Lab Problem – 01 The managing director of a ~~fiml~~ firm claims that his firm produces 110 items on average daily. A random sample of 15 days gives the following data set :

110, 118, 130, 140, 142, 146, 112, 100, 95, 98, 96, 122, 123, 124, 130.

It is known that the number of items produced by the firm follows normal distribution with variance 300.

Can we conclude at 5% level of significance that the average daily production of items of that firm is

(a) 110 items

(b) More than 110 items

(c) Less than 110 items?

(d) Compute p-value for each case.

Solution

(a) Steps Involved in testing the hypothesis will be followed in this case :

(1) ~~fist~~ first we have to formulate null and alternative hypothesis. It is two tailed test. Since if the

average number of items produced by the firm is more or less than 110 to some ~~extend~~ extent, then the claim of the managing director will be proved as false. In that case the claim would be rejected.

So, the null hypothesis $H_0: \mu = 110$

Alternative hypothesis $H_a: \mu \neq 110$

(ii) Level of significance is $\alpha = 0.05$

(iii) Here the sample is taken a normal population with known variance. The appropriate test statistics is

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

where, $\bar{x}$ = The sample mean

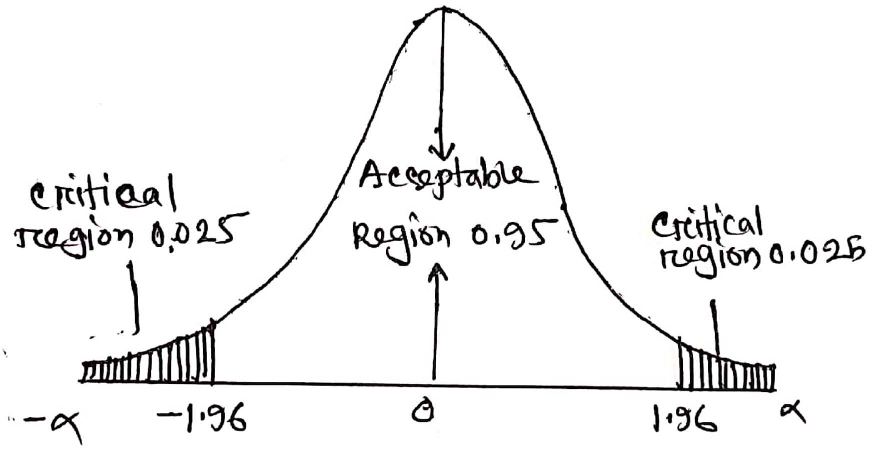
μ = Population mean

σ = Standard deviation of the population

$\sigma/\sqrt{n}$ = Standard error of the sample mean $\bar{x}$.

(iv) Here $\alpha = 0.05$ and it is a two tailed test, the critical region will be on both sides of curve & in such a way ~~of~~ that the critical ~~reas~~ region will ~~comprom~~ comprise 2.5% or 0.025 area at the right end and 2.5% at the left end. From the table of area of standard normal distribution, we see that ~~these~~ value of z are ± 1.96 that means the critical region are $z < -1.96$ (at the left end) and $z > 1.96$ (at the right end), i.e critical region is given by

$$|z_{0.025}| > 1.96$$



(v) Under the null hypothesis, the value of z is

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

(vi) Here $\bar{x} = \frac{\sum x_i}{n} = \frac{1286}{15} = 119.07$, $\mu_0 = 110$

$\sigma^2 = 300$ and $\sigma = 12.32$ and $n = 15$

$$\therefore z = \frac{119.07 - 110}{12.32/\sqrt{15}} = 2.03$$

(vii) It is found that the observed value of $z = 2.03$, which is greater than the right tail critical value 1.96, hence it falls in the upper critical region.

(vii) Since the observed value of the test statistics falls in the critical region, so, we fail to accept the null hypothesis at 5% level of significance.

(ix) Conclusion: Hence we can not accept the claim of the managing director at 5% level of significance.

