-

Stochastic population models

Steinar Engen
Department of Mathematical Science
Norwegian University of Science and Technology
7491 Trondheim Norway

Revised: November 29, 2011

Contents

PREFACE

These are notes that will serve as a text for the course in 'Stochastic population models' at the Department of Mathematical Sciences at NTNU, Norway. The department initiated a new bachelor program in biomathematics in 2003 and this is a course planned for students the third year choosing the direction of 'Modelling in biology'. However, this course will also be quite useful for students going into 'Statistics in medicine' and 'Theoretical biology' as well as master student in biology with required mathematical background since basic knowledge on how to do stochastic modelling of biological populations or systems of populations is an important part of quantitative studies in life science in general.

The mathematical and statistical background required is that obtained during the first two years of the program. The first year these students has two basic courses in mathematical analysis, one in probability, one in statistics and one course in biological computations which focuses on using computer software (R) to do statistics and stochastic simulations. The second year they have a course in applied statistics, an introductory course in stochastic processes as well as one in mathematical genetics. In addition to courses given by the department these students also have courses given by the department of biology and the medical faculty.

These lectures focus on modelling of population dynamics, mostly dealing with one single species. Stochastic modelling of systems of two species are rare in the literature and leads often to rather difficult mathematical problems, although there is a large literature on deterministic predator prey models an competition systems. Two species systems are therefore not dealt with in much details. On the other hand, modelling of communities with many species has a long scientific history and there is a growing interest in stochastic models. Such models also have some analogs in population genetics. Some interesting results for communities can be dealt with in a rather simple way basing on results for a single species. The reason why this may be simpler than two species systems is that the interactions between the species now can be summarized in stochastic terms instead of going into a detailed description of how each particular species interact with all the others, which would tend to end up in a description with too many parameters to be of practical interest.

This is mainly meant to cover a basic course in stochastic population models and is not a course in statistics. The models and results presented here are, however, important for doing correct statistical analysis of population data. Some simple examples of statistical methods are given, but only in cases where the population size is estimated without error. Statistical inference based on data with sampling errors, which is the most common type of data, can be done using methods constructed for this purpose like state-space models, Kalman filtering, or the bayesian approach analyzed by Markov chain Monte Carlo methods. These methods are dealt with in general courses in statistics.

Steinar Engen
Department of Mathemtical Sciences
The Norwegian University of Science and Technology
Trondheim, Norway

Chapter 1

Populations without density-regulation

1.1 The deterministic product model

In this chapter we shall deal with populations reproducing once a year which is a realistic assumption for a large number of organisms. However, populations reproducing continuously in time (for example humans or species in the tropics) may be studied by performing a sensus once a year leading to the same kind of data. The basic unit in population dynamics is the individual. Each individual contribute to the population density and to the change in population size from one generation to the next. The basic vital rates determining population changes are the individuals survival into the next year and their reproduction. If nothing is mentioned about the sex of individuals we will always deal only with the female segment of the population. For large populations, however, we may often obtain a realistic description of the dynamics without going into details on the individuals vital rates. In a large populations the differences between the individuals may not be important by the law of large numbers, only their mean values across the population.

If the vital rates are not affected by the density of individuals we say that the population is not density-regulated. For large populations living in a stable environment, a realistic description may be that the population size the next year simply is given by a multiplication of the size the previous year by a constant factor

$$N_{t+1} = \lambda N_t$$

where N_t denotes the population size at time t and λ is the constant determining the growth of the population. The population is increasing, constant or decreasing according to the value of λ being > 1, one or < 1. Starting with time $t = 0$ we then find simply by recursion

$$N_t = \lambda^t N_0 = N_0 e^{rt}$$

where $r = \ln \lambda$ is called the population growth rate. The population is increasing, constant or decreasing for $r > 0$, $r = 0$ or $r < 0$, respectively. If $r < 0$ the population will eventually go extinct. If we consider the population to be extinct as the population size reaches size one, the time to extinction, say T, is for negative growth rates determined by $N_0 e^{rT} = 1$ giving $T = -\ln(N_0)/r$. For example, a population of 1000 individuals with $\lambda = 0.99$ and $r = -0.01005...$ will go extinct after 687 years, while if $\lambda = 0.9$ it is extinct after 66 years.

Since effects operating on populations often are modelled by multiplications (multiplicative effects) we often get simpler mathematical relations by working on logarithmic scale. Here we will always use natural logarithms. Writing $X_t = \ln(N_t)$ for the log population size, the above multiplicative model takes the simple form $X_{t+1} = X_t + r$ giving

$$X_t = X_0 + rt$$

so that log population size is a straight line with slope r when plotted against time.

Although this deterministic description is not realistic for real populations it is a good approximation for large populations in a very stable environment, for example growing of plankton populations in a laboratory up to the time when the population reaches a density large enough for density regulation to operate.

1.2 Environmental effects on population growth

The factor λ is determined by the mean survival and reproduction rates of the individuals. In natural populations these rates will almost always be affected by a number physical or biological factors. Which factors that are important, and how large effect they have, varies a lot between species. For example, snow depth and temperature during winter may have only a little effect on a brown bear, a larger effect on small rodents, and perhaps even a larger effect on some bird species. In other words, some species may be almost unaffected by the environment, while others will show large changes in population size generated by fluctuations in the environment. Clearly, temporal fluctuations in the environment are also very different from location to location. Typically, fluctuation in the tropics are smaller than fluctuations in areas with large seasonal effects. Even death rates in human populations may be affected by the environment. For example, different contagious diseases that may affect the death rate, especially among old people, are not equally common each year and varies geographically.

Even if we often can isolate some few factors as the major factors affecting the vital rates of a population it is generally a difficult task. Such studies are, however, an important part of population dynamics, for example in studies of which effects climate changes is expected to have on natural populations. However, we can generally write symbolically the environment affecting a population as an environmental vector $\mathbf{z} = (z_1, z_2, \ldots)$ where each component is some factor, physical or biological, that may affect the vital rates. We may obtain large insight into the dynamics of populations without studying each component of $\mathbf{z}$ separately, but rather just plug into our model that the factor λ in practice is some function of the environmental vector. The environmental vector is typically stochastic, varying between years in a more or less unpredictable way. Writing Λ_t for the factor operating at time t which is formally some function of $\mathbf{z}$, we can now forget the environmental vector and concentrate on just modelling the sequence Λ_t as a time series, keeping in mind that its properties are generated by fluctuations in the environments.

In analogy with the deterministic model we write $S_t = \ln \Lambda_t$ giving

$$X_t = X_0 + \sum_{u=0}^{t-1} S_u.$$

In many cases it may be realistic to assume that the sequence of stochastic rates S_t is a sequence of independent identically distributed random variables. In other cases, however, the sequence may be autocorrelated, either due to autocorrelations in the underlying vector $\mathbf{z}$ or some properties of the populations like age structure or migration.

1.3 The lognormal distribution

Since multiplicative effects are common in biological systems, the above multiplicative model being an example, the lognormal distribution has an important role in biology in general. As we have seen, performing a log transformation leads to additive models. When stochastic variable are added we know, through different versions of the central limit theorem, that the sums are approximately normally distributed. Transforming back to the original scale we then obtain the lognormal distribution.
Formally, let Y be normally distributed with mean μ and variance σ^2. Shortly we then write that Y is $N(\mu, \sigma^2)$. Let Y be the log transform of a variable V, that is $Y = \ln V$ and $V = \exp(Y)$. We then say that V is lognormally distributed with parameters μ and σ^2 and write shortly , V is $LN(\mu, \sigma^2)$.
The probability density of Y is the Gaussian curve

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

on the real axis. Applying the transformation formula for stochastic variables we find that the distribution of V is the corresponding lognormal distribution shown in Fig.1.1

$$f_V(v) = \frac{1}{\sqrt{2\pi}\sigma v} e^{-\frac{(\ln v-\mu)^2}{2\sigma^2}}$$

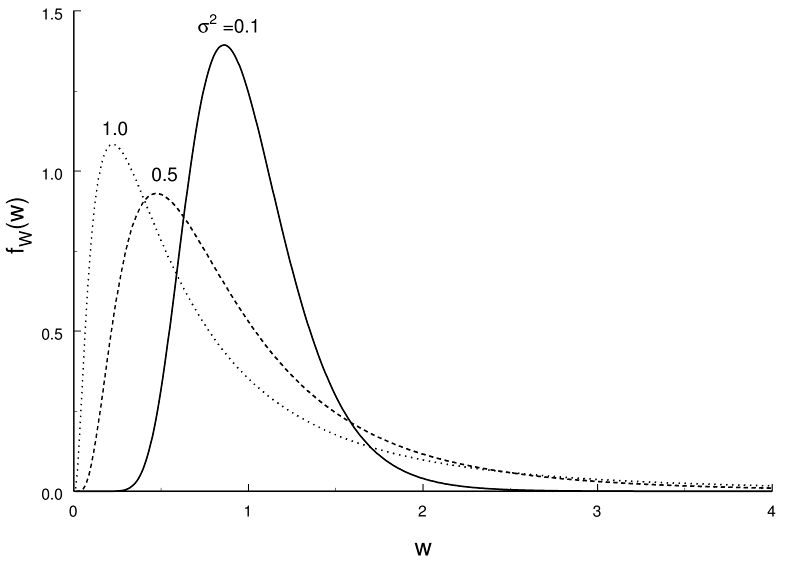

Figure 1.1: The lognormal distribution with mean 1, that is $\mu = -\sigma^2/2$, for three values of σ^2.

for $v > 0$. Notice that since $aY + b$ is $N(a\mu + b, a^2\sigma^2)$, it follows that $e^{aY+b} = e^b V^a$ is $LN(a\mu + b, a^2\sigma^2)$, so that cV^a is $LN(a\mu + \ln c, a^2\sigma^2)$. In particular we see by inserting $a = 1$ that cV is $LN(\mu + \ln c, \sigma^2)$, that is, the variance parameter σ^2 is not affected by a change of scale.
The mean and variance of the lognormal is most easily found from the moment generating function of Y which is known to be

$$M_Y(t) = \mathrm{E}e^{Yt} = e^{\mu t + \sigma^2 t^2/2}.$$

Writing ν and τ^2 for the mean and variance of the lognormal distribution we then find

$$\nu = \mathrm{E}V = \mathrm{E}e^Y = e^{\mu + \sigma^2/2},$$

where we have plugged in $t = 1$ in the moment generating function. In the same way we find

$$\mathrm{E}V^2 = \mathrm{E}e^{2Y} = e^{2\mu + 2\sigma^2}$$

and finally the variance

$$\tau^2 = \mathrm{E}V^2 - (\mathrm{E}V)^2 = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1).$$

Hence, the squared coefficient of variation for the lognormal distribution is

$$CV^2 = \tau^2/\nu^2 = e^{\sigma^2} - 1.$$

1.4 The stochastic growth rate and environmental variance

Let us first assume that the sequence Λ_t and S_t are both sequences of independent variables with constant mean and variance. Then, conditioning on the initial population size $N_0 = n_0$ we find that the expected population size after time t is

$$\mathrm{E}N_t = n_0\lambda^t = n_0 e^{rt}$$

where we now have generalized interpretation of λ to be the the expectation $\mathrm{E}\Lambda$ (we omit the subscript for Λ since they all have the same distribution). The expected population size then grow exponentially with a rate $r = \ln\lambda = \ln(\mathrm{E}\Lambda)$, which we now shall call the *deterministic growth rate*.

On the log scale we can express the mean slope of the curve on the interval from 0 to t as

$$\frac{1}{t}(X_t - X_0) = \frac{1}{t}\sum_{u=0}^{t-1} S_u = \bar{S}.$$

Hence, the mean and variance of this slope is $s = \mathrm{E}S$ and σ_s^2/t, respectively, where $\sigma_s^2 = \mathrm{var}(S)$. As t tends to infinity the slope then approaches the constant s which is called the *stochastic growth rate* of the population.

Now, let us compare the stochastic growth rate s with the deterministic growth rate $r = \ln\lambda$ for the expected population size. Let us first assume that the S_t are $N(s, \sigma_s^2)$. Then Λ is $LN(s, \sigma_s^2)$. From the properties of the lognormal distribution it then follows that

$$s = r - \sigma_s^2/2.$$

In the same model we find

$$\mathrm{var}(N_{t+1}|N_t = n) = n^2\sigma_e^2$$

where $\sigma_e^2 = \mathrm{var}(\Lambda)$ is called the *environmental variance*. For the lognormal model we see, again using properties of the lognormal distribution, that

$$\sigma_s^2 = \ln(1 + \sigma_e^2/\lambda^2).$$

More generally we can use the Taylor expansion

$$\ln\Lambda = \ln\lambda + \ln\left[1 + \left(\frac{\Lambda - \lambda}{\lambda}\right)\right] = r + \left(\frac{\Lambda - \lambda}{\lambda}\right) - \frac{1}{2}\left(\frac{\Lambda - \lambda}{\lambda}\right)^2 + \ldots.$$

This leads to (exercise 5) $\sigma_s^2 \approx \sigma_e^2/\lambda^2$ and $s \approx r - \frac{1}{2}\sigma_e^2/\lambda^2$ provided that the fluctuations in $(\Lambda - \lambda)/\lambda$ are small. Further, if $\lambda \approx 1$, the two variances are approximately equal so that $s \approx r - \sigma_e^2/2$.

The relation between the stochastic and the deterministic growth rate has some interesting and rather surprising consequences. As an example consider a population with deterministic growth rate $r = 0.01$. The expected population size will then tend to infinity as t approaches infinity. If the environmental variance is $\sigma_e^2 = 0.04$, then the stochastic growth rate is $s = -0.01$. Hence, as time approaches infinity the slope approaches the constant value -0.01 which means that the population is certain to go extinct although its expected value approaches infinity. In Fig.1.2 we show simulations of this example. The result is definitely not just a mathematical artifact but a highly real effect of stochasticity. If the stochasticity of the environment increases without affecting the mean vital rates so that $r = \ln(\mathrm{E}\Lambda)$ is kept constant, then the stochastic growth rate will decrease while the expected population sizes are unaffected. In order to understand this result intuitively we have to look deeper into the lognormal distribution. This is a very skew distribution and as t increases the skewness also increases towards infinity. For large values of t it is therefore possible that there is some very small probability that the population is extremely large, actually large enough to give a large expected value. The whole probability mass, however, except this very small proportion approaching zero, may still be concentrated at smaller values and represent extinction with probability 1.

The above approximation have been derived using the first terms of the Taylor expansion. There will in general be a decrease in the stochastic growth rate as the stochasticity increases. A more accurate approximation than the one derived here from the normal distribution is given in 1.8.1.

1.5 Estimation and prediction in the multiplicative model

We consider the situation where the population size is known at $n + 1$ subsequent points of time $t = 0, 1, 2, \ldots, t_n$. Let the observed values of log population size be $X_0, X_1, \ldots, X_n$ and assume that the S_i are approximately

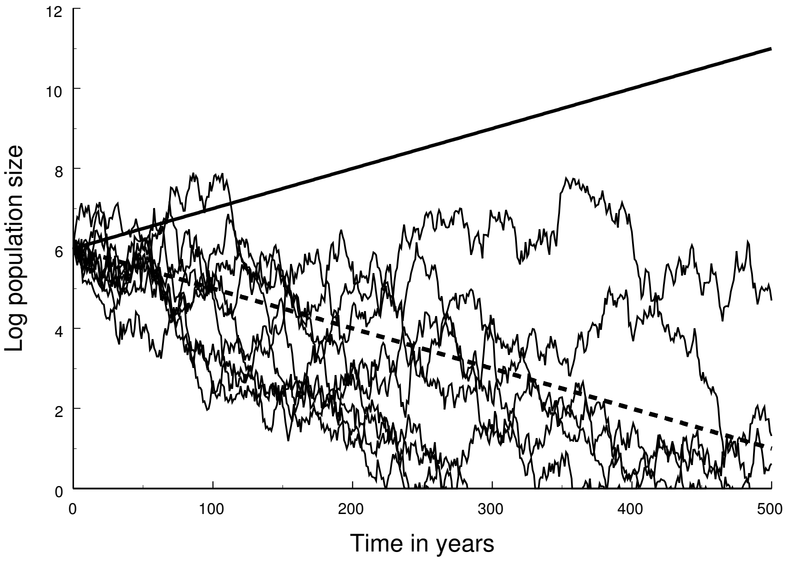

Figure 1.2: Simulation of 10 sample paths using the above stochastic model with $r = 0.01$ and $\sigma_e^2 = 0.04$. The solid straight line shows the deterministic growth, while the dotted line is the mean stochastic growth.

normally distributed. The process X_i is then a random walk with normally distributed increments and the differences $D_i = X_i - X_{i-1}$, for $i = 1, 2, \ldots n$ are independent normal variables with mean s and variance $\sigma_s^2 \approx \sigma_e^2$. Hence, the maximum likelihood estimate of s is the mean value

$$\hat{s} = \frac{1}{n}\sum_{i=1}^{n} D_i = (X_n - X_0)/n.$$

Hence, we se that for this model the estimator for the stochastic growth rate depends only on the first and last observation, the others being redundant. According to standard normal theory this estimator is $N(s, \sigma_s^2/n)$. Further, the unbiased and sufficient estimator for the environmental variance is

$$\hat{\sigma_s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \hat{s})^2$$

which is independent of s^* and distribution given by the fact that $\hat{\sigma_s}^2(n-1)/\sigma_s^2$ is χ^2-distributed with $n-1$ degrees of freedom. We can further find confidence intervals for s and σ_s^2 based on Student's T-distribution and the χ^2-distribution exactly as in the analysis of a single sample from a normal distribution with unknown mean and variance.

These results can be utilized to find prediction intervals for future population sizes provided that the population size is large enough for the possibility of extinction to be ignored. Conditioned on the last observation X_n the population size X_{n+m} at a future time $n + m$, which can be written on the form $X_n + \sum_{i=n+1}^{n+m} D_i$, is $N(X_n + ms, m\sigma_s^2)$. We can find a prediction interval for X_{n+m} by first observing that $X_{n+m} - X_n - m\hat{s}$ is normally distributed with zero mean and variance $m\sigma_s^2 + m^2\sigma_s^2/n = \sigma_s^2(m + m^2/n)$. It follows from Student-Fisher's well known result that

$$T_{n-1} = \frac{X_{n+m} - X_n - m\hat{s}}{\hat{\sigma_s}\sqrt{m + m^2/n}}$$

has Student's T-distribution with $n-1$ degrees of freedom so that

$$P(-t_{n-1,\alpha/2} < T_{n-1} < t_{n-1,\alpha/2}) = 1 - \alpha$$

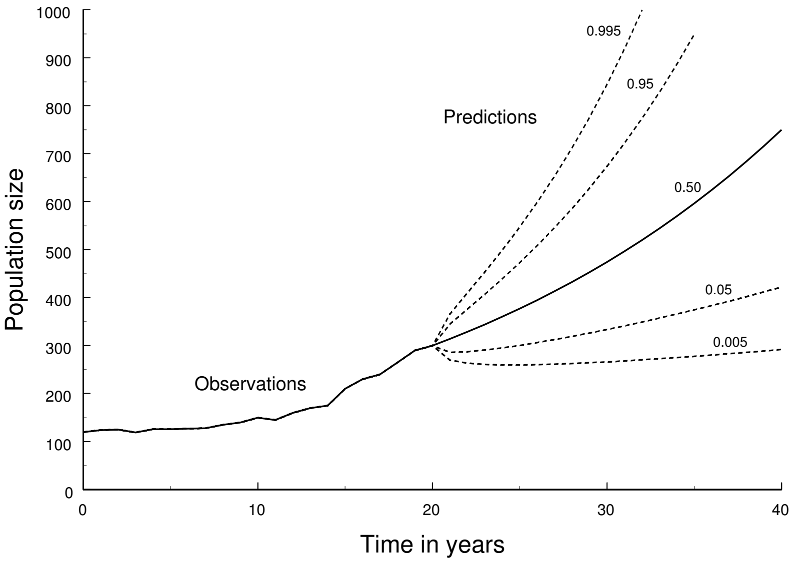

Figure 1.3: Time series observations of a population of Sea Eagles over 21 years together with 20 prediction intervals (90 and 99 %) and predicted median 20 years ahead.

where $t_{n-1,\alpha/2}$ denotes the $\alpha/2$-quantile of the T-distribution with $n-1$ degrees of freedom. Plugging in the expression for T_{n-1} and solving the inequalities with respect to the future observation X_{n-m}, we finally find the prediction interval for the log population size

$$P(X_n+m\hat{s}-t_{n-1,\alpha/2}\hat{\sigma_s}\sqrt{m+m^2/n} < X_{n+m} < X_n+m\hat{s}+t_{n-1,\alpha/2}\hat{\sigma_s}\sqrt{m+m^2/n}) = 1-\alpha.$$

Applying the exponential function over the whole inequality then yields the corresponding prediction interval for the population size.
In practical applications it is informative to plot the observed population size and the interval limits as functions of time for different values of α. An example of this is shown in Fig.1.3.

1.6 Demographic stochasticity

We now return to a more thorough analysis of the multiplicative factors Λ_t. When populations are small, the stochastic fluctuations in survival and fecundity between individuals within years can not be ignored, and we now proceed to analyze the consequences of this kind of stochasticity which is called demographic stochasticity.

The changes in population size from one generation to the next is determined by the survival or death of each individual, as well as the number of surviving offspring each individual contribute with into the next generation. We shall later deal with age-structured models, assuming that the vital rates vary between age-classes. Here we deal with the simpler situation where individuals reach the adult state during a year and the population is sensused just before reproduction. All individuals are then adults with the same mean vital rates. Notice however, that this is a model with overlapping generation since each individual may have an adult survival close to one and have a lifetime of many generations.

For a given population size N at one generation there are N contributions from the individuals adding up to give the population size $N+\Delta N$ in the next generation. The contribution from one particular individual is the number of offspring it produces that survive into the next generation plus 1 if the individual itself survives. Writing $w_1, w_2, \ldots, w_N$ for these contributions, which is the individual fitness for the individuals in the population, we have

$$N + \Delta N = \sum_{i=1}^{N} w_i = N\mathrm{E}w + \sum_{i=1}^{N} d_i,$$

where $\mathrm{E}w$ is the mean of the w_i and the $d_i = w_i - \mathrm{E}w$ are stochastic variables with zero means. Since $N + \Delta N = \Lambda N$ we see that $\Lambda = \frac{1}{N}\sum w_i = \bar{w}$ and $\mathrm{E}w = \lambda$. In older literature on birth and death processes one usually assumes that the contributions within a season are stochastically independent for a given population size. This is a realistic assumption if there is no environmental vector $\mathbf{z}$ generating between years fluctuations in the mean contributions, that is, if we assume $\mathbf{z}$ to be a constant. In such models the variance $\mathrm{var}(w_i) = \mathrm{var}(d_i) = \sigma_d^2$ is called the *demographic variance* of the process. In a density regulated population the parameter $\mathrm{E}w$ as well as σ_d^2 may depend on the population size N.

Using the assumption that all N contributions are stochastically independent for a constant $\mathbf{z}$, that is, there is no common factor acting on the D_i, we find

$$\mathrm{var}(N + \Delta N|N) = \mathrm{var}(\Delta N|N) = N\sigma_d^2$$

by the law of the variance of a sum of independent variables. For the stochastic factor $\Lambda(N)$ we find

$$\mathrm{var}[\Lambda] = \mathrm{var}[\Delta N/N] = \sigma_d^2/N.$$

As an example let the w_i be Poisson distributed with means $\mathrm{E}w = \lambda$. Since the variance of the Poisson distribution is the same as the mean, that is $\sigma_d^2 = \mathrm{E}w = \lambda$, we also have $\mathrm{var}(\Delta N|N) = N\mathrm{E}w = N\lambda$. Notice in particular that for a population fluctuating around some stable equilibrium σ_d^2 must be close to 1 since the mean contribution from the individuals (λ) are then close to 1. We emphasize that this statement is not valid in general and must be considered a property of the Poisson model. However, the Poisson model gives us some idea of the order of magnitude we should expect to find for the demographic variance of real populations.

Notice also that for small fluctuations we have the approximation (see section 1.4 and exercise 5)

$$\mathrm{var}(\Delta \ln N) = \mathrm{var}(\Delta X) = \mathrm{var}(\ln \Lambda) \approx \lambda^{-2}\sigma_d^2/N = \lambda^{-2}\sigma_d^2 e^{-X}.$$

When working on the log scale one commonly includes the factor λ^{-2} in the definition of the demographic variance. In chapter 4 we shall analyze age-structured models and study the dynamics on the log scale. To obtain variance formulas that are consistent with most formulas appearing in the literature we shall then include the factor λ^{-2}. However, when dealing with decomposition of stochasticity in simple models as well as age-structured models it is mathematically simpler to operate on the absolute scale and not including this factor, which is what we do in the present chapter.

1.7 Demographic and environmental stochasticity acting together

If there are fluctuations in the stochastic vector $\mathbf{z}$ between years the individual contributions a given year are no longer independent. We then decompose the contributions into variance components writing

$$w_i = \mathrm{E}w + e + d_i$$

where $e = \mathrm{E}(w|\mathbf{z}) - \mathrm{E}w$ and $d_i = w_i - \mathrm{E}(w|\mathbf{z})$. Then $e, d_1, d_2, ..., d_x$ are stochastic variables with zero means and $\mathrm{cov}(e, d_i) = 0$ (exercise 7).

Fig.1.4 shows distributions of individual fitness w to the next generation of two bird species. We see that there is a considerable variance within each year, corresponding to a between individual variation in d_i. However, the histograms are also very different between years. In 1988 the Song Sparrow has very small contributions compared to the best year which is 1990. Mathematically this means that the component e is small in 1988 and large in 1990. It seems to be a smaller between years variation in e for the Great Tits.

The d_i are the demographic components varying between the individuals of the population a given year, whereas e is an environmental effect which is common for all individuals a given year, but vary stochastically between years. Consequently, applying the general formula for the variance of a sum of correlated variables, we obtain (exercise 9)

$$\mathrm{var}(\Delta N|N) = \mathrm{var}(\sum_{i=1}^{N} w_i|N) = N(\sigma_d^2 - \tau) + N^2(\sigma_e^2 + \tau),$$

where $\sigma_d^2 = \mathrm{var}(d_i)$, $\sigma_e^2 = \mathrm{var}(e)$ and $\tau = \mathrm{cov}(d_i, d_j)$, $i \neq j$. The parameter σ_e^2 is now a generalization of the previously defined concept of *environmental*

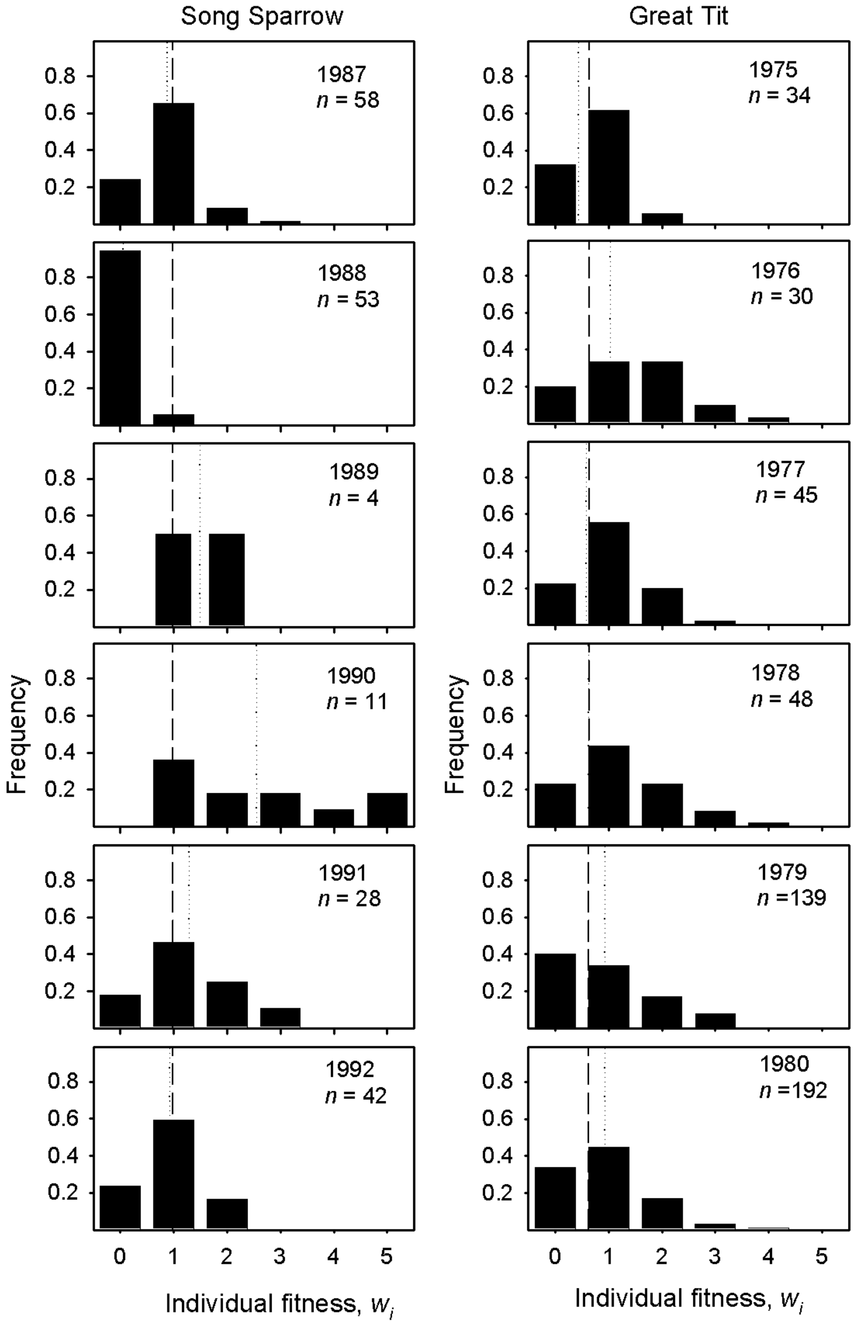

Figure 1.4: Annual variation in the distribution of contributions w to the next generation for two passerine species, the Song Sparrow Melospiza melodia on Mandarte island and the Great Tit in Wytham Wood. The dashed line indicates the mean values across all years and the dotted line the mean contribution a single year.

variance while τ is called the *demographic covariance.* For the multiplicative factor Λ dealt with in the previous section we then have accordingly

$$\mathrm{var}[\Lambda|N] = (\sigma_e^2 + \tau) + (\sigma_d^2 - \tau)/N.$$

Assuming small fluctuations we also have the relation

$$\mathrm{var}(\Delta X|N) \approx \lambda^{-2}(\sigma_e^2 + \tau) + \lambda^{-2}(\sigma_d^2 - \tau)/N.$$

Hence, we see that the demographic variance can be ignored when the population is large.

Notice that the environmental component e expresses how much the mean contributions a given year deviate from the mean contributions over all years. The demographic effects d_i express how each particular individual's contribution deviates from the mean contribution the same year. An alternative way of expressing the parameters defining the stochasticity is (exercise 10)

$$\begin{aligned}\sigma_d^2 &= \mathrm{E}[\mathrm{var}(w_i|\mathbf{z})]\\ \sigma_e^2 &= \mathrm{var}[\mathrm{E}(w_i|\mathbf{z})]\\ \tau &= \mathrm{E}[\mathrm{cov}(w_i, w_j|\mathbf{z})].\end{aligned}$$

The demographic covariance τ is created by interactions between the individuals, such as intra-specific competition. It is defined as the covariance between any two of the demographic contributions d_i. Although these covariances are likely to be different from zero, they are not usually taken into account in population models. This may, however, be seen as a redefinition of demographic and environmental variances as $\sigma_d^2 - \tau$ and $\sigma_e^2 + \tau$ respectively. This latter definition, for which the demographic and environmental variance are the coefficients of N and N^2 in the expression for $\mathrm{var}(\Delta N)$, is equivalent to defining the demographic variance as $\mathrm{var}(w_i) - \mathrm{cov}(w_i, w_j)$ and the environmental variance as $\mathrm{cov}(w_i, w_j)$, where $i \neq j$ (exercise 11). In general, theoretical models as well as empirical findings show that the demographic variance, the environmental variance and the demographic covariance may all depend on N (in density-regulated populations).

1.8 * Quantifications of the effects of stochasticity

1.8.1 Reduction in growth due to stochasticity

We have seen that increasing stochasticity reduces the stochastic growth rate of a population. In general it follows from Jensen's inequality that $\mathrm{E}\ln\Lambda < \ln(E\ln\Lambda)$ since the logarithm is a convex function. In particular, if Λ is lognormally distributed this reduction is half the variance of $\ln\Lambda$ which is approximately half the environmental variance.

We can explore this reduction in more detail by looking at the cumulant generating function of the stochastic variable $\ln\Lambda$ which is defined as

$$K(v) = \ln \mathrm{E}e^{v\ln\Lambda} = k_1 v + \frac{1}{2!}k_2 v^2 + \frac{1}{3!}k_3 v^3 + \ldots,$$

where k_i is the i'th cumulant of $\ln\Lambda$. Observing that $K(1) = \ln \mathrm{E}\Lambda = r$ and $k_1 = \mathrm{E}\ln\Lambda = s$ we find, by inserting $v = 1$ in the definition, that the reduction in the growth due to stochasticity in general is given by relation

$$s = r - (\frac{1}{2!}k_2 + \frac{1}{3!}k_3 + \ldots)$$

If $\ln\Lambda$ is normally distributed, then $k_j = 0$ for $j \geq 3$, giving the reduction of $\frac{1}{2}k_2 = \frac{1}{2}\sigma^2$, since k_2 is the variance of the variable $\ln\Lambda$. Using the first 4 cumulants we find

$$s \approx r - (\frac{1}{2}\sigma^2 + \frac{1}{6}\gamma_3\sigma^3 + \frac{1}{24}\sigma^4\gamma_4)$$

where $\gamma_3 = k_3/k_2^{3/2}$ and $\gamma_4 = k_4/k_2^2$ is the skewness and curtosis of $\ln\Lambda$, respectively.

1.8.2 Stochastic Alle-effect

As N gets smaller, the variance of $\ln\Lambda = \ln[\mathrm{E}\omega + e + \bar{d}] = \ln[\lambda + e + \bar{d}]$ increases. Defining generally the stochastic growth rate at population size N as $s(N) = \mathrm{E}(\ln\Lambda|N) = \mathrm{E}(S|N)$ we find using the normal approximation for $\Delta\ln N$ (exercise 18)

$$s(N) \approx r - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2.$$

In population biology, the term Alle-effect is used for different kinds of effects that makes it difficult for populations at small densities to reproduce for example due

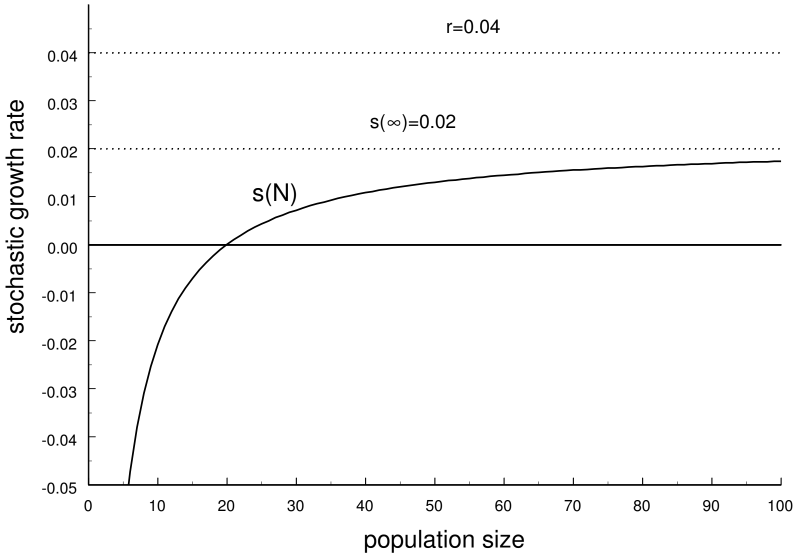

Figure 1.5: The stochastic growth rate as function of the population size. Parameter values are $r = \ln \mathrm{E}w = \ln \lambda = 0.04$, $\sigma^2 = 0.04$, and $\sigma_d^2 = 1$, giving $N^* \approx 20$. The dotted lines show the value of r and $s(\infty)$.

to the difficulty of finding a mate. Such effects may even lead to negative growth-rates and eventual extinction when the population size passes below a certain threshold (an unstable equilibrium point). One interesting effect of demographic stochasticity is that this stochasticity alone may produce a kind of Allee-effect, a stochastic Allee-effect, since we may have an unstable equilibrium point N^* so that $s(N) < 0$ for $N < N^*$ and $s(N) > 0$ for $N > N^*$. As an example, assume that $\lambda + e$ is lognormally distributed with $\mathrm{var}[\ln(\lambda + e)] = \sigma^2$, but add the assumption that $\bar{d}$ is normal with variance σ_d^2/N, which is approximately correct by the central limit theorem.

Fig.1.5 shows numerical values of $s(N)$ for this model, with $r = \ln \mathrm{E}w = 0.04$, $\sigma_e^2 = 0.04$, and $\sigma_d^2 = 1$, giving $N^* \approx 20$.

In Fig.1.6 we show 12 simulations of this process with $N_0 = 50$, from which we can actually get an impression of this stochastic Allee-effect with an unstable equilibrium point at $N^* \approx 20$. We also see that the paths tend to be under the

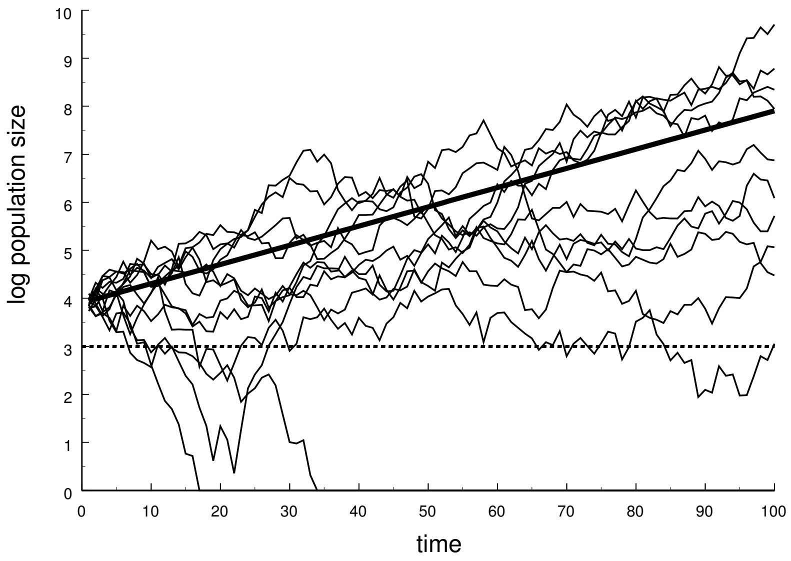

Figure 1.6: Simulations of the process described in Fig.1.5 with initial population size $N_0 = 50$ giving $\ln N_0 \approx 4$.

corresponding deterministic process with growth rate $r = 0.04$, showing that the reduction in the growth rate due to stochasticity is not just an artifact of the model.

1.8.3 Temporal correlations in the environmental noise

We close this section by considering the effect of correlations between the $\ln \Lambda_t$ expressed by the autocorrelation function

$$\rho(h) = \rho(-h) = \mathrm{corr}(\ln \Lambda_t, \ln \Lambda_{t+h}),$$

ignoring demographic stochasticity and assuming that $\ln \Lambda_t$ is a stationary process with $\sum_{i=1}^{\infty} i\rho(i) < \infty$. Such correlations may be generated by time delayed effects on survival and reproduction, by age structure, or by dependence between components of the environmental vectors $\mathbf{z}_t$ and $\mathbf{z}_{t+h}$. From the general formula for the variance of a sum we find

$$\mathrm{var}(\ln N_t|N_0) = t\sigma^2 + 2(t-1)\rho(1)\sigma^2 + 2(t-1)\rho(2)\sigma^2 + \ldots + 2\rho(t-1)\sigma^2,$$

which can be written as

$$\mathrm{var}(\ln N_t|N_0) = \sigma^2 t \sum_{i=-(t-1)}^{t-1} \rho(i) - 2\sigma^2 \sum_{i=1}^{t-1} i\rho(i).$$

Hence, as t approaches infinity we have

$$\frac{\mathrm{var}(\ln N_t|N_0)}{t} \to \sigma^2 \sum_{-\infty}^{\infty} \rho(i).$$

Notice that the autocorrelations neither affect $r = \ln \mathrm{E}\Lambda$ nor $s = \mathrm{E}\ln \Lambda$, but it may have a large effect on $\mathrm{E}N_t \approx \exp[ts + \frac{1}{2}\mathrm{var}(\ln N_t)]$ through the effect of increasing $\mathrm{var}(\ln N_t)$. In order to analyze how the expected population size $\mathrm{E}N_t$ changes with t we may define the growth-rate for the expected population size as

$$u = \lim_{t\to\infty} \frac{1}{t} \ln(\mathrm{E}N_t/N_0),$$

which expressed by the sequence Λ_t is

$$u = \lim_{t\to\infty} \frac{1}{t} \ln \mathrm{E}(\Lambda_1\Lambda_2, \ldots \Lambda_t).$$

If the Λ_t are independent and identically distributed we have $\mathrm{E}(\Lambda_1\Lambda_2 \ldots \Lambda_t) = (\mathrm{E}\Lambda)^t$, giving $u = r$. In the case of a sequence with autocorrelations, however,

u and r are different. In particular, if the $\ln \Lambda_t$ are multinormally distributed variables we find (exercise 19)

$$s = r - \frac{1}{2}\sigma^2 = u - \frac{1}{2}\sigma^2 \sum_{-\infty}^{\infty} \rho(i)$$

1.9 Fitness in a stochastic environment

1.9.1 The stochastic growth rate as a measure of fitness

Although there are some different definitions of fitness, the concept of fitness used in biology refers usually to deterministic models. In a stochastic environment, the stochastic growth rate is the fitness measure that is most informative in predicting the fate of different genotypes. For simplicity we assume no density regulation and consider a haploid organism with two genotypes that differ on one locus. Let $N_t(A)$ and $N_t(B)$ denote the number of individuals of type A and B at time t and write $Q_t = N_t(A)/[N_t(A)+N_t(B)]$ for the frequency of type A. For a given initial frequency Q_0 we shall see that the probability that Q_t is larger than any proportion p approaches one as t approaches infinity if the stochastic growth rate of the subpopulation of type A individuals is greater than that for type B. This can be expressed by the log of the population sizes,

$$P(Q_t > p|Q_0) = P[N_t(A)(1-p) > N_t(B)p] = P[X_t(A)-X_t(B) > \ln(p)-\ln(1-p)]$$

where $X_t = \ln N_t$ for each genotype. If the populations are large enough for demographic stochasticity to be ignored, the distribution of $X_t(A)$ is normal with mean $X_0(A) + s(A)t$ and variance $\sigma_s^2(A)t$, where $\sigma_s^2(A)$ is the environmental variance $\sigma_e^2(A)$ of type A if there are no autocorrelations in the noise, and otherwise the more general expression given in 1.8.3. The stochastic growth rate for A is denoted $s(A)$. Using the same notation for type B and assuming in the general case that there may be some correlation

ρ between the noise terms of the two processes the same year, we find, writing $w = \ln[p/(1-p)]$, that

$$P(X_t(A) - X_t(B) > w) = \Phi\{\frac{X_0(A) - X_0(B) + t[s(A) - s(B)] - w}{\sqrt{t}\sqrt{(\sigma_s^2(A) + \sigma_s^2(B) - 2\rho\sigma_s(A)\sigma_s(B))}}\},$$

where $\Phi(\cdot)$ is the standard normal integral. Hence, this probability tends to one for any value of w, and hence for any value of p in the interval between zero and one, as t approaches infinity, provided that $s(A) > s(B)$. If $s(A) < s(B)$ it approaches zero. This demonstrates that the stochastic growth rate is the most relevant measure of fitness for populations in a stochastic fluctuating environment. Notice that the probability tends to one if $s(A) > s(B)$ even if we have the opposite relation $r(A) < r(B)$ for the corresponding deterministic growth rates defined by the values of $\ln \mathrm{E}\Lambda$ for the two genotypes.

1.9.2 * Bet-hedging

The fact that selection acts on s rather than r has many interesting evolutionary effects. One of these is that so-called *bet-hedging* may be an optimal strategy, which means that there is not necessarily one single strategy that is the best in a stochastic environment, but rather a stochastic strategy. For example, a female could choose to lay different numbers of eggs with different probabilities. Such a stochastic strategy may work fairly well in good as well as bad years, and actually be the best strategy in the long run, in particular better than any strategy of laying a constant number of eggs. As an illustration of this concept we consider a continuous range of strategies defined by a variable U which is normally distributed with mean μ and variance σ^2. The female chooses her strategy by choosing the mean μ and the variance σ^2. For a given constant environment Z the strategy U has fitness $\lambda(U, Z)$. Writing $f(u)$ for the normal density of U, we see that the population size is changed by a factor $\Lambda(z) = \int_{-\infty}^{\infty} \lambda(u, z) f(u) du$ from one year to the next when the environmental variable is equal to z. Hence, if $g(z)$ is the density of the environmental variable Z, we have

$$s = \mathrm{E} \ln \Lambda = \int \ln[\int_{-\infty}^{\infty} \lambda(u, z) f(u) du] g(z) dz,$$

where the integration with respect to z is taken over all possible values of Z. Let the fitness function be of the Gaussian form

$$\lambda(u, z) = \lambda_0(z) \exp[-\frac{1}{2\tau^2}(u - z)^2].$$

Then, for a constant environment Z, the optimal strategy would be to keep U constant equal to Z, which is obtained by choosing $\mu = Z$ and $\sigma^2 = 0$. Solving the above integral in the case when U is a normal variate we find

$$s = \mathrm{E} \ln \Lambda = \mathrm{E} \ln \lambda_0(Z) + \ln \tau - \frac{1}{2} \ln(\tau^2 + \sigma^2) - \frac{\sigma_z^2 + (\mu_z - \mu)^2}{2(\tau^2 + \sigma^2)},$$

where μ_z and σ_z^2 is the mean and variance of Z. This stochastic growth rate is maximized for

$$\mu = \mu_z$$

and

$$\sigma^2 = \begin{cases} \sigma_z^2 - \tau^2 & \text{for } \sigma_z^2 > \tau^2 \\ 0 & \text{otherwise.} \end{cases}$$

We see that bet-hedging, now interpreted as $\sigma^2 > 0$, is an optimal strategy if the stochasticity of the environment is large enough, more precisely, if $\sigma_z^2 > \tau^2$. If $\sigma_z^2 < \tau^2$, then the constant strategy $U = \mu_z$ is optimal. The stochastic growth rate obtained by this optimization is

$$s = \begin{cases} \mathrm{E} \ln \lambda_0(Z) - \frac{1}{2} \ln(\sigma_z^2/\tau^2) - \frac{1}{2} & \text{for } \sigma_z^2 > \tau^2 \\ \mathrm{E} \ln \lambda_0(Z) - \frac{1}{2}\sigma_z^2/\tau^2 & \text{otherwise.} \end{cases}$$

1.10 Exercises

1. For the stochastic model in Fig.1.2 find an expression for the probability that the population size is larger than the expected population size as a function of time. Calculate this probability for t=10, 50, 100, 200 and 500 years.
2. Ignoring the possibility of extinction apply the central limit theorem to show that $P(N_t < N_0 e^{st})$ approaches 1/2 as t increases.
3. Derive an expression for the skewness $\mathrm{E}(V - \mathrm{E}V)^3/\mathrm{var}(V)^{3/2}$ of the lognormal distribution using the moment generating function for the normal distribution.
4. Plot the skewness of the population size against time for the model shown in Fig.1.2

5. Use Taylor expansion to show that $s \approx r - \frac{1}{2}\sigma_e^2$ and $\sigma_e^2 \approx \sigma_r^2/\lambda^2$ in general when σ_e^2 is small.
6. For the stochastic model in Fig.1.2 find an expression for the mean and standard deviation of the population size as a function of time. Make a graph of the mean plus minus one standard deviation as function of time.
7. Make the same graph as in exercise 6 on the log scale.
8. Consider the decomposition in 1.7 of the contributions $w_i = \mathrm{E}w + e + d_i$, where $e = \mathrm{E}(w|\mathbf{z}) - \mathrm{E}w$ and $d_i = w_i - \mathrm{E}(w|\mathbf{z})$. Show that $e, d_1, d_2, ..., d_x$ has zero means and that $\mathrm{cov}(e, d_i) = 0$.
9. Use the decomposition in 1.7 to show that $\mathrm{var}(\Delta N|N) = \mathrm{var}(\sum_{i=1}^N w_i|N) = N(\sigma_d^2 - \tau) + N^2(\sigma_e^2 + \tau)$, where $\sigma_d^2 = \mathrm{var}(d_i)$, $\sigma_e^2 = \mathrm{var}(e)$ and $\tau = \mathrm{cov}(d_i, d_j)$, $i \neq j$.
10. Show that $\sigma_d^2 = \mathrm{E}[\mathrm{var}(w_i|\mathbf{z})]$, $\sigma_e^2 = \mathrm{var}[\mathrm{E}(w_i|\mathbf{z})]$ and $\tau = \mathrm{E}[\mathrm{cov}(w_i, w_j|\mathbf{z})]$.
11. Show that the coefficient $\sigma_d^2 - \tau$ of N in $\mathrm{var}(\Delta N|N)$ is $\mathrm{var}(w_i) - \mathrm{cov}(w_i, w_j)$, and that the coefficient $\sigma_e^2 + \tau$ of N^2 is $\mathrm{cov}(w_i, w_j)$ where $i \neq j$.
12. What is the demographic variance for a population where no individuals are born and the adults die independently of each other with probability p?
13. What is the environmental and demographic variance in the population in exercise 12 when p varies between years with mean μ_p and variance σ_p^2?
14. A female contributes with a Poisson distributed number of offspring with mean ν if she survives. If she dies none of her offspring survive to the next generation. The females survive with probability p and there is no variation in these parameters between generations. Show that the growth rate is $\lambda = (\nu + 1)p$ that the demographic variance is $\sigma_d^2 = \nu p + (\nu + 1)^2 p(1 - p)$.
15. Consider the same model as in 14. Assume that the adult survival p is constant while the parameter ν varies between years with mean value μ and variance σ^2. Show that $\sigma_e^2 = p^2\sigma^2$ and $\sigma_d^2 = p\mu + (\sigma^2 + \mu^2 + 2\mu + 1)p(1 - p)$.
Hint: You may choose ν to be the environmental variable z in the text. Then use the relations $\sigma_d^2 = \mathrm{E}\mathrm{var}(w|z)$ and $\sigma_e^2 = \mathrm{var}\mathrm{E}(w|z)$.
16. Consider a population where the maximum obtainable value of Λ is θ and assume that all positive values of Λ below θ are equally likely, that is, Λ is uniformly distributed on $[0, \theta]$. Find λ, r, σ_e^2, σ_r^2 and s expressed by θ. Why do the approximations $s \approx r - \frac{1}{2}\sigma_e^2$ and $\sigma_r^2 \approx \sigma_e^2/\lambda^2$ break down for this model? Assuming an initial population size of 1000 individuals find the mean and median of the population size after 100 years expressed by θ. What happens if $\theta = 2.2$? Show that if $2 < \theta < e$ then the expected population size approach infinity while
$P(N_t < a)$ approach 1 for any $a > 0$. Discuss this result.
17. Consider an individual producing B offspring a given year and let J be the indicator variable for her survival, that is, $J = 1$ if she survives and otherwise 0. Her contribution to the next generation is then $w = B + J$. Discuss how the

covariance between B and J affects the environmental and demographic variance.
18. Show that the stochastic growth rate of a population of size N with environmental and demographic stochasticity is approximately $r - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2$ provided that $\Delta \ln N$ is approximately normally distributed.

19. Assume that the $\ln \Lambda_t$ are multinormally distributed with constant variance and $\mathrm{corr}(\Lambda_t, \Lambda_{t+h}) = \rho(h) = \rho(-h)$. Show that $s = r - \frac{1}{2}\sigma^2 = u - \frac{1}{2}\sigma^2 \sum_{-\infty}^{\infty} \rho(i)$ where u is the growth rate of the expected population size as defined in 1.8.3.

Chapter 2

Density-regulated populations

2.1 The concept of density-regulation

In chapter 1 we assumed that the vital rates determining the dynamics were unaffected by the population size N. This is only realistic for species that are not limited in growth by food or space. Most populations, however, will grow to reach population densities that are so large that within species competition for resources will affect the vital rates. Then, the contribution from a single individual to the next generation must have a distribution that depends on the population size N. In fisheries, for example, it is common to divide the population into two parts, the spawning stock (N_S) which is the reproducing fragment of the population, and the recruitment (N_R), which is their production of new individuals which will enter the spawning stock when they reach their age of maturity. The two most common models for density regulation are the Beverton-Holt model and the Ricker model that expresses the expected recruitment as functions of the spawning stock. The Beverton-Holt model is given by $N_R = \alpha N_S/(1 + \beta N_S)$ while the Ricker model is $N_R = \alpha N_S e^{-\beta N_S}$, where α and β are constants that must be estimated from data.

In the kind of models dealt with in chapter 1, density-regulation is most appropriately introduced by assuming that the expected contributions to the next generation, $\lambda = \mathrm{E}w$, depends on the population size, hence writing

$\lambda = \lambda(N)$. Correspondingly, the deterministic growth rate is then $r(N) = \ln \lambda(N)$. If ΔN is small compared to N we then find

$$s(N) = \mathrm{E}[\Delta \ln(N)|N] \approx \mathrm{E}[\frac{\Delta N}{N}|N] = \lambda(N) - 1.$$

Density regulation is commonly introduced by assuming that $r(N)$ and $\lambda(N)$ are decreasing functions of the population size N. When analyzing stochastic models, however, we will usually work with the stochastic growth rate. The carrying capacity K of the population is then defined as the stable equilibrium point given by $s(K) = 0$. For a deterministic model there is no distinction between r and s giving $r(K) = 0$.

It is important to notice that a density-regulated population is defined as a population for which the change in population size from one year to the next depends on the population the previous year. In chapter 1 we dealt with populations without density-regulation which we interpreted as the distribution of $\Delta \ln N$ being independent of N. As a further illustration of the concept of density-regulation consider a population producing a large number of eggs each season, out of which only a small fraction can survive to enter the population as adults. If the environment is independent between year the population sizes may also be independent. For example let us assume that $\ln N_t$ is a sequence of independent normally distributed variable with mean μ and variance σ^2. Then

$$s(N) = \mathrm{E}(\Delta \ln N|N) = \mathrm{E}[\ln(N + \Delta N)|N] - \ln(N) = \mu - \ln(N)$$

and $\sigma_e^2 = \mathrm{var}(\Delta \ln N|N) = \sigma^2$. It appears that this population is strongly density-regulated with carrying capacity at $K = e^{\mu}$ although this may seem surprising since the populations sizes are independent between years.

2.2 Return time to equilibrium and strength of density-regulation

An important parameter in deterministic models with density-regulation is the *characteristic return time to equilibrium* T_R, which is closely related to the *strength of the density regulation* at K. Defining the relative deviation $\varepsilon = (N-K)/K$, the dynamic equation for small values of ε is given by the first order approximation $\Delta\varepsilon = [\lambda(N)-1]N/K \approx \lambda(N)-1 \approx K\lambda'(K)\varepsilon$, giving $\varepsilon_t \approx \varepsilon_0 \exp[K\lambda'(K)t]$. The time T_R is defined as the time required for the deviation to reach a fraction $1/e$ of its original value, giving $T_R \approx -1/[K\lambda'(K)] = -1/[Kr'(K)]$. Hence, if the negative slope of $\lambda(N)$ or $r(N)$ is large at $N = K$, the characteristic return time to equilibrium is small. We shall later show that this concept of characteristic return time also has an interesting interpretation in stochastic models for populations fluctuating around a carrying capacity. For such models the autocorrelations between population sizes at two different points of time drops to approximately $1/e$ when the time difference is T_R. A natural measure of the strength of density-regulation is now $\gamma = 1/T_R = -K\lambda'(K) = -Kr'(K)$ which is large if the return time to equilibrium is small and visa versa.

Writing as before $X = \ln N$ we have

$$\frac{d\lambda}{dN} = \frac{d\lambda}{dX}\frac{dX}{dN} = \frac{d\lambda}{dX}\frac{1}{N}$$

so that $K\lambda'(K)$ alternatively can be written as $d\lambda/dX$ evaluated at $N = K$. Further, since $d\ln\lambda/dX = \lambda^{-1}d\lambda/dX$ and $\lambda(K) = 1$ by the definition of K, we see that the strength of density-regulation can be written in different ways, actually as

$$\gamma = -K\lambda'(K) = -\frac{d\lambda}{dX} = -\frac{d\ln\lambda}{dX} = -\frac{d\ln\lambda}{d\ln N},$$

where all derivatives are evaluated at $N = K$.

In the deterministic analog of the previous model with independent lognormally distributed population sizes subsequent years, we have $\mathrm{E}(N+\Delta N) = \exp(\mu+\sigma^2/2)$ giving $\ln\lambda(N) = \mu+\sigma^2/2-\ln N$. Hence, the strength of density regulation γ as well as the return time to equilibrium T_R is one in this model. Notice that this result does not depend on the assumption that the population sizes are lognormally distributed. For any distribution of population sizes we have that $\ln\lambda = \ln(\mathrm{E}N) - \ln N$, where $\mathrm{E}N$ does not depend on N, giving $T_R = \gamma = 1$.

2.3 The deterministic logistic model

For populations with small fluctuations around K the dynamics may be described by a linear approximation to $\lambda(N)$ around K writing

$$\lambda(N) \approx \lambda(K) + \lambda'(K)(N-K) = \lambda(K) + \gamma(1 - N/K)$$

giving

$$\Delta N/N = \lambda(N) - 1 \approx \gamma(1 - N/K).$$

This model, which is called the logistic model, or the logistic type of density regulation, may often be realistic for all values of N. Notice then that as N approaches zero $\Delta N/N$ approaches $\lambda(0) - 1$ so that the model also can be written

$$\Delta N = [\lambda(0) - 1]N(1 - N/K) \approx r(0)N(1 - N/K).$$

An alternative formulation which is almost equivalent to this is

$$\Delta(\ln N) = r(0)(1 - N/K)$$

or equivalently

$$r(N) = r(0)(1 - N/K).$$

We see that the logistic model has the property that the strength of density regulation is determined by the growth rate at small population sizes and is not at all affected by the carrying capacity. More precisely, using the definition $\gamma = -K\lambda'(K)$ we see that γ is the growth rate r at $N = 0$.

2.4 The log-linear model and Gompertz type of density-regulation

If we rather than approximating $\lambda(N)$ by a linear expression in N perform the linearization in $X = \ln N$ we arrive at the model

$$\lambda(N) \approx \lambda(K) + \frac{d\lambda}{dX}|_{N=K}(X - \ln K) = \lambda(K) + \gamma(\ln K - \ln N),$$

giving

$$\Delta N \approx \gamma N \ln K \left(1 - \frac{\ln N}{\ln K}\right).$$

In this form the model is often called the Gompertz type of density-regulation. If we rather write the model in terms of $X = \ln N$ we get

$$\Delta X \approx \gamma k(1 - X/k)$$

where $k = \ln K$ is the carrying capacity on the log scale. We see that this is a linear model in X and is accordingly called a log-linear model. If we add a stochastic term with zero mean and constant variance to the left of the equation we obtain what in statistics is called a first order autoregressive model.

2.5 The theta-logistic model

A general class of functions defining different types of density regulation is the so-called theta-logistic class of models

$$r(N) = r_0[1 - (\frac{N}{K})^\theta]$$

corresponding to the logistic model if $\theta = 1$.
In order to make this model valid for any value of θ, including $\theta \leq 0$, the parameter r_0 must be chosen as a function of θ. One way of doing this is to choose the growth rate at abundance $N = 1$, say r_1, as a free parameter, giving

$$r_1 = r_0(1 - K^{-\theta})$$

and

$$r(N) = \frac{r_1}{1 - K^{-\theta}}[1 - (\frac{N}{K})^\theta] = r_1[1 - \frac{N^\theta - 1}{K^\theta - 1}]$$

for $\theta \neq 0$, whereas for $\theta = 0$ we obtain the limit

$$r(N) = r_1(1 - \frac{\ln N}{\ln K}).$$

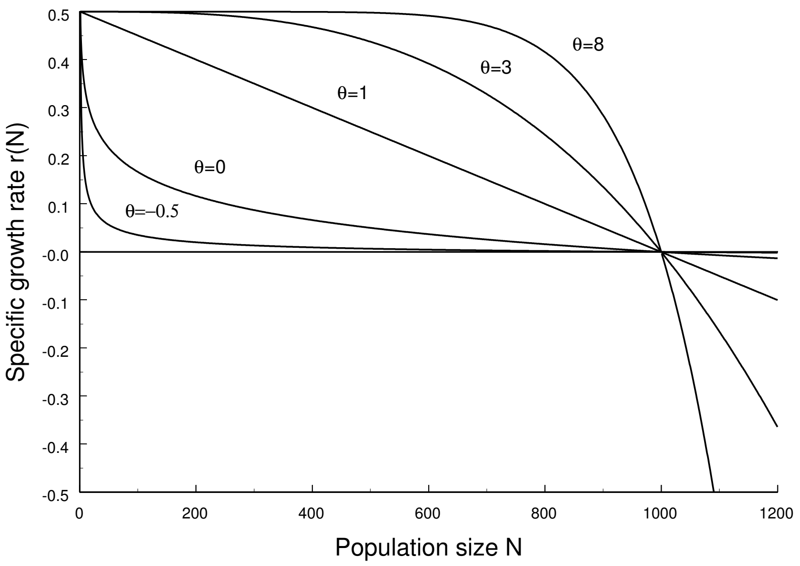

Figure 2.1: The deterministic growth rate $r(N)$ as a function of N for different values of θ in the thetalogistic model. The other parameters are $r_1 = 0.5$ and $K = 1000$.

Fig.2.1 shows the function $r(N)$ for different values of θ .

For the characteristic return rate to equilibrium we find for this class of models that $T_R = 1/(r_0\theta) = (1 - K^{-\theta})/(r_1\theta)$ for $\theta \neq 0$ and $T_R = \ln K/r_1$ for $\theta = 0$. Notice that the characteristic return time depends on r_0 and θ only through their product.

We see in Fig.2.1 that the theta-logistic model corresponds to very different types of density-regulation for different values of θ. For a small θ the regulation starts to act already at very small population sizes. If θ is large, however, there is practically no density-regulation when the population size is smaller than K, but the regulation gets strong when the population size approaches K. We can summarize this by considering 4 special cases of the model:

Type I: $\theta = -1$. This leads to $\Delta \ln N \approx \Delta N/N = r_1(1 - \frac{1/N-1}{1/K-1})$ giving $\Delta N \approx r_1\frac{K-N}{K-1}$. Hence, we see that this value of θ simply corresponds to a linear model in N.

Type II: $\theta = 0$. We have already seen that this case may be written as $\Delta(\ln N) = r_1(1 - \ln N/\ln K)$, so now the model is linear in $\ln N$.

Type III: $\theta = 1$. This give the logistic model $\Delta N \approx r_1N(1 - N/K)$.

Type IV: $\theta = \infty$: In the limit as θ approaches infinity the model approaches $\Delta \ln N = r_1$ for any $N < K$, that is, there is no density-regulation below K. Immediately above K ΔN approaches $-\infty$. Hence, K plays the role of a ceiling for the population size. Accordingly, this kind of model is often called a ceiling model.

2.6 Stochasticity and density-regulation

Obtaining a deterministic analog to a stochastic model can be done simply by just considering the expected values of ΔN or $\Delta \ln N$, or even using some other transformation of N. But choosing these different transformations leads to different deterministic analogs of a stochastic model. From the discussion of stochastic growth rate in chapter 1, however, we have seen that it is preferable to use the expected value on the log scale, that is, the stochastic growth rate, since the growth on the log scale actually tends to a constant as the time intervals get large in the case of no density-regulation. Going in the opposite direction, from a deterministic to a stochastic model, however, leaves us with lots of choices. A deterministic model may have several parameters and may be parametrized in different ways. Each parameter may be subject to temporal fluctuation, and additional stochastic terms with zero mean or stochastic factors with mean 1 may also be included. The overall goal, however, is to formulate a model that is realistic for the population we deal with.

The most common way of formulating a stochastic model is to assume that the parameter expressing the population growth rate at small densities fluctuates in time. But this assumption alone is not enough to uniquely define a model. Take for example the logistic model which we can write as $\Delta \ln N = r(1 - K/N)$ or as $\Delta \ln N = r - \beta N$, where $\beta = r/K$. Although these two deterministic models are equivalent, replacing r by a temporally fluctuating parameter $r(t)$ in the two models will give rather different models. Here, we shall chose the second formulation since it turns out to be most realistic (exercise 1). More generally, we choose to write deterministic models on the form

$$\Delta \ln N = r - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2 - g(N),$$

where we already have taken into account the reduction in the stochastic growth rate due to environmental and demographic stochasticity. Replacing r by a variable $r(t)$ fluctuating in time, we obtain the stochastic model

$$S_t = \Delta \ln N = r(t) - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2 - g(N).$$

Notice that in this formulation there is no stochasticity in the density regulating term. Hence, the variance in the change in the logarithm of the population size during one time step conditioned on the population size the previous year is $\mathrm{var}(S_t|N) = \mathrm{var}[r(t)] \approx \mathrm{var}(\Delta N/N) = \sigma_e^2 + \sigma_d^2/N$.

For the above discrete theta-logistic model this approach leads to

$$S_t = r_1(t) - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2 - \bar{r}_1 \frac{N^\theta - 1}{K^\theta - 1}$$

where $\bar{r}_1$ is the mean of $r_1(t)$. In this formulation $g(N) = -\bar{r}(N^\theta - 1)/(K^\theta - 1)$.

As θ approaches zero we obtain the type II model in 2.2, which is linear in $\ln N$. If $\sigma_d^2 = 0$ and $\Delta \ln N$ is normally distributed this is a first order autoregressive time series for which the mean of future values, the autocorrelation function and the stationary distribution is well known. These results can then be used to find corresponding results for the population size N (exercise 8 and 9).

As mentioned above, there are many different stochastic generalizations of a deterministic model. For example, in the theta-logistic model we may let all three parameters be stochastic processes $r_1(t)$, $K(t)$ and $\theta(t)$. Examples of more general models will be given in section 2.7 and in chapter 3.

Until now we have assumed that r (and λ), for any value of N, are stochastic variables with distributions depending on N. In the next section we show some examples of modelling the stochastic change in the population size from first principles, considering the stochastic contribution of each individual to the next generation. One advantage of doing this is that it may give some insight into how $\text{var}(\Delta N|N)$ is expected to depend on N, a relationship that may be crucial when it comes to analyzing population fluctuations over long time intervals with the possibility of large as well as small population sizes. Above we have assumed that σ_e^2 and σ_d^2 do not depend on the population size.

