INTERNATIONAL JOURNAL OF MEDICAL AND APPLIED SCIENCES

Earthjournals Publisher

E-ISSN:2320-3137
www.earthjournals.org

REVIEW ARTICLE

Applications of Parametric and Non Parametric Statistic in Medical and Agricultural Sciences

Ajay S. Singh

Department of AEM, Faculty of Agriculture and Consumer Sciences, University of Swaziland, Luyengo M205, Swaziland.

Corresponding Author: Dr. Ajay S. Singh

Publication history: Received on 13/3/2016 Published online 17/04/2016

ABSTRACT:
Advance applied science researches have experienced a dramatic change in knowledge and an exponential increase in technology. A lot of these technical developments involve medical and agricultural researches and these researches deal with groups rather than individual cases and usually experimental and field study. The goal of applied research is to provide data to support existing knowledge by filling information gaps or develop new ideas. Medical research requires proper study design, management, data collection and analysis to obtain statistically sound results.

Applied statistical techniques have significant role in analysis of observational facts in medical sciences, agricultural sciences, social sciences, and business management. Many investigations are based on survey research. The inferential statistics especially describe the sampling, estimation and testing of hypothesis. Testing of hypothesis have crucial role in survey-based research problems in applied sciences. Hypothesis is used for specific research problems because one technique may not be appropriate for all problems. Similarly, if the hypothesis is not appropriate it may lead to erroneous conclusions. This paper first gives an idea of hypothesis and associated terms type of error, level of significance, degree of freedom, power of the test, confidence interval etc. The present paper mainly describes the applications of parametric and non parametric statistic in simple manner with hypothetical data. This paper is useful for medical as well as applied science researcher for better understanding about the applications and interpretation of results based on observations.

Key words: Power of the test, confidence interval, level of Significance, parametric test, non parametric test.

INTRODUCTION

Statistics is important in the area of social science, agriculture, medical, engineering, etc. because it provides tools to analyseand interpret the observational facts. Researchers frequently use statistics to analyseobservations. Statistics provides scientific methods for appropriate tools for data collection, analysis and summarization of observation and inferential statistical methods for drawing conclusions in the face of uncertainty. Statistical methodologies have wide applicability to almost any branch of science dealing with the study of uncertain phenomena.

Statistics can help understand a phenomenon by accepting or rejecting a hypothesis. It is vital to know how we acquire information to most scientific manners. The statistical methodof data analysis depends upon the objective for collection of observations. This is not only more convenient, but would offer some insight into the appropriate decisions for the planning and management.

In the 17th century Captain John Graunt of London found the origin of vital statistics, known as father of vital statistics. Captain John was the first man to study the statistics of birth and death to calculate life expectancy and he also constructed the life table [1].

Limitations of Statistics:

1. Statistics is not more appropriate to the study of qualitative phenomenon. Statistics being a science with a set of numerical observations associated with quantitative measurement.
2. Statistics does not study individuals. Statistics deals with an aggregate of objects and does not give any special importance to the individual of a series.
3. Statistical laws are not exact like as physical or natural law of sciences, statistical analysis is only in terms of probability and chance not an exact.
4. Statistics is liable to be misused. The most important limitation of statistics is that it must be used by experts
5. Statistical methods are more dangerous tools in the hand of the inexpert.

Basic steps in the proposals of research:

1. Definition of research problems,
2. Formulation of objectives and hypothesis.
3. Methodology of research for the particular problems.
4. Selection of variables for the research study.
5. Coverage of all possible subject matters related with research objectives.
6. Well defined tools and techniques for data analysis.
7. Well defined study population, sample, control, sample size and time coverage.
8. Formulation of analytical methods for the data and planning for resources.
9. Anticipation and estimation of possible errors and evolving appropriate actions to rectify the errors.

Statistical Inference:

Inferential statistics or statistical inference includes the testing of hypothesis which is essential and important part of research investigations. In traditional statistical hypothesis testing, the statistician starts with a null hypothesis and an alternative hypothesis, performs an experiment, and then decides whether to reject the null hypothesis in favour of the alternative. In other words hypothesis is a numerical statement about the parameter [2].

POPULATION, RESEARCH AND HYPOTHESIS:

Population:

In statistics, a population is a set of similar items or units which is of interest for some objective or experiment. A statistical population can be defined as a group of existing objects or a hypothetical and potentially infinite group of objects.A main aim of statistical analysis is to produce information about some selected or defined population.

