Лекция 2. Средние величины

План

- Средняя арифметическая
- Мода
- Медиана
- Средние показатели динамики

Средняя арифметическая

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n},$$ где

$X_ь$, X_2, X_3... $X_п$, - варианты признака;

n - число единиц наблюдения.

Например:

Даны сведения о заработной плате шести работников (в условных единицах) - 90, 120, 108, 206, 160, 184. Определить средний размер заработной платы данной совокупности работников.

$$\overline{X} = \frac{90 + 120 + 108 + 206 + 160 + 184}{6} \approx 144{,}67$$

Средняя арифметическая

$$\overline{X} = \frac{\sum XP}{\sum P}, \quad \text{где}$$

X - варианты признака;

P - частота вариант (в статистической литературе для обозначения частоты используют букву "f", в последние годы это обозначение проникает и в историческую литературу).

Распределение футбольных матчей высшей лиги России по числу забитых мячей за игру в 1992 г.

число забит, мячей	0	1	2	3	4	5	6	7
число матчей	21	41	42	37	19	10	6	3

Определить среднее число забитых голов за одну игру.

$$\overline{X} = \frac{0 \cdot 21 + 1 \cdot 41 + 2 \cdot 42 + 3 \cdot 37 + 4 \cdot 19 + 5 \cdot 10 + 6 \cdot 6 + 7 \cdot 3}{21 + 41 + 42 + 37 + 19 + 10 + 63} \approx 2{,}34$$

Мода

Мода. (Мо) представляет наиболее часто встречающееся значение признака в упорядоченной совокупности, наиболее типичное среднее значение.

Для вычисления моды в интервальном ряду сначала определяется модальный класс, т.е. интервал с наибольшей частотой. Затем Мо вычисляется по формуле:

$$Mo = X_o + K\frac{P_2 - P_1}{2P_2 - P_1 - P_3},$$ где

X_o - нижняя граница модального интервала;

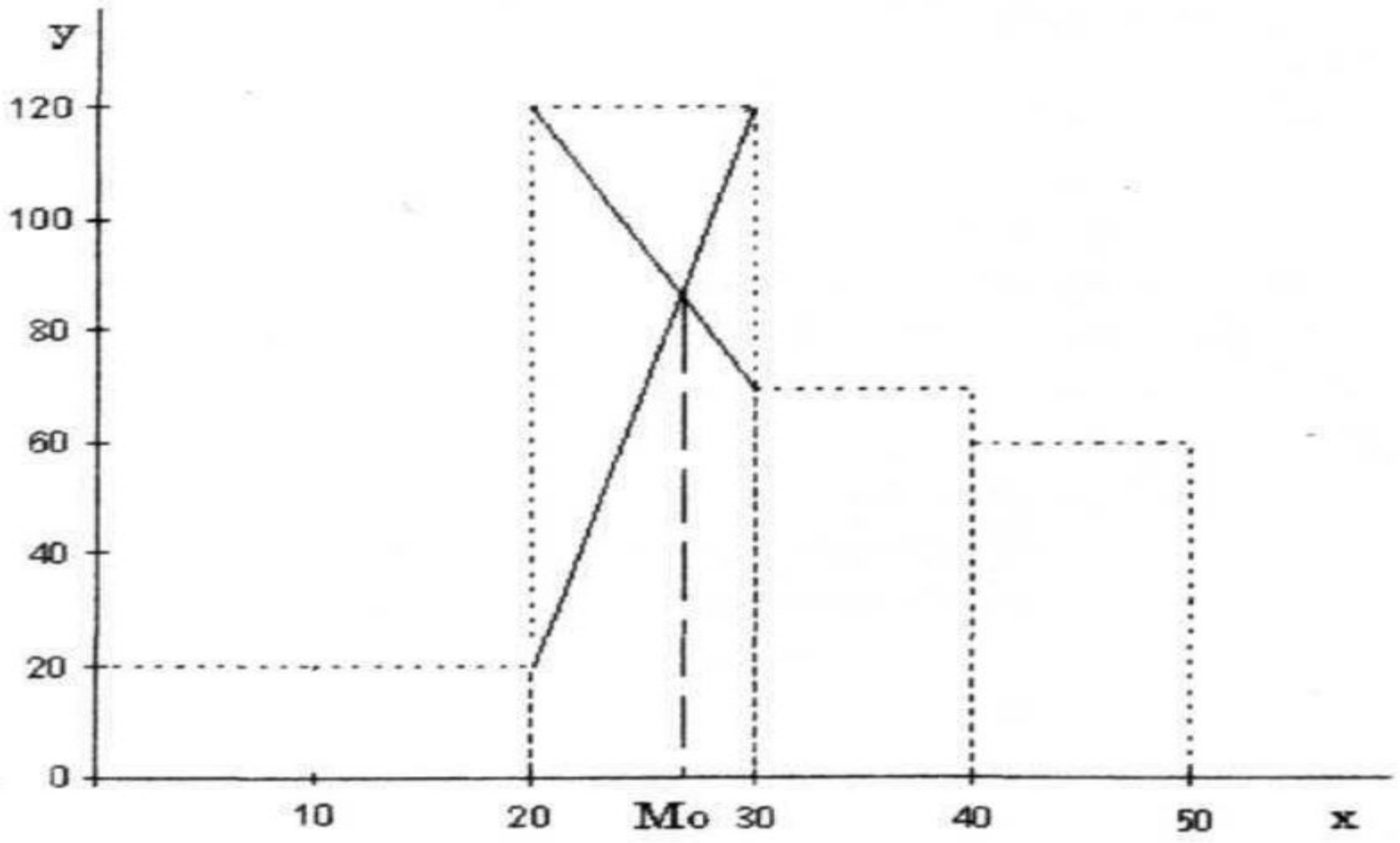
K - величина интервала;

