**Lab 1: For Ungrouped Data**

Enter the following values in SPSS and calculate mean, s.d, ranges, mode, median.

Weights: 25, 35, 45, 55, 65, 75

**Basic Steps for SPSS**

1. Select Analyze ->Descriptive statistics ->Descriptives

2. Click the Descriptives-> Move Midvalue into Variable(s).

3. Click the Option. select Mean, s.d, ranges, mode, median

4. Click Ok

**Working Expression:**

1. **Mean:** Mean= Sum of all values divided by the number of values

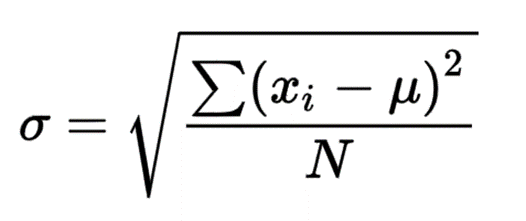
Mean= ∑x/n​

1. **Median:** Arrange the data in ascending order.

If n is odd: Median=Middle value

If n is even: Median=(Middle two values sum)/2

1. **Standard Deviation:** Square root of the average of squared differences from the mean



1. **Range:** Range=Maximum value - Minimum value

**Output:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Weights | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 25.00 | 1 | 14.3 | 16.7 | 16.7 |
| 35.00 | 1 | 14.3 | 16.7 | 33.3 |
| 45.00 | 1 | 14.3 | 16.7 | 50.0 |
| 55.00 | 1 | 14.3 | 16.7 | 66.7 |
| 65.00 | 1 | 14.3 | 16.7 | 83.3 |
| 75.00 | 1 | 14.3 | 16.7 | 100.0 |
| Total | 6 | 85.7 | 100.0 |  |
| Missing | System | 1 | 14.3 |  |  |
| Total | | 7 | 100.0 |  |  |

|  |  |  |
| --- | --- | --- |
| Statistics | | |
| Weights | | |
| N | Valid | 6 |
| Missing | 1 |
| Mean | | 50.0000 |
| Median | | 50.0000 |
| Mode | | 25.00a |
| Std. Deviation | | 18.70829 |
| Range | | 50.00 |
| a. Multiple modes exist. The smallest value is shown | | |

**Interpretation:**

1. Mean=50, mean the average value of weights is 50.
2. Median= 50, means the 50% of the wight is above 50 and 50% is below 50.
3. Mode=25, means the majority of times the weight is 25.
4. Std Deviation=18.708, means 18.708 is the average dispersion from mean.
5. Range= 50, means the difference between highest and lowest weight is 50.

**Lab 2: For Ungrouped Data**

|  |  |  |
| --- | --- | --- |
| Weight | Midvalue | Frequency |
| 20-30 | 25 | 4 |
| 30-40 | 35 | 6 |
| 40-50 | 45 | 7 |
| 50-60 | 55 | 21 |
| 60-70 | 65 | 23 |
| 70-80 | 75 | 2 |

Enter the following values in SPSS and calculate mean, s.d., ranges, mode, median.

**Basic Steps**:

1. Enter the Data Editor Window.

2. Select Data-> Weight Cases.

3. Move Frequency into Frequency Variable.

4. Click Ok. select Analyze ->Descriptive statistics-> Frequencies

5. Click the frequencies-> Move Midvalue into Variable(s)

6. Click the Statistics. select Mean.

7. Click continue. Click ok.

**Working Expression:**

1. **Mean:** Mean= Sum of all values divided by the number of values

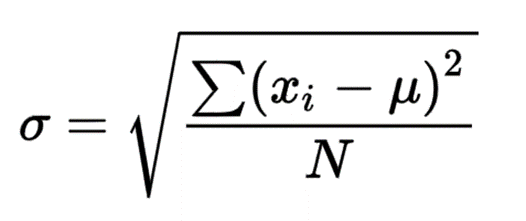
Mean= ∑x/n​

1. **Median:** Arrange the data in ascending order.

