Topic 5: Hypothesis Testing

Optional Reading: Chapter 9

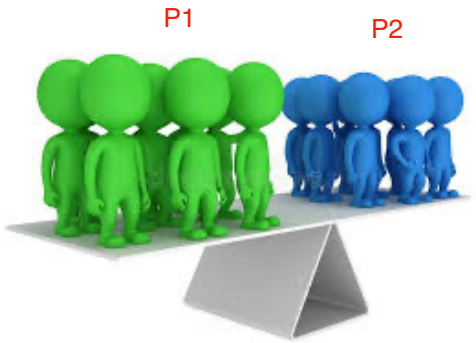
Xiner Zhou

Department of Statistics

University of California, Davis

- **Test about a population mean**
- **Test about a population proportion**
- **Test about the difference between two population means**
- **Test about the difference between two population proportions**

Hypothesis Testing about the difference between two population proportions



Summary: Hypothesis Testing about the difference between two population proportions

Step 1: null hypothesis
$H_0 : p_1 = p_2$

Step 2: Alternative hypothesis

One-sided
$H_a : p_1 > p_2$ or $H_a : p_1 < p_2$

Two-sided
$H_a : p_1 \neq p_2$

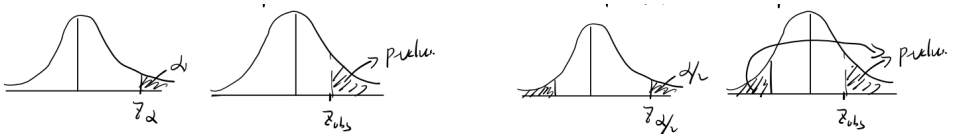
Step 3: calculate test statistic

$$Z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Where $\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$ is the pooled sample proportion

Step 4: critical value or p-value approach

Critical value | P-value | Critical value | P-value



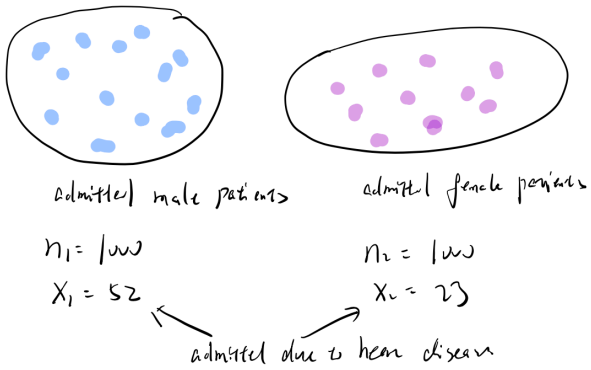
Step 5: Decision and Conclusion

Reject H_0 if:

$Z_{obs} \geq Z_\alpha$ | p-value $\leq \alpha$ | $|Z_{obs}| \geq Z_\alpha$ | p-value $\leq \alpha$

The hospital sees a larger number of male patients than female patients.
So one of the cardiologists is interested to find out whether men are more likely to have heart disease than women, in this hospital.

Does the data present sufficient evidence that there is a higher rate of heart disease among men versus women who have been admitted to this hospital? Use significance level 0.05.



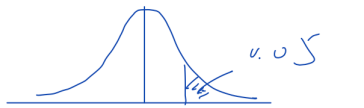
Step 1: $H_0: p_1 - p_2 = 0$ vs $H_a: p_1 - p_2 > 0$

Step 2: $$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{\frac{51}{100} - \frac{23}{100}}{\sqrt{\frac{74}{200}\left(1 - \frac{74}{200}\right)\left(\frac{1}{100} + \frac{1}{100}\right)}} = 3.41$$

Step 3:

Critical value approach:



$z_{0.05} = 1.64$

Acceptance region | rejection region

p-value approach: $p\text{-value} = P(z \geq z_{obs})$

$= P(z \geq 3.41)$

$= 0.0003$

Step 4: conclusion and decision:

Critical value approach:

Since the observed test statistic falls in the rejection region,
We reject H_0 and conclude that, there is significantly higher rate of heart disease among men than women who have been admitted to this hospital.

P-value approach:

Since p-value =0+ < 0.05,
We reject H_0 and conclude that, there is significantly higher rate of heart disease among men than women who have been admitted to this hospital.

Hormone Therapy and Alzheimer's Disease A 4-year experiment involving 4532 women, reported in The Press Enterprise, was conducted at 39 medical centers to study the benefits and risks of hormone replacement therapy (HRT). Half of the women took placebos and half took Prempro, a widely prescribed type of hormone replacement therapy. There were 40 cases of dementia in the hormone group and 21 in the placebo group. Is there sufficient evidence to indicate that the risk of dementia is higher for patients using Prempro? Test at the 1% level of significance.

