Это таблица из 1 лабораторной (которую я сделала)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| i | x | y | y\*x | $$x^{2}$$ | $$y^{2}$$ |
| 1 | 107,8 | 30,4 | 3277,12 | 11620,84 | 924,16 |
| 2 | 126,5 | 31,6 | 3997,4 | 16002,25 | 998,56 |
| 3 | 138,1 | 31,9 | 4405,39 | 19071,61 | 1017,61 |
| 4 | 157,8 | 34,4 | 5428,32 | 24900,84 | 1183,36 |
| 5 | 166,2 | 30,7 | 5102,34 | 27622,44 | 942,49 |
| 6 | 175,4 | 34,5 | 6051,3 | 30765,16 | 1190,25 |
| 7 | 197,1 | 35,1 | 6918,21 | 38848,41 | 1232,01 |
| 8 | 227,7 | 40,7 | 9267,39 | 51847,29 | 1656,49 |
| 9 | 243,9 | 37,4 | 9121,86 | 59487,21 | 1398,76 |
| 10 | 264,8 | 42,6 | 11280,48 | 70119,04 | 1814,76 |
| 11 | 301,2 | 45,4 | 13674,48 | 90721,44 | 2061,16 |
| 12 | 72,5 | 35,3 | 2559,25 | 5256,25 | 1246,09 |
| 13 | 316,8 | 51,8 | 16410,24 | 100362,2 | 2683,24 |
| 14 | 347,3 | 63,4 | 22018,82 | 120617,3 | 4019,56 |
| Σ | 2843,1 | 545,2 | 119512,6 | 667242,3 | 22368,5 |

По данным Лабораторной работы №1 рассчитать параметры ***параболической***парной регрессии и её характеристики.

1. ***Парабола второго порядка*** ****.

Применив МНК для определения параметров параболы второго порядка (т.е. взяв первые производные по параметрам и приравняв их к нулю), получим систему нормальных уравнений:



Система решается методом определителей: .

**Лабораторная работа 2**

По данным 30 наблюдений построить модель множественной регрессии. Рассчитать все характеристики согласно заданию.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№** | **Валовой продукт, млн. руб.** | **Балансовая стоимость оборудования, млн. руб.** | **Объем промышленного производства, млн. руб.** | **Количество занятых, тыс.чел.** |
| ***i*** | ***y***  | ***x1*** | ***x2*** | ***x3*** |
| 1 | 11269 | 49834 | 4406 | 169 |
| 2 | 8740 | 39381 | 3518 | 180 |
| 3 | 12992 | 67977 | 3471 | 268 |
| 4 | 5744 | 31643 | 1777 | 127 |
| 5 | 4769 | 26884 | 1353 | 91 |
| 6 | 7957 | 39625 | 2843 | 153 |
| 7 | 10625 | 40718 | 5610 | 154 |
| 8 | 7654 | 32934 | 2209 | 137 |
| 9 | 7192 | 38551 | 2743 | 122 |
| 10 | 9425 | 53544 | 2726 | 166 |
| 11 | 10012 | 50181 | 3998 | 194 |
| 12 | 12092 | 56124 | 4776 | 163 |
| 13 | 6746 | 29025 | 2647 | 89 |
| 14 | 12284 | 52154 | 3893 | 149 |
| 15 | 5416 | 22855 | 1510 | 105 |
| 16 | 13727 | 61445 | 5280 | 174 |
| 17 | 13581 | 64618 | 2780 | 246 |
| 18 | 15887 | 69583 | 5723 | 278 |
| 19 | 22837 | 95113 | 5340 | 431 |
| 20 | 12328 | 50083 | 4749 | 185 |
| 21 | 9084 | 49066 | 3324 | 186 |
| 22 | 18397 | 59871 | 5820 | 243 |
| 23 | 15962 | 73762 | 4235 | 289 |
| 24 | 8505 | 38833 | 3033 | 156 |
| 25 | 5666 | 28139 | 1249 | 104 |
| 26 | 18306 | 85022 | 3537 | 272 |
| 27 | 12422 | 60576 | 2695 | 236 |
| 28 | 7061 | 28439 | 1015 | 114 |
| 29 | 18285 | 53924 | 4975 | 231 |
| 30 | 7021 | 44981 | 997 | 112 |

***Теоретические сведения:***

1. **Основные формулы для расчета.**

*Множественная регрессия* – это уравнение связи с несколькими независимыми переменными: ,

где *y* – зависимая переменная (*результативный признак*); *х*1, *х*2, …, *х*m – независимые переменные (*признаки-факторы*).

