LOGNORMAL DISTRIBUTIONS

THEORY AND APPLICATIONS

edited by

Edwin L. Crow
Institute for Telecommunication Sciences
National Telecommunications and
Information Administration
U.S. Department of Commerce
Boulder, Colorado

Kunio Shimizu
Department of Information Sciences
Faculty of Science and Technology
Science University of Tokyo
Noda City, Chiba, Japan

MARCEL DEKKER, INC. New York and Basel

12

Applications in Ecology

BRIAN DENNIS College of Forestry, Wildlife, and Range Sciences, University of Idaho, Moscow, Idaho

G. P. PATIL* Center for Statistical Ecology and Environmental Statistics, Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania

1. INTRODUCTION

Perhaps the lognormal distribution finds the widest variety of applications in ecology. Ever since Malthus and Darwin, biologists have been acutely aware that populations of animals and plants grow multiplicatively. Studying the consequences arising from the enormous potential for increase possessed by most species on earth forms a major component of modern ecological research. Whenever quantities grow multiplicatively, the lognormal becomes a leading candidate for a statistical model of such quantities.

In this chapter, we discuss some of the theoretical and descriptive modeling studies in ecology that have featured the lognormal. We focus primarily on the lognormal as a model of the abundances of species and not as a model of the size growth of individual organisms. We review and critique several of the more important ecological modeling approaches related to the lognormal; in some cases, we display new results or offer thoughts

---

*Prepared in part as a Visiting Professor of Biostatistics, Department of Biostatistics, Harvard School of Public Health and Dana-Farber Cancer Institute, Harvard University, Boston, Massachusetts.

on future statistical and ecological research problems. The material in this paper is divided into three sections, for which we here provide the following summaries.

In Section 2, we reexamine the lognormal as a theoretical model of population abundance. The traditional multiplicative growth model is recast as a stochastic differential equation. Population size then becomes a diffusion process, that is, a Markov process with sample paths that are continuous functions of time. The transition distribution of the process is lognormal; various other statistical properties such as time-dependent measures of central tendency are obtained. The model is of limited practical usefulness in ecology, as it is essentially just a stochastic version of exponential growth. The model could only describe growth of a species for a short time interval, since all species eventually encounter environmental limits to growth.

However, we also describe a different stochastic growth model leading to the lognormal. The model is a stochastic differential equation based on the Gompertz growth equation. The model contains an underlying deterministic stable equilibrium for population size, representing the outcome of growth regulated by limiting environmental resources. The transition distribution for population size, as well as the equilibrium distribution, is lognormal in form. Thus, ecologists can regard the lognormal not only as a model of unbounded exponential growth, but also as a model of population regulation in the presence of an environmental carrying capacity.

An interpretive problem arises when using stochastic differential equations. A given stochastic differential equation represents two different diffusion processes, depending on whether Ito or Stratonovich stochastic integrals are used. A main reason for using stochastic differential equations is to approximate more complicated stochastic models. The details of the approximation process determine which type of stochastic integral to use. We show that the statistical properties of both lognormal models under the Stratonovich interpretation are easily recovered from those properties under the Ito interpretation, and vice versa, using the concept of weighted distributions.

We review in Section 3 the role of the lognormal as a model of species frequencies. The lognormal is confined to representing a single species in Section 2; by contrast, in Section 3, the lognormal represents patterns displayed by ecological communities with dozens, even hundreds, of species. Ecological and statistical research on quantitative species abundance patterns began in earnest with the introduction of the logseries model in the early 1940s. Shortly afterward, the lognormal model was proposed in reaction to the logseries, since many data sets did not appear J-shaped when plotted on a logarithmic scale. Numerous ecological studies have incorporated the lognormal model. Unfortunately, ecologists have not paid enough attention to sampling considerations and proper inference methods in these studies. As a result, whether the logseries, the lognormal, or some other distribution will be more widely applicable in species abundance studies is an open question. We try to clarify the problems of making statistical inferences for species frequency models, and we describe one promising inference approach that has been proposed but seldom used.

One intriguing aspect of the lognormal species frequency distribution is the so-called Canonical Hypothesis of species abundance. This hypothesis arose from an empirical pattern that had been noticed on logarithmic species frequency plots. The plots indicated that a randomly selected individual organism in the community would most likely come from a species whose log-abundance was in the same class as that of the largest species. The hypothesis received considerable attention in the ecological literature, and a "canonical" lognormal distribution was proposed having parameters constrained in such a fashion as to fix this abundance relationship. We review this hypothesis in Section 3, and we raise a cautionary note to the effect that the Canonical Hypothesis has seldom been formally tested in any way known to be statistically valid. The hypothesis has been studied more recently in the statistics literature, though. The results, which we summarize in Section 3, suggest that a new level of statistical awareness should be injected into the empirical studies of species frequency patterns.

Possibly the main role of the lognormal in ecology is simply to serve as the handiest adjustable wrench in the toolbox of statistical distributions. Ecological abundance data are intrinsically positive, with a few enormously high data points typically arising in every study. The lognormal distribution is an ideal descriptor of such data, with a positive range, right skewness, heavy right tail, and easily computed parameter estimates.

Ecological data sets, however, sometimes contain complicating factors which rule out the use of a simple two-parameter lognormal. In Section 4 we review three typical modifications of the lognormal. First, ecological data sets often consist of count data. The Poisson-lognormal distribution represents a discrete version of the lognormal potentially applicable to such cases. Second, ecological abundance surveys often contain an overly large number of samples with abundances of zero. The delta-lognormal, formed as a finite mixture of an ordinary lognormal distribution and a degenerate (spike) distribution at zero, offers advantages when estimating mean abundance is the objective of the surveys. Third, ecological abundances observed in samples sometimes grew from random numbers of initial propagules in each sample. We review a compound distribution structure recently proposed for such data; the structure also provides a degenerate component for added zeros.

We adopt the following notation throughout the paper. If $X = \log N$ has a normal distribution with probability density function (pdf) given by

$$f_X(x) = \frac{1}{(\sigma^2 2\pi)^{1/2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{1.1}$$

where $-\infty < x < +\infty$, that is, if

$$X \sim \text{normal}(\mu, \sigma^2) \tag{1.2}$$

then $N = e^X$ has a lognormal distribution with pdf

$$f_N(n) = \frac{1}{n(\sigma^2 2\pi)^{1/2}} \exp\left[-\frac{(\log n - \mu)^2}{(2\sigma^2)}\right] \tag{1.3}$$

and we write

$$N \sim \text{lognormal}(\mu, \sigma^2) \tag{1.4}$$

For additional information on discrete and continuous statistical distributions, including those appearing in this paper, we refer the reader to Patil et al. (1984a, 1984b, 1984c).

2. POPULATION GROWTH MODELS

2.1 Multiplicative Population Growth

MacArthur (1960) quantified for ecologists the notion that the abundance of a single species should have, under certain circumstances, a lognormal distribution. His reasoning was more intuitive than mathematical, but was nonetheless adopted by ecologists as a principal explanation of observed lognormal abundance patterns (May, 1975). MacArthur assumed that the growth rate of a species could be represented by an *ordinary differential equation* (ODE) of the form

$$\frac{dn(t)}{dt} = r(t)n(t) \tag{2.1}$$

where $n(t)$ is population abundance (typically measured in numbers of individuals or biomass per unit area or volume) at time t, and $r(t)$ is the per individual (or per unit biomass) growth rate. This ODE integrates to

$$\log n(t) = \log n_0 + \int_0^t r(\tau)\, d\tau \tag{2.2}$$

where $n_0 = n(0)$. MacArthur noted that the function $r(t)$ might vary randomly in time for some species due to fluctuations of environmental factors. The integral in (2.2) could then be regarded as the accumulated sum of random variables. MacArthur invoked the Central Limit Theorem to predict that $\log n(t)$ would have a normal distribution.

