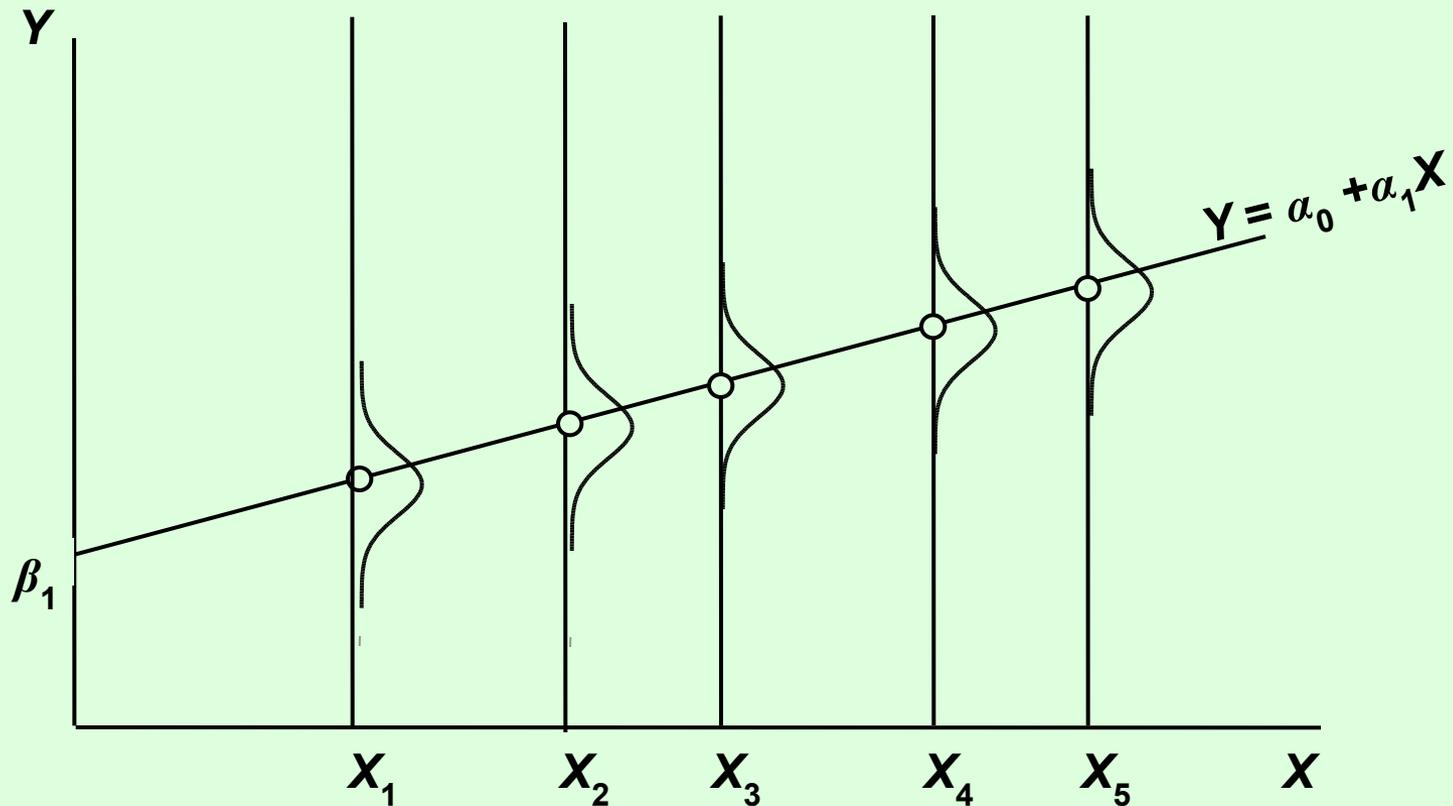
