ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»

Кафедра «Статистики и экономического анализа**»**

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Б1.В.ДВ.06.01 Эконометрика

Направление подготовки 38.03.02 Менеджмент

Профиль подготовки Производственный менеджмент

Форма обучения очная

СОДЕРЖАНИЕ

1. КОНСПЕКТ ЛЕКЦИЙ

Лекция № 1

<u>Тема: «Введение в эконометрику»</u>

1. Вопросы лекции:

1. Эконометрика как наука
2. Предмет эконометрики
3. Задачи эконометрики в области социально-экономических исследований
4. Информационные технологии на базе ПВЭМ в эконометрических исследованиях
5. Классификация переменных в эконометрических исследованиях
6. Понятие спецификации и идентифицируемости модели

2. Краткое содержание вопросов

1. Эконометрика как наука

В журнале «Эконометрика», основанном в 1933 г. Р. Фришем (1895–1973), он дал следующее определение эконометрики: «Эконометрика – это не то же самое, что экономическая статистика. Она не идентична и тому, что мы называем экономической теорией, хотя значительная часть этой теории носит количественный характер. Эконометрика не является синонимом приложений математики к экономике. Как показывает опыт, каждая из трех отправных точек – статистика, экономическая теория и математика - необходимое, но не достаточное условие для понимания количественных соотношений в современной экономической жизни. Это – единство всех трех составляющих. И это единство образует эконометрику».

Таким образом, эконометрика – это наука, которая дает количественное выражение взаимосвязей экономических явлений и процессов. Нельзя утверждать, что достигнуто однозначное определение эконометрики. Так, Э. Маленво придерживался широкого понимания, интерпретируя эконометрику как «любое приложение математики или статистических методов к изучению экономических явлений».

2. Предмет эконометрики

В дословном переводе слово эконометрика означает «экономические измерения». Это очень широкое толкование данного понятия. Как правило, термин эконометрика применяется в более узком смысле. А именно, под эконометрикой понимается раздел науки, изучающий конкретные количественные и качественные взаимосвязи экономических объектов и процессов с помощью математических и статистических методов и моделей (БСЭ).

3. Задачи эконометрики в области социально-экономических исследований

С помощью эконометрики решается очень широкий круг задач. Наиболее общими задачами эконометрики являются:

1) обнаружение и анализ статистических закономерностей в экономике;

2) построение на базе выявленных эмпирических экономических зависимостей эконометрических моделей.

Данные задачи делятся на более конкретные подзадачи, которые можно классифицировать по трём признакам:

1) классификация задач по конечным прикладным целям:

а) прогноз социально-экономических показателей, определяющих состояние и развитие изучаемой системы;

б) моделирование возможных вариантов социально-экономического развития системы для выявления факторов, изменение которых оказывает наиболее мощное влияние на состояние системы в целом;

2) классификация задач по уровню иерархии:

а) задачи, решаемые на макроуровне (страна в целом);

б) задачи, решаемые на мезоуровне (уровень отраслей, регионов);

в) задачи, решаемые на микроуровне (уровень фирмы, семьи, предприятия);

3) классификация задач по профилю изучаемой экономической системы:

а) рынок;

б) инвестиционная, социальная, финансовая политика;

в) ценообразование;

г) распределительные отношения;

д) спрос и потребление;

е) отдельно выделенный комплекс проблем.

4. Информационные технологии на базе ПВЭМ в эконометрических исследованиях

Определение ПППП. Виды ППП. Проблемно-ориентированные ППП. Автоматизации проектирования (или САПР). ППП общего назначения. Офисные ППП. Настольные издательские системы. Системы искусственного интеллекта. Статистический пакет. STATGRAPHICS, SPSS, SYSTAT, BMDP,SAS, CSS, STATISTICA, S-plus, и др. STADIA, ЭВРИСТА, МЕЗОЗАВР, ОЛИМП: Стат-Эксперт, Статистик-Консультант, САНИ, КЛАСС-МАСТЕР и др. Специализированные пакеты. Пакеты общего назначения.

5. Классификация переменных в эконометрических исследованиях

В эконометрических моделях в основном используются данные трёх типов:

1) пространственные данные (cross-sectional data);

2) временные ряды (time-series data);

3) панельные данные (panel data).

Пространственными данными называется совокупность экономической информации, которая характеризует различные объекты, однако полученной за один и тот же период или момент времени.

Пространственные данные являются выборочной совокупностью из некоторой генеральной совокупности. Примером пространственных данных может служить комплекс экономической информации по какому-либо предприятию (численность работников, объём производства, размер основных фондов), объёмах потребления продукции определённого вида, данные о ВВП различных стран в каком-либо конкретном году и т. д.

Временными данными называется совокупность экономической информации, которая характеризует один и тот же объект, но за разные периоды времени.

6. Понятие спецификации и идентифицируемости модели

Помимо выбора спецификации модели не менее важно также правильное описание структуры модели, в частности, для зависимости данных о временном ряде. Значение результативного признака может зависеть не от фактического значения объясняющей переменной, а от значения, которое ожидалось в предыдущем периоде. Тогда, если ожидаемое и фактическое значения тесно связаны, то будет казаться, что между результативным признаком и объясняющей переменной имеется зависимость, хотя в действительности это всего лишь приближение (аппроксимация) и расхождение опять будет связано с наличием остаточного случайного члена.

Лекция № 2

<u>Тема: «Парная линейная регрессия»</u>

1. Вопросы лекции:

1.1. Метод наименьших квадратов для построения линейной модели
1.2. Линейный коэффициент корреляции
1.3. Коэффициент детерминации
1.4. Проверка значимости регрессионной модели, проверка качества параметров уравнения и построение доверительных интервалов коэффициентов регрессии
1.5. Интерпретация уравнения регрессии
1.6. Средняя ошибка аппроксимации

2. Краткое содержание вопросов

1. Метод наименьших квадратов для построения линейной модели

Классический подход к оцениванию параметров линейной регрессии основан на *методе наименьших квадратов* (МНК), разработанный Гауссом.

Метод наименьших квадратов позволяет получить такие оценки параметров a и b, при которых сумма квадратов отклонений фактических значений результативного признака y от расчетных (теоретических) минимальна:

$$S = \sum_{i=1}^{n} \left(y_i - \tilde{y}_i\right)^2 \longrightarrow \min \quad \text{т.е} \quad \sum \varepsilon_i^2 \to \min$$

Другими словами, из множества линий линия регрессии выбирается так, чтобы сумма квадратов расстояний по вертикали между точками и этой линией была бы минимальной.

Для линейной парной зависимости:

$$S = \sum_{i=1}^{n} \left[y_i - (a + b \cdot x)\right]^2 \longrightarrow \min .$$

2. Линейный коэффициент корреляции

Уравнение регрессии всегда дополняется показателем тесноты связи. Для линейной регрессии таким показателем является **линейный коэффициент корреляции**.

Был впервые введен Карлом Пирсоном (1857 - 1936).

В теории разработаны и на практике применяются различные модификации формул расчета:

$$r = b\frac{\sigma_x}{\sigma_y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y},$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{\sum x_i \cdot f_i}{\sum f_i}; \quad \bar{y} = \frac{\sum y_i}{n} = \frac{\sum y_i \cdot f_i}{\sum f_i};$$

$$\overline{x \cdot y} = \frac{\sum x_i \cdot y_i}{n};$$

$$\sigma_x = \sqrt{\frac{\sum\left(x_i - \bar{x}\right)\cdot f_i}{\sum f_i}} = \sqrt{\frac{\sum\left(x_i - \bar{x}\right)^2}{n}} = \sqrt{\frac{\sum x_i^2}{n} - \left(\bar{x}\right)^2}$$

$$\sigma_y = \sqrt{\frac{\sum\left(y_i - \bar{y}\right)\cdot f_i}{\sum f_i}} = \sqrt{\frac{\sum\left(y_i - \bar{y}\right)^2}{n}} = \sqrt{\frac{\sum y_i^2}{n} - \left(\bar{y}\right)^2}$$

Линейный коэффициент корреляции находится в пределах:

$$-1 \le r \le 1$$

Чем ближе r по абсолютной величине к 1, тем теснее связь между признаками.

