WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

Lecture 04 – The Mayfield method & Link functions

Part I – The Nesting Model

Readings

Mayfield, H. 1975. Suggestions for calculating nest success. Wilson Bull. 87:456-466.

Dinsmore, S. J., G. C. White, and F. L. Knopf. 2002. Advanced techniques for modeling avian nest survival. Ecology 83:3476-3488.

Other references

Bart, J., and D. S. Robson. 1982. Estimating survivorship when the subjects are visited periodically. Ecology 63:1078-1090.

Mayfield, H. 1961. Nesting success calculated from exposure. Wilson Bull. 73:255-261

Johnson, D. J. 1979. Estimating nest success: the Mayfield method and an alternative. Auk 96:651-661.

Johnson, D. J., and T. L. Shaffer. 1990. Estimating nest success: when Mayfield wins. Auk 107:595-600.

Rotella, J.J., S. J. Dinsmore, T.L. Shaffer. 2004. Modeling nest-survival data: a comparison of recently developed methods that can be implemented in MARK and SAS. Animal Biodiversity and Conservation 27:187-204

The Mayfield Method for estimating DSR

Although it was developed for estimating nest success Mayfield's method has been commonly applied to studies of animal survival throughout the biological literature. The calculations for the basic Mayfield estimator like those for the Kaplan-Meier estimator are fairly simple and straightforward. However, the assumptions are more rigorous and the estimate is not distribution free. In this section we will cover the assumptions and development of the basic Mayfield method and an example.

Mayfield's (1961, 1975) method germinated from observations made during a study of Kirtland's warbler. The approach is most commonly used when examining the success rates of nests, but like the known fate models it is frequently extended to the study of survival of marked or sessile animals. Mayfield's observation was that the typical measure of nesting success, hatching rate of active nests, (number of nests that hatched eggs/number active nests found; hereafter referred to as apparent success), increased for nests found later in the season. This phenomenon resulted from the fact that not all nests were found as soon as they were initiated, and led to an upward bias in estimates of nest success. The true nature of the bias is associated with the exclusion of lower detection rate for nests that fail before discovery. Mayfield's solution was to estimate daily survival probabilities as opposed to hatching rates.

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

Mayfield also insisted that nests not followed to termination (i.e., hatch or failure) should be used in estimates of nest success, and that failure to do so biased survival rates downward. Subsequently, he demonstrated that reasonable estimates of nest success could be derived from daily survival probabilities by monitoring the status of a sample of nests periodically without the prerequisites of discovery at the time of initiation or knowledge of their ultimate fate. Finally, Mayfield's method was also applicable when the exact failure time was not known.

1. Study design and assumptions

    a. Individuals (nests) marked or uniquely identifiable

    b. Individuals are periodically monitored to determine status

    c. Subjects can be relocated without failure

    d. Censoring is possible

    e. Staggered entry is possible

    f. Data requirements: record of monitoring history for each individual including id, date, time (if pertinent), and status.

    g. Assumptions

        1) Animals in the population of interest have been sampled randomly – sex, age, location, etc. (i.e., the sample is representative)

        2) Survival times are independent for different animals (nests). If not, the variance will be biased low and survival rates may also be biased.

        3) Time of death is known exactly

        4) Capturing, marking, and observing do not influence survival

        5) Censoring is random (i.e., unrelated to fate – either survival or mortality)

        6) "Staggered entry" assumes that newly tagged animals have the same survival function as those previously tagged

        7) Constant daily survival probability over time and across individuals.

        8) Detection probabilities for survival and mortality are equal

2. Basic formulas

d – number of deaths

    exposure –time (in nest days) at risk.

    t – duration of the study period.

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

Daily survival rate (probability) *dsr* – is the portion of nests expected to survive a single day.

$$\hat{dsr} = 1 - \left(\frac{d}{exposure} \right)$$

Survival rate ($\hat{S}$)– is the portion of nests will survive the period of study. Similar to the product limit method

$$\hat{S} = \left(\hat{dsr} \right)^{t}.$$

Thus, a constant *dsr* must be assumed using this calculation. Studies are frequently divided into strata that correspond with different stages in life history or important seasons and separate *dsr* can be calculated for each. However, within periods of interest, variation in *dsr* leads to biased estimates of $\hat{S}$.

