



# 28

## POPULATION VIABILITY ANALYSIS

*In collaboration with James Gibbs*

### Objectives

- Explore properties of the normal curve.
- Develop a Monte Carlo population model in which growth rates are stochastic.
- Estimate the probability of extinction for a population.

*Suggested Preliminary Exercises: Geometric and Exponential Population Models; Statistical Distributions*

### INTRODUCTION

Environmental variation dominates many ecological processes and interactions. Rainfall, for example, varies not only from year to year but also from place to place. Many natural populations have birth and death rates that are tied to some sort of environmental factor. A classic example is a species of Darwin's finch, *Geospiza fortis*, which inhabits the Galápagos Islands. Among *G. fortis* populations, the number of births and deaths on an island can be roughly predicted from the amount of rainfall the island receives (Grant 1986). In periods following droughts, death rates are high and birth rates are low. If drought conditions persist for long periods, the species may be in danger of extirpation or even extinction. If conditions reverse, however, the population may respond with higher growth rates and subsequently increase in number.

Population ecologists are often interested in addressing the question "What is the smallest the population can be and still have a reasonable probability of surviving for some time into the future?" The technique used to answer this question is called a **population viability analysis**, or **PVA**. The results of a PVA can be used to make statements such as "A *Geospiza fortis* population needs to have at least  $x$  individuals if it is to have a 95% probability of surviving for 500 years, given current conditions," or "A population of 100 *Geospiza fortis* has an  $x\%$  probability of surviving 500 years if current conditions persist."

What data are needed to develop a PVA? Conceptually, a PVA would include all the random variation in parameters that measure habitat quality such as climate, nutrients, water, cover, pollutants, and relationships with other species that might be prey, predators, competitors, parasites, and pathogens. Obviously, this is a tall order! Although a PVA can be quite complicated, most PVAs include only one or two factors that are thought to most strongly affect growth rates (Boyce 1992). In species with age or stage structure (see Exercises 11 and 12), it is helpful

to know how each factor affects the birth, death, immigration, and emigration rates of the various groups of individuals.

Often, however, the only information available to the ecologist is growth rates of the population over time and their variation; the factors directly affecting growth rates may not be understood. We will use such a simple scenario that involves just population size and population growth rate and its variance to develop a PVA model. You may want to review Exercise 7, “Geometric and Exponential Population Growth,” to remind you of the derivation of the formula for  $R$ , the geometric rate of natural increase:

$$R = b - d \quad \text{Equation 1}$$

where  $b$  and  $d$  are the per capita birth and death rates, respectively. If  $R$  is positive, the population will grow. If  $R$  is negative, the population will decline. And if  $R$  is 0, the population will remain stable over time. If we start with a population of size  $N$  at time  $t$ , at time  $t + 1$  the population size will be

$$N_{t+1} = N_t + RN_t \quad \text{Equation 2}$$

We can factor  $N_t$  out of the terms on the right-hand side, to get

$$N_{t+1} = (1 + R)N_t$$

The quantity  $(1 + R)$  is often given its own symbol,  $\lambda$  (lambda), and so we can write

