TESTING OF HYPOTHESIS

Any population can describe with the help of p.d.f. There may be some unknowns in population density. The unknowns about the population are called parameters. In theory of statistics it is necessary to know the value of parameters. e.g. in case of $P(\lambda)$

$P(X=0) = e^{-\lambda}$ if we dont know the value of paramete λ then prob. given above is not understandable. Hence the value of the unknown parameter λ is very important.

We have studied some methods for estimating the parameters. Estimated value may be acceptable or not, hence we require some technique to know about the acceptance of the estimated value. such technique is called <u>Testing of hypothesis</u>.

There is another way of looking at testing of hypothesis. We assign some value to the parameter or we make some assumption about or statement about parameter then we use some technique (Hypothesis) called as testing of Hypothesis in which we check or test our assumption is true or false.

- Parameter : It is unknown for the population or unknown in p.d.f. e.g. in $\exp(\lambda)$ $dist^n$, λ is parameter.

- Hypothesis : It is an assumption or statement made regarding the parameter of the density function of population. denoted by H e.g. H: pop^n mean is 50

  or H : $\mu = 50$.

There are two types of Hypothesis. ⇒

1) Null Hypothesis : (H_0)

Any statement made about the popⁿ parameter by adapting null and neutral attitude, by using unbiased approach is called null hypothesis. It is the statement which is belived to be truth. The best way to design null hypothesis is, set it as, 'Hypothesis of no difference'.

2) Alternative hypothesis (H_1) :

The hypothesis considered to be truth when null hypothesis i.e. H_0 is rejected is called alternative hypothesis. i.e. H_1 is accepted when H_0 is rejected.

Both H_0 and H_1 can be of one of the following forms i.e. ways of Hypothesis are.

1) simple Hypothesis :

It is the hypothesis under which parameter takes single value and under which population is specified completely

e.g. $H : \lambda = 5$

2) Composite Hypothesis :

It is the hypothesis under which parameter takes more than one value or parameter is not specified

e.g. $H : \lambda = 5, 7$

or $H : \lambda > 4$

e.g. $X \sim N(\mu, \sigma^2)$

$H : \mu = 10$ is composite hypothesis (σ^2 is unknown)

$X \sim N(\mu, 25)$

$H : \mu = 10$ is simple Hypothesis (σ^2 is known)

In testing procedures we draw a sample from the population about which decision is to be taken. Then we construct a statistic which can be use for taking the decision about H_0

e.g. if $x_1, x_2, \cdots x_n$ is random sample then statistic $T(x_1, x_2, \cdots x_n) = \sum_{i=1}^{n} x_i$ $i = 1, 2, \cdots n$ is statistic test can be reject H_0 if $\sum x_i < C$

Here $\sum x_i$ is called test statistic ∴ In general test statistic is ~~further~~ function of sample values used to take the decision about null hypothesis.

Error and its types:

In testing of hypothesis decision about the hypothesis made for the population is taken on the basis of few ~~attention~~ attempts drawn from the population i.e. sample. As the decision is taken from sample, decision can go wrong. whenever wrong decision is taken we say that there is error

Following are the two types of error.

H_0:

	True	False
Reject	✗ Type I error	✓
Accept	✓	✗ Type II error

It's seen that, two decisions are errors, these two errors are called as type I error and type II error.

Type I error :

Error committed in rejection null hypothesis when null hypothesis is true. then it is called type I error. Probability of rejecting H_0 when H_0 is true is called as size of type I error.

size of type I error $= P(\text{Reject } H_0 / H_0 \text{ true})$

Type II error :

Error committed in accepted null hypothesis when null hypothesis is false, then it is called as type II error. Probability of accepting H_0 when H_0 is false is called as size of type II error.

size of type II error $= P(\text{accept } H_0 / H_0 \text{ false})$
$= P(\text{accept } H_0 / H_1 \text{ true})$

Level of significance (α) :

Maximum probability of type I error is called as level of significance. Usually it is taken as 1%, 5% etc. It is to be fixed before experimentation.

$\therefore$ level of significance denoted by α given as

$$\alpha = \max(P(\text{type I error}))$$
$$= \max(P(\text{Reject } H_0 / H_0 \text{ true}))$$

and

If H_0 is simple then $\alpha = P(\text{type I error})$

Power of the test

It is the probability of correct decision, probability of (Reject H_0 / H_0 false) is called as power of the test.

