Module - IV

Hypothesis

1

Meaning : (Saravanavel)

1 According to George A. Lundberg

" A hypothesis is a tentative generalisation, the validity of which remains to be ~~tested~~ tested ".

a In most elementary stage the hypothesis may be any

→ hunch

→ guess,

→ imaginative idea

which becomes the basis for action or investigation.

2 According to Webster " A hypothesis is a proportion, condition or principle which is assumed, ~~for~~ perhaps without belief, and inorder to ~~draw~~ draw its logical consequences (conclusion.

(2)

According to M.H. Gopal "A hypothesis is a tentative solution, that is a possible solution to the problem".

So generally speaking hypothesis is a
→ a provisional formulation or
→ possible solution or
→ tentative explanation or
→ suggested answers to the problem facing the scientist or researcher.

Classification of Hypotheses:

1. Simple versus Complex Hypothesis

i) Simple - States the relationship between two variables.

ii) Complex: States the relationship between two or more independent variables and two or more dependent variables.

(3)

2. Directional versus Non-directional Hypothesis:

i) Directional form – States the direction of the relationship between the variables

ii) Non-directional form – States a relationship but not the direction.

3. On the basis of Types: Null Alter

Hypothesis

Null Hypothesis

- Research Hypothesis
  - Descriptive
  - Relational
    - Correlational form
    - Difference form
  - Causal
- Research Question from Hypothesis RQ form

from above types we know

i) Each of these types can be in simple ~~and~~ or complex form

ii) By its very nature, null hypothesis will be non-directional

iii) Research hypothesis will be directional

iv) Though RQ form are usually non-directional, they can also be directional.

Testing (of) a Hypothesis :

Example :

empirically testing a hypothesis :

→ There is two variables.

i) Teacher &

ii) Students

Hypothesis assumption :-

→ A Teacher is interested in investigating the hypothesis that (praise) or encouragement (5)

→ result in heightened motivation on the part of the students.

→ If this hypothesis is correct it should be logical to assume that teachers encouraging comments on test papers (praise) would be followed by

⇒ improvement in student performance

→ which implies the assumption that heightened motivation is indicated by improved test performance.

& What is the important point is

→ There is a correlation between heightened motivation of teachers and improvement in students performance, i.e. improved test performance.

Step-1 : (Hypothesis)

The deduced implication is stated as follows:

Teachers comments on students paper result in an improvement in pupil's performance on tests

It is the relationship between the two variables - teachers' comments and pupil performance that will be investigated.

Step-2 : Null Hypothesis

For statistical testing the above research hypothesis must be ~~transformed~~ transformed into a null hypothesis -

The Null hypothesis is stated.

(P.T.O)

7

Teachers' comments on students papers will not result in a significant improvement in pupils performance on tests.

$$\boxed{H_0 : \mu_A = \mu_B}$$

Where

μ_A = the mean gain in performance for the population of students receiving teachers comments on papers.

μ_B = the mean gain in performance for the population of students not receiving teachers' comments on papers

→ So the Null hypothesis is rejected.

($H_0 : \mu_1 = \mu_2$) rejected

8

Step - 3 :

Null hypothesis : (H_0)

i) Null hypothesis means

i) There is no relationship between variables

ii) There is no effect between variables

iii) There is no different between variables etc

That is relationship between variables is Negative.

So null hypothesis is

$$\boxed{H_0 = M_1 = M_2} \text{ or,}$$

$$\boxed{H_0 = M_A = M_B}$$

~~H_0 = means negative~~

~~H_1 = Teacher~~

Examples: (9)

i) If there is no difference of Variable in Cost between to different size of farm holders (m, s, m & Large).

ii) Propensity to save for labourers is not lesser than that of farmers.

Saving: Labourers is not < than to farmers. - no difference -

iii) There is no significance difference between government employees income and private employees income

So $\boxed{H_0 = \mu_1 = \mu_2}$ is accepted

$\boxed{\text{If } \mu_1 \neq \mu_2}$ that is Labours saving < farmers it is Rejected.

