

Министерство образования и науки Украины
Севастопольский национальный технический университет



**«ПОСТРОЕНИЕ И АНАЛИЗ РЕГРЕССИОННЫХ МОДЕЛЕЙ С
ГЕТЕРОСКЕДАСТИЧНОЙ СЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ»**

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к проведению практического занятия №2

по дисциплине «Эконометрия»

для студентов специальности
6.050201 - «Менеджмент организаций»
дневной формы обучения

Севастополь
2008



УДК 658

«Анализ регрессионных моделей с гетероскедастичной случайной составляющей»
методические указания к проведению практического занятия №2 по дисциплине «Эконометрия» для студентов специальности 8.050201 – «Менеджмент организаций» дневной формы обучения / Сост. А.В. Цуканов, Т.А. Кокодей. – Севастополь: Изд-во СевНТУ, 2008г. – 18 с.

Целью методических указаний является закрепление теоретических знаний и получение практических навыков анализа регрессионных моделей с гетероскедастичной случайной составляющей. Методические указания предназначены для студентов специальности «Менеджмент организаций» дневной формы обучения.

Методические указания утверждены на заседании кафедры менеджмента и экономико-математических методов, (протокол № 5 от 23.01.2008 г.).

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент:

Фисун С.Н., канд. техн. наук, доцент кафедры «Кибернетика и вычислительная техника».

СОДЕРЖАНИЕ

1. Цель работы.....	4
2. Теоретические основы и примеры расчётов.....	4
2.1. Гомоскедастичность и гетероскедастичность.....	4
2.2. Обнаружение гетероскедастичности.....	5
2.3. Использование взвешенного метода наименьших квадратов (ВМНК) для оценки моделей с гетероскедастичностью.....	10
3. Индивидуальные расчётно-практические задания.....	13
4. Содержание отчета о практическом занятии.....	17
Библиографический список	18

1. ЦЕЛЬ РАБОТЫ

Целью данной работы является получение навыков построения и анализа регрессионных моделей с гетероскедастичностью, что является составляющей процесса подготовки и принятия решений менеджерами компаний в условиях рыночной конкуренции.

2. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ПРИМЕРЫ РАСЧЁТОВ

2.1. Гомоскедастичность и гетероскедастичность

Регрессионный анализ функции (1), основанный на обычном или одношаговом методе наименьших квадратов (1МНК) должен удовлетворять четырем условиям Гаусса—Маркова:

1. Математическое ожидание случайной составляющей, $M(u_i)$ в любом наблюдении должно быть равно нулю.
2. Дисперсия случайной составляющей $\sigma_{u_i}^2$ должна быть постоянна для всех наблюдений.
3. Отсутствие систематической связи (автокорреляции) между значениями случайной составляющей u_i в любых двух наблюдениях.
4. Случайный характер остатков. Случайная составляющая должна быть распределена независимо от переменных y_i и x_i .

$$y = \alpha_0 + \alpha_1 \cdot x_1 + \dots + \alpha_n x_n + u, \quad (1)$$

где $y, x_1, \dots, x_n$ - зависимая и независимые переменные;

α_i – параметр регрессионной модели

u – аддитивная случайная составляющая (остаток).

Условия 1 и 2 известны как гомоскедастичность, что означает "одинаковый разброс" $\sigma_{u_i}^2 = const$, и представлены на рисунке 1 (график №1). Другими словами, вероятность того, что величина u_i примет какое-то данное положительное (или отрицательное) значение, будет одинаковой для всех наблюдений.

Однако, в экономических и социальных исследованиях (во временных и пространственно-временных выборках) условия Гаусса-Маркова часто не выполняются. Если в данных выборках нарушается условие 2, т.е. для различных наблюдений величина u_i распределена по-разному: вероятность того, что в i -ом наблюдении величина примет какое-то конкретное значение отличается от вероятности принятия этого же значения в j -ом наблюдении. Данное свойство называется гетероскедастичностью, что означает "неодинаковый разброс" $\sigma_{u_i}^2 \neq const$, и представлено на рисунке 1 (графики №2 и №3).

