

STAT 106: Biostatistics (Exercises)

Supervised by
Dr.Hanan Ali
Dr.Rae'esa

Prepared by
Amal Al-Mohisen
Kholoud Al-Sultan

Chapter

One & Two

Organizing and Displaying Data
&
Basic Summary Statistics

Answer the following questions:

[1] The variable is a

- (a) Subset of the population.
- (b) Parameter of the population.
- (c) Relative frequency.
- (d) Characteristic of the population to be measured.
- (e) Class interval.

[2] Which of the following is an example of a discrete variable?

- (a) The number of students taking statistics in this term at KSU.
- (b) The time to exercise daily.
- (c) Whether or not someone has a disease.
- (d) Height of certain buildings.
- (e) Level of education.

[3] Which of the following is not an example of a discrete variable?

- (a) The number of students at the class of statistics.
- (b) The number of times a child cry in a certain day.
- (c) The time to run a certain distance.
- (d) The number of buildings in a certain street.
- (e) Number of educated persons in a family.

[4] Which of the following is an example of a qualitative variable?

- (a) The blood pressure level.
- (b) The number of times a child brush his/her teeth.
- (c) Whether or not someone fail in an exam.
- (d) Weight of babies at birth.
- (e) The time to run a certain distance.

[5] Which of the following is an example of a parameter?

- (a) The mean age of every one living in Saudi Arabia.
- (b) The median of glucose level for all diabetic patients.
- (c) The variance of the birth weights of all babies born in Riyadh.
- (d) All of (a) to (c).
- (e) None of (a) to (c).

[6] Which of the following is an example of a statistic?

- (a) The mean age of all people in Saudi Arabia.
- (b) The number of children in Saudi Arabia.
- (c) The variance for the birth weights of all babies born in Riyadh.
- (d) The median glucose level for a sample of diabetic patients.
- (e) The type of disease a person has.

[7] The continuous variable is a

- (a) Variable with a specific number of values.
- (b) Variable which can't be measured.
- (c) Variable takes on values within intervals.
- (d) Variable with no mode.
- (e) Qualitative variable.

[8] The width of the interval equals

- (a) The number of intervals.
- (b) The distance from the lower limit of an interval to the lower limit of the next one.
- (c) The distance from the lower limit of an interval to the upper limit of it.
- (d) The difference between the frequency of an interval and the frequency of the next one.
- (e) The relative frequency.

[9] Which of the following is an example of a continuous variable?

- (a) The number of visitors of the clinic yesterday.
- (b) The time to finish the exam.
- (c) Whether or not the answer is true.
- (d) The number of patients suffering from a certain disease.
- (e) Level of education.

[10] The y-axis in the graph of the histogram may represent the

- (a) Frequency.
- (b) Relative frequency.
- (c) Either (a) or (b).
- (d) Neither (a) nor (b).
- (e) We can't say.

[11] The discrete variable is a

- (a) Qualitative variable.
- (b) Variable with no mode.
- (c) Variable takes on values within intervals.
- (d) Variable with a specific number of values.
- (e) Variable which cannot be measured.

[12] Which of the following is an example of a nominal variable?

- (a) Age of visitors of the clinic.
- (b) The time to finish the exam.
- (c) Whether or not the person is infected by influenza.
- (d) Weight of a sample of girls.
- (e) The number of educated patients.

[13] The nominal variable is a

- (a) Variable with a specific number of values.
- (b) Qualitative variable which can't be ordered.
- (c) Variable takes on values within intervals.
- (d) Variable with no mode.
- (e) Quantitative variable.

[14] Which of the following is an example of an ordinal variable?

- (a) The number of persons who are injured in accidents.
- (b) The time to finish the exam.
- (c) Whether or not the medicine is effective.
- (d) Blood type of a sample of patients.
- (e) Socio-economic level.

[15] The ordinal variable is a

- (a) Variable with a specific number of values.
- (b) Variable that takes on values within intervals.
- (c) Qualitative variable which can be ordered.
- (d) Variable with several modes.
- (e) Quantitative variable.

[16] We calculate statistic instead of parameters because

- (a) Population are too large to measure the parameters.
- (b) Statistics are more popular than parameters.
- (c) Parameters are very large but statistics are small.
- (d) Statistics give more information than parameters.
- (e) We are only interested in values of statistics.

[17] Which of the following is an example of a statistic?

- (a) The population variance.
- (b) The sample median.
- (c) The population mean.
- (d) The population mode.
- (e) None of these.

106 Stat

SPSS Exercises

إعداد

أ.سناء أبونصره

أ. العنود الزعبي

أ. أمل المحيسن

أ.ربا اليافي



أ. وئام الهدلق

Introduction:

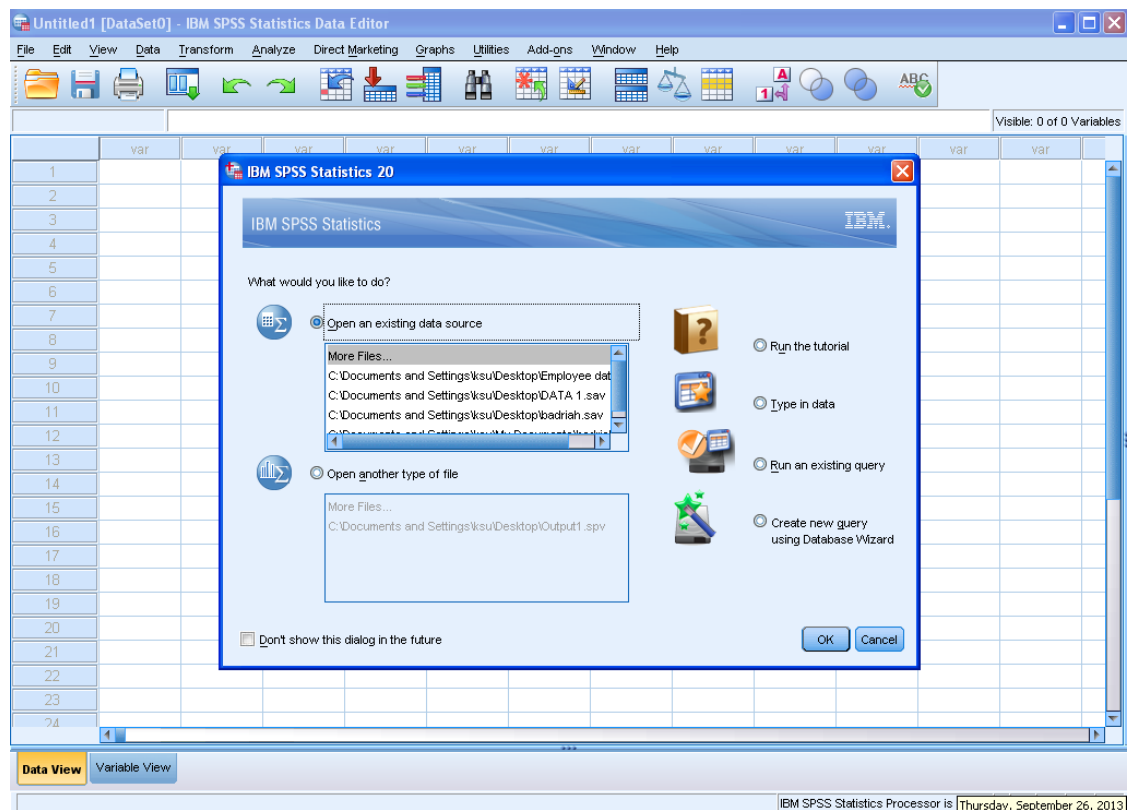
SPSS is a Windows based program that can be used to perform data entry and analysis and to create tables and graphs. In this course we introduce basic necessary to run SPSS program (version 20). It includes how to create a data file and run descriptive statistics.

Opening SPSS:

Depending on how the computer you are working on is structured, you can open SPSS in one of two ways.

1. If there is an **SPSS shortcut** like this  on the desktop, simply put the cursor on it and double click the left mouse button.
2. Click the left mouse button on the  button on your screen, then put your cursor on **All Programs** and left click the mouse. Select **SPSS Inc** by clicking the left mouse button.

The **dialog box** that appears offers choices of running the tutorial, typing in data, or opening an existing data source. The window behind this is the **Data Editor** window which is used to display the data from whatever file you are using. You could select any one of the options on the start-up dialog box and click **OK**, or you could simply hit **Cancel**. If you hit Cancel, you can either enter new data in the blank Data Editor or you could open an existing file using the **File** menu bar as explained later.

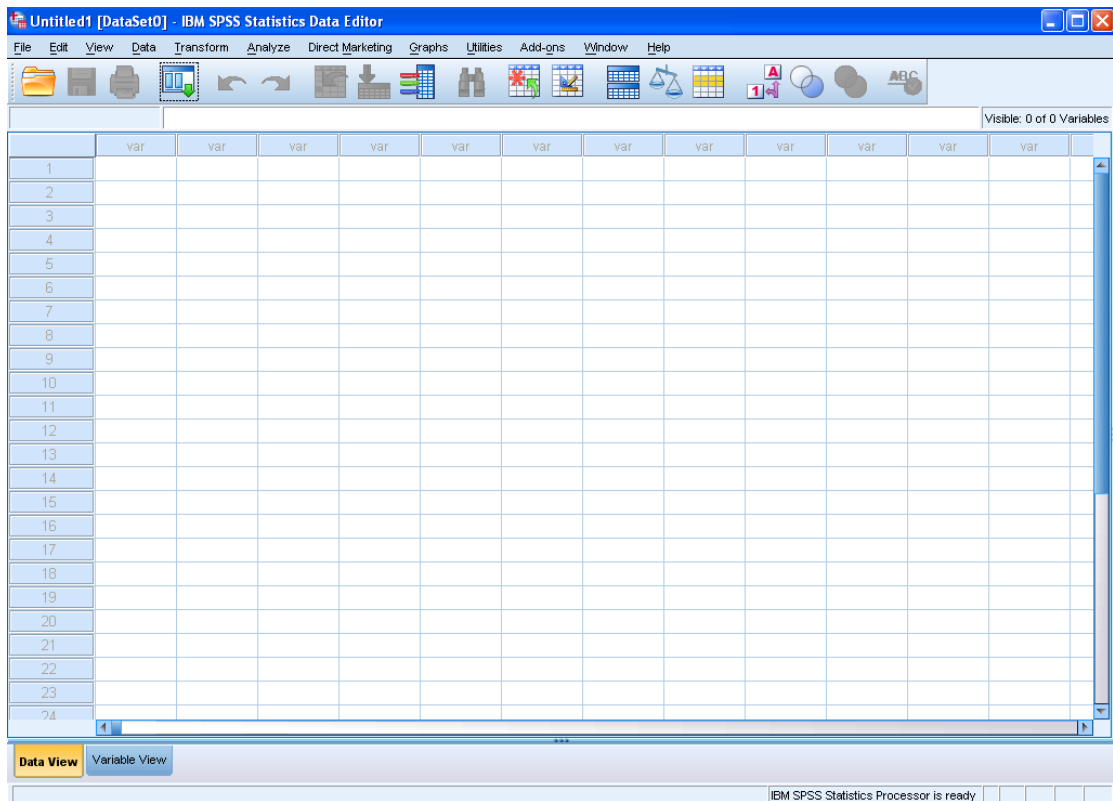


Click **Cancel**, and we'll get acquainted with the layout of SPSS.

