МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ГУМАНИТАРНО-ПЕДАГОГИЧЕСКАЯ АКАДЕМИЯ (ФИЛИАЛ)
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «КРЫМСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ В.И. ВЕРНАДСКОГО» В г. ЯЛТЕ

УТВЕРЖДАЮ
Заместитель директора Гуманитарно-педагогической академии (филиал) ФГАОУ ВО «КФУ имени В.И. Вернадского» в г. Ялте
_____________ Н. В. Горбунова
«____» ___________ 2019 г.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО СТАТИСТИКЕ

Статистические показатели: виды средних величин, мода, медиана, среднеквадратическое отклонение, коэффициент вариации.

Направление подготовки: 38.03.01 «Экономика»
Профиль подготовки: «Финансы и кредит»
Квалификация (степень) выпускника: бакалавр

Ялта – 2019 г.

Оглавление

Цель темы:

Статистическая грамотность является неотъемлемой составной частью профессиональной подготовки каждого экономиста, финансиста, социолога, политолога, а также любого специалиста, имеющего дело с анализом массовых явлений, будь то социально-общественные, экономические, технические, научные и другие. Работа этих групп специалистов неизбежно связана со сбором, разработкой и анализом данных статистического (массового) характера. Нередко им самим приходится проводить статистический анализ различных типов и направленности либо знакомиться с результатами статистического анализа, выполненного другими.

В настоящее время от работника, занятого в любой области науки, техники, производства, бизнеса и прочее, связанной с изучением массовых явлений, требуется, чтобы он был, по крайней мере, статистически грамотным человеком. В конечном счете, невозможно успешно специализироваться по многим дисциплинам без усвоения какого-либо статистического курса. Поэтому большое значение имеет знакомство с данной дисциплиной.

Студент должен уметь:

- выявлять основную закономерность изучаемого признака путем вычисления средних величин;

- обосновывать методику применения критериев разнообразия вариационного ряда;

- давать характеристику разнообразия вариационного ряда;

- делать выводы о типичности обобщающей характеристики признака в изучаемой совокупности, используя критерии разнообразия вариационного ряда;

- рассчитывать средние величины и критерии вариационного ряда, используя мастер функций MS Excel.

Студент должен знать:

- методику расчета средних величин и критериев разнообразия вариационного ряда;

- основные понятия темы (вариационный ряд, средние величины, среднеквадратическое отклонение, коэффициент вариации, правило трех сигм, нормальное распределение Гаусса);

- методику анализа средних величин: значение среднеквадратического отклонения и коэффициента разнообразия для оценки вариабельности изучаемого признака и типичности средней величины;

- нормальное распределение вариационного ряда и его значение для оценки статистических показателей;

- область применения характеристик вариационного ряда.

Место проведения:

Аудитории кафедры финансов и кредита и кафедры информационных технологий.

Оснащение занятия:

- Мультимедийный проектор;

- Наглядный материал в виде мультимедийных презентаций;

- Персональные компьютеры.

Информационный материал:

При изучении банковских операций (например, вкладов на депозиты), анализе деятельности организаций за год (получение прибыли и др.), оценке работы персонала предприятия (выработка на одного рабочего и др.) часто возникает необходимость получить представление о размерах изучаемого признака в совокупности для выявления его основной закономерности.

Оценить размер признака в совокупности, изменяющегося по своей величине, позволяет лишь его обобщающая характеристика, называемая средней величиной.

Для более детального анализа изучаемой совокупности, по какому либо признаку помимо средней величины необходимо также вычислить критерии разнообразия признака, которые позволяют оценить, насколько типична для данной совокупности ее обобщающая характеристика.

Определение вариационного ряда.

Вариационный ряд - это числовые значения признака, представленные в ранговом порядке с соответствующими этим значениям частотами.

Основные обозначения вариационного ряда:

V— варианта, отдельное числовое выражение изучаемого признака;

р — частота ("вес") варианты, число ее повторений в вариационном ряду;

n — общее число наблюдений (т.е. сумма всех частот, n = Σр);

Vmax и Vmin — крайние варианты, ограничивающие вариационный ряд (лимиты ряда);

А — амплитуда ряда (т.е. разность между максимальной и минимальной вариантами, А = V max — min).

Виды вариационных рядов:

а) простой — это ряд, в котором каждая варианта встречается по одному разу (р=1);

б) взвешенный — ряд, в котором отдельные варианты встречаются неоднократно (с разной частотой).

Назначение вариационного ряда:

Вариационный ряд необходим для определения средней величины (М) и критериев разнообразия признака, подлежащего изучению (σ , С).

Средняя величина - это обобщающая характеристика размера изучаемого признака. Она позволяет одним числом количественно охарактеризовать качественно однородную совокупность.

Применение средних величин:

- для оценки финансового состояния предприятия — например, результатов деятельности(средняя заработная плата, средняя выработка на одного работника, и др.);

- для оценки работы банка, а также отдельных банковских операций (средняя величина вкладов на депозиты и др.);

- для оценки состояния финансовых организаций.

Средняя величина - это обобщающий показатель статистической совокупности, который погашает индивидуальные различия значений статистических величин, позволяя сравнивать разные совокупности между собой.

Существует 2 класса средних величин: степенные и структурные.

К структурным средним относятся мода и медиана, но наиболее часто применяются степенные средние различных видов.

Таблица 1

Степенные средние различных видов

Вид степенной средней	Показатель степени	Формула расчета	
		простая	взвешенная
Гармоническая	-1	$\overline{x} = \dfrac{n}{\sum \frac{1}{x}}$,	$\overline{x} = \dfrac{\sum m}{\sum \frac{m}{x}}$, где $m = x \cdot f$
Геометрическая	0	$\overline{x} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot x_n}$	$\overline{x} = \sqrt[\sum f]{x_1^{f_1} \cdot x_2^{f_2} \cdot \ldots \cdot x_n^{f_n}}$.
Арифметическая	1	$\overline{x} = \dfrac{\sum x}{n}$	$\overline{x} = \dfrac{\sum xf}{\sum f}$
Квадратическая	2	$\overline{x} = \sqrt{\dfrac{2x^2}{n}}$	$\overline{x} = \sqrt{\dfrac{\sum x^2 f}{\sum f}}$
Кубическая	3	$\overline{x} = \sqrt[3]{\dfrac{\sum x^3}{n}}$	$\overline{x} = \sqrt[3]{\dfrac{\sum x^3 f}{\sum f}}$

Степенные средние величины:

Степенные средние могут быть простыми и взвешенными.

Простая средняя величина рассчитывается при наличии двух и более не сгруппированных статистических величин, расположенных в произвольном порядке по следующей общей формуле:

$$\overline{X} = \sqrt[m]{\frac{\sum X^m}{N}}$$

Взвешенная средняя величина рассчитывается по сгруппированным статистическим величинам с использованием следующей общей формулы:

$$\overline{X} = \sqrt[m]{\frac{\sum X^m f}{\sum f}}$$

Общая формула взвешенной степенной средней величины

где X – значения отдельных статистических величин или середин группировочных интервалов;

m - показатель степени, от значения которого зависят следующие виды степенных средних величин:

при m = -1 средняя гармоническая;

при m = 0 средняя геометрическая;

при m = 1 средняя арифметическая;

при m = 2 средняя квадратическая;

при m = 3 средняя кубическая.

Используя общие формулы простой и взвешенной средних при разных показателях степени m, получаем частные формулы каждого вида, которые будут далее подробно рассмотрены.

Средняя арифметическая:

Средняя арифметическая - это самая часто используемая средняя величина, которая получается, если подставить в общую формулу m=1. Средняя арифметическая простая имеет следующий вид:

$$\overline{X}_{ар.прост.} = \frac{\sum X}{N}$$

где X - значения величин, для которых необходимо рассчитать среднее значение; N - общее количество значений X (число единиц в изучаемой совокупности).

Например, студент сдал 4 экзамена и получил следующие оценки: 3, 4, 4 и 5. Рассчитаем средний балл по формуле средней арифметической простой: (3+4+4+5)/4 = 16/4 = 4.

Средняя арифметическая взвешенная имеет следующий вид:

$$\bar{X}_{ар.взвеш} = \frac{\sum Xf}{\sum f}$$

Средняя арифметическая взвешенная:

где f - количество с одинаковым значением X (частота).

Например, студент сдал 4 экзамена и получил следующие оценки: 3, 4, 4 и 5. Рассчитаем средний балл по формуле средней арифметической взвешенной: (3*1 + 4*2 + 5*1)/4 = 16/4 = 4.

Если значения X заданы в виде интервалов, то для расчетов используют середины интервалов X, которые определяются как полсуммы верхней и нижней границ интервала. А если у интервала X отсутствует нижняя или верхняя граница (открытый интервал), то для ее нахождения применяют размах (разность между верхней и нижней границей) соседнего интервала X.

Например, на предприятии 10 работников со стажем работы до 3 лет, 20 - со стажем от 3 до 5 лет, 5 работников - со стажем более 5 лет. Тогда рассчитаем средний стаж работников по формуле средней арифметической взвешенной, приняв в качестве X середины интервалов стажа (2, 4 и 6 лет):

(2*10+4*20+6*5)/(10+20+5) = 3,71 года.

Средняя арифметическая применяется чаще всего, но бывают случаи, когда необходимо применение других видов средних величин. Рассмотрим такие случаи далее.

Средняя гармоническая:

Средняя гармоническая применяется, когда исходные данные не содержат частот f по отдельным значениям X, а представлены как их произведение Xf. Обозначив Xf=w, выразим f=w/X, и, подставив эти обозначения в формулу средней арифметической взвешенной, получим формулу средней гармонической взвешенной:

$$\bar{X}_{гарм.взв.} = \frac{\sum w}{\sum \frac{w}{X}} = \frac{w_1 + w_2 + \ldots + w_N}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \ldots + \frac{w_N}{x_N}}$$

Таким образом, средняя гармоническая взвешенная применяется тогда, когда неизвестны частоты f, а известно w=Xf. В тех случаях, когда все w=1, то есть индивидуальные значения X встречаются по 1 разу, применяется формула средней гармонической простой:

$$\bar{X}_{гарм.прост.} = \frac{1+1+...+1}{\frac{1}{x_1}+\frac{1}{x_2}+...+\frac{1}{x_N}} = \frac{N}{\sum\frac{1}{X}}$$

Например, автомобиль ехал из пункта А в пункт Б со скоростью 90 км/ч, а обратно - со скоростью 110 км/ч. Для определения средней скорости применим формулу средней гармонической простой, так как в примере дано расстояние w1=w2 (расстояние из пункта А в пункт Б такое, же как и из Б в А), которое равно произведению скорости (X) на время (f). Средняя скорость = (1+1)/(1/90+1/110) = 99 км/ч.

