### ANNEX 68 The statistical indicators and methods of data analysis

September, 2015

#### Statistical indicators

- System of statistical indicators set of related indicators, who have one level and crossover structure and concrete task.
- Statistical indicators are split by 2 groups:
  - Individual indicators
  - Summary indicators

Examples?

#### Statistical indicators

- According to the form of expression:
  - Absolute indicators
  - Relative indicators
  - Average indicators
- According to the time factor:
  - In concrete time moment Moment indicators;
  - For defined time period (for longer period) Interval indicators

Examples????

#### Statistical indicators

- According to the defined territory:
  - Whole territorial indicators
  - Regional indicators
  - Local indicators
- Statistical indicators also are split:
  - Accounting valuation indicators
  - Analytical indicators

#### Absolute quantities

#### Absolute quantities

- Initial expression of data is in form of **absolute quantities** 
  - Absolute quantities define sizes of fact, direct sizes, number of units, spread in concrete time and place.
  - Absolute quantities have a certain unit of measure.
- Absolute quantities split by:
  - Individual

amount of set, for example., salary of 1 employee, import from the Russia

– Summary

wage bill in enterprise, number of cares in enterprise, etc.

### Absolute quantities

#### Methods of calculation of absolute quantities:

- Balance method
- Method of economically functional connection
- Normative method
- Method of sampling
- Experts method

#### Balance method



## If we know 3 of variables we can calculate the fourth

## Method of economically functional connection

• Variable are calculated by formula (equation), which includes economically linked quantities. *Example, turnover of goods, in EUR* 

#### **Price of good x** sold units = Turnover

If we know the price and the number of sold goods, then can calculate the turnover of goods in financial terms for the concrete period of time. Not need to make the direct observation.

#### Normative method

Variable are calculated, used multiplication with linked quantity by defined coefficient.

 Coefficients can be results of previous period calculations or in nature existing connection.

For example, to calculate the weight of animal , based on changes in size.

### Method of sampling

**Data have been gotten** to survey a part of group. This part have been selected, based on defined parameters.

For example, to get position of society about some process or thing, usually survey 1000 people.

#### Experts method

Data have been getting, based on opinion of competent people from the corresponding industry, to observe the absolute quantities.

#### Units of measure of absolute quantities

#### • Natural units of measures

natural quantity – number, mass, volume, power etc.

- Can be several units of measures, for example m un  $m^2$
- Can be complex units of measures, for example. milliontonnkilometers

#### • Money units of measure

Union of natural un value, only in financial terms–GDP, profit, turnover etc.

#### • Working units of measure

*work consumption in man-hours, person-days, with number of employed + other unit of measure* 

#### Relative quantities

### Relative quantities

- **Relative quantities express the numeric relation of things**, they characterize the level of things. They let to rate the qualitative character of things. More stabile in comparing with absolute quantities
  - Easer memorizing
  - Correspond to measurement and characterization of development speed of things

## Relative quantities

Without relative quantities we can not to measure:

- Composition of researched thing
- Development intensity of researched thing in time
- To rate the level of development of one thing on the background of linked thing
- To make the territorial comparison, included international

## Conditions of formation of relative quantities:

- Base must be stabile (quantity come in normal conditions)
- Opposite quantities must be correlative
- Sessional things must be taken in account
- Comparability of territorial subordination must be provided
- Consistent unit of measure must be provided (EUR, kg, etc.)

#### Relative quantities are expressed:

• In coefficients;

• Ratio of two numbers= k, times

- In percentages
- In per miles;
- In per decimals;
- In percentage points.

- k \* 100 = %
- k \* 1000 = ‰
- $k * 10\ 000 = \%00$
- x% y% = % p

For example, coefficient 2,543 is the same as:

- 2,5 times,
- 254,3 %,
- 2543 ‰
- 25430 ‰<sub>0</sub>.

