Business Analytics using Statistical Modeling
Assignment 7

Question 1

Your colleague, a data analyst in your organization, is working on a hypothesis test where he has sampled product satisfaction ratings from customers who are using a new digital watch. He wishes to test whether the mean ($\mathbf{x_i}$) of the satisfaction ratings is higher than the satisfaction rating the company's new products usually get (μ_0)

H_{null}: The mean satisfaction rating of the digital watches is no more than the usual ratings our products get H_{alt}: The mean satisfaction rating is greater than the usual rating our products get diff: the difference we wish to test (x_i - $\bar{x}$) sd: the standard deviation (s) from our sample data n: the size of our sample alpha: the significance level (α) of our test

After collecting data from 50 customers, he informs you that he has found: diff = 0.3 and sd = 2.9, at n = 50 and alpha = 5%. Your colleague estimates that we cannot reject the null hypothesis.

Visualize the null and alternative distributions of t

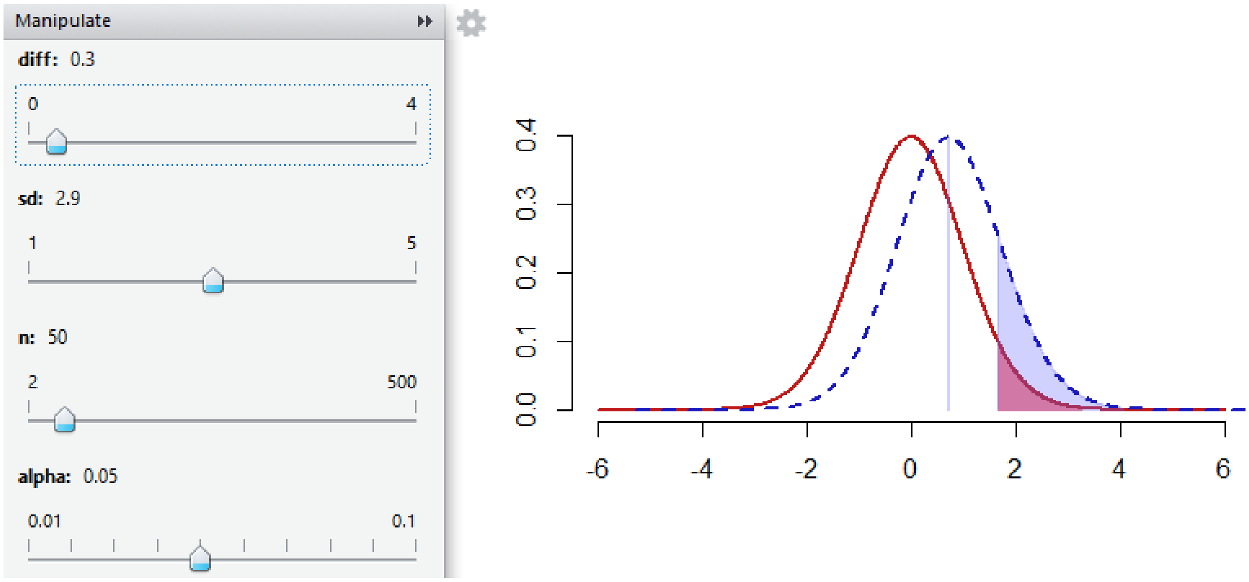


Figure 1: Null and Alternative Distributions of t

Since the mean of the alternative distribution (dashed blue line) is less than the cutoff, we do not reject the null hypothesis.

a. You discover that your colleague wanted to target the general population of Taiwanese users of the product. However, he sent out a survey on popular college social networks, and missed many older customers who you suspect might be more satisfied with the product.

i. Would this scenario create systematic or random error (or neither)?

This scenario would create a systematic error.

ii. Which part of the t-statistic (diff, sd, n, alpha) would be affected?

`diff` would be affected.

iii. Will it increase or decrease our power to reject the null hypothesis?

It will increase our power to reject the null hypothesis.

iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Type I error becomes more likely because of this scenario.

b. You find that 20 of the respondents don't seem to have bought the watch, so they should be removed from the data. These 20 people are just like the others in every other respect.

