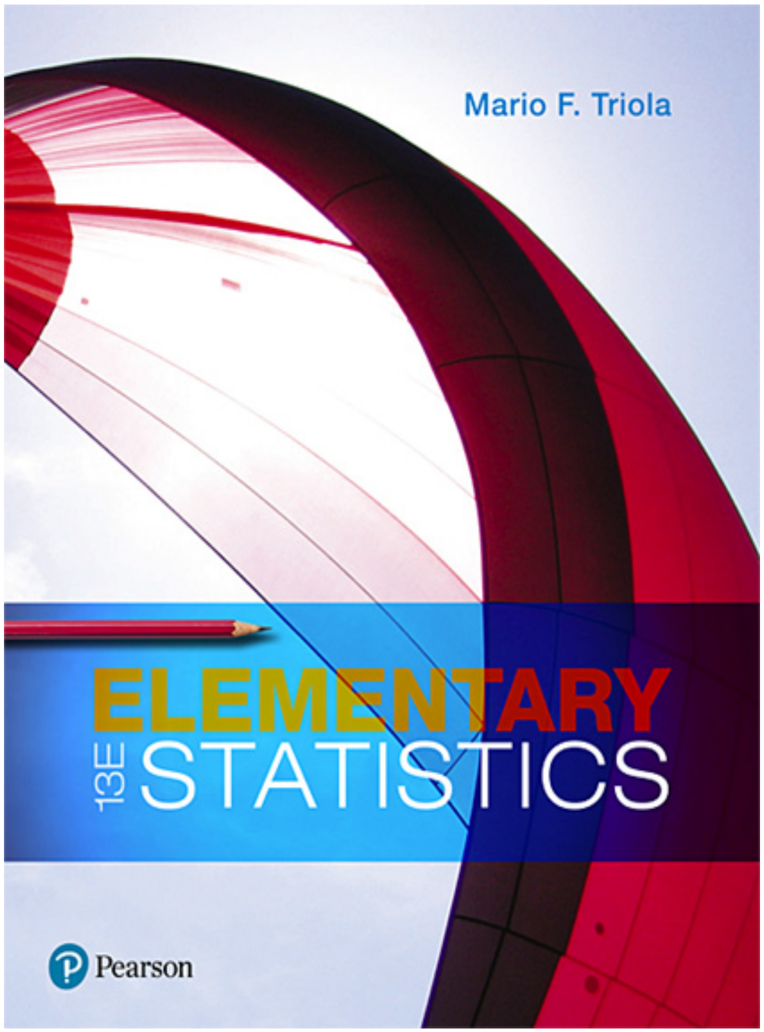
Elementary Statistics

Thirteenth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

8-1 Basics of Hypothesis Testing

8-2 Testing a Claim about a Proportion

8-3 Testing a Claim About a Mean

8-4 Testing a Claim About a Standard Deviation or Variance

Key Concept (1 of 2)

This section describes a complete procedure for testing a claim made about a population proportion p. We illustrate hypothesis testing with the P-value method, the critical value method, and the use of confidence intervals. The methods of this section can be used with claims about population proportions, probabilities, or the decimal equivalents of percentages.

Key Concept (2 of 2)

There are different methods for testing a claim about a population proportion. Part 1 of this section is based on the use of a normal approximation to a binomial distribution, and this method serves well as an introduction to basic concepts, but it is not a method used by professional statisticians. Part 2 discusses other methods that might require the use of technology.

Testing a Claim About a Population Proportion (Normal Approximation Method): Objective

Objective

Conduct a formal hypothesis test of a claim about a population proportion *p*.

Testing a Claim About a Population Proportion (Normal Approximation Method): Notation

Notation

n = sample size or number of trials

p = population proportion (p is the value used in the statement of the null hypothesis)

$\hat{p} = \frac{x}{n}$ (**sample** proportion)

$q = 1 - p$

Testing a Claim About a Population Proportion (Normal Approximation Method): Requirements (1 of 2)

Requirements

1. The sample observations are a simple random sample.
2. The conditions for a **binomial distribution** are satisfied:
   - There is a fixed number of trials.
   - The trials are independent.
   - Each trial has two categories of “success” and “failure.”
   - The probability of a success remains the same in all trials.

Testing a Claim About a Population Proportion (Normal Approximation Method): Requirements (2 of 2)

Requirements

3. The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied, so **the binomial distribution of sample proportions can be approximated by a normal distribution with** $\mu = np$ and $\sigma = \sqrt{npq}$ (as described in Section 6-6).