## Forms of Relative quantities

- Relative quantities of dynamics;
- Relative quantities of projections (task) and realization of projection
- Relative quantities of structure
- Relative quantities of coordination
- Relative quantities of comparison
- Relative quantities of intensity

## Relative quantities of dynamics

The Development of things in time is Expressed (dynamic)

## Fact in reporting period has been divided by quantity of the same indicator in the base period.

• If quantity expressed by coefficient, that it is increasing coefficient (k), if expressed by percentage – growth rate(T).

k \* 100 % = T, %

• Changes of relative quantity of dynamics – rate of increase (T <sub>p</sub>).

T % - 100 % = T p, %

Rate of increase can be negative

### Relative quantities of dynamics

• Example : GDP value in actual prices in 2011 3rd quarter was 3 608 993 thous. EUR, but in 2012 3rd quarter was 3 901 152 thous. EUR.

$$T = 3 \ 901 \ 152 : 3 \ 608 \ 993 = 1.08 \ * \ 100 \ \% = 108 \ \%$$
  
$$T \ p = 108 \ \% \ - \ 100 \ \% = 8 \ \%$$

We can conclude, that GDP in actual prices was increased by 8 %.

#### Relative quantities of dynamics

**Example** : The number of socially insurance persons in 2008 was 940 000 persons, but in 2009 was 890 000 persons.

 $\begin{array}{l} T = 890\ 000: 940\ 000 = 0,947\ *\ 100 = 94,7\ \% \\ T\ p = 94.7\ \ \%\ -\ 100\ \% = -\ 5.3\ \% \end{array}$ 

We can conclude, that number of socially insurance persons during the year decreased by 5.3 %.

940 000:890 000 = 1,056 or 105,6% or by 5,6%

- Used in market economy.
- Relative quantity of projection shows for how many times or by how many percentage we have to increase or decrease the indicator in projection.

#### **Example:**

Latvias export to the country N in the base year was EUR 528 thous. There were projection to increase the export in the reporting year for the EUR 63 thous. The actual value of export in reporting year was 544 thous EUR.

Relative quantity of task of projection = (528+63) : 528 = 591:528 = 1,119 x 100 = 111.9%

The amount of export was projected to increase by 11.9% (111.9% -100%)

Relative quantity of the realization of projection =  $544:591 = 0.920 \times 100 = 92.0\%$ 

Calculation shows that projection is not realized by 8% (92,0% – 100,0%)

#### **Continuation of Example:**

Between the slide showed indicators we can see the connection:

- Coefficient of growth (relative quantity of dynamics) = task of projection x realization of projection: 1.119 \*0.920 = 1,030
- The growth rate (relative quantity of dynamics) = 544 : 528 = 1.030 x100 = 103%
- **Testing of connection**: 1.030 = 1.119 x 0.920

The correlation between these two indicators give the possibility to calculate the third:

- Relative quantity of task of projection = growth rate : relative quantity of realization of projection = 1.030 :0.920 = 1.119
- Relative quantity of the realization of projection = growth rate : relative quantity of task of projection = 1.030 : 1.119 = 0.920

If we would like to **project the growth of the current level**, than we have divided the actual level by the projected level (in the reviewed period)

#### **Example:**

Ministry of Welfare projected to increase the average amount of pensions by 2%. Actually the average amount of pensions increased by 2,5%.

**Realization of projection in percentage** =  $(102,5: 102,0) \ge 100,5\%$ 

If we would like to project the **relative decrease of the current level**, than we have the projected level divide by the actual level (in the reviewed period) **Example:** 

Ministry of Welfare projected decrease the number of old age pensioners by 0,4%. Actually the number of pensioners decreased by 0,6%.

**Realization of projection in percentage** = (99,6: 99,4) x 100 = 100,2%

The level of realization of projection are calculated with precision one tenth (0,1%), but if realization of projection is in interval 99,0-100% (for example, 99,86%), then – with precision hundredth (0,01%). Not to allowed to approximate till full 100,0%.

