

## Генеральная совокупность и выборка

**Бурицев Владимир Олегович**

*студент ФГОУ ВО*

*Курский государственный университет,  
колледж коммерции, технологий и сервиса,*

*Россия, г. Курск*

**Ефимцева Ирина Борисовна,**

*Научный руководитель, канд. тех. наук.*

*преподаватель ФГОУ ВО*

*Курский государственный университет,  
колледж коммерции, технологий и сервиса,*

*Россия, г. Курск*

Основной задачей математической статистики является исследование некоторой совокупности объектов или некоторой случайной величины (нескольких случайных величин).

Выборочной совокупностью или просто выборкой называется совокупность случайно отобранных объектов. Генеральной совокупностью называется совокупность объектов, из которых производится выборка. Повторной называется выборка, при которой отобранный объект возвращается в генеральную совокупность. Если же отобранный объект не возвращается, то выборка называется бесповторной. Например, пусть в лабораторию на исследование направляются из цеха детали. Если после экспертизы деталь отправляется обратно в цех и не исключена возможность, что её снова случайно отправят на экспертизу, то в этой ситуации мы имеем дело с повторной выборкой. Если же из цеха сразу отправляют полное количество испытуемых деталей и отправляют их сразу все обратно после проверки, то здесь уже речь идёт о бесповторной выборке.

Выборка называется репрезентативной (представительной), если каждый объект генеральной совокупности имеет одинаковую вероятность попасть в выборку. Только в случае репрезентативной выборки можно получить объективную информацию о генеральной совокупности. Для

достижения репрезентативности применяют, в зависимости от условий, различные способы отбора: простой случайный бесповторный, простой случайный повторный, типический, механический, серийный. Ограничимся рассмотрением лишь первого, наиболее простого.

Простым случайным называется такой отбор, при котором объекты извлекают по одному из всей генеральной совокупности. Если выборка очень большая, то для обеспечения случайности используют готовые таблицы случайных чисел.

Часто под генеральной совокупностью понимают и исследуемую случайную величину. Для исследования случайной величины при постоянных условиях выполняются испытания. Совокупность полученных значений также называется выборкой и обрабатывается статистически. Методы статистической обработки аналогичны в обоих случаях. Если объём генеральной совокупности большой или неизвестен, то её можно также представлять как некую случайную величину, извлечение объекта из генеральной совокупности представлять как испытание этой случайной величины.

Целью статистических наблюдений (испытаний) является выяснение вероятностных свойств совокупности: распределения, числовых характеристик и т.д. При исследовании объектов можно измерять значение одного или нескольких признаков. Соответственно говорят об одномерной, двумерной, трёхмерной и т.д. выборках. Вначале рассмотрим обработку одномерных выборок. Выбор объекта из генеральной совокупности и измерение значения признака называется статистическим наблюдением. Результаты наблюдений фиксируют в протоколе или дневнике наблюдений в порядке их появления. Составленные таким образом по результатам наблюдений таблицы называются неупорядоченными или не сгруппированными выборками. Рассмотрим пример.

Пример 1. По результатам  $N = 75$  измерений диаметра ( в мм ) шейки плунжера после шлифовки составлена табл. 1.1. Выборка будет намного наглядней, если все её элементы упорядочить по возрастанию или по убыванию. Но в выборке одно значение (вариант) может встречаться несколько

раз, и поэтому целесообразно результаты записать в виде таблицы, в первом столбце которой находятся всевозможные значения (варианты)  $x_i$  генеральной совокупности (или случайной величины)  $X$ , а во втором – числа  $n_i$ , т.е. частоты появления  $i$ -го значения. Таковую таблицу называют вариационной таблицей или дискретным вариационным рядом. Выборку, записанную в виде вариационного ряда, называют сгруппированной и упорядоченной. Для составления вариационного ряда нужно найти в таблице наблюдений наименьшее и наибольшее значения, упорядочить различные варианты по возрастанию и для каждой из них сосчитать по таблице, сколько раз она встречается. Для контроля нужно проверить, чтобы объём не сгруппированной выборки был равен сумме частот вариантов. Для больших выборок эта задача поручается специальным программам на ЭВМ.

Пример 2. Построить вариационный ряд для выборки, взятой из табл. 1.1. Пусть  $X$  – диаметр шейки плунжера. Вариационный ряд представлен табл. 1.2. Проверяем контрольную сумму (объём выборки)  $N = 75$ .

$$1 + 1 + 7 + 5 + 3 + 4 + 14 + 8 + 6 + 12 + 5 + 1 + 4 + 2 + 1 + 1 = 75.$$

Таблица 1.1

Диаметр шейки плунжера

№ набл.	Диаметр	№ набл.	Диаметр	№ набл.	Диаметр
1	6.75	26	6.77	51	6.77
2	6.77	27	6.75	52	6.83
3	6.77	28	6.74	53	6.76
4	6.73	29	6.71	54	6.76
5	6.76	30	6.70	55	6.82
6	6.74	31	6.78	56	6.77
7	6.70	32	6.76	57	6.71
8	6.75	33	6.81	58	6.74
9	6.71	34	6.69	59	6.77
10	6.72	35	6.80	60	6.75
11	6.77	36	6.80	61	6.74
12	6.79	37	6.77	62	6.75
13	6.71	38	6.68	63	6.77
14	6.78	39	6.74	64	6.72
15	6.73	40	6.70	65	6.74
16	6.70	41	6.70	66	6.80
17	6.73	42	6.74	67	6.75

18	6.80	43	6.77	68	6.75
19	6.72	44	6.74	69	6.74
20	6.78	45	6.74	70	6.76
21	6.70	46	6.77	71	6.70
22	6.75	47	6.73	72	6.81
23	6.78	48	6.74	73	6.74
24	6.78	49	6.77	74	6.74
25	6.76	50	6.74	75	6.71

Пример 3. Построить полигон частот для вариационного ряда из примера 2. График полигона представлен на рис. 1.1.

