

13.3 The German Tank Problem

WORKSHEET

The question

A hostile naval power possesses a large, unknown number N of submarines. Interception of radio signals yields a small number n of their identification numbers X_i ($i = 1, 2, \dots, n$), which are taken to be independent and uniformly distributed over the continuous range from 0 to N . Show that Z_1 and Z_2 , defined by

$$Z_1 = \frac{n+1}{n} \max\{X_1, X_2, \dots, X_n\} \quad \text{and} \quad Z_2 = \frac{2}{n} \sum_{i=1}^n X_i,$$

both have means equal to N .

Calculate the variance of Z_1 and of Z_2 . Which estimator do you prefer, and why?

Hints

Try the question on your own first. If you get stuck, read the hints *one at a time*, and go back to the question after each one. Hints marked *Pause and try* are the key steps.



Ask yourself If M is the largest of n independent uniform values on $[0, N]$, what is $P(M \leq m)$?

STAGE 2 Their variances



Ask yourself Find $E(M^2)$ and hence $\text{Var}(M)$. For Z_2 , what is the variance of a single uniform variable on $[0, N]$?

STAGE 3 Which is better?



Ask yourself Compare $\frac{1}{n(n+2)}$ with $\frac{1}{3n}$. How do they behave as n grows?

Try this next

- 1 Mean squared error.** The biased estimator M has $E(M - N)^2 = \text{Var}(M) + (E(M) - N)^2$. Compute it and compare with $\text{Var}(Z_1)$. Find the constant c that minimises $E(cM - N)^2$.
- 2 The minimum too.** Show that $(n + 1) \min\{X_i\}$ is also an unbiased estimator of N , and find its variance. Is it any good?
- 3 The discrete formula.** For serial numbers $1, \dots, N$, sample k without replacement, with maximum m . Show that $P(\max \leq m) = \binom{m}{k} / \binom{N}{k}$, and verify that $m + \frac{m}{k} - 1$ is unbiased when $k = 1$.
- 4 Tanks.** If 5 captured tanks have serial numbers 19, 40, 42, 60, 98, estimate the number of tanks made.

Further reading

- **Richard Ruggles & Henry Brodie**, “An empirical approach to economic intelligence in World War II”, *Journal of the American Statistical Association* 42 (1947). The original account.
- **John Rice**, *Mathematical Statistics and Data Analysis* (Duxbury), on bias, variance and sufficiency.
- **David Spiegelhalter**, *The Art of Statistics* (Pelican).

The full guided solution, and what the question leads to at university, is in Section 13.3:
https://keithmcnulty.github.io/step_book_mcnulty/ch13/german-tanks.html