Ch. 8 Hypothesis Testing

8.1 Foundations of Hypothesis Testing

Definitions

- In statistics, a hypothesis is a claim about a property of a population.
- A hypothesis test is a standard procedure for testing a claim about a population.
- A hypothesis test is often called a test of significance where we are testing whether or not a sample statistic is significantly different from some assumed value.

Examples

1. A survey showed that of 150 randomly selected commuters, 55% use cell-phones while they drive. We want to claim that most commuters use cell-phones when they drive.

   That is; we claim that $p > 0.50$, where p is the proportion of all commuters who use cell-phones while they drive.

2. In a sample of 100 healthy adults, the mean body temperature is 98.3 degrees with a standard deviation of 0.9 degrees. We want to claim that the population mean body temperature is less than the commonly accepted value of 98.6 degrees.

   That is, we claim that $\mu < 98.6$ where μ is the mean body temperature of all healthy adults.

Game Plan

1. **We will make an original assumption that the population has a given property.**

Example 1: $p = 0.5$ Example 2: $\mu = 98.6$

2. **We find the probability of observing the sample statistic.**

Example 1: $p = 0.55$ Example 2: $\bar{x} = 98.3$

3. **If the probability from part (2) is sufficiently small given our original assumption, we have evidence to reject that assumption and support a proposed alternate assumption.**

The Players: You must be able to express these in mathematical (symbolic) form

<u>The claim</u> is the statement that you want to support.

- Example 1: claim: $p > 0.5$
- Example 2: claim: $\mu < 98.6$

<u>The Null Hypothesis</u>, denoted $\mathbf{H_o}$, is the original assumption about the population parameter.

- Example 1: $H_o : p = 0.5$; Example 2: $H_o : \mu = 98.6$

Often (but not always) you want to reject this hypothesis.

<u>The Alternate Hypothesis</u>, denoted $\mathbf{H_1}$, is a different assumption about the parameter.

- Example 1: $H_1 : p > 0.5$ Example 2: $H_1\ \mu < 98.6$

Often (but not always) you want this hypothesis to support your claim.

Choosing the Null and Alternate Hypotheses

The key to conducting a successful hypothesis test is to start by choosing the null and alternative hypotheses appropriately.

1. Identify the claim you want to support and express it symbolically. This will determine the remaining hypotheses.
2. The null hypothesis (H_o) must impose equality (=).
3. The alternate hypothesis (H_1) must be different from H_o. If possible, have it match your claim.

The following table demonstrates some examples.

	Example 1	Example 2	Two-Tailed (A)	Two-Tailed (B)
Claim	$p > 0.5$	$\mu < 98.6$	$\mu = 12$	$\mu \neq 12$
H_o	$p = 0.5$	$\mu = 98.6$	$\mu = 12$	$\mu = 12$
H_1	$p > 0.5$	$\mu < 98.6$	$\mu \neq 12$	$\mu \neq 12$

Note: If the alternate hypothesis (H_1) contains a not equal symbol, the test is called two-tailed, otherwise it is called a one-tailed test (explained later).

Your Turn: Express each claim symbolically, then find H_o and H_1.

- Less than 10% of Americans use public transportation.

- The mean volume of fluid in all 12 ounce Coke cans is 12 ounces.

- Most people love pizza

Game Plan

1. **We will make an original assumption that the population has a given property.**

   This is the null hypothesis H_0

2. **We find the probability of observing the sample statistic.**

3. **If the probability from part (2) is sufficiently small given our original assumption, we have evidence to reject that assumption and support a proposed alternate assumption.**

   The alternate hypothesis is H_1

Your Turn: Express each claim symbolically, then find H_o and H_1.

- Less than 10% of Americans use public transportation.
  - H_0: $p = 0.10$
  - H_1: $p < 0.10$ (H_1 would support your claim)
- The mean volume of fluid in all 12 ounce Coke cans is 12 ounces.
  - H_0: $\mu = 12$ (H_0 would support your claim)
  - H_1: $\mu \neq 12$ (this is a 2-tailed test)

The Test Statistic

The test statistic is the value used in making a decision about the null hypothesis

For population parameters, p,

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

For population means, μ

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ if } \sigma \text{ is known}$$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \text{ if } \sigma \text{ is unknown}$$

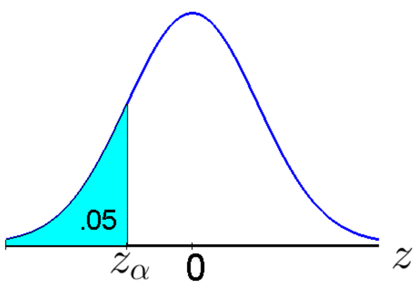
More definitions

- **Critical Region** (or rejection region) is the set of all values that cause us to reject the NULL hypothesis.
- **Significance level** (α) is the probability that the test statistic will fall in the rejection region even though the null hypothesis is true (causing you to make a mistake in your conclusion)
- **A critical value** is any value on the boundary of the rejection region (values denoted by z_{α} or $z_{\alpha/2}$ depending on if it is a 1 or 2-sided test or t_{α} or $t_{\alpha/2}$)

Critical Values - example

Using a significance level of = 0.05, find the critical z values (or t values if σ is unknown) for each of the following alternative hypotheses. Sketch the normal curve and shade the rejection region.

