***Приложение 1***

**СТАТИСТИКА И ДИЗАЙН ИНФОРМАЦИИ**

**1. Изучение нового материала**

Наш XXI в. характеризуют различным образом: век генной инженерии, век новых технологий (в частности, нанотехнологий), век астрофизики (проверка основополагающих космогонических теорий, большой андронный коллайдер) и т. д. Если вдуматься, все определения объединяет прежде всего, получение принципиально новой информации. Поэтому правильнее назвать наш век веком информации. Буквально за несколько последних лет появились  сверхмощные компьютеры, Интернет, различные поисковые системы, разрабатываются и совершенствуются методики обработки информации и т. п.
Многие из нас участвуют в переписи населения, выборах, опросах и т. д. При этом появляется определённая информация. Задача статистики – отражение этой информации и её обработка. Для этого необходимо ввести некоторые статистические характеристики. Рассмотрим следующий пример.

**Пример**. В финал конкурса «Мисс факультета» Вышли 10 студенток, за которых болели и голосовали 90 студентов. В таблице приведены результаты голосования за участниц с номерами 1-10.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № участницы | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Число голосов | 7 | 3 | 14 | 15 | 7 | 4 | 3 | 7 | 20 | 10 |

Прежде всего, возникает вопрос о наглядном отражении результатов голосования. Из курса алгебры  вы знаете, что графическая информация нагляднее табличной. Поэтому применяют три вида графического отражения информации – **диаграммы**.

Первый вид  диаграммы  – **линейная диаграмма** (или многоугольник распределения) строится как обычный график. По оси абцисс откладываются номера участниц, по оси ординат – число голосов, отданных за данную участницу, т. е. точки (1,7), (2,3), (3,14) и т. д. Для наглядности отмеченные точки соединены отрезками.



 Второй вид диаграммы **– столбчатая диаграмма** (или гистограмма распределения) строится следующим образом. В окрестности каждой отмеченной точки по оси абцисс строят прямоугольник, высота которого равна соответствующей ординате. При этом обычно ширину прямоугольников делают одинаковой. Достаточно часто прямоугольники изображаются таким образом, что два соседних имеют общую сторону. При этом прямоугольники могут штриховаться (см. учебник)



 Третья диаграмма – **круговая** (или камамбер, по названию французского сыра) представляет собой круг, разделённый на 10 секторов с различными центральными углами. Так как всего было подано 90 голосов, то каждому голосу соответствует 360о : 90 = 4о. Далее легко пересчитать углы секторов. Например, для первой участницы строим сектор с углом 4о **.** 7 =  28о. Каждый сектор маркируется номером соответствующей участницы. На практике применяют все три вида диаграмм.

Обсудим некоторые характеристики рассматриваемого примера.

***Объём измерения*** – количество источников информации  (т.е. число опрошенных или число голосов) в данном случае 90.

***Размах измерения*** – разница между наибольшим и наименьшим значениями результатов измерения. В данном случае 20 – 3 = 17, так как наибольшее число голосов 20, наименьшее 3.

***Мода измерения*** – наиболее часто встречающийся результат. В данном случае 9, так как за участницу №9 было подано 20 голосов (наибольшее количество).

***Среднее (или среднее арифметическое)*** – частное от деления суммы всех результатов измерения на объём измерения. В данном случае получают:



Обычно результатами измерений являются некоторые числа. Каждое  число, встретившееся в конкретном измерении, называют ***вариантой измерения***. В конкретном измерении его варианты могут быть никак не связаны (например, билетики с результатами голосования). Однако обычно результаты обрабатываются. Если записать все варианты измерения в некотором порядке (например, по времени поступления голосов  в жюри), то получится ряд данных измерения. Обычно упорядочивание проходит определённым образом. Запишем полученные варианты в порядке их возрастания (точнее, неубывания). Получим сгруппированный ряд данных:



Среднюю варианту  в сгруппированном ряде данных в случае нечётного количества чисел или средне арифметическое,  двух стоящих посередине вариант в случае чётного количества чисел называют ***медианой измерения***. В нашем случае средних варианты две, это варианты 45 и 46. Каждая из них равна 5, значит, и медиана равна 5. (5+5): 2=5

 В нашем примере ответ 1 встретился 7 раз (за участницу №1 проголосовали 7 человек). Поэтому говорят, что абсолютная частота (или кратность) варианты 1 равна семи. Поэтому ( в другой терминологии ранее приведённая таблица имеет вид:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Варианта | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Сумма |
| Кратность | 7 | 3 | 14 | 15 | 7 | 4 | 3 | 7 | 20 | 10 | 90 |

Таким образом, получаем таблицу распределения данных измерения. Графа «Сумма» добавляется для контроля: число в этой графе обязательно равняется объёму измерения. Заметим, что при вычислении среднего арифметического в неявном виде уже использовалось понятие кратности варианты. Введём ещё понятие ***частоты*** данной ***варианты*** – частное от деления кратности варианты на объём измерения. Например для варианты 1 находим частоту  . Частоту варианты можно выразить в ***процентах***. Тогда получим: 

****

**(Таблица распределения частот)**