1 Completely Randomized Designs (CRD)

Setup

- g: the number of treatment group
- N: the number of experimental units
- n_i: the number of observations in each treatment group
  - $i = 1, \cdots, g$
  - $n_1 + n_2 + \cdots + n_g = N$

Notation

Symbol	Meaning	Formula
$y_{i\cdot}$	Sum of all values in group i	$y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$
$\overline{y}_{i\cdot}$	Mean of group i	$\overline{y}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_{i\cdot}} y_{ij} = \frac{1}{n_i} y_{i\cdot}$
$y_{\cdot\cdot}$	Sum of all observations	$y_{\cdot\cdot} = \sum_{i=1}^{g} \sum_{j=1}^{n_i} y_{ij}$
$\overline{y}_{\cdot\cdot}$	Overall mean	$\overline{y}_{\cdot\cdot} = \frac{y_{\cdot\cdot}}{N}$

Parameter Estimation

- Residual/error sum of squares

$$SS_E = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i\cdot})^2$$

- Mean squared error

$$\hat{\sigma}^2 = MS_E = \frac{1}{N-g} SS_E = \frac{1}{N-g} \sum_{i=1}^{g} (n_i - 1) s_i^2$$

  - s_i^2: empirical variance in group i
  - This is an unbiased estimator for σ^2 ($N - g$ instead of N in the denominator)
  - The error estimate has $N - g$ degrees of freedom (N observations, g parameters)

$$N - g = \sum_{i=1}^{g} (n_i - 1)$$

Estimation Accuracy

Parameter	Estimator	Standard Error
μ	$\overline{y}_{\cdot\cdot}$	$\sigma \sqrt{\frac{1}{N}}$
μ_i	$\overline{y}_{i\cdot}$	$\sigma \sqrt{\frac{1}{n_i}}$
α_i	$\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot}$	$\sigma \sqrt{\frac{1}{n_i} - \frac{1}{N}}$
$\mu_i - \mu_j = \alpha_i - \alpha_j$	$\overline{y}_{i\cdot} - \overline{y}_{j\cdot}$	$\sigma \sqrt{\frac{1}{n_i} - \frac{1}{n_j}}$

- 95% confidence interval for α_i is given by

$$\hat{\alpha}_i \pm q t_{0.975; N-g} \cdot \hat{\sigma} \sqrt{\frac{1}{n_i} - \frac{1}{N}}$$

One-Way ANOVA Table

Source	df	Sum of squares (SS)	Mean squares (MS)	F-ratio
Treatment	$g - 1$	SS_{Trt}	$MS_{Trt} = \frac{SS_{Trt}}{g-1}$	$\frac{MS_{Trt}}{MS_E}$
Error	$N - g$	SE_E	$MS_E = \frac{SS_E}{N-g}$	

- $E[MS_{Trt}] = \sigma^2 + \sum_{i=1}^{g} n_i \alpha_i / (g-1)$
- F follows an F-distribution with $g-1$ and $N-g$ degrees of freedom: $F_{g-1,N-g}$

F-Distribution

- The F-distribution has two degrees of freedom parameters: One from the numerator and one from the denominator mean square (treatment and error)

$$F_{n,m} = \frac{\frac{1}{n}(X_1^2 + \cdots + X_n^2)}{\frac{1}{m}(Y_1^2 + \cdots + Y_m^2)} \quad \text{where } X_i, Y_j \text{ are i.i.d. } \mathcal{N}(0,1)$$

- It holds that $F_{1,n} = t_n^2$ (= the square of a t_n-distribution). F-test for the case $g = 2$ is nothing else than the squared t-test

2 Specific Differences and Multiple Testing

Sum of Squares a Contrast

- Associate sum of squares

$$SS_c = \frac{(\sum_{i=1}^{g} c_i \overline{y}_{i\cdot})^2}{\sum_{i=1}^{g} \frac{c_i^2}{n_i}}$$

  - SS_C has one degree of freedom, hence $MS_c = SS_c$
  - This is nothing else than the square of the t-statistic of our null hypothesis $H_0 : \sum_{i=1}^{g} c_i \cdot \mu_i = 0$ (without the MS_E factor)
  - $H_0 : \dfrac{MS_c}{MS_E} \sim F_{1,N-g}$

