

МАТЕМАТИЧЕСКАЯ ВЕРТИКАЛЬ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И СТАТИСТИКА. 8 КЛАСС (2 ч./нед.)

Урок 3 – 4. Повторение. Диаграммы рассеивания. Случайная изменчивость

Документ содержит задачи для вводного повторения материала по статистике за 7 класс. Учитель может на свое усмотрение использовать материалы целиком или частично наряду с собственными разработками и материалами учебника¹. Авторы будут благодарны за замечания и предложения по структуре и содержанию сценариев.

Материал для повторения

Диаграмма рассеивания помогает определить примерный характер связи между двумя числовыми величинами при её наличии. По облаку данных можно строить гипотезы о наличии положительной или отрицательной связи между величинами, судить о силе этой связи, быстро находить выбросы данных.

В жизни в основном мы имеем дело с изменчивыми величинами. Чаще всего причины изменчивости известны лишь частично. Поэтому говорят о **случайной изменчивости**.

Рассмотрим измерения (например, взвешивание какого-нибудь товара или измерение какого-то расстояния). Изменчивость результатов при измерении вызывается систематическими и случайными погрешностями.

Систематическая погрешность не создает разброса данных, но приводит к отличию среднего значения от номинального (массы, объема, напряжения в сети и т.п.).

Случайные ошибки разнонаправлены — они дают отклонения то в меньшую, то в большую сторону, компенсируя друг друга. Поэтому случайные ошибки создают рассеивание значений, но мало влияют на среднее.

Точность измерения для разных величин должна быть разной. Слишком большая точность может быть экономически невыгодна, и даже вредна.

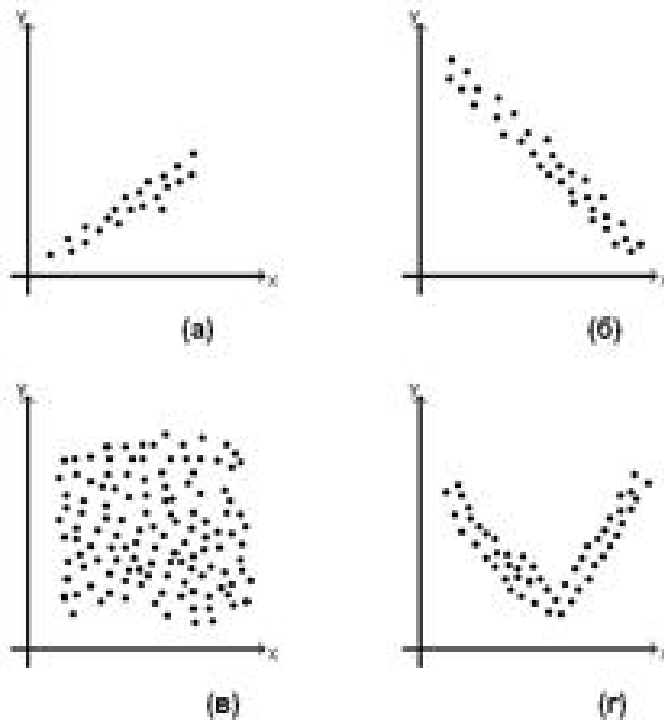
Один из способов представления изменчивой величины — **гистограмма** (диаграмма частот). По форме гистограммы можно указать интервалы наиболее правдоподобных значений величины. Иногда, сравнивая гистограммы двух величин, можно высказать предположение о том, что эти величины имеют сходный или различный характер изменчивости.

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.

