Session 15:

Additional Exercise

Problem Statement 1:

In each of the following situations, state whether it is a correctly stated hypothesis testing problem and why?

1. H0:μ =25, H1 :$\mu \neq 25$

2. H0:σ >10, H1 : σ = 10

3. H0: $\bar{x}$ = 50, H1: $\bar{x} \neq 50$

4. H0: p = 0.1, H1: p = 0.5

5. H0: s = 30, H1: s > 30

Solution:

1. H0 : μ =25, H1 : $\mu \neq 25$

The hypothesis stated above is correct.

Justification:

Hypothesis testing for μ:

H0: $\mu = \mu 0$

Ha: $\mu > \mu 0$, $\mu < \mu 0$, μ 6= $\mu 0$ (we can use any one of these!)

2. H0: σ > 10, H1: σ =10

The given hypothesis statement is valid since if standard deviation is > than 10 it can be either equal to or can range between < than to > 10.

H0 - Null Hypothesis: Standard deviation of the given data is greater than 10

H1 - Alternate Hypothesis: Standard Deviation of the given data is equal to 10

3. H0: $\bar{x}$ = 50, H1 : $\bar{x} \neq 50$

The hypothesis stated above is correct.

Justification:

H0 - Null Hypothesis: Mean of the given data is 50 **$\bar{x}$ = 50**

H1 - Alternate Hypothesis: Mean of the given data is less than or greater than 50 $\rightarrow$ **$\bar{x} \neq 50$**

4. H0: p = 0.1, H1: p = 0.5

The hypothesis stated above is not correct.

Justification:

Hypothesis testing should be for p:

H0: p = p0

Ha: p > p0, p < p0, p ≠ p0 (we can use any one of these 3!)

5. H0: s = 30, H1: s > 30

The hypothesis stated above is correct.

Justification:

Hypothesis testing should be for s:

H0: s = s0

Ha: s > s0, s < s0, s ≠ s0 (we can use any one of these 3!) but, here the value is completely different hence it is not correct.

Problem Statement 2:

The college bookstore tells prospective students that the average cost of its textbooks is Rs. 52 with a standard deviation of Rs. 4.50. A group of smart statistics students thinks that the average cost is higher. To test the bookstore's claim against their alternative, the students will select a random sample of size 100. Assume that the mean from their random sample is Rs. 52.80. Perform a hypothesis test at the 5% level of significance and state your decision.

Solution:

Hypothesis for the above condition:

Ho: $\mu \leq 52$

Ha : $\mu > 52$

<u>Significance Level:</u>

The significance level $\alpha = 0.05$ or 5%

<u>Determination of the suitable test static:</u>

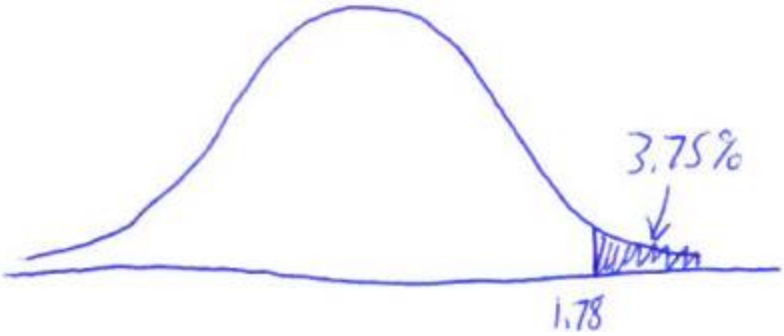
$\mu = 52$

X = Rs. 52.80

s = Rs.4.50

n = 100

We compute the z statistic :



$$z^* = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$$= \frac{52.80 - 52}{4.50 / \sqrt{100}}$$

$$= 1.78$$

Reject Ho if z>1.645

Z = [(52.80-52] /(4.50√100 = 1.78

Reject Ho: conclude that the average cost of textbooks is statistically significantly higher than the bookstore's claim.

p = P(z>1.78) = .5 - .4625 = 0.0375 (Note that this is close to the significance level, so it is not very, very strong evidence against the null hypothesis.)

Problem Statement 3:

A certain chemical pollutant in the Genesee River has been constant for several years with mean μ = 34 ppm (parts per million) and standard deviation σ = 8 ppm. A group of factory representatives whose companies discharge liquids into the river is now claiming that they have lowered the average with improved filtration devices. A group of environmentalists will test to see if this is true at the 1% level of significance. Assume \ that their sample of size 50 gives a mean of 32.5 ppm. Perform a hypothesis test at the 1% level of significance and state your decision.

