Statistics for Business and Economics

Chapter 10

Hypothesis Testing: Additional Topics

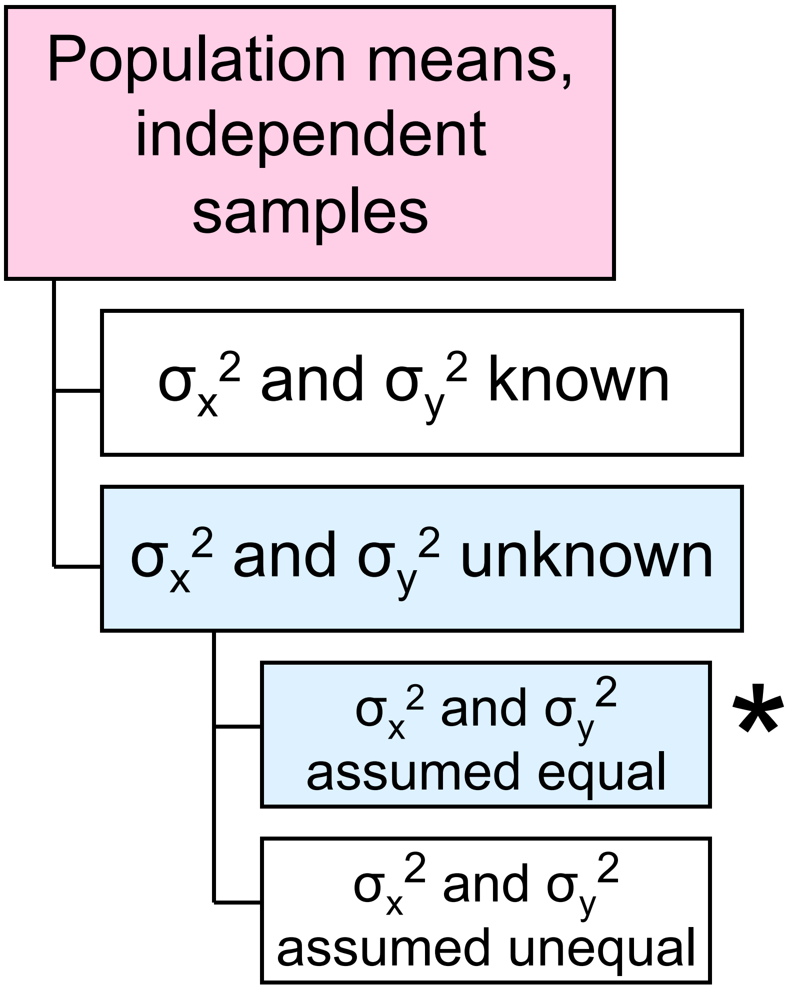
10.2 Difference Between Two Means

Population means, independent samples

Goal: Form a confidence interval for the difference between two population means, $\mu_x - \mu_y$

- Different populations
  - Unrelated
  - Independent
    - Sample selected from one population has no effect on the sample selected from the other population
  - Normally distributed