Exposure period—the total time the subjects were at risk. Usually measured in days, but may be any convenient unit of time. Nesting and survival studies relying on the Mayfield method generally measure exposure in days.

Mayfield calculated *exposure* as the number of nest days during the period of study. *Exposure* is simply the number of nests under observation multiplied by the number of days at risk. This presents no problem when nests survive the observation interval. However, the periods between nest-visits are often longer than one day, and when nests or animals are destroyed during such an interval, the exact date of nest destruction is not known. Mayfield used the midpoint of the interval. For example, if there were 4 days in the interval he assumed that the nest survived 2 days and thus 2 days were added to the calculation of *exposure*.

a. Example Mayfield's Method

Table 1. Survival histories and exposure via the Mayfield method of three hypothetical nests (1 – indicates and active nest on the visit date; 0 – indicates a nest that was destroyed during the previous interval; blank – no data).				
Nest No.	1 May	8 May	15 May	exposure (days)
1	1	1	1	14 (2*7)
2	1	0		3.5 (0.5*7)
3	1	1	0	10.5 (1*7+0.5*7)
TOTAL				28

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$\hat{dsr} = 1 - \left(\frac{d}{exposure} \right)$$
$$= 1 - \left(\frac{2}{28} \right)$$
$$= 0.9286$$

$$\hat{S} = \hat{dsr}^{14}$$
$$= (0.9286)^{14}$$
$$= 0.3543$$

b. Variance estimation and Confidence Intervals

Because Johnson (1979) was able to demonstrate that Mayfield's estimator was a MLE he was able to derive a simple estimate of the variance from the 2nd derivative of the log-likelihood function.

$$\text{var}(\hat{dsr}) = \frac{(exposure - d) \times d}{(exposure)^3}$$

Furthermore, 95% CI's could be estimated as:

$$\hat{dsr} \pm 1.96 \left(\sqrt{\text{var}(\hat{dsr})} \right)$$

and the 95% CI's for $\hat{S}$ could be estimated by raising the corresponding $\hat{dsr}$ to the appropriate power of t.

Thus, referring back to Table 1:

$$\text{var}(\hat{dsr}) = \frac{(28-2) \times 2}{(28)^3},$$
$$= 0.0023688$$

$$\hat{dsr} \pm 1.96 \left(\sqrt{\text{var}(\hat{dsr})} \right) = 0.9286 \pm 1.96(0.04867),$$
$$= (0.833, 1.024)$$

If these are duck nests with an average exposure period of 34 days then

$$\hat{S} = \hat{dsr}^{34}$$
$$= 0.9286^{34}$$
$$= 0.0806$$

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

and 95% CI's for $\hat{S}$ are

$$LCL = 0.833^{34}$$
$$= 0.002$$

$$UCL = 1.024^{34}$$
$$= 2.240$$

3. Example of Mayfield estimator compared to apparent hatch rates (# successful/# found) — Miller and Johnson (1978)

Table 3. (Adapted from Miller and Johnson (1978) Table 1) Observed hatch rates, Mayfield and Mayfield—40% estimates of success for blue-winged teal nests found in North and South Dakota, 1967-72.

Stage		Days to hatch	Number of nests		Hatch rate	Mayfield			Mayfield - 40		
			Hatched	Destroyed		Exposure	dsr	$\hat{S}$	Exposure	dsr	$\hat{S}$
Laying	1	34	9	29	0.24	799	0.96		700.4	0.96	
	2	33	9	23	0.28	676.5	0.97		600.6	0.96	
	3	32	9	29	0.24	752	0.96		659.2	0.96	
	4	31	9	37	0.20	852.5	0.96		737.8	0.95	
	5	30	27	57	0.32	1665	0.97		1494.0	0.96	
	6	29	23	75	0.23	1754.5	0.96		1537.0	0.95	
	7	28	29	79	0.27	1918	0.96		1696.8	0.95	
	8	27	26	93	0.22	1957.5	0.95		1706.4	0.95	
	9	26	34	76	0.31	1872	0.96		1674.4	0.95	
	10	25	40	87	0.32	2087.5	0.96		1870.0	0.95	
	11	24	53	81	0.40	2244	0.96		2049.6	0.96	
Incubating	4	20	120	166	0.42	4060	0.96		3728.0	0.96	
	8	16	128	144	0.47	3200	0.96		2969.6	0.95	
	12	12	57	33	0.63	882	0.96		842.4	0.96	
	16	8	93	24	0.79	840	0.97		820.8	0.97	
	20	4	111	27	0.80	498	0.95		487.2	0.94	
	22	2	28	2	0.93	58	0.97		57.6	0.97	
	Pipped	1	19	1	0.95	19.5	0.95		19.4	0.95	
Hatched		0	13	0	1.00						
TOTAL			837	1063	0.44	26136	0.96	0.23	23651.2	0.96	0.20