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

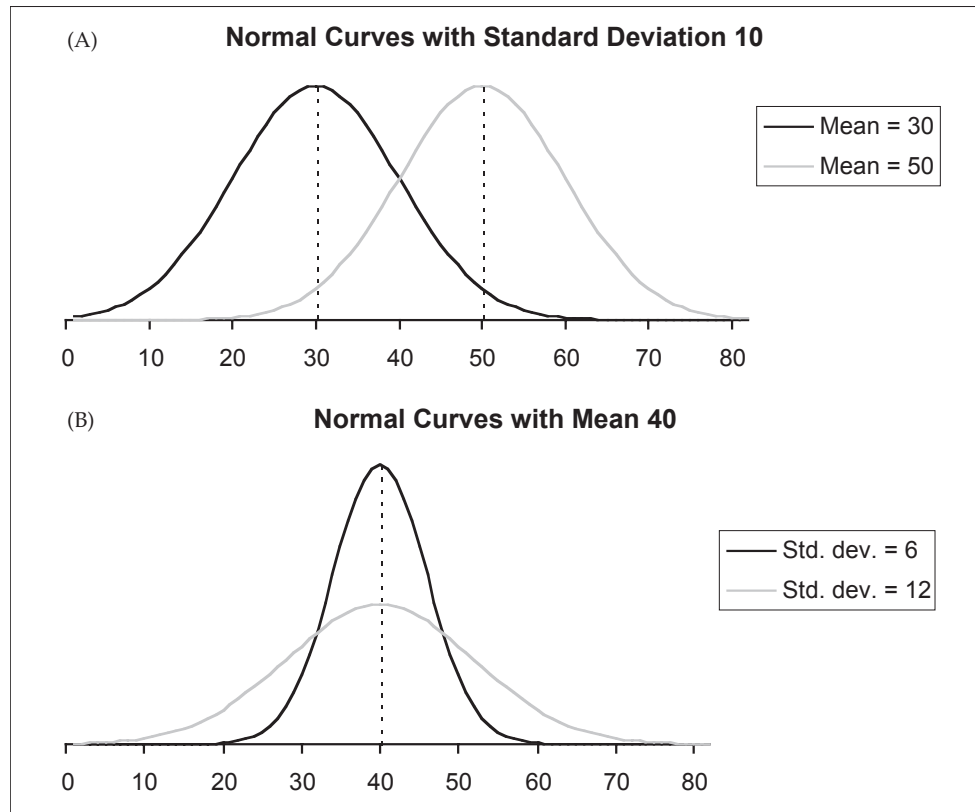
where  $\lambda$  is the **finite rate of increase** and describes the growth of a population in discrete (as opposed to continuous) time. The value of  $\lambda$  is always a positive number that measures the proportional change in population size from one year to the next (Gotelli 2001). It can be calculated as

$$\lambda = \frac{N_{t+1}}{N_t} \quad \text{Equation 3}$$

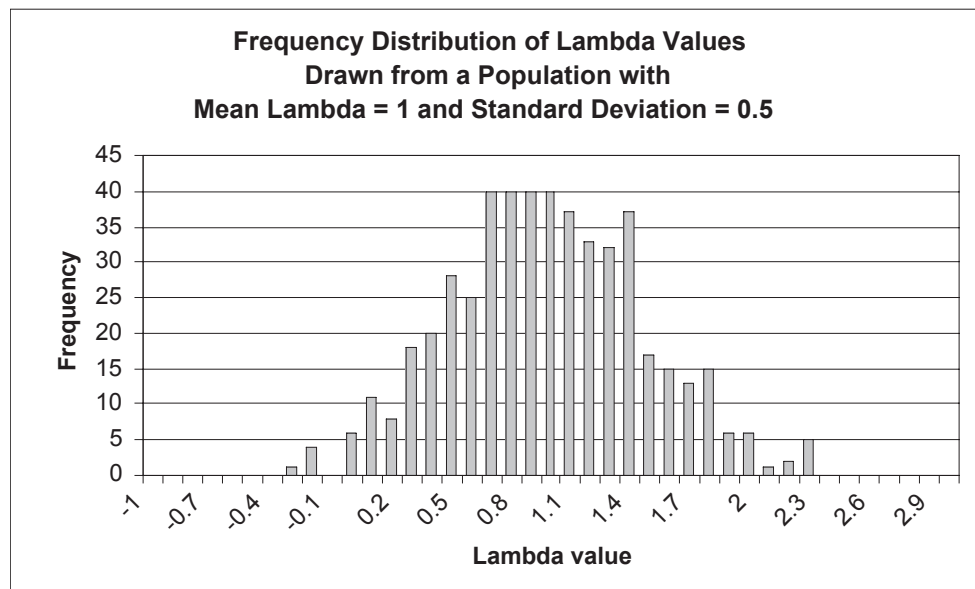
We will use Equation 3 to develop a spreadsheet population viability model.

