Assignment -12

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: 𝜇 ≠ 25

2. 𝐻0: 𝜎 > 10, 𝐻1: 𝜎 = 10

3. 𝐻0: 𝑥 = 50, 𝐻1: 𝑥 ≠ 50

4. 𝐻0: 𝑝 = 0.1, 𝐻1: 𝑝 = 0.5

5. 𝐻0: 𝑠 = 30, 𝐻1: 𝑠 > 30

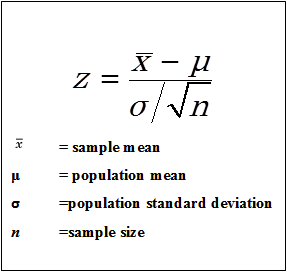
Solution :

1 statement is correct because A **Hypothesis Test** evaluates two mutually exclusive statements about a population to determine which statement is best supported by the sample **data**.

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 :



H0 = µ = 52

Ha = µ ≠ 52

µ = 52

s =4.50

n = 100

x = 52.8

α = 5%

S.E = 4.50 / √100

= 4.50 / 10

= 0.45

Z = 52.8 -52 / 0.45

= 1.78

Z (0.05) = +-1.96

We accept the null hypothesis Ho hence , the average cost of textbook is Rs. 52

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.

Solution :

H0 = µ = 34

Ha = µ ≠ 34

µ = 34

s =8

n = 50

x = 32.5

α = 1%

S.E = 8 / √50

= 8 / 7.07

= 1.13

Z = 32.5 -34 / 1.13

= -1.33

Z(0.01) = +-2.33

We accept the null hypothesis Ho hence , they have lowered the average with improved filtration devices .

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 :

H0 = µ = 1135

Ha = µ ≠ 1135

µ = 1135

s = 240.37

n = 22

x = 1031.32

α = 0.05

S.E = 240.37 / √22

= 240.37 / 4.6

= 52.25

Z = 1031.32-1135 / 52.25

= -103.68 / 51.25

=-2.02

Z(0.05) = +-1.96

So , we accept alternative hypothesis Ha hence that the average a regional dental association expenses for the population is not accurate .

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?

H0 = µ = 48432

Ha = µ ≠ 48432

µ = 48432

s = 2000

n = 400

x = 48574

α = 0.05

S.E = 2000 / √400

= 2000 / 20

= 100

Z = 48574 - 48432 / 100

= 142 / 100

= 1.42

Z(0.05) = +-1.96

So , we accept null hypothesis.

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 :

H0 = µ = 32.28

Ha = µ ≠ 32.28

µ = 32.28

s = 1.29

n = 19

x = 31.67

α = 0.05

S.E = 1.29 / √19

= 1.29 / 4.35

= 0.29

Z = 31.67-32.28 / 0.29

= - 0.61/ 0.29

=-2.10

Z(0.05) = +-1.96

So , we accept alternative hypothesis Ha

Problem Statement 7:

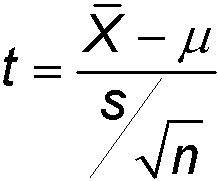
Fill in the blank spaces in the table and draw your conclusions from it.

Acceptance region = 48.5< x <51.5 , sample size =10 find the value of α , β at 52 , β at 50.5

Solution : -

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 :-

H0 = µ = 10

Ha = µ ≠ 10

µ = 10

s = 1.5

n = 16

x = 12

α = 0.05

S.E = 1.5 / √16

= 1.5/ 4

= 0.375

t = 12-10 / 0.375

= 2/ 0.375

=5.3

t(0.05) = +-1.753

so , hence we reject null hypothesis H0

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 :-

T =(0.01)

n-1 =15

1 –α =0.99 , α = 0.01

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 (−𝑡0.05 <𝑡<𝑡0.10).

