MATH 5900 – ADVANCED DATA ANALYTICS

ASSIGNMENT 6

TAIWO JEGEDE – E00755963

APPLIED DATA SCIENCE

SOFTWARE USED: R

QUESTION 1

I

Power Function

The power function is a function that gives us the probability that a statistical test will reject the null hypothesis when the alternative hypothesis is true. This probability is known as the statistical power of the test.

Power Function: $\pi(\theta) = P(\text{Reject Ho} \mid \Theta = \theta)$

What it means to have a range of power values according to this function

When we talk about a range of power values according to this function, we're referring to the different power values that the test can have under different conditions. For example, if we vary the sample size or the effect size, the power of the test will change. This range of power values can give us important information about how sensitive our test is to changes in these parameters. For instance, it can tell us how much we can expect to increase our power by increasing our sample size.

II

Power under a specific alternative

When we refer to the “power under a specific alternative”, we are talking about the statistical power of a test under a specific alternative hypothesis. In hypothesis testing, we start with a null hypothesis (often denoted as Ho), that represents a statement of no effect or no difference. The alternative hypothesis (often denoted as Ha) is what we might believe to be true if we have sufficient evidence to reject the null hypothesis.

The power of a test is the probability that it correctly rejects the null hypothesis when the alternative hypothesis is true. So, when we talk about the “power under a specific alternative”, we’re talking about the probability of correctly rejecting the null hypothesis given that this specific alternative hypothesis is true.

III

Power is a statistical concept that refers to the ability of a test to correctly reject the null hypothesis when the alternative hypothesis is true. In a situation where by we are carrying out a study, we always consider the probability that if there is a true effect or difference, our study will be able to detect it.

IV

Aside from sample size, there are several other factors that affect the power of a statistical test and some of them are

1. Effect Size: This is the magnitude of the difference between the null hypothesis and the true value. A larger effect size means a larger difference to detect, which increases the power of the test.

2. Decision Rate or Significance Level: This is the probability of rejecting the null hypothesis when it is true (Type I error). A lower significance level makes the test more stringent, which decreases the power because it's harder to reject the null hypothesis.

V

Post-hoc power analysis, is also sometimes called observed power because it is a method that estimates the power of a statistical test after the data has been collected and analyzed, So it is based on the observed effect size, the actual sample size, and the significance level used in the study.

Why it is a misleading calculation.

Post-hoc power is directly related to the p-value obtained in the study. If the result is statistically significant, post-hoc power will be high; if the result is not significant, post-hoc power will be low. Therefore, it does not provide any additional information beyond what the p-value has already told us.

