Biostats VHM 802 Course
Winter 2006, Atlantic Veterinary College, PEI
Henrik Stryhn

Index of Lecture 9

Practical Information

Today's lecture:

- random effects models:
  * course "highlight" !
  * really only an introduction (comprehensive topic!),
  * in both VHM 802 and VHM 802/812; different focus:
    - today: ANOVA-based analysis, balanced data, continuous outcome,
    - VHM 812: likelihood-based analysis, unbalanced data, both continuous and discrete outcomes,
- split-plot or hierarchical design:
  * a special type of design which is modelled/analysed by random effects models.
- Stata do-file: at website only (no Stata today).

Textbook reading:

- order 12.4–1–2 (*skip* 12.3 on multivariate ANOVA),
- supplementary notes on linear mixed models (random effects models), part of course curriculum.

Other news / practical information:

- move lecture March 21 $\rightarrow$ March 22, 1:30-4:30pm ?
- time to start working on your projects...

Introduction to Random Effects Models

Fixed and random effects

- "fixed" effect: modelled by usual parameters (β's: fixed, unknown constants $\sim$ factors or regressors),
- "random" effect: modelled by random variables.

Terminology and relationships:

- $\left.\begin{array}{l}\text{mixed}\\ \text{random effects}\\ \text{variance component}\end{array}\right\}$ models — the same,
- "mixed" $\sim$ containing both fixed and random effects,
- $\left.\begin{array}{l}\text{multi-level}\\ \text{hierarchical}\end{array}\right\}$ models — the same,
  and special type of mixed models,
- variance components are mathematical constructs used in mixed models.

Motivations for random effects (decreasing importance):

- data structure, e.g. hierarchical (later slide),
- correct analysis of explanatory variables allocated to larger experimental units ("split-plot" idea: later slides),
- factor where interest is in variation between levels rather than specific levels in study:
  - levels may be randomly selected,
  - levels should represent "population",

  Examples (1-way/2-way ANOVA models): strips, laboratories.

RANDOM EFFECTS 1-WAY ANOVA

Strips example: (textbook 5.1.2 and 12.4.1)

- electrical char. of ceramic components for a phonograph,

strip i	outcome/response variable y_{ij}						
1	17.3	15.8	16.8	17.2	16.2	16.9	14.9
2	16.9	15.8	16.9	16.8	16.6	16.0	16.6
3	15.5	16.6	15.9	16.5	16.1	16.2	15.7
4	13.5	14.5	16.0	15.9	13.7	15.2	15.9

  notation: y_{ij} = characteristic for piece j of strip i,

- random effects 1-way ANOVA model:

  $y_{ij} = \mu + A_i + \varepsilon_{ij}, \quad i = 1, \ldots, a(4); \; j = 1, \ldots, N(7),$

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

- interpretations:
  - A_i = random effect of ith strip (in principle, assumed drawn from a suitable population of strips),
  - σ_A^2 = variability between strips (in that population),
  - random variation : *between strips* and *between pieces*,
  - $\mathrm{Var}(y_{ij}) = \sigma_A^2 + \sigma^2$ (sum of two *variance components*),

- new issues for statistical analysis:
  - estimation of σ_A^2: $\hat{\sigma}_A^2 = (\mathrm{MSA} - \mathrm{MSE})/N$, (derived from the math. fact: $\mathrm{E}(\mathrm{MSA}) = \sigma^2 + N\sigma_A^2$),
  - estimation of μ: $\hat{\mu} = \bar{y}_{\cdot\cdot}$, $\mathrm{SE}(\hat{\mu}) = \mathrm{MSA}/(aN)$,
  - extra model check: normality of A_i's (using $\bar{y}_{i\cdot}$'s).

