

식생활 조사 방법 및 통계 기법  
(survey 방법을 중심으로)

오세영 교수 (경희대학교)

## INTRODUCTION

- 1 Survey 연구는 사람으로부터 정보를 수집하는데 이용됨
- 2 주로 설문지 (mailed or self-administered)나 면담(face-to-face or telephone)을 통해 정보 수집

- ☛ 설문지나 면담을 이용한 연구시 고려해야 할 사항들
- ① questions: fixed choice vs open ended items
- ② instructions
- ③ sampling and design
- ④ data processing and analysis: You must think ahead to how you plan to analyze the survey's data. (i.e. Compare groups?, Look for relationships?)
- ⑤ pilot testing
- ⑥ response rate: "higher is better"
- ⑦ reporting results

- 3 Criteria for selecting among mailed, self-administered questionnaires and face-to-face and telephone interviews

- ① Reliability and validity
- ② Usefulness or credibility of results
- ③ Cost
- ④ Anonymity
- ⑤ Convenience
- ⑥ Complexity of Information

## THE SURVEY FORM: questions, scales and appearance

- ① The content is the message
- ② Define terms: operationalization (concepts vs variables)
- ③ Select your information needs or hypotheses
- ④ Make sure you can get the information you need.
- ⑤ Do not ask for information unless you can act on it
- ⑥ Writing items: open- and closed-ended questions (Should consider how to organize survey items)

### 1. Rules for writing survey items with forced choices

- . Each question should be meaningful to respondents
- . Use standard language
- . Make questions concrete
- . Avoid biased words and phrases
- . Check your own biases

- . Do not get to personal
  - . Each question should have just one thought
2. Types of responses for survey items with forced choices

- 1) Yes and No answers (cf. don't know)
- 2) Checklists (ex. health food list)
- 3) Rating scales
  - . Nominal (ex. gender)
  - . Ordinal (ex. level of education)
  - . Interval (ex. income)
  - . Ratio (ex. weight, height \*has a true zero)

4) Additive scales (Probably just for the experts)

. Refers to the way the responses are organized for an individual question and to a score that represents a person's view on many items.

- (1) Summative scale (i.e. Likert scale)
  - \* Assign a numerical weight to each response category)
- (2) Cumulative scale (i.e. Guttman scale)

**THE SURVEY IS PUT ON TRIAL**

☛ Checklist

- . Will the survey provide the needed information? Are certain words or questions redundant or misleading?
- . Are the questions appropriate for the people who will be surveyed?
- . Will information collectors be able to use survey forms properly? Can they administer, collect and report information using any written directions or special coding forms?
- . Are the procedures standardized? Can everyone collect information in the same way?
- . How consistent is the information obtained by the survey?
- . How accurate is the information obtained from the survey?

1. Reliability and validity: The quality of your survey

Reliable instrument if it yields the same results every time it is used to measure the same object, assuming the object itself has not changed.

Valid instrument if it provides accurate measure.

\* A valid survey is always reliable one, but reliable one is not always valid.

\* *Conceptual framework for evaluating the adequacy of a research design*  
(by Cook and Campbell Quasi-Experimentation, 1979)

1. Statistical conclusion validity
 

The extent to which the data permit the investigator to conclude that a statistical relationship does or does not exist, regardless of what the relationship may mean.

Threats: insufficient numbers of subjects, low reliability of assessments of independent or dependent variables or both, high sampling error due to heterogeneity of sample or (lack of uniformity in administration of treatment or tests)

## 2. Internal validity

The extent to which significant relationship between independent (i.e. malnutrition) and dependent (i.e. behavior) can be interpreted specifically as a nutrition (independent variable) effect.

Threats: selection effect, statistical regression, history, maturation, subject mortality, instrumentation, -- less likely by the use of random assignment of subjects to treatment and the use of one or more appropriate control groups  
diffusion or imitation of treatment, compensatory equalization of treatment, compensatory rivalry, resentful demoralization

## 3. External validity

The extent to which the results may be generalized to different samples, settings and time period.

Threats: setting-, population-treatment interactions

## 4. Construct validity

The conceptual validity of the interpretations of findings:

the adequacy of the operationalizations of variables, appropriateness of the inferences about the relationships between variables.

Threats: monoperation-, monomethod biases, the interaction of treatment with prior testing, subject or investigator biases, investigator's failure to fully consider the theoretical implications of the research.

## 2. Ensuring quality : Selecting ready-to use surveys

1) Check the results, and ask if the survey has proven its reliability

**Stability:** Computed by administering a survey to the same group on two different occasions and then correlating the scores from one time to the next (i.e. test/retest reliability)

2) If it has more than one form, ask if the forms are equivalent.

**Equivalent:** Computed by giving two or more forms of the same survey to the same groups of people on the same day, or by giving different forms of the survey to two or more groups that have been randomly selected.

3) If you are concerned with respondents' scores on several items combined, check that the survey is homogeneous or internally consistent.

**Homogeneity:** Computed by dividing the survey into two parts and correlate the scores on one half with the scores on the other half (i.e. split-half reliability)

## Some questions to ask about a published survey's validity

**Predictive validity:** a measure of the survey's ability to forecast performance

**Concurrent validity:** means that the survey and some other measures agree

Validated by comparing a survey against a known and accepted measure.

Content validity: the accuracy with which the questions represent the characteristics they are supposed to survey

Usually established by asking experts whether the items are representative samples of the attitudes and traits you want to survey

Construct validity: Experimentally obtained proof that a survey that is intended to measure something (i.e. a specific attitude or behavior) truly measures it.

