1 PARAMETRIC TESTS

1.1 **STEPS OF HYPOTHESIS** : In statistics, parametric and nonparametric methodologies refer to those in which a set of data has a normal vs. a non-normal distribution, respectively. Parametric tests make certain assumptions about a data set; namely, that the data are drawn from a population with a specific (normal) distribution. Non-parametric tests make fewer assumptions about the data set. The majority of elementary statistical methods are parametric, and parametric tests generally have higher statistical power. If the necessary assumptions cannot be made about a data set, non-parametric tests can be used.

1.1.1 RESEARCH HYPOTHESIS: Parametric Hypothesis tests are frequently used to measure the quality of sample parameters or to test whether estimates on a given parameter are equal for two samples. Parametric Hypothesis tests set up a null hypothesis against an alternate hypothesis, testing, for instance, whether or not the population mean is equal to a certain value, and then using an appropriate statistic to calculate the probability that the null hypothesis is true. You then reject or accept the null hypothesis based on this calculated probability.

1.1.2. IMPORTANCE OF THE HYPOTHESIS: A hypothesis is a theory or proposition set forth as an explanation for the occurrence of some observed phenomenon, asserted either as a provisional conjecture to guide investigation, called a working hypothesis, or accepted as highly probable in lieu of the established facts. A scientific hypothesis can become a theory or ultimately a law of nature if it is proven by repeatable experiments. Hypothesis testing is common in statistics as a method of making decisions using data. In other words, testing a hypothesis is trying to determine if your observation of some phenomenon is likely to have really occurred based on statistics.

Statistical Hypothesis Testing

- Statistical hypothesis testing, also called confirmatory data analysis, is often used to decide whether experimental results contain enough information to cast doubt on conventional wisdom. For example, at one time it was thought that people of certain races or color had inferior intelligence compared to Caucasians. A hypothesis was made that intelligence is not based on race or color. People of various races, colors and cultures were given intelligence tests and the data was analyzed. Statistical hypothesis testing then proved that the results were statistically significant in that the similar measurements of intelligence between races are not merely sample error.

Null and Alternative Hypotheses

- Before testing for phenomena, you form a hypothesis of what might be happening. Your hypothesis or guess about what's occurring might be that certain groups are different from each other, or that intelligence is not correlated with skin color, or that some treatment has an effect on an outcome measure, for examples. From this, there are two possibilities: a "null hypothesis" that nothing happened, or there were no differences, or no cause and effect; or that you were correct in your theory, which is labeled the "alternative hypothesis." In short, when you test a statistical hypothesis, you are trying to see if something happened and are comparing against the possibility that nothing happened. Confusingly, you are trying to disprove that nothing happened. If you disprove that nothing happened, then you can conclude that something happened.

Importance of Hypothesis Testing

- According to the San Jose State University Statistics Department, hypothesis testing is one of the most important concepts in statistics because it is how you decide if something really happened, or if certain treatments have positive effects, or if groups differ from each other or if one variable predicts another. In short, you want to proof if your data is statistically significant and unlikely to have occurred by chance alone. In essence then, a hypothesis test is a test of significance.

Possible Conclusions

- Once the statistics are collected and you test your hypothesis against the likelihood of chance, you draw your final conclusion. If you reject the null hypothesis, you are claiming that your result is statistically significant and that it did not happen by luck or chance. As such, the outcome proves the alternative hypothesis. If you fail to reject the null hypothesis, you must conclude that you did not find an effect or difference in your study. This method is how many pharmaceutical drugs and medical procedures are tested.

1.2 PARAMETRIC STATISTICS: Parametric statistics is a branch of statistics which assumes that the data has come from a type of probability distribution and makes inferences about the parameters of the distribution. Most well-known elementary statistical methods are parametric.

Generally speaking parametric methods make more assumptions than non-parametric methods. If those extra assumptions are correct, parametric methods can produce more accurate and precise estimates. They are said to have more statistical power. However, if assumptions are incorrect, parametric methods can be very misleading. For that reason they are often not considered robust. On the other hand, parametric formulae are often simpler to write down and faster to compute. In some, but definitely not all cases, their simplicity makes up for their non-robustness, especially if care is taken to examine diagnostic statistics.

Example

Suppose we have a sample of 99 test scores with a mean of 100 and a standard deviation of 1. If we assume all 99 test scores are random samples from a normal distribution we predict there is a 1% chance that the 100th test score will be higher than 102.365 (that is the mean plus 2.365 standard deviations) assuming that the 100th test score comes from the same distribution as the others. The normal family of distributions all have the same shape and are *parameterized* by mean and standard deviation. That means if you know the mean and standard deviation, and that the distribution is normal, you know the probability of any future observation. Parametric statistical methods are used to compute the 2.365 value above, given 99 independent observations from the same normal distribution.

A non-parametric estimate of the same thing is the maximum of the first 99 scores. We don't need to assume anything about the distribution of test scores to reason that before we gave the test it was equally likely that the highest score would be any of the first 100. Thus there is a 1% chance that the 100th is higher than any of the 99 that preceded it.

History

Statistician Jacob Wolfowitz coined the statistical term "parametric" in order to define its opposite in 1942:

"Most of these developments have this feature in common, that the distribution functions of the various stochastic variables which enter into their problems are assumed to be of known functional form, and the theories of estimation and of testing hypotheses are theories of estimation of and of testing hypotheses about, one or more parameters. . ., the knowledge of which would completely determine the various distribution functions involved. We shall refer to this situation as the parametric case, and denote the opposite case, where the functional forms of the distributions are unknown, as the non-parametric case.

1.2.1 DESCRIPTION OF PARAMETRIC TESTS: Parametric tests are defined as conventional statistical methods. For instance, in parametric testing, a sample value is obtained to approximate the population parameters. These tests require parametric assumptions as the values used are sample statistics.