Orthogonal Contrasts

- Two contrasts c and c^* are called orthogonal, if $\sum_{i=1}^{g} c_i \cdot c_i^* / n_i = 0$
- Orthogonal contrasts contain independent information
- If there are g groups, one can find $g-1$ different orthogonal contrasts (1 dimension already used by global mean $(1, \cdots, 1)$)
- Decomposition of Sum of Squares

$$SS_{c^{(1)}} + SS_{c^{(2)}} + \cdots + SS_{c^{(g-1)}} = SS_{Trt}$$

Multiple Comparisons

- Type I error: falsely rejecting H_0
- Perform m tests $H_{0,j}$ where $j = 1, \cdots, m$
- If all $H_{0,j}$ are true and if all tests are independent, the probability of making at least one false rejection is given by

$$1 - (1-\alpha)^m$$

  - where α is the individual significance level

- The more tests we perform, the more likely we are getting some significant result

Different Error Rates

	H_0 true	H_0 false	Total
Significant	V	S	R
Not significant	U	T	$m-R$
Total	m_0	$m-m_0$	m

- Consider testing m hypotheses, whereof m_0 are true
  - V: Type I errors
  - T: Type II errors
  - R: Discoveries
- Comparison-wise error rate is type I error rate of an individual test
- Family-wise error rate is the probability of rejecting at least one of the true H_0's

$$\text{FWER} = P(V \geq 1)$$

- The false discovery rate (FDR) is the expected fraction of false discoveries

$$\text{FDR} = E\left[\frac{V}{R}\right]$$

Confidence Intervals

- Typically, each H_0 corresponds to a parameter
- We can construct confidence intervals for each of them
- We call these confidence intervals simultaneous at level $(1-\alpha)$ if the probability that all intervals cover the corresponding true parameter is $1-\alpha$

Scheffe

- A method which controls for the search over any possible contrast
- These p-values are honest
- Theory
  - $SS_c \leq (g-1)MS_{Trt}$ for any contrast c (because $SS_{Trt} = SS_c + \cdots$)
  - $\frac{SS_c}{MS_E} \leq (g-1)\frac{MS_{Trt}}{MS_E}$
  - $\max_c \frac{SS_c/(g-1)}{MS_E} \leq \frac{MS_{Trt}}{MS_E} \sim F_{g-1,N-g}$
  - The price for the nice properties are low power (meaning: test will not reject often when H_0 is not true)
- If F-test is not significant: Don't even have to start searching
- R
  - Calculate F-ratio $\frac{MS_c}{MS_E}$ as if "ordinary" contrast
  - Use $(g-1)\cdot F_{g-1,N-g,1-\alpha}$ as critical value (instead of $F_{1,N-g,1-\alpha}$)

Pairwise Comparison

- A pairwise comparison is nothing else than comparing two specific treatments
- This is a multiple testing problem because there are $g \cdot \frac{g-1}{2}$ possible comparisons (basically a lot of two-sample t-tests)
  - Simplest solution: apply Bonferroni correction
  - Better (more powerful): Tukey Honest Significant Difference

HSD

- Tukey HSD is better (more powerful) than Bonferroni if all pairwise comparisons are of interest
- Re-consider Bonferroni if only a subset of comparisons are of interest

Multiple Comparison with a Control (MCC)

- Do $g-1$ (pairwise) comparisons with the control group
- Dunnett procedure constructs simultaneous confidence intervals for differences $\mu_i - \mu_g$, $i = 1, \cdots, g-1$ (assuming group g is control group)

3 Factorial Treatment Structure

Setup

- Factor A with a levels
- Factor B with b levels
- $n > 1$ replicates for every combination (i.e. balanced design)
- Total of $N = a \cdot b \cdot n$ observations

Factorial Model
The two-way ANOVA model with interaction

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

- α_i is the main effect of factor A at level i
- β_j is the main effect of factor B at level j
- $(\alpha\beta)_{ij}$ is the interaction effect between A and B (not the product $\alpha_i\beta_j$)
- ϵ_{ijk} are i.i.d. $\mathcal{N}(0, \sigma)$ errors
- Typically, sum-to-zero constraints are used
  - $\sum_{i=1}^{a} \alpha_i = 0, \sum_{j=1}^{b} \beta_i = 0 \implies a-1$ and $b-1$ degrees of freedom
  - $\sum_{i=1}^{a} (\alpha\beta)_{ij} = 0, \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0 \implies (a-1)\cdot(b-1)$ degrees of freedom

