Introduction to Inference

Example With One-Proportion z-Test

1

Example

- In a given year, 13.5% of employed people in the US reported belonging to a union. Officials from a large city contacted a random sample of 2000 city workers and 240 claimed union membership. Is there sufficient evidence to conclude that the proportion of workers in this city who are union members is different from the national rate?

2

Example

Step 1: Identify population **Parameter**, state the null and alternative **Hypotheses**, determine what you are trying to do (and determine what the question is asking).

We want to know whether the proportion of union workers in a particular city is different from the national rate. The **parameter** of interest is the **proportion of union workers in a particular city**. The population is all workers in this particular city. We assume that the proportion is the same as the national rate.

$H_0: p = 0.135$ The proportion of union members is 0.135

$H_A: p \neq 0.135$ The proportion of union members is NOT 0.135

3

Example

Step 2: Verify the **Assumptions** by checking the conditions

Independence Assumption

•**Randomization condition:** We are told that the sample was randomly selected.

•**10% condition:** We can reasonably assume that we observed fewer than 10% of all workers in the city; although if there are not more than 20,000 workers in the city, our results may not be valid.

•**Plausible independence condition:** There is no reason to believe that independence is violated

4

Example

Step 2: Verify the **Assumptions** by checking the conditions

Normal Assumption (Large Enough Sample Assumption)

•**Success/Failure condition:**

$$np_0 \geq 10 \text{ and } nq_0 \geq 10$$
$$np_0 = 2000(.135) = 270 \geq 10$$
$$nq_0 = 2000(.865) = 1730 \geq 10$$

5

Example

Step 3: If conditions are met, **Name the inference procedure,** find the **Test statistic**, and **Obtain the p-value** in carrying out the inference:

Since the conditions are satisfied, it is appropriate to model the sampling distribution with the Normal distribution with ~N(p_0, SE)

- We will use a one-proportion z-test **(this is the name of the inference)**

$$SE_{\hat{p}} = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{(.135)(.865)}{2000}} \approx .0076$$
$$\hat{p} = \frac{240}{2000} = .12$$

6

Example

Step 3: If conditions are met, **Name the inference procedure,** find the **Test statistic**, and **Obtain the p-value** in carrying out the inference:

Test Statistic: $z = \frac{\hat{p} - p_0}{SE_p} = \frac{0.12 - 0.135}{0.0076} \approx -1.97$

p-Value: $P(z < -1.97 \text{ or } z > 1.97) \approx 2(0.0244) = 0.0488$

7

Example

Make a decision (reject or fail to reject H_0). **State your conclusion** in context of the problem using p-value.

The p-value is small enough to reject the null hypothesis in favor of the alternative at the 0.05 alpha level. We reject the null hypothesis **(this is our decision)**. There is sufficient evidence, p-value = 0.049, to conclude that the true proportion of workers who are union members is different from the national value **(conclusion)**.

8

Example

- A company claims to have invented a hand-held sensor that can detect the presence of explosives inside a closed container. Law enforcement and security agencies are very interested in purchasing several of the devices if they are shown to perform effectively. An independent laboratory arranged to perform a study. They placed 4 boxes in a room and randomly placed explosives in one of the boxes. Out of 50 trials, they successfully found the explosives 16 times. Does this indicate that the device is effective in sensing the explosives?

9

Example

Step 1: Identify population **Parameter**, state the null and alternative **Hypotheses**, determine what you are trying to do (and determine what the question is asking).

We want to know whether the proportion found explosives is the same as guessing. The **parameter** of interest is the proportion of detecting explosives. The population is all explosives in a box. We assume that the proportion is the same as guessing.

$H_0 : p = 0.25$ The true proportion is the same as guessing

$H_A : p > 0.25$ The true proportion is better than guessing

10

Example

Step 2: Verify the **Assumptions** by checking the conditions

Independence Assumption

•**Randomization condition:** We are told that the sample was randomly selected.

•**10% condition:** We can reasonably assume that we observed fewer than 10% of explosives that can be in a box.

•**Plausible independence condition:** There is no reason to believe that independence is violated

11

Example

Step 2: Verify the **Assumptions** by checking the conditions

Normal Assumption (Large Enough Sample Assumption)

•**Success/Failure condition:**

$$np_0 \ge 10 \text{ and } nq_0 \ge 10$$
$$np_0 = 50(.25) = 12.5 \ge 10$$
$$nq_0 = 50(.75) = 37.5 \ge 10$$

12

Example

Step 3: If conditions are met, **Name the inference procedure,** find the **Test statistic**, and **Obtain the p-value** in carrying out the inference:

Since the conditions are satisfied, it is appropriate to model the sampling distribution with the Normal distribution with ~N(p_0, SE)

- We will use a one-proportion z-test **(this is the name of the inference)**

$$SE_{\hat{p}} = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{(.25)(.75)}{50}} \approx .0612$$

$$\hat{p} = \frac{16}{50} = .32$$

13

Example

Step 3: If conditions are met, **Name the inference procedure,** find the **Test statistic**, and **Obtain the p-value** in carrying out the inference:

Test Statistic: $z = \frac{\hat{p} - p_0}{SE_{\hat{p}}} = \frac{0.32 - 0.25}{0.0612} \approx -1.14$

p-Value: $P(z < -1.97 \;\; or \;\; z > 1.97) \approx 2(0.0244) = 0.13$

14

Example

Step 4: Make a decision (reject or fail to reject H_0). **State your conclusion** in context of the problem using p-value.

The p-value is not small enough to reject the null hypothesis in favor of the alternative at the 0.05 alpha level. We fail to reject the null hypothesis **(this is our decision)**. There is insufficient evidence, p-value = 0.13, to conclude that the true proportion of detected explosive is different from guessing **(conclusion)**.

15