Box 1: Key Concepts in Frequentist Statistics

- A *sampling distribution* is the distribution of sample statistics computed using different samples of the same size from the same population.
- A *bootstrap distribution* is a distribution of statistics computed using different samples of the same size from the same *estimated* population formed by merging many copies of the original sample data. Alternatively, the sample data may be used to estimate parameters of a statistical distribution, and then this distribution can be used to generate new samples. This alternative is termed the *parametric bootstrap*.
- A *null* or *randomization distribution* is a collection of statistics from samples simulated assuming the null hypothesis is true.
- The *standard error* of a statistic is the standard deviation of the sampling distribution. When forming confidence intervals, we estimate the standard error using the standard deviation of a bootstrap distribution. When calculating p-values, we estimate the standard error using the standard deviation of the randomization distribution.
- *2 SE rule*: when statistics have bell-shaped (i.e., approximately Normal) sampling distributions, we expect roughly 95% of sample statistics to be within 2 standard deviations of the mean of the sampling distribution
- A *confidence interval* for a parameter is an interval computed from data using a method that will capture the parameter for a specified proportion of all samples (e.g., 95% of the time for a 95% confidence interval)
- The *p-value* is the chance of obtaining a sample statistic as extreme (or more extreme than) the observed sample statistic, if the null hypothesis is true.