MEDICAL UNIVERSITY – PLEVEN
FACULTY OF PUBLIC HEALTH

DEPARTMENT OF PUBLIC HEALTH SCIENCES
CENTRE FOR DISTANT LEARNING

LECTURE No2

DISTRIBUTIONS. DESCRIPTIVE STATISTICS FOR QUALITATIVE DATA

Assoc. Prof. G. Grancharova, MD. PhD

Plan of the lecture

Part 1. Distributions. Normal distribution. Standard scores and standard normal curve. Asymmetric distributions.

Part 2. Simple descriptive statistics for categorical data.

Part 1

DISTRIBUTIONS

VARIABLE DISTRIBUTION

- Some of variable values are more frequent than the others.

- The way how frequent the values of a particular variable are is called ***probability distributions or frequency distributions.***

FREQUENCY DISTRIBUTIONS

We could roughly classify the distributions into two groups:

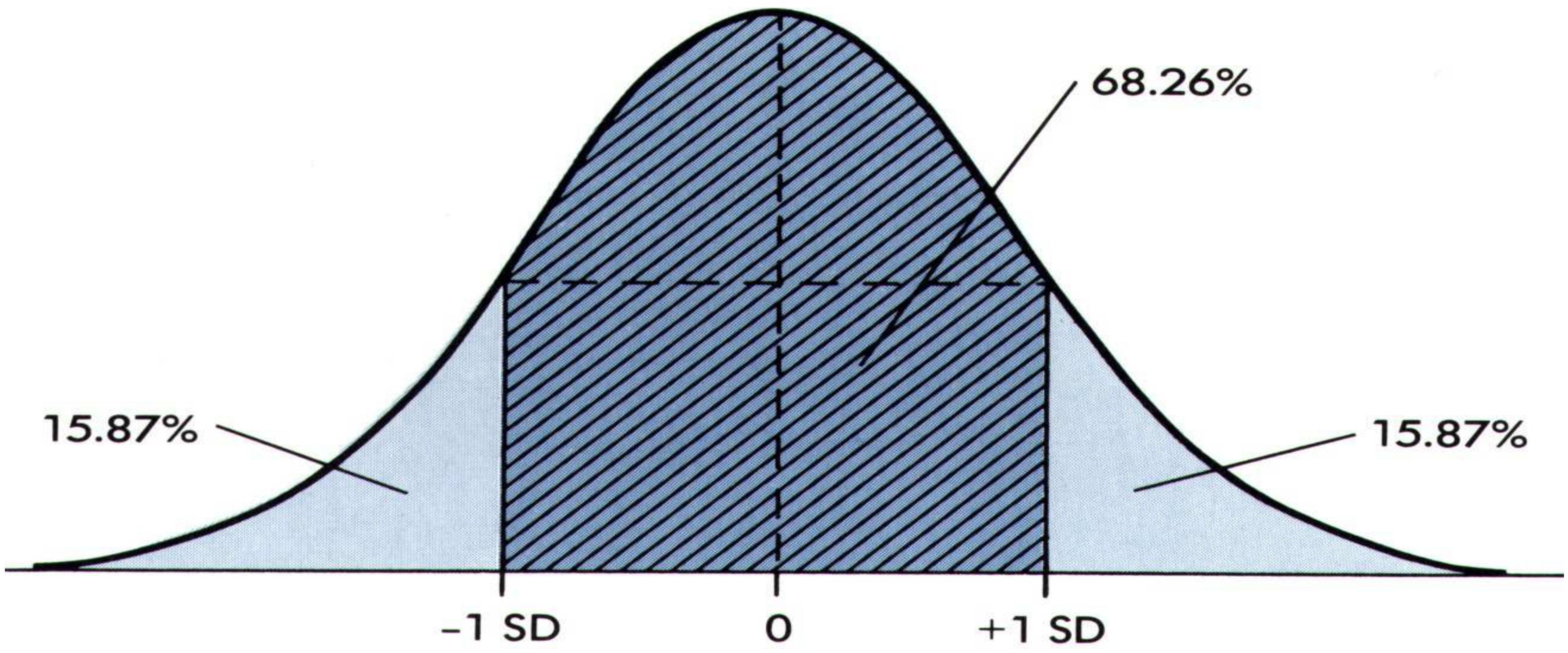
- ***Empirical probability distributions* – distributions observed in real situation;**
- ***Theoretical (mathematical) probability distributions* – mathematical idealization of distributions observed in real situations.**

FREQUENCY DISTRIBUTIONS

The most important theoretical probability distribution is known as ***Normal or Gaussian distribution.***

Other important theoretical distributions are:

- ***Binomial distribution,***
- ***Chi square distribution,***
- ***Poisson distribution,***
- ***Fisher's F distribution.***



Normal distribution and standard normal curve

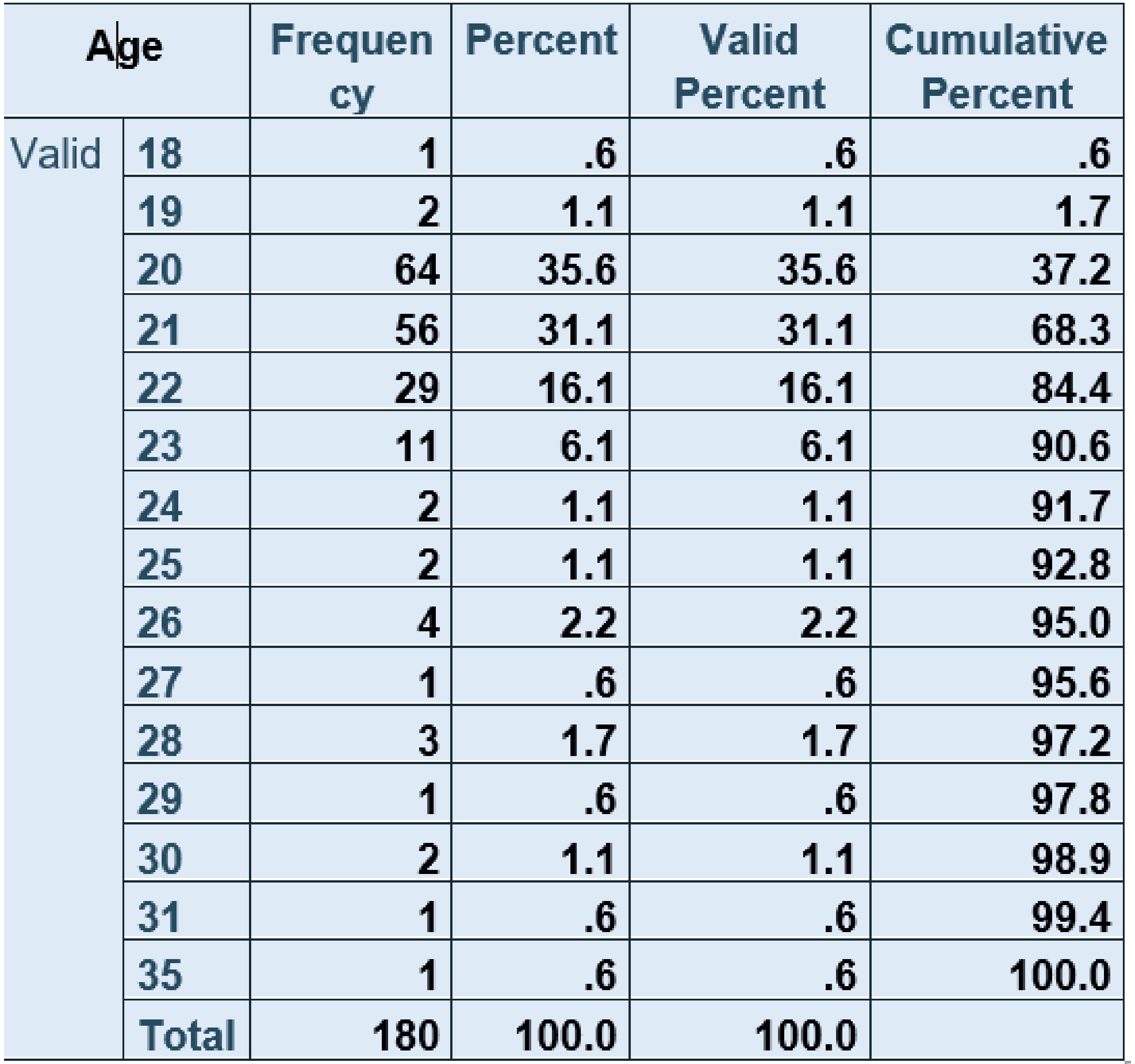
FREQUENCY DISTRIBUTIONS

- ***Frequency distribution*** - the range of all values is divided into ordered classes, and the number of observations into each class is determined.
- ***Absolute frequency (f) -*** the actual number of subjects with a certain score of whose score fall between a particular class interval.
- ***Relative frequency distribution*** – it may be obtained by dividing the absolute frequencies by the total number of observations.
- ***Percentage frequency distribution*** – the same as a relative frequency distribution expressed in percentages.

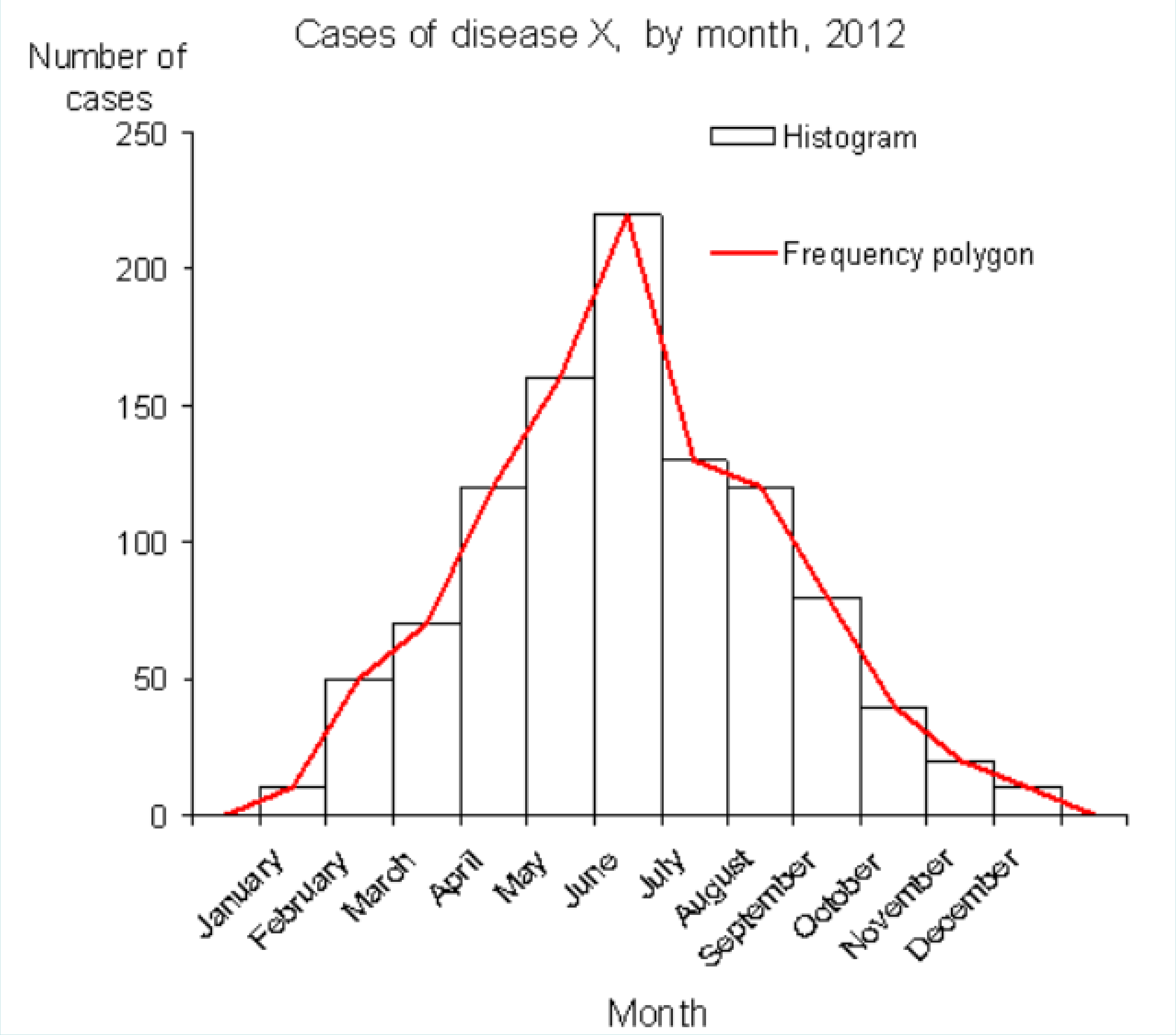
FREQUENCY DISTRIBUTIONS

- ***Cumulative frequency distribution*** – it results from adding up each successive percent in the relative frequency column.
- ***Graphing: histogram and polygon***
- ***Histogram -*** a type of bar graph in which all bars are linked to each other. It is usually used with interval and ratio-type data. When the class intervals are equal in width, the columns (or bars) are all the same width.
- ***Frequency polygon*** - a line graph to represent quantitative data which is used to compare sets of data or to display a cumulative frequency distribution. It emphasizes the overall pattern in the data.

