Unit 4 Testing of Statistical Hypothesis

1. Statistical Hypothesis
2. Level of significance
3. Differences between parametric and non-parametric test
4. Use of Z - distribution in hypothesis testing of population mean and population proportion in one sample case

Meaning and definition of Hypothesis

- Hypothesis is guess statement on something which needs to be tested
- Hypotheses are one type of propostion - a sufficiently concrete statement that an be tested(Sullivan 2001)
- Hypotheses are testable statement of presumed relationships between two or more variables. For example, “cigarette smoking is the cause of long cancer”
- A Hypothesis is a conjectured(educated guess) statement that implies or states a relationship between two or more variables(Kerlinger 2000)

1. Statistical Hypothesis

- A statistical hypothesis is an assumption about a population parameter . This assumption may or may not be true
- A statistical hypothesis is a hypothesis concerning the parameters or from of the probability distribution for a designated population or populations, or, more generally, of a probabilistic mechanism which is supposed to generate the observations
- A statistical hypothesis is a hypothesis that is testable on the basis of observed data modelled as the realised values taken by a collection of random variables(Stuart and Arnold 1999)
- A statistical hypothesis test is a method of statistical inference

- Hypotheses are linked to more abstract theory
- Dependent and independent variables in Hypothesis be clearly specified
- Precise nature and direction of relationship between variables be specified in the Hypothesis
- Hypothesis should be stated that they can be verified or refuted
- All oncepts and comparisons in Hypothesis must be clearly stated

Types ohypothesis

1. Descriptive and relational: Descriptive hypothesis contains only one variable. For example, Nepal government faces budget defficiency. Relational hypothesis describes the relationships between to or more varables
2. Directional and non-directional: Directional hypothesis indicates the particular direction of the expected relationship between variables. For example, younger teachers are less motivated than older teachers. Non-directional hypothesis is formulated when there are no clues available about the positive or negative relationships between two variables. For example, there is difference between work attitudes and of government and Bank employee

3. Null and Alternative hypothesis

- Null hypothesis: Null hypothesis states that no relationships exist between two variables in the population. For example, there is no relation between gender and income. Income is not the cause of gender descrimination
- Alternative hypothesis: Alternative hypothesis states that there is a relationship between two variables in the population. For xample, there is relationship between gender and income. Income is the cause of gender descrimination

2. Level of significance

- Significance level is the criterian used for rejecting the Null hypothesis
- It is used in testing as follows

1. First, the fifference between the results of the experiment and the Null hypothesis is determined. Then, assuming the Null hypothesis is true, the probaility of a difference that large or larger is computed
2. Finally, this probability is less than or equal to the significance level, then the Null hypothesis is rejected and the outcome is said to be statistically significant

- Traditionally, experimenters have used either 0.05 level(sometimes called 5% level) or the 0.01(1%), although, the choice level is largely subjective
- The lower the significance level, the more the data must diverge from the Null hypothesis to be significant
- Value is often denoted Alpha ad is also called the significance level

3. Differences between parametric and non-parametric test

Parametric test

- Parametric statistical test is one that makes assumptions about the parameters(defining properties of the population distribution(s) from which one's data are drawn
- Parametric as reffering to test such as test and analysis of variance that assume the underlying source population(s) to be nomrally distributed
- Parametric data tends to include ratio or intervals
- The mean or average is the best measure of the midpoint of parametric data
- Parametric data always have the homogeneous variance

Non-parametric test

- Non-parametric test is one that makes no such assumptions. In this strict sense. non-parametric test is essentislly a null category, since vitually all statistical tests assume one thing or another about the properties of the sorce population(s)
- Non-parametric as reffering to tests that do not make on these particular assumptions , e.g. Chi square test
- Non-arametric data is either ordinal or nominal
- Medium is more useful for non-parametric data
- Non- parametric data can have any distribution or variance(heterogeneous data)

4. Use of Z - distribution in hypothesis testing of population mean and population proportion in one sample case

- Z - test was developed by Prof. R.A.Fisher
- It is based on the normal distribution
- It is widely used for testing the significance of several statistics such as mean, median, mode, coefficient of correlation and others
- This test is used even when binominal distribution or t dstribution is applicable on the presumption that such a distribution tends to approximate normal disrtibution as the sample size (n) becomes larger

How to Run a One Sample Z Test

A one sample z test is one of the most basic types of hypothesis test. In order to run a one sample z test, you work through several steps:

Step 1: State the Null Hypothesis. This is one of the common stumbling blocks–in order to make sense of your sample and have the one sample z test give you the right information you must make sure you have written the null hypothesis and alternate hypothesis correctly. For example, you might be asked to test the hypothesis that the mean weight gain of pregnant women was more than 30 pounds. Your null hypothesis would be: H0: μ = 30 and your alternate hypothesis would be H,sub>1: μ > 30.

Step 2: Use the z-formula to find a z-score.

One Sample Z Test

$$Z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$$

All you do is put in the values you are given into the formula. Your question should give you the sample mean ($\bar{x}$), the standard deviation (σ), and the number of items in the sample (n). Your hypothesized mean is $\mu 0$.

Example:

1,500 women followed the Atkin's diet for a month. A random sample of 29 women gained an average of 6.7 pounds. Test the hypothesis that the average weight gain per woman for the month was over 5 pounds. The standard deviation for all women in the group was 7.1.

$Z = 6.7 - 5 / (7.1/\sqrt{29}) = 1.289.$

Step 3: Look up your z score in the z-table. The sample score above gives you an area of 0.8997. This area is your probability up to that point (i.e. the area to the left of your z-score). For this one sample z test, you want the area in the right tail, so subtract from 1:

1 – 0.8997 = 0.1003.

If you have difficulty figuring out if you have a left- or-right tailed test, see:

Left Tailed Test or Right Tailed Test? How to Decide in Easy Steps.