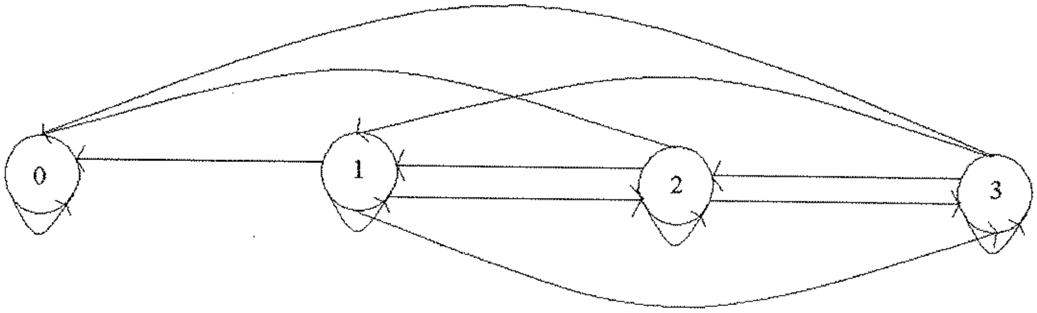
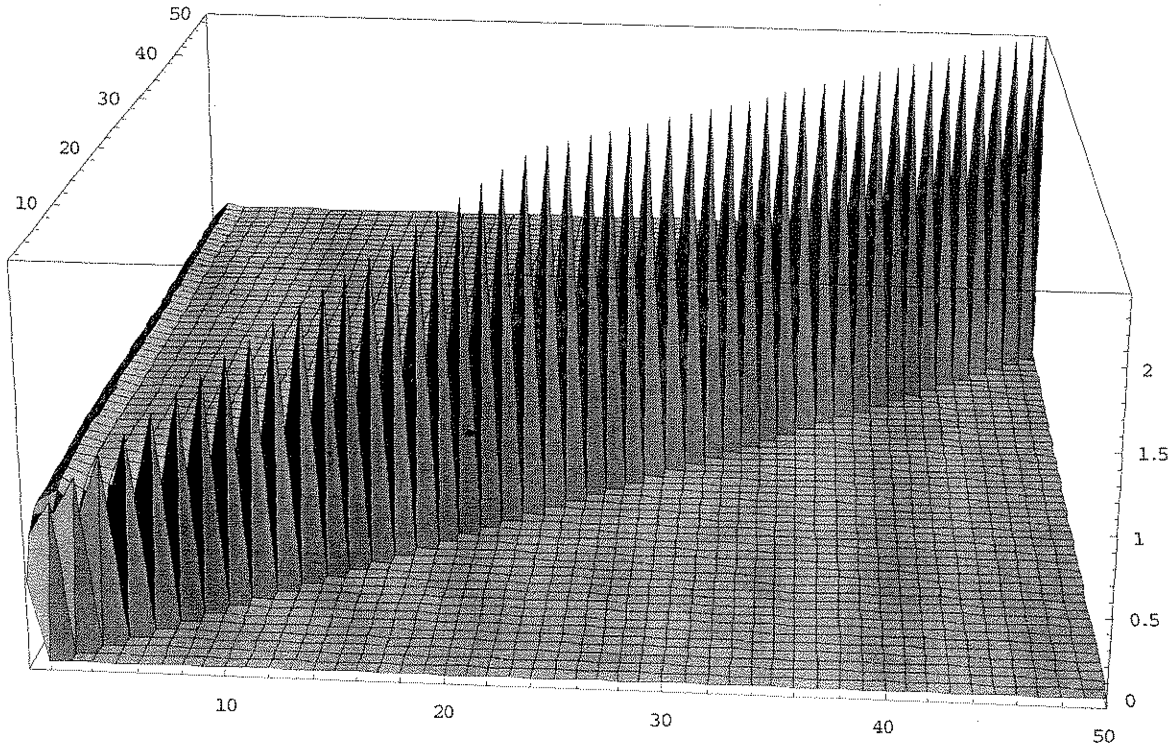
Jacob C. Warren
Georgia Southern University

Northern Arizona University REU 2001
Project Advisor: Terence Blows, Ph.D.



The Impact of Density Dependence upon Quasi-Extinction Risk in Markov Chains



Population modeling is simple in concept, solely concerned with predicting future population sizes. It is, however, very difficult to implement in such a way as to accurately predict the seemingly unpredictable hand of nature. Population models are implemented in forms from statistical regression to differential equations. No matter what model is offered, however, due to the laws of biological entropy there will always be a way to improve it.

Population models can be discrete, i.e. measured in finite steps (breeding cycles, usually) or they can be continuous in nature. Within both branches, there are two major types of models: density independent and density dependent. If a model is density independent, the growth rate (the multiplier of the current population that predicts the next timestep's population) does not depend upon the current population size. The classic example of density independent growth is exponential growth. In a density dependent model, the growth rate *does* depend upon the population size. Most biological research shows density dependent growth, with the most famous exception being the human population which exhibits near-perfect exponential growth. It is generally accepted that a density dependent model has greater potential to correctly model the natural world.

Widely used in discrete mathematical modeling are Markov Chains, with applications from population modeling to music.[i] The basic construction of a Markov Chain is as follows: a set of states (representing population sizes, age-classes, musical notes, etc.) conjoined with a transition matrix, **M**, such that $\mathbf{M}_{ij}$ represents the probability of ending in state i, given the process started in state j. This paper will consider the application of Markov Chains to population modeling; therefore, $\mathbf{M}_{ij}$ in this paper will represent the probability that a population of initially j breeding females will become a population of i breeding females after one discrete timestep. Only females are considered due to the fact that the female population is more crucial to the perpetuation of the species than the male. One major assumption in all proceeding research is that there are always enough males present to mate with the j females. As the main application of Markov Chains to population modeling is in wildlife management, this assumption, though restrictive, is not unreasonable.

Because Markov chains represent probabilistic processes, it is necessary for the probabilities to sum to 1. Thus, since each column of **M** represents a complete probabilistic process, each column of **M** must sum to 1.

Traditional construction of Markov Chains for population modeling has been recursive in nature, defined as:

Equation 1

$$M_{ij} = \sum_{k=0}^{n} M_{k1} M_{i-k\,j-1}$$

where there are n states. Thus, $\mathbf{M} \in \mathbb{R}^{+(n+1 \times n+1)}$. Markov chains constructed by this formula are density *independent*. One of the later foci of this paper will be to examine the impact of building two famous density dependences (both Beverton-Holt and Ricker) into the Markov model.

A very important observation is that due to biological law state 0, the extinction state, is an absorbing state. This simply means that once in state 0, the probability of staying in state 0 is 1, and hence the probability of going from state 0 to any state other than itself is 0. State n represents a population ceiling; it is assumed that the environment cannot support a population of greater than n breeding females. Any females born above n die before the next breeding season. This reduces the functional population down to n. Thus, Equation 1 is modified with Equation 2, the condition that

Equation 2

$$M_{nj} = 1 - \sum_{k=0}^{n-1} M_{kj}$$

Another important observation is that raising **M** to powers represents timesteps, and thus $\mathbf{M}^t_{ij}$ represents the probability that a population initially of j females will, after t timesteps, have i females.

The application under consideration in this paper is to quasi-extinction risk, defined as the probability that the population will go extinct after t years given it had a certain initial population. The obvious biological application is to wildlife management. A particularly suited case is endangered species which are by definition assumed to tend to extinction without intervention.

The quasi-extinction risk at time t given the population started in state j can be formulated as:

Equations 3

$$QE(t, j) = M^t_{0j}$$

or

$$QE(t, j) = (M^t * V_j)_0$$

where V_j represents a zero vector with a 1 in the jth position, representing the initial population size of j.

It has been established that 0 is an absorbing state. One question that begs to be answered, especially by the wildlife manager, is whether or not there are other possible absorbing states. Before that question is addressed, to get a general feel for Markov chains, examine a simple Markov model. It will generally be assumed, unless otherwise stated, that a population begins in the maximal population state, n. This initial condition vector will be referred to as V*.

The first two columns of a Markov Matrix constructed via Equations 1 & 2 must be calculated manually. The first column is always a zero vector with a 1 in the zeroth position. This represents the absorbing state.

A particular species will have to be examined to get numerical entries for the rest of the matrix. Assume that for the species under consideration, each female has a breeding area of 2000 units, and will not tolerate another female in her breeding area. In any given year, a female has a 0.4 chance of dying. She has a 0.35 chance of having no female offspring, a 0.55 chance of having one female offspring, and a 0.1 chance of having two female offspring. Define these assumptions as **Parameters A**. Assume a forest of 10000 units for population. What is the quasi-extinction risk after 25 years if the forest is initially stocked with five females of breeding age and enough mates to perpetuate the species? This seemingly complicated question is easily answered with a Markov Chain. All calculations and simulations in this paper were run using Mathematica v4. Sample code for a general Markov Matrix is contained in Appendix A.