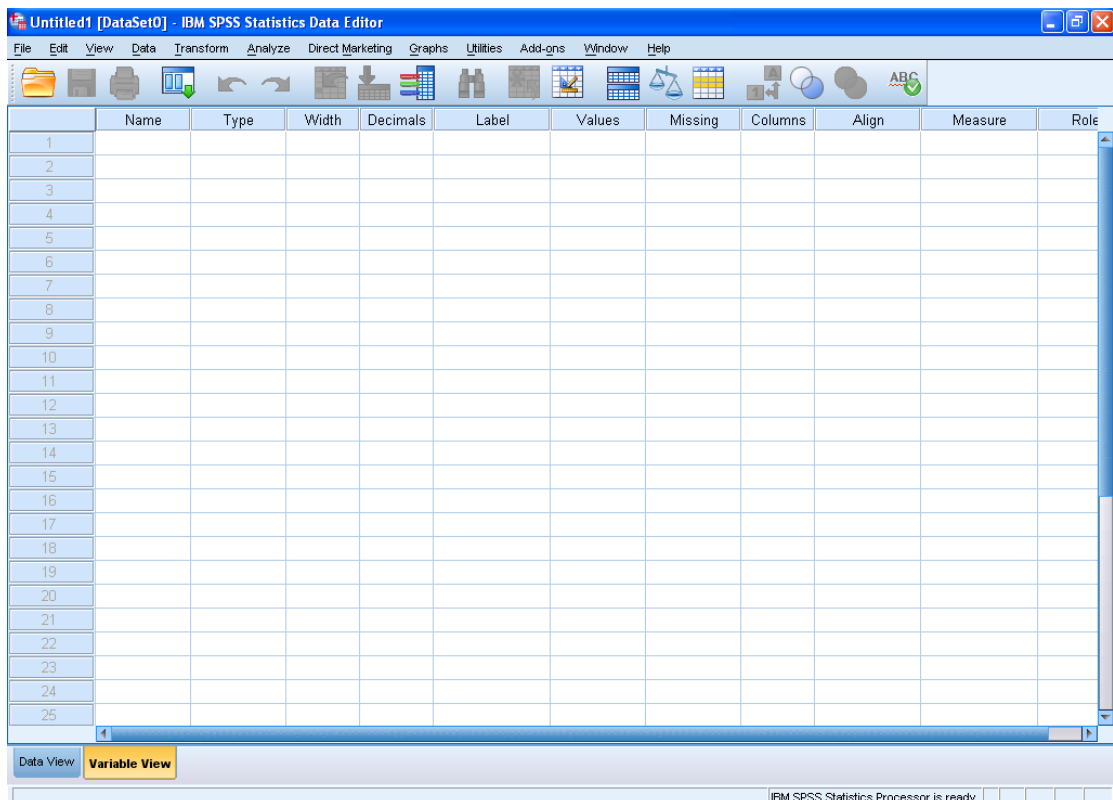
Layout of SPSS:

The **Data Editor** window has two views that can be selected from the lower left hand side of the screen.

Data View is where you see the data you are using, each white row represents a case, and each column represents a variable.



Variable View is where you can specify the format of your data when you are creating a file or where you can check the format of a pre-existing file. The data in the Data Editor is saved in a file with the extension **.save**.



Name

- Each variable name must be unique; duplication is not allowed.
- Start with a letter.
- Variable names cannot end with a period.
- No spaces.

Type

- The basic types are numeric and string.
- String variables may contain letters or numbers.
- For string values a blank is considered a valid value.
- Numeric operations on the string variables will not be allowed, e.g. finding the mean, variance, standard deviation, etc...

Width

- The number of characters SPSS will allow to be entered for the variable.
- For a numerical value with decimals, this total width has to include a spot for each decimal, as well as one for the decimal point.

Decimals

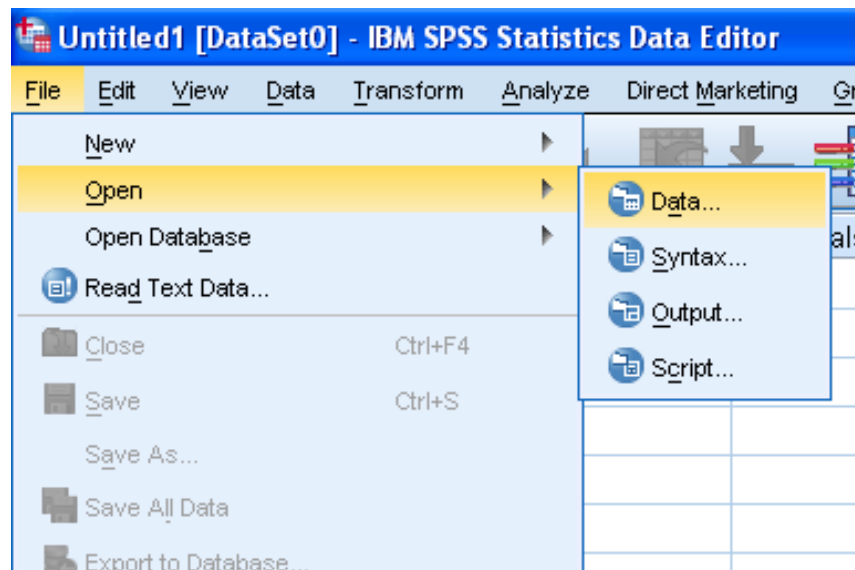
- If more decimals have been entered or computed by SPSS, the additional information will be retained internally but not displayed on screen.

The other most commonly used SPSS window is the **SPSS Viewer** window which displays the output from any analyses that have been run and any error messages. Information from the Output Viewer is saved in a file with the extension **.spv**.

SPSS Menus and Icons:

Now, let's review the most useful menus and icons.

File includes all of the options you typically use in other programs, such as open, save, exit. Notice, that you can open or create new files of multiple types as illustrated in the following.




Transform includes several options to change current variables. For example, you can change continuous variables to categorical variables, add a constant to variables, etc.

Analyze includes all of the commands to carry out statistical analyses and to calculate descriptive statistics.

Graphs include the commands to create various types of graphs including box plots, histograms, line graphs, and bar charts.

Window can be used to select which window you want to view (i.e., Data Editor, or Output Viewer).

Exiting SPSS:

To close SPSS, you can either left click on the close button  located on the upper right hand corner of the screen or select Exit from the **File** menu.

Using SPSS in Analyzing Quantitative discrete Data

For quantitative discrete variables one can obtain the frequency distribution table, represent the frequencies using bar chart, calculate different descriptive statistics.

Example 1:

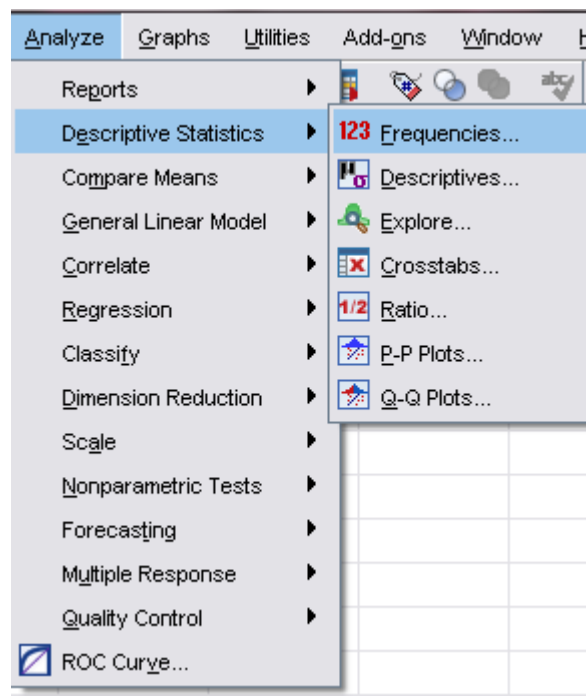
20 women were selected randomly and were asked about the number of children that they have. We get the following data:

5, 2, 4, 2, 3, 2, 4, 2, 3, 1, 2, 4, 3, 2, 3, 5, 2, 4, 1, 2.

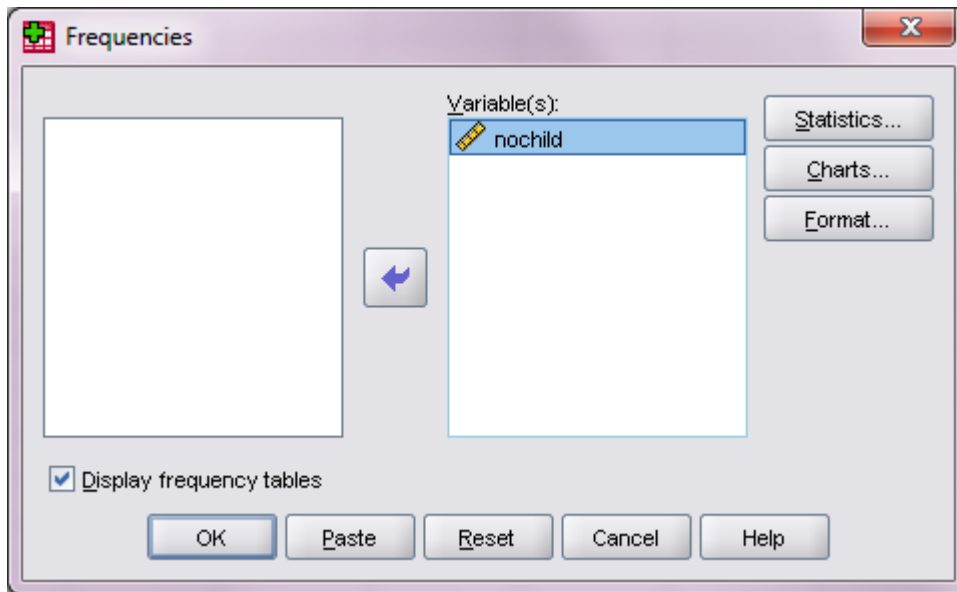
This data will be entered vertically and we will name the variable “Nochild”.

To obtain the frequency table we use the following steps:

- Analyze → Descriptive Statistics → Frequencies.



- The frequencies dialog box will appear, select the variable to be depicted in the frequency table by moving them from the left- to the right-hand box.



- Click on OK in the frequencies dialog box. In the SPSS Output Viewer, you will see two tables. The first shows the number of observations (which is 20 and that means that there are no missing observations)

Statistics

Nochild

| | | |
|---|---------|----|
| N | Valid | 20 |
| | Missing | 0 |

The second table is the frequency table and it lists the values of the variable “nochild” and the frequency of occurrence of each:

Nochild

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------|-----------|---------|---------------|--------------------|
| Valid | 1.00 | 2 | 10.0 | 10.0 | 10.0 |
| | 2.00 | 8 | 40.0 | 40.0 | 50.0 |
| | 3.00 | 4 | 20.0 | 20.0 | 70.0 |
| | 4.00 | 4 | 20.0 | 20.0 | 90.0 |
| | 5.00 | 2 | 10.0 | 10.0 | 100.0 |
| | Total | 20 | 100.0 | 100.0 | |

The output has five columns:

1. The first column represent the values the variable “nochild” takes, which takes the value from 1 to 5, this means that the women under the study have number of children between one and five.

2. The second column represent the frequencies, i.e. :
 - Two women have one child.
 - Eight women have two children.
 - Four women have three children.
 - Four women have four children.
 - Two women have five children.

3. The third column “percent” represent the percentages of the frequencies, i.e. :
 - 10 % women have one child.
 - 40 % women have two children.
 - 20 % women have three children.
 - 20 % women have four children.
 - 10 % women have five children.

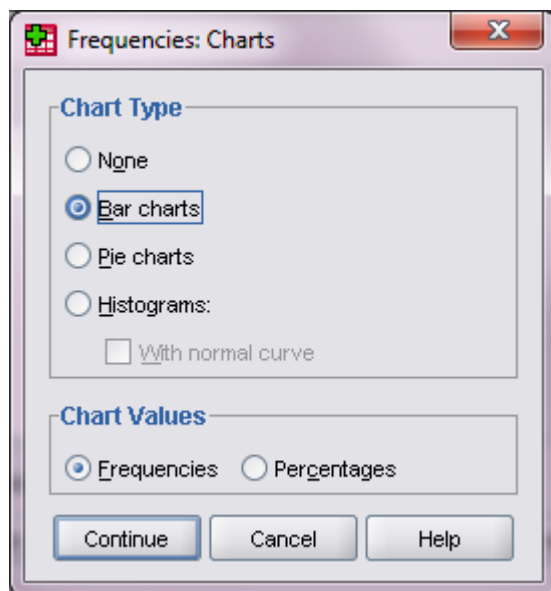
4. The fourth column is the “valid percent”. In our example this column is equal to the “percent” column since we have no missing observation. If we have a missing observation for example if a women under the study refuses to give any information, or we were unable to reach her and ext, in this cases the “valid percent” column will be different than the “percent” column.
5. Finally, the fifth column is the “cumulative percent”, it represent the cumulative frequency percentage by adding in succession the percentage frequencies.

