POLS 5377 Scope & Method of Political Science

Week 12 Hypothesis Testing II

The Two-Sample Test

Healey. (2016) *Statistics: A Tool for Social Research,* Chapter 8 & 9

Key Questions:

* How to test a hypothesis with a small sample size?
* What is the basic logic of the two-sample test?
* How to conduct the two-sample test for means (large and small samples)?
* How to conduct the two-sample test for proportions (large samples)?

Outline

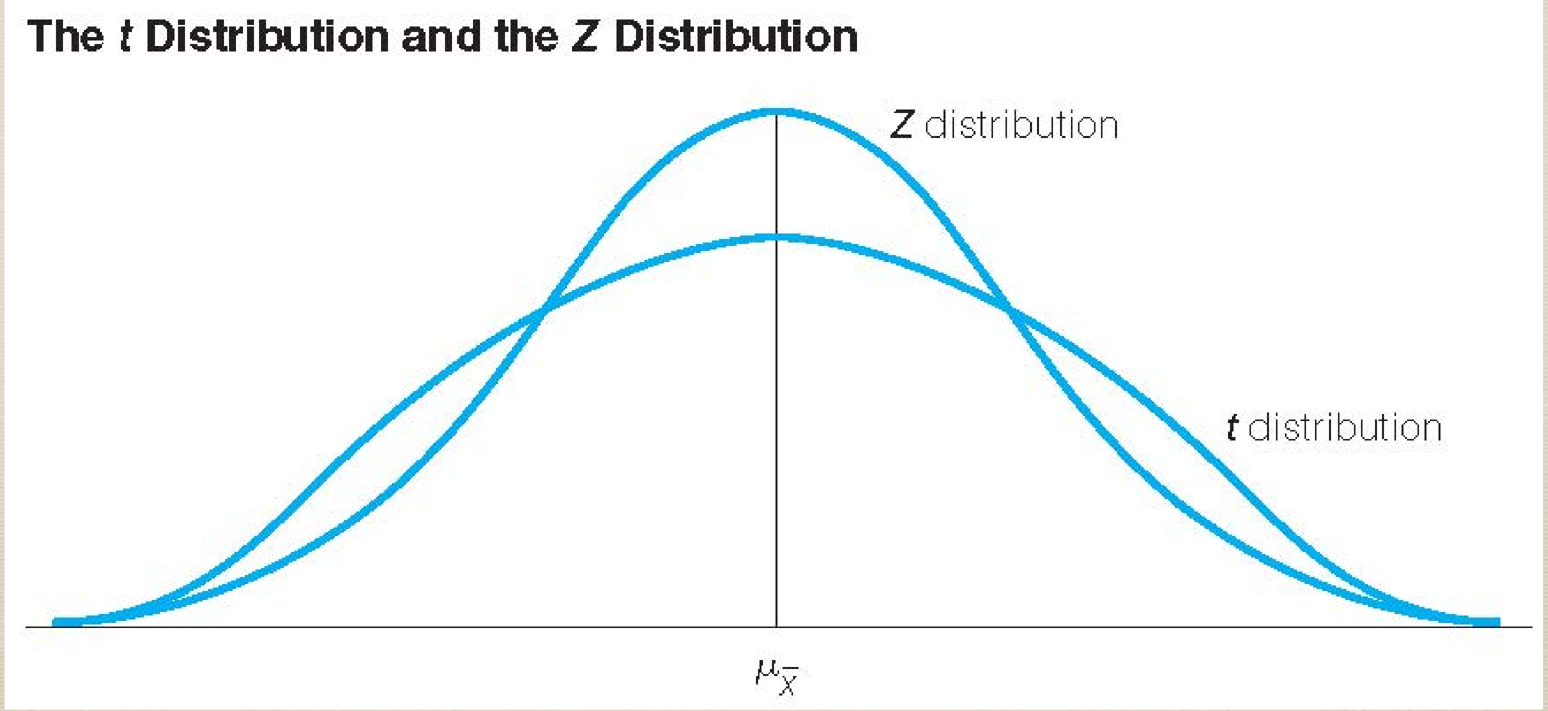
* The Student's t Distribution
* Logic of Two-Sample Case
* Two-Sample Test of Means (Large Samples)
* Two-Sample Test of Means (Small Samples)
* Two-Sample Test of Proportions (Large Samples)

