Problem set_ M5 Probability and Hypothesis Testing

Pablo E. Gutiérrez-Fonseca

2023-08-12

Welcome to our first BIG problem set where we begin to explore the concepts and methods of normality, probability and z scores. You will need to download both the instruction document (which you are looking at right now) and the Sleeping Mammals.xls spreadsheet. Take your time and know that this should probably take 2-3 hrs to complete. Once again, I encourage you to work in groups to discuss these questions. However, you must ultimately do the work on your own and turn in your answers and interpretations independently. HAVE FUN!

Consider the data on sleep patterns in various mammalian species in the spreadsheet **Sleeping Mammals.csv**. This dataset contains information on sleep patterns for comparison to factors such as species body weight, brain size, life span, gestation period, and three class measures of risk: predation, danger, and exposure. In this problem set, we will be practicing our transformations and probabilities associated with standard normal distributions.

We are interested in comparing sleep time spent in dreaming and non-dreaming phases for mammals of different sizes. **We hypothesis that large mammals exhibit more time spent in dreaming sleep than small mammals**. Before we can test this hypothesis, we need to examine the distribution of our data.

- Import your data table into R.
- Analyze the distribution of both the Dreaming and Non-Dreaming variables (Analyze > Distribution). We'll start by just examining some of the descriptive statistics and what they tell us about this data.

1. Based on the summary results in JMP, mammals spend most of their time in which type of sleep phase (dreaming or non-dreaming)?

2. Which type of sleep phase has the highest variability across the species included here (dreaming or non-dreaming)?

Now we will be testing distributions and various transformations. We will examine one variable at a time. We will start with the **Dreaming variable**:
- Fit a normal distribution and test for Normality using the Goodness of Fit test.

3. Enter the p-value for the goodness of fit test for the **Dreaming variable** (NOTE. . . .just enter the number (no letters, symbols, equal signs etc). . . also be careful of decimal places.

4. Based on this goodness of fit test, is the dreaming variable normally distributed?

Let's see if any outliers are skewing our distribution.

5. Enter the new p-value for the goodness of fit test for the **Dreaming variable** when the outliers are excluded (NOTE. . . .just enter the number (no letters, symbols, equal signs etc). . . also be careful of decimal places.