Московский институт международных экономических отношений
Специальность: "Бухгалтерский учет, анализ и аудит"
Курсовая работа
по учебной дисциплине: "Эконометрика"
на тему: "Модель множественной регрессии"
Выполнила: Кристостурян Светлана Саркисовна
Москва - 2006
Содержание

Введение

Экономисты используют количественные данные для наблюдения за ходом развития экономики, ее анализа и прогнозов. Набор статистических методов, используемых для этих целей, называется в совокупности эконометрикой. Для успешного применения этих методов требуется точное (или хотя бы приблизительно верное) моделирование поведения экономических агентов, необходимо также понимание процессов, породивших имеющиеся данные, и насколько эти данные отражают те явления, которые мы пытаемся исследовать.
Эконометрика как наука расположена где-то между экономикой, статистикой и математикой. Один из ответов на вопрос, что такое эконометрика, может звучать так: это наука, связанная с эмпирическим выводом экономических законов. То есть мы используем данные или "наблюдения" для того, чтобы получить количественные зависимости для экономических соотношений. Данные, как правило, не являются

экспериментальными, так как в экономике мы не можем проводить (многократные) эксперименты.
Но это - только малая часть работы эконометриста. Он также формулирует экономические модели, основываясь на экономической теории или на эмпирических данных, оценивает неизвестные величины (параметры) в этих моделях, делает прогнозы (и оценивает их точность) и дает рекомендации по экономической политике.
Поскольку наши модели неполны, а данные несовершенны, значительная часть эконометрики посвящена методам; которые могли бы работать с такими моделями и данными. В конце концов, эконометрика является не более чем набором инструментов, хотя и очень полезных. Качество ингредиентов (моделей и данных) и то, как мы их используем, определяют результаты нашего анализа. Но и хорошие инструменты анализа также необходимы. Эконометрика является одновременно нашим телескопом и нашим микроскопом для изучения окружающего экономического мира.

1. Основные положения регрессионного анализа

Если зависимость между двумя переменными такова, что каждому значению одной переменной соответствует определенное условное математическое ожидание (среднее значение) другой, то такая статистическая зависимость называется корреляционной.
Иначе, корреляционной зависимостью между двумя переменными называется функциональная зависимость между значениями одной из них и условным математическим ожиданием другой.
Корреляционная зависимость может быть представлена в виде:
(1)
где
В регрессионном анализе рассматриваются односторонняя зависимость случайной переменной Y от одной (или нескольких) неслучайной независимой переменной X. При этом зависимую переменную Y называют также функцией отклика, объясняемой, выходной, результирующей, эндогенной переменной, результативным признаком, а независимую переменную X - объясняющей, входной, предсказывающей, предикторной, экзогенной, переменной, фактором, регрессором, факторным признаком.
Уравнение (1) называется модельным уравнением регрессии (или просто уравнением регрессии), а функция модельной функцией регрессии (или просто функцией регрессии), а ее график - модельной линией регрессии (или просто линией регрессии).
В силу воздействия неучтенных случайных факторов и причин отдельные наблюдения переменной Y будут в большей или меньшей мере отклоняться от функции регрессии . В этом случае уравнение взаимосвязи двух переменных (парная регрессионная модель) может быть представлено в виде:
где - случайная переменная (случайный член), характеризующая отклонение от

функции регрессии. Эту переменную будем называть возмущающей или просто возмущением (либо ошибкой). Таким образом, в регрессионной модели зависимая переменная Y есть некоторая функция с точностью до случайного возмущения .
Рассмотрим линейный регрессионный анализ, для которого функции линейна относительно оцениваемых параметров:
(2)

Предположим, что для оценки параметров линейной функции регрессии (2) взята выборка, содержащая п пар значении переменных , где I = 1, 2,..., п. В этом случае линейная парная регрессионная модель имеет вид:
(3)

Отметим основные предпосылки регрессионного анализа.
1. В модели (3) возмущение (или зависимая переменная) есть величина случайная, а объясняющая переменная - величина неслучайная.
2. Математическое ожидание возмущения равно нулю:
(или математическое ожидание зависимой переменной - равно линейной функции регрессии:
.
3. Дисперсия возмущения постоянна для любого i:
,
(или) - условие гомоскедастичности или равноизменчивости возмущения (зависимой переменной)).
4. Возмущения и (или переменные) не коррелированы:
.
5. Возмущение есть нормально распределенная случайная величина.
В этом случае модель (3) называется классической нормальной линейной регрессионной моделью.
6. Векторы значений объясняющих переменных, или столбцы матрицы плана X, должны быть линейно независимыми, т. е. ранг матрицы X - максимальный:
r (X)=p+1.

