Design Tableau: An aid to specifying the linear mixed model for a comparative experiment

Alison Smith and Brian Cullis

Centre for Bioinformatics and Biometrics
NIASRA, University of Wollongong, Australia

AASC18, December 2018
Rotorua, NZ

Design Tableau

Overview of talk

- Motivation
- Calf-feeding example
  - Randomisation distribution
  - Analysis of Variance (ANOVA)
  - Linear Mixed Model (LMM)
- Design Tableau approach
  - Definitions
  - Essential steps
  - Application to calf-feeding example
- Design Tableau for non-orthogonal designs
  - All steps
  - Application to frost expression example
- Summary

Motivation

A personal history of ANOVA and REML/LMM

- A major focus over last 30 years: linear mixed models (LMM) for data from plant improvement programmes
- Comparative experiments: aim is to select “best” varieties
- Developed LMM to maximise accuracy of selection
- Often involve complex variance and correlation structures
  - separable autoregressive processes for field trend
  - factor analytic models for variety by environment interaction
  - genetic relatedness using pedigree or marker information
- How did we get here?

Motivation

A personal history of ANOVA and REML/LMM

- We trained and worked as young biometricians when analysis of variance (ANOVA) was primary method for comparative experiments
- GENSTAT was tool of trade so
  - "Block" and "Treatment" structures
  - Wilkinson and Rogers (1973) notation

  ingrained in our statistical thinking
- Despite complexity of our LMM we (attempt to) maintain these fundamental concepts, especially the link between design and analysis
- Are we outliers?

Motivation

Mis-use of LMM for comparative experiments

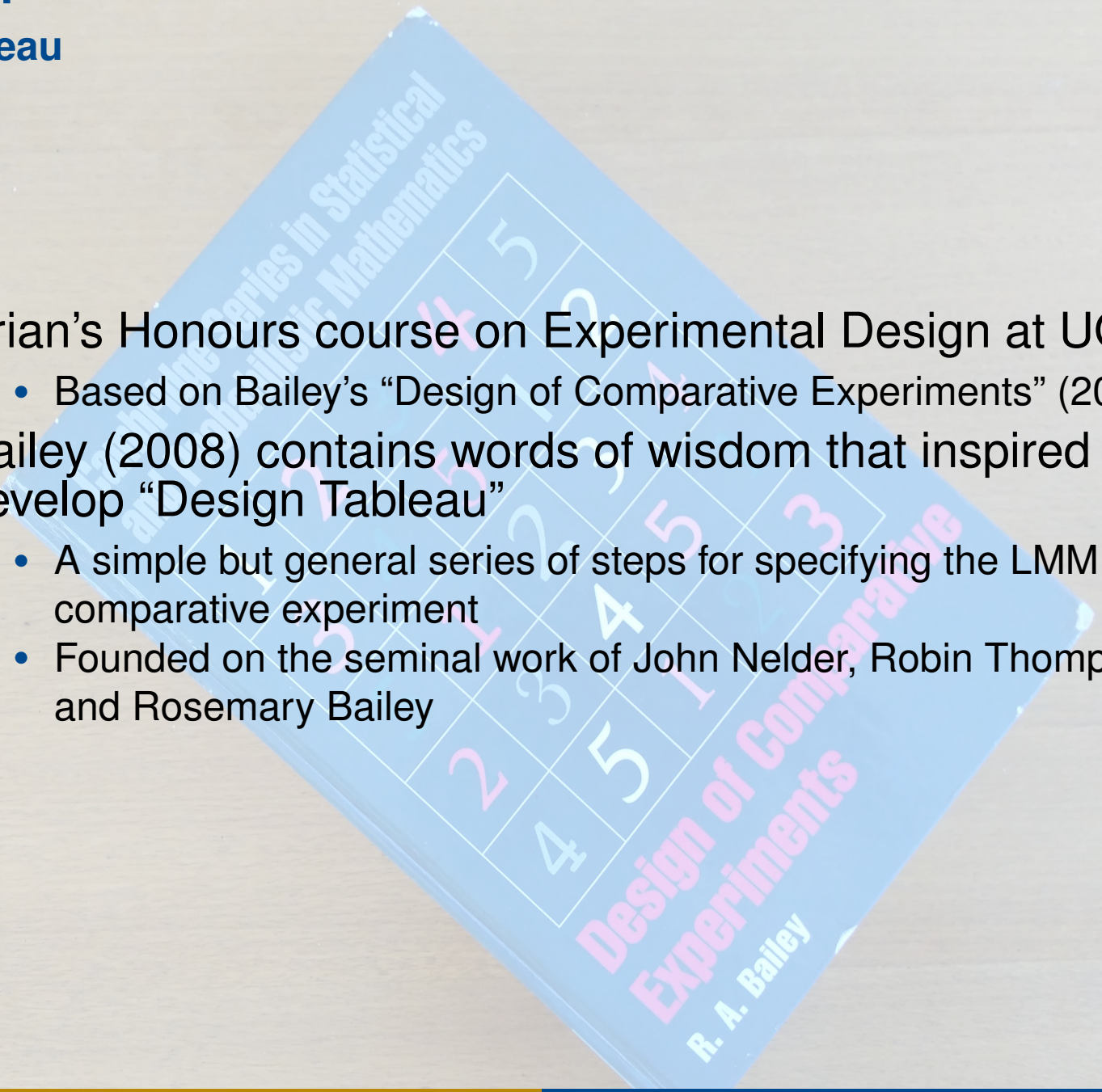
- With proliferation of LMM software (ASReml-R, SAS, lme, . . .), a move away from ANOVA techniques
- Literature full of examples of the mis-use of LMM for comparative experiments. Some common flaws include
  - failing to recognise pseudo (or false) replication
  - testing/dropping model terms that define strata
  - providing standard errors for means (not contrasts)
  - failing to recognise the need for negative estimates of variance components
  - failing to provide sufficient detail for reader to uncover some of these flaws!