If n is odd: Median=Middle value

If n is even: Median=(Middle two values sum)/2

1. **Standard Deviation**: Square root of the average of squared differences from the mean



1. **Range:** Range=Maximum value - Minimum value

|  |  |  |  |
| --- | --- | --- | --- |
| Statistics | | | |
|  | | midvalue | frequency |
| N | Valid | 6 | 6 |
| Missing | 0 | 0 |
| Mean | | 50.0000 | 10.5000 |
| Median | | 50.0000 | 6.5000 |
| Mode | | 25.00a | 2.00a |
| Std. Deviation | | 18.70829 | 9.09395 |
| Range | | 50.00 | 21.00 |
| a. Multiple modes exist. The smallest value is shown | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| midvalue | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 25.00 | 1 | 16.7 | 16.7 | 16.7 |
| 35.00 | 1 | 16.7 | 16.7 | 33.3 |
| 45.00 | 1 | 16.7 | 16.7 | 50.0 |
| 55.00 | 1 | 16.7 | 16.7 | 66.7 |
| 65.00 | 1 | 16.7 | 16.7 | 83.3 |
| 75.00 | 1 | 16.7 | 16.7 | 100.0 |
| Total | 6 | 100.0 | 100.0 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| frequency | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 2.00 | 1 | 16.7 | 16.7 | 16.7 |
| 4.00 | 1 | 16.7 | 16.7 | 33.3 |
| 6.00 | 1 | 16.7 | 16.7 | 50.0 |
| 7.00 | 1 | 16.7 | 16.7 | 66.7 |
| 21.00 | 1 | 16.7 | 16.7 | 83.3 |
| 23.00 | 1 | 16.7 | 16.7 | 100.0 |
| Total | 6 | 100.0 | 100.0 |  |

**Interpretation.**

For Mid-value.

1. Mean=50, mean the average value of weights is 50.
2. Median= 50, means the 50% of the wight is above 50 and 50% is below 50.
3. Mode=25, means the majority of times the weight is 25.
4. Std Deviation=18.708, means 18.708 is the average dispersion from mean.
5. Range= 50, means the difference between highest and lowest weight is 50.

For Frequency.

1. Mean=10.5, mean the average value of weights is 10.5.
2. Median= 6.50, means the 50% of the wight is above 6.50 and 50% is below 6.50.
3. Mode=2, means the majority of times the weight is 2.
4. Std Deviation=9.093, means 9.093 is the average dispersion from mean.
5. Range= 21, means the difference between highest and lowest weight is 21.

**Lab 3: Confidence Interval for Population Mean μ,(ϭ2 Unknown and large n)**

Enter the following values in SPSS and create a confidence interval

assuming normal distribution:

Length: 125, 120, 121, 123, 122, 130, 124, 122, 120, 122, 118, 119,

123, 124, 122, 124, 121, 122, 138, 149, 123, 128, 122, 130, 120, 122,

124, 134, 137, 128, 122, 121, 125, 120, 132, 130, 128, 130, 122, 124.

**Basic Steps for SPSS**

1. Enter the data.

2. Select Analyze ->Compare Means-> One sample T test.

3. Click Options-> Type % (90, 95, 99) confidence interval

4. Click on Continue and then Click OK.

**Working Expression**

95%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| One-Sample Statistics | | | | |
|  | N | Mean | Std. Deviation | Std. Error Mean |
| Length | 40 | 125.2750 | 6.14770 | .97204 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| One-Sample Test | | | | | | |
|  | Test Value = 0 | | | | | |
| t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Length | 128.879 | 39 | .000 | 125.27500 | 123.3089 | 127.2411 |

90%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| One-Sample Statistics | | | | |
|  | N | Mean | Std. Deviation | Std. Error Mean |
| Length | 40 | 125.2750 | 6.14770 | .97204 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| One-Sample Test | | | | | | |
|  | Test Value = 0 | | | | | |
| t | df | Sig. (2-tailed) | Mean Difference | 90% Confidence Interval of the Difference | |
| Lower | Upper |
| Length | 128.879 | 39 | .000 | 125.27500 | 123.6372 | 126.9128 |

99%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| One-Sample Statistics | | | | |
|  | N | Mean | Std. Deviation | Std. Error Mean |
| Length | 40 | 125.2750 | 6.14770 | .97204 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| One-Sample Test | | | | | | |
|  | Test Value = 0 | | | | | |
| t | df | Sig. (2-tailed) | Mean Difference | 99% Confidence Interval of the Difference | |
| Lower | Upper |
| Length | 128.879 | 39 | .000 | 125.27500 | 122.6428 | 127.9072 |

**Interpretation:**

1. For 95% confidence level confidence interval is (123.3089, 127.2411).
2. For 90% confidence level confidence interval is (123.6372, 126.9128).
3. For 99% confidence level confidence interval is (122.6428, 127.9072).

