Исследование ряда динамики

18 октября 2021 г. 11:52

Основные определения

Статистический ряд – определенным образом упорядоченная последовательность статистических признаков

Рядом динамики - последовательность упорядоченных во времени значений одноименного статистического показателя.

Классификация статистических признаков

Характер выражения	Характер вариации	**Отношение ко времени**	Характер взаимосвязи
Атрибутивные Количественные	Альтернативные Дискретные Непрерывные	**Моментные** **Интервальные**	Факторные Результативные

- *моментные* признаки - характеризуют единицы совокупности на критический момент времени;
- *интервальные* признаки - характеризуют явление за определённый временной период (год, квартал, месяц и т.д.).

Виды рядов динамика

В зависимости от формы представления показателя времени:

- *Моментные ряды* - отображают состояние изучаемого явления на критический момент времени (начало месяца, квартала, года и т. д.);
- *Интервальные ряды* – отображают состояние изучаемого явления за определение интервалы времени (сутки, месяц, квартал и т. д.).

Назначение рядов динамики

- определение системы характеристик динамического ряда,
- разложение динамического ряда на компоненты,
- выявление основной тенденции развития явления – выявление статистической закономерности,
- прогнозирование на основе выделенного тренда,
- выявление и оценка сезонности.

Элементы рядов динамики

- показателя времени t_i
- соответствующий уровень развития явления y_i , $i = \overline{1; n}$

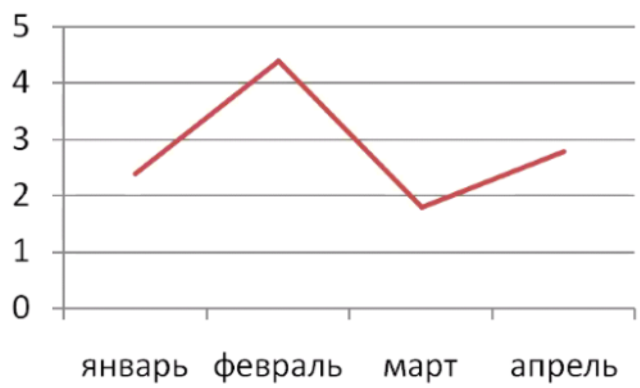
Уровень ряда динамики - численное значение статистического показателя в определенный момент (за определенный период) времени.

Описание ряда динамики

1. Табличное

t_i	t_1	t_2	...	t_n
y_i	y_1	y_2		y_n

2. Графическое



3. Аналитическое

- подбор математического аналога, построение тренда

Этапы исследования ряда динамики:

1. Предварительная обработка ряда динамики – определение системы динамических характеристик ряда
2. Разложение ряда динамики на компоненты:

$$\vec{T} - \textit{тренд}; \vec{S} - \textit{сезонная}; \vec{R} - \textit{периодическая}; \vec{\varepsilon} - \textit{случайная}$$

3. Прогнозирование
4. Выводы

Система характеристик ряда динамики включает:

- ***индивидуальные (частные) характеристики интенсивности:***

- абсолютный прирост ;
- темп роста (коэффициент роста);
- темп прироста (коэффициент прироста);
- абсолютное значение одного процента прироста .

- ***сводные (обобщающие) характеристик интенсивности:***

- Средний уровень ряда;
- Средний абсолютный прирост;
- Средний коэффициент роста;
- Средний коэффициент прироста;
- Дисперсия;
- Среднеквадратичное отклонение;
- Коэффициент вариации.

