

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:

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

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:

$g(\underline{X})$	$E[g(\underline{X})]$	$\text{Var}(g(\underline{X}))$
\bar{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 $\text{MSE} = \text{Variance} + \text{Bias}^2$.)

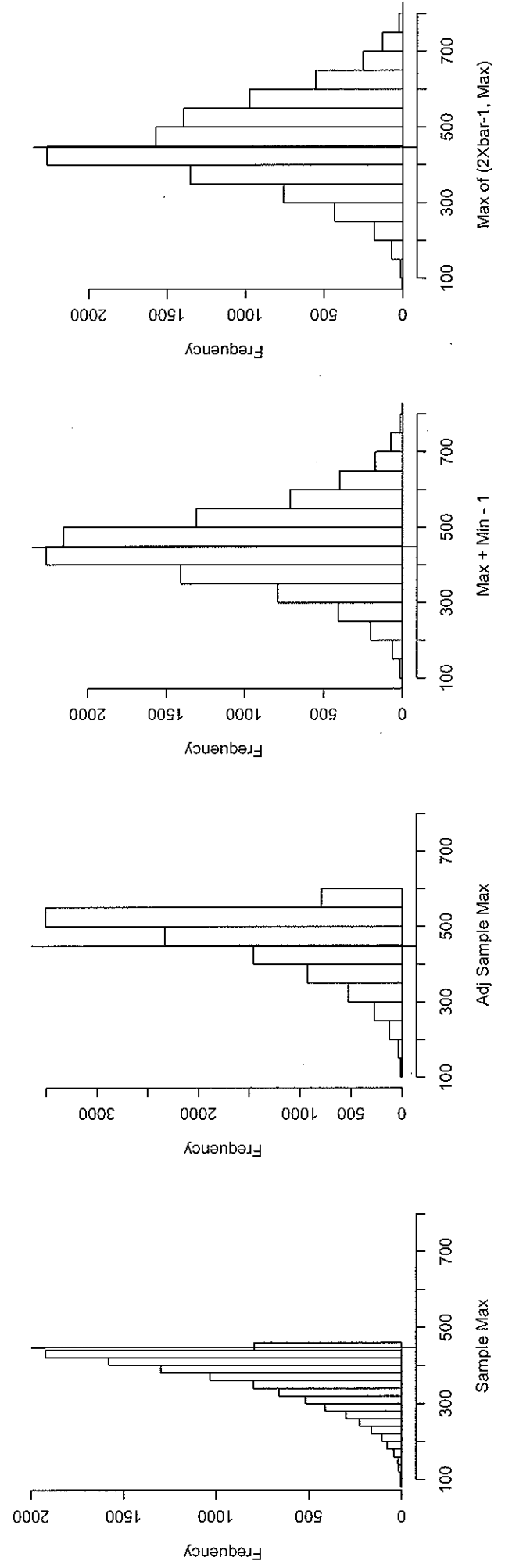
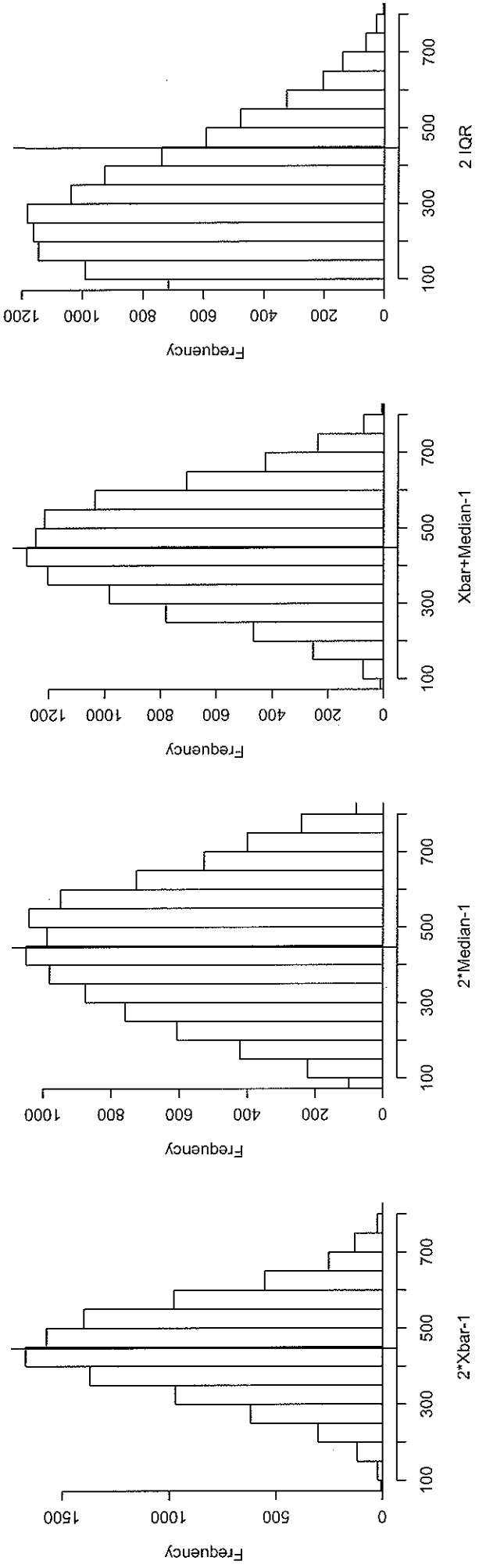
$$\begin{aligned} \text{MSE}(2 \cdot \bar{X} - 1) &= \frac{4(N+1)(N-1)}{12n} + \left(2\left(\frac{N+1}{2}\right) - 1 - N\right)^2 \\ &= \frac{4(N+1)(N-1)}{12n} \end{aligned} \quad (1)$$

