**Task 1:**

1. Blood glucose levels for obese patients have a mean of 100 with a standard deviation of 15. A

researcher thinks that a diet high in raw cornstarch will have a positive effect on blood glucose

levels. A sample of 36 patients who have tried the raw cornstarch diet have a mean glucose

level of 108. Test the hypothesis that the raw cornstarch had an effect or not.

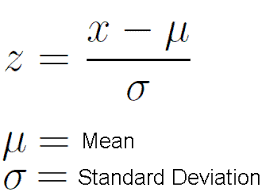
**Solution**

**Step-1:** State the hypotheses. The population mean is 100.

H0: μ= 100  
H1: μ > 100

**Step-2:** Set up the significance level. It is not given in the problem so let’s assume it as 5% (0.05).

**Step-3:** Compute the random chance probability using Z score and Z-table.



Z = (108-100) / (15/√36) = 3.20

Looking at Z- table and p-value associated with 3.20 is 0.9993 i.e. probability of having value less than 108 is 0.9993 and more than or equals to 108 is (1 - 0.9993) = 0.0007.

**Step-4:** It is less than 0.05 so we will reject the Null hypothesis i.e. there is raw cornstarch effect.

**2.** In one state, 52% of the voters are Republicans, and 48% are Democrats. In a second state,

47% of the voters are Republicans, and 53% are Democrats. Suppose a simple random sample

of 100 voters are surveyed from each state.

What is the probability that the survey will show a greater percentage of Republican voters in

the second state than in the first state?

**Solution**

Let’s consider

P1 = the proportion of Republican voters in the first state

P2 = the proportion of Republican voters in the second state

p1 = the proportion of Republican voters in the sample from the first state,

p2 = the proportion of Republican voters in the sample from the second state.

The number of voters sampled from the first state (n1) = 100

The number of voters sampled from the second state (n2) = 100

**Step 1**

Make sure the sample size is big enough to model differences with a normal population.

n1P1 = 100 \* 0.52 = 52,

n1(1 - P1) = 100 \* 0.48 = 48

n2P2 = 100 \* 0.47 = 47

n2(1 - P2) = 100 \* 0.53 = 53

These are each greater than 10, the sample size is large enough.

**Step 2**

Find the mean of the difference in sample proportions:

E(p1 - p2) = P1 - P2

= 0.52 - 0.47

= 0.05.

**Step 3**

Find the standard deviation of the difference.

σ = sqrt{ [ P1(1 - P1) / n1 ] + [ P2(1 - P2) / n2 ] }   
 = sqrt{ [ (0.52)(0.48) / 100 ] + [ (0.47)(0.53) / 100 ] }   
 = sqrt (0.002496 + 0.002491)

= sqrt(0.004987)

= 0.0706

**Step 4**

To find the probability that p1 is less than p2, which is equivalent to finding the probability that p1 - p2 is less than zero. To find this probability, we need to transform the random variable (p1 - p2) into a Z-score. That transformation appears below.

Z p1 - p2 = (x - μ p1 - p2 ) / σd

= (0 - 0.05)/0.0706

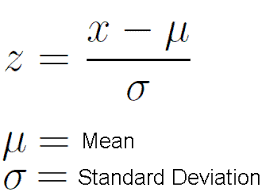
= -0.7082

Z-score being -0.7082 or less is 0.24.

Therefore, the probability that the survey will show a greater percentage of Republican voters in the second state than in the first state is 0.24.

3. You take the SAT and score 1100. The mean score for the SAT is 1026 and the standard deviation is 209. How well did you score on the test compared to the average test taker?

**Solution**



Z = (1100 – 1026) / 209

= 0.354

Z-score of .354 is .6368 or 63.68%.

So, 63.68%.of test-takers scored below 1100.

**TASK 2 :**

1. Is gender independent of education level? A random sample of 395 people were surveyed

and each person was asked to report the highest education level they obtained. The data that

resulted from the survey is summarized in the following table:

High School Bachelors Masters Ph.d. Total

Female 60 54 46 41 201

Male 40 44 53 57 194

Total 100 98 99 98 395

**Question**: Are gender and education level dependent at 5% level of significance? In other

words, given the data collected above, is there a relationship between the gender of an

individual and the level of education that they have obtained?

