DOWNLOAD PDF USING NON-PARAMETRIC TESTS TO MEASURE THE DIFFERENCE BETWEEN INDEPENDENT GROUPS

Chapter 1 : nonparametric - Repeated measures, between-groups non parametric test? - Cross Validated

Preferred measure of variation is the standard deviation of variance which describes the distribution of data. All obs are used & average distance from the theinnatdunvilla.comation is as follows: compare first obs with mean, +ve and -ve on either side.

For example, you could use the Mann-Whitney U test to understand whether attitudes towards pay discrimination, where attitudes are measured on an ordinal scale, differ based on gender i. Alternately, you could use the Mann-Whitney U test to understand whether salaries, measured on a continuous scale, differed based on educational level i. The Mann-Whitney U test is often considered the nonparametric alternative to the independent t-test although this is not always the case. These conclusions can range from simply stating whether the two populations differ through to determining if there are differences in medians between groups. These different conclusions hinge on the shape of the distributions of your data, which we explain more about later. In our enhanced Mann-Whitney U test guide, we take you through all the steps required to understand when and how to use the Mann-Whitney U test, showing you the required procedures in SPSS Statistics, and how to interpret and report your output. You can access this enhanced Mann-Whitney U test guide by subscribing to the site here. Before we show you how to do this, we explain the different assumptions that your data must meet in order for a Mann-Whitney U test to give you a valid result. We discuss these assumptions next. SPSS Statistics Assumptions When you choose to analyse your data using a Mann-Whitney U test, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using a Mann-Whitney U test. You need to do this because it is only appropriate to use a Mann-Whitney U test if your data "passes" four assumptions that are required for a Mann-Whitney U test to give you a valid result. In practice, checking for these four assumptions just adds a little bit more time to your analysis, requiring you to click a few more buttons in SPSS Statistics when performing your analysis, as well as think a little bit more about your data, but it is not a difficult task. Before we introduce you to these four assumptions, do not be surprised if, when analysing your own data using SPSS Statistics, one or more of these assumptions is violated i. This is not uncommon when working with real-world data rather than textbook examples, which often only show you how to carry out a Mann-Whitney U test when everything goes well! Even when your data fails certain assumptions, there is often a solution to overcome this. Your dependent variable should be measured at the ordinal or continuous level. Examples of ordinal variables include Likert items e. Examples of continuous variables include revision time measured in hours , intelligence measured using IQ score , exam performance measured from 0 to , weight measured in kg , and so forth. You can learn more about ordinal and continuous variables in our article: Your independent variable should consist of two categorical, independent groups. Example independent variables that meet this criterion include gender 2 groups: You should have independence of observations, which means that there is no relationship between the observations in each group or between the groups themselves. For example, there must be different participants in each group with no participant being in more than one group. This is more of a study design issue than something you can test for, but it is an important assumption of the Mann-Whitney U test. If your study fails this assumption, you will need to use another statistical test instead of the Mann-Whitney U test e. If you are unsure whether your study meets this assumption, you can use our Statistical Test Selector , which is part of our enhanced content. A Mann-Whitney U test can be used when your two variables are not normally distributed. However, in order to know how to interpret the results from a Mann-Whitney U test, you have to determine whether your two distributions i. To understand what this means, take a look at the diagram below: However, in the diagram on the right, even though both distributions have the same shape, they have a different location i. However, if your two distribution have a different shape, you can only use the Mann-Whitney U test to compare mean ranks. Therefore, when carrying out a Mann-Whitney U test, you must also use SPSS Statistics to determine whether your two distributions have the same shape or a different shape.

DOWNLOAD PDF USING NON-PARAMETRIC TESTS TO MEASURE THE DIFFERENCE BETWEEN INDEPENDENT GROUPS

This requires a few more procedures in SPSS Statistics, but it is an easy step-by-step process that we show you how to do in our enhanced Mann-Whitney U test guide. In this "quick start" guide, we show you how to carry out a Mann-Whitney U test assuming that your two distributions do not have a similar shape, such that you can only compare mean ranks and not medians. Just remember that if you do not check assumption 4, you will not know whether you are correctly comparing mean ranks or medians, and the results you get when running a Mann-Whitney U test may not be valid. This is why we dedicate a number of sections of our enhanced Mann-Whitney U test guide to help you get this right. You can learn more about assumption 4 and what you will need to interpret in the Assumptions section of our enhanced Mann-Whitney U test guide, which you can access by subscribing to the site here. In the Test Procedure in SPSS Statistics section of this "quick start" guide, we illustrate the SPSS Statistics procedure to perform a Mann-Whitney U test assuming that your two distributions are not the same shape and you have to interpret mean ranks rather than medians.

DOWNLOAD PDF USING NON-PARAMETRIC TESTS TO MEASURE THE DIFFERENCE BETWEEN INDEPENDENT GROUPS

Chapter 2 : Differences between Groups

Non-Parametric Test for Independent Measures Between Two groups: Mann-Whitney test John Foxx/Stockbyte/Getty Images The Mann-Whitney Test is used to compare the means between two groups of ordinal (thus, non-parametric) data.