The Student's *t* Distribution

- When the population standard deviation (σ) is known:
  - Conduct z test with the formula: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{N}}$
- when the population standard deviation (σ) is unknown:
  - For **large** samples (N ≥ 100), we can use the sample standard deviation (s) as an estimator of the population standard deviation (σ) (Formula: $Z = \frac{\bar{X}-\mu}{S/\sqrt{N-1}}$) Use standard (Z) normal distribution
  - For **small** samples (N < 100), sample standard deviation (*s)* is too biased an estimator of σ,
    - **Do not** use standard normal distribution (Z distribution)
    - **Use Student's t distribution**
    - Formula: $t = \frac{\bar{X}-\mu}{S/\sqrt{N-1}}$

The Student's *t* Distribution

- Compare the Z distribution to the Student's *t* distribution:



The Student's *t* Distribution

- Compare the Z distribution to the Student's *t* distribution:

Choosing a Sampling Distribution When Testing Single-Sample Means for Significance

If Population Standard Deviation (σ) Is	Sampling Distribution Is the
Known	*Z* distribution
Unknown and sample size (*N*) is large	*Z* distribution
Unknown and sample size (*N*) is small	*t* distribution

The Student’s *t* Distribution

- Use Appendix B for the *t* distribution
- How *t* table differs from Z table?
  - Column at left for degrees of freedom (df)
    - df = N – 1
  - Alpha levels along top two rows: one- and two-tailed
  - Entries in table are actual scores: t(critical)
    - Mark beginning of critical region, not areas under the curve

The Student's *t* Distribution

- A researcher wants to know whether commuter students are different from the general student body in terms of academic achievement.
- Based on the record, the average student GPA is 2.50. She randomly collected a sample of 30 commuter students, and learned the average GPA of the 30 students is 2.78 with a standard deviation of 1.23.
- Based on these information, is there **a difference** between commuter students and the general students? Make decision with the confidence level at 99%.

The Student’s *t* Distribution

Step 1: Make Assumptions and Meet Test Requirements

- Random sampling
- Level of measurement is interval-ratio
  - Student academic achievement is measures in GPA
- Sampling distribution is normal

The Student's *t* Distribution

Step 2: State the Null and Alternative hypothesis

- H_0: $\mu = 2.5$ There is no difference between commuter students and the general students in terms of academic achievement.
- H_1: $\mu \neq 2.5$ There is a difference between commuter students and the general students in terms of academic achievement.

- (Based on the question statement, the researcher didn't predict a specific direction of the difference. Therefore, we will run a two-tailed test.)

The Student's *t* Distribution

Step 3: Select Sampling Distribution and Establish the Critical Region

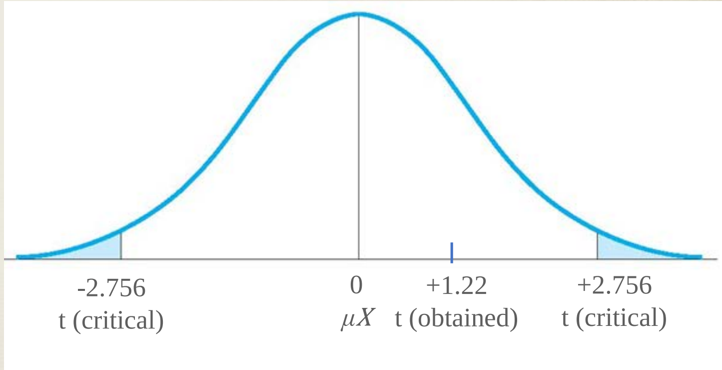
- Sampling distribution = *t* distribution
  - because σ is unknown and N=30 (<100)
  - Alpha (α) = 0.01 for 99% confidence level (two-tailed test)
  - df = N-1= 30-1= 29
- Critical region begins at ± 2.756
  - This is the critical *t* score associated with a two-tailed test and α equal to 0.01
  - If the *t* score of the obtained sample mean falls in the critical region, reject H_0

df	Two-Tailed Test		
	0.02	**0.01**	**0.001**
28	2.467	2.763	3.674
29	2.462	**2.756**	3.659
30	2.457	2.750	3.646