Now that we have a handle on  $\lambda$ , we will let its value vary stochastically from time step to time step. We'll assume that the values  $\lambda$  can assume for a population are **normally distributed**. As you learned in Exercise 3, “Statistical Distributions,” a normal distribution means the raw data are distributed in a bell-shaped curve that is completely symmetrical and has tails that approach the  $x$ -axis. The shape and position of the normal curve is determined by the mean,  $\mu$ , and the standard deviation,  $\sigma$ :  $\mu$  sets the position of the curve, and  $\sigma$  determines the spread of the curve. Some forms that normal curves may take are shown in Figure 1. Note in Figure 1 that when  $\sigma$  is small, most of the data are distributed close to the mean, and when  $\sigma$  is large, the curve “flattens out” because the data vary more from the mean.

A property of normal curves is that the most of the data fall in the middle of the curve around the mean. For normal distributions, approximately 68% of the observations will fall between the mean and  $\pm 1$  standard deviation. If we assume, for example, that the average  $\lambda$  for our population is 1 and that  $\sigma$  is 0.5, this means that ~68% of the  $\lambda$  values will fall between 0.5 and 1.5 (which is the mean,  $1, \pm 0.5$ , which is 1 standard deviation). And ~95% of the observations will fall between the mean and  $\pm 2$  standard deviations. Since our  $\lambda$  values are normally distributed, approximately 95% of them will fall between 0 and 2 (or  $\pm 2$  standard deviations from the mean). Figure 2 shows  $\lambda$  values that were actually drawn from this distribution; the data were generated with the **NORMINV** spreadsheet function. Figure 2 shows the number of times each  $\lambda$  value, called a **deviate**, was drawn.



**Figure 1** (A) Two normal curves with different means ( $\mu$ ) but the same standard deviation ( $\sigma$ ); they are similar in shape but are positioned in different locations along the  $x$ -axis. (B) These two normal curves have the same mean, but their standard deviation from the mean is different. Note that the larger value of  $\sigma$  produces a “flatter” curve.



**Figure 2** You can see that most of the data are centered on  $\lambda = 1$ , but there is quite a lot of variation around the mean. You might also notice that some of the  $\lambda$  deviates are negative. Recall that  $\lambda$  is always positive, so another spreadsheet function must be used to adjust these negative deviates.

## PROCEDURES

The purpose of this exercise is to introduce you to some basics of risk analysis and population viability analysis. We will use spreadsheets to develop a simple, unstructured demographic model. The population will consist of 10 individuals, and will “grow” for 5 years. How much it grows will be determined by  $\lambda$ , which will vary stochastically from one time step to the next. At the end of 5 years, we will examine whether the population went extinct or not. We will repeat the population growth simulation 10 times so that we can compute the probability of extinction. Once the model is developed, you can easily alter the model’s parameters to examine how changes in a population’s parameters affect its probability of extinction. As always, save your work frequently to disk.

## INSTRUCTIONS

### A. Set up the model population.

1. Open a new spreadsheet and label column headings as shown in Figure 3. (Use a symbol font for the letter “ $\lambda$ ” to generate the  $\lambda$  symbol.)

2. Enter the value 10 in cell B4.

3. Enter an arbitrary value for average  $\lambda$  in cell C4.

4. Enter an arbitrary value for  $\sigma$ , the standard deviation, in cell D4.

## ANNOTATION

	A	B	C	D	E	F
1	<b>Population Viability Analysis</b>					
2						
3		<b><math>N(0)</math></b>	<b>Mean <math>\lambda</math></b>	<b>Std (<math>\lambda</math>)</b>	<b><math>\lambda</math> value</b>	<b><math>p</math> (extinction)</b>
4		10	1	0.5		

Figure 3

$N(0)$  will be the number of individuals in the starting population. For example, assume there are 10 finches on a small island in the Galapagos archipelago, and we are interested in developing a PVA for this population.

