# Homework 2

### Generating our random samples

We create an empty list to store the nth order statistic for samples of size 10, 30, 50, 100, 200.

(100,000 / n) samples are taken for each sample size n. We save the maximum observed value in a sample (nth order statistic) and append it to a vector. Finally, this vector is saved as an element of a list.

This way our list consists of possible MLE's of a uniform (0, 3) distribution.

```
set.seed(303)

samples <- list()

for(n in c(10, 30, 50, 100, 200)) {
  total_no_of_samples <- 100000 / n

  maxes_vector <- c()
  for (i in 1:total_no_of_samples) {
    x_n <- max(runif(n, 0, 3))
    maxes_vector <- c(maxes_vector, x_n)
  }
  samples[[as.character(n)]] <- maxes_vector
}
```

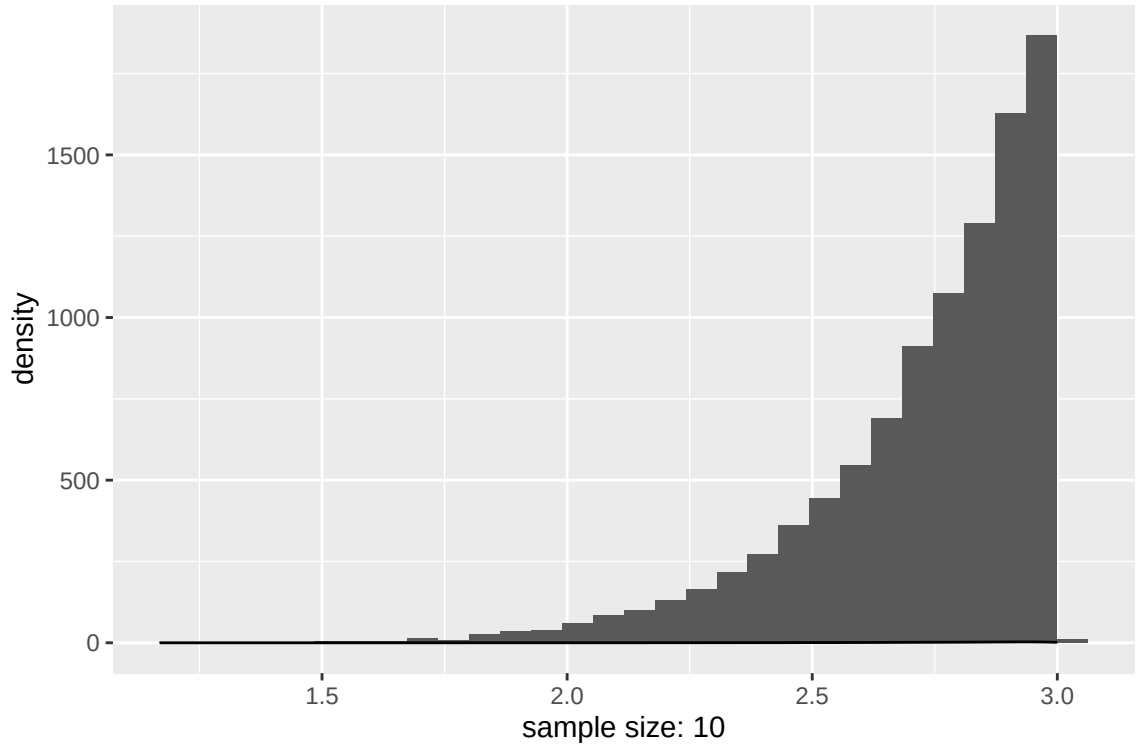
### Visualization:

We will save the sampling distribution for each sample size. (least to greatest) BAN stands for Best Asymptotic Normal. As sample size n increases, we see an improvement in our estimator. Furthermore, mean approaches 3. Variance decreases, and Normality seems to be forming. It seems to be a BAN estimator as we see all across improvements as n increases.

