This histogram shows the distribution of 100,000 values is approximately normal.
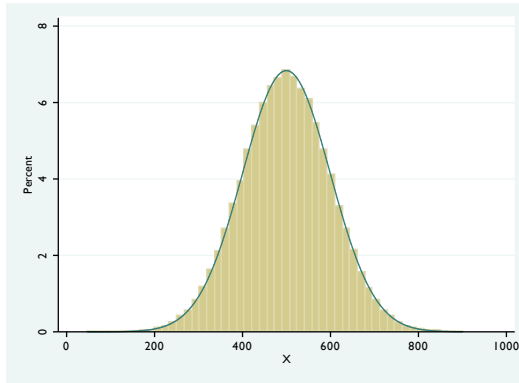
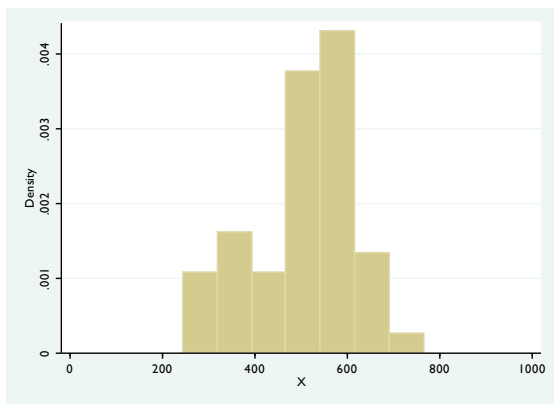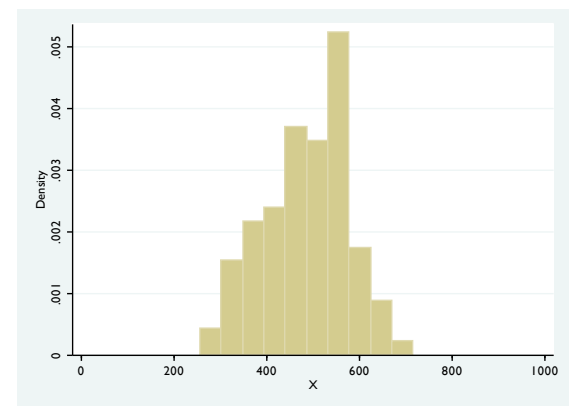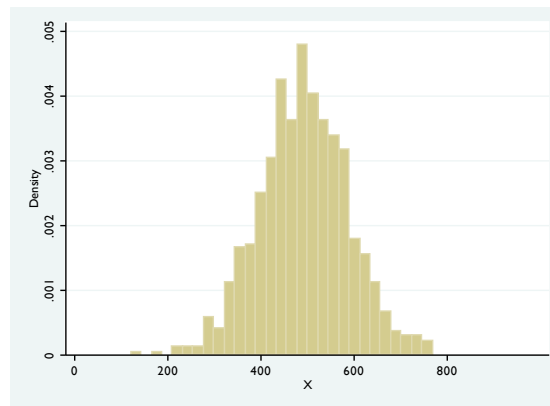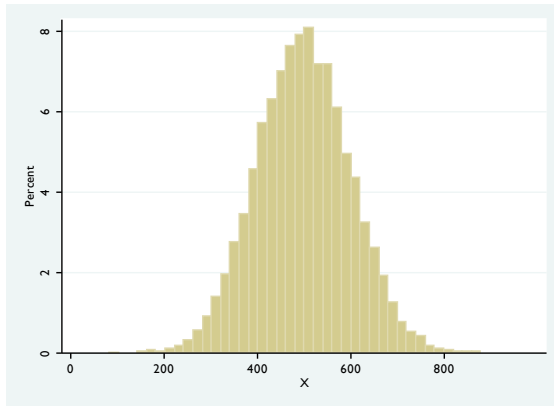Look at the following histograms and determine if you would reasonably assume the values come from a normal distribution:
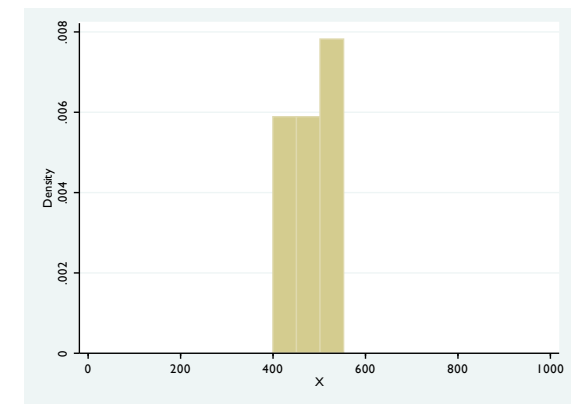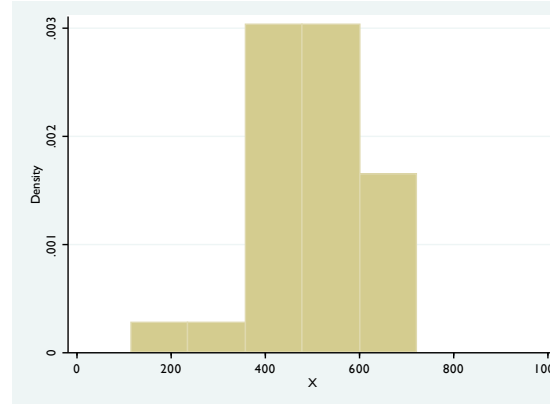
Here's the Stata code that generated these histograms:

- `set obs 100000`
- `generate X = rnormal(500,100)`
- `sample [n], count`
- `histogram X, xscale(range(0 1000)) xlabel(0(200)1000) xmtick(0(200)1000)`

In plain English, here's what that code was doing:

- `Generate 100,000 blank observations that will contain values from a variable named X`
- `Create a variable named X that has a normal distribution with mean=500 and std. dev=100`
- `Take a random sample of n from the 100,000 observations`
- `Create a histogram of these sample values (with the x-axis going from 0 to 1000)`

Each of the histograms displayed on the previous page represents a random sample from the original normal distribution. What does this mean about our ability to look at histograms to assume samples come from a normal distribution?
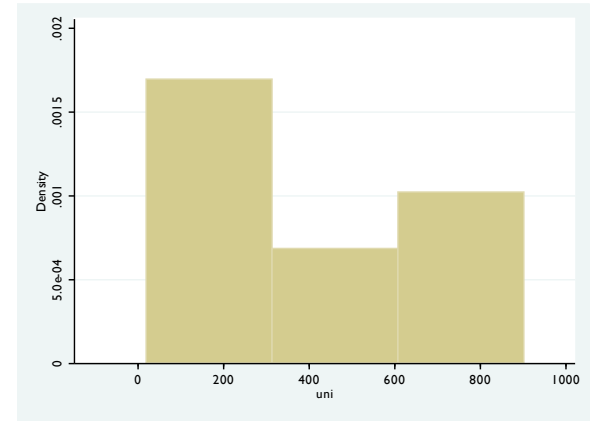
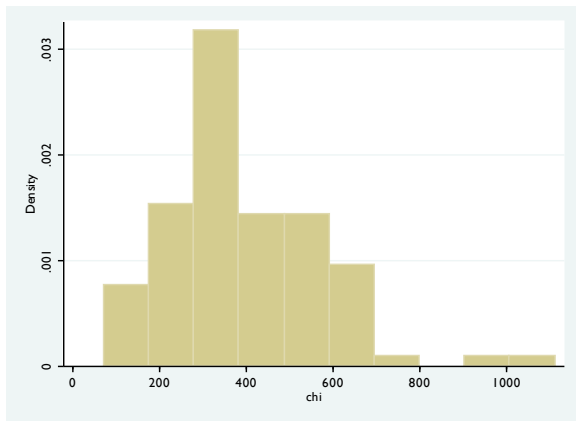Could we assume the following samples come from populations with approximately normal distributions?
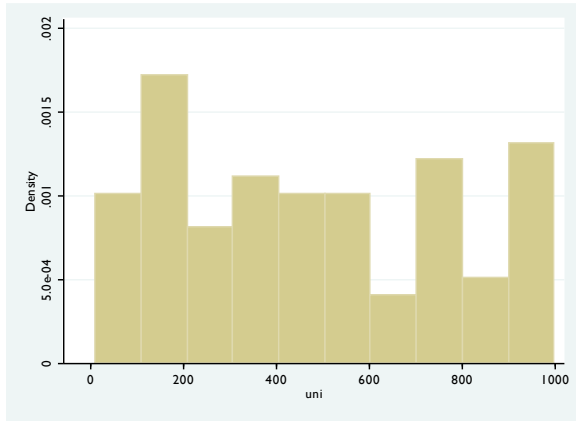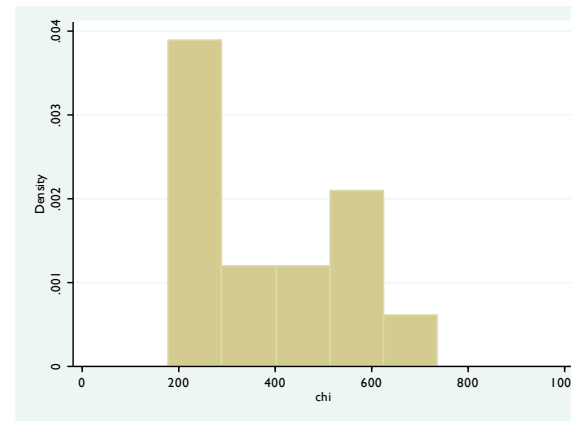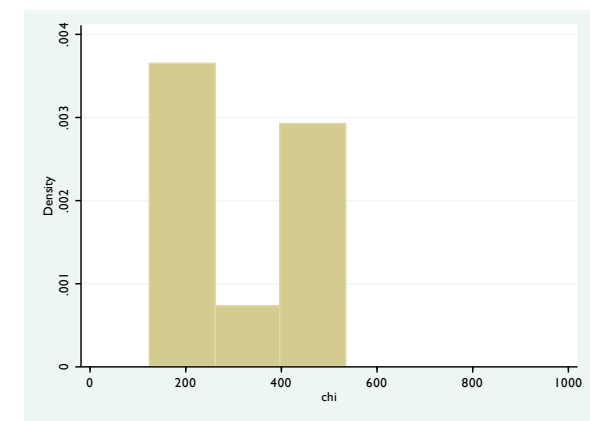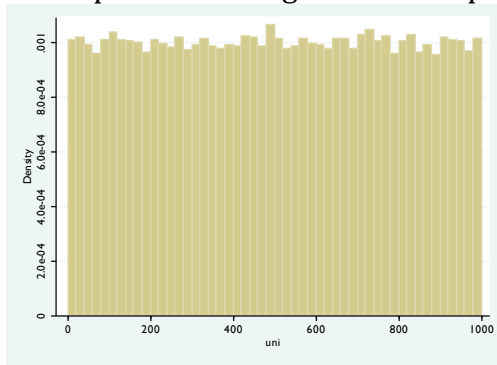