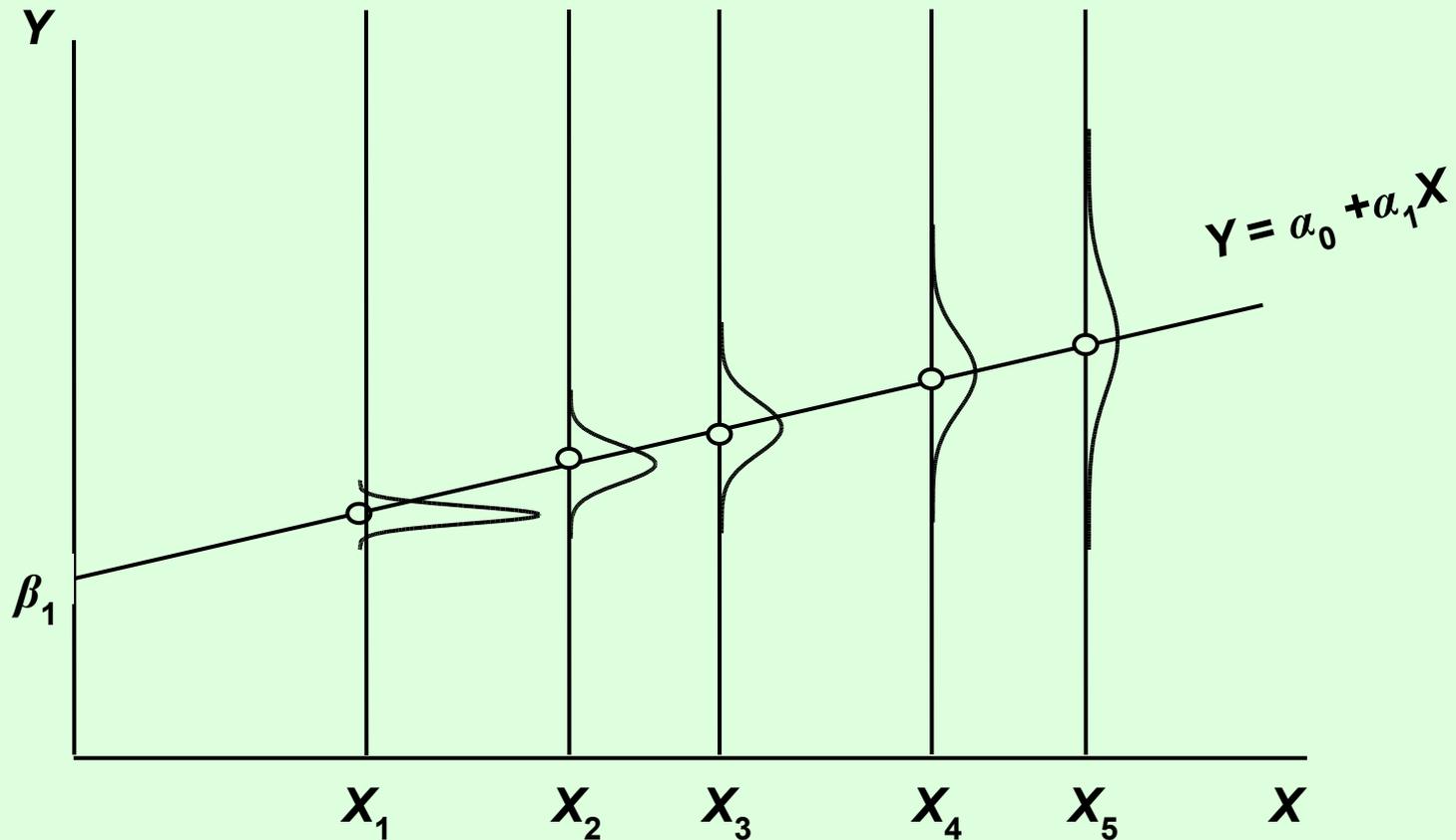


