

Аликова А.М.

## К ВОПРОСУ СОСТАВА СОДЕРЖАНИЯ ЭЛЕМЕНТОВ СТОХАСТИКИ В ШКОЛЕ

Формирование представления о вероятностной природе окружающей действительности, знакомства со случайными явлениями – процесс не скорый. На начальном этапе изучения содержательной линии мы считаем, что следует дать осознать ощущение степени случайности в явлениях окружающей действительности, производить оценку шансов событий, когда число элементарных исходов испытания очевидно и невелико, осознать равновозможность исходов испытания. В процессе ознакомления со случайными явлениями следует изучать и свойства этих явлений, целенаправленно накапливать первоначальный статистический опыт.

Следующим этапом изучения данной линии, мы считаем, целенаправленное формирование вероятностно-статистических представлений. На этом этапе следует обобщение, систематизация первоначальных вероятностно-статистических знаний. Переход от интуитивных вероятностных представлений к дальнейшему формированию правильных представлений о статистических исследованиях, обработке и анализе данных, интерпретации результатов, восприятию статистического подхода к понятию вероятности, что будет создавать благоприятную почву для глубокого усвоения элементов теории вероятностей и математической статистики на этапе внутримодельного изучения. Здесь естественным образом возникает вопрос о систематизации, классификации и анализе информации о большем количестве объектов, представление полезного обзора и резюме основных характеристик статистических данных, об изучении не только объема совокупности, но и наиболее характерные ее черты: структуру, типичные проявления, изменчивость признака, следовательно рассмотрение более сложных видов таблиц и столбчатых диаграмм, гистограмм.

Органично должно вводиться понятия таблицы частот и гистограмм. В больших статистических данных для систематизации и классификации следует подсчет количество проявляемых единиц совокупности, частоты появления тех или иных объектов. Чтобы получить широкое представление об изучаемом явлении используются особые таблицы частот. Таблица частот будет состоять из целых неотрицательных чисел, показывающих сколько раз то или иное значение (или группа значений) было отмечено в опыте или наблюдении. Они могут быть образованы как по качественным, так и по количественным признакам и особенно, полезна для представления данных перечней чисел с относительно небольшим количеством различных значений. Визуальное представление о многих основных свойствах, конфигурации, характерных особенностях, для иллюстрации различных способов группировки набора данных в

целом дает графическое изображение статистических данных с помощью особого вида столбчатой диаграммы – гистограммы. Гистограмма позволит ответить на ряд вопросов следующего вида: –Какие значения типичны для этого набора данных? Как различаются между собой значения? Сконцентрированы ли данные вокруг некоторого типичного значения? Есть ли в этом наборе такие значения, которые настолько сильно отличаются от остальных, что требуют специальной обработки? Гистограммы демонстрируют частоты в виде диаграммы из столбиков, которые расположены над числовой осью и показывают, насколько часто различные значения встречаются в наборе данных. По горизонтальной оси откладываются измеренные значения из набора данных, по вертикальной частоту встречаемости этих значений. В дальнейшем гистограммы будут использоваться для характеристики меры возможности наступления события. Кроме гистограмм для графического представления результатов следует использовать и полигоны. Полигоны это ломанные, абсцисса вершин – результаты статистических исследований, а ординаты – соответствующие им числа, который позволит судить о значениях величины в определенных точках и характере их изменения. Эта ломаная, изображенная в декартовой системе координат, позволит охватить одним взглядом изображенное статистическое явление, отмечая при этом общие черты и особенности.

На этом этапе, следует ознакомить с понятийным аппаратом области знаний – статистики, необходимой каждому современному человеку. Здесь должно мягко осуществиться переход от раннее полученных знаниях о статистических опытах, способах представлений, статистических характеристиках – к комплексным статистическим исследованиям. Организации их, должны отражаться в доступных для учащихся примерах из различных областей жизни и науки. Естественным образом в ходе описаний исследований вводятся новые понятия, отражающие специфику данного исследования. Это следующие понятия. *Генеральная совокупность* – большие статистические объекты, которые интересуют нас, но мы не можем провести полное исследование. Так как полное исследование часто требует больших организационных усилий и финансовых затрат. Кроме того, осуществить полное исследование просто невозможно, так как это привело бы к порче или уничтожению продукции. Например, при исследовании консервированного продукта, выпущенной заводом, невозможно проверить всю партию, так как это просто привело бы к ее уничтожению. Чтобы получить полезное понимание ситуации, вполне достоверную информацию по интересующему нас вопросу в этой

генеральной совокупности, можно отобрать небольшую случайную группу (выборку) и провести обследование. *Случайная выборка* – один из лучших способов извлечения выборки из генеральной совокупности, все объекты генеральной совокупности имеют равные шансы быть отобранным. При проведении статистического исследования, увидеть информацию в статистических данных, расположенных в виде набора чисел – трудно. В подобных случаях числовые данные ранжируют, располагая их в порядке возрастания. Ранжированные числа представляют уже осязаемый определенный результат исследования. По ним легче строить таблицы и гистограммы. Если в ряду статистических данных из чисел одинаковые значения встречаются редко, а числа различных вариантов довольно велико, то ранжирование не позволит выявить характерные черты ряда данных. В таких случаях для обработки данных строят *интервальный ряд*. Для построения такого ряда находят разность между наибольшим и наименьшим значениями ряда (размах). Эту разность делят на выбранную длину ряда и весь промежуток значения ряда данных разбивается на несколько интервалов. На основании данных по интервальному ряду строят таблицу частот и гистограмму.

