Lab 9 Handout – The Robust Design

The appropriate reference material to this lab are Chapter 13 in Powel and Gale and Chapter 15 in The Gentle Introduction, as well as the citations I've provided in the syllabus as supplemental references this week's discussion of robust design models.

Overview: This semester we've covered mark-recapture analyses that have an open structure (CJS, multi-state, nest survival) where parameters of the model explicitly allow for animals to 'exit' the sample population through the mortality process. We've also covered closed capture models, where the model does not accommodate an open structure but does allow for estimation of abundance, which is facilitated by the assumption of closure and repeated captures. Today we will be working with an analysis type that is a hybrid of these two general model forms: the Robust Design.

The key characteristic of the Robust Design is that it breaks sampling occasions into two components: primary and secondary. The primary occasions correspond with the open component of the model and are most analogous to our CJS and multistate analyses. These are the occasions which are separated by intervals during which the sample is assumed to be open and where animals can exit, e.g. due to mortality. For many mark-recapture analyses, primary occasions are the individual study years.

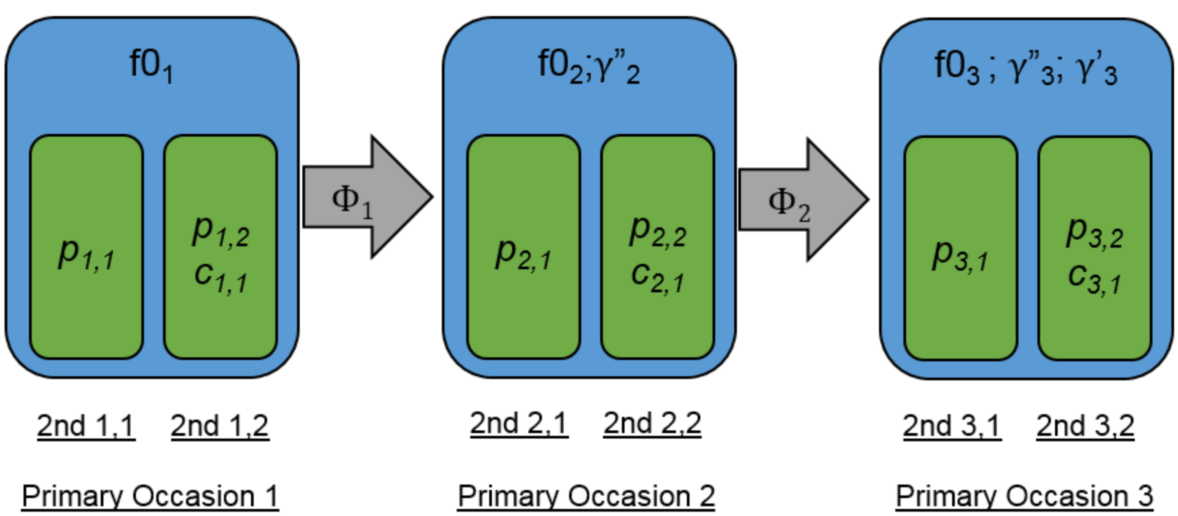
Secondary occasions occur within the primary occasions, and rely on the same assumptions of closure inherent to closed capture models. By dividing the primary occasions into secondary occasions and assuming closure, we can estimate detection probabilities from within each primary occasion (e.g. year) and need not rely on repeated observations among primary occasions to estimate detection. This effectively removes the confounding between phi and p that was present in the CJS model, and means that you will obtain "robust" estimates of phi for all intervals of your study, rather than *t*-1 intervals as with any CMR method based solely on an open capture structure.

With the robust design study design becomes very important, because your sampling much match the assumptions of both the open and closed components of the model. This is probably the most important for the secondary occasions in order to match the assumptions of closure. Your repeated sampling should therefore be relatively fine-scale in nature with respect to the biology of your study organism. For example if you are studying small mammals, secondary occasions based on nightly captures within a 1-week trapping session are probably appropriate, whereas if you were trapping only one day every month you'd certainly have mortality occurring in between sessions and would violate the closure assumption. However, see section 15.8 in the Gentle Introduction for a description of the Open Robust Design model, which can relax some of the closure assumptions. Also of importance is the presence of repeated captures among your primary occasions; if you only rarely encounter an animal twice within a single primary occasion, you will have issues estimating detection and the robust design may be of limited utility.

