Mathematical models of cladogenesis

David M. Raup

Abstract.—The evolutionary pattern of speciation and extinction in any biologic group may be described by a variety of mathematical models. These models provide a framework for describing the history of taxonomic diversity (clade shape) and other aspects of larger evolutionary patterns. The simplest model assumes time homogeneity: that is, speciation and extinction probabilities are constant through time and within taxonomic groups. In some cases the homogeneous model provides a good fit to real world paleontological data, but in other cases the model serves only as a null hypothesis that must be rejected before more complex models can be applied. In cases where the homogeneous model does not fit the data, time-inhomogeneous models can be formulated that specify change, regular or episodic, in speciation and extinction probabilities. An appendix provides a list of the most useful equations based on the homogeneous model.

David M. Raup.  Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637

Accepted: September 6, 1984

Introduction

Much of evolutionary paleobiology has been and continues to be concerned with the basic task of reconstructing phylogenetic trees: Who begat whom and when? As a result of many years of good biostratigraphy and systematics, the phylogenetic relationships of many biologic groups are reasonably well understood. Recent advances in molecular phylogenetic techniques and in the methods of phylogenetic systematics have brought us yet closer to an accurate knowledge of the total evolutionary tree. It has thus become possible to analyze the evolutionary dynamics of large segments of the tree with some rigor.

Any monophyletic group, or clade, owes its existence to the interplay of two processes: lineage branching (speciation) and lineage termination (species extinction). If the incidence of branching exceeds that of termination, the clade will survive and perhaps flourish, but if termination exceeds branching for a sufficient time, extinction of the clade is inevitable. These obvious facts have led evolutionary paleobiologists since the late nineteenth century to try to classify, describe, and model the “shapes” of clades, and this has produced a number of fundamental questions. Does a clade pass through a predictable sequence of stages over time? That is, does a clade have an “ontogeny” that can be defined and studied? Do the major groups of plants and animals differ fundamentally in rates of speciation and extinction? Are speciation and extinction rates inherent properties of an organism, perhaps as part of the genome, or are these rates determined largely by environmental factors? Are rates of speciation and extinction sufficiently regular and constant for the mean rates to have meaning? Are the rates sufficiently regular and constant for the application of simple mathematical models?

A great variety of answers to all these questions can be found in the literature. For example, there is the strongly held view that clades do have definable ontogenies (Thomson 1976, 1977; Gilinsky and Bambach 1984). Müller (1961, 1970, 1974) proposed and tested a number of what he called cladogenetic rules. One such is the Formen-Maxima rule, which states that geologically older taxa of a larger group (such as the Agnatha among the fishes) show a pattern of taxonomic diversity through time that is “regressive,” meaning that the point of maximum generic diversity occurs early in the history of the taxon. Geologically younger taxa in the same group show the diversity maximum later in their own histories, the extreme case being the “progressive” taxon which shows maximum diversity at the end of its time span, or at the Recent for extant organisms (such as teleost fishes). Stoyan (1980) and Stoyan and Stoyan (1980) modeled this rule mathematically using...
a time-inhomogeneous branching process. Their model allows rates to vary regularly over time and from group to group. The general Müller-Stoyan approach is highly deterministic: the cladogenetic process is constrained to play out a predetermined course in which there are few surprises—although it is granted that major perturbations such as the late Permian mass extinction do disturb the system (Stoyan 1980).

At the opposite extreme, a series of papers by the MBL Group (Raup et al. 1973; Gould et al. 1977, for example) explored the proposition that diversity variation through time is simply a result of random variation under conditions of constant probabilities of speciation and extinction. In this view, it is only a matter of chance whether a clade is regressive (low center of gravity, in the terminology of Gould et al. 1977) or progressive (high center of gravity).

There are many intermediate views between the two extremes just presented. For example, Gould et al. (1977) noted that early in an evolutionary radiation, clades tend to be regressive because speciation rates are unusually high, whereas in a more normal interval of evolutionary history the time of maximum diversity is randomly distributed. Thus, Gould et al. (1977) are allowing a deterministic explanation of clade shape in some instances but suggesting stochastic explanations in others. Also, Raup (1981) showed that clade shape in trilobites cannot be explained by a purely random model, and Stanley et al. (1981) noted several problems with application of such models to real world data.

My basic purpose in this review will be to summarize the variety of approaches that can be and have been applied to the analysis of branching patterns and to suggest some avenues for future research. The basic question will be, Is there enough order and regularity in the cladogenetic process to make rigorous mathematical analysis both feasible and informative?