In inferential statistics, a subset of the population (sample) is chosen to represent the population in a statistical analysis. If a sample is selectedappropriately, characteristics of the entire population that the sample is drawn from can be estimated from corresponding characteristics of the sample [3].

Research:

Research is an inseparable part of human knowledge. Life would lose its taste without research. Research is process of analyzing and solving problem, which add new idea and develop the mechanism as well as collect the information to test the significance for generalization. Research should never be treated as compilation of work, research is always expected some thing new, original, advance and innovative. The main objective of applied research is to find out the solutions of existing problems through scientific approaches as well as highlight the truth which is hidden and which has not be known.

Hypothesis:

Hypothesis is a proposition temporarily accepted as true in the light of what is, at the time, known about a phenomenon. It is a tool for action in the search of truth. Lundberg defines hypothesis as "a tentative generalization, the validity of which remains to be tested". In basic stages, the hypothesis may be any hunch, guess, imaginative idea which becomes a basis for investigation [3].

Characteristics of Hypothesis [3]:

The important characteristics of usable hypothesis are given below-
i. The hypothesis should be empirically testable
ii. The hypothesis should be conceptually clear and simple
iii. The hypothesis should be specific.
iv. The hypothesis should be related to the body of theory
v. The hypothesis should be related to available techniques

Utility of Hypothesis:

i. Hypothesis acts as a guide.
ii. Hypothesis spells out the differences between precision and haphazard, between fruitful and fruitless research.
iii. It provides the direction to research, identifying of which factor is relevant, it also prevent irrelevantinformation and literature.
iv. It links up related factors and information in fully understandable.
v. Hypothesis serves as a framework for drawing meaningful conclusions.

Formulation of statistical hypothesis:

The first step in hypothesis testing is to state the null hypothesis ($H_{o)}$ which follows logically from alternative hypothesis (H_1) [4, 5]. Alternative hypothesis define the research statement in positive terms [4]. Acceptance or rejection of null hypothesis based on our statistical testing parametric or non parametric methodologies [2, 5, 6]. If null hypothesis (H_0) is accepted, then H_1 must be rejected and vice versa due to that hypothesis are mutually exclusive. If H_0 is accepted, this concludes that no statistical differences exist and if any differences in groups or observations are due to only chance or due to sampling fluctuations. In other hand, if H_0 is rejected or H_1 is accepted this indicates that a significant difference exits and the differences are not only due to chance or sampling fluctuations.

Hypothesis Testing and Statistical Error:

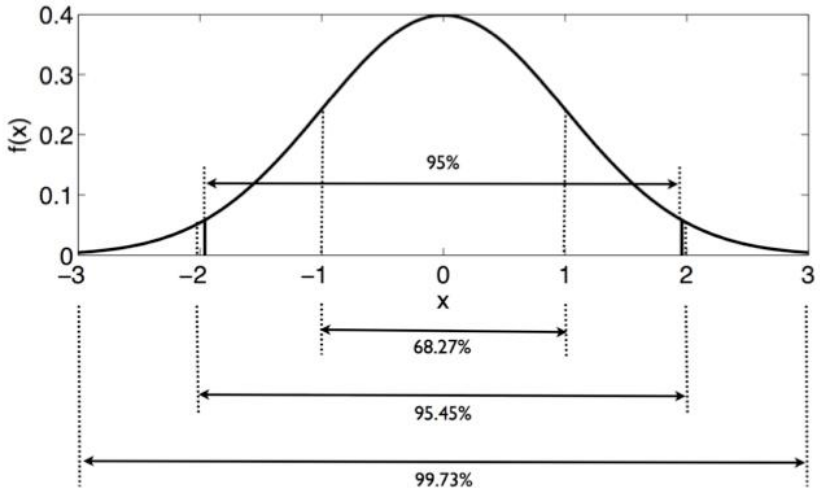
There are two types of error or incorrect conclusions possible in hypothesis testing and possibilities in which the statistical test falsely indicates that significant differences exists between the two or more groups and also analogously to a wrong positive results. Rejection of null hypothesis ($H_{0)}$ when it is true is called as Type I error and acceptance of null hypothesis (H_0)when it is false and it is known as Type II error and Type II error is more harmful then Type I error [3-5].

The probability of Type I error is known as level of significance (α) and the probability of Type II error is known as the power of the test β or (1- α) [3-5]. By convention, statistical significance is generally accepted if the probability of making type I error is less than 0.05, which is commonly denoted as $p < 0.05$ [3, 6-7]. The probability of type II error is more difficult to derive than probability of type I error, actually it is not one single probability value. The probability of type II error (β) is often ignored by researcher [3, 8]. The probability of type I error (α) and probability of type II error (β) are inter-related. As α arbitrarily decreased, β is increased. Similarly, α is increased, β is decreased [2-4].

