8 Tests of Hypotheses Based on a Single Sample

8.2 Tests About a Population Mean

Tests About a Population Mean

Confidence intervals for a population mean μ focused on three different cases.

We now develop test procedures for these cases.

3

Case I: A Normal Population with Known σ

4

Case I: A Normal Population with Known σ

Although the assumption that the value of σ is known is rarely met in practice, this case provides a good starting point because of the ease with which general procedures and their properties can be developed.

The null hypothesis in all three cases will state that μ has a particular numerical value, the *null value*, which we will denote by μ_0 . Let $X_1, \ldots, X_n$ represent a random sample of size n from the normal population.

Case I: A Normal Population with Known σ

Then the sample mean $\overline{X}$ has a normal distribution with expected value $\mu_{\overline{X}} = \mu$ and standard deviation $\sigma_{\overline{X}} = \sigma/\sqrt{n}$.

When H_0 is true, $\mu_{\overline{X}} = \mu_0$. Consider now the statistic Z obtained by standardizing $\overline{X}$ under the assumption that H_0 is true:

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

Case I: A Normal Population with Known σ

Substitution of the computed sample mean $\overline{x}$ gives z, the distance between $\overline{x}$ and μ_0 expressed in "standard deviation units."

For example, if the null hypothesis is H_0: $\mu = 100$, $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 10/\sqrt{25} = 2.0$, and $\overline{x} = 103$, then the test statistic value is $z = (103 – 100)/2.0 = 1.5$.

That is, the observed value of $\overline{x}$ is 1.5 standard Deviations (of $\overline{X}$) larger than what we expect it to be when H_0 is true.

Case I: A Normal Population with Known σ

The statistic Z is a natural measure of the distance between $\overline{X}$, the estimator of μ, and its expected value when H_0 is true. If this distance is too great in a direction consistent with H_a, the null hypothesis should be rejected.

Suppose first that the alternative hypothesis has the form H_a : $\mu > \mu_0$. Then an $\overline{x}$ value less than μ_0 certainly does not provide support for H_a.

Such an $\overline{x}$ corresponds to a negative value of z (since $\overline{x} – \mu_0$ is negative and the divisor $\sigma/\sqrt{n}$ is positive).

Case I: A Normal Population with Known σ

Similarly, an $\bar{x}$ value that exceeds μ_0 by only a small amount (corresponding to z, which is positive but small) does not suggest that H_0 should be rejected in favor of H_a.

The rejection of H_0 is appropriate only when $\bar{x}$ considerably exceeds μ_0—that is, when the z value is positive and large. In summary, the appropriate rejection region, based on the test statistic Z rather than $\bar{X}$, has the form $z \geq c$.

9

Case I: A Normal Population with Known σ

As we have discussed earlier, the cutoff value c should be chosen to control the probability of a type I error at the desired level α.

This is easily accomplished because the distribution of the test statistic Z when H_0 is true is the standard normal distribution (that's why μ_0 was subtracted in standardizing).

The required cutoff c is the z critical value that captures upper-tail area α under the z curve.

10

Case I: A Normal Population with Known σ

As an example, let c = 1.645, the value that captures tail area .05($z_{.05}$ = 1.645). Then,

$\alpha = P(\text{type I error}) = P(H_0 \text{ is rejected when } H_0 \text{ is true})$

$= P(Z \geq 1.645 \text{ when } Z \sim N(0,1)) = 1 - \Phi(1.645) = .05$

More generally, the rejection region $z \geq z_\alpha$ has type I error probability α.

The test procedure is *upper-tailed* because the rejection region consists only of large values of the test statistic.

11

Case I: A Normal Population with Known σ

Analogous reasoning for the alternative hypothesis H_a: $\mu < \mu_0$ suggests a rejection region of the form $z \leq c$, where c is a suitably chosen negative number ($\bar{x}$ is far below μ_0 if and only if z is quite negative).

Because Z has a standard normal distribution when H_0 is true, taking $c = -z_\alpha$ yields P(type I error) = α.

