Тема 4: Одномерные временные ряды

Временной ряд – эконометрическая модель, которая строится по временным данным (в отличие от пространственных).

Временной ряд (ряд динамики) – последовательность значений показателя $\boldsymbol{y}$, упорядоченных по значениям переменной t (по времени)

y_t – **уровень** ряда

Величина каждого уровня складывается под влиянием различных факторов, которые можно разбить на **3 группы**:

- факторы, формирующие **тенденцию** ряда
$$f(t)$$
- факторы, формирующие циклические (периодические) **колебания** $c(t)$
- **случайные** факторы ε_t

Как правило, уровень ряда содержит все эти компоненты

Модель временного ряда

- **Аддитивная** модель:

$$y_t = f(t) + c(t) + \varepsilon_t$$

- **Мультипликативная** модель:

$$y_t = f(t) \cdot c(t) \cdot \varepsilon_t$$

Задача: определение наличия и количественная оценка каждой составляющей

Автокорреляция уровней временного ряда

Автокорреляция – корреляционная зависимость между последовательными уровнями ряда

$r_{y_t y_{t-1}} = r_1$ – коэффициент автокорреляции 1-го порядка (**лаг** = 1):

$$r_1 = \frac{\sum_{t=2}^{n} (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^{n} (y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^{n} (y_{t-1} - \bar{y}_2)^2}}$$

$r_{y_t y_{t-2}} = r_2$ – коэффициент автокорреляции

2-го порядка (**лаг** = 2):

$$r_2 = \frac{\sum_{t=3}^{n}(y_t - \bar{y}_3)(y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^{n}(y_t - \bar{y}_3)^2 \cdot \sum_{t=3}^{n}(y_{t-2} - \bar{y}_4)^2}}$$

$$\bar{y}_3 = \frac{1}{n-2}\sum_{t=3}^{n} y_t \qquad \bar{y}_4 = \frac{1}{n-2}\sum_{t=3}^{n} y_{t-2}$$

Совокупность r_k различных порядков k называется **автокорреляционной функцией** временного ряда. Её график – **коррелограмма**.

Пример: y_t – данные о средних расходах на конечное потребление за 8 лет

t	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
y_t	7	8	8	10	11	12	14	16
y_{t-1}	-	7	8	8	10	11	12	14
y_{t-2}	-	-	7	8	8	10	11	12

Определение структуры временного ряда

- Высокое значение r_1 свидетельствует о наличии линейной тенденции. При ***увеличении лага*** связь ослабевает
- Если r_1 – наиболее высокий коэффи-циент, то ряд содержит **только тенден-цию** (линейную)
- Если наиболее высокий коэффициент – r_m, то ряд содержит **циклические колебания** с периодом m

- Если нет статистически значимых коэф-фициентов, то:
  - либо ряд не содержит тенденции и циклических колебаний, т. е. включает только **случайную составляющую** ε_t (стационарный ряд);
  - либо ряд содержит сильную **нелинейную тенденцию**

Моделирование тенденции временного ряда

Метод – **аналитическое выравнивание** (определение функции $f(t)$) с помощью МНК.

Определение типа тенденции:

- построение и визуальный анализ ***графика***
- расчёт и анализ ***показателей динамики***
- расчёт и анализ ***коэффициентов авто-корреляции*** исходных и преобразованных уровней

Наиболее распространённые функции трендов:

- $\hat{y}_t = b_0 + b_1 \cdot t$ $\qquad \hat{y}_t = b_0 + b_1 \cdot t + b_2 \cdot t^2$

- $\hat{y}_t = b_0 + b_1 / t$ $\qquad \hat{y} = b_0 \cdot t^{b_1}$

- $\hat{y}_t = e^{b_0 + b_1 \cdot t}$ $\qquad \hat{y}_t = \dfrac{L}{1 + e^{b_0 + b_1 \cdot t}}$

Если ряд содержит нелинейную тенденцию, то выбор наилучшего уравнения тренда производится методом перебора

на основе критерия максимума скорректированного индекса детерминации R^2_{adj}

(либо минимума стандартной ошибки оценки $S_{ст.}$)

Моделирование ряда с циклическими (сезонными) колебаниями

Тип модели выбирается в зависимости от характера колебаний:

- если амплитуды колебаний примерно одинаковы, используется аддитивная модель временного ряда;
- если амплитуды увеличиваются или уменьшаются, используется мультипли-кативная модель

Алгоритм определения сезонной составляющей:

1. **Выравнивание** исходного ряда методом скользящей средней (y^*) по интервалу, равному периоду колебаний (это устраняет сезонную компоненту $c(t)$)
2. **Расчёт значений $c(t)$**
   - для аддитивной модели: $c(t) = y_t - y_t^*$
   - для мультипликативной модели: $c(t) = y_t / y_t^*$

и их усреднение по годам.

Алгоритм определения сезонной составляющей:

3. **Устранение сезонной компоненты** из исходных данных:
   - для аддитивной модели: $y_t - \overline{c}(t) = f(t) + \varepsilon_t$
   - для мультипликативной модели: $y_t / \overline{c}(t) = f(t) \cdot \varepsilon_t$
4. **Определение тенденции** – расчёт $\hat{f}(t)$
5. **Расчёт прогнозных значений**:

$$\hat{y}_t^{np} = \hat{f}_{np}(t) + \overline{c}(t) \qquad \hat{y}_t^{np} = \hat{f}_{np}(t) \cdot \overline{c}(t)$$

Расчёт и анализ ошибок

- **Абсолютные ошибки:** $\varepsilon_t = y_t - \hat{y}_t$

$$\hat{y}_t = \hat{f}(t) + \overline{c}(t) \quad \text{или} \quad \hat{y}_t = \hat{f}(t) \cdot \overline{c}(t)$$

- **Относительные ошибки:**

$$\varepsilon_t = y_t \Big/ \left(\hat{f}(t) \cdot \overline{c}(t)\right)$$

Отношение суммы квадратов абсолютных ошибок к общей сумме квадратов отклонений характеризует стандартную ошибку оценки аддитивной модели