MH4511 Sampling & Survey

Tutorial 10 AY2025/26 Semester 1

Problem 10.1

Suppose that an SRS of 500 fish is caught from a lake; each is marked and released, and the fish are allowed to remix with the other fish in the lake. A second sample of 300 fish has 120 marked fish.

a) Estimate the total number of fish in the lake using the ratio estimator, along with a 95% confidence interval.
b) Estimate the total number of fish in the lake using the contingency table approach, along with a 95% confidence interval.

Problem 10.2

A study was conducted by asking moose hunters to collect fecal samples from brown bear. Each sample was genotyped, and the number of distinct individuals was found. In year 2001, 311 unique genotypes were obtained, and in 2002, the procedure was repeated with 239 unique genotypes were obtained. It was found that 165 of the individuals sampled in 2001 were also sampled in 2002.

a) Estimate the number of bears, along with a 95% confidence interval, treating the samples from 2001 and 2002 as independent SRSs. What assumptions are needed to use capture and recapture estimators of population size?

At the same time of the study, another survey was conducted using radio transmitters.

b) Fifty-six (56) bears in the area in 2001 had been followed with radio transmitters; 36 of these bears were represented in the 311 genotypes from the 2001 feces samples. Estimate the number of bears in 2001, along with a 95% confidence interval.
c) In 2002, 57 bears had radio transmitters, and 28 of them were among the 239 genotypes from the 2002 feces samples. Estimate the number of bears in 2002, along with a 95% confidence interval.

Problem 10.3

In a lake of N fish, n_1 of them tagged, the probability of obtaining m recaptured and $n_2 - m$ previously uncaught fish in a simple random sample of size n_2 is

$$L(N) = \frac{\binom{n_1}{m}\binom{N-n_1}{n_2-m}}{\binom{N}{n_2}}$$

The maximum likelihood estimator $\widehat{N}$ of N is the value which maximizes $L(N)$. Find the maximum likelihood estimator of N.

Hint: find the range of N such that $L(N) \geq L(N-1)$.

Note that $\widehat{N}$ is the value that make the observed value of m appear most probable if we know n_1 and n_2.