Hypothesis Testing

- Developing Null and Alternative Hypotheses
- Type I and Type II Errors
- One-Tailed Tests About a Population Mean:
- Large-Sample Case
- Two-Tailed Tests About a Population Mean:
- Large-Sample Case
- Tests About a Population Mean:
- Small-Sample Case
- Tests About a Population Proportion

Developing Null and Alternative Hypotheses

- Hypothesis testing can be used to determine whether a statement about the value of a population parameter should or should not be rejected.
- The null hypothesis, denoted by H_0 , is a tentative assumption about a population parameter.
- The alternative hypothesis, denoted by H_a, is the opposite of what is stated in the null hypothesis.

Developing Null and Alternative Hypotheses

- **Testing Research Hypotheses**
  - Hypothesis testing is proof by contradiction.
  - The research hypothesis should be expressed as the alternative hypothesis.
  - The conclusion that the research hypothesis is true comes from sample data that contradict the null hypothesis.

Developing Null and Alternative Hypotheses

- **Testing the Validity of a Claim**
  - Manufacturers' claims are usually given the benefit of the doubt and stated as the null hypothesis.
  - The conclusion that the claim is false comes from sample data that contradict the null hypothesis.

Developing Null and Alternative Hypotheses

- **Testing in Decision-Making Situations**
  - A decision maker might have to choose between two courses of action, one associated with the null hypothesis and another associated with the alternative hypothesis.
  - Example: Accepting a shipment of goods from a supplier or returning the shipment of goods to the supplier.

A Summary of Forms for Null and Alternative Hypotheses about a Population Mean

- The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population mean μ must take one of the following three forms (where μ_0 is the hypothesized value of the population mean).

$H_0: \mu = \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$
One-tailed	One-tailed	Two-tailed

Example: Metro EMS

- **Null and Alternative Hypotheses**

A major west coast city provides one of the most comprehensive emergency medical services in the world. Operating in a multiple hospital system with approximately 20 mobile medical units, the service goal is to respond to medical emergencies with a mean time of 12 minutes or less.

The director of medical services wants to formulate a hypothesis test that could use a sample of emergency response times to determine whether or not the service goal of 12 minutes or less is being achieved.

Example: Metro EMS

- **Null and Alternative Hypotheses**

Hypotheses	Conclusion and Action
H_0: $\mu = 12$	The emergency service is meeting the response goal; no follow-up action is necessary.
H_a: $\mu > 12$	The emergency service is not meeting the response goal; appropriate follow-up action is necessary.

Where: μ = mean response time for the population of medical emergency requests.

Type I and Type II Errors

- Since hypothesis tests are based on sample data, we must allow for the possibility of errors.
- A Type I error is rejecting H_0 when it is true.
- The person conducting the hypothesis test specifies the maximum allowable probability of making a

  Type I error, denoted by α and called the level of significance.

Type I and Type II Errors

- A Type II error is accepting H_0 when it is false.
- Generally, we cannot control for the probability of making a Type II error, denoted by β.
- Statistician avoids the risk of making a Type II error by using "do not reject H_0" and not "accept H_0".

Example: Metro EMS

- **Type I and Type II Errors**

Conclusion	Population Condition: H_0 True ($\mu = 12$)	Population Condition: H_a True ($\mu > 12$)
Accept H_0 (Conclude $\mu = 12$)	Correct Conclusion	Type II Error
Reject H_0 (Conclude $\mu > 12$)	Type I Error	Correct Conclusion

Using the p-Value

- The p-value is the probability of obtaining a sample result that is at least as unlikely as what is observed.
- The p-value can be used to make the decision in a hypothesis test by noting that:
  - if the p-value is less than the level of significance α, the value of the test statistic is in the rejection region.
  - if the p-value is greater than or equal to α, the value of the test statistic is not in the rejection region.
- Reject H_0 if the p-value < α.

Steps of Hypothesis Testing

1. Determine the null and alternative hypotheses.
2. Specify the level of significance α.
3. Select the test statistic that will be used to test the hypothesis.

Using the p-Value

4. Collect the sample data and compute the value of the test statistic.
5. Use the value of the test statistic to compute the p-value.
6. Reject H_0 if p-value $< \alpha$.