Знак указывает на направление связи

Если $|r| \ge 0{,}7$, то считают связь сильной;

$0{,}5 \le |r| \le 0{,}7$ – средней тесноты;

$|r| < 0{,}5$ – слабой.

3. Коэффициент детерминации

Квадрат линейного коэффициента корреляции называется ***коэффициентом детерминации***. Этот коэффициент характеризует долю общей вариации результативного признака, которая объяснятся вариацией факторного признака. (Например, рассмотрели группировку производительности труда по квалификации рабочих и получили r^2=45,4%. Фактор квалификации рабочих объясняет 45,4% вариации производительности труда, а неучтенные факторы – 54,6%.).

4. Проверка значимости регрессионной модели, проверка качества параметров уравнения и построение доверительных интервалов коэффициентов регрессии

Оценка значимости параметров уравнения регрессии дается с помощью *t*-критерия Стьюдента. Выдвигается нулевая гипотеза о равенстве оцениваемого параметра нулю.

- <u>*коэффициента регрессии в*</u>

H_0: b=0

1. Стандартная ошибка коэффициента регрессии

$$m_b = \sqrt{\frac{\sum (y - \tilde{y})^2}{\sum (x - \bar{x})^2 \cdot (n-2)}}$$

2. Фактическое значение t-критерия Стьюдента

$$t_b = \frac{b}{m_b}$$

3. Определяется $t_{табл}$ (*α, df = n-2*) при определенном уровне значимости и числе степеней свободы.

$t_b > t_{табл}$, H_0 отклоняется и параметр b неслучайно отличается от нуля и сформировался под влиянием систематически действующего фактора *x* (статистически значим с заданной вероятностью).

$t_b < t_{табл}$ – H_0 принимается и признается случайная природа формирования b, т.е. связь надежно не установлена, статистически не значима.

Доверительный интервал для коэффициента регрессии определяется как $b \pm t \cdot m_b$. Поскольку коэффициент регрессии имеет четкую экономическую интерпретацию, доверительные границы интервала не должны содержать противоречивой информации, в частности включать значение 0 в этот интервал.

5. **Интерпретация уравнения регрессии**

Для прогнозирования возможных значений результативного признака

- автогрегрессионное прогнозирование по тренду и колеблемости;

- факторное прогнозирование, основанное на изучении и количественном измерении взаимосвязи между признаками.

Основным условием прогнозирования на основании регрессионного уравнения является стабильность или, по крайней мере, малая изменчивость других факторов и условий изучаемого процесса, не связанных с ними. Если резко изменится «внешняя среда» протекающего процесса, прежнее уравнение регрессии потеряет свое значение.

Прогнозирование по уравнению регрессии проводится в два этапа:

1. Вычисляется «точечный прогноз».
2. Определяется доверительный интервал с достаточно большой вероятностью (интервальный прогноз).

6. **Средняя ошибка аппроксимации**

Важнейшей характеристикой качества модели является средняя относительная ошибка аппроксимации:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{\widehat{y}_i - y_i}{y_i}\right|\cdot 100$$

$\widehat{y}_i$, y_i - соответственно расчетное и фактическое значения;

n – число наблюдений.

$\overline{A} < 10\text{-}13\%$ - высокая точность модели;

$10\% < \overline{A} < 20\%$ - хорошая точность модели;

$20\% < \overline{A} < 50\%$ - удовлетворительная.

Лекция № 3

Тема: «Нелинейные модели регрессии и линеаризация»

1. Вопросы лекции:

1. Виды нелинейных зависимостей, поддающиеся линеаризации
2. Подбор линеаризирующего преобразования
3. Производственные функции и их анализ

2. Краткое содержание вопросов

1. **Виды нелинейных зависимостей, поддающиеся линеаризации**

Рассмотрим примеры линеаризующих преобразований:

1) Полиномиальная модель: $\widehat{y} = a + b_1x + b_2x^2 + \cdots + b_px^p$.

Соответствующая линейная модель: $\widehat{y} = a + b_1z_1 + b_2z_2 + \cdots + b_pz_p$, где $z_1 = x, z_2 = x^2, \ldots, z_p = x^p$.

2) Гиперболическая модель: $\widehat{y} = a + \frac{b}{x}$.

Соответствующая линейная модель: $\widehat{y} = a + bz$, где $z = \frac{1}{x}$.

3) Логарифмическая модель: $\hat{y} = a + b \cdot \ln x$.

Соответствующая линейная модель: $\hat{y} = a + bz$, где $z = \ln x$.

Следует отметить и недостаток такой замены переменных, связанный с тем, что вектор оценок получается не из условия минимизации суммы квадратов отклонений для исходных переменных, а из условия минимизации суммы квадратов отклонений для преобразованных переменных, что не одно и то же.

2. Подбор линеаризирующего преобразования

Примеры внутренне линейных моделей и их линеаризация:

1) Мультипликативная степенная модель: $\hat{y} = a x_1^{b_1} x_2^{b_2} \dots x_p^{b_p}$.

Линеаризующее преобразование:

$\ln \hat{y} = \ln a + b_1 \ln x_1 + b_2 \ln x_2 + \dots + b_p \ln x_p$

или

$\hat{Y} = A + b_1 z_1 + b_2 z_2 + \dots + b_p z_p$,

где $\hat{Y} = \ln \hat{y}, \quad A = \ln a, \quad z_1 = \ln x_1, \quad z_2 = \ln x_2, \dots, \quad z_p = \ln x_p$.

2) Экспоненциальная модель: $\hat{y} = e^{a + b_1 x_1 + b_2 x_2 + \dots + b_p x_p}$.

Линеаризующее преобразование: $\ln \hat{y} = a + b_1 x_1 + b_2 x_2 + \dots + b_p x_p$.

3) Обратная регрессионная модель: $\hat{y} = \dfrac{k}{a + b_1 x_1 + b_2 x_2 + \dots + b_p x_p}$.

Линеаризующее преобразование: $\dfrac{1}{\hat{y}} = \dfrac{a}{k} + \dfrac{b_1}{k} x_1 + \dfrac{b_2}{k} x_2 + \dots + \dfrac{b_p}{k} x_p$.

К моделям, полученным после проведения линеаризующих преобразований можно применять обычные методы исследования линейной регрессии. Но поскольку в них присутствуют не фактические значения изучаемого показателя, то оценки параметров получаются несколько смещенными. При анализе линеаризуемых функций регрессии, следует особенно тщательно проверять выполнение предпосылок метода наименьших квадратов.

3. Производственные функции и их анализ

Среди класса нелинейных функций, параметры которых без особых затруднений оцениваются МНК, следует назвать хорошо известную в эконометрике равностороннюю гиперболу: $\tilde{y}_x = a + \dfrac{b}{x} + \varepsilon$.

Классическим ее примером является **кривая Филлипса,** характеризующая нелинейное соотношение между нормой безработицы x и процентом прироста заработной платы *y*. Английский экономист А. В. Филлипс, анализируя данные более чем за 100-летний период, установил обратную зависимость процента прироста заработной платы от уровня безработицы.

Лекция № 4

<u>Тема: «Линейная модель множественной регрессии»</u>

1. Вопросы лекции:

1. Частные и парные коэффициенты корреляции

2. Множественный коэффициент корреляции
3. Гетероскедастичность: суть, последствия, методы смягчения
4. Мультиколлинеарность: суть, последствия, методы устранения
5. Обобщенная линейная модель множественной регрессии и обобщенный метод наименьших квадратов

2.Краткое содержание вопросов

1. Частные и парные коэффициенты корреляции

<u>Парный коэффициент корреляции</u> характеризует тесноту линейной зависимости между двумя переменными на фоне действия всех остальных показателей, входящих в модель.

$$r_{jl} = \frac{1/n\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)(x_{il}-\bar{x}_l)}{\sigma_j \cdot \sigma_l},\ -1 \le r_{jl} \le 1 .$$

Если r_{jl} близок к ±1, то связь между переменными сильная, если r_{jl}=0 – линейная связь отсутствует.