   Note that p used here is the **assumed** proportion used in the claim, not the sample proportion $\hat{p}$.

Testing a Claim About a Population Proportion (Normal Approximation Method): Test Statistic for Testing a Claim about a Proportion (1 of 2)

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Testing a Claim About a Population Proportion (Normal Approximation Method): Test Statistic for Testing a Claim about a Proportion (2 of 2)

- ***P*-values:** *P*-values are automatically provided by technology. If technology is not available, use the standard normal distribution (Table A-2) and refer to Figure 8-3 on page 364.
- **Critical values:** Use the standard normal distribution (Table A-2).

Equivalent Methods

When testing claims about proportions, the confidence interval method is not equivalent to the *P*-value and critical value methods, so the confidence interval method could result in a different conclusion. (Both the *P*-value method and the critical value method use the same standard deviation based on the **claimed proportion** *p*, so they are equivalent to each other, but the confidence interval method uses an estimated standard deviation based on the **sample proportion**.)

Recommendation: Use a confidence interval to **estimate** a population proportion, but use the *P*-value method or critical value method for **testing a claim** about a proportion.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (1 of 16)

1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% (or 545) of them responded with "no." Use these results to test the claim that most consumers are uncomfortable with drone deliveries. We interpret "most" to mean "more than half" or "greater than 0.5."

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (2 of 16)

Requirement Check

We first check the three requirements.

1. The 1009 consumers are randomly selected.
2. There is a fixed number (1009) of independent trials with two categories (the subject is uncomfortable with drone deliveries or is not).

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries

Requirement Check

3. The requirements $np \geq 5$ and $nq \geq 5$ are both satisfied with $n = 1009$, $p = 0.5$, and $q = 0.5$. [The value of $p = 0.5$ comes from the claim. We get $np = (1009)(0.5) = 504.5$, which is greater than or equal to 5, and we get $nq = (1009)(0.5) = 504.5$, which is also greater than or equal to 5.]

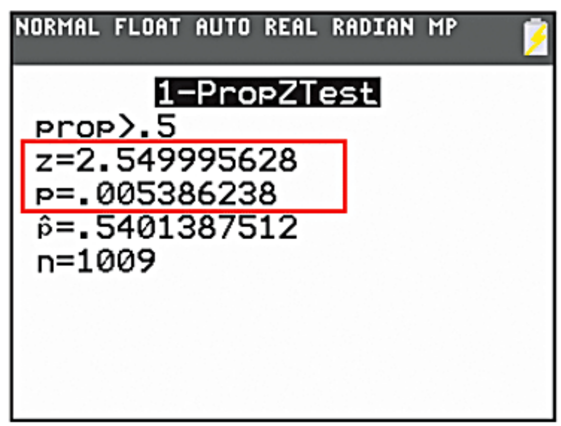
The three requirements are satisfied.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (4 of 16)

Solution: *P*-Value Method

Technology: Computer programs and calculators usually provide a *P*-value, so the *P*-value method is used. See the accompanying TI-83/84 Plus calculator results showing the alternative hypothesis of “prop > 0.5,” the test statistic of $z = 2.55$ (rounded), and the *P*-value of 0.0054 (rounded).

TI-83/84 Plus



Pearson

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (5 of 16)

Solution: *P*-Value Method

Table A-2: If technology is not available, Figure 8-1 on page 360 lists the steps for using the *P*-value method. Using those steps from Figure 8-1, we can test the claim as follows.

Step 1: The original claim is that most consumers are uncomfortable with drone deliveries, and that claim can be expressed in symbolic form as $p > 0.5$.

Step 2: The opposite of the original claim is $p \leq 0.5$.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries

Solution: *P*-Value Method

Step 3: Of the preceding two symbolic expressions, the expression $p > 0.5$ does not contain equality, so it becomes the alternative hypothesis. The null hypothesis is the statement that p equals the fixed value of 0.5. We can therefore express H_0 and H_1 as follows:

H_0: $p = 0.5$

H_1: $p > 0.5$ (original claim)

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (7 of 16)

Solution: *P*-Value Method

Step 4: For the significance level, we select $\alpha = 0.05$, which is a very common choice.

Step 5: Because we are testing a claim about a population proportion p, the sample statistic $\hat{p}$ is relevant to this test. The sampling distribution of sample proportions $\hat{p}$ can be approximated by a normal distribution in this case (as described in Section 6-3).

