Министерство науки и высшего образования

Российской Федерации

Федеральное Государственное

Автономное Образовательное Учреждение

Высшего Образования

Национальный ядерный университет «МИФИ»

Кафедра: «Финансовый мониторинг»

Отчет по Лабораторной работе №5:

Студент Монастырский М. О.

Группа С21-703

Проверила: Домашова Д. В.

Оглавление

Элементы оглавления не найдены.

**Введение**

Одним из самых сложных этапов спецификации модели регрессии – параметризация, заключающаяся в выборе параметрического семейства функций , в рамках которого ищется неизвестная функция регрессии. Иногда подбор параметрического класса удается провести из соображений содержательного (экономического) характера, других соображений. Естественны попытки свести выбранную модель к линейной в целях упрощения оценки и исследование параметров модели.

В случае невозможности линеаризации модели оценка параметров модели может быть осуществлена методом наименьших квадратов, приводящего к решению нелинейной оптимизационной задачи МНК. В вычислительном плане, в настоящее время, это не представляет трудностей, но остаются проблемы с изучением статистических свойств оценок.

Большой интерес представляет построение линеаризуемой нелинейной зависимости в случаях, когда последняя неизвестна, методом Бокса-Кокса. При этом построенную нелинейную степенную зависимость можно рассматривать и как хорошую аппроксимацию достаточно широкого класса функций более общего вида. Освоение приемов подбора нелинейной регрессионной зависимости является целью предлагаемой работы.

**1 Описание лабораторной работы**

Лабораторная работа включает следующие этапы:

- постановку задачи;

- ознакомление с порядком выполнения работы в пакете Statistica 6.0;

- выполнение расчетов для индивидуальных задач;

- подготовку письменного отчета;

- защиту лабораторной работы.

**2 Постановка задачи**

По данным Приложения Б провести регрессионный анализ:

1 Из экономических или других соображений подобрать параметрический класс нелинейных зависимостей для модели регрессии.

2 Линеаризовать модель, оценить параметры и провести содержательный анализ.

3 Подобрать нелинейную модель, используя подход Бокса-Кокса и провести анализ модели.

1. **Порядок работы**

**3.1 Подбор и построение модели множественной регрессии**

**3.1.1 Линейная модель множественной регрессии**

|  |  |
| --- | --- |
| **Признак** | **Значение** |
| **Y2** | Ожидаемая продолжительность жизни при рождении (лет) |
| **X3** | Потребление сахара и кондитерских изделий на душу населения в кг |
| **X4** | Потребление картофеля на душу населения в кг |
| **X5** | Потребление масло растительного и других жиров на душу населения в кг |
| **X6** | Потребление мяса и мясных продуктов на душу населения в кг |
| **X7** | Потребление овощей и бахчевых на душу населения в кг |
| **X8** | Потребление фруктов и ягод на душу населения в кг |
| **X13** | Валовой региональный продукт на душу населения (тыс рублей) |

выявим зависимость между результативным признаком и объясняющими переменными. Вначале построим линейную функцию регрессии:



Multiple Regression Results

Dependent: Y2 Multiple R = ,72347295 F = 12,08079

R?= ,52341311 df = 7,77

No. of cases: 85 adjusted R?= ,48008703 p = ,000000

Standard error of estimate: 1,708779411

Intercept: 65,870050916 Std.Error: 1,656768 t( 77) = 39,758 p = 0,0000

X3 b\*=,229 X4 b\*=-,02 X5 b\*=-,13

X6 b\*=-,17 X7 b\*=,620 X8 b\*=-,08

X13 b\*=,385

(significant b\* are highlighted in red)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y2 (Sheet1 in данные лабы 5) R= ,72347295 R?= ,52341311 Adjusted R?= ,48008703 F(7,77)=12,081 p | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(77) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 65,87005 | 1,656768 | 39,75817 | 0,000000 |
| |  | | --- | | X3 | | 0,229087 | 0,097355 | 0,10301 | 0,043777 | 2,35311 | 0,021171 |
| |  | | --- | | X4 | | -0,022009 | 0,088773 | -0,00427 | 0,017243 | -0,24793 | 0,804852 |
| |  | | --- | | X5 | | -0,127484 | 0,096749 | -0,16055 | 0,121846 | -1,31768 | 0,191517 |
| |  | | --- | | X6 | | -0,170314 | 0,090982 | -0,03024 | 0,016152 | -1,87196 | 0,065010 |
| |  | | --- | | X7 | | 0,619567 | 0,126195 | 0,07944 | 0,016181 | 4,90960 | 0,000005 |
| |  | | --- | | X8 | | -0,081283 | 0,130588 | -0,01366 | 0,021945 | -0,62244 | 0,535493 |
| |  | | --- | | X13 | | 0,384996 | 0,089305 | 0,00027 | 0,000062 | 4,31101 | 0,000048 |

Поскольку можно допустить, нормальный характер распределения регрессионных остатков

После исключения мультиколлинеарности методом включения переменных

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: Y2 (Sheet1 in данные лабы 5) R= ,72154752 R?= ,52063082 Adjusted R?= ,49029100 F(5,79)=17,160 p | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(79) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 65,82588 | 1,585595 | 41,51493 | 0,000000 |
| |  | | --- | | X7 | | 0,560059 | 0,088519 | 0,07181 | 0,011350 | 6,32697 | 0,000000 |
| |  | | --- | | X13 | | 0,374203 | 0,084774 | 0,00026 | 0,000059 | 4,41412 | 0,000032 |
| |  | | --- | | X6 | | -0,181604 | 0,087074 | -0,03224 | 0,015458 | -2,08562 | 0,040241 |
| |  | | --- | | X3 | | 0,220622 | 0,095530 | 0,09921 | 0,042957 | 2,30944 | 0,023532 |
| |  | | --- | | X5 | | -0,128926 | 0,092522 | -0,16237 | 0,116522 | -1,39346 | 0,167387 |