## Relative quantities of structure

- Structure placement and connection of elements of statistical set.
- Relative quantities have been calculated from the grouped data and they characterize the proportion of the separate parts of researched things in the same tings total.
- Relative quantities of structure have been calculated as absolute quantities of separate element of researched thing divided by total of the researched thing.
- Usually relative quantities of structure have been expressed in percentage. Sum of relative quantities of all elements = 100%

### Relative quantities of structure

#### Example

#### Pension expenditures, in mln. EUR

Type of pensions	Code	In actual prices, 2015, I half of year	In percentage from the total
Old age	А	787.4	89.4%
Disability	В	14.3	1.6%
Survival	С	6.4	0.7%
Service	D	73.0	8.3%
Total		881.1	100.0%

#### Relative quantities of structure

#### **Continuation of Example**

- A = 787.4 : 881.1 x 100 = 89.4%
- $B = 14.3 : 881.1 \ge 100 = 1.6\%$
- $C = 6.4 : 881.1 \times 100 = 0.7\%$
- $D = 73.0 : 881.1 \ge 100 = 8.3\%$

89.4% + 1.6% + 0.7% + 8.3% = 100%

#### Relative quantities of coordination

- Relative quantities of coordination characterize the relation between two elements of the same set.
- Relative quantities of coordination have been calculated as one element of set divided by the other element in the same absolute or relative set. Elements are related between each other.

## Relative quantities of coordination

#### Example

The amount of LV export in December 2013 was 810,1 mln.EUR; amount of import – 993,5 mln. EUR

(810,1 : 993,5) x 100% = 81.5% - export was 81,5% of import

993,5:810,1=1.2 – amount of import was 1,2 times higher as amount of export

#### Relative quantities of comparison

- Relative quantities of comparison characterize relation between two or more monotone objects in the fixed moment or period of time.
- Relative quantities of comparison are calculated as ratio between two absolute or relative quantities

## Relative quantities of comparison

#### Example

- Territory of Latvia 64,6 thous. km<sup>2</sup>; of US 9363,5 thous. km<sup>2</sup>; of Russia – 17075,4 thous. Km
- 9363,5 : 64,6 = 144,9 US territory is 144.9 times larger as Latvias
- 17075,5 : 64,6 = 264,3 Russia territory is 264,3 times larger as Latvias
- Or
- (64,6:9363,5)x 100 = 0,7% Latvias territory is 0,7% of US territory and (64,6%:17075,4) x 100 = 0,4% or Latvias territory is 0,4% of Russia territory

#### Relative quantities of intensity

- Relative quantities of intensity characterize spread of one thing in other related thing.
- Relative quantities of intensity have been calculated as dividing the absolute quantity of researched thing by absolute quantity of the environment where thing developed or spread.

# Relative quantities of intensity

Example

- Realization of Industry production in LV in 2006 I quarter in actual prices was 1240 mln. EUR; number of inhabitants was 2294,6 thous. People
- Relative quantities of intensity (amount of industry production per one inhabitant) =
- 1 240 000 : 2 294.6 = 540,40 EUR

# Absolute differences of relative quantities

- Differences of relative quantities showed importance of absolute differences of relative characteristics, have an analytic importance.
- The form of expression of absolute differences of relative quantities are points, which can be percentages points %p (% -%); promiles points %p (% -%); promiles points

# Absolute differences of relative quantities

#### Example

Realization of projection of Realization of projection of average number of pensioners

Type of pensions	C		Realization of projection,
	projected	%	
Old age	472,0	474,3	100.5%
Disability	72,0	73,2	101.7%
Total	544	547.5	100.6%

101,7 - 100,5 = 1,2101,7% : 100,5% = 1,01

#### Distribution orders

## Distribution orders

- Distribution order consists of two elements:
  - Numeral meaning of indication
  - References for how many units have that or other indication of distribution or how significant is frequency of variants
- Variant numeral indication of meaning
- Ranging
  - In increasing sequence (direct range)
  - In decreasing sequence (turned range)