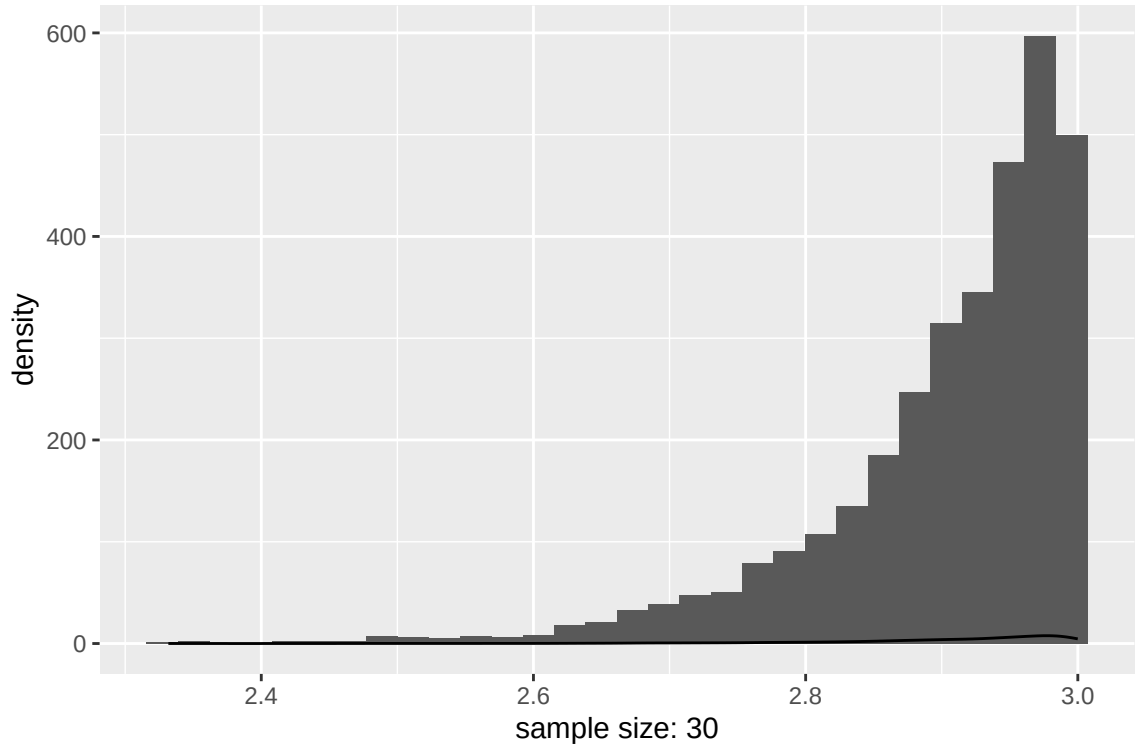
```
library(ggplot2)

pi <- 1 # to set element names
sample_sizes <- c(10, 30, 50, 100, 200)
plots <- list()
for (samples_of_size_n in samples) {
  df <- as.data.frame(samples_of_size_n)
  plots[[as.character(pi)]] <- ggplot(df, aes(samples_of_size_n)) +
    geom_histogram() +
    geom_density() +
    xlab(paste("sample size:", min(sample_sizes)))
  sample_sizes <- sample_sizes[-which(sample_sizes == min(sample_sizes))]
  pi <- pi + 1
}
```

