

КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
им. И. АРАБАЕВА
КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ им. И. РАЗЗАКОВА

Диссертационный совет Д 05.23.689

На правах рукописи
УДК 681.142.35

Давлятова Бузира

ПРОБЛЕМЫ АВТОКОРРЕЛЯЦИИ ОСТАТКОВ И
МУЛЬТИКОЛЛИНЕАРНОСТИ В РЕГРЕССИОННЫХ
МОДЕЛЯХ

05.13.16 - применение вычислительной техники, математического
моделирования и математических методов в научных исследованиях

Автореферат диссертации на соискание ученой степени кандидата
технических наук

Бишкек – 2024

Работа выполнена на кафедре прикладной информатики Кыргызского государственного университета им. И. Арабаева.

Научный руководитель: **Курманбек уулу Талантбек**, доктор технических наук, проректор по учебной работе Кыргызского государственного университета им. И. Арабаева.

Официальные оппоненты: **Абдылдаев Мукаш Юнусалиевич**, доктор технических наук, профессор кафедры прикладной информатики института новых информационных технологий Кыргызского государственного университета им. И. Арабаева

Султанов Райымбек Касымович, кандидат физико-математических наук, доцент, руководитель Отделения компьютерной инженерии, Кыргызско-Турецкий университет Манас

Ведущая организация: Кафедра прикладная информатика и программирования Таразского регионального университета им. М.Х. Дулати, Республики Казахстан (H02T7A5, г.Тараз, ул.Сулейменова, 7)

Защита диссертации состоится 29 марта 2024 года в 14.00 часов на заседании диссертационного совета Д 05.23.689 по защите диссертаций на соискание ученой степени доктора (кандидата) технических наук и кандидата физико-математических наук при Кыргызском государственном университете имени И. Арабаева и Кыргызском государственном техническом университете имени И. Раззакова по адресу: 720026, г. Бишкек, ул. Раззакова 51-А, конференц-зал КГУ им. И. Арабаева.

Идентификационный код онлайн трансляции защиты диссертации:

<https://vc.vak.kg/b/d05-xy8-8iw-xsw>

С диссертацией можно ознакомиться в библиотеках Кыргызского государственного университета им. И. Арабаева (720026, г. Бишкек, ул. И. Раззакова, 51-А) и Кыргызского государственного технического университета им. И. Раззакова (720044, г. Бишкек, пр. Ч. Айтматова 66) и на сайте НАК Кыргызской Республики (<https://vak.kg>).

Автореферат разослан 28- февраля 2024 года

Ученый секретарь
диссертационного совета,
к.ф.-м.н.

Асанбекова Н.О.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. При построении эконометрических моделей используется выборочный метод, т.е. конечное множество статистических данных исследуемых переменных, к тому же экономические переменные являются случайными величинами, также в зависимости от постановки задачи, в модели учитываются не все независимые переменные и т.д. В силу этих причин получение качественных моделей является важным моментом в эконометрическом моделировании. Только качественная модель может быть использована для анализа и прогнозирования.

В данной работе анализированы некоторые проблемы, влияющие на качества регрессионных моделей: обнаружение и устранение автокорреляции остатков. Предлагаются предложения автора по указанным проблемам.

Кроме того, наличие мультиколлинеарности во множественных моделях также является проблемой в построении эконометрических моделей. В данной диссертации тщательно исследованы подбор независимых переменных множественных регрессионных моделей, проверка качества этих моделей и сформулированы предложения по решению некоторых проблем в построении многофакторных моделей в условиях мультиколлинеарности.

Известно, что свойства оценок эконометрических моделей, полученных методом наименьших квадратов, достаточно сильно зависят от случайной составляющей модели, поэтому пристальное внимание уделяется поведению случайного члена модели.

Связь темы диссертации с приоритетными научными направлениями, основными научно – исследовательскими работами. Диссертационная работа выполнена в соответствии с планом научно-исследовательских работ кафедры «Прикладная информатика» КГУ им. И. Арабаева.

Цели и задачи исследования. Целью исследования является развитие методологии эконометрического моделирования по методам проверки и получения качественных моделей. Для достижения поставленной цели были поставлены следующие задачи:

1. Исследовать методы проверки по обнаружению автокорреляции остатков модели и разработать свои предложения;
2. Исследовать методы по устранению автокорреляции остатков и разработать свои предложения;
3. Исследовать методы по подбору независимых переменных множественных регрессионных моделей и разработать свои предложения;
4. Исследовать построения множественных регрессионных моделей в условиях мультиколлинеарности и разработать свои предложения;
5. Разработать автоматизированную систему построения регрессионных моделей.

Научная новизна полученных результатов. Получены следующие результаты исследования:

1. Разработан новый тест по обнаружению автокорреляции остатков модели.
2. Предложен новый подход при выборе значения коэффициента автокорреляции между остатками при применении авто регрессионного преобразования $AR(1)$ для устранения автокорреляции остатков;
3. Предлагается новый подход при выборе независимых переменных во множественных регрессионных моделях в случае несовпадения знаков парных коэффициентов корреляции и соответствующих частных коэффициентов корреляции;
4. Предлагается новый подход для определения влияния каждой объясняющей переменной на зависимую переменную множественных регрессионных моделей в условиях мультиколлинеарности;
5. Разработана автоматизированная информационная система построения регрессионных моделей, включающая полное исследование качества моделей.

Практическая значимость полученных результатов. Полученные результаты, применяемые при исследовании качества эконометрических моделей, способствуют улучшению качества моделей, увеличивают возможности получения адекватных моделей.

Разработанная автоматизированная система построения регрессионных моделей позволяет построение и исследование качества однофакторных, многофакторных, нелинейных регрессионных моделей, а также моделей с применением временных рядов с сезонными колебаниями. На данную систему получено свидетельство № 410 14 июля 2016 г. в Государственной патентной службе КР. Данная система внедрена в учебный процесс КГТУ им. И. Раззакова, подтверждается актом внедрения.

Экономическая значимость полученных результатов. Экономическая значимость полученных результатов заключается в улучшении качества эконометрических моделей, что позволяет получить максимально достоверные результаты анализа и прогнозы при применении разработанных методов, а также в построении и полном исследовании регрессионных моделей с помощью разработанной автоматизированной системы.

