ONE-SIDED TESTS

In testing the null hypothesis that $\mu = \mu_0$, we have chosen a test that calls for rejection when $\bar{X}$ is far from μ_0. That is, a very small value of $\bar{X}$ or a very large value appears to make it unlikely that μ (which $\bar{X}$ is estimating) could equal μ_0. However, what happens when the alternative hypothesis to $H_0 : \mu = \mu_0$ is $H_1 : \mu > \mu_0$? In this latter case we would not want to reject H_0 when $\bar{X}$ is small (since a small $\bar{X}$ is more likely when H_0 is true than when H_1 is true). Thus, in testing

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu > \mu_0 \quad (10.29)$$

we should reject H_0 when $\bar{X}$, the point estimate of μ_0, is much greater than μ_0. That is, the critical region should be of the following form:

$$C = \{(X_1, X_2, \ldots, X_n) : \bar{X} - \mu_0 > c\} \quad (10.30)$$

Since the probability of rejection should equal α when H_0 is true, we require that c be such that

$$P_{\mu_0}\left(\bar{X} - \mu_0 > c\right) = \alpha \quad (10.31)$$

But since

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{\sigma} \quad (10.32)$$

has a standard normal distribution when H_0 is true, Equation (10.31) is equivalent to

$$P\left(Z > \frac{c\sqrt{n}}{\sigma}\right) = \alpha \quad (10.33)$$

But since

$$P\left(Z > z_\alpha\right) = \alpha \quad (10.34)$$

$$\Rightarrow \quad c = \frac{z_\alpha \sigma}{\sqrt{n}} \quad (10.35)$$

Hence, the test of the hypothesis (10.29) is to reject H_0 if $\bar{X} - \mu_0 > z_\alpha \sigma / \sqrt{n}$, and accept otherwise; or, equivalently, to

$$\text{accept } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) \le z_\alpha \tag{10.36}$$

$$\text{reject } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) > z_\alpha \tag{10.37}$$

This is called a *one-sided critical region* (since it calls for rejection only when $\bar{X}$ is large). Correspondingly, the hypothesis testing problem

$$H_0 : \mu = \mu_0 \tag{10.38}$$

$$H_1 : \mu > \mu_0$$

is called a *one-sided testing problem* (in contrast to the two-sided problem that results when the alternative hypothesis is $H_1 : \mu \neq \mu_0$).

To compute the *p*-value in the one-sided test, we first use the data to determine the value of the statistic $\sqrt{n}\left(\bar{X}-\mu_0\right)/\sigma$. The *p*-value is then equal to the probability that a standard normal would be at least as large as this value.

Exercise 5. Suppose in Exercise 1 that we know in advance that the signal value is at least as large as 8. The same signal value is independently sent five times and the mean value received is $\bar{X} = 9.5$. What can be concluded in this case?

To see if the data are consistent with the hypothesis that the mean is 8, we test

$$H_0 : \mu = 8$$

against the one-sided alternative: $H_1 : \mu > 8$

The value of the test statistic is

$$\sqrt{n}\left(\bar{X}-\mu_0\right)/\sigma = \sqrt{5}\left(9.5-8\right)/2 = 1.68$$

and the *p*-value is the probability that a standard normal would exceed 1.68, namely,

$$p-\text{value} = P\left(Z > 1.68\right) = 1-\phi\left(1.68\right) = 0.0465$$

the test would call for rejection at all significance levels greater than or equal to 0.0465, it would, for instance, reject the null hypothesis at the $\alpha = 0.05$ level of significance.

The *p*-value is the probability, assuming the null hypothesis, of seeing data at least as extreme as the experimental data (the value of the test statistic).

The *operating characteristic function* of the one-sided test,

$$\beta(\mu) = P_\mu(\text{accepting } H_0) \tag{10.39}$$

can be obtained as follows:

$$\beta(\mu) = P_\mu\left(\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \le z_\alpha\right) \tag{10.40}$$

$$= P\left(\frac{\bar{X} - \mu - \mu_0}{\sigma / \sqrt{n}} \le z_\alpha - \frac{\mu}{\sigma / \sqrt{n}}\right)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \le z_\alpha + \frac{\mu_0 - \mu}{\sigma / \sqrt{n}}\right)$$

$$= P\left(Z \le z_\alpha + \frac{\mu_0 - \mu}{\sigma / \sqrt{n}}\right), \quad Z \sim N(0,1) \tag{10.41}$$

where the last equation follows since $\sqrt{n}\left(\bar{X} - \mu\right)/\sigma$ has a standard normal distribution. We can write

$$\beta(\mu) = \phi\left(\frac{\mu_0 - \mu}{\sigma / \sqrt{n}} + z_\alpha\right) \tag{10.42}$$

Since ϕ, being a distribution function, is increasing in its argument, it follows that $\beta(\mu)$ decreases in μ, which is intuitively pleasing since it certainly seems reasonable that the larger the true mean μ, the less likely it should be to conclude that $\mu \le \mu_0$.