Motivation

Mis-use of LMM for comparative experiments

- Perhaps an unintentional lapse in transitioning from ANOVA to LMM
- Perhaps a lack of exposure to traditional methods of analysis for comparative experiments
- In recent years, we have made it a priority to fill in this gap for young statistical colleagues in CBB at UOW
  - Link between ANOVA and REML/LMM
  - How to derive LMM that reflect experimental design, no matter how complex
- Non-trivial mentoring exercise!! Tried various approaches but no great success, until . . .

Motivation
Design Tableau



- Brian's Honours course on Experimental Design at UOW
  - Based on Bailey's "Design of Comparative Experiments" (2008).
- Bailey (2008) contains words of wisdom that inspired us to develop "Design Tableau"
  - A simple but general series of steps for specifying the LMM for a comparative experiment
  - Founded on the seminal work of John Nelder, Robin Thompson and Rosemary Bailey

- Design Tableau can be used for classical analyses of experiments with orthogonal designs
- But also (and more typically) for
  - complex experiments with non-orthogonal designs (eg. multi-environment trials, longitudinal data)
  - complex variance modelling (model based analysis)
- Let's start at the beginning . . .

Text-book example (Bailey, 2008)

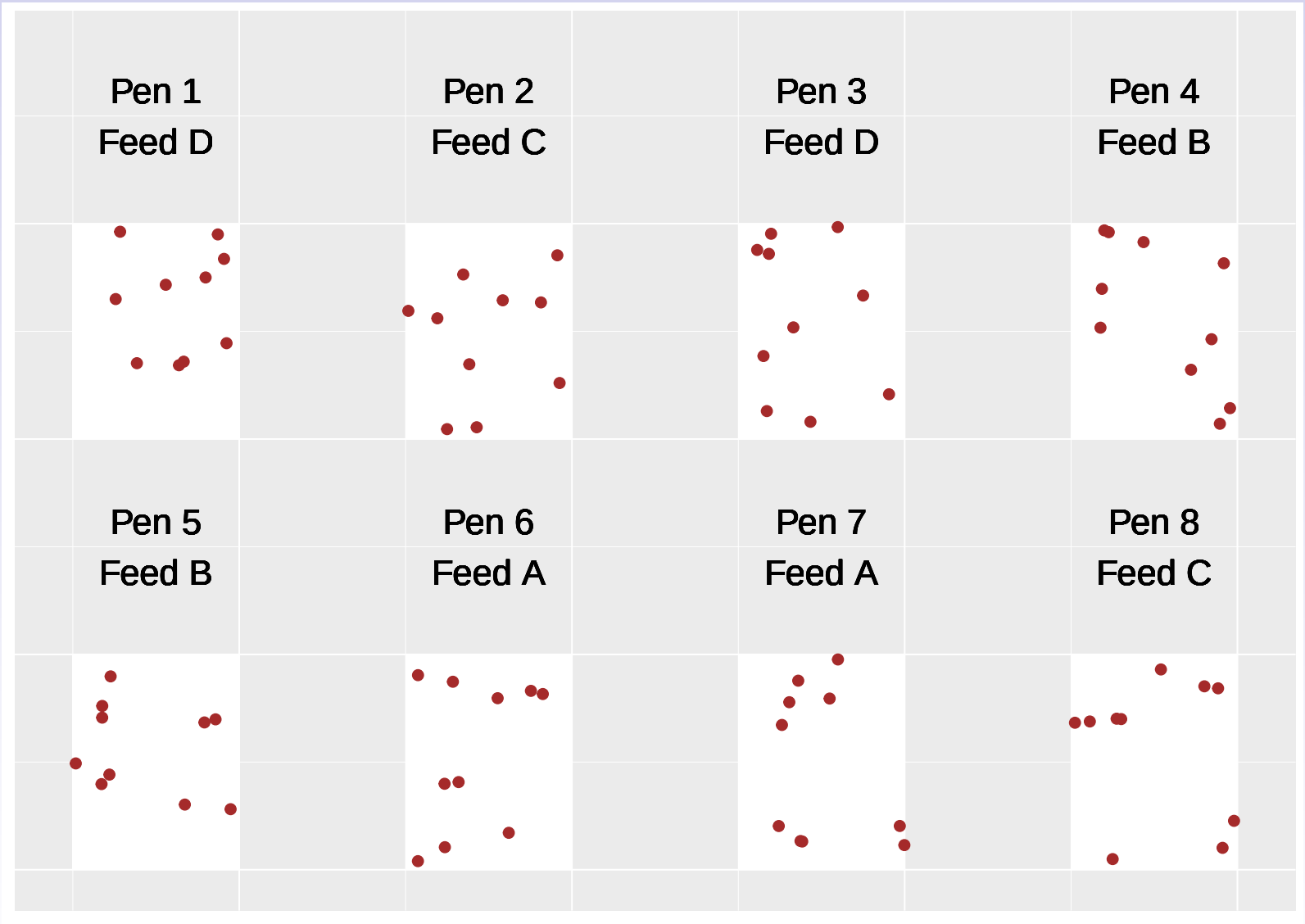
Calf feeding experiment

- Four ($t = 4$) feed treatments (A,B,C,D) are to be compared using $n = 80$ calves
- The calves are housed in $m = 8$ pens with $k = 10$ calves per pen so that $n = mk$
- Each pen allocated one of the four feeds (all calves within the pen consume the same feed)
- Calves are weighed individually at birth then at several times thereafter
- For illustrative purposes we assume variable to be analysed is average daily weight gain for each calf:

$$y = \frac{\text{final calf weight} - \text{initial calf weight}}{\text{number of days}}$$

