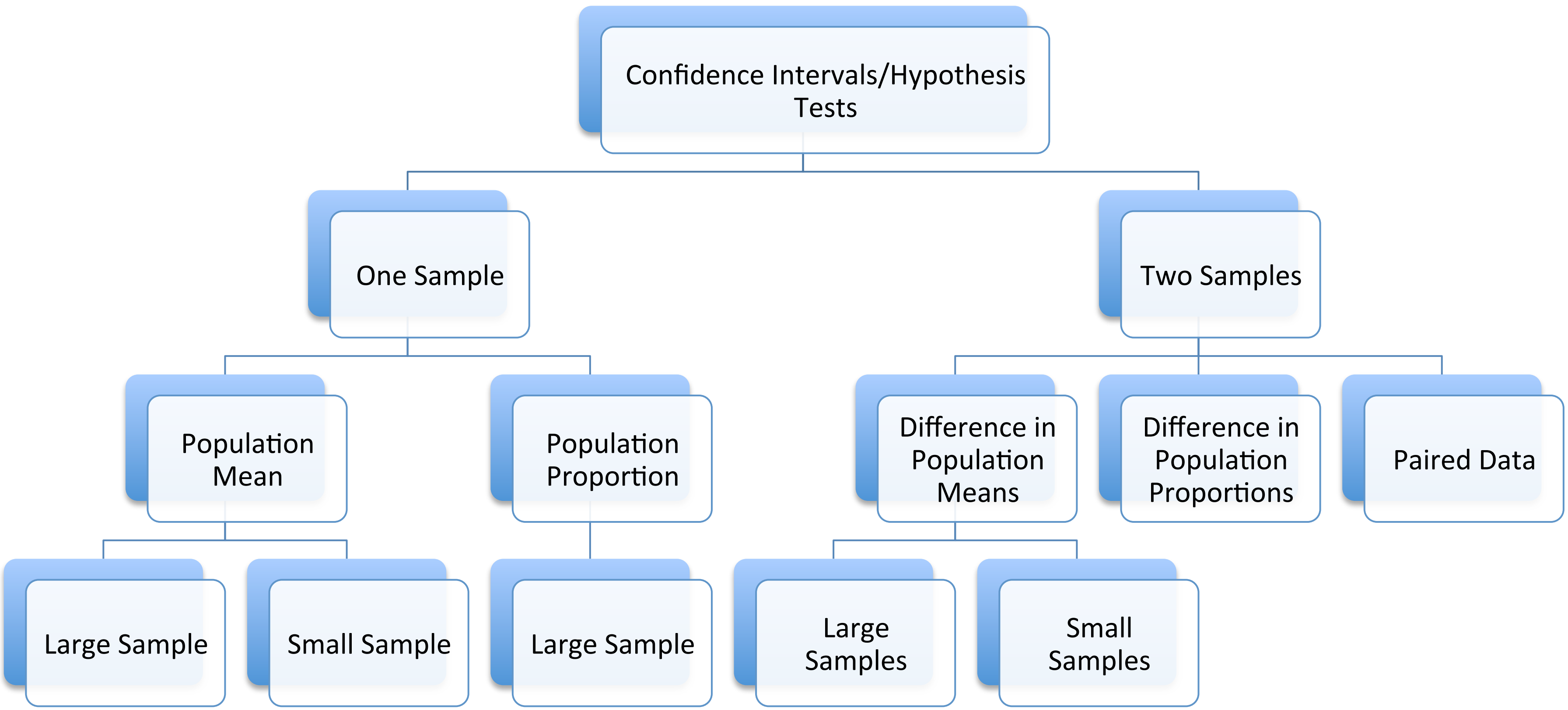
More Hypothesis Testing: Small-Sample Mean and Proportions

Keegan Korthauer
Department of Statistics
UW Madison

Outline


Confidence Intervals/Hypothesis Tests
One Sample
Two Samples
Population Mean
Population Proportion
Difference in Population Means
Difference in Population Proportions
Paired Data
Large Sample
Small Sample
Large Sample
Large Samples
Small Samples

