Measurement Data collection Sampling

Lecture 3.2 – module 4 & 5

28 November 2023

Learning objectives module 4

- Define measurement theory concepts (operationalization, measurement scales, validity, reliability)
- Operationalize a theoretical concept into a simple measurement scale
- Evaluate the measurement scale in terms of validity, reliability and statistical properties

- Distinguish primary and secondary data collection approaches
- Discuss pros and cons of different interview methods
- Explain the structure of a questionnaire
- Critique the format of questionnaire items

Measurement

Non-metric

Metric

Scale	Explanation	Example
Nominal	Identifies groups / classes	Gender; Industries
Order	Ranks groups	Education level
Inter-val	Differences reflect those in real magnitude	Temperature (Celcius)
Ratio	Absolute zero = complete absense of variable	Temperature (Kelvin); Age; Price

Assignment

We conducted a survey with 404 participants to explore data sovereignty impacts on the data economy. This investigation revolves around organizational-level data sovereignty within the context of meta-platforms. We expect that data sovereignty results in improved trust and reduced risk perception, especially for participants who have experience with data sharing. We also expect that improved trust and reduced risk perception lead to a higher willingness to share data.

For the construct `willingness to share data', formulate a measurement item (=question + answer categories) of the form:

1. Semantic differential
2. Likert scale
3. Binary

Menti

(Abbas et al., in preparation)

Assignment

Willingness to share data (WTSD)	I intend to share sensitive data through this meta-platform.	Completely agree – Completely disagree (5-point)
	I predict that I will share sensitive data through this meta-platform in the future.	Completely agree – Completely disagree (5-point)
	It is likely that I will share sensitive data through this meta-platform in the near future.	Completely agree – Completely disagree (5-point)

Suggest a way to evaluate the validity of the measurements:

1. *Concurrent validity*: measure distinguishes objects that are known to be different
2. *Predictive validity*: differentiates objects in relation to future characteristic(s)
3. *Face validity*: items look like they measure the concept they're supposed/intended to

Menti

(Abbas et al., in preparation)

Assignment

Willingness to share data (WTSD)	I intend to share sensitive data through this meta-platform.	Completely agree – Completely disagree (5-point)
	I predict that I will share sensitive data through this meta-platform in the future.	Completely agree – Completely disagree (5-point)
	It is likely that I will share sensitive data through this meta-platform in the near future.	Completely agree – Completely disagree (5-point)
Perceived risk (RISK)	I feel that sharing sensitive data through the meta-platform is risky.	Completely agree – Completely disagree (5-point)
	There will be uncertainty associated with sharing sensitive data through this meta-platform.	Completely agree – Completely disagree (5-point)
	I feel that sharing sensitive data through the meta-platform will negatively affect me.	Completely agree – Completely disagree (5-point)

Suggest a way to evaluate the validity of the measurements:

1. Convergent validity: scores for measuring the same object/concept from two instruments are highly associated.
2. Discriminant validity: scores for measuring two different objects/concepts (from a theoretical perspective) with an instrument are uncorrelated

Menti

(Abbas et al., in preparation)

Assignment

Willingness to share data (WTSD)	I intend to share sensitive data through this meta-platform.	Completely agree – Completely disagree (5-point)
	I predict that I will share sensitive data through this meta-platform in the future.	Completely agree – Completely disagree (5-point)
	It is likely that I will share sensitive data through this meta-platform in the near future.	Completely agree – Completely disagree (5-point)

Suggest a way to evaluate the reliability of the measurements:

1. Test-retest reliability: measure produces consistent results (when same entities are tested at two different points in time or location)
2. Inter-item reliability: Are respondents consistent in their answers to all the items that make up a measure?

Menti

(Abbas et al., in preparation)

Learning objectives module 5

- Explain different probability and non-probability techniques regarding sampling
- Explain the difference between population and sample
- Design and reflect upon a sampling procedure for a given case
- … apply core concepts of inferential statistics
  - Type-1 and type-2 errors
  - Power and significance
- … apply the five steps of inferential hypotheses testing
- … select a proper test statistic given a hypotheses and measurement level, based on a test selection table that is given to you

Assignment

We conducted a survey with 404 participants to explore data sovereignty impacts on the data economy. This investigation revolves around organizational-level data sovereignty within the context of meta-platforms. We expect that data sovereignty results in improved trust and reduced risk perception, especially for participants who have experience with data sharing. We also expect that improved trust and reduced risk perception lead to a higher willingness to share data.

