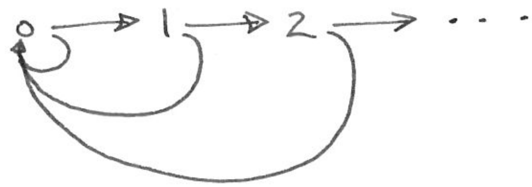
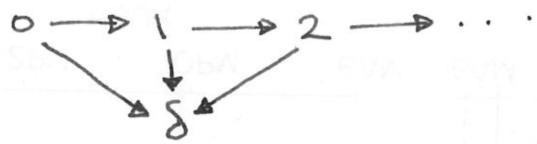


Progressive Multistage Processes



Brian S. Yandell
Biostatistics Group, UCB

Fall 1979
Stat 252 Stochastic Processes
A.J. Thomasian

Introduction

This report investigates some properties of a particular class of stochastic processes which I will call progressive multistage models. An individual subject to such a process can progress from early to late life stages, with the possibility of full recovery (Fig 1a) or death (Fig 1b).

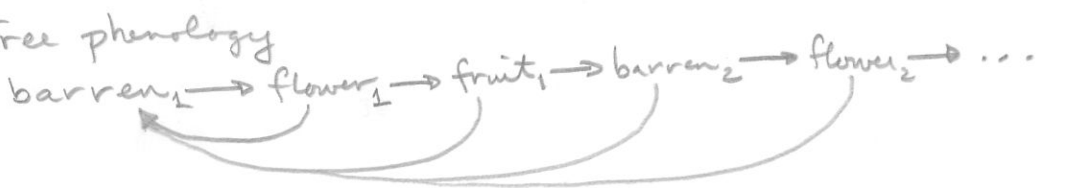
Fig 1. Progressive multistage processes

a) with recovery b) with death

My interest in these processes stems from contact with research problems in several areas of ecology and public health. For example, some people are interested in studying the progress of skin cancer and breast cancer in humans, contraction of diseases in laboratory mice, or the spacing of children. Others study the foraging behavior of birds, ontogenetic growth patterns in animals or plants, or development of insects through larval stages. I wish to concentrate on two problems in plant ecology in which I have participated.

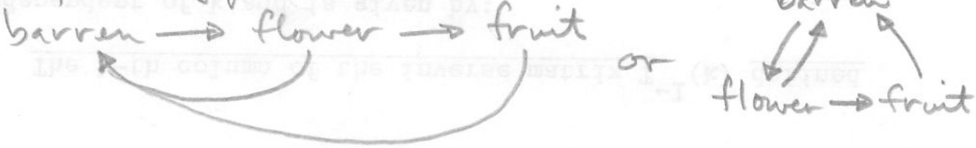
The first example is the flowering phenology, or timing, of a tree. In the tropics, some trees may flower once, several times, or not at all during a year. Once a tree flowers, it may set fruit, or it may lose its flowers with no fruiting. This inspires the following model:

Fig 2a) Tree phenology



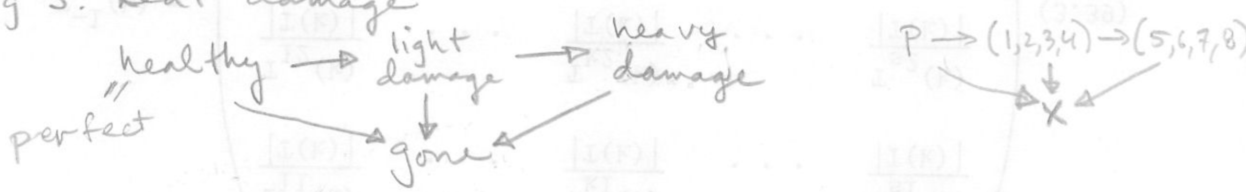
Here I have assumed that a a tree must fruit before it can reach the state $barren_2$, corresponding to a tree which is not flowering but has successfully set fruit once, etc. A simpler, finite state model would assume no difference between $barren_1$ and $barren_2$; i.e. that a barren tree responds similarly to any other barren tree at a particular time, irregardless of flower/fruiting history. In terms of graph,

Fig 2b) Tree phenology



The second example is leaf damage. An individual leaf begins life in a healthy state. As time passes, it is subject to chance attacks by insects which may eat all or part of its surface. Young leaves seem more subject to damage than old ones, and insect population densities vary over the course of a year. The simplified graph for this process is given in figure 3.

Fig 3. Leaf damage



Several general areas of questions seem of interest for these processes. Frankie, Baker and Opler (1974) described seasonal variation and differences among 2 areas and 2 types of trees in phenological studies. One might also be interested in expected time between fruiting and the probability of fruiting as a function of time, as well as formal tests of difference between populations. Suzanne Koptur (pers. comm.) is looking at leaf damage on several species located in several areas. Similar questions about time patterns and differences between species and areas arise. However, I have found it difficult to translate these important ecological questions into well-posed mathematical/statistical problems. For the purposes of this report I will limit attention to certain mathematical properties of a stochastic processes, and in particular will only consider a single population subject to the same forces, or intensities, at the same "age", or time.

I will primarily investigate the following quantities:

1) transition probabilities P and intensities V
2) expected number of visits N
3) reachability of states H
4) mean first passage time M

The inspiration for these comes from Kemeny, Snell and Knapp (1976), with considerable augmentation in approach and notation following Chiang (1968, 1980). Some mention of estimation will be made, though that is not a principal concern here.

Real research problems seldom yield perfect information on the processes studies. I will consider two types of imperfection: state and time. The state of a process may not be known precisely. One may only have a ~~lab~~ test to indicate a condition rather than direct observation. In the leaf damage process the state is determined by quick visual observation in the field, allowing for misclassification. The "exact" time of transitions between states is seldom known; one usually must observe the process continually (child spacing is one exception). The two plant ecology examples have once-a-month monitoring schemes. Thus one only knows the time to 1-month precision.

I will begin with the "basic example" presented in Kemeny, Snell and Knapp (1976), an infinite state Markov chain. I will modify this example in several ways, introducing a death state and considering finite state chains. Then I will examine continuous time Markov processes, following the notation of Chiang (1968, 1980). I will attempt a nonstationary time approach as much as possible.

