Chapters (6 + 7 + 8 + 10 + 11 + 12)

Statistical Applications using Minitab

Biostatistics

Prof. Dr. Moustafa Omar Ahmed Abu-Shawiesh

Professor of Statistics

Department of Mathematics
Faculty of Science
Jordan University

First Semester 2023-2024

Finding a Confidence Interval for the Mean (μ)

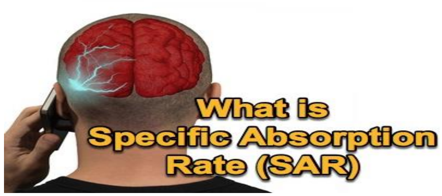
(a) The Z-Confidence Interval for μ

The MINITAB will calculate a confidence interval for the population mean (μ) given the raw data or given the statistics from a sample.

Example

The Specific Absorption Rate (SAR) for a cell phone measures the amount of radio frequency energy absorbed by the user's body when using the handset. The following SAR data are collected from 30 phones:

1.11	1.48	1.43	1.30	1.09	0.455	1.41	0.82	0.78	1.25
1.36	1.34	1.18	1.30	1.26	1.29	0.36	0.52	1.60	1.39
0.74	0.50	0.40	0.867	0.68	0.51	1.13	0.30	1.48	1.38



Use the sample to find the 95% confidence interval of the mean SAR level (μ) for all phones given that population is normal with standard deviation σ = 0.337?

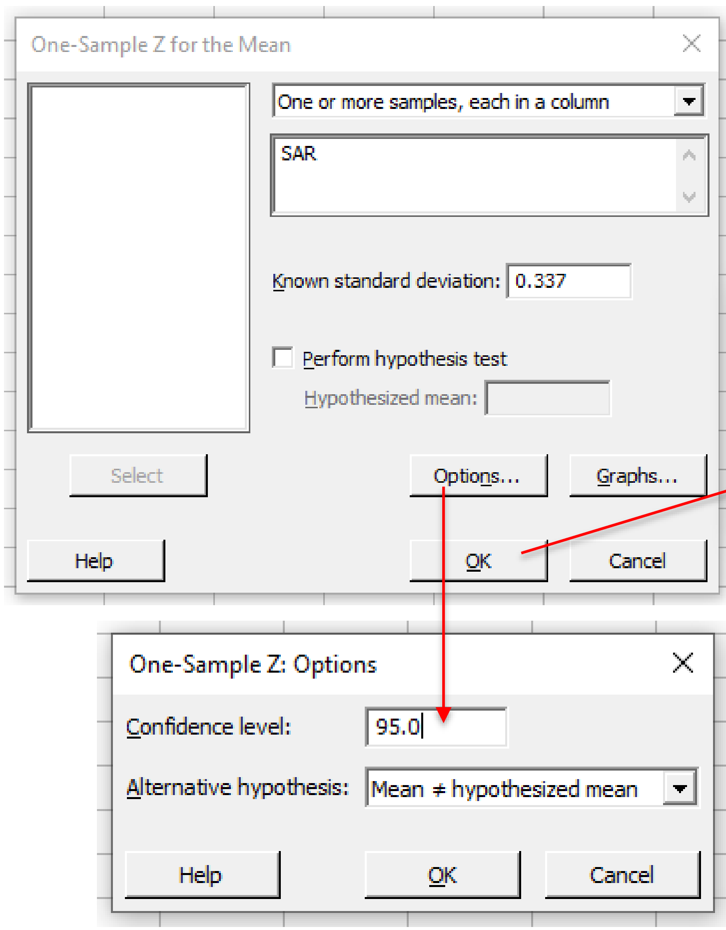
Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into column C1 of MINITAB worksheet.
3. From menu bar Select **File > Save Worksheet As…** .
4. In **File name** write the name SAR and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.
5. Select **Stat > Basic Statistics > 1- Sample Z…** .
6. Double-click C1 for the variable SAR.
7. Click in the box for **Known standard deviation:** and enter **0.337**.

Notation

If the standard deviation for the population **(σ)** is unknown, calculate the sample standard deviation **(S)** and use it.

8. Click in the **[Options]** button to make sure the **Confidence level:** is 95%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
9. Click **OK**.
10. The results will be displayed in a Session Window.



One-Sample Z: SAR

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
30	1.0237	0.4029	0.0615	(0.9031, 1.1443)

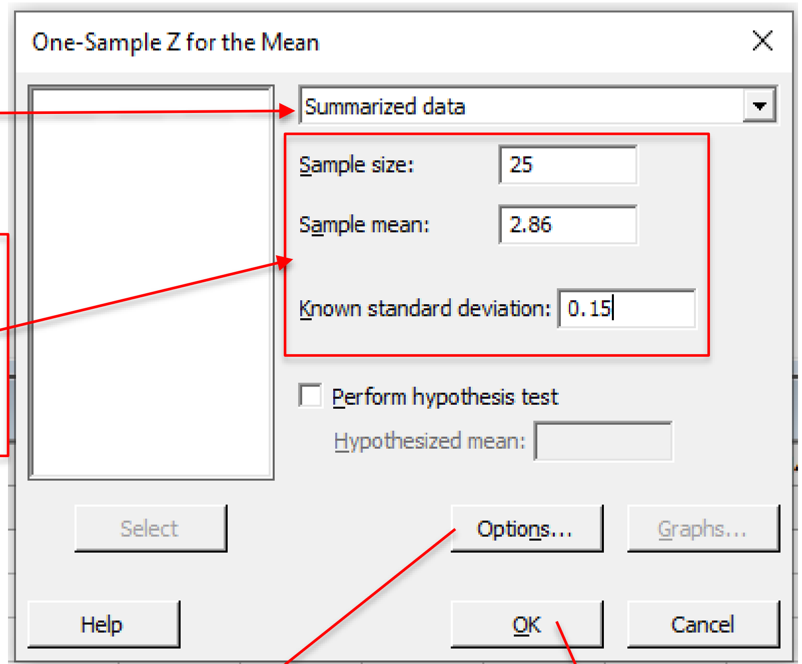
μ: mean of SAR
Known standard deviation = 0.337

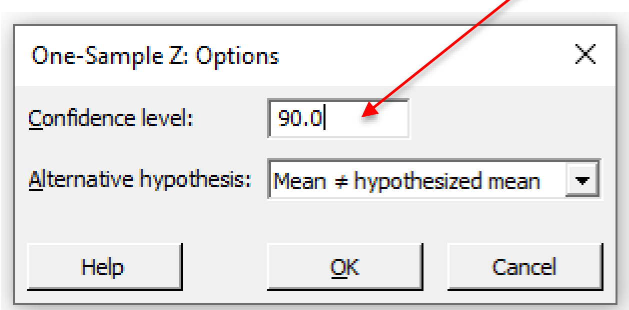
Example

A random sample of 25 economics students selected from HU had a grade point average with a mean of 2.86. Past studies have shown that the standard deviation is 0.15 and the population is normally distributed. Construct a 90% confidence interval for the population mean grade point average (μ)?

Choose this

$$n = 25$$
$$\bar{X} = 2.86 \quad , \sigma = 0.15$$
$$(1 - \alpha)\,100\,\% = 90\,\%$$





One-Sample Z

Descriptive Statistics

N	Mean	SE Mean	90% CI for μ
25	2.8600	0.0300	(2.8107, 2.9093)

μ: mean of Sample
Known standard deviation = 0.15

(b) The t-Confidence Interval for μ

Example

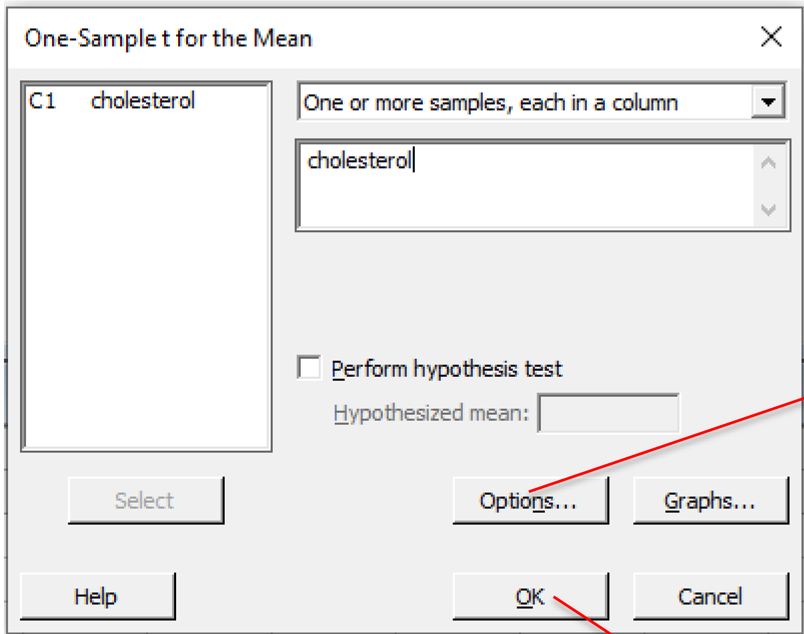
A doctor conducts a small survey with a random sample of size 20 of his patients, measuring their cholesterol levels. Here is his data (the measurements are in m.mol/L):

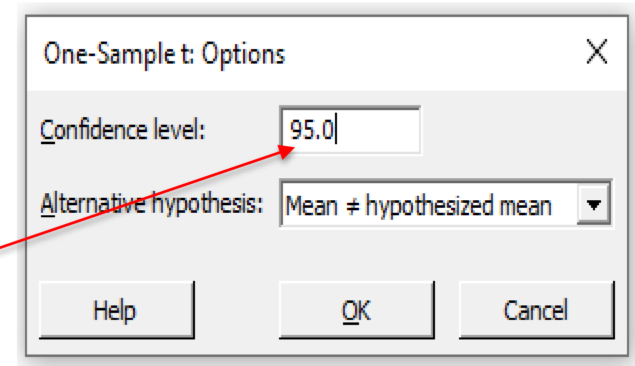
3.6	6.9	5.1	4.2	5.5	7.2	3.0	5.8	4.9	9.9
7.1	5.4	6.2	4.5	6.3	8.2	5.7	4.4	7.9	3.2

Use the sample to find the 95% confidence interval for the mean cholesterol level of his patients given that population is normal?

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into column C1 of MINITAB worksheet.
3. From menu bar Select **File > Save Worksheet As…** .
4. In **File name** write the name cholesterol levels and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.
5. Select **Stat > Basic Statistics > 1- Sample t…** .
6. Double-click C1 for the variable cholesterol.
7. Click in the **[Options]** button to make sure the **Confidence level:** is 95%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
8. Click **OK**.
9. The results will be displayed in a Session Window.