Two features of MacArthur's intuitive derivation are noteworthy: (a) the idea of random fluctuations in the per individual growth rate, and (b) the time-dependence of the normal distribution for $\log n(t)$. The fluctuations, for the derivation to hold, must be of such a nature that the sum of random variables given by $X_1 + X_2 + \cdots + X_k$, where

$$X_i = \int_{a_{i-1}}^{a_i} r(\tau)\, d\tau \tag{2.3}$$

and $0 = a_0 < a_1 < a_2 < \ldots < a_k = t$, conforms to one of the various Central Limit Theorem schemes. The resulting normal distribution for $\log n(t)$ would have a mean that essentially grows linearly with t and a variance that grows proportional to t. Thus, two ecological conditions underlying this derivation become apparent: (a) Any autocovariance of the fluctuations must decay rapidly for the Central Limit Theorem to hold. (b) The time t must be relatively early in the population's growth trajectory, before state-dependent changes in r, due to crowding or food limitation, become important. MacArthur pointed out that the model would only apply to opportunistic species, or species colonizing unutilized resources.

2.2 Stochastic Differential Equations

It is useful to derive various statistical properties for stochastic models such as MacArthur's, in order to test them with ecological data. The analysis is greatly simplified by using *stochastic differential equations* (SDEs). SDEs, known also as diffusion processes, can serve as approximations to many stochastic processes, including stochastic difference equations, branching processes, and birth-death processes (see Karlin and Taylor, 1981, p. 168). The approach to SDEs and lognormal growth models taken here follows that of Dennis and Patil (1984). See also Patil (1984).

An SDE model for the growth of a single species may be written as

$$dN(t) = N(t)g(N(t))\,dt + \sigma N(t)\,dW(t) \tag{2.4}$$

Here $N(t)$ is population abundance (now in upper case to denote a stochastic process) at time t, and $g(N(t))$ is the per unit abundance growth rate, which in general may depend on the population abundance. Also, $W(t)$ is a standard Wiener process ($W(t) \sim \text{normal}(0, t)$; $dW(t) \sim \text{normal}(0, dt)$) and σ is a positive scale constant. The form of (2.4) arises from an assumption that the per unit abundance growth rate, $g(N(t))$, is perturbed by unpredictable environmental factors. Mathematically, the differential $dN(t)$ is defined in terms of an Ito or a Stratonovich stochastic integral (e.g., Karlin and Taylor, 1981, p. 346). The quantity $N(t)$ becomes a diffusion process, a type of Markov process having continuous sample paths with probability one. Two functions, the infinitesimal mean and the infinitesimal variance, characterize most of the statistical properties of a diffusion process. They are defined respectively by

$$m_N(n) = \lim_{h \to 0} \frac{1}{h} E[\Delta N \mid N(t) = n] \tag{2.5}$$

$$v_N(n) = \lim_{h \to 0} \frac{1}{h} E[(\Delta N)^2 \mid N(t) = n] \tag{2.6}$$

where $\Delta N = N(t+h) - N(t)$. A standard result is that the infinitesimal mean and variance of the process $N(t)$ defined by the SDE (2.4) are

$$m_N(n) = ng(n) + \omega n \tag{2.7}$$

$$v_N(n) = \sigma^2 n^2 \tag{2.8}$$

Here ω is an indicator variable which depends on the type of stochastic integral being used to define the SDE (2.4): $\omega = 0$ if the Ito integral is used, and $\omega = \sigma^2/2$ if the Stratonovich integral is used.

Whether the Ito or Stratonovich integral is appropriate depends on the interpretation of (2.4) as an approximation to some underlying stochastic process. If (2.4) is seen as an approximation to a stochastic difference equation with uncorrelated noise, then the Ito interpretation should be used. If, however, $N(t)$ is viewed as an approximation to some process produced by integrating along a sample path of a smooth Gaussian process, then the Stratonovich interpretation of (2.4) should apply. These points are developed further by Ricciardi (1977), Karlin and Taylor (1981), and Horsthemke and Lefever (1984).

The fact that the Ito and Stratonovich interpretations of (2.4) produce different quantitative predictions has caused some consternation in the ecological literature (Feldman and Roughgarden, 1975; Turelli, 1977). The controversy has diminished in more recent years, as ecological modelers now do not take SDEs of form (2.4) too literally, but merely regard the SDEs as mathematically convenient approximations to more detailed, underlying processes. We show below that two SDE models producing lognormal distributions have qualitatively similar predictions under the Ito and Stratonovich interpretations, and that any differences are easily sorted out using weighted distributions.

One of the most useful features of diffusion processes is the transformation property. If $N(t)$ is a diffusion process, then $X(t) = h(N(t))$ is also a diffusion process, provided h is a continuous, strictly increasing (decreasing) function. The infinitesimal mean and variance of $X(t)$ are given by

$$m_X(x) = \frac{v_N(n)h''(n)}{2} + m_N(n)h'(n), \tag{2.9}$$

$$v_X(x) = v_N(n)[h'(n)]^2 \tag{2.10}$$

(if h' and h'' are uniformly continuous functions) with $n = h^{-1}(x)$ (Karlin and Taylor, 1981, p. 173). This property often permits the transformation of a novel diffusion process into a known process with well-studied statistical properties.

2.3 Stochastic Exponential Growth Model

MacArthur's intuitive model can be recast in statistical terms as an SDE of the form (2.4) with a constant per unit abundance growth rate:

$$dN(t) = rN(t)\,dt + \sigma N(t)\,dW(t) \tag{2.11}$$

The Stratonovich version of this model was extensively analyzed by Capocelli and Ricciardi (1974) (see also Tuckwell, 1974). A discrete time stochastic version of the exponential growth model was studied by Lewontin and Cohen (1969), and is essentially recaptured in the Ito version of (2.11). The differences in the Ito and Stratonovich versions were studied by Gray and Caughey (1965), Feldman and Roughgarden (1975), Ricciardi (1977), and Braumann (1983). The main properties of the model are found using the transformation $X(t) = \log N(t)$ and the formulas (2.9) and (2.10), producing

$$m_X(x) = r + \omega - \frac{\sigma^2}{2} \tag{2.12}$$

$$v_X(x) = \sigma^2 \tag{2.13}$$

These are the infinitesimal moments of a Wiener process with drift. A well-known result (e.g. Ricciardi, 1977, p. 58) gives a normal transition distribution for $X(t)$:

$$X(t) \sim \text{normal}\left(x_0 + \left(r + \omega - \frac{\sigma^2}{2}\right) t, \sigma^2 t\right) \tag{2.14}$$

Equivalently, the distribution of $N(t)$ becomes lognormal:

$$N(t) \sim \text{lognormal}\left(\log n_0 + \left(r + \omega - \frac{\sigma^2}{2}\right) t, \sigma^2 t\right) \tag{2.15}$$