 на основании отчета о результатах регрессионного анализа, делаем вывод:

* модель регрессии значима (p-value=0,016<0,05);

|  |  |
| --- | --- |
| **Признак** | **Значение** |
| **Y2** | Ожидаемая продолжительность жизни при рождении (лет) |
| **X3** | Потребление сахара и кондитерских изделий на душу населения в кг |
| **X4** | Потребление картофеля на душу населения в кг |
| **X5** | Потребление масло растительного и других жиров на душу населения в кг |
| **X6** | Потребление мяса и мясных продуктов на душу населения в кг |
| **X7** | Потребление овощей и бахчевых на душу населения в кг |
| **X8** | Потребление фруктов и ягод на душу населения в кг |
| **X13** | Валовой региональный продукт на душу населения (тыс рублей) |

* Переменные, существенно влияющие на результативный признак – Ожидаемая продолжительность жизни (лет)

X6 Потребление мяса и мясных продуктов на душу населения в кг

X7 Потребление овощей и бахчевых на душу населения в кг

X3 Потребление сахара и кондитерских изделий на душу населения в кг

X13 Валовой региональный продукт на душу населения (тыс рублей)

* коэффициент детерминации составил 0,52.

Оценка уравнения регрессии выглядит следующим образом:

Ожидаемая продолжительность жизни снижается на 0,032 года при увеличении Потребления мяса и мясных продуктов на душу населения на 1 кг

Ожидаемая продолжительность жизни увеличивается на 0,072 года при увеличении Потребление овощей и бахчевых на душу населения на 1 кг

Ожидаемая продолжительность жизни увеличивается на 0,1 года при увеличении Потребления сахара и кондитерских изделий на душу населения в кг на 1 кг

Ожидаемая продолжительность жизни увеличивается на года при увеличении Валовой региональный продукт на душу населения (тыс рублей) на 1 тыс рублей

На основании графического анализа проверим гипотезу о наличии положительной автокорреляции (Но: автокорреляция отсутствует) и проверим ее с помощью критерия Дарбина – Уотсона.

   

|  |  |  |
| --- | --- | --- |
|  | Durbin-Watson d (Sheet1 in данные лабы 5) and serial correlation of residuals | |
| |  | | --- | | Durbin- Watson d | | |  | | --- | | Serial Corr. | |
| |  | | --- | | Estimate | | 1,403521 | 0,284676 |

Для расчета критического значения критерия воспользуемся таблицей значений статистики Дарбина-Уотсана. Для n=87, k=3 получаем dн=1,55; dв=1,75 Значение статистики Дарбина-Уотсона для модели находится в интервале DW ∈ (0; dн ) = (0; 1,55), то гипотеза об отсутствии автокорреляции отклоняется, принимается гипотеза о наличии положительной автокореляции.

**Подбор нелинейной модели**

Анализируя данные, можно предположить, что функцию можно искать в форме Кобба-Дугласа:

Модель нелинейной регрессии:

где - регрессионные остатки.

Линеаризуем модель логарифмированием:

или

где

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=85 | Regression Summary for Dependent Variable: lnY (Sheet1 in данные лабы 5) R= ,67916755 R?= ,46126856 Adjusted R?= ,41229297 F(7,77)=9,4183 p | | | | | |
| |  | | --- | | b\* | | |  | | --- | | Std.Err. of b\* | | |  | | --- | | b | | |  | | --- | | Std.Err. of b | | |  | | --- | | t(77) | | |  | | --- | | p-value | |
| |  | | --- | | Intercept | |  |  | 3,944751 | 0,103786 | 38,00835 | 0,000000 |
| |  | | --- | | lnX3 | | 0,148159 | 0,106530 | 0,027483 | 0,019761 | 1,39077 | 0,168300 |
| |  | | --- | | lnX4 | | -0,089598 | 0,096278 | -0,013790 | 0,014818 | -0,93062 | 0,354960 |
| |  | | --- | | lnX5 | | -0,098910 | 0,106814 | -0,017236 | 0,018614 | -0,92601 | 0,357334 |
| |  | | --- | | lnX6 | | -0,189775 | 0,099578 | -0,043671 | 0,022915 | -1,90579 | 0,060410 |
| |  | | --- | | lnX7 | | 0,693749 | 0,134039 | 0,119335 | 0,023057 | 5,17572 | 0,000002 |
| |  | | --- | | lnX8 | | -0,099396 | 0,142477 | -0,016155 | 0,023157 | -0,69763 | 0,487512 |
| |  | | --- | | lnX13 | | 0,241563 | 0,102738 | 0,007074 | 0,003008 | 2,35125 | 0,021269 |



то на основании отчета делаем выводы:

* модель регрессии значима;
* коэффициент детерминации составил 0,46.

В результате получили следующее уравнение регрессии:

|  |  |  |
| --- | --- | --- |
|  | Durbin-Watson d (Sheet1 in данные лабы 5) and serial correlation of residuals | |
| |  | | --- | | Durbin- Watson d | | |  | | --- | | Serial Corr. | |
| |  | | --- | | Estimate | | 1,392928 | 0,289632 |

Для расчета критического значения критерия воспользуемся таблицей значений статистики Дарбина-Уотсана. Для n=85, k=2 получаем dн=1,6; dв=1,7 Значение статистики Дарбина-Уотсона для модели находится в интервале DW ∈ (0; dн ) = (0; 1,6), то гипотеза об отсутствии автокорреляции отвергается. Принимается гипотеза о наличии положительной автокорелляции