Why sampling matters

Prof. Mark NK Saunders
Learning outcomes…

You will…

- be aware of the relationship between research question, population, the sample selected and the implications of sample size.
- understand the implications of using different probability and non-probability sampling techniques and the need to combine techniques within a research project.
- be able to explain sample selection precisely in their method and justify their selection for their own research project.
- have the opportunity to ask questions
Aldi’s “Like Aldi… like the price” campaign

What is the population?

a) Supermarket customers
b) Aldi customers
c) Special K likers?
d) Special K likers who shop at Aldi
e) Other?

Source: http://www.likealdi.co.uk/home/
Population, sample and individual case

- Population (179)
- Sample (72% = 129)
- Individual case or element

But what characteristics do these “Special K likers” have? Are they an entire population or a sample?

How was the population actually defined?
Sampling questions from Aldi

- How has the population from which the sample is selected been defined?
- Why has the population been defined in this way?
- What are the population’s characteristics?
- How does the population relate to the research question being addressed?

- How was the sample selected?
- What technique(s) were used to select the sample?
- What are the implications of the technique(s) for the characteristics of the sample?
Yves St Laurent’s skin cream

'Holy grail' of skin cream found

Yves Saint Laurent, the beauty and fashion firm, will this week launch a product which actually rejuvenates the skin.

“A small test of 50 women found almost nine out of 10 of them said their skin looked more luminous after using it, while 72 per cent said fine lines appeared less visible.”
Probability sampling - and sample size?

The problem of making statistical inferences from a small probability sample

Sample size for 95% confidence level with 5% margin of error

<table>
<thead>
<tr>
<th>Population</th>
<th>5%</th>
<th>2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000,000</td>
<td>384</td>
<td>2395</td>
</tr>
<tr>
<td>10,000,000</td>
<td>384</td>
<td>2400</td>
</tr>
<tr>
<td>34,100,000</td>
<td>384</td>
<td>2401</td>
</tr>
</tbody>
</table>

Yet the total population of women in UK is 31.4 million (Source: 2011 Census)
Additional sampling questions from YSL

- Was a probability sample (or a quota sample) used? (need to use such sampling techniques to make statistical estimates)

- Which precise sampling technique was used to select the sample, and what were the implications for the nature of the sample?

Assuming a sample of 50, a confidence level of 95% and a population of 34.1 million this gives a 14% margin of error:

somewhere between 58% and 86% said ‘fine lines appeared less visible’…

- What are the implications of sample size for utility of findings?
Sampling matters!

- Total population (and its characteristics) must be defined in relation to the research question to be answered.

- Relationship between total population and sample selection technique (probability or non-probability) impact on research question that can be answered.

- Selection of appropriate probability and/or non-probability sample selection techniques depends on the research question being answered, *but* which should be used when?

- Sample size is important, *but* how many?
A true story...
Possible implied research question: How have people managed their careers in order to be successful?
Data collected by …

- Catching commuter train from Guildford to London
- 2nd class ticket
- 5 weeks – Monday to Friday
- Interviewing 5 people on average each journey

Unclear regarding:
- Time of year
- How selecting respondents
Characteristics of population from which sample drawn

- Catching commuter train (commuters?)
- Living near Portsmouth – Guildford – London line
- Working in London
- Working “9 to 5”

What else?
Representativeness

To what extent does the sample represent exactly the population from which it is drawn?

- Problems …
  - Total population unknown
  - Sampling frame unknown
  - Only c.125 interviews

- Must have used non-probability sampling

**Non-probability sample:** where the chance of each case being selected is not known

**Probability sample:** where the chance of each case being selected is known and not zero
Generalisability

How the way in which the sample is drawn affects the extent to which the findings can be applied (generalised) or other settings

Sample from:
- Commuters
- Living south of London
- Unlikely to be senior managers
- Working “9 to 5”
- Unlikely to be employed in manufacturing

Research Question
- People
- Living anywhere
- With a career
- Doing any type of job
Unclear how non-probability sample was drawn

Different non-probability techniques have differing implications
Non probability sampling – what sample size?