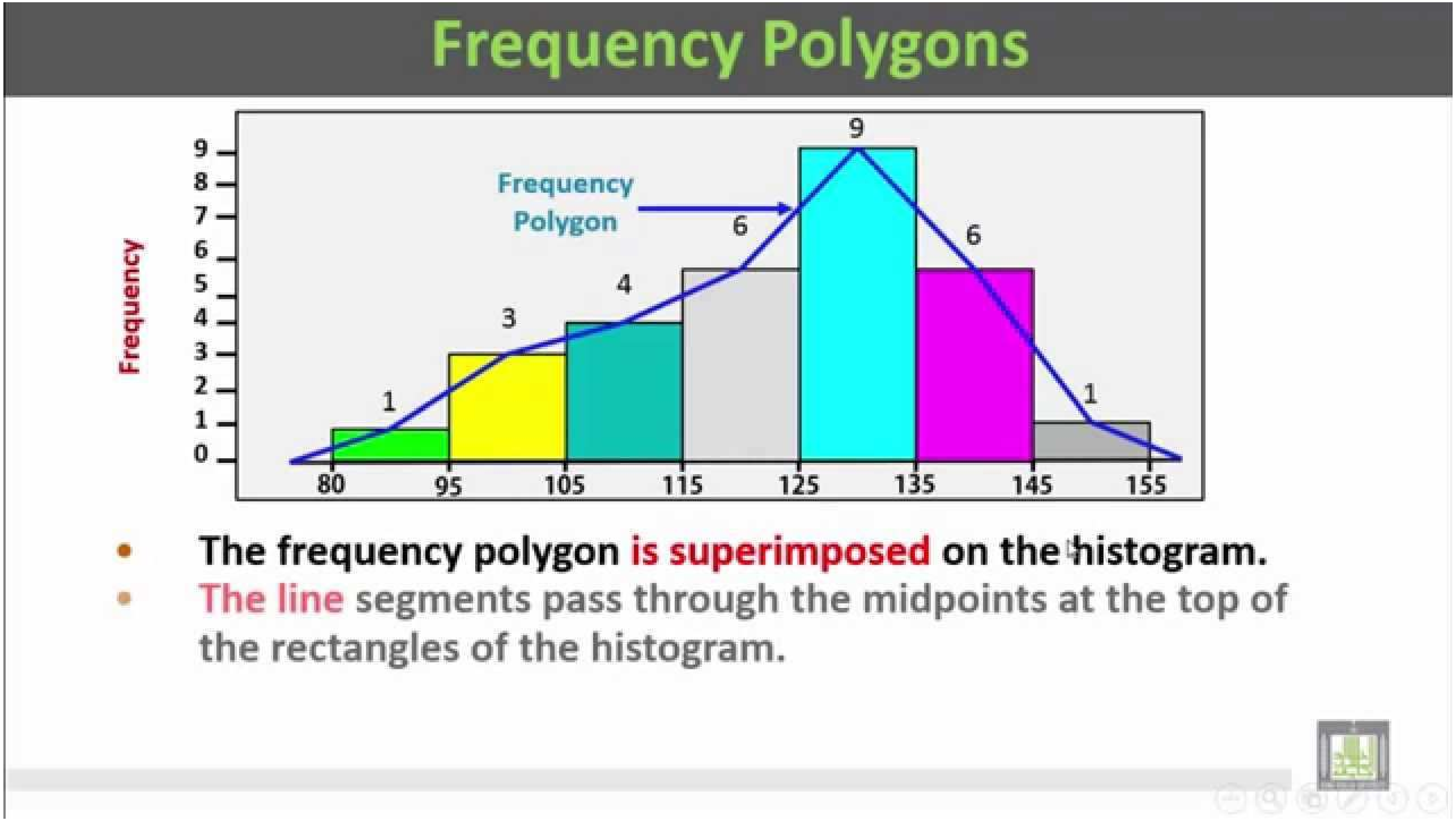
Absolute frequency distribution, percentage and cumulative percentage distribution of 180 students by age



Age		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18	1	.6	.6	.6
	19	2	1.1	1.1	1.7
	20	64	35.6	35.6	37.2
	21	56	31.1	31.1	68.3
	22	29	16.1	16.1	84.4
	23	11	6.1	6.1	90.6
	24	2	1.1	1.1	91.7
	25	2	1.1	1.1	92.8
	26	4	2.2	2.2	95.0
	27	1	.6	.6	95.6
	28	3	1.7	1.7	97.2
	29	1	.6	.6	97.8
	30	2	1.1	1.1	98.9
	31	1	.6	.6	99.4
	35	1	.6	.6	100.0
	Total	180	100.0	100.0	



Histogram and frequency polygon

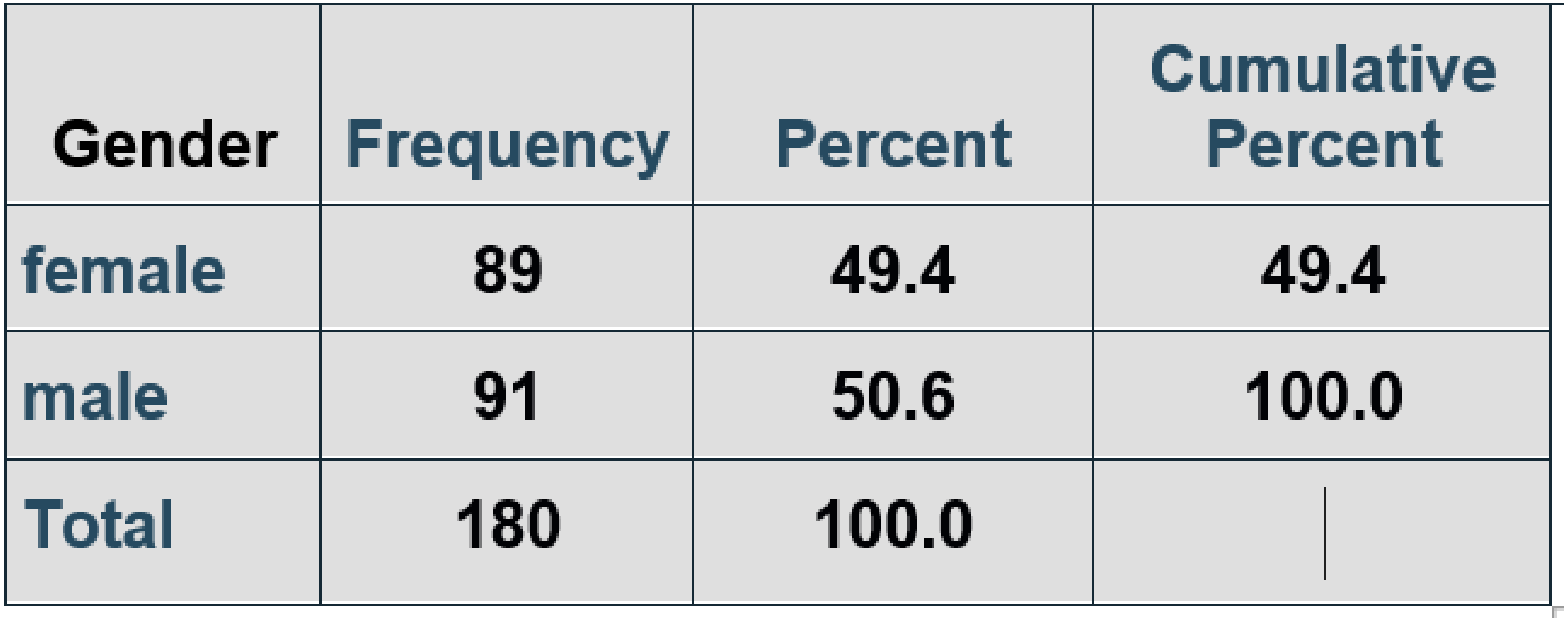


Frequency Polygons

- The frequency polygon is superimposed on the histogram.
- The line segments pass through the midpoints at the top of the rectangles of the histogram.

Examples of frequency distributions for categorical variables

Absolute frequency distribution, percentage and cumulative percentage distribution of 180 students by gender



Gender	Frequency	Percent	Cumulative Percent
female	89	49.4	49.4
male	91	50.6	100.0
Total	180	100.0	

At the table you can see three types of distributions:

1. Binominal frequency distribution – by females and males

2. Percentage frequency distribution

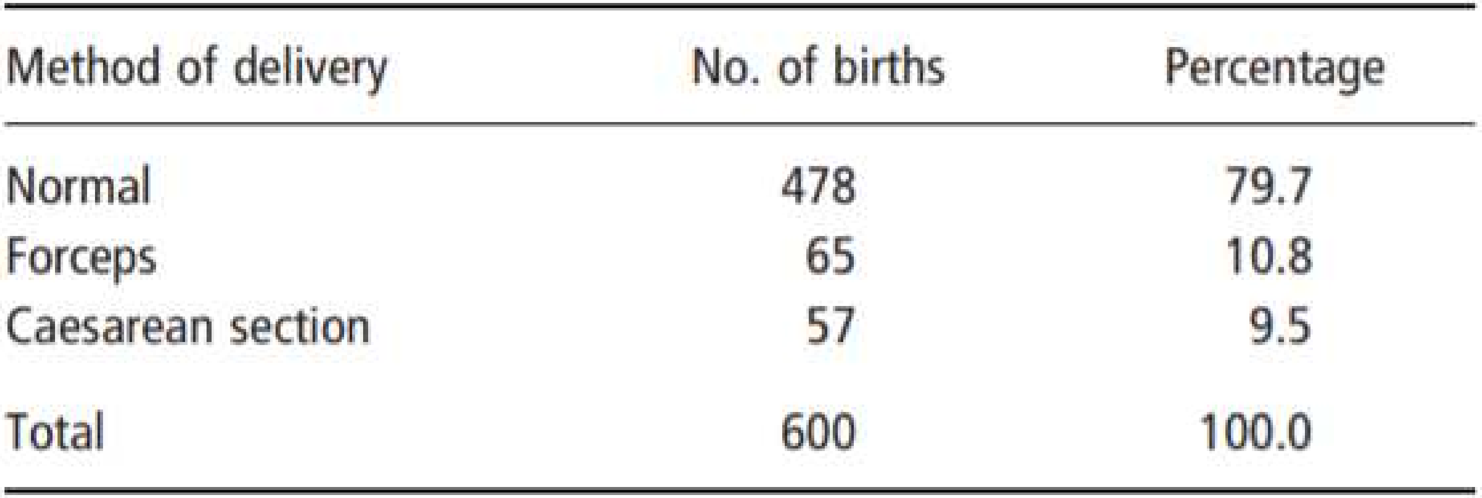
3. Cumulative frequency distribution

The above examples of distributions are produced by IBM SPSS Statistics but it is not always to have at your hand such a powerful statistical package.

Summarizing categorical variables is straightforward, the main task being to count the number of observations in each category. These counts are called **frequencies.**

They are often also presented as relative frequencies; that is as **proportions or percentages** of the total number of individuals.

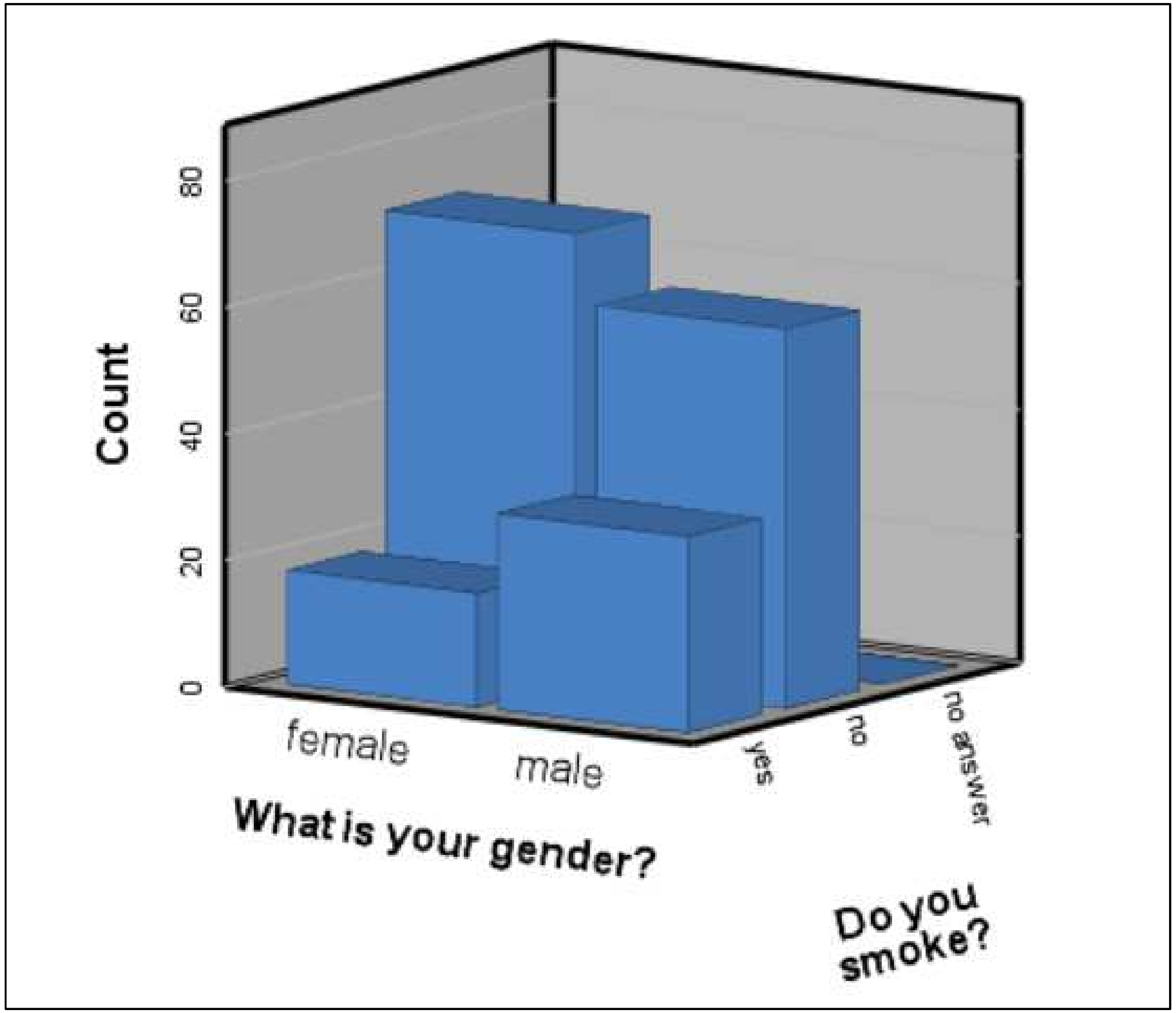
Example: Frequency distribution of delivery of 600 babies born in a hospital by the method of delivery



Method of delivery	No. of births	Percentage
Normal	478	79.7
Forceps	65	10.8
Caesarean section	57	9.5
Total	600	100.0

