

ВОПРОСЫ К ЭКЗАМЕНУ ПО ЭКОНОМЕТРИКЕ

1. Основные понятия эконометрики. Этапы решения эконометрических задач.
2. Парная линейная регрессия. Метод наименьших квадратов (МНК).
3. Парная нелинейная регрессия.
4. Множественная регрессия. Выбор формы множественной регрессии.
5. Оценка коэффициентов множественной регрессии.
6. Отбор факторов при построении множественной регрессии.
7. Уравнение множественной регрессии в стандартизованном масштабе.
8. Оценка значимости уравнения регрессии в целом с помощью F – критерия Фишера.
9. Оценка значимости отдельных коэффициентов уравнения регрессии помощью t–статистики Стьюдента.
10. Доверительные интервалы прогноза для линейного уравнения регрессии.
11. Фиктивные переменные в регрессионных моделях.
12. Система одновременных уравнений. Основные понятия.
13. Косвенный метод наименьших квадратов.
14. Двухшаговый МНК.
15. Динамические модели. Временные ряды. Лаги (запаздывания) в динамических моделях.
16. Полиномиально – распределенные лаги Ш. Алмон.
17. Теорема Гаусса – Маркова.
18. Тесты на гетероскедастичность. Тест ранговой корреляции Спирмена.
19. Тест на гетероскедастичность Голдфелда – Кванта.
20. Тест на гетероскедастичность Уайта.
21. Устранение гетероскедастичности. Взвешенный МНК.
22. Модель бинарного выбора.
23. Оценивание коэффициента β и вероятностей в бинарной модели методом максимального правдоподобия (ММП).
24. Модель Койка.
25. Эконометрика финансовых рынков. Портфельные задачи.

Заключительный экзамен

Часть 1. (30 минут). В каждом из 12 вопросов обведите кружком ровно один ответ, который вы считаете правильным.

1. Укажите среди названных здесь статистических тестов те тесты, которые могут выполняться только для больших выборок.

- I. Тест Чоу на неудачу предсказания;
- II. Тест Спирмена на ранговую корреляцию;
- III. Тест χ^2 для LR-статистики при оценивании по методу максимума правдоподобия;
- IV. Тест Дарбина для уравнения с лаговой объясняющей переменной;
- V. Тест Голдфелда-Квандта.

- 1) Только II, III, IV.
- 2) Только II, IV.
- 3) Только I, III.
- 4) Только IV, V.
- 5) Только I, II, IV.

2. Гипотеза о том, что случайный член не подвержен автокорреляции первого порядка, отвергается в пользу односторонней альтернативной гипотезы о его отрицательной автокоррелированности, если:

- 1) статистика Дарбина-Уотсона d близка к 2;
- 2) статистика Дарбина-Уотсона d близка к 1;
- 3) статистика Дарбина-Уотсона d близка к 0 или 4;
- 4) статистика Дарбина-Уотсона d близка к 4, но не к 0;
- 5) статистика Дарбина-Уотсона d близка к 0, но не к 4.

3. Переменная y_t зависит от переменной x_t : $y_t = \alpha + \beta x_t + u_t$. Оценивается зависимость y_t от переменной z_t : $y_t = a + bz_t + e_t$, где $z_t = x_t + v_t$ (v_t – ошибка измерения, являющаяся случайной величиной с нулевым математическим ожиданием и постоянной дисперсией). Если истинное значение β равно 0,4, а теоретическая дисперсия x_t в три раза больше теоретической дисперсии v_t , то предел по вероятности оценки b равен:

- 1) 0,1; 2) 0,2; 3) 0,267; 4) 0,3; 5) 0,333.

4. После преобразования Койка модель геометрически распределенного лага $y_t = \alpha + \beta x_t + \beta \delta x_{t-1} + \beta \delta^2 x_{t-2} + \beta \delta^3 x_{t-3} + \dots + u_t$ приобретает вид:

- 1) $y_t - \delta y_{t-1} = \alpha (1 - \delta) + \beta x_t + u_t - \delta u_{t-1}$,
- 2) $y_t = \alpha (1 - \delta) + \beta x_t + u_t - \delta u_{t-1}$,
- 3) $y_t = \alpha + \beta x_t + u_t - \delta u_{t-1}$,

- 4) $y_t - \delta y_{t-1} = \alpha + \beta x_{t-1} + u_{t-1}$,
 5) ни один из перечисленных.

