1. предмет эконометрики

Специфической особенностью деятельности экономиста является работа в условиях недостатка информации и неполноты исходных данных. Анализ такой информации требует специальных методов, которые составляют один из аспектов эконометрики. Центральной проблемой эконометрики является построение эконометрической модели и определение возможности ее использования для описания, анализа, прогнозирования реальных экономических процессов. Эконометрика – быстроразвивающаяся отрасль науки, цель которой состоит в том, чтобы передать количественные меры экономическим отношениям. Термин эконометрика был впервые введен бухгалтером Цьемпой в 1910г. слово «эконометрика» состоит из 2 слов: «экономика» и «метрика». Сам термин подчеркивает специфику науки, т.е. количественное выражение тех связей и отношений, которые раскрыты и обоснованы экономической теорией. Эта наука возникла в результате взаимодействия 3 компонентов: экономической теории, математических методов, статистических методов. В последствии к ним присоединились развитие вычислительной техники. В настоящее время эконометрика располагает огромным разнообразием моделей от больших макроэкономических, включающих несколько сот или тысяч уравнений до малых уравнений, предназначенных для решения специфических проблем.

1. Особенности эконометрического метода

Становление и развитие эконометрического метода на методах вычислительной статистики: - на методах парной и множественной корреляции; - выделение тренда и др. компонентов временного ряда; - на статистическом оценивании.

Потребность в причинном объяснении корреляции привела к созданию путевого анализа, - основан на изучении всей структуры причинной связи между переменными, т.е. на построении графа. Его основным положением является то, что оценки стандартизированных коэффициентов и рекурсивной системы уравнений, которые называются коэффициентами влияния, рассчитываются на основе коэффициентов парной корреляции. При работе с временными рядами разных показателей и при изучении взаимосвязи между ними была осознана проблема ложной корреляции, которая возникла под влиянием фактора ЛАГА, т.е. сдвига во времени. Большое внимание в эконометрики уделяется проблеме данных, т.е. специальным методом работы при наличии данных с пропусками, влияние обобщения данных и т.д. информация может отсутствовать по отд. единицам совокупности и быть на уровни только прежней, информация идет не по отд. организациям, а по районам. Результаты могут сильно отличаться. К проблеме данных относится также проблема селективной выборки в микроэкономике. Типичное направление в этой области: рынок труда; выявление факторов, влияющих на решение о выборке работы; какие экономические стимулы влияют на принятие решения о получении образования.

При этом выборка может быть не случайной, а ограничена какими-то определенными ситуациями, а не всеми возможными. Эффект самоселекции возникает тогда, когда объективный отбор подменяется «удобной выборкой». Эконометрическое исследование включает в себя решение сл. Проблем:

1. качественный анализ связей экономических переменных – выделение зависимых Уi и не зависимые Хк переменных. 2. подбор данных. 3. спецификация моделей связи между переменными. 4. оценка параметров модели. 5. проверка гипотез о свойствах распределения вероятностей для случайных компонентов: гипотезы о средней; дисперсии; ковариации. 6. введение фиктивных переменных. 7. выявление автокорреляции, лагов. 8. выявление тренда, циклической и случайной компоненты. 9. проверка остатков на гетероскедастичность (отсутствия норм распределения для регрессионной функции). 10. анализ структуры связей и построение системы одновременных уравнений. 11. моделирование на основе системы временных рядов. 12. построение рекульсивной модели. 13. проблема и идентификация и оценивания параметров.

Эконометрическая модель основана на диалектическом предположении о круге взаимосвязанных переменных. При всем стремлении к наилучшему описанию связи приоритет отдается качественному анализу.

Этапы эконометрического анализа.

1. построение проблемы. 2. получение данных и анализ их качества. 3. спецификация проблем. 4. оценка параметров. 5. интерпретация результатов.
2. Измерения в эконометрики

Понятие эконометрика включает эконометрические измерения. При этом измерение понимаются по-разному. Признаками измерения считают: получение, сравнение, упорядочивание информации.

Измерение предполагает выделение некоторого свойства, по которым производится сравнивание объекта. Др. понимание измерения исходит из числового выражения результатов, т.е. измерение понимается как операция, в результате которой получается числовое выражение величины, причем числа должны соответствовать наблюдаемым свойствам, качествам, закономерностям науки и т.д

Первый подход связан с наличием эталона, это определение измерения в узком смысле. Первый низший уровень изучения предполагает сравнение объектов по наличию или отсутствию исследуемого свойства. На этом уровне используются термины « нумерация», «классификация», « номинация» и т.д.

Второй уровень предполагает сравнение объектов по интенсивности проявляемых свойств. Здесь используются термины «шкалирование», « топология», и «упорядочивание».

Третий уровень – сравнение объектов с эталонами. Здесь термины « измерение» и «квантификация». Все понятия измерения могут быть объединены на базе определения шкалы измерения. Тип шкалы измерения определяется в допустимом преобразовании- преобразование, при котором сохраняется неизменным отношение между элементами системы. Для определения любой шкалы измерения надо дать название объекту, отождествить объект некоторым свойством или группой свойств. Если это преобразование оказывается единственным, то шкала называется шкалой наименования (номинальной). Измерением в этой шкале можно считать любую классификацию, по которой класс объектов получает наименование. Числа на этой шкале играют роль ярлыков и к ним не применимы правила арифметики. Номинальная шкала обладает только свойствами симметричности и транзитивности.

Симметричность означает, что отношения между градациями Х1 и Х2 сохраняются и между Х2 и Х1.

Транзитивность означает, что если Х1=Х2, а Х2=Х3, то Х1=Х3.

Шкала, в которой порядок элементов по уровню проявления некоторого свойств существенен, а количественное выражение не существенно называется порядковой (ранговой). Шкала порядка допускает операции «=», «≠», «>», «<». Порядковые данные возникают, например при определении предпочтений избирателей, экспертиз качества, при оценке землетрясений, оценке уровня интеллекта. Кроме порядковой и номинальной используется интервальная шкала. Измерение в ней более совершенно, чем в порядковой.

Примером интервальной шкалы могут служить измерения большинства эк. Параметров, т.к. производительность труда, себестоимость, рентабельность и т.д. для измерения эк параметров характерны специфические представления о точности. Точность измерения- его адекватность, т.е. соответствие реальным условиям. Проблема точности связана со сл. Проблемами:

1. определение понятия экономической величины.
2. определение эк.показателей.
3. разработка принципов измерения, конструирования, измерителей.
4. основание выбора типа шкал.
5. разработка правил формирования систем показателей.
6. выявление типов и определение методов устранения ошибок измерений.
7. выявление условий сравнимости эк. Величин.

Основной базой для эконометрических исследований служат данные официальной статистики или б.у.

4. парная регрессия и корреляция эконометрических исследований. спецификация моделей

В зависимости от количества факторов, включаемых в уравнение регрессии принято различать парную (простую) и множественную регрессии.

Парная регрессия- зависимость между 2 переменными Х и У, т.е. модель вида , где у зависимая переменная (результативный признак), х – независимая переменная (факторный признак).

Множественная регрессия – зависимости между 2 и более числом факторов и переменной У, т.е. модель вида: .

Любое эконометрическое исследование начинается со спецификации модели, т.е. формулировки вида модели. При этом парная регрессия достаточная, если имеется доминирующий фактор, который используется в качестве объясняющей переменной Х.

Уравнение парной регрессии характеризует связь между 2 переменными, которая проявляется как некоторая закономерность в целом по совокупности наблюдений. Практически же в каждом отдельном случае величина У складывается из 2 слагаемых , где Уj фактическое значение результативного признака, теоретические значение результативного признака исходя из соответствующей матем. функции, Ej случайная величина, характеризуется отклонением реального значения результативного признака от теоретического, найденного из уравнения регрессии, Е- возмущение и включает в себя влияние неучтенных в модели факторов. Ее присутствие в модели порождено 3 источниками:

1. спецификация модели.
2. выборочный характер исходных данных.
3. особенности измерения переменных.