$r_{jl} > 0$ – связь между переменными прямая;

$r_{jl} < 0$ – связь обратная.

Для многомерной корреляционной модели важную роль играют частные и множественные коэффициенты корреляции, детерминации (квадраты соответствующих коэффициентов корреляции).

<u>Частный коэффициент корреляции</u> характеризует тесноту линейной связи между двумя переменными при исключении влияния всех остальных признаков, входящих в модель.

Обладает всеми свойствами парного коэффициента корреляции.

Частный коэффициент корреляции k-2-го порядка между признаками x_1 и x_2 при фиксированном воздействии переменных x_3, x_4, …, x_k может быть определен по следующей формуле:

$$r_{x_1x_2}(x_3, x_4, ..., x_k) = r_{12/3,4,...,k} = \frac{-R_{12}}{\sqrt{R_{11}R_{22}}}$$

R_{jl} — алгебраическое дополнение матрицы R к элементу r_{jl}.

$$r_{yx_j}(x_1,...,x_{j-1},x_{j+1},...,x_k) = r_{yx_j/1,...,j-1,j+1,...,k} = \sqrt{1-\frac{1-R^2_{yx_1...x_k}}{R^2_{yx_1...x_{j-1},x_{j+1}...x_k}}},$$

$R^2_{yx_1...x_k}$ - множественный коэффициент детерминации всего комплекса *k* факторов с результатом;

$R^2_{yx_1...x_{j-1},x_{j+1},...x_k}$ - тот же показатель детерминации, но без введения в модель фактора x_j.

Порядок частного коэффициента корреляции определяется количеством факторов, влияние которых исключается. В практических исследованиях предпочтение отдают показателям частной корреляции самого высокого порядка.

Частный коэффициент корреляции между x_1 и x_2 при фиксированном воздействии переменной x_3 рассчитывается по формуле:

$$r_{x_1x_2}(x_3) = \frac{-R_{12}}{\sqrt{R_{11}R_{22}}} = \frac{r_{12}-r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

2. Множественный коэффициент корреляции

<u>*Множественный коэффициент корреляции*</u> характеризует степень линейной связи между одной переменной (результативной) и остальными, входящими в модель, и может быть рассчитан по формуле:

$$r_y = r_y(x_1, x_2, ..., x_k) = \sqrt{1 - \frac{\det R}{R_{yy}}}$$

$$r_y = r_y(x_1, x_2) = \sqrt{1 - \frac{\det R}{R_{yy}}} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{x_1x_2} r_{yx_1} r_{yx_2}}{1 - r_{x_1x_2}^2}}$$

Если $r_y = 1$, то связь между y и двумерной переменной (x_1, x_2) является функциональной, линейной. Если $r_{x_3} = 0$, то; линейной связи нет.

Из формулы r_y следует, что $r_y > | r_{yx1}|$, $r_y > | r_{yx2}|$, $r_y > |r_{yx1|x2}|$, $r_y > | r_{yx2|x1}|$.

3. Гетероскедастичность: суть, последствия, методы смягчения

Допущение о постоянстве дисперсии остатков $\left(D(\varepsilon_i) = \sigma^2\right)$ известно как допущение о *гомоскедастичности.* Если это допущение нарушено и дисперсия остатков не является постоянной, то говорят, что оценки *гетероскедастичны.*

На практике, для каждого i-го наблюдения определяется единственное значение ε_i, но мы говорим об определении дисперсии остатков, т.е. о множестве ε_i для каждого i-го наблюдения. Это объясняется тем, что мы имеем дело с выборочной совокупностью, а априори ε_i могли принимать любые значения на основе некоторых вероятностных распределений.

Гетероскедастичность приводит к тому, что коэффициенты регрессии не являются оценками с минимальной дисперсией, следовательно, они больше не являются наиболее эффективными коэффициентами. Вследствие, выводы, получаемые на основе t и F-статистик, а также интервальные оценки будут ненадежными. Дисперсии и, следовательно, стандартные ошибки этих коэффициентов будут смещенными. Если смещение отрицательно, то оценочные стандартные ошибки будут меньше, чем они должны быть, а критерий проверки — больше чем в реальности. Таким образом, можно сделать вывод, что коэффициент значим, когда он таковым не является. И наоборот если смещение положительно, то оценочные ошибки будут больше чем они должны быть, а критерии проверки — меньше. Значит, возможно ошибочное принятие нулевой гипотезы.

4. Мультиколлинеарность: суть, последствия, методы устранения

Одним из вопросов спецификации модели является отбор факторов, включаемых в модель.

Включение в уравнение множественной регрессии того или иного набора факторов связано, прежде всего, с представлениями исследователя о природе взаимосвязи моделируемого показателя с другими экономическими явлениями.

Факторы, включаемые во множественную регрессию должны отвечать следующим требованиям:

1. Они должны быть количественно измеримы.
2. Отсутствие коллинеарности и мультиколлинеарности факторов.

Одним из основных препятствий эффективного применения множественного регрессионного анализа, является мультиколлинеарность.

Мультиколлинеарность - линейная взаимосвязь двух или нескольких объясняющих переменных.

5. Обобщенная линейная модель множественной регрессии и обобщенный метод наименьших квадратов

Если известна причина (и, соответственно, форма гетероскедастичности), то для ее устранения можно воспользоваться *обобщенным методом наименьших квадратов* (ОМНК).

Предположим, что дисперсия регрессионных остатков связана с некоторой переменной z_t зависимостью вида

$$D(\varepsilon_i) = \sigma^2 z_i^2.$$

В качестве переменной может быть:

1) среднее квадратическое отклонение σ_t (если она известна), в этом случае получают взвешенный метод наименьших квадратов (ВМНК);

2) x_i или $\sqrt{x_i}$, т.е. дисперсия отатков пропорциональна либо x_i либо x_i^2 .

Для того чтобы избавиться от гетероскедастичности, необходимо разделить каждый член регрессионного уравнения

$$\frac{y_i}{z_i} = b_0 \frac{1}{z_i} + b_1 \frac{x_{1i}}{z_i} + b_2 \frac{x_{2i}}{z_i} + \nu_i, \quad (*)$$

где $\nu_i = \dfrac{\varepsilon_i}{z_i}$ случайная ошибка.

Поскольку $D(\varepsilon_i) = \sigma^2 z_i^2$, то $D(\nu_i) = D\left(\dfrac{\varepsilon_t}{z_i}\right) = \dfrac{D(\varepsilon_i)}{z_i^2} = \dfrac{\sigma^2 z_i^2}{z_i^2} = \sigma^2$.

Таким образом, ошибки в уравнении (*) будут гомоскедастичными.

Однако на практике часто не удается с уверенностью определить причину и форму гетероскедастичности. В этом случае можно либо перевести все переменные в логарифмическую форму (однако необходимо помнить, что этот прием неприменим, если переменные модели могут принимать нулевые или отрицательные значения), либо воспользоваться специальными робастными методами оценки.

Лекция № 5

Тема: «Модели стационарных временных рядов»

1. Вопросы лекции:

1. Построение линейной модели по неоднородным регрессионным данным
2. Введение фиктивных переменных в линейную модель регрессии
3. Проверка регрессионной однородности двух групп наблюдений

2.Краткое содержание вопросов

1. Построение линейной модели по неоднородным регрессионным данным

Рассмотрим формальное определение стационарности.

Стохастический процесс Y_t называется *стационарным в сильном смысле (строго стационарным или стационарным в узком смысле)*, если совместное распределение вероятностей всех переменных $y_{t1}, y_{t2}, ..., y_{tn}$ точно то же самое, что и для переменных $y_{t1+\tau}, y_{t2+\tau}, ..., y_{tn+\tau}$.