Null and alternative hypotheses for the independent t-test The null hypothesis for the independent t-test is that the population means from the two unrelated groups are equal: Most commonly, this value is set at 0. What do you need to run an independent t-test? In order to run an independent t-test, you need the following: One continuous dependent variable. Unrelated groups Unrelated groups, also called unpaired groups or independent groups, are groups in which the cases e. Often we are investigating differences in individuals, which means that when comparing two groups, an individual in one group cannot also be a member of the other group and vice versa. An example would be gender - an individual would have to be classified as either male or female â€" not both. Join the 10,s of students, academics and professionals who rely on Laerd Statistics. Technically, it is the residuals that need to be normally distributed, but for an independent t-test, both will give you the same result. You can test for this using a number of different tests, but the Shapiro-Wilks test of normality or a graphical method, such as a Q-Q Plot, are very common. However, the t-test is described as a robust test with respect to the assumption of normality. This means that some deviation away from normality does not have a large influence on Type I error rates. The exception to this is if the ratio of the smallest to largest group size is greater than 1. Assumption of homogeneity of variance The independent t-test assumes the variances of the two groups you are measuring are equal in the population. If your variances are unequal, this can affect the Type I error rate. This test for homogeneity of variance provides an F-statistic and a significance value p-value. We are primarily concerned with the significance value â€" if it is greater than 0. In all reality, you will probably never have heard of these adjustments because SPSS Statistics hides this information and simply labels the two options as "Equal variances assumed" and "Equal variances not assumed" without explicitly stating the underlying tests used. However, you can see the evidence of these tests as below: The effect of not being able to assume equal variances is evident in the final column of the above figure where we see a reduction in the value of the t-statistic and a large reduction in the degrees of freedom df. This has the effect of increasing the p-value above the critical significance level of 0. In this case, we therefore do not accept the alternative hypothesis and accept that there are no statistically significant differences between means. This would not have been our conclusion had we not tested for homogeneity of variances. The format of the test result is: Therefore, for the example above, you could report the result as t 7. It was found that after the two interventions, cholesterol concentrations in the dietary group 6.

Chapter 3 : When to Use a Nonparametric Test

A popular nonparametric test to compare outcomes between two independent groups is the Mann Whitney U test. The Mann Whitney U test, sometimes called the Mann Whitney Wilcoxon Test or the Wilcoxon Rank Sum Test, is used to test whether two samples are likely to derive from the same population (i.e., that the two populations have the same shape).

Definitions[edit] The statistician Larry Wasserman has said that "it is difficult to give a precise definition of nonparametric inference". The first meaning of nonparametric covers techniques that do not rely on data belonging to any particular distribution. These include, among others: As such it is the opposite of parametric statistics. It includes nonparametric descriptive statistics , statistical models , inference and statistical tests. Order statistics , which are based on the ranks of observations, is one example of such statistics and these play a central role in many nonparametric approaches. For example, the hypothesis a that a normal distribution has a specified mean and variance is statistical; so is the hypothesis b that it has a given mean but unspecified variance; so is the hypothesis c that a distribution is of normal form with both mean and variance unspecified; finally, so is the hypothesis d that two unspecified continuous distributions are identical. It will have been noticed that in the examples a and b the distribution underlying the observations was taken to be of a certain form the normal and the hypothesis was concerned entirely with the value of one or both of its parameters. Such a hypothesis, for obvious reasons, is called parametric. Hypothesis c was of a different nature, as no parameter values are specified in the statement of the hypothesis; we might reasonably call such a hypothesis non-parametric. Hypothesis d is also non-parametric but, in addition, it does not even specify the underlying form of the distribution and may now be reasonably termed distribution-free. Notwithstanding these distinctions, the statistical literature now commonly applies the label "non-parametric" to test procedures that we have just termed "distribution-free", thereby losing a useful classification. The second meaning of non-parametric covers techniques that do not assume that the structure of a model is fixed. Typically, the model grows in size to accommodate the complexity of the data. In these techniques, individual variables are typically assumed to belong to parametric distributions, and assumptions about the types of connections among variables are also made. These techniques include, among others: Applications and purpose[edit] Non-parametric methods are widely used for studying populations that take on a ranked order such as movie reviews receiving one to four stars. The use of non-parametric methods may be necessary when data have a ranking but no clear numerical interpretation, such as when assessing preferences. In terms of levels of measurement , non-parametric methods result in ordinal data. As non-parametric methods make fewer assumptions, their applicability is much wider than the corresponding parametric methods. In particular, they may be applied in situations where less is known about the application in question. Also, due to the reliance on fewer assumptions, non-parametric methods are more robust. Another justification for the use of non-parametric methods is simplicity. In certain cases, even when the use of parametric methods is justified, non-parametric methods may be easier to use. Due both to this simplicity and to their greater robustness, non-parametric methods are seen by some statisticians as leaving less room for improper use and misunderstanding. The wider applicability and increased robustness of non-parametric tests comes at a cost: In other words, a larger sample size can be required to draw conclusions with the same degree of confidence. Non-parametric models[edit] Non-parametric models differ from parametric models in that the model structure is not specified a priori but is instead determined from data. The term non-parametric is not meant to imply that such models completely lack parameters but that the number and nature of the parameters are flexible and not fixed in advance. A histogram is a simple nonparametric estimate of a probability distribution. Kernel density estimation provides better estimates of the density than histograms. Nonparametric regression and semiparametric regression methods have been developed based on kernels , splines , and wavelets. Data envelopment analysis provides efficiency coefficients similar to those obtained by multivariate analysis without any distributional assumption. KNNs classify the unseen instance based on the K points in the training set which are nearest to it. A support vector machine with a Gaussian kernel is a nonparametric large-margin classifier. Methods[edit] Non-parametric or distribution-free inferential statistical methods are mathematical procedures for statistical hypothesis testing which, unlike parametric statistics , make no assumptions about the probability distributions of the variables being assessed. The most frequently used tests include Andersonâ€”Darling test:

Chapter 4 : Mann-Whitney U Test in SPSS Statistics | Setup, Procedure & Interpretation | Laerd Statistics

Use this test to compare differences between two independent groups when dependent variables are either ordinal or continuous. Mood's Median test. Use this test instead of the sign test when you have two independent samples.

Choosing a Test In terms of selecting a statistical test, the most important question is "what is the main study hypothesis? For example, in a prevalence study there is no hypothesis to test, and the size of the study is determined by how accurately the investigator wants to determine the prevalence. If there is no hypothesis, then there is no statistical test. It is important to decide a priori which hypotheses are confirmatory that is, are testing some presupposed relationship , and which are exploratory are suggested by the data. No single study can support a whole series of hypotheses. A sensible plan is to limit severely the number of confirmatory hypotheses. Although it is valid to use statistical tests on hypotheses suggested by the data, the P values should be used only as guidelines, and the results treated as tentative until confirmed by subsequent studies. A useful guide is to use a Bonferroni correction, which states simply that if one is testing n independent hypotheses, one should use a significance level of 0. Note that, since tests are rarely independent, this is a very conservative procedure â€" i. The investigator should then ask "are the data independent? Thus results from a crossover trial, or from a case-control study in which the controls were matched to the cases by age, sex and social class, are not independent. Analysis should reflect the design, and so a matched design should be followed by a matched analysis. Results measured over time require special care. One of the most common mistakes in statistical analysis is to treat correlated variables as if they were independent. For example, suppose we were looking at treatment of leg ulcers, in which some people had an ulcer on each leg. We might have 20 subjects with 30 ulcers but the number of independent pieces of information is 20 because the state of ulcers on each leg for one person may be influenced by the state of health of the person and an analysis that considered ulcers as independent observations would be incorrect. For a correct analysis of mixed paired and unpaired data consult a statistician. The next question is "what types of data are being measured? The choice of test for matched or paired data is described in Table 1 and for independent data in Table 2. Table 1 Choice of statistical test from paired or matched observation It is helpful to decide the input variables and the outcome variables. For example, in a clinical trial the input variable is the type of treatment - a nominal variable - and the outcome may be some clinical measure perhaps Normally distributed. The required test is then the t-test Table 2. However, if the input variable is continuous, say a clinical score, and the outcome is nominal, say cured or not cured, logistic regression is the required analysis. A t-test in this case may help but would not give us what we require, namely the probability of a cure for a given value of the clinical score. As another example, suppose we have a cross-sectional study in which we ask a random sample of people whether they think their general practitioner is doing a good job, on a five point scale, and we wish to ascertain whether women have a higher opinion of general practitioners than men have. The input variable is gender, which is nominal. The outcome variable is the five point ordinal scale. Note, however, if some people share a general practitioner and others do not, then the data are not independent and a more sophisticated analysis is called for. Note that these tables should be considered as guides only, and each case should be considered on its merits. Table 2 Choice of statistical test for independent observations a If data are censored. However, they require certain assumptions and it is often easier to either dichotomise the outcome variable or treat it as continuous. Parametric tests are those that make assumptions about the parameters of the population distribution from which the sample is drawn. This is often the assumption that the population data are normally distributed. Table 3 shows the non-parametric equivalent of a number of parametric tests. Table 3 Parametric and Non-parametric tests for comparing two or more groups Non-parametric tests are valid for both non-Normally distributed data and Normally distributed data, so why not use them all the time? It would seem prudent to use non-parametric tests in all cases, which would save one the bother of testing for Normality. Parametric tests are preferred, however, for the following reasons: We are rarely interested in a significance test alone; we would like to say something about the population from which the samples came, and this is best done with estimates of parameters and confidence intervals. It is difficult to do flexible modelling with non-parametric tests, for example allowing for confounding factors using multiple regression. Parametric tests usually have more statistical power than their non-parametric equivalents. In other words, one is more likely to detect significant differences when they truly exist. Do non-parametric tests compare medians? It is a commonly held belief that a Mann-Whitney U test is in fact a test for differences in medians. However, two groups could have the same median and yet have a significant Mann-Whitney U test. Consider the following data for two groups, each with observations. Only if we are prepared to make the additional assumption that the difference in the two groups is simply a shift in location that is, the distribution of the data in one group is simply shifted by a fixed amount from the other can we say that the test is a test of the difference in medians. However, if the groups have the same distribution, then a shift in location will move medians and means by the same amount and so the difference in medians is the same as the difference in means. Thus the Mann-Whitney U test is also a test for the difference in means. How is the Mann- Whitney U test related to the t-test? If one were to input the ranks of the data rather than the data themselves into a two sample t-test program, the P value obtained would be very close to that produced by a Mann-Whitney U test. Statistics at Square One 11th ed.

