

PART THREE

SAMPLING AND EXPERIMENTAL DESIGN

Ecologists are more and more confronted with the need to do their work in the most efficient manner. To achieve these desirable practical goals an ecologist must learn something about sampling theory and experimental design. These two subjects are well covered in many statistical books such as Cochran (1977), Cox (1958), and Mead (1988), but the terminology is unfortunately foreign to ecologists and some translation is needed. In the next four chapters I discuss sampling and experimental design from an ecological viewpoint. I emphasize methods and designs that seem particularly needed in ecological research. All of this should be viewed as a preliminary discussion that will direct students to more comprehensive texts in sampling and experimental design.

Sampling and experimental design are statistical jargon for the three most obvious questions that can occur to a field ecologist: *where* should I take my samples, *how* should I collect the data in space and time, and *how many* samples should I try to take? Over the past 90 years statisticians have provided a great deal of practical advice on how to answer these questions and - for those that plan ahead - how *not* to answer these questions. Let us begin with the simplest question of the three: how many samples should I take?

CHAPTER 7

SAMPLE SIZE DETERMINATION AND STATISTICAL POWER

(Version 4, 14 March 2013)

Page

7.1 SAMPLE SIZE FOR CONTINUOUS VARIABLES	277
7.1.1 Means From A Normal Distribution	278
7.1.2 Comparison of Two Means	286
7.1.3 Variances From A Normal Distribution	292
7.2 DISCRETE VARIABLES	293
7.2.1 Proportions and Percentages	294
7.2.2 Counts From A Poisson Distribution	297
7.2.3 Counts From A Negative Binomial Distribution	298
7.2.4 Counts From A Lognormal Distribution	300
7.3 SPECIALIZED ECOLOGICAL VARIABLES	303
7.3.1 Mark-and-Recapture Estimates	303
7.3.2 Line Transect Estimates	306
7.3.3 Distance Methods	309
7.3.4 Change-In-Ratio Methods	310
7.4 STATISTICAL POWER ANALYSIS	310
7.4.1 Estimation of Effect Size for Continuous Variables	311
7.4.2 Effect Size for Categorical Variables	313
7.4.3 Power Analysis Calculations	314
7.5 WHAT TO DO IF NOTHING WORKS	316
7.6 SUMMARY	320
SELECTED REFERENCES	321

A recurrent practical problem in ecological research projects is *how much work do I need to do?* The statistician translates this question into the more easily answered question - *how large a sample should I take?* In spite of the close

appearance of these two questions, there is a great gap between them. The detailed objectives of the research program need to be specified for a complete answer to the first question, and a statistician is not very helpful at this stage. For example, you may wish to estimate the density of ring-necked pheasants on your study area. Given this overall objective, you must specify much more ecological detail before you see your statistician. You must decide, for example, how often in each year is estimation required, the type of sampling gear you will use, and whether your study area contains a closed population or is subject to extensive immigration and emigration. All of this information and more must be used to help decide on the method of population estimation you should apply. Do not expect a statistician to help you with these difficult decisions - go back and read the previous three chapters! But then, when you have decided, for example, to use the Petersen method of mark-recapture estimation, it is useful to ask a statistician *how large a sample should I take?*

Throughout the last three chapters we have periodically come across this question and tried to provide a partial answer. This chapter will draw together a series of general methods that one can apply to any situation, and will guide you toward more complex methods when necessary. There is an excellent review of the sample size problem in population ecology by Eberhardt (1978) and general surveys by Kraemer and Thiemann (1987), Mace (1964) and Cochran (1977).

We will start here with the simplest question of statistical *estimation*, and determine how to determine the sample size necessary to obtain a specified level of precision. For example, we might ask how many fish we need to weigh to obtain a mean weight with 95% confidence limits of ± 2.5 grams. From this simple question we will move later in the chapter to discuss the more complex questions of the sample sizes required for statistical *inference* in which, for example, we wish to compare the size of fish in two or more different lakes. To test hypotheses about statistical populations we must clearly understand the concept of statistical power.

7.1 SAMPLE SIZE FOR CONTINUOUS VARIABLES

We will begin by considering variables measured on the interval or ratio scale which can show a normal distribution. How can we decide how big a sample to take?

7.1.1 Means From A Normal Distribution

If you need to measure a sample to estimate the average length of whitefish in a particular lake, you can begin with a very simple statistical assumption that lengths will have a normal, bell-shaped frequency distribution. The two-sided confidence interval for a normal variable is:

$$\text{Probability } [(\bar{x} - t_{\alpha} s_{\bar{x}}) < \mu < (\bar{x} + t_{\alpha} s_{\bar{x}})] = 1 - \alpha$$

where:

$$t_{\alpha} = \text{Student's } t\text{-value for } n - 1 \text{ degrees of freedom for } \alpha/2$$

$$s_{\bar{x}} = \text{Standard error of mean} = \sqrt{\frac{\text{Variance of mean}}{\text{Sample size}}}$$

$$n = \text{Sample size}$$

This confidence interval formulation is usually written in statistical shorthand as:

$$\bar{x} \pm t_{\alpha} s_{\bar{x}}$$

The width of the confidence interval depends on the t -value and the standard error. To proceed further some information is required, and since the steps involved are general, I will describe them in some detail.

Three general steps must be taken to estimate sample size:

Step 1. *Decide what level of precision you require.* Do you want your estimate of the mean to be accurate within $\pm 10\%$ of the true population mean? Within $\pm 1\%$ of the true mean? You must decide what the desired error limits will be for your sample. This is not a statistical question but an ecological one, and it will depend on what theory you are testing or what comparisons you will make with your data.

The desired level of precision can be stated in two different ways and these can be most confusing if they are mixed up. You can define the *absolute* level of precision desired. For example, you may wish the 95% confidence interval to be ± 2.8 mm or less for your whitefish. Alternatively, you can define the *relative* level of precision desired. For example, you may wish the 95% confidence interval to be $\pm 6\%$ of the mean or less. These two are related simply:

$$\text{Percent relative error desired} = \left(\frac{\text{Absolute error desired}}{\bar{x}} \right) 100$$

I will use the absolute level of precision to begin but for some purposes as we proceed we will find it easier to use relative error as a measure of precision.

A second potential source of confusion involves how to specify the level of precision you require. Some authors define the level of precision in terms of ± 1 *standard error* of the mean. Others define the level of precision in terms of the *width of the confidence interval*, which for 95% confidence gives approximately ± 2 standard errors of the mean. In this chapter I will always define the desired level of precision in terms of the width* of the confidence interval. If you use other statistical texts, be sure to check which definition of "desired precision" they adopt.

Step 2. *Find some equation that connects sample size (n) with the desired precision of the mean.* This equation will depend on the kind of sampling being done and the type of variable being measured. The purpose of this chapter is to give you these equations. For the case of the mean from a normal distribution we have the equation:

$$\text{Desired absolute error} = d = t_{\alpha} s_{\bar{x}}$$

or, expanding:

$$d = \frac{t_{\alpha} s}{\sqrt{n}}$$

rearranging:

$$n = \left(\frac{t_{\alpha} s}{d} \right)^2 \tag{7.1}$$

where:

* Technically speaking, I should say the "half-width" of the confidence interval; but this cumbersome wording is omitted. I use the term to indicate that in general we think of confidence intervals as a parameter \pm width of the confidence interval.

n = Sample size needed to estimate the mean
 t_{α} = Student's t value for $n - 1$ degrees of freedom
 for the $1 - \alpha$ level of confidence
 s = Standard deviation of variable
 d = Desired absolute error

This is the formula to estimate sample size for a mean from a normal distribution.

Step (3) *Estimate or guess the unknown parameters of the population that you need to solve the equation.* In this simple case we need to have an estimate of s , the standard deviation of the variable measured. How can we get an estimate of the standard deviation? There are four ways you may get this for the variable you are studying:

- (a) *By previous sampling of a similar population:* You may know, for example, from work done last year on the same whitefish in another lake that the standard deviation was about 6.5 mm.
- (b) *By the results of a pilot study.* You may wish to spend a day or two sampling to get an estimate of the standard deviation in your population.
- (c) *By guesswork:* Often an experienced person may have implicit knowledge of the amount of variability in a variable. From this type of information on the *range* of measurements to be expected (maximum value - minimum value), you can estimate the standard deviation of a measure that has a normal distribution by the formulae given in Table 7.1. For example, if you know that in a sample of about 50 fish, length will vary from 15-70 mm, from Table 7.1:

$$\begin{aligned} \text{Estimated standard deviation} &\approx (0.222)(\text{range}) \\ &\approx (0.222)(55) = 12.2 \text{ mm} \end{aligned}$$

- (d) *By two-stage sampling:* If it is feasible for you to sample in two steps, you can first take a sample of size n_1 and calculate from these measurements a preliminary estimate of the standard deviation. Cochran (1977 p. 79) shows that in the second sample you need to take additional measurements to make a total sample size of:

TABLE 7.1 ESTIMATION OF THE STANDARD DEVIATION (s) OF A VARIABLE FROM KNOWLEDGE OF THE RANGE (w) FOR SAMPLES OF VARIOUS SIZES^a

Sample size	Conversion factor	Sample size	Conversion factor
2	0.886	19	0.271
3	0.591	20	0.268
4	0.486	25	0.254
5	0.430	30	0.245
6	0.395	40	0.231
7	0.370	50	0.222
8	0.351	60	0.216
9	0.337	70	0.210
10	0.325	80	0.206
11	0.315	90	0.202
12	0.307	100	0.199
13	0.300	150	0.189
14	0.294	200	0.182
15	0.288	300	0.174
16	0.283	500	0.165
17	0.279	1000	0.154
18	0.275		

^a Range = maximum value - minimum value. Multiply the observed range by the tabled values to obtain an unbiased estimate of the standard deviation. A normal distribution is assumed. *Source:* Dixon and Massey, 1983

$$n = \frac{(t_{\alpha} s_1)^2}{d^2} \left(1 + \frac{2}{n_1} \right) \quad (7.2)$$

where:

n = Final total sample size needed

n_1 = Size of first sample taken

s_1 = Standard deviation estimated from first sample taken

t_{α} = Student's t with $n - 1$ d.f for $1 - \alpha$ level of confidence

d = Desired absolute error

Since you are using data from the first sample to estimate total sample size, there is an element of statistical circularity here, but the important point is to remember that these procedures are approximations to help you to decide the appropriate sample size needed.

An example will illustrate the application of these three steps with equation (7.1). Assume you wish to measure whitefish to obtain a 95% confidence interval of ± 2.8 mm. From previous sampling it is known that the standard deviation of lengths is about 9.4 mm. Thus, from equation (7.1):

$$n = \left(\frac{t_{\alpha} s}{d} \right)^2 = \left[\frac{t_{\alpha} (9.4)}{2.8} \right]^2$$

We immediately have another problem because to look up a t_{α} -value we need to know the sample size n , so we are trapped in a statistical Catch-22. For those who desire precision, equation (7.1) can be solved iteratively by trial and error. In practice it is never worth the effort because t-values for 95% confidence limits are almost always around 2 (unless n is very small), and we will use in equation (7.1) the approximation $t_{\alpha} = 2$ when we need 95%* confidence limits.