It is interesting to compare various measures of central tendency with the deterministic solution of the ODE $dn(t)/dt = rn(t)$ given by

$$n(t) = n_0 e^{rt} \tag{2.16}$$

The mean, geometric mean, and harmonic mean of $N(t)$ are, respectively,

$$E[N(t)] = n_0 e^{(r+\omega)t} \tag{2.17}$$

$$\exp\{E[\log N(t)]\} = n_0 e^{[r+\omega-(\sigma^2/2)]t} \tag{2.18}$$

$$1/E[1/N(t)] = n_0 e^{(r+\omega-\sigma^2)t} \tag{2.19}$$

The expectations are conditioned on $N(0) = n_0$. Also, the median and mode of the distribution of $N(t)$ are found to be

$$\text{median}(N(t)) = n_0 e^{[r+\omega-(\sigma^2/2)]t} \tag{2.20}$$

$$\text{mode}(N(t)) = n_0 e^{[r+\omega-(3\sigma^2/2)]t} \tag{2.21}$$

For smaller values of r, some of the central tendency measures increase exponentially while others decay to zero. In fact, if $r < (\sigma^2/2) - \omega$, the probability that $N(t)$ is arbitrarily close to zero approaches 1 as t becomes large:

$$\begin{aligned} P[0 < N(t) \le \varepsilon] &= P[-\infty < X(t) \le \log \varepsilon] \\ &= P\left\{-\infty < \frac{X(t) - E[X(t)]}{[\text{var}(X(t))]^{1/2}} \le \frac{\log \varepsilon - E[X(t)]}{[\text{var}(X(t))]^{1/2}}\right\} \longrightarrow 1 \end{aligned} \tag{2.22}$$

as $t \to \infty$, since the last expression is the probability that a standard normal random variable is less than or equal to $[\log \varepsilon - x_0]/(\sigma\sqrt{t}) - [r + \omega -$

$(\sigma^2/2)]\sqrt{t}/\sigma$, which increases without limit as t increases. A similar result was described by Lewontin and Cohen (1969) for a discrete time process, and by Capocelli and Ricciardi (1974) for the Stratonovich version of this model.

Note that in the above central tendency measures, the geometric mean equals the median, which is a general property of the lognormal distribution. Also, the deterministic trajectory (2.16) equals the arithmetic mean for the Ito version, while for the Stratonovich version the deterministic trajectory equals the geometric mean. This point was stressed by Braumann (1983) in asserting that a main practical difference in the Ito and Stratonovich versions is the semantic interpretation of (2.16) as a mean. In fact, a cascade of interrelationships between the central tendency measures of the two versions exists (Dennis and Patil, 1984). Let $N_I(t)$ and $N_S(t)$ denote, respectively, the Ito and Stratonovich versions of $N(t)$. Then,

$$E[N_I(t)] = \exp\{E[\log N_S(t)]\} = n(t) \tag{2.23}$$

$$\exp\{E[\log N_I(t)]\} = 1/E[1/N_S(t)] \tag{2.24}$$

$$1/E[1/N_I(t)] = \mathrm{mode}(N_S(t)) \tag{2.25}$$

More generally, the generalized means of the Ito and Stratonovich versions are related. The θth moment about the origin of $N(t)$ is

$$E[(N(t))^\theta] = n_0^\theta e^{[r+\omega-\sigma^2/2]\theta t+(\sigma^2/2)\theta^2 t} \tag{2.26}$$

The generalized mean of $N(t)$ is then:

$$\{E[(N(t))^\theta]\}^{1/\theta} = n_0 e^{[r+\omega+(\theta-1)(\sigma^2/2)]t} \tag{2.27}$$

The Ito-Stratonovich interrelationship becomes

$$\{E[(N_I(t))^\theta]\}^{1/\theta} = \{E[(N_S(t))^{\theta-1}]\}^{1/(\theta-1)} \tag{2.28}$$

These interrelationships are derived from a general property of the exponential growth SDE: the Stratonovich transition lognormal distribution is a weighted Ito transition lognormal distribution. The statistical concept of weighted distributions was defined by Rao (1965) and has been investigated by Patil and Ord (1976), Patil and Rao (1977, 1978) and Mahfoud and Patil (1982). As pointed out by Dennis and Patil (1984), the Ito and

Stratonovich lognormal transition pdfs are related by

$$f_S(n,t \mid n_0) = \frac{n^{1/2} f_I(n,t \mid n_0)}{E[(N_I(t))^{1/2}]} \tag{2.29}$$

We will further point out here that the Stratonovich pdf can be obtained in this model as a scale-transformed Ito pdf. It is a property of the lognormal distribution that

$$\frac{n^\beta f(n)}{E[N^\beta]} = e^{-\lambda\beta} f(e^{-\lambda\beta} n) \tag{2.30}$$

where $f(n)$ is the pdf of a lognormal (μ, λ) random variable, N. Thus, letting $\beta = 1/2$, the Stratonovich pdf can be obtained from the Ito pdf through the scale transformation $N_S(t) = \exp(\sigma^2 t/2) N_I(t)$.

2.4 Stochastic Gompertz Growth Model

The stochastic growth model given by the SDE (2.11) could not be used indefinitely to represent a population's abundance. An increasing population would eventually encounter limits on nutrient supply, space, or other resources necessary for growth. This situation is frequently modeled in ecology with an ODE of the form

$$\frac{dn(t)}{dt} = n(t) g(n(t)) \tag{2.31}$$

(e.g. Freedman, 1980), where $g(n(t))$ is assumed to be a decreasing function of $n(t)$ with the following properties:

$$g(\bar{n}) = 0 \tag{2.32}$$

for some $\bar{n}$ such that $0 < \bar{n} < \infty$, and

$$g'(\bar{n}) < 0 \tag{2.33}$$

Population abundance for such models increases (or decreases) from n_0 to the stable equilibrium value given by $\bar{n}$.

One particular form of (2.31) is the Gompertz growth model:

$$\frac{dn(t)}{dt} = an(t) \log\left[\frac{\bar{n}}{n(t)}\right] \tag{2.34}$$

This ODE integrates readily to

$$n(t) = \exp\left\{\log \bar{n} + \left[\log \frac{n_0}{\bar{n}}\right] e^{-at}\right\} \tag{2.35}$$

This growth trajectory is a sigmoid curve with an inflection point at $\bar{n}/e$.

Such models can often be approximated quite well by the logistic growth model. The procedure approximates $g(n)$ with a linear function using a Taylor series expansion around $\bar{n}$ (Dennis and Patil, 1984):

$$\frac{dn(t)}{dt} = n(t)\left[r - \frac{r}{\bar{n}} n(t)\right] \tag{2.36}$$

Here $r = -\bar{n}g'(\bar{n})$. For the Gompertz model, the logistic approximation has $r = a$. This approximation has a trajectory that starts at n_0, and levels off at $\bar{n}$ (like the Gompertz), but the inflection point occurs at $\bar{n}/2$.

Stochastic versions of such models can be built as SDEs of the form (2.4). In the resulting stochastic models, population abundance does not level off at a stable equilibrium. Rather, the distribution for $N(t)$ may approach a limiting stationary distribution that is independent of the initial conditions as well as t. The stationary distribution, when it exists, has the following pdf (see Dennis and Patil, 1984):

$$f(n) = \psi \exp\left\{\frac{2}{\sigma^2}\int \frac{g(n)}{n}\,dn - 2\left(1 - \frac{\omega}{\sigma^2}\right)\log n\right\} \tag{2.37}$$

The constant ψ is evaluated by setting the area under $f(n)$ equal to one.

This stationary pdf is a member of the log-exponential family of pdfs defined by Patil and Ord (1976). A feature of this family is the "form-invariance" property: the size-biased version always retains the same form as the original pdf. Patil and Rao (1978) provide a general discussion of the properties of size-biased distributions.