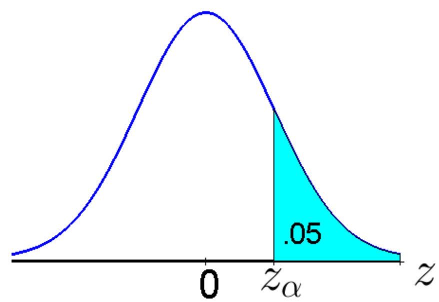
One sided test



$$H_1 : \mu < 98.6$$

$$z_\alpha = -1.645$$

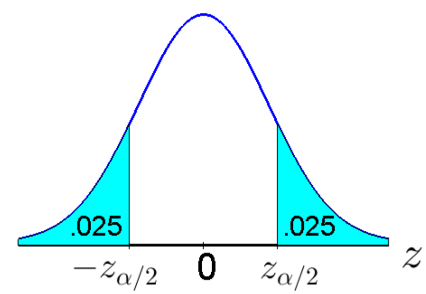
One sided test



$$H_1 : \mu > 98.6$$

$$z_\alpha = 1.645$$

Two-sided test



$$H_1 : \mu \neq 98.6$$

$$z_{\alpha/2} = \pm 1.96$$

Your turn

Using a significance level of $\alpha = 0.01$, find the critical z (or t) values for each of the following alternative hypotheses. Sketch the normal curve and shade the rejection region.

$H_1 : \mu < 98.6$ (left-tailed test)

$H_1 : \mu > 98.6$ (right-tailed test)

$H_1 : \mu \neq 98.6$ (two-tailed test)

P-value (probability value)

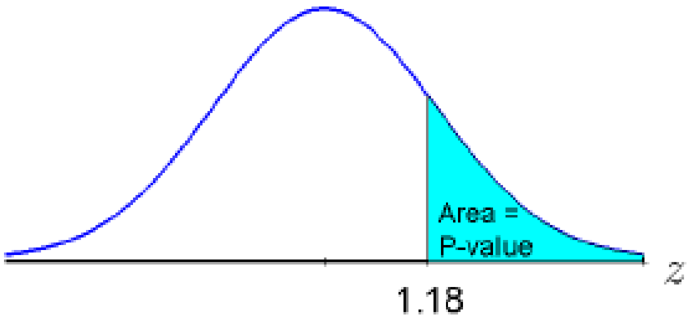
- The P-value of a test statistic is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data. (similar to the “unusually low” or “unusually high”)

Claim that p > 0.25 and the sample data produces a test statistic of z = 1.18.

$H_o : p = 0.25 \qquad H_1 : p > 0.25$

Right-Tailed Test

$$P\text{-value} = P(z > 1.18) = 1 - 0.8810 = \mathbf{0.119}$$

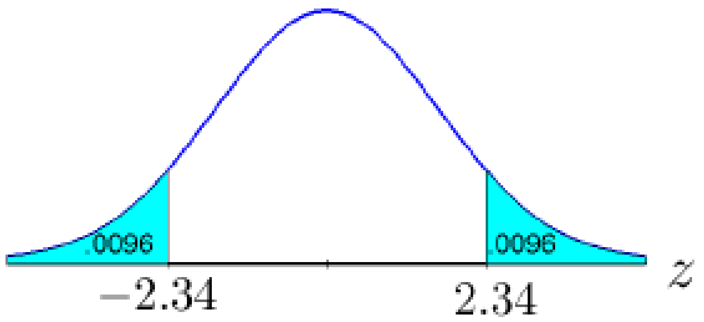


Claim that p ≠ 0.25 and the sample data produces a test statistic of z = 2.34.

$H_o : p = 0.25 \qquad H_1 : p \neq 0.25$

Two-Tailed Test

$$P\text{-value} = P(z < -2.34) + P(z > 2.34)$$
$$= 0.0096 + .0096 = \mathbf{0.0192}$$



Your turn:

Claim that $p < 0.25$ and the sample data produces a test statistic of $z = -1.85$.

Decisions: 2 methods

- Critical-Value Method:

  Reject H_o if the test statistic falls within the rejection region.

  Fail to reject H_o if the test statistic does not fall in the rejection region.

- P-value Method

  Reject H_o if the P-value $\leq \alpha$ (where α is the significance level).