One-Tailed Tests about a Population Mean: Large-Sample Case (n > 30)

- **Hypotheses**

$$H_0:\ \mu = \mu_0 \qquad H_a:\ \mu > \mu_0$$

or

$$H_0:\ \mu = \mu_0 \qquad H_a:\ \mu < \mu_0$$

- **Test Statistic**

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

$$z = \frac{\overline{x} - \mu_0}{s / \sqrt{n}}$$

Example: Metro EMS

- **One-Tailed Test about a Population Mean: Large n**

Let n = 40, $\bar{x}$ = 13.25 minutes, s = 3.2 minutes

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{13.25 - 12}{3.2 / \sqrt{40}} = 2.47$$

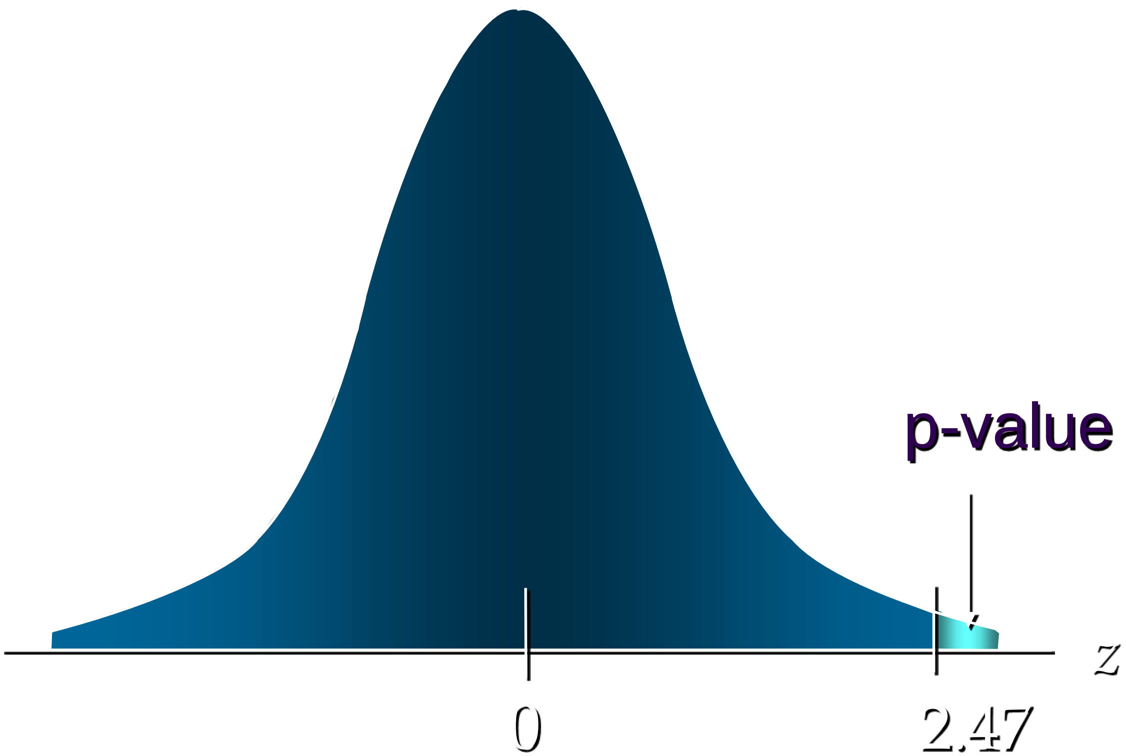
Then p-value = P(Z > 2.47) = .0068.

Since p-value = 0.0068 < 0.05 = α, we reject H_0.

Conclusion**:** We are 95% confident that Metro EMS is not meeting the response goal of 12 minutes; appropriate action should be taken to improve service.

Example: Metro EMS

p-value = P(Z > 2.47) = .0068



Two-Tailed Tests about a Population Mean: Large-Sample Case ($n \geq 30$)

- Hypotheses

$$H_0\colon\ \mu = \mu_0$$
$$H_a\colon\ \mu \neq \mu_0$$

- Test Statistic

σ Known

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

σ Unknown

$$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Example: Glow Toothpaste

- **Two-Tailed Tests about a Population Mean: Large n**

The production line for Glow toothpaste is designed to fill tubes of toothpaste with a mean weight of 6 ounces.

Periodically, a sample of 30 tubes will be selected in order to check the filling process. Quality assurance procedures call for the continuation of the filling process if the sample results are consistent with the assumption that the mean filling weight for the population of toothpaste tubes is 6 ounces; otherwise the filling process will be stopped and adjusted.

Example: Glow Toothpaste

- **Two-Tailed Tests about a Population Mean: Large n**

  A hypothesis test about the population mean can be used to help determine when the filling process should continue operating and when it should be stopped and corrected.

  - Hypotheses

$$H_0: \mu = 6$$

$$H_a: \mu \neq 6$$

Example: Glow Toothpaste

- **Two-Tailed Test about a Population Mean: Large n**

  Assume that a sample of 30 toothpaste tubes provides a sample mean of 6.1 ounces and standard deviation of 0.2 ounces.

  Let n = 30, $\bar{x}$ = 6.1 ounces, s = .2 ounces

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{6.1 - 6}{.2/\sqrt{30}} = 2.74$$

Example: Glow Toothpaste

- **Using the p-Value for a Two-Tailed Hypothesis Test**

Suppose we define the p-value for a two-tailed test as **double** the area found in the tail of the distribution.

