Key

Statistics
Notes for 10.4
Proportions

*Some things to remember from Chapter 9*

*A. Sample Size < 1/10 of Population*

*B.* $np \geq 10$ *and* $nq \geq 10$

1. Time magazine reported the result of a telephone poll of 800 adult Americans. The question posed of the Americans who were surveyed was: "Should the federal tax on cigarettes be raised to pay for health care reform?" The results of the survey were:

| Non Smokers | Smokers |
|---|---|
| N = 605 | N = 195 |
| 361 said "yes" | 41 said "yes" |

Is there significant evidence to conclude that the proportion of Non-Smokers that support raising the tax is different than the proportion of Smokers. Test at the level of significance of 5%.

A. Is this a hypothesis test or confidence interval?

TEST

B. Is this a proportions or average problem? How can you tell?

Proportions, there is no average or standard deviation in the problem

C. What is the assumption for the Null Hypothesis?

Both of the proportions for the populations are equal

D. What is the Alternative Hypothesis for this situation?

$P_{ns} \neq P_s$ (The question states "... "is different..."

E. Write up Step 1. Include all 3 requirements.

$H_0: P_{ns} = P_s$

$P_{ns}$ → proportion of non-smokers that support the tax

$P_s$ → proportion of smokers that support the tax

$H_a: P_s \neq P_s$

$\alpha = .05$

F. What are the 3 major components that must be checked for Step 2?

Random, Size, Distribution

G. In one of the components for Step 2, I've asked you to multiply the sample size by 10. What is the product (the number multiplied by 10) compare to?

You compare that value to the population. The product should be smaller than the population

H. In a hypothesis test between two proportions we POOL the proportions together. How is that value found? (Equation)

$$P_p = \frac{x_1 + x_2}{n_1 + n_2} \rightarrow \frac{361 + 41}{605 + 195} = .5025$$

I. Once the pooled proportion is found, what do you multiply it by? In addition, once the product is found what value do you compare it to?

You multiply $P_p$ by each sample size to check the distribution $n_1 P_p \geq 10$ $n_2 P_p \geq 10$ $n_1 q_p \geq 10$ $n_2 q_p \geq 10$

J. What is the compliment (opposite) of Proportion Pooled ($P_p$) and how is it found?

$$q_p = 1 - P_p \rightarrow 1 - .5025 = .4975$$

K. Complete the table below for Step 2

| | Non-Smokers | Smokers |
|---|---|---|
| 1. Rand | No | No |
| 2. Size | $10(605) = 6050$ * There is more than 6050 non-smokers ... good | $10(195) = 1950$ * There is more than 1950 smokers |
| 3. Dist | $605(.5025) = 304.0125$ $605(.4975) = 300.9875$ ✓ good | $195(.5025) = 97.9875$ $195(.4975) = 97.0125$ ✓ good |

L. What are the 4 required items in Step 3 in a hypothesis test?

curve, work, p-val, result.

M. What is the calculator function used for this hypothesis test?

2 prop z-test

N. When you create the curve for this step, what value goes in the middle of the curve?

0

O. On your calculator screen, define the following variables

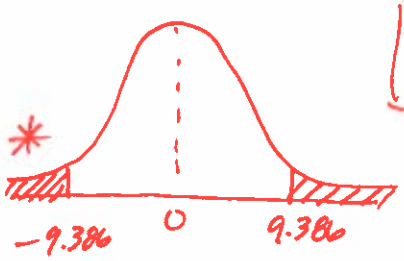
| | |
|---|---|
| $p$ | p-val |
| $\hat{p}_1$ | proportion for sample 1 |
| $\hat{p}_2$ | proportion for sample 2 |
| $\hat{p}$ | proportion pooled |

← I put ns here

← I put s here

P. Complete all the components for Step 3 below.

* Since $H_a: p_{ns} \neq p_s$, must shade both sides




$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p_p \cdot q_p \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \Rightarrow \frac{.5967 - .2103}{\sqrt{(.3025)(.4975)\left(\frac{1}{605} + \frac{1}{195}\right)}} = 9.386$$

p-val = $6.356 \times 10^{-21}$ (very small)

p-val $< \alpha$ Reject

Q. When can the null hypothesis be rejected?

when p-val $< \alpha$

R. What is your conclusion after this test? (Step 4)

I do have sufficient evidence to reject the Null Hypothesis. I can say that the proportion of non-smokers that support the tax is different than smokers that support the tax.

The test could be biased because both samples were not obtained randomly

2. According to the Insurance Institute for Highway Driving, the percentage of male drivers (age 16-19) that died in an auto accident is much lower now than it was in the 1970's. The data is below.

| Year | Total number of Teen Fatalities | Total number of Male Fatalities |
|---|---|---|
| 1975 | 8,748 | 6,532 |
| 2016 | 2,820 | 1,858 |

Estimate with 95% confidence the difference in proportion between male drivers (age 16-19) that were in a fatal accident in 1975 against 2016.

A. Is this a Hypothesis Test or Estimation?

B. What ~~are the two words you need~~ phase is necessary to include in your statement if you are doing an estimation for this problem?

difference in proportion

C. What are the two population proportions you are comparing in this problem?

proportion of male teen drivers that passed away in the years of 1975 and 2016

D. Do Step 1 for this problem

I am estimating with 95% confidence the difference in proportion for male teen drivers that died in an auto accident between the years of 1975 and 2016.

E. What are the 3 required items to check in Step 2?

Random, Size, Distribution

F. Step 2 for Hypothesis Test and Estimation is very similar. What is the main difference in Step 2 that separates Hypothesis Testing and Estimation?

TEST

$np_0 \geq 10$

$nq_0 \geq 10$

Estimation

$n\hat{p} \geq 10$

$n\hat{q} \geq 10$

G. Complete Step 2 for this problem.

| | 1975 | 2016 |
|---|---|---|
| 1 Rand | No | No |
| 2 Size | $8748 < \frac{1}{10}$ population? * No, 8748 is the population | $2820 < \frac{1}{10}$ population? * No, 2820 is the population |
| 3 Dist | $n_1\hat{p}_1 \geq 10$ and $n_1\hat{q}_1 \geq 10$<br>$8748\left(\frac{6532}{8748}\right) = 6532$ ✓<br>~~2820~~ $8748\left(\frac{2216}{8748}\right) = 2216$ Normal | $n_2\hat{p}_2 \geq 10$ and $n_2\hat{q}_2 \geq 10$<br>$2820\left(\frac{1858}{2820}\right) = 1858$<br>$2820\left(\frac{962}{2820}\right) = 962$ |

I. What are the 2 components necessary for Estimation between 2 proportions?

work + result

J. What is the equation for Estimation between 2 proportions?

$$\hat{p}_1 - \hat{p}_2 \pm z^* \cdot \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

K. Regarding the up arrow and down arrow for the Result, what must you do to determine the direction of the arrows for your population of choice?

you must know which way you are subtracting. For this problem I am...

$\hat{p}_{1975} - \hat{p}_{2016}$

L. Complete Step 3

$$(.7467 - .6567) \pm 1.96 \sqrt{\frac{(.7467)(.2533)}{8748} + \frac{(.6567)(.3433)}{2820}}$$

$.0702$ to $.1097$

1975↑ 1975↑

M. Now complete Step 4 based on your result from Step 3. What must you always include in Step 4...if it exists?

I am 95% confident that the difference in proportion between the teen fatalities for auto accidents for the years 1975 and 2016 is .0702 to .1097 higher for 1975.