Гетероскедостичность и ее последствия



Наличие случайного возмущения приводит к размытости значений Y независимо от X . Для случайного возмущения предполагается выполнение ряда требований: условий теоремы Гаусса-Маркова.

Гетероскедостичность и ее последствия



Распределение u для каждого наблюдения имеет нормальное распределение и нулевое ожидание, но дисперсия распределений различна.

Гетероскедастичность и ее последствия

- **Условия обеспечивающие гомоскедастичность (однородность) случайных возмущений:**
 1. Нормальное распределение случайных возмущений для всех наблюдений.
 2. Средние значения случайных возмущений в каждом наблюдении равно нулю.
 3. Распределения одинаковы для всех наблюдений.
- **Последствия нарушения условия гомоскедастичности случайных возмущений:**
 1. Потеря эффективности оценок коэффициентов регрессии, т.е. можно найти другие, отличные от МНК и более эффективные оценки.
 2. Смещенность стандартных ошибок коэффициентов в связи с некорректностью процедур их оценки.

Гетероскедостичность и ее последствия

В связи с тем, что оценка всех параметров модели, включая вид параметры закона распределения случайного возмущения, проводится по результатам случайной выборки, то справедливо говорить только о статистических гипотезах относительно выполнения предпосылок теоремы Гаусса-Маркова.

Вспомним, как производится проверка статистических гипотез.

Методика проверки статистических гипотез

Определение. Под статистической гипотезой понимается любое предположение о виде закона распределения случайной величины или значениях его параметров.

Примеры статистических гипотез:

H_0 : (U имеет нормальный закон распределения).

H_0 : (параметр $a_0 = 0$)

H_1 : (параметр $a_0 = 1$)

Гипотезы H_0 и H_1 называются основной и альтернативной.

Методика проверки статистических гипотез

Алгоритм проверки статистических гипотез.

1. Формулируется статистическая гипотеза H_0 .
2. Искусственно формируется случайная величина « Z », закон распределения которой известен $[P_z(t, a_1, a_2)]$, которая тесно связана с гипотезой.
3. Область допустимых значений Z делится на две части: Ω_0 в которой гипотеза принимается и, Ω в которой она отклоняется. Граница этих областей определяется из условия, что Z попадает в область Ω_0 с заданной вероятностью « ρ ».
4. По данным выборки вычисляется значение случайной величины Z и проверяется ее принадлежность области Ω_0 .

Методика проверки статистических гипотез

Примеры. В схеме Гаусса-Маркова случайные переменные:

$$t = \frac{|a_i - c|}{\sigma(a_i)}; \quad t = \frac{|\tilde{y} - c|}{\sigma(y)}$$

называются дробью Стьюдента и подчиняются закону распределения Стьюдента. Критическое значение дроби Стьюдента находится из уравнения:

$$P(|t| < t_{кр}) = \int_{-t_{кр}}^{t_{кр}} P_t(q) dq = P_{дов}$$

Здесь: $P_t(q)$ функция плотности вероятности распределения Стьюдента, $t_{кр}$ – двусторонняя квантиль распределения, $P_{дов}$ - значение доверительной вероятности, как правило $P_{дов} = 0.95/0.99$ 7

Методика проверки статистических гипотез

В схеме Гаусса-Маркова переменная:

$$f = \frac{\sum_{i=1}^n \boldsymbol{\varepsilon}_i / n}{\sum_{j=1}^m \boldsymbol{\varepsilon}_j / m}$$

подчиняется закону распределения Фишера и критическое значение этой дроби вычисляется из условия:

$$P(f < f_{кр}) = \int_0^{f_{кр}} p_f(q) dq = P_{дов}$$

Методика проверки статистических гипотез

Возможные ошибки при проверке статистических гипотез.

Ошибка первого рода. Когда справедливая гипотеза отклоняется.

Ошибка второго рода. Когда ложная гипотеза принимается.

Тест Готвальда-Квандта

Данный тест предназначен для того, чтобы проверить гипотезу об отсутствии гетероскедастичности случайных возмущений в схеме Гаусса-Маркова.

Случай уравнения парной регрессии.

Имеем спецификацию модели в виде:

$$Y_t = a_0 + a_1 x_t + u_t$$

Имеем выборку в объеме n наблюдений за переменными Y_t и x_t для оценки параметров этой модели.

Задача: проверить гипотезу об отсутствии гетероскедастичности в полученной модели.

Тест Готвальда-Квандта

В основе теста лежат два предположения.

1. Случайные возмущения подчиняются нормальному закону распределения.
2. Стандартные ошибки случайных возмущений $\sigma(u_t)$ пропорциональны значениям регрессора x_t .

Алгоритм применения теста Готвальда-Квандта

Шаг 1. Имеющаяся выборка из n наблюдений сортируется по возрастанию значений регрессора x .

Шаг 2. Полученная в результате сортировки выборка делится на три примерно равные части.

Шаг 3. Для первой и третьей частей выборки строятся модели парной регрессии, т.е. для них вычисляются оценки параметров a_0 и a_1 .

В результате получаются две модели парной регрессии (для каждой части общей выборки):

$$Y^1 = \tilde{a}_0^1 + \tilde{a}_1^1 x + u^1 \quad (9.1)$$

$$Y^3 = \tilde{a}_0^3 + \tilde{a}_1^3 x + u^3 \quad (9.2)$$

Исходя из принятых допущений, считается, что, если ошибки случайных возмущений в «первой» и «третьей» частях выборки будут равны, то условие гомоскедостичности выполняется.

