Testing of Hypothesis: Large Sample Tests

Lesson 3: Test of Hypothesis for Single Population Proportion p

For an infinite population, let p be the proportion (or probability) of success and q = (1 – p) is the proportion (or probability) of failure. Let us consider that all possible samples of size n are drawn from the population. Then the sampling distribution of proportions (R) will have mean $\mu_R = p$ and standard deviation $\sigma_R = \sqrt{\frac{pq}{n}}$, where $R = \frac{x}{n}$ is the proportion of success in the sample [x = number of favorable cases for the concerned event].

For a large sample, though the population is binomially distributed, the sampling distribution will be normally distributed with test statistic

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{R - p}{\sqrt{\frac{pq}{n}}} - - - - (1)$$

The steps for testing are:

(i) Formulate null hypothesis, $H_0 : p = p_1$: There is no significant difference between sample proportion and population proportion

(ii) Formulate alternative hypothesis, $H_1 : p \neq p_1$ (or otherwise) : There is a significant difference between sample proportion and population proportion

(iii) Level of significance is : α

(iv) From the two tailed test, the critical region is : $|z| \geq z_{\alpha/2}$

(v) Compute the test statistic z from equation (1)

(vi) Reject H_0 if the computed value of z falls in the critical region, otherwise accept H_0

Note:

If population size is finite (N), then $\mu_R = p$ and $\sigma_R = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$

Problems:

Ex.1. In a sample of 400 parts manufactured by a factory, the number of defective parts is found to be 30. The company however claims that only 5% of their product is defective. Is the claim plausible? (Test at 5% level)

Solution:

Let us define that the event 'the parts manufactured by the factory is defective' is a success.

The sample size n = 400 (≥ 30), hence large sample test is to be applied

The sample proportion R = $\frac{30}{400}$

The population proportion p = 0.05, since 5% of the population is defective

(i) Null hypothesis, H_0 : p = 0.05, which means 5% of the population is defective, i.e, the claim is plausible

(ii) Alternative hypothesis, H_1 : p > 0.05, which means that the actual proportion of defective parts is more than 5%, i.e, the claim is not plausible

(iii) α = 0.05

(iv) The critical region at 5% level of significance is $z \geq 1.645$

(v) $z = \frac{\mathrm{R-p}}{\sqrt{\frac{\mathrm{pq}}{\mathrm{n}}}} = \frac{\frac{3}{40} - 0.05}{\sqrt{\frac{0.05 \times 0.95}{400}}} = 2.29$

(vi) As computed z > 1.645, i.e, the test statistic lies in the critical region, so null hypothesis H_0 is rejected

Hence, we conclude that the claim is not plausible and the actual proportion of the defective parts is more than 5%.

Ans.

Ex.2. In a big city, 325 men out of 600 were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers?

Solution:

Let us define that the event 'the number of male smokers in the city' is a success.

The sample size n = 600 (≥ 30), hence large sample test is to be applied

The sample proportion R = $\frac{325}{600}$

The population proportion is taken as p = 0.5, since we are to test whether the majority (> 50%) of men in the city are smokers

(i) Null hypothesis, H_0 : p = 0.5, which means that only 50% men in the city are smokers

(ii) Alternative hypothesis, H_1 : p > 0.5, i.e, more than 50% of men in the city are smokers

(iii) α = 0.05 [*since nothing is mentioned, 5% significance level is chosen]

(iv) Critical region at 5% level of significance is $z \geq 1.645$

(v) $z = \frac{\mathrm{R-p}}{\sqrt{\frac{\mathrm{pq}}{\mathrm{n}}}} = \frac{\frac{325}{600} - 0.5}{\sqrt{\frac{0.5\times 0.5}{600}}} = 2.043$

(vi) Since computed z > 1.645, i.e, z lies in the critical region, so null hypothesis H_0 is rejected

Hence we conclude that the given information from the sample supports that the majority of men in the city are smokers.

Ans.

Ex.3. During testing in a sample of 300 chips, 10 have been found to be defective. Can the manufacturers claim that "only 2% of the chips are defective" be accepted at 1% significant level?

Solution:

Let us define that the event 'the chips manufactured by the factory is defective' is a success.

Here, the sample size n = 300 (≥ 30), hence large sample test is to be applied

The sample proportion R = $\frac{10}{300}$

The population proportion p = 0.02

(i) Null hypothesis, H_0 : p = 0.02, the claim can be accepted

(ii) Alternative hypothesis, H_1 : p $>$ 0.02, i.e, the proportion of defective chips in the population is actually more than 2%. That means the claim cannot be accepted

(iii) α = 0.01

(iv) Critical region at 1% level of significance is $z \geq 2.33$

(v) $z = \frac{\mathrm{R-p}}{\sqrt{\frac{\mathrm{pq}}{\mathrm{n}}}} = \frac{\frac{10}{300} - 0.02}{\sqrt{\frac{0.02\times 0.98}{300}}} = 1.649$

(vi) Since computed $z < 2.33$, i.e, the test statistic does not lie in the critical region, hence H_0 is accepted

Therefore, we conclude that the claim can be accepted.

Ans.