Основные зависимости, относящиеся к парной регрессии



от правильной спецификации зависит величина случайной ошибки. От тем меньше, чем в большей мере теоретические значения подходят к фактическим данным.

для получения хорошего результата из совокупности обычно исключают единицы с аномальными значениями результативного признака. В парной регрессии выбор вида моделей или математической функции возможен 3 способами:

1. графический. 2. аналитический, т.е. исходя из теории изучаемой связи. 3. экспериментальный

при изучении взаимосвязи между 2 переменными графический способ подбора вида уравнений основан на поликорреляции ( исходные данные, обозначенные на плоскости ХОУ).

Основные типы кривых, используемые при количественной оценке связи между 2 переменными.

Аналитический способ типа уравнений основан на изучении материальной природы связи исследуемых признаков.

Например, потребность предприятия в электроэнергии у зависит от объема выполняемой продукции х и всю потребленную энергию можно разделить на 2 части:

1. несвязанную напрямую с производством продукции (а)
2. связанную непосредственно с объемом выпускаемой продукции, которая возрастает пропорционально увеличению объема выпуска (b).

Связь можно изобразить в виде: у=а+bx. При использовании компьютеров для обработки информации выбор вида уравнения осуществляется экспериментальным способом, т.е. путем сравнения величины остаточной дисперсии Дост, которая вычисляется по формуле: где n количество наблюдений исследуемого признака, у- фактические данные,  - теоретические данные, полученные по уравнению регрессии. Если уравнение проходит через все точки корреляционного поля, то фактическое значение совпадают с теоретическими. Дост=0.

Практически исследование имеет место некоторая рассеянная точка относительно линии регрессии. Это рассеяние обусловлено влиянием изученных моделей факторов. При экспериментальном способе перебираются разные математические функции в автоматическом режиме и из них выбирается та функция, у которой Дост минимально. Если же Д ост оказывается примерно одинаковой для нескольких функций, то предпочтение отдается более простым функциям.

5. линейная регрессия и корреляция: смысл и оценка параметров

Линейная регрессия сводится к нахождению уравнения вида:

Уравнение вида (1) позволяет по заданным значениям фактора Х найти теоретическое значение результативного признака, представляя в уравнение фактическое значение фактора Х. построение линейной регрессии сводится к оценке этих параметров основан на методе наименьших квадратов (МНК) – позволяет получить также оценки параметров а и b при которых сумма квадратов отклонений теоретических значений результативного признака от фактического значения минимальна, т.е

Это означает, что из всех линий регрессии на графике выбирается так, чтобы сумма квадратов между точками и этой линией по вертикали была минимальна.

Чтобы найти минимум функции нужно вычислить частные производные по каждому из неизвестных параметров a и b и приравнять их к нулю.



Решение системы будут следующие уравнения.



Параметр b называется коэффициентом регрессии если а больше 0, то относительное изменение результата У происходит медленнее чем изменение фактора Х. если а меньше нуля, то происходит опережение изменения результата под изменением фактора.

Уравнение регрессии всегда дополняется коэффициентом или показателем тесноты связи.

При использовании линейной регрессии в качестве показателя тесноты связи используется коэффициент корреляции, который обозначается:



Величина коэффициента корреляции находится в пределах единицы

Если b>0 то коэффициент корреляции [-1;0]. Величина линейного коэффициента корреляции оценивает тесноту связи признака Х и У в линейной форме. Но это не означает ,что если коэффициент корреляции равен 0, то между Х и У связи нет. Это означает, что нужно пользоваться др. спецификацией. Для оценки качества подбора линейной функции рассчитывается квадрат линейной корреляции. - коэффициент детерминации. Он обозначает долю депрессии результативного признака У, который объясняется регрессией в общей депрессии результативного признака. Т.е.  1- - величина характеризует долю дисперсии, вызванную влиянием остальных неучтенных в регрессии факторов. Служит одним из критериев для оценки качества линейной модели, т.е. чем больше доля объясненной вариации, тем меньше модель хорошо аппроксимирует исходные данные. Следовательно, можно использовать для прогнозирования результат. Признака.

6. оценка существования параметров линейной регрессии и корреляции

после того, как найдено уравнение регрессии проводится оценка значимости его параметров, а также уравнения в целом. Оценка значимости уравнений проводится с помощью F критерия Фишера. Для этого выдвигается гипотеза Но, которая говорит, что b=0, что при Х не оказывае6т влияние на У. непосредственно расчету критерия предшествует анализ дисперсии. Центральное место в этом анализе занимает разложение общей суммы квадратов на 2 составляющие: объясненную и необъясненную.



первая сумма- общая сумма квадратов отклонений результативного признака от среднего уровня. Вторая сумма – сумма квадратов отклонений, объясненная регрессией (факторная).третья сумма- остаточная сумма отклонений, необъясненная часть.

Если фактор Х не оказывает влияния на результат У, то линия регрессии на графике параллельна ОХ и . это означает что вся дисперсия результативного признака обусловлена воздействием прочих неучтенных регрессией факторов. И тогда общая сумма квадратов отклонений совпадает с остаточной. Если же кр факторы не влияют на результат, то У и Х связаны функционально и остаточная равна нулю. в этом случае общая сумма квадратов отклонений совпадает с суммой квадратов отклонений объясненной регрессией. Т.к. не все точки полекорреляциии лежат на линии регрессии, то всегда имеет место их разброс, вызванный влиянием пр. факторов. Сумма квадратов отклонений связана с числом степеней свободы, т.е. с числом свободы независимого варьирования признака. Число степеней свободы связано с числом единиц совокупности n и числом опр. у ней констант. Существует равенство между степенями свободы общей факторной и остаточной суммы квадратов отклонений. N-1=1+(n-2). Разделив каждую сумму квадратов на соотв. Степени свободы получим средний квадрат отклонений или дисперсию на одну. Степень свободы



Сопоставляя факторную и остаточную дисперсии в расчете на одну степень свободы получаем величину F критерия



После нахождения величины F для определения верности гипотезы Но она сравнивается с табличным значением F-критерия. Fтабличное зависит от соотв. Степени свободы и уровня значимости. Fтабл больше F фактической, то гипотеза Но не может быть отвергнута, т.к. есть риск неправильного вывода о наличии связи. В этом случае уравнение считается статистически незначимым, если выполняется обратное неравенство, то гипотеза Но – отвергается и уравнение считается статистически значимым и надежным. Кроме выяснения значимости уравнения в линейной регрессии оценивается так4же значимость параметров. С этой целью по каждому из параметров вычисляется стандартная ошибка.



S- остаточная сумма квадратов на одну степень свободы или остаточная дисперсия. Величина стандартной ошибки совместна с t- распределением Стьюдента, поэтому для оценки существенности параметра b его величина сравнивается со стандартной ошибкой и вычисляется значение и оно сравнивается см табличным значением t критерия. Выводы такие же как при использовании F критерия. Доверительный интервал для коэффициента регрессии в этом случае определяется следующим образом Стандартная ошибка параметра a

.

процедура оценивания существенности параметра ф аналогично процедуре оценивания параметра b. Значимость линейного коэффициента корреляции проверяется на основе величины ошибки коэффициента корреляции.Фактическое значение t-критерия Стьюдента.



7. интервалы прогнозов по линейному уравнению регрессии

В прогнозах расчета по уравнению регрессии определяется предсказываемое значение Ур при подстановке в уравнение регрессии соотв. Значений Хр=Хк. При подстановке Хр получаем точечный прогноз, который явно нереален. Поэтому он дополняется интервальным прогнозом У\*.



