МАТЕМАТИЧЕСКАЯ ВЕРТИКАЛЬ

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

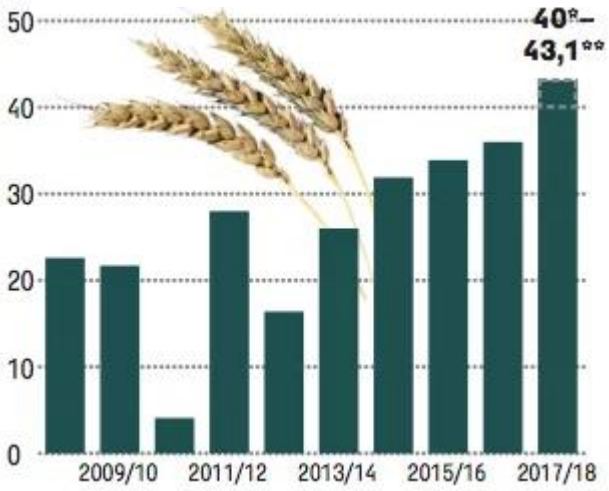
Урок 24. Определение и свойства медианы

Примерный сценарий урока по теме «Определение и свойства медианы». Учитель может на свое усмотрение использовать сценарий целиком или частично, используя фрагменты наряду с собственными разработками и материалами учебника[1]. Авторы будут благодарны за замечания и предложения по структуре и содержанию сценариев.

Цель урока – познакомить учащихся с теоретическим определением медианы, научить проверять определение для разных чисел, продемонстрировать устойчивость медианы на примерах и научиться находить медианных представителей.

Повторение (5 мин). На столбиковой диаграмме показаны данные об экспорте зерна из России с 2008 года (для 2018 года сделаны прогнозы 40 – 43,1 млн т). Рассмотрите диаграмму и выполните задания (устно).

Диагр.1. Экспорт зерна из РФ



1. Как вы думаете, почему объем экспорта считается не по календарным годам, а по концу года – началу следующего (как учебные годы). Что, например, означает экспорт в 2009/2010 г?
2. Чем может объясняться резкое падение экспорта зерна в 2010/2011 году?
3. Какая тенденция наблюдается в последние 6 лет? Придумайте правдоподобное объяснение.
4. Можно ли оценить эту тенденцию количественно – каков примерный средний ежегодный прирост экспорта за последние 6 лет?

Ответы, к которым нужно подвести учащихся в ходе обсуждения. Экспорт (продажа за рубеж) пшеницы производится из запасов собранного урожая. Урожай собирают осенью. Поэтому экспорт этого урожая продолжается в конце текущего года и в начале следующего – до следующего урожая. В 2010

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

году на хлебородных территориях европейской части России почти все лето стояла жестокая засуха. Большая часть урожая сгорела. Запасов собранного зерна оказалось меньше, чем обычно, поэтому экспорт резко сократился. Последние шесть лет экспорт устойчиво растет. Вероятно, этому способствует снижение курса рубля в 2014 году – российское зерно на мировых рынках оказалось дешевле зерна из других стран. Спрос на него возрос, прибыль увеличилась, и дополнительный доход был направлен на оснащение сельского хозяйства, что привело к дальнейшему росту производства и увеличению экспорта. За шесть лет экспорт вырос примерно на $40-16=24$ млн т, то есть в среднем на 4 млн т ежегодно.

Новый материал Определение и свойства медианы

На предыдущем уроке мы познакомились со способом вычисления медианы. Напомним: если чисел в наборе нечетное количество, то медиана – серединное число в упорядоченном ряду чисел. Если чисел четное количество, то медиана – любое из двух серединных чисел или любое число между ними. Эти два случая можно объединить в один следующим **определением:**

Число m называется медианой числового набора, если в этом наборе хотя бы половина чисел *не больше* числа m и хотя бы половина чисел *не меньше* числа m.

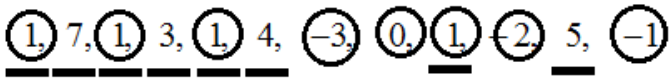
Пример 1. Дан числовой набор

1, 7, 1, 3, 1, 4, -3, 0, 1, -2, 5, -1.