Recap: 5 Steps to Perform a HT

1. Define H_0 and H_1
2. State the level of significance α
3. Construct the test statistic
4. Assume H_0 is true and evaluate the test statistic by finding the p-value
5. Make a conclusion based on the p-value

Use these steps in your HW and on the exam

General Form of Test Statistic

- Recall the general form for the CI:

$$\text{point estimate} \pm \text{critical value} \times \text{standard deviation}$$

- We can write the general form of a test statistic as:

$$\text{test statistic} = \frac{\text{point estimate} - \text{hypothesized value}}{\text{standard deviation}}$$

Recap: HT for Large-Sample Mean

Let $X_1,\ldots,X_n$ be a large ($n > 30$) sample from any population with mean μ and standard deviation σ. Approximate σ with *s* when unknown.

1. Set up the null H_0 and alternative H_1 hypotheses (see table below)
2. State the level of significance α you will use
3. Calculate the **z-score** (test statistic):

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

Use *s* when unknown!

4. Assume H_0 is true and calculate the P-value:

H_0	H_1	P-value
$\mu \le \mu_0$	$\mu > \mu_0$	Area to the right of *z*
$\mu \ge \mu_0$	$\mu < \mu_0$	Area to the left of *z*
$\mu = \mu_0$	$\mu \ne \mu_0$	Area to the left of -z plus area to the right of z

5. Make a conclusion based on the P-value

What About Small Samples?

- We just learned how to conduct a HT involving a sample mean based on a large sample
  - the test statistic is normally distributed by the CLT so the p-value is found by finding areas under the normal curve
- What if we want to test a hypothesis about a mean **based on a small sample?**
  - CLT no longer applies
  - if the population is approximately normal then the mean will be approximately normal
  - sample standard deviation s no longer approximates σ well

→ Make use of the t distribution!

Small-sample HT for Population Mean

- If the population is approximately normal, then

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}} \sim t_{n-1}$$

- Then we can use the following test statistic to measure the evidence against H_0: $\mu = \mu_0$ using the mean and standard deviation of our small sample of size n:

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

- Assuming H_0 is true, t will have a t distribution with n-1 degrees of freedom ➡ find p-value using areas under t curve

Small-Sample Tests for Population Mean

Let $X_1,\ldots,X_n$ be a small ($n < 30$) sample from a **normal** population with mean μ (unknown standard deviation σ).

If known, use the large-sample procedure

1. Set up the null H_0 and alternative H_1 hypotheses (see table below)
2. State the level of significance α you will use
3. Calculate the test statistic: $$t = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$$
4. Assume H_0 is true and calculate the P-value using areas under the t curve with n-1 degrees of freedom:

H_0	H_1	P-value
$\mu \leq \mu_0$	$\mu > \mu_0$	Area to the right of *t*
$\mu \geq \mu_0$	$\mu < \mu_0$	Area to the left of *t*
$\mu = \mu_0$	$\mu \neq \mu_0$	Area to the left of -t plus area to the right of t

5. Make a conclusion based on the P-value

Note on Calculating P-values

- Using the t-table, we can typically only say that the p-value is between two values
- Example: find the p-value for the two-sided test statistic $t = 2.245$ where the sample size is $n = 16$
  - Right-tail area for $t = 2.602$, $df = 15$ is 0.01
  - Right-tail area for $t = 2.131$, $df = 15$ is 0.025
  - The sum of the right and left tail areas for $t = 2.245$, $df = 15$ will be between 0.02 and 0.05
- On the exam, this would be the final answer for the p-value
- On the homework, use R to evaluate the p-value
  - In the example above use

```
2*pt(2.245, df=15, lower.tail=FALSE)
[1] 0.04027284
```

Example 6.7

- Before a substance can be deemed safe for landfilling, its chemical properties must be characterized
- A sample of 6 replicates of sludge from a New Hampshire wastewater treatment plant had a mean pH of 6.68 with a standard deviation of 0.20
- Assume the pH of all samples is approximately normally distributed
- Can we conclude that the mean pH is less than 7?

Test for a Population Proportion

- Say we counted X defective components in a sample of n from a large lot and are interested in whether we can conclude that the proportion of failures in the lot p is greater than some value p_0

- We would formulate the hypotheses as:

$$H_0: p \leq p_0$$

$$H_1: p > p_0$$

- What is the test statistic, and how do we find the p-value?

