Math 1710 Class 27

Cl’s and Hypothesis Testing
Dr. Back

Oct. 30, 2009
Smoke Detectors

CPSC ’96: 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector.
Smoke Detectors

**CPSC ’96:** 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector. Is this strong evidence the local rate is greater than the national rate?
Smoke Detectors

CPSC ’96: 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector. Is this strong evidence the local rate is greater than the national rate?

So \( \hat{p} = \frac{376}{400} = .94. \)
HT Vocabulary List

- Null Hypothesis $H$
- Alternative Hypothesis $H$
  (One Sided vs Two Sided)
- $SD(\hat{p})$ vs. $SE(\hat{p})$
- z-statistic
- Tail Probability
- P-value
- Significance Level
- Retain Null $H$ vs. Accept Alt $H$ (equiv. reject Null $H$)
Null Hypothesis $H_0$ - retained unless disproven.
Null Hypothesis $H_0$ - retained unless disproven. So when it is not clear which hypothesis is likely true, one retains the null.
Null Hypothesis $H_0$ - retained unless disproven. So when it is not clear which hypothesis is likely true, one retains the null. Our textbook will always write this as $p = p_0$. Here $p_0$ is some particular numerical value that we are interested in. Perhaps the historical value of the proportion we are now studying.
Null Hypothesis $H_0$ - retained unless disproven. So when it is not clear which hypothesis is likely true, one retains the null. Our textbook will always write this as $p = p_0$. Here $p_0$ is some particular numerical value that we are interested in. Perhaps the historical value of the proportion we are now studying. Or maybe a specification.
Null Hypothesis $H_0$ - retained unless disproven. So when it is not clear which hypothesis is likely true, one retains the null. Our textbook will always write this as $p = p_0$. Here $p_0$ is some particular numerical value that we are interested in. Perhaps the historical value of the proportion we are now studying. Or maybe a specification. In the 1-sided case, many would write $p \leq p_0$ or $p \geq p_0$ for $H_0$. 
Alternative Hypothesis $H_A$ - Only thing which possibly can be proven in the HT.
**Alternative Hypothesis** $H_A$ - Only thing which possibly can be proven in the HT.

*The only two-sided possibility is $p \neq p_0$.*

Typical verbal form: “strong evidence . . . of a change . . .”
Alternative Hypothesis $H_A$ - Only thing which possibly can be proven in the HT.

*The only two-sided possibility is $p \neq p_0$.*

Typical verbal form: “strong evidence . . . of a change . . .”

*Two one-sided possibilities:*

- $p < p_0$
- $p > p_0$

Typical verbal form: “strong evidence . . . of an increase . . .”
**Alternative Hypothesis** $H_A$ - Only thing which possibly can be proven in the HT.

*The only* two-sided possibility is $p \neq p_0$.

Typical verbal form: “strong evidence . . . of a change . . .”

Two one-sided possibilities:

- $p < p_0$
- $p > p_0$

Typical verbal form: “strong evidence . . . of an increase . . .”

So accepting the alternative means a statistically significant result; something was proven.

Retaining the null just means the evidence was not strong enough to convincingly prove the alternative.
**HT Vocabulary Details**

**z-statistic** - the z-score of $\hat{p}$ in relation to the sampling distribution of $\hat{p}$ assuming $H_0$ is true.
z-statistic - the z-score of $\hat{p}$ in relation to the sampling distribution of $\hat{p}$ assuming $H_0$ is true.
**HT Vocabulary Details**

**z-statistic** - the z-score of \( \hat{p} \) in relation to the sampling distribution of \( \hat{p} \) assuming \( H_0 \) is true.

\[
z = \frac{\hat{p} - p_0}{SD(\hat{p})}
\]
**HT Vocabulary Details**

**z-statistic** - the z-score of \( \hat{p} \) in relation to the sampling distribution of \( \hat{p} \) assuming \( H_0 \) is true.

\[
z = \frac{\hat{p} - p_0}{SD(\hat{p})}
\]

“…assuming \( H_0 \) is true” is why we use \( SD(\hat{p}) \) rather than \( SE(\hat{p}) \) in calculating our z-statistic.
Tail Probability - The probability of a $\hat{p}$ as extreme or more so and on the same side of $p_0$ than the $\hat{p}$ we actually observed. This is again assuming $H_0$ is true.
**Tail Probability** - The probability of a \( \hat{p} \) as extreme or more so and on the same side of \( p_0 \) than the \( \hat{p} \) we actually observed. This is again assuming \( H_0 \) is true.