Средняя геометрическая:

Средняя геометрическая применяется при определении средних относительных изменений, о чем сказано в теме Ряды динамики. Геометрическая средняя величина дает наиболее точный результат осреднения, если задача стоит в нахождении такого значения X, который был бы равноудален как от максимального, так и от минимального значения X.

$$\bar{X}_{геом} = \sqrt[N]{X_1 \cdot X_2 \cdot ... \cdot X_N} = \sqrt[N]{\prod X}$$

Например, в период с 2012 по 2015 годы индекс инфляции в России составлял: в 2012 году - 1,109; в 2013 - 1,090; в 2014 - 1,119; в 2015 - 1,133. Так как индекс инфляции - это относительное изменение (индекс динамики), то рассчитывать среднее значение нужно по средней геометрической: (1,109*1,090*1,119*1,133)^(1/4) = 1,1126, то есть за период с 2012 по 2015 ежегодно цены росли в среднем на 11,26%. Ошибочный расчет по средней арифметической дал бы неверный результат 11,28%.

Средняя квадратическая:

Средняя квадратическая применяется в тех случая, когда исходные значения X могут быть как положительными, так и отрицательными, например при расчете средних отклонений.

$$\bar{X}_{\text{кв}} = \sqrt{\frac{\sum X^2}{N}}$$

Главной сферой применения квадратической средней является измерение вариации значений X, о чем пойдет речь позднее в этой лекции.

Средняя кубическая:

Средняя кубическая применяется крайне редко, например, при расчете индексов нищеты населения для развивающихся стран (ИНН-1) и для развитых (ИНН-2), предложенных и рассчитываемых ООН.

$$\bar{X}_{\text{куб}} = \sqrt[3]{\frac{\sum X^3}{N}}$$

Структурные средние величины:

К наиболее часто используемым структурным средним относятся статистическая мода и статистическая медиана.

Статистическая мода:

Статистическая мода - это наиболее часто повторяющееся значение величины X в статистической совокупности.

Если X задан дискретно, то мода определяется без вычисления как значение признака с наибольшей частотой. В статистической совокупности бывает 2 и более моды, тогда она считается бимодальной (если моды две) или мультимодальной (если мод более двух), и это свидетельствует о неоднородности совокупности.

Например, на предприятии работает 16 человек: 4 из них - со стажем 1 год, 3 человека - со стажем 2 года, 5 - со стажем 3 года и 4 человека - со стажем 4 года. Таким образом, модальный стаж Мо=3 года, поскольку частота этого значения максимальна (f=5).

Если X задан равными интервалами, то сначала определяется модальный интервал как интервал с наибольшей частотой f. Внутри этого интервала находят условное значение моды по формуле:

$$Mo = X_{HMo} + h_{Mo}\frac{f_{Mo} - f_{Mo-1}}{2f_{Mo} - f_{Mo-1} - f_{Mo+1}}$$

где Мо – мода;

XHMo – нижняя граница модального интервала;

hMo – размах модального интервала (разность между его верхней и нижней границей);

fMo – частота модального интервала;

fMo-1 – частота интервала, предшествующего модальному;

fMo+1 – частота интервала, следующего за модальным.

Например, на предприятии 10 работников со стажем работы до 3 лет, 20 - со стажем от 3 до 5 лет, 5 работников - со стажем более 5 лет. Рассчитаем модальный стаж работы в модальном интервале от 3 до 5 лет: Мо = 3 + 2*(20-10)/(2*20-10-5) = 3,8 (года).

Если размах интервалов h разный, то вместо частот f необходимо использовать плотности интервалов, рассчитываемые путем деления частот f на размах интервала h.

Статистическая медиана:

Статистическая медиана – это значение величины X, которое делит упорядоченную по возрастанию или убыванию статистическую совокупность на 2 равных по численности части. В итоге у одной половины значение больше медианы, а у другой - меньше медианы.

Если X задан дискретно, то для определения медианы все значения нумеруются от 0 до N в порядке возрастания, тогда медиана при четном числе N будет лежать посередине между X с номерами 0,5N и (0,5N+1), а при нечетном числе N будет соответствовать значению X с номером 0,5(N+1).

Например, имеются данные о возрасте студентов-заочников в группе из 10 человек - X: 18, 19, 19, 20, 21, 23, 23, 25, 28, 30 лет. Эти данные уже упорядочены по возрастанию, а их количество N=10 - четное, поэтому медиана будет находиться между X с номерами 0,5*10=5 и (0,5*10+1)=6, которым соответствуют значения X5=21 и X6=23, тогда медиана: Me = (21+23)/2 = 22 (года).

Если X задан в виде равных интервалов, то сначала определяется медианный интервал (интервал, в котором заканчивается одна половина частот f и начинается другая половина), в котором находят условное значение медианы по формуле:

$$Me = X_{HMe} + h_{Me}\frac{0{,}5\sum f - \sum f_{Me-1}}{f_{Me}}$$

где Me – медиана;

XHMe – нижняя граница медианного интервала;

hMe – размах медианного интервала (разность между его верхней и нижней границей);

fMe – частота медианного интервала;

знак суммыfMe-1 – сумма частот интервалов, предшествующих медианному.

В ранее рассмотренном примере при расчете модального стажа (на предприятии 10 работников со стажем работы до 3 лет, 20 - со стажем от 3 до 5 лет, 5 работников - со стажем более 5 лет) рассчитаем медианный стаж. Половина общего числа работников составляет (10+20+5)/2 = 17,5 и находится в интервале от 3 до 5 лет, а в первом интервале до 3 лет - только 10 работников, а в первых двух - (10+20)=30, что больше 17,5, значит интервал от 3 до 5 лет - медианный. Внутри него определяем условное значение медианы: Me = 3+2*(0,5*30-10)/20 = 3,5 (года).

Также как и в случае с модой, при определении медианы если размах интервалов h разный, то вместо частот f необходимо использовать плотности интервалов, рассчитываемые путем деления частот f на размах интервала h.

Среднеквадратическое отклонение измеряется в единицах измерения самой случайной величины и используется при расчёте стандартной ошибки среднего арифметического, при построении доверительных интервалов, при статистической проверке гипотез, при измерении линейной взаимосвязи между случайными величинами. Определяется как квадратный корень из дисперсии случайной величины.

Среднеквадратическое отклонение:

Дисперсия случайной величины – мера разброса данной случайной величины, то есть её отклонения от математического ожидания.

Дисперсия простая:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Стандартное отклонение (оценка среднеквадратического отклонения случайной величины x относительно её математического ожидания на основе несмещённой оценки её дисперсии):

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

где σ^2 – дисперсия; x_i – i-й элемент выборки; n – объём выборки; $\bar{x}$– среднее арифметическое выборки:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i = \frac{1}{n}(x_1 + \ldots + x_n).$$

Следует отметить, что обе оценки являются смещёнными. В общем случае несмещённую оценку построить невозможно. Однако оценка на основе оценки несмещённой дисперсии является состоятельной.

Дисперсия взвешенная:

$$\sigma^2 = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 f_i}{\sum_{i=1}^{k} f_i}$$

Более удобно вычислять дисперсию по формуле:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

Данная формула получается из основной путем несложных преобразований. В этом случае средний квадрат отклонений равен средней из квадратов значений признака минус квадрат средней.

Для этого расчета нужны будут:

$$\{x_{i,} f_i, w_i\}\ n$$

xi –распределение ряда;

fi – абсолютные частоты, значение ряда;

wi – относительная частота, аналог вероятности или рассчитаная вместо fi;

n – количество объектов;

Произведем вычисления среднего квадрата отклонения:

$$Dx = \sum (x_i - x)^2 * w_i = \sum x_i^2 * w_i - 2x \sum x_i w_i + xcp^2 * w_i$$

$\sum$wi = 1 или 100%

$\sum$xwi = x среднее

$\sum x_i^2 * w_i$ = x средний квадрат

Суммирование подразумевается по всем возможными «i». Дисперсия вычисляется в квадратных единицах.

Средний квадрат и квадрат среднего это разные величины.

Пример:

X1 = 2

X2 = 6

X среднее = (x1+x2) / 2= (2+6)/2 = 4 x_{cp}^2 = 16

Средний квадрат = ($x1^2 + x2^2$) / 2 = (4+36) / 2 = 20

Dx= 20-16 = 4 квадратных единицы

Среднее квадратичное отклонение = сигма. В теории вероятностей и статистике наиболее распространённый показатель рассеивания значений случайной величины относительно её математического ожидания.

Среднее квадратичное отклонение = $\sqrt{Dx} = \sqrt{4} = 2$

Очень редко случайная величина отклоняется от своего среднего значения больше чем на 3 сигма. Правило трёх сигм (3σ) – практически все значения нормально распределённой случайной величины лежат в интервале $(\bar{x} - 3\sigma; \bar{x} + 3\sigma)$. Более строго - приблизительно с 0,9973 вероятностью значение нормально распределённой случайной величины лежит в указанном интервале (при условии, что величина $\bar{x}$ истинная, а не полученная в результате обработки выборки).

Если же истинная величина $\bar{x}$ неизвестна, то следует пользоваться не σ, а s. Таким образом, правило трёх сигм преобразуется в правило трёх s.

Применение среднеквадратического отклонения:

- для суждения о колеблемости вариационных рядов и сравнительной оценки типичности (представительности) средних арифметических величин. Это необходимо в дифференциальной диагностике при определении устойчивости признаков;

- для реконструкции вариационного ряда, т.е. восстановления его частотной характеристики на основе правила "трех сигм";

- для выявления "выскакивающих" вариант (при сопоставлении реального и реконструированного вариационных рядов);

- для определения параметров нормы и патологии с помощью сигмальных оценок; о для расчета коэффициента вариации;

- для расчета средней ошибки средней арифметической величины.

График плотности вероятности нормального распределения и процент попадания случайной величины на отрезки, равные среднеквадратическому отклонению (рис.1.):

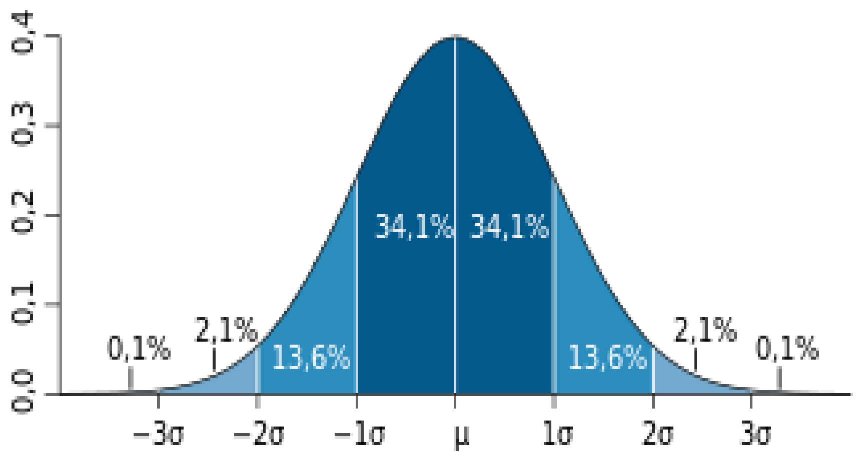


Рис.1. График плотности нормального распределения

Коэффициент вариации (С) – это процентное отношение среднеквадратического отклонения к среднеарифметической величине:

$$C = \sigma / M \times 100\%.$$

Коэффициент вариации – это относительная мера колеблемости вариационного ряда.