Hypothesis for the above condition:

Ho: $\mu \leq 34$

Ha: $\mu > 34$

Significance Level:

The significance level $\alpha = 0.01$ or 1%

Determination of the suitable test static:

$\mu = 34$

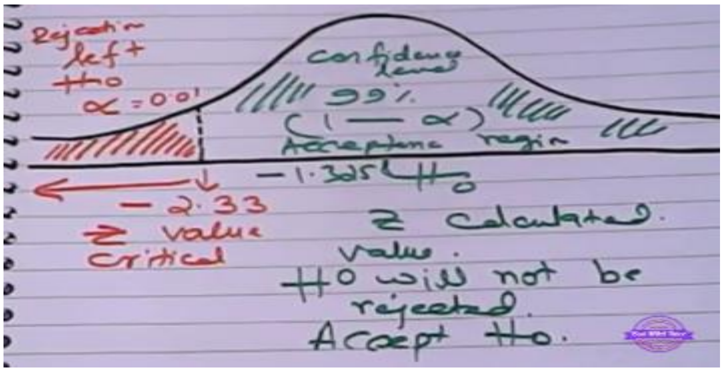
X = Rs. 52.80

s = Rs.4.50

n = 100

We compute the z statistic

$$z^* = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$



$$= \frac{32.5 - 34}{8/\sqrt{50}}$$

$$= -1.5 / 1.13$$

$$= -1.3258$$

Inference: since the calculated z value is in the acceptable region we accept the null hypothesis.

Problem Statement 4:

Based on population figures and other general information on the U.S. population, suppose it has been estimated that, on average, a family of four in the U.S. spends about $1135 annually on dental expenditures. Suppose further that a regional dental association wants to test to determine if this figure is accurate for their area of country. To test this, 22 families of 4 are randomly selected from the population in that area of the country and a log is kept of the family's dental expenditure for one year. The resulting data are given below. Assuming, that dental expenditure is normally distributed in the population, use the data and an alpha of 0.5 to test the dental association's hypothesis.

1008, 812, 1117, 1323, 1308, 1415, 831, 1021, 1287, 851, 930, 730, 699, 872, 913, 944, 954, 987, 1695, 995, 1003, 994

Solution:

Hypothesis for the above condition:

Ho: $\mu = 1135$

Ha: $\mu \neq 1135$

Significance Level:

The significance level $\alpha = 0.05$ or 5%

Determination of the suitable test static:

Population Mean μ = Rs. 1135

Sample Mean $\bar{x}$ = Rs. 1031.32 (Mean of the given data)

s = Rs.240.37 (Standard Deviation of the data)

n = 22

We compute the test statistic :

$$z^* = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$$= \frac{1031.32 - 1135}{240.37 / \sqrt{22}}$$

= -103. 68 / 51.25

= - 2.02

Inference:

Since the calculated z value is > 1.96 it falls under the critical region. Hence, the average dental expenses for the population is higher than 1135. Very strongly evident that the null hypothesis is not acceptable.

Problem Statement 5:

In a report prepared by the Economic Research Department of a major bank the Department manager maintains that the average annual family income on Metropolis is $48,432. What do you conclude about the validity of the report if a random sample of 400 families shows and average income of $48,574 with a standard deviation of 2000?

Hypothesis for the above condition:

Ho: μ = $ 48432

Ha: μ > $ 48432

Large Sample hence, we follow central limit theorem

Significance Level:

The significance level $\alpha = 0.05$ or 5% - reject H0 if the $z < -1.96$ or $z > +1.96$

Determination of the suitable test static:

Population Mean μ = $ 48432

Sample Mean $\bar{x}$ = $ 48574

s = $ 2000

n = 400

$$z^* = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$$= \frac{48574 - 48432}{2000 / \sqrt{400}}$$

= -142 / 100

= - 1.42

Inference: We can accept the null hypothesis at confidence interval of 0.05 since it is inacceptable region.

P-Value p (z>1.42) +p(z<-1.42) = .156

Problem Statement 6:

Suppose that in past years the average price per square foot for warehouses in the United States has been \$32.28. A national real estate investor wants to determine whether that figure has changed now. The investor hires a researcher who randomly samples 19 warehouses that are for sale across the United States and finds that the mean price per square foot is \$31.67, with a standard deviation of \$1.29. Assume that the prices of warehouse footage are normally distributed in population. If the researcher uses a 5% level of significance, what statistical conclusion can be reached? What are the hypotheses?

