Section 8.1-8.4 -- Hypothesis Testing

In statistics, a **hypothesis** is a claim or statement about a property of a population.

A **hypothesis test** is a procedure for testing the validity of a claim about a property of a population.

Null and Alternative Hypotheses...

The **null hypothesis**, H_0, is a statement that the value of a population parameter is *equal to* some claimed value.

- $H_0: p = 0.5$

The **alternative hypothesis**, H_1, is the statement that the parameter value somehow differs from that claimed by the null hypothesis.

- $H_1: p \neq 0.5$
- $H_1: p > 0.5$
- $H_1: p < 0.5$

Identifying the null and alternative hypotheses

1. Identify the specific claim to be tested and express it in symbolic form.
2. Write the symbolic form that must be true if the original claim is false.
3. Using the two symbolic expressions obtained so far
    - H_1 is the symbolic expression that does not contain the equality.
    - H_0 is the symbolic expression that the parameter equals the fixed value being considered.

See exercises 5-8 on page 371.

Computing Test Statistics...

A **test statistic** is a value used in making a decision about the null hypothesis.

- Test statistic for proportion: $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$
- Test statistic for mean (normal): $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$
- Test statistic for mean (Student t): $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$
- Test statistic for standard deviation: $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

Tools for Assessing the Test Statistic...

The **critical region** is the set of values of the test statistic that cause us to reject the null hypothesis.

The **significance level**, α, is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true. If the test statistic falls in the critical region, we reject the null hypothesis. α is the probability of making the mistake of rejecting the null hypothesis when it is true. The most common choice for α is 0.05 (corresponding to a confidence level of 95%).

A **critical value** is any value that separates the critical region from the values of the test statistic that lead to rejection of the null hypothesis. Critical values depend on the nature of the null hypothesis, the sampling distribution that applies, and the significance level.

The $\boldsymbol{P}$**-value** is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data.

- Critical region in left tail: P-value = area to the left of the test statistic
- Critical region in right tail: P-value = area to the right of the test statistic
- Critical region in two tails: P-value = twice the area beyond the test statistic

A **one-tailed test** indicates that the null hypothesis should be rejected when the test statistic is in the critical region *on one side* of the mean, depending on the direction of the inequality in the alternative hypothesis.

In a **two-tailed test**, the null hypothesis should be rejected when the test statistic is in the critical region *on either side* of the mean.

When performing a hypothesis test, our conclusion will always be one of the following:

1. Reject the null hypothesis.
2. Fail to reject the null hypothesis.

Errors in Hypothesis Tests...

- **Type I error:** The mistake of rejecting the null hypothesis when it is actually true. The symbol α is typically used to represent the probability of a type I error.
- **Type II error:** The mistake of failing to reject the null hypothesis when it is actually false. The symbol β is typically used to represent the probability of a type II error.

There are several methods for performing a hypothesis test. We will look at three methods.

Traditional Method

1. Identify the null and alternative hypotheses.
2. Select the significance level α based on the seriousness of a type I error. Make α small if the consequences of rejecting a true H_0 are severe. The values of 0.05 and 0.01 are very common.
3. Determine the appropriate test statistic and its sampling distribution.
4. Find the critical values and the critical region. Draw a graph illustrating the important values.
5. Reject H_o if the test statistic is in the critical region.
6. State your conclusion.

P-Value Method

1. Identify the null and alternative hypotheses.
2. Select the significance level α based on the seriousness of a type I error. Make α small if the consequences of rejecting a true H_0 are severe. The values of 0.05 and 0.01 are very common.
3. Determine the appropriate test statistic and its sampling distribution.
4. Find the P-value. Draw a graph illustrating the important values.
5. Reject H_o if the P-value is less than or equal to α.
6. State your conclusion.

Confidence Interval Method

1. Identify the null and alternative hypotheses.
2. Select the significance level α based on the seriousness of a type I error. Make α small if the consequences of rejecting a true H_0 are severe. The values of 0.05 and 0.01 are very common.
3. For a two-tailed test, construct a confidence interval with confidence level $1 - \alpha$. For a one-tailed test, construct a confidence interval with confidence level $1 - 2\alpha$.
4. Reject H_o if the population parameter has a value that is not included in the confidence interval.
5. State your conclusion.

Example

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 the level $\alpha = 0.05$?

Example

A telephone company representative estimates that 40% of its customers have call-waiting service. To test this hypothesis, she selected a sample of 100 customers and found that 37% had call waiting. At the level $\alpha = 0.01$, is there enough evidence to reject the claim?

Example

A medical investigation claims that the average number of infections per week at a certain hospital is 16.3. A random sample of 10 weeks had a mean number of 17.7 infections. The sample standard deviation is 1.8. Is there enough evidence to reject the investigator's claim at the level $\alpha = 0.05$?

Example

A cigarette manufacturer wishes to test the claim that the variance of the nicotine contents of its cigarettes is 0.644. Nicotine content is measured in milligrams, and assume that it is normally distributed. A random sample of 20 cigarettes has a standard deviation of 1.00 mg. Is there enough evidence to reject the manufacturer's claim at the level $\alpha = 0.05$?

Example

A hospital administrator believes that the standard deviation in the number of people using outpatient surgery per day is greater than 8. Assuming the numbers of people are normally distributed, a random sample of 15 days is selected. The data are shown below. Is there enough evidence to reject the administrator's claim at the level $\alpha = 0.10$?

25	30	5	15	18	42	16	9
10	12	12	38	8	14	27	