Абсолютный прирост	цепные	$\Delta y_i^{ц} = y_i - y_{i-1}, \quad i = \overline{2;n}$	$\sum\limits_{i=2}^{n} \Delta y_i^{ц} = \Delta y_n^{б}$
	базисные	$\Delta y_i^{б} = y_i - y_1, \quad i = \overline{2;n}$	
Коэффициент роста	цепные	$K_i^{ц} = \dfrac{y_i}{y_{i-1}}, i = \overline{2;n}$	$\prod\limits_{i=2}^{n} K_i^{ц} = K_n^{б}$
	базисные	$K_i^{б} = \dfrac{y_i}{y_1}, i = \overline{2;n}$	
Коэффициент прироста	цепные	$\Delta K_i^{ц} = \dfrac{\Delta y_i^{ц}}{y_{i-1}}, i = \overline{2;n}$	$\Delta K_i = K_i - 1$
	базисные	$\Delta K_i^{б} = \dfrac{\Delta y_i^{б}}{y_1}, i = \overline{2;n}$	
Абсолютное значение 1% прироста – показывает, какое абсолютное значение соответствует одному проценту прироста		$A_i = \dfrac{\Delta y_i^{ц}}{T_i^{ц}} = \dfrac{y_i - y_{i-1}}{\dfrac{y_i - y_{i-1}}{y_{i-1}} 100\%} = \dfrac{y_{i-1}}{100\%}$	

Индивидуальные показатели ряда динамики:

Базисные – индивидуальные показатели динамики, для расчет которых используется постоянная база сравнения, т.е. каждый уровень ряда сравнивается с одним и тем же базисным уровнем.

Цепные - индивидуальные показатели динамики, для расчет которых используется переменная база сравнения, т.е. каждый последующий уровень ряда сравнивается с предыдущим.

Таблица 1: Индивидуальные динамические показатели ряда

Хронологический показатель времени, t_i	Уровень ряда, y_i	Абсолютный прирост		Коэффициент роста		Коэффициент прироста		Абсолютное значение одного процента прироста A_i
		цепные $\Delta y_i^{ц}$	базисные $\Delta y_i^{б}$	цепные $K_i^{ц}$	базисные $K_i^{б}$	цепные $\Delta K_i^{ц}$	базисные $\Delta K_i^{б}$	

Обобщающие показатели ряда динамики

Средний уровень ряда	$\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$ Интервальный; $\bar{y} = \dfrac{1/2\,y_1 + y_2 + \ldots + y_{n-1} + 1/2\,y_n}{n-1}$ моментный
Средний абсолютный прирост	$\overline{\Delta} y = \dfrac{\sum \Delta y_{\text{ц}i}}{n-1} = \dfrac{\Delta y_{\text{б}n}}{n-1}$
Средний коэффициент роста	$\overline{\text{К}} = \sqrt[n-1]{\prod_{i=2}^{n} \text{К}_{\text{ц}i}} = \sqrt[n-1]{\text{К}_{\text{б}n}}$
Средний коэффициент прироста	$\Delta\overline{\text{К}} = \overline{\text{К}} - 1$
Дисперсия	$\sigma_y^2 = \dfrac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}$
Среднеквадратичное отклонение	$\sigma_y = \sqrt{\sigma_y^2}$
Коэфф. вариации	$V_y = \dfrac{\sigma_y}{\bar{y}}$

Обобщающие показатели ряда динамики

Средний уровень ряда	характеризует наиболее типичную величину уровней, центр ряда.
Средний абсолютный прирост	показывает, на сколько в среднем за единицу времени изменяется уровень ряда,
Средний коэффициент роста	показывает, во сколько раз в среднем за единицу времени изменился уровень динамического ряда, используется в качестве показателя развития явления за длительный период времени
Средний коэффициент прироста	характеризует среднюю относительную скорость изменения уровней в единицу времени
Дисперсия, Среднеквадратичное отклонение, Коэфф. вариации	используются для оценки уровня вариации уровней.

Методы выявления тренда

Тренд **-** плавное и устойчивое изменение уровней явления во времени, свободное от случайных колебаний.

Этапы изучения тренда:

1. проверка ряда на наличие тренда;

- Метод средних - ряд динамики разбивается на два интервала, для каждого из которых определяется средняя величина, выдвигается гипотеза о существенном различии рассчитанных средних.
- Графический метод- используется графическое представление ряда динамики и его визуальный анализ, позволяющий подтвердить наличие или отсутствие тренда.

2. выравнивание ряда динамики и непосредственное выделение тренда.

