Large-Sample Tests of Hypotheses

Outline

- A statistical test of hypothesis
- Large sample test about a population mean
- Large-sample test of hypothesis for the difference between two population means
- Large-sample test of hypothesis for a binomial proportion
- Large-sample test of hypothesis for the difference between two binomial proportions

A statistical test of hypothesis

Five components of a statistical test

(1) The **null hypothesis, H_0**

(2) The **alternative hypothesis, H_a**

(3) The **test statistic** and its **p-value**

(4) The **rejection region**

(5) The **conclusion**

A statistical test of hypothesis

- (1) The **null hypothesis, H_0**

  The hypothesis contradicting H_a, e.g. H_0: $\mu = \$456$
- (2) The **alternative hypothesis, H_a**

  The hypothesis that we wish to support, for example

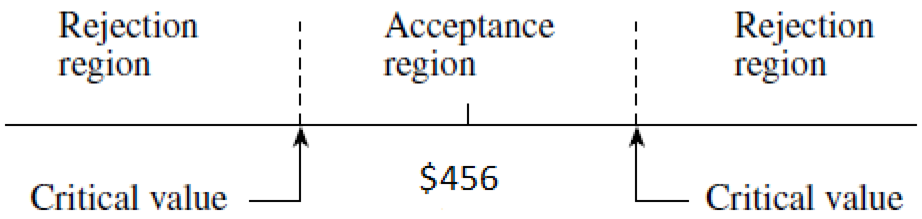
  Example 1: H_a: $\mu \neq \$456$ (2-tailed test of hypothesis)

  Example 2: H_a: $\mu < \$456$ (1-tailed test of hypothesis)

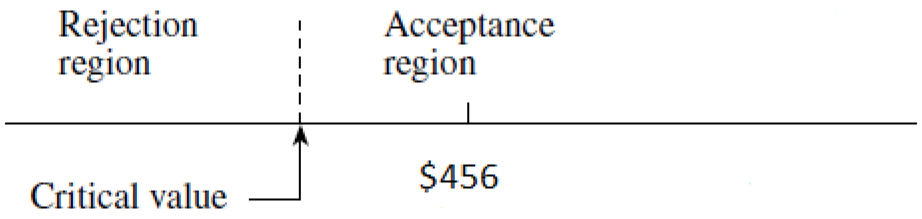
  Example 3: H_a: $\mu > \$456$ (1-tailed test of hypothesis)

A statistical test of hypothesis

- (3) **Test statistic** is a single value calculated from the sample data and **p-value** is a probability of observing an example as large (or as small) as the test statistic.
- (4) The set of possible values of test statistic can be divided into 2 regions
  - **Rejection region** – includes values that support the alternative hypothesis H_a and rejects the null hypothesis H_0
  - **Acceptance region** – includes values that support the null hypothesis H_0



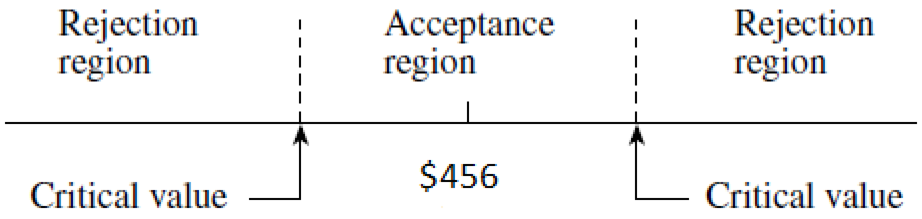
H_a: $\mu \neq \$456$



H_a: $\mu < \$456$

A statistical test of hypothesis

- (5) **Conclusions** – we *always* begin with assuming that the null hypothesis is true, then use sample data as evidence to decide one of the 2 conclusions
  - Reject H_0 and conclude H_a is true
  - Not enough evidence to reject H_0 - the test is inconclusive



- The critical values are decided based on the **significance level α**, which represents the probability of rejecting H_0 when it is true.
- **Type I error** – the error of rejecting the null hypothesis when it is true.

A large-sample test about a population mean

Example – The average monthly income of people in HCMC is $456. A random sample of n=51 IT professionals in HCMC showed that average income $\bar{x} = \$500$, with standard deviation $s = \$155$. Do IT professionals have higher monthly income than the city average? Test the hypothesis with significance level $\alpha = .05$ (or 5%).