Применение коэффициента вариации:

- сильное разнообразие ряда свидетельствует о малой представительности (типичности) соответствующей средней величины и, следовательно, о нецелесообразности ее использования в практических целях;

- для сравнительной оценки разнообразия (колеблемости) разноименных вариационных рядов и выявления более и менее стабильных признаков, что имеет значение в дифференциальной диагностике.

РАСЧЕТ ПОКАЗАТЕЛЕЙ ВАРИАЦИОННОГО РЯДА, ИСПОЛЬЗУЯ МАСТЕР ФУНКЦИЙ (MS EXCEL)

Пример 1. Рассмотрим размер заработной платы у 20 работников и рассчитаем средние величины. Введем в ячейку А1 – А20 значения величины заработной платы рабочих: 14500,15200,17000, 16300, 15100, 14900, 17100, 19400, 20000, 18700, 16500, 17000, 15500, 14700, 13000, 12000, 11000, 14000, 16800, 19000.

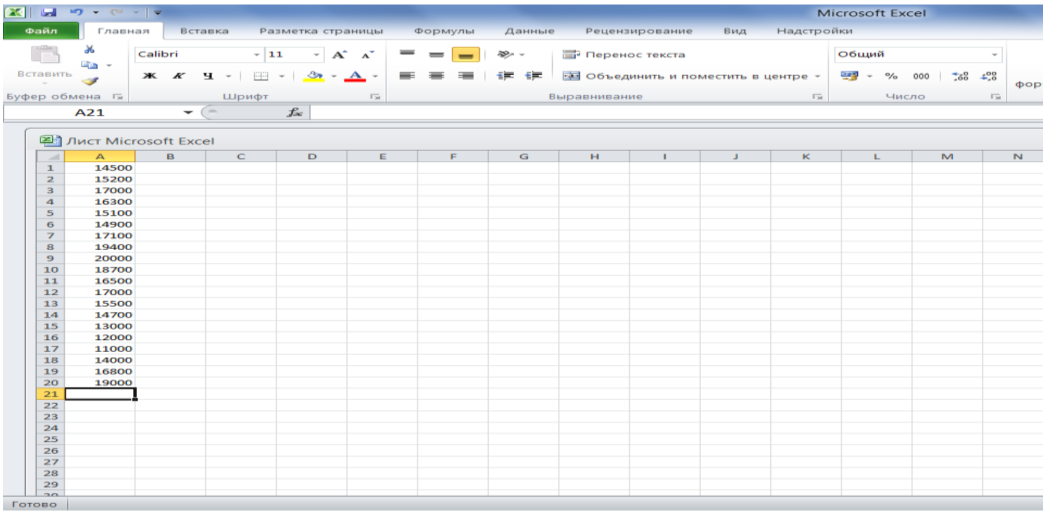


Рис. 2 Ввод данных

Для расчета Моды, открываем Мастер функций. Для этого определяем курсором ячейку В22, щелкаем мышью на кнопке в строке состояния и открывается окно Мастер функций. В категории выбираем Статистические, в перечне функций находим функцию Мода и нажимаем ОК (Рис. 3, 4).

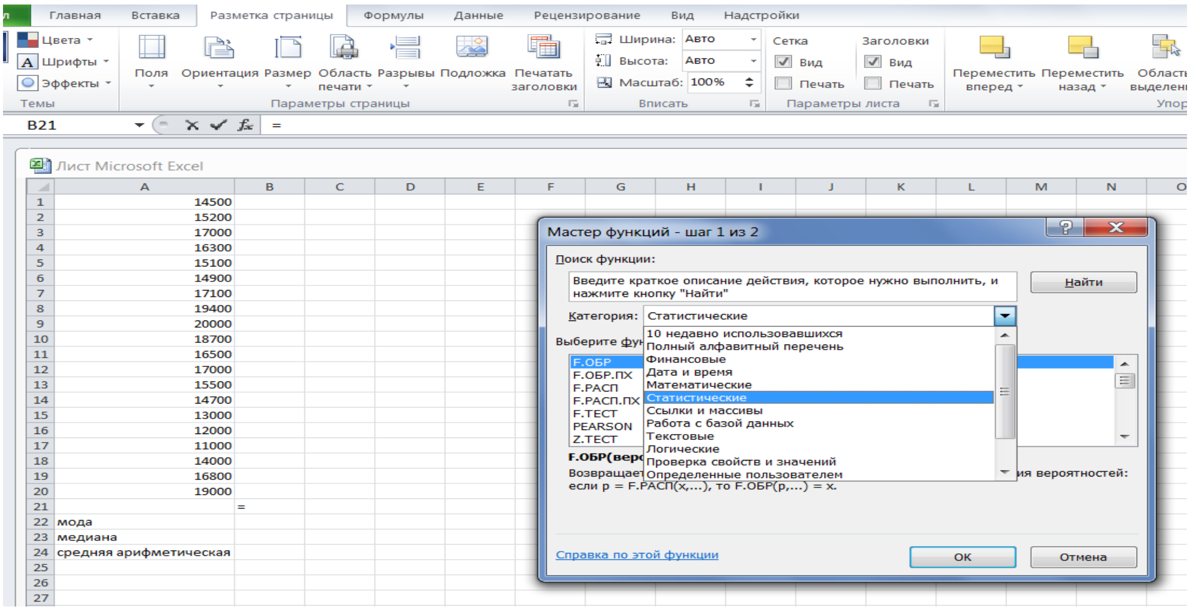


Рис. 3 Выбор категории

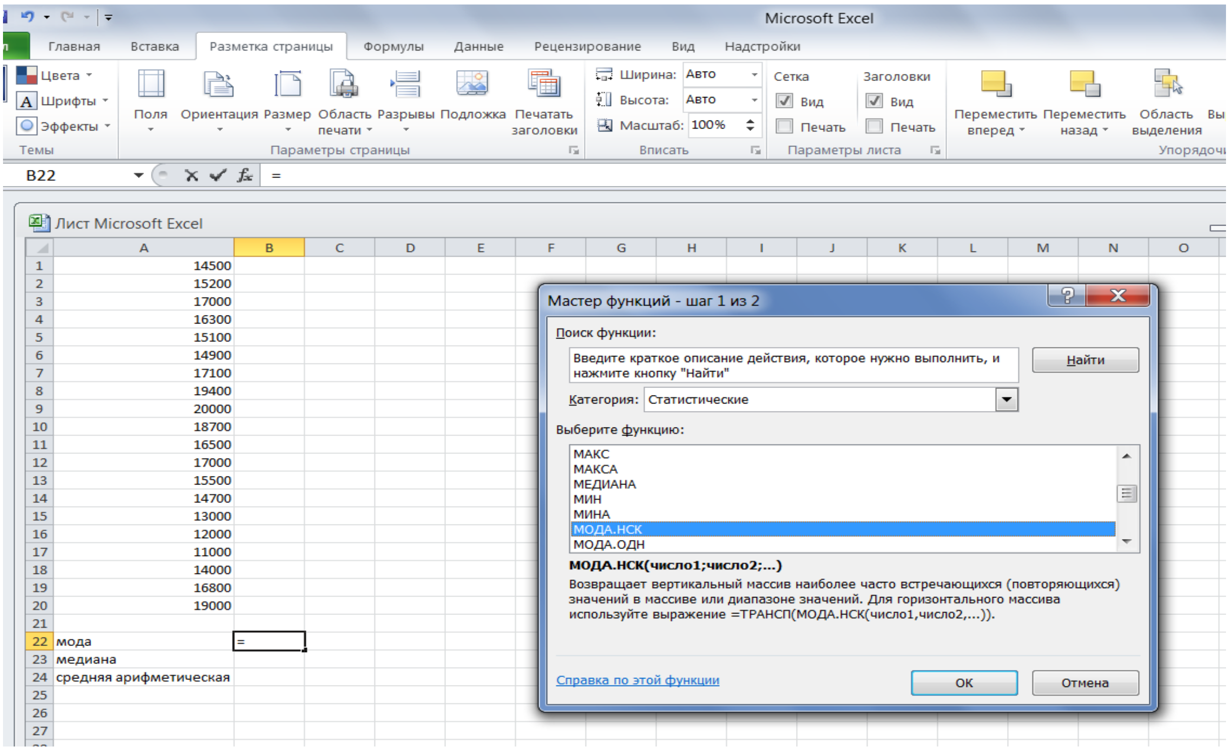


Рис. 4 Выбор функции Мода

В появившемся диалоговом окне Аргументы функции определяем массив А1:А20 и нажимаем ОК (Рис. 5).

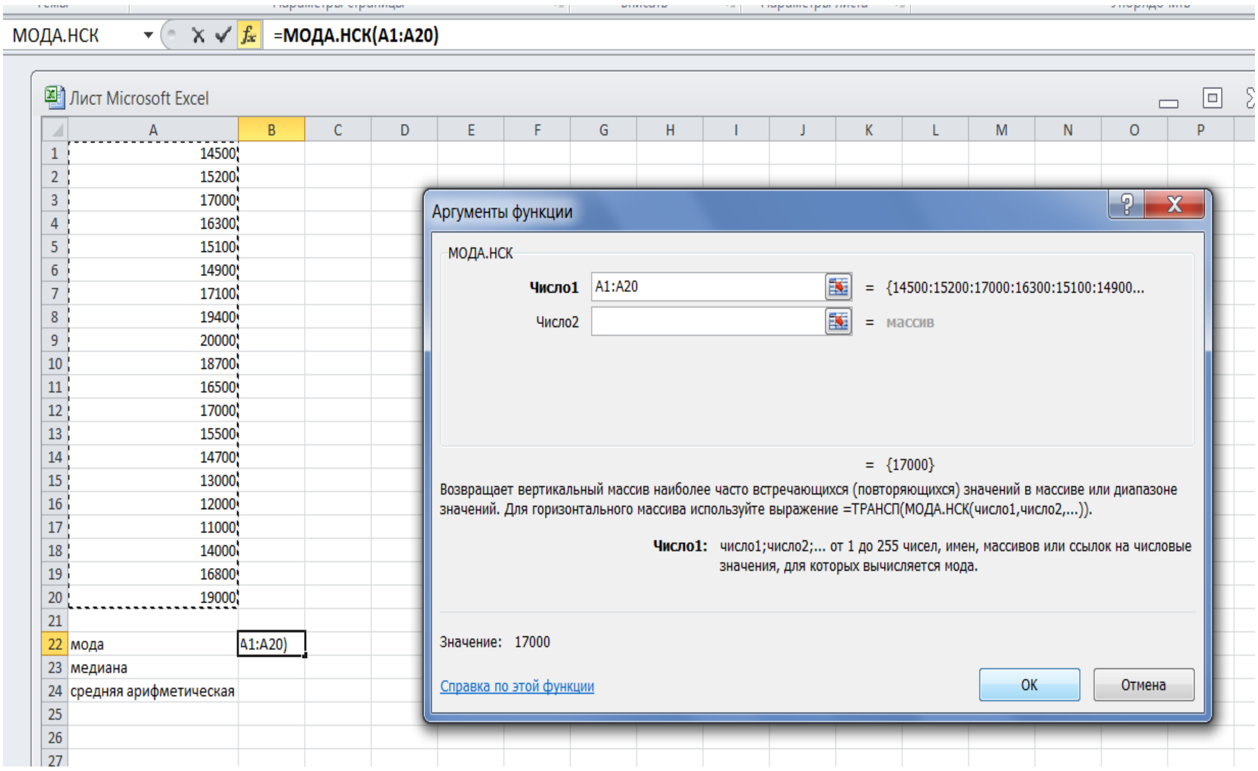


Рис. 5 Выбор массива

В нашем примере Мода равна 17000.

