Информация и ее роль в управлении системами

Основные вопросы лекции:

1. Информация и требования к ней. Роль информации в процессе моделирования и управления сложными системами.
2. Статистика и ее задачи.
3. Основы первичной обработки статистических данных.
4. Вариационные ряды и их графическое изображение.
5. Характеристики положения и рассеяния массива статистических данных.
6. Случайные события. Вероятность случайного события и ее свойства. Сумма и произведение вероятностей.
7. Случайные величины и способы их задания.
8. Основные числовые характеристики случайных величин.

Вопрос 1. Информация и ее свойства. Роль информации в процессе моделирования и управления сложными системами

Для решения вопросов, связанных с организацией ГПС, определением ее численности во всех подсистемах, органах управления и подразделениях, эффективностью организационных структур и всей деятельности пожарной охраны, планированием подготовки кадров и многих других, необходимо прежде всего располагать соответствующей информацией.

Информация – любые сведения и данные об окружающем нас мире, относящиеся (или нет) к интересующей нас проблеме, которые можно накапливать, хранить, перерабатывать, которыми можно обмениваться и пр.

Термин "информация" происходит от латинского слова "informatio", что означает сведения, разъяснения, изложение.

В обиходе информацией называют любые данные или сведения, которые кого-либо интересуют. "Информировать" в этом смысле означает "сообщить нечто, неизвестное раньше".

Информация — сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

Одно и то же информационное сообщение (статья в газете, объявление, письмо, телеграмма, справка, рассказ, чертеж, радиопередача и т.п.) может содержать разное количество информации для разных людей — в зависимости от их предшествующих знаний, от уровня понимания этого сообщения и интереса к нему.

Так, сообщение, составленное на японском языке, не несет никакой новой информации человеку, не знающему этого языка, но может быть высокоинформативным для человека, владеющего японским. Никакой новой информации не содержит и сообщение, изложенное на знакомом языке, если его содержание непонятно или уже известно.

Информация есть характеристика не сообщения, а соотношения между сообщением и его потребителем. Без наличия потребителя, хотя бы потенциального, говорить об информации бессмысленно.

Подходы к определению информации

Традиционный (обыденный):

Информация – это сведения, знания, сообщения о положении дел, которые человек воспринимает из окружающего мира с помощью органов чувств (зрения, слуха, вкуса, обоняния, осязания).

Вероятностный (используется в теории об информации):

Информация – это сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределенности и неполноты знаний.

Знания делят на две группы:

- декларативные – от слова декларация (утверждения, сообщения) начинаются со слов «Я знаю, что …»;
- процедурные – определяют действия для достижения какой-либо цели, начинаются со слов «Я знаю, как …».

Информация может существовать в виде:

- текстов, рисунков, чертежей, фотографий;
- световых или звуковых сигналов;
- радиоволн;
- электрических и нервных импульсов;
- магнитных записей;
- жестов и мимики;
- запахов и вкусовых ощущений;
- хромосом, посредством которых передаются по наследству признаки и свойства организмов, и т. д.

Носители информации – среда или физическое тело для передачи, хранения и воспроизведения информации.

Что можно делать с информацией:

создавать	принимать	комбинировать	хранить
передавать	копировать	обрабатывать	искать
воспринимать	формализовать	делить на части	измерять
использовать	распространять	упрощать	разрушать
запоминать	преобразовыва ть	собирать	и т. д.

Все эти процессы, связанные с определенными операциями над информацией, называются **информационными процессами**.

Информационные процессы - это процессы, связанные с получением, хранением, обработкой и передачей информации (т.е. действия, выполняемые с информацией); процессы, в ходе которых изменяется содержание информации или форма ее представления.

Классификация информации

по способам восприятия:

- визуальная
- аудиальная
- тактильная
- обонятельная
- вкусовая

по формам представления:

- текстовая
- числовая
- графическая
- музыкальная
- комбинированная

по общественному значению:

Массовая - обыденная, общественно-политическая, эстетическая
Специальная - научная, техническая, управленческая, производственная
Личная – наши знания, умения, интуиция

Свойства информации

объективность – не зависит от чьего-либо мнения

достоверность – отражает истинное положение дел

полнота – достаточна для понимания и принятия решения

актуальность – важна и существенна для настоящего времени

ценность (полезность, значимость) – обеспечивает решение поставленной задачи, нужна для того чтобы принимать правильные решения

понятность (ясность)– выражена на языке, доступном получателю

доступность – соответствие уровню восприятия получателя

краткость (лаконичность) – сжатость, опускание несущественных деталей

Вопрос 2. Статистика и ее задачи

Термин "статистика" происходит от латинского слова status - состояние. В XVIII в., когда статистика начала оформляться в научную дисциплину, термин "статистика" связывался с системой описания фактов, характеризующих состояние государства (данные о населении, промышленности, сельском хозяйстве и т.д.), так как без этого

Статистика - это наука, изучающая количественные соотношения и закономерности массовых явлений с учетом их качественного своеобразия.

Задача статистики - обеспечение необходимой информацией соответствующих государственных органов для разработки планов и анализа хода их выполнения, т. е. статистика служит одним из инструментов организации управления в общественной жизни.

Пожарная статистика – раздел статистической науки, изучающий количественные отношения и закономерности таких массовых явлений, как пожары и связанные с ними процессы и явления.

Широко распространенное представление о том, что с помощью статистики можно доказать все, что угодно, является слишком преувеличенным, однако несомненно и то, что статистические методы могут применяться с целью ввести людей в заблуждение.