Parameter Estimates

Parameter	Estimator
μ	$\hat{\mu} = \overline{y}_{...}$
α_i	$\hat{\alpha}_i = \overline{y}_{i..} - \overline{y}_{...}$
β_j	$\hat{\beta}_j = \overline{y}_{.j.} - \overline{y}_{...}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \overline{y}_{ij.} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$

Sum of Squares

Source	df	Sum of squares
A	$a-1$	$\sum_{i=1}^{a} b \cdot n \cdot \hat{\alpha}_i^2$
B	$b-1$	$\sum_{j=1}^{b} a \cdot n \cdot \hat{\beta}_j^2$
AB	$(a-1)\cdot(b-1)$	$\sum_{i=1}^{a}\sum_{j=1}^{b} n \cdot \widehat{(\alpha\beta)}_{ij}^2$
Error	$(n-1)\cdot ab$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \overline{y}_{ij.})^2$
Total	$abn-1$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \overline{y}_{...})^2$

ANOVA Table

Source	df	SS	MS	F
A	$a-1$	SS_A	$\frac{SS_A}{a-1}$	$\frac{MS_A}{MS_E}$
B	$b-1$	SS_B	$\frac{SS_B}{b-1}$	$\frac{MS_B}{MS_E}$
AB	$(a-1)\cdot(b-1)$	SS_{AB}	$\frac{SS_{AB}}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	$ab\cdot(n-1)$	SS_E	$\frac{SS_E}{(n-1)ab}$	

Single Replicates

- If we have a factorial experiment with only one observation per factor-level combination, we cannot fit a full model anymore
- Reason: Perfect fit! All residuals are zero (i.e. # parameters = # observations)

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

  - cannot distinguish between $(\alpha\beta)_{ij}$ and ϵ_{ij} in $n = 1$ situation
- We can still fit a model without interaction term, i.e. main effects only (= additive effects)
- If there is an underlying interaction term, we get an error estimate that is biased upwards (because it contains the error and the interaction term)
- Tests will be conservative (p-values will be too large), which is OK

- If we have no replicates and more than two factors, we would typically remove some of the higher-order interaction terms.
- This means: we put them into the error term (the df's of the error term will thus increase)
- Often: transformations of the response help getting rid of interactions

Source	df	Sum of squares (SS)
A	$a-1$	$\sum_{i=1}^{a} b \cdot \hat{\alpha}_i^2$
B	$b-1$	$\sum_{j=1}^{b} a \cdot \hat{\beta}_j^2$
Error	$(a-1)\cdot(b-1)$	$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-(\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j))^2$
Total	$ab-1$	$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\overline{y}_{..})^2$

- Alternative: Tukey one degree of freedom model for interaction

$$Y_{ij} = \mu + \alpha_i + \beta_j + \lambda\alpha_i\beta_j + \epsilon_{ij}$$

  - interaction is actually the product of the main effects
  - $H_0 : \lambda = 0$

Unbalanced Data

- We cannot estimate the parameters one at a time anymore
- Parameters have to be estimated simultaneously using the principle of least squares
- Sum of squares cannot be uniquely partitioned into different sources anymore
- Alternative to decomposition of sum of squares: using model comparison approach

Different Types of Sum of Squares
Notation: $SS(B \mid 1, A)$ is the reduction of the residual sum of squares when comparing the models $(1, A, B)$ with $(1, A)$

- Type I: Sequential sum of squares
  - Sequentially build up model
  - $SS(A \mid 1) \implies (1)$ vs. $(1, A)$
  - $SS(B \mid 1, A) \implies (1, A)$ vs. $(1, A, B)$
  - $SS(AB \mid 1, A, B) \implies (1, A, B)$ vs. $(1, A, B, AB)$
  - Depends on ordering of factors
  - R: `aov`
- Type II: Hierarchical/partially sequential approach
  - Control for the influence of the largest hierarchical model not including the term of interest
  - $SS(A \mid 1, B) \implies (1, B)$ vs. $(1, A, B)$
  - $SS(B \mid 1, A) \implies (1, A)$ vs. $(1, A, B)$
  - $SS(AB \mid 1, A, B) \implies (1, A, B)$ vs. $(1, A, B, AB)$
  - R: Function `Anova` in package `car`
- Type III: Fully adjusted/marginal approach
  - Control for all other terms
  - $SS(A \mid 1, B, AB) \implies (1, B, AB)$ vs. $(1, A, B, AB)$
  - $SS(B \mid 1, A, AB) \implies (1, A, AB)$ vs. $(1, A, B, AB)$
  - $SS(AB \mid 1, A, B) \implies (1, A, B)$ vs. $(1, A, B, AB)$
  - R: `drop1`

