

Section 8.2 Notes
Testing a Proportion

Basic Definitions:

- In Statistics, a _____ is a claim or statement about a property of a population.
- A _____ (or test of significance) is a standard procedure for testing a claim about a property of a population.
- A sample proportion is said to be _____ if it isn't a reasonably likely outcome when the proposed standard is true.

COMPONENTS OF A FORMAL HYPOTHESIS TEST

#1: Give the name of the test and check the conditions for its use.

For proportions, three conditions must be met:

- The sample was a _____ from a binomial population
- Both np and $n(1-p)$ are at least _____
- The size of the population is at least _____ the size of the sample.

#2: State the hypotheses, defining any symbols.

Null and Alternative Hypotheses

- The _____ (denoted _____) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is _____ some claimed value.
 - We test the null hypothesis directly: we assume that it is true and reach a conclusion to either _____ or _____
- The _____ (denoted _____) is the statement that the parameter has a value that somehow differs from the null hypothesis (usually involves an inequality)

A few notes about H_0 and H_a :

- We conduct the hypothesis test by assuming the parameter is _____ some specific value so that we can work with a single distribution having a _____.
- If you are conducting a study and want to use a hypothesis test to _____ your claim, the claim must be worded so that it becomes the _____ hypothesis. You can never support a claim that some parameter is _____ some specified value.
- Note that the original statement could become the _____, it could become the _____, or it might not correspond exactly to either.

#3: Compute the test statistic, z , and find the critical value, z^* , and the P-value. Include a sketch that illustrates the situation.

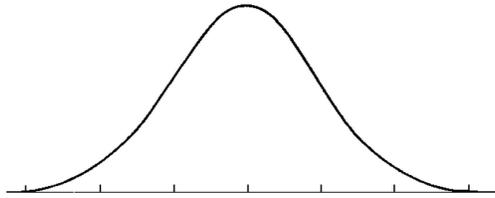
Test Statistic

- The _____ is a value used in making a decision about the null hypothesis, and it is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.
- Formula:

Critical Region, Significance Level, Critical Value, and P-Value

- The _____ (or rejection region) is the set of all values of the test statistic that causes us to reject the null hypothesis.
- The _____ (denoted by _____) 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 _____ the null hypothesis, so _____ is the probability of making the mistake of rejecting the null hypothesis when it is true.

- A _____ is any value that separates the critical region from the values of the test statistic that do not lead to rejection of the null hypothesis.



- The _____ is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true.

#4: Write a conclusion. (Two parts)

- Determine whether to reject or fail to reject the null hypothesis, linking the reason to the P-value or to the critical values.
- Tell what your conclusion means in the context of the situation.

Decisions and Conclusions

The decision to reject or fail to reject the null hypothesis can be made by using any of the following methods:

- _____
 - _____ if the test statistic falls within the critical region.
 - _____ if the test statistic does not fall within the critical region.
- _____
 - _____ if the p-value $\leq \alpha$ (where α is the significance level, such as 0.05)
 - _____ if the p-value $> \alpha$.

AP Exam-Hypothesis Test Rubric	
1) Hypothesis	<ul style="list-style-type: none"> • State the Null Hypothesis • State Alternative Hypothesis (with proper tails) • Use correct notation (define if non-standard)
2) Test	<ul style="list-style-type: none"> • Check assumptions and conditions (not just list) • Specify the model • Name the test
3) Mechanics	<ul style="list-style-type: none"> • Show work (statistics, values subbed into formula, shaded sketch of model, etc.) • Report test statistic (z, t, χ^2, df) • Report P-value
4) Conclusions	<ul style="list-style-type: none"> • State the decision (Reject, Fail to Reject) • With linkage to the P-value ("because the P-value is so low..." or something like that) • Interpret the result in context

Example:

Ships arriving in the US ports are inspected by Customs officials by contaminated cargo. Assume for a certain port, 20% of the ships arriving in the previous year contained cargo that was contaminated. A random selection of 50 ships in the current year included 5 that had contaminated cargo. Does the data suggest that the proportion of ships arriving in the port with contaminated cargoes has decreased in the current year? Use $\alpha = 0.01$.