Алгоритм применения теста Готвальда-Квандта

Шаг 4. Для уравнений (9.1) и (9.2) вычисляются значения ESS_1 и ESS_3 .

Где $ESS = \Sigma(u_i^2) = \Sigma(y_i - \tilde{a}_0 - \tilde{a}_1 x_i)^2$

Шаг 5. Проверяется гипотеза о равенстве σ_u^1 и σ_u^3 .

5.1. Формируется случайная переменная GQ в виде:

$$GQ = \frac{ESS^1 / n_1}{ESS^3 / n_3}$$

В схеме Гаусса-Маркова переменная GQ имеет закон распределения Фишера.

Алгоритм применения теста Готвальда-Квандта

5.2. Вычисленное значение GQ сравнивается с критическим

значением $F_{кр}(P_{дов}, n_1, n_3)$:

Если $GQ \leq F_{кр}(P_{дов}, n_1, n_3)$

и $1/GQ \leq F_{кр}(P_{дов}, n_1, n_3)$,

то гипотеза о гомоскедастичности случайных возмущений принимается.

Случай уравнения множественной регрессии.

$$Y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + a_3 x_{3t} + u_t$$

Сортировка проводится по величине $z = |x_1| + |x_2| + |x_3|$.

Если тест дает отрицательный результат, алгоритм повторяется для каждого регрессора.

В результате обнаруживается регрессор вызывающий гетероскедастичность.

Алгоритм применения теста Готвальда-Квандта

Государственные расходы на образование в различных странах

Применение ф-ии «ЛИНЕЙН»

Страна	Расходы	ВВП	Страна	Расходы	ВВП
Люксембург	0,34	5,67	Швейцария	5,31	101,7
Уругвай	0,22	10,1	С.Аравия	6,4	116
Сингапур	0,32	11,3	Бельгия	7,15	119,5
Ирландия	1,23	18,9	Швеция	11,22	124,2
Израиль	1,81	20,9	Австралия	8,66	141
Венгрия	1,02	22,2	Аргентина	5,56	153,9
Н.Зеландия	1,27	23,8	Нидерланды	13,41	169,4
Португалия	1,07	24,7	Мексика	5,46	186,3
Гонконг	0,67	27,6	Испания	4,79	211,8
Чили	1,25	27,6	Бразилия	8,92	249,7
Греция	0,75	40,2	Канада	18,9	261,4
Финляндия	2,8	51,6	Италия	15,95	395,5
Норвегия	4,9	57,7	Англия	29,9	535
Югославия	3,5	63	Франция	33,59	655,3
Дания	4,45	66,3	ФРГ	38,62	815
Турция	1,6	67	Япония	61,61	1040
Австрия	4,26	76,9	США	181,3	2586

0,0414	0,0821	0,0711	-8,187
0,0123	0,3276	0,0027	2,4453
0,5309	0,518	0,9854	6,2309
11,318	10	674,45	10
3,0371	2,6835	26185	388,24

$$GQ = \frac{2.68}{388.24} = 0.0069$$

F_{кр}=3.0

$$1/GQ = \frac{1}{0.0069} = 144.7$$

Модель:

$$Y = -2.32 + 0.067X \quad (9.4)$$

Метод исправления гетероскедастичности

Имеем:

1. Спецификацию модели:

$$Y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + a_3 x_{3t} + u_t \quad (9.1)$$

2. Выборку наблюдений за переменными
 $\{Y, x_1, x_2, x_3\}$

3. Модель по этим данным гетероскедастична.

4. Известны значения $\sigma(u_t)$ в каждом наблюдении.

Задача: преобразовать модель так, чтобы случайные возмущения были гомоскедастичны.