For this study, what is the

1. Population
2. Element
3. Sample
4. Subject

(Abbas et al., in preparation)

Assignment

We conducted a survey with 404 participants to explore data sovereignty impacts on the data economy. This investigation revolves around organizational-level data sovereignty within the context of meta-platforms. We expect that data sovereignty results in improved trust and reduced risk perception, especially for participants who have experience with data sharing. We also expect that improved trust and reduced risk perception lead to a higher willingness to share data.

Assume we have a database of all individuals in the United Kingdom, by collaborating with the UK Office of National Statistics. The database contains characteristics such as age, sex and ethnicity.

- How would you create a representative sample of 404 participants from the database?

Menti

(Abbas et al., in preparation)

Assignment

We conducted a survey with 404 participants to explore data sovereignty impacts on the data economy. This investigation revolves around organizational-level data sovereignty within the context of meta-platforms. We expect that data sovereignty results in improved trust and reduced risk perception, especially for participants who have experience with data sharing. We also expect that improved trust and reduced risk perception lead to a higher willingness to share data.

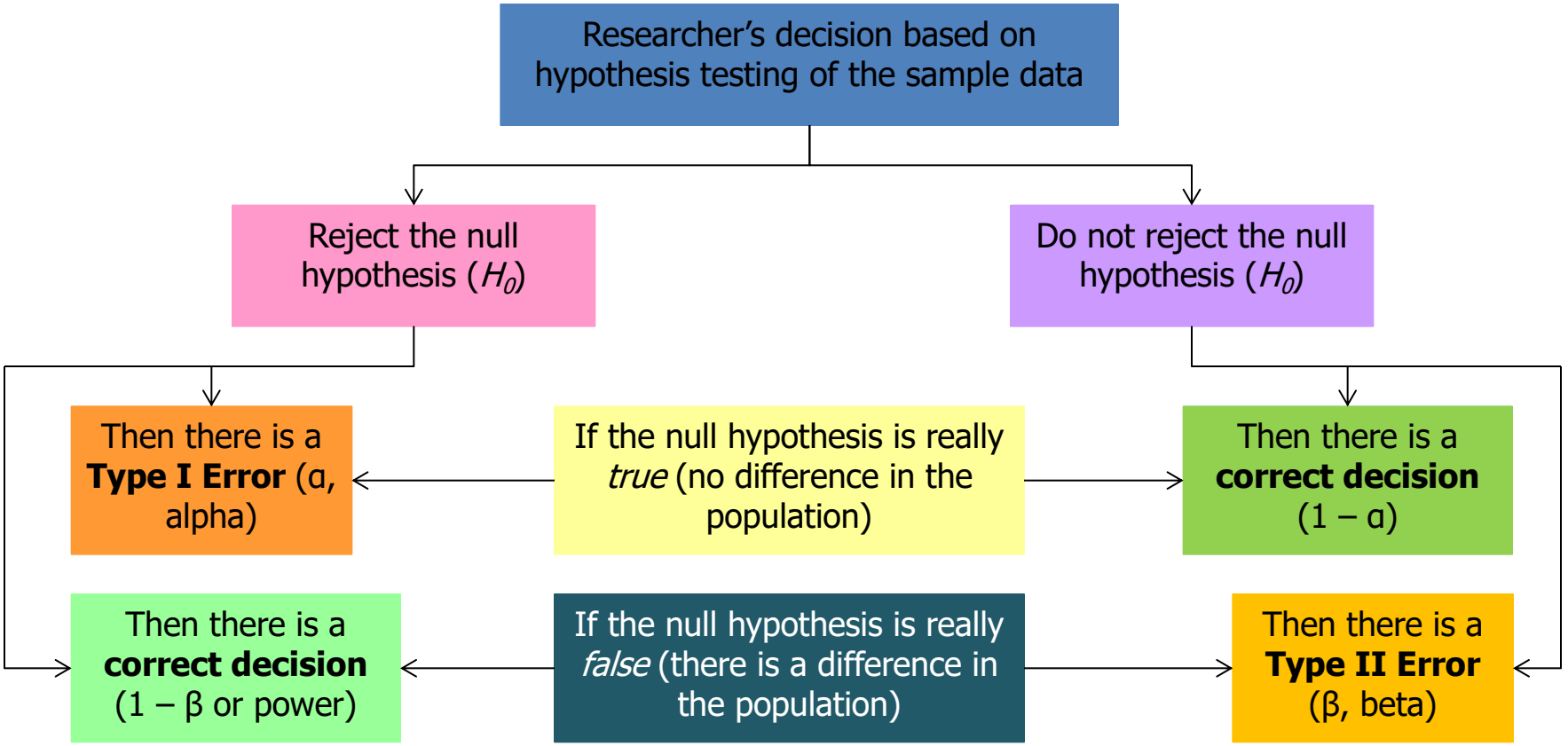
Assume we have a database of all individuals in the United Kingdom, by collaborating with the UK Office of National Statistics. The database contains characteristics such as age, sex and ethnicity.