Кроме того, полагают, что число имеющихся наблюдений (значений) каждой из объясняющих и зависимой переменных превосходит ранг матрицы X, т. е. n > r (X) или п >p+1, ибо в противном случае в принципе невозможно получение сколько-нибудь надежных статистических выводов.
Воздействие неучтенных случайных факторов и ошибок наблюдений в модели (3) определяется с помощью дисперсии возмущений (ошибок) или остаточной дисперсии . Несмещенной оценкой этой дисперсии является выборочная остаточная дисперсия.

где - групповая средняя, найденная по уравнению регрессии;
- выборочная оценка возмущения или остаток регрессии.

2. Модель множественной регрессии

2.1 Классическая нормальная линейная модель множественной регрессии. Основные гипотезы

Экономические явления, как правило, определяются большим числом одновременно и совокупно действующих факторов. В связи с этим часто возникает задача исследования зависимости одной зависимой переменной Y от нескольких объясняющих переменных Эта задача решается с помощью множественного регрессионного анализа.
Обозначим i-е наблюдение зависимой переменной а объясняющих переменных -
Тогда модель множественной линейной регрессии можно представить в виде:
 (4)

где i = 1, 2,..., n; удовлетворяет приведенным выше предпосылкам 2 и 4 регрессионного анализа.
Модель (4), в которой зависимая переменная возмущения и объясняющие переменные удовлетворяют предпосылкам регрессионного анализа, называется классической нормальной линейной моделью множественной регрессии.
Основные гипотезы лежащие в основе модели множественной регрессии:
i = 1, 2,..., n; п - спецификация модели;
 - детерминированные величины. Векторы линейно независимы в .
 - не зависит от i.
,
при - статистическая независимость (некоррелированность) ошибок для разных наблюдений.
Включение в регрессионную модель новых объясняющих переменных усложняет получаемые формулы и вычисления. Это приводит к целесообразности использования матричных обозначений. Матричное описание регрессии облегчает как теоретические концепции анализа, так и необходимые расчетные процедуры.
Введем обозначения: - матрица-столбец, или вектор, значений зависимой переменной размера ;
- матрица значений объясняющих переменных, или матрица плана размера:
;
 - матрица-столбец, или вектор, параметров размера ;
 - матрица-столбец, или вектор, возмущений (случайных ошибок, остатков) размера п.
Тогда в матричной форме модель (4) примет вид:
.
Оценкой этой модели по выборке является уравнение:
Основные гипотезы лежащие в основе модели множественной регрессии в матричной записи выглядят следующим образом:
- спецификация модели;

- X - детерминированная матрица, имеет максимальный ранг p;

.

2.2 Метод наименьших квадратов. Теорема Гаусса-Маркова

Для оценки вектора неизвестных параметров применим метод наименьших квадратов. Целью метода является выбор вектора оценок b, минимизирующего сумму квадратов остатков (т. е. квадрат длины вектора остатков e). Так как произведение транспонированной матрицы на саму матрицу e

то условие минимизации остаточной суммы квадратов запишется в виде:
 (5)

Учитывая, что при транспонировании произведения матриц получается произведение транспонированных матриц, взятых в обратном порядке, т.е. (Xb)'=b'X;

после раскрытия скобок получим:
Произведение Y'Xb есть матрица размера

.

Поэтому условие минимизации (5) примет вид:
На основании необходимого условия экстремума функции нескольких переменных представляющей (5), необходимо приравнять нулю частные производные по этим переменным или в матричной форме - вектор частных производных:
Для вектора частных производных доказаны следующие формулы:
где b и c - вектор столбцы; A - симметрическая матрица, в которой элементы, расположенные симметрично относительно главной диагонали, равны.
Поэтому, полагая c = X'Y, а матрицу A = X'X (она является симметрической), найдем:
откуда получаем систему нормальных уравнений в матричной форме для определения вектора b:
 (6)

Решением этого уравнения является вектор:
 (7)

где - матрица, обратная матрице коэффициентов системы (6), - матрица-столбец, или вектор, ее свободных членов.
Теорема Гаусса-Маркова. Предположим, что:

;

X - детерминированная матрица, имеющая максимальный ранг p;

.

