Introduction to hypothesis testing

Chrysafis Vogiatzis

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

Lectures 24-25

Hypothesis testing

1. 94% of UIUC's College of Engineering graduates secure employment or go to graduate school within a year of graduation.
2. The average starting salary for these Engineering graduates is $78,159.
3. Electrical Engineering or Construction Management? Electrical engineers earn more in the start of their careers.
4. Electrical Engineering or Construction Management? The top 10% construction management professionals earn more than the top 10% electrical engineering professionals.
5. The majority of customers prefers Coke to Pepsi.
6. People with a dog in the house live longer.

All of the above are *claims* waiting to be *tested* and eventually *rejected* or *accepted*.

ILLINOIS

Hypothesis testing

1. 94% of UIUC's College of Engineering graduates secure employment or go to graduate school within a year of graduation.
2. The average starting salary for these Engineering graduates is $78,159.
3. Electrical Engineering or Construction Management? Electrical engineers earn more in the start of their careers.
4. Electrical Engineering or Construction Management? The top 10% construction management professionals earn more than the top 10% electrical engineering professionals.
5. The majority of customers prefers Coke to Pepsi.
6. People with a dog in the house live longer.

All of the above are *claims* waiting to be *tested* and eventually *rejected* or *accepted*.

Formulating a hypothesis

Formally, we define:

- **Statistical hypothesis**: a claim about the unknown parameters or distributions of a population.
  - The mean grade of a student in a class is a B+.
  - The proportion of students that end up with an A in a class is 25%.
  - The grade of a student in a class is normally distributed.
- **Null hypothesis**, H_0: the hypothesis/claim that is being tested.
- **Alternative hypothesis**, H_1: the opposite of or simply an alternative to the hypothesis/claim.

Or in statistical terms:

$$H_0 : \text{null hypothesis}$$
$$H_1 : \text{alternative hypothesis}$$

Formulating a hypothesis

Formally, we define:

- **Statistical hypothesis**: a claim about the unknown parameters or distributions of a population.
  - The mean grade of a student in a class is a B+.
  - The proportion of students that end up with an A in a class is 25%.
  - The grade of a student in a class is normally distributed.
- **Null hypothesis**, H_0: the hypothesis/claim that is being tested.
- **Alternative hypothesis**, H_1: the opposite of or simply an alternative to the hypothesis/claim.

Or in statistical terms:

$$H_0 : \text{null hypothesis}$$
$$H_1 : \text{alternative hypothesis}$$

ILLINOIS

Formulating a hypothesis

Formally, we define:

- **Statistical hypothesis**: a claim about the unknown parameters or distributions of a population.
  - The mean grade of a student in a class is a B+.
  - The proportion of students that end up with an A in a class is 25%.
  - The grade of a student in a class is normally distributed.
- **Null hypothesis**, H_0: the hypothesis/claim that is being tested.
- **Alternative hypothesis**, H_1: the opposite of or simply an alternative to the hypothesis/claim.

Or in statistical terms:

$$H_0 : \text{null hypothesis}$$
$$H_1 : \text{alternative hypothesis}$$

ILLINOIS

Formulating a hypothesis

Formally, we define:

- **Statistical hypothesis**: a claim about the unknown parameters or distributions of a population.
  - The mean grade of a student in a class is a B+.
  - The proportion of students that end up with an A in a class is 25%.
  - The grade of a student in a class is normally distributed.
- **Null hypothesis**, H_0: the hypothesis/claim that is being tested.
- **Alternative hypothesis**, H_1: the opposite of or simply an alternative to the hypothesis/claim.

Or in statistical terms:

$$H_0 : \text{null hypothesis}$$
$$H_1 : \text{alternative hypothesis}$$

ILLINOIS

Formulating a hypothesis

Formally, we define:

- **Statistical hypothesis**: a claim about the unknown parameters or distributions of a population.
  - The mean grade of a student in a class is a B+.
  - The proportion of students that end up with an A in a class is 25%.
  - The grade of a student in a class is normally distributed.
- **Null hypothesis**, H_0: the hypothesis/claim that is being tested.
- **Alternative hypothesis**, H_1: the opposite of or simply an alternative to the hypothesis/claim.

