

---

## Lecture 5 In-Class Simulations

---

1. Create a dataset containing the population ( $N=10$ ) on slide 5 in the lecture. Use the `bootstrap` command to simulate 50,000 random samples of size  $n = 8$  (with replacement), retaining the sample mean  $\bar{x}$  on each iteration. Create a histogram showing the resulting sampling distribution. What are the mean and standard deviation of your sampling distribution?
2. Repeat #1 but draw samples of size  $n = 2$ .
3. Create a dataset containing a population ( $N=200$ ) of random draws from the  $N(10, 5)$  distribution. Use a loop with temporary files to simulate 100 random samples of size  $n = 25$  (with replacement), retaining the sample mean  $\bar{x}$  on each iteration.
  - Create a histogram showing the resulting sampling distribution of  $\bar{x}$ .
  - What are the mean and standard deviation of your sampling distribution?
4. Simulation of Application 1-2 (slide 31ff): Rather than draw from a given dataset (like parts 1-3), create a program and use the `simulate` command to draw 1,000 random samples of size  $n = 16$  from the  $N(15, 3)$  population. Calculate the sample mean  $\bar{x}$  on each iteration and retain the results.
  - Create a histogram showing the resulting sampling distribution of  $\bar{x}$ .
  - What are the mean and standard deviation of your sampling distribution?
  - Using your simulation results, what proportion of  $\bar{x}$  values are 13 or less? 14 or less? between 13.5 and 16.5?
5. Repeat #4 but using a loop and the `postfile` commands.
6. Simulation of Application 4 (slide 54ff): Use a loop and the `postfile` commands to draw 1,000 random samples of size  $n = 100$  from the Bernoulli distribution with  $\pi = 0.46$ . Hint: you can use the random number function for binomial, with only 1 trial. Calculate the sample proportion  $\hat{\pi}$  on each iteration and retain the results.
  - In what percent of samples is  $\hat{\pi} \geq 0.50$ ?
7. Repeat #4, but retain the sample variance and an incorrect sample variance that divides by  $n$  rather than  $n - 1$ . Compare the sampling distributions of the “wrong” and “right” sample variances.