P-value: from the table of the standard normal distribution we find that $P(z > 2.03) = 0.0212$, since it is two tailed test then the p-value $= 2 \times 0.0212 = 0.0424$.

That means the smallest level of significance at which the hypothesis may be rejected is approximately 4.34%

(b)

Here, Null hypothesis $H_0 : \mu = 110$

Alternative hypothesis $H_a : \mu > \cancel{110}\ 110$

$\alpha = 0.05$

P.T.O.

The appropriate z test statistics is the same as (a)

That is $$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Here $\alpha = 0.05$ and it is one sided right tailed test, the critical value $z_{0.05}$ will be found in such a way that $P[z > z_{0.05}] = 0.05$. It is found that the standard normal distribution that $z_{0.05} = 1.645$.

The calculated value of z under the null hypothesis is 2.03 which is greater than 1.645. That is observed value of z lies in the rejection region. Hence we have no reason to accept the null hypothesis.

P-value: From the table of standard normal distribution we find that $p = P(z > 2.03) = 0.0212$

That means the smallest level of significance at which the hypothesis may be rejected is approximately 2.12%.

(c)

Null hypothesis is the same as (a)

Null hypothesis $H_0: \mu = 110$

Alternative hypothesis $H_a: \mu < 110$

$\alpha = 0.05$

The appropriate test statistics is the same as (a).

That is

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Here $\alpha = 0.05$ and it is one sided left tailed test, the critical value $z_{0.05}$ will be found in such a way that $P[z < -z_{0.05}] = 0.05$. It is found from the standard normal distribution that for the left tail $z_{0.05} = -1.645$

The calculated value of z under the null hypothesis is 2.03 lies outside the region $z < -1.645$. This observed the value is in acceptable region.

P-value: from the table of the standard normal distribution we find that $p = P(z > 2.03) = 0.0212$

This is a left tailed test, so, the p-value is

$$P = P(z > -2.03) = 1 - 0.0212 = 0.9788$$

Source code by R language:

```
##a##
## H0:mu=110 vs H1: mu not equal to 110##
n=15
item<-c(110,118,130,140,142,146,112,100,95,98,96,122,
        123,124,130)
x_bar<-sum(item)/15
sigma<-sqrt(300)
z<-(x_bar-110)/(sigma/sqrt(n)) ## z cal=2.03##
z_tab_lower<-qnorm(0.025) ## -1.96
z_tab_upper<-qnorm(0.975) ## 1.96
## H0 is rejected since calculated z falls in the critical region
## P value##
P_value<-2*pnorm(z, lower.tail= FALSE) ## 0.0424

##b##
z=2.03
z_tab_R<-qnorm(0.05, lower.tail= FALSE) ## 1.64
p_value<-pnorm(z, lower.tail= FALSE) ## 0.021
## H0 is rejected since calculated z falls in the critical regio
```

```
##c##
z = 2.03
z_tab_L <- qnorm(0.05, lower.tail = TRUE) ## -1.64
P_value <- pnorm(z) ## 0.9786

## Construct the 95% Confidence Interval of mu##
CI <- c(x_bar + z_tab_lower * sigma/sqrt(n),
        x_bar + z_tab_upper * sigma/ sqrt(n))
## CI = (110.3014 to 127.8319)
```

Lab problem-02

A wholesaler knows that on average sales in it's store is 20% higher in December than in November. for the current year a random sample of six store was taken. Their percentage of sales increased in December was found to be 19.2, 18.4, 19.8, 20.2, 20.4, 19.0. Assuming that the sample has been drawn from a normal population with mean μ and variance σ^2.

(a) test the null hypothesis at 10% level of significance whether the true mean percentage sales increase is 20%, against the two sided alternative.