n = 100
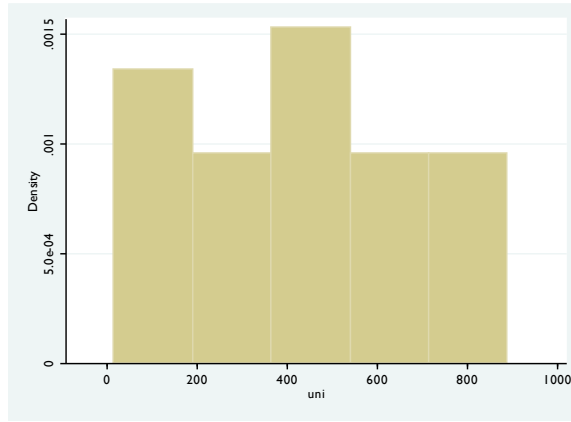
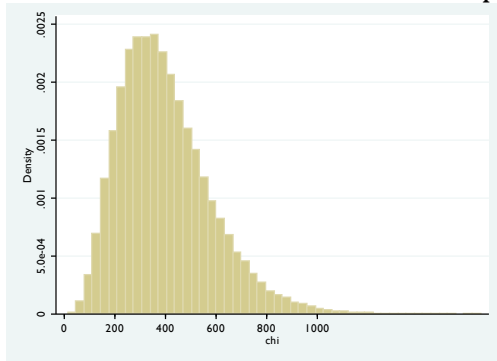n = 30

n = 10

n = 100

n = 30

n = 10

The top row of histograms on the previous page came from a uniform distribution:

The bottom row were random samples from a chi-square distribution with 10 degrees of freedom (each value multiplied by 40)

Look, once again, at our three distributions of n=30. Can we easily determine which came from the normal distribution?