Sec. 10.1 A Brief Introduction to Hypothesis Testing

- In our study of confidence intervals, we sought some interval that was considered a set of "reasonable" values that a parameter might be.
- Frequently in scientific studies, the researcher presents a specific claim about the value of some parameter of interest.
- Then data are gathered, and based on these sample data we test the claim (or <u>hypothesis</u>) about the parameter.

<u>Example 1</u>: We gather experimental data to test the claim that drug 1 is equally effective, on average, as drug 2.

<u>Example 2</u>: We gather survey data to test the claim that no more than 50% of registered voters support the governor's latest policy.

- These are examples of <u>hypothesis testing</u>.

10.2 Elements of a Statistical Test of Hypotheses

- Statistical hypotheses are statements about one or more parameters.
- We have two kinds of hypothesis:

① The ________ hypothesis (or ________ hypothesis, denoted H_a) represents a theory that the researcher suspects to be true, or seeks evidence to "prove."

② The ______ hypothesis (denoted H_0) is the converse (opposite) of H_a.

- H_0 often represents some "previously held belief," "status quo," or "lack of effect."

Example 1 again: The parameter of interest is

Example 2 again: The parameter of interest is

- If we gather a set of sample data and it would be highly unlikely to observe such data if H_0 were true, then we have evidence against ____ and in favor of ____.
- A common approach to hypothesis testing is to calculate a test statistic (some relevant function of the data $Y_1, \dots, Y_n$) and see whether it falls in the rejection region (a set of values for the test statistic that are highly improbable if H_0 were true).
- This approach is covered in detail in STAT 513.

Errors in Hypothesis Testing

- "Proving" the research hypothesis H_a does not imply removing all doubt (as in a mathematical proof).
- If H_0 is true, it is still possible to observe a "highly unlikely" test statistic value (which would cause the researcher to reject H_0 and conclude H_a).

- If H_0 is rejected when H_0 is in fact true, this is called a ______________.

- We set up the rejection region so that the probability of a Type I error is some small number α, like:

- This value α is chosen by the researcher and is called the __________ __________.

(A Type II error occurs when H_0 is not rejected, yet H_0 is false.)

10.5 Relationship Between Hypothesis Testing and Confidence Intervals

- Because confidence intervals represent a set of reasonable values for a parameter, we can often use a CI to determine whether to reject H_0 and conclude H_a.

"Two-Sided" Hypothesis Tests

- For any number θ_0, we can test

at significance level α by constructing a $100(1-\alpha)\%$ CI $[\hat{\theta}_L, \hat{\theta}_u]$ for θ. We reject H_0 (and conclude H_a) <u>if</u> <u>and</u> <u>only</u> <u>if</u>:

- Otherwise, we fail to reject H_0.

<u>Proof</u>:

"One-Sided" Hypothesis Tests

- We can similarly test

at significance level α via a $100(1-\alpha)\%$ <u>lower</u> <u>confidence</u> <u>bound</u> (or one-sided interval): $[\hat{\theta}_L, \infty)$. Again, reject H_0 and conclude H_a if and only if:

- To test

at significance level α, we can use the $100(1-\alpha)\%$ <u>upper confidence bound</u>. We reject H_0 and conclude H_a if and only if

<u>Example 1</u>: Assuming the alcohol percentages of beers are approximately normal, test (at $\alpha = .05$) whether the mean alcohol percentage of (non-light) beers exceeds 5 percent. A random sample of 16 beers resulted in $\bar{Y} = 5.34$ and $s = 0.8483$.

We test:

<u>Example 1(a)</u>: For the beer data, suppose our research question had been to test (at $\alpha = .05$) whether the population <u>variance</u> differed from 2.

We test:

From Chapter 8,

<u>Example 2</u>: Recall exercise in Section 8.6 notes. Is the true mean number of ships per day in the summer different from the mean in the winter? Use $\alpha = .10$.

- This equivalency of CIs and hypothesis tests does not exactly work for tests about proportions (because of standard error part).