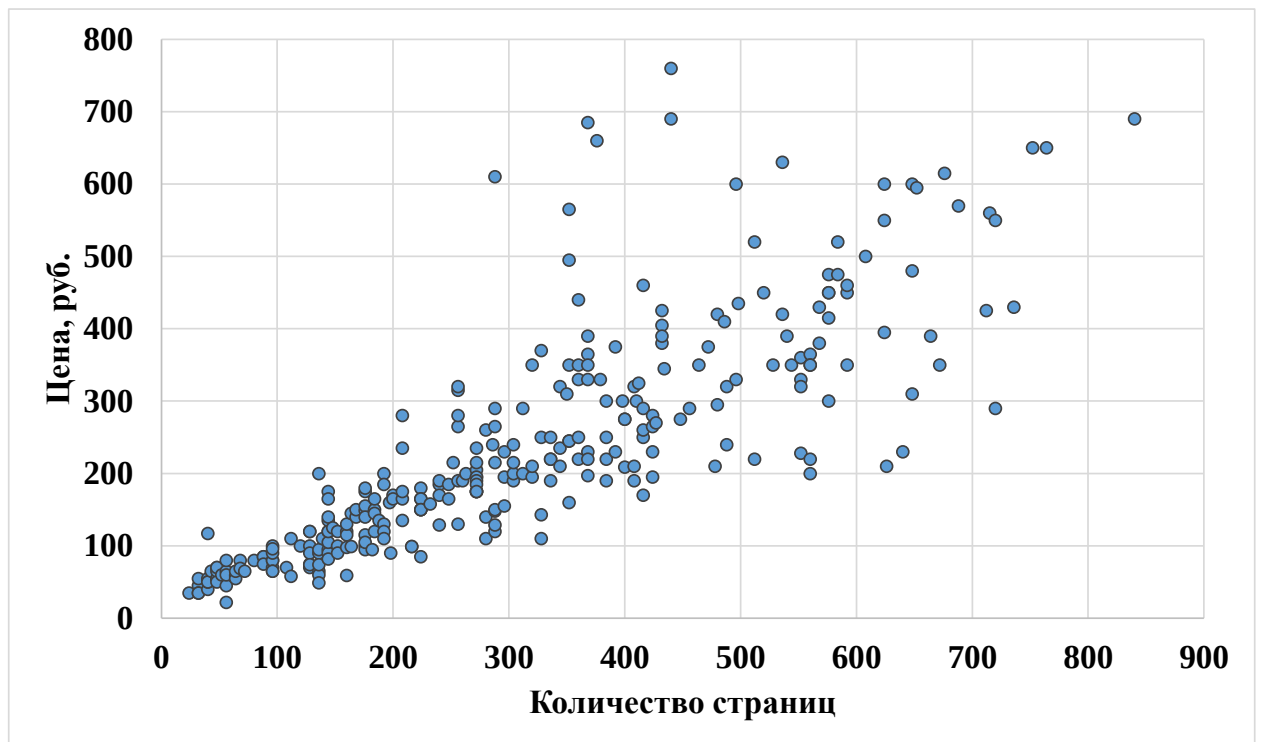
Задачи

Диаграммы рассеивания.

1. По диаграммам рассеивания определите, есть ли связь между наблюдаемыми величинами. Если да – охарактеризуйте её.



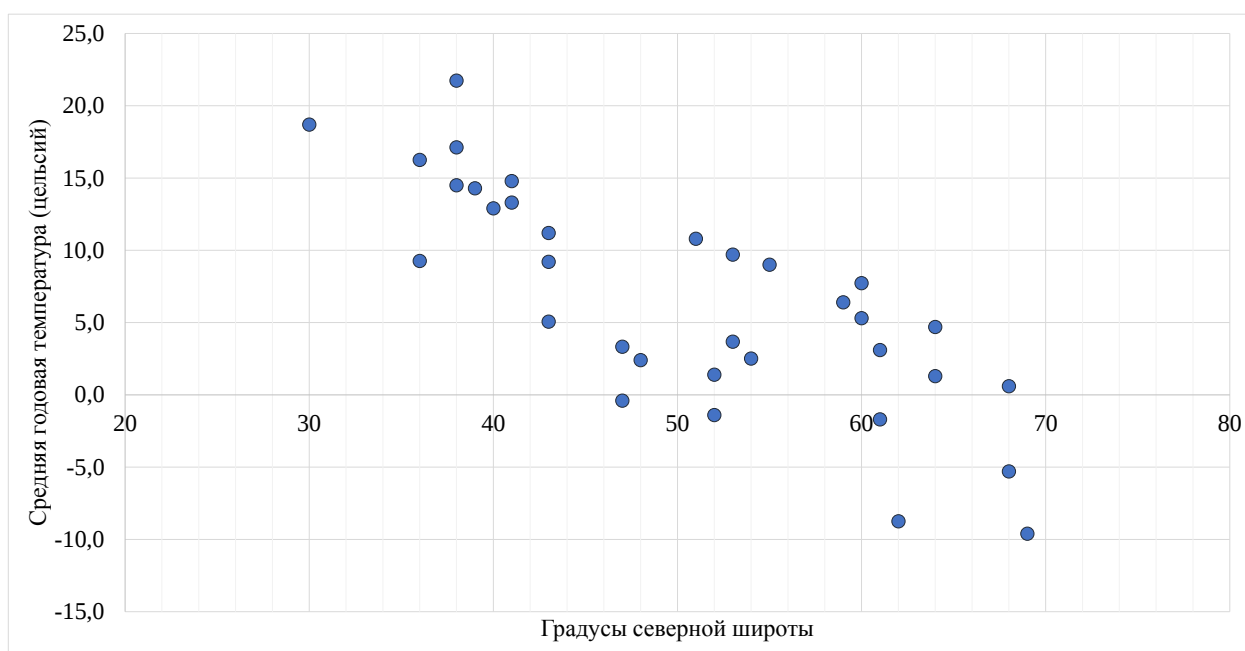
2. В книжной лавке продаётся около 300 наименований книг и журналов. По данным о цене книг и количестве страниц в них построена диаграмма рассеивания.



- а) Можно ли предположить, что есть связь между ценой книги и числом её страниц?
- б) Какие факторы, с вашей точки зрения, влияют на цену книги?
- в) Определите по диаграмме размах цены книг, у которых 250 ± 50 страниц.

3. По данным о средней годовой температуре и широте 34 городов северного полушария построена диаграмма рассеивания.

Город	Сев. шир., °	Ср годовая Т, °С	Город	Сев. шир., °	Ср годовая Т, °С
Берген	60	7,7	Воркута	68	-5,3
Токио	36	16,3	Чита	52	-1,4
Владивосток	43	5,1	Пекин	40	12,9
Южно-Сахалинск	47	3,3	Тбилиси	41	13,3
Петропавловск-Камчатский	53	3,7	Улан-Батор	47	-0,4
Катманду	30	18,7	Хабаровск	48	2,4
Самарканд	39	14,3	Владикавказ	43	9,2
Ланьчжоу	36	9,3	Мурманск	68	0,6
Душанбе	38	14,5	Осло	59	6,4
Якутск	62	-8,8	Архангельск	64	1,3
Алма-Ата	43	11,2	Петрозаводск	61	3,1
Ашхабад	38	17,1	Копенгаген	55	9
Исламабад	38	21,7	Лондон	51	10,8
Ташкент	41	14,8	Дублин	53	9,7
Петропавловск	54	2,5	Рейкьявик	64	4,7
Иркутск	52	1,4	Норильск	69	-9,6
Хельсинки	60	5,3	Сургут	61	-1,7



- а) Найдите координаты центра облака данных;
- б) Проведите ось диаграммы и найдите уравнение оси.

