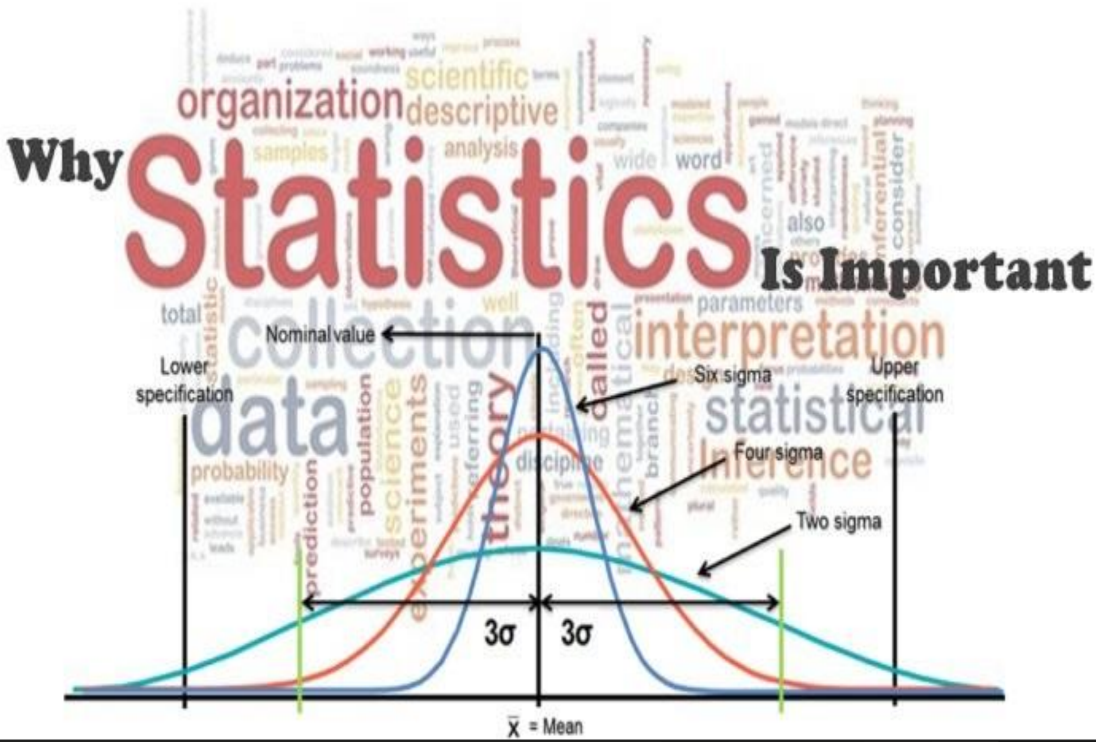
Salahaddin University- Erbil
College of Administration and Economics
Department of Statistics and Informatics



Statistical Method

First Stage
Second Semester

Lecturer: Shakar Maghdid Azeez
Email: shaker.azeez@su.edu.com

Years: 2023-2024

Chapter One “1“
Measures of central tendency
پێوەرەکانی ڕووکردنەچەق

Definition: A measure of central tendency is a value at the center or middle of a data set this value represents all data of the group.

مەبەست لە پێوەرەکانی ڕووکردنەچەق بریتیە لە وەسفکردنی چەند پێدراوێك یان دیاردەیەك بە تاکە ژمارەیەك.

There are some important types of measures of central tendency such as.

1. Arithmetic mean (Average, Mean). ناوەندە ژمێرەیی
2. Harmonic mean. ناوەندە هاوکۆکی
3. Quadratic mean. ناوەندە دووجایی
4. Geometric mean. ناوەندە ئەندازەیی
5. Mode. باو
6. Median. ناوەراست

1. Arithmetic mean (Average, Mean): ناوەندە ژمێرەیی

Definition: The arithmetic mean (generally called mean) is the sum of values divided by the total number of values. The symbol $\bar{x}$ (read "X bar") represents the sample mean and μ represents the population mean.

A. Arithmetic mean for ungrouped data ناوەندە ژمێرە بۆ داتا نارێزکراوەکان

The mean for a population consisting N observations is:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

The mean of sample for (n) observation (values) $x_1, x_2, x_3, \dots\dots\dots, x_n$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Where: x_i = the i^{th} observation,
n = the size of the data.

Example:
Find the ***Mean*** of following data (weight): -
45, 54, 36, 61, 27, 44, 73, 48

Solution:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8}{8}$$
$$= \frac{45+54+36+61+27+44+73+48}{8} = \frac{388}{8} = 48.5$$

B. Arithmetic mean for grouped data ناوەندە ژمێرە بۆ داتا ڕێزکراوەکان

If $x_1, x_2, x_3, \dots, x_n$ are the center of classes and $f_1, f_2, f_3, \dots, f_n$ are the frequency of data then

ئەگەر هاتوو خشتەی دووبارەیمان هەبوو ناوەندە ژمێرە بەپێ ئەو یاسایەی خوارەوە دەدۆزینەوە

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

Where: f_i = the frequency in the i^{th} class,
x_i = the center of classes in the i^{th} class.

Steps of Method mean from grouped data:

بۆ دوزینەوەی ناوەندە ژمێرە ئەم هەنگاوانە جی بەجی دەکەین

1. Find the total of frequency. $\sum f_i$ کۆی دووبارەکان دەدۆزینەوە
2. Find the center of classes by $\frac{U.L+L.L}{2}$. ناوەندە توێژ دەدۆزینەوە
3. Multiply center of classes by frequency. جارانی ناوەندە توێژ لەگەل دووبارەکان
4. Divided multiply by set of frequency.

Example:

Calculate ***Mean*** for the following frequency distribution:

Classes	f_i
2-3	3
4-5	1
6-7	2
8-9	1
10-11	1

Solution:

Classes	f_i	x_i	$f_i x_i$
2-3	3	2.5	7.5
4-5	1	4.5	4.5
6-7	2	6.5	13
8-9	1	8.5	8.5
10-11	1	10.5	10.5
	$\sum_{i=1}^{n} f_i = 8$		$\sum_{i=1}^{n} f_i x_i = 44$

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{44}{8} = 5.5$$

Homework 1/ Find the Arithmetic mean?

Classes	Fi
10 _20	2
20_30	4
30_40	6
40_50	8
50_60	5

Homework 2: A student from Statistics & Informatics department measured her/his waiting time (in minutes) for the Azadi bus on 30 mornings, and obtained the following results:

Classes (Time taken)	Frequency (fi)
1 –	3
5 –	6
9 –	10
13 –	7
17 – 21	4
Total	30

Find the average waiting time.

Merits of the Arithmetic Mean:

1. It is rigidly defined by algebraic formula;
2. It is easy to calculate and simple to understand;
3. It is based upon all the observations;
4. It is capable of further algebraic treatments and hence it is widely used in statistical analysis;
5. It is affected least by fluctuations of sampling.

Demerits of the Arithmetic Mean:

1. It can neither be determined by inspection or by graphical location;
2. It cannot be obtained if a single observation is missing or lost;
3. It is affected very much by extreme values (**outliers**);
4. It cannot be computed for **qualitative data** like data on religious affiliation, Gender and Level of education, etc;
5. It cannot be computed when class intervals have open ends.

C. Weighted mean for ungrouping data.: ($\bar{x}_w$) ناوەندە ژمێرەی کێشکراو

زۆر جار دەبینین هەندیك گوراو گرنگیەکی کێشی (وزن) هەیە لە چاو گۆراوەکانی تر بۆیە ئەگەر هاتوو ئەو (وزن)ی هەژمار نەکەین لەکاتی دۆزینەوەی ناوەندە ژمێرە ئەوا ئەنجامەکەمان زۆر وورد دەرناچێت بۆیە یاسای ناوەندە ژمێرەی کێشکراو بەکاردەهێنرێت

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_n w_n}{w_1 + w_2 + \ldots + w_n}$$

where: w_i = the weight of i^{th} observation.