σ_x^2 and σ_y^2 Unknown, Assumed Equal

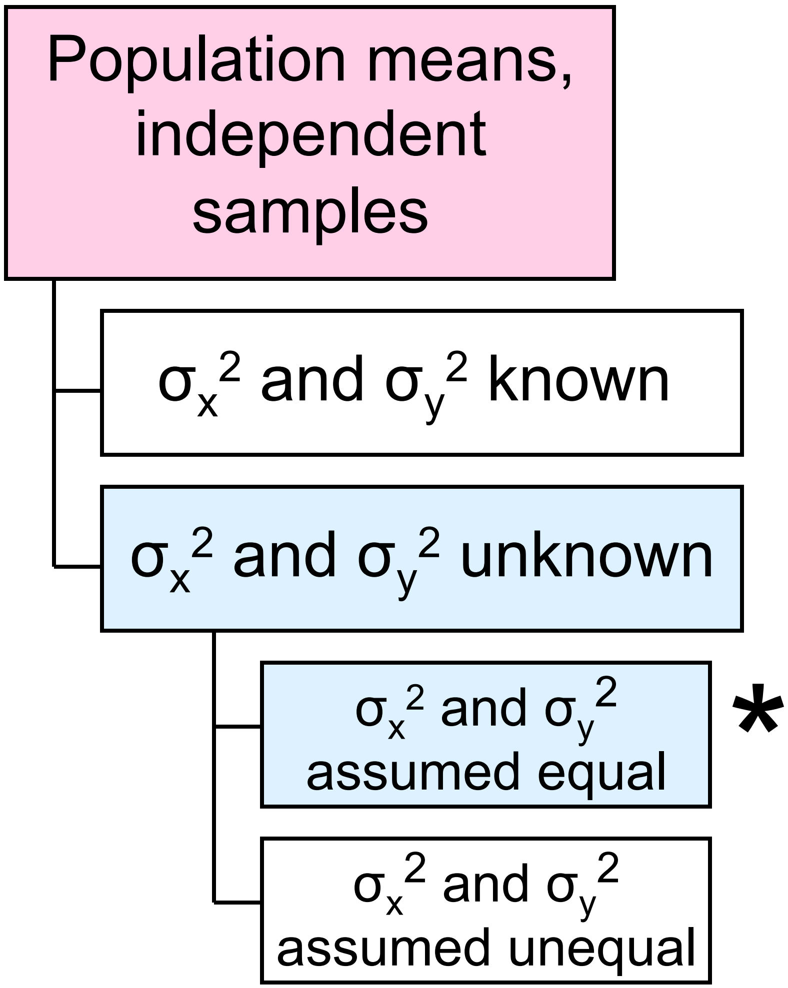


Assumptions:

- Samples are randomly and independently drawn
- Populations are normally distributed
- Population variances are unknown but assumed equal

σ_x^2 and σ_y^2 Unknown, Assumed Equal

(continued)



- The population variances are assumed equal, so use the two sample standard deviations and pool them to estimate σ
- use a t value with $(n_x + n_y - 2)$ degrees of freedom

Test Statistic, σ_x^2 and σ_y^2 Unknown, Equal

σ_x^2 and σ_y^2 unknown

- σ_x^2 and σ_y^2 assumed equal *
- σ_x^2 and σ_y^2 assumed unequal

The test statistic for $\mu_x - \mu_y$ is:

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}}$$

Where t has $(n_1 + n_2 - 2)$ d.f.,

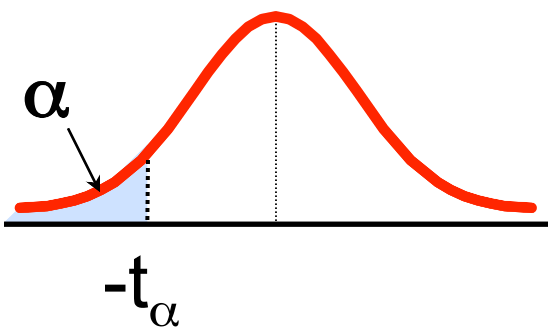
and

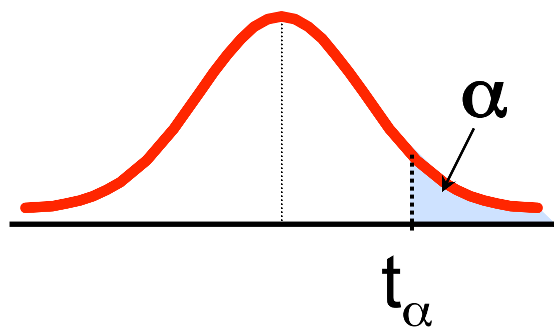
$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$$

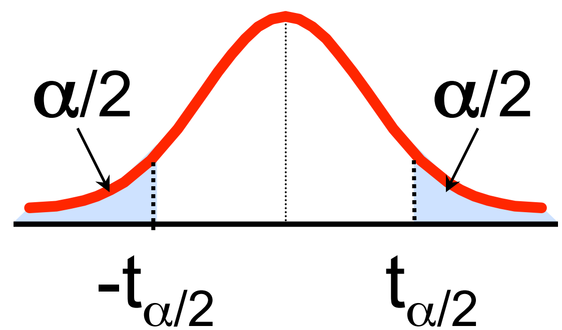
Decision Rules

Two Population Means, Independent Samples, Variances Unknown

Lower-tail test:	Upper-tail test:	Two-tail test:
$H_0: \mu_x - \mu_y \geq 0$ $H_1: \mu_x - \mu_y < 0$	$H_0: \mu_x - \mu_y \leq 0$ $H_1: \mu_x - \mu_y > 0$	$H_0: \mu_x - \mu_y = 0$ $H_1: \mu_x - \mu_y \neq 0$







