Hypothesis Testing

Introduction to Hypothesis Testing

I. Terms and Concepts:

1. 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.
2. 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.
3. The hypothesis actually to be tested is usually given the symbol H0, 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.
4. 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 HA or H1. 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.
5. In general, it is most convenient to always have the null hypothesis contain an equals sign, e.g.
   H0: $\mu = 100$
   HA: $\mu > 100$
6. The true value of the population parameter should be included in the set specified by H0 or in the set specified by HA. Hence, in the above example, we are presumably sure μ is at least 100.
7. A statistical test in which the alternative hypothesis specifies that the population parameter lies entirely above or below the value specified in H0 is a one-sided (or one-tailed) test, e.g.
   H0: $\mu = 100$

HA: μ > 100

8. An alternative hypothesis that specified that the parameter can lie on either side of the value specified by H0 is called a two-sided (or two-tailed) test, e.g.

   H0: μ = 100

   HA: μ <> 100

9. 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.

10. In practice, a 1-tailed test such as
    H0: μ = 100
    HA: μ > 100
    is tested the same way as
    H0: μ # 100
    HA: μ > 100

    For example, if we conclude that μ > 100, we must also conclude that μ > 90, μ > 80, etc.

II. The decision problem.

A. How do we choose between H0 and HA? The standard procedure is to assume H0 is true - just as we presume innocent until proven guilty. Using probability theory, we try to determine whether there is sufficient evidence to declare H0 false.

B. We reject H0 only when the chance is small that H0 is true. Since our decisions are based on probability rather than certainty, we can make errors.

C. Type I error - We reject the null hypothesis when the null is true. The probability of Type I error = α. Put another way, α = Probability of Type I error = P(rejecting H0 | H0 is true) Typical values chosen for α are .05 or .01. So, for example, if α = .05, there is a 5% chance that, when the null hypothesis is true, we will erroneously reject it.

D. Type II error - we accept the null hypothesis when it is not true. Probability of Type II error = ß. Put another way, ß = Probability of Type II error = P(accepting H0 | H0 is false)

E. EXAMPLES of type I and type II error:
H0: μ = 100
HA: μ <> 100

Suppose μ really does equal 100. But, suppose the researcher accepts HA instead. A type I error has occurred. (or). Suppose μ = 105 - but the researcher accepts H0. 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 ß 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.