1. What is the sampling frame?
2. How would you do systematic sampling?
3. What is an example of a relevant stratum?

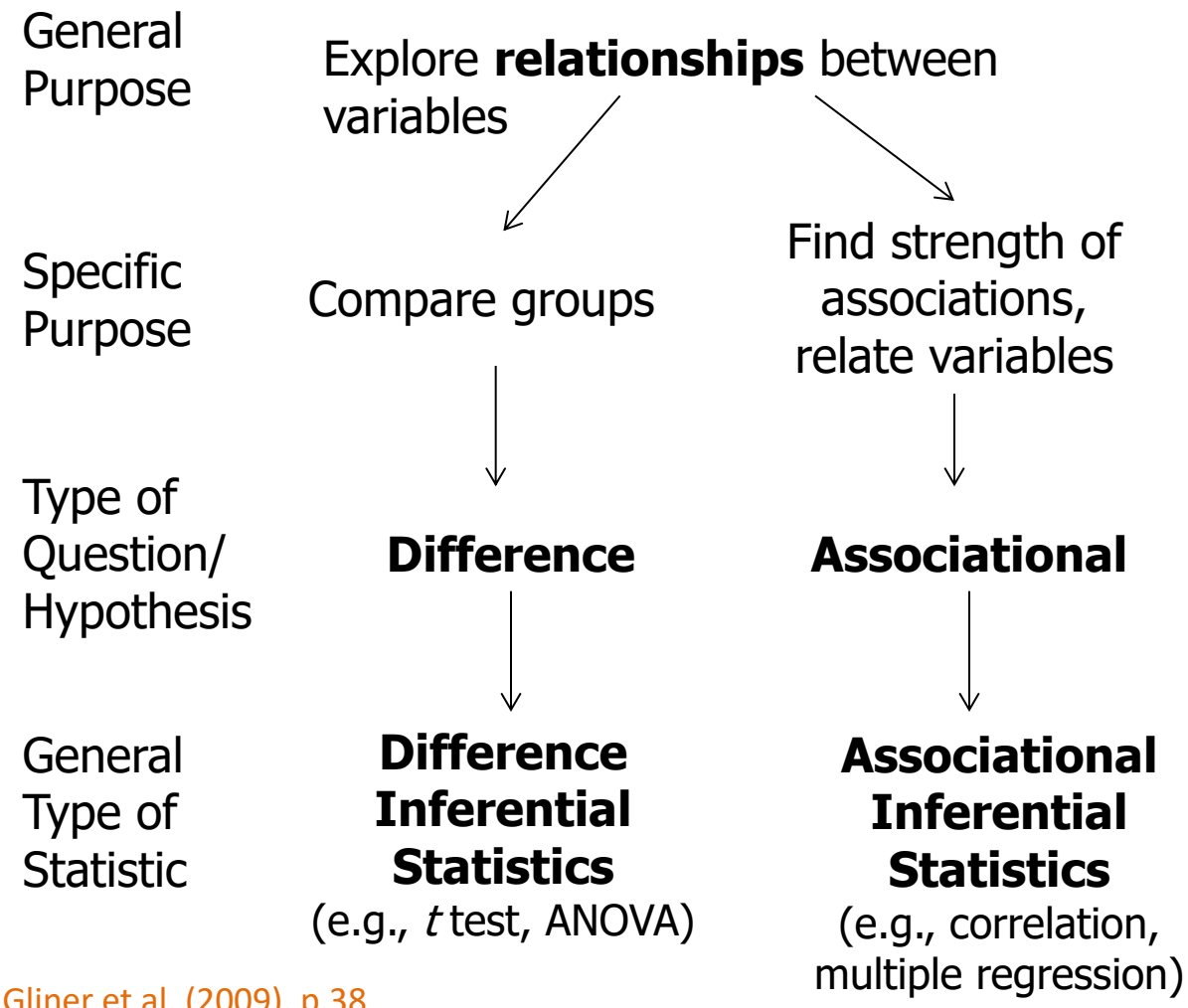
Menti

(Abbas et al., in preparation)