One-Sample T: cholesterol

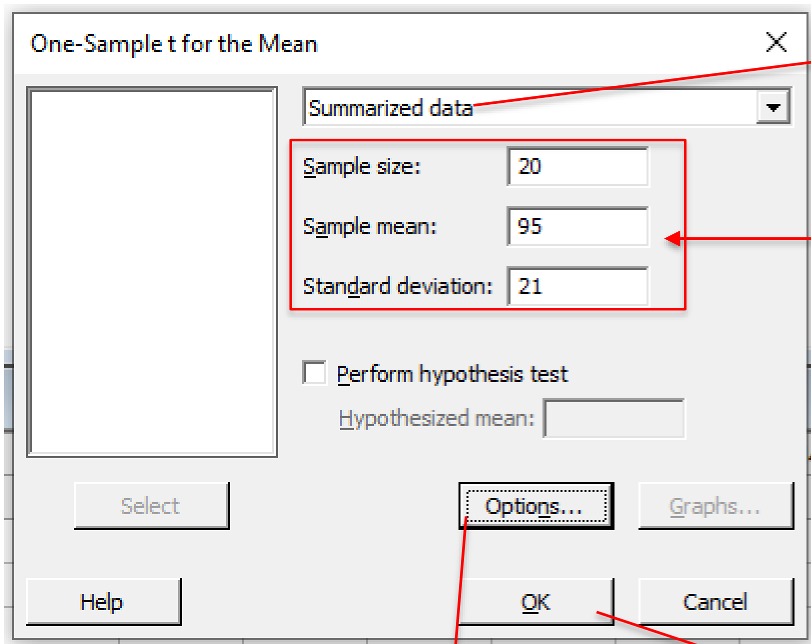
Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
20	5.750	1.768	0.395	(4.922, 6.578)

μ: mean of cholesterol

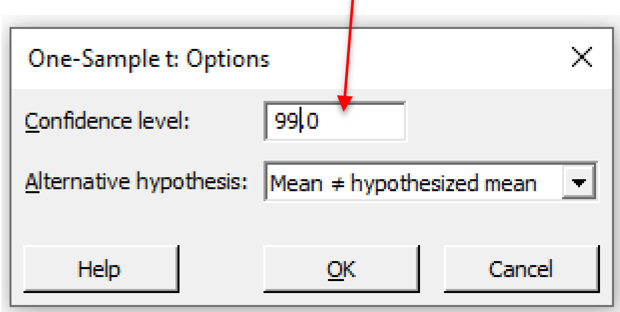
Example

In a random sample of 20 customers at a given supermarket in Jordan, the mean waiting time to get service is 95 seconds, and the standard deviation is 21 seconds. Assume the wait times are normally distributed, then construct a 99% confidence interval for the mean wait time of all customers (μ)?



Choose this

$$n = 20$$
$$\bar{X} = 95 \quad , S = 21$$
$$(1 - \alpha)\ 100\ \% = 99\ \%$$



One-Sample T

Descriptive Statistics

N	Mean	StDev	SE Mean	99% CI for μ
20	95.00	21.00	4.70	(81.57, 108.43)

μ: mean of Sample

Finding a Confidence Interval for the Proportion (p)

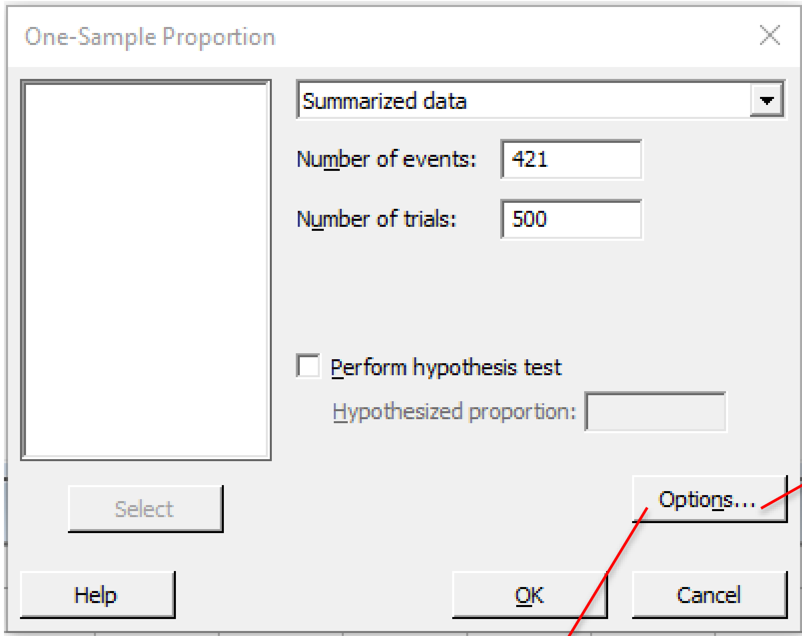
The MINITAB will calculate a confidence interval for the population proportion (p) given the raw data or given the statistics from a sample.

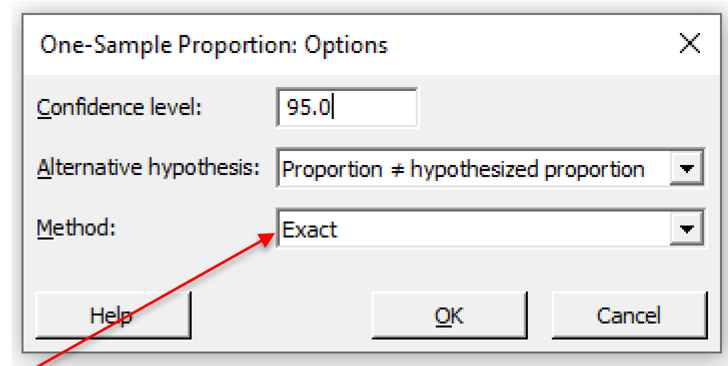
Example

Suppose that a ministry of health in a certain country is interested to estimate the percent of adults living in a large city who have COVID-19. A random sample of 500 adult residents in this city are tested to determine whether they have COVID-19. Suppose that out of the 500 people tested, 421 are infected. Construct a 95% confidence interval for the true proportion (p) of adult residents of this city who have COVID-19?

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Select **Stat > Basic Statistics > 1 Proportion….**
3. Click on the button for **Summarized data**.
4. Click in the box for **Number of events** and enter **421**.
5. In the **Number of trials** box enter **500**.
6. Click in the **[Options]** button to make sure the **Confidence level:** is 95%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
7. Click **OK**.
8. The results will be displayed in a Session Window.





Test and CI for One Proportion

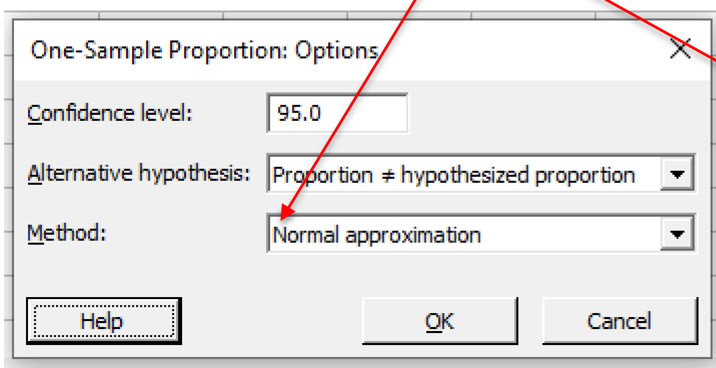
Method

p: event proportion

Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
500	421	0.842000	(0.807022, 0.872869)



Test and CI for One Proportion

Method

p: event proportion

Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
500	421	0.842000	(0.810030, 0.873970)

Hypothesis Test for the Mean (μ)

(a) The Z-Distribution

The MINITAB will show how to calculate the **test statistic** and the **p-value**. The **p-value** does not require a critical value from the table. If the **p-value is** $\leq$ α (Significance Level), the null hypothesis H_0 is rejected, otherwise do not reject (accept) H_0.

Example

A sociologist in Jordan wishes to see if it is true that, for a certain group of professional women, the average age in years at which they have first child (μ) is 28.6. A random sample of size 36 women is selected and their ages at the birth of their first children are recorded as follows:

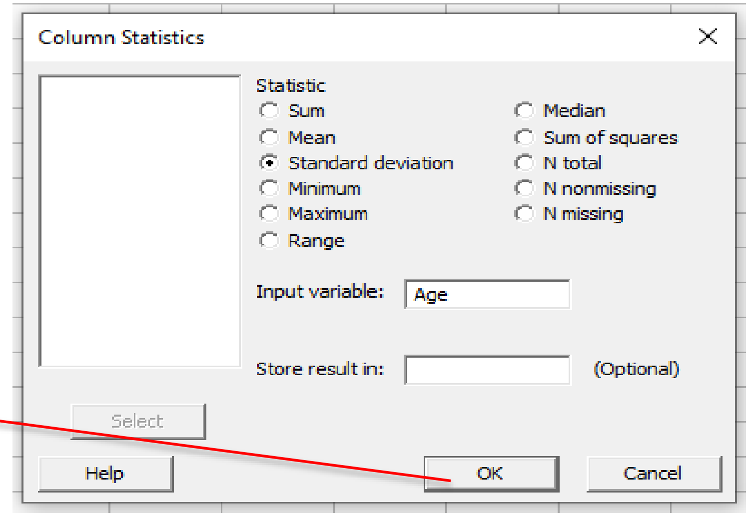
32	28	26	33	35	34
29	24	22	25	26	28
28	34	33	32	30	29
30	27	33	34	28	25
24	33	25	37	35	33
34	36	38	27	29	26

At α = 0.05, does the sociologist's claim is true assuming that the distribution is normal?

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into column C1 of MINITAB worksheet and name the column Age.
3. From menu bar Select **File > Save Worksheet As...** .
4. In **File name** write the name Birth and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.

5. In this example sigma (σ) is unknown. The standard deviation for the sample (**S**) will be calculated and used as an estimate for the population standard deviation (**σ**).
   (a) Select **Calc > Column Statistics...** .
   (b) Check the button for **Standard deviation**. You can only do one of these statistics.
   (c) Use Age for the **Input variable:**.
   (d) Click **OK**.