1.3 INFERENTIAL STATISTICS: With inferential statistics, you are trying to reach conclusions that extend beyond the immediate data alone. For instance, we use inferential statistics to try to infer from the sample data what the population might think. Or, we use inferential statistics to make judgments of the probability that an observed difference between groups is a dependable one or one that might have happened by chance in this study. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in our data.

Here, I concentrate on inferential statistics that are useful in experimental and quasi-experimental research design or in program outcome evaluation. Perhaps one of the simplest inferential test is used when you want to compare the average performance of two groups on a single measure to see if there is a difference. You might want to know whether eighth-grade boys and girls differ in math test scores or whether a program group differs on the outcome measure from a control group. Whenever you wish to compare the average performance between two groups you should consider the t-test for differences between groups.

Most of the major inferential statistics come from a general family of statistical models known as the General Linear Model. This includes the t-test, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA), regression analysis, and many of the multivariate methods like factor analysis, multidimensional scaling, cluster analysis, discriminant function analysis, and so on. Given the importance of the General Linear Model, it's a good idea for any serious social researcher to become familiar with its workings. The discussion of the General Linear Model here is very elementary and only considers the simplest straight-line model. However, it will get you familiar with the idea of the linear model and help prepare you for the more complex analyses described below.

One of the keys to understanding how groups are compared is embodied in the notion of the "dummy" variable. The name doesn't suggest that we are using variables that aren't very smart or, even worse, that the analyst who uses them is a "dummy"! Perhaps these variables would be better described as "proxy" variables.

Essentially a dummy variable is one that uses discrete numbers, usually 0 and 1, to represent different groups in your study. Dummy variables are a simple idea that enable some pretty complicated things to happen. For instance, by including a simple dummy variable in an model, I can model two separate lines (one for each treatment group) with a single equation. To see how this works, check out the discussion on dummy variables.

1.4

1.4.1 APPROACH OF THE PROBLEM

1.4.2 DEVELOP A MODEL: DEVELOPING A PARAMETRIC MODEL

The development of a parametric estimating model can appear to be a daunting task; however, the use of modern computer technology (including popular spreadsheet programs) can make the process tolerable, and much easier than it would have been many years ago. The process of developing a parametric model should generally involve the following steps:

• **COST MODEL SCOPE DETERMINATION:** The first step in developing a parametric model is to establish its scope. This includes defining the end use of the model, the physical characteristics of the model, the cost basis of the model, and its critical components and cost drivers.

• **DATA COLLECTION:** Data collection and development for a parametric model require a significant effort. The quality of the resulting parametric model can be no better than the quality of the data it is based upon. Both cost and scope information must be identified and collected. For the PROCEP model, it was decided to collect data from recent projects where the cost and scope information for process controls could be identified readily and clearly. The collection involved both complete projects (where process controls is a part of the overall project scope) and projects that were primarily process control projects (upgrades or replacements of process control systems). The type of data to be collected was decided upon in cooperation with the process control engineering community. A data input form was developed to aid in collecting the data. Over time, this form has been revised as the data needs were better identified, and the parametric model was revised.

• **DATA NORMALIZATION:** After the data have been collected, the next step in the process of developing a parametric model is to normalize the data before the data analysis stage. Normalizing the data refers to making adjustments to the data to account for differences between the actual basis for each project and a desired standard basis of data to be used for the parametric model. Typically, data normalization implies making adjustments for:

- escalation;
- general location;
- site conditions;
- system specifications; and estimate title information;
- total digital and analog I/O to be addressed by the controller hardware;
- total digital and analog I/O to be purchased by the project;
- total control valves to be purchased;
- total digital and analog I/O to be installed;
- hardware type (distributed, programmable logic controller, etc.);
- redundant hardware/spare capacity information; and
- type of process being controlled.

- **DATA ANALYSIS**: The next step in the development of a parametric model is data analysis. There are many diverse methods and techniques that can be used in data analysis, and they are too complex to delve into here. Typically, data analysis consists of performing regression analysis of costs versus selected design parameters to determine the key costs drivers for the model. Most spreadsheet applications now provide regression analysis and simulation functions, which are reasonably simple to use.

- **DATA APPLICATION:** The data application stage involves establishing the user interface and presentation form for the parametric cost model. Using the mathematical and statistical algorithms developed in the data analysis stage, the various inputs to the cost model are identified, and an interface is developed to provide the user with an easy and straightforward way to enter this information. Electronic spreadsheets provide an excellent mechanism to accept user input, calculate costs based upon algorithms, and display the resulting output.

- **TESTING**: One of the most important steps in developing a cost model is to test its accuracy and validity. When using regression analysis, one of the key indicators of how well a resulting algorithm explains the data is a term called $R2$. This is the coefficient of determination, and it provides a measure of how well the resulting algorithm predicts the calculated costs. An $R2$ value of 1 indicates a perfect fit, while an $R2$ value of .89 indicates an 89 percent confidence that the regression equation explains the variability in cost (based on the data used in the development of the model). However, a high $R2$ value by itself does not imply that the relationships between the data inputs and the resulting cost are statistically significant.

• **DOCUMENTATION:** The resulting cost model and estimating application must be documented thoroughly. A user manual for the estimating application should be prepared that shows the steps involved in preparing an estimate using the cost model, and clearly describes the required inputs to the cost model. The data used to create the cost model should be documented, including a discussion of how the data were adjusted or normalized for use in the data analysis stage. It is usually desirable to make the actual regression data sets available, along with the resulting regression and test results. All assumptions and allowances designed into the cost model should be documented, as should any exclusion. The range of applicable input values, and the limitations of the model's algorithms, also should be explained.