

MARKOV CHAIN MODELS OF FECUNDITY IN SHEEP

D. S. Robson and Luis Iniguez

BU-801-M

January, 1983

Abstract

Estrous cycling and fertility in sheep are influenced by the seasonal changes in photoperiod and temperature. Ignoring age effects and assuming continual presence of males, we have modeled the estrus-to-estrus transitions of a ewe as a Markov chain on a cyclical state space consisting of the 365 calendar dates. Additionally, we distinguish between "success" and "failure" of an estrus; a success results in a pregnancy and consequently an extended waiting period until the next estrus, while a failure results in resumption of estrus cycling. In either case the waiting time (in days) to her next heat is governed by a probability distribution unique to that season of the year. Parameters of this model thus consist of 365 "success" probabilities and $2 \times 365 = 730$ waiting time distributions. Derived distributions of primary interest are then "conception recurrence time" distributions and conditional distributions of the number of estruses required to achieve conception. Algorithms are given for calculating such distributions and, more simply, for calculating their means. Steady state mixtures of these distributions are also given.

* In the Biometrics Unit series, Cornell University, Ithaca, NY 14853.

MARKOV CHAIN MODELS OF FECUNDITY IN SHEEP

D. S. Robson and Luis Iniguez

BU-801-M

January, 1983

Introduction

Fecundity of an idealized, ageless adult ewe may be viewed as a circular random walk in a clockwise direction through the seasons of the year. The steps of this walk, of variable length measured in days, take the ewe from one estrus to the next. Conception normally results in a long step while a failure to conceive will result in a shorter step to the next heat, though the length of either step may be dependent upon the season of the year and may be subject to chance variation. We assume that the seasonal effects upon estrus cycling are photoperiodic and hence constant from year to year, while chance variations will be assumed Markovian; i.e., the previous steps leading to estrus on a particular day of the year have no bearing on the fecundity of the ewe from that day onward. We thus propose to model fecundity in sheep as a finite Markov chain on a two-dimensional state-space, one dimension representing the day of the year on which estrus occurs and the other dimension representing success or failure in the attempt to conceive.

Stochastic Model for Estrus-to-Estrus Transitions

Assuming that chances for success may vary with the seasonal changes in photoperiod and temperature, let p_i denote this conditional probability of conception for a ewe in heat on day i ; for $i = 1, 2, \dots, 365$, let

$$p_i = P(\text{conception} | \text{estrus on day } i) = 1 - q_i \quad (1)$$

In the event that a failure occurs on day i , then a_{ij} will denote the

conditional probability that her next estrus will occur j^* days later, while if a success does occur then b_{ij^*} will denote the conditional probability that her next heat will occur j^* days later. The expected waiting time until the next estrus is then

$$\alpha_i = \sum_{j^*} j^* a_{ij^*} \quad (2)$$

in the event of failure and

$$\beta_i = \sum_{j^*} j^* b_{ij^*} \quad (3)$$

if conception occurs. On average, the waiting time from estrus on year-day i to the next estrus is therefore

$$M_i = q_i \alpha_i + p_i \beta_i \quad (4)$$

(Step lengths j^* may be allowed to exceed 365 days with positive probability, but thereby introduce a notational inconvenience. In the sequel we assume that both a_{ij^*} and b_{ij^*} are zero for $j^* \geq 365$.)

Conception Recurrence Time

Let ξ_i denote the expected waiting time to conception following a conception failure on day i , and let η_i denote the expected waiting time to next conception following a conception success on day i . The average waiting time ω_i to next conception following a conception attempt on day i is, as in (4),

$$\omega_i = q_i \xi_i + p_i \eta_i$$

or in matrix notation with D_p denoting the diagonal matrix with $d_{ii} = p_i$,

$$\underline{\omega} = D_q \underline{\xi} + D_p \underline{\eta} \quad (5)$$

where $D_q = I - D_p$. The components $\underline{\xi}$ and $\underline{\eta}$ of $\underline{\omega}$ are seen to be determined by the linear relations

$$\xi_i = \alpha_i + \sum_j a_{ij} q^{(i+j)} \xi_{(i+j)} \quad \eta_i = \beta_i + \sum_j b_{ij} q^{(i+j)} \xi_{(i+j)} \quad (6)$$

where the subscript $(i+j)$ is cyclically reduced by the remainder formula

$$(i+j) = i+j - 365 \left[\frac{i+j-1}{365} \right]$$

where $[\cdot]$ denotes "integer part of." With this understanding we may use (6) to reexpress (5) as

$$\underline{\omega} = \underline{M} + P D_q \underline{\xi} \quad (7)$$

where P is the square matrix of transition probabilities

$$p_{i,(i+j)} \stackrel{\text{def'n}}{=} q_1 a_{ij} + p_i b_{ij} \quad (8)$$

between year-days of estrus; i.e., for a ewe in heat on year-day i , p_{ik} is the probability that her next estrus will occur on year-day k .

