



MATH 166: Statistics and Probability

Dr. Eric Nimako Aidoo

en.aidoo@yahoo.com

19-01-2018

Data Collection

➤ General Introduction

- ✓ Scientific problems can be solved only on the basis of accurate and reliable **data**.
- ✓ **Data** are values that a variable can assume
- ✓ It is therefore the responsibility of the researcher or investigator to set up a research design including appropriate field strategies in order to obtain the required data to enable him find answers to research questions or problem at hand.
- ✓ Good data is expensive and takes a lot of time to collect.
- ✓ It is only good data collection technique that can collect data which can reflect the whole population. Designing ways to collect good data is an important job in statistical studies.

Data Collection

➤ Definition

- ✓ Data collection is the process of obtaining relevant information on subjects/units of interest in a study.
- Data can be collected in a variety of ways. It is the most important part of statistical procedure because valid conclusions can only result from data which has been collected properly.
- If proper procedure is used to collect data then the issue of representativeness can be guaranteed. The use of wrong or faulty data collection methods would result in wrong conclusions because no good statistical tool can produce good results from wrongly collected data.

Source of Data

- There are two broad sources of data, these are:
 - ✓ Primary source of data
 - ✓ Secondary source of data
- **Primary sources of Data:** This refers to data collected expressly for a specific purpose e.g. the Collection of fact and figures relating to the population in a census. One important advantage of obtaining a primary data is that the exact information wanted is obtained.
- **Secondary Data Sources:** This applies to data already collected and compiled for some other purpose, most importantly for administrative usage. For example, school drop-out rates at the basic school level are compiled by the Ghana Education Service for the entire country. This data is obtained purposively for administrative use. It constitutes a secondary source of data if a researcher wants to use it for his/her studies.

Methods Data Collection

- There are many methods of data collection but the most popular one is called **survey**.
- Survey is an activity that collects data about characteristics of interest from some or all units of a population using well-defined concept, methods and procedures.
- Survey seeking for data on an attribute of a population using all the elements in the population is called **census** (or complete survey).
- Survey seeking for data on an attribute of a population using a portion of the elements in the population is called **sampled survey**.
- Survey data collection is conducted using an instrument called **questionnaire**

The Questionnaire

- **Questionnaire** is a structured document which seeks to collect facts or opinion or both from an individual or group of individuals. It provides researchers a standard procedure for collecting primary data that are comparable irrespective of who collect them.
- Researchers who plan to use questionnaires usually start by writing the questions themselves. After a rough draft is created, the researcher then analyzes their questions to see which ones are related to their variables list. The variables list contains the key concepts or theoretical constructs that are contained in the research question and/or hypotheses. Care is taken to ensure that questions cover every concept.

Structure of a Questionnaire

- Questionnaire whether administered personally, by post or by electronic mail must comply with certain standards generally accepted in modern day research. The questionnaire must as a matter of necessity include three important elements making it structure.
- These include:
 - ✓ Cover letter
 - ✓ Instruction
 - ✓ Main body

Structure of a Questionnaire

➤ Instruction:

- ✓ The Instruction about how to fill the questionnaires is normally given in the cover letter and /or in the questionnaires itself. Instruction provide guidelines as to how to respond to the questions, that all questions should be answered, that there are no right or wrong answers, that all completed questionnaires should be sent to the researcher by a certain date etc.
- Instructions are expected to be written in an easy language. Inadequate instructions are one of the major reasons of non-response and should be avoided.

Structure of a Questionnaire

➤ The main body

- ✓ This includes the questions that are to be answered by the respondents. The questions must be written down clearly with easy language and the right format, need to be adopted in the construction of question. This is important because this is the part of the questionnaire that allows the researcher to collect the data required for the study
- The choice of the questionnaire in preference to other data collection methods is generally a matter of weighting its advantage and disadvantages against the other methods. Questionnaires, however, are unique in terms of structure and place in research. They provide a type and amount of information that other methods cannot provide. Among its specific advantage are;

Sample of a Questionnaire

CSIR-BUILDING AND ROAD RESEARCH INSTITUTE, FUMESUA, GHANA

QUESTIONNAIRE ON:

Commuter's Satisfaction with Public Transport Services in Ghana: A Case Study of Kumasi – Accra Route

Your responses will be used purely for research purpose; hence confidentiality and anonymity are assured.