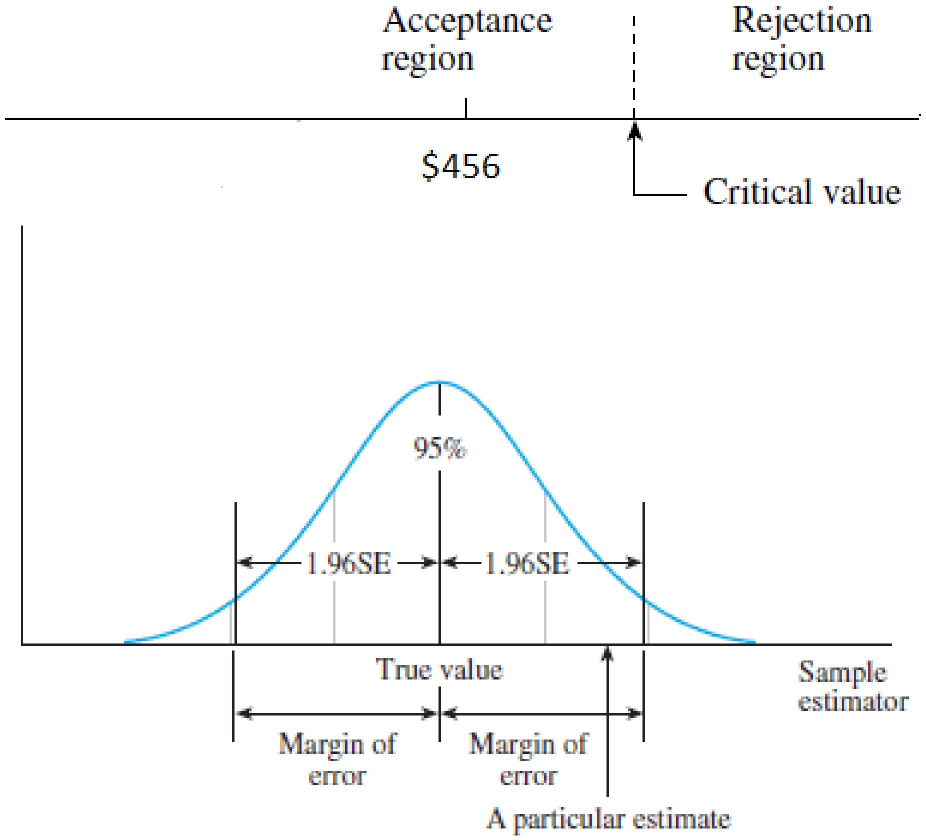
- (1) The **null hypothesis, H_0**: $\mu = \$456$
- (2) The **alternative hypothesis, H_a**: $\mu > \$456$

A large-sample test about a population mean

- Because n is fairly large, the sample mean $\bar{x} = \$500$ is the best estimate of the true average income μ of IT professionals in HCMC (the Central Limit Theorem).

- How large $\bar{x}$ needs to be compared to $\mu_0 = \$456$ for us to reject the null hypothesis?

- Because the sampling distribution of $\bar{x}$ follows a normal distribution, the mean of which is μ, if μ_0 is *many standard errors* (SEs) away from μ we can fairly sure that the probability to see μ_0 is very low, i.e. μ_0 does not equal μ.



A large-sample test about a population mean

- But how many SEs are enough? We need to rely on the significance level α.
- Standard error of $\bar{x}$, $SE = \frac{s}{\sqrt{n}} = \frac{155}{\sqrt{51}} = \$21.9$
- (3) **Test statistic**: The number of SEs $\mu_0 = \$456$ is away from $\bar{x}$ is calculated by

$$\mathrm{z} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{500-456}{21.9} = 2.03$$

In other words, $\bar{x} = \mu_0 + 2.03 * SE$.

- (4) Rejection region: For significance level $\alpha = .05$, the corresponding z-score is 1.64. Any observed z-value larger than this will be in the rejection region.
- (5) Conclusions: Because the test statistic z = 2.03 is larger than the critical value of 1.64, we reject the null hypothesis, and conclude that the *average monthly income of IT professionals is higher than the city average*.
- The probability of this conclusion being wrong is $\alpha = 5\%$.