The test given by Equation (10.36-37),

$$\text{accept } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) \le z_\alpha$$

$$\text{reject } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) > z_\alpha$$

which was designed to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$, can also be used to test, at level of significance α, the one-sided hypothesis

$$H_0 : \mu \le \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0 \qquad (10.43)$$

REMARK

We can also test the one-sided hypothesis

$$H_0 : \mu = \mu_0 \quad (\text{or } \mu \ge \mu_0) \text{ versus } H_1 : \mu < \mu_0 \qquad (10.44)$$

at significance level α by

$$\text{accepting } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) \ge -z_\alpha \qquad (10.45)$$

$$\text{rejecting } H_0 \text{ if } \frac{\sqrt{n}}{\sigma}\left(\bar{X}-\mu_0\right) \le -z_\alpha \qquad (10.46)$$

This test can alternatively be performed by first computing the value of the test statistic $\sqrt{n}\left(\bar{X}-\mu_0\right)/\sigma$. The *p*-value would then equal the probability that a standard normal would be less than this value, and the hypothesis would be rejected at any significance level greater than or equal to this *p*-value.

Exercise 6. All cigarettes presently on the market have an average nicotine content of at least 1.6 mg per cigarette. A firm that produces cigarettes claims that it has discovered a new way to cure tobacco leaves that will result in the average nicotine content of a cigarette being less than 1.6 mg. To test this claim, a sample of 20 of the firm's cigarettes were analyzed. If it is known that the standard deviation of a cigarette's nicotine content is 0.8 mg, what conclusions can be drawn, at the 5 percent level of significance, if the average nicotine content of the 20 cigarettes is 1.54?

We must first decide on the appropriate null hypothesis. The rejection of the null hypothesis is a *strong statement* about the data not being consistent with this hypothesis. In the preceding example we would like to approve the producer's claims only when there is substantial evidence for it, we should take this claim as the alternative hypothesis. That is, we should test

$$H_0 : \mu \geq 1.6 \quad \text{versus} \qquad H_1 : \mu < 1.6$$

Now, the value of the test statistic is

$$\sqrt{n}\left(\bar{X} - \mu_0\right) / \sigma = \sqrt{20}\left(1.54 - 1.6\right) / 0.8 = -0.336$$

and so the *p*-value is given by

$$p - \text{value} = P\left\{Z < -0.336\right\} = 0.368 \ , \quad Z \sim N\left(0,1\right)$$

Since this value is greater than 0.05, the foregoing data do not enable us to reject, at the 0.05 percent level of significance, the hypothesis that the mean nicotine content exceeds 1.6 mg. In other words, the evidence, although supporting the cigarette producer's claim, is not strong enough to prove that claim. ■

REMARKS

(a) There is a direct analogy between *confidence interval estimation* and *hypothesis testing*.

For instance, for a normal population having mean μ and known variance σ^2, we have shown that a $100(1-\alpha)$ percent confidence interval for μ is given by

$$\mu \in \left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) \qquad (10.47)$$

where $\bar{x}$ is the observed sample mean. More formally, the confidence interval statement is equivalent to

$$P\left\{ \mu \in \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) \right\} = 1 - \alpha \qquad (10.48)$$

Hence, if $\mu = \mu_0$, then the probability that μ_0 will fall in the interval

$$\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) \qquad (10.49)$$

is $1-\alpha$, implying that a significance level α test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ is to reject H_0 when

$$\mu_0 \notin \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) \tag{10.50}$$

(b) A Remark on Robustness

A test that performs well even when the underlying assumptions on which it is based are violated is said to be *robust*. For instance, the tests of were derived under the assumption that *the underlying population distribution is normal* with known variance σ^2. However, in deriving these tests, this assumption was used only to conclude that $\bar{X}$ also has a normal distribution. But, by the *central limit theorem*, it follows that for a reasonably large sample size, $\bar{X}$ will approximately have a normal distribution no matter what the underlying distribution. Thus we can conclude that these tests will be relatively *robust* for any population distribution with variance σ^2.