With z = 2.74, the standard normal probability table shows there is a 1.000 - .9969 = .0031 probability of a difference larger than .1 in the upper tail of the distribution.

Considering the same probability of a larger difference in the lower tail of the distribution, we have

p-value = 2(.0031) = .0062

The p-value .0062 is less than α = .05, so H_0 is rejected.

Example: Glow Toothpaste

- **Two-Tailed Test about a Population Mean: Large n**

Conclusion: We are 95% confident that the mean filling weight of the toothpaste tubes is not 6 ounces. The filling process should be stopped and the filling mechanism adjusted.

Confidence Interval Approach to a Two-Tailed Test about a Population Mean

- Select a simple random sample from the population and use the value of the sample mean $\bar{x}$ to develop the confidence interval for the population mean μ.

- If the confidence interval contains the hypothesized value μ_0, do not reject H_0. Otherwise, reject H_0.

Example: Glow Toothpaste

- **Confidence Interval Approach to a Two-Tailed Hypothesis Test**

The 95% confidence interval for μ is

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 6.1 \pm 1.96(.2/\sqrt{30}) = 6.1 \pm .0716$$

or 6.0284 to 6.1716

Since the hypothesized value for the population mean, $\mu_0 = 6$, is not in this interval, the hypothesis-testing conclusion is that the null hypothesis,

H_0: $\mu = 6$, can be rejected.

Tests about a Population Mean: Small-Sample Case (n < 30)

- **Test Statistic**

σ Known

$$t = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

σ Unknown

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

This test statistic has a t distribution with n - 1 degrees of freedom.

p -Values and the t Distribution

- The format of the t distribution table provided in most statistics textbooks does not have sufficient detail to determine the exact p-value for a hypothesis test.
- However, we can still use the t distribution table to identify a range for the p-value.
- An advantage of computer software packages is that the computer output will provide the p-value for the t distribution.

Example: Highway Patrol

- **One-Tailed Test about a Population Mean: Small n**

A State Highway Patrol periodically samples vehicle speeds at various locations on a particular roadway. The sample of vehicle speeds is used to test the hypothesis

H_0: $\mu = 65$, H_a: $\mu = 65$.

The locations where H_0 is rejected are deemed the best locations for radar traps.

At Location F, a sample of 16 vehicles shows a mean speed of 68.2 mph with a standard deviation of 3.8 mph. Use an $\alpha = .05$ to test the hypothesis.

Example: Highway Patrol

- **One-Tailed Test about a Population Mean: Small n**

Let $n = 16$, $\bar{x} = 68.2$ mph, $s = 3.8$ mph

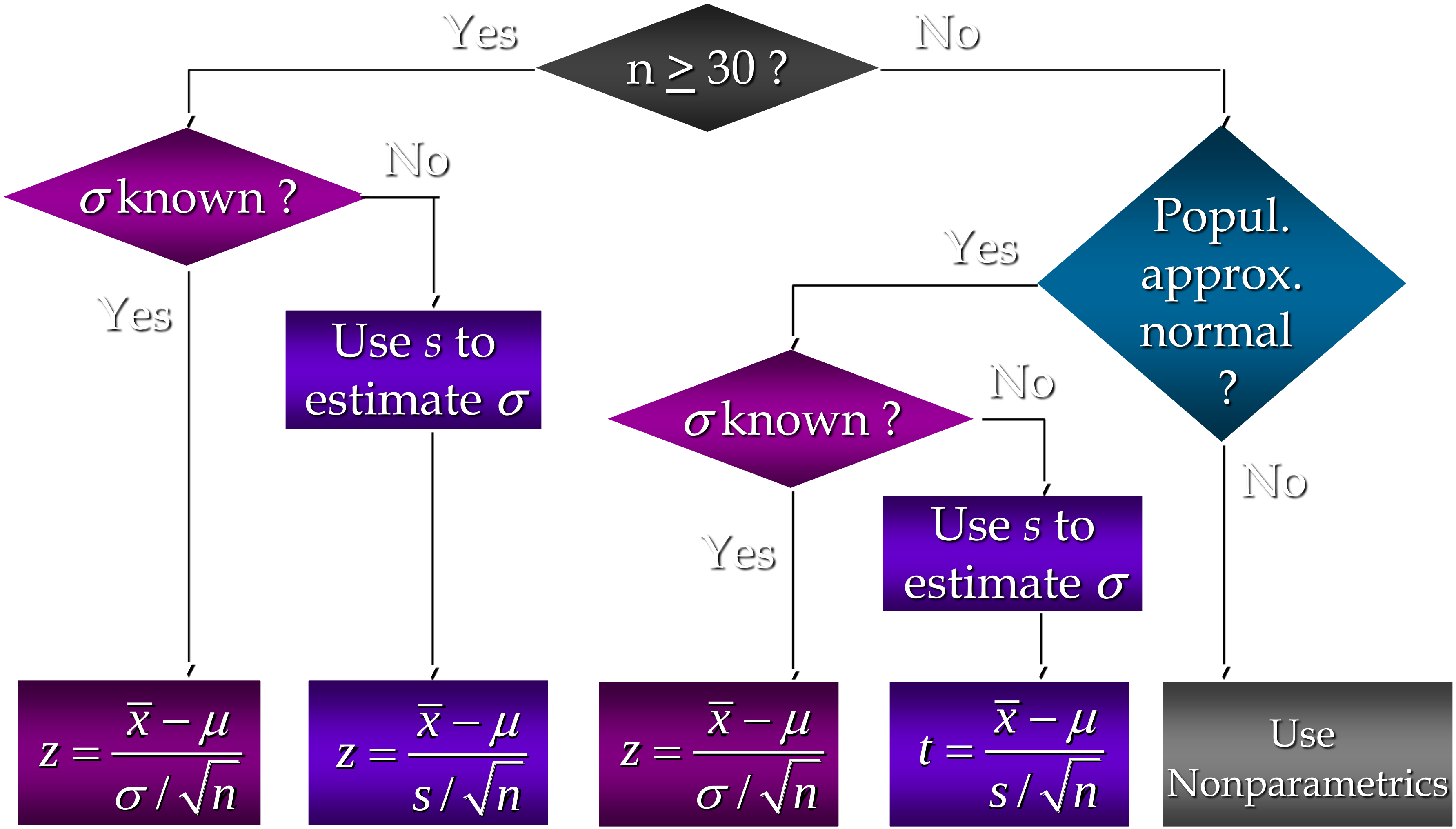
$\alpha = .05$, d.f. = 16-1 = 15, $t_a = 1.753$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{68.2 - 65}{3.8/\sqrt{16}} = 3.37$$