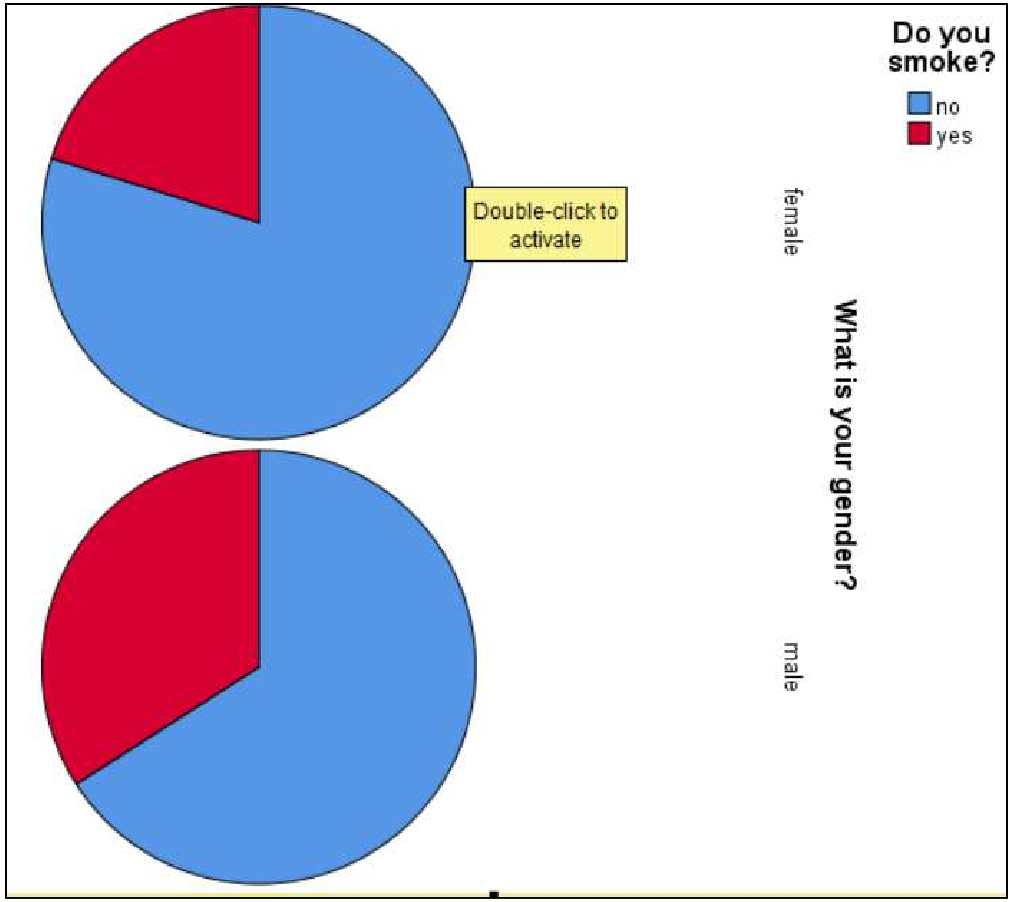
The variable of interest is the method of delivery, a categorical variable with three categories: normal delivery, forceps delivery, and caesarean section.

Frequencies and relative frequencies are commonly illustrated by **a bar chart** (also known as a bar diagram) or by **a pie chart**. In a bar chart the lengths of the bars are drawn proportional to the frequencies.



Alternatively the bars may be drawn proportional to the percentages in each category; the shape is not changed, only the labelling of the scale.

In a **pie chart** the circle is divided so that the areas of the sectors are proportional to the frequencies, or equivalently to the percentages.



Examples of frequency distributions for numerical variables

If there are more than about 20 observations, a useful first step in summarizing a numerical (quantitative) variable is to form a **frequency distribution**.

This is a table showing the number of observations at different values or within certain ranges.

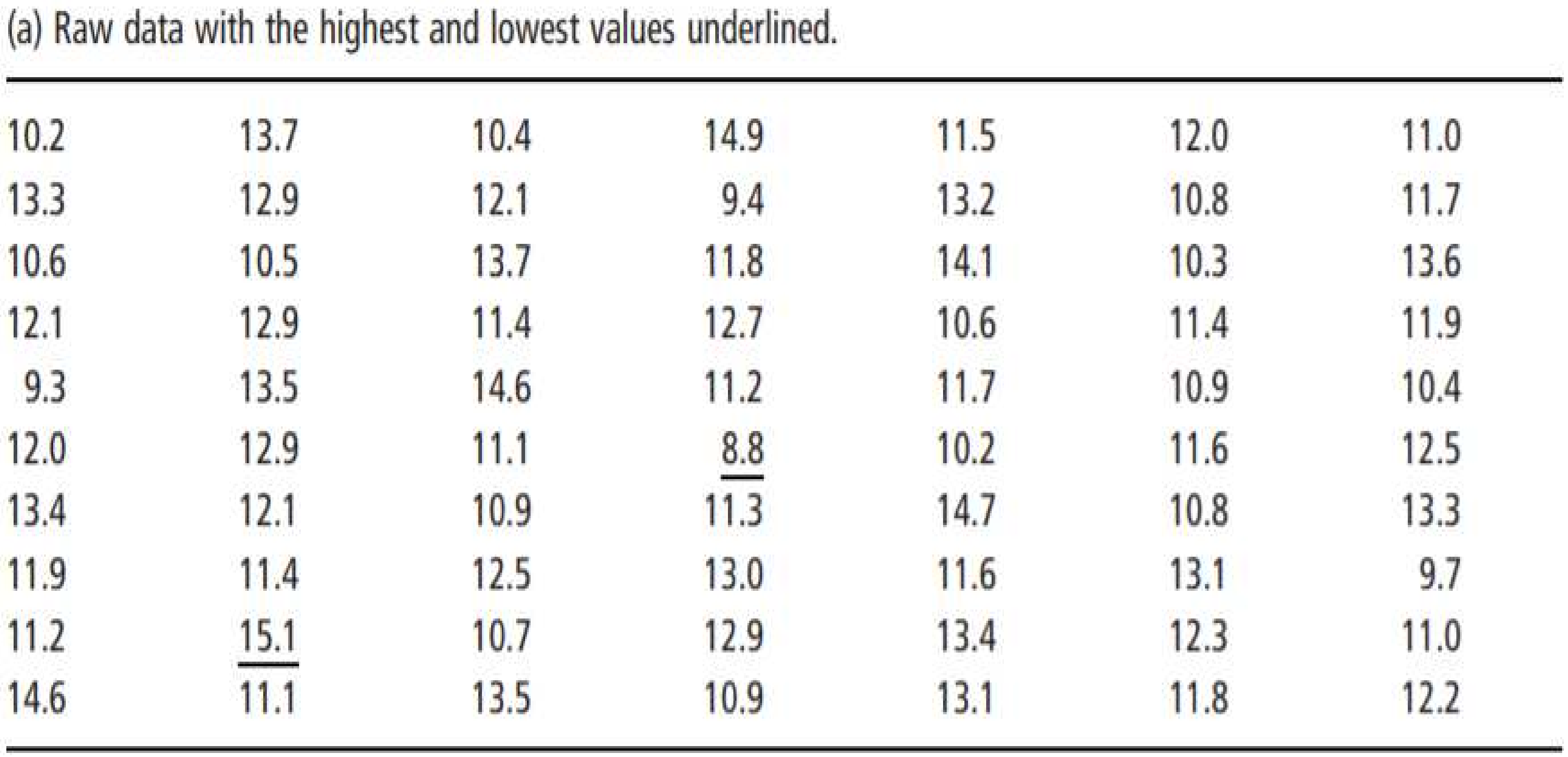
For discrete variables the frequencies may be tabulated either for each value of the variable or for groups of values.

With continuous variables, groups have to be formed. An example is given in the next slide, where haemoglobin has been measured **in g/100 ml for 70 women.**

When forming **a frequency distribution:**

1. Firstly, we have **to count the number of observations** and **identify the lowest and highest values**.

Example: Haemoglobin levels in g/100 ml for 70 women



(a) Raw data with the highest and lowest values underlined.

10.2	13.7	10.4	14.9	11.5	12.0	11.0
13.3	12.9	12.1	9.4	13.2	10.8	11.7
10.6	10.5	13.7	11.8	14.1	10.3	13.6
12.1	12.9	11.4	12.7	10.6	11.4	11.9
9.3	13.5	14.6	11.2	11.7	10.9	10.4
12.0	12.9	11.1	8.8	10.2	11.6	12.5
13.4	12.1	10.9	11.3	14.7	10.8	13.3
11.9	11.4	12.5	13.0	11.6	13.1	9.7
11.2	15.1	10.7	12.9	13.4	12.3	11.0
14.6	11.1	13.5	10.9	13.1	11.8	12.2

2. Secondly, we have to decide whether the data should be grouped and what grouping interval should be used.

As a rough guide we may have 5–20 groups, depending on the number of observations.

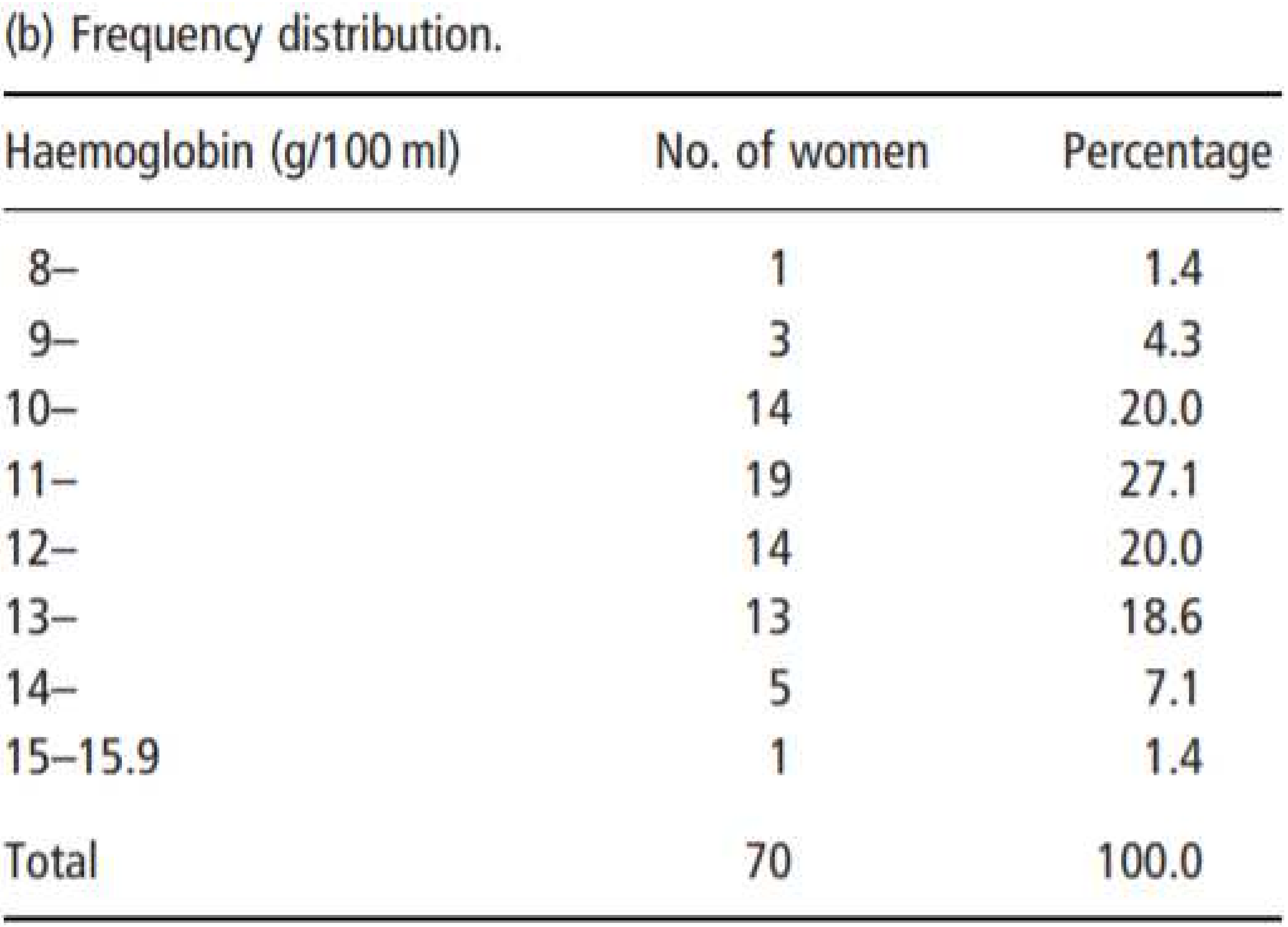
If the interval chosen for grouping the data is too wide, too much detail will be lost, while if it is too narrow the table will be unwieldy.

The starting points of the groups should be round numbers and, whenever possible, all the intervals should be of the same width. There should be no gaps between groups. The table should be labelled so that it is clear what happens to observations that fall on the boundaries.

In the above example, there are 70 haemoglobin measurements. **The lowest value is 8.8 and the highest 15.1 g/100 ml.**

The most appropriate for this example are intervals of width 1 g/100 ml and thus, we can establish 8 groups, labelling them as 8–, 9–, . . . and so on. An acceptable alternative would have been 8.0–8.9, 9.0–9.9 and so on.

3. Once the format of the table is decided, **the numbers of observations (frequencies)** in each group should be counted.