Тогда оценка метода наименьших квадратов

является наиболее эффективной (в смысле наименьшей дисперсии) оценкой в классе линейных несмещенных оценок.
Доказательство. Оценка метода наименьших квадратов есть несмещенная оценка для вектора , т. е. . Любую другую оценку вектора без ограничения общности можно представить в виде:

где С - некоторая матрица размера:
.
Так как рассматриваемые в теореме оценки относятся к классу несмещенных оценок, то и или
.
Учитывая, что матрица в квадратных скобках - неслучайная, а в силу предпосылки 2 регрессионного анализа , получим:
откуда следует, что СХ = 0.
Далее
Так как СХ = 0,
Теперь с помощью преобразований, найдем, что ковариационная матрица вектора оценок , т. е.
Диагональные элементы матрицы неотрицательны, ибо они равны суммам квадратов элементов соответствующих строк этой матрицы. А так как диагональные элементы матриц и есть дисперсии компонент векторов оценок и , то дисперсия:
.
Это означает, что оценки коэффициентов регрессии, найденных методом наименьших квадратов, обладают наименьшей дисперсией, что и требовалось доказать.
На практике часто бывает необходимо сравнение влияния на зависимую переменную различных объясняющих переменных, когда последние выражаются разными единицами измерения. В этом случае используют стандартизированные коэффициенты регрессии и коэффициенты эластичности :

Стандартизированный коэффициент регрессии показывает, на сколько величин изменится в среднем зависимая переменная Y при увеличении только j-й объясняющей переменной на , а коэффициент эластичности - на сколько процентов (от средней) изменится в среднем Y при увеличении только на 1 %.
2.3 Коэффициенты детерминации и

В модели множественной регрессии общая вариация Q - сумма квадратов

отклонений зависимой переменной от средней может быть разложена на две составляющие:
где - соответственно сумма квадратов отклонений, обусловленная регрессией, и остаточная сумма квадратов, характеризующая влияние неучтенных факторов.
Уравнение множественной регрессии значимо (иначе - гипотеза о равенстве нулю параметров регрессионной модели, т. е. : отвергается), если:
т = р+ 1,
(8)

где - табличное значение F-критерия Фишера-Снедекора.
Коэффициент детерминации R2 - одна из наиболее эффективных оценок адекватности регрессионной модели, мера качества уравнения регрессии, характеристика его прогностической силы.
Коэффициент детерминации (или множественный коэффициент детерминации) R2 определяется по формуле:
(9)

R2 характеризует долю вариации зависимой переменной, обусловленной регрессией или изменчивостью объясняющих переменных; чем ближе R2 к единице, тем лучше регрессия описывает зависимость между объясняющими и зависимой переменными.
В какой степени допустимо использовать критерий R2 для выбора между несколькими регрессионными уравнениями? Следующие два замечания побуждают не полагаться только на значение R2.
1. R2, вообще говоря, возрастает при добавлении еще одного регрессора.
2. R2 изменяется даже при простейшем преобразовании зависимой переменной, поэтому сравнивать по значению R2 можно только регрессии с одинаковыми зависимыми переменными.
Если взять число регрессоров равным числу наблюдений, всегда можно добиться того, что R2 = 1, но это вовсе не будет означать наличие содержательной (имеющей экономический смысл) зависимости у от регрессоров. регрессионный анализ множественная мультиколлинеарность
Вместе с тем использование только одного коэффициента детерминации R2 для выбора наилучшего уравнения регрессии может оказаться недостаточным. На практике встречаются случаи, когда плохо определенная модель регрессии может дать сравнительно высокий коэффициент R2.
Недостатком коэффициента детерминации R2 является то, что он, вообще говоря, увеличивается при добавлении новых объясняющих переменных, хотя это и не обязательно означает улучшение качества регрессионной модели. В этом смысле предпочтительнее использовать скорректированный (адаптированный, поправленный) коэффициент детерминации , определяемый по формуле:
или с учетом (9)
(10)

