

## Тема 1. ВЫБОРОЧНЫЙ МЕТОД ОЦЕНКИ СОВОКУПНОСТИ. РАСПРЕДЕЛЕНИЕ ЧАСТОТ И ЕГО ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ

**Объяснение.** Всякое массовое явление, например, группа сосен в лесу, стая перелетных птиц, представляет собой совокупность объектов, т. е. некоторых условных единиц, каждая из которых в отдельности строго индивидуальна и отличается от других рядом признаков – высотой, массой, содержанием каких-либо веществ и т. д. Каждый из признаков может иметь у различных особей разную степень выраженности, поэтому говорят, что признак варьирует. Свойство условных единиц – растений, животных, проб воды из разных водоемов, проб почвы или воздуха в разных местах – отличаться друг от друга даже в однородных совокупностях называется *изменчивостью или варьированием*.

Изменчивость – это свойство, присущее всем объектам и явлениям природы. Варьирование возникает вследствие того, что даже родственные живые организмы могут отличаться своей наследственностью, кроме того, формирование их часто протекает в относительно различных условиях внешней среды.

Изменчивость признаков создает известные трудности в случаях, когда требуется дать общую характеристику определенной варьирующей группе (совокупности) растений, животных, почв и т. п. по отдельным признакам или сравнить две такие группы и найти различия между ними. Практически никогда не представляется возможным исследовать по тому или иному признаку всю совокупность. В таком случае прибегают к изучению части совокупности, по которой делают общее заключение. Такой метод называется выборочным и считается основным при статистическом изучении совокупности.

Всю группу объектов, подлежащую изучению, называют совокупностью или *генеральной совокупностью*, а ту часть объектов, которая попала на проверку, исследование, – *выборочной совокупностью* или выборкой. Число элементов в генеральной совокупности и выборке называют их *объемом*.

Главная цель выборочного метода – по статистическим показателям выборочной совокупности как можно точнее охарактеризовать всю генеральную совокупность объектов.

В результате наблюдений получают сведения о численной величине изучаемого признака у каждого члена данной

выборочной совокупности. Возможные значения варьирующего признака  $X$  называют *вариантами* и обозначают  $X_1, X_2, \dots, X_n$ . Полученный таким образом ряд варьирующих величин можно упорядочить – расположить значения признака (варианты) в порядке их возрастания или убывания. Такое упорядочение ряда называется *ранжированием*. После ранжирования можно заметить, что каждое значение признака встречается неодинаковое число раз – одни редко, другие часто. Числа, которые характеризуют, сколько раз повторяется каждое значение признака у членов данной совокупности, называют *частотами признака* и обозначают  $f$ . Сумма всех частот ( $\sum f$ ) равна объему выборки, т. е. числу членов ряда  $n$ .

В результате такой обработки первичных наблюдений получают так называемый вариационный ряд. *Вариационным рядом* называют такой ряд данных, в котором указаны возможные значения варьирующего признака в порядке возрастания или убывания и соответствующие им частоты.

Результаты экспериментов обычно представляют собой обширный цифровой материал, восприятие которого облегчается при систематизации и представлении в виде таблиц и графиков. При небольшой вариации признака, принимающего целочисленные значения, данные достаточно сгруппировать по значениям вариант с последующим построением вариационной кривой. При значительной вариации признака прибегают к разнесению данных по классам, которые охватывают сразу несколько значений вариант.

**Пример.** Допустим, что в результате подсчета количества жуков долгоносиков в 100 сосудах с пищевым субстратом были получены следующие данные (в штуках):

90	109	99	100	115	68	70	72	73	70
76	82	80	68	69	74	72	69	80	79
79	84	84	108	83	84	99	98	102	101
45	59	60	63	78	87	94	91	88	90
72	68	80	81	84	77	79	81	84	76
70	67	100	103	69	72	74	66	67	72
79	78	83	92	93	81	82	86	89	93
77	76	88	89	94	82	80	81	77	80
92	91	76	79	73	84	79	84	79	84
89	85	93	90	79	83	91	87	89	94

Для того, чтобы проанализировать число долгоносиков в сосудах с пищевым субстратом необходимо сгруппировать значения  $X_1, X_2, \dots, X_n$  в  $k$  групп с интервалом  $i$ . Ориентировочно

число групп равно корню квадратному из объема выборки, которое, однако, не должно быть меньше 5 и больше 20.

Величину интервала групп определяют по соотношению:

$$i = \frac{X_{max} - X_{min}}{\text{число групп}} = \frac{R}{k}$$

В данном примере целесообразно взять 7 групп, тогда величина интервала будет равна целому числу:

$$i = \frac{R}{k} = \frac{115 - 45}{7} = \frac{70}{7} = 10 \text{ шт.}$$

При выборе границ групп следует обращать внимание на то, чтобы верхняя граница группы была меньше, чем нижняя граница прилегающей соседней группы на единицу измерения, в нашем примере на 1 особь (штуку). Группировка осуществляется в такой последовательности:

1. Определяем размах варьирования результатов измерений, т. е. разность между наибольшим и наименьшим значением ряда измерений:

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

2. Устанавливаем число групп  $k$  и размер интервала группировки  $i$ .

3. Готовим макет таблицы сгруппированного распределения частот результатов измерений (табл. 1). В первой колонке записываем интервал группировки, во второй – число измерений, входящих в данный интервал (частота), в третьей – среднее значение в группе (для этого сумму минимального и максимального значений в группе делят на два).

**Таблица 1 – Сгруппированное распределение частот по данным подсчета числа долгоносиков в 100 сосудах с пищевым субстратом**

Группы (интервал группировки)	Частота	Среднее значение в группе
45–54	1	50
55–64	3	60
65–74	21	70
75–84	40	80
85–94	23	90
95–104	9	100
105–115	3	110

Сумма частот должна быть равна объему совокупности  $\Sigma f = n = 100$ .

Для получения графического изображения данных строят гистограмму и вариационную кривую. Это позволит сразу охватить

важнейшие закономерности распределения наблюдений. Для этого на горизонтальной линии (оси абсцисс) наносят значения интервала группировки, а по вертикали (ось ординат) – численности этих значений (их частоты). Масштаб выбирают такой, чтобы весь график имел удобную и легко обозримую форму. Ступенчатый график в виде столбиков, имеющих высоту, пропорциональную частотам, а ширину, равную интервалам группировки, называется *гистограммой*, из которой легко получить *кривую распределения или вариационную кривую*, соединив линией средние значения групп. В конце работы делают вывод о наиболее часто встречающейся численности долгоносиков в сосудах. Такое значение в статистике называется модой (либо, соответствующий класс – модальным). Мода – это значение, наиболее часто встречающееся в выборке.

**Задание.** После рассмотрения примера необходимо сгруппировать результаты эксперимента, представить графическое изображение (приложения 4, 5, 6).