

## Tank Problem

There is a population of  $N$  enemy tanks. Each tank is labeled with an integer  $1 \dots N$ . We "capture" 7 tanks at random and observe their numbers. We want to estimate  $N$  --- the number of enemy tanks.

There are many different estimation procedures. Here are three: the mean of the sample times 3; the median of the sample times 3, the min plus the max.

```
oursample <- sample(tanks, 7)
#produces a sample of 7 tanks from the population (without replacement)

ourestimator <- 3* mean(oursample) # for one possible estimator

# to evaluate the performance of this estimator, we simulate doing this 1000 times:
estimator1 <- function(reps){
  store.estimator1 <- c()
  for (i in 1:reps) {
    samp <- sample(tanks,7)
    tempest <- 3*mean(samp)
    store.estimator1 <- c(store.estimator1, tempest)}
  store.estimator1
}

estimator2 <- function(reps){
  store.estimator2 <- c()
  for (i in 1:reps) {
    samp <- sample(tanks,7)
    tempest <- 3*median(samp)
    store.estimator2 <- c(store.estimator2, tempest)}
  store.estimator2
}

estimator3 <- function(reps){
  store.estimator3 <- c()
  for (i in 1:reps) {
    samp <- sample(tanks,7)
    tempest <- min(samp) + max(samp)
    store.estimator3 <- c(store.estimator3, tempest)}
  store.estimator3
}
*****
```

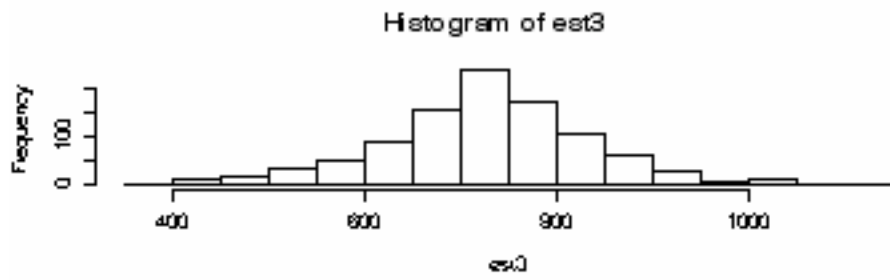
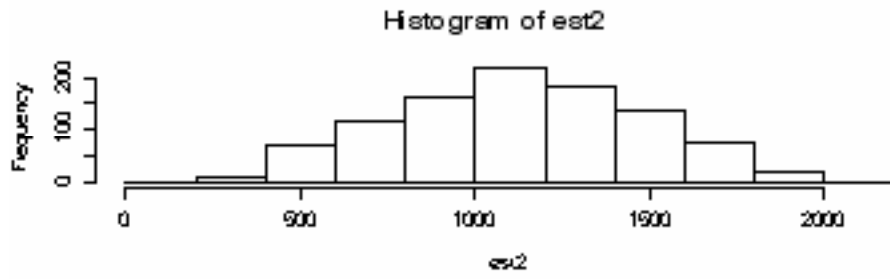
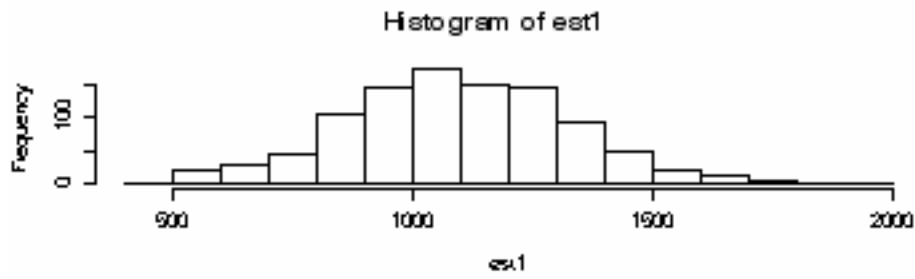
Now we can use these three functions to evaluate the performance of the three estimators.

```
est1 <- estimator1(1000)
```

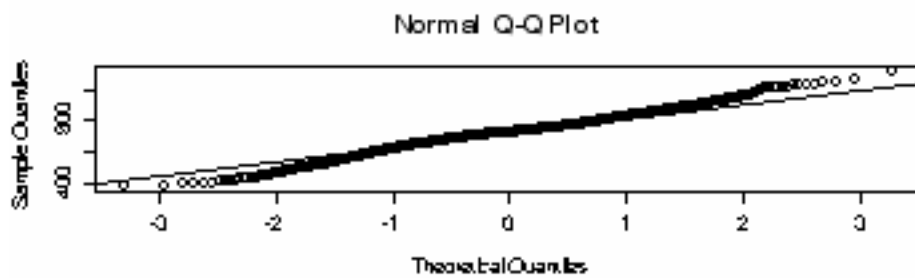
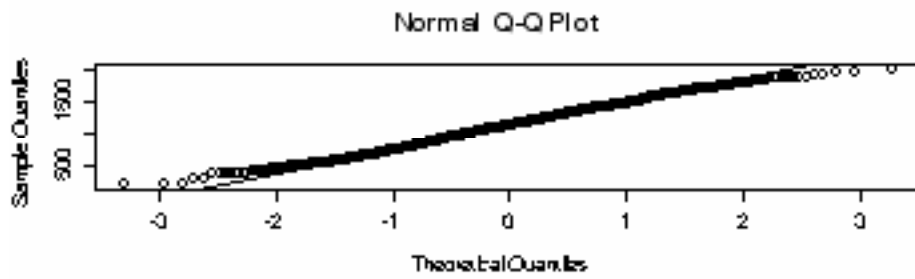
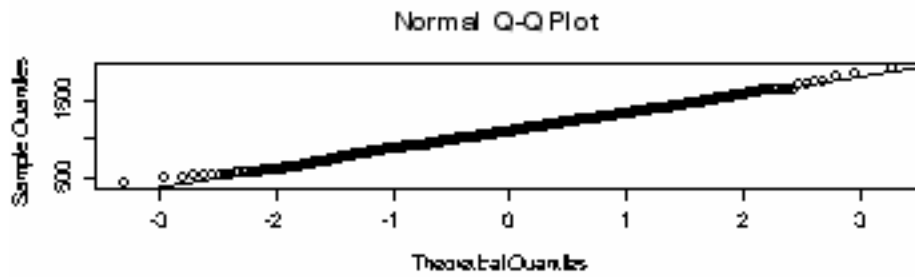
```
est2 <- estimator2(1000)
est3 <- estimator3(1000)
```

Which is best? Which is most accurate? Which most precise? How would you describe their (estimated) sampling distributions? Here are some performance statistics:

```
> summary(est1)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 408.9  930.3 1090.0 1093.0 1256.0 1920.0
> summary(est2)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   174    864   1122   1119   1378   2028
> summary(est3)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   378    662    726    723    785   1115
> sd(est1)
[1] 235.3477
> sd(est2)
[1] 358.1269
> sd(est3)
[1] 112.7787
```



```
par(mfrow=c(3,1))  
hist(est1)  
hist(est2)  
hist(est3)
```



```
qqnorm(est1)  
> qqline(est1)  
> qqnorm(est2)  
> qqline(est2)  
> qqnorm(est3)  
> qqline(est3)
```

And the answer?  $N = 724$