Statistical Science Workshop 6-7 March 2025 University of Waikato – Tauranga
10-11 March University of Auckland – Leigh Marine Lab

Instructor: David Schneider. Memorial University, St. John's, Canada

Session 5. The General Linear Mixed Model GLMM

In this workshop you have learned to

- Translate a research question into a statistical model
- Execute the model and apply the model-checking loop
- Calculate a measure of evidence (the likelihood ratio)
- Calculate a measure of uncertainty on the likelihood ratio (p-value / confidence limit))
- Report effect sizes with a measure of uncertainty
- Interpret parameter estimates in light of the research question

> *..statistics must be relevant to making inferences in science and technology. The subject should be renamed statistical science and be focused on the experimental cycle, design-execute-analyse-predict.*
> *John Nelder 1999*

Learning goals of the fifth session.

Distinguish random and fixed factors in the design and analysis of experiments.

Write the expected Sums of Squares to determine the nested likelihood ratio (and F-ratio) for a mixed model.

Statistical Science Workshop

Session 5 – GLMM. The General Linear Mixed Model (Experimental Design)

Distinguishing random and fixed factors.

In all examples in Session 1 the explanatory variables were fixed.

1. Does the phosphorus content of corn increase when organic soil phosphorus is increased ?
   Based on nutrient requirements, we expect a positive relation. We wish to estimate the relation.

2. Does inversion heterozygosity (HZYG) change with elevation above sea level (Hsl) in *Drosophila persimilis*)?
   Hsl is a fixed effect because we expect a decrease in *Hzyg* in harsher environments at higher elevations.
   with elevation (altitude above sea level)

3. Does length depend on treatment (control versus 4 different sugars with auxin present) ?
   Treatment levels are fixed by design. We are inferring only to the levels in the experiment
   TRT is a fixed effect because we are interested in the contrast of 4 means, relative to the control.
   β_{TRT} is a set of unknown fixed effect contrasts.

4. Does change in inversion heterozygosity (HZYG) with elevation above sea level (Hsl) differ in 2 species ?
   The explanatory variable is fixed for several reasons.
   First of all, we expect a decrease.
   Second the study has a repeatable protocol. It can be repeated at the same elevations.
   The categorical variable is fixed because we are only inferring to these two species.
   We are not inferring to all species of fruit fly, based on a sample of only two species.
   Inferring to all species (taking species as random) is too much of a stretch.

In experimental designs, we have both random and fixed factors.
With fixed factors we infer only to the experimental units in the design, as described above.
With random factors we infer beyond the units measured, to a population of units.

Session 5 – GLMM. The General Linear Mixed Model (Experimental Design)

Examples of random factors.	Demonstration. Crossed design.	Hours of extra sleep	Ch13.3
	Demonstration. Nested design	Flies in cages	Ch13.6
	EMS to F-ratios	Flies in cages	
	Execution	Flies in cages	
	Execution	Wheat yields	

Wheat Yield Example
Cornell (1971) reported wheat yields (g) for 4 fertilizer treatments
No treatment
Straw
Straw + PO4
Straw + PO4 + lime
There were 3 pots/ treatment, three wheat plants/pot.
Does wheat yield depend on treatment?
  How good is the evidence? (LR)
  How certain is the evidence (p-value)
If so, devise three *a priori* hypotheses to identify the best treatment.