Бенджамин Дизраэли (англ. Benjamin Disraeli, родился 21.12.1804 в Лондоне; умер 19.04.1881 там же) - британский политический деятель и романист. Премьер-министр в 1868 и в 1874-1880. Создатель современной консервативной партии Великобритании.

«Есть три вида лжи: ложь, наглая ложь и статистика».

Другие цитаты:

- Когда про Вас начинают ходить анекдоты, пора на покой.
- Почти все великое сделано молодыми.
- Без сильной оппозиции не устоит ни одно правительство.
- В цивилизованной стране перемены неизбежны.

Этапы статистического исследования

1) формулирование задачи (определение цели исследования, множества единиц, составляющих статистическую совокупность; отбор характеристик единиц статистической совокупности, которые включаются в исследование);

2) сбор статистического материала (статистическое наблюдение или эксперимент);

3) обработка статистических данных (группировка, вычисление обобщающих показателей, отыскание соответствующих форм зависимости);

4) всесторонний качественный анализ изучаемых явлений и процессов.

Первичные статистические данные в основном получают путем наблюдений. Каждое наблюдение состоит в измерении (оценке) значений определенных характеристик у одной из единиц статистической совокупности.

Совокупность наблюдений образует массив статистических данных.

Всякое измерение производится по определенной шкале, определяющей уровень измерения.

Конфу́ций (около 551 до н. э. — 479 до н. э.)

Китайский мыслитель и философ.

«Человек измеряется не с ног до головы, а от головы до неба»

В статистике различают признаки качественные (атрибутивные) и количественные.

Качественные признаки не имеют количественного выражения, варианты этого признака различаются качественным содержанием и выражаются словесно (например, цвет, форма).

Количественные признаки измеряются какой-то мерой, варианты таких признаков отличаются друг от друга по величине.

Количественные признаки делятся на дискретные (прерывные) и непрерывные.

Варианты дискретных признаков отличаются друг от друга на некоторую конечную величину (чаще всего на целое число).

Значения непрерывного признака могут отличаться друг от друга на сколь угодно малую величину и принимать любое промежуточное значение в границах варьирования признака.

Шкалы измерений

Вид данных	Качественные данные		Количественные данные	
Тип шкалы измерений	**Шкала наименовани й**	**Шкала порядка**	**Шкала интервалов**	**Шкала отношений**
Допустимые вопросы к данным	Кто? Что? Какой? Где?			
		Что больше? Что лучше?		
			На сколько А больше В?	
				Во сколько раз А больше В?
Примеры данных	цвет, имя, пол, № игрока, причина и место пожара	сорт, балл, воинское звание, степень огнестойкост и	температура по Цельсию, время и дата вызова на пожар	возраст, вес, рост, цена, объем, длина, площадь пожара
Среднее по совокупност и	**Мода** – наиболее часто встречающееся значение			
		Медиана – середина ра ного ряда		
			Средне тическое	

$$x = \sum_{i=1}^{N} \frac{x_i}{N}$$

Вопрос 3. Основы первичной обработки статистических данных

Множество данных, полученных в результате статистического наблюдения, называется статистической совокупностью.

Пусть, например, ставится задача оценить частоту поступления вызовов пожарных подразделений в сутки. Выпишем для этого из диспетчерского журнала последовательность целых чисел, означающих число вызовов пожарных подразделений за сутки. Допустим, она имеет вид :

4, 1, 2, 0, 0, 3, 0, 1, 5, 2, 0, 0, 1, 0, 3, 4, 1, 0, 2, 1

Всего здесь перечислено 20 элементов. Число элементов называют объемом статистической совокупности. Такая последовательность чисел, отражающая случайный характер поступления вызовов (или возникновения пожаров в городе) во времени, а в более общем случае, влияние на некоторую величину случайных обстоятельств, неконтролируемо меняющихся от наблюдения к наблюдению, называется статистическим рядом.

Если отдельные наблюдения расположить в порядке возрастания (убывания) указанных выше значений, то получим ранжированный ряд:

0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 3, 3, 4, 4, 5

Ранжированный ряд более удобен для анализа, чем неупорядоченный, но и такая запись в целом малонаглядна и занимает много места.

Для компактного представления массива статистических данных, а также для систематизации и структуризации содержащихся в них сведений, применятся прием, называемый группировкой.

Группировка – разделение статистической совокупности на группы, в каждую из которых включаются единицы, имеющие одинаковые или сходные значения изучаемых характеристик.

Простая (одномерная) группировка	производится на основе одной характеристики (например, распределение числа пожаров по месяцам года или по дням недели)
Комбинационная (многомерная) группировка	производится на основе нескольких характеристик (например, совместное распределение числа пожаров по причинам и по объектам)

Если статистическая совокупность, содержащая *M* единиц, разделена на *n* групп, то представляет интерес расчет следующих характестик:

1) частота m_k (k=1, 2, …, n) – число единиц, попавших в k-ю группу;

2) частость ω_k (или относительная частота) – доля, которую составляют частоты от общего числа единиц *M*:

$$\omega_k = \frac{m_k}{M}$$

Имеют место соотношения:

$$\sum_{k=1}^{n} m_k = M$$

$$\sum_{k=1}^{n} \omega_k = \sum_{k=1}^{n} \frac{m_k}{M} = 1(100\%)$$

Частотное распределение для статистической совокупности представляет собой перечень различных значений (или групп значений) изучаемой характеристики с указанием соответствующих абсолютной m_k и относительной ω_k частот (обычно в виде таблицы).