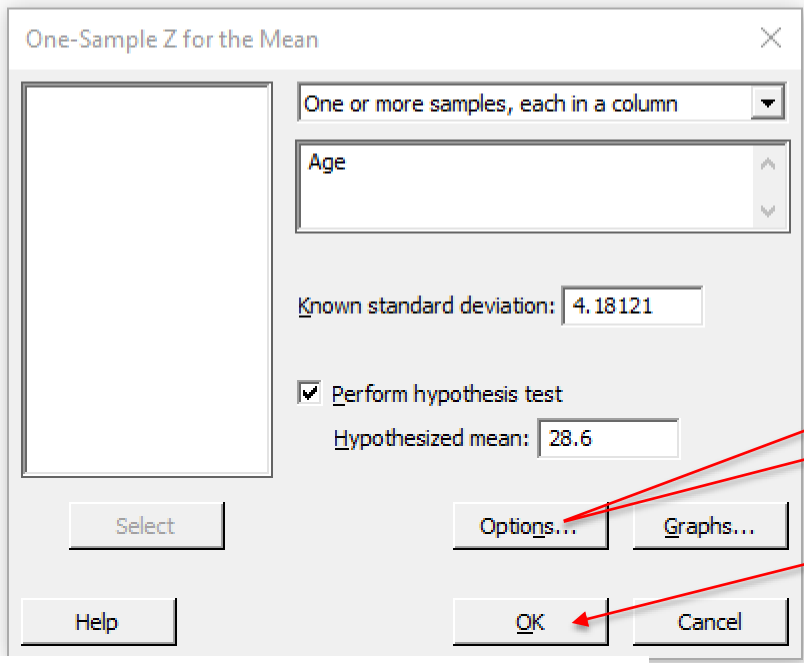


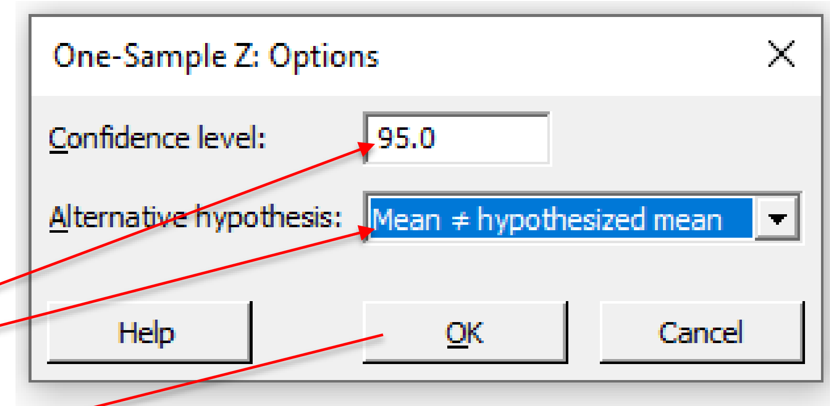
Standard Deviation of Age

Standard deviation of Age = 4.18121

6. Select **Stat > Basic Statistics > 1- Sample Z...** .
7. Choose the Age variable.
8. Click in the text box for **Known standard deviation:** and type in the sample standard deviation, calculated in step 5.
9. Click the button for **Perform hypothesis test** and enter the **Hypothesized mean** value **28.6** in the box.
10. Click in the **[Options]** button to check the form of the **Alternative Hypothesis** and to make sure that the value of the **Confidence level** is 95%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
11. Click **OK**.

12. The results will be displayed in a Session Windows as shown below:





Test

Null hypothesis H_0: μ = 28.6
Alternative hypothesis H_1: μ ≠ 28.6

Z-Value	P-Value
2.09	0.037

13. Decision

(a) Critical Value Approach

We have:

$$|Z| = 2.09 > Z_{1-(\alpha/2)} = 1.96$$

then we reject H_0.

(b) P-Value Approach

Since the p-value = 0.037 is less than $\alpha = 0.05$, then we reject H_0.

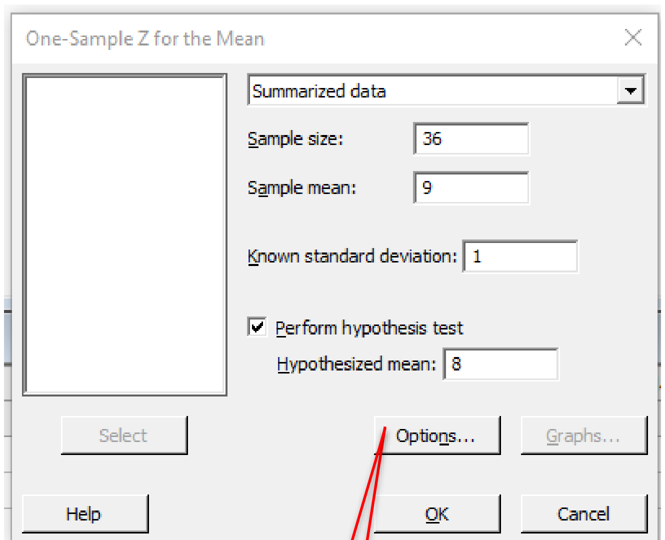
14. Conclusion

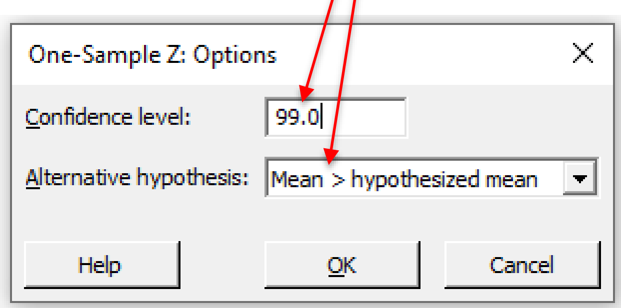
There is enough evidence in the sample to conclude that the average age in years for this certain group of professional women at which they have first child (μ) is not equal to 28.6 year, that is $\mu \neq 28.6$ year. At $\alpha = 0.05$, the sociologist's claim is NOT true.

Example

A recent study stated that if a person chewed gum, the average number of sticks of gum he or she chewed daily (μ) was 8. To test the claim, a researcher selected a random sample of 36 gum chewers and found that the mean number of sticks of gum chewed per day is 9 and the standard deviation is 1. At $\alpha = 0.01$, is the number of sticks of gum a person chews per day actually greater than 8 assuming that the distribution is normal?

Solution





One-Sample Z

Descriptive Statistics

N	Mean	SE Mean	99% Lower Bound for μ
36	9.000	0.167	8.612

μ: mean of Sample
Known standard deviation = 1

Test

Null hypothesis $H_0: \mu = 8$
Alternative hypothesis $H_1: \mu > 8$

Z-Value	P-Value
6.00	0.000

Conclusion

There is enough evidence to support the claim that the average is more than eight ($\mu > 8$).

Decision

(a) Critical Value Approach

We have:

$Z = 6 > Z_{1-\alpha} = 2.33$ then we reject H_0.

(b) P-Value Approach

Since the p-value = 0.000 is less than $\alpha = 0.01$, then we reject H_0.

(b) The t-Distribution

Example

A coach in a health club claims that the average salary of employees in health clubs in Jordan (μ) is less than JD 60 per week. A random sample of size 8 health clubs is selected and the weekly salaries in JD are recorded as follows:

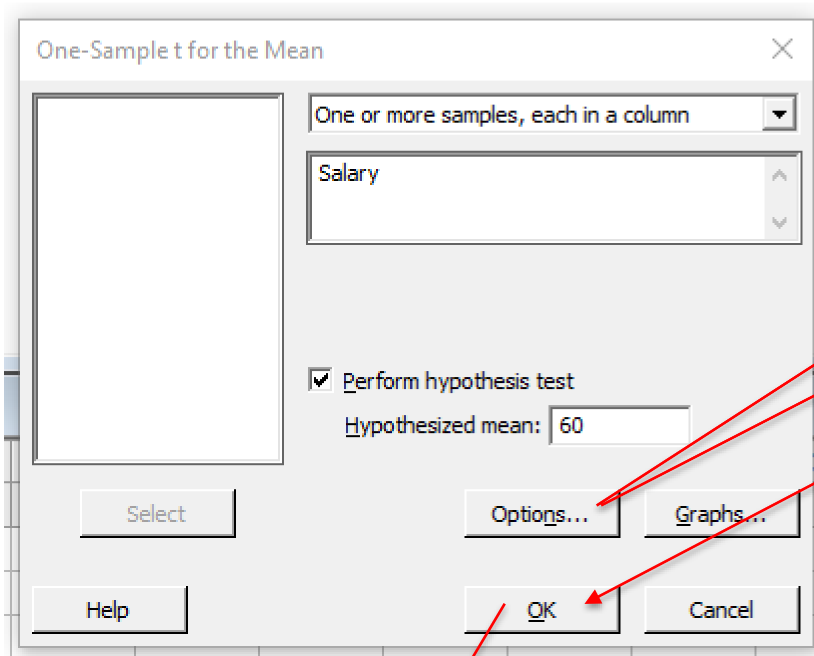
60, 56, 60, 55, 70, 55, 60, 55

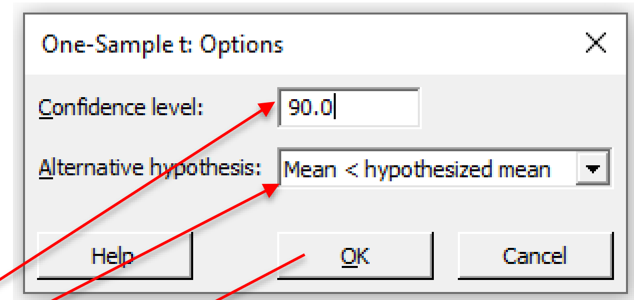
At α = 0.10, is there enough evidence to support the coach's claim assuming that the distribution is normal?

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into column C1 of MINITAB worksheet and name the column Salary.
3. From menu bar Select **File > Save Worksheet As…** .
4. In **File name** write the name Weekly Salary and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.
5. Select **Stat > Basic Statistics > 1- Sample t…** .
6. Double-click C1 for the variable Salary.
7. Click the button for **Perform hypothesis test** and enter the **Hypothesized mean** value **60** in the box.
8. Click in the **[Options]** button to check the form of the **Alternative Hypothesis** and to make sure that the value of the **Confidence level** is 90%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
9. Click **OK**.

10. The results will be displayed in a Session Windows as shown below:





One-Sample T: Salary

Descriptive Statistics

N	Mean	StDev	SE Mean	90% Upper Bound for μ
8	58.88	5.08	1.80	61.42

μ: mean of Salary

Test

Null hypothesis H_0: μ = 60

Alternative hypothesis H_1: μ < 60

T-Value	P-Value
-0.63	0.276

11. Decision

(a) Critical Value Approach

We have:

$$t = -0.63 > -t_{(0.1\,,\,7)} = -1.895$$

then we do not reject H_0.

(b) P-Value Approach

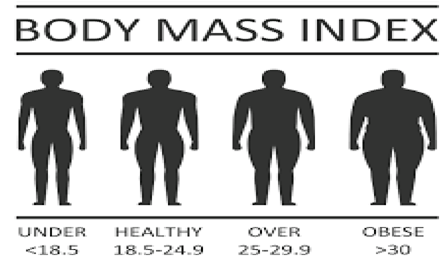
Since the p-value = 0.276 is more than α = 0.10, then we do not reject H_0.

12. Conclusion

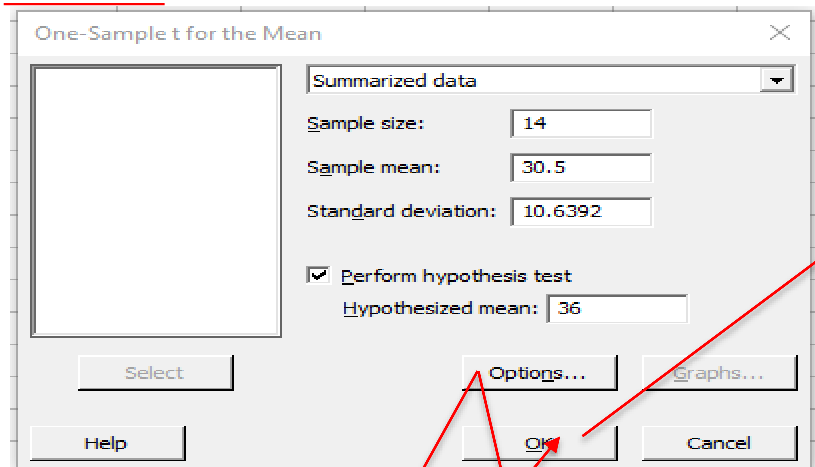
There is not enough evidence to support the coach's claim that the average salary (μ) of employees of health clubs in Jordan is less than JD 60 per week, that is $\mu = 60$ JD. At α = 0.10, the coach's claim is NOT true.