Стандартная ошибка прогноза вычисляется по формуле



Величина стандартной ошибки достигает минимума когда Хк=Х ср и возрастает при удалении от Х ср в любом направлении. При прогнозировании на основе уравнения регрессии следует помнить, что величина прогноза зависит не только от стандартной ошибки но и от точности прогнозного значения фактора Х ( т.е. от Хр). Его величина может задаваться на основе анализа других моделей исходя из конкретной ситуации.

8. корреляция для нелинейной регрессии

Также как и в линейной регрессии используется линейный коэффициент корреляции. В нелинейной регрессии служит индекс корреляции.



Р [0;1]. Чем ближе к 1 том больше связь между рассматриваемыми признаками и тем надежнее уравнение. Коэффициент детерминации- используется для проверки существенности уравнения линейной регрессии по f-критерию Фишера.



где м –число параметров при переменной х, n –число наблюдений.

Число m характеризует число степеней свободы для факторной суммы квадратов. Величина n-m-1 характеризует число степеней свободы для остаточной суммы квадратов. Для степенной функции  для параболы

 y=a+bx+cx2 .

9. нелинейная регрессия

Нелинейная регрессия бывает 2 видов:

- регрессии нелинейные, относительно включаемых в анализ объясняющих переменных, но линейные относительно оцениваемых параметров.

- регрессии нелинейные по параметрам.

Примером первого вида являются полиномы различных степеней:

y=a+bx+cx2, y=a+bx+cx2+dx3, y=a+bx+cx2+dx3+…+zxn.’

Сюда же можно отнести равностепенную гиперболу y=a+b/x. Ко второму типу относятся степенная, показательная, экспоненциальная функции. Полиномы любого порядка сводятся к линейной регрессии с ее методами проверки гипотез и оцениванием параметров. Параболу целесообразно применять для оцениваемого интервала значения фактора X, когда меняется характер среди исследуемых признаков. При этом применяется МНК для оценивания неизвестных параметров a,b,c. В результате получается система из 3 линейных уравнений с 3 неизвестными.



Решение данной системы возможно методом Крамера.

   

Если b>0, c<0, то кривая симметрична относительно высшей точки, т.е. точки перелома кривой, изменяющей направление связи, а именно рост сменяется падением. Такую функцию используют при изучении зависимости з\п работников от возраста. Если b<0, c>0, то кривая симметрична относительно низшей точки и это позволяет определить минимум в точке, меняющей направление связи, а именно падение сменяется ростом.

Такая функция используется при изучении зависимости объема выпуска производства от затрат на производство. Ввиду симметричности кривой второго порядка ее не всегда удобно использовать в конкретных исследованиях, поэтому если на диаграмме рассеивания нет четко выраженной параболы, то использовать ее не нужно, а заменить степенной функцией. Среди нелинейных функций параметры которой можно найти с помощью МНК y-a+b/x можно назвать равносторонней гиперболой. Примером такой функции является кривая Филипса, которая характеризует соотношение между нормой безработицы Х и приростом з\п У. регрессии нелинейные оцениваемые по параметрам делятся также на 2 вида:

- нелинейные модели внутренне линейные.

- Нелинейные модели внутренне нелинейные.

Если модель внутренне линейна, то она с помощью некоторых преобразований может быть приведена к линейному виду. Если модель внутренне нелинейная, то она не может быть приведена к линейному виду. Внутренне линейной можно назвать y=axb. чтобы привести ее к линейному виду нужно ее линеаризовать с помощью логарифмирования.lny=ln(ax)b; lny=Y, lna=c; blnx=X. Далее в помощью МНК находится коэффициенты С и В c=lna; a=ec. к таким моделям относятся показательная, логистическая функции. В экономических исследованиях степенная функция используется для определения коэффициента эластичности.

.

эконометрика регрессия прогноз ошибка

Коэффициенты эластичности для различных математических функций.

|  |  |  |
| --- | --- | --- |
| Вид функции У | Первая производная | Коэффициент эластичности |
| 1. линейная y=a+b\*x | B |  |
| 2. парабола 2 порядка y=a+b\*x+c\*x2 | b+2cx |  |
| 3. гипербола y=a+b\x | -b\x2 |  |
| 4. показательная y=a\*bx | a\*bx\*lnb |  |
| 5. степенная y=a\*xb | a\*b\*xb-1 |  |
| 6. полулогарифмическая y=a+b\*lnx | b/x |  |
| 7.логистическая y=a\(1+b\*e-cx) |  |  |
| 8. обратная y=1/(a+bx) | -b\(a+bx)2 |  |

Если модель внутренне нелинейная, то для оценки параметров используются итеративные процедуры. Решение такого типа задач реализовано в стандартных пакетах прикладных программ.

10. средняя ошибка аппроксимации

Фактические значения результативного признака отличаются от теоретических, рассчитанных по уравнению регрессии. Чем меньше это отличие, тем ближе теоретические значения подходят к эмпирическим данным, тем лучше модель. Величина отклонений фактических данных от теоретических, т.е (y-y^) – ошибка аппроксимации, а т.к. эта величина может быть +\-, то ошибка аппроксимации для каждого наблюдения определяется в % по модулю.

.

Допустимый уровень 8-10%.

11. множественная регрессия и корреляция. спецификация модели

Парная регрессии может дать хороший результат при моделировании, если влиянием др. факторов, воздействующих на объект исследования, можно пренебречь. Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчетах и т.д. основная цель: построить модель с большим числом факторов, определив при этом влияние каждого фактора в отдельности, а также совокупное влияние на моделирование показателей. Множественная регрессия в общем виде можно записать сл. Уравнением:

y=f(x1,x2,…,xn).

12. отбор факторов множественной регрессии

факторы, включаемые в уравнение множественной регрессии должны удовлетворять сл. Требованиям: - должны быть количественно измеримы, не должны быть интеркоррелированы ( т.е. не должны быть связаны друг с другом и тем более не находиться в функциональной зависимости). Если между факторами существует большая корреляция, то нельзя определить их изолированное влияние на результативный признак. Тогда параметры регрессии оказываются не интерпретируемыми. Насыщение модели линейными факторами не снижает величину остаточной дисперсии и не увеличивает показатель детерминации, зато приводит к статистической не значимости уравнения регрессии. И хотя уравнение множественной регрессии позволит получить большое количество факторов, практически необходимости в этом нет, поэтому отбор факторов производится на основе качественного теоретико-экономического анализа. Отбор факторов обычно производится в 2 стадии. На первой выбираются факторы исходя из сущности проблем. На второй стадии на основе матрицы показателей корреляции определяется t- статистики Стьюдента для параметров уравнения регрессии. Это позволяет исключить из модели дублирующие факторы. Считается, что 2 переменные явно коллинеарные, т.е. зависимы, если коэффициент корреляции между ними больше или равен 0,7. 2 переменные дублируют друг друга, поэтому от одной из них необходимо избавиться. В этом случае предпочтение отдается переменной, для которой наблюдается связь с результатом в наименьшей степени. Наибольшие трудности в использовании множественной регрессии возникают при наличии мультиколлинеарности, когда более чем 2 фактора связаны между собой. Включают в модель мультиколлинеарности факторов т.к.: затрудняется интерпретация параметров (теряют смысл); оценки параметров не надежны и обнаруживают большие нестандартные ошибки. Для оценки мультиколлинеарности используется определитель матрицы парных коэффициентов корреляции между факторами. Например, уравнение множественной регрессии имеет вид:

y=a+b1x1+b2x2+b3x3+E.

если rxixi=1, означает что факторы не колленируют между собой. Если между факторами существует полная линейная зависимость, то det R=0. чем он ближе к нулю, тем сильнее мультиколлинеарность факторов и не надежнее уравнение регрессии. Самый простой способ устранения мультиколлинеарности состоит в исключении одного из факторов из модели. Другой подход связан с преобразованием факторов при котором снижается корреляция между ними. Чтобы учесть внутреннюю корреляцию факторов иногда переходят к совмещенным уравнениям.

