Ecological Applications, 6(4), 1996, pp. 1067–1076

UNCERTAINTY AND THE ASSESSMENT OF EXTINCTION PROBABILITIES[1,2]

DONALD LUDWIG
Departments of Mathematics and Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z2

Abstract. A proper assessment of the probability of early collapse or extinction of a population requires consideration of our uncertainty about crucial parameters and processes. Simple simulation approaches to assessment consider only a single set of parameter values, but what is required is consideration of all more or less plausible combinations of parameters. Bayesian decision theory is an appropriate tool for such assessment. I contrast standard (frequentist) and Bayesian approaches to a simple regression problem. I use these results to calculate the probability of early population collapse for three data sets relating to the Palila, Laysan Finch, and Snow Goose. The Bayesian results imply much higher risk of early collapse than maximum likelihood methods. This difference is due to large probabilities of early collapse for certain parameter values that are plausible in light of the data. Because of simplifying assumptions, these results are not directly applicable to assessment. Nevertheless they imply that maximum likelihood and similar methods based upon point parameter estimates will grossly underestimate the risk of early collapse.

Key words: Bayesian inference; probability of extinction; uncertainty; viability analysis.

INTRODUCTION

Assessments of population viability or probability of extinction are important parts of efforts to mitigate or control human impacts upon natural populations (Soulé 1987). The main objective of this work is to point out the large differences between the results of calculations of extinction probabilities that ignore our uncertainty and those that take uncertainty into account. I argue below that proper assessments must take account of uncertainty and that Bayesian decision theory is an appropriate means to this end.

In ecological investigations the difficulties in performing properly controlled experiments over appropriate scales in space and time and the complexity of the underlying genetics, behavior, and component processes preclude clear inference in most situations. We are faced with the problem of making assessments when several or many competing hypotheses may be plausible in light of the data. A simple approach to inference via hypothesis testing is often inconclusive, since the null hypothesis is seldom rejected but the power of our tests to reject is often low. Under such circumstances it is foolish to adopt an estimate for the probability of extinction based upon a null hypothesis when it has not been properly challenged. A second possible approach is to select the most likely hypothesis via a least squares or maximum likelihood estimation and calculate a probability of extinction based upon the "best estimate." Such an approach ignores our uncertainty. It is not a reliable guide to policy since it fails to take other plausible hypotheses into account (Lindley 1985, Morgan and Henrion 1990). A third approach to decision might be to attempt to minimize some risk or loss under the most unfavorable circumstances imaginable. Since there is no limit to the fertility of the human imagination, such a policy may be prisoner to the wildest current fear. It is better to attempt to assess the relative plausibilities of the various hypotheses and weight them accordingly. The statistical theory of decision was devised for this purpose (Chernoff and Moses 1959, Winkler 1972, Berger 1985, Lindley 1985). I shall follow that approach.

The fundamental difference between standard (frequentist) statistical theory and statistical decision theory (Bayesian statistics) is in the interpretation of the term "probability" and the associated random variables. To a frequentist, "probability" can only refer to the result of an infinite series of trials under identical conditions, while a Bayesian interprets "probability" to refer to the observer's degree of belief. Weather forecasts or betting odds are often stated in terms of Bayesian probabilities. In such cases there is no question of an infinite series of trials under identical conditions, but rather an organized appraisal in light of previous experience. In the present context where only limited data are available and uncertainty is large, the Bayesian interpretation is more appropriate.

Some scientists have difficulty in accepting Bayesian methods and interpretations in view of their apparent arbitrariness and subjectivity. These objections deserve a more lengthy discussion than can be given here (Lindley 1971). A short answer is that nothing in Bayesian theory is as arbitrary as the choice of 5% or 1% significance levels or the choice of null hypotheses. The "subjectivity" in Bayesian inference lies in the choice

[1] Manuscript received 14 August 1995; revised 12 February 1995; accepted 1 March 1996; final version received 5 April 1996.

[2] For reprints of this group of papers on Bayesian inference, see footnote 1, p. 1034.

of the prior density for unknown parameters, and in differing amounts and types of information available to various observers. There are systematic methods to choose or elicit priors (e.g., Wolfson et al. 1996). They have the advantage that the corresponding assumptions are explicit, while often similar assumptions about sampling distributions in frequentist inference are unexamined (Berger 1985). The fact that Bayesian theory takes account of differences between observers is an advantage. It allows us to examine the differences in assessments and decisions due to varying amounts of information, and hence to assess the value of additional information.

In this work I show how Bayesian methods may be used to calculate the probability of extinction from population census data. In order to simplify the exposition, I have neglected a number of important features, such as possible inaccuracy in the census data and random "catastrophes." These neglected features are important, but the simplified theory is adequate to assess the importance of including uncertainty. I develop a population dynamics model, a formula for the probability of early collapse, and a statistical model for estimation of the population parameters. I then contrast Bayesian linear regression theory with a frequentist treatment. These results are used to obtain the Bayes posterior distribution for the parameters. Finally, I contrast the Bayesian result with a conventional maximum likelihood approach for three data sets.

Assessment of Population Viability from Data

Small or declining populations are an obvious focus of concern when conservation decisions are made. We would like to concentrate efforts and resources on populations or systems whose survival or integrity are threatened and where there is a fair prospect that feasible actions may make a substantial difference. Such decisions depend upon an assessment of the likely consequences of various actions, including inaction.