Or in statistical terms:

$$H_0 : \text{null hypothesis}$$
$$H_1 : \text{alternative hypothesis}$$

Examples

"The average grade of a student in a class is 84%."

Null hypothesis:

- $H_0 : \mu = 84\%$.

Alternative hypothesis:

- $H_1 : \mu \neq 84\%$.

"More than half of the population prefers Coke to Pepsi."

Null hypothesis:

- $H_0 : p = 0.5$.

Alternative hypothesis:

- $H_1 : p < 0.5$.

"There is no life expectancy change by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 \neq 0$.

"There is no life expectancy increase by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 > 0$.

ILLINOIS

Examples

"The average grade of a student in a class is 84%."

Null hypothesis:

- $H_0 : \mu = 84\%$.

Alternative hypothesis:

- $H_1 : \mu \neq 84\%$.

"More than half of the population prefers Coke to Pepsi."

Null hypothesis:

- $H_0 : p = 0.5$.

Alternative hypothesis:

- $H_1 : p < 0.5$.

"There is no life expectancy change by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 \neq 0$.

"There is no life expectancy increase by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 > 0$.

ILLINOIS

Examples

"The average grade of a student in a class is 84%."

Null hypothesis:

- $H_0 : \mu = 84\%$.

Alternative hypothesis:

- $H_1 : \mu \neq 84\%$.

"More than half of the population prefers Coke to Pepsi."

Null hypothesis:

- $H_0 : p = 0.5$.

Alternative hypothesis:

- $H_1 : p < 0.5$.

"There is no life expectancy change by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 \neq 0$.

"There is no life expectancy increase by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 > 0$.

Examples

"The average grade of a student in a class is 84%."

Null hypothesis:

- $H_0 : \mu = 84\%.$

Alternative hypothesis:

- $H_1 : \mu \neq 84\%.$

"More than half of the population prefers Coke to Pepsi."

Null hypothesis:

- $H_0 : p = 0.5.$

Alternative hypothesis:

- $H_1 : p < 0.5.$

"There is no life expectancy change by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0.$

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 \neq 0.$

"There is no life expectancy increase by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0.$

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 > 0.$

ILLINOIS

Examples

"The average grade of a student in a class is 84%."

Null hypothesis:

- $H_0 : \mu = 84\%$.

Alternative hypothesis:

- $H_1 : \mu \neq 84\%$.

"More than half of the population prefers Coke to Pepsi."

Null hypothesis:

- $H_0 : p = 0.5$.

Alternative hypothesis:

- $H_1 : p < 0.5$.

"There is no life expectancy change by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 \neq 0$.

"There is no life expectancy increase by eating vegetables."

Null hypothesis:

- $H_0 : \mu_1 - \mu_2 = 0$.

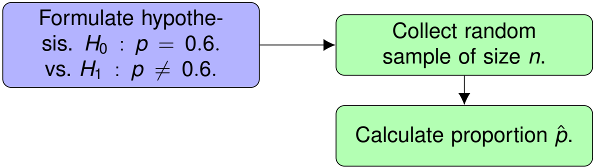
Alternative hypothesis:

- $H_1 : \mu_1 - \mu_2 > 0$.

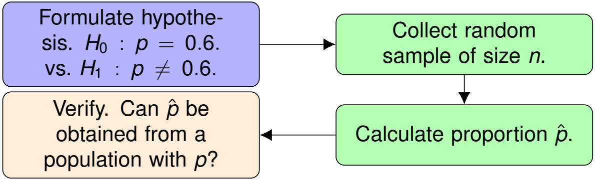
Performing a hypothesis test

Formulate hypothesis. $H_0 : p = 0.6$. vs. $H_1 : p \neq 0.6$.