Y=a+b1x1+b2x2+b3x3+b12x1x3+b23x2x3+E.

Подходы к отбору факторов на основе показателей корреляции различны, что приводит к построению уравнения множественной регрессии разного вида. Наибольшее распространение получили 3 подхода:

1. метод исключения (отсев факторов из полного его набора).

2. метод включения ( дополнительное введение факторов).

3. шаговый регрессионный анализ ( исключение ранее введенного фактора).

13. выбор формы уравнения регрессии

Как и в парной регрессии возможны различные виды: линейные и нелинейные.

Линейные уравнения множественной регрессии имеют вид: y=a+b1x1+b2x2+…+bpxp, где x1,x2,…,xp –факторы, а b1,b2,…,bp- параметры регрессии, b1,…,bp – коэффициенты чистой регрессии. Эти коэффициенты, стоящие перед переменными Х характеризуют средние изменения результативного признака с изменением соответствующего фактора при неизменных значениях др фактора.

 Нелинейные: y=ax1b1x2b2….xpbp степенная множественной регрессии. Параметры bi – коэффициенты эластичности. Они показывают изменении результата с изменением соответствующего фактора на 1% при неизменности др. факторов. Такой вид уравнений множественной регрессии используется в производственных функциях, а также в исследовании спроса и предложения. Для построения множественной регрессии используется также функции:y=e в степени a+b1x1+b2x2+…+bpxp – экспонента.

Y=1\( a+b1x1+b2x2+…+bpxp) – обратная (гипербола).

Стандартные компьютерные программы имеют возможность перебирать возможные функции и выбрать из всех только ту, для которой остаточная дисперсия минимальна и ошибка аппроксимации тоже минимальна. Коэффициент детерминации должен быть приближен к 1. если исследователя не устраивает предполагаемый набор функций регрессии, то можно использовать любые др. функции, приводимые к линейным с помощью преобразования. Однако, чем сложнее функция, тем менее интерпретируемы ее параметры, поэтому использование номинальных моделей очень высокого порядка или сложных функций нежелательно.

1. оценка параметров уравнения множественной регрессии

параметры уравнения множественной регрессии как и для парной регрессии находятся с помощью МНК. При его применении строится система нормальных уравнений, решение которых позволяет получить оценки параметров для уравнения множественной регрессии. Для уравнения множественной регрессии линейного вида получается система нормальных уравнений:



в системе р+1 уравнение и р+1 неизвестная. Решение этой системы возможно методом Крамера. При нелинейной зависимости уравнение множественной регрессии необходимо привести к линейному виду, чтобы затем использовать МНК для нахождения. Например использовать метод линеаризации:

y=ax1b1x2b2….xpbp ; lny= ln(ax1b1x2b2….xpbp); lny= lna+b1lnx1+b2lnx2+…+bplnxp; Y=C+b1X1+b2X2+…+bpXp

1. частные уравнении множественной регрессии

частные линейные уравнения множественной регрессии имеют вид:



Если ввести новое обозначение, то получим



На основе частных уравнений регрессии определяются частные коэффициенты эластичности:



1. множественная корреляция

показатели множественной корреляции характеризуют тесноту связи, рассматриваемого набора фактора с исследуемым признаком, т.е. оценивает тесноту связи совместного влияния фактора на результат. Независимо от формы связи показатель множественной корреляции может быть найден как индекс множественной корреляции:



Индекс множественной корреляции как корень лежит в пределах [0;1]. Чем ближе к 1, тем теснее связь результативного признака со всем набором исследуемых факторов. При правильном включении фактора в уравнение множественной регрессии величина индекса множественной корреляции будет существенно отличаться от индекса корреляции парной регрессии. Если же дополнительное включение фактора второстепенно, то индекс множественной корреляции будет практически совпадать с индексом корреляции парной зависимости. Расчет индекса множественной корреляции предполагает уравнение регрессии и на его основе остаточной дисперсии.можно пользоваться следующей формулой для индекса множественной корреляции:



1. частная корреляция

частные индексы корреляции характеризуют тесноту связи исследуемого признака и одним из факторов при устранении влияния остальных факторов, включенных в модель. Эти показатели представляют собой отношение сокращения остаточной дисперсии за счет включения доп. Факторов. Если рассматриваемая регрессия с числом факторов Р, то возможны коэффициенты корреляции первого, второго и т.д. Р-1 порядков, т.е.

пример: действие влияния Х1 можно оценить при разных условиях независимого действия др. факторов: ryx1x2 при постоянном действии фактора Х2, ryx1x2x3 при постоянном действии факторов Х2 и Х3. формула в общем виде имеет вид:



18. предпосылки МНК

После построения уравнения множественной регрессии проводится проверка наличия у оценок (y=a+b1x1+b2x2+…+bpxp+E) тех свойств, которые предполагаются при МНК. Это связано с тем, что оценки параметров для уравнения регрессии должны отвечать определенным критериям, а именно: д.б. эффективными, несмещенными, состоятельными.

Несмещенность оценки означает, что математическое ожидание остатков равно нулю. Оценка считается эффективной если она характеризуется наименьшей дисперсией. Состоятельность оценок характеризует увеличение их точности с увеличением объема выборки.

Условия, необходимые для получения оценок удовлетворяет этим 3 критериям представляет собой предпосылки МНК:

1. случайный характер остатка.
2. нулевая средняя величина остатков, не зависящая от Xi.
3. гомоскедастичность – дисперсия каждого отклонения одинаково для всех факторов.
4. отсутствие автокорреляции Еi распределены независимо друг от друга.
5. остатки подчиняются нормативному закону.

Если все 5 предпосылок выполняются, то оценки, полученные МНК считаются хорошими. Если не выполняется хотя бы одна предпосылка, то следует корректировать модель.

1).прежде всего проверяется случайный характер остатков Еi. С этой целью строится график зависимости остатков Ei от теоретических значений результативного признака.

А) возможны следующие варианты, если на графике получена горизонтальная полоса, то остатки представляют собой случайные Величины и МНК оправдан, т.е. теоретические значения хорошо аппроксимируют фактические данные.

Б)

остатки неслучайны. В) остатки не имеют постоянной дисперсии.

Г) остатки носят систематический характер. В этом случае отрицательное значение Еi относится к низким значениям y^x, соответственно, положительное значение Ei относится к высоким значениям y^x.

В случаях 2,3,4 необходимо либо применять другую функцию, либо вводить дополнительную информации. А затем строить уравнение регрессии до тех пор, пока остатки не станут случайными величинами.

2) МНК относительно нулевой средней величины остатка означает, сумма разностей фактических и теоретических значений равна нулю . Это выполнимо для линейных моделей и моделей нелинейных относительно включенных переменных. Для выяснения того, что остатки соответственно второй предпосылки строиться график зависимости остатков от факторов включенных в регрессию.

Если на графике получается горизонтальная полоса, то остатки Еi не зависит от Xi. Если график показывает зависимость, то модель неадекватна.

3) в соответствии с 3 предпосылкой МНК требуется, чтобы дисперсия остатков была гомоскедастичной. Это означает, что для каждого фактора Xi остатки Ei имеют одинаковую дисперсию. Если это условие не выполняется, то имеет место гетероскедастичность.

Примеры гетероскедастичности.

А) дисперсия остатков возрастает по мере увеличения Х.

Б) дисперсия остатков достигает максимальной величины при средних значениях величины Х и уменьшения при минимальном значении Х. В)

максимальная дисперсия остатков при таком значении Х и дисперсия однородна по мере увеличения Х.

Для множественной регрессии строится зависимость от Xi и по графику визуально определяется гомоскедастичность.