Под стационарным процессом *в слабом смысле* (в широком смысле) понимается стохастический процесс, для которого среднее и дисперсия независимо от рассматриваемого периода времени имеют постоянное значение, а автокорреляция зависит только от длины лага между рассматриваемыми переменными:

$$\mu(y_t) = \mu(y_{t+\tau}) = \mu;$$
$$D(y_t) = D(y_{t+\tau}) = \mu(y_t - \mu)^2 = \mu(y_{t+\tau} - \mu)^2 = D_0 = const;$$
$$\operatorname{cov}(y_t, y_{t+\tau}) = \mu[(y_t - \mu)(y_{t+\tau} - \mu)] = \operatorname{cov}(\tau).$$

Из этого следует, что автоковариация будет зависеть только от сдвига по времени τ и не будет зависеть от t.

При анализе изменения $\operatorname{cov}(\tau)$ в зависимости от временного сдвига τ принято говорить об автоковариационной функции.

С понятием автоковариационной функции тесно связано понятие автокорреляционной функции (АКФ):

$$\rho(\tau) = \frac{\operatorname{cov}(y_t, y_{t+\tau})}{D(y_t)} = \frac{\operatorname{cov}(\tau)}{D_0}$$

Значения АКФ также характеризуют тесноту статистической связи между уровнями временного ряда, разделенными τ временными тактами. Однако, в отличие от значений автоковариационной функции они не зависят от масштаба измерения уровней временного ряда и подчиняются ограничению: $|\rho(\tau)| \le 1$.

Выборочная оценка коэффициента автокорреляции $r(\tau)$ может быть определена следующим образом:

$$r(\tau) = \frac{\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(y_t - \bar{y})(y_{t+\tau} - \bar{y})}{\frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})},$$

где n - длина временного ряда;
τ - временной сдвиг (лаг);
$\bar{y}$ - среднее значение временного ряда.

Числитель выражения представляет выборочную оценку коэффициента автоковариации.

Для стационарного временного ряда с увеличением τ АКФ должна демонстрировать свойство монотонного убывания по абсолютной величине, т.к. взаимосвязь между уровнями ряда с ростом τ ослабевает. Однако условие может нарушаться для выборочных АКФ.

2. Введение фиктивных переменных в линейную модель регрессии

До сих пор в качестве факторов рассматривались экономические переменные, принимающие количественные значения в некотором интервале. Вместе с тем может оказаться необходимым включить в модель фактор, имеющий два или более качественных уровней. Это могут быть разного рода атрибутивные признаки, такие, например, как профессия, пол, образование, климатические условия, принадлежность к определенному региону. Чтобы ввести такие переменные в регрессионную модель, им должны быть присвоены те или иные цифровые метки, т.е. качественные переменные преобразованы в количественные. Такого вида сконструированные переменные в эконометрике принято называть фиктивными переменными.

3. Проверка регрессионной однородности двух групп наблюдений

Автокорреляционная и частная автокорреляционная функции оказывают существенную помощь в выборе функции, описывающей стационарный процесс.

С помощью ЧАКФ измеряется корреляция между уровнями ряда y_t и $y_{t+\tau}$, разделенными τ временными тактами, при исключении влияния на эту взаимосвязь всех промежуточных уровней ряда $y_{t+1}, y_{t+2}, \ldots, y_{t+\tau-1}$.

Например, коэффициент частной автокорреляции $\rho_ч$ при τ =2 будет определять корреляцию между уровнями временного ряда, разделенными двумя тактами времени, при условии, что значения промежуточных уровней зафиксированы на среднем уровне.

Формула для расчета выборочной оценки частного коэффициента автокорреляции:

$$r_{ij.k} = \frac{r_{ij} - r_{ik} \cdot r_{jk}}{\sqrt{\left(1 - r_{ik}^2\right)\left(1 - r_{jk}^2\right)}}.$$

Например, $r_ч$ 1- го порядка между y_t и y_{t+2} при устранении влияния y_{t+1}:

$$r_ч(2) = r_{02.1} = \frac{r(2) - r(1) \cdot r(1,2)}{\sqrt{\left(1 - r^2(1)\right)\left(1 - r^2(1,2)\right)}},$$

где $r(2)$ - коэффициент корреляции между y_t и y_{t+2};

$r(1)$- коэффициент корреляции между y_t и y_{t+1};

$r(1,2)$- коэффициент корреляции между y_{t+1} и y_{t+2}.

Примером стационарности служит «белый шум».

«*Белым шумом*» называется чисто случайный процесс, т.е. ряд независимых, одинаково распределенных случайных величин ε_t. Главные свойства «белого шума»:

$$\mu(\varepsilon_t) = 0;$$

$$D(\varepsilon_t) = \sigma_0^2 = const;$$

$$\operatorname{cov}(\varepsilon_t, \varepsilon_{t+\tau}) = \begin{cases} \sigma_0^2, \ при \ \tau = 0 \\ 0, \quad при \ \tau \neq 0 \end{cases}$$

В качестве примера белого шума можно привести остатки, рассматриваемые в классической линейной модели.

В практической аналитической работе стационарность временного ряда означает отсутствие:

- тренда;
- систематических изменений дисперсии;
- строго периодичных флуктуаций;
- систематически изменяющихся взаимосвязей между элементами временного ряда.

Лекция № 6

Тема: «Модели нестационарных временных рядов»

1. Вопросы лекции:

1. Основные элементы временного ряда
2. Модели нестационарных временных рядов и их идентификация
3. Модель авторегрессии – проинтегрированного скользящего среднего
4. Модели рядов, содержащих сезонную компоненту

2.Краткое содержание вопросов

1. Основные элементы временного ряда

ВР (ряд динамики, динамический ряд) – это последовательность упорядоченных во времени числовых показателей, характеризующих уровень состояния и изменения изучаемого явления.

ВР состоят из двух элементов:

1. периода времени, за который или по состоянию на который приводятся числовые значения (t);
2. числовых значений того или иного показателя, называемых уровнями ряда (y).

В практике исследования динамики явлений и прогнозирования принято считать, что значения уровней временных рядов могут содержать следующие компоненты (структурообразующие компоненты):

– Тренд (u_t);
– Сезонную компоненту (S_t);
– Циклическую компоненту (V_t);
– Случайную компоненту (ε_t).

2. Модели нестационарных временных рядов и их идентификация

Общий вид моделей: $Y = T + S + E$, $Y = T \cdot S \cdot E$.

Выбор одной из 2-х моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно const, строят аддитивную модель, в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда.

Процесс построения модели включает в себя следующие шаги:

1. Выравнивание исходного ряда методом скользящей средней.
2. Расчет значений сезонной компоненты *S*.
3. Устранение сезонной компоненты из исходных уравнений ряда и получение выровненных данных ($T+E$) в аддитивной или ($T \cdot E$) в мультипликативной модели.
4. Аналитическое выравнивание уровней ($T+E$) или ($T \cdot E$) и расчет значений *T* с использованием полученного уравнения тренда.
5. Расчет полученных по модели значений ($T+E$) или ($T \cdot E$).
6. Расчет абсолютных и/или относительных ошибок.

Если полученные значения ошибок не содержат автокорреляции, или можно заметить исходные уровни ряда и в дальнейшем использовать временной ряд ошибок *E* для анализа взаимосвязи исходного ряда и других временных рядов.

3.Модель авторегрессии – проинтегрированного скользящего среднего

В авторегрессии каждое значение ряда находится в линейной зависимости от предыдущих значений. Если анализируемый динамический процесс зависит от значений, отстоящих на *p* временных лагов назад, то авторегрессионный процесс порядка *p*, т.е. AR (*p*):

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_p y_{t-p} + \varepsilon_t,$$

где ε_t - «белый шум» с $\mu_\varepsilon = 0$.

α_0 - свободный член (часто приравнивается к нулю (опускается)).

Для выполнения условия стационарности все корни характеристического уравнения $1 - \alpha_1 z - \alpha_2 z^2 - ... - \alpha_p z^p = 0$ должны быть по модулю больше 1 и различны, т.е. $|z| > 1$. Если $|z| = 1$, процесс называется процессом *единичного корня* и является нестационарным.

Рассмотрим простейший вариант линейного авторегрессионного процесса – *модель авторегрессии 1-го порядка – AR(1), или марковский процесс.*

Эта модель может быть представлена в виде:

$$y_t = \alpha y_{t-1} + \varepsilon_t,$$

где α - числовой коэффициент, $|\alpha| < 1$,

ε_t - последовательность случайных величин, образующих белый шум.

