ENV 710: Lab 5

Jiahuan Li

'Feb 20, 2023

Write your report in R Markdown, including your answers and your R-code, and submit it to the class Sakai site under the folder Assignments.

Problem 1

Analyze the following data. Write a brief paragraph that includes the following information: (a) your null and alternate hypotheses, (b) the type of test required (one-sample, two-sample, paired), (c) your critical statistic t_c, (d) your correctly reported results (e.g., ($t = 2.10$, df $= 32$, $p = 0.22$), and (e) a clear statement of your conclusion from the analysis. (f) If you made an error in statistical hypothesis testing, what type of error would you have made and why?

A research study was conducted to examine the differences between older and younger adults on perceived life satisfaction. A pilot study was conducted to examine this hypothesis. Ten older adults (over the age of 50) and ten younger adults (between 20 and 30) were given a life satisfaction test (known to have high reliability and validity). Scores on the measure range from 0 to 60 with high scores indicative of high life satisfaction, and low scores indicative of low life satisfaction. The data are presented below. Compute the appropriate t-test.

```r
# Problem 1
young <- c(34, 22, 15, 27, 37, 41, 24, 19, 26, 36)
old <- c(41, 24, 46, 39, 21, 33, 43, 40, 50, 41)

# check for nomality and equal variance
shapiro.test(young)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  young
## W = 0.95959, p-value = 0.7812
```

```r
shapiro.test(old)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  old
## W = 0.90375, p-value = 0.2408
```

```r
var.test(young, old)
```

```
##
##  F test to compare two variances
##
## data:  young and old
## F = 0.85802, num df = 9, denom df = 9, p-value = 0.8233
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.213120 3.454383
## sample estimates:
## ratio of variances
##          0.8580199
```

```r
# critical t. alpha = 0.05, two-sided
t_crit = qt (p = 0.025, df = 18)
print(paste("The critical t value is", t_crit))
```

```
## [1] "The critical t value is -2.10092204024104"
```

```r
t.test(x = young, y = old, var.equal = T, alternative = "two.sided")
```

```
##
##  Two Sample t-test
##
## data:  young and old
## t = -2.4399, df = 18, p-value = 0.02527
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.052468  -1.347532
## sample estimates:
## mean of x mean of y
##      28.1      37.8
```

For the data, the null hypothesis is that there is no significant differences between older and younger adults on perceived life satisfaction. On the contrary, the alternative one is that the significant differences do exist. The t test type is two-sample and two-sided. The critical statistic t_c for a significant level alpha equal to 0.05 is -2.1 or 2.1 (two-tail). The t-test result is ($t =$ -2.44, $df = 18$, $p = 0.025$). Since the p value is lower than the significant level, we can safely reject the null hypothesis. That is, there are significant differences between older and younger adults on perceived life satisfaction. The type 1 error may be made in this process since the p value is 0.025. It reflects there is a probability of 0.025 that the null hypothesis is wrongly rejected.

Problem 2

Using the life satisfaction data from Problem 1, now do the following: (a) calculate the effect size, d, for the study, and (b) conduct a power analysis and report the statistical power. Write a sentence explaining what the power tells you.

```r
# Problem 2

d1 <- abs(mean(young) - mean(old)) / sqrt((sd(young)^2 + sd(old)^2)/2)
print(paste("The effect size is", d1))
```

```
## [1] "The effect size is 1.0911434482251"
```

```r
power <- pwr.t.test(n=10, d=d1, sig.level=0.05, alternative="two.sided")
print(paste("The statistical power is", power$power))
```

```
## [1] "The statistical power is 0.636241512216147"
```

The statistical power means that the probability of correctly rejecting the false null hypothesis is around 0.64. It is the ability to detect a difference when a difference exists. In this case, it is not a very good one since it is less than 0.8, which may be due to the small sample size.