The reader will notice few direct references back to the aforementioned plant ecology examples. They served to motivate my interest in the following processes, but I have yet to come full circle back to the questions: what do I do with the data? how do I answer the ecological questions? These require more work and time than I have found thus far possible, though I wish eventually to do these some justice.

This is hardly a polished document. I have chosen to pursue breadth in directions rather than a depth in any one aspect. Several sections appear to be, and in fact are, incomplete.

Basic example

Description

The state space consists of $S = \{0, 1, 2, \ldots\}$, with transitions at discrete times $t \in T = \{0, 1, 2, \ldots\}$.

The transitions with positive probabilities are $(i-1)$ to i and i to 0. The transition probabilities are denoted

$$P_{ij} = P_r\{\text{state } j \text{ at time } t+1 \mid \text{state } i \text{ at time } t\}$$

$$= \begin{cases} p_{i+1} & \text{if } j = i+1 \quad j \geq 1 \\ q_{i+1} & \text{if } j = 0 \quad i \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad p_i + q_i = 1$$

The transition matrix is

$$P = \{P_{ij}\} = \begin{array}{c|l} i \backslash j & \;\; 0 \quad\; 1 \quad\; 2 \quad\; 3 \;\; \cdots \\ \hline \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{pmatrix} q_1 & p_1 & 0 & 0 & \cdots \\ q_2 & 0 & p_2 & 0 & \\ q_3 & 0 & 0 & p_3 & \\ \vdots & & & & \ddots \end{pmatrix} \end{array}$$

Normally we assume $0 < p_i < 1$ unless otherwise stated.

It will be convenient to have the row vector

$\beta = (\beta_0, \beta_1, \beta_2, \ldots)$, with

$$\beta_0 = 1 \quad \beta_i = \prod_{j \leq i} p_j = P_r\{i \text{ "successes" in a row} \mid \text{start at } 0\}$$
$$= P_r\{\text{State } i \text{ at time } t+i \mid \text{state } 0 \text{ at time } t\}$$

Denote $\beta_\infty = \lim_{i \to \infty} \beta_i$.

Here a success is the event $i-1 \to i$.

We may represent the process as a graph as follows:

$$\begin{tikzcd} 0 \arrow[r, "p_1"] \arrow[loop left, "q_1"] & 1 \arrow[r, "p_2"] \arrow[l, bend left, "q_2"] & 2 \arrow[r, "p_3"] \arrow[ll, bend left, "q_3"] & 3 \arrow[r] \arrow[lll, bend left] & \cdots \end{tikzcd}$$

Properties

The basic example process is stationary. It is recurrent or transient depending upon whether $\beta_\infty = 0$ or not. If it is recurrent then all states communicate, and the chain is in every state infinitely often with probability one. If it is transient then the chain will eventually leave any finite subset of S with probability one.

At this point some notation is needed. Let

$H_{ij} = \Pr\{\text{reach state } j \text{ eventually} \mid \text{start in state } i\}$

$H = \{H_{ij}\}$

$N_{ij} =$ expected no. of times in state j | start in state i

$N = \{N_{ij}\}$

For the recurrent process we have easily.

$$H_{ij} = 1 \qquad N_{ij} = \infty \qquad \forall i \forall j$$

The transient process has (since $\beta_\infty > 0$)

$$H_{ij} = \begin{cases} 1 & i \le j \\ 1 - \dfrac{\beta_\infty}{\beta_i} & i > j \end{cases} \qquad N_{ij} = \begin{cases} \dfrac{\beta_j}{\beta_\infty} & i \le j \\ \dfrac{\beta_j}{\beta_\infty} - \dfrac{\beta_j}{\beta_i} & i > j \end{cases}$$

(see Kemeny, Snell and Knapp 1976). Thus the process eventually reaches any state further along than a given state i, but with positive probability will not reach an earlier state if the process is transient.

Let M_{ij} = mean first passage time
= expected time to first passage into j if the process begins in i

This may be infinite, as will be the case for M_{ii} if the process in null recurrent. In other words, we may define

Basic example

the probability of being in state i, denoted α_i, by

$$\alpha_i = \begin{cases} \dfrac{\beta_i}{\sum_{j=0}^{\infty}\beta_j} & \text{if } \sum_{j=0}^{\infty}\beta_j < \infty \text{ i.e. nonnull recurrent} \\ 0 & \text{if } \sum_{j=0}^{\infty}\beta_j = \infty \text{ i.e. null recurrent} \end{cases}$$

Then, for the nonnull recurrent process, with $0 < p_i < 1$ $\forall i$ we have that the process is ergodic, (i.e. aperiodic nonnull recurrent), and

$$M_{ii} = \frac{1}{\alpha_i} = \frac{\sum \beta_j}{\beta_i}$$

We can derive also (see Kemeny Snell & Knapp 1976)

$$M_{0i} = \frac{1}{\beta_i}\sum_{k<i}\beta_k \qquad M_{i0} = \frac{1}{\beta_i}\sum_{k\geq i}\beta_k\ ;\ i>0$$

$$M_{ij} = \begin{cases} M_{0j} - M_{0i} & 0\leq i<j \\ M_{i0} + M_{0j} & i>j\geq 0 \end{cases}$$

For the null recurrent process we see that

$$M_{ij} = M_{ii} = \infty \text{ for } i>j\ ;\ M_{ij} < \infty \text{ for } i<j.$$

Thus in a recurrent process, the expected time to reach j from i is finite if $i<j$ always, and finite if $i \geq j$ only in the ergodic case.

The transient process would have the same formulae for M as does the null recurrent process.

It is interesting to note that if the process is ergodic, then $\alpha = (\alpha_0, \alpha_1, \ldots)$ is a stationary initial density. That is, $\alpha P = \alpha$. Also it is easily seen that $P^n \to \begin{pmatrix}\alpha\\ \alpha\\ \vdots\end{pmatrix}$ by noting that $H_{0i} = 1$ $\forall i \geq 0$ and $(P^n)_{ij} \to (HP^n)_{ij} \to (P^n)_{0j} \to \alpha_j$

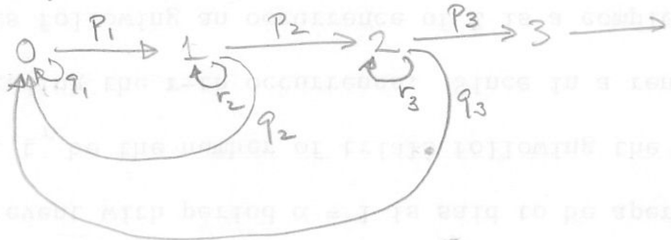
[Here the arrows represent a loose interpretation of tending to the limit, in which each arrow takes us further out the sequence of n's. This is hueristic and is not a proof.] Thus by definition of stability (Thomasian 1969), P is stable.