Solution :

Problem Statement 11:

Two-tailed test for difference between two population means Is there evidence to conclude that the number of people travelling from Bangalore to Chennai is different from the number of people travelling from Bangalore to Hosur in a week, given the following: Population 1: Bangalore to Chennai

n1 = 1200

x1 = 452

s1 = 212

Population 2: Bangalore to Hosur

n2 = 800

x2 = 523

s2 = 185

solution :

Ho= µ(x1) = µ(x2)

Ha = µ(x1) ≠ µ(x2)

S.D = √212 +185 /1200+800-2

= √397/1998

=0.45

S.E = 0.45/44.69

=0.010

T = 452-523/0.010

=-71/0.010

So , we accept alternative hypothesis Ha .

Problem Statement 12:

Is there evidence to conclude that the number of people preferring Duracell battery is different from the number of people preferring Energizer battery, given the following:

Population 1: Duracell

n1 = 100

x1 = 308

s1 = 84

Population 2: Energizer

n2 = 100

x2 = 254

s2 = 67

solution :

Ho= µ(x1) = µ(x2)

Ha = µ(x1) ≠ µ(x2)

S.D = √84 +67 /100+100-2

= √151/198

=0.87

S.E = 0.87/14.07

=0.061

T = 308-254/0.061

=885.2

So , we accept alternative hypothesis Ha .

Problem Statement 13:

Pooled estimate of the population variance Does the data provide sufficient evidence to conclude that average percentage increase in the price of sugar differs when it is sold at two different prices? Population 1: Price of sugar = Rs. 27.50

n1 = 14

x1 = 0.317%

s1 = 0.12%

Population 2: Price of sugar = Rs. 20.00

n2 = 9

x2 = 0.21%

s2 = 0.11%

Problem Statement 14:

The manufacturers of compact disk players want to test whether a small price reduction is enough to increase sales of their product. Is there evidence that the small price reduction is enough to increase sales of compact disk players?

Population 1: Before reduction

n1 = 15

x1 = Rs. 6598

s1 = Rs. 844

Population 2: After reduction

n2 = 12

x2 = RS. 6870

s2 = Rs. 669

solution :

Ho= µ(x1) = µ(x2)

Ha = µ(x1) ≠ µ(x2)

S.D = √844 +669 /15+12-2

= √1513/25

=7.8

S.E = 7.8/5

=1.56

T = 6598-6870/1.56

=-174.3

So , we accept alternative hypothesis Ha .

Problem Statement 15:

Comparisons of two population proportions when the hypothesized difference is zero Carry out a two-tailed test of the equality of banks’ share of the car loan market in 1980 and 1995. Population 1: 1980

n1 = 1000

x1 = 53

𝑝 1 = 0.53

Population 2: 1985

n2 = 100

x2 = 43

𝑝 2= 0.53

Solution :

H0 = p1 = p2

Ha = p1 ≠ p2

p = (p1 \* n1 + p2 \* n2) / (n1 + n2)

= (0.53 \* 1000 + 0.53 \* 100) / (1000+100)

= 583 / 1100

=0.53

SE = sqrt{ p \* ( 1 - p ) \* [ (1/n1) + (1/n2) ] }

=√{0.53 \* (1- 0.53) \* [1/1000 +1/100]}

= √0.0027

= 0.05

z = (x1 – x2) / SE

= 53 – 43 / 0.05

= 200

Z (0.05 ) = +- 1.96

So , we accept alternative hypothesis Ha

Problem Statement 16:

Carry out a one-tailed test to determine whether the population proportion of traveler’s check buyers who buy at least $2500 in checks when sweepstakes prizes are offered as at least 10% higher than the proportion of such buyers when no sweepstakes are on. Population 1: With sweepstakes

n1 = 300

x1 = 120

𝑝 1= 0.40

Population 2: No sweepstakes

n2 = 700

x2 = 140

𝑝 2= 0.20

H0 = p1 = p2

Ha = p1 ≠ p2

p = (p1 \* n1 + p2 \* n2) / (n1 + n2)

= (0.40\* 300 + 0.20 \* 700) / (300+700)