5. По 39 точкам оценена следующая формула производственной функции, в которой отдельно рассмотрены две составляющие затрат основного капитала: K_1 - здания и сооружения, и K_2 - машины и оборудование; а также две составляющие затрат труда: L_1 - затраты квалифицированного труда, и L_2 - затраты неквалифицированного труда; Y – выпуск:

$$\ln(Y) = -4,3 + 0,35 \ln(K_1) + 0,26 \ln(K_2) + 0,63 \ln(L_1) + 0,58 \ln(L_2); R^2 = 0,92; DW = 1,74$$

(1,4) (0,03) (0,05) (0,41) (0,38)

(в скобках приведены стандартные ошибки коэффициентов).

Какой из выводов и дальнейших шагов представляется Вам верным?

- 1) Нужно исключить фактор L (переменные L_1 и L_2), т.к. он оказался незначимым;
- 2) Имеет место мультиколлинеарность, поэтому нужно объединить факторы K_1 и K_2 ;
- 3) Имеет место мультиколлинеарность, поэтому нужно объединить факторы L_1 и L_2 ;
- 4) Отклонения e_i автокоррелированы, нужно изменить формулу зависимости;
- 5) Формула зависимости приемлема по всем приведенным параметрам, и изменения не нужны.

6. Эконометрическая модель описана следующими уравнениями:

$$(1) \quad y_1 = \alpha + \beta y_2 + \gamma x_1 + \varphi x_2 + u_1$$

$$(2) \quad y_2 = \delta + \tau x_1 + \lambda x_2 + u_2$$

где y_1 и y_2 – эндогенные переменные, x_1 и x_2 – экзогенные переменные, и u_1 и u_2 – случайные члены, удовлетворяющие условиям Гаусса-Маркова. Выберите неверное утверждение:

- 1) при непосредственном оценивании по МНК оценки коэффициентов уравнения (1) будут несмещенными;
- 2) при непосредственном оценивании по МНК оценки коэффициентов уравнения (2) будут несмещенными;
- 3) при непосредственном оценивании по МНК оценки коэффициентов уравнения (1) будут состоятельными;

- 4) при непосредственном оценивании по МНК оценки коэффициентов уравнения (2) будут состоятельными;
- 5) косвенный метод наименьших квадратов позволяет получить несмещенные оценки коэффициентов обоих уравнений.

7. Дана следующая макроэкономическая модель:

$$Y = C + I + G \quad \text{- макроэкономическое тождество}$$

$$C = a + bY \quad \text{- функция потребления}$$

$$I = d - e \cdot R_{-1} + fY \quad \text{- функция инвестиций}$$

$$(M/P) = gY - hR \quad \text{- уравнение денежного рынка}$$

где эндогенными переменными являются доход Y , потребление C , инвестиции I и процентная ставка R . Переменные G (государственные расходы) и (M/P) (реальная денежная масса) – экзогенные. Выберите верное утверждение из следующих.

- 1) функция инвестиций неидентифицируема;
- 2) функция потребления неидентифицируема;
- 3) функция потребления однозначно идентифицируема;
- 4) функция инвестиций однозначно идентифицируема;
- 5) функция инвестиций сверхидентифицируема.

8. Коэффициент Тейла в виде $U = \frac{\sqrt{\frac{1}{m} \sum (\hat{y}_{T+p} - y_{T+p})^2}}{\sqrt{\frac{1}{m} \sum \hat{y}_{T+p}^2} + \sqrt{\frac{1}{m} \sum y_{T+p}^2}}$ позволяет сравнить

качество сделанного на период $(T+1, \dots, T+m)$ прогноза \hat{y}_{T+p} с качеством следующего прогноза:

- 1) $\hat{y}_{T+p} = 0$ для всех $p=1, \dots, m$.
- 2) $\hat{y}_{T+p} = y_T$ для всех $p=1, \dots, m$.
- 3) прогнозом по линейной функции, оцененной за период $(1, \dots, T)$.
- 4) прогнозом по экспоненте, оцененной за период $(1, \dots, T)$.
- 5) прогнозом по логарифмической функции, оцененной за период $(1, \dots, T)$.