Various indicators of population viability have been proposed, including (1) the probability of extinction assuming a certain starting population size and known parameters in the statistical model, (2) the probability of survival for a certain period under the same conditions as (1), and (3) the expected time to extinction under the same conditions as (1). The difficulty with (1) as a measure of threat is that according to the theory of stochastic processes, the probability of eventual extinction is always unity if extinction is possible. Hence the more complicated measure (2) may be more appropriate. The expected time to extinction (3) is simpler to compute than (2), but it may be distorted by the appearance of long survival times with low probability (Ludwig 1996). Here I use (4) the probability of collapse to a very low (unsafe) population size from a small one, before reaching a large (safer) population size. This indicator is as easy to compute as (3). It focusses on the short or medium term, and it is not distorted by averaging as (3) is. I shall call it "the probability of early collapse."

Now suppose that the consequences of various proposed actions are to be assessed in terms of the chosen indicator. Nothing is more common in the management of resources than surprise, usually unpleasant surprise (Ludwig et al. 1993). We may attempt to guard against such surprises by taking account of our uncertainty about the "true" parameter values or indeed the "true" population dynamics model. What is required of the biological and statistical sciences is an organization of our current knowledge of the system in order to provide a guide for action. Part of that organization is an assessment of our uncertainty about the system and the consequences of that uncertainty. I pointed out in the *Introduction* that point estimates or hypothesis testing may ignore important components of our uncertainty about the population dynamics that actually apply.

One way to take account of uncertainty is to simulate the population process by incorporating a pseudorandom number generator in computer models of the process (Burgman et al. 1993). One can estimate probabilities such as (2), (3), and (4) by collecting the results of many such simulations. In their simplest form, such simulations assume fixed (known) parameter values. Such simulations are unlikely to be adequate because we seldom or never have enough information to determine the influential parameters (such as intrinsic growth rate at low densities) with much precision. In response to this objection, we may perform "sensitivity analysis" by assessing the change in the indicators of interest when parameters or combinations of parameters are varied. One must choose a variety of more or less plausible parameter combinations and simulate with each of them. Such a procedure may consume much more computer time than the simpler version, in view of the great variety of parameter combinations that may be plausible in light of the data.

There are numerous difficulties with such an approach. If one is going to consider "plausible" parameter combinations, how does one assess their plausibility? One ought to apply statistical theory in order to obtain information about the pattern of variability of parameter estimates, but such considerations apparently do not enter into the prescriptions of Burgman et al. (1993). A frequentist approach to the sensitivity problem would use the sampling distribution of parameter estimates, assuming fixed, known parameters. A procedure of this sort will reveal poorly determined parameters and combinations of parameters. Such a procedure is carried out by Dennis et al. (1991). They use a local approximation near the point of maximum likelihood to estimate the variance of quantities of interest. They also mention bootstrap methods or "the approach of computing profile likelihoods and associated interval estimates" (Kalbfleisch 1986). The latter method is very close to a Bayesian method with a flat prior, which I employ below. All such analyses except for the last one rest upon the assumption of known parameters and many independent data sets, but instead we have one or a few data sets and the parameters are unknown. With a frequentist approach, we cannot even pose questions such as "What is the probability that the average net population growth rate is negative?," since the parameters and functions of the parameters are not random. According to frequentist statistics, the growth rate is either negative or it is not.

The Bayesian approach may be thought of as a systematic procedure for performing sensitivity analysis. There is a considerable literature bearing on decision making under uncertainty. Much of it has been motivated by economic or business decisions. The more theoretical work is known as "statistical decision theory." Some pertinent references are cited in the *Introduction.* In essence the theory measures "plausibility" of parameter values in terms of subjective or Bayesian probability. Decisions are compared and evaluated in terms of the "Bayes risk," which is a weighted sum of the utility or indicator function with weights given by the Bayes posterior distribution.

One might object that probabilities cannot depend upon an observer or the amount of information available. That is certainly true of the measures (2), (3), and (4). But if one performs a sensitivity analysis in order to account for our uncertainty about the underlying dynamics, the result will surely depend upon the assumptions that are made about which parameter combinations are plausible and how less plausible values are weighted. It is not a question of whether or not to make such judgements, but rather whether all of these judgements and assumptions are to be explicitly stated and examined, as in the Bayesian approach.

Bayesian Assessment of the Probability of Quasi-extinction

The Bayesian assessment of the probability of early collapse consists of two steps: (1) the probability of early collapse must be computed as a function of the parameters of an underlying statistical model (or possibly obtained by simulations), and (2) that probability must be considered over the posterior distribution for these parameters. The Bayes risk is given by the integral of that quantity over the posterior distribution. I argued above that the second step is required to account for our uncertainty about the true parameter values. If we knew the true parameter values, then the first step would be sufficient.

Statistical model

Assume that the population size N_t satisfies an equation of the form

$$N_{t+1} = N_t F(N_t) e^{\varepsilon_t}, \qquad (1)$$

where the function $F(N)$ describes the per capita net growth rate and the ε_t are normally distributed variations in the logarithm of the growth rate with mean zero and unknown variance σ^2. If

$$X_t = \log N_t, \qquad (2)$$

then Eq. 1 takes the form

$$X_{t+1} = X_t + f(X_t) + \varepsilon_t, \qquad (3)$$

where $f(X) = \log F(e^X)$. In order to simplify later results, I shall assume that $F(N)$ has the Gompertz form, and hence

$$f(X) = \beta_1 + \beta_2 X. \qquad (4)$$

Other forms could be incorporated, but then the probabilities of interest might have to be computed numerically instead of analytically. The parameter β_1 may be identified with the intrinsic growth rate r, and the carrying capacity k is given by the ratio $\exp(-\beta_1/\beta_2)$. The form of Eq. 4 is more natural for statistical analysis than one involving r and k since β_1 and β_2 appear as regression coefficients. In fact, if we define

$$Y_t = X_{t+1} - X_t, \qquad (5)$$

and assume the form of Eq. 4, then Eq. 3 takes the form of a linear regression.