2.7 Density-dependence in the demographic and environmental variances

2.7.1 The distribution of vital rates

We have defined a population to be density-regulated if the expected relative change in the population size from one year to the next depends on the population size. The population growth, however, is determined by the vital rates, the reproduction and survival, of the individuals in the population. Generally, if the population size has an effect on the expected value of $\Delta N/N$ or $\Delta \ln N$, it will also affect the variances included in the model so that the demographic and environmental variance both are functions of N, say $\sigma_e^2(N)$ and $\sigma_d^2(N)$. For example, if J is the indicator of survival for an individual in the environment $\mathbf{z}$ with expected value $p(\mathbf{z})$, then $\mathrm{E}J = \mathrm{E}p(\mathbf{z})$ and $\mathrm{var}(J) = \mathrm{var}[p(\mathbf{z})] + \mathrm{E}p(\mathbf{z})[1 - p(\mathbf{z})] = \mathrm{E}p(\mathbf{z}) - \mathrm{E}p(\mathbf{z})^2$, indicating a relationship between the mean and variance of J. Below we illustrate this possible relationship through some theoretical examples of fluctuating and density-regulated vital rates.

2.7.2 A logistic model with Poisson distributed contributions

Assume that, for a given environment $\mathbf{z}$, the contributions are independent Poisson variates with means $\lambda(\mathbf{z}, N)$. Then the expected population size the next year conditioned on the environment is $N\lambda(\mathbf{z}, N)$, giving $\mathrm{E}(\Delta N|\mathbf{z}, N) = N[\lambda(\mathbf{z}, N) - 1]$, and the unconditional expectation $\mathrm{E}(\Delta N|N) = N[\mathrm{E}\lambda(\mathbf{z}, N) - 1]$, where the expectation of $\lambda(\mathbf{z}, N)$ is taken with respect to the environmental variable $\mathbf{z}$. We then obtain the logistic model with carrying capacity K if we choose $\mathrm{E}\lambda(\mathbf{z}, N) = 1 + \gamma(1 - N/K)$, where $\gamma = r$ expresses the strength of the density regulation. The environmental variance in this model is now $\sigma_e^2 = \mathrm{var}[\mathrm{E}(w_i|\mathbf{z}, N)] = \mathrm{var}[\lambda(\mathbf{z}, N)]$. As a further demonstration of this concept let us choose two different models, one where the stochasticity in $\lambda(\mathbf{z}, N)$ is defined as a multiplicative effect, and one where it is additive. The multiplicative model takes the form

$$\lambda(\mathbf{z}, N) = A(\mathbf{z})(1 + \gamma(1 - N/K)),$$

where the mean of $A(\mathbf{z})$ over years is 1. Writing $\sigma^2 = \mathrm{var}[A(\mathbf{z})]$, we then find

$$\sigma_e^2(N) = \sigma^2[1 + \gamma(1 - N/K)]^2$$

which decreases with N. On the other hand, choosing an additive model of the type

$$\lambda(\mathbf{z}, N) = 1 + \gamma(1 - N/K) + \varepsilon(\mathbf{z})$$

where $\mathrm{E}\varepsilon(\mathbf{z}) = 0$, we find that the environmental variance is equal to $\mathrm{var}[\varepsilon(\mathbf{z})]$ and hence independent of N if this variance is constant.

Using the fact that the variance of the Poisson is the same as the mean we find for both models

$$\sigma_d^2 = \mathrm{E}[\mathrm{var}(w_i|\mathbf{z}, N)] = 1 + \gamma(1 - N/K).$$

Notice that this is a decreasing function which is approximately 1 when the population size is close to the carrying capacity.

Fig.2.2 shows simulations of the multiplicative process for $K = 1000$, $\gamma = r = 0.1$ and $\sigma^2 = 0.01, 0.04$.

2.7.3 * Environmental fluctuations in r and K

Consider now the discrete logistic model on the form

$$N_{t+1} = N_t \exp[r(1 - N_t/K)],$$

where r is the specific growth rate for small population sizes and K is the caring capacity. Environmental fluctuations may be introduced by assuming that r as well as K depend on the environment, giving the model

$$\Delta N = N \exp[r(\mathbf{z})(1 - N/K(\mathbf{z}))] + \sum_{i=1}^{N} d_i$$

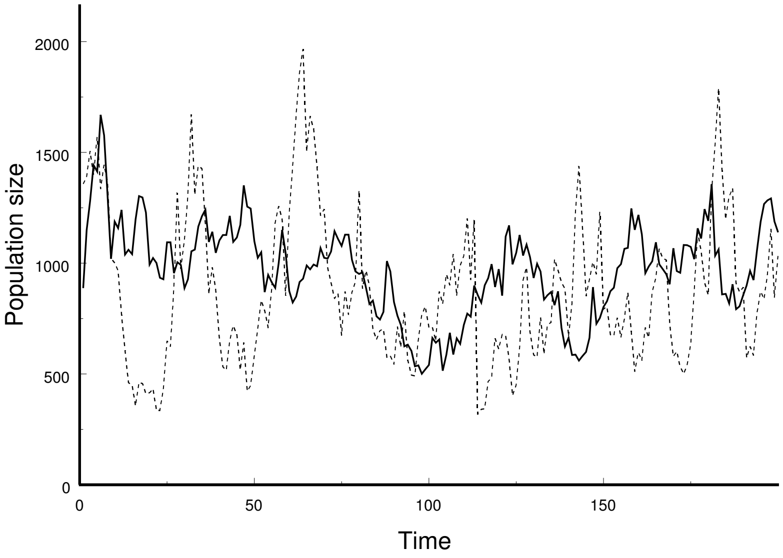

Figure 2.2: Simulations of the multiplicative process with $\sigma^2 = 0.01$ (solid line) and 0.04 (dotted line). The other parameters are $K = 1000$ and $\gamma = r = 0.1$, corresponding to a return time to equilibrium $T_R = 1/\gamma = 10$.

where the d_i are the demographic components which are independent with variance σ_d^2. It follows from this assumption that

$$\mathrm{E}(w_i|N,\mathbf{z}) = E[(\Delta N + N)/N|N,\mathbf{z}] = \exp[r(\mathbf{z})(1 - N/K(\mathbf{z}))]$$

implying that

$$\sigma_e^2(N) = \mathrm{var}\{\exp[r(\mathbf{z})(1 - N/K(\mathbf{z}))]|N\}$$

which can then be evaluated if the bivariate distribution of $r(\mathbf{z})$ and $K(\mathbf{z})$ is known. As a simple illustration assume $r(\mathbf{z})/K(\mathbf{z})$ to be a constant, say β, not depending on the environment, and let $r(\mathbf{z})$ be normally distributed with mean $\bar{r}$ and variance σ_r^2. Then,

$$\sigma_e^2 = \exp[-2\beta N]\mathrm{var}\{\exp[r(\mathbf{z})]\}$$

giving

$$\sigma_e^2 = \exp[2(\bar{r} - \beta N)]\exp(\sigma_r^2)[\exp(\sigma_r^2) - 1],$$

which for small σ_r^2 and $N \approx K$ is actually equal to σ_r^2 to the first order.

2.7.4 * A model with density-regulated fecundity

Consider a population where the adults survive with a constant probability p, and produces exactly one offspring that survives with probability $q\exp(-\alpha N)$, where α is a positive parameter. The only parameter depending on the environment is $q = q(\mathbf{z})$, which is the juvenile survival at small densities ($N = 0$). The strength of density regulation is determined by α. Actually, since $\mathrm{E}(w) = p + \bar{q}\exp(-\alpha N)$, where $\bar{q} = \mathrm{E}[q(\mathbf{z})]$, and the carrying capacity is the value of N giving $\mathrm{E}(w) = 1$, we find

$$\alpha = K^{-1}\ln(\frac{\bar{q}}{1-p}).$$

Since $N + \mathrm{E}\Delta N = N\mathrm{E}w_i$, we find

$$\mathrm{E}(\Delta N|N) = N\{p + \bar{q}^{\;1-N/K}(1-p)^{N/K} - 1\},$$

or equivalently

$$\lambda(N) = \{p + \bar{q}^{\;1-N/K}(1-p)^{N/K}\}.$$

Let us write $\sigma_q^2 = \mathrm{var}[q(\mathbf{z})]$, for the variance of q between years. If we assume that survival and fecundity are independent, and independent between individuals for a given $\mathbf{z}$, then the demographic covariance τ is zero. The conditional mean and variance of the contributions are

$$\mathrm{E}(w_i|\mathbf{z}) = p + q(\mathbf{z})(\frac{1-p}{\bar{q}})^{N/K}$$

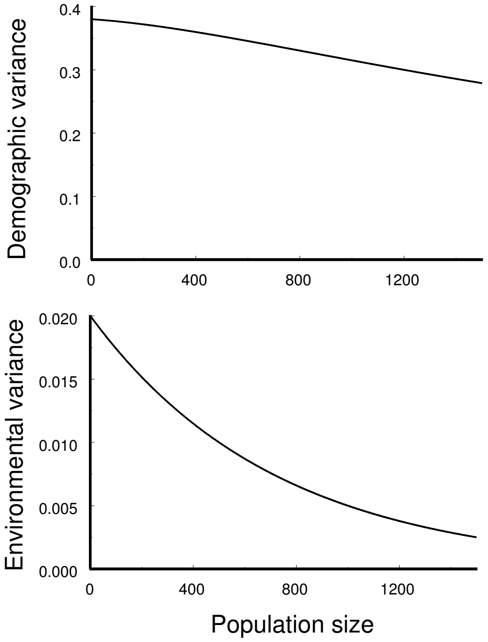

Figure 2.3: The demographic variance σ_d^2 (upper panel) and σ_e^2 (lower panel) as functions of N for the model given in 2.7.4 with parameters $K = 1000$, $p = 0.8$, $\bar{q} = 0.4$ and $\sigma_q^2 = 0.02$.

$$\mathrm{var}(w_i|\mathbf{z}) = p(1-p) + q(\mathbf{z})(\frac{1-p}{\bar{q}})^{N/K}[1 - q(\mathbf{z})(\frac{1-p}{\bar{q}})^{N/K}]$$

from which we find

$$\sigma_d^2(N) = \mathrm{E}[\mathrm{var}(w_i|\mathbf{z})] = p(1-p) + \bar{q}(\frac{1-p}{\bar{q}})^{N/K} - (\sigma_q^2 + \bar{q}^2)(\frac{1-p}{\bar{q}})^{2N/K}$$

$$\sigma_e^2(N) = \mathrm{var}[\mathrm{E}(w_i|\mathbf{z})] = \sigma_q^2(\frac{1-p}{\bar{q}})^{2N/K}$$

Notice that the demographic as well as the environmental variance in this model are functions of the population size N. An example for a given set of parameters is given in Fig.2.3.

2.7.5 * An example of demographic covariance

In the previous examples there was no covariance between the individual contributions for given environmental conditions, in which case the demographic covariance is zero. The following is a theoretical example of a model with demographic covariances that may serve as a further demonstration of the concepts defining the stochasticity.

Suppose that each individual of a population independently produce a Poisson distributed number of eggs with mean much larger than one. Suppose that the offspring selected to constitute the next generation are chosen at random from the total amount of eggs. The environmental variable $\mathbf{z}$ is assumed to have an effect on the population only through its effect on the population size each year. We assume further that this relationship is deterministic, that is, $N = N(\mathbf{z})$. The fluctuations in N between years are then fully determined by the fluctuations in the environmental vector $\mathbf{z}$. If the sequence of environmental vectors are independent, then the population size the next year is independent of the population size the year before, provided that the population is always large enough to produce the number of eggs required for a population of a size determined by the environmental conditions.

Let us consider one year with population size N producing the next years population $N + \Delta N = \sum_{i=1}^{N} c_i$ determined uniquely by some $\mathbf{z}$. Conditional on $N + \Delta N$, the contributions are then multinomially distributed with parameters $(N+\Delta N, 1/N, 1/N, \ldots, 1/N)$. Writing shortly $\mathrm{E}N = \mu$ and $\mathrm{var}(N) = \sigma^2$ we find $E(w_i|\mathbf{z}) = (N + \Delta N)/N$, $\mathrm{var}(w_i|\mathbf{z}) = (N + \Delta N)(1 - 1/N)/N$ and $\mathrm{cov}(w_i, w_j|\mathbf{z}) = -(N + \Delta N)/N^2$. Hence $\sigma_d^2 = \mu(1 - 1/N)/N$, $\sigma_e^2 = \sigma^2/N^2$ and $\tau = -\mu/N^2$. Notice that, if the fluctuations in the population size are small, then $N \approx \mu$, giving $\sigma_d^2 \approx 1 - 1/\mu$, $\sigma_e^2 \approx \sigma^2/\mu^2$ and $\tau \approx -1/\mu$. Generally, however, all three parameters may depend strongly on N.

The coefficients of N and N^2 in the expression for $\mathrm{var}(\Delta N|N)$ are $\sigma_d^2 - \tau = \mu/N$ and $\sigma_e^2 + \tau = (\sigma^2 - \mu)/N^2$ giving $\mathrm{var}(\Delta N|N) = \sigma^2$ in accordance with the assumptions. Notice that the coefficient of N^2, $\sigma_e^2 + \tau = (\sigma^2 - \mu)/N^2$, is negative if $\sigma^2 < \mu$ and otherwise positive, and that both coefficients decreases with N.

2.8 Estimation of demographic and environmental components.

For most populations one will only have a sequence of estimates of the population size each year, say $\hat{N}_1, \hat{N}_2, \ldots, \hat{N}_k$. Often, the uncertainty in the estimates are quite large, which has the consequence that the fluctuations due to demographic and environmental variances may simply disappear in the fluctuations in the estimates due to the uncertainty in the estimation procedure. In such cases one will have to use more advanced statistical methods constructed for taking into account sampling errors, for example using the bayesian approach analyzed by Markov chain Monte Carlo Methods.

In this section we consider only the case of a full census being carried out each year so that the observations are actually $N_1, N_2, \ldots, N_k$. Even if such data are available, there may be quite unrealistic to believe that the demographic and environmental component may be separated from time series of realistic length. We therefore suggest first trying to estimate the demographic component from general biological insight and facts known for the species. In particular, if data on individual survival and fecundity are available, estimation of the demographic component may be done each year. Such an approach have the advantage that it also will give some information on how the demographic stochasticity depends on the population size.

For practical applications, such as a population viability analysis, we will only be interested in the coefficients of N and N^2 in the expression for $\mathrm{var}(\Delta N|N)$, that is, $\sigma_d^2 - \tau$ and $\sigma_e^2 + \tau$. Using the decomposition $w_i = \mathrm{E}w + e + d_i$ given in chapter 1, we see that $\mathrm{var}(w_i) = \sigma_e^2 + \sigma_d^2$ and $\mathrm{cov}(w_i, w_j) = \sigma_e^2 + \tau$. Writing V for the unconditional (not conditioned on $\mathbf{z}$) variance of the contributions w_i and C for the unconditional covariance of any two contributions within a year, we then have simply (exercise 12)

$$\mathrm{var}(\Delta N|N) = (V - C)N + CN^2,$$

which we write as

$$\mathrm{var}(\Delta N|N) = \theta_1(N)N + \theta_2(N)N^2.$$

Suppose a random sample of individual fitness a given year, say $w_1, w_2, \ldots, w_n$ have been recorded. Then (exercise 13)

$$E\{\frac{1}{n-1}\sum(w_i - \bar{w})^2\} = \theta_1(N)$$

where $\bar{w}$ is the mean of the observed contributions. Hence, the standard sum of square estimate $1/(n-1)\sum(w_i - \bar{w})^2$ computed from a sample of contributions within a year, is an unbiased estimator for the coefficient of N for the N-value that year. If such estimators are available for a sequence of years, the estimates may be plotted against N to reveal how the coefficient depends on N, as illustrated in Fig.2.4 (upper panel). Notice that this result is quite different from the corresponding result for independent variables used to estimate variances. Although the sum of squares are identical, our observations are stochastically dependent and the aim is not to estimate V, but actually $\theta_1 = V - C$.

The environmental variance may be estimated from the fluctuations of the population size, provided that the demographic component is known and can be subtracted. Suppose we have a parametric model for $\mathrm{E}(\frac{\Delta N}{N}|N) = h(\alpha, N)$, where α is some vector of unknown parameters . We can estimate α by maximum likelihood or by minimizing $\sum[\Delta N_t/N_t - h(\alpha, N_t)]^2$, with respect to α. Since $\mathrm{var}(\Delta N/N) = \mathrm{E}[\Delta N/N - h(\alpha, N)^2] = \theta_1(N)/N + \theta_2(N)$, we find that

$$\theta_2(N) = \mathrm{E}\{[\Delta N/N - h(\alpha, N)]^2 - \theta_1(N)/N\}.$$

Hence, if we replace α by the estimate $\hat{\alpha}$ we obtain an approximately unbiased estimate

$$\hat{\theta}_2(N_t) = [\Delta N_t/N_t - h(\hat{\alpha}, N_t)]^2 - \hat{\theta}_1(N_t)/N_t$$

which may be plotted against N_t to check whether $\theta_2(N)$ is constant or changes with N (Fig.2.4, lower panel).

Fig. 2.4 shows annual estimates of σ_d^2 and σ_e^2 for a Great Tit population at Oxford. Notice that each yearly estimate is rather uncertain. However, fitting some parametric function by regression leads to much smaller uncertainties. When individual data are available as in this case, one will often have a rather large number of individual counts of reproduction and survival each year. Hence, the demographic variance can be estimated with rather high precision. The length of time series, however, is usually not very large. A 20 years time series is in biology considered to rather long. But this corresponds

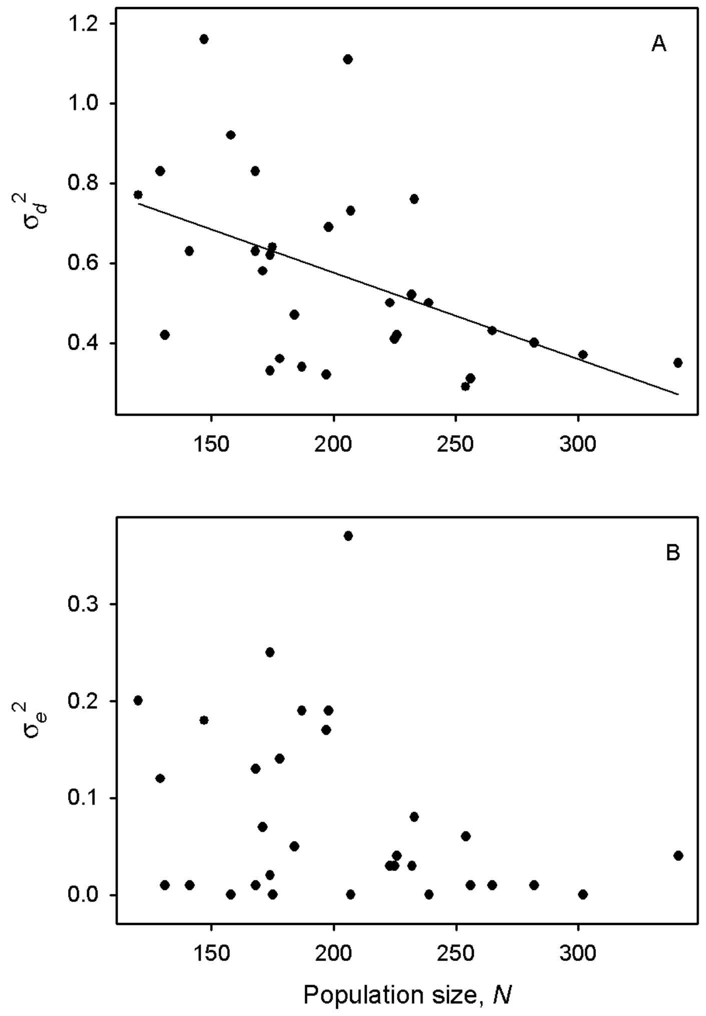

Figure 2.4: Annual estimates of the demographic and environmental variance as functions of population size N for the Great Tit in Wytham Wood at Oxford, England.

only to 20 different values of the environment, so one will always end up with a rather uncertain estimate of σ_e^2. If there are sampling errors, the problem of estimating the environmental variance is even bigger.

2.9 Exercises

1. The deterministic logistic model may be written in the form $\Delta N = rN(1 - N/K)$. If r fluctuates in time a stochastic model my be formulated as $\Delta N = r(t)N(1 - N/K)$. Look at the properties of this stochastic model. Is it a realistic model?
2. A discrete deterministic model can be approximated by replacing $\mathrm{E}(\Delta \ln N)$ by the derivative $d \ln N/dt$ since the time step is one. Solve this differential equation for the logistic model assuming that the population size is large enough to ignore the demographic variance to find the expected value of log population size as a function of time starting at $N(0) = n_0$.
3. Solve the problem in exercise 2 for the theta-logistic model.
4. Verify the results for the four special cases of the theta-logistic model given in the text, that is, for $\theta = -1, 0, 1$ and ∞.
5. As $\theta \to 0$ the theta-logistic model approaches a linear model in $\ln N$. If the population is large enough for the demographic variance to be ignored this is what is called a first order autoregressive model in time series analysis. Ignoring the possibility of extinction and assuming that $\Delta \ln N$ is normally distributed, find the stationary distribution of N expressed by r_1, K and σ_e^2.
Hint: remember that the stochastic growth rate at $N = 1$ is $r_1 - \frac{1}{2}\sigma_e^2$.
6. Consider the model in exercise 5. Suppose that extinctions still can be ignored but not the demographic variance. How does the results differ from those in exercise 5 if the fluctuations around the carrying capacity are so small that the demographic stochastic term can be approximated by a term with constant variance?
7. Consider models on the form $\mathrm{E}(\Delta \ln N) = r(t) - \frac{1}{2}\sigma_e^2 - \frac{1}{2N}\sigma_d^2 - g(N)$. Linearize the expression for $\mathrm{E}(\Delta \ln N)$ in the neighborhood of $N = K$ and use this to find an approximation to the stationary variance when the demographic stochastic term can be approximated by a term with constant variance.
8. The theta-logistic model with $\theta = 0$, $\sigma_d^2 = 0$ and normally distributed $\Delta \ln N$ is a first order autoregressive model. Use this to find $\mathrm{E}(\ln N_{t+h}|N_t)$, $\mathrm{corr}(\ln N_t, \ln N_{t+h})$ and the stationary distribution of $\ln N$ expressed by r_1, K and σ_e^2.
9. For the model in exercise 8 utilize well known properties of the lognormal distribution to find $\mathrm{E}(N_{t+h}|N_t)$, $\mathrm{corr}(N_t, N_{t+h})$ and the stationary distribution of N expressed by r_1, K and σ_e^2.

10. Let B denote the number of offspring of an individual and let J be the indicator variable for its survival. For a given population size N we write $\mathrm{E}(B|z) = z\nu(N)$, where $\nu(N)$ is a decreasing function expressing density regulation and z is an environmental variable with $\mathrm{E}z = 1$ and $\mathrm{var}(z) = v_z$ expressing the environments effect on the fecundity. For a given environment z we further assume that $\mathrm{E}J = p$ (not affected by the environment), the coefficient of variation of B is c_b, and the correlation between B and J is ρ, so that the only parameter depending on the environment is the expected fecundity. Find expressions for λ, σ_e^2 and σ_d^2 as functions of N for this model. Write a computer program to calculate these three functions. Plug in some reasonable values for the parameters and the function $\nu(N)$ and make a graph showing how λ, σ_e^2 and σ_d^2 depends on N in this model.
11. Write a computer program simulating the theta-logistic model choosing $\Delta \ln N$ to be normally distributed. Put $\sigma_d^2 = 0$, $\sigma_e^2 = 0.01$, $K = 1000$, $\theta = 0.5$ and $\bar{r}_1 = 0.1$ and find the variance of the stationary distribution of $\ln N$ by simulations. Compare this to what you find using the linearization around K (exercise 7).
12. Writing $V = \mathrm{var}(w_i)$ and $C = \mathrm{cov}(w_i, w_j)$ for $i \neq j$, show that $\mathrm{var}(\Delta N|N) = \theta_1 N + \theta_2 N^2 = (V - C)N + CN^2$.
13. Suppose a random sample of individual fitness a given year, say $w_1, w_2, \ldots, w_n$ have been recorded. Show that

$$E\{\frac{1}{n-1}\sum(w_i - \bar{w})^2\} = \theta_1(N)$$

where $\bar{w}$ is the mean of the observed contributions.

Chapter 3

Diffusion theory

3.1 Introduction

In models without density-regulation and age-structure we have seen that the population process on the log-scale can be approximated by a random walk when the population size is large enough to ignore demographic stochasticity. In this case future population sizes is a sum of independent identically distributed random variables, and it follows from the central limit theorem that the log population size after some years will be approximately normally distributed. It is well known that this is often a remarkably accurate approximation remembering that we only use the mean and variance of the increments and not any other properties of their distribution. For density-regulated populations we still add together the increments in log value between years, but now these will have distribution depending on the (log) population size the previous year. Hence, the central limit theorem is no longer applicable and we cannot expect the log population sizes to be normally distributed, nor can we expect to find a simple way of calculating the mean and variance. However, there is still possible to approximate the process by a process that is continuous in space and time and only uses the expectation and variance, conditioned on the size the previous year, of the between years change in population size like we do in applications of the central limit theorem. These continuous processes are called *diffusion processes* and a number of results for these processes do play the same important role as the central limit theorem in the sense that they are based only on means and variances and often give remarkably accurate approximations to the real process.

Another class of processes often use in population dynamics are *time series models*. These are models that can be used to analyze stationary processes. One advantage of the time series approach is that the theory also include incorporation of so called time-lags, that is, the change in population size do not only depend on the size last year, but also on the values for one or more previous year. Ecological processes for a single species may often be of this type, for examples due to interactions with other species, temporal autocorrelations in the environmental noise, or age-structure.

Time series analysis, however, do not provide any theory for first passage times, that is, for calculating properties of the time it takes for the population size to reach a certain value. The most important problem of this type in population dynamics is the analysis of extinction processes. Diffusion theory provides a number of results relating to this, for example methods for calculating the distribution of the first passage time in some relatively simple cases, and more generally methods for calculating its expectation. Such methods are particularly useful when it comes to performing population viability analysis and judging to which extent species are threatened or vulnerable. It turns out that also processes with time lags and autocorrelated noise in many cases can be approximated with diffusions. In chapter 4, for example, we shall deal with the accurate diffusion approximation to age-structured populations where there are strong time lag effect generated by fluctuations in the age-structure.

3.2 The mean and variance function for discrete processes

Let us consider the stochastic geometric growth model $N_{t+1} = \Lambda_t N_t$, or equivalently the random walk $X_{t+1} = X_t + S_t$, where the Λ_t and $S_t = \ln \Lambda_t$

are independent with the same distribution. With initial population size $N_0 = \exp(X_0)$ at time 0 we find simply

$$X_t = X_0 + S_0 + S_1 + \ldots + S_{t-1}.$$

Hence, by the central limit theorem the distribution of X_t for a given X_0 is approximately normal with mean $X_0 + \mu t$ and variance νt, where $\mu = \mathrm{E}S_t$ and $\nu = \mathrm{var}(S_t)$. It is well known that this approximation is remarkably good, even for moderate values of t, which means that the form of the distribution of the S_t has practically no effect on the process X_t, only the expectation μ and the variance ν. We obtain a more general class of models of the type $X_{t+1} = X_t + S_t$, by allowing the distribution of S_t to depend on X_t. For this process let us write $\mu(x) = \mathrm{E}(S_t|X_t = x)$ and $\nu(x) = \mathrm{var}(S_t|X_t = x)$. In accordance with our remarks on the simple model with constant μ and ν, where the properties of the process is practically determined by these two parameters, we should expect that the functions $\mu(x)$ and $\nu(x)$ contains most of the information of the behavior of the general process. As an illustration we consider three processes with the same mean and variance functions, but with rather different distributions of the S_t for given values of X_t. Let the models be of the discrete logistic type $\mathrm{E}(\ln N_{t+1}|N_t = n) = \ln N_t + r(1 - n/K)$ or equivalently $\mathrm{E}(\Delta X_t|N_t = n) = \mathrm{E}(\Delta X_t|X_t = \ln n) = r(1 - n/K)$ giving

$$\mu(x) = r(1 - e^x/K)$$

where $x = \ln n$. We have seen in section 2.4 that for populations with small fluctuations around a carrying capacity which is not too small, the variance of ΔX_t conditioned on X_t is approximately equal to the environmental variance. In accordance with this we assume

$$\nu(x) = \sigma_e^2.$$

The model may now be written as

$$\Delta X = \mu(x) + \sqrt{\nu(x)}U$$

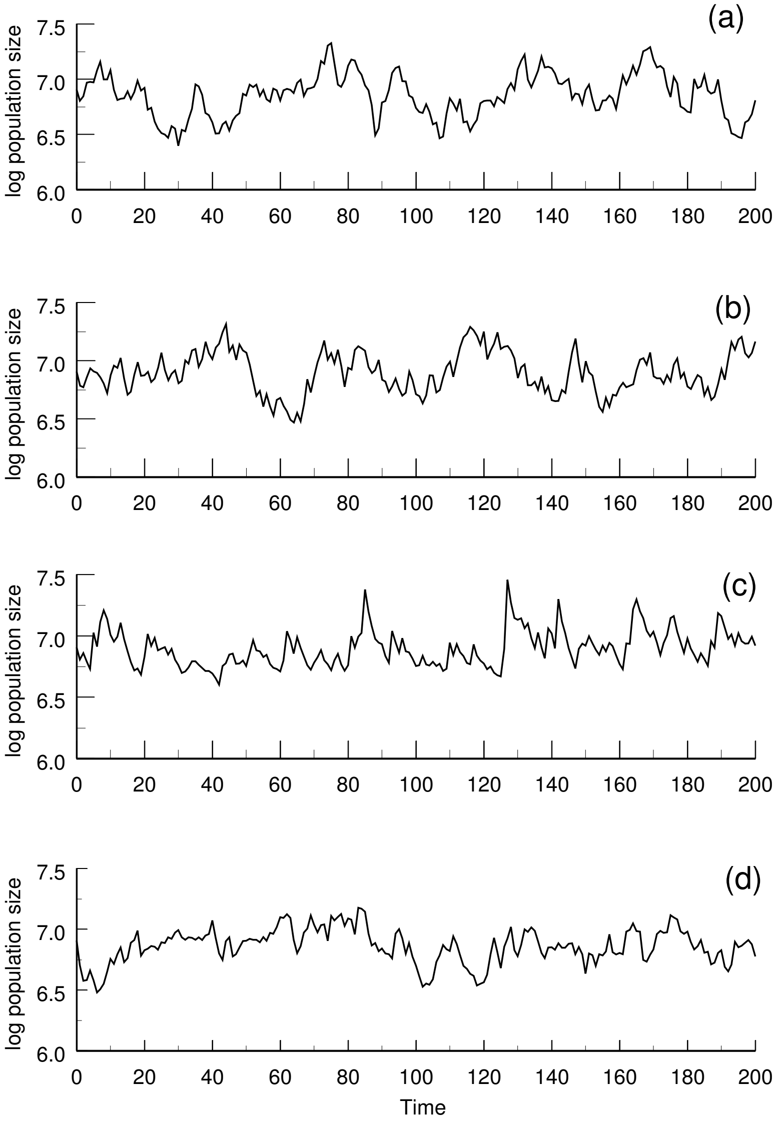

Figure 3.1: Population fluctuations for three models with the same discrete logistic type of dynamics with parameters $r = 0.2$, $K = 1000$, $\sigma_e^2 = 0.01$. The increments are modelled by different distributions: Normal distribution (a), Rectangular distribution (b), Exponential distribution (c) and the diffusion approximation recorded at discrete values with increments 1 (d).

where U is standardized so that $\mathrm{E}U = 0$ and $\mathrm{var}(U) = 1$.
Fig.3.1a-c shows simulations of this process when U is standardized normal, rectangular and exponential, respectively. Even if there is only one simulated process for each distribution we do get the impression that the fluctuations look fairly similar, especially for the normal and the rectangular, which both has zero skewness. For the exponential, which is skewed to the right, there is some tendency that the increases are somewhat quicker and the decreases somewhat slower than for the other distributions.
In the next section, we define diffusion processes, which is a class of processes that are continuous in the state variable as well as in time. The properties of such processes will be completely defined by the functions $\mu(x)$ and $\nu(x)$ which are called the *infinitesimal mean* and *infinitesimal variance* of the process. Together with possible boundary conditions, these functions completely define the diffusion process. It turns out that discrete processes often can be accurately approximated by diffusions with infinitesimal mean and variance equal to the mean and variance function of the discrete process.

3.3 The infinitesimal mean and variance of a diffusion

When the mean and variance functions are constants we have seen that the expectation as well as the variance of $X_t - X_0$ for a given X_0 are proportional to t, more precisely $\mathrm{E}(X_t - X_0 | X_0) = \mu t$ and $\mathrm{var}(X_t - X_0 | X_0) = \nu t$. The basic assumption of diffusions, apart from the Markov property the future depends only on the previous state), is that these relations hold for very small values of t, that is, for a small time interval Δt we assume $\mathrm{E}(\Delta X_t | X_t = x) \approx \mu(x)\Delta t$ and $\mathrm{var}(\Delta X_t | X_t = x) \approx \nu(x)\Delta t$. As Δt actually approaches zero we see that the last relation is equivalent to $\mathrm{E}[(\Delta X_t)^2 | X_t = x] \approx \nu(x)\Delta t$ because $[\mathrm{E}(X_t | X_t = x)]^2$ is of order (Δt) and vanish compared to terms of order Δt as Δt approaches zero. The precise mathematical definitions are that the limit of $\mathrm{E}(\Delta X_t | X_t = x)/\Delta t$ as Δt approaches zero is the infinitesimal mean $\mu(x)$, while the limit of $\mathrm{E}[(\Delta X_t)^2|X_t = x]/\Delta t$ is the infinitesimal variance $\nu(x)$. Together with some boundary conditions, for example an extinction barrier, these two functions $\mu(x)$ and $\nu(x)$ completely define the diffusion process.

Diffusion processes may be simulated by using small discrete time steps. For population processes time steps $\Delta t = 0.1$ seasons usually gives sufficient accuracy. If the state at time t is $X_t = x$ we simulate

$$X_{t+\Delta t} = x + \mu(x)\Delta t + U_t\sqrt{\nu(x)\Delta t},$$

where the U_t are independent standard normal variates. By this method we obtain $\mathrm{E}(\Delta X_t|X_t = x) = \mu(x)\Delta t$ and $\mathrm{var}(\Delta X_t|X_t = x) = \nu(x)\Delta t$.

As an illustration Fig.3.1d shows one simulation of this process, that is $\mu(x) = r(1 - e^x/K)$ and $\nu(x) = \sigma_e^2$, serving as an approximation to all three processes shown in Fig.3.1a-c. The diffusion approximation constructed in this way, by choosing the mean and variance functions of the discrete process as the infinitesimal mean and variance, is commonly referred to as the Ito approximation. More precisely, the method is based on first expressing the process by a stochastic differential equation and using the stochastic integral called Ito integration when solving the equation, which is equivalent to dealing with the above diffusion.

Suppose now that we rather than working with $X_t = \log N_t$ considered the diffusion approximation to N_t. This diffusion approximation would then have infinitesimal mean $\mu_N(n) = \mathrm{E}(\Delta N|N = n)$ and variance $\nu_N(n) = \mathrm{var}(\Delta N|N = n)$. It turns out that these two diffusions are not quite identical, but in practice fairly close if the changes in population size between years are not too large. In order to deal with processes on different scales, for example the log-scale and exponential scale, we need to know in some detail what happens when we change the scale of a diffusion. This is dealt with in some detail in section 3.5.

3.4 Boundary conditions

A diffusion is fully defined by its infinitesimal mean and variance together with some boundary conditions. In biology, the most actual boundary condition is defined by introducing an absorbing barrier at some value of N where the population actually goes extinct. Usually this extinction barrier is chosen at $N = 1$ or $N = 0$. When the population trajectory reaches the extinction barrier the population remains in this state. Hence, extinction barriers should only be used when modelling closed populations with no immigration from other populations.

Sometimes population models may also be defined by introducing a reflecting barrier. This barrier can never be crossed. Rather than crossing, the process is immediately reflected. Mathematically, a reflecting barrier at, say $n = a$, can be modelled by defining the infinitesimal mean and variance to be symmetric around $n = a$. More precisely, for $n > a$ we use the infinitesimal mean $\mu(n) = -\mu(2a-n)$ and variance $\nu(n) = \nu(2a-n)$ and treat the process as having no barrier at $n = a$. If the state of this process is $N_t > a$, we simply interpret this as if the state of the real process with reflecting barrier were $2a - N_t$.

Returning to the theta-logistic model we observe that as θ approaches infinity, the infinitesimal mean approaches $r_1 n$ for $n < K$, and the density dependence tends to be infinitely strong at, or immediately above K. Hence, in the limit $\theta = \infty$ this is a model for geometric growth with a reflecting barrier at $N = K$, a model we shall use to analyze the scaling of the time to extinction (see section 3.9). Fig.3.2 shows a simulation of this diffusion with a reflecting barrier at population size 1000.

3.5 Transformations

If N_t is a diffusion process, then also $X_t = g(N_t)$ is a diffusion for any function g which is twice differentiable. The most common change of scale in biology is the log transformation $g(n) = \ln n$, but other transformations, for

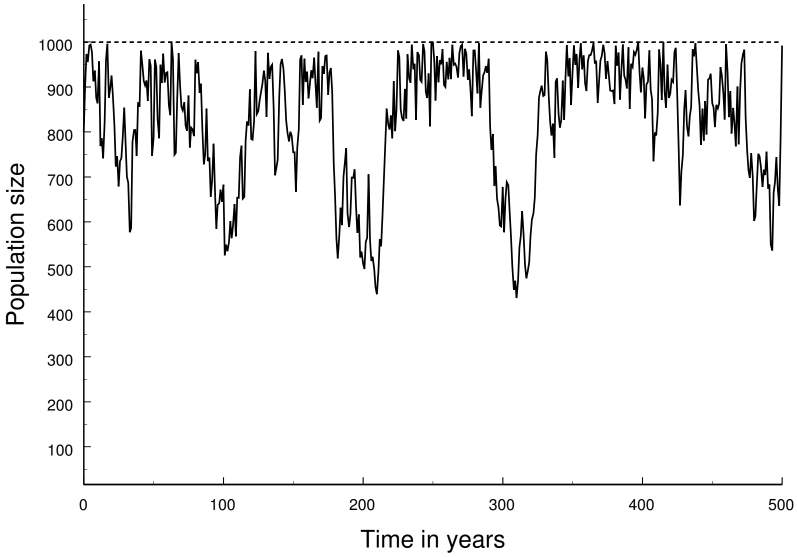

Figure 3.2: Simulation of the ceiling model with reflecting barrier at population size 1000. The parameters are $r_1 = 0.02$, $\sigma_e^2 = 0.01$.

example transformations that stabilizes the infinitesimal variance, may also be of interest (see section 3.6.3). Let $\mu(n)$ and $\nu(n)$ be the infinitesimal mean and variance of the diffusion N_t. Then, the infinitesimal mean and variance of $X_t = g(N_t)$ are

$$\mu_X(x) = g'(n)\mu(n) + \frac{1}{2}g''(n)\nu(n)$$

$$\nu_X(x) = g'(n)^2\nu(n)$$

where $n = g^{-1}(x)$, which is the inverse function or the solution of the equation $x = g(n)$ with respect to n.

For the log transformation we insert $g(n) = \ln n$, $g'(n) = 1/n$ and $g''(n) = -1/n^2$ giving infinitesimal mean $\frac{1}{n}\mu(n) - \frac{1}{2n^2}\nu(n)$ and variance $\frac{1}{n^2}\nu(n)$.

As a simple example of log transformation we consider the geometric Brownian motion which is a process growing exponentially with constant specific growth rate, that is, $\mu(n) = rn$. The environmental variance σ_e^2 is constant, and there is no demographic variance, corresponding to $\nu(n) = \sigma_e^2 n^2$. This is the diffusion approximation for the discrete process $N_{t+1} = \Lambda_t N_t$, where the Λ_t are independent identically distributed. More precisely, $r = \mathrm{E}\Lambda_t - 1$ and $\sigma_e^2 = \mathrm{var}\Lambda_t$. Inserting this infinitesimal mean and variance into the general formula for the log transformation given above, we find that the diffusion $X_t = \ln N_t$ is given by

$$\mu_X(x) = r - \frac{1}{2}\sigma_e^2$$

$$\nu_X(x) = \sigma_e^2.$$

This process is an example of a Brownian motion, which is a diffusion with constant infinitesimal mean and variance. In section 2.7 we defined the stochastic growth-rate of a discrete process as $\mathrm{E}\Delta X = \mathrm{E}\ln\Lambda$. Since $\mathrm{E}\Delta X \approx \mu_X(x)$, we see that the infinitesimal mean of X_t is the natural definition of the stochastic growth rate of a diffusion. Notice that the stochastic growth-rate of the geometric Brownian motion is not the growth rate r of N_t, but actually $s = r - \frac{1}{2}\sigma_e^2$. This reduction in the growth rate of $\frac{1}{2}\sigma_e^2$ due to stochasticity, is the diffusion analogy to the reduction in the growth on log-scale that we found for models in discrete time (see section 1.4 and Fig.1.2).

For the same geometric growth model with constant demographic as well as environmental variance we have $\mathrm{var}(\Delta N|N) = \sigma_d^2 N + \sigma_e^2 N^2$. The corresponding diffusion process has infinitesimal variance $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$. Applying the transformation formulas we find (exercise 1) that the infinitesimal mean and variance for the process $X_t = \ln N_t$ is $r - \frac{1}{2}\sigma_e^2 - \frac{1}{2n}\sigma_d^2$ and $\sigma_e^2 + \sigma_d^2/n$, or if expressed by x

$$\mu_X(x) = r - \frac{1}{2}\sigma_e^2 - \frac{1}{2}\sigma_d^2 e^{-x}$$
$$\nu_X(x) = \sigma_e^2 + \sigma_d^2 e^{-x}.$$

Hence, we find the same kind of reduction in the stochastic growth-rate due to environmental and demographic stochasticity as we found for discrete processes in 1.4. In this model the reduction due to environmental stochasticity is $\sigma_e^2/2$, while the reduction due to demographic stochasticity is $\sigma_d^2/(2n)$, which increases as n decreases in accordance with Fig.1.2. The infinitesimal variance increases as the populations size decreases provided that the demographic variance is positive. One important effect of this decrease in growth rate and increase in stochasticity for small populations is that the probability of extinction strongly increases with the demographic variance.

Notice that for populations with positive stochastic growth-rate for large values of n, that is, $s = r - \frac{1}{2}\sigma_e^2 > 0$, we may still have negative stochastic growth-rate for sufficiently small population sizes, actually an Allee effect, due to the term $\frac{1}{2n}\sigma_d^2$. The unstable equilibrium point at which the stochastic growth-rate changes from positive to negative is (exercise 2)

$$n^* = \frac{\sigma_d^2/2}{r - \sigma_e^2/2} = \frac{\sigma_d^2}{2s}.$$

Further examples of transformations are given in section 3.6.

3.6 * Some examples of transformations

3.6.1 Log-transformation of a model with Gompertz type of density-regulation

The diffusion approximation for the theta-logistic model for $\theta = 0$ when ignoring demographic stochasticity is

$$\mu(n) = r_1 n(1 - \ln n/\ln K)$$

$$\nu(n) = \sigma_e^2 n^2.$$

If the carrying capacity is large so that extinction is unlikely, it is legitimate to ignore the demographic variance. The density-regulation for this model is of the Gompertz type. Performing the log-transformation we find (exercise 3)

$$\mu_X(x) = r_1(1 - \ln n/\ln K) - \sigma_e^2/2$$

$$\nu_X(x) = \sigma_e^2,$$

where $n = e^x$. The infinitesimal mean may be written on the form

$$\mu_X(x) = \alpha - \beta x$$

where $\alpha = r_1 - \sigma_e^2/2$ and $\beta = r_1/\ln K$. This model, which has constant infinitesimal variance and a linear expression with negative coefficient of x for the infinitesimal mean, is called the Ornstein-Uhlenbeck process (OU-process). Some important results for this process will be given in section 3.8.4 and 3.12.

3.6.2 Transformations of the theta-logistic model

Ignoring demographic stochasticity the theta-logistic model may be written on the form

$$\mu(n) = rn(1 - \frac{n^\theta}{K^\theta})$$

$$\nu(n) = \sigma_e^2 n^2$$

for $\theta \neq 0$. We introduce the transformation $X_t = g(N_t) = N_t^\theta$. Using $g'(n) = \theta n^{\theta-1}$, $g''(n) = \theta(\theta-1)n^{\theta-2}$, and applying the transformation formulas, we find (exercise 4)

$$\mu_X(x) = r\theta x[1 + \frac{1}{2}(\theta - 1)\sigma_e^2 - \frac{x}{K^\theta}]$$

$$\nu_X(x) = \theta^2\sigma_e^2 x^2,$$

which is a logistic model with constant environmental variance $\sigma_e^2\theta^2$, growth-rate $r\theta[1+\frac{1}{2}(\theta-1)\sigma_e^2]$, and carrying capacity $K^\theta[1+\frac{1}{2}(\theta-1)\sigma_e^2]$.
Another interesting transformation of the theta-logistic model, still assuming $\theta \neq 0$, is $Y_t = N_t^{-\theta}$ that leads to the diffusion (exercise 5)

$$\mu_Y(y) = \frac{r\theta}{K^\theta} - y\theta[r - \frac{1}{2}(\theta+1)\sigma_e^2]$$
$$\nu_Y(y) = \theta^2\sigma_e^2 y^2.$$

We shall see later (section 3.10) that the simple linear form of the infinitesimal mean is of particular interest since it leads to a very simple expression for the autocorrelation function of the process.

3.6.3 Isotrophic scale transformation.

A more general expression for the stochastic Allee effect due to demographic stochasticity was derived by Lande (1998). Rather than dealing with the stochastic growth rate obtained by performing the log transformation, Lande considered the transformation to a scale with isotropic noise, or a constant infinitesimal variance. From the transformation formula $\nu_X(x) = g'(n)^2\nu(n)$ we see that we obtain $\nu_X(x) = 1$ by choosing $g'(n) = 1/\sqrt{\nu(n)}$ giving

$$g(n) = \int^n \frac{dz}{\sqrt{\nu(z)}}.$$

which inserted into the transformation formula for the infinitesimal mean gives

$$\mu_X(x) = [\mu(n) - \frac{1}{4}\nu'(n)]/\sqrt{\nu(n)}.$$

Hence, the growth-rate on this scale changes sign at $\tilde{N}$ given by

$$\mu(\tilde{N}) = \nu'(\tilde{N})/4.$$

Since we are here interested primarily in small population effects we consider a population well below its carrying capacity so that $\mu(n) = rn$. If the demographic and environmental variances are constants we have $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$ giving $\nu'(n) = \sigma_d^2 + 2n\sigma_e^2$. Inserting this in the equation for $\tilde{N}$, we find

$$\tilde{N} = \frac{\sigma_d^2/4}{r - \sigma_e^2/2} = \frac{\sigma_d^2}{4s},$$

where $s = r - \sigma_e^2/2$ is the stochastic growth rate for large population sizes, which is exactly half the value of the unstable equilibrium point based on the stochastic growth-rate defined by the log-transformation for the same model.
For an explicit expression for the transformation $X = g(N)$ see exercise 7.

3.7 Populations modeled by Brownian motions and OU-process

A Brownian motion is a diffusion with constant infinitesimal mean and variance. A number of results, such as the probability of reaching barriers, the distribution of future states, the distribution of the time to extinction, and predictors for extinction, are available for this model, and will be dealt with in section 3.8.2, 3.8.3, and 3.9.

We have previously seen that the diffusion approximation for the discrete multiplicative population model with no density-regulation is

$$\mu(n) = rn$$

$$\nu(n) = \sigma_e^2 n^2,$$

provided that the population size is large enough for the demographic stochasticity to be ignored. Performing the log-transformation $X_t = \ln N_t$ we have seen that X_t is a diffusion with infinitesimal mean $r - \sigma_e^2/2$ and variance σ_e^2. Since these are constants not depending on n the process X_t is a Brownian motion and the results for these processes are available for analyzing the process X_t.

Another class of diffusions already mentioned for which a number of interesting results are available (see section 3.8.4 and 3.12) is the Ornstein-Uhlenbeck process defined by a constant infinitesimal variance and a linear expression for the infinitesimal mean. This type of process may serve as good approximations to a number of models for populations fluctuating around a carrying capacity that is large enough for the possibility of extinctions to be ignored. For models with constant environmental variance we have seen that the infinitesimal variance of $\ln X_t$ may be approximately constant for large population sizes, actually equal to σ_e^2, and the infinitesimal mean is $\mu(n)/n - \sigma_e^2/2$. If we now replace the infinitesimal mean by its linearization around $x = \ln K$, where K is defined by $\mu(K) = 0$, we obtain an Ornstein-Uhlenbeck process that can be used as an approximation (exercise 8)

$$\mu_X(x) \approx \mu'(K)(x - \ln K) - \sigma_e^2/2$$

$$\nu_X(x) \approx \sigma_e^2.$$

As an example let us perform this approximation for the theta-logistic model

$$\mu(n) = r_1 n[1 - \frac{n^\theta - 1}{K^\theta - 1}]$$

giving

$$\mu'(K) = -\frac{r_1\theta}{1 - K^{-\theta}} = -r\theta$$

and the linear approximation

$$\mu_X(x) \approx (r\theta \ln K - \sigma_e^2/2) - r\theta x.$$

We see that to this order of approximation the process depends on r and θ only by the product $r\theta$. As a consequence, we cannot expect to be able to estimate r and θ separately from time series observations from such populations. Actually, some more detailed knowledge of the populations behavior at small population sizes will often be required.

Notice that in the limit as θ approaches zero we find $\mu'(K) = r_1/\ln K$, giving exactly the same process as we found in 3.6.1, which should be no surprise since the process X_t for $\theta = 0$ is exactly the Ornstein-Uhlenbeck process.

Fig.3.3 shows a simulation of the theta-logistic diffusion together with the approximation by the Ornstein-Uhlenbeck process.

3.8 Computations in diffusion models

3.8.1 The Green function and related functions

Some important results for diffusions may be expressed by the Green function $G(x, x_0)$ for the process. The physical interpretation of this function is the following: Consider a diffusion X_t with infinitesimal mean and variance $\mu(x)$

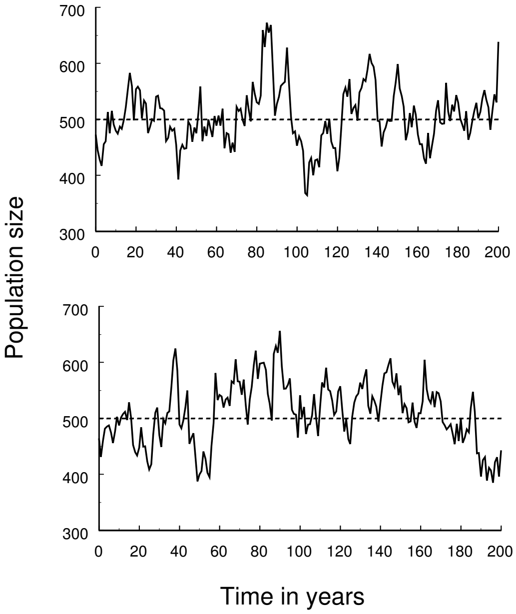

Figure 3.3: Simulation of the theta-logistic model with $r_1 = 0.1$, $K = 500$, $\sigma_e^2 = 0.005$ and $\theta = 2$ (upper panel) and the same process approximated by an Ornstein-Uhlenbeck process (lower panel).

and $\nu(x)$, an absorbing barrier at a and b, $a < b$, and initial state $X_0 = x_0$ between a and b. Then, the expected time the process spends in the interval $(x, x + \Delta x)$ before it goes extinct is $G(x, x_0)\Delta x$ in the limit as Δx tends to zero. Hence, by the additive property of expectations, the expected time the process spends in an interval (c, d) is $\int_c^d G(x, x_0)dx$.

The Green function may in general be expressed by the functions

$$s(x) = \exp[-2\int^x \frac{\mu(z)}{\nu(z)}dz]$$

$$S(x) = \int^x s(z)dz,$$

where the lower limit of the integration may be chosen arbitrarily since the Green function will not depend on them. In order of simplifying the notation we also define the function

$$m(x) = \frac{1}{\nu(x)s(x)}$$

The general expression for the Green function is then

$$G(x, x_0) = \begin{cases} 2\frac{[S(x)-S(a)][S(b)-S(x_0)]}{S(b)-S(a)}m(x) & \text{for } a \leq x \leq x_0 \leq b \\ 2\frac{[S(b)-S(x)][S(x_0)-S(a)]}{S(b)-S(a)}m(x) & \text{for } a \leq x_0 \leq x \leq b. \end{cases}$$

For realistic biological models with density regulation $\mu(x)$ is negative for large values of x so that $S(b)$ approaches infinity as b increases towards infinity. Choosing $b = \infty$ and the extinction barrier a as the lower limit of integration we obtain the much simpler expression

$$G(x, x_0) = \begin{cases} 2m(x)S(x) & \text{for } a \leq x \leq x_0 \\ 2m(x)S(x_0) & \text{for } a \leq x_0 \leq x. \end{cases}$$

3.8.2 The probability of ultimate extinction

Starting at some x_0 between a and b, the probability that the process reaches b before a is

$$u(x_0) = \frac{S(x_0) - S(a)}{S(b) - S(a)}.$$

Consider first density dependent models with exponential growth with positive growth-rate close to the extinction barrier a, positive environmental variance, and no demographic variance. For such models the probability $u(x_0)$ will tend to 1 as a approaches zero for any finite b (exercise 9). This means that this process can never go extinct if $a = 0$. On the other hand, if we either include a positive demographic variance, or choose $a = 1$ or some other positive value, then $u(x_0)$ approaches zero as b tends to infinity for density-regulated populations, which means that ultimate extinction is a certain event (exercise 10).

For density independent models the situation is a different one. Generally, if the stochastic growth-rate is negative, the ultimate extinction is certain. For positive growth-rates the process may actually be absorbed at infinity, and the probability of ultimate extinction can be found from the above expression for $u(x_0)$. For models with constant demographic and environmental variances the solutions are fairly simple.

Consider the geometric Brownian motion N_t with infinitesimal mean $\mu_N(n) = rn$ and variance $\nu_N(n) = \sigma_e^2 n^2$. We have seen that $X_t = \ln N_t$ then is a Brownian motion with drift parameter $\mu(x) = s = r - \sigma_e^2/2$ and infinitesimal variance $\nu(x) = \sigma_e^2$. We choose the extinction barrier for N_t at 1 which corresponds to choosing $a = 0$ for the process X_t. For this model (exercise 12) it is straightforward to perform the integrations, using $a = 0$ as the lower integration limit, giving

$$s(x) = e^{-\frac{2s}{\sigma_e^2}x},$$

which integrates to (exercise 12)

$$S(x) = \begin{cases} \frac{\sigma_e^2}{2s}(1 - e^{-\frac{2s}{\sigma_e^2}x}) & \text{for } s \neq 0 \\ x & \text{for } s = 0, \end{cases}$$

Then, as the upper absorbing barrier tends to infinity we find

$$u(x_0) = \begin{cases} 0 & \text{for } s \leq 0 \\ 1 - e^{-\frac{2s}{\sigma_e^2}x_0} & \text{for } s > 0, \end{cases}$$

which means that this process actually may be absorbed at infinity if s is positive. Transforming back to the original geometric Brownian motion N_t with initial state n_0 we see that ultimate extinction at $N = 1$ happens with probability 1 if $s \leq 0$, and with probability $n_0^{-\frac{2s}{\sigma_e^2}}$ if $s > 0$.

For the more general diffusion model for exponential growth that also takes demographic stochasticity into account (exercise 13) we use $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$, giving

$$s(n) = (\frac{\sigma_d^2 + \sigma_e^2}{\sigma_d^2 + \sigma_e^2 n})^{\frac{2r}{\sigma_e^2}},$$

which is easily integrated to give

$$S(n) = \begin{cases} \frac{\sigma_d^2+\sigma_e^2}{2s}[1 - (\frac{\sigma_d^2+\sigma_e^2 n}{\sigma_d^2+\sigma_e^2})^{-\frac{2s}{\sigma_e^2}}] & \text{for } s \neq 0 \\ \frac{\sigma_d^2+\sigma_e^2}{2s} \ln(\frac{\sigma_d^2+\sigma_e^2 n}{\sigma_d^2+\sigma_e^2}) & \text{for } s = 0. \end{cases}$$

Inserting this in the general expression with $b = \infty$, we still find that the probability of ultimate extinction is 1 if $s \leq 0$, while for $s > 0$ we find that ultimate extinction occur with probability $(\frac{\sigma_d^2+\sigma_e^2}{\sigma_d^2+\sigma_e^2 n_0})^{2s/\sigma_e^2}$. For $\sigma_d^2 = 0$ this expression is equivalent to what we found for the geometric Brownian motion, as expected. The effect of σ_d^2 on the probability of ultimate extinction is illustrated in Fig.3.4.

If the initial population size is small, the environmental variance may sometimes be ignored. The probabilities of ultimate extinction may then be found by considering the limits of the above expressions as σ_e^2 tends to zero, which is 1 for $s < 0$ and otherwise $e^{-2s(n_0-1)/\sigma_d^2}$ (exercise 14).

3.8.3 The expected time to extinction and some related results

Since $G(x, x_0)\Delta x$ expresses the expected time the process X_t is in the interval $(x, x+\Delta x)$ before it is absorbed at a or b, the expected time to absorbtion is found by adding the contributions from all intervals. As Δx approaches zero

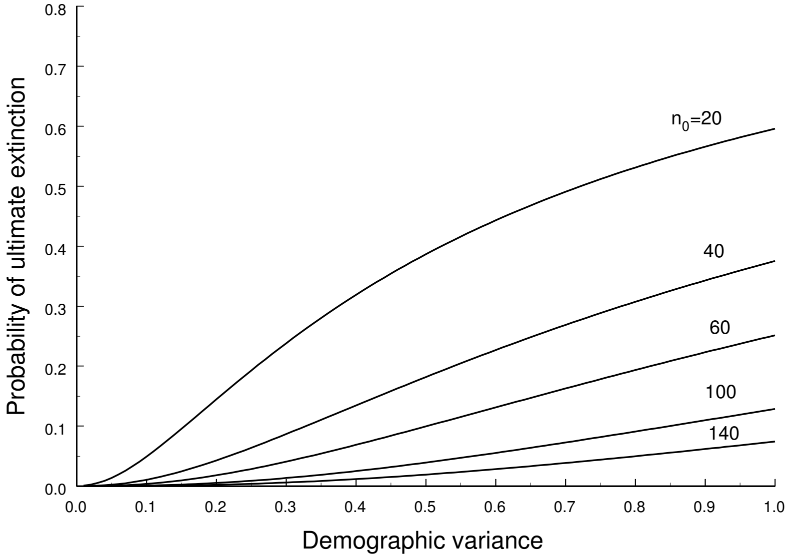

Figure 3.4: The probability of ultimate extinction as function of the demographic variance for different initial population sizes. The diffusion has infinitesimal mean $\mu(n) = rn$ with $r = 0.015$ and constant environmental and demographic variance, $\sigma_e^2 = 0.01$.

we then obtain the exact expression for the expected time to extinction as

$$\mathrm{E}T = \int_a^b G(x, x_0)dx.$$

This integral can be solved analytically only for some very special models, and must otherwise be solved numerically. For the Brownian motion with drift s and variance σ_e^2 we find for $a = 0$ and $b = \infty$ that $s(x) = \exp(-\alpha x)$ and $S(x) = [1 - \exp(-\alpha x)]/\alpha$, where $\alpha = 2s/\sigma_e^2$. For $s \geq 0$ we find that the integral of the Green function diverge, corresponding to $\mathrm{E}T = \infty$. For $s < 0$ we find

$$G(x, x_0) = \begin{cases} \frac{2}{\sigma_e^2 \alpha}(e^{\alpha x} - 1) & \text{for } x \leq x_0 \\ \frac{2}{\sigma_e^2 \alpha}(1 - e^{-\alpha x_0})e^{\alpha x} & \text{for } x > x_0, \end{cases}$$

The integral from zero to infinity of this Green function is $-2x_0/(\sigma_e^2 \alpha)$ or

$$\mathrm{E}T = -x_0/s.$$

For biological populations with density regulation we plug in $a = 1$, $b = \infty$ in the simplified expression for the Green function given in section 3.8.1 giving

$$\mathrm{E}T = \int_1^{n_0} 2m(n)S(n)dn + \int_{n_0}^{\infty} 2m(n)S(n_0)dn,$$

which can be evaluated by numerical integrations. The analysis of the time to extinction is dealt with in more detail in section 3.11.

The Green function may also be applied to find a more general result that may be quite useful in some practical applications. Let h be any function and consider the expectation

$$v(x_0) = \mathrm{E}[\int_0^T h(X_t)dt],$$

where x_0 is the initial state at time $t = 0$, and T is the time at which the process is absorbed. Then $v(x_0)$ may be calculated from the Green function (exercise 15) giving

$$v(x_0) = \int_a^b h(x)G(x, x_0)dx.$$

Notice that if $h(x) = 1$, then $v(x_0) = \mathrm{E}T$, so the result for ET given above is a special case. We shall apply this result in chapter 5 when we deal with the effects of harvesting. Consider a diffusion N_t and assume that at a small time step $(t, t + \Delta t)$ we harvest an amount $h(N_t)\Delta t$. Then the total harvest before the population goes extinct is $\int_0^T h(N_t)dt$, and the expected value of this harvest may be evaluated using the above result.

3.8.4 Predictions and stationary distributions

For a given diffusion X_t with known infinitesimal mean and variance $\mu(x)$ and $\nu(x)$, suppose that $X_0 = x_0$ at time zero has been observed. Then the population size X_t at some future point of time is a stochastic variable with distribution, say $f_t(x; x_0)$. For most models simple analytical expressions for this distribution are not available. However, for the Brownian motion and the Ornstein-Uhlenbeck process, the distribution is normal with known mean and variance provided that the possibility of extinction is ignored. .

Using our previous notation for the logarithm of a geometric Brownian motion, we consider the Brownian motion with drift parameter s and infinitesimal variance σ_e^2. Then, if there are no absorbing barriers, the distribution of X_t given $X_0 = x_0$ is normal with mean $x_0 + st$ and variance $\sigma_e^2 t$. This distribution is also known in the case of an extinction barrier. This will be dealt with in section 3.11.3.

For the Ornstein-Uhlenbeck process with infinitesimal mean $\alpha - \beta x$, $\beta > 0$, infinitesimal variance $\nu(x) = \sigma^2$, and no extinction barrier, the distribution of X_t is also normal with

$$\mathrm{E}(X_t|X_0 = x_0) = \alpha/\beta + (x_0 - \alpha/\beta)e^{-\beta t}$$

$$\mathrm{var}(X_t|X_0 = x_0) = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}).$$

We see that the expectation as well as the variance tend to a limit as t approaches infinity. Hence, if t is sufficiently large, there is no information left in the observation $X_0 = x_0$, and the process will go on fluctuating forever around the mean α/β with variance $\sigma^2/(2\beta)$. Such processes are called stationary processes, and the distribution of the state for large values of t, $f(x) = f_\infty(x; x_0)$, is called the stationary distribution of the process. Accordingly, the Ornstein-Uhlenbeck process has a stationary distribution which is normal with mean α/β and variance $\sigma^2/(2\beta)$. Notice that large values of β, corresponding to strong density regulation around K when applied to population modeling, has two important effects. Firstly, a large β has the effect that the stationary distribution is reached quickly, and secondly, the variance of the stationary distribution will be small which means that the fluctuations around the mean are small. For processes with fairly small fluctuations around the carrying capacity we have seen that the Ornstein-Uhlenbeck process may serve as an approximation. Hence, the corresponding stationary distribution may be used to analyze how the different population parameters affect the population fluctuations.

Though there exists no simple general expression for the distribution $f_t(x; x_0)$, the general expression for the stationary distribution of a stationary diffusion is fairly simple, actually

$$f(x) = m(x)[C_1 S(x) + C_2],$$

where C_1 and C_2 are constants that must be determined from the boundary conditions, and $m(x)$ and $S(x)$ are defined as before. For the type of population process dealt with here, the integral of $m(x)S(x)$ over all accessable values of x will be infinite, which means that C_1 must be zero since $f(x)$ should integrate to 1. Hence, the stationary distribution is simply proportional to $m(x)$, that is

$$f(x) = m(x) / \int_a^\infty m(z)dz,$$

where a is an the extinction barrier that can never be reached or some natural lower bound for the process. If the integral in the denominator is finite so that this scaling is possible, then the process is stationary. We have previously seen that this is not the case if we include the demographic component of the infinitesimal variance.

In section 3.6.1 we saw that the model for N_t with the Gompertz type of density regulation and no demographic stochasticity could be transformed to an Ornstein-Uhlenbeck process by the log transformation. With the notation given above the parameters in the transformed process were $\alpha = r_1 - \sigma_e^2/2$, $\beta = r_1/\ln K$, and $\sigma^2 = \sigma_e^2$, where r_1 were the specific growth rate at $N_t = 1$. It follows that the stationary distribution of $X_t = \ln N_t$ is normal with mean $[1 - \sigma_e^2/(2r_1)]\ln K$ and variance $\sigma_e^2 \ln K/(2r_1)$. By transforming back we see that the stationary distribution of N_t is the lognormal distribution. Using the expressions for the mean and variance of the lognormal given in 1.3 we find the mean and the coefficient of variation of the stationary distribution is

$$\mathrm{E}N = K^{1-\frac{\sigma_e^2}{4r_1}}$$

$$C(N) = SD(N)/\mathrm{E}N = [K^{\frac{\sigma_e^2}{2r_1}} - 1]^{1/2}.$$

We see that increasing environmental stochasticity decreases the mean and increases the coefficient of variation of the stationary distribution in this model.