Chapter 5 : What Are Parametric and Nonparametric Tests? | Sciencing

Is there any non-parametric test that would allow me to compare (W1) with (W2) between 4 groups of articles (A1 to A4)? So far I used Kruskall-Wallis and there are signifacant differences in W2 between groups A1 to A4, but the problem is it doesn't take into account the fisrt measurenment point.

Boston University School of Public Health Introduction The three modules on hypothesis testing presented a number of tests of hypothesis for continuous, dichotomous and discrete outcomes. Tests for continuous outcomes focused on comparing means, while tests for dichotomous and discrete outcomes focused on comparing proportions. All of the tests presented in the modules on hypothesis testing are called parametric tests and are based on certain assumptions. For example, when running tests of hypothesis for means of continuous outcomes, all parametric tests assume that the outcome is approximately normally distributed in the population. This does not mean that the data in the observed sample follows a normal distribution, but rather that the outcome follows a normal distribution in the full population which is not observed. For many outcomes, investigators are comfortable with the normality assumption i. It also turns out that many statistical tests are robust, which means that they maintain their statistical properties even when assumptions are not entirely met. Tests are robust in the presence of violations of the normality assumption when the sample size is large based on the Central Limit Theorem see page 11 in the module on Probability. When the sample size is small and the distribution of the outcome is not known and cannot be assumed to be approximately normally distributed, then alternative tests called nonparametric tests are appropriate. Learning Objectives After completing this module, the student will be able to: Compare and contrast parametric and nonparametric tests Identify multiple applications where nonparametric approaches are appropriate Perform and interpret the Mann Whitney U Test Perform and interpret the Sign test and Wilcoxon Signed Rank Test Compare and contrast the Sign test and Wilcoxon Signed Rank Test Perform and interpret the Kruskal Wallis test Identify the appropriate nonparametric hypothesis testing procedure based on type of outcome variable and number of samples When to Use a Nonparametric Test Nonparametric tests are sometimes called distribution-free tests because they are based on fewer assumptions e. Parametric tests involve specific probability distributions e. The cost of fewer assumptions is that nonparametric tests are generally less powerful than their parametric counterparts i. It can sometimes be difficult to assess whether a continuous outcome follows a normal distribution and, thus, whether a parametric or nonparametric test is appropriate. There are several statistical tests that can be used to assess whether data are likely from a normal distribution. Each test is essentially a goodness of fit test and compares observed data to quantiles of the normal or other specified distribution. The null hypothesis for each test is H0: Data follow a normal distribution versus H1: Data do not follow a normal distribution. If the test is statistically significant e. It should be noted that these tests for normality can be subject to low power. Specifically, the tests may fail to reject H0: Data follow a normal distribution when in fact the data do not follow a normal distribution. Low power is a major issue when the sample size is small - which unfortunately is often when we wish to employ these tests. The most practical approach to assessing normality involves investigating the distributional form of the outcome in the sample using a histogram and to augment that with data from other studies, if available, that may indicate the likely distribution of the outcome in the population. There are some situations when it is clear that the outcome does not follow a normal distribution. Using an Ordinal Scale Consider a clinical trial where study participants are asked to rate their symptom severity following 6 weeks on the assigned treatment. Symptom severity might be measured on a 5 point ordinal scale with response options: Symptoms got much worse, slightly worse, no change, slightly improved, or much improved. Distribution of Symptom Severity in Total Sample The distribution of the outcome symptom severity does not appear to be normal as more participants report improvement in symptoms as opposed to worsening of symptoms. When the Outcome is a Rank In some studies, the outcome is a rank. APGAR scores generally do not follow a normal distribution, since most newborns have scores of 7 or higher normal range. When There Are Outliers In some studies, the outcome is continuous but subject to outliers or extreme values. For example, days in the hospital following a particular surgical procedure is an outcome that is often subject to outliers. Suppose in an observational study investigators wish to assess whether there is a difference in the days patients spend in the hospital following liver transplant in for-profit versus nonprofit hospitals. The number of days in the hospital are summarized by the box-whisker plot below. Recall from page 8 in the module on Summarizing Data that we used Q In the box-whisker plot above, Limits of Detection In some studies, the outcome is a continuous variable that is measured with some imprecision e. For example, some instruments or assays cannot measure presence of specific quantities above or below certain limits. HIV viral load is a measure of the amount of virus in the body and is measured as the amount of virus per a certain volume of blood. It can range from "not detected" or "below the limit of detection" to hundreds of millions of copies. Thus, in a sample some participants may have measures like 1,, or , copies and others are measured as "not detected. Hypothesis Testing with Nonparametric Tests In nonparametric tests, the hypotheses are not about population parameters e. Instead, the null hypothesis is more general. For example, when comparing two independent groups in terms of a continuous outcome, the null hypothesis in a parametric test is H0: In a nonparametric test the null hypothesis is that the two populations are equal, often this is interpreted as the two populations are equal in terms of their central tendency. Advantages of Nonparametric Tests Nonparametric tests have some distinct advantages. With outcomes such as those described above, nonparametric tests may be the only way to analyze these data. Outcomes that are ordinal, ranked, subject to outliers or measured imprecisely are difficult to analyze with parametric methods without making major assumptions about their distributions as well as decisions about coding some values e. As described here, nonparametric tests can also be relatively simple to conduct. Introduction to Nonparametric Testing This module will describe some popular nonparametric tests for continuous outcomes. Interested readers should see Conover3 for a more comprehensive coverage of nonparametric tests. Parametric tests are generally more powerful and can test a wider range of alternative hypotheses. It is worth repeating that if data are approximately normally distributed then parametric tests as in the modules on hypothesis testing are more appropriate. However, there are situations in which assumptions for a parametric test are violated and a nonparametric test is more appropriate. The techniques described here apply to outcomes that are ordinal, ranked, or continuous outcome variables that are not normally distributed. Recall that continuous outcomes are quantitative measures based on a specific measurement scale e. Some investigators make the distinction between continuous, interval and ordinal scaled data. Interval data are like continuous data in that they are measured on a constant scale i. Differences between interval scores are interpretable, but ratios are not. Temperature in Celsius or Fahrenheit is an example of an interval scale outcome. Ordinal outcomes can be less specific as the ordered categories need not be equally spaced. Symptom severity is an example of an ordinal outcome and it is not clear whether the difference between much worse and slightly worse is the same as the difference between no change and slightly improved. Pain is often measured in this way, from 0 to 10 with 0 representing no pain and 10 representing agonizing pain. Participants are sometimes shown a visual scale such as that shown in the upper portion of the figure below and asked to choose the number that best represents their pain state. Sometimes pain scales use visual anchors as shown in the lower portion of the figure below. Visual Pain Scale In the upper portion of the figure, certainly 10 is worse than 9, which is worse than 8; however, the difference between adjacent scores may not necessarily be the same. It is important to understand how outcomes are measured to make appropriate inferences based on statistical analysis and, in particular, not to overstate precision. Assigning Ranks The nonparametric procedures that we describe here follow the same general procedure. The outcome variable ordinal, interval or continuous is ranked from lowest to highest and the analysis focuses on the ranks as opposed to the measured or raw values. First, the data are ordered from smallest to largest. The lowest value is then assigned a rank of 1, the next lowest a rank of 2 and so on. The observed data and corresponding ranks are shown below:

DOWNLOAD PDF USING NON-PARAMETRIC TESTS TO MEASURE THE DIFFERENCE BETWEEN INDEPENDENT GROUPS

Chapter 6 : Module Comparing 2 Samples: A Nonparametric Test

The Mann-Whitney U test in the tests for two independent samples is an alternative form of the t-test. It is widely used to test whether or not two independent samples are significantly different. In SPSS, the Mann-Whitney U test in the tests for two independent samples is done by selecting.

What statistical analysis should I use? Each section gives a brief description of the aim of the statistical test, when it is used, an example showing the SPSS commands and SPSS often abbreviated output with a brief interpretation of the output. You can see the page Choosing the Correct Statistical Test for a table that shows an overview of when each test is appropriate to use. In deciding which test is appropriate to use, it is important to consider the type of variables that you have i. About the hsb data file Most of the examples in this page will use a data file called hsb2, high school and beyond. This data file contains observations from a sample of high school students with demographic information about the students, such as their gender female , socio-economic status ses and ethnic background race. It also contains a number of scores on standardized tests, including tests of reading read , writing write , mathematics math and social studies socst. You can get the hsb data file by clicking on hsb2. One sample t-test A one sample t-test allows us to test whether a sample mean of a normally distributed interval variable significantly differs from a hypothesized value. For example, using the hsb2 data file , say we wish to test whether the average writing score write differs significantly from We can do this as shown below. The mean of the variable write for this particular sample of students is We would conclude that this group of students has a significantly higher mean on the writing test than One sample median test A one sample median test allows us to test whether a sample median differs significantly from a hypothesized value. We will use the same variable, write, as we did in the one sample t-test example above, but we do not need to assume that it is interval and normally distributed we only need to assume that write is an ordinal variable. Binomial test A one sample binomial test allows us to test whether the proportion of successes on a two-level categorical dependent variable significantly differs from a hypothesized value. Chi-square goodness of fit A chi-square goodness of fit test allows us to test whether the observed proportions for a categorical variable differ from hypothesized proportions. We want to test whether the observed proportions from our sample differ significantly from these hypothesized proportions. Two independent samples t-test An independent samples t-test is used when you want to compare the means of a normally distributed interval dependent variable for two independent groups. For example, using the hsb2 data file , say we wish to test whether the mean for write is the same for males and females. Because the standard deviations for the two groups are similar In other words, females have a statistically significantly higher mean score on writing An overview of statistical tests in SPSS Wilcoxon-Mann-Whitney test The Wilcoxon-Mann-Whitney test is a non-parametric analog to the independent samples t-test and can be used when you do not assume that the dependent variable is a normally distributed interval variable you only assume that the variable is at least ordinal. We will use the same data file the hsb2 data file and the same variables in this example as we did in the independent t-test example above and will not assume that write, our dependent variable, is normally distributed. Why is the Mann-Whitney significant when the medians are equal? Chi-square test A chi-square test is used when you want to see if there is a relationship between two categorical variables. In SPSS, the chisq option is used on the statistics subcommand of the crosstabs command to obtain the test statistic and its associated p-value. Remember that the chi-square test assumes that the expected value for each cell is five or higher. This assumption is easily met in the examples below. The point of this example is that one or both variables may have more than two levels, and that the variables do not have to have the same number of levels. In this example, female has two levels male and female and ses has three levels low, medium and high. Please see the results from the chi squared example above. One-way ANOVA A one-way analysis of variance ANOVA is used when you have a categorical independent variable with two or more categories and a normally distributed interval dependent variable and you wish to test for differences in the means of the dependent variable broken down by the levels of the independent variable. For example, using the hsb2 data file , say we wish to test whether the mean of write differs between the three program types prog. The command for this test would be: The mean of the dependent variable differs significantly among the levels of program type. However, we do not know if the difference is between only two of the levels or all three of the levels. The F test for the Model is the same as the F test for prog because prog was the only variable entered into the model. If other variables had also been entered, the F test for the Model would have been different from prog. From this we can see that the students in the academic program have the highest mean writing score, while students in the vocational program have the lowest.