Text-book example (Bailey, 2008)

Calf feeding experiment



Classical analysis
Randomisation theory

- Classical analysis for comparative experiments is based on randomisation theory (Nelder, 1954)
- Data are re-randomised to the observational units and inferences are based on the observed outcome of the resultant randomisation distribution
- This provides a platform for inference that is distribution-free

Calf feeding experiment

Randomisation distribution: the null experiment

- Consider the *null experiment*: all calves are assumed to receive the same treatment (Nelder, 1954, 1965a)
- To obtain the moments of the randomisation distribution the observed data are considered as given or known
- Let $x_{ij},\ i = 1 \ldots m, j = 1 \ldots k$ be the observed datum from the j^{th} calf in the i^{th} pen
- From these numbers form a set of random variables y_{ij} by
  - Choose a pen at random; re-order members at random to give $y_{11}, \ldots, y_{1k}$
  - Repeat procedure with another pen to give $y_{21}, \ldots, y_{2k}$
  - Repeat for all other pens

Calf feeding experiment

Randomisation distribution: the null experiment

- The (null) distribution of the y_{ij} is such that

$$\begin{aligned}
\mathrm{E}\left(y_{ij}\right) &= \mu_0 \\
\mathrm{var}\left(y_{ij}\right) &= \sigma_y^2 \\
\mathrm{cov}\left(y_{ij},\ y_{ib}\right) &= \rho_1\sigma_y^2 \quad (j \neq b, \text{so 2 calves in same pen}) \\
\mathrm{cov}\left(y_{ij},\ y_{ab}\right) &= \rho_2\sigma_y^2 \quad (i \neq a, \text{so 2 calves in different pens})
\end{aligned}$$

- In vector notation, and assuming that the data are ordered as calves within pens

$$\begin{aligned}
\mathrm{E}\left(\boldsymbol{y}\right) &= \mu_0\mathbf{1}_n \\
\mathrm{var}\left(\boldsymbol{y}\right) &= \sigma_y^2\left[(1-\rho_1)\boldsymbol{I}_m \otimes \boldsymbol{I}_k + \rho_2\boldsymbol{J}_m \otimes \boldsymbol{J}_k + (\rho_1-\rho_2)\boldsymbol{I}_m \otimes \boldsymbol{J}_k\right]
\end{aligned}$$

where $\boldsymbol{J}_m$ is an $m \times m$ matrix with all elements equal to 1

Calf feeding experiment

Randomisation distribution and ANOVA

- Null Analysis of Variance (ANOVA) is built up by forming *strata* which are defined as the eigenspaces of $\mathrm{var}\,(\boldsymbol{y})$
- For calf experiment there are 3 eigenspaces, with dimensions $1,\ (m-1)$ and $m(k-1)$ and eigenvalues
  - $\xi_0 = \sigma_y^2(1-\rho_1) + \sigma_y^2 k(\rho_1 - \rho_2) + \sigma_y^2 mk\rho_2$
  - $\xi_1 = \sigma_y^2(1-\rho_1) + \sigma_y^2 k(\rho_1 - \rho_2)$
  - $\xi_2 = \sigma_y^2(1-\rho_1)$
- These will be called the “mean”, “pens” and “calves” strata

Calf feeding experiment

Randomisation distribution and ANOVA

- We can then re-express $\mathrm{var}\,(\boldsymbol{y})$ as

$$\mathrm{var}\,(\boldsymbol{y}) = \xi_0 \boldsymbol{P}_0 + \xi_1 \boldsymbol{P}_1 + \xi_2 \boldsymbol{P}_2$$

- The $\boldsymbol{P}_s,\ s = 0, 1, 2$ are orthogonal projection matrices that can be written as $\boldsymbol{K}_s \boldsymbol{K}_s^\top$

Stratum	$\boldsymbol{P}_s$	$\boldsymbol{K}_s$
mean	$\boldsymbol{J}_m \otimes \boldsymbol{J}_k/(mk)$	$\mathbf{1}_m/\sqrt{m} \otimes \mathbf{1}_k/\sqrt{k}$
pens	$\boldsymbol{I}_m \otimes \boldsymbol{J}_k/k - \boldsymbol{J}_m \otimes \boldsymbol{J}_k/(mk)$	$(\boldsymbol{I}_m - \boldsymbol{J}_m/m) \otimes \mathbf{1}_k/\sqrt{k}$
calves	$\boldsymbol{I}_m \otimes \boldsymbol{I}_k - \boldsymbol{I}_m \otimes \boldsymbol{J}_k/k$	$\boldsymbol{I}_m \otimes (\boldsymbol{I}_k - \boldsymbol{J}_k/k)$

- The strata define 3 independent linear models that are obtained by applying a one-to-one transformation of the data from $\boldsymbol{y}$ to $\boldsymbol{K}^\top \boldsymbol{y}$ where $\boldsymbol{K} = [\boldsymbol{K}_0\ \boldsymbol{K}_1\ \boldsymbol{K}_2]$

Calf feeding experiment

Randomisation distribution plus treatments

- We now consider the imposition of the treatments so that $\mathrm{E}(y_{ij}) = \mu_A,\ \mu_B,\ \mu_C$ or μ_D
- Thus the first and second moments of the distribution are given by

$$\begin{aligned} \mathrm{E}(\boldsymbol{y}) &= \boldsymbol{\mu} \otimes \mathbf{1}_k \\ \mathrm{var}(\boldsymbol{y}) &= \xi_0 \boldsymbol{P}_0 + \xi_1 \boldsymbol{P}_1 + \xi_2 \boldsymbol{P}_2 \end{aligned}$$

where $\boldsymbol{\mu} = (\mu_D, \mu_C, \mu_D, \mu_B, \mu_B, \mu_A, \mu_A, \mu_C)^\top$