9. Введение фиктивных переменных D_{T+p} , $p=1, \dots, m$ при прогнозировании ($D_{T+p}=1$ при $t=T+p$, $D_{T+p}=0$ при всех других t) позволяет:

- I. Рассчитать ошибки предсказания для $t=T+1, \dots, T+m$.
- II. Рассчитать стандартные ошибки предсказания для $t=T+1, \dots, T+m$.
- III. Выполнить F -тест на стабильность коэффициентов.

- 1) Верно только I.
- 2) Верно только II.
- 3) Верно только I и II.
- 4) Верно только I и III.
- 5) Верно I, II и III.

10. При оценивании линейной регрессии по методу максимума правдоподобия оценки коэффициентов совпадают с полученными по МНК при следующих условиях:

- 1) всегда;
- 2) если случайный член нормально распределен и выполнены условия Гаусса-Маркова;
- 3) если случайный член нормально распределен, независимо от выполнения условий Гаусса-Маркова;
- 4) если случайный член не автокоррелирован, при любом его распределении;
- 5) вообще говоря, эти оценки не совпадают, независимо от выполнения условий Гаусса-Маркова и распределения случайного члена.

11. Логит-оценивание модели $F(Z_i)=p(Degree_i=1 | ASVABC_i, \beta)$ дало результат $Z=-11.96+0.169 \cdot ASVABC$. Вероятность получения академической степени ($Degree=1$) при $ASVABC=60$ равна:

- 1) $\frac{\exp(-11.96 + 0.169 \cdot 60)}{\exp(-11.96 + 0.169 \cdot 60) + 1}$;
- 2) $\frac{\exp(11.96 - 0.169 \cdot 60)}{(1 + \exp(-11.96 + 0.169 \cdot 60))^2} \cdot 0.169$;
- 3) $(\frac{\exp(-11.96 + 0.169 \cdot 60)}{\exp(-11.96 + 0.169 \cdot 60) + 1}) \cdot 0.169$;
- 4) $\frac{\exp(11.96 - 0.169 \cdot 60)}{(1 + \exp(-11.96 + 0.169 \cdot 60))^2}$;
- 5) $\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(11.96 - 0.169 \cdot 60)^2)$.

12. Пробит-оценивание модели $F(Z_i)=p(Degree_i=1 | ASVABC_i, \beta)$ дало результат $Z=-5.787+0.0782 \cdot ASVABC$. Увеличение вероятности $p(Degree_i=1)$ при увеличении $ASVABC$ на один пункт при $ASVABC=60$ равно здесь:

- 1) $\frac{\exp(-11.96 + 0.169 \cdot 60)}{\exp(-11.96 + 0.169 \cdot 60) + 1}$;
- 2) $\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(5.787 - 0.0782 \cdot 60)^2)$;
- 3) $\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(5.787 - 0.0782 \cdot 60)^2) \cdot 0.0782$;
- 4) $F(-5.787+0.0782 \cdot 60)$, где F – функция стандартного нормального распределения;
- 5) $F(-5.787+0.0782 \cdot 60) \cdot 0.0782$, где F – функция стандартного нормального распределения;

Часть 2. (1 час 30 минут). Ответьте письменно на два из трех следующих вопросов.

1. Исследователь располагает данными о продолжительности обучения респондента HGC , результатах теста на познавательные способности $ASVABC$, и о продолжительности обучения матери респондента $HGCM$, для выборки 570 индивидов. Он предполагает, что продолжительность обучения респондента зависит как от результатов теста на познавательные способности, так и от продолжительности обучения матери респондента

$$HGC = \alpha + \beta_1 ASVABC + \beta_2 HGCM + u \quad (1.1)$$

и что познавательные способности, в свою очередь, являются функцией успехов в учебе:

$$ASVBC = \gamma + \delta_1 HGC + v \quad (1.2)$$

1a. [2 балла] Объясните разницу между структурными уравнениями и уравнениями в приведенной форме в модели одновременных уравнений.

1b. [2 балла] Выведите приведенную форму уравнения для HGC .

1c. [6 баллов] Выведите выражение для смещения (для больших выборок) оценки δ_1 , предполагая, что уравнение (1.2) было оценено обычным методом наименьших квадратов, и определите вероятное направление смещения.

