**4. Контрольная работа по эконометрике «Построение парной линейной регрессии».**

**4.1 Задание:**

*Исходные данные**(см. Приложение 1)**представляют собой двумерную выборку , . По выборке необходимо построить парную линейную регрессию и оценить качество построенной модели.*

**4.2 Порядок выполнения работы:**

1. Для заданных исходных данных постройте поле корреляции — диаграмму зависимости показателя от фактора . При построении выберите тип диаграммы «*Точечная*» (без отрезков, соединяющих точки).

Вычислите коэффициенты выборочной линейной регрессии вручную.

1. Запишите найденное уравнение эмпирической регрессии. Дайте интерпретацию коэффициенту  в рамках Вашей задачи. Вычислите по уравнению эмпирической регрессии значения , .
2. Постройте на корреляционном поле прямую выборочной линейной регрессии по точкам , . (При построении выберите тип диаграммы «*Точечная*», на которой значения соединены отрезками.)
3. Найдите величину средней ошибки аппроксимации . Прокомментируйте полученное значение.
4. Вычислите коэффициент детерминации  непосредственно по формуле:



1. Используя построенную модель, рассчитайте значение зависимой переменной при значении фактора ,на 10% превышающего среднее значение .
2. На основании проведенного выше анализа адекватности модели сделайте вывод о правдоподобности прогноза.
3. Используя опцию Excel «Добавить линию тренда», проверьте свои вычисления и построение графика прямой регрессии.

**4.3 Требования к оформлению контрольной работы. Выбор варианта.**

Задания могут выполняться с применением компьютера. Вычисления произ­водятся с точностью до двух знаков после запятой.

При выполнении контрольной работы необходимо придерживаться указанных ниже правил. Работы, выполненные без соблюдения этих правил, не зачитываются и возвращаются студенту для переработки.

1. Студент должен выполнять контрольные задания по варианту, номер которого совпадает с последней цифрой номера его зачетной книжки. Контрольные работы, выполненные ***не по своему варианту, НЕ ЗАСЧИТЫВАЮТСЯ***.
2. Работа сдается на бумажном носителе в печатном виде. Образец оформления работы дан в Приложении 2 настоящих методических указаний.

***Приложение 1. Варианты заданий.***

***Вариант №1***

В таблице представлены статистические данные о размере товарооборота Х и суммы издержек обращения Y по десяти магазинам.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Товарооборот Х** | **480** | **510** | **530** | **540** | **570** | **590** | **680** | **640** | **650** | **660** |
| **Издержки обращения Y** | **31** | **25** | **31** | **28** | **29** | **32** | **36** | **36** | **37** | **36** |

**Образец выполнения работы**

Задача.

По 21 региону страны изучается зависимость розничной продажи телевизоров () от среднедушевого денежного дохода в месяц ().

|  |  |  |
| --- | --- | --- |
| Номер региона | Среднедушевой денежный доход в месяц, тыс. руб., | Объем розничной продажи телевизоров, тыс. шт., |
| 1 | 2 | 28 |
| 2 | 2,4 | 21,3 |
| 3 | 2,1 | 21 |
| 4 | 2,6 | 23,3 |
| 5 | 1,7 | 15,8 |
| 6 | 2,5 | 21,9 |
| 7 | 2,4 | 20 |
| 8 | 2,6 | 22 |
| 9 | 2,8 | 23,9 |
| 10 | 2,6 | 26 |
| 11 | 2,6 | 24,6 |
| 12 | 2,5 | 21 |
| 13 | 2,9 | 27 |
| 14 | 2,6 | 21 |
| 15 | 2,2 | 24 |
| 16 | 2,6 | 24 |
| 17 | 3,3 | 31,9 |
| 18 | 3,9 | 33 |
| 19 | 4 | 35,4 |
| 20 | 3,7 | 34 |
| 21 | 3,4 | 31 |

Решение.

1. Построим поле корреляции.



