9. Hypothesis Tests : An Introduction

9.1 Hypothesis Tests : An Introduction

9.2 Hypothesis Tests About μ : σ known

9.3 Hypothesis Tests About μ : σ not known

9.4 Hypothesis Tests About a population proportion : large samples

Tests of Hypothesis

- Test a certain given theory or belief about a population parameter
- Want to find out, using some sample information, whether or not a given claim about a population mean is true

Example : Soft-drink company

- A soft -drink company claim that, on average, its cans contains 12 ounces of soda.
- A government agency may want to test whether or not such cans contain, on average, 12 ounces of soda.
- Sample 100 cans, measure the amount of soda.
- The mean amount of soda is 11.89 ounces.

Q. Can we state that, on average, all such cans contain less than 12 ounces of soda?

➔ **NO!** It can be the sampling error.

9.1 Hypothesis Tests : An Introduction

Example : Court case

- A person is indicated for committing a crime and is being tried in a court.
- “The person is presumed not guilty”.
- The prosecutor’s tries to prove to this person’s guilty.
- Based on the available evidence, the judge or jury will make one of two possible decisions.
  - Not guilty ➔ Null hypothesis H_0
  - Guilty ➔ alternative hypothesis H_1

9.1 Hypothesis Tests : An Introduction

Null Hypothesis

A null hypothesis is a claim(or statement) about a population parameter that is assumed to be true until it is declared false

Alternative Hypothesis

An alternative hypothesis is a claim about a population parameter that will be true if the null hypothesis is false.