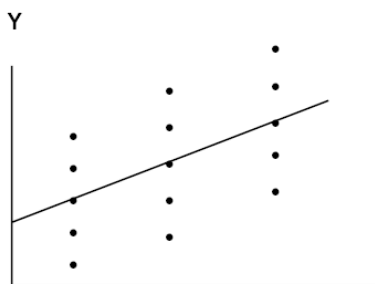
Следовательно, метод 1МНК в данном случае не даёт наилучших из всех возможных результатов: $\alpha_1 \dots \alpha_n$ не будут лучшими оценками коэффициентов,

неправильно оцененная дисперсионно-ковариационная матрица отрицательно скажется на t- и F- тестах, что приведёт к принятию ошибочных гипотез.

Далее будем рассматривать модель линейной регрессии (1), в которой дисперсии случайных составляющих (остатков) u_i не являются постоянными от наблюдения к наблюдению (гетероскедастичны).

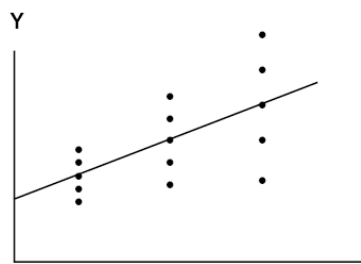
№1 Гомоскедастичность

$$\sigma^2_{u_i} = const$$



№2 Растущая гетероскедастичность

$$\sigma^2_{u_i} \neq const$$



№3 U-образная гетероскедастичность

$$\sigma^2_{u_i} \neq const$$

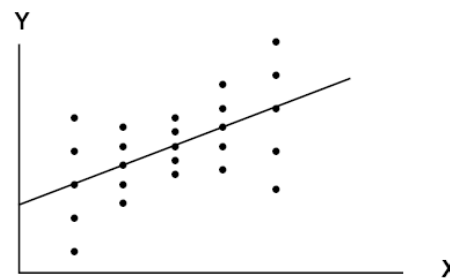


Рисунок 1 – Гомоскедастичные и гетероскедастичные случайные составляющие, u_i

2.2. Обнаружение гетероскедастичности

Для формальной проверки существует большое число статистических процедур (тестов), позволяющих обнаружить гетероскедастичность (например, тест Голдфелда-Квандта, тест Бреуша-Погана, тест Уайта, тест Глейзера и т.д.). В каждом тесте пытаются опровергнуть гипотезу о гомоскедастичности, если это удаётся, то можно сделать вывод, что в модели наблюдается гетероскедастичность.

Рассмотрим наиболее известный тест Голдфелда-Квандта (Goldfeld-Quandt, 1965) для обнаружения гетероскедастичности. При проведении проверки по предложенному ими критерию предполагается, что регрессия однофакторная $y = a_0 + a_1 \cdot x + u$ и стандартное отклонение пропорционально значению x . Метод состоит в оценке двух линий регрессии одношаговым методом наименьших квадратов (МНК). Первая линия строится на основе данных с наименьшими значениями регрессионных остатков (и дисперсии ошибки), вторая – на основе данных с наибольшими значениями остатков (и дисперсии ошибки). Если значения остатков (и дисперсий) обеих регрессий примерно одинаковы, то принимается гипотеза о гомоскедастичности. В противном случае можно считать, что в модели присутствует гетероскедастичность.

Тест состоит из следующих этапов:

1. Формулируются гипотезы H_0 = гомоскедастичность, H_1 = гетероскедастичность.

2. Располагаются данные в порядке возрастания величины x , пропорционально которой изменяется стандартное отклонение случайной составляющей.
3. Исключаются средние d наблюдений. d может быть выбрано, например, как $1/5$ всех наблюдений.
4. Оцениваются две регрессии по формуле (2). Первая из них использует наименьшие значения переменной x , вторая – наибольшие значения этой переменной. Каждая из регрессий построена на $(n-d)/2$ наблюдениях с $[(n-d)/2]-k$ степенями свободы (k равно числу коэффициентов, для однофакторной регрессии $k=2$). Величина d должна быть такой, чтобы гарантировать достаточность степеней свободы для правильной оценки каждой из регрессий.

$$\hat{\alpha} = [X^T \cdot X]^{-1} \cdot X^T \cdot Y, \quad (2)$$

где X - матрица независимых переменных; Y - вектор-столбец зависимой переменной.

