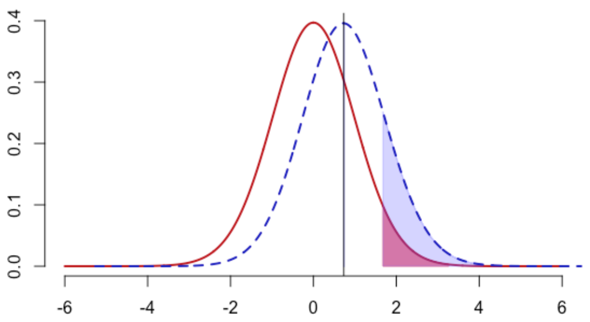
Question I

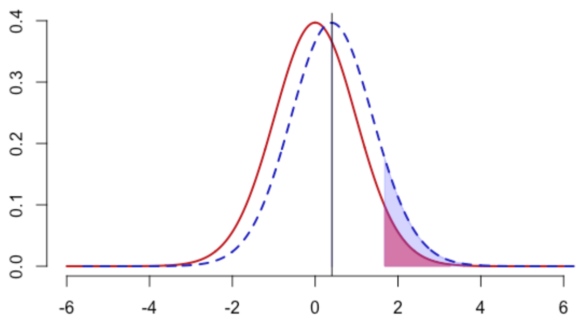
Use the slider bars of the simulation to the values your colleague found, and confirm from the visualization that we cannot reject the null hypothesis at 5% significance. Consider the following scenarios independently using the simulation tool. For each scenario, start with the initial parameters above, then adjust them to answer the following questions:

i. **Would this scenario create systematic or random error (both or neither)?**
ii. **Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?**
iii. **Will it increase or 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?**

a. Your colleague wanted to target the general population of Taiwanese users of the product; however, he only collected data from a pool of young consumers, and missed many older customers who you suspect might use the product much less every day.

i. The lack of data from older consumers will lead to systematic error (bias).
ii. The mean of samples, which take older consumers into consideration, should be less than the original mean. Thus, the diff will decrease; on the other hand, sd may increase due to the samples from older people.
iii. The graph on the left side illustrates the alternative and null distribution when initial parameters are used; on the other hand, the plot on the right side shows the distributions when diff decreases, and sd increases. It can be observed that the power decreases.



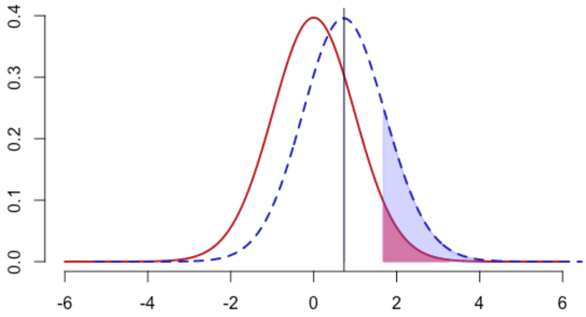


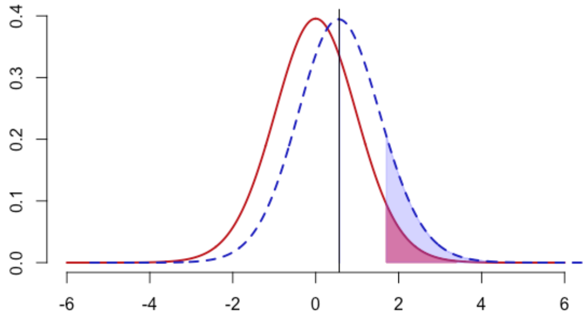
iv. The probability of type I error is alpha; however, the significance does not change, so the type I error will not change. Oppositely, since the power decreases, the type II error will relatively increase. In conclusion, type II error will increase because of this scenario.

b. You find that 20 of the respondents are reporting data from the wrong wearable device, so they should be removed from the data. These 20 people are just like the others in every other respect.

i. Since these 20 respondents are just like the others in every other aspect, I think this scenario will not lead to systematic or random errors.
ii. Also, because the 20 people are just like the others in every other aspect, only n will decrease from 50 to 30.

iii. The graph on the left side illustrates the alternative and null distribution when initial parameters are used; on the other hand, the plot on the right side shows the distributions when n decreases and the other parameters remain the same. It can be observed that the power decreases.