Thus:

$$n \approx \left[\frac{2(9.4)}{2.8} \right]^2 \approx 45.1$$

The recommendation is to measure about 45 fish.

This technique for estimating sample size can also be used with relative measures of precision (Eberhardt 1978). The most convenient measure of relative variability is the *coefficient of variation*:

$$CV = \frac{s}{\bar{x}} \tag{7.3}$$

* If you need to work in 90% confidence intervals, you can use $t_{\alpha} \sim 1.7$. If you use 99% confidence intervals, use $t_{\alpha} \sim 2.7$. These are approximations.

where:

$$s = \text{standard deviation}$$

$$\bar{x} = \text{Observed mean}$$

If you know the coefficient of variation for the variable you are measuring, you can estimate the sample size needed in the following way.

From the formulas given above:

$$\text{Desired relative error} = r = \left(\frac{t_{\alpha} s_{\bar{x}}}{\bar{x}} \right) 100$$

or, expanding:

$$r = \left(\frac{t_{\alpha} s}{\bar{x} \sqrt{n}} \right) 100$$

rearranging:

$$n = \left(\frac{s}{\bar{x}} \right)^2 \left(\frac{t_{\alpha}}{r} \right)^2 100^2$$

or

$$n = \left(\frac{100CV t_{\alpha}}{r} \right)^2 \quad (7.4)$$

We can simplify this equation by assuming for 95% confidence limits that $t_{\alpha} = 2$ so this reduces to:

$$n \cong \left(\frac{200CV}{r} \right)^2 \quad (7.5)$$

where:

$$r = \text{desired relative error (width of confidence interval as percentage)}$$

$$CV = \text{coefficient of variation (equation 7.3)}$$

As an example, suppose you are counting plankton samples and know that the coefficient of variation is about 0.70 for such data. If you wish to have $\pm 25\%$ in the relative precision of your mean estimate, then:

$$n \approx \left[\frac{(200)(0.70)}{25} \right]^2 = 31.4$$

and you require a sample size of about 31 samples to get a confidence interval of approximately $\pm 25\%$ of the mean.

Eberhardt (1978) has drawn together estimates of the coefficient of variation from published ecological data (Table 7.2). Clearly it is an oversimplification to assume that all different sampling regions will have a constant coefficient of variation for any ecological measurement. But the point is that as an *approximate* guide and in the absence of more detailed background data it will give you a general idea of the variation you might expect with different kinds of sampling. Table 7.2 and equation (7.5) can supply a rough guide to the sampling intensity you should provide.

TABLE 7.2 COEFFICIENTS OF VARIATION OBSERVED IN A VARIETY OF POPULATION SAMPLING TECHNIQUES TO ESTIMATE POPULATION SIZE^a

Group of organisms	Coefficient of variation
Aquatic organisms	
<i>Plankton</i>	0.70
<i>Benthic organisms</i>	
Surber sampler, counts	0.60
Surber sampler, biomass or volume	0.80
Grab samples or cores	0.40
<i>Shellfish</i>	0.40
<i>Fish</i>	0.50 to 2.00
Terrestrial organisms	
<i>Roadside counts</i>	0.80
<i>Call counts</i>	0.70
<i>Transects (on foot)</i>	0.50 to 2.00
<i>Fecal pellet counts</i>	1.00

^a Average values compiled by Eberhardt (1978a).

All of the above formulas assume that the population is very large relative to the number sampled. If you are measuring more than 5-10% of the whole population, you

do not need to take so large a sample. The procedure is to calculate the sample size from equation (7.1) or (7.5) above and then correct it with the following *finite population correction*:

$$n^* \cong \frac{n}{1 + (n/N)} \quad (7.6)$$

where:

- n^* = Estimated sample size required for finite population of size N
- n = Estimated sample size from (7.1) and (7.5) for infinite population
- N = Total size of finite population

For example, if we know that the whitefish population described above (page 282) totals only 250 individuals, to get a 95% confidence interval of ± 2.8 mm, we first estimate $n = 45.1$ as above, and then:

$$n^* \approx \frac{n}{1 + n/N} = \frac{45.1}{1 + (45.1/250)} = 38.2 \text{ fish}$$

Note that the required sample size is always *less* when we use the finite population correction, so that less effort is required to sample a finite population.

These equations to estimate sample size are derived from the normal distribution. What happens if the variable we are measuring does not have a normal, bell-shaped distribution? Fortunately, it does not matter much because of the Central Limit Theorem. The Central Limit Theorem states that, *as sample size increases, the means of samples drawn from a population with any shape of distribution will approach the normal distribution**. In practice this theorem means that with large sample sizes ($n > 30$) we do not have to worry about the assumption of a normal distribution. This theorem is among the most important practical findings of theoretical statistics because it means that you do not need to worry if you are measuring a variable that has a skewed distribution - you can still use the approach

* Assuming the variance is finite for the distribution.

outlined above (Sokal and Rohlf 1995). But if the distribution is *strongly* skewed, you should be cautious and apply these formulas conservatively.

7.1.2 Comparison of Two Means

Planning experiments which involve comparisons of two or more means is more complicated than deciding on a single sample size. You may, for example, wish to compare whitefish lengths in two different lakes. After the data are collected, you will evaluate how different the lakes are in this parameter. How can you decide beforehand how big a sample to use?

To do this you must first decide on the smallest difference you wish to be able to detect:

$$d = |\mu_A - \mu_B| \quad (7.7)$$

where:

$$\begin{aligned} d &= \text{Smallest difference you wish to detect} \\ \mu_A &= \text{Mean value for population A} \\ \mu_B &= \text{Mean value for population B} \end{aligned}$$

For example, you may decide you wish to pick up a difference of 8 mm or more in fish lengths between lakes A and B. This difference d must be expressed in units of the standard deviation of the variable being measured:

$$D = \frac{d}{s} \quad (7.8)$$

where:

$$\begin{aligned} D &= \text{Standardized smallest difference you wish to detect} \\ d &= \text{Smallest difference you wish to detect} \\ s &= \text{Standard deviation of variable measured} \end{aligned}$$

The standard deviation is assumed to be the same for both populations A and B, and is estimated from previous knowledge or guesswork. Note that the size of this difference is dictated by your management objectives or the ecological hypothesis you are testing.

The second decision you must make is the probability you will tolerate for making a Type I or a Type II error. These are usually called α and β :

α = Probability of rejecting the null hypothesis of no difference when in fact it is true (Type I error)

β = Probability of accepting the null hypothesis when in fact it is false and the means really do differ (Type II error)

By convention α is often set to 0.05, but of course it can be set to any other value depending on the consequences of making a Type I mistake. The probability β is less well known to ecologists and yet is critical. It is related to the "power" of the statistical test:

$$\text{Power} = 1 - \beta \quad (7.9)$$

When β is very small the test is said to be very powerful, which means that you will not make Type II errors very often. Ecologists are less familiar with β because it is fixed in any typical study after you have specified α , measured a particular sample size, and specified the alternative hypothesis. But if you have some choice in deciding on what sample size to take, you can specify β independently of α . Sokal and Rohlf (2012, pg. 167) and Zar (2010, pg. 81) have a good discussion of the problem of specifying α and β .

Given that you have made these decisions, you can now estimate the sample size needed either by solving the equations given in Mace (1964 page 39) or by the use of Table 7.3 from Davies (1956). In most cases the table is more convenient. For example, suppose for your comparison of fish lengths you have:

$$\begin{aligned} d &= |\mu_A - \mu_B| = 8 \text{ mm} \\ s &= 9.4 \text{ mm} \quad (\text{from previous studies}) \end{aligned}$$

and you decide you wish $\alpha = 0.01$ and $\beta = 0.05$ and are conducting a two-tailed test. Calculate the standardized difference D from equation (7.8):

$$D = \frac{d}{s} = \frac{8}{9.4} = 0.85$$

Table 7.3 NUMBER OF SAMPLES REQUIRED FOR A t-TEST OF DIFFERENCE BETWEEN THE MEANS OF TWO POPULATIONS^a

Two-sided test		Level of t-test																							
		0.01					0.02					0.05					0.1								
		$\alpha = 0.01$					$\alpha = 0.02$					$\alpha = 0.05$					$\alpha = 0.1$								
$\beta =$		0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5				
V A L U E	0.05																					0.05			
	0.10																					0.10			
	0.15																					0.15			
	0.20																			137		0.20			
	0.25														124						88	0.25			
	0.30									123					87						61	0.30			
	0.35					110				90					64					102	45	0.35			
	0.40					85				70					100	50				108	78	45	0.40		
	0.45				118	68				101	55				105	79	39			108	86	62	28	0.45	
	0.50				96	55				106	82	45				106	86	64	32			88	70	51	23
O F D	0.55			101	79	46			106	88	68	38			87	71	53	27	112	73	58	42	19	0.55	
	0.60			101	85	67	39			90	74	58	32	104	74	60	45	23	89	61	49	36	16	0.60	
	0.65			87	73	57	34	104	77	64	49	27	88	63	51	39	20	76	52	42	30	14	0.65		
	0.70	100	75	63	50	29	90	66	55	43	24	76	55	44	34	17	66	45	36	26	12	0.70			
	0.75	88	66	55	44	26	79	58	48	38	21	67	48	39	29	15	57	40	32	23	11	0.75			
	0.80	77	58	49	39	23	70	51	43	33	19	59	42	34	26	14	50	35	28	21	10	0.80			
	0.85	69	51	43	35	21	62	46	38	30	17	52	37	31	23	12	45	31	25	18	9	0.85			
	0.90	62	46	39	31	19	55	41	34	27	15	47	34	27	21	11	40	28	22	16	8	0.90			
	0.95	55	42	35	28	17	50	37	31	24	14	42	30	25	19	10	36	25	20	15	7	0.95			