5. Вычислить сумму квадратов остатков (ошибок) для каждой из регрессий: RSS_1 для малых значений x , и RSS_2 для больших x , формула (3).

$$RSS = Y^T \cdot Y - \hat{\alpha}^T \cdot X^T \cdot Y. \quad (3)$$

6. В предположении, что ошибки имеют нормальное распределение, статистика RSS_2/RSS_1 будет иметь F-распределение с $(n-d-4)/2$ степенями свободы. Опровергается гипотеза о гомоскедастичности в выбранном уровне значимости, если вычисленная статистика превышает соответствующее критическое значение F-распределения.

Тест Голдфелда-Квандта может применяться не только в случае парной линейной регрессии. В случае множественной модели число степеней свободы F-распределения будет $(n-d-2k)/2$.

Пример 1. Проверим однофакторную регрессионную модель $y = -70,4 + 2,52x$, построенную по данным таблицы 1 методом МНК, на гетероскедастичность (тест Голдфелда-Квандта, $\alpha = 1\%$ и 5%), исключив средние $d=N/5$ наблюдений, ($N=20$).

Таблица 1 – Исходные и сортированные данные X и Y

№	X	Y	X- сорт.	Y- сорт.
1	100	500	5	1
2	15	3	10	2
3	10	2	15	3
4	20	4	20	4
5	25	6	25	6
6	5	1	30	8
7	30	8	35	10
8	35	10	40	11
9	75	110	45	9
10	45	9	50	100
11	50	100	55	15
12	95	200	60	2
13	65	7	65	7
14	40	11	70	15
15	70	15	75	110
16	90	115	80	35
17	60	2	85	90
18	55	15	90	115
19	85	90	95	200
20	80	35	100	500

1. Сформулируем гипотезы H_0 = гомоскедастичность, H_1 = гетероскедастичность.
2. Отсортируем данные X и Y по возрастанию X и поместим результаты в таблицу 1 как X-сорт и Y-сорт.
3. Определим $d=4$ и исключим соответствующие данные из середины отсортированной выборки, получим две новые выборки.
4. Оценим параметры и RSS двух соответствующих регрессий по формулам (2 и 3).

Для первой регрессии рассчитаем коэффициенты

$$[X^T \cdot X]^{-1} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 \\ 1 & 10 \\ 1 & 15 \\ 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \\ 1 & 40 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 8 & 180 \\ 180 & 5100 \end{bmatrix}^{-1} = \begin{bmatrix} 0,607143 & -0,021429 \\ -0,021429 & 0,000952 \end{bmatrix}.$$

$$X^T \cdot Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 6 \\ 8 \\ 10 \\ 11 \end{bmatrix} = \begin{bmatrix} 45 \\ 1330 \end{bmatrix}.$$

$$\mathcal{E} = \begin{bmatrix} 0,607143 & -0,021429 \\ -0,021429 & 0,000952 \end{bmatrix} \cdot \begin{bmatrix} 45 \\ 1330 \end{bmatrix} = \begin{bmatrix} -1,1786 \\ 0,3024 \end{bmatrix}.$$

Рассчитаем RSS_1

$$Y^T \cdot Y = 351.$$

$$\mathcal{E}^T \cdot X^T \cdot Y = [-1,1786 \quad 0,3024] \cdot \begin{bmatrix} 45 \\ 1330 \end{bmatrix} = 349,131.$$

$$RSS_1 = 351 - 349,131 = 1,869.$$

Для второй регрессии рассчитаем коэффициенты

$$[X^T \cdot X]^{-1} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 65 & 70 & 75 & 80 & 85 & 90 & 95 & 100 \end{bmatrix} \cdot \begin{bmatrix} 1 & 65 \\ 1 & 70 \\ 1 & 75 \\ 1 & 80 \\ 1 & 85 \\ 1 & 90 \\ 1 & 95 \\ 1 & 100 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 8 & 660 \\ 660 & 55500 \end{bmatrix}^{-1} = \begin{bmatrix} 6,607143 & -0,078571 \\ -0,078571 & 0,000952 \end{bmatrix}.$$

