

## МОДУЛЬ 1. ТОЧЕЧНОЕ И ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ

### Практическое занятие №1.

#### ФОРМЫ ПРЕДСТАВЛЕНИЯ СТАТИСТИЧЕСКИХ ДАННЫХ

**Цель занятия.** Знать выборочную совокупность и вариационный ряд. Научиться строить полигон, гистограмму частот и кумуляту в MS Excel.

**Учебные вопросы:**

1. Дискретный и интервальный статистический ряд
2. Полигон и эмпирическая функция распределения (кумулята).
3. Гистограмма.

**Задание 1**

Таблица 1.1

25	6	7	59	1
18	37	2	28	43
52	17	1	9	29
2	35	5	21	14
1	55	65	1	6
19	6	9	5	1
18	4	24	25	15
4	19	17	3	24
28	11	10	20	15
19	4	7	9	24

Произведено 50 опытов по определению времени обслуживания (в мин) клиентов в Центральном банке города Энска (табл. 1.1). Используя MS Excel, рассчитать относительные и накопленные частоты, построить гистограмму абсолютных частот, полигон и кумуляту времени обслуживания.

*Решение.*

1. В объединенные ячейки A1:E1 (табл. 1.2) вводим слово «Наблюдения», а в диапазон ячеек A2:E11 – время обслуживания.

2. По формуле Стерджесса в ячейке B13 (табл. 1.3) определяем число интервалов  $k$ . Для этого вводим формулу  $=1+3,32*LOG10(СЧЁТ(A2:E11))$ . Получаем, что  $k = 6,6$ . В

ячейке D13 округляем это значение до 7 ( $=ОКРВВЕРХ.ТОЧН(B13)$ ). В ячейках B14 и B15 определяем  $X_{\max} = 65$  ( $=МАКС(A2:E11)$ ),  $X_{\min} = 1$  ( $=МИН(A2:E11)$ ). В ячейке B16 рассчитываем размах выборки  $R$  ( $=B14-B15$ ), а в ячейке B17 – длину интервала  $h$  ( $=B17/C13$ ). В ячейках B19:B25 производим расчет правых границ интервалов (карманов) по формулам:

$$\begin{aligned} &=B15+\$B\$17; =B19+\$B\$17; =B20+\$B\$17; =B21+\$B\$17; =B22+\$B\$17; \\ &=B23+\$B\$17; =B24+\$B\$17. \end{aligned}$$

Таблица 1.2

	A	B	C	D	E
1	Наблюдения				
2	25	6	7	59	1
3	18	37	2	28	43
4	52	17	1	9	29
5	2	35	5	21	14
6	1	55	65	1	6
7	19	6	9	5	1

8	18	4	24	25	15
9	4	19	17	3	24
10	28	11	10	20	15
11	19	4	7	9	24

Так как длина интервала не меняется, то в расчетах используем абсолютный адрес ячейки  $\$B\$17$  (абсолютный адрес ячейки ставится с помощью функциональной клавиши F4 на клавиатуре). Используем автозаполнение ячеек B20:B25. В расчетах за левую границу первого интервала можно брать  $X_{\min}$  или  $X_{\min} - h/2$ .

3. Расчеты абсолютных частот, относительных частот (частностей) и накопленных относительных частот (накопленных частностей) оформим в виде табл. 1.4 (диапазон ячеек – G1:L10). В строке G1:L1 расположим заголовки соответствующих столбцов. В ячейки G3:G9 вводим карманы. Для расчета абсолютных частот используем функцию ЧАСТОТА, которая задается в виде формулы массива и выдает на одно значение больше рассчитанного числа интервалов. Поэтому перед вызовом функции необходимо выделить количество ячеек, достаточное для размещения результата (в данном случае необходимо выделить 8 ячеек).

Таблица 1.3

	А	В	С
13	$k$	6,6	7
14	$X_{\max}$	65	
15	$X_{\min}$	1	
16	$R$	64	
17	$h$	9,1	
18	Карманы		
19	$K1$	10,1	
20	$K2$	19,3	
21	$K3$	28,4	
22	$K4$	37,6	
23	$K5$	46,7	
24	$K6$	55,9	
25	$K7$	65,0	

Выделяем диапазон ячеек Excel H3:H10, вызываем функцию ЧАСТОТА. Эта функция Excel имеет синтаксис: =ЧАСТОТА(массив данных; массив интервалов). Для рассматриваемого примера будет:

=ЧАСТОТА(A2:E11;B19:B25).

