Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования
Санкт-Петербургский горный университет

Кафедра экономики, учета и финансов

ЭКОНОМЕТРИКА (ПРОДВИНУТЫЙ УРОВЕНЬ)

Методические указания к лабораторным работам для студентов магистратуры направления 38.04.01

САНКТ-ПЕТЕРБУРГ
2020

УДК 330.43 (073)

ЭКОНОМЕТРИКА (ПРОДВИНУТЫЙ УРОВЕНЬ): Методические указания к лабораторным работам / Санкт-Петербургский горный университет. Сост. *Н.В. Василенко.* СПб, 2020. 40 с.

В методических указаниях приведены общие требования, критерии оценки, а также рекомендации по выполнению пяти лабораторных работ, охватывающих основные темы дисциплины «Эконометрика (продвинутый уровень)».

Предназначены для студентов магистратуры направления 38.04.01 «Экономика».

Научный редактор проф. *И.Б. Сергеев*

Рецензент доц. *И.Ю. Парик* (ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»)

ЭКОНОМЕТРИКА
(ПРОДВИНУТЫЙ УРОВЕНЬ)

Методические указания к лабораторным работам
для студентов магистратуры направления 38.04.01

Сост. *Н.В. Василенко*

Печатается с оригинал-макета, подготовленного кафедрой
экономики, учета и финансов

Ответственный за выпуск *Н.В. Василенко*

Лицензия ИД № 06517 от 09.01.2002

Подписано к печати 29.06.2020. Формат 60×84/16.
Усл. печ. л. 2,2. Усл.кр.-отт. 2,2. Уч.-изд.л. 2,1. Тираж 50 экз. Заказ 421.

Санкт-Петербургский горный университет
РИЦ Санкт-Петербургского горного университета
Адрес университета и РИЦ: 199106 Санкт-Петербург, 21-я линия, 2

ВВЕДЕНИЕ

Лабораторные работы составляют важную часть профессиональной подготовки студентов. Они направлены на экспериментальное подтверждение теоретических положений и формирование учебных и профессиональных практических умений.

Лабораторные работы по дисциплине «Эконометрика (продвинутый уровень)» представляют собой особый вид аудиторной работы, направленный на обучение студентов практическим навыкам использования эконометрических методов при построении регрессионных моделей, содержательного обоснования полученных моделей регрессии для принятия управленческих решений, применения пакетов прикладных программ.

Трудоемкость выполнения лабораторных работ, согласно учебному плану по магистерской программе «Экономика и управление на предприятиях минерально-сырьевого комплекса» (направление подготовки 38.04.01 «Экономика»), составляет 33 часа, которые включают 21 час аудиторной и 12 часов самостоятельной работы.

Выполнение лабораторных работ по дисциплине «Эконометрика (продвинутый уровень)» вносит ***вклад в формирование следующих компетенций***:

– способность обосновывать актуальность, теоретическую и практическую значимость избранной темы научного исследования (ПК-2);

– способность проводить самостоятельные исследования в соответствии с разработанной программой (ПК-3);

– способность использовать цифровые и компьютерные технологии в организации научно-исследовательских и опытно-конструкторских работ (ПСК-3).

Целью лабораторных работ является анализ эконометрических проблем, выработка умения применять современный математический инструментарий для решения содержательных задач, строить эконометрические модели экономических объектов и процессов, проводить собственные научные исследования и публично излагать полученные результаты.

Выполняя лабораторные работы, студенты развивают умения представлять экономическую задачу в конкретной параметрической форме; получать оценки параметров эконометрической модели и проверять их качество; проводить отбор факторов с целью улучшения спецификации модели; проводить отбор адекватной модели из возможных вариантов; владеть приемами преобразования данных; экономически интерпретировать полученную эконометрическую модель.

В методических рекомендация изложены подходы к выполнению пяти лабораторных работ, охватывающих основные подходы к эконометрическому моделированию пространственных данных:

– ***№1. Парная линейная регрессия*** – приобретение навыка построения и оценки параметров модели парной линейной регрессии, а также анализа ее качества, в том числе при помощи коэффициентов парной линейной корреляции, детерминации, средней ошибки аппроксимации и т.д.;

– ***№2. Множественная линейная регрессия*** – приобретение навыка построения и оценки параметров модели множественной линейной регрессии, а также анализа ее качества, в том числе при помощи парных, частных и множественных коэффициентов корреляции, скорректированного коэффициента детерминации и т.д.;

– ***№3. Множественная линейная регрессия*** – определение набора переменных – приобретение навыка улучшения эконометрической модели на основе уточнения включаемых в нее факторов;

– ***№4. Замена переменных (логарифмические модели)*** – приобретение навыка улучшения эконометрической модели посредством замены переменных при помощи логарифмирования линейной регрессии;

– ***№5. Фиктивные переменные*** – приобретение навыка улучшения эконометрической модели на основе учета качественной неоднородности данных.

Использование предлагаемых методических рекомендаций поможет студентам не только успешно освоить пройденный материал и подготовиться к зачету, но и осмыслить возможности применения эконометрического инструментария для проведения самостоятельных научных исследований.

1. ОБЩИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ И ОФОРМЛЕНИЮ ЛАБОРАТОРНЫХ РАБОТ

Необходимыми структурными элементами лабораторной работы являются:

– самостоятельная деятельность студентов по подготовке к лабораторной работе, нацеленная на повторение основных понятий и методов экономического моделирования;

– проверка теоретической готовности студентов к выполнению задания;

– инструктаж, проводимый преподавателем, включающий содержательную постановку задачи с выдачей задания, а также требования техники безопасности при использовании студентом компьютерного оборудования;

– собственно выполнение студентом лабораторной работы под руководством преподавателя в компьютерном классе;

– оформление студентом отчета о выполнении лабораторной работы;

– обсуждение (защита) итогов выполнения лабораторной работы, в ходе которого преподаватель осуществляет контроль понимания выполненных студентом расчетов, качества интерпретации полученных результатов и хода формирования компетенций по дисциплине.

Отчет о выполнении лабораторной работы оформляется на отдельных листах формата А4 и обязательно должен содержать:

1. Титульный лист.

Титульный лист лабораторной работы должен содержать все необходимые реквизиты: названия университета и кафедры, наименование учебной дисциплины, номер группы, Ф.И.О. студента и преподавателя.

2. Исходные данные.

Исходные данные выдаются вместе с заданием и, как правило, представляются в табличной форме.

3. Результаты эконометрических расчетов.

Статистические и эконометрические расчеты должны быть выполнены с помощью пакета Microsoft Excel.

Результаты расчетов соответствующих показателей по каждой теме лабораторной работы должны быть представлены на бумажном носителе и сопровождаться аналитической запиской с эконометрическим анализом расчетов и экономическими выводами. В выводах обязательно указывать единицы измерения полученных показателей.

4. Основные экономические выводы в соответствии с целью лабораторной работы.

По итогам обсуждения (защиты) лабораторной работы студент может получить одну из двух оценок.

Оценка «незачтено» выставляется, если студент:

– не выполнил лабораторную работу в соответствии с заданием либо выполнил работу с существенными ошибками;

– не владеет либо слабо владеет теоретическими знаниями по изучаемой теме;

– при решении задач, предусмотренных в лабораторной работе, допускает неточности, существенные ошибки в выводах по модели, демонстрирует неумение интерпретировать результаты моделирования;

– необходимые практические компетенции не сформированы.