It is assumed that for all species under consideration, at most two babies can be born to a female in each breeding cycle. The second column is generated via probability theory as follows:

Equations 4

$$M_{01} = \text{pNoBabies} * \text{pDie} + \text{pOneBaby} * \text{pDie}^2 + \text{pTwoBabies} * \text{pDie}^3$$

$$M_{11} = \text{pNoBabies} * (1 - \text{pDie}) + 2 * \text{pOneBaby} * \text{pDie} * (1 - \text{pDie}) + 3 * \text{pTwoBabies} * (1 - \text{pDie}) * \text{pDie}^2$$

$$M_{21} = \text{pOneBaby} * (1 - \text{pDie})^2 + 3 * \text{pTwoBabies} * \text{pDie} * (1 - \text{pDie})^2$$

$$M_{31} = \text{pTwoBabies} * (1 - \text{pDie})^3$$

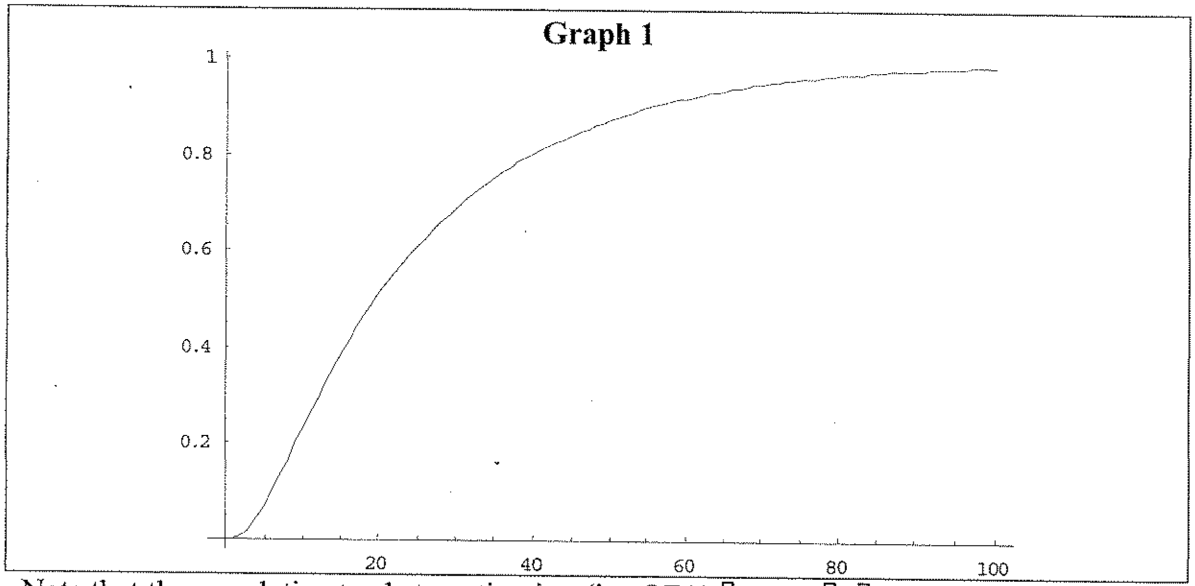
The Markov matrix generated by the above conditions is:

$$\begin{pmatrix} 1 & 0.2344 & 0.0549434 & 0.0128787 & 0.00301877 & 0.0007076 \\ 0 & 0.5028 & 0.235713 & 0.0828766 & 0.0259017 & 0.00758919 \\ 0 & 0.2412 & 0.365882 & 0.217531 & 0.0957661 & 0.0361991 \\ 0 & 0.0216 & 0.252677 & 0.301234 & 0.200252 & 0.101403 \\ 0 & 0 & 0.0798984 & 0.239116 & 0.261768 & 0.185703 \\ 0 & 0 & 0.0108864 & 0.146363 & 0.413294 & 0.668398 \end{pmatrix}$$

Raising the matrix to the 25th power yields:

$$\begin{pmatrix} 1. & 0.846135 & 0.740559 & 0.671709 & 0.630525 & 0.608507 \\ 0. & 0.0210235 & 0.0354179 & 0.044791 & 0.0503922 & 0.053385 \\ 0. & 0.0261827 & 0.0441269 & 0.0558194 & 0.0628097 & 0.0665458 \\ 0. & 0.0271382 & 0.0457569 & 0.0578977 & 0.0651596 & 0.0690419 \\ 0. & 0.0265433 & 0.0447666 & 0.0566556 & 0.063769 & 0.0675727 \\ 0. & 0.0529776 & 0.0893726 & 0.113127 & 0.127344 & 0.134948 \end{pmatrix}.$$

Thus, the quasi-extinction risk after 25 years is .608507. Plotting risk versus year yields the following graph:


Graph 1

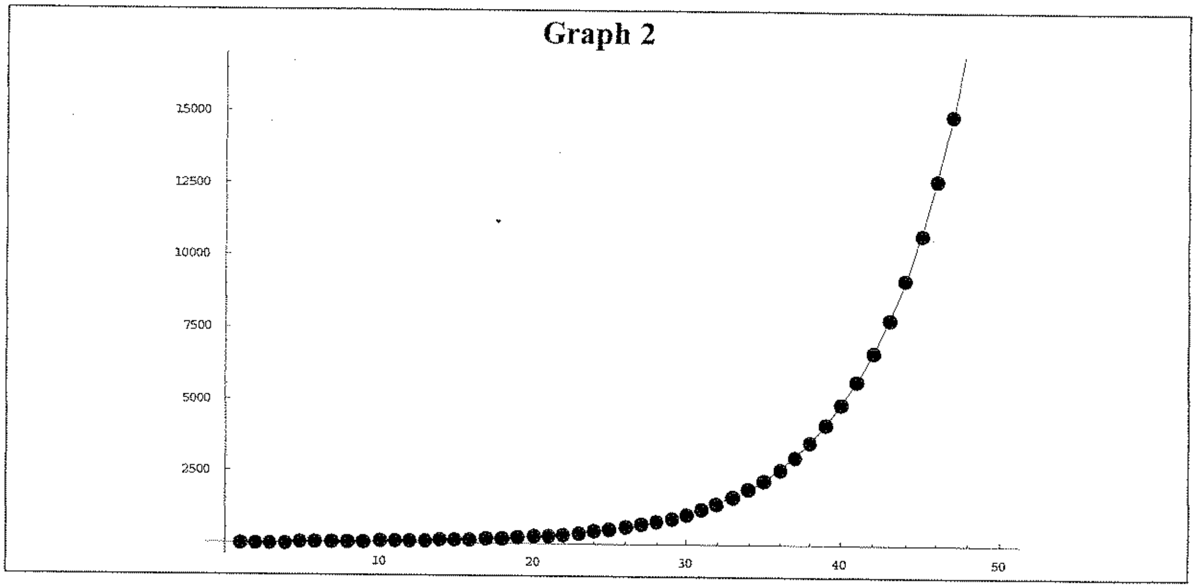
Note that the population tends to extinction (i.e. QE(t) □ 1 as t □ □). Because these values were generated using very specific assumptions on the parameters, now offered is an analysis of varying these parameters.

There are a number of benchmarks that can be used to measure the speed of increase of quasi-extinction risk. For simplicity's sake, define H to be the year in which the quasi-extinction risk rises above .50 for the first time. For Parameters A with n=5, H=20. Thus, after 20 years, it is more probable than not that the species under observation will have become extinct.

Varying only the population ceiling, n, yields:

n	H
5	20
10	55
20	245
30	1046
40	4859
50	24272

As should be expected, as you raise the population ceiling, H increases. Following is a graph of H vs n:



Graph 2

Another possible source of great variation lies in the birthrates. Only extreme cases were examined as it is safely assumable that the correlation is smooth. Varying the birthrates but leaving the growthrate (E[number of births]) the same, and varying n yields:

pNoBabies=.25
pOneBaby=.75
pTwoBabies=0

n	H
10	67
20	355
30	1901
40	11310
50	72529

pNoBabies=0.625
pOneBaby=0
pTwoBabies=.375

n	H
10	36
20	121
30	353
40	1050
50	3290

pNoBabies seems to be the most important parameter, the higher it is, the lower the H value. Regardless, if only growthrate data are available instead of pOneBaby and pTwoBabies values, merely run scenarios in the extreme cases for a feasible region.

In most matrix population models, growth can be observed by examining eigenvalues. Is the same true for **M**? Define the Perron root of a matrix as its dominant eigenvalue (in absolute value or modulus). Observe that if one reduces any Markov matrix of construction from Equations 1 and 2, the 1 in the 0,0 position will remain dominant. Thus, 1 is the Perron root of **M**, regardless of parameter values. Its eigenvector is V_0.