Проверьте, пользуясь определением, является ли медианой этого набора

а) число 1; б) число 0?

Решение. а) Всего в наборе двенадцать чисел. Чтобы число 1 было медианой, нужно, чтобы в наборе было хотя бы шесть чисел, которые не больше числа 1 (то есть меньше или равны числу 1) и хотя бы шесть чисел, которые не меньше числа 1 (то есть больше или равны числу 1). Обведем кружком числа, которые не больше числа 1, и подчеркнем числа, которые не меньше 1:



Заметим, что число 1 попадает в обе группы. Поэтому его нужно и обвести, и подчеркнуть всякий раз, когда оно встречается. Всего обведено 8 чисел и подчеркнуто тоже 8 чисел.

Таким образом, в данном наборе восемь чисел, которые не больше числа 1, и восемь чисел, которые не меньше числа 1. Определение медианы выполнено. Значит, 1 – медиана данного набора.

б) Сделаем то же самое, приняв за границу число 0. Обведем числа, которые не больше нуля, и подчеркнем те, которые не меньше нуля.

1, 7, 1, 3, 1, 4, –3, 0, 1, –2, 5, –1

В наборе четыре числа, которые не больше числа 0, и девять чисел, которые не меньше числа 0. Чисел, которые не больше числа 0, слишком мало: $4 < 6$. Поэтому 0 не является медианой данного набора по определению.

Пример 2. Дан числовой набор

3, 2, 2, 7, 4, 1, 5, 4.

Проверьте, пользуясь определением, является ли медианой этого набора число:

а) 2; б) 3; в) 4; г) 3,22.

Решение. а) В наборе 8 чисел. Не превосходят числа 2 три числа (2, 2, 1). А нужно, чтобы таких чисел было не меньше, чем четыре. Значит, 2 не является медианой.

б) Не превосходят число 3 четыре числа (3, 2, 2, 1). Не меньше, чем число 3, пять чисел (3,7, 4, 5,4). Значит, число 3 – медиана.

Аналогично показываем, что числа 4 и 3,22 являются медианами.

Желательный результат обсуждения. Этот пример показывает нам, что иногда медиана определяется неоднозначно. Вычисляя медиану на практике, мы будем придерживаться правила: если чисел в наборе четное число, медианой является среднее арифметическое двух серединных чисел.

Некоторые свойства медианы. Медиана, так же, как и среднее арифметическое, «двигается» и «сжимается» вместе со всем набором данных.

Пример 3. На протяжении суток каждые 3 часа записывали результат измерения температуры воздуха (в градусах Цельсия).

Табл. 1. Измерения температуры, °C

Время	Темп. °C	Время	Темп. °C	Время	Темп. °C
00:00	18	09:00	14	18:00	22
03:00	14	12:00	19	21:00	27
06:00	12	15:00	23	24:00	16

а) Найдите медиану этих данных.

б) Предположим, для публикации в канадском журнале эти данные нужно перевести в шкалу Фаренгейта. Для перевода температуры в градусах Цельсия в градусы Фаренгейта используется формула:

$$T°\text{F} = 1{,}8 \cdot T°C + 32.$$

Можно ли найти медиану в шкале Фаренгейта, не переводя в шкалу Фаренгейта все девять измерений? Найдите медиану в °F.

Ответ: а) 18°C; б) 64,4°F.

Устойчивость медианы к выбросам[2]

[2] До сих пор мы не обсуждали, как определить, является ли значение выбросом или нет. На этом этапе обучения понятие выброса остается интуитивным. Впоследствии будет предложено некоторое правило.

Пример 4. Предположим, что в числовом наборе 10 чисел. Мы помним, что если увеличивать или уменьшать одно число набора (двигать его), то среднее арифметическое будет двигаться в ту же сторону, но в 10 раз медленнее. А как поведет себя медиана? Проще всего разобраться в этом на примере. Рассмотрим набор из первых 10 натуральных чисел:

1, 2, 3, 4, **5, 6**, 7, 8, 9, 10

Среднее арифметическое и медиана совпадают. Они равны 5,5.

а) Увеличим последнее число на 10, а потом еще на 100. Как изменятся среднее и медиана?

б) Увеличим первое число на 10, а потом еще на 100. Как изменится среднее и медиана в этом случае?

