5. Hypothesis Tests About the Mean of a Normal Population When σ^2 is Not Known

In Section 3 we used the *t*-distribution as a basis for confidence interval estimation for the mean of a normal population when the population variance σ^2 is not known. Similarly, when testing hypothesis, if σ^2 is not known, we use a *t*-statistic. From (3.5) we know that

$$t = \frac{b-\beta}{\hat{\sigma}/\sqrt{T}} \sim t_{(T-1)} \tag{5.1}$$

When testing the null hypothesis $H_0: \beta = c$ against the alternative hypothesis $H_1: \beta \neq c$ the test statistic

$$t = \frac{b-c}{\hat{\sigma}/\sqrt{T}} \sim t_{(T-1)} \tag{5.2}$$

if the null hypothesis is true. Following the same logic as in Section 4, we reject H_0: $\beta = c$ *if* $|t| \geq t_c$, or if the *p*-value is less than the level of significance α. The rejection rules and critical values from the *t*-distribution are shown in Figure 4.

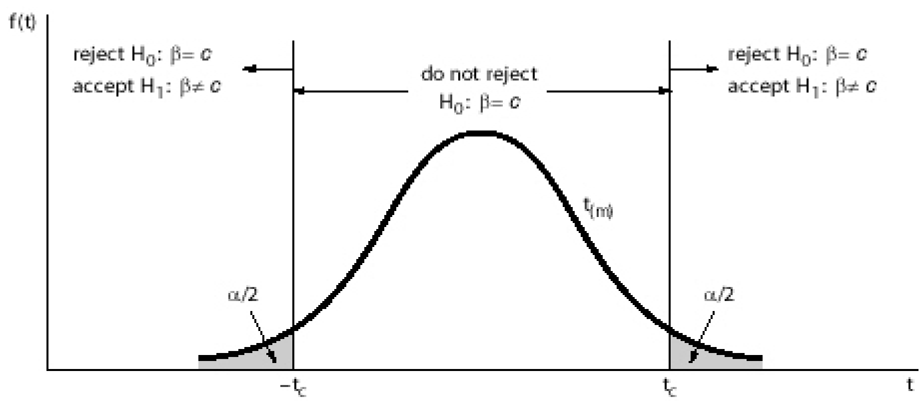

Figure 4 Rejection region for testing $H_0: \beta = c$ against $H_1: \beta \neq c$

(5.1) An Empirical Example of a Two-Tailed Test

Let us illustrate by testing the null hypothesis that the population hip size is 16 inches, against the alternative that it is not, using the hip data. We will follow the steps outlined in the testing format suggested in Section 4.

1. The null hypothesis is H_0: $\beta = 16$. The alternative hypothesis is H_1: $\beta \neq 16$.

2. The test statistic $t = \dfrac{b-16}{\hat{\sigma}/\sqrt{T}} \sim t_{(T-1)}$ *if the null hypothesis is true.*

3. Let us select α=.05. The critical value t_c is 2.01 for a t-distribution with (T–1) = 49 degrees of freedom. Thus we will reject the null hypothesis in favor of the alternative if $t \geq 2.01$ or $t \leq -2.01$, or equivalently , if $|t| \geq 2.01$

4. Using the hip data, the least squares estimate of β is b = 17.158, with estimated variance $\hat{\sigma}^2 = 3.267$, so $\hat{\sigma} = 1.807$. The value of the test statistic is $t = \dfrac{17.158-16}{1.807/\sqrt{50}} = 4.531$.

5. Conclusion: Since t=4.531 > t_c=2.01 we *reject* the null hypothesis. The sample information we have is *incompatible* with the hypothesis that β = 16, or, that the population mean hip size is 16 inches. Equivalently, for this test the p-value is p=.000038 < α=.05 and on this basis we reject the null hypothesis.