(or)

Scanned by CamScanner

iv) There is a positive correlation between Teacher heightened motivation (Teacher comments) and improvement of Students performance on Test.

Null hypothesis:

Teachers' comments and (heightened motivation) on students paper will not result in significant improvement in ~~pup~~ pupils performance on ~~the~~ test.

That is:
~~So~~ No significance between the variable of Teachers motivation and improvement of student performance.

So $\boxed{H_0 = M_1 = M_2}$

2

15 Type-I error & Type II error (β)

15 11

Alternative Hypothesis: (H_1)

→ It is a opposite of null Hypothesis.

→ Alternative hypothesis means

→ There is a relationship between variables

→ There is effect between variables

→ There is different between variables.

→ That is the relationship between variables is ~~positive~~ positive.

→ So it is a opposite of null hypothesis

Alternative hypothesis is

$$\boxed{H_1 = \mu_1 \neq \mu_2}$$ (or)

$$\boxed{H_1 = \mu_A \neq \mu_B}$$ or

$$\boxed{H_1 = B_1 < B_2}$$ or

$$\boxed{H_1 = B_1 > B_2}$$

Example:

(12)

i) There is significant variation in the variable of cost between different size of farm holders.

ii) There is significant variation between in the variables of propensity to consume between labourers and farmers.

iii) There is significant different between the income of government employees and private employees.

iv) There is positive correlation between Teachers high level motivation and improvement of students performance on Test papers.

so $\boxed{H_1 = \mu_A \neq \mu_B}$

13 Criteria of good Hypothesis : (Kothari)

i) Hypothesis should be clear and precise.

If the hypothesis is not clear and precise the inferences drawn on its basis cannot be taken as reliable

ii) Hypothesis should be capable of being tested.

iii) Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.

iv) Hypothesis ~~must~~ should be limited in scope and must be specific.

v) Hypothesis should be stated as far possible in most simple terms so that the same is ~~ea~~ easily understood by all concerned.

(16)

(14)

vi) Hypothesis should be consistent with most known facts.
ie: it must be consistent with a substantial body of established facts.
In other words, it should be one which judges accept as being the most likely.

vii) Hypothesis should be amenable to testing within a reasonable time.

viii) Hypothesis must explain the facts that gave rise to the need for explanation.
2 That is, the hypothesis must actually explain what it claims to explain.

V. It should have empirical reference

15

13

Type-I error & Type II error (β)

Level of Significance

1. In the context of testing of hypothesis we can make two types of errors. They are: 1) Type-I error and 2) Type II error

2. Type I error (α) means rejection of hypothesis (Teacher-student 2. Cost of production - Types of farmers - Income of self and wage employees) which should have been accepted. - That is rejection of hypothesis which is true

3. Type II error (β) means accepting the hypothesis (Relationship between variables) which should have been rejected.

4. That is acceptance of hypothesis which is false

16

4 Decision process

→ The decision to accept or reject the null hypothesis H_0 is made on the basis of the information of the sample data.

5 There are two possible types of errors in the test of a hypothesis.

2 That is Type I & Type II error.

3 They are limited with chance of making error 1% and 5% level of significance.

8 Type I Error : Rejection of hypothesis which is true.

Type II error : Acceptance of hypothesis which is false.

2 So the four possible situations that arise in any testing hypothesis.

That is

MBA

Table

Actual	Decision from Sample	
	Accept H_0	Reject H_0
H_0 is true	Correct decision (No error) probability $= 1-\alpha$	Wrong (Type I error) probability $= \alpha$
H_0 is false	Wrong (Type II error) probability $= \beta$	Correct decision (No error) probability $= 1-\beta$

IT
→ Rejection
→ True
→ ~~~~ Acceptance
→ False.

