КЫРГЫЗСКО-РОССИЙСКИЙ СЛАВЯНСКИЙ УНИВЕРСИТЕТ

Кафедра экономической теории

В.И. Марук

СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ

Конспект лекции и решение типовых задач

Бишкек 2007

В.И. Марук

М 25
СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ: Конспект лекции и решение типовых задач. – Бишкек: КРСУ, 2007. – 16 с.

Задача предлагаемого учебно-методического пособия – ознакомить студентов с сущностью и методами расчета средних величин и показателей вариации, оказать им практическую помощь при выполнении контрольной работы по курсу “Статистика”.

Лекция предназначена для студентов-заочников экономических специальностей.

Печатается по решению
кафедры экономической теории КРСУ

1. СУЩНОСТЬ СРЕДНИХ ВЕЛИЧИН И ПОКАЗАТЕЛЕЙ ВАРИАЦИИ

Статистическая совокупность состоит из отдельных единиц, обладающих индивидуальными особенностями и поэтому они отличаются друг от друга величиной количественного признака. Чтобы одним числом охарактеризовать уровень признака у всех единиц совокупности, когда размер признака у отдельных единиц колеблется (варьирует), исчисляется средняя величина.

Средней величиной в статистике называется обобщающий показатель, характеризующий размер признака, отнесенный к единице совокупности. В большинстве случаев средняя величина исчисляется делением общего объема варьирующего признака совокупности единиц на общую численность этих единиц. В соответствии с законом больших чисел в средней величине взаимопогашаются случайные индивидуальные различия между единицами и отражается то общее, что имеется в каждой единице однородной совокупности. Средние являются именованными числами, они выражаются в тех же единицах измерения, в каких измеряется признак.

Средние показатели подразделяются на типические и системные. Средняя величина, исчисленная для совокупности единиц, качественно однородной по изучаемому признаку, называется типической. Например, типической величиной является средний возраст пенсионеров страны. Средняя величина, исчисленная для единой, но неоднородной совокупности, называется системной. Например, системной величиной является средний возраст всех жителей страны, при расчете которого учитывается возраст младенцев, молодежи, пенсионеров и других возрастных групп. И типические, и системные средние величины широко применяются в статистике, так как специфика предмета статистики постоянно имеет дело с этими обобщающими показателями.

Основные виды средних величин, применяемых в статистике, и методы их расчета рассмотрены во втором и третьем параграфах лекции.

Необходимо отметить, что средняя величина не позволяет достаточно полно охарактеризовать совокупность единиц, так как по ней нельзя судить о том, как отклоняются от средней индивидуальные значения признака, т.е. какова вариация. Однако размер вариации признака позволяет судить о типичности и надежности средней величины. Чем меньше вариация, тем однороднее совокупность и, следовательно, тем типичнее средняя величина. При статистическом анализе социально-

экономических явлений бывает, что средняя величина признака в двух совокупностях единиц одинакова, а вариация признака совершенно разная, и поэтому практическое значение этих средних тоже разное.

Для количественной характеристики колеблемости признака в совокупности единиц применяются показатели вариации. Виды показателей вариации и методы их расчета рассмотрены в четвертом разделе лекции.

2. ВИДЫ СРЕДНИХ И СПОСОБЫ ИХ ВЫЧИСЛЕНИЯ

Наиболее часто встречающиеся в статистике виды средних (арифметическая, гармоническая, квадратическая и геометрическая) относятся к классу степенных средних. Выбор вида средней зависит от сущности решаемой задачи. В зависимости от состояния исходных данных, по которым рассчитывается средний показатель, каждая из этих четырех средних может быть простой и взвешенной. Средние простые рассчитываются по несгруппированным единицам совокупности, а средние взвешенные – по сгруппированным, т.е. для рядов распределения.

Из всех степенных средних в статистике наиболее часто применяется *средняя арифметическая*. Она применяется, когда общий объем варьирующего признака для всей совокупности образуется как сумма значений признака отдельных единиц этой совокупности.

Формула средней арифметической имеет вид:

простая

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\Sigma x}{n},$$

взвешенная

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n} = \frac{\Sigma xf}{\Sigma f},$$

где $\bar{x}$ – средняя величина;

x – индивидуальные значения признака (варианты);

n – численность единиц совокупности;

f – частоты (веса), показывающие, сколько раз встречаются значения признака в совокупности единиц ($\Sigma f = n$).