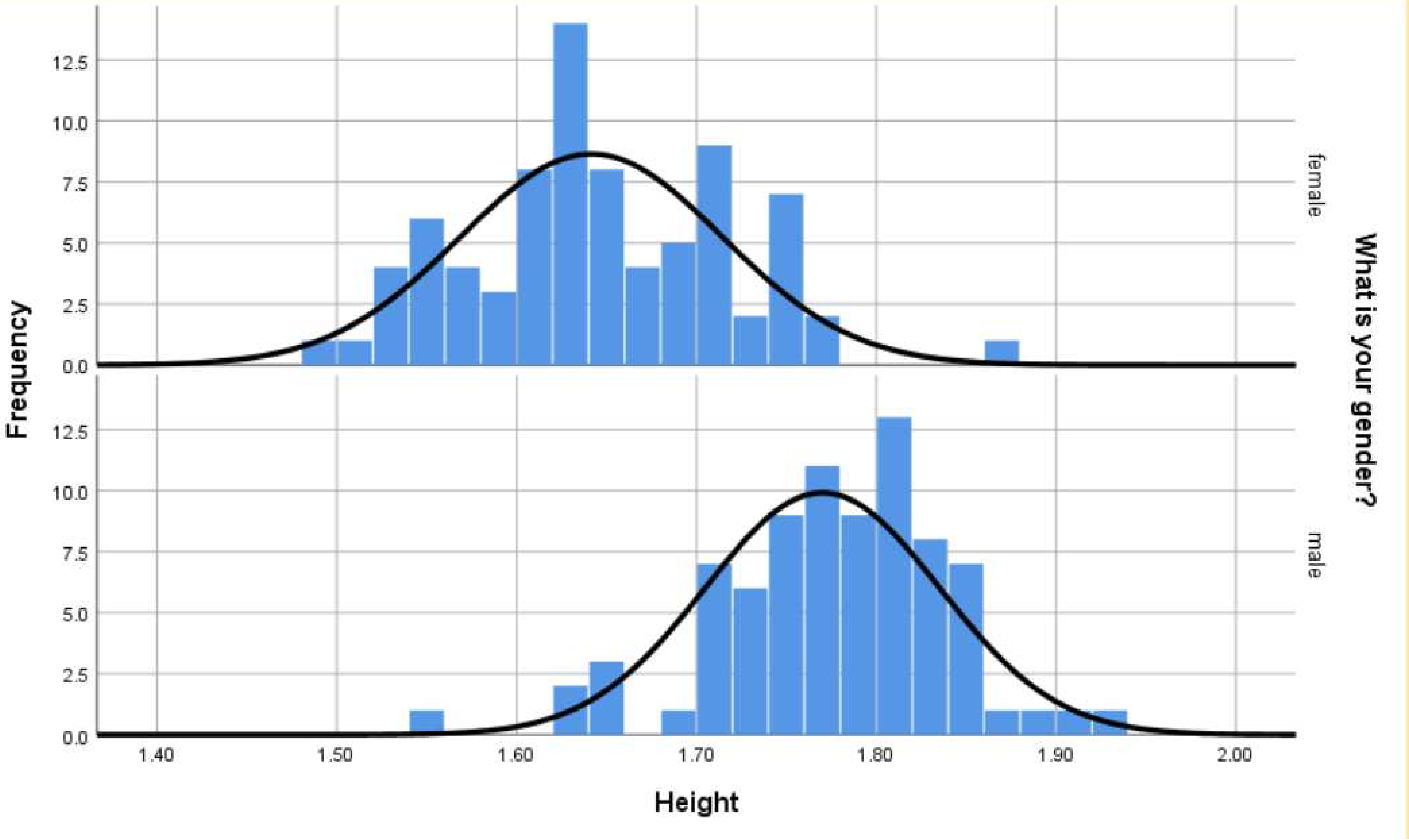


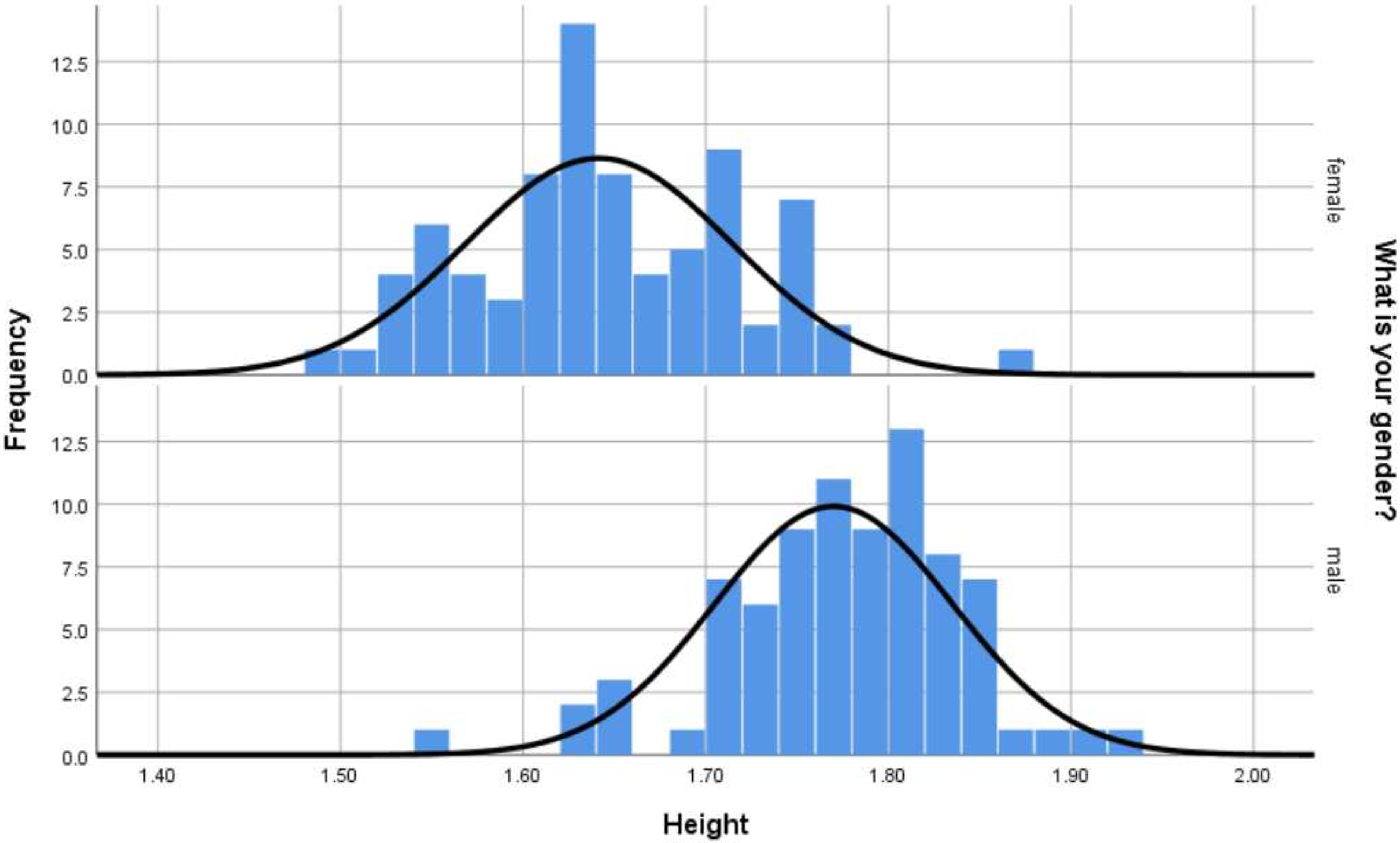
(b) Frequency distribution.

Haemoglobin (g/100 ml)	No. of women	Percentage
8–	1	1.4
9–	3	4.3
10–	14	20.0
11–	19	27.1
12–	14	20.0
13–	13	18.6
14–	5	7.1
15–15.9	1	1.4
Total	70	100.0

Frequency distributions are usually illustrated by histograms. Either the frequencies or the percentages may be used; the shape of the histogram will be the same.

The construction of a histogram is straightforward when the grouping intervals of the frequency distribution are all equal, as is the cases below.

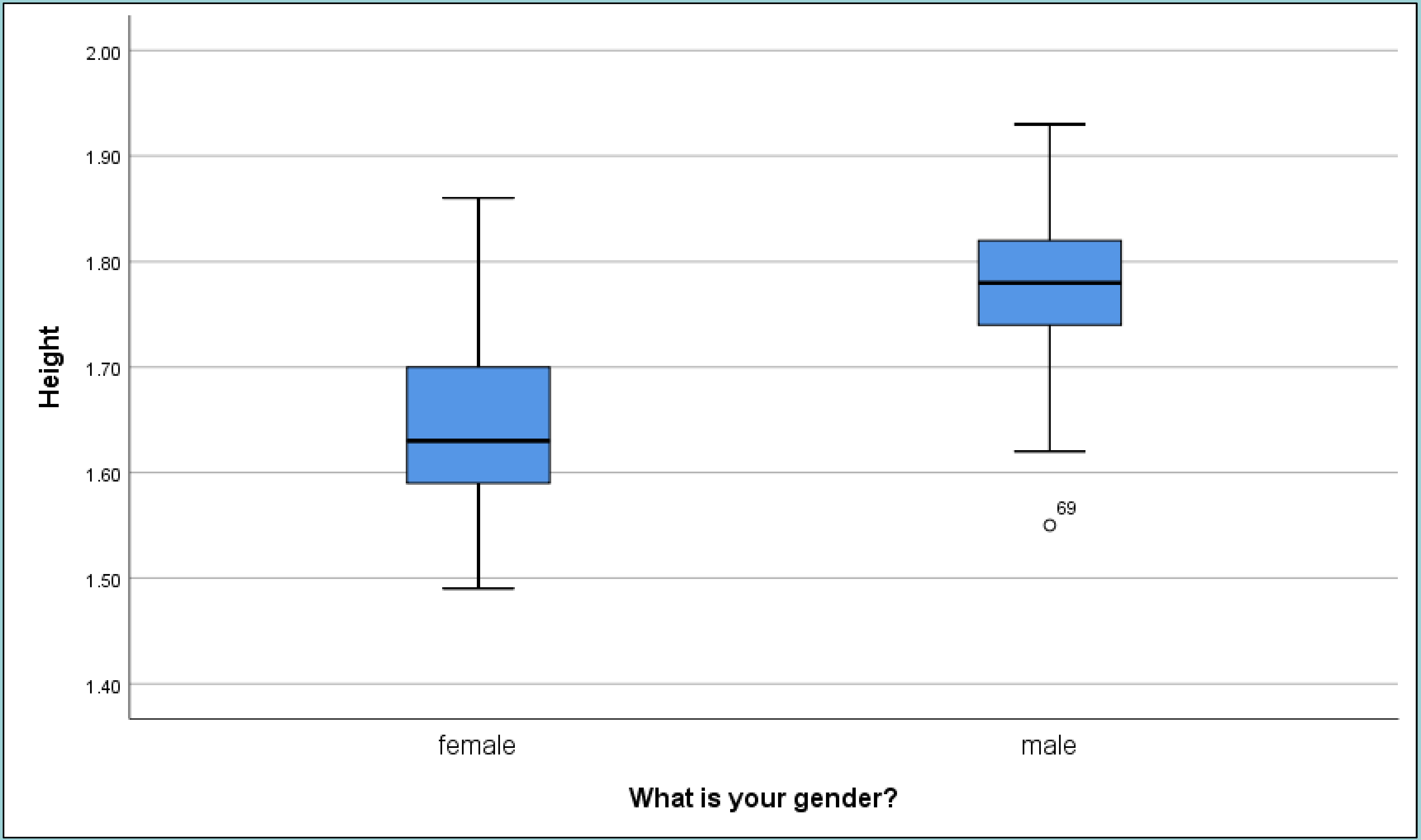

Frequency
12.5
10.0
7.5
5.0
2.5
0.0
female
male
What is your gender?
1.40
1.50
1.60
1.70
1.80
1.90
2.00
Height


Frequency
12.5
10.0
7.5
5.0
2.5
0.0
female
male
What is your gender?
1.40
1.50
1.60
1.70
1.80
1.90
2.00
Height

To display a distribution of a numerical variable some other types of graphs can be used such as:

- **Box-plot** – it is a plot in which a rectangle is drawn to represent the second and third quartiles, usually with a vertical line inside to indicate the median value. The lower and upper quartiles are shown as horizontal lines either side of the rectangle.

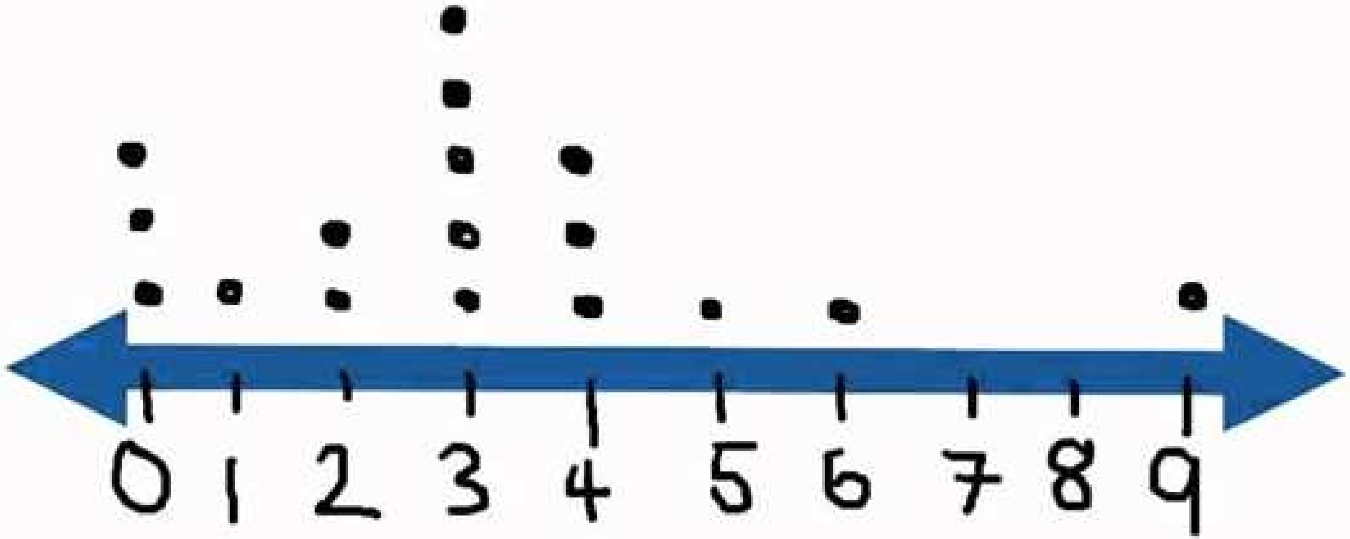
- **Dot-plot** - also called a **dot** chart or strip **plot**, is a type of simple histogram-like chart used in statistics for relatively small data sets where values fall into a number of discrete bins (categories).