Graphical analysis:

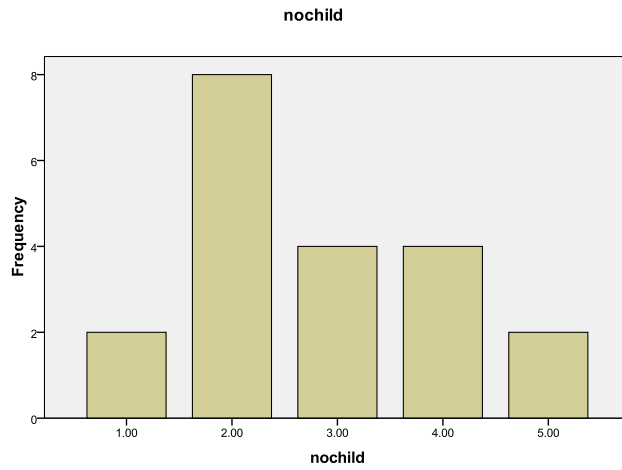
To create a graph, select Analyze → Descriptive Statistics → Frequencies→ charts.

This option creates bar charts, pie charts, or histograms in addition to the frequency table. Specify which chart you want to display by clicking on the Chart button.

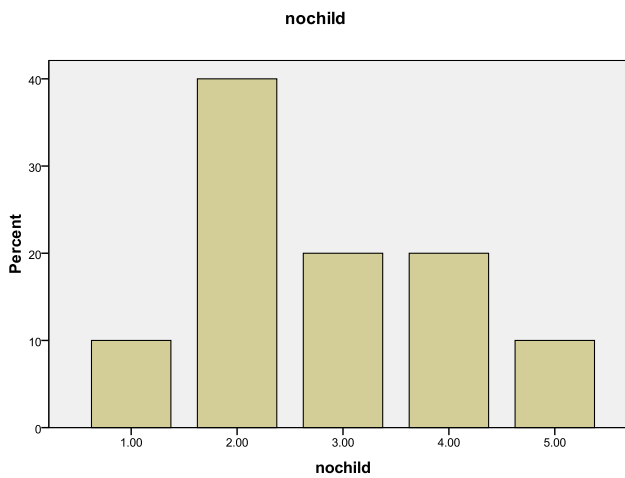
For discrete variables we select bar chat as follows:



- In the chart values dialog box, click on frequencies if we want the vertical axis to present the frequencies as shown in (1), click on percentages if we want the vertical axis to present the percentages as shown in (2). Then press continue→ ok.



(1)



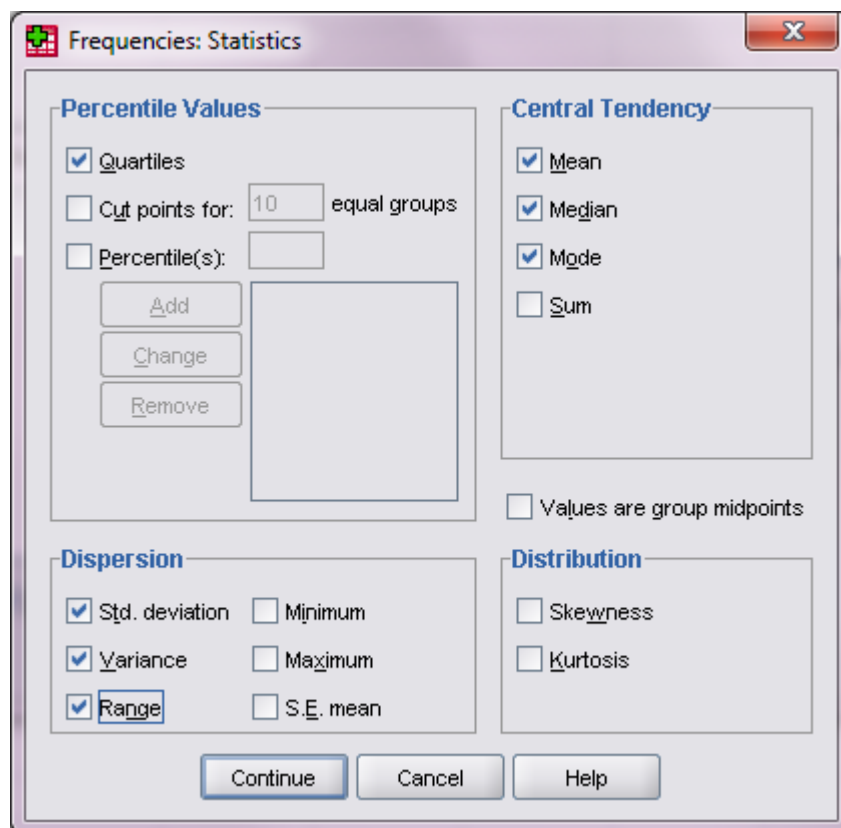
(2)

Descriptive Statistics:

- Convert large sets of data to more meaningful, easier to interpret, chunks or values.
- They summarize the data.
- Ex: mean, median, variance, and range.

SPSS contains a function that will compute many of these statistics.

- Analyze → Descriptive Statistics → Descriptive.



- From the statistics dialog box, click on the desired statistics that you want to perform (e.g. mean, median, mode, standard deviation, variance, range, etc.), click on the Continue button then ok.
- Central tendency (Mean, Median, Mode).
- Dispersion (Std. deviation, Variance, Range).
- Percentile values (Quartiles).

In the SPSS Output Viewer, you will see the requested statistics:

Statistics

Nochild

| | | |
|----------------|---------|---------|
| N | Valid | 20 |
| | Missing | 0 |
| Mean | | 2.8000 |
| Median | | 2.5000 |
| Mode | | 2.00 |
| Std. Deviation | | 1.19649 |
| Variance | | 1.432 |
| Range | | 4.00 |
| Percentiles | 25 | 2.0000 |
| | 50 | 2.5000 |
| | 75 | 4.0000 |

As we can see from the table:

The mean for the number of children is 2.8.

The median is 2.5.

The mode is 2.

The standard deviation is 1.1965.

The variance is 1.4316.

The range is 4.

The first quartile (The 25th percentile) is 2.

The second quartile (The 50th percentile also known as median) is 2.5.

The third quartile (The 75th percentile) is 4.

Continuous Variable

The following table gives the hemoglobin level (in g/dl) of a sample of 45 apparently healthy men aged 20-24. Find the grouped frequency distribution for the data.

| | | | | | | | | |
|------|------|------|------|------|------|------|------|------|
| 17 | 17.1 | 14.6 | 14 | 16.1 | 15.9 | 16.3 | 14.2 | 16.5 |
| 17.7 | 15.7 | 15.8 | 16.2 | 15.5 | 15.3 | 17.4 | 16.1 | 14.4 |
| 15.9 | 17.3 | 15.3 | 16.4 | 18.3 | 13.9 | 15 | 15.7 | 16.3 |
| 15.2 | 13.5 | 16.4 | 14.9 | 15.8 | 16.8 | 17.5 | 15.1 | 17.3 |
| 16.2 | 16.3 | 13.7 | 17.8 | 16.7 | 15.9 | 16.1 | 17.4 | 15.8 |

Using SPSS, enter this data and give name of the variable “Hemoglobin”,

To make a frequency table , first we have to use the intervals of the hemoglobin level as follows

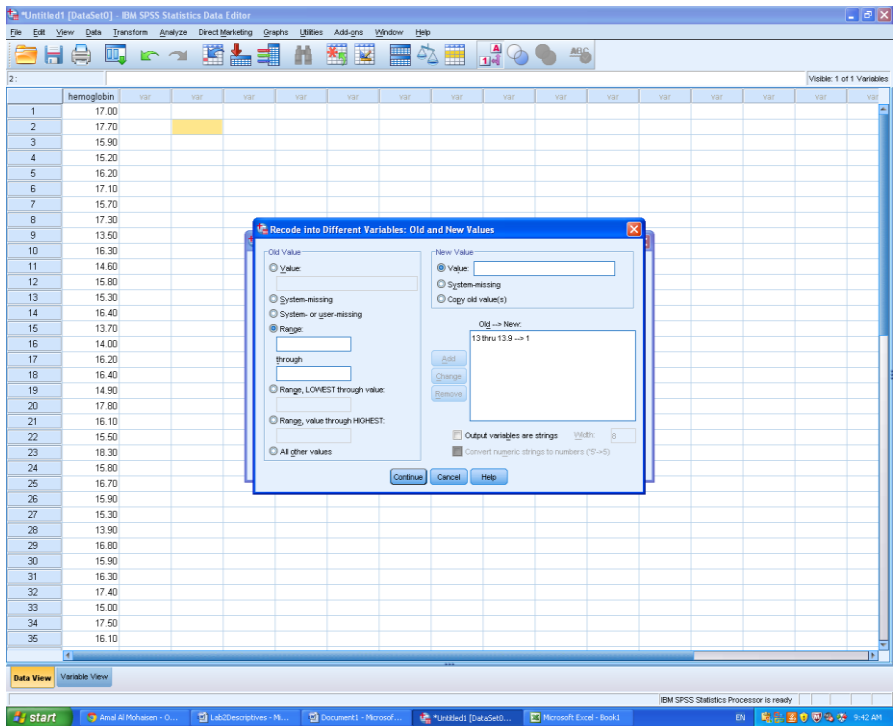
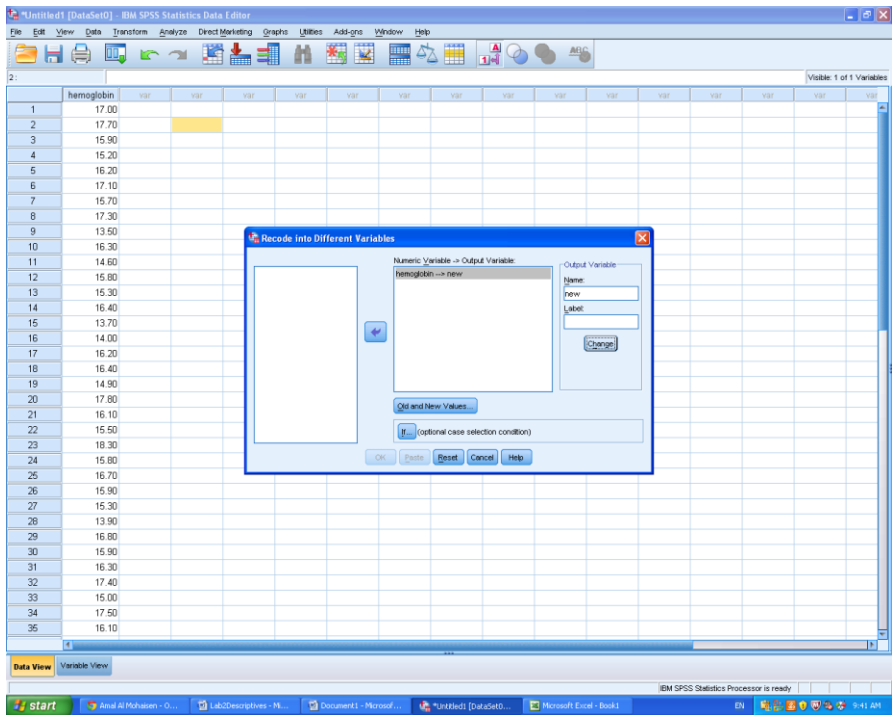
13 - 13.9 , 14 - 14.9 , 15 - 15.9 , 16 - 16.9 , 17 - 17.9 , 18 - 18.9