Problem 3

A researcher hypothesizes that electrical stimulation of the lateral habenula will result in decreased food intake (in grams) in rats. Rats undergo stereotaxic surgery and an electrode is implanted in the right lateral habenula. Following a ten-day recovery period, rats (kept at 80% body weight) are tested for the amount of food consumed during a 10-minute period of time both with and without electrical stimulation. The rats are tested under one of the treatments (with or without electrical stimulation) and then tested under the other treatment after a period of rest. Compute the appropriate t-test for the data provided below and write a brief descriptive paragraph including all relevant information (as in Problem 1).

```r
# Problem 3
nostimulation <- c(18.4, 16.1,  9.2, 32.2, 13.8, 16.1, 27.6, 11.5, 11.5, 18.4,
                   22.3, 21.1, 16.4, 29.5, 27.9)
stimulation <- c(27.6, 16.1,  6.9, 25.3, 18.4, 11.5, 32.2, 16.1, 20.7, 23.0,
                 24.7, 18.1, 26.5, 23.4, 8.3)

# check for nomality
shapiro.test(nostimulation)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  nostimulation
## W = 0.94129, p-value = 0.3989
```

```r
shapiro.test(stimulation)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  stimulation
## W = 0.96765, p-value = 0.8218
```

```r
# critical t. alpha = 0.05, one-sided
t_crit1 = qt (p = 0.05, df = 14)
print(paste("The critical t value is", t_crit1))
```

```
## [1] "The critical t value is -1.76131013577489"
```

```r
t.test(x = stimulation, y = nostimulation, paired = T, alternative = "less")
```

```
##
##  Paired t-test
##
## data:  stimulation and nostimulation
## t = 0.22362, df = 14, p-value = 0.5869
## alternative hypothesis: true mean difference is less than 0
## 95 percent confidence interval:
##      -Inf 4.023932
## sample estimates:
## mean difference
##       0.4533333
```

For the study design, the null hypothesis is that electrical stimulation of the lateral habenula will result in no change in the food intake (in grams) of rats. While the alternative one is that electrical stimulation will decrease the food intake. The t test type is paired and one-sided. The critical statistic t_c for a significant level alpha equal to 0.05 is -1.76. The t-test result is ($t = 0.22$, $df = 14$, $p = 0.59$). Since the p value is higher than the significant level, we cannot safely reject the null hypothesis. That is, there is no statistical significant evidence to support that the electrical stimulation will decrease the food intake of rats. And just like in Problem 1, type 1 error may be made in this process with a probability of 0.59.

Problem 4

For your Master's project you have chosen to study whether exposure of Cardinal eggs to direct sunlight increases the mass of the hatchlings. You plan to locate Cardinal nests and open the vegetation around half of the nests (randomly chosen) to increase exposure to sunlight. A previous study on the nesting of the Eastern Bluebird found the effect size to be 0.4.

