Задание:

1. **Выполните практическую работу:**
2. **Вышлите конспект к 07.12.2021 г,** по адресу **orlova__olga@list.ru**

(между фамилией и именем два подчеркивания)

Практическая работа №12.

Тема: Составление статистического распределения выборки, построение полигона и гистограммы.

Цель:

- сформировать представление о задачах математической статистики и овладеть навыками их решения.

Обеспечение практической работы: методические указания для практической работы, средства вычислительной техники.

Порядок выполнения работы:

1. Записать в тетрадь тему и цели практической работы.
2. Повторить теоретический материал по теме практической работы.
3. Ознакомиться с методикой решения задач.
4. Решить задачи самостоятельно.
5. Ответить на контрольные вопросы.

Сведения из теории.

Генеральная и выборочная совокупности

Пусть имеется некоторая совокупность (множество) N однородных объектов. Назовем ее **генеральной совокупностью**, а число N – ее **объемом**.

Пусть требуется изучить генеральную совокупность относительно качественного или количественного признака. Например, имеется генеральная совокупность объемом N = 1000 деталей.
Примером качественного признака этой совокупности является стандартность деталей, а примером **количественного** признака – размер детали. Количественный признак выражается числом.

При этом можно проводить сплошное обследование всех объектов генеральной совокупности. Часто это неприемлемо, потому что:

- физически трудно;
- экономически нецелесообразно (дорого);
- иногда приводит к разрушению исследуемых объектов.

В таких случаях из генеральной совокупности отбирают некоторую часть объектов – n штук и подвергают их изучению, а выводы, полученные при изучении этой совокупности, распространяют на всю генеральную совокупность.

Определение. **Выборочной совокупностью** или просто **выборкой объема n** называется совокупность n объектов, отобранных из генеральной совокупности.

Выводы, полученные при изучении выборки, называются **эмпирическими**, то есть полученными опытным путем.

Замечание. Часто, если объем генеральной совокупности N достаточно велик, считают, что генеральная совокупность имеет бесконечный объем

Статистическое распределение выборки

Из генеральной совокупности, элементы которой обладают количественным признаком, производится выборка из n элементов. Наблюдаемые значения $x_1, x_2, \ldots, x_n$ количественного признака называется вариантами.

Последовательность значений вариант, записанная в возрастающем порядке, называется **вариационным рядом**. Число наблюдений того или иного значения признака называется частотой и обозначается n_i. Это означает, что варианта x_i наблюдалась n_i раз.

Относительной частотой варианты (признака) в выборке называется отношение:

$$W_i = \frac{n_i}{n},$$

где n - объем выборки.

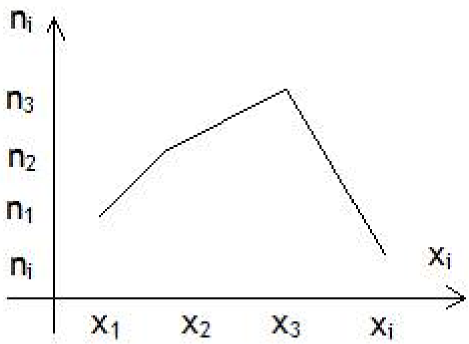
Статистическим распределением выборки называется перечень вариант и соответствующих им частот.

Варианта	x_1	x_2	x_3	...	x_k
Частота	n_1	n_2	n_3	...	n_k

$n_1 + n_2 + n_3 + \ldots + n_k = n$.

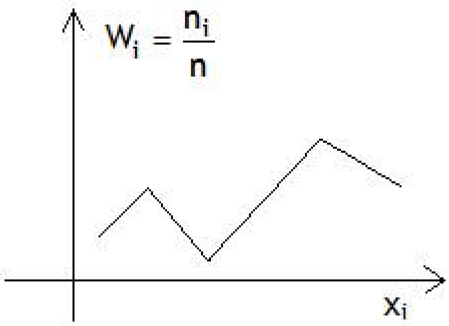
Полигон и гистограмма

Если соединить точки (x_1, n_1), …, (x_k, n_k), то получим ломаную линию, которая называется **полигоном частот**.



Полигон частот (x_i, n_i)

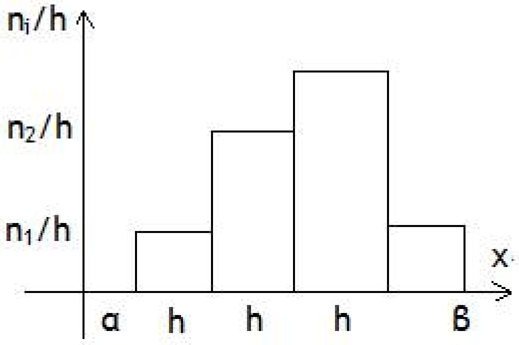
Если по оси ординат откладывать относительную частоту, а по оси абсцисс – значения количественного признака x_i, то получим **полигон относительных** частот.