Для построения уравнения множественной регрессии чаще используются следующие функции:

* линейная – ;
* степенная – ;
* экспонента – ;
* гипербола – ;

Можно использовать и другие функции, приводимые к линейному виду.

В начале построения модели множественной регрессии осуществляют *отсев* факторов, оказывающих *слабое влияние* на результативный признак. Для этого строят ***матрицу парных коэффициентов корреляции***, которые находят по формулам:

,

где .

Факторы с низкими по значению коэффициентами корреляции выводят из модели.

Наряду с парными коэффициентами корреляции для отсева используют ***стандартизированные коэффициенты регрессии*** *βi*, значения которых вычисляются по формулам:

 – из модели исключаются факторы с наименьшим значением *β*.

Стандартизованные коэффициенты регрессии *βi* *можно сравнивать* между собой. Сравнивая их друг с другом, можно ранжировать факторы по силе их воздействия на результат. В этом основное достоинство стандартизованных коэффициентов регрессии в отличие от ***коэффициентов «чистой» регрессии***, которые *несравнимы между собой*.

Кроме того, стандартизированные коэффициенты регрессии *показывают, на сколько своих «сигм» изменится в среднем результат, если соответствующий фактор xi изменится на одну свою «сигму» при неизменном уровне других факторов*.

По стандартизированным коэффициентам строят уравнение регрессии в ***стандартизированном масштабе***:

, где *ty*, *tx1*, *tx2* – ***стандартизированные переменные***: , для которых их среднее значение равно нулю, а среднее квадратическое отклонение равно единице.

После отсева факторов переходят к построению уравнения множественной регрессии в ***натуральном масштабе***. Для оценки его параметров применяют *метод наименьших квадратов* (МНК). Для ***линейной двухфакторной модели*** вида  система уравнений будет иметь вид:



коэффициенты которой определяются методом определителей: .

Так же можно воспользоваться готовыми формулами, которые являются следствием из этой системы:



В линейной множественной регрессии параметры при *x* называются ***коэффициентами «чистой» регрессии***. Они *характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизмененном значении других факторов, закрепленных на среднем уровне*.

Коэффициенты «чистой» регрессии *bi* связаны со стандартизованными коэффициентами регрессии *βi* следующим образом:

 и соответственно .

Поэтому можно переходить от уравнения регрессии в стандартизованном масштабе к уравнению регрессии в натуральном масштабе переменных, при этом параметр *a* определяется как 

***Средние коэффициенты эластичности*** для линейной регрессии рассчитываются по формуле

 и *показывают на сколько процентов в среднем изменится результат, при изменении соответствующего фактора на 1%.*

Средние показатели эластичности *можно сравнивать* друг с другом и соответственно ранжировать факторы по силе их воздействия на результат.

*Тесноту совместного влияния* всех факторов на результат оценивает ***индекс множественной корреляции***:



где  – теоретические (расчетные) значения функции, посчитанные по полученному уравнению множественной регрессии.

Значение индекса множественной корреляции лежит в пределах от 0 до 1 и должно быть больше или равно максимальному парному индексу корреляции:



При *линейной зависимости* линейный коэффициент множественной корреляции называется ***совокупным коэффициентом корреляции*** и его можно определить через матрицы парных коэффициентов корреляции:



где  – определитель матрицы парных коэффициентов;

 –

определитель матрицы межфакторной корреляции, который получается из ∆*r* отбрасыванием 1-й строки и 1-го столбца.

Так же при линейной зависимости признаков формула коэффициента множественной корреляции может быть также представлена следующим выражением:



где *βi* – стандартизованные коэффициенты регрессии; *ryx* – парные коэффициенты корреляции результата с каждым фактором. Или для двухфакторной модели:



*Качество построенной модели в целом* оценивает ***коэффициент*** (***индекс***) ***детерминации***. Коэффициент множественной детерминации рассчитывается как квадрат индекса множественной корреляции .