В ходе описания статистического исследования возникает необходимость введения понятий средних. В сложных ситуациях один из самых эффективных способов увидеть всю картину заключается в обобщении, т.е. использовании одного или нескольких отобранных или рассчитанных значений для характеристики набора данных. Подробное изучение каждого отдельного случая само по себе особой трудности не представляет, но обнаружение и идентификация особенностей, которые в целом характерны для рассматриваемых случаев очень важно и вся информация при этом рассматривается в целом. Свести набор данных к одному числу, которое выражает наиболее фундаментальные свойства данных входит и в цель статистического исследования. Необходимость в таких показателях возникает всегда, когда имеем дело с массой отдельных данных, обладающих индивидуальными различиями, в то же время имеющих существенные общие свойства. Рассматривая средние величины статистических данных, учащиеся будут получать представления о типичных, общих свойствах случайных явлений. Среднее, медиана и мода – это различные способы выбора единственного числа, которые лучше всего описывают все числа в наборе данных. Такой представленный одним числом показатель называется типическим значением или мерой центральной тенденции. *Среднее* (среднее арифметическое) есть условное расчетное значение наблюдаемого признака, при вычислении которого общий объем признака в совокупности (т.е. сумма всех наблюдавшихся значений) сохраняется неизменным. Она не есть характеристика какого-то одного, отдельного наблюдения; она характеризует ряд в целом, сумму всех наблюдавшихся зна-

чений. Учащимся следует понять, что среднее (среднее арифметическое) можно интерпретировать как равномерное распределение суммы всех значений между единицами всей совокупности. Поэтому оно наиболее полезно в качестве обобщающего показателя при отсутствии экстремальных значений (выбросов), когда набор данных представляет собой более или менее однородную группу. Например, средняя заработная плата рабочего завода – это такая сумма денег, которая приходилось бы на каждого, если бы весь фонд оплаты труда был распределен между работниками поровну. Если же заработная плата одного рабочего будет сильно отличаться от второго, то среднее вряд ли можно использовать в качестве обобщающего показателя. Т.е. среднее можно трактовать как центр рассеивания значений наблюдаемого признака, значение около которого колеблются все наблюдаемые значения. Медиану рассматривают как значение, которое расположено посередине; половина элементов в наборе данных больше этого значения, а вторая половина меньше. Таким образом, медиана располагается в центре и дает представление о списке статистических данных. Чтобы найти медиану, данные располагают в порядке возрастания, а затем определяют среднее значение. Если среднее реально учитывает значения экстремальных наблюдений (выбросов сильно отличающихся от линии статистического ряда), то для медианы важно лишь по какую сторону от нее лежит то или иное значение. Третьей, наиболее простой и доступной учащимся средних характеристик является мода. мода представляет собой наиболее распространенное значение, которая чаще всего встречается в наборе данных. Моду найти легко. Необходимо просто выбрать категорию с самым большим количеством. Если на первое место претендуют две или больше категории, то необходимо указать их всех под общим названием «мода», для этого набора статистических данных. Все эти понятия демонстрируются на доступных учащимся примерах и раскрываются их смысл по мере возникновения потребности их применения в процессе анализа ситуации. Все три показателя по-разному характеризуют данные, полученные в результате наблюдений. Поэтому на практике при анализе данных в зависимости от конкретной ситуации используют либо все три показателя, либо некоторые из них. В каждой из ситуации должно возникать вопрос, какой из трех показателей (среднее, медиану или моду) следует использовать в данном случае? Это зависит от того, что можно вычислить, и какой из показателей более полезен. И в каком случае вообще имеет смысл говорить о центральной тенденции. При введении данных понятий важно логическое подведение к нахождению средних величин, анализ, интерпретация результатов решения задач, а не просто решение конкретной задачи, дано ...и найти...?

Характеристика средних величин не показывает, как располагается около нее другие значения, сосре-

доточены ли они вблизи средней или значительно отклоняются от нее. Формирование вероятностно-статистических представлений не должно ограничиваться изучением только средних. В процессе проведения учебных статистических исследований учащимся важно дать понять о том, что статистические данные изменчивы. Если бы данные не изменялись, то ответы на многие вопросы были бы очевидными. Ситуация, в которой присутствует изменчивость, часто содержит риск, поскольку даже использование всей доступной информации не позволяет точно предугадать, что же произойдет, например, в будущем. Поэтому необходимо понимать природу данных, уметь измерять изменчивость. Изменчивость можно определить как степень различий между отдельными значениями. Существует разные способы описания степени изменчивости набора данных, причем каждый из них требует соответствующих числовых значений. Более полное представление об изменчивости (вариации) изучаемого явления учащиеся могут получить при рассмотрении среднего квадратичного отклонения на этапе внутримодельного обучения элементам теории вероятностей и математической статистики. Но на содержательных примерах учащимся следует показать изменчивость (вариацию) набора данных.