Diffusion approximation

As explained at the beginning of this section, I shall calculate the probability of the population hitting a low value x_0 before reaching a high value x_1, as a function of the starting value x. The low value need not be zero: any low value that is deemed to be critical for survival of the population will do. This probability may be thought of as the probability of an early population collapse. In order to simplify its calculation, I shall employ a diffusion approximation. I provide a much more general theory in Ludwig (1996), and in fact I warn there against uncritical use of diffusion approximations. For purposes of illustration, the diffusion approximation is adequate, since it has the correct qualitative behavior.

Diffusion theory is based upon the replacement of Eq. 3 by the stochastic equation in continuous time

$$dX = f(X)dt + dW, \qquad (6)$$

where dW is normally distributed with mean 0 and variance $\sigma^2 dt$. If $u(x)$ denotes the probability of hitting x_0 before x_1 for a population starting at x, then u satisfies the diffusion equation (Ludwig 1974)

$$\frac{1}{2\sigma^2}\frac{d^2u}{dx^2} + (\beta_1 + \beta_2 x)\frac{du}{dx} = 0, \qquad (7)$$

with the boundary conditions

$$u(x_0) = 1,\ u(x_1) = 0. \qquad (8)$$

It is clear from Eqs. 7 and 8 that u depends upon x_0, x_1, β_1, β_2, and σ as well as x, and hence the later formulas will show this dependence. If $\hat{x}$ denotes the logarithm of the carrying capacity,

$$\hat{x} = \frac{-\beta_1}{\beta_2}, \tag{9}$$

then the solution of Eq. 7 with the boundary conditions (Eq. 8) is given by

$$u(x, x_0, x_1, \beta_1, \beta_2, \sigma) = \frac{\int_x^{x_1} \exp[-\beta_2(x' - \hat{x})^2/\sigma^2]\, dx'}{\int_{x_0}^{x_1} \exp[-\beta_2(x' - \hat{x})^2/\sigma^2]\, dx'}. \tag{10}$$

The numerical evaluation of this expression is described in Appendix A.

Uncertainty and Regression

In this section I address the second part of the calculation of a Bayesian assessment of the risk of early population collapse. The population dynamics model Eqs. 1 or 4 takes the form of a linear regression, as pointed out above. Now we must characterize the uncertainty of parameter estimates based upon such a linear regression. The calculations are almost the same as for a frequentist treatment of the problem, but the final result is a probability density for the regression parameters. This density may be used to provide weightings for the hypotheses corresponding to each set of parameter values. Guttman et al. (1982) is a general reference for the methods and results of this section.

I begin with the statistical model

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t, \; t = 1, \ldots, n \tag{11}$$

where Y_t are observations of the dependent variable, X_t are values of the explanatory variable, and ϵ_t are normally distributed errors. In order to simplify the exposition, I assume at first that the variance of ϵ_t is known to be σ^2; the more general case is treated below. Then the logarithm of the likelihood is given by

$$\ell = \frac{-1}{2\sigma^2} \sum_{t=1}^{n} [\beta_1 + \beta_2 X_t - Y_t]^2. \tag{12}$$

A few definitions and transformations simplify the calculations: let the means of the independent and dependent variables be given by

$$\bar{X} = 1/n \sum_{t=1}^{n} X_t, \qquad \bar{Y}_t = 1/n \sum_{t=1}^{n} Y_t. \tag{13}$$

For later convenience also set

$$s^2 = \frac{1}{(n-1)} \sum_{t=1}^{n} (X_t - \bar{X})^2. \tag{14}$$

The final formulas come out more simply in terms of the variables

$$\gamma_1 = \beta_1 + \beta_2 \bar{X}, \qquad \gamma_2 = \beta_2. \tag{15}$$

The choice of γ_1 and γ_2 instead of β_1 and β_2 produces a sum of squares involving γ_1 and γ_2 instead of a more complicated form in β_1 and β_2 in (Eq. 17) below.

The least squares or maximum likelihood estimates of γ_1 and γ_2 are given by

$$\hat{\gamma}_1 = \bar{Y}$$

$$\hat{\gamma}_2 = \frac{1}{(n-1)s^2} \sum_{t=1}^{n} (X_t - \bar{X})(Y_t - \bar{Y}). \tag{16}$$

With the definitions (Eq. 15), Eq. 12 takes the form

$$\ell = \frac{-1}{2\sigma^2}[n(\gamma_1 - \hat{\gamma}_1)^2 + (n-1)s^2(\gamma_2 - \hat{\gamma}_2)^2 + (n-2)S^2], \tag{17}$$

where the residual sum of squares is given by

$$(n-2)S^2 = \sum_{t=1}^{n} [\hat{\gamma}_1 + \hat{\gamma}_2(X_t - \bar{X}) - Y_t]^2. \tag{18}$$

Frequentist interpretation

The frequentist interpretation of the foregoing is that the variables ϵ_t are random, and hence the estimates $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are also random variables, as can be seen by substituting Eqs. 11 and 13 into Eq. 16. One then studies the sampling distribution of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ if repeated samples are taken with the same X_t and fixed values of the parameters γ_1 and γ_2. These estimates are unbiased: the expected values of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are γ_1 and γ_2, respectively. The random variables $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are uncorrelated and normally distributed, and their variances are given by

$$\sigma_1^2 = \frac{\sigma^2}{n}, \qquad \sigma_2^2 = \frac{\sigma^2}{(n-1)s^2}. \tag{19}$$

Bayesian interpretation

As indicated in the *Introduction,* the fundamental difference between a frequentist and a Bayesian viewpoint is in the interpretation of the term "probability" and the associated random variables. In contrast to the preceding frequentist interpretation, the Bayesian interpretation is that the data Y_t and X_t are fixed, since there is only one set of data under consideration. The variables γ_1 and γ_2 are regarded as random since we are uncertain of their true values. The joint probability distribution of γ_1 and γ_2 characterizes our uncertainty.