Рассмотрим применение этих формул.

Задача 2.1

Бригада рабочих состоит из пяти человек. Их месячная заработная плата составляет: 2400, 2500, 2800, 3000 и 3200 сом. Определите среднюю заработную плату рабочего бригады.

Решение

Так как имеются несгруппированные исходные данные, то для расчета средней месячной заработной платы воспользуемся формулой средней арифметической простой:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{2400+2500+2800+3000+3200}{5} = 2780 \text{ сом.}$$

Задача 2.2

Определите среднюю заработную плату рабочих бригады, если известен месячный заработок каждого из них:

Месячная заработная плата рабочего, сом. (x)	Количество рабочих, чел. (f)	Произведение вариантов и частот (xf)
2400	3	7200
2600	5	13000
2800	10	28000
3000	2	6000
Итого	20	54200

Решение

По условию задачи имеется вариационный дискретный ряд, т.е. исходные данные сгруппированы. Поэтому для расчета средней заработной платы воспользуемся формулой средней арифметической взвешенной (сумма из произведений вариантов и частот вычислена в последней графе таблицы):

$$\bar{x} = \frac{\Sigma xf}{\Sigma f} = \frac{54200}{20} = 2710 \text{ сом.}$$

Задача 2.3

Определите среднюю заработную плату рабочих бригады по следующим данным:

Месячная заработная плата рабочего, сом.	Количество рабочих, чел. (f)	Середина интервала (x)	Произведение вариантов и частот (xf)
До 2600	3	2500	7500
2600–2800	7	2700	18900
2800–3000	5	2900	14500
3000–3200	3	3100	9300
Более 3200	2	3300	6600
Итого	20	–	56800

Решение

По условию задачи имеется интервальный вариационный ряд, т.е. исходные данные сгруппированы. Поэтому, как и при решении задачи 2.2, воспользуемся формулой средней арифметической взвешенной. С этой целью преобразуем интервальный вариационный ряд в дискретный, определив для каждого интервала его середину. Для закрытого интервала (вторая, третья и четвертая группы) среднее значение равно полусумме его верхней и нижней границ. Если интервальный ряд имеет открытые границы, то условно величина интервала первой группы приравнивается к величине интервала второй группы, а величина интервала последней группы принимается равной величине интервала предпоследней группы.

Все вспомогательные расчеты выполнены в таблице. Средняя заработная плата рабочих бригады составит:

$$\bar{x} = \frac{\Sigma xf}{\Sigma f} = \frac{56800}{20} = 2840 \text{ сом.}$$

Следует отметить, что исчисленная средняя является неточной, так как при расчете середины интервала предполагалось, что внутри интервалов варианты распределены равномерно, а это не всегда соответствует действительности. Кроме того, то обстоятельство, что границы первого и последнего интервалов вариационного ряда открыты и поэтому их середина найдена условно, делает среднюю еще менее точной.

Средняя гармоническая – это величина, обратная средней арифметической из обратных значений признака. Она применяется, когда общий объем варьирующего признака образуется как сумма обратных значений признака отдельных вариантов.

Формула средней гармонической имеет вид:

простая

$$\bar{x}_h = \frac{n}{\sum \frac{1}{x}},$$

взвешенная

$$\bar{x}_h = \frac{M_1 + M_2 + \ldots + M_n}{\frac{M_1}{x_1} + \frac{M_2}{x_2} + \ldots + \frac{M_n}{x_n}} = \frac{\Sigma M}{\sum \frac{M}{x}},$$

где M – объем изучаемого признака по группам ($M = xf$).

Из сравнения формул средней арифметической взвешенной и средней гармонической взвешенной следует, что последняя применяется, когда неизвестны частоты и приходится взвешивать варианты по объемам признака в каждой группе.

Рассмотрим применение формулы средней гармонической взвешенной на примере[1].