Box-plot of height for 180 medical students by gender

17 students we asked how many text messages they had sent on a particular day.

~~0~~, ~~3~~, ~~2~~, ~~0~~, ~~4~~, ~~5~~, ~~4~~, ~~3~~, ~~3~~, ~~1~~, ~~4~~, ~~6~~, ~~3~~, ~~0~~, ~~3~~, ~~2~~, ~~9~~



Dot-plot for relatively small data set

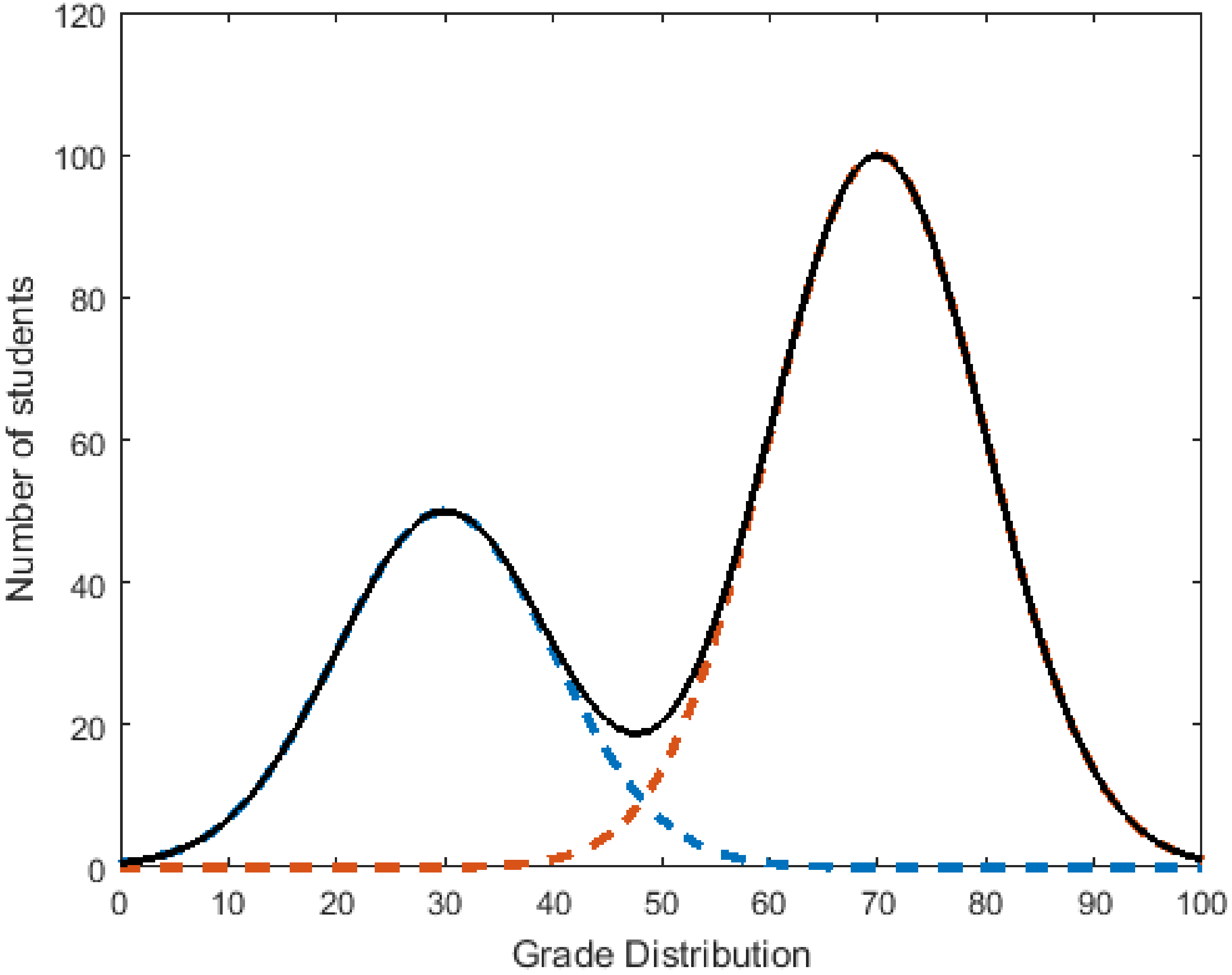
DESCRIBING A DISTRIBUTION

When the graphical presentation of the shape of a distribution is done it should be described. The shape itself depends on the number and features of the place of highest density (peak).

Regarding the number of peaks:

- ***unimodal*** **–** distributions with a single peak,
- ***bimodal*** **–** distributions with two peaks,
- ***polymodal*** – distributions with more than two peaks.

Example of bimodal distribution



DESCRIBING A DISTRIBUTION

Regarding the shape of the peak:

- ***bell shaped*** **–** distributions in which extreme values tend to be less likely than values in the middle of the ordered series,

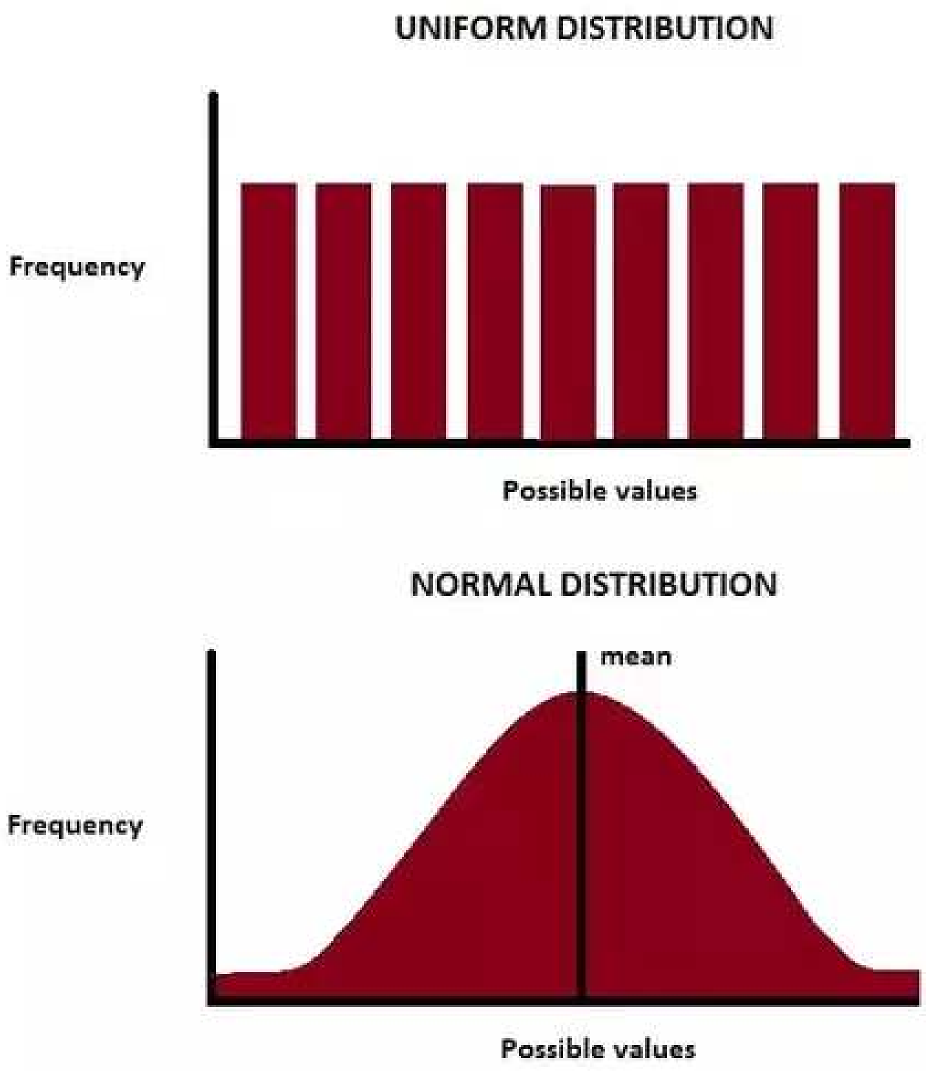
- ***uniform*** **–** sometimes also known as a **rectangular distribution**, is a **distribution** that has constant probability, e.g. in which all values have the same frequency.

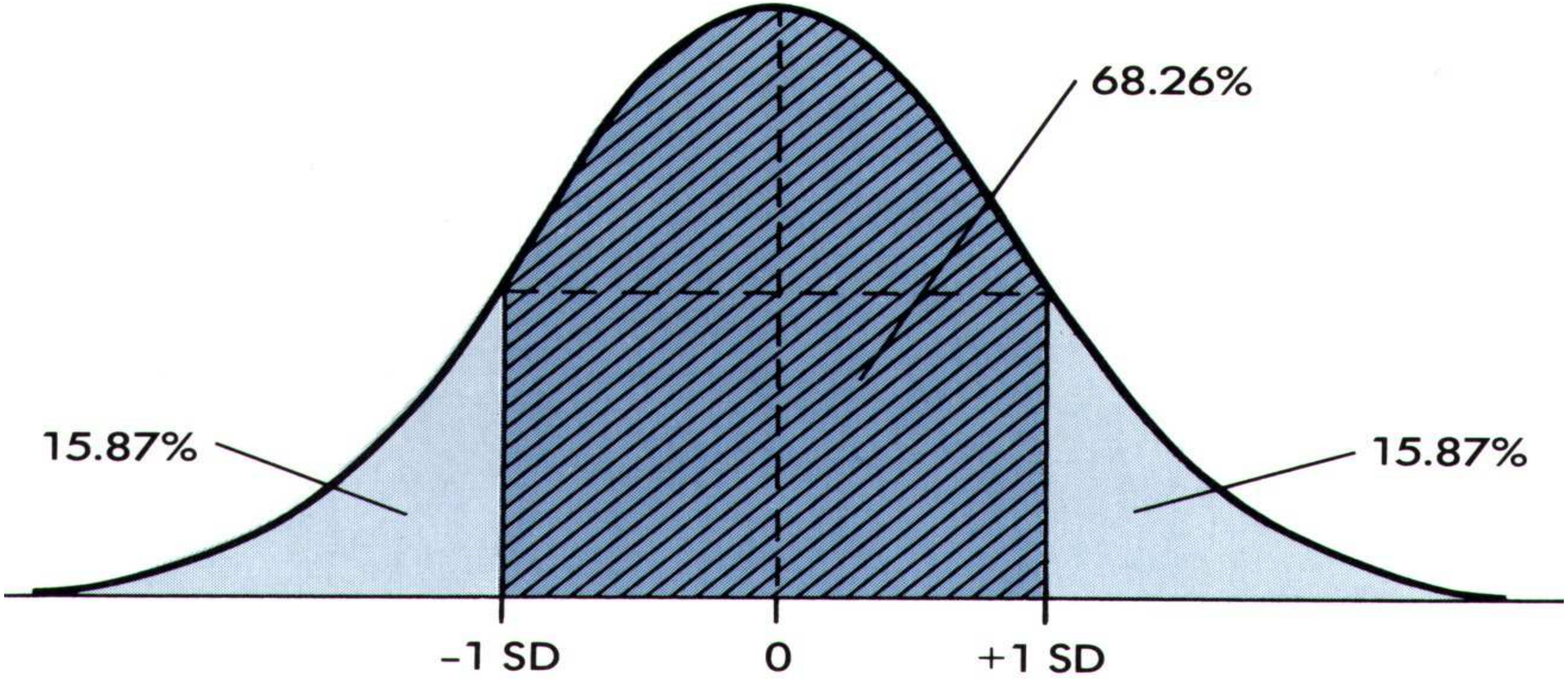
DESCRIBING A DISTRIBUTION

- Regarding the symmetry:

- Normal distribution (symmetrical, bell-shaped, Gaussian) - the mean, median, and mode coincide in the centre

- Non-normal distributions (asymmetrical, skewed)


UNIFORM DISTRIBUTION
Frequency
Possible values
NORMAL DISTRIBUTION
mean
Frequency
Possible values



Normal, bell-shaped, symmetrical or Gaussian distribution

STANDARD SCORES

- **Standard scores** are a way of expressing a score in terms of its relative distance from the mean. Such “transformed” scores are called z scores or standard scores.

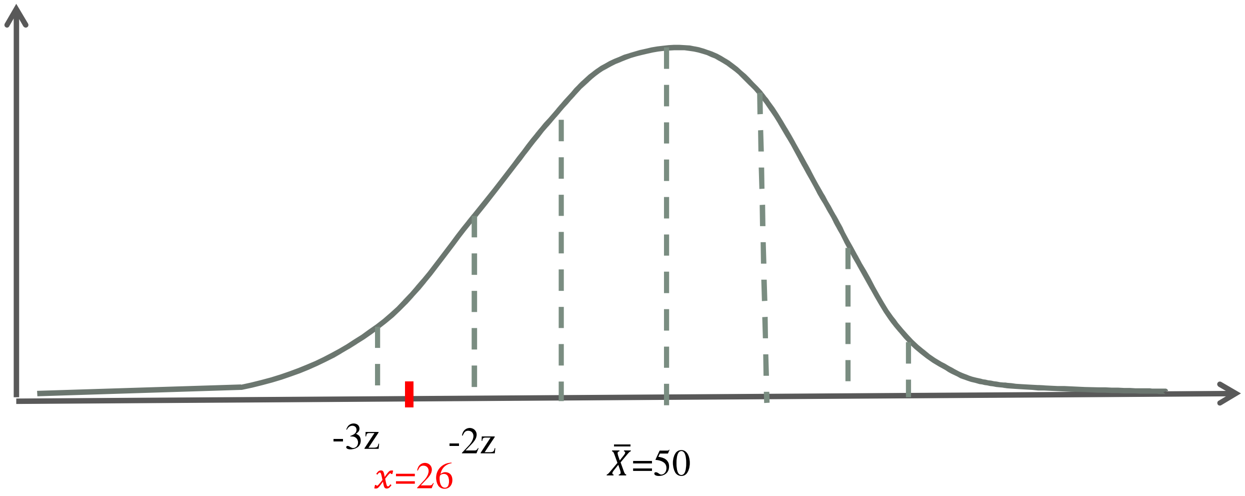
- $Z = \dfrac{x - \overline{X}}{s}$

Z score represents how many standard deviations a given raw score is above or below the mean.

