**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. Correct. The alternate of case Mean equal to 50 is always Mean not equal to 50
2. Correct: This is one tail hypothesis testing(Right tail)
3. Not correct. Hypothesis statement should be taken with statistical data, not with random variable
4. Not correct.Hypothesis statement should be taken with statistical data, not with probality of distribution

5.Correct. This is one tail hypothesis testing(Left tail)

**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.



n = 100

Sample mean = 52.80

Null hypothesis , H0)μ = 52

Alternate hypothesis , H1) μ > 52

Decision rule:

This is right tail test

If Z value is greater than 1.68 then accept H1

Z = (52.80 - 52)/(4.5/sqrt(100)) = 1.78

**Since 1.78 falls outside critical region. So we will accept H1**

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:

Null hypothesis , H0)μ = 34

Alternate hypothesis , H1) μ < 34

Alpha = 0.01

Decision rule:

If Z value is less than -2.3 reject null hypothesis



**Since -1.32 Falls outside critical region, so we will accept null hypothesis**

**Problem Statement 4:**

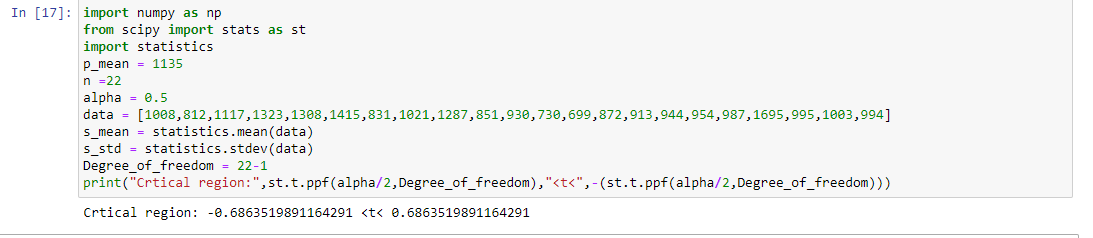
Based on population figures and other general information on the U.S. population, suppose it has been estimated that, on average, a family of four in the U.S. spends about $1135 annually on dental expenditures. Suppose further that a regional dental association wants to test to determine if this figure 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:

Null hypothesis , H0)μ = 1135

Alternate hypothesis , H1) μ ≠ 1135

Decision rule:



If t value is less than -0.69 and greater than 0.69 then reject null hypothesis



The t value is inside critical region and hence we will reject null hypothesis

**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?

Solution:

Null hypothesis , H0)μ = 48,432

Alternate hypothesis , H1) μ ≠ 48,432

Decision rule:

Let alpha = 0.05

Here we use Z stats. Because population size is very large

Decision rule:

If z value is less than -1.96 or greater than 1.96, them we will reject null hypothesis



**Since 1.42 is outside critical region, we will 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:

Null hypothesis , H0)μ = 32.28

Alternate hypothesis , H1) μ ≠ 32.28



**Since t value is inside the critical region, we will reject null hypothesis**

**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.

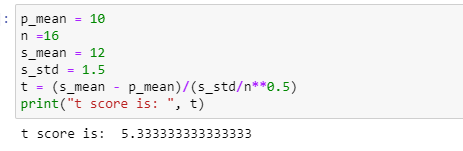
Solution:

n =16

P\_mean = 10

Standard deviation = 1.5

S\_mean = 12



**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:

Alpha = 1-0.99 = 0.01

Degree of freedom = 15

From t table , t score is -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:

Alpha = (1- 0.95)/2 = 0.025

So t score is 2.064



**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:

H0: Mean is same

H1: Mean is not same



Since z score fall inside critical region, we will reject null hypothesis. That means people travelling from Bangalore to Chennai is different from the number of people travelling from Bangalore to Hosur in a week

**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:

H0: Peoples preferring Duracell and Energizer battery is same

H1: Peoples preferring Duracell and Energizer battery is different.



