THE DERIVATION AND CHOICE OF APPROPRIATE TEST STATISTIC (Z, T, F AND CHI-SQUARE TEST) IN RESEARCH METHODOLOGY

Teshome Hailemeskel Abebe

Department of Economics, Ambo University, Ethiopia

ABSTRACT: *The main objective of this paper is to choose an appropriate test statistic for research methodology. Specifically, this article tries to explore the concept of statistical hypothesis test, derivation of the test statistic and its role on research methodology. It also try to show the basic formulating and testing of hypothesis using test statistic since choosing appropriate test statistic is the most important tool of research. To test a hypothesis various statistical test like Z-test, Student's t-test, F test (like ANOVA), Chi square test were identified. In testing the mean of a population or comparing the means from two continuous populations, the z-test and t-test were used, while the F test is used for comparing more than two means and equality of variance. The chi-square test was used for testing independence, goodness of fit and population variance of single sample in categorical data. Therefore, choosing an appropriate test statistic gives valid results about hypothesis testing.*

KEYWORDS: Test Statistic, Z-test, Student's t-test, F test (like ANOVA), Chi square test, research methodology

INTRODUCTION

The application of statistical test in scientific research has increased dramatically in recent years almost in every science. Despite, its applicability in every science, there were misunderstanding in choosing of appropriate test statistic for the given problems. Thus, the main objective of this study is to give a direction which test statistic is appropriate in conducting statistical hypothesis testing in research methodology.Choosing appropriate statistical test may be depending on the types of data (continuous or categorical), i.e. whether a t-test, z-test, F-test or a chi-square test should be used depend on the nature of data. For example the two-sample independent t-test and z-test used if the two samples are independent, a paired z-test and t-test used if two samples are dependent, while F-test can be used if the sample is independent.Therefore, testing hypothesis about the mean of a single population, comparing the means from two populations such as one-way ANOVA, testing the variance of a single population using chi-square test and test of two or several mean and variance using F-test are an important topic in this paper.

DERIVATION OF Z, t, F AND CHI-SQUARE TEST STATISTIC

Theorem 1: if we draw independent random samples, $X_1, X_2, \ldots, X_n$ from a population and compute the mean $\bar{X}$ and repeat this process many times, then $\bar{X}$ is approximately normal. Since assumptions are part of the "if" part, the conditions used to deduce sampling distribution of statistic, then the t, χ^2 and F distributions all depend on normal "parent" population.

Chi-Square Distribution

The chi square distribution is a theoretical distribution, which has wide application in statistical work. The term `chi-square' (denoted by Greek letter $\mathcal{X}^2$ and pronounced with `chi') is used to define this distribution. The chi square theoretical distribution is a continuous distribution, but the χ^2 statistic is obtained in a discrete manner based on discrete differences between the observed and expected values. Chi-square(χ^2) is the sum of independent squared normal random variables with mean $\mu = 0$ and variance $\sigma^2 = 1$ (i.e., standard normal random variables). Suppose we pick $X_1, X_2, \ldots, X_n$ from a normal distribution with mean $E(X) = \mu$ and variance, $var(X) = \sigma^2$, that is $X \sim \mathcal{N}(\mu, \sigma^2)$. Then it turns out that $Z_1, Z_2, \ldots, Z_n$ be independent standard normal variables ($Z_i \sim \mathcal{N}(0,1)$) and if the sum of independent standard random variable ($\sum_{i=1}^{n} Z_i^2 = \chi_v^2$) have a chi-square distribution with K degree of freedom given as follows:

$$Z^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$$

Based on the Central Limit Theorem, the "limit of the χ^2 distribution is normal (as $n \to \infty$). So that the chi-square test is given by

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \qquad (1)$$

The F Distribution

$\mathcal{F}_{v_1,v_2}$ is the ratio of two independent chi-squared random variables each divided by their respective degrees of freedom as given by:

$$\mathcal{F}_{v_1,v_2} = \frac{\chi_{v_1}^2/v_1}{\chi_{v_2}^2/v_2} \qquad (2)$$

Since χ_v^2's depend on the normal distribution, the $\mathcal{F}$ distribution also depends on normal distribution. The limiting distribution of $\mathcal{F}_{v_1,v_2}$ as $v_1 \to \infty$ is $\chi_{v_1}^2/v_1$ because as $v_2 \to \infty$ $\chi_{v_2}^2/v_2 \sim 1$.

Student's t-Distribution

The t distribution is a probability distribution, which is frequently used to evaluate hypothesis regarding the mean of continuous variables. Student's t-distribution is quite similar to normal distribution, but the exact shape of t-distribution depends on sample size, i.e., as the sample size increases then t-distribution approaches normal distribution. Unlike the normal distribution, however, which has a standard deviation of 1, the standard deviation of the t distribution varies with an entity known as the degrees of freedom.