3.9 Some stationary distributions

3.9.1 The logistic model

For the logistic model with no demographic stochasticity the infinitesimal mean is $\mu(n) = rn(1 - n/K)$ and the variance $\nu(n) = \sigma_e^2 n^2$. Since σ_d^2 is ignored $n = 0$ is a natural lower barrier for the process that cannot be reached. From the definition of $s(n)$ we then find

$$s(n) = \exp(-2\int^n \frac{rx(1 - x/K)}{\sigma_e^2 x^2} dx) = kn^{-\frac{2r}{\sigma_e^2}} e^{\frac{2r}{\sigma_e^2 K} n},$$

where k is some constant depending on the lower integration limit. Then, using the definition $m(n) = 1/[s(n)\nu(n)]$ and performing the appropriate scaling, we see that the solution is the gamma distribution

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

where the scale parameter $\alpha = 2r/(\sigma_e^2 K)$ and the shape parameter is $\beta = 2s/\sigma_e^2$, and $s = r - \sigma_e^2/2$ as before. Using the well known expression for the mean and variance of the gamma distribution we find

$$\mathrm{E}N = \beta/\alpha = K(1 - \frac{\sigma_e^2}{2r})$$

$$\mathrm{var}N = \beta/\alpha^2 = \frac{\sigma_e^2}{2r} K^2 (1 - \frac{\sigma_e^2}{2r}).$$

3.9.2 The theta-logistic model

Let us again consider the more general theta-logistic model without demographic stochasticity, writing

$$\mu(n) = rn[1 - (n/K)^\theta]$$

$$\nu(n) = \sigma_e^2 n^2,$$

where $r = r_1/(1 - K^{-\theta})$. Performing the integration in the expression for $s(n)$ we find that the stationary distribution for $\theta \neq 0$ takes the form

$$f(n; K, \alpha, \theta) = C n^{\alpha-1} e^{-\frac{(\alpha+1)}{\theta}(n/K)^\theta}$$

where $\alpha = 2r/\sigma_e^2 - 1 = 2s/\sigma_e^2$, which must be assumed positive. Introducing the new variable $y = (n/K)^\theta$ and integrate we find the complete distribution

$$f(n; K, \alpha, \theta) = \frac{|\theta|(\frac{\alpha+1}{\theta})^{\alpha/\theta}}{K\Gamma(\alpha/\theta)} (n/K)^{\alpha-1} e^{-\frac{(\alpha+1)}{\theta}(n/K)^\theta} \quad \text{for} \quad \theta \neq 0,$$

which has moments

$$\mathrm{E}N^p = \frac{K^p \Gamma(\frac{\alpha+p}{\theta})}{\Gamma(\frac{\alpha}{\theta})(\frac{\alpha+1}{\theta})^{p/\theta}}$$

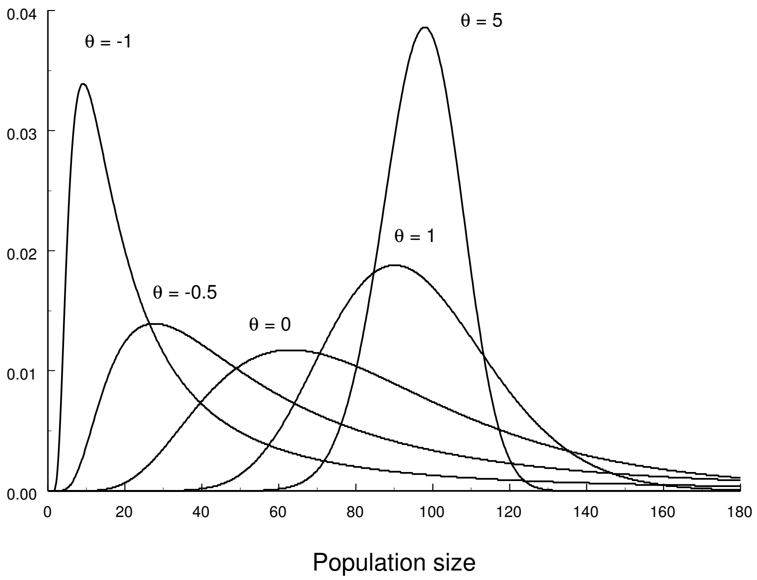

Figure 3.5: The stationary distribution for the theta-logistic model, the generalized gamma distribution, for different values of θ. The other parameters are $K = 100$, $\sigma_e^2 = 0.01$, and $r_1 = 0.1$.

for $p = 1, 2, \ldots$.. From this expression the mean and variance may be computed. This distribution is called the *generalized gamma distribution* and usually written on a different parametric form in the literature.

For $\theta = 1$ this is the gamma distribution with shape parameter α and scale parameter $(\alpha + 1)/K$ in accordance with 3.9.1. Notice that K is a scale parameter of the distribution $f(n; K, \alpha, \theta)$, while α and θ are shape parameters. If $\theta = \alpha$ the distribution is the *Weibull distribution* and if $\theta = -1$ it is the *inverse gamma distribution.*

To confirm that the limiting distribution as θ approaches zero is the lognormal in accordance with 3.8.4 we introduce $z = n/K$, and observe that the distribution of z is proportional to $z^{\alpha-1}e^{-(\frac{\alpha+1}{\theta})z^{\theta}}$. Expanding the exponent at $\theta = 0$ we find

$$\exp[-(\alpha+1)/\theta - 2\ln(z) - \frac{1}{2}(\alpha+1)\theta\ln(z)^2 + ...].$$

Absorbing the constant in the constant factor of the distribution and observing that $(\alpha+1)\theta$ approaches $\beta = \frac{2r_1}{\sigma_e^2 \ln K}$ as θ tends to zero, we find that the limiting distribution is proportional to $(1/z)\exp[-\ln(z) - \frac{1}{2}\beta\ln(z)^2]$ or proportional to $(1/z)\exp[-\frac{1}{2}\beta(\ln(z)+1/\beta)^2]$. Hence, the limiting distribution of z is the lognormal distribution, and the corresponding distribution of $\ln(z)$ is $N[-1/\beta, 1/\beta]$. Finally, since $\ln n = \ln z + \ln K$ we see that the limiting distribution of $\ln(N_t)$ is normal with mean $\ln(K)[1-\sigma_e^2/(2r_1)]$ and variance $\sigma_e^2\ln(K)/(2r_1)$ in accordance with what we previously have found.

In the limit as θ approaches infinity we obtain the model with no density-regulation below K and a reflecting barrier at K. We find by studying the limit (exercise 16) that the stationary distribution for this model is

$$f(n; K, \alpha, \infty) = \frac{\alpha}{K^\alpha} n^{\alpha-1}$$

for $0 \le n \le K$, and otherwise zero.

The stationary distribution for different values of θ, keeping K, r_1 and σ_e^2 constant, are shown in Fig.3.5.

3.9.3 The Beverton-Holt model

The Beverton-Holt model is usually associated with age-structured populations. If individuals reach the adult state after one year and their vital rates are not age-dependent the model assumes the individuals on average produce $\alpha/(1+\beta n)$ individuals the next season when the population size is n. If the adult survival is s the expected population size next year is $n[\alpha/(1+\beta n)+s]$ so that

$$\mathrm{E}(\Delta N|N=n) = \mu(n) = n[\alpha/(1+\beta n) + s - 1].$$

The carrying capacity of the deterministic model determined by $\mu(K) = 0$ is accordingly

$$K = \frac{\alpha+s-1}{\beta(1-s)}.$$

Assuming large population sizes and constant environmental variance σ_e^2 the diffusion approximation has infinitesimal mean and variance $\mu(n)$ and $\nu(n) = \sigma_e^2 n^2$, respectively. This leads to

$$s(n) = \exp\left[-\frac{2}{\sigma_e^2}\int^n \left(\frac{\alpha}{n(1+\beta n)} + \frac{s-1}{n}\right) dn\right].$$

Integrating using partial fractions then gives

$$s(n) = \exp\left[-\frac{2(\alpha+s-1)}{\sigma_e^2}\ln n + \frac{2\alpha}{\sigma_e^2}\ln(1+\beta n)\right]$$

giving a stationary distribution

$$f(n) = cn^{2(\alpha+s-1)/\sigma_e^2}(1+\beta n)^{-2\alpha/\sigma_e^2} n^{-2}.$$

This has the form of a beta distribution of the second kind

$$f(n) = \frac{\Gamma(p+q)\beta^p}{\Gamma(p)\Gamma(q)} \frac{n^{p-1}}{(1+\beta n)^{p+q}}$$

where $p = 2(\alpha + s - \sigma_e^2/2 - 1)/\sigma_e^2$ and $q = 2(1 + \sigma_e^2/2 - s)/\sigma_e^2$. The mean and variance of this distribution is $p/[\beta(q-1)]$ and $p(p+q-1)/[\beta^2(q-1)^2(q-2)]$. Inserting the expressions for p and q we find the expected population size

$$\mathrm{E}N = \frac{\alpha + s - 1 - \sigma_e^2/s}{\beta(1-s)}$$

which equals K for $\sigma_e^2 = 0$ as expected. The squared coefficient of variation is

$$C_N^2 = \frac{\sigma_e^2(\alpha - \sigma_e^2/2)}{2(1-s-\sigma_e^2/2)(\alpha+s-1-\sigma_e^2/2)}.$$

3.10 Quasi-stationary distributions

If the demographic stochasticity is included in the model, or if the extinction barrier is chosen at $N_t = 1$, the probability of ultimate extinction will usually be one, which means that the process is not stationary and no stationary distribution exists. However, most populations are in any case likely to go on fluctuating around the carrying capacity for a very long time. We have seen in section 3.8.1 that the Green function then expresses the expected time the process spends at each state. Further, the expected time until extinction, which is the integral of the Green function from the extinction barrier to infinity, is then finite. If we then scale the Green function by its integral, we obtain a distribution, actually

$$q(n; n_0) = G(n, n_0) / \int_a^\infty G(z, n_0) dz,$$

where a is the extinction barrier. This distribution is called the quasi-stationary distribution of the process. Notice that this distribution, contrary to the stationary distribution, depends on the initial state n_0, though this dependence often in practice will be very week.

The quasi-stationary distribution is not really the distribution of some stochastic variable in the model. However, it expresses the expected properties of the population fluctuations from the initiation at n_0 and up to the time of extinction. Properties of this distribution, such as the mean and the variance are very useful quantities in the characterization of the process. Actually, the interpretation of the distribution is also very close to the practical interpretation of a stationary distribution by the following argument: The ratios $q(n_1; n_0)/q(n_2; n_0)$ and $f(n_1)/f(n_2)$ for some two states n_1 and n_2, both expresses the ratio between the expected time the process spends in the intervals $(n_1, n_1 + \Delta n)$ and $(n_2, n_2 + \Delta n)$ as Δn approaches zero. The difference is that for a stationary process this refers to an infinite time interval while for the quasi-stationary distribution the ratio refers to the finite time interval from initiation and up to extinction. Notice also, by the definition of the Green function for biological processes given in section 3.8.1, that both distributions have exactly the same shape for $n > n_0$ if the parameters other than the extinction barrier are the same, since they are both proportional to $m(n)$. For $n < n_0$, however, the stationary distribution is proportional to $m(n)$ while the quasi-stationary distribution is proportional to $m(n)S(n)$.

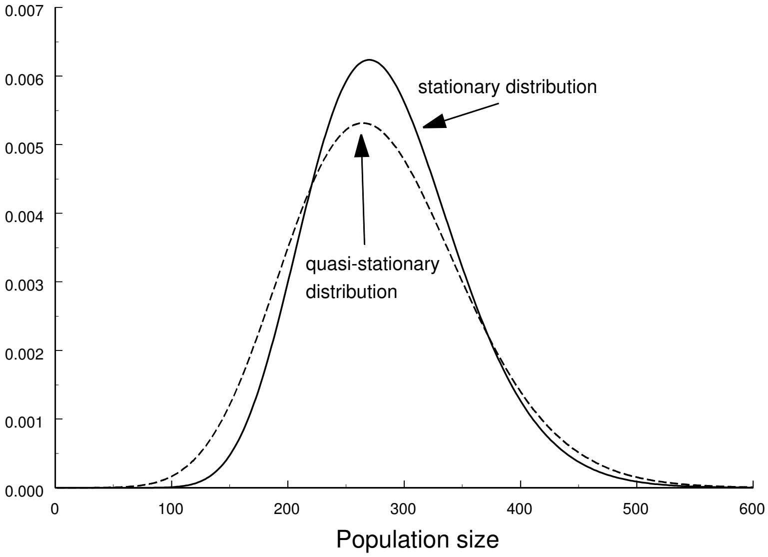

Figure 3.6: The stationary distribution, the gamma distribution, for the logistic model with absorbing barrier at zero and parameters $r = 0.1$, $K = 300$ and $\sigma_e^2 = 0.01$ (solid line) together with the quasi-stationary distribution with $N_0 = K$ obtained by adding demographic variance $\sigma_d^2 = 1$ to the model.

In section 3.9.1 we showed that the stationary distribution for the logistic model with extinction barrier at zero and $\sigma_d^2 = 0$ is the gamma distribution. In Fig.3.6 this stationary distribution is depicted together with the quasi-stationary distributions for the non-stationary case of the same model obtained by choosing $\sigma_d^2 = 1$ and $n_0 = K$.

3.11 Extinction and population viability

3.11.1 Definitions of population viability

Viability of populations is judged according to many different criteria, using all obtainable information of the species. One such criterium is based on considering the probability that the populations goes extinct during a given time interval. Commonly populations are classified as viable if the risk of extinction within one hundred years is less than 10%. If the parameters describing the population fluctuations, the mean specific growth rate and the carrying capacity are known, this criteria for viability may alternatively be expressed by the model parameters and the initial population size. In this section we present some results for the diffusion approximation that may be applied to perform such characterizations. More generally, the above criteria may be expressed as

$$P(T > t_v) > 1 - \alpha_v$$

so that the above criterium is obtained by choosing $t_v = 100$ years and $\alpha_v = 0.1$.

For populations that are far below their carrying capacity, which is often the case when populations are threatened with extinction, a geometric growth model may be applied. The main problem will then be to compute the minimum population size n_0 required for the population to be viable according to the above definition. On the other hand, fragmentation or other human activities may have the effect on the population parameters, in particular the carrying capacity. One may then be interested in linking the concept of viability to this population parameter, assuming that the initial population size is close to K, and compute the minimum value of K required to make the population viable.

We emphasize that we make no statistical analysis in this chapter, but only theoretical characterizations of the concept of viability. A full viability analysis requires that parameters are estimated from time series data and/or data on individual survival and reproduction. Since the amount of data, in particular the number of years a time series is recorded, is usually small, uncertainties in the estimates may be a major problem. For example, the estimator of the minimum K required for viability, or the estimate of the probability of extinction, will have uncertainties that are so large that the characterization of the population as viable or not viable can not really be done in practise. We shall see in chapter 5 that these problems require that the viability analysis in practice should be approached by using the concept of statistical predictions.

3.11.2 The exponential approximation for density regulated populations

We have seen that for populations which are density regulated, the probability of ultimate extinction is one if the demographic stochasticity is included or the extinction barrier is chosen at $N = 1$. The expected time to extinction is then the integral of the Green function

$$\mathrm{E}T = \int_a^\infty G(n, n_0)dn,$$

where a is the extinction barrier and n_0 the initial population size. For most populations the expected time to extinction will be quite large. This is illustrated in Fig.3.7, where the expected time to extinction for the logistic model is shown for different parameter values. Consequently, a population which initially is not very much smaller than K will fluctuate around its carrying capacity and return to K a very large number of times. Since diffusions are Markov processes, the distribution of the time to extinction each time K is passed is the same. Hence, the process will approximately have the property that $P(T > u + t|T > u)$ does not depend on u. Writing $G(t) = P(T > t)$ this leads to

$$G(u + t)/G(u) = \psi(t)$$

for some function ψ with $\psi(0) = 1$. This relation may be written as

$$\frac{G(u + t) - G(u)}{t} = \frac{\psi(t) - \psi(0)}{t} G(u).$$

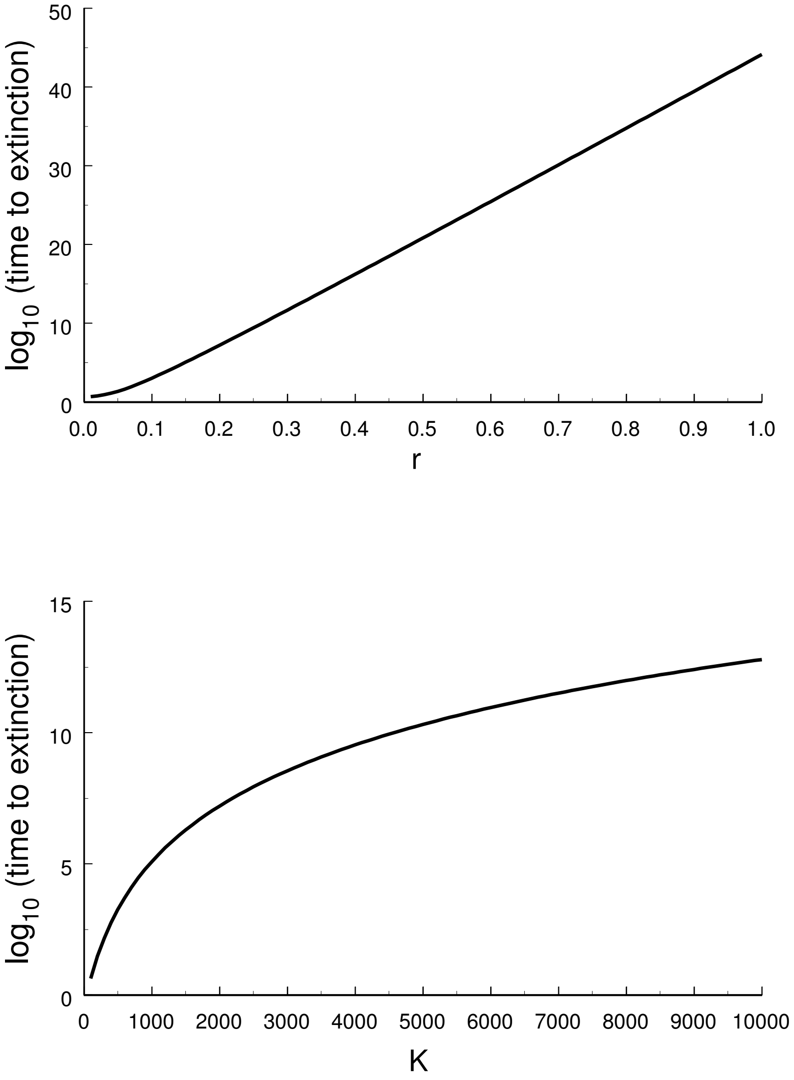

Figure 3.7: The upper panel shows log_{10} of the expected time to extinction at $N = 0$ in the logistic model as a function of r for $K = 500$, $\sigma_e^2 = 0.04$, $\sigma_d^2 = 1$, and $N_0 = K = 500$. The lower panel gives log_{10} of the expected time to extinction at $N = 0$ in the same model as a function of K for $r = 0.05$, $\sigma_e^2 = 0.01$, $\sigma_d^2 = 1$, and $N_0 = K$.

As t approaches zero this leads to $G'(u) = kG(u)$, where $k = \psi'(0)$. As a consequence, $G(u) = \exp(-ku)$, which means that the distribution of the time to extinction is exponential with mean $1/k$, that is

$$P(T < t) \approx 1 - \exp(-t/\mathrm{E}T).$$

An important consequence is that a population may, with fairly large probability, go extinct a long time before the expected time to extinction. As an example, consider a population with expected time to extinction $\mathrm{E}T = 2000$ years. Then, the probability that this population goes extinct before 500 years is approximately $1 - \exp(-500/2000) = 0.2212$, while extinctions before 200 and 100 years occur with probabilities 0.0952 and 0.0488, respectively.
The definition of viability based on the α-quantile of the distribution of time to extinction can now simply be expressed by the expected time to extinction, that is

$$\mathrm{E}T > -t_v/\ln(1 - \alpha_v).$$

Using $t_v = 100$ and $\alpha = 0.1$ we see that populations are viable if $\mathrm{E}T$ is at least 950 years.
Fig.3.8 shows the minimum K required required for viability according to this definition for some different parameter values in the logistic model with demographic and environmental stochasticity.

3.11.3 Extinctions in populations without density regulation

In section 3.8.2 we found the probability of ultimate extinction for the geometric growth model

$$\mu(n) = rn$$

$$\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2.$$

The more general problem of finding the distribution of the time to extinction has fairly simple solutions in the cases where either the demographic or the

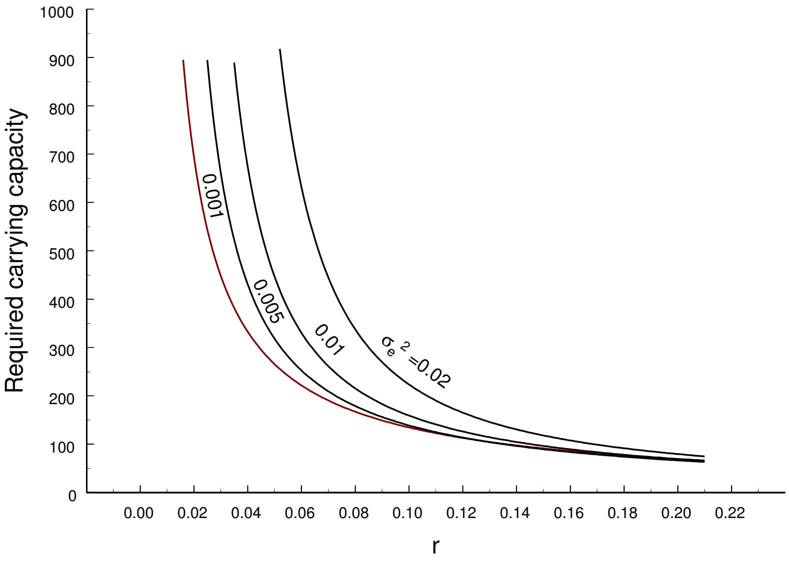

Figure 3.8: The carrying capacity required to obtain a probability of extinction at most 0.1 within a 100 years period for the logistic type of model. There is one graph for four different values of the environmental variance. The demographic variance is 1 and the initial population size is at the carrying capacity.

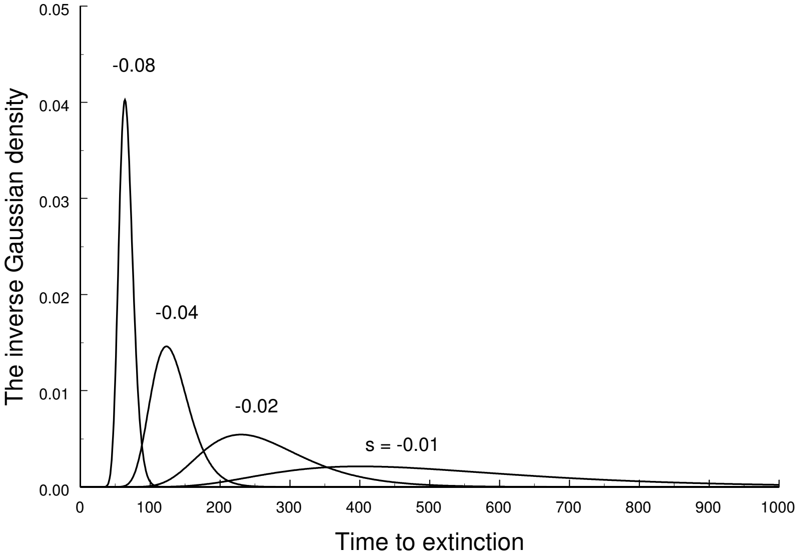

Figure 3.9: The inverse Gaussian distribution of the time to extinction in a Brownian motion for different values of the stochastic growth rate s. The initial population size is $n_0 = 200 = e^{x_0}$ and the environmental variance is 0.01.

environmental variance is zero. If $\sigma_d^2 = 0$ we have seen that $X_t = \ln N_t$ is a Brownian motion with drift parameter $s = r - \sigma_e^2/2$ and infinitesimal variance σ_e^2. If x_0 is the initial state of the process X_t at $t = 0$ and the extinction barrier is at zero, it is a well known result that the distribution of the time to extinction is the inverse Gaussian distribution shown in Fig.3.9

$$g(t) = \frac{x_0}{\sqrt{2\pi\sigma_e^2 t^3}} \exp[-\frac{(x_0 + st)^2}{2\sigma_e^2 t}].$$

If the extinction barrier is at a, then x_0 should simply be replaced by $x_0 - a$. If $s < 0$ this is a proper distribution in the sense that the integral from zero to infinity of $g(t)$ is one. If $s \geq 0$, we have previously seen that the process may be absorbed at infinity in which case the same integral equals the probability of ultimate extinction given in section 3.8.2, which is $\exp(-2sx_0/\sigma_e^2)$. The cumulative distribution (exercise 17) is given by

$$G(t) = P(T \leq t) = \Phi(-\frac{st + x_0}{\sigma_e \sqrt{t}}) + e^{-2sx_0/\sigma_e^2}\Phi(\frac{st - x_0}{\sigma_e \sqrt{t}}),$$

where $\Phi(x) = \int_{-\infty}^{x} \exp(-x^2/2)dx$ is the standard normal integral. Notice that the probability of ultimate extinction, which is $G(\infty)$, is one if $s < 0$ and $\exp(-2sx_0/\sigma_e^2)$ if $s > 0$, in agreement with the result in 3.8.2.

The geometric Brownian motion model is most realistic for viability analysis when stochastic growth rate $s = r - \sigma_e^2/2$ is negative so that the population is unlikely to become large enough for density regulation to occur. In this case the last term of the cumulative inverse Gaussian $G(t_v)$ will be quite small when we insert $t_v = 100$. Then the requirement $P(T < t_v) < \alpha$ is the fulfilled if

$$\Phi(-\frac{st_v + x_0}{\sigma_e \sqrt{t_v}}) < \alpha$$

giving

$$x_0 > \sigma_e \sqrt{t_v} u_\alpha - st_v$$

where u_α denotes the upper α quantile of the standard normal distribution. Inserting $t_v = 100$ and $\alpha = 0.1$ we find simply

$$n_0 = \exp(x_0) > \exp(12.82\sigma_e - 100s).$$

For $\sigma_e^2 = 0.04$, $r = -0.02$ we find $s = -0.04$ giving that the required population size is 709 individuals. However, it is important to be aware of the fact that the demographic component of stochasticity always will have the effect of increasing this threshold.

In case of no extinction barrier we have seen that the distribution of X_t is simply normal with mean $x_0 + st$ and variance $\sigma_e^2 t$. When there is an extinction barrier at $X_t = 0$ this distribution is no longer applicable. At a given time t, the process has either gone extinct, which occur with probability

$G(t)$, or the population is still present with $X_t > 0$. The distribution of X_t in the case of an extinction barrier is also known. With our parameterization this probability density takes the form

$$h(x;t) = \frac{1}{\sqrt{2\pi t}\sigma_e}[1 - e^{-2xx_0/(\sigma_e^2 t)}]e^{-(x-x_0-st)^2/(2\sigma_e^2 t)}.$$

This distribution may be integrated (exercise 18) to give $P(X_t > x) = \int_x^\infty h(z;t)dz$, giving

$$H(x;t) = P(X_t \leq x) = \Phi(\frac{x - x_0 - st}{\sigma_e\sqrt{t}}) + e^{-2sx_0/\sigma_e^2}\Phi(\frac{st - x - x_0}{\sigma_e\sqrt{t}})].$$

Notice that $H(0;t) = P(X_t = 0) = P(T \leq t) = G(t)$ as expected. Notice also that $H(x;t)$ is different from the corresponding cumulative distribution in the case of no extinction barrier (exercise 19).

For some small populations the environmental fluctuations may be ignored compared to those generated by demographic stochasticity giving the model with $\mu(n) = rn$ and $\nu(n) = \sigma_d^2 n$. The distribution of the time to extinction at zero for this model has the simple form

$$P(T < t) = \exp[-\frac{2n_0 re^{rt}}{\sigma_d^2(e^{rt} - 1)}]$$

for $r \neq 0$, and $\exp[-2n_0/(\sigma_d^2 t)]$ for $r = 0$. For $r > 0$ the limit obtained as t approaches infinity gives that the probability of ultimate extinction at zero is $\exp(-2rn_0/\sigma_d^2)$. In section 3.8.2 we found that the probability of ultimate extinction at $N = 1$ for this model was the same expression with n_0 replaced by $n_0 - 1$. For $r \leq 0$ we see that ultimate extinction is a certain event.

No simple expression for the time to extinction is known when the demographic as well as the environmental variance is included, so we have to rely on stochastic simulations of the process. Fig.3.10 shows some examples of how the distribution of the time to extinction is affected by the demographic stochasticity, and in section 3.13 we explain how these simulations may be speeded up when $s > 0$.

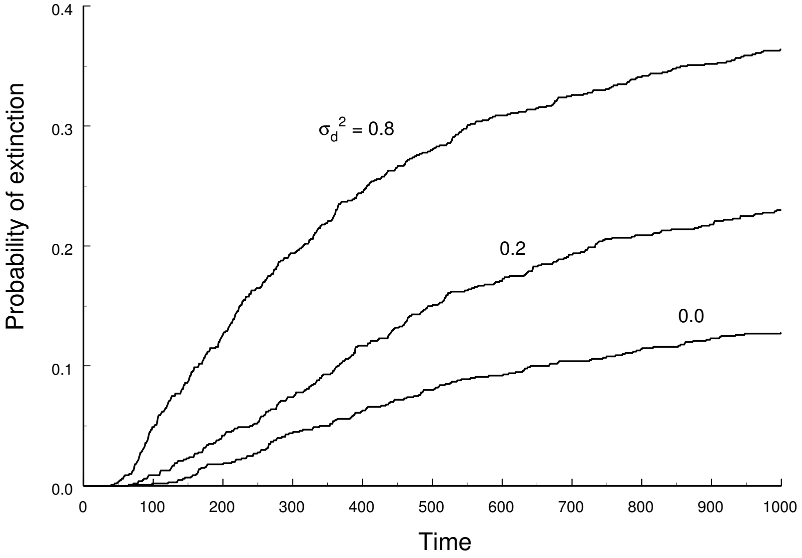

Figure 3.10: The probability of extinction (cumulative distribution) approximated by stochastic simulations of the model with infinitesimal mean $\mu(n) = rn$ and variance $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$ for different values of σ_d^2. The other parameters are $r = 0.025$, $\sigma_e^2 = 0.04$, and $n_0 = 500$. The extinction barrier is at n=1 and the frequencies are base on 1000 simulations.

3.11.4 Some results on the scaling of the time to extinction

We have seen that the Green function involves so many integrations that an analytic expression only can be found in some few special cases, for example models for population growth without density regulation. The expected time to extinction, which is the integral of the Green function, is even more unlikely to reduce to a simple analytical expression. When density regulation is introduced, even quite simple models, such as for example the Ornstein-Uhlenbeck process, leads to complicated expressions for the expected time to extinction that can only be computed by numerical integrations.

On the other hand, it would be of great interest in population biology to have some general insight about the relations between the time to extinction and basic population parameters such as the specific growth rate, the environmental variance, demographic variance, and the carrying capacity. A simplified but still quite informative model that can be used to investigate such relations, is the model with exponential growth up to K, a reflecting barrier at K and extinctions occurring at $N = 1$. Hence, below the carrying capacity the dynamics is exactly the same as we have analyzed for density independent models in section 3.8.2. The simplest way of dealing with the reflecting barrier at K is first to allow the population to exceed K but introduce some density regulation above K so that return to K is certain, and then remove any time interval in which the population path is above K. Clearly, the expected time to extinction in the resulting model with reflection at K must then simply be the integral of the Green function over the interval from the extinction barrier and up to K, since this is the total time the process spends in the same interval. Another way of seeing this result is to start with the theta-logistic model and let θ approach infinity. In the limit we then obtain the same model with reflection at K, and it is easy to see that $s(n)$ tends to infinity for $n > K$, which implies that $m(n)$ and the Green function is zero above K. Consequently, if we choose the initial population size to be at K while the extinction barrier and the lower integration limit in $S(n)$ are at $N = 1$, we have

$$G(n, K) = 2m(n)S(n)$$

for $1 < n < K$ and zero for $n > K$. Using the definitions of $s(n)$ and $S(n)$ we then find (exercise 20) $S(n) = (1+\delta)^{-\gamma} - (n+\delta)^{-\gamma}$, and $s(n) = \gamma(n+\delta)^{-\gamma-1}$, where $\gamma = 2s/\sigma_e^2 = 2r/\sigma_e^2 - 1$, and $\delta = \sigma_d^2/\sigma_e^2$. This leads to the Green function

$$G(n; K) = \frac{1}{sn}[(\frac{n+\delta}{1+\delta})^{\gamma} - 1]$$

for $\sigma_e^2 > 0$. For $\sigma_e^2 = 0$ the solution takes the form

$$G(n; K) = \frac{1}{sn}[e^{2s(n-1)/\sigma_d^2} - 1].$$

Since the Green function does not depend on K other than that the function is zero for $n > K$, we also have the relation

$$\frac{d(\mathrm{E}T)}{dK} = \frac{1}{sK}[(\frac{K+\delta}{1+\delta})^{\gamma} - 1]$$

In Fig.3.11 $sd\mathrm{E}T/dK$ is plotted against K for different values of γ. A simple analytic solution of the integral can only be found when the demographic variance is ignored, that is when $\delta = 0$, giving

$$\mathrm{E}T = \frac{1}{s}[(K^{\gamma} - 1)/\gamma - \ln K].$$

If γ is positive and not very close to zero we find that the expected time to extinction is approximately proportional to K^{γ}.

3.12 Autocorrelations

Though the stationary distribution of a stationary process contains all information of the magnitude of the population fluctuations, it does not tell us anything about how quickly, or how often, the process will return to K, which is an important aspect in practical applications. For example, if we multiply

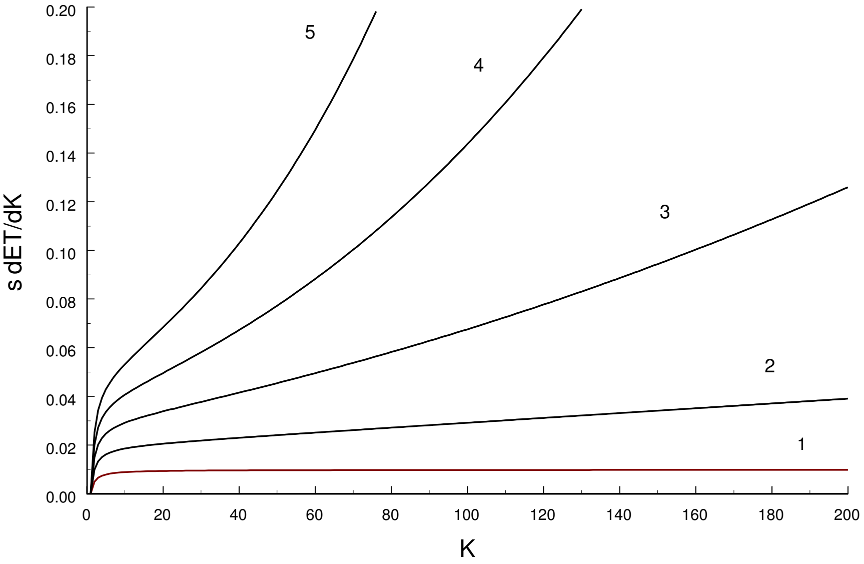

Figure 3.11: Showing $sd\mathrm{E}T/dK$ against K for different values of γ. The ratio $\delta = \sigma_d^2/\sigma_e^2$ is kept constant equal to 100.

the infinitesimal mean and variance by the same constant, the functions $s(x)$ and $S(x)$ are unchanged, and so is the stationary distribution. In fact, this multiplication corresponds to just changing the speed of the process, so it will have a large effect on the return time to equilibrium.

For a stationary process the autocorrelation $\mathrm{corr}[X_t, X_{t+h}] = \rho(h)$ depends on the time step h only. If this correlation is large, X_t and X_{t+h} will be close. In other words, X_t contains much information on X_{t+h} in the sense that if X_t is large, we also expect X_{t+h} to be large. Hence, if $\rho(h)$ is positive and declines slowly with h, the return time to equilibrium is likely to be large. For processes with time-lag, the autocorrelation may also be negative, which indicates a return to equilibrium that is likely to be followed by an amplitude in the opposite direction.

It turns out that stationary diffusions with a linear infinitesimal mean $\mu(x) = \alpha - \beta x$ have an exponentially decreasing autocorrelation regardless the infinitesimal variance. This can be seen by considering the covariance

$$\mathrm{cov}(X_t, X_{t+h+dh}) = \mathrm{cov}(X_t, X_{t+h} + dX_{t+h}),$$

and insert $dX_{t+h} = (\alpha - \beta X_{t+h})dt + \sqrt{\nu(x)dt}U$, where U is a standardized variable which is independent of X_t. This leads to

$$\mathrm{cov}(X_t, X_{t+h+dh}) = (1 - \beta dh)\mathrm{cov}(X_t, X_{t+h}).$$

Dividing by the variance of the stationary distribution we then find

$$[\rho(h + dh) - \rho(h)]/dh = -\beta\rho(h).$$

The left side is the derivative of the autocorrelation function, and the solution of the differential equation with boundary condition $\rho(0) = 1$ is simply

$$\rho(h) = \exp(-\beta h).$$

We see immediately that this applies to the Ornstein-Uhlenbeck process which has linear infinitesimal mean. In section 3.7 we demonstrated how many other processes fluctuating around a carrying capacity may be approximated by such a process. For this approximations we found the coefficient $\beta = -\mu'(K)$, which means that $\exp(\mu'(K)h)$ is often a good approximation to the autocorrelation. The scaling of the autocorrelation may be defined by the time difference T_R that gives a correlation of $1/e$. Using the autocorrelation function based on approximating the process by an Ornstein-Uhlenbeck process we find $T_R \approx -1/\mu'(K)$. For the discrete stochastic model $N_{t+1} = \Lambda(N_t)N_t$ the Ito-approximation gives the diffusion with infinitesimal mean $\mu(n) = n[\lambda(n) - 1]$, where $\lambda(n) = \mathrm{E}[\Lambda(N)|N = n]$. Hence, since $\mu'(K) = K\lambda'(K)$, the scaling of the autocorrelation is $T_R \approx -1/[K\lambda'(K)]$, which is exactly the same as the return time to equilibrium for the analogue deterministic model discussed in section 2.1.

In section 3.6.1 we showed that a model with the Gompertz type of density regulation and no demographic stochasticity transformed to the Ornstein-Uhlenbeck process by the log-transformation. Consequently, using the same notation, we have $\mathrm{corr}(\ln N_t, \ln N_{t+h}) = \exp(-r_1 h/\ln K)$. For the theta-logistic model with $\theta \neq 0$, still ignoring demographic stochasticity, we demonstrated in section 3.6.2 that the transformation $Y_t = N_t^{-\theta}$ lead to a linear infinitesimal mean for Y_t. Consequently we have exactly

$$\mathrm{corr}(N_t^{-\theta}, N_{t+h}^{-\theta}) = \exp(-\theta r h)$$

for this model.

3.13 * Conditional diffusions

Even if the time to extinction often will be extremely large, extinction is the ultimate fate of any population. It may therefore be of some importance to investigate how fast populations approaches the extinction barrier when they finally go extinct. For populations fluctuating around their carrying capacity one may for example consider the time it takes from the last time the population passes the carrying capacity until it goes extinct. And when there is no density regulation, we may be interested in the time to extinction conditional on extinction occurring, that is, considering only those sample paths that do not grow to infinity.

This type of analysis may be performed by defining the so-called conditional diffusion. Consider a diffusion with infinitesimal mean and variance $\mu(n)$ and $\nu(n)$, which is initiated at $X_0 = x_0$ between a and b, $a < b$. One can show that if we only consider the sample paths reaching a before b, we still have a diffusion process, say N_t^*. This process has infinitesimal mean and variance

$$\mu^*(n) = \mu(n) - s(n)\nu(n)/[S(b) - S(n)]$$

$$\nu^*(n) = \nu(n),$$

when the lower integration limit of $S(n)$ is chosen at a.

Consider the geometric growth model with demographic and environmental stochasticity, that is, with $\mu(n) = rn$ and $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$. If the stochastic growth-rate $s = r - \sigma_e^2/2$ is negative, then ultimate extinction is certain, so we consider the case when s is positive. If we chose b to be at infinity, the process N_t^* is the diffusion conditioned on the event of ultimate extinction actually occurring. Choosing the extinction barrier at $a = 1$ and using the expressions for $s(n)$ and $S(n)$ derived in 3.8.2 we find (exercise 21) that the last term to be subtracted in $\mu^*(n)$ is simply $2sn$, giving

$$\mu^*(n) = (-s + \sigma_e^2/2)n,$$

while the infinitesimal variance is unaltered. Hence, we see that the conditional diffusion belongs to the same parametric class of models. The parameter s is simply replaced by $-s$. This is an important result when it comes to simulating the process to extinction in cases where $s > 0$. If A is the event of ultimate extinction and T the time to extinction we have

$$P(T \leq t) = P(T \leq t|A)P(A) + P(T \leq t|A^*)P(A^*),$$

where A^* denotes the compliment of A. If A^* occurs, the process is ultimately absorbed at infinity, so the event $T \leq t$ has probability zero. Hence,

$$P(T \leq t) = P(T \leq t|A)P(A).$$

Here, the conditional probability may be found by simulating the conditional process, that is, the process with s replaced by $-s$, a large number of times and record the frequencies of extinctions. Obviously, we now reach extinction much more often than if we were to simulate the process with positive s. The expression for $P(A)$ is given in 3.8.2.

A well known special case of this result is found for $\sigma_d^2 = 0$. Then s is replaced by $-s$ in the geometric Brownian motion. Transforming to log-scale we see that the drift parameter in the conditional Brownian motion just changes sign as well. For $s < 0$ we have seen that the expected time to extinction at zero is simply $-x_0/s$, where $x_0 = \ln n_0$. Hence, when the drift parameter s is positive, the expected time to extinction, conditional on extinction occurring, is x_0/s.

From the above analysis it appears that the larger the positive growth rate is the quicker will the final decline towards extinction be. This may seem as a paradox, but may to some extent be explained intuitively. The population can only go extinct if a large number of random events accidentally act together to decrease the population size. The mathematical results tell us that when the growth rate is large this must necessarily happen rather quickly if extinction shall occur, while if the growth rate is only a little larger than zero, these random forces may be spread over a longer period of time. The probability that such a final decline should happen, however, decreases strongly as the growth rate increases.

Fig.3.12 shows simulations of the final decline to extinction for a logistic model.

3.14 Stochastic differential equations

For the diffusion process with infinitesimal mean and variance $\mu(x)$ and $\nu(x)$, respectively, we have by definition that $\mathrm{E}(dX) = \mu(X)dt$ and $\mathrm{var}(dX) = \nu(X)dt$. If B_t is a standard Brownian motion the process may be written on

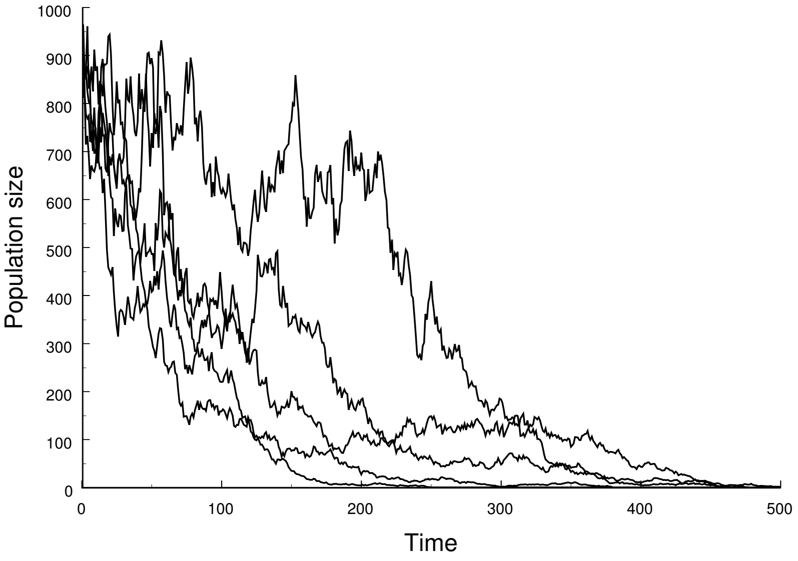

Figure 3.12: Simulations of the final decline from the carrying capacity to extinction in the logistic model with parameters $r = 0.2$, $K = 1000$, $\sigma_d^2 = 1$ and $\sigma_e^2 = 0.01$.

the form

$$dX = \mu(X)dt + \sqrt{\nu(X)}dB,$$

where $dB = B_{t+dt} - B_t$. Notice that this is the continuous analogy to the discrete process in 3.2 that we used to motivate diffusions. Since $\mathrm{E}(dB) = 0$ and $\mathrm{var}(dB) = dt$ we see that this formulation is consistent with the above expressions for $\mathrm{E}(dX)$ and $\mathrm{var}(dX)$. The model may then formally be expressed as

$$dX = \mu(X)dt + \sqrt{\nu(X)}W_t dt$$

where $W_t = dB/dt$ at time t is called *Gaussian white noise*. Formally, for a small dt we find $\mathrm{E}W_t = 0$ and $\mathrm{var}(W_t) = 1/dt$, so the variance actually approaches infinity as dt approaches zero. Hence, the white noise process is not really well defined and this cause some problems when it comes to integration. For Brownian motions we also have that $\mathrm{E}dB_t dB_{t+h} = 0$ implying that $\mathrm{corr}(W_t, W_{t+h}) = 0$. Notice that, by introducing the concept of white noise, the model may formally be written as a differential equation, actually the stochastic differential equation

$$dX/dt = \mu(X) + \sqrt{\nu(X)}W_t.$$

As already mentioned, integration of such an equation, called stochastic integration, is difficult and not uniquely defined. Different methods (Ito-integral, Stratonovich integral) may lead to different results. Here we only show the simple example of solving the Ornstein-Uhlenbeck process, in which case these problems do not occur due to the nice linear form of the solution.
The OU-process written as a stochastic differential equation takes the form

$$dX_t = (\alpha - \beta X_t)dt + \sigma dB(t).$$

The solution of the homogenous equation $dX/dt = \alpha - \beta X$ yields the solution

$$X_t = \alpha/\beta + Ce^{-\beta t}.$$

Inserting a function $C(t)$ for the integration constant C in the original equation we find (exercise 22) assuming that the process starts at x_0 at time zero that

$$X_t = \alpha/\beta + (x_0 - \alpha/\beta)e^{-\beta t} + \sigma \int_0^t e^{-\beta(t-u)} dB(u).$$

Since this solution is linear in $dB(t)$ it follows that X_t is normally distributed because any increment $B(t+h) - dB(t)$ is normally distributed by definition. Taking the expectation and using $\mathrm{E}dB(t) = 0$ we find the expected value

$$\mathrm{E}(X_t | X_0 = x_0) = \alpha/\beta + (x_0 - \alpha/\beta)e^{-\beta t}$$

as given in 3.8.4. To find the variance we use the continuous analogy of the formula for a sum of independent random variables considering $\int e^{-\beta(t-u)} dB(u)$ as a sum of small independent increments during time intervals Δu, $e^{-\beta(t-u)} \Delta B(u)$, and going to the limit. Since $\mathrm{var}[\Delta B(u)] = \Delta u$ we then find

$$\mathrm{var}(X_t | X_0 = x_0) = \sigma^2 \int_0^t e^{-2\beta(t-u)} du = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}).$$

3.15 Autocorrelated noise

3.15.1 Diffusion approximations to discrete models with autocorrelated noise

The environmental variables affecting the stochastic terms in the change in population size at different seasons may have temporal correlation. For example, there may be autocorrelations in the physical environments of the population. If death rates and fecundity vary with the age of the individual, an age-structured population model may be required. If we then only deal with the total population size and not the whole population vector, the model will no longer be a Markov process and age-structure will generate autocorrelations. This will be dealt with in chapter 4.

In 1.8.3 we introduced the concept of autocorrelations in models for exponential growth. These autocorrelations did not have any effect on the stochastic growth rate, but could dramatically change the variance of future population sizes. More precisely, for large values of t we found

$$\frac{\mathrm{var}(\ln N_t|N_0)}{t} \approx \sigma^2 \sum_{-\infty}^{\infty} \rho(i)$$

where $\sigma^2 = \mathrm{var}(\ln \Lambda_t)$ and the autocorrelation function is defined by $\rho(h) = \mathrm{corr}(\ln \Lambda_t, \ln \Lambda_{t+h})$. If there are no autocorrelations the process $X_t = \ln N_t$ can be approximated by a Brownian motion with infinitesimal mean and variance $\mu = \mathrm{E} \ln \Lambda$ and σ^2, respectively. For large values of t, the discrete model as well as the Brownian motion will have the property that X_t is normal with mean $x_0 + \mu t$ and variance $\sigma^2 t$. On the other hand, if there are autocorrelations in the discrete process, then X_t is approximately normal with the same mean but with variance $\sigma^2 t \sum_{-\infty}^{\infty} \rho(i)$. Hence, the appropriate Brownian motion to use as a continuous approximation when there are autocorrelations in the noise has infinitesimal mean μ and infinitesimal variance $\sigma^2 \sum_{-\infty}^{\infty} \rho(i)$. Since the autocorrelation most realistically is positive, we see that autocorrelations must be compensated by an increase in the infinitesimal variance of the Brownian motion.

For the general type of discrete process with autocorrelations in the noise, the same kind of approximation may be used. Consider the discrete model with

$$\Delta X_t = \mu(X_t) + \sqrt{\nu(X_t)} U_t,$$

where U_t is some noise process with $\mathrm{E}U_t = 0$, $\mathrm{var}(U_t) = 1$ and $\mathrm{corr}(U_t, U_{t+h}) = \rho(h)$. In accordance with the previous argument the diffusion approximation to this process has infinitesimal mean $\mu(x)$ and infinitesimal variance $\nu(x) \sum_{-\infty}^{\infty} \rho(i)$. The best approximation is obtained if one works on the log-scale rather than modelling directly the populations sizes. This is because the approximation is constructed for population growth with no density-regulation (constant growth rate) and constant variances, which is approximately correct on the log scale for small moderate population fluctuations around the carrying capacity. Further, one cannot expect a good approximation if the scaling of the autocorrelations for the process is not much larger than the scaling of the autocorrelations of the noise process. Fig.3.13 shows

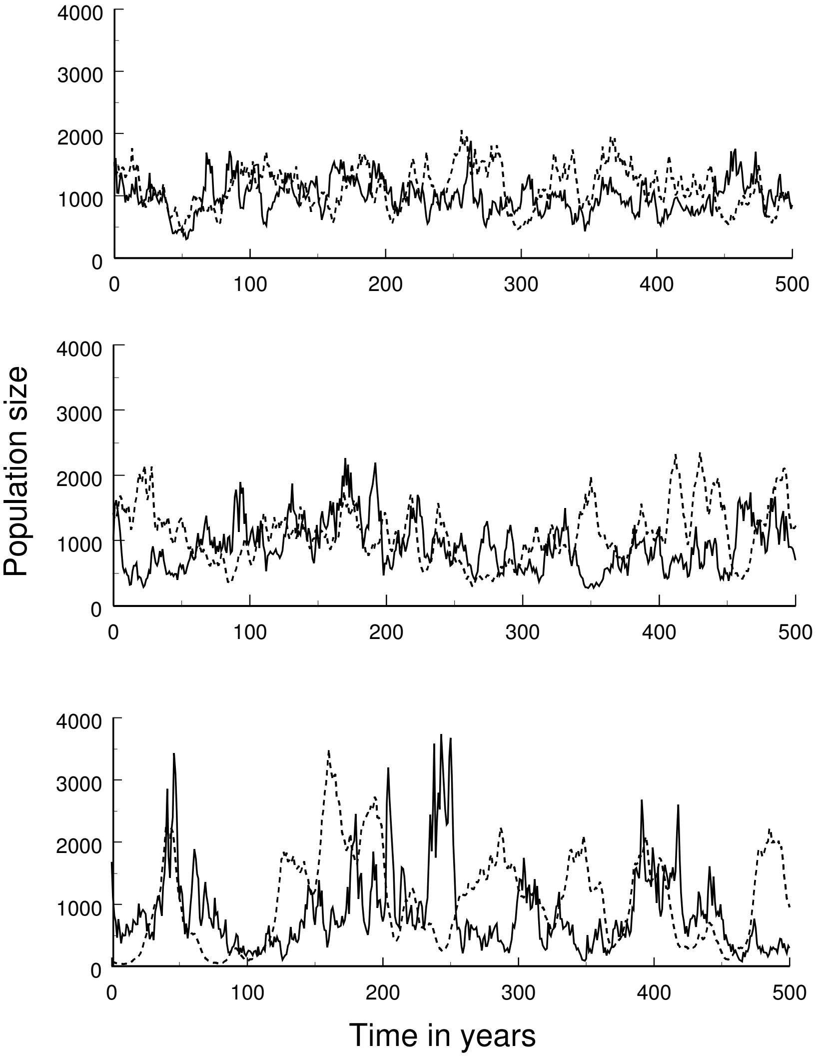

Figure 3.13: Simulations of the discrete autocorrelated process (solid lines) and the diffusion approximation (dashed line). The model is the logistic type with $r = 0.1$, $K = 1000$, $\sigma_e^2 = 0.01$, and $\sigma_d^2 = 0$. The parameter γ (see main text) is 0.3 in the upper panel, 0.5 in the middle and 0.8 in the lower panel.

some simulations of a discrete process with autocorrelations together with some simulations of the relevant diffusion approximation. The noise process is chosen as $U_t = (1-\gamma^2)^{1/2} V_t + \gamma U_{t-1}$ where $-1 < \gamma < 1$ and the V_t are independent and uniformly distributed on $[-1/2, 1/2]$. This gives $\rho(h) = \gamma^{|h|}$ and $\sum_{-\infty}^{\infty} \rho(i) = (1+\gamma)/(1-\gamma)$. As a consequence the infinitesimal variance of the diffusion approximation should be $\nu(x)(1+\gamma)/(1-\gamma)$.

Fig.3.14 shows the same comparison over a period of 1000 years and γ as large as 0.8. We see that even for such a highly correlated noise process the diffusion approximation seems to show the same kind of large scale fluctuations as those for the process with autocorrelated noise, although the diffusion will always have larger fluctuations over small time intervals. These small scale fluctuations of diffusions are actually the nature of any diffusion processes.

3.15.2 Diffusion approximations to continuous models with colored noise

Suppose now that we now replace the white noise W_t by a colored noise process Z_t with $\mathrm{E}Z_t = 0$, $\mathrm{var}(Z_t) = \sigma^2$ and autocorrelation function $\rho(h)$. Writing $V_t = \int_0^t Z_t dt$ we find in analogy with what we did for discrete processes

$$\frac{\mathrm{var} V_t}{\sigma^2 t} = \int_{-t}^{t} \rho(x) dx - \frac{2}{t} \int_0^t x\rho(x) dx.$$

Hence, provided that $\int_0^\infty x\rho(x)dx$ is finite $\mathrm{var}V_t/t$ approaches $\tau^2 = \sigma^2 \int_{-\infty}^{\infty} \rho(x)dx$ as t increases towards infinity. This suggests that V_t can be replaced by the Brownian motion with zero mean and infinitesimal variance τ^2. The corresponding population process would then be

$$dX = \mu(X)dt + \sqrt{\nu(x)}\tau dB.$$

Hence, the infinitesimal variance of the diffusion approximation to the above process with colored noise should be $\nu(x)\tau^2$.

In order to exemplify this, let the noise process Z_t be the Ornstein-Uhlenbeck process with infinitesimal mean and variance $-\beta z$ and ω^2, respectively. For

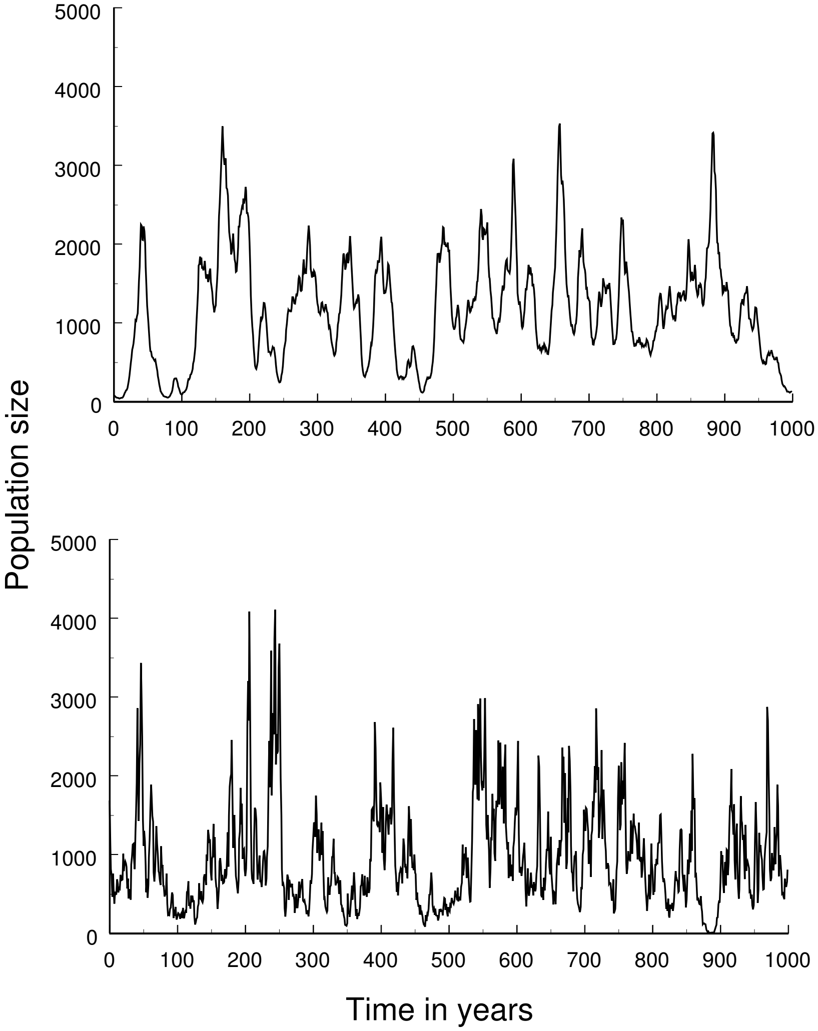

Figure 3.14: The same model as in the lower panel of Fig.3.13 with $\gamma = 0.8$ but showing the processes separately for 1000 years. The upper panel is the process with autocorrelated noise and the lower panel is the diffusion approximation.

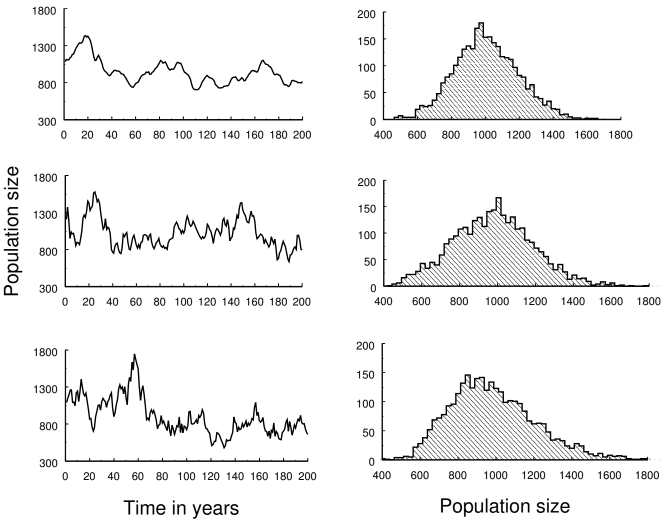

Figure 3.15: Simulations of the continuous processes (left panels) of the logistic type during 200 years with $r = 0.1$, $K = 1000$, and $\sigma_e^2 = 0.01$ with noise process Z_t of the Ornstein-Uhlenbeck type shown together with histograms (right panels) of the stationary distributions based on 30000 years. In the upper panel $\beta = 0.3$, in the middle $\beta = 2$ and in the lower $\beta = 100$. In all simulations $\omega^2 = \beta^2$. In the process with $\beta = 100$ the noise Z_t is practically white noise.

this process we have that $\mathrm{E}Z_t = 0$, $\mathrm{var}(Z_t) = \omega^2/(2\beta)$ and $\rho(h) = \exp(-\beta h)$. This leads to $\tau^2 = \omega^2/\beta^2$. Hence any process with $\omega^2 = \beta^2$, that is with $\tau^2 = 1$, should be approximated by the same diffusion since the noise Z_t of the process can be approximated by the same white noise process W_t. As β approaches infinity $\rho(h)$ approaches zero for $h > 0$. Hence, this limit corresponds to the diffusion provided that ω^2 is kept equal to β^2, that is ω^2 also approaches infinity. Then also $\mathrm{var}(Z)$ approaches infinity in accordance with what we found for the white noise process W_t.

The accuracy of this diffusion approximation depends on the value of β. The approximation will break down for small values of β corresponding to considerable autocorrelations in the noise over large time differences. Fig.3.15 shows the simulations of a logistic process X_t with noise Z_t for different values of $\beta^2 = \omega^2$ including the diffusion, together with histograms for the stationary distributions..

3.16 The accuracy of the diffusion approximation

We see from Fig.3.15 that a process with considerable temporal autocorrelation in the noise may be approximated by a diffusion with white noise. Comparing the histograms for the stationary distributions (left panel) it appears that the diffusion approximation is quite good.

The best way to make comparisons with a given process and its diffusion approximation is to simulate a large number of processes over a long time interval and compare the distribution of population sizes each year. In this way we can also judge to which extent the diffusion approximation can be applied to calculate approximations for the expected time to extinction, distribution of the time to extinction, distribution of future population sizes or the stationary distribution in the case of stationarity. It is important, however, that these comparisons are done correctly, that is, that we use exactly the mean and variance functions for the process as the infinitesimal mean and variance in the diffusion approximation.

Usually the extinction barrier is chosen at $N = 0$ or $N = 1$. This means that the process do show large discrete steps compared to N the last time period before extinction occurs. One may therefore tend to think that the diffusion is no good as an approximation for analyzing extinction processes when the barriers are so small. Furthermore, real populations may frequently show fairly large between years fluctuations. Ignoring the demographic stochasticity, we have seen that the between years fluctuations in the growth rate r has variance σ_e^2 and consequently a standard deviation of σ_e so we can think of the next years population size being the previous population size multiplied by the factor which is often larger than $\lambda + \sigma_e$ and smaller that $\lambda - \sigma_e$. If $\lambda = 1$ and the environmental variance is $\sigma_e^2 = 0.01$, then $\sigma_e = 0.1$, and these two factors are 1.1 and 0.90. Hence, this environmental variance reflects rather large between years fluctuations as we already have demonstrated in a number of graphs. If $\sigma_e^2 = 0.04$ the same factors are 1.2 and 0.8 with a ratio of $1.2/0.8 = 1.5$. Hence there is a very large uncertainty with respect to next years population size.

In Fig.3.16 we see a model in discrete time with normally distributed between years change on log population size together with the diffusion approximation to this model. It appears that the diffusion approximation performs extremely well, actually all the way down to extinction. In Fig.3.17 we show a model in discrete time as well as discrete population sizes, the model defined in section 2.4.2 with Poisson distributed contributions to the next generation. Even if this is a model with discrete population sizes and an environmental variance as large as 0.04, the diffusion approximation still performs well, not only for prediction of future population sizes but also for the distribution of the time to extinction at $N = 1$. This is quite remarkable, taken into account that the demographic contribution to the variance of Λ in this model is also large at small population sizes. Actually $\sigma_d^2 = 1 + r(K - N)/K$ in this model so for a population size $N = 10$ the demographic contribution to $\mathrm{var}(\Lambda)$ is approximately $1.02/10 \approx 0.10$, giving a total standard deviation in Λ as large as $\sqrt{0.04 + 0.10} = 0.66$. In Fig.3.18 the environmental variance is increased to $\sigma_e^2 = 0.09$, corresponding to a standard error of Λ as large as 0.3 excluding the demographic contribution. At population size 10 the total standard deviation of Λ is now 0.5. We see that the diffusion approximation is no longer accurate, although it still gives the time to extinction with error only about 10%.

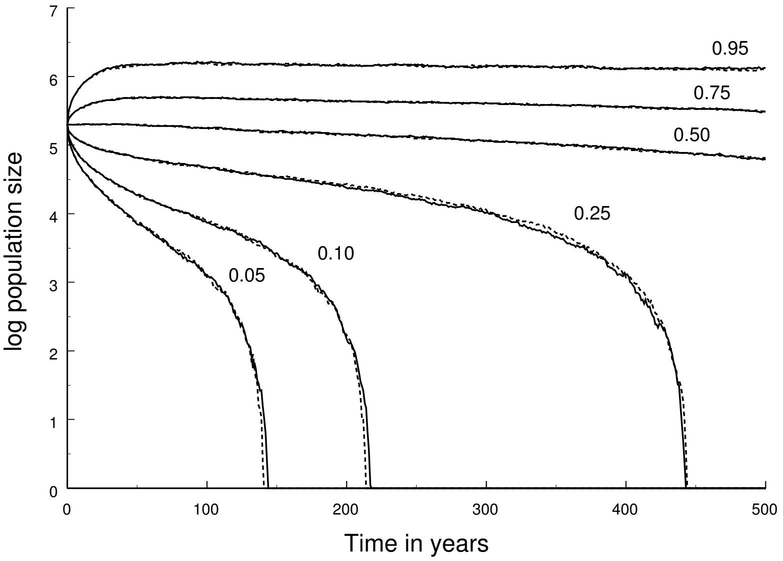

Figure 3.16: Quantiles of the distribution of log population size as function of time found from 10000 stochastic simulations. The model is of the logistic type and the initial population size is at the carrying capacity. The solid lines show the results for the discrete time model with normally distributed changes in $\ln(N)$ from one year to the next, while the dotted lines are simulations of the diffusions. The parameters are $r = 0.02$, $K = 200$, $\sigma_e^2 = 0.01$ and $\sigma_d^2 = 1$.

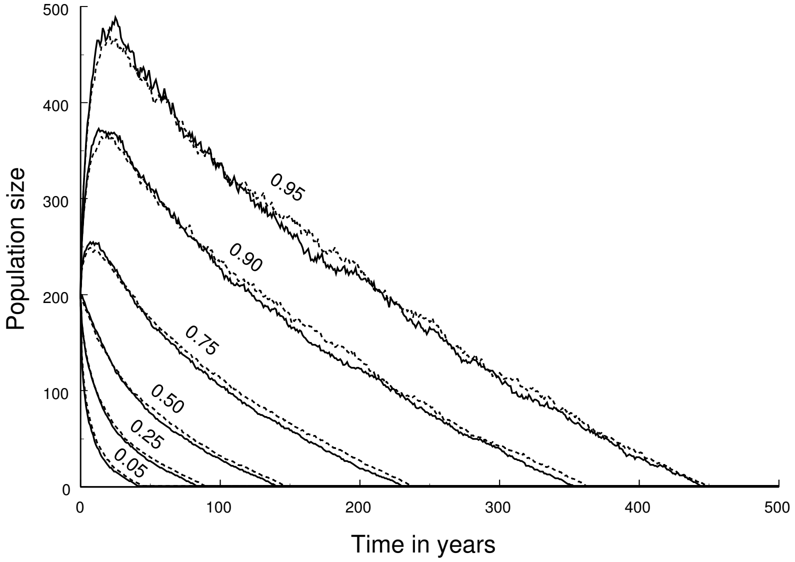

Figure 3.17: The solid lines show quantiles of the distribution of population size as function of time for the logistic Poisson type of model defined in 2.7.2 with additive stochastic effect on $\lambda(\mathbf{z}, N)$. The parameters are $r = 0.02$, $K = 200$ and $\sigma_e^2 = 0.04$. The initial population size is at the carrying capacity. The dotted lines shows the same quantiles for the diffusion approximation. The graphs are based on 10000 simulations for each model.

