ВАРИАНТ 1 (резервный)

Задание 1 (20 баллов)

Даны две независимые выборки $x_1, \dots, x_{40}$ и $y_1, \dots, y_{20}$ из распределений $N(a, \sigma^2)$ и $N(b, \sigma^2)$ соответственно.

Известно, что $\sum x_i = 202, \sum x_i^2 = 1190, \sum y_i = 58, \sum y_i^2 = 250$.

(1) Найдите несмещенную оценку дисперсии, пользуясь только 2[й] выборкой.
(2) Найдите несмещенную оценку дисперсии, пользуясь обеими выборками.
(3) Постройте доверительный интервал надежности 95% для b, пользуясь только 2[й] выборкой.
(4) Постройте доверительный интервал надежности 95% для b, пользуясь обеими выборками.

Решение задания 1

(1) $s_y^2 = \frac{1}{19}\sum(y_i - \bar{y})^2 = \frac{\sum y_i^2}{19} - \frac{20}{19}\bar{y}^2 = \frac{250}{19} - \frac{58^2}{19*20} = 4{,}31$

(2) Аналогично пункту один $s_x^2 = \frac{1}{39}\sum(x_i - \bar{x})^2 = \frac{\sum x_i^2}{39} - \frac{40}{39}\bar{x}^2 = \frac{1190}{39} - \frac{202^2}{39*40} = 4{,}36$
И тогда $s_{xy}^2 = \frac{40-1}{40+20-2}s_x^2 + \frac{20-1}{40+20-2}s_y^2 = 4{,}34$

(3) Из нормальности выборки знаем, что $\bar{y} \sim N\left(b, \frac{\sigma^2}{20}\right)$, поэтому $\frac{\bar{y}-b}{\sigma}\sqrt{20} \sim N(0,1)$. Из теоремы Фишера знаем, что $\frac{19 s_y^2}{\sigma^2} \sim \chi^2(19)$, поэтому $\frac{\bar{y}-b}{s_y}\sqrt{20} \sim St(19)$. Тогда (самый короткий) симметричный относительно среднего ДИ имеет вид $\left(\bar{y} - \frac{s_y t_{кр}}{\sqrt{20}}; \bar{y} + \frac{s_y t_{кр}}{\sqrt{20}}\right)$, где $t_{кр}$ – критическая точка распределения Стьюдента с 19 степенями свободы, такая что распределение выходит за нее по модулю с вероятностью 5%. Из таблицы распределения Стьюдента находим $t_{кр} = 2{,}09$ и окончательно ДИ (0,96;4,84).

(4) Снова пользуемся тем, что $\frac{\bar{y}-b}{\sigma}\sqrt{20} \sim N(0,1)$, а также тем, что $\frac{39 s_x^2}{\sigma^2} + \frac{19 s_y^2}{\sigma^2} \sim \chi^2(58)$, откуда $\frac{\bar{y}-b}{s_{xy}}\sqrt{20} \sim St(58)$. Тогда симметричный относительно среднего ДИ имеет вид $\left(\bar{y} - \frac{s_{xy} t_{кр}}{\sqrt{20}}; \bar{y} + \frac{s_{xy} t_{кр}}{\sqrt{20}}\right)$, где $t_{кр}$ – критическая точка распределения Стьюдента с 58 степенями свободы, такая что распределение выходит за нее по модулю с вероятностью 5%. Из таблицы распределения Стьюдента находим $t_{кр} = 2$ и окончательно ДИ (1,97;3,83).

Задание 2 (20 баллов)

Количество договоров страхования, заключенных со страховой компанией в течение t дней, имеет пуассоновское распределение с математическим ожиданием $10t$. При этом, согласно договору страхования, в случае наступления страхового случая страховые выплаты составляют одну и ту же сумму 1 у.е. Накопленная за предыдущие годы статистика показывает, что количество заключенных договоров и страховые случаи по отдельным договорам из данного портфеля являются независимыми в совокупности, а страховой случай наступает с вероятностью 0,1. Рассмотрим суммарные страховые выплаты страховой компании по такому портфелю, сформированному в течение 2019 года (365 дней).

(1) Найдите распределение суммарных выплат.

(2) Найдите среднее значение суммарных выплат.

(3) Найдите дисперсию суммарных выплат.