We know that $t_v^2 = \left(\frac{\bar{X}-\mu}{s/\sqrt{n}}\right)^2 = \frac{(\bar{X}-\mu)^2 n}{\sum_{i=1}^{n}(X_i-\bar{X})^2/n-1} = \frac{(\bar{X}-\mu)^2 n}{\sum_{i=1}^{n}(X_i-\bar{X})^2/n-1} \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma^2}}$

$$= \frac{(\bar{X}-\mu)^2 n}{\sum_{i=1}^{n}(X_i-\bar{X})^2/n-1} \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma^2}} = \frac{\frac{(\bar{X}-\mu)^2}{\sigma^2/n}}{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2/(n-1)}} = \frac{\frac{(\bar{X}-\mu)^2}{\sigma^2/n}}{\frac{(n-1)\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{(n-1)}}{\sigma^2/(n-1)}}$$

$$= \frac{\frac{(\bar{X}-\mu)^2}{\sigma^2/n}}{\frac{(n-1)s^2}{\sigma^2/(n-1)}} = \frac{\frac{(\bar{X}-\mu)^2}{\sigma^2/n}}{\chi^2/(n-1)} = \frac{z^2}{\chi^2/(n-1)}$$

$$t_v^2 = \frac{z^2}{\chi^2/(n-1)}$$

where $n-1$ is the degree of freedom

A squared t random variable equals the ratio of squared standard normal variable divided by chi-squared with $n-1$ degree of freedom. Since $t_v^2 = \frac{z^2}{\chi^2/n-1}$, the $t_v = \frac{z}{\sqrt{\chi^2/n-1}}$. As $n \to \infty$, $t_{n-1} \sim \mathcal{N}(0,1)$ because $\chi^2/n-1 \to 1$. In this case, squared root of both side results, the Student's t based on normal distribution become: $t_t = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$. Moreover, $z^2 = \chi_1^2$, then the $t^2 = \frac{z^2/1}{\chi_n^2/n-1} = \frac{\chi^2/1}{\chi_n^2/n-1} = \mathcal{F}_{1,n-1}$. For t-distribution, $t_v = \frac{z}{\sqrt{\chi^2/n-1}}$, where $z = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$ and $\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$, then the test statistic:

$$t = \frac{z}{\sqrt{\chi^2/n-1}} \sim t_{n-1} \qquad (3)$$

Summary of relationships

Let $\mathcal{Z} \sim \mathcal{N}(0,1)$

Distribution	Definition	Parent	Limiting
$\mathcal{X}_\mathcal{V}^2$	$\sum_{i=1}^{\mathcal{V}} \mathcal{Z}_i^2$, Independent $\mathcal{Z}'s$	Normal	As $\mathcal{V} \to \infty, \mathcal{X}_\mathcal{V}^2 \to normal$
$\mathcal{F}_{\mathcal{V}_1,\mathcal{V}_2}$	$(\mathcal{X}_{\mathcal{V}_1}^2/\mathcal{V}_1)/(\mathcal{X}_{\mathcal{V}_2}^2/\mathcal{V}_2)$, Independent $\mathcal{X}^{2'}s$	Chi-squared	As $\mathcal{V}_2 \to \infty$, $\mathcal{F}_{\mathcal{V}_1,\mathcal{V}_2} \to \mathcal{X}_{\mathcal{V}_1}^2/\mathcal{V}_1$
t	$\mathcal{Z}/\sqrt{\mathcal{X}^2/\mathcal{V}}$	Normal	As $\mathcal{V} \to \infty, t \to normal$

Note: $\mathcal{F}_{1,\mathcal{V}} = t_\mathcal{V}^2$, also $\mathcal{F}_{1,\infty} = t_\infty^2 = \mathcal{Z}^2 = \mathcal{X}_1^2$

TEST STATISTIC APPROACH OF HYPOTHESIS TESTING

Hypotheses are predictions about the relationship among two or more variables or groups based on a theory or previous research [3].Hence, hypothesis testing is the art of testing if variation between two sample distributions can be explained through random chance or not. The three approaches of hypothesis testing are test statistic, p-value and confidence interval.

In this paper, we focus only test statistic approach of hypotheses testing. A test statistic is a function of a random sample, and is therefore a random variable. When we compute the statistic for a given sample, we obtain an outcome of the test statistic. In order to perform a statistical test we should know the distribution of the test statistic under the null hypothesis. This distribution depends largely on the assumptions made in the model. If the specification of the model includes the assumption of normality, then the appropriate statistical distribution is the normal distribution or any of the distributions associated with it, such as the Chi-square, Student's t, or Snedecor's F.

z-tests (large sample case)

A Z-test is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution [17]. In this paper, the z-test is used for testing significance of single population mean and difference of two population means if our sample is taken from a normal distribution with known variance or if our sample size is large enough to invoke the Central Limit Theorem (usually $n \geq 30$ is a good rule of thumb).

One sample z-test for Mean

A one-sample z-test helps to determine whether the population mean, μ is equal to a hypothesized value when the sample is taken from a normal distribution with known variance σ^2. Recall theorem 1. If $X_1, X_2, \ldots, X_n$ are a random variables, then, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$. In this case, the mean of the sample from the population is not constant rather it varies depend on which sample drawn from the population is chosen. Therefore, the sample mean is a random variable with its own distribution is called sampling distribution.