One implication of this property is that the Ito and Stratonovich versions of the SDE (2.4) have stationary distributions of the same form (Dennis and Patil, 1984). From (2.37), the Stratonovich stationary pdf is found to be a size-biased Ito stationary pdf:

$$f_S(n) = \frac{n f_I(n)}{E[N_I]} \tag{2.38}$$

Thus, if the Ito version of (2.4) predicted a certain type of stationary distribution, such as a lognormal or a gamma, then the Stratonovich version would predict the same type (provided both pdfs exist). An immediate consequence of this relationship (2.38) is that the harmonic mean of the stationary Stratonovich pdf is the mean of the stationary Ito pdf:

$$1/E[1/N_S] = E[N_I] \tag{2.39}$$

The stochastic version of the Gompertz model is in the form (2.4), with $g(N(t)) = a\log[\bar{n}/N(t)]$. The stationary pdf is found from (2.37) to be that of a lognormal random variable:

$$N(\infty) \sim \text{lognormal}\left(\log\bar{n} - \frac{[(\sigma^2/2) - \omega]}{a}, \frac{\sigma^2}{2a}\right) \tag{2.40}$$

This result considerably extends the conceptual use of the lognormal in ecology as a population growth model. The lognormal under the MacArthur-type scenarios was strictly a time-dependent, transient model for a population in the early phase of its growth. By contrast, the lognormal (2.40) is a model for a population fluctuating around a stable equilibrium value.

A stochastic version of the logistic model (2.36) takes the form (2.4), with $g(N(t)) = r - (r/\bar{n})N(t)$. The stationary distribution is a gamma distribution with the following pdf:

$$f(n) = \frac{\alpha^\beta}{\Gamma(\beta)} n^{\beta-1} e^{-\alpha n} \tag{2.41}$$

where $\alpha = 2r/(\sigma^2\bar{n})$ and $\beta = 2r/\sigma^2 - 1 + 2\omega/\sigma^2$. If $\beta > 1$, the gamma has roughly the same shape as the lognormal: unimodal and right-skewed. In many practical instances, ecologists would not be able to distinguish between the lognormal and the gamma on the basis of fit to a given data set. Just as the logistic (2.36) is an approximation to growth models with a stable equilibrium, the gamma can serve as an approximation to the stationary distributions of stochastic growth models of the form (2.4). Dennis and Patil (1984) provide further discussion of the ecological role of the gamma as an abundance model.

The stochastic Gompertz model has the convenient feature that the complete transition pdf can be obtained (Ricciardi, 1977). The transformation $X(t) = \log N(t)$ yields, with the help of formulas (2.9) and (2.10), a diffusion process with infinitesimal moments given by

$$m_X(x) = a\log\bar{n} - \frac{\sigma^2}{2} + \omega - ax \qquad v_X(x) = \sigma^2 \tag{2.42}$$

These are the infinitesimal moments of the well-known Ornstein-Uhlenbeck process (e.g. Karlin and Taylor, 1981, p. 170). The transition distribution for $X(t)$ is normal:

$$X(t) \sim \text{normal}(\mu(t), \lambda(t)) \tag{2.43}$$

where

$$\begin{aligned}\mu(t) = {} & \log \bar{n} - \frac{(\sigma^2/2) - \omega}{a} \\ & + \left\{ \log n_0 - \log \bar{n} + \frac{(\sigma^2/2) - \omega}{a} \right\} e^{-at}\end{aligned} \tag{2.44}$$

$$\lambda(t) = \frac{\sigma^2}{2a}(1 - e^{-2at}) \tag{2.45}$$

Thus, the transition distribution of $N(t)$ is lognormal:

$$N(t) \sim \text{lognormal}(\mu(t), \lambda(t)) \tag{2.46}$$

We have here a time-dependent lognormal distribution that could represent a population's abundance for large values of t as well as small values. The limiting stationary distribution (2.40) is recovered from (2.46) as $t \to \infty$.

Various time-dependent measures of central tendency can be written down:

$$\{E[(N(t))^\theta]\}^{1/\theta} = e^{\mu(t) + \theta\lambda(t)/2} \tag{2.47}$$

$$\text{mode}(N(t)) = e^{\mu(t) - \lambda(t)} \tag{2.48}$$

The arithmetic mean, geometric mean (= median), and harmonic mean are found from (2.47) by setting $\theta = 1$, 0, and -1, respectively. In particular, the geometric mean of $N(t)$ for the Stratonovich version of the Gompertz SDE equals the deterministic growth trajectory given by (2.35).

The relationship between the Ito and Stratonovich versions of the stochastic Gompertz model is manifested in terms of a weighted distribution, as was the case for the stochastic exponential growth model. A property of the lognormal distribution relates the lognormal random variable N with its β-weighted version N_β. Specifically, if

$$N \sim \text{lognormal}(\mu, \lambda) \tag{2.49}$$

with pdf $f(n)$, and if N_β has a pdf given by

$$f_\beta(n) = \frac{n^\beta f(n)}{E[N^\beta]} \tag{2.50}$$

then

$$N_\beta \sim \text{lognormal}(\mu + \beta\lambda, \lambda) \tag{2.51}$$

This β-weighted lognormal is also the distribution of a scale transformation of the original random variable given by $\exp(\beta\lambda)N$ (see (2.30)). The property (2.51) is related to the variance-invariance characterization theorem of the lognormal due to Mahfoud and Patil (1982): N is lognormal iff $\text{Var}(\log N) = \text{Var}(\log N_\beta)$ for all $\beta > 0$. From (2.44), we find that the means of the log-transformed Ito and Stratonovich variables are related by

$$\mu_S(t) = \mu_I(t) + \frac{\sigma^2}{2a}(1 - e^{at}) \tag{2.52}$$

where $\mu_I(t) = E[\log N_I(t)]$, etc. Thus, the Stratonovich and Ito lognormal distributions share precisely the weighted relationship (2.50), where β is a function of time. Specifically,

$$f_S(n, t \mid n_0) = \frac{n^{\beta(t)} f_I(n, t \mid n_0)}{E[(N_I(t))^{\beta(t)}]} \tag{2.53}$$

where

$$\beta(t) = \frac{\sigma^2}{2a}(1 - e^{-at})/\lambda(t) \tag{2.54}$$

As t becomes large, the relationship (2.53) of the transition pdfs approaches the stationary size-biased relationship (2.38), since $\beta(t) \rightarrow 1$.

3. SPECIES FREQUENCY MODELS

3.1 Fisher's Logseries Models

A pivotal, three-part paper by Fisher, Corbet, and Williams (1943) launched four decades of ecological research on quantitative patterns of species abundance. Though it was not used by Fisher et al., the lognormal distribution has been one of the main tools used by ecologists in this research.

C. B. Williams had been studying samples of moths from Great Britain, and A. S. Corbet had been studying butterfly samples from Malaya. Many species were represented in the samples. Corbet and Williams noted that any particular sample usually had a large number of species represented by only a single individual apiece, while a less number of species were represented by two individuals, even less by three individuals, and so on. Corbet had observed that these species frequencies appeared to follow a harmonic series pattern: if N was the number of species with one individual in the sample, then $N/2$ was approximately the number of species with two individuals, $N/3$ was approximately the number of species with three individuals, and so on. Unfortunately, this mathematical model had the inconvenient property that it diverged: the sum given by $N(1+\frac{1}{2}+\frac{1}{3}+\ldots)$, if continued indefinitely, would predict an infinite number of species in the sample.

