

Sampling in Psychology, Sociology and Pedagogy

Content of Manual

Iryna Lupan

Volodymyr Vynnychenko Central Ukrainian
State Pedagogical University (CUSPU), Ukraine



Co-funded by the
Erasmus+ Programme
of the European Union

2018/06/26



Why SAMPLING?

Sampling is the selection of a subset (**a sample**) of individuals from within a **target population** to estimate characteristics of the whole population, because the **census** is

- impossible
- expensive

OR requires long time survey



Introduction

The objectives of our manual are to teach students of basic ideas of sampling from an applied perspective and to support the course.

Traditionally the course would cover the main techniques used in actual sampling practice:

- simple random sampling
- stratification
- systematic sampling
- cluster sampling
- multistage sampling



Co-funded by the
Erasmus+ Programme
of the European Union



Part 1. Elements of the sampling problem

1.1. Population and sample

1.2. Sampling frames, sampling types

1.3. Scheme of sampling. The Design of the Sample Survey

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Example of tasks:

For each survey describe the target population, sampling frame, sampling unit, and observation unit. Discuss any possible sources of selection bias or inaccuracy of responses:

Karras (2008) reports on a survey conducted by *SELF* magazine on prevalence of eating disorders in women. The survey, posted online at *self.com*, obtained responses from 4000 women. Based on these responses, the article reports that 27% of women in the survey “say they would be ‘extremely upset’ if they gained just 5 pounds”; it is estimated that 10% of women have eating disorders such as anorexia or bulimia.

Ann Landers (1976) asked readers of her column to respond to the question: “If you had it to do over again, would you have children?” About 70% of the readers who responded said “No.” She received over 10,000 responses, 80% of those from women.



Co-funded by the
Erasmus+ Programme
of the European Union



Part 2. Simple random sampling

2.1. Estimation of a Population Total

2.2. Estimation of a Population Mean

2.3. Estimation of a Population Proportion

2.4. Sample random sampling with replacement.

2.5. Efficiency of the sampling.

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Example of tasks:

In 2005, the Statistical Society of Canada (SSC) had 864 members listed in the online directory. An SRS of 150 of the members was selected; the sex and employment category (industry, academic, government) was ascertained for each person in the SRS, with results in file `ssc.dat`.

- a) What are the possible causes of selection bias in this sample?
- b) Estimate the percentage of members who are female, and give a 95% CI for your estimate.
- c) Assuming that all members are listed in the online directory, estimate the total number of SSC members who are female, along with a 95% CI.



Co-funded by the
Erasmus+ Programme
of the European Union



Part 3. Systematic sampling

3.1. How to Draw a Systematic Sample.

3.2. Estimation of a Population Mean, Total and Proportion.

3.3. Selecting the Sample Size.

3.4. Comparing Estimates. Effectiveness of systematic sampling.

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Example:

The list of inhabitants of a certain village is written as follows:

H - husband,
W - wife;
d - daughter,
s - son

House number	Inhabitants				
1	H	W	d	s	
2	H	W	d	s	
3	H	W	s	s	d
4	H	W	s	d	d
5	H	W	d	d	
6	H	W	d	s	s
7	H	W	d	s	
8	H	W	s	d	
9	H	W	s		
10	H	W			
11	H	W	s	d	d
12	H	W	d		

Compare variances of systematic sample one person of five with simple random sample of 20% of total number of inhabitants for

- proportion of children;
- proportion of men.



Co-funded by the
Erasmus+ Programme
of the European Union



Part 4. Stratified random sampling

4.1. How to Draw a Stratified Random Sample

4.2. Estimation of a Population Mean, Total and Proportion

4.3. Selecting the Sample Size for Estimating Population Means and Totals

4.4. Allocation of the Sample

4.5. An Optimal Rule for Choosing Strata

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Example. Construct a data set with 3835 observations:

Discipline	Membership	Number Mailed	Valid Returns	Female Members(%)
Literature	9,100	915	636	38
Classics	1,950	633	451	27
Philosophy	5,500	658	481	18
History	10,850	855	611	19
Linguistics	2,100	667	493	36
Political Science	5,500	833	575	13
Sociology	9,000	824	588	26
Totals	44,000	5,385	3,835	

Include three columns:

column 1 is the stratum number (from 1 to 7), column 2 contains the response variable of gender (0 for males and 1 for females), and column 3 contains the sampling weight N_h/n_h for each observation. Using columns 2 and 3 calculate \hat{p}_{str} . Is it possible to calculate $SE(\hat{p}_{str})$ by using only columns 2 and 3, with no additional information? Explain.



Co-funded by the
Erasmus+ Programme
of the European Union



Part 5. Cluster sampling

5.1. How to Draw a Cluster Sample

5.2. Estimation of a Population Mean and Total.
Equal Cluster Sizes

5.3. Comparison to Simple Random Sampling
Selecting the Sample Size for Estimating
Population Means and Totals

5.4. Sampling of Unequal-Sized Clusters

5.5. Selecting the Sample Size for Cluster Sampling

5.6. Two-Stage Cluster Sample

- Questions for self-examination

- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Part 6. Ratio, regression and difference estimation.

6.1. Ratio Estimation Using Simple Random Sampling. Selecting the Sample Size. Ratio Estimation in Stratified Random Sampling.

6.2. Regression Estimation. Difference Estimation. Relative Efficiency of Estimators.

6.3. Two-Stage Cluster Sampling with Probabilities Proportional to Size.

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Part 7. Survey sampling in gender research.