Основные положения диссертации, выносимые на защиту:

1. Новый тест по обнаружению автокорреляции остатков регрессионных моделей.
2. Новое предложение при выборе значения коэффициента корреляции между остатками ρ при применении авто регрессионного преобразования $AR(1)$ для устранения автокорреляции остатков;
3. Новый подход при выборе независимых переменных во множественных регрессионных моделях в случае несовпадения знаков парных коэффициентов корреляции и соответствующих частных коэффициентов корреляции;

4. Предлагается новый подход моделирования для определения влияния каждой объясняющей переменной на зависимую переменную множественных регрессионных моделей в условиях мультиколлинеарности;

5. Автоматизированная информационная система построения регрессионных моделей.

Личный вклад диссертанта состоит в проведении самостоятельных исследований, в получении научных результатов, их анализе и формулировании основных выводов, разработке автоматизированной системы построения регрессионных моделей.

Апробация диссертационной работы. Основные результаты работы докладывались и обсуждались на научных семинарах и конференциях.

Полнота публикации результатов. По результатам исследований и положениям, отражающим основное содержание диссертационной работы, опубликованы в 17 научных работах.

Структура диссертации. Диссертация состоит из введения, четырех глав, 17 рисунков, 11 таблиц.

Автор выражает глубокую признательность научному руководителю д.т.н. Курманбек у.Т. за руководство научной работой.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении дается общая характеристика работы, делается краткий обзор рассматриваемых проблем в построении и получении качественных регрессионных моделей.

В первой главе «Обзор литературы по проблемам автокорреляции остатков и мультиколлинеарности» делается краткий обзор литературы по исследуемым проблемам работы. Анализ научной литературы по исследованию качества эконометрических моделей представляет широкий спектр различных мнений, что свидетельствует о том, что еще предстоит много работы в этом поприще. Эконометрические методы развиваются, появляются новые методы, зачастую требуется более тонкие подходы в конкретной ситуации.

Рассмотрены работы различных авторов, содержащие научные исследования по проблемам автокорреляции остатков и мультиколлинеарности во множественных регрессионных моделях.

Особое внимание заслуживают исследования М.П. Базилевского, А.И Орловой, С.А. Айвазяна, Лапача С.Н., Раденко С.Г., Розенцвайг А.К. и др., которые предлагают оригинальные методы устранения проблем и получения качественных моделей. В их исследованиях также предлагаются свои подходы по обнаружению и устранению автокорреляции остатков, а также решение проблем мультиколлинеарности во множественных моделях.

Среди кыргызстанских ученых можно отметить работы Бийбосунова Б. И., Лукашовой И.В., Кыдыралиева С. по исследованию экономики Кыргызстана с применением эконометрического моделирования.

Вторая глава «Методология и методы исследования». Приведены основные методологии по разработке эконометрических моделей и исследованию их качества. Описаны традиционная и современная методологии эконометрики. Также приведены основные виды эконометрических моделей и методы построения регрессионных моделей.

Третья глава «Исследование автокорреляции остатков модели, гетероскедастичности и мультикоррелированности во множественных регрессионных моделях».

Эконометрические модели строятся с использованием ограниченного количества значений исследуемых переменных, т.е. выборочной совокупности, и они являются эмпирическими или оценочными.

Особое внимание уделяется линейным регрессионным моделям, так как большинство нелинейных моделей строится с приведением их к линейному виду. Пусть рассматривается зависимость переменной Y от m независимых переменных $X=(x_1, x_2, \dots, x_m)$. Теоретическая модель линейной регрессии имеет вид:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon \quad (1)$$

где Y - зависимая переменная; X_1, X_2, \dots, X_m - независимые переменные, ε – случайная ошибка регрессии; $\beta_1, \beta_2, \dots, \beta_m$ – коэффициенты регрессии, β_0 , – свободный член.

Вместо теоретической линейной регрессионной модели (1) строится её оценка - эмпирическая регрессионная модель:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_m X_m + e \quad (2)$$

Для построения модели (2) применяется метод наименьших квадратов, суть этого метода заключается в минимизации суммы квадратов остатков модели:

$$Q = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left[y_i - (b_0 + \sum_{j=1}^m b_j x_{ij}) \right]^2 \Rightarrow \min \quad (4)$$

Полученные при этом коэффициенты $b_0, b_1, b_2, \dots, b_m$ являются оценками соответственно $\beta_0, \beta_1, \beta_2, \dots, \beta_m$. Для дальнейшего анализа регрессионной модели используем следующие предположения об отклонениях для любого x :

1. Математическое ожидание отклонений модели равно нулю для всех наблюдений: $M(\varepsilon_i)=0$.

2. Условие гомоскедастичности - дисперсия отклонений постоянна и одинакова для всех наблюдений: $D(\varepsilon_i) = D(\varepsilon_j) = \sigma^2$, $i, j=1, \dots, n$.

3. Отсутствие автокорреляции остатков модели: $\rho_{\varepsilon_i \varepsilon_j} = 0$ при $i \neq j$.

4. Случайные отклонения модели не зависят от объясняющих переменных: $\rho_{\varepsilon_i x_j} = 0$.

5. Модель является линейной относительно переменных.

6. Остатки модели ε_i нормально распределены.

Условия (1) – (4) называются условиями Гаусса – Маркова.

Для исследования качества эконометрических моделей, т.е. выполнения условий Гаусса – Маркова, проводятся следующие работы:

- проверка статистической значимости свободного члена и коэффициентов регрессии;
- проверка общего качества модели;
- проверка наличия или отсутствия автокорреляции остатков;
- проверка постоянства дисперсий отклонений, т.е. гомоскедастичности.

Для множественных линейных регрессионных моделей добавляется еще выполнение условия - отсутствие мультиколлинеарности объясняющих переменных.

В данной работе рассматриваются вопросы построения качественных эконометрических моделей, такие, как автокорреляция остатков, мультиколлинеарность. Описаны проблемы автокорреляции остатков модели, методы их обнаружения и устранения и проблема мультиколлинеарности независимых переменных во множественных регрессионных моделях, методы их устранения.

