

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение высшего профессионального
образования «Уральский государственный университет им. А.М. Горького»

ИОНЦ «Бизнес-информатика»

Экономический факультет

Кафедра экономического моделирования и информатики

ЭКОНОМЕТРИКА

Экзаменационные билеты

Подпись руководителя ИОНЦ

Дата

Екатеринбург
2007

Промежуточный экзамен.

I. (10 минут).

1.1. В чем заключаются условия Гаусса-Маркова применительно к регрессионной модели? Какого типа модели подразумеваются этими условиями?

1.2. Какими свойствами обладают оценки коэффициентов регрессии в предположении выполнимости каких предпосылок Гаусса-Маркова. В чем прикладной смысл этих свойств?

1.3. Зачем используется дополнительное условие нормальности распределения случайного члена?

II. Из приведенных вариантов выберите 1 (20 минут).

1. Укажите правильное соотношение между свойствами оценок в статистике и эконометрике

- a) несмещенность является достаточным условием эффективности
- b) эффективность является достаточным условием несмещенности
- c) состоятельность является достаточным условием несмещенности
- d) эффективность является необходимым условием состоятельности
- e) несмещенность является необходимым условием эффективности.

2. Используя данные для США за 1959-1983 гг. по затратам на питание и располагаемому личному доходу, было получено уравнение регрессии:

$\hat{y} = 55,3 + 0,093x$ (где y и x измерены в миллиардах долларов США в постоянных ценах 1972 г.). Имеются следующие интерпретации коэффициента при x :

I. Эластичность расходов на питание по доходу составляет 0,093.

II. При возрастании личного располагаемого дохода на сто долларов расходы на питание возрастают на 9,3 долларов.

III. Доля расходов на питание в величине личного располагаемого дохода составляет 9,3%.

Какие из приведенных интерпретаций являются по существу верными?

- a) Только I.
- b) Только II.
- c) Только III.
- d) I и III.
- e) II и III.

3. При исследовании вопроса о значимости парной регрессионной модели, возможны следующие соотношения между t , F , R^2 :

I. F и R^2 свидетельствуют о статистической незначимости модели, а t говорит о статистической незначимости коэффициента регрессии

II. t – велико (значимо), а F , а R^2 – малы, так что регрессия в целом незначима

III. t - и F -тесты всегда дают тождественные результаты, R^2 однозначно определяется их значениями

IV. t , R^2 , – малы, так что применимые к ним тесты незначимы, а F – велико (значимо)

V. t , F , R^2 - значимы, но при различных уровнях доверия

- a) I, II, III, IV
- b) III, V
- c) Только V
- d) Только III
- e) I, II, IV V

4. Величина стандартного отклонения МНК-оценки коэффициентов регрессии, обычно приводимая в статистических пакетах, используется при построении модели множественной регрессии для:

- a) отыскания параметров распределения случайных остатков
- b) проверки гипотезы о значимости модели
- c) проверки гипотезы о нулевом значении соответствующего коэффициента

- d) проверки гипотезы о независимости случайных остатков и соответствующих МНК-оценок
- e) построения F-статистики

5. Исходя из теоретических соображений можно предполагать знак правильно оцененного коэффициента регрессии отрицательным. Однако в уравнении регрессии он оказался положительным и незначимым на уровне 5%. Что следует предпринять?

- a) Отказаться от дальнейшей проверки и считать его незначимым
- b) Можно попытаться проверить его значимость на уровне 1% с помощью одностороннего критерия, используя в качестве альтернативной гипотезу “коэффициент больше нуля”
- c) Можно попытаться проверить его значимость на уровне 5% с помощью одностороннего критерия, используя в качестве альтернативной гипотезу “коэффициент меньше нуля”
- d) Можно попытаться проверить его значимость на уровне 5% с помощью одностороннего критерия, используя в качестве альтернативной гипотезу “коэффициент больше нуля”
- e) В данном случае нельзя ограничиться одним уровнем значимости. Следует проверить его значимость по двухстороннему критерию на уровне 1%

6. Коэффициент регрессии в линейной регрессии совокупного спроса на мобильные телефоны (в тысячах рублей) по цене (в рублях) оказался равным –