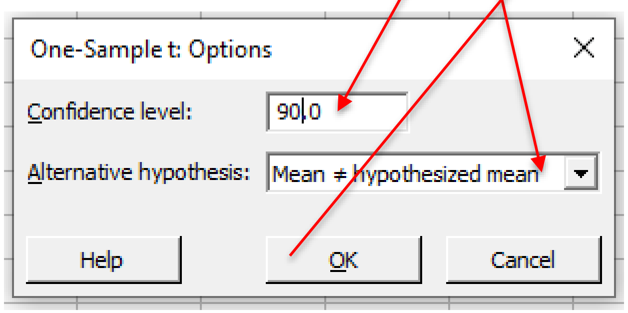
Example

The body mass index (BMI) of a group of 14 healthy adult males has a mean of 30.5 and a standard deviation of 10.6392, can we conclude that the mean BMI of the population is equal to 36 assuming that the population is normally distributed? Use α = 0.10 to test the hypothesis?



Solution





One-Sample T

Descriptive Statistics

N	Mean	StDev	SE Mean	90% CI for μ
14	30.50	10.64	2.84	(25.46, 35.54)

μ: mean of Sample

Test

Null hypothesis $H_0: \mu = 36$
Alternative hypothesis $H_1: \mu \neq 36$

T-Value	P-Value
-1.93	0.075

Decision

(a) Critical Value Approach

We have:

$$|t| = |-1.93| = 1.93 > t_{(0.05\,,\,13)} = 1.771$$

then we reject H_0.

(b) P-Value Approach

Since the p-value = 0.075 is less than α = 0.10, then we reject H_0.

Conclusion: We conclude that the mean BMI of the population is not equal to 36, that is, $\mu \neq 36$.

The Z-Test for the Proportion (p)

The MINITAB will calculate the **test statistic** and the **p-value** for a test of a proportion (p) given the statistics from a sample or given the raw data.

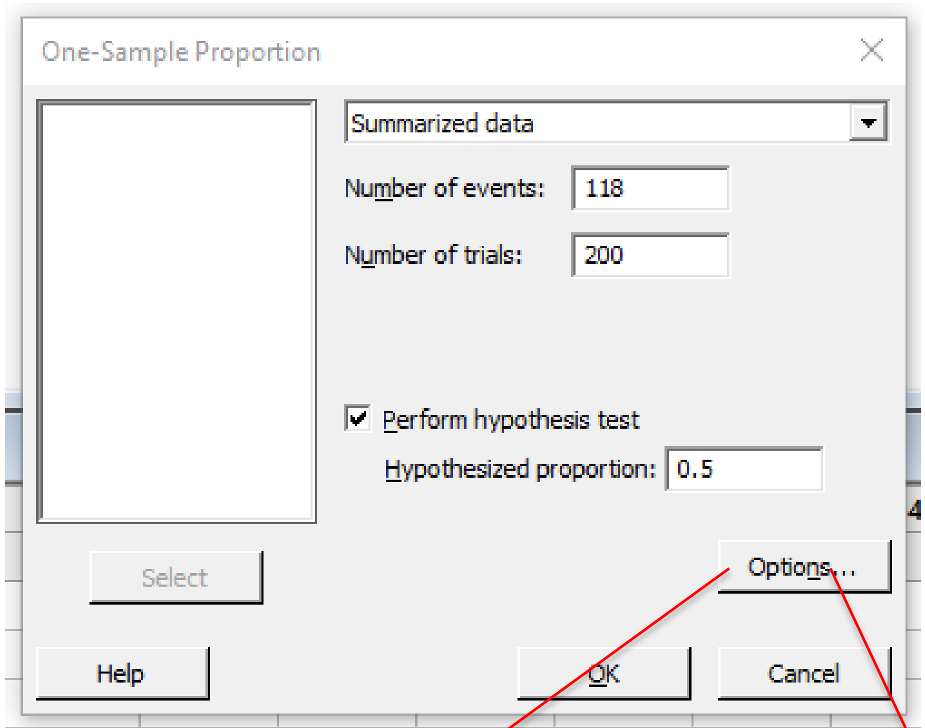
Example

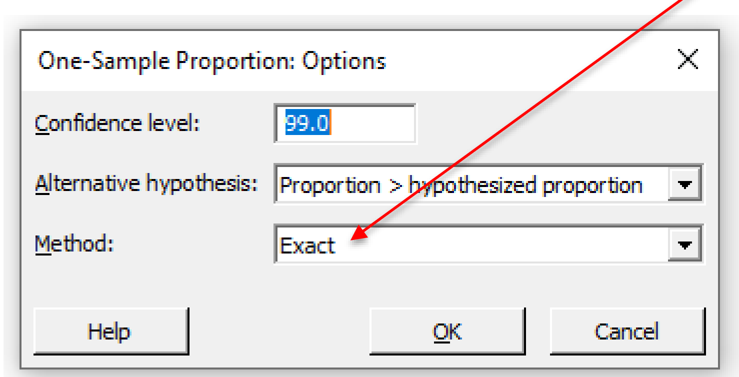
The coach of the national football team in Jordan believes that the chance of the team winning a match is greater than 50%. In a random sample of size 200 matches, the team won 118 times. Is there enough evidence to suggest that the coach believes is correct? Conduct a hypothesis test using alpha = 0.01 assume normal distribution?

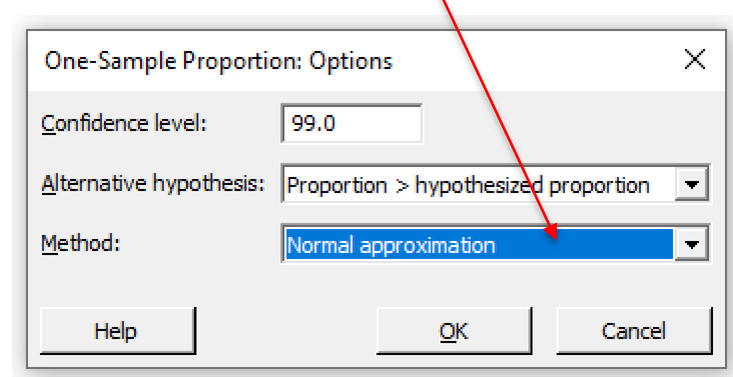
Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Select **Stat > Basic Statistics > 1- Proportion…** .
3. Click on the button **Summarized data**.
4. Click on the box **Number of events** and enter 118.
5. Click on the box **Number of trials** and enter 200.
6. Click the button for **Perform hypothesis test** and enter the **Hypothesized proportion** value **0.5** in the box.
7. Click in the **[Options]** button to check the form of the **Alternative Hypothesis** and to make sure that the value of the **Confidence level** is 99%. You may need to click inside the textbox before you type the new confidence level. Click **OK**.
8. Click **OK**.

9. The results will be displayed in a Session Window.







Test and CI for One Proportion

Method

p: event proportion
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	99% Lower Bound for p
200	118	0.590000	0.505448

Test

Null hypothesis H_0: p = 0.5
Alternative hypothesis H_1: p > 0.5

P-Value
0.007

Test and CI for One Proportion

Method

p: event proportion
Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	99% Lower Bound for p
200	118	0.590000	0.509095

Test

Null hypothesis H_0: p = 0.5
Alternative hypothesis H_1: p > 0.5

Z-Value	P-Value
2.55	0.005

Decision

(a) Critical Value Approach

We have:

$Z = 2.55 > Z_{1-\alpha} = 2.33$ then we reject H_0.

(b) P-Value Approach

Since the p-value = 0.005 is less than α = 0.01, then we reject H_0.

Conclusion

There is enough evidence to support the claim that the proportion is more than 50% (p > 0.5). This means that the coach believes is correct.

Correlation Methods

Example

In a study between age (x) and systolic blood pressure (y) for a random sample of size n = 6 patients selected from King Abdulla University Hospital (KAUH), the following data was obtained:

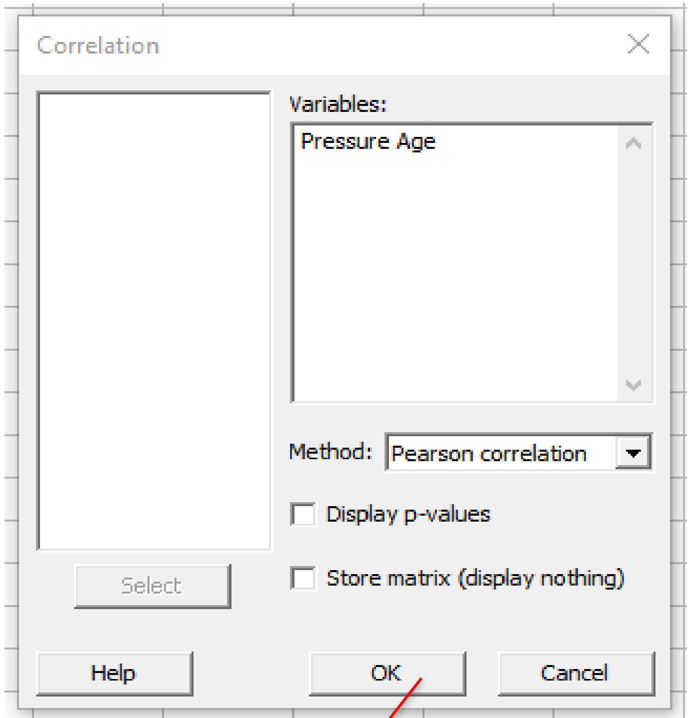
Age (x)	Systolic Blood Pressure (y)
43	128
48	120
56	135
61	143
67	141
70	152

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into two columns C1 (Age) and C2 (Pressure) in the MINITAB worksheet.
3. From menu bar Select **File > Save Worksheet As…** .
4. In **File name** write the name Blood Pressure and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.

↓	C1	C2
	Age	Pressure
1	43	128
2	48	120
3	56	135
4	61	143
5	67	141
6	70	152

5. Select **Stat > Basic Statistics > Correlation…** .
6. Double click on Pressure and double click Age. The **dependent variable** should be first.
7. Click **OK**.