Recall from linear algebra that for any n by n matrix with linearly independent eigenvectors, the eigenvectors form a basis for $\mathbb{R}^n$. Therefore (assuming **M** has linearly independent eigenvectors, which for most parameter choices is true) there is a linear combination of the eigenvectors that will equal the initial condition vector V_j. This reduces to the matrix equation:

Equation 5

$$V_j = P\left(\bar{c}\,\bar{\lambda}\right)$$

where P represents the matrix of eigenvectors (in correspondence to decreasingly absolute dominant eigenvalues), c the coefficient vector, and □ the vector of eigenvalues in decreasing absolute dominance. Since the Perron root is always 1 and its eigenvector has a value, 1, solely in its primary position, Equations 3 and 5 combined yield:

Equation 6

$$QE(t) \approx c_1 + c_2\,\lambda_2{}^t\,X_{2_0} + c_3\,\lambda_3{}^t\,X_{3_0} + \ldots + c_{n+1}\,\lambda_{n+1}{}^t\,X_{n+1_0}$$

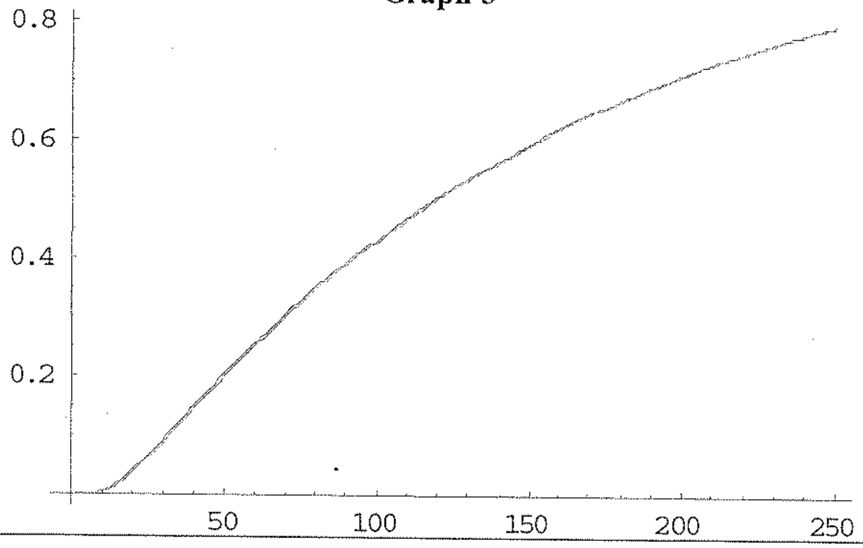
By left-multiplying V_j with the inverse of P and then dividing out (pointwise) the eigenvalue vector, one can solve for the c-vector:

Equation 7

$$\bar{c} = \left(P^{-1} * V_j\right) ./\, \bar{\lambda}$$

Using this c-vector to approximate quasi-extinction risk via Equation 6 versus time with Parameters A and n=15 yields the following graph when combined with the "real" quasi-extinction graph:

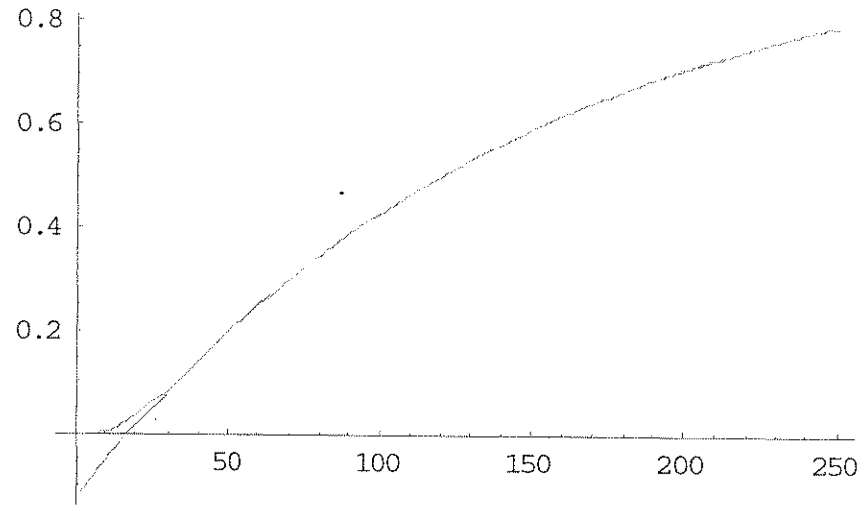
Graph 3



As can be seen, the approximation is very good, and this continues over into other choices for parameters.

Can this approximation be pared back? For example, can one merely consider the second-most absolutely dominant eigenvalue, and its c-value, in conjunction with c1? Such an approximation yields the following graph:

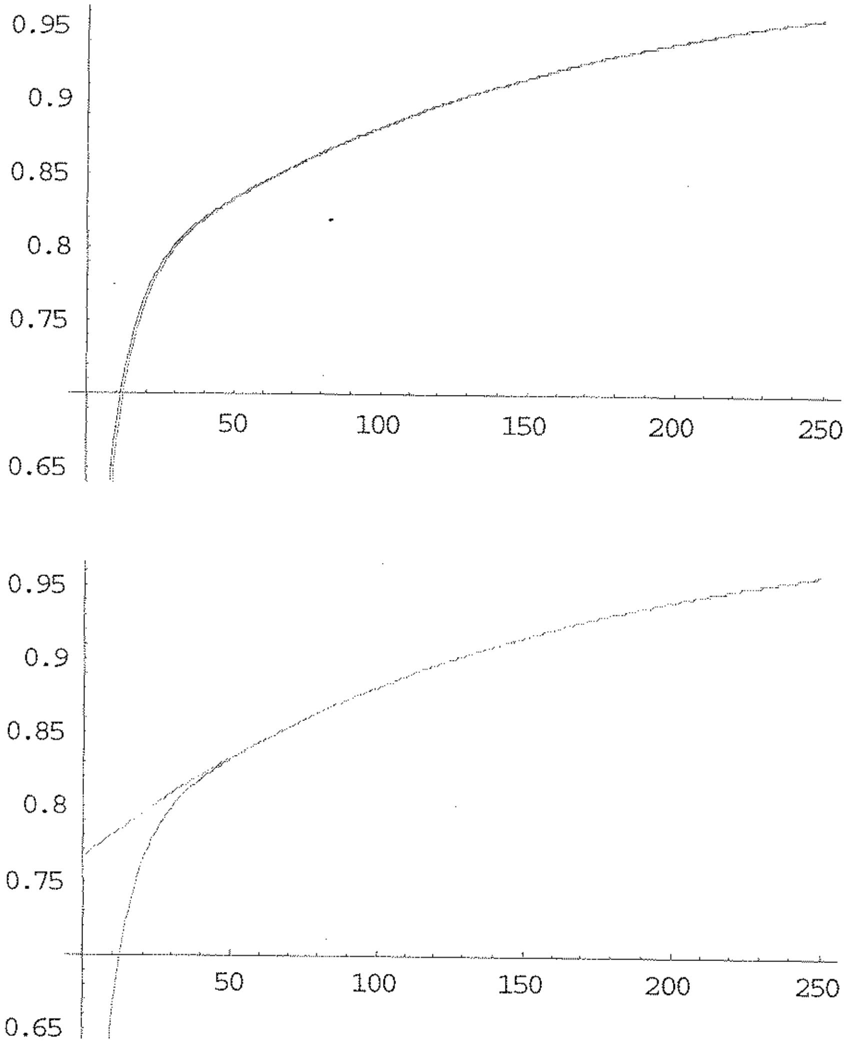
Graph 4



Though the approximation is slightly off at first, as time increases, the individual curves become indistinguishable.

Does this extend to initial conditions other than V*? Graphs with the same set-up as before, but with initial condition V_1, yield:

Graphs 5



Again, initial approximation is bad, but the curves again merge. This leads to the conclusion that the convergence to the extinction rate is driven by the second most absolutely dominant eigenvalue, referred to heretofore as λ_2. As λ_2 increases, quasi-extinction risk decreases. Therefore, if it could be shown that $\lambda_2 \to 1$ as $n \to \infty$, this would show that populations could theoretically be saved by raising n above some threshold.

Unfortunately, though $\lambda_2 \to 1$ as $n \to \infty$, λ_2 will never equal 1, therefore there is no threshold. The proof lies in matrix theory.

Define M` to be the matrix formed when the zeroth row and column of M are removed. Notice M` is no longer a Markov matrix. Also observe that $\square_2$ is the Perron root of M`. When dealing with M and M`, it is important to classify them in traditional nonnegative matrix terminology: it needs to be known if M and M` are irreducible or not. In order to do so, recall a bit of elementary graph theory.

Definition 1
A directed graph, D, is said to be strongly connected if for any ordered pair of distinct vertices, i and j, there is a path in D connecting i to j.

Note that any M with an absorbing state represents an adjacency matrix for a graph that is NOT strongly connected: there is no transition from the 0 state to any other. An important theorem directly yields the result that while M` is irreducible, M is not.

Theorem 1
A nonnegative matrix is irreducible if and only if the associated directed graph is strongly connected.