Оценка «зачтено» выставляется, если студент:

– выполнил курсовую работу полностью в соответствии с заданием или с некоторыми незначительными ошибками и неточностями;

– при защите курсовой работы демонстрирует хорошую теоретическую подготовку;

– успешно справляется с решением задач, предусмотренных в лабораторной работе, в том числе в точки зрения правильных результатов расчетов и экономической интерпретации результатов эконометрического моделирования;

– необходимые практические компетенции сформированы.

2. ЛАБОРАТОРНЫЕ РАБОТЫ

2.1. Парная линейная регрессия. Построение и оценка качестве модели

Исходные данные

При исследовании 8 магазинов получены следующие данные.

Таблица 1

Объем товарооборота и число работников в магазинах

Наблюдение	Объем товарооборота, млн. руб.	Число работников, чел.
1	0,3	74
2	0,6	87
3	1	103
4	1,2	119
5	1,4	123
6	1,7	128
7	1,9	137
8	2,1	151

Задание

Построить регрессионную модель зависимости объема товарооборота от числа работников. Проверить значимость модели и коэффициентов модели. Рассчитать коэффициент эластичности и дать ему экономическую интерпретацию. Построить 95% доверительный интервал для оценки объема товарооборота отдельного магазина со 100 работниками.

Решение

Поскольку необходимо определить **характер зависимости** объема товарооборота от числа работников, то принимаем:

– в качестве Y – ____________________,

– в качестве X – ____________________.

Уровень значимости устанавливаем как _______.

Строим и анализируем **поле корреляции** между Y и X.

Вывод: на основании анализа поля корреляции (*можно / нельзя*) выдвинуть гипотезу (для генеральной совокупности) о том, что связь между всеми возможными значениями X и Y носит *линейный характер*.

Оцениваем **коэффициенты уравнения парной линейной регрессии** (на основе метода наименьших квадратов), имеющего следующий вид

$$Y = \alpha + \beta X + \varepsilon \qquad (1)$$

Заполняем первые шесть столбцов *таблицы*, остальные - по мере накопления результатов расчетов.

Таблица 2

Результаты вспомогательных расчетов

i	X	Y	X^2	Y^2	XY	$Y(X)$	$(Y_i - \bar{Y})^2$	$(Y - Y(X))^2$	$(X_i - \bar{X})^2$	$\lvert Y - Y(X)\rvert : Y$
1										
2										
3										
4										
5										
6										
7										
8										
Σ										

Получаем значения эмпирических коэффициентов линейной регрессии, решая систему нормальных уравнений:

$$\begin{cases} an + b\sum X = \sum Y \\ a\sum X + b\sum X^2 = \sum YX \end{cases} \qquad (2)$$

Вывод: эмпирические коэффициенты регрессии:

b = __________,

a = __________.

Эмпирическое уравнение регрессии:

y =

Экономический смысл коэффициентов уравнения линейной регрессии:

1) b = _________ показывает среднее изменение результативного показателя (в единицах измерения *Y*) с повышением или понижением величины фактора x на единицу его измерения.

Вывод: с увеличением *X* на 1 единицу *Y* повышается в среднем на __.

2) a = ________формально показывает прогнозируемый уровень у, но только в том случае, если *X* = 0 находится близко с выборочными значениями.

Связь между *Y* и *X* определяет знак коэффициента регрессии b (если > 0 – прямая связь, иначе - обратная).

Вывод: В нашей задаче связь _______________.

Определяем **тесноту линейной связи** между признаками при помощи линейного **коэффициента корреляции.** Для этого рассчитываем**:**

– выборочные средние:

$$\bar{X} = \frac{\sum X_i}{n} \tag{3}$$

$$\bar{Y} = \frac{\sum Y_i}{n} \tag{4}$$

$$\overline{XY} = \frac{\sum X_i Y_i}{n} \tag{5}$$

– выборочные дисперсии:

$$S^2(X) = \frac{\sum {X^2}_i}{n} - \overline{X^2} \tag{6}$$

$$S^2(Y) = \frac{\sum {Y^2}_i}{n} - \overline{Y^2} \tag{7}$$

– среднеквадратическое отклонение

$$S(X) = \sqrt{S^2(X)} \tag{8}$$

$$S(Y) = \sqrt{S^2(Y)} \tag{9}$$

Коэффициент линейной парной корреляции может быть определен:

1 способ – через коэффициент регрессии *b*:

$$r_{x,y} = b\frac{S(X)}{S(Y)} \tag{10}$$

2 способ – по формуле:

$$r_{x,y} = \frac{\overline{XY} - \bar{X} \times \bar{Y}}{S(X) \times S(Y)} \tag{11}$$

Линейный коэффициент корреляции принимает значения от –1 до +1. Для оценки силы связи между признаками используем шкале Чеддока: $0{,}1 < r_{xy} < 0{,}3$ – слабая; $0{,}3 < r_{xy} < 0{,}5$ – умеренная; $0{,}5 < r_{xy} < 0{,}7$ – заметная; $0{,}7 < r_{xy} < 0{,}9$ – высокая; $0{,}9 < r_{xy} < 1$ – весьма высокая.

Вывод: В рассматриваемом примере линейная связь между признаком *Y* и фактором *X* ________________________________.

Вычисляем **коэффициент эластичности и бета-коэффициенты**.

Если существует различие единиц измерения результативного показателя *Y* и факторного признака *X*, коэффициенты регрессии (в нашей задаче b) нежелательно использовать для непосредственной оценки влияния факторов на результативный признак. Для этих целей вычисляются коэффициенты эластичности *E* и бета - коэффициенты.

Коэффициент эластичности рассчитываем по формуле:

$$E = \frac{\partial Y}{\partial X} \times \frac{X}{Y} = b\,\frac{\bar{X}}{\bar{Y}} \qquad (12)$$

Коэффициент эластичности *E* показывает, на сколько процентов в среднем по совокупности изменится результат *Y* от своей средней величины при изменении фактора *X* на 1% от своего среднего значения.

Вывод: В рассматриваемом примере коэффициент эластичности _______. Следовательно, при изменении *X* на 1%, *Y* изменится на ______, следовательно, *X* (*влияет / не влияет*) на *Y*.

Бета – коэффициент рассчитываем по формуле:

$$\beta_j = b_j\,\frac{S(X)}{S(Y)} \qquad (13)$$

Бета – коэффициент показывает, что увеличение *X* на величину среднеквадратического отклонения *S(X)* приведет к увеличению среднего значения *Y* на b_j среднеквадратичного отклонения *S(Y)*.

Вывод: В рассматриваемом примере__________________.

Оценим качество уравнения регрессии с помощью **ошибки аппроксимации**, показывающей среднее отклонение расчетных значений от фактических:

$$\bar{A} = \frac{\sum_i |Y_i - Y(X_i)| : Y_i}{n} \times 100\% \qquad (14)$$

Ошибка аппроксимации в пределах 5-7% свидетельствует о хорошем подборе уравнения регрессии к исходным данным.

Вывод: В рассматриваемой задаче ошибка _______________, поэтому данное уравнение (*можно/не следует*) использовать в качестве уравнения регрессии.

