

**Задание для письменного экзамена по дисциплине «Эконометрика и прогнозирование»
зимняя экзаменационная сессия 2009/2010г.**

Экзамен состоит из 16 заданий. Каждое задание оценивается в соответствии с числом баллов, указанном в нем. Максимальная сумма – 100 баллов. При решении задач, пожалуйста, поясняйте и отражайте последовательность вычислений, четко выделяйте (обводите) итоговые результаты.

1. Исследователь по данным 50 стран оценил уравнение регрессии:

$$N = 25.0 + 0.020 \times G \quad R^2 = 0,16 \quad RSS = 4000 \quad F = 4.12$$

(s) (10.0) (0.010)

Здесь N - число газет на душу населения, G - ВВП на душу населения в долларах США.

Объясните (с математическим доказательством), как изменятся следующие компоненты уравнения регрессии, если G будет измерено в фунтах стерлингов (пусть 1 доллар равен 2 фунтам стерлингов):

А) (4 балла) коэффициент при G и свободный член;

Б) (2 балла) остаточная сумма квадратов RSS ;

В) (3 балла) стандартные ошибки коэффициентов;

Г) (2 балла) F - статистика для оценки статистической значимости коэффициента детерминации.

2. По тридцати наблюдениям построена регрессионная модель.

Зависимая переменная: Y			
Переменная	Коэффициент	t-статистика	P-значение
$Const$	3,77	10,75	0,000
X_1	1,25	4,89	0,003
X_2	-0,24	-3,41	0,002
Y_{-1}	5,18	2,96	0,001
R-квадрат	0,8265	Prob F-st	0,0002

DW	2,134	AIC	8,156
Тест White			0,1358
Breusch-Godfrey Тест:			0,0452
Тест Jarque-Bera для остатков			0,6574

А) (4 балла) Запишите и оцените качество построенной модели при $\alpha=0.05$.

Б) (2 балла) Постройте 95% доверительный интервал для коэффициента при переменной X_2 .

В) (2 балла) Оцените, как изменится переменная Y , если X_1 уменьшится на единицу, X_2 увеличится на 2 единицы.

Г) (2 балла) Какой процент вариации переменной Y можно объяснить совокупным влиянием переменных X_{i1} , X_{i2} , Y_{i-1} .

Д) (2 балла) Какой фактор оказывает более значимое влияние на изменение переменной Y . Ответ обоснуйте. Примечание: среднеквадратические (стандартные) отклонения переменных Y , X_1 , X_2 соответственно равны 2.60, 1.30, 8.20

3. (3 балла) Исследователь изучает зависимость заработной платы сотрудников предприятия W от ряда факторов и предполагает построить регрессию вида: $W = \beta_0 + \beta_1 Age + \beta_2 Ed + \beta_3 S + \varepsilon$, где Age - возраст сотрудника, Ed - наличие профильного образования сотрудника, S - стаж работы сотрудника на предприятии. Кроме перечисленных факторов исследователь решает включить в модель общий стаж работы сотрудника FS , который ему неизвестен. Для определения общего стажа он решает использовать соотношение: $FS = Age - 16 - 6$, где 16 определяет количество лет обучения в школе и в вузе. Подробно объясните, какие проблемы могут возникнуть в предполагаемой модели. Предложите решение проблемы.

4. (2 балла) Приведите примеры регрессионной модели, нелинейной по параметрам и нелинейной по объясняющим переменным.

5. Исследователь предполагает, что отношения между двумя переменными Y и X описываются соотношением $Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$, $t = 1, 2, \dots, T$. Для модели выполняются предпосылки МНК: $M\{\varepsilon_i\} = 0$, $D\{\varepsilon_i\} = \sigma^2$, $\text{cov}\{\varepsilon_i, \varepsilon_j\} = 0, (i \neq j)$. По имеющимся выборкам x_1, x_2, \dots, x_T и y_1, y_2, \dots, y_T переменных X и Y соответственно, а также выборке z_1, z_2, \dots, z_T переменной Z , которая не зависит от

переменной Y , исследователь оценивает коэффициент β_1 следующим образом:
$$b_1 = \frac{\sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})}$$
.

А) (6 баллов) Покажите, что оценка b_1 параметра β_1 является несмещенной оценкой.

Б) (6 баллов) Покажите, что оценка b_1 параметра β_1 имеет дисперсию $\sigma_{b_1}^2 = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \times \frac{1}{r_{XZ}^2}$, где r_{XZ} -

коэффициент корреляции между переменными X и Z .

6. Исследователь оценивает зависимость часовой оплаты труда Y от количества лет, потраченных на образование, X_1 и стажа работы X_2 по выборкам из 1774 женщин и 1468 мужчин. Исследователь использует фиктивные переменные: D_1 - пол сотрудника, $D_1 \times X_1$ - произведение переменных X_1 и D_1 , $D_1 \times X_2$ - произведение переменных X_2 и D_1 .