Enter a value of 1 in cell C4.

The growth rate  $\lambda$  of a population is simply  $N_{t+1}/N_t$ , or the number of individuals at time  $t + 1$  divided by the number of individuals at time  $t$ . You can estimate  $\lambda$  by simply counting the number of individuals in two time periods. If you do this over many time periods, you can estimate the average  $\lambda$  for the population. If the average  $\lambda = 1$ , in some years  $\lambda$  will be lower than 1, and other years  $\lambda$  will be higher than 1, but *on average*, populations will remain constant over time ( $\lambda = 1$ ).

Enter a value of 0.5 in cell D4.

The standard deviation captures the variability of  $\lambda$ . (Typing in the letter s and changing the font to the symbol font can generate the  $\sigma$  symbol.) If you measure populations over time and see little variation in  $\lambda$  (i.e., most of the values are around 1 over time), the standard deviation is small. If you see significant variation in  $\lambda$  over time, the standard deviation is large. Note that in both cases, the average  $\lambda$  remains 1.

Our population growth model will take the form  $N_{t+1} = N_t \times \lambda$ , where  $N_t$  is the number of individuals present at time  $t$ ,  $\lambda$  is the finite rate of increase, and  $N_{t+1}$  is the population at the next time step. Think about this formula for a bit and what it means. If  $\lambda$  were always 1, the population would not change in number over time. But the essence of environmental stochasticity is that population growth rates vary, and we need to incorporate this variation into our model. Because  $\lambda$  varies around a mean of 1 with a standard deviation of 0.5, we can now “sample” from this distribution to obtain a random  $\lambda$  value to be used to calculate  $N_{t+1}$ .

5. In cell E4, use the **NORMINV** function to obtain a random deviate of  $\lambda$ , drawn from a distribution whose mean is given in cell C4 and whose standard deviation is given in cell D4.

6. Save your work.

**B. Obtain stochastic  $\lambda$  values for the population model.**

1. Set up column headings as shown in Figure 4.

2. Generate a random deviate of  $\lambda$  in cell C7. Copy this formula into cells C7–F16.

3. Press F9, the calculate key, to draw new deviates from the distribution (Figure 5).

Enter the formula `=NORMINV(RAND(),$C$4,$D$4)` in cell E4.

The **NORMINV** function has the form **NORMINV(probability,mean,standard\_dev)**, where **probability** is a cumulative probability corresponding to the normal distribution, **mean** is the arithmetic mean of the distribution, and **standard\_dev** is the standard deviation of the distribution. This formula tells the spreadsheet to draw a random probability from a distribution with the mean specified in cell C4 and the standard deviation specified in cell D4. The program will convert this probability into an actual data point, that represents a deviate of  $\lambda$ . Note that when you press F9, the calculate key, you generate a new random number, and hence a new deviate of  $\lambda$ .

The last cell in your spreadsheet, cell F4, calculates the risk of extinction, given  $N_0$  and the mean and standard deviation of  $\lambda$ . This risk will be calculated later in the spreadsheet.

Now we will generate several deviates of  $\lambda$  to be used in our population model.

	B	C	D	E	F
6	<b>Trial</b>	<b><math>\lambda</math> value</b>			
7	1				
8	2				
9	3				
10	4				
11	5				
12	6				
13	7				
14	8				
15	9				
16	10				

Figure 4

Use the paste function key to guide your through your formula. Your final formula should have the form: `=NORMINV(RAND(),$C$4,$D$4)`.