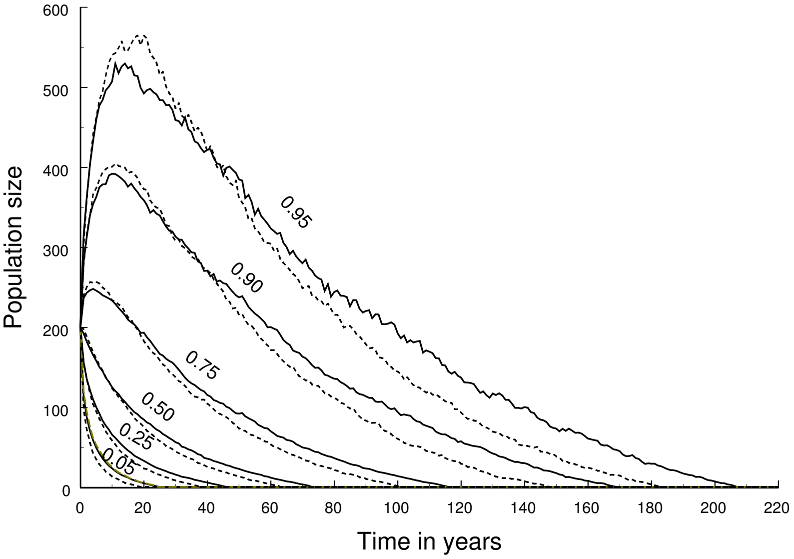

Figure 3.18: The same as Fig.3.17 but with larger environmental variance, $\sigma_e^2 = 0.09$.

3.17 Exercises

1. The diffusion approximation for the discrete process N_t with no density-regulation and constant environmental and demographic variance has infinitesimal mean and variance $\mu(n) = rn$ and $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$. Find the corresponding infinitesimal mean and variance for the process $X_t = \ln N_t$.

2. For the model in exercise 1 find the unstable equilibrium point for the process X_t defined by $\mu_X(x*) = 0$ and the corresponding population size $n* = \exp(x*)$.

3. Consider the diffusion N_t with with infinitesimal mean and variance $\mu(n) = r_1 n(1 - \ln n/\ln K)$ and $\nu(n) = \sigma_e^2 n^2$. Show that $X_t = \ln N_t$ is an Ornstein-Uhlenbeck process and find the parameters of this process expressed by r_1, K and σ_e^2.

4. Let N_t be the diffusion approximation to the theta-logistic model with no demographic variance and constant environmental variance, that is, the infinitesimal mean and variance are $\mu(n) = rn(1 - \frac{n^\theta}{K^\theta})$ and $\nu(n) = \sigma_e^2 n^2$ for $\theta \neq 0$. Find the infinitesimal mean and variance for the process $X_t = N_t^\theta$.

5. For the process in exercise 4 show that $Y_t = N_t^{-\theta}$ has infinitesimal mean of the same linear form as the Ornstein-Uhlenbeck process.

6. For the process N_t with infinitesimal variance $\nu(n) = \sigma_d^2 n$ find a transformation that gives a process with constant infinitesimal variance.

7. Let N_t be a diffusion with infinitesimal variance $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$. Show that the process $X_t = \sigma_e^{-1}\{\ln N_t + 2\ln[1 + \sqrt{1 + \sigma_d^2/(\sigma_e^2 N_t)}] - b\}$ where $b = 2\ln(1 + \sqrt{1 + \sigma_d^2/\sigma_e^2})$ has constant environmental variance.

8. Consider the diffusion for N with infinitesimal mean and variance $\mu(n)$ and $\nu(n) = \sigma_e^2 n^2$. Suppose that $\nu(n)$ is a decreasing function with $\mu(K) = 0$. Find the infinitesimal mean and variance for the process $X = \ln N$. Then linearize the infinitesimal mean for X around $X = \ln K$ $(N = K)$. Show that this linear approximation is the Ornstein-Uhlenbeck process with infinitesimal mean and variance $\mu'(K)(x - \ln K) - \sigma_e^2/2$ and σ_e^2.

9. Consider a diffusion model with absorbing barriers at a and b, $a < b$ which is in state N_0 at time zero. Suppose that the infinitesimal mean for small values of n is rn and that the infinitesimal variance is $\nu(n) = \sigma_e^2 n^2$. Choose 1 as the lower integration limits in the formulas and show that the probability that the process

is absorbed at b tends to one as a tends to zero provided that $s = r - \frac{1}{2}\sigma_e^2 > 0$. How do you interpret this result as b approaches infinity?

10. Consider the model in exercise 9 but add a demographic term $\sigma_d^2 n$ to the infinitesimal variance. Show that the process now may be absorbed at a if $a = 0$ and $s > 0$.

11. For the model in exercise 9 suppose that the population is density-regulated with negative infinitesimal mean above K decreasing with increasing population sizes. Show that extinction now is certain at $a = 1$.

12. Perform the integrations required to find $s(x)$, $S(x)$ and $u(x_0)$ for the Brownian motion given in 3.8.2.

13. Find $s(n)$, $S(n)$ and the probability of ultimate extinction at $n = 1$ for the model $\mu(n) = rn$, $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$ given in 3.8.2 when the initial population size is n_0.

14. Find the probability of ultimate extinction for the model in exercise 13 for $\sigma_e^2 = 0$ by finding the limit as $\sigma_e^2 \to 0$.

15. Use our definition (interpretation) of the Green function to show that $\mathrm{E}[\int_0^T h(X_t)dt] = \int_a^b h(x)G(x, x_0)dx$.

16. Find the limit of the stationary distribution for the theta-logistic model as $\theta \to \infty$.

17. Show that the inverse Gaussian distribution $g(t) = \frac{x_0}{\sqrt{2\pi\sigma_e^2 t^3}} \exp[-\frac{(x_0+st)^2}{2\sigma_e^2 t}]$ has cumulative distribution $G(t) = \Phi(-\frac{st+x_0}{\sigma_e\sqrt{t}}) + e^{-2sx_0/\sigma_e^2}\Phi(\frac{st-x_0}{\sigma_e\sqrt{t}})$.

18. Show that the distribution of log population size in the Brownian motion conditioned on extinction not occurring before time t,
$h(x;t) = \frac{1}{\sqrt{2\pi t}\sigma_e}[1 - e^{-2xx_0/(\sigma_e^2 t)}]e^{-(x-x_0-st)^2/(2\sigma_e^2 t)}$ has cumulative distribution
$H(x;t) = P(X_t \le x) = \Phi(\frac{x-x_0-st}{\sigma_e\sqrt{t}}) + e^{-2sx_0/\sigma_e^2}\Phi(\frac{st-x-x_0}{\sigma_e\sqrt{t}})]$.

19. Consider the geometric brownian motion with $\mu(n) = rn$, $\nu(n) = \sigma_e^2 n^2$ and initial state $n_0 > 0$ at time $t = 0$. Compare the expressions for $P(N_t > n)$ when there is no extinction barrier and when extinction occur at $N = 1$. Discuss the result.

20. For the ceiling model with exponential growth and constant environmental and demographic variance show that $S(n) = (1+\delta)^{-\gamma} - (n+\delta)^{-\gamma}$, and $s(n) = \gamma(n+\delta)^{-\gamma-1}$, where $\gamma = 2s/\sigma_e^2 = 2r/\sigma_e^2 - 1$, and $\delta = \sigma_d^2/\sigma_e^2$, and from this find the Green function $G(n;K) = \frac{1}{sn}[(\frac{n+\delta}{1+\delta})^{\gamma} - 1]$ for $n < K$.

21. Consider the diffusion with infinitesimal mean and variance $\mu(n) = rn$ and $\nu(n) = \sigma_d^2 n + \sigma_e^2 n^2$ and hence stochastic growth rate $s = r - \sigma_e^2/2$ for large population sizes. For $s > 0$ show that the conditional diffusion, conditioning on extinction to occur, has infinitesimal mean $\mu^*(n) = (-s + \sigma_e^2/2)n$.

22. Verify the solution to the OU-process $X_t = \alpha/\beta + (x_0 - \alpha/\beta)e^{-\beta t} + \sigma \int_0^t e^{-\beta t} dB(t)$ by solving the stochastic differential equation.

Chapter 4

Age-structured populations

4.1 Introduction

We have until now dealt with female segments of populations with overlapping generations by assuming that all females at the time of census have identical distributions of vital rates within years. Although individuals may live for a long time we cannot obtain improved predictions or better understanding of the dynamics by separating the population into age-classes when this assumption of equivalent individuals is realistic. However, a large number of organisms will have survival and reproductive rates that change with age. Many species will not be mature until a certain age, reproduction may last for a number of years at changing rate and stop when the individual gets older, as for example for humans. If survival rates were constant, their lifetime would be geometrically distributed, which is only rarely the case. Survival rate will often be small for quite young individuals, larger in an intermediate period and decrease for old individuals.

The dynamics of age-structured populations with changing vital rates will differ from the simplified case of equivalent individuals. The expected change in populations size is no longer just a function of the female population size the previous year, but depends also on the composition of the population, the distribution of individuals among age-classes.

In this chapter we shall deal with the description of the dynamics of age-structured populations, still only considering the female segment, assuming that there is no lack of males. We first consider the classical deterministic theory without density-regulation, which in the stochastic case can be applied to analyze the expected value of future population sizes. Then we go on analyzing stochastic fluctuations in vital rates. It appears that these may be decomposed into environmental end demographic components, and that a stochastic growth rate can be defined for the populations as in the simple case. The diffusion approximation will also often still be applicable for the total population size. These results can then be utilized to find rather accurate predictions of future population size, and to analysis of extinction processes. The theory is extended to include density-regulated populations with stochastic fluctuations around a stable equilibrium.
For many species the determination of age may be difficult. The vital rates may also not be determined by age, but rather by weight or stage. One should then work with stage-structured models rather than age-structure which in general may be more complicated. However, if each stage is reached at a fixed age, the formulation given here, possibly with minor technical modifications, is still applicable.

4.2 Deterministic theory

4.2.1 Population growth rate and stable age-distribution

Let $\mathbf{n} = (n_1, n_2, \ldots, n_k)'$, where $'$ denotes matrix transposition, be the population vector with elements n_i, the number of individuals of age i just prior to reproduction. As k is the last age-class, the survival from age k to $k+1$ is assumed to be zero. The life history is determined by the age-specific reproductive rates, p_i, the proportion surviving from age i to $i+1$, and f_i, the number of female offspring surviving to the next census produced by each female of age i. The deterministic projection equations are then

$$n_1 + \Delta n_1 = \sum_{i=1}^{k} f_i n_i$$

$$n_i + \Delta n_i = p_{i-1} n_{i-1} \quad \text{for} \quad i = 2, 3, \ldots, k.$$

Except for some very special choices of parameter values which hardly occur in practice, this population will approach a stable age-distribution, say $\mathbf{u} = (u_1, u_2, \ldots, u_k)'$ with $\sum u_i = 1$, and total population size growing exponentially in time with multiplicative rate λ. Assuming the stable age-distribution has been reached we then have $u_i p_i = \lambda u_{i+1}$, or $u_{i+1} = u_i p_i / \lambda$ for $i > 1$. Solving these equations recursively gives $u_i = u_1 l_i / \lambda^{i-1}$ for $i = 2, 3 \ldots, k$, where l_i is the probability that an individual survives to age i,

$$l_i = p_1 p_2 \ldots p_{i-1} \quad \text{for} \quad i = 2, 3, \ldots, k$$

and $l_1 = 1$. Inserting this into the equation for reproduction yields

$$\lambda u_1 = \sum_{i=1}^{k} f_i u_i = \sum_{i=1}^{k} f_i u_1 l_i / \lambda^{i-1}.$$

Dividing the equation by λu_1 produces the Euler-Lotka equation

$$\sum_{i=1}^{k} l_i f_i \lambda^{-i} = 1.$$

This polynomial equation of degree k will generally have k different solutions for λ. However, since the expression on the left side is a decreasing function of λ, decreasing from infinity to zero as λ increases from zero to infinity there is only one real root that must be positive. This root, λ_1, can be shown to be the root with the largest modulus. The modulus of the root with the second largest modulus, $|\lambda_2|$, determines the multiplicative rate of approach to the stable age-distribution, $|\lambda_2|/\lambda_1$. From now on the dominant eigenvalue λ_1 will for simplicity be denoted λ. Hence, this parameter has essentially the same interpretation as λ in chapter 1.

Using the fact that $\sum u_i = 1$ we find the stable age distribution

$$u_i = l_i \lambda^{-i} / \sum_{j=1}^{k} l_j \lambda^{-j}.$$

Notice that if the population size is constant, $\lambda = 1$, then the u_i are proportional to l_i. For a growing population λ is greater than 1, giving relatively larger values of u_i for young individuals which means that there are relatively more young individuals, whereas decreasing populations will have relatively more old individuals.

For realistic parameter values the age-distribution usually approaches the stable age-distribution within a few generations as exemplified in Fig.4.1.

The generation time T is defined as the mean age of mothers of newborn individuals when the population has reached the stable age-distribution. Using the above Euler-Votka equation (exercise 2) gives

$$T = \sum_{i=1}^{k} i l_i f_i \lambda^{-i}.$$

4.2.2 Reproductive value

Starting initially with a total population size of $N_0 = n_i$ individuals, all of the same age i at time $t = 0$, the population will asymptotically reach the stable age-distribution and grow exponentially with multiplicative rate λ. Hence the population size N_t at time t will approximately be proportional to λ^t, or more precisely, N_t/λ^t approaches a constant, say $n_i v_i$. The parameter v_i, first introduced by R.A. Fisher, is called the reproductive value of individuals of age i. It expresses these individuals ability to contribute to future generations. For example, if $v_2 = 2v_1$, individuals of age 2 will contribute with twice as many individuals to future generations as those in the first age-class.

Since the individuals in the k'th age-class do not survive, the contribution to the next generation from n_k individual is simply $n_k f_k$ individuals, all of age 1. At time t these individuals contribute with $n_k f_k v_1 \lambda^{t-1}$, giving

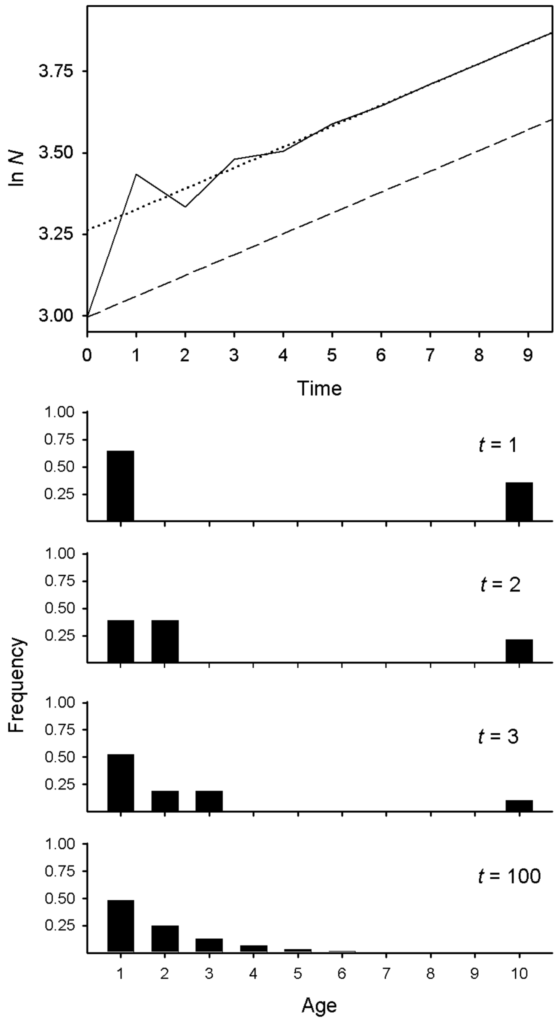

Figure 4.1: Deterministic growth of the total population size (upper panel) and change in age-distribution through time (lower panel) in a deterministic matrix model of the Lefkovitch with 10 stages, corresponding to ages 1-9 and the last stage contains individuals of age 10 and older. Parameter values are $p_i = 0.55$ for all classes, $f_1 = 0$ and $f_i = 1$ for $i > 1$. Initially $n_1 = 20$ and $n_i = 0$ for $i > 1$ at time zero. The dashed line shows the exact exponential growth in the case that the 20 individuals originally were distributed according to the stable age-distribution for the model. The dotted line shows the growth of the reproductive value.

$n_k v_k \lambda^t = n_k f_k v_1 \lambda^{t-1}$ or

$$f_k v_1 = \lambda v_k.$$

Arguing in the same way for the other age-classes we find the relation $n_i v_i \lambda^t = n_i p_i v_{i+1} \lambda^{t-1} + n_i f_i v_1 \lambda^{t-1}$, giving

$$f_i v_1 + p_i v_{i+1} = \lambda v_i$$

for $i = 1, 2, \ldots, k-1$. These equations determine the reproductive values up to a constant factor. One solution is (exercise 3)

$$v_i' = \frac{\lambda^i}{l_i} \sum_{j=i}^{k} l_j f_j \lambda^{-j}.$$

We find the correct scaling of the reproductive values by considering a population initially of size N_0 having the stable age-distribution. Asymptotically the population size will then be $N_0 \sum u_i v_i \lambda^t$. On the other hand, this population grow exactly with multiplicative rate λ, so the population size is also $N_0 \lambda^t$ giving $\sum u_i v_i = 1$. Consequently, the above v_i' must be multiplied by the factor $(\sum u_i v_i')^{-1}$ to give the correct scaling, that is

$$v_i = v_i' / (\sum u_i v_i').$$

The total reproductive value of the population, say $V = \sum n_i v_i$, is the sum of the reproductive values of all individuals. Notice that for a population with the stable age-distribution the total reproductive value is identical to the population size since then $n_i = N u_i$ giving $V = \sum n_i v_i = \sum N u_i v_i = N$, so in this case the population vector is $\mathbf{n} = V\mathbf{u}$. Now, consider an initial population $N_0 = \sum n_i$ with reproductive value $V_0 = \sum n_i v_i$ at time zero. According to the above results this population will asymptotically reach the stable age-distribution and grow exponentially so that N_t / λ^t approaches $\sum n_i v_i = V_0$. This means that V_t / V_0 and hence N_t / V_0 approaches λ^t as t increases. The stronger result that the reproductive value has exactly exponential growth with multiplicative rate λ can be showed by using the linear equations for the reproductive values (exercise 4). We shall show this result in a simpler way in the next section using matrix algebra.

4.2.3 Matrix formulation

The projection equations are most conveniently written on matrix form $\mathbf{n}_{t+1} = \mathbf{l}\mathbf{n}_t$ where $\mathbf{n}_t = (n_1, n_2, \ldots, n_k)'$ is the population vector at time t and $\mathbf{L}$ is the *projection matrix*

$$\mathbf{l} = \begin{bmatrix} f_1 & f_2 & \cdots & f_k \\ p_1 & 0 & \cdots & \\ 0 & p_2 & 0 & \cdots \\ \cdots & & & \\ \cdots & & p_{k-1} & 0 \end{bmatrix}$$

commonly called the Leslie matrix. For a stage-structured population the corresponding matrix is called a Lefkovitch matrix. Using the projection equation recursively we obtain $\mathbf{n}_t = \mathbf{l}^t\mathbf{n}_0$. The asymptotic multiplicative growth rate of the population is the dominant eigenvalue λ of the projection matrix $\mathbf{l}$. The stable age distribution denoted by the column vector $\mathbf{u} = (u_1, u_2, \ldots, u_k)'$ is the right dominant eigenvector defined by $\mathbf{lu} = \lambda\mathbf{u}$ and scaled so that $\sum u_i = 1$. The reproductive values are given by the row vector $\mathbf{v} = (v_1, v_2, \ldots, v_k)$ defined by $\mathbf{vl} = \lambda\mathbf{v}$ and scaled by $\sum u_i v_i = 1$ (exercise 5). The Perron-Frobenius Theorem for matrices with non-negative elements guarantees that the dominant eigenvalue is unique, real, and positive, and that the elements of $\mathbf{u}$ and $\mathbf{v}$ are non-negative.

We can now see immediately that the reproductive value has exactly exponential growth with multiplicative rate λ. The total reproductive value after one generation is $\mathbf{v}(\mathbf{n} + \Delta\mathbf{n}) = \mathbf{vln} = \lambda\mathbf{vn}$ and hence changes exactly by the factor λ from one generation to the next.

The sensitivity of λ with respect to one of the non-zero elements l_{ij} of $\mathbf{l}$ is defined as the partial derivative $\partial\lambda/\partial l_{ij}$. The way we have scaled the eigenvectors a result from matrix algebra gives simply $\partial\lambda/\partial l_{ij} = v_i u_j$. These sensitivities will later play an important part in the derivation of approximations to the stochastic growth rate and environmental and demographic variance of and age-structured process in a stochastic environment (see section 4.3).

4.3 Stochastic age-structured model

4.3.1 Introduction

The deterministic age structured model leading to exponential growth is a generalization of the simple deterministic multiplicative model we dealt with in chapter 1. This model was used to construct stochastic models by analyzing the stochastic properties of the multiplicative factor. Writing $N + \Delta N = \Lambda N$ we saw in the case of no density-regulation that the mean value $\mathrm{E}\Lambda = \lambda$ was constant and that the variance could be decomposed into two components generated by different mechanisms, the demographic and environmental variance, σ_d^2 and σ_e^2, with a total variance $\mathrm{var}(\Lambda) = \sigma_e^2 + \sigma_d^2/N$. The age structured model is, however, more complicated. As shown in Fig.4.1 the population do not grow exactly exponentially so the multiplicative factor is not constant even in the deterministic case. Furthermore, we had previously only two vital rates, survival and fecundity of adults, while the Leslie matrix has $2k - 1$ vital rates. Introducing stochastic vital rates in the age structured model will generate stochastic fluctuations in the multiplicative rate that will act together with the deterministic fluctuations of the type shown in Fig.4.1. Even if the environmental vector acting on the vital rates has no temporal autocorrelation, the age-structure fluctuations will therefore, through the deterministic component, create temporal autocorrelations in the multiplicative factor acting on the total population size. We shall see how this problem can be solved by utilizing the nice properties of the total reproductive rate given in 4.2.2. The two mechanisms generating stochasticity in vital rates, that is, the environment's effect on the mean value a given year and the stochastic variability between individuals, are still operating in age structured populations. We shall se how this enables us to define a stochastic growth rate as well as an environmental and demographic variance also for age structured models in such a way that these three parameters contains almost all information about the process for the total population size in the case of no density-regulation.

4.3.2 Stochastic projection matrices

A population $\mathbf{n} = (n_1, n_2, \ldots, n_k)'$ contributes to the population the next year $\mathbf{n} + \Delta\mathbf{n}$ through each individuals survival and reproduction. For example, the number of individuals of age 2 will be the number of individuals of age 1 the previous year that survive. Defining indicator variables for their survival (1 for survival 0 for death), the next years age-class is simply the sum of these indicators. Writing $\bar{J}_2$ for the mean of these indicators, the sum can be written as $n_2\bar{J}_2$. Hence, the element of the projection matrix, corresponding to p_2 in the deterministic case, is $\bar{J}_2$. Similarly, writing $\bar{B}_i$ for the mean number of offspring produced by individuals of age i, the number of offspring produced by individuals in age-class i is $n_i\bar{B}_i$. Hence, we can still write the change from one year to the next on matrix form, $\mathbf{n} + \Delta\mathbf{n} = \mathbf{Mn}$, where

$$M = \begin{bmatrix} \bar{B}_1 & \bar{B}_2 & \cdots & \bar{B}_k \\ \bar{J}_1 & 0 & \cdots & \\ 0 & \bar{J}_2 & 0 & \cdots \\ \cdots & & & \\ \cdots & & \bar{J}_{k-1} & 0 \end{bmatrix}$$

Writing as before $\mathbf{z}$ for the environmental vector acting on the population, we define F_i as the expected number of offspring produced by an individual of age i in environment $\mathbf{z}$. Writing B_i for the number of offspring produced by a randomly chosen individual of age i, we then have $F_i = \mathrm{E}(B_i|\mathbf{z})$. Similarly for the survivals we define $P_i = \mathrm{E}(J_i|\mathbf{z})$. For a large population, the mean values of the rates within years will be approximately the expected values conditioned on the environment. Hence, the matrix operating on large populations is

$$L(\mathbf{z}) = \mathrm{E}(\mathbf{M}|\mathbf{z}) = \begin{bmatrix} F_1 & F_2 & \cdots & F_k \\ P_1 & 0 & \cdots & \\ 0 & P_2 & 0 & \cdots \\ \cdots & & & \\ \cdots & & P_{k-1} & 0 \end{bmatrix}$$

The unconditional expected value of the matrices $\mathbf{M}$ and the expectation of $\mathbf{L}$ are the same through the theorem on double expectation, $\mathrm{E}\mathbf{M} = \mathrm{E}\mathrm{E}(\mathbf{M}|\mathbf{z}) = \mathrm{E}\mathbf{L}$. The deterministic theory in the previous section can now be interpreted as a theory for the expected population vector by interpreting $\mathbf{l}$ as $\mathrm{E}\mathbf{L} = \mathrm{E}\mathbf{M}$ writing $f_i = \mathrm{E}F_i = \mathrm{E}\mathrm{E}(B_i|\mathbf{z})$ and $p_i = \mathrm{E}P_i = \mathrm{E}\mathrm{E}(J_i|\mathbf{z})$. Then, if the environmental vectors $\mathbf{z}$ are independent between years we have $\mathrm{E}(\mathbf{n}_t|\mathbf{n}_0) = \mathbf{l}^t\mathbf{n}_0$ which means that the projection matrix $\mathbf{l}$ operates on the expected population $\mathrm{E}\mathbf{n}$, so that all results given for the deterministic case are valid replacing the population vector $\mathbf{n}$ by its expectation $\mathrm{E}\mathbf{n}$. The vectors $\mathbf{u}$ and $\mathbf{v}$ are in the following the right and left eigenvalues of the expected matrix $\mathbf{l}$ as defined in 4.2.3.

4.3.3 Reproductive value dynamics

We have seen that there are fluctuations in total population size in the deterministic case. However, if we work with the total reproductive value V of the population rather than the population size, these fluctuations disappear, and the deterministic growth is exactly exponential. We have also seen that if the age-distribution is close to the stable age-distribution defined by the Leslie matrix $\mathbf{l}$, then $V \approx N$. This indicates that stochastic dynamics also may be easier to describe if we work with V_t rather than N_t. The autocorrelations found in time series of age-structured populations are often mainly due to the deterministic components of the fluctuations around the exponential growth curve generated by deviations from the stable age distribution. Since the process V_t does not show these initial fluctuations we should expect the stochastic dynamics of the reproductive value, to be simpler than that of N_t. We then separate the population vector $\mathbf{n}$ into two components, one vector with components exactly proportional to the stable age-distribution, and another vector defining the deviation from this distribution. This latter component we define as

$$\mathbf{x} = \mathbf{n}/V - \mathbf{u}.$$

If the population has the stable age-distribution then $\mathbf{n} = V\mathbf{u}$ so that the corresponding value of $\mathbf{x}$ is the zero vector. The population vector can now be written as

$$\mathbf{n} = V(\mathbf{u} + \mathbf{x}).$$

The total population size N is the sum of the components, that is

$$N = V(1 + X)$$

where the notation X is used for the sum of the components of $\mathbf{x}$.

In order to investigate the dynamics of the reproductive value we write $\mathbf{M}_t = \mathbf{l} + \boldsymbol{\varepsilon}_t$ for then matrix operating at time t, where the stochastic deviation $\boldsymbol{\varepsilon}_t$ has zero expectation. The dynamics of the reproductive value is then given by

$$V_{t+1} = \mathbf{v}\mathbf{n}_{t+1} = \mathbf{v}\mathbf{M_t}\mathbf{n_t} = \mathbf{v}(\mathbf{l} + \boldsymbol{\varepsilon}_t)\mathbf{n}_t = \lambda V_t + \mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{n}_t.$$

If we ignore second order terms by approximating $\mathbf{n}_t$ in the stochastic term by a vector proportional to the stable age distribution, that is $\mathbf{n}_t \approx V_t\mathbf{u}$ we find

$$V_{t+1} = \lambda V_t(1 + \lambda^{-1}\mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{u}),$$

and $\mathrm{var}[V_{t+1}|V_t] \approx V_t^2\mathrm{var}[\mathbf{v}\boldsymbol{\varepsilon}\mathbf{u}] = V_t^2\sigma_V^2$. Writing $\mathbf{v}\boldsymbol{\varepsilon}\mathbf{u} = \sum v_i\varepsilon_{ij}u_j$, the general formula for the variance of a sum of random variables gives

$$\sigma_V^2 = \mathrm{var}[\mathbf{v}\boldsymbol{\varepsilon}\mathbf{u}] = \sum_{ij}\sum_{kl} v_i u_j v_k u_l \mathrm{cov}[\varepsilon_{ij}, \varepsilon_{kl}].$$

Since $\partial\lambda/\partial l_{ij} = v_i u_j$ and $\mathbf{M} = \mathbf{l} + \boldsymbol{\varepsilon}$ this variance can alternatively be written as

$$\sigma_V^2 = \sum_{ij}\sum_{kl} \frac{\partial\lambda}{\partial l_{ij}}\frac{\partial\lambda}{\partial l_{kl}}\mathrm{cov}[M_{ij}, M_{kl}].$$

The stochastic growth rate for the reproductive value, that is, the expected growth on log scale, is

$$s = \mathrm{E}[\ln V_{t+1} - \ln V_t|\ln V_t] \approx r + \mathrm{E}\ln(1 + \lambda^{-1}\mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{u}),$$

where $r = \ln\lambda$ is the growth of the deterministic model defined by $\mathbf{l}$. Assuming that the stochastic term is small and using the second order approximation for the logarithm we find $s \approx r - \frac{1}{2}\lambda^{-2}\mathrm{var}[\mathbf{v}\boldsymbol{\varepsilon}\mathbf{u}]$. Again using the relation $\partial\lambda/\partial l_{ij} = v_i u_j$ we obtain the following formulas for the stochastic growth rate and the variance function,

$$s = r - \frac{1}{2}\sigma_V^2 \qquad \text{and} \qquad \sigma^2 = \lambda^{-2}\sigma_V^2.$$

To this order of approximation the process for the log reproductive value can now be written as

$$\ln V_{t+1} = \ln V_t + s + \xi_t$$

where $\xi_t = \lambda^{-1}\mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{u}$ so that $\mathrm{E}\xi_t = 0$ and $\mathrm{var}[\xi_t] = \sigma^2 = \lambda^{-2}\sigma_V^2$. To the first order, this process has no autocorrelation in the noise ξ_t if the stochastic projection matrices, $\mathbf{M}_t$, are independent between years. Hence the process for the reproductive value is of the same type as the corresponding population process without age-structure, that is, a simple random walk. This implies that the process for the log reproductive value can be approximated by a diffusion with infinitesimal mean s and variance σ^2.

4.3.4 Environmental and demographic variance

The variance function σ_V^2 is a linear combination of covariances $\mathrm{cov}(M_{ij}, M_{kl})$. The elements of $\mathbf{M}$ are all mean values over all individuals in an age-class of their number of offspring or the indicator for their survival. In chapter 1 we decomposed the contribution from each individual w into its expectation and an environmental as well as a demographic component so that the mean fitness had variance $\sigma_e^2 + \sigma_d^2/N$ where $\sigma_d^2 = \mathrm{E}\mathrm{var}(w|\mathbf{z})$ and $\sigma_e^2 = \mathrm{var}\mathrm{E}(w|\mathbf{z})$. The same kind of decomposition can be done for an age structured population using the general formula for the total covariance

$$\mathrm{cov}[M_{ij}, M_{kl}] = \mathrm{E}\mathrm{cov}[M_{ij}, M_{kl}|\mathbf{z}] + \mathrm{cov}[\mathrm{E}(M_{ij}|\mathbf{z}), \mathrm{E}(M_{kl}|\mathbf{z})].$$

Here the first term describes demographic variation generated by variations conditioned on the environmental vector $\mathbf{z}$ while the second is an environmental component generated by temporal variation in $\mathbf{z}$. Using this decomposition and the previous notation $\mathrm{E}(M_{ij}|\mathbf{z}) = L_{ij}$ the expression for $\sigma^2 = \lambda^{-2}\sigma_V^2$ can be decomposed accordingly, giving

$$\sigma^2 = \sum_{ij}\sum_{kl} \lambda^{-2} v_i u_j v_k u_l \mathrm{Ecov}[M_{ij}, M_{kl}|\mathbf{z}] + \sum_{ij}\sum_{kl} \lambda^{-2} v_i u_j v_k u_l \mathrm{cov}[L_{ij}, L_{kl}].$$

The second term is the environmental variance for the age structured process. Writing as before $r = \ln\lambda$ this term can be written on the form

$$\sigma_e^2 = \sum_{ij}\sum_{kl} \frac{\partial r}{\partial l_{ij}} \frac{\partial r}{\partial l_{kl}} \mathrm{cov}[L_{ij}, L_{kl}].$$

Notice that σ^2 is the variance on the log scale. Hence, as mentioned in chapter 1, we now work on the log scale and include the factor λ^{-2} in the definitions of environmental and demographic variances.

In order to analyze the first term describing the demographic stochasticity we introduce the bivariate variables (B_i, J_i), the number of offspring and the indicator of survival for an individual of age i. The non-zero elements of the matrix M_{ij} are the mean fecundities $M_{1j} = \bar{B}_j$ and mean survivals $M_{j+1,j} = \bar{J}_j$. Conditioned on the environment these are the mean values of n_j fecundities and indicators for survival, respectively. We also assume that there are no demographic covariances (see section 2.4.5) so that the vital rates are independent between individuals when conditioned on the environment. Hence, $\mathrm{cov}[M_{ij}, M_{kl}|\mathbf{z}]$ is zero unless $j = l$. Further, $\mathrm{cov}[M_{ij}, M_{kj}|\mathbf{z}]$ is non-zero only if $i = k = 1$, or $i = k = j+1$, or if one of the indices i, k is 1 and the other $j + 1$. Defining the parameters $\sigma_{Bi}^2 = \mathrm{Evar}(B_i|\mathbf{z})$, $\sigma_{Ji}^2 = \mathrm{Evar}(J_i|\mathbf{z}) = \mathrm{E}[P_i(1 - P_i)|\mathbf{z}]$ and $\sigma_{BPi}^2 = \mathrm{Ecov}(B_i, J_i|\mathbf{z})$ we find for example (exercise 6)

$$\mathrm{Ecov}(M_{1j}, M_{1j}|\mathbf{z}) = n_j^{-1}\sigma_{Bj}^2.$$

Using similar expressions for the other contributions we find that the total demographic contribution can be written as (exercise 6)

$$\lambda^{-2}\sum_{ijkl} v_i u_j v_k u_l \mathrm{Ecov}[M_{ij}, M_{kl}|\mathbf{z}] = \lambda^{-2}\sum_{j=1}^{k} n_j^{-1} u_j^2 [v_1^2\sigma_{Bj}^2 + v_{j+1}^2\sigma_{Pj}^2 + 2v_{j+1}v_1\sigma_{BPj}^2].$$

Now, approximating the age distribution by the stable age distribution we have $n_j \approx Nu_j$, showing that the demographic term is approximately σ_d^2/N where σ_d^2 is the demographic variance

$$\sigma_d^2 = \sum_{j=1}^{k} \lambda^{-2} u_j [v_1^2 \sigma_{Bj}^2 + v_{j+1}^2 \sigma_{Pj}^2 + 2v_{j+1} v_1 \sigma_{BPj}^2].$$

Since $N \approx V$ this indicates that the process for the logarithm of reproductive value can be approximated by a diffusion with infinitesimal mean $r - \sigma_e^2/2 - \sigma_d^2/(2V)$ and variance $\sigma_e^2 + \sigma_d^2/V$. For large populations the demographic terms can be ignored and the dynamics can be approximated by a Brownian motion with mean $s = r - \sigma_e^2/2$ and variance σ_e^2.

4.3.5 Simulation examples

As a first example let us consider a population with no environmental stochasticity that at most produce one offspring each year. Then, the variables B_i as well as J_i are simple indicators, so that $\sigma_{Bi}^2 = \mathrm{Evar}(B_i|\mathbf{z}) = \mathrm{var}(B_i) = l_{1i}(1 - l_{1i})$ and $\sigma_{Ji}^2 = \mathrm{Evar}(J_i|\mathbf{z}) = \mathrm{var}(J_i) = l_{i+1,i}(1 - l_{1+1,i})$. Hence, if reproduction and survival are independent, this model has the interesting property that the stochasticity is defined uniquely by the elements of the expected Leslie matrix, the expression for the demographic variance being

$$\sigma_d^2 = \sum_{i=1}^{k} \lambda^{-2} u_i [v_1^2 l_{1i}(1 - l_{1,i}) + v_{i+1}^2 l_{i+1,i}(1 - l_{i+1,i})].$$

We have seen in section 3.11.3 that the cumulative distribution for the diffusion approximation to this model is given by a simple analytical expression. Fig.4.2 shows some simulation examples giving the cumulative distribution for the time to extinction for this age structured process together with the cumulative distribution for the corresponding diffusion approximation assuming $N_t \approx V_t$.

In Fig.4.3 a similar simulation example is shown with environmental as well as demographic stochasticity and 15 age-classes. In the right panel the demographic stochasticity has been ignored. We see that the time to extinction is

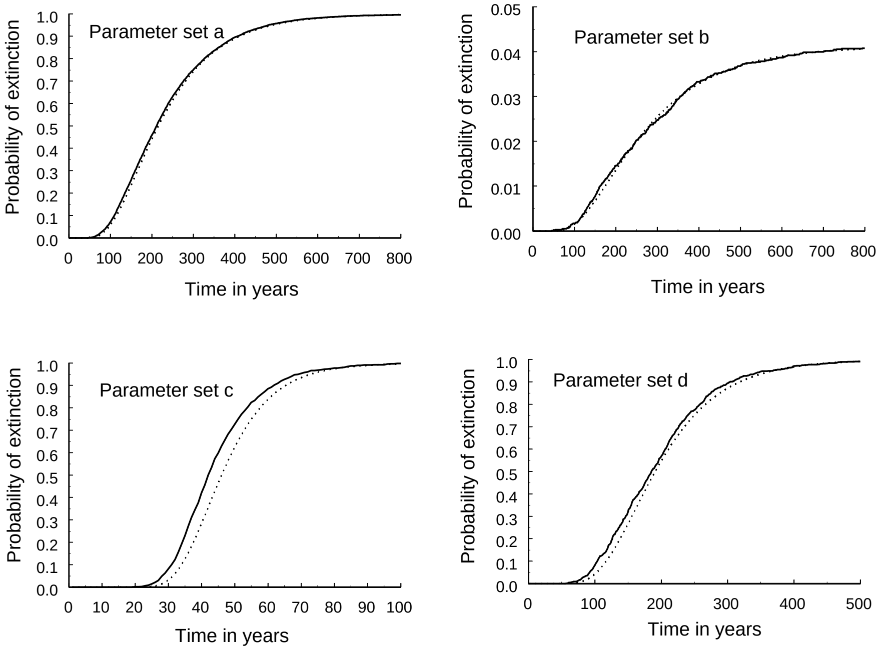

Figure 4.2: Cumulative distributions of time to extinction for processes without environmental stochasticity. Dotted lines are the cumulative distribution for the diffusion approximations, while the solid lines are based on stochastic simulations of the age-structured process. The individual yearly fecundity is restricted to take values 0 or 1. Demographic stochasticity is then uniquely determined by the elements of the projection matrix. The initial population size is 100. Parameter sets a and b are for populations with 4 age classes. For set a the survivals are (0.4, 0.6, 0.9, 0.7), the fecundities (0, 0.6, 0.9, 0.7). For set b the growth rate is slightly positive so that the probability of ultimate extinction is smaller than one. The survival rates in set b are the same as in set a while the fecundities are (0, 0.6, 0.9, 0.8). Parameter sets c and d are populations with 9 age-classes. The survivals in set c are (0.6, 0.7, 0.7, 0.8, 0.9, 0.9, 0.9, 0.7, 0.5), and the fecundities (0, 0, 0, 0.4, 0.5, 0.8, 0.8, 0.9, 0.8). The survivals in set d are (0.5, 0.7, 0.7, 0.8, 0.8, 0.9, 0.9, 0.7, 0.5) and the fecundities are (0, 0, 0, 0, 0.2, 0.5, 0.6, 0.9, 0.8).

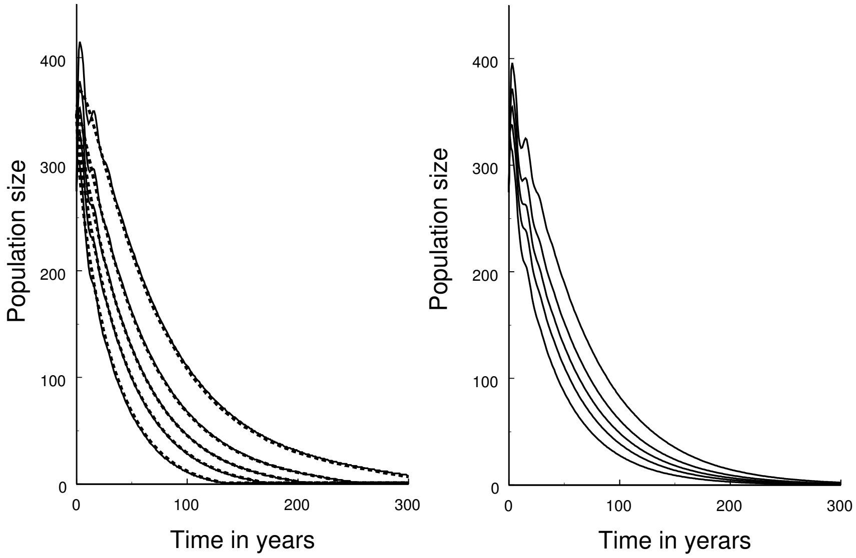

Figure 4.3: Stochastic simulations of an age-structured model with Poisson lognormally distributed number of offspring (exercise 7-9) with variance 3 times the mean value. The only parameter with environmental stochasticity is the first year survival P_1 assumed to have variance 0.04 between years. The solid lines from bottom to top in the left panel are the 0.05, 0.25, 0.50, 0.75 and 0.95 quantiles computed from 100,000 simulations of the full age-structured model while the dotted lines are from the corresponding simulations of the diffusion approximation. The right panel shows the same simulations of the full model ignoring demographic stochasticity, that is, using the matrix L instead of M. There are 15 age classes and the following parameters: $(f_1, f_2, \ldots, f_{15}) = (0, 0, 0, 0, 0, 0, 0, 0, 0.5, 0.6, 0.7,$ $0.9, 1.0, 1.0, 1.0, 1.0)$, and $(p_1, p_2, \ldots, p_{15}) = (0.5, 0.6, 0.7, 0.9, 0.95, 0.95, 0.95,$ $0.95, 0.9, 0.9, 0.85, 0.85, 0.8, 0.8, 0.6)$, and initial population size 20 for classes 1 to 10 and 15 for classes 11 to 15. This gives $\lambda = 0.9813$, $\sigma_e^2 = 0.0011$, and $\sigma_d^2 = 0.2615$.

then considerably larger, demonstrating how the demographic stochasticity also for age structured populations speeds up the extinction process. Notice the fluctuations of the quantiles in the beginning of the simulations. These are generated by the same kind of fluctuations as we saw in the deterministic case in Fig.4.1. In this case these fluctuations are rather large and last for about 30 years because the simulations are started far from the stable age distribution and there are as many as 15 age-classes. Nevertheless, the quantiles for the reproductive value found from the diffusion approximation are very close to the quantiles for the population size after 30 years.

4.3.6 Fluctuations in age-structure

We have seen in some simulation examples that the diffusion approximation for the reproductive value process is often a good approximation to the total population size for the full age structured model. However, we have also seen that the reproductive value may be different from the total population size if the age distribution deviates from the stable age distribution derived from the expected projection matrix. In this section we analyze this deviation in some more details. The deviation has been expressed by the vector $\mathbf{x} = \mathbf{n}/V - \mathbf{u}$. The basic dynamics of the population age vector, $\mathbf{n}_{t+1} = (\mathbf{l} + \varepsilon_t)\mathbf{n}_t$, can be linearized by substituting $\mathbf{n}_t = V_t(\mathbf{u} + \mathbf{x}_t)$ and neglecting second-order terms by assuming that the noise ε_t and the deviation from the stable age distribution $\mathbf{x}_t$ are both small,

$$\mathbf{n}_{t+1} = \mathbf{l}\mathbf{n}_t + V_t \boldsymbol{\varepsilon}_t \mathbf{u}$$

Premultiplying both sides by $V_t^{-1}(\mathbf{I} - \mathbf{u}\mathbf{v})$, where $\mathbf{I}$ is the identity matrix, using the definition of $\mathbf{x}_t$ produces

$$\frac{V_{t+1}}{V_t}\mathbf{x}_{t+1} = \mathbf{l}\mathbf{x}_t + (\mathbf{I} - \mathbf{u}\mathbf{v})\boldsymbol{\varepsilon}_t \mathbf{u}.$$

In 4.3.3 we showed that $V_{t+1} = \lambda V_t(1 + \lambda^{-1}\mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{u})$ or $V_{t+1}/V_t = \lambda + \mathbf{v}\boldsymbol{\varepsilon}_t\mathbf{u}$. Inserting this in the above equation then shows that to first order (exercise

11)

$$\mathbf{x}_{t+1} = \lambda^{-1}\mathbf{l}\mathbf{x}_t + \lambda^{-1}(\mathbf{I} - \mathbf{u}\mathbf{v})\boldsymbol{\varepsilon}_t\mathbf{u}.$$

From the definition of $\mathbf{x}_t = \mathbf{n}_t/V_t - \mathbf{u}$ we see that $\mathbf{u}\mathbf{v}\mathbf{x}_t = \mathbf{0}$. Utilizing this we find that the dynamics of the residual variation around the stable age distribution can be written as (exercise 12)

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \lambda^{-1}(\mathbf{I} - \mathbf{u}\mathbf{v})\boldsymbol{\varepsilon}_t\mathbf{u}$$

where $\mathbf{A} = \lambda^{-1}\mathbf{l} - \mathbf{u}\mathbf{v}$.

One can show that the eigenvalues of $\mathbf{A}$ are $0, \lambda_2/\lambda, \lambda_3/\lambda, \ldots$ (exercise 13). This clarifies that increasing powers of $\mathbf{A}$ approach the zero matrix, such that if $\boldsymbol{\varepsilon}_t$ is a stationary process the fluctuations in $\mathbf{x}_t$ approach a stationary process. Solving recursively produces (exercise 14)

$$\mathbf{x}_t = \lambda^{-1}\sum_{p=1}^{\infty}\mathbf{A}^{p-1}(\mathbf{I} - \mathbf{u}\mathbf{v})\boldsymbol{\varepsilon}_{t-p}\mathbf{u}.$$

Since $N_t = V_t(1 + X_t)$ there are two different properties of X_t that are of great interest when analyzing age structured models. First, in order to check the approximation $N_t \approx V_t$ we can calculate the standard deviation in the stationary distribution of X_t. If this is much smaller than 1 the approximation is likely to be good at any point of time. Secondly, it is interesting to study the memory of the process X_t, that is, how long time it takes before it returns to its equilibrium value which is zero. This last property is given by the above autocovariance function. If the autocorrelation at a given time distance τ is small, X_t contains practically no information about $X_{t+\tau}$ so that the distribution of $X_{t+\tau}$ conditioned on X_t is approximately the stationary distribution of X_t.

Let us assume that the population is large enough for the demographic variance to be ignored so that the $\mathbf{M} = \mathbf{L}$ has elements with constant variances and covariances. To derive the moments of the stationary distribution of X_t in this case, first note that $\mathrm{E}\mathbf{x}_t = 0$ so $\mathrm{E}X_t = 0$. The autocovariance function for X_t, given by $\mathrm{cov}[X_t, X_{t+\tau}]$, can be obtained from the sum of elements of the matrix $\mathrm{E}[\mathbf{x}_t\mathbf{x}'_{t+\tau}]$ where $'$ denotes matrix transposition,

$$\mathrm{E}[\mathbf{x}_t\mathbf{x}'_{t+\tau}] = \lambda^{-2}\sum_{p=1}^{\infty}\mathbf{G}_p\mathbf{S}\mathbf{G}'_{p+\tau}$$

where $\mathbf{G}_p = \mathbf{A}^{p-1}(\mathbf{I} - \mathbf{uv})$ and $\mathbf{S} = \mathrm{E}[\boldsymbol{\varepsilon}\mathbf{uu}'\boldsymbol{\varepsilon}']$ (exercise 15). The elements of $\mathbf{S}$ are given by $S_{ij} = \sum_{kl} u_k u_l \mathrm{cov}[L_{ik}, L_{jl}]$. The above expression is useful for numerical computations of $\mathrm{cov}[X_t, X_{t+\tau}]$ and $\mathrm{corr}[X_t, X_{t+\tau}] = \mathrm{cov}[X_t, X_{t+\tau}]/\mathrm{cov}[X_t, X_t]$.

Fig.4.4 shows the log population size and log reproductive value for a stochastic age structured population. The lower panel shows the process $1 + X_t$. Notice that the fluctuations in X_t are rather small even if there is large stochasticity in the model. The standard deviations of log fecundities are as large as 0.2 while the standard deviation in $\ln(P_1)$ is approximately 0.07. The environmental variance in this model is $\sigma_e^2 = 0.002548$.

4.3.7 Estimating demographic and environmental variance

Let us first consider a population that is large enough for the demographic variance to be ignored. We assume that the population is fully censused with correct determination of age for each individual. If there is no density-regulation we can then first estimate the expected leslie matrix $\mathbf{l}$ just using the mean values of the observed survivals and fecundities for each age class. We then compute the dominant eigenvalue for this matrix which is then the estimate $\hat{\lambda}$ of λ, and the corresponding left and right eigenvalues $\hat{\mathbf{v}}$ and $\hat{\mathbf{u}}$ as estimates of $\mathbf{v}$ and $\mathbf{u}$. From these estimates we can compute the log of the total reproductive, $\ln V_t$, for the time interval with observations, as shown in Fig.4.4. The sequence of differences $\ln \hat{V}_{t+1} - \ln \hat{V}_t = \hat{s} + \hat{\varepsilon}_t$ is the approximately a sequence of independent random variables with variance σ_e^2 which we then estimate by standard sum of square.

If the demographic variance has to be taken into account we need data on individual survival and reproduction as in the case of no age-structure dealt

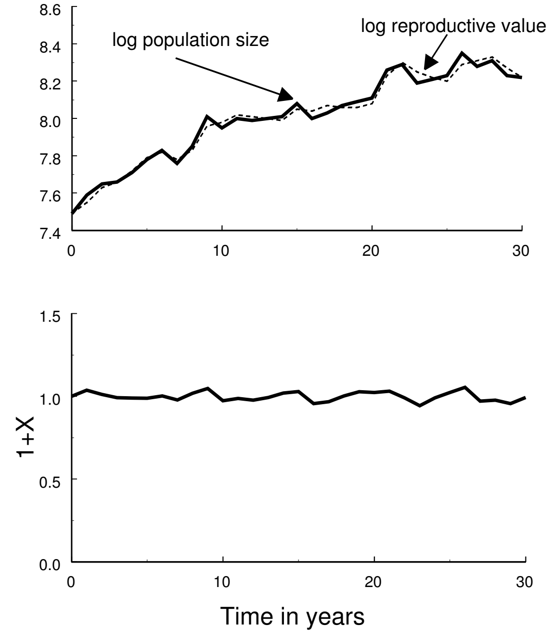

Figure 4.4: The upper panel shows log population size and log reproductive value as function of time. There are 7 age classes. The mean fecundities $f_1 - f_7$ are 0, 0, 0.4, 0.8, 1.0, 0.8, 0.5, and mean survivals $p_1 - p_7$ are 0.5, 0.9, 0.9, 0.8, 0.7, 0.6 and $p_7 = 0$ giving $\lambda = 1.0161$. The only stochastic survival is $P_1 = p_1 \exp(-0.5\sigma_p^2 + \sigma_p z_p)$ where z_p standard normally distributed. All fecundities are stochastic with $F_1 = f_1 \exp(-0.5\sigma_f^2 + \sigma_f z_f)$, where z_f also is a standard normal variate. The correlation between z_p and z_f is 0.4, $\sigma_p^2 = 0.005$ and $\sigma_f^2 = 0.04$. This gives an environmental variance $\sigma_e^2 = 0.002548$.

with in chapter 2.5. To obtain an estimate of σ_d^2 we need to estimate all the components in the expression for σ_d^2, which in addition to λ, $\mathbf{v}$ and $\mathbf{u}$ are the variance components σ_{Bj}^2, σ_{Pj}^2 and σ_{BPj}^2. These components can be estimated by sum of squares over realized rates within years for each age group to obtain an estimate of σ_d^2 valid for the population size that year (exercise 17). This estimate can be plotted against population size.

If the demographic variance has to be taken into account we first estimate the demographic variance form individual data. From the time series of reproductive values we can then estimate the values of $\sigma_e^2 + \sigma_d^2/N$ each year and subtract the demographic term. These yearly estimates should be plotted against N as we did in chapter 2.5.

4.3.8 Density-regulation

In general the term density-regulation is used for the fact that the vital rates, survivals and reproduction, may depend on the population size of the species under consideration. For an age-structured population this means that the distribution of the components of the projection matrix depends on the sizes of the various age-classes. This will in general lead to very complicated models, but for relatively small fluctuations around an equilibrium, a simple linearization may be rather accurate. Assuming that there is a stable equilibrium $\hat{\mathbf{n}}$ with corresponding total populations size K, we write $\hat{\mathbf{l}}$ for the expected projection matrix at the equilibrium. This matrix has leading eigenvalue equal to 1 and corresponding left and right eigenvector given by $\mathbf{v}\hat{\mathbf{l}} = \mathbf{v}$ and $\hat{\mathbf{l}}\mathbf{u} = \mathbf{u}$. The deviation from equilibrium can be denoted as vector $\mathbf{y}$ defined by the relation $\mathbf{n} = K(\mathbf{u} + \mathbf{y}) = \hat{\mathbf{n}} + K\mathbf{y}$, where K is the carrying capacity. For a population $K(\mathbf{u}+\mathbf{y})$ we linearize the expected matrix writing $\mathbf{l}(\mathbf{n}) = \hat{\mathbf{l}} + \boldsymbol{\Delta}\mathbf{l}$ where $\Delta l_{ij} = K \sum_k y_k \frac{\partial l_{ij}}{\partial n_k}$ where the derivatives are evaluated at the equilibrium. Adding a noise matrix $\boldsymbol{\varepsilon}_t$ to the expected matrix as before, we obtain the dynamic model

$$\mathbf{u} + \mathbf{y}_{t+1} = (\hat{\mathbf{l}} + \boldsymbol{\Delta}\mathbf{l} + \boldsymbol{\varepsilon}_t)(\mathbf{u} + \mathbf{y}_t).$$

Ignoring all second order terms and using $\hat{\mathbf{l}}\mathbf{u} = \mathbf{u}$ produces

$$\mathbf{y}_{t+1} = \hat{\mathbf{l}}\mathbf{y}_t + \boldsymbol{\Delta}\mathbf{l}\ \mathbf{u} + \boldsymbol{\varepsilon}_t\mathbf{u}.$$

Finally we can express $\boldsymbol{\Delta}\mathbf{l}\ \mathbf{u}$ as $-\mathbf{D}\mathbf{y}$ where $D_{ij} = -K\sum_k \frac{\partial l_{ik}}{\partial n_j}u_k$ giving the dynamic equation

$$\mathbf{y}_{t+1} = (\mathbf{l} - \mathbf{D})\mathbf{y}_t + \boldsymbol{\varepsilon}_t\mathbf{u}.$$

Transforming back to the population vector by substituting $\mathbf{y} = \mathbf{n}/K - \mathbf{u}$ on both sides then gives

$$\mathbf{n}_{t+1} = \mathbf{A}\mathbf{n}_t + \mathbf{B} + \boldsymbol{\epsilon}_t$$

where $\mathbf{A} = \mathbf{l} - \mathbf{D}$ is a stability matrix, $\mathbf{B} = K(\mathbf{I} - \mathbf{A})\mathbf{u}$ is a constant vector and the noise vector is $\boldsymbol{\epsilon}_t = K\boldsymbol{\varepsilon}_t\mathbf{u}$.

If $\mathbf{A}$ and $\mathbf{B}$ are known or estimated from data, the population vector at equilibrium is given by $K\mathbf{u}$ where $\mathbf{u} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}/K$ (exercise 18). We now write $\mathbf{w}$ for the left eigenvalue of the stability matrix $\mathbf{A}$ and scale it by the relation $\mathbf{w}\mathbf{u} = 1$ we now define the total reproductive value of the population as $V_t = \mathbf{w}\mathbf{n}_t$. Multiplying the dynamic equation by $\mathbf{w}$ and redefining λ to be the dominant eigenvalue of $\mathbf{A}$ then leads to

$$V_{t+1} = \lambda V_t + b + \delta_t$$

where $b = \mathbf{w}\mathbf{B}$ and $\delta_t = \mathbf{w}\boldsymbol{\epsilon}_t = K\mathbf{w}\boldsymbol{\varepsilon}_t\mathbf{u}$.

Although this looks like a simple first order autoregressive model, the situation may in fact be more complicated. When there is no density-regulation the Leslie matrix has a form that ensures the existence of a real dominant eigenvalue. The matrix $\mathbf{A}$, however, may have a number of negative elements, and the dominant eigenvalue may be complex. In that case the reproductive value given above is also complex and the equation for the reproductive value describes a two-dimensional first order autoregressive model, the two dimensions being the real and imaginary part of V_t. Notice however, that we have scaled the reproductive value in such a way that it is exactly equal to the population size K at equilibrium so that the equilibrium value of the imaginary component necessarily is zero. When the density-regulation is week the dominant eigenvalue may still be real so that the process for the total population size can be approximated by the above one-dimensional process for the reproductive value. If the eigenvalue with the second largest modulus has modulus close to that of the dominant one, this approximation may not be very accurate since the fluctuation around the reproductive value then may interfere with the process for the reproductive value giving a more complex dynamics. If λ is real we can define the strength of density-regulation as we did for a population without age-structure, that is, as $\gamma = 1 - \lambda$. In the complex case, however, the process is more complicated, and the density-regulation depends on the imaginary as well as the real part of the dominant eigenvalue.

Estimation may be performed by simple linear regression if population data for all age-classes are available. Then, by the linear dynamic equation the number of individuals in an age-class may be chosen as the dependent variable, while all age-classes the previous year are covariates. One regression must be done for each age-class to obtain an estimate of $\mathbf{A}$ and $\mathbf{B}$. Since the stochastic terms are likely to be dependent, uncertainties are best investigated by bootstrapping, simulating the process from its initial value choosing the noise vector randomly from the set of residual noise vectors obtained after fitting the model.

Fig.4.5 shows an example of a fitted model for an island population of red-deer where all individuals of age larger than 5 years are collected in one last age-class. The real part of λ is estimated to be 0.64, while the imaginary part is -0.39 (the sign of the imaginary part is not unique since the complex eigenvalues occur in complex conjugated pairs). Fig. 4.6 shows the residuals in the complex plane.

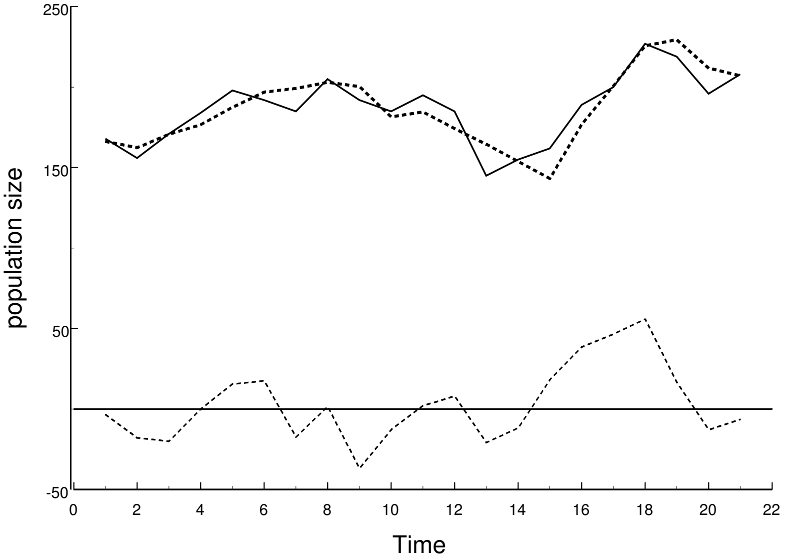

Figure 4.5: The upper graphs are the observed total population size of an island red-deer population (solid line) together with the real part of the reproductive value (dotted line). The lower graph is the imaginary part of the reproductive value fluctuating around zero.

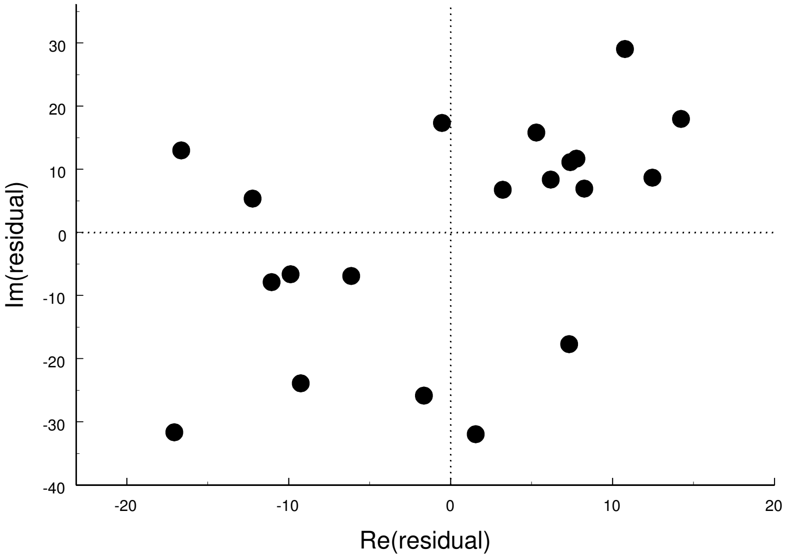

Figure 4.6: The complex residuals for the red-deer, that is, the deviations between the complex V_{t+1} and the prediction $\lambda V_t + b + \delta_t$.

4.4 Exercises

1. Assume that individuals in the k'th age-class has survival p_k and remain in the k'th age-class, which is then a stage rather than an age, if they survive (as in Fig.4.1). Find the equation for the asymptotic growth rate λ in this case using the fact that the population approaches a stable age/stage-distribution with exponential growth at rate λ.

2. Show that the generation time defined as the mean age of mothers of newborn is $T = \sum_{i=1}^{k} i l_i f_i \lambda^{-i}$.

3. Show that $v_i' = \frac{\lambda^i}{l_i} \sum_{j=i}^{k} l_j f_j \lambda^{-j}$ is a solution to the set of linear equations for the reproductive values and that the reproductive values are $v_i = v_i' / (\sum u_i v_i')$.

4. Use the set of linear equations $f_i v_1 + p_i v_{i+1} = \lambda v_i$ (with $p_k = 0$) and the projection equations for $n_i + \Delta n_i$ to show that the reproductive value grows exactly exponentially with multiplicative rate λ, that is, $\sum (n_i + \Delta n_i) v_i = \lambda \sum n_i v_i$.

5. Show that the scaled right and left eigenvectors are identical to the stable age-distribution and the reproductive values as defined in section 4.2.1 and 4.2.2.

6. Write out the details showing that $\sigma_d^2 = \sum_{j=1}^{k} \lambda^{-2} u_j [v_1^2 \sigma_{Bj}^2 + v_{j+1}^2 \sigma_{Pj}^2 + 2 v_{j+1} v_1 \sigma_{BPj}^2]$.

7. Suppose that the distribution of a variable X conditioned on λ is Poisson with mean λ. The unconditional distribution of X is then called a *Poisson mixture*. Find the mean and variance of X expressed by the mean and variance of λ.

8. Find an expression for the distribution of X in exercise 7 when λ has the gamma distribution with scale parameter α and shape parameter k. This distribution is called the *negative binomial distribution*. What is the mean and variance of X expressed by α and k?

9. The Poisson lognormal distribution is the Poisson mixture for which the mean of the Poisson distribution has the lognormal distribution. Find the mean and variance of X expressed by the mean μ and variance σ^2 of $\ln \lambda$.

10. Find the parameters σ^2_{Bi} and σ^2_{Ji} for the simulation model in Fig.4.3.

11. Show that $\mathbf{x}_{t+1} = \lambda^{-1}\mathbf{l}\mathbf{x}_t + \lambda^{-1}(\mathbf{I} - \mathbf{uv})\boldsymbol{\varepsilon}_t\mathbf{u}$ when higher order terms are neglected.

12. Show that the dynamics of the residual variation around the stable age distribution can be written as $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \lambda^{-1}(\mathbf{I} - \mathbf{uv})\boldsymbol{\varepsilon}_t\mathbf{u}$ where $\mathbf{A} = \lambda^{-1}\mathbf{l} - \mathbf{uv}$.

13. Show that the eigenvalues of the matrix $\mathbf{A} = \lambda^{-1}\mathbf{l} - \mathbf{uv}$ are $0, \lambda_2/\lambda, \lambda_3/\lambda, \ldots$. Hint: The Leslie matrix $\mathbf{l}$ can be diagonalized, that is, there exists a matrix $\mathbf{P}$ so that $\mathbf{P^{-1}lP} = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix with $d_{ii} = \lambda_i$, the i'th line of $\mathbf{P}^{-1}$ is the left eigenvector of $\mathbf{l}$ associated with λ_i and the j'th column of $\mathbf{P}$ is right eigenvector associated with λ_j.

14. Solve the dynamic equation in exercise 12 recursively to find $\mathbf{x}_t = \lambda^{-1}\sum_{p=1}^{\infty}\mathbf{A}^{p-1}(\mathbf{I} - \mathbf{uv})\boldsymbol{\varepsilon}_{t-p}\mathbf{u}$.

15. Show that $\mathrm{E}[\mathbf{x}_t\mathbf{x}'_{t+\tau}] = \lambda^{-2}\sum_{p=1}^{\infty}\mathbf{G}_p\mathbf{S}\mathbf{G}'_{p+\tau}$ where $\mathbf{G}_p = \mathbf{A}^{p-1}(\mathbf{I} - \mathbf{uv})$ and $\mathbf{S} = \mathrm{E}[\boldsymbol{\varepsilon}\mathbf{uu}'\boldsymbol{\varepsilon}']$.

16. For the model in Fig.4.4 show that the non-zero elements $\mathrm{cov}(L_{ij}, L_{kl})$ are $\mathrm{cov}(L_{1i}, L_{1j}) = f_i f_j(e^{\sigma_f^2} - 1)$, $\mathrm{cov}(L_{21}, L_{21}) = p_1^2(e^{\sigma_p^2} - 1)$ and $\mathrm{cov}(L_{21}, L_{1j}) = p_1 f_j(e^{\rho\sigma_p\sigma_f} - 1)$.

17. Show how to estimate the component σ^2_{B3}, σ^2_{P3} and σ^2_{BP3} in the expression for the demographic variance on the basis of recorded reproduction and survival for a number of individuals in age-class 3 a given year.