Producer risk

18

1. In Type I error P (Reject H_0 | H_0) which is denoted by α.

2. α : this statistic it is called a producer's risk

3. In Type II error P (Accept H_0 | H_1) which is denoted by β.

4. β : this statistical it is called Consumer's risk.

Scanned by CamScanner

Characteristics of hypothesis: Hypothesis must possess the following characteristics:

(i) Hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.

(ii) Hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programmes have bogged down. Some prior study may be done by researcher in order to make hypothesis a testable one. A hypothesis is testable if other deductions can be made from it which, in turn, can be confirmed or disproved by observation.

(iii) Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.

(iv) Hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.

(v) Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.

(vi) Hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of established facts. In other words, it should be one which judges accept as being the most likely.

(vii) Hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis, if the same cannot be tested in reasonable time for one cannot spend a life-time collecting data to test it.

(viii) Hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must actually explain what it claims to explain; it should have empirical reference.

10.2 BASIC CONCEPTS CONCERNING TESTING OF HYPOTHESIS

Basic concepts in the context of testing of hypotheses need to be explained.

10.2.1 Null Hypothesis and Alternative Hypothesis

In the context of statistical analysis, we often talk about null hypothesis and alternative hypothesis. If we are to compare method *A* with method *B* about its superiority and if we proceed on the assumption that both methods are equally good, then this assumption is termed as the null hypothesis. As against this, we may think that the method *A* is superior or the method *B* is inferior, we are then stating what is termed as alternative hypothesis. The null hypothesis is generally symbolized as H_0 and the alternative hypothesis as H_1. Suppose we want to test the hypothesis that the population mean (μ) is equal to the hypothesised mean $(\mu_0) = 100$. Then we would say that the null hypothesis is that the population mean is equal to the hypothesised mean 100 and symbolically we can express as:

$$H_0: \mu = \mu_0 = 100$$

If our sample results do not support this null hypothesis, we should conclude that something else is true. What we conclude rejecting the null hypothesis is known as alternative hypothesis. In other words, the set of alternatives to the null hypothesis is referred to as the alternative hypothesis. If we accept H_0, then we are rejecting H_1 and if we reject H_0, then we are accepting H_1. For $H_0: \mu = \mu_0 = 100$, we may consider three possible alternative hypotheses as follows:

Table 10.1

Alternative hypothesis	*To be read as follows*
$H_1: \mu \neq \mu_0$	(The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100)
$H_1: \mu > \mu_0$	(The alternative hypothesis is that the population mean is greater than 100)
$H_1: \mu < \mu_0$	(The alternative hypothesis is that the population mean is less than 100)

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn (the researcher must avoid the error of deriving hypotheses from the data that he collects and then testing the hypotheses from the same data). Alternative and null hypotheses are the statements about unknown population parameters. In null hypothesis, we should always have 'equal to' sign. Null hypothesis is the specific statement about the parameter, e.g., $H_0 : \mu = 50$.

Alternative hypothesis is usually the one which one wishes to prove and the null hypothesis is the one which one wishes to disprove. Thus, a null hypothesis represents the hypothesis we are trying to reject, and alternative hypothesis represents all other possibilities.

10.2.2 Type I and Type II Errors

In the context of testing of hypotheses, these are basically two types of errors we can make. We may reject H_0 when H_0 is true and we may accept H_0 when in fact H_0 is not true. The former i known as Type I error and the latter as Type II error. In other words, Type I error means rejection c hypothesis which should have been accepted and Type II error means accepting the hypothes which should have been rejected. All the possibilities in decision making using hypothesis testing a given in the tabular form as below:

Table 10.2

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Accept H_0	No Error	Type II Error
	Probability $= 1 - \alpha$	Probability $= \beta$
Reject H_0	Type I Error	No Error
	Probability $= \alpha$	Probability $= 1 - \beta$

Also, the size of Type I Error is given by the probability of Type I error, P (Reject $H_0 | H_0$) which is denoted by α. In statistical quality control. Similarly, the size of Type II Error is given by the probability of Type II error, P (Accept $H_0 | H_1$) which is denoted by β. In statistical quality control, α is called as producer's risk and β is called as consumer's risk.