Вычислим **коэффициент детерминации.** Коэффициент детерминации представляет собой квадрат (множественного) коэффициента корреляции:

Множественный коэффициент корреляции:

$$R = \sqrt{1 - \frac{\sum_i (Y_i - Y(X_i))^2}{\sum_i (Y_i - \bar{Y})^2}} \tag{15}$$

Коэффициент детерминации показывает долю вариации результативного признака, объясненную вариацией факторного признака, и может быть выражен в процентах.

Вывод: В рассматриваемом примере _________% случаев изменения X приводят к изменению Y. Остальные ________% изменения Y объясняются факторами, не учтенными в модели.

Проанализируем **точность определения оценок коэффициентов регрессии.**

Несмещенной оценкой дисперсии возмущений является величина:

$$S(Y)^2 = \frac{\sum_i (Y_i - Y(X_i))^2}{n-m-1} \tag{16}$$

$S(Y)^2$ – необъясненная дисперсия (мера разброса зависимой переменной вокруг линии регрессии), m – количество переменных (факторов). Для парной линейной регрессии m=1.

Стандартная ошибка регрессии:

$$S_y = \sqrt{S(Y)^2} \tag{17}$$

S_a – стандартное отклонение случайной величины a:

$$S_a = S_y \frac{\sqrt{\sum X^2}}{nS(X)} \tag{18}$$

S_b – стандартное отклонение случайной величины b:

$$S_b = \frac{S_y}{\sqrt{n} \times S(X)} \tag{19}$$

Проверяем **гипотезы относительно коэффициентов линейного уравнения регрессии.**

t-статистика. Критерий Стьюдента. С помощью МНК мы получили лишь оценки параметров уравнения регрессии, которые характерны для конкретного статистического наблюдения (конкретного набора значений X и Y).

Для оценки статистической значимости коэффициентов регрессии и корреляции рассчитываются t-критерий Стьюдента и доверительные интервалы каждого из показателей.

В качестве основной (нулевой) гипотезы H_0 выдвигают гипотезу о незначимом отличии от нуля параметра или статистической характеристики в генеральной совокупности. Наряду с основной (проверяемой) гипотезой выдвигают альтернативную (конкурирующую) гипотезу H_1 о неравенстве нулю параметра или статистической характеристики в генеральной совокупности. В случае если основная гипотеза окажется неверной, мы принимаем альтернативную.

Проверим гипотезу H_0 о равенстве отдельных коэффициентов регрессии нулю (при альтернативе H_1 не равно) на уровне значимости $\alpha = 0.05$. Уровень значимости α – вероятность отвергнуть правильную гипотезу при условии, что она верна.

Для проверки этой гипотезы используется *t-критерий Стьюдента*. Найденное по данным наблюдений значение t-критерия (его еще называют наблюдаемым или фактическим) t_b сравнивается с табличным (критическим) значением $t_{крит}$, определяемым по таблицам распределения Стьюдента в зависимости от уровня значимости (α) и числа степеней свободы, которое в случае линейной парной регрессии равно (n - 2), где n – число наблюдений.

$$t_b = \frac{b}{S_b} \qquad (20)$$

$$t_a = \frac{a}{S_a} \qquad (21)$$

Если фактическое значение t-критерия *по модулю* больше табличного $t_{крит}$ (n-m-1; α/2), то основную гипотезу отвергают и считают, что с вероятностью (1-α) параметр или статистическая характеристика в генеральной совокупности значимо отличается от нуля.

Если фактическое значение t-критерия *по модулю* меньше табличного $t_{крит}$ (n-m-1;α/2), то нет оснований отвергать основную гипотезу, т.е. параметр или статистическая характеристика в генеральной совокупности незначимо отличается от нуля при уровне значимости α.

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии b (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии a (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Определим **доверительные интервалы для коэффициентов уравнения регрессии**.

С вероятностью 95% можно утверждать, что значение параметра b будут лежать в интервале: (b - $t_{крит}$ S_b; b + $t_{крит}$ S_b).

Вывод: В нашей задаче (________________; ________________)

С вероятностью 95% можно утверждать, что значение параметра b будут лежать в интервале: (a - $t_{крит}$ S_a; a + $t_{крит}$ S_a).

Вывод: В нашей задаче (________________; ________________)

***F*-статистика. Критерий Фишера.** Коэффициент детерминации R^2 используется для проверки существенности уравнения линейной регрессии в целом.

Проверка значимости модели регрессии проводится с использованием F-критерия Фишера, расчетное значение которого находится как отношение дисперсии исходного ряда наблюдений изучаемого показателя и несмещенной оценки дисперсии остаточной последовательности для данной модели. Если расчетное значение с k_1=(m) и k_2=(n-m-1) степенями свободы больше табличного при заданном уровне значимости, то модель считается значимой, где m – количество переменных в модели (для парной линейной регрессии m = 1).

Выдвигается нулевая гипотеза о том, что уравнение в целом статистически незначимо: H_0: R^2=0 на уровне значимости α.

Фактическое значение F-критерия определяется по формуле:

$$F = \frac{R^2(n-m-1)}{(1-R^2)m} \qquad (22)$$

Табличное значение $F_{табл}$ определяется по таблицам распределения Фишера для заданного уровня значимости, принимая во внимание, что число степеней свободы для общей суммы квадратов (большей дисперсии) равно 1 и число степеней свободы остаточной суммы квадратов (меньшей дисперсии) при линейной регрессии равно n-2.

$F_{табл}$ $(k_1; k_2)$ – это максимально возможное значение критерия под влиянием случайных факторов при данных степенях свободы и уровне значимости α. Если фактическое значение F-критерия меньше табличного $F_{табл}$, то считают, что нет основания отклонять нулевую гипотезу. Если фактическое значение F-критерия равно или больше табличного $F_{табл}$, нулевая гипотеза отклоняется и с вероятностью (1-α) принимается альтернативная гипотеза о статистической значимости уравнения в целом.

Вывод: Поскольку ______________________, то коэффициент детерминации статистически *(значим / не значим*), следовательно, найденная оценка уравнения регрессии статистически *надежна / не надежна*).

2.2. Множественная линейная регрессия. Построение и оценка качества модели

Исходные данные

По 20 предприятиям региона изучается зависимость выработки продукции на одного работника Y (тыс. руб.) от ввода в действие новых основных фондов X_1 (% от стоимости фондов на конец года) и от удельного веса рабочих высокой квалификации в общей численности рабочих X_2 (%).

Таблица 3

Ввод в действие основных фондов и удельный вес рабочих высокой квалификации в общей численности рабочих по предприятиям

Номер предприятия	Выработки продукции на одного работника, тыс. руб. Y	Ввод в действие основных фондов, % от стоимости фондов на конец года, X_1	Удельный вес рабочих высокой квалификации в общей численности рабочих, %, X_2
1	7	3,5	10
2	7	3,6	12
3	7	3,9	15
4	7	4,1	21
5	8	4,2	23
6	8	4,5	23
7	9	5,3	24
8	9	5,5	24
9	10	5,6	26
10	10	6,1	25
11	10	6,3	27
12	10	6,5	28
13	11	7,2	31
14	12	7,5	31
15	12	7,9	34
16	13	8,2	38
17	13	8,4	40
18	14	8,6	41
19	14	9,5	44
20	15	9,6	45

Задание

Построить линейную модель множественной регрессии. Записать стандартизированное уравнение множественной регрессии.