1. Это означает, что

- a) увеличение цены на 1% снижает спрос на мобильные телефоны на 1%
- b) увеличение цены на 1 рубль снижает спрос на мобильные телефоны на 1%
- c) увеличение цены на 1% снижает спрос на мобильные телефоны на одну тысячу рублей
- d) увеличение цены на 1 рубль снижает спрос на мобильные телефоны на одну тысячу рублей

е) полученное число никак не интерпретируется

7. Коэффициент наклона линии регрессии

- а) всегда находится от -1 до 1
- б) никогда не бывает отрицательным
- в) равен коэффициенту корреляции
- г) не может быть равен нулю
- е) может принимать любое значение

8. Если объясняющая переменная x в модели парной линейной регрессии принимает

среднее в выборке значение $\bar{x} = \sum x_i/n$, то

- а) наблюдаемая величина зависимой переменной y равна $y = \sum y_i/n$;
- б) рассчитанная по уравнению регрессии величина зависимой переменной y равна $y = \sum y_i/n$;
- в) рассчитанная по уравнению регрессии величина зависимой переменной y равна $y = \sum y_i/n$ в среднем, но не в каждом конкретном случае;
- г) равенство $y = a + b x$ свидетельствует об отсутствии корреляции между отклонениями;
- е) равенство $y = a + b x$ свидетельствует о выполнении условий Гаусса-Маркова.

9. Введение линейного ограничения на параметры в регрессионную модель

- а) приводит к незначительному сокращению суммы квадратов отклонений, если ограничение значимо
- б) приводит к значительному сокращению суммы квадратов отклонений, если ограничение значимо
- в) приводит к незначительному сокращению суммы квадратов отклонений, если ограничение незначимо
- г) приводит к значительному сокращению суммы квадратов отклонений, если ограничение незначимо
- е) все перечисленное неверно

10. Оценена линейная регрессионная зависимость величины заработка респондента y от совокупности факторов (x_1, x_2, \dots, x_n) . Исследователь предполагает, что величина заработка зависит также от места проживания респондента, которое может принадлежать к одному из трех типов: крупный город, малый город, сельская местность. Для того, чтобы учесть это в модели, нужно:

- a) ввести фиктивную переменную, которая может принимать три разных значения;
- b) ввести фиктивную переменную, которая может принимать два разных значения;
- c) ввести две фиктивные переменные, каждая из которых может принимать два разных значения;
- d) ввести три фиктивные переменные, каждая из которых может принимать два разных значения;
- e) ввести две фиктивные переменные, каждая из которых может принимать три разных значения.

III. (10 минут).

3.1. Ковариационная матрица для переменных модели выглядит следующим образом.

	Y	X_1	X_2
Y	458.4986	99.97390	96.99030
X_1	99.97390	57.53368	24.81317
X_2	96.99030	24.81317	77.88488

Определите коэффициенты корреляции между переменными модели, взятыми попарно.

Какие выводы можно сделать из этих данных?

3.2. Далее была оценена парная регрессия

$$\hat{Y} = -33.7 + \dots X_1, R^2 = 0.45 \text{ Воспользовавшись имеющимися данными,} \\ (2.7) \dots$$

рассчитайте коэффициент наклона и стандартную ошибку в модели.

IV. (30 минут).

Имеется выборка 570 наблюдений молодых людей по следующим переменным: Y – зарплата (руб/час); S – число лет, потраченных на обучение; уровень образования: среднее, среднее техническое, высшее, второе высшее. Для последнего фактора введены четыре дамми-переменные для каждого уровня соответственно: $EDU1$, $EDU2$, $EDU3$, $EDU4$. Исследователь осуществляет регрессию натурального логарифма Y на

- 1) только S ;
- 2) только на $EDU2$, $EDU3$, $EDU4$;
- 3) на S , $EDU2$, $EDU3$, $EDU4$.

Результаты приведены в таблице (в скобках стандартные ошибки):

	(1)	(2)	(3)
константа	1.359 (0.113)	2.321 (0.027)	1.842 (0.236)
S	0.079 (0.008)		0.040 (0.019)
$EDU2$		-0.173 (0.075)	-0.055 (0.094)
$EDU3$		0.042 (0.074)	0.065 (0.080)
$EDU4$		0.420 (0.047)	0.246 (0.095)
R^2	0.141	0.145	0.15
ESS	132.1	131.48	130.44

Вопросы:

1. (2 минуты) Дайте интерпретацию коэффициента при S в первой регрессии.

