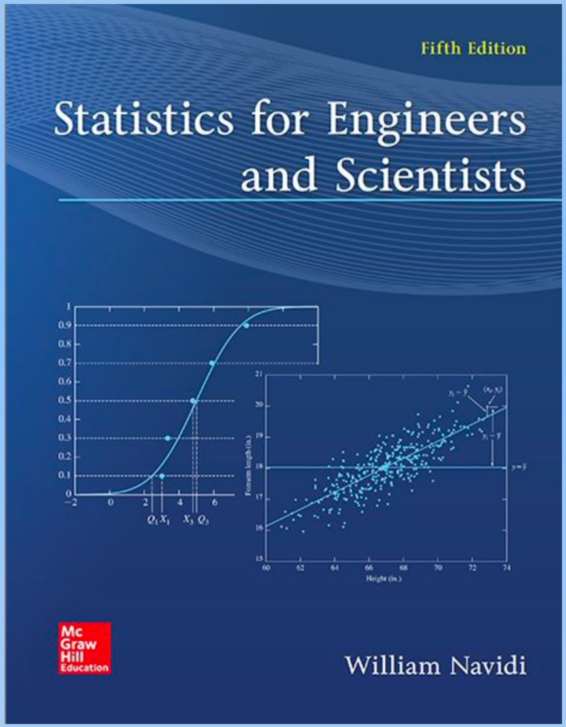


CHAPTER 6

Hypothesis Testing (part 1)

Ch. 6 - Hypothesis Testing

6-1 Large-Sample Tests for a Population Mean

6-2 Drawing Conclusions from the Results of Hypothesis Tests

6-3 Tests for a Population Proportion

6-4 Small-Sample Tests for a Population Mean

6-5 Large-Sample Tests for the Difference Between Two Means

6-7 Small-Sample Tests for the Difference Between Two Means

6-13 Power

Hypothesis Testing

Hypothesis testing is a decision-making process for evaluating claims about a population.

In hypothesis testing, one must

- ✓ define the population under study
- ✓ state the hypotheses that will be investigated
- ✓ give the significance level
- ✓ select a sample from the population
- ✓ collect the data
- ✓ perform the calculations required for the statistical test
- ✓ reach a conclusion.

Hypothesis Testing…

Three methods are used to test hypotheses:

1. **The *P*-value method**
2. The traditional method *(not studied in this course)*
3. The confidence interval method *(not studied in this course)*

P-value Method

Step 1	State the hypotheses and identify the claim.
Step 2	Compute the test value.
Step 3	Find the *P*-value.
Step 4	Make the decision.
Step 5	Summarize the results.

P-value Method - textbook (p.406)

Steps in Performing a Hypothesis Test

1. Define H_0 and H_1.
2. Assume H_0 to be true.
3. Compute a **test statistic**. A test statistic is a statistic that is used to assess the strength of the evidence against H_0.
4. Compute the *P*-value of the test statistic. The *P*-value is the probability, assuming H_0 to be true, that the test statistic would have a value whose disagreement with H_0 is as great as or greater than that actually observed. The *P*-value is also called the **observed significance level**.
5. State a conclusion about the strength of the evidence against H_0.

Steps in Hypothesis Testing

- A **statistical hypothesis** is a statement about a population parameter developed for the purpose of testing. This statement **may or may not be true**.
- **Hypothesis testing** is a procedure, based on sample evidence and probability theory, used to determine whether the hypothesis is a reasonable statement and should not be rejected, or is unreasonable and should be rejected.

Steps in Hypothesis Testing

- The **null hypothesis**, symbolized by H_0, is a statistical hypothesis that states that there is no difference between a parameter and a specific value, or that there is no difference between two parameters.
- The **alternative hypothesis**, symbolized by H_1, is a statistical hypothesis that states the existence of a difference between a parameter and a specific value, or states that there is a difference between two parameters.

Situation A…

A medical researcher is interested in finding out whether a new medication will have any undesirable side effects. The researcher is particularly concerned with the pulse rate of the patients who take the medication. Will the pulse rate **increase**, **decrease**, or **remain unchanged** after a patient takes the medication? The researcher knows that the mean pulse rate for the population under study is 82 beats per minute.