Then follow the steps

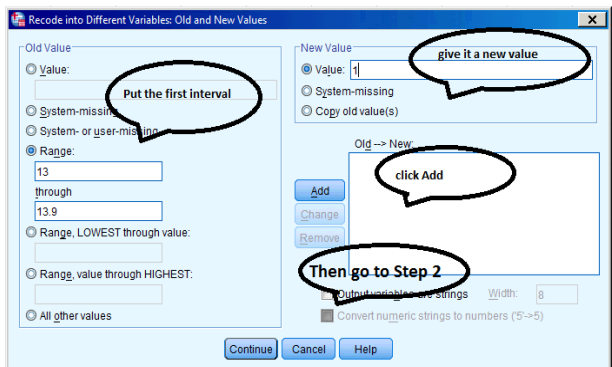
Recoding into different variable:

1. Select **Transform --> Recode into different variables**
2. Move the variable into the “Input Window”
3. Type a name for the new variable for example (“New”)
4. Click “Changes”
5. Click “Old and New Values”
6. Enter the “old” value and the “new value” and click “Add”
7. Click Continue
8. Click OK

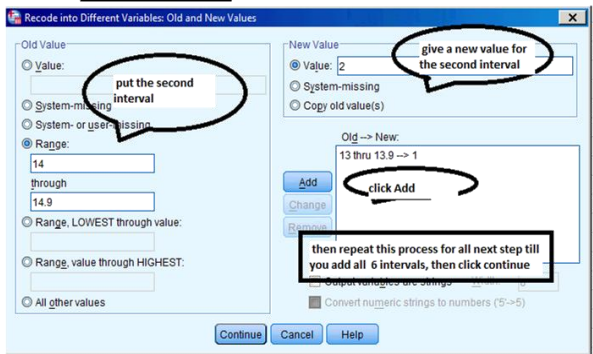
Then you will note that the new variable (“New”) have been added to the data



Step 1



Step 2



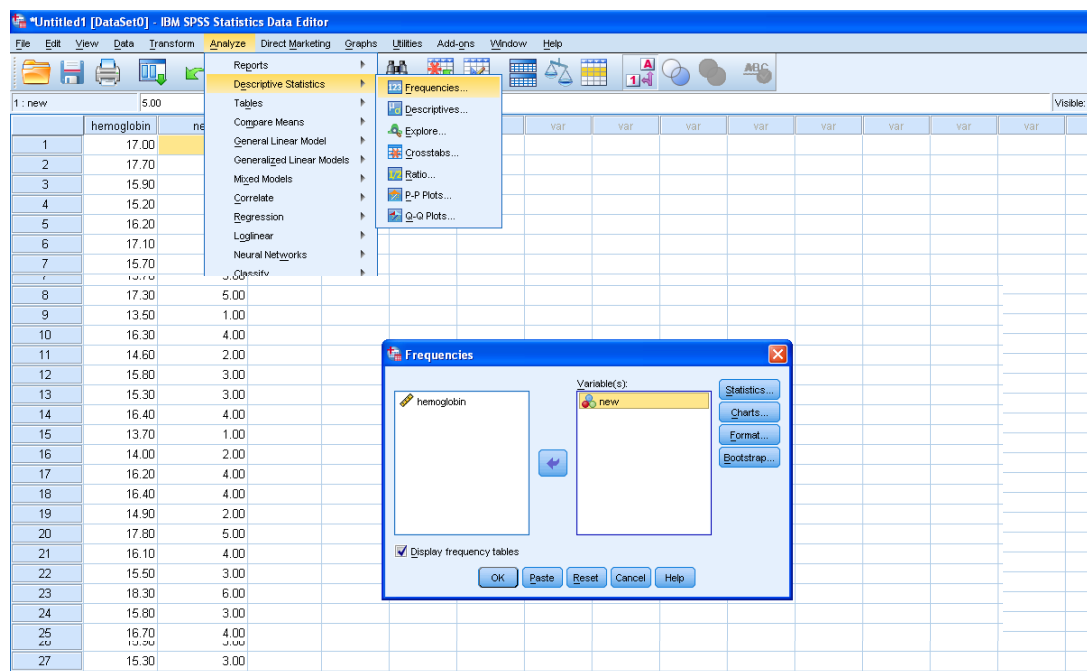
Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

| | hemoglobin | new | var | var | var | var | var |
|----|------------|------|-----|-----|-----|-----|-----|
| 1 | 17.00 | 5.00 | | | | | |
| 2 | 17.70 | 5.00 | | | | | |
| 3 | 15.90 | 3.00 | | | | | |
| 4 | 15.20 | 3.00 | | | | | |
| 5 | 16.20 | 4.00 | | | | | |
| 6 | 17.10 | 5.00 | | | | | |
| 7 | 15.70 | 3.00 | | | | | |
| 8 | 17.30 | 5.00 | | | | | |
| 9 | 13.50 | 1.00 | | | | | |
| 10 | 16.30 | 4.00 | | | | | |
| 11 | 14.60 | 2.00 | | | | | |
| 12 | 15.80 | 3.00 | | | | | |
| 13 | 15.30 | 3.00 | | | | | |
| 14 | 16.40 | 4.00 | | | | | |
| 15 | 13.70 | 1.00 | | | | | |
| 16 | 14.00 | 2.00 | | | | | |
| 17 | 16.20 | 4.00 | | | | | |
| 18 | 16.40 | 4.00 | | | | | |
| 19 | 14.90 | 2.00 | | | | | |
| 20 | 17.80 | 5.00 | | | | | |
| 21 | 16.10 | 4.00 | | | | | |

Frequency table and Chart:

To analyze the data we need to make the frequency table as follows

1. Select **Analyze --> Descriptive Statistics --> Frequencies**
2. Move **the recoded variable (New)** into the “Variable(s)” window.
3. Click “Charts” and put a checkmark next to the histogram.
4. Click OK.



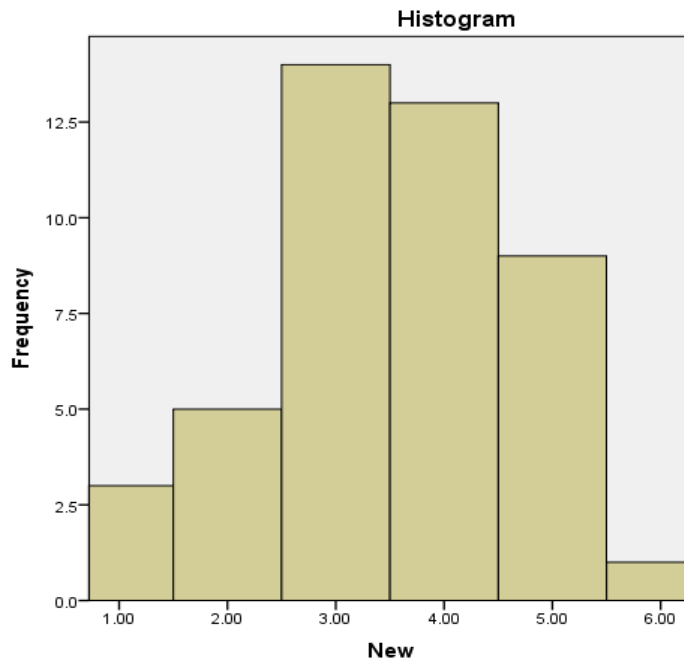
The output shows the Frequency table for the” **New**” variable as follows:

The variable name “New”

New

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------|-----------|---------|---------------|--------------------|
| Valid | 1.00 | 3 | 6.7 | 6.7 | 6.7 |
| | 2.00 | 5 | 11.1 | 11.1 | 17.8 |
| | 3.00 | 14 | 31.1 | 31.1 | 48.9 |
| | 4.00 | 13 | 28.9 | 28.9 | 77.8 |
| | 5.00 | 9 | 20.0 | 20.0 | 97.8 |
| | 6.00 | 1 | 2.2 | 2.2 | 100.0 |
| | Total | 45 | 100.0 | 100.0 | |

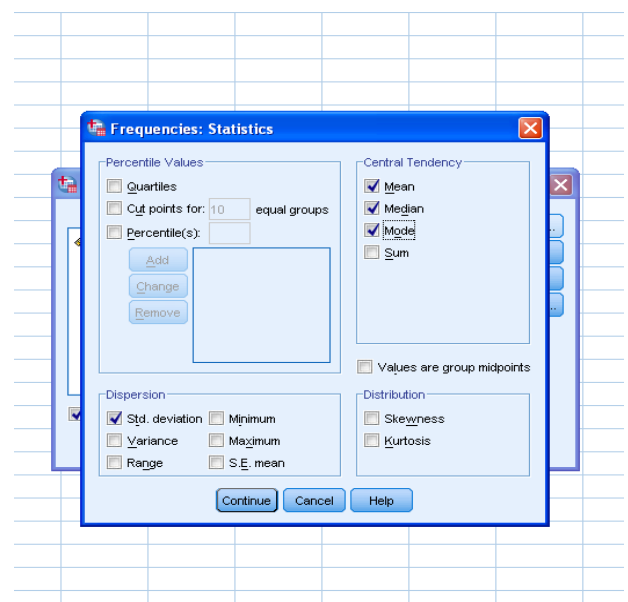
And the Histogram chart:



Statistics

The variable name "New"

1. Select **Analyze --> Descriptive Statistics**
2. Move **only the variable hemoglobin** (the original variable) into the "Variable(s)" window.
3. Click "Statistics" and put a checkmark next to mean, median, mode, variance and St.deviation.
4. Click OK.



The Statistics output:

Statistic
Hemoglobin

The variable name
"hemoglobin"

| | | |
|----------------|---------|--------------------|
| N | Valid | 45 |
| | Missing | 0 |
| Mean | | 15.9622 |
| Median | | 16.1000 |
| Mode | | 15.80 ^a |
| Std. Deviation | | 1.14937 |
| Variance | | 1.321 |

a. Multiple modes exist. The smallest value is shown

Now answer the following:

[1] The variable name is

[2] The number of patients with hemoglobin level between (16 – 16.9) is.....

[3] The number of patients with hemoglobin level less than 15 is.....

[4] The number patients with hemoglobin level more than or equal 17 is.....

[5] The relative frequency of patients with hemoglobin level from 13 to

14.9.....

[6] The percentage of patients with hemoglobin level from 15to 16.9 is.....

[7] The mode is.....

