ЛЕКЦИЯ

Дисциплина «ПРАВОВАЯ СТАТИСТИКА»

Тема № 4-1:

«Абсолютные и относительные показатели, средние величины»

Кафедра Управления и информационно-технического обеспечения деятельности УИС
ОЗЁРСКИЙ СЕРГЕЙ ВЛАДИМИРОВИЧ, к.ф.-м.н., доцент
САМАРА 2014

План лекции

1. Абсолютные и относительные показатели (величины).
2. Средние величины.

Статистические показатели

Статистический показатель – это количественная характеристика социально-экономического явления или процесса, вычисленная с учётом их качественных характеристик

В отличие от признака статистический показатель, чаще всего, получается путем расчета

Статистические показатели в зависимости от способа их вычисления подразделяются на:

Абсолютные –

- Это суммарные обобщающие показатель, характеризующие размеры изучаемых явлений в конкретных условиях места и времени.
- Это исходная, первичная, самая общая форма выражения СП; числа, взятые из таблиц без преобразований.
- Это именованные величины, выраженные через единицы измерения

Относительные –

- Представляют собой производные обобщающие показатели, получаемые в результате деления одних абсолютных показателей на другие.
- Позволяют провести сравнение различных показателей.
- Как правило, измеряются в безразмерных коэффициентах или процентах.

Относительные величины вычисляются как отношение двух чисел:

- Числитель называют *сравниваемой (текущей) величиной*
- Знаменатель называют *базой относительного сравнения (предшествующая величина)*

$$\text{ОСП} = \frac{\textit{сравниваемая величина}}{\textit{база относительного сравнения}}$$

Средние величины

Средняя величина – это обобщающий показатель, характеризующий значение признака, вокруг которого концентрируются наблюдения.

Различают следующие средние величины:

- **Структурные средние величины** (медиана, мода и др.)
- **Аналитические средние величины** (средняя степенная)

Структурные средние величины

Медианой называют вариант, приходящийся на середину вариационного ряда

для дискретного вариационного ряда

$$\widetilde{Me} = \begin{cases} x_j, & \text{если} \quad n = 2j - 1, \\ \frac{1}{2}\left(x_j + x_{j+1}\right), & \text{если} \quad n = 2j \end{cases}$$

x_i	13	14	15	16	17
m_i	12	22	28	30	8

Структурные средние величины

для интервального вариационного ряда

$$\widetilde{Me} = x_{\min} + d\,\frac{\frac{1}{2}\sum_{i=1}^{k} m_i - \sum_{i=1}^{i_{Me}-1} m_i}{m_{Me}};$$

где $d = x_{\max} - x_{\min}$ - ширина медианного интервала

x_i	100-120	120-140	140-160	160-180
m_i	9	16	11	4

Структурные средние величины

Модой $\widetilde{Mo}$ называют вариант, имеющий наибольшую частоту встречаемости m_i вариационного ряда

для дискретного вариационного ряда

x_i	**13**	**14**	**15**	**16**	**17**
m_i	12	22	28	**30**	8

Структурные средние величины

для интервального вариационного ряда

$$\widetilde{Mo} = x_{\min} + d\frac{\left(m_{Mo} - m_{Mo-1}\right)}{\left(m_{Mo} - m_{Mo-1}\right) + \left(m_{Mo} - m_{Mo+1}\right)};$$

где $d = x_{\max} - x_{\min}$ - ширина модального интервала

x_i	100-120	120-140	140-160	160-180
m_i	9	16	11	4

Аналитические средние величины

Средняя степенная

$$\bar{x}_\alpha = \left(\frac{x_1^\alpha + x_2^\alpha + \dots + x_n^\alpha}{n} \right)^{\frac{1}{\alpha}};$$

- α **= 1** – средняя арифметическая;
- α **= -1** – средняя гармоническая;
- α **= 0** – средняя геометрическая;
- α **= 2** – средняя квадратическая.

Аналитические средние величины

Средняя арифметическая величина

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n};$$

Средняя арифметическая взвешенная

$$\bar{x} = \frac{x_1 m_1 + x_2 m_2 + \cdots + x_k m_k}{m_1 + m_2 + \cdots + m_k} = \frac{\sum_{i=1}^{k} x_i m_i}{\sum_{i=1}^{k} m_i};$$

Аналитические средние величины

Средняя геометрическая величина

$$\overline{g} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i};$$

Средняя геометрическая взвешенная

$$\overline{g} = \sqrt[n]{x_1^{m_1} \cdot x_2^{m_2} \cdot \ldots \cdot x_k^{m_k}} = \sqrt[n]{\prod_{i=1}^{k} x_i^{m_i}};$$