Lemma 1. If $\bar{X}$ is the random variable of the sample mean of all the random samples of size n from the population with expected value $E(X)$ and variance $var(X)$, then the expected value and variance of $\bar{X}$ are $E(\bar{X}) = E(X)$ and $var(\bar{X}) = \frac{var(X)}{n}$, respectively.

Proof: consider all of the possible samples of size n from a population with expected value $E(X)$ and variance $var(X)$. If a sample $X_1, X_2, \ldots, X_n$ are chosen, each X_i comes from the same population so each X_i has the same expected values, $E(X)$ and variance $var(X)$.

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} E(X_i)\right) = \frac{1}{n}(n.E(X)) = E(X)$$

$$var(\bar{X}) = var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\left(\sum_{i=1}^{n} var(X_i)\right) = \frac{1}{n^2}(n.var(X) = \frac{var(X)}{n}$$

The first step is to state the null hypothesis which is a statement about a parameter of the population (s) labeled as H_0 and alternative hypothesis which is a statement about a parameter of the population (s) that is opposite to the null hypothesis labeled H_1. The null and alternatives hypotheses (two-tail test) are given by; $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$.

Theorem 2: If $X_1, X_2, \ldots, X_n$ are normally distributed random variable with μ and variance σ^2, then the standard random variable Z(one-sample Z-test) has a standard normal distribution given by:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad \sim N(0,1) \qquad (4)$$

where $\bar{X}$ is the sample mean, μ_0 is the hypothesized population mean under the null hypothesis, σ is the population standard deviation and n is the sample size, recall that $\sigma/\sqrt{n} = \sigma_{\bar{x}}$ (standard error). However, according to Joginder K. [6] if the population standard deviation is unknown, but $n \geq 30$, given the Central Limit Theorem, the test statistic is given by:

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \qquad (5)$$

where$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

Two-Sample z-test(when variances are unequal)
A two-sample test can be applied to investigate the significant of the difference between the mean of two populations when the population variances are known and unequal, but both distributions are normal and independently drawn. Suppose we have independent random sample of size n and m from two normal populations having means μ_X and μ_Y and known variances σ_X^2 and σ_Y^2, respectively, and that we want to test the null hypothesis $\mu_X - \mu_Y = \mu_d$, where μ_d is a given constant, against one of the alternatives $\mu_X - \mu_Y \neq \mu_d$, now we can derive a test statistic as follow.
Lemma 2. If$X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ are independent normally distributed random variables with $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$, then $\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m})$.

Proof: by Lemma 1. $\bar{X} \sim N(\mu_X, \frac{\sigma_X^2}{n})$ and $\bar{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{m})$. Then $\bar{X} - \bar{Y} = \bar{X} + (-)\bar{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + (-1)^2\frac{\sigma_Y^2}{m})$. Then the appropriate test statistic for this test is given by:
$z = \frac{(\bar{X}-\bar{Y})-\mu_d}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}} \sim N(0,1)$,where $\sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ and $\sigma_Y^2 = \frac{1}{m}\sum_{j=1}^{m}(Y_j - \bar{Y})^2$
If we do not know the variance of both distributions, then we may use:
$$z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} \qquad (6)$$
Provided n and m are large enough to invoke the Central Limit Theorem.
where n is the sample size of the one group (X) and $m$is the sample size of the other group (Y).
$\bar{X}$is the sample mean of one group and calculated as $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, while the sample mean of the other group is denoted by $\bar{X}$ and calculated as $\bar{Y} = \frac{\sum_{i=1}^{M} Y_i}{m}$.

Two sample Z-test (when variances are equal) for Mean
It can be applied to investigate the significance of the difference between the mean of two populations when the population variances are known and equal, but both distributions are normal. The null and alternatives hypotheses (two-tail test) are given by $H_0: \mu_X - \mu_Y = \mu_0$and $H_1: \mu_X - \mu_Y \neq \mu_0$, respectively.Then test statistic is:
$z = \frac{(\bar{X}-\bar{Y})-\mu_0}{\sigma^2\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1)$, where standard error of difference is $\sigma^2\sqrt{\frac{1}{n} + \frac{1}{m}}$
If we don't know the variance of both distributions, then we may use:
$$z = \frac{(\bar{X} - \bar{Y}) - (\mu_0)}{s^2\sqrt{\frac{1}{n} + \frac{1}{m}}} \qquad (7)$$
provided$n$ and m are large enough to invoke the Central Limit Theorem.

Paired sample z test

According to Banda Gerald [2], the words like dependent, repeated, before and after, matched pairs, paired and so on are hints for dependent samples. Paired samples also called dependent samples used for testing the population mean of paired differences. A paired z-test helps to determine whether the mean difference between randomly drawn samples of paired observations is significant. Statistically, it is equivalent to a one-sample z-test on the differences between the paired observations. The null and alternative hypotheses (two-tail test) are given by $H_0: \mu_d = d_0$ and $H_1: \mu_d \neq d_0$, respectively.

