>Sum(res) = 0</u>? Yes

<u>Independent</u>? The graph suggests a downward trend for residuals listed in order of blocks within genotype. There is a suspicious negative association going from block 1 to 2, 2 to 3, 3 to 4. There appears to be a negative carryover effect, assuming the experiments are presented in the order in which they were conducted.



3. Evaluate the model.



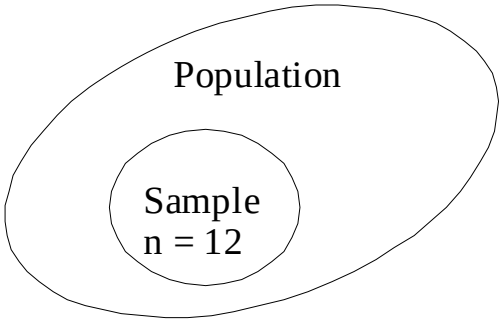
Normal ?

The residuals deviate somewhat from normal, as judged relative to trend line in the normal probability plot. There is some indication of clustering of values around the median value of the residuals, as the plot tends horizontally from the 5th to 8th smallest residual.

If our p-values are close to the criterion of 5%, we will need to consider recomputing them by randomization.

4. State population and whether sample is representative.

In this example the investigator (R.R. Sokal) chose to treat genotype as a fixed effect, so inference is to these three genotypes only.



Block is a random effect. The time period in which each experiment was run, and the other conditions that vary from experiment to experiment were considered immaterial to the outcome.

The population is defined by the experimental protocol: all possible measurements that could have been made on *Tribolium* using the procedural statement for this experiment. This contrasts with an enumerable populations such as a biological population (all organisms of a species alive at any given moment).

5. Decide on mode of inference. Is hypothesis testing appropriate?

Hypothesis testing appropriate because research question is binary (yes/no): does weight depend on genotype ? It is not clear whether the observed differences among genotypes are due to chance variation, so hypothesis testing is appropriate.

6. State H_A H_o pairs, test statistic, distribution, tolerance for Type I error.

Interaction term. There is no H_A for the interaction term, because the interaction was removed by the experimental design: genotypes weighed in random order (S&R95 p351), to remove interaction effect, treatment*block.

In effect, we are going to assume that interaction variance Var(GxB) = 0

> Note: no df left if Var(GxB) estimated
> $df_{total} = 11$ $df_G = 2$ $df_B = 3$ $df_{GxB} = 2*3 = 6$
> $df_{res} = 11 - 2 - 3 - 6 = 0$

Block term (experiment)

There is no H_A about this term. We are not interested in this effect. We merely want to estimate the variance component due to blocks and remove this variance from the error term, so as to have a more sensitive test, and a better chance of detecting main effects.

Experimental term

This is a fixed effect, the means are of interest.

$E(M_{bb})$ is the expected value (true mean) of the weight of genotype bb
$E(M_{+b})$ is the expected value (true mean) of the weight of genotype +b
$E(M_{++})$ is the expected value (true mean) of the weight of genotype ++

H_o: $E(M_I) = E(M_{II}) = E(M_{III})$
H_A: the means differ

The hypothesis pair above is equivalent to the following pair concerning variance.

H_A: $Var(\beta_{Genotype}) > 0$ There is variance due to experimental factor.
H_o: $Var(\beta_{Genotype}) = 0$

Are there more specific hypotheses about parameters? No

State test statistic	F-ratio
Distribution of test statistic	F-distribution
Tolerance for Type I error	5% (conventional level)

7. ANOVA - Calculate then partition df and SS according to model.

```
Model at top of board on left.
ANOVA table at top of board on right.
```

GLM	$M-\beta_o$	=	$\beta_B \cdot X_B$	+ $\beta_G \cdot X_G$	+	ε
Source	Total	=	Block	+ Genotype	+	Resid
df	12–1	=	4 –1	+ 3–1	+	12–3–2
SS	0.0353	=	0.021391	+ 0.009717	+	0.004184

$SS_{tot} = Var(M) \cdot df_{total} = 11 * Var(M) = 11 * 0.0032084 = 0.0353$

$SS_{tot} = \Sigma Y^2 - n^{-1}(\Sigma Y)^2 = 10.914748 - 12^{-1} \cdot 11.426^2 = 0.0353$

SS_{total} computed by Minitab.

```
MTB> let k1 = ssq('weight')
MTB> print k1
```

GLM commands in other packages perform in similar ways, to partition the variance.

GLM commands compute MS and variance ratio F. MS block not computed, there is no interest in testing whether this term is significant. The interest is in estimating it.

```
    Source      df        SS       MS     F---->      p
  blocks         3    0.021391
  gtype          2    0.009717   0.004858       6.97    0.027
  residual       6    0.004184   0.000697
 total          11    0.0353
```

Calculate Type I error.

p = 0.027 calculated from F-distribution with df = 2, 6

Statistical control

Compare this partitioning to that when Block term not included in model

```
MTB > anova 'weights' = 'gtype'
Factor     Type Levels Values
gtype     fixed      3     1     2     3
Analysis of Variance for weights

Source      DF         SS         MS       F      P
gtype        2   0.009717   0.004859    1.71  0.235
Error        9   0.025575   0.002842
Total       11   0.035292   0.003208
```

Compare SS	SS_{error} shrinks from 0.0256 to 0.0042
Compare MS	MS_{error} shrinks from 0.02842 to 0.000697
Compare F-ratio	F-ratio increases from 1.71 to 6.97
Compare p-values	p-value shrinks from 0.235 to 0.027

Statistical control

Because the block effects are estimated and removed, the residual SS is much smaller. This allows smaller genotypic differences to be detected.

Sokal and Rohlf 1995 (p 350), provide a calculation of the increased efficiency of the randomized block design. Statistical contraol to produce a better analysis is one of the key concepts of experimental design.

8. Decide whether to recompute p-value.

Residuals were homogeneous, were not independent, and deviated slightly from normal distribution.

n is small, but p = 0.027 and hence would need to change be a factor 2-fold to change our decision.

Given this information, we would not usually undertake randomization, even though the residuals were not independent.

How good was this judgement?

The p-value via randomization in this case is 128/5000 = 0.0256

The p-value changed by a factor of 0.027 / 0.0256 = 1.05.

Our judgement (no need for randomization) was correct.

Having computed the p-value based on randomization, we report it because it is free of assumptions.

9. Declare decision about terms.

Only one term, the fixed factor, is tested.

Reject H_o and accept H_A: Var(G) > 0

There is significant variation in mean dry weight among genotypes.

$F_{2,6}$ = 6.97 p = 0.026

10. Report and interpret parameters of biological interest.

No parameters are reported for the block term because it is a random factor and so the means are of no interest.

mean(M_{bb}) = 0.957mg	st.err = 0.0104 mg
mean(M_{+b}) = 0.9845 mg	st.err = 0.0339 mg
mean(M_{++}) = 0.915 mg	st. err = 0.0295 mg

The differences in weight among genotypes are small and not detectable with these standard errors. The differences become detectable (as in the ANOVA table) when variation among experiments (blocks) is removed from error term.