

المحاضرة (5)

Introduction to Hypothesis Testing

Terms, Concepts.

A. In general, we do not know the true value of population parameters – they must be estimated. However, we do have hypotheses about what the true values are.

B. The major purpose of hypothesis testing is to choose between two competing hypotheses about the value of a population parameter. For example, one hypothesis might claim that the wages of men and women are equal, while the alternative might claim that men make more than women.

C. The hypothesis actually to be tested is usually given the symbol H_0 , and is commonly referred to as the null hypothesis. As is explained more below, the null hypothesis is assumed to be true unless there is strong evidence to the contrary – similar to how a person is assumed to be innocent until proven guilty.

D. The other hypothesis, which is assumed to be true when the null hypothesis is false, is referred to as the alternative hypothesis,

and is often symbolized by H_A or H_1 . Both the null and alternative hypothesis should be stated before any statistical test of significance is conducted. In other words, you technically are not supposed to do the data analysis first and then decide on the hypotheses afterwards.

E. In general, it is most convenient to always have the null hypothesis contain an equals sign, e.g. $H_0: \mu = 100$ $H_A: \mu > 100$

F. The true value of the population parameter should be included in the set specified by H_0 or in the set specified by H_A . Hence, in the above example, we are presumably sure μ is at least 100.

G. A statistical test in which the alternative hypothesis specifies that the population parameter lies entirely above or below the value specified in H_0 is a one-sided (or one-tailed) test, e.g. $H_0: \mu = 100$ $H_A: \mu > 100$

H. An alternative hypothesis that specified that the parameter can lie on either side of the value specified by H_0 is called a two-sided (or two-tailed) test, e.g. $H_0: \mu = 100$ $H_A: \mu \neq 100$

I. Whether you use a 1-tailed or 2-tailed test depends on the nature of the problem. Usually we use a 2-tailed test. A 1-tailed test typically requires a little more theory.

For example, suppose the null hypothesis is that the wages of men and women are equal. A two-tailed alternative would simply state that the wages are not equal – implying that men could make more than women, or they could make less. A one-tailed alternative would be that men make more than women. The latter is a stronger statement and requires more theory, in that not only are you claiming that there is a difference, you are stating what direction the difference is in.

J. In practice, a 1-tailed test such as $H_0: \mu = 100$ $H_A: \mu > 100$ is tested the same way as $H_0: \mu \neq 100$ $H_A: \mu > 100$ For example, if we conclude that $\mu > 100$, we must also conclude that $\mu > 90$, $\mu > 80$, etc.

Suppose μ really does equal 100. But, suppose the researcher accepts H_A instead. A type I error has occurred.

Or, suppose $\mu = 105$ - but the researcher accepts H_0 . A type II error has occurred.

The following tables from Harnett help to illustrate the different types of error.

		The true situation is	
		H_0 is true	H_a is true
Action	Reject H_0 (Accept H_a)	Type I error	Correct decision
	Reject H_a (Accept H_0)	Correct decision	Type II error

		The true situation is	
		Not guilty (H_0)	Guilty (H_a)
Action	Jury finds guilty (Accept H_a)	Type I error	Correct decision
	Jury finds not guilty (Accept H_0)	Correct decision	Type II error

F. α and β are not independent of each other - as one increases, the other decreases. However, increases in N cause both to decrease, since sampling error is reduced.

G. In this class, we will primarily focus on Type I error. But, you should be aware that Type II error is also important. A small sample size, for example, might lead to frequent Type II errors, i.e. it could be that your (alternative) hypotheses are right, but because your sample is so small, you fail to reject the null even though you should.