

XIII. Sampling Models, 4: Introduction to time-varying models (forward problems)

1 Basic framework

(see Foote 2000, *Paleobiology* Supplement to 26(4):74-102, Foote 2001, *Paleobiology* 27:796 [erratum], and Foote 2003, *Journal of Geology* 111:125-148, 752-753 [erratum])

- 1.1 Let there be n time intervals, each characterized by a set of origination, extinction, and sampling rates: p_i , q_i , and r_i .
- 1.2 Use time series of p and q to predict true time series of N_{bt} , N_{Ft} , N_{FL} , and N_{bt} (and thus N_b and N_t).

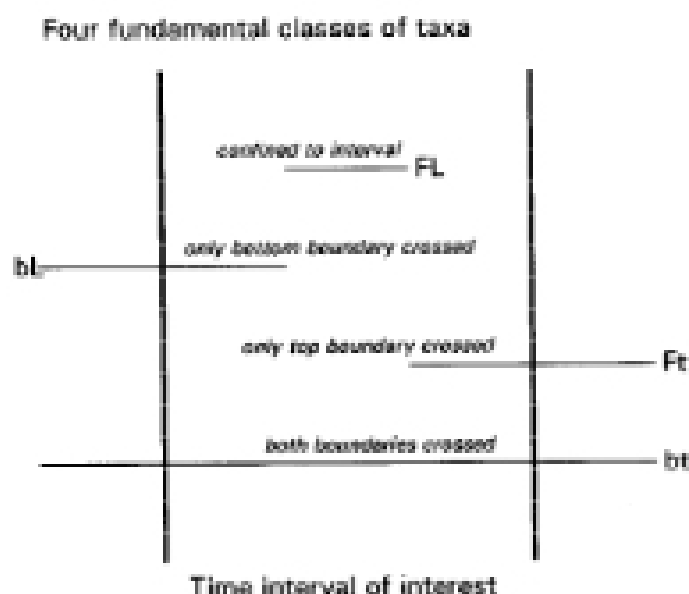


FIGURE 1. Illustration of four fundamental classes of taxa present during a stratigraphic interval. N_{FL} is the number of taxa confined to the interval, N_{bc} is the number that cross the bottom boundary only, N_{tc} is the number that cross the top boundary only, and N_{bc} is the number that cross both boundaries.

- 1.2.1 These quantities all scale to N_{bt} .
- 1.2.2 N_b for continuous-turnover model (q.v.) in which origination and extinction occur at constant per-capita rate within an interval:
 - 1. Let the age of the bottom boundary be at time $t = x$.

2. Let $N_0 = 1$ at $t = 0$.
3. Let p_t and q_t be time-specific rates.
4. Then

$$N_x = \exp \left[\int_0^x (p_t - q_t) dt \right].$$

5. Or, if we divide time into intervals and assume constant p and q *within an interval* (while still varying among intervals):

$$N_{bi} = \exp \left[\sum_{j=1}^{i-1} (p_j - q_j) \right],$$

if the rates p_i and q_i are expressed per lineage per interval. If instead they are expressed per lineage-million-years, and interval durations are given by Δt_i , then we have:

$$N_{bi} = \exp \left[\sum_{j=1}^{i-1} (p_j - q_j) \Delta t_j \right].$$

1.2.3 N_b for pulsed-turnover model (q.v.) in which originations are all at start of interval and extinctions at end of interval (so all lineages extend throughout the interval):

1. In this model, P is the number of new lineages produced per lineage extant at the start of the interval. (So number of originations is equal to $N_b \cdot P$ and total interval diversity is $N_b[1 + P]$.)
2. In this model, Q is the extinction probability. (So number of extinctions is equal to $N_b[1 + P]Q$.)
3. Thus

$$N_{bi} = \prod_{j=1}^{i-1} (1 + P_j)(1 - Q_j).$$

1.3 Use time series of p , q , and r to determine sampling probabilities for given time intervals.

Probability of being sampled in a given interval of time (either before a reference point, after a reference point, or between two reference points) is obtained as the integral, over all possible durations, of the probability of having a certain duration multiplied by the probability of being sampled given that duration.