Basic example with pause

Take the basic example and allow the process to remain in state i with positive probability r_{i+1}. This yield transition matrix

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ \\ \\ \end{array}\begin{pmatrix} q_1 & p_1 & 0 & 0 & \cdots \\ q_2 & r_2 & p_2 & 0 & \\ q_3 & 0 & r_3 & p_3 & \\ \vdots & & & & \ddots \\ \vdots & & & & \end{pmatrix} \qquad \begin{array}{l} q_1 + p_1 = 1 \\ q_i + r_i + p_i = 1 \\ 0 < (q_i, r_i, p_i) < 1 \end{array}$$

The graph of the process looks like



Define $\bar{p}_i = p_i \sum_{k=0}^{\infty} r_i^k = \dfrac{p_i}{1-r_i}$; $\bar{q}_i = 1 - \bar{p}_i = \dfrac{q_i}{1-r_i}$; $i \geq 1$

Then $\bar{p}_i = \Pr\{\text{reach state } i \text{ without going through } 0 \mid \text{in state } i-1\}$

$\bar{q}_i = \Pr\{\text{reach state } 0 \text{ without reaching state } i \mid \text{in state } i-1\}$

If we call the move $i \to i$ a pause, then

$$\bar{m}_i = \sum_{k=1}^{\infty} r_i^k = \frac{r_i}{1-r_i} < \infty$$

is the expected no. of pauses in i for each visit to i.

Also define $\bar{\beta}_i = \prod_{j \leq i} \bar{p}_j$; $\bar{\beta}_0 = 1$; $\bar{\beta}_\infty = \lim_{i \to \infty} \bar{\beta}_i$; $\bar{\beta} = (\bar{\beta}_0, \bar{\beta}_1, \ldots)$

Then we have clearly that this process is recurrent or transient as $\bar{\beta}_\infty = 0$ or not, as before. Also H and N are the same as in the basic example except β is replaced by $\bar{\beta}$. I believe the same goes for α and M, but I have not checked this to my satisfaction. If so we see that this process is essentially the basic example process with the time axis distorted, though this distortion is itself stochastic.

Basic example with death

Consider the variation on the basic example process in which the process may move to the next state or be absorbed. In terms of a graph we have

$$\begin{tikzcd} 0 \arrow[r, "p_1"] \arrow[dr, "q_1"'] & 1 \arrow[r, "p_2"] \arrow[d, "q_2"] & 2 \arrow[r, "p_3"] \arrow[dl, "q_3"] & \cdots \\ & \delta \arrow[loop below] & & \end{tikzcd}$$

The transition matrix becomes

$$P = \begin{array}{c} \delta \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & 0 & \\ q_2 & 0 & 0 & p_2 & 0 & \\ q_3 & 0 & 0 & 0 & p_3 & \\ \vdots & & & & & \ddots \end{pmatrix} \quad \text{with } 0 < p_i < 1 \text{ by assumption.}$$

Define $\beta_i = \prod_{j \le i} p_j \qquad \beta_0 = 1$ as before.

Then we can deduce the reachability matrix as.

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 1-\beta_\infty & 1 & \beta_1 & \beta_2 & \beta_3 & \\ \left(1-\frac{\beta_\infty}{\beta_1}\right) & 0 & 1 & \beta_2/\beta_1 & \beta_3/\beta_1 & \\ \left(1-\frac{\beta_\infty}{\beta_2}\right) & 0 & 0 & 1 & \beta_3/\beta_2 & \\ \left(1-\frac{\beta_\infty}{\beta_3}\right) & 0 & 0 & 0 & 1 & \\ \vdots & & & & & \ddots \end{pmatrix}$$

and the expected number of visits

$$N = \begin{pmatrix} \infty & 0 & 0 & 0 & 0 & \cdots \\ \infty & 1 & \beta_1 & \beta_2 & \beta_3 & \\ \infty & 0 & 1 & \beta_2/\beta_1 & \beta_3/\beta_1 & \\ \infty & 0 & 0 & 1 & \beta_3/\beta_2 & \\ \infty & 0 & 0 & 0 & 1 & \\ \vdots & & & & & \ddots \end{pmatrix}$$

That is $N_{ij} = H_{ij}$ for $i,j \in \{0,1,2,\ldots\} = \mathcal{S} \setminus \{\delta\}$. Each transient state is visited at most once. The steady state probability of being in a state is given by

$$\alpha_\delta = 1 - \beta_\infty, \quad \alpha_i = 0 \quad i = 0,1,2,\ldots$$

Basic example with death

Thus there is probability β_∞ that the process will never reach δ but will continue on indefinitely.

The mean first passage matrix differs more markedly from that of the basic example than H or N did.

$$M_{ij} = \infty \qquad i \geq j$$

$$M_{0i} = i\beta_i \qquad i \geq 1$$

$$M_{ij} = M_{0j} - M_{0i} = j\beta_j - i\beta_i \qquad 0 < i < j$$

$$M_{\delta\delta} = 1$$

$$M_{0\delta} = \sum_{i=0}^{\infty} \beta_i$$

$$M_{i\delta} = M_{0\delta} - M_{0i} = \left(\sum_{j=0}^{\infty} \beta_j\right) - i\beta_i \qquad i \geq 1$$

$$M_{\delta i} = \infty \qquad i \geq 0$$

If we denote $\gamma = \sum_{j=0}^{\infty} \beta_j$, which may be ∞, then

$$M = \begin{array}{c} \delta \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ \\ \end{array} \begin{pmatrix} 1 & \infty & \infty & \infty & \infty & \cdots \\ \gamma & \infty & \beta_1 & 2\beta_2 & 3\beta_3 & \\ \gamma-\beta_1 & \infty & \infty & 2\beta_2-\beta_1 & 3\beta_3-\beta_1 & \\ \gamma-2\beta_2 & \infty & \infty & \infty & 3\beta_3-2\beta_2 & \\ \gamma-3\beta_3 & \infty & \infty & \infty & \infty & \\ \vdots & & & & & \ddots \\ \vdots & & & & & \end{pmatrix}$$
$$\begin{array}{ccccc} \delta & 0 & 1 & 2 & 3 \end{array}$$

I leave an interpretation of the relation between the properties of this process and the basic example process to the reader. Clearly δ looks a lot like 0 in the former, while the states 1, 2, 3, ... look different due to the fact that these are only visited once.

