One sample Z-test for population proportion

Function of the test: The function of this test is to compare observed (or sample) proportion of success with the theoretical proportion of success i.e., to test whether population proportion of success is significantly different from some hypothesized value.

Test Assumptions:

1. Variable of interest is categorical and having two categories one of which is called success and other is called failure.
2. Sample is drawn randomly from population
3. Sample size is relatively large so that sampling distribution of sample proportion of success is approximately normally distributed.

Hypothesis to test

Let π be the true proportion of success in the population and let π_0 be the hypothesized value.

Then null and alternative hypothesis for the test is given below:

Null Hypothesis	Alternative Hypothesis	No. of tails/rejection regions
H_0: $\pi = \pi_0$	H_1: $\pi \neq \pi_0$	Two
H_0: $\pi \geq \pi_0$	H_1: $\pi < \pi_0$	One (left sided)
H_0: $\pi \leq \pi_0$	H_1: $\pi > \pi_0$	One (right sided)

Test Statistic

The appropriate test statistic under H_0 (If H_0 is true) is given by,

$$Z = \frac{p - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

Where,

p = sample proportion of success = X/n

X = No. of successes in sample

n = sample size

π_0 = Hypothesized value of proportion of success

1- π_0 = Hypothesized value of proportion of failure

The quantity $\sqrt{\pi_0(1-\pi_0)/n}$ is called standard error of sample proportion of success 'p'

Distribution of test statistic

The test statistic is distributed as Z i.e. Standard Normal Distribution with mean = 0 and variance = 1.

Decision Rule

We will adopt following decision rule.

Hypothesis	
Case I (Two-sided test)	Reject H_0 if $\lvert \text{cal } Z \rvert \geq Z_{\alpha/2}$ i.e. Accept H_0 if $-Z_{\alpha/2} < \text{cal } Z < +Z_{\alpha/2}$ otherwise reject
Case II (Left sided test)	Reject H_0 if cal $Z \leq -Z_\alpha$
Case III (Right sided test)	Reject H_0 if cal $Z \geq +Z_\alpha$

Problem:

An e-commerce research company claims that 60% or more graduate students have bought merchandise on-line. A consumer group is suspicious of the claim and thinks that the proportion is lower than 60%. A random sample of 80 graduate students show that only 22 students have ever done so. Is there enough evidence to show that the true proportion is lower than 60%? Conduct the test at 5 % Type I error rate, and use the p-value and rejection region approaches.

Solution:

Population is divided into two groups

- college graduate who bought merchandise online (success)
- college graduate who does not bought merchandise online (failure)

Step 1: Null and alternative hypothesis

Let π be the true proportion of success in the population and let π_0 be the hypothesized value.

Null Hypothesis

H_0: $\pi \geq 60\ \%\ or\ 0.60$ (proportion of graduates who bought merchandise online is at least 60 %)

Alternative Hypothesis

H_1: $\pi < 60\ \%\ or\ 0.60$ (proportion of graduates who bought merchandise online is lower than 60 %)

Step 2: Level of significance for the test (choice of α for the test)

The level of significance for the test = prob of type I error = 5 %

Step 3: Test statistic

The appropriate test statistic under H_0 (If H_0 is true) is given by,

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

Where,

p = sample proportion of success

n = sample size

π_0 = Hypothesized value of proportion of success

1- π_0 = Hypothesized value of proportion of failure

The quantity $\sqrt{\pi_0(1-\pi_0)/n}$ is called standard error of sample proportion of success ‘p’

Distribution of test statistic

The test statistic is distributed as Z i.e. Standard Normal Distribution with mean = 0 and variance = 1.

Step 4 : Calculated Z or Observed Z

X = no. of graduate who bought merchandise online in sample = 22

n = no. of graduate students in the sample = 80

p = proportion of graduate students who bought merchandise on line in the sample = X/n = 22/80

= 0.275

q = 1 – p = proportion of graduate students who did not bought merchandise on line in the sample

= 1 - 0.275 = 0.725

The calculated Z is given by,

$$Z = \frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.275-0.60}{\sqrt{\frac{0.60x0.40}{80}}} = -5.933$$

Step 5: Critical Z or tabulated Z

The test is left sided and given level of α employed for the test is 5 %. Hence, the rejection region lies in the left side of Z curve.

The critical value of z from Z table is given by

$Z_C = -1.65$

AR: $Z \geq -1.65$

RR: $Z < -1.65$

Statistical decision: Reject H0 if cal Z is lower than – 1.65 i.e. when cal Z falls in the rejection region

Step 6: Statistical Decision

Since, cal Z = - 5.933 < critical z = - 1.65, we reject H_0 at 5 % level of significance, in favor of alternative hypothesis H_1

Step 7: Conclusion

Proportion of graduates who bought merchandise online is lower than 60 % i.e. claim of ecommerce company is not valid.