Calf feeding experiment

Linear models for strata

- The 3 linear models associated with the strata are defined for $\boldsymbol{K}_s^\top \boldsymbol{y}, \quad s = 0, 1, 2$ with

Stratum	$\mathrm{E}\left(\boldsymbol{K}_s^\top \boldsymbol{y}\right)$	$\mathrm{var}\left(\boldsymbol{K}_s^\top \boldsymbol{y}\right)$
mean	$\bar{\mu}\sqrt{mk}$	ξ_0
pens	$\left(\boldsymbol{\mu} - \bar{\mu}\mathbf{1}_m\right)\sqrt{k}$	$\xi_1 \boldsymbol{I}_{(m-1)}$
calves	$\mathbf{0}$	$\xi_2 \boldsymbol{I}_{m(k-1)}$

  where $\bar{\mu} = \sum_{ij} \mathrm{E}\left(y_{ij}\right)/n$

- ξ_s called stratum variances
- Typically the 3 models are represented using an ANOVA table ...

Calf feeding experiment: ANOVA table

Stratum	Source	df	ms	E(ms)	VR
mean		1			
	Mean	1	ms_M	$f_0(\bar{\mu}) + \xi_0$	
	residual	0			
pens		7			
	Feed	3	ms_F	$f_1(\boldsymbol{\mu} - \bar{\mu}\mathbf{1}_m) + \xi_1$	ms_F/ms_P
	residual	4	ms_P	ξ_1	
calves		72			
	residual	72	ms_R	ξ_2	
	Total	80			

- Using Nelder (1965b) can show that information on
  - Mean entirely in mean stratum. Obtain best linear unbiased estimate (BLUE) of mean within this stratum.
  - Feed treatment contrasts entirely in pens stratum. Obtain BLUEs of contrasts within this stratum.
- Residual mean squares provide unbiased estimates of stratum variances; cannot estimate ξ_0 so arbitrarily set $\xi_0 = \xi_1$

Calf feeding experiment: ANOVA table

Stratum	Source	df	ms	E(ms)	VR
mean		1			
	Mean	1	ms_M	$f_0(\bar{\mu}) + \xi_0$	
	residual	0			
pens		7			
	Feed	3	ms_F	$f_1(\boldsymbol{\mu} - \bar{\mu}\mathbf{1}_m) + \xi_1$	ms_F/ms_P
	residual	4	ms_P	ξ_1	
calves		72			
	residual	72	ms_R	ξ_2	
	Total	80			

- In order to test hypothesis $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$ must assume multivariate Normal distribution, so

$$\boldsymbol{y} \sim N\left(\boldsymbol{\mu} \otimes \mathbf{1}_k, \;\; \xi_0\boldsymbol{P}_0 + \xi_1\boldsymbol{P}_1 + \xi_2\boldsymbol{P}_2\right)$$

- Then test H_0 by comparing the VR with an F-distribution on $(3, 4)$ df

Calf feeding experiment
ANOVA and Linear Mixed Model

- ANOVA model assuming multivariate Normal distribution:

$$\boldsymbol{y} \sim N\left(\boldsymbol{\mu} \otimes \mathbf{1}_k,\ \ \xi_0 \boldsymbol{P}_0 + \xi_1 \boldsymbol{P}_1 + \xi_2 \boldsymbol{P}_2\right)$$

- Except that must set $\xi_0 = \xi_1$ so

$$\begin{aligned} \boldsymbol{y} &\sim N \quad \left(\boldsymbol{\mu} \otimes \mathbf{1}_k,\ \ \xi_1 \left(\boldsymbol{P}_0 + \boldsymbol{P}_1\right) + \xi_2 \boldsymbol{P}_2\right) \\ &\sim N \quad \left(\boldsymbol{\mu} \otimes \mathbf{1}_k,\ \ \xi_1 \boldsymbol{I}_m \otimes \boldsymbol{J}_k/k + \xi_2 \boldsymbol{I}_m \otimes \left(\boldsymbol{I}_k - \boldsymbol{J}_k/k\right)\right) \end{aligned}$$

- We can fit this as a linear mixed model

Linear Mixed Model

- The linear mixed model (LMM) for the data vector $\boldsymbol{y}$ is

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$$

  - $\boldsymbol{\tau}$ is the vector of fixed effects with associated design matrix $\boldsymbol{X}$ (assumed full column rank)
  - $\boldsymbol{u}$ is the vector of random effects with associated design matrix $\boldsymbol{Z}$
  - $\boldsymbol{e}$ is the vector of residuals
- Variance models given by:

$$\text{var}\left(\boldsymbol{u}\right) = \boldsymbol{G} \quad \& \quad \text{var}\left(\boldsymbol{e}\right) = \boldsymbol{R}$$
$$\text{var}\left(\boldsymbol{y}\right) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^{\top} + \boldsymbol{R}$$

- Fitting the LMM $\Rightarrow$
  - Residual Maximum Likelihood (REML) estimates of variance parameters
  - Empirical Best Linear Unbiased Estimates (EBLUEs) of fixed effects
  - Empirical Best Linear Unbiased Predictions (EBLUPs) of random effects

Calf feeding experiment

Equivalence of ANOVA and Linear Mixed Model

- $\boldsymbol{\tau}$ is the $t-$ vector of fixed effects (overall mean and feed treatment effects) with associated design matrix $\boldsymbol{X}$ so that

$$\begin{aligned} \mathrm{E}(\boldsymbol{y}) &= \boldsymbol{X}\boldsymbol{\tau} \\ &\equiv \boldsymbol{\mu} \otimes \mathbf{1}_k \end{aligned}$$