18. In the linear model with density-dependence show that the solution for $\mathbf{u}$ given by solving $\mathbf{n}_{t+1} = \mathbf{A}\mathbf{n}_t + \mathbf{B} + \boldsymbol{\epsilon}_t$ is the right eigenvector of $\hat{\mathbf{l}}$.

Chapter 5

Some applications

5.1 Harvesting

5.1.1 Introduction

We shall now see how some of the results given in earlier chapters may be used to explore the effects of harvesting from populations using different harvesting strategies. When harvesting (fishing hunting) from wild populations one is always faced with the conflict between the goal of hunting a large amount for economic reasons, and that of preserving the population. On the other hand, these two goals are only in conflict with one another under a short time horizon. In order to make an economic profit over a long time interval, sustainability of the population necessarily has to be taken into account. A population that is driven close to extinction cannot lead to economic profit over long time intervals. A well known example is the overfishing of the Norwegian spring spawning Herring towards the end of the sixties.

Fig.5.1 shows estimates of the spawning stock of this species during a 75 years period. A population collapse in 1967-68 due to overfishing also led to a collapse of the herring fisheries that lasted for 30 years.

In this chapter we shall focus on some relatively simple ways of calculating harvesting statistics for population processes that can be approximated by

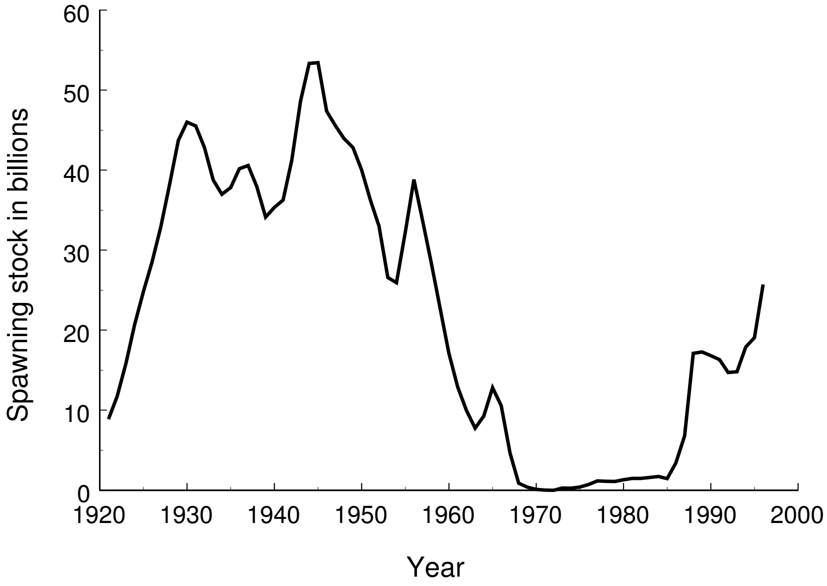

Figure 5.1: The spawning stock of the Norwegian spring spawning Herring during the period 1921 to 1996.

diffusions. These statistics can be used to explore the long term effects of adopting different types of harvesting strategies, such as the mean amount harvested, the distribution of yield between years and the expected time to extinction or quasi-extinction.

Stochastic and deterministic harvesting models turns out to be rather different although there are similarities. If the population growth is deterministic, the population should obviously be kept at the population size where the absolute growth ΔN takes its maximum value. For example, for a discrete logistic deterministic model $\Delta N = rN(1 - N/K)$ the maximum growth is obtained for $N = K/2$ giving $\Delta N = rK/4$ which is then the constant amount that can be harvested each season (exercise 1). However, for the corresponding stochastic model $\Delta N = r_t N - \bar{r}N^2/K$ where r_t is subject to stochastic fluctuations between seasons with mean $\bar{r}$, the stochastic process N_t obtained after subtracting the yield must be studied. The possibility of extinction and long periods with very small population sizes must be explored in detail.

5.1.2 Diffusion model

If a population process can be approximated by a Markov process, harvesting strategies should simply be some function of the population size expressing how much should be harvested each season. If the population estimates are uncertain, harvesting has to be based on the estimates and the uncertainty in the estimates has to be taken into account. Initially we assume that the population size is known. We write $\mu_0(n) = \mathrm{E}(\Delta N | N = n)$ and $\nu_0(n) = \mathrm{var}(\Delta N | N = n)$ for the expectation and variance of the change in population size in the absence of harvesting. Writing $y(n)$ for the yield at population size n, the diffusion approximation has infinitesimal mean and variance

$$\mu(n) = \mu_0(n) - y(n)$$

and

$$\nu(n) = \nu_0(n)$$

provided that the yield can be chosen exactly and do not create any additional stochastic terms. Our aim is now to investigate properties of the stochastic process expressing the population size N_t as well as the process expressing the harvest, that is $y(N_t)$.

Provided that the harvest is not so aggressive that the population is soon driven towards extinction, the quasi-stationary distribution of N and $y(N)$ will give considerable information about the consequences of adopting strategy $y(n)$. The possibility of extinction or quasi-extinction can be explored by calculating the expected time until the process first reaches the chosen barrier.

In chapter 3 we have seen how to calculate the Green function $G(n, n_0)$ for the diffusion defined by $\mu(n)$ and $\nu(n)$ which is initially at population size n_0. Notice that the Green function now depends on the harvesting strategy $y(n)$ since this is subtracted in the infinitesimal mean. Choosing the extinction barrier at $N = 0$ the expected time to extinction is given by

$$T(n_0) = \int_0^\infty G(n, n_0) dn.$$

A statistic that gives considerable insight is the expected total yield obtained before the population eventually goes extinct when the initial population size is n_0 which we denote $Y(n_0)$. Mathematically this quantity is $\mathrm{E} \int_0^\infty y(N_t) dt$. We have seen in chapter 3 that this also can be computed from the Green function as

$$Y(n_0) = \mathrm{E} \int_0^\infty y(N_t) dt = \int_0^\infty y(n) G(n, n_0) dn.$$

The quasi-stationary distribution (see section 3.10) is

$$f(n) = G(n, n_0)/T(n_0).$$

Since the yearly harvest $y(N)$ is just a function of the population size N we can now find the quasi-stationary distribution of the harvest $y(N)$ by a straightforward transformation. In particular, the mean annual harvest relative to the quasi-stationary distribution is

$$\mathrm{E}y(N) = \int_0^\infty y(n)f(n)dn = Y(n_0)/T(n_0)$$

while the variance of the yield is

$$\mathrm{var}[y(N)] = \int_0^\infty [y(n) - \mathrm{E}y(N)]^2 f(n)dn.$$

5.1.3 Some harvesting strategies

Some harvesting strategies can be performed without having any knowledge about the population size. Others require that the population size is known. A third class of strategies are those based on a yearly estimate of the population size and can be applied when such estimates are available, in which case we need some information on the uncertainty of the estimator in order to assess the properties of the harvesting tactic.
Constant harvesting
For the industry it is important to have a rather stable amount available each year. In order to analyze the effect of a constant harvest we can consider the model with infinitesimal mean

$$\mu(n) = \mu_0(n) - y$$

and $\nu(n) = \nu_0(n)$, where y is now the constant yearly harvest independent of population size. In chapters 2 and 3 we described density-regulated populations by writing $\mu_0(n) = \bar{r}n - g(n)$ so that constant harvesting leads to

$$\mu(n) = rn - g(n) - y.$$

This equation immediately indicates what is the major problem with this harvesting tactic. Since $g(n)$ is an increasing function and $g(0) = 0$, we see that the infinitesimal mean becomes negative if $n < y/r$ (see also exercise 2). This means that, if the population may reach values close to this due to harvesting and stochastic effects, then the population will almost certainly reach extinction rather soon. Numerical analysis of this type of model has also shown that the constant harvesting tactic has a destabilizing effect of the population and lead to short expected time to extinction.

So, unless one actually relax this tactic by stopping the harvest when the population size goes below a prescribed value (which is no longer constant harvesting tactic), or we harvest at a very small constant rate, constant harvesting is not sustainable and can not be recommended.

Proportional harvesting

It may often be a realistic assumption that the harvest for a given harvesting effort is proportional to the population size. Then, for a given effort we have

$$\mu(n) = \mu_0(n) - cn = (r - c)n - g(n)$$

and $\nu(n) = \nu_0(n)$, and c is a parameter indicating the harvesting effort. We see that the effect of this harvesting on the population dynamics is found simply by replacing the growth rate at small population sizes r by $r - c$ keeping the density-regulation $g(n)$ unchanged.

We can exemplify this by the stationary logistic model with no demographic stochasticity and extinction barrier at $n = 0$ given in 3.9.1.

Writing the infinitesimal mean as $\mu_0(n) = rn(1 - n/K) = rn - an^2$ where $a = r/K$, we first observe, in the case of no stochasticity ($\sigma_e^2 = 0$) that the optimal value of c is $r/2$ giving maximum constant harvest $Y = rK/4$ (exercise 3).

With constant environmental variance $\sigma_e^2 > 0$ the mean (see 3.9.1) of the stationary distribution when there is no harvesting is

$$\mathrm{E}N = K\left(1 - \frac{\sigma_e^2}{2r}\right) = \frac{r}{a} - \frac{\sigma_e^2}{2a}.$$

Replacing r by $r - c$ and multiplying by c we find the expected harvest

$$\mathrm{E}y = \mathrm{E}(cN) = c\left(\frac{r - c}{a} - \frac{\sigma_e^2}{2a}\right)$$

which is maximized for $c = (r - \sigma_e^2/2)/2 = s/2$ corresponding to the maximum mean harvest

$$\mathrm{E}y_{max} = \frac{s^2K}{4r}$$

in accordance with the results for the deterministic model when $\sigma_e^2 = 0$. We see in this model that increasing stochasticity leads to a reduction in the maximum mean harvest, which is obtained by a more conservative tactic, that is, by reducing the constant harvesting effort c.

For the same model with the Gompertz type of density regulation (exercise 5) dealt with in 3.6.1 and 3.8.4 we have

$$\mu_0(n) = r_1 n(1 - \ln n/\ln K)$$

giving, using the above approach (exercise 4),

$$\mathrm{E}y = c\, e^{\ln K(r_1 - \sigma_e^2/4 - c)/r_1}$$

which is maximized for $c = r_1/\ln K$ giving (exercise 5)

$$\mathrm{E}y_{max} = K^{1-\sigma_e^2/(4r_1)} r_1 e^{-1}/\ln K.$$

Fig.5.2 shows how the maximum mean harvest depends on the carrying capacity for this model.

Threshold harvesting

One can show that a number of different optimization problems for the diffusion model has the so-called threshold tactic as solution. This tactic is formally defined by a threshold c so that the harvest $y(n;c) = 0$ if $n \leq c$ and $y(n;c) = \infty$ if $n > c$. Hence, if the population size is below the threshold no harvesting should be done, and as soon the population exceeds c it should be harvested down to c. For example, one can show that the total expected yield before extinction in general is maximized by choosing the threshold c exactly at the carrying capacity.

In continuous time threshold harvesting implies that the population never exceeds the threshold c. The Green function for the process turns out to be

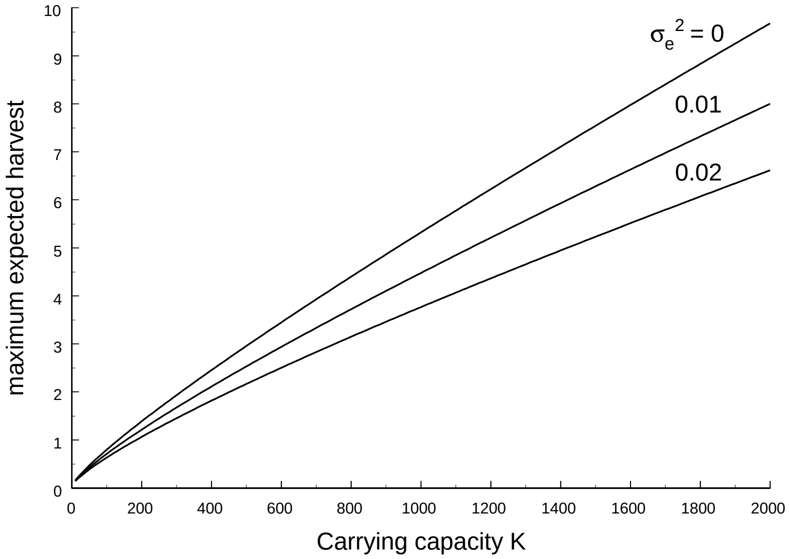

Figure 5.2: Maximum mean harvest in the Gompertz model as function of the carrying capacity for different values of the environmental variance. The growth rate is $r_1 = 0.1$.

equivalent to the Green function for the process without harvest for $n \leq c$ and zero for $n > c$ (exercise 6). Hence, if the Green function without harvesting is $G_0(n, n_0)$, the expected time to extinction at $n = 0$ for a threshold at c is simply given by

$$T(n_0) = \int_0^c G_0(n, n_0)dn.$$

In order to find expressions for the harvesting statistics for this model we first rewrite the expression for the total harvest Y before extinction inserting $y(n) = \mu_0(n) - \mu(n)$ giving

$$Y(n_0) = \int_0^\infty y(n)G(n, n_0)dn = \int_0^\infty \mu_0(n)G(n, n_0)dn - \int_0^\infty \mu(n)G(n, n_0)dn.$$

The last integral is, writing T^* for the stochastic extinction time

$$\int_0^\infty \mu(n)G(n, n_0)dn = \mathrm{E}\left(\int_0^{T^*} \mathrm{E}(\frac{dN}{dt}|N = n)\right) dt = \mathrm{E}[N(T^*)-N(0)] = -n_0.$$

The equation for the total expected yield before extinction then becomes

$$Y(n_0) = n_0 + \int_0^\infty \mu_0(n)G(n, n_0)dn.$$

Hence, for threshold harvesting with a threshold at c the total expected yield before extinction is

$$Y(n_0) = n_0 + \int_0^c \mu_0(n)G_0(n, n_0)dn$$

and the relevant measure of mean yearly harvest in the period before extinction is $\mathrm{E}[y(N)|N_0 = n_0] = Y(n_0)/T(n_0)$.

Proportional threshold harvesting

In practice the population size is usually not known exactly. Then the harvesting tactic must be based on the available population estimate, say $\hat{N}$. If the variance of this estimator is large one can intuitively see that threshold harvesting can perform very badly. Suppose that the estimate one year is much larger than the real population size. Then $\hat{N} - c$ is much larger than

$N - c$ so if we harvest the amount $\hat{N} - c$ the population will after harvesting end up far below c which may as a worst case lead to extinction.
Arguing through the binormal approximation of $(\hat{N}, N)$ this has lead to the more conservative approach of only harvesting a proportion q of the difference $\hat{N} - c$ when $\hat{N} > c$ and otherwise no harvesting, so that the yield is

$$y(\hat{N}) = \begin{cases} 0 & \text{for } \hat{N} < c \\ q(\hat{N} - c) & \text{for } \hat{N} > c \end{cases}$$

The infinitesimal mean and variance in the diffusion approximation for N is then

$$\mu(n) = \mu_0(n) - \mathrm{E}[y(\hat{N})|N = n]$$

$$\nu(n) = \nu_0(n) + \mathrm{var}[y(\hat{N})|N = n].$$

For a given distribution of the estimator $\hat{N}$, say $f(\hat{n}|n)$, the above mean and variance of the yield are given by

$$\mathrm{E}[y(\hat{N})|N = n] = q \int_c^\infty (\hat{n} - c) f(\hat{n}|n) d\hat{n}$$

and

$$\mathrm{var}[y(\hat{N})|N = n] = q^2 \int_c^\infty (\hat{n} - c)^2 f(\hat{n}|n) d\hat{n} - \{\mathrm{E}[y(\hat{N}|N = n)]\}^2.$$

Writing $\hat{n} = nZ$, an unbiased estimator with constant coefficient of variation requires that Z has distribution independent of n with mean 1. Then, if Z is approximately normally distributed the mean and variance can be expressed by the standard normal integral (exercise 7). Similar calculations can be made for other assumptions on how the variance of the estimator depends on the population size (exercise 8 and 9).
In practice, the harvesting statistics can now be computed numerically for a given tactic, that is, a given set of constants (c, q). Finally some numerical maximization procedure must be used to compute the value of (c, q) that maximizes for example the mean annual yield.

5.2 Population viability analysis

5.2.1 Introduction

Natural communities show large variation in species abundance, often approximately lognormally distributed between species. Hence, many species are naturally rare at a given site. Human disturbance of natural communities may therefore have serious effects on a number of species increasing the probabilities of extinction and reducing the expected time to extinction. The four major factors threatening species are habitat destruction, overexploitation (fishing, hunting, captured for live animal pet trade), introduced species and pollution.

The International Union for Conservation of Nature and Natural Resources (IUCN) develop criteria for classifying species into categories as Critically Endangered, Endangered and Vulnerable. IUCN uses a number of different criteria in this work, such as population size, trends in population size, fragmentation of habitats, decrease in geographic range and area species are occupying, that together determine the risk of extinction assigned to a species. Species at risk of extinction according to these criteria are listed in the so-called Red List for each country or region of the world which is updated periodically. In year 2000 23.7% of the mammals and 11.9 % of the bird species were considered threatened.

Since about 1980, IUCN has included the modern concept of Population Viability Analysis (PVA) in their criteria for classification of species. Population viability analysis uses stochastic modelling of population dynamics including concepts like demographic and environmental variances, defining viability through the probability of extinction within a certain time.

5.2.2 Population Prediction intervals

We have seen in chapter 3 how diffusion theory can be applied to determine the probability of extinction within a certain time in some models, while extensive stochastic simulations would have to be used in other more complex models. The probability of extinction will in any case depend on the population parameters and it will be very sensitive to the value of some of these parameters, in particular the population growth rate. Since time series observation always are only available over short time intervals (even 50 years is a short time series in the statistical sense, although biologists consider it to be an extremely long time series), parameter estimates will always be subject to rather large sampling errors. Because of the strong sensitivity of the probability of extinction to some of the parameters, this makes the sampling error in the estimate of the probability of extinction very large. Therefore, the task of estimating the viability of a species through the probability of extinction within a certain time is an almost impossible task. Attempts to find confidence limits for these probabilities has lead some scientist to claim that it is practically impossible to do a proper population viability analysis. Another problem with the above approach is that a confidence interval for a probability of extinction would in practice be hard to interpret for those who in practice are responsible for the management of populations. In particular, this will be difficult if the confidence interval is wide, which it almost always will be.

These problems may be overcome, however, by using the statistical concept of prediction intervals rather than confidence intervals. In statistics confidence intervals are used to draw inference about parameters, while prediction intervals are constructed for making statements about the realized values of stochastic variables that are not yet observed. The most common example in statistical text books is the solution to the problem of predicting an observation for some given value of a covariate in a simple linear regression model using students T-distribution. We have seen an even simpler example in section 1.5 where we constructed prediction intervals for future population sizes in the random walk model with normally distributed increments also using the T-distribution. This example (and the regression example) is particularly simple since they give exact solutions that do not depend on any nuisance parameters which in this case is the variance of the underlying normal distribution.

Generally, writing $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ for a set of observations with distribution depending on parameters $\boldsymbol{\theta}$, our aim is to construct two functions of $\mathbf{X}$ only, say $A(\mathbf{X})$ and $B(\mathbf{X})$ so that $P[A(\mathbf{X}) < Y < B(\mathbf{X})] = \mathbf{1} - \alpha$, where Y is the unobserved stochastic variable with distribution depending on $\boldsymbol{\theta}$ that are to be predicted with confidence $1 - \alpha$.

In more complex models, there is impossible to construct such functions A and B that makes the coverage $1 - \alpha$ constant, that is, not depending on any unknown parameters. Hence, one would have to rely on methods giving approximately correct coverage. In practice, one can try to make the coverage equal to a prescribed value for some value of the unknown parameters believed to be close to the real ones, that is, for the point estimate of the parameters. Computing such intervals will usually require rather extensive stochastic simulations.

5.2.3 Frequentistic population prediction interval

A rather simple way of constructing prediction intervals is to use parametric bootstrap replicates of the unknown parameters. Having estimated the parameters in the model we can simulate parametric bootstrap data sets using the estimated values. For each such set of data a bootstrap-replicate of the parameters are found by performing exactly the same estimation procedure as when the estimation was originally done from the data. Then, the process can be simulated into the future using the bootstrap-replicates, starting from the last observation, one simulation of the process as far as we like for each bootstrap replicate of the parameters. If this is done for example 10000 times we end up with 10000 time series into the future. For each future year we can then use the quantiles of the empirical distribution of the simulated population size each year. The lower 5% quantile would the be series number 500 when they are ordered according to population size. Fig.5.3 shows an example of such intervals. This population is far below its carrying capacity, so a model without density regulation is used. The demographic variance has been estimated from a large number of individual records of number of offspring and survival and is considered as known in this analysis.

There are different ways of performing the above simulations. In this example the simulations has been done in such a way that the last observation of each simulation equals the last observation in the data.

The above method will not give prediction intervals with exactly correct coverage. So, one can go further with these computation and do a large number of simulations to check what the coverage actually is for given parameter values (usually the point estimates). Finally the bounds can be adjusted to give more correct coverage. Such a technique has been applied to produce the graphs in Fig.5.3.

5.2.4 Bayesian population prediction intervals

It is often conceptually and practically simpler to use the Bayesian approach to derive prediction intervals. Although this approach do not lead to correct coverage for a given set of parameters, which is a major point of the frequentistic approach, there are advantages that the frequentistic approach can not compete with. Apart from the simplicity, the Bayesian approach enables us to use prior information about parameters obtained from analysis of the same or closely related species. If one do not want to use such information, or it is not available, one usually use a so-called non-informative prior distribution of the parameters. Another great advantage is that the appoach can easily be generalized to include sampling errors in the observations.

Let $\pi(\boldsymbol{\theta})$ denote the prior distribution of the parameters and let $\mathbf{X}$ be the data. Our model then describes the distribution $f(\mathbf{x}|\boldsymbol{\theta})$ of $\mathbf{X}$. The Bayesian approach is simply to evaluate the posterior distribution of $\boldsymbol{\theta}$ which is its distribution conditioned on the observed value $\mathbf{x}$ of $\mathbf{X}$, which is

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{\mathbf{f}(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int \mathbf{f}(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\mathbf{d}\boldsymbol{\theta}}.$$

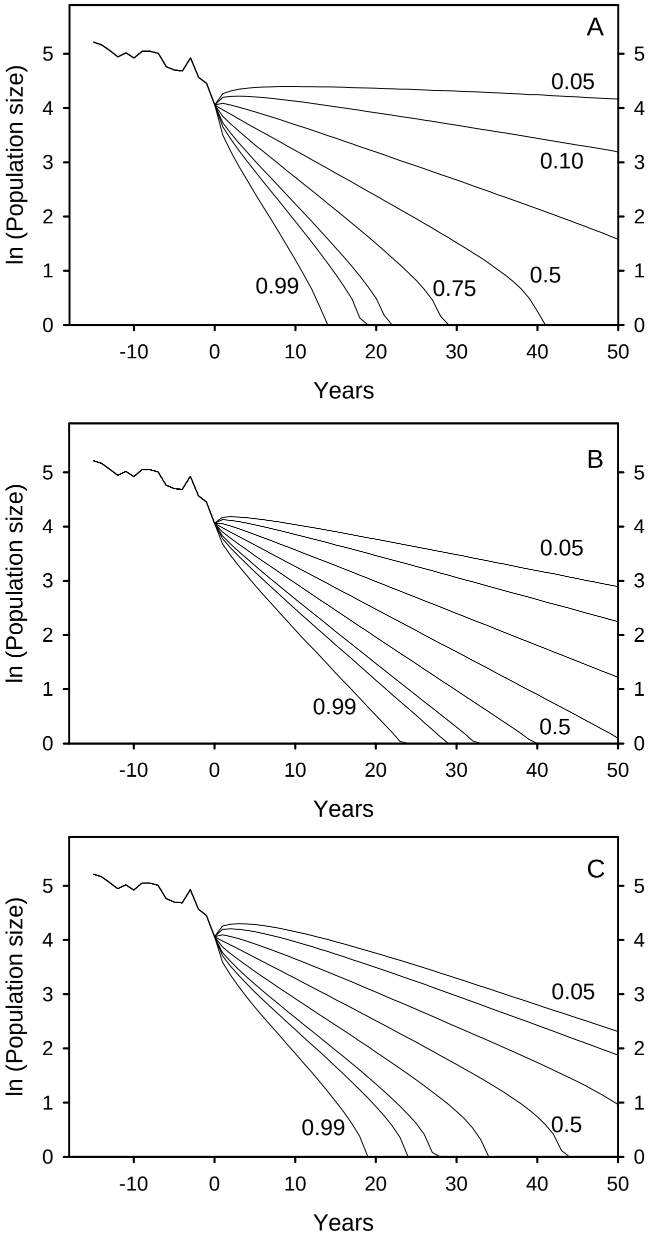

Figure 5.3: Population prediction intervals for the Barn Swallow at Kraghede, (A) accounting for demographic and environmental stochasticity and uncertainty in the parameters, (B) assuming no uncertainty in the parameters, and (C) assuming no demographic stochasticity so that all stochasticity is estimated as environmental stochasticity.

If we now are able to simulate observations of $\boldsymbol{\theta}$ from the posterior distribution we can simply find the posterior distribution of future population sizes by simulating the process for each simulation of the parameters as we did in the frequentistic approach. Finally, prediction intervals comes out using the ordered population sizes as before.

If there are sampling errors, the population sizes $\mathbf{X}$ cannot be observed, but only some estimates, say $\mathbf{Y}$ with distribution say $h(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$. Notice that $\boldsymbol{\theta}$ now may have components expressing distribution of sampling errors. In this case we need the posterior distribution of $(\mathbf{X}, \boldsymbol{\theta})$ which is

$$q(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) = \frac{\mathbf{h}(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})\mathbf{f}(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int \mathbf{h}(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})\mathbf{f}(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\mathbf{dxd}\boldsymbol{\theta}}.$$

5.2.5 A simple example of Bayesian population prediction intervals

As an illustration we consider the linear model on the log scale defined by assuming that X_{t+1} conditioned on X_t is normally distributed with

$$\mathrm{E}(X_{t+1}|X_t = x) = \alpha + \beta x$$

and constant variance $\mathrm{var}(X_{t+1}|X_t = x) = \sigma^2$. We also assume that the log population sizes at $n + 1$ subsequent years $X_0, X_1, \ldots, X_n$ are known. If we are unsure about the process actually starting at the stable age distribution we should condition on the first observation x_0. The joint distribution of the observations $\mathbf{x} = (\mathbf{X_1}, \mathbf{X_2} \ldots, \mathbf{X_n})$ given the parameter values and the observed value x_0 of X_0 are then

$$f(\mathbf{x}|\alpha, \beta, \tau^{\mathbf{2}}, \mathbf{x_0}) = \boldsymbol{\Pi}_{\mathbf{i=1}}^{\mathbf{n}} \frac{\tau}{\sqrt{\mathbf{2}\pi}} \mathbf{e}^{-(\mathbf{x_i}-\alpha-\beta\mathbf{x_{i-1}})^{\mathbf{2}}\tau^{\mathbf{2}}/\mathbf{2}}$$

where we have introduced the parameter $\tau^2 = 1/\sigma^2$ often called the precision in Bayesian theory.

If nothing is known about the parameters apriori, one often use a so-called non-informative prior distribution. Accordingly we assume that α and β

are distributed uniformly on the real axes, which must be considered as a limiting case of the of a uniform distribution ranging from $-c$ to c as c approaches infinity. For the precision one often use the distribution formally proportional to $1/\tau^2$ which can be considered as the limiting case of the gamma distribution when the shape parameter as well as the scale parameter approaches zero. Multiplying with these distribution, that is, multiplying by $1/\tau^2$, we find the joint distribution of $(\mathbf{x}, \alpha, \beta, \boldsymbol{\tau^2})$ up to a constant factor that can be determined by integration. Further, the joint distribution of (α, β, τ^2) conditioned on the observations is also proportional to the same expression. The goal of this Bayesian approach is to simulate observations from this posterior distribution of the unknown parameters conditioned on the observations. One way of doing this is to use the so-called Gibbs sampler, for which the main idea is the following: if the parameters are generated from the wanted distribution we can take out a single component and replace it by a new simulation of this from the conditional distribution of the component given the others. Repeating this procedure subsequently for all parameters (components) we obtain a new simulation of the whole set of parameters.

One problem with this approach is that we do not know how to start the simulations from the correct distribution. However, under rather general conditions the procedure will converge to the correct distribution starting from some reasonable parameter values. The process then produces a Markov chain of parameter values with stationary distribution being the one we want to simulate from.

In the simple example we consider here we see that the distribution of α conditioned on the observations and the other parameters has the form of a normal distribution, being the exponential function of a second degree polynomial in α. The same is the case for β. The distribution of τ^2, on the other hand, is a gamma distribution.

Using a technique given in exercise 10 we find (exercise 10) that the mean and variance of the appropriate conditional distribution of α is

$$\mathrm{E}\alpha = \frac{1}{n}\sum_{i=1}^{n}(x_i - \beta x_{i-1})$$

and

$$\mathrm{var}(\alpha) = 1/(n\tau^2).$$

Using the same approach we also find (exercise 10) that β is normal with

$$\mathrm{E}\beta = \sum_{i=1}^{n} x_{i-1}(x_i - \alpha)(\sum_{i=1}^{n} x_{i-1}^2)^{-1}$$

and

$$\mathrm{var}(\beta) = (\tau^2 \sum_{i=1}^{n} x_{i-1}^2)^{-1}.$$

Finally, the shape parameter in the gamma distribution of τ^2 is simply $n/2$ while the scale parameter is $\sum_{i=1}^{n}(x_i - \alpha - \beta x_{i-1})^2/2$ (exercise 11).
Fig.5.5 shows the Markov chains obtained for the three unknown parameters using the Gibbs sampler as described above. Finally we find the prediction intervals by simulating the time series starting from the last observation and using a sequence of simulated values from the Markov chain as described above. Fig.5.5 shows predictions from the above model using $X = 0$ $(n = 1)$ as extinction barrier.
MCMC-methods can be applied to very complex models. However, using these in practice require a course in modern statistical methods. Statistical software is available for performing the simulations for a number of different models.

5.3 Genetic drift

5.3.1 A two-dimensional diffusion model

The theory on demographic and environmental stochasticity, diffusion, and age-structured population dynamics may be used to analyze genetic drift

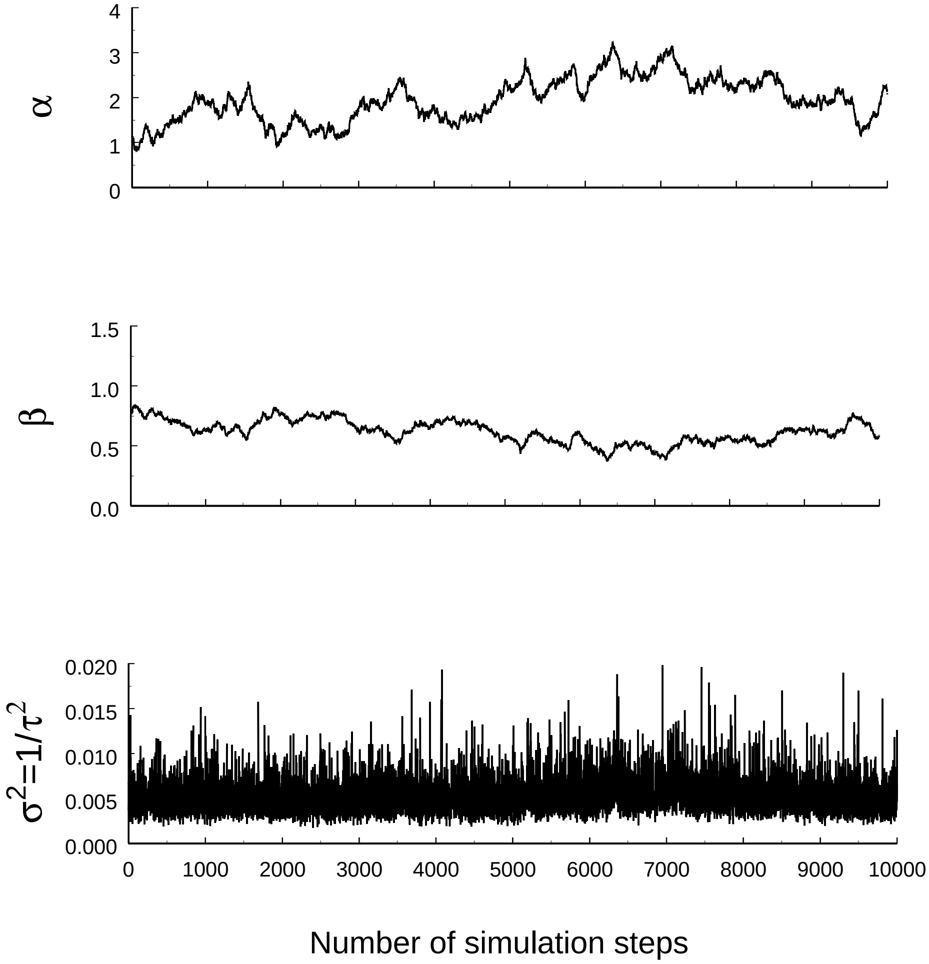

Figure 5.4: Markov Chain Monte Carlo simulations of the posterior distribution of α, β and τ^2. The data is a time series of length 21 ($n = 20$)generated from the model with $\alpha = 1$, $\beta = 0.8$ and $\tau^2 = 100$.

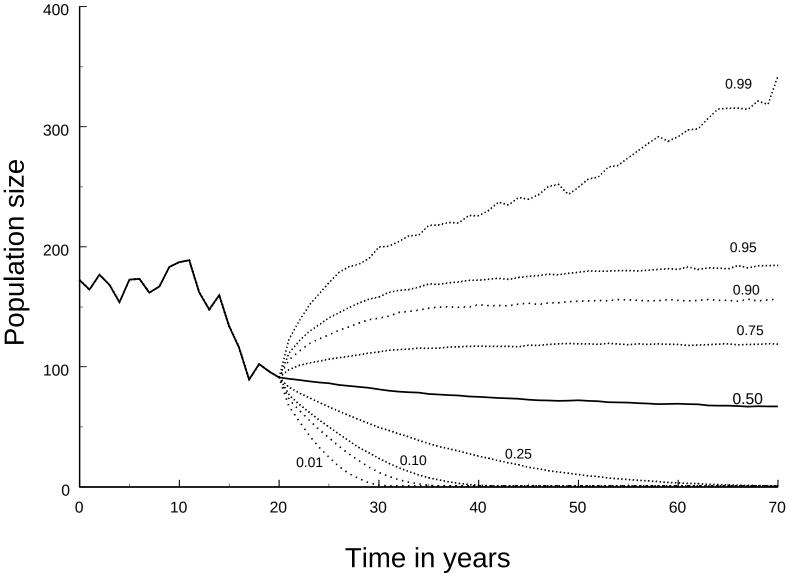

Figure 5.5: Simulated time series data and Bayesian prediction intervals. The data is a time series of length 21 ($n = 20$) generated from the model with $\alpha = 1$, $\beta = 0.8$ and $\tau^2 = 100$.

of neutral genes, that is, to study the temporal changes in gene frequencies in fluctuating populations. This analysis, however, require use of two-dimensional diffusions.

Diffusions in two dimensions are defined by the infinitesimal mean and variance of each component, defined as in the one-dimensional case, but requires also an infinitesimal covariance. For a diffusion (X_t, Y_t) we write $\mu_X(x)$ and $\nu_X(x)$ for the infinitesimal mean and variance of X_t and similarly for Y_t. The infinitesimal covariance is defined as

$$c(x, y) = \mathrm{E}(dXdY)/dt.$$

As in the one dimensional case a transformation into another two-dimensional variable will also lead to a diffusion and the infinitesimal moments of this can be derived by general formulas. Here, we shall only consider the particular transformation

$$N = X + Y$$

$$P = X/(X + Y)$$

so we restrict the presentation to this transformation and are in particular interested in the process P. The infinitesimal change in P is given by

$$dP = \frac{X(1 + dX/X)}{N(1 + dN/N} - \frac{X}{N} = \frac{X}{N}(1 + dX/X)(1 - dN/N + (dN/N)^2 - \ldots - 1)$$

Since diffusions are determined by the first two moments only we can ignore higher order moments giving, inserting $X = NP$ and $Y = N(1 - p)$

$$dP = dX\frac{(1 - P)}{NP} - dY\frac{1}{N} + dX^2\frac{(1 - 1/P)}{N^2} + dY^2\frac{1}{N^2} - dXdY\frac{(2 - 1/P)}{N^2}$$

and

$$\mathrm{E}dP^2 = P^2(\frac{dX}{X} - \frac{dN}{N})^2 = \frac{1}{N^2}[(1 - P)dX - PdY]^2.$$

We shall in particular consider X and Y as two sub-populations with perfectly correlated environmental noise (since they are in the same environment), the same demographic variance and the same exponential growth. Then $\mu_X(x) = rx$, $\mu_Y(y) = ry$, $\nu_x(x) = \sigma_d^2 x + \sigma_e^2 x^2$, $\nu_Y(y) = \sigma_d^2 y + \sigma_e^2 y^2$ and $c(x,y) = \sigma_e^2 xy$. This leads to (exercise 12)

$$\mu_N(n,p)) = rn$$

and

$$\nu_N(n,p)) = \sigma_d^2 n + \sigma_e^2 n^2.$$

Using the above expression for dP and dP^2 we find after some algebra (exercise 13) that

$$\mu_P(n,p) = 0$$

and (exercise 14)

$$\nu_P(n,p) = \frac{\sigma_d^2 p(1-p)}{n}.$$

One can also show (exercise 15) that the infinitesimal covariance for the process (N,P) is zero.

5.3.2 Effective population size

Genetic drift is the change in gene frequency of neutral genes due to randomness in the process that determines how genes are transmitted through generations. The term *neutral genes* is used for genes that are equivalent in their effect on vital rates and the stochasticity in vital rates. So, for two neutral genes a and A there is no natural selection of any of them and the change in their frequency in a population is due to purely random effects only.

Genetic drift of a gene will depend on the frequency of the gene in the population. In order to compare drift of genes with different frequencies, however, the concept of *effective population size* is used, which, although it is independent of gene frequency, contains all information about the genetic drift of any neutral gene in the population. Effective population size is defined by a comparison to a simple one sex model with non-overlapping generations and constant population size where the genes of the next generation are selected by simple binomial random sampling.

Consider a diploid population of N individuals and two neutral genes a and A at a given locus. If there are X genes of type A in the population we say that the gene frequency of A is $P = X/(2N)$ since there are $2N$ genes altogether at this locus. Now, let the next generation be formed by randomly selecting $2N$ genes to form N individuals from an infinitely large gene pool with a fraction P of genes of type A which means that each individual has contributed equally to the gene pool. The number of genes A in the next generation, say $X + \Delta X$ is then binomially distributed with parameters $(2N, P)$ so that

$$\mathrm{var}(\Delta P|P = p) = \mathrm{var}(\frac{X + \Delta X}{2N}|P = p) = \frac{p(1-p)}{2N}.$$

This model is now used as a yardstick for comparison with other more complex models. Quite generally, the variance of ΔP over a time span of one generation will be proportional to $p(1-p)$, say

$$\mathrm{var}(\Delta P|P = p) = cp(1-p) = \frac{p(1-p)}{2N_e}$$

which is then the definition of the effective population size N_e. Notice that N_e determines the drift (the variance in ΔP) for any neutral gene regardless its gene frequency.

To determine the effective population size of populations is therefore extremely important in analyzing changes in gene frequency, reduction of heterozygosity and extinction of genes, all being important components of conservation biology. Quite often N_e may be much smaller than the population size N, which means that the drift will be much larger than the drift in the yardstick model for equal population sizes. We shall now look at some simple examples showing how different properties of vital rates in populations affect the genetic drift.

5.3.3 Haploid model with age-structure

We consider first a haploid organism with two neutral genes a and A. In a haploid model individuals has only one gene at a given locus and transmit this gene to their offspring. The reference genetic drift is accordingly $p(1-p)/n$. Since individuals now just produce copies of themselves the two genotypes can be considered as two separate populations, X and Y individuals of type A and a rerspectively. If there is no density-regulation these populations do not interact, although they are dependent through the fact that they are in the same environment so that the infinitesimal covariance is $\sigma_e^2 xy$. We have seen that even a complex age-structured model may be accurately approximated by a diffusion when there is no density regulation. This diffusion process will be exactly the two-dimensional process we have considered earlier in this chapter. Hence, we see immediately that the genetic drift during one year is determined by

$$\mathrm{var}(\Delta P|P=p,N=n)=\sigma_d^2 p(1-p)/n.$$

The variance of ΔP during one generation T is then approximately $\sigma_d^2 Tp(1-p)/n = p(1-p)/N_e$ so that the effective population size for a population of size N is simply

$$N_e=\frac{N}{\sigma_d^2 T}\ .$$

Although this result may seem very simple, notice that σ_d^2 and the generation time T are more complex functions of the properties of the vital rates in the age-structured model (see chapter 4).

This simple result is interesting in two ways. First, we see that the environmental stochasticty has absolutely no effect on the genetic drift for a given population size. Over longer time intervals, however, the drift is affected by the environmental stochasticity through its effect on the process N. Secondly, we see that the drift during one year is determined by the demographic variance alone. So this is an important example of how the two forms of stochasticity, environmental and demographic, have quite different effects on populations and must be considered separately.

5.3.4 Diploid two-sex model with overlapping generations

When analyzing more complex models we shall do the calculations assuming that the gene A is rare so that P is small and $P(1-P) \approx P$. The advantage of this approach is that an individual with this rare gene always (approximately) is a heterozygote that mate with individuals that do not have this gene, a fact that simplifies the analysis a lot. It is easy to show that the effective population size derived from a small gene frequency actually is valid for any gene frequency (exercise 20). We consider a two-sex model using m and f as subscripts referring to males and females respectively.

We now assume that all individuals of the same sex has the same properties writing b_f and b_m for the mean number of offspring of a single female and male that survive to enter the adult population, s_f and s_m for the adult survival rates, and q for the probability that a newborn individual is a female. So, this is a population with individuals of different ages, but since all individuals of the same sex has the same properties we just call this a population with overlapping generations rather than an age-structured population.

Let $\mathbf{Z}_t = (X_t, Y_t)'$ be the number of adult females and males respectively in the subpopulation with the rare gene, where $'$ denotes matrix transposition as before. Then, if there is no density regulation the expected dynamics are given by

$$\mathrm{E}\mathbf{Z}_{t+1} = \mathbf{M}\mathrm{E}\mathbf{Z}_t$$

where the projection matrix is (exercise 16)

$$\mathbf{M} = \begin{bmatrix} \frac{1}{2}qb_f + s_f & \frac{1}{2}qb_m \\ \frac{1}{2}(1-q)b_f & \frac{1}{2}(1-q)b_m + s_m \end{bmatrix}$$

The factors $\frac{1}{2}$ occurs due to the fact that a mother or father with the given gene transmit the gene to their offspring with probability $\frac{1}{2}$ because the gene is rare and all individuals bearing it therefore are heterozygotes that mate with an individual not bearing the gene.

Let $Z = X + Y$ be the number of individuals of type (aA), both sexes included, and write $W = 2N - Z$ for the number of homozygotes with genotype (aa). The gene frequency if A is $P = Z/(Z + W)$. As before, a common environment acting on the processes Z and W will have no effect on the process P for a given N, and neutrality ensures also here that $\mu_P(p, n) = 0$. Using the previously derived result (with slightly different symbols) $\mathrm{E}dP^2 = \frac{1}{(2N)^2}[(1-P)dZ - PdW]^2$ and remembering that $p \approx 0$ we find

$$\nu_P(p, n) \approx \sigma_{dg}^2 p/(2N)$$

where σ_{dg}^2 is the demographic variance of the process Z, which is affected by individual variation in survival and reproduction as well as the binomial sampling of genes from parents to offspring. So, for a small value of p we see that

$$N_e \approx \frac{N}{\sigma_{dg}^2 T}$$

which must hold in general since the effective population size is independent of gene frequency. So, the effective population size is determined as soon as we know the demographic variance σ_{dg}^2 of the process Z and the mean generation time.

The first step is to analyse the expected dynamics determined by the matrix M. Let $\mathbf{u} = (u_1, u_2)^T$ and $\mathbf{v} = (v_1, v_2)$ denote the left and right eigenvectors of $\mathbf{M}$ associated with the dominant eigenvalue λ defined by $\mathbf{Mu} = \lambda\mathbf{u}$ and $\mathbf{vM} = \lambda\mathbf{v}$. If $\mathbf{u}$ is scaled so that the $u_1+u_2 = 1$, it is the stable sex distribution corresponding to the deterministic model. Since the males and females must have the same expected reproduction (each individual is produced by one male and one female), we have $u_1 b_f = u_2 b_m$ giving $u_1 = b_m/(b_f + b_m)$ and $u_2 = b_f/(b_f + b_m)$. From the equation $\mathbf{Mu} = \lambda\mathbf{u}$ we find (exercise 17) that the dominant eigenvalue λ is $qb_f + s_f$ as well as $(1-q)b_m + s_m$. This relates the mean birth rates for males and females through the relation $b_m = (qb_f + s_f - s_m)/(1-q)$. The right eigenvector $\mathbf{v}$, scaled so that the sum of its components is one, turns out to be $(1-q, q)$ (exercise 18). The sensitivities of this model are $\partial\lambda/\partial M_{ij} = v_i u_j/(\mathbf{vu})$.

In order to find the demographic variance we must consider the stochastic dynamics as we did in chapter 4. Writing G for the stochastic matrix operating a given year so that $\mathbf{Z}_{t+1} = \mathbf{G}\mathbf{Z}_t$, one can show that the expectation of $\mathbf{G}$ is $\mathbf{M}$ and that the non-zero variances and covariances $\text{cov}(G_{ij}, G_{kl})$ of its elements can be expressed by some constants $a_{ij,kl}$ defined by $a_{i1,j1} = X_t \text{cov}(G_{i1}, G_{j1})$ and $a_{i2,j2} = Y_t \text{cov}(G_{i2}, G_{j2})$. These coefficients are

$$a_{11,11} = \frac{1}{4} b_f q(2-q) + \frac{1}{4} q^2 \sigma_f^2 + s_f(1-s_f) + qc_f$$

$$a_{21,21} = \frac{1}{4} b_f (1-q)(1+q) + \frac{1}{4}(1-q)^2 \sigma_f^2$$

$$a_{11,21} = a_{21,11} = -\frac{1}{4} b_f q(1-q) + \frac{1}{4} q(1-q)\sigma_f^2 + \frac{1}{2}(1-q)c_f$$

$$a_{12,12} = \frac{1}{4} b_m q(2-q) + \frac{1}{4} q^2 \sigma_m^2$$

$$a_{22,22} = \frac{1}{4} b_m (1-q)(1+q) + \frac{1}{4}(1-q)^2 \sigma_m^2 + s_m(1-s_m) + (1-q)c_m$$

$$a_{12,22} = a_{22,12} = -\frac{1}{4} b_m q(1-q) + \frac{1}{4} q(1-q)\sigma_m^2 + \frac{1}{2} qc_m$$

where σ_f^2 is the variance of the number of offspring produced by females, c_f is the covariance between a female's number of offspring and the indicator of her survival, and similar for males.

We show how to derive the first of these coefficient and leave the others as an exercise (exercise 19). In order to find $\text{cov}(G_{11}, G_{11}) = \text{var}(G_{11})$ we first consider a single female of type (a, A) producing B_f offspring surviving to the next sensus and having survival indicator I_f which is one if she survives herself and otherwise zero. First conditioning on (B_f, I_f) her B_f offspring fall independently into three different categories. They have genotype (A, A) with probability one half since she mate with a male of type (A, A), and they are female of type (a, A) with probability $q/2$ and males of type (a, A) with probability $(1-q)/2$. Writing W for her contribution of females of type (a, A), we consequently have

$$\mathrm{E}(W|B_f, I_f) = \frac{1}{2}qB_f + I_f$$

and

$$\mathrm{var}(W|B_f, I_F) = B_f\frac{1}{2}q(1-\frac{1}{2}q) = B_f\frac{1}{4}q(2-q).$$

Using the general formulas for the unconditional mean and variance of a stochastic variable we then find

$$\mathrm{E}W = \mathrm{E}\mathrm{E}(W|B_f, I_f) = \frac{1}{2}qb_f + s_f$$

and

$$\mathrm{var}(W) = \mathrm{E}\mathrm{var}(W|B_f, I_f) + \mathrm{var}\mathrm{E}(W|B_f, I_f) = \frac{1}{4}q^2\sigma_f^2 + s_f(1-s_f) + qc_f + \frac{1}{4}b_f q(2-q).$$

The element G_{11} is the mean value of X_f such independent contributions W so that $\mathrm{var}(G_{11}) = X_f^{-1}\mathrm{var}(W)$ leading to the above expression for $a_{11,11}$. Using the approximation for the demographic variance derived in chapter 4 we find that the demographic variance of the process Z is approximately

$$\sigma_{dg}^2 = \lambda^{-2}\sum_{ijk} u_j^{-1}\frac{\partial\lambda}{\partial M_{ij}}\frac{\partial\lambda}{\partial M_{kj}}a_{ij,kj} = (\lambda\mathbf{vu})^{-2}\sum_{ijk} u_j v_i v_k a_{ij,kj} \qquad (5.1)$$

where all three indices ijk run over 1 and 2. The mean generation time for this model is $T = (T_f + T_m)/2 = [\lambda/(\lambda - s_f) + \lambda/(\lambda - s_m)]/2 = \lambda/b_f + \lambda/b_f$. Although these results may seem complicated, it is straightforward to write a simple computer program that can be used to investigate how each parameter determining the dynamics of $\mathbf{Z}$ also affects the effective population size and hence the genetic drift of this population.

The same kind of analysis can be performed on a general age-structured model although more algebra is required to find the results.

5.4 Exercises

1. Find the optimal harvesting tactic for then discrete deterministic theta-logistic model $\Delta N = r_1 N[1 - (N^\theta - 1)/(K^\theta - 1]$, and investigate how the type of density-regulation (defined by θ) affects the solution.
2. Consider a population with constant environmental and demographic variance which is far below its carrying capacity so that $\mu_0(n) = \bar{r}n$. Find the lower population size n^* that gives a positive stochastic growth rate when this population is harvested at a constant rate y. Make a graph of n^* as function of the harvest y when $\sigma_d^2 = 1$, $\sigma_e^2 = 0.01$ and $\bar{r} = 0.02$.
3. For the logistic deterministic model $\Delta N = rN(1 - N/K)$ show that the maximum harvest $Y = rK/4$ is obtained for harvesting rate $r/2$ (that is, harvesting reduces the growth rate ΔN by $Nr/2$).
4. Solve the deterministic problem in exercise 3 for the Gompertz model $\Delta N = r_1 N(1 - \ln N/\ln K)$.
5. Show that the maximum mean harvest in the diffusion approximation for the Gompertz model with no demographic stochasticity and environmental variance σ_e^2 is $\mathrm{E}Y_{max} = K^{1-\sigma_e^2/(4r_1)} r_1 e^{-1}/\ln K$.
6. Show that the Green function for threshold harvesting with threshold c equals the Green function for the process with no harvesting for $n \leq c$ and zero for $n > c$.
7. Let the estimate of the population size be of the form nZ, there Z is normally distributed with mean 1 and variance σ^2. Find the mean and variance of the yield for proportional threshold harvesting expressed by the standard normal integral (cumulative standard normal distribution).
8. Solve the problem in exercise 7 when $\mathrm{var}(Z) = \sigma^2/n$ so that the variance of $\hat{N}$ is proportional to n.
9. Solve the problem in exercise 7 when $\mathrm{var}(Z) = \sigma^2 h(n)$.
10. Write $f(z; \mu, \sigma^2)$ for the normal density with mean μ and variance σ^2. Show that $\partial \ln f/\partial z|_{z=0} = \mu/\sigma^2$ and $\partial^2 \ln f/\partial z^2 = -1/\sigma^2$. Use these results to find the mean and variance of α and β in the Bayesian approach to the linear normal model when conditioned on the observations and the other

parameters.

11. In the same example find the parameters of the gamma distribution of τ^2 when conditioned on the observation and the other parameters.

12. Using the infinitesimal moments of the two-dimensional diffusion (X, Y) defined in the text show that $N = X+Y$ has infinitesimal mean $\mu_N(n, p) = rn$ and infinitesimal variance $\nu_N(n, p) = \sigma_d^2 n + \sigma_e^2 n^2$.

13. For the model used in exercise 12, use the expression for dP given in the text to show that P has zero infinitesimal mean.

14. For the model used in exercise 12, use the expression for dP^2 given in the text to show that the infinitesimal variance of P is $\nu_p(n, p) = \sigma_d^2 p(1-p)/n$.

15. For the model used in exercise 12, use the expression for dP given in the text to show that the infinitesimal covariance for the process (N, P) is zero.

16. Write out the details showing that the two-sex diploid model with overlapping generations has the projection matrix M for the expected dynamics given in the text.

17. Show that the dominant eigenvalue of M is $\lambda = qb_f + s_f = (1-q)b_m + s_m$.

18. Show that the scaled right eigenvector $\mathbf{v}$ of M, scaled so that the sum of its components is one, is $(1-q, q)$.

19. Derive all expectations, variances and covariances of the elements of the stochastic matrix G given by the coefficients $a_{ij,kl}$ defined in the text.

20. Let the neutral genes at a given locus be theoretically subdivided into a large number m of differently labelled alleles $A_1, A_2, \ldots, A_m$ with the same frequency $P_i = 1/m, i = 1, 2 \ldots m$. Define a new allele B consisting of k of the $A_i's$ which then have frequency $Q = k/m$. We have shown in the text that $\text{var}(\Delta P_i) \approx P_i \sigma_{dg}^2/(2N)$. Use the fact that $\text{var}(\sum_{i=1}^m \Delta P_i) = 0$ (why is it so?) to find $\text{cov}(\Delta P_i, \Delta P_j)$ and finally use this result to show that $\text{var}(\Delta Q) \approx \sigma_{dg}^2 Q(1-Q)/(2N)$.

Chapter 6

Spatial models

6.1 Introduction

Until now we have have considered individuals to be members of a population without considering their actual position. Such models can be realistic in many cases when individuals are located within a well defined area, for example an island or some other suitable habitat that is small enough for each individual to have the same relation to any other individual of that population. However, we have also assumed that the populations are closed, that is, there is no interaction with other populations through migration. Although some populations may be rather isolated, this is generally a very strong and often unrealistic assumption, and in many cases completely wrong. Many quite different species, for example fish and insects, migrate a lot, and may not have a well defined home range. Modeling such populations is a difficult task for several reasons. Rather than dealing with a single number of individuals, we will need to keep track of the individuals' position in space at different times. Even in the simplest case of some few interacting populations, this is mathematically a big step away from the simple one-dimensional single population model. The challenge of obtaining relevant statistical information about populations spread out in space is also enormous. Mark recapture experiments is an important component of this research, enabling observations of the same individual at different times giving partial information about its migration.

One question of great interest is the synchrony in population fluctuations. Often one find that population density at rather close distances are spatially correlated, that is, if one density is high at a given point of time, the other density is also expected to be large. This synchrony usually decreases with increasing distance, but synchrony in population fluctuations can also be observed over very large distances. For insects, the synchrony usually disappears already at rather small distances such as some few kilometers, while fish in the ocean may show population synchrony even at thousands of kilometers. Population synchrony can be generated in many different ways. The simplest explanation is that the habitat quality for a species may be distributed with some patchiness in space and that population densities accordingly are distributed spatially according to quality. But there are also many other possible explanations. The stochastic noise acting at two locations may be positively correlated due to spatial correlation in the physical and biological environments collected in what we consider as noise in the models. Further, migration between locations, especially between locations at close distances, must have some impact on the synchrony.

We have seen that most populations are density-regulated. This phenomena is also more complex when it comes to spatially distributed populations as the competition between two individuals must be related to their physical distance. We shall also see that density-regulation also affects population synchrony. If the density-regulation is small, migration may have a large effect on population synchrony even at distances much larger than typical migration distances, while strong local density-regulation seems to break down the spatial correlations in density.

In 1953, the statistician Moran published a paper explaining that spatial correlation in noise alone could synchonize populations. Actually, if two local populations can be described as simple linear time series with the equal parameters determining the density-regulation, then the correlation between the densities (or log densities) will be exactly the same as the correlation between the noise terms. We have seen that such models often can be a realistic description if we work on the log scale and the population size is large enough to ignore demographic stochasticity (see chapter II, theta-logistic model with $\theta = 0$, and exercise 5). The Moran effect may be considered as a kind of null model in spatial population ecology as much attention has been given to estimation and explanation of deviations from the correlations explained by Moran.

Spatial analysis of ecological data was initiated by investigations of statistical patterns in individual counts, such as recordings of number of individuals in quadrates. Most well known are the so-called Taylor's scaling laws, expressing that the relation between the mean and variance in counts are linear on the log scale. Empirically this has been shown to be a good approximation for different ways of changing the mean, such as considering repeated counts through time (years) or studying the effect of increasing sampling size (quadrate size). The law has also been shown to hold approximately for between species comparisons for different taxonomic groups of species. However, theoretical models indicate that these laws are mainly rather rough approximations to the true relationships which has slopes that may vary with the mean, but usually stays between 1 and 2 in a double logarithmic plot.

In chapter III we have studied extinctions and quasi-extinctions of single populations. The concept of extinction, however, is also more complicated for populations spread over large areas. It will now be more appropriate to analyze the probability of local extinction or quasi-extinction since subpopulations that apparently have gone extinct may be recolonized through migration provided that the total population spread out in space has not gone extinct.

6.2 The meta-population approach

Many populations may naturally be considered as a set of discrete sub-populations. These local populations may be dealt with as we until now have dealt with single populations. However, there are two important things that may connect the dynamics of the populations. Usually, populations are not completely isolated and there may be some migration of individuals between them. Secondly, the stochastic terms in the model for each sub-population may not be independent. Populations located nearby to each other are likely to have highly correlated noise since the physical and biological stochastic effects that act on them are likely to be similar. These two effects make the mathematical analysis of such a system, called a meta-population, rather intractable.

However, the meta-population approach has the great advantage that it is conceptually simple, and can easily be studied by extensive stochastic simulations. Although this is a very useful approach with many advantages, it has the general disadvantage of simulation models that it may be hard to find generally valid conclusions due to the large number of parameters that can possibly be varied in such simulations. Nevertheless, simulation studies of properties of meta-population models has been proved to be a very useful tool in the exploration of spatial population dynamics.

Here we shall only analyze the stochastic version of the classical meta-population model that makes some simplifying assumptions. We have seen in chapter 3.11.2 that the time to extinction of a population fluctuating around an equilibrium often can be approximated by the exponential distribution, which is equivalent to assuming the the probability that the population will go extinct during the next small time interval Δt is constant equal to, say $e\Delta t$. We assume that the extinction rate e is the same for all colonized sub-populations, or islands. Further, we make the (strong) assumption that extinctions at different locations are independent events. This would be correct if the noise terms were independent and there was no connections through migration.

Provided that all other islands are empty we assume that a colonized island colonizes one of the others the next small time interval Δt with constant probability $c\Delta t$. If the meta-population then consists of n population altogether and x of these are colonized at time t, the probability that a colonization occur is then $cx(1 - x/n)\Delta t$. This is because there are x islands that may colonize the others, but only a fraction $(1 - x/n)$ of islands that actually can be colonized. Accordingly a transition from x to $x + 1$ in the next infinitesimal interval dt occurs with probability $\alpha_x dt = cx(1 - x/n)dt$, from x to $x - 1$ with probability $\beta_x dt = exdt$ while the number of colonized islands remains unchanged with probability $1 - [\alpha_x + \beta_x]dt$. Writing X_t for the stochastic number of occupied islands at time t we see that X_t is a birth- and death-process with the above transition probabilities. Writing $p_t(x) = P(X_t = x|X_0 = x_0)$ we have accordingly

$$p_{t+dt}(x) = p_t(x)[1 - (\alpha_x + \beta_x)dt] + p_t(x + 1)\beta_{x+1}dt + p_t(x - 1)\alpha_{x-1}dt.$$

Replacing dt by small discrete steps Δt this equation can be used to calculate $p_t(x)$ recursively. In particular it is interesting to find the probability $p_t(0)$ that the meta-population has gone extinct within time t.

Let S_x denote the expected first passage time from state x to $x - 1$. The expected time the process stays in state x before leaving x for the first time is $1/(\alpha_x + \beta_x)$. When a transition occurs, the next state is $x + 1$ and $x - 1$ with probability $\alpha_x/(\alpha_x + \beta_x)$ and $\beta_x/(\alpha_x + \beta_x)$, respectively, giving the recurrence relation

$$S_x = \frac{1}{\alpha_x + \beta_x} + \frac{\alpha_x}{\alpha_x + \beta_x}(S_{x+1} + S_x) + \frac{\beta_x}{\alpha_x + \beta_x} \cdot 0$$

which, after solving for S_x gives

$$S_x = \frac{1}{\beta_x} + \frac{\alpha_x}{\beta_x} S_{x+1}.$$

All S_x can then be calculated subsequently, downwards from n to 1, with initial condition $S_n = 1/\beta_n$. Finally, the expected time to extinction of the meta-population with initial state x_0 is $T = \sum_{x=1}^{x_0} S_x$. If the expected time to extinction is large, it will be approximately independent of x_0 and the time to extinction will be approximately exponentially distributed with parameter $1/T$ in accordance with our derivation in 3.11.1. Hence, the probability that the whole meta-population goes extinct before time t when it initially is at an intermediate state is approximately $\exp(-t/T)$.

Even if extinction is certain for this model for a finite n the process may fluctuate a long time before extinction occur and hence be what we have called a quasi-stationary process. As the expected change in X_t is negative when $X_t > x^* = (1 - e/c)n$ and positive if $X_t < x^*$, the process will tend to fluctuate around x^* provided that $e < c$, that is, the local extinction rate is smaller than the colonization rate. If $e > c$ there is no such quasi-stationary equilibrium point.

We can investigate this in more detail by using the discrete analog of the Green function $G(x, x_0)$. Now we let this express the expected time the process spends in state x before it goes extinct when it initially at time $t = 0$ is in state x_0. The expected time spent in x during time interval $(t, t + dt)$ is $p_t(x, x_0)dt$, so due to the fact that expectations are additive we must have the relation $G(x, x_0) = \int_0^\infty p_t(x, x_0)dt$. Hence, when solving the equation for $p_t(x, x_0)$ numerically, we may in the same calculations add the contributions $p_t(x, x_0)\Delta t$ to find $G(x, x_0)$. The quasi-stationary distribution is then $p(x, x_0) = G(x, x_0)/T(x_0)$. Examples are shown in Fig.6.1.

Fig.6.2 shows a simulation of a process from time $t = 1000$ until it goes extinct at $t \approx 1650$ years.

6.3 The Moran effect

6.3.1 Correlated time series

We have seen in chapter II that the theta-logistic model with $\theta = 0$ (Gompertz type of density regulation) leads to linear models on the log scale. Fur-

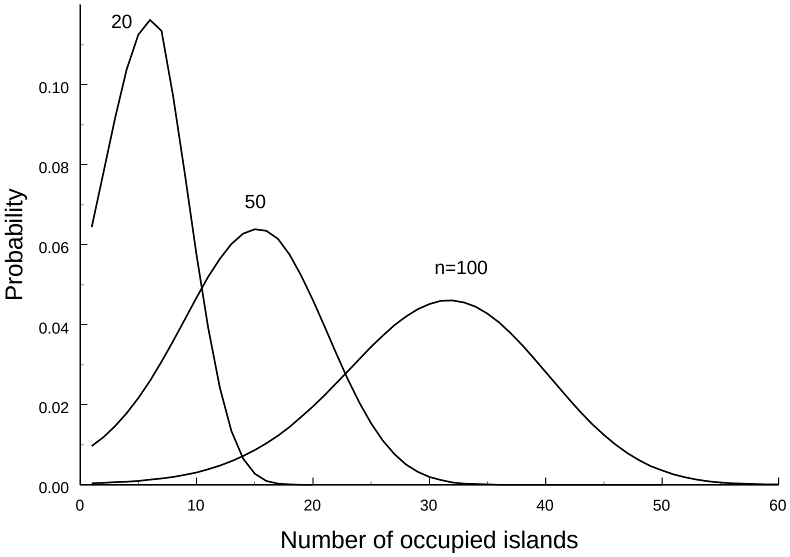

Figure 6.1: The quasi-stationary distribution of number of occupied islands for parameter values $c = 0.06$ and $e = 0.04$ for three different values of metapopulation size n. The time to extinction is approximately 3900 for $n = 20$, 26000 for $n = 50$ and 600000 for $n = 100$. The corresponding values for n^* are 6.7, 16.7 and 33.3.

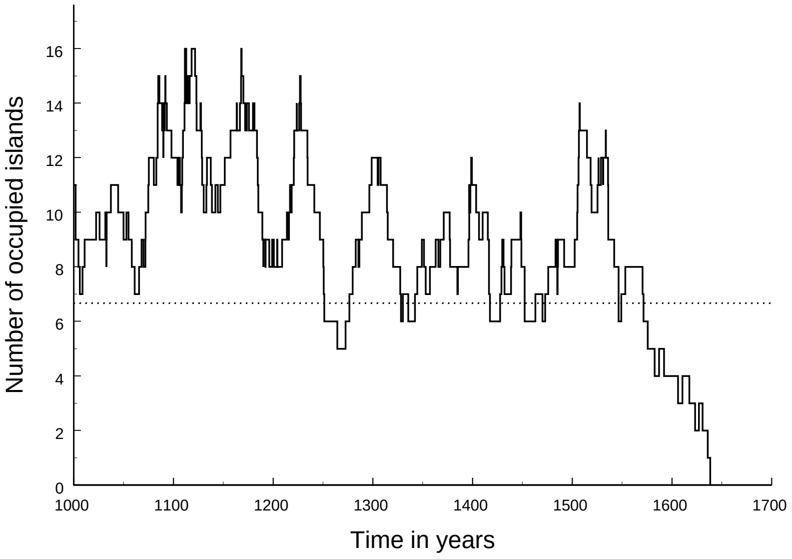

Figure 6.2: Simulation of a meta-population from time $t = 1000$ until it goes extinct at about 1650 years. Parameters are $c = 0.06$, $e = 0.04$, and $n = 20$. The dotted line shows the quasi-stationary equilibrium point $x^* \approx 6.7$.

ther, if the populations are large, the demographic variance can be ignored, and the noise term on the log scale will then have approximately constant variance which actually is the environmental variance σ_e^2. Writing $X = \ln N$ for the log population size we then have

$$X_{t+1} = \alpha_1 + \beta X_t + \varepsilon_t$$

where $\text{var}(\varepsilon_t) = \sigma_{e1}^2$ is the environmental variance and the carrying capacity on the log scale is $k_1 = \alpha_1/(1-\beta)$. Alternatively, we can write the model as $\Delta X = \alpha_1 - \gamma X + \varepsilon$, where $\gamma = 1 - \beta$ is the strength of density-regulation according to the definition made in chapter II. This is a stationary time series when $|\beta| < 1$. If the noise term is normally distributed and there are no temporal autocorrelations in the noise, the stationary distribution is the normal distribution with mean k_1 and variance $\sigma_{e1}^2/(1-\beta^2) = \sigma_{e1}^2/(2\gamma - \gamma^2)$. Notice that if γ is much smaller than one, then the stationary distribution has variance approximately $\sigma_{e1}^2/(2\gamma)$ which is the same as for the analogue continuous model (see 3.8.4). Now, X_{t+1} can be written as a linear combination of previous noise terms (exercise 3)

$$X_{t+1} = k_1 + \sum_{u=0}^{\infty} \beta^u \varepsilon_{t-u}.$$

Let the fluctuations at another location be given by the same model, possibly with a different log carrying capacity k_2, but with the same strength of density-regulation, and noise terms δ_t with variance σ_{e2}^2, giving

$$Y_{t+1} = k_2 + \sum_{u=0}^{\infty} \beta^u \delta_{t-u}.$$

From these assumptions it follows (exercise 4) that if $\text{corr}(\varepsilon_t, \delta_t) = \rho$ then the stationary correlation between the log population sizes $\text{corr}(X_{t+1}, Y_{t+1})$ also equals ρ. This is what is known as the Moran effect.

