IBP 1203 Business Statistics

Sampling Methods

Learning Objectives:

At the end of this chapter the students will be able to:

1. Describe the meaning and usage of statistics,

2. Identify the types of survey methods,

3. Select appropriate methods for sampling methods and data collection,

4. Describe framework and limitation of survey research and

5. Work cooperatively and effectively with team member.

Survey Method

A population and sample are two basic concepts of statistics. A population is the set of measurements corresponding to the entire collection of units for which inferences are to be made. Where as a sample is the part of the population from which information is collected.
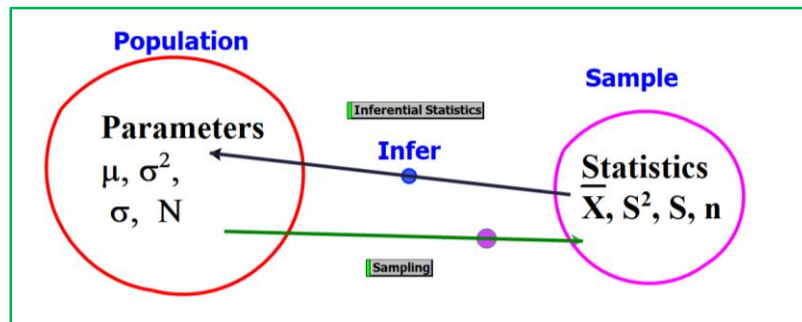
■ Surveys

A survey is a non-experimental, descriptive research method. A survey is a data collection tool used to gather information about individuals. A survey is a kind of a medium which is used by organizations or companies to gather certain data or information about something in particular.

■ Types of Surveys Methods

A survey is one of the most effective methods to get feedback from many different people. To conduct a survey, the researchers or an organization can adopt different methods or techniques. There are many varied types of survey methods and each one is unique. Surveys can be used qualitative data or quantitative data. Data are usually collected through the use of questionnaires, sometimes researchers directly interview subjects.

There are two basic types of surveys:
- cross-sectional surveys and
- longitudinal surveys.

Some of the most common ways to administer survey are as follows:

- *Telephone* - An example of a telephone survey would be a market research call about your experiences with a certain consumer product.

- *Post Mail* - An example might include an alumni survey distributed via direct mail by your alma mater.

- *Online* - Online surveys might focus on your experience with a particular retailer, product or website.

- *At home interviews* - The Census is a good example of an at-home interview survey administration.

Framework and Limitation of Survey

The frameworks of survey research are as following:

1) Rules and reason of survey research.

   The reason for conducting a sample survey is to estimate the value of some attribute of a population.
   - *Population parameter*. A population parameter is the true value of a population attribute.
   - *Sample statistic*. A sample statistic is an estimate, based on sample data, of a population parameter.

2) Quality of survey research and sample statistic values.

   The quality of survey research and sample statistic values such as accuracy, precision, and representativeness are strongly affected by the way that sample observations are chosen.
   The importance questions before conducting survey research are:

   - What is the sampling method to be used?
   - What is the sample size of the sample survey?
   - How are the surveys conducted?
   - How is the sample selected for a survey?
   - How will the sample unit be chosen?
   - What is the budget of conducting sample survey research?
   - What is the duration of conducting sample survey research?
   - What mode will be used to pose questions and collect answers from respondents?

   3) Concern about the *advantages* and *disadvantages* of using surveys research

   ● Advantages of using surveys research are:
   - researchers can collect a large amount of data in a relatively short period,
   - less expensive than many other data collection techniques,
   - surveys are easy to administer,
   - surveys are simple to score and code.
   - surveys determine the values and relations of variables and constructs,
   - responses can be generalized to other members of the population studied and often to other similar populations,
   - surveys research can be used to collect information on a broad range of things, including personal facts, and attitudes,
   - surveys can be reused easily, and provide an objective way of comparing responses over different groups, times, and places,
   - surveys can be used to predict behavior, and
   - surveys can help confirm and quantify the findings of qualitative research.

● Disadvantages of using surveys research are:
  ▪ Surveys are just a snapshot of behavior at one place and time,
  ▪ The researcher has to be careful about assuming the validity in different contexts. In particular, different cultures may produce different results,
  ▪ Poor survey construction and administration can undermine otherwise well-designed studies,
  ▪ The answer choices provided on a survey may not be an accurate reflection of how the participants actually feels, and
  ▪ The response rates can bias the results of a survey.