Задание 3 (20 баллов)

Используя данные по 1000 работников некоторой отрасли, исследователь оценил параметры уравнения (в скобках под оценками коэффициентов указаны соответствующие стандартные ошибки):

$$\widehat{\ln w_i} = \underset{(2,64)}{2,00} + \underset{(0,007)}{0,03}\, exp_i + \underset{(0,001)}{0,60}\, union_i + \underset{(0,002)}{0,04}\, union_i \cdot exp_i, \qquad R^2 = 0,500.$$

w_i — зарплата i-ого работника в тысячах рублей; exp_i — опыт i-го работника в годах; $union_i$ — бинарная переменная, равная единице, если i-ый работник является членом профсоюза.

(1) Вычислите скорректированный коэффициент детерминации R^2_{adj}.

(2) Значим ли коэффициент при переменной exp_i (используйте однопроцентный уровень значимости)? Если да, то дайте его содержательную интерпретацию.

(3) Сравним двух работников, каждый из которых имеет десятилетний стаж работы. Первый работник является членом профсоюза, а второй — нет. На сколько процентов в соответствии с моделью выше зарплата для первого работника по сравнению со вторым?

(4) В соответствии с эконометрической теорией обратная причинно-следственная связь (влияние объясняемой переменной на объясняющую) может негативно влиять на качество модели. Следует ли ожидать возникновения этой проблемы в рассматриваемом примере?

На случай, если ваш калькулятор не вычисляет экспоненты, сообщаем: $e = 2,718$, $e^2 = 7,389$, $e^3 = 20,086$.

Решение задания 3

(1) $R^2_{adj} = R^2 - \frac{k-1}{n-k} * (1 - R^2) = 0,5 - \frac{4-1}{1000-4} * 0,5 = 0,5 * 993/996 = 0,498$

(2) Нулевая гипотеза теста состоит в том, что коэффициент при переменной exp равен нулю. Расчетное значение тестовой статистики составляет 0,03/0,007, что больше, чем критическое. Следовательно, мы отвергаем нулевую гипотезу и делаем вывод о том, что коэффициент значим.

Увеличение опыта работы на один год увеличивает заработную плату работник, **не состоящего в профсоюзе**, на 3%.

(3) Для члена профсоюза с десятилетним стажем предсказанное значение логарифма заработной платы составляет: 2+0,03*10+0,6+0,04*10

Для работника, не состоящего в профсоюзе, предсказанное значение логарифма заработной платы составляет: 2+0,03*10

$$\ln w_1 - \ln w_0 = (2 + 0,03 * 10 + 0,6 + 0,04 * 10) - (2 + 0,03 * 10) = 1$$

$$\frac{w_1}{w_0} = e^1 = 2{,}718$$

Ответ: на (2,718-1)*100%=171,8%

ВАРИАНТ 2 (основной на второй волне 24.07)

Задание 1 (20 баллов)

Для оценки качества изделий, изготовленных двумя стажерами, взяты выборки из 200 изделий первого и 300 изделий второго. В них оказалось по 92 и 148 бракованных изделий соответственно.

(1) Постройте критерий отношения правдоподобия с вероятностью ошибки первого рода, равной 10%, для проверки гипотезы о том, что вероятность брака на втором заводе составляет 45%, против гипотезы о том, что она составляет 50%.

(2) Найдите мощность критерия, построенного на предыдущем шаге.

(3) Предложите критерий для проверки гипотез из п. 1, обладающий единичной мощностью.

(4) Можно ли утверждать на уровне значимости 10%, что вероятность изготовления брака на обоих заводах одинакова?

Решение задания 1

1) Имеем выборки $x_1, \dots, x_{200}$ из распределения Бернулли с параметром p_x и $y_1, \dots, y_{300}$ из распределения Бернулли с параметром p_y. Знаем, что $\sum x_i = 92$, $\sum y_i = 148$. Сформулируем гипотезы
$H_0: p_y = 0{,}45$
$H_1: p_y = 0{,}5$
и выпишем функции правдоподобия и отношение правдоподобия
$L_0(y_1, \dots, y_{300}) = 0{,}45^{\sum y_i} 0{,}55^{300-\sum y_i}$
$L_1(y_1, \dots, y_{300}) = 0{,}5^{\sum y_i} 0{,}5^{300-\sum y_i}$
$\frac{L_1}{L_0} = \left(\frac{0{,}5}{0{,}45}\right)^{\sum y_i} \left(\frac{0{,}5}{0{,}55}\right)^{300-\sum y_i}$
Видим, что неравенство $\frac{L_1}{L_0} > c$, задающее критическую область критерия, равносильно неравенству $\bar{y} > c1$. Найдем константу c1 из условия, что вероятность ошибки первого рода равна 10% (воспользуемся тем, что объем выборки 300 позволяет пользоваться ЦПТ):