Theorem3.If $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are normally distributed random variables, but X_i and Y_i are dependent, then the standard random variable z is:

$$Z = \frac{\bar{d} - d_0}{\sigma_d/\sqrt{n}} \sim N(0,1) \qquad (8)$$

where $d_i = X_i - Y_i$ is a pairwise difference (is the sample mean of the sum of differences between the first results and the second results of participants in the study) also given by $\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$, d_0 is the mean value of pair wise difference under the null hypothesis, and σ_d is the population standard deviation of the pairwise difference. The basic reason is that X_i and Y_i are normal random variable, so $X_i - Y_i = d_i$ is also normal random variable, thus $d_1, d_2, \ldots, d_n$ are thought to be a sequence of random variable.

Student's tt-test

Unfortunately, z-tests require either the population is normally distributed with a known variance, or the sample size is large. However, when the sample size is small ($n < 30$), the $z$value for the area under the normal curve is not accurate. Thus, instead of a z value, the t values are more accurate for small sample sizes since the $t$value depend upon the number of cases in the sample, i.e., as the sample size change; the t value will change [10].However, the t-test needs an assumption that the sample value should be drawn randomly and distributed normally (since the data is continuous).Therefore, the t-test is a useful technique for comparing the mean value of a group against some hypothesized mean (one-sample) or of two separate sets of numbers against each other (two-sample) in case of small samples ($n < 30$) taken from a normal distribution of unknown variance.

A t-test defined as a statistical test used for testing of hypothesis in which the test statistic follows a Student's t-distribution under the assumption that the null hypothesis is true [5].

One-Sample t-test

Banda Gerald [2] suggested that a one-sample t-test used to compares the mean of a sample to a predefined value. Thus, we use one sample t-test to investigate the significance of the difference between an assumed population mean and a sample mean. The population characteristics are known from theory or are calculated from the population. Thus, a researcher can use one-sample t-test to compare the mean of a sample with a hypothesized population mean to see if the sample is significantly different. The null and alternative hypotheses (two-tail test) are given; $H_0: \mu = \mu_0$ and $H_1: \mu \neq \mu_0$, respectively.Then the t-test statistic is:

$$t = \frac{\bar{\mathcal{X}} - \mu_0}{s/\sqrt{n}} \quad \sim \quad t_{(n-1)} \tag{9}$$

where $\bar{\mathcal{X}}$ is the sample mean, μ is the population mean or hypothesized value, n is the sample size and s is the sample standard deviation which is calculated as $s = \frac{\sum(X_i - \bar{X})^2}{n-1}$.

Independent Two sample t-test (variances unknown but equal)

A two-sample t-test is used to compare two sample means when the independent variable is nominal level data while the dependent variable is interval/ratio level data[13]. Thus, independent samples t-test is used to compare two groups whose means are not dependent on one another. Two samples are independent if the sample values selected from one population are not related with the sample values selected from the other population. Thus, an independent sample t-test helps to investigate the significance of the difference between the means of two populations when the population variance is unknown but equal. Let assume a group of independent normal samples given by $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ distributed as $X_i \sim N(\mu_X, \sigma^2)$ and $Y_i \sim N(\mu_Y, \sigma^2)$, respectively. Then, the null and alternative hypotheses (two-tail test) are given by $H_0: \mu_X - \mu_Y = \mu_0$ and $H_1: \mu_X - \mu_Y \neq \mu_0$, respectively. If sample sizes $n \ and \ m$ aren't large enough to invoke the Central Limit Theorem, then the test statistic is:

$$t = \frac{(\bar{X} - \bar{Y}) - \mu_0}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \quad \sim t_{(n+m-2)} \tag{10}$$

where n refres to sample size from the first group (X), m is sample size from the second group (Y), $\bar{X}$ is the sample mean of the first group (X), $\bar{Y}$ is the sample mean of the second group (Y), s_p is the pooled standard deviation for both groups which is calculated as follows:

$$s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n + m - 2}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^m (Y_i - \bar{Y})^2}{n + m - 2}}$$

$\bar{X} = s_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ and $s_Y^2 = \frac{1}{m-1}\sum_{i=1}^m (Y_i - \bar{Y})^2$

Let us derived t-distribution in two sample t-test (variances unknown but equal).

Lemma 3. If $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are normally and independent distributed random variables with $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_i \sim N(\mu_Y, \sigma_Y^2)$, then $\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y, \sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)\right)$.

Proof: by **Lemma 1**. $\bar{X} \sim N(\mu_X, \frac{\sigma_X^2}{n})$ and $\bar{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{m})$. Then $\bar{X} - \bar{Y} = \bar{X} + (-)\bar{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + (-1)^2\frac{\sigma_Y^2}{m})$.