Таблица 1.2

Вариационный ряд

№ i	Значение X	Частота n <sub>i</sub>
1	6.68	1
2	6.69	1
3	6.70	7
4	6.71	5
5	6.72	3
6	6.73	4
7	6.74	14
8	6.75	8
9	6.76	6
10	6.77	12
11	6.78	5
12	6.79	1
13	6.80	4
14	6.81	2
15	6.82	1
16	6.83	1

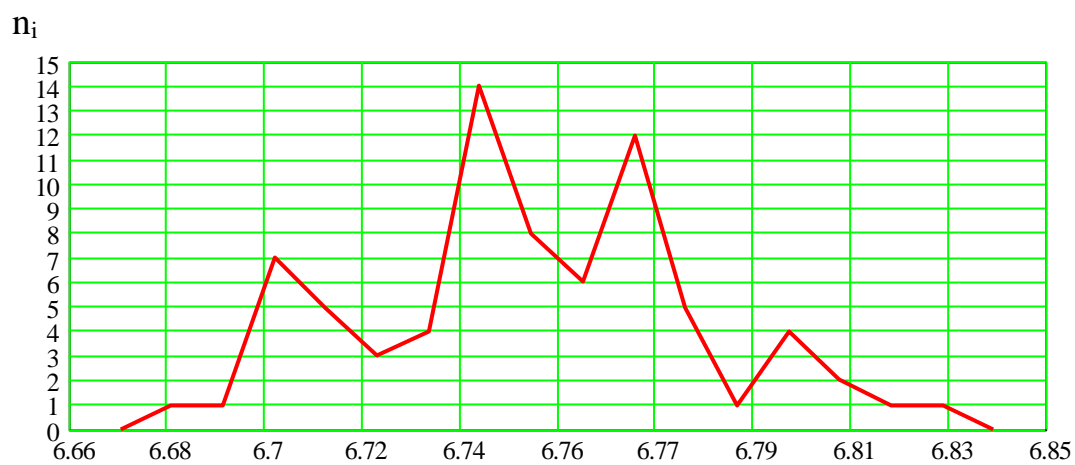


Рис. 1.1. Полигон частот

Если количество различных вариантов слишком велико или близко к объёму выборки, то целесообразно составить вариационный ряд по интервалам (интервальный ряд) значений генеральной совокупности.

Вариационный ряд по интервалам значений составляется так же, как и дискретный вариационный ряд, только в таблице вместо столбца значений записывается другой столбец, в который записывают интервалы значений генеральной совокупности. Обычно (хотя и не обязательно) все интервалы выбирают одинаковой длины так, что начало интервала входит в интервал, а его конец – не входит. При этом минимальное значение  $x_{\min}$  входит в первый интервал, максимальное  $x_{\max}$  – в последний. Число интервалов выбирается обычно так, чтобы в каждый из интервалов попало не менее 5 - 8 значений. Для более ответственных исследований используется формула Стёрджеса.

### Список литературы

1. Васильев, А. А. Теория вероятностей и математическая статистика: учебник и практикум для СПО / А. А. Васильев. — 2-е изд., испр. и доп. — М.: Издательство Юрайт, 2017. — 253 с. — (Серия : Профессиональное образование). — ISBN 978-5-534-05176-6. — Режим доступа: [www.biblio-online.ru/book/61129D36-34CF-4B87-901E-CF4C3D4B056A](http://www.biblio-online.ru/book/61129D36-34CF-4B87-901E-CF4C3D4B056A);
2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для СПО / В. Е. Гмурман. — 11-е изд., перераб. и доп. — М. : Издательство Юрайт, 2017. — 404 с. — (Серия : Профессиональное образование). — ISBN 978-5-534-00935-4. — Режим доступа : [www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A](http://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A);
3. Гмурман, В. Е. Теория вероятностей и математическая статистика: учебник для СПО / В. Е. Гмурман. — 12-е изд. — М. : Издательство Юрайт, 2017. — 479 с. — (Серия : Профессиональное образование). — ISBN 978-5-534-00859-3. — Режим доступа : [www.biblio-online.ru/book/535E35F5-83AD-48A3-833E-DE002FC2268A](http://www.biblio-online.ru/book/535E35F5-83AD-48A3-833E-DE002FC2268A);
4. Теория вероятностей и математическая статистика [Электронный ресурс]: учебное пособие/ Шилова З.В., Шилов О.И.— Электрон. текстовые данные.— Саратов: Ай Пи Ар Букс, 2015.— 158 с.— Режим доступа: <http://www.iprbookshop.ru/33863>.— ЭБС «IPRbooks», по паролю