- **REMEMBER** – making generalisations to theory *not* about a population (except with quota samples)
- Sample size depends upon research question and objectives…
  - 😊 what need to find out
  - 😊 what will be useful
  - 😊 what is credible
  - 😊 what can be done within available resources

Continue until reach data saturation
## Non-probability sample size ...more detail

<table>
<thead>
<tr>
<th>Author</th>
<th>Nature of study</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bertaux (1981)</td>
<td>Qualitative</td>
<td>15</td>
</tr>
<tr>
<td>Kvale and Brinkmann (2009)</td>
<td>Interviews</td>
<td>5-25</td>
</tr>
<tr>
<td>Bernard (2000); Morse (1994)</td>
<td>Ethnographic</td>
<td>35 - 36</td>
</tr>
<tr>
<td>Creswell (1998); Morse (1994)</td>
<td>Grounded theory</td>
<td>20 - 35</td>
</tr>
<tr>
<td>Creswell (1998); Morse (1994)</td>
<td>Phenomenological</td>
<td>5 - 25</td>
</tr>
<tr>
<td>Guest et al. (2006); Kuzel (1992); Romney, Batchelder and Weller (1986)</td>
<td>Considering a homogenous population</td>
<td>4 – 12</td>
</tr>
<tr>
<td>Kuzel (1992); Creswell (2007)</td>
<td>Considering a heterogeneous population</td>
<td>12 - 30</td>
</tr>
</tbody>
</table>

Source: Saunders (2012)
Different probability sampling techniques also have different implications (given known total population)
The bust billboard ever: 'Hello Boys' Wonderbra poster voted most iconic ad image of all time

By DAILY MAIL REPORTER
UPDATED: 12:15, 31 March 2011

They stunned a generation and sent millions of men weak at the knees.

Now Czech model Eva Herzigova’s traffic-stopping billboard for Wonderbra has been declared the most eye-catching advert of the past few decades.

The image, part of the lingerie company’s ‘Hello Boys’ campaign, featured a huge photo of Eva’s cleavage.

Traffic-stopper: Eva Herzigova’s Wonderbra billboards have been declared the most eye-catching ads of the past few decades

The ad sparked a sensation when it was unveiled in 1994 and was blamed for stopping traffic and causing accidents as commuters stared at the huge roadside posters.

The Herzigova ‘Hello Boys’ ads propelled her to stardom around the world, making her a household name.

Now the advert has won a public vote as the favourite ‘iconic’ advertising image in a poll by the Outdoor Media Centre, the trade body for outdoor advertising.
More detail

In February 2011 the UK’s Outdoor Media Centre launched its ‘Hall of Fame’ competition to identify the 100 best advertising posters of all time. Working with the History of Advertising Trust, they generated a list of 500 posters. This was reduced to a shortlist of 228 posters by a committee of media and creative experts together with the editor of weekly magazine - Campaign. These were displayed on a dedicated website www.outdoorhalloffame.co.uk. Creative agencies, media planners, advertisers, media owners and the general public were invited in an article in Campaign to go to the web site, view the advertising campaigns and cast their votes for what they considered to be the best outdoor posters. Each person was able to cast a total of ten votes, the best advertisements being “chosen after more than 10,000 reader votes”.
Sample selection mattered!
Sample selection ... realities in management research

**REALITIES**
- Don’t always have sampling frame
- Sampling frame often restricted
- Difficult to show sample truly random

**CONSEQUENCES**
- Population assumed itself to be a random sample from larger population or
- Inferences (including statistical) confined to actual population from which sample drawn –can’t generalise statistically
- May not be able to answer certain questions... rephrase the question!

Ideally generalise to the theoretical population of managers in UK & Eire

Can gain access to study population of managers taking PG courses at a University

Using the sampling frame of their email addresses

Sample comprises those who respond
Sampling matters…

remember:

- Statistical representativeness requires probability sampling
- Probability sampling requires a sampling frame
- Relatively large sample sizes are needed for statistical representativeness, but beware of effect size
- Generalisability requirements necessitate careful thought at the sample selection stage
- Choice of sample always has implications for the Research Question that can be answered and the data collection technique(s) used
References (1)

References (2)