Example:

In the Final exam's degrees of student the following

Degree: 62 80 75 88 84 86 90

Unit: 2 2 2 3 3 3 3

Find the **Weighted mean**.

Solution:

x_i (نمرە)	w_i (ژمارەی کاتژمێر)	$x_i w_i$
62	2	124
80	2	160
75	2	150
88	3	264
84	3	252
86	3	258
90	3	270
	$\sum_{i=1}^{n} w_i$=18	$\sum_{i=1}^{n} x_i w_i$=1478

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} = \frac{1478}{18} = 82.11$$

The formula for computing weighted arithmetic mean in case of **grouped data** is:

$$\bar{x}_w = \frac{w_1 f_1 x_1 + w_2 f_2 x_2 + w_3 f_3 x_3 + \dots + w_n f_n x_n}{w_1 f_1 + w_2 f_2 + w_3 f_3 + \dots + w_n f_n} = \frac{\sum_{i=1}^{n} w_i f_i x_i}{\sum_{i=1}^{n} w_i f_i}$$

Example: Find the mean from frequency table

classes	fi	Wi
2_	4	6
4_	5	5
6_	6	6
8_	3	4
10_12	2	4

Solution:

classes	fi	Wi	x_i	$w_i f_i$	$w_i x_i f_i$
2_	4	6	3	24	72
4_	5	5	5	25	125
6_	6	6	7	36	252
8_	3	4	9	12	108
10_12	2	4	11	8	88
				$\sum w_i f_i = 105$	$\sum x_i w_i f_i = 645$

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i w_i f_i}{\sum_{i=1}^{n} w_i f_i} = \frac{645}{105} = 6.143$$

properties of the Arithmetic mean

1. The sum of deviations of a set of number from their arithmetic mean is zero.

$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ (ungrouping data)

$\sum_{i=1}^{n} f_i(x_i - \bar{x}) = 0$ (grouping data)

prove/

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - n\bar{x}$$

$$= \sum_{i=1}^{n} x_i - n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)$$

$$= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i$$

$$= 0$$

prove/

$$\sum_{i=1}^{n} f_i(x_i - \bar{x}) = \sum_{i=1}^{n} f_i x_i - \bar{x} \sum_{i=1}^{n} f_i$$

$$= \sum_{i=1}^{n} f_i x_i - \left(\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}\right) \sum_{i=1}^{n} f_i$$

$$= \sum_{i=1}^{n} f_i x_i - \sum_{i=1}^{n} f_i x_i$$

$$= 0$$

2. The Arithmetic mean of sum of two variables value sum of two arithmetic mean of two variable.

If $Z_i = X_i + Y_i$ then $\bar{Z} = \bar{X} + \bar{Y}$

prove/

$$Z_i = X_i + Y_i$$

$$\sum Z_i = \sum X_i + \sum Y_i \quad \Big] \div n$$

$$\frac{\sum Z_i}{n} + \frac{\sum X_i}{n} = \frac{\sum Y_i}{n}$$

$$\bar{Z} = \bar{X} + \bar{Y}$$

2. Harmonic mean: ($\bar{H}$) ناوەندە هاوکۆکی

Harmonic mean is inverse of arithmetic mean of inverse value.

a. Ungrouped data

$$\bar{H} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

b. Grouped data

$$\bar{H} = \frac{\sum_{i=1}^{n} fi}{\sum_{i=1}^{n} \frac{fi}{xi}} = \frac{f1 + f2 + \cdots + fn}{\left(\frac{f1}{x1}\right) + \left(\frac{f2}{x2}\right) + \cdots + \left(\frac{fn}{xn}\right)}$$

x_i : Center of Classes.
f_i : Frequency.

Example:
Find **Harmonic mean** from the frequency data.

x_i= 2, 5, 3, 4, 7, 8, 8

Solution:

x_i	$\frac{1}{xi}$
2	1/2=0.5
5	1/5=0.2
3	1/3=0.33
4	1/4=0.25
7	1/7=0.14
8	1/8=0.13
8	1/8=0.13
	$\sum_{i=1}^{n} \frac{1}{xi} = 1.68$

$$\bar{H} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{7}{1.68} = 4.17$$

Example: Find Harmonic mean from the frequency data.

Classes	Fi
50 – 60	8
60 – 70	10
70 – 80	16
80 – 90	14
90 – 100	10
100 – 110	5
110 –120	2

Solution:

Classes	Fi	xi	$\frac{fi}{xi}$
50 – 60	8	55	8/55=0.145
60 – 70	10	65	10/65=0.154
70 – 80	16	75	16/75=0.214
80 – 90	14	85	14/85=0.165
90 – 100	10	95	10/95=0.105
100 – 110	5	105	5/105=0.048
110 –120	2	115	2/115=0.017
	$\sum_{i=1}^{n} fi$=65		$\sum_{i=1}^{n} \frac{fi}{xi}$=0.847

$$\overline{\boldsymbol{H}} = \frac{\sum_{i=1}^{n} fi}{\sum_{i=1}^{n} \frac{fi}{xi}} = \frac{65}{0.847} = 76.741$$

Prove that/

$\bar{X}.\bar{H} = x_1\ \ x_2$

Solution/

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2}{2}$$

$$\bar{H} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{2}{\frac{1}{x_1} + \frac{1}{x_2}} = \frac{2}{\frac{x_1 + x_2}{x_1 x_2}}$$

$$\bar{X}.\bar{H} = \frac{x_1 + x_2}{2} \cdot \frac{2}{\frac{x_1 + x_2}{x_1\ x_2}} = x_1\ x_2$$

3. Quadratic mean: ($\bar{Q}$) ناوەندە دووجایی

Quadratic mean is quadratic root of arithmetic mean of values squares.

a. Ungrouped data:

$$\bar{Q} = \sqrt{\frac{\sum_{i=1}^{n} x_i^{2}}{n}}$$

b. Grouped data

$$\bar{Q} = \sqrt{\frac{\sum_{i=1}^{n} f_i x_i^{2}}{\sum_{i=1}^{n} f_i}}$$

Example: Find **Quadratic mean** from the following values:
$x_i = 2,\ 4,\ 6,\ 8,\ 5$
Solution:

$$\bar{Q} = \sqrt{\frac{\sum_{i=1}^{n} x_i^{\,2}}{n}}$$

$$= \sqrt{\frac{2^2 + 4^2 + 6^2 + 8^2 + 5^2}{5}}$$

$$= \sqrt{\frac{4 + 16 + 36 + 64 + 25}{5}}$$

$$= \sqrt{\frac{145}{5}} = \sqrt{29} = 5.39$$

Example: From following data find **Quadratic mean**.

Classes	f_i
1 – 3	3
4 – 6	1
7 – 9	2
10 - 12	1
13 – 15	1

Solution:

Classes	f_i	x_i	$x_i^{\,2}$	$f_i x_i^{\,2}$
1 – 3	3	2	4	12
4 – 6	1	5	25	25
7 – 9	2	8	64	128
10 - 12	1	11	121	121
13 – 15	1	14	196	196
	$\sum_{i=1}^{n} f_i = 8$			$\sum_{i=1}^{n} f_i x_i^{\,2} = 482$

$$\bar{Q} = \sqrt{\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i}} = \sqrt{\frac{482}{8}} = \sqrt{60.25} = 7.76$$

4. *Geometric Mean: (G.M)*

The geometric mean of statistical data is defined as the (nth) root of the product of all the n values of the variable.

A.) Geometric Mean for ungrouped data:

$$G.M = \sqrt[n]{\prod_{i=1}^{n} x_i} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

$$G.M = e^{(\frac{1}{n}\sum_{i=1}^{n} Ln(X_i))}$$

Example: Find the Geometric mean of the following data:
(3, 7, 9, 4, 6, 10, 20).