Вопрос 4. Вариационные ряды и их графическое изображение

Число вызовов в сутки	Абсолютная частота (число суток с указанным числом вызовов)	Относительная частота (частость)
0	7	0,35
1	5	0,25
2	3	0,15
3	2	0,10
4	2	0,10
5	1	0,05
Всего:	20	1,00

Такую таблицу называют вариационным рядом. Он характеризует изменение (варьирование) исследуемого количественного признака (числа вызовов в сутки) в исследуемой статистической совокупности и представляет собой ряд частот, соответствующих вариантам (значениям) признака.

Вариационные ряды подразделяют на дискретные и интервальные.

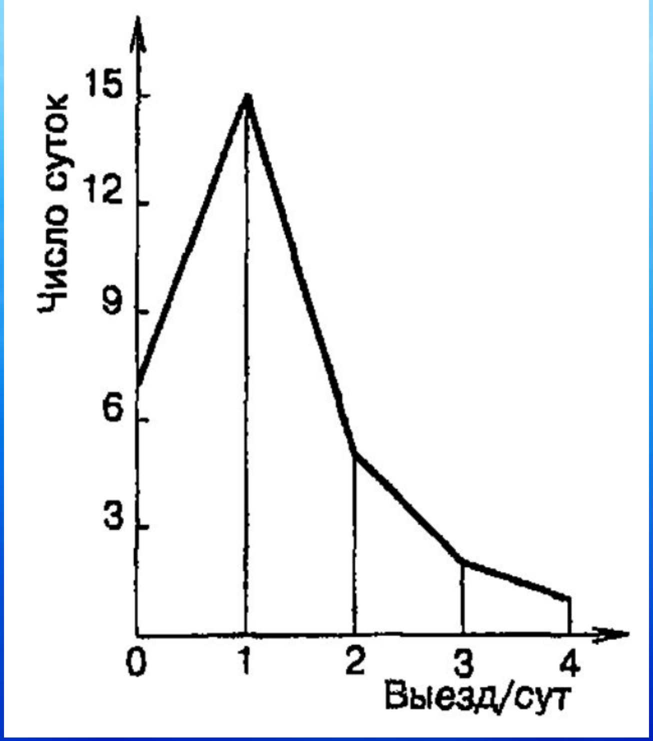
Если множество возможных значений дискретного признака достаточно велико и их трудно (или даже невозможно) перечислить, то всю совокупность разбивают на части (интервалы), равные или неравные, и подсчитывают частоты значений, попадающих в тот или иной интервал. Такой ряд называют интервальным вариационным рядом.

Распределение продолжительности обслуживания вызова

Интервал времени, мин	Абсолютная частота (число вызовов)	Относительная частота (частость)
[0; 30)	237	0,37
[30; 60)	186	0,28
[60; 90)	124	0,19
[90; 120)	75	0,11
[120; 150)	29	0,04
150 и более	8	0,01
Всего:	659	1,00

Вариационные ряды графически могут быть изображены в виде полигона, гистограммы, кумуляты и огивы.

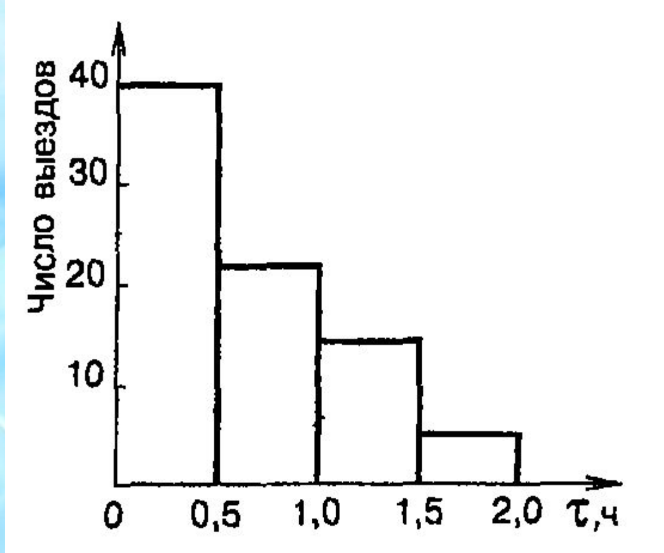
Полигон чаще всего используется при графическом изображении дискретного вариационного ряда.



Число выездов в сутки	0	1	2	3	4
Число суток	7	15	5	2	1

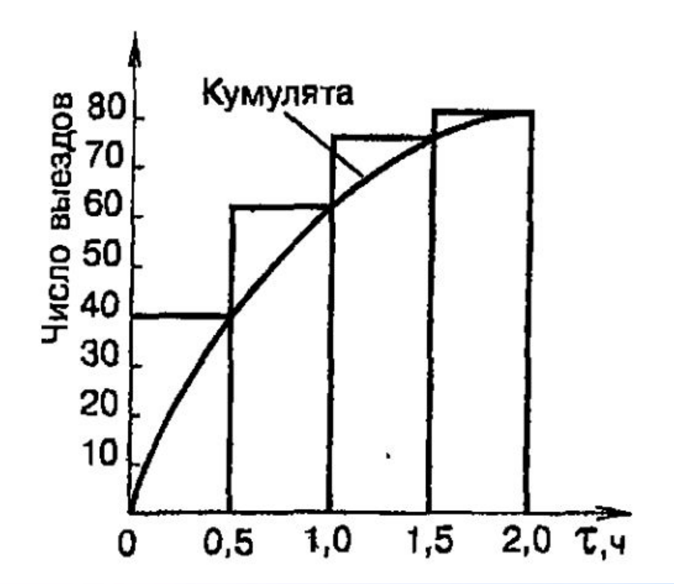
Гистограмма используется только при графическом изображении интервальных вариационных рядов.