Reject H_0 if
$t \leq -t_{(n1+n2-2),\, \alpha}$

Reject H_0 if
$t \geq t_{(n1+n2-2),\, \alpha}$

Reject H_0 if
$t \leq -t_{(n1+n2-2),\, \alpha/2}$ or
$t \geq t_{(n1+n2-2),\, \alpha/2}$

Pooled Variance t Test: Example

You are a financial analyst for a brokerage firm. Is there a difference in dividend yield between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	**NASDAQ**
Number	**21**	**25**
Sample mean	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Assuming both populations are approximately normal with equal variances, is there a difference in average yield (α = 0.05)?

Calculating the Test Statistic

The test statistic is:

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\left(3.27 - 2.53\right) - 0}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = \boxed{2.040}$$

$$S_p^2 = \frac{\left(n_1 - 1\right)S_1^{\,2} + \left(n_2 - 1\right)S_2^{\,2}}{(n_1 - 1) + (n_2 - 1)} = \frac{\left(21 - 1\right)1.30^2 + \left(25 - 1\right)1.16^2}{(21 - 1) + (25 - 1)} = 1.5021$$

Solution

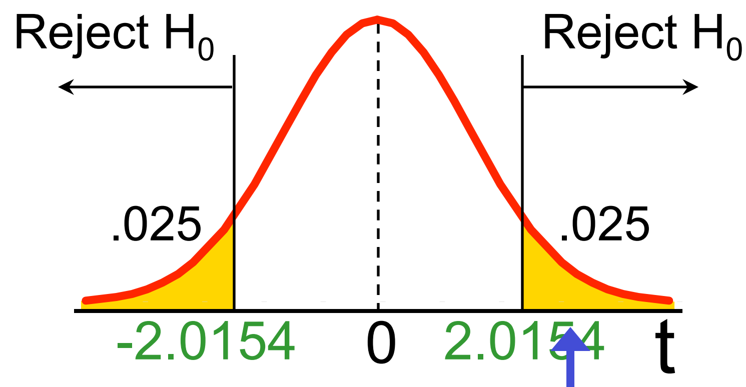
H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 21 + 25 − 2 = 44

Critical Values: t = ± 2.0154



2.040

Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.040$$

Decision:

Reject H_0 at $\alpha = 0.05$

Conclusion:

There is evidence of a difference in means.

10.3

Two Population Proportions

Population proportions

Goal: Test hypotheses for the difference between two population proportions, $p_x - p_y$

Assumptions:

Both sample sizes are large.

Two Population Proportions

(continued)

Population proportions

- The random variable

$$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\dfrac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \dfrac{\hat{p}_y(1-\hat{p}_y)}{n_y}}}$$

is approximately normally distributed

Test Statistic for Two Population Proportions

Population proportions

The test statistic for

$H_0: P_x - P_y = 0$

is a z value:

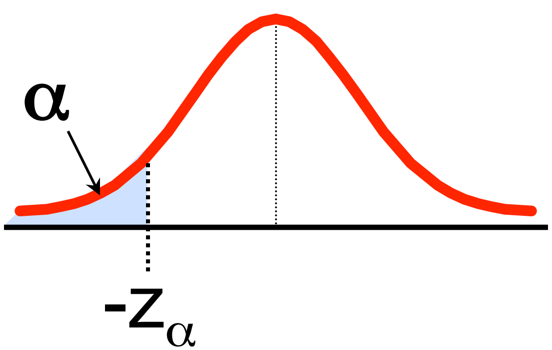
$$z = \frac{\left(\hat{p}_x - \hat{p}_y\right)}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}}$$

