Testing of Hypothesis: Large Sample Tests

Lesson 1: Test of Hypothesis for a Single Population Mean μ (population variance σ^2 is known)

Let $\bar{x}$ is the mean of a sample of size n (≥ 30). Then by the Central Limit Theorem, we know that the sampling distribution of $\bar{x}$ is approximately normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$ and the test statistic for this case is given by

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} - - - - - \;\; - - (1)$$

The steps for testing are:

(i) Formulate null hypothesis, $H_0 : \mu = \mu_1$: There is no significant difference between sample mean and population mean

(ii) Formulate alternative hypothesis, $H_1 : \mu \neq \mu_1$ (or otherwise) : There is significant difference between sample mean and population mean

(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:

1. If σ is unknown, then it can be replaced by sample standard deviation s for large sample tests

2. If samples are drawn from a finite population of size N, then the sampling distribution of $\bar{x}$ will have mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$, where $\sqrt{\frac{N-n}{N-1}}$ is called the finite population correction factor

Problems:

Ex.1. A college claims that its average size of class is 35 students. A random sample of 64 classes has a mean size of 37 students with a standard deviation of 6 students. Test at α = 0.05 level of significance if the claim is too low.

Solution:

Here the population mean μ = 35

The sample size n = 64 (≥ 30), hence large sample test is to be applied

The sample mean $\bar{x}$ = 37

The sample standard deviation s = 6

Since the population standard deviation σ is unknown, hence it will be replaced by sample standard deviation s in computation of the z statistic.

(i) Null hypothesis, H_0 : μ = 35

(ii) Alternative hypothesis, H_1 : μ > 35 [since we have to check whether the claim is too low, hence the alternative hypothesis should be that the mean is actually more than 35]

(iii) α = 0.05 (5% level)

(iv) For right one-tailed test, the critical region at 5% level of significance is $z \geq 1.645$

(v) $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{37-35}{\frac{6}{\sqrt{64}}} = 2.67$

(vi) As computed z is more than 1.645, i.e, the computed value of z falls in the critical region, hence null hypothesis H_0 is rejected

Hence, we conclude that the claim is too low and the average class size of the college is likely to be more than 35.

Ans.

Ex.2. A machine runs on an average of 125 hours/year. A random sample of 49 machines has an annual average use of 126.9 hours with standard deviation 8.4 hours. Does this suggest to believe that machines are used on the average more than 125 hours annually at 5% level of significance?

Solution:

Here the population mean $\mu = 125$

The sample size $n = 49$ (≥ 30), hence large sample test is to be applied

The sample mean $\bar{x} = 126.9$

The sample standard deviation $s = 8.4$

Since the population standard deviation σ is unknown, it will be replaced by sample standard deviation s in computation of the z statistic.

(i) Null hypothesis, $H_0 : \mu = 125$

(ii) Alternative hypothesis, $H_1 : \mu > 125$

(iii) $\alpha = 0.05$

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

(v) $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{126.9-125}{\frac{8.4}{\sqrt{49}}} = 1.583 < 1.645$

(vi) As computed z is less than 1.645, i.e, the computed value of z does not fall in the critical region, hence null hypothesis H_0 is accepted

Hence we conclude that the data does not suggest to believe that machines are used on the average more than 125 hours annually.

Ans.

Ex.3. Sugar is packed in bags by an automatic machine with mean content of bags as 1Kg. A random sample of 36 bags is selected and mean mass has been found to be 1.003 Kg. If a standard deviation of 0.01 Kg is acceptable on all the bags being packed, determine on the basis of sample test whether the machine requires adjustment.

Solution:

Here the population mean $\mu = 1$

The population standard deviation $\sigma = 0.01$

The sample size n = 36 (≥ 30), hence large sample test is to be applied

The sample mean $\bar{x} = 1.003$

(i) Null hypothesis, $H_0 : \mu = 1$

(ii) Alternative hypothesis, $H_1 : \mu \neq 1$

(iii) $\alpha = 0.05$ [since nothing is mentioned, we take 5% level of significance]

(iv) Critical region at 5% level of significance is $|z| \geq 1.96$

(v) $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1.003-1}{\frac{0.01}{\sqrt{36}}} = 1.8$

(vi) Since $|z| < 1.96$, so it does not fall in the critical region. Hence H_0 is accepted

Therefore we conclude that the machine does not require any adjustment.

Ans.