Полигон относительных частот (x_i, n_i /n)

Гистограммой называют ступенчатую фигуру, состоящую из прямоугольников, основания которых – частичные интервалы длиной h, а высоты равны n_i /h (i = 1, …, k).

Гистограммы строят, если количественный признак изменяется непрерывно, например, в интервале (α, β). Для построения **гистограммы частот** интервал (α, β) делят на частичные интервалы длиной h. По оси ординат откладывают высоту прямоугольника, равную n_i/h, где n_i - сумма частот всех вариант, попадающих в данный интервал.



Гистограмма

Площадь гистограммы частот равна сумме всех частот, то есть объему выборки.

Гистограммой относительных частот называется ступенчатая фигура, построенная аналогично, но по оси ординат откладывают плотность относительных частот, то есть n_i/hn. Площадь гистограммы относительных частот равна 1.

Полигоны и гистограммы используют для графического изображения статистических распределений.

Эмпирическая функция распределения

Эмпирическая функция распределения определяется по формуле:

$$F^*(x) = \frac{n_x}{n},$$

где n - объем выборки, n_x - число вариант, меньших x.

Она обладает теми же свойствами, что и рассмотренная в теории вероятностей функция распределения F(x).

Числовые характеристики статистического распределения выборки

Пусть выборка объема *n* состоит из чисел x_i , причем значение x_i встречается n_i раз (i = 1, …, k), $n_1 + n_2 + n_3 + \ldots + n_k = n$. Числовыми характеристиками статистического распределения выборки являются:

1) Выборочная средняя

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i . \qquad (15.1)$$

2) Размах выборки: $R = x_{max} - x_{min}$, где x_{max} и x_{min} – максимальное и минимальное числа в выборке.

3) Выборочная дисперсия:

$$D_E = \frac{1}{n}\sum_{i=1}^{k} (x_i - \overline{x}_E)^2 \cdot n_i .$$

Её удобнее вычислять по формуле:

$$D_E = \frac{1}{n}\left(\sum_{i=1}^{k} n_i x_i^2 - n(\overline{x})^2\right). \qquad (15.2)$$

4) Выборочное среднее квадратическое отклонение:

$$\sigma = \sqrt{D_в} .$$

<u>Пример</u>

Дана таблица распределения данных:

в выборке число x_i встречается n_i раз ($i = 1, \ldots, 6$). Найти $\overline{x}$ и D_E.

i	n_i	x_i
1	5	-2
2	15	-1
3	50	0
4	16	1
5	10	2
6	4	3

Для удобства вычислений добавим столбцы $n_i x_i$, $n_i x_i^2$ и строку Σ (сумм по столбцу):

i	n_i	x_i	n_ix_i	$n_ix_i^2$
1	5	-2	-10	20
2	15	-1	-15	15
3	50	0	0	0
4	16	1	16	16
5	10	2	20	40
6	4	3	12	36
Σ	100		23	127

По формуле (15.1) выборочное среднее:

$$\bar{x} = \frac{1}{100}23 = 0{,}23,$$

а по (15.2) выборочная дисперсия

$$D_E = \frac{1}{100}(127 - 100 \cdot (0{,}23)^2) \approx 1{,}22.$$

Задание для самостоятельного решения

Пример. Работников предприятия попросили с точностью до 10 минут оценить время, которое они тратят на дорогу до работы. Было опрошено 50 человек.

Полученные результаты были следующими: 20, 100, 20, 30, 40, 50, 30, 80, 90, 40, 30, 50, 20, 50, 30, 30, 50, 60, 60, 50, 30, 40, 60, 50, 100, 60, 90, 10, 20, 50, 90, 80, 20, 40, 50, 10, 50, 40, 30, 40, 60, 120, 30, 40, 60, 20, 60, 10, 50, 60.

Составьте таблицу распределения данных.

Найдите числовые характеристики данных измерения:
выборочное среднее $\bar{x}$, размах выборки, выборочную дисперсию, среднее квадратическое отклонение.

Постройте полигон частот.

Контрольные вопросы

1. Объясните, что такое генеральная и выборочная совокупность.
2. Что такое вариационный ряд?
3. Как построить полигон частот?
4. Какие вы знаете числовые характеристики статистического распределения выборки и как их найти?