Проблемы автокорреляции остатков модели. Остатки эконометрических моделей ε_i содержат много информации: ошибку спецификации модели, неточность и ограниченность статистических данных исследуемых переменных и т. п.

Как отмечено выше, одним из условий Гаусса – Маркова, выполнение которого необходимо для эконометрических моделей, является отсутствие автокорреляции остатков ε_i . Отклонения эмпирической модели обозначаются через e_i и они являются оценками ε_i .

Обнаружение автокорреляции остатков. Для обнаружения автокорреляции остатков, на практике чаще используют тест Дарбина-Уотсона. Но этот тест имеют свои недостатки: он применим только для моделей со свободными членами; остатки подчиняются к итерационной схеме - $\varepsilon_i = \rho\varepsilon_{i-1} + v_i$; имеются зоны неопределенности.

Другой метод - метод рядов используется для получения более точных результатов.

Тест Бреуша – Голдффри используется при больших выборках и в основном для авторегрессионных моделей, поэтому в данной работе не рассматриваем.

Предложение диссертанта по обнаружению автокорреляции остатков. Диссертант предлагает использовать выборочный коэффициент корреляции между соседними остатками для обнаружения автокорреляции:

$$r_{i,i-1} = \frac{\sum e_i e_{i-1}}{\sqrt{\sum e_i^2 \sum e_{i-1}^2}}, \quad (5)$$

считая, что $M(e_i)=0$. Эти коэффициенты называются коэффициентами автокорреляции остатков первого порядка.

Действительно, если доказать статистическую значимость или незначимость теоретического коэффициента автокорреляции остатков $\rho_{i,i-1}$, оценкой которого является эмпирический коэффициент автокорреляции $r_{i,i-1}$, то можно ответить на поставленный вопрос.

Приведем теоретическую часть этого предположения. Проверяется статистическая значимость выборочного коэффициента корреляции между остатками модели. Для этого проведем испытание следующей гипотезы при уровне значимости α :

H_0 : Между остатками нет коррелированности, т.е. $\rho_{i,i-1} = 0$.

H_1 : Между остатками модели существует линейная связь: $\rho_{i,i-1} \neq 0$.

Как известно, испытание данной гипотезы проводится с помощью t – критерия. Проверочная статистика:

$$t = \sqrt{\frac{r_{i,i-1}^2 \cdot (n-m-1)}{(1-r_{i,i-1}^2)}}, \quad (6)$$

где n – объем выборки, m – число независимых переменных модели. Сравнивая значение проверочной статистики с критическим значением $t_{\alpha/2;n-m-1}$, сделаем выводы:

- при $t < t_{\alpha/2;n-m-1}$ принимается H_0 , т.е. $\rho_{i,i-1} = 0$, между остатками нет коррелированности.

- при $t > t_{\alpha/2;n-m-1}$ принимается H_1 , т.е. $\rho_{i,i-1} \neq 0$, соседние остатки коррелированы между собой.

- Кроме того, при $r_{i,i-1} \geq 0$, можно говорить о наличии положительной автокорреляции, а при $r_{i,i-1} \leq 0$ об отрицательной автокорреляции.

Изложенный способ был применен для обнаружения автокорреляции остатков при моделировании ВВП в зависимости от объема промышленной продукции и инвестиций в основной капитал с данными КР за 2000 – 2019 гг., при моделировании средней величины заработной платы В КР за 2000 - 2019 гг., также при моделировании расходов на питание от личных располагаемых доходов с данными США и др. Полученные результаты применения этого способа и результаты сравнения с известными методами, приведены в таблице 1.

Таблица 1 – Сравнение результатов проверки автокорреляции разными методами для моделей ВВП (КР) и расходов на питание (США) после устранения автокорреляции

Методы	Модель ВВП (КР)	Результаты обнаружения автокорреляции остатков	Модель расходов на на питание (США)	Результаты обнаружения автокорреляции остатков
Тест Дарбина-Уотсона	$DW = 1.88$, $d_l = 1.05$, $d_u = 1.53$	Автокорреляция отсутствует	$DW = 1.428$, $d_l = 1.18$, $d_u = 1.401$	Автокорреляция отсутствует
Метод рядов	$k = 10$, $k_l = 6, k_2 = 16$	Автокорреляция отсутствует	$k_l = 6, k_2 = 16$, $k = 10$	Автокорреляция отсутствует
По нашему предложению	$r_{i,i-1} = 0.105$, $\alpha = 0.05$, $t = 0.44$, $t_{krum} = 2.11$	Автокорреляция отсутствует	Автокорреляция отсутствует	Автокорреляция отсутствует

Кроме того, экспериментальным путем доказано, что при обнаружении автокорреляции остатков с помощью теста Дарбина – Уотсона и методом рядов, этот факт также обнаруживает с применением теста коэффициента корреляции $r_{i,i-1}$. Приведем результаты проверки для модели расходов на питание с данными США до устранения автокорреляции остатков (табл. 2).

Таблица 2 - Сравнение результатов проверки автокорреляции разными методами для модели расходов на питание до устранения её (США)

Методы	Модель расходов на Питание (США)	Результаты обнаружения автокорреляции остатков
Тест Дарбина-Уотсона	$DW = 0.778$, $d_l = 1.201, d_u = 1.411$	Имеется положительная автокорреляция
Метод рядов	$k_l = 7, k_2 = 17, k = 6$	Имеется положительная автокорреляция
По нашему предложению	$r_{i,i-1} = 0.597, t = 3.33$, $t_{krum} = 2.093, \alpha = 0.05$	Имеется положительная автокорреляция

Теперь рассмотрим результаты для модели ВВП с данными КР до устранения автокорреляции остатков, обнаруженного по тесту Дарбина – Уотсона (табл. 3).

Таблица 3 - Сравнение результатов проверки автокорреляции разными методами для модели ВВП (КР) до устранения её

Методы	Модель ВВП (КР)	Результаты обнаружения автокорреляции остатков
Тест Дарбина-Уотсона	$DW = 1.149$, $d_l = 1.05, d_u = 1.53$	Ситуация неопределенности
Метод рядов	$k_l = 6, k_2 = 17, k = 8$	Автокорреляция отсутствует
По нашему предложению	$r_{i,i-1} = 0.364, t = 1.61$, $t_{krum} = 2.11, \alpha = 0.05$	Автокорреляция отсутствует

Приведем результаты проверки наличия автокорреляции остатков при моделировании средней заработной платы в КР тремя методами в табл. 4.