1.00	50	38	32	26	15	45	33	28	22	13	38	27	23	17	9	33	23	18	14	7	1.00
1.1	42 ^a	32	27	22	13	38	28	23	19	11	32	23	19	14	8	27	19	15	12	6	1.1
1.2	36	27	23	18	11	32	24	20	16	9	27	20	16	12	7	23	16	13	10	5	1.2
1.3	31	23	20	16	10	28	21	17	14	8	23	17	14	11	6	20	14	11	9	5	1.3
1.4	27	20	17	14	9	24	18	15	12	8	20	15	12	10	6	17	12	10	8	4	1.4
1.5	24	18	15	13	8	21	16	14	11	7	18	13	11	9	5	15	11	9	7	4	1.5
1.6	21	16	14	11	7	19	14	12	10	6	16	12	10	8	5	14	10	8	6	4	1.6
1.7	19	15	13	10	7	17	13	11	9	6	14	11	9	7	4	12	9	7	6	3	1.7
1.8	17	13	11	10	6	15	12	10	8	5	13	10	8	6	4	11	8	7	5		1.8
1.9	16	12	11	9	6	14	11	9	8	5	12	9	7	6	4	10	7	6	5		1.9
2.0	14	11	10	8	6	13	10	9	7	5	11	8	7	6	4	9	7	6	4		2.0
2.1	13	10	9	8	5	12	9	8	7	5	10	8	6	5	3	8	6	5	4		2.1
2.2	12	10	8	7	5	11	9	7	6	4	9	7	6	5		8	6	5	4		2.2
2.3	11	9	8	7	5	10	8	7	6	4	9	7	6	5		7	5	5	4		2.3
2.4	11	9	8	6	5	10	8	7	6	4	8	6	5	4		7	5	4	4		2.4
2.5	10	8	7	6	4	9	7	6	5	4	8	6	5	4		6	5	4	3		2.5
3.0	8	6	6	5	4	7	6	5	4	3	6	5	4	4		5	4	3			3.0
3.5	6	5	5	4	3	6	5	4	4		5	4	4	3		4	3				3.5
4.0	6	5	4	4		5	4	4	3		4	4	3			4					4.0

^a The entries in the table show the number of observations needed to test for differences of specified size (D) between two means at fixed levels of α and β . Entries in the table show the sample size needed in each of two samples of equal size. Source: Davies (1956), Table E.1.

From Table 7.3 in the second column, we read

$$n = 51$$

This means that you should measure 51 fish from lake A *and* 51 fish from lake B to achieve your stated goal.

Note from Table 7.3 that the smaller the α and the smaller the β , and the smaller the standardized distance D that you wish to detect, the *larger* the sample size required. This table thus restates the general principle in statistics that if you wish to gain in precision you must increase sample size.

The approach used in Table 7.3 can be generalized for the comparison of 3, 4, 5 ... means from different populations, and consequently this table can be useful for planning relatively complex experiments. Kastenbaum *et al.* (1970) give a more detailed discussion of sample sizes required when several means are to be compared.

If Table 7.3 does not cover your particular needs, you can use the approximate equation suggested by Snedecor and Cochran (1967, pg. 113) for a comparison of two means:

$$n \cong \frac{2(z_{\alpha} + z_{\beta})^2 s^2}{d^2} \quad (7.10)$$

where:

n = Sample size required from *each* of the two populations

z_{α} = Standard normal deviate for α level of probability ($z_{.05}$ is 1.96;
 $z_{.01}$ is 2.576)

z_{β} = Standard normal deviate for the probability of a type II error
(see table below)

s^2 = Variance of measurements (known or guessed)

$d = |\mu_A - \mu_B|$ = Smallest difference between means you wish to
detect with probability $1 - \beta$

Equation (7.10) is only approximate because we are using z -values in place of the more proper t -values, but this is a minor error when $n > 50$. The z_{β} values are

obtained from tables of the standard normal deviate (z). A few examples are (Eberhardt 1978):

Type II error (β)	Power ($1 - \beta$)	Two-tailed z_β
0.40	0.60	0.25
0.20	0.80	0.84
0.10	0.90	1.28
0.05	0.95	1.64
0.01	0.99	2.33
0.001	0.999	2.58

For example, if you wished to detect a difference in fish lengths in the above example of 4 mm, then:

$$d = |\mu_A - \mu_B| = 4 \text{ mm}$$

$$s^2 = (9.4)^2 \text{ mm} \quad (\text{from previous studies})$$

and $\alpha = 0.01$ and $\beta = 0.05$. Thus from equation (7.10):

$$n \cong \frac{2(2.576 + 1.64)^2 (9.4)^2}{4^2} = 196.3 \text{ fish}$$

Note that Table 7.3 does not cover this particular combination of d and s^2 .

An alternative approach is to use the tables in Kastenbaum *et al.* (1970) to estimate the required sample size. Figure 7.1 is derived from these tables. In this case the means are expressed as a ratio you wish to detect:

$$k = \frac{\mu_A}{\mu_B} \tag{7.11}$$

where:

μ_A = Larger mean

μ_B = Smaller mean

k = Ratio you wish to detect between the means

If you have a rough idea of the coefficient of variation in your data (cf. Table 7.2), Figure 7.1 can be used in planning a study involving the comparison of two or more

means. Note that the sample size given in Figure 7.1 is required for *each* of the two populations being compared.

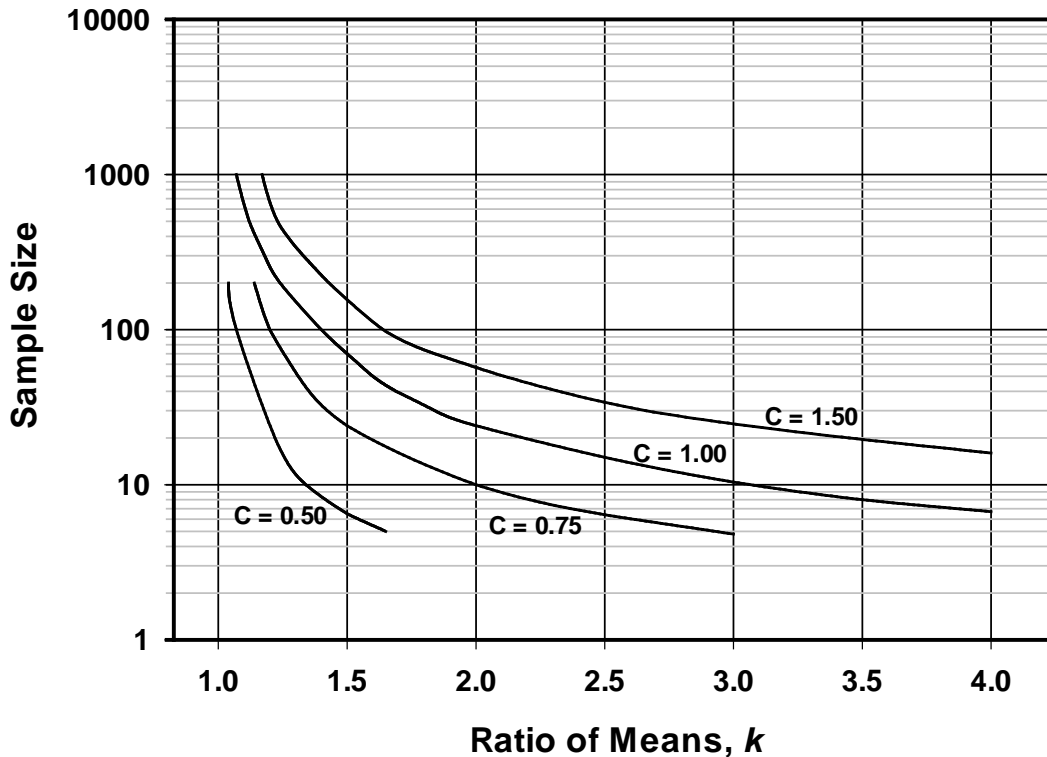


Figure 7.1 Sample sizes required to detect a significant difference between two means expressed as a ratio (k) with $\alpha = 0.05$ and $\beta = 0.20$, plotted for different coefficients of variation (C) that are typical of many population techniques. (From Eberhardt, 1978b.)

7.1.3 Variances From A Normal Distribution

In some cases an ecologist may wish to estimate the variance of a given measurement to a specified level of precision. How can you estimate how large a sample you should take? First, you must specify the allowable limits for the confidence limits on the variance. This is most easily expressed as a percentage ($\pm 25\%$) or as a proportion (± 0.25). Second, you must choose the probability (α) of the confidence interval not including the true variance. Then, if sample sizes are not too small ($n > 30$), from Mace (1964 p. 57):

$$n \cong \frac{3}{2} + z_{\alpha}^2 \left[\frac{1}{v} \left(\frac{1}{v} + \sqrt{\frac{1}{v^2} - 1} \right) - \frac{1}{2} \right] \quad (7.12)$$

where:

n = Approximate sample size required to estimate the variance

v = Allowable limits of error (expressed as a proportion) of the variance

In equation (7.12) substitute $z_{\alpha} = 1.96$ for $\alpha = 0.05$; if you wish to have $\alpha = 0.01$ use $z_{\alpha} = 2.58$. Other constants may be chosen for other α values from tables of the standard normal distribution.

For example, suppose that you wish to estimate the variance of a measurement with allowable limits for the confidence limits of $\pm 35\%$ and you wish to use $\alpha = 0.05$.

Then from equation (7.12):

$$\begin{aligned} n &\cong \frac{3}{2} + 1.96^2 \left[\frac{1}{0.35} \left(\frac{1}{0.35} + \sqrt{\frac{1}{0.35^2} - 1} \right) - \frac{1}{2} \right] \\ &= 60.3 \end{aligned}$$

so you would need to take about 60 samples.

Unfortunately, this procedure is applicable only to variances from a *normal* distribution, and is quite sensitive to departures from normality.

7.2 DISCRETE VARIABLES

Counts of the numbers of plants in a quadrat or the numbers of eggs in a nest differ from continuous variables in their statistical properties. The frequency distribution of counts will often be described by either the binomial distribution, the Poisson distribution, or the negative binomial distribution (Elliott 1977). The sampling properties of these distributions differ, and thus we require a different approach to estimating sample sizes needed for counts.

7.2.1 Proportions and Percentages

Proportions like the sex ratio or the fraction of juveniles in a population are described statistically by the binomial distribution. All the organisms are classified into 2 classes, and the distribution has only two parameters:

$$\begin{aligned} p &= \text{Proportion of } x \text{ types in the population} \\ q &= 1 - p = \text{Proportion of } y \text{ types in the population} \end{aligned}$$

We must specify a margin of error (d) that is acceptable in our estimate of p , and the probability (α) of not achieving this margin of error. If sample size is above 20, we can use the normal approximation to the confidence interval:

$$\hat{p} \pm t_{\alpha} s_{\hat{p}} \quad (7.13)$$

where:

$$\begin{aligned} \hat{p} &= \text{Observed proportion} \\ t_{\alpha} &= \text{Value of Student's } t\text{-distribution for } n-1 \text{ degrees of freedom} \\ s_{\hat{p}} &= \text{Standard error of } \hat{p} = \sqrt{\hat{p}\hat{q}/n} \end{aligned}$$

Thus the desired margin of error is:

$$d = t_{\alpha} s_{\hat{p}} = t_{\alpha} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Solving for n , the sample size required is:

$$n = \frac{t_{\alpha}^2 \hat{p}\hat{q}}{d^2} \quad (7.14)$$

where:

$$\begin{aligned} n &= \text{Sample size needed for estimating the proportion } p \\ d &= \text{Desired margin of error in our estimate} \end{aligned}$$

and the other parameters are defined above.