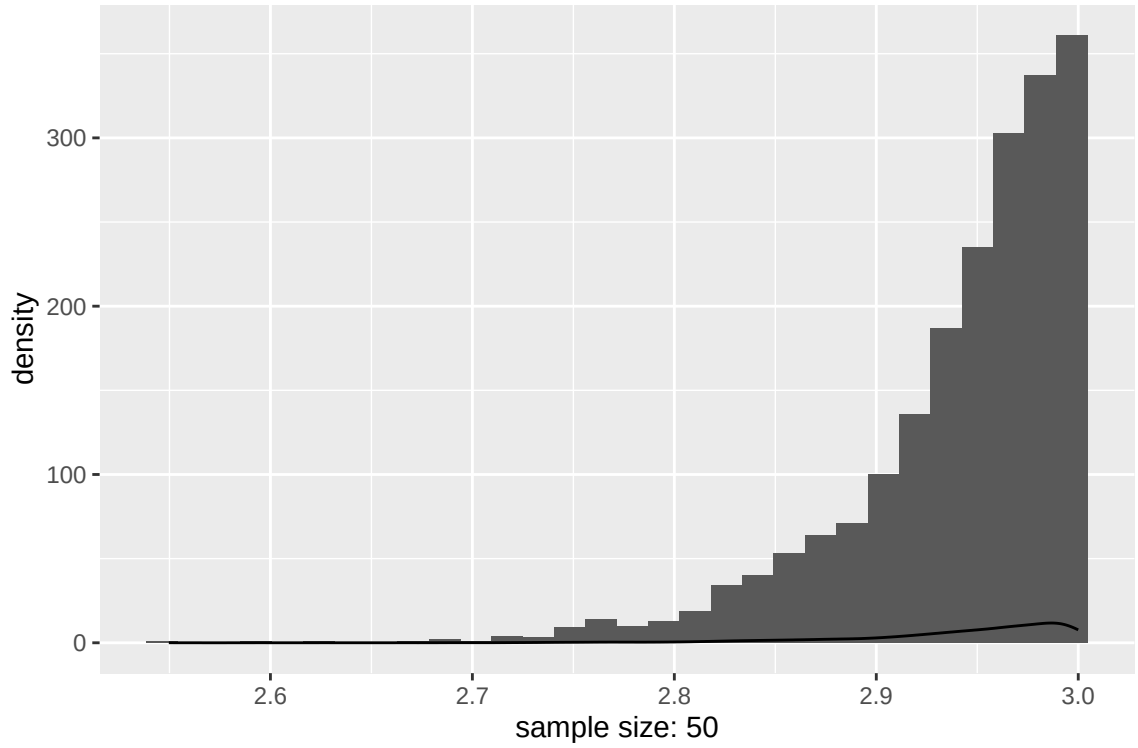
**Sample Size 10:**



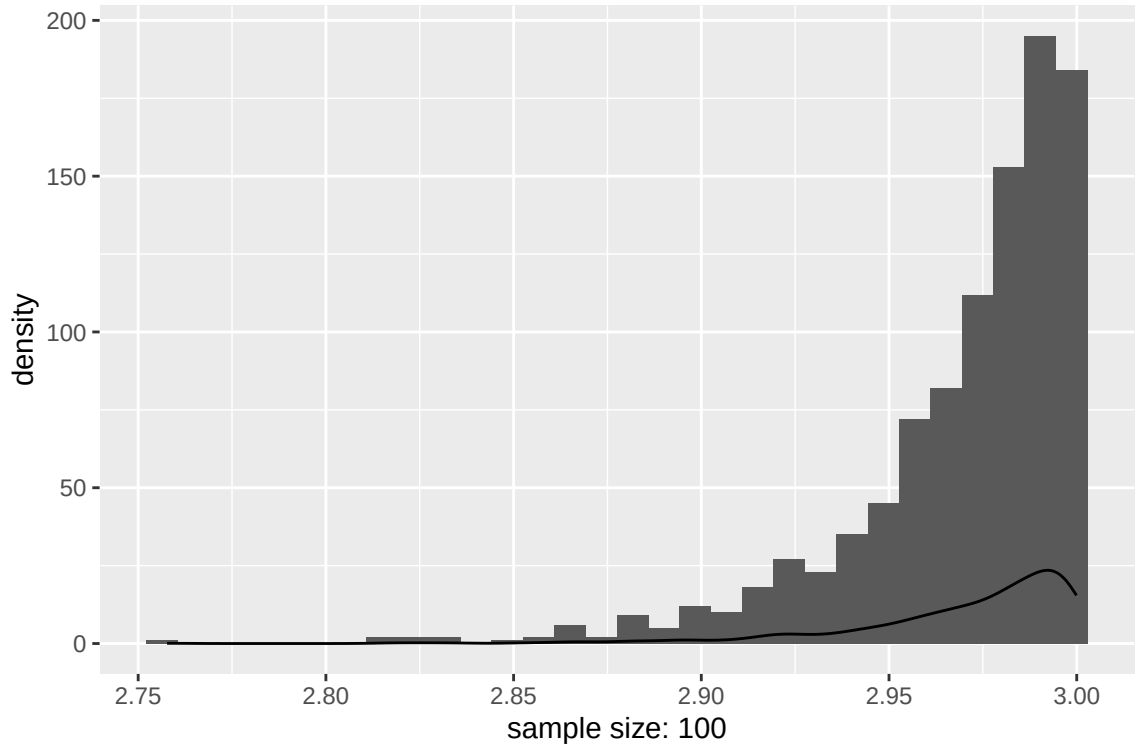

**Sample Size 30:**