Таблица 4 - Сравнение результатов проверки автокорреляции разными методами для модели средней заработной платы (КР)

Методы	Модель средней заработной платы в КР	Результаты обнаружения автокорреляции остатков
Тест Дарбина-Уотсона	$DW=1.569$ $d_l=1.026, d_u=1.669$	Ситуация неопределенности
Метод рядов	$k_1=6, k_2=16, k=7$	Автокорреляция отсутствует
По нашему предложению	$r_{i,j-1}=0,206, t=1,61, t_{krit}=2,12, \alpha=0,05$	Автокорреляция отсутствует

Как видно, результаты проверки по обнаружению автокорреляции остатков по тесту Дарбина – Уотсона, методу рядов и по предложенному в данной работе методу совпадают, кроме ситуации неопределенности по Дарбину – Уотсона.

Одной из причин такого несоответствия могут быть допущения, сделанные для определения границ областей принятия, отклонения нулевой гипотезы и наличия зон неопределенности по тесту Дарбина – Уотсона.

Проведенные исследования показывают, что предложенный метод обнаружения автокорреляции остатков регрессионных моделей с применением коэффициента автокорреляции остатков, может применяться на практике.

Устранение автокорреляции остатков модели. Для устранения автокорреляции остатков модели сначала следует проверить спецификацию модели, либо увеличить объем выборки, уточнить значения данных. Если такие простые способы не устранили автокорреляцию, то применяется авто регрессионная схема первого порядка $AR(1)$.

Для примера возьмём парную линейную регрессионную модель:

$$y_t = b_0 + b_1 x_t + e_t, \quad (7)$$

Составим такое же уравнение для y_{t-1} :

$$y_t - \rho \cdot y_{t-1} = b_0 \cdot (1 - \rho) + b_1 \cdot (x_t - \rho \cdot x_{t-1}) + (e_t - \rho \cdot e_{t-1}). \quad (8)$$

Введя обозначения

$$y_t^* = y_t - \rho \cdot y_{t-1}, \quad b_0^* = b_0 \cdot (1 - \rho), \quad x_t^* = x_t - \rho \cdot x_{t-1}, \quad v_t = e_t - \rho \cdot e_{t-1},$$

получим: $y_t^* = b_0^* + b_1 x_t^* + v_t$, при этом ρ означает коэффициент корреляции между соседними остатками модели. Затем находим $b_0 = \frac{b_0^*}{1 - \rho}$ и подставляя в модель (7) получаем новую модель. Но здесь значение ρ неизвестно, методы его определения рассмотрим ниже.

Предложение по выбору коэффициента корреляции между остатками для авторегрессионного преобразования $AR(1)$. Рассмотрим применение авто регрессионного преобразования $AR(1)$ в случае наличия

автокорреляции остатков. Для этого должно быть известно значение ρ , которое присутствует в уравнении регрессии (8).

Для нахождения значения этого параметра имеются несколько методов: методы Хилдрета – Лу, Кохрана – Оркатта, метод первых разностей.

По методу Хилдрета –Лу, оценка уравнения (8) будет производиться для каждого значения ρ из отрезка $[-1; 1]$ с шагом, например 0,001 или 0,01 и т.д. В качестве окончательного значения ρ берется такое значение ρ , при котором достигается минимальная стандартная ошибка регрессии.

Кохран – Оркатт предлагает найти ρ с помощью авто регрессионной схемы первого порядка $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$. Найденное при построении этой регрессии значение ρ ставится в уравнение (8), далее находится $b_0 = \frac{b_0^*}{1-\rho}$, получим искомое регрессионное уравнение (7). Снова вычисляются ошибки последней модели, и повторяется описанный процесс с новым значением ρ . Процесс повторяется до тех пор, пока не достигнет сходимости значений по ρ .

Метод первых разностей предполагает слишком сильное допущение, сделав $\rho = \pm 1$, поэтому выше перечисленные методы являются более предпочтительными.

Исследование автора на примере моделирования ВВП в зависимости от объема промышленной продукции и инвестиций в основной капитал показало, что значения ρ , найденные по методу Хилдрета – Лу ($\rho = 0,42$) и Кохрана – Оркатта (0,41) почти совпадают (табл. 5).

Таблица 5 – Определение ρ при устранении автокорреляции

Критерий	ρ	Станд.ошибка регрессии	Статистика Дарбина – Уотсона (DW)	Наличие автокорреляции остатков
По Хилдрету-Лу	0,42	$9,1158 \cdot 10^9$	1,7827	Отсутствует автокорреляция остатков
По Кохрану - Оркатту	0,41	$9,1161 \cdot 10^9$	1,7725	Отсутствует автокорреляция остатков

По тесту Дарбина-Уотсона, автокорреляция остатков исчезла и при $\rho = 0,42$ и при $\rho = 0,41$, так как из таблицы Дарбина – Уотсона имеем - $d_l = 1,1$, $d_u = 1,537$, $4 - d_l = 2,9$, $4 - d_u = 2,463$. Разделение областей выглядит: $(0; 1,1)$ – интервал, в котором имеется положительная автокорреляция, $(2,9; 4)$ – интервал, в котором имеется отрицательная автокорреляция, $(1,537; 2,463)$ – интервал отсутствия автокорреляции, $(1,1; 1,537)$ и $(2,463; 2,9)$ – интервалы неопределенности. В этом случае следует выбрать $\rho = 0,42$ с наименьшим значением стандартной ошибки регрессии.