4) Consider of the survey process and preparation of a check list
  ▪ determination of the unit of analysis such as the individual, group, or organization,
  ▪ construct multiple choice questions and use a wider range of question types,
  ▪ creation and use of multi-item scales,
  ▪ have to do the pre-testing and use of pilot data assessment of both content validity, and assessment of reliability,
  ▪ random sampling from a defined sample frame,
  ▪ determination of an appropriate response rate and evaluation of nonresponse bias assessment of whether significant correlations imply real causal relations.

A Poll:
- A poll is a human research survey of public opinion from a particular sample.
- A poll is used to ask one simple question, while a survey is generally used to ask a wide range of questions.
- Surveys are for making decisions and getting insights, polls are for taking a snapshot.
- Polls give you results right-away, with the surveys, you have to wait for the analysis.

1. **Do you have a big decision to make?**
   - **No.** Use *a poll* to get a snapshot or pulse of how people view a topic.
   - **Yes.** Use *a survey* to gather quantitative or statistical data that predicts how people will act.

2. **Do you need to control the respondents?**
   - **Yes**. Use *a survey*. You can ask qualifying questions and only allow those respondents who fit the criteria to finish the survey.
   - **No**. Polls are not as good as survey at describing specific respondent attributes. Your responses will generally be an overall measure.

Sampling Methods

The quality of survey research and sample statistic values such as accuracy, precision, and representativeness are strongly affected by the way that sample observations are chosen.
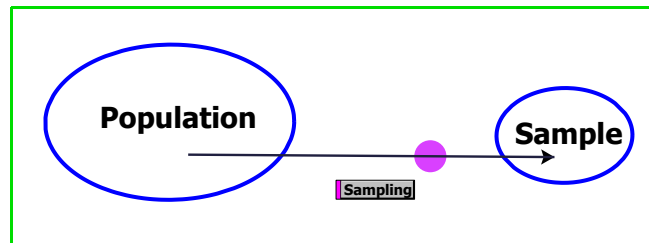
■ Sampling Process

The sampling process comprises several stages:

- Defining the population of concern,
- Specifying a sampling frame, a set of items or events possible to measure,
- Specifying a sampling method for selecting items or events from the frame,
- Determining the sample size,
- Implementing the sampling plan,
- Sampling and data collecting.

■ Sampling Methods

Sampling method refers to the way that observations are selected from a population to be in the sample for a sample survey. The reason for conducting a sample survey is to estimate the value of some attribute of a population.

- **Population parameter**. A population parameter is the true value of a population attribute.
- **Sample statistic**. A sample statistic is an estimate, based on sample data, of a population parameter.

The inference about some specific unknown parameter is based on a statistic. The example of symbols of parameters and statistic values are as in the following table.

	Topic	Symbol	
		Population (Parameter)	Sample (Statistic)
1	Mean	μ	$\bar{x}$
2	Variance	σ^2	S^2 or SD^2
3	Standard Deviation	σ	S or SD
4	Size	N	n

> **Parameter** is to **Population** as **Statistic** is to **Sample**

■ **Probability and Non-Probability Samples Methods**

There are two types of sampling methods:

1. Probability Sample	• Known, non-zero chance of being chosen
2. Non-Probability Sample	• Do not know probability • Might be zero change of selection

The researchers can select sample group and sample unit by using the sampling methods.

The sampling methods consist of two categories:

- **Probability Samples Method**.

  Probability sampling is a sampling technique wherein the samples are gathered in a process that each element in the population has equal chance of being chosen for the sample. Probability sampling is designed to allow extrapolation from a small, highly representative sample, to a larger population.

  Probability sample method will be used when the researcher wants to answer the "**where**" and "**how many**" questions.

- **Non-probability Samples Method**.

  Non-probability sampling is a sampling technique where the samples are gathered in a process that we do not know the probability that each population element will be chosen. Each element in the population has no equal chances of being selected.

  Non-probability sample method will be used when the researcher wants to answer the "**how**" and "**why**" questions.