## **Distribution orders**

#### Example

Distribution orders which are used for the direct range (increasing sequence)

Age , in years	Number of employees in enterprise, persons
18 - 30	100
30 - 50	250
50 and older	86

Distribution orders which are used for the turned range (reducing sequence)

Number of employees, persons	Number of enterprises
200 – 150	54
150 – 100	267
100 and less	589

The main characterized indicators:

- variant, marked as *x*;
- frequency of variants:
  - Absolute quantities, marked as f; f = x1 + x2 + ... + xn
  - Relative quantities, marked as *w*. Equivalent to relative quantities of structure

$$w = \frac{f}{\sum f}$$

- Density of intervals, marked by f(b) and w(b).

#### Example

Distribution of employed by the loads

Load (x)	0.50	0.75	1.00
Employed (f), persons	25	102	78
Relative frequency (w)	0.12	0.50	0.38

How to calculate the relative frequency?

- In addition to the direct frequency (*f* un *w*) in distribution orders also **accumulated frequencies** are used :
  - the accumulated absolute frequencies of variants marks as  $S_f$
  - the accumulated relative frequencies of variants marks as  $S_w$

The accumulated frequencies = consecutive sum of absolute or relative frequencies.

• The accumulated relative and absolute densities are getting as division between accumulated frequencies and accumulated lengths of intervals,  $\Delta$ 

# **Example** of distribution order: age structure of divorce people in 2009

Variants	Frequencies					
(years) x	Direct		Accumulated			
	Absolute f	Relative <i>w</i> , %	Absolute Sf	Relative Sw		
20 - 24	140	2.2	140	2.2		
25 - 29	745	11.7	885	13.9		
30 - 34	1 358	21.4	2 243	35.3		
35 - 39	1 352	21.3	3 595	56.6		
40 - 49	1 842	29.1	5 437	85.7		
50 - 59	690	10.9	6 127	96.6		
60 and more	214	3.4	6 341	100.0		
Total:	6 341	100.0	X	X		

 The accumulated relative and absolute density are getting as division between accumulated frequencies and accumulated lengths of intervals, Δ

# **Example** of distribution order: age structure of divorce people in 2009

Variants	Length of	Freque	Frequencies Densities of Interv		
(years) x	Interval	Absolute f	Relative w	f(b)	w(b)
20 - 24	4	140	2.2	35	0.55
25 - 29	4	745	11.7	186	2.93
30 - 34	4	1358	21.4	340	5.35
35 - 39	4	1352	21.3	338	5.36
40 - 49	9	1842	29.1	205	3.23
50 - 59	9	690	10.9	77	1.21
60 and more	9	214	3.4	23.78	0.38
Total:		6341	100.0	Х	X

# **Example** of distribution order: age structure of divorce people in 2009

Variants (years) x	Length of Interval	Accumulat ed length of	Accumulated		Accumulated densities	
		Interval	Absolute Sf	Relative Sw	Sf(b)	Sw(b)
20 - 24	4	4	140	2.2	35	0.55
25 - 29	4	8	885	13.9	111	1.74
30 - 34	4	12	2243	35.3	187	2.94
35 - 39	4	16	3595	56.6	225	3.53
40 - 49	9	25	5437	85.7	218	3.43
50 - 59	9	34	6127	96.6	180	2.84
60 and more	9	43	6341	100.0	147	2.33

- The most important and widely used in practice average quantities are as follows:
  - arithmetic average
  - geometric average
  - square average
- Depending on the specific characters of processed average quantities are as follows:
  - unweighted (simple) averages, calculated from the ungrouped data;
  - weighted averages, calculated by the grouped data.