Chapter 7 : Parametric and Non-parametric tests for comparing two or more groups | Health Knowledge

Only if we are prepared to make the additional assumption that the difference in the two groups is simply a shift in location (that is, the distribution of the data in one group is simply shifted by a fixed amount from the other) can we say that the test is a test of the difference in medians.

Nonparametric Correlations General Purpose Brief review of the idea of significance testing. To understand the idea of nonparametric statistics the term nonparametric was first used by Wolfowitz, first requires a basic understanding of parametric statistics. Elementary Concepts introduces the concept of statistical significance testing based on the sampling distribution of a particular statistic you may want to review that topic before reading on. In short, if we have a basic knowledge of the underlying distribution of a variable, then we can make predictions about how, in repeated samples of equal size, this particular statistic will "behave," that is, how it is distributed. Now imagine that we take an additional sample in a particular city "Tallburg" where we suspect that people are taller than the average population. Are most variables normally distributed? In the above example we relied on our knowledge that, in repeated samples of equal size, the standardized means for height will be distributed following the t distribution with a particular mean and variance. However, this will only be true if in the population the variable of interest height in our example is normally distributed, that is, if the distribution of people of particular heights follows the normal distribution the bell-shape distribution. For many variables of interest, we simply do not know for sure that this is the case. For example, is income distributed normally in the population? The incidence rates of rare diseases are not normally distributed in the population, the number of car accidents is also not normally distributed, and neither are very many other variables in which a researcher might be interested. For more information on the normal distribution, see Elementary Concepts ; for information on tests of normality, see Normality tests. Another factor that often limits the applicability of tests based on the assumption that the sampling distribution is normal is the size of the sample of data available for the analysis sample size; n. We can assume that the sampling distribution is normal even if we are not sure that the distribution of the variable in the population is normal, as long as our sample is large enough e. However, if our sample is very small, then those tests can be used only if we are sure that the variable is normally distributed, and there is no way to test this assumption if the sample is small. Applications of tests that are based on the normality assumptions are further limited by a lack of precise measurement. For example, let us consider a study where grade point average GPA is measured as the major variable of interest. Is an A average twice as good as a C average? Is the difference between a B and an A average comparable to the difference between a D and a C average? Somehow, the GPA is a crude measure of scholastic accomplishments that only allows us to establish a rank ordering of students from "good" students to "poor" students. This general measurement issue is usually discussed in statistics textbooks in terms of types of measurement or scale of measurement. Without going into too much detail, most common statistical techniques such as analysis of variance and t- tests , regression, etc. However, as in our example, this assumption is very often not tenable, and the data rather represent a rank ordering of observations ordinal rather than precise measurements. Parametric and nonparametric methods. Specifically, nonparametric methods were developed to be used in cases when the researcher knows nothing about the parameters of the variable of interest in the population hence the name nonparametric. In more technical terms, nonparametric methods do not rely on the estimation of parameters such as the mean or the standard deviation describing the distribution of the variable of interest in the population. Therefore, these methods are also sometimes and more appropriately called parameter-free methods or distribution-free methods. To index Brief Overview of Nonparametric Methods Basically, there is at least one nonparametric equivalent for each parametric general type of test. In general, these tests fall into the following categories: Tests of differences between groups independent samples ; Tests of differences between variables dependent samples ; Tests of relationships between variables. Differences between independent groups. Usually, when we have two samples that we want to compare concerning their mean value for some variable of interest, we would use the t-test for independent samples ; nonparametric alternatives for this test are the Wald-Wolfowitz runs test, the Mann-Whitney U test, and the Kolmogorov-Smirnov two-sample test. Differences