На основе стандартизированных коэффициентов регрессии и средних коэффициентов эластичности ранжировать факторы по степени их влияния на результат. Найти и проанализировать коэффициенты парной, частной и множественной корреляции. Найти скорректированные коэффициент детерминации. Сравнить его с нескорректированным (общим) коэффициентом детерминации. С помощью F- критерия Фишера оценить статистическую надежность уравнения регрессии и коэффициента детерминации R^2.

<u>Решение</u>

Заполним строки таблицы 4.

Таблица 4

Результаты вспомогательных расчетов

N	Y	X_1	X_2	X_1^2	X_2^2	X_1Y	X_2Y	X_1X_2	Y^2
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
Σ									
срзнач									

Оцениваем **коэффициенты линейной модели множественной регрессии.** Для этого решить систему трех линейных уравнений с тремя неизвестными b_0, b_1, b_2:

$$\begin{cases} \sum Y_i = nb_0 + b_1 \sum X_{1i} + b_2 \sum X_{2i} \\ \sum X_{1i}Y_i = b_0 \sum X_{1i} + b_1 \sum {X_{1i}}^2 + b_2 \sum X_{1i}X_{2i} \\ \sum X_{2i}Y_i = b_0 \sum X_{2i} + b_1 \sum X_{1i}X_{2i} + b_2 \sum {X_{2i}}^2 \end{cases} \tag{23}$$

Решая систему, находим b_0, b_1, b_2.

Уравнение регрессии имеет вид:

$$Y = b_0 + b_1X_1 + b_2X_2 \tag{24}$$

Определим **парные коэффициенты корреляции,** используя формулы и занесем результаты расчетов в таблицу 5.

Дисперсию для результативного признака *Y* и признаков-факторов X_1 и X_2 рассчитываем по формулам:

$$D(X) = \frac{\sum_i X_i^2}{n} - \bar{X}^2 \tag{25}$$

$$D(Y) = \frac{\sum_i Y_i^2}{n} - \bar{Y}^2 \tag{26}$$

Среднеквадратическое отклонение признака *Y* и признаков-факторов X_1 и X_2 рассчитываем по формулам:

$$S(X) = \sqrt{D(X)} \tag{27}$$

$$S(X) = \sqrt{D(X)} \tag{28}$$

Используя формулу (11) рассчитываем *парные коэффициенты корреляции* для *Y*, X_1 и X_2 (попарно).

Таблица 5

Результаты вспомогательных расчетов для определения парных коэффициентов корреляции

Соответствие в формуле	Дисперсия		Среднеквадратическое отклонение		Парный коэффициент корреляции
	D(X)	*D(Y)*	*S(X)*	*S(Y)*	*r (XY)*
Для Y и X_1					
Для Y и X_2					
Для X_1 и X_2					

Коэффициенты уравнения множественной регрессии можно также рассчитать на основе парных корреляционных коэффициентов:

$$b_1 = \frac{r_{YX_1} - r_{YX_2} \times r_{X_1X_2}}{1 - {r_{X_1X_2}}^2} \times \frac{S(Y)}{S(X_1)} \tag{29}$$

$$b_2 = \frac{r_{YX_2} - r_{YX_1} \times r_{X_1X_2}}{1 - r_{X_1X_2}{}^2} \times \frac{S(Y)}{S(X_2)} \qquad (30)$$

где r_{YX1}, r_{YX2}, r_{X1X2} – коэффициенты парной корреляции между результатом и каждым из факторов и между факторами;

S(X_1), *S*(X_2) – среднее квадратическое отклонение 1-го и 2-го факторов соответственно;

S(Y) – среднее квадратическое отклонение результативного признака.

Параметр b_0 уравнения регрессии можно также определить по формуле:

$$b_0 = \bar{Y} - b_1\overline{X_1} - b_2\overline{X_2} \qquad (31)$$

Построим **стандартизированное уравнение множественной регрессии.**

Расчет β-коэффициентов выполняем на основе парных коэффициентов корреляции:

$$\beta_1 = \frac{r_{YX_1} - r_{YX_2} \times r_{X_1X_2}}{1 - r_{X_1X_2}{}^2} \qquad (32)$$

$$\beta_2 = \frac{r_{YX_2} - r_{YX_1} \times r_{X_1X_2}}{1 - r_{X_1X_2}{}^2} \qquad (33)$$

Стандартизированная форма уравнения регрессии имеет вид:

$$Y^0 = \beta_1 X_1 + \beta_2 X_2 \qquad (34)$$

Проанализируем **обоснованность включения в уравнение регрессии признаков-факторов**. Для этого используем *парные, частные коэффициенты корреляции*, а также *коэффициенты эластичности*.

Заполним таблицу 6.

Таблица 6

Матрица парных коэффициентов корреляции.

-	*Y*	X_1	X_2
Y	1		
X_1		1	
X_2			1

Используя шкалу Чеддока (см. с. 9) проверяем признаки-факторы на *мультиколлинеарность*. В случае обнаружения мультиколлинеарности, определяем, какой признак-фактор следует включить в уравнение регрессии. В случае необнаружения мультиколлинеарности, определяем, какой признак будет включен в уравнение регрессии первым.

Вывод: Наибольшее влияние на результативный признак оказывает фактор _______, так как для него парный коэффициент корреляции с Y равен_________, значит, при построении модели он войдет в регрессионное уравнение первым.

Рассчитаем **частные коэффициенты корреляции** по формулам:

$$r_{YX1|X2} = \frac{r_{YX_1} - r_{YX_2} \times r_{X_1X_2}}{\sqrt{(1-r_{YX_2}{}^2)\times(1-r_{X_1X_2}{}^2)}} \tag{34}$$

$$r_{YX2|X1} = \frac{r_{YX_2} - r_{YX_1} \times r_{X_1X_2}}{\sqrt{(1-r_{YX_1}{}^2)\times(1-r_{X_1X_2}{}^2)}} \tag{35}$$

$$r_{X1X2|Y} = \frac{r_{X_1X_2} - r_{YX_1} \times r_{YX_2}}{\sqrt{(1-r_{YX_1}{}^2)\times(1-r_{YX_2}{}^2)}} \tag{36}$$

Для оценки силы связи между признаками используем шкалу Чеддока (см. с. 9).

Вывод: теснота связи между Y и X_1 при фиксированном X_2 ________________; теснота связи теснота связи между Y и X_2 при фиксированном X_1 ________________; теснота связи теснота связи теснота связи между X_1 и X_2 при фиксированном Y ________________.

Рассчитаем **частные коэффициенты эластичности** по формуле:

$$E_i = b_i \frac{\overline{X_i}}{\overline{Y}} \tag{37}$$

Частный коэффициент эластичности показывает, насколько *процентов* в среднем изменяется признак-результат Y с увеличением признака-фактора X_i на 1% от своего среднего уровня *при фиксированном положении других факторов модели*. Если $|E_i| < 1$, влияние X_i на результативный признак Y незначительно. И наоборот.

Вывод: наибольшее влияние на результат оказывает признак ________, так как Е = ________ > __________.