Proof of these and all other matrix theory theorems used in this paper are available.[ii]

With the result that M` is irreducible, many theorems are available for restricting $\square_2$. The most important is:

Theorem 2
Let M be a nonnegative matrix with maximal eigenvalue r. Define r_i to be the sum of row i of M, and c_i to be the sum of column i of M. Define □ = min{r_i} and R = max{r_i}. Also define □ = min{c_i} and C = max{c_i}. □ □ r □ R and □ □ r □ C. Equality holds if and only if □ = R or □ = C.

Let M be M` and r be $\square_2$. At first, one might think that herein lies the proof of a second absorbing state since all of the columns sum to 1. True, all of the columns of M sum to 1, but no such restriction is upon M`; thus, Theorem 2 proves that while $\square_2$ might tend to 1, it will never equal 1.

All analysis up to this point has been of density independent growth. As was discussed in the beginning, density independent growth is biologically unlikely. How, then, can Markov Chains be made density dependent?

Density dependence is formally defined as follows.

Definitions 2

Let N_t represent the population size at timestep t. A population model is density **dependent** if it is of the form:

$$\mathbf{N_{t+1} = f(N_t)N_t}$$

Similarly, a population model is density **independent** if and only if it is of the form:

$$\mathbf{N_{t+1} = rN_t}$$

There are two classic density dependences, Beverton-Holt and Ricker, that provide different definitions of $f(N_t)$. The nonlinear Beverton-Holt iterated map corresponds to the logistic equation in continuous models. Beverton-Holt density dependence assumes a stable carrying capacity K, i.e. any population beginning below K will increase to K, and any population above K will reduce to K. If the population hits K, it stays there. As the population grows towards K, the growth rate decreases linearly, starting at □ (assumed to be greater than 1) and ending at 1. This is summed up in the equation:

Equation 8

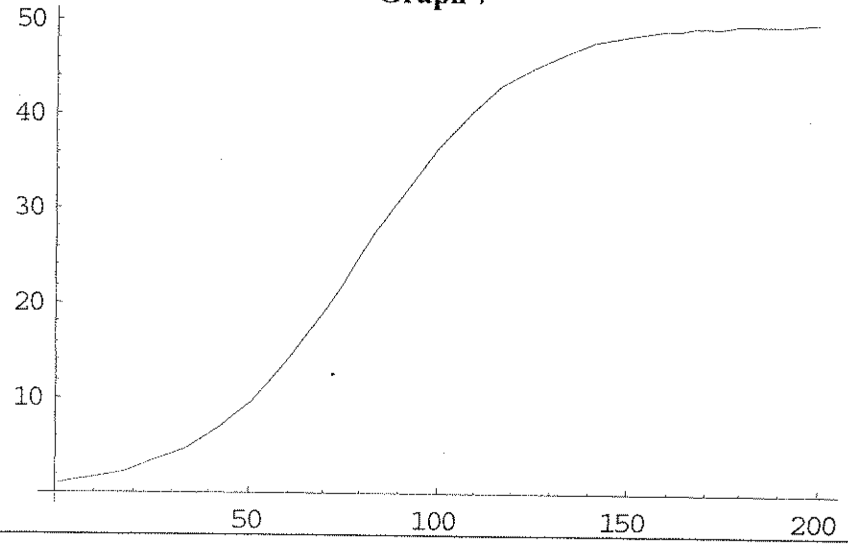
$$N_{t+1} = \frac{\lambda N_t}{1 + \frac{\lambda - 1}{K} N_t}$$

Most nonlinear iterated maps cannot be solved, but the solution to the Beverton-Holt equation is:

Equation 9

$$N_t = \frac{K}{1 + \frac{K - N_0}{N_0} \lambda^{-t}}$$

The general shape of a Beverton-Holt population curve is the same as that of the logistic:



Graph 7

The Ricker Equation is constructed similarly to Beverton-Holt, but with one very different assumption, that N_{t+1} □ 0 as N_t □ □ . This is a quasi-example of a scramble density-dependence, one in which the population grows until it reaches the "carrying capacity" which cannot carry it, so the population immediately goes extinct. The species scrambles for the very limited resources, and as a result of overpopulation, resources are exhausted and nothing is left to sustain the next year's population. The Ricker equation is a scramble dependency only in the limit. The Ricker equation is defined as:

Equation 10

$$N_{t+1} = N_t \, e^{r\left(1-\frac{N_t}{K}\right)}$$

where r represents initial birth rate.

How, though, to build these dependencies into a Markov Chain when Markov Chains are driven off birth and death rates, not growth rates, and the "carrying capacity" of n is unstable? In each case, the equations can be broken into the growth rate functions as follows:

Equations 11

$$\text{BevertonHoltGrowthRate}\,(N_t) = \frac{\lambda}{1 + \frac{\lambda - 1}{K}\,N_t}$$

$$\text{RickerGrowthRate}\,(N_t) = e^{r\left(1 - \frac{N_t}{K}\right)}$$

The growth rate for any population is (1+b)(1-d). By letting b=pOneBaby+2pTwoBabies there is now a method for transferring data from the density dependence into the Chain.

Assume that in the population, birthrates are constant. As the growth rate increases and decreases, the death rate decreases and increases respectively. By forcing equality in the t = 0 case, since there are already assumptions on the values for pDie and b, one can solve Equations 11 for □ and r. This yields:

Equations 12

$$\lambda = -\frac{(1 + b) * (-1 + \text{pDie}) * (-1 + K)}{-1 + b * (-1 + \text{pDie}) + \text{pDie} + K}$$

$$r = \frac{\ln[(1 + \text{pOneBaby} + 2 * \text{pTwoBabies}) * (1 - \text{pDie})]}{1 - \frac{1}{K}}$$

Now, since it is assumed that the death rate varies to fit the growth rate, one can solve for the death rates at population N as follows:

Equations 13

$$\text{BevertonHoltDeathRate}\,(N) = 1 - \frac{\frac{\lambda}{1 + N\,\frac{\lambda - 1}{K}}}{1 + b}$$

$$\text{RickerDeathRate}\,(N) = 1 - \frac{e^{r\left(1 - \frac{N}{K}\right)}}{1 + b}$$

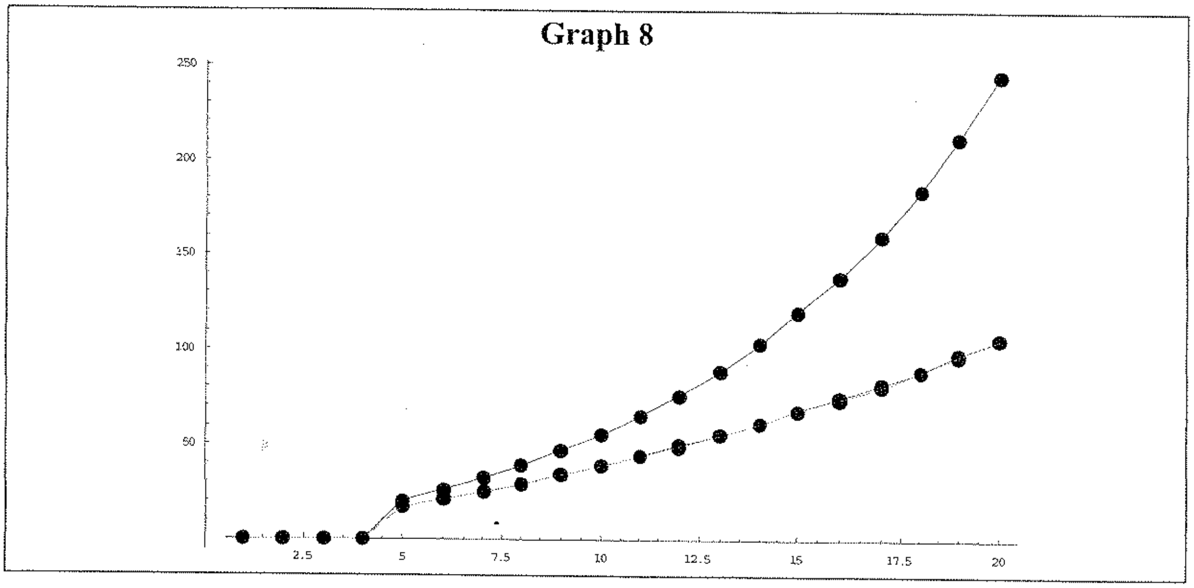
Now, it is easy to build in the density dependence--vary the death rates according to Equations 13 for each column.

Since, unlike density independent Markov Chains, the assumptions behind Beverton-Holt and Ricker allow the functional population to rise above K, scenarios were run in which in the density dependent cases populations were allowed to rise up to 2n, 1.5n, and n. Quasi-extinction risks were lower for the density dependent scenarios than the density independent in the 2n and 1.5n cases, as should be expected. It was decided, however, to analyze in-depth only the case in which the population ceiling was a true ceiling. This remains true to the assumption that any females born above the ceiling die before the next breeding cycle.

