10 Estimating the Population Size

10.1 Introduction

We wish to estimate N, an unknown population size. We can estimate changes in N over a period of time.

Examples:

- How many people attended a political rally?
- How many deer are in a particular region, both before and after hunting season?
- How many large-mouth bass are in a particular reservoir?
- How many Siberian tigers (an endangered species) are in the eastern part of Russia.

One method used for estimating N is based on **dividing** the region into sub-regions and **counting** observations in a sample of these sub-regions (section 10.5).

Another method used for estimating N is often called **capture / recapture** (section 10.2).

10.2 Estimation of a Population Size Using Direct Sampling

Capture / Recapture

Suppose t deer are *tagged* during the *capture* stage in a particular area. What proportion of deer were tagged?

A couple of weeks later, a sample of n deer are *recaptured* from this same area. Among these n deer, s of them were previously tagged from the *capture* stage. What proportion of the *recaptured* deer are tagged?

For large sample sizes, should these two proportions be approximately the same (assuming independence among captures and recaptures)?

Our estimate of N is

$$\hat{N} = t\ n/s. \tag{10.1}$$

The estimated variance can be shown to be

$$\hat{V}(\hat{N}) = \frac{t^2\ n\ (n-s)}{s^3} \tag{10.2}$$

Exercise 10.6, p. 328: A game commission is interested in estimating the number of large-mouth bass in a reservoir. A random sample of 2876 bass is caught. Each bass is marked and released. One month later, a second sample of 2562 is caught. Suppose 678 of the bass have tags on the second sample.

(a) Estimate the total number of bass in the reservoir.

(b) Determine the bound on the error of estimation.

(c) Construct a 95% confidence interval on the total number of bass in the reservoir.

Interpretation: We are 95% confident that the population total number of large-mouth bass lies between 10,151.90 and 11,583.54. □

Homework p. 328: Exercise 10.5

10.5 Estimating Population Density and Size from Quadrat Samples

Divide and Count

Example: A large number of people are scattered relatively uniformly on a football field. We want to estimate the number of people on the field. Divide the field into 120 rectangles of equal areas. Count the number of people in ten of these rectangles, selected at random; 143 people are counted in this sample.

What type of sampling method is used here?

Example, similar to example 10.4, pp. 321-322: Florida has a serious problem with fire ants, and estimating the number of ant hills per unit area (at some key locations) is an important consideration in keeping track of population increases. Fifty sample quadrats of approximately 16 m^2 each gave the results in the table. The total number of quadrats in the population is 1000.

Number of hills	Frequency
0	13
1	8
2	12
3	10
4	5
5	2

(a) Save the data as individual observations in a vector y.

(b) Estimate the population **mean** number of ant hills per **16 m^2,** the bound on the error of estimation, and the confidence interval.

(c) Estimate the population **mean** number of ant hills per $\mathbf{m^2}$**,** the bound on the error of estimation, and the confidence interval.

(d) Estimate the population **total** number of ant hills, the bound on the error of estimation, and the confidence interval.

(e) Repeat parts (b), (c), and (d) using the `survey` package.

□

Homework p. 330: Exercises 10.19, 10.21