

Statistics with R

Exercise Sheet 0

Mohamed Genedy

2580760

1. Example of continuous measurement variable: the speed of the authentication while logging in to Email / Social media account. It is a ratio scale
Example of discrete measurement variable: The satisfaction of users with each authentication method. It is an ordinal scale.
2. When answering the following question:
Is the online studying mode more difficult than the in-person mode?
The population: is all of university students around the world.
The Sample: is Saarland University students.
3. Normally when sampling, there are two types which are (Probability Sampling and non-Probability Sampling) the probability sampling is a random technique for choosing the sample and gives each element of the population the equal chance to be chosen in the sample like (Systematic sampling, simple random sampling, etc). The non-probability sampling depends on the ability of researcher of choosing the sample carefully but it happens for the sample to be biased and it's non random sample like (Convenience sampling, Purposive sampling, etc). so I think the researcher decided for one of the two type and when using the non-probability sampling, there should be requirements for choosing the sample.
4. Yes the sample is totally random because as it clearly explained in the previous question, the random sampling gives each element of the population the equal chance to be chosen in the sample and in my example, each student around the world has the same chance to be a part of the sample, that means if I were in the USA, I would choose an American university instead of Saarland University as a sample.
- 5.

| | Independent variables | Scale: | Dependent variables | Scale: |
|---|--|---------|-------------------------|----------|
| 1 | Transportation mode | Nominal | Time | Interval |
| 2 | population measures for “green revolution” | Ordinal | Number of “YES” or “NO” | Ratio |
| 3 | Caffeine | Ratio | Hunger level | Ordinal |

6.

- a. The research is : “If HTTPS Were Secure, I Wouldn’t Need 2FA”- End User and Administrator Mental Models of HTTPS.
- b. The population was: end user and administrator.
- c. The Sample was: 18 End users and 12 Administrators and this sample was totally random.
- d. I didn’t see any bias because the participants were totally random chosen.
- e. The study was about studying end user and administrator mental models of HTTPS. So they interviewed 18 End users and 12 Administrators and asked them and the HTTPS, their understanding of encryption and cryptography, the expectations that they have when visiting a website using HTTPS, etc
- f. The dependent variable is: the understanding of HTTPS, encryption, authentication and taken security decisions.
- g. The independent variable is: the two mental models of End users and Administrators
- h. They used the P-value test in this research and calculated Krippendorff’s Alpha.
- i. It is not clear why they used this kind of test.

Stats with R

Assignment sheet 0

Ahmad Usman Khattak: 7025483

1. Example of a continuous and discrete measurement variable from your interest area:

Smoking in underage children affect the growth

Variable "Birth year": **discrete measurement variable** as the year of birth is precise and no value can come in between the birth year.

Height: **continuous measurement variable** as we cannot measure the precise height, it can be 4 feet 9 inches 89mm 23nanomemeter and so on.

Birth year = Discrete measurement variable / Interval Scale

Height = Continuous measurement variable /Ratio Scale

2. Give an example of a population and of a sample (in the context of some specific research question). Do you know (from studies you have read or heard about), how the samples are typically chosen in your field?

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1516247/pdf/califmed00057-0028.pdf>

Population: All pregnant women is population

Sample: 5659 pregnancies recorded over six years is sample

how the samples are typically chosen in your field?

Pregnant women who smoke and who don't smoke were chosen over the period of six years

3. In your example, is it really a random sample of the population or not? Why?

No because the sample does not represent the whole population, pregnant women from different continents and age factor should have been there in study.

4. Identify the independent and dependent variable in the following example and mention its scale

Comparing the time taken by three different modes of transportations namely car, bus and train

Modes of transportation: Independent Variable / Nominal Scale

Time taken: Dependent Variable/ Ratio Scale

Calculating the number of "YES" or "NO" votes of population measures for "green revolution"

Population measure: Independent Variable / Ratio Scale

Yes or No: Dependent Variable /ordinal Scale

Whether caffeine affects the appetite of a person by measuring the hunger level in a 5point Likert scale after the caffeine intake

Caffeine: Independent Variable / Ratio Scale

Hunger level: Dependent Variable / Interval Scale

5. Please find a recent research paper form an area you're interested in, which includes a study that reports statistical significance. Write down:

Multidimensional gender discrimination in workplace and depressive symptoms

- a) **Abstract:**
Discrimination is associated with depressive symptoms and other negative health effects, but little is known about the mental health risks of workplace gender discrimination. We aimed to investigate the association of workplace gender discrimination and depressive symptoms among employed women in South Korea.
- b) **Population:** Employed women in South Korea
- c) The sample was random as it includes women from 19-64 years of age and both rural and urban areas and it was stratified multistage sampling design
- d) The collection of data was not bias.
- e) Design of the study was observational as the sample was questioned and then conclusion was drawn without any experiment.
- f) **Dependent Variable:** Workplace gender discrimination
- g) **Independent Variable:** Hiring, promotion, paid wages, work assignments, training opportunities, and firing.
- h) All variables are discrete since ordinal variables are discrete
- i) **Measurement Scale:** All variables are ordinal
- j) multivariate logistic regression model
- k) To measure the association of workplace gender discrimination and depressive symptoms

Reference of paper:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0234415>