R. A. Fisher proposed a modification of the model. Fisher supposed that the expected numbers of species in the sample might be proportional to the terms of a negative binomial distribution. Upon analyzing the data of Corbet and Williams, Fisher found that the estimated values of the parameter k in the negative binomial were invariably small, usually quite close to zero. Fisher reduced the number of parameters in the model by taking the limit $k \to 0$, $s \to \infty$ in such a way that

$$sk \longrightarrow \alpha \tag{3.1}$$

in the negative binomial model. The quantity s is the proportionality constant in the negative binomial terms, representing the number of species in the ecological community being sampled (we must note that the limit (3.1) never explicitly appears in Fisher's discussion (Fisher et al., 1943), but rather is more or less implied). The result was that the expected number of species with r individuals in the sample, m_r, became proportional to the terms of a logseries distribution:

$$m_r = \frac{\alpha q^r}{r} \qquad r = 1, 2, 3, \ldots \tag{3.2}$$

where $\alpha > 0$ and $0 < q < 1$.

The logseries distribution has been used extensively since Fisher et al. (1943) to describe species frequencies, most notably by C. B. Williams

(1964) and R. A. Kempton (Kempton and Taylor, 1974, 1979; Kempton, 1975; Taylor et al., 1976).

3.2 Preston's Lognormal Model

F. W. Preston published an influential objection to the logseries model a few years after the Fisher et al. paper appeared (Preston, 1948). Preston worked with data sets on bird communities as well as moth communities, including some of Williams' data. Preston grouped the data into logarithmic abundance intervals which he called "octaves": the number of species with 1-2 individuals in the samples, with 2-4 individuals, with 4-8 individuals, etc., were displayed as a frequency histogram. Preston observed that the histograms, when drawn on such a logarithmic scale, tended to have modes, and in fact, tended to look quite Gaussian. Preston fitted a (left-truncated) normal curve to the histograms, which seemed to describe the data sets very well. Since Preston's paper, the normal curve has been widely used to "graduate," in Preston's words, species frequency data grouped into logarithmic abundance intervals (see reviews by Whittaker, 1972, and May, 1975).

3.3 Sampling Considerations

These applications of the lognormal as a species frequency model have unfortunately been marred by a lack of statistical rigor. Preston and subsequent investigators in many cases fit the Gaussian curves to the histograms by eye. Later, ecologists employed nonlinear regression routines to find the least-squares fits of the Gaussian curves to the histograms (Gauch and Chase, 1974). Such procedures ignore any probabilistic content of the Gaussian curve, ignore the intrinsically discrete nature of the data, and ignore sampling mechanisms. Lacking an explicit likelihood function, the ecologists are unable to provide valid confidence intervals for the parameters, test for goodness of fit, or tests for differences between samples.

Statisticians, in fact, have been unable to agree on the appropriate sampling model to use in conjunction with either the logseries or the lognormal models (see, for instance, Rao, 1971; Watterson, 1974; Kempton, 1975; Engen, 1979; Lo and Wani, 1983). It is unclear whether Fisher originally had an explicit sampling model in mind for the logseries. It is the authors' opinion that Kempton's (1975) sampling model is likely to find the widest use in species frequency studies, though more statistical and ecological research on this question certainly remains to be done. We will briefly describe Kempton's sampling model here, with attention to the role of the lognormal distribution in this approach.

Let N_r be the number of species with r representatives in the sample, $r = 1, 2, 3, \ldots$. The numbers $N_1, N_2, \ldots,$ are assumed to be independent, but not identically distributed, Poisson random variables. The total number of species in the sample is assumed to be a Poisson random variable with mean s. Also, the number of individuals in the sample of a particular species is assumed to be a Poisson random variable with mean λ. The values of λ differ among species; it is assumed that the λ values arise from a continuous distribution on the positive real line with pdf $f(\lambda)$. The result of these assumptions is that

$$E[N_r] \equiv m_r = s \int_0^\infty \frac{e^{-\lambda}\lambda^r}{r!} f(\lambda)\, d\lambda \tag{3.3}$$

The pdf $f(\lambda)$ would typically be that of either a lognormal or a gamma distribution. Thus, the observed species frequencies, $n_1, n_2, \ldots,$ are realized values of independent Poisson variables, $N_1, N_2, \ldots,$ whose means, $m_1, m_2, \ldots,$ contain a common set of unknown parameters. The unknown parameters are found in (3.3) and consist of s plus the parameters in the pdf $f(\lambda)$.

If the λ values arise from a lognormal (μ, σ^2) pdf, then

$$m_r(s, \mu, \sigma^2) = \frac{s}{r!(\sigma^2 2\pi)^{1/2}} \int_0^\infty \lambda^{r-1} \exp[-\lambda - (\log\lambda - \mu)^2/(2\sigma^2)]\, d\lambda \tag{3.4}$$

In other words, the expected values m_r are proportional to the terms of a discrete Poisson-lognormal distribution (Holgate, 1969; Bulmer, 1974; Kempton and Taylor, 1974; Shaban, this volume). For the gamma model, with $f(\lambda) = [\beta^k/\Gamma(k)]\lambda^{k-1}e^{-\beta\lambda}$, the m_r values are proportional to the terms of a negative binomial distribution:

$$m_r(s, k, \beta) = s\binom{k + r - 1}{r} q^r p^k \tag{3.5}$$

where $q = 1 - p = 1/(1 + \beta)$. Taking Fisher's limit $s \to \infty$, $k \to 0$, and $sk \to \alpha$ here produces the logseries:

$$m_r(\alpha, q) = \frac{\alpha q^r}{r} \tag{3.6}$$

Let the unknown parameters be denoted by the vector θ. The likelihood function becomes the product of Poisson probabilities:

$$l(\theta) = \prod_{r=1}^{\infty} \frac{\exp[-m_r(\theta)][m_r(\theta)]^{n_r}}{n_r!} \tag{3.7}$$

With product-Poisson sampling, iteratively reweighted least squares could be used for calculating maximum likelihood estimates (Jennrich and Moore, 1975). Using the lognormal model, though, requires an additional routine for numerical integration in order to evaluate the Poisson-lognormal terms (3.4).

3.4 Preston's Canonical Hypothesis

Preston (1962) noticed a curious pattern in his lognormal curves of species frequencies. The pattern formed the basis of "Preston's Canonical Hypothesis" of species abundance. The Canonical Hypothesis (CH) essentially states that the species frequency curves observed in nature will be predominantly lognormal, and that the parameter values observed will be found only in a small, constrained region of the parameter space. The CH has attracted considerable attention in the ecological literature (see May, 1975).

Specifically, the CH consists of a lognormal distribution with the following structure. We must first define the so-called individuals curve. If $f(\lambda)$ is the species abundance pdf in (3.3), then the expected number of species with abundances greater than λ would be

$$s \int_{\lambda}^{\infty} f(u)\, du \tag{3.8}$$

It would follow that the expected total abundance of all those species with abundance greater than λ would be

$$s \int_{\lambda}^{\infty} u f(u)\, du \tag{3.9}$$

Because of (3.8) and (3.9), $sf(\lambda)$ is called the species curve, and $s\lambda f(\lambda)$ is called the individuals curve. In economics, if $f(\lambda)$ represents a distribution of wealth among individuals, then (3.8) is the number of individuals with wealth greater than λ, and (3.9) is the total amount of wealth these individuals have cornered. On a logarithmic scale, with $\tau = \log \lambda$, these curves become $se^{\tau}f(e^{\tau})$ and $se^{2\tau}f(e^{\tau})$, respectively. When $f(\lambda)$ is a lognormal pdf, these logarithmic species and individuals curves are of course Gaussian.