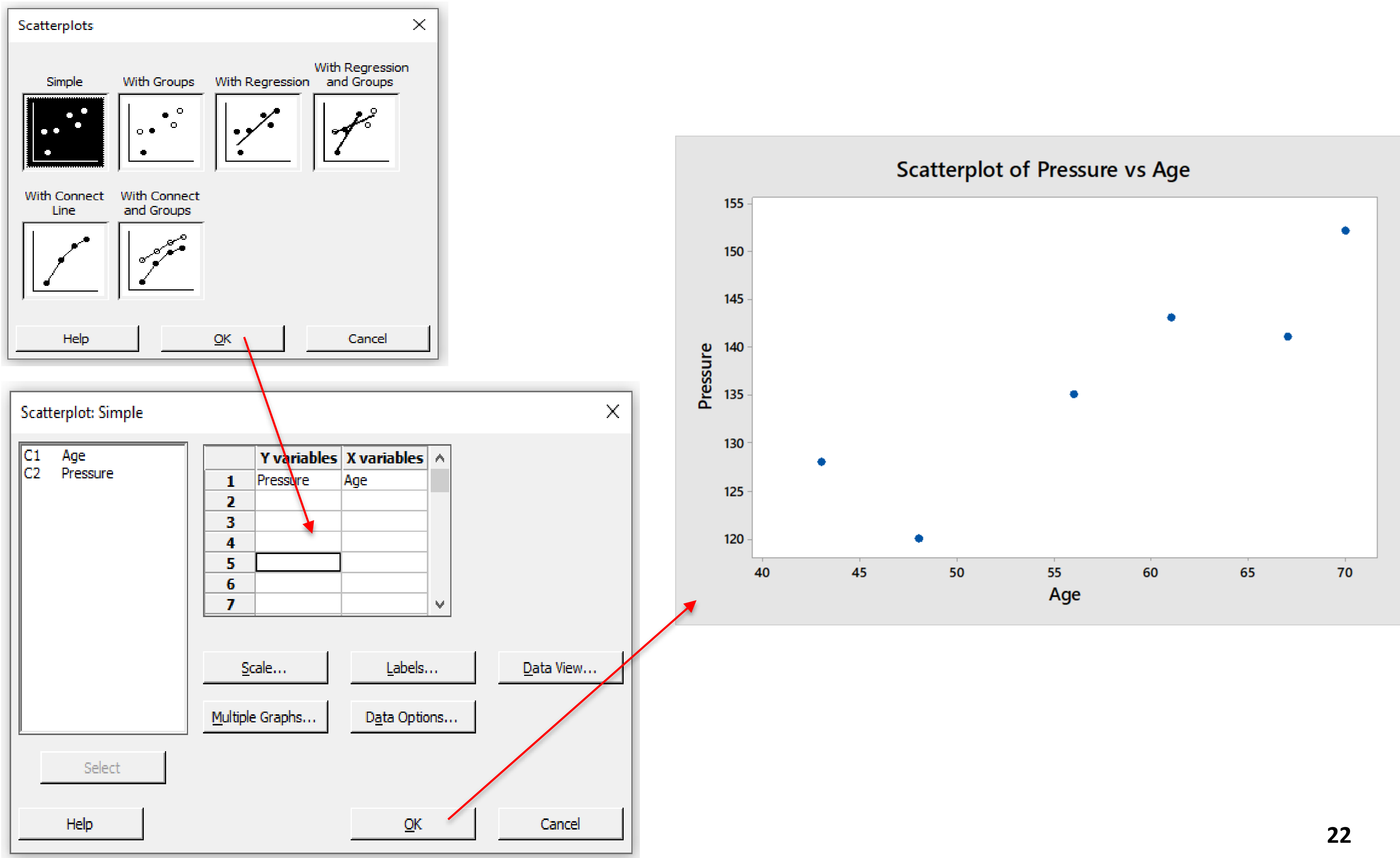


Correlation: Pressure, Age

Correlations

Pearson correlation 0.897

8. To draw a scatter plot select **Graph > Scatterplot…** .
9. Choose **Simple**.
10. Click **Ok**.
11. Double click on Pressure for Y variable and Age for X variable. Click **Ok**..



One-Way Analysis of Variance (ANOVA)

Example

A researcher wishes to try three different techniques to lower the blood pressure of individuals diagnosed with high blood pressure. The subjects are randomly assigned to three groups; the first group takes medication, the second group exercises, and the third group diets. After four weeks, the reduction in each person's blood pressure is recorded. At $\alpha = 0.05$, test the claim that there is no difference among means. The following data was obtained:

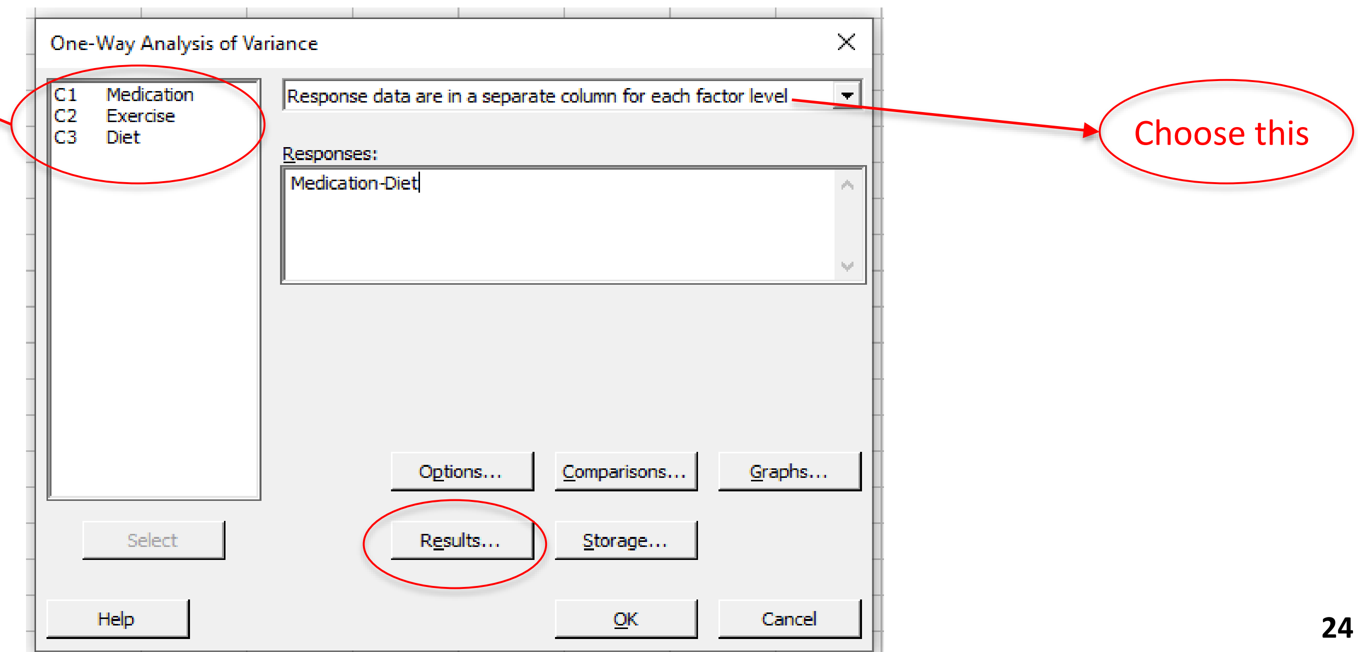
Medication	Exercise	Diet
10	6	5
12	8	9
9	3	12
15	0	8
13	2	4

Steps

1. Start your **Minitab** program by double click on the icon Minitab 17 .
2. Enter the data into three columns C1 (Medication), C2 (Exercise) and C3 (Diet) in the MINITAB worksheet.
3. From menu bar Select **File > Save Worksheet As…** .
4. In **File name** write the name Diet and determine the place where you want to save your data (Desktop, Folder, ….) then press on **Save** to complete the process.

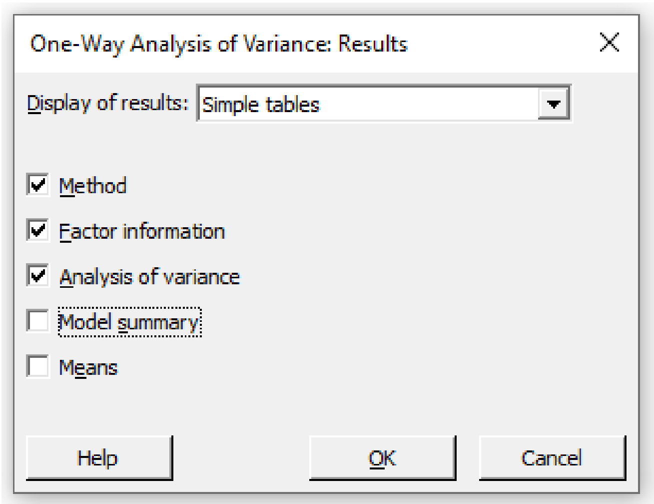
↓	C1	C2	C3
	Medication	Exercise	Diet
1	10	6	5
2	12	8	9
3	9	3	12
4	15	0	8
5	13	2	4

5. Select **Stat > ANOVA > One-Way…** .
6. Drag the mouse over the three columns C1-C3 in the list box and then click **[Select]**.



7. Click on **Results** and choose from it.

8. Click on **OK**.



One-way ANOVA: Medication, Exercise, Diet

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0.05

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	3	Medication, Exercise, Diet

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	160.1	80.067	9.17	0.004
Error	12	104.8	8.733		
Total	14	264.9			

9. Decision

(a) Critical Value Approach

We have F = 9.17 > $F_{(2,12,0.05)}$= 3.89

(b) P-Value Approach

Since the p-value = 0.004 is less than $\alpha = 0.05$, then we reject H_0.

10. Conclusion

There is enough evidence to reject the claim and conclude that at least one mean is (one method) is different from the others.

Fisher's Least Significant Difference (LSD) Method for Multiple Comparisons

Fisher's LSD

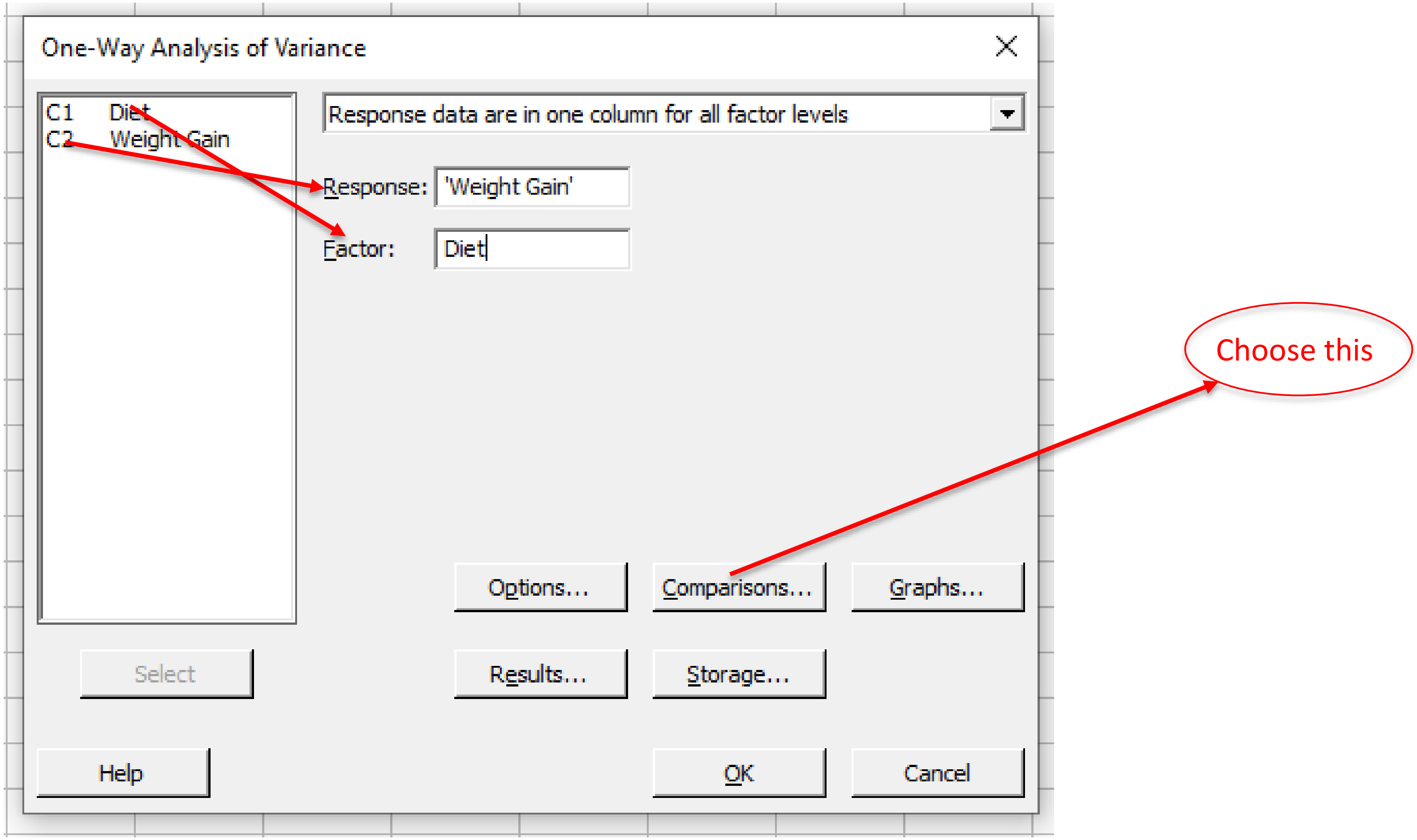
Step (1)