**Solution**

We have to calculate the expected counts:

To get the expected count for female in High school – (Total Female students surveyed \* Total High school students surveyed ) / Total students surveyed

Similarly other expected counts are calculated.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | High School | Bachelors | Masters | Ph.d. | Total |
| Female | 50.886 | 49.868 | 50.377 | 49.868 | 201 |
| Male | 49.114 | 48.132 | 48.623 | 48.132 | 194 |
| Total | 100 | 98 | 99 | 98 | 395 |

So, working this out, χ2=(60−50.886)^2 /50.886 + (54-49.868)^2/49.868 + (46-50.377)^2/50.377 + (41-49.868)^2/49.868 + (40 – 49.114)^2/49.114 + (44 – 48.132)^2/48.132 + (53-48.623)^2 /48.623 + (57−48.132)^2/48.132

=1.632 + 0.342 + 0.38 + 1.577 + 1.691 + 0.355 + 0.394 + 1.634

=8.006

The critical value of χ2 with 3 degree of freedom is 7.815.

Since 8.006 > 7.815, therefore we reject the null hypothesis and conclude that the education level depends on gender at a 5% level of significance.

**2.** Using the following data, perform a oneway analysis of variance using α=.05. Write up the

results in APA format.

[Group1: 51, 45, 33, 45, 67]

[Group2: 23, 43, 23, 43, 45]

[Group3: 56, 76, 74, 87, 56]

**Solution**

**Group 1**

value mean value – mean (value – mean)^2

1 51 48.2 2.8 7.84

2 45 48.2 -3.2 10.24

3 33 48.2 -15.2 231.04

4 45 48.2 -3.2 10.24

5 67 48.2 18.8 353.44

**Group 2**

value mean value – mean (value – mean)^2

1 23 35.4 -12.4 153.76

2 43 35.4 7.6 57.76

3 23 35.4 -12.4 153.76

4 43 35.4 7.6 57.76

5 45 35.4 9.6 92.16

**Group 3**

value mean value – mean (value – mean)^2

1 56 69.8 -13.8 190.44

2 76 69.8 6.2 38.44

3 74 69.8 4.2 17.64

4 87 69.8 17.2 295.84

5 56 69.8 -13.8 190.44

Sum of squared deviations from the mean (SS) for the groups:

612.8 515.2 732.8

s12 = 612.8 / (5-1) = 153.2

s22 = 515.2 / (5-1) = 128.8

s32 = 732.8 / (5-1) = 183.2

MSE = (153.2 + 128.8 + 183.2)/3 = 155.07

group mean grand mean (group mean - grand mean) (group mean - grand mean)^2

48.2 51.13 -2.93 8.58

35.4 51.13 -15.73 247.43

69.8 51.13 18.67 348.57

Sum of square of means = 604.58

Variance of means=604.58/(3−1)=302.29

MSbetween=(302.29)(5)=1511.45

dfgroups=3−1=2

SSgroup=(1511.45)(3−1)=3022.9

H0:μ1 =μ2 =μ3   
H1: Population means are not equal

**Test statistic and critical value**

F=1511.45/155.07=9.75

Fcritical(2,12)=3.89

Since, 9.75>3.885 We Reject the null hypothesis.

Therefore, the population means are not equal.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **source** | **Sum of Squares** | **Degrees of freedom** | **Mean Squares** | **F** |
| group | 3022.9 | 2 | 1511.45 | 9.75 |
| error | 1860.8 | 12 | 155.07 |  |
| total | 4883.7 |  |  |  |

**Effect size**

η2=3022.9/ 4883.7=0.62

**APA writeup**

*F*(2, 12)=9.75, *p* <0.05, η2=0.62.

3. Calculate F Test for given 10, 20, 30, 40, 50 and 5,10,15, 20, 25.

**Solution**

Sample variance s2 = Σ(X - X̄)2/ N-1

For 10, 20, 30, 40, 50

Mean = (10+20+30+40+50)/5 = 30  
s12 = ((10-30) 2+(20-30) 2+(30-30) 2+(40-30) 2+(50-30) 2))/(5-1)

=(400.0 + 100.0 + 0.0 + 100.0 + 400.0)/4

= 250.0

Mean = (5+10+15+20+25)/5 = 15  
s22 = ((5-15) 2+(10-15) 2+(15-15) 2+(20-15) 2+(25-15) 2)) /(5-1)

=(100.0 + 25.0 + 0.0 + 25.0 + 100.0)/4

= 62.5

F-statistic = s12/s22 = 250/62.5 = 4