

МАТЕМАТИЧЕСКАЯ ВЕРТИКАЛЬ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И СТАТИСТИКА. 7 КЛАСС (2 ч./нед.)

Урок 18. Среднее арифметическое группированных данных

Материалы представляют собой примерный сценарий урока по теме «Среднее арифметическое». Учитель может на свое усмотрение использовать сценарий целиком или частично, используя фрагменты наряду с собственными разработками и материалами учебника¹. Авторы будут благодарны за замечания и предложения по структуре и содержанию сценариев.

Цель урока – сформировать умение находить среднее арифметическое данных, сгруппированных по величине. Другая цель – создать представление о частоте события и связи частоты и среднего арифметического. Соответствующие навыки и понятия используются впоследствии при изучении теории вероятностей. При возможности познакомьте учащихся со средствами Excel, предназначенными для поиска средних в электронных таблицах. Это облегчит выполнение последующих практических и лабораторных работ.

Оборудование. Калькуляторы или персональные компьютеры

Вопросы для повторения

1. В таблице даны площади земельных угодий в РФ по категориям. Какого типа диаграмму лучше использовать для представления этих данных?

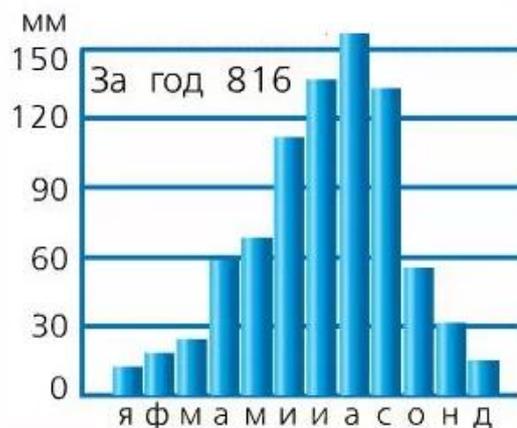
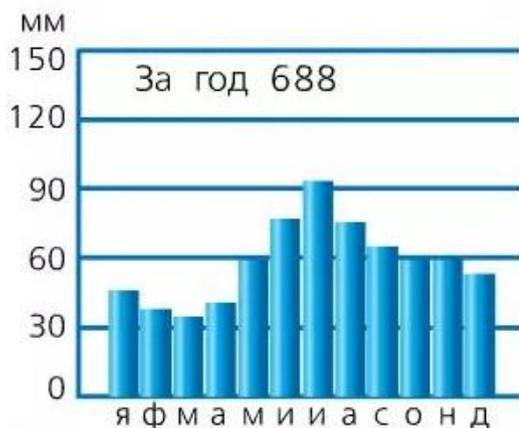
Табл. 1. Распределение земель РФ по категориям

Категория	Леса	Особо охраняемые территории	Пром. предприятия и земли спец. назначения	Населенные пункты	Земли сельхоз. назначения	Водный фонд	Земли запаса
Площадь (млн га)	1122,3	46,8	16,9	20,0	386,5	28,0	89,3

¹ Математика 7-9 класс. Теория вероятностей и статистика / Ю. Н. Тюрин, А. А. Макаров, И. Р. Высоцкий, И. В. Яценко. – 3-е изд., стереотипное. – М.: МЦНМО: ОАО «Московские учебники», 2011. – 256 с.: ил.

Ответ: круговая, поскольку дается в таблице дается распределение долей целого (площади земель Российской Федерации).

2. На диаграммах показано распределение осадков по месяцам в двух городах России: Москве и во Владивостоке. Попробуйте, опираясь на свои знания географии и личный опыт, определить, какая диаграмма какому городу соответствует.



Ответ: в Москве умеренно-континентальный климат. Моря далеко. Осадки летом (дождь) и зимой (снег) не очень сильно различаются по количеству. Во Владивостоке муссонный климат – дождей летом должно быть намного больше, чем осадков зимой, а всего осадков больше, чем в Москве. Скорее всего, левая диаграмма относится к Москве, а правая – к Владивостоку.

Среднее арифметическое группированных данных

Часто, когда данных много, их приходится группировать: указывать в таблице значения и их количества: сколько каких чисел встречается. Рассмотрим самый простой случай.

Пример 1. За контрольную работу по статистике школьникам выставлены отметки (см. табл.2). Нужно найти среднюю отметку.