= 260 / 1000

=0.26

SE = sqrt{ p \* ( 1 - p ) \* [ (1/n1) + (1/n2) ] }

=√{0.26 \* (1- 0.26) \* [1/300 +1/700]}

= 0.030

z = 120 – 140 / 0.030

= -66.6

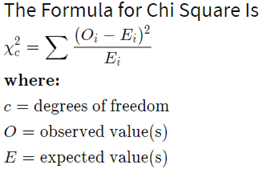
Z (0.05 ) = +- 1.96

So , we accept alternative hypothesis Ha

Problem Statement 17:

A die is thrown 132 times with the following results: Number turned up: 1, 2, 3, 4, 5, 6 Frequency: 16, 20, 25, 14, 29, 28 Is the die unbiased? Consider the degrees of freedom as 𝑝 ^ -1

Solution :-



H0 = The die is unbiased.

Ha = The die is not unbiased.

The frequency of each number to be 132/6 =22

|  |  |  |
| --- | --- | --- |
| Observation o/p | Expected o/p | (O-E)2 |
| 16 | 22 | 36 |
| 20 | 22 | 4 |
| 25 | 22 | 9 |
| 14 | 22 | 64 |
| 29 | 22 | 49 |
| 28 | 22 | 36 |

X2 = (O-E)2 / E

= 198 / 22

= 9

X(0.05) = 11.07

So , we accept the null hypothesis Ho

Problem Statement 18:

In a certain town, there are about one million eligible voters. A simple random sample of 10,000 eligible voters was chosen to study the relationship between gender and participation in the last election. The results are summarized in the following 2X2 (read two by two) contingency table:

|  |  |  |
| --- | --- | --- |
| Gender | Men | Women |
| Voted | 2792 | 3541 |
| Not voted | 1486 | 2131 |

We would want to check whether being a man or a woman (columns) is independent of having voted in the last election (rows). In other words, is “gender and voting independent”?

Solution :

H0 = Gender is independent of Voting

Ha =Gender and Voting are dependent

Observed table

Men Women Total

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Voted 2792 3591 6383

Didn't vote 1486 2131 3617

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4278 5722 10000

Expected table

Men Women Total

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Voted 2731 3652 6383

Didn't vote 1547 2070 3617

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4278 5722 10000

X2  = (2792-2731)2 / 2731 + (3591-3652)2 / 3652 + (1486-1547)2 / 1547 + (2131-2070)2 / 2070

= 1.4+1.0+2.4+1.8

=6.6

X(0.05) = 3.84

So, we reject the null hypothesis Ho

Problem Statement 19:

A sample of 100 voters are asked which of four candidates they would vote for in an election. The number supporting each candidate is given below:

|  |  |  |  |
| --- | --- | --- | --- |
| Higgins | Reardon | White | Chartoon |
| 41 | 19 | 24 | 16 |

Do the data suggest that all candidates are equally popular? [Chi-Square = 14.96, with 3 df, 𝑝 0.05 .

Solution :

H0 =  No preference

Ha = preference

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| O | 41 | 19 | 24 | 16 |
| E | 25 | 25 | 25 | 25 |
| (O-E)2 | 256 | 36 | 1 | 81 |
| (O-E)2 / E | 10.24 | 1.44 | 0.04 | 3.24 |