As a first approximation for $\alpha = .05$ we can use $t_{\alpha} = 2.0$. We need to have an approximate value of p to use in this equation. Prior information, or a guess, should be used, with the only rule-of-thumb that when in doubt pick a value of p closer to 0.5 than you guess - this will make your answer conservative.

As an example, suppose you wish to estimate the sex ratio of a deer population. You expect p to be about 0.40 and you would like to estimate p within an error limit of ± 0.02 with $\alpha = 0.05$. From equation (7.14):

$$n \cong \frac{(2.0^2)(0.40)(1-0.40)}{(0.02)^2} = 2400 \text{ deer}$$

Given this estimate of n , we can recycle to equation (7.14) with a more precise value of $t_\alpha = 1.96$, to get a better estimate of $n = 2305$. So you must classify 2305 deer to achieve this level of precision. If you wish to have $\alpha = 0.01$ ($t_\alpha = 2.576$) you must classify 3981 deer, while if you will permit $\alpha = 0.10$ ($t_\alpha = 1.645$) you must classify only 1624 deer.

You can also use two-stage sampling (page 000) to estimate the proportion p (Cochran 1977, pg. 79). Take a first sample of size n , and calculate a preliminary estimate of \hat{p} . Then in the second sample take additional samples to make a total sample size of:

$$n = \left(\frac{\hat{p}_1 \hat{q}_1}{v} \right) + \left(\frac{3 - 8\hat{p}_1 \hat{q}_1}{\hat{p}_1 \hat{q}_1} \right) + \left(\frac{1 - 3\hat{p}_1 \hat{q}_1}{v n_1} \right) \quad (7.15)$$

where:

$$\begin{aligned} n &= \text{Total final sample size needed to estimate proportion} \\ \hat{p}_1 &= \text{Proportion of } x \text{ types in first sample} \\ \hat{q}_1 &= 1 - \hat{p}_1 \\ v &= d^2 / t_\alpha^2 \text{ (for } \alpha = .05, v \approx d^2 / 3.84; \text{ for } \alpha = .01, v \approx d^2 / 6.64) \\ d &= \text{Desired margin of error in } \hat{p} \end{aligned}$$

If you are sampling a finite population of size N , you may then correct your estimated sample size n by equation (7.6) using the finite population correction to reduce the actual sample size needed. For example, suppose the deer population is known to be only about 1500 animals. From equation (7.6):

$$n^* = \frac{n}{1 + n/N} = \frac{2305}{1 + 2305/1500} = 909 \text{ deer}$$

so a much smaller sample is required in this example if there is a finite population.

Wildlife managers often measure attributes of wildlife populations in a slightly different way with "population ratios". These are expressed, for example, as fawns/100 does, or males/100 females*. For ratios of this type, Czaplewski *et al.* (1983) have presented a useful series of charts and graphs to determine how large a sample size you need to attain a certain level of precision. In all cases they have assumed sampling without replacement. Figure 7.2 shows their recommended sample sizes for a range of total population sizes and for 90% confidence intervals. Note that if the total population is below 2000-4000 animals, the finite population correction (eq. 7.6) reduces the required sample size considerably. If population size exceeds 12,000, for all practical purposes it can be considered infinite.

One example will illustrate the use of Figure 7.2. Assume you have a deer population of about 3000 animals, and you expect a sex ratio of 70 males per 100 females. From Figure 7.2, you should classify at least 520 deer and at most 1350 deer**. In the first case, with $n = 520$, you will achieve an estimate of the population ratio of approximately ± 9 per 100 (i.e. 61 males/100 females to 79 males per 100 females). In the second case, with $n = 1350$, you will achieve an estimate of the population ratio of approximately ± 5 per 100. The general message is that to achieve a high precision in estimating ratios, you need to take large samples.

* The more conventional notation is as a proportion or percentage; e.g., if 70 males per 100 females, then $70/(100+70)$ or 0.412, or more usually 41% males in the population.

** Note that you can get the same result using equation (7.14).

Precision corresponding to curve	Population ratio				
	20:100	40:100	50:100	70:100	100:100
A	± 4:100	± 6:100	± 7:100	± 9:100	± 13:100
B	± 3:100	± 4:100	± 5:100	± 7:100	± 10:100
C	± 2:100	± 3:100	± 4:100	± 5:100	± 7:100

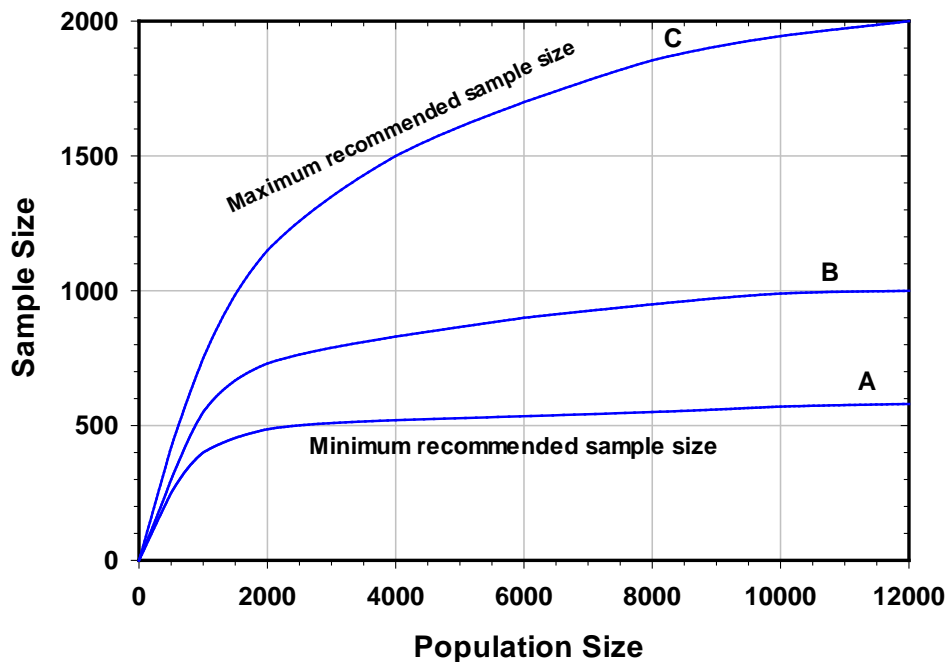


Figure 7.2 Recommended sample sizes for estimating a “population ratio” of two sex or age classes in populations of finite size. Three curves are shown for the three levels of precision given at the top of the graph. Precision is specified here as a 90% confidence interval. A finite population correction is included in these curves, and sampling without replacement is assumed. (From Czaplewski et al., 1983.)

7.2.2 Counts From A Poisson Distribution

Sample size estimation is very simple for any variable which can be described by the Poisson distribution in which the variance equals the mean (Chapter 4). From this it follows that:

$$CV = \frac{s}{\bar{x}} = \frac{\sqrt{s^2}}{\bar{x}} = \frac{\sqrt{\bar{x}}}{\bar{x}}$$

or

$$CV = \frac{1}{\sqrt{\bar{x}}}$$

Thus from equation (7.5) assuming $\alpha = 0.05$:

$$n \approx \left(\frac{200CV}{r} \right)^2 = \left(\frac{200}{r} \right)^2 \frac{1}{\bar{x}} \quad (7.16)$$

where:

$$\begin{aligned} n &= \text{Sample size required for a Poisson variable} \\ r &= \text{Desired relative error (as percentage)} \\ CV &= \text{Coefficient of variation} = 1/\sqrt{\bar{x}} \end{aligned}$$

For example, if you are counting eggs in starling nests and know that these counts fit a Poisson distribution and that the mean is about 6.0, then if you wish to estimate this mean with precision of $\pm 5\%$ (width of confidence interval), you have:

$$n \cong \left(\frac{200}{5} \right)^2 \left(\frac{1}{6.0} \right) = 266.7 \text{ nests}$$

Equation (7.16) can be simplified for the normal range of relative errors as follows:

$$\begin{aligned} \text{For } \pm 10\% \text{ precision} \quad n &\cong \frac{400}{\bar{x}} \\ \text{For } \pm 25\% \text{ precision} \quad n &\cong \frac{64}{\bar{x}} \\ \text{For } \pm 50\% \text{ precision} \quad n &\cong \frac{16}{\bar{x}} \end{aligned}$$

Note that these are all expressed in terms of the width of the confidence interval and all are appropriate only for variables that fit a Poisson distribution.

7.2.3 Counts From A Negative Binomial Distribution

Since many animal and plants have a clumped pattern, quadrat counts will often be described by the negative binomial distribution instead of the Poisson (Chapter 4). To estimate the sample size for a series of counts which fit a negative binomial, you must know two variables and decide two more:

1. Mean value you expect in the data (\bar{x})

2. Negative binomial exponent (k)
3. Desired level of error (r) as a percentage
4. Probability (α) of not achieving your desired level of error

The level of error is set as an expected confidence interval, for example as $\pm 20\%$ of the mean. Since the variance of the negative binomial is given by:

$$s^2 = \bar{x} + \frac{\bar{x}^2}{k}$$

we can substitute this expected variance into equation (7.1) and do some algebra to derive an estimated sample size:

$$n = \frac{(100t_\alpha)^2}{r^2} \left(\frac{1}{\bar{x}} + \frac{1}{k} \right) \quad (7.17)$$

where:

- n = Sample size required for a negative binomial variable
- t_α = Student's t value for $n - 1$ degrees of freedom for α probability
- \bar{x} = Estimated mean of counts
- k = Estimated negative binomial exponent
- r = Desired level of error (percent)

For most practical purposes we can assume $t_\alpha = 2.0$ (for 95% confidence limits). As an example, using the data from Box 4.2 (page 000) on the black-bean aphid with $\bar{x} = 3.46$, $\hat{k} = 2.65$, assume we would like to have confidence limits of $\pm 15\%$ of the mean:

$$n \cong \frac{200^2}{15^2} \left(\frac{1}{3.46} + \frac{1}{2.65} \right) = 118.5 \text{ quadrats (stems)}$$

As with the Poisson distribution, we can simplify equation (7.17) for the normal range of relative errors used in ecological research:

$$\text{For } \pm 10\% \text{ precision} \quad n \cong 400 \left(\frac{1}{\bar{x}} + \frac{1}{k} \right)$$

$$\text{For } \pm 25\% \text{ precision} \quad n \cong 64 \left(\frac{1}{\bar{x}} + \frac{1}{k} \right)$$

$$\text{For } \pm 50\% \text{ precision} \quad n \cong 16 \left(\frac{1}{\bar{x}} + \frac{1}{k} \right)$$

By comparing these formulas with those developed above for the Poisson, you can see the extra sampling required to attain a given confidence interval with the negative binomial. For example, if you had erroneously assumed that the black bean aphids had a Poisson distribution with mean 3.46, you would estimate for 95% confidence limits of $\pm 15\%$ that you would require a sample size of 51 quadrats, rather than the correct value of 118 quadrats calculated above. This is a vivid illustration of why the tests outlined in Chapter 4 are so critical in ecological sampling and the design of experiments.