Решение. а) Если увеличить последнее число 10 на 10, то среднее арифметическое увеличится на 1 и станет равным 6,5. А медиана останется прежней (5,5): она зависит только от двух серединных чисел 5 и 6, которые не изменялись. Если увеличить последнее число еще на 100, то среднее арифметическое вырастет еще на 10 и теперь будет 16,5. Медиана и в этом случае не изменится.

б) Увеличим теперь первое число 1 на 10. Получится 11. Среднее вырастет на 1 и станет равно 6,5. Так как вариационный ряд теперь изменился, медиана изменилась тоже. Серединными числами нового набора

2, 3, 4, 5, **6, 7**, 8, 9, 10, 11

являются не числа 5 и 6, как прежде, а числа 6 и 7. Следовательно, медиана теперь тоже равна 6,5. Но дальнейшее увеличение первого числа уже не изменит медиану.

Если теперь первое число увеличить еще на 100, получится набор

2, 3, 4, 5, **6, 7**, 8, 9, 10, 111,

где среднее стало равно $6{,}5+10=16{,}5$, а медиана уже больше не изменилась. Она равна 6,5.

Желательный результат обсуждения. На простом примере мы увидели, что один выброс может повлиять на медиану, но влияние это слабое. Сильно «сдвинуть» медиану с места одно «гуляющее» значение набора не может. Это свойство называют *устойчивостью медианы по отношению к выбросам*. Именно это свойство делает медиану очень полезной мерой центра в современной статистике.

Медианный представитель. Вспомним задачу из урока 23 про площади озер. Нужно было найти медиану данных о площади поверхности десяти крупнейших озер мира.

Табл. 2. Площадь водной поверхности крупнейших озер мира (кв.км)

	Озеро	Площадь		Озеро	Площадь
1	Каспийское море	376000	6	Аральское море	51100
2	Верхнее	82100	7	Танганьика	32900
3	Виктория	68100	8	Байкал	31500
4	Гурон	60000	9	Бол.Медвежье	31326
5	Мичиган	57800	10	Ньяса	30800

Медиана равна 54450 кв.км. Мы назвали медианными представителями озеро Мичиган и Аральское море, поскольку их площади близки к медиане.

Если даны числовые характеристики каких-то объектов (озер, городов, стран и т.п.), то объекты, у которых характеристика близка к медиане, называют медианными представителями этой характеристики. Медианных представителей может быть один или несколько. Возникает вопрос – что считать значением, «близким к медиане». Однозначного ответа нет. Все зависит от конкретной задачи и целей исследования.

Пример 5. В таблице 3 даны высоты крупнейших вулканов Срединного вулканического пояса Камчатки. Найдите медиану данных о высотах и медианного представителя (или медианных представителей). Подумайте, как можно разумно сформулировать правило или правила для выбора медианного представителя по высоте в данном случае.

Желательный результат обсуждения. Найти медиану легко без вычислений. Вулканы в таблице упорядочены по убыванию высоты. Всего их 62, деление в таблице на правый и левый столбцы проходит точно посередине. Медиана равна 1679,5 м. Ближайшими по высоте являются вулканы Кайленэй и Северный Черпук. Они и являются медианными представителями.

Выводы. Медиана набора числовых данных иногда определяется неоднозначно. Но на практике в таких случаях в качестве медианы следует брать среднее арифметическое двух серединных чисел.

При одинаковом изменении всех данных медиана меняется вместе с ними таким же образом. Однако медиана не меняется или мало меняется (устойчива) при изменении одного-двух элементов массива. Это дает свойство устойчивости относительно выбросов.

Рекомендуемое домашнее задание: см. приложение.

Дополнительные материалы для подготовки к уроку. Учитель на свое усмотрение может свободно использовать таблицы[3], размещенные на сайте «Вероятность в школе» http://ptlab.mccme.ru/node/350.

[3] 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.