The same result also holds for higher order autoregressive models because the state for such models can be written as linear combinations of the noise terms. Hence, provided that the autoregressive coefficients are the same at the two locations, the correlation between the log population sizes will equal the spatial correlation of the noise terms. As this holds for two populations, it also holds for each correlation between a set of sub-populations as long as there is no migration between them.

6.3.2 Correlation in linear models in continuous time

The Ornstein Uhlenbeck process

In 3.7 we used the Ornstein-Uhlenbeck process as an approximation to the dynamics of log population size $X_t = \ln N_t$. The complete solution for this process was given in 3.8.4. Now consider two islands with log population size X_t and Y_t, and dynamics given by the diffusions with means $\alpha_1 - x\beta$ and $\alpha_2 - y\beta$ and variances σ_1^2 and σ_2^2, respectively. Then, $X_{t+dt} = X_t + (\alpha_1 - \beta X_t) + \sigma_1 dB_1(t)$ and similarly for Y_{t+dt}. Writing C for the stationary covariance between X_t and Y_t, which can be put equal to $\text{cov}(X_{t+dt}, Y_{t+dt})$, we obtain

$$C = (1 - \beta dt)^2 C + \sigma_1 \sigma_2 \rho_e dt$$

where $\rho_e = \text{E}[dB_1(t)dB_2(t)]/dt$ is the environmental correlation. Omitting second order terms in the limit we then find $C = \sigma_1\sigma_2\rho_e/(2\beta)$. Using the fact that $\text{var}(X_t) = \sigma_1^2/(2\beta)$ and $\text{var}(Y_t) = \sigma_2^2/(2\beta)$ we see that the Moran effect also holds for this model as $\text{corr}(X_t, Y_t) = \rho_e^2$.

The corresponding models for the population sizes N_x and N_y are non-linear, with the Gompertz type of density-regulation. Using well known properties of the bivariate normal distribution for (X_t, Y_t) wee find (exercise 5) that

$$\text{corr}(N_x, N_y) = \frac{e^{\sigma_1\sigma_2\rho_e} - 1}{\sqrt{(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)}}$$

The absolute value of this correlation is always smaller than the absolute value of the environmental correlation ρ_e (exercise 6). However, we see from Fig.6.3 that the correlation between population sizes stay very close to the correlation between the log of population sizes, which is the environmental correlation, even if the environmental variances are very large ($\sigma_1\sigma_2 = 1$).

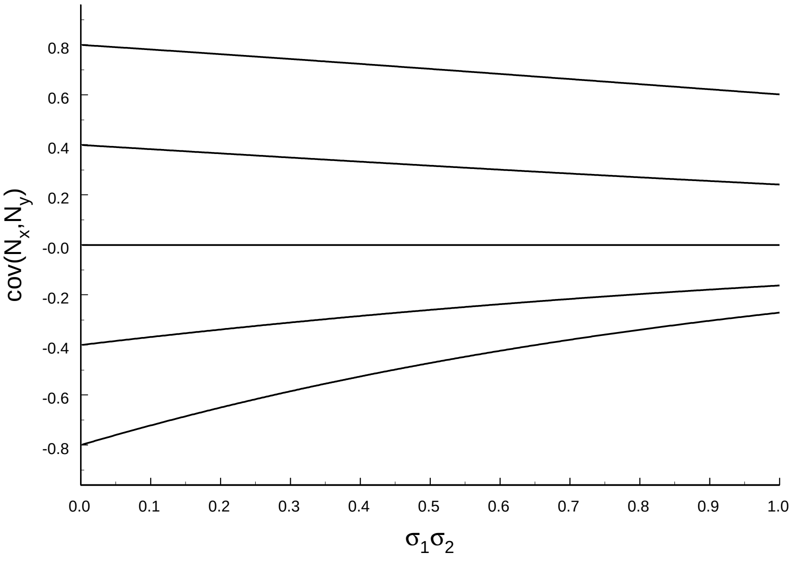

Figure 6.3: The correlation between population sizes $\mathrm{corr}(N_x, N_y)$ plotted as function of $\sigma_1\sigma_2$ when $\sigma_2 = 2\sigma_1$. The values for $\sigma_1\sigma_2 = 0$ is exactly the environmental correlation ρ_e.

6.3.3 Correlation in non-linear models in continuous time

Correlation and measurement scale

We have seen in the previous section that the correlation depends on the scale at which we measure 'population size'. The non-linear Gomperz type of model could be linearized by the log transformation, $X = \ln N$, leading to the linear Ornstein-Uhlenbeck process for which the Moran effect holds. However, for the non-linear process describing the fluctuations in the population size N, a positive correlation in the noise will lead to a positive but smaller correlation between the population sizes. So, as the Moran effect seems to dependent on our choice of measurement scale, the above example indicates that we should look for the Moran effect using a measurement scale that makes the process as close to linear as possible. Other measurement scales are then likely to give smaller correlations. We could even say, as a definition, that the Moran effect holds for N in the Gompertz type of model because the correlation between population sizes on a suitable scale (here the log scale) equals the environmental correlation.

Linearizing the theta-logistic model

We now consider the diffusion with infinitesimal mean $rn[1-(n^\theta-1)/(K^\theta-1)]$ and and (environmental) variance σ^2, that is, the theta-logistic model with population size large enough for the demographic variance to be ignored so that σ^2 is the environmental variance. For this model of population fluctuations we now introduce the transformation

$$X = g(N) = \frac{1 - N^{-\theta}}{\theta}$$

and use the limiting form $X = \ln N$ for $\theta = 0$. Hence, for $\theta = 0$ this is exactly the Gomperz model and the transformation leads to the linear OU-process for X in that case. Using the transformation formulas for diffusions dealt with in chapter 3.5, we find (exercise 7) that the infinitesimal mean and variance for X can be written as $\mu(x) = a - bx$ and $\nu(x) = (1 - \theta x)^2\sigma^2$, where $a = r - (\theta + 1)\sigma^2/2$ and $b = r/g(K) - \theta(\theta + 1)\sigma^2/2$.

In order to find the stationary correlation between two populations described by this model we first need to write the models as stochastic differential equations (see 3.14). However, the simple example we dealt with in 3.14 used a constant infinitesimal variance so that the two different ways of writing up the stochastic differential equation from the diffusion were equivalent. Here we use the approach corresponding to so-called Stratonovich integration, an approach that ensures that the diffusions and the stochastic differential equations are both stationary for the same set of parameters. The general form of the equation is then

$$dX_t = [\mu(X_t) - \frac{1}{4}\nu'(X_t)]dt + \sqrt{\nu(X_t)}dB(t)$$

where $B(t)$ is a standard Brownian motion as before. This stochastic differential equation for the transformed theta-logistic model then takes the form (exercise 8)

$$dX_t = (\alpha - \beta X_t) + (1 - \theta X_t)\sigma dB(t),$$

where $\alpha = r - \sigma^2/2$ and $\beta = r/g(K) - \theta\sigma^2/2$. This linearized theta-logistic model can now be used to study deviations from the Moran effect.

Correlations in the linearized model

Now consider two island which both can be described by the above theta-logistic model, with parameters and variables indexed 1 and 2, and with environmental correlation ρ_e defined by $\mathrm{E}[dB_1(t)dB_2(t)] = \rho_e dt$. The solution is found by requiring that the covariance at stationarity $c_{12} = \mathrm{cov}(X_1, X_2) = \mathrm{cov}(X_1 + dX_1, X_2 + dX_2)$. Writing out these covariances and omitting terms of order $(dt)^2$, we find, utilizing that dB_1 and dB_2 are independent of X_1 and X_2, that

$$\mathrm{cov}(X_1 + dX_1, X_2 + dX_2) = [1 - (\beta_1 + \beta_2)dt]\mathrm{cov}(X_1, X_2) + \omega \qquad (6.1)$$

where

$$\omega = \sigma_1\sigma_2 \mathrm{cov}[(1 - X_1\theta_1)dB_1, (1 - X_2\theta_2)dB_2]. \tag{6.2}$$

Collecting all terms in ω we find (exercise 9)

$$\omega = \sigma_1\sigma_2 \mathrm{E}(1 - X_1\theta_1)(1 - X_2\theta_2)\rho_e dt$$

or

$$\omega = \sigma_1\sigma_2\rho_e[1 - \alpha_1\theta_1/\beta_1 - \alpha_2\theta_2/\beta_2 + \alpha_1\theta_1\alpha_2\theta_2/(\beta_1\beta_2) + \theta_1\theta_2 c_{12}]dt.$$

Finally we require, assuming that both processes are stationary, that $\mathrm{cov}(X_1, X_2) = \mathrm{cov}(X_1 + dX_1, X_2 + dX_2)$. Solving this equation for c_{12} then gives (exercise 10)

$$c_{12} = \frac{\rho_e\sigma_1\sigma_2[1 - \alpha_1\theta_1/\beta_1][1 - \alpha_2\theta_2/\beta_2]}{\beta_1 + \beta_2 - \rho_e\sigma_1\sigma_2\theta_1\theta_2}$$

From this we find the corresponding correlation $\mathrm{corr}(X_1, X_2) = \rho_x = \frac{c_{12}}{\sqrt{c_{11}c_{22}}}$ using the fact that $\mathrm{corr}(X_1, X_1) = \mathrm{corr}(X_2, X_2) = 1$, which turns out to be (exercise 11)

$$\rho_x = \mathrm{corr}(X_1, X_2) = \rho_e \frac{\sqrt{(2\beta_1 - \theta_1^2\sigma_1^2)(2\beta_2 - \theta_2^2\sigma_2^2)}}{\beta_1 + \beta_2 - \rho_e\sigma_1\sigma_2\theta_1\theta_2}. \tag{6.3}$$

Some special cases

Is is often realistic to assume that the local dynamics is the same at two locations, in which case we can omit the parameter indices. Writing $\delta = \theta^2\sigma^2/(2\beta)$ we then find

$$\rho_x = \rho_e \frac{1 - \delta}{1 - \delta\rho_e}. \tag{6.4}$$

The population correlation ρ_x is shown as function of the environmental correlation in Fig.6.4. Since δ is proportional to θ^2 and θ expresses the degree of non-linearity, we can immediately see from the graph that non-linearity in the dynamics tends to reduce the synchrony in population fluctuations. This effect increases with the environmental variance σ^2 and mean return to

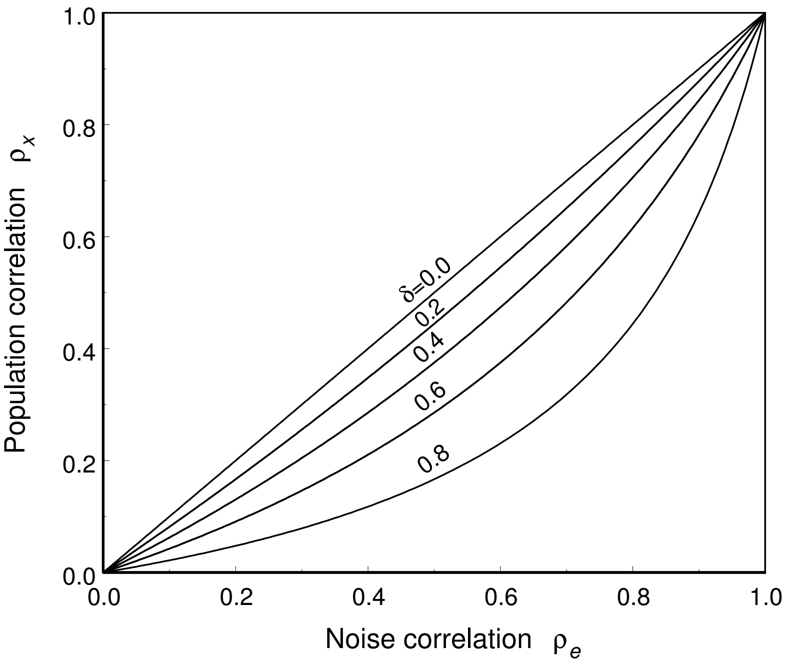

Figure 6.4: The population correlation ρ_x as function of the noise correlation ρ_e for different values of $\delta = \sigma^2\theta^2/(2\beta)$ when the population parameters are constant in space.

equilibrium $1/\beta$. If $\theta = 0$ giving $\delta = 0$ we have the linear model on the log scale, and in accordance with what we have already seen, the Moran effect $\rho_x = \rho_e$ then applies.

In order to illustrate the effect of different dynamics at the two locations we consider the log-linear model, that is, chose $\theta = 0$. The, inserting $s = \beta_1/(\beta_1 + \beta_2)$ in the general formula we find $\rho_x = \rho_e\sqrt{4s(1-s)}$. Since $4s(1-s)$ has maximum 1 for $s = 1/2$ corresponding to $\beta_1 = \beta_2$ we see that any kind of differences in the strength of density-regulation (or return time to equilibrium) also tends to make the synchrony in population fluctuations smaller.

6.4 Continuous spatio-temporal models

Although the meta-population approach of considering separate populations in space is realistic for many organisms, others may be spread over large areas in a way that makes it impossible to define distinct sub-populations. In order to describe such populations it is required that we make a distinction between population density, the mean number of individuals per unit area, and the process determining the actual positioning of each individual. The density can often be considered as a continuous function in space and described mathematically as a stochastic field in space and time when the dynamics are stochastic. The location of the individuals, however, are more realistically described by a point process. These two processes may often have different spatial scales as well as time scales. The point processes determined by individuals clumping in groups or by competition between single individuals typically acts over small distances where densities can be considered as practically constant. Larger distances are required to really see significant changes in population density. And for the temporal aspects, the movement of individuals during short time intervals such as days or weeks may be described by considering the spatial density function as given. The density function, on the other hand, changes on a slower time scale, often the scale of years, as it is yearly reproduction, death, annual migration and environmental noise in these variables that determine its variation in space and time. We now first give some definitions and results for continuous density functions. Then, in 6.5 we add a point process to see how counts are likely to vary all the way from very small to very large sampling units or areas.

6.4.1 Population density function and spatial autocorrelation

Writing $z = (z_1, z_2)$ for points in the plane we define the spatial density function at a given time, $\Lambda(z)$, by the expected number of individuals in an area A being $\int_A \Lambda(z)dz$. For a small area dz at location z the expected number of individuals is then $\Lambda(z)dz$. If two individuals cannot be at the same position, then $\Lambda(z)dz$ also expresses the probability of finding an individual at dz. It is often convenient to work with density functions defined over the whole two-dimensional Euklidian space. Then, choosing two points at random at prescribed distance h gives two densities, $\Lambda(z)$ and $\Lambda(z+h)$. If Λ is a stationary process in space, then the bivariate distribution of these two recordings of Λ at distance h is a function of h only. Then as we define temporal autocovariance functions for stochastic processes in time we define the spatial autocovariance function $c(h) = \text{cov}[\Lambda(z), \Lambda(z+h)]$, where z is interpreted as a point chosen randomly in the field. The corresponding autocorrelation function is then $\rho(h) = \text{corr}[\Lambda(z), \Lambda(z+h)] = c(h)/c(0)$, where $c(0) = \sigma^2$ is the variance of Λ at for random points in space.

6.4.2 Measures of spatial scale

As we have mentioned earlier, different kinds of populations may show synchrony in population fluctuations over very different distances, so it may be convenient to define parameters defining this scale. Some populations, like fish in the ocean, may show highly correlated fluctuations over hundreds of kilometers, while fluctuations in butterfly populations may be almost independent for distances as small as a few kilometers.

Parameters of spatial scale, are properties of the spatial autocorrelation function. It is tempting to choose the scale as the distance at which the correlation has reached zero. However, such a distance is hard to find as the spatial correlation may approach zero but actually not reach that value. And if the function reach zero, that distance may be very hard to estimate accurately from data. It is therefore more convenient to define the scale as the distance where the correlation has reached some given value smaller than one. The value $e^{-1} \approx 0.37$ has been used a lot. Notice that if the correlation function is not isotropic, that is, $\rho(h)$ is not a function of the Euklidian distance $r = \sqrt{h_1^2 + h_2^2}$, then the scale varies with direction. For isotropic models, however, it is a single parameter.

Another measure which is mathematically convenient, is to consider the correlation as function of distance along a line transect, normalize this function so that it integrates to one, and use the standard deviation of that distribution as the scale along the line transect. For the isotropic correlation function of the Gaussian form

$$\rho(h) = e^{-(h_1^2+h_2^2)/(2l_\rho^2)}$$

the parameter l_ρ is exactly this scale in any direction. The measurement defined as the distance where ρ is e^{-1}, however, is $l_\rho\sqrt{2}$.

6.4.3 Gaussian and log-Gaussian density fields

We now consider a linear model (absolute or log scale) defined at each point z in space

$$dX_t(z) = [(\alpha - \beta X_t(z)]dt + \sigma dB(t, z).$$

Then, we have seen in chapter 3 that the stationary distribution at location z is normal with mean $\mu = \alpha/\beta$ and variance $\sigma^2 = \sigma_e^2/(2\beta)$. We have also seen in 3.14 that the solution is linear in $dB(t, z)$. This implies that the stationary distribution for any vector of X-values at some points in space possess a multivariate normal distribution. Further, in 6.3.2 we saw that the correlations are the same as the correlations between the noise terms, that is, the Moran effect holds for this model. The resulting spatial function $X_t(z)$ when stationarity is reached, is, at a given time, a Gaussian spatial field with mean value μ and spatial autocovariance function $\sigma^2\rho(h)$, where $\rho(h)$ is the correlation in the noise terms as function of distance between locations. Mathematically, this is defined by the relation $\mathrm{E}dB_e(t, z)dB_e(t, z + h) = \rho_e(h)dt$. If X is log density, then the field of population densities $\Lambda(z) = e^{X(z)}$ is a log-Gaussian spatial field. As we have seen before, the Moran effect holds for X but not for Λ. Below we show how such fields can be constructed mathematically and simulated.

Construction and simulation of Gaussian and log-Gaussian density fields

A Gaussian field of population densities is a field constructed so that the densities at any finite set of locations has a multivariate normal distribution. Such fields can be constructed by defining analogies to Brownian motions and 'white noise' in space. This can be done by first dividing the plane into square cells by a grid with sides $\Delta z_1 = \Delta z_2$ and area $\Delta z = \Delta z_1 \Delta z_2$. Let $\Delta B(z)$ be independent normal variates defined at each cell center z with $\mathrm{E}\Delta B(z) = 0$ and $\mathrm{var}[\Delta B(z)] = \Delta z$. Then, summing the variables over an area A which is a set of cells is then normal with zero mean and variance A. The variables $W(z) = \Delta B(z)/\Delta z$ are correspondingly independent normal variates with zero mean and variance $1/\Delta z$. The above sum over the area A can then be written as

$$B(A) = \sum_A \Delta B(z) = \sum_A W(z)\Delta z.$$

When passing to the limit as the side length of the squares approach zero we write accordingly dz, $dB(z)$, and $W(z) = dB(z)/dz$. The sum of the $dB(z)$ over A then approaches the integral

$$B(A) = \int_A dB(z) = \int_A W(z)dz$$

which is normally distributed with zero mean and variance A. The process $W(z)$, which is constructed exactly as we have previously constructed 'white noise' on the time scale, is a 'white noise' process in space, while $B(A)$ is a Brownian motion. If we in some way extend the area and plot $B(A)$ against A the graph will be exactly the one-dimensional Brownian motion as we have previously defined it as a process in time, only that time has been replaced by area. Notice that the variance of $W(z)$ is $1/dz$ and hence approach infinity in the limit as for 'white noise' on the temporal scale.

We can now define a Gaussian field $\Lambda(z)$ as

$$\Lambda(z) = \mu + \int g(z-u)dB(u) = \mu + \int g(z-u)W(u)du$$

where the integral runs over the entire plane. As the variance of $\Lambda(z)$ equals $\sigma^2 = \int g(z-u)^2 du = \int g(u)^2 du$ we must require that this integral is finite. Since $\Lambda(z)$ is linear in normal variates it is normally distributed with mean μ and variance σ^2.

Starting with the grid of squares and passing to the limit we see that the autocovariance function $c(h) = \mathrm{cov}[\Lambda(z), \Lambda(z+h)]$ is

$$c(h) = \mathrm{cov}[\int g(z-u)dB(u), \int g(z+h-u)dB(u)] = \int g(z-u)g(z+h-u)du.$$

Changing variable from u to $v = z - u + h$ we get

$$c(h) = \int g(v-h)g(v)dv.$$

Notice that this covariance is the convolution of $g(v)$ with $g(-v)$.

As an example let g have the Gaussian form

$$g(v) = a\phi(v_1/b)\phi(v_2/b)$$

where ϕ is the standard normal density function. Then (exercise 12) we find that $c(h) = a^2\phi(h_1/l_\rho)\phi(h_2/l_\rho)$ where $l_\rho = b\sqrt{2}$. The variance of this stochastic field is $\sigma^2 = c(0) = a^2/(2\pi l_\rho^2)$. Hence, we obtain a stochastic field $\Lambda(z)$ with variance σ^2 and autocorrelation of the Gaussian form with scale

l_ρ, that is, proportional to a Gaussian distribution with standard error l_ρ, by choosing $a = \sqrt{2\pi\sigma^2 l_\rho^2}$ and $b = l_\rho/\sqrt{2}$. The function g now takes the form

$$g(v) = \sqrt{2\pi\sigma^2 l_\rho^2}\phi[v_1/(l_\rho/\sqrt{2})]\phi[v_2/(l_\rho/\sqrt{2})].$$

The problem of finding a function g that leads to a prescribed Gaussian field is most conveniently solved using Fourier transforms. Here we only consider the above case of spatial covariance function with Gaussian form.

The Gaussian field can be simulated using discrete square cells Δz and the relation $\Lambda(z) \approx \mu + \sum g(z-u)\Delta B(u)$, where $\Delta B(u)$ are independent normal variables with zero mean and variance Δz. The most efficient approach, requiring no storing of the $\Delta B(u)$, is to simulate the contribution from a cell at position u to the whole field, and successively add the contributions from all cells. That is, for each u we first simulate $\Delta B(u)$, then add its contribution $g(z-u)\Delta B(u)$ to all cells z in the area A we want a picture of. The u-cells must cover an area which is so large that contributions from outside give no significant contributions to cells in A. If $g(z)$ is approximately zero for distances larger than, say l_0, then we can use the contributions from all u at distance smaller than l_0 from A. Finally the mean μ must be added to each cell. Examples of a simulated Gaussian field is shown in Fig.6.5, while Fig.6.6 shows the corresponding log-Gaussian field with lognormal stationary distribution.

6.4.4 Effect of permanent heterogeneity in the environment

A major reason why population density varies in space is permanent spatial heterogeneity in habitat quality. A general linear dynamic model for this is

$$dX_t(z) = [\alpha(z) - \beta(z)X_t(z)]dt + \sigma(z)dB(t,z).$$

Since this is a linear model we can use results for the theta-logistic model with $\theta = 0$. Then, considering two locations z_1 and z_2 we see that the stationary

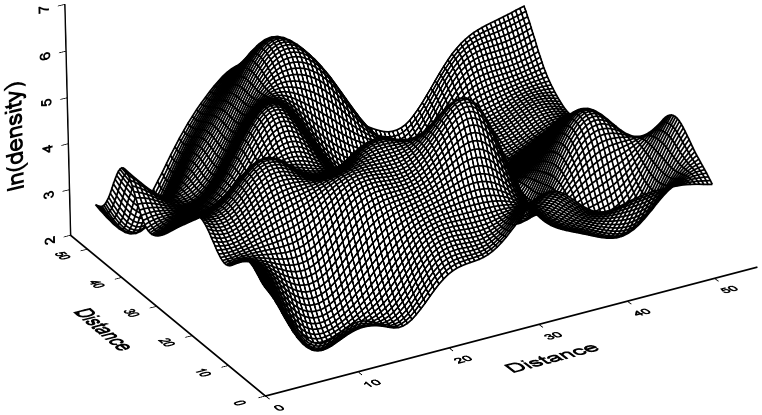

Figure 6.5: A simulation of a Gaussian field with mean $\mu = 4$, variance $\sigma^2 = 1$ and spatial scale $l_\rho = 7$.

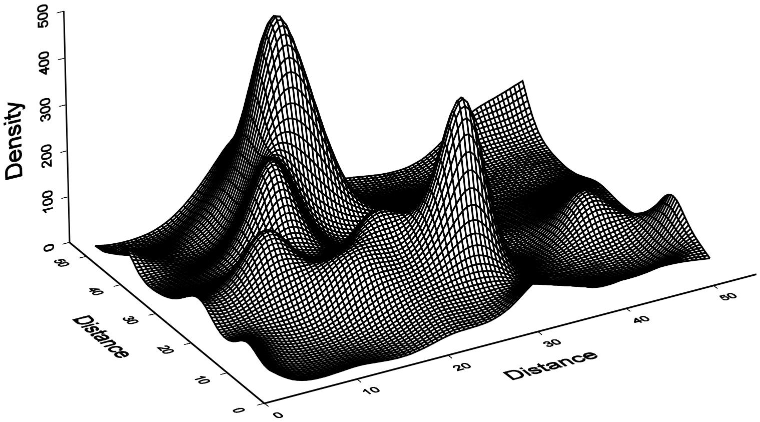

Figure 6.6: Simulation of a log-Gaussian field with log-normal stationary distribution. The logarithm of this field is the Gaussian field shown in Fig.6.5.

distribution for the two densities are binormal (due to the linearity) with mean values $\mu(z) = \alpha(z)/\beta(z)$ and variances $\sigma(z)^2/[2\beta(z)]$, for $z = z_1, z_2$. The correlation is, from equation (6.3)

$$\rho(z_1, z_2) = \rho_e(z_2, z_2)\frac{\sqrt{4\beta(z_1)\beta(z_2)}}{\beta(z_1) + \beta(z_2)}.$$

There are two natural biological interpretations of this linear model. First, the model may be considered as a linearization in population size of some non-linear model, then X is the population size. Using the definition of γ, the strength of density regulation, in 2.3 we see (exercise 13) that $\gamma(z) = \beta(z)$ while $\alpha(z) = \gamma(z)K(z)$, where $K(z)$ is the carrying capacity at location z. Such a linearization will only be valid for rather small fluctuations in population size around $K(z)$. The variance parameter $\sigma^2(z)$ should then be interpreted as the noise at equilibrium, which, expressed by the demographic and environmental variance is

$$\sigma^2(z) = \sigma_d^2(z)K(z) + \sigma_e^2(z)K(z)^2.$$

The other interpretation of the model is to define $X_t(z)$ as log population size, giving what we have called the Gompertz type of model for the population size. In this case the strength of density regulation is still $\gamma = \beta$, but α should be interpreted as the stochastic growth rate and σ^2 as the environmental variance σ_e^2. The demographic variance is now assumed to be zero. Alternatively this can also be viewed as a linear approximation in log population size for small fluctuations in log population size around equilibrium, in which case we have $\sigma^2 = \sigma_e^2 + \sigma_d^2 e^{-K}$. In this interpretation $K = \alpha/\beta$ is the carrying capacity on the log scale.

For simplicity we now assume that noise as well as density regulation are homogeneous in space, while the local carrying capacities vary. Then, the process $Y_t(z) = X_t(z) - K(z)$ takes the form

$$dY_t(z) = -\gamma Y_t(z)dt + \sigma dB(t, z)$$

with spatial autocorrelation $\rho_e(h)$ (the Moran effect), where $\rho_e(h)dt = \mathrm{E}[dB(t,z)dB(t,z+h)]$. The spatial covariance function for Y_t is accordingly $\rho_e(h)\sigma^2/(2\gamma)$, while the covariance function for X_t, choosing z randomly in space, is

$$\mathrm{cov}[X_t(z), X_t(z+h)] = \rho_e(h)\sigma^2/(2\gamma) + \rho_K(h)\sigma_K^2,$$

where σ_K^2 and $\rho_K(h)$ is the variance and spatial autocorrelation for the field $K(z)$. It follows that the spatial autocorrelation for X_t is

$$\rho(h) = p\rho_e(h) + (1-p)\rho_K(h),$$

where $p = \sigma^2/(2\gamma)/[\sigma^2/(2\gamma) + \sigma_K^2]$.
Writing I with appropriate index for the integral of the correlation along the given direction, and put $q = I_e/(I_e + I_K)$, the spatial scale for the process X_t can be written as (exercise 14)

$$l_\rho^2 = \frac{pql_e^2 + (1-p)(1-q)l_K^2}{pq + (1-p)(1-q)}.$$

We see from this that the spatial scale l takes a value between l_e and l_K. Large values of p and q gives a scaling close to l_e. The parameter p is large if the spatial variance in X_t has a large component $\sigma^2/(2\gamma)$ due to noise compared to the component σ_K^2 generated by permanent spatial heterogeneity. Hence, large noise terms as well as weak density regulation makes the scaling approach l_e. The parameter q is large if I_e is large compared to I_K. These integrals are themselves measures of spatial scale of the noise and heterogeneity. Using I as a measure of spatial scale we have the simple relation $I = pI_e + (1-p)I_K$. Hence, the scale of spatial fluctuations in population density depends on the magnitude as well as the scale of the noise and the heterogeneity in carrying capacity.
For autocorrelations of the Gaussian type we find (exercise 15)

$$l_\rho^2 = \frac{pl_e^3 + (1-p)l_K^3}{pl_e + (1-p)l_K}.$$

6.4.5 The effect of dispersal in homogeneous linear models

Dispersal from an area must lead to an immediate reduction in the density while immigration generates an increase. Although there are single individuals that disperse, it is not a convenient approach to model each single individuals contribution to density. As we have defined density, it is a contribution from several individuals, and may also be mean contribution through shorter time intervals reflecting movements of individuals within their home range.

Here we shall only consider the linear model for density and assume that dispersal out from an area is proportional to the density, that is, all individuals have the same chance of migrating out of the area. We write m for the migration rate, that is, the probability that an individual migrate during time dt is mdt. Individuals that migrate are further assumed to do so instantaneously, ending up in some other area. The displacement has a two-dimensional symmetric distribution $g(u) = g(-u)$. If $X_t(z)$ is the density at z at time t, migration from area dz at location z during dt is accordingly $mX_t(x)dzdt$. This migration generates an increase of density at $z+u$ of $mX_t(z)g(u)dzdudt$. Correspondingly, the density at z gets an increase $mX_t(z-u)g(u)dzdudt$ from individuals dispersing from locations $z-u$ to z. The sum of all these contributions is then $\int mX_t(z-u)g(u)dzdtdu = mdzdt\int X_t(z-u)g(u)du$. Dividing through by dz this leads to the dynamic equation

$$dX_t(z) = [r-\gamma X_t(z)]dt - mX_t(z)dt + mdt\int X_t(z-u)g(u)du + K\sigma_e dB_e(t,z), \tag{6.5}$$

where now demographic stochasticity has been ignored so that $\sigma^2 = K^2\sigma_e^2$. If this model leads to a stationary process in space and time the stationary spatial covariance function $c(h) = \operatorname{cov}[X_t(z), X_t(z+h)]$ does not depend on t implying that $\operatorname{cov}[X_t(z), X_t(z+h)] = \operatorname{cov}[X_t(z) + dX_t(z), X_t(z+h) + dX_t(z+h)]$. Using the above dynamic equation and omit terms of order dt^2

that vanish in the limit, this leads to (exercise 16)

$$c(h) = c(h)[1 - 2(\gamma + m)dt] + 2mdt \int c(h-u)g(u)du + K^2\sigma_e^2\rho_e(h)dt,$$

where $\rho_e(h)dt = \mathrm{E}dB(t,z)dB(t,z+h)$ is the spatial autocorrelation in the environmental noise. Dividing by dt then yields

$$2(\gamma + m)c(h) = 2m \int c(h-u)g(u)du + K^2\sigma_e^2\rho_e(h). \tag{6.6}$$

Integrating this equation over the entire space gives the relation (exercise 17) $I_e = 2\gamma I/(K^2\sigma_e^2)$, where I and I_e are the integrals of $c(h)$ and $\rho_e(h)$ respectively. Dividing (6.6) by I then gives

$$2(\gamma + m)f(h) = 2m \int f(h-u)g(u)du + 2\gamma f_e(h),$$

where $f(h) = c(h)/I$ and $f_e(h) = \rho_e(h)/I_e$ are distributions. Multiplying by h_1^2 and integrating along the h_1-axis then gives the relation (exercise 18)

$$l_\rho^2 = l_e^2 + \frac{ml^2}{\gamma}, \tag{6.7}$$

where l_e is the spatial scale of the environmental noise and l is the standard deviation, both along the direction of the h_1-axis. Again, putting $m = 0$ yields the Moran result $l_\rho = l_e$. The last term in (6.7) represents the effect of migration on the spatial scale of population synchrony. Interestingly, we see that even the effect of small short distance migration may be large, provided that there is a weak density-regulation.

Equation (6.6) is an integral equation that is most easily solved by Fourier transforms, but can also be solved by elementary methods using recursion and the fact that the integral is a convolution.

Starting from $c_0(h) = 0$, inserting $c_i(h)$ on the left side giving c_{i+1} on the right side gives the recursion

$$c_{i+1}(h) = \frac{m}{\gamma + m} c_i * g(h) + \frac{\sigma_e^2 K^2}{2(\gamma + m)}\rho_e(h)$$

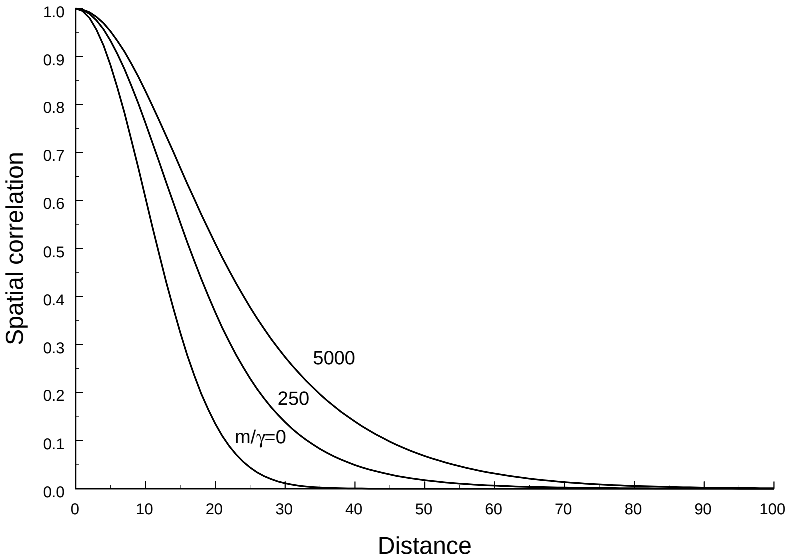

Figure 6.7: The spatial correlation $\rho(h)$ for different values of m/γ against distance h for the isotropic Gaussian model for environmental noise as well as dispersal. Parameter values are $l_e = 10$, $l = 1$.

where $c_i * g(h)$ denotes the convolution of $c_i(h)$ and $g(h)$. Solving recursively then gives (exercise 19)

$$c(h) = \frac{\sigma_e^2 K^2}{2(\gamma + m)} \left[\rho_e(h) + \int \rho_e(h-u) \sum_{i=1}^{\infty} \left(\frac{m}{\gamma + m} \right)^i g^{i*}(u) du \right] \qquad (6.8)$$

where $g^{i*}(h) = \int g(h-u) g^{(i-1)*}(u) du$, with $g^{1*}(h) = g(h)$, is the i-fold convolution of the distribution of dispersal distance with itself. Notice that the spatial autocorrelation function $\rho(h) = c(h)/c(0)$ depends on m and γ only through their ratio m/γ.

If the environmental correlation has the isotropic Gaussian form and the

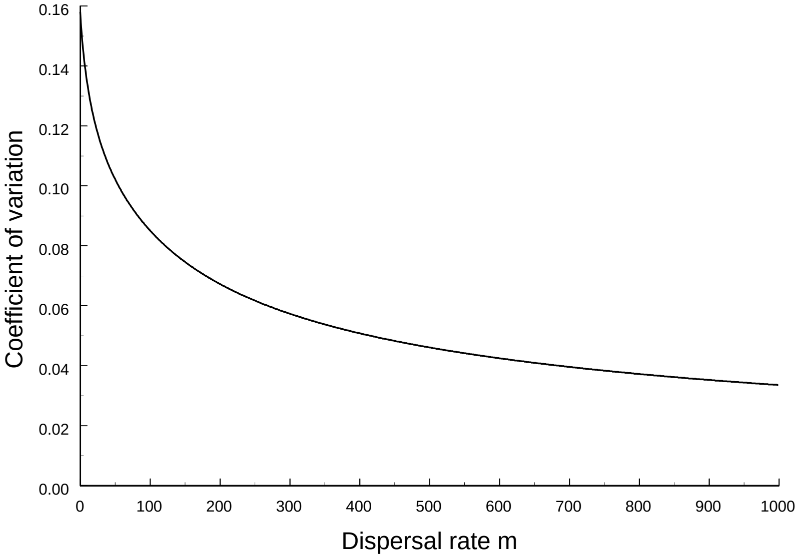

Figure 6.8: Coefficient of variation in density $\sqrt{c(0)}/K$ as function of dispersal rate m. Parameter values are $l_e = 10$, $l = 1$, and $\gamma = 0.2$.

dispersal has an isotropic Gaussian distribution, then (6.8) takes the form (exercise 20)

$$c(h) = \frac{\sigma_e^2 K^2}{2(\gamma + m)} \sum_{i=0}^{\infty} \left(\frac{m}{\gamma + m} \right)^i \phi \left(h/\sqrt{l_e^2 + il^2} \right) \frac{l_e^2}{l_e^2 + il^2} \tag{6.9}$$

where ϕ is the standard normal density.
Fig.6.7 shows an example of how this spatial autocorrelation changes with the migration m, and Fig.6.8 shows how the coefficient of variation of the density $\sqrt{c(0)}/K$ decreases with increasing migration.

6.5 Poisson point process in space

Until now we have only dealt with the density-function, not going down to very small areas where the position of each individual has to be considered. When sampling biological populations in space, however, samples give individual counts, and may be counts of individuals in very small areas. For very small area one may even find only zero or one individual. In order to study the statistical properties of such counts all the way from very small to very large areas, it is required that we also analyze the point processes, that is, give a description of how the position of each individual are distributed in space. Such a description is a point process superimposed on the density function we so far have dealt with.

6.5.1 The homogeneous Poisson process in space

The simplest continuous spatial processes are those with constant density function. The density may vary between years according to some dynamical stochastic model. We write Λ for the spatially constant density a given year. By the definition of density, the expected number of individuals in an area of size A is then ΛA. This function, however, does not alone define how individuals will distribute themselves in space as this may depend on the social organization, individuals tendency of staying close in groups, or the tendency of avoiding one another due to competition for space or other resources. The simplest point process for the constant density function is the Poisson process, where the positioning of each individual is independent of the position of the others. We arrive at this process by assuming that the probability of finding an individual in a small area dA is ΛdA, and that these events are all independent. The moment generating function for the number of points X in dA is then

$$M_{dA}(t) = \mathrm{E}e^{Xt} = (1 - \Lambda dA)e^0 + \Lambda e^t dA = 1 + (e^t - 1)\Lambda dA.$$

As dA symbolizes an area actually approaching zero, the cumulant generating function is

$$K_{dA}(t) = \ln M_{dA}(t) = \Lambda(e^t - 1)dA.$$

Since the cumulant generating function for a sum of variables is the sum of the cumulant generating functions when the variables are independent, we find for any area A, using the fact that the sum of contributions from each dA approaches an integral in the limit, that

$$K_A(t) = \int_A \Lambda(e^t - 1)dA = \Lambda A(e^t - 1),$$

which is the cumulant generating function for the Poisson distribution with mean ΛA. Hence, the number of individuals in an area A is Poisson distributed with mean ΛA. Further, the number of individuals in two disjoint areas must be independent as each term in the sums (integrals) for the two areas are independent. If follows from this that the variance in individual number between equally sized areas must be the variance of the Poisson distribution which equals the mean ΛA. Notice that the coefficient of variation is $1/\sqrt{\Lambda A}$, decreasing with the area A and approaching zero for large areas.

6.5.2 The inhomogeneous Poisson process

The above process may easily be generalized to the case of a general spatial density function. Again, this function may vary between years, but we only consider the distribution of individuals a given year when the function is, say $\Lambda(z)$, where z denotes spatial coordinates. Now, we assume that the probability of an individual being located in a small area $dA = dz$ at location z is $\Lambda(z)dA$. Assuming independence as before we then find that $K_{dA}(t; z) = \Lambda(z)(e^t - 1)dA$. Finally, integrating over the area A gives

$$K_A(t) = \int_A \Lambda(z)(e^t - 1)dA = (e^t - 1)\int_A \Lambda(z)dz.$$

Hence, even in the case of varying density, the number of individuals in disjoint areas are independent Poisson variates, the mean value for an area A being $\int_A \Lambda(z)dz$, that is, the volume under the curve $\Lambda(z)$ in the area A. Alternatively we can write the parameter as $\bar{\Lambda}A$, where $\bar{\Lambda} = A^{-1}\int_A \Lambda(z)dz$ is the mean density in A.

6.6 Point processes with dependence between individuals

The above homogeneous Poisson process and the inhomogeneous Poisson process for a given density field both assume no dependence between individuals. In practice, individuals will often choose positions in a way that depends on positioning of other individuals. Individuals may avoid each other due to intra-specific competition, or they may tend to group together as a social behavior. Point processes for competition leading to under-dispersion relative to the poisson process tends to be mathematically intractable. Also, over-dispersion relative to the Poisson process is more common in biological populations. For the above homogeneous as well as inhomogeneous Poisson process there is no dependence between individual positions for a given density function.

6.6.1 The covariance function for a point process

We now adopt the notation $dN(z)$ for the number of individuals in the infinitesimal area dz at location z. Then, for a given density function Λ we have $\mathrm{E}dN(z) = \Lambda(z)dz$, which is also the probability that the there is one individual at z provided that two individuals can not be located exactly at the same point.

The dependence of each individual's position on other individuals can then be described by the concept of product density at two locations u and v $(u \neq v)$ defined as $\mathrm{E}[dN(u)dN(v)] = f(u,v)dudv$ where the expectation is taken with respect to the point process for some given density function. The spatial covariance function for the point process is defined accordingly as $g(u,v) = f(u,v) - \Lambda(u)\Lambda(v)$, which is the covariance between $dN(u)/du$ and $dN(v)/dv$. For the above homogeneous or inhomogeneous Poisson process the individuals' positions were independent of each other, implying that the the covariance function $g(u,v)$ is zero.

We define the point process covariance function for a random point z and a given displacement $h \neq 0$ as the mean value

$$C_p(h) = \mathrm{E}_z g(z, z+h).$$

where E_z refers to the random choice of z in space. Here we have used the subscript p to emphasize that this component of the spatial covariance function is made up of covariances that are generated by the point process defining local displacement of individuals.

It remains to evaluate the unconditional covariance function, also taking into account the spatial covariance function $C_d(h) = \mathrm{cov}_z[\Lambda(z), \Lambda(z+h)]$ of the field Λ, where d denotes density and cov_z is the covariance evaluated for random choice of z. We define the total spatial covariance function, taking into account the point process as well as the random choice of z, as

$$C(h) = \mathrm{cov}[dN(z), dN(z+h)]/[dzd(z+h)]$$

Using the general formula for unconditional covariances, first conditioning on z, we then find that

$$C(h) = C_p(h) + C_d(h).$$

In the following section we define a process with the property that knowing an individual exists at a given displacement from another individual has an effect that is independent of density. For given positions $\mathbf{u}$ and $\mathbf{v}$ we assume that $P[dN(\mathbf{v}) = 1|dN(\mathbf{u}) = 1] = \Lambda(\mathbf{v})\beta(\mathbf{v} - \mathbf{u})d\mathbf{v}$, where β is a function of the non-zero displacement that is one for the inhomogeneous Poisson process, larger than one if there is a tendency for clumping, and less than one if the individuals are more regularly spread than the Poisson process, e.g. due to competition for resources or space. Using the definition of the covariance function for the point process we then find $g(u, v) = [\beta(u - v) - 1)\Lambda(u)\Lambda(v)dudv$ giving $C_p(h) = [\beta(h) - 1][\mu^2 + C_d(h)]$ showing that C_p in general may depend on C_d.

6.6.2 Overdispersion in the point process defined by log-Gaussian field

To find a process with the above multiplicative property consider a given spatial density field $\Lambda(u)$ and a spatial Gaussian field $\ln \Omega(u)$ with mean $-\tau^2/2$ and spatial covariance function $\tau^2\rho_p(\mathbf{z})$ where the subscript p indicates that this correlation is a property of the point process. We shall use the field Ω to define the positioning of individuals. Hence, as individuals may redistribute themselves quite often, the field Ω may change quickly with time. However, at a given time Ω is a lognormal spatial field with mean 1. Individuals are spread according to an inhomogenious Poisson process with mean $\Lambda(u)\Omega(u)$. At given positions u and $u + z$ the densities, say X and Y, have a bivariate lognormal distribution with mean values $\Lambda(u)$ and $\Lambda(u + z)$, and the log densities are binormally distributed with mean values $\mu_x = \ln \Lambda(u) - \tau^2/2$ and $\mu_y = \ln \Lambda(u + z) - \tau^2/2$, variances τ^2 and correlation $\rho_p(z)$. The marginal distribution of X is the lognormal

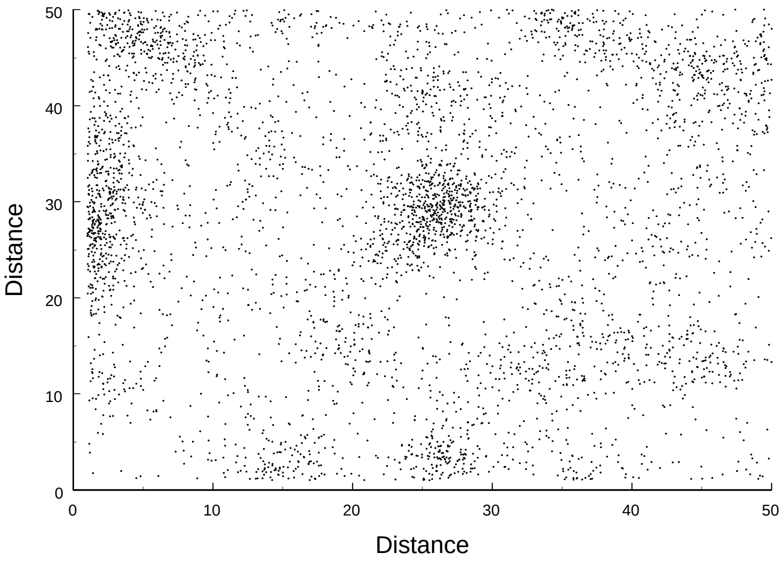

Figure 6.9: A simulation of the point process with constant density equal to one. The field $\ln \Omega(u)$ generating the point process has parameter $\tau = 2$ and the Gaussian type of spatial autocorrelation with standard deviation $l_p = 5$.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\tau}\frac{1}{x}e^{-\frac{1}{2}\left(\frac{\ln x - \mu_X}{\tau}\right)^2}.$$

Fig.6.9 and Fig.6.10 show a realization of such a process with constant and variable density.

In order to give an idea of how individuals in this process relate to each other we now derive an expression for how the knowledge of an individuals in a given position affects the probability of finding one at some nearby location. Let B denote the event that there is an individual at du. Then, considering the conditional probability $P(B|X = x) = xdu$ and using Bayes' formula

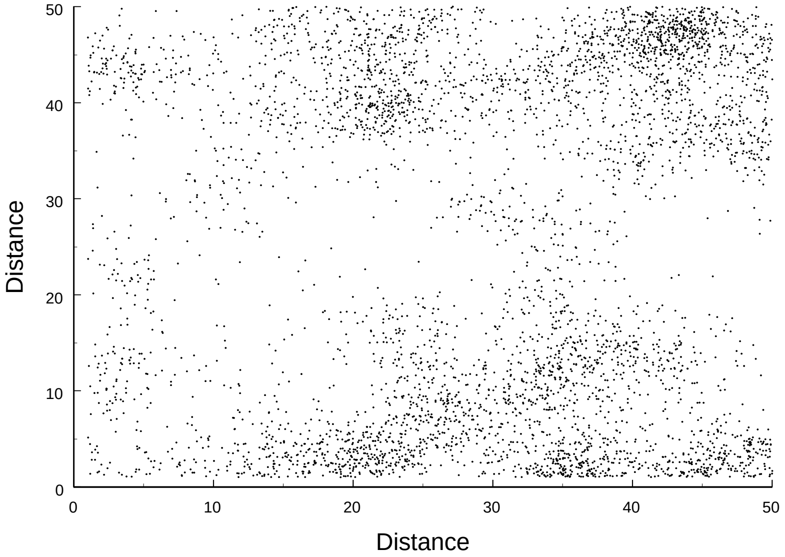

Figure 6.10: A simulation of the point process with log-Gaussian type of spatial density. The variance of $\ln \Lambda(u)$ is 1 with spatial autocorrelation of the Gaussian type with standard deviation $l_d = 20$. The field $\ln \Omega(u)$ generating the point process has parameter $\tau = 2$ and the Gaussian type of spatial autocorrelation with standard deviation $l_p = 5$.

$$f(x|B) = \frac{P(B|X = x)f(x)}{P(B)}$$

where $P(B) = \int P(B|X = x)f(x)dx$, the distribution of X given the event B is seen to be proportional to $xf_X(x)$ (the first moment distribution) which is itself a lognormal distribution $f_X(x)$ with μ_x replaced by $\mu_x + \tau^2/2$ (exercise 21). The distribution of $\ln Y$ conditioned on $X = x$ is normal with mean $\mu_y + \rho_p(z)(\ln x - \mu_x)$ and variance $\tau^2[1 - \rho_p(z)^2]$. Using the conditional distribution $f(x|B)$ we then see that the marginal distribution of $\ln Y$ given B is normal with mean $\mu_y + \rho_p(\mathbf{z})\tau^2/2$ and variance τ^2. The mean of Y given B is accordingly $\exp[\mu_y + \frac{1}{2}\tau^2\rho_p(\mathbf{z}) + \frac{1}{2}\tau^2]$. For a given value y of Y the probability of finding an individual at $d(u + z)$ is $yd(u + z)$, and conditional on B this probability is $\mathrm{E}(Y|B)d(u + z) = \Lambda(u + z)e^{\tau^2\rho_p(z)/2}d(u + z)$. Hence, this point process has the nice property that the information about an individual being at distance z from a point changes the probability of an individual at this point by the factor $\beta(z) = e^{\tau^2\rho_p(z)/2}$. If there is no correlation, that is, at distances large enough for $\rho_p(z)$ to be zero, this factor is 1, and there is no effect of the other individual. On the other hand, an individual at a close distance with $\rho_p \approx 1$ gives a factor $e^{\tau^2/2}$ which is large for large values of τ.

6.6.3 Mean and variance of counts in an area

Let now A be an area of given shape and size and suppose sampling is performed by counting individuals in such areas chosen at random positions in space. Initially let A be the union of small cells Δz with $\Delta N(z)$ individuals so that the total number of individuals in A is $N = \sum_z \Delta N(z)$, where the sum is taken over all cells defining A. Then $\mathrm{E}N = \sum \mathrm{E}\Delta N$ and $\mathrm{var}N = \sum_z \mathrm{var}[\Delta N(z)] + \sum_z \sum_u \mathrm{cov}[\Delta N(z), \Delta N(u)]$. As the cell size approach zero the sum approach the corresponding integrals giving mean value

$$\mathrm{E}N = \int_A \mathrm{E}dN(z) = \int_A \mu dz = \mu A$$

and variance

$$\mathrm{var}(N) = \int\int C(z-u)dzdu + \int_A \mu dz,$$

where the double integral runs over A^2 excluding $z = u$. The last integral is the contribution from the points $z = u$ which is the only term different from zero if the points obey the homogeneous Poisson process in space, that is, the field Λ of population densities is constant and the individuals are randomly distributed. For such a process we have no autocorrelation for either the point process or the population density and find $\mathrm{E}N = \mathrm{var}(N) = \mu A$ in accordance with the theory for the homogeneous Poisson process.

For an isotropic model the above double integral can alternatively be expressed as $A^2\mathrm{E}C(R)$ where R is the distance between two points chosen at random in the area A.

Let us now keep the shape of the sampling unit constant, for example as a quadrat or a circle, and let the area vary. If the area has unit size, $A = 1$, then the distribution of R, say $f_1(r)$, can in principle be derived for any shape. In general, the variance of the number of individuals within an area of size A chosen at random in the field is then

$$\mathrm{var}(N) = \mu A + A^2 \int_0^{r_{max}} C(r\sqrt{A}) f_1(r)dr, \tag{6.10}$$

where r_{max} is the maximum distance between two points in the unit area. For the quadrat with unit area we have accordingly that $r_{max} = \sqrt{2}$.

Here we consider quadrats only, for which the distribution of the distance between two randomly chosen points is (exercise 22)

$$f_1(r) = \begin{cases} 2r(\pi + r^2 - 4r) & \text{for } 0 \le r \le 1 \\ 2r[2\arcsin(2/r^2 - 1) + 4(r^2-1)^{1/2} - 2 - r^2] & \text{for } 1 \le r \le \sqrt{2}. \end{cases} \tag{6.11}$$

The mean and variance can now be computed numerically for any model for population density given by $C_d(h)$ with superimposed point process given by

$C_p(h)$ by plugging $C(h) = C_p(h)+C_d(h)$ into equation (6.10) and performing the integration numerically.

6.7 Relations to Taylor's scaling laws

Taylor's scaling law are empirical laws for the relation between the mean m and variance V of the number of individuals. For a number of different types of sampling in space and time it has been observed that there is approximately a linear relationship between V and m in a double logarithmic plot with a slope usually found to be between 1 and 2. Here, equation (6.10) gives this relationship indirectly for the case that m varies by changing the size of the sampling unit.

6.7.1 General expression for the variance as function of the mean

Substituting $C = C_p + C_d$ we first separate the variance into components generated by the point process and the spatial density function. As the first term μA is also generated by the point process, we include this term and write $\nu(A)\mu$ for the point process variance component, giving

$$\nu(A) = 1 + A\mu^{-1} \int_0^{r_{max}} C_p(r\sqrt{A}) f_1(r) dr. \tag{6.12}$$

This is an expression for the over-dispersion in the point process relative to the inhomogeneous Poisson process. It is approximately independent of the sampling area A if the scale of the point process is small compared to the side length $\sqrt{A}$ of the quadrat used as sampling unit. For the inhomogeneous Poisson process we see that $\nu = 1$ because $C_p = 0$. Since interactions between individuals usually occurs at rather small distances we often have that C_p is approximately zero for distances larger than say l_p. In that case, appreciable contributions to the integral exist only for $r < l_p/\sqrt{A}$. If $l_p/\sqrt{A}$ is much smaller than one it follows from equation (6.11) that $f_1(r) \approx 2\pi r$ for values of r that contribute to the integral. By a simple scale transformation (exercise 23) of the integration variable we then find that $\nu \approx 1 + 2\mu^{-1}\pi \int_0^\infty uC_p(u)du$ and hence, becomes approximately independent of the area.

With this decomposition the relation between the variance and the area takes the form

$$\mathrm{var}(N) = \nu(A)\mu A + A^2 \int_0^{r_{max}} C_d(r\sqrt{A}) f_1(r) dr. \tag{6.13}$$

Expressing the variance as function of $m = \mu A$ then gives

$$V(m) = \nu(m/\mu)m + m^2\mu^{-2} \int_0^{r_{max}} C_d(r\sqrt{m/\mu}) f_1(r) dr. \tag{6.14}$$

6.7.2 Approximations for small and large sampling areas

Small quadrats

The spatial scale of the density covariance function C_d will often be large compared to the side-length of the quadrat. Then $C_d(r\sqrt{A}) \approx C_d(0) = \sigma^2$ for the whole range of the variable r. If ν is approximately constant this gives the quadratic relationship

$$V(m) \approx \nu m + c_v^2 m^2, \tag{6.15}$$

where $c_v = \sigma/\mu$ is the coefficient of variation for densities Λ in space. As the maximum value of r is $\sqrt{2}$ for a quadrat, this relation is a good approximation if $C_d(\sqrt{2A}) = C_d(\sqrt{2m/\mu}) \approx C_d(0) = \sigma^2$.

Large quadrats

For large areas we utilize the known distribution $f_1(r)$ given by (6.11) to find an approximate relationship. The density covariance function $C_d(r)$ usually approaches zero as r approaches infinity. If $C_d(\sqrt{A}) \approx 0$, that is, the zero value is reached approximately within the quadrat, we may use the formula for $f_1(r)$ valid for $r < 1$, giving the approximation (exercise 24)

$$V(m) \approx (\nu + 2\pi\alpha_1/\mu)m - 8\alpha_2\mu^{-1/2}m^{1/2} + 2\alpha_3 \tag{6.16}$$

for large values of $m = A\mu$, where $\alpha_j = \int_0^\infty C_d(r)r^j dr$.
If C_d the Gaussian form with scale l, we have (exercise 25) $\alpha_1 = \sigma^2 l^2$, $\alpha_2 = \sqrt{\frac{\pi}{2}}\sigma^2 l^3$, and $\alpha_3 = 2\sigma^2 l^4$. For an exponentially decreasing function $C_d(r) = \sigma^2 e^{-r/l}$ the corresponding values are (exercise 26) $\alpha_j = j!\sigma^2 l^{j+1}$.

6.7.3 The slope in Taylor's scaling law

Our approximate relation between the variance and the mean does not have the form $V(m) = am^b$ known as Taylor's scaling law, neither for small not for large areas. However, our approximations for small and large areas can be used to investigate the form of this function. From equations (6.15) and (6.16) we see that $\ln V$ is approximately linear in $\ln m$ for very small and very large $\ln m$ (Fig.6.11), with intercept $\ln \nu$ for small values of $\ln m$ and $\ln(\nu + 2\pi\alpha_1/\mu)$ for large quadrats. The quadratic term in (6.15) has the effect that the difference between $\ln V$ and $\ln \nu + \ln m$ increase with $\ln m$, so that $\ln V$ approaches a straight line for large $\ln m$. The value of $\ln V$ computed from equation (6.15) is half way between the two straight lines with slope 1 for $m^* = \pi\alpha_1/(\mu c_v^2)$. At this value of m we should therefore expect the slope to be close to its largest value. For the Gaussian model of spatial autocorrelation this value is $m^* = \pi l^2 \mu$, occurring at sampling area $A^* = m^*/\mu$, giving simply $A^* = \pi l^2$.
The slope of $\ln V$ versus $\ln m$ will therefore be largest when the the quadrat side-length roughly equals the scale l of the spatial covariance function C_d. Fig.6.11 shows examples demonstrating how the functions are affected by the scale l. It appears that the slope first increases from 1 to 2 with increasing $\ln m$, which is a consequence of equation (6.15). For small values of m the first term is dominating, giving slope 1, while the second term dominates for larger values of m so that the slope approximately equals 2. For even

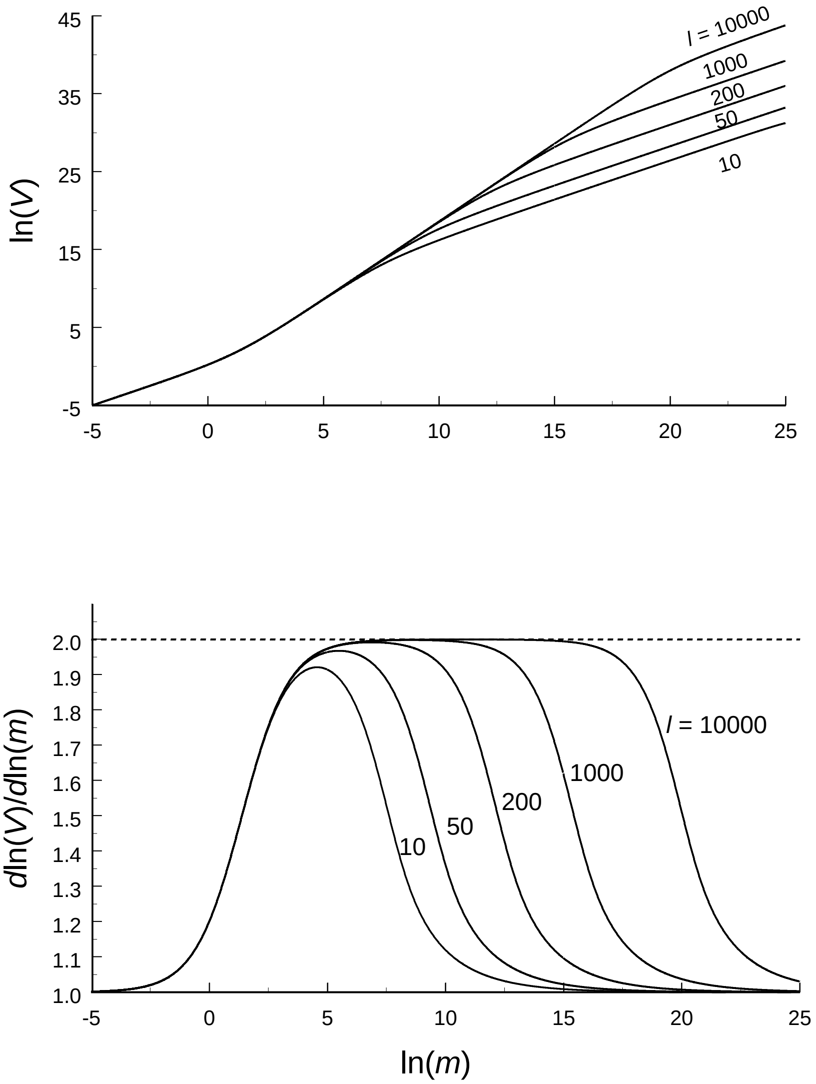

Figure 6.11: Upper panel gives log variance ($\ln V$) as function of log mean ($\ln m$) for increasing size of quadrats and different values of the spatial scale l of the Gaussian type of covariance function $C_d(r)$. Other parameters are $\nu = 1$, $\mu = 1$, and $c_v = 0.5$. The lower panel shows the derivatives of the functions depicted in the upper panel.

larger m, that is, as m gets much larger than $l\sqrt{\pi\mu}$, the curve approaches the large m approximation given by equation (6.16), which again has slope 1 on the log scale. Accordingly, the curves in practice will be indistinguishable from straight lines even over rather large ranges of values of m (Fig.6.11). In empirical data analysis the values of $\ln V$ are usually estimated from sums of squares with rather large uncertainty, so the estimated values will seem to be spread around a straight line.

6.8 Exercises

1. Find the diffusion approximation to the meta-population model in 6.2.
2. Find the return time to the quasi-stationary equilibrium for the meta-population model in 6.2.
3. Show that the state of a first order autoregressive model is a linear combination of noise terms.
4. Show the Moran effect, that the correlation between two log population sizes is the same as the correlation between the noise terms, is correct for two log population sizes described as first order autoregressive models with equal strength of density-regulation.
5. Show that the correlation between the population sizes for the bivariate Ornstein-Uhlenbeck process for log population sizes described in 6.3.2 is $\mathrm{corr}(N_x, N_y) = (e^{\sigma_1\sigma_2\rho_e} - 1)/\sqrt{(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)}$.
6. Show that the absolute value of the correlation in exercise 5 is always smaller than $|\rho_e|$.
Hint: For $\rho_e > 0$ show that $(e^{\sigma_1\sigma_2\rho_e} - 1)/\rho_e$ is an increasing function of ρ_e. Then show that the result holds for $\rho_e = 1$ (if $\sigma_1 \neq \sigma_2$) and hence for all values of $\rho_e > 0$. Use a similar approach for $\rho_e < 0$.
7. Derive the diffusion for the transformed variable $X = g(N) = (1 - N^{-\theta})/\theta$ when the fluctuations in N are given by the theta-logistic diffusion with no demographic variance.
8. Show that the stochastic differential equation $dX_t = [\mu(X_t) - \frac{1}{4}\nu'(X_t)]dt +$

$\sqrt{\nu(X_t)}dB(t)$ for the model in exercise 7 takes the form $dX_t = (\alpha - \beta X_t) + (1 - \theta X_t)\sigma dB(t)$ where $\alpha = r - \sigma^2/2$ and $\beta = r/g(K) - \theta\sigma^2/2$.

9. For two islands described by theta-logistic models with correlated noise derive the expression for the parameter $\omega = \sigma_1\sigma_2\text{cov}[(1 - X_1\theta_1)dB_1, (1 - X_2\theta_2)dB_2$ given in the text.

10. For the model in exercise 9 derive the stationary covariance c_{12} between the 'population sizes' on the X-scale.

11. Derive the expression for the correlation $c_{12}/\sqrt{c_{11}c_{22}}$ given in the text.

12. Consider the equation $c(h) = \int g(v - h)g(v)dv$ and show that $g(v) = \sqrt{2\pi\sigma^2 l_\rho^2}\phi[v_1/(l_\rho/\sqrt{2})]\phi[v_2/(l_\rho/\sqrt{2})]$ leads to the Gaussian autocovariance function $c(h) = 2\pi\sigma^2 l_\rho^2\phi(h_1/l_\rho)\phi(h_2/l_\rho)$, where ϕ is the standard normal density.

13. Consider the model $dX = (\alpha - \beta X)dt + \sigma dB(t)$ as a linearization of a non-linear model interpreting X as population size. Use the definition of γ, the strength density regulation, in 2.3 to show that $\gamma = \beta$.

14. Writing I with appropriate index for the integral of the correlation along the given direction, and put $q = I_e/(I_e + I_K)$, show that the spatial scale for the process X_t given in 6.4.4 can be written as $l_\rho^2 = pql_e^2 + (1 - p)(1 - q)l_K^2$, where $p = \sigma^2/(2\gamma)/[\sigma^2/(2\gamma) + \sigma_K^2]$.