Systems of linear equations analogous to (5) - (7) may also be obtained for higher moments of waiting time to conception, but since these waiting time distributions are quite skewed and somewhat irregular in possessing multiple modes, higher moments would be virtually uninterpretable. The cumulative probability distributions (cdf's) are themselves calculable by more laborious, recursive methods, however, and a graph of the annual cycle in median and quartile waiting times, as well as the daily mean waiting time to conception, would be more readily comprehended and utilized than the higher moments.

The conditional distribution function $G_i(w)$ of waiting time to next conception after an estrus on day i ,

$$G_i(w) = P(\text{conception occurs during the next } w \text{ days after day } i | \text{estrus at } i) ,$$

may be expressed as a mixture of the two conditional cdf's

$$G_i(w) = q_i F_i(w) + p_i H_i(w) \quad (9)$$

where $F_i(w)$ is conditional on a failure to conceive during the estrus on day i and $H_i(w)$, a cumulative distribution of conception recurrence time, is conditional upon conception on day i . The corresponding probability mass functions

$$\begin{aligned} f_i(w) &= F_i(w) - F_i(w-1) = f_i^*(w) p_{(i+w)} \\ h_i(w) &= H_i(w) - H_i(w-1) = h_i^*(w) p_{(i+w)} \end{aligned} \quad (10)$$

may be calculated recursively, either independently for each i as

$$\begin{aligned} f_i^*(w) &= a_{iw} + \sum_{j=1}^{w-1} f_i^*(w-j) q_{(i+w-j)} a_{(i+w-j),j} \\ h_i^*(w) &= b_{iw} + \sum_{j=1}^{w-1} h_i^*(w-j) q_{(i+w-j)} a_{(i+w-j),j} \end{aligned} \quad (11a)$$

or jointly as

$$\begin{aligned} f_i^*(w) &= \sum_j a_{ij} q_{(i+j)} f_{(i+j)}^*(w-j) \\ h_i^*(w) &= \sum_j b_{ij} q_{(i+j)} f_{(i+j)}^*(w-j) \end{aligned} \quad (11b)$$

starting from the initial conditions $f_i^*(1) = a_{i1}$ and $h_i^*(1) = b_{i1}$.

Since the conditional frequency distributions a_{ij} of j = estrus cycle length and the frequency distributions b_{ij} of j = length of the period from

conception to next heat are likely to concentrate their mass within relatively short intervals around their respective means, α_i and β_i , most terms on the right-hand side of (11) will vanish. A computing algorithm for (11) could exploit this feature to greatly reduce the number of arithmetic operations required in calculating $F_i(w)$ and $H_i(w)$. Note also that since these computations proceed sequentially for $w=1,2,3$, etc., the computing process may be terminated when the highest desired percentiles of F_i and H_i are reached. Continued calculations beyond these percentiles are not needed in order to calculate the means ξ_i and η_i since these may be obtained by the linear methods (6).

Steady State Results

In a large flock of ageless ewes the proportion of estruses occurring on year-day i would converge to the equilibrium proportion u_i defined by the eigenvector $\underline{u}' = \underline{u}'P$, obtained as the solution of this system of linear equations supplemented with the linear constraint that \underline{u} sums to unity. For a flock in steady state the proportion of conceptions occurring on day i is then $c_i = u_i p_i / \underline{u}'\underline{p}$, and the frequency distribution of conception recurrence times in the entire flock is then $h(w) = \underline{c}'\underline{h}(w)$ with mean $\eta = \underline{c}'\underline{\eta}$. Note that this grand mean recurrence time of conception may also be calculated by the formula

$$\eta = \frac{\underline{u}'\underline{M}}{\underline{u}'\underline{p}} \quad (12)$$

since, from (5) and (7)

$$\underline{u}'\underline{\omega} = \underline{u}'\underline{D}_q \underline{\xi} + \underline{u}'\underline{D}_p \underline{\eta} = \underline{u}'\underline{M} + \underline{u}'\underline{P}\underline{D}_q \underline{\xi} = \underline{u}'\underline{M} + \underline{u}'\underline{D}_q \underline{\xi}$$

implying

$$\underline{u}'\underline{D}_p \underline{\eta} = \underline{u}'\underline{M} .$$