Для того чтобы не допустить преувеличения тесноты связи, применяется ***скорректированный индекс множественной детерминации***, который содержит поправку на число степеней свободы и рассчитывается по формуле:



где *n* – число наблюдений, *m* – число факторов. При небольшом числе наблюдений нескорректированная величина коэффициента множественной детерминации *R*2 имеет тенденцию переоценивать долю вариации результативного признака, связанную с влиянием факторов, включенных в регрессионную модель.

***Частные коэффициенты*** (или ***индексы***) ***корреляции***, измеряющие влияние на *y* фактора *xi*, при *элиминировании* (*исключении влияния, закреплении*) других факторов, можно определить по формулам для двухфакторной модели:



или по рекуррентным формулам:



Рассчитанные по рекуррентной формуле частные коэффициенты корреляции изменяются в пределах от –1 до +1, а по формулам через множественные коэффициенты детерминации – от 0 до 1. Сравнение их друг с другом позволяет ранжировать факторы по тесноте их связи с результатом. *Частные коэффициенты корреляции дают меру тесноты связи каждого фактора с результатом в чистом виде*.

*Значимость уравнения множественной регрессии в целом* оценивается с помощью ***F-критерия Фишера***:



***Частный F-критерий*** *оценивает статистическую значимость присутствия каждого из факторов в уравнении* *или* *целесообразность включения в модель каждого фактора, а также статистическую значимость коэффициента «чистой» регрессии при факторе хi*. В общем виде для фактора *x* частный *F*-критерий определится как



где  – коэффициент множественной детерминации по всем факторам,  – коэффициент множественной детерминации без включения в модель фактора *хi*.

Фактическое значение частного *F*-критерия сравнивается с табличным при уровне значимости *α* и числе степеней свободы: *k*1 = 1 и *k*2 = *n* – *m* – 1. Если фактическое значение *Fх* **превышает** табличное *F*табл(α, *k*1, *k*2), то дополнительное включение фактора *xi*в модель **статистически оправданно** и коэффициент чистой регрессии *bi* при факторе *xi* **статистически значим**.

Если же фактическое значение *Fх* **меньше** табличного, то дополнительное включение в модель фактора *xi* не увеличивает существенно долю объясненной вариации признака *y*, следовательно, **нецелесообразно его включение в модель**; коэффициент регрессии при данном факторе в этом случае **статистически незначим**.

*Оценка значимости коэффициентов чистой регрессии* проводится по ***t-критерию Стьюдента***. В этом случае, как и в парной регрессии, для каждого фактора используется формула: 

Для уравнения множественной регрессии ***средняя квадратическая ошибка коэффициента регрессии*** может быть определена по формуле:



где  – коэффициент детерминации для зависимости фактора *xi* со всеми другими факторами уравнения множественной регрессии. Для двухфакторной модели (при *m* = 2):



**Лабораторная работа 3**

Взяв данные из таблицы (см. ПРИЛОЖЕНИЕ), по показателям 20 предприятий построить модель вида:



Выполнить анализ системы в соответствии с заданием лабораторной работы.

**ПРИЛОЖЕНИЕ**

Рассматриваются следующие показатели для 20 предприятий:

***Y*1** – производительность труда;

***Y*2** – индекс снижения себестоимости продукции;

***Y*3** – рентабельность;

***x*1** – трудоёмкость единицы продукции;

***x*2** – удельный вес рабочих;

***x*3** – удельный вес покупных изделий;

***x*4** – коэффициент сменности оборудования;

***x*5** – премии и вознаграждения на одного работника;

***x*6** – удельный вес потерь от брака;

***x*7** – фондоотдача;

***x*8** – среднегодовая численность работников;

***x*9** – среднегодовая стоимость основных производственных фондов;

***x*10** – среднегодовой фонд заработной платы работников

***x*11** – непроизводственные расходы.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***i*** | ***Y1*** | ***Y2*** | ***Y3*** | ***x1*** | ***x2*** | ***x3*** | ***x4*** | ***x5*** | ***x6*** | ***x7*** | ***x8*** | ***x9*** | ***x10*** | ***x11*** |
| **1** | 3,78 | 21,9 | 6,24 | 0,51 | 0,62 | 0,20 | 1,47 | 0,24 | 0,23 | 1,54 | 6265 | 17,16 | 11237 | 30,53 |
| **2** | 6,48 | 48,4 | 12,08 | 0,36 | 0,75 | 0,64 | 1,27 | 0,57 | 0,32 | 2,25 | 8810 | 27,29 | 17306 | 17,98 |
| **3** | 10,44 | 173,5 | 9,49 | 0,23 | 0,71 | 0,42 | 1,51 | 1,22 | 0,54 | 1,07 | 17659 | 184,33 | 39250 | 22,09 |
| **4** | 7,65 | 74,1 | 9,28 | 0,26 | 0,74 | 0,27 | 1,46 | 0,68 | 0,75 | 1,44 | 10342 | 58,42 | 19074 | 18,29 |
| **5** | 8,77 | 68,8 | 11,42 | 0,27 | 0,65 | 0,37 | 1,27 | 1,0 | 0,16 | 1,40 | 8901 | 59,40 | 18452 | 26,05 |
| **6** | 7,00 | 60,8 | 10,31 | 0,29 | 0,66 | 0,38 | 1,43 | 0,81 | 0,24 | 1,31 | 8402 | 49,63 | 17500 | 26,20 |
| **7** | 11,06 | 355,6 | 8,65 | 0,01 | 0,84 | 0,35 | 1,50 | 1,27 | 0,59 | 1,12 | 32625 | 391,27 | 7888 | 17,26 |
| **8** | 9,02 | 264,8 | 10,94 | 0,02 | 0,74 | 0,42 | 1,35 | 1,14 | 0,56 | 1,16 | 31160 | 258,62 | 58947 | 18,83 |
| **9** | 13,28 | 526,6 | 9,87 | 0,18 | 0,75 | 0,32 | 1,41 | 1,89 | 0,63 | 0,88 | 46461 | 75,66 | 94697 | 19,70 |
| **10** | 9,27 | 118,6 | 6,14 | 0,25 | 0,75 | 0,33 | 1,47 | 0,67 | 1,10 | 1,07 | 13833 | 123,68 | 29626 | 16,87 |
| **11** | 6,70 | 37,1 | 12,93 | 0,31 | 0,79 | 0,29 | 1,35 | 0,96 | 0,39 | 1,24 | 6391 | 37,21 | 11688 | 14,63 |
| **12** | 6,69 | 57,7 | 9,78 | 0,38 | 0,72 | 0,30 | 1,40 | 0,67 | 0,73 | 1,49 | 11115 | 53,37 | 21955 | 22,17 |
| **13** | 9,42 | 51,6 | 13,22 | 0,24 | 0,70 | 0,56 | 1,20 | 0,98 | 0,28 | 2,03 | 6555 | 32,87 | 12243 | 22,62 |
| **14** | 7,24 | 64,7 | 17,29 | 0,31 | 0,66 | 0,42 | 1,15 | 1,16 | 0,10 | 1,84 | 11085 | 45,63 | 20193 | 26,44 |
| **15** | 5,39 | 48,3 | 7,11 | 0,42 | 0,69 | 0,26 | 1,09 | 0,54 | 0,68 | 1,22 | 9484 | 48,41 | 20122 | 22,26 |
| **16** | 5,61 | 15,0 | 22,49 | 0,51 | 0,71 | 0,16 | 1,26 | 1,23 | 0,87 | 1,72 | 3967 | 13,58 | 7612 | 19,13 |
| **17** | 5,59 | 87,5 | 12,14 | 0,31 | 0,73 | 0,45 | 1,36 | 0,78 | 0,49 | 1,75 | 15283 | 63,99 | 27404 | 18,28 |
| **18** | 6,57 | 108,4 | 15,25 | 0,37 | 0,65 | 0,31 | 1,15 | 1,15 | 0,16 | 1,46 | 20874 | 104,55 | 39648 | 28,23 |
| **19** | 6,54 | 267,3 | 31,34 | 0,16 | 0,82 | 0,08 | 1,87 | 4,44 | 0,85 | 1,60 | 19418 | 222,11 | 43799 | 12,39 |
| **20** | 4,23 | 34,2 | 11,56 | 0,18 | 0,80 | 0,68 | 1,17 | 1,06 | 0,13 | 1,47 | 3351 | 25,76 | 6235 | 11,64 |

***Теоретические сведения:***

1. **Виды систем эконометрических уравнений.**

Сложные экономические процессы описываются с помощью системы взаимосвязанных (одновременных) уравнений.