Two Datasets

Ex. 1: Laboratory testing (Environmental Res. Inst., 1992)

- conc. of 4-methylphenol measured at different labs,
- labs selected among accredited labs for analysis,
- purpose: determine variation within and between labs,
- data for 5 labs, 3 dilutions and 2 samples:

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

Ex. 2: Pig breeding (Snedecor & Cochran, 1967)

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

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 (that is, 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} = conc. measured for jth sample at lab. i,
- model: $y_{ij} = \mu + A_i + \varepsilon_{ij}$, for i=A,B,C,D,E; j=1,2,
- formulas 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, $k \sim$ rep.):
$$y_{ijk} = \mu + A_i + \beta_j + AB_{ij} + \varepsilon_{ijk}$$
where $A_i \sim \mathrm{N}(0, \sigma^2_A)$, $AB_{ij} \sim \mathrm{N}(0, \sigma^2_{AB})$, $\varepsilon_{ijk} \sim \mathrm{N}(0, \sigma^2)$,
- A_i = overall (across dilutions) random effect of lab i,
- AB_{ij} = dilution-specific random effect of lab i (at j),
- ε_{ijk} = within-laboratory (and dilution) error,
- $\mathrm{Var}(y_{ijk}) = \sigma^2_A + \sigma^2_{AB} + \sigma^2$.

Source	DF	SS	MS	EMS	F	P
Labs	4	47.70	11.93	$\sigma^2+2\sigma^2_{AB}+6\sigma^2_A$	MSA/MSAB=8.59	0.005
Dilu.	2	271.70	135.8	$\sigma^2+2\sigma^2_{AB}+\sigma^2_\beta$	MSB/MSAB=98.0	< 0.001
L $*$ D	8	11.11	1.389	$\sigma^2 + 2\sigma^2_{AB}$	MSAB/MSE=8.61	< 0.001
Error	15	2.42	0.161	σ^2		
Total	29	332.93				

- all effects clearly significant $\Rightarrow$ all variance components of importance (+ dilution differences, of course!),
- estimates:

$$\begin{aligned}
\hat{\sigma}^2 &= \text{MSE} = 0.161, \\
\hat{\sigma}^2_{AB} &= (\text{MSAB} - \text{MSE})/2 = 0.614, \\
\hat{\sigma}^2_A &= (\text{MSA} - \text{MSAB})/6 = 1.756, \\
\hat{r} &= 2\sqrt{2} \times \sqrt{\hat{\sigma}^2} = 2.83\sqrt{0.161} = 1.13, \\
\hat{R} &= 2\sqrt{2} \times \sqrt{\hat{\sigma}^2_A + \hat{\sigma}^2_{AB} + \hat{\sigma}^2} = 2.83\sqrt{1.756 + 0.614 + 0.161} = 4.50.
\end{aligned}$$

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 formulas/rules to deal with exceptions,
  - * "likelihood"-based — general method: no explicit formulas, requires good software,
- 2 data situations:
  - * balanced data — simple case, all methods give same results,
  - * unbalanced data — complex analysis, only likelihood-based methods generally work well,
- different types of software/procedures:
  a) ANOVA-based analysis, requires knowledge of specific formulas (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: SAS (`proc mixed`), R/S-Plus (`lme` library), Stata (`xtmixed`), SPSS (version 11).

BUILDING RANDOM EFFECTS MODELS

Multifactorial / Regression-type 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 regressors: possible but *difficult*,
- interaction between fixed and random effects: always random effect,
- interaction between two fixed effects: fixed[1] effect,
- 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),
  - ∗ modelling: no main effect of B (as B is meaningless without A); pig breeding ex.: $y_{ijk} = \mu + \alpha_i + B_{ij} + \varepsilon_{ijk}$.

Analysis of general, balanced random effects models:

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

[1] An interaction between two fixed effects may be taken as random, to allow inference for main effects in presence of significant interaction (!!).

Hierarchical Data Structure

A hierarchical data structure:

- observations grouped at different levels,
- treatments applied at different levels.

may induce clustering in the data, that is, some observations are more alike than others, or put in another way: the observations are no longer independent.

Typical example from veterinary epidemiology:

```
region
  └── farm/herd
        └── flock/pen
              └── litter
                    └── animal
                          └── (time)
```

Note: "time" as a bottom level $\sim$ longitudinal data / repeated measures on the same animal (next lecture), and raises some additional modelling issues.

Random effects models for hierarchical data:

- insert random effect(s) for each hierarchical level (above the bottom level).

Split-Plot Designs

Consider a two-factor design with blocks or replications:

- treatments (factors) A at a levels, and B at b levels,
- ab experimental units in each of c blocks, or replicated c times (assuming a balanced design).