7.2.4 Counts From A Lognormal Distribution

Many ecological variables are not well described by the normal distribution because they represent rates and are lognormal in distribution (see Chapter 15, page 000). By transforming lognormal data with a logarithmic transform, the transformed data become more nearly a normal distribution. The question we address here is how to determine the sample size needed for estimating the mean of a variable that has a lognormal distribution.

We proceed by transforming the original data as follows:

$$Y = \log(X) \tag{7.18}$$

where Y = transformed data

X = original data

Note that the base of logarithms used does not matter, so that \log_e or \log_{10} can be used as convenient. Next, with the Y as data, calculate the usual mean and variance of Y with the conventional formulas. Olsson (2005) showed that the best confidence interval for Y is given by:

$$\bar{Y} + \frac{s^2}{2} \pm t \sqrt{\left(\frac{s^2}{n} + \frac{s^4}{2(n-1)} \right)} \quad (7.19)$$

where s^2 = variance of Y

Convert these confidence limits to the original measurement scale by taking anti-logs.

To use this relationship to produce an estimate of desired sample size once you have an estimate of the mean \bar{Y} and the variance of Y for your data, decide on the width of confidence interval desired and by trial-and-error in Excel or similar program enter sample sizes and determine whether or not they satisfy your desired width of the confidence interval. Box 7.1 provides an example of this empirical approach to sample size determination for a lognormal variable.

Box 7.1 ESTIMATING SAMPLE SIZE FOR A VARIABLE THAT HAS A LOGNORMAL DISTRIBUTION

A forest ecologist in a pilot study measured the number of tree seedlings in 20 plots in a regenerating forest, and wishes to know if 20 plots are sufficient for obtaining estimates of mean density within a lower confidence limit about 25% of the mean. She measured the following numbers from 20 quadrat counts:

Sample quadrat #	Observed count.	Log _e (observed count)
1	199	5.2933
2	38	3.6375
3	75	4.3174
4	81	4.3944
5	93	4.5325
6	23	3.1354
7	357	5.8777
8	275	5.6167
9	218	5.3844
10	199	5.2933
11	70	4.2484
12	726	6.5875
13	401	5.9939

14	308	5.7300
15	210	5.3471
16	213	5.3612
17	253	5.5333
18	654	6.4831
19	41	3.7135
20	39	3.6635
Sum	4473	100.14454
Mean	223.65	5.0073
Variance	38118.87	0.9808

Estimate the 95% confidence limits from equation (7.19):

$$\begin{aligned} \bar{Y} + \frac{s^2}{2} \pm t \sqrt{\left(\frac{s^2}{n} + \frac{s^4}{2(n-1)} \right)} \\ = 5.0073 + \frac{0.9808}{2} \pm 2.093 \sqrt{\left(\frac{.9808}{20} + \frac{(0.9808)^2}{2(20-1)} \right)} \\ = 5.4977 \pm 0.5682 \text{ or } 4.9294 \text{ to } 6.0659 \end{aligned}$$

These log values can be converted back to the original count scale by anti-logs to give a mean of 244.1 with 95% confidence intervals of 138.3 to 430.9: Note that these are always asymmetric confidence intervals (- 43%, +76%).

Given these pilot data, you can now determine what size of sample you require to achieve a lower 95% confidence interval that is 25% of the mean. Solve equation (7.19) iteratively for a hypothetical range of sample sizes, for example $n = 50$:

$$\begin{aligned} \bar{Y} + \frac{s^2}{2} \pm t \sqrt{\left(\frac{s^2}{n} + \frac{s^4}{2(n-1)} \right)} \\ = 5.0073 + \frac{0.9808}{2} \pm 2.093 \sqrt{\left(\frac{.9808}{50} + \frac{(0.9808)^2}{2(50-1)} \right)} \\ = 5.4977 \pm 0.5682 \text{ or } 5.1551 \text{ to } 5.8402 \end{aligned}$$

which gives 95% confidence intervals of 173.3 to 342.8, or - 29% to +41%. Repeat these calculations several times until you converge to an $n = 71$ which gives: 95% confidence limits: 5.2104 to 5.7850, or in the original units 183.2 to 325.4, or in percentages -25% to + 33% of the estimated mean. Further sampling from this or similar forest sites should thus consider sample sizes of approximately 70 quadrats.

7.3 SPECIALIZED ECOLOGICAL VARIABLES

Some variables that ecologists estimate, like population density, are basically continuous variables that are estimated in indirect ways. For example, we might use a line transect technique (page 000) to estimate population density. Such derived variables are more difficult to analyze in order to decide in advance how big a sample you need. But some approximate techniques are available to help you plan a field study.

7.3.1 Mark-and-Recapture Estimates

In Chapter 2 we analyzed the simplest mark-recapture method (the Petersen method) in considerable detail, and we included the Robson and Regier (1964) charts (Figures 2.3, 2.4, page 000) for estimating the number of animals that need to be marked and recaptured to achieve a given level of accuracy in the estimate of population size.

An alternative procedure is to use the coefficient of variation of the estimated population size to estimate the required sample size directly. Seber (1982 p. 60) gives the coefficient of variation of the Petersen population estimate as:

$$CV(\hat{N}) \cong \frac{1}{\sqrt{R}} = \frac{1}{\sqrt{MC/N}} \quad (7.20)$$

where:

- R = Expected number of marked animals to be caught in the second Petersen sample
- M = Number of animals marked and released in the first Petersen sample
- C = Total number of animals caught in the second Petersen sample
- \hat{N} = Estimated population size

This formula can be used in two ways. First, if you have started a Petersen sample and thus know the value of M , and you can guess the approximate population size, you can determine C for any desired level of the coefficient of variation. For example, if you have marked and released 200 fish (M) and you think the population is about 3000 (N), and you wish the coefficient of variation of the estimated population size to

be about $\pm 25\%$ (corresponding to a level of precision of $\pm 50\%$ *), we have from equation (7.18):

$$CV(\hat{N}) = 0.25 = \frac{1}{\sqrt{MC/N}} = \frac{1}{\sqrt{200C/300}}$$

Solving for C, we obtain:

$$C = \frac{3000}{(200)(0.25)^2} = 240 \text{ fish}$$

so we should capture about 240 fish in our second Petersen sample and we would expect in these 240 fish to get about 16 marked individuals (R). This is only a crude approximation because equation (7.18) assumes large sample sizes for R and to be conservative you should probably take a somewhat larger sample than 240 fish.

A second use for this equation (7.18) is to get a rough estimate of the number of marked individuals (R) that should be obtained in the second Petersen Sample. In this case sampling might continue until a prearranged number of recaptures is obtained (Eberhardt 1978). For example, suppose in the fish example above you wish to have a coefficient of variation of (\hat{N}) of 0.05 (5%). From equation (7.18):

$$CV(\hat{N}) = 0.05 = \frac{1}{\sqrt{R}}$$

Solving for R :

$$R = \frac{1}{(0.05)^2} = 400 \text{ fish}$$

If you wish to obtain such a high degree of accuracy in a Petersen estimate you must recapture 400 *marked* fish, so clearly you must mark and release *more* than 400 in the first Petersen sample. These examples assume a large population so that the finite population correction is negligible.

* Note that the coefficient of variation for population estimates is equal to approximately one-half the relative level of precision (r) defined above for $\alpha = 0.05$ (page 246).

The Schnabel method of population estimation utilizes a series of samples (e.g. Table 2.2 page 000) and can be readily adapted to a predetermined level of precision. The coefficient of variation for a Schnabel estimate of population size is, from Seber (1982, p. 190):

$$CV(\hat{N}) \cong \frac{1}{\sqrt{\sum R_t}} \quad (7.21)$$

where:

$CV(\hat{N})$ = Expected coefficient of variation for Schnabel estimate of population size

R_t = Number of marked individuals caught in sample t

This formula is similar to equation (7.18). It can be used as follows to decide when to stop sampling in a Schnabel experiment. Consider the data in Table 2.1 (page 000). After June 11 (sample 10) there was a total of 13 recaptures ($\sum R_t$). Thus:

$$CV(\hat{N}) = \frac{1}{\sqrt{13}} = 0.277$$

and thus the 95% confidence limits would be about twice 0.277 or $\pm 55\%$ of the estimated population size \hat{N} . By the end of this experiment on June 15 there were 24 recaptures, and thus

$$CV(\hat{N}) \cong \frac{1}{\sqrt{24}} = 0.204$$

so the 95% confidence limits would be approximately $\pm 40\%$ of \hat{N} . If you wanted to reduce the confidence interval to approximately $\pm 20\%$ of \hat{N} you would have to continue sampling until $\sum R_t = 100$. These estimates of sample size are only a rough guide because they ignore finite population corrections, and equation (7.19) assumes a normal approximation that is not entirely appropriate when sample sizes are small. The important point is that they provide some guidance in planning a marking experiment using the Schnabel method. There are no simple methods available to estimate sample sizes needed to attain a given level of precision with the

Jolly-Seber model. At present the only approach possible is to use a simulation model of the sampling process to help plan experiments. Program MARK (White 2008, Cooch and White 2010) allows user inputs of certain sampling rates in a Jolly-Seber model and will simulate the estimation so you can see the levels of precision obtained. Roff (1973) and Pollock *et al.* (1990) have emphasized that sampling intensity must be very high to attain good levels of precision with Jolly-Seber estimation.

7.3.2 Line Transect Estimates

If you wish to design a line transect survey, you need to determine in advance what length of transect (L) or what sample size (n) you need to attain a specified level of precision. We discussed in Chapter 5 the general methods of estimation used in line transect surveys and in aerial surveys. How can we estimate the sample size we require in line transect surveys?