Temporary emigration: Under the most general form of the Robust Design we will estimate many of the same parameters we've already become familiar with (phi, p, c, f0), and will also introduce a fundamentally new parameter, γ, which gives the probability of temporary absence from the study area during one or more intervals. This parameter is also commonly referred to as temporary emigration, and we will further sub-divide it into two forms:

> γ'' – read as "Gamma double prime", the parameter represents the probability that an animal will be temporarily absent from the study area, and thus unavailable for detection, given that it was present within the study area during the previous primary occasion AND that it survived the interval to the present primary occasion. Said differently, γ'' reflects the probability of absence for a single primary occasion, and 1- γ'' gives us the probability of presence of an animal within the sampled area during that occasion.
>
> γ' – read as "Gamma prime", this parameter represents the probability that an animal will be temporarily absent given that it was also absent during the previous primary occasion. Again, this is conditioned on the animal surviving all intervals and also being unavailable for detection (due to absence). Said differently, γ' reflects the probability of repeated absence over >1 intervals, and 1- γ' gives us the probability that an animal remains present in the sampling area for two successive occasions.

Estimation of both forms of γ is facilitated by the ability to estimate detection probabilities under both the open and closed form of the model. See Section 15.2 in the Gentle Introduction for a thorough description γ estimation. A conceptual diagram of the robust design for three primary occasions, each with 2 secondary occasions, is provided below, with the estimable parameters identified as associated with the primary occasions (f0, γ'', γ'), secondary occasions (p, c), and open intervals (Φ).



Importantly you will notice that there are only two primary occasions where γ'' is included. This should be somewhat intuitive; if γ'' reflects temporary absence given prior presence, prior presence must first be established (in the first primary occasion) before absence can be estimated. Similarly, γ' is only included for the 3rd primary occasion, and again this should be intuitive; an animal can only be absent following prior absence if it was known to be previously absent, and for that to have occurred it must have been known to be present during the first primary occasion. Thus the take-home message here is that for every *t* primary occasions in your study, you will get *t*-1 estimate of γ'' and *t*-2 estimates of γ'.

Random vs Markovian Process: Having two forms of temporary absence allows us to evaluate the evidence for an additional process – that of random vs Markovian temporary emigration. A Markovian process, generally speaking, is any process where the current system state is dependent on the previous state. For example, changes in animal abundance through time (i.e. population dynamics) are inherently Markovian in nature, because the previous abundance is at least partially responsible for dictating the future abundance as a function of the population's growth rate. Population size is likely to be more similar during two successive years than during two years that are separated by a greater distance in time. In the context of temporary absence, a Markovian process would suggest that the prior presence or absence of an animal in the system would make the animal more or less likely (respectively) to be present during the subsequent occasion. The alternative to Markovian process is one that is random in nature, where prior state does not influence present state. In the context of temporary absence, a random process would indicate that you are as likely to be absent once as you are to be absent multiple time steps in a row.

To test for random vs Markovian structure on temporary absence, we can run alternative model structures where we either treat γ'' and γ' as two distinct components of the model (a Markovian process) or we constrain their model structures to be the same (a random process). If you think about it, when $\gamma'' = \gamma'$ (random process) we are saying that the probability of absence just once (γ'') is the same as being absent for multiple occasions (γ'), which indicates that the previous state is not a determining factor in the probability of absence. In contrast, when $\gamma'' \neq \gamma'$ (Markovian process) we have allowed differing probability of absence depending on previous state (present or absent) which fits the definition of a Markovian process. Using RMark, we will constrain $\gamma'' = \gamma'$ using a similar approach to what we used for constraining p and c when testing for behavioral effects on p (Note – we can also test for these behavioral effects in the Robust Design). For some additional context, below is a screen shot of what these two model approaches would look like in MARK itself using the design matrix.