(b) do you think that the true mean percentages sales increases is ~~than~~ less than 20% at 10% level of significance

(c) do you think that true mean percentage sales increase is less than 20% at 10% level of significance

Solution

Here the population variance is unknown and the sample size is small. The estimate standard error of sample mean $\bar{x}$ is given by,

$se(\bar{x}) = s/\sqrt{n}$ where $s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$

(a)

We want to test the null hypothesis

$H_0: \mu = \mu_0 = 20$ against the alternative

$H_1: \mu = \mu_1 \neq 20$

The value of the test statistics under null hypothesis is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

which is distributed as Student's t with $(n-1)$ degrees of freedom. It is a also two tailed test, so the decision rule: Reject H_0 in favour of H_1 if

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha/2;(n-1)}$$

or $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{\alpha/2;(n-1)}$$

The critical value of t at 10% significance with $n-1=5$ df are $\pm t_{n-1;\alpha/2} = \pm t_{5;.05} = \pm 2.015$

Here, $n=6$, $\bar{x}=19.5$, $s^2 = 0.588$ and $s/\sqrt{n} = 0.29$

$$t = \frac{19.5-20}{0.29} = -1.08$$

Since the observed value $t=-1.08$ lies between -2.015 and 2.015, hence we fail to reject the null hypothesis at 10% level of significance.

$\boxed{b}$

In this case, we have to perform a one-tailed test,
$H_0: \mu = \mu_0 = 20$ against the alternative $H_1: \mu = \mu_1 > 20$
The decision rule is to reject H_0 in favour of H_1 if $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha; n-1}$

The compute value of t is the same as before $t = -1.08$, the critical value of t at 10% level of significance is $t_{0.10;5} = 1.476$

Since the observed value of $t = -1.08$ which is less than the critical value, we fail to reject the null hypothesis at 10% level of significance. which means the average sales increased by more than 20% is not ~~et~~ evident from the given data.

$\boxed{c}$

In this case, we have to perform a one tailed test, given $H_0: \mu = \mu_0 = 20$ against the alternative

$H_1: \mu = \mu_1 < 20$

The decision rule is to reject H_0 in favour of H_1 if $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{\alpha: n-1}$

The compute value of t is the before i.e. $t = -1.08$, the critical value of t at 10% level of significance is $t_{0.10;5} = -1.476$

Since the observed value $t = -1.08$ lies beyond the critical region, then we fail to reject the null hypothesis at 10% level of significance, which means the average sales increased by less than 20% is not evident from the given data.

Lab problem — 03

(Small sample size with unknown and equal population variances) The residence of Dhaka city complains that traffic speeding fines given in their city are higher than the traffic speeding fines that are given in Chittagong city. The appropriate authority agreed to study the problem. To check if the complains were resonable, independent random sample of the amounts paid by the residents for speeding fines in each of two cities over the last three months were obtained by the following data:

Dhaka city	100	125	135	128	140	142	128	137	156	142
Chattagong city	95	87	100	75	110	105	85	95		

Assuming an equal population variance, test

(I) whether there is any significance difference in the mean cost of speeding in these two cities and find the 95% confident interval.

(II) whether the mean speeding cost in Dhaka city ~~and~~ is higher than Chittagong city at 1% level of significance.

Solution

$\langle i \rangle$

Let X_1 be the speeding cost in Dhaka city and X_2 be the speeding cost in Chittagong city.
Here two normal populations with mean μ_1 and μ_2 respectively with common variance σ^2.

Then $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

It is a two tailed test, sample size small and variance unknown.

So,
$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Given that $n_1 = 10$, $n_2 = 8$, So the degrees of freedom is $n_1 + n_2 - 2 = 16$

From the table of t distribution at 5% level of significance with 16 df are ± 2.12.

From the given observation, we have

$\bar{x}_1 = 133.30$, $\bar{x}_2 = 94.00$, $s_1 = 18.20$, $s_2 = 20.68$

$s^2 = 371.98$,

So,
$$t = \frac{\overline{x_1} - \overline{x_2}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{133.30 - 94.00}{9.14} = 4.30$$

The observed ~~valte~~ value of t is greater than the critical value, so, the null hypothesis may be rejected at 5% level of significance.

