BACS_HW_Week9_106071041

106071041

2021/4/25

Question 1

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

systematic. Since he skipped some part of samples intentionally.

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

diff soared and sd dropped. n is still 50 and alpha is still 0.05.

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

increase.

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

It depends how the diff soars and how the sd drops.

b. 20 respondents reporting data from the wrong wearable device

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

systematic. As for what the systematic error, the wrong wearable may makes the statistics all the difference.

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

diff and sd.

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

it depends on the wrong device. For example, if the device is a disaster, diff will be lower and sd could be larger or smaller.

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

It depends. If the wrong wearable device is atually a hit, then the diff will be larger and sd could be lower or larger.

c. suggested relaxing 95% cinfidence to just 90%.

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

random error only.

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

alpha.

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

Increase.

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

Type I

d. underreport usage for younger people who are very active on weekends, whereas it over-reports usage of older users.

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

systematic error. the sample was biased.

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

diff may drop and sd may soar. since the the inactive people’s usage be reported more and the variance become larger.

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

decrease.

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

Type I.

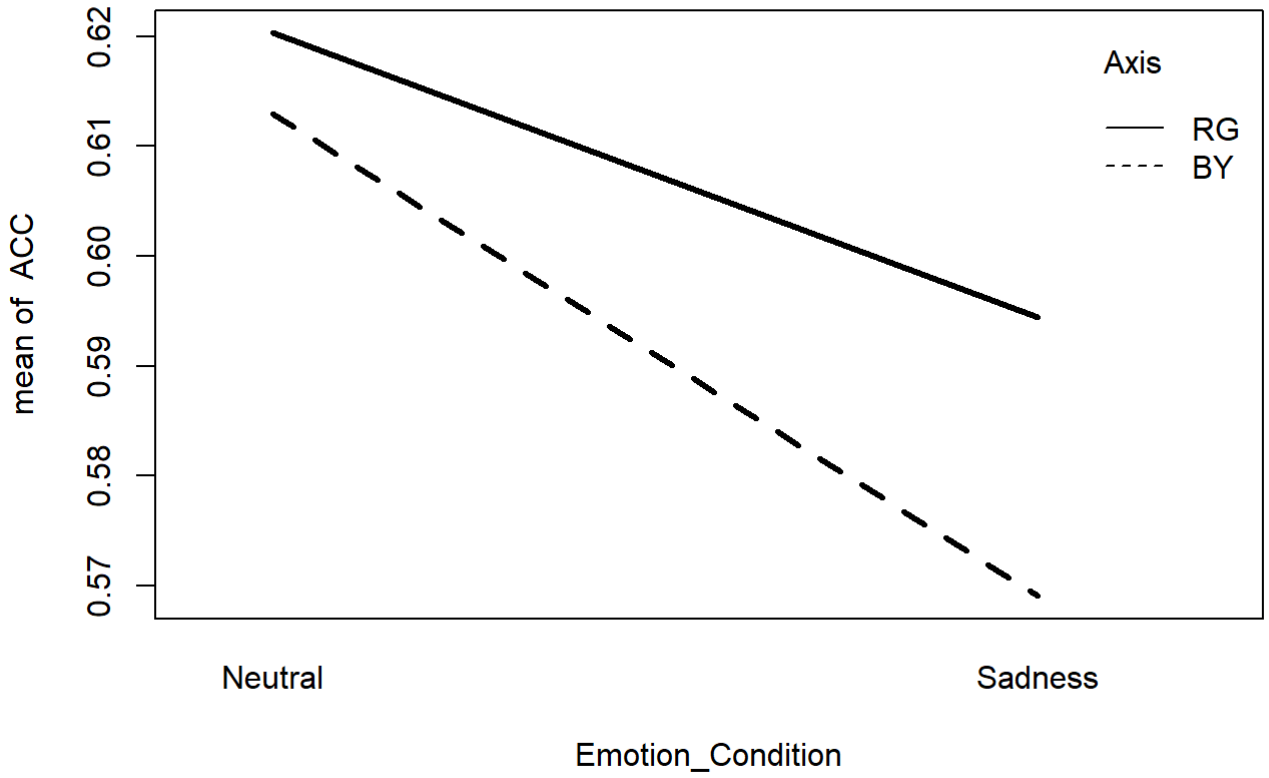
Question 2

```
experiment <- read.csv('study2Data.csv')
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))
```

```
combine <- rbind(BY_data, RG_data)
```

a. Visualize the difference between BY_ACC and RG_ACC for both the sad and neutral viewers

```
with(combine, interaction.plot(x.factor = Emotion_Condition, trace.factor = Axis, response =
ACC, lwd=3))
```



b. t-test/ check if there's a difference between sad and neutral for BY

Yes, there's a significant difference in blue-yellow accuracy between sad and neutral participants at 95% confidenc.

```
Sadness <- subset(BY_data, Emotion_Condition == 'Sadness')
```

```
Neutral <- subset(BY_data, Emotion_Condition == 'Neutral')
```

```
t.test(Sadness$ACC, Neutral$ACC, var.equal = TRUE, conf.level = 0.95)
```

```
##
##  Two Sample t-test
##
## data:  Sadness$ACC and Neutral$ACC
## t = -2.0435, df = 128, p-value = 0.04305
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.086300426 -0.001391882
## sample estimates:
## mean of x mean of y
## 0.5690769 0.6129231
```

c. t-test/ check if there's a difference between sad and neutral for RG

No, there are not.

```
Sadness_RG <- subset(RG_data, Emotion_Condition == 'Sadness')
```

```
Neutral_RG <- subset(RG_data, Emotion_Condition == 'Neutral')
```

```
t.test(Sadness_RG$ACC, Neutral_RG$ACC, var.equal = TRUE, conf.level = 0.95)
```

```
##
##  Two Sample t-test
##
## data:  Sadness_RG$ACC and Neutral_RG$ACC
## t = -0.87491, df = 128, p-value = 0.3833
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08429879  0.03260649
## sample estimates:
## mean of x mean of y
## 0.5944615 0.6203077
```

d. (not graded) Do the above t-tests support a claim that there is an interaction between emotion and color axis?

Yes, while it's not obvious for RG, it's significant for BY.

e. factorial design ANOVA

Emotion Condition can possibly influence color perception accuracy at 90% confidence.

```
summary(aov(formula = ACC ~ Axis + Emotion_Condition + Axis:Emotion_Condition, data = combin
e))
```

```
##                         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
```