STAT 2593

Lecture 037 - Inferences Concerning a Difference Between Population Proportions

Dylan Spicker

Inferences Concerning a Difference Between Population Proportions

Learning Objectives

1. Construct hypothesis tests and confidence intervals for two sample tests of proportions.

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.
- What if we have two samples from binomial distributions?

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.
- What if we have two samples from binomial distributions?
  - X comes from a $\text{Bin}(n, p_1)$ distribution.

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.
- What if we have two samples from binomial distributions?
  - X comes from a $\text{Bin}(n, p_1)$ distribution.
  - Y comes from a $\text{Bin}(m, p_2)$ distribution.

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.
- What if we have two samples from binomial distributions?
  - X comes from a $\mathrm{Bin}(n, p_1)$ distribution.
  - Y comes from a $\mathrm{Bin}(m, p_2)$ distribution.
  - X and Y are independent.

Two Sample Proportion Tests

- We have currently assumed sample means on continuous data from (approximately) normally distributed populations.
- What if we have two samples from binomial distributions?
  - X comes from a $\text{Bin}(n, p_1)$ distribution.
  - Y comes from a $\text{Bin}(m, p_2)$ distribution.
  - X and Y are independent.
- We are interested in the difference in proportions, $p_1 - p_2$.

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.
  - This will have variance
  $$\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}.$$

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.
  - This will have variance
$$\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}.$$
- As long as the normal approximation would apply for each sample individually, we can use the normal approximation here for confidence intervals and hypothesis tests.

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.
  - This will have variance
$$\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}.$$
- As long as the normal approximation would apply for each sample individually, we can use the normal approximation here for confidence intervals and hypothesis tests.
  - Confidence intervals are constructed in exactly the expected way.

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.
  - This will have variance
$$\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}.$$
- As long as the normal approximation would apply for each sample individually, we can use the normal approximation here for confidence intervals and hypothesis tests.
  - Confidence intervals are constructed in exactly the expected way.
  - Hypothesis tests only work if we test $H_0 : p_1 = p_2$.

Estimating the Difference in Proportions

- The estimator given by $\widehat{p}_1 - \widehat{p}_2$ will be unbiased for $p_1 - p_2$.
  - This will have variance
$$\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}.$$
- As long as the normal approximation would apply for each sample individually, we can use the normal approximation here for confidence intervals and hypothesis tests.
  - Confidence intervals are constructed in exactly the expected way.
  - Hypothesis tests only work if we test $H_0 : p_1 = p_2$.
- Under the assumption $p_1 = p_2 = p$, we have variance $p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)$

Summary

- Two populations from independent binomial distributions can have their proportions tested through a difference in sample proportions, so long as we assume independence.

- The approximation is only valid enough for the null hypothesis of equality.