The Student's *t* Distribution

Step 4: Compute the Test Statistic

- Sample size: 30 (<100, small sample)
- Population standard deviation (σ) : unknown
- Population mean (μ) = 2.5
- Sample mean ($\bar{X}$) = 2.78
- Sample standard deviation (s): 1.23
- Formula: $t = \frac{\bar{X}-\mu}{S/\sqrt{N-1}}$



$$t(obtained) = \frac{2.78-2.5}{1.23/\sqrt{30-1}} = \frac{0.28}{1.23/5.39} = \frac{0.28}{0.23} = 1.22$$

t (obtained): 1.22

The Student's *t* Distribution

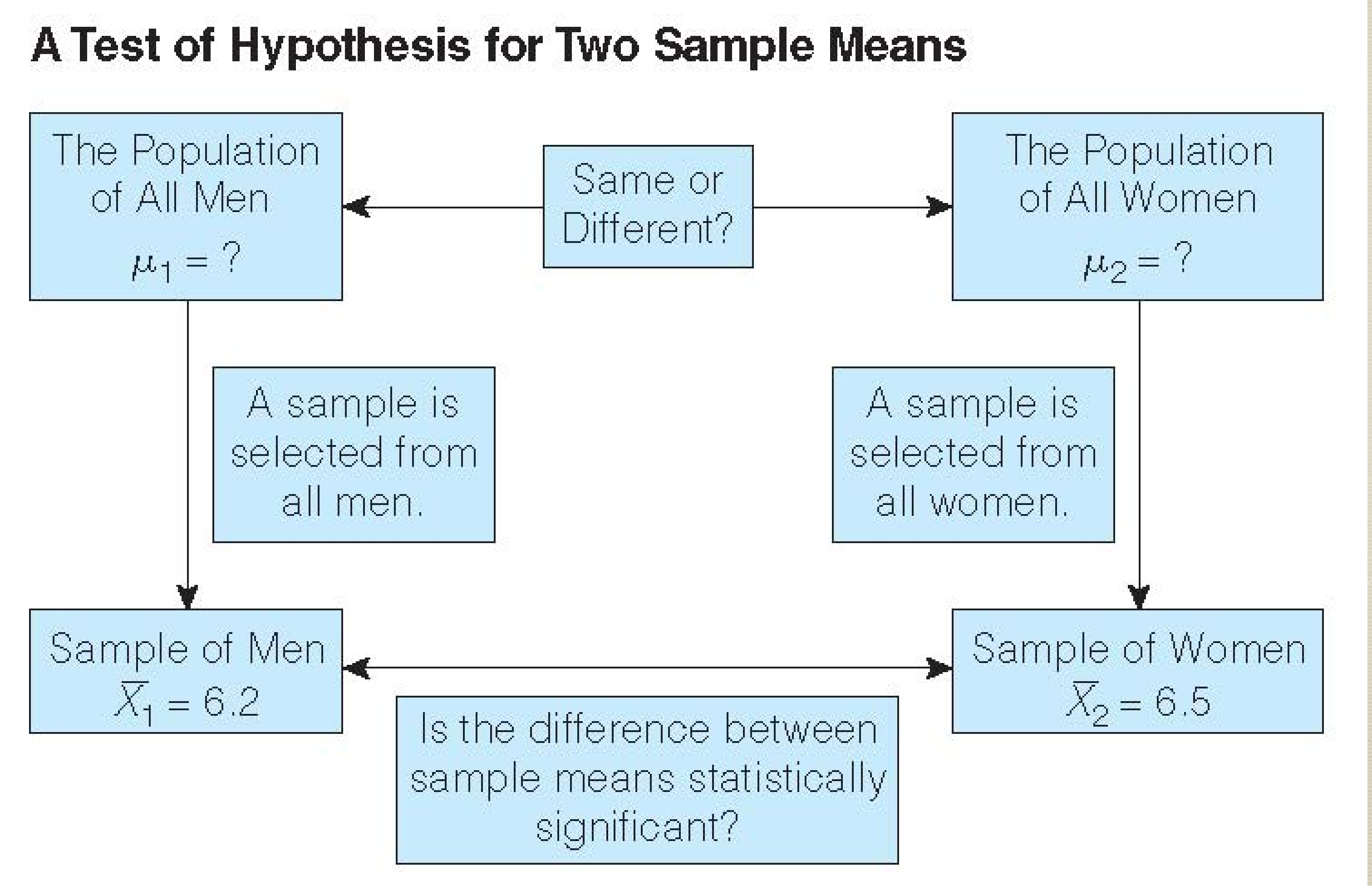
Step 5: Make a Decision and Interpret Results

- The obtained *t* score **doesn't** fall in the critical region so we **fail to reject** the H_0
  - If the H_0 is true, a sample outcome of 2.78 would be likely
  - Therefore, the GPA difference between 2.78 and 2.5 is not significant enough. H_0 could be true, we fail to reject H_0
- Report result:
  - With α equals 0.01, the obtained t score of 1.22, we fail to reject the null hypothesis. The data suggests that there is no significant difference between commuter students and the general students in terms of academic achievement.