В отличие от полигона при построении гистограммы на оси абсцисс берутся не точки, а отрезки, изображающие интервалы, а вместо ординат строят прямоугольники с высотой, пропорциональной частотам, частостям или плотностям интервалов, под которыми понимают отношение частот (частостей) к величине интервала.



Интервал, ч	0 - 0,5	0,5 - 1,0	1,0 - 1,5	1,5 - 2,0
Число выездов	40	22	14	5

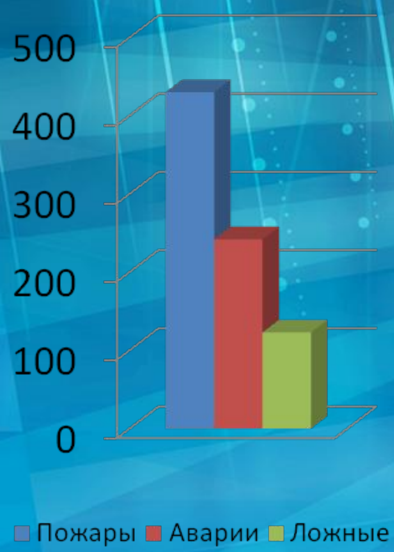
Кумулятивная кривая (кривая сумм - кумулята) получается при изображении вариационного ряда с накопленными частотами (частостями).

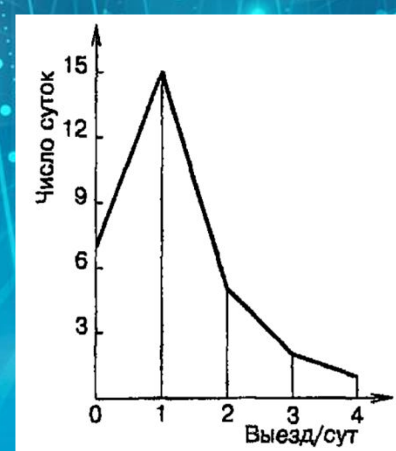


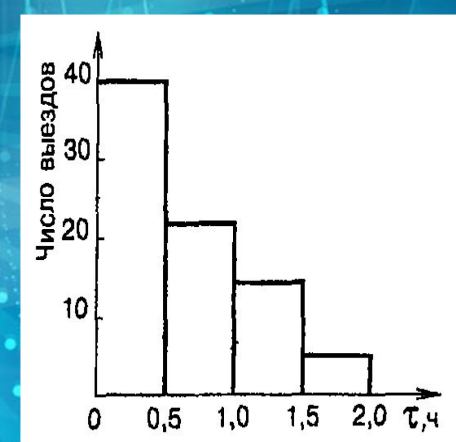
Интервал, ч	0 - 0,5	0,5 - 1,0	1,0 - 1,5	1,5 - 2,0
Число выездов	40	22	14	5
Накопленное число выездов	40	62	76	81

Огива строится аналогично кумуляте с той разницей, что на оси абсцисс наносят накопленные частоты, а на оси ординат - значения признака (т. е. строят график функции, обратной функции, изображаемой кумулятой).

Виды данных	Качественные (атрибутивные)	Количественные	
		дискретные	непрерывные
Способ группировки данных	Атрибутивный ряд	Вариационный ряд	
		дискретный	интервальный
Графическое изображение	Диаграммы	Полигон	Гистограмма







Вопрос 5. Характеристики положения и рассеяния массива статистических данных

На практике часто нет необходимости характеризовать полностью исследуемый признак (случайную величину). Обычно бывает достаточно указать только его отдельные числовые параметры, в известной степени характеризующие существенные черты распределения значений признака.

Такие характеристики, назначение которых - выразить в сжатой форме наиболее существенные особенности распределения - называются числовыми характеристиками признака.

Характеристики положения. Средние величины

В статистике различают средние величины:

- арифметическую
- геометрическую
- гармоническую
- квадратическую
- кубическую

и др.

Средняя арифметическая

Пусть имеем статистическую совокупность объемом n элементов, причем значения исследуемого признака таковы: **x_1, x_2, …, x_n**. Тогда средняя арифметическая данного признака вычисляется по формуле:

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Эта формула дает простую (невзвешенную) среднюю арифметическую и используется для ряда, в котором значения признака не сгруппированы. Если же имеем дискретный вариационный р

x_1	x_2	...	x_k
m_1	m_2	...	m_k

в котором значения признака сгруппированы, то необходимо использовать формулу средней взвешенной арифметической данного дискретного ряда:

$$\overline{x} = \frac{m_1 x_1 + m_2 x_2 + \ldots + m_k x_k}{m_1 + m_2 + \ldots + m_k} = \frac{\sum_{i=1}^{k} m_i x_i}{\sum_{i=1}^{k} m_i}$$

Если использовать частости, то среднюю взвешенную арифметическую можно

представить в виде:

$$\bar{x} = \omega_1 x_1 + \omega_2 x_2 + \ldots + \omega_k x_k = \sum_{i=1}^{k} \omega_i x_i$$

При вычислении взвешенной средней арифметической интервального ряда в качестве значений варианта принимают середины интервалов.