2. (4 минуты) В уравнения не включена переменная $EDU1$ для среднего образования. Исследователь говорит, что он не включил ее в уравнение потому что эта переменная для низшего уровня образования. Что вы ему скажете.
3. (2 минуты) Проинтерпретируйте коэффициенты при дамми-переменных во второй модели.
4. (4 минуты) Проинтерпретируйте все коэффициенты в третьей модели.
5. (4 минуты) Проверьте значимость всех трех регрессий.
6. (6 минут) Проверьте значимость коэффициентов в третьей модели.
7. (4 минуты) Проверьте гипотезу о том, что уровень образования не влияет на заработную плату.
8. (4 минуты) Имеется предположение о том, что коэффициенты для женщин в модели 2 при переменных $EDU3$ и $EDU4$ меньше, чем для мужчин. Что это означает содержательно. Как можно проверить эту гипотезу.

V. (10 минут)

5.1. Уравнение зависимости уровня расходов на косметику от личного располагаемого дохода получилось следующим

$$\hat{Y} = -34,02 + 0,1X \quad . \quad R^2 = 0.95$$

(1.95) (0.02)

Дайте интерпретацию модели и ее коэффициентов. Проверьте значимость регрессии в целом и значимость коэффициентов.

5.2. Если, взять логарифм величины расходов, то уравнение принимает вид

$$\ln(Y) = 0.77 + 0.0014X \quad . \quad R^2 = 0.83$$

(0.1) (0.00013)

Дайте интерпретацию коэффициентов модели. Запишите уравнение, не используя логарифмы.

5.3. Модель в двойных логарифмах принимает вид

$$\ln(Y) = -5.2 + 0.98X \quad . \quad R^2 = 0.9$$

(0.48) (0.07)

Дайте интерпретацию коэффициентов модели. Запишите уравнение, не используя логарифмы.

5.4. Выберите из трех моделей наиболее адекватную

Итоговый экзамен.

1. **Обсудите различие подходов для одной из предлагаемых пар (10 баллов):**

- a) OLS vs. GLS
- b) подход «снизу вверх» vs. «сверху вниз»
- c) обычные стандартные ошибки vs. стандартные ошибки в форме Уайта или Ньюи-Веста.

2. **Ответьте на каждый вопрос каждый – 4 балла, последние два – 8 баллов.**

- a) Объясните, что значит состоятельная оценка, эффективная оценка. Правда ли, что несмещенная оценка всегда предпочтительнее, чем смещенная?
- b) Как вы можете протестировать присутствие гетероскедастичности в модели.
- c) Почему предположение о нормальности ошибки важно в эконометрике. Нужно ли расстраиваться, если ошибки не имеют нормального распределения? Почему.
- d) Объясните, что вы понимаете под тестом Дарбина-Уотсона, Какие предпосылки должны выполняться при использовании этого теста.
- e) Слабо значимая переменная – переменная, p – значение которой лежит между 0,1 и 0,05. Следует ли такую переменную удалять из модели?
- f) Покажите, что $\hat{\beta} = \frac{\bar{Y}}{\bar{X}}$ является несмещенной оценкой коэффициента β в модели $Y = \beta X + \varepsilon$. Какие предположения вы использовали. Какая статистическая оценка является BLUE. При каких условиях.
- g) Как известно, не все регрессионные модели могут быть сведены к линейной. Приведите пример такой модели. Для оценивания параметров

таких моделей применяется нелинейный метод наименьших квадратов. Может быть вы можете выписать функцию, которая минимизируется в этом методе для вашего примера и систему нормальных уравнений (преобразовывать ее не надо). Какие свойства будут у этого метода.

- h) Рассматривается обобщенная линейная регрессионная модель, матрица ковариаций которой известна (для простоты). Тем не менее, кроме обобщенного МНК эта модель еще была оценена МНК без каких-либо коррекций. Ваш друг утверждает, что необходимо брать результаты обычного МНК, т. к. там стандартные ошибки коэффициентов меньше, стандартная ошибка регрессии меньше, а коэффициент детерминации больше. Что вы ему ответите.

Ответе на три вопроса (3 или 4, 5 или 6, 6 или 7). Каждый – 20 баллов.

3. Рассмотрим регрессионную модель $Y_t = \alpha + \beta X_t + \varepsilon_t$, где $\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + u_t$, где $u_t \sim iid(0, \sigma^2)$.