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (8 of 16)

Solution: *P*-Value Method

Step 6: The test statistic $z = 2.55$ can be found by using technology or it can be calculated by using

$\hat{p} = \frac{545}{1009}$ (sample proportion), $n = 1009$ (sample size),

$p = 0.5$ (assumed in the null hypothesis), and

$q = 1 - 0.5 = 0.5$.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{545}{1009} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1009}}} = 2.55$$

Pearson

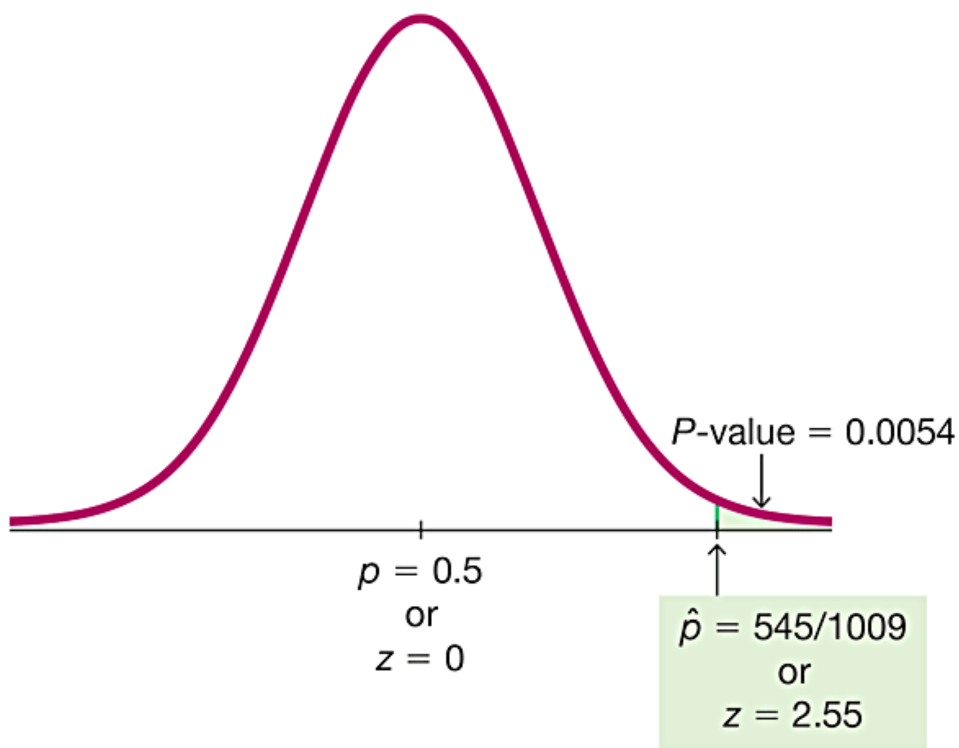
Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (9 of 16)

Solution: *P*-Value Method

Because this hypothesis test is right-tailed with a test statistic of $z = 2.55$. The *P*-value is the area to the right of $z = 2.55$. Referring to Table A-2, the cumulative area to the **left** of $z = 2.55$ is 0.9946, so the area to the right of that test statistic is $1 - 0.9946 = 0.0054$. We get *P*-value = 0.0054.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries

Solution: *P*-Value Method



Pearson

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries

Solution: *P*-Value Method

Step 7: Because the *P*-value of 0.0054 is less than or equal to the significance level of $\alpha = 0.05$, we reject the null hypothesis.

Step 8: Because we reject H_0: $p = 0.5$, we support the alternative hypothesis of $p > 0.5$. We conclude that there is sufficient sample evidence to support the claim that more than half of consumers are uncomfortable with drone deliveries.