Табл.2. Отметки за контрольную работу

Фамилия		Фамилия		Фамилия		Фамилия	
Абдуллаева	3	Грачёва	3	Луданова	3	Обжужан	4
Андреенко	4	Заикина	4	Любчикова М.	2	Ожерельев	4
Антипов	4	Зайцева	4	Любчикова Л.	3	Прошунин	3
Афанасьева	4	Инамов	5	Макаркина	4	Сахаров	4
Балабанова	5	Карпова	5	Малофеев	4	Соколов	4
Болелова	3	Клопов	3	Мельникова	4	Степанов	4
Боровкова	4	Копылова	4	Мохаммад	5	Хромоченков	3
Бочарова	3	Крестиничева	3	Новикова А.	3	Фетисов	4
Воробьева	3	Кудряшов	2	Новикова Ю.	3	Яковкина	4

Здесь 36 человек и, стало быть, 36 отметок. Хорошо то, что все отметки – целые числа от 2 до 5. Это позволяет представить данные гораздо компактнее,

ведь нам не важно, кто именно какую отметку получил. Составим таблицу, где просто укажем, сколько каких отметок.

Табл.3. Распределение отметок

Отметка	2	3	4	5
Количество				
	2	13	17	4

Для заполнения такой таблицы удобно использовать технику ручного подсчета, например, с помощью «заборчиков».

Теперь запишем среднее, сгруппировав в скобках двойки, тройки, четверки и пятерки. В знаменателе запишем сумму чисел из графы «Количество». Это даст общее количество значений.

$$\frac{(2+2) + \left(\frac{13 \text{ троек}}{3+3+\dots+3}\right) + \left(\frac{17 \text{ четверок}}{4+4+\dots+4}\right) + (5+5+5+5)}{2+13+17+4}$$

Можно записать проще:

$$\frac{2 \cdot 2 + 3 \cdot 13 + 4 \cdot 17 + 5 \cdot 4}{2 + 13 + 17 + 4} = \frac{4 + 39 + 68 + 20}{36} = \frac{131}{36} \approx 3,64.$$

Обратим внимание на то, как получается итоговая формула: в числителе сумма произведений значений на их количества, а в знаменателе – сумма количеств (то есть общее количество всех отметок). Такой способ поиска среднего арифметического удобен, когда в наборе много повторяющихся данных.

Пример 2. Международная картографическая компания 2ГИС проводила подсчет средней этажности домов в одном из кварталов Москвы. Данные занесены в таблицу 4. Найдите среднее число этажей в этом квартале. Подумайте, несет ли какую-нибудь полезную информацию такой подсчет.

Табл.4. Этажность зданий в квартале

Число этажей	2	3	5	9	12	14	17
Количество зданий	2	5	24	7	3	3	1

Решение. Среднее равно

$$\frac{2 \cdot 2 + 3 \cdot 5 + 5 \cdot 24 + 9 \cdot 7 + 12 \cdot 3 + 14 \cdot 3 + 17 \cdot 1}{2 + 5 + 24 + 7 + 3 + 3 + 1} = \frac{4 + 15 + 120 + 63 + 36 + 42 + 17}{45} = \frac{297}{45} = 6,6.$$

Желательный итог обсуждения. Пятиэтажные дома обычно имеют 3 – 10 подъездов по 15 или 20 квартир в каждом. Многоэтажные дома часто имеют один подъезд. Само по себе число «среднее число этажей» ничего не говорит

ни о плотности застройки (количество зданий на кв.км), ни о плотности населения (число жителей на кв.км) в квартале. Трудно представить себе, зачем нужна и что в действительности отражает такая информация.

Интервальная группировка. В рассмотренных примерах количество различных значений было небольшим. В примере 1 всего четыре значения – 2, 3, 4 и 5. В примере 2 – семь значений. Бывает так, что значений очень много. Тогда их приходится группировать в интервалы. Близкие значения, как правило, попадают в один и тот же интервал.

Пример 3. Некоторое агентство изучало рост взрослых женщин в России. Измерили рост женщин в случайной выборке 1000 человек. Для целей исследования данные сгруппировали в интервалы по 5 см и занесли в таблицу (см. табл. 4). Вычислите среднее арифметическое значение роста женщин, попавших в выборку.

Табл. 4.

Рост, см	Кол-во
140 – 145	5
145 – 150	14
150 – 155	52
155 – 160	106
160 – 165	178
165 – 170	236
170 – 175	205
175 – 180	123
180 – 185	60
185 – 190	19
190 – 195	2

Табл. 5.