15. For autocorrelations ρ_e and ρ_K of the Gaussian form show that the spatial scale in exercise 14 obeys $l_\rho^2 = [pl_e^3 + (1 - p)l_K^3]/[pl_e + (1 - p)l_K]$.

16. Show that the dynamic equation (6.5) leads to equation (6.6) when the process is stationary in space and time.

17. Use equation (6.6) to show that $I_e = 2\gamma I/(K^2\sigma_e^2)$, where I and I_e are the integrals of $c(h)$ and $\rho_e(h)$ respectively.

18. From the equation $2(\gamma + m)f(h) = 2m\int f(h - u)g(u)du + \gamma f_e(h)$, where $f(h) = c(h)/I$ and $f_e(h) = \rho_e(h)/I_e$, show the scaling result (6.7).

19. Use the recurrence formula $c_{i+1}(h) = \frac{m}{\gamma+m}c_i * g(h) + \frac{\sigma_e^2}{2(\gamma+m)}\rho_e(h)$ to derive equation (6.8).

20. Show that the spatial autocorrelation is given by equation (6.9) when the environmental correlation has the isotropic Gaussian form and the dispersal has an isotropic Gaussian distribution.

21. Let X be a random variable with distribution $f(x)$. Let the probability of an event B conditioned on the event $(X = x)$ be $P(B|X = x) = g(x)$. Use Bayes' formula to show that the distribution of X conditioned on B is then $f(x|B) = g(x)f(x)/\int g(u)f(u)du$. Use this to show that the corresponding distribution in 6.5.3 with $g(x) = xdu$ is the lognormal, and find the parameters of this distribution.

22. Show that the distribution of the distance between two points chosen at random in a unit quadrat is $f_1(r)$ given by equation (6.11).

23. Use the definition given by equation (6.12) to show that the over-dispersion $\nu(A)$ is approximately independent of the area when the scale of C_p is small compared to the side length $\sqrt{A}$ of the quadrat, that is, derive the relation $\nu \approx 1 + 2\mu^{-1}\pi \int_0^\infty uC_p(u)du$.

24. Use the formula for $f_1(r)$ valid for $r < 1$ to show that $V(m) \approx (\nu + 2\pi\alpha_1/\mu)m - 8\alpha_2\mu^{-1/2}m^{1/2} + 2\alpha_3$ for large values of $m = A\mu$, where $\alpha_j = \int_0^\infty C_d(r)r^j dr$.

25. When C_d has the Gaussian form with scale l show that $\alpha_1 = \sigma^2 l^2$, $\alpha_2 = \frac{1}{2}\sigma^2 l^3$, and $\alpha_3 = 2\sigma^2 l^4$, where $\alpha_j = \int_0^\infty C_d(r)r^j dr$.

26. For an exponentially decreasing function $C_d(r) = \sigma^2 e^{-r/l}$ show that $\alpha_j = \int_0^\infty C_d(r)r^j dr = (j-1)!\sigma^2 l^{j+1}$.

Chapter 7

Community models

7.1 Introduction

Until now we have discussed a number of stochastic models for a single species only. However, in this discussion we have not ignored the presence of prey, predators and competitors that may affect the dynamics. Our philosophy, which is often in good agreement with empirical findings, is that other species affect one single species in a very complex way. Such complex multivariate interactions are most conveniently described by collecting all biological effects, together with physical components of the environment, into an environmental vector $\mathbf{z}$ that fluctuates between years and generate environmental stochasticity in population growth rates.

The same philosophy is commonly carried over to studies of communities where a large number of species abundances are recorded jointly. Although there may be interactions, the major part of them are most conveniently described by introducing random effects through stochastic terms rather than trying to characterize all specific interactions by deterministic effects expressing for example the degree of competition between all pairs of species. Such models necessarily require a very large number of parameters even under simple linearizations, and lead to over-parametrization relative to the information that can possibly be obtained through data collection. Therefore, stochasticity plays an even more important role in community dynamics than in single species theory. Even models where almost any temporal and spatial changes are purely stochastic, the so-called neutral community models, have become very popular during the last 20-30 years, mainly through the work of S.P. Hubbell on models describing temporal and spatial dynamics of tropical trees.

Realistic description of communities is definitely a complex task that requires some new concepts in addition to those we have used for a single species. One observation of a single species abundance at a given time will only give us a single number and is hardly interesting at all. So, in order to understand the biological processes involved we need repeated observations in time and space. This fact stands in deep contrast to the study of communities. Actually, the major fraction of papers on ecological communities are dealing just with one single sample. However, the fact that such a sample reflects the abundances of a large number of species, raises many interesting questions that can be analyzed without repeating sampling in time or space. For example, the number of different species recorded is by itself an interesting statistic, as well as measurements expressing how abundances vary among species. Such description of a sample from a community can be carried out in a purely statistical way through estimation of appropriate indices of diversity, often defined using the vector of relative species abundances. This requires some knowledge of the sampling method used, but can be performed without going into complex modeling details concerning spatial and temporal dynamics. However, two samples can be compared using appropriate measurements of similarity, which again can form the bases for statistical analysis of temporal and spatial variation in community composition.

Although we first are going to present some simple descriptive measurements, our main focus will be on dynamic models for species abundance distributions utilizing the theory of single species dynamics presented in chapters 1-3. We will make links between these models and the concepts of diversity, its variation in time and space, as well as the concept of similarity. We first consider neutral models and their connection to neutral gene frequency models in population genetics. Next, we define models with mathematically identical species with some deterministic forcing on the abundance of each species through density regulation. Finally, we go on to heterogeneous models with dynamical parameters varying among species.

7.2 Diversity and similarity

Let $y_1, y_2, \ldots, y_s$ be the abundances of a species in the community. These quantities could be the number of individuals or some measure of density. In any case we assume that the expected number of individuals representing a species with abundance y_i in a random sample is proportional to y_i, with a constant of proportionality depending on the sampling effort. The relative abundances are $p_1, p_2, \ldots, p_s$ where $p_i = y_i / \sum y_j$ so that $\sum p_i = 1$.

There is a large number of diversity indices proposed in the literature. Here we only mention briefly the two most commonly applied measurements, the information index and Simpson's index. The information index, often also called Shannon's index, is borrowed from information theory and defined as

$$H_I = -\sum p_i \ln p_i.$$

It is easy to show (exercise 1) that H_I reaches its maximum $\ln s$ for a given s for equal abundances, that is, for $p_i = 1/s$. As a single p_i approaches one and the others zero, the index approaches zero. Hence, the value depends on the distribution of abundances among the species as well as the number of species in the community.

Simpson's index is defined by considering two individuals sampled at random (with replacement) from the community. The probability that they both belong to species i is then p_i^2, and the probability that they belong to the same species is $\sum p_i^2$. Simpson's diversity index H_S is the probability that they belong to different species, that is

$$H_S = 1 - \sum p_i^2.$$

This index also takes its maximum for equal abundances (exercise 2), the maximum being $(s-1)/s$.

As the number of species increases we see that there is little change in the maximum of Simpson's index compared to the information index. Consequently, the information index is considered to be more sensitive to changes in species number. The value of Simpson's index is mainly determined by the few most abundant species. For example, if $p_1 = 0.9$, so that species 1 is the dominant species in the community, Simpson's index is smaller than $1 - 0.9^2 = 0.19$ regardless how large s is, and is therefore classified as a 'dominance index'. The information index, however, can be made arbitrarily large even if $p_1 = 0.9$ (exercise 3), and is therefor a 'species richness index'. The index most sensitive to species number is, however, the species number s itself.

Unbiased estimation of Simpson's index is straightforward under the assumption of multinomial sampling (exercise 4). Under the same sampling model no unbiased estimator of H_I exists (exercise 5), and estimation may be difficult when there is a large number of rare species in the community, many of which not appearing in a finite sample.

Similarity indices are constructed to compare two communities. The most commonly used indices are the Jaccard and Sørensen index, which are actually equivalent. These indices are only based on counting the species occurring in both communities and only in one of them, and are not affected by the distribution of abundance among species. The Jaccard index is

$$J = \frac{A}{A+B+C},$$

where A is the number of species present in both communities, while B and C are the species number only present in one or the other. The Sørensen index is

$$L = \frac{2A}{2A+B+C}.$$

The equivalence of the indices are given by the simple functional relationship between them, $L = 2J/(J+1)$ or $J = L/(2-L)$.

These indices are usually estimated from samples by replacing the species numbers in the communities by observed number of species. However, these estimates can be extremely biased. General methods for bias correction are proposed in the literature, but even these methods can not be trusted in the case of many rare species. We shall return to these problems when dealing with abundance models.

There has been much focus in the literature on applying indices of diversity and similarity in studies of spatial distribution of species in a community. Since the species abundances and the species composition of the community generally changes in space, the diversity at a given local site is usually smaller than the diversity for the total community occupying a large area. The first one has by R.H. Whittaker been named α-diversity and the last is called γ-diversity. The difference between them, which is related to species turnover rate in space, is called β-diversity. For example, if there is large β-diversity, then similarity indices tend to decrease rapidly with increasing distance between the local communities compared.

7.3 Some history of species abundance models

Around 1920 botanists started to study how the number of species recorded in samples increased with the sampling area. However, these scientists were most interested in species number and did not relate their findings to distribution of abundance among the species in the community. The first empirical studies of such distributions were done around 1940, while the first mathematical model fitted to data on butterflies was developed by R.A. Fisher in a classical paper in 1943. Fisher's approach discussed below was based on using an extremely skew gamma distribution for the abundances and assuming Poisson sampling, that is, for a given abundance the number of recordings of a species is Poisson distributed with mean proportional to its abundance. The distribution of individuals among species in the sample then follows what has become known as Fisher's logarithmic series distribution, a limiting form of the zero-truncated negative binomial distribution.

In 1948 F.W. Preston analyzed several large data sets and plotted his data in a way that revealed the shape of the distribution of log abundance. These histograms looked like samples from a normal distribution, but where censored at low abundances because a large number of rare species are unlikely to appear in the sample. Preston did not attempt to model the sampling process using the Poisson model, as Fisher did 5 years earlier. Actually, the Poisson mixture obtained by using the lognormal distribution of abundances is mathematically more intractable than the mixture based on the gamma distribution, which is known to lead to the negative binomial distribution. A maximum likelihood estimation procedure for the Poisson lognormal distribution was not developed until M. Bulmer did so in 1974. Now, however, simple numerical maximization of the likelihood function can easily be done using standard computer software.

The influential ecologist R.H. MacArthur used a rather different approach in 1957, presenting a very simple model that probably was meant to initiate a discussion rather that being a generally valid species abundance distribution. He used the niche concept, arguing that similar species compete for resources and accordingly divide the niche space among them. His simple model, that actually fits well to many bird communities, was to through $s-1$ point at random on a line segment of unit length and let the lengths of the segments defined by these points be the relative abundances of the s species in the community. This is a way of thinking that is rather different from just defining a distribution of abundances, since MacArthur actually attempted to model the niche sharing process. The sharing process defined in this way may appear to be purely stochastic, but the underlying idea is that species occupy different niches in a way determined by their competitive ability. The fact that one species may end up with large abundance while another is rare, is due to differences in traits and how these relate to properties of the occupied territory. Hence, the subdivision of the niche space defined by MacArthur is meant to be temporarily permanent. Although his model, known as 'MacArthur's broken stick model', was oversimplified, his idea has had large impact, inspiring many scientists to develop more realistic niche sharing models.

The counterpart of MacArthur's model is the concept of neutral community models, where nothing is permanent, first introduced by H. Caswell in 1976, and later developed further by the influential ecologist S.P. Hubbell, studying distributions of abundance among tropical trees. Neutral community models, where all temporal changes in species abundances are purely stochastic and there are no permanent differences among species other than their realized abundances, are inherited from population genetics. Neutral genes are genes that are not subject to selection and are analog to species having the same growth rate and the same stochastic dynamics defined by demographic stochasticity only. The total number of alleles then represents the community size, each allele represents a species, allele frequencies represent relative species abundances, and mutations represent new species entering the community by speciation or colonization. The neutral theory of gene frequencies was initiated by Crow and Kimura in 1964, that is, 21 after Fisher publisher his logarithmic series model. Quite remarkably however, the distribution of relative abundances and gene frequencies in these influential papers are the same. This has later inspired ecologists to go further in utilizing neutral gene frequency models in community ecology.

7.4 Neutral species abundance models

7.4.1 The genetic neutral model with random mutations

We consider a constant population with N individuals of a diploid organism in which the genetic drift is determined by the variance effective population size N_e. At a given locus all alleles are assumed to be neutral. We also make the assumption that mutations occur at rate u, and that all mutations result in alleles not already present in the population, giving what is often called the infinite allele model. As the mutation rate is small, it will practically have no effect on the variance of changes in relative abundances. Consequently, the variance in gene frequency change for an allele A with frequency p is approximately $\mathrm{var}(\Delta p) = p(1-p)/(2N_e)$. The probability that an A will mutate and give a reduction $-1/(2N)$ in the frequency of A is $2Npu$. Accordingly, the expected change in the frequency of A is $\mathrm{E}(\Delta p) = -pu$. Now, considering an allele that initially has frequency $p_0 = 1/(2N)$ and approximating the process for its frequency by a diffusion with infinitesimal mean and variance $\mu(p) = -pu$ and $\nu(p) = p(1-p)/(2N_e)$, the Green function for $p > p_0$ with extinction barrier at $p = 0$ is (exercise 6)

$$G(p, p_0) = \frac{S(p_0)}{s(p)\nu(p)} = cp^{-1}(1-p)^{4N_e u-1},$$

where c is a constant. By definition, the Green function multiplied by Δp expresses the expected time the process is in $[p, p+\Delta p]$ before it goes extinct. After a long time there will be a balance between mutations and extinctions and the multivariate process for the set of all gene frequencies will reach stationarity. Since any allele frequency can be described by the above process, the fraction of a long time interval during which the frequency of any allele present in the population takes a value in $[p, p+\Delta p]$ must be proportional to $G(p, p_0)\Delta p$. Consequently, if an allele is chosen at random from the list of those present in the population at a given time, the distribution of its frequency must be proportional to $p^{-1}(1-p)^{M-1}$, for $p > p_0$, where $M = 4N_e u$. Hence, also the expected number of alleles with frequency in $[p, p+\Delta p]$, say $f(p)\Delta p$, must be proportional to the Green function multiplied by Δp. To find the correct scaling factor we use the fact that the frequencies add up to one. The expected total frequency of alleles with frequency in $[p, p+\Delta p]$ is $pf(p)\Delta p$, so the total expected sum of frequencies is $\int_{p_0}^{1} pf(p)dp = 1$ giving $c \approx M$ (exercise 7). The function $f(p) = Mp^{-1}(1-p)^{M-1}$ is sometimes called the 'frequency spectrum'.

If we choose an allele at random from all $2N$ alleles in the population rather than an allele from the list of alleles present, an allele with frequency p is chosen with probability p, and the distribution of its frequency must be proportional to $M(1-p)^{M-1}$. As we have seen, the integral of this function representing the sum of all gene frequencies equals 1, so this is actually a distribution. Similarly, the sum of squares of gene frequencies, which is the fraction of homozygotes, is (exercise 8)

$$H = \int_0^1 p^2 f(p) dp = 1/(M+1).$$

This is also the probability that two alleles chosen at random are equivalent. Hence, 1 minus this quantity is the probability that they are different. This corresponds to Simpson's index of diversity

$$H_S = \frac{M}{M+1},$$

now used as a measure of genetic diversity.

7.4.2 Fisher's log series distribution

When recording the number of species z_n with n representatives in a large sample of butterflies caught in light traps, the two entomologists A. S. Corbet and C. B. Williams discovered that the sequence z_n for $n = 1, 2 \ldots$, was approximately proportional to the harmonic series $1/n$ for many large data sets. A similar data set of tropical butterflies collected by P.J. DeVries over a period of 5 years is shown in Fig.7.1. This observed pattern led R.A. Fisher in 1943 to propose his classical log-series distribution for species abundances in a sample. He approached the problem by first assuming that the species abundances y followed the gamma distribution $[\rho^k/\Gamma(k)]y^{k-1}e^{-\rho y}$ among species, and that a species with abundance y would be represented by a Poisson-distributed number of individuals in the sample with mean νy, where ν is a measure of sampling intensity. Then, the species actually observed will have abundances in the sample following a zero-truncated negative binomial distribution (exercise 9). In the limit as the number of species s approaches infinity and the shape parameter k approaches zero so that sk approaches a constant α, this distribution becomes Fisher's well known logarithmic series distribution, $[-\ln(1-w)]^{-1}w^n/n$, $n = 1, 2, \ldots$ (exercise 10), where $w = \nu/(\nu+\rho)$, and the expected number of species with n representatives in the sample becomes (exercise 11)

$$\mathrm{E}(z_n) = \alpha w^n/n.$$

.

Notice that $w \approx 1$ for very large samples so that the sequence is practically the harmonic series if n is not extremely large. If the abundances y in the community are scaled so that the total expected abundance is one, then $sk/\rho = 1$ and $\alpha = \rho$. The parameter α, which is the single parameter describing the structure of a population following Fisher's model, was later by Williams called the diversity of the community.

For a positive k the relative abundances $p_i = y_i/\sum_{j=1}^{s} y_j$, follow a Dirichlet distribution, while the marginal distribution for a given species is the beta distribution $[\Gamma(\alpha)/[\Gamma(k)\Gamma(\alpha-k)]p^{k-1}(1-p)^{\alpha-k-1}$ (exercise 12). Hence, we find the expected number of species with relative abundance in the interval $(p, p+dp)$ by multiplying by sdp, which under Fisher's limit becomes $\alpha p^{-1}(1-p)^{\alpha-1}dp$ (exercise 13), showing that the 'frequency spectrum' is the same in Fisher's species abundance model and Crow and Kimura's neutral infinite allele model, Fisher's α corresponding to $M = 4N_e u$ in the genetic model.

It follows from our analysis of the neutral genetic model that the expected value of Simpson's index in Fisher's model must be $\alpha/(1+\alpha)$. We may adopt the same approach to find the expected value of the information index. The expected contribution to the index from species with relative abundances in $(p, p+dp)$ is $f(p)(-p\ln p)dp$, and consequently, adding expected contributions from all segments of length dp gives

$$\mathrm{E}(H_I) = -\int_0^1 \alpha \ln p \; (1-p)^{\alpha-1}dp.$$

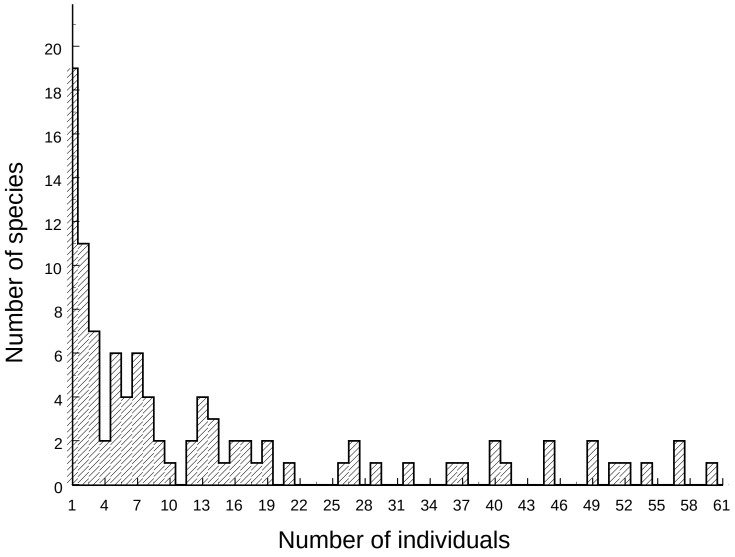

Figure 7.1: The number of species z_n with n representatives in the sample plotted against n for a data set of tropical butterflies collected by P.J. DeVries. There are also 26 species with individual numbers larger than 60, the largest individual number being 756. There are 128 species and 11861 individuals in the sample.

Solving the integral (exercise 14) gives

$$\mathrm{E}(H_I) = \Psi(\alpha + 1) - \Psi(1)$$

where $\Psi(t) = \frac{d}{dt} \ln \Gamma(t)$ is the digamma function for which $\Psi(t+1) = \Psi(t) + 1/t$ and

$$\Psi(t) = \ln t - \frac{1}{2}t^{-1} - \frac{1}{12}5t^{-2} + \frac{1}{120}t^{-4} - \frac{1}{252}t^{-6} + \cdots$$

In particular $\Psi(1) = -\gamma$, where $\gamma = 0.5772\ldots$ is Euler's constant. Consequently, for large values of α we have the approximation $\mathrm{E}(H_I) \approx \ln(\alpha+1) + 0.58$.

7.4.3 Estimation

Fisher did not attempt explaining why abundances seemed to follow the highly skewed gamma distribution. Actually, the way he evaluated uncertainties in his proposed estimator for α indicates that he viewed the gamma model purely as a descriptive tool, giving a realistic description of variation in abundance among species. Uncertainties in his variance estimates were conditioned on species abundances following this pattern, so that the only source of uncertainty was the Poisson sampling.

In 1950 F.J. Anscombe explored the estimation problem using the multinomial distribution for the number of species with different number of individuals in the sample. Hence, in contrast to Fisher, Anscombe also included the stochastic variation among species in the underlying abundances. Quite remarkably, he found that the number individuals and species observed were jointly a sufficient statistic for estimating Fisher's α, the estimation equation for α being $S/\alpha = \ln[(N + \alpha)/\alpha]$ (exercise 15), which is also the estimation equation originally proposed by Fisher. That is, if the observed individual and species number are known, then no additional information is contained in the observed abundances, indicating a kind of 'neutralitiy' in Fisher's model. The parameter w in the logarithmic series $\alpha w^n/n$ given by $w = \nu/(\nu+\rho)$ can

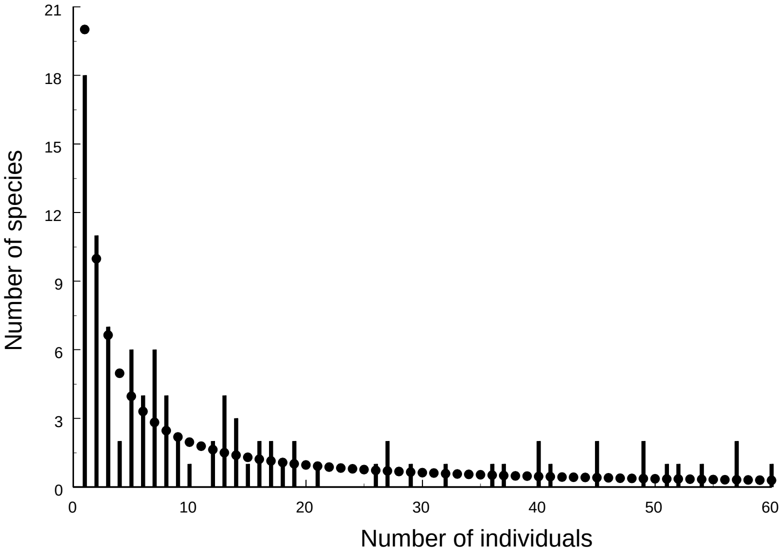

Figure 7.2: The data set shown in Fig.7.1 (bars) together with the fitted logarithmic series model (filled circles). The estimates obtained by Fisher's method are $\hat{\alpha} = 20.05$ and $\hat{w} = 0.99831$.

be estimated by observing that the expected abundance of a random species is $\mathrm{E}(\nu y) = \nu k/\rho$, which means that the expected number of individuals is $s\nu k/\rho$ approaching $\nu\alpha/\rho$ under Fisher's limit. Estimating this quantity by the observed total number of individuals N is equivalent to estimating ν/ρ by N/α and consequently estimating w by $N/(N+\alpha)$, which also was the estimator proposed by Fisher. Fig.7.2 shows the same data as in Fig.7.1 together with the estimated log series sequence of expectations $\mathrm{E}(z_n)$.

Sampling and statistical inference in the neutral model of Crow and Kimura were analyzed by Ewens as well as by Karlin and McGregor in 1972. Their approach is somewhat more complex than Anscombe's approach to Fisher's model. Actually, Ewens considered a sample size of n individuals as given,

implying that the multivariate distribution of counts for different alleles are dependent. Nevertheless, Ewens found in agreement with Anscombe that the number of different alleles in the sample is a sufficient statistic for the estimation of the parameter $M = 4N_e u$ corresponding to Fisher's α.
The exact multivatiate distribution of observed numbers for all alleles in the sample, usually called Ewens sampling formula, was conjected by Ewens and proved by Karlin and McGregor in 1972. The probability of having $S = \sum z_j$ different alleles in the sample with $z_1, z_2 \ldots, z_n$ representatives was found to be

$$P(S, z_1, z_2 \ldots z_n) = \frac{n! M^S}{1^{z_1} 2^{z_2} \ldots n^{z_n} z_1! z_2! \ldots z_n! \prod_{i=1}^{n} (M + i - 1)}, \tag{7.1}$$

where $n = \sum j z_j$ is the number of alleles sampled. Accordingly, the maximum likelihood estimate of M is the solution of the equation (exercise 16)

$$S/M = \sum_{i=1}^{N} (M + i - 1)^{-1}. \tag{7.2}$$

Since the sum on the left side is approximately $\ln[(N+M)/M]$ (exercise 17), the estimator is approximately the same as the estimator found by Fisher and Anscombe, the difference being that Ewens conditioned on n.

7.4.4 Hubble's neutral model

The neutral community model of Hubbell is a meta-community model, dealing with the species abundance distribution in the total community as well as in local communities. The basic idea is that the dynamics of species abundances are described exactly as in the neutral infinite allele model. What in genetics is 'genetic drift' is in community ecology called 'ecological drift', the stochasticity in change in species abundance being purely demographic. Since the meta-community is assumed to be very large, changes due to demographic stochasticity will necessarily be very slow. Local communities, however, are smaller and consequently have much faster changes. This makes it realistic to consider species abundances being constant in the meta-community under investigation of species dynamics in local communities. This assumption is realistic when the local community is only a small fraction of the meta-community.

In addition to assuming neutral dynamics in the meta-community, there is an assumption of a connection to the local community by migration. Hubbell has described this system starting with Ewen's sampling formula as model for the total community. However, as Hubbell points out, when the meta-community is large, we can as well just use the frequency spectrum of Fisher's model assuming that the expected number of species with relative abundance in $(q, q + dq)$ is $f(q)dq = \alpha q^{-1}(1 - q)^{\alpha-1}$. The parameter α corresponds to $4N_e u$ in the neutral genetic model with $2N$ alleles, where N_e is the effective population size and u is the mutation rate. When applying this to ecological drift using Fisher's model, Fisher's diversity parameter is accordingly given by

$$\alpha = 2Nu/\sigma_d^2$$

where N is the community size, u is the speciation rate, and σ_d^2 is a common demographic variance for the species in the community. Hence, the effective population size in the genetic model is replaced by N/σ_d^2 in accordance with our treatment of genetic drift in chapter 5. The ecological drift of species frequencies in the meta-community is then given by the infinitesimal variance $q(1 - q)\sigma_d^2/N$ which is approximately zero for a large N.

Now, following the idea of Hubbell, but using diffusion and continuously distributed abundances, we consider a local community of size n much small than N. The ecological drift of a relative abundance p in this community is defined by the infinitesimal variance $\nu(p) = p(1 - p)\sigma_d^2/n$. The simplest approach is to consider this as just a single island that is so small that it does not affect the large mainland community. Below we also consider the situation where the meta-community consists of k such local communities. In that case we do not automatically obtain a consistent model unless we adjust the parameter α that will then be determined by the dynamics in all local communities.

Let us first consider the local dynamics of a species with relative abundance q in the meta-community and p in the local one. The expected fraction of individuals migrating from the island is mp, where m is a migration rate assumed to have the same value for all islands and all species. The total expected number of individuals migrating within the meta-community is mNq, and a fraction n/N of these is assumed to migrate to the island giving a mean increase mq in the relative abundance on the island. Consequently, if we assume that the speciation rate u is much smaller than the migration rate m, we may ignore the speciation at the island, giving the infinitesimal mean for the relative abundance process $\mu(p) = -m(p-q)$. Also assuming that the stochastic effect of the migration is much smaller than that of the ecological drift we can approximate the variance by the above infinitesimal variance $\nu(p)$. The diffusion for p then have the stationary beta-distribution (exercise 18)

$$g(p;q) = \frac{\Gamma[\tau]}{\Gamma[\tau q]\Gamma[\tau(1-q)]} p^{\tau q-1}(1-p)^{\tau(1-q)-1}$$

where $\tau = 2mn/\sigma_d^2$, with mean value q. This leads to the joint frequency spectrum for p and q being $g(p;q)f(q)$. The interpretation of this spectrum is that the expected number of species with relative abundance in the meta-community in $(q, q+dq)$ and in the local community in $(p, p+dp)$ is $g(p;q)f(q)dpdq$.

The frequency spectrum for the local community is then

$$f_L(p) = \int g(p;q)f(q)dq. \tag{7.3}$$

Notice that we have not used the form of the function $f(q)$ in this derivation of the local spectrum $f_L(p)$, so this formula, transforming the meta-community spectrum to the local one, is general and can be applied to any species abundance distribution for the meta-community. Although we have used an approach that is somewhat different from Hubbell's, utilizing diffusion theory, equation (7.3) contains one major finding of Hubbell, that $f(q)$

and $f_L(p)$ are different distributions. Hubbell found that when $f(q)$ is chosen as Fisher's model describing a neutral community, the species abundance distribution for the local community would tend to look more like Preston's lognormal distribution. This is depicted in Fig.7.3 where we have illustrated how the frequency spectrum on the log scale $x = \ln p$, which is $e^x f_L(e^x)$, is affected by isolation of the local community. For large migration rates m the spectrum is indistinguishable from the Fisher model describing the meta-community, that is, $f_L(p) \approx f(p)$. As the migration rate decreases, the curves tend to look more like Preston's lognormal model, that is, like a Gaussian curve on the log scale. The smallest log frequency shown corresponds to about three individuals of a species in the local community and the areas under the curves are therefore approximately the expected number of species in the local communities.

Now consider the case where the meta-population consists of large number of k islands of size n, so that $N = kn$. We also make the assumption of purely long-distance random migration in the sense that a migrating individual is equally likely to end up in any local community. Then, in order to obtain a consistent model we must ensure that the ecological drift in the total community actually is the drift generated by the assembly of all islands. Since the meta-population abundance is the mean of local abundances, $q = \bar{p}$, the variance of Δq must be

$$\mathrm{var}(\Delta q) = \mathrm{var}(\bar{\Delta} p) = k^{-2} \sum \mathrm{var}(\Delta p_i) \approx \frac{\sigma_d^2}{nk} \mathrm{E}[p(1-p].$$

Using the beta-distribution of p at islands we find that $\mathrm{E}[p(1-p)] = q(1-q)\tau/(\tau+1)$ (exercise 18). Hence, the parameter in Fisher's model for the meta-community must be

$$\alpha = \frac{2Nu(\tau+1)}{\sigma_d^2 \tau}.$$

From equation (7.3) we can also study how the expected diversity of the local community is affected by local community size and migration, an effect determined by the parameter $\tau = 2mn/\sigma_d^2$. It follows from the above result

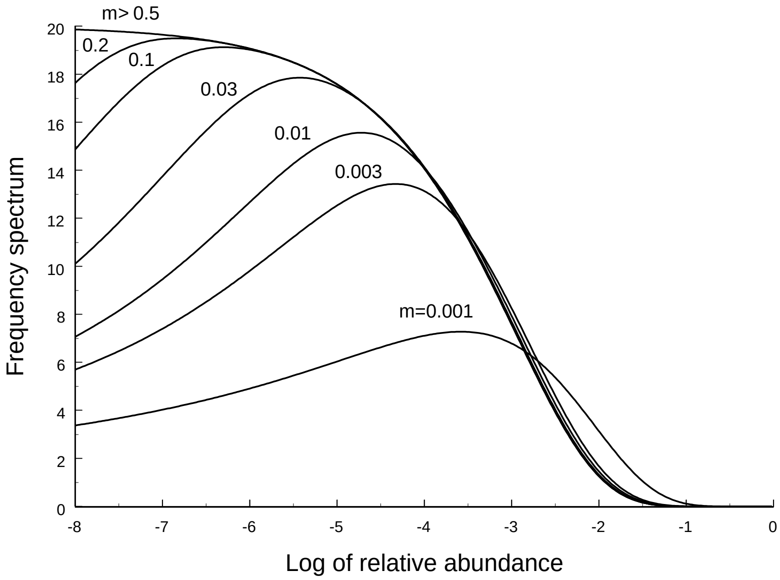

Figure 7.3: The frequency spectrum on the log scale $e^x f_L(e^x)$ for island under varying migration rates m. The other parameters are $N = 10^7$, $n = 10^4$, $u = 10^{-6}$, and $\sigma_d^2 = 1$. Under large migration $\alpha = 20$. For small migration α is slightly adjusted to give consistency. For $m = 0.001$ we have $\tau = 20$ and $\alpha = 19.05$. The area under the curves is approximately the expected number of species in the local community.

for $\mathrm{E}p(1-p)$ that Simpson's index for the local community is $\mathrm{E}(H_{S,L}) = \frac{\tau}{\tau+1}\mathrm{E}(H_S)$, where H_S is the index for the meta-community (exercise 19). So, with the parameters used in Fig.7.3 the smallest migration rate shown ($m = 0.001$) only makes the index about 5% smaller than its value in the meta-community. This is because Simpson's index is a dominance index and the isolation of the island mainly affects the large number of rare species and have little effect on those dominating the community.

Using equation (7.3) we can also find the expected value of the information index in the local community (exercise 21)

$$\mathrm{E}(H_{I,L}) = \int [\Psi(\tau+1) - \Psi(\tau q+1)]\alpha(1-q)^{\alpha-1}dq$$

which is shown in Fig.7.4. We see that this index, which is more sensitive to species number, is rather strongly affected by isolation compared to the effect on Simpson's index.

To study the same kind of problem for islands of different size or for more realistic migration pattern is very complex. For example, one will need to specify how the migration parameter depends on the size of the island. For small islands individuals are more likely to migrate to a location outside than for a very large island, where a migration very well may end up inside the border of the island.

7.5 Independent species dynamics

Neutral models describe community dynamics through several rather strong assumptions. First, all growth rates are assumed to be zero as there is no differences between species and the total community size is kept constant. This means that all temporal changes in species abundances are purely stochastic. The growth rates are zero regardless species abundance, that is, there is no density-regulation. Furthermore, the stochasticity driving species abundance fluctuations is identical for all species. Another important assumption is that the dynamics is driven by demographic stochasticity only.

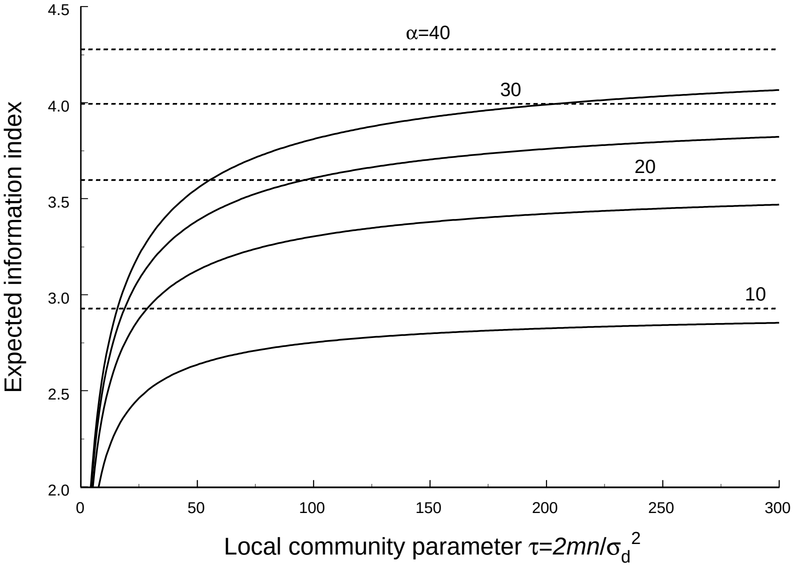

Figure 7.4: The expected value of the information index on islands as function of the parameter $\tau = 2mn/\sigma_d^2$ (solid lines) for different values of Fisher's diversity index α for the meta-community. The dashed lines show the information index for the meta-communities.

Although some of these assumptions may be realistic for some communities such as tropical trees, most communities have species known to be regulated by their own density. Also, estimation of population dynamical parameters has shown that even rather similar species may show differences in dynamics, with different growth rates at small densities as well as different carrying capacities. Most studies of communities using species abundance distributions deal with communities with large species richness. These large communities also have a very large number of individuals. We have seen in chapter 1 that stochastic fluctuations then mainly are driven by the environment so that even very large populations can show fluctuations that are large relative to its size. For example, a population of $N = 10^6$ individuals and demographic variance $\sigma_d^2 = 1$ will after a period of 100 years only have a variance in population change during that period of approximately $100N\sigma_d^2 = 10^8$, which means that the standard deviation is only 1% relative to its initial size. It is highly unlikely that populations show that kind of stability in size over such a long period, so neutral theory is likely to underestimate the changes in the abundant species. Environmental stochasticity, however, is known to have the same relative effect on population fluctuations for all population sizes. If we assume that the above population has environmental variance $\sigma_e^2 = 0.01$, the variance in the change during the same period is $100N^2\sigma_e^2 = 10^{12}$, so that the standard deviation is 100% relative to initial population size.

In the following sections we present some alternative descriptions of dynamical species abundance models including environmental stochasticity and density-dependence. These are, however, also based on a number of simplifying assumptions, such as for example independent dynamics of the species in the community. We first consider homogeneous models, that is, the species are interchangeable in the model, all being described by the same parameter value. In 7.7 we generalize this approach to heterogeneous communities where the parameters are considered to have a given random variation among invading species.

The assumption of independent dynamics may at first seem rather unrealistic. However, a strong common density regulating effect of the total community size ensuring small fluctuations in the overall density of individuals tends to wipe out common noise terms and make species dynamics approximately independent. Let us first illustrate this by an analogy. Consider the diffusion approximation to the Gompertz type of population model for a single species with environmental and demographic variance, $dN = rN(1 - \ln N/\ln K)dt + \sigma_e N dB_e + \sigma_d \sum dB_i$, where dB_e and the dB_i are independent Brownian motions and the sum runs over all individuals. Choosing $K = 1000$, $\sigma_d^2 = 1$ and $\sigma_e^2 = 0.01$, a large $r = \gamma \ln K$ will correspond to strong density regulation given by γ and the population will show small fluctuations around K. These fluctuations can be studied by ignoring σ_d^2 because $\sigma_e^2 K^2$ is much larger than the demographic contribution $\sigma_d^2 K$ for N close to K. One way of looking at these population fluctuations with a small stationary variance $\sigma_e^2/(2\gamma)$ on the log scale, is to realize that the large fluctuations generated by σ_e^2 are immediately canceled by the large density regulation. However, considering the contributions to the next generation from two different species, their demographic components are still independent and much larger than their environmental components. As the demographic components altogether have practically no effect on the stochastic change in population size, the density regulation does not affect these components in the same way as the environmental components are affected, and they remain approximately independent as in the case of no density regulation.

Returning to communities the analogy of species and individual in the above example is now community and species. The analogy of environmental variance affecting all species is a common environmental noise term affecting all species, while the analogy of demographic noise is environmental noise terms that are independent among species. To illustrate our conjecture on approximately independent dynamics we consider a community with large species richness. It is then unlikely that the abundance of one single species will have a major effect on another species. However, all the other species together are likely to affect a single species, adding up the possible small effect of each of them. Let s be the number of species and N_i the individual numbers, $i = 1, 2 \ldots, s$, so that the total community size is $N = \sum N_i$. Let $\mathbf{X} = (X_1, X_2, \ldots, X_s)$ be the vector of log abundances $X_i = \ln N_i$. As an illustration consider the model on the form $\mathrm{E}(\Delta X_i|\mathbf{X}) = r_i - \gamma X_i - D(\mathbf{X})$, that is, species have different growth rates but are density regulated by its own density and the total community vector. Adding noise we write this as a continuous model

$$dX_i = r_i - \gamma X_i - D(\mathbf{X}) + \sigma_E dA + \sigma_e dB_i$$

where dA and the dB_i are independent Brownian motions. The parameter σ_E^2 is an environmental effect common to all species, while σ_e^2 denotes the magnitude of environmental effects that are independent among species. The mean value $\bar{X}$ then have the dynamics

$$d\bar{X} = \bar{r} - \gamma \bar{X} - D(\bar{X}) + \sigma_E dA + \sigma_e d\bar{B}$$

while the dynamics of the deviations $Y_i = X_i - \bar{X}$ is

$$dY_i = r_i - \bar{r} - \gamma Y_i + \sigma_e(dB_i - d\bar{B}).$$

We now make the assumption that the mean value is strongly density regulated through the term $D(\mathbf{X})$. A consequence of this is that $\bar{X}$ only show small fluctuations. Approximating $\bar{X}$ by its mean $\bar{\mu}$ and inserting $Y_i = X_i - \bar{\mu}$ in the above equation then gives the approximation

$$dX_i = (r_i - \bar{r} + \gamma\bar{\mu}) - \gamma X_i + \sigma_e(dB_i - d\bar{B}).$$

For a species rich community the terms $d\bar{B}$ has variance $s^{-1}dt$ and can therefore be ignored compared to dB_i having variance dt. Finally, redefining r_i as the previous $r_i - \bar{r} + \gamma\bar{\mu}$, the dynamic equations are

$$dX_i = r_i - \gamma X_i + \sigma_e dB_i.$$

Comparing this to the initial equation for the dynamics of X_i we see that our assumption of a strong density regulation for the mean $\bar{X}$ had the effect that $D(\mathbf{X}) \approx \sigma_E dA$. This can be understood intuitively as follows: The single species noise terms dB_i are independent and will, by the law of large numbers, have little effect on $d\bar{X}$ compared to the common term dA. The assumption of a strong density dependence for the mean ensures that the mean changes little. Consequently, the common environmental effect of $\sigma_E dA$ is canceled by the density regulation given by $D(\mathbf{X})$. As a consequence, for strong density dependence of the total community, the assumption of independent dynamics of species driven by the environment may be rather realistic. The environmental noise term, however, is not the total environmental effect on the species, but only the species specific components that are independent among species.

We have seen that neutral theory was inspired by the theory for neutral genetic drift. The above model, where changes are driven by environmental stochasticity affecting all individuals of a species in the same way, can in a similar way be compared to the theory of fluctuating selection. Temporal fluctuations in selection coefficients affect all individuals with a given genotype, corresponding to the environmental effect on species.

7.6 Homogeneous community models

7.6.1 Colonizations and extinctions

In chapter 6 we have used the two-dimensional inhomogeneous Poisson process to model spatial distribution of individuals. The process was defined by its intensity, which could be interpreted as the density of points. The Poisson process is characterized by the fact that the number of points in two disjoint areas are independent Poisson variates with mean values equal the integral of the intensity function over the area.

Here we shall use the same inhomogeneous Poisson process, defined on the real line rather than in space, in two different ways. First, the process will be applied to describe the colonization or speciation process, points on the line representing new species entering the community. Secondly, we use the same type of process to define species abundance models at a given time. Then, points on the positive real axis generated by an inhomogeneous Poisson process represent the species abundances in the community.

We consider a very large community without interactions with its surroundings so that all new species are actually generated by an evolutionary speciation process. We assume that new species appear in time according to an inhomogeneous Poisson process with intensity $\omega(t)$. Hence, speciations occurring in non-overlapping time segments are independent events, and the probability of a new species appearing in $(t, t+dt)$ is $\omega(t)dt$. Assuming a constant intensity ω may be realistic, but it is not required for showing that abundances follow an inhomogeneous Poisson process.

Let us consider the distribution of species abundances at time $t = 0$. Let Ω_1 and Ω_2 be two disjoint intervals and let $P_i(t)$ be the probabilities that a species invading at time $-t$ has abundance in Ω_i, $i = 1, 2$, at time zero. Let X_i be the number of species that invaded in $(-t, -t+dt)$ with abundance in Ω_i at $t = 0$ and write $q(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$. Then $q(1, 0) = P_1(t)\omega(-t)dt$, $q(0, 1) = P_2(t)\omega(-t)dt$ and $q(0, 0) = 1 - P_1(t)\omega(-1)dt - P_2(t)\omega(-t)dt$. The joint moment generating function of (X_1, X_2) is accordingly

$$M_t(u_1, u_2) = \mathrm{E}(e^{u_1X_1+u_2X_2}) = 1 + \omega(-t)dt\sum_{i=1}^{2} P_i(e^{u_i} - 1)$$

and the cumulant generating function is

$$K_t(u_1, u_2) = \omega(-t)dt\sum_{i=1}^{2} P_i(e^{u_i} - 1).$$

Finally, adding the contributions from all past time intervals $(-t, -t+dt)$, and assuming that the species in the community have independent dynamics, yields the cumulant generating function for the number of species with abundance in Ω_1 and Ω_2 at time zero as

$$K(u_1, u_2) = \sum_{i=1}^{2} (e^{u_i} - 1) \int_0^\infty \omega(-t) P_i(t) dt.$$

This is the joint cumulant generating function for two independent Poisson variates with mean values $\int_0^\infty \omega(-t) P_i(t) dt$. Accordingly, as our two intervals were arbitrary and disjoint, the abundances at time zero and accordingly at any time, must follow an inhomogeneous Poisson process. This theory covers the case of continuous species abundances, as must be assumed when adopting diffusion theory, as well as the discrete case when abundances for example are the individual counts $0, 1, 2 ...$. In the last case the number of species with abundance equal any non-negative integer is Poisson distributed. We see that this theory leads to the assumption of independent Poisson variates used by Anscombe to find estimators for the parameters in Fisher's model, as outlined in exercise 15.

7.6.2 Homogeneous diffusion models

We now assume that species enter the community with abundance $x_0 = a + \delta$ (for example 2) and go extinct at abundance a (for example 1) and that their dynamics are defined by a diffusion process with infinitesimal mean and variance $\mu(x)$ and $\nu(x)$ respectively.

We further assume a constant rate ω of new species and let each species have the dynamics defined so that the population size fluctuates around some quasi-stationary equilibrium and finally goes extinct. The community will then eventually reach stationarity with long term balance between speciations and extinctions. It follows from the derivation in the previous section that the species abundances at any time then follow an inhomogeneous Poisson process. This kind of dynamics is illustrated schematically in Fig.7.5.

It remains to derive the rate $\lambda(x)$ of the inhomogeneous Poisson process. We have seen from our general derivation that the number of species with abundance in the interval Ω is Poisson distributed with mean $\int_\Omega \lambda(x) dx =$

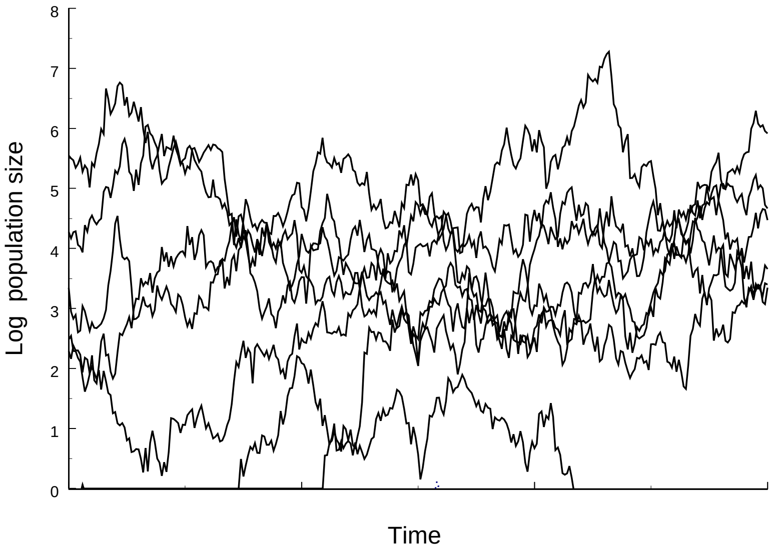

Figure 7.5: Schematic presentation of log population trajectories with speciations and extinctions. The simulated model is homogeneous and the time scale is arbitrarty.

$\omega \int_0^\infty P(t)dt$ where $P(t)$ is the probability that a species entering at time zero has abundance in Ω at time $t > 0$. Now, $P(t)dt$ is the expected time this species has abundance in Ω during the time interval $(t, t + dt)$. Hence $\int_0^\infty P(t)dt$ is the expected total time the species has abundance in Ω which, in a diffusion model, is the integral of the Green function over Ω. Hence

$$\int_\Omega \lambda(x)dx = \omega \int_\Omega G(x, x_0)dx.$$

Since the integrands are non-negative and the equivalence holds for any interval Ω we can conclude that

$$\lambda(x) = \omega G(x, x_0).$$

The inhomogeneous Poisson process has the property (exercise 22) that if the number of points in Ω is N, then the location of the N points conditioned on $N = n$ are n independent observations from the distribution $\lambda(x)/\int_\Omega \lambda(u)du$. Choosing Ω as the interval (a, ∞), we see that the distribution of species abundance in the community given the number of species at a given time is

$$f(x) = G(x, x_0)/T$$

where $T = \int_{x_0}^\infty G(x, x_0)dx$ is the expected time a species stays in the community. Hence, in this model the species abundance distribution $f(x)$ is simply the quasi-stationary distribution of species abundance.

For an extinction barrier at a an initial value $x_0 = a + \delta$ we have for a small δ that $S(x_0) = S(a) + s(a)\delta = \delta$ when the lower integration limits are chosen at a and $S(x)$ and $s(x)$ are the functions used to express the Green function in chapter 3. This gives in the limit as $\omega \to \infty$ and $\delta \to 0$ so that $\omega\delta \to \omega_0$

$$\lambda(x) = \frac{2\omega_0}{\nu(x)s(x)}. \tag{7.4}$$

If δ approaches zero and ω is kept finite we see that all paths goes to extinction immediately since $\lambda(x)$ then also approaches zero. This is due to the fact the diffusion paths show large local fluctuations over short time steps, so with initial value very close to the extinction barrier practically all paths will go extinct immediately. If the speciation rate is defined as the rate at which new species entering at abundance $a+\delta$ reach abundance $a+\varepsilon > a+\delta$, then the speciation rate is in the limit as $\delta \to 0$ is $\omega\delta/\varepsilon$ (exercise 23). This means that ω_0 is approximately the speciation rate for $\varepsilon = 1$, that is, for example the rate at which new species reach abundance 2 when the extinction barrier is at $a = 1$.

Inserting the definition of the function $s(x)$ we find

$$\lambda(x) = \frac{2\omega_0}{\nu(x)} e^{2\int_a^x \mu(u)/\nu(u)du}.$$

We see that the shape of the species abundance distribution for this class of models strongly depends on the form of the density-regulation defined by the infinitesimal mean.

When dealing with distribution of relative abundance among species we have previously used the term 'frequency spectrum' for the function $f(p)$ defined so that $f(p)dp$ is the expected number of species with relative abundance in $(p, p+dp)$. Since $\lambda(x)dx$ is the expected number of species with abundance in $(x, x+dx)$, it could naturally be called the 'abundance spectrum'. We shall occasionally use this term together with the term 'Poisson rate' which is more generally used for inhomogeneous Poisson processes.

The homogeneous gamma model

Consider now the model with logistic density regulation and constant demographic and environmental variance, that is, $\mu(x) = rx - \beta x^2$ and $\nu(x) = \sigma_e^2 x^2 + \sigma_d^2 x$, and let $a = 1$. The carrying capacities in the corresponding deterministic model are then $K = r/\beta$. Equation 7.4 then leads to the Poisson rate

$$\lambda(x) = bx^{-1}(x+c)^{2r(1+c\beta/r)/\sigma_e^2 - 1} e^{-2\beta x/\sigma_e^2}, \tag{7.5}$$

where

$$b = \frac{2\omega_0 e^{2\beta/\sigma_e^2}}{\sigma_e^2 (1+c)^{2r(1+c\beta/r)/\sigma_e^2}}$$

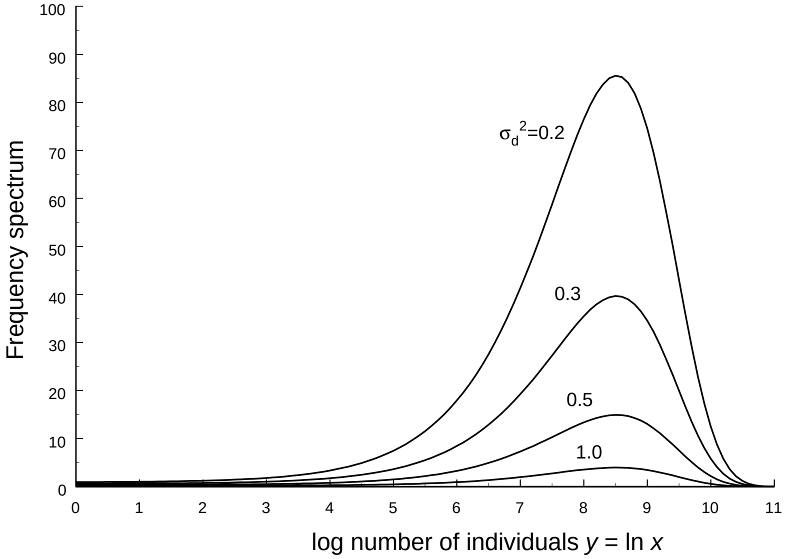

Figure 7.6: The frequency spectrum for log number of individuals, that is $e^y\lambda(e^y)$ as function of log individual number (log abundance) $y = \ln x$ for different values of σ_d^2 for the homogeneous logistic model with environmental and demographic stochasticity. The other parameters are $\omega_0 = 0.1$, $r = 0.01$, $\sigma_e^2 = 0.01$, and $\beta = 10^{-6}$, giving $K = 10^4$.

and $c = \sigma_d^2/\sigma_e^2$. Fig.7.6 shows this abundance spectrum on the log scale, that is $e^y\lambda(e^y)$ as function of $y = \ln x$, for different values of the demographic variance σ_d^2. The curves are skew to the right as the curves for the local community in the neutral model. The area under the curves represents the expected number of species. We see that an increase in demographic variance dramatically reduces the species richness.

In view of Fig.7.6 it may seem very wrong to ignore demographic variance. However, if we are dealing with very large communities all species that appear in the samples, even if only by a single individual, will have a large number of individuals in the community. However, the speciation and extinction rates has to be adjusted in order to include the effect of demographic variance. First, consider the model without such an adjustment, choosing $a = 1$. With $\sigma_d^2 = 0$ the Poisson rate is then proportional to the gamma distribution

$$\lambda(x) = \frac{2\omega_0 e^{2\beta/\sigma_e^2}}{\sigma_e^2} x^{2s/\sigma_e^2-1} e^{-2\beta x/\sigma_e^2},$$

where $s = r - \sigma_e^2/2$ is the stochastic growth rate of each species at small densities. Hence, the species abundance distribution is the gamma distribution with shape parameter $k = 2s/\sigma_e^2$ and scale parameter $2\beta/\sigma_e^2$ with mean approximately $Ks/r = s/\beta$, censored at the smallest possible abundance 1. This Poisson rate is approximately proportional to the rate in equation 7.5 for large abundances, that is, for x much larger than $c = \sigma_d^2/\sigma_e^2$. So, if we want to use the gamma model as an approximation, ignoring species that are very rare and unlikely to be sampled, the rate should be adjusted by a factor so that the two rates are approximately equal for large abundances. In other words, we approximate the rate given by equation 7.5 by a gamma model, which is a good approximation for species that are not very rare. Using the assumption of a very large community (K large) so that $\beta = s/K$ is small, the Poisson rate becomes approximately

$$\lambda(x) = \frac{2\omega_0}{\sigma_e^2(1 + \sigma_d^2/\sigma_e^2)^{2r/\sigma_e^2}} x^{2s/\sigma_e^2-1} e^{-2\beta x/\sigma_e^2}. \tag{7.6}$$

Remembering that the gamma-distribution of species abundances corresponding to this abundance spectrum was used by Fisher to derive his logarithmic series distribution, we see that we actually get Fisher's model for $s = 0$. Accordingly, Fisher's abundance model for large communities may also be generated by a model with environmental and demographic stochasticity in contrast to the infinite allele model with only demographic stochasticity.

The expected total abundance of all species in the community is $\mathrm{E}\sum x \approx \int x\lambda(x)dx$ which is approximately $(1 + \sigma_d^2/\sigma_e^2)^{-1}\omega_0/\beta$ for $s = 0$, provided that we can approximate the integral from 1 by a lower integration limit at zero. This is a valid approximation if β is small. For $s = 0$ we must then require that β is much smaller than $\sigma_e^2/2 = r$, which means that K is much larger than 1, which is an assumption we have already made. Defining rescaled abundances $x(1+\sigma_d^2/\sigma_e^2)\beta/\omega_0$, the expected sum of these abundances are one. Then, we have seen before that the scale parameter in the gamma-distribution is actually Fisher's α, that is

$$\alpha = \frac{\omega_0}{(1 + \sigma_d^2/\sigma_e^2)\beta}\frac{2\beta}{\sigma_e^2} = \frac{2\omega_0}{\sigma_e^2 + \sigma_d^2} \approx \frac{2\omega_0}{\sigma_d^2}.$$

This result also has common features with the result $\alpha = 2Nu/\sigma_d^2$ in the ecological application of the infinite allele model. Fisher's parameter is proportional to the speciation rate in both cases, and inversely proportional to the demographic variance. A major difference, however, is that α depends strongly on the community size N in the infinite allele model, and will approach infinity as N increases. In the homogeneous model with environmental and demographic stochasticity, however, there is no dependence on the community size, except that we have assumed that the carrying capacities are large. Although β determines the expected size of the community, it only appears in the scale parameter of the gamma model and will therefore have no effect on the frequency spectrum for relative abundances defining Fisher's α.

The two models are very different when it comes to estimating speciation rates. Consider a community where α has been estimated from data. The speciation rate in the homogeneous model then becomes $\omega_0 = \alpha\sigma_d^2/2$ and for the neutral model $u = \alpha\sigma_d^2/(2N)$. Hence, for extremely large communities (large N) such as communities of tropical butterflies or beetles, the neutral model leads to an extremely small speciation rate.

Choosing $\alpha = 20$ and $\sigma_d^2 = 1$ we find from the homogeneous model that $\omega_0 = 10$, which means that 10 new species enter the community each year. This may seem a very large rate, but most of these species are likely to go extinct very soon and have abundances so small that they are never seen in practice. If we for example require that a species should have n representatives in the community before we consider it as established, the rate will be $\alpha\sigma_d^2/(2n)$ (exercise 24), so in the above numerical example with n chosen as 100 there are on average 10 years between speciations, while the event that new species reach the abundance of 1000 only occur once in 100 year. This also underlines the importance of a precise definition of speciation. This result, that a large number of new species formed will never reach abundance above some few individuals before they almost immediately go extinct, has an interesting analogy in genetics that will be treated in some detail in chapter 8. Using a rather simple model C.B.S. Haldane showed in 1927 that new beneficial mutations with a small advantage $s > 0$ (multiplicative growth rate $1 + s$) will reach fixation in the population with probability approximately $2s$. That is, a fraction $1 - 2s$ of the mutations will go extinct, and most of them will go extinct after only a few generations. Therefore the rate of fixation of beneficial genes will be much smaller than the rate of mutation. This is an analogy to the above result that most new species, with no or only a small advantage relative to the rest of the community, will go extinct during some few generations.

Now let us assume Poisson sampling, that is, the number of individuals representing a species with abundance x in the sample is Poisson distributed with mean νx, where ν is a measure of sampling intensity. The numbers z_j, $j = 1, 2, \ldots$ of species with j representatives in the sample are the independent Poisson variates (exercise 25) with means proportional to the terms of

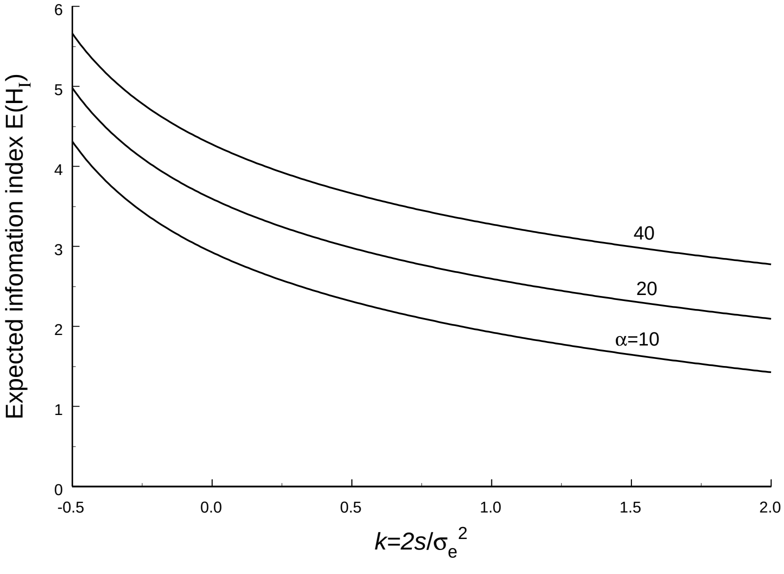

Figure 7.7: The expected information index $\mathrm{E}(H_I)$ in the gamma model as function of the shape parameter $k = 2s/\sigma_e^2$ for different values of $\alpha = 2\omega_0/\sigma_d^2$.

a negative binomial distribution. Accordingly, the observed abundances is a sample from the truncated negative binomial distribution (exercise 26)

$$p_j = \frac{k(1-w)^k}{1-(1-w)^k}\frac{\Gamma(k+n)}{\Gamma(k+1)j!}w^j, \quad j = 1, 2, \ldots \tag{7.7}$$

where $k = 2s/\sigma_e^2$ and $w = \nu/(\nu + 2\beta/\sigma_e^2)$. Notice that this is a well defined distribution also for negative shape parameters $k > -1$. This means that species may invade and lead to a stationary model even if they have negative stochastic growth rates. However, $k = 2s/\sigma_e^2$ must be larger than -1, that is, the corresponding deterministic growth rate $r = s + \sigma_e^2/2$ must be positive. The parameters k and w can be estimated by maximum likelihood using the above zero-truncated negative binomial distribution.

For this extended negative binomial species abundance distribution with $k > -1$, including Fisher's model with $k = 0$, one can show that the frequency spectrum is

$$f(p) = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k)} p^k (1-p)^{\alpha-k-1}.$$

leading to the expected value of Simpson's index (exercise 27)

$$\mathrm{E}(H_S) = 1 - \frac{k+1}{\alpha+1}$$

and the expectation of the information index shown in Fig.7.7 (exercise28)

$$\mathrm{E}(H_I) = \Psi(\alpha+1) - \Psi(k+1).$$

The homogeneous lognormal species abundance model

It is known from chapter 3 that the Ornstein-Uhlenbeck process has normal stationary distribution and by a transformation it is equivalent the model with Gompertz type of density-regulation and constant environmental variance having the lognormal stationary distribution. This indicates that the Gomperz type of density regulation will lead to the lognormal species abundance distribution. For this model it is preferable to work on the log scale $y = \ln x$, using infinitesimal mean and variance $\mu(y) = s - \gamma y$ and $\nu(y) = \sigma_e^2$, where $s = r - \sigma_e^2/2$ is the stochastic growth rate at small densities. Choosing the extinction barrier at $a = 0$ on the log scale we find from equation 7.4

$$\lambda(y) = \frac{2\omega_0}{\sigma_e^2} e^{s^2/(\gamma\sigma_e^2)} e^{-\frac{1}{2}\frac{(y-s/\gamma)^2}{\sigma_e^2/(2\gamma)}} \tag{7.8}$$

which is proportional to the normal density truncated at $y = 0$. Hence, the corresponding species abundance distribution is the lognormal distribution truncated at $x = 1$. From this we find the expected number of species

$$\mathrm{E}S = \int_0^\infty \lambda(x)dx = \frac{2\omega_0}{\sigma_e^2} e^{s^2/(\gamma\sigma_e^2)} \sqrt{\pi\sigma_e^2/\gamma}\,\Phi(s\sqrt{2/\gamma}/\sigma_e)$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution.

As for the gamma model the effect of introducing a demographic variance can be investigated by choosing $\nu(x;\sigma_d^2) = \sigma_e^2 + \sigma_d^2 e^{-x}$. The Poisson rate can then be evaluated by numerical integration. In Fig.7.8 this rate is shown for three different values of σ_d^2 for a model with weak density regulation. Notice that the curves are close to the normal distribution even for $\sigma_d^2 = 1$, the effect of increasing σ_d^2 being mainly a decrease in the number of species in the community. Since the speciation rate ω_0 is a factor in $\lambda(x)$ the number of species is also proportional to this parameter. This means that it will be hard in practice to separate the effects of σ_d^2 and ω_0. This also indicates that it may be a useful approach to work with the model with $\sigma_d^2 = 0$ that gives exactly the (truncated) normal distribution of log abundances, as long as we realize that it is impossible to estimate ω_0 from that model. Actually, ω_0 needs to be chosen very small in order to compensate for the lack of demographic stochasticity that would have driven many rare species to extinction.

Provided that practically the whole normal distribution of log abundance is above zero so that the truncation at zero can be ignored, the mean and variance can be estimated by maximum likelihood under the assumption of Poisson sampling using the Poisson lognormal distribution. This distribution is in general terms defined as

$$p_i = \mathrm{E}\left(\frac{Q^i}{i!}e^{-Q}\right),$$

for $i = 0, 1, ...$, where $\ln Q$ is normally distributed with mean μ and variance σ^2, or

$$p_i(\mu, \sigma^2) = \int_{-\infty}^{\infty} \frac{e^{(\mu+\sigma u)i}}{i!} e^{-e^{\mu+\sigma u}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du,$$

where $u = (\ln Q - \mu)/\sigma$ is the standardization of $\ln Q$. This is the Poisson-lognormal distribution with parameters (μ, σ^2). The special case $\sigma^2 = 0$ is simply the Poisson distribution with mean e^μ.

Over-dispersion in the sampling relative to the Poisson distribution can be modeled by introducing a variable V ensuring that $\ln V$ is normally distributed with mean $-\theta^2/2$ and variance θ^2 so that $\mathrm{E}V = 1$. The values of

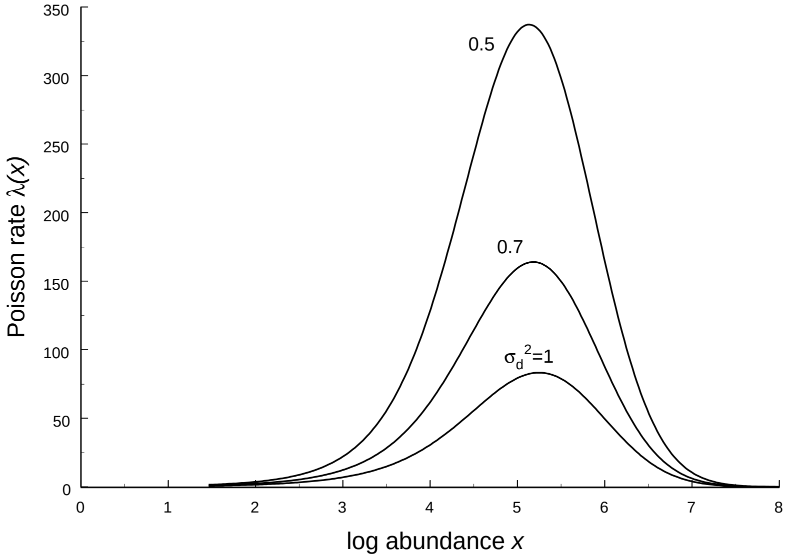

Figure 7.8: The Poisson rate on the log scale for the homogeneous model with the Gompertz type of density regulation for different values of the environmental variance σ_d^2. The other parameters are $s = 0.05$, $\gamma = 0.01$, $\sigma_e^2 = 0.01$ and $\omega_0 = 0.1$.