1d. [3 балла] Объясните, почему $HGCM$ может быть использована в качестве инструмента для HGC во втором уравнении.

1e. [6 баллов] Продемонстрируйте, что оценка метода инструментальных переменных для δ_1 , использующая в качестве инструмента $HGCM$, является состоятельной, предполагая, что модель корректно определена.

1f. [6 баллов] Предположим теперь, что корректной спецификацией второго уравнения является

$$ASVBC = \gamma + \delta_1 HGC + \delta_2 HGCM + v \quad (1.3)$$

Продемонстрируйте, что оценка метода инструментальных переменных для δ_1 будет несостоятельной, выведите смещение для случая большой выборки и установите его направление. (Указание: учтите, что приведенная форма уравнения для HGC в этом случае изменится).

2а. [8 баллов] Случайная переменная принимает два возможных значения 0 и 1 с неизвестной вероятностью p того, что она равна 1. В выборке имеется всего три наблюдения: первое наблюдение равно 1, второе также 1, а третье равно 0. Выведите оценку максимального правдоподобия для p .

Следующая информация относится к остальным вопросам этого задания.

Переменная *WORKING* определена равной 1, если респондент имеет работу, и равной 0 в противном случае. Для оценивания зависимости переменной *WORKING* от уровня образования респондента *HGC* (продолжительность обучения респондента) был использован метод логит-анализа.

Исследование было выполнено отдельно для мужчин и женщин (в приведенных ниже результатах расчетов первая таблица соответствует женщинам, а вторая – мужчинам, указание на число проведенных итераций не приводится).

logit working hgc if male=0
Logit Estimates

Number of obs = 2726
chi2(l) = 70.42
Prob > chi2 = 0.0000

Log Likelihood = -1586.5519

Pseudo R2 = 0.0217

working	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]	
hgc 1	.1511872	.0186177	8.121	0.000	.1146971	.1876773
cons 1	-1.049543	.2448064	-4.287	0.000	-1.529355	-.5697314

logit working hgc if male=1
Logit Estimates

Number of obs = 2573
chi2(l) = 75.03
Prob > chi2 = 0.0000
Pseudo R2 = 0.0446

Log Likelihood = -802.65424

working	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]	
hgc 1	.2499295	.0306482	8.155	0.000	.1898601	.3099989
cons 1	-.9670268	-.3775658	-2.561	0.010	-1.707042	-2270113

95 процентов респондентов имеют значение *HGC* в диапазоне 9-18 лет со средним значением *HGC* равным 13,3 и 13,2 лет для женщин и мужчин соответственно.

Исходя из логит-анализа, предельный эффект *HGC* на вероятность наличия работы в среднем оценивается как 0,030 и 0,020 для женщин и мужчин соответственно.

Оценка регрессии метода наименьших квадратов *WORKING* по *HGC* дает коэффициент наклона 0,029 и 0,020 для женщин и мужчин соответственно.

Логистическая функция $F(z)$ и ее производная $f(z)$ даются выражениями

$$F(z) = \frac{1}{1 + e^{-z}} \quad \text{и} \quad f(z) = \frac{e^{-z}}{(1 + e^{-z})^2},$$

где z является линейной функцией характеристик респондента.

На графиках (см. следующую страницу) представлены вероятность наличия работы и предельное влияние HGC на вероятность наличия работы, для мужчин и женщин.

