Natural Resources Data Analysis – Lecture Notes
Brian R. Mitchell

II. Week 2:

A. Questions from last week

1. Do multi-strata models reduce to a logistic model?

a) Multi-strata (or multi-state) models do not generally reduce to a logistic regression model, because they estimate fundamentally different quantities.

b) ***Logistic regression models*** estimate the probability of an observation being in one of two states, given knowledge of certain predictor variables. Note that multinomial logistic regression models can estimate probability of membership when there are more than 2 possible states.

c) ***Multi-strata models*** estimate the probability of an observation moving among states, given knowledge of the current state and certain predictor variables. Multi-strata models represent a first-order ***Markov process*** (i.e. the state at the next interval is dependent on the state at the current interval).

2. What are the ***degrees of freedom*** for a ***likelihood ratio*** χ^2 test based on the coin flip binomial data?

a) The appropriate statistic is actually ***2ln(LR)***, where LR is the likelihood ratio; this statistic is χ^2 distributed with ***1 degree of freedom***.

b) So for the case where L = 1.5, the χ^2 statistic is 0.811 and p = 0.368. When L = 20.5, the χ^2 statistic is 6.04, and p = 0.014.

B. Presentation of model sets

C. Sample size and pseudoreplication

1. Hurlbert (1984)

a) I think this paper is important as a good review of some basics of experimental design, as well as the still common mistake of pseudoreplication.

b) The major application of this paper, for the purposes of this course, is to get everyone thinking about what the true ***sample size*** is for their experiment, since this is needed for AIC_c and $QAIC_c$.

c) This paper is also important in its argument against using inferential statistics when they are not valid, and for arguing that appropriate interspersion trumps randomization.

d) ***Pseudoreplication*** is the use of inferential statistics to test for treatment effects when the treatments were not replicated or the replicates were not independent.

e) A big part of analyzing data is not necessarily knowing what the answer is, but knowing exactly what the question is that the statistic answers. This is the ***"Deep Thought" effect*** (from Hitchhiker's Guide to the Galaxy… getting the answer was relatively easy; knowing the question was much harder). Example: experiment to test difference in leaf decomposition at 1-m and 10-m in a lake. Used 8 bags at a 1-m location and 8 at a 10-m location. A significant t-test only shows that the two points differed; it does NOT mean that the difference is due to 1-m versus 10-m; would have to suitably sample both depths to determine if the difference is due to depth.

f) ***Simple pseudoreplication*** occurs when an experiment is not replicated; sometimes replication is not possible, but it is objectionable when tentative conclusions from unreplicated treatments are given a veneer of rigor by inappropriately using statistics. It is often advisable to take multiple samples per experimental unit to get more precise results; but don't confuse this with replication.

g) ***Temporal pseudoreplication*** occurs when multiple samples are taken from each experimental unit sequentially over time, and the sampling dates are taken to represent replicated treatments.

h) ***Sacrificial pseudoreplication*** involves true replication, but the data is pooled prior to statistical analysis, or where multiple measurements are taken from each replicate and treated as independent.

i) ***Implicit pseudoreplication*** occurs when authors do not present significance tests, but discuss effect sizes and confidence intervals as if significance was detected.

2. Oksanen (2001)

a) So, is pseudoreplication a pseudoissue? My personal feeling is that Oksanen misinterprets the fundamental nature of pseudoreplication – i.e. that pseudoreplication is 1) a failure to grasp the true question being answered by the statistics, and 2) a mistake in the determination of sample sizes for inferential

statistics. His insistence on using inferential statistics for unreplicated experiments strikes me as egregious.

b) What do people think about this comment: "Choosing two systems and assigning them randomly to a treatment and a control is normally an adequate design for a deductive experiment"? (Deduction = reasoning from general principles to predict specific results.) I think that even if you are using a deductive approach, you need to be very careful to match the systems, and that the results will be highly sensitive to random effects.

c) How about this comment: "Whether the experiment is replicated or not, inferential statistics should always be used, to enable the reader to judge how well the apparent patterns in samples reflect real patterns in statistical populations"? I think that inferential statistics should only be used when there are clearly defined populations being compared, and the comparison is worthwhile (i.e. no "silly nulls").

d) And this one: "Replicated experiments with compound treatments should never be referred to as pseudoreplicated, because all treatments are inherently compound"? I think that Oksanen misunderstood what Hurlbert was referring to. Oksanen seems to be referring to the fact that treatments always have some sort of unintended side effect (e.g. fence effect). Hurlbert would probably not doubt this, and would say that the treatment includes some unintended effects; the problem is in knowingly connecting things that should be separate. Hurlbert, for example, suggested independent filters for each tank, or one filter for all tanks.

e) Are people comfortable with Oksanen's faith in meta-analysis to make up for lack of replication?

3. Hurlbert's rebuttal (2004)

a) "Pseudoreplication in any of its various guises is simply an error of statistical analysis and interpretation."

b) "Confusion between effects of procedures used to impose treatments and effects of chance events impinging on an experiment (= non-demonic intrusion, Hurlbert 1984) is introduced in Oksanen's (2001) discussion of what he calls ''compound treatments''…. This general problem of how one controls for procedure effects has, however, nothing to do with pseudoreplication."

c) "Hurlbert (1984) concerns itself in no way… with the relative roles or importance of induction versus deduction, but only with

whether analyses and interpretations of experiments are concordant with the way experiments were designed and conducted."

d) "One must question the notion that 'our collective rate of progress' in large scale ecology will be maximized by allocating resources to large numbers of experiments lacking treatment replication and relying on meta-analysis rather than allocating the same resources to a smaller number of more expensive experiments with modestly replicated treatments. Meta-analysis is far from a methodological panacea that can compensate for the weaknesses of studies fed into it."

D. Computer exploration:

1. I have provided an Excel spreadsheet that calculates AIC, AIC_c, QAIC, $QAIC_c$, and BIC.

2. The data on the spreadsheet is from my dissertation; it is a confirmatory model set exploring factors that contributed to coyote approaches to a playback site. There are 34 models for a sample size of 448.

3. Explore the following questions:

a) How does changing the sample size affect AIC_c, $QAIC_c$, and BIC?

b) What is the effect of forgetting a parameter (across all models – e.g. you used least squares and forgot to add sigma)?

c) What if there were many fewer parameters (remove model 34, reduce K by 9 for all models, and then look at the difference a parameter makes).

d) What is the effect of different c-hat values?

e) What if the model set was smaller? Try all of the above using models 4 – 15 only.

4. Since we ran out of time in class this week, we will revisit this spreadsheet next week and discuss what we found.