- Методы механического выравнивания:

метод укрупнения интервалов;

метод скользящей средней и др.

- Метод аналитического выравнивания

Характеристика метода укрупнения интервалов:

- Ряд динамики, состоящий из мелких интервалов, заменяется рядом, состоящим из более крупных интервалов;
- При укрупнении интервалов влияние разнонаправленных факторов нивелируется, и основная тенденция проявляется более отчетливо;
- **Расчет среднего значения уровня по укрупненному интервалу осуществляется по формуле простой средней арифметической.**

Недостаток метода: кратное сокращение числа уровней исходного ряда

В л/р выявить основную тенденцию методом укрупнения интервалов. Исходные интервалы увеличить в 3 раза, т.е. период укрупнения $\alpha = 3$

Таблица определения средних по методу укрупнения интервалов и по методу скользящих средних

<table>
<tr><td rowspan="2">Хронологический показатель времени, ti</td><td rowspan="2">Уровень ряда, yi</td><td colspan="3">Метод укрупнения интервалов</td><td colspan="2">Метод скользящей средней</td></tr>
<tr><td>Новый показатель времени, t_i</td><td>Сумма уровней за интервал</td><td>Среднее значение уровня</td><td>Скользящая сумма</td><td>Скользящая средняя</td></tr>
<tr><td>1</td><td></td><td rowspan="3">I</td><td rowspan="3"></td><td rowspan="3"></td><td>-</td><td>-</td></tr>
<tr><td>2</td><td></td><td></td><td></td></tr>
<tr><td>3</td><td></td><td></td><td></td></tr>
<tr><td>4</td><td></td><td rowspan="3">II</td><td rowspan="3"></td><td rowspan="3"></td><td></td><td></td></tr>
<tr><td>5</td><td></td><td></td><td></td></tr>
<tr><td>6</td><td></td><td></td><td></td></tr>
<tr><td>…</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>n</td><td></td><td></td><td></td><td></td><td>-</td><td>-</td></tr>
</table>

Заполнить таблицу, выявленные тенденции представить графически

Характеристика метода скользящей средней

исходный ряд динамики ***заменяется*** теоретическим, уровни которого рассчитываются по формуле ***скользящей средней***:

$$\hat{y}_i = \frac{y_{i-1} + y_i + y_{i+1}}{3}, \quad i = 2, n-1 \quad при \quad \alpha = 3$$

α -число уровней, по которым укрупняется интервал, называется диапазоном укрупнения, интервалом или периодом сглаживания .

- *При нечетном периоде сглаживания* (3, 5 и т.д.) полученное среднее значение уровня закрепляется за серединой расчетного интервала.
- *При четном периоде сглаживания* возникает проблема центрирования, для решения которой необходимо осуществить сдвиг сглаженных уровней.

недостатки метода:

- первые и последние уровни ряда теряются (не сглаживаются),
- метод применим лишь для рядов, имеющих линейную тенденцию

Виды временных функций :

- *<u>линейная</u>* $\hat{y}=a+b\cdot t$ *- при постоянных абсолютных цепных приростах* $\Delta y_i^{ц} \approx const$

b *- коэффициент регрессии (показывает, насколько в среднем изменится уровень ряда при изменении времени на единицу).*

- *<u>параболическая</u>* $\hat{y}=a+b\cdot t+c\cdot t^2$ *- для рядов динамики, в которых меняется направление развития*

с – коэффициент регрессии (характеризует изменение интенсивности развития в единицу времени)

- *<u>показательная функция</u>* $\hat{y}=a\cdot b^t$ *- для рядов динамики со стабильными цепными темпами роста* $T_i^{ц} = const$

b *- коэффициент регрессии* (интерпретируется как средний коэффициент изменения уровней изучаемого явления).

- и др.