Review Questions

Q18) For a sample of women, here are the numbers of their children

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 5 | 2 | 1 | 4 | 3 | 5 | 4 | 0 | 1 | 2 | 6 | 4 | 0 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

Using SPSS, answer the following questions

1. The type of the variable is

| | | | | | |
|----------------|---------------|-----------|-----------|-------------|--------------|
| a) qualitative | b) continuous | c) number | d) normal | e) discrete | f) statistic |
|----------------|---------------|-----------|-----------|-------------|--------------|

2. The sample size

| | | | | | |
|------|-------|------|------|------|------|
| a) 5 | b) 15 | c) 3 | d) 2 | e) 4 | f) 0 |
|------|-------|------|------|------|------|

3. The number of women having 3 children

| | | | | | |
|------|------|------|-------|------|------|
| a) 2 | b) 3 | c) 5 | d) 15 | e) 1 | f) 6 |
|------|------|------|-------|------|------|

4. The percent of women having less than 4 children

| | | | | | |
|---------|--------|--------|---------|--------|------|
| a) 0.26 | b) 0.4 | c) 0.6 | d) 0.13 | e) 0.3 | f) 1 |
|---------|--------|--------|---------|--------|------|

5. The percentage of women having 5 or more children

| | | | | | |
|--------|--------|-------|-------|--------|---------|
| a) 99% | b) 40% | c) 3% | d) 0% | e) 20% | f) 100% |
|--------|--------|-------|-------|--------|---------|

6. The number of women having between 2 and 5 children

| | | | | | |
|------|------|------|-------|------|------|
| a) 2 | b) 3 | c) 5 | d) 15 | e) 1 | f) 6 |
|------|------|------|-------|------|------|

7. The sample mean is

| | | | | | |
|------|------|---------|--------|--------|------|
| a) 3 | b) 6 | c) 7.84 | d) 2.8 | e) 2.5 | f) 0 |
|------|------|---------|--------|--------|------|

8. The sample median is

| | | | | | |
|------|------|--------|--------|--------|--------|
| a) 3 | b) 2 | c) 3.5 | d) 7.5 | e) 2.5 | f) 1.5 |
|------|------|--------|--------|--------|--------|

9. The sample mode is

| | | | | | |
|------|--------|----------|------------|------------|------|
| a) 6 | b) 2,4 | c) 2,3,4 | d) 0,1,3,5 | e) no mode | f) 5 |
|------|--------|----------|------------|------------|------|

10. The sample standard deviation is

| | | | | | |
|--------|----------|----------|----------|----------|------|
| a) 166 | b) 3.457 | c) 1.364 | d) 1.796 | e) 1.859 | f) 1 |
|--------|----------|----------|----------|----------|------|

11. The coefficient of variation is

| | | | | | |
|-----------|--------|------------|------------|-----------|---------|
| a) 18.59% | b) 15% | c) 123.42% | d) 150.62% | e) 66.39% | f) 100% |
|-----------|--------|------------|------------|-----------|---------|

12. The first quartile is

| | | | | | |
|------|------|--------|------|------|------|
| a) 1 | b) 2 | c) 2.5 | d) 3 | e) 5 | f) 0 |
|------|------|--------|------|------|------|

13. The value that 25% of data before is

| | | | | | |
|------|------|--------|------|------|------|
| a) 1 | b) 2 | c) 2.5 | d) 3 | e) 5 | f) 0 |
|------|------|--------|------|------|------|

Q19) Suppose we measure the number of times a child has been to a doctor in the past month .For a sample of Saudi children age 2-6 years

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 3 | 4 | 0 | 5 | 6 | 2 | 2 | 5 | 4 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|

Using SPSS , answer the following questions

1. The type of the variable is

| | | | | | |
|---------------|--------------|----------|----------|------------|-------------|
| a)qualitative | b)continuous | c)number | d)normal | e)discrete | f)statistic |
|---------------|--------------|----------|----------|------------|-------------|

2. The sample size

| | | | | | |
|-----|------|-----|-----|-----|-----|
| a)2 | b)12 | c)6 | d)0 | e)1 | f)3 |
|-----|------|-----|-----|-----|-----|

3. The number of children who has been to a doctor less than 3 times in the past month is

| | | | | | |
|-----|-----|-----|-----|------|-----|
| a)1 | b)2 | c)6 | d)7 | e)12 | f)5 |
|-----|-----|-----|-----|------|-----|

4. The percent of children who did not go to a doctor in the past month is

| | | | | | |
|--------|-----|-------|-------|-----|--------|
| a)0.08 | b)0 | c)0.2 | d)0.6 | e)1 | f)0.12 |
|--------|-----|-------|-------|-----|--------|

5. The sample mean is

| | | | | | |
|---------|---------|---------|-----|-----|-------|
| a)2.917 | b)1.881 | c)3.538 | d)3 | e)2 | f)2.5 |
|---------|---------|---------|-----|-----|-------|

6. The sample median is

| | | | | | |
|---------|---------|---------|-----|-----|-------|
| a)2.917 | b)1.881 | c)3.538 | d)3 | e)2 | f)2.5 |
|---------|---------|---------|-----|-----|-------|

7. The sample mode is

| | | | | | |
|---------|---------|---------|-----|-----|-------|
| a)2.917 | b)1.881 | c)3.538 | d)3 | e)2 | f)2.5 |
|---------|---------|---------|-----|-----|-------|

8. The sample standard deviation is

| | | | | | |
|---------|---------|---------|-----|-----|-------|
| a)2.917 | b)1.881 | c)3.538 | d)3 | e)2 | f)2.5 |
|---------|---------|---------|-----|-----|-------|

9. The coefficient of variation is

| | | | | | |
|------|----------|----------|----------|--------|----------|
| a)0% | b)64.49% | d)15.62% | e)66.39% | f)100% | d)10.62% |
|------|----------|----------|----------|--------|----------|

10. The third quartile is

| | | | | | |
|-----|-------|-------|-----|--------|-----|
| a)2 | b)2.5 | c)3.5 | d)4 | e)4.75 | f)5 |
|-----|-------|-------|-----|--------|-----|

11. The value that 50% of data before is

| | | | | | |
|-----|-------|-------|-----|--------|-----|
| a)2 | b)2.5 | c)3.5 | d)4 | e)4.75 | f)5 |
|-----|-------|-------|-----|--------|-----|

Q20) Here the weights in kg of 50 college students

| | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 42 | 62 | 46 | 54 | 41 | 37 | 54 | 44 | 32 | 45 | 47 | 50 | 58 | 49 | 51 | 42 | 46 |
| 37 | 42 | 39 | 54 | 39 | 51 | 58 | 47 | 64 | 43 | 48 | 49 | 48 | 49 | 61 | 41 | 40 |
| 58 | 49 | 59 | 57 | 57 | 34 | 56 | 38 | 45 | 52 | 46 | 40 | 63 | 41 | 51 | 41 | |

Use these class intervals to obtain frequency table and answer the following questions

31-<35 36-<40 41-<45 46-<50 51-<55 56-<60 61-<65

1. The type of the variable is

| | | | | | |
|---------------|--------------|----------|----------|------------|-------------|
| a)qualitative | b)continuous | c)number | d)normal | e)discrete | f)statistic |
|---------------|--------------|----------|----------|------------|-------------|

2. The sample size

| | | | | | |
|-----|------|-----|------|-----|-----|
| a)5 | b)15 | c)3 | d)50 | e)4 | f)0 |
|-----|------|-----|------|-----|-----|

3. The number of students have weight between 36 and 45 kg is

| | | | | | |
|-----|------|------|------|------|------|
| a)6 | b)12 | c)18 | d)20 | e)50 | f)16 |
|-----|------|------|------|------|------|

4. How many students have weight 51 kg or more

| | | | | | |
|-----|------|------|------|------|------|
| a)6 | b)12 | c)18 | d)20 | e)50 | f)16 |
|-----|------|------|------|------|------|

5. Which weight class has the lowest percentage of students

| | | | | | |
|-----------|-----------|-----------|-----------|-----------|-----------|
| a) 31-<35 | b) 36-<40 | c) 41-<45 | d) 46-<50 | e) 51-<55 | f) 56-<60 |
|-----------|-----------|-----------|-----------|-----------|-----------|

6. The probability that a student is of weight less than 40 kg is

| | | | | | |
|--------|-----|-------|--------|--------|-----|
| a)0.04 | b)1 | c)0.4 | d)0.16 | e)0.18 | f)0 |
|--------|-----|-------|--------|--------|-----|

7. The sample mean is

| | | | | | |
|---------|---------|--------|-----|---------|------|
| a)47.50 | b)47.94 | c)7.97 | d)0 | e)63.53 | f)50 |
|---------|---------|--------|-----|---------|------|

8. The sample median is

| | | | | | |
|---------|---------|--------|-----|---------|------|
| a)47.50 | b)47.94 | c)7.97 | d)0 | e)63.53 | f)50 |
|---------|---------|--------|-----|---------|------|

9. The sample mode is

| | | | | | |
|------|------|-----------|------|---------|------|
| a)45 | b)50 | c)no mode | d)30 | e)41,49 | f)65 |
|------|------|-----------|------|---------|------|

10. The sample standard deviation is

| | | | | | |
|---------|---------|--------|-----|---------|------|
| a)47.50 | b)47.94 | c)7.97 | d)0 | e)63.53 | f)50 |
|---------|---------|--------|-----|---------|------|

11. The coefficient of variation is

| | | | | | |
|----------|----------|---------|-----|----------|-------|
| a)47.50% | b)16.62% | c)7.97% | d)0 | e)63.53% | f)50% |
|----------|----------|---------|-----|----------|-------|

Chapter 1 & Chapter 2

الأسئلة النظرية

Q21: For a sample of children with frequent toothache, we measure the number of times a child has been to a doctor in the last year as follows:

| | | | | | |
|--------------------|----|----|---|---|---|
| Number of times | 1 | 2 | 3 | 4 | 5 |
| number of children | 15 | 10 | 8 | 4 | 3 |

[1] The name of the variable.....

[2] The type of the variable

[3] The sample size.....

[4] The number of children have been to a doctor 3 times is.....

[5] The number of children have been to a doctor from 2 to 4 times is.....

[6] The relative frequency of children have been to a doctor 2 times is.....

[7] The percentage of children has been to a doctor less than 3 times is

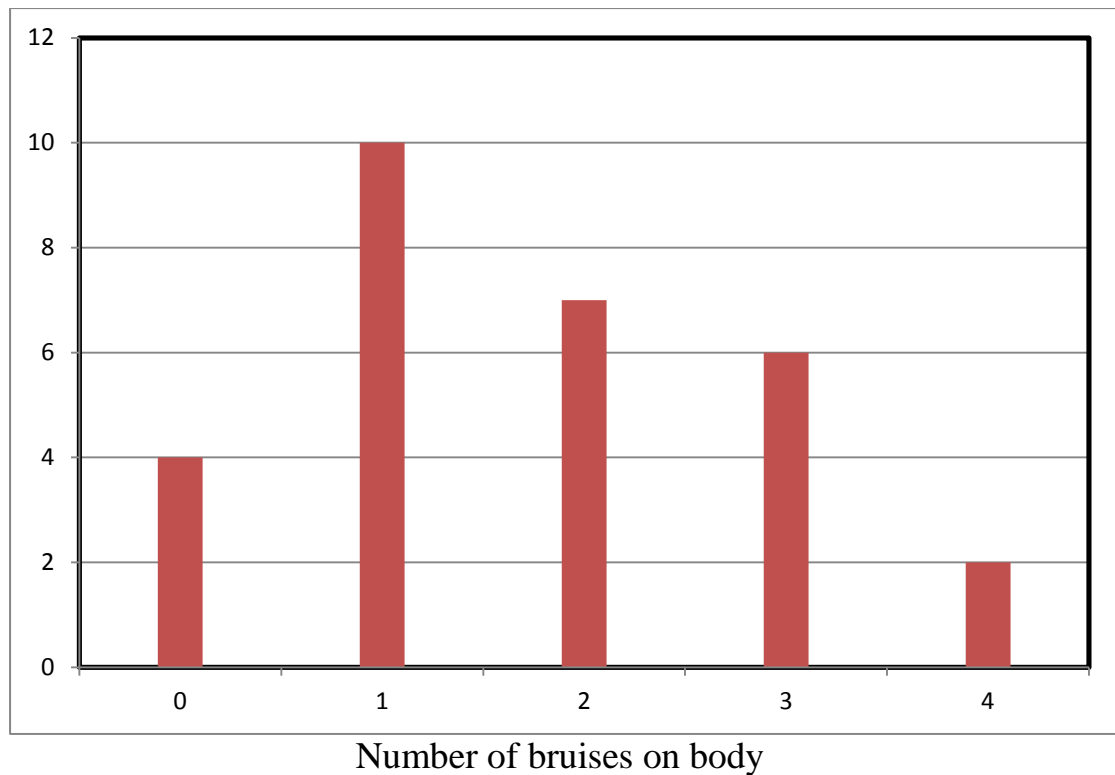
.....

[8] The largest value of the variable is

[9] The number of times a child has been to a doctor with the highest Percentage is

.....

Q22: For a sample of children age 5, we obtain the following graph



[1] The name of the variable is.....

[2] The name of the graph is.....

[3] The type of the variable

[4] The sample size.....

[5] The number of children had 3 or more bruises is

[6] The percentage of children has at most 2 bruises is.....

[7] The smallest value of the variable is

[8] The number of bruises has the largest percentage.....