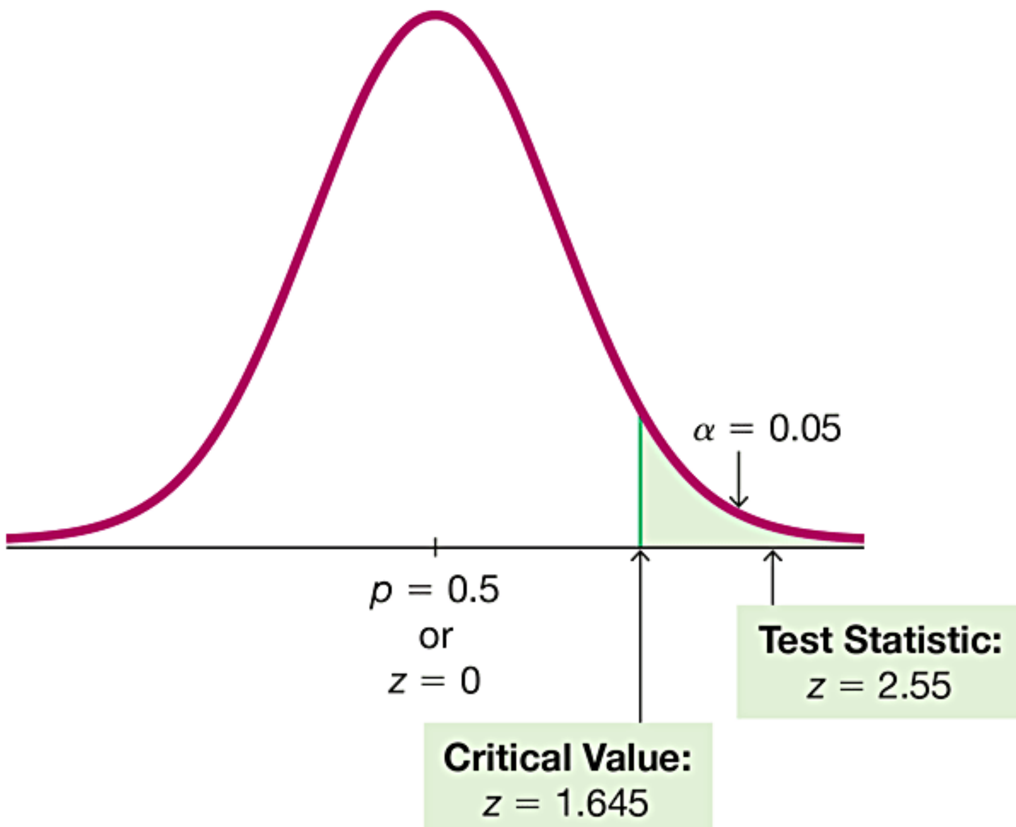
Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (12 of 16)

Solution: Critical Value Method (steps 1 to 5 are the same as the previous method)

Step 6: The test statistic is computed to be $z = 2.55$. With the critical value method, we now find the critical values. This is a right-tailed test, so the area of the critical region is an area of $\alpha = 0.05$ in the right tail. Referring to Table A-2 and applying the methods of Section 6-1, we find that the critical value is $z = 1.645$, which is at the boundary of the critical region.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (13 of 16)

Solution: Critical Value Method



Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (14 of 16)

Solution: Critical Value Method

Step 7: Because the test statistic does fall within the critical region, we reject the null hypothesis.

Step 8: Because we reject H_0: $p = 0.5$, we conclude that there is sufficient sample evidence to support the claim that most (more than half) consumers are uncomfortable with drone deliveries.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (15 of 16)

Solution: Confidence Interval Method

The claim "The Majority of Consumers Are Not Comfortable with Drone Deliveries," can be tested with a 0.05 significance level by constructing a 90% confidence interval.

Example: Claim – Most Consumers Uncomfortable with Drone Deliveries (16 of 16)

Solution: Confidence Interval Method

The 90% confidence interval estimate of the population proportion p is found using the sample data of $n = 1009$ and $\hat{p} = \frac{545}{1009}$.

Using the methods of Section 7-1 we get: $0.514 < p < 0.566$. The entire range of values in this confidence interval is greater than 0.5. We are 90% confident that the limits of 0.514 and 0.566 contain the true value of p, the sample data appear to support the claim that most (more than 0.5) consumers are uncomfortable with drone deliveries.

Finding the Number of Successes *x*

When using technology for hypothesis tests of proportions, we must usually enter the sample size n and the number of successes x, but in real applications the sample proportion $\hat{p}$ is often given instead of x. The number of successes x can be found by evaluating $x = n\hat{p}$.

Note that in the next example, the result of 5587.712 adults must be rounded to the nearest whole number of 5588.

Example: Finding the number of Successes *x* (1 of 2)

A study of sleepwalking or “nocturnal wandering” was described in **Neurology** magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked. What is the actual number of adults who have sleepwalked?