Recall: Normal Approximation to Binomial

- Recall that if X ~ Bin(n, p), then we can write X as a sum of independent and identically distributed RVs from a Bernoulli(p) population:

$$X = Y_1 + \ldots + Y_n$$

where $Y_1, \ldots, Y_n$ ~ Bern(p) (with mean p and variance p(1-p))

- Also note that $\hat{p} = \frac{X}{n} = \frac{Y_1 + \ldots + Y_n}{n} = \overline{Y}$

- Then by the CLT if n is large enough,

$$\hat{p} \sim N(p,\ p(1-p)/n) \text{ and } X \sim N(np,\ np(1-p))$$

(approximately)

Test Statistic for Population Proportion

- Then we can use the following test statistic to measure the evidence against H_0: $p \leq p_0$ using the sample proportion, sample size n, and hypothesized value p_0:

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

- Assuming H_0 is true, z will have a standard normal distribution (it is a z-score)

→ The p-value will correspond to areas under the standard normal curve

Tests for a Population Proportion

Let X be the number of successes in n independent Bernoulli trials, each with (unknown) success probability p

1. Set up the null H_0 and alternative H_1 hypotheses (see table below)
2. State the level of significance α you will use
3. Calculate the test statistic (z-score): $$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$
4. Assume H_0 is true and calculate the P-value using areas under the standard normal curve:

H_0	H_1	P-value
$p \le p_0$	$p > p_0$	Area to the right of *z*
$p \ge p_0$	$p < p_0$	Area to the left of *z*
$p = p_0$	$p \ne p_0$	Area to the left of -z plus area to the right of z

5. Make a conclusion based on the P-value

Note on Sample Size

- Since the normal approximation to the binomial relies on the CLT, we have to have some restriction on the sample size

- The HT given on the previous slide is valid when:

$$np_0 \geq 10 \text{ and } n(1-p_0) \geq 10$$

Example 6.6

- We are interested in a new method for measuring orthometric heights above sea level
- In a sample of 1225 baseline measurements, 926 gave results that were within the class C spirit leveling tolerance limits
- Can we conclude that this method produces results within the tolerance limits more than 75% of the time?

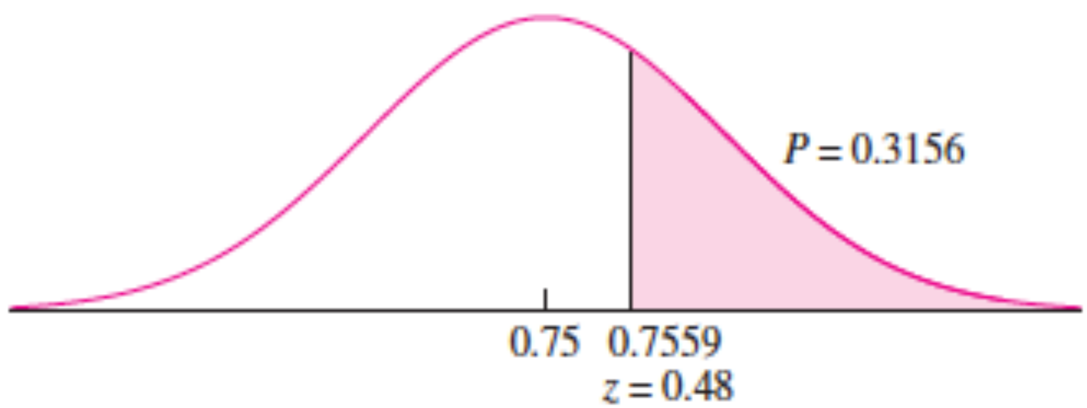


FIGURE 6.6 The null distribution of $\widehat{p}$ is $N(0.75,\ 0.0124^2)$. Thus if H_0 is true, the probability that $\widehat{p}$ takes on a value as extreme as or more extreme than the observed value of 0.7559 is 0.3156. This is the P-value.