Sol//

$$G.M = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[7]{\prod_{i=1}^{7} x_i} = \sqrt[7]{(3\times 7\times 9\times 4\times 6\times 10\times 20)}$$

$$= \sqrt[7]{907200} = 7.097$$

$Or\ /$

$$G.M = e^{(\frac{1}{n}\sum_{i=1}^{n} LnX_i)}$$

$$= (\frac{1}{7}(\text{Ln}\,3 + \text{Ln}\,7 + \text{Ln}\,9 + \text{Ln}\,4 + \text{Ln}\,6 + \text{Ln}\,10 + \ln 20))$$

$$= e^{(\frac{1}{7}(1.99 + 1.95 + 2.19 + 1.38 + 1.79 + 2.30 + 2.99))}$$

$$= e^{(\frac{1}{7}(13.72))} = e^{(1.96)} = 7.097$$

B.) Geometric Mean for grouped data:

$$G.M = e^{\left(\frac{1}{\sum f_i}\sum_{i=1}^{n} f_i LnX_i\right)}$$

Example: Find the Geometric mean for the frequency distribution:

classes	1-	3-	5-	7-	9-	11-13
fi	5	7	15	4	2	1

Sol //

classes	Fi	Xi	Ln Xi	fi Ln Xi
1-	5	2	0.6931	3.4657
3-	7	4	1.3863	9.7041
5-	15	6	1.7918	26.876
7-	4	8	2.0794	8.3178
9-	2	10	2.3026	4.6052
11-13	1	12	2.4849	2.4849
	34			55.454

$$G.M = e^{\left(\frac{1}{\sum f_i}\sum_{i=1}^{n} f_i LnX_i\right)}$$

$$G.M = e^{\left(\frac{1}{34}\times 55.454\right)} = e^{(1.631)} = 5.11$$

5. The Mode: (Mo) باو

The mode is the measurement that occurs most frequently in the data set.

باو بریتیە لەو بەهایەی کە زۆر دووبارە دەبێتەوە لە نێو کۆمەڵە بەهایەك

a. Ungrouped data. باو بۆ داتا نارێزکراوەکان

باو= زۆرترین بەهای دووبارووبونەوە

Example: Find the mode from a following data:

Xi =2, 4, 2, 2, 2, 7, 2 mode = 2

Xi = 2, 2, 7, 2, 4, 4, 3, 4 mode = 2, 4

Xi = 2, 2, 2, 7, 7, 7, 4, 4, 4 no mode Or mode= 2,7,4

b. Grouped data. باو بۆ داتا رێزکراوەکان

1. Discrete: ئەگەر هاتوو جۆری داتاکەمان لە جۆری پچراو بوو

بۆ دۆزینەوەی باو ناوەندە توێژ دەدۆزینەوە پاشان باو دەکاتە بەهای ئەو ناوەندە توێژەی کە دەکەوێتە بەرامبەر زۆرترین دووبارەبوونەوە لە دابەشکردنەکە.

Example: Find **Mode** for the following frequency distribution

Classes	fi
60_74	4
75_89	7
90_104	14
105_119	9
120_134	5

Solution:

Classes	fi	Xi
60_74	4	67
75_89	7	82
90_104	14	97
105_119	9	112
120_134	5	127

Mo=97

2. Continuous ئەگەر هاتوو جۆری داتاکان بەردەوام بوو

سەرەتا تویژی باو دەدۆزینەوە کە بریتیە لەو تویژەی زۆرترین دووبارەیی هەیە ئینجا ئەو یاسایەی خوارەوە بەکاردەهێنین

You calculated (Mo) in grouped data by:

$$\textbf{Mo} = l_k + \left(\frac{\Delta 1}{\Delta 1+\Delta 2}\right) \times w$$

l_k = lower limit of mode classes. نزمترین ڕادەی توێژی باو

$\Delta 1$= different of mode frequency over frequency of next lower classes.

جیاوازی یەکەم = (دوبارەی توێژی باو − دووبارەی پێشوو)

$\Delta 2$= different of mode frequency over frequency of next higher classes.

جیاوازی دووەم = (دوبارەی توێژی باو − دووبارەی دواتر)

$w =$ Width of mode all class. درێژی توێژی باو

A. Equal Length of Classes:

Example: Find **Mode** for the following frequency distribution.

Classes	fi
10_20	10
20_30	15
30_40	30
40_50	12
50_60	8
60_70	6

Solution:

	Classes	fi	
	10_20	10	
	20_30	15	→ دووبارەی پێشوو
l_k توێژی باو →	30_40	30	→ گەورەترین دووبارەبونەوە
	40_50	12	→ دووبارەی دواتر
	50_60	8	
	60_70	6	

l_k = 30
Δ1 = 30-15 =15
Δ2 = 30-12 =18
W = 10

$$Mo = l_k + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2}\right) * w$$

$$Mo = 30 + \left(\frac{15}{15+18}\right) * 10$$

$$= 34.54$$

B. Not Equal Length of Classes:

$$Average\ Frequency = \frac{Frequency}{Width\ of\ Classes}$$

Example: Find **Mode** for the following frequency distribution.

Classes	fi
5-10	2
10-15	4
15-25	10
25-35	22
35-50	27
50-60	11

Solution:

Classes	fi	Width of Classes	Average Frequency
5-10	2	5	2/5=0.4
10-15	4	5	4/5=0.8
15-25	10	10	10/10=1
25-35	22	10	22/10=2.2
35-50	27	15	27/15=1.8
50-60	11	10	11/10=1.1

$l_k = 25$
$\Delta 1 = 2.2 - 1 = 1.2$
$\Delta 2 = 2.2 - 1.8$
$w = 35 - 25 = 10$

$$Mo = l_k + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2}\right) * w$$

$$Mo = 25 + \left(\frac{1.2}{1.2 + 0.4}\right) * 10$$

$$= 25 + 0.75 * 10 \quad = 32.5$$

Merits and Demerits of the Mode:

#	Merits	Demerits
1.	In many cases it can be found by inspection.	It is not based upon all the observations.
2.	It is not affected by extreme values.	It is not capable of further mathematical treatment.
3.	It can be calculated for distributions with open end classes.	It is much affected by sampling fluctuations.
4.	It can be located graphically (Histogram).	It is not always possible to find a clearly defined mode (2 modes,3 modes)
5.	It can be used for qualitative data.	

6. The median: (Me) ناوەراست

The median of a set number arranged in order of size in the middle values or the mean of the two middle values.

ناوەراستە بریتیە لەو بەهایەی کە دەکەوێتە ناوەراستەوە پاش ڕێزکردنی بەشێوەی بەرزبووە یان نزم بووە.

a. Ungrouped data: ناوەڕاست بۆ داتا نارێزکراوەکان

If $X_1, X_2, \ldots, X_n$ observation arranged in order from the smallest to the largest, if

1- داتاکان ڕێزدەکەین بەشێوەی نزم بووە یان بەرزبووە

2- ناوەراست دەدۆزینەوە بەپێی ئەم یاسایانە:

- If n is odd number $\left(\frac{n+1}{2}\right)$ th.

ئەگەر هاتوو (n)ژمارەی تاك بوو ئەوا ناوەراستە بریتیە لە

$$Me = \text{بەهای ژمارە } \left(\frac{n+1}{2}\right)$$

- If n is even number, we take the average of $\left(\frac{n}{2}, \frac{n}{2}+1\right)$ th.

ئەگەر هاتوو (n) ژمارەی جووت بوو ئەوا ناوەراستە بریتیە لە :

$$Me = \frac{\text{بەهای ژمارە } \left(\frac{n}{2}\right) + \text{بەهای ژمارە } \left(\left(\frac{n}{2}+1\right)\right)}{2}$$

Example: find the median of the series.

a= 5 ,8 ,1 ,9, 2

b= 8, 2, 11, 2, 10, 6, 4, 2

Solution: Arranged data

a= 1, 2, 5, 8, 9

$\frac{n+1}{2} = \frac{5+1}{2} = 3\ th$

$\therefore Me = 5$

b= 2, 2, 2, 4, 6, 8, 10, 11

$\frac{n}{2}$, $\frac{n}{2}+1=\frac{8}{2}$, $\frac{8}{2}+1=(4{,}5)th$

$$\therefore \text{Me} = \frac{4+6}{2} = 5$$

Home work : Find the **median** of the series .