0	0	0	0	0	0	0	0	0	0	0	6:S	0
1	1	0	0	0	0	0	0	0	0	0	7:Gamma"	0
1	0	1	0	0	0	0	0	0	0	0	8:Gamma"	0
1	0	0	1	0	0	0	0	0	0	0	9:Gamma"	0
1	0	0	0	1	0	0	0	0	0	0	10:Gamma"	0
1	0	0	0	0	1	0	0	0	0	0	11:Gamma"	0
1	0	0	0	0	0	0	0	0	0	0	12:Gamma"	0
0	0	0	0	0	0	1	1	0	0	0	13:Gamma'	0
0	0	0	0	0	0	1	0	1	0	0	14:Gamma'	0
0	0	0	0	0	0	1	0	0	1	0	15:Gamma'	0
0	0	0	0	0	0	1	0	0	0	1	16:Gamma'	0
0	0	0	0	0	0	1	0	0	0	0	17:Gamma'	0
0	0	0	0	0	0	0	0	0	0	0	18:p Session 1	1

Design Matrix example for a Markovian model structure where $\gamma'' \neq \gamma'$. The model includes a time structure for both Gamma terms.

0	0	0	0	0	0	6:S	0
1	1	0	0	0	0	7:Gamma"	0
1	0	1	0	0	0	8:Gamma"	0
1	0	0	1	0	0	9:Gamma"	0
1	0	0	0	1	0	10:Gamma"	0
1	0	0	0	0	1	11:Gamma"	0
1	0	0	0	0	0	12:Gamma"	0
1	0	1	0	0	0	13:Gamma'	0
1	0	0	1	0	0	14:Gamma'	0
1	0	0	0	1	0	15:Gamma'	0
1	0	0	0	0	1	16:Gamma'	0
1	0	0	0	0	0	17:Gamma'	0
0	0	0	0	0	0	18:p Session 1	0

Design Matrix example for a Markovian model structure where $\gamma'' = \gamma'$. The model includes a time structure that is shared, along with the model intercept, for both Gamma terms. Notice the shift in the first γ', which does not occur until the second time interval (see diagram and explanation of this above).

The Robust Design is a very data hungry analysis type, as you might imagine for a model that contains 6 different parameter types and where model complexity is compounded as you add more sampling occasions to your study. For that reason it often requires much more substantial datasets to implement a Robust Design that converges across all parameters, especially when modeling group and time structures. There are a few tricks you can use to implement the Robust Design a bit more effectively. One is to use the Huggins specification of the Robust Design, which removes f0 from the model likelihood (thus saving parameters) and instead estimates abundance as a derived parameter based on p* (see page 15-31 in the Gentle Introduction, with further explanation in Chapter 14 on closed capture models).

A second useful trick is to implement a model where we include a difference between γ"and γ', but allow them to share the same time structure. This is accomplished in MARK as follows:

0	0	0	0	0	0	0	6:S	0
1	1	1	0	0	0	0	7:Gamma"	0
1	1	0	1	0	0	0	8:Gamma"	0
1	1	0	0	1	0	0	9:Gamma"	0
1	1	0	0	0	1	0	10:Gamma"	0
1	1	0	0	0	0	1	11:Gamma"	0
1	1	0	0	0	0	0	12:Gamma"	0
1	0	0	1	0	0	0	13:Gamma'	0
1	0	0	0	1	0	0	14:Gamma'	0
1	0	0	0	0	1	0	15:Gamma'	0
1	0	0	0	0	0	1	16:Gamma'	0
1	0	0	0	0	0	0	17:Gamma'	0

You should notice that this approach is effectively the same as incorporating a group plus time structure, and it will serve to allow for a situation where the relative values for the two forms of absence differ from each other, but reflect a hypothesis where a shared mechanism drives the background temporal variability. As of the writing of this lab (8:24 pm on April the 2nd) I have not figured out how to implement this structure in RMark, but there is probably a way.

Lab Exercise and Homework:

Today's lab will implement a Robust Design model in RMark, with script provided on Blackboard. The script will show you how to set up the analysis, incorporate time intervals with differing numbers of secondary occasions, and how to run both a random and Markovian structure on temporary emigration.

The assignment for the week will be to complete this analysis using a comprehensive and defensible approach to model selection that explores 1) temporal and behavioral effects on detection probability, 2) temporal effects on survival, 3) temporal and Markovian process in temporary emigration, and 4) generates model-averaged estimates of N.