Ingredients: Lineages, Speciation, Extinction, and Clades

For the purposes of this paper, a lineage will be defined as the evolutionary line consisting of an ancestor-descendant sequence of populations (Fig. 1). At any instant in time, the lineage is made up of the population(s) comprising a single biological species. This use of the term "lineage" is far more restrictive than most. In the present context, one could not speak of the "dinosaur lineage" unless there was only one species of dinosaur living at the time in question.

A count of lineages at a point in time represents species diversity (species richness). Whether a species constituting a lineage changes by anagenesis (phyletic transformation) over time need not concern us because we are only interested in the number of coexisting lineages. In other words, we need not be concerned with the question of stasis versus phyletic change. Even if there is enough anagenesis to transform one species into another through the duration of a lineage, the count of lineages is not changed.

Given this framework, a lineage can have only two fates in an interval of geologic time: it can persist or go extinct. Extinction is the complete termination of the lineage, and anything short of this (such as pseudoextinction by anagenesis) need not be considered. Given a collection of lineages, a convenient way to express the extinction rate is the number of extinction events divided by the sum of the time spans of all lineages, the latter being expressed in lineage million years (Lm). Thus, five lineages, each persisting for 10 ma before going extinct, would show an extinction rate of $5/50 = 0.1$, or an average of one extinction every 10 Lm.

A lineage can also branch by speciation to give off a new lineage. We need not be concerned with whether the speciation is allopatric, parapatric, or sympatric as long as the process adds one to the number of coexisting, genetically separate lineages. It is assumed as a virtual tautology that the ancestral lineage survives, at least for a time, after the branching event. A branching event that is absolutely coincident with the extinction of the ancestral lineage constitutes pseudoextinction and is not relevant in the present discussion. Branching rate (speciation rate) can be expressed also in terms of the number of events per lineage million years.

The foregoing is completely straightforward as long as the definitions are understood and followed. In any clade, the history of diversity is the inevitable consequence of the sequence of extinction and branching events. Problems lie, however, in the definition of clades. A single lineage that has no branches can be called a clade with total objectivity. At the other extreme, we
can consider all life on earth, descended directly or indirectly from a single ancestor, to be in one large clade, and this has been called a bioclade (Raup and Valentine 1983). Between these extremes, there is considerable room for honest confusion and semantic argument. To the phylogenetic systematist a clade is a group of lineages consisting of all the descendants of a single, common ancestor—that is, a truly monophyletic group. By this definition, the dinosau- rian order Saurischia does not constitute a true clade if, as is generally agreed, the birds evolved from a lineage that branched off from these dinosaurs. Because of the evolution of birds, the group of lineages commonly called Saurischia would be paraphyletic. The birds could be placed in a single monophyletic group (clade) that could be part of a larger such group that also includes the dinosaurs, but the Saurischian dinosaurs could not be placed in a clade by themselves.

In contrast to the cladistic treatment above, the traditional evolutionary taxonomist is quite comfortable placing the dinosaurs in one or more groups separate from the birds as long as the members are monophyletic to the extent that the dinosaurian condition did not arise from more than one ancestral lineage. That is, the evolutionary taxonomist accepts paraphyly but not polyphylly. The phylogenetic systematist insists on both (see Ashlock [1971] for further discussion). Whereas I agree with the phylogenetic systematists on this question where the issue is the logic of classification or the use of classification in determining relationships. I will choose a position closer to that of the evolutionary taxonomist for the present discussion of the mathematics of cladogenesis.

I will use the term "paraclade" for groups of lineages that, while not polyphyletic, may be monophyletic or not. That is, the lineages making up a paraclade are descended from a common ancestor, but the group need not contain all of the descendants of that ancestor. Para-
clades are usually definable morphologically because they represent concentrations or clusters of species in a larger morphospace. Such clusters are due to symplesiomorphy, which produces a region in the evolutionary tree in which there is a relative lack of change from lineage to lineage. Probably most genera, families, and orders in the paleontological literature are paraclades.

Two kinds of lineage branching events can occur within a paraclade. The first produces a species in the same small region of morphospace and thus contributes to the survival of the paraclade. The second produces a species sufficiently different from its immediate ancestor (farther away in morphospace) that it becomes the founder of a new paraclade and therefore does not contribute to the survival of the parent paraclade. Yule (1924) recognized this distinction when he noted that some speciation events yield species congeneric with the ancestral species whereas other speciation events initiate new genera.

The definitions of lineages and paraclades and of branching and extinction rates just given provide a basic framework with which one can build a mathematical description of cladogenetic patterns. This framework has been used in whole or in part by almost everyone who has attempted to model the cladogenetic process (Galton and Watson 1875; Yule 1924; and numerous workers in the 1970s and 1980s).