Again the 95% confident interval for $(\mu_1 - \mu_2)$ is given

$$(\overline{x_1} - \overline{x_2}) - t_{16;0.025} \times s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < (\mu_1 - \mu_2) < (\overline{x_1} - \overline{x_2}) + t_{16;0.25} \times s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

or $39.30 \pm 2.12 \times 4.30 = 39.30 \pm 9.12$

$= (30.18, 48.42)$

P value: From the table of critical value of t distribution, we find that $t_{16;0.005} = 2.921$, that means $2 \times 0.005 = 0.010 = 1\%$ can be considered as the smallest level of significance at which the null is still rejected, so the required p value is 0.01

<11>

In this case, we have to perform one tailed test given by,

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 > \mu_2$

Here the critical value of t at 1% level of significance with 16 df is $t_{16:0.01} = 2.583$

The observed value of t is same as previous one,

so, $t = 4.30$

Here the observed t falls in the critical region, so, the null hypothesis may be rejected at 1% level of significance.

P-value: In this case the p-value is 0.005

because $p(t > 4.30) = 0.005$.

Lab problem-04

(Matched observations or paired sample) A study was conducted by a pharmaceutical company to compare the difference in effectiveness of two particular drugs in cholestrol levels. The company used paired sample approache to control variation in reduction that might be due to factors other than the drugs itself. Each member of pair was matched by age, weight, lifestyle and other pertinent factors. Drug X was given to one person randomly selected from each pair and drug Y was given to other individual in the pair. After a specific period of time each person's cholestrol level was measured again. Suppose a random sample of eight pair of patients with known cholestrol ~~prob~~ problems is selected from the large population of participants. The following table gives the numbers of point by which each person cholestrol level was reduced.

P.T.O.

Table: 1. Reduction levels of cholestrol by drugs:

Pair	1	2	3	4	5	6	7	8
Drug X	29	32	31	32	32	29	31	30
Drug Y	26	27	28	27	30	26	33	36

(i) The test whether there is any significant different between the mean reduction of cholestoral level by two drugs at 1% level of significance.

(ii) Find 99% confidente interval for the difference between the population means.

Solution

<1>

Let us formulate the null and alternative hypothesis as

$$H_0: \mu_x = \mu_y \text{ against } H_1: \mu_x \neq \mu_y$$

The appropriate test statistics is $t = \dfrac{\bar{d}}{se(\bar{d})} \sim t_{n-1}$

which is distribibuted as t with n-1 df,

where, $\bar{d} = \dfrac{\Sigma d}{n}$, $S_d^2 = \dfrac{\Sigma(d-\bar{d})^2}{n-1}$ and $se(\bar{d}) = \dfrac{S_d}{\sqrt{n}}$

Here $n=8$, degrees of freedom is $n-1=7$

It is two tailed test, so at 1% level of significance the critical region is

$$|t| > t_{n-1;\alpha/2} = t_{7;0.005} = 3.499$$

Table-02 Calcalation the mean and standard deviation of d

Pair	1	2	3	4	5	6	7	8
Drug X	29	32	31	32	32	29	31	30
Drug Y	26	27	28	27	30	26	33	36
$d = X - Y$	3	5	3	5	2	3	-2	-6
$(d - \bar{d})$	1.38	3.38	1.38	3.38	0.38	1.38	-3.63	-7.63
$(d-\bar{d})^2$	1.89	11.39	1.89	11.39	0.14	1.89	13.14	58.14

We have $\bar{d} = 1.625$, $sd^2 = 14.27$, $sd = 3.777$

So, $t = \frac{\bar{d}}{se(\bar{d})} = \frac{\bar{d}}{sd/\sqrt{n}} = \frac{1.625}{3.777/\sqrt{8}} = 1.22$

The observed value of t is smaller than absolute value of critical value, so, we fail to reject null hypothesis.

