

## Семинары №6-7.

### Проверка линейных гипотез. Тест Чоу.

1. С помощью модели  $\ln Y = \beta_0 + \beta_1 \ln L + \beta_2 \ln K + \varepsilon$   
по данным для 30 фирм была оценена зависимость выпуска  $Y$  от труда  $L$  и капитала  $K$ :

$$\hat{\ln Y} = 1.2 + 0.6 \ln L + 0.4 \ln K, \quad F\text{-statistic} = 200.24$$

(0.3)    (0.12)            (0.08)

В скобках указаны значения стандартных ошибок.

На уровне значимости 5 % отвергаются гипотезы

- 1)  $H_0 : \beta_1 = 0$    2)  $H_0 : \beta_2 = 0$    3)  $H_0 : \beta_1 = \beta_2 = 0$    4)  $H_0 : \beta_1 = 0.5$    5)  $H_0 : \beta_2 = 0.5$

2. Оценка производственной функции Кобба - Дугласа с помощью модели

$$\ln Q = \beta_0 + \beta_1 \ln K + \beta_2 \ln L + \varepsilon,$$

где  $Q$  – выпуск,  $K$  – капитал,  $L$  – труд,

по 40 наблюдениям дала следующие результаты (в скобках указаны стандартные ошибки коэффициентов регрессии):

$$\ln \hat{Q} = 1.37 + 0.632 \ln K + 0.452 \ln L, \quad R^2 = 0.98, \quad \text{cov}(\hat{\beta}_1, \hat{\beta}_2) = -0.044$$

(0.257)                      (0.219)

На уровне значимости 5% проверить гипотезы

- а) о равенстве эластичностей по труду и капиталу,  
б) о постоянной отдаче от масштаба

3. По данным для 27 фирм оценили производственную функцию с помощью трех моделей:

$$\ln Y = \beta_0 + \beta_1 \ln L + \beta_2 \ln K \quad (1)$$

$$\ln Y = \beta_0 + \beta_1 \ln(LK) \quad (2)$$

$$\ln Y / K = \beta_0 + \beta_1 \ln L / K \quad (3)$$

Суммы квадратов остатков в этих моделях оказались соответственно равны

$$RSS_1 = 8.51, \quad RSS_2 = 8.94,$$

$$RSS_3 = 9.01,$$

Объяснить, почему вторая и третья модели являются ограниченными версиями первой, выписать соответствующие ограничения на коэффициенты регрессии и проверить их выполнение.

4. а) Изучая зависимость длительности обучения индивида НГС от его способностей ASVABC, характеризуемых результатами трех тестов (см. п.в), длительности обучения

матери индивида  $HGCM$ , длительности обучения отца индивида  $HGCF$ , исследователь получил следующие регрессии:

$$HGC = \underset{(0.52)}{5} + \underset{(0.0099)}{0.115} ASVABC + \underset{(0.039)}{0.12} HGCM + \underset{(0.029)}{0.1} HGCF, \text{RSS} = 2100.646, R^2 = 0.336$$

$$HGC = \underset{(0.48)}{6.5} + \underset{(0.009)}{0.14} ASVABC, \text{RSS} = 2267.587$$

Исходя из полученных результатов, можно ли считать, что длительность обучения индивида зависит только от его способностей?

б) Была оценена также регрессия

$$HGC = \underset{(0.502)}{5.22} + \underset{(0.0099)}{0.115} ASVABC + \underset{(0.016)}{0.109}(HGCM + HGCF), \text{RSS} = 2100.962.$$

Исходя из полученного результата, можно ли считать, что родители в равной степени влияют на длительность обучения индивида?

5. Исследователя интересует зависимость среднегодового прироста работающих  $E$  от прироста валового национального продукта  $X$  ( $E$  и  $X$  измеряются в процентах).