Finite Basic example #1

We reconsider the basic example but allow $p_i = 0$ for $i \geq n$. The process becomes

and the transition matrix is

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ n-2 \\ n-1 \end{array} \begin{pmatrix} q_1 & p_1 & 0 & 0 & \cdots & 0 \\ q_2 & 0 & p_2 & 0 & \cdots & 0 \\ q_3 & 0 & 0 & p_3 & & \\ \vdots & & & & \ddots & 0 \\ q_{n-1} & 0 & \cdots & & 0 & p_{n-1} \\ 1 & 0 & \cdots & & 0 & 0 \end{pmatrix}$$

For simplicity let $n=3$, then $P = \begin{pmatrix} q_1 & p_1 & 0 \\ q_2 & 0 & p_2 \\ 1 & 0 & 0 \end{pmatrix}$ $\quad 0 < p_i < 1 \quad i = 1,2$

The results generalize easily for all n.

The row vector becomes: $\beta = (1, p_1, p_1 p_2)$, and trivially $\beta_\infty = 0$

Clearly the process is recurrent and ergodic, hence

$$H_{ij} = 1 \qquad N_{ij} = \infty \qquad \forall i,j < n$$

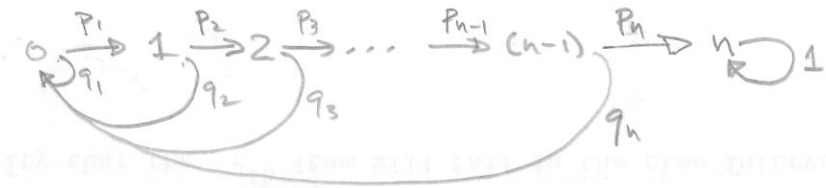
$$\alpha_i = \frac{\beta_i}{\sum_{j=0}^{n-1} \beta_j} \qquad \alpha = \left(\frac{4}{7}, \frac{2}{7}, \frac{1}{7}\right) \text{ if } p_1 = p_2 = \tfrac{1}{2}$$

and the mean first passage time is given by (for $n=3$)

$$M = \begin{pmatrix} 1 + p_1 + p_1 p_2 & \dfrac{1}{p_1} & \dfrac{1+p_1}{p_1 p_2} \\[2ex] 1 + p_2 & 1 + \dfrac{1 + p_1 p_2}{p_1} & \dfrac{1 + p_1 - p_2}{p_1 p_2} \\[2ex] 1 & 1 + \dfrac{1}{p_1} & 1 + \dfrac{1+p_1}{p_1 p_2} \end{pmatrix} = \begin{pmatrix} 1.75 & 2 & 6 \\ 1.5 & 3.5 & 4 \\ 1 & 3 & 7 \end{pmatrix} \text{ if } p_1 = p_2 = \tfrac{1}{2}$$

Finite Basic example #2

Another variation on the theme. Consider the process



In other words, state n is absorbing state, such as death following a possibly subterminal stage. The transition matrix looks like

$$P = \begin{array}{c} 0 \\ 1 \\ \\ n\text{-}2 \\ n\text{-}1 \\ n \end{array}\begin{pmatrix} q_1 & p_1 & 0 & \cdots & 0 & 0 \\ q_2 & 0 & p_2 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \\ q_{n-1} & 0 & 0 & \cdots & p_{n-1} & 0 \\ q_n & 0 & 0 & & 0 & p_n \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \qquad \beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_n, \beta_n)$$

$$\beta_i = \prod_{j \le i} p_j \qquad \beta_0 = 1$$

We may view this as the basic example with

$$p_i = 1, \ \beta_i = \beta_{n-1} \ \text{for } i > n; \ \beta_\infty = \beta_n = \prod_{j \le n} p_j$$

The process is clearly transient, and has the appropriate H and N matrices with H_{ij} N_{ij} defined as in the basic example for $0 \le i, j < n$. In addition

$$\left.\begin{array}{ll} H_{nj} = 0 & H_{jn} = 1 \\ N_{nj} = 0 & N_{jn} = \infty \end{array}\right\} \ 0 \le j < n; \quad H_{nn} = 1, \ N_{nn} = \infty$$

The steady state probability vector α is

$$\alpha = (0, 0, \ldots, 0, 1) \quad \text{i.e. } \alpha_i = \begin{cases} 0 & i < n \\ 1 & i = n \end{cases}$$

Therefore $M_{ii} = \left\{\begin{array}{ll} \infty & i < n \\ 1 & i = n \end{array}\right\}$

$$M_{0i} = \frac{1}{\beta_i} \sum_{k < i} \beta_k \qquad 0 < i \le n$$

$$M_{ij} = M_{0j} - M_{0i} \qquad 0 < i < j \le n$$

$$M_{ij} = \infty \qquad 0 < j < i \le n$$

Continuous time Markov process

The continuous time Markov process can be handled in a variety of ways. I choose to allow a very general nonstationary process for the specific model, and add restrictive assumptions only as needed for mathematical tractability in order to get results of a desired nature.