- $\boldsymbol{u}$ is the $m-$ vector of random pen effects with associated design matrix $\boldsymbol{Z} = \boldsymbol{I}_m \otimes \mathbf{1}_k$
- Variance models given by:

$$\begin{aligned} \mathrm{var}(\boldsymbol{u}) &= \sigma_p^2 \boldsymbol{I}_m \;\; \& \;\; \mathrm{var}(\boldsymbol{e}) = \sigma^2 \boldsymbol{I}_{mk} \\ \mathrm{var}(\boldsymbol{y}) &= \sigma_p^2 \boldsymbol{I}_m \otimes \boldsymbol{J}_k + \sigma^2 \boldsymbol{I}_m \otimes \boldsymbol{I}_k \\ &\equiv \xi_1 \boldsymbol{I}_m \otimes \boldsymbol{J}_k/k + \xi_2 \boldsymbol{I}_m \otimes (\boldsymbol{I}_k - \boldsymbol{J}_k/k) \end{aligned}$$

where $\xi_1 = k\sigma_p^2 + \sigma^2$ and $\xi_2 = \sigma^2$

Calf feeding experiment

Equivalence of ANOVA and Linear Mixed Model

- Variance parameter estimates:

ANOVA	LMM	note/proviso
$\hat{\xi}_1,\ \hat{\xi}_2$	$\hat{\sigma}^2_p,\ \hat{\sigma}^2$ $\hat{\xi}_1 = k\hat{\sigma}^2_p + \hat{\sigma}^2$ $\hat{\xi}_2 = \hat{\sigma}^2$	allow $\hat{\sigma}^2_p < 0$

- Treatment effect estimates and inference:

ANOVA	LMM	note/proviso
$\hat{\mu}_i,\quad i{=}A,B,C,D$ se$(\hat{\mu}_i - \hat{\mu}_j)$ F test, df	$\hat{\mu}_i,\quad i{=}A,B,C,D$ se$(\hat{\mu}_i - \hat{\mu}_j)$ Wald test, df	se$(\hat{\mu}_i)$ not valid (ξ_0 not estimable) allow $\hat{\sigma}^2_p < 0$; use Kenward & Roger (1997) for Wald df

- How can we derive an appropriate LMM for a comparative experiment?
- We use an approach that we have called “Design Tableau”
- It can be used for quite complex non-orthogonal experiments, with the aim that it reproduces an ANOVA in orthogonal cases
- Design Tableau requires some definitions . . .

Comparative experiments

Some key definitions (Bailey, 2008)

- An *experimental unit* is the smallest unit to which a treatment can be applied
- A *treatment* is the entire description of what can be applied to an experimental unit
- An *observational unit* is the smallest unit on which a response will be measured. It is often called a *plot*.

Comparative experiments

Some key definitions (Bailey, 2008)

- All designs have three components:
  - A plot structure: meaningful ways of dividing up the set of all plots
  - A treatment structure: meaningful ways of dividing up the set of all treatments
  - A design function: manner in which treatments are allocated to plots
- Plot and treatment structures described using *factors*
  - Universal factor (a single level): must be both a treatment (“1”) and plot factor (“U”)
  - Aliasing of factors: “F” and “G” aliased if the same apart from names of their levels
  - A factor may occur in either the plot or treatment structure, but not both (Welham, pers comm)

Design Tableau

Essential steps

Step 0 Talk to researcher and draw a picture of the experimental layout!

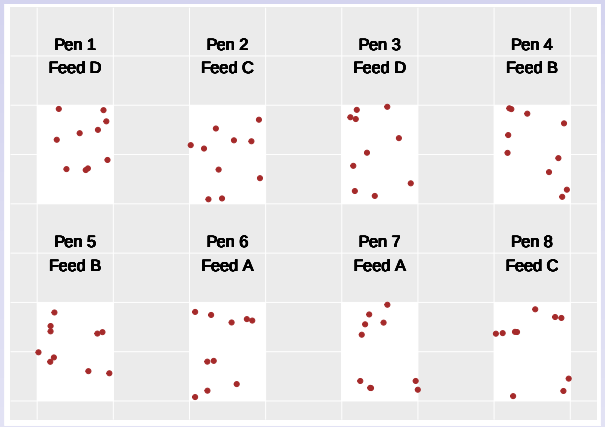
Step 1 Define treatments; list treatment factors

Step 2 Define plots (observational units); list plot factors

Step 3 Describe design function (how treatments are allocated to plots); thence define experimental unit

Calf feeding example

Step 0



Step 1

- treatments = feeds (4 treatments)
- treatment factors = { 1, Feed (4 levels) }

Step 2

- plots (observational units) = calves (80 units)
- plot factors = { U, Pen (8 levels), Calf (10 levels) }

Step 3

- design function: feeds allocated to calves such that all 10 calves within a pen receive same feed. Experimental units = pens

Design Tableau

Essential steps

Step 6 Use treatment factors to construct model formula for treatment structure (Wilkinson and Rogers, 1973, notation)

- Universal factor “1” included by default
- Terms in formula will be included in LMM as fixed effects

Step 7 Use plot factors to construct model formula for plot structure

- Combinations of levels of factors must completely index observational units
- Universal factor “U” included by default
- Terms in formula will be included in LMM as random effects, each set with IID variance structure

Design Tableau
Essential steps

Step 8 Identify obvious aliasing of factors in treatment structure with factors in plot structure eg. “1” and “U” (re-write as “1[U]” or “U[1]”)

Step 9 Construct a table (Design Tableau) listing all terms in the treatment model formula followed by terms in plot model formula

Step 11 Fit LMM commensurate with Design Tableau

Calf feeding example

Design Tableau

Step 6 treatment structure model formula:
1/Feed $\equiv$ 1 + Feed

Step 7 plot structure model formula:
U/Pen/Calf $\equiv$ U + Pen + Pen:Calf

Step 8 Aliasing: "1" and "U". Write "1[U]" where fitted as fixed and "U[1]" where fitted as random