Visualize the null and alternative distributions of t

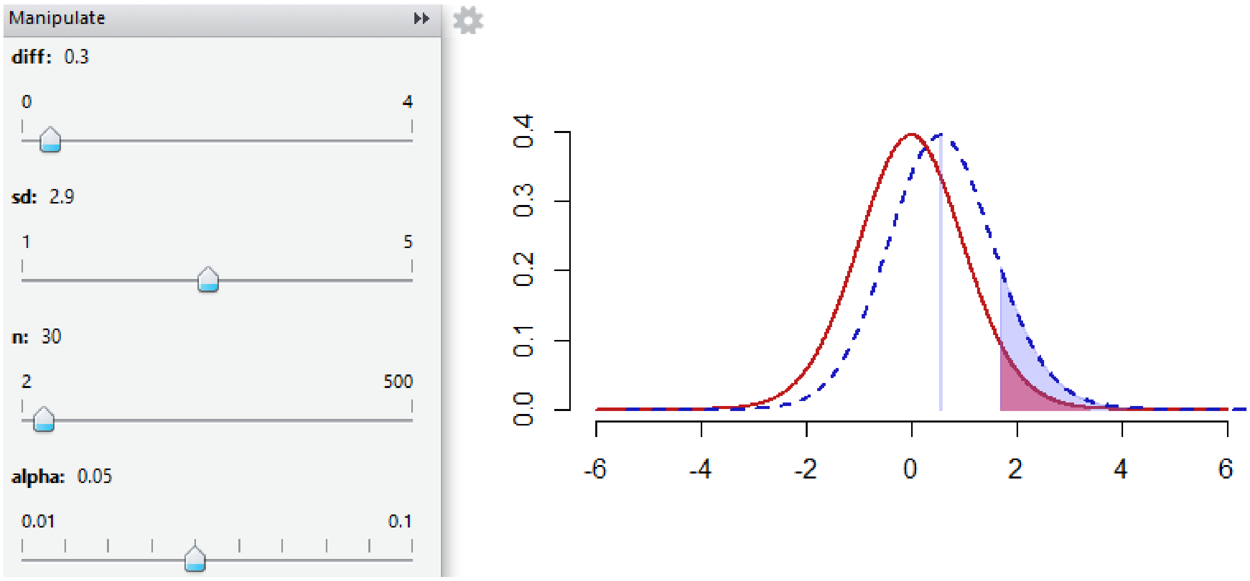


Figure 2: Null and Alternative Distributions of t (b)

i. Would this scenario create systematic or random error (or neither)?

This scenario would create random error.

ii. Which part of the t-statistic (diff, sd, n, alpha) would be affected?

`n` would be affected.

iii. Will it increase or decrease our power to reject the null hypothesis?

It will decrease our power to reject the null hypothesis.

iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Type II error becomes more likely because of this scenario.

c. A very annoying professor visiting your company has criticized your colleague's "95% confidence" criteria, and has suggested relaxing it to just 90%.