QUESTION 2

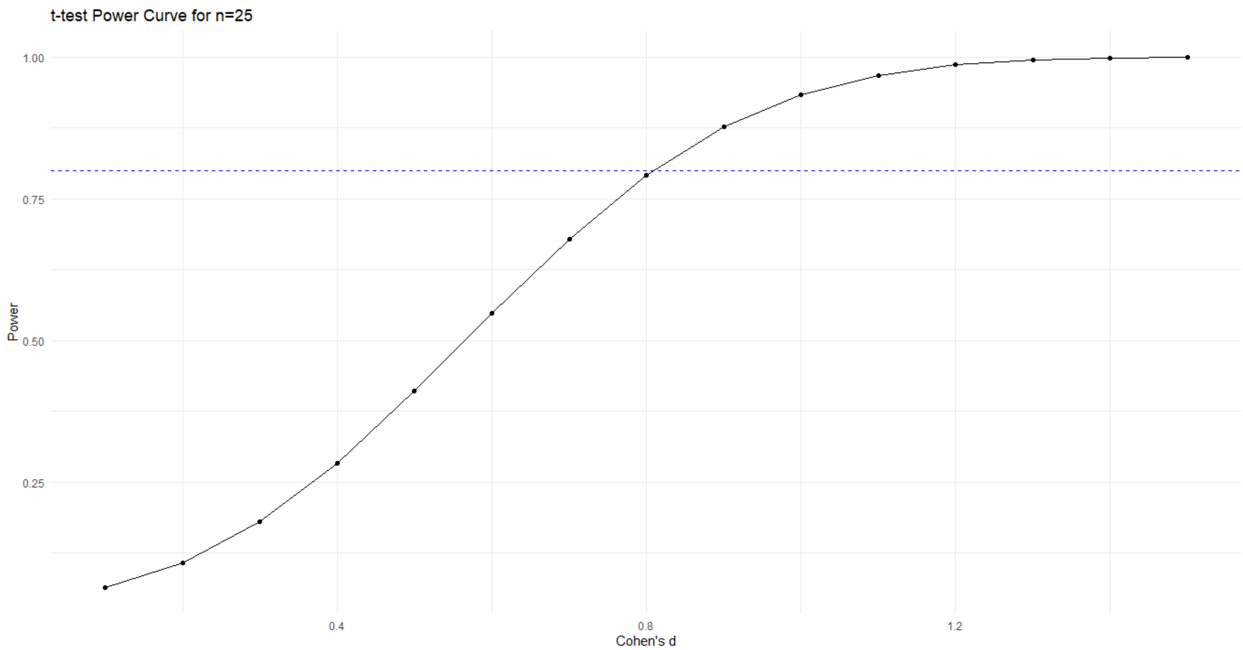
I

An appropriate test we could use is a t-test, specifically a 2-sample t-test. So we use the 2-sample t-test when we have two measurements like this one, the self-efficacy scores before and after the intervention, on the same item or individual and we want to see if there is a significant difference between them.

In this case, the null hypothesis (Ho) would be that there is no difference in the mean self-efficacy score before and after the intervention. The alternative hypothesis (Ha) would be that there is a difference in the mean self-efficacy score before and after the intervention.

The test will compare the means of the two sets of scores and determine whether any observed difference is statistically significant. If the p-value obtained from the test is less than the chosen significance level (often 0.05), then the null hypothesis is rejected and it can be concluded that there is a significant difference in self-efficacy before and after the intervention.

II



In this plot, the power curve for a sample size of 25 shows that the test has a power of 0.8 for a difference of 0.8. As the difference approaches 0, the power of the test decreases and approaches α (also called the significance level), which is 0.05 for this analysis.

III

To estimate the power of a statistical test, in addition to the testing procedure, we would need the effect size, the significant level and the sample size.

The effect size is the magnitude of the different that we would expect to see between the pre-intervention and post-intervention self-efficacy scores.

The significance level, which is denoted by α, is the probability of rejecting the null hypothesis when it is true, often set at 0.05. It represents the risk you're willing to take of making a Type I error (false positive).

The Sample Size, which is the number of students (or pairs of pre- and post-intervention measurements) in the study. More participants increase the power of the test.

IV

After running my R code, the estimated power of the study is approximately 0.46.

This means that if there is indeed a true difference in self-efficacy scores before and after the intervention, your study has about a 46% chance of correctly detecting this difference at a significance level of 0.05.

It's worth noting that a power of 0.8 is often considered the standard target in research, indicating an 80% chance of detecting a true effect. A power of 0.46 is relatively low, which suggests that the study might not be sufficiently powered to detect the expected difference in self-efficacy scores.

V

Sample Size: If the sample size increases while all other factors remain the same, the power of the test will increase. This is because with more data, we have a better chance of detecting a true effect if one exists.

Difference Between Means (Effect Size): If the difference between means (or the effect size) increases while all other factors remain the same, the power of the test will also increase. A larger effect size means a larger difference to detect, which makes it easier to reject the null hypothesis if it is false.

Response Variation (Standard Deviation): If the response variation or standard deviation increases while all other factors remain the same, the power of the test will decrease. Greater variability in the data makes it harder to detect a true effect, reducing the power. Conversely, if the standard deviation decreases, the power of the test will increase.

Significance Level (α): If the significance level (α) increases while all other factors remain the same, the power of the test will increase. A higher significance level makes the test less stringent, making it easier to reject the null hypothesis. However, it’s important to note that increasing the significance level also increases the probability of a Type I error (false positive), so this needs to be considered carefully.

QUESTION 3

I

The steps involved when we want to perform a power/sample size analysis are as follows:

Define the Effect Size: This is the minimum difference you want to be able to detect, which could be based on previous studies, theoretical considerations, or practical significance.

Choose the Significance Level (α): This is the probability of rejecting the null hypothesis when it is true, often set at 0.05 or 0.01.

Specify the Power: This is the probability of correctly rejecting the null hypothesis when the alternative hypothesis is true. A common target is 0.80 or 0.90.

Estimate the Variability: This could be based on previous studies, pilot data, or other relevant information.

Calculate the Required Sample Size: Use a statistical formula or software to calculate the sample size needed to achieve the desired power

II

Using α = 0.05, provide minimum values of sample size associated with power levels

of 0.70, 0.75, 0.80, 0.85, 0.90, and 0.95.

