Quantitative research methods: Statistical thinking and research process

Shuhei Nomura, PhD
s-nomura@keio.jp
https://researchmap.jp/nom3.shu

Associate Professor, Keio University, Japan
PhD in Biostatistics and Epidemiology
from Imperial College London

- Global Burden of Disease (GBD)  
  GBD Scientific Council (Seattle)

- Global Health Policy
  Bill & Melinda Gates Foundation Tokyo Office (Tokyo), SEEK Development (Berlin),
  2023 Hiroshima G7 Global Health Taskforce (Tokyo)

- Domestic Health Policy

- Health Emergency and Disaster Risk Management (Health-EDRM)
  World Health Organization Centre for Health Development (Kobe)

- Nutrition Science and Policy  
  Global Nutrition Report (GNR) (London)

- Pandemic prevention, preparedness, and response (PPR)
  Google, LINE, Yahoo! Japan, Ministry of Health, Labour and Welfare
Outline

• Statistical thinking and research process
• Case study
Why learn statistics?

• The reason you are here is because you possess an inquiring mind!
  o What are the most in-demand medical procedures following an earthquake?
  o What potential health risks does climate change introduce?
  o To what extent does a building's earthquake resistance mitigate health damage?
  o Does a lockdown policy effectively reduce the risk of COVID-19 infection?

• To answer the questions of interest, statistical thinking is necessary.
What is statistical thinking?

- **Statistical thinking** is an approach to problem-solving and decision-making that relies on statistics. It involves using data and statistical techniques to gain insights, make informed decisions, and draw valid conclusions.
The quantitative research process

**Data**
- Initial observation (research question)
- Generate hypotheses
- Collect data to test hypotheses
- Analysis data

**Identify variables**
- Measure variables
- Graph data
- Fit a model

Statistical thinking plays a role in all these processes
Case study – the Fukushima nuclear disaster on 11 March 2011

< 20km, mandatory evacuation zone, 12 March

20–30km, voluntary evacuation zone, 15 March
Case study – the Fukushima nuclear disaster on 11 March 2011

The quantitative research process

- **Data**
  - Initial observation (research question)

- **Identify variables**
  - Generate hypotheses

- **Measure variables**
  - Collect data to test hypotheses

- • Graph data
  • Fit a model
  - Analysis data
Case study – the Fukushima nuclear disaster on 11 March 2011

The quantitative research process

- Data
- Initial observation (research question)
- Generate hypotheses
- Collect data to test hypotheses
- Fit a model

• A total of 328 individuals from all five nursing facilities in Minamisoma were evacuated within two weeks.

• However, reports have emerged of deaths occurring shortly after the evacuation.

• [Question] Has there been an increase in the mortality rate among facility residents after the evacuation compared to before?

• Data on residents of the nursing facilities.
Case study – the Fukushima nuclear disaster on 11 March 2011

The quantitative research process

- Data
  - Initial observation (research question)
- Identify variables
  - Generate hypotheses
- Measure variables
  - Collect data to test hypotheses
  - Analysis data
  - • Graph data
  - • Fit a model

- [Hypothesis] Evacuation was related to mortality.
- More specifically, the duration of stay until death for facility residents differed before and after the evacuation.
- A variable indicating whether facility residents have survived or passed away is required.
- To be precise, this refers to the duration of each facility resident's stay, including prior to the disaster, and their mortality status at the end of the observation period – in other words, time to event.
Case study – the Fukushima nuclear disaster on 11 March 2011

The quantitative research process

- Coordinate with the original facilities and collect data on their residents, including information on past residents, date of evacuation, etc.
- Coordinate with the facilities that served as evacuation destinations to collect survival information on the evacuees.
- Monitor the evacuees, verify the dates of their deaths post-evacuation, and calculate the number of days they survived.

- Data
- Identify variables
- Measure variables
- Initial observation (research question)
- Generate hypotheses
- Collect data to test hypotheses
- Graph data
- Fit a model
- Analysis data
Case study – the Fukushima nuclear disaster on 11 March 2011

The quantitative research process

- Data
  - Initial observation (research question)
- Identify variables
  - Generate hypotheses
  - Collect data to test hypotheses
- Measure variables
  - Analysis data
  - Graph data
  - Fit a model

- Survival analysis, or more generally, time-to-event analysis.
- Next slide!
Case study – the Fukushima nuclear disaster on 11 March 2011

Estimated pre- and post-disaster (evacuation) survival
(using the Kaplan-Meier product limit method)

- Graph data
- Fit a model

Case study – the Fukushima nuclear disaster on 11 March 2011

Multiple regression model of survival
(using the Cox proportional hazards regression method)

<table>
<thead>
<tr>
<th>Disaster (evacuation)</th>
<th>Hazard ratio</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>1.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>After</td>
<td>2.88</td>
<td>1.74 to 4.76</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Adjusted for age, gender, facility id, care level. CI: confidence interval

Case study – the Fukushima nuclear disaster on 11 March 2011

• Implications for guiding policy
  o Evacuation of older adults carries health risks.
  o Preparations should be made to ensure safe evacuation - such as securing evacuation destinations and means of transport.
  o Temporary indoor evacuation is also an important option.
Benchmarking outcomes is difficult

- Health measures and metrics are rarely comparable – across different data types, various locations, and/or timeframes.

- It poses a challenge to isolate differences in health performance to true differences in health outcomes – and not simply differences in measurement methods.

- Collecting data during disasters proves challenging, and there exist limitations to the kind and amount of data that can be acquired.
Summary

• **Importance of statistical thinking:** Statistical thinking plays a vital role in objectively analyzing and interpreting information, thus facilitating data-driven decision-making processes.

• **Understanding research processes:** The four main processes - observation, hypothesis formulation, data collection, and analysis - form the backbone of any quantitative research endeavor.

• **Lessons from case study:** While the Fukushima case offered valuable insights, it also shed light on how statistical thinking helped decipher complex situations.

• **Cooperation with statistical professionals:** Regardless of your own expertise, working with statistical professionals from the start can greatly enhance the research process and outcomes.