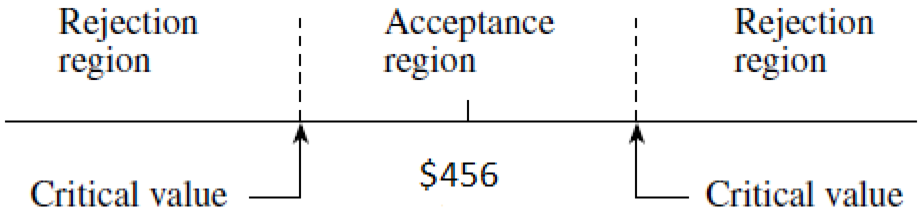
A large-sample test about a population mean

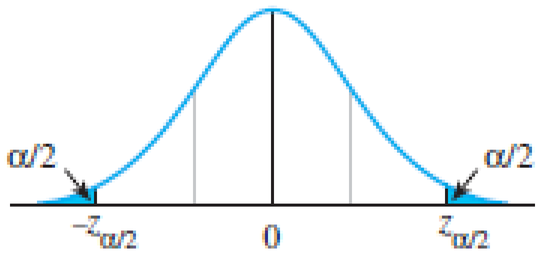
Example – The average monthly income of people in HCMC is \$456. A random sample of n=51 IT professionals in HCMC showed that average income $\bar{x} = \$500$, with standard deviation $s = \$155$. Do IT professionals have monthly income **different to** the city average? Test the hypothesis with significance level $\alpha = .05$ (or 5%).

- (1) The **null hypothesis, H_0**: $\mu = \$456$
- (2) The **alternative hypothesis, H_a**: $\mu \neq \$456$

A large-sample test about a population mean

- (3) **Test statistic** – We use the same reasoning as before and come up with the test statistic z = 2.03
- (4) **Rejection region** – In 2 tailed test using significance level $\alpha = .05$, the critical values separating the rejection region and the acceptance region corresponds to $\alpha/2 = .025$ to the right and left of the tail of the standardized normal distribution. These values are $\mathrm{z} = \pm 1.96$. The rejection region includes z < -1.96 of z > 1.96.





- (5) **Conclusion** – Because z = 2.03 is larger than 1.96, we ignore the null hypothesis and conclude that the average monthly income of IT professionals *is different to the city average*. The probability of making the wrong decision is $\alpha = 5\%$.

A large-sample test about a population mean

In summary:

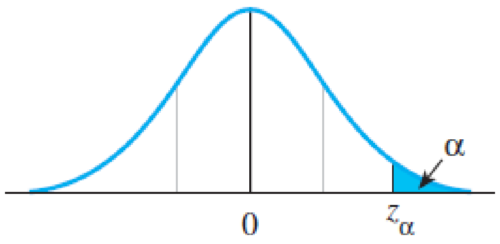
1. Null hypothesis: $H_0 : \mu = \mu_0$
2. Alternative hypothesis:

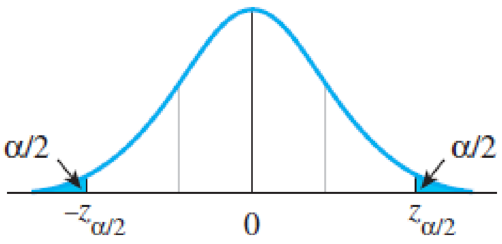
One-Tailed Test	**Two-Tailed Test**
$H_a : \mu > \mu_0$ (or, $H_a : \mu < \mu_0$)	$H_a : \mu \neq \mu_0$

3. Test statistic: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ estimated as $z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