Например:

Интервал	$[x_1, x_2)$	$[x_2, x_3)$	...	$[x_{k-1}, x_k)$
Частота	m_1	m_2	...	m_{k-1}

имеем:

$$\bar{x} = \frac{m_1 \frac{x_1 + x_2}{2} + m_2 \frac{x_2 + x_3}{2} + \ldots + m_{k-1} \frac{x_{k-1} + x_k}{2}}{m_1 + m_2 + \ldots + m_{k-1}} = \frac{\sum_{i=1}^{k-1} m_i \frac{x_i + x_{i+1}}{2}}{\sum_{i=1}^{k-1} m_i}$$

Медиана

Рассмотрим ранжированный вариационный ряд. Его медианой называют значение признака, которое находится в середине этого ряда.

Очевидно, если в дискретном вариационном ряду содержится нечетное число 2m + 1 возможных значений признака, то медианным будет (m+1)-е значение признака, т. е.

$$Me = x_m + 1.$$

Если же ряд имеет четное число 2m значений, то:

$$Me = \frac{x_m + x_{m+1}}{2}$$

Основное свойство медианы состоит в том, что сумма абсолютных величин отклонений вариантов (т. е. значений изучаемого признака) от медианы меньше, чем от любой другой величины (в том числе и от средней арифметической):

$$\sum_{i=1}^{n} \left| x_i - Me \right| = \min \left\{ \left| x_i - \bar{x}_{гм} \right|, \left| x_i - \bar{x}_{ap} \right|, \ldots \right\}$$

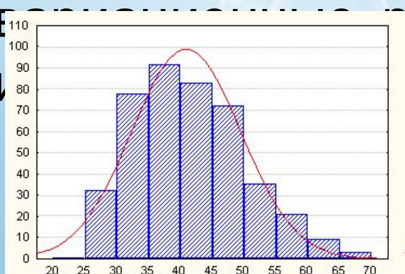
Мода

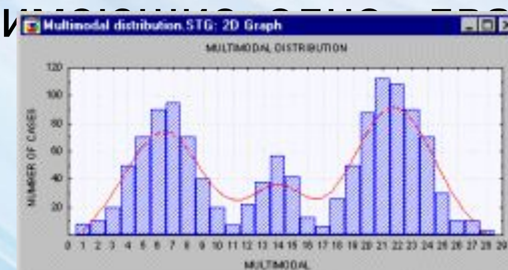
Модой **Мо** называют значение признака, наиболее часто встречающееся в данном вариационном ряду. Мода, как средняя величина, употребляется чаще для данных имеющих нечисловую природу. Среди перечисленных причин пожаров — поджог, детская шалость с огнем, поджог, нарушение ПУЭ, грозовой разряд, поджог — мода будет равна поджогу.

Для дискретного ряда мода, являющаяся характеристикой вариационного ряда, определяется по частотам вариантов и соответствует варианту с наибольшей частотой.

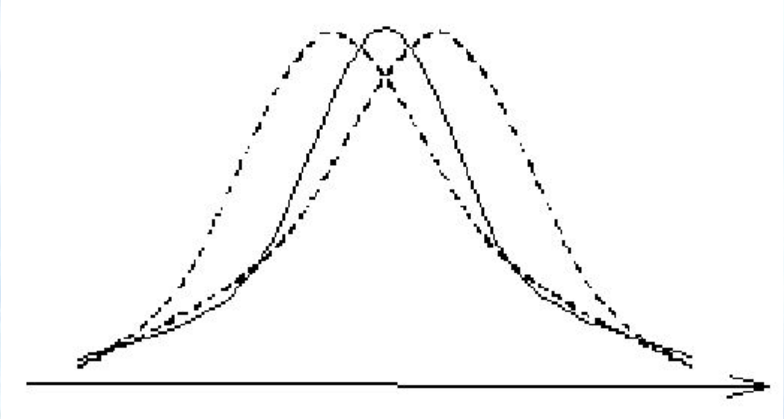
Для интервального вариационного ряда с равными интервалами модальный интервал (т. е. интервал, содержащий моду) тоже определяется по наибольшей частоте, а при неравных интервалах - по наибольшей плотности.

Бывают унимодальные, бимодальные и полимодальные ...ряды, т. ...ряды, и... ...одальны...





Симметричный вариационный ряд - это ряд, в котором частоты вариантов, равноотстоящих от средней, равны между собой. В противном случае вариационный ряд называют асимметричным (скошенным). Различают левостороннюю и правостороннюю скошенность.



Необходимым (но недостаточным) условием симметричности является:

$$\overline{x}_{ap} = Mo = Me$$

Характеристики рассеяния

Средние величины, характеризуя вариационный ряд одним числом, не учитывают вариацию (разброс значений) признака, между тем как она очень существенна для характеристики вариационного ряда.

Необходимы специальные числовые характеристики рассеяния значений признака около средней.

Вариационный размах

Простейшей мерой рассеяния является вариационный размах (иногда говорят, широта распределения), который представ-ляет собой разность между наибольшим и наименьшим значениями варьируемого признака в ряду его значений:

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

Эта мера рассеяния (размах) весьма неустойчива и может менять свои значения при каждом новом наблюдении за данным признаком. Она может служить лишь ориентировочной характеристикой вариации признака.