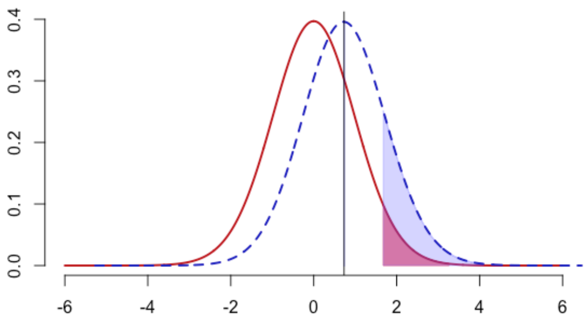
iv. Type II error will increase because of this scenario, and type I error will remain the same. The detailed reason is the same as scenario (a).

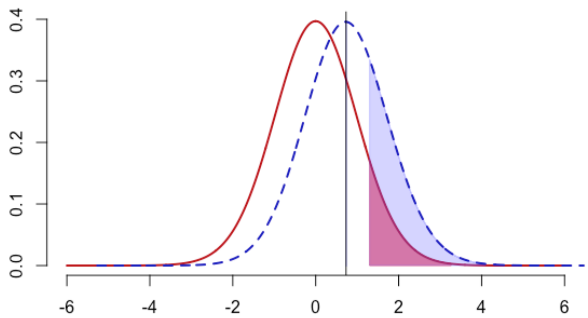
c. A very annoying professor visiting your company has criticized your colleague's 95% confidence criteria, and has suggested relaxing it to just 90%.

i. Since only significance level changes, I think this scenario will not result in both systematic and random error.

ii. Alpha changes from 5% to 10%, the others remain the same.

iii. The graph on the left side illustrates the alternative and null distribution when initial parameters are used; on the other hand, the plot on the right side shows the distributions when alpha increases from 5% to 10%, and the other parameters remain the same. It can be observed that both the purple area (the power to reject the null hypothesis) and the red area increase.





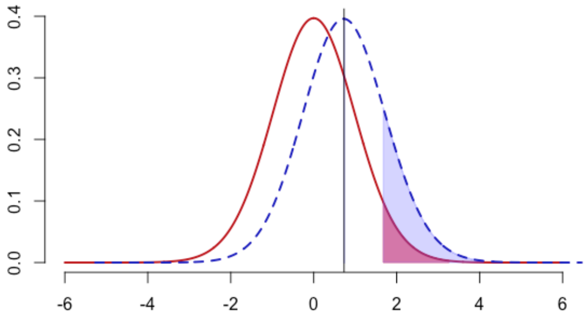
iv. The probability of type I error is alpha; thus, the type I error will increase because of this scenario. According to the plot above, the power increases, which indicates that the type II error will decrease due to this scenario.

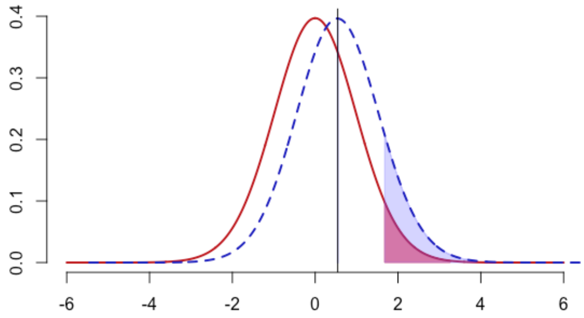
d. Your colleague has measured usage times on five weekdays and taken a daily average. But you feel this will underreport usage for younger people who are very active on weekends, whereas it overreports usage of older users.

i. In this scenario, the samples from younger people are underestimated and the samples from older people are overestimated. Thus, this scenario may lead to random error.

ii. I guess that diff may not differ greatly from its initial value since the underestimation and overestimation occur at the same time. However, the wrong estimation results in random error, which makes sd greater than the initial sd.

iii. The graph on the left side illustrates the alternative and null distribution when initial parameters are used; on the other hand, the plot on the right side shows the distributions when sd increases from 2.9 to 3.9, and the other parameters remain the same. It can be observed that the power decreases.

