36. **Тест Голдфелда-Квандта гомоскедастичности случайного возмущения в ЛММР.**

Тест для проверки второй предпосылки теоремы Гаусса-Маркова. Предпосылка означает независимость дисперсий случайных возмущений от значений объясняющих переменных. (гомоскедостичность)

$2)\ Var\left(u_1^2\right) \ = Var\left(u_2^2\right) \ = \ \dots \ = Var\left(u_n^2\right) = \sigma_n^2;$

Шаги теста предпосылки №2 Голдфелда-Кванта

Шаг 1. Составляется система уравнений наблюдений объекта

$$\begin{cases} y_1 = a_0 + a_1 \cdot x_{1,1} + a_2 \cdot x_{2,1} + \dots a_k \cdot x_{k,1} + u_1 \\ y_2 = a_0 + a_1 \cdot x_{1,2} + a_2 \cdot x_{2,2} + \dots a_k \cdot x_{k,2} + u_2 \\ \dots \\ y_n = a_0 + a_1 \cdot x_{1,n} + a_2 \cdot x_{2,n} + \dots a_k \cdot x_{k,n} + u_n \end{cases}$$

Шаг 2. Уравения наблюдений упорядычеваются по возрастанию сумм абсолютных значений объясняющих перменных

$$\sum_{j=1}^{k} |x_{ji}|$$

Шаг 3. По первым n_1 упорядоченным уравениям оцениваются методом наименьших квадратов параметры модели и запоминается значения ESS_1. Количество n_1 выбирается согласно следующим двум условиям:

$$a)\ n_1 \ \approx \frac{1}{3}n, \ \ b)\ n_1 \ > \ k \ + \ 1$$

Аналогично оценивается модель по последним n_1 уравениям и запоминается значение ESS_2.

Шаг 4. Вычисляется по следующему правилу дробь:

$$GQ = \frac{ESS_1}{ESS_2} \sim P_F(q)$$

Эта дробь является <u>статистикой</u> критерия проверяемой гипотезы о гомоскедастичности случайного возмущения. Величина GQ имеет распредение Фишера с кол-ом степеней свободы m, n.

Шаг 5. Гипотеза о гомоскедастичности принимается как не противоречащая реальным данным, если оказываеются справедливыми следующие два неравенства:

$$\begin{cases} GQ \overset{?}{\leq} F_{\text{крит}} \\ \dfrac{1}{GQ} \overset{?}{\leq} F_{\text{крит}} \end{cases}$$

Где символом $F_{\text{крит}}$ мы обозначаем квантиль распределения Фишера заданного уровня $1 - \alpha$, например $1 - \alpha \ = 0.95$.

Итог: экономисты тестируют все предпосылки в частности предпосылка №2 тестируется тестом Голдфелда-Кванта.

37. **Тест Дарбина–Уотсона отсутствия автокорреляции у случайного возмущения в ЛММР.**

Проверяемая гипотеза в этом тесте имеет вид:

$$H_0 : Cov(u_{t+1}, u_t) = 0$$

Альтернативная гипотеза заключается в положительном значении ковариации u_{t+1}, u_t:

$$H_1 : Cov(u_{t+1}, u_t) > 0$$

Альтернатива имеет наиболее важное для практики значение. Если справедлива данная альтернатива, то причина этого обстоятельства чаще всего заключается в ошибочной спецификации модели.

Тест $DW(\text{Дарбина} - \text{Уотсона})$ проводится в итоге следующих шагов:

Шаг 1. По уравнениям наблюдений оценивается модель и вычисляется по правилу

$$DW = \frac{\sum_{i=1}^{n-1} \left(\widetilde{u}_{i+1} - u_i\right)^2}{\sum_{i=1}^{n} \left(\widetilde{u}_i\right)^2}$$

статистика критерия гипотезы H_0.