Рост, см	Сред. инт-ла	Кол-во
140 – 145	142,5	5
145 – 150	147,5	14
150 – 155	152,5	52
155 – 160	157,5	106
160 – 165	162,5	178
165 – 170	167,5	236
170 – 175	172,5	205
175 – 180	177,5	123
180 – 185	182,5	60
185 – 190	187,5	19
190 – 195	192,5	2

Решение. Прежде чем заняться вычислениями, предложите ученикам попробовать сказать без подсчетов, каким примерно должно быть среднее арифметическое.

Если бы в таблице были конкретные значения, то средний рост вычислялся бы так же, как в примере про отметки. Но там интервалы, а находить «среднее арифметическое интервалов» мы не умеем. Поэтому на каждом интервале нужно выбрать какое-нибудь число². Проще всего (обычно так и делают) взять середину каждого интервала (см. табл.5).

Если есть компьютер, то не важно, с какими числами работать. Если же вычисления делаются вручную или с помощью калькулятора, удобнее работать с небольшими числами, желательно – с целыми. Можно воспользоваться свойствами среднего арифметического, изученными на предыдущем уроке: среднее значение «двигается» и «сжимается» вместе со всем набором данных.

Рассмотрим один из способов упрощения вычислений. Вычтем из каждого числа 142,5 – наименьшее из средних по интервалам, а результаты разде-

² При этом, конечно, расчет получается менее точным, чем если бы мы усредняли 1000 исходных значений. Но погрешность мала, а выигрыш в скорости расчета велик.

лим на 5 (это удобно, так как каждое из полученных чисел кратно пяти). Результаты этих действий отражены в таблице 6. Рассчитав среднее по полученным данным, легко будет найти среднее исходных.

Табл.6

Рост, см	142,5	147,5	152,5	157,5	162,5	167,5	172,5	177,5	182,5	187,5	192,5
Рост – 142,5	0	5	10	15	20	25	30	35	40	45	50
$\frac{\text{Рост}-142,5}{5}$	0	1	2	3	4	5	6	7	8	9	10
Количество	5	14	52	106	178	236	205	123	60	19	2

Среднее равно

$$\frac{0 \cdot 5 + 1 \cdot 14 + 2 \cdot 52 + 3 \cdot 106 + 4 \cdot 178 + 5 \cdot 236 + 6 \cdot 205 + 7 \cdot 123 + 8 \cdot 60 + 9 \cdot 19 + 10 \cdot 2}{1000} = 5,09.$$

Теперь с полученным числом нужно сделать обратные действия: сначала умножить на 5, а затем прибавить 142,5:

$$5,09 \cdot 5 + 142,5 = 167,95.$$

Мы получили приближенное среднее арифметическое. Почему приближенное? Потому что мы знали не точный рост каждой из участниц выборки, а только приближенное значение (а именно – интервал, в который попал ее рост). Но потеря точности невелика.

Частоты значений. Вернемся к примеру 2 про дома и этажи. Значение 9 (девятиэтажный дом) встречается 7 раз из 45. Скажем, что число $\frac{7}{45} \approx 0,156$ – это *частота* значения 9.

Пример 4. Найдите частоты значений 5 и 12 из примера 2.

Ответ: $\frac{24}{45}$ и $\frac{3}{45}$, то есть приблизительно 0,533 и 0,067.

Среднее арифметическое группированных данных можно вычислять с помощью частот. Сначала занесем частоты в таблицу, округлив их до сотых.

Табл.7. Этажность зданий в квартале с частотами

Число этажей	2	3	5	9	12	14	17
Количество зданий	2	5	24	7	3	3	1
Частота	0,044	0,111	0,533	0,156	0,067	0,067	0,022

Составим сумму произведений значений и частот:

$$2 \cdot 0,44 + 3 \cdot 0,111 + 5 \cdot 0,533 + 9 \cdot 0,156 + 12 \cdot 0,067 + 14 \cdot 0,067 + 17 \cdot 0,022 = 6,6.$$

Обсудите с учениками, почему такая формула вычисления среднего верна. Чему равна сумма частот всех этажей?

Желательный результат обсуждения. Прежняя формула и формула с частотами – одно и то же. Если в последней формуле вынести за скобки общий множитель $1/45$, получится обычная формула для среднего арифметического. Сумма частот равна единице.