	B	C	D	E	F
6	<b>Trial</b>	<b><math>\lambda</math> value</b>			
7	1	1.4157869	0.69563078	0.430788193	0.390610128
8	2	1.6358459	-0.1551856	1.232694219	0.838216809
9	3	1.3694845	1.33481228	1.346010438	0.806152573
10	4	1.3400447	0.8803662	0.430931439	1.259032049
11	5	0.6482688	1.74465106	0.851280563	0.735405254
12	6	0.539599	-0.0557869	1.285403985	0.889758442
13	7	1.4307401	0.49157655	1.218484502	1.430346745
14	8	0.0483275	0.31749244	0.110721092	0.073602339
15	9	0.1716595	0.86723992	0.997122018	0.503590971
16	10	0.9940314	1.78597168	1.368504516	1.633492618

Figure 5

4. Save your work.

**C. Model the population over 100 years.**

1. Set up column headings as shown in Figure 6.

	A	B	C	D	E	F	G
18		Year					Persist?
19	Trial	0	1	2	3	4	0 = Yes, 1 = No

Figure 6

Although only years 0–4 are shown, we will ultimately follow the fate of the population for 100 years. And you will run many different “trials” for your PVA. For now though, we will focus on developing the model through year 4 and will run 10 trials. After you have gone through the entire exercise, it will be fairly straightforward to extend your model to incorporate 100 years and 100 different trials.

2. Set up a linear series from 1 to 10 in cells A20–A29.

3. Use a formula to enter the initial population size in cell B20. Reference cell \$B\$4.

4. Copy this formula down to cell B29 to denote a starting population size of 10 for each of your 10 trials.

5. In cell C20, enter a formula to project population size for year 1, trial 1. Use one of the random deviates in cells C7–F16.

6. Copy cell C20 into cells D20–F20. This completes your projection for trial 1.

Enter the value 1 in cell A20.

In cell A21, enter `=1+A21`. Copy your formula down to cell A29.

Enter the formula `=B$4` in cell B20.

Note that this is an *absolute formula* (see Exercises 1 and 2). Later in the exercise you will change  $N_0$ . This formula establishes the initial population in your trials by directly linking  $N_0$  to cell B20.

Enter the formula `=ROUND(IF(B20*C7<1,0,B20*C7),0)` in cell C20.

To keep things simple, we will reference the  $\lambda$  deviates in order of appearance, so that the  $\lambda$  value for year 1, trial 1 is cell C7. The formula tells the spreadsheet to multiply B20, which is  $N_0$  for trial 1, by C7, which is a randomly obtained  $\lambda$ . This is the spreadsheet version of  $N_t \times \lambda$ , and generates the population size for time step 1. (Remember that our model is based on the equation  $N_{t+1} = N_t \times \lambda$ .) We phrased the formula as an **IF** to keep the population from going below zero. If the product is less than 1, the spreadsheet will return a 0, indicating the population has gone extinct. If the product is not less than 1, the spreadsheet will return the value  $B20 \times C7$ . Because there is no such thing as a decimal individual, we used the **ROUND** function to round the result of the **IF** formula to an ordinal number.

Your spreadsheet should now look something like Figure 7. Your results will be different, however, because your random  $\lambda$  values will be different than ours, resulting in different population sizes over time.

	A	B	C	D	E	F
18		Year				
19	Trial	0	1	2	3	4
20	1	10	14	10	4	2

Figure 7

7. Select cells C20–F20 and copy the formulae down to cells C29–F29.

**D. Create graphs.**

1. Graph your projections using the Line Graph option.

With the results of your first trial in place, you can now project the population sizes in other trials. Because your formulae in cells C20–F20 reference the cells to the left in terms of population numbers (a relative reference), and because the formula uses a relative reference for the random deviates, you can simply copy the formulae in C20–F20 down to obtain projections for trials 2–10.

Since the data are given in rows, select the Rows option in the Data Range tab (Figure 8). Make sure you fully label your axes. Select the Series tab and select cells B19–F19 as Category (x) axis labels. Your graph should resemble Figure 9.