Logic of Two-Sample Case

- One-sample case: we draw a random sample, compare it to the population, to decide if the sample mean is significantly different from the population mean.
- Two-sample case: we draw two random samples, compare the two sample means to generalize about a difference between the two respective population means.
- Example of two-sample case:
  - We may want to examine the difference in support for gun control between men and women.
  - We can draw two sets of sample, men and women, and compare the difference between the two sample means.
  - With the test result, we can decide if *all* U.S. men and *all* US women differ on this issue.

Logic of Two-Sample Case



Logic of Two-Sample Case

- In two-sample case, the H_0 always states that the two populations are the same:
  - There is no difference between the parameters (features) of the two populations

- If the difference between the sample statistics is large enough, or, if a difference of this size is *unlikely*, assuming that the H_0 is true, we will *reject* the H_0 and conclude there is a difference between the populations

The Five Step Model

Changes from the one-sample case:

- Step 1: In addition to samples selected according to EPSEM principles, samples must be selected independently of each other: **independent random sampling**
  - Independent random sampling: the two samples should be selected randomly and separately from each other. In our case, selecting a specific male would not affect the probability of selecting any particular female.
- Step 2: Null hypothesis statement will state that the two populations are not different.
- Step 3: Sampling distribution refers to **difference between the sample statistics.**

The Five Step Model

Changes from the one-sample case:

- Step 3: Sampling distribution refers to the distribution of all possible mean differences between two samples.
  - for one-sample test, z obtained $= \frac{\bar{X}-\mu}{\sigma/\sqrt{N}}$ or $\frac{\bar{X}-\mu}{S/\sqrt{N-1}}$
  - for two-sample test, z obtained= $\frac{(\bar{X}_1-\bar{X}_2)-\underbrace{(\mu_1-\mu_2)}_{=0}}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1-\bar{X}_2}{\sigma_{\bar{X}_1-\bar{X}_2}}$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{N_1}+\frac{\sigma_2^2}{N_2}} \quad \text{(when } \sigma \text{ is known)}$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_1^2}{N_1-1}+\frac{s_2^2}{N_2-1}} \quad \text{(when } \sigma \text{ is unknown)}$$

Hypothesis Test for Means

Two-Sample Test of Means (Large Samples)

- Based on the information generated by the two samples

Sample 1 (Men)	Sample 2 (Women)
Mean ($\bar{X}_1$)= 6.2	Mean ($\bar{X}_2$)= 6.5
Standard deviation (S_1)= 1.3	Standard deviation (S_2)= 1.4
Sample size (N_1)= 324	Sample size (N_2)= 317

- Questions: Do men and women significantly differ on their support of gun control? With 95% confidence level.

Two-Sample Test of Means (Large Samples)

Step 1: Make Assumptions and Meet Test Requirements

- Independent random sampling
  - The samples must be independent of each other
- Level of measurement is interval-ratio
  - Support of gun control is assessed with an interval-ratio level scale, so the mean is an appropriate statistic
- Sampling distribution is normal in shape
  - Total $N \geq 100$ ($N_1 + N_2 = 324 + 317 = 641$) so the Central Limit Theorem applies and we can assume a normal shape

Two-Sample Test of Means (Large Samples)

Step 2: State the Null and Alternative Hypothesis

- H_0: μ1 = μ2
  - There is no difference between the men and the women in terms of support of gun control.
  - The null hypothesis asserts there is no difference between the populations
- H_1: μ1 ≠ μ2
  - There is a difference between the men and the women in terms of support of gun control.
  - The alternative hypothesis contradicts the H_0 and asserts there is a difference between the populations

Two-Sample Test of Means (Large Samples)

Step 3: Select the Sampling Distribution and Establish the Critical Region

- Sampling Distribution = Z distribution
  - Because sample size (N) is large ($N = N_1 + N_2 = 324 + 317 = 641 > 100$)
  - Alpha (α) = 0.05 (two-tailed)
  - Z (critical) = ±1.96

Two-Sample Test of Means (Large Samples)