Table 10.1 summarizes the tests of this section.

H_0	H_1	Test Statistic TS	Significance level α	p-value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/\sigma$	Reject if $\lvert TS\rvert > z_{\alpha/2}$	$2P\left(Z \geq \lvert t\rvert\right)$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/\sigma$	Reject if $TS > z_\alpha$	$P\left(Z \geq t\right)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/\sigma$	Reject if $TS < -z_\alpha$	$P\left(Z \leq t\right)$

10.3.2 *Case of Unknown Variance: The t-Test*

Up to now we have supposed that the only unknown parameter of the normal population distribution is its mean. However, *the more common situation is one where the mean μ and variance σ^2 are both unknown*. Let us suppose this to be the case and again consider a test of the hypothesis that the mean is equal to some specified value μ_0. That is, consider a test of

$$H_0 : \mu = \mu_0 \tag{10.51}$$

versus the alternative $H_1 : \mu \neq \mu_0$

As before, it seems reasonable to reject H_0 when the sample mean $\bar{X}$ is far from μ_0. However, how far away it need be to justify rejection will depend on the variance σ^2. Recall that when the value of σ^2 was known, the test called for rejecting H_0 when $|\bar{X}-\mu_0|$ exceeded $z_{\alpha/2}\sigma/\sqrt{n}$ or, equivalently, when

$$\left|\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma}\right| > z_{\alpha/2} \qquad (10.52)$$

Now when σ^2 is no longer known, it seems reasonable to estimate it by

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1} \qquad (10.53)$$

and then to reject H_0 when

$$\left|\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S}\right| \qquad (10.54)$$

is large. To determine how large a value of this statistic to require for rejection, in order that the resulting test have significance level α, we must determine the *probability distribution of this statistic* when H_0 is true. However, as shown, the statistic T, defined by

$$T = \frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S} \qquad (10.55)$$

has, when $\mu=\mu_0$, a t-distribution with $n-1$ degrees of freedom. Hence,

$$P_{\mu_0}\left(-t_{\alpha/2,n-1} \le \frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S} \le t_{\alpha/2,n-1}\right) = 1-\alpha \qquad (10.56)$$

where $t_{\alpha/2,n-1}$ is the $100\alpha/2$ upper percentile value of the t-distribution with $n-1$ degrees of freedom. That is

$$P\left(T_{n-1} \ge t_{\alpha/2,n-1}\right) = P\left(T_{n-1} \le -t_{\alpha/2,n-1}\right) = \alpha/2 \qquad (10.57)$$

when T_{n-1} has a t-distribution with $n-1$ degrees of freedom. From Equation (10.56) we see that the appropriate significance level α test of

$$H_0: \mu=\mu_0 \qquad \text{versus} \qquad H_1: \mu \ne \mu_0 \qquad (10.58)$$

is, when σ^2 is unknown, to

$$\text{accept } H_0 \text{ if } \left|\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S}\right| \le t_{\alpha/2,n-1} \qquad (10.59)$$

$$\text{reject } H_0 \text{ if } \left|\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S}\right| > t_{\alpha/2,n-1} \qquad (10.60)$$

The test defined by Equation (10.59-60) is called a *two-sided t-test*. It is pictorially illustrated in Figure 10.4:

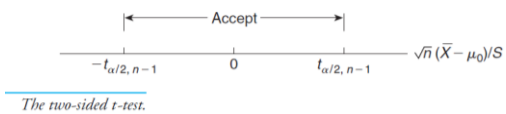

Figure 10.4

If we let t denote the observed value of the test statistic $T=\sqrt{n}\left(\bar{X}-\mu_0\right)/S$, then the p-value of the test is the probability that $|T|$ would exceed $|t|$ when H_0 is true. That is, the p-value is the probability that the absolute value of a t-random variable with $n-1$ degrees of freedom would exceed $|t|$. The test then calls for rejection at all significance levels higher than the p-value and acceptance at all lower significance levels.

Program R computes the value of the test statistic and the corresponding p-value. It can be applied both for one- and two-sided tests.