7.1. International gender studies.

7.2. Survey of different gender aspects.

- Questions for self-examination
- Tasks



Co-funded by the
Erasmus+ Programme
of the European Union



Main problems of gender sampling

- it is impossible to create a population frame
so
- nonprobability sampling would be created
 - spontaneous (haphazard) sampling
 - snow-ball sampling
 - quota sampling
 - judgemental or expert choice sample



Co-funded by the
Erasmus+ Programme
of the European Union



Example of spontaneous sampling

Strategic transformation team. Gender impact survey. Interim Report. (Submitted by The Center for Gender and Development Studies The University of the West Indies, May, 2007, 68 pages)

While the sample is large enough to be considered representative, it was hoped that a full population survey be completed. Consequently, because the population completing the survey was self selected and not randomly selected by the research team, findings may not be applicable to the overall UWI Mona population.

Several reasons were given by staff members for not completing the instrument. These included:

- a. A general indifference to the relevance and usefulness of the exercise
- b. The complexity and length of the instrument, which was considered discouraging
- c. Irrelevance of gender concerns to them (Particularly male respondents)
- d. Concerns about the confidentiality of the responses recorded
- e. Insufficient time to complete the instrument



Co-funded by the
Erasmus+ Programme
of the European Union



Example of mixed-method sampling

Bauer, G. R., Braimoh, J., Scheim, A. I., \ Dharma, C. (2017)

Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. PLoS ONE, 12(5), e0178043

<http://doi.org/10.1371/journal.pone.0178043>



Co-funded by the
Erasmus+ Programme
of the European Union



Recruitment strategy, data collection and study samples

Survey data were collected between October 2015 and March 2016 using an online questionnaire.

After first reading the letter of information, participants indicated their consent to participate by clicking on a button. Anyone who was 14 years old or older, lived in Canada, and was able to complete an English-language survey was eligible to participate. Consent from a parent or guardian was not required for minors. Methods for this study were approved by the Research Ethics Board at The University of Western Ontario. The study was promoted through Facebook ads, Facebook postings, and e-mails to listservs chosen to generate a diverse sample with high frequencies of sexual and gender minorities; **588 individuals completed the first of two randomly assigned sets of survey items** (approximately 5 ± 10 minutes) and provided contact information for follow-up. These participants were contacted via e-mail one to three weeks after participation, with an invitation to either complete the remaining measures (3 minutes) or participate in both the follow-up survey and an immediate semi-structured cognitive interview. Interview participants received a \$50 gift card as an honorarium; no honorarium was provided for survey participation.



Co-funded by the
Erasmus+ Programme
of the European Union



II stage

Interview participants were selected to maximize demographic variation with regard to transgender status, sexual orientation, age, province of residence, immigration status, linguistic background, Indigenous identity, race/ethnicity, education, religiosity, and religious affiliation. Interviews were conducted via telephone or Skype by two of the authors (JB, CD) and were audio recorded; they ranged in length from 24 to 81 minutes, and covered measurement of race/ethnicity and sexual orientation as well as sex/gender.

This mixed-methods analysis is based on participants who completed both surveys ($n = 311$) and the subgroup who participated in individual interviews ($n = 79$).



Co-funded by the
Erasmus+ Programme
of the European Union



Example of sample

National Transgender Discrimination Survey Report on health and health care

Findings of a Study by the National Center for Transgender Equality and the National Gay and Lesbian Task Force By Jaime M. Grant, Ph.D., Lisa A. Mottet, J.D., and Justin Tanis, D.Min. With Jody L. Herman, Ph.D., Jack Harrison, and Mara Keisling October 2010

http://www.thetaskforce.org/static_html/downloads/resources_and_tools/ntds_report_on_health.pdf



Co-funded by the
Erasmus+ Programme
of the European Union



Some references for Part 7

Sampling tools for Multiple Indicator Cluster Surveys (MICS) <http://mics.unicef.org/tools>

Sex and Gender Diversity Among Transgender Persons in Ontario, Canada: Results From a Respondent-Driven Sampling Survey. / [Ayden I. Scheim](#) [Greta R. Bauer](#) //Journal of Sex Research, 2015 Jan 2; 52(1): 1–14. – Published online 2014 Apr 21. – <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4299544/>



Co-funded by the
Erasmus+ Programme
of the European Union



Some references for Part 7

Fox, M. F. Scientists Gender, Family Characteristics, and Publication Productivity among scientists.// Social Studies of Science, 35/1 (February 2005). – 131-150 pp.
– <http://journals.sagepub.com/doi/pdf/10.1177/0306312705046630>

Does Gender Influence Online Survey Participation? –
<http://files.eric.ed.gov/fulltext/ED501717.pdf>



Co-funded by the
Erasmus+ Programme
of the European Union



Example of task for Part 7

The university's rector, in which $N = 4,000$ students study, instructed the assistant to conduct a survey to find out if students were satisfied with life in the campus. The table shows the distribution of students by sex and course.

	Course				
Gender	1	2	3	4	Total
Female	700	520	500	480	2200
Male	560	460	400	380	1800
Total	1260	980	900	860	4000

The assistant should form a probabilistic sample of $n = 200$ and spread the results to the general population.



Co-funded by the
Erasmus+ Programme
of the European Union



a) What type of sample can be used if the students' lists are organized in alphabetical order?

b) what are the advantages of simple random sampling during this survey?

c) what are the benefits of systematic sampling?

d) which type of sample should be selected if the student lists are based on the alphabetical order in eight sheets in accordance with the article and the course as shown in the table?

e) let each of the registered 4,000 students live in one of the 20 hostels. Each dormitory has four floors, and each floor has 50 beds, that is 200 students in the hostel. The administration is trying to bring together students of one sex who are studying on the same course on each floor. What kind of sample should be formed if the distribution of students on dormitories and floors is the basis?



Co-funded by the
Erasmus+ Programme
of the European Union



Thank you!



Co-funded by the
Erasmus+ Programme
of the European Union