Statistical Power:
Statistically power indicates mathematically the probability of not making a type II error. Statistical Power is defined as ($1-\beta$). β indicates the probability of making II error and if sample size increases, power increases [3,5,7].
Power is analogous to sensitivity in hypothesis testing. Sensitivity indicates the probability that the diagnostic test can detect disease when it present. Power indicates the probability that the statistical test can detect significant differences, when in fact such differences truly exist.

p - value:
The p value is the probability to observe effects as big as those seen in the study if there is really no difference between the groups or treatments. The reasoning of hypothesis testing and p values is convoluted. The p values helps to assessing whether this apparent effect is likely to be actual or could just by chance or sampling fluctuation.The p value gives the magnitude of difference present between populations. In calculation of p values, first assume that no true difference between the two groups/treatments.The p values allow the assessment of findings that are significantly different or not statistically different. If the p value is small, the findings are unlikely to have arisen by chance or sampling fluctuation, reject the null hypothesis. If the p is large, the observed difference is plausibly chance finding, we do not reject the null hypothesis. By convention, p value of less than 5% is considered small or significant. Sometimes p value is less than 1% or 0.01, called as highly significant [3, 5, 9].

Confidence Interval:
Confidence interval, like p values, provides a guide to help the interpretation of research findings in the light of the probability. Confidence interval describes the different information from that arising in the hypothesis test. Confidence interval provides a range about the observed effect size. The formal definition of confidence interval is a range of values for a variable of interest constructed so that this range has a specified probability is called the confidence level, and the end points of confidence interval are called the confidence limits [10]. By conventional, confidence interval at the 95% corresponds to hypothesis testing with p values, with a cut off for p is less than 0.05 [3, 5, 11].Graph clearly showing the concept of confidence interval.



One tailed and two tailed test[3]:

A test of any statistical hypothesis where the alternative hypothesis is one tailed (right tailed/ left tailed) is called a one tailed test. For example, A test for testing the mean of a population
Null Hypothesis (H_0): $\mu = \mu_0$ against the alternative hypothesis
Alternative Hypothesis (H_1): $\mu > \mu_0$ (Right tailed test) or H_1: $\mu < \mu_0$ (Left tailed test)
A test of statistical hypothesis where the alternative hypothesis is two tailed such as
H_0: $\mu = \mu_0$ against the alternative hypothesis
H_1: $\mu \neq \mu_0$ (two tailed test) where $\mu < \mu_0$ and $\mu > \mu_0$

Degree of freedom:

The number of independent variates which make up the test statistic is known as degree of freedom. The number of degree of freedom, in general, is the total number of observations less the number of independent constraints imposed on the observations.

Parametric test And Non parametric test:

Parametric statistics is a part of inferential statistics that assumes that the data have come from a type of probability distribution and makes inferences about the parameters of the distribution [12]. Most well-known elementary statistical methods are parametric [13]. Parametric test have more statistical power. Generally speaking parametric methods have more assumptions than non-parametric methods[14]. If the extra assumptions are correct, parametric methods can provide more accurate and precise estimates. However, if the assumptions are incorrect, parametric methods can provide very misleading results. On the other hand, parametric formulas are often simpler to write down and faster to compute. In some cases, but definitely not all cases, their simplicity makes up for their non-robustness, especially if care is taken to examine diagnostic statistics [15].

Non-parametric statistics:

Nonparametric statistics are those data that do not assume a prior distribution. When an experiment is performed or observations collected for some objective, it is usually assumed that it fits some given probability distribution, generally the normal distribution. This is the basis on which the data is interpreted. When these assumptions are not made, it becomes nonparametric statistics.

There are many advantages of using nonparametric statistics. As can be expected, since there are fewer assumptions that are made about the sample being studied, nonparametric statistics are usually wider in scope as compared to parametric statistics that actually assume a distribution. This is mainly the case when we do not know a lot about the sample we are studying and making a priori assumptions about data distributions might not give us accurate results and interpretations. This directly translates into an increase in robustness.

There are also some disadvantages of nonparametric statistics. The main disadvantage is that the degree of confidence is usually lower for these types of studies. This means for the same sample under consideration, the results obtained from nonparametric statistics have a lower degree of confidence than if the results were obtained using parametric statistics. Of course, this is assuming that the study is such that it is valid to assume a distribution for the sample.

