ASSIGNMENT 3

Submitted By

Zajiba Sadia Islam

Student ID: 501279357

Group-09

Submitted for

Program: MEng in Biomedical Engineering

Course: Stats for the Health Sciences

BP8101, Winter 2024

Class Timing: Wednesday 3:00-6:00 PM

Hypothesis Testing (Large Samples)

**1.Ans:**

Null hypothesis (H0): The actual percentage of type A donations is equal to 40%. Alternative hypothesis (H1): The actual percentage of type A donations is not equal to 40%.

Significance level (α) of 0.01 for this test.

Given:

* Sample size (n) = 150
* Number of type A donations = 82
* Expected percentage of type A donations = 40%

The binomial distribution is used to conduct this test, approximated by the normal distribution since the sample size is reasonably large.

The expected number of type A donations under the null hypothesis:

Expected number of type A donations = (expected percentage / 100) \* sample size = (40 / 100) \* 150 = 0.40 \* 150 = 60

Standard error = sqrt[(p \* (1 - p)) / n] where p is the expected proportion of type A donations.

Standard error = sqrt[(0.40 \* 0.60) / 150] ≈ 0.04123

z = (Observed - Expected) / Standard error

z = (82 - 60) / 0.04123 ≈ 53.275

For a two-tailed test, the critical z-value is approximately ±2.576.

Since the absolute value of our calculated z-score (|53.275|) is much larger than the critical z-value (2.576), the null hypothesis is rejected.

Therefore, at a significance level of 0.01, the actual percentage of type A donations differs from 40%.

If a significance level of 0.05 had been used, the critical z-value would have been approximately ±1.96. In that case, the null hypothesis would still be rejected because the calculated z-score is much larger than 1.96. Therefore, the conclusion would have been the same.

**2. Ans:**

The confidence interval for the difference between two population means is:

(*x*ˉ1​−*x*ˉ2​)±*t*×*SE*

where:

* *x*ˉ1​ and *x*ˉ2​ are the sample means of the two groups (male and female workers, respectively).
* *SE* is the standard error of the difference between means.
* *t* is the critical value from the t-distribution corresponding to the desired confidence level.

The standard error of the difference between means:

SE=

where:

* *s*1​ and *s*2​ are the sample standard deviations of the two groups.
* *n*1​ and *n*2​ are the sample sizes of the two groups.

Given:

* For males: *n*1​=152, *x*ˉ1​=5.5, and *SE*1​=0.3 (standard error).
* For females: *n*2​=86, *x*ˉ2​=3.8, and *SE*2​=0.2 (standard error).

Calculation of the standard error of the difference between means:

SE=

*SE*≈0.0325

Assuming a 95% confidence level, which corresponds to a t-value with 152+86−2=236 degrees of freedom. From a t-table, the t-value for a 95% confidence level with 236 degrees of freedom is approximately 1.972.

Constructing the confidence interval for the difference between the true average blood lead levels for male and female workers:

(*x*ˉ1​−*x*ˉ2​)±Margin of Error

(5.5−3.8)±1.972×0.0325

1.7±0.0641

So, the 95% confidence interval for the difference between the true average blood lead levels for male and female workers is approximately 1.6359 to 1.7641.

This means there is 95% confidence that the true difference in average blood lead levels between male and female hazardous-waste workers falls between 1.6359 and 1.7641 units.

**3.** **Ans:** Given:

* Sample size (*n*): 20
* Population standard deviation (*σ*): 4
* Sample mean (*x*ˉ): 8
* Significance level (*α*): 0.05
* Null hypothesis (*H*0​): *μ*=7
* Alternative hypothesis (*H*1​): *μ* ≠7

Using the Z-test for the population mean:

*Z*= ​*x*ˉ−*μ/(*​ *σ/√ n*​)

Substituting the given values:

*Z*=​8−7/(4/√20)​

*Z*≈1.118

The critical Z-value for a two-tailed test at a significance level of 0.05 corresponds to a Z-value of approximately ±1.96.

Since the calculated Z-value (1.118) does not exceed the critical Z-value (1.96) in absolute terms, there is failure in rejecting the null hypothesis *H*0​.

In conclusion, at a significance level of 0.05, there is not enough evidence to reject the null hypothesis that the population mean is equal to 7.

* 1. **Ans:**

**a)** Null hypothesis (*H*0​): The average miles driven per vehicle in Chicago is the same as the national average (*μ*=11.1 thousand miles).