Homogeneous Branching Models

If one is willing to assume, or can verify, that probabilities of branching and extinction (expressed as events per time) are constant during the evolution of a clade or paraclade, then a substantial arsenal of equations is available for prediction of the history of diversity, or clade shape sensu Müller and Gould. The most basic of these equations predicts species diversity as a function of branching probability ($\lambda$), extinction probability ($\mu$), and time ($t$):

$$m_t = a e^{\lambda t - \mu t}.$$  \hspace{1cm} (1)

In this equation, $a$ is the initial number of species and $m_t$ is the number of species after some elapsed time, $t$. The quantity $a$ will be one if the clade is considered from the time of its founding or some higher value if the analysis starts after some branching has occurred. Because $\lambda$ and $\mu$ are probabilities, the calculated $m$, is an expected value subject to random error. The variance of this estimate is given by Bailey (1964, Eq. [8.49]). This general model is termed "time-homogeneous" because $\lambda$ and $\mu$ do not change through time.

Eq. (1) describes exponential growth if branching exceeds extinction ($\lambda > \mu$) or exponential decay if extinction exceeds branching ($\lambda < \mu$). It should be emphasized that change in any branching system of the sort described here is exponential, and one should avoid the vernacular use of the term exponential, reserved as it usually is for rapid growth or decay. As long as $\lambda > \mu$, a doubling time can be computed as $\ln 2/(\lambda - \mu)$. This is equivalent to Stanley's $t_2$ (Stanley 1979, Eq. [5.3]). If $(\lambda - \mu)$ is barely larger than zero, growth of diversity will be slow and doubling time large. As $(\lambda - \mu)$ increases in numerical value, doubling time shrinks and the rate of cladogenetic diversification approaches the vernacular "exponential" typical of adaptive radiations.

Eq. (1) has been used in a slightly different form by many authors to describe evolutionary diversification. The quantity $(\lambda - \mu)$ is replaced by a single constant for net rate of growth, often denoted by $R$ (Sepkoski 1978; Gillespie and Ricklefs 1979; Stanley and Newman 1980; Wilson 1983).

Because branching processes of the type described by Eq. (1) have proven extremely useful in a variety of fields, the literature contains a large number of equations derived from the basic process that can be applied directly to evolutionary problems. If one starts with a clade of a given size (one or more lineages), one can then estimate the fate of the clade at any time $t$ in the future. For example, one can calculate the probability of survival of the clade, the expected number of species at time $t$ if the clade survives, and the total progeny or the number of lineages produced between time 0 and time $t$. Also, one can calculate the expected frequency distribution of lineage durations. Perhaps more important, one can consider a whole collection of clades (or paraclades) operating under the same values of $\lambda$ and $\mu$. In this case, one can calculate the expected size distributions of clades, analogous to the commonly constructed frequency distributions of the number of species per genus in some
larger taxonomic group (the “hollow curve” of many authors). In short, if one is able or willing to assume constancy of branching and extinction probabilities, it is possible to predict (or postdict) the complete geometry and fabric of an evolutionary tree. A number of the most useful equations have been collected in the Appendix to this paper.

Many of these equations have been used implicitly or explicitly by numerous authors in attempting to describe cladogenetic patterns. In his landmark paper, Yule (1924) calculated generic size distributions for certain groups of angiosperms, beetles, and reptiles and made basic calculations of total species diversity which he then compared with observed values from the fossil record and from the extant biota. Yule’s treatment of extinction departed, however, from the homogeneous model, as will be discussed below. Van Valen’s (1973) analysis of taxonomic survivorship in the fossil record, leading to Van Valen’s law (Raup 1975), was based implicitly on the model described above. In fact, most of the survivorship analyses that followed Van Valen are similarly based on the fundamental birth-death process of Eq. (1).

Inhomogeneous Branching Models

The analysis of Müller’s cladogenetic rules by Stoyan (1980) and by Stoyan and Stoyan (1980), cited above, uses branching models that are elaborations of the model just presented. By Müller’s view of the evolution of major fossil groups, branching and extinction rates change through time and the model thus becomes time inhomogeneous. For a regressive clade, \( \lambda - \mu \) is strongly positive at first, leading to rapid exponential increase in diversity, and then \( \lambda - \mu \) becomes negative, leading to a gradual decline in diversity. Geologically younger paraclades in the same larger (monophyletic) clade follow the same course except that the reversal of the sign of \( \lambda - \mu \) comes later. In the final, progressive paraclade, \( \lambda - \mu \) remains positive to the end of the time span being considered. Stoyan was able to model this rather complex mixture of changing branching and extinction rates by modifying and expanding the basic equations. The justification for this elaboration is the contention, based on empirical data, that highs and lows in paraclade diversity are not randomly distributed in time. In other words, it is claimed that the simpler homogeneous model is not an accurate description of the real world.

