KHARKIV NATIONAL MEDICAL UNIVERSITY



MEAN VALUES,

METHODS OF THEIR CALCULATION

Methodical instructions

for students to the practical lesson

on the course ***“Social medicine, public health (biostatistics)”***

for students in the specialty:

– 222 “Medicine”

– 228 “Pediatrics”,

– 221 “Dentistry”.

Kharkiv

2019

MINISTRY OF PUBLIC HEALTH OF UKRAINE

KHARKIV NATIONAL MEDICAL UNIVERSITY

DEPARTMENT OF PUBLIC HEALTH AND HEALTHCARE MANAGEMENT

MEAN VALUES,

METHODS OF THEIR CALCULATION

Methodical instructions

for students to the practical lesson

on the course ***“Social medicine, public health (biostatistics)”***

for students in the specialty:

– 222 “Medicine”

– 228 “Pediatrics”,

– 221 “Dentistry”.

*Затверджено вченою радою Харківського національного медичного університету.*

*Протокол № 5 від 18.05.2017*

Kharkiv

KhNMU

2019

УДК 311.21

Mean values, methods of their calculation: methodical instructions for students to the practical lesson on the course “Social medicine, public health (biostatistics)” for students in the specialty 222 “Medicine”, 228 “Pediatrics, 221 “Dentistry” / V.A. Ohniev, S.G. Usenko, I.A. Chukhno. – Kharkiv : KhNMU, 2019. – 16 p.

|  |  |
| --- | --- |
| Drafters: | Ohniev V.A. |
|  | Usenko S.G. |
|  | Chukhno I.A. |

**RECOMMENDATIONS FOR STUDYING THE TOPIC**

**The aim of the class:** to learn the definition of mean values and methods of their calculation.

**Need to know:**

* ***program questions:***
* the definition of arithmetic mean values;
* the use of arithmetic mean values in clinical and epidemiological studies;
* practical value of arithmetic means;
* types of arithmetic mean values;
* methods of arithmetic mean values calculation;
* features of arithmetic mean values use.

**Need to be able to:**

* evaluate and analyze statistical indicators and parameters of statistical aggregates;
* master the methods of arithmetic mean values calculation.

**Recommended literature**

**Basis literature**

1. Біостатистика / за заг. ред. чл.-кор. АМН України, проф. В.Ф. Москаленка. – К. : Книга плюс, 2009. − С. 12-31.

2. Социальная медицина и организация здравоохранения / под общ. ред. Ю.В. Вороненка, В.Ф. Москаленко. – Тернополь : Укрмедкнига. 2000. –
С. 23-32.

3. Социальная гигиена и организация здравоохранения / под ред. Н.Ф. Серенко, В.В. Ермакова. – М. : Медицина, 1984. – С. 102-104.

4. Тестовые задачи по социальной медицине, организации здравоохранения и биостатистике : учеб.пособ. для студентов мед. ф-тов / под ред. В.А. Огнева. – Харьков : Майдан, 2005. – С. 9-14.

5.Лекционный курс кафедры.

**Additional literature**

1. Альбом А. Введение в современную эпидемиологию / А. Альбом, С. Норелл. – Таллинн, 1996. – 122 с.

2. Власов В.В. Введение в доказательную медицину / В.В. Власов. – М. : Медиа Сфера, 2001. – 392 с.

3. Герасимов А. Н. Медицинская статистика / А.Н. Герасимов. – М. : ООО «Мед.информ. агентство», 2007. – 480 с.

4. Зайцев В.М. Прикладная медицинская статистика / В.М. Зайцев, В.Г. Лифляндский, В.И. Маринкин. – СПб. : ООО «Изд-во ФОЛИАНТ», 2003. – 432 с.

5. Общая теория статистики: учебник / под ред. чл.-корр. РАН И.И. Елисеевой. − 4-е изд., перераб. и доп. − М. : Финансы и Статистика, 2000. − 480 с.

