**Практическая работа № 7**

 **«Построение графической диаграммы выборки, расчёт характеристик выборки»**

**Учебная цель:** научиться строить графические диаграммы выборки, рассчитывать характеристики выборки

**Краткие теоретические и учебно-методические материалы по теме практической работы**

В самых различных областях производственной и научной деятельности приходится проводить изучение (обследование, измерение, проверку) объектов, принадлежащих некоторой совокупности, по какому-либо признаку. При этом иногда приходится исследовать каждый объект совокупности, т. е. проводить сплошное исследование. Однако на практике гораздо чаще применяется выборочное исследование. При выборочном исследовании из всей совокупности отбирают некоторым образом определенное число объектов и только их подвергают исследованию. При этом совокупность всех исследуемых объектов называют *генеральной совокупностью.*

 *Выборкой* называют совокупность случайно отобранных объектов из генеральной совокупности. Под *случайным отбором* при образовании выборки понимают такой отбор, при котором все объекты генеральной совокупности имеют одинаковую вероятность попасть в выборку.

 Выборку можно проводить двумя основными способами. При первом способе объект извлекают из генеральной совокупности, исследуют и возвращают в исходную генеральную совокупность; затем снова извлекают некоторый объект, исследуют и возвращают в генеральную совокупность и т. д. Полученную таким образом выборку называют *повторной.* При втором способе после исследования объекты в генеральную совокупность не возвращают, и выборку в этом случае называют *бесповторной*.

 Число объектов выборочной или генеральной совокупности называют *объемом выборки*. Например, если из 10 000 изделий для контроля отобрано 100 изделий, то объем генеральной совокупности N=10 000, а объем выборки n=100.

 Для того чтобы по выборке можно было с определенной уверенностью судить о всей генеральной совокупности, выборка должна достаточно полно отражать изучаемое свойство объектов генеральной совокупности,т.е быть *репрезентативной*. Для этого необходимо, чтобы отбор объектов в выборку осуществлялся действительно случайно и чтобы изучаемому свойству была присуща статистическая устойчивость .

 Пусть для изучения количественного (дискретного или непрерывного) признака *Х* из генеральной совокупности извлечена выборка

 *x1, х2, x3, …  xn* , (1)

Разность между наибольшим значением числовой выборки и ее наименьшим значением называют *размахом выборки*.

 Наблюдавшиеся значения *хi* признака *Х* называются *вариантами*, а неубывающую последовательность вариант называют *вариационным* *рядом.*

 Пусть при исследовании некоторой генеральной совокупности получена числовая выборка объема *n*, причем значение *х1* встретилось в выборке *n1* раз, значение *х2 - n2* раз, ..., значение *хk — nk* раз. Числа *n1 , n2* , …, *nk* называют *частотами*, а их отношения к объему выборки, т. е. отношения  - *относительными частотами* соответствующих значений *x1, х2, x3, … хk*  выборки. Очевидно, что сумма частот равна объему выборки, а сумма относительных частот равна единице, т. е.

 *n1 + n2* + …+ *nk* = *n ,*  (2)

 Последовательность пар (*х1 ; n1 ); (х2 ; n2 ); ( х3 ; n3 ); … (хk ; nk )*

называют *статистическим рядом.*Обычно статистический ряд записывают в виде таблицы:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  *х1* | *х2* | *х3* | … | *xi* | … | *хk* |
|  *n1* | *n2* | *n3* | … | *ni* | … | *nk* |

 (3)

 Следующей таблицей задается так называемое *выборочное распределение*, в которой указываются все значения выборки и их соответствующие относительные частоты:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *х1* | *х2* | *х3* | … | *xi* | … | *хk* |
|  |  |  | … |  | … |  |

 (4)

**Графические изображения выборки. Полигон и гистограмма**

 Для наглядного представления о выборке часто используют различные графические изображения выборки. Простейшими изображениями выборки являются полигон и гистограмма. Пусть выборка задана вариационным рядом: (*х1 ; n1 ); (х2 ; n2 ); ( х3 ; n3 ); … (хk ; nk ) .* *Полигоном частот* называют ломаную с вершинами в указанных точках.

 *Полигоном относительных частот* называют ломаную с вершинами в точках

 (*х1 ;*  *); (х2 ;*  *); ( х3 ;*  *); … (хk ;*  *)*

 Ясно, что полигон относительных частот получается из полигона частот сжатием вдоль оси ординат в *n*  раз, где *n* — объем выборки.