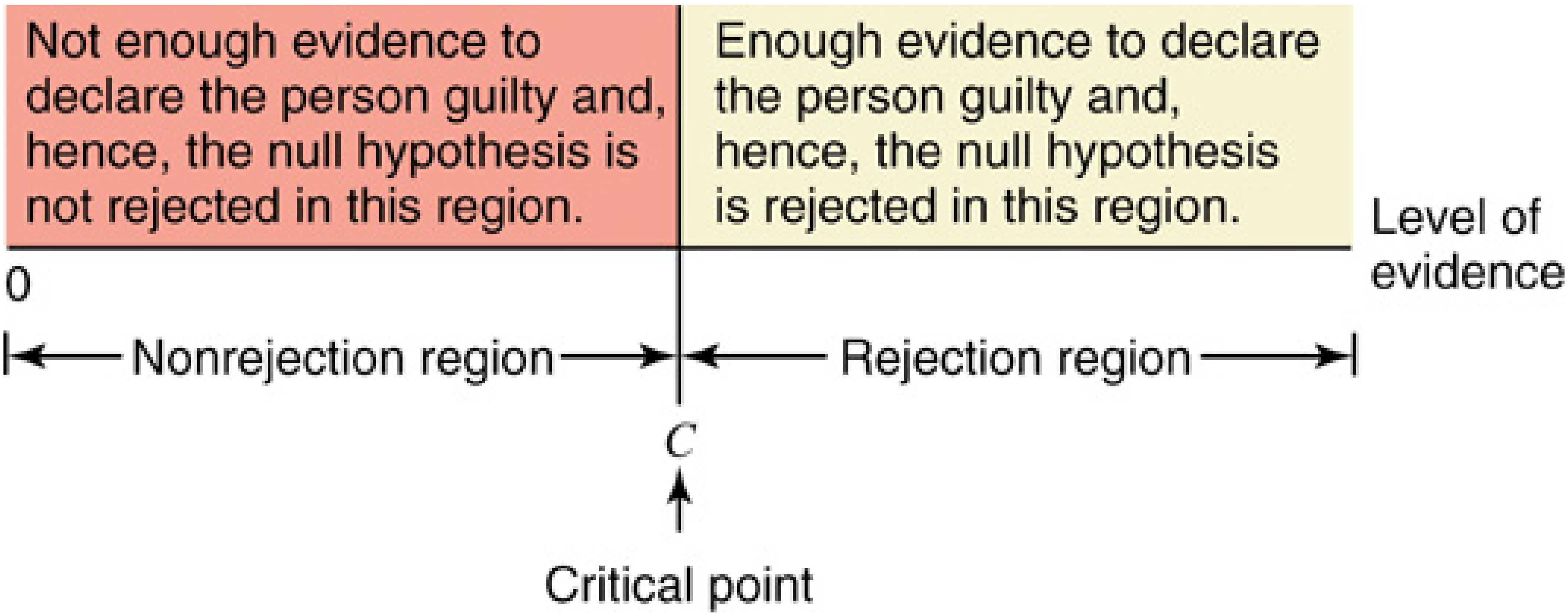
Example : Soft-drink company

- Initially assume that the company's claim is true

$$H_0 : \mu = 12 \quad (\mu \geq 12)$$

$$H_1 : \mu < 12$$

Rejection and Non-rejection regions



- "Do not reject"

  ⟺ "There is not enough evidence to reject H_0"

9.1 Hypothesis Tests : An Introduction

Two types of Error

Type I Error

A Type I error occurs when a true null hypothesis is rejected. The value of α represents the probability of committing this type of error; that is,

$$\alpha = P(H_0 \text{ is rejected } \mid H_0 \text{ is true})$$

The value of α represents the significance level of the test.

Type II Error

A Type II error occurs when a false null hypothesis is not rejected. The value of β represents the probability of committing a Type II error; that is,

$$\beta = P(H_0 \text{ is not rejected } \mid H_0 \text{ is false})$$

The value of 1- β is called the power of the test. It represents the probability of not making a Type II error.

9.1 Hypothesis Tests : An Introduction

Example : Court case

- A person is declared "Guilty"
  - The person has not committed the crime but is declared guilty (because of what may be false evidence)

    → Type I error
  - The person has committed the crime and is rightfully declared guilty.

    → correct decision
- A person is declared "Not Guilty"
  - The person has not committed the crime and is declared not guilty.

    → correct decision
  - The person has committed the crime, but, because of the lack of enough evidence, is declared not guilty

    → Type II error

9.1 Hypothesis Tests : An Introduction

Example : Court case

Table 9.1

		Actual Situation	
		The Person Is Not Guilty	The Person Is Guilty
Court's decision	The person is not guilty	Correct decision	Type II or β error
	The person is guilty	Type I or α error	Correct decision

*** In statistical hypothesis testing procedures, we need to consider**

1) Type I error

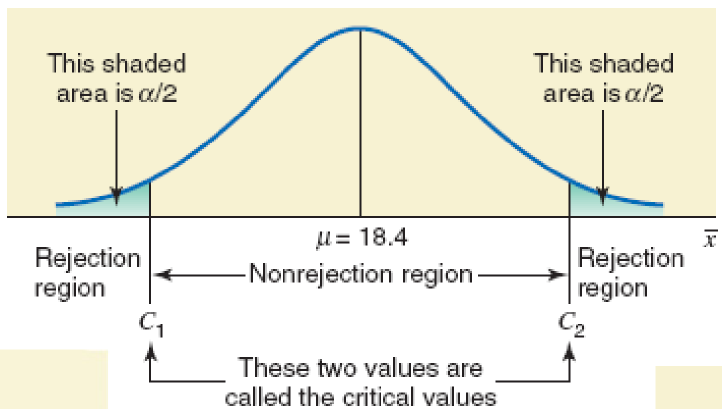
2) rejection region

9.1 Hypothesis Tests : An Introduction

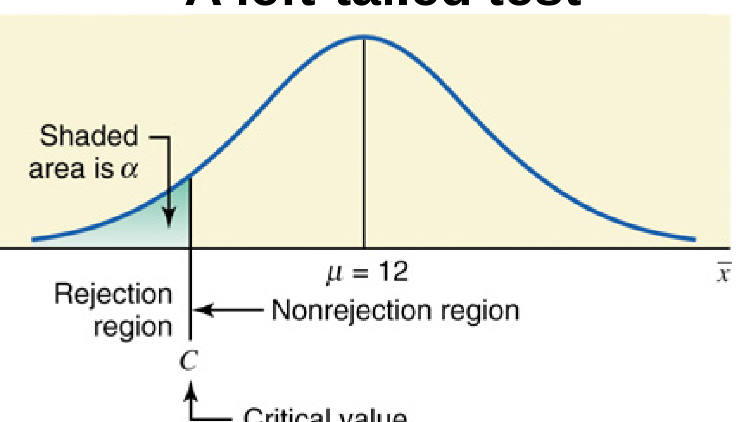
Tails of the Test

A two-tailed test has rejection regions in both tails, a left-tailed test has the rejection region in the left tail, and a right-tailed test has the rejection region in the right tail of the distribution curve

