

Эконометрика, 2018-2019, 3 модуль

Семинар 3

28.01.19 для

Группы Э_Б2016_Э_3

Семинарист О.А.Демидова

Гетероскедастичность

Упражнение 10.1. (Учебник Демидова, Малахов)

По данным файла `clothing.dta`, содержащем данные о продажах одежды в 400 голландских магазинах мужской одежды

База данных `clothing` (файл `clothing.dta`).

В файле `clothing.dta` содержатся данные о продажах одежды в 400 немецких магазинах одежды, позаимствованные с сайта издательства книги Марно Вербика “A guide to modern econometrics”, <http://wileyurope.com/go/verbeek2ed>.

Переменные:

`tsales` – среднегодовые продажи в гюльденах,
`sales` - продажи в расчете на квадратный метр,
`margin` – маржинальная валовая прибыль,
`pown` – количество собственников (менеджеров),
`nfull` – количество полностью занятых,
`npart` - количество частично занятых,
`paux` – количество временно работающих,
`hoursw` – общее число отработанных часов,
`hourspw` – количество отработанных часов в расчете на одного работающего,
`inv1` – капиталовложения в помещения,
`inv2` - капиталовложения в автоматизацию,
`ssize` – размер магазина в м²,
`start` – год открытия магазина.

Оцените коэффициенты уравнения регрессии

$$sales = \beta_0 + \beta_1 hoursw + \beta_2 ssize + u.$$

- 1) Проведите тесты Уайта, Бройша – Пагана, Голдфелда-Квандта, Глейзера на выявление гетероскедастичности возмущений.
- 2) Если гетероскедастичность будет выявлена, то проведите коррекцию.

Методические рекомендации

- 1) Открыв файл `clothing.dta` в статистическом пакете STATA, оцените необходимую регрессию с помощью команды:

```
reg sales hoursw ssize
```

- 2) Для проведения теста Уайта после оценки регрессии необходимо набрать в командном окне:

```
estat imtest, white
```

- 3) Для проведения теста Бройша – Пагана необходимо набрать в командном окне

```
estat hettest, rhs mtest
```

- 4) Для проведения теста Глейзера необходимо сохранить остатки регрессии с помощью команды

```
predict res, resid
```

Создать новые переменные можно $|e|$, \sqrt{hoursw} , $\frac{1}{hoursw}$ с помощью команд:

```
gen modres = abs(res)
```

```
gen sqrthoursw = sqrt(hoursw)
```

```
gen invhoursw = 1/hoursw
```

Оценить коэффициенты трех регрессий:

$$|e| = \alpha + \beta hoursw + u, \quad |e| = \alpha + \beta \sqrt{hoursw} + u, \quad |e| = \alpha + \frac{\beta}{hoursw} + u$$

Можно с помощью команд

```
reg modres hoursw
```

```
reg modres sqrthoursw
```

```
reg modres invhoursw
```

- 5) Для проведения теста Голдфелда – Квандта необходимо упорядочить наблюдения по переменной `hoursw` с помощью команды

```
sort hoursw
```

Для того, чтобы оценить параметры уравнения регрессии по первой и последней трети (приблизительно) наблюдений и сохранить соответствующие RSS, наберите в командном окне

```
reg sales hoursw ssize in 1/133
```

```
scalar rss1 = e(rss)
```

```
reg sales hoursw ssize in 268/400
```

```
scalar rss2 = e(rss)
```

Используя RSS в последних двух оцененных регрессиях, рассчитать тестовую статистику

можно по формуле $F = \frac{RSS_2 / (n_2 - k)}{RSS_1 / (n_1 - k)} = \frac{RSS_2}{RSS_1} =$

С помощью команды:

```
scalar F = rss2/rss1
```

```
display F
```

а соответствующее критическое значение F-статистики с помощью команды:

```
display invFtail(130,130,0.05)
```

б) Применим два способа для решения проблемы гетероскедастичности:

Первым способом является преобразование исходных переменных.

Для этого создадим новые переменные

$$sales_new = \frac{sales}{hoursw}, \quad cons_new = \frac{1}{hoursw}, \quad size_new = \frac{size}{hoursw}$$

с помощью команд

```
gen sales_new = sales/hoursw
```

```
gen cons_new = 1/hoursw
```

```
gen size_new = ssize/hoursw
```

Оценить коэффициенты новой регрессии можно с помощью команды

```
reg sales_new cons_new size_new
```

Нужно быть внимательным при интерпретации оценок коэффициентов новой регрессии! Вторым способом решения проблемы гетероскедастичности возмущений является использование оценок Уайта (White, 1980) для дисперсий коэффициентов.

Их можно получить с помощью команды `reg sales hoursw ssize, robust`

Проверка нормальности распределения остатков регрессии

По данным файла `clothing.dta`, содержащем данные о продажах одежды в 400 голландских магазинах мужской одежды

1) Оцените коэффициенты уравнения регрессии

$$sales = \beta_0 + \beta_1 hoursw + \beta_2 size + u.$$

2) Постройте гистограмму распределения остатков регрессии и `qqplot`.

С помощью тестов Колмогорова-Смирнова, Харке-Бера, Шапиро-Уилка проверьте гипотезу о нормальности распределения случайных ошибок.

3) Рассчитайте Стьюдентизированные остатки регрессии, постройте их график и с его помощью определите возможные наблюдения, которые являются выбросами. Если такие наблюдения будут выявлены, то оцените регрессию

а) без наблюдений, являющихся выбросами,

б) R-регрессию Хубера.

4) Сравните результаты оценки регрессий, оцененных с помощью МНК (с выбросами и без), с оценками Уайта для стандартных ошибок, R-регрессии Хуберта.

Методические рекомендации

1) Открыв файл `clothing.dta` в статистическом пакете STATA, оцените необходимую регрессию с помощью команды:

```
reg sales hoursw ssize
```

2) Сохраните остатки регрессии с помощью команды

```
predict res, resid
```

3) Постройте их гистограмму с помощью команды `hist res`

Qq plot с помощью команды

```
qnorm res
```

4) Проведите тест Колмогорова-Смирнова с помощью команды

```
sum res
```

```
ksmirnov res = normal((res-r(mean))/r(sd))
```

5) Проведите тест Харке-Бера с помощью команды `sktest res, noadjust`

6) Проведите тест Шапиро-Уилка с помощью команды `swilk res`

7) Постройте график студентизированных остатков регрессии с помощью команды

```
predict residst, rstudent
```

```
twoway(scatter residst number, sort)
```

```
list number if abs(residst) > 2
```

Оценить регрессию без наблюдений-выбросов можно с помощью команды

```
reg sales hoursw ssize if abs(residst) < 2
```

8) Оценить R-регрессию Хубера можно с помощью команды

```
rreg sales hoursw ssize
```

9) Сохранить результаты оценки регрессии можно с помощью команды

```
est store reg1 (и т.д.)
```

Для сравнения результатов удобно сформировать общую таблицу с помощью команды

```
est tab reg1 (и т.д.), star (0.1 0.05 0.01) b(%7.3f)
```