**LAB 4:** **Testing of Hypothesis**

The following values are the lengths of 40 steel rods selected for lab test from a factory.

Length: 125, 120, 121, 123, 122, 130, 124, 122, 120, 122, 118, 119,

123, 124, 122, 124, 121, 122, 138, 149, 123, 128, 122, 130, 120, 122,

124, 134, 137, 128, 122, 121, 125, 120, 132, 130, 128, 130, 122, 124.

Test whether this sample of size 40 has come from a population whose mean length is 125 cm.

**Basic Steps in SPSS**

1. Enter the data in the data editor.

2. Select Analyze ->Compare Means-> One sample T test. Type in Test Value box.

3. Click Options-> Type 95 in confidence interval percentages box.

4. Click on Continue and then Ok.

**Working Expression:**

We wish to test the hypothesis that the samples differ significantly from a hypothesized population mean height of 125 cm. So, we have

**Step 1: Null Hypothesis(H0):** μ = 125

i.e. There is no difference between sample mean and population mean.

**Step 2: Alternate Hypothesis(H1)**: μ ≠ 125

i.e. There is significant difference between sample mean and population mean.

**Step3: Test statistics**

Under H0 Test statistics is given by,

p- value of test statistics (sig. (2-tailed)) = 0.00, compared to α= 0.05.

**Step4: Decision and Conclusion**

P< α, accept H1 and reject H0.

**Conclusion:** There is significant difference between sample mean and population mean.

**Output:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| One-Sample Statistics | | | | |
|  | N | Mean | Std. Deviation | Std. Error Mean |
| Length | 40 | 125.2750 | 6.14770 | .97204 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| One-Sample Test | | | | | | |
|  | Test Value = 0 | | | | | |
| t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Length | 128.879 | 39 | .000 | 125.27500 | 123.3089 | 127.2411 |

**Interpretation:** There is significant difference between sample mean and population mean.

**Lab 5: Hypothesis Testing between two population Means for Matched Paired Samples**

The sales of a product of a company after and before advertisement are as follows: Is advertisement effective at 5 % ?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Month | 1 | 2 | 3 | 4 | 5 | 6 |
| Before X | 120 | 140 | 160 | 140 | 180 | 190 |
| After Y | 200 | 210 | 150 | 200 | 220 | 240 |

**Basic steps in SPSS**

1.Enter the data into Data Editor

2. Select Analyze-> Compare Means ->Paired- Samples T test.

3. Click Options-> Continue-> Ok.

**Working expression:**

**Step 1: Null Hypothesis(H0)**: μx = μy

I.e There is no difference between before and after.

**Step 2: Alternate Hypothesis(H1)**: μx < μy

I.e There is significant difference between before and after.

**Step3: Test statistics**

Under H0 Test statistics is given by,

Tcal=0.465

p- value of test statistics (sig. (2-tailed)) = 0.00, compared to α= 0.05.

**Step4: Decision and Conclusion**

P< α, accept H1 and reject H0.

**Conclusion:** There is significant difference between sample mean and population mean.

**Output:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Paired Samples Statistics | | | | | |
|  | | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | BeforeX | 155.0000 | 6 | 26.64583 | 10.87811 |
| AfterY | 203.3333 | 6 | 30.11091 | 12.29273 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Paired Samples Correlations | | | | |
|  | | N | Correlation | Sig. |
| Pair 1 | BeforeX & AfterY | 6 | .374 | .465 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Paired Samples Test | | | | | | | | | |
|  | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
| Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Pair 1 | BeforeX - AfterY | -48.33333 | 31.88521 | 13.01708 | -81.79481 | -14.87186 | -3.713 | 5 | .014 |