Since z score fall in the critical region, we will reject null hypothesis

**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%

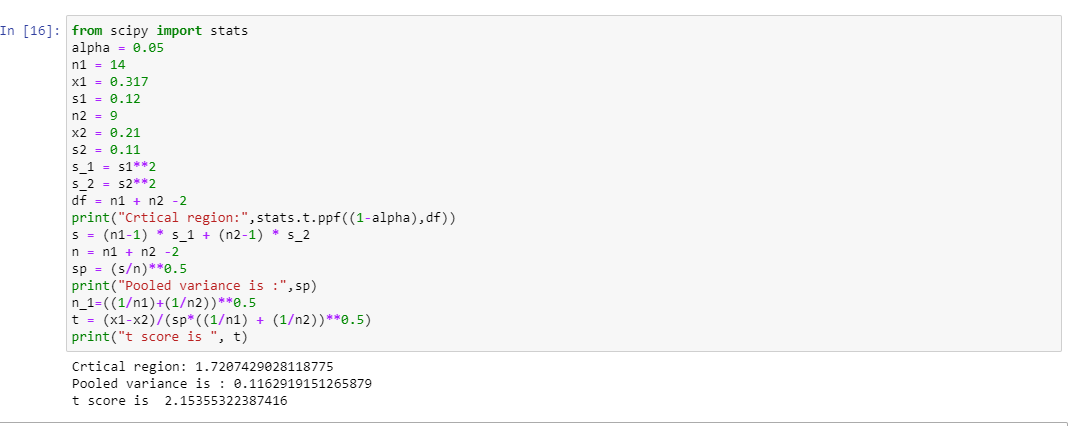
Population 2: Price of sugar = Rs. 20.00 n2 = 9

x2 = 0.21%

s2 = 0.11%

H0 : Average Sugar price has no change(Mean1 = Mean2)

H1 : Sugar price has decrease(Mean1 > Mean2)



Since t score is in critical region, we will reject null hypothesis

**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

H0: There is no change in sales(Mean1 = Mean2)

H1: Sales increase after reduction(Mean1 < Mean2)



T score falls outside critical region, we will accept null hypothesis

**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

p1 = 0.53

Population 2: 1985

n2 = 100

x2 = 43

p2= 0.53

H0: Proportions are same

H1: Proportions are not same



Since z score is outside critical region, we will accept null hypothesis

**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

p1= 0.40

Population 2: No sweepstakes n2 = 700

x2 = 140

p2= 0.20

H0: P1-P2 < 0.10

H1: P1-P2> 0.10



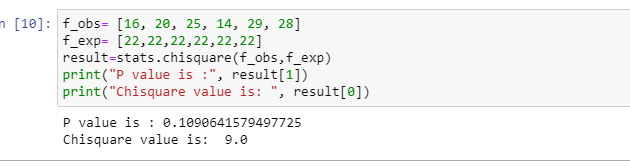
Since z score is in the critical region we will reject null hypothesis

**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 .



Since chi square value is in the critical region, we can say that die is unbiased

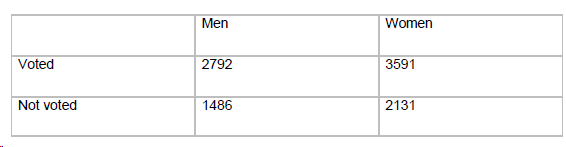
**Problem Statement 18:**

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:



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 dependant on voting

H1: Gender is not dependant on voting



We will reject null hypothesis

**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:

Do the data suggest that all candidates are equally popular? [Chi-Square = 14.96,

with 3 df, 0.05 .

Solution:

H0: All candidates are equally popular

H1: All candidates are not equally popular



So all candidates are note equally popular

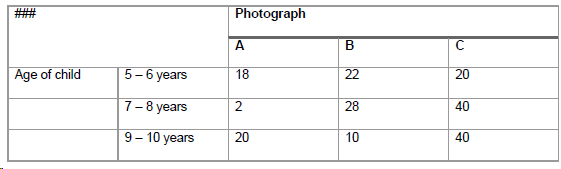
**Problem Statement 20:**

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: < 0.05].