The simple arithmetic average shall be calculated using the formula:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_1}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

where  $X_1$  - variant; n – number of units in set

**The harmonious average** shall be calculated using the formula:

$$\overline{X} \quad \mathbf{h} = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_{1}}} = \frac{n}{\frac{1}{X_{1}} + \frac{1}{X_{2}} + \dots + \frac{1}{X_{n}}}$$
  
where  $X_{1}$  - variant;

n – number of units in set

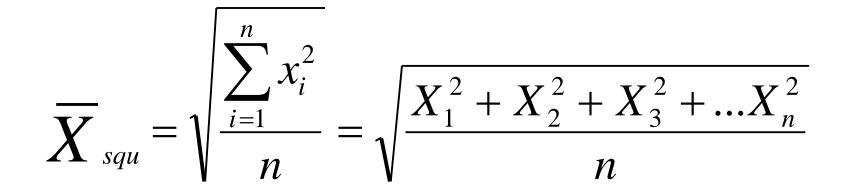
#### Example – harmonious average.

5 employees make one kind of product, but with different working usage for making one product. The first employee use 2,1 h, second -1,5 h, third -2 h, fourth -1,6 h, fifth -2,3 h.

The average labour intensity (t) =

$$h = \frac{1+1+1+1+1}{\frac{1}{2.1}+\frac{1}{1.5}+\frac{1}{2.0}+\frac{1}{1.6}+\frac{1}{2.3}} = \frac{5}{0.48+0.67+0.5+0.63+0.77} = \frac{5}{3.05} = 1.64 \text{ h}$$

Square average formula:



Geometric average formula:

$$\overline{X}_{geom} = \sqrt[n]{\prod_{i=1}^{n} X_i} = \sqrt[n]{X_1 * X_2 * X_3 * ... X_n}$$

#### Example

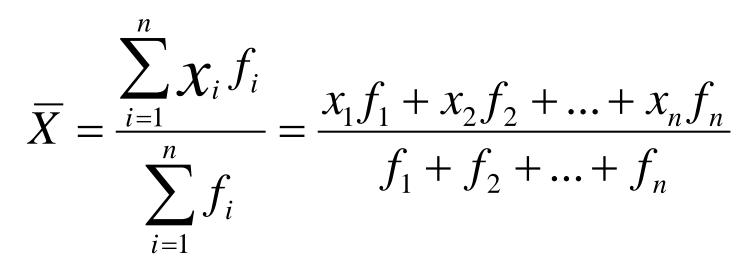
Latvias inflation in February 2006 compared to January was 126.9%; in March compared with February 127.2%; in April compared with March 127.9%; in May compared with April 129.5%.

Average monthly inflation:

$$\overline{X}_{inf} = \sqrt[4]{1,269 * 1,272 * 1,279 * 1,295} = 127,9\%$$

- If the data is grouped, then use the **weighted** averages
- They are calculated by taking in account of each variant repetition quantity (statistical weights)
- In statistics the weighing is called as variants multiplication with their frequencies

#### Weighted arithmetical average formula:



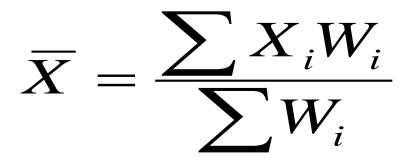
Where f –frequencies of variants

#### Example

Number of employees, people, (f)	1	2	6	5	3	2	1
Average wage, EUR (X)	280	355	475	540	690	780	960

$$\overline{X} = \frac{1 \times 280 + 2 \times 355 + 6 \times 475 + 5 \times 540 + 3 \times 690 + 2 \times 780 + 1 \times 960}{20} = 556,50$$

In calculating the **weighted arithmetic average** instead of absolute values can use the **proportions** of variants or **relative frequencies** of variants (w).