В качестве зависимой переменной используется логарифм часовой оплаты труда. Результаты регрессий представлены в таблице. В скобках записаны стандартные ошибки коэффициентов.

	(1)	(2)	(3)	(4)	(5)
константа	5,165 (0,054)	5,283 (0,083)	5,166 (0,068)	5,111 (0,052)	5,166 (0,074)
X_1	0,094 (0,003)	0,099 (0,004)	0,094 (0,005)	0,097 (0,003)	0,094 (0,005)
X_2	0,046 (0,002)	0,042 (0,003)	0,039 (0,002)	0,040 (0,002)	0,039 (0,003)
D_1	-	-	-	2,34 (0,16)	0,117 (0,108)
$D_1 \times X_1$	-	-	-	-	0,005 (0,007)
$D_1 \times X_2$	-	-	-	-	0,003 (0,004)
R^2	0,319	0,277	0,363	0,359	0,359
RSS	714,6	411,0	261,6	672,8	672,5
n	3242	1774	1468	3242	3242

А) Исследователь предположил, что регрессионные модели логарифма часовой оплаты труда различаются для мужчин и женщин. Проверьте это предположение при $\alpha=0.05$

(1 балл) используя только регрессию (4),

(3 балла) сравнивая регрессии (1) и (5).

Б) (4 балла) Другой исследователь предположил, что для ответа на вопрос пункта А можно использовать тест Чоу. Поясните прав ли исследователь. Если да, то проверьте предположение о различии функции часовой оплаты для мужчин и женщин, используя тест Чоу при $\alpha=0.05$.

7. (2 балла) Для регрессионной модели $Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \varepsilon_t$ выполняются следующие условия: $M\{\varepsilon_t\} = 0$, $D\{\varepsilon_t\} = \sigma^2 X_{t1} \times X_{t2}$. Подробно объясните, какой метод необходимо использовать для оценки регрессионной модели, чтобы получить BLUE -оценки коэффициентов.

8. По данным за 40 лет об экспорте, импорте и ВВП получены две регрессионные модели (1) и (2) соответственно:

$$\begin{aligned} \text{(S)} \quad \text{Exp}_t &= 1,22 + 0,64 \text{Imp}_t + 0,82 \text{GDP}_t & R^2 &= 0,983 & \ln e_t^2 &= 2,35 + 1,05 \ln \text{GDP}_t & R^2 &= 0,67 \\ & (0,64) \quad (0,21) \quad (0,03) & DW &= 2,01 & (t) & (1,21) \quad (4,12) \end{aligned}$$

А) (2 балла) Как можно интерпретировать две приведенные регрессионные модели.

Б) (2 балла) Подробно сформулируйте и объясните, какие предположения сделаны от модели (1) к модели (2). Подтвердились ли эти предположения при $\alpha=0.05$, ответ обоснуйте.

9. (3 балла) Запишите общий вид *LPM* - модели. Какие недостатки и достоинства *LPM* - модели вы знаете?

10. Рассматривается следующая модель:

$$\begin{cases} Y_{1t} = \beta_0 + b_1 Y_{4t} + b_2 Y_{2t} + \varepsilon_{1t} \\ Y_{2t} = \gamma_0 + \gamma_1 Y_{3t} + \gamma_2 Y_{5t} + \varepsilon_{2t} \\ Y_{3t} = a_0 + a_1 Y_{2t} + a_2 Y_{4t} + a_3 Y_{6t} + \varepsilon_{3t} \\ Y_{4t} = Y_{1t} + Y_{2t} + Y_{7t} \end{cases}$$

А) (2 балла) Запишите в общем виде систему в приведенной форме.

Б) (2 балла) Определите, является ли данная система идентифицируемой.

В) (2 балла) Будет ли система идентифицируемой, если в уравнении для Y_{3t} инвестиций исключить Y_{6t} .

Г) (2 балла) Каким методом и почему могут быть оценены параметры исходной системы уравнений.