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop
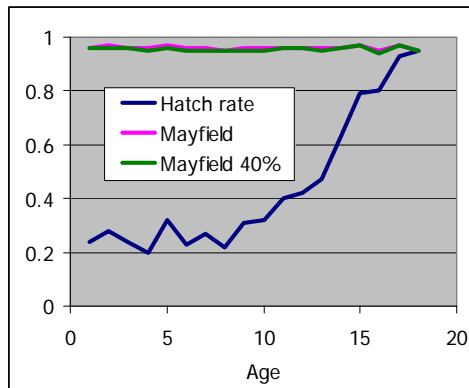
4. When Mayfield wins out – (Schaffer and Johnson 1990)

    In 1990, Shaffer and Johnson published the results of an analysis for simulated Mallard nests that examined when success estimates based on Mayfield's method were more accurate than apparent success and vice-versa. They determined that apparent success was accurate under certain conditions (below), and Mayfield's method performed poorly under conditions when nest destruction was catastrophic (e.g., floods or other catastrophic weather events).

    Conditions under which Apparent Success of nests performs well:

    - Detection rates of nests are very high.
      For example, colonial nesting or island nesting species.
    - Nest failures are catastrophic (i.e., *dsr* are not constant).
    - Nest searches are frequent (detection rates must still be high).
      Frequency of searches must increase with increasing asynchrony of nest initiation.
    - Nest success is very high.

5. Mayfield MLE

    The value of estimating the MLE of *dsr* is that this approach does not invoke Mayfield's midpoint assumption about the interval during which mortality occurs. Remember, in general, fewer assumptions is better.

    Recall the **Binomial likelihood** where y_i are the outcomes (success = 1 or failure = 0), n is the number of trials and p is the probability of success (model). The likelihood of p given y and n is the product of the probability of the individual trials:

    $$\mathcal{L}(p \mid y, n) = \prod_{i=1}^{n} \left((p)^{y_i} (1-p)^{(1-y_i)} \right)$$

    Now combine this with the concept of daily survival rate (*dsr*), recognizing that the probability of surviving a period of *t* days is just dsr^t. Substituting *dsr* for the model (p), fate (f, survived = 1 and died = 0) for success (y), and adding *ti* the length of the n exposure intervals. We get the Mayfield MLE,

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$\mathcal{L}(dsr \mid f,n,t) = \prod_{i=1}^{n}\left(\left(dsr^{t_i}\right)^{f_i}\left(1-dsr^{t_i}\right)^{(1-f_i)}\right),$$

as long as the dsr are constant. Note that the exposure intervals are the functional sampling units. Taking the log of the equation yields:

$$\ln\left(\mathcal{L}(dsr \mid f,n,t)\right) = \sum_{1}^{n} f_i t_i \ln(dsr^{t_i}) + \sum_{1}^{n}(1-f_i)\ln(1-dsr^{t_i}).$$

6.   Dinsmore's model – the nesting model in MARK

Dinsmore (2002) further generalized the Mayfield MLE to allow dsr to vary within the observation intervals and employed the use of the logit link function.