$$0{,}1 = P(H_1|H_0) = P(\bar{y} > c1|H_0) = 1 - P\left(\frac{\bar{y} - 0{,}45}{\sqrt{0{,}45 * 0{,}55}}\sqrt{300} < \frac{c1 - 0{,}45}{\sqrt{0{,}45 * 0{,}55}}\sqrt{300}|H_0\right)$$
$$\to 1 - \Phi\left(\frac{c1 - 0{,}45}{\sqrt{0{,}45 * 0{,}55}}\sqrt{300}\right) = 0{,}5 - \Phi_0\left(\frac{c1 - 0{,}45}{\sqrt{0{,}45 * 0{,}55}}\sqrt{300}\right)$$

откуда $\frac{c1-0{,}45}{\sqrt{0{,}45*0{,}55}}\sqrt{300} = \Phi_0^{-1}(0{,}4) = 1{,}28$
Окончательно критерий имеет вид: принимать H_1, если $\bar{y} > 0{,}48$
В нашем случае $\bar{y} = \frac{148}{300}$, поэтому мы принимаем H_1

2) $1 - \beta = 1 - P(\bar{y} < c1|H_1) \to 1 - \Phi\left(\frac{c1-0{,}5}{\sqrt{0{,}5*0{,}5}}\sqrt{300}\right) = 0{,}75$

3) Всегда принимать H_1

4) Сформулируем гипотезы
$H_0: p_x = p_y$
$H_1: p_x \neq p_y$
Если верна H_0, то наилучшей оценкой для общего параметра будет $\hat{p} = \frac{\sum x_i + \sum y_i}{200+300} = 0{,}48$

Объемы выборок достаточно велики, чтобы пользоваться стандартным критерием с критической областью $|u_{\text{расч}}| > u_{\text{кр}}$, где $u_{\text{расч}} = \frac{\frac{\sum x_i}{200} - \frac{\sum y_i}{300}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{200}+\frac{1}{300}\right)}} = -0{,}73$, $u_{\text{кр}} = \Phi_0^{-1}\left(\frac{0{,}9}{2}\right)$=1,65. Поэтому на наших выборках мы принимаем H_0, т.е. можем считать, что вероятность изготовления брака на обоих заводах одинакова.

Задание 2 (20 баллов)

Магазин закупает некоторый товар у трех поставщиков в пропорции $3: 2: 1$, при этом доля брака у k-го поставщика равняется $0{,}1 \cdot k,\ k = 1{,}2{,}3$.

(1) Какова вероятность того, что клиенту, купившему данный товар в магазине, достанется бракованный экземпляр?

(2) Покупатель совершил покупку данного товара в магазине и, придя домой, обнаружил брак. Какова вероятность того, что этот товар был поставлен в магазин первым поставщиком?

Задание 3 (20 баллов)

Используя данные по 400 компаниям некоторой отрасли, исследователь при помощи обычного метода наименьших квадратов оценил параметры двух уравнений:

$$\widehat{\ln TC_i} = \underset{(0{,}12)}{1{,}00} + \underset{(0{,}08)}{0{,}80} \ln Q_i, \qquad R^2 = 0{,}2, \tag{1}$$

$$\widehat{\ln TC_i} = \underset{(0{,}44)}{1{,}00} + \underset{(0{,}03)}{1{,}50} \ln Q_i - \underset{(0{,}08)}{1{,}00} D_i - \underset{(0{,}07)}{0{,}50} D_i \cdot \ln Q_i, \qquad R^2 = 0{,}6. \tag{2}$$

Здесь TC_i — общие издержки i-ой фирмы в тысячах рублей; Q_i — объем выпуска i-ой фирмы в тысячах штук; D_i — бинарная переменная, равная единице, если i-ая фирма использует энергосберегающую технологию, и равная нулю в противном случае.

Гипотезы в пунктах (1)-(2) проверяйте при уровне значимости 5%.