4. Модели рядов, содержащих сезонную компоненту

Выравнивание исходного ряда методом скользящей средней:

1) Для этого суммируется, уровни ряда последовательно за каждые 4 квартала со сдвигом на 1 момент времени и определяются условные годовые объемы, уровни и т.д.

2) Разделив полученные суммы на 4, найдем скользящие средние. Полученные т.о. выровненные значения уже не содержат сезонной компоненты.

3) Приведем эти значения в соответствии с фактическими моментами времени, для чего найдем средние значения из двух последовательных скользящих средних – центрированные скользящие средние.

Оценки сезонной компоненты найдем как разность между фактическими уровнями ряда и центрированными скользящими средними. Используем эти оценки для расчета значений сезонной компоненты *S*. Для этого найдем средние за каждый квартал оценки сезонной компоненты S_i. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. Например, в аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю.

В мультипликативной модели взаимопогашаемость сезонных воздействий выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна числу периодов в цикле. Например, при сезонных колебаниях число периодов одного цикла (год) равно 4 (4 квартала).

В аддитивной модели: $S_i = \overline{S}_i - K,$ где *K* – корректирующий коэффициент, $K = \sum_{i=1}^{4} S_i / 4$ и тогда $\sum S_i = 0$;

В мультипликативной модели: $S_i = \overline{S}_i \cdot K, i = \overline{1,4}$, $K = 4 \Big/ \sum_{i=1}^{4} S_i$ и тогда $\sum_{i=1}^{4} S_i = 4.$

В аддитивной модели вычитаем ее значения из каждого уровня исходного временного ряда. Получим *T+E =Y-S.*

В мультипликативной модели разделим каждый уровень исходного ряда на соответствующие значения сезонной компоненты. Тем самым получим $T \cdot E = Y \div S$, которые содержат только тенденцию и случайную компоненту.

Определение компоненты *T* данной модели.

Для этого проводится аналитическое выравнивание ряда *T+E* $(T \cdot E)$ с помощью линейного тренда $T = a + bt$.

Для этого посчитаем *a* – const, *b* – коэффициент регрессии, стандартную ошибку коэффициента регрессии R^2, число наблюдений и число степеней свободы. С помощью них определяем значимость регрессии.

Найдем значения уровней ряда по T_i – полученным по теоретической (аналитической) формуле и S_i – значениям сезонной компоненты для соответствующих кварталов.

$$E = Y - (T + S)$$

$$E = Y \div (T \cdot S)$$

Это абсолютные значения (абсолютные ошибки). По аналогии с моделью регрессии для оценки качества построенной модели или для выбора наилучшей модели можно применять сумму квадратов полученных абсолютных ошибок.

$\left(1 - \dfrac{\sum E_i^2}{\sum (y_i - \bar{y})^2}\right) \cdot 100\% -$ это доля факторной дисперсии уровней ряда объясняет полученное количество процентов от общей вариации уровней временного ряда.

Лекция № 7

Тема: «Система линейных одновременных уравнений»

1. Вопросы лекции:

1.1. Структурная и приведенная формы модели систем одновременных уравнений.
1.2. Рекурсивные системы одновременных уравнений.
1.3.Модель спроса – предложения как пример системы одновременных уравнений.
1.4.Основные структурные характеристики моделей.
1.5. Условия идентифицируемости уравнений системы.
1.6.Идентификация рекурсивных систем

2.Краткое содержание вопросов

1. Структурная и приведенная формы модели систем одновременных уравнений

Система совместных, одновременных уравнений (или структурная форма модели) обычно содержит эндогенные и экзогенные переменные.

Эндогенные переменные обозначены в приведенной ранее системе одновременных уравнений как *y.* Это зависимые переменные, число которых равно числу уравнений в системе.

Экзогенные переменные обозначаются обычно как *x.* Это предопределенные переменные, влияющие на эндогенные переменные, но не зависящие от них.

Структурная форма модели позволяет увидеть влияние изменений любой предопределенной переменной на значения эндогенной переменной. Целесообразно в качестве экзогенных переменных выбирать такие переменные, которые могут быть объектом регулирования.

Структурная форма модели в правой части содержит при эндогенных и экзогенных переменных коэффициенты b_l и a_i (b_l -коэффициент при эндогенной переменной, a_i — коэффициент при экзогенной переменной), которые называются *структурные коэффициенты модели.* Все переменные в модели выражены в отклонениях от среднего уровня, т. е. под x подразумевается $x-\bar{x}$, а под y — соответственно $y-\bar{y}$. Поэтому свободный член в каждом уравнении системы отсутствует.

Использование МНК для оценивания структурных коэффициентов модели дает, как принято считать в теории, смещенные и несостоятельные оценки. Поэтому обычно для определения структурных коэффициентов модели структурная форма модели преобразуется в приведенную форму модели.

Приведенная форма модели представляет собой систему линейных функций эндогенных переменных от экзогенных:

$$\begin{cases} y_1 = \delta_{11}x_1 + \delta_{12}x_2 + \ldots + \delta_{1m}x_{m1} \\ y_2 = \delta_{21}x_1 + \delta_{22}x_2 + \ldots + \delta_{2m}x_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = \delta_{n1}x_1 + \delta_{n2}x_2 + \ldots + \delta_{nm}x_m \end{cases}$$

где δ_j — коэффициенты приведенной формы модели.

2.Рекурсивные системы одновременных уравнений

Каждое уравнение системы независимых уравнений может рассматриваться самостоятельно. Для нахождения его параметров используется метод наименьших квадратов. По существу каждое уравнение этой системы является уравнением регрессии. Поскольку никогда нет уверенности, что факторы полностью объясняют зависимые переменные, в уравнениях присутствует свободный член a_{0i}. Так как фактические значения зависимой переменной отличаются от теоретических на величину случайной ошибки, в каждом уравнении присутствует величина случайной ошибки.

Однако если зависимая переменная y включает в каждое последующее уравнение в качестве факторов все зависимые переменные предшествующих уравнений наряду с набором собственно факторов x, то имеем модель в виде *системы рекурсивных уравнений:*

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m + \varepsilon_1 \\ y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m + \varepsilon_2 \\ y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 + \ldots + a_{3m}x_m + \varepsilon_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m + \varepsilon_n \end{cases}$$

3..Модель спроса – предложения как пример системы одновременных уравнений

Многие экономические взаимосвязи допускают моделирование одним уравнением. Однако ряд экономических процессов моделируются не одним, а несколькими уравнениями, содержащими как повторяющиеся, так и собственные переменные. В силу этого возникает необходимость использования систем уравнений. Например, при оценке эффективности производства нельзя руководствоваться только моделью рентабельности. Она должна быть дополнена моделью производительности труда и моделью себестоимости единицы продукции. Особенно эта необходимость возрастает при исследовании процессов на макроуровне. Это связано с тем, что макроэкономические показатели, являясь обобщающими показателями состояния экономики, чаще всего взаимозависимы. Так, расходы на конечное потребление в экономике зависят от валового национального дохода, и вместе с тем, величина ВНД рассматривается как функция инвестиций.

4.Основные структурные характеристики моделей

Проверим каждое уравнение системы на необходимое и достаточное условия идентификации. Для этого заполним следующую таблицу коэффициентов при отсутствующих в первом уравнении переменных, в которой определитель матрицы (detA) коэффициентов отличен от нуля.

Матрица коэффициентов (1)

Уравнение	Переменные	
	x_3	x_4
2 3	a_{23} 0	a_{24} 0

detA=0, а ранг матрицы равен 1. Следовательно, достаточное условие идентификации не выполняется и первое уравнение можно считать не идентифицируемым.

Коэффициенты при отсутствующих во втором уравнении переменных составят.

Матрица коэффициентов (2)

Уравнение	Переменные	
	y_3	x_1
1 3	b_{13} -1	a_{11} a_{31}

Согласно таблице detA≠0, а ранг матрицы равен 2, второе уравнение идентифицируемо.

Таблица коэффициентов при переменных, отсутствующих в третьем уравнении, в которой detA ≠ 0.