Step 9 Design Tableau

Source	Term in model	Fixed or Random	Variance model
1[U]	1	F	
Feed	Feed	F	
U[1]	-	R	
Pen	Pen	R	$\sigma_p^2 \boldsymbol{I}_m$
Pen:Calf	Pen:Calf (= residual)	R	$\sigma^2 \boldsymbol{I}_{mk}$

Calf feeding example

Step 11 using ASReml-R (Butler et al, 2009)

- Fit linear mixed model:
  calf.asr <- asreml(y $\sim$ 1 + Feed, random = $\sim$ Pen, residual = $\sim$ units, data= ...)
  - **1 + Feed**: fixed model formula, includes overall mean **1** by default
  - **random = $\sim$ Pen**: random model formula, default IID variance model, default constrained positive
  - **residual = $\sim$ units**: residual model formula, default IID variance model for **units** (factor with n levels)
- Estimates, $\hat{\mu}_i$, and sed for feed means:
  predict(calf.asr, classify="Feed")
- Test hypothesis $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$
  Wald(calf.asr, denDF="algebraic")

Design Tableau for comparative experiments

Summary: orthogonal designs

- Have demonstrated how Design Tableau can be used to derive a LMM that is a surrogate for randomisation-based ANOVA for experiments with orthogonal designs.
- Some provisos . . .
  - Allow negative estimates of variance components so can reproduce strata for valid inference
  - Use Kenward & Roger (1997) df adjustments so can use correct reference distribution for F-tests
- Note that we do not attempt to structure Design Tableau table like an ANOVA (strata, sources within strata) since in non-orthogonal cases this is not possible

Design Tableau for comparative experiments

Summary: non-orthogonal designs

- Very few of the experiments we analyse use orthogonal designs!
- Also typically complex (unbalanced multi-environment trials; longitudinal data; multi-phase experiments with composite sampling . . .)
- But we always start with Design Tableau to obtain the terms that reflect the randomisation used in the experiment. This provides safe-guard against false replication, omission of strata, . . .
- For most experiments, Design Tableau provides base-line "working model" which we may extend in various ways eg. incorporate spatial correlation models for field trials, factor analytic models for variety by environment effects, . . .

Design Tableau

All steps (so far!)

Step 0 Picture of the experimental layout

Step 1 Define treatments; list treatment factors

Step 2 Define plots (observational units); list plot factors

Step 3 Describe design function; define experimental unit

Step 4 List *anatomical* variables, if any

Step 5 List *extraneous* variables, if any

Step 6 Use treatment factors and anatomical variables to construct model formula for treatment structure

Design Tableau

All steps (so far!)

Step 7 Use plot factors to construct model formula for plot structure

Step 8 Identify obvious aliasing between factors in treatment and plot structures

Step 9 Construct a table (Design Tableau) listing all terms in the treatment model formula followed by terms in plot model formula

Step 10 Possibly modify "working" table from **Step 9** eg.

- Incorporate more complex variance structures for random effects
- Selection experiments: move treatment effects from fixed to random

Step 11 Fit LMM commensurate with final Design Tableau

Example: non-orthogonal design

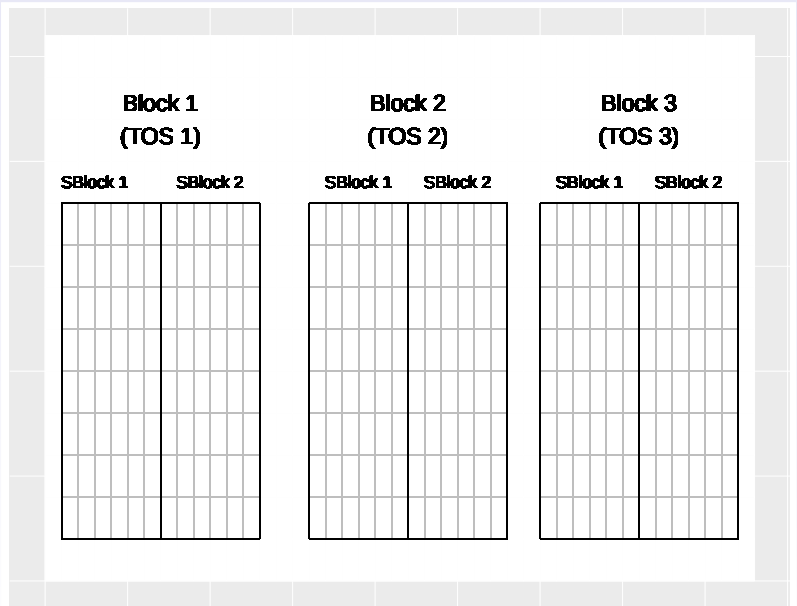
Frost expression experiments

- Frost damage a key issue for Australian wheat growers
- Frost expression experiments (FEEs) conducted at sites across Australia to provide information for growers on tolerance of commercial and near release varieties
- FEEs are field trials in which varieties exposed to natural frost events
- Variable of interest, frost induced sterility (FIS), obtained after frost events: ratio of number of sterile grains to total grains for individual tillers