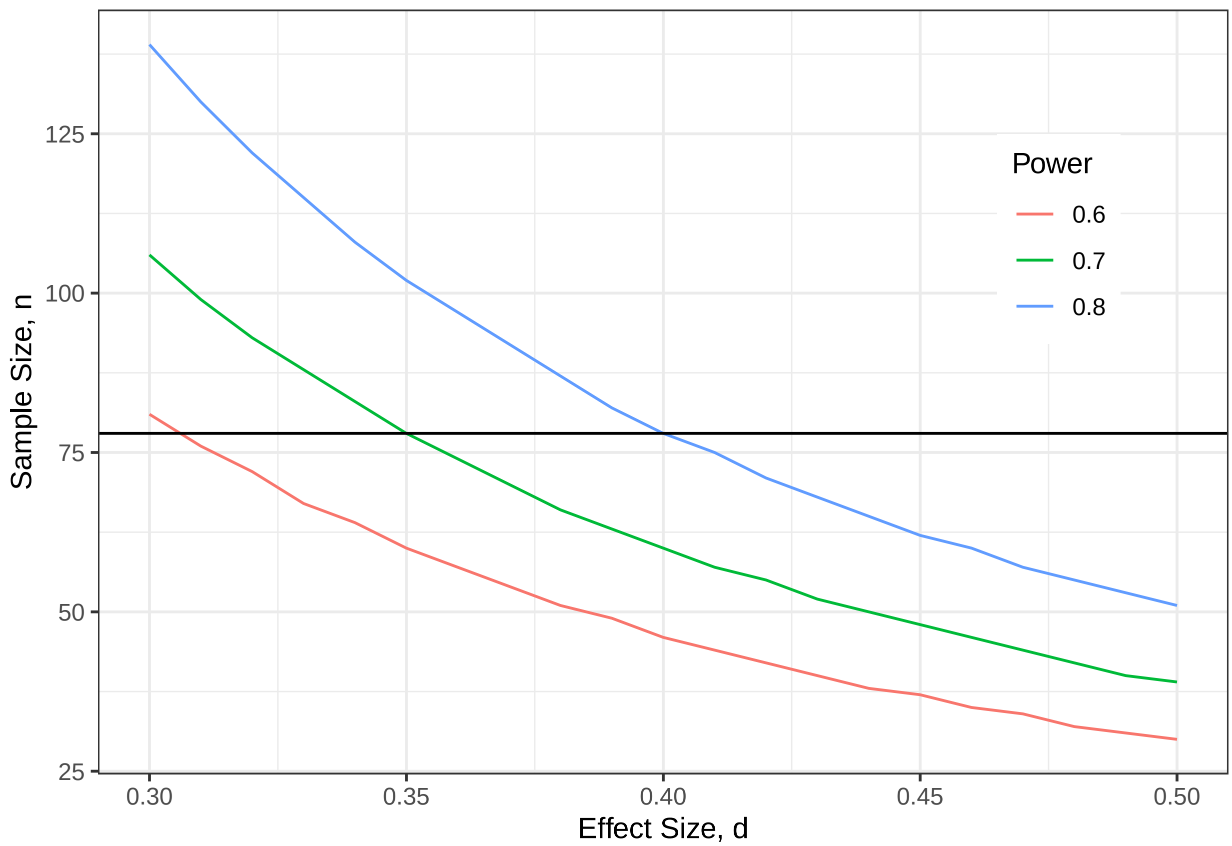
What sample size should you aim for to have sufficient power (0.80) to detect a difference? Alter the above power graph to highlight the needs of your project, assuming that the effect size could vary between 0.3 and 0.5 and that your desired statistical power is between 0.60 and 0.80. Remember that your hypothesis is a one-sided test, and so you need to take this into account in your graph. Add a horizontal black line on your power graph that crosses the power curves at the required sample size. Make sure that sample size is on the y-axis and effect size is on the x-axis of your graph.

```r
# Problem 4
power1 <- pwr.t.test(n=NULL, d = 0.4, sig.level = .05, power = 0.8, alternative = "g")
a <- ceiling(power1$n)
print(paste("The expected sample size is", a))
```

```
## [1] "The expected sample size is 78"
```

```r
library(ggplot2)
# Set the range of effect sizes, d
d <- seq(from = 0.3, to = 0.5, by = 0.01)
nd <- length(d)
# Set the range of power values, p
p <- seq(from = 0.6, to = 0.8, by = 0.1)
np <- length(p)
# Loop over 6 power values and 81 effect sizes to calculate the
```

```r
# sample sizes
samsize <- array(numeric(nd*np), dim=c(nd,np))
for (i in 1:np){
  for (j in 1:nd){
    result <- pwr.t.test(n=NULL, d = d[j], sig.level = .05,
    power = p[i], alternative = "greater")
    samsize[j,i] <- ceiling(result$n)
  }
}
# Turn samesize into a dataframe for ggplot
pwr_dt <- data.frame(cbind(deff=d, ss=as.vector(samsize), n = rep(1:np, each = nd)))
ggplot(data = pwr_dt) +
geom_line(aes(x=deff, y=ss, color = as.factor(n))) +
ylab("Sample Size, n") + xlab("Effect Size, d") +
guides(color = guide_legend(order = -1)) +
scale_color_discrete(labels = c("0.6", "0.7", "0.8"),
name = "Power") + theme_bw() +
theme(legend.position = c(.85, .7)) +
geom_hline(yintercept=78)
```



Problem 5

Change the power curve graph in the lab text so that it plots the range of statistical power with sample sizes from 10 to 100. Make sure that sample size is on the x-axis and effect size is on the y-axis of your graph. What conclusion can you draw about your ability to conduct experiments with 80% power?