[9] The mode is

Q23: For a sample of 60 women having children, we measure the age(in year)

| Age | Frequency | Relative frequency | cumulative frequency |
|---------|-----------|--------------------|----------------------|
| 15 - 20 | 3 | 0.05 | 3 |
| 21 - 26 | 12 | ---- | 15 |
| 27 - 32 | 18 | 0.3 | 33 |
| 33 - 38 | ---- | 0.283 | 50 |
| 39 - 44 | 10 | 0.166 | 60 |
| Total | 60 | | |

Complete the table, then answer

[1] The variable is

[2] The number of women with age between (33 – 38) is

[3] The number of women with age less than or equal 26 is

[4] The number of women with age more than 32 is

[5] The relative frequency of women with age from 21 to 38 is

[6] The percentage of women with age from 33 to 44 is

[7] The age class had the largest percentage is

[8] The width of the interval is

Q24: for the sample of 200 mothers at delivery ,we measure a certain serum level (in Mg/I).

| Time with doctor | Frequency | Relative frequency | cumulative frequency |
|------------------|-----------|--------------------|----------------------|
| 5 - 24 | 50 | 0.25 | 50 |
| 25 - 44 | 63 | 0.315 | 113 |
| 45 - 64 | ---- | 0.305 | 174 |
| 65 - 84 | 24 | ----- | 198 |
| 85 – 104 | 2 | 0.01 | 200 |
| Total | 200 | | |

Complete the table, then answer

[1] The variable is

[2] The number of mothers with time with doctor between (25-84) is

.....

[3] The number of mothers with time with doctor less than or equal 64 is

.....

[4] The number of mothers with time with doctor more than 44 is

.....

[5] The relative frequency of mothers with time with doctor from 65 to 84 is

.....

[6] The percentage of mothers with time with doctor more than or equal 45 is

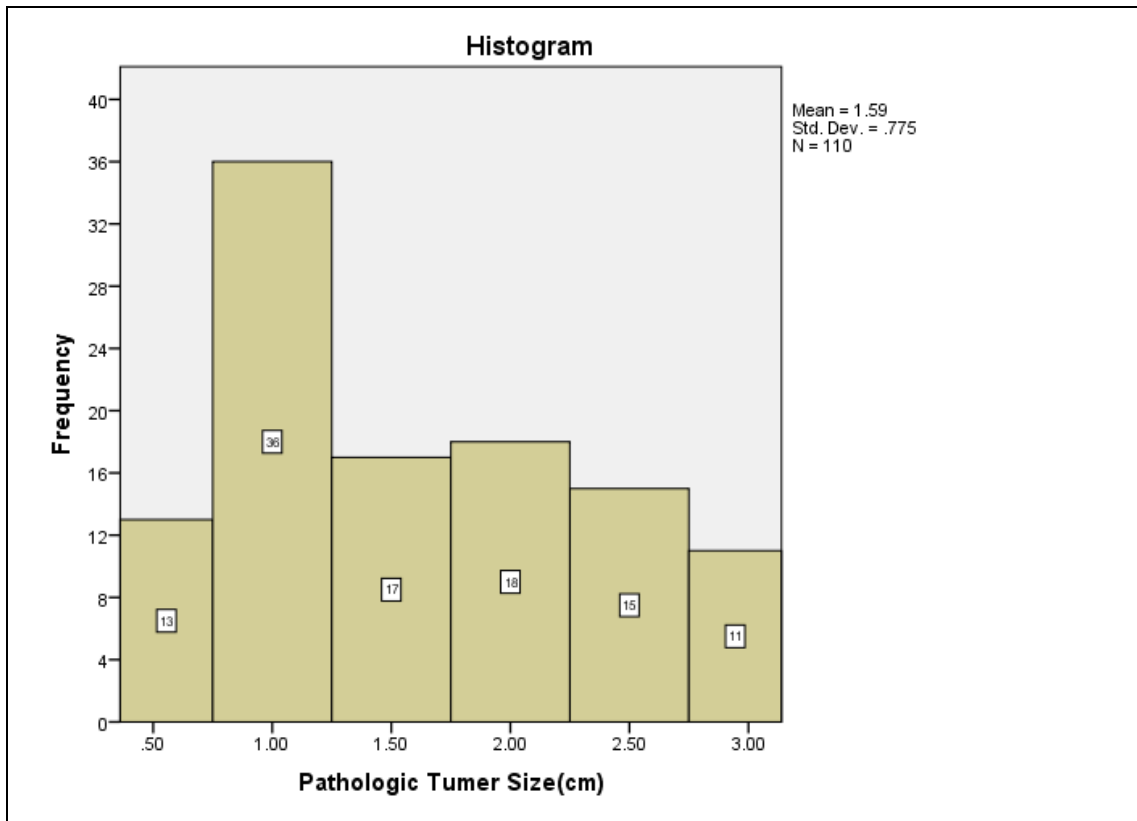
.....

[7] The mode is

[8] The sample size is

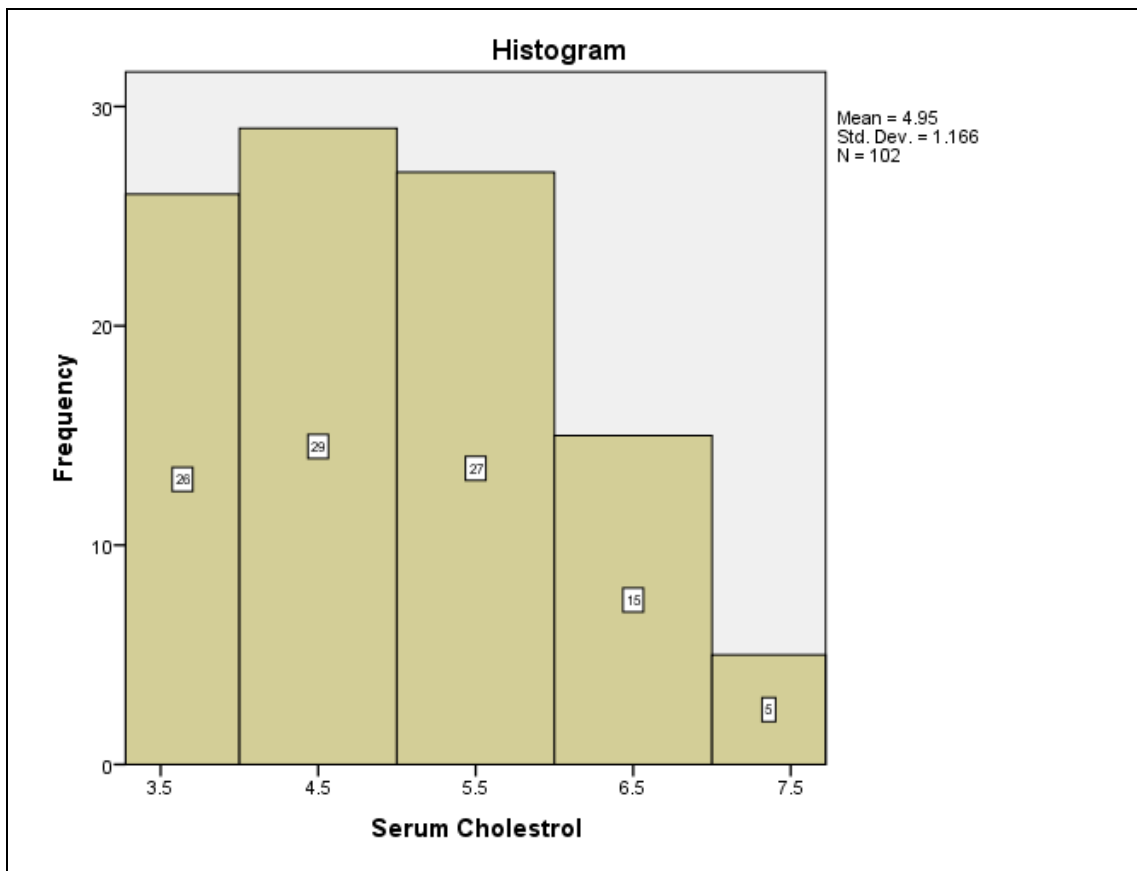
[9] The width of the interval is

Q25: The following graph shows the frequency distribution of pathologic tumor size (In cm) for a sample of cancer patient



- 1) The name of the graph is
 - (a) Bar chart
 - (b) histogram
 - (c) polygon
 - (d) curve
- 2) The sample size is
 - (a) 102
 - (b) 40
 - (c) 29
 - (d) 110
- 3) The number of cancer patient with the lowest pathologic tumor size is
 - (a) 3
 - (b) 11
 - (c) 13
 - (d) 15
- 4) The percent of cancer patient with approximate level of pathologic tumor size = 2 cm is
 - (a) 18%
 - (b) 16.36%
 - (c) 32.72%
 - (d) 50%
- 5) The approximate size of pathologic tumor with the highest percentage of patients is
 - (a) 3
 - (b) 1
 - (c) 36
 - (d) 110
- 6) The mode takes a value near
 - (a) 3
 - (b) 1
 - (c) 36
 - (d) 110

Q26: For the sample of Saudi women ,we obtain the following graph for the serum cholesterol level(in mmol/l)



- 1) The name of the graph is
 - (a) Bar chart
 - (b) histogram
 - (c) polygon
 - (d) curve
- 2) The sample size is
 - (a) 102
 - (b) 40
 - (c) 29
 - (d) 100
- 3) The number of women with the highest serum cholesterol level is
 - (a) 82
 - (b) 5
 - (c) 26
 - (d) 29
- 4) The approximate serum cholesterol level with the lowest percentage of women is
 - (a) 5.5
 - (b) 4.5
 - (c) 7.5
 - (d) 29
- 5) The relative frequency of women with the approximate serum cholesterol level of 4.5 is
 - (a) 0.083
 - (b) 0.263
 - (c) 0.147
 - (d) 0.2
- 6) The mode takes a value near
 - (a) 5.5
 - (b) 4.5
 - (c) 7.5
 - (d) 29

Q27: The following table gives the results of a survey to study the weights of students of two schools

| | Mean weights | Standard deviation |
|------------|--------------|--------------------|
| School- A | 30 | 6 |
| School – B | 60 | 10 |

Determine whether the weights of students from school –A are more than the variable weights of students from school –B

- (a) School-A more than School-B (b) School-B more than School-A (c) cannot say
-

Q28: If we measure the hemoglobin level and RBC for a sample of women

| | Mean weights | Standard deviation |
|------------------|--------------|--------------------|
| hemoglobin level | 13.7 | 1.2 |
| RBC | 4.5 | 0.5 |

Determine which one of them is less variable?

