Testing of Hypothesis

Some important definitions

1) Hypothesis:

Any statement or ascertion regarding a population or its parameter is called a hypothesis.

Eg: Let, 'p' be the probability of getting head in a single toss of a coin. Then, $H: p = \frac{1}{2}$ is a hypothesis.

2) Simple Hypothesis:

A hypothesis that specifies a population completely, i.e. the probability distribution & all the parameters are known, is called a simple hypothesis.

Eg: Let, 'p' be the probability of getting head in a single toss of a coin. Then the hypothesis $H: p = \frac{1}{2}$ is a simple hypothesis.

3 Composite Hypothesis:

A hypothesis which fails to specify the population completely, i.e either the form of the probability function or some of the parameters remain unknown, is called a composite hypothesis.

Eg: Let, 'p' be the probability of getting head in a single toss of a coin. Then the hypothesis $H: p > \frac{1}{2}$ composite hypothesis.

4 Null Hypothesis

A statistical hypothesis which is set up (i.e assumed) & whose validity is tested for possible rejection on the basis of the sample observations is called a null hypothesis. It is denoted by H_0 & tested against the possible alternatives. Tests of hypothesis deal with rejection or acceptance of null hypothesis only.

5) Alternative Hypothesis:

A statistical hypothesis which contradicts the null hypothesis is called an alternative hypothesis & is denoted by H_1. The alternative hypothesis is not tested, but its acceptance or rejection depends on the possible rejection or acceptance of the null hypothesis. The choice of an appropriate critical region depends on the type of the alternative hypothesis, viz, whether both sided or ~~one~~ one sided (right or left).

(6) Testing of Hypothesis

By testing of hypothesis, we mean a procedure which specifies a set of 'rules for decision' regarding whether to 'accept' or 'reject' the null hypothesis.

(7) Test of Hypothesis

By test of a hypothesis, we mean a set of 'rules for decision' regarding whether to accept or reject the null hypothesis.

(8) Test Statistic

A function of sample observations (ie. statistic), whose computed value determines the final decision regarding the acceptance or ~~r~~ rejection of H_0, is called a test statistic. The appropriate test statistic has to be chosen very carefully & a knowledge of its sampling distribution under H_0 (ie when the null hypothesis is true) is essential in framing the decision rules. If the values of the test statistic falls in the critical region, the null hypothesis is rejected.

(9) Critical Region

The set of values of the test statistic which leads to the rejection of the null hypothesis, is called the critical region of the test. The probability with which a true null hypothesis is rejected by the test, is often is often referred to as the 'size' of the critical region.

Geometrically, a sample $x_1, x_2, \ldots, x_n$ of size 'n' is looked upon as just a point x, called the ~~p~~ sample point within the region of all possible samples, called the sample space (W). The critical region, is then defined as a subset (w) of those sample points which lead to the rejection of null hypothesis.

Eg: Let, 'p' be the probability of getting head in a single toss of a coin. Suppose, for the purpose of this testing, we toss the coin 10 times & let x be the number of heads obtained in these 10 tosses. Then, x may be regarded as our test statistic. Further, suppose we decide to reject H_0 if the no. of heads is more than 7 or less than 3. The set of all possible values that can be taken up by the test statistic is called the sample space & is denoted by $\mathcal{X}$. Here, $\mathcal{X} = \{0, 1, 2, \ldots, 9, 10\}$. Here, the ~~criti~~ critical region is $W = \{0, 1, 2, 8, 9, 10\}$. The complement of the critical region i.e W^c is called the acceptance region of the test.

(10) Two types of error:

In any testing problem one may commit two types of errors. The first type is the error of rejecting a true null hypothesis termed as type-I error & the 2nd type is the error of accepting false null hypothesis ~~is~~ called the type-II error. The situations may be represented by the following table.

True state	Decision Taken	
	Accept H_0	Reject H_0
H_0 is true	Correct Decision	Type-I error
H_0 is false	Type-II Error	Correct decision

Let, T be the test statistic & $T(\underset{\sim}{X})$ be the value of T for a set of sample observations. Then,

$$P(\text{Type}-\text{I error}) = P\{T(\underset{\sim}{X}) \in W \mid H_0\}.$$

$$P(\text{Type}-\text{II error}) = P\{T(\underset{\sim}{X}) \in W^c \mid H_1\}.$$

(11) Level of Significance :

The maximum probability, with which a true null hypothesis is rejected, is known as the level of significance of the test.

It would be an ideal situation, if a test could minimise the probabilities of both the types of error, simultaneously. However given a fixed sample size, the reduction in the probability of one kind of error results in an increment of the other kind of error. For this reason, we give an upper bound to the probability of type-I error, so that it is not allowed exceed a predetermined low level of probability. & subject to this constraint, we try to choose a test ~~statistic~~ which minimises the probability of type-II error, as much as possible. The predetermined low level of probability of type-I error is known as the level of significance of the test & is denoted by α. One usually takes $\alpha = 0.05, 0.01$ & so on.

The actual probability of type I error is known as the size of the test. [Size $\leq$ level of significance] or the size of the critical region.

12 Power:

The power of a test is the probability of rejecting a false null hypothesis. Thus, power is the probability of taking a correct decision. We thus have,

~~Power~~ Power $= P\{T(\underset{\sim}{X}) \in W \mid H_1\} = 1 - P\{T(\underset{\sim}{X}) \in W^c \mid H_1\}$

$= 1 - P\{\text{Type-II error}\}$.

13 One Sided test:

Let us consider the following testing problems:

a) $H_0: \mu = \mu_0$ against $H_1: \mu > \mu_0$.

b) $H_0: \mu = \mu_0$ against $H_1: \mu < \mu_0$,

μ being the parameter under consideration. Here, the desired level of significance is α.

In the above problems, the alternative hypothesis are one sided. In order to test H_0, we shall be looking for only large values [in problem (a)] or only small values [in problem (b)] of the suitable test statistic - T in order to reject H_0. Whether the value of T is significant or not is determined by α.

Here, one tries to see whether T assumes a significant value on one side or one tail of its distribution. Such tests are known as one sided or one tailed test.

14. Two sided test:

Let us consider the following testing problem:

$H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$,

'μ' being the parameter under consideration. Here the desired level of significance is α;

Here, the alternative hypothesis is two sided. In order to test H_0, we shall be looking for ~~at~~ large as well as small values of the suitable test statistic 'T' inorder to reject H_0. Wether the value of T is significant or not is determined by α.

Here, one tries to see whether T assumes a significant value on either side or tail of its distribution. Such test is known as ~~one~~ two sided or two tailed tests.

Definition

The ***P*-value** is the smallest level of significance that would lead to rejection of the null hypothesis H_0.