Таким образом, можно дать два определения среднего арифметического числового набора.

1. Сумма значений, деленная на их количество.
2. Сумма произведений значений на их частоты.

Примечание. Если есть возможность выполнять задания на компьютере в программе Excel, разумно использовать следующие функции³.

= СУММ(массив) – вычисляет сумму чисел массива;

= СУММПРОИЗВ(массив1, массив2) – вычисляет сумму произведений соответствующих чисел массива. Пример выполнения задания из примера 3 показан на рисунке.

Обратите внимание на вычисление среднего с использованием функций СУММПРОИЗВ и СУММ:

$$= \text{СУММПРОИЗВ}(\text{значения}; \text{количества}) / \text{СУММ}(\text{количества}).$$

C7		=СУММПРОИЗВ(C4:M4;C5:M5)/СУММ(C5:M5)											
	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2	1	Рост, см	142,5	147,5	152,5	157,5	162,5	167,5	172,5	177,5	182,5	187,5	192,5
3	2	Рост -142,5	0	5	10	15	20	25	30	35	40	45	50
4	3	(Рост -142,5)/5	0	1	2	3	4	5	6	7	8	9	10
5		Количество	5	14	52	106	178	236	205	123	60	19	2
6													
7		Среднее в строке 3	5,09										
8		Средний рост	168										
9													

Вывод. Среднее арифметическое можно успешно находить, не только когда у нас есть "сырые" данные эксперимента. Иногда значения сгруппированы в интервалы, и если группировка удачная, то найти среднее арифметическое можно проще, чем имея дело с исходным массивом данных.

Важно! При группировке большую роль играет выбор интервалов.

³ Более подробно действие функций и примеры их применения будут изучены в ходе выполнения лабораторной работы по темам описательной статистики.

Если интервалы группировки очень узкие, то их много. Поэтому ошибка вычислений маленькая, но вычислений много. Если интервалы очень широкие, то вычислений мало, зато сильно падает точность усреднений.

При представлении и описании данных стараются подобрать интервалы таким образом, чтобы их было не очень много (удобно видеть), но при этом ошибки в вычислениях не были бы слишком большими (результаты можно использовать).

Рекомендуемое домашнее задание: в Приложении.

Дополнительные материалы для подготовки к уроку. Учитель на свое усмотрение может свободно использовать таблицы⁴, размещенные на сайте «Вероятность в школе» <http://ptlab.mccme.ru/node/350>.

При проведении урока в классе, оборудованном компьютерами на столах учащихся, учитель может включить в урок вычисления в электронных таблицах непосредственно на рабочих столах учащихся.

⁴ Все данные взяты из открытых источников и представлены в формате .xls или .xlsx. Таблицы можно использовать как целиком, так и делая из них необходимые выдержки.

Приложение

Домашнее задание

1. В течение года в школе прошли олимпиады по 12 учебным предметам. В таблице показано, сколько учащихся 7-х классов приняло участие в одной, в двух олимпиадах и т.д. Например, 15 школьников не участвовали ни в одной из олимпиад. С помощью калькулятора или компьютера найдите среднее число олимпиад, в которых участвовал семиклассник этой школы.

Количество предметов	Количество учащихся
0	15
1	53
2	103
3	76
4	81
5	53
6	34
7	12
8	4
9	2
10	0
11	0
12	1

2. В таблице дано распределение баллов ЕГЭ по математике в трех школах. Данные сгруппированы в интервалы по 10 баллов. В трех правых графах показано, сколько выпускников школы получили балл, попадающий в соответствующий интервал.

а) Найдите приближенно средний балл в школе 3.

б) Можно ли утверждать, что истинный средний балл выпускников школы 1 выше, чем истинный средний балл выпускников школы 2?

в) Можно ли утверждать, что истинный средний балл в школе 1 выше, чем в школе 3?

Обоснуйте свои ответы на вопросы б) и в).

Баллы ЕГЭ	Количество выпускников		
	Школа 1	Школа 2	Школа 3
1 – 10	2	3	7
11 – 20	5	7	9
21 – 30	11	10	14
31 – 40	10	15	17
41 – 50	23	19	21
51 – 60	19	22	11
61 – 70	27	18	9
71 – 80	8	5	4
81 – 90	0	3	2
91 – 100	2	1	0
Средний балл по школе	49,9	47,1	