**Interpretation:** There is significant difference between sample mean and population mean.

**Lab 6: Hypothesis Testing When raw data for Independent Samples is given**

The monthly advertising cost of a company for two products X and Y were as follows during 6-month period:

Is there sufficient evidence to conclude that average cost on advertising on product Y is more than on product X.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Cost I (X) | 220 | 240 | 160 | 240 | 280 | 290 | - |
| Cost II (Y) | 100 | 110 | 150 | 100 | 120 | 140 | 145 |

**Basic Steps for SPSS**

1. Enter the data into Data editor.

2. Select Analyze ->Compare means ->Independent samples T Test

3. Move value into Test variable(s) and type into grouping variable

4. Click Define groups and type 1 and 2 into group 2

5. Click Options-> Continue ->Ok.

**Working Expression:**

**Step 1: Null Hypothesis(H0):** μ 1 = μ 2

i.e. There is no significant difference between population mean of two group.

**Step 2: Alternate Hypothesis(H1):** μ 1 < μ 2

i.e. There is significant difference between population mean of two group.

**Step3: Test statistics**

Under H0 Test statistics is given by,

Tcal=5.862

**Step4: Critical Value**

For α=5% level of significance and (n1+n2-2) degree of freedom

Ttab= 1.796

**Step4: Decision and Conclusion**

Ttab<Tcal, accept H1 and reject H0.

**Conclusion:** There is significant difference between population mean of two groups.

**Output:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Group Statistics | | | | | |
|  | group | N | Mean | Std. Deviation | Std. Error Mean |
| cost | 1.00 | 6 | 238.3333 | 46.65476 | 19.04673 |
| 2.00 | 7 | 123.5714 | 21.35304 | 8.07069 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Independent Samples Test | | | | | | | | | | | |
|  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | |
| F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| cost | Equal variances assumed | 1.357 | .269 | 5.862 | 11 | .000 | 114.76190 | 19.57600 | 71.67541 | 157.84840 |
| Equal variances not assumed |  |  | 5.548 | 6.775 | .001 | 114.76190 | 20.68608 | 65.51535 | 164.00846 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Independent Samples Effect Sizes | | | | | |
|  | | Standardizera | Point Estimate | 95% Confidence Interval | |
| Lower | Upper |
| cost | Cohen's d | 35.18658 | 3.262 | 1.502 | 4.964 |
| Hedges' correction | 37.83666 | 3.033 | 1.396 | 4.617 |
| Glass's delta | 21.35304 | 5.374 | 2.213 | 8.512 |
| a. The denominator used in estimating the effect sizes.  Cohen's d uses the pooled standard deviation.  Hedges' correction uses the pooled standard deviation, plus a correction factor.  Glass's delta uses the sample standard deviation of the control group.  **Interpretation:** There is significant difference between population mean of two groups. | | | | | |

**Lab 7: Run Test**

In 30 toss of a coin the following sequence of heads (H) and tail (T) is obtained. H T T H T H H H T H H T T H T H T H H T H T T H T H H T H T

Test at 0.05 level of significance level whether the sequence is random.

**Basic Steps for SPSS:**

1.Enter the data in data editor window.

2. Analyze/ Nonparametric test / Legacy Dialogs / Runs

3. Click Options Select Descriptive and continue

4. Click OK

**Working Expression:**

**Step 1: Null Hypothesis(H0):**

I.e. Sample observation is random.

**Step 2: Alternate Hypothesis(H1):** μ 1 < μ 2

I.e. Sample observation is not random.

**Step3: Test statistics:**

Under H0 Test statistics is given by,

Number of runs(R)=22

Number of head(n1)=14

Number of tails(n2)=16

**Step 4: Critical value**

For α= 0.05 level of significance, n1=16 and n2=14 tabulated value is

Rtab=(10,22)

**Step4: Decision and Conclusion**

R=22∈ (10,22), accept H0 and reject H1.