Randomized (block) design:

- units allocated randomly to treatments (A$\times$B),
- all treatments estimated with same precision.

Split-plot design:

- A varies between (is applied to) larger units than B (for convenience or of necessity),
- conceptual two-step design construction:
  - i) A is allocated randomly to large units ("whole plots"),
  - ii) units for A are split into sub-units ("sub-plots"), onto which B is allocated randomly,
- A = "whole plot factor" and B = "sub-plot factor",
- (textbook): B–units *subsampled* from the A–units,
- implication: A estimated with less precision than B and the interaction A$*$B (not intuitively obvious!).

Analysis of split-plot design:

- include random effects for A–units in model!

Absorption Example

Garner's dynamic absorption data (textbook 12.1.1):

- moisture absorption by water-repellant cloth after laundry treatment of 6 yard cloths, which are subsequently divided into 1.5 yard cloths and measured at 4 different laboratories,

	Repl. i=A				Repl. i=B			
	Laboratory k				Laboratory k			
Laundry j	A	B	C	D	A	B	C	D
1	7.20	11.70	15.12	8.10	9.06	11.79	14.38	8.12
2	2.40	7.76	6.13	2.64	2.14	7.76	6.89	3.17
3	2.19	4.92	5.34	2.47	2.69	1.86	4.88	1.86
4	1.22	2.62	5.50	2.74	2.43	3.90	5.27	2.31

  notation: y_{ijk} = moisture absorption measured at lab. k for laundry j and in replication i,

  * whole plots = 6 yard strips (for each repl×laundry),
  * whole plot factor = laundry,
  * sub-plots = 1.5 yard strips (for each lab×repl×laundry),
  * sub-plot factor = laboratory (test in RC),
  * block = replication,

- split-plot model with blocks:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (AB)_{ij} + \gamma_k + (\beta\gamma)_{jk} + \varepsilon_{ijk},$$

  * $\mathrm{Var}(\varepsilon_{ijk}) = \sigma^2$ (variation between sub-plots),
  * $\mathrm{Var}(AB_{ij}) = \sigma^2_{AB}$ (extra variation btw. whole plots).

Absorption Example Results

ANOVA table:

Source	DF	SS	MS	EMS	F	P
Laundry	3	298.33	99.4	$\sigma^2+4\sigma^2_{AB}+\sigma^2_\beta$	MSB/MSAB = 125	< 0.001
Lab	3	102.92	34.3	$\sigma^2+\sigma^2_\gamma$	MSC/MSE = 60.1	< 0.001
L ∗ L	9	26.87	3.0	$\sigma^2+\sigma^2_{\beta\gamma}$	MSBC/MSE = 5.23	0.005
Block	1	0.01	0.0	$\sigma^2+4\sigma^2_{AB}+\sigma^2_\alpha$	MSA/MSAB = 0.01	0.933
Whole plot	3	2.38	0.8	$\sigma^2+4\sigma^2_{AB}$	MSAB/MSE = 1.39	0.293
Subplot	12	6.85	0.6	σ^2		
Total	31	437.36				

- clearly significant effects of both treatments and their interaction ⇒ interaction plot,
- no significance of blocks or whole-plot variation; keep in model anyway because part of design,
- variance components:
  $\hat\sigma^2 = \text{MSE} = 0.571, \quad \hat\sigma^2_{AB} = (\text{MSAB} - \text{MSE})/4 = 0.056,$

Standard errors of means, and pairwise comparisons:

- whole-plot factor (lab): usual formulae, using MSAB instead of MSE,
- subplot factor (laundry): pairwise comparisons by usual formulae using MSE,
- interaction: within whole-plot factor pairwise comparisons by usual formulae using MSE,

— otherwise by approximation method (details in notes).