Среднее линейное отклонение

Для оценки рассеяния можно также брать среднее значение абсолютной величины отклонений значений x_i признака от его среднего значения:

$$\rho = \frac{\sum_{i=1}^{n} \left| x_i - \bar{x} \right|}{n}$$ или, учитывая частоты: $$\rho_{взв} = \frac{\sum_{i=1}^{k} \left| x_i - \bar{x} \right| m_i}{\sum_{i=1}^{k} m_i}$$

Дисперсия

Дисперсией называют среднее арифметическое квадратов отклонений значений признака изучаемой совокупности от их среднего значения:

$$D = \sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

Недостаток дисперсии заключается в том, что ее размерность равна квадрату размерности исследуемого признака.

Среднее квадратическое отклонение (стандарт)

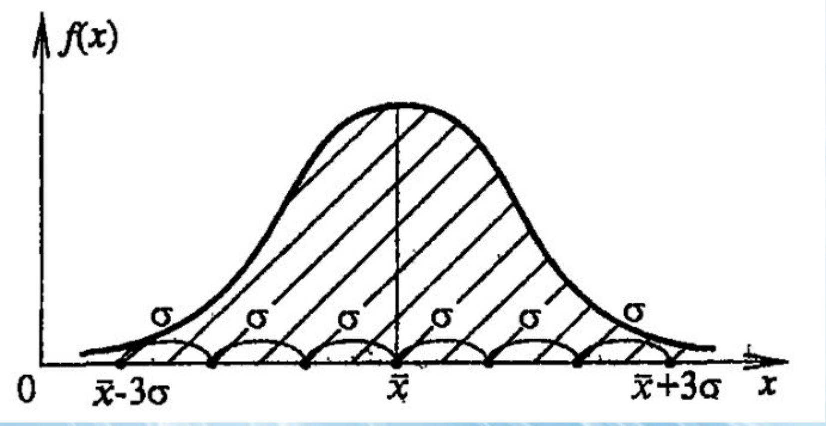
Средним квадратическим отклонением (**стандартом**) называют квадратный корень из дисперсии:

$$\sigma = \sqrt{D}$$

Среднее квадратическое отклонение - очень удобная и наглядная, характеристика рассеяния. Она сразу же дает представление о размахе колебаний значений признака около среднего значения.

Для большинства встречающихся на практике распределений значений признаков с достоверностью можно утверждать, что эти значения отклоняются от своего среднего арифметического не больше чем на ***3σ***.

Это утверждение, наиболее точно выполняющееся для так называемого нормального распределения, носит название правила "трех сигма".



Коэффициент вариации (изменчивости)

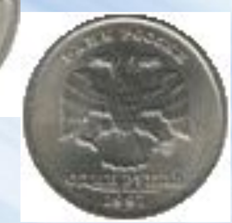
Коэффициентом вариации называют отношение среднего квадратического отклонения к среднему арифметическому, выраженное в процентах:

$$\upsilon = \frac{\sigma}{\bar{x}} 100\%$$

Коэффициент вариации служит для сравнения величин рассеяния вариационных рядов, причем из двух вариационных рядов тот имеет большее рассеяние, у которого коэффициент вариации больше.

Вопрос 6. Случайные события. Вероятность случайного события и ее свойства. Сумма и произведение вероятностей.

Случайным событием называют такое событие, которое обладает статистической устойчивостью при многократном воспроизведении данного комплекса условий, но при однократной его реализации событие может как произойти, так и не произойти.



Классическое определение вероятности

Пусть данный комплекс условий имеет ***n*** равновозможных и исключающих друг друга исходов (результатов опыта), из которых m_A - число случаев, благоприятствующих событию ***A*** (означающих его появление). Тогда вероятностью случайного события ***A*** называют отношение:

$$P(A)=\frac{m_A}{n}$$

Статистическое определение вероятности

Пусть производится достаточно большое число определенных испытаний, в каждом из которых может появиться или не появиться событие ***A***. Предположим, что провели серию из ***n*** испытаний, в которой событие ***A*** появилось m_A раз. Тогда ту постоянную величину ***P(A)***, вокруг которой будет колебаться относительная частота частость) m_A/n при все увеличивающемся числе испытаний ***n***, называют статистической вероятностью случайного события ***A***. Условно это можно обозначить так:

$$P(A)=\frac{m_A}{n} \quad при \quad n\to\infty$$

Таким образом, при увеличении числа опытов *n* частость события приближается к его вероятности, но не с абсолютной достоверностью, а с большой вероятностью, тем большей, чем большее число опытов произведено. В таких случаях говорят, что имеет место сходимость по вероятности, а относительную частость m_A/n называют эмпирической оценкой вероятности $P(A)$, имея в виду, что:

$$P(A) \approx \frac{m_A}{n}$$ при достаточно большом *n*

Геометрический подход к понятию вероятности

Пусть система испытаний заключается в следующем. Из некоторого оружия вслепую, наудачу (скажем, в темноте) производится стрельба по достаточно большой плоской мишени, попадание в которую гарантировано (т. е. промахнуться невозможно). Случайное событие *A* заключается в том, что пуля (снаряд, ракета) попадет не просто в большую мишень, а в некоторую ее часть (определенную зону). Очевидно, чем больше площадь зоны S_o по сравнению с площадью мишени S, тем вероятнее попадание в эту зону (и наоборот). В таком случае естественно под вероятностью случайного события *A* понимать отношение площадей S_o и S:

$$P(A) = \frac{S_0}{S}$$

Свойства вероятности случайного события

1) Вероятность любого случайного события **A** удовлетворяет следующему условию:

$$0 \le P(A) \le 1$$

2) Вероятность достоверного события **D** всегда равна 1:

$$P(D) = 1$$

Достоверным событием называют такое событие, которое всегда произойдет при данном комплексе условий. В этом случае $m_A = n$ для любой серии испытаний.