P_1 - частота интервала, предшествующего модальному;

P_2 - частота модального интервала;

P_3 - частота интервала, последующего за модальным.

Графическое представление моды



Медиана

Медиана. (Ме) - величина, определяющая значение признака, находящегося в середине упорядоченной совокупности. Медиана делит изучаемую совокупность так, что число единиц с большим и меньшим, чем медиана значением признака, одинаково.

$$Me = X_o + K\frac{\sum P/2 - \sum_{m-1}}{P_m}, \text{ где}$$

X_o - нижняя граница медианного интервала;

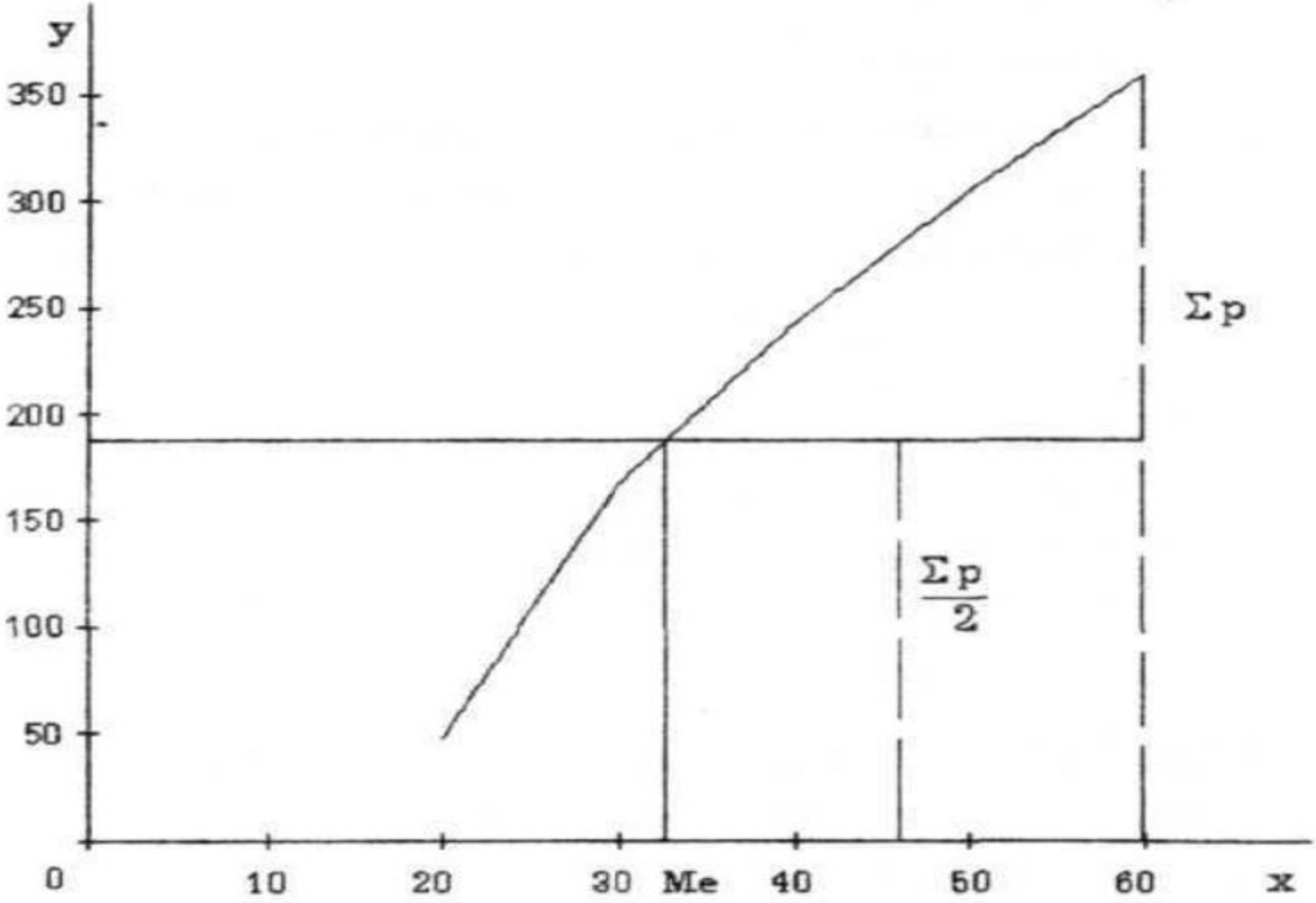
K - величина медианного интервала;