(5.2) A Relationship Between Hypothesis Testing and Interval Estimation

There is an *algebraic* relationship between two-tailed hypothesis tests and confidence interval estimates that is sometimes useful. Suppose that we are testing the null hypothesis $H_0 : \beta = c$ against the alternative $H_1 : \beta \neq c$. If we *fail to reject* the null hypothesis at the α level of significance, then the value c will fall *within* a (1–α)×100% confidence interval estimate of β. Conversely, if we reject the null hypothesis, then c will fall *outside* the (1–α)×100% confidence interval estimate of β. This algebraic relationship is true because we fail to reject the null hypothesis when $-t_c \leq t \leq t_c$, or when

$$-t_c \leq \frac{b-c}{\hat{\sigma}/\sqrt{T}} \leq t_c$$

which, when rearranged becomes

$$b - t_c \frac{\hat{\sigma}}{\sqrt{T}} \leq c \leq b + t_c \frac{\hat{\sigma}}{\sqrt{T}}$$

The endpoints of this interval are the same as the endpoints of a (1–α)×100% confidence interval estimate of β. Thus for any value of c within the interval we do not reject $H_0 : \beta = c$ against the alternative

$H_1 : \beta \neq c$. For any value of c outside the interval we reject $H_0 : \beta = c$ and accept the alternative $H_1 : \beta \neq c$.

This relationship can be handy if you are given only a confidence interval and want to determine what the outcome of a two-tailed test would be. You can verify that the test statistics z in Table 4 lead us to reject the null hypothesis for the same samples in Table 2 which do not contain the value $\beta = 10$.

We have two comments about this relationship between interval estimation and hypothesis testing:

1. The relationship is between confidence intervals and *two-tailed* tests. It does not apply to one-tailed tests.
2. A confidence interval is an *estimation* tool; that is, it is an interval *estimator*. A hypothesis test about one or more parameters is a completely separate form of inference, with the only connection being that the test statistic incorporates the least squares estimator. To test hypotheses you should carry out the steps outlined in Section 4 and ***should not*** compute and report an interval estimate.

(5.3) One-Tailed Tests

We have focused so far on testing hypotheses of the form H_0: $\beta = c$ against the alternative H_1: $\beta \neq c$. This kind of test is called a two-tailed test since portions of the rejection region are found in both tails of the test statistic's distribution. One-tailed tests are used to test H_0: $\beta = c$ against the alternative H_1: $\beta > c$, or H_1: $\beta < c$.

The logic of one-tailed tests is identical to that for the two-tailed tests that we have studied. The test statistic is the same, and given by (5.2). What *is* different is the selection of the rejection region and the computation of the p-values. For example, to test H_0: $\beta = c$ against the alternative H_1: $\beta > c$ we select the rejection region to be values of the test statistic t that support the *alternative* hypothesis and which are *unlikely* if the null hypothesis is true. *Large* values of the t-statistic are unlikely if the null hypothesis is true. We define the rejection region to be values of t greater than a critical value t_c, from a t-distribution with T–1 degrees of freedom, such that $P(t \geq t_c) = \alpha$, where α is the level of significance of the test and the probability of a Type I error. See Figure 5 below.

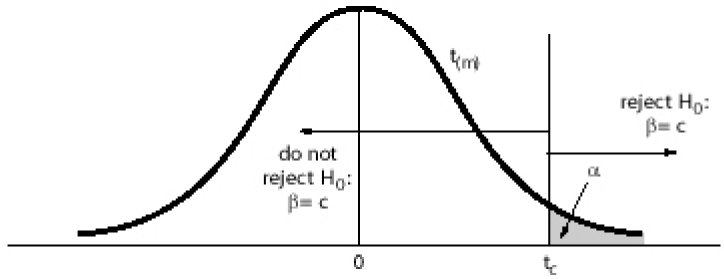

Figure 5 Critical Value for One-Tailed Test H_0: β = c vs. H_1: β > c

The decision rule for this one-tailed test is, "Reject H_0: β = c and accept the alternative H_1: β > c if $t \geq t_c$." If $t < t_c$ then we do not reject the null hypothesis.