Были оценены регрессии:

- по выборке для 27 развитых стран:

$$E = \underset{0.58}{-2.74} + \underset{(0.15)}{0.5} X \quad \text{RSS}_1 = 18.63$$

- по выборке для 23 развивающихся стран:

$$E = \underset{0.42}{-0.85} + \underset{(0.15)}{0.78} X \quad \text{RSS}_2 = 25.23$$

- по общей выборке:

$$E = \underset{0.53}{-0.56} + \underset{(0.16)}{0.24} X \quad \text{RSS}_3 = 121.61.$$

На уровне значимости 5% проверить гипотезу о том, что зависимость прироста работающих от прироста ВВП едина для развитых и развивающихся стран.

### Фиктивные (дамми) переменные.

6. Зависимость расходов на продукты питания от располагаемого дохода  $X$  имеет вид:

$$Y = 2 + 0.6X + 0.07D_1X,$$

где  $D_1$  – фиктивная переменная, равная 1 для городских и 0 для сельских жителей.

а) Коэффициент наклона в линейной зависимости для сельских жителей равен

- 1) 0,67      2) 0,6      3) 0,53      4) 2

б) Если вместо  $D_1$  использовать переменную  $D_2$ , равную 0 для городских и 1 для сельских жителей, то зависимость примет вид:

- 1)  $Y = 2 + 0.67X - 0.07D_2X$

2)  $Y = 2 + 0.67X + 0.07D_2X$

3)  $Y = 2 + 0.6X - 0.07D_2X$

4)  $Y = 2.07 + 0.6X - 0.07D_2X$ .

7. Оцененная зависимость  $Y$  - расходов потребителей на газ и электричество в США в 1977 – 1999 г.г. в постоянных ценах I квартала 1977г. от времени  $t = 1$  для 1977 г.,  $t = 2$  для 1978 г. и т.д. с учетом сезонных факторов ( $D_i = 1$ , если наблюдение относится к  $i$ -му кварталу и 0 иначе,  $i = 1, \dots, 4$ ) имеет вид:

$$Y = 8 + 0.1t - 3D_2 - 2.6D_3 - 2D_4$$

Если в качестве выделенной категории будет выбран не первый квартал, а второй, то уравнение регрессии примет вид:

1)  $Y = 5 + 0.1t + 3D_1 + 0.4D_3 + D_4$

2)  $Y = 8 + 0.1t - 3D_1 - 2.6D_3 - 2D_4$

3)  $Y = 5 + 0.1t - 3D_1 - 2.6D_3 - 2D_4$

4)  $Y = 5 + 0.1t - 3D_2 - 0.4D_3 - D_4$ .

8. Оцененная зависимость почасовой оплаты труда американцев  $Y$  (измеряется в долларах в час) от длительности обучения  $X$  (измеряется в годах) и расовой принадлежности, описываемой с помощью фиктивной переменной  $D$ , равной 1 для светлокожих и 0 для темнокожих американцев, имеет вид:  $Y = 5 + 0.7X + 0.04DX$ .

Все коэффициенты являются значимыми при уровне значимости 1%.

Каждый дополнительный год обучения приводит к увеличению почасовой оплаты труда темнокожих американцев на

- 1) 0.74 \$    2) 0.7 \$    3) 0.66 \$    4) 0.74 %

### **Интерпретация коэффициентов различных форм уравнений МНК. Ошибки спецификации.**

*При не включении в регрессию существенных переменных оценки коэффициентов оказываются смещенными. Последствием этой ошибки может явиться автокорреляция.*

*Диагностику этой ошибки можно осуществить с помощью теста Рамсея.*

*При включении в регрессию несущественных переменных оценки коэффициентов остаются несмещенными, но теряют эффективность. Последствием этой ошибки может явиться*

*квазимультиколлинеарность. Диагностику этой ошибки можно осуществить с помощью F-теста на группу лишних переменных.*

9. Зависимость расходов на непродовольственные товары  $Y$  от располагаемого дохода  $X$  имеет вид:  $\ln \hat{Y} = -1.2 + 0.02X$  (все коэффициенты регрессии значимы). При увеличении дохода на 1 единицу расходы увеличатся на

