10E 2SamProp

Contents

Textbook: Devore 8e

E Two-Sample Inference on Proportion

[ToC]

E.1 Do the same with $\hat{p}_1$ and $\hat{p}_2$.

We know that

$$\hat{p}_1 \sim N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \qquad \text{and} \qquad \hat{p}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

Then, since they are independent,

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \quad \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

Again, this looks kinda like in one-sample case,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

E.2 Two-sample CI for $p_1 - p_2$.

From above characteristics, $100(1-\alpha)\%$ Confidence Interval for $\mu_1 - \mu_2$ can be derived as

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

E.3 Two-sample test for $p_1 - p_2$.

To test the null hypothesis of $H_0 : p_1 - p_2 = 0$ against alternatives

$$\begin{aligned} H_A : & p_1 - p_2 > 0 && \text{(Upper-tailed alternative)} \\ H_A : & p_1 - p_2 < 0 && \text{(Lower-tailed alternative)} \\ H_A : & p_1 - p_2 \neq 0 && \text{(Two-tailed alternative)} \ , \end{aligned}$$

Use the test statistic

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\hat{p}_1 = \frac{X}{n_1} \qquad \hat{p}_2 = \frac{Y}{n_2} \qquad \hat{p} = \frac{X+Y}{n_1+n_2}.$$

This is a z-test.

E.4 Ex: Polio vaccine

1954 Salk polio-vaccine double-blind experiment.

Out of 201,229 people who was not vaccinated, 110 got polio.
Out of 200,745 people who was not vaccinated, 33 got polio.

E.5 Polio vaccine numbers

1954 Salk polio-vaccine double-blind experiment.

Out of 201,229 people who was not vaccinated, 110 got polio.
Out of 200,745 people who was not vaccinated, 33 got polio.

$$\hat{p}_1 = \frac{110}{201,229} = 0.00054664, \quad \hat{p}_2 = \frac{33}{200,745} = 0.00016438$$

Is this a significant difference?

E.6 (Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$ Perform z-test with

E.7 (Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$ Perform z-test with

$$\hat{p} = \frac{33 + 110}{200,745 + 201,229} = 0.00035574.$$

$$\hat{p}_1 = 0.00054664, \qquad n_1 = 201,229$$

$$\hat{p}_2 = 0.00016438 \qquad n_2 = 200,745$$

E.8 (Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$ Perform z-test with

$$\hat{p} = \frac{33 + 110}{200,745 + 201,229} = 0.00035574.$$

$$\hat{p}_1 = 0.00054664, \qquad n_1 = 201,229$$

$$\hat{p}_2 = 0.00016438 \qquad n_2 = 200,745$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = -6.4258$$

E.9 Ex: Two methods to Check

Suppose that method 1 resulted in 20 unacceptable transistors out of 100 produced; whereas method 2 resulted in 12 unacceptable transistors out of 100 produced. Can we conclude from this, at the 10 percent level of significance, that the two methods are equivalent?