Math 152, Fall 2018
Jo Hardin

Tank Estimators

How can a random sample of integers between 1 and N (with N unknown to the researcher) be used to estimate N?

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.

   Our estimates of N are:

   Our rules or formulas for the estimators of N based on a sample of n (in your case 5) integers are:

Assuming the random variables are distributed according to a discrete uniform. (Tbh, this 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?

Mean Squared Error

Most of our estimators are made up of four basic functions of the data: the mean, the median, the min, and the max. Fortunately, we know something about the moments of these functions:

$g(\underline{X})$	$E[\, g(\underline{X})\,]$	$Var(\, g(\underline{X})\,)$
$\overline{X}$	$\frac{N+1}{2}$	$\frac{(N+1)(N-1)}{12n}$
$\text{median}(\underline{X}) = M$	$\frac{N+1}{2}$	$\frac{(N-1)^2}{4n}$
$\min(\underline{X})$	$\frac{(N-1)}{n} + 1$	$\left(\frac{N-1}{n}\right)^2$
$\max(\underline{X})$	$N - \frac{(N-1)}{n}$	$\left(\frac{N-1}{n}\right)^2$

Using this information, we can calculate the MSE for 4 of the estimators that we have derived. (Remember that MSE = Variance + Bias^2.)

$$\begin{aligned} \text{MSE }(2\cdot\overline{X}-1) &= \frac{4(N+1)(N-1)}{12n} + \left(2\left(\frac{N+1}{2}\right) - 1 - N\right)^2 \\ &= \frac{4(N+1)(N-1)}{12n} \qquad (1) \end{aligned}$$

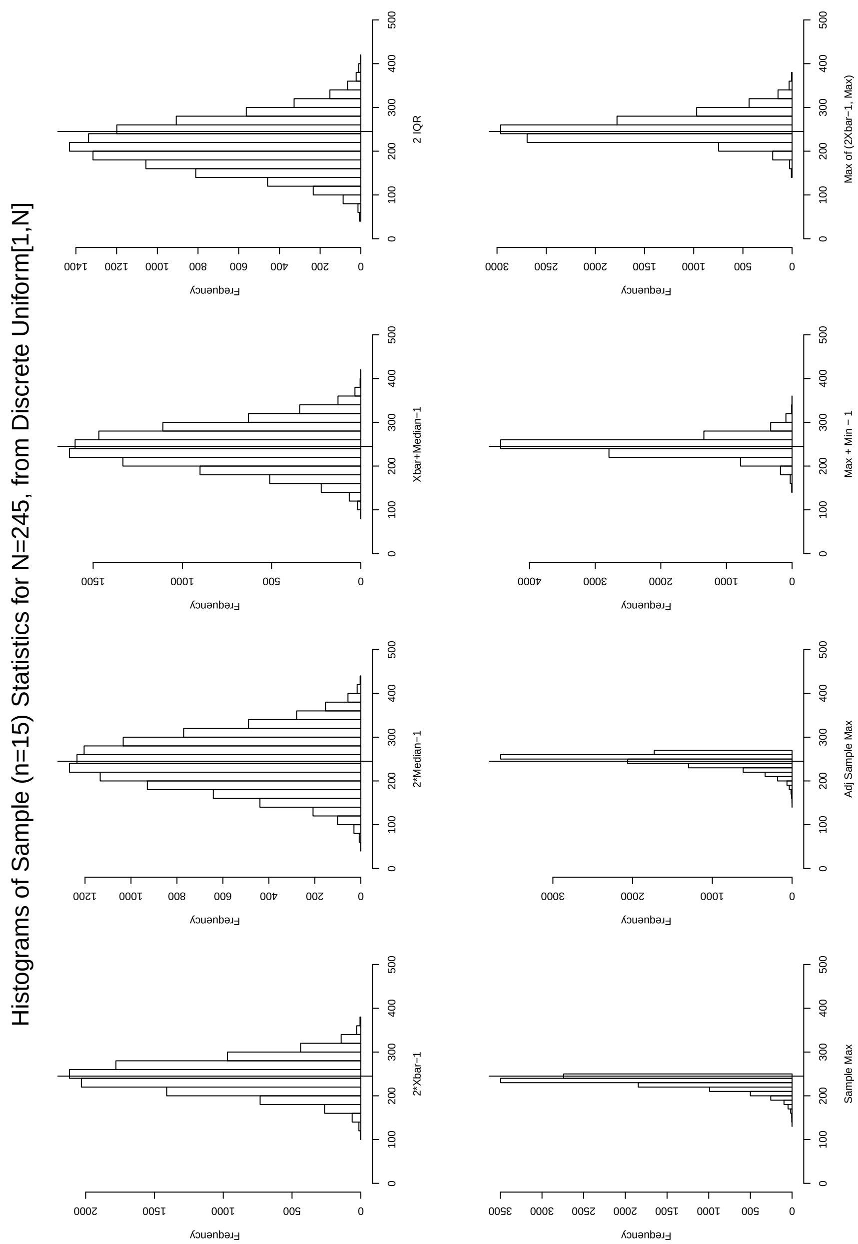
$$\begin{aligned} \text{MSE }(2\cdot M-1) &= \frac{4(N-1)^2}{4n} + \left(2\left(\frac{N+1}{2}\right) - 1 - N\right)^2 \\ &= \frac{4(N-1)^2}{4n} \qquad (2) \end{aligned}$$