При большом объеме выборки более наглядное представление о ней дает *гистограмма*. Чтобы построить гистограмму частот, промежуток от наименьшего значения выборки до наибольшего ее значения разбивают на несколько частичных промежутков длины *h*. Для каждого частичного промежутка вычисляют сумму *si*  частот значений выборки, попавших в этот промежуток. Значение *xi* выборки, совпавшее с правым концом промежутка, относят к следующему промежутку (если *xi* — не наибольшее значение выборки). Затем на каждом частичном промежутке, как на основании, строят прямоугольник с высотой . Объединение всех построенных таким образом прямоугольников называют *гистограммой частот*. Итак, гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых являются частичные промежутки длины *h*, а высотами — отрезки длины  , где *si* — сумма частот значений выборки, попавших

в *i*-й промежуток.

 Из определения гистограммы ясно, что ее площадь равна объему выборки.

При решении задач в зависимости от объема выборки в большинстве случаев целесообразно брать 10-20 частичных промежутков.

 Аналогично определяют и строят гистограмму относительных частот.

*Гистограммой относительных частот* называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых являются частичные промежутки длины *h*, а высотами — отрезки длины , где *wi -*  суммы относительных частот значении выборки, попавших в *i*-й промежуток. Площадь гистограммы относительных частот, очевидно, равна единице.

 Пусть имеется некоторая выборка объема *n: x1, х2, x3, …  xn* . *Выброчной средней* называется среднее арифметическое значений выборки:

  (5)

 Если выборка задана статистическим рядом (3) или выборочным распределением (4), то формулу (5) естественно записать в следующем виде:

  (6)

*Выборочной дисперсией* называется среднее арифметическое квадратов отклонений значений выборки от выборочной средней.

 (7)

Если выборка задана статистическим рядом (3) или выборочным распределением (4), то формулу (7) можно записать так:

 (8)

Формулы (7) и (8) можно преобразовать к более удобному для вычислений виду:

 (9)

т. е. выборочная дисперсия равна среднему квадратов значений выборки без квадрата выборочной средней.

 *Исправленной выборочной дисперсией* называется

  (10)

где *S0* — выборочная дисперсия, *п* — объем выборки. Отсюда, используя формулу (7),

  (11)

**Примеры по выполнению практической работы**

 **Пример 1.** Составить для выборки 1, 10, -2, 1,0, 1, 10, 7, -2, 10, 10, 7

вариационный ряд и найти ее размах.

Решение: записав заданную выборку в виде неубывающей последовательности, получим вариационный ряд

 -2,-2,0, 1, 1, 1,7,7, 10, 10, 10, 10.

Размах данной выборки равен 10 - (-2) =12.

 **Пример 2** Для выборки 3,8,-1,3, 0, 5,3,-1,3, 5 определить объем и размах. Записать выборку в виде вариационного ряда и в виде статистического ряда. Найти выборочное распределение. Построить полигон частот.

Решение: Объем выборки n = 10, ее размах равен 8 - (-1) = 9. Записав значения выборки в виде неубывающей последовательности получим вариационный ряд

 -1,-1,0, 3,3, 3,3, 5, 5, 8.

Статистический ряд можно записать в виде последовательности пар чисел - (-1;2), (0;1), (3;4), 5;2), (8;1) или в виде таблицы

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -1 | 0 | 3 | 5 | 8 |
| 2 | 1 | 4 | 2 | 1 |

 Для контроля находим сумму частот: 2+ 1 +4 + 2+ 1 = 10 и убеждаемся в том, что она равна объему выборки.

 Вычислив относительные частоты, найдем выборочное распределение:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -1 | 0 | 3 | 5 | 8 |
|  |  |  |  |  |

 Для контроля убеждаемся в том, что сумма относительных частот равна единице:

 ++++=1.

Полигон частот для заданной выборки имеет вид:

 

 **Пример 3.**  При измерении напряжения в электросети получена следующая выборка:

 218, 221, 215, 225, 225, 217,

 224, 220, 220, 219, 221, 219,

 222, 227, 218, 220, 223, 230,

 223, 216, 224, 227, 220, 222

(данные выражены в вольтах). Построить гистограмму частот, если число частичных промежутков равно 5.

Решение: наименьшее значение выборки равно 215, наибольшее — 230.

Находим длину частичных промежутков . Подсчитываем с учетом кратности число значений выборки, попавших в каждый промежуток.