Alternative hypothesis (*H*1​): The average miles driven per vehicle in Chicago is different from the national average (*μ* ≠11.1 thousand miles).

(b) Here, the sampling distribution of the sample mean will be used. Since the sample size is large (*n*=36, n>30). Therefore, a Z-test for the population mean can be used.

*Z*= ​*x*ˉ−*μ/(*​ *s/√ n*​)

Given:

* Sample size (*n*): 36
* Sample mean (*x*ˉ): 10.8
* Population mean (*μ*): 11.1
* Standard deviation (*s*): less than 600

Standard Error=600/√36=100

Z-score:

Z=(10.8−11.1)/100​=-0.003

(c) The p-value corresponding to the calculated Z-value, from the standard normal distribution table. At z=-0.003, the P value is 0.9976

(d) Since the p-value is greater than the significance level (*α*=0.05), there is failure to reject the null hypothesis.

(e ) Hence, there is not enough evidence to suggest that the average miles driven per vehicle in Chicago is different from the national average.

**Small Samples TestingTop of Form**

**2.** We know,

Confidence Interval=*x*ˉ±*Z*(s/√n)​

Where:

* *x*ˉ is the sample mean,
* *s* is the sample standard deviation,
* *n* is the sample size, and
* *Z* is the Z-score corresponding to the desired confidence level (in this case, 95%).

For the intervention group:

* *x*ˉ1​ = 54.8 grams
* *s*1​ = 28.1 grams
* *n*1​ = 156

For the control group:

*x*ˉ2​ = 69.5 grams

*s*2​ = 34.7 grams

*n*2​ = 148

For a 95% confidence level, the Z-score is approximately 1.96.

Calculating the confidence intervals for each group:

For the intervention group:

Lower Bound=54.8−1.96(28.1/√156​)

Upper Bound= =54.8+1.96(​28.1/√​156)

​=54.8±3.818

For the control group:

Lower Bound=69.5−1.96(34.7/√148​)

Upper Bound= =69.5+1.96(34.7/√148​)

=69.5±5.684

So, the 95% confidence intervals for the true mean fat intakes of men in each group are approximately:

* Intervention group: (51.982, 57.618)
* Control group: (63.816, 75.184)

To calculate the 95% confidence interval for the true difference in population means:

Confidence Interval=(*x*ˉ1​−*x*ˉ2​)±*Z*

The confidence interval: −14.7±3.529

So, the 95% confidence interval for the true difference in population means is approximately (-18.229, -11.171).

(d) Testing the null hypothesis that the two populations have the same mean carbohydrate intake. A two-sample t-test can be done for this purpose.

* Intervention group:
  + Sample mean (*x*ˉ1​) = 172.5 grams
  + Sample standard deviation (*s*1​) = 68.8 grams
  + Sample size (*n*1​) = unknown
* Control group:
  + Sample mean (*x*ˉ2​) = 185.5 grams
  + Sample standard deviation (*s*2​) = 69.0 grams
  + Sample size (*n*2​) = unknown

Since the sample sizes (*n*1​ and *n*2​) are unknown, the t-statistic cannot be calculated directly. Therefore, the hypothesis test cannot be performed.

**4. (a)** It is appropriate to use a Student’s t-distribution because the population standard deviation (*σ*) is not known, and the sample size is small (less than 30).

Since the sample size is 10, the degrees of freedom (*df*) for the t-distribution would be *n*−1=10−1=9.

(b) Null hypothesis (*H*0​): The mean age of a car when the fuel injection system fails is 48 months (*μ*=48).

Alternative hypothesis (*H*1​): The mean age of a car when the fuel injection system fails is less than 48 months (*μ*<48).

(c) This is a single tail (left-tailed) test because it is being tested whether the mean age is less than 48 months.

(d) The sample test statistic (*t*) can be calculated using the formula:

*t*= ​*x*ˉ−*μ/(*​ *s/√ n*​)

Given:

Sample mean (*x*ˉ): 44.2 months

Population mean (*μ*): 48 months

Sample standard deviation (*s*): approximately 8.61 months

Sample size (*n*): 10

*t*= ​*44.2-48/(*​ *8.61/√ 10*​)

*t*≈−1.397

(e)The p-value at t-value (-1.397) with 9 degrees of freedom is 0.0979

(f) Since the p value is greater than the significance of 0.05, H0 is failed to be rejected.

(g)Interpretation: Since there is failure to reject *H*0​, it means there is not enough evidence to conclude that the mean age is less than 48 months.

Top of Form