The hypotheses for this situation are

$$H_0:\ \mu = 82 \qquad H_1:\ \mu \neq 82$$

This is called a **two-tailed** hypothesis test.

Situation B…

A chemist invents an additive to **increase** the life of an automobile battery. The mean lifetime of the automobile battery without the additive is 36 months.

The hypotheses for this situation are

$$H_0: \mu = 36 \qquad H_1: \mu > 36$$

This is called a **right-tailed** hypothesis test.

Situation C…

A contractor wishes to **lower** heating bills by using a special type of insulation in houses. If the average of the monthly heating bills is $78, her hypotheses about heating costs with the use of insulation are

$$H_0: \mu = 78 \qquad H_1: \mu < 78$$

This is called a **left-tailed** hypothesis test.

Claim

Two-tailed test	Right-tailed test	Left-tailed test
$H_0: \mu = k$	$H_0: \mu = k$	$H_0: \mu = k$
$H_1: \mu \neq k$	$H_1: \mu > k$	$H_1: \mu < k$

Claim

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\mu = k$	H_0: $\mu = k$	H_0: $\mu = k$
H_1: $\mu \neq k$	H_1: $\mu > k$	H_1: $\mu < k$

When a researcher conducts a study, he or she is generally looking for evidence to support a **claim**. Therefore, the claim should be stated as the alternative hypothesis, or **research hypothesis**.

A claim, though, can be stated as either the null hypothesis or the alternative hypothesis; however, the statistical evidence can only *support* the claim if it is the alternative hypothesis. Statistical evidence can be used to ***reject*** the claim if the claim is the null hypothesis.

These facts are important when you are stating the conclusion of a statistical study.

Example

State the null and alternative hypotheses for each conjuncture.

a. A researcher studies gambling in young people. She thinks those who gamble spend more than \$30 per day.

b. A researcher wishes to see if police officers whose spouses work in law enforcement have a lower score on a work stress questionnaire than the average score of 120.

c. A teacher feels that if an online textbook is used for a course instead of a hardback book, it may change the students' scores on a final exam. In the past, the average final exam score for the students was 83.

SOLUTION

a. H_0: $\mu = \$30$ and H_1: $\mu > \$30$

b. H_0: $\mu = 120$ and H_1: $\mu < 120$

c. H_0: $\mu = 83$ and H_1: $\mu \neq 83$

Hypothesis Testing…

- After stating the hypothesis, the researcher's next step is to design the study. The researcher selects the **correct statistical test**, chooses an appropriate **level of significance**, and formulates a plan for conducting the study.

Hypothesis Testing…

- A **statistical test** uses the data obtained from a sample to make a decision about whether the null hypothesis should be rejected.
- The numerical value obtained from a statistical test is called the **test value**.
- In the hypothesis-testing situation, there are **four** possible outcomes.

Hypothesis Testing…

	H_0 true	H_0 false
Reject H_0	**Error** Type I	Correct decision
Do not reject H_0	Correct decision	**Error** Type II

Hypothesis Testing…

- In reality, the null hypothesis may or may not be true, and a decision is made to reject or not to reject it on the basis of the data obtained from a sample.
- A **type I error** occurs if one rejects the null hypothesis when it is true.
- A **type II error** occurs if one does not reject the null hypothesis when it is false.

Hypothesis Testing Example – Jury Trial

H_0: The defendant is innocent.
H_1: The defendant is not innocent.

The results of a trial can be shown as follows:

	H_0 true (innocent)	H_0 false (not innocent)
Reject H_0 (convict)	1. Type I error	2. Correct decision
Do not reject H_0 (acquit)	3. Correct decision	4. Type II error

Hypothesis Testing…

- The **level of significance** is the maximum probability of committing a type I error. This probability is symbolized by α (alpha). That is,

  P(type I error) = α

Likewise,

P(type II error) = β (beta)

Hypothesis Testing

- Typical significance levels are:

  0.10, 0.05, and 0.01

- For example, when α **= 0.10**, there is a 10% chance of rejecting a true null hypothesis.

Summarizing Results

- It is important to summarize the results of a statistical study correctly.
- The following table will help you when you summarize the results.