в) При помощи полученной формулы связи между средней годовой температурой и широтой оцените среднюю температуру в Бангкоке (Таиланд), расположенного на уровне 13,5 градусов северной широты.

Ответы и решения задач

1. На диаграммах а) и б) точки группируются как бы вдоль прямых, поэтому возможны предположения о существовании линейной зависимости между X и Y : а) положительная, б) отрицательная связь. На диаграмме в) связь между X и Y практически отсутствует. На диаграмме г) вероятно между X и Y существует некоторая зависимость, но не линейная.

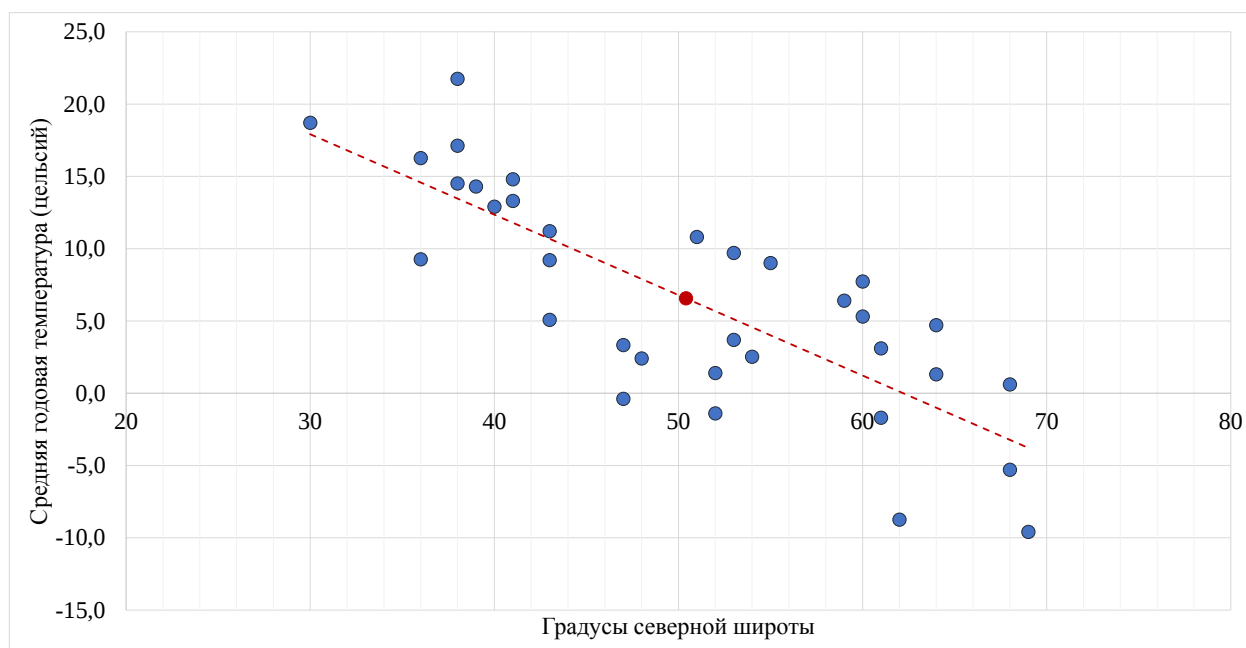
2. а) Облако данных очевидно наклонено вправо-вверх, что говорит о наличии возможной зависимости между ценой книги и количеством страниц.

б) На цену книги влияет качество бумаги, наличие иллюстраций, качество переплёта, число страниц, количество использованных цветов, покупательский спрос, талант автора и другие факторы.

в) Размах можно оценить в 250 рублей, если не учитывать выброс (книга, стоимостью более 600 рублей). Если его учитывать, то размах приблизительно равен 520 рублям.

3. а) (50,4; 6,6);

б) Пример оси диаграммы:



Для этой оси найдем уравнение.

График линейной функции проходит через точку (62;0) и центр облака.

$$\begin{cases} 6,6 = 50,4k + b \\ 0 = 62k + b \end{cases} \rightarrow 6,6 = -11,6k \rightarrow k \approx -0,57, \text{ откуда } b = -62k \approx -62 \cdot (-0,57) = 35,3.$$

Уравнение оси: $y = -0,57x + 35,5$.

в) По имеющемуся уравнению² $y = -0,57 \cdot 13,5 + 35,5 \approx 27,8^\circ\text{C}$.

² Получившиеся значения будут зависимости от того, как ученик провёл ось. Значение должно получиться близкое к средней годовой температуре в Бангкоке: $+28,7^\circ\text{C}$.

Случайная изменчивость

1. В таблице представлены данные о численности населения Москвы за 7 лет.

Год	2013	2014	2015	2016	2017	2018	2019
Население, тыс. чел	11 979,5	12 108,3	12 197,6	12 330,1	12 380,7	12 506,5	12 615,3

- а) С какой точностью представлены данные о численности населения столицы?
- б) Является ли население Москвы в разные годы постоянной величиной?
- в) Какие случайные факторы могут влиять на численность населения столицы?
- г) Постройте диаграмму данных, какая тенденция численности населения имеет место в последние годы?

2. Подумайте, какие из предложенных измерений можно сделать непосредственно, а какие — только косвенно, и почему:

- а) Измерение массы Сатурна;
- б) Измерение высоты дерева;
- в) Измерение площади поверхности Тихого океана;
- г) Измерение массы арбуза;
- д) Измерение высоты полёта самолета;
- е) Измерение площади тетрадного листа.