Established experimentally by trying the survey on people whom the experts say do and do not exhibit the behavior associated with the construct.

### 3. Guidelines for pilot testing

- 1) Try to anticipate the actual circumstances in which the survey will be conducted and make plans to handle them.
- 2) You can start by trying out selected portions of the survey in very informal fashion.
- 3) Choose respondents similar to the ones who will eventually complete the survey.
- 4) Enlist as many people in the trial as seems reasonable without wasting your resources.
- 5) For reliability, focus on the clarity of the questions and the general format of the survey.
- 6) For validity, make sure that all relevant topics have been included in the survey.
- 7) Test your ability to get a range of responses.

### 4. Ethic, Privacy and Confidentiality

## **SAMPLING**

### 1. Sample size and response rate: who and how many

Should consider the followings

- 1) How quickly are the data needed?
- 2) What type of survey is planned?
- 3) What are your resources?
- 4) How credible will your findings be?
- 5) How familiar are you with sampling methods?

### 2. Probability sampling methods

#### 1) Simple random sampling

. Each person has an equal chance of being selected for participation in a survey

. Simplest of all sampling methods and easiest to conduct

. Produces greater errors in the results (greater "standard errors") than do other sampling methods; Cannot be used if you want to break respondents into subgroups or strata.

#### 2) Stratified random sampling (ex. male 40%, female 60%)

. Can be more precise than simple random sampling (more homogeneous groups).

. Requires more efforts and larger sample size

#### 3) Simple random cluster sampling (ex. Guatemala study)

. Can be used when it is inconvenient or unethical to randomly select individuals; administratively simple.

. Not mathematically efficient

4) Narrowing the margin of error: Minimize "standard error of the mean(SEM)"

3. Nonprobability sampling

1) Systematic sampling

ex. Select every fifth name on a list

. bias?

2) Accidental samples

. Select ones who are available.

. convenient, but may be a very biased.

3) Purposive samples

4. How large should your sample be

1) Statistical methods

Sampling errors (SEM), Stratification, Confidence Levels

5. Response rate: higher is better

\* Need to prove that the loss of data from nonrespondents do not harm or bias the survey's findings.

## **SURVEY DESIGN: Environmental Control**

When?, how often?, how many groups?

1. Cross-sectional survey design :Data are collected at a single point in time.

Question: If only one program was available, which would you choose?

Sample: A cross section of 500 people, randomly selected, who attended an educational program this year.

2. Longitudinal Surveys

1) Trend designs: surveys a particular groups over time

Question: What do participants know about something?

Sample: Random samples of 500 participants attending Dine-out in 1990, 500 in 1993, and 500 in 1996. (A general population, but different persons in each survey)

2) Cohort designs: survey a particular group over time, but the people in the group may vary.

Question: How have attitude toward diet changed since 1970?

Samples: A different random sample of participants attending a nutrition program in 1970 is surveyed every five years.

3) Panel designs: collect data from the same sample over time.

3. Comparison group survey designs:

Quasi-(assignment is not random) and true experiments

\* True-experiment and longitudinal design is among the most sophisticated and will enable you to make sound inferences.

4. Other survey designs

1) Normative survey designs: Compared to 'norm' or 'model' group.

- . Two groups are compared, but only one is actually survey.
- 2) Case control designs
  - . Generally used for testing a specific hypothesis
  - . Need two groups, and so, a major concern is their selection.

## **ANALYZING DATA FROM SURVEYS**

1. Descriptive statistics
  - . mean, mode, median, numbers, percentage, range, standard deviation
  - . PROC MEANS, PROC UNIVARIATE in SAS
2. Correlations
  - . Spearman rank-order (ordinal scale), Pearson product-moment(interval or ratio scale)
  - . PROC CORR in SAS
3. Comparisons
  - . Mann-Whitney U (PROC NPARIWAY, ordinal scale), chi-square (PROC FREQ, nominal scale), t-test (PROC TTEST, interval or ratio scale), analysis of variance (PROC ANOVA, PROC GLM)
4. Trends
  - . repeated measures analysis of variance (PROC ANOVA or GLM with REPEATED option)
  - McNemar test(PROC FREQ with AGREE option)

### Assessment of household organization and structure

#### 1. Component of the basic household interview

##### 1) household composition

Format:name,relationship to household head, sex, age, education, occupation/employment

\* The problem of non-resident members

##### 2)Material resources

Economic resources related to dietary patterns and nutritional status

### Assessment of dietary intake

24hr recall, food record, food frequency questionnaire, diet history

usual intake, intra- and inter-individual variation

## **References**

1. Babbie ER. Survey Research Methods. Belmont, California: Wadworth Publishing Company, 1973.
2. Bernard HR. Research Methods in Anthropology: qualitative and quantitative approaches. 2nd ed. Thousand Oaks: Sage Publications, 1994.
3. Cody RP, Smith JK. Applied Statistics and the SAS Programming Language, 3rd ed. New York: North-Holland, 1991.
4. Cook TD, Campbell DT. Quasi-Experimentation: Design and analysis issues for field settings. Chicago:Rand McNally, 1979.
5. Fink A, Kosecoff J. How to conduct surveys: A step-by-step guide. Newbury Park: Sage Publications, 1985.

6. Pedhazur EJ, Schmelkin LP. Measurement, Design, and Analysis:An integrated approach Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1991.
7. SAS Institute. SAS/STAT User's Guide, Version 6.0.
8. Willett W. Nutritional Epidemiology. New York:Oxford University Press, 1990.