Preston found, through examining many of his logarithmic histogram diagrams, that the mode of the logarithmic individuals curve tended to fall in the octave of the largest species. In other words, a randomly picked individual (dollar) would most likely come from the logarithmically largest species (wealthiest individual), rather than from, say, a group of species with intermediate logarithmic abundance (middle class). This pattern occured repeatedly in Preston's eye-fitted curves, leading Preston to propose a "canonical" lognormal distribution in which the parameters are constrained so as to fix this mode = max relationship.

Ecologists, judging from their literature, have practically come to regard the CH as an established empirical law of nature. Sugihara (1980), for instance, states: "Few propositions in ecology have as much empirical support as Preston's (1962) canonical hypothesis of species abundance." Sugihara goes on to propose a refinement of the lognormal sequential breakage model (see Aitchison and Brown, 1957; Pielou, 1975) which produces a canonical lognormal distribution: the pieces being broken are niches in a multidimensional niche space, and a breakage corresponds to the evolutionary splitting of a species or a successful invasion of a niche occupied by another species. Preston himself regards departures from the canonical lognormal distribution as indicative of defective, nonrandom sampling, of sampling heterogeneous ecological communities, or of sampling overpacked communities with more species than niches (Preston, 1980).

The enthusiasm ecologists have for this hypothesis must be judged from a statistical standpoint as premature. The studies supporting the CH are based on data sets analyzed with dubious parameter estimation methods having no known statistical validity. By contrast, Kempton's extensive analyses of British moth communities incorporated the explicit sampling model and likelihood function described earlier. These studies reported no evidence that the canonical lognormal is the best fitting distribution; in fact, the logseries model tended to outperform the full lognormal model for many of the moth collections (Kempton and Taylor, 1974; Taylor et al., 1976).

3.5 Statistics of Preston's Canonical Hypothesis

Patil and Taillie (1979a) have defined the CH in statistically precise terms. Their work provides formal statistical hypotheses concerning the CH that potentially could be tested for any data set on species frequencies.

Patil and Taillie define the predicted abundance of the largest species as

$$\lambda_{\max} = \bar{F}^{-1}\left(\frac{1}{s+1}\right) \tag{3.10}$$

where $\bar{F}(\lambda) = 1 - F(\lambda)$, and F is the cumulative distribution function given by

$$F(\lambda) = \int_0^\lambda f(u)\,du \tag{3.11}$$

The idea arises from the fact that $E[\bar{F}(\Lambda_{\max})] = 1/(s+1)$, where $\Lambda_{\max}$ is the largest observation from a random sample of size s from $f(\lambda)$. Then $\lambda_{\max}$ is a convenient, tractable approximation to $E[\Lambda_{\max}]$. The mode, $\tilde{\tau}$ of the logarithmic individuals curve is found by setting $d\log[se^{2\tau}f(e^\tau)]/d\tau = 0$. The CH is then formally stated by Patil and Taillie as $\log\lambda_{\max} \approx \tilde{\tau}$, or

$$\bar{F}^{-1}\left(\frac{1}{1+s}\right) \approx e^{\tilde{\tau}} \tag{3.12}$$

Using a lognormal (μ, σ^2) pdf for $f(\lambda)$, the CH becomes

$$\left[\bar{\Phi}^{-1}\left(\frac{1}{s+1}\right)\right]^2 \approx \sigma^2 \tag{3.13}$$

where $\bar{\Phi}$ is the right tail of a standard normal distribution. This statement of the CH amounts to a constraint on the parameters s and σ^2. If s is large, as is the case for most species abundance studies, then

$$2\log s - \log\log s - \log(4\pi) \approx \sigma^2 \tag{3.14}$$

provides a very good approximation to the relationship (3.13).

Patil and Taillie further point out that other distributions besides the lognormal could be used for $f(\lambda)$ in the CH (3.12). For instance, the gamma model with $f(\lambda) = [\beta^k/\Gamma(k)]\lambda^{k-1}e^{-\beta\lambda}$ yields the following version of the CH:

$$\frac{\Gamma(k, k+1)}{\Gamma(k)} \approx \frac{1}{s+1} \tag{3.15}$$

Here $\Gamma(k,x)$ is the incomplete gamma function defined by

$$\Gamma(k,x) = \int_{x}^{\infty} t^{k-1}e^{-t}\,dt \tag{3.16}$$

Thus, the CH for the gamma implies a constraint between the parameters s and k.

The parameters k in the gamma and σ^2 in the lognormal are related to the degree of evenness of the species abundances. The coefficient of variation in the gamma is $1/\sqrt{k}$, while in the lognormal it is $[\exp(\sigma^2)-1]^{1/2}$. Large k, or small σ^2, corresponds to a small coefficient of variation in the species abundances. The species would tend to have similar λ values under such circumstances, resulting in greater evenness of the abundances in the community. Taillie (1979) and Patil and Taillie (1979b) have formalized this notion of evenness in species curves using the concept of Lorenz ordering from economics. They have shown that k and σ^2 completely determine the Lorenz ordering for the gamma and lognormal models.

The CH constraints (3.13) and (3.15), as pointed out by Patil and Taillie (1979a), imply an inverse relationship exists between species richness and evenness in an ecological community. For the gamma model, in fact, the relationship (3.15) between k and s is well-approximated by

$$sk \approx 4.56 \tag{3.17}$$

for large s and small k. This is found by dividing both sides of (3.15) by k, taking the limit $s \to \infty$, $k \to 0$, $sk \to \alpha$, and then numerically evaluating the integral. So Preston's CH applied to the gamma model turns out to be a special case ($\alpha \approx 4.56$) of Fisher's limiting logseries!

A type of limiting lognormal model can be derived using the CH, in analogy with the logseries as a limiting form of the gamma model. The CH constraint (3.13) is approximately a linear relationship between σ and s for large $s : \sigma \approx a + bs$. In the Poisson-lognormal model (3.4) for species frequencies, one can substitute $\sigma = a + bs$ and take the limit as $s \to \infty$, producing

$$\lim_{s\to\infty} m_r(s,\mu,(a+bs)^2) = \frac{\gamma}{r} \tag{3.18}$$

for $r = 1, 2, \ldots$, with $\gamma = 1/[b(2\pi)^{1/2}]$. In a sense, we have come full circle in recovering Corbet's original harmonic series model for species frequencies as a limiting lognormal model. We point out that Patil and Taillie (1979a) studied a somewhat different divergent series as a limiting lognormal model obtained using a different limiting scheme.

The topic of species frequency distributions, from an ecological standpoint, would now benefit from some large-scale, serious data analysis. There is presently no reason to draw any more sweeping conclusions based on makeshift estimation techniques and eyeball testing. Claims concerning which distributions fit best, changes in distributions or parameters following ecological disturbance, or the CH should now be rigorously examined through careful attention to appropriate statistical modeling of sampling procedures. It is exciting to contemplate what patterns in nature remain to be discovered through a healthy injection of statistical thinking into species abundance studies.

4. MODIFIED LOGNORMAL MODELS AS DESCRIPTIVE ABUNDANCE MODELS

The lognormal is commonly used in ecology in a purely descriptive role as a model of abundance of a single species present in different samples. If many samples are taken across time or space, the abundance of a species typically varies greatly from sample to sample. The lognormal is used to describe these abundances mostly for convenience. Parameter estimates for the lognormal are easy to compute; and, an added attraction for ecologists is the theoretical underpinning of the lognormal as a single species growth model. (Section 2).