Рассчитаем **множественный коэффициент корреляции.**

Способ 1 – на основе известных значений линейных коэффициентов парной корреляции и β-коэффициентов:

$$R = \sqrt{\sum r_{YXi}\beta_{YXi}} = \sqrt{r_{YX1}\beta_{YX1} + r_{YX2}\beta_{YX2}} \quad (38)$$

Способ 2 – на основе известных значений линейных коэффициентов парной корреляции и частных коэффициентов корреляции:

$$R = \sqrt{1 - (1 - r^2_{YX1})(1 - r^2_{YX2|X1})} \quad (39)$$

Определим и проанализируем **общий и скорректированный коэффициенты детерминации.**

Общий коэффициент детерминации представляет собой квадрат (множественного) коэффициента корреляции $\boldsymbol{R^2}$**.**

Более объективной оценкой является *скорректированный коэффициент детерминации*, который можно рассчитать по формуле:

$$\overline{R^2} = 1 - (1 - R^2)\frac{(n-1)}{(n-m-1)} \quad (40)$$

Добавление в модель новых объясняющих переменных осуществляется до тех пор, пока растет скорректированный коэффициент детерминации.

Вывод: так как коэффициент детерминации: R^2 = ___________, построенное уравнение регрессии объясняет ________% зависимости между признаком Y и факторами X_i.

Оценка статистической надежности уравнения регрессии и коэффициента детерминации $\boldsymbol{R^2}$ с помощью $\boldsymbol{F}$-критерия Фишера.

Проверим гипотезу об одновременном равенстве нулю всех коэффициентов регрессии при объясняющих переменных: H_0: $R^2 = 0$; $\beta_1 = \beta_2 = ... = \beta_m = 0$. H_1: $R^2 \neq 0$.

Фактическое значение F-критерия определяется по формуле (22). При правосторонней проверке $F_{\text{крит}} = F(k_1; k_2)$ определяется по таблицам Фишера. $\alpha = 0{,}05$; $k_1 = m$; $k_2 = n\text{-}m\text{-}1$. Если $F < F_{\text{крит}}$, то нет оснований для отклонения гипотезы H_0. Если фактическое значение $F > F_{\text{крит}}$, то коэффициент детерминации статистически значим и уравнение регрессии статистически надежно.

Вывод: Поскольку _____________________, то коэффициент детерминации статистически *(значим / не значим*) (найденная оценка уравнения регрессии статистически (*надежна / не надежна*).

2.3. Множественная линейная регрессия. Определение набора переменных

Исходные данные

Приведены данные по ВВП Y (млрд долл.), потреблении X_1 (млрд долл.), а также объеме инвестиций X_2 (млрд долл.) в 9 странах мира в текущих ценах.

Таблица 7

ВВП, потребление и инвестиции по странам мира, млрд долл.

Год	ВВП, Y	Объем потребления, X_1	Объем инвестиций, X_2
1	16331,97	771,92	176,64
2	16763,35	814,28	173,15
3	17492,22	735,6	151,96
4	18473,83	788,54	171,62
5	19187,64	853,62	192,26
6	20066,25	900,39	198,71
7	21281,78	999,55	227,17
8	22326,86	1076,37	259,07
9	23125,9	1117,51	259,85

Задание

С помощью частных *F*-критериев Фишера оценить целесообразность включения в уравнение множественной регрессии фактора X_1 после X_2 и фактора X_2 после X_1. Составить уравнение линейной парной регрессии, оставив лишь один значащий фактор.

Решение

Для принятия решения о значащих факторах используем **частные *F*-критерии Фишера**.

Таблица 8

Результаты вспомогательных расчетов для определения парных коэффициентов корреляции

Соответствие в формуле	Дисперсия		Средне квадратическое отклонение		Парный коэффициент корреляции	Квадрат парного коэффициента корреляции
	$D(X)$	$D(Y)$	$S(X)$	$S(Y)$	r_{XY}	r^2_{XY}
Для Y и X_1						
Для Y и X_2						

Рассчитаем *парные коэффициенты корреляции* для результативного признака Y и признаков-факторов X_i, используя формулы (11), (25)-(28) и заполним таблицу 8.

Определяем множественный коэффициент детерминации R^2, используя формулу (38) или (39).

Частные критерии F_{X1} и F_{X2} оценивают статистическую значимость включения факторов X_1 и X_2 в уравнение множественной регрессии и целесообразность включения в уравнение одного фактора после другого. Если фактическое значение $F > F_{крит}$ (см с. 20), то коэффициент F для соответствующего признак-фактора статистически значим, т.е. прирост факторной дисперсии за счет этого признака фактора является существенным.

F_{X1} оценивает целесообразность включения в модель фактора X_1 после включения фактора X_2 и рассчитывается по формуле:

$$F_{X1} = \frac{(R^2 - r^2_{YX2})(n-m-1)}{(1-R^2)m} \quad (41),$$

где R^2 – множественный коэффициент детерминации,

r_{YX1} – парный коэффициент корреляции между Y и X_1,

n – число наблюдений,

m – число признаков-факторов.

Вывод: В нашей задаче ____________, т.е. коэффициент F_{X1} статистически. (*значим/не значим*), следовательно, (*целесообразно / нецелесообразно*) включать в уравнение X_1 после включения в него фактора x_2.

F_{X2} указывает на целесообразность включения в модель фактора X_2 после включения фактора X_1 и рассчитывается по формуле:

$$F_{X2} = \frac{(R^2 - r^2_{YX1})(n-m-1)}{(1-R^2)m} \quad (42),$$

где r_{YX2} – парный коэффициент корреляции между Y и X_2.

Вывод: В нашем случае ____________, т.е. коэффициент F_{X2} статистически (*значим/не значим*), следовательно, (*целесообразно / нецелесообразно*) включать в уравнение X_1 после включения в него фактора X_2.

В уравнение регрессии **включаем один из двух значащий фактор**. Принимаем:

Y – ____,

X – ____.

Строим уравнение линейной парной регрессии со значащим фактором

$$Y = b_0 + b_1 X \qquad (43),$$

Заполняем расчетную таблицу 9.

Таблица 9

Результаты вспомогательных расчетов

N	*Y*	*X*	X^2	Y^2	*XY*
1					
2					
3					
4					
5					
6					
7					
8					
9					
Σ					

Вычисляем **коэффициенты уравнения парной линейной регрессии,** используя систему нормальных уравнений (2).

Вывод: эмпирические коэффициенты регрессии:

b = __________, a =

Записываем эмпирическое уравнение регрессии:

Y =

Проводим проверку на статистическую значимость коэффициентов a и b при помощи t-критерия (см. с. 11-12), и уравнения в целом при помощи F-критерия (см. с. 13-14).

2.4. Замена переменных (логарифмические модели)

Исходные данные

При исследовании 8 промышленных предприятий получены следующие данные.