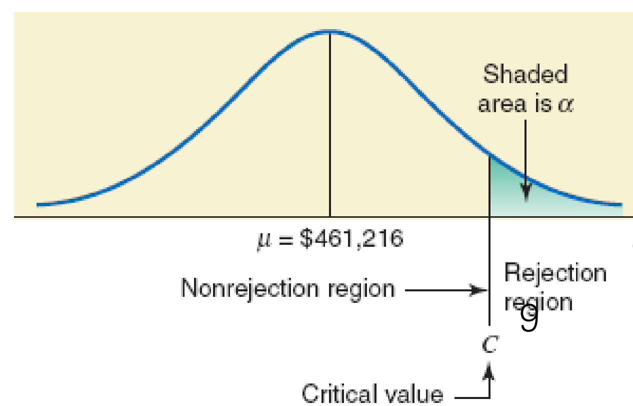
A two-tailed test



A left-tailed test



A right-tailed test



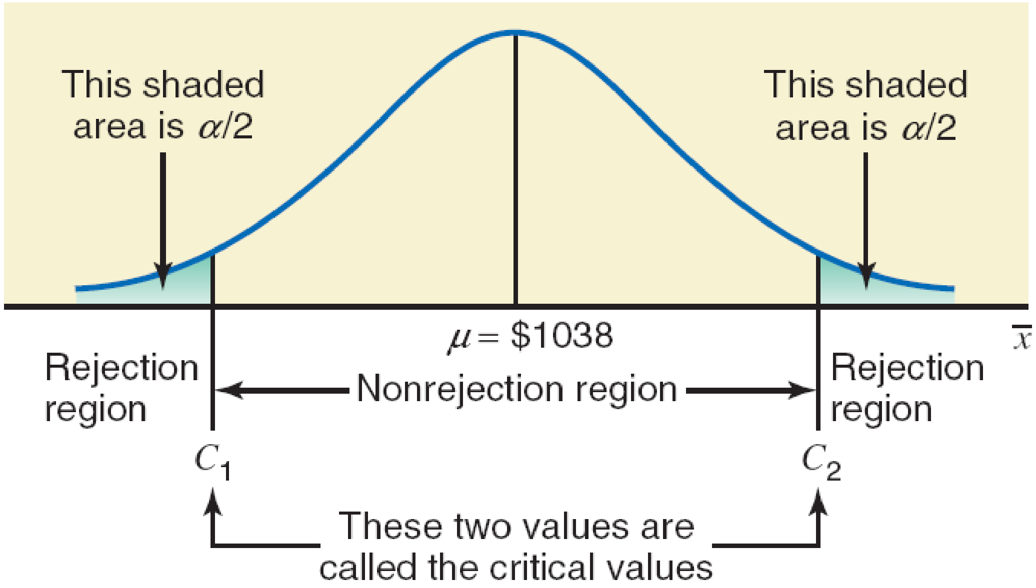
9.1 Hypothesis Tests : An Introduction

A two-tailed test

According to the US Bureau of labor Statistics, people in the United States who has a bachelor's degree and were employed earned an average of $ 1038 a week in 2010. Suppose an economist wants to check whether this mean changed since 2010.