Аналогичным способом вычисляем Медиану и среднюю арифметическую (Рис. 6 - 9).

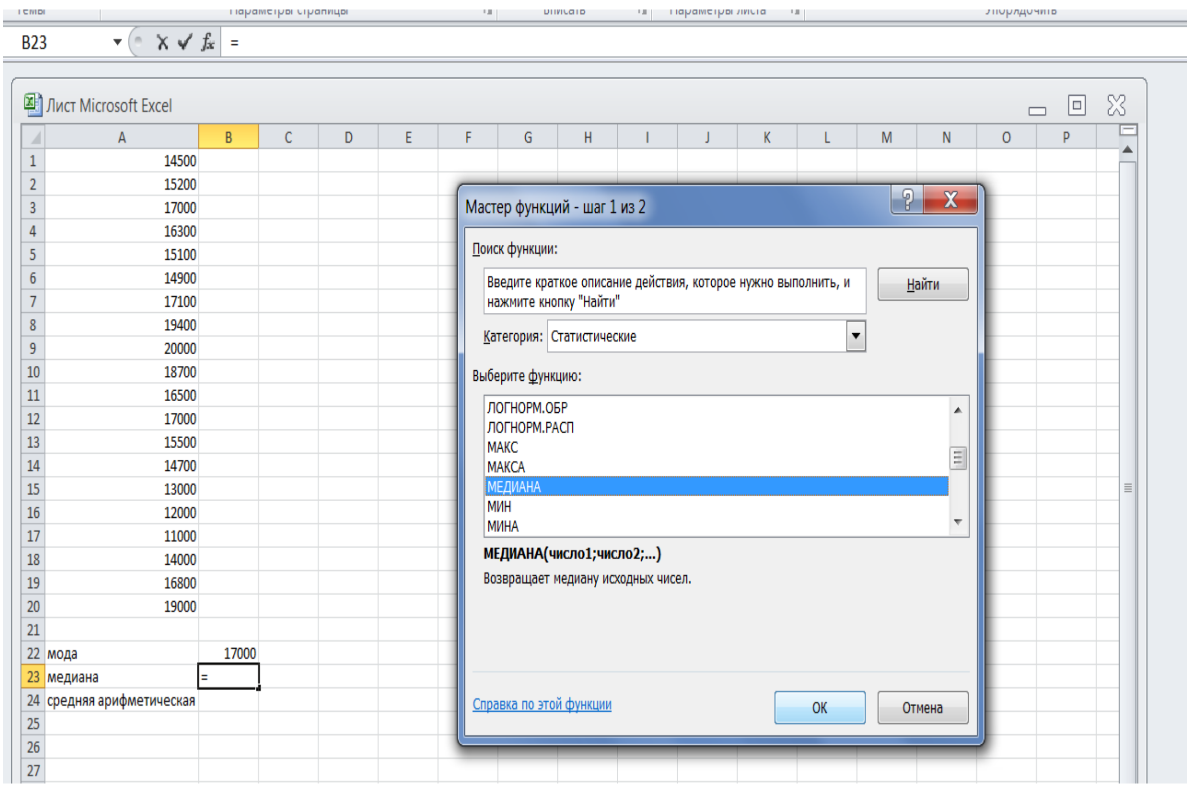


Рис. 6 Выбор функции Медиана.

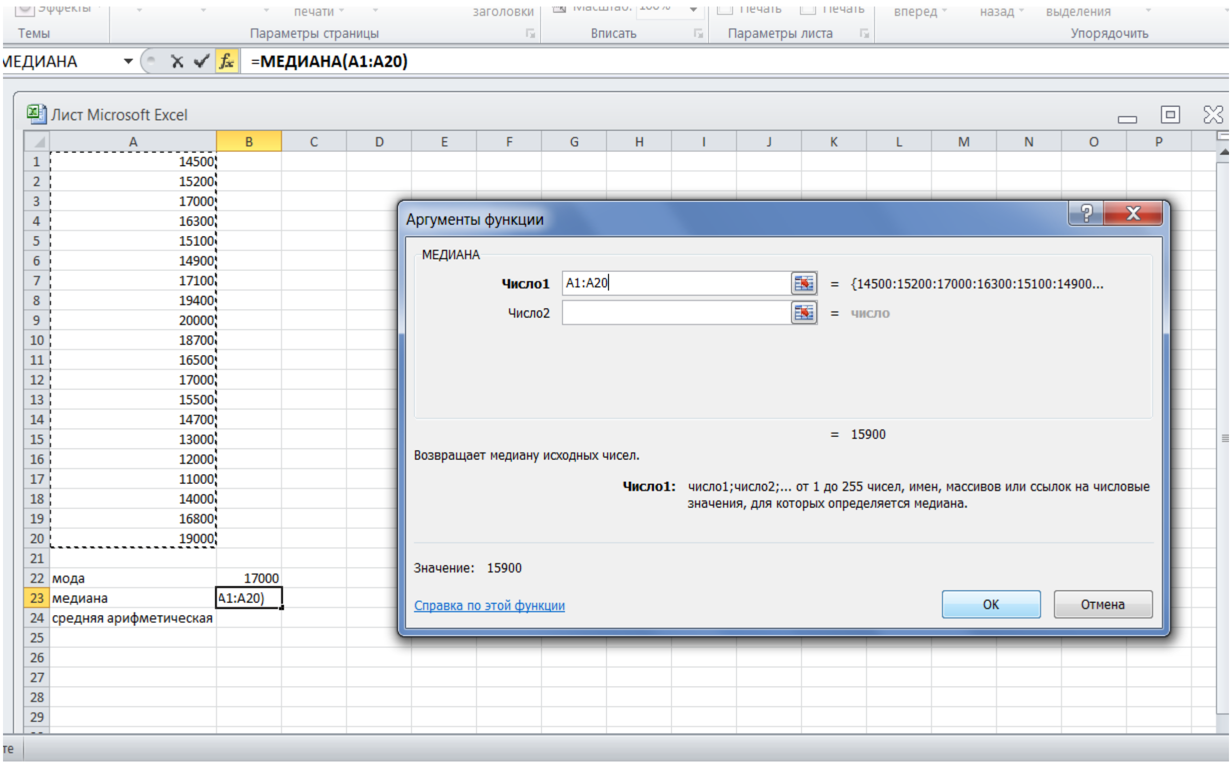


Рис. 7 Выбор массива.

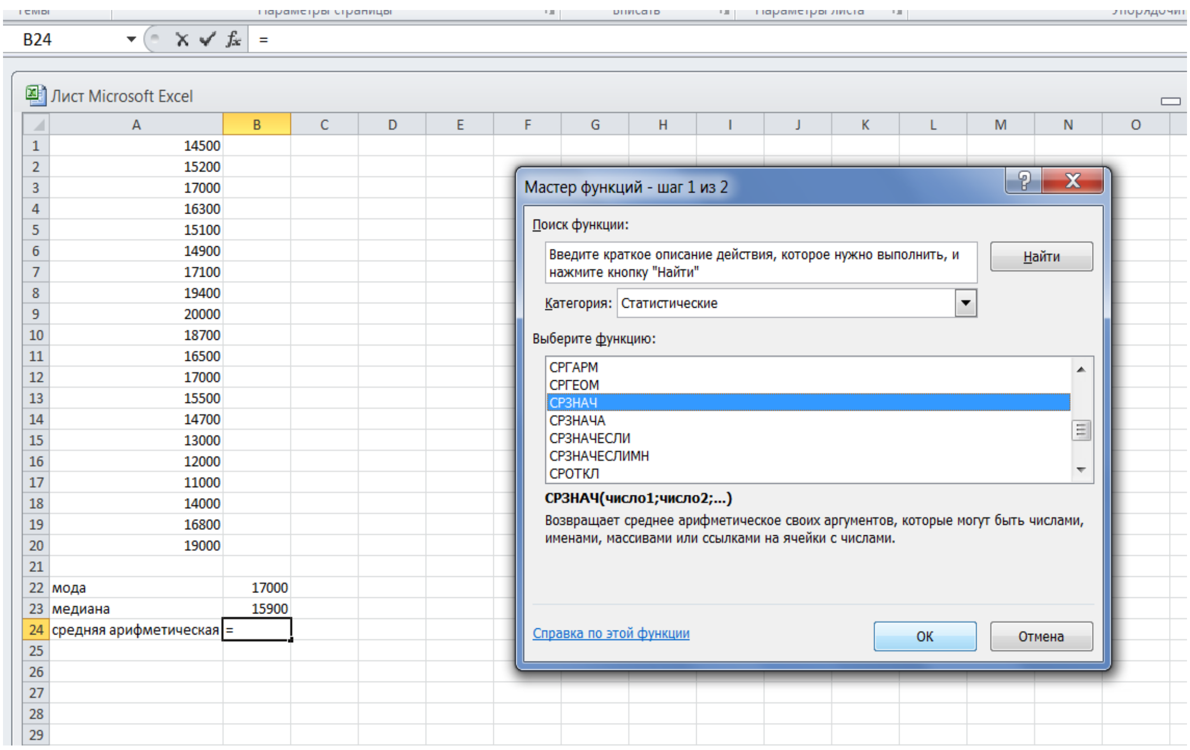


Рис. 8 Выбор функции Срзнач.