Subsequent calculations require the integration of a posterior density for γ_1 and γ_2 over the γ_1, γ_2 plane. It is necessary to start with a "prior density" for these variables to characterize the state of our knowledge before examining the data. Since I am assuming no information about γ_1 and γ_2 other than that contained in the data, I shall adopt a "flat" prior for γ_1 and γ_2, i.e., assume that all values of γ_1 and γ_2 are equally likely a priori. Since this density is integrated over the whole γ_1, γ_2 plane, this density is "improper": it leads to an infinite integral. The choice of prior is discussed in Jeffreys (1961) and Berger (1985).

According to Bayes' theorem, the posterior density for γ_1 and γ_2 is the normalized product of the prior density and the likelihood function. With the adoption of a flat prior, the density for the Bayes posterior is proportional to the likelihood function. After normalization to have an integral of unity, the posterior density is

$$p(\gamma_1, \gamma_2)\, d\gamma_1\, d\gamma_2 = \sqrt{\frac{n(n-1)s^2}{4\pi^2\sigma^4}} \exp\left[-\frac{n}{2\sigma^2}(\gamma_1 - \hat{\gamma}_1)^2 - \frac{(n-1)s^2}{2\sigma^2}(\gamma_2 - \hat{\gamma}_2)^2\right] d\gamma_1\, d\gamma_2. \quad (20)$$

Note that the term involving S^2 in Eq. 17 disappears when the density is normalized, since S^2 is independent of γ_1 and γ_2. The form Eq. 20 implies that γ_1 and γ_2 have a bivariate normal distribution: they are uncorrelated, they have means $\hat{\gamma}_1$ and $\hat{\gamma}_2$, respectively, and their variances are given by Eq. 19.

Exactly the same formulas appear in the frequentist version as in the Bayesian version, but their interpretation is quite different. In the frequentist version the statistics $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are normally distributed random variables since the data are random, while the parameters γ_1 and γ_2 are fixed. In the Bayesian version the data are fixed and hence the statistics $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are also fixed. The variables γ_1 and γ_2 are random, since they are uncertain. The two interpretations are incompatible.

The case of unknown variance

The main objective of this work is to characterize the distribution of functions of the parameters such as the probability of early population collapse given by Eq. 10. This quantity depends strongly upon the variance σ^2, which is unknown. Therefore the preceding discussion must be extended to describe the uncertainty about σ^2. The main complication that results is the necessity of choosing a prior density for σ. Note that σ must be positive, and hence the prior should be nonzero only on the positive axis. On the other hand, the logarithm of σ ranges from $-\infty$ to ∞, and a flat prior for $\log \sigma$ is consistent with the earlier choice of flat prior for γ_1 and γ_2. A flat prior for $\log \sigma$ is equivalent to a prior density for σ that is proportional to $d\sigma/\sigma$, and that is the choice that I make here. This point is discussed at length in Jeffreys (1961) and Berger (1985).

The posterior density $p(\gamma_1, \gamma_2, \sigma)d\gamma_1 d_2 d\sigma$ takes the following form since σ is now an unknown parameter:

$$p(\gamma_1, \gamma_2, \sigma)\, d\gamma_1\, d\gamma_2\, d\sigma = \frac{1}{\sigma^{n+1}} e^{\ell}\, d\gamma_1\, d\gamma_2\, d\sigma, \quad (21)$$

where ℓ is given by Eq. 17. Note that ℓ depends upon γ_1 and γ_2 only through a combination involving their squares. This motivates the introduction of polar coordinates in the γ_1, γ_2 plane. The radial coordinate will be denoted by ρ and the angular corrdinate will be denoted by θ. They are defined by

$$\rho^2 = \frac{n}{(n-2)S^2}(\gamma_1 - \hat{\gamma}_1)^2 + \frac{(n-1)s^2}{(n-2)S^2}(\gamma_2 - \hat{\gamma}_2)^2 \quad (22)$$

$$\tan \theta = \sqrt{\frac{(n-1)s^2}{n}}\left(\frac{\gamma_2 - \hat{\gamma}_2}{\gamma_1 - \hat{\gamma}_1}\right). \quad (23)$$

It is not sufficient to consider ρ alone, because the probability of early extinction depends upon θ as well as ρ, and we will have to integrate that probability over the ρ, θ plane. In place of σ^2, we may introduce w, defined by

$$w = \frac{(n-2)S^2(1+\rho^2)}{2\sigma^2}. \quad (24)$$

With these definitions, the posterior density takes the relatively simple form

$$p(\gamma_1, \gamma_2, \sigma)\, d\gamma_1\, d\gamma_2\, d\sigma = \frac{(n-2)\rho\, d\rho}{(1+\rho^2)^{n/2}}\left(\frac{1}{\Gamma(n/2)} e^{-w} w^{n/2-1}\, dw\right)\frac{1}{2\pi}\, d\theta. \quad (25)$$

This equation has been normalized so that each of the separate components integrates to unity. The numerical evaluation of integrals over the posterior density is described in Appendix B.