The expected number of attempts required to achieve conception may also be calculated from the equilibrium distribution u . For a ewe conceiving on day i let $k_{ij}(n)$ denote the conditional probability that her n 'th subsequent estrus occurs on day j . The conditional probability of conception at this n 'th estrus is then

$$P(\text{conception at } n\text{'th estrus} \mid \text{initial conception on day } i) = \sum_j k_{ij}(n) p_j$$

and the corresponding unconditional probability π_n is

$$P(\text{conception at } n\text{'th estrus} \mid \text{conception at initial estrus}) \stackrel{\text{def'n}}{=} \pi_n$$

$$= \sum_i c_i \sum_j k_{ij}(n) p_j .$$

Letting n get large and noting that $k_{ij}(n)$ is the n -step transition probability $p_{ij}^{(n)} \rightarrow u_j$, we find

$$\lim_{n \rightarrow \infty} \pi_n = \sum_i c_i \sum_j u_j p_j = \tilde{u}' \tilde{p} .$$

Applying the ergodic theorem of Kolmogorov (Theorem XIII.3 of Feller, Vol. 1) then gives $1/\tilde{u}'\tilde{p}$ as the mean number of estruses required to achieve conception.

Number of Estruses Required to Achieve the Next Conception

The probability distribution and the mean number of estruses required to achieve the next conception may be calculated by somewhat simpler procedures than those used in obtaining waiting time distributions and means. The greater simplicity is a consequence of the reduction to equal-sized steps from one estrus to the next that is achieved by simply counting steps rather than measuring their length. By analogy to (10) and (11), we define

the following conditional probability distributions for a ewe in heat on day i :

$$\begin{aligned}
 r_i(n) &= P\{\text{next conception at } n\text{'th subsequent estrus} \mid \text{failure at } i\} \\
 r_i^*(n) &= P\{\text{no conception before } n\text{'th subsequent estrus} \mid \text{failure at } i\} \\
 s_i(n) &= P\{\text{next conception at } n\text{'th subsequent estrus} \mid \text{conception at } i\} \\
 s_i^*(n) &= P\{\text{no conception before } n\text{'th subsequent estrus} \mid \text{conception at } i\}
 \end{aligned} \tag{13}$$

where now we have the simplifying relations with the cdf's:

$$\begin{aligned}
 R_i(n) &= \sum_{k=1}^n r_i(k) = 1 - r_i^*(n+1) \\
 S_i(n) &= \sum_{k=1}^n s_i(k) = 1 - s_i^*(n+1) .
 \end{aligned} \tag{14}$$

Only the starred functions therefore need to be computed, and they are seen to satisfy the initial conditions $r_i^*(1) = s_i^*(1) = 1$ and the recursions:

$$r_i^*(n+1) = \sum_j a_{ij} q_{(i+j)} r_{(i+j)}^*(n), \quad s_i^*(n+1) = \sum_j b_{ij} q_{(i+j)} r_{(i+j)}^*(n) . \tag{15}$$

Letting ρ_i and σ_i denote the means of the distributions $R_i(\cdot)$ and $S_i(\cdot)$, respectively,

$$\rho_i = 1 + \sum_{n=1}^{\infty} \{1 - R_i(n)\} \quad \sigma_i = 1 + \sum_{n=1}^{\infty} \{1 - S_i(n)\}$$

and applying (14) to (15) we obtain the linear equations

$$\rho_i = 1 + \sum_j a_{ij} q_{(i+j)} \rho_{(i+j)}, \quad \sigma_i = 1 + \sum_j b_{ij} q_{(i+j)} \rho_{(i+j)} .$$

(From the previous section we note that \underline{g} will also satisfy the equation

$$\underline{u}' \underline{D}_p \underline{g} = 1 .)$$

Numerical Example

The following numerical example utilizes a three-day year instead of a 365-day year, and has the range of $b_{i,j}$ exceeding one "year" in order to illustrate the fact that the earlier restriction $j < 365$ was not critical.