1. 55 , 62 , 53 , 70 , 68 , 65 , 63 , 79 , 80
2. 80 , 82 , 76 , 84 , 87 , 63

b. **Grouped data** ناوەڕاست بۆ داتا ڕێزکراوەکان

1. Discrete variable: - ئەگەر هاتوو جۆری داتاکەمان لە جۆری پچڕاو بوو

1- خشتەی دووبارەیی کۆکراوەی بەرزبووە (ACF) دەدۆزینەوە.

2- ڕێزبەندی ناوەڕاست هەژماردەکەین: $(\frac{\sum_{i=1}^{n} fi}{2})$

3- ئەوا بۆ دۆزینەوەی ناوەڕاست ناوەندە تویژ دەدۆزینەوە بۆ توێژی ناوەراست $F_{k-1} < \frac{\sum_{i=1}^{n} fi}{2} < F_k$

Example: From this table find the median.

class	f_i	ACF (F_i)	X_i
2_4	6	6	3
5_7	9	15	6
8_10	12	27	9
11_13	20	47	12
14_16	14	61	15
17_19	11	72	18
20_22	8	80	21

$$\frac{\sum_{i=1}^{n} fi}{2} = \frac{80}{2} = 40$$

$$F_{k-1} < \frac{\sum_{i=1}^{n} fi}{2} < F_k$$

27 < 40 < 47

Me =12

Homework: From this table find the median.

Classes	f_i
100_119	3
120_139	7
140_159	14
160_179	20
180_199	18
200_219	12
220_239	6

2. Continuous variable: ئەگەر هاتوو جۆری داتاکان بەردەوام بوو

$$\mathbf{Me} = l_k + \left(\frac{\frac{\sum_{i=1}^{n} fi}{2} - F_{k-1}}{f_k}\right) w$$

l_k: lower limit of median class . نزمترین رادەی تویژی ناوەراست

w: width of median class. درێژی تویژی ناوەراست

f_k: Frequency of the median class. دووبارەیی تویژی ناوەراست

F_{k-1} : Ascending Cumulative Frequency before the median class

دووبارەیی کۆکراوەیی بەرزبووەی پێشوو بۆ توێژی ناوەراست.

Example:

The following table gives the weekly income ($) of 100 families find **Median**.

Classes	f_i	ACF (F_i)
100_120	3	3
120_140	7	10
140_160	14	24
160_180	20	44
180_200	18	62
200_220	12	74
220_240	6	80

Solution:

$$\frac{\sum_{i=1}^{n} fi}{2} = \frac{80}{2} = 40$$

$$F_{k-1} < \frac{\sum_{i=1}^{n} fi}{2} < F_k$$

$24 \quad < \quad 40 \quad < \underline{44}$

$l_k = 160$

$w = 20$

$$Me = l_k + \left(\frac{\frac{\sum fi}{2} - F_{k-1}}{f_k} \right) * w$$

$$= 160 + (\frac{40-24}{20}) * 20$$

$$= 160 + (0.8)20$$

$$= 160 + 16$$

$$= 176$$

Home work : From this table find the median .

Classes	f_i
2_5	6
5_8	9
8_11	12
11_14	20
14_17	14
17_20	11
20_23	8

Relation between (Mean, Mode and Median):

Mode = 3 Median – 2 Mean

$$M_o = 3M_e - 2\overline{X}$$

$\bar{x}$ = ?

Me =?

Example: If the mean and median of moderately asymmetrical series are 26.8 and 27.9 respectively, what would be its most probable mode?

Solution:

$$M_o = 3M_e - 2\overline{X}$$

= 3(27.9) – 2(26.8) = 30.1

Home Work

median = 20.6

mode = 26

Find mean.

Chapter Two "2"
Measures of Dispersion & Shape & Location

پێوەرەکانی پەرتبوون و شێوە و شوێن

- *Measures of Dispersion Variation* — *پێوەرەکانی پەرت و بلاّوی*
  - *Measures of Absolute Variation* — پێوەرەکانی بەربلاّوی رووت
  - *Measures of Relative Variation* — پێوەرەکانی بەربلاّوی رێژەیی
- *Measures of Shape* — پێوەرەکانی شێوە
- *Measures of Location* — پێوەرەکانی شوێن

Measures of Dispersion variation *پێوەرەکانی پەرت و بلاّوی*

Description or variation: is meaning near or for between observation value of any variables.

مەبە ست لە پەر شوبلاّوی بریتی یە لە دووری و نزیکی یەی هەیە لە نێوان کۆمەلێك داتای گۆڕاوێکی دیاریکراو، وە هەر چەندە پێوانەی پەر شوبلاّوی گەورە بێت ئەوا ئاماژە بەوە دەکات کە داتاکان لێکچوون نین وە بەپێچەوانەوەش راستە.

Types of Measures of Dispersion (variation)

A) Measures of absolute variation

Many measures are using to determine the variations, such as: Range, Quartile deviation, Mean deviation, Standard deviation, Variance. All these measures used when the variables have the same units of measurement.

B) Measures of relative variation

When the variables have different units of measurement, measures of relative variation are used. Coefficient of variation is the main measure.

Objectives of measures of dispersion

The main objective is to know the homogeneity of the values for a data set, or to compare between the values for two or more than two data set.

1- Range: (R) مەودا (ungrouped data)

The simplest measure of absolute variation is the range which calculated by subtracting the lowest value from the highest value of a data set.

مەودا سادەترین جۆری پێوەرەکانی پەر شوبلاًوییە، وە مەودا پێنا سەدەکرێت کە بریتی یە لە جیاوازی نێوان گەورەترین و بـ چووکترین بەها لە کۆمەلێك داتا ، ئەگەر X_L بریتی بێت لە گەورەترین بەها وە X_S بریتی بێت لە بچوکترین بەها لە کۆمەلێك داتا ئەو کاتە مەودا بۆ ئەم لە کۆمەلە بریتی یە لە:

R=highest value – lowest value

$$R = X_L - X_S$$

Example: Find the range for the following data.

1. 7, 8, 9, 10, 11
2. 3, 6, 9, 12, 15
3. 1, 5, 9, 13, 17

Solution:

1. $R = x_L - x_s = 11 - 7 = 4$
2. $R = x_L - x_s = 15 - 3 = 12$
3. $R = x_L - x_s = 17 - 1 = 16$

Home Work:

Find the range for the following data.

1) 53 , 55 , 40 , 63 , 65 , 68 , 70 , 79 , 80
2) 100 , 35 , 87 , 43 , 90 , 22 , 54 , 38 , 76 , 84 , 30 , 52 , 49 , 52
3) 83 , 65 , 70 , 20 , 31 , 19 , 52 , 99 , 42, 80 ,62

Solution:

1-

2-

3-

(Grouped data)

R = Upper limit last class – Lower limit first class

R = بەرزترین رادەی تویژی کۆتایی - نزمترین رادەی تویژی یەکەم

Example: Find the range for the following data.