4) при построении уравнения регрессии важно соблюдать 4 предпосылку, т.е. отсутствие автокорреляции остатков. Автокорреляция остатков означает, наличие корреляции между остатками текущих и предыдущих наблюдений. В этом случае определяется коэффициент корреляции между Ei и E(i+1) как обычный коэффициент корреляции. Если этот коэффициент оказывается отличным от нуля, то остатки автокоррелированы. Выполнение этой предпосылки особенно актуально при построении регрессионных моделей по рядам динамики, где последующие уровни зависят от предыдущих (касается временных рядов).

19. оценка надежности результата множественной регрессии и корреляции

значимость уравнения множественной регрессии в целом как и в парной оценивается с помощью F критерия Фишера



где R2 коэффициент множественной детерминации как квадрат индекса множественной корреляции, n- число наблюдений, m – число параметров при переменной Х. кроме оценивания уравнения во множественной регрессии оценивается также значимость фактора дополнительно включенную в регрессионную модель. Необходимость такой оценки связана с тем, что не каждый фактор, включенный в уравнение множественной регрессии, будет существенной увеличивать долю объясненной регрессии. Ввиду корреляции между факторами значимость одного и того же фактора может быть разной в зависимости от последовательности введения его в модель. Мерой для оценки включенного фактора в модель служит частный F критерий Fxi. В общем виде для фактора Хi частный критерий определяется по формуле:



20. фиктивные переменные во множественной регрессии

при построении уравнения множественной регрессии может оказаться необходимым включение в модель фактора, имеющего 2 и более качественного уровня. Например, это атрибутивные признаки – пол, профессия, образование, климатические условия и т.д. чтобы ввести такие переменные в регрессионную модель им присваиваются цифровые метки, т.е. качественные переменные преобразуются в количественные. Такого вида структурированные переменные называются фиктивные.

Пример, по группе Х м и ж пола изучается линейная зависимость потребления кофе от цены, у- потребление кофе, х – цена.

Y=a+bx; y1=a1+b1x+E1-Mужчины,

y2=a2+b2x+E2-женщины.

Из этих 2 уравнений нужно получить 1 уравнение.

Y=a1z1+a2z2+bx+E Z1= Z2=

В отдельном случае, может оказаться необходимость введения 2 и более фиктивных переменных, тогда модель представляет собой сумму

y=a1z1+a2z2+a2s3+a4s4+bx+E

Фиктивные переменные для оценки сезонных различий потреблений. Фиктивные переменные могут вводиться не только в линейные, но и не в линейные модели, но приводимые к линейным с помощью некоторых преобразований.

21. основные элементы временных рядов

Построить эконометрическую модель можно, используя 2 типа данных:

1. данные, характеризуют совокупность объектов в определенный момент или период времени.
2. данные, характеризующие один объект за несколько последовательных моментов или периодов времени.

Модели, построенные по данным первого типа, называются пространственными моделями.

модели, построенные по данным 2 типа, называются моделями временных рядов.

Временной ряд- совокупность значений какого-либо показателя за несколько моментов или периодов времени.

Каждый уровень временного ряда формируется под воздействием большого числа факторов, которые условно можно разделить на 3 группы:

1. факторы, формирующие тенденцию ряда.
2. фактора, формирующие циклические колебания ряда.
3. случайные факторы.

При различных состояниях изучаемого явления этих факторов зависимость уровня ряда от времени может быть различие. Во-первых, большинство временных рядов экономических показателей имеет тенденцию, характеризующую совокупное долговременное воздействие множества факторов на исследуемый показатель. Во-вторых, изучаемый показатель может быть подвержен циклическим колебаниям. Эти колебания могут носить сезонный характер, т.к. экономическая деятельность ряда отраслей экономики зависит от времени года. Некоторое временные ряды не содержат тенденции и циклические компоненты. А их каждый следующий уровень образуется как сумма следующего уровня ряда и некоторого положительной или отрицательной компоненты. В большинстве случаев фактический уровень временного ряда может представлять собой сумму или произведение трендовой, циклической и случайной компонент.

Модель, в которой временной ряд представлен как сумма перечисленных компонент называется аддитивной. Модель, в которой временной ряд представляет собой произведение 3 компонент называется мультипликативной.

Основные компоненты временного ряда.

Тенденция циклическая случайная

Основная задача эконометрического исследования временных рядов- выявление и предание количественного выражения каждой из перечисленных компонент с тем, чтобы использовать полученную информацию для прогнозирования будущих значений ряда.

22. автокорреляция уровней временного ряда и выявление его структуры

При наличии во временном ряде тенденции и циклических колебаний значения каждого последующего уровня ряда зависит от предыдущего. Корреляционная зависимость между последовательными уровнями временного ряда называют автокорреляцией временного ряда. Количественно она измеряется с помощью линейного коэффициента корреляции между уровнем исходного ряда и уровнями этого же ряда, сдвинутого на несколько шагов во времени. Число периодов, по которым рассчитываются коэффициенты автокорреляции называются лагом. С его увеличением число пар значений по которым рассчитывается коэффициент автокорреляции уменьшается. Последовательность коэффициентов автокорреляции уровней первого, второго и т.д. порядков называют автокорреляционной формулой временного ряда графику зависимости ее значений от величины ряда называется коррелограммой. Если наиболее большим оказался коэффициент автокорреляции первого порядка, то исследуемый ряд содержит только тенденцию. Если наиболее большим оказался коэффициент корреляции порядка тетта, то ряд содержит циклические колебания с периодичностью тетта-моментов времени. Если ни один из коэффициентов автокорреляции не оказался значимым, то можно сделать одно из 2 предположений относительно структуры исследуемого ряда:

-либо он не содержит тенденции и циклических колебаний и имеет структуру, похожую на структуру ряда из рисунка 3 в параграфе 4.1.

- било ряд содержит сильную тенденцию для появления которой нужно провести дополнительный анализ.

23. моделирование тенденции временного ряда

Одни из наиболее распространенных способов моделирования тенденции временного ряда – построение аналитической функции, характеризующей зависимость уровня ряда от времени и тренда. Этот способ называется аналитическим выравниванием временного ряда. Для построения тренда чаще всего используются сл. Функции:



параметры каждого из перечисленных трендов можно определить МНК, используя в качеств независимой переменной время t=1…n, а в качестве зависимой переменной фактические уровни временного ряда. Для нелинейных трендов проводят стандартную процедуру линеаризации. К числу наиболее распространенных способов относятся также качественный анализ исследуемого объекта, а также построение и визуальный анализ графика зависимости уровней временного ряда от времени.

24. моделирование сезонных и циклических колебаний

1. аддитивная и мультипликативная модели временного ряда.

Простейший подход к анализу структуры временного ряда – расчет значений сезонных колебаний методом входящей средней и построение аддитивной модели временного ряда. Общий вид мультипликативной модели: Y=TSE, где T –тренд, S- сезонная компонента, E – случайная компонента. Аддитивная модель: Y=T+S+E. Построение аддитивной и мультипликативной моделей сводится к расчету T,S,E каждого уровня временного ряда.

Процесс построения аддитивной и мультипликативной моделей.

1. выравнивание исходных уровней ряда методом входящей средней.
2. расчет сезонной компоненты S.
3. устроение сезонной компоненты из исходных уровней ряда, получение выровненных данных T+Е для аддитивной и ТЕ для мультипликативной моделей.
4. аналитическое выравнивание уравнений Т+Е и ТЕ и расчет значений тренда Т с использованием полученного уравнения тренда.
5. расчет полученных по модели значений Т+S и TS.
6. расчет абсолютных или относительных ошибок.

Если значение ошибок не содержит автокорреляции, то ими можно исходные уровни временного ряда и в дальнейшем использовать временной ряд ошибок Е.

2. применение фиктивных переменных для моделирования сезонных колебаний.

