

## Unit 3 Homework Problems

### **Homework Problem:** Analyzing the Effectiveness of a New Training Method

Smith Manufacturing Company is a small manufacturing firm in the midwestern US. The company makes customized machined parts for use in machinery and aircraft owned by the US military. The company is interested in determining whether a new employee training method improves employee performance as rated by their employee reviews. They randomly select 50 employees from one of their Assembly & Fitting areas and randomly assign 25 employees to use the new training method (treatment group) and 25 employees to use the traditional training method (control group). After the training course is concluded and the employee receives a review, the review scores are recorded for each employee.

The employee review scores for the treatment group are: 85, 92, 78, 88, 95, 83, 90, 87, 91, 84, 82, 89, 93, 79, 81, 86, 94, 77, 80, 96, 85, 92, 88, 90, 83. The employee review scores for the control group are: 80, 85, 76, 82, 90, 78, 87, 83, 89, 79, 77, 84, 88, 75, 81, 86, 91, 74, 79, 93, 80, 87, 82, 85, 78.

In order to complete this assignment in R, you'll need to run the following code first:

```
install.packages("moments")
install.packages("tidyverse")
install.packages("effectsize")

library(tidyverse)
library(moments)
library(effectsize)
```

Remember, if you have already installed moments, tidyverse, or effectsize you DO NOT need to run the "install.packages()" commands again. In such a case, just run the library commands.

### **Questions**

1. Import the data above into R and create two vectors, `treatment\_scores` and `control\_scores`, containing the exam scores for the treatment and control groups, respectively. Be prepared to answer questions regarding these vectors.
2. Calculate the mean, median, range, and standard deviation of exam scores for both the treatment and control groups.
3. Perform a two-sample t-test to determine if there is a significant difference in exam scores between the treatment and control groups. Use a significance level of 0.05. Be prepared to interpret the results.
4. Determine the effect size (Cohen's d) of the difference in exam scores between the treatment and control groups. Be prepared to interpret the results.
5. Conduct a power analysis for a two-sample t-test using the following parameters:
  - a. Significance level: 0.05
  - b. Power: 0.8

- c. Effect size (Cohen's d): Use the effect size calculated in question 4
- d. Sample size: 25 per group

Be prepared to interpret the results. Is the current sample size sufficient to detect the observed effect size with a power of 0.8?

6. Create a boxplot in R to visualize the distribution of exam scores for the treatment and control groups. Describe any notable observations from the plot.
7. Calculate the 95% confidence interval for the difference in mean exam scores between the treatment and control groups.
8. Perform a Welch's t-test, which does not assume equal variances between the two groups, to compare the exam scores of the treatment and control groups. Use a significance level of 0.05. Be prepared to interpret the results. How do these results compare to the two-sample t-test performed in question 2?
9. Create a new variable called ``study_method`` that indicates whether a student belongs to the treatment or control group. Combine the exam scores from both groups into a single vector called ``exam_scores``. Create a data frame called ``study_data`` with the variables ``study_method`` and ``exam_scores``. Provide the R code and display the first few rows of the data frame.
10. Using the ``study_data`` data frame from question 9, create a scatter plot with ``study_method`` on the x-axis and ``exam_scores`` on the y-axis. Add a title, x-label, and y-label to the plot. Provide the R code and describe any differences you observe in the distribution of exam scores between the two study methods.

Once you have completed these calculations, you are ready to begin the Unit 3 Homework Quiz.