Расположение точек на диаграмме дает нам право предположить, что переменные связаны линейной зависимостью. Рассчитаем выборочные коэффициенты корреляции. Для этого проведем промежуточные вычисления, по формулам ( ) и поместим результаты вычислений в таблицу:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Номер региона* |  |  |  |  |  |
| 1 | 2 | 28 | 56 | 4 | 784 |
| 2 | 2,4 | 21,3 | 51,12 | 5,76 | 453,69 |
| 3 | 2,1 | 21 | 44,1 | 4,41 | 441 |
| 4 | 2,6 | 23,3 | 60,58 | 6,76 | 542,89 |
| 5 | 1,7 | 15,8 | 26,86 | 2,89 | 249,64 |
| 6 | 2,5 | 21,9 | 54,75 | 6,25 | 479,61 |
| 7 | 2,4 | 20 | 48 | 5,76 | 400 |
| 8 | 2,6 | 22 | 57,2 | 6,76 | 484 |
| 9 | 2,8 | 23,9 | 66,92 | 7,84 | 571,21 |
| 10 | 2,6 | 26 | 67,6 | 6,76 | 676 |
| 11 | 2,6 | 24,6 | 63,96 | 6,76 | 605,16 |
| 12 | 2,5 | 21 | 52,5 | 6,25 | 441 |
| 13 | 2,9 | 27 | 78,3 | 8,41 | 729 |
| 14 | 2,6 | 21 | 54,6 | 6,76 | 441 |
| 15 | 2,2 | 24 | 52,8 | 4,84 | 576 |
| 16 | 2,6 | 24 | 62,4 | 6,76 | 576 |
| 17 | 3,3 | 31,9 | 105,27 | 10,89 | 1017,61 |
| 18 | 3,9 | 33 | 128,7 | 15,21 | 1089 |
| 19 | 4 | 35,4 | 141,6 | 16 | 1253,16 |
| 20 | 3,7 | 34 | 125,8 | 13,69 | 1156 |
| 21 | 3,4 | 31 | 105,4 | 11,56 | 961 |
| ***Сумма*** | ***57,4*** | ***530,1*** | ***1504,46*** | ***164,32*** | ***13926,97*** |

Составляем систему уравнений:



и решаем ее по формулам Крамера:







Тогда, согласно теореме Крамера,

2. Получаем уравнение регрессии:



Величина коэффициента регрессии  означает, что увеличение среднедушевого месячного дохода на 1 тыс. руб. приведет к увеличение объема розничной продажи в среднем на 7 540 телевизоров. Коэффициент  в данном случае не имеет содержательной интерпретации.

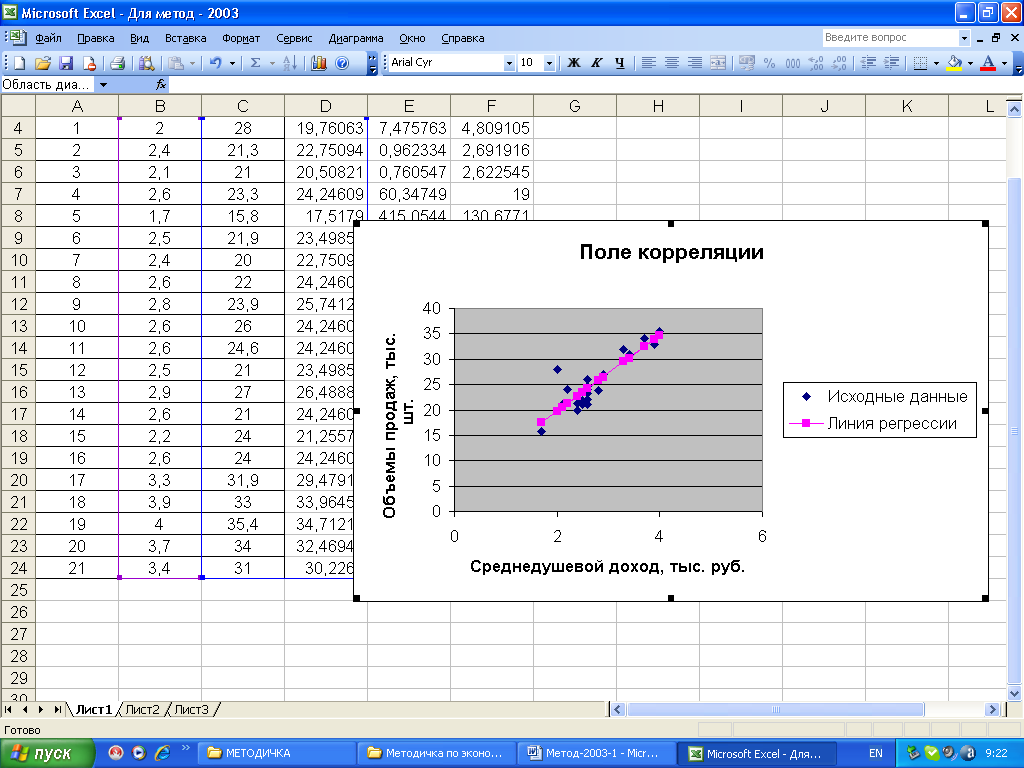
1. Нанесем построенную линию регрессии на диаграмму. Для этого рассчитаем значения , , по формуле:



Результаты вычислений запишем в таблицу:

|  |  |  |  |
| --- | --- | --- | --- |
| *Номер региона* |  |  |  |
| 1 | 2 | 28 | 19,76 |
| 2 | 2,4 | 21,3 | 22,75 |
| 3 | 2,1 | 21 | 20,51 |
| 4 | 2,6 | 23,3 | 24,25 |
| 5 | 1,7 | 15,8 | 17,52 |
| 6 | 2,5 | 21,9 | 23,50 |
| 7 | 2,4 | 20 | 22,75 |
| 8 | 2,6 | 22 | 24,25 |
| 9 | 2,8 | 23,9 | 25,74 |
| 10 | 2,6 | 26 | 24,25 |
| 11 | 2,6 | 24,6 | 24,25 |
| 12 | 2,5 | 21 | 23,50 |
| 13 | 2,9 | 27 | 26,49 |
| 14 | 2,6 | 21 | 24,25 |
| 15 | 2,2 | 24 | 21,26 |
| 16 | 2,6 | 24 | 24,25 |
| 17 | 3,3 | 31,9 | 29,48 |
| 18 | 3,9 | 33 | 33,96 |
| 19 | 4 | 35,4 | 34,71 |
| 20 | 3,7 | 34 | 32,47 |
| 21 | 3,4 | 31 | 30,23 |

Наносим на диаграмму точки из последнего столбца таблицы (*Линия регрессии*):



4. Для оценки тесноты линейной зависимости рассчитаем коэффициент детерминации. Для этого необходимо провести ряд дополнительных вычислений.

Прежде всего, найдем ***выборочное среднее***  по формуле:



Теперь произведем расчет остальных вспомогательных величин:

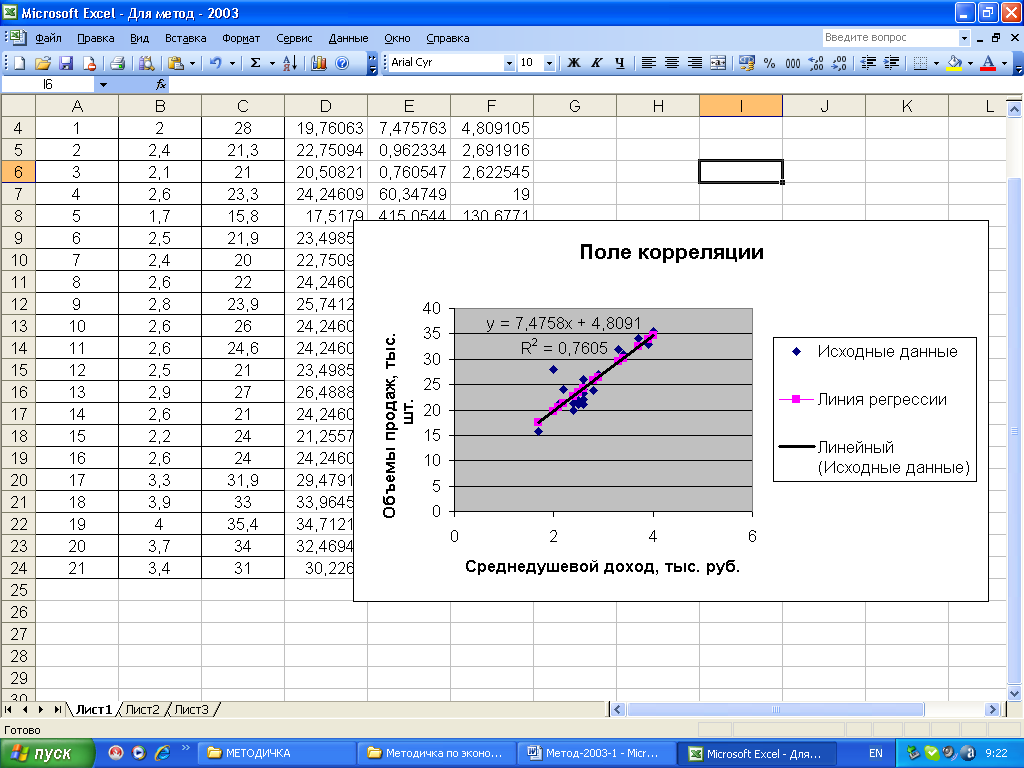
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Номер региона* |  |  |  |  |  |  |  |
| 1 | 2 | 28 | 19,76 | 8,24 | 67,89 | 2,76 | 7,60 |
| 2 | 2,4 | 21,3 | 22,75 | -1,45 | 2,11 | -3,94 | 15,55 |
| 3 | 2,1 | 21 | 20,51 | 0,49 | 0,24 | -4,24 | 18,00 |
| 4 | 2,6 | 23,3 | 24,25 | -0,95 | 0,90 | -1,94 | 3,77 |
| 5 | 1,7 | 15,8 | 17,52 | -1,72 | 2,95 | -9,44 | 89,17 |
| 6 | 2,5 | 21,9 | 23,50 | -1,60 | 2,56 | -3,34 | 11,17 |
| 7 | 2,4 | 20 | 22,75 | -2,75 | 7,57 | -5,24 | 27,49 |
| 8 | 2,6 | 22 | 24,25 | -2,25 | 5,04 | -3,24 | 10,52 |
| 9 | 2,8 | 23,9 | 25,74 | -1,84 | 3,39 | -1,34 | 1,80 |
| 10 | 2,6 | 26 | 24,25 | 1,75 | 3,08 | 0,76 | 0,57 |
| 11 | 2,6 | 24,6 | 24,25 | 0,35 | 0,13 | -0,64 | 0,41 |
| 12 | 2,5 | 21 | 23,50 | -2,50 | 6,24 | -4,24 | 18,00 |
| 13 | 2,9 | 27 | 26,49 | 0,51 | 0,26 | 1,76 | 3,09 |
| 14 | 2,6 | 21 | 24,25 | -3,25 | 10,54 | -4,24 | 18,00 |
| 15 | 2,2 | 24 | 21,26 | 2,74 | 7,53 | -1,24 | 1,54 |
| 16 | 2,6 | 24 | 24,25 | -0,25 | 0,06 | -1,24 | 1,54 |
| 17 | 3,3 | 31,9 | 29,48 | 2,42 | 5,86 | 6,66 | 44,32 |
| 18 | 3,9 | 33 | 33,96 | -0,96 | 0,93 | 7,76 | 60,17 |
| 19 | 4 | 35,4 | 34,71 | 0,69 | 0,47 | 10,16 | 103,17 |
| 20 | 3,7 | 34 | 32,47 | 1,53 | 2,34 | 8,76 | 76,69 |
| 21 | 3,4 | 31 | 30,23 | 0,77 | 0,60 | 5,76 | 33,14 |
| ***Сумма*** | ***57,4*** | ***530,1*** |  |  | ***130,68*** |  | ***545,73*** |

