Index of Lecture 9: Cross-over designs and Random effects models

PRACTICAL INFORMATION

Today's lecture:

- cross-over designs: based on lecture notes and references:[1]
  - GO Sections 13.3.1, 13.3.4, 13.3.6, 16.7,
  - [1] Senn (2005), Crossover Designs, In: *Encyclopedia of Biostatistics* (Moodle access),
  - [2] Toutenburg (2002), *Statistical Analysis of Designed Experiments*, Chapter 9 (Moodle access),
- start of random effects models:
  - in both VHM 802 and VHM 802/812; different focus:
    - today: ANOVA-based analysis, balanced data, continuous outcome,
    - VHM 802/812 (next week): likelihood-based analysis, unbalanced data, continuous and discrete outcomes,
    - next VHM 802 session (in two weeks): split-plot or hierarchical design,
  - textbook reading:
    - Chapter 11 (skip Sections 4, 6 and 8-9),
    - supplementary notes[2] on linear mixed models (random effects models).

Home assignments: # 3 returned today (with solution), # 4 posted (due March 17).

---

[1] Extra references at UPEI library: Senn (2002): Cross-Over Trials in Clinical Research; Jones & Kenward (1989): Design and Analysis of Cross-Over Designs.

[2] The notes contain essentially the same topics as the textbook, but omit detailed formulae and calculations for EMS-based estimates and tests, which are too technical for us; the notes are part of the course curriculum.

INTRODUCTION TO CROSS-OVER DESIGNS

Definition of a cross-over trial:

> "a trial in which individual subjects are given sequences of treatments with the object of studying differences between individual treatments (or subsequences of treatments)" [1]

- idea: each subject forms a block (or one's own control)
  ⇒ reduced variation (by eliminating between-subject variation),
- advantages:
  - (potentially) (much) higher efficiency than completely randomized design,
  - individuals' reactions to treatments can be studied,
- drawbacks:
  - longer execution time ⇒ logistical challenges, greater risk of subject dropouts,
  - more complex data analysis (due to added risk of bias caused by carry-over effects), and really a special case of repeated measures,
- simple 3 × 3 Latin square example: Bioequivalence (GO Example 13.6).

Common terminology:

- periods: occasions where subjects are treated (and measured),
- carry-over effect: residual effect from one period to the next,
- wash-out period: interval between tx periods to eliminate/reduce carry-over effects.

SIMPLEST DESIGN: AB/BA

- two treatments (A and B),
- two periods (1 and 2),
- two groups of subjects:

	Period 1	Period 2
subject group 1	A	B
subject group 2	B	A

Data example: Scents and learning[3] — can pleasant (floral) aromas help a student learn better?

- completion times for pencil and paper mazes for 21 subjects with scented (S) and unscented (U) masks,
- 11 subjects used unscented masks first, 10 subjects scented masks first (1 subject excluded):

Subject	Sequence	mean U	mean S	diff U−S	diff 1−2
1	US	30.6	38.0	-7.4	-7.4
2	SU	48.4	51.6	-3.2	3.2
3	US	60.8	56.7	4.1	4.1
4	SU	36.1	40.5	-4.4	4.4
5	US	68.5	49.0	19.5	19.5
6	SU	32.4	43.2	-10.8	10.8
...	...	...	...		

---

[3] Reduced data set (Mazes and smells) available at: `https://dasl.datadescription.com`.

ANALYSIS OF AB/BA DESIGN

Assuming no carry-over effects, simple approaches will work:

- compare treatments by two-sample analysis (e.g. t-test) for differences $1-2$,[4]
- compare periods by two-sample analysis (e.g. t-test) for differences A−B.

Combined analysis of response y_{ijk} in period j for k'th subject with treatment sequence i, using the model

$$y_{ijk} = \mu_{ij} + s_{ik} + \varepsilon_{ijk}, \qquad \text{where}$$

- μ_{ij} is the mean for sequence i in period j, given by:

sequence i	period 1 $(j=1)$	period 2 $(j=2)$
AB	$\mu+\alpha_A+\beta_1$	$\mu+\alpha_B+\beta_2+\lambda_{AB}$
BA	$\mu+\alpha_B+\beta_1$	$\mu+\alpha_A+\beta_2+\lambda_{BA}$

  $\alpha_A, \alpha_B \sim$ treatment effects; $\beta_1, \beta_2 \sim$ period effects; $\lambda_{AB}, \lambda_{BA} \sim$ carry-over effects,

- s_{ik} is the effect of subject k with treatment sequence i,
- ε_{ijk} is the error term $\sim \mathrm{N}(0, \sigma^2)$,
- same inference as above if $\lambda_{AB} = \lambda_{BA} = 0$,
- complex analysis ([1],[2]) if λ's $\neq 0$.