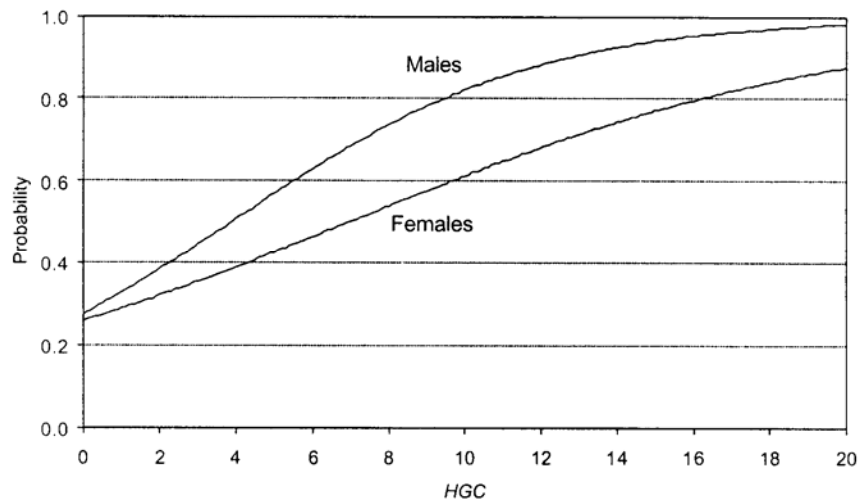


Рис. 1. Вероятность наличия работы как функция HGC .

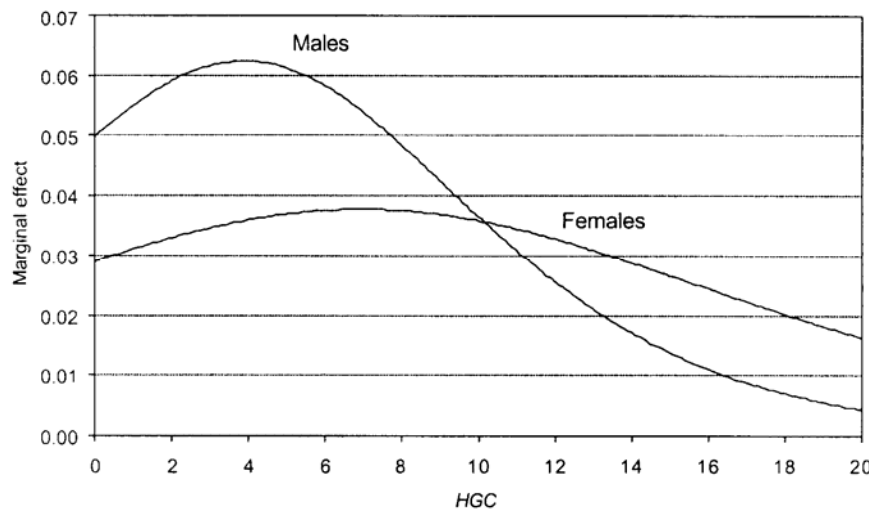


Рис. 2. Предельное влияние HGC на вероятность наличия работы.

2b. [4 балла] Объясните, почему логит-модель может рассматриваться как подходящая математическая спецификация для такого рода моделей.

2c. [3 балла] Объясните, как оценивается предельное влияние продолжительности обучения на вероятность наличия работы на основе логит-регрессии.

2d. [3 балла] Как видно из Рис. 2, предельное влияние продолжительности обучения ниже для мужчин для значительной части диапазона ($HGC > 9$). Обсудите приемлемость такого результата (указание: никаких математических выкладок в этой части не требуется и баллы за них не добавляются).

2e. [3 балла] Как можно видеть из рис. 2, предельное влияние продолжительности обучения убывает с ростом HGC как для мужчин, так и для женщин в диапазоне $HGC > 9$. Обсудите приемлемость такого результата (указание: никаких математических выкладок в этой части не требуется и баллы за них не добавляются).

2f. [4 балла] Сравните оценки предельного влияния продолжительности обучения, полученные с помощью логит-анализа, с оценками метода наименьших квадратов.

3. Исследователь изучает прогностические свойства моделей регрессии по данным о затратах на образование в зависимости от времени (переменная $TIME$, равная 1 для 1959 года), личном располагаемом доходе DPI и индексе цен на образовательные услуги.

3a. [2 балла] Объясните разницу между предсказанием и прогнозом применительно к рассматриваемым данным.