Let's consider a process with 3 states $\Delta = \{0, 1, 2\}$ represented by the graph, which can easily be generalized to any number of states.

$$\begin{tikzcd} 0 \arrow[r, "\nu_{01}"] & 1 \arrow[r, "\nu_{12}"] \arrow[l, bend left, "\nu_{10}"] & 2 \arrow[ll, bend left=60, "\nu_{20}"] \end{tikzcd}$$

Define the following probabilities

$$P_{ij}(t_1, t_2) = Pr\{\text{in state } j \text{ at time } t_2 \mid \text{in state } i \text{ at time } t_1\}; \quad i, j \in \Delta, \quad 0 \le t_1 < t_2$$

and intensities

$$\nu_{ij}(t) = \text{transition intensity at time } t; \ i \ne j.$$

We *assume* that over a short period of time, say $(t, t+dt)$ that the probability of transition $i \to j$, with $i \ne j$ is proportional to $\nu_{ij}(t)$. In other words

$$P_{ij}(t, t+dt) \doteq \nu_{ij}(t)\,dt; \quad i \ne j \quad 0 \le t$$

NOTE: '$\doteq$' is used to mean almost equal in the sense that the difference is $o(dt)$.

In order to complete the definition we need

$$\nu_{ii}(t) = -\sum_{j \ne i} \nu_{ij}(t)$$

and

$$P_{ii}(t, t+dt) \doteq 1 + \nu_{ii}(t)\,dt = 1 - \sum_{j \ne i} \nu_{ij}(t)\,dt.$$

Thus we have

$$1 = P_{ii}(t, t+dt) + \sum_{j \ne i} P_{ij}(t, t+dt) \qquad \forall i.$$

Now we can construct the instantaneous transition matrix for transitions during the instant $(t, t+dt)$. For the above process we have

$$\underline{V}(t) = \begin{pmatrix} 1 + \nu_{00}(t)dt & \nu_{01}(t)dt & 0 \\ \nu_{10}(t)dt & 1 + \nu_{11}(t)dt & \nu_{12}(t)dt \\ \nu_{20}(t)dt & 0 & 1 + \nu_{22}(t)dt \end{pmatrix}$$

One immediately notices the similarities with the basic example with pause in which the process could remain in the present state ($r_i \sim 1 + \nu_{ii}(t)\,dt$), move to the next state ($p_i \sim \nu_{i,i+1}(t)\,dt$) or revert back to the initial state ($q_i \sim \nu_{i0}(t)\,dt$). In that process we could determine the transition matrix for a given number n of steps by calculating P^n. Analogously, here we are interested in transitions over an interval of time and use the product integral (see Cox 1972; Aalen 1978).
[One can make an argument similar to the binomial $\rightarrow$ Poisson in the limit.]

$$\underline{P}(t_1, t_2) = \text{transition probability matrix over interval } [t_1, t_2)$$
$$= \prod_{t_1}^{t_2} \underline{V}(t)$$

Let $H_{i}(t_1, t_2) = P_r\{\text{remain in state } i \text{ during } (t_1, t_2) \mid \text{in state } i \text{ at } t_1\}$
$$= e^{\int_{t_1}^{t_2} \nu_{ii}(t)\,dt}.$$

Then we see that

$$\prod_{t_1}^{t_2}(1+\nu_{ii}(t)\,dt) = \exp \int_{t_1}^{t_2} \log(1+\nu_{ii}(t)\,dt) \doteq \exp \int_{t_1}^{t_2} \nu_{ii}(t)\,dt = H_i(t_1, t_2)$$

The first equality requires some regularity in $\nu_{ii}(\cdot)$ which will be assumed and ignored from now on. From this one may visualize a particular path. From here on we will restrict attention to the time interval $[0,1]$ and assume transitions between states occur only at times $0 < t_1 < t_2 < t_3 < \dots < t_k < 1$. A particular realization would be

state $\quad 0 \longrightarrow 1 \longrightarrow 0 \longrightarrow 1 \rightarrow 1$ –

time $\quad 0 \quad t_1 \quad t_2 \quad t_3 \quad 1$

In other words, the process is in state 0 until time t_1, and moves to state 1 during $(t_1, t_1 + dt_1)$, etc. Although much of the following [illegible]
[illegible]

II introduce some notation to describe this realization

For fixed $k > 0$ let.

$$\underset{\sim}{x} = (x_0, x_1, \dots, x_{k-1}) \qquad \underset{\sim}{t} = (t_0, t_1, \dots, t_k)$$

with $x_\lambda \in \mathcal{S}$, $0 \le t_0 < t_1 < \dots < t_k$.

Define

$$H_{\underset{\sim}{x}}(\underset{\sim}{t})\,d\underset{\sim}{t} = H_{x_0}(t_0, t_1) \prod_{\lambda=1}^{k-1} \left[\nu_{\lambda,\lambda+1}(t_\lambda)\,dt_\lambda\, H_{x_\lambda}(t_\lambda, t_{\lambda+1}) \right]$$

Then we see that

$$H_{(0,1,0,1)}(0, t_1, t_2, t_3, 1)\,dt_1\,dt_2\,dt_3 = P_r\{\text{realization depicted on previous page}\}$$

$$= H_0(0,t_1)\,\nu_{01}(t_1)\,dt_1\,H_1(t_1,t_2)\,\nu_{10}(t_2)\,dt_2\,H_0(t_2,t_3)\,\nu_{01}(t_3)\,dt_3\,H_1(t_3,1)$$

Further, the probability of a realization with the state order as given for this realization $\underset{\sim}{x} = (0,1,0,1)$ is.

$$\bar{H}_{\underset{\sim}{x}}(t_0, t_k) = \int H_{\underset{\sim}{x}}(\underset{\sim}{t})\,d\underset{\sim}{t}, \text{ and } \bar{H}_{\underset{\sim}{x}}(0,\infty) = \bar{H}_{\underset{\sim}{x}}.$$

For this example it would be

$$\bar{H}_{(0,1,0,1)}(0,1) = \int_0^1 \int_0^{t_3} \int_0^{t_2} H_{(0,1,0,1)}(0, t_1, t_2, t_3, 1)\,dt_1\,dt_2\,dt_3$$

Let $A_{\lambda j}$ = collection of all paths from λ to j over time $[0,1]$

$$= \{ \underset{\sim}{x} \mid (\underset{\sim}{x} = (x_0, x_1, \dots, x_{k-1}) \text{ for some } k, x_0 = \lambda, x_k = j, x_\ell \in \mathcal{S}\ 0 \le \ell \le k \}$$

Then one can see readily that $A_{\lambda j}$ is a countable set, and

$$P_{\lambda j}(s,t) = \sum_{\underset{\sim}{x} \in A_{\lambda j}} \bar{H}_{\underset{\sim}{x}}(s,t) \qquad 0 \le s < t$$

in which the sum is over distinct elements of $A_{\lambda j}$.