Скорректированный коэффициент обладает следующими свойствами:
но может принимать значения < 0.
В определенной степени использование скорректированного коэффициента детерминации более корректно для сравнения регрессий при изменении количества регрессоров.
Из формулы (10) следует, что чем больше число объясняющих переменных p, тем меньше по сравнению с R2. В отличие от R2 скорректированный коэффициент может уменьшаться при введении в модель новых объясняющих переменных, не оказывающих существенного влияния на зависимую переменную. Однако даже увеличение скорректированного коэффициента детерминации при введении в модель новой объясняющей переменной не всегда означает, что ее коэффициент регрессии значим (это происходит только в случае, если соответствующее значение t-статистики больше единицы (по абсолютной величине)), т. е. Другими словами, увеличение еще не означает улучшения качества регрессионной модели.
Если известен коэффициент детерминации R2, то критерий значимости (8) уравнения регрессии может быть записан в виде:
,
ибо в уравнении множественной регрессии вместе со свободным членом оценивается т = p+1 параметров.
2.4 Обобщенная линейная модель множественной регрессии

При моделировании реальных экономических процессов мы нередко сталкиваемся с ситуациями, в которых условия классической линейной модели регрессии оказываются нарушенными. В частности, могут не выполняться предпосылки 3 и 4 регрессионного анализа о том, что случайные возмущения (ошибки) модели имеют постоянную дисперсию и не коррелированы между собой. Для линейной множественной модели эти предпосылки означают, что ковариационная матрица вектора возмущений (ошибок) имеет вид:
.
В тех случаях, когда имеющиеся статистические данные достаточно однородны, это допущение вполне оправдано. Однако в других ситуациях оно может оказаться неприемлемым.
Обобщенная линейная модель множественной регрессии:

, (11)

в которой переменные и параметры определены следующей системой соотношений и условий:
 - случайный вектор; X - неслучайная (детерминированная) матрица;
;
,
где - положительно определенная матрица;
где p - число объясняющих переменных; п - число наблюдений.

Сравнивая обобщенную модель с классической, видим, что она отличается от классической только видом ковариационной матрицы: вместо
для классической модели имеем для обобщенной. Это означает, что в отличие от классической, в обобщенной модели ковариации и дисперсии объясняющих переменных могут быть произвольными. В этом состоит суть обобщения регрессионной модели.
Для оценки параметров модели (11) можно применить обычный метод наименьших квадратов.
Оценка
полученная ранее и определенная соотношением (7), остается справедливой и в случае обобщенной модели. Оценка b по-прежнему несмещенная и состоятельная. Однако полученная ранее формула для ковариационной матрицы вектора оценок оказывается неприемлемой в условиях обобщенной модели. Это означает, что обычный метод наименьших квадратов в обобщенной линейной регрессионной модели дает смещенную оценку ковариационной матрицы вектора оценок b.
2.5 Обобщенный метод наименьших квадратов

Вопрос об эффективности линейной несмещенной оценки вектора для обобщенной регрессионной модели решается с помощью следующей теоремы.
Теорема Айткена. В классе линейных несмещенных оценок вектора для обобщенной регрессионной модели оценка:
 (12)

имеет наименьшую ковариационную матрицу.
Доказательство. Убедимся в том, что оценка является несмещенной. Учитывая (12), представим ее в виде:
Математическое ожидание оценки , т. е. , ибо , т. е. оценка есть несмещенная оценка .
Для доказательства оптимальных свойств преобразуем исходные данные - матрицу X, вектор Y и возмущение к виду, при котором выполнены требования классической модели регрессии.
Из матричной алгебры известно, что всякая невырожденная симметричная матрица A допускает представление в виде , где P - некоторая невырожденная матрица.
Поэтому существует такая невырожденная матрица P, что:
 (13)

Учитывая свойства обратных квадратичных матриц, т. е.
 и ,
(14)
Следует заметить, что если обе части равенства (13) умножить слева на матрицу , а справа - на матрицу , то в произведении получим единичную матрицу.
Действительно,
 (15)
Теперь, умножив обе части обобщенной регрессионной модели

на матрицу слева, получим:
 (16)
 (17)