(ii)

The 99% confidence interval for the difference of population average reduction of cholestoral level is given by

$\bar{d} - t_{n-1;\alpha/2} \cdot se(\bar{d}) < \mu_x - \mu_y < \bar{d} + t_{n-1;\alpha/2} \cdot se(\bar{d})$

$\doteq (-3.05, 6.30)$

CamScanner

Lab problem - 05

For the following questions carry out the test of significance of population proportions at 5% level of significance (where x represents the number of things of particular catagory)

(i) $H_0: \pi = 0.25$, $H_1: \pi \neq 0.25$; $n = 100$, $x = 40$

(ii) $H_0: \pi = 0.40$, $H_1: \pi > 0.40$; $n = 200$, $x = 150$

(iii) $H_0: \pi = 0.30$, $H_1: \pi < 0.30$; $n = 400$, $x = 100$

Solution

The statistics to be used for testing the given hypothesis:

$$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$ where p is the estimate of the proportion.

(i)

This is two tailed test, So, the decision rule is:
Reject H_0 in favour of alternative at 5% level if $|z| > z_{0.005}$ or $|z| > 1.96$

Here $n=100$, so, $p=\frac{40}{100}=0.40$ and $\pi_0=0.025$, so the compute value of test statistics is

$$z=\frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}=\frac{0.40-0.25}{\sqrt{\frac{0.25(1-0.25)}{100}}}=3.46$$

Here $|z|>1.96$, the compute value of z falls the critical region, so, we fail to accept null hypothesis.

(ii)

This is right test, so the decision rule is:
Reject H_0 in favour of alternative at 5% level if $z>z_{0.05}=1.645$.

Here $n=200$, so, $p=\frac{100}{200}=0.50$ and $\pi_0=0.40$

$$\therefore z=\frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}=\frac{0.50-0.40}{\sqrt{\frac{0.40(1-0.40)}{200}}}=2.89$$

The observed value of $z>1.645$, which falls in the critical region, so, we fail to accept null hypothesis.

Includes:

(iii)

This is a left tailed test, so the decision rule is:

Reject H_0 in favour of alternative at 5% level if

$z < -z_{0.05} = -1.645$

Here, $n = 400$, so, $p = \frac{100}{400} = 0.25$ and $\pi_0 = 0.30$

$$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.25 - 0.30}{\sqrt{\frac{0.30(1-0.30)}{400}}} = -2.18$$

The observed value of z is $z < -1.645$, which falls the critical region, so, we fail to accept null hypothesis.

Lab problem — 6

A company is considering two different television advertisements for promotion of a new products. Manager believes that advertisement A is more effective than B. Two test market areas with virtually identical consumer characteristics are selected: Advertisement A is used in one area and advertisement B is used in other area. In a random sample of 60 customer who saw the advertisement A, 18 had tried to buy the product, on the other hand, in a random sample of 100 customers who saw the advertisement B, 22 had tried to buy the product. Does this indicate that the advertisement A is more efficient than advertisement B, if level of significants is 5% ?

Solution

Let, π_1 and π_2 be the population proportions of customers who had tried to buy the products after seeing the advertisement A and B.

Then we consider the null hypothesis as both advertisements are equally effective, that means,

$H_0: \pi_1 = \pi_2$ against the alternative $H_1: \pi_1 > \pi_2$

~~This is one tailed test~~. Under the null hypothesis, the appropriate test statistics is

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

Where p is the pooled estimate of proportion.

Given $\alpha = 0.05$ and it is a right tailed test. So, the decision rule is: Reject H_0 in favour of H_1 if $z > z_{0.05}$ or $z > 1.645$.

$n_1 = 60$, proportion, $p_1 = \frac{18}{60} = 0.30$ and $n_2 = 100$,

$p_2 = \frac{22}{100} = 0.22$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{60 \times 0.30 + 100 \times 0.22}{60 + 100} = 0.25$$

$$\text{Thus, } z = \frac{0.30 - 0.22}{\sqrt{0.25(1-0.25)\left(\frac{1}{60} + \frac{1}{100}\right)}} = 1.131$$

So, the observed value z is less than the critical value 1.645, we fail to reject null hypothesis at 5% level of significance.