Visualize the null and alternative distributions of t

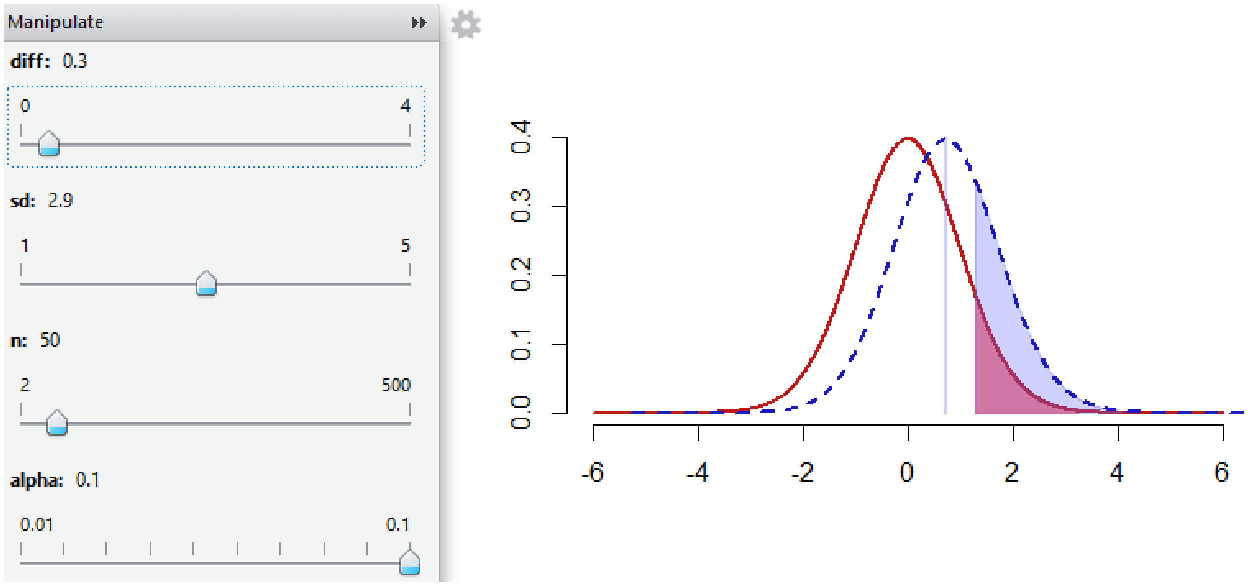


Figure 3: Null and Alternative Distributions of t (c)

i. Would this scenario create systematic or random error (or neither)?

It would create neither systematic error nor random error.

ii. Which part of the t-statistic (diff, sd, n, alpha) would be affected?

`alpha` would be affected.

iii. Will it increase or decrease our power to reject the null hypothesis?

It will increase our power to reject the null hypothesis.

iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Type I error becomes more likely because of this scenario.

d. Your colleague has measured satisfaction as "i would consider buying another digital watch for a friend or family member". But you feel that this measure will score very low for teenagers without disposable incomes, whereas older people might exaggerate their high intentions to buy another watch.

Visualize the null and alternative distributions of t

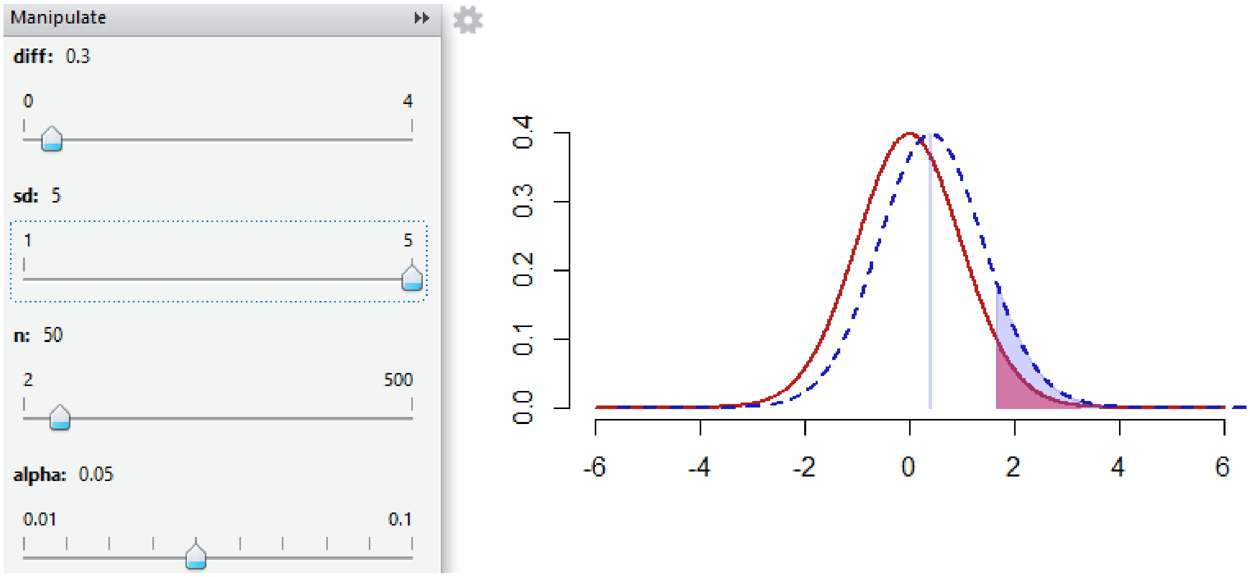


Figure 4: Null and Alternative Distributions of t (c)

i. Would this scenario create systematic or random error (or neither)?