Аналитические средние величины

Средняя гармоническая величина

$$\bar{h} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_k}};$$

Средняя гармоническая взвешенная

$$\bar{h} = \frac{\sum_{i=1}^{k} m_i}{\frac{m_1}{x_1} + \frac{m_2}{x_2} + \dots + \frac{m_k}{x_k}} = \frac{\sum_{i=1}^{k} m_i}{\sum_{i=1}^{k} \frac{m_i}{x_i}};$$

Аналитические средние величины

Средняя квадратическая величина

$$\bar{s} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}};$$

Средняя квадратическая взвешенная

$$\bar{s} = \sqrt{\frac{x_1^2 m_1 + x_2^2 m_2 + \dots + x_k^2 m_k}{m_1 + m_2 + \dots + m_k}} = \sqrt{\frac{\sum_{i=1}^{k} x_i^2 m_i}{\sum_{i=1}^{k} m_i}};$$

Пример использования средних величин (*средняя скорость движения*)

Пример использования средних величин (*средний темп роста*)

Показатели	Годы				
	2004	**2005**	**2006**	**2007**	**2008**
	1	2	3	4	5
Число экономических преступлений	105	116	118	124	160
Темпы роста	-	**1.1048**	**1.0172**	**1.0508**	**1.2903**

$$\overline{T}_p = \overline{x} = \frac{1.1048 + 1.0172 + 1.0508 + 1.2903}{4} = 1.1158.$$

$$\overline{T}_p = \overline{g} = \sqrt[4]{1.1048 \cdot 1.0172 \cdot 1.0508 \cdot 1.2903} = 1.1110.$$

Свойство мажорантности средних величин

$$\bar{h} \le \bar{g} \le \bar{x} \le \bar{s}$$

$$x^{-1} \le x^{0} \le x^{1} \le x^{2}$$

Показатели вариации

Размахом вариационного ряда называют абсолютную величину разности между максимальными и минимальными значениями (вариантами) изучаемого признака

$$R = x_{\max} - x_{\min}$$

дискретный вариационный ряд

x_i	13	14	15	16	17
m_i	12	22	28	30	8

интервальный вариационный ряд

x_i	100-120	120-140	140-160	160-180
m_i	9	16	11	4

Показатели вариации

Средним линейным отклонением вариационного ряда называют среднюю арифметическую абсолютных величин отклонений вариантов от их средней арифметической

$$d = \frac{|x_1 - \bar{x}|m_1 + |x_2 - \bar{x}|m_2 + \dots + |x_k - \bar{x}|m_k}{m_1 + m_2 + \dots + m_k} = \frac{\sum_{i=1}^{k} |x_i - \bar{x}|m_i}{\sum_{i=1}^{k} m_i}$$

Показатели вариации

Дисперсией вариационного ряда называют среднюю арифметическую квадратов отклонений вариантов от их средней арифметической

$$s^2 = \frac{(x_1 - \bar{x})^2 m_1 + (x_2 - \bar{x})^2 m_2 + \dots + (x_k - \bar{x})^2 m_k}{m_1 + m_2 + \dots + m_k} = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 m_i}{\sum_{i=1}^{k} m_i}$$

учитывая, что $w_i = \dfrac{m_i}{\sum_{i=1}^{k} m_i}$, получим:

$$s^2 = (x_1 - \bar{x})^2 w_1 + (x_2 - \bar{x})^2 w_2 + \dots + (x_k - \bar{x})^2 w_k = \sum_{i=1}^{k} (x_i - \bar{x})^2 w_i$$

Показатели вариации

Среднее квадратическое отклонение:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 m_i}{\sum_{i=1}^{k} m_i}}.$$

Коэффициент вариации:

$$\bar{v} = \frac{s}{\bar{x}} \cdot 100\% \quad (\bar{x} \neq 0).$$

Понятие генеральной и выборочной совокупности

Совокупность всех мысленно возможных объектов того или иного вида, над которыми проводятся наблюдения с целью получения конкретных значений определённой случайной величины, или совокупность результатов всех мыслимых наблюдений, проводимых при неизменных условиях над одной из случайных величин, связанных с данным видом объектов, называют ***генеральной совокупностью***.

Понятие генеральной и выборочной совокупности

Часть отобранных объектов из генеральной совокупности (результаты наблюдений над ограниченным числом объектов из этой совокупности) называется ***выборочной совокупностью*** или просто ***выборкой***.

Способы отбора статистических данных

- ***Собственно случайный отбор***, при котором объекты выбираются путём жеребьевки.
- ***Механический отбор***, при котором генеральную совокупность «*механически*» делят на столько групп, сколько объектов должно войти в выборку, а из каждой группы выбирают один объект.
- ***Серийный отбор***, при котором объекты отбираются из генеральной совокупности не по одному, а сериями, которые подвергаются сплошному обследованию.
- ***Типический*** (***районированный***) ***отбор***, при котором объекты отбираются не из всей генеральной совокупности, а из каждой ее *типической* (*однородной*) части.

