Math 152, Fall 2022
Jo Hardin
WU # 6
Tuesday 9/20/22

Your name: ________________________________

Names of people you worked with: ________________________________

Task:

How can a random sample of integers between 1 and N (with N unknown to the researcher) be used to estimate N? This problem is known as the German tank problem and is derived directly from a situation where the Allies used maximum likelihood to determine how many tanks the Axes had produced. See `https://en.wikipedia.org/wiki/German_tank_problem`.

1. The tanks are numbered from 1 to N. Working with your group, randomly select five tanks, without replacement, from the bowl. The tanks are numbered:

2. Think about how you would use your data to estimate N. (Come up with at least 3 estimators.) Come to a consensus within the group as to how this should be done. One person from your group will report out after the warm-up is over. Ideally, the person to report out will be someone who has not yet spoken in class this semester. Step-up if you haven't yet spoken. Step back if you speak regularly.

   Our estimates of N are:

   Our rules or formulas for the estimators of N based on a sample of n (here $n = 5$) integers are:

Assume the random variables are independently and identically distributed according to a discrete uniform. (Tbh, the *iid* model is with replacement, but the answers you get aren't much different than without replacement if $n << N$.)

$$X_i \sim P(X = x|N) = \frac{1}{N} \qquad x = 1, 2, \ldots, N \quad i = 1, 2, \ldots, n$$

3. What is the method of moments estimator of N?

4. What is the maximum likelihood estimator of N?

Solution:

1. Everyone has different tanks!
2. So many good answers!! See `https://m152-stat-theory.netlify.app/handout/tank_152.pdf` for some ideas.
3. To find the method of moments estimator, one needs notice that the first moment ($E[X]$) is a function of N. Solve for N as a function of the sample moment to find $\hat{N}$.

$$\begin{aligned} E[X] &= \overline{X} \\ \frac{N+1}{2} &= \overline{X} \\ \widehat{N} &= 2 \cdot (\overline{X} - 1) \end{aligned}$$

4. First, write down the joint likelihood for n observations:

$$\begin{aligned} f(x_i|N) &= \frac{1}{N} \quad\quad x_i \leq N \\ &= \frac{1}{N} I_{[1,N]}(x_i) \\ f(\underline{x}|N) &= \prod_{i=1}^{N} \frac{1}{N} I_{[1,N]}(x_i) \\ &= \frac{1}{N^n} I_{[1,N]}(\max(x_i)) \end{aligned}$$

Sketch $f(\underline{x}|N)$ as a function of N and notice that the function (i.e., the likelihood!) goes to zero immediately after $N > \max(x_i)$. Therefore, the MLE of N is $\max(x_i)$.