$$X^T \cdot Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 65 & 70 & 75 & 80 & 85 & 90 & 95 & 100 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 15 \\ 110 \\ 35 \\ 90 \\ 115 \\ 200 \\ 500 \end{bmatrix} = \begin{bmatrix} 1072 \\ 99555 \end{bmatrix}.$$

$$\mathcal{E} = \begin{bmatrix} 6,607143 & -0,078571 \\ -0,078571 & 0,000952 \end{bmatrix} \cdot \begin{bmatrix} 1072 \\ 99555 \end{bmatrix} = \begin{bmatrix} -739,3214 \\ 10,5857 \end{bmatrix}.$$

Рассчитаем RSS_2

$$Y^T \cdot Y = 324924.$$

$$\mathcal{E}^T \cdot X^T \cdot Y = \begin{bmatrix} -739,3214 & 10,5857 \end{bmatrix} \cdot \begin{bmatrix} 1072 \\ 99555 \end{bmatrix} = 261308,2.$$

$$RSS_2 = 324924 - 261308,2 = 63615,79.$$

5. Рассчитаем $\frac{RSS_2}{RSS_1} = \frac{63615,79}{1,869} = 34037$, а также число степеней свободы

$$(n-d-4)/2 = (20-4-4)/2 = 6$$

Для уровня значимости $\alpha = 1\%$

$$F_{\text{табл}} = 8,47$$

Для уровня значимости $\alpha = 5\%$

$$F_{\text{табл}} = 4,28$$

Поскольку вычисленное значение 34037 больше табличных значений F – критериев при выбранных уровнях значимости 1% и 5%, то в обоих случаях принимается гипотеза H_1 о наличии гетероскедастичности. Данный вывод подтверждается графически (рисунок 2).

Regression Plot

$$Y = -70.4105 + 2.52496X$$

$$R\text{-Sq} = 0.410$$

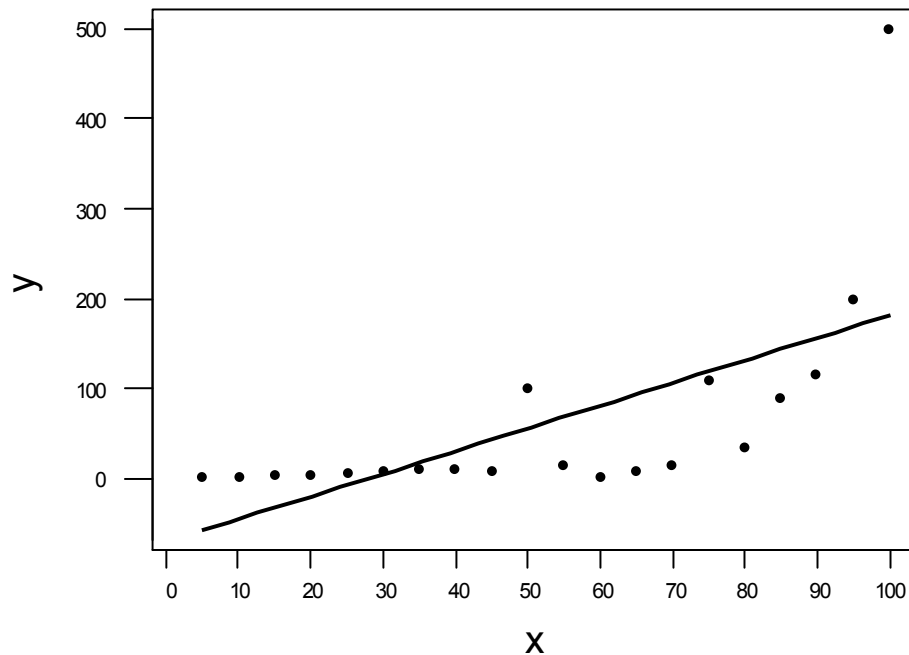


Рисунок 2 – Иллюстрация модели с гетероскедастичностью остатков

2.3. Использование взвешенного метода наименьших квадратов (ВМНК) для оценки моделей с гетероскедастичностью