$$H_0 : \mu = \$1038$$

$$H_1 : \mu \neq \$1038$$

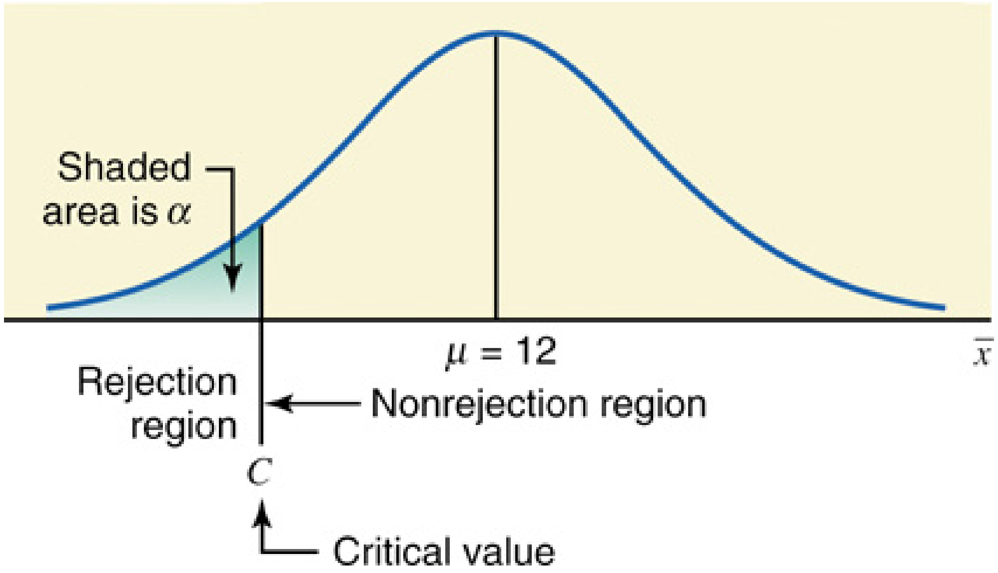


9.1 Hypothesis Tests : An Introduction

A left-tailed test

Soft-drink company example

$$H_0 : \mu = 12oz$$
$$H_1 : \mu < 12oz$$



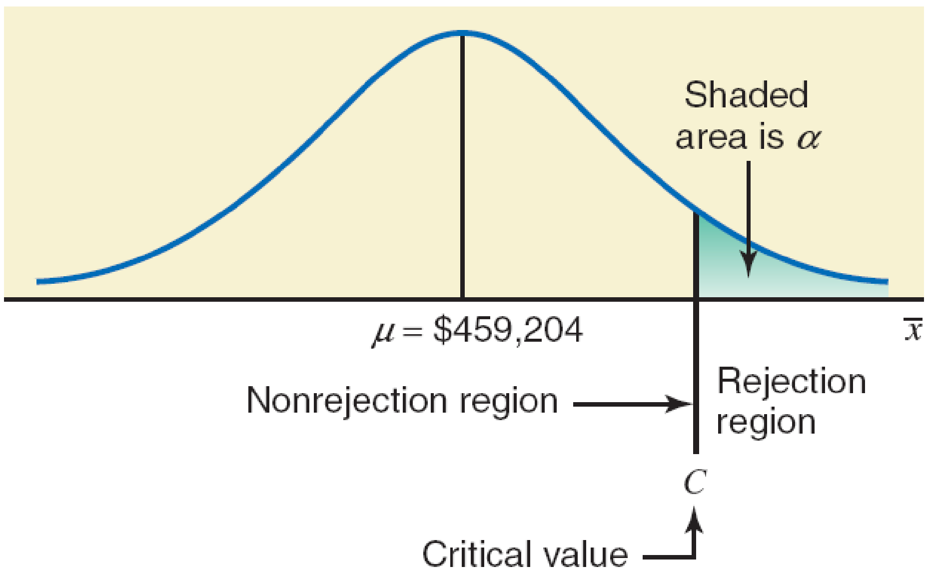
9.1 Hypothesis Tests : An Introduction

A right-tailed test

According to www.city-data.com, the average price of homes in West Orange, NJ, was \$459,204 in 2009. Suppose a real estate researcher wants to check whether the current mean price of homes in this town is higher than \$459,204.

$$H_0 : \mu = \$461{,}216$$

$$H_1 : \mu > \$461{,}216$$



Signs in H_0 and H_1 and Tails of a test

	Two-Tailed Test	Left-Tailed Test	Right-Tailed Test
Sign in the null hypothesis H_0	$=$	$=$ or $\geq$	$=$ or $\leq$
Sign in the alternative hypothesis H_1	$\neq$	$<$	$>$
Rejection region	In both tails	In the left tail	In the right tail

9.1 Hypothesis Tests : An Introduction

Two procedures to make test of hypothesis

1. **The p-value approach**
   - calculate what is called the p-value for the observed value of the sample statistic.
   - compare the p-value with a predetermined significance level
   - make a decision.

