

STATISTICS FOR DATA SCIENCE HYPOTHESIS and INFERENCE

Dr. Deepa NairDepartment of Science and Humanities



UNIT-4 HYPOTHESIS and INFERENCE

Session-6

Tests for a Population Proportion

Dr. Deepa Nair

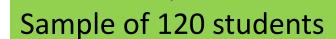
Department of Science and Humanities

Tests for a Population Proportion

PES UNIVERSITY

Example:

More than 85% of the students across the campuses participated in the revision session taken on 7/10/2020





Found 75 attended



Can we accept the claim?

Tests for a Population Proportion



The Sample Size Must Be Large:

• The test just described requires that the sample proportion be approximately normally distributed.

Tests for a Population Proportion



- This assumption will be justified whenever both $n\,p_0>10$ and $n(1-p_0)>10$. where p_0 is the population proportion specified in the null distribution.
- Then the z —score can be used as the test statistic, making this a z test.

Tests for a Population Proportion



- Let X be the number of successes in n independent Bernoulli trials, each with success probability p; in other words, let $X \sim Bin(n,p)$.
- Null distribution of \widehat{p} , $\widehat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$
- p is approximated by using p_0 value.
- To test a null hypothesis of the form

$$H_0: p \leq p_0, H_0: p \geq p_0, or H_0: p = p_0,$$

•

Tests for a Population Proportion



- Assuming that both np_0 and $n(1-p_0)$ are greater than 10:
- Compute the *z*-score:

$$z = \frac{\widehat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Tests for a Population Proportion

Compute the P —value.

• The *P* —value is an area under the normal curve, which depends on the alternate hypothesis as follows:



Tests for a Population Proportion



Alternate Hypothesis	P-value
$H_1: p > p_0$	Area to the right of z
$H_1: p < p_0$	Area to the left of z
H_1 : $p = p_0$	Sum of the areas in the tails cut off by z and $-z$

Tests for a Population Proportion



Example:

• The article "Refinement of Gravimetric Geoid Using GPS and Leveling Data" (W. Thurston, *Journal of Surveying Engineering*, 2000:27–56) presents a method for measuring orthometric heights above sea level.

Tests for a Population Proportion



Example:

- For a sample of 1225 baselines, 926 gave results that were within the class C spirit leveling tolerance limits.
- Can we conclude that this method produces results within the tolerance limits more than 75% of the time?

Tests for a Population Proportion



Solution:

$$\alpha = 0.05 \ H_0: \ p \le 0.75 \ versus \ H_1: \ p > 0.75$$

$$\widehat{p} = \frac{926}{1225} = 0.7559 \quad n = 1225$$

$$z = \frac{\widehat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

$$z = \frac{0.7559 - 0.7500}{0.0124}$$

$$= 0.48$$

Tests for a Population Proportion



Solution:

- The P -value is 0.3156 > 0.05.
- We cannot conclude that the method produces good results more than 75% of the time.

Tests for a Population Proportion



Example:

- A commonly prescribed drug for relieving nervous tension is Believed to be only 60% effective.
- Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief.
- Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance.

Tests for a Population Proportion



Soliution:

$$H_0: p = 0.6$$

$$H_1: p > 0.6$$

$$\hat{p} = \frac{70}{100} = 0.7$$
 $n = 100$
$$z = \frac{0.70 - 0.6}{\sqrt{\frac{0.6 * 0.4}{100}}} = 2.04$$

$$p(z > 2.04) = 0.0207 < 0.05$$

So reject H_0 and conclude that the new drug is superior.

Tests for a Population Proportion



Example:

- If in a random sample of 600 cars making a right turn at a certain traffic junction 157 drove into the wrong lane, test whether actually 30% of all drivers make this mistake or not at this given junction.
- Use (a) 0.05 (b) 0.01 L.O.S.

Tests for a Population Proportion



Solution:

$$H_0: p = 0.3$$
 $H_1: p \neq 0.3$ $\widehat{p} = \frac{157}{600} = 0.262$ $n = 600$ $z = \frac{0.30 - 0.7}{\sqrt{\frac{0.3 \times 0.7}{600}}} = -2.03$

$$p(z < -2.03) = 0.0212$$

So the P value is 0.0424

When
$$\alpha=0.05$$
 , $0.0424<0.05$ we need to reject H_0

When
$$lpha=0.01$$
 , $\ 0.0424>0.01$ we need to reject H_0

Tests for a Population Proportion



Relationship with Confidence Intervals for a Proportion:

• A level $100(1-\alpha)\%$ confidence interval for a population mean μ contains those values for a parameter for which the P-value of a hypothesis test will be greater than α .

Tests for a Population Proportion



Relationship with Confidence Intervals for a Proportion:

- For the confidence intervals for a proportion presented earlier and the hypothesis test presented here, this statement is only approximately true.
- The reason for this is that the methods presented earlier are slight modifications (that are much easier to compute) of a more complicated confidence interval method for which the statement is exactly true.



Dr. Deepa Nair

Department of Science and Humanities

deepanair@pes.edu