---

[4] Treatment comparison by a paired two-sample analysis for measurements for A and B (effectively a one-sample analysis for differences A−B) is only valid when no period effects exist.

MORE TREATMENTS AND PERIODS

Examples of extensions of AB/BA design:

- 2 tx, > 2 periods: enables modelling of carry-over effects in analysis, e.g. for sequences AABB and BBAA:

	Sequence				Sequence			
Carry-over model	A	A	B	B	B	B	A	A
"prev. period"	–	λ_A	λ_A	λ_B	–	λ_B	λ_B	λ_A
"change only"	–	0	λ_{AB}	0	–	0	λ_{BA}	0
"prev. + present"	–	λ_{AA}	λ_{AB}	λ_{BB}	–	λ_{BB}	λ_{BA}	λ_{AA}

- 3 tx, 3 periods: to ensure balancedness of tx's in periods traditionally laid out in Latin squares, e.g.,

$$\boxed{\begin{array}{ccc} A & B & C \\ B & C & A \\ C & A & B \end{array}} \qquad \boxed{\begin{array}{ccc} A & C & B \\ C & B & A \\ B & A & C \end{array}}$$

  where rows $\sim$ periods, columns $\sim$ subjects, symbols $\sim$ tx,

  - simple analysis when no carry-over effects,
  - desirable to include all sequences by combining two different Latin squares (as shown above),

- g tx, g periods: use (multiple) $g \times g$ Latin squares,
- g tx, $k < g$ periods: use incomplete block design, preferably BIBD (g, b, k, r, λ).

COMBINING LATIN SQUARES

Multiple Latin squares in same design:

- increases the error degrees of freedom ⇒ larger power,
- several extra options for modelling, depending on data context.

Notation/Model:
$$\begin{cases} y_{ijkl} = \text{outcome for tx } i \text{ in row } j \text{ and column } k \text{ in square } l, \\ y_{ijkl} = \mu + \alpha_i + \beta_{j(l)} + \gamma_{k(l)} + \varepsilon_{ijkl}, \end{cases}$$

- row ($\beta_{j(l)}$) and column ($\gamma_{k(l)}$) effects "nested in" (separate for) squares ∼ different effects across squares,[5]
- a block effect may be assumed the same in all squares, e.g. for periods:
  - same effects if all subjects go through same periods,
  - different effects if periods are not the same (e.g. due to different ages of subjects),

Modelling refinements:

- square type[6] interactions: carry-over effects may show up as interactions between square type and periods,
- residual effects may be modelled directly to split each tx effect into "direct" and "residual" effects.

---

[5] See 9L−14 for details about nesting, and examples for specification in software.

[6] The square type is determined by the carry-over combinations it contains.

Latin Square Cross-Over Trial Examples

Bioequivalence trial (GO Example 13.10) with 12 subjects in 4 Latin squares.

Milk production example ∼ cross-over trial (GO Example 13.12):

- milk yield of cows during three periods with different diets,

  $$\begin{aligned} y_{ijkl} &= \text{yield for cow } k \text{ in square } l \text{ in period } j \text{ on diet } i \\ i &= \text{A,B,C} \sim \text{diets (roughage, limited grain, full grain)} \\ j &= 1,2,3 \sim \text{period (for each cow)}, \\ k &= 1,\ldots,3 \sim \text{cow number (within squares)} \\ l &= 1,\ldots,6 \sim \text{Latin square number.} \end{aligned}$$

	Cow			Cow			Cow		
Period	1	2	3	7	8	9	13	14	15
1	A	B	C	A	B	C	A	B	C
2	B	C	A	B	C	A	B	C	A
3	C	A	B	C	A	B	C	A	B

	Cow			Cow			Cow		
Period	4	5	6	10	11	12	16	17	18
1	A	B	C	A	B	C	A	B	C
2	C	A	B	C	A	B	C	A	B
3	B	C	A	B	C	A	B	C	A

- 6 separate Latin squares,
  - two types of Latin squares (top/bottom), 3 replicates of each,
  - top ∼ diet order AB, BC, CA; bottom ∼ AC, CB, BA,
- **basic statistical model** (additive, no square effects),

  $$\begin{aligned} y_{ijkl} &= \mu + \alpha_i + \beta_j + \gamma_{k(l)} + \varepsilon_{ijkl}, \text{ or} \\ y_i &= \mu + \alpha_{\text{diet}(i)} + \beta_{\text{per}(i)} + \gamma_{\text{cow}(i)} + \varepsilon_i, \end{aligned}$$

  where the errors are i.i.d. and $\sim \mathrm{N}(0, \sigma^2)$.

INTRODUCTION TO RANDOM EFFECTS MODELS

Fixed and random effects:[7]

- "fixed" effect: modelled by usual parameters (e.g., βs) $\sim$ factors or regressors,
- "random" effect: modelled by random variables.