Таблица 10

Выработка и затраты по предприятиям

Номер предприятия	Количество деталей, тыс шт.	Затраты, тыс. руб.
1	22,000	3221
2	30,000	3589
3	30,000	3783
4	26,000	3856
5	30,000	3876
6	41,000	4158
7	38,000	4325
8	36,000	4489

Задание

Построить модели зависимости количества деталей от объема затрат на основе **парной линейной** регрессии

$$Y = \beta X + \alpha + \varepsilon \qquad (44),$$

и **логарифмической** модели

$$Y = \alpha X^{\beta} + \varepsilon \qquad (45),$$

Оценить и проверить значимость коэффициентов каждой модели с помощью критерия Стьюдента (α=0,05). Рассчитать коэффициент детерминации и проверить значимость регрессионной модели в целом при помощи ошибки аппроксимации и критерия Фишера (α=0,05). Определение предпочтительной модели зависимости количества деталей от объема затрат и обоснование модели.

Решение

Поскольку необходимо найти зависимость количества деталей от объема затрат, то принимаем в качестве:

Y – ____________________,

X – ______________________.

Построение модели парной линейной регрессии.

Заполняем первые шесть столбцов таблицы 11, остальные – по мере накопления результатов расчетов

Таблица 11

Результаты вспомогательных расчетов

i	X	Y	X^2	Y^2	XY	$Y(X)$	$(Y_i-\bar{Y})^2$	$(Y-Y(X))^2$	$(X_i-\bar{X})^2$	$\lvert Y-Y(X)\rvert : Y$
1										
2										
3										
4										
5										
6										
7										
8										
Σ										

Получаем значения эмпирических коэффициентов линейной регрессии, решая систему нормальных уравнений (2).

Вывод: эмпирические коэффициенты регрессии:

b = __________,

a = __________.

Эмпирическое уравнение регрессии:

y =

Экономический смысл коэффициентов уравнения линейной регрессии:

1) b = __________ показывает среднее изменение результативного показателя (в единицах измерения *Y*) с повышением или понижением величины фактора х на единицу его измерения.

Вывод: с увеличением *X* на 1 единицу *Y* повышается в среднем на __.

2) a = _________формально показывает прогнозируемый уровень у, но только в том случае, если *X* = 0 находится близко с выборочными значениями.

Связь между *Y* и *X* определяет знак коэффициента регрессии b (если > 0 – прямая связь, иначе - обратная).

Вывод: В нашей задаче связь ________________.

Определяем **тесноту линейной связи** между признаками при помощи линейного **коэффициента корреляции.** Для этого рассчитываем**:**

– выборочные средние (3)-(5),

– выборочные дисперсии (6)-(7),

– среднеквадратическое отклонение (8)-(9).

Коэффициент линейной парной корреляции может быть определен по формуле (10) либо (11).

Для оценки силы связи между признаками используем шкале Чеддока (см. с.10).

Вывод: В рассматриваемом примере линейная связь между признаком *Y* и фактором *X* ____________________________________.

Вычисляем **коэффициент эластичности** по формуле (12). Коэффициент эластичности *E* показывает, на сколько процентов в среднем по совокупности изменится результат *Y* от своей средней величины при изменении фактора *X* на 1% от своего среднего значения.

Вывод: В рассматриваемом примере коэффициент эластичности ________. Следовательно, при изменении *X* на 1%, *Y* изменится на ______, следовательно, *X* (*влияет / не влияет*) на *Y*.

Вычисляем **бета – коэффициент** по формуле (13). Бета – коэффициент показывает, что увеличение *X* на величину среднеквадратического отклонения *S(X)* приведет к увеличению среднего значения *Y* на b_j среднеквадратичного отклонения *S(Y)*.

Вывод: В рассматриваемом примере___________________.

Оценим качество уравнения регрессии с помощью **ошибки аппроксимации** (14).Ошибка аппроксимации в пределах 5-7% свидетельствует о хорошем подборе уравнения регрессии к исходным данным.

Вывод: В рассматриваемой задаче ошибка ________________, поэтому данное уравнение (*можно/не следует*) использовать в качестве уравнения регрессии.

Вычислим **коэффициент детерминации** как квадрат коэффициента корреляции по формуле (15). Коэффициент детерминации показывает долю вариации результативного признака, объясненную вариацией факторного признака, и может быть выражен в процентах.

Вывод: В рассматриваемом примере _________% случаев изменения X приводят к изменению Y. Остальные ________% изменения Y объясняются факторами, не учтенными в модели.

Проанализируем **точность определения оценок коэффициентов регрессии,** вычисляя

– несмещенную оценку дисперсии (16),

– стандартная ошибка регрессии (17),

– стандартное отклонение случайной величины a (18),

– стандартное отклонение случайной величины b (19).

Проверяем **гипотезы относительно коэффициентов линейного уравнения регрессии** при помощи **t-статистики** на уровне значимости $\alpha = 0.05$. Уровень значимости α – вероятность отвергнуть правильную гипотезу при условии, что она верна (см. с. 11-12).

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии b (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии a (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Проверяем существенность уравнения линейной регрессии в целом при помощи ***F*-статистики** (см. с. 13-14).

Вывод: Поскольку _______________________, то коэффициент детерминации статистически *(значим / не значим)*, следовательно, найденная оценка уравнения регрессии статистически *надежна / не надежна*).

Построение логарифмической регрессионной модели.

Производим преобразование исходных данных

$$Y = aX^b \qquad (46),$$

следующим образом:

$$\ln(Y) = \ln(a) + b ln(X) \qquad (47),$$

и замену переменных:

$$Y' = \ln(Y) \qquad (48),$$

$$a' = \ln(a) \qquad (49),$$

$$Z = \ln(X) \quad (50),$$

Получим новую форму уравнения регрессии:

$$Y' = a' + bZ \quad (51).$$

Заполняем первые шесть столбцов таблицы 12, остальные – по мере накопления результатов расчетов

Таблица 12

Результаты вспомогательных расчетов

i	X	Y	X^2	Y^2	XY	$Y(X)$	$(Y_i - \bar{Y})^2$	$(Y - Y(X))^2$	$(X_i - \bar{X})^2$	$\lvert Y - Y(X)\rvert : Y$
1										
2										
3										
4										
5										
6										
7										
8										
Σ										

Получаем значения эмпирических коэффициентов парной линейной регрессии (51), решая систему нормальных уравнений (2).

Вывод: эмпирические коэффициенты регрессии:

b = __________,

a' = _________.

Эмпирическое уравнение регрессии:

Z =

Экономический смысл коэффициентов уравнения **логарифмической регрессионной модели**

1) b = ____________ показывает среднее изменение результативного показателя (в процентах с повышением или понижением величины фактора х на 1%.

В данном примере с увеличением ___________________на 1 % __________________повышается в среднем на ____________%.

2) a' = ____________ формально показывает прогнозируемый уровень у, но только в том случае, если х=0 находится близко с выборочными значениями.

Определяем **тесноту линейной связи** между признаками в (51) при помощи линейного **коэффициента корреляции.** Для этого рассчитываем**:**

– выборочные средние (3)-(5),

– выборочные дисперсии (6)-(7),

– среднеквадратическое отклонение (8)-(9).

Коэффициент линейной парной корреляции может быть определен по формуле (10) либо (11).

Для оценки силы связи между признаками используем шкале Чеддока (см. с.10).

Вывод: В рассматриваемом примере (51) линейная связь между признаком *Y* и фактором *X* ___________________________.

Вычисляем **коэффициент эластичности** по формуле (12). Коэффициент эластичности *E* показывает, на сколько процентов в среднем по совокупности изменится результат *Y* от своей средней величины при изменении фактора *X* на 1% от своего среднего значения.