(Tail probability shaded in gray.)
P-value: The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true.
P-value: The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true.

So P-value =

(single) tail probability in 1-sided test case
twice the (single) tail probability in 2-sided test case.
**P-value:** The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true. 

So \( P-value = \)  
(single) tail probability in 1-sided test case  
	twice the (single) tail probability in 2-sided test case. 

\[ H_a : p > p_0 \]  
P-value = tail probability
**P-value:** The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true.

So P-value =

*(single) tail probability in 1-sided test case twice the (single) tail probability in 2-sided test case.*

\[ H_a : p \neq p_0 \quad \text{P-value} = 2(\text{tail probability}) \]
HT Vocabulary Details

**P-value:** The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true.

**Significance Level $\alpha$:** The cutoff between ”small” and ”big” P-values.
**P-value:** The probability of a z-statistic as extreme as the given one or more so (and in support of the alternative hypothesis) under the assumption that the null hypothesis is true.

**Significance Level \( \alpha \):** The cutoff between "small" and "big" P-values.

Small P-value means reject the null.

Big P-value means retain the null.
Logic of Hypothesis Testing

Procedure:

1 ... 

2 ... 

3 Consider the sampling distribution of $\hat{p}$ which would hold if $H_0$ were true.

4 Retain $H_0$ if $\hat{p}$ is reasonably consistent with this sampling distribution. Otherwise reject $H_0$.

5 ...
Logic of Hypothesis Testing

Procedure:

1. . .
2. . .
3. Consider the sampling distribution of $\hat{p}$ which would hold if $H_0$ were true.
4. Retain $H_0$ if $\hat{p}$ is reasonably consistent with this sampling distribution. Otherwise reject $H_0$.
5. . .

Three ways to carry out step 4:

1. Calculate a z–statistic and determine a p-value.
2. Calculate a z-statistic and compare with a critical value $z^*$.
3. Use a confidence interval.
CPSC Example

Let $p$ be the true proportion of smoke detectors homes in our city have.
CPSC Example

Let $p$ be the true proportion of smoke detectors homes in our city have.

$H_0 : p = 0.9, \quad H_A : p > 0.9$
Let \( p \) be the true proportion of smoke detectors homes in our city have.

\[ H_0 : p = 0.9, \quad H_A : p > 0.9 \]

\[
SD(\hat{p}) = \sqrt{\frac{0.9 \cdot 0.1}{400}} = 0.015.
\]
Let $p$ be the true proportion of smoke detectors homes in our city have.

$H_0 : p = .9, \quad H_A : p > .9$

$$SD(\hat{p}) = \sqrt{\frac{.9 \cdot .1}{400}} = .015.$$

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$
z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

z-statistic is the z-score of \( \hat{p} \) on the samp dist centered at the hypothesized value of \( p \)
z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

z-statistic is the z-score of \( \hat{p} \) on the samp dist centered at the hypothesized value \( p_0 \) of \( p \).
CPSC Example

z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

Two primary ways to continue:
**CPSC Example**

- z-statistic
  \[ z = \frac{.94 - .9}{.015} = 2.67. \]

**Two primary ways to continue:**

**Method 1:** Tail Prob = \( P(\hat{p} > .94) = P(Z > 2.67) = .0038 \)

Since the test is 1-sided, P-value = tail probability = .0038.
CPSC Example

z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

Two primary ways to continue:
Method 1: Tail Prob = \( P(\hat{p} > .94) = P(Z > 2.67) = .0038 \)
Since the test is 1-sided, P-value=tail probability=.0038.
This is small, so conclusion is we reject \( H_0 \).
**CPSC Example**

z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

Two primary ways to continue:
Method 2: *Say significance level is* $\alpha = .05$. 
z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

Two primary ways to continue:
Method 2: Say significance level is \( \alpha = .05 \).
Since this is a 1-sided test with \( \alpha = .05 \), the appropriate critical values is \( z^* = 1.645 \).
CPSC Example

z-statistic

\[ z = \frac{.94 - .9}{.015} = 2.67. \]