STANDARD SCORES

Example: As a simple application, what portion of a normal distribution with a mean of 50 and a standard deviation of 10 is below 26? Applying the formula, we obtain the following result:

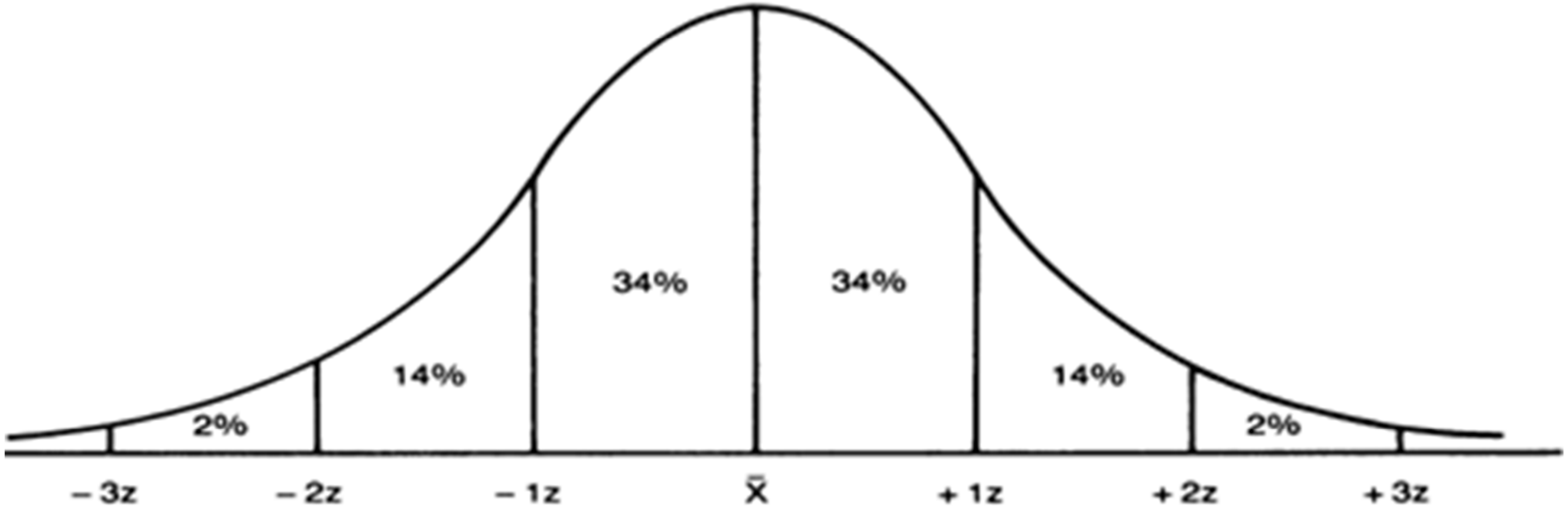
$$Z = \frac{x - \bar{X}}{s} = \frac{26 - 50}{10} = -2{,}4$$



STANDARD SCORES

Example: Infant A walked unaided at the age of 40 weeks, while infant B is 65 weeks old but still cannot walk. What sense can we make of these measurements?

- We need additional information to compare these data with norms for other children. Suppose that $\overline{X}$=50 weeks and s=5 weeks;
- Infant’s A score is 2s below the mean
  - (40-50) : 5 = - 2, e.g. z = - 2
- Infant’s B score is 3s above the mean
  - (65-50) : 5 = 3, e.g. z = 3

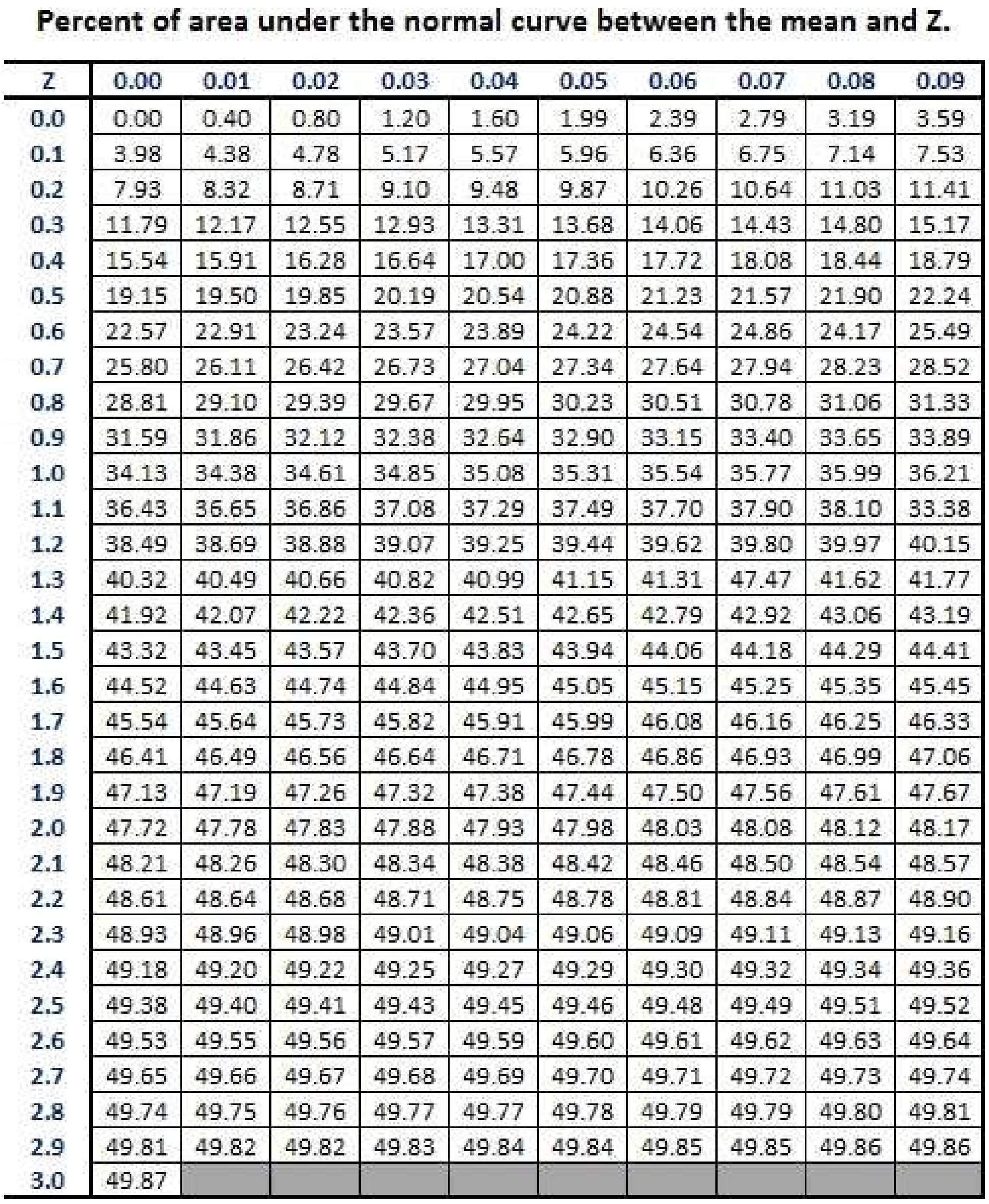

34%
34%
14%
14%
2%
2%
– 3z
– 2z
– 1z
$\bar{X}$
+ 1z
+ 2z
+ 3z

THE NORMAL CURVE

- Normal curve - it is a theoretically perfect frequency polygon in which the mean, median, and mode all coincide and which takes the form of a symmetrical bell-shaped curve.
- Characteristics of the normal curve:
- 1. Most of the cases fall close to the mean;
- 2. Relatively few cases fall into the high or low values of x.

STANDARD NORMAL CURVE

- 3. We can use appropriate tables to estimate the area under the standard normal curve for any given z scores.
- 4. The area under the curve between any two points is directly proportional to the percentage of cases falling between those two points.
- All these properties underlie the calculations of the limits for ‘norms’ and is used in clinical practice to determine the so called “normative groups”.



Percent of area under the normal curve between the mean and Z.

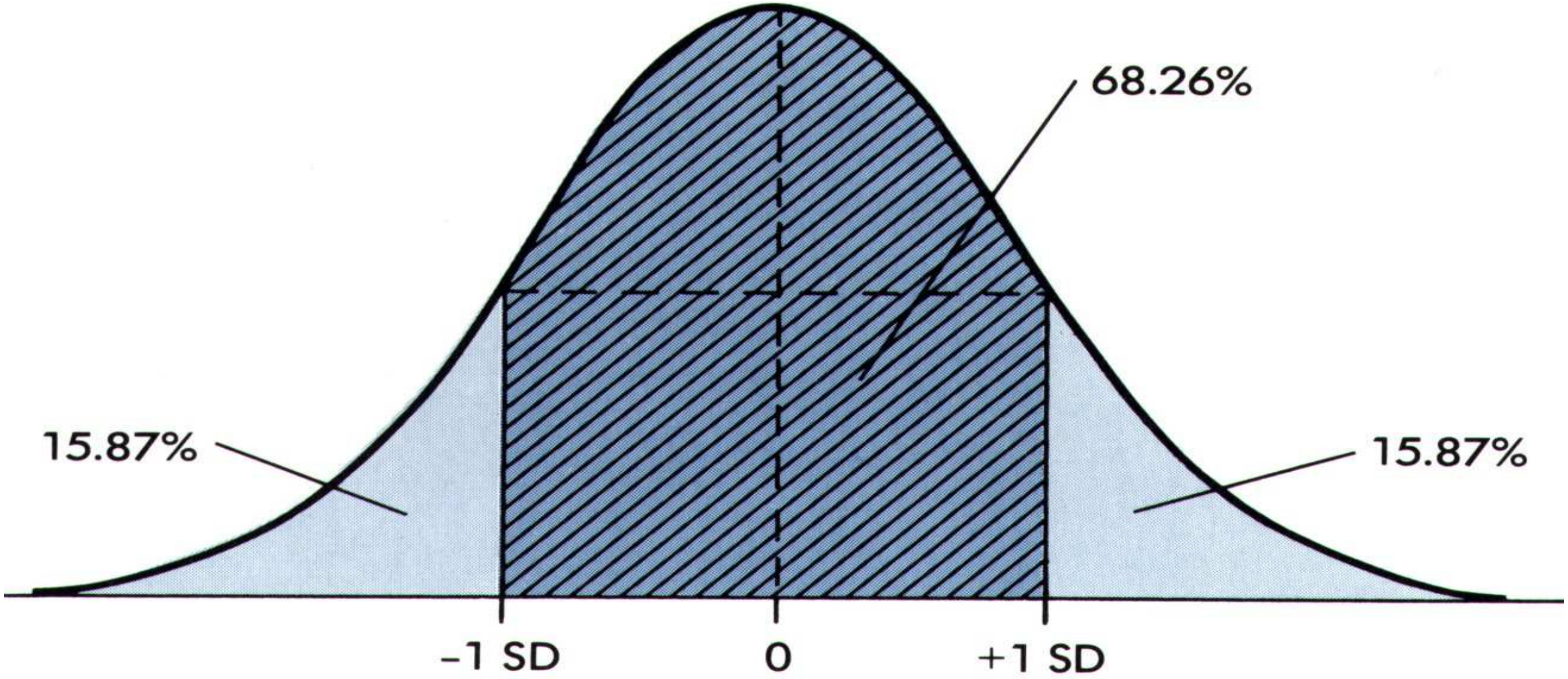
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00	0.40	0.80	1.20	1.60	1.99	2.39	2.79	3.19	3.59
0.1	3.98	4.38	4.78	5.17	5.57	5.96	6.36	6.75	7.14	7.53
0.2	7.93	8.32	8.71	9.10	9.48	9.87	10.26	10.64	11.03	11.41
0.3	11.79	12.17	12.55	12.93	13.31	13.68	14.06	14.43	14.80	15.17
0.4	15.54	15.91	16.28	16.64	17.00	17.36	17.72	18.08	18.44	18.79
0.5	19.15	19.50	19.85	20.19	20.54	20.88	21.23	21.57	21.90	22.24
0.6	22.57	22.91	23.24	23.57	23.89	24.22	24.54	24.86	24.17	25.49
0.7	25.80	26.11	26.42	26.73	27.04	27.34	27.64	27.94	28.23	28.52
0.8	28.81	29.10	29.39	29.67	29.95	30.23	30.51	30.78	31.06	31.33
0.9	31.59	31.86	32.12	32.38	32.64	32.90	33.15	33.40	33.65	33.89
1.0	34.13	34.38	34.61	34.85	35.08	35.31	35.54	35.77	35.99	36.21
1.1	36.43	36.65	36.86	37.08	37.29	37.49	37.70	37.90	38.10	33.38
1.2	38.49	38.69	38.88	39.07	39.25	39.44	39.62	39.80	39.97	40.15
1.3	40.32	40.49	40.66	40.82	40.99	41.15	41.31	47.47	41.62	41.77
1.4	41.92	42.07	42.22	42.36	42.51	42.65	42.79	42.92	43.06	43.19
1.5	43.32	43.45	43.57	43.70	43.83	43.94	44.06	44.18	44.29	44.41
1.6	44.52	44.63	44.74	44.84	44.95	45.05	45.15	45.25	45.35	45.45
1.7	45.54	45.64	45.73	45.82	45.91	45.99	46.08	46.16	46.25	46.33
1.8	46.41	46.49	46.56	46.64	46.71	46.78	46.86	46.93	46.99	47.06
1.9	47.13	47.19	47.26	47.32	47.38	47.44	47.50	47.56	47.61	47.67
2.0	47.72	47.78	47.83	47.88	47.93	47.98	48.03	48.08	48.12	48.17
2.1	48.21	48.26	48.30	48.34	48.38	48.42	48.46	48.50	48.54	48.57
2.2	48.61	48.64	48.68	48.71	48.75	48.78	48.81	48.84	48.87	48.90
2.3	48.93	48.96	48.98	49.01	49.04	49.06	49.09	49.11	49.13	49.16
2.4	49.18	49.20	49.22	49.25	49.27	49.29	49.30	49.32	49.34	49.36
2.5	49.38	49.40	49.41	49.43	49.45	49.46	49.48	49.49	49.51	49.52
2.6	49.53	49.55	49.56	49.57	49.59	49.60	49.61	49.62	49.63	49.64
2.7	49.65	49.66	49.67	49.68	49.69	49.70	49.71	49.72	49.73	49.74
2.8	49.74	49.75	49.76	49.77	49.77	49.78	49.79	49.79	49.80	49.81
2.9	49.81	49.82	49.82	49.83	49.84	49.84	49.85	49.85	49.86	49.86
3.0	49.87									

Normal Distribution

The normal distribution is a continuous probability distribution which is very important in many fields of science, and especially in medicine.

