В программе Microsoft Excel следует создать файл данных с именем STAT

В открывшееся окно редактора следует занести все данные из выборки.

Указание: Заносить данные в файл следует в столбец, в дробных числах ставится запятая. После окончания набора данных рекомендуется сохранить данные.

Построение гистограмм.

Построение гистограммы осуществляется по следующему алгоритму.

• В выборке данных, занесенных в созданный файл, отыскивается наибольшее и наименьшее значения оцениваемого параметра Y: Y_{max} и Y_{min} соответственно.

Пример:

=MИH(A2:A101)

=MAKC(A2:A101)

где (А2:А101) – координаты диапазона значений данных выборки параметра Ү.

• Рассчитывается диапазон варьирования данных в выборке

- Осуществляется разбиение диапазона варьирования DY на интервалы ΔY . Число интервалов K и величина каждого интервала ΔY_i должны быть получены двумя различными способами:
 - рассчитаны математически по эмпирической формуле.
 - заданы пользователем согласно заданию.

Для расчета числа интервалов по эмпирической формуле используется формула Стерджесса

$$K = \varepsilon(1 + 3{,}322 \lg N),$$

где N - число данных в выборке; ϵ - операция округления до ближайшего целого.

Пример:
$$K = OKPY\Gamma \Pi((1+3,2*LOG(E10)); 0)$$
 где (E10) - N - число данных в выборке

Величина интервала ΔY_i рассчитывается как ΔY_i =DY/K. В этом случае все интервалы ΔY_i имеют одинаковую длину.

| Ymin | Ymax | | |
|--------|--------|---|--------|
| 341,79 | 649,19 | | |
| | | | |
| N | DY | K | ΔΥί |
| 100 | 307,4 | 8 | 38,425 |
| | | | |

• Формируется столбец границ интервалов.

В каждой последующей ячейке в столбце шаг за шагом прибавляется полученное значение величины интервала ΔYi : сначала к наименьшему значению оцениваемого параметра Ymin, затем в следующей ячейке ниже — к полученной сумме и т.д. Так постепенно доходя до максимального значения Ymax.

| Ymin | Ymax | | |
|--------|------------|---|--------|
| 341,79 | 649,19 | | |
| N | DY | K | ΔΥί |
| 100 | 307,4 | 8 | 38,425 |
| | _ | | |
| | Границы | | |
| № п/п | интервалов | | |
| 0 | 341,79 | | |
| 1 | 380,215 | | |
| 2 | 418,64 | | |
| 3 | 457,065 | | |
| 4 | 495,49 | | |
| 5 | 533,915 | | |
| 6 | 572,34 | | |
| 7 | 610,765 | | |
| | | | |

Интервалом считается следующий диапазон: (i-1; i] или i-1<значения<=i (нестрогая верхняя граница интервала — это значение в ячейке, нижняя строгая граница — значение в предыдущей ячейке).

• Определяется число данных выборки, попадающих в каждый из интервалов ΔY_i .

Пользователь выделяет столбец рядом со столбцом границ интервалов, не выделяя нулевую по порядку ячейку. Нажимает «F2» и вводит функцию:

= ЧАСТОТА(Массив данных; Массив интервалов)

Массив_данных это множество данных, для которых вычисляются частоты. Массив_интервалов это столбец границ интервалов, за исключением максимального и минимального значения.

Пример: = ЧАСТОТА(A2:A101; G15:G22)

После требуется нажать «Ctr+Shift+Enter» или, удерживая нажатыми клавиши Ctrl-Shift щелкнуть мышкой «Ok».

| | Границы | Количество |
|-------|------------|------------|
| № п/п | интервалов | точек |
| 0 | 341,79 | |
| 1 | 380,215 | 4 |
| 2 | 418,64 | 4 |
| 3 | 457,065 | 8 |
| 4 | 495,49 | 10 |
| 5 | 533,915 | 13 |
| 6 | 572,34 | 12 |
| 7 | 610,765 | 16 |
| 8 | 649,19 | 33 |
| | | |
| | | |

• Рассчитывается частота попадания в каждый интервал P_i = n_i/N , где n_i - число данных в i-том интервале.

| | Границы | Количество | Частота | |
|-------|--------------|------------|---------|--|
| № п/п | интервалов | точек | Pi | |
| (| 341,79 | | | |
| 1 | 380,215 | 4 | 0,04 | |
| 2 | 418,64 | 4 | 0,04 | |
| | 457,065 | 8 | 0,08 | |
| 4 | 495,49 | 10 | 0,1 | |
| | 533,915 | 13 | 0,13 | |
| (| 572,34 | 12 | 0,12 | |
| | 610,765 | 16 | 0,16 | |
| 8 | 649,19 | 33 | 0,33 | |
| | | | | |
| | Всего точек: | 100 | | |
| | | | | |

• На основании полученной таблицы с помощью стандартного инструмента для построения диаграмм «Мастер диаграмм» строится гистограмма распределения исследуемого показателя.



Тип диаграммы — гистограмма; диапазон данных — столбец «Частота P_i »; подписи по оси X — столбец «Границы интервалов»

Следует отметить, что на графике отражены только нижние границы интервалов!!!

После построения первой гистограммы, в которой число интервалов рассчитывается по эмпирической формуле, следует построить вторую гистограмму, где число интервалов задается пользователем согласно заданию.