Характеристика методов аналитического выравнивания

Основная тенденция развития рассчитывается как ***временная функция*** $\hat{y}_i = f(t_i^{усл.})$, где $\hat{y}_i$ - теоретические уровни (уровни динамического ряда, вычисленные по соответствующему аналитическому уравнению на момент времени t_i)

Цель аналитического выравнивания динамического ряда: определение аналитической или графической зависимости $f(t_i)$

Подбор функции и нахождение статистических оценок параметров осуществляется ***методом наименьших квадратов (МНК)***
в соответствии с которым наилучшим образом тренд описывает временная функция, обеспечивающая минимальную величину суммы квадратов отклонений эмпирических уровней ряда от соответствующих уровней теоретического ряда: $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \rightarrow \min$

где y_i - фактические уровни;
$\hat{y}_i$ - выровненные по временной функции уровни ряда.

Решение системы уравнений для нахождения параметров

Параметры полинома первой степени $\hat{y} = a + b \cdot t$

$$\begin{cases} n \cdot a + b\sum t = \sum y, \\ a \cdot \sum t + b\sum t^2 = \sum yt \end{cases}$$

Параметры полинома второй степени $\hat{y} = a + b \cdot t + c \cdot t^2$

$$\begin{cases} n \cdot a + b \cdot \sum t + c \cdot \sum t^2 = \sum y, \\ a \cdot \sum t + b \cdot \sum t^2 + c \cdot \sum t^3 = \sum yt, \\ a \cdot \sum t^2 + b \cdot \sum t^3 + c \cdot \sum t^4 = \sum yt^2. \end{cases}$$

Параметры показательной функции $\hat{y} = a \cdot b^t$

$$\ln \hat{y} = \ln ab^t = \ln a + t \cdot \ln b = A + Bt$$

$$\ln a = A$$

$$\ln b = B$$

Метод отсчета от условного нуля

Для получения статистических оценок параметров временных функций необходимо решить проблему преобразования хронологических показателей времени в условные показатели.

Метод отсчета от условного нуля:

начало координат по признаку времени переносится в середину ряда так, чтобы $\sum_{i=1}^{n} t_i^{усл.} = 0$

Т.о. **хронологические показатели времени заменяются условными (числовыми аналогами)**

Таблица: Трендовая модель, построенная на основании полинома 1-ой степени (2-ой степени, показательной функции)

Хронологический показатель, t_i	Условный показатель времени, $t_i^{усл}$	Уровень ряда, y_i	Теоретический уровень ряда, $\hat{y}_i$

Заполнить таблицу, выявленную тенденцию представить графически

$$\sum_{i=1}^{n} t_i^{усл} = 0$$

Оценка надежности трендовой модели
критерий Фишера (F-критерий)

Если $F_{\text{факт}} > F_{\text{кр}}$ трендовая модель признается надежной

$$F_{\text{факт}} = \frac{\eta_{\text{т}}^2}{1-\eta_{\text{т}}^2} \cdot \frac{n-m}{m-1} \qquad \eta_{\text{т}}^2 = 1 - \frac{\sigma_{y-\hat{y}}^2}{\sigma_y^2}$$

$\eta_{\text{т}}^2$ - теоретический коэффициент детерминации,

m - количество параметров трендового уравнения

$$\sigma_{y-\hat{y}}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}$$ - остаточная дисперсия уровней

Критическое значение F-критерия выбирается по таблицам распределения Фишера- Снедекора (F-распределения)

при числе степеней свободы $V_1 = m$-1 и $V_2 = n$-m

и уровне значимости α = 0,05

Значения критерия Фишера (F-критерия) для уровня значимости q =5%

ν_1 – число степеней свободы большей дисперсии; ν_2 – число степеней свободы меньшей дисперсии

ν_2	ν_1											
	1	2	3	4	5	6	7	8	9	10	15	20
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	245,9	248,0
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43	19,45
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12

17

Статистический прогноз - статистические оценки состояния явления в будущих периодах.

Экстраполяция - статистическое прогнозирование основанное на предположении, что закономерность развития, основная тенденция, действующая в прошлом, сохранится и в будущем.

Точность прогноза зависит от сроков прогнозирования: чем они короче, тем надежнее результат экстраполяции

Делаем прогноз на l=1 и l=3 года.