2. **The critical-value approach**
   - find the critical value(s) from a table
   - Find the value of the test statistic for the observed value of the sample statistic
   - Compare these two values
   - Make a decision

9.2 Hypothesis Tests About μ : σ known

Case I : use Normal distribution

1. The population standard deviation σ is known
2. The sample size is small ($n < 30$)
3. The population from which the sample is selected is normally distributed

Case II : Use Normal distribution

1. The population standard deviation σ is known
2. The sample size is large ($n \geq 30$)

Case III : Use nonparametric method

1. The population standard deviation σ is known
2. The sample size is small ($n < 30$)
3. The population from which the sample is selected is not normally distributed (or its distribution is unknown)

1. The p-value Approach

- find a probability value such that a given null hypothesis is rejected for any α (significance level) greater than this value and it is not rejected for any α less than this value.

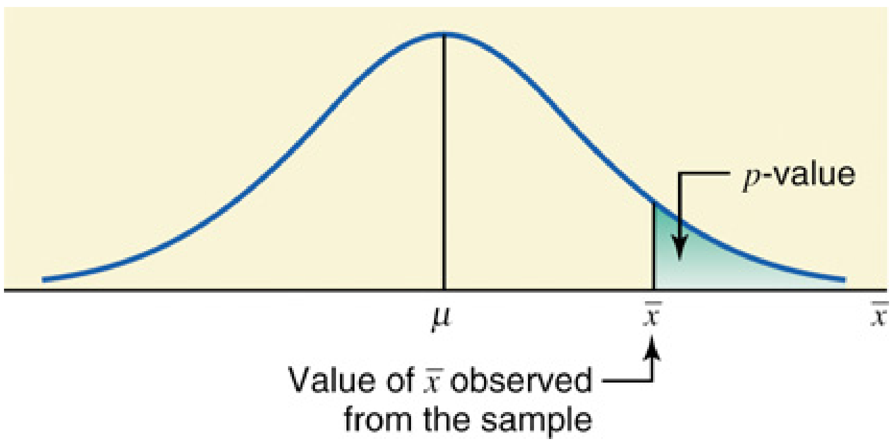
P-value

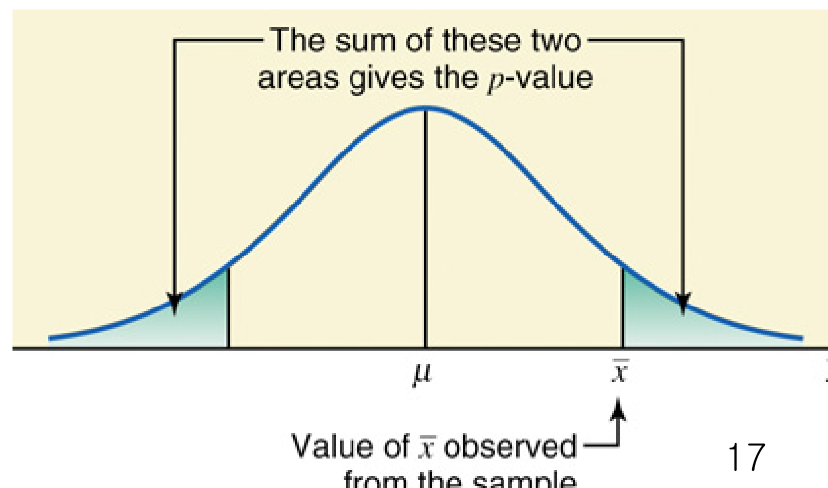
Assuming that the null hypothesis is true, the p-value can be defined as the probability that a sample statistic (such as the sample mean) is at least as far away from the hypothesized value in the direction of the alternative hypothesis as the one obtained from the sample data under consideration. Note that the p-value is the smallest significance level at which the null hypothesis is rejected.

9.2 Hypothesis Tests About μ : σ known

- Reject H_0 if $p - value < \alpha$

- Do not reject H_0 if $p - value \geq \alpha$





9.2 Hypothesis Tests About μ : σ known

Calculating the z value for $\bar{x}$

When using the normal distribution, the value of z for $\bar{x}$ for a test of hypothesis about μ is computed as follows;

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \quad \text{where} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

The value of z calculated for $\bar{x}$ using this formula is also called the observed value of z.