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means that there are about 5 chances in 100 that we will reject H_0 when H_0 is true. We can control Type I error just by fixing it at a lower level. For instance, if we fix it at 1 per cent, we will say that the maximum probability of committing Type I error would only be 0.01.

But with a fixed sample size, n, when we try to reduce Type I error, the probability of committing Type II error increases. Both types of errors cannot be reduced simultaneously. There is a trade-off between two types of errors which means that the probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error. To deal with this trade-off in business situations, decision-makers decide the appropriate level of Type I error by examining the costs or penalties attached to both types of errors.

10.2.3 Level of Significance

As discussed, level of significance (α) is the probability of Type I error. This is very important concept in the context of hypothesis testing. It is always some percentage (usually 5%) which should be chosen wit great care, thought and reason. In case we take the significance level at 5 per cent, then this implies that H_0 will be rejected when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if H_0 is true. In other words, the 5 per cent level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H_0) happens to be true. Thus the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

(ii) $H_0 : \mu = \mu_0$ Against $H_a : \mu > \mu_0$ or $H_0 : \mu \leq \mu_0$ Against $H_1 : \mu > \mu_0$

(iii) $H_0 : \mu = \mu_0$ Against $H_a : \mu < \mu_0$ or $H_0 : \mu \geq \mu_0$ Against $H_1 : \mu < \mu_0$

Based on the sign in alternative hypothesis ($\neq$, >, or <), we have three different tests. When we have '$\neq$' sign in alternative hypothesis, we have two-tailed test; when we have '>' sign in alternative hypothesis, we have right-tailed test; and for '<' sign in alternative hypothesis, we have left-tailed test.

10.3 TESTING THE HYPOTHESIS

Given a hypothesis H_0 and an alternative hypothesis H_1, we make a rule which is known as decision rule according to which we accept H_0 (i.e., reject H_1) or reject H_0 (i.e. accept H_1). For example, suppose we want to examine that the mean age of the people in a city is 40 years. In order to conduct the hypothesis testing, we need to be a bit more specific if we wish to examine that

(i) the mean age of the people in a city is 40 years or not; or

(ii) the mean age of the people in a city is 40 years or higher; or

(iii) the mean age of the people in a city is 40 years or lower.

For testing above claims, we first setup the hypotheses which are given as below:

(i) $H_0 : \mu = 40$ Against $H_1 : \mu \neq 40$ (for claim (i))

(ii) $H_0 : \mu = 40$ Against $H_1 : \mu > 40$ (for claim (ii))

(iii) $H_0 : \mu = 40$ Against $H_1 : \mu < 40$ (for claim (iii))

Now we draw a probabilistic (or random) sample of a size from the aforesaid population. Size of the sample is already known. Since we are testing the claim about population mean, we obtain sample mean as sample mean is a "good" estimate of population mean. Suppose sample mean comes to be 20 years. This is significantly lower than the claimed mean population age 40 years. If the claim (H_0) is true, the probability of getting such a different sample mean would be very small. Getting a sample mean of 20 is very unlikely. So, when we get sample mean as 20, we do not believe on the claim (H_0). If the sample mean is close to the assumed population mean, H_0 is accepted. If the sample mean is far-off from the assumed population mean, H_0 is rejected. How far is "far enough" to reject H_0? The concept of critical value is used to decide on this (to be discussed later).

Also, in a hypothesis test, we initially assume that the null hypothesis is true and we proceed to try to reject null hypothesis using the sample. In case, when we cannot reject the null hypothesis it only means that sample has insufficient information to reject null hypothesis at given level of significance. It does not mean that the parametric statement under alternative hypothesis is true. Therefore, whenever we say that the null hypothesis is accepted, it only means that null hypothesis cannot be rejected as there is no statistical evidence against it.

...CAL REGION