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

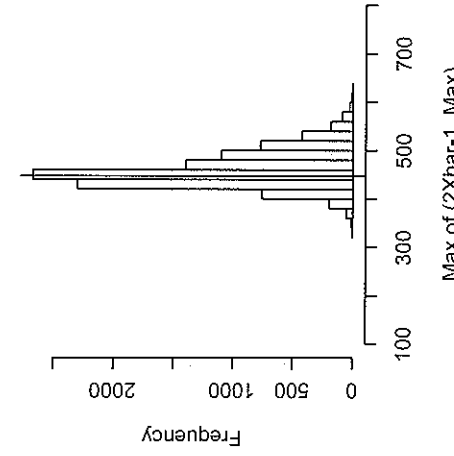
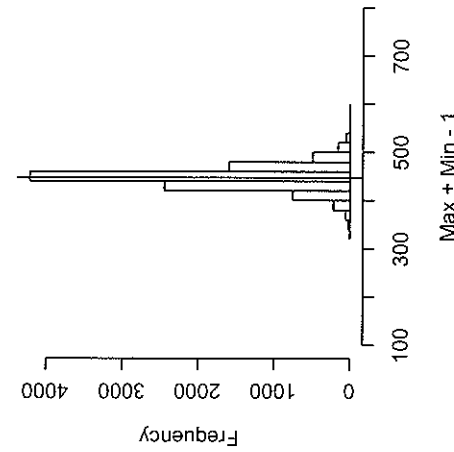
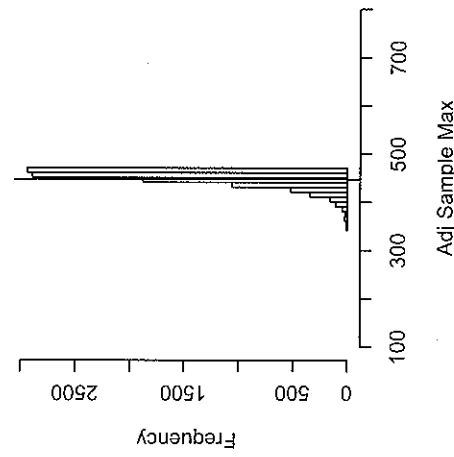
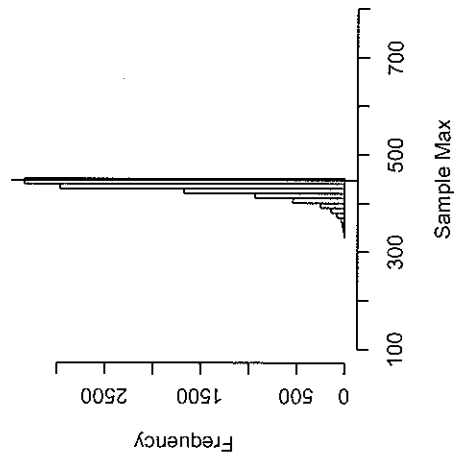
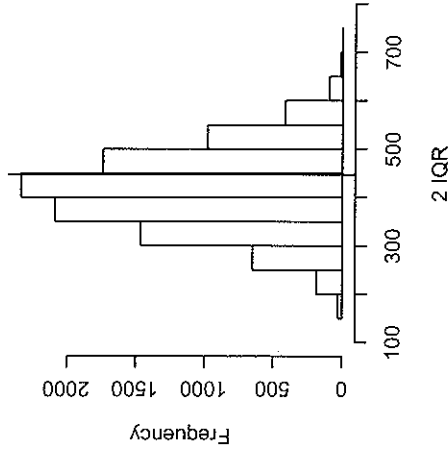
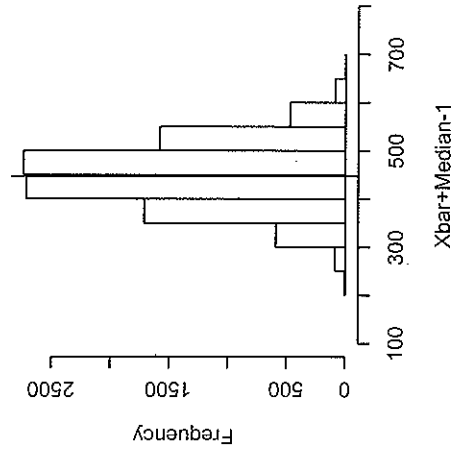
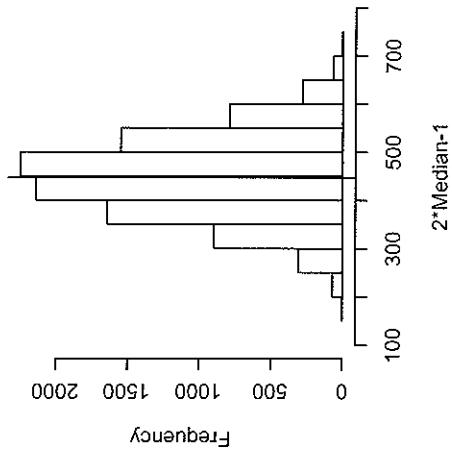
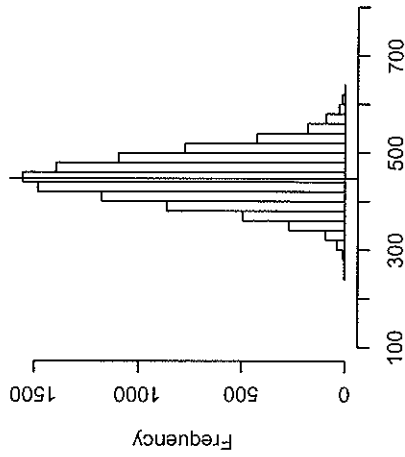
$$\begin{aligned} \text{MSE}(\max(\underline{X})) &= \left(\frac{N-1}{n}\right)^2 + \left(N - \frac{(N-1)}{n} - N\right)^2 \\ &= \left(\frac{N-1}{n}\right)^2 + \left(\frac{N-1}{n}\right)^2 = 2 * \left(\frac{N-1}{n}\right)^2 \end{aligned} \quad (3)$$