6. Основы доказательноймедицины / под ред. М. П. Скакун. – Тернополь : Укрмедкнига, 2005. – 244 с.

7. Реброва О.Ю. Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA / О.Ю. Реброва.–М. : Медиа Сфера, 2002. – 312с.

8. Сергиенко В.И. Математическая статистика в клинических исследованиях / В.И. Сергиенко, И.Б. Бондарева. – М. : ГЭОТАР-МЕД, 2001. – 256 с.

**Information resources**

1.Население Украины. Демографический ежегодник. – К. : Госкомстат Украины –[www.ukrstat.gov.ua](http://www.ukrstat.gov.ua)

2.U.S. National Library of Medicine –Национальная медицинская библиотека США– <http://www.nlm.nih.gov/>

3.Государственная научно-педагогическая библиотека Украины им. В.О. Сухомлинского–<http://www.dnpb.gov.ua/>

4.Научная библиотека Харьковского національного медицинского университета – <http://libr.knmu.edu.ua/index.php/biblioteki>

5.Научная педагогическая библиотека им. К.Д. Ушинского Российской академии образования – <http://www.gnpbu.ru/>

6.Национальная библиотека Украины им. В.И. Вернадского –<http://www.nbuv.gov.ua/>

7.Национальная научная медицинская библиотека Украины –<http://www.library.gov.ua/>

8.Харковская государственная научная библиотека им. В.Г. Короленка – http://korolenko.kharkov.com

9.Центральная библиотека Пущинского научного центра РАН –<http://cbp.iteb.psn.ru/library/default.html>

10.Центральная научная медицинская библиотека Первого Московского государственного медицинского университетаим. И.М. Сеченова–<http://elibrary.ru/defaultx.asp>

**BASIC THEORETICAL MATERIAL**

**FOR PREPARATION FOR THE LESSON**

**1. The essence and types of mean values**

**The mean value** is the derivative of the statistical aggregate which characterizes a whole series of observations with one number. It is a summary of the entire statistical aggregation on a certain quantitative attribute. The mean value is a kind of expression of the general property that a **variable** of the aggregation is characterized by.

The main requirements for practical and scientific use *of mean values are next:*

*1) the mean values* should be determined on the basis of a summarizing of the facts;

2) *the mean values* should be applied to the qualitatively homogeneous aggregations.

The mean values cannot be determined if the totality of the studied attributes, processes, phenomena is composed of inhomogeneous elements. Moreover, a sufficient number of observations should be taken into account.

Thus, to calculate the mean values, two basic requirements must be complied:

* homogeneity of the studied quantitative **variable;**
* sufficient number of observations.

The statistics section that studies the mean values is called ***descriptive statistics***.

***The mean values are widely used in the health care system:***

* to characterize the activities of the health care facilities (average occupancy of the hospital bed, average duration of the patient`s hospital treatment, number of visits per inhabitant, sanitary and hygienic standards, etc.);
* to determine the physiological parameters of an organism (pulse rate, blood pressure level, biochemical tissue composition, enzyme activity, etc.);
* to characterize the indicators of the physical development (length, body weight, head circumference, etc.);
* to evaluate data of the sanitary-hygienic surveys (average food ration norms, level of radiation pollution, etc.).

**Types of the mean/average values:**

1. Mode ($Mo$).

2. Median ($Me$).

3. Arithmetic mean/average $(\overbar{X})$.

4. Geometric mean/average.

5. Root/quadratic mean square / average.

**Mode** $(Mo)$is the value of the variable that occurs more often than others in this variation series or statistical aggregation. This is the observation $(x\_{i})$ that has the greatest frequency ($f\_{i}$)*.*

The **median** $(Me)$ is the value of the variable that occupies the middle position in the variation series . It divides the series into two equal parts according to the number of observations. In the case of an odd number of observations **in the simple** variation series the median is determined as follows:

$$ Me=x\_{i}, where i\_{Me}=\frac{n+1}{2}, (1)$$

And for an even number of observations we take as a median an arithmetic mean of 2 central observations.