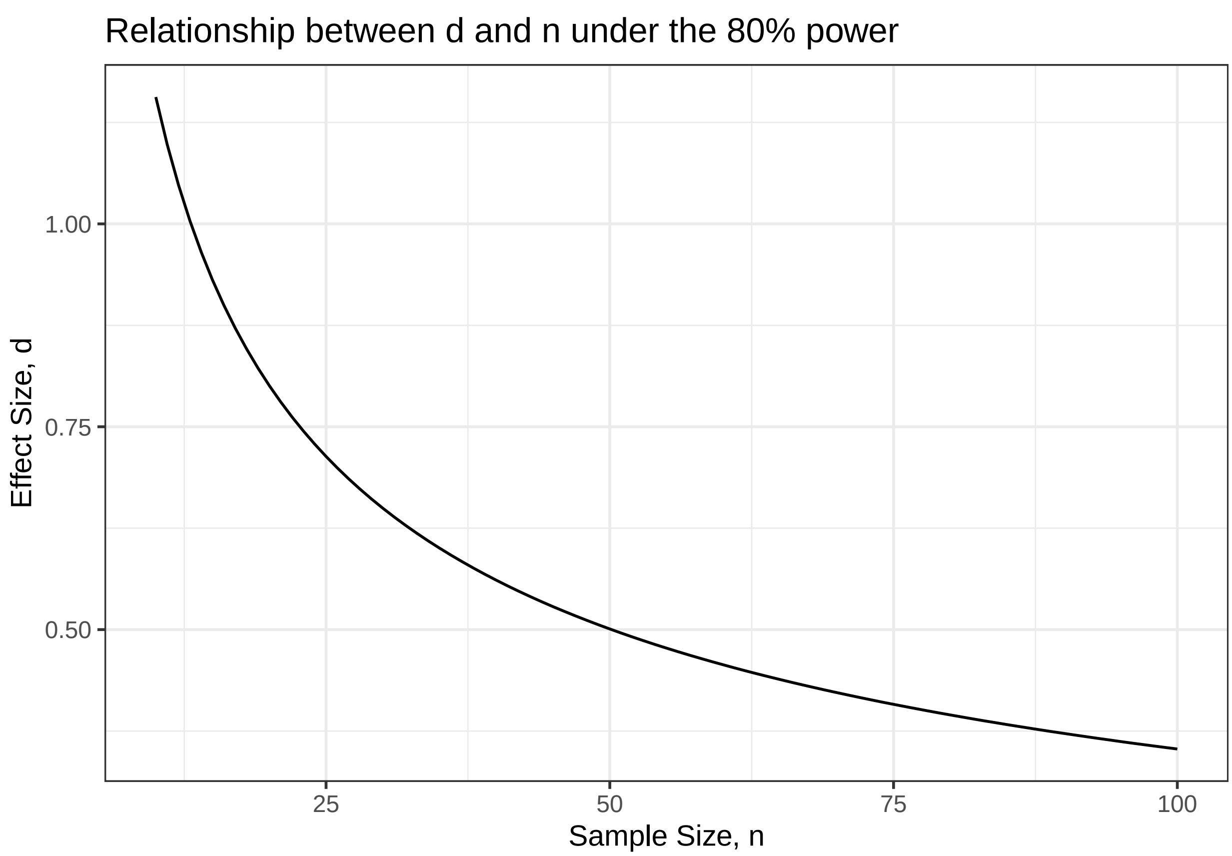
```r
# Problem 5
n <- seq(from = 10, to = 100, by = 1)
nn <- length(n)

ef <- numeric(nn)

for (i in 1:nn){
  result1 <- pwr.t.test(n=n[i], d = NULL, sig.level = .05,
  power = 0.8, alternative = "greater")
  ef[i] <- result1$d
}

pwr_dt1 <- data.frame(cbind(n,ef))

ggplot(data = pwr_dt1) +
  geom_line(aes(x=n, y=ef)) +
  ylab("Effect Size, d") + xlab("Sample Size, n") + theme_bw() +
  ggtitle("Relationship between d and n under the 80% power")
```



To conduct experiments with 80% power, the larger the sample size, the smaller the effect size is needed. As shown in the graph, if we assume the effect size is 0.5 as usual, the required sample size will be 50. And the sample size of around 100 can ensure the 80% power even when d is small.