Для улучшения качества оценки коэффициентов регрессионных моделей с неодинаковой дисперсией (гетероскедастичной случайной составляющей) используется взвешенный метод наименьших квадратов (ВМНК). На данную модель как и на классическую распространяется правило об отсутствии зависимости между случайными составляющими любых последовательных наблюдений. Метод ВМНК, как и МНК, применим к однофакторной и множественной линейной регрессии и использует то же правило минимизации суммы квадратов остатков (ошибок), RSS , однако, во ВМНК вместо одинаковых весов для каждого наблюдения им приписываются значения, обратные соответствующим дисперсиям ошибки, формула (4).

$$RSS = \sum_{i=1}^n w_i (y_i - \tilde{y}_{x_i})^2 = \sum_{i=1}^n \frac{1}{\sigma_{u_i}^2} (y_i - \tilde{y}_{x_i})^2, \quad (4)$$

где $y_i, \tilde{y}_{x_i}$ - фактическое и модельное i -е значение зависимой переменной;

$\sigma_{u_i}^2$ - дисперсия i -й случайной составляющей (ошибки);

w_i – вес i -го наблюдения, определяемый по формуле (5).

$$w_i = \frac{1}{\sigma_{u_i}^2}, \quad (5)$$

Тогда коэффициенты линейной регрессии найдём по формуле (6).

$$\mathcal{E} = [X^T \cdot W \cdot X]^{-1} \cdot X^T \cdot W \cdot Y, \quad (6)$$

где $W = \text{diag}\{w_1, \dots, w_n\}$ - диагональная матрица весов;
n- число наблюдений.

В редких случаях когда известны $\sigma_{u_i}^2$ данная процедура оценки предполагает составление матрицы весов W по известным значениям $\sigma_{u_i}^2$ и используется формула (7) для расчёта коэффициентов.

В более частом случае, дисперсия ошибки $\sigma_{u_i}^2$ неизвестна, но между ней и значением какой-либо объясняющей переменной x_k в регрессионной модели существует линейная зависимость вида (7).

$$\sigma_{u_i}^2 = c \cdot x_{1i}^2, \quad (7)$$

где c - ненулевая константа;

x_{1i} - значение объясняющей переменной x_1 (в данном примере) в i -ом наблюдении.

В случае подобного соотношения $\sigma_{u_i}^2$ можно считать известными, т.к. постоянная величина «с» не влияет на взвешенную процедуру (сокращается в процессе вычисления коэффициентов). Тогда значения весов w_i определяем по формуле (8) для регрессионной модели (9).

$$w_i = \frac{1}{c \cdot x_{1i}^2}, \quad (8)$$

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_p x_{pi} + u_i. \quad (9)$$

Составив таким образом диагональную матрицу весов W , также используем формулу (6) для расчёта коэффициентов.

Пример 2.

Рассмотрим пример расчёта коэффициентов линейной однофакторной регрессии $y_i = \alpha_0 + \alpha_1 x_{1i} + u_i$ с гетероскедастичностью по данным таблицы 2,

$$\sigma_{u_i}^2 = c \cdot x_{1i}^2.$$

Таблица 2 – Величины X и Y

X	Y
10	1
20	3
30	4
80	11

Составим матрицу весов W .

$$W = \begin{bmatrix} \frac{1}{c \cdot 100} & 0 & 0 & 0 \\ 0 & \frac{1}{c \cdot 400} & 0 & 0 \\ 0 & 0 & \frac{1}{c \cdot 900} & 0 \\ 0 & 0 & 0 & \frac{1}{c \cdot 6400} \end{bmatrix}.$$

В дальнейшем опустим константу «с», поскольку она сокращается в процессе расчётов. Оценим коэффициенты по формуле (6).