However, ecological data are frequently not so cooperative. Ecological studies can contain complicated factors, and the lognormal distribution often requires some modification for use as a descriptive model of abundance. We will not dwell here on reviewing standard descriptive uses of the lognormal in ecology. Rather, we will mention here a few of the typical modifications to the lognormal that are in use.

4.1 Poisson-Lognormal

When plankton are sampled using replicated net hauls or other methods, the frequency distribution of sampled abundances tends to be a unimodal, right-skewed distribution resembling a lognormal (Barnes and Marshall, 1951; Barnes, 1952). However, plankton samples are typically count data, representing numbers of particles suspended in a unit volume of water. A given sample could be assumed to have a Poisson distribution with mean parameter λ. Additional between-sample variability could then be induced by a mixing distribution with pdf $f(\lambda)$. The plankton count distribution is Poisson-lognormal if the mixing distribution is lognormal. The probabilities would then be

$$P[X = x] = \frac{1}{x!(\sigma^2 2\pi)^{1/2}} \int_0^\infty \lambda^{x-1} \exp[-\lambda - (\log \lambda - \mu)^2/(2\sigma^2)]\, d\lambda \quad (4.1)$$

where X is the number of particles in a unit volume of water. This distribution was discussed earlier in the entirely different context of species frequency models. Cassie (1962) gives an extensive discussion of the Poisson-lognormal as a plankton abundance model, with particular attention to its differences from the negative binomial. Further statistical properties and applications as a plankton model are developed by Reid (1981). Readers are also referred to the article in this volume by S. A. Shaban on the Poisson-lognormal distribution.

4.2 Delta-Lognormal

Data from surveys on abundances of marine organisms, including plankton, often contain a large proportion of zeros. The lognormal distribution typically provides a reasonable description of abundances for samples in which organisms are present. The spatial distribution of marine organisms tends to be patchy, though; samples are drawn from a mosaic of areas where organisms are present and areas where organisms are absent. When the objective of such surveys is to estimate mean abundance, there are advantages to using a modified lognormal distribution with an added discrete probability mass at zero (Pennington, 1983). Such a distribution is called a delta-distribution by Aitchison and Brown (1957) and a delta-lognormal distribution by K. Shimizu in Chapter Two of the present volume. The delta-lognormal has "pdf" given by

$$g(x) = \alpha\delta(x) + (1-\alpha)f(x) \quad (4.2)$$

where $f(x)$ is a lognormal pdf, and $\delta(x)$ is the Dirac delta function defined by

$$\int_a^b \delta(x)\, dx = \begin{cases} 1 & a < 0 < b \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

and $0 \leq \alpha \leq 1$. Estimation for distributions of the general form (4.2) was studied by Aitchison (1955).

The purpose of the marine abundance surveys often is to estimate $E[X] = \kappa$, which for the delta-lognormal becomes

$$\kappa = E[X] = (1 - \alpha)e^{\mu + (\sigma^2/2)} \tag{4.4}$$

Suppose a random sample of size n drawn from the delta-lognormal has m nonzero values, and suppose $\bar{y}$ and s^2 are the sample mean and sample variance, respectively, of the log-transformed nonzero values in the sample. One unbiased estimate of κ is of course the ordinary sample mean of the observations, zeros and all. However, $\bar{y}$, s^2, and m/n are joint complete sufficient statistics for μ, σ^2, and α, and this fact can be exploited to produce a much better estimate. Aitchison (1955) obtained the minimum variance unbiased estimate of κ; Pennington (1983) obtained the MVUE for the variance of the estimate and applied the results to fish and plankton survey data. Pennington noted that the MVUE for κ is considerably more efficient than the ordinary sample mean under the high variability conditions encountered in marine abundance surveys. These results and generalizations are contained in Section 3.1 of Chapter Two of the present volume.

We might remark here that it would be useful to study the addition of extra zeros to the Poisson-lognormal, in connection with marine surveys involving count data.

4.3 Delta-Compound-Lognormal

When terrestial plant communities are sampled with quadrats, the data often consist of large proportions of quadrats with no plants, and continuous, right-skewed distributions of plant abundances among quadrats where plants are present. Plant abundance in such studies is typically measured in terms of cover. The situation is more difficult than the preceding marine surveys in which the delta-lognormal could be used, for two reasons: (a) Plant cover present in a quadrat arises from a random number of initial propagules (seeds, rhizomes, etc.). (b) Plant cover typically grows as a function of time.

Steinhorst et al. (1985) proposed distribution models to describe plant cover development in forest communities following clearcutting and burning. The models consist of a randomly stopped sum of continuous iid random variables, plus an additional probability mass at zero. Such a model would have a Laplace-Stieltjes transform given by

$$\phi_Y(s) = \alpha + (1 - \alpha)\phi_N(-\log \phi_X(s)) \tag{4.5}$$

where Y is the total cover on a quadrat, N is the number of initial propagules on the quadrat (a discrete random variable on the non-negative integers), X is the size attained by a plant at the time of sampling, and $\phi_X(s) = E[e^{-sX}]$, $\phi_N(s) = E[e^{-sN}]$, $\phi_Y(s) = E[e^{-sY}]$. The parameters in the distribution of X were assumed to be functions of time such that $E[X]$ would follow a growth law like the logistic. Steinhorst et al. consider models in which N is either Poisson or negative binomial, and X is either gamma or normal. They were able to compute maximum likelihood estimates for data sets on various species of shrubs. The estimates were computed using the EM algorithm for those species which grow from rhizomes, since N is then an unobservable variable ("individual" plants not being distinguishable).

While Steinhorst et al. (1985) did not explicitly discuss using the lognormal as a distribution model for X, they have now investigated its use and are studying statistical inference problems for models in the form (4.5) in more generality (Steinhorst, manuscript in preparation). The lognormal would seem to be a promising candidate for modeling the size attained by a plant at a given time. The stochastic Gompertz model (see Section 2.4), for instance, would provide an explicit lognormal model with a mean that evolves according to a well-known growth law. The lognormal, however, does not have a convenient Laplace-Stieltjes transform, and so writing the model in the form (4.5) may not be very useful.

REFERENCES

Aitchison, J. (1955). On the distribution of a positive random variable having a discrete probability mass at the origin, *J. Amer. Statist. Assoc., 50*, 901–908.

Aitchison, J. and Brown, J. A. C. (1957). *The Lognormal Distribution*, Cambridge University Press, Cambridge, Massachusetts.

Barnes, H. and Marshall, S. M. (1951). On the variability of replicate plankton samples and some applications of 'contagious' series to the statistical distribution of catches over restricted periods, *J. Marine Biol. Assoc., 30*, 233–263.

Barnes, H. (1952). The use of transformations in marine biological statistics, *J. Cons. Explor. Mer., 18*, 61–71.

Braumann, C. A. (1983). Population growth in random environments, *Bull. Math. Biol., 45*, 635–641.

Bulmer, M. G. (1974). On fitting the Poisson lognormal distribution to species-abundance data, *Biometrics, 30*, 101–110.

Capocelli, R. M. and Ricciardi, L. M. (1974). A diffusion model for population growth in random environment, *Theor. Popul. Biol., 5*, 28–41.

Cassie, R. M. (1962). Frequency distribution models in the ecology of plankton and other organisms, *J. Anim. Ecol., 31*, 65–92.

Dennis, B. and Patil, G. P. (1984). The gamma distribution and weighted multimodal gamma distributions as models of population abundance, *Math. Biosci.*, *68*, 187–212.

Engen, S. (1979). Abundance models: sampling and estimation, *Statistical Distributions in Ecological Work* (J. K. Ord, G. P. Patil, and C. Taillie, eds.), International Co-operative Publishing House, Fairland, Maryland, pp. 313–332.