Матрица коэффициентов (3)

Уравнение	Переменные	
	x_3	x_4
1 2	0 a_{23}	0 a_{24}

Условие идентификации невыполняется. Уравнение неидентифицируемо. Следовательно, рассматриваемая в целом структурная модель неидентифицируемая.

В эконометрических моделях часто наряду с уравнениями, параметры которых должны быть статистически оценены, используются балансовые тождества переменных, коэффициенты при которых равны ±1. В этом случае, хотя само тождество и не требует проверки на идентификацию, ибо коэффициенты при переменных в тождестве известны, в проверке на идентификацию собственно структурных уравнений системы тождества участвуют.

5. Условия идентифицируемости уравнений системы

Необходимое и достаточное условие идентифицируемости: Уравнение идентифицируемо, тогда и только тогда, когда по отсутствующим в нем переменным (эндогенным и экзогенным) можно из коэффициентов при них в других уравнениях системы получить матрицу, определитель которой не равен нулю, а ранг матрицы не меньше, чем число эндогенных переменных в системе без одного (n-1).

6.Идентификация рекурсивных систем

При переходе от приведенной формы модели к структурной исследователь сталкивается с проблемой идентификации. Под проблемой идентификации понимается возможность численной оценки параметров структурных уравнений по оценкам коэффициентов приведенных уравнений.

С позиции идентифицируемости структурные модели можно подразделить на три вида:

- идентифицируемые;
- неидентифицируемые;
- сверхидентифицируемые.

Модель идентифицируема, если все структурные ее коэффициенты определяются однозначно, единственным образом по коэффициентам приведенной формы модели, т. е. если число параметров структурной модели равно числу параметров приведенной формы модели. В этом случае структурные коэффициенты модели оцениваются через параметры приведенной формы модели и модель идентифицируема.

$$\begin{cases} \hat{y}_1 = b_{12}y_2 + a_{11}x_1 + a_{12}x_2 \\ \hat{y}_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3 \end{cases}$$

Модель неидентифицируема, если число приведенных коэффициентов меньше числа структурных коэффициентов, и в результате структурные коэффициенты не могут быть оценены через коэффициенты приведенной формы модели. Структурно модель в полном виде, содержащая n эндогенных и m предопределенных переменных в каждом уравнении системы, всегда неидентифицируема.

$$\begin{cases} \hat{y}_1 = b_{12}y_2 + a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ \hat{y}_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{cases}$$

Модель сверхидентифицируема, если число приведенных коэффициентов больше числа структурных коэффициентов. В этом случае на основе коэффициентов приведенной формы можно получить два или более значений одного структурного коэффициента.

$$\begin{cases} \hat{y}_1 = b_{12}y_2 + a_{11}x_1 + a_{12}x_2 \\ \hat{y}_2 = b_{21}y_1 + a_{23}x_3 \end{cases}$$

Лекция № 8

Тема: «Идентификация систем одновременных уравнений»

1. Вопросы лекции:

1. Статистическое оценивание неизвестных значений параметров.
2. Двухшаговый метод наименьших квадратов оценивания структурных параметров отдельного уравнения системы.
3. Трехшаговый метод наименьших квадратов одновременного оценивания всех параметров системы уравнений.
4. Другие методы оценивания систем одновременных уравнений.

2.Краткое содержание вопросов

1. Статистическое оценивание неизвестных значений параметров

КМНК применяется в случае точно идентифицируемой структурной модели. Процедура состоит из следующих этапов:

а) структурная модель преобразуется в приведённую форму модели;

б) для каждого уравнения приведённой формы обычным МНК оцениваются β_{ij} коэффициенты;

в) с помощью алгебраических преобразований от коэффициентов приведённой формы переходят - возвращаются к параметрам структурной модели, получая тем самым оценки этих параметров.

Применение КМНК предполагает, что система уравнений содержит в правой части в каждом уравнении как экзогенные, так и эндогенные переменные. Между тем, могут быть системы, в которых в одном или нескольких уравнениях, например, отсутствуют экзогенные переменные. Для такой модели непосредственной получение структурных коэффициентов невозможно. В этом случае, сначала определяется система приведённой формы модели, решаемая обычным МНК, а затем путём алгебраических преобразований переходя к коэффициентам структурной модели.

2. Двухшаговый метод наименьших квадратов оценивания структурных параметров отдельного уравнения системы

Двухшаговый МНК (ДМНК). Если система сверхидентифицируема, то используется ДМНК. Этапы выполнения этой процедуры:

а) на основе приведённой формы модели получают для сверхидентифицируемого уравнения теоретические значения эндогенных переменных, содержащихся в правой части уравнения;

б) далее, подставив их вместо фактических значений, можно применить обычный МНК к структурной форме сверхидентифицируемого уравнения.

В этом методе дважды используется МНК: на первом этапе, при определении приведённой формы модели и нахождении на её основе оценок теоретических значений эндогенной переменной: $\hat{y}_i = \delta_{i1}x_1 + \delta_{i2}x_2 + ... + \delta_{ij}x_j$ и на втором шаге, применительно к структурному сверхидентифицируемому уравнению при определении структурных коэффициентов модели по данным теоретических (расчётных) значений эндогенных и экзогенных переменных

3. Трехшаговый метод наименьших квадратов одновременного оценивания всех параметров системы уравнений

Трехшаговый метод наименьших квадратов применяется для оценки параметров системы одновременных уравнений в целом. Сначала к каждому уравнению применяется двухшаговый метод с целью оценить коэффициенты и погрешности каждого уравнения, а затем построить оценку для ковариационной матрицы погрешностей, После этого для оценивания коэффициентов всей системы применяется обобщенный метод наименьших квадратов.

4. Другие методы оценивания систем одновременных уравнений

1. Методы оценивания параметров

В прикладной статистике используются разнообразные параметрические модели. Термин «параметрический» означает, что вероятностно-статистическая модель полностью описывается конечномерным вектором фиксированной размерности. Причем эта размерность не зависит от объема выборки.

2. Метод моментов

С какой оценки начинать? Одним из наиболее известных и простых в употреблении методов является метод моментов. Название связано с тем, что этот метод опирается на использование выборочных моментов

$$M_{nm} = \frac{1}{n}\sum_{i=1}^{n} x_i^m, \quad m = 1,2,...$$

где $x_1, x_2, \ldots, x_n$ – выборка, т.е. набор независимых одинаково распределенных случайных величин с числовыми значениями.

3. Метод максимального правдоподобия

В работах, предназначенных для первоначального знакомства с математической статистикой, обычно рассматривают оценки максимального правдоподобия (сокращенно ОМП):

$$\theta_0(n) = \theta_0(n; x_1, x_2, \ldots, x_n) = \underset{\theta \in \Theta}{Arg\min} \prod_{i=1}^{n} f(x_i; \theta).$$

Таким образом, сначала строится плотность распределения вероятностей, соответствующая выборке. Поскольку элементы выборки независимы, то эта плотность представляется в виде произведения плотностей для отдельных элементов выборки. Совместная плотность рассматривается в точке, соответствующей наблюденным значениям. Это выражение как функция от параметра (при заданных элементах выборки) называется функцией правдоподобия. Затем тем или иным способом ищется значение параметра, при котором значение совместной плотности максимально. Это и есть оценка максимального правдоподобия.

2. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ЛАБОРАТОРНЫХ РАБОТ

2.1 Лабораторная работа №1

Тема: Парная линейная регрессия

2.1.1 Цель работы: научиться рассчитывать параметры уравнения парной линейной регрессии.