In other words the non-parametric covers techniques that do not rely on data belonging to any particular distribution. Meaning that non parametric statistic is distribution free methods, which do not rely on assumptions that the data are drawn from a given probability distribution. As such it is the opposite of parametric statistics. It includes non-parametric statistical models, inference and statistical tests.

Non-parametric statistic, whose interpretation does not depend on the population fitting of any parametric distributions. Statistics based on the ranks of observations are one example of non parametric statistics and ranks also play a central role in many non-parametric approaches. These techniques include, Chi-square test, Wilcoxon Signed Rank test, Mann Whitney U test.

Parametric Tests and Application:

Student's t- Test

In 1908, 'Student' derived a new distribution and test statistic known as t. The value of t is dependent upon the sample size 'n' and for each value of n-1, (degree of freedom used for estimating the standard deviation of sample). The student's t-test is a statistical method that is used to test if two sets of data differ significantly.

(i) t-test or t statistics is calculated as a ratio of the difference between the two means to the standard error of the difference. The t-test is applicable for small samples ($n < 30$) and for quantitative data.

t- test = (Mean of X – Mean of Y) / s_{md} where X and Y are two sample mean.

s_{md} is the standard error of the difference of two sample means

where: $s_{md} = s_d / \sqrt{n}$

In the actual research experiment, the observations may be carried out on two independent samples one known as the control group and the other known as the treated group. In such cases the comparisons are defined as unpaired comparison.

(ii) t-test for comparing paired observation- In this case t-statistics is [Mean of d / s_{md}] with (n-1) degree of freedom. Whereas the mean of 'd' is the difference in the values of the variable before and after exposure or treatment and n is the number of observations in the sample.

$S_{md} = s_{d\,/}\sqrt{n}$ wheres_d is the standard deviation of the values of d_i. For the calculation of $s_d = \sqrt{[(d_i - \text{Mean of d})^2 / (n-1)]}$

After the computation of t-statistics compare the value with t distribution table at α (conventionally, 1% or 5%) level of significance with (n-1) degree of freedom.

In the other situation, for comparing the means of two independent sample t-statistics is equal to (Mean of X – Mean of Y) / s_{md} with (n_1 + n_2 -2) degree of freedom. Here $s_{md}$is the estimated standard error of the difference between the two sample means.
$S_{md} = \sqrt{}[(n_1 + n_2)/ (n_1 . n_2)]*[\{(n_1 – 1) s_1^2 + (n_2 – 1) s_2^2 \}/(n_1 + n_2 – 2)]$ where s_1^2 and s_2^2 are the standard deviations of the two samples and n_1 and n_2 are their respective sample sizes.
If in experiment two sample sizes are equal ($n_1 = n_2$). Therefore,
$s_{md} = [(s_1^2 + s_2^2)/ n]$ and degree of freedom is (2 n -2).

Applications on hypothetical observations:

(**i**)A hypothetical observations of ten plantsfrom the large group of plants (as a population) with mean length is = 0.7 and mean length of sample is 0.742 and standard deviation is 0.4. To test the differences between populationmean length and sample mean length of plants are significant or insignificant.
PopulationMean = 0.7 and Sample Mean = 0.74 and S.D. = 0.4
Null Hypothesis (H_0) : μ = 0.7 and Alternative Hypothesis (H_1) : μ ≠ 0.7
$t_{cal.}$= {Mean of x – μ }/ SQRT [s^2 / n – 1]
Degree of freedom (d.f.) = 9 and Level of Significance (α) = 0.05
Therefore,$t_{cal.}$ = 1.05according to t-statistic and $t_{tab.}$with 9 d.f. at α=5% = 1.83
On the basis of comparison between $t_{cal..}$and$t_{tab.}$, mean length differences are insignificant.
(**ii**)Hypothetical observations of 12 anaemic patients, how to test the differences in Hb. levels of patients before and after treatment.

Patient No.	1	2	3	4	5	6	7	8	9	10	11	12
Hb level(Before Treatment)	4.8	6.5	7.5	5.6	3.5	4.5	6.7	8.0	7.5	4.4	5.6	6.2
Hb, level (After Treatment)	9.4	11.0	7.5	10.2	6.0	7.5	8.0	8.0	8.0	8.4	10.0	9.6
Difference (d_i)	4.6	4.5	0	4.6	2.5	3.0	1.3	0	0.5	4.0	4.4	3.4

Null hypothesis (H_0): No significant impact of treatment on anaemic patients
Mead difference in Hb level=2.73, Standard Deviation of the difference = 1.83 and Standard Error of the mean difference is 2.73/12= 0.53 Therefore, $t_{cal.}$ = 5.2 and $t_{tab.}$with 11 degree of freedom at 5% level of significance is 1.8. On the basis of comparison of $t_{cal..}$and$t_{tab..}$, null hypothesis is not accepted and conclude that treatment have significant impact on anaemic patients.