 Для первого промежутка [215; 218) это число равно 3, для второго [218; 221) оно равно 8, для третьего [221; 224) — 6, для четвертого [224; 227) — 5, для пятого [227; 230] — 2. Следовательно, высоты прямоугольников (слева направо), образующих гистограмму, равны По полученным данным строим гистограмму

 

Для контроля убеждаемся в том, что площадь гистограммы равна объему выборки:

 

**Пример 4.** На основании данных о средней заработной плате работников в области в тыс. руб., которые помещены в интервальный вариационный ряд в таблицу, построить гистограмму распределения частот зарплаты работников:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Заработнаяплата | 1-3 | 3-5 | 5-7 | 7-9 | 9-11 | 11-13 |
| ЧислоРаботников | 12 | 23 | 37 | 19 | 15 | 9 |

Решение: при построении гистог­раммы по оси абсцисс откладываются значения изучаемого признака (границы интервалов), а по оси у – соответствую­щие частоты, в том случае, если интервалы одинаковой величины. Используя мастер диаграмм в MS Excel, получим гистограмму:

 Гистограмма распределения частот зарплаты работников

**Пример 3**. Для выборки 4,5,3,2, 1,2,0,7,7,3 найти выборочную среднюю , выборочную дисперсию *S0*, исправленную выборочную дисперсию *S*.

Решение: объем выборки *п* = 10. По формуле (5) находим выборочную среднюю:

 

Чтобы найти выборочную дисперсию, воспользуемся формулой (9). Для этого вычислим среднее квадратов значений выборки:

 

Теперь по формуле (9) находим *S0* = 16,6 -3,42= 5,04. Наконец, используя формулу (10), вычисляем исправленную выборочную дисперсию:

 

**Пример 4.** Для выборки 3,8-1,3,0,5,3,4,3,5 найти выборочную среднюю , выборочную дисперсию *S0*, исправленную выборочную дисперсию *S*.

Решение: статистический ряд для для данной выборки имеет вид

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -1 | 0 | 3 | 5 | 8 |
| 2 | 1 | 4 | 2 | 1 |

Объем выборки *п*=10. Выборочную среднюю найдем по формуле (6):



Вычислим среднее квадратов значений выборки:



Согласно формуле (9) находим выборочную дисперсию:

*S0* = 15,2-2,82= 7,36.

Для вычисления исправленной выборочной дисперсии воспользуемся формулой (10):

 

**Задания для практического занятия:**

 **Вариант 1**

 1.Для выборки 1,1,2,-5,4,3,3,8,8,1 определите объем и размах. Запишите выборку в виде вариационного ряда и в виде статистического ряда. Найдите выборочное распределение.

 2. Для выборки, заданной статистическим рядом

|  |  |  |  |
| --- | --- | --- | --- |
| 2 | 4 | 6 | 8 |
| 5 | 2 | 1 | 3 |

постройте 1) полигон частот; 2) полигон относительных частот;

 3. Для выборки, заданной вариационным рядом -5, -5, 2, 3, 5,10,15, 15, 20, 20, найдите выборочную среднюю ; выборочную дисперсию *S0*, несмещенную выборочную дисперсию *S*.

4. Произведено выборочное обследование коммерческих фирм по затратам на ре

кламу, результаты которого представлены в таблице:

|  |  |
| --- | --- |
| Затраты на рекламу (усл. ден. ед.) | Кол-во фирм |
| 40-60 | 4 |
| 60-80 | 3 |
| 80-100 | 5 |
| 100-120 | 6 |
| 120-140 | 2 |

По данным выборочного обследования постройте гистограмму частот, используя мастер диаграмм в MS Excel.

  **Вариант 2**

 1.Для выборки -3,1,2,4,3,4,4,1,2,1 определите объем и размах. Запишите выборку в виде вариационного ряда и в виде статистического ряда. Найдите выборочное распределение.

 2. Для выборки, заданной статистическим рядом

|  |  |  |  |
| --- | --- | --- | --- |
| -1 | 1 | 3 | 7 |
| 1 | 3 | 4 | 2 |

постройте 1) полигон частот; 2) полигон относительных частот

 3. Для выборки, заданной статистическим рядом

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| -1 | 1 | 3 | 5 | 7 | 9 |
| 2 | 2 | 1 | 3 | 1 | 1 |

найдите выборочную среднюю ; выборочную дисперсию *S0*, несмещенную

выборочную дисперсию *S*;

4. В результате выборочного обследования коммерческих банков о размере прибыли за год получено следующее распределение:

|  |  |
| --- | --- |
| Размер прибыли (млн руб.) | Число банков |
| 10- 2020-3030-4040-50 50-60 | 510201510 |

По данным выборочного обследования постройте гистограмму частот, используя мастер диаграмм в MS Excel.