Стандартное отклонение – это число, описывающее, насколько значения данных обычно отличаются от среднего. Это понятие является очень важным, поскольку оно представляет собой основной инструмент определения степени случайности в изучаемых ситуациях. Этот показатель является мерой случайности отклонения отдельных значений от их среднего.

Наиболее простым показателем изменчивости является размах, которая определяется как разность между самым большим и самым малым значением данных. Размах определяет предел изменения данных в наборе. Однако размах дает несколько поверхностное представление об изменчивости данных, не отражает типичную изменчивость, так как учитывает только два крайних значения, а разброс значений заключается между ними, поэтому имеет и ограниченное применение.

Дальнейшее изучение данной линии на этом этапе – это органичное введение понятия относительной частоты. Мотивация в изучении понятия возникает естественным образом, в процессе проведения учебных статистических исследований. Следует оценить насколько часто, наблюдается интересующее нас событие по результатам проведенного исследования. Каждый раз при проведении статистических исследований мы наблюдаем, наступает определенное нас интересующее событие или нет. Каждое событие соответствует некоторому представляющему интерес свойству. Для того, чтобы найти одним числом это свойство, находим отношение числа экспериментов, в которых это событие произошло, к общему числу экспериментов. Это от-

ношение называется относительной частотой. В ходе статистических исследований можно установить, что при многократном повторении определенных опытов или наблюдений относительная частота появления ожидаемого события может оставаться примерно одинаковой, незначительно отличаясь от некоторого числа. Для описания этого свойства следует строить таблицы относительных частот. Они позволят наблюдать устойчивость относительных частот. В дальнейшем изучении данной линии учащиеся подобно таблице относительных частот будут строить ряд распределения случайных величин. Поэтому таблицу относительных частот можно назвать эмпирическим прообразом ряда распределения случайных величин. Стабилизации относительных частот около одного числа можно наглядно изобразить с помощью линейных графиков. При статистических исследованиях и рассмотрении относительных частот следует обратить внимание учащихся на то, что в процессе проведения экспериментов, интересующие нас случайные события, происходят с разной относительной частотой: одни чаще, другие реже. Те события, которые происходят чаще, имеют большую возможность появления, а те которые реже – меньшую, т.е. каждое случайное событие имеет свою меру возможности появления - вероятность. Нужно знать насколько правдоподобно конкретное событие. Это правдоподобность, мера может быть выражена числом, которая является значением вероятности события. Здесь следует содержательно описать понятие вероятности, обозначить, откуда берутся числа, характеризующие вероятность, показать, как вероятность указывает приблизительное количество появления некоторого события при многократном повторении случайного эксперимента, оценить вероятность случайного события по частоте, осуществить статистический подход к понятию вероятности. Каждому событию соответствует число в интервале от 0 до 1, которые являются значением вероятности события. Вероятность показывает насколько правдоподобно наступление данного события при каждом выполнении случайного эксперимента. Вероятность, равная 0, означает, что соответствующее событие, по сути, никогда не наблюдается, а вероятность, составляющая 1, свидетельствует о том, что данное событие происходит практически всегда. В целом вероятность показывает приблизительную долю случаев, в которых ожидается наступления события. Относительные частоты могут использоваться для оценки (наилучшая приблизительная оценка) значения вероятности в случае наличия информации основанной на предшествующем опыте. То положение, что относительная частота появления некоторого события можно использовать в качестве «наилучшей приблизительной оценки» вероятности, на основе имеющихся данных, будет знакомством на содержательном уровне с законом больших чисел. Если в серии экспериментов со случайными исходами относительная частота принимает достаточно устойчивое

значение, т.е. наблюдаемые значения относительной частоты группируются около некоторого постоянного числа, то это число и есть вероятность случайного события. Свойство «устойчивости частот», многократно проверенное экспериментально и подтверждающееся всем опытом практической деятельности человечества, есть одна из наиболее характерных закономерностей, наблюдаемых в случайных явлениях. Знакомство школьников с простейшей формой закона больших чисел будет стержнем в обучении содержательной линии на данном этапе. Учащиеся должны осознать, что связь между относительной частотой и его вероятностью – глубокая, органическая связь. И, что когда мы оцениваем степень возможности какого – либо события, мы неизбежно связываем эту оценку с большей или меньшей частотой появления аналогичных событий на практике. Ха-

рактизуя вероятность события каким–то числом, мы не можем придать этому числу иного реального значения и иного практического смысла, чем относительная частота появления данного события при большом числе опытов. Численная оценка степени возможности события посредством вероятности имеет практический смысл именно потому, что более вероятные события происходят в среднем чаще, чем менее вероятные. И если практика определения указывает на то, что при увеличении числа опытов частота события имеет тенденцию выравниваться, приближаясь сквозь ряд случайных отклонений, к некоторому постоянному числу, естественно предположить, что это число и есть вероятность события. Учащиеся должны знать, что многочисленные опыты, производившиеся, со времен возникновения теории вероятностей, подтверждает это предположение.