	Claim	
Decision	**Claim is H_0**	**Claim is H_1**
Reject H_0	There is enough evidence to reject the claim.	There is enough evidence to support the claim.
Do not reject H_0	There is not enough evidence to reject the claim.	There is not enough evidence to support the claim.

Large-Sample Tests for a Population Mean

The ***z* test** is a statistical test for the mean of a population. It can be used when ***n* ≥ 30**, or when the population is normally distributed and **σ is known**.

The formula for the *z test is*

$$\boxed{z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}}$$

where

$\bar{X}$ = sample mean

μ = hypothesized population mean

σ = population standard deviation

n = sample size

P-value Method

- Typical significance levels are:

  0.10, 0.05, and 0.01

- The ***P-value*** (or probability value) is the probability of getting a sample statistic (such as the sample mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.

P-value Method…

Step 1	State the hypotheses and identify the claim.
Step 2	Compute the test value.
Step 3	Find the *P*-value.
Step 4	Make the decision.
Step 5	Summarize the results.

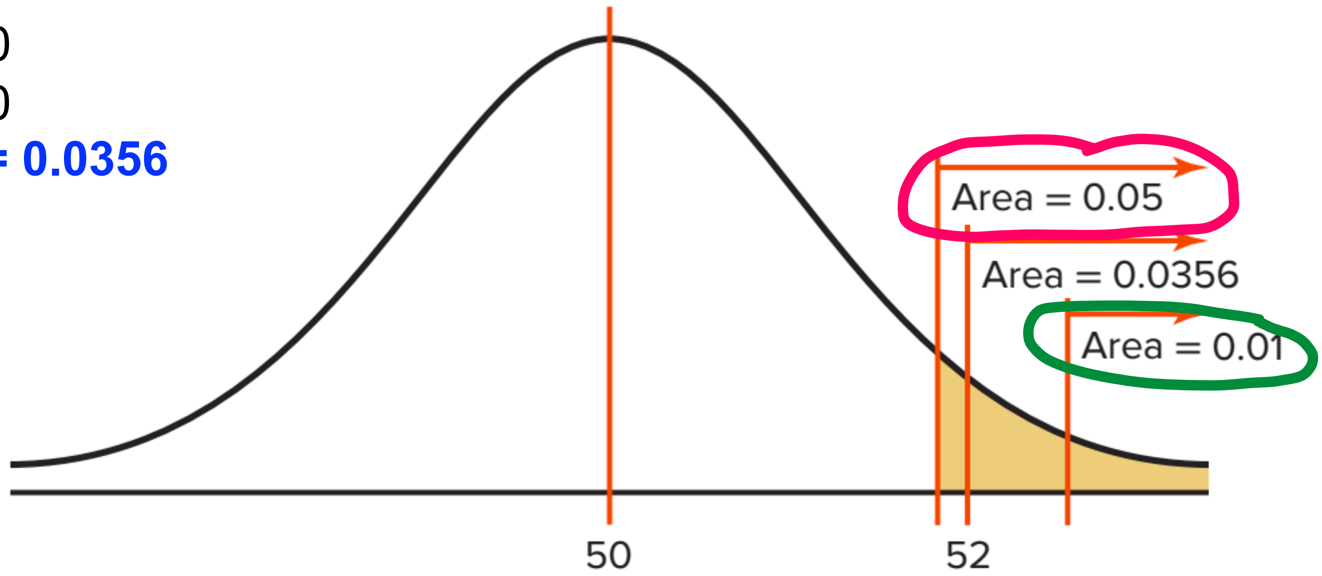
P-value Method…

<u>Example</u>

H_0: μ = 50

H_1: μ > 50

P-value = 0.0356



Decision Rule When Using a *P*-Value

If *P*-value ≤ α, reject the null hypothesis.

If *P*-value > α, do not reject the null hypothesis.

EXAMPLE 8–6 Cost of College Tuition

A researcher wishes to test the claim that the average cost of tuition and fees at a four-year public college is greater than \$5700. She selects a random sample of 36 four-year public colleges and finds the mean to be \$5950. The population standard deviation is \$659. Is there evidence to support the claim at $\alpha = 0.05$? Use the *P*-value method.

Source: Based on information from the College Board.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0: \mu = \$5700 \quad \text{and} \quad H_1: \mu > \$5700 \text{ (claim)}$$

Step 2 Compute the test value.

