

LECTURE - (21)

Agenda:

- ① Formal definition of a statistical test
- ② Test statistic and rejection region.

In hypothesis testing, the objective is to test a hypothesis concerning the values of one or more unknown parameters of a population. Here are a few examples.

- ① A medical researcher wants to test/validate a hypothesis that a new drug is more effective than the existing drug to treat a disease, for a given population.
- ② A pollster wants to test/validate a claim made by one of the candidates, say candidate A, that more than 50% of the voters in a given population want to vote for him.
- ③ A quality control engineer wants to test/validate a hypothesis that a new assembly method produces less than 5% defective items.

The most common, scientific, and cost-effective way to ~~research~~ validate such hypothesis is to obtain a ~~small~~ sample from the population, and use the sample to examine whether there is enough evidence to support the hypothesis.

OFTEN, THE PROPOSED HYPOTHESIS IS AN ALTERNATIVE, OR A CHALLENGER, TO AN EXISTING, WELL-ESTABLISHED HYPOTHESIS.

Definition: The existing, well-established hypothesis is known as the NULL HYPOTHESIS (H_0).

Definition: The alternative, challenging hypothesis is known as the ALTERNATIVE HYPOTHESIS (H_A).

Examples (continued):

- ① H_0 : Both the new and existing drug have the same effect.
 H_A : The new drug works better.
- ② H_0 : Candidate A is evenly tied with the other candidate.
 H_A : Candidate A has more than 50% of the votes.

③ H_0 : The new assembly method still produces 5% defective items

H_A : The new assembly method produces less than 5% defective items

How do we decide based on the observed sample, if there is enough evidence to support the alternative hypothesis H_A ?

Definition: The TEST STATISTIC is a function of the data on which the decision of the statistical test is based.

SINCE THE DATA IS DRAWN FROM THE POPULATION BY A RANDOM MECHANISM, THE TEST STATISTIC IS A RANDOM QUANTITY.

Definition: The REJECTION REGION specifies the values of the test statistic for which the null hypothesis H_0 is to be rejected in favor of the alternative hypothesis H_A .

Hence, a formal STATISTICAL TESTING PROBLEM has 4 ingredients.

1. Null hypothesis (H_0).
2. Alternative hypothesis (H_A).
3. Test statistic.
4. Rejection region.

Let us study Example 2 (the voting example) in detail. We want to check if the proportion of voters, say p , supporting candidate A is more than 0.5. To do this, we just pick a random sample of n voters and ask each of them if they support candidate A.

Data: Y_1, Y_2, \dots, Y_n , where

$$Y_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ sampled voter supports} \\ & \text{candidate A,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, each Y_i has a Bernoulli(p) distribution.

Assume that the population is large enough, so that Y_1, Y_2, \dots, Y_n are independent random variables.

H_0 : Both candidates are evenly placed, or $p = 0.5$.

H_A : Candidate A has more than 50% votes, or $p > 0.5$.

The most natural statistic to decide about the population proportion p is the sample proportion

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i. \text{ Clearly, large values of } \bar{Y}$$

should favor the alternative hypothesis. But we have to decide on a cutoff. How large should \bar{Y} be such that we have sufficient evidence that the alternative hypothesis should be true?

As of now, there seems to be no perfect answer to this question. Let us make the following subjective choice. If $\bar{Y} > 0.6$, then we will accept H_A , otherwise we will accept H_0 . This is ^{slightly} a conservative choice, as we really want to see strong evidence before accepting candidate A's claim. Hence the rejection region for our test is

$$\{ \bar{Y} > 0.6 \}.$$

As with any procedure based on a random variable (here, the test statistic) there is always a chance that we make an error. Here, we are concerned with two types of error.

TYPE I ERROR: We choose H_A when H_0 is true.

$$\alpha \triangleq P(\text{We accept } H_A \mid H_0 \text{ is true})$$

In our example,

$$\alpha = P(\bar{Y} > 0.6 \mid H_0 \text{ is true})$$

$$= P\left(\sum_{i=1}^n Y_i > 0.6n \mid H_0 \text{ is true}\right)$$

$$= P\left(\sum_{i=1}^n Y_i > 0.6n \mid p = 0.5\right)$$

Note that $\sum_{i=1}^n Y_i$ is a Binomial $(n, 0.5)$ random variable if $p = 0.5$. Hence,

$$\alpha = \sum_{j > 0.6n} \binom{n}{j} (0.5)^j (1-0.5)^{n-j}$$

TYPE II ERROR: We choose H_0 when H_A is true.

$$\beta \triangleq P(\text{We accept } H_0 \mid H_A \text{ is true})$$

In our example, suppose $p = 0.6$. For this particular alternative

$$\beta_{p=0.6} = P(\bar{Y} \leq 0.6 \mid p = 0.6) = \sum_{j \leq 0.6n} \binom{n}{j} (0.6)^j (1-0.6)^{n-j}$$

Of course, there is a wide range of possible values of p under the alternative hypothesis, and there is a TYPE II ERROR ^{probability} β for each such possible p .

TAKE AWAY: If the alternative hypothesis

claims the parameter lies in an interval (in our example, $H_A: p \in (0.5, 1]$), then we need to look at the TYPE II ERROR probability β for each possible parameter value in this interval.