- Почему в выражении для ошибки отсутствует свободный коэффициент?
- Получите выражения для $Var(\varepsilon_t)$, $Cov(\varepsilon_t, \varepsilon_{t-1})$, $Cov(\varepsilon_t, \varepsilon_{t-2})$, $ACF(1)$, $ACF(2)$ (подсказка – возникнет система для нахождения ковариаций (она называется система Юла-Уолкера), зная ковариации, можно будет найти дисперсию – дисперсия это тоже ковариация).
- Объясните, как вы будете проверять гипотезу $H_0 : \rho_1 = \rho_2 = 0$
- Опишите последствия применения МНК к уравнению $Y_t = \alpha + \beta X_t + \varepsilon_t$.
- Получите наилучшую линейную несмещенную оценку для α и β (о первых наблюдениях можно не беспокоиться (для скольких первых?)).

4. Рассмотрим регрессионную модель $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$, $\varepsilon_i \sim iidN(0, \sigma^2)$.

а) Объясните, как вы будете тестировать следующие гипотезы (обратите внимание на число степеней свободы, статистический тест, регрессии, которые должны посчитать)

<p>i) $H_0 : \beta_1 = \beta_2, \beta_3 = -2$ $H_a : \beta_1 \neq \beta_2 \text{ (или)} \beta_3 \neq -2$</p> <p>ii) $H_0 : \beta_1 + \beta_2 + \beta_3 = 1$ $H_a : \beta_1 + \beta_2 + \beta_3 = c \quad c \neq 1$</p> <p>iii) $H_0 : \beta_1 + \beta_2 = 1$ - проверить численно $H_a : \beta_1 + \beta_2 < 1$</p> <p>($N = 50$)</p>	<p>Дана матрица ковариаций вектора $\hat{\beta}_{OLS}$:</p> $\Sigma_{\hat{\beta}_{OLS}} = s^2 (X'X)^{-1} = \begin{pmatrix} 2.25 & -0.5 & 0.2 & 1.2 \\ -0.5 & 3.1 & -0.35 & 2.4 \\ 0.2 & -0.35 & 0.7 & 0.15 \\ 1.2 & 2.4 & 0.15 & 1.4 \end{pmatrix}$ <p>$\hat{\beta}_1 = 0,56, \hat{\beta}_2 = 0,48$</p>
---	--

б) Как вычислить матрицу ковариаций $\hat{\beta}_{OLS}$, в случае обобщенной регрессионной модели (если знаете формулы – замечательно, нет – опишите процесс)? А если применяем обобщенный МНК?

5. Экономист располагает следующими данными: спрос на труд L , совокупный объем промышленной продукции в текущих ценах Y , средняя зарплата в текущих ценах W , индекс цен P для промышленного сектора некоторой страны за 1975-2004 гг. Он оценил следующие регрессии (в скобках стандартные ошибки):

$$\log L = -3.12 + 0.42 \log Y - 0.34 \log W - 0.11 \log P, \text{ ESS} = 1.99 \quad (1)$$

(0,13) (0,09) (0,1) (0,06)

$$\log L = -2.56 + 0.46 \log \frac{Y}{P} - 0.32 \log \frac{W}{P}, \text{ ESS} = 2,03 \quad (2)$$

(0,13) (0,07) (0,1)

$$\log L = -3.12 + 0.42 \log \frac{Y}{P} - 0.34 \log \frac{W}{P} - 0.03 \log P, \text{ ESS} = 1.99 \quad (3)$$

(0,13) (0,09) (0,1) (0,04)

- a) дайте экономическую интерпретацию коэффициентов уравнения в обеих регрессиях;
- b) Объясните почему вторая регрессия является регрессией с ограничениями для первой. Сформулируйте ограничение.
- c) Проверьте его F – тестом.
- d) проверьте его t – тестом.
- e) Могут ли F -тест и t -тест дать разные результаты. Ответ поясните.
- f) Какую спецификацию вы предпочтете по результатам теста. Почему.
- g) У регрессии (2) хуже качество подгонки, чем у (1). Это важно?