He started by writing the probability function for an individual nest, given that a new may be observed several times during the course of a study. It may also survive for some time and then fail during the period between observations. If we stick with the notation above, the probability of a nest surviving a period between visits ti days apart can be written as:

$$dsr^{t_i} = \left(\prod_{j=1}^{t_i} dsr\right).$$

Likewise, the probability of **not** surviving a period between visits ti days apart can be written as:

$$\left(1-dsr^{t_i}\right) = \left(1-\prod_{j=1}^{t_i} dsr\right)$$

Now, what if the dsr differ? We could write the probability of a nest surviving a period between visits ti days where dsr was different for each nest as:

$$dsr_i^{t_i} = \left(\prod_{j=1}^{t_i} dsr_i\right).$$

Likewise, the probability of **not** surviving a period between visits t_i days for these nests with different dsr can be written as:

$$\left(1-dsr_i^{t_i}\right) = \left(1-\prod_{j=1}^{t_i} dsr_i\right)$$

Now think about the observations on this nest written as a sort of capture history consisting of the time (date) the nest was found (k), the time when the nest was last known to be surviving (l), the last or terminal time of the nest (m), and the ultimate fate of the nest (f, where 1 is failure, 0 is not failure). We know the nest survived the period from $k - l$, so the probability of this occurring is:

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$\left(\prod_{i=k}^{l-1} dsr_i\right)$$

The probability of failure during $l - m$, is:

$$\left(1 - \prod_{i=l}^{m-1} dsr_i\right)^f,$$

and the probability of surviving $l - m$, is:

$$\left(\prod_{i=l}^{m-1} dsr_i\right)^{(1-f)}$$

Dinsmore developed the likelihood by combining those three things:

$$\mathcal{L}(dsr_{ij} \mid n, f_j, k_j, l_j, m_j) = \prod_{i=1}^{n}\left(\prod_{j=1}^{k_j-1} dsr_{ij}\right)\left(1 - \prod_{j=l_j}^{m_j-1} dsr_{ij}\right)^{f_j}\left(\prod_{j=l_j}^{m_j-1} dsr_i\right)^{(1-f_j)}$$

So, for example a nest found on day 1, alive and visited again on day 7, alive and visited again on day 10, and found destroyed on day 14. We would prepare the capture history:

1 10 14 0 1;

and the likelihood of this observation is:

$$\mathcal{L}(dsr_i \mid f_i, k_i, l_i, m_i) = \left(\prod_{i=1}^{9} dsr_i\right)\left(1 - \prod_{i=10}^{(14-10)} dsr_i\right)^{(1-0)}\left(\prod_{i=10}^{(14-10)} dsr_i\right)^{(0)}$$

It should be pretty obvious that we can now insert a link function (e.g. logit) in place of dsr_i

$$\mathcal{L}(dsr_i \mid f_i, k_i, l_i, m_i) = \left(\prod_{i=1}^{9} \frac{e^{\mathbf{X_i}\beta}}{\left(1 - e^{\mathbf{X_i}\beta}\right)}\right)\left(1 - \prod_{i=10}^{(14-10)} \frac{e^{\mathbf{X_i}\beta}}{\left(1 - e^{\mathbf{X_i}\beta}\right)}\right)^{(1-0)}\left(\prod_{i=10}^{(14-10)} \frac{e^{\mathbf{X_i}\beta}}{\left(1 - e^{\mathbf{X_i}\beta}\right)}\right)^{(0)}$$

As you already know, the data (**X**), can include covariates that vary among groups, individuals, or within individuals over time, and they may be continuous covariates (e.g., size, weight, or elevation), or discrete (such as habitat type).

7. Other link functions

There are 6 commonly used link functions. Logit and the first 3 listed below constrain estimates to the interval [0,1]. The default in MARK is the sine function. The last two link functions do not constrain the probability to the interval [0, 1], so occasionally cause numerical problems when

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

optimizing the likelihood. Other more complex link functions are beyond the scope of this class – see the program MARK documentation.

    a. Sine

$$\hat{S} = \frac{(\sin(X\beta)+1)}{2}$$

    b. Loglog

$$\hat{S} = e^{-e^{X\beta}}$$

    c. Complimentary loglog (Cloglog)

$$\hat{S} = 1 - e^{-e^{X\beta}}$$

    d. Log

$$\hat{S} = e^{X\beta}$$

    e. Identity

$$\hat{S} = X\beta$$