Steps to perform a test of hypothesis using the p-value approach

1. State the null and alternative hypothesis
2. Select the distribution to use
3. Calculate the p-value
4. Make a decision

Example 9.1

At Canon Food Corporation, it took an average of 90 minutes for new workers to learn a food processing job. Recently the company installed a new food processing machine. The supervisor at the company wants to find if the mean time taken by new workers to learn the food processing procedure on this new machine is different from 90 minutes. A sample of 20 workers showed that it took, on average, 85 minutes for them to learn the food processing procedure on the new machine. It is known that the learning times for all new workers are normally distributed with a population standard deviation of 7 minutes. Find the p-value for the test that the mean learning time for the food processing procedure on the new machine is different from 90 minutes. What will your conclusion be if $\alpha=0.01$?

Example 9.2

The management of Priority Health club claims that its members lose an average of 10 pounds or more within the first month after joining the club. A consumer agency that wanted to check this claim took a random sample of 36 members of this health club and found that they lost an average of 9.2 pounds within the first month of membership. The population standard deviation is known to be 2.4 pounds. Find the p-value for this test. What will your decision be if $\alpha=0.01$? What if $\alpha=0.05$?

2. The Critical-Value Approach

- the traditional approach
- use a predetermined value of the significance level α

Steps to perform a test using the Critical-Value approach

1. State the null and alternative hypothesis
2. Select the distribution to use
3. Determine the rejection and non-rejection regions
4. Calculate the value of the test statistic
5. Make a decision

9.2 Hypothesis Tests About μ : σ known

Test Statistic

In tests of hypotheses about μ using the normal distribution, the random variable

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{x}}} \qquad \text{where} \qquad \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Is called the test statistic. The test statistic can be defined as a rule or criterion that is used to make the decision whether or not to reject the null hypothesis.

Example 9.3

The TIV Telephone Company provides long-distance telephone service in an area. According to the company's records, the average length of all long-distance calls placed through this company in 2004 was 12.44 minutes. The company's management wanted to check if the mean length of the current long-distance calls is different from 12.44 minutes. A sample of 150 such calls placed through this company produced a mean length of 13.71 minutes. The standard deviation of all such calls is 2.65 minutes. Using the 2%significance level, can you conclude that the mean length of all current long-distance calls is different from 12.44 minutes?

Example 9.4

The mayor of a large city claims that the average net worth of families living in this city is at least \$300,000. a random sample of 25 families selected from this city produced a mean net worth of \$288,000. Assume that the net worths of all families in this city have a normal distribution with the population standard deviation of \$80,000. Using the 2.5% significance level, can you conclude that the mayor's claim is false?

9.3 Hypothesis Tests About μ : σ Not Known

Case I : use t distribution

1. The population standard deviation σ is unknown
2. The sample size is small ($n < 30$)
3. The population from which the sample is selected is normally distributed

Case II : Use t distribution

1. The population standard deviation σ is unknown
2. The sample size is large ($n \geq 30$)

Case III : Use nonparametric method

1. The population standard deviation σ is unknown
2. The sample size is small ($n < 30$)
3. The population from which the sample is selected is not normally distributed (or its distribution is unknown)

9.3 Hypothesis Tests About μ : σ Not Known

Test Statistic

The value of the test statistic t for the sample mean $\bar{x}$ is computed as

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \qquad \text{where} \qquad s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

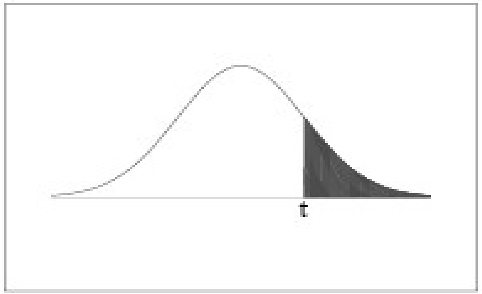
The value of t calculated for $\bar{x}$ by using this formula is also called the observed value of t.