$\sum P$ - сумма частот;

$\sum_{m-1}$ - сумма частот интервалов, предшествующих медианному (накопленная частота в интервале, предшествующем медианному).

P_m - частота медианного интервала.

Графическое представление медианы



Меры разброса признака

Простейшим способом изучения вариации признака в совокупности является размах вариации или ее амплитуда (R) Величина R определяется как разность между максимальным и минимальным значениями признака в изучаемой совокупности.

$$R = X_{max} - X_{min}$$

нака. На практике чаще всего прибегают к изучению ***среднего квадратического отклонения*** (стандартного отклонения) конкретных значений признака от его средней величины. Оно обозначается σ (сигма) или S и позволяет определить границы, в которых изменяются конкретные значения признака. Величина, насколько в среднем каждое значение признака отличается от его средней арифметической, находится по формуле:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}, \qquad \text{где}$$

x- конкретные значения признака;

$\bar{x}$ - средняя арифметическая ;

n- число наблюдений.

Стандартное квадратическое отклонение, возведенное в квадрат, называется дисперсией.

Меры разброса признака

В том случае, если мы имеем дело с группировкой, с интервальным рядом, формула видоизменяется:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 p}{\sum p}}, \qquad \text{где}$$

x- конкретные значения признака (для интервальной группировки - срединные значения признака);

$\bar{x}$ - средняя арифметическая;

p- частота признака в группировке.

Достаточно просто вычисляется среднее квадратическое отклонение для определения размаха вариации качественных (альтернативных) признаков. Формула выглядит так:

$$\sigma = \sqrt{\frac{p_1 p_2}{n}}, \qquad \text{где}$$

P_1 - частота первой варианты признака;

P_2 - частота второй варианты признака;

n - число наблюдений.

Меры разброса признака

Дисперсии при интерпретации выражаются в тех же единицах, что и сами признаки. Это приводит, к тому, что будучи выражены в разных единицах измерения, средние квадратические отклонения несравнимы. То есть, нельзя сравнивать количество детей с земельной площадью. В случае необходимости пользуются коэффициентом вариации (V), определяемым как отношение стандартного отклонения к средней арифметической.

$$V = \sigma / \bar{x}$$

Полученную величину можно выразить в процентах. Сопоставление коэффициентов вариации нескольких признаков расширяет возможности исследователя при анализе и интерпретации распределений признаков - их равномерности, нормальности, колеблемости.