At a considerably larger taxonomic scale, the kinetic model of Sepkoski (1978, 1979) describes the behavior of groups of clades (faunas) over Phanerozoic time. The Sepkoski model involves changing speciation and extinction rates as density- (diversity-) dependent phenomena. This then represents another kind of time-inhomogeneous model for the cladogenetic process but one that still involves regular and persistent changes in rates through time. Density dependence will be discussed more thoroughly in a later section.

There is yet another type of departure from time homogeneity that must be considered. This can be illustrated by the set of hypothetical survivorship curves in Fig. 2. Each of these curves describes taxonomic survivorship in the manner of Raup (1978), Hoffman and Kitchell (1984), and other authors. The form of plotting is that loosely termed “cohort” survivorship. The highest point on the left is the number of species that either originated or happened to be in existence at some particular time in the geologic past. The descending curves show the size of the “cohort” through time as it loses species by extinction. Unlike other survivorship curves (Van Valen 1973, for example), the horizontal axis in a cohort plot is the true geologic time scale, so that if the starting point on the left is, for example, a point in the mid-Cretaceous, decrease in cohort size can be seen for specific times following the mid-Cretaceous. The slope of the survivorship curve for any interval of time reflects the intensity of extinction during that interval. This can be seen by considering a variant on Eq. (1),

\[
S_t = S_0 e^{-\mu t},
\]

which, in logarithmic form, becomes

\[
\ln(S_t) = \ln(S_0) - \mu t,
\]

where \( S_0 \) and \( S_t \) represent the number of species remaining in the cohort at time 0 and time \( t \), respectively, and \( \mu \) is the extinction rate (or probability) as before. This equation is effectively the same as Eq. (1) except that speciation (\( \lambda \)) is absent because we are concerned only with the decay of a cohort through extinction. Note that
the equation is mathematically identical to that for radioactive decay, with $\mu$ being equivalent to the decay constant.

If the process is perfectly homogeneous (constant $\mu$), a plot of survivorship yields a straight line where the ordinate for plotting is a logarithmic scale. From Eq. (3), we see that the slope of the straight line is $-\mu$ and the intercept is the natural logarithm of $S_0$. Such a plot is illustrated in Fig. 2, A and represents perfect adherence to Van Valen’s law.

Three kinds of departure from time homogeneity, all of which have been found or claimed for paleontologic data, are shown in Fig. 2, B–D. In Fig. 2, B I represent a Stoyan-type inhomogeneous model: the convex curve might be typical of an early cohort of a regressive paraclade wherein the extinction rate is low at first and increases through time, indicated by the regularly increasing slope of the survivorship curve. It should be emphasized that survivorship curves show only the extinction half of the cladogenetic process; the Müller model of diversity history could also be accomplished by varying branching rate under conditions of constant extinction rate or by a combination of changes in both.

Patterns of cohort survivorship of a type that has appeared in a number of analyses of real world fossil data are shown in Fig. 2, C–D. In Fig. 2, C, the descent of the survivorship curve is in linear segments separated by sharp drops. This is the condition where a constant rate of background extinction is punctuated by sudden spurts of extinction. In the extinction episodes, the curve becomes essentially vertical indicating a near infinite extinction rate. The overall or average slope of the curve in Fig. 2, C is the same as the straight line in Fig. 2, A but contains a mixture of two radically different modes of extinction.

**Figure 2.** Four hypothetical taxonomic survivorship curves illustrating modes of extinction. Each diagram shows the decay of a cohort of species through geologic time. A, linear survivorship (perfect adherence to Van Valen’s law). B, survivorship curve indicating steady increase in extinction rate through the life of the cohort. C, linear survivorship interrupted by episodes of sudden increase in extinction. D, zero extinction rate punctuated by episodes of high extinction. The dashed lines in B, C, and D are copies of the linear survivorship curve of A.
The segments between the mass extinctions in Fig. 2, D, unlike those in Fig. 2, C, have no slope; that is, extinction rate is zero and the species are apparently at no risk of extinction. In Fig. 3 nested polycohort survivorship curves for some Cenozoic species of planktonic Foraminifera are shown (Hoffman and Kitchell 1984), and a mixture of the patterns of Fig. 2, C–D can be seen even though only one of the extinction events in the time span would conventionally qualify as a major extinction: the late Middle Miocene extinction of Raup and Sepkoski (1984) at about 13 ma B.P. The general appearance of the curves in Fig. 3 is that of a series of stairsteps with vertical or nearly vertical risers and horizontal to gently sloping treads. The horizontal portions do not indicate a lack of data for spans of several million years: the horizontal portions are documented by ample data, but the extinctions during these intervals are either lacking or negligible.