INSTRUCTIONS:

Please read the following questions carefully – Tick [✓] appropriate answer in the boxes provided and specify where necessary. Thank you

1. Gender

[1] Female

[2] Male

2. What is your age category?

[1] Below 18 years

[2] 18 – 30 years

[2] 31 – 40 years

[3] 41 – 50 years

[4] Above 50

Merits of using a Questionnaire

- The respondent does not have to make instantaneous decisions as in the case of personal interview.
- It enhanced the accuracy of recording.
- It is useful where the respondents must remain anonymous.
- They produce quick results cutting down the time used for the study.
- The questionnaire is a convenient method; they can be completed at the respondent's convenience.
- They help to avoid bias or errors caused by the presence or attitudes of the interviewer.
- Questionnaires are stable, consistent and uniform measure without variation.
- The use of questionnaires promises to wider coverage, since they can approach respondents more easily than other methods.

Demerits of using a Questionnaire

- Low response rate.
- Due to lack of supervision, partial response is quite possible.
- The identity of the respondent and the conditions under which the questionnaires was answered are not known.
- Respondents who find problem on/or ambiguities are not readily able to rectify the situation.
- Illiteracy problem.
- Questionnaires do not provide an opportunity to collect additional information while they are being completed. There is no researcher present for instance, to make observation while the questions are being answered.
- Questionnaires do not permit probing, prompting an clarification of question especially when a respondent is in a difficulty.

Questionnaire Designing

➤ *Five Important steps in Questionnaire Designing*

- ✓ Determine the questions to ask.
- ✓ Selecting question type.
- ✓ Determining question wording
- ✓ Determining sequencing of questions
- ✓ Determining on overall questionnaire layout.

Determine the Questions to Ask

- When determining questions for your questionnaire, it is necessary to ask you self on number of questions such as
 - **Is the question necessary?** – i.e. does the question contribute to the objectives of the study?
 - ✓ Basic questions
 - ✓ Classification questions
 - ✓ Identification questions
 - **How many questions do I need to cover on a particular issue?**
 - **Does the question demand a lot of effort to answer?**
 - **Is the question sensitive?**

Selecting Question Type

- The type of questions on the questionnaire has been grouped in two main types. These are open questions (item) and closed questions (item). Every questionnaire may take one of this form or both.
- **Open questions:** are questions that respondents answer in their own words. No alternative answers are provided for the respondent to select from.
- **Closed questions:** are question with a series of alternative answers for the respondents to select from.

Selecting Question Type

➤ Advantages of using open ended questions

- ✓ They are useful in exploratory in pilot survey.
- ✓ It provides the opportunity for the respondents to express selves freely.
- ✓ The researcher gets the right information.
- ✓ Answer provided by the respondent may provide an insight to the researcher on some information that the researcher has not thought off.

➤ Disadvantages of using open ended questions

- ✓ It takes longer time to complete.
- ✓ Less articulate respondents find them more difficult to answer.
- ✓ It results in few responses.
- ✓ The analysis may be affected

Selecting Question Type

➤ Advantages of using closed ended questions

- ✓ Responses are early to analyze.
- ✓ It takes shorter time to complete.
- ✓ Specific information is obtained.
- ✓ The answers are easier to process.
- ✓ All the respondent answer to same question using the same set of alternative answers.

➤ Disadvantages of using closed ended questions

- ✓ Preparing the alternative answers is time consuming.
- ✓ Respondents are restricted because there may not be adequate information.
- ✓ The respondent may be force to choose an answer even though that may not be his/her answer.

Selecting Question Types

➤ Classification of closed ended questions

Multiple-choice question: In this type of question bias a number of alternative answer are provided and the respondent i.e. ask to select one or more from of this alternatives.

Dichotomous question: This type of question has only two alternative responses. Eg. Yes or No, agree, or disagree, accept or don't accept.

Ranking question: In this question the respondent followed to select two or more alternative as answer to a single question and rank those answers in order of increasing or decreasing imports.

Rating question (Likert scale): This increasing the level of importance attaches to each attribute separately using the scale of measurement.