	Probability Sample	Non-Probability Sample
Convenience	-	✔
Cost	✔	✔
Representative	✔	-
Analysis	✔	-

■ **Probability Samples Methods**

The main types of probability sampling methods are simple random sampling, stratified sampling, cluster sampling, systematic sampling and multistage sampling. The key benefit of probability sampling methods is that they guarantee that the sample chosen is representative of the population. This ensures that the statistical conclusions will be valid. The main types of probability sample methods are as follows.

1. Simple Random Sampling

> A simple random sample (SRS) of size **n** consists of **n** individuals from the population of size **N**, chosen in such a way that every set of **n** individuals has an equal chance to be the sample actually selected.

Simple random sampling refers to any sampling method that has the following properties.

- The population consists of **N** objects.
- The sample consists of **n** objects.
- If all possible samples of **n** objects are equally likely to occur, the sampling method is called simple random sampling.

> Simple Random sampling is a procedure for sampling from a population in which:
>
> - the selection of a sample unit is based on chance, and
> - every element of the population has a known, non-zero probability of being selected.

Simple random sampling is the simplest form of probability sampling. To select a simple random sampling, you need to:
- make a numbered list of all the units in the population from which you want to draw a sample or use an already existing one (sampling frame),
- decide on the size of the sample,
- select the required number of sampling units, using a 'lottery' method or a table of random numbers, and
- draw a unit.

The simple random sample is the basic sampling method assumed in statistical methods and computations. For example,
let's say you have a population of 1,000 people and you wish to choose a simple random sample of 50 people.
1) each person is numbered 1 through 1,000,
2) generate a list of 50 random numbers, using Lottery method, or Random Digits table,
3) the individuals assigned those numbers are the ones you include in the sample.

How to select a random sample:
1) *Lottery method*.

- Each of the **N** population members is assigned a unique number.
- The numbers are placed in a bowl and thoroughly mixed.
- The researcher selects a blind-folded by random **n** numbers.

2) *Random Digits Table*:
   - Close your eyes and choose the random start digit,
   - From the random start digit, choose the digits in the same row or column. These digits are the assigned number for the sample unit,
   - generate a list of the sample size (**n)** random numbers.

Part of the Table of Random Digits are shown as following.

The Table of Random Digits*

|									
00000	10097	32533	76520	13586	34673	54876	80959	09117	39292 74945
00001	37542	04805	64894	74296	24805	24037	20636	10402	00822 91665
00002	08422	68953	19645	09303	23209	02560	15953	34764	35080 33606
00003	99019	02529	09376	70715	38311	31165	88676	74397	04436 27659
00004	12807	99970	80157	36147	64032	36653	98951	16877	12171 76833
00005	66065	74717	34072	76850	36697	36170	65813	39885	11199 29170
00006	31060	10805	45571	82406	35303	42614	86799	07439	23403 09732
00007	85269	77602	02051	65692	68665	74818	73053	85247	18623 88579
00008	63573	32135	05325	47048	90553	57548	28468	28709	83491 25624
00009	73796	45753	03529	64778	35808	34282	60935	20344	35273 88435
00010	98520	17767	14905	68607	22109	40558	60970	93433	50500 73998
00011	11805	05431	39808	27732	50725	68248	29405	24201	52775 67851
00012	83452	99634	06288	98083	13746	70078	18475	40610	68711 77817
00013	88685	40200	86507	58401	36766	67951	90364	76493	29609 11062
00014	99594	67348	87517	64969	91826	08928	93785	61368	23478 34113
00015	65481	17674	17468	50950	58047	76974	73039	57186	40218 16544
00016	80124	35635	17727	08015	45318	22374	21115	78253	14385 53763
00017	74350	99817	77402	77214	43236	00210	45521	64237	96286 02655
00018	69916	26803	66252	29148	36936	87203	76621	13990	94400 56418
00019	09893	20505	14225	68514	46427	56788	96297	78822	54382 14598
00020	91499	14523	68479	27686	46162	83554	94750	89923	37089 20048
00021	80336	94598	26940	36858	70297	34135	53140	33340	42050 82341
00022	44104	81949	85157	47954	32979	26575	57600	40881	22222 06413
00023	12550	73742	11100	02040	12860	74697	96644	89439	28707 25815
00024	63606	49329	16505	34484	40219	52563	43651	77082	07207 31790
00025	61196	90446	26457	47774	51924	33729	65394	59593	42582 60527
00026	15474	45266	95270	79953	59367	83848	82396	10118	33211 59466
00027	94557	28573	67897	54387	54622	44431	91190	42592	92927 45973
00028	42481	16213	97344	08721	16868	48767	03071	12059	25701 46670
00029	23523	78317	73208	89837	68935	91416	26252	29663	05522 82562
00030	04493	52494	75246	33824	45862	51025	61962	79335	65337 12472
00031	00549	97654	64051	88159	96119	63896	54692	82391	23287 29529
00032	35963	15307	26898	09354	33351	35462	77974	50024	90103 39333
00033	59808	08391	45427	26842	83609	49700	13021	24892	78565 20106
00034	46058	85236	01390	92286	77281	44077	93910	83647	70617 42941