It is interesting that Yule (1924) perceived extinction in the manner of Fig. 2, D. Although he expressed evolutionary diversification as a mathematically homogeneous process (constant \( \lambda \)), he saw extinction as being episodic and used an explicitly episodic model in his mathematical formulation. To quote from Yule (1924, p. 23), “How then, it is naturally asked, do you account for the number of extinct species in the geological record? The answer is by their having been killed out; in the main, as it seems to me, by agencies which act less continuously than spasmodically and may fairly be described as cataclysmic . . . .”

The survivorship curves of Figs. 2, C–D, and 3 represent clear departures from the homogeneous model discussed earlier and also from the continuous but inhomogeneous models used by Stoyan. Extinction rates may be constant or changing gradually during fairly long intervals, only to be interrupted by short-lived extinction events. Furthermore, there is considerable evidence, at least for the larger mass extinctions of the Phanerozoic, that branching rates increase markedly in the recovery period following an extinction event (Sepkoski 1979). Appropriate plots of data on speciation should thus be compatible with Fig. 3.

Strict linearity of the survivorship curve is the prediction of Van Valen’s law, and such linearity has been found in a number of species level analyses, most notably by Van Valen (1973) for Cenozoic diatoms, planktonic foraminifera, and dinoflagellates. It has been shown, however, by Raup (1975) that if survivorship is linear for species, it should be nonlinear (concave upward) for higher taxa. This follows directly from the equations for clade survival \( (P_o) \) in the Appendix. The higher taxon equivalent of Fig. 2, C would thus have gently curved segments between the sharp drops and would illustrate the pattern described for genera of Phanerozoic invertebrates (Raup 1978).

Interactions between Branching and Extinction Rates

Eq. (1) and the discussion of it assume tacitly that the branching and extinction rates \( (\lambda, \mu) \) are independent quantities. Given this independence and given \( \lambda > \mu \), diversity should grow without limit. Although there has been debate for several years about whether speciation and extinction are diversity dependent, there is considerable evidence for such dependence when an entire biota is considered. Sepkoski (1978), in particular, has developed this idea for Phanerozoic diversity in the marine realm. In Sepkoski’s view, the exponential growth implied by Eq. (1) can continue only until use of resources or other space approaches saturation. The growth curve then tends to follow a logistic function until diversity is in dynamic equilibrium about some saturation level or carrying capacity. This clearly requires a time-inhomogeneous model for the cladogenetic process.

One can also develop a circumstantial argument for a link between branching and extinction rates simply by looking at the total history of the two quantities for the whole biota of the earth (our bioclade). Although there have been many fluctuations in both rates, the long-term average branching rate has been only a few percent higher than the long-term average extinction rate. This is clear from the fact that nearly all species that have ever lived are extinct; the increase in total number of species over geologic time is due only to the relatively slight excess of speciations over extinctions. Yet the two processes are as different as birth and death. It is difficult to conceive that the number of speciations and of extinctions could have remained so
similar were there not some connection between them. And diversity dependence is the most likely candidate to provide the link.

In spite of the foregoing, it is possible to argue that a relatively small paraphyletic operating in a much larger system of clades will not see total biotic diversity. The evolution of a genus or family of brachio pods, for example, may involve such small numbers of species that the large-scale, density-dependent effects may not have influence over the relatively short time spans of a genus or family. If this is true, the simple exponential of Eq. (1) may be allowed to stand.

The question of interactions between branching and extinction rates is by no means settled, and considerably more analysis is needed.

Testing Models with Empirical Data

It is clear that a range of mathematical models is available to test real data from the fossil record. The models include the ultra-simple homogeneous type where the only parameters are the constant rates of branching and extinction, the Stoyan-type inhomogeneous models where the two rates are allowed to vary over time and from clade to clade, and the more complex models where the system is perturbed occasionally by sudden extinction and/or radiation events. The perturbations could be incorporated into the mathematical formulation without too much difficulty, but this has been done only by Yule (1924), and his formulation was somewhat primitive. The timing of the perturbations could be entered either as a Poisson distributed random variable or as a periodic function, the latter being suggested by the apparently periodic nature of the larger extinctions (Fischer and Arthur 1977; Raup and Sepkoski 1984).

Naturally, we would like to know which of the several models most accurately describes the real world of evolutionary phenomena. If data bases were ideal, the model showing the best fit statistically could be identified. But the data are rarely good enough for this because of the noise produced by the many biases of sampling and preservation and by uncertainties in taxonomy and dating. For example, survivorship curves have been published in recent years where the noise level is so high that discrimination among the four alternative modes illustrated in Fig. 2 is impossible.