Terminology and relationships:

- mixed/random effects/variance component[8] models – the same,
- "mixed" $\sim$ containing both fixed and random effects,
- multi-level/hierarchical models – the same, and special type of mixed models.

Motivations for random effects (decreasing importance):

- data structure, e.g. hierarchical (next lecture),
- correct analysis of explanatory variables allocated to larger experimental units ("split-plot" idea: next VHM 802 lecture),
- factor with interest in variation between factor levels rather than actual levels in study:
  - levels may be randomly selected,
  - levels should represent "population".

  examples (1-way/2-way ANOVA models): cartons, laboratories.

---

[7] Standard notation (but not GO!): Greek letters for fixed effect parameters, Latin letters for random effect variables.

[8] Variance components are mathematical constructs used in mixed models.

RANDOM EFFECTS 1-WAY ANOVA

Carton example: (GO Section 11.1, Example 11.2)

- strength of cardboard carton produced by 10 machines (1 operator, 1 glue),

Outcome:	Machine (i)						
y_{ij} = strength	1	2	3	4	…	9	10
of carton j	157.97	160.28	156.81	156.26	…	149.88	153.33
from machine i	157.16	162.31	153.56	157.22	…	160.86	153.30

- random effects 1-way ANOVA model:

$$y_{ij} = \mu + A_i + \varepsilon_{ij}, \quad i = 1, \dots, a\,(10);\ j = 1, \dots, n\,(2),$$

  where the ε_{ij}'s are i.i.d. and $\sim \mathrm{N}(0, \sigma^2)$, and the A_i's are i.i.d. and $\sim \mathrm{N}(0, \sigma_A^2)$,

- interpretations:
  - A_i = random effect of ith machine (in principle, assumed drawn from a suitable population of machine effects),
  - σ_A^2 = variability between machines (in that population),
  - $\mathrm{Var}(y_{ij}) = \sigma_A^2 + \sigma^2$ (sum of two variance components $\sim$ sources of variation[9]),

- new issues for statistical analysis:
  - estimation of σ_A^2: $\hat{\sigma}_A^2 = (\mathrm{MS}_A - \mathrm{MS}_E)/n$,[10]
  - estimation of μ: $\hat{\mu} = \bar{y}_{..}$, $\mathrm{SE}(\hat{\mu}) = \sqrt{\mathrm{MS}_A/(an)}$,
  - extra model check: normality of A_i's (using $\bar{y}_{i.}$'s).

---

[9] The model has two sources of random variation: between machines and within machines (or between cartons).

[10] The formula comes from the mathematical fact: $\mathrm{E}(\mathrm{MS}_A) = \sigma^2 + n\sigma_A^2$.

Two Datasets (from the Notes)

Example 1: Laboratory testing (Environmental Research Institute, 1992)

- concentration of 4-methylphenol measured at different labs, selected among accredited labs for analysis,
- **purpose**: determine variation within and between labs,
- data for 5 labs,
  3 dilutions, and
  2 replicates per lab:

phenol (μg/l)	Dilution					
Laboratory	1		2		3	
A	5.5	4.7	9.8	10.3	11.6	11.8
B	7.7	7.5	12.4	12.5	16.4	17.0
C	7.4	7.1	12.5	11.8	15.9	16.2
D	6.5	7.1	10.0	9.4	12.6	12.7
E	6.5	7.0	11.0	9.9	13.5	12.7

Example 2: Pig breeding (Snedecor & Cochran, 1967)

- weight gain for pigs bred from different sires and dams,
- data for 5 sires,
  2 dams per sire,
  and 2 pigs per dam:

weight gain	Sire				
Dam	1	2	3	4	5
1	2.77	2.28	2.36	2.87	2.74
	2.38	2.22	2.71	2.46	2.56
2	2.58	3.01	2.72	2.31	2.50
	2.94	2.61	2.74	2.24	2.48

REPEATABILITY AND REPRODUCIBILITY

Fact: laboratory analyses do not give same result when replicated (on identical samples).

Sources of variation (assuming same laboratory method):

- laboratory, technician / equipment, "day" (time),
- pure replication error.

Idea: compute statistic to summarize variation — the value *not* exceeded with probability 95% by the difference between two measurements taken:

- under identical conditions → repeatability r,
- under "similar" conditions → reproducibility R.