Where  $\sum Wi = 100,0\%$ 

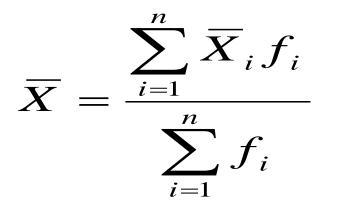
#### **Example**:

Apartment house has such division of apartments:

- 22% one room flat
- 50% two rooms ...
- 19% three rooms ...
- 9% four rooms ...

$$\overline{X} = \frac{1*22 + 2*50 + 3*19 + 4*9}{22 + 50 + 19 + 9} = 2,15rooms$$

**Averages** can be calculated from the variants which are themselves already averages:



Where  $\overline{X}_i$  - weighted arithmetical average of group i  $f_i$  - statistical weight of group i

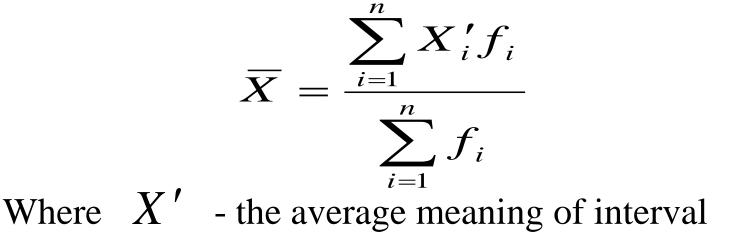
#### **Example:**

Companies	l	II	III
Average wage, EUR	574	618	821
Number of employed, people	202	135	49

$$\overline{X} = \frac{574 * 202 + 618 * 135 + 821 * 49}{202 + 135 + 49} = 620,74EUR$$

# Calculation of arithmetic average for the interval distribution order

#### The **arithmetic average for the interval distribution order** is calculated using the formula:



#### Calculation of arithmetic average for the interval distribution order Example:

Age, years (X)	Number of employees, people (f)	Structure of employed, % (W)	Centers of intervals $(X'_i)$
until 25	6	7,7	22,5
25-30	20	25,6	27,5
30-35	16	20,5	32,5
35-40	30	38,5	37,5
40-45	4	5,1	42,5
45 and more	2	2,6	47,5
	78	100,0	

How to calculate the average age of all employees?

# Calculation of arithmetic average for the interval distribution order

#### Example

Average age of employed by absolute quantities:

$$\overline{X} = \frac{6*22,5+20*27,5+16*32,5+30*37,5+4*42,5+2*47,5}{6+20+16+3+4+4} = \frac{2595}{78} = 33,3 \text{ years}$$

Average age of employed by relative quantities:  $\overline{X} = \frac{7,7*22,+25,6*27,5+20,5*32,5+38,5*37,5+5,0*42,5+2,6*47,5}{7,7+25,6+20,5+38,5+5,1+2,6} = \frac{3327,5}{100,0} = 33,3 \text{ years}$ 

Indicators of distribution center are:

- Arithmetical average
- Mode
- Median

• Average quantities **of discrete distribution** order is calculated by the weighted arithmetic average formula:

$$\overline{X} = \frac{\sum_{i=1}^{n} \chi_{i} f_{i}}{\sum_{i=1}^{n} f_{i}} = \frac{x_{1} f_{1} + x_{2} f_{2} + \dots + x_{n} f_{n}}{f_{1} + f_{2} + \dots + f_{n}}$$

• Average quantities **of interval distribution order** is calculated by the formula:

$$\overline{X} = \frac{\sum_{i=1}^{n} \chi_{i} f_{i}}{\sum_{i=1}^{n} f_{i}} = \frac{\chi_{1}' f_{1} + \chi_{2}' f_{2} + \dots + \chi_{n}' f_{n}}{f_{1} + f_{2} + \dots + f_{n}}$$
  
Where  $\chi_{1}'$  - average meaning of interval

**Mode** in the **discrete** distribution order is the highest frequently met number in the number group

**Example**: mode in the number group: 2, 3, 3, 5, 7 and 10 is 3

Mode in the interval distribution order is calculated by formula:

$$M_{0} = X_{0} + \Delta * \frac{f_{mo} - f_{mo-1}}{(f_{mo} - f_{mo-1}) + (f_{mo} - f_{mo+1})}$$

Where  $M_0$  - mode

- the lowest border of the modal interval  $X_0$
- the space of the modal interval  $\boldsymbol{\Lambda}$ 
  - the frequency of the modal interval
- $f_{mo}$  the frequency of the modal inter- $f_{mo-1}$  prior modal interval frequency
- $f_{mo+1}$  after modal interval frequency

- Median is the statistical quantity of indication of the unit placed in the center of ranged distribution order.
- **Ranged** is the distribution order containing all the units listed in increasing or reducing order.