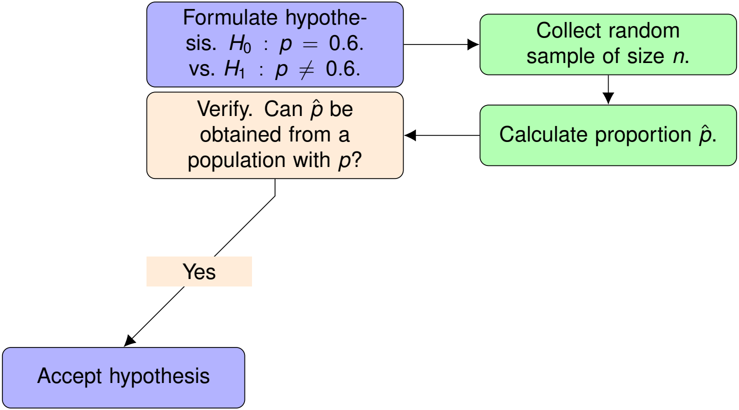
Performing a hypothesis test



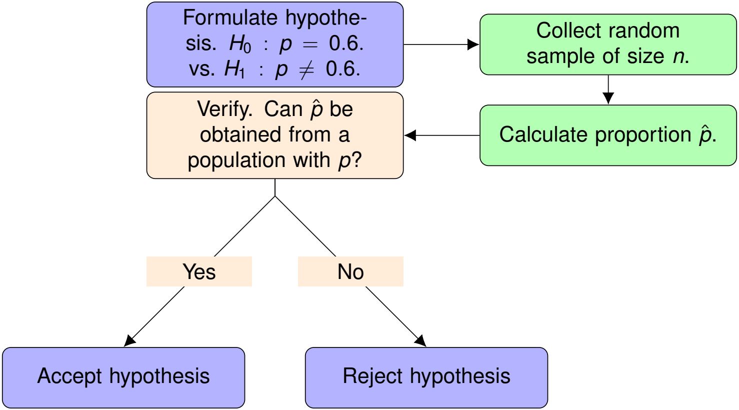
Performing a hypothesis test



Performing a hypothesis test



Performing a hypothesis test



Outcomes of a hypothesis test

Accepting and **rejecting** a hypothesis are not the correct terms for the outcomes of a hypothesis test. Instead, we say:

- **Reject** the hypothesis.
  - **Strong conclusion**.
  - Implies the existence of sufficient evidence agains the hypothesis.
  - We are quite certain that H_0 is wrong.
- **Fail to reject** the hypothesis.
  - **Weak conclusion**.
  - Implies the lack of sufficient evidence agains the hypothesis.
  - It does not mean that H_0 is true! It merely implies that we are uncertain about either H_0 or H_1 being true.

Can you think of a parallel to this thought process? How about a trial and the fact that a defendant is presumed not guilty?

Outcomes of a hypothesis test

Accepting and **rejecting** a hypothesis are not the correct terms for the outcomes of a hypothesis test. Instead, we say:

- **Reject** the hypothesis.
  - **Strong conclusion**.
  - Implies the existence of sufficient evidence agains the hypothesis.
  - We are quite certain that H_0 is wrong.
- **Fail to reject** the hypothesis.
  - **Weak conclusion**.
  - Implies the lack of sufficient evidence agains the hypothesis.
  - It does not mean that H_0 is true! It merely implies that we are uncertain about either H_0 or H_1 being true.

Can you think of a parallel to this thought process? How about a trial and the fact that a defendant is presumed not guilty?

ILLINOIS

Outcomes of a hypothesis test

Accepting and **rejecting** a hypothesis are not the correct terms for the outcomes of a hypothesis test. Instead, we say:

- **Reject** the hypothesis.
  - **Strong conclusion**.
  - Implies the existence of sufficient evidence agains the hypothesis.
  - We are quite certain that H_0 is wrong.
- **Fail to reject** the hypothesis.
  - **Weak conclusion**.
  - Implies the lack of sufficient evidence agains the hypothesis.
  - It does not mean that H_0 is true! It merely implies that we are uncertain about either H_0 or H_1 being true.