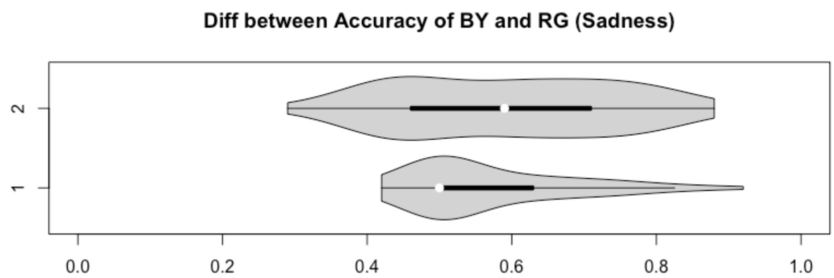


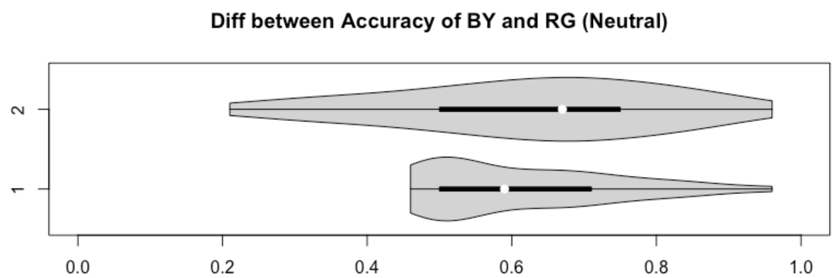


iv. Due to the decline in the power, type II error will increase relatively. The type I error will remain the same, and the detailed reason is the same as scenario (a) and (b).

Question II

a. Visualize the differences between blue-yellow accuracy and red-green accuracy for both the sad and neutral viewers. (You are free to choose any visualization method you wish, but only report the most useful visualizations and any first impressions.)





```
# Read datas
experiment <- read.csv('study2Data.csv', header=TRUE)
BY_data <- with(experiment, data.frame(Subject, Axis='BY',
Emotion_Condition, ACC=BY_ACC, SAD_ESRI))
RG_data <- with(experiment, data.frame(Subject, Axis='RG',
Emotion_Condition, ACC=RG_ACC, SAD_ESRI))

# Use violin plot to distinguish the difference
install.packages("vioplot")
library(vioplot)
par(mfrow=c(2,1))
vioplot(BY_data$ACC[BY_data$Emotion_Condition == "Sadness"],
      RG_data$ACC[RG_data$Emotion_Condition == "Sadness"],
      main = "Diff between Accuracy of BY and RG (Sadness)",
      col="lightgray", horizontal = TRUE, ylim = c(0,1))

vioplot(BY_data$ACC[BY_data$Emotion_Condition == "Neutral"],
      RG_data$ACC[RG_data$Emotion_Condition == "Neutral"],
      main = "Diff between Accuracy of BY and RG (Neutral)",
      col="lightgray", horizontal = TRUE, ylim = c(0,1))
```

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

```
t.test(neu_BY,sad_BY,alternative = "greater",var.equal = TRUE)
    Two Sample t-test

data:  neu_BY and sad_BY
t = 2.0435, df = 128, p-value = 0.02153
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.008296958         Inf
sample estimates:
mean of x mean of y
0.6129231 0.5690769
```

According to the result above, the p-value is less than 0.05. That is, there is a significant difference in blue-yellow accuracy between sad and neutral participants.

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

```
t.test(neu_RG,sad_RG,alternative = "greater",var.equal = TRUE)
    Two Sample t-test

data:  neu_RG and sad_RG
t = 0.87491, df = 128, p-value = 0.1916
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.02309932        Inf
sample estimates:
mean of x mean of y
0.6203077 0.5944615
```

According to the result above, the p-value is far greater than 0.05. That is, there is no significant difference in red-green accuracy between sad and neutral participants.

d. **(Not graded) Do the above t-test support a claim that there is an interaction between emotion and color axis? (i.e., does people's accuracy of color perception along different color-axes depend on their emotion?)**

I think the above t-test can support the claim that there is actually an interaction between emotion and people's accuracy of color perception along a specific color axis (like Blue-yellow). Because there is no significant difference in red-green accuracy between sad and neutral participants.

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?**

```
# Data organization for factorial experiment
total_data <- rbind(BY_data,RG_data)

# ANOVA
aov_result = aov(ACC ~ Axis + Emotion_Condition + Axis:Emotion_Condition,
data=total_data)
summary(aov_result)
                        Df Sum Sq Mean Sq F value Pr(>F)
Axis                     1  0.017 0.01745   0.806 0.3703
Emotion_Condition        1  0.079 0.07893   3.644 0.0574 .
Axis:Emotion_Condition   1  0.005 0.00526   0.243 0.6224
Residuals              256  5.545 0.02166
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The result above shows that these three factors (emotion/color-axis/interaction) don’t significantly influence color perception accuracy at both 95% and 99% confidence level.