Возвращаясь к исходным наблюдениям X и Y и учитывая (14), получим:
т. е. выражение (12), что и требовалось доказать.
Нетрудно проверить, что в случае классической модели, оценка обобщенного метода наименьших квадратов (12) совпадает с оценкой "обычного" метода b (7).
Следует отметить, что для обобщенной регрессионной модели, в отличие от классической, коэффициент детерминации, вычисленный по формуле:
,
где - оценка обобщенного метода наименьших квадратов), не является удовлетворительной мерой качества модели. В общем случае может выходить даже за пределы интервала , а добавление (удаление) объясняющей переменной не обязательно приводит к его увеличению (уменьшению).
Поэтому коэффициент детерминации в обобщенной модели может использоваться лишь как весьма приближенная характеристика качества модели.
Подчеркнем еще раз, что для применения обобщенного метода наименьших квадратов (ОМНК) необходимо знать матрицу , что на практике бывает крайне редко.
Поэтому вполне естественным кажется такой способ: оценить (каким-нибудь образом) матрицу , а затем использовать эту оценку в формуле (12) вместо . Этот подход составляет суть так называемого доступного обобщенного метода наименьших квадратов.
Если же считать все элементов симметричной ковариационной матрицы неизвестными параметрами обобщенной модели, то общее число параметров значительно превысит число наблюдений п, что сделает оценку этих параметров неразрешимой задачей. Поэтому для практической реализации ОМНК необходимо вводить дополнительные условия на структуру матрицы . Так мы приходим к практически реализуемому (или доступному) ОМНК.

3. Различные аспекты множественной регрессии

3.1 Мультиколлинеарность

Под мультиколлинеарностью понимается высокая взаимная коррелированность объясняющих переменных. Мультиколлинеарность может проявляться в функциональной (явной) и стохастической (скрытой) формах.
При функциональной форме мультиколлинеарности, по крайней мере, одна из парных связей между объясняющими переменными является линейной функциональной зависимостью. В этом случае матрица особенная, так как содержит линейно зависимые векторы-столбцы и ее определитель равен нулю, т. е. нарушается предпосылка 6 регрессионного анализа. Это приводит к невозможности решения соответствующей системы нормальных уравнений и получения оценок параметров регрессионной модели.

Однако в экономических исследованиях мультиколлинеарность чаще проявляется в стохастической форме, когда между хотя бы двумя объясняющими переменными существует тесная корреляционная связь. Матрица в этом случае является неособенной, но ее определитель очень мал.
Выделим некоторые наиболее характерные признаки мультиколлинеарности.
1. Небольшое изменение исходных данных (например, добавление новых наблюдений) приводит к существенному изменению оценок коэффициентов модели.
2. Оценки имеют большие стандартные ошибки, малую значимость, в то время как модель в целом является значимой (высокое значение коэффициента детерминации R2 и соответствующей F-статистики).
3. Оценки коэффициентов имеют неправильные с точки зрения теории знаки или неоправданно большие значения.
Точных количественных критериев для определения наличия или отсутствия мультиколлинеарности не существует. Тем не менее, имеются некоторые эвристические подходы по ее выявлению.
Один из таких подходов заключается в анализе корреляционной матрицы между объясняющими переменными и выявлении пар переменных, имеющих высокие коэффициенты корреляции (обычно больше 0,8). Если такие переменные существуют, то говорят о мультиколлинеарности между ними.
Полезно также находить множественные коэффициенты детерминации между одной из объясняющих переменных и некоторой группой из них. Наличие высокого множественного коэффициента детерминации (обычно больше 0,6) свидетельствует о мультиколлинеарности.
Что же делать, если по всем признакам имеется мультиколлинеарность? Однозначного ответа на этот вопрос нет, и среди эконометристов есть разные мнения на этот счет. Существует даже такая школа, представители которой считают, что и не нужно ничего делать, поскольку "так устроен мир" (см. Kennedy, 1992). Однако для устранения или уменьшения мультиколлинеарности используется ряд методов. Самый простой из них (но далеко не всегда возможный) состоит в том, что из двух объясняющих переменных, имеющих высокий коэффициент корреляции (больше 0,8), одну переменную исключают из рассмотрения. При этом какую переменную оставить, а какую удалить из анализа, решают в первую очередь на основании экономических соображений. Если с экономической точки зрения ни одной из переменных нельзя отдать предпочтение, то оставляют ту из двух переменных, которая имеет 1 больший коэффициент корреляции с зависимой переменной.
Другой метод устранения или уменьшения мультиколлинеарности заключается в переходе от несмещенных оценок, определенных по методу наименьших квадратов, к смещенным оценкам, обладающим, однако, меньшим рассеянием относительно оцениваемого параметра, т. е. меньшим математическим ожиданием квадрата отклонения оценки от параметра или .
Оценки, определяемые вектором (7), обладают в соответствии с теоремой Гаусса-Маркова минимальными дисперсиями в классе всех линейных несмещенных оценок,