This is a *lower-tailed* test. For example, $z_{.10}$ = 1.28 implies that the rejection region $z \leq -1.28$ specifies a test with significance level .10.

12

Case I: A Normal Population with Known σ

Finally, when the alternative hypothesis is H_a: $\mu \neq \mu_0$, H_0 should be rejected if $\bar{x}$ is too far to either side of μ_0. This is equivalent to rejecting H_0 either if $z \geq c$ or if $z \leq -c$. Suppose we desire $\alpha = .05$. Then,

$.05 = P(Z \geq c$ or $Z \leq -c$ when Z has a standard normal distribution)

$= \Phi(-c) + 1 - \Phi(c) = 2[1 - \Phi(c)]$

13

Case I: A Normal Population with Known σ

Thus c is such that $1 - \Phi(c)$, the area under the z curve to the right of c, is .025 (and not .05!).

From Appendix Table A.3, $c = 1.96$, and the rejection region is $z \geq 1.96$ *or* $z \leq -1.96$.

For any α, the *two-tailed* rejection region $z \geq z_{\alpha/2}$ *or* $z \leq -z_{\alpha/2}$ has type I error probability α (since area $\alpha/2$ is captured under each of the two tails of the z curve).

14

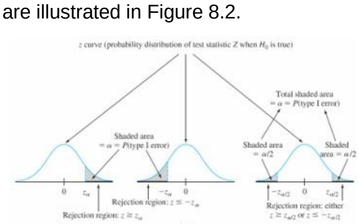
Case I: A Normal Population with Known σ

Again, the key reason for using the standardized test statistic Z is that because Z has a known distribution when H_0 is true (standard normal), a rejection region with desired type I error probability is easily obtained by using an appropriate critical value.

15

Case I: A Normal Population with Known σ

The test procedure for case I is summarized in the accompanying box, and the corresponding rejection regions are illustrated in Figure 8.2.



Rejection regions for z tests: (a) upper-tailed test; (b) lower-tailed test; (c) two-tailed test

Figure 8.2

16

Case I: A Normal Population with Known σ

Null hypothesis: $H_0: \mu = \mu_0$

Test statistic value : $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Alternative Hypothesis	**Rejection Region for Level α Test**
$H_a: \mu > \mu_0$	$z \geq z_\alpha$ (upper-tailed test)
$H_a: \mu < \mu_0$	$z \leq -z_\alpha$ (lower-tailed test)
$H_a: \mu \neq \mu_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$ (two-tailed test)

17

Case I: A Normal Population with Known σ

Use of the following sequence of steps is recommended when testing hypotheses about a parameter.

1. Identify the parameter of interest and describe it in the context of the problem situation.

2. Determine the null value and state the null hypothesis.

3. State the appropriate alternative hypothesis.

18

Case I: A Normal Population with Known σ

4. Give the formula for the computed value of the test statistic (substituting the null value and the known values of any other parameters, but *not* those of any samplebased quantities).

5. State the rejection region for the selected significance level α.

6. Compute any necessary sample quantities, substitute into the formula for the test statistic value, and compute that value.

19

Case I: A Normal Population with Known σ

7. Decide whether H_0 should be rejected, and state this conclusion in the problem context.

The formulation of hypotheses (Steps 2 and 3) should be done before examining the data.

20

Example 6

A manufacturer of sprinkler systems used for fire protection in office buildings claims that the true average system-activation temperature is 130°.

A sample of $n = 9$ systems, when tested, yields a sample average activation temperature of 131.08°F.

If the distribution of activation times is normal with standard deviation 1.5°F, does the data contradict the manufacturer's claim at significance level $\alpha = .01$?

21

Example 6

cont'd

1. Parameter of interest: μ = true average activation temperature.

2. Null hypothesis: H_0: $\mu = 130$ (null value = $\mu_0 = 130$).

3. Alternative hypothesis: H_a: $\mu \neq 130$ (a departure from the claimed value in *either* direction is of concern).

4. Test statistic value:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 130}{1.5/\sqrt{n}}$$

22

Example 6

cont'd

5. Rejection region: The form of H_a implies use of a two-tailed test with rejection region *either* $z \geq z_{.005}$ *or* $z \leq -z_{.005}$. From Appendix Table A.3, $z_{.005} = 2.58$, so we reject H_0 if either $z \geq 2.58$ or $z \leq -2.58$.