Solution:

Hypothesis for the above condition:

Ho: $\mu = \$ 32.28$

Ha: $\mu > \$ 31.67$

Significance Level:

The significance level $\alpha = 0.05$ or 5% - reject H0 if the $z < -1.96$ or $z > +1.96$

Determination of the suitable test static:

Population Mean $\mu = \$ 32.28$

Sample Mean $\bar{x} = \$ 31.67$

$s = \$ 1.29$

$n = 19$

$$z^* = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$$= \frac{31.67 - 32.28}{1.29 / \sqrt{19}}$$

$$= -0.61 / 0.29$$

$$= -2.1$$

The Critical value of z is -1.96 and +1.96

The Critical value of $z = \pm 1.96$ for a two-tailed test at 5% level of significance. Since, the compound value of $z = -2.1$ falls in rejection region, we reject the null hypothesis. Hence, the average price per square foot for warehouses has changed now.

Problem Statement 7:

Fill in the blank spaces in the table and draw your conclusions from it.

Acceptance region	Sample size	α	B at μ = 52	β at μ = 50. 5
$48.5 < \bar{x} < 51.5$ 10	10		26.43 (-4.43, -0.63)	89.05 (-2.53 , 1,26)
$48 < \bar{x} < 52$ 10	10		50.00% (-5.06, 0.00)	97.05 (-3.16, 1.90)
$48.81 < \bar{x} < 51.9$ 16	16		43.64 (-5.10, -0.16)	98.40 (-2.70, 2.24)
$48.42 < \bar{x} < 51.58$ 16	16		25.14 (-5.73, -0.67)	95.78 (-3.13, 1.73)

Problem Statement 8:

Find the t-score for a sample size of 16 taken from a population with mean 10 when the sample mean is 12 and the sample standard deviation is 1.5.

Solution:

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

Population Mean- $\mu = 10$

Sample Mean - $\bar{x} = 12$

Standard Deviation- $s = 1.5$

$n = 16$

$$t = \frac{12-10}{1.5 / \sqrt{16-1}}$$

$$= \frac{2*3.87}{1.5}$$

$$= 5.33$$

Problem Statement 9:

Find the t-score below which we can expect 99% of sample means will fall if samples of size 16 are taken from a normally distributed population.

Solution:

$1 - \alpha = 0.99$

$\alpha = 0.01$

df = n-1

$\therefore$ df = 16-1 = 15

$t_{0.99} = -t_{0.01} = -2.602$

Problem Statement 10:

If a random sample of size 25 drawn from a normal population gives a mean of 60 and a standard deviation of 4, find the range of t-scores where we can expect to find the middle 95% of all sample means. Compute the probability that $P(-t_{0.05} < t < t_{0.10})$.

Solution:

n = 25

Sample mean = 60

sd = 4

$\alpha = 0.05$

Since the population is normally distributed, the sample is small, and the population standard deviation is unknown, the formula that applies is

$$\bar{x} \pm t_{\alpha/2} \, (s / \sqrt{n})$$

Confidence level 95% means that $\alpha = 1 - 0.95 = 0.05$,

so $\alpha/2 = 0.025$.

Since the sample size is n = 25, there are n−1 = 24 degrees of freedom.

"Critical Values of" $t_{0.025} = 2.145$. Thus

$$\bar{x} \pm t_{\alpha/2} \, (s / \sqrt{n}) = 60 \pm 2.145 \; (4/\sqrt{25}) = 60 \pm 1.716$$

One may be 95% confident that the true value of μ is contained in the interval (61.716, 58.284)

$\mu = 60$

$\sigma = 4$

$P(-t_{0.05} < t < t_{0.10}) = ?$

$\Rightarrow$ using 23 as degrees of freedom

$P(-1.713) < t < 1.319)$

$\Rightarrow P(-1.716) < \frac{t-60}{4\sqrt{25}} < 1.319$

$\Rightarrow -1.716 < \frac{t-60}{4\sqrt{25}} < 1.319$

$\Rightarrow -1.716 \times 5 < \frac{t-60}{4} < 1.319 \times 5$

$\Rightarrow -8.58 < \frac{t-60}{4} < 6.595$