(1) Используя F-тест для сравнения «короткой» и «длинной» регрессий, сделайте выбор между первой и второй моделями.

(2) Значим ли коэффициент при переменной $\ln Q_i$ в модели (2)? Если да, то дайте его содержательную интерпретацию.

(3) Предположим, исследователь принял решение использовать для оценки второй модели не обычный, а обобщенный метод наименьших квадратов. Опишите процедуру оценки параметров модели при помощи этого метода.

(4) В соответствии с эконометрической теорией обратная причинно-следственная связь (влияние объясняемой переменной на объясняющую) может негативно влиять на качество модели. Следует ли ожидать возникновения этой проблемы в рассматриваемом примере?

На случай, если ваш калькулятор не вычисляет экспоненты, сообщаем: $e = 2{,}718$, $e^2 = 7{,}389$, $e^3 = 20{,}086$.

Решение задания 3

(1) Нулевая гипотеза состоит в том, что коэффициенты при переменных D и DlnQ одновременно равны нулю. Расчетное значение тестовой статистики составляет:

$$F_{\text{расч}} = \frac{R_{UR}^2 - R_R^2}{1 - R_{UR}^2} * \frac{n-k}{q} = \frac{0{,}6 - 0{,}2}{1 - 0{,}6} * \frac{400 - 4}{2} = 198$$

Это больше критического значения. Следовательно, мы отвергаем нулевую гипотезу и делаем вывод в пользу «длинной» (второй) модели.

(2) Нулевая гипотеза теста состоит в том, что коэффициент при переменной lnQ равен нулю. Расчетное значение тестовой статистики составляет 1,50/0,03, что больше, чем критическое. Следовательно, мы отвергаем нулевую гипотезу и делаем вывод о том, что коэффициент значим.

Увеличение выпуска на 1% увеличивает издержки фирмы, **не использующей энергосберегающую технологию**, на 1,5%.

(3) Описание процедуры можно найти в пятой главе учебника:

Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. — 6-е изд., перераб. и доп. — М.: Дело, 2004.

ВАРИАНТ 3 (основной на первой волне 25.06)

Задание 1 (20 баллов)

Ежедневный объем продаж некоторого товара (в тыс. кг) в торговой сети имеет равномерное распределение на отрезке $[a, b]$. За 100 дней наблюдения объем продаж составил в среднем 19 тыс. кг, а среднее значение квадрата объема продаж составило 410 тыс. кг^2.

(1) С помощью метода моментов оцените параметры a и b.

(2) С помощью метода моментов оцените объем продаж, на который торговая сеть может рассчитывать с вероятностью 90%.

(3) Будет ли оценка параметра b, найденная вами в п.1, несмещенной, завышенной или заниженной?

(4) Пусть стало известно точное значение параметра b, равное 30. Постройте асимптотический доверительный интервал для параметра a надежности 90% (указание: воспользуйтесь асимптотической нормальностью выборочного среднего арифметического).

Решение задания 1

1)

Имеем выборку $x_1, \dots, x_{100}$ объема n=100 из распределения $U[a, b]$, $\bar{x} = 19, \overline{x^2} = 410$. Найдем несмещенную оценку дисперсии: $s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2 = \frac{n\overline{x^2}}{n-1} - \frac{n}{n-1}\bar{x}^2 = 49{,}5$. Приравнивая теоретические моменты $Ex_i = \frac{a+b}{2}$ и $Dx_i = \frac{(b-a)^2}{12}$ к их выборочным аналогам $\bar{x}$ и s^2, получим

$$\hat{a} = \bar{x} - \sqrt{3s^2} = 6{,}81$$
$$\hat{b} = \bar{x} + \sqrt{3s^2} = 31{,}19$$

2)

Нужно оценить такой параметр q, что $P(x_i < q) = 0{,}9$, откуда $q = a + 0{,}9(b - a)$ и $\hat{q} = \hat{a} + 0{,}9(\hat{b} - \hat{a})$ =28,75 тыс. кг.

3)

$E\hat{b} = E\bar{x} + E\sqrt{3s^2} = Ex_i + \sqrt{3}E\sqrt{s^2} < \{\text{по неравенству Йенсена}\} < Ex_i + \sqrt{3}\sqrt{Es^2} = \frac{a+b}{2} + \sqrt{3}\sqrt{\frac{(b-a)^2}{12}} = b$, поэтому оценка занижена.

