1.3 DESCRIPTIVE STATISTICS

When social scientists conduct a study, they usually collect a great deal of information or data about the problem at hand. The data may take a variety of forms: the number of medical schools in each state, the SAT scores of a group of college students, or responses to an attitude survey in a questionnaire. As collected in their original form, these data are usually a confusing hodgepodge of scores, frequency counts, and so forth. In performing the descriptive function, the statistician employs rules and procedures for presenting information about a variable or the relationship between variables in a more usable and meaningful form.

Rules are also followed when calculating various statistics from masses of raw data. Imagine that a team of social scientists administered a questionnaire concerning drug usage and attitudes to a group of high school students. What are some of the things they might do with the resulting information?

1. They might rearrange the scores and group them in various ways in order to see at a glance an overall picture of the data (Chapter 3, "Frequency Distributions and Graphing Techniques"). For example, how many of the students use drugs regularly? How many have never used drugs? Which drugs are used most frequently? What is the age distribution of the students? How do student attitudes toward drugs vary? What is the extent of multiple drug use?

2. They might construct tables, graphs, and figures to permit visualization of the results (Section 3.2, "Graphing Techniques," in Chapter 3). A pie chart might summarize the drugs most frequently used by the students, whereas a trend chart would provide a concise picture of the extent of drug use by year in school.

3. They might convert raw scores to other types of scores that are more useful for specific purposes. Thus, these scores could be used to rank the students, compare a student with the class average and so on. Other types of conversion will also be described in the text (Chapter 4, "Percentiles," and Chapter 7, "The Standard Deviation and the Standard Normal Distribution").

4. They might calculate averages to learn something about the typical attitudes of the students (Chapter 5, "Measures of Central Tendency"). What is the mean age when the students first used marijuana? What is the most frequently held attitude toward the legalization of marijuana?

5. Using the average as a reference point, they might describe the spread of scores about this central point. Statistics that measure this spread of scores are known as measures of variability or measures of dispersion (Chapter 6, "Measures of Dispersion"). Are the attitudes held by freshmen more or less homogeneous than those of seniors? Are males more likely to be multidrug users than females?

6. They may obtain a relationship between two different variables. Statistics are available for describing the relationship between two variables. The statistics for describing the extent of the relationship are referred to as *measures of association.*

Such statistics are extremely useful to the social scientist. For example, it might be important to determine the relationship between the number of books in the home and classroom grades or that between students' attitudes toward drugs and the extent to which they reported having committed delinquent acts. Once these relationships are known, the social scientist may use information from one variable to predict another (Chapter 8, "An Introduction to Contingency Tables," Chapter 9, "Correlation," and Chapter 10, "Regression and Prediction").

1.4 INFERENTIAL STATISTICS

As a social scientist, your task is not nearly over when you have completed the descriptive function. In fact, you are often nearer to the beginning than to the end of the task. The reason for this is clear when we consider that the purpose of the research is often to explore hypotheses of a general nature rather than simply to compare limited samples.

Imagine that you are a sociologist who is interested in determining the effectiveness of a videotape lecture approach compared with the usual approach of the professor delivering lectures in person. Consequently, you design a study involving two conditions: *experimental* and *control*. The question you wish to answer is which approach results in better student performance on a final sociology examination. The students in the experimental group view videotapes of the lectures attended by the control group. The students in the control group serve as a basis for comparison and are taught by means of the usual lecture approach. After all students have been tested, you might find that "on the average" the experimental group that viewed the

videotapes did not perform as well on the final examination as the control group. In other words, the average of the experimental group on the final examination was lower than that of the control group. You then ask the question: Can I conclude that the teaching technique produced the difference between the two groups? To answer this question, it is not sufficient to rely solely upon *descriptive statistics*.

After all, you reason, even if both techniques worked, it is highly unlikely that the final examination averages of the two groups would have been *identical*. *Some* difference would have been observed. The operation of chance factors such as the time at which the classes are scheduled is certain to produce some differences in the final examination scores between the two groups. The critical question from the point of view of inferential statistics becomes: Is the difference between groups great enough to rule out chance factors in the experiment as a sufficient explanation? Stated another way: If you were to repeat the experiment, would you be able to predict with confidence that the same differences (i.e., the control group would consistently perform better than the experimental group) would systematically occur? Ultimately, we wish to infer to the population from which the study participants were drawn. Thus, we want to use the results of this experiment to reach conclusions about which technique would be better in similar sociology courses.