Computation of the p-value is similarly confined to one tail of the distribution of the test statistic, though its interpretation is exactly as before. For testing H_0: β = c against the alternative H_1: β > c the p-value is computed by finding the probability that the test statistic is greater than or equal to the computed sample value of the test statistic.

(5.4) An Empirical Example of a One-Tailed Test

Let us illustrate by testing the null hypothesis that the population hip size is 16 inches, against the alternative that it is *greater* than 16 inches, using the hip data. We will follow the steps outlined in the testing format suggested in Section 4.

6. The null hypothesis is H_0: β = 16. The alternative hypothesis is H_1: β > 16
7. The test statistic $t = \dfrac{b-16}{\hat{\sigma}/\sqrt{T}} \sim t_{(T-1)}$ *if the null hypothesis is true.*
8. Let us select α=.05. The critical value t_c is 1.68 for a t-distribution with (T–1) = 49 degrees of freedom. Thus we will reject the null hypothesis in favor of the alternative if $t \geq 1.68$.
9. Using the hip data, the least squares estimate of β is b = 17.158, with estimated variance $\hat{\sigma}^2 = 3.267$, so $\hat{\sigma} = 1.807$. The value of the test statistic is $t = \dfrac{17.158-16}{1.807/\sqrt{50}} = 4.531$.
10. Conclusion: Since t=4.531 > t_c=1.68 we *reject* the null hypothesis. The sample information we have is *incompatible* with the hypothesis that β = 16, or, that the population mean hip size is 16 inches. Thus we accept the alternative that the population mean hip size is greater than 16 inches, based on

our test which has α=.05 probability of a Type I error. Equivalently, for this test the *p*-value is p=.000019 < α=.05 and on this basis we reject the null hypothesis.

(5.5) A Comment on Stating Null and Alternative Hypotheses

We have noted in the previous sections that a statistical test procedure cannot prove the truth of a null hypothesis. When we fail to reject a null hypothesis, all the hypothesis test can establish is that the information in a sample of data is *compatible* with the null hypothesis. On the other hand, a statistical test can lead us to *reject* the null hypothesis, with only a small probability, α, of rejecting the null hypothesis when it is actually true. Thus rejecting a null hypothesis is a stronger conclusion than failing to reject it.

The null hypothesis is usually stated in such a way that if our theory is correct, then we will reject the null hypothesis. For example, our airplane seat designer has been operating under the assumption (the maintained or null hypothesis) that the population mean hip width is 16 inches. Casual observation suggests that people are getting larger all the time. If we are larger, and if the airline wants to continue accommodate the same percentage of the population, then the seat widths must be increased. This costly change should be undertaken only if there is statistical evidence that the population size is indeed larger. When using a hypothesis test we would like to find out that there is statistical evidence against our current "theory," or if the data are compatible with it. With this goal, we set up the null hypothesis that the population mean is 16 inches, H_0: $\beta = 16$against the alternative that it is greater than 16 inches, H_1: $\beta >$ 16.

You may view the null hypothesis to be too limited in this case, since it is feasible that the population mean hip width is now smaller than 16 inches, or that $\beta < 16$. The hypothesis testing procedure for the testing the null hypothesis that H_0: $\beta \leq 16$ against the alternative hypothesis H_1: $\beta > 16$ is *exactly the same* as testing H_0: $\beta = 16$ against the alternative hypothesis H_1: $\beta > 16$. The test statistic, rejection region and *p*-value are exactly the same. For a one-tailed test you can form the null hypothesis in either of these ways. What counts is that the alternative hypothesis is properly specified.

Finally, it is important to set up the null and alternative hypotheses *before* you carry out the regression analysis. Failing to do so can lead to errors in formulating the alternative hypothesis. Suppose that we wish to test whether $\beta > 16$ and the least squares estimate of β is b= 15.5. Does that mean we should set up the alternative $\beta < 16$, to be consistent with the estimate? The answer is *no*. The alternative is formed to state the conjecture that we wish to establish, $\beta > 16$.