4)

Имеем выборку $x_1, \dots, x_{100}$ объема n=100 из распределения $U[a, 30]$. При таком объеме можем пользоваться асимптотическими свойствами выборочного среднего: $\bar{x} \sim \text{ас}N\left(\frac{a+30}{2}, \frac{(30-a)^2}{12n}\right)$, поэтому $\tilde{a} = 2\bar{x} - 30 \sim \text{ас}N\left(a, \frac{(30-a)^2}{3n}\right)$ и $\frac{\tilde{a}-a}{30-a}\sqrt{3n} \sim \text{ас}N(0{,}1)$. Поскольку выборочное среднее состоятельно, то и $\tilde{a}$ – состоятельная оценка, и замена a на $\tilde{a}$ не нарушит асимптотической нормальности: $\frac{\tilde{a}-a}{30-\tilde{a}}\sqrt{3n} \sim \text{ас}N(0{,}1)$ (можно такую замену не делать и выразить а напрямую, такой вариант, конечно, тоже засчитываем, при этом интервал получится уже). Тогда $P\left(-u_{\text{кр}} < \frac{\tilde{a}-a}{30-\tilde{a}}\sqrt{3n} < u_{\text{кр}}\right) \to 0{,}9$, где $u_{\text{кр}} = \Phi_0^{-1}\left(\frac{0{,}9}{2}\right)$=1,65.

Окончательно асимптотический доверительный интервал надежности 90% имеет вид

$$\tilde{a} \pm \frac{u_{\text{кр}}}{\sqrt{3n}}(30 - \tilde{a}) = 2\bar{x} - 30 \pm \frac{u_{\text{кр}}}{\sqrt{3n}}(60 - 2\bar{x}) = 8 \pm 5$$

Задание 2 (20 баллов)

Один из портфелей страховой компании состоит из n полисов, страховые выплаты по которым являются независимыми и имеют показательное распределение с параметром 0,5.

(1) Найти математическое ожидание и дисперсию суммарных страховых выплат по данному портфелю.

(2) Найти предельное при $n \to \infty$ значение вероятности того, что суммарные страховые выплаты по данному портфелю превысят $\frac{2n}{1-1/\sqrt{n}}$.

(3) Найти все возможные значения констант $a, b \in \mathbb{R}$, такие что предельное при $n \to \infty$ значение вероятности того, что суммарные страховые выплаты по данному портфелю превысят $an + b\sqrt{n}$, равняется 0,1.

Решение

(1)

Пусть X_k – выплаты по k-ому полису, тогда $S_n = X_1 + \dots + X_n$ – суммарные страховые выплаты по портфелю. Тогда $\mathbb{E}X_k = 2, DX_k = 4, \mathbb{E}S_n = 2n, DS_n = 4n$.

(2)

В силу ЦПТ и равномерной сходимости

$$P\left(S_n > \frac{2n}{1-\frac{1}{\sqrt{n}}}\right) = P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{DS_n}} > \frac{1}{2\sqrt{n}}\left(\frac{2n}{1-\frac{1}{\sqrt{n}}} - 2n\right)\right) =$$

$$= P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{DS_n}} > 1 + O\left(\frac{1}{\sqrt{n}}\right)\right) \to 1 - \Phi(1) = \frac{1}{\sqrt{2\pi}}\int_1^{\infty} e^{-\frac{t^2}{2}}dt = 0.1586 \dots$$

(3)

Аналогично пункту (2):

$$P\left(S_n > an + b\sqrt{n}\right) = P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{DS_n}} > \frac{an + b\sqrt{n} - 2n}{2\sqrt{n}}\right) \to$$

$$\to 1 - \lim_{n\to\infty} \Phi\left(\left(\frac{a}{2} - 1\right)\sqrt{n} + \frac{b}{2}\right) = \begin{cases} 1, a < 2 \\ 1 - \Phi\left(\frac{\mathrm{b}}{2}\right), a = 2 \\ 0, a > 2 \end{cases} = 0.1 \Leftrightarrow$$

$$\Leftrightarrow a = 2, b = 2\Phi^{-1}(0.9) = 2 * 1.2815 \dots = 2.5631 \dots$$

Задание 3 (20 баллов)