Нажать комбинацию клавиш Ctrl + Shift + Enter (если вывелось только одно значение, то снова выделить ячейки, нажать функциональную клавишу F2, в первой ячейке диапазона появится функция =ЧАСТОТА(A2:E11;B19:B25), снова нажать комбинацию клавиш Ctrl +

Shift + Enter).

Таблица 1.4

	Г	Н	І	Ј	К	Л
1	Время, мин	Абсолютные частоты	Интервалы	Абсолютные частоты по интервалам	Частности	Накопленные частности
2	1				0	0
3	10,1	22	(1–10)	22	0,44	0,44

4	19,3	11	(10–9)	11	0,22	0,66
5	28,4	9	(19–28)	9	0,18	0,84
6	37,6	3	(28–38)	3	0,06	0,9
7	46,7	1	(38–47)	1	0,02	0,92
8	55,9	2	(47–56)	2	0,04	0,96
9	65,0	1	(56–65)	2	0,04	1
10		1		50		

4. В ячейках I3:I9 запишем интервалы в обычном виде (1–10) и т. д. В ячейках J3:J9 поместим абсолютные частоты по данным интервалам, при этом сложим два последних значения функции =ЧАСТОТА (A2:E11;B19:B25), т. е. в ячейке J9 расположена формула =H9+H10. Просуммируем значения в ячейках J3:J9 и получим в J10 общую сумму частот  $n = 50$ .

5. Заполним столбец относительных частот. В ячейку K3 вводим формулу =J3/\$J\$10. Делаем автозаполнение ячеек K4:K9.

6. Рассчитываем столбец накопленных частностей. В ячейку L3 копируем значение относительной частоты из ячейки K3. В ячейку L4 введем формулу =L3+K4. Делаем автозаполнение ячеек L5:L9. Если все сосчитано правильно, то в ячейке L9 будет 1 (общая сумма накопленных относительных частот). В ячейках K2 и L2 для удобства построения графиков записать нули.

7. Построим с помощью **Мастера диаграмм MS Excel** гистограмму абсолютных частот, полигон и кумуляту времени обслуживания. Вначале выделяем диапазон ячеек I3:J9 и вызываем **Мастер диаграмм**. Выбираем вкладку **Гистограмма с группировкой**, через контекстное меню **Выбрать данные**. **Изменить** даем имя ряду «Абсолютные частоты». Получаем гистограмму (рис. 1.1).

### Гистограмма

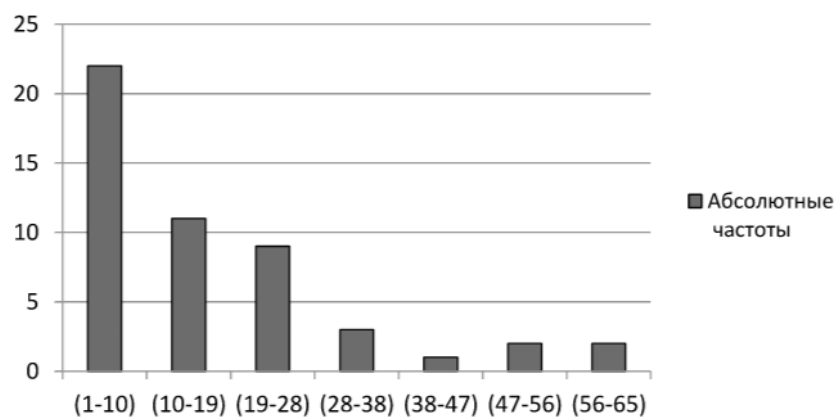


Рис. 1.1. Гистограмма

В математической статистике гистограммы изображаются в виде смежных прямоугольных областей, поэтому столбики гистограммы на рис. 1.5

расширяются до соприкосновения друг с другом. Для этого удобно воспользоваться контекстным меню **Формат ряда данных**. В диалоговом окне на вкладке **Параметры** установить **Перекрытие рядов** с минимальным зазором и **Боковой зазор** со значением без зазора. Можно задать при необходимости заливку и цвет границы.