Различают следующие ***виды эконометрических систем***:

* системы независимых уравнений;
* системы рекурсивных уравнений;
* систем взаимозависимых уравнений.

***Система независимых уравнений*** – каждая зависимая переменная ***у*** рассматривается как функция одного и того же набора фактора ***х***:



Каждое уравнение такой системы может рассматриваться самостоятельно. Для нахождения его параметров используется метод наименьших квадратов (МНК).

***Система рекурсивных уравнений*** – в каждое последующее уравнение в качестве факторов включаются все зависимые переменные ***у*** предшествующих уравнений и набор фактора ***х***:



В таких моделях уравнения оцениваются последовательно (от первого уравнения к последнему) с использованием МНК.

***Система взаимозависимых уравнений*** – одни и те же зависимые переменные в одних уравнениях входят в левую часть, а в других – в правую часть системы:



***Структурной формой модели*** или ***системой одновременных уравнений*** (***СФМ)*** называется система уравнений, в каждом из которых аргументы содержат не только *объясняющие* переменные, но и *объясняемые* переменные из других уравнений.

Уравнения, составляющие исходную модель, называются ***структурными уравнениями модели***.

В процессе оценки параметров одновременных уравнений различают *эндогенные* и *экзогенные* переменные.

***Эндогенными*** («*эндо*» – внутренний) называются переменные, значения которых определяются *внутри* модели. Это зависимые переменные, число которых равно числу уравнений системы.

***Экзогенными*** («*экзо*» – внешний) называются такие переменные, значения которых определяются *вне* модели. Это заранее заданные переменные, влияющие на эндогенные переменные, но не зависящие от них.

В качестве экзогенных могут рассматриваться значения эндогенных переменных за предшествующий период времени (так называемые ***лаговые переменные***).

Простейшая структурная форма модели имеет вид:



где *уi* и *хi* – зависимая и независимая переменные, ε*i* – случайные ошибки, а *ai* и *bi* – параметры модели.

Параметры структурной формы модели называются ***структурными коэффициентами.*** Причем для зависимых переменных, как правило, используется коэффициент *b*, а для независимых – коэффициент *а*. Первый индекс указывает на номер объясняемой переменной (т.е. той зависимой переменной, которая стоит в левой части уравнения), а второй индекс соответствует объясняющей переменной. Например, структурный коэффициент *а*21 соответствует переменной *х*1, объясняющей зависимую переменную *у*2.

Если в каждом уравнении *отсутствует свободный член*, то значит, что все переменные в модели *выражены в отклонения от своего среднего уровня*, т.е. под *х* понимается , а под *у* – соответственно . Если же свободные члены присутствуют, то значит, в системе уравнений используются абсолютные значения всех переменных.

Использование МНК для оценки структурных коэффициентов модели дает, как правило, смещенные и несостоятельные оценки. Поэтому обычно для определения структурных коэффициентов модели структурная форма преобразуется в *приведенную форму модели*.

***Приведенной формой модели*** (***ПФМ)*** называется система линейных уравнений, в каждом из которых эндогенные переменные выражаются только через экзогенные и случайные ошибки:



где δ – ***приведенные коэффициенты модели***.

По своему виду приведенная форма представляет собой *систему независимых уравнений*, параметры которой оцениваются традиционным МНК. Применяя МНК, оцениваются приведенные коэффициенты модели δ, а затем осуществляется оценка эндогенных переменных через экзогенные (***косвенный метод наименьших квадратов КМНК*** см. далее).

Коэффициенты приведенной формы модели представляют собой ***нелинейные зависимости от коэффициентов структурной формы модели***.

Структурная модель в *полном виде* содержит ***n*·(*n* – 1 + *m*)** параметров, а приведенная форма модели в *полном виде* содержит ***n·m*** параметров, где *n* – количество *эндогенных* переменных, а *m* – количество *экзогенных* переменных в системе.