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{5950 - 5700}{659/\sqrt{36}} = 2.28$$

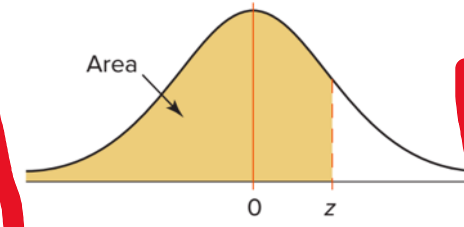
Step 3 Find the *P*-value. Using Table E in Appendix A, find the corresponding area under the normal distribution for $z = 2.28$. It is 0.9887. Subtract this value for the area from 1.0000 to find the area in the right tail.

$$1.0000 - 0.9887 = 0.0113$$

TABLE E *(continued)*

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916



Step 3 Find the P-value. Using Table E in Appendix A, find the corresponding area under the normal distribution for $z = 2.28$. It is 0.9887. Subtract this value for the area from 1.0000 to find the area in the right tail.

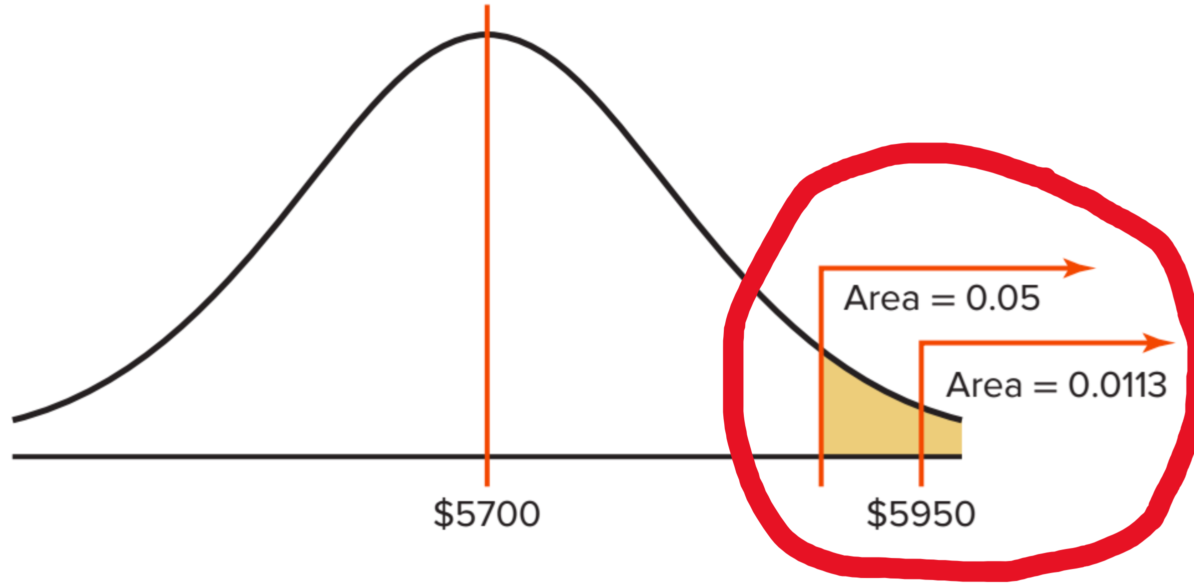
$$1.0000 - 0.9887 = 0.0113$$

Hence, the P-value is 0.0113.

Step 4 Make the decision. Since the P-value is less than 0.05, the decision is to reject the null hypothesis. See Figure 8–17.

FIGURE 8–17

P-Value and α Value for Example 8–6



Step 5 Summarize the results. There is enough evidence to support the claim that the tuition and fees at four-year public colleges are greater than $5700.

Note: Had the researcher chosen $\alpha = 0.01$, the null hypothesis would not have been rejected since the P-value (0.0113) is greater than 0.01.