* Retrieved from http://research-advisors.com

2. Stratified Random Sampling

Stratified Random Sampling is possible when it makes sense to partition the population into groups based on a factor that may influence the variable that is being measured. These groups are then called **strata**. An individual group is called a **stratum**. With stratified sampling the researcher should:

- partition the population into groups (strata)
- obtain a simple random sample from each group (stratum)
- collect data on each sampling unit that was randomly sampled from each group (stratum)
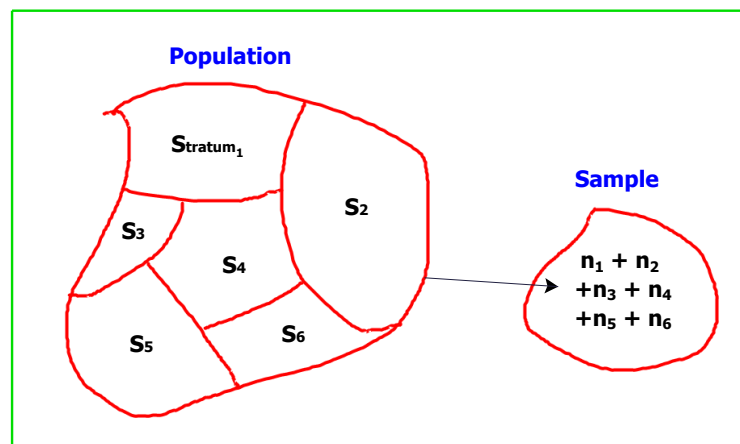
Stratified random sampling refers to a sampling method that has the following properties.

- The population consists of **N** elements,
- The population is divided into **H** groups (called **strata**),
- Random select elements from each stratum and combine into a sample. The researcher obtains a probability sample from each stratum,
- The number of observations within each stratum n_h is known, and

$$\text{Sample size}(\mathbf{n}) = n_1 + n_2 + n_3 + ... + n_H.$$

Stratified sampling works best when a **heterogeneous** population is split into fairly **homogeneous** groups. Under these conditions, stratification generally produces more precise estimates of the population percent than estimates that would be found from a simple random sample.

As a example, suppose we conduct a national survey. We might divide the population into groups or strata, based on geography - north, east, south, and west. Then, within each stratum, we might randomly select survey respondents.

3. Cluster Sampling

Cluster sampling. With cluster sampling, every member of the population is assigned to one, and only one, group. Each group is called
a cluster. A sample of clusters is chosen, using a probability method (often simple random sampling). Only individuals within sampled clusters are surveyed.
The difference between cluster sampling and stratified sampling.

- *With cluster sampling*, the sample includes elements only from sampled clusters.
- *With stratified sampling*, the sample includes elements from *each* stratum.
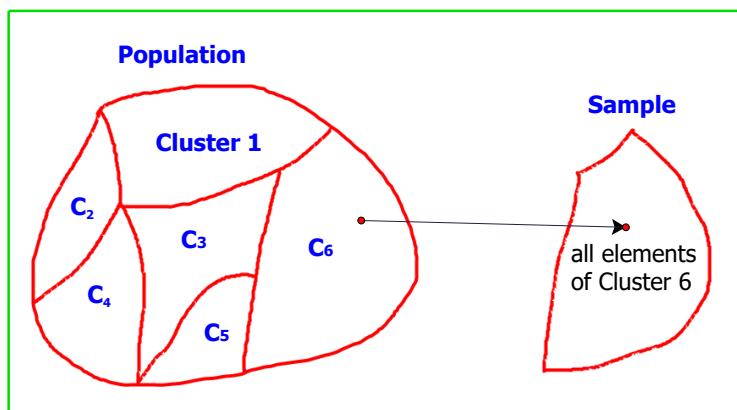
4. Systematic Random Sampling

Systematic random sampling is a type of probability sampling method in which sample members from a larger population are selected according to a *random starting point* (**r**) and a *fixed periodic interval*. This interval, called the *sampling interval* (**k**), is calculated by dividing the population size by the desired sample size.