Comparison of the Bayes Assessment and Maximum Likelihood Results

When performing an assessment on the basis of census data, two possible methods are (1) calculation of a quasi-extinction probability using maximum likelihood estimates for the parameters β_1, β_2, and σ or (2) calculation of an integral of the quasi-extinction probability over the Bayes posterior distribution for the parameters.

Maximum likelihood estimates

The likelihood function as given by Eqs. 17 or 21 is maximized if $\gamma_1 = \hat{\gamma}_1$ and $\gamma_2 = \hat{\gamma}_2$: this is also the result obtained by a least squares estimation. The variance σ^2 is generally estimated by the unbiased estimator S^2. It is noteworthy that these values (including the estimate S^2) correspond to the mode of the posterior density as given by (25) at $\rho = 0$ and $u = n/2 - 1$.

Table 1 shows the maximum likelihood statistics for three data sets. Data for the Palila (*Loxioides balleui*) and the Laysan Finch (*Telespyza cantans*) were taken from graphs in Dennis et al. (1991), and data for the Snow Goose (*Anser caerulescens caerulescens*) were taken from a graph in Cooch and Cooke (1991). In each case, the numbers were divided by 1000 before taking logarithms to obtain the corresponding values of x.

TABLE 1. Maximum likelihood statistics as defined in Eqs. 13–18.

Species	n	$\bar{X}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	s^2	S^2	$\hat{\beta}_1$	$\hat{\beta}_2$
Palila	8	1.03	0.032	−0.485	0.210	0.206	0.530	−0.485
Laysan Finch	24	2.30	0.000	−0.932	0.114	0.119	2.144	−0.932
Snow Goose	17	1.50	0.0564	−0.217	0.176	0.0445	0.383	−0.217

Integrals of the probability of early collapse

The probabilities of early collapse were obtained as follows: The least squares estimate for the carrying capacity is

$$\hat{N} = 1000 \exp\left(\frac{-\hat{\beta}_1}{\hat{\beta}_2}\right). \quad (26)$$

In order to compare similar quantities, I choose x_0, x_1, and x by setting the quasi-extinction threshold at $N_0 = 0.1\hat{N}$, the upper limit at $N_1 = 0.8\hat{N}$, and the starting value $N = 0.2\hat{N}$. Then $x_0 = \log(N_0/1000)$, $x_1 = \log(N_1/1000)$, and $x = \log(N/1000)$. The corresponding values of N_0, N_1, and N are shown in Table 2. The probabilities shown in that table are the probabilities of a population reaching the lower value N_0 before the upper value N_1, starting at size N, as obtained from Eqs. 29, 32, and 33. The values in the column headed "Maximum likelihood" were obtained by substituting the maximum likelihood estimates into those formulas. The values in the column headed "Bayes posterior" were obtained by applying Eq. B.12 to the probabilities of early collapse with $N_\theta = N_\rho = N_w = 20$. In each case, the Bayes posterior values are far larger than those obtained by maximum likelihood. In order to aid understanding of this phenomenon, I provide additional details about the posterior distribution of the probability of early collapse.

Palila data

Fig. 1 shows the cumulative distribution functions (using the Bayesian posterior distribution for the parameters) for the the probability of early collapse of the Palila population, and its logarithm to the base 10. The cumulative distribution function is approximated by sorting the values associated with the grid points in a numerical integration scheme: the associated weights are one in the present scheme. The horizontal dotted lines in each figure indicate the quantiles for 5, 25, 50, 75, and 95%, respectively. The solid curve in the right-hand panel intersects the 50% quantile line somewhat below −2. Hence the median of the posterior distribution of the probability of early collapse is fairly close to 0.00262, which is the maximum likelihood estimate from Table 2. The portion of the horizontal axis between the outermost dotted vertical lines gives the Bayesian credibility interval for the probability of early collapse: we expect that the true value of that probability lies within that interval with a posterior probability of 90%. In contrast, the corresponding frequentist confidence interval is a random interval that is expected to contain the true parameter value for 90% of a set of random data drawn under the same postulated conditions as the actual data. We may also use these curves to check the plausibility of the computed value for the probability of early collapse. In the left-hand panel of Fig. 1 the solid curve is fairly straight above the 75% quantile line. The mean of that portion of the distribution is ≈0.6, which yields a lower estimate for the mean of the entire distribution of 0.25 × 0.6 = 0.15. Hence the value of 0.171 in Table 2 is plausible.

Laysan Finch data

The Laysan Finch data of Dennis et al. (1991) provide an interesting contrast to the Palila data. The ratio of the maximum likelihood and Bayes results is even more extreme than before, but not enough to produce a very substantial posterior estimate for the probability of early collapse. The left-hand panel of Fig. 2 shows the cumulative distribution function for the logarithm to the base 10 of that probability. If one examines the interval between the 5 and 95% quantiles, it corresponds to a multiple of over 10^{10} in the probabilities. Of course, the very small probabilities are of no great significance: the probability of early collapse is small no matter how the various possible parameter values are weighted.

Snow Goose data

A corresponding plot for the Snow Goose data is shown in the right-hand panel of Fig. 2. This case is intermediate between the other two.

If restricted to a single summary statistic, the integral of the probability of early collapse over the Bayes pos-

TABLE 2. Quasi-extinction probabilities. Here N_0, N_1, and N represent 10, 20, and 80% of estimated carrying capacity, respectively.