11. По годовым данным 1981-2005гг. для валового национального продукта *GDP* и агрегата денежной массы M_1 построены следующие модели регрессии:

$$\begin{aligned} \ln M_{1t} &= -10.2571 + 1.597 \ln \text{GDP}_t & R^2 &= 0.9463 & DW &= 1.7399 \\ (t) & (-12.39) \quad (25.88) \end{aligned} \tag{1}$$

$$\begin{aligned} \Delta e_t &= -0.1958 e_{t-1} & R^2 &= 0.1118 & DW &= 1.4767 \\ (t = \tau) & (-2.2521) \end{aligned} \tag{2}$$

здесь e_t представляют собой остатки модели (1).

$$\begin{aligned} \Delta \ln M_{1t} &= 0.084 + 0.734 \Delta \ln \text{GDP}_t - 0.0811 e_{t-1} & R^2 &= 0.1066 & DW &= 1.669 \\ (t) & (0.046) \quad (3.0636) \quad (-2.8537) \end{aligned} \tag{3}$$

А) (6 баллов) Подробно проинтерпретируйте и оцените представленные регрессии (1) - (3). Можно использовать «грубое правило»

Б) (3 балла) Что можно сказать о стационарности временного ряда остатков модели (1). Ответ обоснуйте.

12. Для временного ряда X_t построена регрессионная модель $X_t = 2.54 - 0.90X_{t-1} - 0.20X_{t-2} + \varepsilon_t$.

А) (2 балла) Определите тип модели, с помощью которой описан временной ряд X_t . Как выглядит модель в общем виде.

Б) (3 балла) Запишите и проверьте условие стационарности модели.

13. (3 балла) Используя сведения о временных рядах ежемесячных показателей, приведенных в таблице, предложите не менее трех моделей (включаящих в совокупности все представленные показатели), согласованных с теорией временных рядов и экономической теорией.

<i>CPI</i> , индекс потребительских цен, %	I(0)
<i>GDP</i> , валовый внутренний продукт, млн. руб.	I(1)
<i>IMPORT</i> , импорт товаров и услуг, млн. руб.	I(1)
<i>WAGE</i> , оплата труда, млн. руб.	I(2)
<i>CONSUM</i> , денежные расходы населения, млн. руб.	I(1)
<i>INCOME</i> , денежные доходы населения, млн. руб.	I(1)
<i>EX RATE</i> , реальный курс национальный валюты, руб.	I(0)

14. (4 балла) Для изучения зависимости объемов материальных запасов Y от объемов продаж X промышленного сектора США исследователь предполагает построить регрессионную модель вида:

$$Y_i = \alpha + \beta_0 X_i + \beta_1 X_{i-1} + \beta_2 X_{i-2} + \beta_3 X_{i-3} + \dots + \varepsilon_i \quad (1)$$

При этом предполагается, что коэффициенты β_i имеют геометрическую структуру лагов. Используя преобразование Койка, получено уравнение вида $Y_i = (1 - \lambda) \times \alpha + \beta_0 X_i + \lambda Y_{i-1} + v_i$.

По данным 1955-1974гг. оценено следующее уравнение регрессии:

$$Y_i = 7.5 + 6.64 X_i + 0.25 Y_{i-1} \quad (2)$$

$$(S) (1.9) (0.16) (0.04) \quad R^2 = 0.9961$$

По оценкам коэффициентов уравнения (2) определите оценки коэффициентов уравнения (1) и запишите полученное уравнение.

15. (4 балла) Покажите, что модель частичной корректировки является авторегрессионной моделью.

16. (10 баллов) Кратко поясните почему, следующие утверждения являются истинными или ложными.

1) Гетероскедастичность появляется в регрессионной модели, если случайное отклонение коррелировано с одной из объясняющих переменных.

2) Расширенный тест Дики-Фуллера используется для определения коинтегрированности временных рядов.

3) Обычный МНК при автокорреляции случайных отклонений дает несмещенные, но несостоятельные оценки.

4) Применение обычного МНК для структурной системы одновременных уравнений позволяет получить BLUE-оценки.

5) При построении модели ARIMA обязательным является выделение детерминированного тренда.

6) Суть метода наименьших квадратов состоит в минимизации суммы квадратов коэффициентов уравнения регрессии.

7) Использование полиномиальной структуры лагов Алмон позволяет свести модель с распределенными лагами к авторегрессионной модели.

8) Для устранения мультиколлинеарности используют авторегрессионное преобразование.

9) Построение регрессионных моделей по стационарным временным рядам может привести к установлению ложных регрессионных зависимостей.

10) При использовании двухшагового МНК для оценивания системы одновременных уравнений возникает проблема идентифицируемости.

Для справки:

$$t_{\frac{0.05}{2}, 3237} = 1.96 \quad t_{\frac{0.05}{2}, 26} = 2.06 \quad t_{\frac{0.05}{2}, 38} = 2.02 \quad F_{0.05, 3, 3236} = 2.607 \quad F_{0.05, 3, 26} = 2.98$$

$$\text{для } \alpha=0.05 \quad n = 30, \quad m = 3: \quad D_L = 1.214 \quad D_U = 1.650$$

$$t = \tau(N) = -1.95 \quad t = \tau(C) = -3.01 \quad t = \tau(CT) = -3.62$$