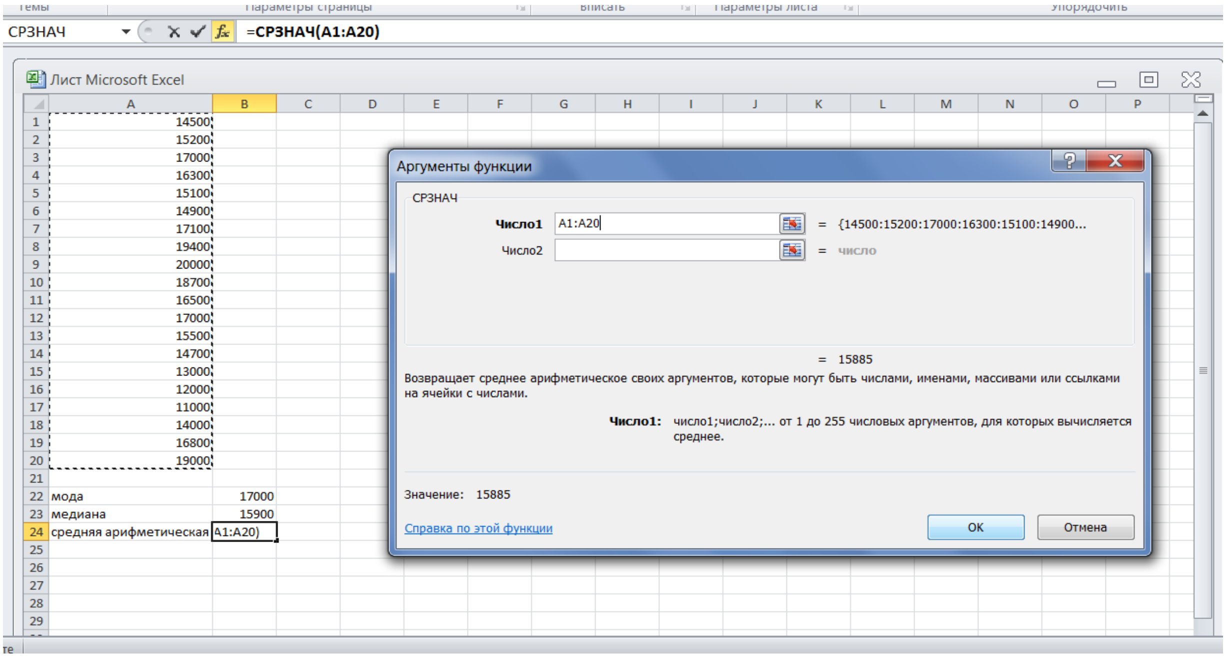


Рис. 9 Выбор массива.

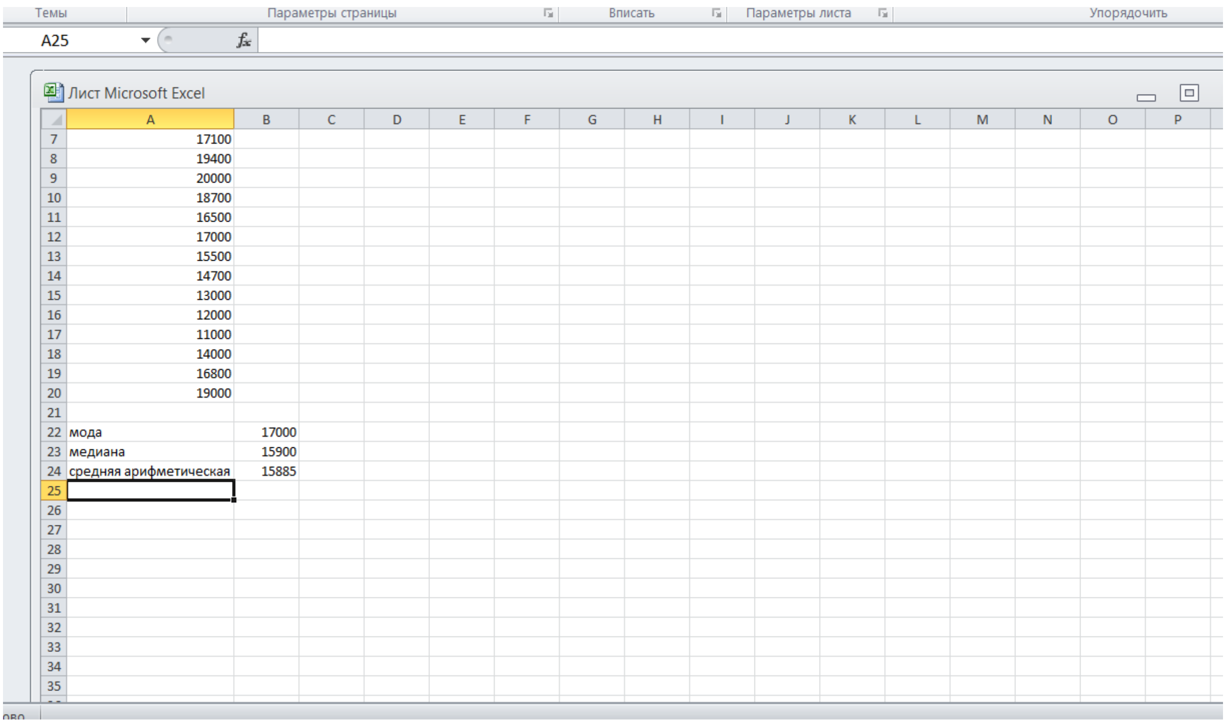


Рис. 10 Результат анализа.

Итого, Медиана равна 15900, а средняя арифметическая – 15885 (Рис. 10).

Среднее значение или арифметическое среднее наиболее широко используется в статистике. Это одно значение может использоваться для представления некоторого набора данных. В этом случае среднее значение можно назвать "центром тяжести" этого набора. Среднее значение вычисляется следующим образом: складываются все значения выборки и результат делится на общее число значений. Средние величины позволяют сравнивать уровни одного и того же признака в различных совокупностях и находить причины этих расхождений.

Для характеристики структуры статистической совокупности применяются показатели, которые называют структурными средними. К ним относятся мода и медиана.

Мода (Мо) – чаще всего встречающийся вариант. Модой называется значение признака, которое соответствует максимальной точке теоретической кривой распределений.

Мода представляет наиболее часто встречающееся или типичное значение. Мода применяется в коммерческой практике для изучения покупательского спроса и регистрации цен.

В дискретном ряду мода – это варианта с наибольшей частотой. В интервальном вариационном ряду модой считают центральный вариант интервала, который имеет наибольшую частоту (частность).

Медиана (Ме – это величина, которая делит численность упорядоченного вариационного ряда на две равные части: одна часть имеет значения варьирующего признака меньшие, чем средний вариант, а другая – большие. Медиана – это элемент, который больше или равен и одновременно меньше или равен половине остальных элементов ряда распределения.

Свойство медианы заключается в том, что сумма абсолютных отклонений значений признака от медианы меньше, чем от любой другой величины. Применение медианы позволяет получить более точные результаты, чем при использовании других форм средних.

Порядок нахождения медианы в интервальном вариационном ряду следующий: располагаем индивидуальные значения признака по ранжиру; определяем для данного ранжированного ряда накопленные частоты; по данным о накопленных частотах находим медианный интервал:

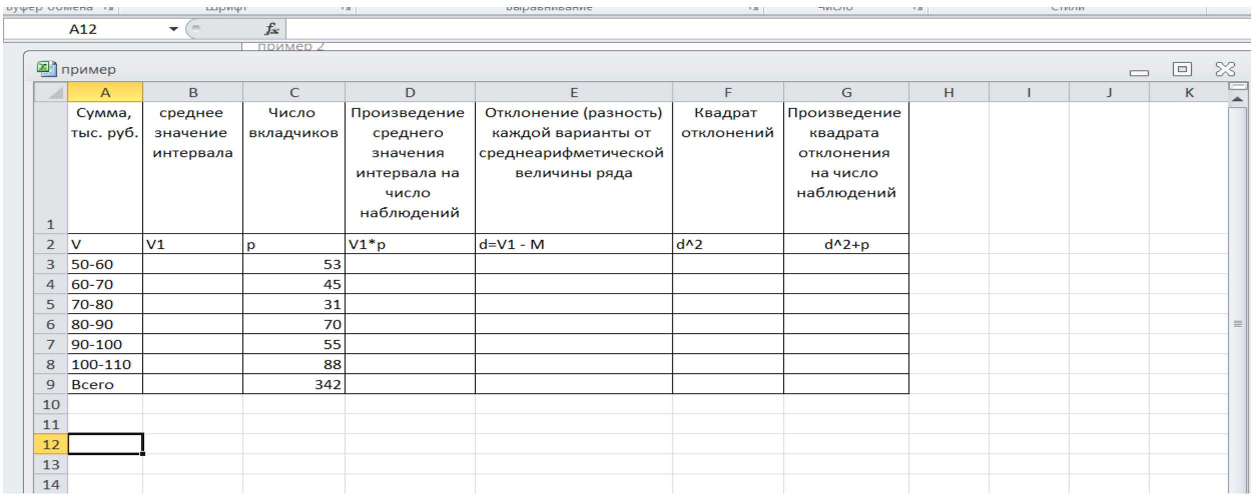
Пример 2. Определить среднюю сумму вкладов на депозит в банк, граждан, среднеквадратическое отклонение, коэффициент вариации.

Таблица 2

Входные данные

Сумма, тыс.руб. (V)	Количество человек (p)
50-60	53
60-70	45
70-80	31
80-90	70
90-100	55
100-110	88
Всего	342

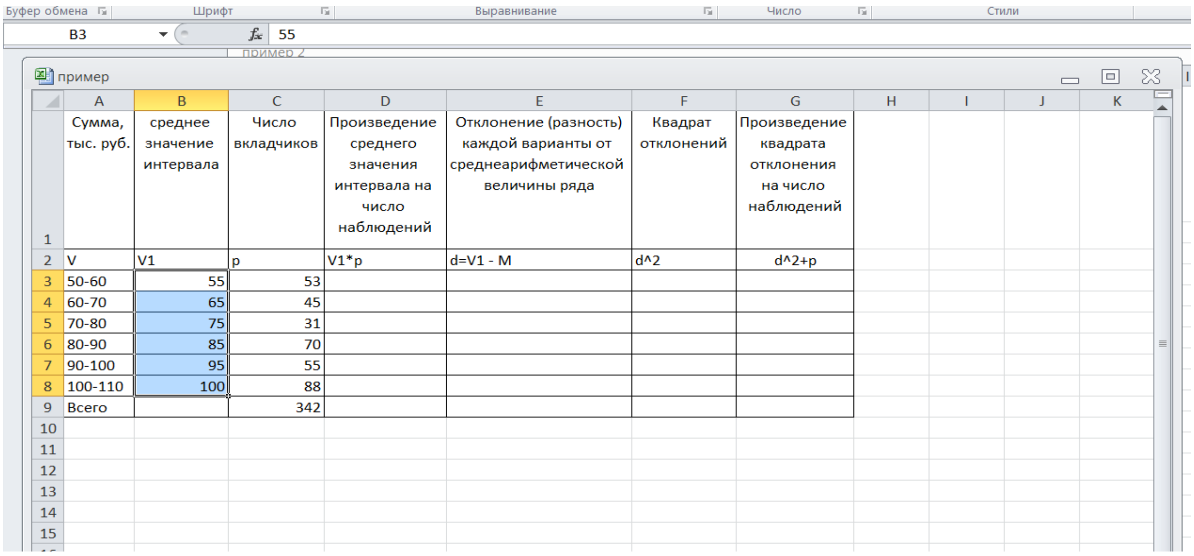
Введем данные в таблицу Excel (Рис. 11).