3. На пачке масла написана масса. В паспорте золотого изделия тоже указана масса. Очень важно знать массу самолета. Каждый человек тоже знает свой вес. Упорядочите по возрастанию точности измерения массы:



4. На фабрике по производству моющих средств стоят два одинаковых дозирующих автомата. Каждый автомат должен отмерять порции жидкого мыла объемом 240 мл. При проверке оборудования инженеры измерили с большой точностью объем 10 порций, отмеренных каждым автоматом. Результаты измерений представлены в таблице.

Автомат	Номер пробы	1	2	3	4	5	6	7	8	9	10
1	Объем, мл	236	233	235	233	234	231	236	235	232	230
2	Объем, мл	242	249	231	240	233	245	243	237	240	250

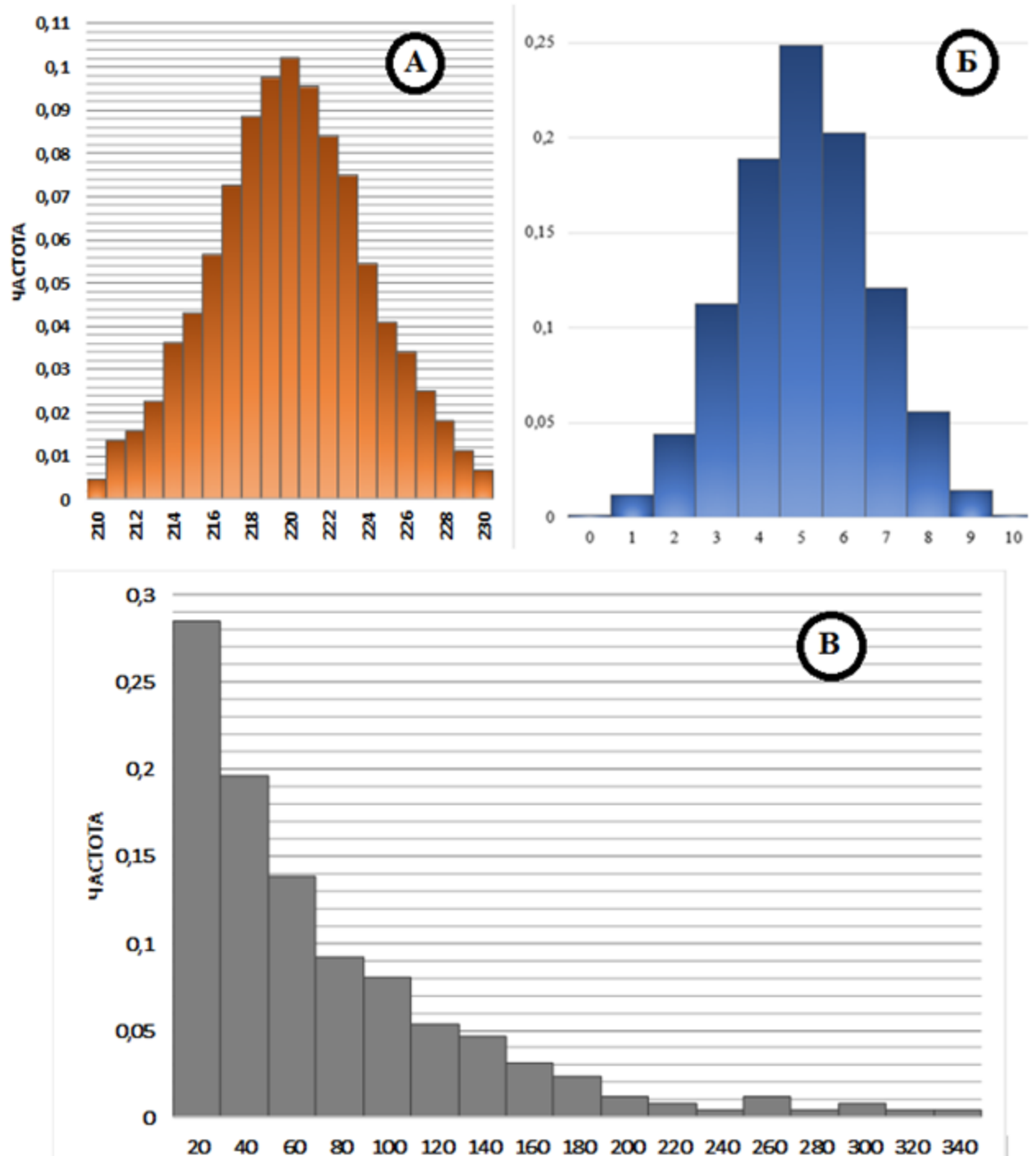
Попробуйте сделать выводы о состоянии дозирующих автоматов. Какой из них в первую очередь нуждается в ремонте?

5. Результаты многократных измерений трёх изменчивых величин представлены на гистограммах А, Б и В. Определите, какой величине какая из гистограмм соответствует. Объясните свое решение.