Frost expression experiments

Protocol for single FEE

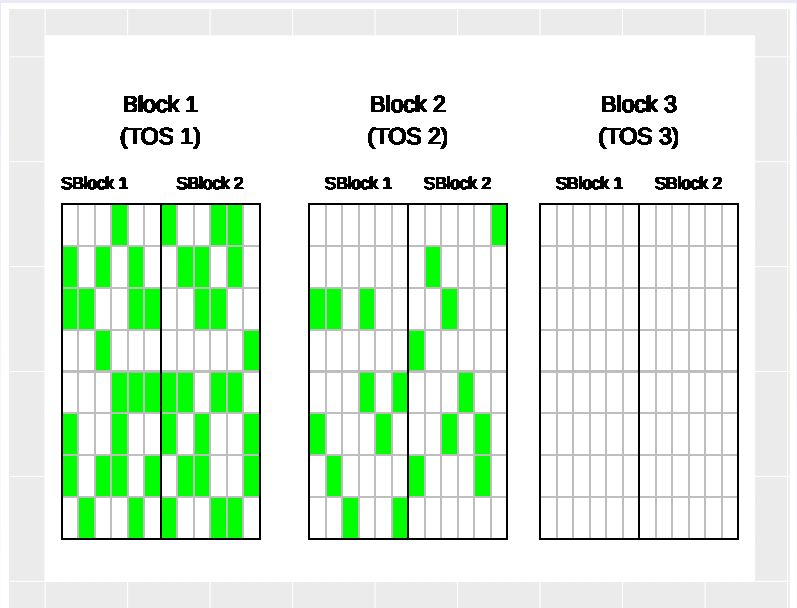
- Each variety grown at several times of sowing (TOS) to ensure some tillers at correct stage of development (SOD) when frost event occurs
- All plots for a TOS grouped together in a block
- Within each TOS block a randomised complete block design for varieties
- An illustrative example: 3 TOS blocks; 48 varieties; 2 replicates of varieties within each TOS block



Frost expression experiments

Protocol for single FEE

- After a frost event, researchers walk through the trial
  - Visually assses if any tillers in a plot are at an SOD of interest (flowering and ear peep)
  - If so, tag these tillers (up to a maximum of 30 per stage per plot), but leave the plant to continue growing
  - About 2 weeks after frost event, tagged tillers are cut and individually bagged; grains counted to provide FIS
  - Highly unbalanced: only a subset of plots measured for a single frost event (and number varies between TOS blocks); number of tillers measured in a plot varies between plots



Frost expression experiments

Some key issues

- Data for single frost event highly unbalanced
- Typically multiple frost events so potential for repeated measurements on a plot. Even more imbalance (number of repeated measurements per plot varies and may be 0)
- Aim is to assess variety tolerance but expect variety by TOS (careful!), variety by SOD and possibly variety by TOS by SOD interactions
- Finally there are 11 FEEs so a multi-environment trial analysis required to examine interactions with environment
- Where to begin?

- Where to begin?
- Start with Design Tableau for a single trial and frost event
- Illustrate some key points using simple example and assuming complete balance: 33 tillers measured in every plot (total of 9504 observational units)
- We can use ANOVA for this . . .

Frost expression experiments

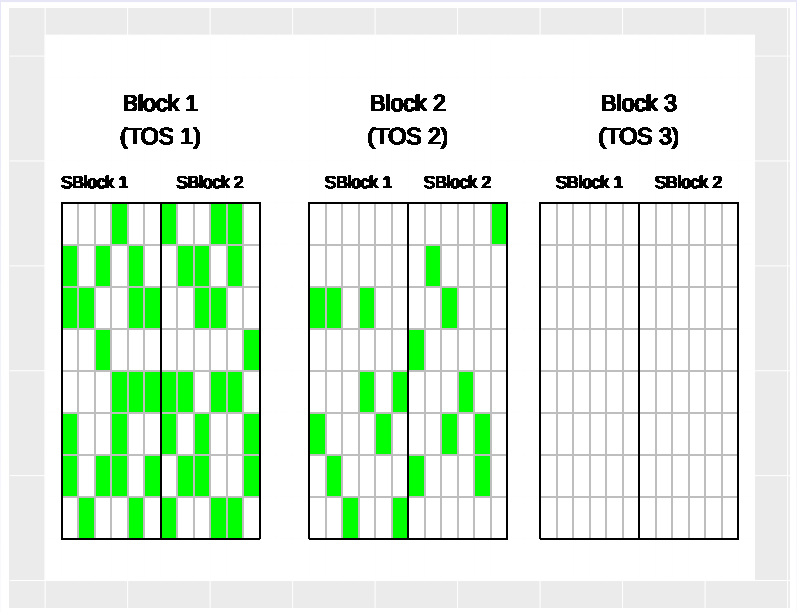
Single FEE and single frost event (balanced): ANOVA table

Stratum	Source	df
mean		1
	Mean	1
	residual	0
Block		2
	TOS	2
	residual	0
Block:SBlock	Block:SBlock	3
Block:SBlock:Plot		282
	Variety	47
	TOS:Variety	94
	residual	141
Block:SBlock:Plot:Tiller	residual	9216
Total		9504

Frost expression experiments

Single FEE and single frost event: ANOVA to Design Tableau

- TOS information is in Block stratum and no residual df (so no inference for testing)
- Variety information is in Block:SBlock:Plot stratum (not Block:SBlock:Plot:Tiller)
- How to capture all of this when we move into unbalanced scenario?
- How to add SOD?