A	B	C	D	E	F	G
Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
V	V1	p	V1*p	d=V1 - M	d^2	d^2+p
50-60		53				
60-70		45				
70-80		31				
80-90		70				
90-100		55				
100-110		88				
Всего		342				

Рис. 11 Ввод данных для расчета

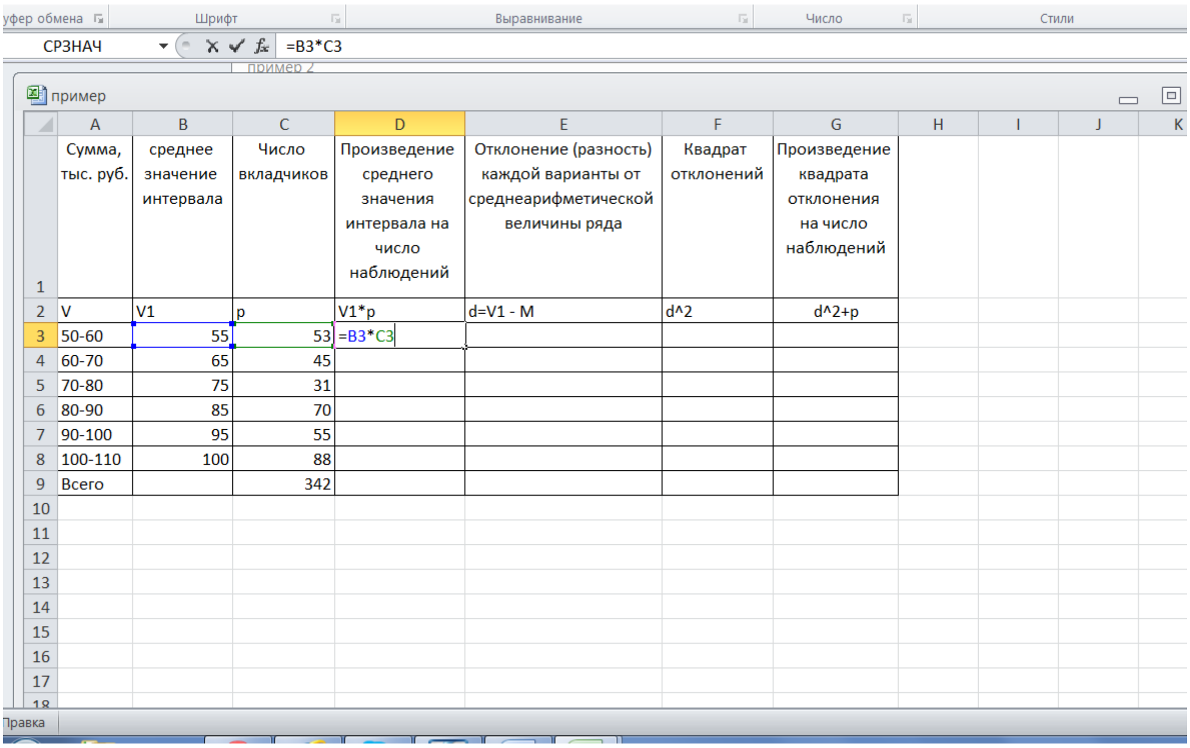
В ячейки B3:B8 внесем показатели центров интервалов роста (V1) (Рис. 12).



A	B	C	D	E	F	G
Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
V	V1	p	V1*p	d=V1 - M	d^2	d^2+p
50-60	55	53				
60-70	65	45				
70-80	75	31				
80-90	85	70				
90-100	95	55				
100-110	100	88				
Всего		342				

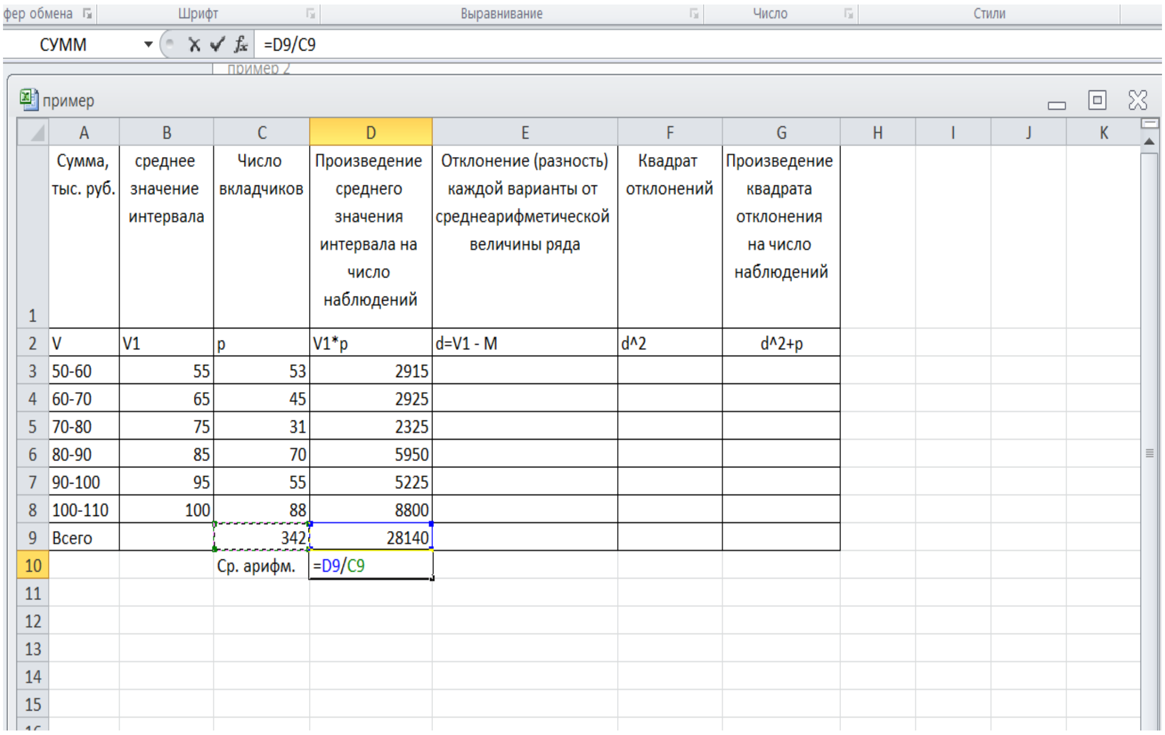
Рис. 12 Ввод среднего значения интервала

В ячейку D3 вводим формулу = B3*C3 (Рис. 13), копируем на остальные ячейки D4:D8 и полученные данные суммируем в ячейке D9. А в ячейку D10 введем формулу расчета средней арифметической = D9/C9, после чего и получаем искомую величину взвешенной средней арифметической (Рис. 14).



	A	B	C	D	E	F	G
1	Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
2	V	V1	p	V1*p	d=V1 - M	d^2	d^2+p
3	50-60	55	53	=B3*C3			
4	60-70	65	45				
5	70-80	75	31				
6	80-90	85	70				
7	90-100	95	55				
8	100-110	100	88				
9	Всего		342				

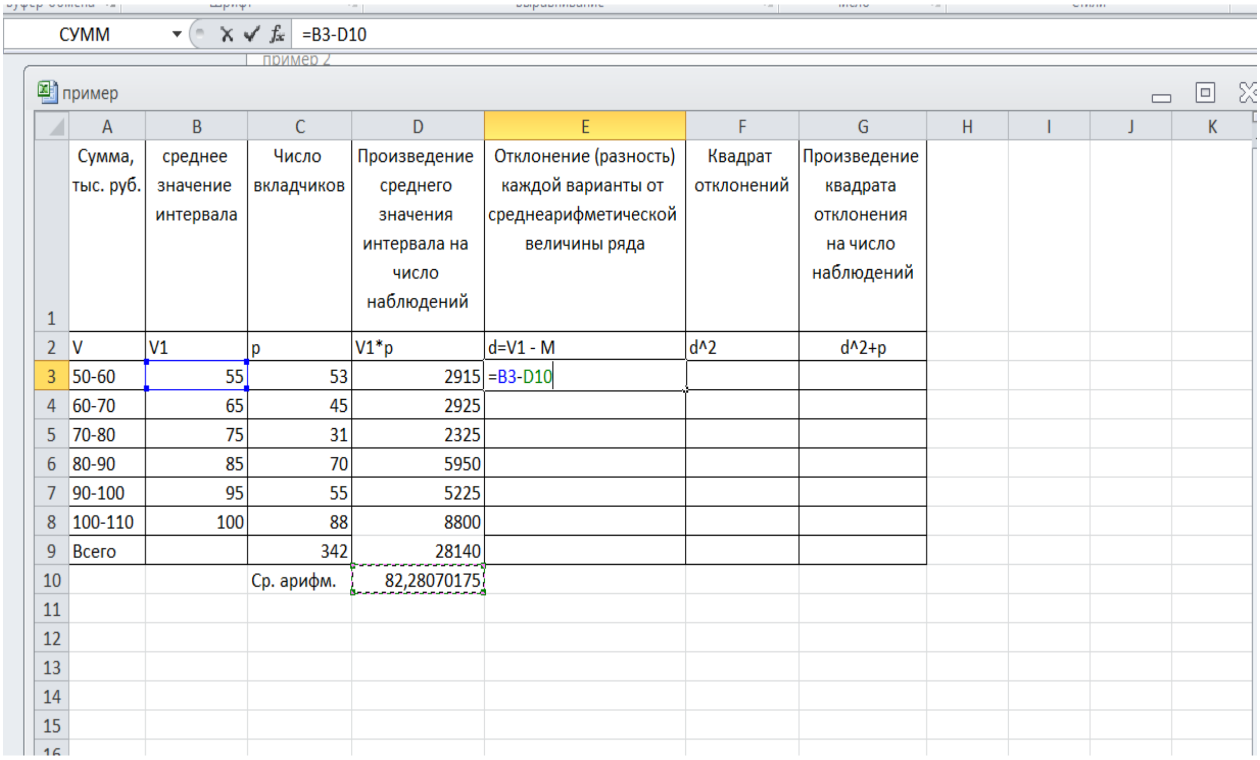
Рис. 13 Ввод формулы



	A	B	C	D	E	F	G
1	Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
2	V	V1	p	V1*p	d=V1 - M	d^2	d^2+p
3	50-60	55	53	2915			
4	60-70	65	45	2925			
5	70-80	75	31	2325			
6	80-90	85	70	5950			
7	90-100	95	55	5225			
8	100-110	100	88	8800			
9	Всего		342	28140			
10			Ср. арифм.	=D9/C9			

