COA 620 Introduction to Bayesian Statistics in Ecology

Lab 1

February 5, 2026

Due: February 12, 2026

Instructor: Wei Wu

We will learn how to use simulated data to test the accuracy of the Central Limit Theorem (CLT): When one takes random samples from any distribution with true mean μ and standard deviation σ, the distribution of the sample means will follow a normal distribution with mean μ and standard deviation of $\frac{\sigma}{\sqrt{n}}$.

1. Generate large numbers of random samples from a particular distribution. Here we simulate 10,000 samples of size n=50 from a normal distribution with mean of 1 and standard deviation of 2.
   X <- rnorm(500000, 1, 2)
   X <- matrix(X, nrow=10000, ncol=50)
2. Calculate mean for each sample.
   mean.X <- apply(X, 1, "mean")
3. Calculate mean and standard deviation of the sample means and plot histogram of sample means. Then discuss whether the simulated samples follow CLT.
4. We just test CLT using simulated data from a normal distribution. What if the underlying distribution is not normal distribution? Test CLT using random samples from gamma distribution with shape parameter equal to 2 and scale parameter equal to 3. Check the notes for calculating mean and standard deviation of gamma distribution from its shape and scale parameters.

Bonus points: Find out the minimum sample size for CLT to apply.