3) Вероятность невозможного события **H** всегда равна нулю:

$$P(H) = 0$$

Невозможным событием называют событие, которое никогда не произойдет при данном комплексе условий, т. е. здесь всегда $m_A = 0$.

Замечание. Если событие невозможно, то его вероятность равна нулю, но обратное утверждение неверно.

Вспомогательные понятия, относящиеся к случайным событиям

Полная группа событий. Говорят, что несколько событий при данном комплексе условий образуют полную группу, если в результате реализации этого комплекса (проведения эксперимента) неизбежно должно появиться хотя бы одно из этих событий.

Несовместные события. Несколько событий в данном опыте называются несовместными, если никакие два из них не могут осуществиться вместе.

Независимые события. Событие А называют независимым от события В, если его вероятность не зависит от того, произошло событие В или нет.

Противоположные события. Два события называют противоположными, если появление одного из них означает невозможность появления другого.

Противоположные события обычно обозначают A и $\overline{A}$ (не A). Следовательно, наступление события A означает невозможность наступления противоположного события $\overline{A}$. Очевидно, что противоположные события образуют полную группу событий и являются несовместными.

“Алгебра” событий

Суммой двух событий А и В называют событие С=А+В, состоящее в выполнении события А или В, или обоих событий вместе.

Суммой нескольких событий называют событие, состоящее в выполнении хотя бы одного из этих событий. Очевидно, что сумма противоположных событий является достоверным событием D:

$$A+\overline{A}=D$$

Произведением двух событий А и В называют событие F=AB, состоящее в совместном выполнении события А и события В.

Произведением нескольких событий называют событие, состоящее в совместном выполнении всех этих событий.

Существуют правила сложения и умножения вероятностей случайных событий, позволяющие вычислять вероятности сложных событий, если известны вероятности событий их составляющих.

Правило сложения вероятностей. Пусть случайные события А и В имеют вероятности появления Р(А) и Р(В). Предположим, что они несовместные. Тогда справедливо равенство :

$$P(A+B)=P(A)+P(B)$$

означающее, что вероятность суммы двух несовместных событий равна сумме их вероятностей.

Правило умножения вероятностей. Пусть случайные события А и В - независимы и имеют вероятности появления Р(А) и Р(В). Тогда справедливо равенство:

$$P(AB)=P(A)\cdot P(B)$$

означающее, что для независимых случайных событий вероятность произведения событий равна произведению их вероятностей.

Вопрос 7. Случайные величины и способы их задания

Случайной величиной называют такую величину, которая в результате эксперимента (реализации данного комплекса условий) примет одно и только одно значение из множества возможных значений, однозначно предсказать которое заранее невозможно, но имеет смысл рассматривать вероятность принятия данного значения.

Случайные величины можно подразделить на два основных класса в зависимости от структуры множества их возможных значений:

дискретные - возможные значения случайной величины представляют собой конечное или счетное множество (элементы которого можно перенумеровать), причем все значения изолированы и отличаются друг от друга на определенную величину;

непрерывные - множество возможных значений случайной величины сплошь занимает какой-то отрезок (интервал) числовой оси, конечный или бесконечный.

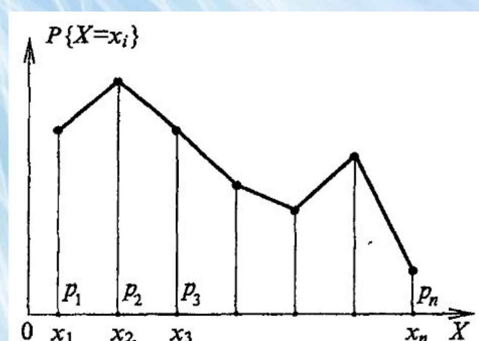
Способ задания дискретных случайных величин

Предположим, нас интересует дискретная случайная величина X. Для того чтобы полностью описать ее, достаточно указать все ее возможные значения $x_1, x_2, ..., x_n$ (здесь п - заданное целое число) и вероятности $P\{X=x_i\}=p_i$ где $i=1, 2,..., п$, с которыми эти значения принимаются. Обычно все эти значения записываются в виде таблицы.

X	x_1	x_2	...	x_n
$P\{X=x_i\}=p_i$	p_1	p_2	...	p_n

Таблицу называют законом распределения дискретной случайной величины.

Графически закон распределения дискретной случайной величины можно представить в виде ломаной линии – полигона.



Нормирующее условие:

$$\sum_{i=1}^{n} p_i = 1$$

Общий способ задания любых случайных величин

Способ задания дискретных случайных величин принципиально непригоден для описания непрерывных случайных величин.

Поэтому в теории вероятностей разработан универсальный способ задания любых случайных величин. Он заключается в следующем.

Пусть X - случайная величина и x – произвольное действительное число (точка на числовой оси). Рассмотрим неравенство $X<x$, означающее, что в данном опыте случайная величина X приняла значение меньше, чем x (лежащее левее на числовой оси, чем точка x).

Очевидно, что выполнение неравенства $X < x$ можно рассматривать как случайное событие, имеющее определенную вероятность реализации $P\{X<x\}$. Естественно, эта вероятность должна вполне определенным образом зависеть от значений x (т. е. каждому значению x будет соответствовать определенное значение вероятности $P\{X < x\}$).