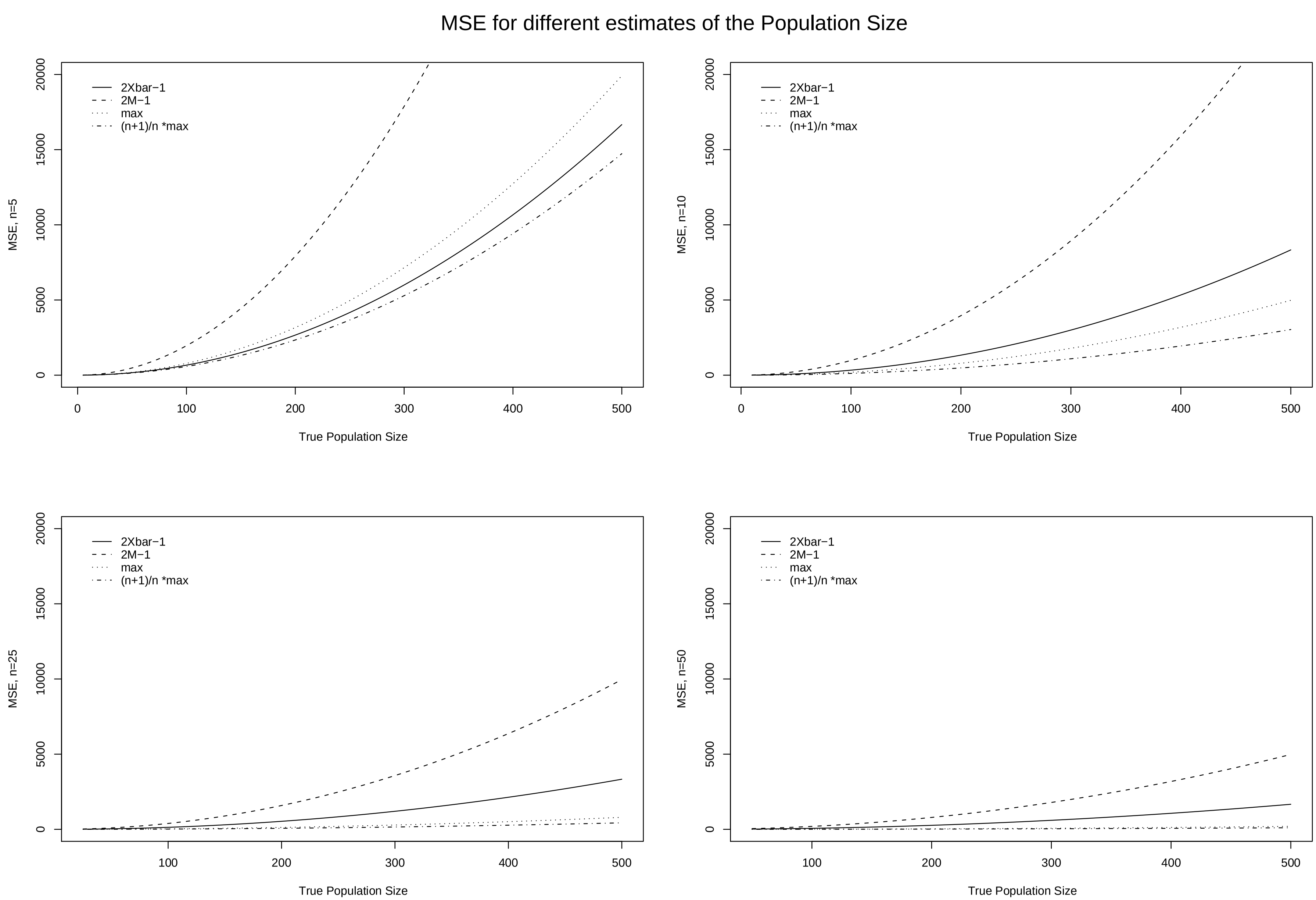
$$\begin{aligned} \text{MSE }(\max(\underline{X})) &= \left(\frac{N-1}{n}\right)^2 + \left(N - \frac{(N-1)}{n} - N\right)^2 \\ &= \left(\frac{N-1}{n}\right)^2 + \left(\frac{N-1}{n}\right)^2 = 2*\left(\frac{N-1}{n}\right)^2 \qquad (3) \end{aligned}$$

$$\text{MSE }\left(\left(\frac{n+1}{n}\right)\max(\underline{X})\right) = \left(\frac{n+1}{n}\right)^2\left(\frac{N-1}{n}\right)^2 + \left(\left(\frac{n+1}{n}\right)\left(N - \frac{N-1}{n}\right) - N\right)^2 \qquad (4)$$

	xbar2	med2	xbarmed	iqr2	max	adjmax	max.min	maxmax
mean	244.6	244.1	244.3	215.5	230.6	247.1	244.9	251.7
median	244.5	245.0	244.0	215.0	235.0	251.8	245.0	245.0
sd	35.8	58.4	45.6	53.5	13.9	14.8	20.2	28.2
min	111.3	53.0	92.1	47.0	139.0	148.9	147.0	148.3
max	378.2	429.0	400.8	401.0	245.0	262.5	359.0	378.2

Table 1: Sample statistics for 10,000 reps taken from a population with $N = 245$ and $n = 15$.





MSE for different estimates of the Population Size
2Xbar−1
2M−1
max
(n+1)/n *max
MSE, n=5
MSE, n=10
MSE, n=25
MSE, n=50
True Population Size
0
5000
10000
15000
20000
100
200
300
400
500