X2  = 10.24+ 1.44 + 0.04 + 3.24

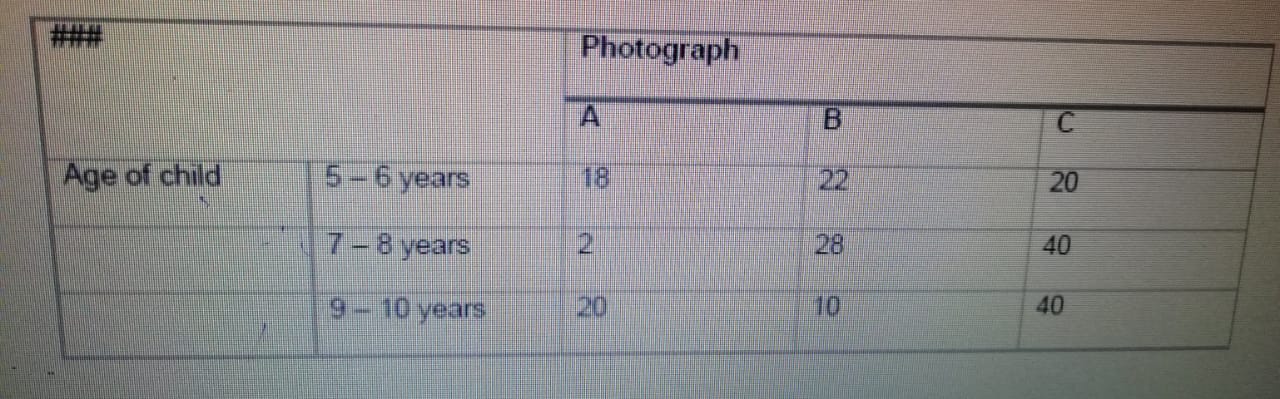
= 14.96

X(0.05) =   13.28

So , we reject the null hypothesis Ho .

Problem Statement 20:

Children of three ages are asked to indicate their preference for three photographs of adults. Do the data suggest that there is a significant relationship between age and photograph preference?



What is wrong with this study? [Chi-Square = 29.6, with 4 df: 𝑝 < 0.05].

Solution :

H0 =  not relation

Ha = relation



|  |  |  |
| --- | --- | --- |
| Observation | Expected | (O-E)2 / E |
| 18 | 12 | 3 |
| 22 | 18 | 0.89 |
| 20 | 30 | 3.33 |
| 2 | 14 | 10.29 |
| 28 | 21 | 2.33 |
| 40 | 35 | 0.71 |
| 20 | 14 | 2.57 |
| 10 | 21 | 5.76 |
| 40 | 35 | 0.71 |
|  |  |  |

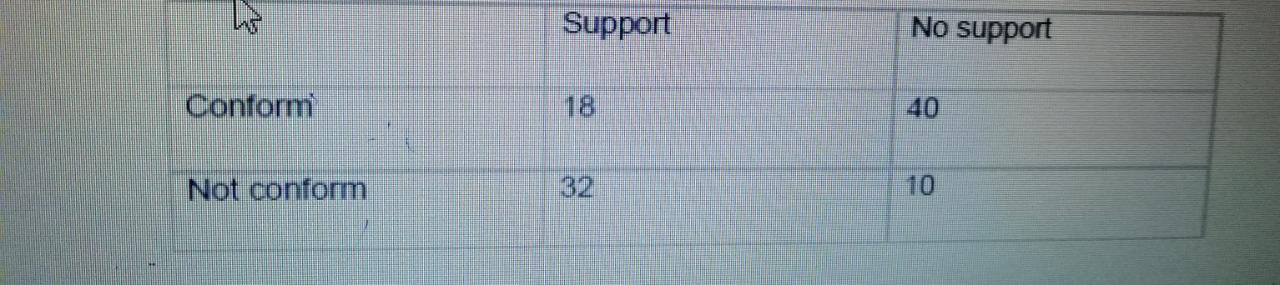
X 2  = 29.60

X (0.001) = 18.46

So, we accept alternative hypothesis Ha

Problem Statement 21:

A study of conformity using the Asch paradigm involved two conditions: one where one confederate supported the true judgement and another where no confederate gave the correct response.



Is there a significant difference between the "support" and "no support" conditions in the frequency with which individuals are likely to conform? [Chi-Square = 19.87, with 1 df: 𝑝 < 0.05].