8. Для добавления полигона на график гистограммы нужно через контекстное меню задать **Выбрать данные. Добавить. Значения**, указать тот же диапазон ячеек, что и для гистограммы, но с именем «Полигон». Через контекстное меню **Изменить тип диаграммы для ряда** задать для полигона **График с маркерами**.

9. Для построения накопленной функции частот (кумуляты) задать **Выбрать данные. Добавить. Значения**, указать диапазон накопленных частот и имя ряда «Накопленные частоты». В меню **Формат ряда данных. Параметры ряда** указать **Построить ряд по вспомогательной оси**. Можно задать вторую горизонтальную ось со значениями правых границ интервалов (для которых рассчитываются накопленные частоты). Графики представлены на рис. 1.2.

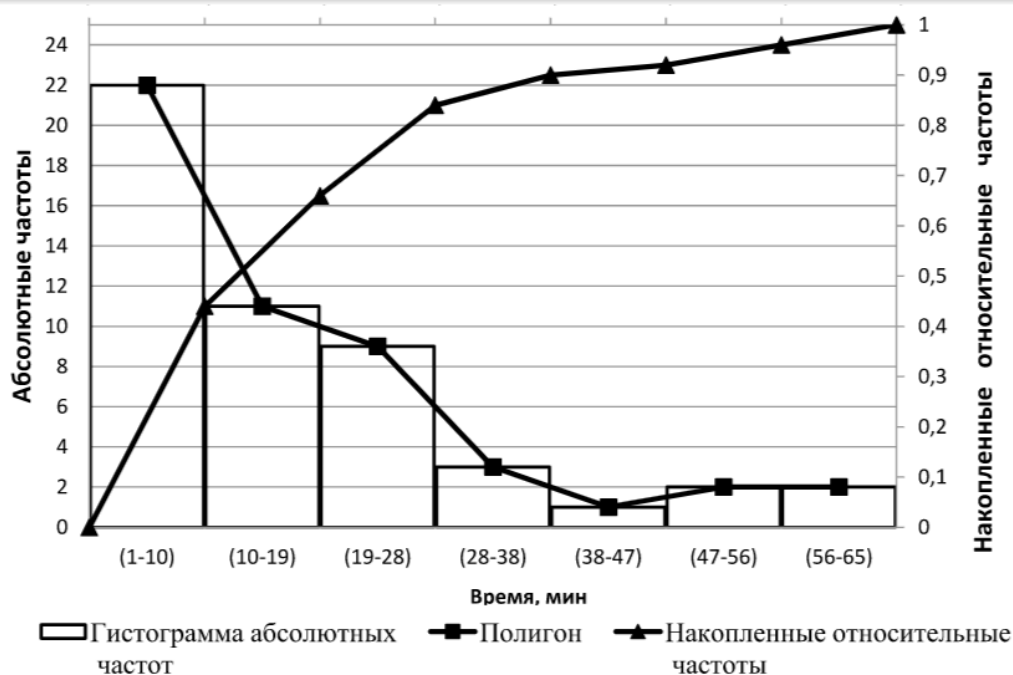


Рис. 1.2. Гистограмма абсолютных частот, полигон и кумулята

Далее нужно отформатировать оси: названия осей, минимальное и максимальное значения, цену делений, где будут подписи оси, тип, цвет линии и т. д. Все операции по форматированию графика можно сделать не только через контекстное меню, но и с помощью меню **Макет** (форматирование названия

диаграммы, осей, легенды, подписей данных, осей, сетки, области построения).

## Задание 2

Для данных предыдущей задачи (ячейки A1:E11) построить эмпирические распределения (гистограмму, накопленный процент частот) с помощью надстройки Microsoft Excel **Анализ данных. Гистограмма**.

*Решение.*

1. Для вызова процедуры **Гистограмма** выбрать меню **Данные. Анализ данных**. В окне **Инструменты анализа** выбрать **Гистограмма**.