Example 9.5

A psychologist claims that the mean age at which children start walking is 12.5 months. Carol wanted to check if this claim is true. She took a random sample of 18 children and found that the mean age at which these children started walking was 12.9 months with a standard deviation of 0.8 months. It is known that the ages at which all children start walking are approximately normally distributed. Find the p-value for the test that the mean age at which all children start walking is different from 12.5 months. What will your conclusion be if the significance level is 1%?

9.3 Hypothesis Tests About μ : σ Not Known

t-Distribution Table



The shaded area is equal to α for $t = t_\alpha$.

df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878

Example 9.6

Grand Auto Corporation produces auto batteries. The company claims that its top-of-the-line Never Die batteries are good, on average, for at least 65 months. A consumer protection agency tested 45 such batteries to check this claim. It found that the mean life of these 45 batteries is 63.4 months and the standard deviation is 3 months. Find the p-value for the test that the mean life of all such batteries is less than 65 months. What will your conclusion be if the significance level is 2.5%?

Example 9.7

Refer to Example 9-5. A psychologist claims that the mean age at which children start walking is 12.5 months. Carol wanted to check if this claim is true. She took a random sample of 18 children and found that the mean age at which these children started walking was 12.9 months with a standard deviation of 0.8 months. Using the 1% significance level, can you conclude that the mean age at which all children start walking is different from 12.5 months? Assume that the ages at which all children start walking have an approximately normal distribution.

Example 9.8

The management at Massachusetts Savings Bank is always concerned about the quality of service provided to its customers. Within the old computer system, a teller at this bank could serve, on average, 22 customers per hour. The management noticed that with this service rate, the waiting time for customers was too long. Recently the management of the bank installed a new computer system in the bank, expecting that it would increase the service rate and consequently make the customers happier by reducing the waiting time. To check if the new computer system is more efficient than the old system, the management of the bank took a random sample of 70 hours and found that during these hours the mean number of customers served by tellers was 27 per hour with a standard deviation of 2.5. Testing at the 1% significance level, would you conclude that the new computer system is more efficient than the old computer system?

9.4 Hypothesis Tests About a Population Proportion : Large Sample

Test Statistic

The value of the test statistic z for the sample proportion $\hat{p}$ is computed as

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} \qquad \text{where} \qquad \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

The value of p used in this formula is the one used in the null hypothesis. The value of z calculated for $\hat{p}$ using the above formula is also called the observed value of z

9.4 Hypothesis Tests About a Population Proportion : Large Sample

Example 9.9

In a 2011 National Institute on Alcohol Abuse and Alcoholism survey, 33% of American adults said that they had never consumed alcohol (USA TODAY, November 17, 2011). Suppose that this result is true for the 2011 population of American adults. In a recent random sample of 2300 adult Americans, 35% said that they had never consumed alcohol. Find the p-value to test the hypothesis that the current percentage of American adults who have never consumed alcohol is different from 33%. What is your conclusion if the significance level is 5%?

9.4 Hypothesis Tests About a Population Proportion : Large Sample

Example 9.10

When working properly, a machine that is used to make chips for calculators does not produce more than 4% defective chips. Whenever the machine produces more than 4% defective chips, it needs an adjustment. To check if the machine is working properly, the quality control department at the company often takes samples of chips and inspects them to determine if they are good or defective. One such random sample of 200 chips taken recently from the production line contained 12 defective chips. Find the p-value to test the hypothesis whether or not the machine needs an adjustment. What would you conclusion be if the significance level is 2.5%?

Example 9.11

Refer to Example 9-9. In a 2011 National Institute on Alcohol Abuse and Alcoholism survey, 33% of American adults said that they had never consumed alcohol (USA TODAY, November 17, 2011). Suppose that this result is true for the 2011 population of American adults. In a recent random sample of 2300 adult Americans, 35% said that they had never consumed alcohol. Using a 5% significance level, can you conclude that the current percentage of American adults who have never consumed alcohol is different from 33%?

Example 9.12

Direct Mailing Company sells computers and computer parts by mail. The company claims that at least 90% of all orders are mailed within 72 hours after they are received. The quality control department at the company often takes samples to check if this claim is valid. A recently taken sample of 150 orders showed that 129 of them were mailed within 72 hours. Do you think the company's claim is true? Use a 2.5% significance level.