На практике не всегда удается устранить автокорреляцию остатков модели после применения авто регрессионного преобразования $AR(1)$. В

качестве примера рассмотрим моделирование расходов на питание (Y) в зависимости от личных располагаемых доходов (X) на примере данных США за 1959 – 1980 гг. Были получены следующие результаты:

Таблица 6 – Результаты проверки автокорреляции после применения $AR(1)$

Критерий	ρ	Станд.ошибка регрессии	Статистика Дарбина – Уотсона (DW)	Наличие авто-корреляции остатков
По Хилдрету-Лу	0,62	2,412	1,385	Ситуация неопределенности
По Кохрану-Оркатту	0,62	2,412	1,385	Ситуация неопределенности

В связи с этим возникает вопрос: Не выбрать ли значение ρ не только при минимальной стандартной ошибке регрессии по методу Хилдрета – Лу или по методу Кохрана – Оркатта по сходимости значений ρ , но и при отсутствии автокорреляции остатков и при выполнении других условий к качеству модели?

Естественно, при добавлении указанного условия при выборе ρ , стандартная ошибка регрессии может стать на немного больше. Если при этом коэффициенты регрессии останутся статистически значимыми и качество модели не ухудшится, то можно сделать вывод о правильности такого подхода.

Рассуждая, таким образом, на примере моделирования расходов на питание от личных располагаемых доходов США получены следующие результаты (в данном случае $d_l = 1,221$, $d_u = 1,42$).

Таблица 7 – Сравнение значений ρ , полученных разными методами

Критерий	ρ	Станд.ошибка регрессии	Статистика Дарбина – Уотсона (DW)	Наличие авто-корреляции остатков
По Хилдрету-Лу	0,62	2,412	1,385	Ситуация неопределенности
По Кохрану - Оркатту	0,62	2,412	1,385	Ситуация неопределенности
По нашему предложению	0,67	2,416	1,428	Отсутствует автокорреляция остатков

При этом не ухудшилось качество модели, так как свободный член и коэффициент регрессии статистически значимы: $t_{b_0} = 7.64$, $t_{b_1} = 9.92$ и $t_{krut} = 2.093$ при уровне значимости $\alpha = 0.05$. Также общее качество модели остается хорошим: статистика $F = 98.35$, $F_{krut} = 4.38$. Стандартная ошибка модели увеличилась незначительно.

Идея изложенного подхода соответствует современной эконометрической методологии и правильность этого подхода подтверждена на других примерах.

Проблемы мультиколлинеарности в многофакторных моделях. Одной из проблем в построении множественной регрессии является наличие линейной связи между объясняющими переменными модели, т.е. мультиколлинеарность. Все объясняющие (независимые) переменные множественной регрессионной модели должны быть независимы между собой.

Последствия мультиколлинеарности:

1. Большие дисперсии, следовательно, и стандартные ошибки полученных оценок.
2. Уменьшение t – статистик коэффициентов регрессии, в результате чего можно сделать неправильный вывод о влиянии соответствующей объясняющей переменной на зависимую переменную.
3. Неустойчивость полученных оценок.
4. Трудности в определении вклада каждой объясняющей переменной на поведение независимой переменной.
5. Возможно даже получение неверного знака у коэффициента регрессии.

Признаки, по которым можно определить наличие мультиколлинеарности:

1. Некоторые из коэффициентов регрессии статистически незначимы, т.е. они имеют низкие t – статистики, хотя коэффициент детерминации R^2 достаточно высок.
2. Частные коэффициенты корреляции между объясняющимися переменными достаточно высоки.
3. Парные коэффициенты корреляции между малозначимыми объясняющимися переменными достаточно высоки.

При изучении множественной регрессии иногда парные коэффициенты корреляции могут дать неправильные представления о характере связи между двумя переменными. Для этого существует коэффициент корреляции между двумя переменными, очищенный от влияния других переменных, называемый частным коэффициентом корреляции.

Рассмотрим частный коэффициент корреляции в общем случае, т.е. при наличии m независимых переменных. Корреляционная матрица из парных коэффициентов корреляции и ее обратная матрица имеют вид:

$$R = \begin{bmatrix} 1 & r_{x_1x_2} & r_{x_1x_3} & \dots & r_{x_1x_m} \\ r_{x_2x_1} & 1 & r_{x_2x_3} & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{x_mx_1} & r_{x_mx_2} & r_{x_mx_3} & \dots & 1 \end{bmatrix}, \quad C^* = R^{-1} \cdot \begin{bmatrix} c_{11}^* & c_{12}^* & c_{13}^* & \dots & c_{1m}^* \\ c_{21}^* & c_{22}^* & c_{23}^* & \dots & c_{2m}^* \\ \dots & \dots & \dots & \dots & \dots \\ c_{m1}^* & c_{m2}^* & c_{m3}^* & \dots & c_{mm}^* \end{bmatrix}.$$

Тогда частный коэффициент корреляции между переменными x_i и x_j , очищенный от влияния всех остальных $m-2$ переменных

$$r_{x_i x_j \cdot x_1 x_2 \dots x_{i-1} x_{i+1} \dots x_{j-1} x_{j+1} \dots x_m} = \frac{-c_{ij}^*}{\sqrt{c_{ii}^* c_{jj}^*}}. \quad (9)$$

Устранение мультиколлинеарности не потребуется в том случае, если целью исследования является прогнозирование значений зависимой переменной, а множественный коэффициент корреляции достаточно высок ($R^2 \geq 0,9$), что свидетельствует о высоком общем качестве модели. Наличие мультиколлинеарности в этом случае не мешает получить надежные прогнозы.

Мультиколлинеарность является проблемой, если целью исследования является определение степени влияния каждой независимой переменной на зависимую переменную в отдельности, так как она может привести к увеличению стандартных ошибок и исказить реальные зависимости между переменными.

Не существует строгого единого метода устранения мультиколлинеарности, так как причины и последствия неоднозначны.

Способы устранения мультиколлинеарности: исключение коррелированных переменных, уточнение выборки, уточнение спецификации модели, преобразование переменных.

Предложения по построению множественных регрессионных моделей. При построении множественных линейных регрессионных моделей, как сказано выше, на разных этапах могут возникнуть проблемные ситуации, решение которых требует большой осторожности и с точки зрения требований к качественным эконометрическим моделям, и с точки зрения экономического смысла рассматриваемых переменных и их связей между собой. Ниже перечислим эти возможные проблемы:

- несоответствие выводов по парным коэффициентам корреляции и по частным коэффициентам корреляции ввиду несоответствия знаков этих коэффициентов;
- несоответствие вывода об исключении независимой переменной из модели по наличию мультиколлинеарности или не исключению по экономическому смыслу;
- при наличии мультиколлинеарности независимых переменных, как определить влияние каждой из этих коррелированных переменных на зависимую переменную?