2. В появившемся окне **Гистограмма** заполнить поля:

входной диапазон – введите диапазон исследуемых данных (A2:E11);

интервал карманов – G3:G9;

выходной интервал – указывается одна ячейка, ниже и правее которой будут размещены выходные данные (таблица, график);

поставить флажки **Интегральный процент**, **Вывод графика**.

В результате появляются таблица и не отформатированная гистограмма (рис. 1.3).

Карман	Частота	Интегральный %
10,1	22	44,00 %
19,3	11	66,00 %
28,4	9	84,00 %
37,6	3	90,00 %
46,7	1	92,00 %
55,9	2	96,00 %
65,0	1	98,00 %
Еще	1	100,00 %

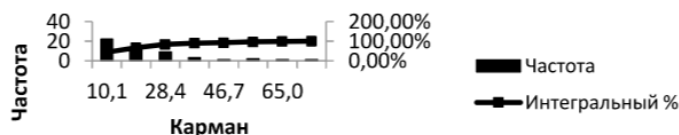


Рис. 1.3. Таблица и гистограмма

Теперь нужно привести графики к принятому в статистике виду. Частоту кармана «Еще» прибавляем к значению в предыдущем кармане. Расширяем столбики гистограммы до их соприкосновения друг с другом. Делаем подписи, максимальные значения, пересечения с осью, цену деления по осям. Добавляем вспомогательную горизонтальную ось для накопленной функции частот. Делаем нужные маркеры для кумуляты (рис. 1.4).

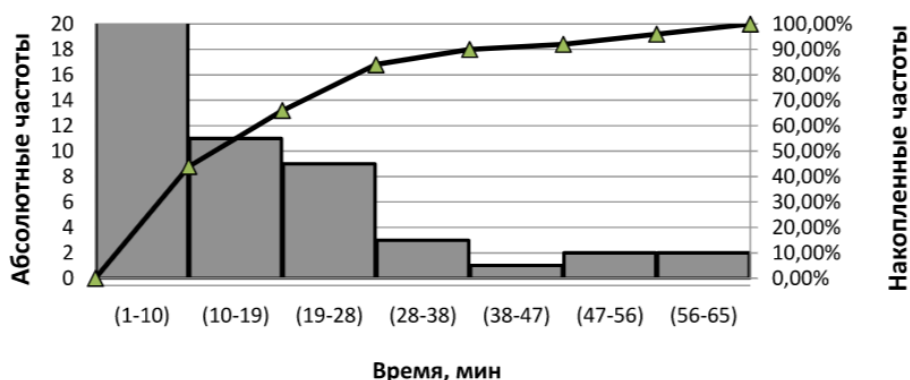


Рис. 1.4. Гистограмма абсолютных частот и кумулята

При использовании надстройки **Анализ данных. Гистограмма** можно не вводить диапазон карманов. Результат показан на рис. 1.5.

Карман	Частота	Интегральный %
1	5	10,00 %
22,333333	30	70,00 %
43,666667	11	92,00 %
Еще	4	100,00 %

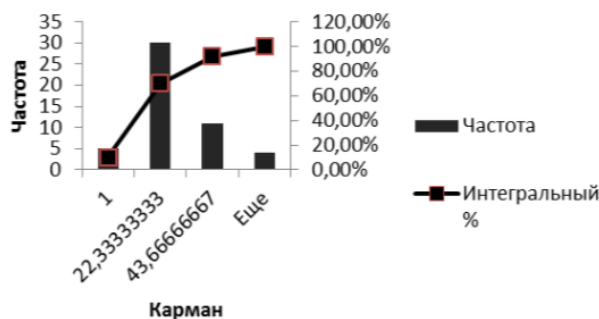


Рис. 1.5. Гистограмма абсолютных частот и кумулята

Если диапазон карманов не был введен, то будет создан набор интервалов, равномерно распределенных между минимальным и максимальным значениями данных. Границы интервалов определяются в этом случае автоматически, а полученные результаты будут отличаться от полученных ранее как по интервалам, так и по частотам. Гистограмма и кумулята выглядят по-другому. Данное обстоятельство следует обязательно учитывать при использовании надстройки **Анализ данных. Гистограмма**.