Табл.3. Высоты вулканов СВП Камчатки, м над уровнем моря

	Вулкан	Высота, м над у.м.		Вулкан	Высота, м над у.м.
1	Ичинская сопка	3621	32	Северный Черпук	1679
2	Хувхойтун	2618	33	Гречишкина	1651
3	Алней	2581	34	Ука	1643
4	Острая Сопка	2539	35	Междусопочный	1641
5	Чашаконджа	2526	36	Титила	1559
6	Шишель	2525	37	Уксичан	1550
7	Иктунул	2300	38	Бонгапчи	1550
8	Снежный	2211	39	Лангуткин	1534
9	Снеговой	2172	40	Кэбеней	1529
10	Спокойный	2171	41	Большая Кытэпана	1502
11	Кевеэйтунул	2133	42	Еловский	1381
12	Белый	2080	43	Иэттунул	1340
13	Вулкан Горного института	2024	44	Большой Чекчебонай	1338
14	Новограбленова	2000	45	Леутонгей	1333
15	Хангар	2000	46	Мутный	1315
16	Малый Паяллан	1980	47	Большой	1301
17	Южный Черпук	1962	48	Теклетунул	1290
18	Лагерный	1961	49	Тильмыг	1265
19	Ульванэй	1954	50	Малый Чекчебонай	1261
20	Северный	1936	51	Плоский	1255
21	Чинейнейн	1922	52	Малая Кытэпана	1230
22	Большой Паяллан	1906	53	Воямпольский	1225
23	Апнгей	1856	54	Кахтана	1217
24	Анаун	1828	55	Тунупилянум	1200
25	Паяллан	1811	56	Ламутский	1198
26	Лелякина	1770	57	Слюнина	1175
27	Атласова	1764	58	Озерной	1021
28	Сергеева	1759	59	Шлен	1001
29	Каменистый	1758	60	Федотыч	965
30	Айнелькан	1725	61	Терпук	765
31	Кайленэй	1680	62	Киненин	583

Приложение

1. Юный статистик записывал в электронную таблицу результаты взвешивания шоколадных батончиков массой 50 г.

	A	B	C	D
1		Масса шок. батончиков, г		
2	49,8	53,1	47,8	52,4
3	50,2	50,4	49,3	4902
4	51,2	47,9	48,3	49,9
5	52,8	59,2	49,2	50,1
6				
7		Среднее:		354
8				50,15

Потом он посмотрел на среднее арифметическое (345 г) и очень удивился.
а) Почему среднее арифметическое получилось таким большим? Есть ли в данных выброс?
б) Как вы думаете, в чем может быть причина появления этого выброса? Как исправить ситуацию?
в) Найдите медиану данных до и после исправления ошибки.
г) Можно ли считать, что медиана ошибочных данных существенно отличается от медианы данных с исправленной ошибкой?
д) Вместо того, чтобы исправлять ошибку, удалите ошибочное число из расчетов. Останется 15 значений. Найдите теперь среднее арифметическое и медиану. Сильно ли они отличаются, по вашему мнению?

2. В таблице 4 дана плотность населения стран Южной Европы.

Табл. 4. Плотность населения стран Южной Европы

Страна	Плотность населения, чел/кв.км	Страна	Плотность населения, чел/кв.км
Черногория	46	**Албания**	105
Босния и Герцеговина	76	**Португалия**	117
Хорватия	79	**Андорра**	167
Македония	83	**Италия**	205
Греция	86	**Сан-Марино**	350
Испания	95	**Мальта**	1315
Словения	98	**Ватикан**	1818
Сербия	103	**Монако**	16500

а) Найдите медиану и среднее арифметическое данных. Какое среднее лучше отражает плотность населения страны Южной Европы?
б) В чем причина сильного отличия значений найденных средних?
в) Укажите медианного представителя (или представителей).

Ответы. 1. а) Есть выброс 4902 г. Это неправдоподобно много. б) Причиной выброса послужила, скорее всего, ошибка в записи. Вероятно, юный статистик случайно вместо запятой нажал клавишу с цифрой 0. в) Медианы равны 50,15 г и 50 г соответственно. г) Разность 0,15 г вряд ли можно считать существенной. д) Если удалить значение 4902, то среднее будет 50,77 г, а медиана 50,1 г. Разность составляет 0,67 г, что очень немного.
2. Среднее арифметическое 1327,7 чел/кв.км. Медиана 104 чел./кв.км. Такая разница из-за выброса 16500 (Монако). Возможно, Мальту и Ватикан тоже следует считать странами с особенно высокой плотностью. Медиана намного лучше описывает плотность населения типичной южноевропейской страны.