Задача 2.4

Определите среднюю заработную плату рабочих бригады по следующим данным:

Месячная заработная плата рабочего, сом. (x)	Месячный заработок группы рабочих, сом. (M)	Расчетная численность рабочих, чел. $\left(\frac{M}{x}\right)$
2400	7200	3
2600	13000	5
2800	28000	10
3000	6000	2
Итого	54200	20

Решение

По условию задачи имеются варианты и суммарные объемы признака (месячный заработок по группам рабочих), а информации о частотах (численности рабочих по группам) нет. Поэтому определим среднюю месячную заработную плату рабочих бригады по формуле средней гармонической взвешенной (сумма из отношений M/x вычислена в последней графе таблицы):

$$\bar{x}_h = \frac{\Sigma M}{\sum \frac{M}{x}} = \frac{54200}{20} = 2710 \text{ сом.}$$

Средняя квадратическая исчисляется, когда общий объем варьирующего признака образуется как сумма квадратов отдельных вариантов. Формула средней квадратической имеет вид:

простая

$$\bar{x}_q = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}} = \sqrt{\frac{\Sigma x^2}{n}} ,$$

[1] Средняя гармоническая простая в статистике практически не применяется.

взвешенная

$$\bar{x}_q = \sqrt{\frac{x_1^2 f_1 + x_2^2 f_2 + \ldots + x_n^2 f_n}{f_1 + f_2 + \ldots + f_n}} = \sqrt{\frac{\Sigma x^2 f}{\Sigma f}} .$$

Средняя квадратическая применяется в статистике при исчислении показателей вариации.

Средняя геометрическая исчисляется, когда общий объем варьирующего признака образуется как произведение отдельных вариантов. Формула средней геометрической имеет вид:

простая

$$\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \ldots x_n} = \sqrt[n]{\prod x} ,$$

взвешенная

$$\bar{x}_g = \sqrt[\Sigma f]{x_1^{f_1} \cdot x_2^{f_2} \ldots x_n^{f_n}} = \sqrt[\Sigma f]{\prod x^f} .$$

Средняя геометрическая применяется в статистике при исчислении средних темпов роста.

3. МОДА И МЕДИАНА

Кроме степенных средних, в статистике применяются структурные средние, к которым относятся мода и медиана. В отличие от степенной средней, являющейся абстрактной величиной, мода и медиана в дискретном вариационном ряду – это конкретные значения отдельных вариантов.

Модой называется величина признака, которая наиболее часто встречается в данной совокупности единиц. Мода применяется, например, для характеристики наиболее распространенного размера заработной платы на предприятии; модальной является цена, по которой на рынке было продано наибольшее количество данного товара.

Медианой называется величина признака у той единицы совокупности, которая находится в середине ранжированного ряда. Медиана делит вариационный ряд на две части, в одной из них все единицы имеют значение признака меньшее, чем у медианы, а в другой – большее. Медиану целесообразно применять вместо средней арифметической при наличии в вариационном ряду открытых интервалов. Из-за условно установленных границ открытых интервалов ряда средняя арифметическая является неточной, а на точность расчета медианы это не влияет. Медиана применяется при статистических методах контроля качества продукции. Совместное использование средней арифметической, моды и медианы при анализе вариационных рядов позволяет охарактеризовать тип распределения совокупности единиц.

В дискретном вариационном ряду модой является вариант, которому соответствует наибольшая частота. Для нахождения медианы в дискретном вариационном ряду необходимо определить накопленные частоты. Порядковый номер медианы ряда с нечетным числом вариантов (n) равен $\frac{n+1}{2}$. Если число вариантов четное, то медиана равна средней из двух вариантов, находящихся в середине ряда, т.е. имеющих порядковые номера $\frac{n}{2}$ и $\frac{n+2}{2}$.

Мода и медиана интервального вариационного ряда могут быть исчислены лишь приближенно. Сначала необходимо определить модальный и медианный интервалы. Модальным называется интервал, которому соответствует наибольшая частота. Медианным называется первый интервал, накопленная частота которого превышает половину общей суммы частот.

Модальная величина признака для интервального вариационного ряда определяется по формуле:

$$Mo = x_0 + h\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)},$$

где Mo – мода;
x_0 – нижняя граница модального интервала;
h – величина модального интервала;
f_2 – частота модального интервала;
f_1 – частота интервала, предшествующего модальному;
f_3 – частота интервала, следующего за модальным.

Медиана для интервального вариационного ряда рассчитывается по формуле:

$$Me = x_0 + h\frac{0{,}5\Sigma f - S_{m-1}}{f_m},$$

где Me – медиана;
x_0 – нижняя граница медианного интервала;
h – величина медианного интервала;
Σf – сумма частот ряда;
S_{m-1} – сумма частот интервалов, предшествующих медианному;
f_m – частота медианного интервала.