Abrasion Resistance Example

Abrasion resistance data (textbook 11.3.1, 12.2.1, 13.3.1):

- fabric resistance computed as weight losses of pieces of fabric during 0–1000, 1000–2000 and 2000–3000 rotations (revolutions),

Surf. treat.	Filler type	Filler prop. 25% Rotations 1000	2000	3000	Filler prop. 50% Rotations 1000	2000	3000	Filler prop. 75% Rotations 1000	2000	3000
yes	A	194	192	141	233	217	171	265	252	207
		208	188	165	241	222	201	269	283	191
	B	239	127	90	224	123	79	243	117	100
		187	105	85	243	143	110	226	125	75
no	A	155	169	151	198	187	176	235	225	166
		173	152	141	177	196	167	229	270	183
	B	137	82	77	129	94	78	155	76	91
		160	82	83	98	89	48	132	105	67

$$
\begin{aligned}
y_{ijklr} &= \text{resist. of fabric } l \text{ of type } (i,j,k) \text{ after } r \text{ rotations,} \\
i &= 1,2 \sim \text{surface treatment (yes,no)} \\
j &= 1,2 \sim \text{type of filler (A,B),} \\
k &= 1,2,3 \sim \text{proportion of filler (25\%, 50\%, 75\%),} \\
l &= 1,2 \sim \text{replication,} \\
r &= 1,2,3 \sim \text{rotation interval (1000,2000,3000).}
\end{aligned}
$$

- *may be considered* as a split-plot design with 2 replications,
  - ∗ whole plots = pieces of fabric,
  - ∗ whole plot factors = all treatments (surf,fill,prop),
  - ∗ sub-plots = weight loss measurements,
  - ∗ sub-plot factor = rotation intervals.

  actually *repeated measures* on each piece of fabric. . .

Abrasion: Hierarchical Model

$$\begin{aligned} y_{ijklr} \;=\; & \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} \\ & + \delta_r + (\alpha\delta)_{ir} + (\beta\delta)_{jr} + (\gamma\delta)_{kr} + (\alpha\beta\delta)_{ijr} + (\alpha\gamma\delta)_{ikr} \\ & + (\beta\gamma\delta)_{jkr} + (\alpha\beta\gamma\delta)_{ijkr} + (ABC)_{ijkl} + \varepsilon_{ijklr}, \end{aligned}$$

where the whole plot errors $ABC_{ijkl} \sim \mathrm{N}(0, \sigma_w^2)$ and the subplot errors $\varepsilon_{ijklr} \sim \mathrm{N}(0, \sigma^2)$, and all variables are independent.

Analysis:

- model specification in Minitab: `rep(surf fill prop)`, and random `rep` effect,
- variance components: $\hat{\sigma}_w^2 = 40.7$, $\hat{\sigma}^2 = 189.7$,
- residuals: subplots ok, whole plots from analysis of whole plot means also ok,
- interaction `surf*fill*rota` clearly signif. $\Rightarrow$ look for patterns:
  - `surf*fill` at different `rota`-levels: present only at `rota`=1000,
  - `surf*rota` at different `fill`-levels: present only at `fill`=B and `rota`=1000,
- interaction `fill*prop*rota` "close" to significant, but nothing interesting seen in the two interaction plots (RC: linear and quadratic contrasts for both factors),
- interaction `prop*rota`: "close" to significant, but nothing interesting seen in the interaction plot,
- interaction `prop*fill`: interaction plot shows approximately linear effect of `prop`, and that the slope is non-zero (positive) only for `fill`=A. (note: linear effect difficult to analyse in Minitab)
- presentation of results: separate for `fill=A` and `fill=B` (because fill in both interactions) by graphs/lsmeans with standard errors.

Final Remarks on Split-Plot Designs

Split-plot vs. hierarchical data structure:
much the same thing. . . :

- sub-plots constitute a level below the mainplots:

```
mainplot                        herd
  └── sub-plot   similarly to     └── animal
```

- in a hierarchical structure, upper level units $\sim$ mainplots and lower level units $\sim$ sub-plots,
- explanatory variables may vary at upper or lower levels (possibly not randomised as in a split-plot design).

Two types of split-plot designs:

- with blocks (absorption example),
- with replicates instead of blocks (resistance example).

How to identify a "split-plot" structure?:

- *measurement* units are smaller than *experimental* units (for some treatment factors),
- expl. variables vary between "units" of different size,
- some units have been *subsampled*
  $\Rightarrow$ subsamples $\sim$ subplots, units $\sim$ mainplots,
- ask yourself: how many exper. units for this factor?
- draw a sketch of the experimental layout!