Classes	Fi
10_20	2
20_30	4
30_40	8
40_50	4
50_60	2
60_70	2

Solution:

R =70 – 10 =60

2.Mean Deviation (M.D)

The (M.D.) of statistical data as the (A.M.) of number value of numerical value of deviation of items from average.

$$M.D = \frac{\sum_{i=1}^{n} |X_i - \overline{X}|}{n}$$ For ungroup data

$$M.D = \frac{\sum_{i=1}^{n} f_i |X_i - \overline{X}|}{\sum_{i=1}^{n} f_i}$$ For group data

Example2: find the (M.D) for the following data. 4, 6, 10, 12, 18

Soluion:

X_i	$\mid X_i - \bar{X} \mid$
4	|4-10|=6
6	|6-10|=4
10	|10-10|=0
12	|12-10|=2
18	|18-10|=8
$\sum_{i=1}^{n} X_i = 50$	$\sum_{i=1}^{n} \mid X_i - \bar{X} \mid = 20$

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{50}{5} = 10$$

$$M.D = \frac{\sum_{i=1}^{n} \left| X_i - \overline{X} \right|}{n} = \frac{20}{5} = 4$$

Example: Find the (M.D) for the following frequency distribution:

classes	0 -	10 -	20 -	30 -	40 -	50 - 60
fi	12	18	27	20	17	6

Solution:

Classes	f_i	X_i	f_iX_i	$\mid X_i - \bar{X}\mid=\mid X_i - 28\mid$	$f_i \mid X_i - \bar{X}\mid$
0_10	12	5	60	$\mid 5-28\mid=23$	12*23=276
10_20	18	15	270	$\mid 15-28\mid=13$	18*13=234
20_30	27	25	675	$\mid 25-28\mid=3$	27*3=81
30_40	20	35	700	$\mid 35-28\mid=7$	20*7=140
40_50	17	45	765	$\mid 45-28\mid=17$	17*17=289
50_60	6	55	330	$\mid 55-28\mid=27$	6*27=162
	$\sum f_i$ =100		$\sum f_iX_i$=2800		$\sum_{i=1}^{n} f_i \mid X_i - \bar{X}\mid$ =1182

$$\bar{x} = \frac{\sum fixi}{\sum fi} = \frac{2800}{100} = 28$$

$$M.D = \frac{\sum_{i=1}^{n} f_i \left|X_i - \bar{X}\right|}{\sum_{i=1}^{n} f_i} = \frac{1182}{100} = 11.82$$

3- Variance جیاکاری

It is one of the measures of absolute variation. The variance can be calculated by taking the average of the square of the distance from the mean for each value.

بەیەکێك لە پێوەرەکانی پەرشوبلاوی دادەنرێت لەبواری جێبەجێکاریدا زۆر بەکاردێت،وەبریتیە لە سەرجەمی دووجای لادانی داتاکان لە ناوەندە ژمێرە

The formula for the **population** variance (σ^2) for ungroup data is:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Where:
x= individual value
μ = population mean
N = population size.

Also the formula for the **sample** variance (S^2)

(Ungrouped data)

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

(Grouped data)

$$S^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \overline{x})^2}{\sum_{i=1}^{n} f_i - 1}$$

Where:

- $\overline{x}$ = sample mean
- n = sample size.

Properties of sample variance:

1) $\boldsymbol{S^2 \geq 0}$
2) If $\boldsymbol{y_i = ax_i \rightarrow S_Y^2 = a^2 . S_X^2}$, where **a** is a constant. **(Prove that)**
3) If $\boldsymbol{y_i = x_i + b \rightarrow S_Y^2 = S_X^2}$, where **b** is a constant. **(Prove that)**
4) If X and Y are independent variables and Z=X+Y, then the variance of Z is: $\boldsymbol{S_Z^2 = S_X^2 + S_Y^2}$

Example:

Find the **variance** of the following data.

5, 7, 9, 11, 13, 15, 17

Solution:

xi	(xi -x̄)	$(\text{xi} - \bar{\text{x}})^2$
5	5-11 = -6	36
7	7-11 = -4	16
9	9-11 = -2	4
11	11-11 = 0	0
13	13-11 = 2	4
15	15-11 = 4	16
17	17-11 = 6	36
$\sum xi$=77		$\sum_{i=1}^{n}(xi - \bar{x})^2 = 112$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

$$s^2 = \frac{\sum_{i=1}^{n}(xi - \bar{x})^2}{n-1} = \frac{112}{7-1} = \frac{112}{6} = 18.6$$

Example:

Find the **Variance** of the following data.
xi = 8, 2, 11, 2, 10, 6, 4, 2

Solution:

(Grouped data)

Example: find the **Variance** of following table.

Classes	fi
0_10	12
10_20	18
20_30	27
30_40	20
40_50	17
50_60	6

Solution:

Classes	fi	xi	fi xi	(xi-$\bar{x}$)	$(xi - \bar{x})^2$	$fi(xi - \bar{x})^2$
0_10	12	5	60	-23	529	6348
10_20	18	15	270	-13	169	3042
20_30	27	25	675	-3	9	243
30_40	20	35	700	7	49	980
40_50	17	45	765	17	289	4913
50_60	6	55	330	27	729	4374
	$\sum fi$ =100		$\sum fixi$ =2800			$\sum_{i=1}^{n} f_i(xi - \bar{x})^2$ =19900

$$\bar{x} = \frac{\sum fixi}{\sum fi} = \frac{2800}{100} = 28$$

$$s^2 = \frac{\sum_{i=1}^{n} f_i(xi - \bar{x})^2}{\sum_{i=1}^{n} fi - 1} = \frac{19900}{100 - 1} = 201.01$$

Example: find the **Variance** of following table.

Classes	f_i
2_5	6
5_8	9
8_11	12
11_14	20
14_17	14

Solution:

4- Standard deviation (S) لادانی پێوانەیی

Standard deviation is the most important and most widely used measure of absolute variation.

S : is the square root of variance . لادانی پێوانەیی بریتیە لەرەگی دووجای جیاکاری

The formula for the **population** Standard deviation (σ) for ungroup data is:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(xi-\mu)^2}{N}}$$

Where:
x= individual value
μ = population mean
N = population size.

Also the formula for the **sample** Standard deviation (S)

(Ungrouped data)

$$s = \sqrt{\frac{\sum_{i=1}^{n}(xi-\bar{x})^2}{n-1}}$$

(Grouped data)

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i (xi-\bar{x})^2}{\sum_{i=1}^{n} fi - 1}}$$

Example:

Find the **Standard deviation** of the following data.
xi = 5 , 7 , 9 , 11 , 13 , 15 , 17
Solution:

xi	(xi -$\bar{x}$)	$(xi - \bar{x})^2$
5	5-11 = -6	36
7	7-11 = -4	16
9	9-11 = -2	4
11	11-11 = 0	0
13	13-11 = 2	4
15	15-11 = 4	16
17	17-11 = 6	36
$\sum xi$=77		$\sum_{i=1}^{n}(xi - \bar{x})^2 = 112$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}(xi - \bar{x})^2}{n-1}} = \sqrt{\frac{112}{7-1}} = 4.32$$

Example:

Find the **Standard deviation** of the following data.
xi = 8, 2, 11, 2, 10, 6, 4, 2

Solution:

(Grouped data)

Example: find the **Standard deviation** of following table.

Classes	fi
0_10	12
10_20	18
20_30	27
30_40	20
40_50	17
50_60	6

Solution:

Classes	fi	xi	fi xi	$(xi-\bar{x})$	$(xi-\bar{x})^2$	$fi(xi-\bar{x})^2$
0_10	12	5	60	-23	529	6348
10_20	18	15	270	-13	169	3042
20_30	27	25	675	-3	9	243
30_40	20	35	700	7	49	980
40_50	17	45	765	17	289	4913
50_60	6	55	330	27	729	4374
	$\sum fi$ =100		$\sum fixi$ =2800			$\sum_{i=1}^{n} f_i(xi-\bar{x})^2$ =19900

$$\bar{x} = \frac{\sum fixi}{\sum fi} = \frac{2800}{100} = 28$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i(xi-\bar{x})^2}{\sum_{i=1}^{n} fi-1}} = \sqrt{\frac{19900}{100-1}} = \sqrt{201.01} = 14.18$$

Home Work: Find **Standard deviation** for the following frequency distribution.

Classes	fi
2_5	6
5_8	9
8_11	12
11_14	20
14_17	14

Solution:

5. Coefficient of variation (C.V) معامل الاختلاف

The (C.V) is the (S) divided by the mean. The result is expressed as a percentage .it is use for compare two or more series for variability.