Feldman, M. W. and Roughgarden, J. (1975). A population's stationary distribution and chance of extinction in a stochastic environment with remarks on the theory of species packing, *Theor. Popul. Biol.*, *7*, 197–207.

Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population, *J. Anim. Ecol.*, *12*, 42–58.

Freedman, H. I. (1980). *Deterministic Mathematical Models in Population Ecology*, Marcel Dekker, New York.

Gauch, H. G. and Chase, G. B. (1974). Fitting the Gaussian curve to ecological data, *Ecology*, *55*, 1377–1381.

Gray, A. H. and Caughey, T. K. (1965). A controversy in problems involving random parametric excitation, *J. Math. and Phys.*, *44*, 288–296.

Holgate, P. (1969). Species frequency distributions, *Biometrika*, *56*, 651–660.

Horsthemke, W. and Lefever, R. (1984). *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology*, Springer-Verlag, Berlin.

Jennrich, R. I. and Moore, R. H. (1975). Maximum likelihood estimation by means of nonlinear least squares, *Am. Statist. Assoc. Proc. Statist. Comput. Sec.*, 57–65.

Karlin, S. and Taylor, H. M. (1981). *A Second Course in Stochastic Processes*, Academic Press, New York.

Kempton, R. A. (1975). A generalized form of Fisher's logarithmic series, *Biometrika*, *62*, 29–37.

Kempton, R. A. and Taylor, L. R. (1974). Log-series and log-normal parameters as diversity discriminants for the Lepidoptera, *J. Anim. Ecol.*, *43*, 381–399.

Kempton, R. A. and Taylor, L. R. (1979). Some observations on the yearly variability of species abundance at a site and the consistency of measures of diversity, *Contemporary Quantitative Ecology and Related Ecometrics* (G. P. Patil and M. Rosenzweig, eds.), International Co-operative Publishing House, Fairland, Maryland, pp. 3–22.

Lewontin, R. C. and Cohen, D. (1969). On population growth in a randomly varying environment, *Proc. Nat. Acad. Sci.*, *62*, 1056–1060.

Lo, H. and Wani, J. K. (1983). Maximum likelihood estimation of the parameters of the invariant abundance distributions, *Biometrics*, *39*, 977–986.

MacArthur, R. H. (1960). On the relative abundance of species, *Am. Nat.*, *94*, 25–36.

Mahfoud, M. and Patil, G. P. (1982). On weighted distributions, *Statistics and Probability: Essays in Honor of C. R. Rao* (G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh, eds.), North-Holland, pp. 479–492.

May, R. M. (1975). Patterns of species abundance and diversity, *Ecology and Evolution of Communities* (M. L. Cody and J. M. Diamond, eds.), Belknap Press, Cambridge, Massachusetts, pp. 81–120.

Patil, G. P. (1984). Studies in statistical ecology involving weighted distributions, in *Proceedings of the Indian Statistical Institute Golden Jubilee International Conference in Statistics: Applications and New Directions* (J. K. Ghosh and J. Roy, eds.), Statistical Publishing Society, Calcutta, India, pp. 478–503.

Patil, G. P., Boswell, M. T., Joshi, S. W., and Ratnaparkhi, M. V. (1984a). *Dictionary and Bibliography of Statistical Distributions in Scientific Work. Vol. 1: Discrete Models.* International Co-operative Publishing House, Fairland, Maryland.

Patil, G. P., Boswell, M. T., and Ratnaparkhi, M. V. (1984b). *Dictionary and Bibliography of Statistical Distributions in Scientific Work. Vol. 2: Continuous Univariate Models.* International Co-operative Publishing House, Fairland, Maryland.

Patil, G. P., Boswell, M. T., Ratnaparkhi, M. V., and Roux, J. J. (1984c). *Dictionary and Bibliography of Statistical Distributions in Scientific Work. Vol. 3: Multivariate Models.* International Co-operative Publishing House, Fairland, Maryland.

Patil, G. P. and Ord, J. K. (1976). On size-biased sampling and related form-invariant weighted distributions, *Sankhya B*, *38*, 48–61.

Patil, G. P. and Rao, C. R. (1977). The weighted distributions: a survey of their applications, *Applications of Statistics* (P. R. Krishnaiah, ed.), North-Holland, pp. 383–405.

Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families, *Biometrics*, *34*, 179–189.

Patil, G. P. and Taillie, C. (1979a). Species abundance models, ecological diversity, and the canonical hypothesis, *Bull. Int. Statist. Inst.*, *44*, 1–23.

Patil, G. P. and Taillie, C. (1979b). An overview of diversity, *Ecological Diversity in Theory and Practice* (J. F. Grassle, G. P. Patil, W. Smith, and C. Taillie, eds.), International Co-operative Publishing House, Fairland, Maryland, pp. 3–27.

Pennington, M. (1983). Efficient estimators of abundance, for fish and plankton surveys, *Biometrics*, *39*, 281–286.

Pielou, E. C. (1975). *Ecological Diversity*, John Wiley, New York.

Preston, F. W. (1948). The commonness, and rarity, of species, *Ecology*, *29*, 254–283.

Preston, F. W. (1962). The canonical distribution of commonness and rarity, *Ecology*, *43*, 185–215 and 410–432.

Preston, F. W. (1980). Noncanonical distributions of commonness and rarity, *Ecology*, *61*, 88–97.

Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment, *Classical and Contagious Discrete Distributions* (G. P. Patil, ed.), Statistical Publishing Soc., Calcutta, pp. 320–332.

Rao, C. R. (1971). Some comments on the logarithmic series distribution in the analysis of insect trap data, *Spatial Patterns and Statistical Distributions* (G. P. Patil, E. C. Pielou, and W. E. Waters, eds.), The Pennsylvania State University Press, University Park, Pennsylvania pp. 131–142.

Reid, D. D. (1981). The Poisson lognormal distribution and its use as a model of plankton aggregation, *Statistical Distributions in Scientific Work, Vol. 6* (C. Taillie, G. P. Patil, and B. A. Baldessari, eds.), D. Reidel, Dordrecht, Holland, pp. 303–316.

Ricciardi, L. M. (1977). *Diffusion Processes and Related Topics in Biology*, Springer-Verlag, Berlin.

Steinhorst, R. K., Morgan, P., and Neuenschwander, L. F. (1985). A stochastic-deterministic simulation model of shrub succession, *Ecological Modelling*, *29*, 35–55.

Sugihara, G. (1980). Minimal community structure: an explanation of species abundance patterns, *Am. Nat.*, *116*, 770–787.

Taillie, C. (1979). Species equitability: a comparative approach, *Ecological Diversity in Theory and Practice* (J. F. Grassle, G. P. Patil, W. Smith, and C. Taillie, eds.), International Co-operative Publishing House, Fairland, Maryland, pp. 51–62.

Taylor, L. R., Kempton, R. A., and Woiwood, I. P. (1976). Diversity statistics and the log-series model, *J. Anim. Ecol.*, *45*, 255–272.

Tuckwell, H. C. (1974). A study of some diffusion models of population growth, *Theor. Popul. Biol.*, *5*, 345–357.

Turelli, M. (1977). Random environments and stochastic calculus, *Theor. Popul. Biol.*, *12*, 140–178.

Watterson, G. A. (1974). Models for the logarithmic species abundance distributions, *Theor. Popul. Biol.*, *6*, 217–250.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, *21*, 213–251.

Williams, C. B. (1964). *Patterns in the Balance of Nature*, Academic Press, New York.