Table 1. Expressions Used to Calculate Survivorship Probabilities

Quantity/model	Expression
N_k :	
•C	$N_k e^{-\tau}$
•P	$N_k(1 - q)$
N_k :	
•C	$N_k(1 - e^{-\tau})$
•P	$N_k q$
N_k :	
CC	$N_k e^{-\tau}(1 - e^{-\tau})$
PP	$N_k p(1 - q)$
CP	$N_k e^{-\tau}(1 - q)(1 - e^{-\tau})$
PC	$N_k p e^{-\tau}$
N_k :	
CC ¹	$N_k(e^{-\tau} + p - 1)$ if $p = q$, $N_k \frac{qe^{k+\tau} + (p - q)e^{-\tau} - p}{p - q}$ if $p \neq q$
••	
PP	$N_k pq$
CP	$N_k q(e^{\tau} - 1)$
PC	$N_k p(1 - e^{-\tau})$
$P_d(i)$:	
•C	$\sum_{k=1}^{\infty} \left((e^{-\tau_{k+1}^i} (1 - e^{-\tau}) (1 - (e^{-\tau_{k+1}^i} (1 - P_{tot}(k)))) + (e^{-\tau_{k+1}^i} (1 - (e^{-\tau_{k+1}^i} (1 - P_d(i)))) \right)$
•P	$\sum_{k=1}^{\infty} \left(\left(\prod_{m=1}^{k-1} (1 - q_m) \right) q_k (1 - (e^{-\tau_{k+1}^i} (1 - P_{tot}(k)))) + \left[\prod_{m=1}^{\infty} (1 - q_m) \right] (1 - (e^{-\tau_{k+1}^i} (1 - P_d(i)))) \right)$
$P_d(i)$:	
C	$\sum_{k=1}^{\infty} \left((e^{-\tau_{k+1}^i} (1 - e^{-\tau}) (1 - (e^{-\tau_{k+1}^i} (1 - P_{tot}(k)))) + (e^{-\tau_{k+1}^i} (1 - (e^{-\tau_{k+1}^i} (1 - P_d(i)))) \right)$
P	$\sum_{k=1}^{\infty} \left(\left(\prod_{m=1}^{k-1} \frac{1}{1 + p_m} \right) \left(\frac{P_k}{1 + p_k} \right) (1 - (e^{-\tau_{k+1}^i} (1 - P_{tot}(k)))) + \left(\prod_{m=1}^{\infty} \frac{1}{1 + p_m} \right) (1 - (e^{-\tau_{k+1}^i} (1 - P_d(i)))) \right)$
P_{tot} :	
•	$1 - e^{-\tau}$
P_{tot} :	
•C	$\frac{(r + qe^{k+\tau})/(q + r) - e^{-\tau}}{1 - e^{-\tau}}$
•P	$1 - e^{-\tau}$
P_{tot} :	
C ²	$\frac{(r + pe^{k+\tau})/(p + r) - e^{-\tau}}{1 - e^{-\tau}}$
P	$1 - e^{-\tau}$
P_{tot} :	
CC ¹	$\frac{N_k p \left(\frac{r}{p + r} - \frac{1 - e^{-\tau}}{p} - \frac{p(1 - e^{k+\tau})}{(p + r)^2} \right)}{N_k}$ if $p = q$, $\frac{N_k \left(\frac{pr(e^{k+\tau} - 1)}{(q + r)(p - q)} + \frac{pqe^{k+\tau}(e^{k+\tau} - 1)}{(p + r)(q + r)} - e^{-\tau}(e^{\tau} - 1) \right)}{N_k}$ if $p \neq q$
PP	$1 - e^{-\tau}$
CP	$\frac{r(e^{\tau} - 1) + p(e^{-\tau} - 1)}{p(r + p(e^{\tau} - 1))}$
PC ²	$\frac{r(e^{\tau} - 1) + q(e^{-\tau} - 1)}{q(r + q(e^{\tau} - 1))}$

Note. In the two-character code for model, the first character denotes origination and the second extinction, C = continuous, P = pulsed. A bullet means the expression applies to either model for the corresponding process. N_k is the true standing diversity at the start of the interval, because all relevant numbers scale to N_k , this can be arbitrarily set to unity.

¹ Foote 2000a, eqq. (6b) and (6c).

² Foote 2000a, eq. (27b).

³ Foote 2000a, eq. (28b).

⁴ Foote 2000a, eqq. (29b) and (29c).

⁵ Let z represent time within an interval of duration t , where $z = 0$ and $z = t$ are the beginning and end of the interval, respectively. By assumption, there is no extinction until the end of the interval. Thus, the density of origination at time z is equal to $e^z/(e^t - 1)$ (cf. Foote 2001a, eq. [3]). Because all lineages originating within the interval extend to the end, the probability of preservation, given origin at z and extinction at t , is equal to $1 - e^{-t-k}$. It is necessary to integrate the density of origination times the probability of preservation over all values of z . Thus, $P_{tot} = \int_0^t (1/e^t - 1) [e^z (1 - e^{-t-k})] dz$, which is equal to the expression in the table once t is set to unity.

⁶ Derived as in the foregoing footnote, with origination and extinction reversed.