Основные характеристики генеральной и выборочной совокупности

Характеристика	Генеральная совокупность	Выборочная совокупность
Объём совокупности	N	n
Кол. единиц, обладающих заданным признаком	M	m
Доля единиц, обладающих заданным признаком	$p = \frac{M}{N}$	$W = \frac{m}{n}$
Среднее значение	$\bar{x} = \frac{\sum x_i}{N}$	$\tilde{x} = \frac{\sum x_i}{n}$
Дисперсия	$\sigma_{\bar{x}}^2 = \frac{\sum (x_i - \bar{x})^2}{N}$	$\sigma_{\tilde{x}}^2 = \frac{\sum (x_i - \tilde{x})^2}{n}$
Среднее квадратическое отклонение	$\sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}$	$\sigma_{\tilde{x}} = \sqrt{\frac{\sum (x_i - \tilde{x})^2}{n}}$

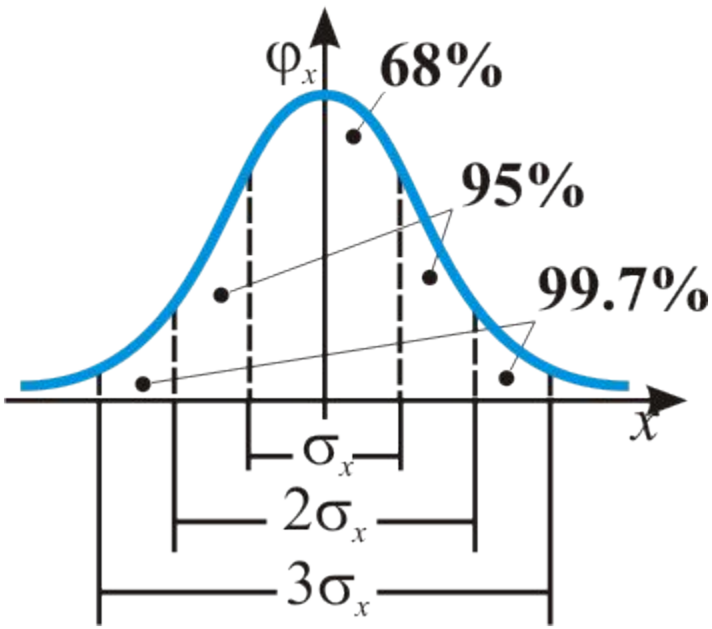
Основные характеристики генеральной и выборочной совокупности

Характеристика	Генеральная совокупность	Выборочная совокупность
Дисперсия доли	$\sigma_p^2 = pq$	$\sigma_W^2 = W(1-W)$
Среднее квадратическое отклонение доли	$\sigma_p = \sqrt{pq}$	$\sigma_W = \sqrt{W(1-W)}$

Определение необходимой численности выборочной совокупности

Допущение, принимаемое при собственно случайном отборе:

Объекты изучаемой совокупности подчиняются нормальному закону распределения случайной величины.



Правило трёх сигм:

$$P(|X-a|>3\sigma)=1-P(|X-a|\leq 3\sigma)=1-0.997=0.003$$

Определение необходимой численности выборочной совокупности

Предельная ошибка выборки

Объём выборки

$$\Delta = t\sqrt{\frac{\sigma^2}{n}} \quad \Rightarrow \quad \Delta^2 = \frac{t^2\sigma^2}{n} \quad \Rightarrow \quad n = \frac{t^2\sigma^2}{\Delta^2}$$

Определение необходимой численности выборочной совокупности

Пример. Для определения среднего возраста 50 тыс. человек, совершивших экономические преступления в России, необходимо провести выборочное обследование методом механического отбора. При проведении предыдущего подобного обследования величина дисперсии составила $\sigma^2_{\tilde{x}}$ = 75. Определите необходимую численность выборки, чтобы с вероятностью 0.997 предельная ошибка выборки не превышала бы $\Delta_x \leq$ 2.5 года.

Определение необходимой численности выборочной совокупности

Решение. Для данной численности N = 50 000 чел. и величины доверительной вероятности α = 0.997, имеем t = 3. Используя формулу для определения необходимой численности выборки, получаем:

$$n_x = \frac{t^2 \sigma_{\tilde{x}}^2}{\Delta_x^2} = \frac{3^2 \cdot 75}{2.5^2} = 108.$$

Ответ. Необходимая численность выборки составляет 108 чел.