Therefore we see that although $P_{\lambda j}(0,1)$ is well defined the set $A_{\lambda j}$ will be infinite, and the calculation of $P_{\lambda j}(0,1)$ without further assumptions may be rather involved. I will now step back from this task to examine $H_{\underset{\sim}{x}}(\underset{\sim}{t})$ and $\bar{H}_{\underset{\sim}{x}}(s,t)$, assuming that the sequence of states $\underset{\sim}{x}$ and/or times $\underset{\sim}{t}$ are known. In order to motivate this, let's consider a slightly different process.

Let's return to the example of leaf damage. Suppose the leaf can be in one of 4 states, healthy (0) slight damage (1), heavy damage (2), or gone (δ). thus $\mathcal{S} = \{0, 1, 2, \delta\}$ and the graph of the process is

$$\begin{tikzcd} 0 \arrow[r, "\nu_{01}"] \arrow[dr, "\nu_{0\delta}"'] & 1 \arrow[r, "\nu_{12}"] \arrow[d, "\nu_{1\delta}"] & 2 \arrow[dl, "\nu_{2\delta}"] \\ & \delta & \end{tikzcd}$$

NOTE: traditionally (Chiang 1968) we set $\mu_1(t) = \nu_{1\delta}(t)$.

Let $A = \{(0), (0,1), (0,1,2), (0,\delta), (0,1,\delta), (0,1,2,\delta), (1), (1,2), (1,\delta), (1,2,\delta), (2), (2,\delta), (\delta)\}$

= collection of all possible paths

Thus A_{ij} is finite, and in fact contains at most one element unless $j = \delta$ in which case it may contain up to 3 elements. The introduction of the absorbing state δ greatly simplifies the "path space". Note in particular that

$$P_{11}(0,1) = \bar{H}_1(0,1) = e^{\int_0^1 \nu_{11}(t)\,dt}.$$

Chiang (1979) has treated this problem in great depth provided one accepts his proportionality assumption

$$\nu_{ij}(t) = \nu_{ij}\,\beta(t)\;;\quad \int_0^t \beta(x)\,dx \longrightarrow \infty \text{ as } t \to \infty$$

For Chiang's analysis, one needs to know the times of transition as well as the time of death. Short of his approach one is left with terms of the form

$$P_{00}(0,t_1)\,\nu_{01}(t_1)\,P_{11}(t_1,t_2)\,\nu_{12}(t_2)\,P_{22}(t_2,t_3)\,\mu_2(t_3)$$

which indicate that the identifiable quantities are

$$\mu_i(t),\ \nu_{i,i+1}(t) \text{ and } \int_s^t \nu_{ii}(x)\,dx.$$

Odd Aalen (1976, 1978) indicated a possible direction for nonparametric estimation of these quantities, though it is beyond the scope of this paper. One may also notice similarities with the general birth and death process (Chiang 1968, 1980).

Properties of continuous time Markov Process

Consider ~~the~~ leaf damage process.

$$\begin{tikzcd} 0 \arrow[r] \arrow[dr] & 1 \arrow[r] \arrow[d] & 2 \arrow[dl] \\ & \delta & \end{tikzcd}$$

We have clearly that reachability matrix is

$$H_{\lambda\pi} = \begin{array}{c} 0 \\ 1 \\ 2 \\ \delta \end{array} \begin{pmatrix} 1 & H^*(0,1) & H^*(0,2) & \bar{H}(0,\delta) \\ 0 & 1 & H^*(1,2) & \bar{H}(1,\delta) \\ 0 & 0 & 1 & \bar{H}(2,\delta) \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with $H^*(0,1) = \int_0^\infty P_{00}(0,t)\,\nu_{01}(t)\,dt$

whereas $\bar{H}_{(0,1)} = \int_0^\infty P_{00}(0,t)\,\nu_{01}(t)\,dt\;P_{11}(t,\infty)$,

and similarly for $H^*(0,2)$, $H^*(1,2)$.

Notice that $\bar{H}(\lambda,\gamma) \leq H^*(\lambda,\gamma)$

The number of times matrix is the same. That is,

$N = H$.

Here we count one entry into state i for each time the process enters i from another state. Since each state can only be entered once, the equality $N=H$ holds trivially. This definition differs from that for N in the discrete case.

The mean first passage time matrix is

$$M = \begin{pmatrix} \infty & M_{01} & M_{02} & M_{0\delta} \\ 0 & \infty & M_{12} & M_{1\delta} \\ 0 & 0 & \infty & M_{2\delta} \\ 0 & 0 & 0 & \infty \end{pmatrix}$$

with $M_{01} = \int_0^\infty t\,P_{00}(0,t)\,\nu_{01}(t)\,dt$, etc.

Let's assume that the intensities are of the form

$\nu_{ij}(t) = \nu_{ij}\beta(t)$, with $\alpha(t) = \int_0^t \beta(x)\,dx \longrightarrow \infty$ as $t \to \infty$.

(see Chiang 1979). Then we have

$$P_{ii}(t_1, t_2) = e^{\nu_{ii}\alpha(t)}$$

$$H^*(i, i+1) = \int_0^\infty e^{\nu_{ii}\alpha(t)} \nu_{i,i+1}\beta(t)\,dt = \frac{-\nu_{i,i+1}}{\nu_{ii}} \qquad i = 0, 1$$

$$H^*(0,2) = \int_0^\infty \int_{t_2}^\infty e^{\nu_{00}\alpha(t_1)} \nu_{01}\beta(t_1)\, e^{\nu_{11}(\alpha(t_2) - \alpha(t_1))} \nu_{12}\beta(t_2)\,dt_1\,dt_2$$

$$= \frac{\nu_{01}\nu_{12}}{\nu_{00}} \frac{1}{\nu_{00} - \nu_{11}}$$

$$\bar{H}(2,\delta) = \frac{\mu_2}{\nu_{22}}$$

And so on.

– In order to get M_{ij} in terms of intensities one might assume also $\beta(t) \equiv 1$. I will not elaborate further here.