1. Enter the category labels in Column C1.
2. Enter the corresponding data value in Column C2.

↓	C1	C2
	Diet	Weight Gain
1	1	16
2	1	15
3	1	13
4	1	21
5	1	15
6	2	18
7	2	22
8	2	20
9	2	16
10	2	24
11	3	26
12	3	31
13	3	24
14	3	30
15	3	24

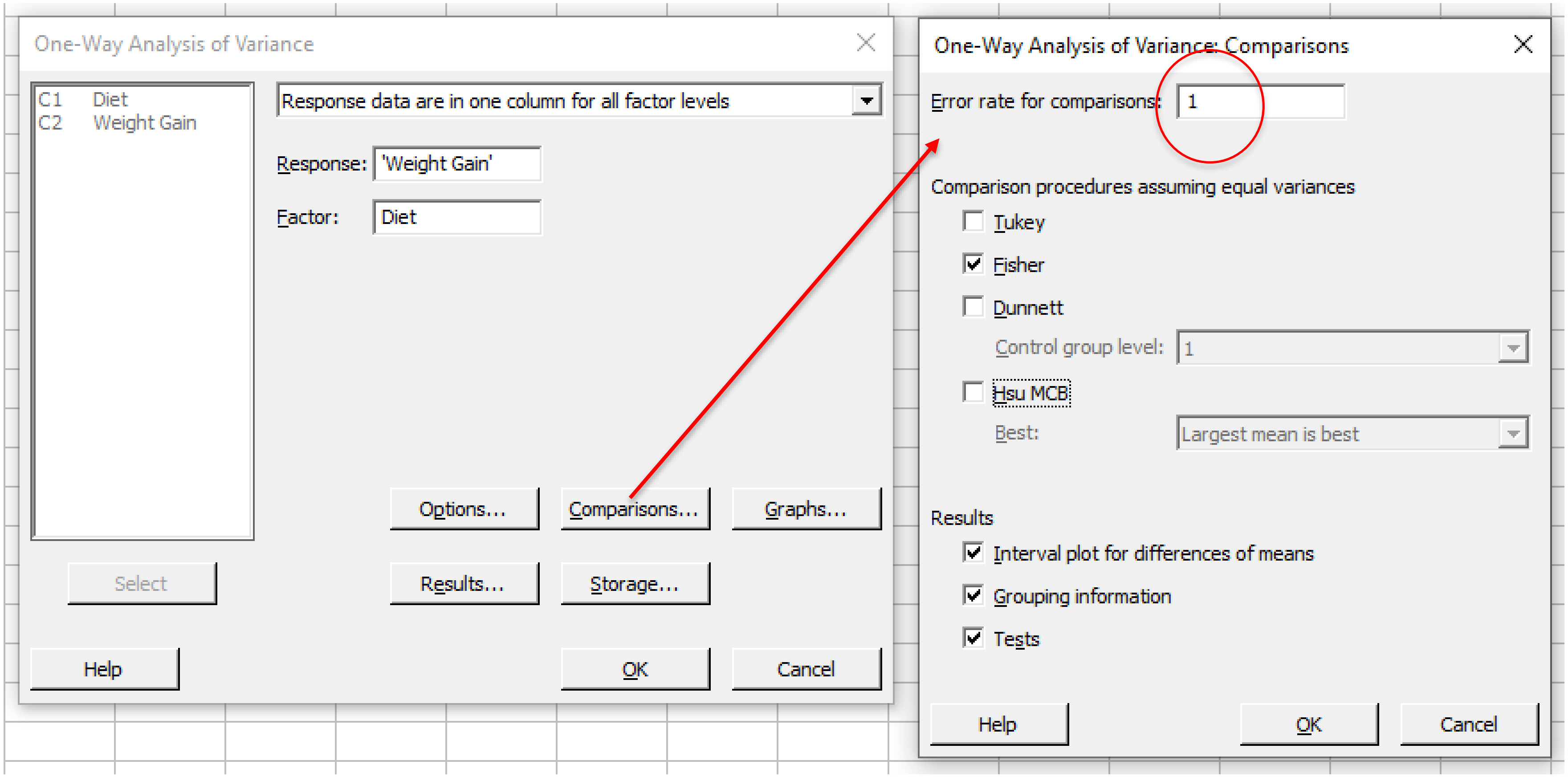
Step (2)

1. Choose Stat, ANOVA, and One-Way.
2. Enter the data (C2) for the Response.
3. Enter the categories (C1) for the Factor.
4. Click on Comparisons.

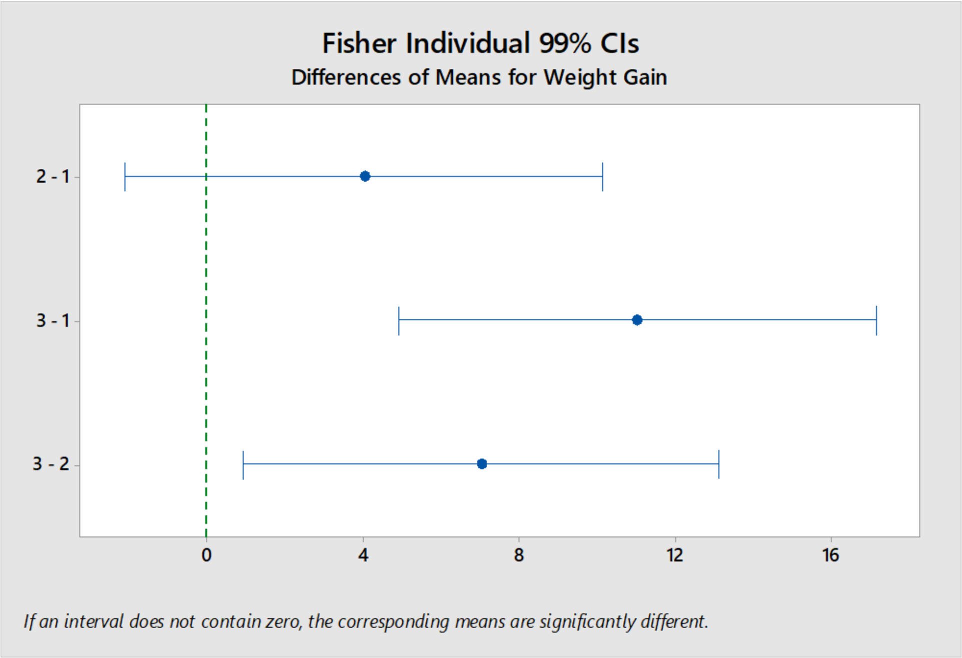


Step (3)

1. Check the box for Fisher and for Tests .
2. Observe the results of Fisher's LSD test.



Step (4): Results of Fisher's LSD test.



Fisher Individual Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	99% CI	T-Value	Adjusted P-Value
2 - 1	4.00	2.00	(-2.11, 10.11)	2.00	0.069
3 - 1	11.00	2.00	(4.89, 17.11)	5.50	0.000
3 - 2	7.00	2.00	(0.89, 13.11)	3.50	0.004

Fisher Pairwise Comparisons

Grouping Information Using the Fisher LSD Method and 99% Confidence

Diet	N	Mean	Grouping	
3	5	27.00	A	
2	5	20.00		B
1	5	16.00		B

Means that do not share a letter are significantly different.

(a) The Paired t Test and Interval Estimation

Example

The weights (in kgs) for a random sample of six patients selected from the Jordan University Hospital (JUH) before and after special exercise program are recorded in the following table:

Patient Number	1	2		3	4	5	6
Before	65	75		82	90	105	98
After	68	70		72	85	95	9

Answer the following:

(a) Construct the 95% confidence interval (CI) for the mean μ_d of the population paired differences?

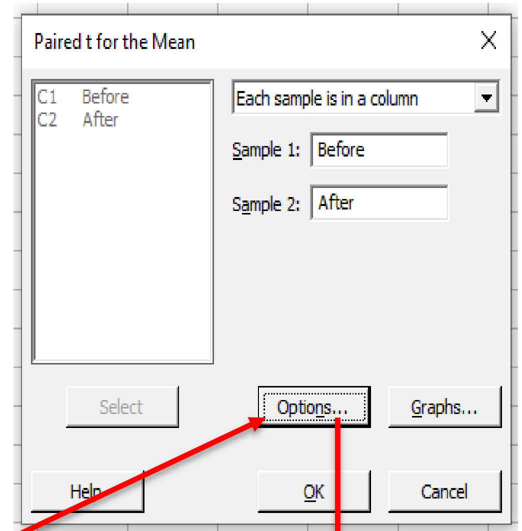
(b) Can we conclude that there is a difference in weights of patients before and after the exercise program? Test using $\alpha = 0.01$?

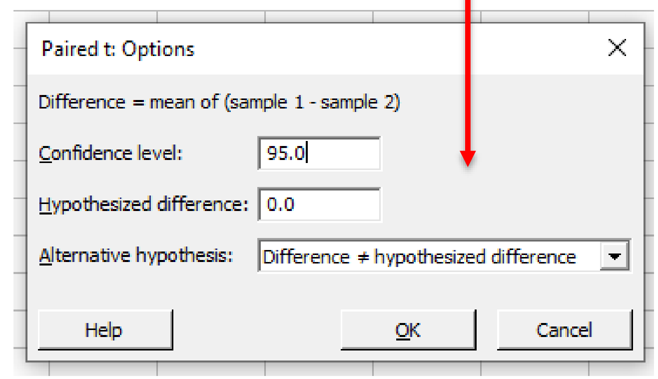
(c) Calculate the p-value for the test in (b)?

Step (1)

1. Open Minitab and enter the data, then save it.
2. Choose Stat > Basic Statistics > Paired t.

+	C1	C2
	Before	After
1	65	68
2	75	70
3	82	72
4	90	85
5	105	95
6	98	95





Step (2)

1. From the drop-down list Select Each sample is in a column.
2. In Sample 1, enter Before.
3. In Sample 2, enter After.

Step (3): Click OK.