As you can see, *inference* is used in much the same way that logical reasoning is used when talking with friends. This notation of logical reasoning is sometimes employed to establish community standards in pornography litigation. Because the tolerance for pornography differs greatly from community to community, a survey of residents is often conducted to determine what sorts of sexually explicit activities should be available commercially. A well-drawn sample survey of 600 Milwaukee adults could be used by prosecution or defense lawyers in a pornography case to conclude that Milwaukee adults hold a particular attitude. If 78% of those interviewed believed that nude nightclubs should not be allowed, the prosecutors would seek to convince the jury members that the sample survey results allow for the inference that basically the same attitude is held by about three out of four adult residents in the entire population of Milwaukee.* In this instance, the sample was used to infer to the population from which it was drawn.

Most population parameters can never be known directly except in cases in which the population is small or if we have sufficient funds to study the population in its entirety. Counting election day ballots and taking the census are examples of measuring the population. We could also argue that the students in your statistics class are a population. Your instructor measures the population's knowledge of statistics each time he or she administers an exam. In most cases, however, we will almost certainly never know specific population parameters such as the percentage of people afflicted with acquired immune deficiency syndrome (AIDS). However, sampling from popula-

2020-4-2 14:12

প্রসঙ্গে Mueller & Schuessler[1] বলেন : "The concept of statistics may therefore be defined in two related senses : (1) the factual data themselves, such as vital statistics, statistics of trade, production, and the like, and (2) The methods, theories, and techniques by means of which the collected descriptions are summarized and interpreted". Elifson et al[2] পরিসংখ্যান পদ্ধতির সংজ্ঞা এইভাবে নির্ধারণ করেন : "Statistics is a method for dealing with data and involves the organization and analysis of numerical facts or observations that are collected in accordance with a systematic plan", ক্রিয়াগত দিক থেকে পরিসংখ্যান পদ্ধতি দুটি ভাগে বিভক্ত থাকে—**বর্ণনামূলক পরিসংখ্যান পদ্ধতি (Descriptive statistics) এবং সিদ্ধান্তমূলক পরিসংখ্যান পদ্ধতি (Inferential statistics)**।

বর্ণনামূলক পরিসংখ্যান পদ্ধতি (Descriptive Statistics) : সমাজবিজ্ঞানীরা সামাজিক সমস্যা সম্পর্কে গবেষণা সূত্রে বিপুল তথ্যসম্ভার সংগ্রহ করে থাকে। সাধারণত এইসব তথ্য প্রাপ্তাঙ্ক (scores) এবং পরিসংখ্যার (frequency) এক বিভ্রান্তিকর অবিন্যস্ত সংকলন হয়ে থাকে। এই অবিন্যস্ত তথ্যাবলী বোধগম্য বা অর্থবহ করে তোলা গবেষকের কাছে এক জটিল সমস্যা হয়ে থাকে। এই সমস্যা সমাধানে বর্ণনামূলক পরিসংখ্যান পদ্ধতি প্রয়োগ বিশেষ ফলপ্রদ হয়। এই পদ্ধতি দুইভাবে সাহায্য করে। এই পদ্ধতির দ্বারা তথ্যাবলী সংক্ষেপায়ন (reduction) করা যায় এবং একাধিক চলের (variable) মধ্যে সম্পর্ক নির্ধারণ করা যায়। এ প্রসঙ্গে Joseph F. Healey[3] বলেন : "Descriptive statistics is relevant (i) when the researcher needs to summarize or describe the distribution of a single variable and (ii) when the researcher wishes to understand the relationship between two or more variables" তথা সংক্ষেপায়নের জন্য যে সব প্রক্রিয়া প্রয়োগ করা হয় তাদের মধ্যে বিশেষ উল্লেখযোগ্য হল পরিসংখ্যা বণ্টন সারণী, শতাংশে হার নির্ণয়, লেখচিত্র, কেন্দ্রীয় প্রবণতা ও মান সমূহের বিস্তৃতি নির্ণয়। একাধিক চলের মধ্যে সম্পর্ক নির্ণয়ের ক্ষেত্রে যে সব প্রক্রিয়া প্রয়োগ করা হয় তাদের মধ্যে বিশেষ উল্লেখযোগ্য হল Pearson-এর 'r', Spearman-এর Rho, লামডা (λ) ইত্যাদি। এইসব পদ্ধতি প্রয়োগে দুটি চলের মধ্যে সম্পর্ক আছে কিনা, সম্পর্ক সদর্থক না নঞর্থক এবং সম্পর্কের মান নির্ণয় করা যায়।