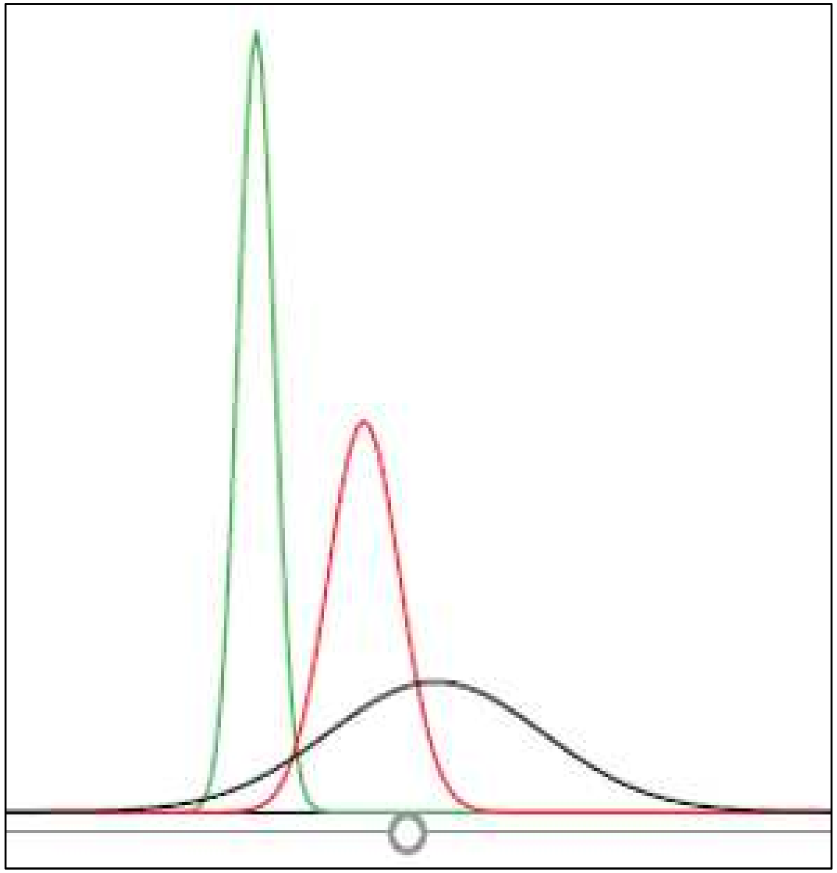
It is also called **Gaussian distribution** because it was discovered by Carl Friedrich Gauss.

Normal distributions are a family of distributions of the same general form. They may differ in their location and scale: the mean ("average") of the distribution defines its location, and the standard deviation ("variability") defines the scale as it can be seen in the next slide.


68.26%
15.87%
15.87%
−1 SD
0
+1 SD

Normal Distribution

Normal distributions can differ in their means and in their standard deviations. The diagram shows three normal distributions. The green (left-most) distribution has a mean of -3 and a standard deviation of 0.5, the distribution in red (the middle distribution) has a mean of 0 and a standard deviation of 1, and the distribution in black (right-most) has a mean of 2 and a standard deviation of 3. These as well as all other normal distributions are symmetric with relatively more values at the center of the distribution and relatively few in the tails.

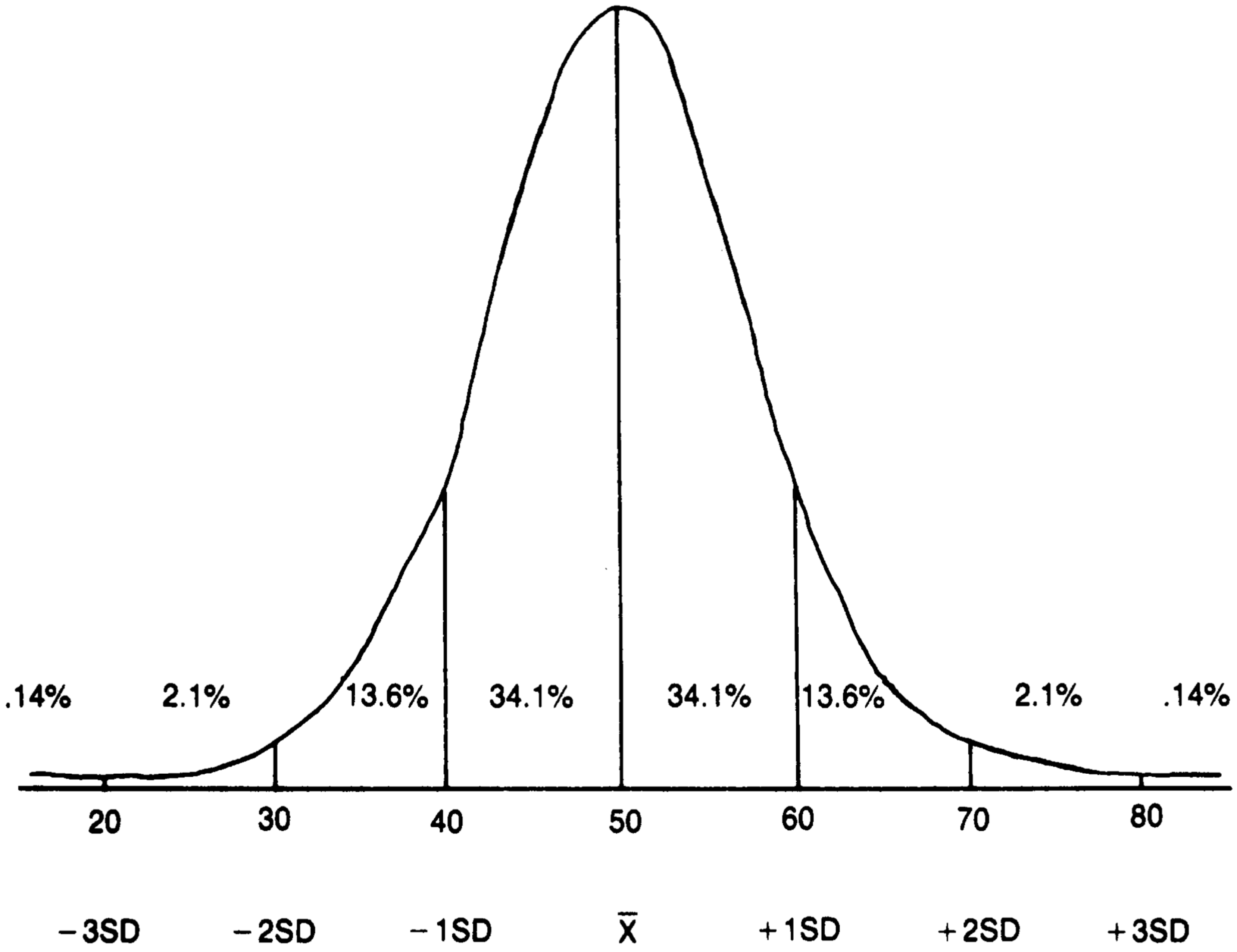


Normal distributions differing in their means and standard deviations

Normal Distribution

The mean and the standard deviation in the normal distribution can be used in interpreting individual scores within a distribution.

Using the basic principle of normal distribution, we can determine exactly where a particular score is situated and what percentages constitute the area under the normal curve from the mean and this score. Based on this principle the limits of different groups of normality can be determined (see the next slide).


.14%
2.1%
13.6%
34.1%
34.1%
13.6%
2.1%
.14%
20
30
40
50
60
70
80
−3SD
−2SD
−1SD
$\bar{X}$
+1SD
+2SD
+3SD

Normal Distribution

Basic features of the normal distributions:

1. Normal distributions are symmetric around their mean

2. The mean, median, and mode of normal distribution are equal

3. The area under the normal curve is equal to 1.0

4. Normal distributions are denser in the center and less dense in the tails

5. Normal distributions are defined $\bar{\boldsymbol{x}}$ by two parameters, the mean $\bar{\boldsymbol{x}}$

and the standard deviation **s.**

6. 68% of the area of a normal distribution is within one standard deviation of the mean

7. Approximately 95% of the area of a normal distribution is within two standard deviations of the mean.

8. Approximately 99,7% of the area of a normal distribution is within three standard deviations of the mean.

31.10.2019 г.

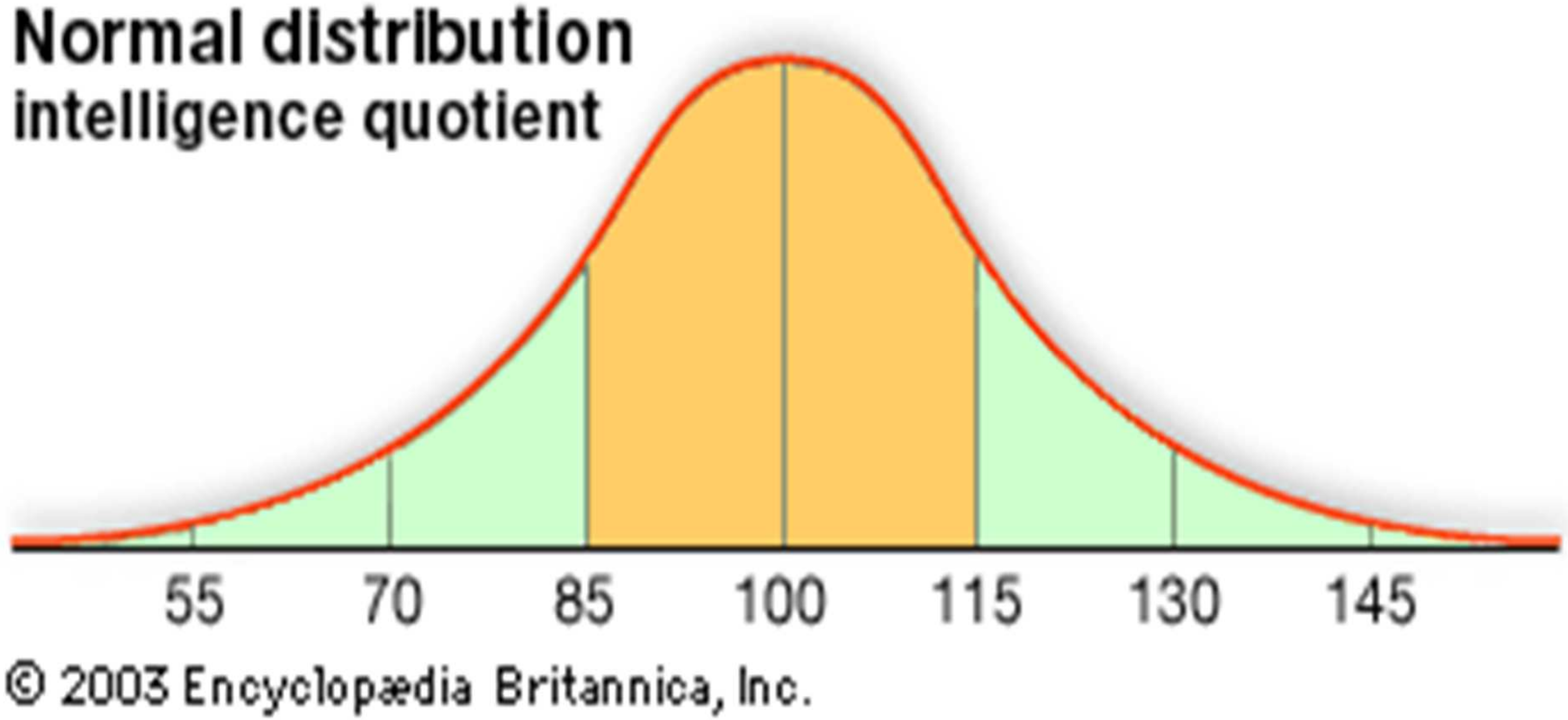
Normal Distribution

Basic principle of a normal distribution

Number of standard deviations (z or t) from the mean	*% of results lying inside* $\overline{x} \pm s$	*% of results lying outside* $\overline{x} \pm s$
0,5	38,2	61,4
1	68,2	31,8
1,96	95	5
2,58	99	1
3,00	99,7	0,3
3,29	99,9	0,1

31.10.2019 г.