Species	N_0	N_1	N	Maximum likelihood	Bayes posterior
Palila	299	2390	598	0.00262	0.171
Laysan Finch	1000	7988	1998	8.77×10^{-10}	1.94×10^{-4}
Snow Goose	583	4660	1166	2.60×10^{-6}	0.0442

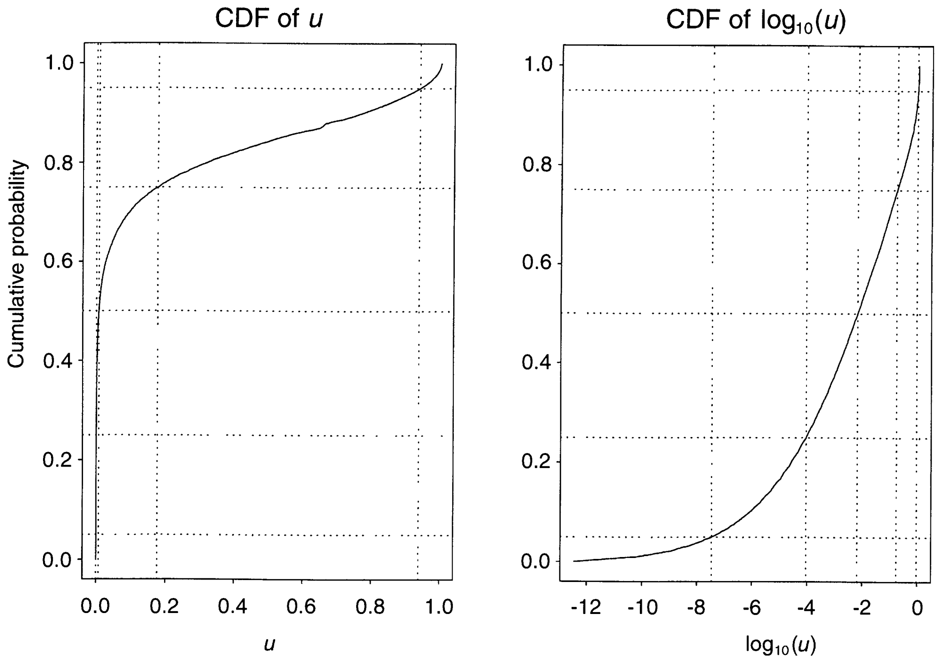

FIG. 1. The posterior cumulative distribution function (CDF) for the probability of early collapse, given by Eq. 10, based upon data for the Palila. The horizontal dotted lines correspond to the quantiles for 5, 25, 50 (median), 75, and 95%, respectively. The right-hand panel has the horizontal axis transformed by taking the logarithm to base 10.

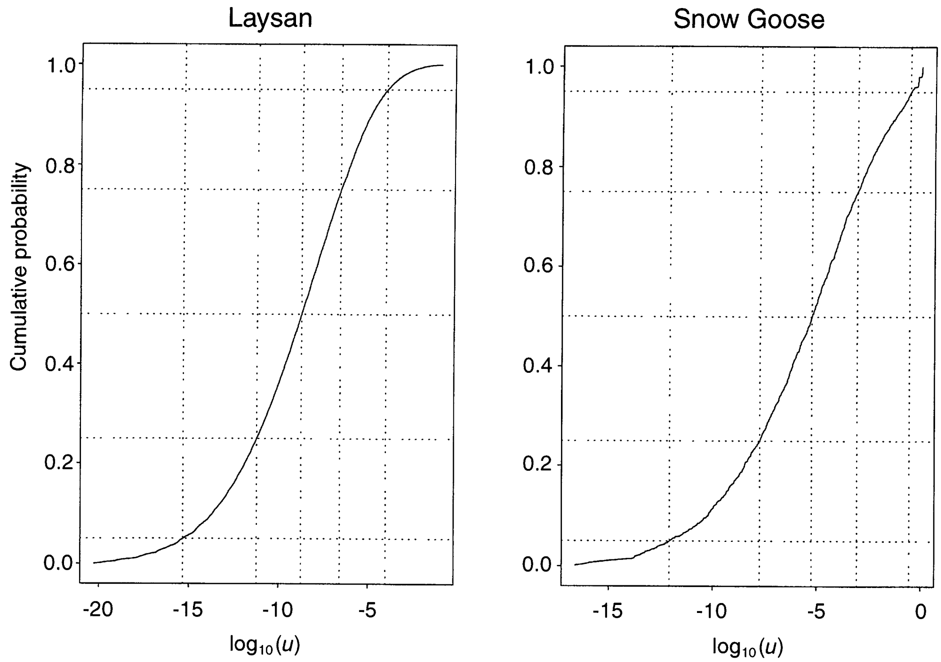

FIG. 2. The posterior cumulative distribution function (CDF) for the probability of early collapse given by Eq. 10, based upon data for the Laysan Finch and Snow Goose, respectively. The horizontal axes have been transformed by taking the logarithm to base 10.

terior seems to be a reasonable choice. However, Figs. 1 and 2 show the very wide range of plausible values for this probability. These figures demonstrate that such short data sets are not very informative about the population dynamics. One would also expect that integrals over the posterior density would be quite sensitive to the choice of prior density. Such sensitivity to choice of prior is a diagnostic for poor information about the quantity of interest. Therefore these results demonstrate the need for caution in performing assessments.

Concluding Remarks

The major conclusion of this work is that the great difference between maximum likelihood estimates of extinction probabilities and corresponding Bayes estimates should not be ignored in a proper assessment. However, the present work cannot be used to make a proper assessment because it neglects some important effects. It ignores errors and inaccuracies in the census data. These are known to cause severe bias in estimates and to lead to underestimation of the uncertainty in the resulting estimates (Ludwig and Walters 1981, Walters and Ludwig 1981). The calculations given here ignore the effects of occasional catastrophes upon extinction probabilities. This neglect is known to lead to underestimates of the probability of extinction (Ludwig 1996). The Gompertz model and the diffusion approximation were used here in order to simplify later calculations: they should not be used without an investigation of their accuracy (Ludwig 1996). Nevertheless, these defects do not affect the validity of the conclusion that point estimates are inappropriate for assessment of extinction probabilities.