Type-1 and type-2 errors

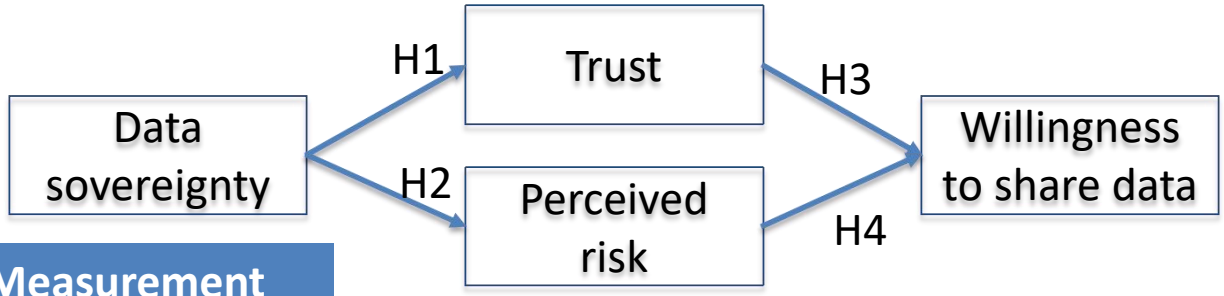


Selecting the proper test statistic



Source: Gliner et al. (2009), p.38

Assignment

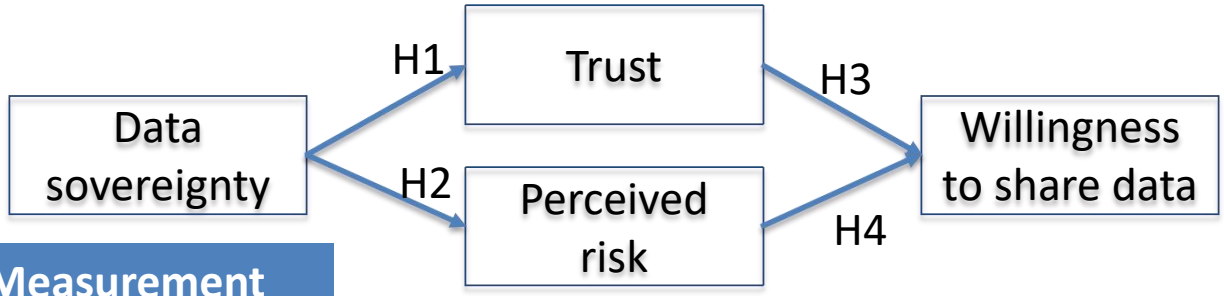


Construct	Measurement type
Data sovereignty	**Assumption:** Binary
Trust	Likert-scale
Control	Likert-scale
Willingness to share data	Likert-scale

1. Please indicate the type of hypothesis (difference / association)
H1:
H2:
H3:
H4:
2. Under which conditions can we use a parametric test to evaluate H3?

(Abbas et al., in preparation)

Assignment



Construct	Measurement type
Data sovereignty	**Assumption:** Binary
Trust	Likert-scale
Control	Likert-scale
Willingness to share data	Likert-scale

While testing H3, the p-value turns out to be .045

1. What is your conclusion: is the hypothesis supported or not?
2. Interpret: Is there an effect of trust on willingness to share data?
3. By reaching this conclusion, is there a risk of a type-1 or type-2 error?
4. The researchers decide to increase the sample size with 300 more participants. Do you expect the p-value to decrease, increase or stay the same?

(Abbas et al., in preparation)

To do before Monday

- Read S&B: Ch 7, 9, 12, 13
- Weekly assignment 3 (deadline Thursday)

- Monday:
  - Wrap-up module 4 & 5
    - Q&A
    - Generic feedback
    - Quizzes
  - Intro to module 6 & 7