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

Histograms of Sample (n=5) Statistics for N=447, from Discrete Uniform[1,N]



Histograms of Sample ($n=25$) Statistics for $N=447$, from Discrete Uniform[1, N]



Math 152- Fall 2008
 Jo Hardlin
 Tank Activity

n=5, N=147

	xbar2	med2	xbarmed	iqr2	max	adjmax	max.min	maxmax
mean	446.1	445.9	447.0	299.3	373.2	466.6	446.3	455.5
median	446.2	447.0	447.2	284.0	390.0	487.5	447.0	446.6
sd	115.1	168.7	137.7	158.6	62.8	78.5	97.3	106.2
min	94.6	25.0	75.4	4.0	71.0	88.8	86.0	99.0
max	808.2	879.0	830.4	822.0	447.0	558.8	793.0	808.2

n=25, N = 447

	xbar2	med2	xbarmed	iqr2	max	adjmax	max.min	maxmax
mean	447.3	447.6	447.4	413.6	430.8	448.7	447.1	459.2
median	447.2	449.0	447.8	412.0	436.0	454.2	447.0	447.2
sd	50.4	83.6	64.9	82.6	16.0	16.6	23.4	37.4
min	256.7	187.0	226.8	150.0	331.0	344.8	331.0	331.0
max	627.0	723.0	662.0	728.0	447.0	465.6	581.0	627.0

MSE for different estimates of the Population Size