2.1.2 Задачи работы:

1. Рассчитайте оценки параметров уравнения парной линейной регрессии.
2. Оцените тесноту связи между признаками с помощью выборочного коэффициента корреляции. Проверьте значимость коэффициента корреляции ($\alpha = 0{,}05$).
3. Рассчитайте выборочный коэффициент детерминации. Сделайте экономический вывод.
4. Проверьте значимость оценки коэффициента регрессии с помощью критерия Стьюдента при уровне значимости $\alpha = 0{,}05$.
5. Постройте доверительный интервал для коэффициента регрессии. Дайте экономическую интерпретацию.
6. Составьте таблицу дисперсионного анализа.
7. Оцените с помощью F- критерия Фишера значимость уравнения линейной регрессии ($\alpha = 0{,}05$).
8. Рассчитайте выручку организации, если затраты составят 65 тыс.руб. Постройте доверительный интервал для прогнозного значения объясняемой переменной. Сделайте экономический вывод.
9. Рассчитайте средний коэффициент эластичности ($\overline{Э}$). Сделайте экономический вывод.
10. Определить среднюю ошибку аппроксимации.
11. На поле корреляции постройте линию регрессии.

2.1.3 Перечень приборов, материалов, используемых в лабораторной работе:

1. Рабочая тетрадь по «Эконометрике»
2. Excel

2.1.4 Описание (ход) работы:

Имеются данные о выручке сельскохозяйственной организации (у), млн.руб. и затратах на реализацию – (х), тыс.руб.

у	8	5	4,9	4	3,8	3,5	3,8	3,7	3,6	3,5	3,4	3	3
х	5	10	12	15	20	22	25	30	35	36	40	50	60

РЕШЕНИЕ:

1. Оценки параметров уравнения парной линейной регрессии:

Рассчитаем оценки параметров парной линейной регрессии, где х – ______________________________, а у –________________________________.

Параметры уравнения линейной регрессии рассчитываются с использованием системы нормальных уравнений относительно a и b:

$\tilde{y}_x = a + bx$ - уравнение прямой

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

На основе вспомогательной таблицы 1 записываем систему нормальных уравнений:

$$\begin{cases} \\ \\ \end{cases}$$

Решаем ее методом определителей:

$$\Delta = \begin{vmatrix} & \\ & \end{vmatrix} =$$

$$\Delta \grave{a} = \begin{vmatrix} & \\ & \end{vmatrix} =$$

$$\Delta b = \begin{vmatrix} & \\ & \end{vmatrix} =$$

$$a = \frac{\Delta a}{\Delta} =$$

$$b = \frac{\Delta b}{\Delta} =$$

Получаем уравнение регрессии: $\widetilde{y}_x =$

Этот же результат можно получить, используя следующие формулы.

$$b = \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sigma_x^2} \qquad \bar{\tilde{o}} = \qquad \bar{o} = \qquad \overline{\tilde{o}o} =$$

$$\sigma_x^2 = \frac{\sum x^2}{n} - \overline{x}^2 \qquad \sigma_x^2 =$$

$$b \;=$$

$$a = \overline{y} - b\overline{x}$$

$$a =$$

Получаем уравнение регрессии: $\widetilde{y}_x =$

ВЫВОД____

____

____

____

____

Таблица 1 – Вспомогательная таблица для расчета параметров уравнения линейной регрессии

Номер	x	y	xy	x^2	y^2	$\tilde{y}_x$	$y-\tilde{y}_x$	$(y-\tilde{y}_x)^2$	$(y-\bar{y})^2$	$(\tilde{y}_x-\bar{y})^2$	$(x-\bar{x})^2$	$\left\lvert\frac{y-\tilde{y}_x}{y}\right\rvert\cdot 100$
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
Сумма												
Среднее значение						х	х	х	х	х	х	х

2.Оценка тесноты связи между признаками. Проверка значимости коэффициента корреляции (α=0,05)

Линейный коэффициент корреляции: $r_{yx} = \dfrac{\overline{yx} - \bar{y} \cdot \bar{x}}{\sigma_x \cdot \sigma_y}$ или $r_{yx} = b\dfrac{\sigma_x}{\sigma_y}$

$\sigma_x -$ среднеквадратическое отклонение факторного признака х

$$\sigma_x = \sqrt{\bar{x}^2 - (\bar{x})^2}$$

$\sigma_y -$ среднеквадратическое отклонение результативного признака у

$$\sigma_y = \sqrt{\bar{y}^2 - (\bar{y})^2}$$

$\bar{õ} =$ $\bar{ó} =$ $\bar{õ}^2 =$ $\bar{ó}^2 =$

$\sigma_{õ} =$

$\sigma_{ó} =$

$r_{yx} =$

ВЫВОД____

____

____

____

Для проверки гипотезы о не значимости коэффициента корреляции используют статистику критерия: $t_r = \dfrac{r}{m_r}$

где m_r – стандартная ошибка для линейного парного уравнения регрессии

$m_r = \sqrt{\dfrac{1-r^2}{n-2}}$ $m_r =$

$t_r = \dfrac{r}{m_r}$ $t_r =$

ВЫВОД____

____

____

____

3. Расчет коэффициента детерминации

Коэффициент детерминации составит: $r_{yx}{}^2 = (\_\_\_\_)^2 =$

ВЫВОД____

____

____

____

____

4. Проверка значимости оценки коэффициента регрессии с помощью критерия Стьюдента при уровне значимости α=0,05

С помощью МНК получены оценки параметров уравнения регрессии. Чтобы проверить значимость параметров используют статистические методы проверки гипотезы, оценка существенности коэффициента регрессии сводится к проверке следующих гипотез:

H_0: коэффициент регрессии является статистически незначимым, т.е. b=0;

H_1 : коэффициент регрессии статистически значим, т.е. b≠0.

Стандартная ошибка для коэффициента регрессии m_b:

$$m_b = \sqrt{\frac{\sum\left(y-\tilde{y}_x\right)^2}{(n-2)\sum\left(x-\bar{x}\right)^2}}$$

$m_b =$

Далее вычисляем значения t – критерия Стьюдента: $t_b = \frac{|b|}{m_b}$

$t_b =$

ВЫВОД____

5. Определение доверительного интервала для коэффициента регрессии

Предельная ошибка параметра b: $\Delta_b = t_{табл} \cdot m_b$

$\Delta_b =$

Доверительные интервалы: $\gamma_b = b \pm \Delta_b =$, т.е.

___________ $\le b \le$ ____________

Таким образом, коэффициент регрессии может находиться в интервале (____________;____________).

6. Таблица дисперсионного анализа

Таблица 2 – Таблица дисперсионного анализа

Вариация результата	Число степеней свободы	Сумма квадратов отклонений	Дисперсия на одну степень свободы	F - критерий	
				факт.	табл.
Общая					
Факторная					
Остаточная					

7.Оценка с помощью F – критерия Фишера значимости уравнения линейной регрессии

В силу того, что $F_{факт}$=________ $F_{табл}$=_________, гипотеза о случайности различий факторной и остаточной дисперсий ______________________________

8. Рассчитайте выручку организации, если затраты составят 65 тыс.руб. Постройте доверительный интервал для прогнозного значения объясняемой переменной. Сделайте экономический вывод.

Полученные оценки уравнения регрессии позволяют использовать его для прогноза. Если x=65, то точечный прогноз:

$y^p =$

Чтобы получить интервальный прогноз, найдем стандартную ошибку предсказываемого значения цены на программу m_p.

$$m_p = S\sqrt{1+\frac{1}{n}+\frac{\left(x^p-\bar{x}\right)^2}{\sum\left(x-\bar{x}\right)^2}};$$

где $S=\sqrt{\frac{\sum\left(y-\tilde{y}_x\right)^2}{n-2}}$ - стандартная ошибка регрессии.

$$S =$$

$$m_p =$$

Предельная ошибка прогнозируемой составит:

$$\Delta_{y^p} = t_{табл}\cdot m_p =$$

Доверительный интервал прогнозируемой величины составит:

$$y_p =$$

ВЫВОД____

9. Рассчитайте средний коэффициент эластичности ($\bar{Э}$)