بریتیه لەو پێوەرەی کەپشت نابەستێت بە یەکەکان وەبلاوترین پێوەری جیاوازیە کە بریتیە لە دابەشی لادانی پێوانەیی لەسەر ناوەندە ژمێرە

The formals of (C.V) are as follow.

Sample	Population
$C.V. = \frac{S}{\bar{x}} . 100$	$C.V. = \frac{\sigma}{\mu} . 100$

Example: find the C.V of following two data sets.

a.) 5, 7, 9, 11, 13, 15, 17 ,
b.) 5, 6, 7, 11, 15, 16, 17

Solution:

a. 5, 7, 9, 11, 13, 15, 17

C.V= $\frac{S}{\bar{x}} * 100$

xi	(xi -$\bar{x}$)	$(xi - \bar{x})^2$
5	5-11 = -6	36
7	7-11 = -4	16
9	9-11 = -2	4
11	11-11 = 0	0
13	13-11 = 2	4
15	15-11 = 4	16
17	17-11 = 6	36
$\sum xi$=77		$\sum_{i=1}^{n} (xi - \bar{x})^2 = 112$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}(xi - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{112}{7-1}} = 4.32$$

$$C.V = \frac{4.32}{11} * 100 = \%\ 39.27$$

b. 5, 6 , 7 , 11 , 15 , 16 , 17

C.V = $\frac{S}{\bar{x}} * 100$

xi	(xi -$\bar{x}$)	$(xi - \bar{x})^2$
5	5-11 = -6	36
6	6-11 = -5	25
7	7-11 = -4	16
11	11-11 = 0	0
15	15-11 = 4	16
16	16-11 = 5	25
17	17-11 = 6	36
$\sum xi$=77		$\sum_{i=1}^{n}(xi - \bar{x})^2 = 154$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}(xi - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{154}{7-1}} = 5.07$$

$$C.V = \frac{S}{\bar{x}} * 100$$

$$C.V = \frac{5.07}{11} * 100 = 46.09\%$$

The first data is most homogenous.

Example: find the **Coefficient of variation** of following table.

Classes	fi
0_10	12
10_20	18
20_30	27
30_40	20
40_50	17
50_60	6

Solution:

$$C.V = \frac{S}{\bar{x}} * 100$$

S=? $\bar{x}$ =?

Classes	fi	xi	fi xi	(xi-$\bar{x}$)	$(xi - \bar{x})^2$	$fi(xi - \bar{x})^2$
0_10	12	5	60	-23	529	6348
10_20	18	15	270	-13	169	3042
20_30	27	25	675	-3	9	243
30_40	20	35	700	7	49	980
40_50	17	45	765	17	289	4913
50_60	6	55	330	27	729	4374
	$\sum fi$ =100		$\sum fixi$ =2800			$\sum_{i=1}^{n} f_i(xi - \bar{x})^2$ =19900

$$\bar{x} = \frac{\sum fixi}{\sum fi} = \frac{2800}{100} = 28$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i (xi - \bar{x})^2}{\sum_{i=1}^{n} fi - 1}} = \sqrt{\frac{19900}{100-1}} = \sqrt{201.01} = 14.18$$

$$C.V = \frac{S}{\bar{x}} * 100$$

$$= \frac{14.18}{28} * 100 = \%\ 50.64$$

Exercises:

1) Prove that $S^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$

2) Find the **Coefficient of variation** of the following data.
xi = 8, 2, 11, 2, 10, 6, 4, 2

3) A data set consisting of 10 observations has a mean equal to zero and a variance equal to (d). Find $\sum x_i^2$.

4) Which of the two students is better from looking at their marks (homogeneity)?

Student I	90, 50, 50, 60, 95, 55
Student II	67, 73, 80, 70, 50, 55

5) Mean and (S.D) for marks of students have obtained in Math exam are 69 and 19.3 respectively, when the marks in Statistic exam have 75 and 25.5. Which of them are homogeneity in marks?

6) Find the **Variance** and **Coefficient of variation** for the following frequency distribution.

Classes	fi
10_20	10
20_30	15
30_40	30
40_50	12
50_60	8
60_70	6

7) Answer the following questions.

A- If the mean of five values is 64, find the sum of the values?

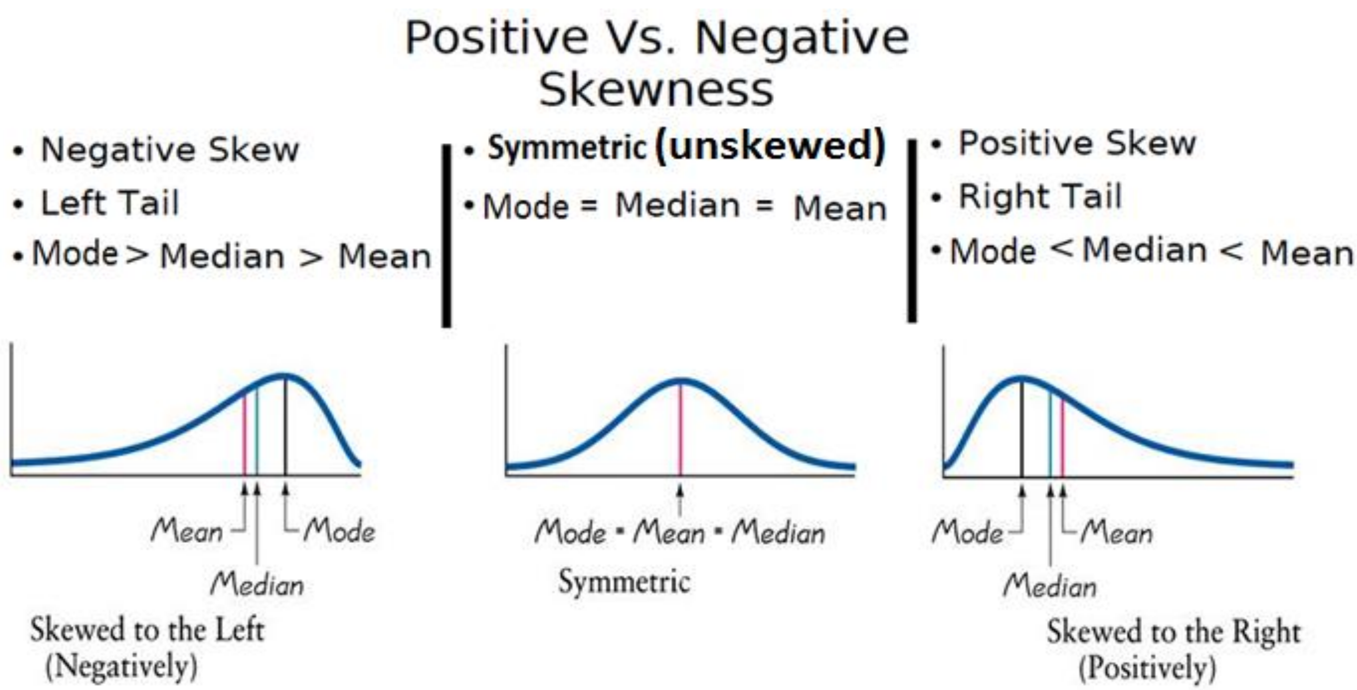
B- Why might the range not be the best estimate of variability?

C- If the mean of five values is 8.2 and four of the values are 6, 10, 7, 12, find the fifth value.

Measures of Shape
Measures of Skewness and Kurtosis

1- Skewness:

It is the degree of asymmetry or departure from symmetry of a distribution of data. If the frequency curve of a distribution has a longer tail to the right of the central maximum than to the left, the distribution is said to be skewed to the right, or to have positive skewness. If the reverse is true, it is said to be skewed to the left, or to have negative skewness. The word “skew” means a tail, so that distributions that have a large tail of outlying values on the right-hand-side or left-hand-side are called skewed.