Решение типовых задач

Задача 3.1

Имеются следующие данные о распределении студентов по результатам сдачи экзамена (табл.). Определите моду и медиану.

Оценка	2	3	4	5
Количество студентов, чел.	28	32	23	22
Накопленные частоты	28	60	83	105

Решение

В этой таблице модальной оценкой является тройка, так как большинство студентов (32 человека) получило именно эту оценку. В середине ранжированного ряда находится студент с порядковым номером 53 $\left(\frac{105+1}{2}\right)$. Студент с этим порядковым номером относится ко второй группе, так как в эту группу попали студенты с порядковыми номерами от 29 до 60. Следовательно, медиана равна трем.

Задача 3.2

Имеются следующие данные о заработной плате работников фирмы (табл.). Определите моду и медиану заработной платы работников.

Заработная плата работника, сом.	2800–2900	2900–3000	3000–3100	3100–3200	3200–3300
Количество работников, чел.	12	22	28	44	10
Накопленные частоты	12	34	62	106	116

Решение

Модальным является интервал 3100–3200, так как ему соответствует наибольшая частота (44 человека). Подставим данные задачи в формулу моды:

$$Mo = 3100 + 100\frac{44-28}{(44-28)+(44-10)} = 3132 \text{ сом.}$$

Следовательно, у работников фирмы наиболее часто встречается заработная плата в размере 3132 сом.

Чтобы определить медианный интервал, в последней строке таблицы подсчитаны накопленные частоты. Первым интервалом, накопленная частота которого превышает половину общей суммы частот (58), является интервал третьей группы. Поэтому медианным будет интервал 3000–3100. Подставив соответствующие данные в формулу медианы, получим:

$$Me = 3000 + 100\frac{0{,}5\times 116 - 34}{28} \approx 3086 \text{ сом.}$$

Следовательно, половина работников фирмы получает меньшую заработную плату (3086 сом.), а другая половина – большую.

4. ПОКАЗАТЕЛИ ВАРИАЦИИ ПРИЗНАКА

Чтобы охарактеризовать вариацию количественного признака в совокупности единиц применяются следующие показатели: размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение, коэффициент вариации и другие. Наиболее простым из перечисленных показателей является размах вариации (R), который представляет собой разность между максимальным и минимальным значениями признака в изучаемой совокупности.

$R = x_{\max} - x_{\min}$.

Чем больше значение R, тем больше вариация. Недостаток этого показателя заключается в том, что его величина зависит только от двух крайних значений признака и не учитывает как часто встречается тот или иной вариант в данной совокупности единиц, т.е. не учитывает частоты отдельных вариантов. Следующие показатели позволяют более объективно охарактеризовать вариацию признака.

Среднее линейное отклонение представляет собой среднюю арифметическую из абсолютных значений отклонений отдельных вариантов от их средней арифметической. Среднее линейное отклонение ($\bar{d}$) вычисляется по формулам:

для несгруппированных данных

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n};$$

для сгруппированных данных (для вариационного ряда)

$$\bar{d} = \frac{\sum |x - \bar{x}| f}{\sum f}.$$

Чем больше значение ($\bar{d}$), тем больше вариация.

Наибольшее применение среди показателей вариации получили дисперсия и среднее квадратическое отклонение. Значительное распространение этих показателей в качестве меры вариации признака объясняется их математическими свойствами. Дисперсией называется средний квадрат отклонений вариантов от их средней арифметической. Дисперсия признака (σ^2) вычисляется по формулам:

для несгруппированных данных

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n};$$

для сгруппированных данных

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}.$$

Ввиду важности этого показателя рассмотрим порядок его расчета по сгруппированным данным:

1) определяется отклонение каждого варианта от средней арифметической: $(x - \bar{x})$;

2) каждое отклонение возводится в квадрат: $(x - \bar{x})^2$;

3) каждый квадрат отклонения умножается на соответствующую частоту: $(x - \bar{x})^2 f$;

4) все найденные произведения суммируются: $\sum (x - \bar{x})^2 f$;

5) сумма произведений делится на сумму частот: $\dfrac{\sum (x - \bar{x})^2 f}{\sum f}$.