Еще один способ для моделирования сезонных колебаний – построение уравнения регрессии, с включением фактора времени и фиктивных переменных. При этом количество фиктивных переменных должно быть на 1 меньше числа моментов времени внутри одного цикла колебаний. Каждая фиктивная переменная отражает сезонную компоненту для какого-било одного периода и одна равна 1 для данного периода и 0 для всех остальных периодов. Пусть имеется временной ряд, содержащий циклические колебания с периодичностью 2, тогда модель регрессии с фиктивными переменными для этого ряда будет иметь вид:

y=a+bt+b1x1+b2x2+bjxj+…+b(k-1)x(k-1)+E, где

Xj=

25. применение систем эконометрических уравнений

Под системой эконометрических уравнений обычно понимается система одновременных совместных уравнений. Ее применение имеет ряд сложностей, связанных с ошибками спецификации модели. В виду большого числа факторов, влияющих на экономические переменные исследователь как правило не уверен в точности предлагаемой модели. Кроме того, набор эндогенных и экзогенных переменных, соответствующих теоретическому представлению, исследователя об исследуемом объекте. Это представление сложилось по данным моделей, но может меняться. Соответственно, может меняться и вид модели с точки зрения ее идентифицируемости. Сверхидентифицируемую модель можно превратить в точно идентифицируемую модель путем добавления некоторых переменных или отбрасывания ограничений на некоторые параметры. Не исключено, что при правильной спецификации модель может быть неидентифицируема, поэтому от нее переходят к идентифицируемой, использует сверхидентифицируемую модель. Наиболее широко системы одновременных моделей используются для построения макроэкономических моделей функционирования экономики той или иной страны. Большинство из таких моделей - модели кейнсианского типа.

26. моделирование тенденции временного ряда при наличии структуры изменений. тест Чоу

От сезонных и циклических колебаний следует отличать единовременные изменения характера тенденции временного ряда, которые вызваны структурными изменениями в экономике или иными факторами. В этом случае, с некоторого времени t\* происходят изменения характера динамики изучаемого параметра тренда, описывающего эту динамику. Схематично эта ситуация выглядит сл.образом:

В момент времени t\* сопровождается значительными изменениями ряда факторов, оказывающих сильное воздействие на изучаемый показатель Yt. Чаще всего эти изменения вызваны изменениями общей экологической ситуацией или факторами глобального характера. Если исключить временной ряд включающий в себя соответствующий момент времени, то далее основной является задача, а значит повлияют эти структурные изменения на характер тенденции. Если влияние окажется значимым, то моделирование тенденции данного временного ряда используют кусочно-линейные модели регрессии. Для этого исходная совокупность делится на 2 подсовокупности: до момента t\* и после него, а затем отдельно по каждой совокупности строится отдельное уравнение регрессии. Если структурные изменения незначительно повлияют на характер тенденции, то Yt можно описать с помощью одного уравнения регрессии для всей совокупности данных.

Каждый из этих подходов имеет свои положительные и отрицательные формулы. При построении кусочно-линейной модели происходит снижение остаточной сумы квадратов по сравнению с единым уравнением, но разделение исходной совокупности на 2 части ведет к потери числа наблюдений, а соответственно числа степеней свободы. Построение единого уравнения наоборот позволяет сохранить число наблюдений, но остаточная сумма квадратов будет увеличиваться по сравнению с кусочно-линейной моделью. Поэтому выбор между 2 моделями будет зависеть от соотношения между степенями остаточной дисперсии и потерей числа степеней свободы. Формальный тест для оценки этого соотношения был предложен Грегори Чоу. Применение теста предполагает расчет параметра уравнения трендов, графики которых обозначены (1),(2), (3).

Условные обозначения для алгоритма теста Чоу.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| №уравнения | Вид уравнения | Число набл сов-ти | Остаточная сумма квадратов | Число параметр. в Ур-ии | Число степеней свободы-ост дисп |
|  Кусочно-линейная модель |
| (1)(2) | Y=a1+b1tYa2+b2t | n1n2 | C1 остС2 ост | K1K2 | n1-k1n2-k2 |
|  Уравнение тренда по всей совокупности |
| (3) | Y=a3+b3t | n | С3 ост | К3 | n-k3=(n1+n2)-k3 |

В соответствии с методикой Чоу определятся фактическое значение Fкритерия Стьюдента по следующим дисперсиям на одну степень свободы.

  

Найденное значение Fфакт сравнивается с табличным значением, полученным по таблице распределения Фишера для уровня значимости 2 и степеням свободы (k1+k2)-k3 и n-k1-k2. если Fфакт больше Fтабл то гипотеза о структуре стабильности отклонений и влияние структурных изменений на динамику изучаемого показателя признается значимой. В этом случае моделирование тенденции временного ряда нужно делать с помощью кусочно-линейной модели. Если же фактическая будет меньше табличной, то нет основания отклонить гипотезу о структурной стабильности тенденции, значит ее моделирование можно осуществить для всей совокупности уравнения тренда.

Особенности применения теста Чоу.

1) если число параметров во всех уравнениях (1),(2), (3) одинаково и равно К, то формула для вычисления Fкритерия упрощается



2) тест Чоу позволяет сделать вывод о наличии или отсутствии структурной стабильности в исследуемом временном ряде. Если Fфакт <F табл, означает что уравнения (1), (2)описывают одну и ту же тенденцию, а различие статистических оценок а1 и а2, b1и b2 являются статистически незначимыми. Если Fфакт>Fтабл, то гипотеза о структурной стабильности отклоняется, означает что статистическая значимость различий оценок а1 и а2, b1 и b2.

3) применение теста Чоу предполагает соблюдение предпосылок о нормальном распределении остатков в уравнениях (1), (2).

Возможны следующие сочетания изменения численных оценок параметров уравнений (1),(2)6

- изменение численной оценки свободного члена уравнения тренда а2 по сравнению с а1, при условии, что различия между b1 и b2 незначимы. Геометрически это означает, что прямые (1) и (2) параллельны. Экономически это говорит о многообразном изменении ряда Yt в момент времени t\* при неизменном среднем абсолютном приросте за период.

- изменение численной оценки параметра b2 по сравнению с b1 при условии, что различия между а1 и а2 статистически незначимы. Геометрически это означает, что прямые (1) и (2) пересекают ось ординат в одной точке. Экономически это связано с изменением среднего абсолютного прироста временного ряда начиная с момента времени t\* при неизменном начальном уровне временного ряда при t=0.

- изменение численных оценок а1 и а2, а также b1 и b2. геометрически это означает, что прямые (1) и (2) пересекаются в точке с абсциссой t\*. Экономически можно говорить , что изменение тенденции сопровождается как изменением начального уровня ряда, так и среднего за период абсолютного прироста.

1. общее понятие о системах уравнений, используемых в эконометрике

Объектом статистического изучения являются сложные системы. При использовании отдельных уравнений регрессии с большинстве случаев предполагается, что аргументы или факторы можно измерять независимо друг от друга. На практике изменение одной переменной как правило влечет за собой изменений др, поэтому отд. взятое уравнение множественной регрессии не характеризуют истинные влияния отд. признаков. Система уравнений в эконометрике может быть построена по-разному:

1. возможна система независимых уравнений, когда каждая зависимая переменная у рассматривается как функция одного набора факторов х.



при этом, в каждом уравнении набор факторов xi может варьировать, т.е. система независимых уравнений может считаться и

.

отсутствие того или иного фактора в уравнении системы может быть следствием как экономически нецелесообразного включения его в уравнение, так и не существенности влияния его на результативный признак. Каждое уравнение такой системы может рассматриваться самостоятельно, поэтому для нахождения его параметров используется обычный МНК, т.к. каждое уравнение представляет собой уравнение множественной регрессии.

1. если зависимая переменная у одного уравнения выступает в виде фактора х в др. уравнении, то можно строить систему рекурсивных уравнений.



В данной системе зависимая переменная у включается в каждое последующее уравнение в качестве фактора + полный набор факторов х. как и в предыдущей системе каждое уравнение может рассматриваться самостоятельно, а параметры этих уравнений снижаются с помощью МНК.