3b. [6 баллов] Исследователь строит регрессию расходов на образование по личному располагаемому доходу и индексу цен (расчеты выполнены в программе EViews 3.1)

Dependent Variable: EDUC

Method: Least Squares

Sample: 1959 1990

Included observations: 32

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-14.24189	2.512958	-5.667378	0.0000
DPI	0.027975	0.001521	18.39055	0.0000
PEDUC	-0.287045	0.055257	-5.194752	0.0000
R-squared	0.987355	Mean dependent var	57.79063	
Adjusted R-squared	0.986483	S.D. dependent var	18.18783	
S.E. of regression	2.114590	Akaike info criterion	4.424659	
Sum squared resid	129.6733	Schwarz criterion	4.562072	

Log likelihood	-67.79454	F-statistic	1132.176
Durbin-Watson stat	0.462621	Prob(F-statistic)	0.000000

Предсказание значений *EDUC* (переменная *EDUCF*), выполненное по построенной модели для четырех последующих лет, дало следующие результаты

	EDUCF	EDUC
1991	83.91171	90.30000
1992	86.12100	93.10000
1993	86.32668	95.10000
1994	88.15283	97.00000

Результаты оценивания той же регрессии для 1959-1994 годов по методу Салкевера с фиктивными переменными *d91*, *d92*, *d93*, *d94*, определенными равными 1 только для соответствующих лет, выглядят следующим образом.

Dependent Variable: EDUC
Method: Least Squares
Sample: 1959 1994
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-14.24189	2.512958	-5.667378	0.0000
DPI	0.027975	0.001521	18.39055	0.0000
PEDUC	-0.287045	0.055257	-5.194752	0.0000
D91	6.388293	2.396289	2.665910	0.0124
D92	6.978997	2.438620	2.861863	0.0077
D93	8.773318	2.512594	3.491737	0.0016
D94	8.847170	2.561074	3.454477	0.0017
R-squared	0.991302	Mean dependent var	61.80000	
Adjusted R-squared	0.989503	S.D. dependent var	20.63906	
S.E. of regression	2.114590	Akaike info criterion	4.508265	
Sum squared resid	129.6733	Schwarz criterion	4.816171	
Log likelihood	-74.14877	F-statistic	550.8716	
Durbin-Watson stat	0.615401	Prob(F-statistic)	0.000000	

Dependent Variable: EDUC

Постройте доверительные интервалы (выбрав уровень значимости самостоятельно) для предсказания значения *EDUC* для 1991-1994 годов. На этой основе обсудите свойства полученного предсказания.

Зс. [4 балла] Обсудите сравнительные преимущества и недостатки различных подходов к оценке качества прогноза, включая коэффициенты Тейла.

$$U_1 = \sqrt{\frac{\frac{1}{m} \sum (\Delta \hat{y}_{T+p} - \Delta y_{T+p})^2}{\frac{1}{m} \sum (\Delta y_{T+p})^2}} \quad \text{и} \quad U_2 = \frac{\sqrt{\frac{1}{m} \sum (\hat{y}_{T+p} - y_{T+p})^2}}{\sqrt{\frac{1}{m} \sum \hat{y}_{T+p}^2 + \frac{1}{m} \sum y_{T+p}^2}}$$

3d. [5 баллов] Далее исследователь строит регрессию расходов на образование по времени

Dependent Variable: EDUC

Method: Least Squares

Date: 12/30/02 Time: 13:28

Sample: 1959 1990

Included observations: 32

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	26.23528	1.100393	23.84175	0.0000
TIME	1.912445	0.058198	32.86089	0.0000
R-squared	0.972969	Mean dependent var	57.79063	
Adjusted R-squared	0.972068	S.D. dependent var	18.18783	
S.E. of regression	3.039708	Akaike info criterion	5.121861	
Sum squared resid	277.1947	Schwarz criterion	5.213470	
Log likelihood	-79.94978	F-statistic	1079.838	
Durbin-Watson stat	0.150184	Prob(F-statistic)	0.000000	

Соответствующие прогнозные значения представлены в таблице

	EDUCF	EDUC
1991	89.34597	90.30000
1992	91.25841	93.10000
1993	93.17086	95.10000
1994	95.08330	97.00000

Опишите процедуру построения доверительных интервалов для предсказания расходов на образование.

3e. [4 баллов] Объясните сущность теста Чоу на неудачу предсказания. Выполните тест Чоу для множественной регрессии *EDUC* по переменным *DPI* и *PEDUC*, учитывая, что значение суммы квадратов остатков соответствующей регрессии, выполненной для всего периода 1959-1994, составляет 228,3452.

3f. [4 балла] Объясните сущность F-теста на стабильность коэффициентов. Какая информация необходима для его выполнения в случае регрессии, рассмотренной в пункте **3b** (расчеты проводить не нужно)?