Step 4: Compute the Test Statistic

- Calculate the pooled estimate of the standard error

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_1^2}{N_1-1}+\frac{s_2^2}{N_2-1}} = \sqrt{\frac{1.3^2}{324-1}+\frac{1.4^2}{317-1}}$$

$$= \sqrt{\frac{1.69}{323}+\frac{1.96}{316}} = \sqrt{0.0052+0.0062} = \sqrt{0.0114} = 0.107$$

- Calculate the obtained Z score

$$Z(obtained) = \frac{\bar{X}_1-\bar{X}_2}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{6.2-6.5}{0.107} = \frac{-0.3}{0.107} = -2.80$$

Two-Sample Test of Means (Large Samples)

Step 5: Make a Decision

- The obtained test statistic ($Z = -2.80$) falls in the critical region so reject the null hypothesis
  - The difference between the sample means is so large that we can conclude that a difference exists between the populations represented by the samples
  - The difference between men's and women's support of gun control is **significant**.

- Report result:
  - With α equals 0.05, the obtained Z score of -2.80, we reject the null hypothesis. The data suggest that there is a difference between men's and women's support of gun control.

Hypothesis Test for Means

Two-Sample Test of Means (Small Samples)

- Based on the information generated by the two samples

Sample 1 (suburbs)	Sample 2 (center-city)
Mean ($\bar{X}_1$)= 2.37	Mean ($\bar{X}_2$)= 2.78
Standard deviation (S_1)= 0.63	Standard deviation (S_2)= 0.95
Sample size (N_1)= 42	Sample size (N_2)= 37

- Question: Do families that reside in the center-city have more children than families that reside in the suburbs? With 95% confidence level.

Two-Sample Test of Means (Small Samples)

Step 1: Make Assumptions and Meet Test Requirements

- Independent random sampling
  - The samples must be independent of each other
- Level of measurement is interval-ratio
  - Number of children can be treated as interval-ratio, so the mean is an appropriate statistic
- Sampling distribution is normal in shape
  - Because we have two small samples ($N < 100$), we have to add the next assumption in order to meet this assumption
- Population variances are equal
  - As long as the two samples are approximately the same size, we can make this assumption

Two-Sample Test of Means (Small Samples)

Step 2: State the Null and Alternative Hypothesis

- H_0: μ1 = μ2
  - There is no difference between the suburbs and the center-city in terms of number of children in a family.
  - The null hypothesis asserts there is no difference between the populations
- H_1: μ1 < μ2
  - Families that reside in the center-city have more children than families that reside in the suburbs.
  - The alternative hypothesis contradicts the H_0 and asserts there is a difference between the populations

Two-Sample Test of Means (Small Samples)

Step 3: Select the Sampling Distribution and Establish the Critical Region

- Sampling Distribution = *t* distribution (N is small)
- Alpha (α) = 0.05 (one-tailed)
- Degrees of freedom = $N_1 + N_2 - 2 = 42+37-2 = 77$
- *t* (critical) = -1.671

df	One-Tailed Test		
	0.10	0.05	0.025
40	…	…	…
60		**1.671**	…
120	…	…	…

Two-Sample Test of Means (Small Samples)

Step 4: Compute the Test Statistic

- Calculate the pooled estimate of the standard error

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}}\sqrt{\frac{N_1 + N_2}{N_1 N_2}} = \sqrt{\frac{(42)(0.63)^2 + (37)(0.95)^2}{42 + 37 - 2}}\sqrt{\frac{42 + 37}{(42)(37)}}$$

$$= \sqrt{\frac{(42)(0.40) + (37)(0.90)}{77}}\sqrt{\frac{79}{1554}} = \sqrt{\frac{50.10}{77}}\sqrt{\frac{79}{1554}} = \sqrt{0.65}\sqrt{0.05} = (0.81)(0.22)$$

$$= \mathbf{0.18}$$

- Calculate the obtained *t* score

$$t(obtained) = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{2.37 - 2.78}{0.18} = \frac{-0.41}{0.18} = -2.28$$

Two-Sample Test of Means (Small Samples)

Step 5: Make a Decision

- The obtained test statistic (t = -2.28) falls in the critical region so **reject** the null hypothesis

  t (critical) = -1.671

  - The difference between the sample means is so large that we can conclude that a difference exists between the populations represented by the samples
  - The difference between city-center and suburbs in terms of number of children in a family is **significant**.

- Report result:
  - With α equals 0.05, the obtained Z score of -2.28, we reject the null hypothesis. The data suggest families that reside in the center-city have more children than families that reside in the suburbs.