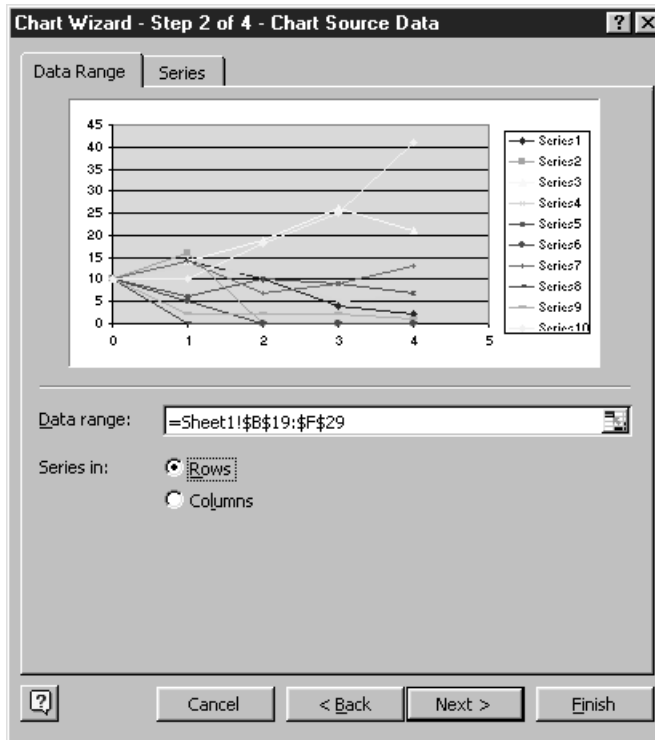


Figure 8

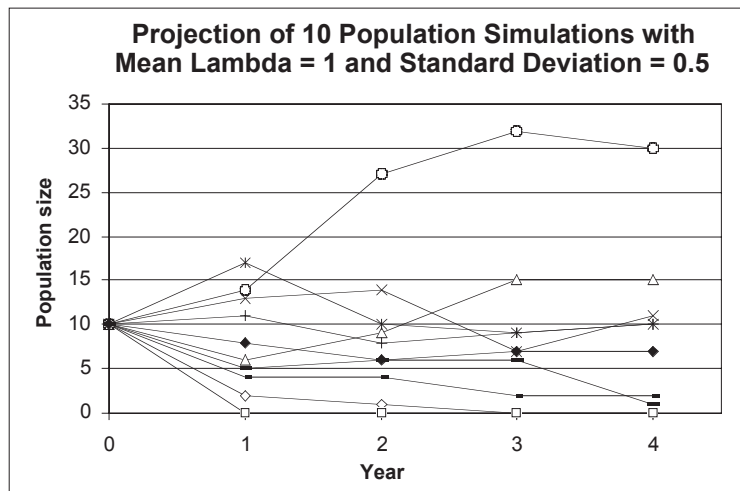


Figure 9



2. Press F9, the calculate key, to run new projections. Save your work.

**E. Calculate the risk of extinction.**

1. In cells G20–G29, enter an **IF** formula to determine which simulations (trials) went extinct or persisted.

2. In cell G30, enter a formula to obtain the proportion of trials in which the population persisted.

3. Enter a formula in cell F4 to calculate the extinction probability.

4. Extend your model to encompass 100 trials.

5. Answer the questions at the end of the exercise.

You have probably noticed that when you press F9, your results vary substantially. In some cases, the population persists over time for all 10 trials. For other cases, the population goes extinct for at least some of the 10 trials. How can you characterize this “risk of extinction?” The steps in the next section will help us quantify this risk.

To calculate the risk of extinction, you must first determine whether each of the simulations ended at zero (the population went extinct by year 4) or some number greater than 0 (the population persisted). Then, you need to determine the overall percentage of trials the ended with extinction or persistence.