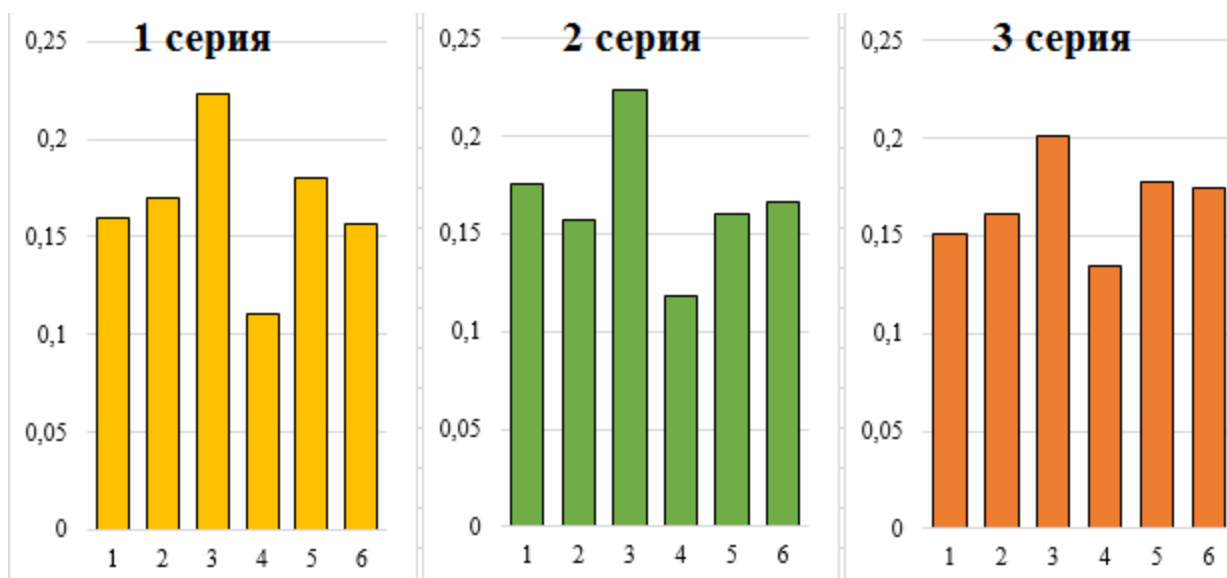
ЭКСПЕРИМЕНТЫ

- 1) напряжение в бытовой электросети;
- 2) длительность телефонных разговоров абонента (в секундах);
- 3) число орлов, выпавших при бросании 10 монет.

ГИСТОГРАММЫ

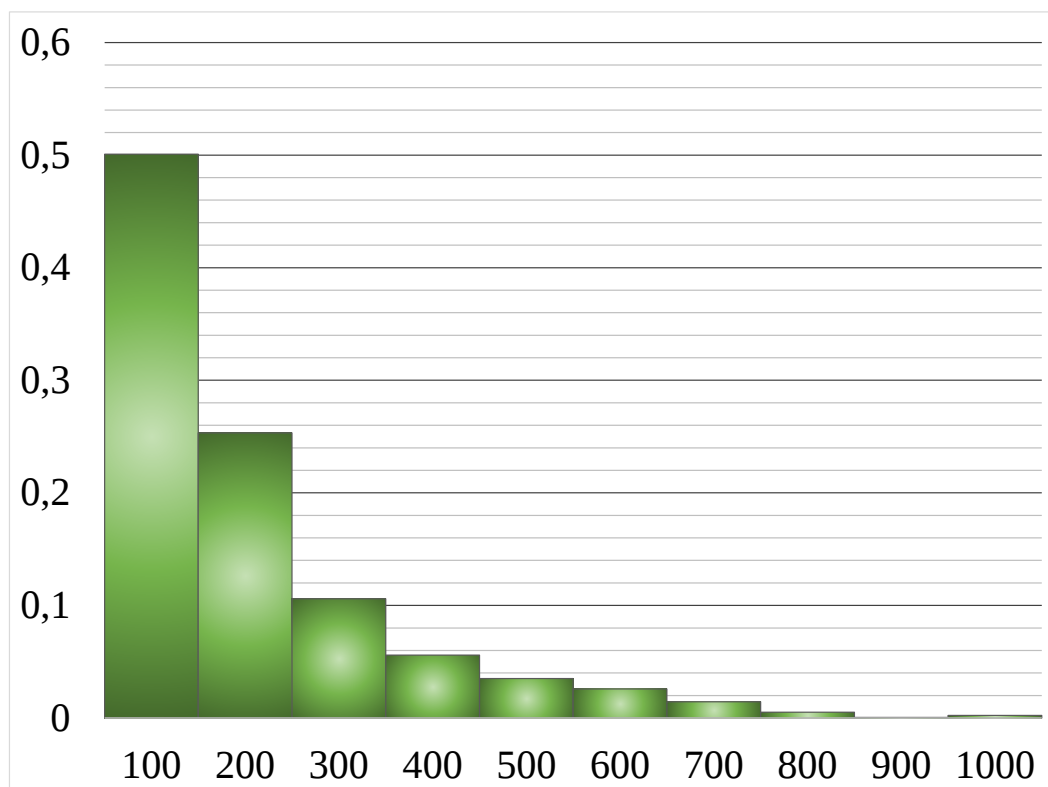


6. Коля решил проверить, симметричен ли игральный кубик, который он слепил из хлебного мякиша. Он провёл три серии испытаний, в каждой из которых он многократно бросал кубик и подсчитывал частоты выпадения каждой грани. Результаты всех трёх серий представлены на гистограммах.