If we measure the radial distances (r_i) to each animal seen (see Figure 5.2 page 000), Eberhardt (1978a) has shown that the coefficient of variation of the density estimate is given by:

$$CV(\hat{D}) \approx \sqrt{\frac{1}{n} \left[1 + CV^2 \left(\frac{1}{r_i} \right) \right]} \quad (7.22)$$

where:

$$\begin{aligned} CV(\hat{D}) &= \text{Coefficient of variation of the line transect density estimate } D \\ &\quad \text{of equation (5.2)} \\ n &= \text{Sample size} \\ CV^2(1/r_i) &= (\text{Coefficient of variation of the reciprocals of the radial} \\ &\quad \text{distances})^2 \end{aligned}$$

There are two possibilities for using this equation. The coefficient of variation of the reciprocals of the radial distances can be estimated empirically in a pilot study, and this value used to solve equation (7.22). Alternatively, Eberhardt (1978) and Seber (1982, p. 31) suggest that the coefficient of variation of the reciprocals of radial distances is often in the range of 1-3, so that equation (7.20) can be simplified to:

$$CV(\hat{D}) = \sqrt{\frac{b}{n}} \quad (7.23)$$

where $b =$ a constant (typically 1-4)

For any particular study, you could assume b to be in the range of 1-4 and be conservative by estimating it to be relatively high. For example, suppose that you wish to have a coefficient of variation of density of about $\pm 10\%$ so that the 95% confidence interval is about $\pm 20\%$ of the population size. A pilot study has determined that $b = 2$. From equation (7.23):

$$CV(\hat{D}) = \sqrt{\frac{2}{n}} \quad (\text{assumes } b = 2)$$

$$0.10 = \sqrt{\frac{2}{n}} \quad \text{or} \quad n = 200$$

Alternatively, if the perpendicular distances (x_j) are used in a line transect estimator (such as equation 5.7), Eberhardt (1978) suggests that the approximate coefficient of variation is:

$$CV(\hat{D}) = \sqrt{\frac{4}{n}} \quad (\text{assumes } b = 4) \quad (7.24)$$

which is exactly double that given above for radial distances.

One alternative strategy to use with line transect data is two-step sampling in which the first sample is used as a pilot survey. Burnham *et al.* (1980, pp. 31-37) discuss the details of planning a line transect study by two-stage sampling. They provide a rule-of-thumb equation to predict the length of a line transect you would need to achieve a given precision:

$$\hat{L} = \frac{b}{[CV(\hat{D})]^2} \left(\frac{L_1}{n_1} \right) \quad (7.25)$$

where:

- \hat{L} = Length of total survey line required
 $b \approx$ A constant between 1.5 and 4 (recommend $b = 3$)
 $CV(\hat{D})$ = Desired coefficient of variation in the final density estimate
 L_1 = Length of pilot survey line
 n_1 = Number of animals seen on pilot survey line

There is some disagreement about the value of b in ecological line transect studies (Eberhardt 1978a, Burnham *et al.* 1980). It may be between 1.5 and 4 but is more likely to be about 3, so Burnham *et al.* (1980, p. 36) recommend using $b = 3$ to be conservative or $b = 2.5$ to be more liberal. The ratio (L_1/n_1) could be known from earlier studies and will vary greatly from study to study. For example, in surveys of duck nests it could be 10 nests/km. If the ratio (L_1/n_1) is not known already, you can estimate it from a pilot study.

Equation (7.25) can also be used backwards to estimate what the coefficient of variation in density will be if you can sample a line of length L (due to financial constraints or time). To illustrate, consider the data in Box 5.1 (page 000) as data from a pilot line transect survey. From equation (7.25), assuming you wish to have a coefficient of variation in the final density of $\pm 10\%$ (so that a 95% confidence interval would be about $\pm 20\%$), and given that you have seen 12 animals in a line transect of 10 km and wish to be conservative so assume that $b = 3$:

$$\hat{L} = \frac{3}{(0.10)^2} \left(\frac{10}{12} \right) = 250 \text{ km}$$

Consequently you should plan a total transect of 250 km if you wish to achieve this level of precision. But suppose that you only have money to sample 60 km of transect. From equation (7.23) with a conservative estimate of $b = 3$:

$$\hat{L} = 60 = \frac{3}{[CV(\hat{D})]^2} \left(\frac{10}{12} \right)$$

Solving for $CV(\hat{D})$:

$$CV(\hat{D}) = 0.204$$

Thus you could achieve $\pm 20\%$ precision (or a 95% confidence interval of $\pm 40\%$) in your final density estimate if you could do only 60 km of transect in this particular example.

7.3.3 Distance Methods

There has been relatively little analysis done of the sample size requirements for population density estimates that use plotless sampling methods. Seber (1982 p. 43) suggests that the coefficient of variation of density estimates from distance sampling will be approximately:

$$CV(\hat{N}) = \frac{1}{\sqrt{sr - 2}} \quad (7.26)$$

where:

$CV(\hat{N})$ = Coefficient of variation for plotless sampling estimate of population density

s = Number of random points or random organisms from which distances are measured

r = Number of measurements made from each point (i.e., if only nearest neighbor $r = 1$)

Thus in sampling with plotless methods you can achieve the same level of precision with (say) 100 random points at which you measure only nearest-neighbors, as with 50 random points at which you measure the distance to the first and second nearest-neighbors.

For example, if you wish to achieve a coefficient of variation of $\pm 20\%$ for population density estimated from distance sampling (which will give a 95% confidence interval of about $\pm 40\%$), and you are measuring only nearest neighbors ($r = 1$), you have from equation (7.26):

$$CV(\hat{N}) = 0.20 = \frac{1}{\sqrt{s - 2}}$$

Solving for s :

$$s = 27 \text{ random points}$$

As with all estimates of population size, this should be taken as an approximate guideline to achieve the desired precision. If you are lucky, this sample size will give you better precision than you expect, but if you are unlucky, you may not achieve the precision specified.

7.3.4 Change-In-Ratio Methods

To estimate population density from the change-in-ratio estimator discussed previously in Chapter 3 (page 000), you should consult Paulik and Robson (1969) and Figures 3.1 and 3.2. These graphs will allow you to select the required sample size directly without the need for any computations.

7.4 STATISTICAL POWER ANALYSIS

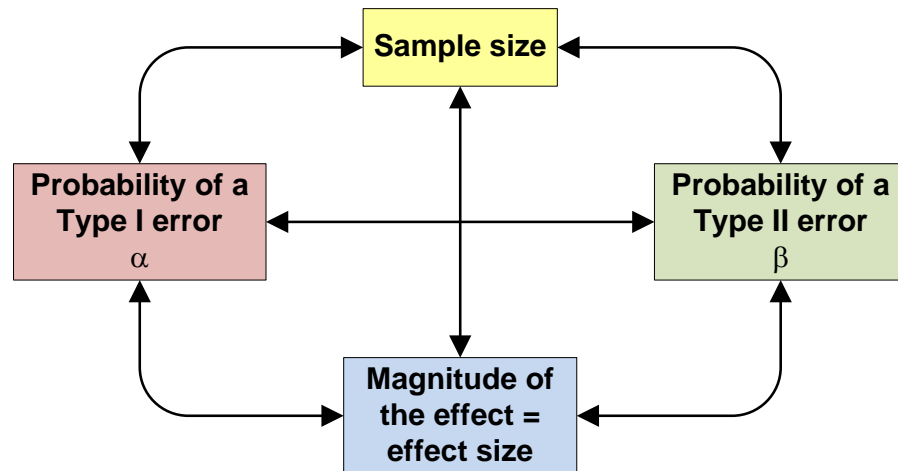
Much of the material presented in this chapter has been discussed by statisticians under the rubric of power analysis. Statistical power can be viewed most clearly in the classical decision tree involved in Neyman-Pearson statistical inference (Hurlbert and Lmobardi 2009):

State of real world	Decision	
	Do not reject null hypothesis	Reject the null hypothesis
Null hypothesis is actually true	Possibly correct decision but suspend judgment (probability = $1-\alpha$)	Type I error (probability = α)
Null hypothesis is actually false	Possibly a Type II error (probability = β)	Correct decision (probability = $(1-\beta)$ = power)

We will discuss in Chapter 15 the problems of the null hypothesis. Suffice to say for the present that the null hypothesis need not be 'no effect', and there can be a whole set of alternative hypotheses that are a range of values (e.g. that fish lengths differ by more than 10 mm, more than 20 mm, etc.). Most ecologists worry about α , the probability of a Type I error, but there is abundant evidence now that we should worry just as much or more about β , the probability of a Type II error (Peterman 1990, Fairweather 1991).

Power analysis can be carried out before you begin your study (*a priori* or prospective power analysis) or after you have finished (retrospective power analysis). Here we discuss *a priori* power analysis as it is used for the planning of experiments. Thomas (1997) discusses retrospective power analysis.

The key point you should remember is that there are four variables affecting any statistical inference:



These four variables are interconnected and once any 3 of them are fixed, the fourth is automatically determined. Looked at from another perspective, given any 3 of these you can determine the fourth. We used this relationship previously when we discussed how to determine sample size (page 278). We can now broaden the discussion and in particular discuss effect size, a key item in all ecological research.

7.4.1 Estimation of Effect Size for Continuous Variables

Effect size is a general term that is interpreted in different ways by ecologists and statisticians. To an ecologist the effect size is typically the difference in means between two populations. Fish in lake A are 2.5 kg heavier than fish in lake B. But statisticians prefer to convert these biological effect sizes to standardized effect sizes so they can discuss the general case. So for an ecologist the first thing you need to learn is how to translate biological effect sizes into statistical effect sizes.

The most common situation involves the means of two populations which we would like to test for equality. One measure of effect size is the standardized effect size, defined as:

$$D = \frac{\bar{X}_A - \bar{X}_B}{s} \quad (7.27)$$

where

D = Standardized effect size or standardized difference in means
 \bar{X}_A, \bar{X}_B = Postulated mean values for group A and group B
 s = Standard deviation of the variable being measured

D is a dimensionless number, so it can be used for any continuous variable that has an approximately normal distribution. Note that by scaling our means in relation to the standard deviation we implicitly bring the variability of the data into our power analysis. Note also that this is not the only possible effect size measure that can be used (Hurlbert 1994). It may be more meaningful ecologically to use $\bar{X}_A - \bar{X}_B$ as the effect size, or $\frac{\bar{X}_A - \bar{X}_B}{\bar{X}_B}$ as better measures

The effect size that is ecologically significant will depend on the specific situation, and is part of the overall problem of defining biological significance in contrast to statistical significance. Cohen (1988) has provided very general guidelines for effect sizes for comparison of two means. It is important to measure effects with respect to specific ecological situations or specific ecological hypotheses you are testing. Depending on the questions you are investigating, 'small' effects may be critically important, or 'large' effects may be relatively unimportant.

For the analysis of variance in which we need to compare more than two means, we need to define a more general effect size. Cohen (1988) defines the statistical effect size for several means as:

$$f = \frac{s_m}{s} \quad (7.28)$$

where

f = Statistical effect size for several means
 s_m = Standard deviation of the population means
 s = Standard deviation within the populations

In the usual analysis of variance, the standard deviation s is assumed to be the same in all groups being compared and is computed as the square-root of the “error” variance. The standard deviation of the population means is easy to determine when all groups have equal sample sizes (a balanced design). From Cohen (1988, p. 275):

$$\hat{s}_m = \sqrt{\frac{\sum_k (\bar{x}_i - \bar{x})^2}{k}} \quad (7.29)$$

where

- k = Number of means being compared
- \bar{x}_i = Mean value for group i
- \bar{x} = Grand mean value for all groups being compared

When several means are being compared, the spread of the means can be highly variable. For example, means may be equally spaced from high to low, or several means may be equal and only one divergent. This variation is measured by s_m .