Для выяснения этих проблем было рассмотрено моделирование линейной множественной регрессионной модели ВВП КР, фактического конечного потребления КР, средней величины заработной платы КР и др. В этих работах были решены проблемы отбора регрессоров моделей, получения качественных моделей для определения влияния каждого регрессора на зависимую переменную и моделей для прогнозирования.

Получены следующие предложения по указанным выше вопросам

- при отборе независимых переменных многофакторной модели при несовпадении знаков парных и частных коэффициентов корреляции, сделать

вывод по частным коэффициентам корреляции и по экономическому смыслу связи рассматриваемых переменных;

- при наличии мультиколлинеарности во множественной регрессионной модели, сделать вывод об исключении переменных следует не только по коррелированности между независимыми переменными, но и по экономическому смыслу;

- при наличии мультиколлинеарности во множественной регрессионной модели, для определения влияния каждой независимой переменной на зависимую переменную, следует построить парные регрессионные модели в зависимости от каждой объясняющей переменной.

В заключении следует подчеркнуть, что в условиях мультиколлинеарности, в каждом конкретном случае надо подходить индивидуально в зависимости от ситуации.

В четвертой главе приведено описание математических методов, применяемых в автоматизированной системе построения регрессионных моделей: метод наименьших квадратов, метод взвешенных наименьших квадратов построения парных и множественных регрессионных моделей; методы исследования статистической значимости коэффициентов регрессии, общего качества модели; методы обнаружения и устранения мультиколлинеарности независимых переменных модели; методы обнаружения и устранения автокорреляции остатков и гетероскедастичности, определения прогнозных качеств модели; методы моделирования с использованием временных рядов с сезонными колебаниями.

Также в этой главе дается общее описание автоматизированной информационной системы построения регрессионных моделей и исследования их качества: от однофакторной до семи факторной, нелинейных моделей, моделей с временными рядами с сезонными колебаниями.

Общая схема автоматизированной системы построения регрессионных моделей выглядит следующим образом (рисунок 2):

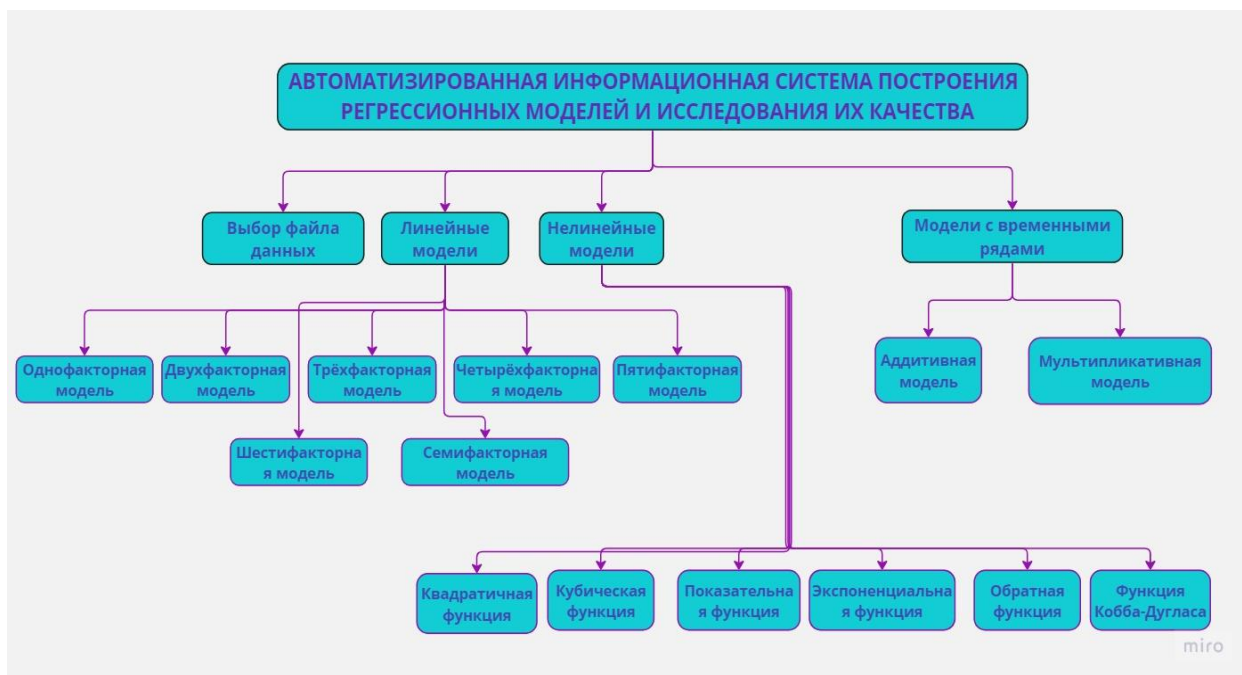


Рисунок 2 - Общая схема автоматизированной системы построения регрессионных моделей

ЗАКЛЮЧЕНИЕ

В диссертационной работе проведены исследования проблемы построения эконометрических моделей, такие как автокорреляции остатков модели, проблемы гетероскедастичности, проблемы мультиколлинеарности в многофакторных моделях. Сделаны следующие выводы:

1. По обнаружению автокорреляции остатков модели предлагается применять наряду с известными тестами, тест с применением коэффициента корреляции $r_{i,i-1}$ между остатками.

2. Для определения ρ в авторегрессии первого порядка $\varepsilon_t = \rho \cdot \varepsilon_{t-1} + v_t$ для устранения автокорреляции остатков, следует опираться не только на наименьшей стандартной ошибке регрессии или на сходимости ρ при итерации, но и на отсутствии автокорреляции при соблюдении условий к качеству модели.

