

Статистические ряды распределения

Лекция 4

Математическая статистика- это раздел математики, посвященный методам сбора, анализа и обработки результатов статистических данных наблюдений для научных и практических целей.

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить стандартность детали, а количественным- контролируемый размер детали.

Исследование:

- сплошное
- выборочное

Основная задача математической статистики заключается в исследовании всей совокупности по выборочным данным в зависимости от поставленной цели, например, для принятия управленческих решений в условиях неопределенности.

Основные понятия

- ▣ **Генеральная совокупность** – это совокупность объектов, из которой производится выборка.
- ▣ **Выборочная совокупность (выборка)** – это совокупность случайно отобранных объектов.
- ▣ **Объем совокупности** – это число объектов этой совокупности. Объем генеральной совокупности обозначается N , выборочной – n .



Выборка

Для того, чтобы по данным выборки можно было достаточно уверенно судить об интересующем признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. Выборка должна правильно представлять пропорции генеральной совокупности. Выборка должна быть **репрезентативной (представительной)**.



Большинство встречающихся на практике величин принимают неодинаковые значения у различных членов совокупности

- ▣ **Статистический ряд распределения** – это упорядоченное распределение единиц совокупности на группы по определенному варьирующемуся признаку (стаж работы, возраст, пол и т.д.)

Виды статистических рядов распределения и их элементы

Атрибутивный ряд

Вариационный ряд

В зависимости от характера
вариации

Дискретный ряд

Интервальный ряд



Атрибутивный ряд

- Ряд построенный по атрибутивному признаку (пол, занятость, национальность, профессия и пр.)

Распределение студентов I курса экономического факультета по полу

Группа студентов, пол	Число студентов	Удельный вес в общей численности, %
Женщины	90	60,0
Мужчины	60	40,0
Всего	150	100,0



Вариационный ряд

- Ряд построенный по количественному признаку (в порядке возрастания или убывания)

Число детей в семье	Количество семей	Удельный вес в общей численности, %
1	700	70,0
2	250	25,0
Более 2	50	5,0
Всего	1000	100,0



Характеристики вариационных рядов:

Варианты – это числовые значения

количественного признака в вариационном ряду
распределения

Частоты – это численности отдельных вариантов

или каждой группы вариационного ряда, т.е.

числа, показывающие насколько часто

встречаются те или иные варианты в ряду

распределения



Дискретный вариационный ряд

- В основе этого ряда лежит дискретный (прерывный) признак, т.е. имеющий только целые значения (число студентов в группе, размер обуви)

Число детей в семье	Количество семей	Удельный вес в общей численности, %
1	700	70,0
2	250	25,0
Более 2	50	5,0
Всего	1000	100,0



Интервальный вариационный ряд

- В основе этого ряда лежит непрерывный признак, который может принимать любые значения (температура воздуха, объем выручки)

Численность работающих, чел.	Число торговых предприятий	Удельный вес, % к итогу
50-100	24	15,00
100-150	36	22,50
150-200	50	31,25
200-250	28	17,50
250 и выше	22	13,75
Всего	160	100,00

Числовые характеристики рядов

1. Выборочная средняя

Дискретный ряд:

$$\bar{x}_v = \frac{\sum_{i=1}^k x_i \cdot n_i}{n} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \dots + x_k \cdot n_k}{n_1 + n_2 + \dots + n_k}$$

Интервальный ряд:

x_i - середины интервалов $x_i = \frac{a+b}{2}$

2. Дисперсия

$$D_{\sigma} = \overline{x^2} - (\overline{x})^2$$

$$\overline{x^2} = \frac{\sum_{i=1}^k x_i^2 \cdot n_i}{n}$$



3. Среднее квадратичное отклонение

$$\sigma = \sqrt{D_v}$$



Пример

Интервалы	(20; 26)	(26; 32)	(32; 38)	(38; 44)	(44; 50)	(50; 56)	(56; 62)	(62; 68)
Частоты	1	4	20	43	60	44	23	5

Для данного интервального ряда
распределения требуется вычислить
числовые характеристики выборки



Решение

1. Строим вспомогательную таблицу, вычисляем середины интервалов.

Интервалы	x_i	n_i	$x_i \cdot n_i$	x_i^2	$x_i^2 \cdot n_i$
20-26	23	1	23	529	529
26-32	29	4	116	841	3364
32-38	35	20	700	1225	24500
38-44	41	43	1763	1681	72283
44-50	47	60	2820	2209	132540
50-56	53	44	2332	2809	123596
56-62	59	23	1357	3481	80063
62-68	65	5	325	4225	21125
Сумма	-	200	9436	-	458000

$$\bar{x}_g = \frac{9436}{200} = 47,18$$

$$\bar{x}_g^2 = \frac{458000}{200} = 2290$$

$$D_g = 2290 - 47,18^2 = 64,0476 \approx 64,05$$

$$\sigma = \sqrt{64,05} \approx 8$$