Z-Test

Z-test is a statistical test where normal distribution is applied and is basically used for dealing with problems relating to large samples when n ≥ 30 where n = sample size

Uses of Z-Test's for Different Purposes

There are different types of Z-test each for different purpose. Some of the popular types are outlined below:

1. z test for single proportion is used to test a hypothesis on a specific value of the population proportion.

Statistically speaking, we test the null hypothesis H_0: $p = p_0$ against the alternative hypothesis H_1: $p \neq p_0$ where p is the population proportion and p_0 is a specific value of the population proportion we would like to test for acceptance.

2. z test for difference of proportions is used to test the hypothesis that two populations have the same proportion.

3. z -test for single mean is used to test a hypothesis on a specific value of the population mean.

Statistically, to test the null hypothesis H_0: $\mu = \mu_0$ against the alternative hypothesis H_1: $\mu \neq \mu_0$ where μ is the population mean and μ_0 is a specific value of the population that we would like to test for acceptance.

Unlike the t-test for single mean, this test is used if $n \geq 30$ and population standard deviation is known.

4. z test for single variance is used to test a hypothesis on a specific value of the population variance.

Statistically speaking, we test the null hypothesis H_0: $\sigma = \sigma_0$ against H_1: $\sigma \neq \sigma_0$ where σ is the population mean and σ_0 is a specific value of the population variance that we would like to test for acceptance.

In other words, this test enables us to test if the given sample has been drawn from a population with specific variance σ_0. Unlike the chi square test for single variance, this test is used if $n \geq 30$.

5. z test for testing equality of variance is used to test the hypothesis of equality of two population variances when the sample size of each sample is 30 or larger.

Assumptions

Irrespective of the type of z-test used, it is assumed that the populations from which the samples are drawn are normal.

Applications on hypothetical observations:

(i)

Mean of X_1 = 2000 S.D.$_1$ = 192 n_1 = 1000

Mean of X_2 = 2110 S.D.$_2$ = 225 n_2 = 1000

Null Hypothesis (H_0): No difference in the sample mean

Alternative Hypothesis (H_1) :Mean of $X_1 \neq$ Mean of X_2

S. E. of Mean of X_1 – Mean of X_2 = SQRT [σ^2 { $(\sigma_1^2/n1)$ - $(\sigma_2^2/n2)$}] = 9.35

Z = (2000 – 2110) / 9.35 = - 11.76

On the basis of comparison, null hypothesis (H_0) is rejected. Therefore, differences in two means are insignificant.

(ii)

Rural community are consuming water from two different resources. Cases of diarrhoea reported in this community. Find out the impact of water on proportion of diarrhoea cases reported in the community.

Source of Water	No. of people consuming water	No. of cases reported
Source-I	800	35
Source-II	2400	120
Total	3200	155

Null Hypothesis (H_0): Differences in proportion of cases in the two groups are insignificant.

Proportion of cases in Source-I (p_1)= 0.0444 q_1= 1-p_1= 0.96
Standard error of the proportion p_1 (SE_1)= √ {p_1 q_1 /n_1}= 0.0072
Proportion of cases in Source-II (p_2)= 0.050 q_2= 1-p_2= 0.95
Standard error of the proportion p_2 (SE_2)= √ {p_2 q_2 /n_2} = 0.0044
Difference between the proportions=0.006 ----------------------------------- (1)
Standard error of the difference = $\sqrt{(SE_1)^2 + (SE_2)^2}$ = 0.0085 --------- (2)
Therefore, Z = (1)/(2) = 0.71
On the basis of comparison with value of Z from normal distribution, null hypothesis is accepted and conclude that differences are insignificant.

F-test:
Any statistical test that uses F-distribution can be called a F-test. It is used when the sample size is small i.e. $n < 30$.
However one assumption of t-test is that the variance of the two populations is equal- here two populations are the population of heights of male and female students. Unless this assumption is true, the t-test for difference of means cannot be carried out.
The F-test can be used to test the hypothesis that the population variances are equal.
F-test for different purposes
There are different types of t-tests each for the different purposes. Some of the popular types are outlined below.