3. При построении множественных линейных регрессионных моделей для решения некоторых проблемных ситуаций, предлагаются следующие пути решения:

3.1. При отборе независимых переменных многофакторной модели при несовпадении знаков парных и частных коэффициентов корреляции, сделать вывод по частным коэффициентам корреляции и по экономическому смыслу связи рассматриваемых переменных;

3.2. При наличии мультиколлинеарности во множественной регрессионной модели, сделать выбор переменных следует не только по коррелированности между независимыми переменными, но и по экономическому смыслу переменных;

3.3. При наличии мультиколлинеарности во множественной регрессионной модели, для определения влияния каждой независимой переменной на зависимую переменную, следует построить парные регрессионные модели в зависимости от каждой объясняющей переменной.

4. Разработана автоматизированная система построения и исследования качества регрессионных моделей.

5. Разработаны макроэкономические модели экономических показателей КР.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Давлятова, Б.Д. Автоматизированная система построения регрессионных моделей [программы для ЭВМ] / Б.Д. Давлятова. - Свидетельство № 410 в государственном реестре программ для ЭВМ КР, Б., Кыргызпатент, 14.07.2016.
2. Давлятова, Б.Д. Исследование экономического роста Кыргызстана с помощью эконометрического моделирования [Текст] / Б.Д. Давлятова. - Б.: Известия КГТУ им. И.Раззакова. Вып. 1(49), 2019. - С. 55-59.
3. Давлятова, Б.Д. Проблемы мультиколлинеарности и гетероскедастичности в модели конечного потребления на примере Кыргызской Республики [Текст] / Б.Д. Давлятова. - М.: Вестник науки и образования, №10(64). Часть 2. 2019. – С. 65-70.
4. Давлятова, Б.Д. Моделирование с помощью временных рядов на примере Кыргызской Республики [Текст] / Б.Д. Давлятова. - М.: Вестник науки и образования. №14(68). Часть 2. 2019. - С. 27-32.
5. Давлятова Б.Д. Моделирование структурных диспропорций экономики Кыргызской Республики [Текст] / Б.Д. Давлятова, Бийбосунов Б.И., Чорое К.Ч. и др. –М.: Фундаментальные исследования. №7. 2019. – С.21-26.
6. Давлятова, Б.Д. Моделирование средней величины заработной платы по данным Кыргызской Республики [Текст] / Б.Д. Давлятова. - М.: Проблемы современной науки и образования. №10(143). 2019. – С. 39-45.
7. Давлятова, Б.Д. Применение информационных технологий для решения региональных экономических задач [Текст] / Б.Д. Давлятова, Н.О. Асанбекова, У.Т. Керимов. – Б.: Известия вузов Кыргызстана, №5, 2016. - С.135- 138.
8. Давлятова, Б.Д. Балансовые модели женской занятости с учетом эмиграции из Кыргызстана [Текст] / Б.Д. Давлятова, С.К. Бийбосунова, У.Т. Керимов. М.: Actualscience, Том 2, №5, 2016. - С.74-78.
9. Давлятова, Б.Д. Информационная система и математические модели для региональных экономик [Текст] / Б.Д. Давлятова, С.К. Бийбосунова, Н.О. Асанбекова. – Б.: Современные проблемы механики, №26(4), 2016. –С.75-82.
10. Давлятова, Б.Д. Экономико – математическое моделирование системы оплаты труда в сфере государственной службы. [Текст]/ Б.И. Бийбосунов, Б.Д. Давлятова. – Б.: Известия вузов Кыргызстана, №5, 2017. - С.64-68.
11. Давлятова, Б.Д. Математическая модель оптимального распределения инвестиционного вложения между отраслями [Текст] / Б.Д. Давлятова, Б.Р. Сабитов, А. Сейтбеков и др. – М.: Фундаментальные исследования, №9, 2017. - С. 53-57.

12. Давлятова, Б.Д. Построение сетевой информационной системы автоматизации АКП [Текст] / Б.Д. Давлятова, Б.Р. Сабитов. – М.: Экономика и предпринимательство, №8, 2017. –С. 72-80.
13. Давлятова, Б.Д. Моделирование ВВП в зависимости от инвестиций в основной капитал на примере Кыргызской Республики [Текст] / Б.Д. Давлятова. М.: Вестник науки и образования, №12(90), 2020. – С.33-37.
14. Давлятова, Б.Д. Одна из моделей уровня безработицы на примере Кыргызской Республики [Текст] / Б.Д. Давлятова, Ж.Б.Саматова. М.: - Вестник науки и образования, №13 (91), 2020.- С.40-44.
15. Давлятова, Б.Д. Исследование ВВП во времени на примере данных Кыргызской Республики [Текст] / Б.Д. Давлятова. М.: - Научный журнал № 6 (61), М., 2021. –С.20-23.
16. Давлятова, Б.Д. Об автокорреляции остатков регрессионных моделей [Текст] / Б.Д. Давлятова, Т. Курманбек уулу. - Б.: Вестник КГУ им. И.Арабаева №2, 2023. – С.471 -475.
17. Давлятова, Б.Д. Моделирование и прогнозирование ВВП Кыргызской Республики с помощью временного ряда за 2000 – 2020гг [Текст] / Б.Д. Давлятова, Т. Курманбек уулу. –Б.: Современные проблемы механики, №50, (4), 2022. – С.19-26.

05.13.16 – «Илимий изилдөөлөрдө эсептөө техникаларын, математикалык моделдөөнү жана математикалык методдорду колдонуу» адистиги боюнча техника илимдеринин кандидаты илимий даражасы үчүн Давлятова Бузира Давлятовнанын «Регрессиялык моделдердин калдыктарынын автокорреляциясынын жана мультиколлинеардуулугунун проблемалары жөнүндө» деген темадагы диссертациялык изилдөөлөрүнүн

РЕЗЮМЕСИ

Негизги сөздөр: регрессия, эң аз квадраттар, салмактуу эң аз квадраттар, регрессия коэффициенти, корреляция коэффициенти, автокорреляция, гетероскедастика, гомоскедастика, мультиколлинеардуулук, аддитивдик моделдер, мультипликативдик моделдер.

Изилдөөнүн объектиси: эконометрикалык регрессиялык моделдөө жана регрессиялык моделдердин сапаты.