This scenario would create systematic error.

ii. Which part of the t-statistic (diff, sd, n, alpha) would be affected?

`sd` would be affected.

iii. Will it increase or decrease our power to reject the null hypothesis?

It will decrease our power to reject the null hypothesis.

iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Type II error becomes more likely because of this scenario.

Question 2

A psychological research paper has published an experiment to see if emotion affects our perception of color on different color-axes. Participants viewed one of two videos: either the famous death scene in the Lion King, or a video of a desktop screensaver — let's call these the sad and neutral conditions, respectively. Afterwards, participants performed a color discrimination task requiring them to classify colors along the red-green color-axis and blue-yellow color axis. The dependent measures are the accuracy in each of the color conditions (red-green and blue-yellow). The researchers found some potential difference in the blue-yellow accuracy of sad versus neutral participants, but not so for red-green accuracy.

```r
library(data.table)
```

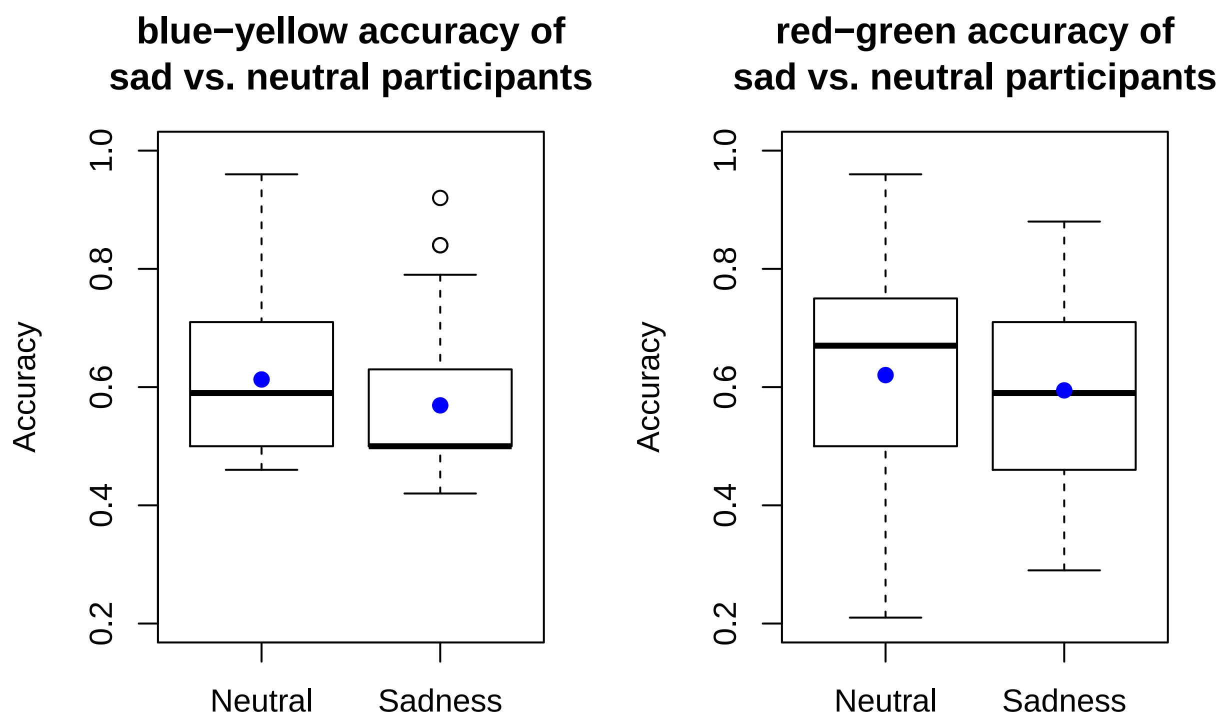
```
## Warning: package 'data.table' was built under R version 3.2.5
```

```r
study2data <- fread('../7-study2Data.csv')
sad_by <- study2data$BY_ACC[study2data$Emotion_Condition == 'Sadness']
neutral_by <- study2data$BY_ACC[study2data$Emotion_Condition == 'Neutral']
sad_rg <- study2data$RG_ACC[study2data$Emotion_Condition == 'Sadness']
neutral_rg <- study2data$RG_ACC[study2data$Emotion_Condition == 'Neutral']
```