Note on the HT:CI Relationship

- Recall that for large-samples, the values contained in a two-sided 100(1-α)% CI for a population mean are exactly those that have p-values greater than α in a two-sided HT
  - also for the one-sided case
- This relationship is also true for the population mean of small samples
- It is **not** true for the population proportion
  - The test statistic (which relies on larger sample size) does not exactly correspond to the CI for proportions for which we used the modern 'plus-four' method (which was valid for smaller samples)

MORE EXAMPLES

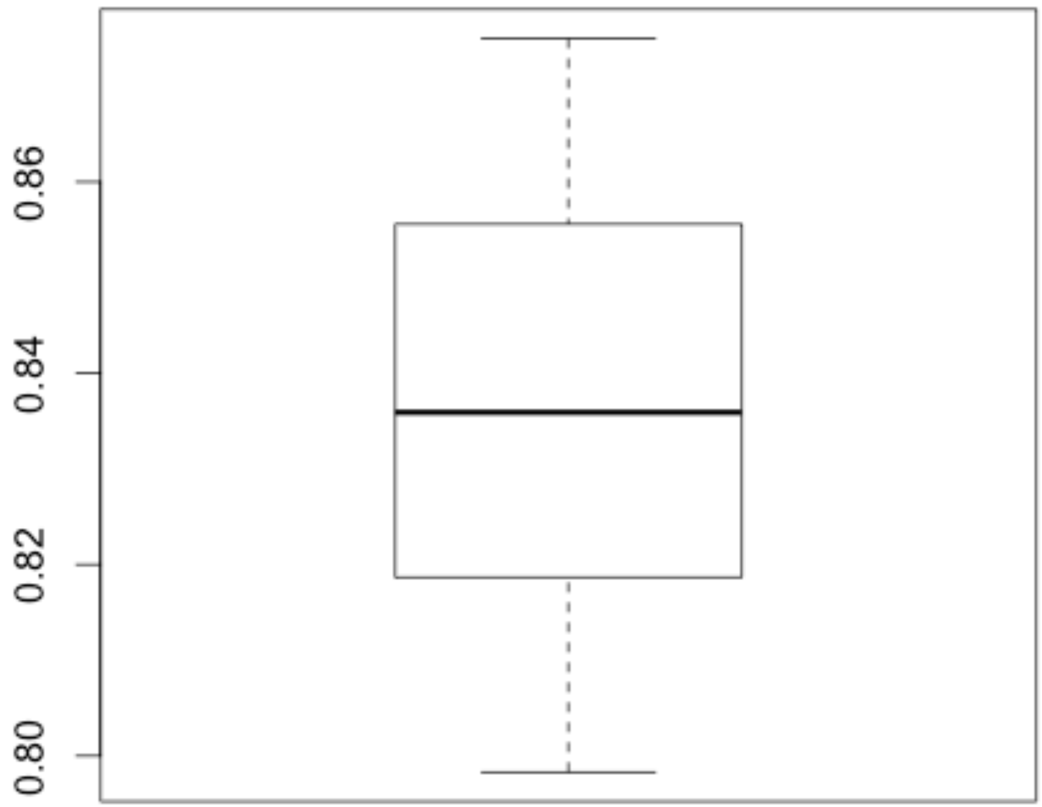
Example – Golf Club Design

- We are interested in the 'spring-like' quality of golf clubs, measured as the ratio of a ball's outgoing to ingoing velocities (called the **coefficient of restitution**)
- 15 drivers produced by a particular club maker were selected at random and their coefficients of restitution measured
- We want to determine if there is evidence (with α=0.01) to support a claim that the mean coefficient of restitution exceeds 0.82.
- The observations follow:

0.8411, 0.8580, 0.8042, 0.8191, 0.8532,

0.8730, 0.8182, 0.8483, 0.8282, 0.8125,

0.8276, 0.8359, 0.8750, 0.7983, 0.8660


0.80
0.82
0.84
0.86

Example – Automobile Engine Controller

- A semiconductor manufacturer produces controllers used in automobile engine applications
- The customer requires that the fraction of defective controllers be less than 0.05 and that the latter be demonstrated using a significance level of 0.05
- The manufacturer takes a random sample of 200 devices and finds 4 defective
- Will the customer be able to conclude that their requirements are met?

Next

- Two-sample tests for population means and proportions