Determining Question Wording

- Use simple vocabulary
- Ask precise questions
- Avoid technical words or jargons
- Avoid ambiguous words
- Avoid leading questions

Determining sequencing of questions

- This is the ordering of question in a logical sequence in the questionnaire. The key factors being the subject matter, the sensitivity of the various questions and the difficulty level of the questions.
- The following table may be useful

Question	identification	classification	Basic
Easy	1	1	2
Difficult	2	2	3
Sensitive	3	3	3

Questionnaire Layout

- Name of the organization conducting the survey
- A short but self explanatory title of the survey
- This authority or bases for the survey.
- Serial number of questionnaire (identification mark)
- Print questions on good quality paper.
- Font style for instructions, question and for specific question must be different.
- Number the question especially when branching questions are used.
- The questionnaire should have in professional appearance.

Piloting the Questionnaire

- A new questionnaire must first be tested on a section of the target population. The information obtained from pilot survey may suggest rephrasing the questions in order to attract more appropriate responses.

Questionnaire Administration

- There are various ways of administering questionnaires but the most commonly used methods are:
 - ✓ By post.
 - ✓ By personal (face-to-face) interview.
 - ✓ By telephone interview
 - ✓ By electronic medium (email or website)

Questionnaire administration by post

- **Questionnaire administration by post:** this involves sending a copy of the questionnaire to the selected respondent through mail (by post) and asking them to complete and return it to you through the same means.
- **Merits**
 - ✓ It allows respondent to consult relevant documents for more information.
 - ✓ There is no interviewer's bias
 - ✓ It is a good method to use in research which involves questions bordering on personal and sensitive issues
 - ✓ Cost effective, since there is no need to hire enumerators.

Questionnaire administration by post

➤ Demerits

- ✓ The research has no opportunity to probe respondents for further details.
- ✓ The respondent is unable to seek clarifications or further help from the researcher in completing the questionnaire
- ✓ Low response rate.

Personal (face-to-face) Interview

- **Personal (face-to-face) interview:** this method involves the researcher interviewing the respondents and obtaining the information from them face to face. Usually the researcher poses the questions one by one as written on the questionnaire and fill in the appropriate answers as maybe given by respondent.
- **Merits**
 - ✓ There is high response rate
 - ✓ Timely response
 - ✓ Reliability of response
 - ✓ Further clarification can be made by both the respondents and the interviewers.
 - ✓ Completeness of the interviewers also guaranteed. The fact that the interviewer presents the questions guarantees that all questions will be attempted.

Personal (face-to-face) Interview

➤ Demerits

- In spite of the numerous advantages of the interview method elaborated above, interviewing is limited by some factors that cannot be overlooked.
 - ✓ Instantaneous decisions are made by the respondents.
 - ✓ The method is expensive (in terms of interviewers training and allowance)
 - ✓ It suffer from both respondent and interviewers bias.
 - ✓ It may not be applicable in certain situations (sensitive issues).
 - ✓ It offers less anonymity than other method since the interviewer knows the identity, residence, type of housing, family conditions and other personal details of the respondent.

Telephone Interview

- **Telephone Interview:** this method involves the researcher calling the respondents on the telephone, reading out the questions to them and recording their responses.
- **Merit**
 - ✓ It is relatively economical.
 - ✓ Direct contact over the telephone enables the interviewer to clarify any issue of understanding as the interview procedure.
 - ✓ There is uniformity of approach if one interviewer is involved.
 - ✓ None responds is not a problem since refusals are ignored and another unit selected.
 - ✓ It produces quick answers
 - ✓ It offers more anonymity than other techniques particularly when random digit dialing techniques are used.

Telephone Interview

➤ Demerit

- ✓ Low response rate.
- ✓ It can cause an aggravation if the interviewer is not skillful.
- ✓ Interviewer influence.
- ✓ It has an inbuilt bias since only people with telephones can be
- ✓ It is not possible to control the interview completely
- ✓ It may not be applicable in certain situations

Electronic medium (e-mail or website)

- **Electronic medium:** this method involves the researcher sending a copy of the questionnaire to the selected respondent as an attachment via email for completion or placing the questionnaire on a website and then inviting and directing the potential respondent to complete it.

- **Merit**
 - ✓ It is relatively economical.
 - ✓ None responds is not a problem since refusals are ignored and another unit selected.
 - ✓ It produces quick answers

Electronic medium (e-mail or website)

➤ Demerit

- ✓ Low response rate.
- ✓ It has an inbuilt bias since only people with email or access to internet can take part
- ✓ It is not possible to control the process completely
- ✓ It may not be applicable in certain situations