LAB PROJECT — COMPARISON OF ESTIMATORS

Setting

Suppose two armies are at war. For the sake of discussion, let's call them the Red Army and Blue Army. Placing yourself in the position of a general in the Blue Army, suppose you are faced with the task of estimating the number of tanks in the Red Army's tank fleet. Through intelligence, you learn that the Red Army tanks have sequential seriel numbers $1, 2, 3, ..., N$, and therefore N is the unknown number (i.e. unknown parameter) you need to estimate.

Data

A spy for the Blue Army will cross enemy lines and return with a random sample of n tank seriel numbers. We will assume that the random sampling of the n seriel numbers is conducted with replacement, and therefore the n numbers can be viewed as a collection of independent and identically distributed random variables. Let $X_1, X_2, ..., X_n$ denote the sample of n seriel numbers.

Estimating the Value of N — Theoretical Development

1 Find the method of moments estimator U of N.

Answer: $U = 2\overline{X} - 1$

2 Find $\mathrm{E}(U)$ and $\mathrm{Var}(U)$.

Answer: $\mathrm{E}(U) = N$, $\mathrm{Var}(U) = \frac{N^2-1}{3n}$

3 Show for large n that U has approximately a Normal$(N, \sqrt{\frac{N^2-1}{3n}})$ distribution.

4 Find the maximum likelihood estimator T of N.

Answer: $T = \max\{X_1, X_2, ..., X_n\}$

5 Find the cumulative distribution function, $F(t) = \mathrm{P}(T \leq t)$, for the variable T.

Answer:

$$F(t) = \begin{cases} 0, & \text{if } t < 1, \\ \left(\frac{\lfloor t \rfloor}{N}\right)^n, & \text{if } 1 \leq t \leq N, \\ 1, & \text{if } t > N \ , \end{cases}$$

where $\lfloor t \rfloor$ denotes the greatest integer function, sometimes called the floor function.

6 With an eye toward finding the value of $\mathrm{E}(T)$, show that if X is a random variable taking on non-negative integer values, then

$$\mathrm{E}(X) = \sum_{k=1}^{\infty} \mathrm{P}(X \geq k) .$$

7 Combine **5** and **6** to show that

$$\mathrm{E}(T) = \sum_{k=1}^{N} \left[1 - \left(\frac{k-1}{N} \right)^n \right] .$$

8 Using the theory of Riemann sums from calculus, demonstrate that $\mathrm{E}(T)/N$ is a left Riemann sum and furthermore show

$$\frac{\mathrm{E}(T)}{N} \approx \int_0^1 (1 - x^n)\, dx .$$

9 Using **8**, show that

$$\mathrm{E}(T) \approx \frac{n}{n+1} N .$$

Look at $n = 1, 2, 3$. Does the above approximation for $\mathrm{E}(T)$ make intuitive sense in these cases?

10 Now, we turn our attention toward the derivation of the variance of T. Show that if X is a random variable taking on non-negative integer values, then

$$\mathrm{E}(X^2) = \sum_{k=1}^{\infty} (2k-1) \mathrm{P}(X \geq k) .$$

11 Similar to the method of **7**, show that

$$\mathrm{E}(T^2) = \mathrm{E}(T) + N \sum_{k=1}^{N} 2 \left(\frac{k-1}{N} \right) \left[1 - \left(\frac{k-1}{N} \right)^n \right] .$$

12 Using the theory of Riemann sums again, show that

$$\mathrm{E}(T^2) \approx \frac{n}{n+1} N + \frac{n}{n+2} N^2 ,$$

and then find an approximate form for $\mathrm{Var}(T)$.

13 Problem **9** shows that T is a biased estimator for N. Create a new estimator of N, we'll call this estimator V, that is a modification of T and that is unbiased for N.

14 Find the variance for V and explore the relative sizes of variances of U, T, and V by taking ratios of their variances.

15 Summarize your findings of the characteristics of the three estimators of N we've studied. Which estimator would you use to estimate the size of the enemy's tank fleet? Rigorously justify your choice.