  Fail to reject H_o if the P-value $> \alpha$.

Phrasing of Conclusions

- If the alternate hypothesis supports the original claim then if
  1. we reject the null hypotheses, we say
     The data supports the claim that ...
  2. we fail to reject the null hypothesis, we say
     There is not enough data to support the claim that ...
- If the null hypothesis supports the original claim then if
  1. we reject the null hypotheses, we say
     There is enough data to justify rejection of the claim that ....
  2. we fail to reject the null hypothesis, we say
     There is not enough data to justify rejection of the claim that ...

What we haven't done!

- If we fail to reject the null hypothesis, we **did not show that is was correct**. We merely did not have enough evidence to reject it. Therefore, we do not accept the null hypotheses.

- Courtroom analogy: If someone is not convicted of a crime, they are not found innocent by the jury, they are found not-guilty. In this analogy, the jury's null hypothesis is Ho: defendant is innocent. If there is enough evidence, the jury rejects this hypothesis and concludes the defendant is guilty. However, if there is insufficient data to reject the null hypothesis, then the defendant is found not-guilty. They did not conclude that the defendant was innocent. That's a different problem.

Summary

1. Define the claim, H_0, H_1
2. List information
3. Is this left, right, or two-tailed test?
4. Find the critical value and sketch critical region.
5. Make decision
    1. Based on critical region
    2. Based on p-value
6. Make concluding statement about the claim.

Try one:

- A farmer produces corn with an average of $\mu=30$ kernels per row on the cob. He has invented a new ``breed'' of corn he claims gives him a higher average number of kernels per row. Let X denote number of kernels per row on the new corn and suppose X is normal with $\sigma=1.35$ and μ unknown. Suppose that a sample of $n=50$ ears of the new corn yielded an average of 31.7 kernels per row. Test the farmer's claim to a significance level of 0.10.

Try another:

Test the following claim at the $\alpha = 0.05$ significance level.

A survey showed that of 150 randomly selected commuters, 40% have used a cell-phone while they drive.

We want to claim that most commuters DO NOT use cell-phones when they drive.

One more

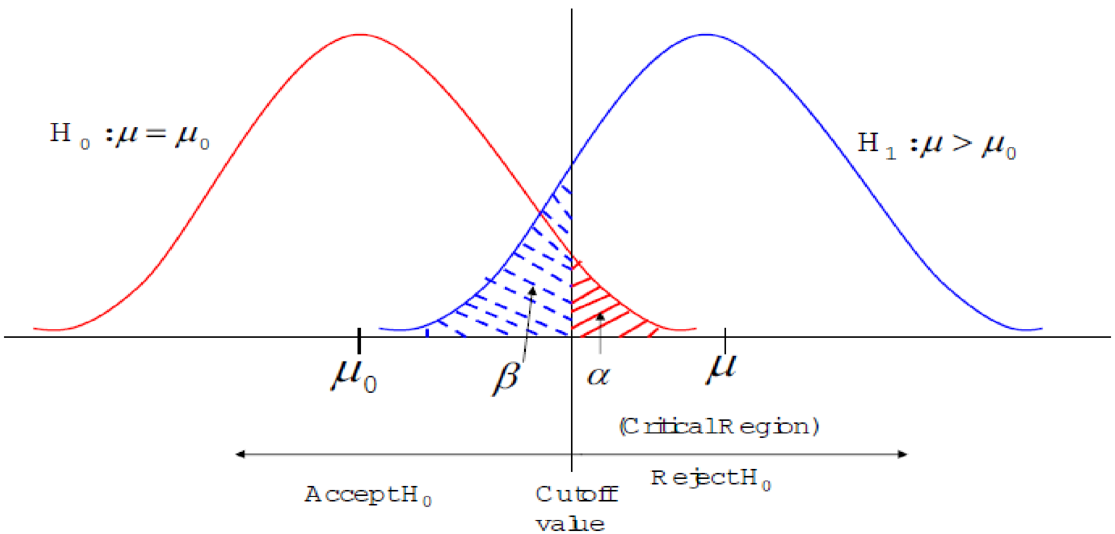
- The Spinach Society reported that only one-quarter of all people eat enough spinach. The Broccoli Bunch claims that this is not true. In a survey of 50 people 15 reported they ate enough spinach. Who is right? Test the claims to a significance level of 0.01.

Errors

- Type I error: Reject the null hypothesis when it is true.
- Type II error: Fail to reject the null hypothesis when it is false.

	Reject H_0	**Fail to reject H_0**
H_0 true	Type I error	Correct decision
H_0 false	Correct decision	Type II error

Errors



$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

P(Type I error) = α (significance level)
P(Type II error) = β

We call (1- β) the "power" of the test. It is the probability of (correctly) rejecting the null hypothesis when it is false.