Example: Finding the number of Successes *x* (2 of 2)

Solution

The number of adults who have sleepwalked is 29.2% of 19,136, or $0.292 \times 19,136 = 5587.712$, but the result must be a whole number, so we round the product to the nearest whole number of 5588.

Example: Fewer Than 30% of Adults have Sleepwalked? (1 of 8)

Using the same sleepwalking data from the previous example (n = 19,136 and $\hat{p}$ = 29.2%), would a reporter be justified in stating that "fewer than 30% of adults have sleepwalked"?

Let's use a 0.05 significance level to test the claim that for the adult population, the proportion of those who have sleepwalked is less than 0.30.

Example: Fewer Than 30% of Adults have Sleepwalked? (2 of 8)

Solution

Requirement Check

(1) The sample is a simple random sample.

(2) There is a fixed number (19,136) of independent trials with two categories (a subject has sleepwalked or has not).

(3) The requirements $np \geq 5$ and $nq \geq 5$ are both satisfied with $n = 19{,}136$ and $p = 0.30$. [We get $np = (19{,}136)(0.30) = 5740.8$, which is greater than or equal to 5, and we also get $nq = (19{,}136)(0.70) = 13{,}395.2$, which is greater than or equal to 5.]

The three requirements are all satisfied.

Example: Fewer Than 30% of Adults have Sleepwalked? (3 of 8)

Solution

Step 1: The original claim is expressed in symbolic form as $p < 0.30$.

Step 2: The opposite of the original claim is $p \geq 0.30$.

Step 3: Because $p < 0.30$ does not contain equality, it becomes H_1. We get

H_0: $p = 0.30$ (null hypothesis)

H_1: $p < 0.30$ (alternative hypothesis and original claim)

Example: Fewer Than 30% of Adults have Sleepwalked? (4 of 8)

Solution

Step 4: The significance level is $\alpha = 0.05$.

Step 5: Because the claim involves the proportion p, the statistic relevant to this test is the sample proportion $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.

Example: Fewer Than 30% of Adults have Sleepwalked? (5 of 8)

Solution

Step 6: Technology If using technology, the test statistic and the *P*-value will be provided. See the accompanying results from StatCrunch showing that the test statistic is $z = -2.41$ (rounded) and the *P*-value = 0.008.

StatCrunch

Hypothesis test results:
p : Proportion of successes
$H_0 : p = 0.3$
$H_A : p < 0.3$

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	5588	19136	0.29201505	0.0033127149	-2.4103945	0.008

Example: Fewer Than 30% of Adults have Sleepwalked? (6 of 8)

Solution

Table A-2 If technology is not available, proceed as follows to conduct the hypothesis test using the *P*-value method.

The test statistic $z = -2.41$ is calculated as follows:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{5588}{19{,}136} - 0.30}{\sqrt{\frac{(0.30)(0.70)}{19{,}136}}} = -2.41$$

Example: Fewer Than 30% of Adults have Sleepwalked? (7 of 8)

Solution

For this left-tailed test, the *P*-value is the area to the left of the test statistic. Using Table A-2, we see that the area to the left of $z = -2.41$ is 0.0080, so the *P*-value is 0.0080.

Step 7: Because the *P*-value of 0.0080 is less than or equal to the significance level of 0.05, we reject the null hypothesis.

Example: Fewer Than 30% of Adults have Sleepwalked? (8 of 8)

Interpretation

Because we reject the null hypothesis, we support the alternative hypothesis. We therefore conclude that there is sufficient evidence to support the claim that fewer than 30% of adults have sleepwalked.

Exact Methods for Testing Claims About population Proportion *p*

Identify the sample size n, the number of successes x, and the claimed value of the population proportion p; then find the P-value by using technology for finding binomial probabilities as follows:

- Left-tailed test: P-value = P(x or fewer successes among n trials)
- Right-tailed test: P-value = P(x or more successes among n trials)
- Two-tailed test: P-value = twice the smaller of the preceding left-tailed and right-tailed values

Example: Using the Exact Method (1 of 6)

In testing a method of gender selection, 10 randomly selected couples are treated with the method, they each have a baby, and 9 of the babies are girls. Use a 0.05 significance level to test the claim that with this method, the probability of a baby being a girl is greater than 0.75.

Example: Using the Exact Method (2 of 6)

Solution

Requirement Check

The normal approximation method described in Part 1 of this section requires that $np \geq 5$ and $nq \geq 5$, but $nq = (10)(0.25) = 2.5$, so the requirement is violated. The exact method has only the requirements of being a simple random sample and satisfying the conditions for binomial distribution and those two requirements are satisfied.

