Unit 7: Hypothesis testing

Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. First, a tentative assumption is made about the parameter or distribution. This assumption is called the null hypothesis and is denoted by H_0. An alternative hypothesis (denoted H_a), which is the opposite of what is stated in the null hypothesis, is then defined. The hypothesis-testing procedure involves using sample data to determine whether or not H_0 can be rejected. If H_0 is rejected, the statistical conclusion is that the alternative hypothesis H_a is true.

For example, assume that a radio station selects the music it plays based on the assumption that the average age of its listening audience is 30 years. To determine whether this assumption is valid, a hypothesis test could be conducted with the null hypothesis given as H_0: $\mu = 30$ and the alternative hypothesis given as H_a: $\mu \neq 30$. Based on a sample of individuals from the listening audience, the sample mean age, $\bar{x}$, can be computed and used to determine whether there is sufficient statistical evidence to reject H_0. Conceptually, a value of the sample mean that is “close” to 30 is consistent with the null hypothesis, while a value of the sample mean that is “not close” to 30 provides support for the alternative hypothesis. What is considered “close” and “not close” is determined by using the sampling distribution of $\bar{x}$.

Ideally, the hypothesis-testing procedure leads to the acceptance of H_0 when H_0 is true and the rejection of H_0 when H_0 is false. Unfortunately, since hypothesis tests are based on sample information, the possibility of errors must be considered. A type I

error corresponds to rejecting H_0 when H_0 is actually true, and a type II error corresponds to accepting H_0 when H_0 is false. The probability of making a type I error is denoted by α, and the probability of making a type II error is denoted by β.

In using the hypothesis-testing procedure to determine if the null hypothesis should be rejected, the person conducting the hypothesis test specifies the maximum allowable probability of making a type I error, called the level of significance for the test. Common choices for the level of significance are $\alpha = 0.05$ and $\alpha = 0.01$. Although most applications of hypothesis testing control the probability of making a type I error, they do not always control the probability of making a type II error. A graph known as an operating-characteristic curve can be constructed to show how changes in the sample size affect the probability of making a type II error.

A concept known as the *p*-value provides a convenient basis for drawing conclusions in hypothesis-testing applications. The *p*-value is a measure of how likely the sample results are, assuming the null hypothesis is true; the smaller the *p*-value, the less likely the sample results. If the *p*-value is less than α, the null hypothesis can be rejected; otherwise, the null hypothesis cannot be rejected. The *p*-value is often called the observed level of significance for the test.

A hypothesis test can be performed on parameters of one or more populations as well as in a variety of other situations. In each instance, the process begins with the formulation of null and alternative hypotheses about the population. In addition to the population mean, hypothesis-testing procedures are available for population parameters such as proportions, variances, standard deviations, and medians.

Hypothesis tests are also conducted in regression and correlation analysis to determine if the regression relationship and the correlation coefficient are statistically significant (see below Regression and correlation analysis). A goodness-of-fit test refers to a hypothesis test in which the null hypothesis is that the population has a

specific probability distribution, such as a normal probability distribution. Nonparametric statistical methods also involve a variety of hypothesis-testing procedures.

Comprehension Exercises

Choose a, b, c or d which best completes each item.

1) The _____ hypothesis is a tentative assumption about the parameter or distribution of the population.

a) zero b) test c) null d) alternative

2) A _____ error corresponds to rejecting H_0 when H_0 is actually true.

a) testing b) sampling c) type II d) type I

3) The level of significance for a hypothesis testing procedure is the _____ allowable probability of making a type I error.

a) minimum b) maximum c) mean d) median

4) The *p*-value is a measure of how _____ the sample results are, assuming the null hypothesis is true.

a) likely b) unlikely c) small d) large

Words to Learn

Find the Persian equivalents of the following terms and expressions.

hypothesis	test	sufficient	conceptually
assumption	tentative	evidence	ideally
null	type I error	consistent	actually
alternative	type II error	information	unfortunately
opposite	level	concept	audience
reject	operating	otherwise	valid
accept	characteristic	regression	false
rejection	curve	correlation	conclusion
acceptance	*p*-value	nonparametric	acceptable
significance	variety	goodness-of-fit	allowable