Median **unit serial number** is found by increasing the number of members per one unit and the result obtained by dividing with number two.

$$M_{e(No)} = \frac{n+1}{2}$$
 or  $M_{e(No)} = \frac{\sum f+1}{2}$ 

## **Median** in the **interval** distribution order is calculated by formula:

$$Me = X_0 + \Delta * \frac{\frac{\sum f}{2} - S_{fme-1}}{f_{me}}$$

Where *Me* - median

- $X_0$  the lowest border of the median interval
- $\Delta$   $\,$  the space of the median interval
- $\sum f$  the sum of distribution order members
- $S_{fme-1}$  prior median interval accumulated frequencies
  - $f_{me}$  median interval frequency

#### Example

Average wage, EUR (X)	Number of employed, people (f)	Structure of employed, % (W)	Accumulated frequencies	
			Absolute(Sf)	Relative (Sw)
180-220	10	10.3	10	10.3
220-350	28	28.9	38	39.2
350-450	25	25.8	63	64.9
450-600	18	18.6	81	83.5
600-800	13	13.4	94	96.9
800 and				
more	3	3.1	97	100.0

Mode of average wage per month

- Calculated used absolute indicators:

$$M_o = 220 + 130 * \frac{28 - 10}{(28 - 10) + (28 - 25)} = 220 + 130 * \frac{18}{21} = 331,43EUR$$

- Calculated used relative indicators:

$$M_o = 220 + 130 * \frac{28,9 - 10,3}{(28,9 - 10,3) + (28,9 - 25,8)} = 331,43EUR$$

Median of average wage per month

- Calculated used absolute indicators:

$$Me = 350 + 100 * \frac{48,5 - 38}{25} = 350 + 100 * \frac{10,5}{25} = 392,0EUR$$

- Calculated used relative indicators:

$$Me = 350 + 100 * \frac{50 - 39,2}{25,8} = 391,86EUR$$

#### Indicators of variations

## Indicators of variations

- The main absolute indicators of variations are:
  - Amplitude
  - Average linear error/deviation
  - Dispersion
  - Average quadratic error/deviation (standard error)
- The main relative indicators of variations
  - Coefficient of variation
  - Oscillation coefficient
  - Relative linear error

**Amplitude** is the difference between maximal and minimal border of variable and calculated by formula:

Rv = Xmax - Xmin,

where Rv – amplitude;

Xmax – the maximal value of indication; Xmin – the minimum value of indication.

#### Average linear error calculated by formula:

$$\alpha = \frac{\sum \left| X - \overline{X} \right| f}{\sum f}$$

Where f-frequency of variant

- $\sum f$  number of units of set
- $\overline{X}$  arithmetical average of indication

#### **Dispersion calculated by formula:**

$$S^{2} = \delta^{2} = \frac{\sum (X - \overline{X})^{2} f}{\sum f}$$

# Average quadratic error/deviation (standard error) calculated by formula:

$$S = \delta = \sqrt{\frac{\sum (X - \overline{X})^2 f}{\sum f}}$$

#### Relative indicators of variations

**Coefficient of variation** calculated by formula:

$$K_{\rm var} = \frac{\delta}{\overline{X}} * 100$$

#### Relative indicators of variations

**Oscillation coefficient** calculated by formula:

$$K_{osc} = \frac{R_v}{\overline{X}} * 100$$

#### Relative indicators of variations

**Relative linear error** calculated by formula:

$$\alpha_{rel} = \frac{\alpha}{\overline{X}} * 100$$

## Thank you