Рис. 14 Расчет средней арифметической величины

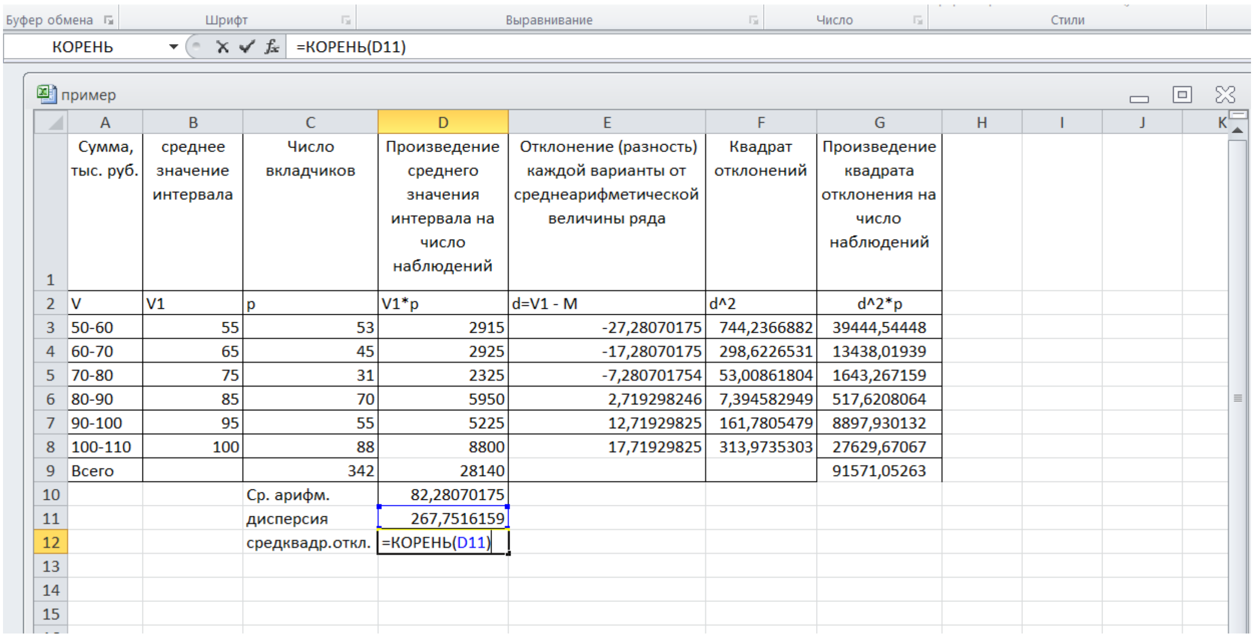
Для вычисления отклонения (разности) каждой варианты от средней арифметической величины вводим формулу в ячейку E3=B3-D10 (Рис. 15) и аналогично в остальные ячейки E4:E8. Находим квадрат отклонения и умножаем полученное число на число наблюдений.



	A	B	C	D	E	F	G
1	Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
2	V	V1	p	V1*p	d=V1 - M	d^2	d^2+p
3	50-60	55	53	2915	=B3-D10		
4	60-70	65	45	2925			
5	70-80	75	31	2325			
6	80-90	85	70	5950			
7	90-100	95	55	5225			
8	100-110	100	88	8800			
9	Всего		342	28140			
10			Ср. арифм.	82,28070175			

Рис. 15 Ввод формулы.

В ячейке G9 суммируем полученные произведения квадрата отклонения на число наблюдений, делим на общее число наблюдений и находим дисперсию (D11 =G9/C9). В ячейке D12 вводим формулу расчета среднеквадратического отклонения – извлечение квадратного корня из величины дисперсии (Рис 16). Для получения квадратного корня щелкаем мышью на кнопке в строке состояния, в открывшемся окне Мастера функций выбираем категорию Математические, находим функцию Корень. Нажимаем кнопку ОК и в открывшемся диалоговом окне указываем ячейку G9, из которой следует извлечь квадратный корень.

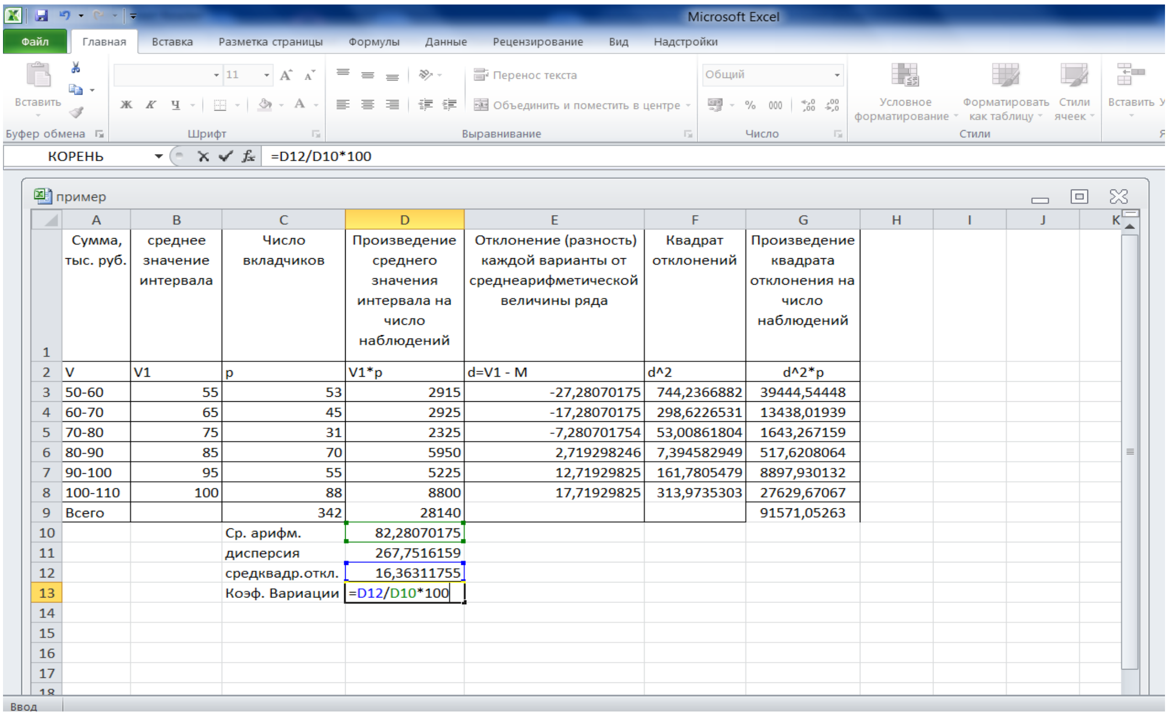


КОРЕНЬ =КОРЕНЬ(D11)

	A	B	C	D	E	F	G
1	Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
2	V	V1	p	V1*p	d=V1 - M	d^2	d^2*p
3	50-60	55	53	2915	-27,28070175	744,2366882	39444,54448
4	60-70	65	45	2925	-17,28070175	298,6226531	13438,01939
5	70-80	75	31	2325	-7,280701754	53,00861804	1643,267159
6	80-90	85	70	5950	2,719298246	7,394582949	517,6208064
7	90-100	95	55	5225	12,71929825	161,7805479	8897,930132
8	100-110	100	88	8800	17,71929825	313,9735303	27629,67067
9	Всего		342	28140			91571,05263
10			Ср. арифм.	82,28070175			
11			дисперсия	267,7516159			
12			среднквадр.откл.	=КОРЕНЬ(D11)			

Рис. 16 Ввод формулы

Для получения коэффициента вариации в ячейке D13 вводим формулу =D12/D10*100 (Рис. 17).



КОРЕНЬ =D12/D10*100

	A	B	C	D	E	F	G
1	Сумма, тыс. руб.	среднее значение интервала	Число вкладчиков	Произведение среднего значения интервала на число наблюдений	Отклонение (разность) каждой варианты от среднеарифметической величины ряда	Квадрат отклонений	Произведение квадрата отклонения на число наблюдений
2	V	V1	p	V1*p	d=V1 - M	d^2	d^2*p
3	50-60	55	53	2915	-27,28070175	744,2366882	39444,54448
4	60-70	65	45	2925	-17,28070175	298,6226531	13438,01939
5	70-80	75	31	2325	-7,280701754	53,00861804	1643,267159
6	80-90	85	70	5950	2,719298246	7,394582949	517,6208064
7	90-100	95	55	5225	12,71929825	161,7805479	8897,930132
8	100-110	100	88	8800	17,71929825	313,9735303	27629,67067
9	Всего		342	28140			91571,05263
10			Ср. арифм.	82,28070175			
11			дисперсия	267,7516159			
12			среднквадр.откл.	16,36311755			
13			Коэф. Вариации	=D12/D10*100			

Рис. 17 Ввод формулы

Выводы:

1. Средний размер вкладов составляет 82,28 тыс. руб.;
2. σ = ± 16,36 (руб.).
3. Величина коэффициента вариации, равная 19,88 % свидетельствует о незначительном разнообразии признака.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что такое вариационный ряд?
2. Для чего используются средние величины?
3. По каким критериям можно оценить разнообразие признака?
4. В каких случаях применяют среднеквадратическое отклонение?
5. Каково назначение коэффициента вариации?
6. Как оценить величину коэффициента вариации?
7. Правило нормального распределения Гаусса.

ТЕСТОВЫЕ ЗАДАНИЯ:

1. К какому классу относятся средняя арифметическая, средняя гармоническая, средняя геометрическая?

А) к классу структурных средних

Б)к классу порядковых средних

В)к классу степенных средних

Г)к классу промежуточных средних.

2. Понятие средней величины.

А)обобщающий показатель, характеризующий структурные сдвиги

Б)частный показатель, характеризующий индекс цен

В)частный показатель, характеризующий развитие явления

Г)обобщающий показатель, характеризующий типичный уровень явления в конкретных условиях места и времени

3. Основное условие правильного расчета средних величин.

А)рассчитываются для качественно однородных совокупностей

Б)рассчитываются для разнокачественных совокупностей

В)рассчитываются для качественно не однородных совокупностей

Г)рассчитываются для разнородных совокупностей по существенным признакам

4. Статистический ряд распределения это…

А)бессистемное распределение единиц изучаемой совокупности

Б)упорядоченное распределение единиц изучаемой совокупности на группы по определенному варьирующему признаку

В)хаотичное распределение единиц изучаемой совокупности

Г)упорядоченное распределение единиц изучаемой совокупности по баллам

5. Дисперсия представляет собой:

А) разность между максимальным и минимальным значениями группировочного признака

Б) среднюю арифметическую абсолютных значений отдельных вариантов от их средней арифметической

В) средний квадрат отклонений вариантов от их средней величины

6. Медианой называется…

А) среднее значение признака в ряду распределения;

Б) наиболее часто встречающееся значение признака в данном ряду;

В) значение признака, делящее совокупность на две равные части;

Г) наиболее редко встречающееся значение признака в данном ряду.

7. Модой называется…

А) наиболее часто встречающееся значение признака в данном ряду;

Б) значение признака, делящее данную совокупность на две равные части;

В) наиболее редко встречающееся значение признака в данном ряду

Г) среднее значение признака в данном ряду распределения.

8. К относительным показателям вариации относятся...

А) размах вариации

Б) дисперсия

В) коэффициент вариации

Г) среднее линейное отклонение

9. Для следующих значений признака: 3, 3, 3, 4, 4, 6, 7, 9, 9 мода…

А) отсутствует

Б) равна 3

В) равна 9

Г) равна 4

10. Если все значения признака увеличить в 16 раз, то дисперсия …

А) увеличится в 16 раз

Б) увеличится в 4 раза

В) не изменится

Г) увеличится в 256 раз

11. Вариационный ряд - это ряд распределения, построенный по ... признаку

А) количественному

Б) качественному

В) количественному и качественному

Г) непрерывному

12. При увеличении всех значений признака в 2 раза средняя арифметическая ...