В вашем распоряжении имеются следующие данные о 540 работниках (270 мужчин и 270 женщин): l_EARNINGS — **логарифм** текущего часового заработка в долларах США, S — продолжительность обучения (число полных лет обучения), l_EXP — **логарифм** общего стажа работы после окончания учебы, FEMALE — пол респондента (0 для мужчин, 1 для женщин). Цель исследователя состояла в том, чтобы выявить влияние опыта работы и образования на доход индивида. Ниже представлены результаты оценивания одной из моделей:

Модель 2: МНК, использованы наблюдения 1-540
Зависимая переменная: l_EARNINGS

	Коэффициент	*Ст. ошибка*	*t-статистика*	*P-значение*
const	**«A»**	0,0200	1,96	0,0500
l_EXP	0,3215	**«B»**	5,356	0,0001
S	0,1243	0,0087	14,22	0,0001
S*FEMALE	−0,0204	0,0030	−6,682	0,0001

Среднее зав. перемен	2,791993	Ст. откл. зав. перемен	0,588554
R-квадрат	**«C»**	Испр. R-квадрат	0,315195

Тест Бреуша-Пагана (Breusch-Pagan) на гетероскедастичность -
Тестовая статистика: LM = 7,77811
p-значение = P(Хи-квадрат(3) > 7,77811) = 0,0508271

(1) Заполните отмеченные буквами «A», «B», «C» пропуски в таблице. Обоснуйте ответ.
(2) Дайте содержательную интерпретацию коэффициента при переменной l_EXP.
(3) Интерпретируйте представленные результаты теста Бреуша-Пагана, используя пятипроцентный уровень значимости. Опишите процедуру проведения этого теста.

Решение

(1) Найдем значение параметров A и B на основе информации столбцов «Коэффициент», «Ст. ошибка» и «t-статистика (t-расчетное)»:

$$t_{\text{расч}} = \frac{\hat{\beta}_i}{\widehat{se}(\hat{\beta})}$$

$$A: \frac{A}{0{,}02} = 1{,}96 \Rightarrow A = 0{,}0392$$

$$B: \frac{0{,}3215}{B} = 5{,}356 \Rightarrow B = 0{,}06$$

R^2 найдем по имеющемуся R^2-исправленному:

$$R^2_{adj} = R^2 - (1 - R^2)\frac{k-1}{n-k}$$

$$0{,}315195 = R^2 - (1 - R^2)\frac{4-1}{540-4}$$

$$R^2 = 0{,}319$$

(2) Пусть при прочих равных опыт работника растет на 1% ($EXP_1 = 1{,}01EXP_0$), тогда:

$$lnEARNINGS_1 - lnEARNINGS_0 = 0{,}3215(lnEXP_1 - lnEXP_0)$$
$$\frac{EARNINGS_1}{EARNINGS_0} = (1{,}01)^{0{,}3215}$$

Таким образом, увеличение опыта работника на 1% при прочих равных приводит к увеличению часового заработка в долларах США на 0,32%.

(3) Опишем процедуру теста.

Исходная модель: $y_i = \beta_1 + \beta_2 x_i^{(2)} + \cdots + \beta_k x_i^{(k)} + \varepsilon_i$

$$H_0\colon \sigma_1^2 = \cdots = \sigma_n^2$$
H_1: дисперсия ε_i зависит от заданного набора переменных,

$$\text{то есть } \sigma_i^2 = \gamma_0 + \gamma_1 z_i^{(1)} + \cdots + \gamma_p z_i^{(p)}$$

Это могут быть переменные из модели или какие-то другие.

Оцениваем исходную модель и сохраняем из нее остатки $e_i = y_i - \hat{y}_i$.

Вычисляем $\tilde{\sigma}^2 = \frac{1}{n}\sum e_i^2$.

Оцениваем параметры модели

$$\frac{e_i^2}{\tilde{\sigma}^2} = \gamma_0 + \gamma_1 z_i^{(1)} + \cdots + \gamma_p z_i^{(p)} + u_i$$

Расчетное значение статистика равно $\frac{RSS}{2}$, где RSS взят из оцененной на предыдущем шаге модели.

Табличное значение $\chi^2(p)$.

В нашем случае оказалось, что p-значение для этого теста равно 0,0508271, то есть больше 0,05, соответствующего заданному уровню значимости. Таким образом, принимается гипотеза H_0, то есть в модели нет проблемы гетероскедастичности.