Biological Stats 2: Lab 3

Dr. Gavin Fay

2025-02-05

Within your class Project, open a new R markdown file. Save it. (name it lastname_lab3.Rmd or something similar)
At the top of the script, add comments with your name and lab 3.
Work in pairs or individually.
Submit your .Rmd file via myCourses before class on Tuesday next week.

Lab exercise 1/3

a. Generate 1,000 random normal numbers with mean 24 and standard deviation 10. Find the proportion of those random numbers that are $\geq$ 2 standard deviations from the sample mean.
b. Flip a (fair) coin six times.
c. Find the probability of getting six heads on those six flips (i.e. $P(X = 6)$ given $n = 6$).
d. How much more likely is it to get three heads than six?
e. For a standard normal random variable, find the number x such that $P(-x \leq X \leq x) = 0.24$.
f. The mean rate of arrival of alewives at a weir is 3.5 per hour. Plot the probability distribution function for the number of alewife arrivals in an hour.
g. Find the 95% confidence interval for the number of alewives arriving per day.

Lab exercise 2/3

1. Extract the residuals from the RIKZ_lm2 model
2. Are the linear regression assumptions met? Explain your reasoning
3. Summarize the results of the model. What are the parameter estimates telling you about species richness on these beaches?

Lab exercise 3/3

1. Fit a linear regression using 2007 gapminder data of the form `lm(gdpPercap ~ continent`, where gdpPercap is the new outcome variable `y`. Get information about the best-fitting line from the regression table. How do the regression results match up with those of an analysis of life expectancy by continent?
2. Extract the model coefficients and their 95 percent confidence intervals.
3. Plot the residuals vs the fitted values and comment on their distribution and patterns.
4. Identify the five countries with the five most negative residuals? What do these negative residuals say about their life expectancy relative to their continents life expectancy?
5. Repeat this process, but identify the five countries with the five most positive residuals. What do these positive residuals say about their life expectancy relative to their continents life expectancy?