Can you think of a parallel to this thought process? How about a trial and the fact that a defendant is presumed not guilty?

Outcomes of a hypothesis test

Accepting and **rejecting** a hypothesis are not the correct terms for the outcomes of a hypothesis test. Instead, we say:

- **Reject** the hypothesis.
  - **Strong conclusion**.
  - Implies the existence of sufficient evidence agains the hypothesis.
  - We are quite certain that H_0 is wrong.
- **Fail to reject** the hypothesis.
  - **Weak conclusion**.
  - Implies the lack of sufficient evidence agains the hypothesis.
  - It does not mean that H_0 is true! It merely implies that we are uncertain about either H_0 or H_1 being true.

Can you think of a parallel to this thought process? How about a trial and the fact that a defendant is presumed not guilty?

Outcomes of a hypothesis test

Accepting and **rejecting** a hypothesis are not the correct terms for the outcomes of a hypothesis test. Instead, we say:

- **Reject** the hypothesis.
  - **Strong conclusion**.
  - Implies the existence of sufficient evidence agains the hypothesis.
  - We are quite certain that H_0 is wrong.
- **Fail to reject** the hypothesis.
  - **Weak conclusion**.
  - Implies the lack of sufficient evidence agains the hypothesis.
  - It does not mean that H_0 is true! It merely implies that we are uncertain about either H_0 or H_1 being true.

Can you think of a parallel to this thought process? How about a trial and the fact that a defendant is presumed not guilty?

Errors

Decision	H_0 is true	H_0 is false
Reject H_0	incorrect decision	correct decision
Fail to reject H_0	correct decision	incorrect decision

Type I error: $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha$,

Type II error: $P(\text{fail to reject } H_0 | H_0 \text{ is not true}) = \beta$.

- $1 - \alpha$ is also the significance of the test.
- $1 - \beta$ is also the power of the test.

Errors

Decision	H_0 is true	H_0 is false
Reject H_0	incorrect decision	correct decision
Fail to reject H_0	correct decision	incorrect decision

Type I error: $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha$,

Type II error: $P(\text{fail to reject } H_0 | H_0 \text{ is not true}) = \beta$.

- $1 - \alpha$ is also the significance of the test.
- $1 - \beta$ is also the power of the test.

Errors

Decision	H_0 is true	H_0 is false
Reject H_0	incorrect decision	correct decision
Fail to reject H_0	correct decision	incorrect decision

Type I error: $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha$,

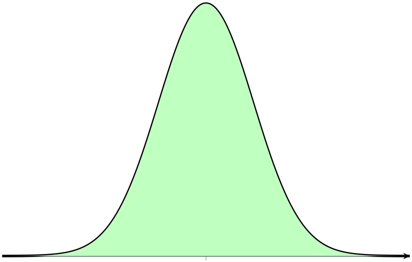
Type II error: $P(\text{fail to reject } H_0 | H_0 \text{ is not true}) = \beta$.

- $1 - \alpha$ is also the significance of the test.
- $1 - \beta$ is also the power of the test.

ILLINOIS

Hypothesis testing for proportions

First, assume that $H_0 : p = p_0$ is true. Then, the observed proportion $\hat{p}$ is distributed as $\mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.



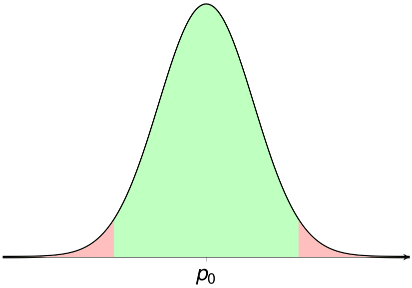
Additionally, we can put α in the mix.