Output

```
> # Define the power levels
> power_levels <- c(0.70, 0.75, 0.80, 0.85, 0.90, 0.95)
> # Calculate the required sample sizes for alpha = 0.05
> for (power in power_levels) {
+   n <- pwr.t.test(d = d, power = power, sig.level = 0.05, type = "paired")$n
+   print(paste("For power =", power, "and alpha = 0.05, the required sample size is", round
(n)))
+ }
[1] "For power = 0.7 and alpha = 0.05, the required sample size is 57"
[1] "For power = 0.75 and alpha = 0.05, the required sample size is 64"
[1] "For power = 0.8 and alpha = 0.05, the required sample size is 73"
[1] "For power = 0.85 and alpha = 0.05, the required sample size is 83"
[1] "For power = 0.9 and alpha = 0.05, the required sample size is 97"
[1] "For power = 0.95 and alpha = 0.05, the required sample size is 119"
```

III

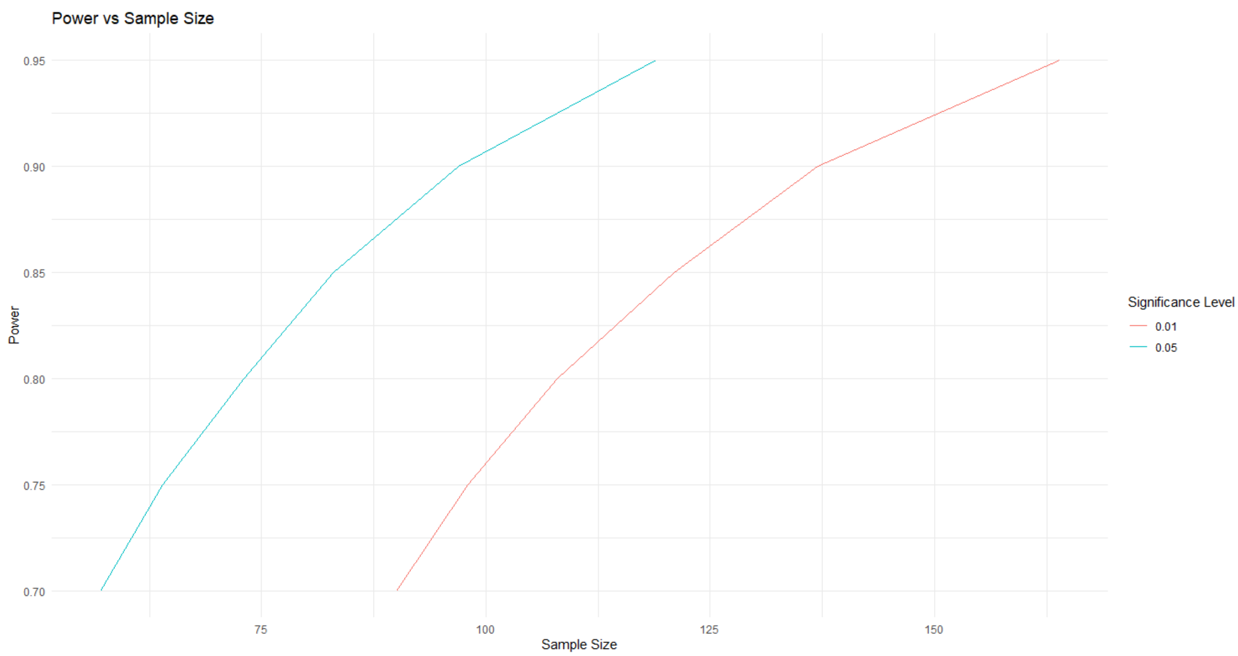
Using α = 0.01, provide minimum values of sample size associated with power levels

of 0.70, 0.75, 0.80, 0.85, 0.90, and 0.95.

Output:

```
> # Calculate the required sample sizes for alpha = 0.01
> for (power in power_levels) {
+   n <- pwr.t.test(d = d, power = power, sig.level = 0.01, type = "paired")$n
+   print(paste("For power =", power, "and alpha = 0.01, the required sample size is", round
(n)))
+ }
[1] "For power = 0.7 and alpha = 0.01, the required sample size is 90"
[1] "For power = 0.75 and alpha = 0.01, the required sample size is 98"
[1] "For power = 0.8 and alpha = 0.01, the required sample size is 108"
[1] "For power = 0.85 and alpha = 0.01, the required sample size is 121"
[1] "For power = 0.9 and alpha = 0.01, the required sample size is 137"
[1] "For power = 0.95 and alpha = 0.01, the required sample size is 164"
```

IV



V

Based on the power analysis results conducted for two different significance levels (α = 0.05 and α = 0.01), we can observe that as the desired power level increases, the required sample size also increases.