1. F-test for testing equality of variance is used to test the hypothesis of equality of two population variances.
2. F-test for testing equality of several means. Test for equality of several means is carried out by the technique named Analysis of Variance (ANOVA).
3. To test if there are significant differences among the three levels of the drug in terms of efficacy, the ANOVA technique has to be applied. The test used for this purpose is the F-test.
4. F-test for testing significance of regression is used to test the significance of the regression model. The appropriateness of the multiple regression models as a whole can be tested by this test. A significant F indicates a linear relationship between Y and at least one of the X's.

Assumptions
Irrespective of the type of F-test used, one assumption has to be met. The populations from which the samples are drawn have to be normal. In the case of F-test for equality of variance, a second assumption has to be satisfied in that the larger the sample variance has to be placed in the numerator of the test statistic.
Like t-test, F-test is also a small sample test and may be considered for use if sample size is < 30.
Testing
In attempting to reach decisions, we always begin by specifying the null hypothesis against a complementary hypothesis called alternative hypothesis. The calculated value of the F-test with its associated p-value is used to infer whether one has to accept or reject a null hypothesis.
If the associated p-value is small i.e. (<0.05) we say that the test is significant at 5% and one may reject the null hypothesis and accept the alternative one.

On the other hand if associated p-value of the test is >0.05, one may accept the null hypothesis and reject the alternative. Evidence against the null hypothesis will be considered very strong if p-value is less than 0.01. In that case the test is significant at 1%.

Uses

The main use of F-distribution is to test whether two independent samples have been drawn for the normal populations with the same variance, or if two independent estimates of the population variance are homogeneous or not, since it is often desirable to compare two variances rather than two averages. For instance, college administrators would prefer two college professors grading exams to have the same variation in their grading. For this, the F-test can be used, and after examining the p-value, inference can be drawn on the variation.

Assumptions

In order to perform F-test of two variances, it is important that the populations from which the two samples are drawn are normally distributed. The two populations are independent of each other.

If the two populations have equal variances, then s_1^2 and s_2^2 are close in value and F is close to 1. But if the two population variances are very different, s_1^2 and s_2^2 tend to be very different, too.

Non parametric tests and applications[3, 5, 9]**:**

Chi-Squared Test

Chi square test is based on χ^2 distribution. It has large number of application in applied research. Generally it is used to test the goodness of fit, to test the independence of attributes and to test homogeneity of independent estimates of the population variance.

It has to be noted that the Chi square goodness of fit test and test for independence of attributes depend only on the set of observed and expected frequencies and degrees of freedom. These two tests do not need any assumption regarding distribution of the parent population from which the samples are taken.

Since these tests do not involve any population parameters or characteristics, they are also termed as non parametric or distribution free tests.

A Chi-Squared test gives an estimate of the agreement between a set of observed data and a random set of data that you expected the measurements to fit.

Chi Squared (χ^2)

The Chi squared calculation involves summing the distances between the observed and random data.

$\chi^2 = \Sigma$ [(Observed – Expected)^2 / Expected]

Calculation of expected frequencies in 2 x 2 table.

			Total
	a E(a)	b E(b)	(a+b)
	c E (c)	d E(d)	(c + d)
Total	(a+c)	(a +d)	N=a+b+c+d

E(a) = (a+b)(a+c) / N   E(b) = (a+b)(b+d) / N

E (c) = (a+c)(c+d) / N   E(d) = (b+d)(c +d) /N

χ^2 test ={ [O(a) – E(a)]^2/E(a)} + {[O(b)- E(b)]^2/E(b) +{ [O(c) – E(c)]^2 /E(c)}+{ [O(d) – E(d)]^2 / E(d)}

with(2-1)(2-1) degree of freedom with α (Conventionally, 0.05 or 0.01) level of significance.

In other derivation,

Gender	With disease	Without Disease	Total
Male	a	b	a+b

Female	c	d	c+d
Total	a+c	b+d	G

χ^2 test = { (a d – b c)^2 G}/ (a+b)(c +d)(a+c)(b+d) at 1 degree of freedom.-----(i)

Application on hypothetical observations:

Gender	With disease	Without Disease	Total
Male	28	237	265
Female	20	222	242
Total	48	459	507

On the basis of simple application of (i), χ^2 test = 0.78 on the basis comparison of χ^2 tabulated value with 1 degree of freedom at 5% level of significance is 3.84, the null hypothesis is accepted and conclude that the difference in prevalence of disease is insignificant. Means prevalence of disease not associated with gender.

Signed Test:

The significance of the difference between the two procedures can be tested using by usual χ^2 test calculated as follows.

$\chi^2 = (|a\text{-}b| -1)^2 / n$ with one degree of freedom.