As usual in R, the functions pt, dt, qt, rt correspond to cdf, pdf, quantiles, and random sampling for a t distribution. Remember that you can type ?dt in RStudio to view the help file specifying the parameters of dt. For example, pt(1.65,3) computes the probability that x is less than or equal 1.65 given that x is sampled from the t distribution with 3 degrees of freedom, i.e. $P\left(x \le 1.65\right)$ given that $x \sim T_3$.

EXAMPLE 7. Among a clinic's patients having blood cholesterol levels ranging in the medium to high range (at least 220 milliliters per deciliter of serum), volunteers were recruited to test a new drug designed to reduce blood cholesterol. A group of 50 volunteers was given the drug for 1 month and the changes in their blood cholesterol levels were noted. If the average change was a reduction of 14.8 with a sample standard deviation of 6.4, what conclusions can be drawn?

Let us start by testing the hypothesis that the change could be due solely to chance—that is, that the 50 changes constitute a normal sample with mean 0. Because the value of the *t*-statistic used to test the hypothesis that a normal mean is equal to 0 is

$$T=\sqrt{n}\,\frac{\bar{X}-0}{S}=\sqrt{50}\,\frac{14.8}{6.4}=16.352$$

is clear that we should reject the hypothesis that the changes were solely due to chance. Unfortunately, however, we are not justified at this point in concluding that the changes were due to the specific drug used and not to some other possibility. For instance, it is well known that any medication received by a patient (whether or not this medication is directly relevant to the patient's suffering) often leads to an improvement in the patient's condition — the so-called placebo effect. Also, another possibility that may need to be taken into account would be the weather conditions during the month of testing, for it is certainly conceivable that this affects blood cholesterol level. Indeed, it must be concluded that the foregoing was a very poorly designed experiment, for in order to test whether a specific treatment has an effect on a disease that may be affected by many things, we should try to design the experiment so as to neutralize all other possible causes. The accepted approach for accomplishing this is to divide the volunteers at random into two groups—one group to receive the drug and the other to receive a placebo (that is, a tablet that looks and tastes like the actual drug but has no physiological effect). The volunteers should not be told whether they are in the actual or control group, and indeed it is best if even the clinicians do not have this information (the so-called *double-blind test*) so as not to allow their own biases to play a role. Since the two groups are chosen at random from among the volunteers, we can now hope that on average all factors affecting the two groups will be the same except that one received the actual drug and the other a placebo. Hence, any difference in performance between the groups can be attributed to the drug.

EXAMPLE 8. A public health official claims that the mean home water use is 350 gallons a day. To verify this claim, a study of 20 randomly selected homes was instigated with the result that the average daily water uses of these 20 homes were as follows: 340 344 362 375

356 386 354 364

332	402	340	355
362	322	372	324
318	360	338	370

Do the data contradict the official's claim?

To determine if the data contradict the official's claim, we need to test

$$H_0 : \mu = 350 \qquad \text{versus} \qquad H_1 : \mu \neq 350$$

This can be accomplished by running Program R or, by noting first that the sample mean and sample standard deviation of the preceding data set are

$$\bar{X} = 353.8 \; , \qquad S = 21.8478$$

Thus, the value of the test statistic is

$$T = \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{S} = \frac{\sqrt{20}\left(353.8 - 350\right)}{21.8478} = 0.7778$$

```
wa<-c(340,344,362,375,356,386,354,364,332,402,340,355,362,322,372,324,318,360
,338,370)
> mean(wa)
[1] 353.8
> sd(wa)
[1] 21.8478
> t.test(wa,alternative = c("two.sided"),mu=350,conf.level = 0.95)

        One Sample t-test
```

$\alpha = 0.05$

```
data:  wa
t = 0.77784, df = 19, p-value = 0.4462
alternative hypothesis: true mean is not equal to 350
95 percent confidence interval:
 343.5749 364.0251
sample estimates:
mean of x
    353.8
t.test(wa,alternative = c("two.sided"),mu=350,conf.level = 0.90)

        One Sample t-test
```

$\alpha = 0.1$

```
data:  wa
t = 0.77784, df = 19, p-value = 0.4462
alternative hypothesis: true mean is not equal to 350
90 percent confidence interval:
 345.3526 362.2474
sample estimates:
mean of x
    353.8
```

Because this is less than $t_{0.05,19} = 1.73$, the null hypothesis is accepted at the 10 percent level of significance. Indeed, the *p*-value of the test data is

$$p-\text{value} = P\left\{|T_{19}| > 0.7778\right\} = 2P\left\{T_{19} > 0.7778\right\} = 0.4462$$

indicating that the null hypothesis would be accepted at any reasonable significance level, and thus that the data are not inconsistent with the claim of the health official.