6. Substituting $n = 9$ and $\bar{x} = 131.08$,

$$z = \frac{131.08 - 130}{1.5/\sqrt{9}} = \frac{1.08}{.5} = 2.16$$

That is, the observed sample mean is a bit more than 2 standard deviations above what would have been expected were H_0 true.

23

Example 6

cont'd

7. The computed value $z = 2.16$ does not fall in the rejection region ($-2.58 < 2.16 < 2.58$), so H_0 cannot be rejected at significance level .01. The data does not give strong support to the claim that the true average differs from the design value of 130.

24

Case I: A Normal Population with Known σ

β and Sample Size Determination The z tests for case I are among the few in statistics for which there are simple formulas available for β, the probability of a type II error.

Consider first the upper-tailed test with rejection region $z \geq z_\alpha$.

This is equivalent to $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$, so H_0 will not be rejected if $\bar{x} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$.

25

Case I: A Normal Population with Known σ

Now let μ' denote a particular value of μ that exceeds the null value μ_0. Then,

$$\begin{aligned}
\beta(\mu') &= P(H_0 \text{ is not rejected when } \mu = \mu') \\
&= P(\bar{X} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n} \text{ when } \mu = \mu') \\
&= P\left(\frac{\bar{X} - \mu'}{\sigma/\sqrt{n}} < z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}} \text{ when } \mu = \mu'\right) \\
&= \Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

26

Case I: A Normal Population with Known σ

As μ' increases, $\mu_0 - \mu'$ becomes more negative, so $\beta(\mu')$ will be small when μ' greatly exceeds μ_0 (because the value at which Φ is evaluated will then be quite negative).

Error probabilities for the lower-tailed and two-tailed tests are derived in an analogous manner.

If σ is large, the probability of a type II error can be large at an alternative value μ' that is of particular concern to an investigator.

27

Case I: A Normal Population with Known σ

Suppose we fix α and also specify β for such an alternative value. In the sprinkler example, company officials might view $\mu' = 132$ as a very substantial departure from H_0: $\mu = 130$ and therefore wish $\beta(132) = .10$ in addition to $\alpha = .01$.

More generally, consider the two restrictions $P(\text{type I error}) = \alpha$ and $\beta(\mu') = \beta$ for specified α, μ' and β.

28

Case I: A Normal Population with Known σ

Then for an upper-tailed test, the sample size n should be chosen to satisfy

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

This implies that

$$-z_\beta = \begin{matrix} z \text{ critical value that} \\ \text{captures lower-tail area } \beta \end{matrix} = z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}$$

29

Case I: A Normal Population with Known σ

It is easy to solve this equation for the desired n. A parallel argument yields the necessary sample size for lower- and two-tailed tests as summarized in the next box.

Alternative Hypothesis	**Type II Error Probability for a Level a Test**
H_a: $\mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

30

Case I: A Normal Population with Known σ

where $\Phi(z)$ = the standard normal cdf.

The sample size n for which a level α test also has $\beta(\mu') = \beta$ at the alternative value μ' is

$$n = \begin{cases} \left[\dfrac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{for a one-tailed (upper or lower) test} \\ \left[\dfrac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{for a two-tailed test (an approximate solution)} \end{cases}$$

31

Example 7

Let μ denote the true average tread life of a certain type of tire.

Consider testing H_0: $\mu = 30{,}000$ versus H_a: $\mu > 30{,}000$ based on a sample of size $n = 16$ from a normal population distribution with $\sigma = 1500$.

A test with $\alpha = .01$ requires $z_\alpha = z_{.01} = 2.33$.