Normal Distribution

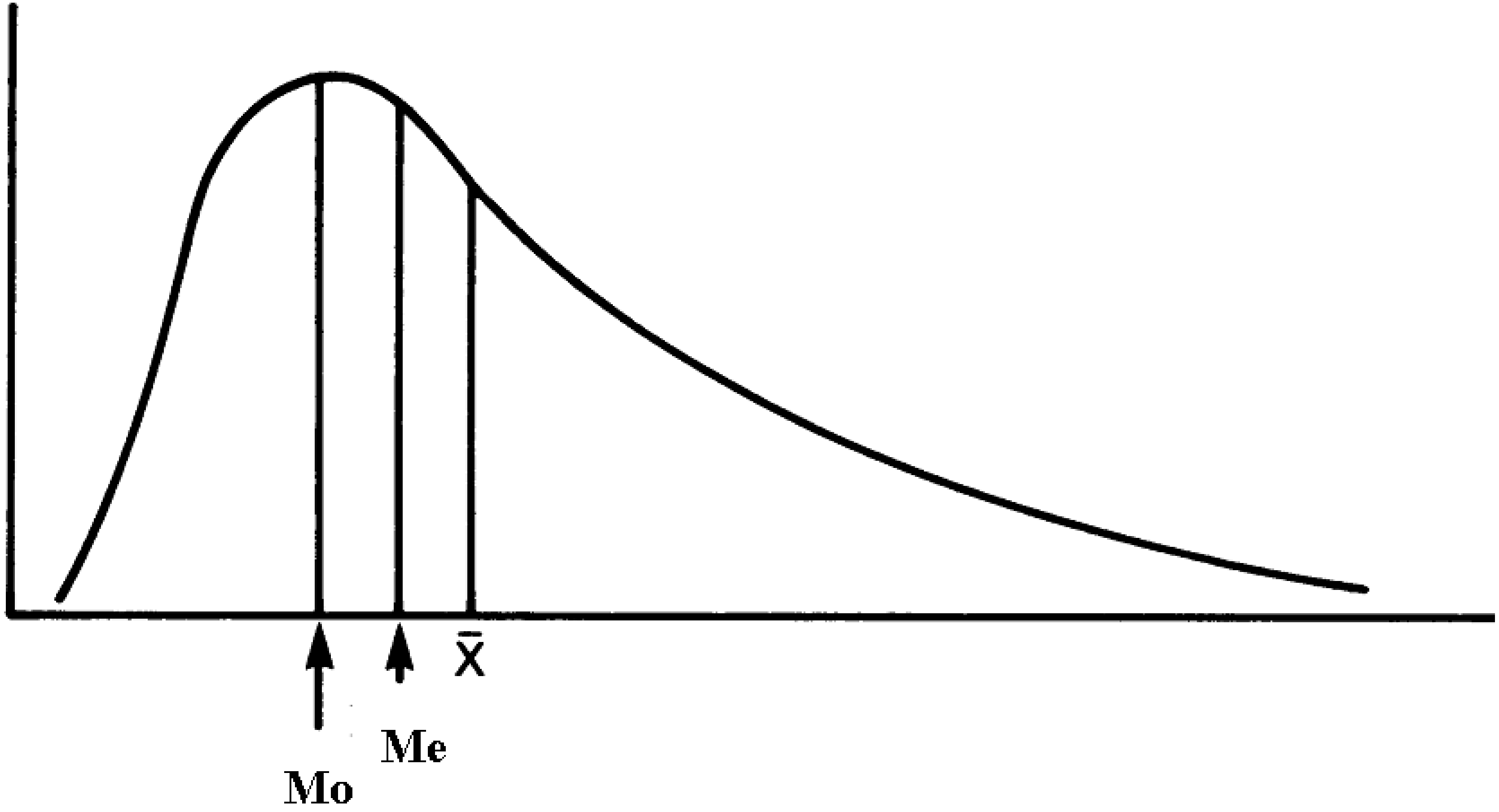


ASYMETRIC DISTRIBUTIONS

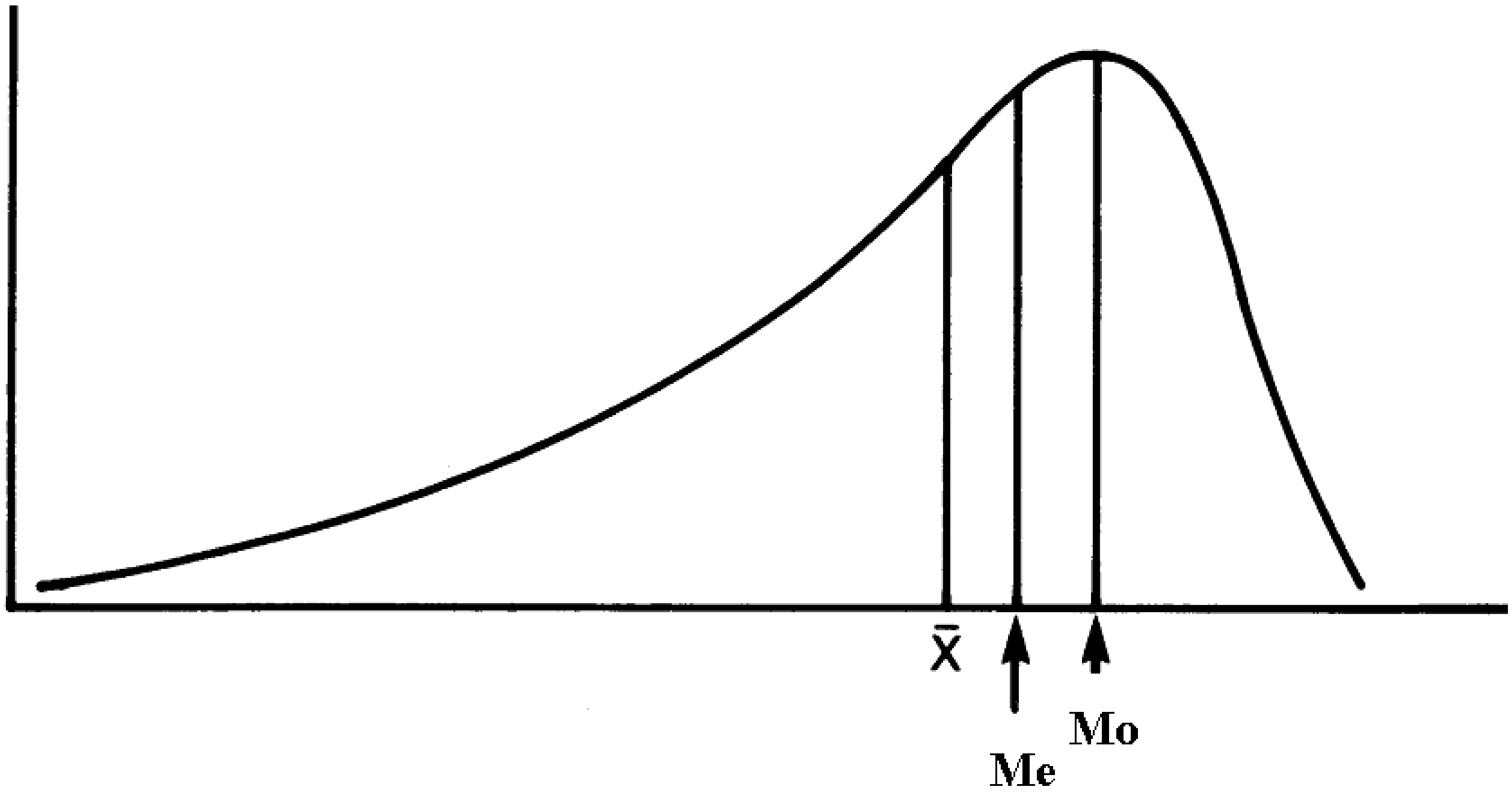
Regarding the inclination of the peak or skewness:

- ***positive skewness*** – distributions with an extended right hand tail (lower values more likely);
- ***negative skewness*** – distributions with an extended left hand tail (higher values more likely).

- **POSITIVELY SKEWED - with most of the scores being low, but with some scores spreading out towards the upper end of the distribution; the tail is directed to the right or to the positive side of the distribution**
- **mode<median<mean**


$\bar{X}$
Me
Mo

- **NEGATIVELY SKEWED - with most of the scores being high, but with some scores spreading out towards the lower end of the distribution; the tail is directed to the left or negative side of the distribution;**
- **mean<median<mode**


$\bar{X}$
Me
Mo

Important:

The type of the distribution determines the statistical tests to be used for descriptive or inferential statistics.

Part 2

SIPMLE DESCRIPTIVE STATISTICS FOR QUALITATIVE (CATEGORICAL) DATA

SIMPLE DESCRIPTIVE STATISTICS FOR CATEGORICAL DATA

1. Ratios - statistics which express the relative frequency of one set of frequencies, A, in relation to another, B.

- $\text{Ratio} = \dfrac{A}{B}$ e.g. $M/F=10/20=0.5$

Ratios are useful in the health sciences when we are interested in the distribution of illnesses or symptoms or the categories of subjects requiring or benefiting from some treatment.

SIMPLE DESCRIPTIVE STATISTICS FOR CATEGORICAL DATA

2. Proportions - **the frequency of one category over that of the total numbers in the sample or the population**

$$\text{Proportion of A} = \frac{A}{A + B} = M/M+F=0.33$$

3. Percentages - **the same as the proportions but multiplied by 100; in the above example - 33%**

SIMPLE DESCRIPTIVE STATISTICS FOR CATEGORICAL DATA

4. Rates - they are used in epidemiology to represent the level at which a disorder is present in a given population.

Incidence rate - represents the number of new cases of a disorder reported within a time period.

$$IR = \frac{\text{Number of new cases of a disorder}}{\text{total population at risk of the disorder}} \times \text{base}$$

SIMPLE DESCRIPTIVE STATISTICS FOR CATEGORICAL DATA

Prevalence rate - represents the total number of cases suffering from a particular disorder.

$$PR = \frac{\text{number of existing cases of a disorder}}{\text{total population at risk of the disorder}} \text{ x base}$$

The base for transforming the rates depends on the magnitude of the rates, conventionally 1000, 10000, 100000.

Important rule: The rarer the event, the bigger multiplier is used.

Test examples

1. ***z scores express how many standard deviations a particular score is from the mean.***

   A. True **B.** False

2. ***The total area under the standard normal curve is always 1.0.***

   A. True **B.** False

3. ***The area of a normal curve between any two designated z scores expresses the proportion or percentage of cases falling between the two points.***

   A. True **B.** False

4. ***About 10% of scores fall 3 standard deviations above the mean***.

   A. True **B.** False

31.10.2019 г.

5. ***50% of scores fall between z = 0.5 and z = - 0.5.***

A. True **B.** False

6. ***In a normal curve, approximately 34% of the scores fall between z = 0 and z = - 1.***

A. True **B.** False

7. ***Numerous human characteristics are distributed approximately as a normal curve.***

A. True **B.** False

8. The height of the rectangle in a histogram is proportional to class frequency and class width.

A. True **B.** False

31.10.2019 г.

9. Which of the following statements is true?

A. A z score indicates how many standard deviations a raw score is above or below the mean.

B. The mean of a standard normal distribution is always 0 (zero).

C. All the above statements are true.

10. In an anatomy test, your result is equivalent to z score of - 0.2. What does this z score imply?

A. You performed very well when compared to others.

B. Your result was slightly above average.

C. Your result was slightly below average.

11. State whether the data reflecting the age at death of individuals in the general population are likely to be skewed to the right, skewed to the left or symmetrical.

A. Symmetrical.

B. Skewed to the right (positively skewed)

C. Skewed to the left (negatively skewed)

12. Select the statement which you believe to be true. The Normal distribution:

A. Is a family of distributions which can have a variety of means and standard deviations.

B. Is the distribution of a variable measured on healthy individuals.

C. Has a mean of zero and a standard deviation of one.

D. Is skewed to the right.

13. Frequency distribution is another expression for a bar chart.

A. True B. False

14. A histogram can be used instead of a pie chart to display categorical data.

A. True B. False

15. A histogram Is similar to a bar chart but there are no gaps between the bars.

A. True B. False

16. A histogram can be used to display either a frequency or a relative frequency distribution.

A. True B. False

17. A histogram Is used to show the relationship between two variables.

A. True **B.** False

18. A bar chart Is used to display categorical data.

A. True **B.** False

19. A bar chart can only be used to display data which have a symmetrical distribution.

A. True **B.** False

20. A bar chart contains separate bars, with the length of each bar being proportional to the relevant frequency or relative frequency.

A. True **B.** False

21. Select all of the following type(s) of figures that would be appropriate for illustrating the distribution of heights of children in a class.

A. Histogram

B. Box-plot

C. Dot-plot

D. All listed types of figures are appropriate

22. Select all of the following type(s) of figures that would be appropriate for illustrating the distribution of blood groups in a sample of adults.

A. Bar chart

B. Pie chart

C. Both types are appropriate

31.10.2019 г.

23. Select all of the following type(s) of figures that would be appropriate for illustrating the number of fruit and vegetable portions consumed in a week by the 60 first year medical students in a medical school.

A. Bar

B. Pie

C. Box-plot

D. Dot-plot

E. All listed figures would be appropriate

24. State whether the data reflecting the salaries of all employees in an industrial company are likely to be skewed to the right, skewed to the left or symmetrical.

A. Skewed to the right

B. Skewed to the left

C. Symmetrical

25. State whether the data reflecting the heights of individuals in the general population are likely to be skewed to the right, skewed to the left or symmetrical.

A. Skewed to the right

B. Skewed to the left

C. Symmetrical

26. State whether the data reflecting the degree of flexion in a knee joint, expressed as a percentage of the maximum possible flexion in the joint are likely to be skewed to the right, skewed to the left or symmetrical.

A. Skewed to the right

B. Skewed to the left

C. Symmetrical

27. State whether the data reflecting the number of visits to a GP made in a year by individuals living in one particular region are likely to be skewed to the right, skewed to the left or symmetrical.

A. Skewed to the right

B. Skewed to the left

C. Symmetrical

28. Select all of the following statements which you believe to be true. The Normal distribution:

A. Has its mean equal to its median.

B. Is a continuous probability distribution.

C. Both statements are true.

31.10.2019 г.

29. Select all of the following variables that are likely to follow a Normal distribution.

A. Heights of individuals in the population.

B. The number of hospital attendances in a year in a sample of adults from the general population.

C. The ages of first year medical students.

30. Which distribution is the proportion of individuals with a disease who are successfully treated with a new drug likely to follow?

A. Normal distribution

B. Binominal distribution

C. None of the two

31.10.2019 г.

Answers:

1-A
2-A
3-A
4-B
5-B
6-A
7-A
8-A
9-C
10-C
11-C
12-A
13-B
14-B
15-A
16-A
17-B
18-A
19-B
20-A
21-C
22-C
23-E
24-A
25-C
26-B
27-A
28-C
29-A
30-B