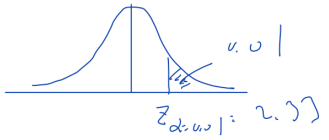
Step 1 : $H_0: p_1 - p_2 = 0$ vs $H_a: p_1 - p_2 > 0$

Step 2 : $$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{\frac{40}{2266} - \frac{21}{2266}}{\sqrt{\frac{40+21}{4532}\left(1 - \frac{40+21}{4532}\right)\left(\frac{1}{2266} + \frac{1}{2266}\right)}} = 2.45$$

Step 3 :

critical value approach :



Acceptance region rejection region

p-value approach : $p\text{-value} = P(z \geq z_{obs})$

$= P(z \geq 2.45)$

$= 0.007$

Step 4: conclusion and decision:

Critical value approach:

Since the observed test statistic falls in the rejection region,
We reject H_0 and conclude that, the risk of dementia is significantly higher for patients who took the HRT.

P-value approach:

Since p-value =0+ < 0.05,
We reject H_0 and conclude that, the risk of dementia is significantly higher for patients who took the HRT.

Female Models In a study to assess various effects of using a female model in automobile advertising, 100 men were shown photographs of two automobiles matched for price, color, and size, but of different makes. One of the automobiles was shown with a female model to 50 of the men (group A), and both automobiles were shown without the model to the other 50 men (group B). In group A, the automobile shown with the model was judged as more expensive by 37 men; in group B, the same automobile was judged as the more expensive by 23 men. Do these results indicate that using a female model influences the perceived cost of an automobile? Use a one-tailed test with $\alpha = .05$.

Step 1: $H_0: p_A - p_B = 0$ vs $H_a: p_A - p_B > 0$

Step 2: $$z_{obs} = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

$$= \frac{\frac{37}{50} - \frac{23}{50}}{\sqrt{\frac{37+23}{100}\left(1 - \frac{37+23}{100}\right)\left(\frac{1}{50} + \frac{1}{50}\right)}} = 2.86$$

Step 3:

critical value approach:

$z_{\alpha=0.05} = 1.64$

Acceptance region | rejection region

p-value approach: $p\text{-value} = P(z \geq z_{obs})$

$= P(z \geq 2.86)$

$= 0.002$

Step 4: conclusion and decision:

Critical value approach:

Since the observed test statistic falls in the rejection region,
We reject H_0 and conclude that, using female model does influence the perceived cost of an automobile, specifically, using female model significantly increases the likelihood of being perceived as more expansive.

P-value approach:

Since p-value =0+ < 0.05,
We reject H_0 and conclude that, using female model does influence the perceived cost of an automobile, specifically, using female model significantly increases the likelihood of being perceived as more expansive.

Adolescents and Social Stress In a study to compare ethnic differences in adolescents' social stress, researchers recruited subjects from three middle schools in Houston, Texas. [22] A tabulation of student responses to a question regarding their socioeconomic status (SES) compared with other families in which the students chose one of five responses (much worse off, somewhat worse off, about the same, better off, or much better off) resulted in the tabulation that follows.

	European American	African American	Hispanic American	Asian American
Sample Size	144	66	77	19
About the Same	68	42	48	8

1. Do these data support the hypothesis that the proportion of adolescent African Americans who state that their SES is "about the same" exceeds that for adolescent Hispanic Americans?

2. Find the p-value for the test.

3. If you plan to test using $\alpha = .05$, what is your conclusion?

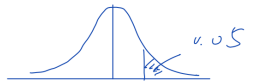
Step 1 : $H_0: p_{AA} - p_{His} = 0$ vs $H_a: p_{AA} - p_{His} > 0$

Step 2 : $z_{obs} = \dfrac{\hat{p}_{AA} - \hat{p}_{His}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_{AA}} + \frac{1}{n_{His}}\right)}}$

$= \dfrac{\frac{42}{66} - \frac{48}{77}}{\sqrt{\frac{42+48}{66+77}\left(1-\frac{42+48}{66+77}\right)\left(\frac{1}{66}+\frac{1}{77}\right)}} = 0.16$

Step 3 :

critical value approach :



Acceptance region rejection region

rejection region

p-value approach : $p\text{-value} = P(z \geq z_{obs})$

$= P(z \geq 0.16)$

$= 0.44$

Step 4: conclusion and decision:

Critical value approach:

Since the observed test statistic falls in the acceptance region,
We do not reject H_0 and conclude that, the proportion of African Americans who state their SES as “about the same” does not exceed that for Hispanic Americans.

P-value approach:

Since p-value =0.44> 0.05,
We do not reject H_0 and conclude that, the proportion of African Americans who state their SES as “about the same” does not exceed that for Hispanic Americans.