EXAMPLE 8–7 Wind Speed

A researcher claims that the average wind speed in a certain city is 8 miles per hour. A sample of 32 days has an average wind speed of 8.2 miles per hour. The standard deviation of the population is 0.6 mile per hour. At $\alpha = 0.05$, is there enough evidence to reject the claim? Use the *P*-value method.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0: \mu = 8 \text{ (claim)} \qquad \text{and} \qquad H_1: \mu \neq 8$$

Step 2 Compute the test value.

$$z = \frac{8.2 - 8}{0.6/\sqrt{32}} = 1.89$$

Step 3 Find the *P*-value. Using Table E, find the corresponding area for $z = 1.89$. It is 0.9706. Subtract the value from 1.0000.

$$1.0000 - 0.9706 = 0.0294$$

Since this is a two-tailed test, the area of 0.0294 must be doubled to get the *P*-value.

$$2(0.0294) = 0.0588$$

Step 4 Make the decision. The decision is to not reject the null hypothesis, since the P-value is greater than 0.05. See Figure 8–18.

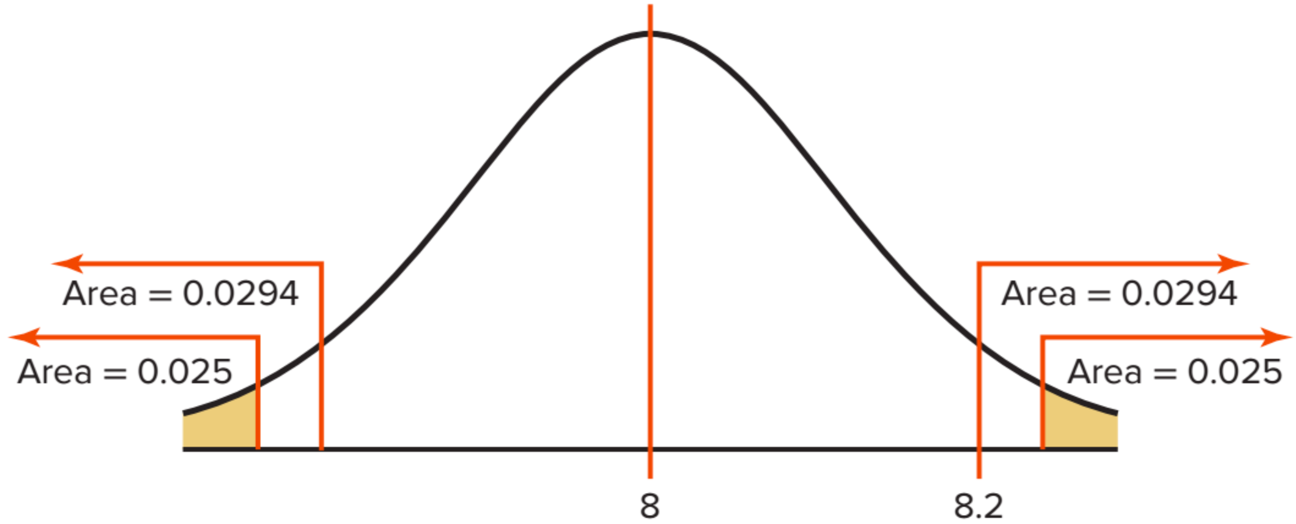


FIGURE 8–18
P-Values and α Values for Example 8–7

Step 5 Summarize the results. There is not enough evidence to reject the claim that the average wind speed is 8 miles per hour.

Decision Rule When Using a *P*-Value

If P-value $\leq \alpha$, reject the null hypothesis.

If P-value $> \alpha$, do not reject the null hypothesis.

Summary

Let $X_1, \ldots, X_n$ be a *large* (e.g., $n > 30$) sample from a population with mean μ and standard deviation σ.

To test a null hypothesis of the form $H_0: \mu \le \mu_0$, $H_0: \mu \ge \mu_0$, or $H_0: \mu = \mu_0$:

- Compute the z-score: $$z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}.$$
  If σ is unknown it may be approximated with s.
- Compute the P-value. The P-value is an area under the normal curve, which depends on the alternate hypothesis as follows:

Alternate Hypothesis	***P*-value**
$H_1: \mu > \mu_0$	Area to the right of z
$H_1: \mu < \mu_0$	Area to the left of z
$H_1: \mu \ne \mu_0$	Sum of the areas in the tails cut off by z and $-z$

p.409 (textbook)

Comments on P-value Method

- The P-values given on calculators and computers are slightly different from those found in tables (for example Table A.2).
- A clear distinction between the α-value and the P-value should be made.
- The **α value is chosen** by the researcher **before** the statistical test is conducted.
- The **P-value is computed** after the sample mean has been found.