Two primary ways to continue:
Method 2: Say significance level is \( \alpha = .05 \).
Since our z-statistic of 2.67 is more extreme than \( z^* = 1.645 \) (and supports \( H_a \)), we reject \( H_a \) at \( \alpha = .05 \).
Method 3: **Using a CI. Still $\alpha = .05$.**

The picture shows that a 90% CI has the same critical value as an $\alpha = .05$ HT.
Method 3: Using a CI. Still $\alpha = .05$.

\[
SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119
\]
Method 3: Using a CI. Still $\alpha = .05$.

$$SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119$$

A 90% CI is

$$.94 \pm 1.645 \cdot .0119 = .94 \pm .0196 = (.9204, .9596)$$
Method 3: Using a CI. Still $\alpha = .05$.

$$SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119$$

A 90% CI is

$$.94 \pm 1.645 \cdot .0119 = .94 \pm .0196 = (.9204, .9596)$$

Since all the values of $p$ in our CI support $H_a$, the result of the HT is to reject the null.

Had even one been consistent with $H_0$, we would have retained the null.
Method 3: Using a CI. Still $\alpha = .05$.

$$SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119$$

A 90% CI is

$$.94 \pm 1.645 \cdot .0119 = .94 \pm .0196 = (.9204, .9596)$$

Because of the extra error associated with SE instead of SD, Method 3 is not arithmetically equivalent to Methods 1 and 2. Method 3 could reject the null while method 1 retained or vice-versa.
CPSC Example

Method 3: Using a Cl. Still $\alpha = .05$.

$$SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119$$

A 90% Cl is

$$.94 \pm 1.645 \cdot .0119 = .94 \pm .0196 = (.9204, .9596)$$

Because of the extra error associated with SE instead of SD, Method 3 is not arithmetically equivalent to Methods 1 and 2. Method 3 could reject the null while method 1 retained or vice-versa.

This small discrepancy is not of practical importance. It could be corrected by a more elaborate Cl formula.
CPSC Example

Method 3: Using a CI. Still $\alpha = .05$.

$$SE(\hat{p}) = \sqrt{\frac{.94 \cdot .06}{400}} = .0119$$

A 90% CI is

$$.94 \pm 1.645 \cdot .0119 = .94 \pm .0196 = (.9204, .9596)$$

Because of the extra error associated with SE instead of SD, Method 3 is not arithmetically equivalent to Methods 1 and 2. Method 3 could reject the null while method 1 retained or vice-versa.

This small discrepancy is not of practical importance. It could be corrected by a more elaborate CI formula.

The general principle that many HT’s can be done via a CI is an important fact.
Three Kinds of $H'_a$s

$H_a$ for a prop. HT can be any of the three possibilities

- **Two-Sided** $p \neq p_0$
- **One-Sided** $p > p_0$
- **One-Sided** $p < p_0$
Three Kinds of $H_a's$

$H_a$ for a prop. HT can be any of the three possibilities

- **Two-Sided** $p \neq p_0$
- **One-Sided** $p > p_0$
- **One-Sided** $p < p_0$

One chooses among these based on the question being studied.
Three Kinds of $H_a$'s

$H_a$ for a prop. HT can be any of the three possibilities

- **Two-Sided** $p \neq p_0$
- **One-Sided** $p > p_0$
- **One-Sided** $p < p_0$

One chooses among these based on the question being studied.

A question like “Is there strong evidence that $p$ has changed . . .” would point to 2-sided.
Three Kinds of $H_a’s$

$H_a$ for a prop. HT can be any of the three possibilities

- **Two-Sided** $p \neq p_0$
- **One-Sided** $p > p_0$
- **One-Sided** $p < p_0$

One chooses among these based on the question being studied.

A question like “Is there strong evidence that $p$ has increased …” would point to 1-sided.
Three Kinds of $H_a'$s

$H_a$ for a prop. HT can be any of the three possibilities

**Two-Sided**  $p \neq p_0$

**One-Sided**  $p > p_0$

**One-Sided**  $p < p_0$

One chooses among these based on the question being studied.