Paired T-Test and CI: Before, After

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Before	6	85.83	14.82	6.05
After	6	80.83	12.48	5.10

Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for μ_difference
5.00	4.86	1.98	(-0.10, 10.10)

μ_difference: mean of (Before - After)

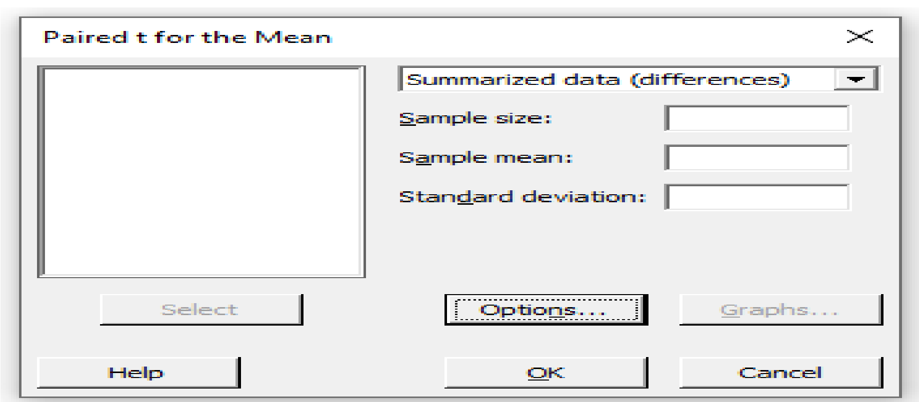
Test

Null hypothesis H_0: μ_difference = 0

Alternative hypothesis H_1: μ_difference ≠ 0

T-Value	P-Value
2.52	0.053

Notation



(b) Two-Sample t Test for Independent Samples (Equal Variances)

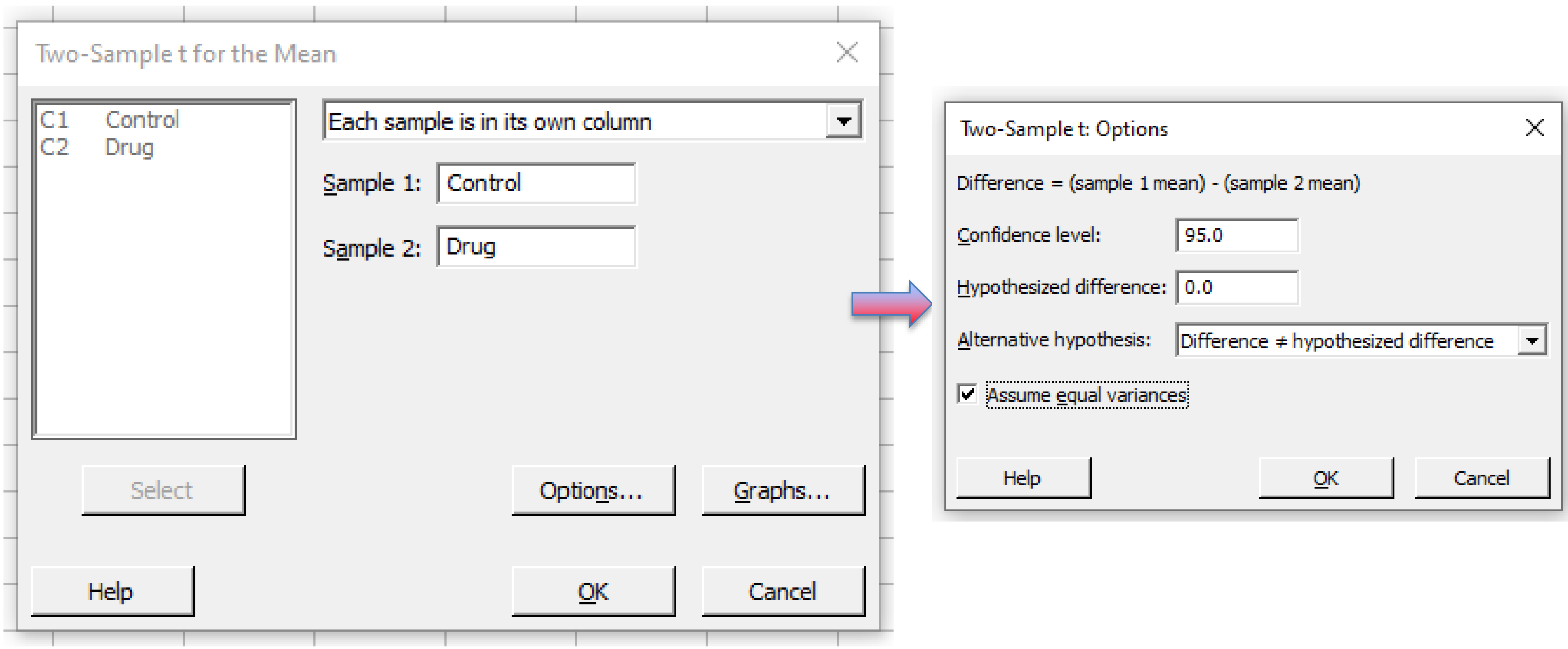
Example

To investigate the effect of a new hay fever drug on driving skills, a researcher studies 24 individuals with hay fever: 12 who have been taking the drug and 12 who have not. All participants then entered a simulator and were given a driving test that assigned a score to each driver as shown on the table:

Control	Drug
23	16
15	21
16	16
25	11
20	24
17	21
18	18
14	15
12	19
19	22
21	13
22	24

Steps

1. Open Minitab.
2. Choose Stat > Basic Statistics > 2- Sample t.
3. From the drop-down list, select Each samples in its own column.
4. In Sample1, enter Control.
5. In Sample2, enter Drug.
6. Click OK.



Two-Sample T-Test and CI: Control, Drug

Method

μ_1: mean of Control
μ_2: mean of Drug
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Control	12	18.50	3.90	1.1
Drug	12	18.33	4.23	1.2

Estimation for Difference

Difference	Pooled StDev	95% CI for Difference
0.17	4.07	(-3.28, 3.61)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
0.10	22	0.921

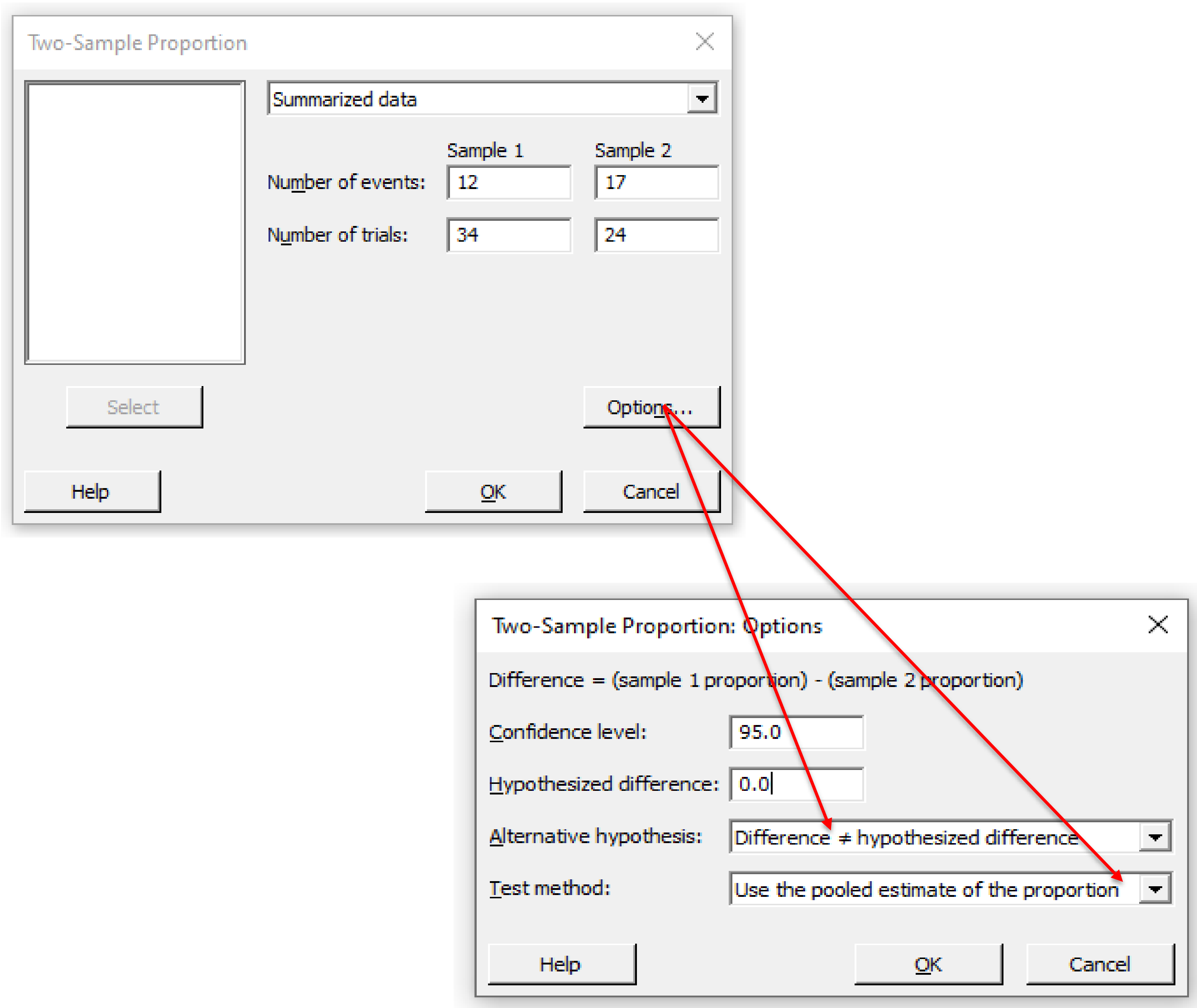
(c) Two-Sample Test for Binomial Proportions

Example

In the nursing home study, the researchers found that 12 out of 34 small nursing homes had a resident vaccination rate of less than 80%, while 17 out of 24 large nursing homes had a vaccination rate of less than 80%. At $\alpha = 0.05$, test the claim that there is no difference in the proportions of the small and large nursing homes with a resident vaccination rat of less than 80%?

Minitab Steps

1) Select Stat> Basic Statistics> 2 Proportions.
2) Click the button for Summarized data.
3) Enter required data.
4) Click Options.
5) Choose pooled estimate of p for test.
6) Click OK.


Two-Sample Proportion
Summarized data
Sample 1
Sample 2
Number of events:
12
17
Number of trials:
34
24
Select
Options...
Help
OK
Cancel
Two-Sample Proportion: Options
Difference = (sample 1 proportion) - (sample 2 proportion)
Confidence level:
95.0
Hypothesized difference:
0.0
Alternative hypothesis:
Difference ≠ hypothesized difference
Test method:
Use the pooled estimate of the proportion
Help
OK
Cancel

Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	34	12	0.352941
Sample 2	24	17	0.708333

Estimation for Difference

Difference	95% CI for Difference
-0.355392	(-0.598025, -0.112759)

CI based on normal approximation

Test

Null hypothesis	$H_0: p_1 - p_2 = 0$
Alternative hypothesis	$H_1: p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Normal approximation	-2.67	0.008
Fisher's exact		0.016

The pooled estimate of the proportion (0.5) is used for the tests.