A similar contrast between results from point estimates such as maximum likelihood and the Bayes estimates may be expected whenever the quantity of interest (in the present case, the probability of early population collapse) varies strongly and asymmetrically with one or more of the unknown parameters (Ellison 1996). Such quantities as the expected time to extinction or a minimum viable population may be expected to have similar properties, but I do not recommend their use for assessment (Ludwig 1996).

One might wonder whether a proper assessment is feasible in light of the difficulties noted above. I believe that all of the difficulties listed above can be overcome, but the calculations will require much more computing resources than the present work (D. Ludwig, *unpublished manuscript*) and very extensive data sets. This leaves open the issues of age- or stage-structured populations, spatial effects, genetics, and a host of species and ecosystem-specific considerations. We may never achieve a proper assessment that includes all that should be included. However, present attempts at management of natural populations typically make a systematic underestimation of hazards (Ludwig et al. 1993). We can strive to eliminate this unfortunate bias.

Acknowledgment

This work was supported by the Natural Science and Engineering Research Council of Canada under Grant Number A-9239. The routine DAW by W. J. Cody in the "specfun" library is available through the "netlib" on the internet.

Literature Cited

Abramowitz, M., and I. A. Stegun, editors. 1964. Handbook of mathematical functions. National Bureau of Standards, Washington, D.C., USA.

Berger, J. O. 1985. Statistical decision theory and Bayesian analysis. Springer Verlag, New York, New York, USA.

Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in conservation biology. Chapman and Hall, London, England.

Chernoff, H., and L. E. Moses. 1959. Elementary decision theory. (Reprinted by Dover [1986]). John Wiley and Sons, New York, New York, USA.

Cooch, E. G., and F. Cooke. 1991. Demographic changes in a Snow Goose population: biological and management implications. Pages 168–179 *in* C. M. Perrins, J.-D. Lebreton, and G. J. M. Hirons, editors. Bird population studies. Oxford University Press, Oxford, England.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. Ecological Monographs **6**:115–143.

Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications **6**:1036–1046.

Guttman, I., S. S. Wilks, and J. S. Hunter. 1982. Introductory engineering statistics. Third edition. John Wiley & Sons, New York, New York, USA.

Jeffreys, H. 1961. The theory of probability. Oxford University Press, London, England.

Kalbfleisch, J. D. 1986. Pseudolikelihood. Pages 324–327 *in* S. Kotz, N. L. Johnson, and C. B. Read, editors. Encyclopedia of statistical sciences. Volume 7. John Wiley & Sons, New York, New York, USA.

Lindley, D. V. 1971. Bayesian statistics, a Review. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

———. 1985. Making decisions. Second edition. Wiley, New York, New York, USA.

Ludwig, D. 1974. Stochastic population theories. Lecture Notes in Biomathematics 3. Springer-Verlag, Berlin, Germany.

———. 1996. The distribution of population survival times. American Naturalist **147**(4):506–526.

Ludwig, D., R. Hilborn, and C. J. Walters. 1993. Uncertainty, resource exploitation and conservation: lessons from history. Science **260**:17,36.

Ludwig, D., and C. J. Walters. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences **38**:711–720.

Morgan, M. G., and M. Henrion. 1990. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge, England.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. Numerical recipes: the art of scientific computing. Cambridge University Press, Cambridge, England.

Soulé, M. E., editor. 1987. Viable populations for conservation. Cambridge University Press, Cambridge, England.

Walters, C. J., and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences **38**: 704–710.

Winkler, R. L. 1972. Introduction to Bayesian inference and decision. Holt, Reinhart, and Winston, New York, New York, USA.

Wolfson, L. J., J. B. Kadane, and M. J. Small. 1996. Bayesian environmental policy decisions: two case studies. Ecological Applications **6**:1056–1066.

APPENDIX A

Evaluation of the Probability of Early Collapse

In this Appendix I describe the numerical method to evaluate Eq. 10.

Evaluation of u *if* $\beta_2 < 0$

If $\beta_2 < 0$ (as is usually the case), then evaluation of the integrals in Eq. 10 can be reduced to evaluation of the Dawson integral (Abramowitz and Stegun 1964)

$$D(x) = e^{-x^2} \int_0^x e^{t^2}\,dt. \tag{A.1}$$

Let

$$v = \sqrt{\frac{-\beta_2}{\sigma^2}}(x - \hat{x}), \tag{A.2}$$

with corresponding definitions for v', v_1 and v_2. Then

$$u(x, x_0, x_1, \beta_1, \beta_2, \sigma) = \frac{\int_v^{v_1} e^{t^2}\,dt}{\int_{v_0}^{v_1} e^{t^2}\,dt} = \frac{e^{v_1^2}D(v_1) - e^{v^2}D(v)}{e^{v_1^2}D(v_1) - e^{v_0^2}D(v_0)}. \tag{A.3}$$

The Dawson integral can be evaluated numerically by use of routine DAW by W. J. Cody in the "specfun" library (see *Acknowledgments*).