Следовательно, вероятность $P\{X<x\}$ является некоторой функцией от x:

$$P\{X<x\}=F(x)$$

Функция F(x) называется функцией распределения случайной величины X и позволяет описать любую случайную величину (и дискретную, и непрерывную). Отсюда следует, что общим способом задания случайных величин является описание их с помощью соответствующих функций распределения.

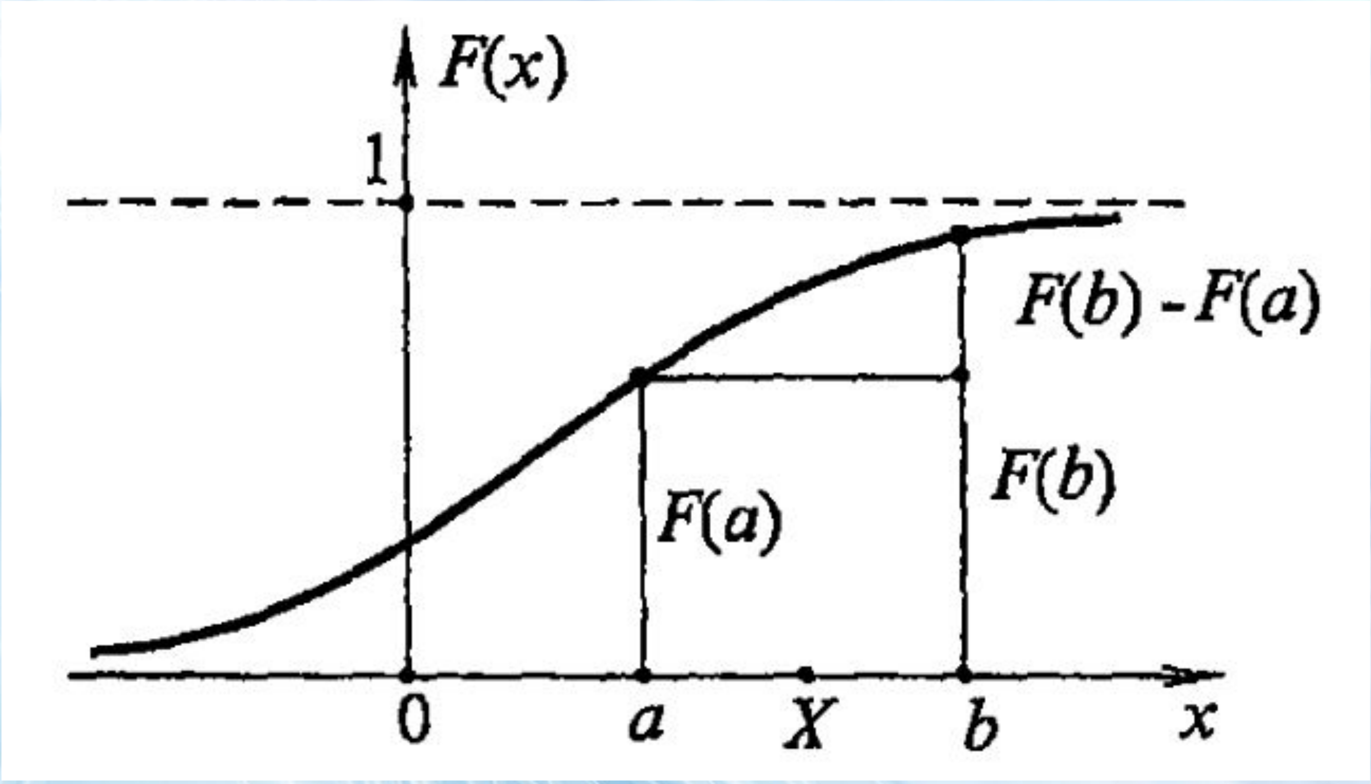


График функции распределения F(x)

Вопрос 8. Основные числовые характеристики случайных величин

Математическое ожидание характеризует положение случайной величины на числовой оси, т. е, указывает то среднее значение, около которого группируются все возможные ее значения.

Рассмотрим дискретную случайную величину X, принимающую конечное множество значений и заданную своим законом распределения:

X	x_1	x_2	...	x_{n-1}	x_n
$P\{X = x_i\}$	p_1	p_2	...	p_{n-1}	p_n

Математическим ожиданием дискретной случайной величины X называют выражение вида:

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n = \sum_{i=1}^{n} x_i p_i$$

При большом количестве испытаний выполняется условие:

$$M(X) \approx \bar{x}$$

Поэтому математическое ожидание случайной величины называют также ее средним значением.

Дисперсия характеризует степень рассеяния, разбросанности значений случайной величины около ее математического ожидания.

Дисперсией случайной величины X называют математическое ожидание квадрата отклонения этой величины от ее математического ожидания:

$$D(X) = M[X - M(X)]^2$$

Для дискретной случайной величины X с конечным множеством возможных значений это определение приводит к узнаваемому выражению:

$$D(X) = \sum_{i=1}^{n} (x_i - \bar{x})^2 p_i$$

В случае непрерывной случайной величины X дисперсия имеет вид интеграла:

$$D(X) = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx$$

Среднее квадратическое отклонение. Дисперсия имеет размерность квадрата случайной величины, но часто в качестве характеристики рассеяния удобнее пользоваться величиной, имеющей ту же размерность, что и сама случайная величина.

Поэтому в теории вероятностей, как и в статистике, вводят понятие среднего квадратического (квадратичного) отклонения, равного квадратному корню из дисперсии:

$$\sigma(X) = \sqrt{D(X)}$$