32

Example 7

cont'd

The probability of making a type II error when μ = 31,000 is

$$\beta(31{,}000) = \Phi\left(2.33 + \frac{30{,}000 - 31{,}000}{1500/\sqrt{16}}\right)$$

$$= \Phi(-.34)$$

$$= .3669$$

33

Example 7

cont'd

Since $z_{.1}$ = 1.28, the requirement that the level .01 test also have β(31,000) = .1 necessitates

$$n = \left[\frac{1500(2.33 + 1.28)}{30{,}000 - 31{,}000}\right]^2$$

$$= (-5.42)^2$$

$$= 29.32$$

The sample size must be an integer, so n = 30 tires should be used.

34

Case II: Large-Sample Tests

35

Case II: Large-Sample Tests

When the sample size is large, the z tests for case I are easily modified to yield valid test procedures without requiring either a normal population distribution or known σ.

Earlier we used the key result to justify large-sample confidence intervals:

A large n implies that the standardized variable

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has *approximately* a standard normal distribution.

36

Case II: Large-Sample Tests

Substitution of the null value μ_0 in place of μ yields the test statistic

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

which has approximately a standard normal distribution when H_0 is true.

37

Case II: Large-Sample Tests

The use of rejection regions given previously for case I (e.g., $z \geq z_\alpha$ when the alternative hypothesis is H_a: $\mu > \mu_0$) then results in test procedures for which the significance level is approximately (rather than exactly) α.

The rule of thumb $n > 40$ will again be used to characterize a large sample size.

38

Example 8

A dynamic cone penetrometer (DCP) is used for measuring material resistance to penetration (mm/blow) as a cone is driven into pavement or subgrade.

Suppose that for a particular application it is required that the true average DCP value for a certain type of pavement be less than 30.

The pavement will not be used unless there is conclusive evidence that the specification has been met.

39

Example 8

cont'd

Let's state and test the appropriate hypotheses using the following data ("Probabilistic Model for the Analysis of Dynamic Cone Penetrometer Test Values in Pavement Structure Evaluation," *J. of Testing and Evaluation*, 1999: 7–14):

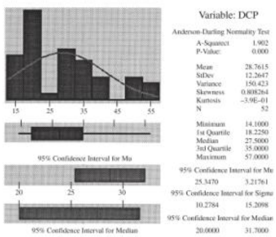
14.1	14.5	15.5	16.0	16.0	16.7	16.9	17.1	17.5	17.8
17.8	18.1	18.2	18.3	18.3	19.0	19.2	19.4	20.0	20.0
20.8	20.8	21.0	21.5	23.5	27.5	27.5	28.0	28.3	30.0
30.0	31.6	31.7	31.7	32.5	33.5	33.9	35.0	35.0	35.0
36.7	40.0	40.0	41.3	41.7	47.5	50.0	51.0	51.8	54.4
55.0	57.0								

40

Example 8

cont'd

Figure 8.3 shows a descriptive summary obtained from Minitab.



Descriptive Statistics

Minitab descriptive summary for the DCP data of Example 8

Figure 8.3

Example 8

cont'd

The sample mean DCP is less than 30. However, there is a substantial amount of variation in the data (sample coefficient of variation = $s/\bar{x} = .4265$).

The fact that the mean is less than the design specification cutoff may be a consequence just of sampling variability.

Notice that the histogram does not resemble at all a normal curve (and a normal probability plot does not exhibit a linear pattern), but the large-sample z tests do not require a normal population distribution.

Example 8

cont'd

1. μ = true average DCP value

2. H_0: $\mu = 30$

3. H_a: $\mu < 30$(so the pavement will not be used unless the null hypothesis is rejected)

4. $z = \dfrac{\bar{x} - 30}{s/\sqrt{n}}$

Example 8

cont'd

5. A test with significance level .05 rejects H_0 when $z \leq -1.645$ (a lower-tailed test).

6. With $n = 52$, $\bar{x} = 28.76$, and $s = 12.2647$,

$$z = \frac{28.76 - 30}{12.2647/\sqrt{52}} = \frac{-1.24}{1.701} = -.73$$

7. Since $-.73 > -1.645$, H_0 cannot be rejected. We do not have compelling evidence for concluding that $\mu < 30$; use of the pavement is not justified.