When sample sizes are not equal in all groups, weighted means must be used in equation (7.29) as described in Cohen (1988, p. 359). This general approach can be applied to all analysis of variance problems including interaction effects. Cohen (1988) gives the details.

7.4.2 Effect Size for Categorical Variables

In the simplest case we wish to compare two proportions, for example the proportion of males in two lizard populations. The simplest measure of effect size for a biologist would be the expected difference in the proportions, e.g. 0.45 vs. 0.55 for the two populations. But there is a statistical problem in using this intuitive difference as a measure of effect size because the variance of the binomial distribution changes with the p -values, becoming maximal at $p = 0.5$. The statistical solution is to transform the proportions to φ values:

$$\varphi = 2 \arcsin \sqrt{p} \quad (7.30)$$

We can then define the effect size h as:

$$h = \varphi_1 - \varphi_2 \quad (7.31)$$

For purposes of data presentation this transformed effect size can be back-transformed to a proportion if desired.

7.4.3 Power Analysis Calculations

The calculations involved in power analysis are complex and are not presented in detail here. Figure 7.3 illustrates graphically how power analysis works in principle. In this case we have a fixed effect size and a fixed sample size, and the power of the test is the area of non-overlap of the expected statistical distributions. As with any statistical test, power calculations involve probabilities and we cannot be certain for any particular set of observations that we will in fact achieve the intended result.

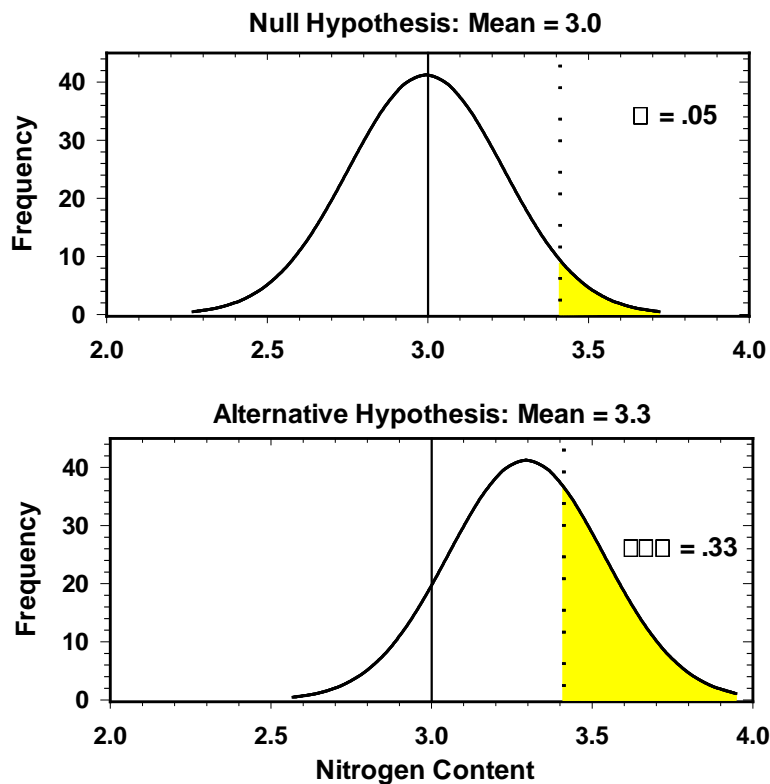


Figure 7.3 An example of how power calculations can be visualized. In this simple example a *t*-test is to be carried out to determine if the plant nitrogen level has changed from the base level of 3.0% (the null hypothesis) to the improved level of 3.3% (the alternative hypothesis). Given $n = 100$, $s^2 = 2.5$, and $\alpha = 0.05$ we can see that under the alternative hypothesis the distribution of possible means will follow a *t*-distribution that is shifted to the right. The shaded area is the power ($1 - \beta$) that can be expected in this study, and consequently the likelihood of a type II error is relatively high.

A power analysis consists of fixing four of the five variables – α , β , sample size, standard deviation, and magnitude of effect– and then calculating the fifth. It is often desirable to investigate a range of values rather than just one. For example, we may wish to know for a particular system the relationship between sample size and power for a given α and a specified magnitude of effect. Figure 7.4 illustrates one example for a study of parasitism rates.

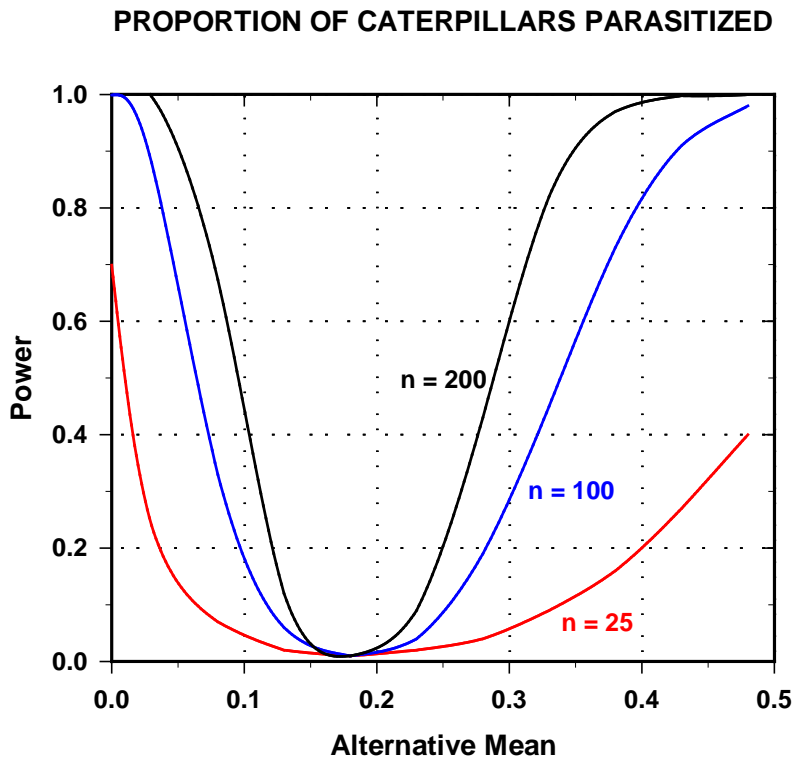


Figure 7.4 Power curves for a study of caterpillar parasitism rates in which the expected average parasitism rate is 18% and $\alpha = 0.01$. Power curves for sample sizes of $n = 25$, $n = 100$, and $n = 200$ are shown for a spectrum of alternative rates. Clearly for this ecological situation sample sizes of $n = 25$ cannot achieve high statistical power. Calculations were carried out with program PASS 6.0 (c.f. Thomas and Krebs 1997).

Power analysis is a relatively recent form of analysis in ecological studies but it has been discussed by statisticians for over 50 years. Only recently have ecologists begun to worry about Type II errors arising from studies in which the power to reject the null hypothesis is very low. Weak studies, studies with low power, may be difficult to avoid if money or time are limited, but we should be aware of the relative precision of our decisions. Failing to reject the null hypothesis should not be confused with

accepting the truth of the null hypothesis. Power analysis is useful precisely because it forces ecologists to provide specific alternative hypotheses in order to estimate power, and these alternative hypotheses must reflect ecological theory or ecologically significant effects.

Power is relatively easy to compute now even for complex experimental designs because of the personal computer revolution. Power calculations can be obtained from a variety of specific programs that are available (G*Power 3, nQuery, PASS 12) commercially or as freeware from a variety of sites. Search the web to see what is currently available. These programs can facilitate the planning of ecological studies and should be a part of every ecologist's repertoire (Thomas and Krebs 1997).

7.5 WHAT TO DO IF NOTHING WORKS

Some situations will arise in any ecological study in which none of the methods outlined above seem to work. This may be because you have a complex sampling problem or because you do not have any of the background information needed to use the techniques given in this chapter. In this case you can do one of two things. First, consult a professional statistician. There are many clever techniques that statisticians have developed for estimating sample size in complex experimental designs, and they are known to professionals but beyond the scope of this book. In many cases it is possible to use computer-intensive randomization procedures like bootstrapping to calculate the statistical power of particular experimental designs. This is a rapidly developing area in statistics and particularly useful for ecologists.

Alternatively, you can adopt a completely empirical approach and decide the question in a step-wise manner. Every variable you can measure has an appropriate variance estimate attached to it, so that you can compute 95% confidence limits after each set of samples. This crude, empirical approach is more feasible now that computers can be taken directly into the field complete with programs to compute tedious formulae. The general procedure is simply:

1. Collect samples for 1 day or 1 week or whatever time unit is relevant to study.
2. Compute at the end of that period the mean estimate and its 95% confidence interval for all the data gathered to date.

3. Decide on the basis of your desired level of precision whether the confidence interval is
 - (a) sufficiently small - go to **4**;
 - (b) still too large - go to **1**.
4. Stop and compute final estimates.

Figure 7.5 illustrates this simple procedure for counts of red alder trees in 4 m^2 quadrats. Clearly as sample size grows, the confidence interval will shrink, and the only decision you must make is when to stop. Note that you must *not* use this procedure to calculate a test of significance (like the t -test) and continue sampling until you get the answer you want.

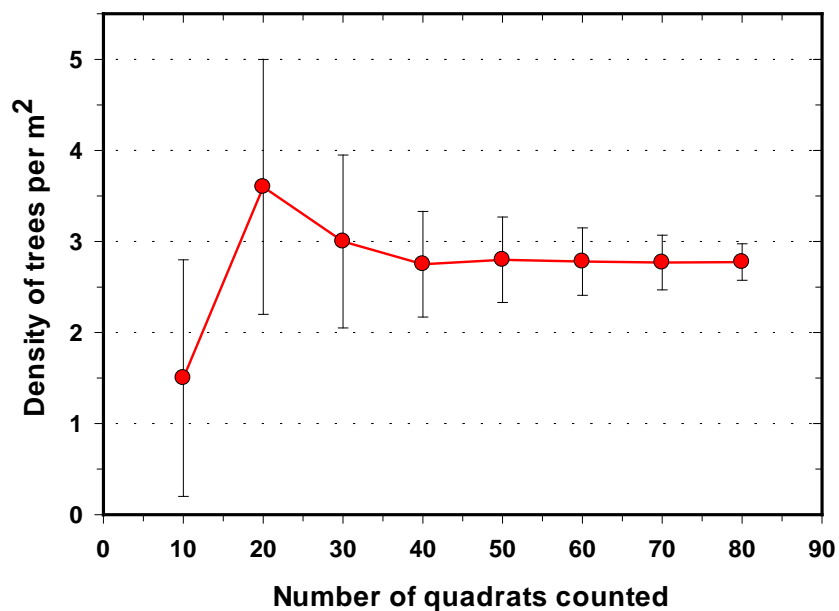


Figure 7.5 An empirical approach to determining how large a sample to take. An ecology class counted red alder trees on an area undergoing secondary succession. After each 10 quadrats were counted, the mean and 95% confidence interval were plotted. Sampling was continued until the confidence interval was judged to be sufficiently small.