Lemma 4. If $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are normally and independent distributed random variables with $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_i \sim N(\mu_Y, \sigma_Y^2)$, then $\frac{(\bar{X} - \bar{Y}) - \mu_0}{\sigma^2\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1)$.

Theorem 4.If $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are normally and independent distributed random variables with $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_i \sim N(\mu_Y, \sigma_Y^2)$, then the test statistic is given as:

$$t = \frac{(\bar{X} - \bar{Y}) - \mu_0}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \tag{11}$$

where $s_p = \sqrt{\frac{(n-1)s_X^2+(m-1)s_Y^2}{n+m-2}} = \sqrt{\frac{\sum_{i=1}^n(X_i-\bar{X})^2+\sum_{i=1}^m(Y_i-\bar{Y})^2}{n+m-2}}$ has a t-distribution with $n+m-2$ degree of freedom.

Proof. $t = \frac{(\bar{X}-\bar{Y})-\mu_0}{\sqrt{\frac{(n-1)s_X^2+(m-1)s_Y^2}{n+m-2}}\sqrt{\frac{1}{n}+\frac{1}{m}}} \frac{\sigma}{\sigma} = \frac{(\bar{X}-\bar{Y})-\mu_0}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} / \sqrt{\left(\frac{(n-1)s_X^2}{\sigma^2}+\frac{(m-1)s_Y^2}{\sigma^2}\right)\frac{1}{n+m-2}}$

Let $z = \frac{(\bar{X}-\bar{Y})-\mu_0}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}$, by lemma 4. $z \sim N(0,1)$ and if $U = \left(\frac{(n-1)s_X^2}{\sigma^2}+\frac{(m-1)s_Y^2}{\sigma^2}\right)$. Then $U \sim \chi^2_{n+m-2}$ since

$\frac{(n-1)s_X^2}{\sigma^2} \sim \chi^2_{n-1}$ and $\frac{(m-1)s_Y^2}{\sigma^2} \sim \chi^2_{m-1}$. Therefore, $t = \frac{Z}{\sqrt{\frac{U}{n+m-2}}} \sim t_{(n+m-2)}$ which has a t-distribution

with $n+m-2$ degree of freedom.

Independent Two sample t-test (variances unknown but unequal)
However, if the variances are unknown and unequal, then we use a t-test for two population mean given the null hypothesis of the independent sample t-test (two-tail test) $H_0: \mu_X - \mu_Y = \mu_0$, the test statistic is

$$t = \frac{(\bar{X} - \bar{Y}) - \mu_0}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} \quad \sim t_{(n+m-2)} \tag{12}$$

where $s_X^2 = \frac{1}{n-1}\sum_{i=1}^n(X_i - \bar{X})^2$ and $s_Y^2 = \frac{1}{m-1}\sum_{i=1}^m(Y_i - \bar{Y})^2$

Paired (dependent) sample t test
Two samples are dependent (or consist of matched pairs) if the members of one sample can be used to determine the members of the other sample. A paired t-test is used to compare two population means when we have two samples in which observations in one sample can be paired with observations in the other sample under Gaussian distribution [7], [16].When two variables are paired, the difference of these two variables, $d_i = X_i - Y_i$, is treated as if it were a single sample. This test is appropriate for pre-post treatment responses. The null hypothesis is that the true mean difference of the two variables is d_0 as given by; $H_0: \mu_d = d_0$. Then the test statistic:

$$t_{\bar{d}} = \frac{\bar{d} - d_0}{s_{\bar{d}}} \quad \sim t_{(n-1)} \tag{13}$$

where $\bar{d} = \frac{\sum d_i}{n}$, $s_d^2 = \frac{\sum(d_i-\bar{d})^2}{n-1}$, $s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$, μ_0 is the mean of the pairwise difference under the null hypothesis and s_d is sample standard deviation of the pairwise differences.
Proof: Let X_i and Y_i are normal random variable. Therefore, $d_i = X_i - Y_i$ is also a normal random variable.Then $d_1, d_2, \ldots, d_n$ can be thought as the sequence of random variable, $t \sim t_{n-1}$.

Effect Size

Both independent and dependent sample t-test will give the researcher an indication of whether the difference between the two groups is statistically significant, but not indicates the size (magnitude) of the effect. However, effect size explains the degree to which the two variables are associated with one another. Thus, effect size is the statistics that indicates the relative magnitude of the differences between means or the amount of total variance in the dependent variable that is predictable from the knowledge of the levels of independent variable.

3.2.5.1.   Effect Size of an Independent Sample t-test

An effect size is a standardized measure of the magnitude of observed effect. The Pearson's correlation coefficient r is used as common measures of effect size. Correlation coefficient is used to measure the strength of relationship between two variables. From the researcher's point of view, correlation coefficient of 0means there is no effect, and a value of 1 means that there is a perfect effect. The following is the formula used to calculate the effect size;

$$r = \sqrt{\frac{t^2}{t^2 + n + m - 2}} \qquad (14)$$

where t is the calculated value of the independent sample t-test, $n \ and \ m$ is the sample size, if $r = 0.1$refers to small effect, $r = 0.3$refers to medium effect and $r = 0.5$ refers to large effect.