In the case of **weighted variation series** , the **median** is determined as a variable ($x\_{i}$) with cumulative frequency equal or greater than half of observations:

$$Me=x\_{i} with \sum\_{i=1}^{i}f\_{i}\geq \frac{n}{2} (2)$$

The basic advantage of the mode and median in describing data compared to the mean is that they are not skewed so much by extremely large or small values, and so they may give a better idea of a “typical” value.

**Arithmetic mean (**$\overbar{X}$**)** is more reliable value compared to “mode” and “median”. It relies on all the observations that were made. Arithmetic mean can be calculated in several ways depending on the number of observations, character of the variation series and the availability of computing technologies.

**2. Methods of mean values calculation**

***Arithmetic mean*** is the most common type of mean values. For its calculation, two methods are used:

1) the arithmetic average method;

2) the method of moments.

***The arithmetic average method:*** for a simple variation series, in which each observation occurs only once, the **simple arithmetic mean** is calculated as the ratio of the sum of the observation’s values to the total number of observations:

$$\overbar{X}=\frac{\sum\_{i=1}^{n}x\_{i}}{n} (3)$$

denote: $x\_{i}$ – the value of the each observation,

$n$ – total number of observations.

For weighted variation series in which at least one of the observations occurs two or more times, the **weighted arithmetic mean** is calculated due to the formula:

$$\overbar{X}=\frac{\sum\_{i=1}^{n}x\_{i}×f\_{i}}{\sum\_{i=1}^{n}f\_{i}}=\frac{\sum\_{i=1}^{n}x\_{i}×f\_{i}}{n} (4)$$

denote: $x\_{i}$ – the values of the $i$-th observation;

$f\_{i}$ – the frequency of the $i$-th observation;

$n$ – total number of observations.

***The method of moments***. If there is a grouped or interval variation series and the observations consist of multi-digit numbers and the aggregation consists of a large number of observations, it is expedient to calculate the arithmetic mean with the ***method of moments***. Its formula:

$$\overbar{X}=A×i×\frac{\sum\_{i=1}^{n}\left(a\_{i}×f\_{i}\right)}{n} (5)$$

denote: $A$ – is the conditional mean value (most often, the mode (Mo) is conditionally taken as an average one)

$i$ – an interval;

$a\_{i}$ – the conditional deviation of each observation (in intervals) from the conditional mean value (mode): $a\_{i}=x\_{i}-A$;

$a\_{i}×f\_{i}$ – the product of the deviation ($a\_{i}$) by the frequency ($f\_{i}$).

The method of moments is based on one of the features of the arithmetic mean: the sum of the deviations of all the observations from the arithmetic mean is zero.

**The arithmetic mean is characterized by the following properties:**

1. The arithmetic mean is abstract.
2. The arithmetic mean occupies the middle position in a strictly symmetric variation series.
3. The sum of the deviations of all the observations from the arithmetic mean is 0.
4. The product of the arithmetic mean by the sum of the frequencies is always equal to the sum of the product of the observations by the frequency.
5. If any number is deducted from each of the observations, new arithmetic mean decreases by the same number.
6. *If any number is added to each observation, then the arithmetic mean increases by the same number.*
7. *If each observation is divided by any number, then the arithmetic mean decreases by the same factor.*
8. *If each observation is multiplied by any number, then the arithmetic mean increases by the same factor.*
9. *If all the frequencies are multiplied or divided by any number, the arithmetic mean does not change – if we increase or decrease the equivalent frequency of the observations, we do not change the weight of each individual observation in the series.*

The general properties of the mean are used to facilitate the technique of determining the arithmetic mean using a variation series.