Frost expression experiments

Design Tableau for single FEE and single frost event

- Now allow for unequal number of tillers measured per plot (assume max of 50); unequal number of plots measured (max of 48 per sub-block); introduce SOD

Step 1

- treatments = TOS x Variety combinations (144 treatments)
- treatment factors = { 1, TOS (3), Variety (48) }

Step 2

- plots (observational units) = tillers (n units)
- plot factors = { U, Block (3), SBlock (2), Plot (48), Tiller (50) }

Frost expression experiments

Design Tableau for single FEE and single frost event

Step 3

- design function: treatments allocated to plots so that
  - all tillers in same plot relate to same variety and TOS
  - each plot within sub-block allocated a different variety but same TOS
  - each sub-block within a block contains a single replicate of each variety and a single TOS
  - each block receives a different TOS

Step 4

- Anatomical variables: SOD (factor with 2 levels)

Step 6 treatment structure model formula:
1/(TOS*Variety*SOD)

Step 7 plot structure model formula:
U/Block/SBlock/Plot/Tiller

Step 8 Aliasing

- "1" and "U"
- TOS and Block (write TOS[Block] where fitted as fixed and Block[TOS] where fitted as random)

Step 9 Design Tableau (working table)

Source	Term in model	Fix/Ran	
1[U]	1	F	
TOS[Block]	TOS	F	
Variety	Variety	F	
TOS[Block]:Variety	TOS:Variety	F	
SOD	SOD	F	
TOS[Block]:SOD	TOS:SOD	F	
Variety:SOD	Variety:SOD	F	
TOS[Block]:Variety:SOD	TOS:Variety:SOD	F	
U[1]	-	R	
Block[TOS]	-	R	
Block[TOS]:SBlock	Block:SBlock	R	$\sigma_s^2 \boldsymbol{I}$
Block[TOS]:SBlock:Plot	Block:SBlock:Plot	R	$\sigma_p^2 \boldsymbol{I}$
Block[TOS]:SBlock:Plot:Tiller	Block:SBlock:PLot:Tiller (=residual)	R	$\sigma^2 \boldsymbol{I}$

Frost expression experiments

Design Tableau

- Have shown Design Tableau for single trial and frost event
- Extend to DT for single trial and multiple frost events
- Extend to DT for multiple trials and multiple frost events
- Finally modify LMM with complex variance models to accommodate multi-environment and longitudinal aspects
- See Cocks, March, Biddulph, Smith & Cullis (under revision) for full discussion, but here is final DT . . .

Frost expression experiments

Final Design Tableau

Source	Term in model	Fix/Ran	
1[U]	1	F	
Env[Expt]	Env	F	
Variety	-	R	
Env[Expt]:Variety	Env:Variety	R	$(\boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}) \otimes \boldsymbol{I}$
SOD	SOD	F	
Env[Expt]:TOS[Block]	at(Env,...):TOS	F	
Env[Expt]:SOD	at(Env,...):SOD	F	
Env[Expt]:TagEvent[Time]	at(Env,...):Time	R	$\oplus \boldsymbol{G}_{1i}$
Env[Expt]:TOS[Block]:Variety	at(Env,...):TOS:Variety	R	$\oplus \boldsymbol{G}_{2i}$
Env[Expt]:TOS[Block]:SOD	at(Env,...):TOS:SOD	F	
Env[Expt]:TOS[Block]:TagEvent[Time]	at(Env,...):TOS:Time	R	$\oplus \boldsymbol{G}_{3i}$
Env[Expt]:Variety:SOD	at(Env,...):Variety:SOD	R	$\oplus \boldsymbol{G}_{4i}$
Env[Expt]:Variety:TagEvent[Time]	at(Env,...):Variety:Time	R	$\oplus \boldsymbol{G}_{5i}$
Env[Expt]:TOS[Block]:Variety:SOD	at(Env,...):TOS:Variety:SOD	R	$\oplus \boldsymbol{G}_{6i}$
Env[Expt]:TOS[Block]:Variety:TagEvent[Time]	at(Env,...):TOS:Variety:Time	R	$\oplus \boldsymbol{G}_{7i}$
Env[Expt]:Tagger	at(Env,...):Tagger	R	$\oplus \boldsymbol{G}_{8i}$
Env[Expt]:Counter	at(Env,...):Counter	R	$\oplus \boldsymbol{G}_{9i}$
U[1]	-	R	
Expt[Env]	-	R	
Expt[Env]:Block[TOS]	-	R	
Expt[Env]:Block[TOS]:SBlock	at(Env,...):Block:SBlock	R	$\oplus \boldsymbol{G}_{10i}$
Expt[Env]:Block[TOS]:SBlock:Plot	at(Env,...):Block:SBlock:Plot	R	$\oplus \boldsymbol{G}_{11i}$
Expt[Env]:Block[TOS]:SBlock:Time[TagEvent]	at(Env,...):Block:SBlock:Time	R	$\oplus \boldsymbol{G}_{12i}$
Expt[Env]:Block[TOS]:SBlock:Plot:Time[TagEvent]	at(Env,...):Block:SBlock:Plot:Time	R	$\oplus \boldsymbol{G}_{13i}$
Expt[Env]:Block[TOS]:SBlock:Plot:Time[TagEvent]:Tiller	residual	R	$\oplus \boldsymbol{R}_i$

Frost expression experiments

Impact of Design Tableau

- Previous analyses of these data did not use our approach and failed to identify key issues; a loss of faith in results by industry
- With use of Design Tableau and close association with researchers we have regained industry and grower confidence in the results. Complete acceptance.

Design Tableau for comparative experiments

Summary

- We and our colleagues in CBB at UOW have been using Design Tableau for 12 months
- General consensus is that it is intuitive, straight-forward and helpful!
- Also useful for writing up statistical methods: reports for clients and journal papers
- We have used it for a wide range of (weird and wonderful) problems, including METs (with and without pedigree), GS, QTL detection, multi-phase

Design Tableau for comparative experiments

Summary

- Design Tableau can also be used for designs generated using model-based techniques (our typical paradigm)
- Even the most experienced biometricians can miss key features when using LMM to analyse comparative experimets
- We believe Design Tableau provides a framework to safeguard against this

Design Tableau for comparative experiments

Key references

- Bailey, R.A. 2008. *The design of comparative experiments*. Cambridge University Press.
- Nelder, J.A. 1954. *The Interpretation of negative components of variance*. Biometrika
- Nelder, J.A. 1965a. *The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance*. Proceedings of the Royal Society, A.
- Nelder, J.A. 1965b. *The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance*. Proceedings of the Royal Society, A.