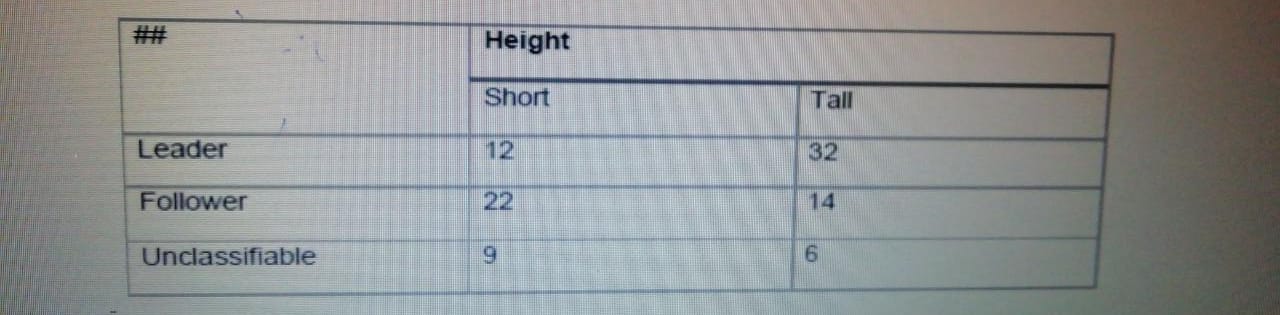
Solution :-

O: 18 40 32 10

E: 29 29 21 21

Problem Statement 22:

We want to test whether short people differ with respect to their leadership qualities (Genghis Khan, Adolf Hitler and Napoleon were all stature-deprived, and how many midget MP's are there?) The following table shows the frequencies with which 43 short people and 52 tall people were categorized as "leaders", "followers" or as "unclassifiable". Is there a relationship between height and leadership qualities? [Chi-Square = 10.71, with 2 df: 𝑝 < 0.01].



Solution :

H0 =  not relation

Ha = To be a relationship between height and leadership qualities

leader: 12 (19.92) , 32 (24.08)

follower: 22 (16.29) , 14 (19.71)

unclassifiable: 9 (6.79) 6 , (8.21) 15

X2  = 3.146 + 2.602 + 1.998 + 1.652 + 0.720 + 0.595

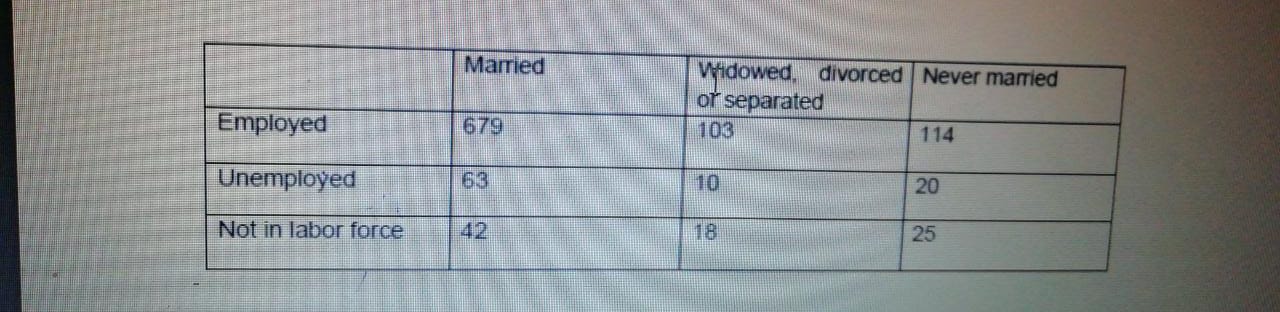
= 10.712

X(0.01) =9.21

So, we accept the null hypothesis Ha

Problem Statement 23:

Each respondent in the Current Population Survey of March 1993 was classified as employed, unemployed, or outside the labor force. The results for men in California age 35- 44 can be cross-tabulated by marital status, as follows:



Men of different marital status seem to have different distributions of labor force status. Or is this just chance variation? (you may assume the table results from a simple random sample.)

Solution :-

Ho: Rows and columns are independent.

Ha: There is a relationship between rows and columns.

Obs 1 = 679 103 114

Exp 1 = 654.06 109.29 132.65

Obs 2 = 63 10 20

Exp 2 = 67.89 11.34 13.77

Obs 3 = 42 18 25

Exp 3 = 62.05 10.37 12.58

X2  = 31.613103

X (0.05) = 9.487

So , we reject the null hypothesis