4. Rejection region: Reject H_0 when

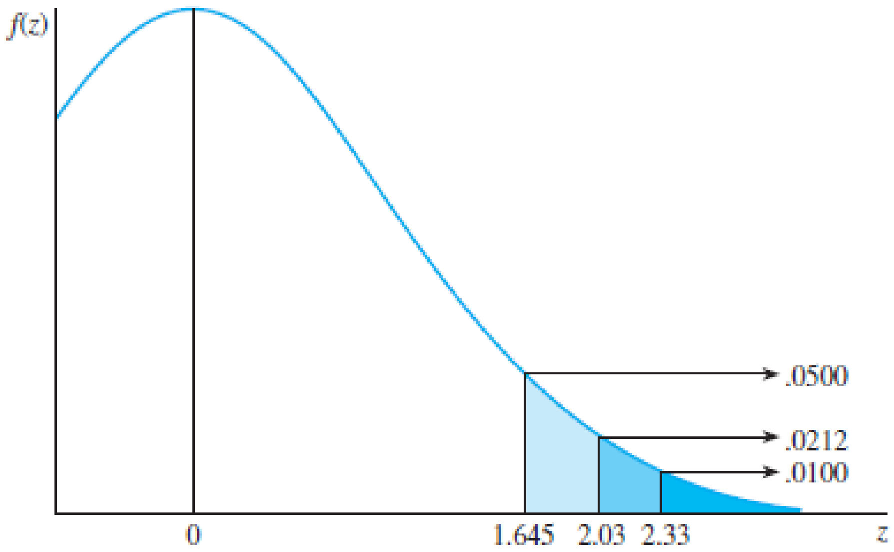
One-Tailed Test	**Two-Tailed Test**
$z > z_\alpha$ (or $z < -z_\alpha$ when the alternative hypothesis is $H_a : \mu < \mu_0$)	$z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$





A Large-sample test about a population mean

- In previous examples, the decision to reject a null hypothesis was based on value of z determined from a significance level α.
- In Example 1, $\alpha = .05$, the critical value of z is 1.64. We rejected the null hypothesis because the observed value of z_0 = 2.03 is larger the critical value.
- However if $\alpha = .01$, the critical value of z is 2.33, we do not reject the null hypothesis because z_0 = 2.03 is smaller the critical value. (The conclusion in this case is that the *average monthly income of IT professionals is not higher than the city average*)



A Large-sample test about a population mean

- The smallest critical value that we can use to reject H_0 is 2.03. The probability of this reject decision being wrong is P(z>2.03) = .0212, which if the p-value for the test.
- Smaller p-value means larger z_0, which means larger distance between $\mu_0 = \$456$ and sample mean $\bar{x} = \$500$, which means higher chance of rejecting the null hypothesis.

- p-value can also be compared *directly* with the significance level α.
  - *If p-value $\leq \alpha$, we reject the null hypothesis and report that the results are* ***statistically significant*** *at level α.*

- In example 1 (one-tailed test),

  p-value = P(z>2.03) = .0212
- In example 2 (two-tailed test),

  p-value = P(z>2.03) + P(z<-2.03) = .0212 + .0212 = .0424

A large-sample test about a population mean

- In example 1, if we set the significant level $\alpha = .01$, because p-value = .0212 is larger than α, we do not reject the null hypothesis and conclude that the average monthly income of IT professionals is not higher than the city average.
- Note that we do NOT say that we *accept* the null hypothesis, i.e. we do NOT conclude that the average monthly income of IT professionals *equals* the city average.
- This is because if we choose to accept the null hypothesis, we need to know the probability of error associate with such a decision.
- ***Type II error*** *for statistical test is the error of accepting the null hypothesis when it is false and an alternative hypothesis is true, represented by a probability* β.

	Null Hypothesis	
Decision	True	False
Reject H_0	Type I error	Correct decision
Accept H_0	Correct decision	Type II error

Large-sample test of hypothesis for the difference between two population means

Assumptions. Given two *large* samples (n_1>30 and n_2>30) *randomly and independently drawn* from two populations.