3. наиболее распространенной системой эконометрических уравнений является система взаимосвязанных уравнений. В такой системе одни и те же зависимые переменные в одних уравнениях входят в левую часть, в др. в правую часть.



Эта система носит также название системы совместных одновременных уравнений ( структурной формы модели). В отличие от двух предыдущих систем каждое уравнение такой системы не может рассматриваться как самостоятельное, а следовательно, МНК для нее не применим. Для оценки параметров этой системы используется специальные приемы оценивания.

28.структурная и приведенная формы модели

Система одновременных уравнений обычно содержит эндогенные и экзогенные элементы. Эндогенные переменные обозначены в приведенной ранее системе как у. это зависимые переменные, число которых равно числу уравнений в системе. Экзогенные переменные обозначаются обычно х. это предопределенные переменные, влияющие на зависящие от них. Простейшая структурная форма модели имеет вид:

 (1)

классификация переменных на экзо- и эндогенные зависит от теоретической концепции данной модели. Экономические переменные в одних моделях могут выступать как эндогенные переменные, а в др. как экзогенные. Внешнеэкономические переменные (климатические условия) выступают только в качестве экзогенных. Кроме того, в качестве экзогенных переменных могут выступать значения эндогенных переменных за предшествующий период времени. Такие переменные называются лаговыми. Структурная форма модели в правой части содержит коэффициенты при эндогенных переменных aj и bj, которые называются структурными коэффициентами модели. Использование МНК для оценивания структурных коэффициентов дает смещение и несостоятельные оценки, поэтому обычно для определении структуры коэффициентов структурная форма модели преобразуется в приведенную. Приведенная форма модели представляет собой систему линейных функций эндогенных переменных от экзогенных. Приведенная форма модели имеет вид:





По своему виду эта система ничем не отличается от системы независимых уравнений, параметры которой находятся с помощью МНК, поэтому применяя МНК можно найти коэффициент ۟δi, а затем через них оценить значения экзогенных и эндогенных переменных. Коэффициент приведенной формы модели представляет собой функций коэффициентов структурной модели. Рассмотрим это на примере. Для структурной формы модели (1) приведенная форма модели имеет вид:



Выразим у2 из второго уравнения первой системы:



Чтобы найти δ21 и δ22 нужно выразить у1 из второго уравнения первой системы и прировнять к правой части первого уравнения первой системы.

1. проблема идентификации

При приведенной формы модели к структурной исследователь сталкивается с проблемой идентификации. Идентификация – единственность соответствия между приведенной структурной формами модели. С позиции идентифицируемости структурные модели можно подразделить: идентифицируемые, неидентифицируемые, сверхидентифицируемые.

1. если все ее структурные коэффициенты определяются однозначно, т.е. единственным образом по коэффициентам приведенной формы. Это означает, число параметров структурной модели равно числу параметров приведенной формы модели. В этом случае, структурные коэффициенты можно оценить через параметры приведенной формы.
2. Если число приведенных коэффициентов меньше числа структурных коэффициентов, которые могут быть оценены через коэффициенты приведенной формы. Структурная форма модели в полном виде и эндогенными переменными и экзогенными всегда неидентифицируема.
3. Если число приведенных коэффициентов больше числа структурных коэффициентов. В этом случае, на основе приведенных коэффициентов можно получить 2 и более значения структурных коэффициентов. В такой модели число структурных коэффициентов меньше числа приведенных коэффициентов.
4. Сверхидентифицируемая модель в отличии от неидентифицируемой практически решаема, но для этого требуются специальные методы. Модель считается идентифицируема, если каждое уравнение системы идентифицируется.. если хотя бы одно из уравнений системы не идентифицируется, то модель считается неидентифицируемой. Сврерхидентифицируемая модель содержит хотя бы одно сверхидентифицируемое уравнение. Если обозначить число эндогенных уравнений в j уравнении через н, а число экзогенных тип содержаться в системе, но не входят в данное уравнение через Д, то условие идентифицируемости можно записать в сл. Виде: Д+1=н - идент,Д+1<н- неидент, Д+1>н – сврерхидент. для определения структурной модели система должна быть идентифицируема или Сверхидентифицируемая.

30. Специфика статистической оценки взаимосвязи 2 временных рядов

В предыдущей главе было показано, что временной ряд содержит 3 основных компоненты: тенденцию, циклические и сезонные колебания и случайные компоненты. Наличие этих компонент сказывается на результатах корреляционно-регрессионного анализа временных рядов данных. Предварительный этап такого анализа заключается в выявлении структуры изучаемых временных рядов. Если на этом этапе выявлено, что временные ряды содержат циклические или сезонные колебания, то перед проведением дальнейшего исследования необходимо устранить сезонную или циклическую составляющую из уровня ряда. Это устранение можно проводить в соответствии с методикой построения аддитивной или мультипликативной модели. Пусть изучается зависимость между временными рядами Х и У. для количественной характеристики такой зависимости используются линейные коэффициенты корреляции. Для того, чтобы получить коэффициенты корреляции нужно избавиться от «ложной корреляции». Она означает, наличие тенденции в каждом ряде. Обычно это осуществляется с помощью метода исключения тенденции. Наличие тенденции в каждом временном ряде означает, что на зависимую переменную Yt и независимую переменную Xt оказывает влияние фактор времени, который в модели непосредственно неучтен. Yt=a+bXt+Et. фактор времени учитывается в корреляционной зависимости между значениями остатка Et за текущий момент времени. Такое влияние получило название «автокорреляцией в остатках» - нарушение одной из основных предпосылок МНК, в которой говорится о случайности остатков, полученных по уравнению регрессии. Возможный путь решения проблемы - применение обобщенного метода МНК.

31. методы исключения тенденции

Сущность всех таких методов заключается Вт том, что устранить или зафиксировать воздействие фактора времени на формирование уровня ряда. Основные методы исключения тенденции можно разделить на 2 группы:

1 группа. Методы, основанные на преобразовании исходных уровней ряда в новые переменные, не содержащие тенденции. Полученные переменные используются дальше для анализа взаимосвязи изучаемых временных рядов. Такие методы предполагают непосредственное устранение трендовой компоненты из каждого уровня временного ряда. В этой группе 2 основных метода: - метод последовательных разностей; - метод отклонений от тренда.

2 группа. Методы, основанные на изучении взаимосвязи исходных уровней временных рядов при устранении воздействия фактора времени на зависящую и независимую переменные модели. К таким методам относятся – метод включения в модель регрессии фактора времени.

Метод отклонений от тренда. Пусть имеется 2 временных ряда Хt и Yt. Каждый из которых содержит трендовую переменную t и случайную компоненту Е. после проведения аналитического выравнивания можно найти параметры соответствующих уравнений тренда и определить расчетные по тренду значения и , соответствующие исходным временным рядам. Эти значения можно принять за оценку трендовой компоненты t каждого ряда, тогда влияние тенденции можно устранить путем вычитания расчетных значений из фактических уровней ряда. Эту процедуру преодолевают для каждого временного ряда, а разностей Xt- иYt- при условии, что эти отклонения не содержит тенденции.

Метод последовательных разностей. В ряде случаев вместо аналитического выравнивания для устранения тенденции используется простой метод – метод последовательных разностей. Если временной ряд содержит ярко выраженную линейную тенденцию, то ее можно устранить путем замены исходных уровней ряда с цепными абсолютными приростами (первыми разностями). Пусть временной ряд Yt содержит тренд и случайную компоненту Еt. =a+bt. Тогда первая разность



Коэффициент b – компонента, не зависящая от времени при наличии сильной тенденции остатки Et и E(t-1) малы и в соответствии с предпосылками МНК носят случайный характер, поэтому первые разности ∆t не зависят от переменной времени и их можно использовать для дальнейшего анализа. Если временной ряд содержит тенденцию в форме параболы, то для ее устранения используются разности второго порядка, которые считаются через разности первого порядка =a+bt+сt2. . Если временной ряд содержит экспоненциальный или степенной тренд, то метод последовательный разностей не применим не к исходным уровням ряда, а к их логарифмам. Методы разностей при своей простоте имеет 2 недостатка: - его изменение связано с сокращением числа пар наблюдений по которым строится уравнение регрессии. Это ведет к потере числа степеней свободы. – использованием вместо исходных уровней их прироста приводит к потере информации, содержащейся в исходных данных.