The value of $\hat{p}$ never plays a role in formulating hypotheses!
Three Kinds of $H'_a$s

$N(0,1)$ and the critical value $z^*$
Three Kinds of $H_a's$

$H_a : p \neq p_0$ (2-sided), $\alpha = .10$

N(0,1)
Three Kinds of $H_a's$

$H_a : p > p_0$ (1-sided), $\alpha = .05$

$N(0,1)$
Three Kinds of $H_a's$

$N(0,1)$

$H_a : p < p_0 \text{ (1-sided)}, \alpha = .05$
Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:
Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

random sampling -
could be critical;
might be ok if "representative"
representative hard/impossible to define
Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

**plausible independence** -
could be critical
sometimes just a working hypothesis
Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

10% condition - results in overestimation of samp. dist. st dev gradual breakdown in formulas, not method
Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

success/failure -
progressive reduction of accuracy
accuracy varies regardless for smaller values of n
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

\[ MOE = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} \]
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

\[ MOE = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} \]

Margin of error picture.
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

\[ MOE = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} \]

Algebra gives

\[ n = \left( \frac{z^*}{MOE} \right)^2 \hat{p}\hat{q} \]
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Algebra gives

\[ n = \left( \frac{z^*}{\text{MOE}} \right)^2 \hat{p}\hat{q} \]

Here we have

\[ n = \left( \frac{1.96}{.001} \right)^2 \hat{p}\hat{q} = 1960^2 \hat{p}\hat{q}. \]
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Algebra gives

$$n = \left( \frac{z^*}{\text{MOE}} \right)^2 \hat{p}\hat{q}$$

Here we have

$$n = \left( \frac{1.96}{.001} \right)^2 \hat{p}\hat{q} = 1960^2 \hat{p}\hat{q}.$$  

The challenge is we don’t know $\hat{p}$ until we conduct the sample.
Sample Size for a Given MOE

We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Algebra gives

\[ n = \left( \frac{z^*}{\text{MOE}} \right)^2 \hat{p}\hat{q} \]

Here we have

\[ n = \left( \frac{1.96}{.001} \right)^2 \hat{p}\hat{q} = 1960^2 \hat{p}\hat{q}. \]

Three Approaches:
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Algebra gives

$$n = \left( \frac{z^*}{\text{MOE}} \right)^2 \hat{p}\hat{q}$$

Here we have

$$n = \left( \frac{1.96}{.001} \right)^2 \hat{p}\hat{q} = 1960^2 \hat{p}\hat{q}.$$ 

Three Approaches:

1. **Conservative**: Just use $\hat{p} = .5$ since that gives the biggest possible value of $\hat{p}\hat{q}$.

(Think about the graph of $y = x(1-x)$.)
Sample Size for a Given MOE

We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Here we have

\[ n = \left( \frac{1.96}{.001} \right)^2 \hat{p} \hat{q} = 1960^2 \hat{p} \hat{q}. \]

Three Approaches:
1. Conservative: Just use \( \hat{p} = .5 \) since that gives the biggest possible value of \( \hat{p} \hat{q} \).

Here we get

\[ n = 1960^2(.5)(.5) = 980^2 = 960,400 \]

which explains why surveys DO NOT seek an MOE of .1%.
We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use? Here we have

\[ n = \left( \frac{1.96}{0.001} \right)^2 \hat{p} \hat{q} = 1960^2 \hat{p} \hat{q}. \]

Three Approaches:
1. **Conservative:** Just use \( \hat{p} = .5 \) since that gives the biggest possible value of \( \hat{p} \hat{q} \).
Sample Size for a Given MOE

We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Here we have

\[ n = \left( \frac{1.96}{.001} \right)^2 \hat{p}\hat{q} = 1960^2 \hat{p}\hat{q}. \]

Three Approaches:

1. **Conservative**: Just use \( \hat{p} = .5 \) since that gives the biggest possible value of \( \hat{p}\hat{q} \).
2. **Use an estimate of \( \hat{p} \):**
Sample Size for a Given MOE

We wish to produce a 95% CI with a margin of error of .1%. What sample size should we use?

Here we have

\[ n = \left( \frac{1.96}{.001} \right)^2 \hat{p} \hat{q} = 1960^2 \hat{p} \hat{q}. \]

Three Approaches:
1. **Conservative**: Just use \( \hat{p} = .5 \) since that gives the biggest possible value of \( \hat{p} \hat{q} \).
2. Use an estimate of \( \hat{p} \):
3. If you expect \( \hat{p} \) to be in a range, use the most demanding \( \hat{p} \) within that range.