3.2.5.2.   Effect Size of Dependent Sample t-test

The most common used method for calculating effect size independent sample t-test is the Eta Squared [11]. The formula for simplified Eta Squared is given by:

$$Eta \ Squared = \frac{1}{1 + \frac{n-1}{t^2}} \qquad (15)$$

where$t$ is the calculated value of the dependent sample t-test, n is the sample size. The calculated value of the Eta Squared should be between 0 and 1. $Eta \ Squared = 0.01, 0.06,$ and $0.14$are small, medium, and large effect, respectively.

Chi-Square Test Statistic

Karl Pearson in 1900 proposed the famous chi-squared test for comparing an observed frequency distribution to a theoretically assumed distribution in large-sample approximation with the chi-squared distribution. According to Pearson, chi-square test is a nonparametric test used for testing the hypothesis of no association between two or more groups, population or criteria (i.e. to check independence between two variables) and how likely the observed distribution of data fits with the distribution that is expected (i.e., to test the goodness-of-fit) in categorical data. In this case, we apply chi-square test for testing population variance, independence and goodness of fit.

Chi-Square Test of Population Variance (One-sample test for variance)

The chi-square test can be used to judge if a random sample has been drawn from a normal population (interval and ratio data) with mean (μ) and with specified variance (σ^2) [15]. Thus, it used to compare the sample variance s^2 to a hypothesized varianceσ_0^2 (i.e. it is a test for a single sample variance). Assume the data $X_1, X_2, .., X_n$are independent normal samples with$X_i \sim N(\mu, \sigma^2)$

and we know that $\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_{n-1}^2$ if $z \sim \mathcal{N}(0,1)$. The null and alternative hypotheses (two-tail) given by $H_o: \sigma^2 = \sigma_0^2$ and $H_1: \sigma^2 \neq \sigma_0^2$, respectively. Then the test statistic:

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_{n-1}^2$$

However, we don't know μ, so we use $\bar{X} = \frac{\sum X_i}{n}$ as an estimate of μ and $s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$. Thus, sampling distribution of test statistic if H_o is true:

$$\chi_v^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi_{n-1}^2 \tag{16}$$

The decision rule is to compare the obtained test statistic to the chi-squared distribution with $v = n - 1$ degree of freedom. If we construct the critical region of size α in the two-tail case, our critical region of size α will be all values of χ^2 such that either $\chi^2 \leq \chi_{1-\alpha/2}^2, n - 1$ or $\chi^2 \geq \chi_{\alpha/2}^2, n - 1$.

Chi-Square Test of Independence
The chi-square test of independence is a nonparametric statistical test used for deciding whether two categorical (nominal) variables are associated or independent [1], [5]. Let the two variables in the cross classification are X and Y, then the null H_0: no association between X and Y against alternative H_1: some association between X and Y_i. The chi-square statistic used to conduct this test is the same as in the goodness of fit test given by:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_{1-\alpha[(r-1)(c-1)]}^2 \tag{17}$$

where χ^2 is the test statistic that asymptotically approaches a chi-square distribution, O_{ij} is the observed frequency of the ith row and j^{th} column, representing the numbers of respondents that take on each of the various combinations of values for the two variables. $E_{ij} = \frac{(\text{sum of row } I) \times (\text{sum of column } J)}{\text{sample size } n}$ is the expected (theoretical) frequency of the ith row and j^{th} column, where r is the number of rows in the contingency table, c is the number of columns in the contingency table.

The statistical decision is if $\chi^2 > \chi_{1-\alpha[(r-1)(c-1)]}^2$, reject the null hypothesis and conclude that there is a relationship between the variables. If $\chi^2 < \chi_{1-\alpha[(r-1)(c-1)]}^2$, the null hypothesis cannot be rejected and conclude that there is insufficient evidence of a relationship between the variables.

Chi-Square Test for Goodness of Fit
A chi-square test for goodness of fit is used to compare the frequencies (counts) among multiple categories of nominal or ordinal level data for one-sample [13]. Moreover, a chi-square test for goodness of fit used to compares the expected and observed values to determine how well the experimenter's predictions fit the data [14].
The null hypothesis ($H_0: O_{ij} = E_{ij}$) is stated as the population distribution of the variable is the same as the proposed distribution against the alternative ($H_1: O_{ij} \neq E_{ij}$) the distributions are different. The chi square statistic is defined as:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \sim \chi^2_{1-\alpha[K-1]} \qquad (18)$$

Large values of this statistic lead rejection of the null hypothesis, meaning the observed values are not equal to theoretical values (expected).

The F-test statistic

F-tests are used when comparing statistical models that have been fit to a data set to determine models that best fit the population from the data being sampled [4]. The F distribution is a joint distribution of two independent variables, with each having a chi-square distribution [9].The F Distribution needs both parent populations to be normal, have the same variance and independent sample. F-test is applicable for testing equality of variances and testing equality of several means.