Imperfect time information in continuous process

Suppose that we know a particular path $\underset{\sim}{x} = (x_0, x_1, \ldots, x_{k-1})$ occurs over the interval $[0,1]$, but we know only that the process is in x_0 at 0 and x_{k-1} at 1. We lack some information, and this inspires two questions: (1) how might we estimate the unknown times $t_1, t_2, \ldots, t_{k-1}$, and (2) what can we say about probabilities if we do not try to estimate these quantities?

From our earlier work we have

$$L(\underset{\sim}{t}) = \frac{H_{\underset{\sim}{x}}(\underset{\sim}{t})}{\bar{H}_{\underset{\sim}{x}}(0,1)} = \text{likelihood of times } \underset{\sim}{t} = (0, t_1, \ldots, t_{k-1}, 1)$$

provided $\bar{H}_{\underset{\sim}{x}}(0,1) > 0$. Therefore a natural procedure would be to use maximum likelihood estimation. However, before doing this let us note that

$$\bar{H}_{\underset{\sim}{x}}(0,1) = \Pr\{\text{realize path } \underset{\sim}{x} \text{ over interval } [0,1]\}$$

takes care of the second question above, and cannot, in my understanding, be improved upon.

Let's examine the simple path $\underset{\sim}{x} = (0,1)$. Then in order to maximize the likelihood we need only maximize $H_{\underset{\sim}{x}}(\underset{\sim}{t})$ with $\underset{\sim}{t} = (0, t, 1)$, $0 \le t \le 1$. Assume for the present that we know $\nu_{00}(\cdot), \nu_{01}(\cdot), \nu_{11}(\cdot)$. Then we have

$$H_{\underset{\sim}{x}}(\underset{\sim}{t}) = e^{\int_0^t \nu_{00}(x)dx} \nu_{01}(t) e^{\int_t^1 \nu_{11}(x)dx}$$

$$\frac{\partial \log H}{\partial t} = H_{\underset{\sim}{x}}(\underset{\sim}{t}) \left[(\nu_{00}(t) - \nu_{11}(t)) + \frac{\nu_{01}'(t)}{\nu_{01}(t)} \right] \quad \text{provided } \nu_{01}(t) > 0$$

Note that if $\nu_{01}(t) = 0$ then $H_{\underset{\sim}{x}}(\underset{\sim}{t}) = 0$ and this cannot be the maximum. The maximum must occur either at 0, 1, or at some point where $\frac{\partial \log H}{\partial t} = 0$.

At this point we must make some <u>assumptions</u> about the transition intensities. If

$$\nu_{ij}(t) = \nu_{ij}\beta(t)$$

then we need to solve

$$\nu_{11} - \nu_{00} \overset{?}{=} \frac{\beta'(t)}{[\beta(t)]^2} = -\frac{\partial}{\partial t}\left(\frac{1}{\beta(t)}\right).$$

If we further assume $\beta(t) \equiv 1$ then we have

$$\nu_{11} \overset{?}{=} \nu_{00}$$

independent of t. In other words, if $\nu_{11} = \nu_{00}$ then t has a uniform $(0,1)$ distribution — all values are equally likely. Also,

if $\nu_{11} > \nu_{00}$ then $\hat{t}_{mle} = 0$

if $\nu_{11} < \nu_{00}$ then $\hat{t}_{mle} = 1$.

Another approach would be the following. Assume that

$$\left|\frac{\nu_{01}'(t)}{\nu_{01}(t)}\right| << \left|\nu_{00}(t) - \nu_{11}(t)\right|$$

which will hold if $\nu_{01}(\cdot)$ is constant or nearly so. Then the $\hat{t}_{mle}$ may occur at $0, 1$, or very near a t with

$$\nu_{00}(t) \doteq \nu_{11}(t).$$

One sees readily that these assumptions are rather artificial considering what we would like to accomplish. However if the intensities are moderately well behaved (which I will not define here) then the mle may be one of the interval endpoints, and probably the one with the smaller intensity.

The intensities are usually unknown and must be estimated along with the t_i's, which would greatly complicate this picture. However I have not examined this yet.

Imperfect state information

I consider two types of imperfect information. The first, which I will call truncation, results from collapsing the state space into a subspace $S' \subset S$. More specifically let
$f_i = i \pmod{n}$ for some fixed n.

The flower phenology provides an excellent example. We may suppose that each return to the barren state brings the tree back to the same absolute state. However, some trees, especially in the tropics, may flower several times in one year, and the quality of flowering may depend on the success of previous attempts. Thus the process may be

$$\begin{tikzcd}
B_1 \arrow[r] & Fl_1 \arrow[r] \arrow[l, bend left] & Fr_1 \arrow[r] \arrow[ll, bend left] & B_2 \arrow[r] \arrow[lll, bend left] & Fl_2 \arrow[r] \arrow[llll, bend left] & Fr_2 \arrow[r] \arrow[lllll, bend left] & \cdots
\end{tikzcd}$$

in which the process reverts to B_1, the initial barren state if there is any failure. However, an ecologist may view this is

$$\begin{tikzcd}
B \arrow[r] & Fl \arrow[r] & Fr \arrow[ll, bend left]
\end{tikzcd}$$

Setting $n=2$ above yields this truncated process, and will be explored below.

The second type of imperfect information arises from misclassification. There is a very real chance that an observer of leaf damage might classify a lightly damaged leaf as severely damaged, and so on, since this is a rather quick in-the-field observation process. Graphically we have

$$\begin{tikzcd}
\text{actual} & H \arrow[r] \arrow[d] \arrow[dr] \arrow[rrr, bend left] & LD \arrow[r] \arrow[d] \arrow[dl] \arrow[dr] & HD \arrow[r] \arrow[d] \arrow[dl] & GONE \arrow[d] \\
\text{observed} & H & LD & HD & GONE
\end{tikzcd}$$

We might interpret these observed states as fuzzy sets (see Zadeh 1973), though I will not pursue that analogy here.

These two imperfect state processes fall in a general class called partially observeable Markov chains (Smallwood and Sondik 1973; see also Bertsekas 1976). They lay out a useful notation which I will adopt, though they proceed with a Markov decision problem approach.