Evaluation of u *if* $\beta_2 > 0$

If $\beta_2 > 0$, then Eq. 10 may be evaluated in terms of the complementary error function (Abramowitz and Stegun 1964):

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\,dt. \tag{A.4}$$

If

$$v = \sqrt{\frac{\beta_2}{\sigma^2}}(x - \hat{x}), \tag{A.5}$$

and v_0 and v_1 have similar definitions, then

$$u(x, x_0, x_1, \beta_1, \beta_2, \sigma) = \frac{\int_v^{v_1} e^{-t^2}\,dt}{\int_{v_0}^{v_1} e^{-t^2}\,dt} = \frac{\mathrm{erfc}(v) - \mathrm{erfc}(v_1)}{\mathrm{erfc}(v_0) - \mathrm{erfc}(v_1)}. \tag{A.6}$$

The complementary error function is available as a built-in function for many FORTRAN or C compilers. Algorithms for its evaluation may also be obtained from Press et al. (1986). If the arguments of the integrals in Eq. 32 are large in absolute value, the numerator or denominator may vanish to machine precision, or the result may be inaccurate. To avoid this, the following approximations may be used:

$$u(x, x_0, x_1, \beta_1, \beta_2, \sigma) \approx \begin{cases} \exp(v_0^2 - v_2)\dfrac{v_0}{v}\dfrac{1 - \exp[2v(v - v_1)]}{1 - \exp[2v_0(v_0 - v_1)]} & \text{if } v_0 > 5 \\ \dfrac{1 - \exp[-2v_1(v - v_1)]}{1 - \exp[-2v_1(v_0 - v_1)]} & \text{if } v_1 < -5. \end{cases} \tag{A.7}$$

APPENDIX B

Integration over the Posterior Distribution

The integration will be carried out in the polar coordinates given in Eqs. 22–24. The quantities ρ, θ, and w defined there must be related to the original variables β_1, β_2, and σ. The appropriate relations are

$$\beta_1 = \gamma_1 - \gamma_2\bar{X} \tag{B.1}$$

$$\beta_2 = \gamma_2 \tag{B.2}$$

$$\gamma_1 = \hat{\gamma}_1 + \sqrt{\frac{(n - 2)S^2}{n}}z_1 \tag{B.3}$$

$$\gamma_2 = \hat{\gamma}_2 + \sqrt{\frac{(n - 2)S^2}{(n - 1)s^2}}z_2 \tag{B.4}$$

$$z_1 = \rho \cos\theta \tag{B.5}$$

$$z_2 = \rho \sin\theta \tag{B.6}$$

$$\sigma^2 = \frac{(n - 2)S^2(1 + \rho^2)}{2w}. \tag{B.7}$$

In order to calculate the Bayes risk of an early collapse of the population, it is necessary to integrate $u(x, x_0, x_1, \beta_1, \beta_2, \sigma)$ (given by Eqs. A.3 and A.6) over the posterior density. In order to carry out the numerical integration, the integral is replaced by a sum over appropriate regions in the ρ, θ, w space. The first step is to subdivide the ranges of ρ, θ, and w. For efficiency in the evaluation, I divide these coordinate axes into intervals that have equal probabilities. This procedure is analogous to a stratified sampling scheme: Monte Carlo trials and the usual sensitivity analysis often yield less accuracy per unit of calculation than such stratified sampling. In the present case, the posterior density is known explicitly and so design of an efficient stratified sampling scheme is easy.

The subdivisions of the axes are determined as follows: the points ρ_i corresponding to the midpoints of such a subdivision of the ρ axis satisfy

$$\int_0^{\rho_i} \frac{(n - 2)\rho\,d\rho}{(1 + \rho^2)^{n/2}} = \frac{i - 1/2}{N_\rho}, \qquad i = 1, \ldots, N_\rho, \tag{B.8}$$

where N_ρ is the number of subdivisions. Note that the choice of midpoints also avoids the difficulty due to an infinite interval of integration. The integral in Eq. B.8 may be evaluated to yield

$$\rho_i = \sqrt{[N_\rho/(N_\rho - i + 1/2)]^{2/(n-2)} - 1}.$$
$$i = 1, \ldots, N_\rho. \qquad \text{(B.9)}$$

In a similar fashion it follows that

$$\theta_j = 2\pi(j - 1/2)/N_\theta, \qquad j = 1, \ldots, N_\theta. \qquad \text{(B.10)}$$

The mesh is uniform in θ since in fact the posterior density is independent of θ. The analogous relation for w is

$$\frac{1}{\Gamma(n/2)} \int_0^{w_k} e^{-w} w^{n/2-1}\, dw = \frac{k - 1/2}{N_w},$$
$$k = 1, \ldots, N_w. \qquad \text{(B.11)}$$

This last equation cannot be solved explicitly, but w_k may be found by using a root-finding algorithm and a numerical approximation for the incomplete Gamma function (Abramowitz and Stegun 1964, Press et al. 1986). With these definitions, the following approximation results:

$$\int_0^\infty \int_0^{2\pi} \int_0^\infty \phi(\rho, \theta, w) \frac{(n-2)\rho}{(1+\rho^2)^{n/2}} \cdot \frac{1}{\Gamma(n/2)} e^{-w} w^{n/2-1} \cdot \frac{1}{2\pi}\, dw\, d\theta\, d\rho$$
$$\approx \frac{1}{N_\rho N_\theta N_w} \sum_{i=1}^{N_\rho} \sum_{j=1}^{N_\theta} \sum_{k=1}^{N_w} \phi(\rho_i, \theta_j, w_k). \qquad \text{(B.12)}$$

The function ϕ in Eq. B.12 is a shorthand for $u(x, x_0, x_1, \beta_1, \beta_2, \sigma)$ with the substitutions of Eqs. B.1–B.7.