**Conclusion:** Sample observation is random.

**Output:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Descriptive Statistics | | | | | | | |
|  | N | Mean | | Std. Deviation | | Minimum | Maximum |
| Runs | 30 | 10.9333 | | 6.32964 | | 1.00 | 22.00 |
| Runs Test | | | | |
|  | | | Runs | |
| Test Valuea | | | 10.50 | |
| Cases < Test Value | | | 15 | |
| Cases >= Test Value | | | 15 | |
| Total Cases | | | 30 | |
| Number of Runs | | | 2 | |
| Z | | | -5.017 | |
| Asymp. Sig. (2-tailed) | | | .000 | |

**Interpretation:** Sample observation is random.

**Lab 8: Binomial Test**

Test whether the coin is unbiased from following observations.

Tail Head Head Tail Head Tail Head Head Tail

Head Head Head Tail Head Head Head Head Head Head

Tail Tail Tail Tail Head Tail Tail Tail Tail Tail

Tail Tail Head Tail Tail Tail Head Tail Tail Tail

Tail Head Tail Tail Head Tail Head Tail Tail Tail

**Basic Steps in SPSS**

1.Enter the data in data editor window.

2. Analyze/ Nonparametric test / Legacy Dialogs / Binomial test

3. Click Options Select Descriptive and continue

4. Click OK

**Working Expression:**

**Step 1: Null Hypothesis(H0):** P=1/2

I.e. Head and Tails are equally likely.

**Step 2: Alternate Hypothesis(H1):** P ≠ 1/2

I.e. Head and Tails are not equally likely.

**Step3: Test statistics**

Under H0 Test statistics is given by,

Number of Toss(n)=50

Number of head(n1)=20

Number of tails(n2)=30

**Step 4: Critical value**

Pvalue=0.203

Step4: Decision and Conclusion

Since Pvalue = 0.203>0.05, accept H0 and reject H1.

**Conclusion:** Heads and tails are equally likely.

**Output:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Binomial Test | | | | | | |
|  | | Category | N | Observed Prop. | Test Prop. | Exact Sig. (2-tailed) |
| Toss | Group 1 | Tail | 30 | .60 | .50 | .203 |
| Group 2 | Head | 20 | .40 |  |  |
| Total |  | 50 | 1.00 |  |  |

**Interpretation:** Heads and tails are equally likely.

**Lab 9: One Sample K-S Test**

The number of disease infected tomato plants in 10 different plots of equal size are given

below. Test whether the disease infected plants are uniformly distributed over the entire area

use Kolmogorov Smirnov Test.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plot no. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| No. of infected plants | 8 | 10 | 9 | 12 | 15 | 7 | 5 | 12 | 13 | 9 |

**Basic Steps in SPSS**

1. Start the SPSS program. In the Data editor window, type in the data.

2. Select Analyze-> Nonparametric tests-> Legacy Dialogs-> 1-Samples K-S.

3. Move X into Test Variable List. Then, Click OK.

**Working Expression**:

**Step 1: Null Hypothesis(H0)**: Fe(x)=F0(x)

I.e. disease infected plants are uniformly distributed over the entire area.

**Step 2: Alternate Hypothesis(H1):** Fe(x) ≠F0(x)

I.e. disease infected plants are not uniformly distributed over the entire area.

**Step3: Test statistics:**

Under H0 Test statistics is given by,

Number of Toss(n)=50

Number of head(n1)=20

Number of tails(n2)=30

**Step 4: Critical value**

Pvalue=0.200

Step4: Decision and Conclusion

Since Pvalue = 0.200>0.05, accept H0 and reject H1.

**Conclusion:** Disease infected plants are uniformly distributed over the entire area.

**Output:**

|  |  |  |  |
| --- | --- | --- | --- |
| One-Sample Kolmogorov-Smirnov Test | | | |
|  | | | Noofinfectedplants |
| N | | | 10 |
| Normal Parametersa,b | Mean | | 10.0000 |
| Std. Deviation | | 3.01846 |
| Most Extreme Differences | Absolute | | .146 |
| Positive | | .130 |
| Negative | | -.146 |
| Test Statistic | | | .146 |
| Asymp. Sig. (2-tailed)c | | | .200d |
| Monte Carlo Sig. (2-tailed)e | Sig. | | .786 |
| 99% Confidence Interval | Lower Bound | .775 |
| Upper Bound | .796 |
| a. Test distribution is Normal. | | | |

**Interpretation:** Disease infected plants are uniformly distributed over the entire area.