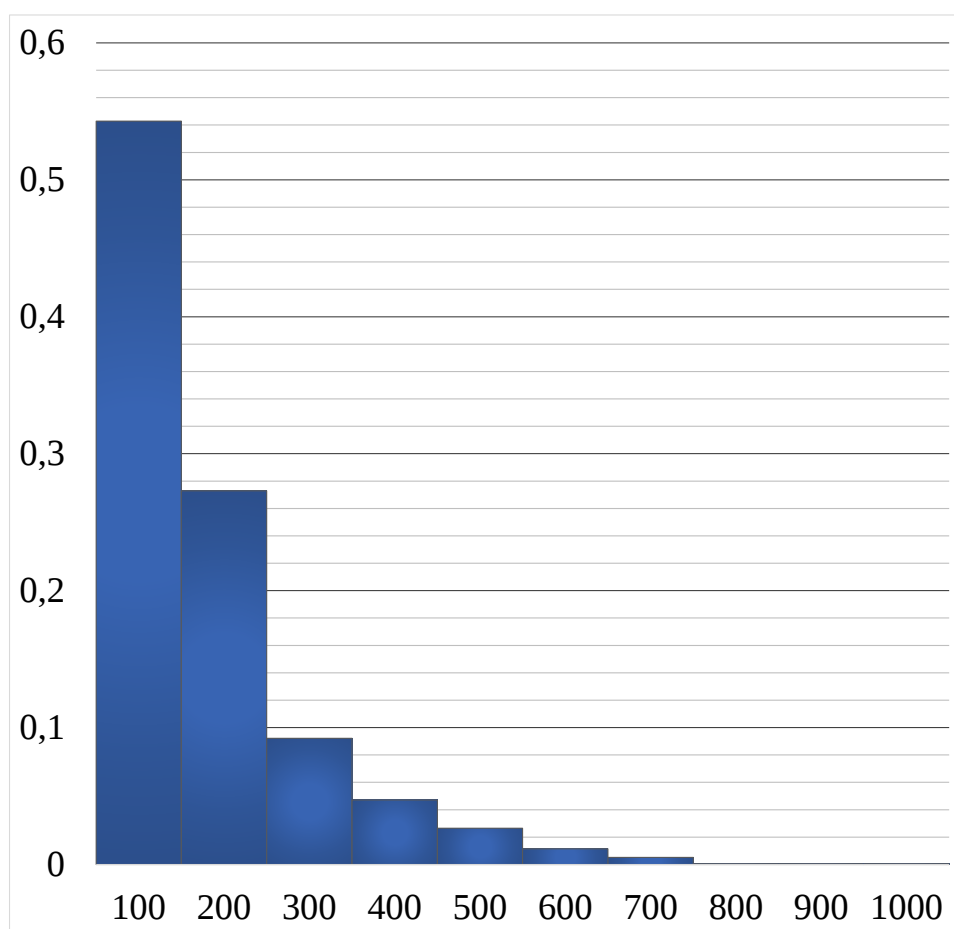
Коля посчитал кубик несимметричным. Как мог рассуждать Коля? Попробуйте найти аргументы «за» и «против» гипотезы «кубик симметричный».

7. В книжном магазине продаётся примерно 1700 различных книг. На гистограмме представлены данные о распределении различных книг по количеству страниц (шаг группировки — 100 страниц).



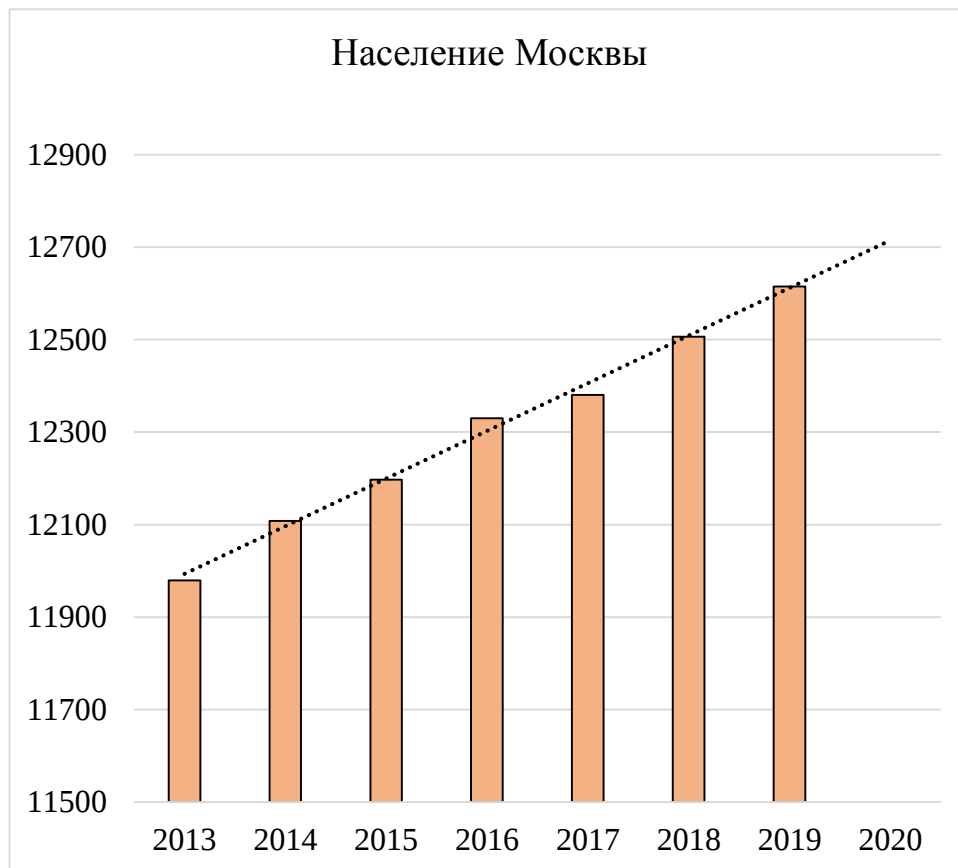
Ответьте на вопросы:

- а) Верно ли, что книг, в которых не более 100 страниц, приблизительно столько же, сколько тех, в которых страниц больше 100?
- б) Верно ли, что книг, в которых страниц менее 300, около 10%?
- в) Найдите примерно частоту события «в книге не больше 200 страниц».
- г) Найдите примерно частоту события «в книге больше 200 страниц».
- д) На гистограмме ниже представлено распределение цен на книги в том же магазине (шаг группировки — 100 рублей). Можно ли сказать, что величины «цена» и «объём» схожи по характеру изменчивости? Приведите доводы за и против этого предположения.