Case II: Large-Sample Tests

Determination of β and the necessary sample size for these large-sample tests can be based either on specifying a plausible value of σ and using the case I formulas (even though s is used in the test) or on using the methodology to be introduced shortly in connection with case III.

45

Case III: A Normal Population Distribution

46

Case III: A Normal Population Distribution

When n is small, the Central Limit Theorem (CLT) can no longer be invoked to justify the use of a large-sample test.

Our approach here will be the same one used there: We will assume that the population distribution is at least approximately normal and describe test procedures whose validity rests on this assumption.

47

Case III: A Normal Population Distribution

If an investigator has good reason to believe that the population distribution is quite nonnormal, a distribution-free test can be used.

Alternatively, a statistician can be consulted regarding procedures valid for specific families of population distributions other than the normal family. Or a bootstrap procedure can be developed.

48

Case III: A Normal Population Distribution

The key result on which tests for a normal population mean are based was used to derive the one-sample t CI:
If $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution, the standardized variable

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t distribution with $n - 1$ degrees of freedom (df).

49

Case III: A Normal Population Distribution

Consider testing against H_0: $\mu = \mu_0$ against H_a: $\mu > \mu_0$ by using the test statistic $T = (\overline{X} - \mu_0)/(S/\sqrt{n})$.

That is, the test statistic results from standardizing $\overline{X}$ under the assumption that H_0 is true (using $S/\sqrt{n}$, the estimated standard deviation of $\overline{X}$, rather than $\sigma/\sqrt{n}$).

When H_0 is true, the test statistic has a t distribution with $n - 1$ df.

50

Case III: A Normal Population Distribution

Knowledge of the test statistic's distribution when H_0 is true (the "null distribution") allows us to construct a rejection region for which the type I error probability is controlled at the desired level.

In particular, use of the upper-tail t critical value $t_{\alpha,n-1}$ to specify the rejection region $t \geq t_{\alpha,n-1}$ implies that

P(type I error) = P(H_0 is rejected when it is true)

= $P(T \geq t_{\alpha,n-1}$ when T has a t distribution with $n - 1$ df)

= α

51

Case III: A Normal Population Distribution

The test statistic is really the same here as in the large-sample case but is labeled T to emphasize that its null distribution is a t distribution with $n - 1$ df rather than the standard normal (z) distribution.

The rejection region for the t test differs from that for the z test only in that a t critical value $t_{\alpha,n-1}$ replaces the z critical value z_α .

Similar comments apply to alternatives for which a lower-tailed or two-tailed test is appropriate.

52

Case III: A Normal Population Distribution

The One-Sample t Test

Null hypothesis: H_0: $\mu = \mu_0$

Test statistic value: $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Alternative Hypothesis	**Rejection Region for a Level α Test**
H_a: $\mu > \mu_0$	$t \geq t_{\alpha,n-1}$ (upper-tailed)
H_a: $\mu < \mu_0$	$t \leq -t_{\alpha,n-1}$ (lower-tailed)
H_a: $\mu \neq \mu_0$	either $t \geq t_{\alpha/2,n-1}$ or $t \leq -t_{\alpha/2,n-1}$ (two-tailed)

53

Example 9

Glycerol is a major by-product of ethanol fermentation in wine production and contributes to the sweetness, body, and fullness of wines.

The article "A Rapid and Simple Method for Simultaneous Determination of Glycerol, Fructose, and Glucose in Wine" (*American J. of Enology and Viticulture,* 2007: 279–283) includes the following observations on glycerol concentration (mg/mL) for samples of standard-quality (uncertified) white wines: 2.67, 4.62, 4.14, 3.81, 3.83. Suppose the desired concentration value is 4.

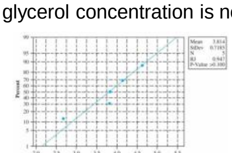
Does the sample data suggest that true average concentration is something other than the desired value?

54

Example 9

cont'd

The accompanying normal probability plot from Minitab provides strong support for assuming that the population distribution of glycerol concentration is normal.



Normal probability plot for the data of Example 9

Figure 8.4

Let's carry out a test of appropriate hypotheses using the one-sample *t* test with a significance level of .05.