Изилдөөнүн предмети - регрессия моделдерин куруунун жана сапатын изилдөө методдору, регрессиялык моделдерди түзүүнүн автоматташтырылган системасын иштеп чыгуу.

Изилдөө методдору. Эконометрикалык моделдөө, регрессиялык моделдерди түзүү жана алардын сапатын изилдөө ыкмалары, маалыматтык технологиялар, программалоо.

Диссертациялык иштин максаты - регрессиялык моделдердин сапатын изилдөө методдорун өркүндөтүү, көп факторлуу сызыктуу регрессиялык моделдерди түзүүнүн кээ бир проблемаларын чечүү, регрессия

моделдерин куруунун жана алардын сапатын текшерүүчү автоматташтырылган системасын иштеп чыгуу.

Алынган жыйынтыктар жана алардын илимий жанылыктары. Калдык автокорреляция коэффициентин ($r_{i,i-1}$) колдонуу менен регрессиялык моделдердин калдыктарынын автокорреляциясын аныктоо үчүн жаңы тест иштелип чыкты. Калдыктардын автокорреляциясын жок кылууда биринчи даражадагы авторегрессияда колдонулуучу ρ нын маанисин аныктоо үчүн жаңы ыкма сунушталат. Көптүк регрессиялык моделдердеги түшүндүрүүчү өзгөрмөлөрдүн мультиколлинеардуулугу менен байланышкан маселелерди чечүүдө моделдин көз карандысыз өзгөрмөлөрүн тандоо, ар бир регрессордун изилденип жаткан өзгөрмөгө таасирин аныктоо үчүн моделдөө боюнча тыянактар чыгарылган. Регрессиялык моделдерди куруунун автоматташтырылган системасы иштелип чыккан.

Практикалык мааниси. Эконометрикалык моделдердин сапатын изилдөөдө алынган натыйжалар моделдердин сапатын жогорулатууга көмөктөшөт, адекваттуу моделдерди алуу мүмкүнчүлүктөрүн кеңейтет. Регрессиялык моделдерди куруунун автоматташтырылган системасы моделдерди иштеп чыгууну камсыз кылат.

Колдонуу чөйрөсү. Изилдөөнүн натыйжаларын прикладдык экономикалык проблемаларды изилдөөдө кеңири колдонууга болот.

РЕЗЮМЕ

диссертации Давлятовой Бузиры Давлятовны на тему: «Проблемы автокорреляции остатков и мультиколлинеарности в регрессионных моделях» на соискание учёной степени кандидата технических наук по специальности: 05.13.16 – «Применение вычислительной техники, математического моделирования и математических методов в научных исследованиях»

Ключевые слова: регрессия, метод наименьших квадратов, метод взвешенных наименьших квадратов, коэффициент регрессии, коэффициент корреляции, автокорреляция, гетероскедастичность, гомоскедастичность, мультиколлинеарность, аддитивные модели, мультипликативные модели.

Объект исследования: эконометрическое регрессионное моделирование и качество регрессионных моделей.

Предметом исследования являются методы построения и исследования качества регрессионных моделей, разработка автоматизированной системы построения регрессионных моделей.

Методы исследования. Эконометрическое моделирование, методы построения регрессионных моделей и исследования их качества, информационные технологии, программирование.

Целью диссертационной работы является усовершенствование методов исследования качества регрессионных моделей, решение проблем

построения многофакторных линейных регрессионных моделей, разработка автоматизированной системы построения регрессионных моделей.

Полученные результаты и их научная новизна. Разработан новый тест для обнаружения автокорреляции остатков регрессионных моделей с применением коэффициента автокорреляции остатков $r_{i,i-1}$. Предложен новый подход для определения ρ в авторегрессии первого порядка $\varepsilon_t = \rho \cdot \varepsilon_{t-1} + v_t$, применяемого при устранении автокорреляции остатков. При решении проблем, связанных с мультиколлинеарностью объясняющих переменных во множественных регрессионных моделях, получены выводы по отбору независимых переменных модели, моделированию для определения влияния каждого регрессора на исследуемую переменную. Разработана автоматизированная система построения регрессионных моделей с полным исследованием качества моделей.

Практическая значимость. Полученные результаты, применяемые при исследовании качества эконометрических моделей, способствуют улучшению качества моделей, расширяют возможности получения адекватных моделей.

Разработанная автоматизированная система построения регрессионных моделей позволяет построение и исследование качества регрессионных моделей.

Область применения. Результаты исследования могут иметь широкое применение в прикладных экономических задачах.

RESUME

of the dissertation of Davliatova Buzira Davliatovna on the topic "Problems of autocorrelation of residuals and multicollinearity in regression models" for the degree of candidate of technical sciences in the specialty: 05.13.16 - the use of computer technology, mathematical modeling and mathematical methods in scientific research

Key words: regression, least squares method, weighted least squares method, regression coefficient, correlation coefficient, autocorrelation, heteroscedasticity, homoscedasticity, multicollinearity, additive models, multiplicative models, Cobb-Douglas function, automated system.

Object of study: econometric regression modeling and the quality of regression models.

The subject of the research is the methods of building and studying the quality of regression models, the development of an automated system for building regression models.

Research methods. Econometric modeling, methods for constructing regression models and studying their quality, information technology, programming.

The purpose of the dissertation work is to improve the methods for studying the quality of regression models, building multifactorial linear regression models, developing an automated system for building regression models and checking their quality.

The results obtained and their scientific novelty. A new test has been developed to detect autocorrelation of residuals of regression models using the residual autocorrelation coefficient. A new approach is proposed for determining ρ in first-order autoregression, used when eliminating autocorrelation of residuals. When solving problems associated with multicollinearity of explanatory variables in multiple regression models, conclusions were drawn on the selection of independent variables of the model, modeling to determine the influence of each regressor on the variable under study. An automated system for constructing regression models with a complete study of the quality of the models has been developed.

Practical significance. The results obtained, used in the study of the quality of econometric models, contribute to the improvement of the quality of models, expand the possibilities of obtaining adequate models.

The developed automated system for building regression models allows the construction and study of the quality of linear single-factor, multi-factor, nonlinear regression models, as well as models using time series with seasonal fluctuations.

Application area. The results of the study can be widely used in applied economic problems.