Ответы и решения задач

1. а) С точностью до ста человек; б) нет, население Москвы — изменчивая величина; в) естественная миграция населения, рождаемость и смертность и т.п.; г) пример диаграммы:



Наблюдается рост населения столицы.

2. Непосредственно можно измерить высоту небольшого дерева, массу арбуза, площадь тетрадного листа. Размеры объектов и их доступность это позволяют. Недоступные или очень большие объекты измеряются косвенно (масса Сатурна, площадь поверхности Тихого океана, высота полёта самолёта).

3. По возрастанию точности (уменьшению допустимой погрешности) измерения массы:

- 1) Самолет;
- 2) Человек;
- 3) Пачка масла;
- 4) Золотое кольцо.

4. Найдем средние значения для каждого автомата. У первого автомата среднее равно 233,5 мл (отличается от номинала на 6,5 мл), а у второго 241 мл (отличается от номинала на 1 мл). Казалось бы, со вторым автоматом все в порядке — он дает небольшую ошибку в среднем. Но, чтобы разобраться в ситуации, представим измерения наглядно, отметив для каждого автомата полученные значения точками на числовой прямой (средние отмечены красным). Видно, что порции мыла, выдаваемые вторым автоматом, сильно разбросаны по объему. Иными словами, случайная величина «объем порции» у второго автомата имеет большое рассеивание. Такая случайная изменчивость скорее всего вызвана износом деталей, так что второй автомат определённо требует ремонта. У первого автомата разброс данных небольшой

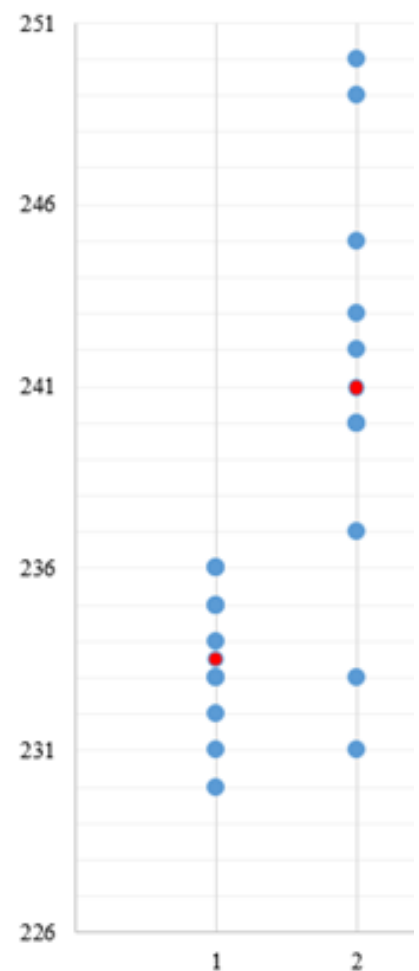
(это говорит о хорошем состоянии механизма в целом), такую систематическую ошибку можно исправить регулировкой.

5. 1 – А, 2 – В, 3 – Б.

6. В каждой серии тройка выпадает чаще других чисел, а противоположная грань — четверка — реже. Есть основание предположить, что центр тяжести кубика смещён, и он не является правильным. У симметричного кубика частоты выпадения граней могут значительно различаться от опыта к опыту, но здесь во всех трёх опытах мы видим преобладание одних и тех же граней над другими. Без вычислений можно предположить, что такое совпадение маловероятно.

7. а) верно; б) неверно; в) пригл. 0,81; г) пригл. 0,19.

д) Гистограммы внешне похожи, можно предположить схожий характер изменчивости или зависимость — чем больше в книге страниц, тем она в среднем дороже. Но это только предположение, наверняка утверждать нельзя.



Самостоятельная работа (10 мин)

Вариант 1

1. Найдите среднее арифметическое и размах числового набора:

–4,7 3,2 0,7 4,5 –6,2

2. В таблице представлены данные о населении и о годовой выработке электроэнергии за 2017 год девяти стран, которые являются мировыми лидерами по производству электроэнергии.

Страна	Население, тыс. чел.	Годовая выработка электроэнергии, млрд кВт·ч
Китай	1395814	6529
Индия	1359741	1541
США	333337	4251
Бразилия	209737	585
Россия	146781	1090
Япония	126220	1101
Германия	83214	653
Южная Корея	53733	579
Канада	36086	712

а) Найдите медиану годовой выработки электроэнергии в представленных странах.

б) Найдите медианного представителя — страну, в которой годовая выработка электроэнергии ближе всего к медиане.

в) Средняя годовая выработка электроэнергии в этих девяти странах составляет приблизительно 1893,4 млрд кВт·ч. Какой из показателей — среднее арифметическое или медиана — лучше характеризует годовую выработку типичного крупного производителя электроэнергии? Кратко обоснуйте свое мнение.