This approach can be considered a type of *sequential sampling* which will be described more fully in Chapter 9.

The approach used in Figure 7.5 will not work when it takes a long time to process samples, as, for example, with plankton counts taken on oceanic cruises. In these cases you must obtain prior information to plan your sampling, either from

earlier studies, or from a pilot experiment whose sole purpose is to give initial estimates of statistical parameters needed in the equations listed in this chapter.

In some ecological areas enough wisdom has accumulated to allow empirical rules to be formulated for sampling even though it may take a long time to process samples. For example, Downing (1979) suggested that for benthic organisms in freshwater lakes and large rivers - no matter what the species, substrate, or type of sampling machinery - you can predict the standard error of a series of samples from the multiple regression equation:

$$\text{S.E.} = \frac{\text{antilog}(0.581 + 0.696 \log \bar{x} - 2.82 \times 10^{-4} A)}{\sqrt{n}} \quad (7.32)$$

where:

- S.E. = Standard error of mean density estimate
- \bar{x} = Mean density in numbers per m²
- A = Area removed by benthic sampler in cm² (sample area)
- n = Sample size

You can solve this equation for sample size n for any particular application to benthic sampling, or consult Table 7.4 which gives sample sizes needed to achieve a confidence interval of about $\pm 40\%$ of the mean. This type of approximate rule-of-thumb might be useful for other ecological sampling methods in which some prior information is useful in planning field studies (Taylor 1980).

Similarly, for marine and freshwater zooplankton samples, Downing *et al.* (1987) found that data from over a thousand sets of replicate samples could all be described by the general relationship:

$$s^2 = 0.745m^{1.622}v^{-0.267} \quad (7.33)$$

where:

- s^2 = Estimated variance among replicate zooplankton samples
- m = Mean number of organisms per liter
- v = Volume of the sampler you are using (in liters)

TABLE 7.4 NUMBER OF REPLICATE SAMPLES NEEDED TO SAMPLE THE BENTHIC INVERTEBRATES OF LAKES AND LARGE RIVERS IN
--

RELATION TO SIZE OF SAMPLER AND DENSITY OF ORGANISMS STUDIED ^a							
---	--	--	--	--	--	--	--

Density (number/m ²)	Size of sampler (cm ²)						
	20	50	100	250	500	750	1000
30	45	43	40	33	24	17	12
50	33	32	30	24	18	13	9
100	22	21	19	16	12	8	6
300	11	11	10	8	6	4	3
500	8	8	7	6	4	3	2
1,000	5	5	5	4	3	2	<2 ^b
5,000	<2	<2	<2	<2	<2	<2	<2
10,000	<2	<2	<2	<2	<2	<2	<2

^a These sample sizes are designed to achieve a precision of about $\pm 40\%$ of the mean for the 95% confidence interval.

^b Clearly you should always take at least 2 samples even at high densities.

Source: Downing (1979).

This empirical relationship can be used to estimate the number of samples you need to take to achieve a specified level of precision:

$$\hat{n} = 0.754m^{-0.378}v^{-0.267}p^{-2} \quad (7.34)$$

where:

- \hat{n} = Estimated sample size needed for zooplankton sampling
- m = Estimated number of organisms per liter
- v = Volume of the sampler you are using (in liters)
- p = S.E./ m = Relative level of precision desired (as a proportion)
- S.E. = Estimated standard error of the zooplankton counts (from equation 7.33)

Once you have decided on the level of precision you need (p), you must guess the density of zooplankton you wish to sample (m) to obtain an estimate of the sample size you will need. Downing *et al.* (1987) discuss how much time and money can be saved by adopting this approach to zooplankton sampling in both marine and in freshwater environments.

The approach captured in Table 7.4 and equation (7.33) should be used in the preliminary planning stages of a zooplankton study, and the assumptions should be checked for any particular ecological study.

There is no doubt money wasted in ecological research by the use of poorly designed studies and that is what this chapter attempts to correct. In some studies too much work is done and estimates are accurate to $\pm 5\%$ when $\pm 20\%$ will do. Much more frequently, however, too little work is done and the data have such wide confidence limits as to be useless. If you are trying to test a hypothesis that predicts a 10% change in metabolism association with water quality, and your sample sizes are fixed so that 95% confidence limits of $\pm 40\%$ are expected, you might as well not do the experiment. So the message is simple: *plan ahead*.

To assist in planning, Program SAMPLE in Appendix 2 (page 000) will compute the sample sizes required for a specified level of precision for all the formulas given in this chapter.

7.6 SUMMARY

The most common question in ecological research is *how large a sample should I take?* This chapter attempts to give a general answer to this question by providing a series of equations from which sample size may be calculated. It is always necessary to know something about the population you wish to analyze and this leads into a statistical Catch-22 unless you use guesswork or prior observations. You must also make some explicit decision about how much "error" you will allow in your estimates (or how small a confidence interval you wish to have).

For continuous variables like weight or length, we can assume a normal distribution and calculate the required sample sizes for means and for variances quite precisely.

For counts, we need to know the underlying statistical distribution - binomial, Poisson, or negative binomial - before we can specify sample sizes needed. Patchy patterns described by the negative binomial require much larger sample sizes to describe precisely than do random patterns.

Estimates of population size and density from mark-recapture methods can be achieved to a specified level of precision by the use of graphs, and a few simple rule-of-thumb approximations. Similar approximate methods for deciding on sample sizes are available for line transect studies, plotless sampling, and change-in-ratio estimators of population size.

Power analysis explores the relationships between the four interconnected variables α (probability of Type I error), β (probability of Type II error), effect size, and sample size. Fixing three of these automatically fixes the fourth, and ecologists should explore these relationships before they begin their experiments. Significant effect sizes should be specified on ecological grounds before a study is begun.

If all else fails, a completely empirical trial-and-error method can be used to decide on sample size by taking samples in steps and computing the confidence limits after each step. Computer intensive randomization methods can be used to investigate the statistical power of simple to complex experimental designs. The important message is always: *plan ahead*.

SELECTED REFERENCES

- Cochran, W.G. 1977. *Sampling Techniques*, Chapter 4, The estimation of sample size, pp. 72-88, John Wiley, New York.
- Cohen, J. 1992. A power primer. *Psychological Bulletin* **112**: 155-159.
- Downing, J.A. 1979. Aggregation, transformation, and the design of benthos sampling programs. *Journal of the Fisheries Research Board of Canada* **36**: 1454-1463.
- Eberhardt, L.L. 1978. Appraising variability in population studies. *Journal of Wildlife Management* **42**: 207-238.
- Kraemer, H.C. and S. Thiemann. 1987. *How many subjects?* Sage Publications, Newbury Park, California.
- Nakagawa, S., and Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews* **82**(4):591-605.
- O'Brien, C., van Riper, C., and Myers, D.E. 2009. Making reliable decisions in the study of wildlife diseases: using hypothesis tests, statistical power, and observed effects. *Journal of Wildlife Diseases* **45**(3): 700-712.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries and Aquatic Sciences* **47**: 2-15.

Smith, D.R., Hardy, I.C.W., and Gammell, M.P. 2011. Power rangers: no improvement in the statistical power of analyses published in *Animal Behaviour*. *Animal Behaviour* **81**(1): 347-352.

QUESTIONS AND PROBLEMS

- 7.1.** Estimate the sample size required for quadrat counts of goldenrod (*Solidago spp*) in an old field in which a random spatial pattern occurs, and the expected density on 1 m² quadrats is 3.32 plants/m². Assume that you wish to have a 95% confidence interval of $\pm 1\%$ of the mean.
- What changes in sampling must occur if you count 4 m² quadrats instead of 1 m² quadrats?
 - If you count 16 m² quadrats?
- 7.2.** A mammal ecologist wishes to compare the mean litter size of ground squirrels on a control area with that of squirrels living in a radiation-disposal site. She specifies the probability of a Type I error as 0.01 and the probability of a Type II error as 0.05. The standard deviation of litter sizes from previous ground squirrel studies has been 0.90 to 1.65. She wishes to pick up a difference between her two populations of 1.0 embryos or more. What sizes of samples should she take?
- What recommendation can you make to her if it is much more difficult to sample from site B than from site A, so that 3/4 of the total samples will come from site A?
- 7.3.** A limnologist wishes to estimate the size of a *Daphnia magna* population in a lake. He knows that replicate samples fit a negative binomial distribution and from past experience k is approximately 3.8. He wishes to know the mean density within a 99% confidence interval of $\pm 40\%$, and he expects the mean density to be about 18.0 animals/sample. How many samples should he take?
- If he can accept a 99% confidence interval of $\pm 50\%$, how much work can he save?
- 7.4.** A marine ecologist is doing an aerial census of porpoises off California. He expects to see about 0.04 porpoises per km of travel, and he needs to know population density within a 80% confidence belt of $\pm 30\%$. How long a transect should he plan?
- 7.5** A wildlife ecologist wishes to estimate the population ratio of fawns/100 does in an antelope population of northern Utah. She thinks there are between 8000 and 10000 antelope on her study area. She expects the ratio to be

approximately 80 fawns/100 does and she would like to achieve a 95% confidence belt of ± 5 fawns/100 does. How many animals should she count and classify?

(a) How would the sample size be affected if there were really only 4000 antelope in the area?

- 7.6.** A marine ecologist is planning a mark-recapture program using the Schnabel method to estimate the size of a beluga whale population. How many recaptures are needed to achieve a coefficient of variation of the population estimate of $\pm 20\%$? Discuss several possible ways of achieving this goal of getting the number of recaptures required.
- 7.7.** Compare the relative efficiency of using quadrats to estimate population density and using plotless sampling. What ecological factors and what statistical factors will influence your decision about which technique to use?
- 7.8.** Construct a power curve relating effect size (X -axis) to sample size (Y -axis) to answer the question: what sample size is required to recognize a specific effect of fertilizer on plant growth at $\alpha = 0.05$ and $\beta = 0.10$, when the expected mean is 4.8 g with expected standard deviation 0.5 ? Express the effect size in biological units (g). Table 7.3 can be utilized, or more specific power software packages.
- 7.9.** Discuss in general the question of how large a benthic sampling device you should use, given the recommendations in Table 7.4. Refer back to Chapter 3 and discuss exactly how you would make this decision statistically. What practical matters might influence you decision about size of sampler?