between dependent groups. If the variables of interest are dichotomous in nature i. If there are more than two variables that were measured in the same sample, then we would customarily use repeated measures ANOVA. Cochran Q is particularly useful for measuring changes in frequencies proportions across time. To express a relationship between two variables one usually computes the correlation coefficient. Nonparametric equivalents to the standard correlation coefficient are Spearman R , Kendall Tau , and coefficient Gamma see Nonparametric correlations. If the two variables of interest are categorical in nature e. In addition, a simultaneous test for relationships between multiple cases is available: Kendall coefficient of concordance. This test is often used for expressing inter-rater agreement among independent judges who are rating ranking the same stimuli. For example, in the area of psychometrics it is well known that the rated intensity of a stimulus e. In this example, the simple mean rating sum of ratings divided by the number of stimuli is not an adequate summary of the average actual intensity of the stimuli. In this example, one would probably rather compute the geometric mean. Nonparametrics and Distributions will compute a wide variety of measures of location mean , median , mode , etc. When to Use Which Method It is not easy to give simple advice concerning the use of nonparametric procedures. Each nonparametric procedure has its peculiar sensitivities and blind spots. For example, the Kolmogorov-Smirnov two-sample test is not only sensitive to differences in the location of distributions for example, differences in means but is also greatly affected by differences in their shapes. The Wilcoxon matched pairs test assumes that one can rank order the magnitude of differences in matched observations in a meaningful manner. If this is not the case, one should rather use the Sign test. In general, if the result of a study is important e. On the other hand, nonparametric statistics are less statistically powerful sensitive than their parametric counterparts, and if it is important to detect even small effects e. Large data sets and nonparametric methods. Nonparametric methods are most appropriate when the sample sizes are small. When the data set is large e. Elementary Concepts briefly discusses the idea of the central limit theorem. In a nutshell, when the samples become very large, then the sample means will follow the normal distribution even if the respective variable is not normally distributed in the population, or is not measured very well. Thus, parametric methods, which are usually much more sensitive i. However, the tests of significance of many of the nonparametric statistics described here are based on asymptotic large sample theory; therefore, meaningful tests can often not be performed if the sample sizes become too small. Please refer to the descriptions of the specific tests to learn more about their power and efficiency. To index Nonparametric Correlations The following are three types of commonly used nonparametric correlation coefficients Spearman R , Kendall Tau , and Gamma coefficients. Note that the chi-square statistic computed for two-way frequency tables , also provides a careful measure of a relation between the two tabulated variables, and unlike the correlation measures listed below, it can be used for variables that are measured on a simple nominal scale. Spearman R can be thought of as the regular Pearson product moment correlation coefficient , that is, in terms of proportion of variability accounted for, except that Spearman R is computed from ranks. Kendall tau is equivalent to Spearman R with regard to the underlying assumptions. It is also comparable in terms of its statistical power. However, Spearman R and Kendall tau are usually not identical in magnitude because their underlying logic as well as their computational formulas are very different. Siegel and Castellan express the relationship of the two measures in terms of the inequality: More importantly, Kendall tau and Spearman R imply different interpretations: Spearman R can be thought of as the regular Pearson product moment correlation coefficient, that is, in terms of proportion of variability accounted for, except that Spearman R is computed from ranks. Kendall tau, on the other hand, represents a probability, that is, it is the difference between the probability that in the observed data the two variables are in the same order versus the probability that the two variables are in different orders. In terms of the underlying assumptions, Gamma is equivalent to Spearman R or Kendall tau; in terms of its interpretation and computation it is more similar to Kendall tau than Spearman R. In short, Gamma is also a probability; specifically, it is computed as the difference between the probability that the rank ordering of the two variables agree minus the probability that they disagree, divided by 1 minus the probability of ties. Thus, Gamma is basically equivalent to Kendall tau, except that ties are explicitly taken into account.