The steps to create a group from systematic random sampling:

1. Create a list of the population.
2. Calculate a sampling interval (k)

$$\mathbf{k} = \frac{\text{Population size}}{\text{Sample size}}$$

3. Random a number to be a starting point (r); $1 \le \mathbf{r} \le k$
4. Gather a list based on the interval number.

For example, you want to create a systematic random sample of 100 people from a population of 1,000.

sampling interval (k) $= \dfrac{1000}{100} = 10$

let random starting point $= 7$

The sample, then, would be composed of persons numbered

		7,
10 + 7	=	17,
17 + 10	=	27,
27 + 10	=	37,
37 + 10	=	47,

………………………………….

and so on down the line until you have reached 100 smples.

4. Multistage Sampling

Multistage sampling. With multistage sampling, we select a sample by using combinations of different sampling methods.
For example, in Stage 1, we might use cluster sampling to choose clusters from a population. Then, in Stage 2, we might use simple random sampling to select a subset of elements from each chosen cluster for the final sample.

Non- Probability Sampling

Non-probability sampling is a sampling technique where the samples are gathered in a process that does not give all the individuals in the population equal chances of being selected.

In any form of research, true random sampling is always difficult to achieve. Most researchers are bounded by time, money and workforce and because of these limitations, it is almost impossible to randomly sample the entire population and it is often necessary to employ another sampling technique, the non- probability sampling methods.

Subjects in a non-probability sample are usually selected on the basis of their accessibility or by the purposive personal judgment of the researcher.

Some survey research do not know the list of Population for their research,
With the non- probability sampling method:
may be the entire population was not sampled.

This entails that the sample may or may not represent the entire population accurately. Therefore, the results of the research cannot be used in generalizations pertaining to the entire population.

Types of Non-Probability Sampling*

1.   Convenience Sampling

Convenience sampling is probably the most common of all sampling techniques. With convenience sampling, the samples are selected because they are accessible to the researcher. Subjects are chosen simply because they are easy to recruit. This technique is considered easiest, cheapest and least time consuming. To sample friends, co-workers, or shoppers at a single mall, are all examples of convenience sampling. Such samples are biased because researchers may unconsciously approach some kinds of respondents and avoid others and respondents who volunteer for a study may differ in unknown but important ways from others.

2.   Quota Sampling

Quota Sampling is a non-probability sampling technique wherein the researcher ensures equal or proportionate representation of subjects depending on which trait is considered as basis of the quota. For example, if basis of the quota is college year level and the researcher needs equal representation, with a sample size of 100, he must select 25 of 1st year students, another 25 of 2nd year students, 25 of 3rd year and 25 of 4th year students. The bases of the quota are usually age, gender, education, race, religion and socioeconomic status.

3.   Judgment Sampling

Judgment Sampling is more commonly known as purposive sampling. In this type of sampling, subjects are chosen to be part of the sample with a specific purpose in mind. With judgmental sampling, the researcher believes that some subjects are more fit for the research compared to other individuals. This is the reason why they are purposively chosen as subjects.

4.   Snowball sampling

Snowball sampling is usually done when there is a very small population size. In this type of sampling, the researcher asks the initial subject to identify another potential subject who also meets the criteria of the research. The downside of using a snowball sample is that it is hardly representative of the population. Snowball sampling - The first respondent refers an acquaintance. The friend also refers a friend, and so on. Such samples are biased because they give people with more social connections an unknown but higher chance of selection (Berg 2006), but lead to higher response rates.

* Retrieved from https://explorable.com/probability-sampling

Sample Size

When choosing a sample size, we must consider the following issues:

- What population parameters we want to estimate?
- What is the cost of sampling?
- How much is already known?
- Spread (variability) of the population.
- Practicality: how hard is it to collect data?