$\therefore$ Power of the test $= P(\text{reject } H_0 / H_0 \text{ false})$

$= P(\text{reject } H_0 / H_1 \text{ true})$

$= 1 - P(\text{accept } H_0 / H_1 \text{ true})$

$= 1 - P(\text{type II error})$

$= 1 - \beta$

Hence powe of the test is $1-\beta$.

$\Rightarrow$ If we minimize β, the power will increase

Power function : if H_1 is simple we get numerical value for the above probability, but if H_1 is composite we can not calculate above probability. In that case we will get probability interms of parameter such probability is called as power function

$\therefore$ Power function $= P(\text{reject } H_0 / H_1 \text{ is composite})$

$= K(\theta)$

Power curve : If we plot values of power function $K(\theta)$ against θ, then we get power curve.

[* $K(\theta)$ = power function $= P(\text{reject } H_0)$]

For poisson (λ) $dist^n$, $\lambda=3$ is rejected i.e. $H_0: \lambda=3$ is rejected in favour of $H_1: \lambda=4$ if single observation drawn from the pop^n is below 3, find sizes of both error and power of the test. if $H_1: \lambda>4$ find $K(\lambda)$ and construct power curve.

→ $H_0: \lambda=3$

$H_1: \lambda=4$

let X be single observation such that

$$x \sim P(\lambda) \therefore f(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & x=0,1,2\ldots \\ 0 & o.w \end{cases}$$

size of type I error = P(reject H_0 / H_0 is true)

$= P(x<3 / \lambda=3)$

$= \frac{17}{2} e^{-3}$

size of type II error = P(accept H_0/ H_1 is true)

$= P(X>3 / \lambda=4)$

$= 1 - P(X<3 / \lambda=4)$

$= 1 - e^{-4}(13)$

Power function = P(reject H_0)

$K(\lambda) = P(X<3)$

$= e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}\right)$ //

A urn contains 7 marble out of which some are black if θ denotes no. of black marble in order to test $H_0 : \theta = 2$ against $H_0 : \theta = 4$, 3 marbles are drawn at random if number of black marbl in sample is almost 2 then Ho is accepted. Find sizes of both errors

→ we have $H_0 : \theta = \cancel{3}\ 2$

$H_1 : \theta = 4$

size of type I error = P(type I error)

= P(Reject Ho / Ho true)

$= P(X \geq 3 / \theta = 2)$ (∵ X is no of black marble)

$= 1 - P(X \leq 2 / \theta = 2)$

$= 1 - P(x = 0, 1, 2 / \theta = 2)$

$$= 1 - \left\{ \frac{\binom{2}{0} \times \binom{5}{3}}{\binom{7}{3}} + \frac{\binom{2}{1}\binom{5}{2}}{\binom{7}{3}} + \frac{\binom{2}{2}\binom{5}{1}}{\binom{7}{3}} \right\}$$

$= 1 - 0.999985714$

$= 0.000014285714$

size of type II error = P(type II error)

= P(accept Ho / Ho is false)

= P(accept Ho / H_1 is true)

$= P(\text{accept Ho} / \theta = 4)$

$= P(X \leq 2 / \theta = 4)$

$$= \left\{ \frac{\binom{4}{0}\binom{3}{3}}{\binom{7}{3}} + \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}} + \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} \right\}$$

$= 0.887142$

In testing of hypothesis we draw a sample and construct a statistic on the basis of which we take desesion about null hypothesis. The statistic use for taking decision about H_0 is called test statistic.

The set of all possible values of test statistic is called as sample space. The subset of the sample space which leads us to reject H_0 when observed value of the test statistic falls in it, is called critical ~~range~~ region.

size of C.R. : (c)

It is defined as prob. that sample points falls in C.R. when null hypothesis true, Hence

size of C.R $= P(\text{sample point} \in C) / H_0$ true

$= P((X_1, X_2, \ldots X_n) \in C) / H_0$ true.

$= P(\text{Reject } H_0 / H_0 \text{ true})$

$=$ level of significance

$\therefore c = \alpha$