সিদ্ধান্তমূলক পরিসংখ্যান পদ্ধতি (Inferential Statistics) : বর্ণনামূলক পরিসংখ্যান পদ্ধতি সাধারণত নমুনা সংশ্লিষ্ট তথ্য থেকে নমুনার বৈশিষ্ট্যাবলী (statistics) নির্ণয় করে থাকে। সিদ্ধান্তমূলক পরিসংখ্যান পদ্ধতি নমুনার বৈশিষ্ট্য থেকে সমগ্রকের (population) বৈশিষ্ট্যাবলী (parameter) নির্ণয়ে সহায়ক হয়ে থাকে। Blalock[4] বলেন এই প্রক্রিয়ার গুরুত্বপূর্ণ কাজ হল : "Inferring properties of a population on the basis of known sample results", এছাড়া, দুটি নমুনার বৈশিষ্ট্যের মধ্যে গুরুত্বপূর্ণ পার্থক্য বা সমতা নির্ণয়ের ক্ষেত্রে এই পদ্ধতি বিশেষ কার্যকরী হয়। বিশেষ করে পরীক্ষামূলক গবেষণার ক্ষেত্রে গবেষণামূলক গোষ্ঠী এবং নিয়ন্ত্রিত গোষ্ঠীর গবেষণা সংশ্লিষ্ট চলের প্রভাবসাপেক্ষে বৈ

2020-4-2 14:13

This way. Work

2020-4-2 14:13

বিষমতা নির্ণয় করা যায়। এক্ষেত্রে তথ্যাবলীর পরিমাপক-এর স্তর (leveles of measurement) অনুযায়ী ক্ষেত্রবিশেষে X^2, t-test, Z-test ইত্যাদি প্রয়োগ করা হয়ে থাকে।

সামাজিক গবেষণায় পরিসংখ্যান পদ্ধতির ভূমিকা (Role of Statistics in Social Research) :

সামাজিক গবেষণায় গবেষণামূলক প্রশ্ন নির্ধারণ, প্রকল্প নির্মাণ, প্রকল্প অন্তর্ভুক্ত ধারণাবলীর কার্যকরীকরণ করার পর গবেষক সুপরিকল্পিতভাবে তথ্য সংগ্রহ করে থাকে। এই স্তরে বিশদ তথ্যাবলী অর্থবহ করা গবেষকের কাছে এক সমস্যা হয়ে থাকে। পরিসংখ্যান পদ্ধতির প্রয়োগ এই স্তরে গবেষণাকার্যে বিশেষ সহায়ক হয়ে থাকে। Blallock[1] বলেন : "Actually, statistical considerations enters only into the analysis stage of the research process, after all data have been collected" (বর্ণনামূলক পরিসংখ্যান কৌশল যেমন, পরিসংখ্যা বিভাজন, লেখচিত্র, কেন্দ্রীয় মান ও বিস্তৃতি, প্রক্রিয়া প্রয়োগে অবিন্যস্ত বিভ্রান্তিকর তথ্যমালা সংক্ষেপায়িত করে অর্থবহ করা যায়। গবেষক দ্বিচল বা বহুচলের সম্পর্ক নির্ধারণে অসহায় বোধ করলে পরিসংখ্যানের দ্বিচল এবং বহুচল বিশ্লেষণ প্রক্রিয়া (multivariate technique) প্রয়োগে বিভিন্ন চলের মধ্যে সম্পর্ক নির্ধারণ করা যায়। পরীক্ষামূলক গবেষণায় (experimental research) কোনো চলের প্রভাব নির্দেশ করে কোনো তাত্ত্বিক ধারণাকে প্রতিষ্ঠিত বা খণ্ডন করার ক্ষেত্রে t-test প্রক্রিয়া প্রয়োগ করা যায়। এই প্রক্রিয়া দ্বারা পরীক্ষামূলক গোষ্ঠী (experimental group) এবং নিয়ন্ত্রিত গোষ্ঠীর (control group) বৈশিষ্ট্যের মধ্যে গুরুত্বপূর্ণ পার্থক্য (significant difference) থাকা বা না থাকা নির্দেশ করে সংশ্লিষ্ট চলের (variable) প্রভাব নির্ণয় করা যায়, এবং এই সূত্রে কোনো তাত্ত্বিক ধারণা সমর্থন বা খণ্ডন করা হয়ে থাকে। এছাড়া, নিরীক্ষামূলক গবেষণায় কোনো নমুনার বৈশিষ্ট্য (statistics) থেকে সমগ্রকের বৈশিষ্ট্য (parameter) অনুমানিক নির্ধারণ (estimation) করা যায়। Joseph Healey[2] বলেন : "Statistics permit us to analyse data, to identify and probe trends and relationships, to develop generalizations, and to revise and improve our theories" তথ্য সংগ্রহের পর পরিসংখ্যান পদ্ধতি ব্যবহার করা হয় গবেষণা প্রক্রিয়ার পরবর্তী অন্যান্য স্তরে বিশেষ করে নমুনায়ন কাজে পরিসংখ্যানের বিভি