A simple formula for skewness using Karl Pearson’s coefficient of skewnessis as follows:

$$C.S_k = \frac{(\bar{x} - Mo)}{S} = \begin{cases} - & Negative\ skewness \\ 0 & Symmetric \\ + & Positive\ skewness \end{cases}$$

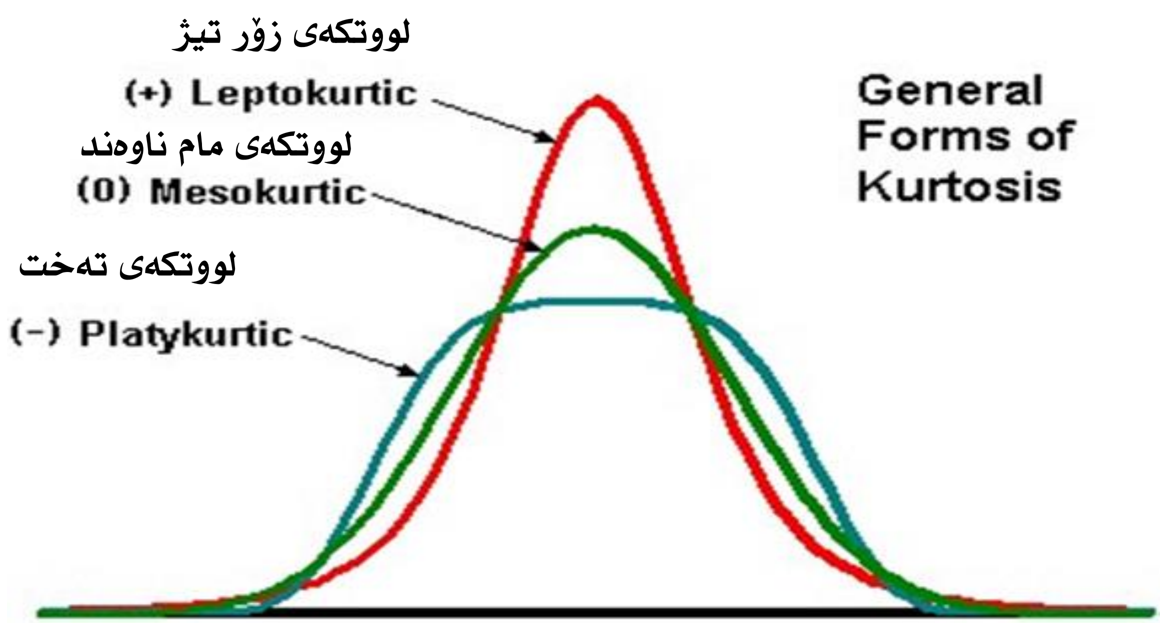
OR

$$C.S_k = \frac{3(\bar{x} - Me)}{S} = \begin{cases} - & Negative\ skewness \\ 0 & Symmetric \\ + & Positive\ skewness \end{cases}$$

2- Kurtosis:

It is the degree of peakedness of a distribution, usually taken relative to symmetric distribution. To find the kurtosis, it will be used the following formula:

$$Kurtosis = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}\right)^2} - 3 \quad = \begin{cases} - & Platykurtic \\ 0 & Mesokurtic \\ + & Leptokurtic \end{cases}$$



Example: **find the skewness and kurtosis** of the following 2 data sets:

a.) 5, 7, 9, 11, 13, 15, 17

b.) 5, 6, 7, 11, 15, 16, 17

Solution: Since median and mean for each data set (a, b) are equal to 11, then

a) xi *=5, 7, 9, 11, 13, 15, 17*

Skewness:

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

5, 7, 9, 11, 13, 15, 17

$$\frac{n+1}{2} = \frac{7+1}{2} = 4$$ القيمة الرقم

$$\therefore Me = 11$$

$$C.S_k = \frac{3(\bar{x} - Me)}{S}$$

$$= \frac{3(11-11)}{S} = 0 \quad \text{is Symmetric}$$

$$\textbf{Kurtosis} = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}\right)^2} - 3$$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

xi	(xi - $\bar{x}$)	$(xi - \bar{x})^2$	$(xi - \bar{x})^4$
5	5-11=-6	$(-6)^2$=36	$(-6)^4$=1296
7	7-11=-4	$(-4)^2$=16	$(-4)^4$=256
9	9-11=-2	$(-2)^2$=4	$(-2)^4$=16
11	11-11=0	$(0)^2$=0	$(0)^2$0
13	13-11=2	$(2)^2$=4	$(2)^4$=16
15	15-11=4	$(4)^2$=16	$(4)^4$=256
17	17-11=6	$(6)^2$=36	$(6)^4$=1296
$\sum xi$=77		$\sum_{i=1}^{n}(xi - \bar{x})^2$ =112	$\sum_{i=1}^{n}(xi - \bar{x})^2$ =3136

$$Kurtosis = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n})^2} - 3$$

$$= \frac{\frac{3136}{7}}{(\frac{112}{7})^2} - 3 = \frac{448}{16^2} - 3 = -1.25 \qquad is\ Platykurtic$$

***b) xi** =5, 6, 7, 11, 15, 16, 17*

Skewness:

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

5, 6, 7, 11, 15, 16, 17

$\frac{n+1}{2} = \frac{7+1}{2} = 4$ القيمة الرقم

$\therefore Me = 11$

$$C.S_k = \frac{3(\bar{x} - Me)}{S}$$

$$= \frac{3(11 - 11)}{S} = 0 \quad is\ Symmetric$$

$$\boldsymbol{Kurtosis} = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n})^2} - 3$$

$$\bar{x} = \frac{\sum xi}{n} = \frac{77}{7} = 11$$

xi	(xi - x̄)	$(xi - \bar{x})^2$	$(xi - \bar{x})^4$
5	5-11=-6	$(-6)^2$=36	$(-6)^4$=1296
6	6-11=-5	$(-5)^2$=25	$(-5)^4$=625
7	7-11=-4	$(-4)^2$=16	$(-4)^4$=256
11	11-11=0	$(0)^2$=0	$(0)^2$0
15	15-11=4	$(4)^2$=4	$(4)^4$=256
16	16-11=5	$(5)^2$=25	$(5)^4$=625
17	17-11=6	$(6)^2$=36	$(6)^4$=1296
$\sum xi$=77		$\sum_{i=1}^{n}(xi - \bar{x})^2$ =154	$\sum_{i=1}^{n}(xi - \bar{x})^2$ =4354

$$\boldsymbol{Kurtosis} = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n})^2} - 3$$

$$= \frac{\frac{4354}{7}}{(\frac{154}{7})^2} - 3 = \frac{622}{22^2} - 3 = -1.71 \qquad is\ Platykurtic$$

H.W

1) Calculate a Skewness and kurtosis-using form the data represents the length of 13 students in a high school (in centimeter).

(120, 130, 110, 125, 105, 130, 160, 140, 150, 175, 145, 210, 140)

2) Calculate a Skewness and kurtosis-using form the data represents the price (in dollars) for sample of round-trip flights from Erbil to Dubai.

Xi	872	432	397	427	388	782	397

Measures of Location پێوەرەکانی شوێن

The value of variable which partition a series to equal parts are called (partition values).