Enter the formula `=IF(F20>0,1,0)` in cell G20. Copy the formula down to cell G29 to obtain results for all 10 of your trials.

If the last year of the trial had a positive number of individuals, you want to return the value “1” to indicate the population persisted. If it ended in 0, you want to return the value “0” to indicate that the population went extinct.

Enter the formula `=SUM(G20:G29)/10` in cell G30.

We’re interested in knowing in how many trials out of 10 the population persisted. We can get this by adding cells G20 to G29 and dividing by 10. This equals the probability that a trial did not result in extinction.

Enter the formula `1-G$30` in cell F4.

Extinction probability is the flip side of the value in cell G30. *The value returned by the formula in cell F4 is the key output of the model.*

Now that you have the basic idea, extend your simulation to encompass 100 years and run 100 different trials. You will want to create thousands of new deviates ( $100 \times 100$ ) to run your model. Make sure your formulae reflect what it is you want the spreadsheet to do. Be aware that extending the spreadsheet may slow down your computer’s computation speed a bit.

By completing this exercise, you will have made 100 simulations with randomly varying values of  $\lambda$  and determined the probability of extinction for a geometric model of population growth. This is a legitimate model using Monte Carlo simulation to estimate extinction risks. It also makes some useful predictions.

## QUESTIONS

1. Population viability analyses can yield information that allows an ecologist to make statements such as “A population of 10 *Geospiza fortis* has an  $x\%$  probability of surviving 100 years if current conditions persist.” Based on your PVA, what values can you enter for  $x$ ?
2. What value, if any, is there to running your model for 100 years, rather than just 4 years? What value, if any, is there to running 100 trials rather than just 10? Are there any “risks” to running your model for short versus long time periods?
3. We assume that models that include random variation in their parameters (that is, **stochastic models**) are more realistic than those that do not (that is, **deterministic models**) and hence are more useful in conservation applications. Whatever the case, stochastic and deterministic models should have the same underlying structure. Therefore, a stochastic model should generate the same predictions when its variability parameters are set to zero as does a deterministic model that has no variability parameters. Test this property of your model



by setting  $\sigma$  to zero. (You might have to set your value to a very small value, such as 0.001, instead of 0 to run your model.) What do you see?

4. Contrast the fate of this population in terms of extinction probability when the growth rate is held constant at 1 and the standard deviation in growth rate is increased from 0 to 1, by 0.1 increments. Set up column headings as shown for your 100 trials/100 year simulation, and record the probability of extinction. Why does the extinction probability vary when the growth rate is constant? If  $\lambda$  is held at 1, why doesn't the population simply stay stable?

	I	J
16	100 years, 100 trials	
17	$\sigma$	$p$ (extinction)
18	0	
19	0.1	
20	0.2	
21	0.3	
22	0.4	
23	0.5	
24	0.6	
25	0.7	
26	0.8	
27	0.9	
28	1	

5. For a model with identical growth parameters (say,  $\lambda = 1$ ,  $\sigma = 0.25$ ), vary population size from 100 to 1000 by some standard interval (like 100s). How does the extinction probability change as population size changes? Make a graph of population size ( $x$ -axis) versus extinction probability ( $y$ -axis) to show this relationship. Why is the relationship not constant? In other words, if all the model parameters don't change on average, then why should initial population size have any influence on extinction probability? Is there a critical minimum initial size (a threshold) above which extinction probability changes more slowly? Might this serve as a minimum viable population size for setting management objectives? If so, what are the assumptions of the model and, in particular, the characteristics of individuals you have modeled?

### LITERATURE CITED

- Boyce, M. S. 1992. Population viability analysis. *Annual Review of Ecology and Systematics* 23: 481–506.
- Gotelli, N. 2001. *A Primer of Ecology*, 3rd Ed. Sinauer Associates, Sunderland, MA.
- Grant, P. 1986. *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton, NJ.