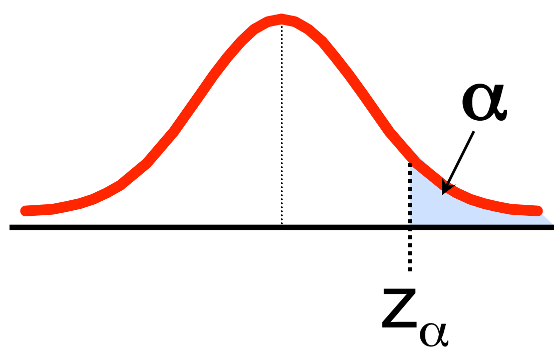
Where $\hat{p}_0 = \frac{n_x\hat{p}_x + n_y\hat{p}_y}{n_x + n_y}$

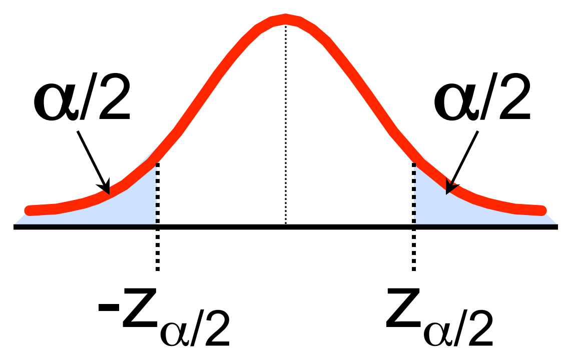
Decision Rules: Proportions

Population proportions

Lower-tail test:	Upper-tail test:	Two-tail test:
$H_0: p_x - p_y \geq 0$ $H_1: p_x - p_y < 0$	$H_0: p_x - p_y \leq 0$ $H_1: p_x - p_y > 0$	$H_0: p_x - p_y = 0$ $H_1: p_x - p_y \neq 0$







Reject H_0 if $z \leq -z_{\alpha}$

Reject H_0 if $z \geq z_{\alpha}$

Reject H_0 if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$

Example: Two Population Proportions

Is there a significant difference between the proportion of men and the proportion of women who will vote Yes on Proposition A?

- In a random sample, 36 of 72 men and 31 of 50 women indicated they would vote Yes
- Test at the .05 level of significance

Example: Two Population Proportions

(continued)

- The hypothesis test is:

  $H_0: p_M - p_W = 0$ (the two proportions are equal)
  $H_1: p_M - p_W \neq 0$ (there is a significant difference between proportions)

- The sample proportions are:
  - Men: $\hat{p}_M = 36/72 = .50$
  - Women: $\hat{p}_W = 31/50 = .62$

- The estimate for the common overall proportion is:

$$\hat{p}_0 = \frac{n_M\hat{p}_M + n_W\hat{p}_W}{n_M + n_W} = \frac{72(36/72) + 50(31/50)}{72 + 50} = \frac{67}{122} = .549$$

Example: Two Population Proportions

(continued)

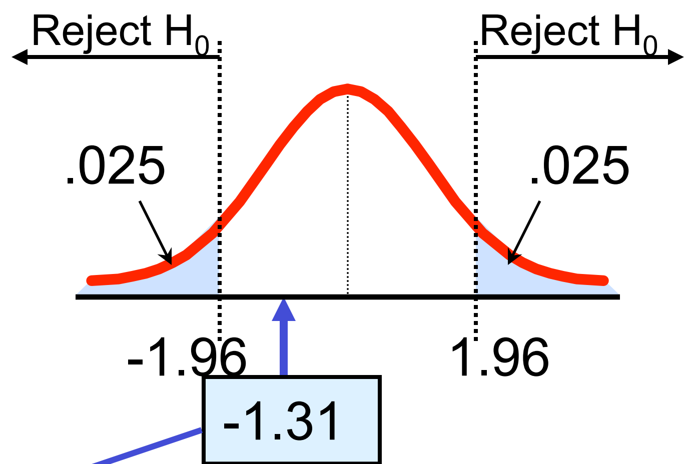
The test statistic for $P_M - P_W = 0$ is:

$$z = \frac{(\hat{p}_M - \hat{p}_W)}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_1} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_2}}}$$

$$= \frac{(.50 - .62)}{\sqrt{\left(\frac{.549(1-.549)}{72} + \frac{.549(1-.549)}{50}\right)}}$$

$$= -1.31$$

Critical Values = ±1.96
For α = .05



Decision: Do not reject H_0

Conclusion: There is not significant evidence of a difference between men and women in proportions who will vote yes.