(Ungrouped data)

1. **Quartiles:** چواریەك

$Q_i = \text{size of } i\left(\frac{n}{4}\right)^{th} \text{ Value} \qquad i = 1,2,3$

2. **Deciles:** دەیەك

$D_i = \text{size of } i\left(\frac{n}{10}\right)^{th} \text{ Value} \qquad i = 1,2,3 \ldots \ldots,\ 9$

3. **Percentiles:** سەدیەك

$P_i = \text{size of } i\left(\frac{n}{100}\right)^{th} \text{ Value} \qquad i = 1,2,3, \ldots \ldots, 99$

Example :

Calculate part value ($Q_1, Q_3, D_5, D_7, P_{22}$)

(4, 7, 3 , 9 , 4, 12, 7, 12 ,6 ,18 ,19 ,20 ,4 ,7 ,6 ,4 ,17 ,16 ,19 ,22)

Solution : arrange data ascending

(3 , 4 , 4 , 4 , 4 ,6 ,6 ,7 ,7 ,7 ,9 , 12 ,12 ,16 ,17 , 18 , 19 , 19 , 20, 22) n=20

c. $Q_1 = i\left(\frac{n}{4}\right) = 1\left(\frac{20}{4}\right) = (5)^{th}$ Value = 4

2. $Q_3 = i\left(\frac{n}{4}\right) = 3\left(\frac{20}{4}\right) = (15)^{th}\text{Value} = 17$

3. $D_5 = i\left(\frac{n}{10}\right) = 5\left(\frac{20}{10}\right) = \frac{100}{10} = (10)^{th}$ Value=7

4. $D_7 = i\left(\frac{n}{10}\right) = 7\left(\frac{20}{10}\right) = \frac{140}{10} = (14)^{th}$ Value = 16

5. $P_{22} = i\left(\frac{n}{100}\right) = 22\left(\frac{20}{100}\right) = (4.4)^{th}$ Value = 4^{th} Value = 4

Home Work// Q_2? D_3? D_9? P_{50} ?

(Quartiles, Deciles and Percentiles for Grouped data) داتا ریزکراوەکان

Discrete variable: - ئەگەر هاتوو جۆری داتاکەمان لە جۆری پچراو بوو

To find the quartiles, or deciles, or percentiles we follow the same procedure to fined the median.

Step 1 Find an ascending cumulative frequency.

Step 2 If numbering arrangement of quartiles, deciles, and percentiles is fraction then its value is for the number greater than it, if true number the value is the mean of its and the greater numbers.

Example: Find the Q_1, Q_3 , D_3 , D_8 P_{28} *and* P_{65} for the following table.

class	f_i
2_4	6
5_7	9
8_10	12
11_13	20
14_16	14
17_19	11
20_22	8

Solution:

class	f_i	ACF (F_i)	X_i
2_4	6	6	3
5_7	9	15	6
8_10	12	27	9
11_13	20	47	12
14_16	14	61	15
17_19	11	72	18
20_22	8	80	21
	$\sum_{i=1}^{n} fi = 80$		

Position for $Q_1 = i\left(\frac{n}{4}\right) = \frac{80}{4} = 20$

$F_{k-1} < i\left(\frac{n}{4}\right) < F_k$

15 < 20 < <u>27</u>

$Q_1 = 9$

Position for $Q_3 = i\left(\frac{n}{4}\right) = 3(\frac{80}{4}) = 60$

$F_{k-1} < i\left(\frac{n}{4}\right) < F_k$

47 < 60 < <u>61</u>

$Q_3 = 15$

Position for $D_3 = i\left(\frac{n}{10}\right) = 3(\frac{80}{10}) = \frac{240}{10} = 24$

$F_{k-1} < i\left(\frac{n}{10}\right) < F_k$

15 < 24 < <u>27</u>

$D_3 = 9$

Position for $D_8 = i\left(\frac{n}{10}\right) = 8(\frac{80}{10}) = 64$

$F_{k-1} < i\left(\frac{n}{10}\right) < F_k$

61 < 64 < <u>72</u>

$D_8 = 18$

Position for $P_{28} = i\left(\frac{n}{100}\right) = 28(\frac{80}{100}) = 22.4$

$F_{k-1} < i\left(\frac{n}{100}\right) < F_k$

$15 < 22.4 < \underline{27}$

$P_{28} = 9$

Position for $P_{65} = i\left(\frac{n}{100}\right) = 65(\frac{80}{100}) = \frac{240}{100} = 52$

$F_{k-1} < i\left(\frac{n}{100}\right) < F_k$

$47 < 52 < \underline{61}$

$P_{65} = 15$

(Quartile s for Grouped data) داتا ڕیزکراوەکان

Continuous variable: ئەگەر هاتوو جۆری داتاکان بەردەوام بوو

$$Q_i = l_k + \left(\frac{i\left(\frac{n}{4}\right) - F_{k-1}}{f_k}\right) w$$

l_k: lower limit of quartile class .
w: width of quartile class.
f_k: Frequency of the quartile class.
F_{k-1} : Ascending Cumulative Frequency before the quartile class

Example:

The following table gives the weekly income ($) of 100 families find **Quartiles**.

Classes	f_i
100_120	3
120_140	7
140_160	14
160_180	20
180_200	18
200_220	12
220_240	6

Solution:

Classes	f_i	ACF (F_i)
100_120	3	3
120_140	7	10
140_160	14	24
160_180	20	44
180_200	18	62
200_220	12	74
220_240	6	80

Position for $Q_2 = i\left(\frac{n}{4}\right) = 2(\frac{80}{4}) = 40$

$F_{k-1} < i\left(\frac{n}{4}\right) < F_k$

$24 < 40 < \underline{44}$

$$Q_2 = l_k + \left(\frac{2\left(\frac{n}{4}\right) - F_{k-1}}{f_k}\right) w$$

$$= 160 + \left(\frac{40-24}{20}\right) 20 = 176$$

(Deciles for Grouped data) داتا ڕیزکراوەکان

Continuous variable: ئەگەر هاتوو جۆری داتاکان بەردەوام بوو

$$D_i = l_k + \left(\frac{i\left(\frac{n}{10}\right) - F_{k-1}}{f_k}\right) w$$

l_k: lower limit of quartile class .
w: width of quartile class.
f_k: Frequency of the quartile class.
F_{k-1} : Ascending Cumulative Frequency before the quartile class

Example:

The following table gives the weekly income ($) of 100 families find **Quartiles**.

Classes	f_i
100_120	3
120_140	7
140_160	14
160_180	20
180_200	18
200_220	12
220_240	6

Solution:

Classes	f_i	ACF (F_i)
100_120	3	3
120_140	7	10
140_160	14	24
160_180	20	44
180_200	18	62
200_220	12	74
220_240	6	80

Position for $D_6 = i\left(\frac{n}{10}\right) = \frac{6*80}{10} = \frac{480}{10} = 48$

$F_{k-1} < i\left(\frac{n}{10}\right) < F_k$

$44 < 48 < \underline{62}$

$$D_6 = l_k + \left(\frac{6\left(\frac{n}{10}\right) - F_{k-1}}{f_k}\right) w$$

$$= 180 + \left(\frac{48-44}{18}\right)20$$

$$= 180 + \left(\frac{4}{18}\right)20$$

$$= 184.4$$

(Percentiles for Grouped data) داتا ڕیزکراوەکان

Continuous variable: ئەگەر هاتوو جۆری داتاکان بەردەوام بوو

$$P_i = l_k + \left(\frac{i\left(\frac{n}{100}\right) - F_{k-1}}{f_k}\right) w$$

l_k: lower limit of percentile class .

w: width of percentile class.

f_k: Frequency of the percentile class.

F_{k-1} : Ascending Cumulative Frequency before the percentile class.

Example:

The following table gives the weekly income ($) of 100 families find **percentile.**

Classes	f_i
100_120	3
120_140	7
140_160	14
160_180	20
180_200	18
200_220	12
220_240	6

Solution:

Classes	f_i	ACF (F_i)
100_120	3	3
120_140	7	10
140_160	14	24
160_180	20	44
180_200	18	62
200_220	12	74
220_240	6	80

Position for $P_{50} = i\left(\frac{n}{100}\right) = \frac{50 * 80}{100} = \frac{4000}{100} = 40$

$F_{k-1} < i\left(\frac{n}{100}\right) < F_k$

$24 < 40 < \underline{44}$

$$P_{50} = l_k + \left(\frac{50\left(\frac{n}{100}\right) - F_{k-1}}{f_k}\right) w$$

$$= 160 + \left(\frac{40 - 24}{20}\right) 20$$

$$= 176$$