55

Example 9

cont'd

1. μ = true average glycerol concentration

2. H_0: $\mu = 4$

3. H_a: $\mu \neq 4$

4. $t = \dfrac{\bar{x} - 4}{s/\sqrt{n}}$

5. The inequality in H_a implies that a two-tailed test is appropriate, which requires $t = (3.814 - 4)/.321 = -.58$. Thus H_0 will be rejected if either $t \geq 2.776$ or $t \leq -2.776$.

56

Example 9

cont'd

6. $\Sigma x_i = 19.07$, and $\Sigma x_i^2 = 74.7979$, from which $\bar{x}$ = 3.814 s = .718, and the estimated standard error of the mean is $s/\sqrt{n} = .321$. The test statistic value is then t = (3.814 – 4)/.321 = –.58.

7. Clearly t = –.58 does not lie in the rejection region for a significance level of .05.

It is still plausible that μ = 4. The deviation of the sample mean 3.814 from its expected value 4 when H_0 is true can be attributed just to sampling variability rather than to H_0 being false.

Example 9

cont'd

The accompanying Minitab output from a request to perform a two-tailed one sample t test shows identical calculated values to those just obtained.

```
Test of mu = 4 vs not = 4
Variable   N   Mean  StDev  SE Mean  95% CI              T      P
glyc conc  5  3.814  0.718    0.321  (2.922, 4.706)  -0.58  0.594
```

The fact that the last number on output, the "*P*-value," exceeds .05 (and any other reasonable significance level) implies that the null hypothesis can't be rejected.

Case III: A Normal Population Distribution

β and Sample Size Determination

The calculation of β at the alternative value μ' in case I was carried out by expressing the rejection region in terms of $\bar{x}$ (e.g., $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$) and then subtracting μ' to standardize correctly.

An equivalent approach involves noting that when $\mu = \mu'$ the test statistic $z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ still has a normal distribution with variance 1, but now the mean value of Z is given by $(\mu' - \mu_0)(\sigma/\sqrt{n})$. That is, when $\mu = \mu'$, the test statistic still has a normal distribution though not the standard normal distribution.

Case III: A Normal Population Distribution

Because of this, $\beta(\mu')$ is an area under the normal curve corresponding to mean value $(\mu' - \mu_0)/\sigma/\sqrt{n}$ and variance 1. Both α and β involve working with normally distributed variables.

The calculation of $\beta(\mu')$ for the t test is much less straightforward. This is because the distribution of the test statistic $T = (\bar{X} - \mu_0)/(S/\sqrt{n})$ is quite complicated when H_0 is false and H_a is true. Thus, for an upper-tailed test, determining

$$\beta(\mu') = P(T < t_{\alpha,n-1} \text{ when } \mu = \mu' \text{ rather than } \mu_0)$$

involves integrating a very unpleasant density function. This must be done numerically.

Case III: A Normal Population Distribution

The results are summarized in graphs of β that appear in Appendix Table A.17.

There are four sets of graphs, corresponding to one-tailed tests at level .05 and level .01 and two-tailed tests at the same levels.

To understand how these graphs are used, note first that both β and the necessary sample size n in case I are functions not just of the absolute difference $|\mu_0 - \mu'|$ but of $d = |\mu_0 - \mu'|/\sigma$.

61

Case III: A Normal Population Distribution

Suppose, for example, that $|\mu_0 - \mu'| = 10$.

This departure from H_0 will be much easier to detect (smaller β) when $\sigma = 2$, in which case μ_0 and μ' are 5 population standard deviations apart, than when $\sigma = 10$.

The fact that β for the t test depends on d rather than just $|\mu_0 - \mu'|$ is unfortunate, since to use the graphs one must have some idea of the true value of σ.

62

Case III: A Normal Population Distribution

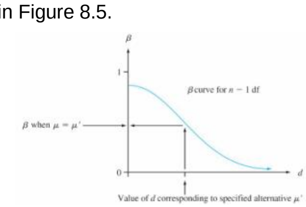
A conservative (large) guess for σ will yield a conservative (large) value of $\beta(\mu')$ and a conservative estimate of the sample size necessary for prescribed α and $\beta(\mu')$.