Comments

- With balanced data, we always get the same result, no matter what type we use.
- For main effects only models, Type II and Type III coincide.
- If there is a significant interaction, tests of the corresponding main effects are typically difficult to interpret (better to use individual models)

4 Randomized Complete Block Designs (RCBD)

Paired t-Test

- Compare two different eye-drops
- Every subject gets both treatments (meaning: one per eye at the same time)
- At the end, measure redness on quantitative scale in every eye
- For every patient, calculate the difference
- Perform standard one-sample t-test with these differences

Setup

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

- α_i: treatment effect
- β_j: block effect
- By only using main effects, we implicitly assume that the effects are additive
- Due to the balanced design, we can use our standard estimates (one at a time) and sums of squares

Randomized Complete Block Designs (RCB)

- A blocking design uses a restricted randomization scheme. Each block gets its own randomization
- Blocking exists at the time of randomization
- We call a blocking design complete if every treatment is used in every block (every block contains all treatments)
- In the standard setup, we observe every treatment only once in every block, hence we have a total of r (the number of blocks) observations per treatment
- Therefore, we have no replicates (every combination of treatment and block is only observed once)
- If we only have one observation per treatment and block combination, we could only detect interaction effects having the form of Tukey's one degree of freedom interaction

Factorials in Complete Block Designs

Source	df
Block	$r-1$
A	$a-1$
B	$b-1$
AB	$(a-1)\cdot(b-1)$
Error	$(ab-1)\cdot(r-1)$
Total	$rab-1$

- We can test the interaction AB even if we only have one replicate per AB combination per block

How Much Does Blocking Increase Precision

- Squared standard errors for treatment means are
  - Randomized Complete Block Design (RCB): $\frac{\sigma^2_{RCB}}{r}$
  - Completely Randomized Design (CRD): $\frac{\sigma^2_{CRD}}{n}$
- If we want to have the same precision, we have to ensure that

$$\frac{\sigma^2_{RCB}}{r} = \frac{\sigma^2_{CRD}}{n}$$

- If we knew σ_{RCB} and σ_{CRD}, we would have to use a ratio of

$$\frac{n}{r} = \frac{\sigma^2_{CRD}}{\sigma^2_{RCB}}$$

- σ^2_{RCB} is estimated by MS_E of our RCB

- CRD can be estimated using a properly weighted average of MS_E and MS_{Block}

$$\hat{\sigma}^2_{CRD} = w \cdot MS_{Block} + (1-w) \cdot MS_E$$

- Relative efficiency is defined as

$$RE = \frac{\hat{\sigma}^2_{CRD}}{\hat{\sigma}^2_{RCB}} = \frac{n}{r}$$

  - Sometimes multiply with correction factor for df's, only relevant for small samples
  - A CRD would need $\frac{n}{r}$ as many experimental units to achieve the same efficiency

Latin Squares

- A Latin Square needs to have
  - g treatments
  - two block factors having g levels each (the rows and the columns)
  - a total of g^2 experimental units
- Each treatment appears exactly once in each row and exactly once in each column
  - A Latin Square is nothing else than an assignment of treatments to units with the side constraints
- We see only g^2 out of g^3 possible combinations (but the subset is selected in a smart, balanced way)

Source	df
Rows	$g-1$
Columns	$g-1$
Treatments	$g-1$
Error	$(g-1)(g-2)$
Total	g^2-1