For Parameters A and n = 10, the following H values result:

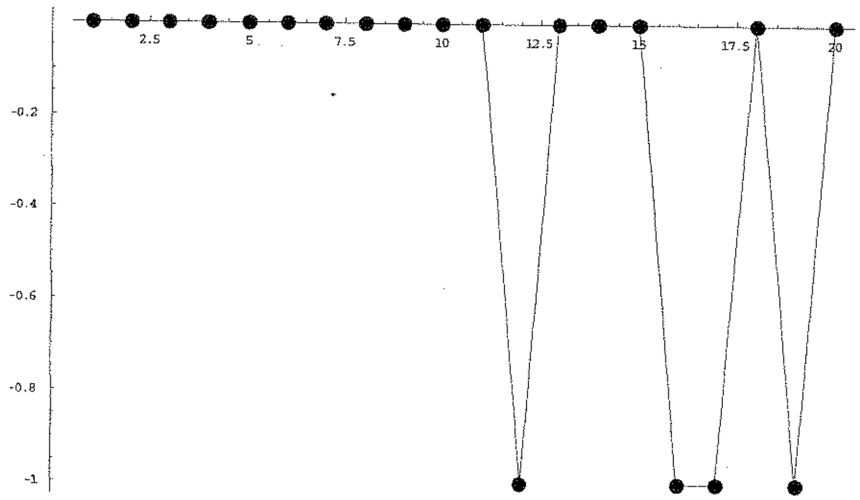
Density Dependence	Independent	Beverton-Holt	Ricker
H value	55	39	39

Plotting H curves for Parameters A and varying n from 1 to 20 yields:


Graph 8

The Beverton-Holt and Ricker H curves, on the bottom, are nearly identical, only ever differing by one. Graphing the difference of the two H values vs increasing n:

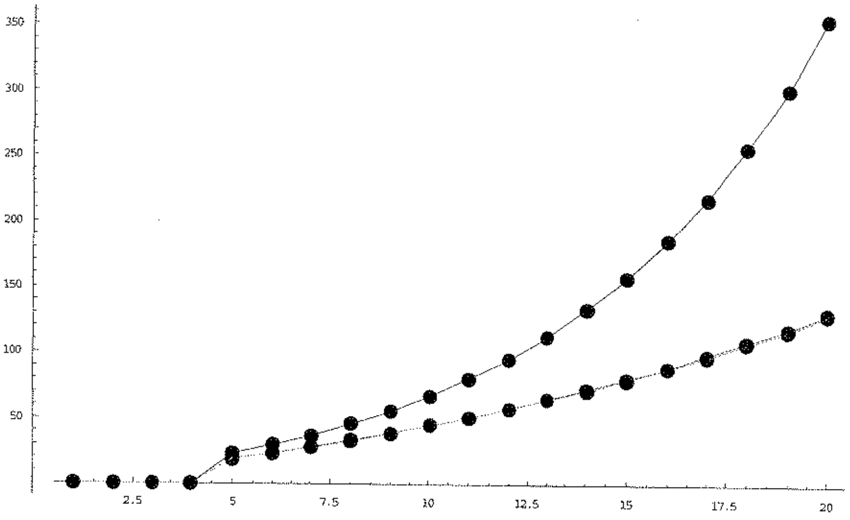
Graph 9

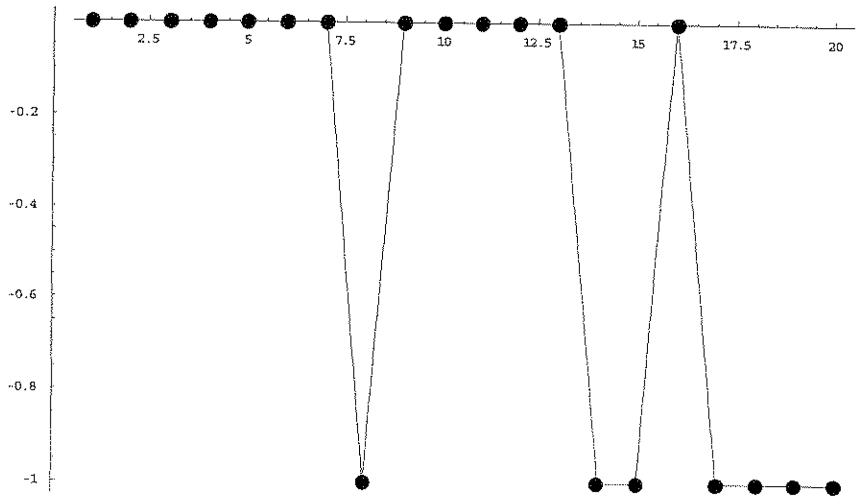


Varying parameters as before, but leaving the growth rate constant yields:

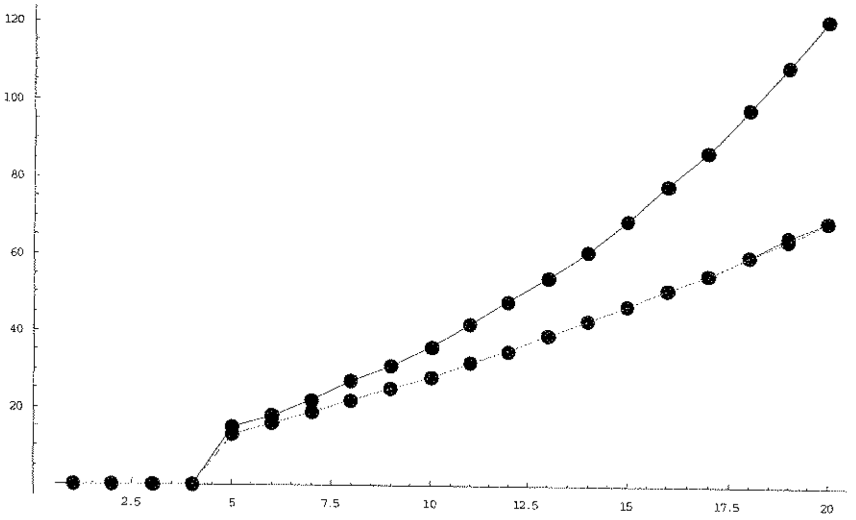
Graphs 10

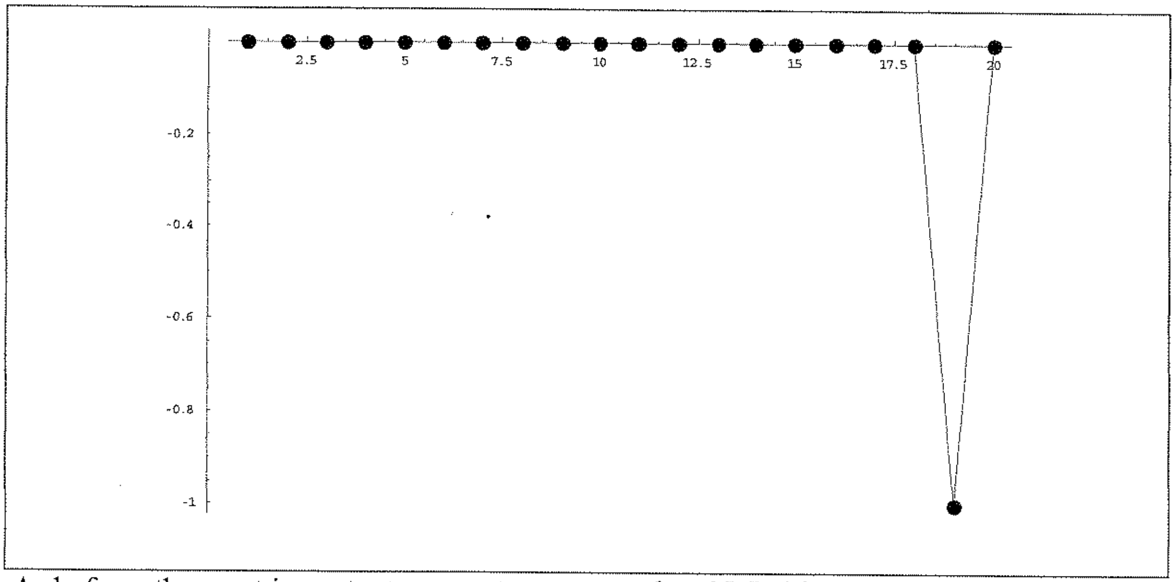
pNoBabies=.25
pOneBaby=.75
pTwoBabies=0





pNoBabies=0.625
pOneBaby=0
pTwoBabies=.375



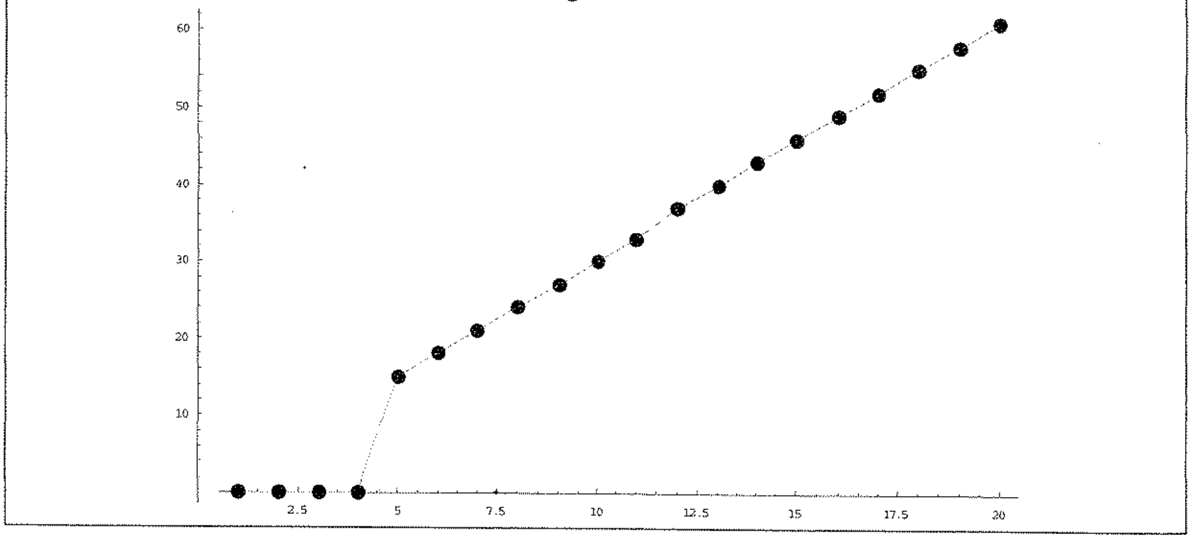


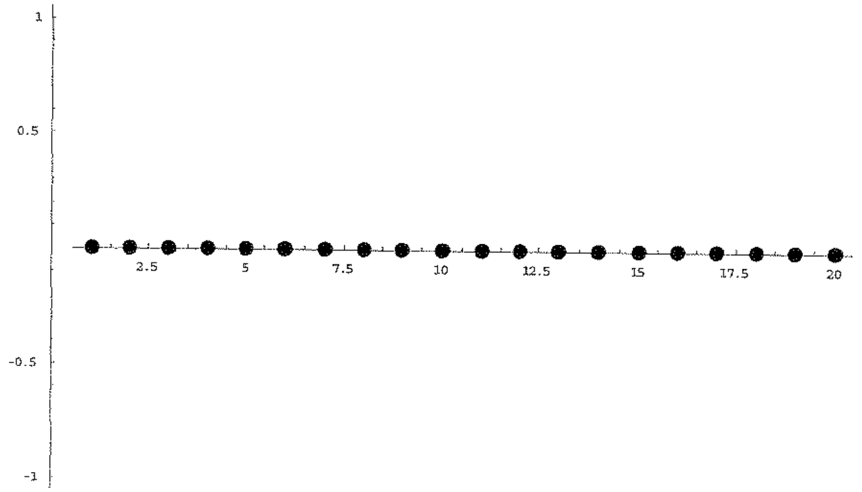
As before, the most important parameter seems to be pNoBabies.

What happens, though, if the growth rate, g, is varied as well? As the growth rate is increased, the change is assumed to be reflected in pOneBaby and pTwoBabies by the originally assumed ratio: .55/.75 going to pOneBaby. If pOneBaby rises above 1, it is reduced and pTwoBabies is increased until all parameters are within the range from 0 to 1. Notice the growth rate cannot be greater than 2, due to the fact that at most two babies can be born to a female each breeding cycle. Following are graphs for g=1 and g=1.1.

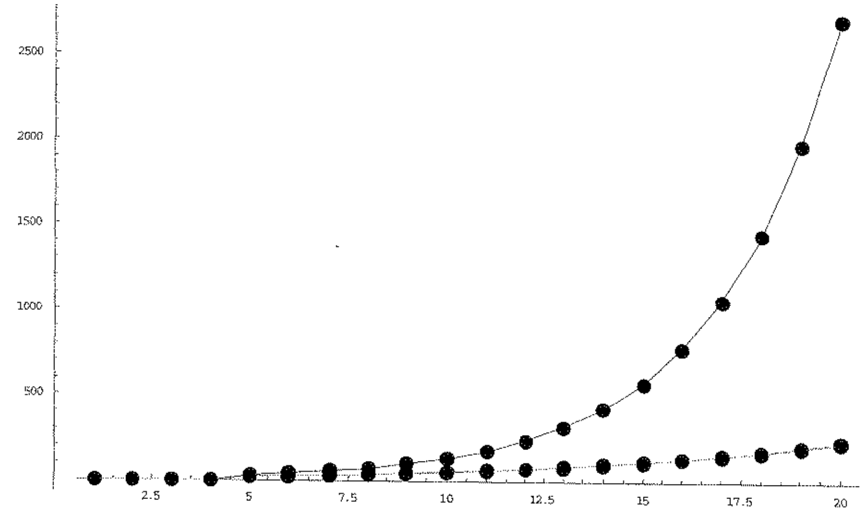
Graphs 11

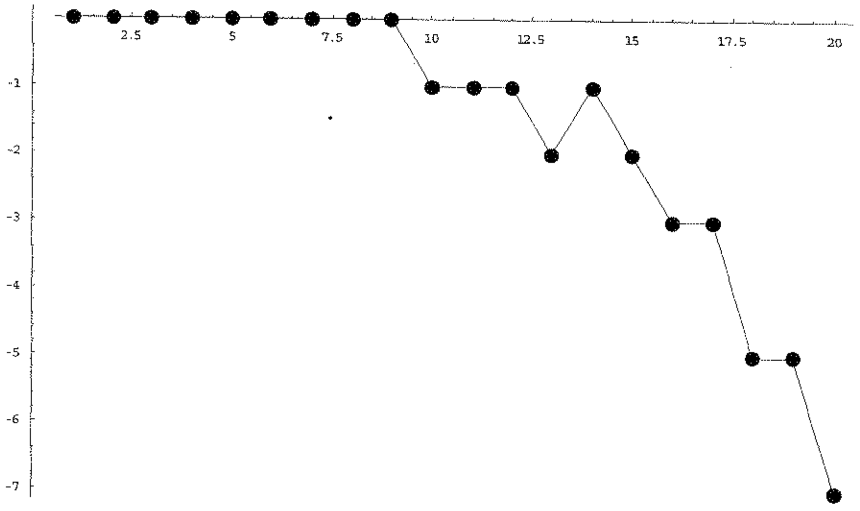
$g = 1$




1
0.5
-0.5
-1
2.5
5
7.5
10
12.5
15
17.5
20

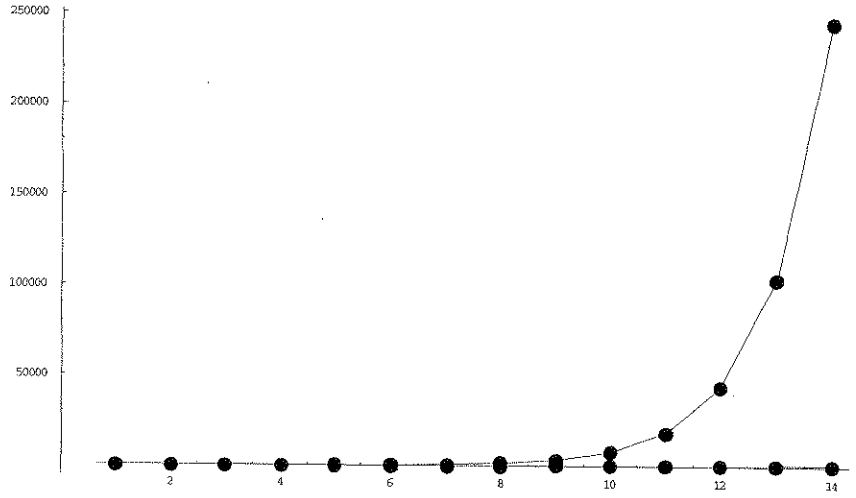
g = 1.1


2500
2000
1500
1000
500
2.5
5
7.5
10
12.5
15
17.5
20


2.5
5
7.5
10
12.5
15
17.5
20
-1
-2
-3
-4
-5
-6
-7

For g=1, since there is nowhere for the density dependences to taper off to, they are identical to the density independent values. The difference becomes more pronounced as g increases. As is expected, H values increase as g increases. At first scenarios were run for g values 1.25, 1.5, etc., but in the original Markov Chain case, H values skyrocketed off towards infinity:

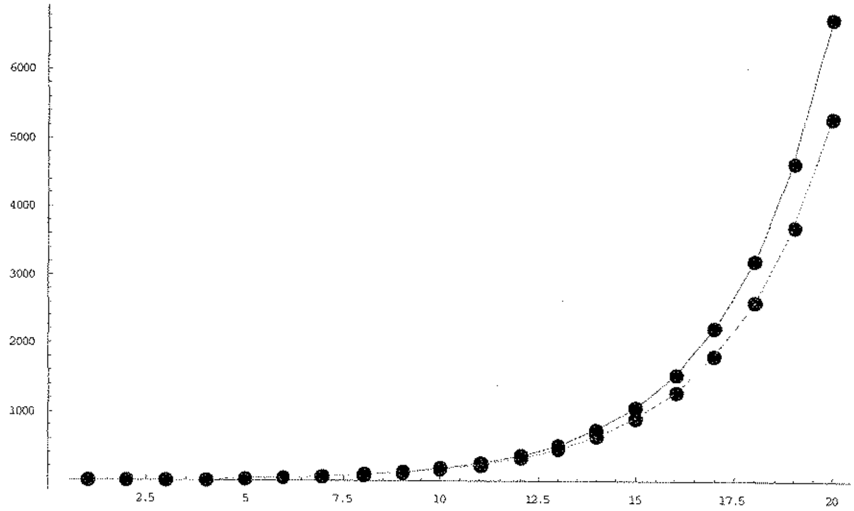
Graph 12
g = 1.25



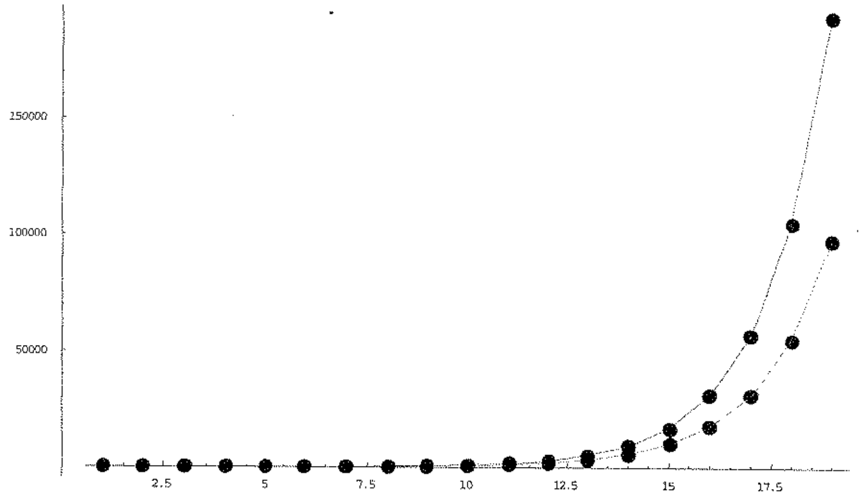
As can be seen, calculation at high g values for the original Markov H values becomes difficult to complete. At g = 1.25 and n=14, for example, H=245420. To greatly speed up calculation, as higher g values make the gap more pronounced, the density independent case was thrown out of the simulations. Following are graphs of how the Ricker and Beverton-Holt differ for various g values. In each case, Ricker predicts a higher H value, and thus Beverton-Holt yields a more pessimistic view of the extinction chances.

Graphs 13

g = 1.25



g = 1.5



The difference for g = 1.5 and n=19 is approximately 80000. g values above 1.5 force parameters outside of possible ranges. As would be expected, as the growth rate increases, the two density dependences grow further and further apart. The reason for this is obvious if one re-examines the definition of the growth rates in Equations 13.

In conclusion, for all growth rates, density dependence makes a difference, and this difference becomes more prominent as n and g increase. Also, for small g values, Ricker and Beverton-Holt predict very similar quasi-extinction curves, with the Beverton-Holt density dependence yielding higher quasi-extinction risks. As g increases, the impact of the density dependencies becomes more and more pronounced, predicting much more severe H values. Again, also as g increases, the two density dependent curves are forced apart. This is also true for increasing n values. Beverton-Holt predicts more pessimistic quasi-extinction curves. Therefore, if concerned with worst-case scenarios, use a Beverton-Holt density dependence.

Thus Markov Chains support the need for wildlife management. Since humankind infringes increasingly upon the habitat of species, the available breeding area for these animals is continually shrinking. Thus, many species are entering the dangerous scenarios of small n values. Also, as pollution and other contaminants increase in the environment, g values continue to drop dangerously low. This was seen recently with the bald eagle population. Bald eagles were already limited in their breeding area when DDT was introduced. Too late it was realized that DDT caused eggshell fragility. This collapsed the g value for the bald eagle population. DDT is now banned in the United States, but not Mexico. As previously stated, density dependent growth is widely accepted as the biological standard for species not our own, so it is a better model. It predicts higher extinction risks; thus, wildlife management will continue to be necessary until we take a more responsible stance towards the environment.

Appendix A

Sample Mathematica Code

The matrix generated via this code must be transposed to fit the Markov definition of Equations 1 and 2.

```
<< LinearAlgebra`MatrixManipulation`
maxFemales = 5;
pNoBabies = .35;
pOneBaby = .55;
pTwoBabies = .10;
pDie = .4;
Clear[Q2, M2];
m = 0; n = 0;
While[m ≤ maxFemales, While[n ≤ maxFemales, Q2[m][n] = 0; n++]; n = 0; m++];
Q2[0][0] = 1;
Q2[1][0] = pNoBabies * pDie + pOneBaby * pDie^2 + pTwoBabies * pDie^3;
Q2[1][1] = pNoBabies * (1 - pDie) + 2 * pOneBaby * pDie * (1 - pDie) + 3 * pTwoBabies * (1 - pDie) * pDie^2;
Q2[1][2] = pOneBaby * (1 - pDie)^2 + 3 * pTwoBabies * pDie * (1 - pDie)^2;
Q2[1][3] = pTwoBabies * (1 - pDie)^3;
m = 1; n = 4;
While[n ≤ maxFemales, Q2[1][n] = 0; n++];
m = 2; n = 0;
                                                      n                                                         n-1
While[m ≤ maxFemales, While[n < maxFemales, Q2[m][n] = ∑ Q2[1][k] * Q2[m - 1][n - k]; n++]; Q2[m][n] = 1 - ∑ Q2[m][k];
                                                     k=0                                                        k=0
 n = 0; m++];
m = 0; n = 0;
M2 = ZeroMatrix[maxFemales + 1];
m = 0; n = 0;
While[m ≤ maxFemales, While[n ≤ maxFemales, Part[M2, m + 1, n + 1] = Q2[m][n]; n++]; n = 0; m++];
initialState = maxFemales;
OriginalMarkovExtinction[n_] := Part[MatrixPower[M2, n], maxFemales + 1, 1]
Plot[OriginalMarkovExtinction[n], {n, 1, 250}, PlotPoints → 250];
MatrixForm[Transpose[M2]]
```