These can be formulated as discrete- or continuous-time processes. I choose the latter as it is more general. Let

$x_0, x_1, \ldots$ be an enumeration of the actual states

$z_0, z_1, \ldots$ be an enumeration of the observed states.

Define the probabilities

$$P_j(t) = \Pr\{\text{in state } x_j \text{ at time } t\}$$

$$P_{jk}(s,t) = \Pr\{\text{in state } x_k \text{ at time } t \mid \text{in state } x_j \text{ at time } s\}$$

$$r_{jk}(t) = \Pr\{\text{observe } z_k \text{ at time } t \mid \text{in state } x_j\}$$

Then we can derive the following probabilities, where well defined.

$$a_k(t) = \Pr\{\text{observe } z_k \text{ at time } t\} = \sum_j P_j(t)\, r_{jk}(t)$$

$$a_{k\ell}(s,t) = \Pr\{\text{observe } z_\ell \text{ at time } t \mid \text{observe } z_k \text{ at time } s\} = \frac{\sum_\lambda \sum_j r_{\lambda k}(s)\, P_\lambda(s)\, P_{\lambda j}(s,t)\, r_{j\ell}(t)}{a_k(s)}$$

$$b_{kj}(t) = \Pr\{\text{in state } x_j \text{ at time } t \mid \text{observe } z_k\} = \frac{P_j(t)\, r_{jk}(t)}{a_k(t)}$$

[Note: throughout this section I assume there is perfect time information. that is, times of transition are known.

We can rewrite

$$a_{k\ell}(s,t) = \sum_j b_{kj}(s)\, P_{\lambda j}(s,t)\, r_{j\ell}(t)$$

However, these do not guarantee that $Z_t\ t \geq 0$ is Markov. In fact, at least for discrete time, $I_t, t \geq 0$ is Markov, with $I_t = \{Z_s, 0 \leq s \leq t\}$. (Bertsekas 1976)

Does this carry over to continuous time?

Thus I should consider the entire history when analysing the observed process. Below I only speak in terms of the z_t's.

The truncation process translates readily into these terms. We have

$$r_{jk}(t) = \left\{ \begin{array}{ll} 1 & \text{if } x_j = k \pmod n \\ 0 & \text{otherwise} \end{array} \right\}$$

Thus, let $A_k = \{ j \mid x_j = k \pmod n \}$ and hence

$$a_k(t) = \sum_{j \in A_k} p_j(t)$$

$$b_{kj}(t) = \left\{ \begin{array}{ll} p_j(t)/a_k(t) & \text{if } x_j = k \pmod n \\ 0 & \text{otherwise} \end{array} \right\}$$

$$a_{k\ell}(s,t) = \sum_{j \in A_\ell} b_{kj}(s)\, p_{xj}(s,t)\, r_{j\ell}(t)$$

Thus when $n=3$ we reduce to a 3-state process and can only infer back to the original process through the $b_{kj}(\cdot)$ probabilities. However, we cannot identify the $b_{kj}(\cdot)$ probabilities, or the $p_j(\cdot)$ or $p_{jk}(\cdot,\cdot)$, with only information about the observed z-process. This seems analogous to the "aliasing" problem encountered in spectral analysis of times series, though I don't know whether it merits pursuit.

The classification process may be visualized with the aid of the graph

$$\begin{tikzcd} 0 \arrow[r] \arrow[dr] & 1 \arrow[r] \arrow[d] & 2 \arrow[dl] \\ & \delta & \end{tikzcd}$$

which serves for actual and observed processes. Let $x_i = z_i = i$, $i = 0,1,2,\delta$ for ease of notation. By design we have

$$0 \le r_{jk}(\cdot) \le 1 \quad j,k \in \{0,1,2\}$$

$$r_{\delta\delta}(\cdot) = 1 \quad r_{j\delta}(\cdot) = r_{\delta k}(\cdot) = 0 \quad j,k \in \{0,1,2\}$$

Also, leaf damage increases, so a misclassification $1 \to 2$ cannot later be followed by $2 \to 1$. In other words

$$\left. \begin{array}{l} p_{jk}(s,t) = 0 \\ a_{jk}(s,t) = 0 \end{array} \right\} \text{if } j > k \;\; \forall\, 0 \le s \le t, \quad j,k \in \{0,1,2,\delta\} \quad \text{with } \delta > 2 \text{ by convention.}$$

This places a nonobvious (to me) time dependence on the $a_k(\cdot)$ and $a_{k\ell}(\cdot,\cdot)$ probabilities, and through them back on the $r_{jk}(\cdot)$ misclassification probabilities. Here we see the need to consider the vector I_t instead of z_t.

I have not taken this any further...

Bibliography

1. Aalen, O. (1976) Nonparametric inference in connection with multiple decrement models.
Scand. J. Statist. 3:15-27.

2. Aalen O. (1978) Nonparametric estimation of partial transition probabilities in multiple decrement models.
Annals of Stat. 6(3): 534-545.

3. Bertsekas D.P. (1976) Dynamic Programming and Stochastic Control.
Academic Press: 397 pp.

4. Chiang C.L. (1968) Introduction to Stochastic Processes in Biostatistics. Wiley.

5. Chiang C.L. (1979) Survival and stages of disease.
Math. Biosci 43: 159-171.

6. Chiang C.L. (1980) An Introduction to Stochastic Processes and their Applications. (to appear in book form)

7. Cox D.R. (1972) Regression models and life tables (with discussion). J. Roy. Stat. Soc. Ser B 34: 187-220.

8. Frankie G.W., Baker H.G., and Opler P.A. (1974) Comparative phenological studies of trees in tropical wet and dry forests in the lowlands of Costa Rica.
J. Ecol. 62: 881-919.

9. Kemeny J.G., Snell J.L. and Knapp A.W. (1976) Denumerable Markov Chains. Graduate Texts in Mathematics 40; Springer-Verlag.

10. Smallwood R.D and Sondik E.J. (1973) The optimal control of partially observable Markov processes over a finite horizon. Operations Research 21: 1071-1088.

11. Thomasian A.J. (1969) The Structure of Probability Theory with Applications. McGraw-Hill (out-of-print).

12. Zadeh L.A. (1973) Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man, and Cybernetics, vol SMC-3(1): 28-44.