Much depends also on the temporal and taxonomic scale at which one chooses to work. Consider Fig. 2, C: the average survivorship is indicated by the dashed line drawn through the irregular stagerepp step pattern. This average represents the mean rate of extinction, with both kinds of extinction (background and mass) contributing to the average. For many questions of clade development, the mean rate might be quite satisfactory or might be the only accessible estimate of rates under conditions of noisy data. Furthermore, where time resolution is coarse, as was the case in analysis of generic survivorship at the system level (Raup 1978), the data points may not be close enough together on the time scale to resolve the stagerepp step pattern. In the system-level analysis just cited, resolution was good enough to see the effect of the major extinctions at the ends of the Permian and Cretaceous but not sufficient to see lesser extinction events. This is legitimate as long as it is recognized that the pattern observed is in some sense a function of the resolution of the biologic and geologic data.
Macroevolutionary Implications

Cladogenetic patterns are a critical ingredient in our understanding of the larger-scale evolutionary process. With a full understanding of cladogenesis and of the pattern and process of morphological change, it should be possible to describe and understand the process of macroevolution. The contrasts among the models presented in this paper are of nontrivial importance. At one extreme, the development of clades (or paraclades) is a continuous process wherein irregularities or departures from expectations are minimal. At the other extreme, the process is chaotic in the sense of being interrupted by relatively rare but sudden changes in the dynamics. And mass extinction represents the most severe and probably the most significant of these sudden changes. In fact, mass extinction and its aftermath may have a greater impact on evolution than the processes operating between the extinctions. Furthermore, if the Hoffman and Kitchell (1984) survivorship curves (Fig. 3) are typical, the steplike pattern may be present on a small scale between the several mass extinctions normally identified in the Phanerozoic record. Given high resolution data, it may be possible to show that cladogenesis is punctuational at almost all temporal scales. If this is indeed the case, time-homogeneous models may be useful, but only as a simplification of a more complex process.

Acknowledgments

Many individuals have contributed to the ideas and mathematical formulations described in this paper. I am especially grateful to student groups at the University of Rochester and at the University of Chicago.

Literature Cited


Appendix

Some Equations Predicting the Rates of Lineages and Paradacies Under the Time-Homogeneous Model

Notation

\( \lambda \) = speciation events per lineage millions years (Lmya)

\( \mu \) = speciation extinction events per Lmya

\( t \) = time in millions of years from some arbitrary starting point

\( a \) = number of coexisting lineages (species) at time 0

If the rates of speciation and extinction are viewed as probabilities, the equations yield only the statistical expectations. The uncertainty of this estimate, in the form of the variance, is given in a few critical cases. In other cases, equations for the variance can usually be found in the source literature or constructed by the user using monte-carlo techniques. As indicated, derivations of many of the equations can be found in the works of Kendall (1948) and Bailey (1964).

Two basically different kinds of applications of these equations are possible. The normal approach is to search for cases where real world data conform to the predictions of the equations. The alternative is to search for genuine departures from the predictions of the model. Where it can be shown that the time-homogeneous model does not apply, conservative arguments can be developed for alternative inhomogeneous models. The homogeneous model is thus used as a null hypothesis that should be rejected before it is legitimate to apply the more complex models (see, for example, Raup 1981).

1) Expected fate of a single lineage:

probability of survival at least until time \( t \):

\[
P_{0,s} = e^{-\mu t},
\]

(A1)

probability that a lineage will go extinct in a given time interval \( (t \rightarrow t+1) \):

\[
P_{0} = e^{-\mu t} - e^{-\mu(t+1)},
\]

(A2)

mean lineage duration:

\[
T_m = 1/\mu;
\]

(A3)

median lineage duration (half-life):

\[
T_m = -\ln(0.5)/\mu.
\]

(A4)

Note.—These four equations apply whether or not time is measured from the time of origin (speciation) of the lineage. That is, the time designated as \( t = 0 \) can be any point in the history of a lineage. Thus, if Eq. (A4), for example, is applied to a lineage that has already been in existence for several million years, \( T_m \) estimates half the time remaining for that lineage. Several of these equations have been applied in various forms to paleobiologic data by Raup (1975, 1978), Salthe (1975), Gillespie and Ricklefs (1979). Eq. (A1) is the basic equation for species survivorship if extinction rate is constant, and this forms the basis for Van Valen’s law (Van Valen 1973; Raup 1975).

2) Probability of the ultimate extinction of a paraclade (without limit of time):

where \( \lambda \leq \mu \):

\[
P_{0} = 1
\]

(Bailey, eq. [8.59]);

where \( \lambda > \mu \):

\[
P_{0} = (\mu/\lambda)^{a}
\]

(Bailey, eq. [8.59]).