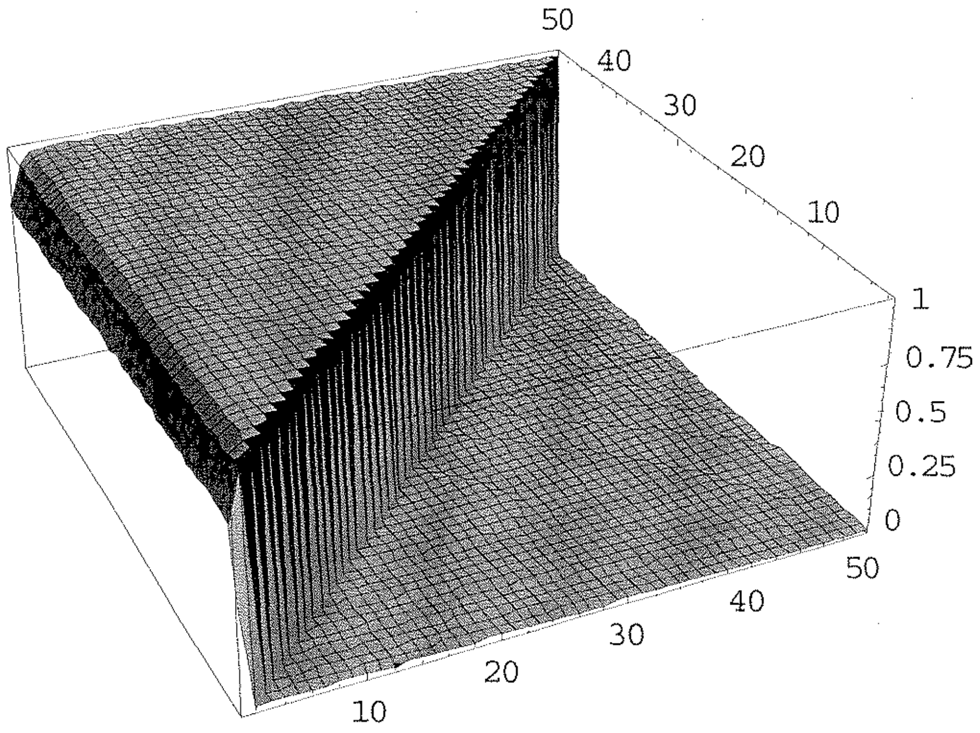
Appendix B

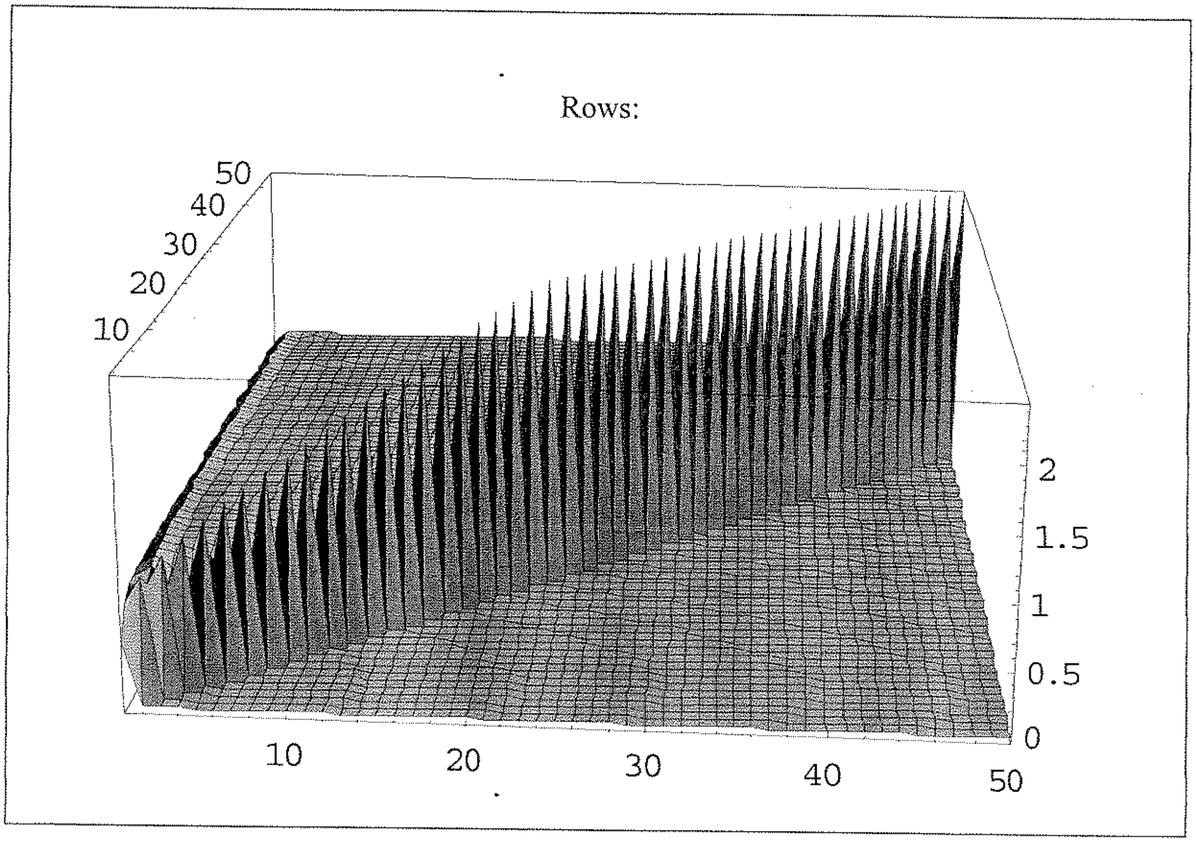
Theorems and their restrictions upon $\square_2$

If Theorem 2 is used, the following graph of bounds is generated with Parameters A and n varying from 1 to 50, the highest n value Mathematica can readily handle:

Graphs A

Columns:





The extreme values of the graph yield:

Restrictions A

$.7656 < \square_2 < 1$

$.709814 < \square_2 < 2.57995$

Another useful restrictive theorem is:

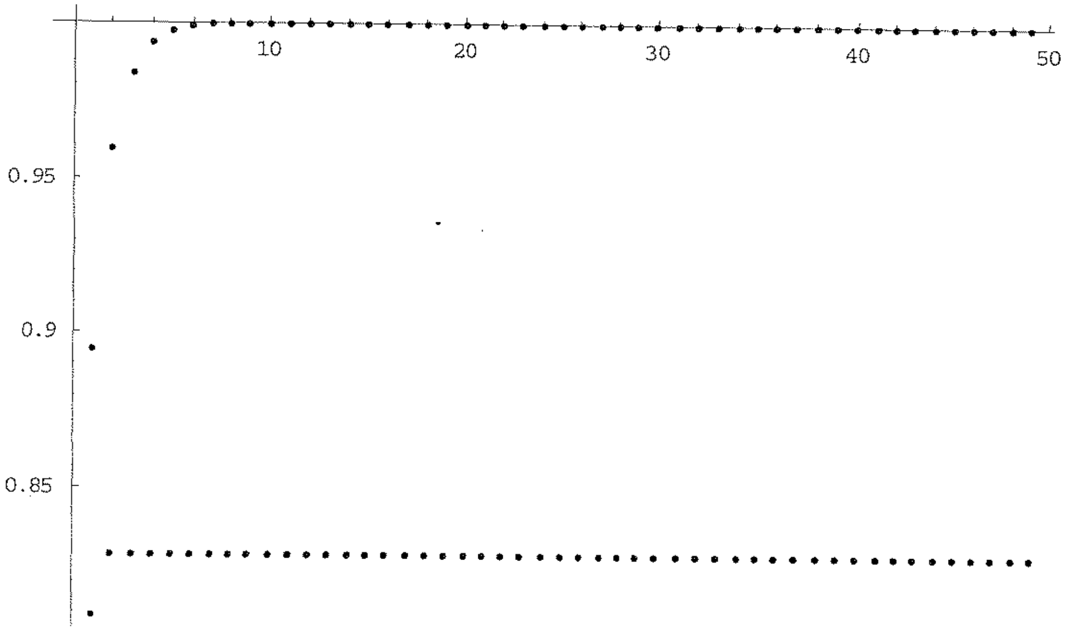
Theorem A

Define M, r, and r_i as in Theorem 2.

$$\min\left\{\frac{1}{r_i}\sum_{t=1}^{n} M_{it}\, r_t\right\} \le r \le \max\left\{\frac{1}{r_i}\sum_{t=1}^{n} M_{it}\, r_t\right\}$$

Again, using Parameters A and varying n over the feasible values yields:

Graph B



The restrictions at n=50 are:

Restriction B

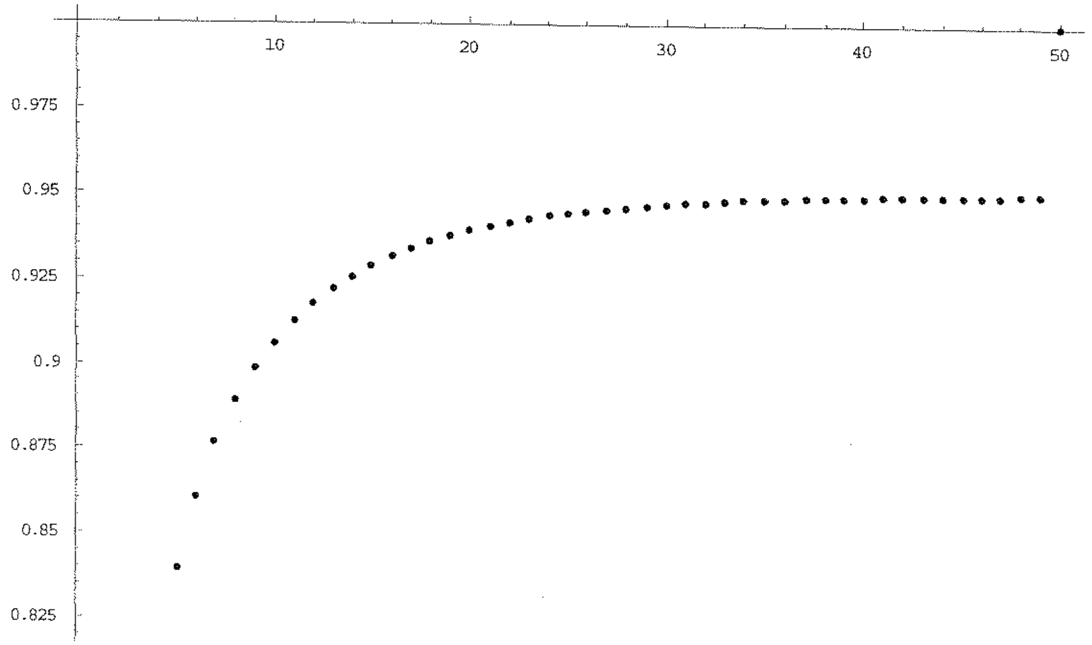
$.828387 \leq \lambda_2 \leq 1$

The third useful theorem is

Theorem B

The maximal eigenvalue of an irreducible matrix is greater than the maximal eigenvalue of any of its principal submatrices.

This yields the following graph:



Graph C

Yielding

Restriction C

$.950234 < \square_2$

Combining all Restrictions yields:

Restriction D

$.950234 < \square_2 < 1$

---

[i] http://freespace.virgin.net/kevin.perry/PCMusic/Papers/paper4.htm

[ii] Minc, Henryk. Nonnegative Matrices. John Wiley and Sons, NY. 1988