Comments on P-value Method…

- There are two schools of using P-values.
- Some researchers do not choose an α value but report the P-value and allow the reader to decide whether the null hypothesis should be rejected.

Guidelines for *P*-Values

If P-value ≤ 0.01, reject the null hypothesis. The difference is highly significant.

If P-value > 0.01 but P-value ≤ 0.05, reject the null hypothesis. The difference is significant.

If P-value > 0.05 but P-value ≤ 0.10, consider the consequences of type I error before rejecting the null hypothesis.

If P-value > 0.10, do not reject the null hypothesis. The difference is not significant.

Comments on P-value Method…

- Others decide on the α value in advance and use the P-value to make the decision.
- Caution: If a researcher selects $\alpha = 0.01$ and the P-value is 0.03, the researcher may decide to change the α value from 0.01 to 0.05 so that the null hypothesis will be rejected.
- This, of course, **should not be done**. If the α level is selected in advance, it should be used in making the decision.

Drawing Conclusions from the Results of Hypothesis Tests

Summary

- The smaller the P-value, the more certain we can be that H_0 is false.
- The larger the P-value, the more plausible H_0 becomes, but we can never be certain that H_0 is true.
- A rule of thumb suggests to reject H_0 whenever $P \leq 0.05$. While this rule is convenient, it has no scientific basis.

➢ Whenever the P-value is less than a particular threshold, the result is said to be "statistically significant" at that level.

➢ For example, if $P \leq 0.05$, the result is statistically significant at the 5% level.

p.413 (textbook)

Drawing Conclusions from the Results of Hypothesis Tests…

Summary

Let α be any value between 0 and 1. Then, if $P \leq \alpha$,

- The result of the test is said to be statistically significant at the $100\alpha\%$ level.
- The null hypothesis is rejected at the $100\alpha\%$ level.
- When reporting the result of a hypothesis test, report the P-value, rather than just comparing it to 5% or 1%.

- Reporting the P-value gives more information about the strength of the evidence against the null hypothesis and allows each reader to decide for himself or herself whether to reject.
- Software packages always output P-values.
- These should be included whenever the results of a hypothesis test are reported.

p.414 (textbook)

The Relationship Between Hypothesis Tests and Confidence Intervals

- Both confidence intervals and hypothesis tests are concerned with determining plausible values for a quantity such as a population mean μ.
- In a hypothesis test for a population mean μ, we specify a particular value of μ (the null hypothesis) and determine whether that value is plausible.
- In contrast, a confidence interval for a population mean μ can be thought of as the collection of all values for μ that meet a certain criterion of plausibility, specified by the confidence level $100(1 - \alpha)\%$.

The Relationship Between Hypothesis Tests and Confidence Intervals…

- The values contained within a two-sided level 100(1 − α)% confidence interval for a population mean μ are precisely those values for which the P-value of a two-tailed hypothesis test will be greater than α.
- Example (p. 417)

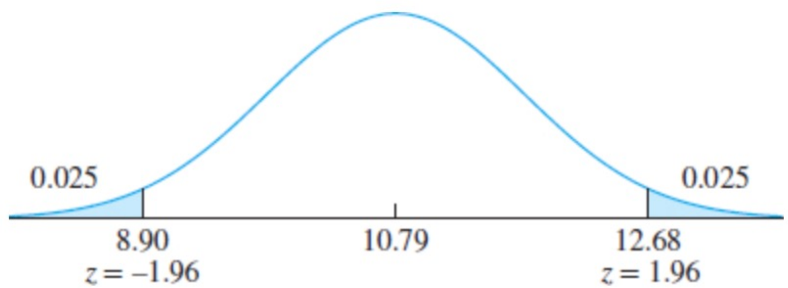


FIGURE 6.4 The sample mean $\overline{X}$ is equal to 12.68. Since 10.79 is an endpoint of a 95% confidence interval based on $\overline{X} = 12.68$, the *P*-value for testing H_0: $\mu = 10.79$ is equal to 0.05.