6. Регрессия совокупного потребления (C) на совокупный доход (X) и уровень безработицы (U) по данным за 1985 – 2003 для США дала следующие результаты (в скобках стандартные ошибки):

$$\hat{C} = 17880 + 0.7527X + 0.93U, R^2 = 0.988$$

(28170) (0,026) (0,798)

Таблица коэффициентов корреляции между переменными.

	C	X	U
C	1	0.996	0.783
X	0.996	1	0,771
U	0.783	0,771	1

- a) Проверьте гипотезы о значимости регрессии в целом и значимости коэффициентов уравнения. Какие предположения необходимы для корректного выполнения тестов.
- b) Постройте 95% доверительные интервалы для коэффициентов уравнения.
- c) Проинтерпретируйте результаты оценивания и приведенную таблицу коэффициентов корреляции. Проблема ли здесь мультиколлинеарность? Объясните. Что делать, если это проблема?

d) Если осуществить регрессию потребления только на доход, как вы думаете, что произойдет с коэффициентом при доходе? Объясните.

7. Исследователь располагает данными о расходах на книги домашнего хозяйства (B), Доходе домашнего хозяйства (X) и числа лет образования главы семьи (S) для 100 домашних хозяйств в 2003 г. Он предположил, что переменные связаны следующим соотношением: $\log B = \beta_0 + \beta_1 \log X + \beta_2 \log S + \varepsilon$. Однако исследователь не уверен в значимости S и оценивает еще короткую модель $\log B = \beta_0 + \beta_1 \log X + \varepsilon$.

Результаты оценивания приведены в таблице (в скобках стандартные ошибки):

	(1)	(2)
$\log X$	1.1	2.1
	(0.69)	(0.35)
$\log S$	0.59	-
	(0.35)	
const	-6.89	-3.37
	(2.28)	(0.89)
R^2	0.29	0.27

- Предположив, что (1) это правильная спецификация, покажите с математическим доказательством почему ожидается, что коэффициент при $\log X$ будет больше, чем в регрессии (2).
- Предположив, что (2) – правильная спецификация, объясните, почему в регрессии (1) коэффициент ниже. А вы бы что предположили?
- Предположив, что (1) – правильная спецификация, поясните, будут ли стандартные ошибки в регрессии (2) являться «правильными» оценками.

d) Предположив, что (1) – правильная спецификация, поясните, будут ли стандартные ошибки в регрессии (2) являться «правильными» оценками (что значит «правильные»).

e) Предположив, что (2) – правильная спецификация, поясните, будут ли стандартные ошибки в регрессии (1) являться «правильными» оценками.

8. Исследуется влияние на часовую зарплату (Y) числа лет, потраченных на обучение (S) и стажа работы по специальности (X). Имеются данные для мужчин и женщин ($MALE$ – дамми переменная с 1 для мужчин). Результаты регрессий приведены в таблице (в скобках стандартные ошибки). Зависимая переменная – $\log Y$.

	(1)	(2)	(3)	(4)	(5)
S	0.094	0.099	0.094	0.097	0.094
	(0.003)	(0.004)	(0.005)	(0.003)	(0.005)
X	0.046	0.042	0.039	0.04	0.039
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
$MALE$	-	-	-	0.234	0.117
				(0.016)	(0.108)
$SMALE$					0.005
					(0.007)
$XMALE$					0.003
					(0.004)
$const$	5.165	5.283	5.166	5.111	5.166
	(0.054)	(0.083)	(0.068)	(0.052)	(0.074)
R^2	0.319	0.277	0.363	0.359	0.359
ESS	714.6	411	261.6	672.8	672.5
N	3.242	1.774	1.468	3.242	3.242

Коэффициенты корреляции между *MALE* и *SMALE* – 0,96, *MALE* и *XMALE* – 0,96, *SMALE* и *XMALE* – 0,93.

- a) Проинтерпретируйте коэффициенты при *S* и *SMALE* в регрессии (5)
- b) Проинтерпретируйте коэффициент при *MALE* в регрессии (4) и (5)
- c) Гипотеза, что модели зарплаты для мужчин и женщин разные. Проверьте эту гипотезу при помощи регрессии (4).
- d) Проверьте гипотезу при помощи регрессии (1) и (5)
- e) Объясните разницу в тесте в пунктах c) и d)
- f) На семинаре кто-то предположил, что тест Чоу прольет свет на эту гипотезу. Это верно?
- g) Объясните, какая спецификация (1), (4) или (5) более предпочтительна.