V are independent between species and samples and the number of individuals observed of a species with abundance $x = e^y$ are Poisson distributed with mean $\nu x V$, when conditioned on V. The unconditional mean is then νx but the variance is larger than the mean, increasing with θ^2. The distribution of the number of individuals among species in the sample is now the Poisson-lognormal distribution with parameters $(x + \ln \nu - \theta^2/2, \theta^2)$.
Applying this sampling distribution to to the lognromal model given by equation 7.8 with mean $\eta = s/\gamma$ and variance $\rho^2 = \sigma_e^2/(2\gamma)$ and writing z_i for the number of species with i representatives in the sample, we find (exercise 29)

$$\mathrm{E}z_i = p_i(\eta + \ln \nu - \theta^2/2, \rho^2 + \theta^2)\mathrm{E}S.$$

As for the gamma model the observed abundances constitute a sample from the corresponding zero-truncated distribution $p_i/(1 - p_0)$. The parameters, $\eta + \ln \nu - \theta^2/2$ and $\rho^2 + \theta^2$ of this distribution can then be estimated by maximizing the likelihood function numerically. So, in order to find the parameters η and ρ^2 of the abundance distribution we need to have some estimates of the over-dispersion θ^2 as well as the sampling intensity ν. The expected number of species in the community, $\mathrm{E}S$, can be estimated as $S/(1 - p_0)$, where S is the observed number of species.

7.7 Heterogeneous models

We have seen that there are two very different approaches to analyzing the distribution of species abundances in communities. The first one, introduced by R.H. MacArthur, was the niche theory, where species divide the available recourses between them according to their competitive ability. The relative abundances of species are then determined by all traits that act together to determine the carrying capacity of each species. Ignoring stochastic fluctuations in species abundance through time, the set of relative abundances are then just carrying capacities relative to the total carrying capacity of the community. The second approach was the neutral theory, where every temporal change in species abundances are purely random, actually pure ecological drift determined only by demographic stochasticity. There are no carrying capacity for a given species and no differences in traits that make their dynamics differ. The total community size, however, is kept constant, so there is a density dependence acting on the total number of individuals in the community. The homogeneous models presented in the previous section included environmental stochasticity as well as density regulation, but was still based on the assumption of all species having the same kind of dynamics. Apparently, none of these approaches can be quite realistic descriptions of communities since species are known to have different traits affecting their dynamics, as well as often being subject to rather large temporal fluctuation that most conveniently are described by environmental as well as demographic stochastic terms. Now we define a class of dynamic stochastic abundance models that are naturally classified between the two above most extreme types, including the niche concept indirectly by open up for species having different dynamics and different carrying capacities, but also including temporal stochastic fluctuations in species abundance.

Suppose that species entering the community do not have the same dynamic parameters, but a set θ of parameters generated independently at invasion for each species by some distribution $\pi(\theta)$. Then, the speciation rate for species with θ in some small (multidimensional) region $(\theta, \theta + d\theta)$ is $\omega_0\pi(\theta)d\theta$. The abundance of these species at time $t = 0$ then follow an inhomogeneous Poisson process with rate, say $\lambda(x;\theta)\pi(\theta)d\theta = (\omega_0\pi(\theta)d\theta)G(x, x_0;\theta)$, where $G(x, x_0;\theta)$ is the Green function for the diffusion with parameters θ. Since Poisson processes are additive, the Poisson rate is the sum of the Poisson rates for all θ (exercise 30), which in the continuous case is

$$\lambda(x, x_0) = \int \lambda(x, x_0;\theta)\pi(\theta)d\theta. \tag{7.9}$$

Fig.7.9 shows a schematic presentation of some pathes in such a heterogeneous model. This type of model can be used to study the effect of het-

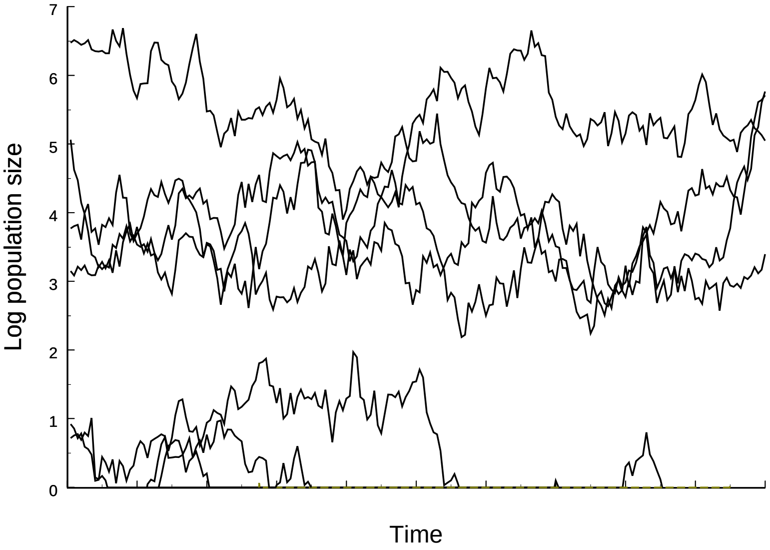

Figure 7.9: Schematic presentation of log population trajectories with speciations and extinctions. The simulated model is heterogeneous and the time scale is arbitrarty.

erogeneity in dynamic parameters, a heterogeneity determined by variation in traits among species. For example, even closely related species are often known to have rather different growth rates and density regulation, as well as different demographic and environmental variances. This makes it possible to study a number of different types of species abundance models. Here we illustrate this by considering the heterogeneous lognormal abundance model generated by normally distributed growth rates among species. Some other examples are left as exercises (31 and 32).

The heterogeneous lognormal model

Now we choose θ in equation 7.9 as the stochastic growth rate s in equation

7.8 and assume that species entering the community has stochastic growth rates generated from a normal distribution with mean s_0 and variance τ^2. This leads to the Poisson rate (exercise 32) on the log scale

$$\lambda(y) = \frac{2\omega_0}{\sigma_e^2} e^{s_0\eta/\sigma_e^2} e^{-(y-\eta)^2/(2\rho^2)}, \tag{7.10}$$

where $\eta = \psi s_0/\gamma$, $\rho^2 = \psi\sigma_e^2/(2\gamma)$ and $\psi = [1 - 2\tau^2/(\gamma\sigma_e^2)]^{-1}$. This model is stationary only if $\tau^2 \leq \frac{1}{2}\sigma_e^2\gamma$. Otherwise, the number of species and individuals in the community will approach infinity. However, as we have discussed previously in this chapter, the overall density regulation of the total community size will necessarily regulate the parameters so that stationarity requirement is fulfilled.

Although the distribution of s among invading species is normal with mean s_0 and variance τ^2, the distribution of s among species present in the community is different. This is due to the fact that the value of s strongly affects the lifetime of the species. Species with a large stochastic growth rate are more likely to stay in the community for a long period of time. It turns out (exercise 34) that the distribution of s in the community is still the normal distribution, but with mean $s_0\psi$ and variance $\tau^2\psi$. Both of these parameters are larger than the corresponding parameters at invasion.

The carrying capacity of a species with stochastic growth rate s is $K = e^{s/\gamma}$, which is lognormally distributed among species at invasion as well as in the stationary community. Using well known properties of the lognormal distribution we find that even the coefficient of variation of K among species is larger in the community than at invasion (exercise 35). The fact that nature tends to select the species with the largest fitness can here be interpreted as choosing those with large values of s through this parameter's effect on the time to extinction. The above result shows that this 'selection process' does not necessarily reduce the heterogeneity of the community, and may actually make the community much more heterogeneous than the heterogeneity recorded at invasion.

7.8 Species area curves

7.8.1 Introduction

More than 20 years before Fisher's log series model was published, the botanists H.A. Gleason and O. Arrhenius investigated community patterns by plotting the number of plant species recorded within areas of different size against their size. Typically, the number of species increases with the size of the area, and two different mathematical curves were soon proposed, species number and log species number, respectively, being proportional to log area. These initial contributions were based on choosing areas within a much larger area containing the total community. Accordingly, one can claim that the areas are a kind of random samples of areas within the community. However, since species do not distribute themselves at random but may show large degree of clustering as we have discussed in chapter 6, different species having small and large density at different locations, these samples are not random samples of individuals from the community. Random samples of individuals will produce somewhat different curves, sometimes called rarefaction curves. Even the sampling of areas may be performed in different ways that may give somewhat different curves. One may start with a small sample and enlarge this by adding new areas to it, or one may sample many different disjoint areas of different sizes. A fourth way of producing curves of species number increasing with area or individual number relates to subdivided populations, such as for example island. Curves are then produces by plotting species number of each sub-population against its area or individual number.

The first attempt to establish a mathematical connection between the species abundance distribution and the species-individual curve was done already in Fisher's original paper in 1943. He showed that the curve based of random sampling of individuals, such as the sampling by many light traps as performed by his colleges, led to a number of recorded species being approximately proportional to the log of the number of individuals sampled. A number of large empirical studies based on different sampling techniques has later been performed by ecologists, and many theoretical models have been prosed to explain observed patterns. Here we consider a few such models, including some that relate to the stochastic theory of species abundance distributions.

One problem related to these curves is how to estimate the total number of species in the community. Since most communities of interest are extremely large, one can only hope to find a small fraction of its individuals in samples. This fact, combined with the empirical observation that one usually find a large number of species that are rare with only one or two individuals in the samples, makes this problem very difficult. Fisher's model, for example, says that the number of singletons actually stays constant approximately equal to α as the sample size increases, all the time revealing species not previously observed. The parameter α may sometimes be as large as 40. Although modeling of species abundance distributions is important in trying to reveal the true number of species, estimates of species numbers will always remain rather uncertain. The uncertainty is a combination of statistical standard errors in the estimator of species number and the uncertainty in what is the correct shape of the left tail of the species abundance distribution.

Although it is extremely difficult to estimate the number of species, it is possible to get a rather reliable estimate of the total abundance of the species not represented in the sample. To se this we consider a community with s species with relative species abundances $p_1, p_2, ..., p_s$ and assume Poisson sampling. Then, the number of representatives of species i, $i = 1, 2, ..., s$, in the sample, are independent Poisson variates X_i with mean νp_i, where ν now is the total expected number of individuals in the sample, $\nu = \mathrm{E}N$. Accordingly, N is an estimator of ν.

Let I_i be an indicator of the event that species i is not represented in a large sample, while J_i indicates that it is represented by exactly one individual. Then, the unknown total abundance of unobserved species, which we want to predict, is the stochastic variable $U = \sum p_i I_i$. Now, the variable $\hat{U} = \sum J_i/\nu$ has the same mean as U (exercise 36). Accordingly $\sum J_i/N = z_1/N$, where

z_1 is the number of singletons in the sample, can be used to predict U. In order to assess the uncertainty in this prediction we evaluate the variance of $(\hat{U} - U)$, which turns out to be $\sigma^2 = (2\mathrm{E}z_2 + \mathrm{E}z_1)/\nu^2$ (exercise 37), which can be estimated by $\hat{\sigma}^2 = (2z_2 + z_1)/N^2$. Hence, an approximately 95% prediction interval for the total relative abundance of the unobserved species is $[z_1/N - 2\hat{\sigma}, z_1/N + 2\hat{\sigma}]$.

7.8.2 Rarefaction

The simplest species-individual curves are the so-called rarefaction curves produced by calculating the expected number of species in a random sub-sample of size, say n, from a real sample of size $N \geq n$. This gives an estimate of the species-individual curve under random sampling of individuals from the community because the sub-samples are also random samples from the community, but the computations are only feasible for $n \leq N$. Let the number of individuals in the sample be $X_1, X_2, \ldots, X_S$, all of which are at least one, $\sum X_i = N$, and let J_i be an indicator of the event that species i is represented in a sub-sample of size n and write $S_n = \sum J_i$ for the number of species in the sub-sample. Then

$$\mathrm{E}J_i = P(J_i = 1) = 1 - P(J_i = 0) = 1 - \frac{\binom{N-X_i}{n}}{\binom{N}{n}}$$

giving the expected number of species

$$\mathrm{E}(S_n) = S - \sum_{i=1}^{S} \frac{\binom{N-X_i}{n}}{\binom{N}{n}}.$$

Notice, however, that this is not the expected number of species in a random sample of size n from the community, but the expectation conditioned on the real sample. Hence, unconditionally it is an unbiased estimator of the number of species in a random sample of size n. Using the same technique we can calculate the variance of S_n given the sample (exercise 38), but again, this is not the variance of the estimator but the conditional variance given the sample.

7.8.3 Observed species number under random sampling using abundance models

Now consider the abundance model with Poisson rate $\lambda(x)$, that is, the number of species in the community with abundance in the interval $(x, x + dx)$ is $\lambda(x)dx$. Further, we assume Poisson sampling with intensity ν so that a species with abundance x is represented by a Poisson distributed number of individuals in the sample with mean νx. Using the previous notation z_j for the number of species with j representatives in the sample, the z_j are independent Poisson variates with means

$$\mathrm{E}(z_j) = \int_a^\infty \frac{(\nu x)^j}{j!} e^{-\nu x} \lambda(x) dx,$$

where a is the extinction barrier. The number of species in the sample is also Poisson distributed, with mean

$$\mathrm{E}(S) = \int_a^\infty (1 - e^{-\nu x}) \lambda(x) dx, \tag{7.11}$$

while the expected number of individuals sample is $\mathrm{E}(N) = \nu \int_a^\infty x\lambda(x)dx$. If the abundances are scales so that the total expected abundance of all species in the community is one, then $\mathrm{E}(N) = \nu$, in which case ν can be estimated by N.

This species-individual curve takes a simple form for the gamma model with $a = 0$. Scaling the total abundance to having unit expectation we have for $k > -1$ and $\alpha > 0$

$$\lambda(x) = \frac{\alpha^{k+1}}{\Gamma(k+1)} x^{k-1} x^{-\alpha x}.$$

Performing the integration (exercise 39) then gives for $k \neq 0$

$$\mathrm{E}(S) = \frac{\alpha}{k} \left[1 - \frac{\alpha^k}{(\alpha + \nu)^k} \right] \tag{7.12}$$

and

$$\mathrm{E}(S) = \alpha \ln\left(\frac{\alpha + \nu}{\alpha}\right) \tag{7.13}$$

for $k = 0$. Now, replacing ν by its estimate N we see that the expected number of species is approximately proportional to log number of individuals for Fisher's model ($k = 0$) when the sample size is large. For shape parameters k between -1 and 0 we see that the expected number of species is approximately proportional to N^{-k} for large samples. Hence, in the case of random sampling of individuals, the gamma model explains the linear logarithmic relationship being approximately $\alpha \ln N$ for $k = 0$ and the the linear relationship in a double logarithmic plot being approximately $\alpha^{k+1}/(-k)N^{-k}$ for $-1 < k < 0$.

For other models such as for example the lognormal species abundance model, the curves produced under random sampling can be evaluated by numerical integration.

7.8.4 Island size curves

The dynamic models we have presented earlier, the neutral model with local communities as well as the homogeneous and heterogeneous models with environmental stochasticity and density regulation, are all models that can be used to calculate the (expected) number of species. However, the problem that needs to be resolved in order to derive species-area or species-individual curved is to find realistic descriptions of how the parameters in the models depends on the area. For example, it seems realistic to assume that the stochastic growth rate s at small densities is not affected by area. This assumption may also be realistic for the environmental and demographic variance. The two parameters most likely to be affected is the strength of density regulation and the speciation or invasion rate.

Hubbell's neutral model

Fig.7.3 shows examples of the abundance spectrum for Hubbell's neutral model. The area under these curves are the expected number of species in

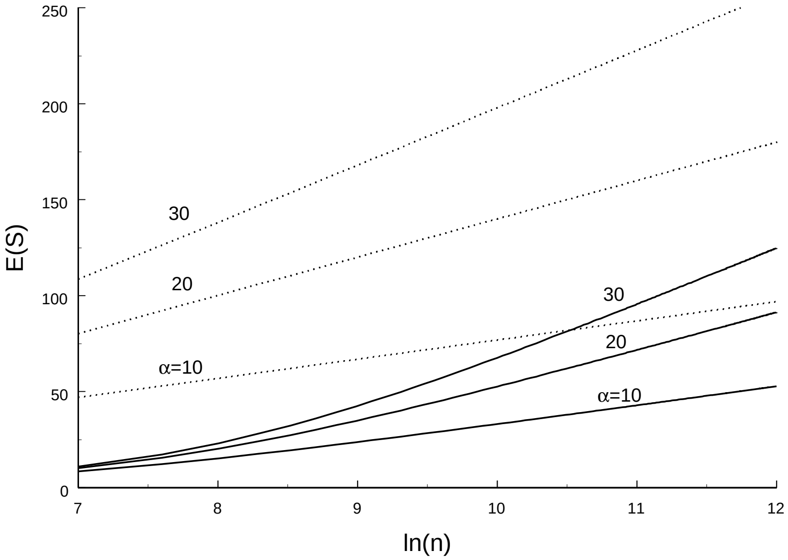

Figure 7.10: The solid lines are species-individual curves, expected number of species ES against log local community size $\ln n$, for Hubbell's neutral model. The curves are given for three different values of the parameter α in Fisher's model describing the total meta-community. The dotted lines are the corresponding curves for random sampling of the same number of individuals from the meta-community, that is E$S = \alpha \ln[(\alpha + n)/\alpha]$ with slope approximately equal to α for large values if n. The other parameters are $N = 10^7$, $\sigma_d^2 = 1$, and $m = 0.001$.

the local community. Hence, by numerical integration we can easily find the expected number of species as function of the size n of the local community. Some examples of such curves, showing the expected number of species as function of $\ln n$, are shown in Fig.7.10.

Notice that the curves are approximately straight lines with slope α for relatively large values of n. This is the same kind of linear relationship that Fisher found under random sampling of individuals from the total community, which was $\mathrm{E}(S) = \alpha \ln[(\alpha + N)/\alpha]$. It appears from the graph that the curves has the same slope approximately equal to α, but the curves for random sampling have much larger number of species for the same number of individuals. This may be understood intuitively by observing that Hubbell's local community is a kind of random sampling of individuals from the meta-community. Actually, the new individuals entering the local community constitute such a sample, but the species dynamics, especially the rare ones, will lead to a large number of local extinctions. Accordingly, we should expect the species number to be much smaller than under random sampling, as demonstrates in Fig.7.10.

The homogeneous gamma model

In the gamma model, for example, with infinitesimal mean $rx - \beta x^2 = rx(1 - x/K)$, with $K = r/\beta$, the area is a major limitation of resources. Therefore, we must expect that the carrying capacity K increases with increasing area. As long as we keep the stochastic growth rate $s = r - \sigma_e^2/2$ constant, the density-regulating parameter β in this model must decrease with area.

A complete modeling, however, require that we also include information about how area affect the invasion rate. Most likely, the rate will increase slightly with area, but only extensive data on invasion that is very hard to find, can tell what is a realistic functional relationship. Here, we only consider models with constant rates. However, it will appear from the derivation that any functional form for the invasion rate can easily be used to derive mode general relationships.

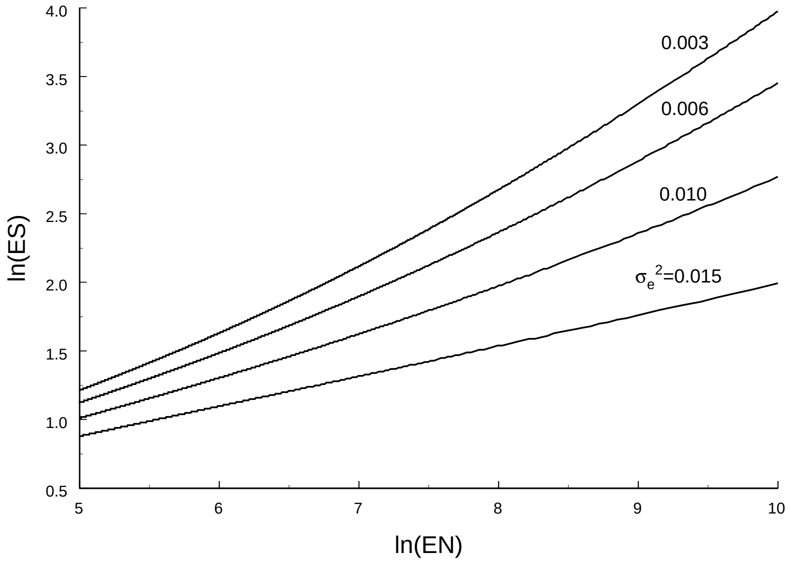

Figure 7.11: Species individual curves for the gamma model given by equation 7.5 in a double logarithmic plot. The curves are generated by varying the parameter β determining the carrying capacities of each species. The curves are shown for some different values of the environmental variance σ_e^2. The other parameters are $r = 0.01$, $\sigma_d^2 = 0.5$, and $\omega_0 = 0.1$. The extinction barrier is chosen at $a = 1$.

The derivation of species-individual curves under the assumption that only the density-regulation depends on the area is very simple in principle. The expected number of species $\mathrm{E}(S) = \int \lambda(x)dx$ and individuals $\mathrm{E}(N) = \int x\lambda(x)dx$ both depends on the parameter determining the density-regulation and the hence carrying capacities. Therefore, simply by varying this parameter only, pairs of numbers $[\mathrm{E}(S), \mathrm{E}(N)]$ can be found defining $\mathrm{E}(S)$ as function of $\mathrm{E}(N)$. In Fig.7.11 we show such curves in a double logarithmic plot produced from the gamma model with environmental and demographic stochasticity producing the Poisson rate given by equation (7.5) using extinction barrier at $a = 1$.

Notice that the curves look approximately as straight lines over rather large ranges of N, for example over ranges in $\ln(N)$ of about 2.3 corresponding to a 10-fold change in N. The slope of these curves is in the range 0.2 to 0.4 which is quite typical for slopes estimated from empirical data.

The homogeneous lognormal model

Fig.7.12 shows species individual curves in a double logarithmic plot for the homogeneous lognormal model produced in the same way as for the gamma model. It appears that the curves are approximately straight lines and the slope is approximately the same for different values of σ_d^2. This, together with a number of other numerical calculations, suggests that we can analyze the slope using the simplified model with no demographic stochasticity given by equation (7.8), or the corresponding heterogeneous model with normally distributed stochastic growth rates among species given by equation (7.10).

The heterogeneous lognormal model

Now considering equation (7.10) for the Poisson rate in the heterogeneous lognormal model and introducing the parameter $\delta = \gamma - 2\tau^2/\sigma_e^2$ the Poisson rate takes the form

$$\lambda(y) = \frac{2\omega_0}{\sigma_e^2} e^{s_0^2/(\delta\sigma_e^2)} e^{-(y-s_0/\delta)^2/[2\sigma_e^2/(2\delta)]},$$

which is equivalent to the rate in the homogeneous model given by equation (7.8) with the stochastic growth rate s replaced by the mean s_0 and γ replaced

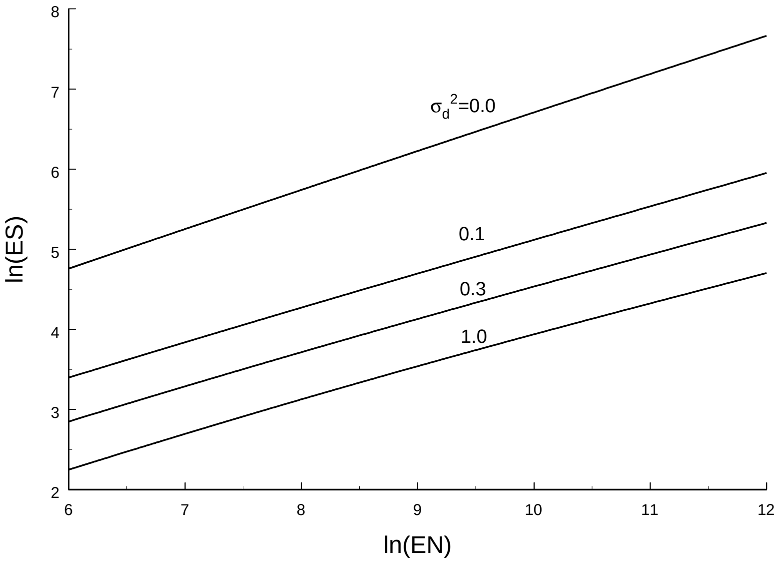

Figure 7.12: The upper curve is the species individual curves for the homogeneous lognormal model given by equation 7.8 in a double logarithmic plot. The curves are generated by varying the parameter γ determining the strength of density regulation. The other parameters are $s = 0.01$, $\sigma_e^2 = 0.001$, and $\omega_0 = 0.1$. The three lower graphs shows the same curves for three different values of the demographic variance. The extinction barrier is chosen at $a = 1$.

by δ. The mean and variance in this Gaussian function are $\eta = s_0/\delta$ and $\rho^2 = \sigma_e^2/(2\delta)$. Since the strength of density-regulation only affects δ, the species-individual curves for the heterogeneous model can be produced simply by varying δ. The mean number of species in the community is

$$\mathrm{E}S = \int_0^\infty \lambda(x)dx = \frac{2\omega_0}{\sigma_e^2} e^{s_0^2/(\delta\sigma_e^2)} \sqrt{\pi\sigma_e^2/\delta}\,\Phi(\eta/\rho),$$

while the mean number of individuals is

$$\mathrm{E}N = \int_0^\infty e^x \lambda(x)dx = \frac{2\omega_0}{\sigma_e^2} e^{s_0^2/(\delta\sigma_e^2)} \sqrt{\pi\sigma_e^2/\delta}\, e^{(s_0+\sigma_e^2/4)/\delta}\Phi(\eta/\rho + \rho).$$

To find an expression for the slope in the double logarithmic plot we use the variable $\kappa = -\ln\delta$ so that $\eta = s_0 e^\kappa$ and $\rho^2 = \sigma_e^2 e^\kappa/2$. We can now study the relation between $\mathrm{E}S$ and $\mathrm{E}N$ as κ, and hence the density regulation given by γ, varies. Expressing $\ln \mathrm{E}S$ and $\ln \mathrm{E}N$ by η, ρ^2 and κ and observing that $d\eta/d\kappa = \eta$, $d\rho^2/d\kappa = \rho^2$, and then using the fact that η/ρ^2 does not depend on κ, we get (exercise 39)

$$\frac{d\ln \mathrm{E}S}{d\kappa} = \frac{1}{2}\left[1 + \eta^2/\rho^2 + G(\eta/\rho)\right]$$

and (exercise 41)

$$\frac{d\ln \mathrm{E}N}{d\kappa} = \frac{1}{2}\left[1 + \eta^2/\rho^2 + G(\eta/\rho + \rho)\right] + \eta + \frac{1}{2}\rho^2.$$

Here $G(x) = x\phi(x)/\Phi(x)$ where ϕ is the standard normal density, that is, the derivative of Φ. From this we find the slope

$$z = \frac{d\ln \mathrm{E}S}{d\ln \mathrm{E}N} = \frac{1 + \eta^2/\rho^2 + G(\eta/\rho)}{1 + \eta^2/\rho^2 + G(\eta/\rho + \rho) + 2\eta + \rho^2}. \tag{7.14}$$

An interesting conclusion from this derivation is that the slope in the double logarithmic plot depends on the parameters in the heterogeneous lognormal model only through the mean and variance of the underlying normal distribution of log abundances. Rather surprisingly, information about the heterogeneity parameter τ^2 is not required. Therefore, since the mean and variance can be estimated from one single observation from a community by fitting the zero-truncated Poisson lognormal model, then also z can be estimated. We have previously seen that this requires that we have information about the sampling intensity as well as a possible over-dispersion in the sampling relative to the Poisson distribution.

7.8.5 Curves produced by quadrat sampling

In chapter 6 we studied the spatial distribution of individuals. The simplest model for a given species is given by the homogeneous Poisson process, that is, the underlying density is constant in space and individuals are randomly scattered according to the given density of the species. If all species in the community are described by this simple model, then for any species the number of individuals within any given area will be Poisson distributed with mean value proportional to its density as well as the area. Accordingly, the number of species within an area can be considered as a Poisson sample of individuals from the community with a given sampling intensity. Increasing the area has the same effect as increasing the sampling intensity. As a consequence, the species-individual curves produced by varying the area are exactly the curves we earlier have produced by random sampling of individuals.

On the other hand, if there is spatial variation in population densities described by a spatial auto-covariance function so that individuals to some extent are clustered in space, the curves will necessarily be different. However, it is difficult to make general models that cover all the different possibilities. For example, species with different abundances may not have the same type of spatial scaling. Here we exemplify by considering Fisher's model and the extended gamma model for the overall mean density and assuming that the spatial coefficients of variation and autocorrelations are the same for all species.

Let $x|\mu(z)$, $z = (z_1, z_2)$, denote the mean abundance of a species at location z where $\mu(z)$ is a stochastic spatial field with mean $\mathrm{E}\mu(z) = 1$. Accordingly, the abundance x is scaled so that it is the expected number of individuals in a unit area. Notice that this assumption makes the coefficient of variation of density, as well as the spatial scaling, the same for all species. Assuming an inhomogeneous Poisson process for the location of individuals, the number of individuals of a species with abundance x within an area A is Poisson distributed with mean $W = x \int_A \mu(z)dz = xU$ for a given $\mu(z)$. For a randomly chosen area, however, U will have some distribution , say $f(u|A)$ depending on the stochasticity of the field $\mu(z)$. The expected value of $U = \int_A \mu(z)dz$ must be A because the field has mean 1. Writing σ^2 for the variance of $\mu(z)$ we have seen in chapter 6 that the variance of U is $\sigma^2 A^2 \bar{\rho}(A)$, where $\bar{\rho}(A)$ is the mean spatial correlation between for two points chosen at random within A. The probability that this species is present in A is therefore $\int(1-e^{-ux})f(u|A)du$, giving the mean number of species in A for this spatial model

$$\mathrm{E}S_{spatial}(A) = \int\int(1 - e^{-ux})f(u|A)\lambda(x)dudx. \tag{7.15}$$

For homogeneous pattern with $\mu(z) = 1$ we have simply $U = A$ and the corresponding species-area curve is simply the same as the curve produced by random sampling of individuals as pointed out above. With the whole probability mass of U at A equation (7.15) is the same as equation (7.11) with sampling intensity $\nu = A$. Hence, writing $\mathrm{E}S(\nu)$ for the the expected number of species in the sample under random sampling of individuals expressed by equation (7.11), we find

$$\mathrm{E}S_{spatial}(A) = \int \mathrm{E}S(u)f(u|A)du,$$

where $f(u|A)$ denotes the distribution of $\int_A \mu(z)$ depending of the properties of the field $\mu(z)$. This expression is the sum of the expected contributions from each species and is therefor unaffected by a possible dependence between species. We arrive at the same expression for the curves when the fields $\mu(z)$ are independent or dependent among species. The variance of the species number found in areas of given sizes is, however, depends strongly on the covariances between the indicator variables for the presence of each species. Now, consider Fisher's model with the species-individual curve given by equation (7.13) as an example. Using the second order Tailor expansion around $u = A$ we have

$$\mathrm{E}S_{spatial}(A) \approx \alpha \ln\left(\frac{\alpha + A}{\alpha}\right) - \frac{\alpha\sigma^2 A^2 \bar{\rho}(A)}{2(\alpha + A)^2}.$$

The factor $A^2/(\alpha + A)^2$ in the last term increases from zero to 1 as A increases, while $\bar{\rho}(A)$ usually decreases from 1 toward zero. Consequently, for areas smaller than some intermediate value A^* the curve has smaller slope than the curve produced by random sampling, while the slope is large for values above this. This is related to the concept of β-diversity. We see that the value A^* depends on the spatial scaling of the density field $\mu(z)$. The spatial effect is demonstrated in Fig.7.13 using a Gaussian form for the spatial autocorrelation of $\mu(z)$ and a function $\bar{\rho}(A)$ computed for squares A using equation (6.11). We see that the spatial aggregation of individuals makes the curve have small slope in the beginning and a large slope when $\ln(A)$ becomes larger than about $\ln(l^2)$, where l is the standard error of for the correlation function scaled to integrate to one. Notice that the coefficient of variation in the spatial variation of $\mu(z)$ here is as large as $\sqrt{3}$, which means that the densities of all species are very small in some areas and quite large in others.

Using the same approach for the random sampling curve from the extended gamma model with approximately linear relation in a double logarithmic plot for large areas A given by equation (7.12), we find

$$\mathrm{E}S_{spatial} \approx \frac{\alpha}{k}\left[1 - \left(\frac{\alpha}{\alpha + A}\right)^k\right] - \frac{\alpha^{k+1}\sigma^2 A^2 \bar{\rho}(A)}{2(\alpha + A)^{k+2}}$$

which is exemplified in Fig.7.14. Notice that the curves in the case of spatial patchiness are very close to straight lines for $\ln(A)$ between 5 and 9. However, they have much larger slopes than the line for random sampling, which has

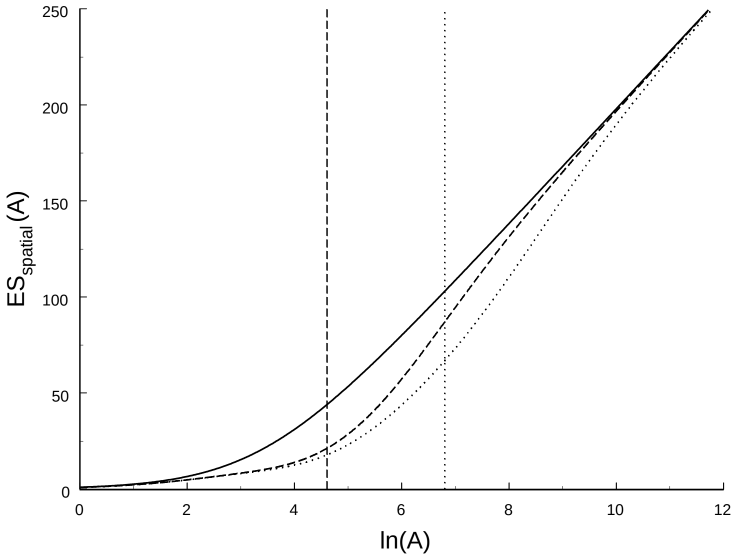

Figure 7.13: The second order approximation to the species area curve when the spatial correlation for the density of all species is of the Gaussian form $\rho(r) = e^{-r^2/(2l^2)}$, where r is the distance. The solid line is the species area curve for constant densities for Fisher's model given by equation (7.13). The dashes line is the curve for scale $l = 10$, and the dotted line for scale $l = 30$. The corresponding vertical lines are at $\ln(A) = \ln(l^2)$. The other parameters are $\sigma^2 = 3$ and $\alpha = 30$.

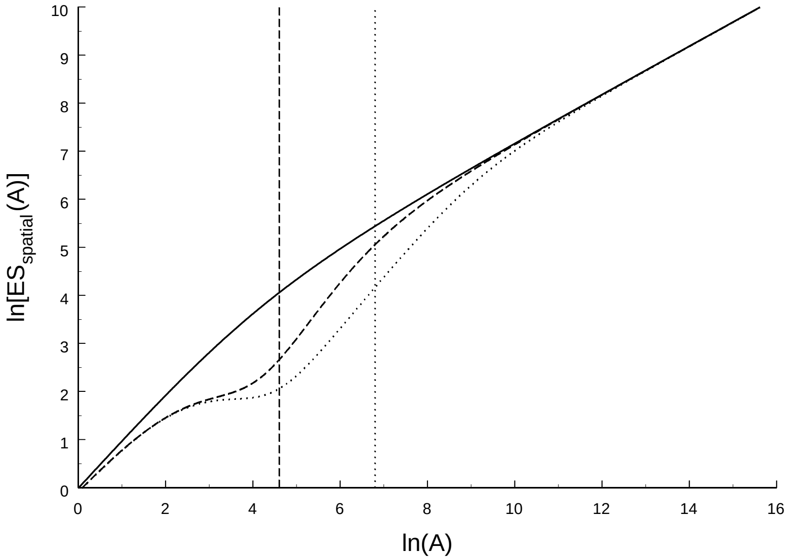

Figure 7.14: The second order approximation to the species area curve when the spatial correlation for the density of all species is of the Gaussian form $\rho(r) = e^{-r^2/(2l^2)}$, where r is the distance. The solid line is the species area curve for constant densities for the extended gamma model with $k = -0.5$ given by equation (7.12). The dashes line is the curve for scale $l = 10$, and the dotted line for scale $l = 30$. The corresponding vertical lines are at $\ln(A) = \ln(l^2)$. The other parameters are $\sigma^2 = 3$ and $\alpha = 20$.

slope $-k = 0.5$ for large areas. This indicates that large slopes found over large ranges in empirical studies may to some extend be a result of spatial segregation, that is, a result of β-diversity.

7.9 Temporal and spatial analysis of similarity

7.9.1 Introduction

In the introduction to this chapter we defined Jaccard and Sørensen's index for similarity between two communities. Since these two indices, that actually are equivalent, are based on counting the number of species in the two communities, it is obvious from our previous discussion of species-area curves that the estimation of the indices from samples must be difficult due to many unobserved rare species in both communities. On the other hand, if we know that the community can be described by a certain species abundance model, the number of species can be estimated by fitting a species abundance distribution taking the sampling into account, and calculate the number of species from the estimated parameters in the model. This indicates strongly the possibility of also relating indices of similarity like Jaccard and Sørensen's index in some way to abundance distributions, which we discuss further in 7.9.4.
Let (x_i, y_i), $i = 1, 2...s$ be the abundances of species i in two communities, either two communities at the same time but at different locations, at different times at the same location, or different times at different locations. When dealing with abundance models we have considered $(x_1, x_2, ...x_s)$ as well as $(y_1, y_2, ...y_s)$ as samples from a distribution, which is called the species abundance distribution. These samples, however, are not in general independent. If the time differences and spatial distances between the communities are small, the two-dimensional species abundance distribution of (x_i, y_i) is a bivariate distribution expected to have large correlation. In statistics the correlation is the most widely used measurement for quantifying the similar-

ity between samples like this. Consequently, any measurement of correlation can be used as index of similarity between two communities described by species abundance distributions.

Below be return to the heterogeneous lognormal model and discuss contributions to the correlation between abundances on the log scale from correlations between noise terms, heterogeneity, density-regulation and migration. This decomposition of the variance in the lognormal model can be utilized to perform statistical analysis of spatial and temporal community data.

7.9.2 Spatio-dynamical species abundance models

The heterogeneous lognormal species abundance model given by equation (7.10) was derived by assuming that the stochastic growth rates of species were generated by a normal distribution at speciation/colonization. The stochastic growth rates for the resident species when stationarity is reached is also normal, although the parameters are different (exercise 33). Since speciations are rare, we now assume stationarity and ignore speciations and extinctions during some time interval under investigation. Then consider a site $z = (z_1, z_2)$ with log abundance $x(t, z)$ for some species at time t, and dynamics given by

$$dx(t, z) = [s - \gamma x(t, z)]dt + \sigma_e dB(t, z)$$

where $dB(t, z)$ is a temporal Brownian motion specific for site z. The stochastic growth rate s is normally distributed among species with mean, say s_0 and variance τ^2 (a redefinition of the notation $s_0\psi$ and $\tau^2\psi$ in exercise 33), and the noise terms are independent among species. The noise terms are, however, correlated in space with correlation depending on the distance between sites given by $\mathrm{E}[dB(t, z)dB(t, z+w)] = \rho_e(w)dt$ (exercise 42). We have seen in 6.3.2 that the Moran effect holds for this model so that the stationary distribution of $[x(t, z), x(t, z + w)]$ is the bivariate normal distribution with correlation $\rho_e(w)$ and variances $\sigma^2 = \sigma_e^2/(2\gamma)$. More generally we may consider the same species at two different times as well as different locations and use the assumption that the stochastic growth rates are normally distributed with variance τ^2. The correlation is then (exercise 43)

$$\rho(h, w) = \mathrm{corr}[x(t, z), x(t + h, z + w)] = \frac{\rho_e(w)e^{-\gamma h}\sigma^2 + \tau^2}{\sigma^2 + \tau^2}. \tag{7.16}$$

7.9.3 Decomposition of the variance

The lognormal model

The correlation $\rho(w, h)$ between two communities can be estimated from data under the assumption of Poisson sampling or Poisson lognormal sampling from both communities with over-dispersion parameter θ^2 as defined in section 7.6.2. This over-dispersion adds a term θ^2 to the variance parameter for both marginal distributions, but have no effect on the covariance. Accordingly, evaluating the terms of the bivariate Poisson lognormal distribution (exercises 44 and 45) for the data from these two communities, the correlation parameter that can be estimated by maximum likelihood (exercises 46 and 47) is

$$\rho^*(h, w) = \frac{\rho_e(w)e^{-\gamma h}\sigma^2 + \tau^2}{\sigma^2 + \tau^2 + \theta^2} \tag{7.17}$$

while the variance parameter is

$$v = \sigma^2 + \tau^2 + \theta^2.$$

If samples are available from a large number of communities at different locations and different times, the correlations $\rho^*(w, h)$ can be estimated from each pair of communities (exercise 47). The variance parameter v can be chosen as the mean of all estimates of v. The correlation estimates are preferably smoothed, for example by assuming a parametric form of $\rho_e(w)$, to find estimates of γ, $\rho^*(0, 0)$ and $\rho^*(z, \infty)$. The decomposition of the common variance v of the lognormal species abundance distributions, that is, the components σ^2, τ^2 and θ^2, are then found from the relations

$$\theta^2 = v([1 - \rho^*(0, 0)],$$

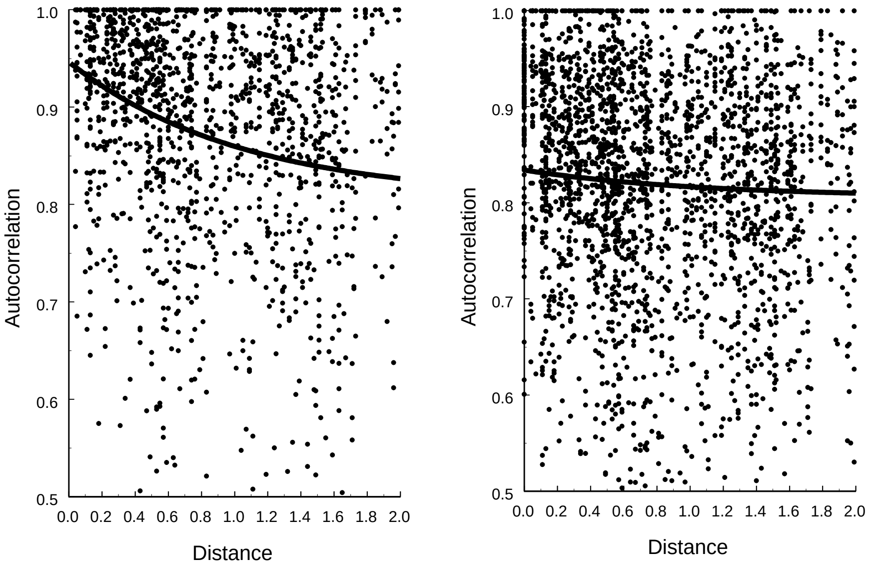

Figure 7.15: Correlation estimates for samples of tropical butterfly communities from 25 sites over 5 years. The left panel shows correlation estimates for zero time difference, while the right one has time differences one year. The estimates are found by maximizing the likelihood function based on the bivariate Poisson lognormal distribution given in exercises 43-48.

$$\sigma^2 = v[\rho^*(0,0) - \rho^*(z,\infty)],$$

and

$$\tau^2 = v\rho^*(z,\infty),$$

following from equation (7.17). The spatial scale of the noise is given by $l = 1/\gamma$. Examples of smoothed estimates are shown in Fig.7.15.

We have in the above analysis assumed a given environmental correlation in the noise expressed by $\rho_e(r)$ that alone generates the same equal spatial correlation in log population size. However, we have seen in chapter 6 that other effects, in particular migration and permanent spatial heterogeneity, also may have large effect on the field $X(z)$. This will make the expression for the correlation, such as the one given by equation (7.16), more complex.

Hubbell's neutral model

It is impossible to compare the above spatio-temporal results for the log-normal model to Hubbell's neutral model because we have no spatial theory for that model unless we also model distance-dependent migration between islands. However, it is possible at least to make a comparison with the result from the pure temporal analysis given by equation (7.16) with $w = 0$ and $\rho_e(w) = 1$.

In the neutral model the infinitesimal mean for a relative abundance p of some species in the local community is $\mu(p) = -m(p-q)$, where q is its fixed abundance in the meta-community. Also, due to immigration, the diffusion for p is stationary. Then we can apply the result for linear infinitesimal means shown in chapter 3.12, saying that the temporal autocorrelation is simply $\rho(t) = e^{-mt}$. This is the autocorrelation for p as well as the absolute variance $x = np$ because the local community size is assumed to be constant. To find the unconditional correlation not conditioned on q, we first need to derive the distribution of q for species that are present in the local community. We can find this from the two-dimensional frequency spectrum $g(p;q)f(q)$ discussed just before equation (7.3), expressing the expected number of species with relative abundance in $(p, p+dp)$ in the local community and in $(q, q+dq)$ in the meta-community when multiplied by $dpdq$. Since the smallest abundance for p indicating presence is $1/n$ the distribution conditioned on presence is

$$f^*(q) = cf(q)\int_{1/n}^{\infty} g(p;q)dp,$$

where c it the required scaling factor. Using this distribution for q we find the unconditional correlation

$$\rho^*(t) = \frac{\sigma^2 e^{-mt} + \tau^2}{\sigma^2 + \tau^2}, \tag{7.18}$$

where now $\sigma^2 = \mathrm{Evar}^*(p|q)$ and $\tau^2 = \mathrm{var}^*(q)$, where * indicates reference to the distribution $f^*(q)$ (exercise 49). Although these are the correlations on the absolute scale while equation (7.16) refers to the log scale, the functions appear to have very similar forms. The most interesting difference, however, is the exponential temporal functions $e^{-\gamma t}$ in the heterogeneous lognormal model and e^{-mt} in the neutral model. That is, the temporal scale is $1/\gamma$ and $1/m$, respectively. A rapid decrease will only occur in the lognormal model under strong density-regulation, while in the neutral model only under large migration (small degree of isolation from the main community).

7.9.4 Correlation and indices of similarity

We have seen that the correlation parameter in the bivariate Poisson lognormal model can be estimated from data without having any information about the sampling intensity for the sampling from the two communities. This is a very important observation because sampling intensities are often unknown, or partly unknown. In statistics correlations are generally defined for any bivariate distribution, so for all two-dimensional species abundance models, we can use the correlation as an index of similarity. Using the normal distribution as an approximation when constructing estimates of correlation is also a well known technique that usually leads to reliable estimates. The same approach can be used for communities. Even if the binormal distribution does not give precise description of the system, a number of empirical studies has shown that it often is a good approximation. Therefore, fitting the bivariate Poisson lognormal distribution and using it to estimate the correlation will rather generally give quite a useful measurement of similarity.

The Jaccard and Sørensen's indices are often estimated directly from data by substituting the number of species in the communities by the numbers actually observed. Estimates obtained in this way are of course very sensitive to sampling efforts. In order to investigate this in some detail it is convenient to count species that are 'likely to be observed' rather than the observed numbers. This will practically make no difference, but it facilitates the analysis. We do this by defining a species as likely to be present if its abundance exceeds a given threshold. Hence, defining thresholds for both communities there is a unique expected number of species 'likely to be observed' in both or in one of the communities corresponding to some unique fractions of revealed species. The Sørensen index was defined as $L = 2A/(2A + B + C)$, where A was the number of species present in both and B and C the numbers present in only one or the other community. Replacing A, B, and C by the above numbers of species with abundances exceeding the thresholds then leads to an index not depending on any sampling assumptions other than the expected fraction of species in the samples.

Returning to the bivariate lognormal species abundance model it is most convenient to define thresholds referring to standardized log abundances. A species with log abundance X is likely to be present if $U = (X - \mu_x)/\sigma_x > \alpha$. The fraction of species exceeding the threshold is then $P(U > \alpha) = 1 - \Phi(\alpha)$, where Φ is the standard cumulative normal distribution. Consequently, this definition of 'presence' through α refers to a sampling effort expected to reveal a fraction $p = 1 - \Phi(\alpha)$ of the species abundance distribution. In the same way we define 'presence' in the other community by β so that $V = (Y - \mu_y)/\sigma_y > \beta$, with corresponding fraction $q = 1 - \Phi(\beta)$ expected to

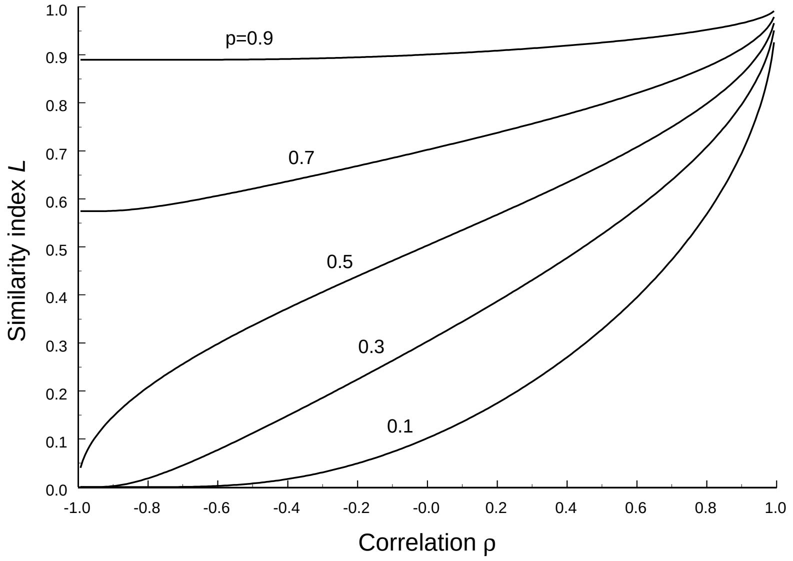

Figure 7.16: Sørensen's similarity index as function of community correlation ρ in the bivariate lognormal community model. Definition of the index is here based on $p = q$, that is, sampling reveals the same expected fraction of species in both communities. The curves are shown for different values of $p = q$.

be seen.

The two-dimensional distribution of (U, V) among species is the standard binormal distribution with correlation ρ,

$$f(u, v; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{u^2-2\rho uv+v^2}{2(1-\rho^2)}}$$

having marginal distributions with zero means and unit variances. Hence, with s species altogether the quantities entering the expression for Sørensen's index are $A = s\int_\alpha^\infty \int_\beta^\infty f(u, v; \rho)dudv$, $B = s\int_\alpha^\infty \int_{-\infty}^\beta f(u, v; \rho)dudv$ and $C = s\int_{-\infty}^\alpha \int_\beta^\infty f(u, v; \rho)dudv$. Inserting these into the expression for L, the species number s disappears and the index, say $L_{p,q}$, appears to be a function of the community correlation ρ only (exercise 50)

$$L_{p,q} = \frac{2G(\alpha,\beta;\rho)}{2G(\alpha,\beta;\rho) + G(-\alpha,\beta;-\rho) + G(\alpha,-\beta;-\rho)}.$$

where $G(\alpha,\beta;\rho) = P(U > \alpha, V > \beta)$, which can be written as the univariate integral

$$G(\alpha,\beta) = \int_{\beta}^{\infty} [1 - \Phi(\frac{\alpha - \rho v}{\sqrt{1-\rho^2}})]\phi(v)dv.$$

This index for different thresholds $p = q$ is depicted as function of the community correlation ρ Fig.7.16. The special case $p = q = 1/2$, that is, when one half of the species is expected to be observed in both communities, Sørensens index takes the simple form (exercise 51),

$$L_{0.5,0.5} = 1/2 + \arcsin(\rho)/\pi$$

which is zero for $\rho = -1$, 0.5 for $\rho = 0$ and 1 for $\rho = 1$. This function is the middle line in Fig.7.16.

7.10 Exercises

1. Show that the information index for a given number of s species takes its maximum value $\ln s$ for $p_i = 1/s$, $i = 1, 2 \ldots, s$.

2. Show that Simpson's index for a given number of s species takes its maximum value $(s-1)/s$ for $p_i = 1/s$, $i = 1, 2 \ldots, s$.

3. Show that the information index is unbounded even if there is a dominant species with relative abundance 0.9.

4. Let Y_i be the number of individuals in a sample representing species i, $i = 1, 2, \ldots, s$, $\sum Y_i = N$. Under the assumption that $(Y_1, Y_2 \ldots, Y_s)$ is multinomially distributed with parameters $(N, p_1, p_2, \ldots, p_s)$, find an unbiased estimator for Simpson's index.

Hint: Find first an unbiased estimator for p_i^2, using the fact that Y_i is binomially distributed.

5. Show that no unbiased estimator exists for the information index.

Hint: Consider a general estimator $\hat{H}_I(N, Y_1, Y_2, \ldots, Y_s)$ and study the mathematical form of its expectation.

6. Consider the diffusion with infinitesimal mean and variance $\mu(p) = -up$ and $\nu(p) = p(1-p)/(2N_e)$ and initial state $p_0 = 1/(2N)$. Show that the Green function is proportional to $p^{-1}(1-p)^{M-1}$, where $M = 4N_eu$.

7. For a 'frequency spectrum' $f(p) = cp^{-1}(1-p)^{M-1}$, show that $c \approx M$.

8. Show that the expected fraction of homozygotes in the neutral infinite allele model is $1/(M+1)$, where $M = 4N_eu$, and that the expected value of Simpson's index for genetic diversity in the same model is $M/(M+1)$.

9. Let the species abundances y in the community follow the gamma distribution $[\rho^k/\Gamma(k)]y^{k-1}e^{-\rho y}$, and assume that a species with abundance y is represented by a Poisson-distributed number of individuals in the sample with mean νy. Then, show that the observed number of individuals for the species represented in the sample follow a zero-truncated negative binomial distribution.

10. Derive Fisher's logarithmic series distribution from the zero-truncated negative binomial distribution derived in exercise 9.

11. Consider the model in exercise 9 and show that the expected number of species with n representatives in the sample approaches Fisher's log-series $\alpha w^n/n$ as the shape parameter k approaches zero and the number of species s in the community approaches infinity so that $ks \to \alpha$, where $w = \nu/(\nu+\rho)$.

12. For the model in exercise 9 show that the marginal distribution of the relative abundance of a given species is the beta-distribution $\Gamma(\alpha)/[\Gamma(k)\Gamma(\alpha-k)]p^{k-1}(1-p)^{\alpha-1}$, where $\alpha = ks$.

13. Use the result in exercise 12 to show that the 'frequency spectrum' in Fisher's model is the same as in Crow and Kimura's neutral infinite allele model.

Hint: Use the relation $k\Gamma(k) = \Gamma(k+1)$.

14. In order to find $\mathrm{E}(H_I)$ in Fisher's model first define the function $G(t) = \int_0^1 p^t(1-p)^{M-1}dp$.
Hint: Use the formula $\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and evaluate $\frac{d}{dt}\ln G(t)$ at $t = 0$.
15. Assume that the number of species with j representatives in the sample, $j = 1, 2, \ldots$, are independent Poisson variates with means $r_j = \alpha w^j/j$. Show that the total number of observed species S and individuals N are jointly sufficient for α and w and that the maximum likelihood estimator of α is the solution of the equation $S/\hat{\alpha} = \ln[(N + \hat{\alpha})/\hat{\alpha}]$ and the estimator of w is $\hat{w} = N/(N + \hat{\alpha})$.
16. Find the maximum likelihood equation for M from Ewens' sampling formula.
17. Show that the likelihood equation for M from Ewens' formula is approximately the same as Anscombe's formula for α in Fisher's model, that is, the solution is approximately the solution of $S/\hat{M} = \ln[(N + \hat{M})/\hat{M}]$.
18. Let $\lambda(x)$ be the intensity of an inhomogeneous Poisson process and let N be the number of Points in the interval Ω for some realization of the Process. Show that, conditionally on $N = n$, the points in Ω are n independent observations from the distribution $\lambda(x)/\int_\Omega \lambda(u)du$.
19. Show that the stationary distribution for the diffusion with infinitesimal mean $\mu(p) = -m(p-q)$ and variance $p(1-p)\sigma_d^2/n$ is the beta-distribution with parameters τq and $\tau(1-q)$, where $\tau = mn/\sigma_d^2$. Find the mean and variance of this distribution as well as $\mathrm{E}[p(1-p)]$.
20. Use the result in exercise 19 to how that the expected value of Simpson's index for the local community is the index for the meta-community multiplied by $\tau/(\tau+1)$.
21. Use equation (7.3) to show that the expected value of the information index on islands in Hubbell's model is
$\mathrm{E}(H_{I,L}) = \int[\Psi(\tau+1) - \Psi(\tau q+1)]\alpha(1-q)^{\alpha-1}dq$.
Hint: First find the expected value of $-p\ln p$ on the island for a species with relative abundance q in the meta-community. Use the technique similar to

that in exercise 14.

22. For an inhomogeneous Poisson process with intensity $\lambda(x)$ on the interval $[a, b]$ consider the realization of points $x_1, x_2, \ldots, x_N$. Show that, conditioned on N, the distribution of $x_1, x_2, \ldots, x_N$ is a sample from the distribution $\lambda(x)/\int_a^b \lambda(u)du$.

23. Consider a diffusion with initial state $x_0 = a+\delta > a$. Find the probability that the path reaches $a + \varepsilon > a + \delta$ before reaching a is δ/ε as $\varepsilon \to 0$ (and consequently $\delta \to 0$). Let ω be the rate of diffusions starting at $a + \delta$ and let the speciation rate be defined as the rate at which these diffusions reach $a + \varepsilon$ for the first time. Show that the speciation rate in the above limit then approaches $\omega\delta/\varepsilon$.

24. Consider a diffusion model with environmental variance only where the density regulation is so small that it can be ignored for $x < n$. If we define speciation to occur as a new species reach abundance $x = n$, show that the speciation rate in the homogeneous gamma model is ω_0/n.

25. Under the assumption of Poisson sampling show that the number of species z_j with j representatives in the sample are independent Poisson distributed variables.

26. Show that the Poisson rate given by equation 7.6 combined with Poisson sampling leads to the zero-truncated negative binomial distribution of observed abundances.

27. Find the expected value of Simpson's diversity index for the model defined by equation 7.6.

28. Find the expected value of the information index for the model defined by equation 7.6.

29. For the lognormal model given by equation 7.8 with mean $\eta = s/\gamma$ and variance $\rho^2 = \sigma_e^2/(2\gamma)$, and sampling given by the Poisson lognormal distribution with overdispersion parameter θ and intensity ν, show that the expected number of species with i representatives in the sample, $i = 0, 1, ...$, is proportional to the terms of the Poisson lognormal distribution with parameters $(\eta + \ln \nu - \theta^2/2, \rho^2 + \theta^2)$.

30. Let $\lambda_i(x)$, $i = 1, 2, \ldots n$ be independent inhomogeneous Poisson processes. Show that the set of points generated by all n processes together is an inhomogeneous Poisson process with rate $\sum_{i=1}^{n} \lambda_i(x)$.
31. Construct a heterogeneous model from the gamma model given by equation 7.6 by assuming that the parameter β, expressing the strength of density regulation, is gamma distributed among species at invasion. Show that the species abundance distribution then is the beta-distribution of the second kind, that is

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{b^p x^{p-1}}{(1+bx)^{p+q}}.$$

Find the parameters p, q and b expressed by the parameter of the heterogeneous dynamic species abundance model.
32. Construct a heterogeneous model by assuming that the parameter σ_e^2 in the model given by equation 7.8 has an inverse gamma distribution among species at invasion. Show that the resulting species abundance distribution on the log scale is Student's T-distribution (the general form where the degrees of freedom is not required to be an integer). Find the parameters of the distribution expressed by the parameters of the underlying dynamical model. Also find the species abundance distribution on the absolute scale and compare it to the lognormal distribution.
33. Use equation 7.9 to find the lognormal species abundance model resulting from assuming that the stochastic growth rate s in equation 7.8 is normally distributed among species at invasion with mean s_0 and variance τ^2.
34. For the model in exercise 32 show that the distribution of s among the species in the community is normal with mean $s_0\psi$ and variance $\tau^2\psi$ when stationarity is reached, where $\psi = [1 - 2\tau^2/(\gamma\sigma_e^2)]^{-1}$.
35. For the model given in exercise 33 show that the coefficient of variation in the carrying capacities among species is larger when stationarity is reached than at invasion.
36. Consider a community with relative abundances $p_1, p_2, \ldots, p_s$, and assume that the number of observed species X_i of species i, for $i = 1, 2, \ldots, s$

are independent Poisson variates with mean νp_i. Let I_i be an indicator of the event that species i is not represented in the sample, and let J_i indicate that it is represented by exactly 1 individual in the sample. Show that the total relative abundance of unobserved species $U = \sum p_i I_i$ has the same expectation as $\hat{U} = \sum J_i/\nu$.

37. For the model in exercise 35 show that $\mathrm{var}(U - \hat{U}) = (2\mathrm{E}z_2 + \mathrm{E}z_1)/\nu^2$, where z_j is the number of species represented by j individuals in the sample.

38. Consider a sample $(X_1, X_2, \ldots X_S)$ of S species from a community (that is, $X_i > 0$) and let J_i be an indicator of the event that species i is represented in a random sub-sample of size $n \leq N$. Find an expression for the variance of the number of species $S_n = \sum J_i$ in the sub-sample conditioned on the sample.

Hint: The covariances between the J_i must be included. These can be calculated by using indicator variables for events $(J_i = 1, J_j = 1) = (J_i J_j = 1)$ for $i \neq j$.

39. In order to derive the expected number of species in a random sample under the extended gamma model first consider $\mathrm{E}(S) = g(\nu)$ as function of the sampling intensity ν. Find an expression for the derivative of $g(\nu)$ and evaluate it by integration. Finally use the fact that $g(0) = 0$ to find the general expression for $\mathrm{E}(S)$ valid for $k > -1$.

40. In the heterogeneous lognormal model given by equation (7.10) define $\kappa = -\ln \delta$ so that $\eta = s_0 e^{\kappa}$ and $\rho^2 = \sigma_e^2 e^{\kappa}/2$. Then show that

$$\frac{d \ln \mathrm{E}S}{d\kappa} = \frac{1}{2}\left[1 + \eta^2/\rho^2 + G(\eta/\rho)\right]$$

where $G(x) = x\phi(x)/\Phi(x)$ and ϕ is the standard normal density.

41. For the model in exercise 39 show that

$$\frac{d \ln \mathrm{E}N}{d\kappa} = \frac{1}{2}\left[1 + \eta^2/\rho^2 + G(\eta/\rho + \rho)\right] + \eta + \frac{1}{2}\rho^2.$$

42. Consider two Brownian motions at location z and $z + w$ and assume that $dB(t,z)dB(t,z+w) = \rho_e(w)dt$. Show that $\rho_e(w)$ then is the correlation between $dB(t,z)$ and $dB(t,z+w)$.

Hint: Consider first a discrete time step Δt. Then go to the limit.

43. Consider Ornstein-Uhlenbeck processes on the form $dX(t,z) = [s - \gamma X(t,z)]dt + \sigma_e dB(t,z)$ and correlated noise given by $\mathrm{E}[dB(t,z)dB(t,z+w)] = \rho_e(w)dt$. Show that

$$\rho(h,w) = \mathrm{corr}[x(t,z), x(t+h,z+w)] = \frac{\rho_e(w)e^{-\gamma h}\sigma^2 + \tau^2}{\sigma^2 + \tau^2}$$

when s is normally distributed with variance τ^2.

44. Assume that the log abundances (X,Y) of species, is a sample from the bivariate normal distribution with parameters $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. Let the sampling process be given by the Poisson distribution in both communities so that the number of individuals of a species in the samples, say (N_x, N_y), under unit sampling effort are Poisson distributed with means e^X and e^Y when conditioned on the abundances, and define the function

$$h_n(\mu, \sigma^2, u) = \frac{\exp(u\sigma n + \mu n - e^{-(u\sigma+\mu)})}{n!}.$$

Then show that the bivariate Poisson lognormal distribution $P(N_x = n_x, N_y = n_y)$ can be written as

$$q(n_x, n_y; \mu_1, \sigma_x^2, \mu_2, \sigma_y^2, \rho) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_{n_x}(\mu_x, \sigma_x, u)h_{n_y}(\mu_y, \sigma_y, v)f(u,v;\rho)dudv$$

where $f(u,v;\rho)$ denotes the binormal distribution with zero means, unit variances and correlation ρ.

45. Consider the bivariate Poisson lognormal model given in exercise 44 but assume that the sampling effort for the two communities are ν_x and ν_y, that is, N_x and N_y are Poisson distributed with means $\nu_x e^X$ and $\nu_y e^Y$ when conditioned on the abundances. Show that the distribution of (N_x, N_y) is now the bivariate Poisson lognormal distribution $q(n_x, n_y; \mu_x + \ln\nu_x, \mu_y + \ln\nu_y, \sigma_x^2, \sigma_y^2, \rho)$.

46. Generalize the sampling model in exercise 44 by assuming that N_x conditioned on the abundance X and a variable V is Poisson distributed with

mean $V\nu_x e^X$, where $\ln V$ are normally distributed with mean $-\theta^2/2$ and variance θ^2 so that $\mathrm{E}V = 1$. The sampling variable V takes independent values for each Poisson sampling. With the same assumptions for N_y, show that the bivariate distribution of (N_x, N_y) then is the Poisson lognormal distribution $q(n_x, n_y; \mu_x + \ln \nu_x, \mu_y + \ln \nu_y, \sigma_x^2 + \theta^2, \sigma_y^2 + \theta^2, \rho^*)$, where

$$\rho^* = \frac{\rho\sigma_x\sigma_y}{\sqrt{(\sigma_x^2 + \theta^2)(\sigma_y^2 + \theta^2)}}.$$

47. Use the result in exercise 42-45 to derive equation (7.16).

48. In the one-dimensional case we usually consider the species abundance distribution truncated by omitting the zero class since the number of species in this class is unknown. Discuss how unobserved species should be treated in the estimation procedure in the two-dimensional case.

49. Derive equation (7.18).

50. Derive the expression for Sørensen's index when the standardized log abundances for two communities (U, V) has the standard binormal distribution with correlation ρ.

Hint: Use $P(U > \alpha, V < \beta) = P(U > \alpha, -V > -\beta)$ and the fact and the symmetry property that $(U, -V)$ is standard bivariate normal with correlation $-\rho$.

51. A classical result for the standard bivariate normal distribution, due to W.F. Sheppard in 1898, is that the probability of an observation taking a value in the first quadrant is $1/4 + \arcsin(\rho)/(2\pi)$. Use this result to find an analytic expression for Sørensen's index for $p = q = 1/2$.