Testing for Equality of Variances

The chi-square test used for testing a single population variance, while $\mathcal{F}$-test for testing the equality of variances of two independent populations (comparing two variances). Thus, it helps to make inferences about a population variance or the sampling distribution of the sample variance from a normal population given the null$H_o: \sigma_X^2 = \sigma_Y^2$ against the alternativehypothesis$H_1: \sigma_X^2 \neq \sigma_Y^2$. Suppose two independent random samples drawn from each population (normally and identically distributed population) as given by$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.Moreover, let s_X^2 and s_Y^2 represent the sample variances of two different populations. If both populations are normal and the population variances $\sigma_X^2$and σ_Y^2 are equal, but unknown, then the sampling distribution of $\mathcal{F}$ is called an $\mathcal{F}$-distribution.

The first samples$X_1, X_2, \ldots, X_n$ are normally and independent distributed random variables, $V = \frac{(n-1)s_X^2}{\sigma_X^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma_X^2}$ with $n - 1$ d.f.

The second samples $Y_1, Y_2, \ldots, Y_m$ are normally and independent random variables$U = \frac{(m-1)s_Y^2}{\sigma_Y^2} = \frac{\sum_{j=1}^{m}(Y_j - \bar{Y})^2}{\sigma_Y^2}$ with $m - 1$ d.f.Let assume $V \: and \: U$ are two independent random variables & given the null hypothesis ($H_o: \sigma_X^2 = \sigma_Y^2$) holds true, then$\mathcal{F}$-test for population variance is given by:

$$\mathcal{F} = \frac{U/m-1}{V/n-1} = \frac{s_Y^2/\sigma_Y^2}{s_X^2/\sigma_X^2} = \frac{s_Y^2}{s_X^2} \sim \mathcal{F}_{v_1, v_2} = \mathcal{F}_{m-1, n-1} \qquad (19)$$

where$n$ and m are samples from population X and Y,$v_1 = m - 1$ = numerator degrees of freedom, $v_2 = n - 1$ = denominator degrees of freedom.

Testing equality of several means (ANOVA)

Analysis of variance (ANOVA) is statistical technique used for comparing means of more than two groups(usually at least three) given the assumption of normality, independency and equality of the error variances. In one-way ANOVA, group means are compared by comparing the variability between groups with that of variability within the groups. This is done by computing an F-statistic. The F-value is computed by dividing the mean sum of squares between the groups by the mean sum of squares within the groups [5], [8].

A one-way analysis of variance (ANOVA) is given by:

$$Y_{ij} = \alpha + \tau_i + \varepsilon_{ij}$$

where Y_{ij} is the jth observation for the ith income level, α is a constant term, τ_i is the effect on the ith income level, $\varepsilon_{ij} \sim N(0, \sigma_i^2)$, $j = 1,2, \ldots, n_i$. The null hypothesis given by $H_0: \tau_1 = \tau_2 \ldots = \tau_\lambda = \tau = 0$ against $H_1: at\ least\ one\ \tau_i \neq \tau, i = 1,2, \ldots, \lambda$. Total variation of the data is partitioned as variation within levels (groups) and between levels (groups).Mathematically, $\bar{X} = \sum_{i=1}^{K} \sum_{j=1}^{J} \frac{X_{ij}}{n}$ and portion of total variance, $\sum_{i=1}^{K} \sum_{j=1}^{J} (X_{ij} - \bar{X})^2 = \sum_{j=1}^{J} n_i \cdot (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^{K} \sum_{j=1}^{J} (X_{ij} - \bar{X}_i)^2$, where $\sum_{i=1}^{K} \sum_{j=1}^{J} (X_{ij} - \bar{X})^2 = TSS$ is sum of square variation about the grand mean across all n observations, $\sum_{j=1}^{J} n_i \cdot (\bar{X}_i - \bar{X})^2 = ESS$ is sum of squared deviation for each group mean about the grand mean, $\sum_{i=1}^{K} \sum_{j=1}^{J} (X_{ij} - \bar{X}_i)^2 = RSS$ is sum of squared deviation for all observation within each group from that group mean, summed across all groups. Then the mean sum of square between or $MSSB = \frac{\sum_{j=1}^{J} n_i \cdot (\bar{X}_i - \bar{X})^2}{K-1}$ denotes the sum of squares that represents the variation among the different samples, while within mean sum square or $WMSS = \frac{\sum_{i=1}^{K} \sum_{j=1}^{J} (X_{ij} - \bar{X}_i)^2}{n-K}$ which refers to the sum of squares that represents the variation within samples that is due to chance. Then the test statistic is the ratio of group and error mean squares:

$$\mathcal{F} = \frac{MSSB}{MSSW} \sim \mathcal{F}_{n-k,k-1} \qquad (20)$$

One-Way ANOVA Table

Source of Variation	Sum of square	DF	Mean squares	$\mathcal{F}$
Between (between group)	SSB	k-1	MSSB=$\frac{SSB}{k-1}$	$\dfrac{MSSB}{MSSw}$
Within (error)	SSW	n-k	MSSW=$\frac{MSW}{n-k}$	
Total	TSS	n-1		

Reject H_0 if the calculated $\mathcal{F}$ is greater than the tabulate value ($\mathcal{F}_{n-k,k-1}$), meaning variability between groups is large compared to variation within groups, while fail to reject H_0 if the calculated $\mathcal{F}$ is less than the tabulate value ($\mathcal{F}_{n-k,k-1}$), meaning the variability between groups is negligible compared to variation within groups.