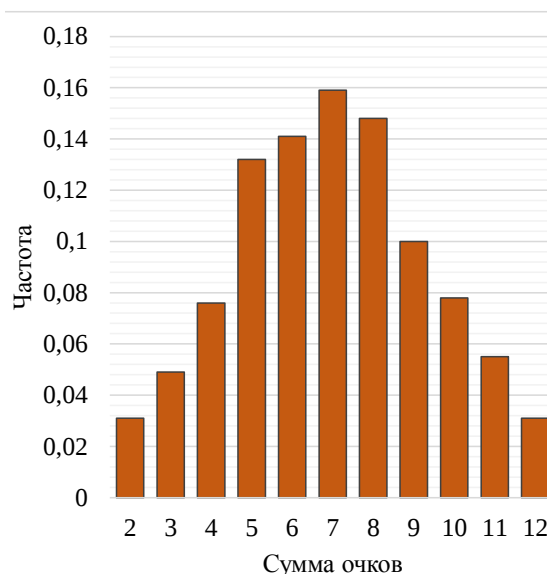
г) Найдите количество электроэнергии, вырабатываемой на душу населения (в кВт·ч/чел.) в Южной Корее. Округлите результат до целых.

3. **Эксперимент.** Бросают два игральных кубика и считают сумму выпавших очков. Действие повторяется многократно. По результатам эксперимента построена гистограмма.

Определите по гистограмме приблизительно частоту случаев, когда:

а) сумма выпавших очков равна 7;

б) сумма выпавших очков меньше 5.



Вариант 2

1. Найдите среднее арифметическое и размах числового набора:

9,1 -0,6 1,7 2,9 -2,6

2. В таблице представлены данные о населении и о годовой добыче нефти в 2018 году в девяти странах, которые являются мировыми лидерами по нефтедобыче.

Страна	Население, тыс. чел.	Добыча нефти, млн. баррелей
Китай	1395814	1533
США	333337	4271
Бразилия	209737	1022
Россия	146781	3906
Мексика	133141	1095
Иран	82979	1497
Саудовская Аравия	42248	3869
Канада	36086	1424
ОАЭ	9541	1168

а) Найдите медиану годовой добычи нефти в представленных данных.

б) Найдите медианного представителя — страну, в которой годовая нефтедобыча ближе всего к медиане.

в) Средняя годовая добыча нефти в этих девяти странах в 2018 г. составила приблизительно 2198 млн. баррелей. Какой из показателей — среднее арифметическое или медиана — лучше характеризует годовую добычу нефти типичной крупной нефтяной державы? Кратко обоснуйте свое мнение.

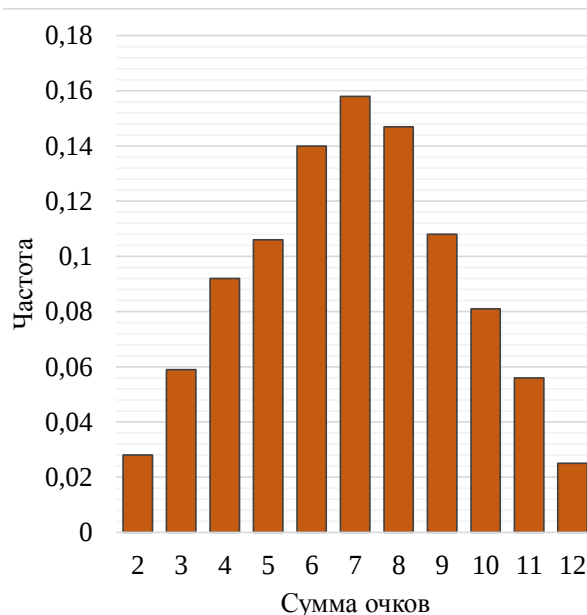
г) Найдите годовой объём нефтедобычи на душу населения (баррелей/чел.) в Бразилии. Округлите результат до десятых.

3. Эксперимент: Бросают два игральных кубика и считают сумму выпавших очков. Действие повторяется многократно. По результатам эксперимента построена гистограмма.

Определите по гистограмме приблизительно частоту случаев, когда:

а) сумма выпавших очков была равна 7;

б) сумма выпавших очков была больше 9.



Ответы:

№	Вариант 1	Вариант 2
1.	Среднее: $-0,5$ Размах: $10,7$	Среднее: $2,1$ Размах: $11,7$
2.	а) $1090 \cdot 10^9$ кВт·ч; б) Россия; в) Медиана $1090 \cdot 10^9$ кВт·ч сильно отличается от среднего $1893,4 \cdot 10^9$ кВт·ч. Страны, где выработка электроэнергии близка к $1893,4 \cdot 10^9$ кВт·ч, в таблице нет. Среднее слишком большое, потому что на него повлияли две страны с очень высокой выработкой электроэнергии — Китай и США. Годовую выработку типичного крупного производителя электроэнергии лучше характеризует медиана. г) $10\,776$ кВт·ч/чел.	а) 1497 млн. баррелей; б) Иран; в) Медиана 1497 млн. баррелей сильно отличается от среднего 2198 млн. баррелей. Страны, где нефтедобыча близка к 2198 млн. баррелей, в таблице нет. Среднее слишком большое, потому что на него повлияли три страны с очень большим количеством добываемой нефти — Россия, Саудовская Аравия и США. Годовую добычу нефти типичной крупной нефтяной державы лучше характеризует медиана. г) $4,9$ баррелей/чел.
3.	а) $0,159$; б) $0,076 + 0,048 + 0,032 = 0,156$;	а) $0,158$; б) $0,081 + 0,056 + 0,024 = 0,161$;