BACS HW9

Bijon Setyawan Raya

4/22/2021

Question 1

a. You discover that 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.

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

In this scenario, it will be neither systematic nor random error. However, the data that he has are not representing the actual population being measured.

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

Considering that he has a small data set, he has to gather more data. Once the number of data is increased, it will affect $diff$ and sd.

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

Since Taiwan's population is considered an aging population, my colleague might have missed a great number of older users. If the older users are included in the data set, the possibility of accepting the null hypothesis is high.

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

asd

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.

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

asd

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

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

- Which kind of error (Type I or Type II) 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%.

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

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

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

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

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

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

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

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

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

Question 2

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). You are free to choose any visualization method you wish, but only report the most useful visualizations and any first impressions.

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.

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.

d. (not graded) Do the above t-tests 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? Here, accuracy is an outcome variable, while color-axis and emotion are independent factors)

e. Run a factorial design ANOVA where color perception accuracy is determined by