a. Visualize the differences between blue-yellow accuracy (BY_ACC) and red-green accuracy (RG_ACC) for both the sad and neutral viewers (Emotion_Condition).

```
par(mfrow = c(1, 2))
boxplot(BY_ACC ~ Emotion_Condition, data = study2data, ylab = 'Accuracy',
        ylim = c(0.2, 1), main = 'blue-yellow accuracy of\nsad vs. neutral participants')
by_means <- tapply(study2data$BY_ACC, study2data$Emotion_Condition, mean)
points(by_means, col = 'blue', pch = 19)

boxplot(RG_ACC ~ Emotion_Condition, data = study2data, ylab = 'Accuracy',
        ylim = c(0.2, 1), main = 'red-green accuracy of\nsad vs. neutral participants')
rg_means <- tapply(study2data$RG_ACC, study2data$Emotion_Condition, mean)
points(rg_means, col = 'blue', pch = 19)
```



b. Run a t-test (traditional or bootstrapped) to check if there is a significant difference in blue-yellow accuracy between sad and neutral participants at 95% confidence.

H_0: $\mu_{\text{sad\_by}} = \mu_{\text{neutral\_by}}$ vs. H_a: $\mu_{\text{sad\_by}} \neq \mu_{\text{neutral\_by}}$

```
t.test(sad_by, neutral_by, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  sad_by and neutral_by
## t = -2.0435, df = 125.61, p-value = 0.04309
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.086308149 -0.001384159
## sample estimates:
## mean of x mean of y
## 0.5690769 0.6129231
```

Since the 95% confidence intervals do not contain 0, we can reject the null hypothesis.

c. Run a t-test (traditional or bootstrapped) to check if there is a significant difference in red-green accuracy between sad and neutral participants at 95% confidence.

H_0: $\mu_{\text{sad\_rg}} = \mu_{\text{neutral\_rg}}$ vs. H_a: $\mu_{\text{sad\_rg}} \neq \mu_{\text{neutral\_rg}}$

```
t.test(sad_rg, neutral_rg, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  sad_rg and neutral_rg
## t = -0.87491, df = 121.98, p-value = 0.3833
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08432635  0.03263405
## sample estimates:
## mean of x mean of y
## 0.5944615 0.6203077
```

Since the 95% confidence intervals contain 0, we cannot reject the null hypothesis.

d. (not graded)

e. Run a factorial design ANOVA where color perception accuracy is determined by emotion (sad vs. neutral), color-axis (RG vs. BY), and the interaction of emotion and color-axis. Are any of these three factors (emotion/color-axis/interaction) possibly influencing color perception accuracy at any meaningful level of confidence?

```
emot_axis_accu1 <- data.table(Subject = study2data$Subject,
                              Emotion_Condition = study2data$Emotion_Condition,
                              Color_Axis = 'RG', Accuracy = study2data$RG_ACC)
emot_axis_accu2 <- data.table(Subject = study2data$Subject,
                              Emotion_Condition = study2data$Emotion_Condition,
                              Color_Axis = 'BY', Accuracy = study2data$BY_ACC)
emot_axis_accu <- rbind(emot_axis_accu1, emot_axis_accu2)
fact_des_anova <- aov(Accuracy ~ Emotion_Condition + Color_Axis +
                        Emotion_Condition:Color_Axis, data = emot_axis_accu)
summary(fact_des_anova)
```

```
##                               Df Sum Sq Mean Sq F value Pr(>F)
## Emotion_Condition              1  0.079 0.07893   3.644 0.0574 .
## Color_Axis                     1  0.017 0.01745   0.806 0.3703
## Emotion_Condition:Color_Axis   1  0.005 0.00527   0.243 0.6224
## Residuals                    256  5.545 0.02166
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

At 90% level of confidence, only `Emotion_Condition` factor that is influencing color perception accuracy.