но при наличии мультиколлинеарности эти дисперсии могут оказаться слишком большими, и обращение к соответствующим смещенным оценкам может повысить точность оценивания параметров регрессии.
Для устранения мультиколлинеарности может быть использован переход от исходных объясняющих переменных связанных между собой достаточно тесной корреляционной зависимостью, к новым переменным, представляющим линейные комбинации исходных. При этом новые переменные должны быть слабокоррелированными либо вообще некоррелированными. В качестве таких переменных берут, например, так называемые главные компоненты вектора исходных объясняющих переменных, изучаемые в компонентном анализе, и рассматривают регрессию на главных компонентах, в которой последние выступают в качестве обобщенных объясняющих переменных, подлежащих в дальнейшем содержательной (экономической) интерпретации.
Ортогональность главных компонент предотвращает проявление эффекта мультиколлинеарности. Кроме того, применяемый метод позволяет ограничиться малым числом главных компонент при сравнительно большом количестве исходных объясняющих переменных.
Еще одним из возможных методов устранения или уменьшения мультиколлинеарности является использование пошаговых процедур отбора наиболее информативных переменных. Например, на первом шаге рассматривается лишь одна объясняющая переменная, имеющая с зависимой переменной Y наибольший коэффициент детерминации. На втором шаге включается в регрессию новая объясняющая переменная, которая вместе с первоначально отобранной образует пару объясняющих переменных, имеющую с Y наиболее высокий (скорректированный) коэффициент детерминации. На третьем шаге вводится в регрессию еще одна объясняющая переменная, которая вместе с двумя первоначально отобранными образует тройку объясняющих переменных, имеющую с Y наибольший (скорректированный) коэффициент детерминации, и т. д.
Процедура введения новых переменных продолжается до тех пор, пока будет увеличиваться соответствующий (скорректированный) коэффициент детерминации (более точно - минимальное значение).

3.2 Фиктивные переменные

До сих пор мы рассматривали регрессионную модель, в которой в качестве объясняющих переменных (регрессоров) выступали количественные переменные (производительность труда, себестоимость продукции, доход и т. п.). Однако на практике достаточно часто возникает необходимость исследования влияния качественных признаков, имеющих два или несколько уровней (градаций). К числу таких признаков можно отнести: пол (мужской, женский), образование (начальное, среднее, высшее), фактор сезонности (зима, весна, лето, осень) и т. п.
Качественные признаки могут существенно влиять на структуру линейных связей между переменными и приводить к скачкообразному изменению параметров регрессионной модели. В этом случае говорят об исследовании регрессионных

моделей с переменной структурой или построении регрессионных моделей по неоднородным данным.
Например, нам надо изучить зависимость размера заработной платы Y работников не только от количественных факторов , но и от качественного признака (например, фактора "пол работника").
В принципе можно было получить оценки регрессионной модели:
 (18)

для каждого уровня качественного признака (т. е. выборочное уравнение регрессии отдельно для работников-мужчин и отдельно - для женщин), а затем изучать различия между ними.
Есть и другой подход, позволяющий оценивать влияние значений количественных переменных и уровней качественных признаков с помощью одного уравнения регрессии. Этот подход связан с введением так называемых фиктивных (манекенных) переменных, или манекенов.
В качестве фиктивных переменных обычно используются дихотомические (бинарные, булевы) переменные, которые принимают всего два значения: "0" или "1" (например, значение такой переменной по фактору "пол": для работников-женщин и - для мужчин).
В этом случае первоначальная регрессионная модель (18) заработной платы изменится и примет вид:
 (19)

Таким образом, принимая модель (19), мы считаем, что средняя заработная плата у мужчин на выше, чем у женщин, при неизменных значениях других параметров модели. А проверяя гипотезу , мы можем установить существенность влияния фактора "пол" на размер заработной платы работника.
Следует отметить, что в принципе качественное различие можно формализовать с помощью любой переменной, принимающей два разных значения, не обязательно "0" или "1". Однако в экономической практике почти всегда используются фиктивные переменные типа "0-1", так как при этом интерпретация полученных результатов выглядит наиболее просто.
Рассматриваемая выше регрессионная модель (19) и отражает влияние качественного признака (фиктивных переменных) только на значения переменной Y, т. е. на свободный член уравнения регрессии. В более сложных моделях может быть отражена также зависимость фиктивных переменных на сами параметры при переменных регрессионной модели.
Таким образом:
1. Для исследования влияния качественных признаков в модель можно вводить фиктивные переменные, которые, как правило, принимают значение "1", если данный качественный признак присутствует в наблюдении, и значение "0" при его отсутствии;
2. Способ включения фиктивных переменных зависит от априорной информации

относительно влияния соответствующих качественных признаков на зависимую переменную и от гипотез, которые проверяются с помощью модели;
3. От способа включения фиктивной переменной зависит и интерпретация оценки коэффициента при ней.