The **harmonic mean** $(H)$ is calculated in cases when the data about the dividend are known and the data about the divisor are unknown.

$$H=\frac{n}{\sum\_{i=1}^{n}\frac{1}{x\_{i}}} or H=\frac{\sum\_{i=1}^{m}n\_{i}}{\sum\_{i=1}^{m}\frac{n\_{i}}{x\_{i}}} \left(6\right)-(7)$$

For example, it is known that 5 doctors conduct the reception for 8 hours, and we need to determine the average time used for a one patient reception. Each doctor spends on the admission of one patient 20; 16; 20; 15; 24 min on average, respectively. The calculation is as follows:



**Explanation**: Each doctor works 8 hours or 480 minutes every day, so totally 5 doctors work 2400 minutes per day.

The load for each doctor can be calculated as duration of doctor’s working day over the time for 1 patient’s acceptance. Thus for the 1st doctor – 480/20 = 24 patients per day; for the 2nd – 480/16 = 30 patients; for the 3rd – 480/20 = 24 patients; for the 4th – 480/15 = 32 patients; for the 5th – 480/24 = 20 patients. So totally all doctors accept 130 patients per day.

Thus average time taken for 1 patient's acceptance can be calculated as total working time over the amount of accepted patients.

The **geometric mean** $(G)$ is defined for those parameters which value’s change happens in a geometric progression (a change in the number of people in the period between the census, results of titration of vaccines, an increase in the body weight of newborn babies during certain months of life, etc.). The formula for calculating a simple geometric mean is as follows:

$$G=\sqrt[n]{\prod\_{i=1}^{n}x\_{i}}=\sqrt[n]{x\_{1}×x\_{2}×…×x\_{n}} (8)$$

The role of mean values in medicine is extremely high. On the one hand, they are used to characterize phenomena as a whole, on the other, they are necessary for the evaluation of individual values. When comparing the individual values with the means, valuable characteristics are obtained for each of them. The use of mean values requires the strict adherence to the principle of homogeneity of the population. Violation of this principle leads to a distorted picture of real processes.

*Therefore*, we must be sure that mean values characterize homogeneous statistical populations in order to use them correctly.

**PRACTICAL TASK**

To perform an independent work, student must compose a grouped variation series, basing on the data given below, and then, on the basis of this series, calculate the arithmetic mean (X) using the arithmetic mean method. Make conclusions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit of observation** | **Considered variable: body weight (kg)** |  | **Unit of observation** | **Considered variable: body weight (kg)** |
| 1 | 77 |  | 19 | 80 |
| 2 | 88 |  | 20 | 80 |
| 3 | 82 |  | 21 | 81 |
| 4 | 78 |  | 22 | 80 |
| 5 | 81 |  | 23 | 81 |
| 6 | 82 |  | 24 | 81 |
| 7 | 84 |  | 25 | 85 |
| 8 | 82 |  | 26 | 83 |
| 9 | 83 |  | 27 | 84 |
| 10 | 86 |  | 28 | 81 |
| 11 | 85 |  | 29 | 81 |
| 12 | 84 |  | 30 | 82 |
| 13 | 83 |  | 31 | 83 |
| 14 | 83 |  | 32 | 83 |
| 15 | 82 |  | 33 | 85 |
| 16 | 79 |  | 34 | 85 |
| 17 | 79 |  | 35 | 86 |
| 18 | 79 |  | 36 | 85 |

**Based on the data presented in the table,**

**we construct a grouped variation series:**

|  |  |
| --- | --- |
| **Considered variable: body weight** | **Frequency of the phenomenon studied** |
| $x\_{i}$ | $$f\_{i}$$ |
| 77 | 1 |
| 78 | 1 |
| 79 | 3 |
| 80 | 3 |
| 81 | 6 |
| 82 | 5 |
| 83 | 6 |
| 84 | 3 |
| 85 | 5 |
| 86 | 2 |
| 87 | 0 |
| 88 | 1 |
| **Total** | **n** =36 |