Note.—The main value of these two equations is heuristic. Together, they say that unless branching probability is higher than extinction probability, the paraclade will inevitably become extinct. Even if the branching probability is higher, the paraclade has a nonzero probability of becoming extinct, with the numerical value of this probability being estimated by Eq. (A6) and hinging on the relative magnitudes of \( \lambda \) and \( \mu \) and on the number of lineages \( (a) \) at the starting point \( (t = 0) \).

3) Expected median duration of a paraclade (half-life):

where \( \lambda = \mu \) and \( a = 1 \):

\[
T_m = 1/\mu
\]

(Kendall, p. 7);

where \( \lambda = \mu \) and \( a > 1 \):

\[
T_m = (1/\lambda)/(2^{1/a} - 1)
\]

(Kendall, eq. (288));

where \( \lambda < \mu \) and \( a = 1 \):

\[
T_m = \frac{1}{\mu - \lambda} \ln \left( \frac{2}{\lambda} \right)
\]

(Kendall, eq. (262));

where \( \lambda < \mu \) and \( a > 1 \):

\[
T_m = \frac{1}{\mu - \lambda} \ln \left( \frac{2^{1/a} - 1}{\lambda} \right)
\]

(Kendall, eq. (293)).

Note.—Eq. (A7) for the half-life of a paraclade is algebraically identical to Eq. (A3) for the mean duration of a lineage but this is coincidental; Eq. (A9) has been applied to Paleobiologic data by Raup (1978).

4) Probability of paraclade extinction at or before time \( t \):

where \( \lambda = \mu \) and \( a = 1 \):

\[
P_{0,s} = (\lambda t)/(1 + \lambda t)
\]

(Bailey, eq. [8.53]);

where \( \lambda = \mu \) and \( a > 1 \):

\[
P_{0} = \left( \frac{\lambda t}{1 + \lambda t} \right)^{a}
\]

(Bailey, eq. [8.58]);

where \( \lambda \neq \mu \) and \( a = 1 \):

\[
P_{0} = \frac{\mu(\mu^{a-1} - 1)}{\lambda \mu^{a} - \mu}
\]

(Bailey, eq. [8.46]);

where \( \lambda \neq \mu \) and \( a > 1 \):

\[
P_{0} = \alpha^{a}
\]

(Bailey, eq. [8.47]);

where \( a = P_{0,s} \) from Eq. (A13) above.

Note.—These four equations can be “reversed” to provide estimates of the probability (\( P_{0,s} \)) of survival of a dae until time \( t \): \( P_{0,s} = 1 - P_{0} \). As such, they provide basic equations for clad survivorship and have been applied to various paleobiologic problems by Raup (1978), Ricklefs (1980), and Holman (1983). Also, Raup (1981) used Eq. (A14) to assess the proportion that the tribolites of the Paleozoic had values of \( \lambda \) and \( \mu \) typical of all invertebrates but went extinct by chance. \( P_{0,s} \) was calculated using \( a = 6000 \) (estimated number of tribolite species living at a point in the Cambrian), \( t = 350 \) ma (the remaining time before the end of the Permain), and values of \( \lambda \) and \( \mu \) estimated for all Paleozoic invertebrates. The calculated value of \( P_{0,s} \) was extremely low and provided convincing evidence that the tribolites must have gone extinct by having fundamentally different cladogenetic rates from other invertebrates of the time: either lower speciation rates or higher species extinction rates.

5) Probability that a paraclade will have exactly \( a \) species (where \( a \geq 1 \)) at time \( t \):

where \( \lambda = \mu \) and \( a = 1 \):

\[
P_{0,s} = \frac{(\lambda t)^{a-1}}{(1 + \lambda t)^{a+1}}
\]

(Bailey, eq. [8.53]);

where \( \lambda = \mu \) and \( a > 1 \):

\[
P_{0,s} = \frac{(\lambda t)^{a-1}}{1 + \lambda t} \sum_{j=1}^{\infty} \left( \frac{\mu}{\lambda} \right)^{j-1} \left( 1 - \frac{\mu}{\lambda} \right)^{j-1}
\]

(Stanley et al. 1981);

where \( \lambda \neq \mu \) and \( a = 1 \):

\[
P_{0,s} = (1 - a)(1 - \beta)\beta^{a-1}
\]

(Bailey, eq. [8.46]);

where \( \lambda \neq \mu \) and \( a > 1 \):

\[
P_{0,s} = \sum_{j=1}^{\infty} \left( \frac{a}{a + \lambda - 1} \right)^{a+j-1} \alpha^{a-j} \beta^{j-1} (1 - \alpha - \beta)^{j-1}
\]

(Bailey, eq. [8.47]);

where \( a \) and \( \beta \) are as in Eq. (A17) above.