Вывод: В рассматриваемом примере коэффициент эластичности _______. Следовательно, при изменении *X* на 1%, *Y* изменится на ______, следовательно, *X* (*влияет / не влияет*) на *Y*.

Вычисляем **бета – коэффициент** по формуле (13). Бета – коэффициент показывает, что увеличение *X* на величину среднеквадратического отклонения *S(X)* приведет к увеличению среднего значения *Y* на b_j среднеквадратичного отклонения *S(Y)*.

Вывод: В рассматриваемом примере___________________.

Оценим качество уравнения регрессии с помощью **ошибки аппроксимации** (14). Ошибка аппроксимации в пределах 5-7% свидетельствует о хорошем подборе уравнения регрессии к исходным данным.

Вывод: В рассматриваемой задаче ошибка ________________, поэтому данное уравнение (*можно/не следует*) использовать в качестве уравнения регрессии.

Вычислим **коэффициент детерминации** как квадрат коэффициента корреляции по формуле (15). Коэффициент детерминации показывает долю вариации результативного признака, объясненную вариацией факторного признака, и может быть выражен в процентах.

Вывод: В рассматриваемом примере _________% случаев изменения *X* приводят к изменению *Y*. Остальные ________% изменения Y объясняются факторами, не учтенными в модели.

Проанализируем **точность определения оценок коэффициентов регрессии (51),** вычисляя

– несмещенную оценку дисперсии (16),

– стандартная ошибка регрессии (17),

– стандартное отклонение случайной величины a (18),

– стандартное отклонение случайной величины b (19).

Проверяем **гипотезы относительно коэффициентов линейного уравнения регрессии** (51) при помощи **t-статистики** на уровне значимости α = 0.05. Уровень значимости α – вероятность отвергнуть правильную гипотезу при условии, что она верна (см. с. 11-12).

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии b (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Вывод: Поскольку ___________________________, то статистическая значимость коэффициента регрессии a (*подтверждается / не подтверждается*), следовательно, *(отвергаем / принимаем)* гипотезу о равенстве нулю этого коэффициента.

Проверяем существенность уравнения линейной регрессии (51) в целом при помощи ***F*-статистики** (см. с. 13-14).

Вывод: Поскольку ________________________, то коэффициент детерминации статистически *(значим / не значим*), следовательно, найденная оценка уравнения регрессии статистически *надежна / не надежна*).

Обоснование выбора в пользу лучшей модели.

Вывод: Для регрессионного анализа зависимости количества деталей от объема затрат с использованием первичных данных предпочтительнее (*модель парной линейной регрессии (50) / логарифмическая модель (51*).

Обоснование: ____________________________.

2.5. Фиктивные переменные

Исходные данные

По 15 предприятиям региона изучается зависимость величины спроса Y (тыс. шт.) от модификации товара X_1 (отражается фиктивной переменной: новый товар – 1, старый товар - 0) и цены товара X_2 (руб.).

Таблица 13

Содержание и трудоемкость разделов практики

Номер наблюдения	Величина спроса на товар, тыс. шт., Y	Модификация товара, X_1	Цена товара, руб., X_2
1	75,11	новый	13,39
2	74,28	новый	13,51
3	72,71	старый	13,58
4	70,15	старый	13,79
5	71,92	старый	13,84
6	70,84	старый	13,92
7	74,15	новый	14
8	70,25	новый	14,21
9	69,70	старый	14,22
10	71,00	новый	14,25
11	69,72	новый	14,61
12	66,88	старый	14,63
13	69,06	новый	14,9
14	66,06	старый	14,99
15	66,01	новый	15,06

Задание

1. Построить линейную модель множественной регрессии с **фиктивной переменной сдвига** и определить влияние качественного признака на параметры модели. Найти и проанализировать коэффициенты парной, частной и множественной корреляции. С помощью F- критерия Фишера оценить статистическую надежность уравнения регрессии и коэффициента детерминации R^2.

2. Построить линейную модель множественной регрессии с **фиктивной переменной наклона** и определить влияние качественного признака на параметры модели. Найти и проанализировать коэффициенты парной, частной и множественной корреляции. С по-

мощью F- критерия Фишера оценить статистическую надежность уравнения регрессии и коэффициента детерминации R^2.

Решение

В линейной модели множественной регрессии с **фиктивной переменной сдвига**

$$Y = b_0 + b_1X_1 + b_2X_2 \tag{52}$$

можно определить степень влияния качественного признака на значение b_0, то есть при изменении качества (новизны) товара можно говорить о том, что *изменится спрос* и *кривая спроса сдвинется вправо или влево*.

Определим:

Y –___________________________________,

X_1 – фиктивная переменная, принимающая значение 1 для нового товара и значение 0 для старого товара.

X_2 – ___________________________________.

Заполним строки таблицы 14.

Таблица 14

Результаты вспомогательных расчетов

N	Y	X_1	X_2	X_1^2	X_2^2	X_1Y	X_2Y	X_1X_2	Y^2
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
Σ									
срзнач									

Для определения **коэффициентов линейной модели множественной регрессии с фиктивной переменной наклона** составим и решим систему линейных уравнений с тремя неизвестными b_0, b_1, b_2 (23).

Решая систему, находим b_0, b_1, b_2.

Уравнение регрессии имеет вид:

$Y =$

Учитывая, что значение фиктивной переменной X_1 равно 1 для «нового» товара и 0 для «старого» товара, находим уравнение регрессии для старого товара $Y =$
и для нового товара $Y =$ ________________________________.

*Вывод***: в модели множественной регрессии с фиктивной переменной сдвига** изменение характеристик товара приводит к изменению в уравнении для нового товара значения коэффициента ______ на ___________ Это соответствует ______________ кривой спроса (*влево/вправо*).

Определим **парные коэффициенты корреляции,** используя формулы и занесем результаты расчетов в таблицу 15.

Дисперсию для результативного признака Y и признаков-факторов X_1 и X_2 рассчитываем по формулам (25)-(26).

Среднеквадратическое отклонение признака Y и признаков-факторов X_1 и X_2 рассчитываем по формулам (27)-(28).

Используя формулу (11) рассчитываем *парные коэффициенты корреляции* для Y, X_1 и X_2 (попарно).

Таблица 15

Результаты вспомогательных расчетов для определения парных коэффициентов корреляции

Соответствие в формуле	Дисперсия		Среднеквадратическое отклонение		Парный коэффициент корреляции
	D(X)	*D(Y)*	*S(X)*	*S(Y)*	*r (XY)*
Для Y и X_1					
Для Y и X_2					
Для X_1 и X_2					

Проанализируем **обоснованность включения в уравнение регрессии признаков-факторов**. Для этого используем *парные, частные коэффициенты корреляции*.

Заполним таблицу 16.

Таблица 16

Матрица парных коэффициентов корреляции.

-	Y	X_1	X_2
Y	1		
X_1		1	
X_2			1

Рассчитаем **частные коэффициенты корреляции** по формулам (34)-(36).

Для оценки силы связи между признаками используем шкалу Чеддока (см. с. 9).