А) не изменится

Б) увеличится в 2 раза

В) уменьшится в 2 раза

Г) увеличится более чем в 2 раза

13. При уменьшении значений частот в средней арифметической взвешенной в 2 раза значение средней величины признака ...

А) не изменится

Б) уменьшится в 2 раза

В) увеличится в 2 раза

Г) увеличится более чем в 2 раза

14. Сумма отклонений индивидуальных значений признака от их средней величины…

А) больше нуля

Б) меньше нуля

В) равна нулю

Г) больше или равна нулю

15. Средняя величина признака равна 20, а коэффициент вариации 25%. Дисперсия признака равна

А) 20

Б) 25

В) 125

Г) 45

16. К относительным показателям вариации относятся...

А) размах вариации

Б) среднее линейное отклонение

В) дисперсия

Г) коэффициент вариации

17. Для значений признака: 3, 5, 6, 9, 11, 12, 13 Мода…

А) отсутствует

Б) равна 11

В) равна 9

Г) равна 3

18. Средний квадрат отклонений вариантов от средней величины – это:

А) размах вариации

Б) коэффициент вариации

В) дисперсия

Г) среднее квадратическое отклонение

19. Средний квадрат индивидуальных значений признака равен 625, а его дисперсия - 400. Величина средней равна:

А) 15

Б) 25

В) 80

Г) 40

20. Абсолютные величины могут выражаться в…

А) натуральных единицах измерения

Б) процентах

В) денежных единицах измерения

Г) виде простого кратного отношения

Ответы к тестам:

1- в; 2-г; 3-а; 4-б; 5-в; 6-в; 7-а; 8-в; 9-б; 10-г; 11-а; 12-б; 13-а; 14-в; 15-б; 16-г; 17-а; 18-в; 19-а; 20-а,в.

ЗАДАЧИ НА САМОСТОЯТЕЛЬНОЕ ВЫПОЛНЕНИЕ:

Задача 1. Определить среднюю сумму вкладов на депозит в банк, граждан, среднеквадратическое отклонение, коэффициент вариации.

Сумма, тыс.руб. (V)	Количество человек (р)
35-45	12
45-55	45
55-65	65
65-75	38
75-85	67
Всего	…

Вариант	Р
1	20,34, 56,78,43
2	45,23,19,67,80
3	56,23,48,59,36
4	45,76,36,13,46
5	12,34,65,27,48
6	56,76,17,34,49
7	23,41,60,19,59
8	47,28,38,71,65
9	56,23,72,77,22
10	44,76,13,37,75
11	32,67,33,76,82
12	45,13,47,86,37
13	43,64,86,96,13
14	36,15,65,86,27
15	55,86,35,18,73
16	44,76,36,66,31
17	65,31,75,49,10
18	48,81,37,55,21
19	35,80,47,33,64
20	55,21,16,23,41

21	34,28,50,12,11
22	27,48,22,53,24
23	60,37,67,21,32
24	34,56,70,32,12
25	55,43,29,28,33
26	30,44,53,70,60
27	78,83,89,72,19
28	99,45,66,87,54
29	23,34,88,54,34
30	65,87,95,44,30

Задача 2. Рассмотрим заработную плату на предприятии и рассчитаем средние величины.

Вариант	**з/п**
1	13000,20100,15500,10300,23400,16700,19000,17220,14250
2	19500,21400,21000,18900,17000,15700,15300,14900,17250
3	15000,13800,20200,17800,13900,17400,16450,19200,15000
4	17600,18900,14900,17000,14300.18900,15300,14770,13650
5	16700,25700,23800,25780,24700,23490,27000,19050,16800
6	18500,17870,13590,16390,12800,17640,15900,12380,20000
7	19500,21400,21000,18900,24700,23490,27000,19050,16800
8	12800,17640,15900,12380,20000,17600,18900,14900,17000
9	13000,20100,15500,10300,23400,16700,15900,12380,20000
10	17000,15700,15300,14900,17250,10300,23400,14900,17000
11	21400,21000,18900,24700,23400,16700,15900,12380,20000
12	10300,23400,16700,10300,23400,16700,15900,12380,20000
13	16700,10300,23400,16700,15900,12380,20000,13590,16390
14	15300,14900,17250,10300,23400,14900,17000,16700,10300
15	20100,15500,10300,10300,23400,16700,15900,13590,16390
16	20100,15500,10300,23400,16700,15900,12380,23400,16700
17	21400,21000,18900,24700,23400,16700,15900,12380,23400
18	10300,10300,23400,16700,15900,17250,21000,18900,15300
19	23400,16700,15900,17250,12380,20000,13590,16390,23450
20	16200,13500,18000,16700,14300,15000,15500,18000,17500
21	14100,13500,16600,15000,16500,20000,12700,18100,18200
22	16700,15900,17220,14250,16700,15900,18900,14900,16400
23	12380,20000,12500,18900,16200,17800,12300,20000,17000
24	16000,14000,20000,13300,16500,14000,17800,13900,21000
25	17000,14300.18900,23400,16700,13800,20500,22000,23100
26	13900,17400,15900,12380,18500,17870,15700,15300,23000
27	23800,25780,15900,12380,15300,14900,17640,15900,22400
28	15500,10300,23400,16200,13500,20200,17800,21000,18900
29	15000,15500,18000,16700,10300,21400,21000,18900,22000
30	15900,12380,20000,17800,13900,12380,20000,13590,21100

Задача 3. Имеются следующие данные о распределении предприятий по стоимости основных средств. Определить моду и медиану по вариантам, среднее значение и дисперсию. Сделать выводы по полученным результатам. Необходимо использовать подряд 10 строк, начиная от номера варианта.

Вариант	Стоимость ОС, млн. руб.	Количество предприятий, %

1	До 500	10
2	500-550	12
3	550-600	19
4	600-650	31
5	650-700	23
6	700-750	45
7	750-800	54
8	800-850	19
9	850-900	21
10	900-950	47
11	950-1000	66
12	1000-1050	79
13	1050-1100	33
14	1100-1150	47
15	1150-1200	65
16	1200-1250	15
17	1250-1300	20
18	1350-1400	15
19	1400-1450	17
20	1500-1550	40
	1550-1600	55
	1650-1700	43
	1750-1800	77
	1800-1850	75
	1850-1900	43
	1900-1950	21
	1950-2000	33
	2000-2050	67
	2050-2100	86
	2100-2150	43

Задача 4. Имеются данные о распределении рабочих предприятий по выполнению норм выработки. Определить моду и медиану по вариантам, среднее значение и дисперсию. Сделать выводы по полученным результатам и представить распределение графически. Необходимо использовать подряд 10 строк, начиная от номера варианта.

Вариант	Выполнение норм выработки, %	Численность рабочих
1	10-15	14
2	15-20	18
3	20-25	22
4	25-30	20
5	30-35	43
6	35-40	24
7	40-45	57
8	45-50	45
9	50-55	32
10	55-60	53
11	60-65	64
12	65-70	22
13	70-75	56

14	75-80	78
15	80-85	32
16	85-90	78
17	90-95	86
18	95-100	43
19	100-105	23
20	105-110	42
	110-115	12
	115-120	32
	120-125	21
	125-130	31
	130-135	43
	135-140	34
	140-145	56
	145-150	89
	150-155	76
	155-160	44

Литература:

1. Капитоненко В.В. Учебник- М.: Финансы и статистика, 2007. -256 с.

2. Цымбаленко Т.Т. Статистика финансов в АПК [Электронный ресурс]: учебное пособие/ Цымбаленко Т.Т., Цымбаленко С.В., Герасимов А.Н.— Электрон. текстовые данные.— М.: Финансы и статистика, 2014.— 160 с.

3. Тимофеева Т.В. Практикум по финансовой статистике [Электронный ресурс]: учебное пособие/ Тимофеева Т.В., Снатенков А.А.— Электрон. текстовые данные.— М.: Финансы и статистика, 2014.— 320 с.— Режим доступа: http://www.iprbookshop.ru/18830.— ЭБС «IPRbooks»

4. Шерстнева Г.С. Учебное пособие по финансовой статистике [Электронный ресурс]/ Шерстнева Г.С.— Электрон. текстовые данные.— Саратов: Научная книга, 2012.— 159 с.— Режим доступа: http://www.iprbookshop.ru/6274.— ЭБС «IPRbooks»

5. Цымбаленко Т.Т. Статистика финансов в АПК [Электронный ресурс]: учебное пособие/ Цымбаленко Т.Т., Цымбаленко С.В., Герасимов А.Н.— Электрон. текстовые данные.— М.: Финансы и статистика, 2014.— 160 с.

6. Волкова Н.В. Статистика [Электронный ресурс]: учебное пособие/ Волкова Н.В., Каурова О.В.— Электрон. текстовые данные.— М.: Палеотип, 2009.— 400 с.— Режим доступа: http://www.iprbookshop.ru/10252.

7. Божко В.П. Информационные технологии в статистике [Электронный ресурс]: учебник/ Божко В.П.— Электрон. текстовые данные.— М.: Финансы и статистика, 2011.— 152 с.— Режим доступа: http://www.iprbookshop.ru/12430

8. Ефимова М.Р. Практикум по общей теории статистики [Электронный ресурс]: учебное пособие/ Ефимова М.Р., Ганченко О.И., Петрова Е.В.— Электрон. текстовые данные.— М.: Финансы и статистика, 2011.— 369 с.— Режим доступа: http://www.iprbookshop.ru/12441.— ЭБС «IPRbooks»

9. Улитина Е.В. Статистика [Электронный ресурс]: учебное пособие/ Улитина Е.В., Леднева О.В., Жирнова О.Л.— Электрон. текстовые данные.— М.: Московский финансово-промышленный университет «Синергия», 2013.— 320 с.— Режим доступа: http://www.iprbookshop.ru/17045

10. Ловцов Д.А. Статистика [Электронный ресурс]: учебное пособие/ Ловцов Д.А., Богданова М.В., Михайлов М.А.— Электрон. текстовые данные.— М.: Российская академия правосудия, 2010.— 120 с.— Режим доступа: http://www.iprbookshop.ru/1872.

11. Илышев А.М. Общая теория статистики [Электронный ресурс]: учебник/ Илышев А.М.— Электрон. текстовые данные.— М.: ЮНИТИ-ДАНА, 2012.— 536 с.— Режим доступа: http://www.iprbookshop.ru/10504.

12. Тюрин Ю.Н. Многомерная статистика. Гауссовские линейные модели [Электронный ресурс]: учебное пособие/ Тюрин Ю.Н.— Электрон. текстовые данные.— М.: Московский государственный университет имени М.В. Ломоносова, 2011.— 136 с.— Режим доступа: http://www.iprbookshop.ru/13143.

13. Бурханова И.В. Теория статистики [Электронный ресурс]: учебное пособие/ Бурханова И.В.— Электрон. текстовые данные.— Саратов: Научная книга, 2012.— 159 с.— Режим доступа: http://www.iprbookshop.ru/8229.