Type I and Type II Errors

	The null hypothesis is true	The null hypothesis is false
We decide to reject the null hypothesis	Type I Error (rejecting a true null hypothesis) α	Correct Decision
We fail to reject the null hypothesis	Correct Decision	Type II Error (failing to reject a false null hypothesis), β

Example: Identify the type I and type II error that corresponds to the given hypothesis.
The proportion of LHS students that are seniors is 0.27.

Type I Error:

- To decrease the probability of a Type I Error, make _____. Changing the sample size has _____ on the probability of a Type I Error.
- If the null hypothesis is false, you can't make a Type I Error.

Type II Error:

- To decrease the probability of a Type II Error, take a _____ or make _____
- If the null hypothesis is true, you can't make a Type II Error.

Power:

- Power: the probability of rejecting a null hypothesis.
- When the null hypothesis is _____, you want to reject it and therefore you want the power to be _____
- To increase power, you can either take a _____ or make _____.

Section 8.3 Notes
CI for Difference of Two Proportions

Conditions that must be met:

- The two samples are taken _____ and _____ from two populations.
- Each population is at least _____ as large as its sample size
- _____, _____, _____, _____ are all at least _____.

CI for the difference $p_1 - p_2$:

$$\text{statistic} \pm \text{critical value} \cdot \text{standard deviation of the statistic}$$

Example:

Suppose the Cartoon Network conducts a nationwide survey to assess viewer attitudes toward Superman. Using a simple random sample, they select 400 boys and 300 girls to participate in the study. Forty percent of the boys say that Superman is their favorite character, compared to thirty percent of the girls. What is the 90% confidence interval for the true difference in attitudes toward Superman?

Section 8.4 Notes
Significance Test for Difference of Two Proportions

Conditions that must be met:

- The two samples are taken _____ and _____ from two populations
- Each population is at least _____ as large as its sample size
- _____, _____, _____, _____ are at least _____

H_0 and H_a :

Forms of H_0 : _____

Forms of H_a : _____

Test Statistic:

Where $\hat{p} = \frac{\text{total number of successes in both samples}}{n_1 + n_2}$

Example:

A survey was conducted of students from Cincinnati Public Schools system to determine if the incidence of hungry children was consistent in two schools located in lower-income areas. A random sample of 80 elementary students from school A found that 23% did not have breakfast before coming to school. A random sample of 180 elementary students from school B found that 7% did not have breakfast before coming in school.

Section 8.5 Notes
Inferences for Experiments

EXPERIMENTS

Conditions (for both CI and HT):

- The two treatments are _____ to the population of available experimental units.
- _____, _____, _____, _____ are all greater than _____

Example:

In 1954, the largest medical experiment of all times was carried out to test whether the newly developed Salk vaccine was effective in preventing polio. This study incorporated all three characteristics of an experiment: use of a control group of children who received a placebo injection, random assignment of children to either the placebo injection group or the Salk vaccine injection group, and assignment of each treatment to several hundred thousand children. Of the 200,745 children who received the Salk vaccine, 82 were diagnosed with polio. Of the 201,229 children who received the placebo, 162 were diagnosed with polio.

OBSERVATIONAL STUDY

- Treatments are NOT randomly assigned
- DO NOT allow you to make _____ conclusions or _____ in a rigorous way, but they can provide evidence of a possible association.
 - Why? _____

Example:

Researchers wanted to study the effects of exercising on dementia. They followed a group of people age 65 or older from 1994 to 2003. Among the 1185 free of dementia at the end of this time, 77% reported exercising three or more times a week. Among the 158 who showed signs of dementia at the end of the period, 67% reported exercising three or more times per week.