3.3 Частная корреляция

В том случае, когда имеются одна независимая и одна зависимая переменные, естественной мерой зависимости (в рамках линейного подхода) является (выборочный) коэффициент корреляции между ними. Использование множественной регрессии позволяет обобщить это понятие на случай, когда имеется несколько независимых переменных. Корректировка здесь необходима по следующим очевидным соображениям. Высокое значение коэффициента корреляции между исследуемой зависимой и какой-либо независимой переменной может, как и раньше, означать высокую степень зависимости, но может быть обусловлено и другой причиной. А именно, есть третья переменная, которая оказывает сильное влияние на две первые, что и служит, в конечном счете, причиной их высокой коррелированности. Поэтому возникает естественная задача найти "чистую" корреляцию между двумя переменными, исключив (линейное) влияние других факторов. Это можно сделать с помощью коэффициента частной корреляции.
Выборочным частным коэффициентом корреляции (или просто частным коэффициентом корреляции) между переменными Xi и Xj при фиксированных значениях остальных (p - 2) переменных называется выражение:

(20)
где и - алгебраические дополнения элементов и матрицы выборочных коэффициентов корреляции:
В частности, в случае трех переменных (п=3) из (20) следует, что:
Частный коэффициент корреляции , как и парный коэффициент, может принимать значения от -1 до +1. Кроме того, , вычисленный на основе выборки объема п, имеет такое же распределение, как и , вычисленный по наблюдениям. Поэтому значимость частного коэффициента корреляции оценивают так же, как и обычного коэффициента корреляции r, но при этом полагают

Заключение

В практике экономических исследований имеющиеся данные не всегда можно считать выборкой из многомерной нормальной совокупности, когда одна из рассматриваемых переменных не является случайной или когда линия регрессии явно не прямая и т. п. В этих случаях пытаются определить кривую (поверхность), которая дает наилучшее (в смысле метода наименьших квадратов) приближение к исходным данным. Соответствующие методы приближения получили название регрессионного анализа.
Задачами регрессионного анализа являются установление формы зависимости между переменными, оценка функции регрессии, оценка неизвестных значений

(прогноз значений) зависимой переменной.
Экономические явления, как правило, определяются большим числом одновременно и совокупно действующих факторов. В связи с этим часто возникает задача исследования зависимости одной зависимой переменной Y от нескольких объясняющих переменных Эта задача решается с помощью множественного регрессионного анализа.
В своей работе я постаралась показать основные направления изучения классической линейной модели множественной регрессии. Привела оценки параметров модели и проверку статистических гипотез о регрессии.
Так же мною были раскрыты некоторые проблемы, связанные с практическим использованием модели множественной регрессии. К ним относятся: мультиколлинеарность, ее причины и методы устранения; использование фиктивных переменных при включении в регрессионную модель качественных объясняющих переменных; вопросы частной корреляции между переменными.
В последние десятилетия эконометрика как научная дисциплина стремительно развивается. Растет число научных публикаций и исследований с применением эконометрических методов. Свидетельством всемирного признания эконометрики является присуждение за наиболее выдающиеся разработки в этой области Нобелевских премий по экономике Р. Фришу и Я. Тинбергу (1969), Л. Клейну (1980), Т. Хаавельмо (1989), Дж. Хекману и Д. Макфаддену (2000).
Язык эконометрики все больше становится языком математики, а экономику все чаще называют одной из наиболее математизированных наук.
Список использованной литературы

1. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов / Под ред. проф. Н.Ш. Кремера. - М.: ЮНИТИ-ДАНА, 2003. - 311 с.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. - 4-е изд. - М.: Дело, 2000. - 400 с.
3. Колемаев В.А., Староверов О.В., Турундаевский В.Б. "Теория вероятностей и математическая статистика" / М., 1991.
4. Головина Л.И. Линейная алгебра и некоторые ее приложения. - М.: Наука, 1985.
5. Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа: Пер. с нем. - М.: Финансы и статистика, 1982.