Hypothesis Test for Proportions

Two-Sample Test of Proportions (Large Samples)

- One-sample test of proportions: $Z(obtained) = \frac{P_s - P_u}{\sqrt{P_u(1-P_u)/N}}$
- Two-sample test of proportions: $Z(obtained) = \frac{(P_{s1} - P_{s2}) - \boxed{(P_{u1} - P_{u2})}}{\sigma_{p-p}}$ =0

$$Z(obtained) = \frac{(P_{s1} - P_{s2})}{\sigma_{P-P}}$$

$$\sigma_{p-p} = \sqrt{P_u(1-P_u)}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

$$P_u = \frac{N_1 P_{s1} + N_2 P_{s2}}{N_1 + N_2}$$

- σ_{p-p}=the standard deviation of the sampling distribution of the difference between sample proportions
- $P_{s1} - P_{s2}$= the difference between the sample proportions
- $P_{u1} - P_{u2}$= the difference between the population proportions

Hypothesis Test for Proportions

Two-Sample Test of Proportions (Large Samples)

- Random samples of black and white senior citizens have been selected and classified as high or low in terms of their number of memberships. The proportions of each group classified as “high” in participation and sample size for both groups are:

Sample 1 (Black senior citizens)	**Sample 2 (White senior citizens)**
Sample Proportion (P_{s1})= 0.34	Sample Proportion (P_{s2})= 0.25
Sample size (N_1)= 83	Sample size (N_2)= 103

- Questions: Do Black and White senior citizens differ in their number of memberships in clubs and organizations? With 95% confidence level.

Two-Sample Test of Proportions (Large Samples)

Step 1: Make Assumptions and Meet Test Requirements

- Independent random sampling
  - The samples must be independent of each other
- Level of measurement is nominal
  - We have measured the proportion of each group classified as having a “high” level of membership
- Sampling distribution is normal in shape
  - Total $N \geq 100$ ($N_1 + N_2 = 83 + 103 = 186$) so the Central Limit Theorem applies and we can assume a normal shape

Two-Sample Test of Proportions (Large Samples)

Step 2: State the Null and Alternative Hypothesis

- H_0: $P_{u1} = P_{u2}$
  - There is no difference between Black and White senior citizens in terms of the proportion of classified as “high” in participation memberships
  - The null hypothesis asserts there is no difference between the populations
- H_1: $P_{u1} \neq P_{u2}$
  - There is a difference between Black and White senior citizens in terms of the proportion of classified as “high” in participation memberships
  - The research hypothesis contradicts the H_0 and asserts there is a difference between the populations

Two-Sample Test of Proportions (Large Samples)

Step 3: Select the Sampling Distribution and Establish the Critical Region

- Sampling Distribution = Z distribution
  - Because sample size (N) is large ($N = N_1 + N_2 = 83 + 103 = 186 > 100$)
  - Alpha (α) = 0.05 (two-tailed)
  - Z (critical) = ± 1.96

Two-Sample Test of Proportions (Large Samples)

Step 4: Compute the Test Statistic

- Calculate the pooled estimate of the standard error
- $P_u = \frac{N_1 P_{s1} + N_2 P_{s2}}{N_1 + N_2} = \frac{(83)(0.34)+(103)(0.25)}{83+103} = 0.29$
- $\sigma_{p-p} = \sqrt{P_u(1-P_u)}\sqrt{\frac{N_1+N_2}{N_1 N_2}} = \sqrt{0.29(0.71)}\sqrt{\frac{83+103}{(83)(103)}} = (0.45)(0.15) = 0.07$
- Calculate the obtained Z score

$$Z(obtained) = \frac{(P_{s1} - P_{s2})}{\sigma_{P-P}} = \frac{0.34 - 0.25}{0.07} = 1.29$$

Two-Sample Test of Proportions (Large Samples)

Step 5: Make a Decision

- The obtained test statistic (Z = +1.29) does not fall in the critical region so ***fail to reject*** the null hypothesis
  - The difference between the sample proportions is so large that we can conclude that a difference exists between the populations represented by the samples
  - The difference between the memberships of Black and White senior citizens is **not significant**.

- Report result:
  - With α equals 0.05, the obtained Z score of 1.29, we fail to reject the null hypothesis. The data suggest that there is no difference between Black and White senior citizens in their number of memberships in clubs and organizations.

After this lecture:

You should learn the following key concepts:

- How to conduct *t* test
- The differences between one-sample and two-sample cases
- How to conduct two-sample test of means with large and small samples
- How to conduct two-sample test of proportion with large samples