Метод исправления гетероскедастичности

Способ 1. Делится каждое уравнение наблюдений на свое $\sigma(u_t)$ и получается:

$$\frac{y_t}{\sigma(u_t)} = \frac{a_0}{\sigma(u_t)} + a_1 \frac{x_{1t}}{\sigma(u_t)} + a_2 \frac{x_{2t}}{\sigma(u_t)} + a_3 \frac{x_{3t}}{\sigma(u_t)} + \frac{u_t}{\sigma(u_t)} \quad (9.2)$$

Тогда дисперсия случайного возмущения в каждом уравнении наблюдений есть:

$$\sigma^2\left(\frac{u_t}{\sigma(u_t)}\right) = \frac{1}{\sigma^2(u_t)} \sigma^2(u_t) = 1$$

Модель (9.2) в каждом уравнении наблюдения имеет одинаковые дисперсии случайного возмущения равные 1.

Недостаток способа – оценить $\sigma(u_t)$ не возможно!

Метод исправления гетероскедастичности

Способ 2.

Предполагаем, что $\sigma(u_t) = \lambda x_{kt}$, где x_{kt} регрессор «вызывающий» гетероскедастичность. Пусть для примера это регрессор x_{2t} .

Уравнение (9.1) делится на значение этого регрессора.

$$\frac{y_t}{x_{2t}} = a_0 \frac{1}{x_{2t}} + a_1 \frac{x_{1t}}{x_{2t}} + a_2 + a_3 \frac{x_{3t}}{x_{2t}} + \frac{u_t}{\left(\frac{\sigma(u_t)}{\lambda}\right)} \quad (9.3)$$

Дисперсия случайного возмущения при этом есть:

$$\sigma^2 \left(\frac{u_t}{\frac{\sigma(u_t)}{\lambda}} \right) = \lambda^2 \frac{\sigma^2(u_t)}{\sigma^2(u_t)} = \lambda^2$$

Уравнения модели (9.3) имеют постоянную дисперсию случайного возмущения равную λ^2 .

Метод исправления гетероскедастичности

Рассмотренные способы устранения гетероскедастичности носят название «Взвешенный метод наименьших квадратов».

Теорема. Если в схеме Гаусса-Маркова не выполняется предпосылка о гомоскедастичности случайных возмущений, то наилучшей линейной процедурой оценки параметров модели является:

$$\hat{a} = (X^T P^{-1} X)^{-1} X^T P^{-1} Y$$

где: P матрица ковариаций случайных возмущений в уравнения наблюдений:

$$P = \begin{pmatrix} \sigma^2(u_1) & 0 & \dots & 0 \\ 0 & \sigma^2(u_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma^2(u_n) \end{pmatrix}$$

Метод исправления гетероскедастичности

Относительные расходы на образование в различных странах

Страна	Y/ВВП	1/ВВП	Страна	Y/ВВП	1/ВВП
Люксембург	0,060	0,1764	Швейцария	0,052	0,0098
Уругвай	0,022	0,0987	С.Аравия	0,055	0,0086
Сингапур	0,028	0,0882	Бельгия	0,060	0,0084
Ирландия	0,065	0,0530	Швеция	0,090	0,0081
Израиль	0,086	0,0478	Австралия	0,061	0,0071
Венгрия	0,046	0,0451	Аргентина	0,036	0,0065
Н.Зеландия	0,053	0,0420	Нидерланды	0,079	0,0059
Португалия	0,043	0,0405	Мексика	0,029	0,0054
Гонконг	0,024	0,0363	Испания	0,023	0,0047
Чили	0,045	0,0363	Бразилия	0,036	0,0040
Греция	0,019	0,0249	Канада	0,072	0,0038
Финляндия	0,054	0,0194	Италия	0,040	0,0025
Норвегия	0,085	0,0173	Англия	0,056	0,0019
Югославия	0,056	0,0159	Франция	0,051	0,0015
Дания	0,067	0,0151	ФРГ	0,047	0,0012
Турция	0,024	0,0149	Япония	0,059	0,0010
Австрия	0,055	0,0130	США	0,070	0,0004

Применение ф-ии «ЛИНЕЙН»

0,0297	0,0438	-2,628	0,0585
0,1453	0,0105	2,5862	0,0098
0,0042	0,021	0,0936	0,0179
0,0417	10	1,0329	10
2E-05	0,0044	0,0003	0,0032

$$GQ = \frac{0.0044}{0.0032} = 1.377$$

$$1/GQ = \frac{1}{1.377} = 0.726$$

$F_{кр} = 3.0$

Модель:

$$Y = -0.066 + 0.053X \quad (9.5)$$

Метод исправления гетероскедастичности

Диаграмма рассеяния и графики моделей с гетероскедастичными и гомоскедастичными случайными возмущениями.