Note.—These equations are potentially very useful for estimating an expected frequency distribution of paraclade size and thus can be used to generate artificial or predicted “hollow curves.” By solving any of the equations for all reasonable values of \( a \), one generates a probability density distribution for paraclade size that is, in effect, a prediction of the paraclade size frequency distribution.

6) Expected paraclade size (standing diversity) at time \( t \), including the possibility of prior extinction:

\[
\lambda < \mu \text{ and } a > 1:
\]

\[
T_{a} = \frac{1}{\mu - \lambda} \ln \left( \frac{2^{1/a} - 1}{\lambda} \right)
\]

(Kendall, eq. (293)).
where \( \lambda = \mu \) and \( a \geq 1 \):

\[
m_i = a
\]  
(A19)

(variance\( m_i \) = \( 2a \lambda t \))

(A20)

(Bailey, eq. [8.54];

where \( \lambda \neq \mu \) and \( a \geq 1 \):

\[
m_i = ae^{(\lambda - \mu)t}
\]  
(A21)

(Bailey, eq. [8.48];

(variance\( m_i \) = \( \frac{a(\lambda + \mu)}{\lambda - \mu} \cdot e^{(\lambda - \mu)t} (e^{(\lambda - \mu)t} - 1) \))

(A22)

(Bailey, eq. [8.49].)

Notes.—Eq. (A19) indicates that if speciation and extinction rates are identical, the statistical expectation is that the number of species (standing diversity) will not change over time. But note that the variance of this estimate can be extremely high. Eq. (A21) is the fundamental growth equation described in the text.

7) Expected parasite size (standing diversity) at time \( t \), assuming that the parasite survives at least until time \( t \):

\[
m'_t = 1 + \lambda t
\]  
(A23)

(variance\( m'_t \) = \( \lambda t (1 + \lambda t) \))

(A24)

where \( \lambda \neq \mu \) and \( a = 1 \):

\[
m'_t = \frac{\lambda e^{(\lambda - \mu)t} - \mu}{(\lambda - \mu)}
\]  
(A25)

(variance\( m'_t \) = \( \frac{\lambda (\lambda - \mu) e^{(\lambda - \mu)t} (1 - e^{-(\lambda - \mu)t})}{(\lambda - \mu)^2 e^{2(\lambda - \mu)t}} \))

(A26)

where \( \lambda \neq \mu \) and \( a \geq 1 \):

\[
m'_t = \frac{ae^{(\lambda - \mu)t}}{1 - \alpha^t}
\]  
(A27)

where \( a = P_0 \), from Eq. (A13) above.

Notes.—These equations are from T. Maruyama (pers. comm. 1976). In estimating expected parasite size after a given passage of time, it is very important to decide whether it is appropriate to assume the basic survival of the parasite (Eq. [A23]-[A27]) or not (Eq. [A19]-[A22]). The answer to this question depends on the nature of the problem.

8) Estimated total "progeny," including all lineages produced (surviving and extinct) up to time \( t \), including the possibility of prior extinction of the whole parasite:

\[
M_t = \alpha (1 + \beta t)
\]  
(Kendall, p. 10)

(A28)

where \( \lambda < \mu \) and \( a \geq 1 \):

\[
M_t = \alpha \left( \frac{\mu - \lambda e^{(\lambda - \mu)t}}{\mu - \lambda} \right)
\]  
(Kendall, p. 10, eq. [45])

(variance\( M_t \) = (see Kendall eq. [47]).)

9) Estimated total "progeny," including all lineages produced (surviving and extinct) up to time \( t \), assuming that the parasite survives at least until time \( t \):

\[
M'_t = 1 + \left[ \frac{\lambda}{1 - P_{0,t}} \cdot \left[ 1 + \sum_{i=1}^{t-1} \left( 1 - \mu + \lambda \right)i - P_{0,t} \cdot \prod_{i=1}^{t} \left[ 1 - \mu - \lambda + 2\lambda P_{0,i} \right] \right] \right]
\]  
(A29)

where \( P_{0,t} = \mu + P_{0,i} \cdot \left[ 1 - \mu - \lambda + 2\lambda P_{0,i} \right] \)

(A30)

with \( P_{0,1} = \mu \).

Notes.—Although not yet applied to paleobiologic data, Eq. (A30) has considerable potential in answering the question, How many species have there ever been in a given biologic group? As in the earlier total progeny case (Eq. [A28]-[A22]), it is assumed that \( a = 1 \). Eq. (A30) was developed from a class project by Paul Wolfson at the University of Rochester.