Solution:

H0: There is a no significant relation

H1: There is a significant relation



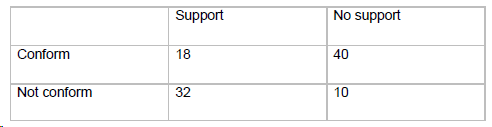
So we will reject null hypothesis

**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

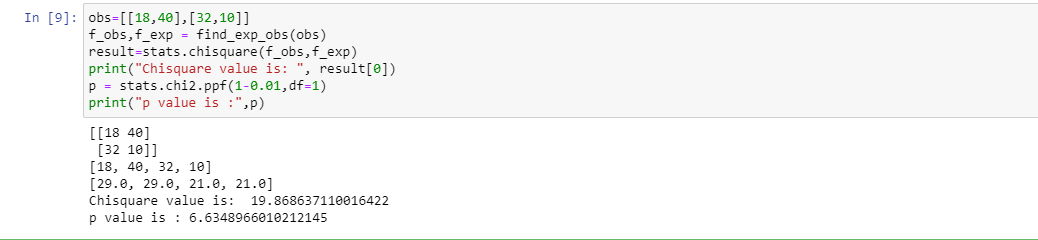
frequency with which individuals are likely to conform? [Chi-Square = 19.87, with 1 df:

< 0.05].

Solution:

H0: There is no significant difference

H1: There is a significant deference



We will reject null hypothesis

**Problem Statement 22:**

We want to test whether short people differ with respect to their leadership qualities

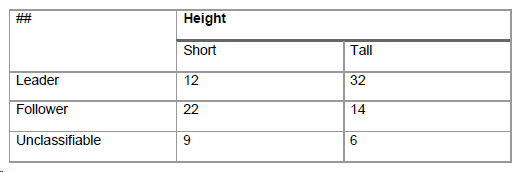
(Genghis Khan, Adolf Hitler and Napoleon were all stature-deprived, and how many midget

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

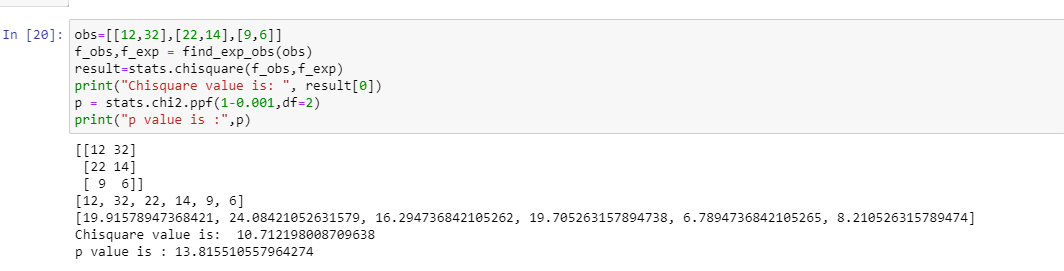
relationship between height and leadership qualities?

[Chi-Square = 10.71, with 2 df: < 0.01]



H0: There is no relationship

H1: There is a relationship



We will accept null hypothesis

**Problem Statement 23:**

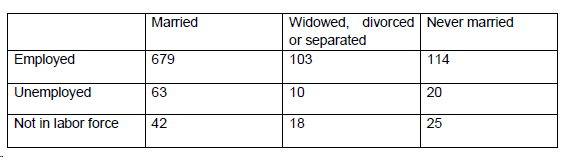
Each respondent in the Current Population Survey of March 1993 was classified as

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

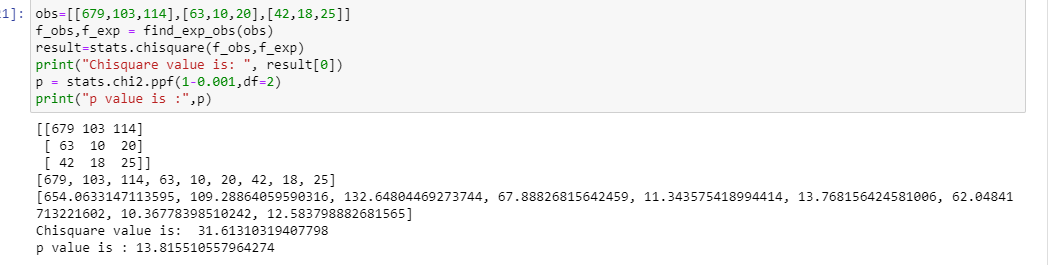
sample.)



Solution:

H0: Men of different marital status has same distribution

H1: Men of different marital status has different distribution



We will reject null hypothesis