Since p-value = $P(t_{15} > 3.37) = .0021 < 0.05 = \alpha$, we reject H_0.

<u>Conclusion</u>: We are 95% confident that the mean speed of vehicles at Location F is greater than 65 mph. Location F is a good candidate for a radar trap.

Summary of Test Statistics to be Used in a Hypothesis Test about a Population Mean



A Summary of Forms for Null and Alternative Hypotheses about a Population Proportion

- The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population proportion π must take one of the following three forms (where π_0 is the hypothesized value of the population proportion).

One-tailed	One-tailed	Two-tailed
$H_0: \pi = \pi_0$	$H_0: \pi = \pi_0$	$H_0: \pi = \pi_0$
$H_a: \pi < \pi_0$	$H_a: \pi > \pi_0$	$H_a: \pi \neq \pi_0$

Tests about a Population Proportion: Large-Sample Case (np ≥ 5 and n(1 - p) ≥ 5)

- **Test Statistic**

$$z = \frac{(p - \pi_0)}{\sigma_p}$$

where:

$$\sigma_p = \sqrt{\frac{\pi_0(1 - \pi_0)}{n}}$$

Tests about a Population Proportion: Large-Sample Case ($np \geq 5$ and $n(1 - p) \geq 5$)

- **Desicion Rule**

H_a: $\pi < \pi_0$ Reject H_0 if p-value = $P(z < z_{cal}) < \alpha$

H_a: $\pi > \pi_0$ Reject H_0 if p-value = $P(z > z_{cal}) < \alpha$

H_a: $\pi \neq \pi_0$ Reject H_0 if p-value = $2P(z < |z_{cal}|) < \alpha$

Example: NSC

- **Two-Tailed Test about a Population Proportion: Large n**

For a Christmas and New Year's week, the National Safety Council estimated that 500 people would be killed and 25,000 injured on the nation's roads. The NSC claimed that 50% of the accidents would be caused by drunk driving.

A sample of 120 accidents showed that 67 were caused by drunk driving. Use these data to test the NSC's claim with $\alpha = 0.05$.

Example: NSC

- **Two-Tailed Test about a Population Proportion: Large n**
  - **Hypothesis**

$$H_0: \pi = .5$$
$$H_a: \pi \neq .5$$

  - **Test Statistic**

$$\sigma_p = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} = \sqrt{\frac{0.5(1-0.5)}{120}} = 0.045644$$

$$z = \frac{(p-\pi_0)}{\sigma_p} = \frac{\left(\frac{67}{120}\right) - 0.5}{0.045644} = 1.278$$

Example: NSC

- **Two-Tailed Test about a Population Proportion: Large n**
  - **Decision Rule**

    Reject H_0 if p-value < α

  - **Conclusion**

    Do not reject H_0. For z = 1.278, the p-value is .201.