- (a) hemoglobin level (b)RBC (c)Same variable
-

Q29: For a sample of 15 women, we have the following results of the number of their children

| | N | | Mean | St. deviation | Mode | Percentiles | | | Range |
|--------------------|-------|---------|------|---------------|------|-------------|----|----|-------|
| | valid | missing | | | | 25 | 50 | 75 | |
| Number of children | 15 | 0 | 2.8 | 1.859 | 4 | 2 | 3 | 4 | 6 |

Then,

1) The type of the variable is

- (a) qualitative (b) continuous (c) discrete (d) statistic

2) The sample mean is

- (a) 3 (b) 2.8 (c) 4 (d) 1.859 (e) 3.456 (f) 15 (h) 6

3) The sample mode is

- (a) 3 (b) 2.8 (c) 4 (d) 1.859 (e) 3.456 (f) 15 (h) 6

4) The sample standard deviation is

- (a) 3 (b) 2.8 (c) 4 (d) 1.859 (e) 3.456 (f) 15 (h) 6

5) The sample variance is

- (a) 3 (b) 2.8 (c) 4 (d) 1.859 (e) 3.456 (f) 15 (h) 6

6) The sample median is

- (a) 3 (b) 2.8 (c) 4 (d) 1.859 (e) 3.456 (f) 15 (h) 6

7) The coefficient of variation is

- (a) 18.59% (b) 15% (c) 66.39% (d) 123.42% (e) 100%

Q30: We measure the number of asthma cases seen in the past months for a sample of hospital

| | N | | Mean | variance | Mode | Median |
|------------------------|-------|---------|-------|----------|------|--------|
| | valid | missing | | | | |
| number of asthma cases | 12 | 0 | 23.25 | 128.94 | 20 | 20 |

Then,

1) The type of the variable is

- (a) qualitative (b) continuous (c) discrete (d) statistic

2) The sample mean is

- (a) 12 (b) 23.25 (c) 128.94 (d) 20 (e) 11.355 (f) 20

3) The sample mode is

- (a) 12 (b) 23.25 (c) 128.94 (d) 20 (e) 11.355 (f) 20

4) The sample standard deviation is

- (a) 12 (b) 23.25 (c) 128.94 (d) 20 (e) 11.355 (f) 20

5) The sample variance is

- (a) 12 (b) 23.25 (c) 128.94 (d) 20 (e) 11.355 (f) 20

6) The sample median is

- (a) 12 (b) 23.25 (c) 128.94 (d) 20 (e) 11.355 (f) 20

7) The coefficient of variation is

- (a) 204.76% (b) 48.84% (c) 554.5% (d) 84.82%

Chapter

Three

Some Basic Probability Concepts

Q31: For any two events, A and B:

[1] $P(A \cap B) =$

- (a) $P(A)P(B)$ (b) $P(B) - P(B \cap A^c)$ (c) $P(A) + P(B)$
 (d) $P(A) - P(B \cap A^c)$ (e) none of these

[2] $P(A \cup B) =$

- (a) 1 (b) $1 - P(A^c \cap B^c)$ (c) $P(A) + P(B)$
 (d) $P(A)P(B)$ (e) none of these

[3] If A and B are independent, then $P(B|A) =$

- (a) $P(B)$ (b) $P(B \cap A)$ (c) $P(A)$
 (d) $P(B \cap A) / P(B)$ (e) none of these

[4] $P(B) = P(A \cap B) +$

- (a) $P(B^c)$ (b) $P(A \cup B)$ (c) $P(A^c \cap B)$ (d) $P(A - B)$ (e) none of these

[5] $P(A \cap B) = 0$, then we know that A and B are

- (a) empty (b) disjoint (c) independent (d) events (e) none of these

Q32: if $\Omega = \{A, B, C, D\}$, and the outcomes are equally likely, then

[1] $P(A) =$ (a) 1 (b) 1/2 (c) 1/4 (d) 1/3 (e) none of these

[2] $2P(D) =$ (a) $P(A) - P(B)$ (b) $P(A) / 2$ (c) $P(C)$ (d) $P(A) + P(B)$ (e) none of these

Q33: A and B are events defined on the same sample space.

If $P(\bar{A}) = 0.6, P(B) = 0.5$ and $P(A \cap B) = 0.1$,

Find: (i) $P(A \cup B)$

(ii) $P(A \cap \bar{B})$

(iii) $P(\bar{A} \cap B)$

(iv) $P(\bar{A} \cap \bar{B})$

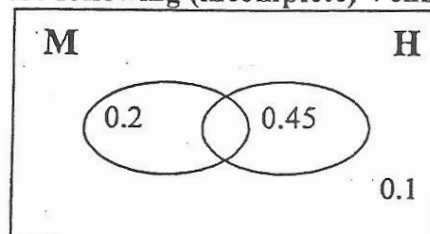
Q34: If $P(A \cap \bar{B}) = 0.3, P(A \cap B) = 0.2$ and $P(\bar{A} \cap \bar{B}) = 0.1$,

Find: (i) $P(A)$

(ii) $P(\bar{A} \cap B)$

(iii) $P(A \cup B)$

Q35: In a population of adult patients with a certain disease, let M = "is a man" and H = "has a heart disease". We have the following (incomplete) Venn diagram:



If we randomly choose one patient, find the probabilities that the patient chosen:

[1] is a man and had a heart disease:

- (a) 0.75 (b) 0.20 (c) 0.25 (d) 0.45 (e) 0.10

[2] is a women :

- (a) 0.55 (b) 0.70 (c) 0.45 (d) 0.10 (e) 0.80

[3] does not have a heart disease : (a) 0.55 (b) 0.10 (c) 0.65 (d) 0.30 (e) 0.70

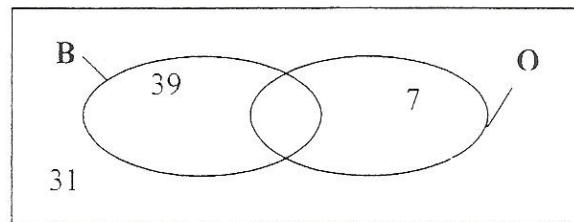
[4] is a man or has a heart disease : (a) 0.45 (b) 0.70 (c) 0.55 (d) 0.25 (e) 0.90

[5] she has heart disease : (a) 0.45 (b) 0.55 (c) 0.6429 (d) 0.8182 (e) 0.70

[6] if we know that the patient has heart disease . what effect does this have on the probability that he is a man ? (a) decreases (b) increases (c) has no effect

- (d) it is the probability of a man (e) person has heart disease

Q36: From a population of 80 babies in a certain hospital the last month, let B = "is a boy" and O = "is over weight". We have the following incomplete Venn diagram:



If we randomly choose a baby born in this hospital, find the probabilities that the baby chosen:

[1] is a boy

- (a) 0.4875 (b) 0.525 (c) 0.6125 (d) 0.9286 (e) 0.5

[2] is not an over weight baby

- (a) 0.875 (b) 0.9125 (c) 0.0875 (d) 0.4875 (e) 0.3875

[3] is a girl or an over weight baby

- (a) 0.475 (b) 0.125 (c) 0.875 (d) 0.5125 (e) 0.6125

[4] is an over weight baby knowing that he is a boy

- (a) 0.0714 (b) 0.525 (c) 0.125 (d) 0.0375 (e) 0.3

Q37: The following table classifies 400 people according to their smoking habits and whether or not they have cancer

| | Smoker (A) | Non-Smoker(\bar{A}) |
|-----------------------------------|------------|-------------------------|
| Has cancer (C) | 200 | 50 |
| Dose not have cancer(\bar{C}) | 50 | 100 |

If an individual is selected at random from this group, find the probability that he/she is

(a) a smoker and has cancer,

(b) a smoker or has cancer

(c) a non-smoker or has cancer

(d) If an individual is selected at random from the group, find the probability that the person selected:

(i) has cancer given that he/she is a smoker.

(ii) is not a smoker given that he/she dose not have cancer

(e) determine whether smoking and having cancer are independent

Chapter

Four

Probability Distribution

Q38: If X is a discrete random variable with probability distribution:

| | | | | | | |
|----------|------|------|------|------|-----|-----|
| X | 0 | 1 | 2 | 3 | 4 | 5 |
| $P(X=x)$ | 0.05 | 0.15 | 0.15 | 0.25 | 0.3 | 0.1 |

- 1) What value of X has the highest probability? (a) 1 (b) 2 (c) 3 (d) 4 (e) 5
 2) $P(X < 3) =$ (a) 0.25 (b) 0.35 (c) 0.4 (d) 0.6 (e) 0.65
 3) $P(1 \leq X < 4) =$ (a) 0.1 (b) 0.15 (c) 0.3 (d) 0.55 (e) 0.85
 4) $\mu =$ (a) 2.5 (b) 2.9 (c) 3.0 (d) 3.5 (e) 11.3

Q39: For a population of families, $x =$ the number of children in primary school. We randomly choose one and the cumulative distributed is given below:

| | |
|-----|---------------|
| X | $P(X \leq x)$ |
| 0 | 0.12 |
| 1 | 0.36 |
| 2 | 0.72 |
| 3 | 0.95 |
| 5 | 1 |

- (1) $P(X = 2) =$ (a) 0.24 (b) 0.36 (c) 0.48 (d) 0.64 (e) 0.72
 (2) $P(X = 3) =$ (a) 0.23 (b) 0.36 (c) 0.47 (d) 0.59 (e) 0.95
 (3) $P(1 < X \leq 4) =$ (a) 0 (b) 0.36 (c) 1.08 (d) 0.59 (e) 0.72
 (4) The expected number of children in primary school for a family from this population:
 (5) (a) 1.9 (b) 1.09 (c) 0.19 (d) 9.65 (e) we can't find it

Q40: For a probability of children, $X =$ the number of sweets eaten on a certain day. We randomly chose a child and the number of sweets he ate was from 0 to 4 sweets.

The following probabilities are given:

$$P(X=0) = 0.08, P(X=2) = 0.31, P(X>2) = 0.42, P(1 < X \leq 3) = 0.63.$$

Then,

- [1] $P(X = 1) =$ (a) 0.09 (b) 0.23 (c) 0.42 (d) 0.19 (e) 0
 [2] $P(X = 3) =$ (a) 0.68 (b) 0.22 (c) 0.32 (d) 0.51 (e) 0
 [3] $P(X \leq 2) =$ (a) 0.92 (b) 0.58 (c) 0.99 (d) 0.27 (e) 1
 [4] The expected number of sweets eaten daily by a child from this population is:
 (a) 1 (b) 2 (c) 0.217 (d) 4.7089 (e) 2.17

Q41: In a large population of people, 25% of them use dental floss regularly. If we randomly choose 15 of these people and let $X =$ the number of these 15 people who use dental floss regularly, then

[1] The probability distribution of X , is $P(X = x) =$

- (a) $\binom{15}{x} (0.75)^x (0.25)^{15-x}$ (b) $\binom{15}{x} (0.25)^x (0.75)^{15-x}$ (c) $\binom{25}{x} (0.15)^x (0.85)^{25-x}$
 (d) $e^{-15} (15)^x / x!$ (e) $e^{-0.25} (0.25)^x / x!$

[2] The values that x takes are:

- (a) 1, 2, ..., 15 (b) 1, 2, ..., 25 (c) 0, 1, ..., 15
 (d) 0, 1, ..., 25 (e) 0, 1, ..., ∞

[3] $P(X = 4) =$

- (a) 0.0634 (b) 0.2252 (c) 0.0202
 (d) 0.1651 (e) none of these

[4] The probability that at most one person uses dental floss regularly?

- (a) 0.9198 (b) 0.0134 (c) 0.0668
 (d) 0.0802 (e) none of these

[5] The expected value of $X =$

- (a) 3.75 (b) 11.25 (c) 15 (d) 25 (e) none of these

Q42: In a large population of young Arab children, 11 percent of them have trouble saying some letters. If we randomly choose 8 children and let X = the number of these 8 children who have trouble saying some letters, then

[1] The probability distribution of X , is $P(X = x) =$

- (a) $\binom{8}{x} (0.11)^x (0.89)^{8-x}$ (b) $\binom{11}{x} (0.08)^x (0.92)^{11-x}$ (c) $\binom{8}{x} (0.11)^x (0.89)^{8-x}$
 (d) $e^{-11} (11)^x / x!$ (e) $e^{-8} (8)^x / x!$

[2] The values that x takes are:

- (a) 1,2,...,8 (b) 1,2,...,11 (c) 0,1,...,89
 (d) 0,1,...,8 (e) 0,1,...,∞

[3] $P(X = 4) =$

- (a) 0.057 (b) 0.0064 (c) 0.0303
 (d) 0.0075 (e) none of these

[4] The probability that at most there are one children who have trouble saying some letters =

- (a) 0.7829 (b) 0.2171 (c) 0.3892
 (d) 0.3937 (e) none of these

[5] The expected value of $X =$

- (a) 11 (b) 8 (c) 0.88 (d) 7.12 (e) none of these

Q43: At a certain hospital, X = the number of kidney transplants in a year has a Poisson(5.4) distribution.

[1] The formula for $P(X = 4)$ is:

- (a) $\binom{4}{x} (0.54)^x (0.46)^{4-x}$ (b) $e^{-4} (5.4)^4 / 4!$ (c) $\binom{4}{x} (0.46)^x (0.54)^{4-x}$
 (d) $e^{-4} (4)^{5.4} / 4!$ (e) $e^{-5.4} (5.4)^4 / 4!$

[2] $P(X \geq 2) =$

- (a) 0.0045 (b) 0.9711 (c) 0.54 (d) 0.0289 (e) none of these

[3] What is the distribution of Y = the number of kidney transplants in 3 years.