Для вычисления коэффициента детерминации воспользуемся формулой ( ):



Значение коэффициента детерминации позволяет сделать предварительный вывод о том, что у нас имеются основания использовать модель линейной регрессии в данной задаче, поскольку .

5. Нанесем теперь уравнение регрессии на диаграмму, используя специальные средства Excel («*Добавить линию тренда*»).



Линия регрессии, построенная нами ранее, совпала с данной линией регрессии. Нетрудно убедиться, что уравнение регрессии и коэффициент детерминации тоже совпадают с полученными ранее вручную.

6. Найдем теперь среднюю ошибку аппроксимации для оценки погрешности модели. Для этого нам потребуется вычислить еще ряд промежуточных величин:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Номер региона* |  |  |  |  |  |
| 1 | 2 | 28 | 19,76 | 8,24 | 0,29 |
| 2 | 2,4 | 21,3 | 22,75 | -1,45 | 0,07 |
| 3 | 2,1 | 21 | 20,51 | 0,49 | 0,02 |
| 4 | 2,6 | 23,3 | 24,25 | -0,95 | 0,04 |
| 5 | 1,7 | 15,8 | 17,52 | -1,72 | 0,11 |
| 6 | 2,5 | 21,9 | 23,50 | -1,60 | 0,07 |
| 7 | 2,4 | 20 | 22,75 | -2,75 | 0,14 |
| 8 | 2,6 | 22 | 24,25 | -2,25 | 0,10 |
| 9 | 2,8 | 23,9 | 25,74 | -1,84 | 0,08 |
| 10 | 2,6 | 26 | 24,25 | 1,75 | 0,07 |
| 11 | 2,6 | 24,6 | 24,25 | 0,35 | 0,01 |
| 12 | 2,5 | 21 | 23,50 | -2,50 | 0,12 |
| 13 | 2,9 | 27 | 26,49 | 0,51 | 0,02 |
| 14 | 2,6 | 21 | 24,25 | -3,25 | 0,15 |
| 15 | 2,2 | 24 | 21,26 | 2,74 | 0,11 |
| 16 | 2,6 | 24 | 24,25 | -0,25 | 0,01 |
| 17 | 3,3 | 31,9 | 29,48 | 2,42 | 0,08 |
| 18 | 3,9 | 33 | 33,96 | -0,97 | 0,03 |
| 19 | 4 | 35,4 | 34,71 | 0,69 | 0,02 |
| 20 | 3,7 | 34 | 32,47 | 1,53 | 0,05 |
| 21 | 3,4 | 31 | 30,23 | 0,77 | 0,02 |

Просуммируем теперь элементы последнего столбца и разделим полученную сумму на 21 – общее количество исходных данных:

.

Итак, средняя ошибка аппроксимации . Величина ошибки оказалась около 8%, что говорит о небольшой погрешности построенной модели. Данную модель, с учетом неплохих характеристик ее качества, вполне можно использовать для прогноза – одной из основных целей эконометрического анализа.

7. Рассчитаем значение фактора, для которого необходимо построить прогноз. Для этого необходимо вычислить выборочное среднее значение  по формуле:

.

Для нашей задачи среднее значение среднедушевого месячного дохода:

.

Рассчитаем теперь значение .

Подставим теперь полученное значение фактора  в уравнение регрессии и найдем прогнозируемое значение:

.

Таким образом, если среднедушевой месячный доход в некотором регионе составит 3 003 руб., количество продаваемых телевизоров составит в среднем 27 450 шт. в месяц.