Example: Using the Exact Method (3 of 6)

Solution

Here are the null and alternative hypotheses:

H_0: $p = 0.75$ (null hypothesis)

H_1: $p > 0.75$ (alternative hypothesis and original claim)

Example: Using the Exact Method (4 of 6)

Solution

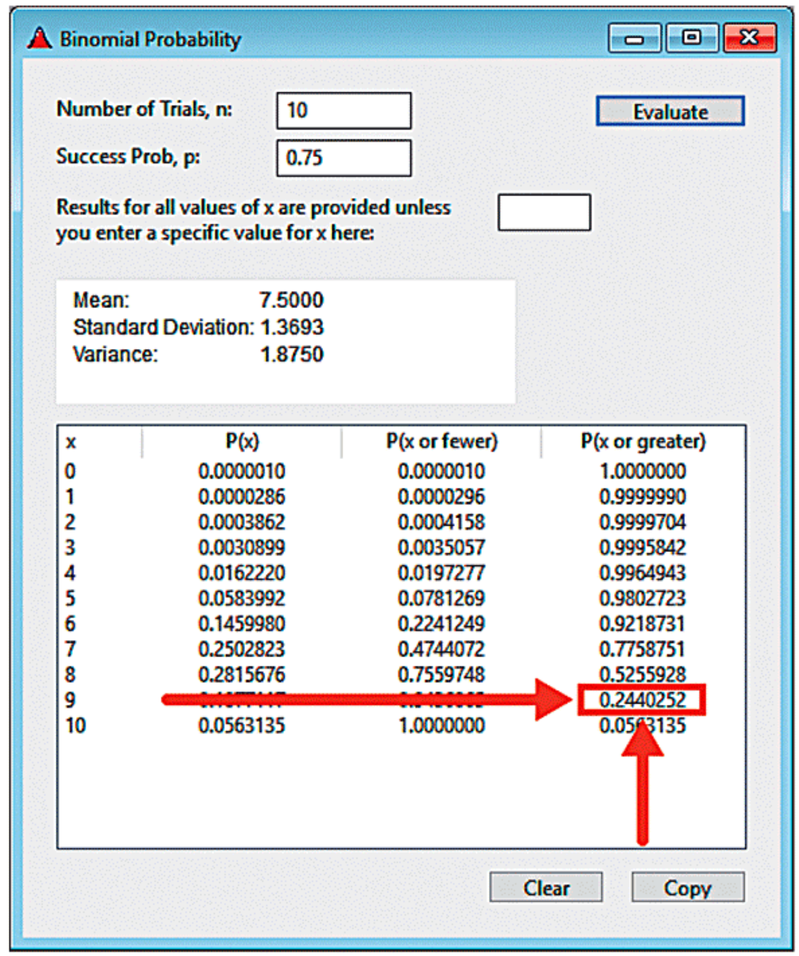
Instead of using the normal distribution, we use technology to find probabilities in a binomial distribution with $p = 0.75$. Because this is a right-tailed test, the *P*-value is the probability of 9 or more successes among 10 trials, assuming that $p = 0.75$. See the Statdisk display of exact probabilities from the binomial distribution on the next slide.

Example: Using the Exact Method (5 of 6)

Solution

This Statdisk display shows that the probability of 9 or more successes is 0.2440252, so the *P*-value is 0.2440252. The *P*-value is high, so we fail to reject the null hypothesis. There is not sufficient evidence to support the claim that with the gender selection method, the probability of a girl is greater than 0.75.

Statdisk



Pearson

Example: Using the Exact Method (6 of 6)

With the exact method, the **actual** probability of a type I error is less than or equal to α, which is the **desired** probability of a type I error.

Improving the Exact Method

A **simple continuity correction** improves the conservative behavior of the exact method with an adjustment to the *P*-value that is obtained by subtracting from it the value that is one-half the binomial probability at the boundary.

Simple Continuity Correction

Simple Continuity Correction to the Exact Method

Left-tailed test: P-value = $P(x \text{ or fewer}) - \frac{1}{2}P(\text{exactly } x)$

Right-tailed test: P-value = $P(x \text{ or more}) - \frac{1}{2}P(\text{exactly } x)$

Two-tailed test: *P*-value = twice the smaller of the preceding left-tailed and right-tailed values