Where a and b are the number of + (sign) and - (sign) respectively and n=(a+b). All zero differences are omitted for calculation, therefore n always equal to (a+b). Probability p obtained from χ^2 distribution table and conclusion about significance will be made.

Application with hypothetical observations:

Twelve patients were clinically tested with two different diagnostic approaches. Positive case indicated by 1 and negative indicated by 0 of the same patients.

Patient No.↓											
1	2	3	4	5	6	7	8	9	10	11	12
Serological Test ↓											
1	1	0	1	1	0	1	1	0	0	0	1
Parasitological Test ↓											
0	0	1	0	0	0	0	0	0	1	1	1
Difference ↓											
+	+	-	+	+	0	+	+	0	-	-	0

Therefore,

Number of (+) = a =6

Number of (-) = b = 3

n = 9, χ^2 = 0.44, Therefore, on the basis of χ^2 and p values, test indicate that the differences between diagnostic procedure is not significant.

Wilcoxon Signed Rank Test

The Wilcoxon Signed Rank Test is a non-parametric statistical test for testing hypothesis on median. The test has two versions: "single sample" and "paired samples / two samples".

Single Sample

The first version is the analogue of independent one sample t-test in the non parametric context. It uses a single sample and is recommended for use.

Paired Samples

The second version of the test uses paired samples and is the non parametric analogue of dependent t-test for paired samples.

This test uses two samples but it is necessary that they are paired. Paired samples imply that each individual observation of one sample has a unique corresponding member in the other sample.

However the test has certain assumption notable among them being normality. If this normality assumption is not satisfied, one would have to go for the non parametricWilcoxon Signed Rank Test.

Wilcoxon Signed Rank Test is that it neither depends on the form of the parent distribution nor on its parameters. It does not require any assumptions about the shape of the distribution.For this reason, this test is often used as an alternative to t test's whenever the population cannot be assumed to be normally distributed. Even if the normality assumption holds, it has been shown that the efficiency of this test compared to t-test is almost 95%.

Test statistic $z = (|\mu\text{-}T| - 1/2)/\sigma$

Where, T= smaller rank sum, $\mu = n (n+ 1)/4$, $\sigma = \sqrt{(2n +1)/4}$ and n=number of pairs of observations.

Probability p is obtained from normal distribution table corresponding to the value of calculated z for the comparison and significance.

Mann-Whitney U-Test:

The Mann-Whitney U-test is used to test whether two independent samples of observations are drawn from the same or identical distributions. An advantage with this test is that the two samples under consideration may not necessarily have the same number of observations.

This test is based on the idea that the particular pattern exhibited when 'm' number of X random variables and 'n' number of Y random variables are arranged together in increasing order of magnitude provides information about the relationship between their parent populations.

Assumptions

The test has two important assumptions. First the two samples under consideration are random, and are independent of each other, as are the observations within each sample. Second the observations are numeric or arranged by ranks.

Calculation:

First the observations in the samples are arranged in the order of magnitude taking all the observations in both the sample together. Proper tag is made to distinguish the observations of the two samples separately. Then the ranks are assigned to the combined observations. Whenever there are common observations, average of the ranks is given to them. Then the sum of ranks of each of the samples is calculated separately. The smaller rank sum out of the above two is referred to the prepared table (Mann-Whitney Table [18]) which gives the maximum sum of ranks required for the rejection of null hypothesis, under the different probability levels. If the calculated smaller rank sum is less than tabulated value the null hypothesis is rejected. When the two samples are of unequal size the smaller ranks sum is corrected as,

$T_2 = n_1 (n_1 + n_2 + 1) – T_1$, where T_1 is the rank sum of sample with smaller number of observations, n_1 and n_2 are the number of observations in bigger sample. T, the smaller of T_1

and T_2 is referred in the mentioned table. For values of n_1 and n_2 greater than 20, normal approximation is followed and Z is calculated as follow,

$Z = (|\mu - T| - 1/2) / \sigma$ where, $\mu = n_1(n_1 + n_2 + 1)/2$ and n_1 is smaller sample size than n_2. The calculated value is referred to the table of normal distribution and p value is obtained [9,18].