The limits can be found as $p_0 \pm z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$.

So how to reject or fail to reject? See where $\hat{p}$ is!

Hypothesis testing for proportions

First, assume that $H_0 : p = p_0$ is true. Then, the observed proportion $\hat{p}$ is distributed as $\mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.



Additionally, we can put α in the mix.

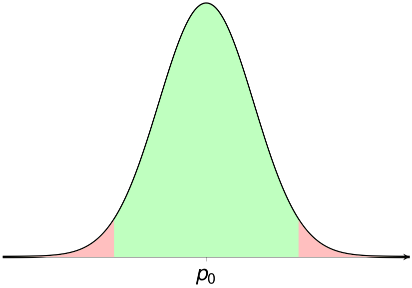
The limits can be found as $p_0 \pm z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$.

So how to reject or fail to reject? See where $\hat{p}$ is!

ILLINOIS

Hypothesis testing for proportions

First, assume that $H_0 : p = p_0$ is true. Then, the observed proportion $\hat{p}$ is distributed as $\mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.



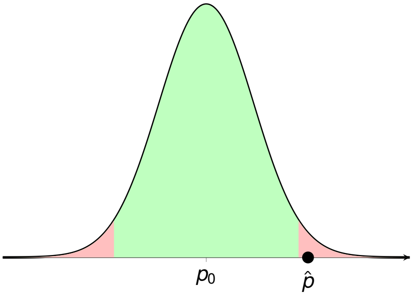
Additionally, we can put α in the mix.

The limits can be found as $p_0 \pm z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$.

So how to reject or fail to reject? See where $\hat{p}$ is!

Hypothesis testing for proportions

First, assume that $H_0 : p = p_0$ is true. Then, the observed proportion $\hat{p}$ is distributed as $\mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.



Additionally, we can put α in the mix.

The limits can be found as $p_0 \pm z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$.

So how to reject or fail to reject? See where $\hat{p}$ is!

Procedure

1. Select the desired α and set up your hypothesis test:

$$H_0 : p = p_0$$
$$H_1 : p \neq p_0.$$

2. Compute test statistic or simply $\hat{p}$:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \quad \text{or} \quad \hat{p}$$

3. Check whether Z_0 is below $z_{\alpha/2}$ or above $z_{\alpha/2}$, or check whether $\hat{p}$ is below $p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$ or above $p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$.

4. If yes, reject the hypothesis; otherwise, fail to reject it.

ILLINOIS

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \leq \hat{p} \leq U)$$

  where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0: p = p_0$ vs. $H_1: p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right)$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \leq \hat{p} \leq U)$$

where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0 : p = p_0$ vs. $H_1 : p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right).$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \leq \hat{p} \leq U)$$

  where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0 : p = p_0$ vs. $H_1 : p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right).$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \leq \hat{p} \leq U)$$

where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0 : p = p_0$ vs. $H_1 : p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right).$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.

ILLINOIS

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \leq \hat{p} \leq U)$$

  where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0 : p = p_0$ vs. $H_1 : p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0\left(1-p_0\right)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2}\sqrt{\frac{p_0\left(1-p_0\right)}{n}}\right).$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.

ILLINOIS

Calculating α and β

- α: typically given! But, if not, then it can be found as

$$P(L \le \hat{p} \le U)$$

where $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

- β: requires a specific alternative. For example, let $H_0 : p = p_0$ vs. $H_1 : p = p_1$. Then:
  - If H_1 is true, $\hat{p} \sim \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n}\right)$.
  - And hence:

$$\beta = P\left(p_0 - z_{\alpha/2}\sqrt{\frac{p_0\,(1-p_0)}{n}} \le \hat{p} \le p_0 + z_{\alpha/2}\sqrt{\frac{p_0\,(1-p_0)}{n}}\right).$$

- Finally, increasing the sample size n will improve (decrease) both α and β. Keeping n constant, then improving α will worsen β and vice versa.