We calculate the arithmetic mean with the arithmetic mean method (formula 4):

$$\overbar{X}=\frac{\sum\_{i=1}^{n}x\_{i}×f\_{i}}{n}$$

Then we substitute the obtained data in the formula:

$$\overbar{X}=\frac{77∙1+78∙1+79∙3+80∙3+81∙6+82∙5+83∙6+84∙3+85∙5+86∙2+88∙1}{36}$$

$$\overbar{X}=\frac{2963}{36}=82,31 (kg)$$

| $x\_{i}$ | $$f\_{i}$$ | $$x\_{i}×f\_{i}$$ |
| --- | --- | --- |
| 77 | 1 | 77 |
| 78 | 1 | 78 |
| 79 | 3 | 237 |
| 80 | 3 | 240 |
| 81 | 6 | 486 |
| 82 | 5 | 410 |
| 83 | 6 | 498 |
| 84 | 3 | 252 |
| 85 | 5 | 425 |
| 86 | 2 | 172 |
| 88 | 1 | 88 |
| Total | $n=\sum\_{i=1}^{n}f\_{i}=36$  | $$\sum\_{i=1}^{n}x\_{i}×f\_{i}=2963$$ |

**Conclusion**: The arithmetic mean value of body weight is 82,31 kg.

**TEST TASKS**

|  |  |
| --- | --- |
| 1. | The mean values are widely used in medicine and health care. Determine which of the following statements is related to the area of practical application of mean values. |
|  | А | To characterize the qualitative variables  |
|  |  В | To characterize the incidence rate |
|  | С | To compare absolute data |
|  | D | To characterize the natural movement |
|  | \* Е | To determine the physiological parameters of the body |
| 2. | The mean values are widely used to characterize medico-biological data. What requirements must the primary material for their calculation should meet? |
|  | А | Sufficient amplitude of the series |
|  | \* В | Sufficient number of observations |
|  | С | The presence of group variables |
|  | D | The presence of a “null” hypothesis |
|  | Е | The absence of fluctuations in the variables that are being studied |
| 3. | The arithmetic mean values are computed on the basis of the primary statistical material, which must correspond to certain requirements. Determine which of the following applies to these requirements. |
|  | А | Sufficient amplitude of the series |
|  | \*В | Qualitatively homogeneous population |
|  |  С | The presence of group variables |
|  | D | The presence of a “null” hypothesis |
|  | Е | Absence of fluctuations in the variables that are being studied |
| 4. | To determine the physical development of students the average growth and body weight were calculated. Determine what requirements must the primary statistical material for calculating the mean values meet. |
|  | А | Absence of fluctuation of the studied trait |
|  | \* В | Qualitatively homogeneous series |
|  | С | The smallest number of observations |
|  | D | The presence of a “null” hypothesis |
|  | Е | The presence of group variables |
| 5. | In health care, the mean values are used quite widely. Determine which of the following indicators of the health status of the population is estimated by the average mean. |
|  | А | Demographic indicators |
|  | В | Disability |
|  | С | Morbidity of the population |
|  | \* D | Physical development |
|  | Е | Prenosological status |
| 6. | Mean values are widely used in medicine and health care administration. Determine which of the following statements should be variabled to the field of practical application of mean values. |
|  | А | To compare absolute data |
|  | В | To characterize the qualitative variables |
|  | **\*** С | To characterize the organization of work of health institutions |
|  | D | To characterize the natural movement |
|  | Е | To characterize the incidence rate |
| 7. | Variation series are the basis for determining the mean values. Determine which average arithmetic can be calculated from the variation series, in which the frequency of each observation is equal to one. |
|  | А | Amplitude |
|  | В | Limit |
|  | С | Median |
|  | D | Grouped arithmetic mean |
|  | \* Е | Simple arithmetic mean |
| 8. | Variation series is the basis for determining the mean values. Determine which average arithmetic can be calculated from the variation series, where each observation has its own certain frequency. |
|  |  А | Simple arithmetic mean |
|  | **\*** В | Grouped arithmetic mean |
|  | С | Harmonic mean |
|  | D | Geometric mean. |
|  | Е | Root mean square |
| 9. | To assess the medical and biological data obtained during medical examinations, various types of averages are used. Determine which of the following most exhaustively characterizes these data. |
|  | А | Amplitude |
|  | \* В | Arithmetic mean |
|  | С | Median |
|  | D | Mode |
|  | Е | Limit |
| 10. | In the analysis of the variation series, the following data were obtained: the conditional arithmetic mean =A, the values of deviation of observations from the conditional mean , the values of the products of deviations of the observation by their frequencies , the sum of the products , the sum of the frequencies in the variation series , and the value of the interval between the observations . Determine in the form of a formula the order of calculating the arithmetic mean of the variation series with the method of moments. |
|  | А |  |
|  | B |  |
|  | \*C |  |
|  | D |  |
|  | E |  |
| 11. | The variation series has the following form:

|  |  |  |
| --- | --- | --- |
| V | Р | VР |
| 1 | 2 | 2 |
| 2 | 3 | 6 |
| 3 | 4 | 12 |
| 4 | 3 | 12 |
|  | 2 | 10 |
|  |  |  |

Determine the weighted arithmetic mean of the given variation series. |
|  | A | 2,0. |
|  | \*B | 3,0. |
|  | C | 3,5. |
|  | D | 4,0. |
|  | E | 4,5. |
| 12. | What the main property of mean value is used to calculate the arithmetic mean with the method of moments? |
|  | A | All answers are correct |
|  | B | Change of each observation by a certain number increases the mean value by this number |
|  | C | The mean is the generalizing quantity |
|  | D | The mean value occupies the middle position in any variation series |
|  | \* E | The sum of the deviations of all observations from the arithmetic mean is zero |

**CONTROL QUESTIONS**

1. What is the mean value in statistics?
2. What kinds of mean values do you know?
3. What is “mode” and “median”?
4. What are the requirements for the research material, and the statistical totality for determining the mean value?
5. What are the main areas of practical application of mean values?
6. What the main properties of the mean arithmetic value do you know?
7. What is the procedure for calculating a simple arithmetic mean?
8. What is the procedure for calculating the weighted mean of the arithmetic mean?
9. What is the procedure for calculating the arithmetic mean by the method of moments?
10. What is the main property of the mean arithmetic value used to calculate it by the method of moments?
11. Why is the arithmetic mean the most reliable characteristic of the population in terms of the mean value?
12. Is it possible to compare statistical sets with different variations of the quantitative trait with the help of mean values?

**CONTENTS**

|  |  |
| --- | --- |
| Recommendations for studying the topic …………..…………………………. | 3 |
| Basic theoretical material for preparation for the lesson ……………………… | 5 |
| 1. The essence and types of mean values…………..…………………… | 5 |
| 2. Methods of mean values calculation ………………………………… | 6 |
| Practical task ...………………………………………………………………… | 9 |
| Test tasks .……………………………………………………………………… | 12 |
| Control questions………………………………………………………………. | 15 |

*Educational publication*

**MEAN VALUES, METHODS OF THEIR CALCULATION**

Methodical instructions

for students to the practical lesson on the topic:

 ***“Social medicine, public health (biostatistics)”***

for students in the specialty:

222 “Medicine”, 228 “Pediatrics, 221 “Dentistry”.

|  |  |
| --- | --- |
| Drafters: | ***Ohniev Viktor*** |
|  | ***Usenko Svitlana*** |
|  | ***Chukhno Inna*** |

*Translated by Chukhno I., Tregub Y., Ivahnenko D.*

Responsible to sign-off *Ognev V.*

Формат А5. Ризографія. Ум.друк. арк. 1,0.

Тираж 150 прим. Зам. № 17-33397.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Редакційно-видавничий відділ

ХНМУ, пр. Науки , 4, м. Харків, 61022

izdatknmu@mail.ru, izdat@knmu.kharkov.ua

Свідоцтво про внесення суб’єкта видавничої справи до Державного реєстру видавництв,

виготівників і розповсюджувачів видавничої продукції

серії ДК № 3242 від 18.07.2008 р.