(1) Null hypothesis $\quad H_0: \mu_1 - \mu_2 = D_0$

(2) Alternative hypothesis $\quad H_a: \mu_1 - \mu_2 > D_0$, or Ha: $\mu_1 - \mu_2 < D_0$ (one-tailed test)

$H_a: \mu_1 - \mu_2 \neq D_0$ (two-tailed test)

(3) Test statistic

$$\mathrm{z} = \frac{(\mu_1 - \mu_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx \frac{(\mu_1 - \mu_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(4) Rejection region

$\mathrm{z} > z_\alpha$ for $H_a: \mu_1 - \mu_2 > D_0$

$\mathrm{z} < -z_\alpha$ for $H_a: \mu_1 - \mu_2 < D_0$

$\mathrm{z} > z_{\alpha/2}$ or $\mathrm{z} < -z_{\frac{\alpha}{2}}$ for $H_a: \mu_1 - \mu_2 \neq D_0$

OR $\mathrm{pValue} < \alpha$ (for any $\mathrm{H_a}$)

Large-sample test of hypothesis for the difference between two population means

Example. A 2018 survey from 100 randomly selected foreign tourists in Hanoi and 100 randomly selected tourists in HCMC revealed that the average stay in Hanoi and in HCMC were 3.4 days and 3.67 days respectively. The standard deviation of the Hanoi sample is 1.2 days and of the HCMC sample is 1 day. Was there enough evidence to conclude that the average length of stay of foreign tourists are different between the 2 cities? Use $\alpha = .05$

Calculate the 95% confidence intervals of the difference between 2 population means.

Large-sample test of hypothesis for a binomial proportion

Assumptions. Given a *large* number of n identical trials *randomly drawn* from a binomial population, i.e. $\mathrm{n}\hat{p} > 5$ and $\mathrm{n}(1 - \hat{p}) > 5$

- (1) Null hypothesis: $\mathrm{H_0}: p = p_0$
- (2) Alternative hypothesis: $\mathrm{H_a}: p < p_0$ or $\mathrm{H_a}: p > p_0$ (one-tailed test)

  $\mathrm{H_a}: p \neq p_0$ (two-tailed test)
- (3) Test statistic: $z = (\hat{p} - p_0)/\sqrt{p_0(1 - p_0)/n}$

  where $\hat{p}$ is sample proportion
- (4) Rejection region: $z > z_\alpha$ for $\mathrm{H_a}: p > p_0$

  $z < -z_\alpha$ for $\mathrm{H_a}: p < p_0$

  $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$ for $\mathrm{H_a}: p \neq p_0$

OR $\mathrm{pValue} < \alpha$ (for any $\mathrm{H_a}$)

Large-sample test of hypothesis for a binomial proportion

Example. A survey in the first half of 2019 observed that out of 500 random visitors to HCMC, 123 were foreigners. Is this evidence sufficient to conclude that the proportion of foreign visitors to HCMC has increased from 2018, which was approximately 20.4%? Use $\alpha = .05$

Large-sample test of hypothesis for the difference between 2 binomial proportions

Assumption. Given two samples *independently and randomly* drawn from two binomial populations, and that each sample has *large* number trials, i.e. $n_1\widehat{p_1}$, $n_2\widehat{p_2}$, $n_1(1-\widehat{p_1})$, and $n_2(1-\widehat{p_2})$ larger than 5.

(1) Null hypothesis $\quad$ H_0: $p_1 - p_2 = D_0$

(2) Alternative hypothesis $\quad$ H_a: $p_1 - p_2 > D_0$, or H_a: $p_1 - p_2 < D_0$ (one-tailed test)

H_a: $p_1 - p_2 \neq D_0$ (two-tailed test)

(3) Test statistic $\quad$ $z = (\widehat{p_1} - \widehat{p_2} - D_0) \Big/ \sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1} + \frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}}$

where $\widehat{p_1}$ and $\widehat{p_2}$ are proportion of sample 1 and sample 2 respectively

(4) Rejection region $\quad$ $z > z_\alpha$ for H_a: $p_1 - p_2 > D_0$

$z < -z_\alpha$ for H_a: $p_1 - p_2 < D_0$

$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$ for H_a: $p_1 - p_2 \neq D_0$

OR pValue $< \alpha$ (for any H_a)

Large-sample test of hypothesis for the difference between 2 binomial proportions

Example. A 2018 survey observed that 102 out of random 500 tourists visiting HCMC were foreigners. The same survey observed that only 98 out of random 500 tourists visiting Hanoi were foreigners. Is there enough evident to conclude that HCMC has a higher proportion of foreign visitors compared to Hanoi? Use $\alpha = .05$

Calculate the 95% confidence interval of the difference between the proportions of foreign visitors of the 2 cities.