We can use a one-sided *t*-test to test the hypothesis

$$H_0 : \mu = \mu_0 \qquad (\text{or } H_0 : \mu \le \mu_0)$$

against the one-sided alternative

$$H_1 : \mu > \mu_0$$

The significance level α test is to

$$\text{accept } H_0 \text{ if } \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{S} \le t_{\alpha,n-1}$$

$$\text{reject } H_0 \text{ if } \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{S} > t_{\alpha,n-1}$$

If $\sqrt{n}\left(\bar{X} - \mu_0\right)/S = v$, then the *p*-value of the test is the probability that a *t*-random variable with $n-1$ degrees of freedom would be at least as large as *v*.

The significance level α test of

$$H_0 : \mu = \mu_0 \qquad (\text{or } H_0 : \mu \ge \mu_0)$$

against the one-sided alternative

$$H_1 : \mu < \mu_0$$

is

$$\text{accept } H_0 \text{ if } \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{S} \ge -t_{\alpha,n-1}$$

$$\text{reject } H_0 \text{ if } \frac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{S} < -t_{\alpha,n-1}$$

The *p*-value of this test is the probability that a *t*-random variable with $n-1$ degrees of freedom would be less than or equal to the observed value of $\sqrt{n}\left(\bar{X}-\mu_0\right)/S$.

EXAMPLE 9. The manufacturer of a new fiberglass tire claims that its average life will be at least 40,000 miles. To verify this claim a sample of 12 tires is tested, with their lifetimes (in 1,000s of miles) being as follows:

Tire	1	2	3	4	5	6	7	8	9	10	11	12
Life	36.1	40.2	33.8	38.5	42	35.8	37	41	36.8	37.2	33	36

Test the manufacturer's claim at the 5 percent level of significance.

To determine whether the foregoing data are consistent with the hypothesis that the mean life is at least 40,000 miles, we will test

$$H_0: \mu \geq 40000 \qquad \text{versus} \qquad H_1: \mu < 40000$$

A computation gives that

$$\bar{X} = 37.2833 \,, \quad S = 2.7319$$

and so the value of the test statistic is

$$T = \frac{\sqrt{12}\left(37.2833-40\right)}{2.7319} = -3.4448$$

Since this is less than $-t_{0.05,11} = -1.796$, the null hypothesis is rejected at the 5 percent level of significance. Indeed, the *p*-value of the test data is

$$p-\text{value} = P\left\{T_{11} < -3.4448\right\} = P\left\{T_{11} > 3.4448\right\} = 0.0028$$

indicating that the manufacturer's claim would be rejected at any significance level greater than 0.003.

Using R:

```
tire<-c(36.1,40.2,33.8,38.5,42,35.8,37,41,36.8,37.2,33,36)
> mean(tire)
[1] 37.28333
> sd(tire)
[1] 2.73191
```

```
> t.test(tire,alternative = c("less"),mu=40,conf.level = 0.95)

        One Sample t-test

data:  tire
t = -3.4448, df = 11, p-value = 0.002739
alternative hypothesis: true mean is less than 40
95 percent confidence interval:
     -Inf 38.69963
sample estimates:
mean of x
 37.28333
```

Table 10.2 summarizes the tests of this section.

$X_1, X_2, \ldots, X_n$ is a sample from a $N\left(\mu, \sigma^2\right)$ population σ^2 is unknown,

$\bar{X} = \sum_{i=1}^{n} X_i / n$, $S^2 = \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 / (n-1)$

H_0	H_1	Test Statistic TS	Significance level α	p-value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/S$	Reject if $\lvert TS \rvert > t_{\alpha/2, n-1}$	$2P\left\{T_{n-1} \geq \lvert t \rvert\right\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/S$	Reject if $TS > t_{\alpha, n-1}$	$P\left\{T_{n-1} \geq t\right\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}\left(\bar{X} - \mu_0\right)/S$	Reject if $TS < -t_{\alpha, n-1}$	$P\left\{T_{n-1} \leq t\right\}$

T_{n-1} is a *t*-random variable with $n-1$ degrees of freedom , $P\left\{T_{n-1} > t_{\alpha, n-1}\right\} = \alpha$