For a significance level of α = 0.05, the required sample sizes range from 57 to 119, while for α = 0.01, the required sample sizes range from 90 to 164.

In recommending a reasonable sample size for this study, a larger sample size might be desirable for increased precision, but it could also be more costly and time-consuming. Also, the effect size plays an important role, A larger effect size generally requires a smaller sample size to achieve the same power level.

The chosen significance level (α) determines the risk of Type I errors (false positives). A lower α level (e.g., 0.01) reduces the likelihood of falsely rejecting the null hypothesis, but it may necessitate larger sample sizes to maintain power.

I think aiming for a higher power level (e.g., 0.90 or higher) with a moderately sized sample could provide sufficient sensitivity to detect meaningful effects while controlling for Type I error rate.

CODES

```
library('pwr')

library('ggplot2')

library(pwr)

library(ggplot2)


power.curve <- function(n){

  cd <- seq(.1,1.5,.1)  #Vector of effect size

  samp.out <- NULL

  for(i in 1:length(cd)){

    power <-  pwr.t.test(d=cd[i],n=n,sig.level=.05,type="two.sample")$power

    power <-  data.frame(effect.size=cd[i],power=power)

    samp.out <- rbind(samp.out,power)
```

```
  }
  ggplot(samp.out, aes(effect.size,power))+
    geom_line() +
    geom_point() +
    theme_minimal() +
    geom_hline(yintercept = .8,lty=2, color='blue') +
    labs(title=paste0("t-test Power Curve for n=", n), x="Cohen's d", y="Power")
}


n <- 25
power.curve(n)
n <- 23  # sample size
mu1 <- 12.3  # pre-intervention mean
mu2 <- 14.8  # post-intervention mean
sd <- 7.5  # standard deviation
sig.level <- 0.05
# Calculate the effect size using Cohen's d
d <- (mu2 - mu1) / sd
# Estimate the power
```

```
power = pwr.t.test(n=n,d=d,sig.level= sig.level,type="two.sample",alternative="two.sided")

power <- pwr.t.test(n = n, d = d, sig.level = sig.level, type = "paired", alternative =
"greater")$power

# Print the estimated power

print(power)

# Set your parameters

mu1 <- 12.3  # pre-intervention mean

mu2 <- 14.8  # post-intervention mean

sd <- 7.5  # standard deviation


# Calculate the effect size using Cohen's d

d <- (mu2 - mu1) / sd

# Define the power levels

power_levels <- c(0.70, 0.75, 0.80, 0.85, 0.90, 0.95)

# Calculate the required sample sizes for alpha = 0.05

for (power in power_levels) {

  n <- pwr.t.test(d = d, power = power, sig.level = 0.05, type = "paired")$n

  print(paste("For power =", power, "and alpha = 0.05, the required sample size is", round(n)))

}
```

```r
# Calculate the required sample sizes for alpha = 0.01

for (power in power_levels) {

  n <- pwr.t.test(d = d, power = power, sig.level = 0.01, type = "paired")$n

  print(paste("For power =", power, "and alpha = 0.01, the required sample size is", round(n)))

}

# Install and load the ggplot2 package for creating plots

install.packages("ggplot2")

library(ggplot2)

# Define the power levels and corresponding sample sizes

power_levels <- c(0.70, 0.75, 0.80, 0.85, 0.90, 0.95)

sample_sizes_005 <- c(57, 64, 73, 83, 97, 119)  # Replace with your values

sample_sizes_001 <- c(90, 98, 108, 121, 137, 164)  # Replace with your values

# Create a data frame

df <- data.frame(

  Power = rep(power_levels, 2),

  SampleSize = c(sample_sizes_005, sample_sizes_001),

  Alpha = rep(c("0.05", "0.01"), each = length(power_levels))

)

# Create the plot
```

```
ggplot(df, aes(x = SampleSize, y = Power, color = Alpha)) +
  geom_line() +
  labs(title= "Power vs Sample Size", x = "Sample Size", y = "Power", color = "Significance Level") +
  theme_minimal()
```