# University of the People

MATH 1281 - Statistical Inference

Unit 2 Written Assignment

Liang Xiao

# MATH 1281 - Unit 2 Assignment

## Part 1a

Problem:

A group of 441 adults who did not have a college degree and were not currently enrolled in school were randomly selected. 38% said they did not attend college because they could not afford it.

Task: Conduct a hypothesis test to determine if there is strong evidence supporting the statement that less than 50% of adults who decide not to attend college do so because of affordability issues.

Step 1: State the Hypotheses

- Null Hypothesis (H0): p = 0.50

- Alternative Hypothesis (Ha): p < 0.50

Where p = true proportion of adults who decide not to attend college because they cannot afford it.

Step 2: Validate Conditions

- Independence: Random sampling is given; sample size <10% of population.

- Success-Failure Condition:

- np = 441 × 0.5 = 220.5 > 10

- n(1-p) = 441 × 0.5 = 220.5 > 10

Thus, conditions are satisfied.

Step 3: Compute Test Statistic

Observed sample proportion (p̂) = 0.38

Standard error (SE) = sqrt( p0 (1 - p0) / n ) = sqrt( 0.5 × 0.5 / 441 ) = 0.0237

Z = (p̂ - p0) / SE = (0.38 - 0.5) / 0.0237 ≈ -5.06

Step 4: Find p-value

Using the Z-table:

- p-value for Z = -5.06 is approximately 0.

Step 5: Conclusion

Since p-value < 0.05, we reject H0.

Interpretation: There is strong evidence that less than 50% of adults who do not attend college cite affordability as the reason.

---

## Part 1b

Problem:

Suppose we wanted the margin of error for a 90% confidence level to be about 1.5%. How large of a survey would you recommend?

Step 1: Identify known values

- Margin of Error (ME) = 0.015

- Z\* for 90% confidence = 1.645

- Use conservative estimate for p = 0.5 (max variability)

Step 2: Use Formula

n = (Z\* / ME)^2 × p(1-p)

n = (1.645 / 0.015)^2 × 0.5 × 0.5

n = (109.67)^2 × 0.25

n = 12023.4 × 0.25

n = 3005.85

Step 3: Conclusion

Recommended sample size = 3006.

---

## Part 2a

Problem:

A random sample study: 13,270 Texas residents and 4,681 Dallas residents.

- 7.0% (Texas) vs 6.8% (Dallas) report insufficient rest.

Task: Calculate a 95% confidence interval for the difference between proportions.

Step 1: Validate Conditions

- Independence: Samples are random and <10% of populations.

- Success-Failure:

- Texas: 13,270 × 0.07 = 928.9 successes

- Dallas: 4,681 × 0.068 = 318.3 successes

Both >10, thus condition is met.

Step 2: Compute Confidence Interval

Proportions:

- p1 (Texas) = 0.07

- p2 (Dallas) = 0.068

Difference (p1 - p2) = 0.002

Standard Error (SE) = sqrt( (p1(1-p1)/n1) + (p2(1-p2)/n2) )

SE ≈ 0.00429

Z\* = 1.96 (for 95% confidence)

Margin of Error = 1.96 × 0.00429 = 0.0084

Thus, the 95% CI is:

(0.002 - 0.0084, 0.002 + 0.0084)

(-0.0064, 0.0104)

Step 3: Interpretation

We are 95% confident that the true difference between the proportions lies between –6.4% and 1.04%. Since the interval contains 0, there is no significant difference between Texas and Dallas residents.

---

## Part 2b

Task: Conduct a hypothesis test to determine if there is strong evidence the rate of sleep deprivation differs between Texas and Dallas.

Step 1: State the Hypotheses

- H0: p1 - p2 = 0

- Ha: p1 - p2 ≠ 0

Step 2: Test Statistic

Pooled proportion (p̂):

p̂ = (928.9 + 318.3) / (13270 + 4681) ≈ 0.0695

Standard Error (pooled) ≈ 0.00424

Z = (p1 - p2) / SE = (0.002) / 0.00424 ≈ 0.472

Step 3: Find p-value

For Z = 0.472, two-tailed p-value ≈ 0.637.

Step 4: Conclusion

Since p-value > 0.05, we fail to reject H0.

Interpretation:

There is no statistically significant difference in the sleep deprivation rates between Texas and Dallas residents.