1. **Проблема идентификации.**

Приведенная форма модели аналитически уступает структурной форме, так как в ней отсутствуют оценки взаимосвязи между эндогенными переменными. Поэтому при переходе от приведенной формы модели к структурной возникает *проблема идентификации*.

***Идентификация*** – это единственность соответствия между приведенной и структурной формами модели. Поэтому структурная форма модели должна быть сначала *идентифицирована*.

1. СФМ является ***точно идентифицируемой***, когда число параметров СФМ ***равно*** числу параметров ПФМ.
2. СФМ ***сверхидентифицируема***, когда число приведенных коэффициентов ***больше*** числа структурных коэффициентов.
3. СФМ ***неидентифицируема***, когда число приведенных коэффициентов ***меньше*** числа структурных.

Причем модель считается *идентифицируемой*, если каждое ее *уравнение идентифицируемо*. Если хотя бы одно из уравнений системы *неидентифицируемо*, то и вся модель считается *неидентифицируемой*. Сверхидентифицируемая модель содержит хотя бы одно сверхидентифицируемое уравнение.

Выполнение условия идентифицируемости модели проверяется для каждого уравнения системы с помощью ***счетного правила***.

Пусть ***D*** – число ***экзогенных*** переменных системы, которые ***отсутствуют*** в данном уравнении, а ***H*** – число ***эндогенных*** переменных, ***присутствующих*** в данном уравнении. Тогда:

**Необходимое условие идентификации:**

* Уравнение ***точно идентифицировано***, если ***D* + 1 = *H***;
* Уравнение ***сверхидентифицировано***, если ***D* + 1 > *H***;
* Уравнение ***неидентифицировано***, если ***D* + 1 < *H***.

**Достаточное условие идентификации:**

Уравнение *идентифицируемо*, если ***ранг матрицы***, составленной из коэффициентов при *отсутствующих* в уравнении эндогенных и экзогенных переменных, ***не меньше*** ***n* – 1** (где *n* – количество эндогенных переменных системы), а ***определитель*** этой матрицы ***не равен нулю***.

1. **Оценка параметров структурной модели.**

Коэффициенты структурной модели могут быть оценены различными способами в зависимости от вида системы одновременных уравнений. Наибольшее распространение получили методы:

1. косвенный метод наименьших квадратов (КМНК);
2. двухшаговый метод наименьших квадратов (ДМНК).

**Косвенный метод наименьших квадратов.**

КМНК применяется в случае ***точно идентифицируемой структурной модели***. Выделяют следующие *этапы* КМНК:

1. Структурная модель преобразовывается в приведенную форму модели.
2. Для каждого уравнения приведенной формы модели традиционным МНК оцениваются приведенные коэффициенты δ*ij*.
3. Коэффициенты ПФМ преобразовываются в параметры СФМ методом подстановок.

**Двухшаговый метод наименьших квадратов.**

ДМНК используется, когда СФМ ***сверхидентифицируема***. Сверхидентифицируемая структурная модель может быть двух типов:

* все уравнения системы *сверхидентифицируемы*. И тогда для оценки структурных коэффициентов каждого уравнения используется ДМНК;
* система содержит наряду со сверхидентифицируемыми *точно идентифицируемые* уравнения, тогда структурные коэффициенты по ним находятся из системы приведенных уравнений.

*Этапы* ДМНК:

1. Составляется ПФМ и обычным МНК определяются численные значения параметров каждого ее уравнения.
2. Выявляются эндогенные переменные, находящиеся в правой части СФМ, параметры которой определяются ДМНК. Находят расчетные значения этих эндогенных переменных по соответствующим уравнениям ПФМ.
3. Обычным МНК определяют параметры структурного уравнения, используя в качестве исходных данных фактические значения объясняющих переменных и расчетные значения эндогенных переменных, стоящих в правой части данного структурного уравнения.

В случае **неидентифицируемости** модели, чтобы она имела статистическое решение, в нее вводятся либо дополнительные экзогенные переменные, приводящие к идентифицируемости уравнения, либо ненулевое ограничение, задающее соотношение между структурными коэффициентами неидентифицируемого уравнения, что также приводит к его идентифицируемости.