Средний коэффициент эластичности для линейной регрессии рассчитывается по формуле:

$$\bar{Э} = b\cdot\frac{\bar{x}}{\bar{y}} =$$

$$\bar{Э} =$$

ВЫВОД____

10. Определить среднюю ошибку аппроксимации

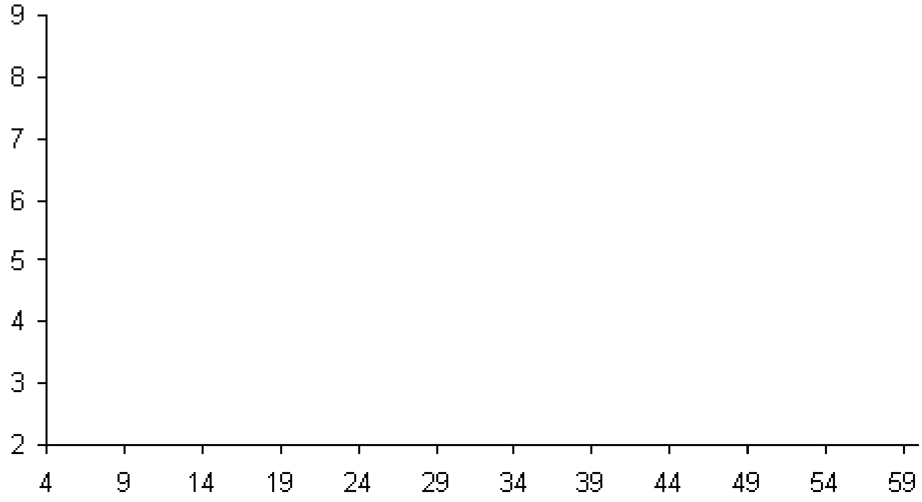
Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$A = \frac{1}{n}\sum\left|\frac{y-\tilde{y}_x}{y}\right|\cdot 100\%$$

$$A =$$

ВЫВОД____
____
____

10. Построение линии регрессии на поле корреляции



2.2 Лабораторная работа №2

Тема: Линейная модель множественной регрессии

2.2.1 Цель работы: научиться рассчитывать параметры уравнения множественной линейной регрессии.

2.2.2 Задачи работы:

1. Построить уравнение множественной линейной регрессии в стандартизованном масштабе и в естественной форме.
2. Рассчитайте частные коэффициенты эластичности.
3. Рассчитать линейные коэффициенты частной корреляции и коэффициент множественной корреляции.
4. Оцените значимость уравнения регрессии в целом с помощью F – критерия Фишера.

2.2.3 Перечень приборов, материалов, используемых в лабораторной работе:

1. Рабочая тетрадь по «Эконометрике»
2. Excel

РЕШЕНИЕ:

1. Построить уравнение множественной линейной регрессии в стандартизованном масштабе и в естественной форме

Линейное уравнение множественной регрессии: $\tilde{y}_x = a + b_1 x_1 + b_2 x_2$

Уравнение в стандартизированном масштабе: $t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2}$

Расчет β – коэффициентов выполним по формулам:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{1 - r_{x_1x_2}^{\ 2}} =$$

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1} r_{x_1x_2}}{1 - r_{x_1x_2}^{\ 2}} = .$$

Получим уравнение $t_y =$ ________ $t_{x_1} -$ ________ t_{x_2} .

Для построения уравнения в естественной форме рассчитаем b_1 и b_2, используя формулы для перехода от β_i к b_i:

$$\beta_i = b_i \frac{\sigma_{x_i}}{\sigma_y}; \quad b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}};$$

$b_1 =$ $b_2 =$

Значение а определим из соотношения:

$$a = \overline{y} - b_1 \cdot \overline{x_1} - b_2 \cdot \overline{x_2} =$$

Получим линейное уравнение множественной регрессии:

$\tilde{y}_{x_1x_2} =$ __________ $-$ __________ __ $x_1 -$ _______ x_2

2. Рассчитайте частные коэффициенты эластичности

Рассчитаем средние коэффициенты эластичности для определения относительной силы влияния x_1 и x_2 на у:

$$\overline{Э}_{yx_j} = b_j \cdot \frac{\overline{x}_j}{\overline{y}}$$

$\overline{\acute{Y}}_{yx_1} =$ $\overline{\acute{Y}}_{yx_2} = .$

ВЫВОД____
____
____
____
____

3. Рассчитать линейные коэффициенты частной корреляции и коэффициент множественной корреляции

Линейные коэффициенты частной корреляции здесь рассчитываются по рекуррентной формуле:

$$r_{yx_1/x_2} = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^{\ 2})(1 - r_{x_1x_2}^{\ 2})}} =$$

$$r_{yx_2/x_1} = \frac{r_{yx_2} - r_{yx_1} r_{x_1x_2}}{\sqrt{(1-r_{yx_1}^{\ 2})(1-r_{x_1x_2}^{\ 2})}} =$$

ВЫВОД____

____

____

Расчет линейного коэффициента множественной корреляции выполним с использованием коэффициентов r_{yx_j} и β_i:

$$R_{yx_1x_2} = \sqrt{r_{yx_1} \cdot \beta_1 + r_{yx_2} \cdot \beta_2} =$$

ВЫВОД____

____

____

4. Оцените значимость уравнения регрессии в целом с помощью F – критерия Фишера

Общий F – критерий проверяет гипотезу H_0 о статистической значимости уравнения регрессии и показателя тесноты связи (R^2=0):

$$F_{\text{факт}} = \frac{R^2_{yx_1x_2}}{1-R^2_{yx_1x_2}} : \frac{m-1}{n-m};$$

$$F_{\text{табл}} = \qquad ; \ \alpha = .$$

ВЫВОД____

____

____

____

____

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ

ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

3.1 Практическое занятие 1 (ПЗ-1) - Введение в эконометрику

3.1.1 Задание для работы:

1. Эконометрика как наука.
2. Предмет эконометрики
3. Задачи эконометрики
4. Особенности эконометрического моделирования

3.1.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике

2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.1.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

1.2 Практическое занятие 2 (ПЗ-2) - Парная линейная регрессия

3.2.1 Задание для работы:

1. Суть метода наименьших квадратов
2. Коэффициенты тесноты связи
3. Проверка значимости уравнения регрессии
4. Интерпретация уравнения регрессии

3.2.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике

2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.2.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

3.3 Практическое занятие 3 (ПЗ-3) - Нелинейные модели регрессии и линеаризация

3.3.1 Задание для работы:

1. Логарифмическая модель
2. Полулогарифмическая модель
3. Обратная модель
4. Степенная модель
5. Показательная модель
6. Выбор формы модели

3.3.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике

2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.3.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

<u>3.4 Практическое занятие 4 (ПЗ-4) - Линейная модель множественной регрессии</u>

3.4.1 Задание для работы:

1. Определение параметров уравнения регрессии
1. Коэффициенты тесноты связи
2. Гетероскадестичность модели
3. Мультиколлинеарность модели

3.4.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике

2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.4.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

<u>3.5 Практическое занятие 5 (ПЗ-5) - Модели стационарных временных рядов</u>

3.5.1 Задание для работы:

1. Понятие стационарного временного ряда
2. Необходимость использования фиктивных переменных

3.5.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике

2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.5.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

<u>3.6 Практическое занятие 6 (ПЗ-6) - Модели нестационарных временных рядов</u>

3.6.1 Задание для работы:

1. Модель авторегрессии – проинтегрированного скользящего среднего
2. Модели рядов, содержащих сезонную компоненту

3.6.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике
2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.6.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

<u>3.7 Практическое занятие 7 (ПЗ-7) - Система линейных одновременных уравнений</u>

3.7.1 Задание для работы:

1. Структурная и приведенная формы модели систем одновременных уравнений.
2. Рекурсивные системы одновременных уравнений.

3.7.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике
2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.7.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.

3.8 <u>Практическое занятие 8 (ПЗ-8) - Идентификация систем одновременных уравнений</u>

3.8.1 Задание для работы:

1. Двухшаговый метод наименьших квадратов оценивания структурных параметров отдельного уравнения системы
2. Трехшаговый метод наименьших квадратов одновременного оценивания всех параметров системы уравнений.
3. Другие методы оценивания систем одновременных уравнений.

3.8.2 Краткое описание проводимого занятия:

1. Работа с рабочей тетрадью по эконометрике
2. С помощью устного опроса и (или) тестирования оценить уровень усвоения студентами изученного материала.

3.8.3 Результаты и выводы:

Усвоение студентами знаний по теме практического занятия.