1) 0.02 единицы    2) 2 единицы    3) 2%    4) 0.02%    5) 20%

10. Зависимость расходов на продовольственные товары  $Y$  от располагаемого дохода  $X$  имеет вид:  $\ln \hat{Y} = -1.2 + 2 \ln X$  (все коэффициенты регрессии значимы). При увеличении дохода на 1% расходы увеличатся на

1) 0.02 единицы    2) 2 единицы    3) 2%    4) 0.02%    5) 20%

11. По ежемесячным данным за 3 года была оценена зависимость количества потребляемых безалкогольных напитков  $Y$  от температуры  $T$  и реальных денежных доходов населения  $IR$  с помощью трех моделей:

$$\hat{Y} = -1877.04 + 21.97 T + 42.84 IR \quad R^2 = 0.28 \quad (1)$$

(149414)      (8.64)      (20.19)

$$\hat{Y} = -1226 + 35.54 IR \quad R^2 = 0.09 \quad (2)$$

(162486)      (22.06)

$$\hat{Y} = 1285.665 + 19.36 T \quad R^2 = 0.15 \quad (3)$$

(10369)      (9.13)

причем  $\hat{c\hat{o}r}(IR, T) = -0.14$ .

Какую модель Вы предпочтете и почему? Объясните изменения в коэффициентах и их стандартных ошибках в остальных моделях.

12. Исследователь оценил зависимость уровня активности в теневой экономике  $Y$  от уровня налогообложения  $X$  и государственных расходов на борьбу с теневой экономикой  $Z$  (все переменные измеряются в миллионах долларов США) с помощью трех моделей  
а) по данным для 30 развитых стран:

$$\ln Y = -1.137 + 0.699 \ln X - 0.646 \ln Z, R^2 = 0.44$$

(0.863)      (0.154)      (0.162)

$$\ln Y = -1.065 + 0.201 \ln X, R^2 = 0.1$$

(1.069)      (0.112)

$$\ln Y = 1.23 - 0.053 \ln Z, R^2 = 0.01$$

(0.896)      (0.162)

б) и данным для 30 развивающихся стран:

$$\ln Y = -1.122 + 0.806 \ln X - 0.091 \ln Z, R^2 = 0.71$$

(0.873)      (0.137)      (0.117)

$$\ln Y = -1.024 + 0.727 \ln X, R^2 = 0.7$$

(0.859)      (0.09)

$$\ln Y = 2.824 + 0.427 \ln Z, R^2 = 0.33$$

(0.835)      (0.116)

Какую из моделей Вы предпочтете в случаях а и б? Объясните изменения в коэффициентах и их стандартных ошибках в остальных моделях.

### Задание для выполнения на компьютерах

1. Откройте файл **housing.dta**. Описание переменных содержится в файле **housing.txt**. Посмотрите на описательные статистики переменных.
2. Постройте уравнение регрессии вида:  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{11} X_{11} + \varepsilon$ , где  $Y$  - цена продажи дома ( в канадских долларах),  $X_1, \dots, X_{11}$  все остальные переменные, которые есть в файле.
3. Проверьте адекватность регрессии в целом и значимость коэффициентов по отдельности. Дайте экономическую интерпретацию полученным результатам.
4. Попробуйте исключить из модели какую-нибудь значимую переменную. Что изменилось? Что произошло с коэффициентами модели?
5. Проверьте гипотезу о совместной незначимости группы переменных.
6. Исключите из модели незначимые переменные. Что изменилось?
7. Проведите тест Рамсея для проверки гипотезы о существовании упущенных переменных для вашей модели. Сделайте вывод.
8. Создайте логарифм переменной `price` (цена продажи дома). Оцените полулогарифмическую модель.  $\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{11} X_{11} + \varepsilon$ . Проинтерпретируйте полученные результаты. Что показывают коэффициенты в такой модели?
9. На основании теста Бокса-Кокса сделайте вывод о том, какая модель лучше: линейная или полулогарифмическая.
10. Постройте прогноз  $\hat{Y}$  для выбранной вами модели. Получите ряд остатков. Постройте графики для прогноза и для остатков. Какой можно сделать вывод на основании этих графиков.