$$\begin{array}{c}
 A = \begin{array}{c|ccc}
 & j & 1 & 2 & 3 \\
 i & & & & \\
 \hline
 1 & & 0 & \frac{2}{3} & \frac{1}{3} \\
 2 & & 0 & 1 & 0 \\
 3 & & 0 & 1 & 0
 \end{array}
 \quad
 B = \begin{array}{c|cccc}
 & j & 1 & 2 & 3 & 4 \\
 i & & & & & \\
 \hline
 1 & & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
 2 & & 0 & 0 & 1 & 0 \\
 3 & & 0 & 0 & 1 & 0
 \end{array}
 \quad
 D_p = \begin{array}{c|ccc}
 & j & 1 & 2 & 3 \\
 i & & & & \\
 \hline
 1 & & .9 & 0 & 0 \\
 2 & & 0 & .8 & 0 \\
 3 & & 0 & 0 & .7
 \end{array}
 \end{array}$$

Applying (2) - (4):

$$\begin{array}{c|ccc}
 i & \alpha_i & \beta_i & M_i \\
 \hline
 1 & 7/3 & 3.5 & 10.15/3 \\
 2 & 2 & 3 & 2.8 \\
 3 & 2 & 3 & 2.7
 \end{array}$$

Equations (6) become:

$$\begin{aligned}
 \xi_1 &= \frac{7}{3} + \frac{2}{3}(.3\xi_3) + \frac{1}{3}(.1\xi_1) = \frac{8.44}{2.888} & \eta_1 &= 3.5 + \frac{1}{2}(.1\xi_1) + \frac{1}{2}(.2\xi_2) = \frac{11.192}{2.888} \\
 \xi_2 &= 2 + .1\xi_1 = 6.62/2.888 & \eta_2 &= 3 + .2\xi_2 = 9.988/2.888 \\
 \xi_3 &= 2 + .2\xi_2 = 7.1/2.888 & \eta_3 &= 3 + .3\xi_3 = 10.794/2.888
 \end{aligned}$$

and the transition probability matrix P is:

$$P = \begin{bmatrix} \frac{1}{3}(.1) + \frac{1}{2}(.9) & \frac{1}{2}(.9) & \frac{2}{3}(.1) \\ .2 & .8 & 0 \\ 0 & .3 & .7 \end{bmatrix} \Rightarrow \tilde{u} = \begin{bmatrix} 180/685 \\ 465/685 \\ 40/685 \end{bmatrix}$$

where $\tilde{u} = P'\tilde{u}$.

Applying (10) and (11) gives:

w	$f_1^*(w)$	$F_1(w)$	$f_2^*(w)$	$F_2(w)$	$f_3^*(w)$	$F_3(w)$
1	0	0	0	0	0	0
2	$\frac{2}{3}$.4 $\bar{6}$	1	.9	1	.8
3	$\frac{1}{3}$.7 $\bar{6}$	0	.9	0	.8
4	.2	.92 $\bar{6}$.2/3	.94 $\bar{6}$.2	.98
5	.2/9	.94 $\bar{2}$.1/3	.97 $\bar{6}$	0	.98
6	.46/9	.988 $\bar{2}$.02	.992 $\bar{6}$.04/3	.989 $\bar{3}$
7	.02/3	.993 $\bar{5}$.000 $\bar{2}$.994 $\bar{2}$.1/3	.995 $\bar{3}$
8	.092/27	.9959 $\bar{407}$.005 $\bar{1}$.9988 $\bar{2}$.004	.9985 $\bar{3}$
9	.082/27	.99867 $\bar{407}$.000 $\bar{6}$.9993 $\bar{5}$.004/9	.9988 $\bar{4}$
10	.0092/9	.99949 $\bar{185}$.0010 $\bar{2}/3$.99959 $\bar{407}$.0092/	.99976 $\bar{4}$

with mean waiting times as calculated earlier,

$$\xi_1 = 2.9224\bar{376} \quad \xi_2 = 2.29224\bar{376} \quad \xi_3 = 2.45844875 .$$

An example of (11b) is:

$$f_1^*(6) = a_{12}q_3f_3^*(4) + a_{13}q_1f_1^*(3) = \left(\frac{2}{3}\right)(.3)(.2) + \left(\frac{1}{3}\right)(.1)\left(\frac{1}{3}\right) = .46/9$$

and

$$\begin{aligned} h_1(6) &= [b_{13}q_{(1+3)}f_{(1+3)}^*(6-3) + b_{14}q_{(1+4)}f_{(1+4)}^*(6-4)]p_{(1+6)} \\ &= \left[\left(\frac{1}{2}\right)(.1)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)(.2)(1)\right](.9) = .105 . \end{aligned}$$