I. STUDENT VERSION

Title: German tank problem and its applications

Summary of the project:

The German Tank Problem refers to allied attempts during World War II, to estimate the monthly rate of German tank production based on a small sample of captured tanks. Using a small sample. the allied forces realized that several parts on the German Panzer IV, Panzer V, and Tiger I had sequential serial numbers on various parts including the gearbox and tires. Historical sources show that by the end of the war and through the use of the developed formula [1] and [5] that the allies had a fairly accurate picture of the number of tanks produced by German army. For example [4][, using conventional intelligence sources, allied estimated that 1400 tanks between June 1940 and September of 1942 was produced per month. Applying the formula to the serial number of captured tanks, the number was estimated to be 246 per month. The actual number based on German records was 245. In this project students will investigate the accuracy of the formula used to estimate the number of German tanks and will explore, through use of simple simulation, the utility of the method for estimating the size of a population for other applications.

Instructions to students:

- **Background information**: During World War II, the allied forces realized that several parts on the German Panzer IV, Panzer V, and Tiger I had sequential serial numbers on various parts including the gearbox and tires. Work of Statistician in Royal academy in England led to the following formula:
  If we let the number of captured tanks be represented by the letter K, and the maximum serial number we have seen be represented by the letter M, then we can estimate the maximum number of tanks with the equation $M + \frac{M}{K} - 1$.

- **Individual or group project**: This project will be assigned to teams of students

Procedure/plan of action to complete the project:

Basic study:

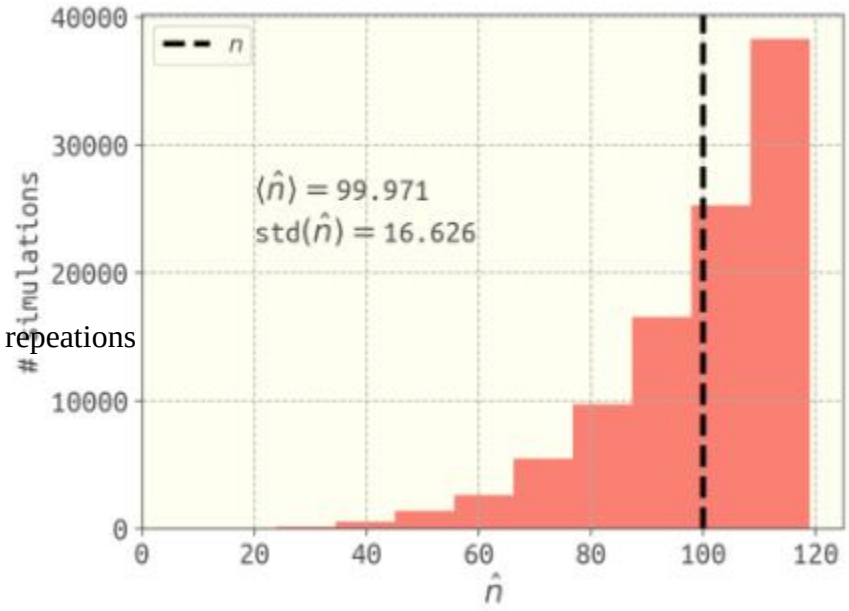
a. Look over the references [4] and [5] to become familiar with the problem and the solution(s).
b. Do a literature search to find more sources on the application, verification, and proof of the formula. See [6] for a simple proof.
c. Do a simple simulation: create a population size 100, for example use ping pong balls or marbles with numbers on them. Pick random small samples (3 or 4), estimate the size of

the population using German tank formula. Repeat the selection process at least 50 times. Form a histogram of the estimates for the size of the population. Is it symmetric or skewed? Explain why. What is the mean and standard deviation of the maximum(using German tank formula)? Is mean of the estimates close to actual mean?

d. Repeat (c), this time use a larger sample size(twice as part (c)).
e. Look for the applications of the formula on other fields.

Example(s):

Your histogram (part c) should look something like this (x-axis; estimate of the population size, y-axis: number of repetitions):



Expanded study:

Instead of a physical object(ping pong balls or marbles), use a software (see [2] and [3]) to pick random samples of size 3 or 4 to repeat parts (c) and (d)

Deliverables and evaluation:

-Students are expected to be prepared for the pre-project questions.
- Each team will present their results in a five-minute presentation.
-Students are expected to be prepared for the post-project discussions.

References

[1] "Statistics Example from World War 2." *Math Encounters Blog*, 11 Dec. 2015,

www.mathscinotes.com/2015/04/statistics-example-from-world-war-2/.

[2] Ensemble, The Simon. "The Data Generating Process (Dgp)." *Porous Cage Molecule*, simonensemble.github.io/2019-11/German -tank-problem.

[3] Johnson, Gavin R. "A Program to Help Quickly Simulate 'Tank Problem' Scenarios." *Gist*, 8 July 2021, gist.github.com/garJohnsonTLU/44b98ee9a5df13a8bb0059c3aaef2d47.

[4] "German Tank Problem." *Wikipedia*, Wikimedia Foundation, 27 June 2021, en.wikipedia.org/wiki/German tank problem.

[5] Grajalez, Carlos Gómez, et al. "Great Moments in Statistics." *Significance*, vol. 10, no. 6, 2013, pp. 28–28., doi:10.1111/j.1740-9713.20.

[6] Johnson, Roger, estimating size of a population, Carleton college, Minnesota https://web.archive.org/web/20140223104835/http://www.rsscse-edu.org.uk/tsj/wp-content/uploads/2011/03/johnson.pdf