DOWNLOAD PDF USING NON-PARAMETRIC TESTS TO MEASURE THE DIFFERENCE BETWEEN INDEPENDENT GROUPS

Chapter 8 : How to Analysis Data with Low Quality or Small Samples, Nonparametric Statistics

Reasons to Use Parametric Tests. Reason 1: Parametric tests can perform well with skewed and nonnormal distributions. This may be a surprise but parametric tests can perform well with continuous data that are nonnormal if you satisfy the sample size guidelines in the table below.

These computations only make sense when the data are numeric. In this topic we present statistical tests that only use the ordinal information in the data. These tests do not involve adding, subtracting, dividing or multiplying the numbers in the data. They only involve using the order of the numbers. We discuss two tests: This test is used when we obtain ordinal data in the independent groups situation. Wilcoxon Signed-Ranks Test This test is used when we obtain ordinal data in the paired samples situation. Mann-Whitney U-Test The Mann-Whitney U-test evaluates the difference between two treatments two populations using data from the independent measures design. A real difference between two treatments should make the scores in one group generally larger than those in the other. If there is no effect, then large and small scores should be mixed together. All observations from the two groups are combined and rank ordered. The Mann-Whitney U is calculated. Each observation in Group A gets a point for every observation in Group B that it is larger than. The total number of points for Group A is calculated. Each observation in Group B gets a point for every observation in Group A that it is larger than. The total number of points for Group B is calculated. The smaller of these two sums is the Mann-Whitney U. The hypothesis test for the Mann-Whitney U is performed. The two groups are identically distributed. The two groups are not identically distributed. The value of U is looked up in a table to determine its significance, or the computer calculates the significance. These data are about the manual dexterity of three-year-old children. A sample of 13 children was obtained, 4 boys and 8 girls. They were asked to place a set of blocks into a specified pattern. The time in seconds required by each child to arrange the blocks was recorded. These data are from Gravetter and Walnau, Ed. These data were analyzed with the Univariate Analysis module of ViSta. The report is shown below. The Mann-Whitney suggests that the difference between the two groups is border-line significant. Wilcoxon Signed-Ranks Test The Wilcoxon Signed-ranks test evaluates the difference between two treatments, using data from a paired-samples repeated-measures design. Since we have paired samples, we calculate the difference scores. If the treatment had an effect, the scores in one treatment would be consistently larger than those in the other, producing difference scores that are consistently positive or consistently negative. If there is no effect, then we would expect positive and negative differences to be intermixed evenly. The Wilcoxon test uses the signs and ranks of the differences to decide on the significance of the differences. Calculations Rank the absolute values of the difference scores. Separate the ranks into those associated with positive differences and those associated with negative differences. Sum the ranks for the positive differences and sum the ranks for the negative differences. The smaller sum is the Wilcoxon signed ranks test statistic, identified as the Wilcoxon T not to be confused with the regular T-Test T value. The hypothesis test for the Wilcoxon T is performed. The value of T is looked up in a table to determine its significance, or the computer calculates the significance. These data are about an intensive campaign conducted by the Red Cross to increase blood donations. The campaign concentrated on 10 businesses. In each company the goal was to increase the percentage of employees who participated. The data are the percentage participation. These data are from Gravetter and Wallnau 4th Ed. The Wilcoxon test suggests that the difference between the two groups is significant.

Chapter 9 : Nonparametric Tests

For example, when comparing two independent groups in terms of a continuous outcome, the null hypothesis in a parametric test is H 0: Î¼ 1 =Î¼ 2. In a nonparametric test the null hypothesis is that the two populations are equal, often this is interpreted as the two populations are equal in terms of their central tendency.

All Modules When to Use a Nonparametric Test Nonparametric tests are sometimes called distribution-free tests because they are based on fewer assumptions e. Parametric tests involve specific probability distributions e. The cost of fewer assumptions is that nonparametric tests are generally less powerful than their parametric counterparts i. It can sometimes be difficult to assess whether a continuous outcome follows a normal distribution and, thus, whether a parametric or nonparametric test is appropriate. There are several statistical tests that can be used to assess whether data are likely from a normal distribution. Each test is essentially a goodness of fit test and compares observed data to quantiles of the normal or other specified distribution. The null hypothesis for each test is H0: Data follow a normal distribution versus H1: Data do not follow a normal distribution. If the test is statistically significant e. It should be noted that these tests for normality can be subject to low power. Specifically, the tests may fail to reject H0: Data follow a normal distribution when in fact the data do not follow a normal distribution. Low power is a major issue when the sample size is small - which unfortunately is often when we wish to employ these tests. The most practical approach to assessing normality involves investigating the distributional form of the outcome in the sample using a histogram and to augment that with data from other studies, if available, that may indicate the likely distribution of the outcome in the population. There are some situations when it is clear that the outcome does not follow a normal distribution. Using an Ordinal Scale Consider a clinical trial where study participants are asked to rate their symptom severity following 6 weeks on the assigned treatment. Symptom severity might be measured on a 5 point ordinal scale with response options: Symptoms got much worse, slightly worse, no change, slightly improved, or much improved. Distribution of Symptom Severity in Total Sample The distribution of the outcome symptom severity does not appear to be normal as more participants report improvement in symptoms as opposed to worsening of symptoms. When the Outcome is a Rank In some studies, the outcome is a rank. APGAR scores generally do not follow a normal distribution, since most newborns have scores of 7 or higher normal range. When There Are Outliers In some studies, the outcome is continuous but subject to outliers or extreme values. For example, days in the hospital following a particular surgical procedure is an outcome that is often subject to outliers. Suppose in an observational study investigators wish to assess whether there is a difference in the days patients spend in the hospital following liver transplant in for-profit versus nonprofit hospitals. The number of days in the hospital are summarized by the box-whisker plot below. Recall from page 8 in the module on Summarizing Data that we used Q In the box-whisker plot above, Limits of Detection In some studies, the outcome is a continuous variable that is measured with some imprecision e. For example, some instruments or assays cannot measure presence of specific quantities above or below certain limits. HIV viral load is a measure of the amount of virus in the body and is measured as the amount of virus per a certain volume of blood. It can range from "not detected" or "below the limit of detection" to hundreds of millions of copies. Thus, in a sample some participants may have measures like 1,, or , copies and others are measured as "not detected. Hypothesis Testing with Nonparametric Tests In nonparametric tests, the hypotheses are not about population parameters e. Instead, the null hypothesis is more general. For example, when comparing two independent groups in terms of a continuous outcome, the null hypothesis in a parametric test is H0: In a nonparametric test the null hypothesis is that the two populations are equal, often this is interpreted as the two populations are equal in terms of their central tendency. Advantages of Nonparametric Tests Nonparametric tests have some distinct advantages. With outcomes such as those described above, nonparametric tests may be the only way to analyze these data. Outcomes that are ordinal, ranked, subject to outliers or measured imprecisely are difficult to analyze with parametric methods without making major assumptions about their distributions as well as decisions about coding some values e. As described here, nonparametric tests can also be relatively simple to conduct.