Once the alternative μ' and value of σ are selected, d is calculated and its value located on the horizontal axis of the relevant set of curves.

63

Case III: A Normal Population Distribution

The value of β is the height of the $n - 1$ curve above the value of d (visual interpolation is necessary if $n - 1$ is not a value for which the corresponding curve appears), as illustrated in Figure 8.5.



A typical β curve for the t test

Figure 8.5

64

Case III: A Normal Population Distribution

Rather than fixing n (i.e., $n - 1$ and thus the particular curve from which β is read), one might prescribe both α (.05 or .01 here) and a value of β for the chosen μ' and σ.

After computing d, the point (d, β) is located on the relevant set of graphs.

The curve below and closest to this point gives $n - 1$ and thus n (again, interpolation is often necessary).

Example 10

The true average voltage drop from collector to emitter of insulated gate bipolar transistors of a certain type is supposed to be at most 2.5 volts.

An investigator selects a sample of n = 10 such transistors and uses the resulting voltages as a basis for testing H_0: μ = 2.5 versus H_a: μ > 2.5 using a t test with significance level α = .05.

If the standard deviation of the voltage distribution is σ = .100, how likely is it that H_0 will not be rejected when in fact μ = 2.6?

Example 10

cont'd

With $d = |2.5 - 2.6|/ .100 = 1.0$, the point on the β curve at 9 df for a one-tailed test with α = .05 above 1.0 has a height of approximately .1, so $\beta \approx .1$.

The investigator might think that this is too large a value of β for such a substantial departure from H_0 and may wish to have β = .05 for this alternative value of μ.

Since d = 1.0, the point (d, β) = (1.0,.05) must be located. This point is very close to the 14 df curve, so using n = 15 will give both α = .05 and β = .05 when the value of μ is 2.6 and σ = .10.

Example 10

cont'd

A larger value of σ would give a larger β for this alternative, and an alternative value of σ closer to 2.5 would also result in an increased value of β.

Case III: A Normal Population Distribution

Most of the widely used statistical software packages are capable of calculating type II error probabilities.

They generally work in terms of **power,** which is simply $1 - \beta$. A small value of β (close to 0) is equivalent to large power (near 1).

A *powerful* test is one that has high power and therefore good ability to detect when the null hypothesis is false.

69

Case III: A Normal Population Distribution

As an example, we asked Minitab to determine the power of the upper-tailed test in Example 8.10 for the three sample sizes 5, 10, and 15 when $\alpha = .05$, $\sigma = .10$ and the value of μ is actually 2.6 rather than the null value 2.5—a "difference" of.

2.6 – 2.5 = .1. We also asked the software to determine the necessary sample size for a power of .9 ($\beta = .1$) and also .95.

70

Case III: A Normal Population Distribution

Here is the resulting output:

```
Power and Sample Size
Testing mean = null (versus > null)
Calculating power for mean = null + difference
Alpha = 0.05 Assumed standard deviation = 0.1
            Sample
Difference    Size    Power
0.1            5      0.579737
0.1           10      0.897517
0.1           15      0.978916
               Sample Target
                                Actual
Difference    Size    Power     Power
0.1           11      0.90      0.924489
0.1           13      0.95      0.959703
```

71

Case III: A Normal Population Distribution

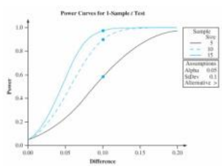
The power for the sample size $n = 10$ is a bit smaller than .9. So if we insist that the power be at least .9, a sample size of 11 is required and the actual power for that n is roughly .92.

The software says that for a target power of .95, a sample size of $n = 13$ is required, whereas eyeballing our β curves gave 15. When available, this type of software is more reliable than the curves.

72

Case III: A Normal Population Distribution

Finally, Minitab now also provides power curves for the specified sample sizes, as shown in Figure 8.6. Such curves show how the power increases for each sample size as the actual value of μ moves further and further away from the null value.



Power curves from Minitab for the *t* test of Example 10

Figure 8.6