Среднее квадратическое отклонение (σ) представляет собой корень квадратный из дисперсии:

для несгруппированных данных

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}};$$

для сгруппированных данных

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}.$$

Среднее квадратическое отклонение, как и среднее линейное отклонение, показывает, на сколько в среднем отклоняются индивидуальные варианты от их средней величины. Среднее квадратическое отклонение и среднее линейное отклонение выражаются в тех же единицах измерения, что и варианты.

Для сравнения вариации одного и того же признака в разных совокупностях или для сравнения вариации различных признаков в одной совокупности используется относительный показатель колеблемости – коэффициент вариации.

Коэффициент вариации (V) представляет собой процентное отношение среднего квадратического отклонения к средней арифметической величине изучаемого признака:

$$V = \frac{\sigma}{\bar{x}} 100 .$$

При измерении вариации альтернативного признака дисперсия определяется по формуле:

$$\sigma^2 = pq ,$$

где p – доля единиц совокупности, обладающих изучаемым признаком;

q – доля единиц совокупности, не обладающих изучаемым признаком.

Поскольку $p + q = 1$, то формулу дисперсии альтернативного признака можно представить в виде:

$$\sigma^2 = p(1 - p) .$$

Решение типовых задач

Задача 4.1

По следующим данным определите показатели вариации возраста рабочих бригады:

Возраст рабочих, лет (x)	17	18	21	24	26
Количество рабочих, чел. (f)	2	2	6	3	1

Решение

Прежде всего необходимо определить средний возраст рабочих бригады. Для этой цели воспользуемся формулой средней арифметической взвешенной, так как по условию задачи известны варианты и частоты. Вспомогательные расчеты, необходимые для вычисления средней и показателей вариации, выполнены в таблице:

x	f	xf	$\lvert x-\bar{x} \rvert$	$\lvert x-\bar{x} \rvert f$	$\lvert x-\bar{x} \rvert^2 f$
17	2	34	4	8	32
18	2	36	3	6	18
21	6	126	0	0	0
24	3	72	3	9	27
26	1	26	5	5	25
Итого	14	294	–	28	102

1. Средний возраст рабочих бригады:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{294}{14} = 21 \text{ год.}$$

2. Среднее линейное отклонение:

$$\bar{d} = \frac{\sum \lvert x-\bar{x} \rvert f}{\sum f} = \frac{28}{14} = 2 \text{ года.}$$

3. Дисперсия

$$\sigma^2 = \frac{\sum (x-\bar{x})^2 f}{\sum f} = \frac{102}{14} = 7{,}3$$

4. Среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2 f}{\sum f}} = \sqrt{7{,}3} = 2{,}7 \text{ года.}$$

5. Коэффициент вариации:

$$V = \frac{\sigma}{\bar{x}} 100 = \frac{2{,}7}{21} 100 = 12{,}9\% .$$

Задача 4.2

При обследовании 850 единиц изготовленной продукции оказалось, что 102 единицы имели различные дефекты и были забракованы. Определите дисперсию доли бракованной продукции.

Решение

Определим долю бракованной продукции среди обследованных единиц:

$$p = \frac{102}{850} = 0{,}12 .$$

Дисперсия доли бракованной продукции:

$$\sigma^2 = p(1 - p) = 0{,}12(1 - 0{,}12) = 0{,}106 .$$

ЛИТЕРАТУРА

1. *Гусаров В.М.* Статистика. – М.: ЮНИТИ, 2002.
2. Практикум по статистике / Под ред. В.М. Симчеры. – М.: Финстатинформ, 1999.
3. Практикум по теории статистики / Под ред. Р.А. Шмойловой. – М.: Финансы и статистика, 2004.
4. Теория статистики / Под ред. Р.А. Шмойловой. – М.: Финансы и статистика, 2006.

СОДЕРЖАНИЕ

Виктор Иванович Марук

СРЕДНИЕ ВЕЛИЧИНЫ
И ПОКАЗАТЕЛИ ВАРИАЦИИ

Конспект лекции и решение типовых задач

Редактор *Т. П. Вязьмина*
Технический редактор *О.А. Матвеева*
Корректор *Е.И. Полихова*
Компьютерная верстка *Н.А. Лапиной*

Подписано к печати 10.05.07. Формат 60×84 $^{1}/_{16}$
Офсетная печать. Объем 1 п.л.
Тираж 130 экз. Заказ 63.

Отпечатано в типографии КРСУ
720000, Бишкек, ул. Шопокова, 68