CONCLUSION

In this paper, the derivation and choice of appropriate a test statistic were reviewed since application of right statistical hypothesis tests to the right data is the main role of statistical tools in empirical research of every science. Hence, the researcher apply z-test, which is appropriate to test the existence of population mean difference in the case of large sample size and the t-test is for small sample size. Moreover, $\mathcal{F}$-test is used for test of several mean and variance, while Chi-square test is for testing single variance, goodness of fit and independence.

References

[1] Alena Košťálová. *(2013). Proceedings of the 10th International Conference "Reliability and Statistics in Transportation and Communication" (RelStat'10), 20–23 October 2010, Riga, Latvia, p. 163-171. ISBN 978-9984-818-34-4 Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia.*

[2] Banda Gerald. (2018). A Brief Review of Independent, Dependent and One Sample t-test. *International Journal of Applied Mathematics and Theoretical Physics*, 4(2): 50-54. doi: 10.11648/j.ijamtp.20180402.13.

[3] David, J. Pittenger. (2001). Hypothesis Testing as a Moral Choice. *Ethics & Behavior*, 11(2): 151-162, DOI: 10.1207/S15327019EB1102_3.

[4] DOWNWARD, L. B., LUKENS, W. W. & BRIDGES, F. (2006).A Variation of the F-test for Determining Statistical Relevance of Particular Parameters in EXAFS Fits.*13th International Conference on X-ray Absorption Fine Structure*, 129-131.

[5] J. P. Verma. (2013). Data Analysis in Management with SPSS Software, DOI 10.1007/978-81-322-0786-3_7, *Springer*, India.

[6] Joginder Kaur. (2013). Techniques Used in Hypothesis Testing in Research Methodology A Review. *International Journal of Science and Research (IJSR)* ISSN (Online): 2319-7064 Index Copernicus Value: 6.14 | Impact Factor: 4.438.

[7] Kousar, J. B. and Azeez Ahmed. (2015). The Importance of Statistical Tools in Research Work. *International Journal of Scientific and Innovative Mathematical Research (IJSIMR) Volume 3, Issue 12, PP 50-58 ISSN 2347-307X (Print) & ISSN 2347-3142 (Online)*

[8] Liang, J. (2011). Testing the Mean for Business Data: Should One Use the Z-Test, T-Test, F-Test, The Chi-Square Test, Or The P-Value Method? *Journal of College Teaching and Learning,* 3(7), doi:10.19030/tlc.v3i7.1704.

[9] LING, M. (2009b).Compendium of Distributions: Beta, Binomial, Chi-Square, F, Gamma, Geometric, Poisson, Student's t, and Uniform. The Python Papers Source Codes 1:4.

[10] MCDONALD, J. H. (2008). *Handbook of Biological Statistics.* Baltimore, Sparky House Publishing.

[11] Pallant, J. (2007). SPSS Survival Manual: A Step by Step to Data Analysis Using SPSS for Windows (Version 15), Sydney, Allen and Unwin.

[12] Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables that arisen from random sampling. *Philos. Mag.* 1900, *50*, 157-175.Reprinted in K. Pearson (1956), pp. 339-357.

[13] Philip, E. Crewson.(2014). *Applied Statistics*, 1st Edition, United States Department of Justice, https://www.researchgate.net/publication/297394168.

[14] Sorana, D. B., Lorentz, J., Adriana, F. S., Radu, E. S. and Doru, C. P. (2011). Pearson-Fisher Chi-Square Statistic Revisited. *Journal of Information, 2:* 528-545; doi: 10.3390/info2030528; ISSN 2078-2489, www.mdpi.com/journal/information

[15] Sureiman Onchiri.(2013).Conceptual model on application of chi-square test in education and social sciences. *Academic Journals*, *Educational Research and Reviews,* 8(15):1231-1241; DOI: 10.5897/ERR11.0305; ISSN 1990-3839.http://www.academicjournals.org/ERR.

[16] Tae Kyun Kim. (2015). T test as a parametric statistic. *Korean Society of Anesthesiologists,* ISSN 2005-6419, eISSN 2005-7563.

[17] Vinay Pandit. (2015). A Study on Statistical Z Test to Analyses Behavioral Finance Using Psychological Theories. *Journal of Research in Applied Mathematics, 2(1): 01-07 ISSN(Online), 2394-0743 ISSN (Print):2394-0735.*www.questjournals.org.