Lab example (dilution 1):

- concentration of phenol in identical samples submitted to 5 laboratories (A-E),
- notation: y_{ij} = concentration measured for jth sample at laboratory i,
- model: $y_{ij} = \mu + A_i + \varepsilon_{ij}$, for i =A,B,C,D,E; $j = 1, 2$,
- formulae for repeatability and reproducibility:

$$\begin{aligned} \hat{r} &= 2\sqrt{2}\sqrt{\hat{\sigma}^2} = 2.83\sqrt{0.138} = 1.05, \\ \hat{R} &= 2\sqrt{2}\sqrt{\hat{\sigma}_A^2 + \hat{\sigma}^2} = 2.83\sqrt{0.852 + 0.138} = 2.82. \end{aligned}$$

Random Effects 2-Way ANOVA

Lab example (full data) — statistical model ($i \sim$ lab, $j \sim$ dilution, $k \sim$ replicate):

$$y_{ijk} = \mu + A_i + \beta_j + AB_{ij} + \varepsilon_{ijk}$$

- where $A_i \sim \mathrm{N}(0, \sigma_A^2)$, $AB_{ij} \sim \mathrm{N}(0, \sigma_{AB}^2)$, $\varepsilon_{ijk} \sim \mathrm{N}(0, \sigma^2)$,
- A_i = overall (across dilutions) random effect of lab i,
- AB_{ij} = dilution-specific random effect of lab i (at dilution j),
- ε_{ijk} = within-laboratory (and dilution) error,
- $\mathrm{Var}(y_{ijk}) = \sigma_A^2 + \sigma_{AB}^2 + \sigma^2$.

Source	DF	SS	MS	EMS	F	P
Lab	4	47.70	11.93	$\sigma^2+2\sigma_{AB}^2+6\sigma_A^2$	MSA/MSAB=8.59	0.005
Dilution	2	271.70	135.8	$\sigma^2+2\sigma_{AB}^2+\sigma_\beta^2$	MSB/MSAB=97.8	< 0.001
L * D	8	11.11	1.389	$\sigma^2 + 2\sigma_{AB}^2$	MSAB/MSE=8.61	< 0.001
Error	15	2.42	0.161	σ^2		
Total	29	332.93				

all effects significant ⇒ all variance components important (and dilutions!)

σ_β^2 = constant × $\sum_j \beta_j^2$

- estimates:

$$\begin{aligned} \hat\sigma^2 &= \mathrm{MS}_E = 0.161, \quad \hat\sigma_{AB}^2 = (\mathrm{MS}_{AB} - \mathrm{MS}_E)/2 = 0.614, \quad \hat\sigma_A^2 = (\mathrm{MS}_A - \mathrm{MS}_{AB})/6 = 1.756, \\ \hat r &= 2\sqrt{2} \times \sqrt{\hat\sigma^2} = 2.83\sqrt{0.161} = 1.13, \\ \hat R &= 2\sqrt{2} \times \sqrt{\hat\sigma_A^2 + \hat\sigma_{AB}^2 + \hat\sigma^2} = 2.83\sqrt{1.756 + 0.614 + 0.161} = 4.50. \end{aligned}$$

OVERVIEW: ANALYSIS OF RANDOM EFFECTS MODELS

- 2 types of statistical analysis:
  - ANOVA-based – "linear model methods + exceptions": *same* ANOVA table except for ***F***-tests, explicit formulae/rules to deal with exceptions,
  - "likelihood"-based – general method: no explicit formulae, requires good software,
- 2 data situations:
  - balanced data – simple case, both methods give same results,[11]
  - unbalanced data – complex analysis, only likelihood-based methods generally work well,
- different types of software/procedures:
  - a) ANOVA-based analysis, requires knowledge of specific formulae (essentially, a manual analysis): Minitab, Stata (`anova` command), SAS (`proc ANOVA`),
  - b) ANOVA-based analysis but automatic analysis given the design: Minitab (`General Linear Model`) and SAS (`proc glm`), except for problems with standard errors for fixed parameter estimates,
  - c) likelihood-based analysis and automatic analysis given the design: Minitab (`Mixed Effects Model`), SAS (`proc mixed`), R/S-Plus (`lme` library), Stata (`mixed`).

---

[11] When using the REML version of likelihood-based analysis.

BUILDING RANDOM EFFECTS MODELS

Multifactorial random effects models? yes, sure — some guidelines and ideas:

- several random effects: indeed possible (lab testing example) but increase complexity of model and analysis,
- random effects for predictors: possible but difficult,
- interaction between fixed and random effects: always random effect,
- interaction between two fixed effects: usually fixed effect (advanced: random possible),
- nesting: (of one factor within another)
  - a (random) factor B is nested within A, if there is no relation between levels of B across levels of A,
  - opposite of crossed factors A and B,
  - model formula notation: B(A) (Stata: B|A ; R: A/B),
  - modelling: no main effect of B (because B is meaningless without A); pig breeding example: $\boldsymbol{y_{ijk} = \mu + \alpha_i + B_{ij} + \varepsilon_{ijk}}$.

Analysis of general, balanced random effects models:

- ***F*-tests** and variance component estimation based on E(MS) formulae: use "clever" software or textbook tables,
- residuals for random effects: can be checked by looking at residuals in analysis of means for corresponding factor.