Показатели вариации раскрывают уровень репрезентатив-

Средняя квадратическая

Средняя квадратическая. Историки иной раз сталкиваются с проблемой определения средней земельной площади (площадь посевов; территория археологических раскопов; регион, охваченный восстанием и т.п.), но при этом известны не площади осредняемых участков, а только их линейные параметры. Путем качественного анализа проверяется возможность допущения, что каждый из рассматриваемых земельных участков имеет вид квадрата. Если проверка дала положительный результат, то для расчета средней площади имеющихся участков применяется формула средней квадратической:

$$\overline{X}_{кв} = \sqrt{\frac{X_1^2 + X_2^2 + X_3^2 + \ldots + X_n^2}{n}}$$

Средние показатели динамики. К средним показателям динамики относятся средний уровень ряда, средние абсолютные изменения и ускорения, средний темп роста. Все они выступают характеристиками тенденции.

Средний ***уровень*** $(\bar{y})$ интервального динамического ряда определяется как простая средняя арифметическая из уровней за равные промежутки времени или как средневзвешенная из уровней за неравные промежутки времени, длительность которых выступает в качестве "весов".

Средние показатели динамики

В моментном ряду средняя величина характеризует обобщенное значение признака между начальным и конечным моментом наблюдения. Следовательно, начальный и конечный уровни лишь наполовину относятся к изучаемому отрезку времени, а на половину к прошлому и будущему периодам. Это обстоятельство определило формулу ***средней хронологической.***

$$\bar{y}_{хр} = \frac{y_1/2 + y_2 + y_3 + \ldots + y_n/2}{n-1}$$

В случае, если промежутки между датами моментного ряда не равны, хронологическая средняя вычисляется по формуле:

$$\bar{y}_{хр} = \frac{y_1 T_1 + y_2(T_1 + T_2) + y_3(T_2 + T_3) + \ldots y_n T_{n-1}}{2(T_1 + T_2 + T_3 + \ldots + T_{n-1})},$$ где

T - промежутки между датами;

$y_1, y_2, y_3 \ldots y_n$ - уровни ряда;

n - количество уровней.

Средние показатели динамики

Один из наиболее важных средних показателей динамического ряда - ***средний темп изменения*** (роста или сокращения). Он определяется по формуле средней геометрической. В литературе до сих пор нет универсального символа для обозначения этого показателя. Так, у Т.И.Славко используется G, у И.Д.Ковальченко - K_p, у И..И.Елисеевой - $\bar{k}$, у А.Я.Боярского - $\bar{T}$.

$$\bar{T} = \sqrt[n-1]{\frac{У_n}{У_1}},$$ где

$У_1$ - начальный уровень динамического ряда;

$У_п$ - последний уровень диинамического ряда;

n - число временных промежутков.

Средние показатели динамики

С помощью средней геометрической величины мы получили среднюю скорость изменения признака. В нашем распоряжении есть еще один средний показатель - **средний абсолютный прироста** (абсолютное значение). Он показывает на какую величину (в единицах измерения уровней ряда) показатель одного временного периода больше или меньше любого предшествующего. При возрастании уровней абсолютное изменение принимает положительное, а при уменьшении - отрицательное значения.

Он определяется по формуле:

$$\Delta\overline{У} = \frac{У_n - У_1}{n},$$ где

$У_1$ - начальный уровень динамического ряда;

$У_n$ - конечный уровень динамичческого ряда;

n - число временных промежутков (число осредняемых отрезков времени).

Средние показатели динамики

По среднему темпу роста можно без труда определить ***средний темп прироста,*** вычитая из значения $\overline{T}$ единицу. В нашем примере $\overline{T} = 1{,}16$. Тогда средний темп прироста: 1,16-1 = 0,16. Полученное значение можно выразить в процентах, умножив его на 100% (у нас 0,16 100% = 16%).

Разделив абсолютный прирост на темп прироста (за соответствующий период) получим среднее абсолютное значение прироста $(\overline{\alpha})$:

$$\overline{\alpha} = \overline{\Delta y} / \overline{T}_{np}$$

Встречаются ситуации, когда темпы роста и прироста, а также абсолютные приросты по годам снижаются, в то же время абсолютные значения 1% прироста возрастают. Может быть и обратный процесс.