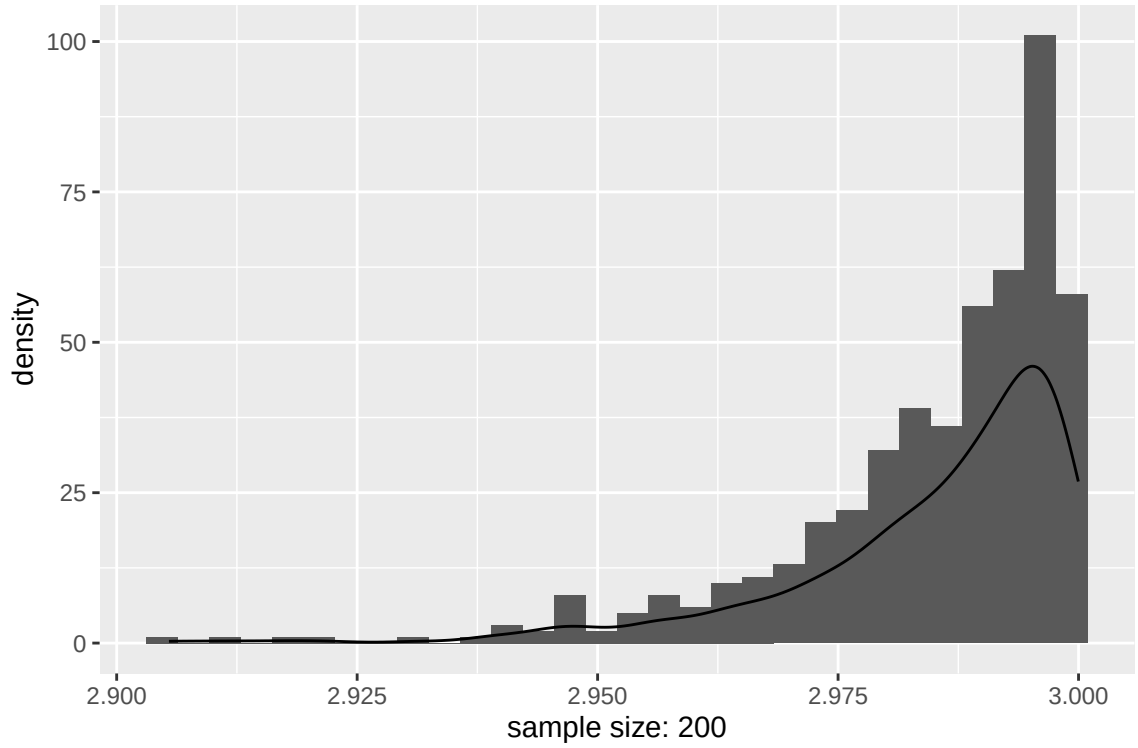
**Sample Size 50:**




**Sample Size 100:**



**Sample Size 200:**




**Approaches normality as sample size increases**