$$\begin{aligned} [X^T \cdot W \cdot X]^{-1} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 10 & 20 & 30 & 80 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{100} & 0 & 0 & 0 \\ 0 & \frac{1}{400} & 0 & 0 \\ 0 & 0 & \frac{1}{900} & 0 \\ 0 & 0 & 0 & \frac{1}{6400} \end{bmatrix} \cdot \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 80 \end{bmatrix}^{-1} = \begin{bmatrix} 239,252336 & -11,7133956 \\ -11,7133956 & 0,82346838 \end{bmatrix}. \\ [X^T \cdot W \cdot Y] &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 10 & 20 & 30 & 80 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{100} & 0 & 0 & 0 \\ 0 & \frac{1}{400} & 0 & 0 \\ 0 & 0 & \frac{1}{900} & 0 \\ 0 & 0 & 0 & \frac{1}{6400} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 4 \\ 11 \end{bmatrix} = \begin{bmatrix} 0,023663194 \\ 0,520833333 \end{bmatrix}. \end{aligned}$$

$$\mathcal{E} = [X^T \cdot W \cdot X]^{-1} \cdot X^T \cdot W \cdot Y = \begin{bmatrix} 239,2523 & -11,71334 \\ -11,7134 & 0,823468 \end{bmatrix} \cdot \begin{bmatrix} 0,023663 \\ 0,520833 \end{bmatrix} = \begin{bmatrix} -0,43925 \\ 0,151713 \end{bmatrix}.$$

Полученная модель: $y = -0,43925 + 0,151713x$.

3. ИНДИВИДУАЛЬНЫЕ РАСЧЁТНО-ПРАКТИЧЕСКИЕ ЗАДАНИЯ

Задание 1. Ответить на следующие вопросы, выбрав один или несколько вариантов ответа:

1. Гетероскедастичность можно обнаружить с помощью:

- ☐ Теста Вальда
- ☐ Теста Глейзера
- ☐ Теста Голдфелда-Квандта

2. Гетероскедастичность - это:

- ☐ свойство нелинейной модели
- ☐ отсутствие случайной составляющей в регрессии
- ☐ разные значения дисперсий в разных наблюдениях

3. Гетероскедастичность характеризуется тем, что случайная составляющая:

- ☐ имеет одинаковое распределение в каждом наблюдении
- ☐ имеет разное распределение в каждом наблюдении

4. Гетероскедастичность означает:

- ☐ "одинаковый разброс"
- ☐ "неодинаковый разброс"
- ☐ "разное среднее значение"

5. Известны ли исследователю заранее величины дисперсий случайной величины в каждом наблюдении?

- ☐ да
- ☐ нет

6. Какой из тестов более применим к малым выборкам?

- ☐ тест Глейзера
- ☐ тест Голдфелда-Квандта

7. Оценка гетероскедастичной модели МНК-методом является:

- ☐ несостоятельной, но эффективной
- ☐ состоятельной, но неэффективной
- ☐ несостоятельной и неэффективной

8. Гетероскедастичность - это:

- ☐ нарушение условия нормальности случайной составляющей
- ☐ нарушение одинаковой распределённости случайной составляющей

9. Важно ли знать вид зависимости σ_i от X_i для исправления гетероскедастичности?

- ☐ да
- ☐ нет

10. Коррелированы ли случайные составляющие при гетероскедастичности?

- ☐ да
- ☐ нет

11. Известно, что в модели $y_i = \alpha + \beta x_i + u_i$ ошибки линейно зависят от некоторой величины Z_i . Какая из записей будет верной:

- ☐ σ^2
- ☐ $\sigma^2 Z_i$
- ☐ $\sigma^2 Z_i^2$
- ☐ $\sigma^2 Z_i^3$

12. Является ли гетероскедастичность нарушением условий теоремы Гаусса-Маркова?

- ☐ да
- ☐ нет

13. Может ли тест Голдфелда-Квандта обнаружить гетероскедастичность вида

$$\sigma_i = \frac{C}{X_i} ?$$

- ☐ да
- ☐ нет

14. Какой из тестов не требует нормальности распределения случайной составляющей?