Contingency-Table Method

(a) A 2 × 2 Contingency Table

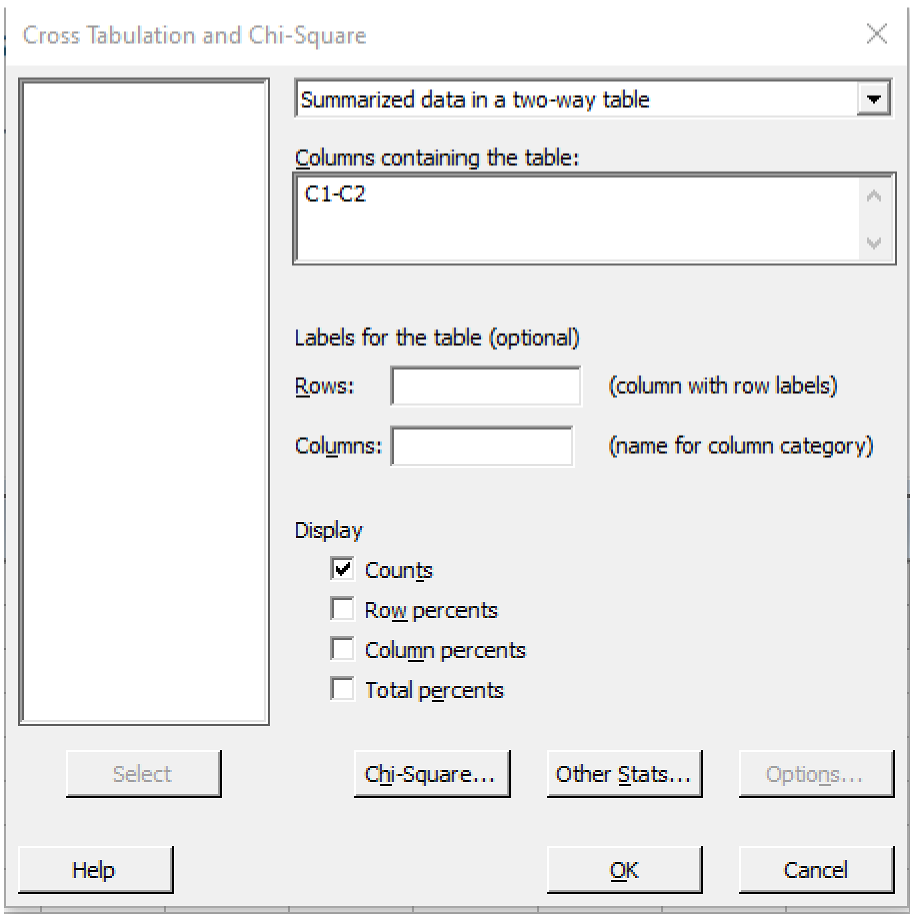
1. Click Stat → Tables → Cross Tabulation and Chi-Square.
2. A new window named "Cross Tabulation and Chi-Square" pops up.
3. Select "Results" as "For rows."
4. Select "Supplier" as "For columns."
5. Select "Count" as "Frequencies."
6. Click the "Chi-Square" button.

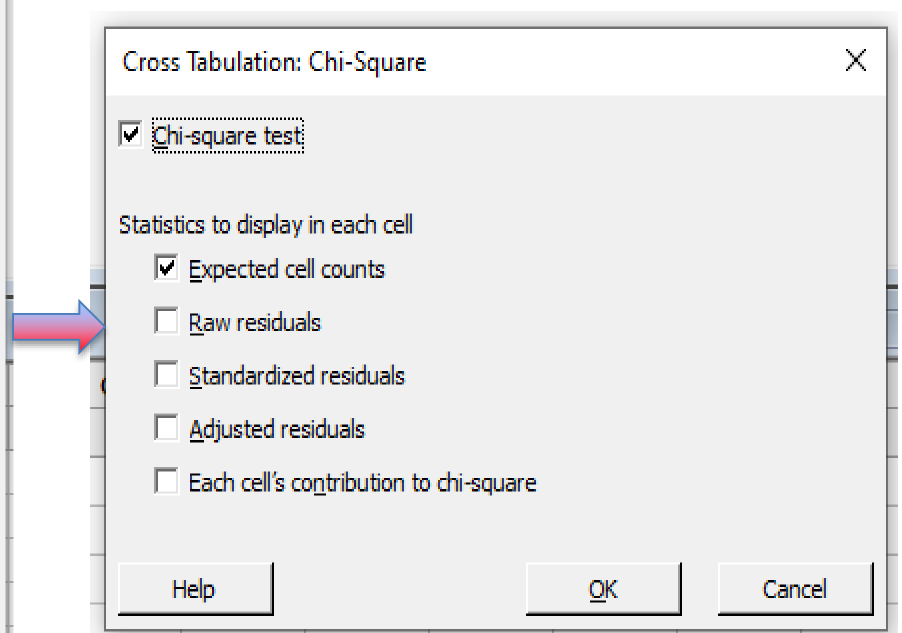
Example

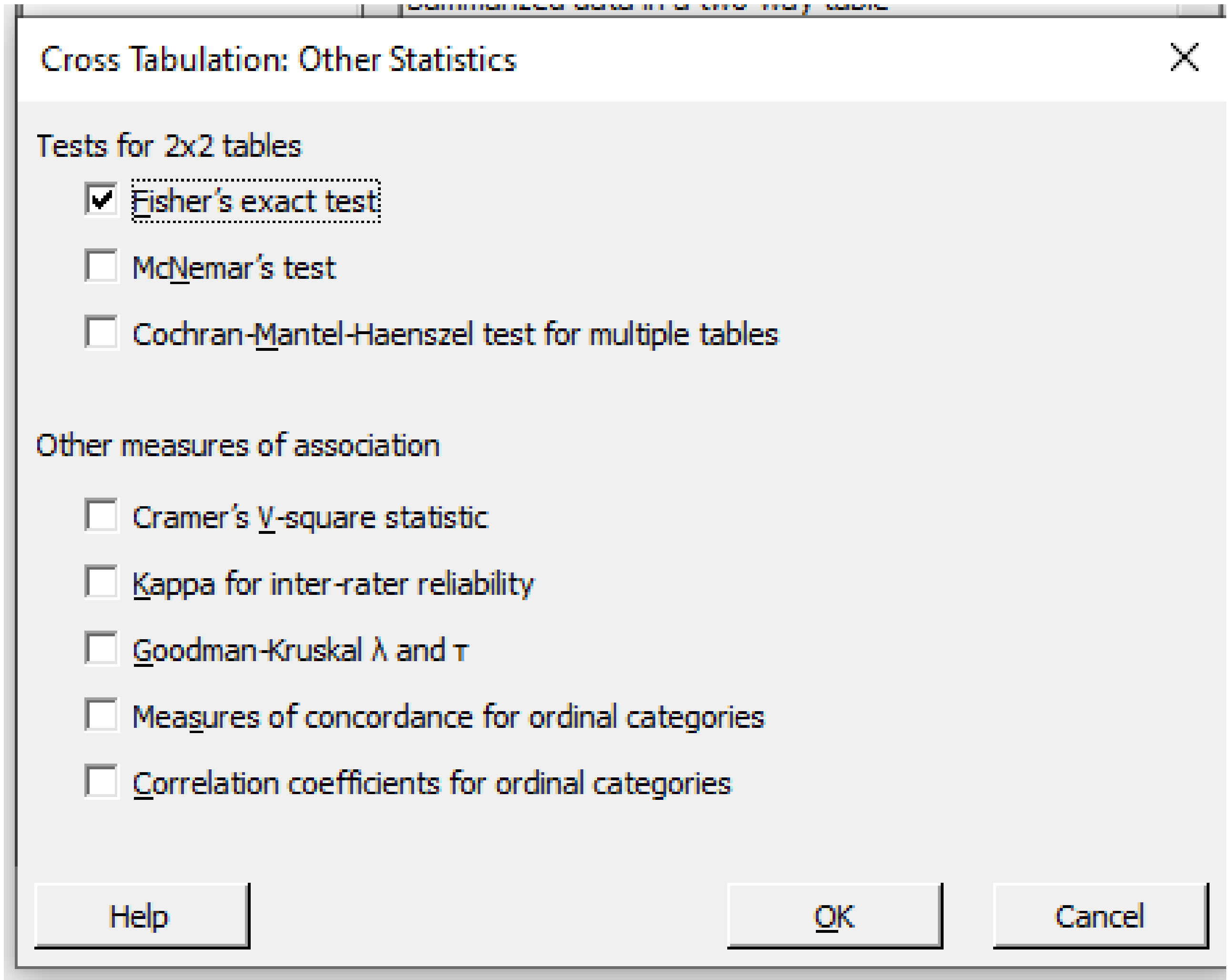
A sample of 50 randomly selected men with high triglyceride levels consumed 2 tablespoons of oat bran daily for six weeks. After six weeks, 60% of the men had lowered their triglyceride level. A sample of 80 men consumed 2 tablespoons of wheat bran for six weeks. After six weeks, 25% had lower triglyceride levels. By using a 2 x 2 contingency-table approach can we conclude that there is a significance difference in the two proportions at $\alpha = 0.01$?

Observed Table

Triglyceride level	Type of consumed food for six weeks		Total
	Oat bran	Wheat bran	
Lowered	30	20	50
Non-Lowered	20	60	80
Total	50	80	130






Cross Tabulation: Other Statistics
Tests for 2x2 tables
Fisher's exact test
McNemar's test
Cochran-Mantel-Haenszel test for multiple tables
Other measures of association
Cramer's V-square statistic
Kappa for inter-rater reliability
Goodman-Kruskal λ and τ
Measures of concordance for ordinal categories
Correlation coefficients for ordinal categories
Help
OK
Cancel

Tabulated Statistics: Worksheet rows, Worksheet columns

Rows: Worksheet rows   Columns: Worksheet columns

	C1	C2	All
1	30	20	50
	19.23	30.77	
2	20	60	80
	30.77	49.23	
All	50	80	130

Cell Contents
Count
Expected count

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	15.925	1	0.000
Likelihood Ratio	15.958	1	0.000

Fisher's Exact Test

P-Value
0.0000914

Example

A researcher wishes to determine whether there is a relationship between the gender (sex) of an individual and the amount of headache medications consumed. A sample of 69 people is selected, and the data in the following contingency table are obtained:

Contingency Table

Gender	Headache Consumption			Total
	Low	Moderate	High	
Male	10	9	8	27
Female	13	16	12	41
Total	23	25	20	68

At $\alpha = 0.10$, can the researcher conclude headache consumption is related to gender?

10	9	8	
13	16	12	

Chi-Square Test for Association: Worksheet rows, Worksheet columns

Rows: Worksheet rows Columns: Worksheet columns

	C1	C2	C3	All
1	10	9	8	27
	9.132	9.926	7.941	
2	13	16	12	41
	13.868	15.074	12.059	
All	23	25	20	68

Cell Contents
Count
Expected count

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	0.281	2	0.869
Likelihood Ratio	0.281	2	0.869