Виды прогноза

1. *Точечный прогноз* - конкретное численное значение уровня в прогнозируемый период (момент) времени. $\hat{y}_{n+l} = f(t_i) \qquad i = n + l$

2. *Интервальный прогноз* – диапазон численных значений, предположительно содержащий прогнозируемое значение уровня.

$$\hat{y}_{n+l} = \pm t_\alpha \cdot \sigma_{y_i - \hat{y}_i} \qquad \sigma_{y_i - \hat{y}_i} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-m}}$$

t_α - коэффициент доверия по распределению Стьюдента при $\nu = n - 2; \quad \alpha = 0{,}05$

$\sigma_{y_i - \hat{y}_i}$ - среднеквадратическая ошибка тренда

l - период упреждения прогноза (срок экстраполяции)

n – число уровней исходного ряда

m – число параметров трендового уравнения

Методы экстраполяции (прогнозирования)

– *на основе среднего абсолютного прироста* $\overline{\Delta y}$

– *на основе среднего коэффициента роста* $\overline{K}$

– *на основе аналитического выравнивания ряда*

I. Метод прогнозирования на основе среднего абсолютного прироста применяется в том случае, если уровни ряда динамики изменяются равномерно (линейно).

Прогнозируемое значение уровня определяется по формуле:

$$\hat{y}_{n+l} = y_n + \overline{\Delta y}$$

$\hat{y}_{n+l}$ - экстраполируемый уровень;

y_n - конечный уровень ряда динамики;

l - период упреждения прогноза (срок экстраполяции).

II. Прогнозирование по среднему коэффициенту роста применяется, если общая тенденция характеризуется экспотенциальной кривой. В этом случае экстраполируемый уровень определяется по формуле:

$$\hat{y}_{n+l} = y_n \cdot (\overline{K})^l$$

III. Прогнозирование на основе аналитического выравнивания

Для прогноза используется аналитическое выражение тренда, в модели необходимо продолжить значение условного показателя времени t_n до t_{n+l}

Экстраполяция, в общем виде, может быть представлена зависимостью

$$\hat{y}_{n+l} = f\left(y_i,\ l,\ a_j\right)$$

$\hat{y}_{n+l}$ - прогнозируемое значение уровня;

y_i - текущий уровень ряда динамики;

l - период упреждения прогноза;

a_j - параметры уравнения тренда.

- Рассчитать относительную ошибку точечного прогноза:

$$E_{пр} = \frac{y_{iф} - \hat{y}_i}{\hat{y}_i} \cdot 100\%$$

Ошибка след месяца должна быть больше предыдущего

Характеристика точности прогноза

Интерпретация показателя относительной ошибки E, %

- E<10 - Высокая
- E [10–20) - Хорошая
- E [20–50) - Удовлетворительная
- E>50 - Неудовлетворительна

Интервальный прогноз на основе метода аналитического выравнивания

Интервальный прогноз – диапазон численных значений, предположительно содержащий прогнозируемое значение уровня.

Интервальный прогноз имеет значительные преимущества перед точечным – они учитывает вероятность свершения прогноза

$$\hat{y}_{n+l} = \pm t_{\alpha} \cdot \sigma_{y_i - \hat{y}_i} \qquad \hat{y}_{n+l} - t_{\alpha} \cdot \sigma_{y_i - \hat{y}_i} \le \hat{y}_{n+l} \le \hat{y}_{n+l} + t_{\alpha} \cdot \sigma_{y_i - \hat{y}_i}$$

$$\sigma_{y_i - \hat{y}_i} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-m}}$$

t_{α} - коэффициент доверия по распределению Стьюдента $\nu = n-2; \quad \alpha = 0{,}05$

$\sigma_{y_i - \hat{y}_i}$ -среднеквадратическая ошибка тренда

n – число уровней исходного ряда

m – число параметров трендового уравнения

l - период упреждения прогноза (срок экстраполяции).

Рассчитать точечные и интервальные прогнозы при периодах упреждения *l*=1 и *l*=3.

- Сделать выводы о характере динамики изучаемого явления, качестве построенных прогнозов.