5 Random Effects

Random Effects Model

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

- α_i i.i.d.$\sim \mathcal{N}(0, \sigma^2_\alpha)$
- ϵ_{ij} i.i.d.$\sim \mathcal{N}(0, \sigma^2)$
- $\mathrm{Var}(Y_{ij}) = \sigma^2_\alpha + \sigma^2$
- $\mathrm{Cor}(Y_{ij}, Y_{kl}) = \begin{cases} 0 & i \neq k \quad \text{different machines} \\ \sigma^2_\alpha/(\sigma^2_\alpha + \sigma^2) & i = k, j \neq l \quad \text{same machine, intraclass correlation} \\ 1 & i = k, j = l \quad \text{same machine} \end{cases}$
  - Reason: Observations from the same machine "share" the same random value α_i and are therefore correlated
- Conceptually, we could also put all the correlation structure into the error term and forget about the α_i's, i.e. $Y_{ij} = \mu + \epsilon_{ij}$

Random vs. Fixed Models

Term	Fixed effects model	Random effects model
α_i	fixed, unknown constant	α_i i.i.d.$\sim \mathcal{N}(0, \sigma^2_\alpha)$
Side constraint on α_i	needed	not needed
$E[Y_{ij}]$	$\mu + \alpha_i$	μ, but $E[Y_{ij} \mid \alpha_i] = \mu + \alpha_i$
$\mathrm{Var}(Y_{ij})$	σ^2	$\sigma^2_\alpha + \sigma^2$
$\mathrm{Corr}(Y_{ij}, Y_{kl})$	0 if $i \neq k$ or $j \neq l$ 1 if $i = k$ or $j = l$	

- A note on the sampling mechanism
  - Fixed: Draw new random errors only, everything else is kept constant
  - Random: Draw new "treatment effects" and new random errors

Notes:

- Hierarchy is typically less problematic in random effects model
- More modern and flexible approach for estimating variance: Restricted Maximum-Likelihood estimator (REML)
- General rule: Variances are difficult to estimate in the sense that you will need a lot of observations to have some reasonable accuracy
  item Approximate confidence intervals (or tests) can be obtained by calling the function `confint`
- Exact tests (simulation based) for variance components can be found in the package `RLRsim`
- Typically, we are more interested in the accuracy of the variance component estimates (confidence intervals) than in tests of the form ($H_0 : \sigma_\alpha^2 = 0$ vs. $H_A : \sigma_\alpha^2 > 0$)
- We can also get "estimates" (more precisely: conditional means) of the random effects α_i with the function `ranef`
- As a new addition, we can also do QQ-plots of the α_i's (not only of the residuals)

ANOVA for Random Effects Models
One-way ANOVA (A random, n observations per cell)

Source	df	SS	MS	E[MS]
A	$g-1$	$\cdots$	$\cdots$	$\sigma^2 + n\sigma_\alpha^2$
Error	$N-g$	$\cdots$	$\cdots$	σ^2

Two-way ANOVA (A, B, AB random, n observations per cell)

Source	df	SS	MS	E[MS]
A	$a-1$	$\cdots$	$\cdots$	$\sigma^2 + b\cdot n\cdot\sigma_\alpha^2 + n\cdot\sigma_{\alpha\beta}^2$
B	$b-1$	$\cdots$	$\cdots$	$\sigma^2 + a\cdot n\cdot\sigma_\beta^2 + n\cdot\sigma_{\alpha\beta}^2$
AB	$(a-1)(b-1)$	$\cdots$	$\cdots$	$\sigma^2 + n\cdot\sigma_{\alpha\beta}^2$
Error	$ab(n-1)$	$\cdots$	$\cdots$	σ^2

- In a fixed effects model, the sum (or mean) of these interaction terms is zero by definition.
- In the random effects model, this is only true for the expected value, but not for an individual realization!
- We can test the interaction AB even if we only have one replicate per AB combination per block. Why?

6 Nested and Mixed Effects

Fully Nested Design

$$Y_{ijklm} = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + \delta_{l(ijk)} + \epsilon_{m(ijkl)}$$

ANOVA Table for Fully Nested Design (Balanced Design)

$$SS_{Total} = SS_A + SS_{B(A)} + SS_{C(AB)} + SS_{D(ABC)} + SS_E$$

Random effects and a balanced design

Source	df	E[MS]
A	$a-1$	$\sigma + n\sigma_\delta^2 + nd\sigma_\gamma^2 + ncd\sigma_\beta^2 + nbcd\sigma_\alpha^2$
$B(A)$	$a(b-1)$	$\sigma + n\sigma_\delta^2 + nd\sigma_\gamma^2 + ncd\sigma_\beta^2$
$C(AB)$	$ab(c-1)$	$\sigma + n\sigma_\delta^2 + nd\sigma_\gamma^2$
$D(ABC)$	$abc(d-1)$	$\sigma + n\sigma_\delta^2$
Error	$abcd(n-1)$	σ