- ☐ тест Уайта
- ☐ тест Бреуша-Погана

15. Для оценки модели с гетероскедастичностью применяют:

- ☐ метод взвешенной суммы наименьших квадратов
- ☐ метод исключения переменных

16. Каким из способов можно обнаружить гетероскедастичность:

- ☐ построение диаграммы рассеяния
- ☐ тест Глейзера
- ☐ МНК-оценка параметров

17. Значение t-статистики коэффициента, не превышающее критическое значение свидетельствует об:

- ☐ неправильном вычислении коэффициента
- ☐ не значимости коэффициента в модели
- ☐ гетероскедастичности в модели

18. Какой из тестов кроме подтверждения или опровержения гетероскедастичности дает вид зависимости σ_i от X_i ?

- ☐ тест Глейзера
- ☐ тест Уайта
- ☐ тест Бреуша-Погана

19. Что означает величина F-статистики в тесте Глейзера?

- ☐ эффективность оценки
- ☐ значимость объясняющих переменных

20. При коррекции регрессии на гетероскедастичность нужно оценить модель вида:

- ☐ $\frac{Y_i}{\sigma_i} = \alpha \frac{1}{\sigma_i} + \beta \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i}$
- ☐ $\frac{Y_i}{\sigma_i} = \delta + \alpha \frac{1}{\sigma_i} + \beta \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i}$
- ☐ $\frac{Y_i}{\sigma_i} = \alpha + \beta \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i}$

21. Какая гипотеза в тестах Уайта, Голдфелда-Квандта и Бреуша-Погана принимается за нулевую?

- ☐ гипотеза о гомоскедастичности
- ☐ гипотеза о гетероскедастичности

23. При использовании обобщенного МНК-метода имеется ли в зависимости случайная составляющая?

- ☐ да
- ☐ нет

24. В каком случае гетероскедастичность будет иметь большее значение?

- ☐ значения объясняющих переменных сильно варьируются
- ☐ значения объясняющих переменных слабо варьируются

Задание 2. Проверить модель $y = -83 + 2,52x$, построенную по данным таблицы 3 с использованием МНК, на гетероскедастичность (тест Голдфелда-Квандта $\alpha = 5\%$). Исключить средние $d=N/5$ наблюдений. Построить график.

Таблица 3. – Исходные и сортированные данные X и Y.

№	X	Y	X- сорт.	Y- сорт.
1	105	500		
2	20	3		
3	15	2		
4	25	4		
5	30	6		
6	10	1		
7	35	8		
8	40	10		
9	80	110		
10	50	9		
11	55	100		
12	100	200		
13	70	7		
14	45	11		
15	75	15		
16	95	115		
17	65	2		
18	60	15		
19	90	90		
20	85	35		

Задание 3.

Оценить коэффициенты линейной однофакторной регрессии $y_i = \alpha_0 + \alpha_1 x_i + u_i$ с гетероскедастичностью методом ВМНК по данным таблицы 4, $\sigma_{u_i}^2 = c \cdot x_{ii}^2$.

Таблица 4 – Величины X и Y

X	Y
30	1
40	3
50	4
100	11

4. СОДЕРЖАНИЕ ОТЧЕТА О ПРАКТИЧЕСКОМ ЗАНЯТИИ

- 1) Название и цель работы.
- 2) Условия индивидуальных заданий.
- 3) Ход решения индивидуальных заданий.
- 3) Выводы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Замков О.О. Взвешенный метод наименьших квадратов. Математические методы в экономике / О.О. Замков, А.В. Толстопятенко, Р.Н. Черемных. – М.: Дис, 1997. – 230 с.
2. Glesjer H. A New Test for Heteroscedasticity. Journ. Am. Stat. Assoc., 1969, vol. 64, pp. 316-323.
3. White H. A Heteroscedasticity-Consistent Matrix Estimator and a Direct Test for Heteroscedasticity. Econometrica, vol. 48 (1980), № 4, pp. 817-838.

Заказ № _____ от « _____ » _____ 2008. Тираж _____ экз.

Изд-во СевНТУ