*Вывод***:** теснота связи между Y и X_1 при фиксированном X_2 _______________; теснота связи теснота связи между Y и X_2 при фиксированном X_1 _______________; теснота связи теснота связи теснота связи между X_1 и X_2 при фиксированном Y _______________.

Рассчитаем **множественный коэффициент корреляции** по формуле (39).

Определим и проанализируем **общий и скорректированный коэффициенты детерминации.**

Общий коэффициент детерминации представляет собой квадрат (множественного) коэффициента корреляции R^2**.**

Скорректированный коэффициент детерминации, который можно рассчитать по формуле (40).

*Вывод***:** так как коэффициент детерминации: R^2 = __________, построенное уравнение регрессии объясняет _______% зависимости между признаком Y и факторами X_i.

Оценим **статистическую надежности уравнения регрессии и коэффициента детерминации R^2 с помощью F-критерия Фишера** (см. с. 20).

Вывод: Поскольку ______________________, то коэффициент детерминации статистически (*значим / не значим*) (найденная оценка уравнения регрессии статистически (*надежна / не надежна*).

В линейной модели множественной регрессии с **фиктивной переменной наклона** (52) можно определить степень влияния качественного признака на значение b_2, то при изменении качества (новизны) товара можно говорить о том, что *изменится эластичность спроса* и соответственно *наклон кривой спроса*.

Определим:

Y –________________________________,

X_1 – фиктивная переменная, принимающая значение 1 для нового товара и значение 0 для старого товара.

X_2 – ________________________________.

Обозначим $Z=X_1X_2$.

Заполните строки таблицы 17.

Таблица 17

Результаты вспомогательных расчетов

N	Y	Z	X_2	Z^2	X_2^2	$Z\,Y$	X_2Y	$Z\,X_2$	Y^2
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
Σ									
срзнач									

Для определения **коэффициентов линейной модели множественной регрессии с фиктивной переменной наклона** составим и решим систему линейных уравнений с тремя неизвестными b_0, b_1, b_2 (23).

Решая систему, находим b_0, b_1, b_2.

Уравнение регрессии имеет вид (с учетом $Z=X_1X_2$):

Y =

Учитывая, что значение фиктивной переменной X_1 равно 1 для «нового» товара и 0 для «старого» товара, находим уравнение регрессии для старого товара Y =
и для нового товара Y = ________________________________.

Вывод: в **модели множественной регрессии с фиктивной переменной наклона** изменение характеристик товара приводит к изменению в уравнении для нового товара значения коэффициента ______ на ________________. Это соответствует изменению ______________________________ кривой спроса.

Определим **парные коэффициенты корреляции,** используя формулы и занесем результаты расчетов в таблицу 18.

Дисперсию для результативного признака Y и признаков-факторов X_1 и X_2 рассчитываем по формулам (25)-(26).

Среднеквадратическое отклонение признака Y и признаков-факторов X_1 и X_2 рассчитываем по формулам (27)-(28).

Используя формулу (11) рассчитываем *парные коэффициенты корреляции* для Y, X_1 и X_2 (попарно).

Таблица 18

Результаты вспомогательных расчетов для определения парных коэффициентов корреляции

Соответствие в формуле	Дисперсия		Среднеквадратическое отклонение		Парный коэффициент корреляции
	D(X)	*D(Y)*	*S(X)*	*S(Y)*	*r (XY)*
Для Y и X_1					
Для Y и X_2					
Для X_1 и X_2					

Проанализируем **обоснованность включения в уравнение регрессии признаков-факторов**. Для этого используем *парные, частные коэффициенты корреляции*.

Заполним таблицу 19.

Таблица 19

Матрица парных коэффициентов корреляции.

-	Y	X_1	X_2
Y	1		
X_1		1	
X_2			1

Рассчитаем **частные коэффициенты корреляции** по формулам (34)-(36).

Для оценки силы связи между признаками используем шкалу Чеддока (см. с. 9).

*Вывод***:** теснота связи между Y и X_1 при фиксированном X_2 _______________; теснота связи теснота связи между Y и X_2 при фиксированном X_1 _______________; теснота связи теснота связи теснота связи между X_1 и X_2 при фиксированном Y _______________.

Рассчитаем **множественный коэффициент корреляции** по формуле (39).

Определим и проанализируем **общий и скорректированный коэффициенты детерминации.**

Общий коэффициент детерминации представляет собой квадрат (множественного) коэффициента корреляции $\boldsymbol{R^2}$**.**

Скорректированный коэффициент детерминации, который можно рассчитать по формуле (40).

*Вывод***:** так как коэффициент детерминации: R^2 = __________, построенное уравнение регрессии объясняет _______% зависимости между признаком *Y* и факторами $X_{i.}$

Оценим **статистическую надежности уравнения регрессии и коэффициента детерминации** $\boldsymbol{R^2}$ **с помощью** $\boldsymbol{F}$**-критерия Фишера** (см. с. 20).

Вывод: Поскольку _____________________, то коэффициент детерминации статистически *(значим / не значим*) (найденная оценка уравнения регрессии статистически (*надежна / не надежна*).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

Основной:

1. *Бабешко, Л.О.* Эконометрика и эконометрическое моделирование : учебник / Л.О. Бабешко, М.Г. Бич, И.В. Орлова. - Москва : Вузовский учебник : ИНФРА-М, 2018. 385 с. Режим доступа: URL: https://znanium.com/catalog/product/968797

2. *Базилевский, М.П., Гефан, Г.Д.* Эконометрика (продвинутый уровень): лабораторный практикум / М.П. Базилевский, Г.Д. Гефан.– Иркутск: ИрГУПС, 2016. 77 с. https://docplayer.ru/86929453-Ekonometrika-prodvinutyy-uroven.html

3. Крянев, А.В. Эконометрика (продвинутый уровень) [Электронный ресурс]: Конспект лекций / А.В. Крянев. М.: ИНФРА-М, 2017. 62 с. Режим доступа: http://znanium.com/bookread2.php?book=767248

Дополнительный:

1. *Бородич, С.А.* Эконометрика. Практикум : учеб. пособие / С.А. Бородич. — Минск : Новое знание ; Москва : ИНФРА-М, 2018. 329 с. Режим доступа: URL: https://znanium.com/catalog/product/988809

2. *Орлова, И.В., Филонова, Е.С.* Эконометрика (продвинутый уровень) [Электронный ресурс]. Методические указания повыполнению контрольной и лабораторной работ / И.В. Орлова, Е.С. Филонова. М.: ВЗФЭИ, 2011. 108 с. Режим доступа: URL: https://znanium.com/catalog/product/453448

3. *Елисеева, II.Я.* Эконометрика: Учебник / II.Я. Елисеева. М.: Юрайт, 2012. 464 с.

4. *Кремер, Н.Ш., Путко, Б.А.* Эконометрика. Учебник для вузов / Н.Ш. Кремер, Б.А. Путко; под ред. проф. Н.Ш. Кремера. М: ЮНИТИ-ДАНА, 2010. 328 с.

5. *Магнус, Я.Р., Катышев, П.К, Пересецкий, А.А.* Эконометрика. [Электронный ресурс] / Я.Р. Магнус, П.К. Катышев, А.А. Пересецкий. М: Дело, 2007. 504с. Режим доступа: http://math.isu.ru/ru/chairs/me/files/books/magnus.pdf.

СОДЕРЖАНИЕ