7 Split-Plot Designs

- A split-plot design is a special case of a design with factorial treatment structure.
- The standard split-plot design consists of two experiments with different experimental units of different "size".
  - In the split-plot would, whole plots act as blocks
  - Each experiment has its own randomization
  - Each experiment has its own idea of experimental unit
- Split-plot designs can arise in more complicated forms

  - There can be more than one whole-plot factor
  - There can be more than one factor on the split-plot level
  - For each level (whole plot or split plot) of the experiment, we have to introduce a corresponding random effect which acts as the experimental error on that level
- The main effect of the whole-plot factor is estimated less precisely, and the test is less powerful (compared to the split-plot level)
  - Price for laziness on the whole-plot level (only a few observations)

8 Incomplete Block Designs

- In a complete block design, we could estimate the difference in each block with the same precision
- We call a design disconnected if we can build two groups of treatments such that it never happens that we see members of both groups together in the same block
  - In a disconnected design, it is not possible to estimate all treatment differences (with a fixed effects model)

Balanced Incomplete Block Designs (BIBDs)

- BIBD: all treatment pairs occur together in the same block equally often (we denote this number by λ).
- The precision (variance) of the estimated treatment differences $\alpha_i - \alpha_j$ is the same no matter what combination of i and j we are considering.
  - We can estimate all treatment differences with the same accuracy!
- Setup:
  - g: number of treatments
  - b: number of blocks
  - k: number of units per block with $k < g$ (i.e. block size)
  - r: number of replicates per treatment
  - N: total number of units

$$N = b \cdot k = g \cdot r$$

- A necessary but not sufficient condition for a BIBD to exist

$$\lambda \stackrel{?}{=} \frac{r \cdot (k-1)}{g-1}$$

  - $\frac{r \cdot (k-1)}{g-1}$ must be a whole number
  - Even if the condition is fulfilled, it might be the case that you cannot find a BIBD

Unreduced BIBDs

- We can always find a BIBD for every setting of $k < g$
- Unreduced BIBD
  - The number of combinations is $\binom{g}{k}$ (i.e. $\frac{g!}{k!(g-k)!}$)

(B)IBD: Intra-Block Analysis

- We do not observe all treatment$\times$block combinations, the "usual" estimates are not working, and we need computer to find the least squares estimates
- We use Type III sum of squares to test treatment effects adjusted for block effects

(B)IBD: Inter-Block Analysis

- It is possible to recover some information about the treatment by comparing different blocks
- The analysis when treating the block factor as random

Partially Balanced Incomplete Block Designs

- Some treatment Paris occur together more often than other pairs

Row-Column Incomplete Block Designs

- A row-column incomplete block design is a design where we block on rows and columns and one or both are incomplete blocks.

Row-Orthogonal Design

- The rows are complete blocks,while the columns form a BIBD.

Youden Square

- A Youden Square is rectangular such that
  - columns (rows) form a BIBD
  - rows (columns): each treatment appears equally often in each row (column)
- Columns form a BIBD, while rows form an RCBD

9 Power

- Error rates
  - Type I error: Reject H_0 even though it is true
  - Type II error: Fail to reject H_0 even though H_A holds
  - The probability of a type I error is controlled by the significance level α
  - The probability of a type II error is denoted by β (it is not being controlled)
- The power of a statistical test is defined as

$$P[\text{reject } H_0 \text{ given that a certain setting in the alternative } H_A \text{ holds}] = 1 - \beta$$

- Calculating power is like a "thought experiment": We do not need data, but a precise specification of the parameter setting under H_A that we believe in
- General rule: The more observations you have, the larger the power

10 Exams

- In an RCB, we can never test interactions between block and treatment. (False)
- Blocking can increase precision, even if the p-value corresponding to the block factor is not significant. (True)
- If we change the (family-wise) significance level from 5% to 10%, all the confidence intervals would be shorter and thus some treatment differences that were not significant may become significant.
- Type I sum of squares and Type II sum of squares of the main effects do typically not coincide for one main effect if the design is unbalanced.