- (a) Poisson(1.87) (b) Poisson(3) (c) Poisson(16.2)
 (d) Poisson(8.4) (e) Binomial(3,0.54)

Q44: in a certain population, it is known that the average number of deaths per year from cancer is 3 if the number of deaths from this disease has the Poisson distribution, find the probability that, from this disease, there are:

- (a) not more than 2 deaths in a year,
 (b) at least two deaths in a year,
 (c) exactly 3 deaths in a year,
 (d) exactly 4 deaths in 4 months.

Q45: At a certain hospital in Riyadh, X = the number of food poisoning cases in a month has a Poisson(5.6) distribution.

[1] The possible values of X are:

- (a) $x = 0,1,2$ (b) $x = 0,1,2,\dots,6$ (c) $x = 1,2,\dots,30$
 (d) $x = 0,1,2,3,4,\dots$ (e) $x = 1,2$

[2] $P(X = 7) =$

- (a) 0 (b) 0.0312 (c) 0.1267 (d) 0.4109 (e) none of these

[3] What is the expected number of food poisoning cases in 3 months?

- (a) 0.56 (b) 1.87 (c) 5.6
 (d) 7.47 (e) 16.8

[4] What is the distribution of $Y =$ the number of food poisoning cases in 1/2 a month?

- (a) Poisson(0.373) (b) Poisson(2.8) (c) Poisson(5.6)
(d) Poisson(11.2) (e) Poisson(84)

Q46: In a certain population, an average of 6 new cases of breast cancer are diagnosed each year. If the number of new diagnosed cases of this disease in the population follows the Poisson distribution, find the probability that, in this population, there are:

- (a) no new cases of the disease in a year,
(b) at most two new cases of the disease in 6 months.

Q47: If Z a standard normal distribution, then

- [1] $P(z > 1.34) =$ (a) 0.0901 (b) 0.34 (c) 0.3669 (d) 0.6331
[2] $P(z \leq -0.78) =$ (a) 0.0375 (b) 0.2177 (c) 0.4681 (d) 0.7823
[3] $P(-1.89 \leq z < 2.56) =$ (a) 0.7486 (b) 0.9112 (c) 0.9654 (d) 0.9929

Q48: If Z is $N(0,1)$, find

- (a) $P(Z \leq 1.36)$ (b) $P(Z \geq 2.4)$ (c) $P(Z < 1.81)$
(d) $P(Z > 2.7)$ (e) $P(-1.2 < Z < 2.1)$
(f) $P(Z = 1.4)$ (g) $P(-2.36 < Z < 1.45)$

Q49: Let Z have the standard normal distribution. The value of z such that $P(Z \leq z) = \alpha$, is defined by z_α . Find the following z values:

- (a) $z_{0.5}$ (b) $z_{0.95}$ (c) $z_{0.025}$ (d) $z_{0.975}$

Q50: In population of people, $X =$ the body mass index (in kg/m^2) is normally distributed with mean $\mu = 25$ and standard deviation $\sigma = 2$. For a randomly chosen person,

- [1] $P(X < 21) =$
(a) 0.9772 (b) 0.4772 (c) 0.9821 (d) 0.0228 (e) none of these
[2] $P(19 < X < 28) =$
(a) 0.9332 (b) 0.9319 (c) 0.9345 (d) 0.5332 (e) none of these
[3] The value of x such $P(X > x) = 0.2578$ is:
(a) 0.65 (b) 1.3 (c) 26.3 (d) 23.7 (e) none of these

Q51: In a population of heat stroke victims, $X =$ the cooling time (in min.) is normally distributed with $\mu = 40$ and $\sigma^2 = 100$. For a randomly chosen victim,

- [1] $P(X \leq 25) =$
(a) 0.0062 (b) 0.0668 (c) 0.4404 (d) 0.5596 (e) 0.9332
[2] $P(20 \leq X \leq 70) =$
(a) 0.1587 (b) 0.8413 (c) 0.9544 (d) 0.9759 (e) 1.0215
[3] The value of x such $P(X > x) = 0.791$ is:
(a) 0.81 (b) 18.4 (c) 31.9 (d) 47.9 (e) 48.1

Chapter

Five

Statistical Inference

Q52: Answer 1 to 3 directly from the question. Answer 4 to 6 using new information as given

Statement: In a sample of 49 Saudis living in villas, the mean Vitamin D level was 16 with a variance of 12. Assume that the distribution is normal. **Question:** Find and interpret a 90% confidence interval for the average.

[1] The number 12 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] the value of α is:

- (a) 0.01 (b) 0.10 (c) 0.45 (d) 0.95 (e) unknown

[3] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
 (c) normal, σ^2 unknown (d) not normal, σ^2 unknown, n large
 (e) none of these

[4] Suppose that the assumption are "normal, σ^2 known" then the correct formula to answer the question is:

- (a) $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (b) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

[5] using $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$, calculate the confidence interval with $\alpha = 0.05, s = 3.8$ and $n = 49$

- (a) (14.9066, 17.0934) (b) (14.9096, 17.0904) (c) (14.936, 17.064)
 (d) (15.107, 16.893) (e) (15.3041, 16.6959)

[6] If a 90% confidence interval for the average gives (15.2, 16.9), what part is missing in the indicated place in the interpretation below:

Interpretation: we are 90% sure that the (a) the (b) Vitamin D level (c) of (d) Saudis living in villa: between 15.2 and 16.9.

- (a) If the mean was 16 (b) average (c) with a variance of 12
 (d) 49 (e) all of these are missing

Q53: The random variable X, representing the lifespan of a certain light bulb is distributed normally with mean of 400 hours ,and standard deviation of 10 hours.

1. What is the probability that a particular light bulb will last for more than 380 hours?
2. What is the probability that a particular light bulb will last for exactly 399 hours?
3. What is the probability that a particular light bulb will last for between 380 and 420 hours?
4. The mean is
5. The variance is.....
6. The standard deviation.....

Q54: The tensile of a certain type of thread is approximately normally distributed with standard deviation of 6.8 Kg. A sample of 20 pieces of the thread has an average strength of 72.8 Kg. Then

1. A point estimate of the population mean of tensile strength μ is
(a) 72.8 (b) 20 (c) 6.8 (d) 46.24 (e) none of these
2. A 98% Confident interval for mean of tensile strength μ ,the lower bound equal to:
(a) 68.45 (b) 69.26 (c) 71.44 (d) 69.68 (e) none of these
3. A 98% Confident interval for mean of tensile strength μ ,the upper bound equal to:
(a) 74.16 (b) 77.15 (c) 75.92 (d) 76.34 (e) none of these

Q55: Suppose that we are interested in making some statistical inferences about the mean μ of normal population with standard deviation 2.0. Suppose that a random sample of size $n = 49$ from this population gave the sample mean 4.5

1. The distribution of \bar{X} is
(a) $N(0,1)$ (b) $t(48)$ (c) $N(\mu,(0.02857)^2)$ (d) $N(\mu,2.0)$

2. A good point estimate for μ is

- (a) 4.5 (b) 2 (c) 2.5 (d) 7 (e) 1.125

3. Assumptions is

- (a) Normal, σ known (b) Normal, σ unknown
(c) not Normal, σ known (d) not Normal, σ unknown

4. A 95% confident interval for μ is

- (a) (3.44,5.56) (b) (3.34,5.66) (c) (4.444, 4.556) (d) (3.94,5.05) (e) (3.04,5.96)

Q56: An electronics company wanted to estimate in monthly operating expenses riyals (μ). Assume that the population variance equals 0.584.

Suppose that a random sample of size 49 is taken and found that the sample mean equals 5.47. Find

- (1) Point estimate for μ
- (2) The distribution of \bar{X} is
- (3) The assumptions?
- (4) A 90% confident interval for μ .

Q57: In a random sample of 500 homes in a certain city, it is found that 114 are heated by oil. Let π be the proportion of homes in this city that are heated by oil.

1. Find a point estimate for π .

2. Construct a 98% confidence interval for π .

Q58: In a study involved 1200 car drivers, it was found that 50 car drivers do not use seat belt.

1. A point estimate for the proportion of car drivers who do not use seat belt is:
(A) 50 (B) 0.0417 (C) 0.9583 (D) 1150 (E) None of these
2. The lower limit of a 95% confidence interval of the proportion of car drivers not using seat belt is
(A) 0.0322 (B) 0.0416 (C) 0.0304 (D) -0.3500 (E) None of these
3. The upper limit of a 95% confidence interval of the proportion of car drivers not using seat belt is
(A) 0.0417 (B) 0.0530 (C) 0.0512 (D) 0.4333 (E) None of these

Q59: A study was conducted to make some inferences about the proportion of female employees (π) in a certain hospital. A random sample gave the following data:

| | |
|-------------------|-----|
| Sample size | 250 |
| Number of females | 120 |

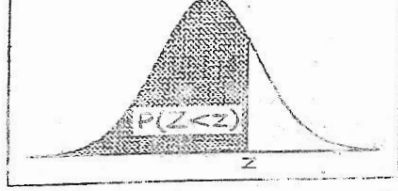
1. Calculate a point estimate (p) for the proportion of female employees (π).
2. Construct a 90% confidence interval for p .

Q60: A random sample of 200 students from a certain school showed that 15 students smoke. let π be the proportion of smokers in the school.

1. Find a point estimate for π
2. Find 95% confidence interval for π

Q61: A researcher was interested in making some statistical inferences about the proportion of females (π) among the students of a certain university. A random sample of 500 students showed that 150 students are female.

1. A good point estimate for π is
(A) 0.31 (B) 0.30 (C) 0.29 (D) 0.25 (E) 0.27
2. The lower limit of a 90% confidence interval for π is
(A) 0.2363 (B) 0.2463 (C) 0.2963 (D) 0.2063 (E) 0.2663
3. The upper limit of a 90% confidence interval for π is
(A) 0.3337 (B) 0.3137 (C) 0.3637 (D) 0.2937 (E) 0.3537



Areas Under The Standard Normal Curve

| Z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -3.4 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0002 |
| -3.3 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0003 |
| -3.2 | 0.0007 | 0.0007 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0005 | 0.0005 | 0.0005 |
| -3.1 | 0.0010 | 0.0009 | 0.0009 | 0.0009 | 0.0008 | 0.0008 | 0.0008 | 0.0008 | 0.0007 | 0.0007 |
| -3.0 | 0.0013 | 0.0013 | 0.0013 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |
| -2.9 | 0.0019 | 0.0018 | 0.0018 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| -2.8 | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| -2.7 | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| -2.6 | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0041 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| -2.5 | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| -2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| -2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| -2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| -2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| -2.0 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| -1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| -1.8 | 0.0359 | 0.0351 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| -1.7 | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| -1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| -1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| -1.4 | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0721 | 0.0708 | 0.0694 | 0.0681 |
| -1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| -1.2 | 0.1151 | 0.1131 | 0.1112 | 0.1093 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.9985 |
| -1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| -1.0 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| -0.9 | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| -0.8 | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| -0.7 | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2296 | 0.2266 | 0.2236 | 0.2206 | 0.2177 | 0.2148 |
| -0.6 | 0.2743 | 0.2709 | 0.2676 | 0.2643 | 0.2611 | 0.2578 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| -0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| -0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| -0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| -0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3974 | 0.3935 | 0.4286 | 0.4247 |
| -0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9987 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

Appendix for chapter 5:

The values of standard normal $Z_{1-\alpha/2}$ that used I (1- α)% confident interval

| Level of confidence | $Z_{1-\alpha/2}$ |
|----------------------------|------------------------------------|
| 99% | 2.575 |
| 98% | 2.33 |
| 95% | 1.96 |
| 90% | 1.645 |
| 80% | 1.28 |