Метод включения в модель регрессии фактора времени. В корреляционно- регрессионном анализе устранить воздействие какого-либо фактора можно, если зашифровать это воздействие на результат и др включить в модель фактора. Этот прием широко используется в анализе временных рядов, когда тенденция фиксируется через включение фактора времени в качестве независимой переменной в модель: 

 +b3x(t-1)+biY(t-1)

Такая модель включает число независимых переменных больше 1. кроме этого в нее могут быть включены не только текущие, но и лаговые значения независимой переменной, а также лаговые значения результативных переменных. Преимущество данной модели, по сравнению с 2 предыдущими состоит в том, что она позволяет учесть всю информацию, содержащуюся в рядах Xt и Yt. Кроме того, модель строится по всей совокупности данных за рассмотренный период, а значит не ведет к потери числа степеней свободы. Параметры a и b определяются МНК, включая фактор времени.

32. автокорреляция остатка. Критерий Дарбина-Уотсона

Рассмотрим уравнение множественной регрессии вида



где к- число, независимых переменных. Для каждого момента времени t=1,n остатки



 Если рассмотреть последовательность остатка как временной ряд, то можно построить зависимость от времени. В соответствии с предпосылками МНК, остатки должны носить случайный характер.

Но при моделировании временных рядов возможны ситуации, когда имеется тенденция (2),(3) или циклические колебания. Последние 3 ситуации говорят о том, что каждое последующее значение зависит от предыдущего. В этом случае говорят об автокорреляции в остатках.

Причины автокоррнек5ляции в остатках.

* 1. связана с исходными данными и вызвана наличием ошибок измерения в значениях результативного признака Yt.
	2. В формулировке модели, т.е модель может не включать фактор оказывающий сильное воздействие на результат. Это воздействие сказывается на остатках, из-за чего они становятся автокоррелированными. Чаще всего таки м фактором является фактор времени t.

Существует 2 основных метода определения автокорреляции в остатках.

1. построение графика зависимости остатка времени и визуальное определение наличия или отсутствия автокорреляции.
2. Использование критерия Дабрина-Уотсона. В этом случае рассчитывается величина

d=.

После вычисления коэффициента d выдвигается гипотеза Ho об отсутствии автокорреляции в остатках. Альтернативные гипотезы H1 и Н1\* наличие положительной или отрицательной автокорреляции. Далее по спец. Таблицам значения критерия Дарбина-Уотсона определяется критическое значение dL и du для количественного наблюдения и числа независимых переменных к. кроме этого задается уровень значимости L по значениям dLи du числовой промежуток разбивается на 5 отрезков. При этом принятые гипотезы Но с вероятностью 1-l определяется с помощью схемы:

 В зависимости от того в какую область попадает фактическое значение критерия Д-У d можно сделать сл. Выводы. Если в области 1 существует положительная автокорреляция Но отклоняется в с вероятностью 1-l принимается гипотеза Н1. в область 3- нет основания отвергать нулевую гипотезу, т.е автокорреляция остатков отсутствует. В область 5 – существует отрицательная автокорреляция в остатках. Но отклоняется и также с вероятностью 1-l принимается Н1\*.в область 2 или 4 – зоны неопределенности. В этом случае на практике предполагают существование автокорреляции остатка и отклоняют гипотезу Но.

Ограничения применения критерия Д-У.

1. он не применим к моделям, включающим в качестве независимых переменных лаговые значения результативного признака, т.е. к моделям автокорреляции.
2. методика расчета и использования критерия Д-У направлена на выявление автокорреляции только перового порядка. При проверке остатка на автокорреляцию более высокого порядка применяются др. методы.
3. критерий дает достоверные результаты только для больших выборок.

 33. оценивание параметров уравнения регрессии при наличии автокорреляции в остатках.



к этому уравнении. Применим несколько допущений:

1. пусть У и Х не содержат тенденции.
2. пусть a и b найдены МНК.
3. пусть критерий Д-У показал наличие автокорреляции в остатках.

автокорреляция в остатках первого порядка – каждый сл. Уровень остатков зависит от предыдущего. Следовательно, существует модель регрессии вида где c и d неизвестные параметры, остатки остатков. В соответствии с рабочими формулами МНК c и d сл. Образом:

с=

после того, как найдены c и d можно сказать, что уравнение регрессии между Х и У приобретают вид:



Это означает, что Yt зависит не только от Х, но и от Е. потому, чтобы избавиться от автокорреляции в остатках необходимо использовать обобщенный МНК. Для его реализации необходимо выполнить сл. Условия:

1. исходные переменные Yt и Xt преобразуем к виду:



1. применив обычный МНК уравнения рассчитываем параметры a’ и b.
2. рассчитать параметр а исходного уравнения по формуле



1. записать уравнение с найденными a и b.

основная проблема, связанная с применением данного алгоритма заключается в том, чтобы построить оценку. Основанной способ вычисления этого коэффициента: как оценка по непосредственным остаткам, которые получаются из исходного уравнения регрессии и получение его приближенного значения из соотношения между коэффициентом корреляции перового порядка и критерием

Д-У.

* 1. оценивание параметров структурной модели

Коэффициенты структурной модели могут быть оценены разными способами, в зависимости от вида системы одновременных уравнений. Наибольшее распространение получили сл. Элементы:

-косвенный МНК (КМНК).

- двухшаговый МНК. (ДМНК).

- трехшаговый МНК.

- Метод мах правдоподобия с полной информацией.

- метод мах правдоподобия при ограниченной информации.

Косвенный МНК. Применяется для идентифицируемой модели одновременных уравнений. Он достаточно легко реализуем и предполагает выполнение сл. Этапов: структурная форма модели преобразуется в приведенную, для каждого уравнения приведенной формы определяются коэффициенты δij обычно МНК коэффициенты приведенной модели трансформируются в параметры структурной модели.

ДМНК. Если модель Сверхидентифицируемая, то использовать КМНК нельзя, т.е. в этом случае, однозначных оценок для параметров структурной модели не получится. Основная идея ДМНК на основе приведенной формы модели - получить для сверхидентифицируемого уравнения теоретические значения эндогенных переменных, содержащихся в правой части уравнения. Далее подставив эти значения можно применить обычный МНК в структурной форме сверхидентифицируемого уравнения. Метод получил название двухшагового, т.к. дважды используется обычный МНК. На первом шаге при определении приведенной формы модели и нахождения на ее основе оценок теоретических значений эндогенных переменных и на втором этапе применительно сверхидентифицируемой. Сверхидентифицируемые модели делятся на 2 вида: - все уравнения в системе сверхидентифицируемые; - содержат как идентифицируемые, так и сверхидентифицируемые уравнения. Если все уравнения системы сверхидентифицируемые, то для оценок структурных коэффициентов каждого уравнения применяется ДМНК. Если в системе есть идентифицируемые уравнения, то структурные коэффициенты по ним находятся из системы приведенных уравнений.

Метод мах правдоподобия используется как наиболее распространенный метод оценивания, результаты которого при нормальном распределении совпадаю с МНК. Но при большом числе уравнений системы этот метод приводит к достаточно сложным вычислительным процедурам, поэтому в качестве модификации используется метод мах правдоподобия с ограниченной информацией. В отличии от метода мах правдоподобия в общем виде, в этом методе сняты ограничения на параметры, связанные с функционированием системы в целом.