The sample size needs to be *statistically significant*. This means it is not chosen by chance. The accuracy level of the estimates that the researcher proposes to present and the error margin he would allow for the estimates determines the sample size.

The researcher can employ the formula of Sample Size Determination used by Krejcie & Morgan in their 1970 article "Determining Sample Size for Research Activities, Educational and Psychological Measurement. PP, 607-610. (Retrieved from http://research-advisors.com). The Sample Size Determination Table is shown as follows.

The Sample Size Determination Table*

Pop. size	Confidence = 95%				Confidence = 95%			
	Margin of Error				Margin of Error			
	5%	3.5%	2.5%	1%	5%	3.5%	2.5%	1%
10	10	10	10	10	10	10	10	10
20	19	20	20	20	19	20	20	20
30	28	29	29	30	29	29	30	30
50	44	47	48	50	47	48	49	50
75	63	69	72	74	67	71	73	75
100	80	89	94	99	87	93	96	99
150	108	126	137	148	122	135	142	149
200	132	160	177	196	154	174	186	198
250	152	190	215	24	182	211	229	246
300	169	217	251	291	207	246	270	295
400	196	265	318	384	250	309	348	391
500	217	306	377	475	285	365	421	485
600	234	340	432	565	315	416	490	579
700	248	370	481	653	341	462	554	672
800	260	396	526	739	363	503	615	763

* Retrieved from http://research-advisors.com

Questionnaire Design

It is important to design questions very carefully. A poorly designed questionnaire renders results meaningless. There are many factors to consider as following:

- Ask only one clear questions, avoid double-barreled questions,
- Make items clear (do not assume the person you are questioning knows the terms you are using),
- Make sure that the respondents are able to answer the questions,
- Questions should be relevant (don't ask questions on topics that respondents don't care about or haven't thought about),
- Short items are best, so that they may be read, understood, and answered quickly),
- Avoid negative items,
- Avoid biased items and terms (be sensitive to the effect of your wording on respondents),
- Whenever possible, develop consistent response methods,
- As an ordinary rule, sequence questions from the general to the specific,
- Make the questions as easy to answer as possible,
- When unique and unusual terms need to be defined in questionnaire items, use very clear definitions,
- Use an attractive questionnaire format that conveys a professional image, and
- If Open-ended questions are employed, try to develop extremely clearly.

Exercises

1. Identify the sampling technique used in the following problems. (simple random sampling , cluster, stratified, or systematic sampling):

1) Every fifth person boarding a plane is searched thoroughly.
2) At ABC College, five math classes are randomly selected out of 20 and all of the students from each class are interviewed.
3) A researcher randomly selects and interviews twenty male and fifty female students.
4) A researcher of ABC Airline Co. interviews all of the passengers on five randomly selected flights.
5) Based on 1,250 responses from 4,200 surveys, the university sent questionnaire to its alumni, the university researcher estimated that the annual salary of its alumni was 52,500.
6) The researcher interviews everyone in Airline Business Degree program in ABC College to determine the percentage of students that own a car.
7) A market researcher randomly selected 200 drivers under 35 years of age and 100 drivers over 35 years of age.
8) All of the teachers from 85 randomly selected from schools in Bangkok were interviewed.
9) To avoid working late, the quality control manager inspects the last 10 items produced that day.
10) The names of 70 contestants are written on 70 cards. The cards are placed in a bag, and three names are picked from the bag.

2. .If you want to collect data and select sampling unit for 30 units, describe how to get your sampling unit by using Non-probability sampling such as Convenience, haphazard or accidental sampling.

3. Mr. Patana wants to survey about the choice of foods and the satisfaction with the eating facilities in the RB restaurant in his College. He wants to obtain a random sample and he proposes the following method:

Interviewers should stand at the entrance of the College and select the first 50 students who walk by before 9.00am.

a) What kind of sample would the Mr. Patana really get?
b) In what way might this sampling method be biased?
c) Describe how Mr. Patana could find a sample of 50 students by using a probability sampling method.
d) If Mr. Patana want to use stratified random sampling, what strata would you recommend that he choose?
e) Describe the advantage and disadvantage of using a random sample method in Item (a) and Item (c).

_____