CONCLUSION:

Applied research is the study of the distribution and determinants of the associated factors in the specific area and time. Medical and Agricultural research is directly associated with development of nation. Medical Science is the study of the distribution and determinants of the health related events in the specific population. Health research is directly associated with, through the collection of information related to health and family welfare. Medical data analysis makes a significant contribution to emerging population-based health management

Modern medical and agricultural research linked with better living management. It requiresmultiple set of skills with different specialties. In the advance medical and agricultural data base research required appropriate statistical tools and research designs to provide the unbiased results, conclusions and appropriate interpretation. In this paper, describe the fundamentals of statistical concepts and techniques for testing of hypothesis with the help of parametric and non parametric techniques and their applications, have been emphasized scientifically in simple manner. These tests are is valid for any kind of data and especially for quantitative data with checked quality.These have brought difficult concept of fundamentals as well as parametric and non parametric techniques within the reach of all research scientists. The availability of computing power naturally makes it all the more important that the researcher applies statistical approaches in correct way.

REFERENCES:

[1]Yates Daniel S., Moore, David S. Starnes and Daren S.,The Practice of Statistics (2nd ed.). New York: Freeman, 2003.

[2] Hopkins K.D., Glass, G.V. Basic statistics for the behavioural sciences, Englewood Cliffs, New Jersey, Prentice Hall, USA, 1978.

[3] Singh A.S. and Masuku, M.B.,An insight statistical techniques and design in agricultural and applied research, World Journal of Agricultural Sciences,2012,8(6): 568-84**.**

[4] Keppel G., Design and analysis. A researcher's handbook, Englewood Cliffs, New Jersey, Prentice Hall, USA, 1978.

[5] Gupta S.C. and Kapoor V.K., Fundamental of mathematical statistics,S.C. Publication, New Delhi, India 1970.

[6]Elenbaas R.M., Elenbass J.K. and Cuddy P.G., Evaluating the medical literature Part II: Statistical Analysis, Ann. Emerg. Med., 1983, 12, 610-613

[7]Sokal R. R., Rohlf F.I.,Biometry (ed. 2), New York, WH Freeman and Co., 1981.

[8] Freeman J.A. , Chalmers T.C. and Smith H., The importance of beta, the type II error, and sample size in the design and interpretation of randomized clinical trial. New Engl. Jr. of Med., 1978, 299, 690-694.

[9]Rao N. S. N., Elements of Health Statistics, First edition, R. Publication, Varanasi, India, 1985.

[10] Last J. M. A., dictionary of epidemiology, Oxford: International Jr. of Epidemiology, 1988.

[11] Gardner M.J. and Altman D.G., Confidence intervals rather than p values: estimation rather than hypothesis testing, British Medical Journal, 1986, 292; 746-750.

[12]Seymour G. and Johnson W. M., Modes of Parametric Statistical Inference, John Wiley & Sons, 2006.

[13]Cox D. R.,Principles of Statistical Inference, Cambridge University Press, 2006.

[14]Corder and Foreman, Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach, John Wiley & Sons, 2009.

[15]David Freedman, Statistical Models: Theory and Practice, Cambridge University Press, 2005.

[16] Nutter, Jr., F.W., Understanding the interrelationships between botanical, human and veterinary epidemiology: the Ys and Rs of it all. Ecosys Health, 1999,**5** (3): 131–40.

[17] ClaytonD. and Michael H., Statistical Models in Epidemiology, Oxford University Press, USA, 1993.

[18]Snedecor, G.W. and Cochran, W.G., Statistical Methods,Oxford and IBH publishing House, New Delhi, India, 1967.

BIOGRAPHY

Dr. Ajay S. Singh:

Dr. Ajay S. Singh received his primary school education in Bihar, India and completed his intermediate studies (I. Sc.) from Uday Pratap College, Varanasi, U.P., India. He obtained his bachelor degree and master degree from the Banaras Hindu University (BHU), Varanasi, India. He received his Ph. D. degree from the Institute of Medical Sciences, BHU, India, completing his doctoral thesis on 'Human fertility behaviour through analytical modelling' in 1992. He worked as a Research Officer in the Indian Council of Medical Research Scheme in the Postgraduate Department of Pathology, S.N. Medical College, Agra for 'Oral cancer prevention program' and also served as medical data analyst for cancer registry in the same institution. He also worked as a lecturer (biostatistics) in the National Pharmaceutical Education Research (NIPER), India. Presently, he is working as a Senior lecturer in the University of Swaziland, Swaziland. He has supervised and co-supervised many undergraduate and graduate level students. Dr. Singh is the member of the editorial boards of over half a dozen international journals and reviewer for many reputable international journals. He is also member of several reputed professional organizations**.**

Paper cited as: Ajay S. Singh. Applications of Parametric and Non Parametric Statistic in Medical and Agricultural Sciences. International Journal of medical and applied Sciences, 5(1), 2016, pp.16-29