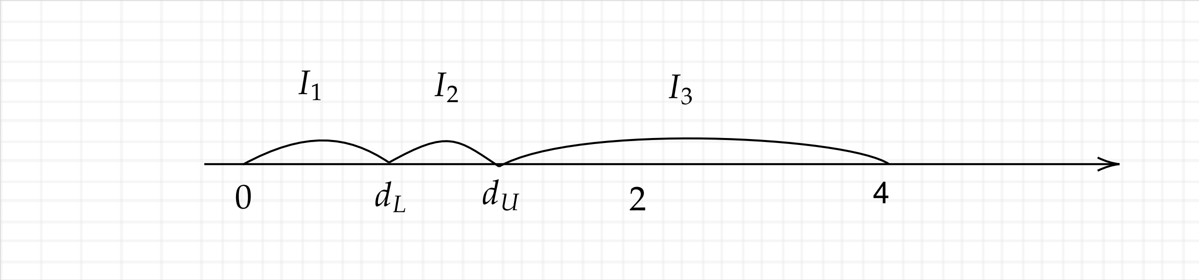
Шаг 2. По таблицам Дарбина-Уотсена

n	$k^1 = 1$		$k^1 = 2$		$k^1 = 3$		$k^1 = 4$		$k^1 = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,61	1,40	–	–	–	–				
7	0,70	1,36	0,47	1,90	–	–				
8	0,76	1,33	0,56	1,78	0,37	2,29				
9	0,82	1,32	0,63	1,70	0,46	2,13				
10	0,88	1,32	0,70	1,64	0,53	2,02				
11	0,93	1,32	0,66	1,60	0,60	1,93				
12	0,97	1,33	0,81	1,58	0,66	1,86				
13	1,01	1,34	0,86	1,56	0,72	1,82				
14	1.05	1.35	0.91	1.55	0.77	1.78				
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83

Figure 1: Значения статистики Дарбина - Уотсона

Выбираются две величины $d_L,\ d_U$ используя два входа $n,\ k$.

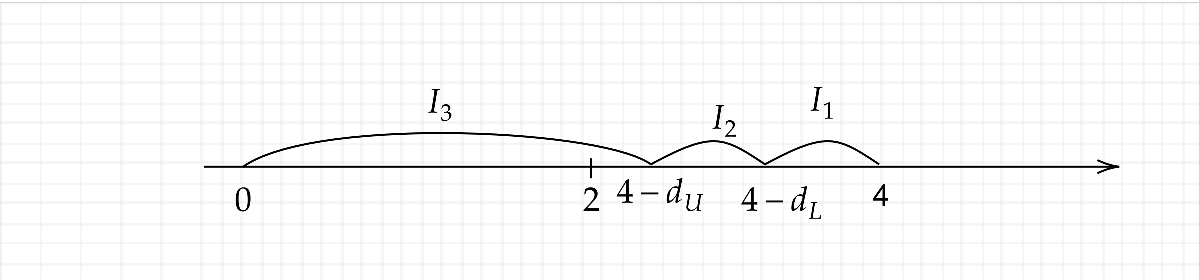
Шаг 3. Определяется один из трёх интервалов в который попадает статистика DW.



Если DW попало в I_3, то H_0 принимается, если в I_1, то гипотеза H_0 отвергается в пользу гипотезы H_1; если в интервал I_2, то ничего сказать нельзя - это *интервал неопределённости*.

Проверка гипотезы при альтернативе $Cov(u_{t+1}, u_t) < 0$

Первые два **шага** остаются без изменений, а чертёж с интервалами выглядит так:



Если статистика DW попадает I_1, то гипотеза H_0 отклонятся в пользу гипотезы H_1 (очень редкий случай), если в I_2 то ничего сказать нельзя - это *интервал неопределённости;* Если DW попадает в I_3, то H_0 принимается.

38. **Коэффициент детерминации как мерило качества спецификации эконометрической модели. Скорректированный коэффициент детерминации и его использование для модификации ЛММР.**

Оценивание эконометрической модели осуществляется на 3-ем этапе схемы её построения. Простешей характеристикой качества служит коэффициент детерминации модели R^2. R^2 – это доля эндогенной переменной модели, которая объясняется предопределёнными переменными модели.

Выведем формулу для величины R^2

Шаг 1. По уравнениям наблюдений рассчитываем оценки случайных возмущений

$$\widetilde{u}_i = y_i - \left(\widetilde{a}_0 + \widetilde{a}_1 \cdot x_i\right) = y_i - \widetilde{y}_i$$

Перепишем следующим образом уравнения наблюдений

$$\begin{cases} y_1 = \widetilde{y}_1 + \widetilde{u}_1 \\ y_2 = \widetilde{y}_2 + \widetilde{u}_2 \\ \dots \\ y_n = \widetilde{y}_n + \widetilde{u}_n \end{cases} \qquad (4.14)$$

В уравнении (4.14) первое слагаемое в правой части объясняются перменной x, а вторые слагаемые необъясняются x-ами. Справедливо, следующая теорема:

$$\sum (y_i - \overline{y})^2 = \sum \left(\widetilde{y}_i - \overline{\widetilde{y}}\right)^2 + \sum \widetilde{u}_i^2 \tag{4.15}$$

В левой части тождества размещается характеристика изменчивости эндогенной переменной - *волатильность.* Первое слагаемое в правой части обозначим его символом RSS объясняется изменчивостью предопределённых значений и полность объясняется x. А второе слагаемое пораждено неучтёнными факторами $ESS = \sum \widetilde{u}_i^2$, левую часть $TSS = \sum \left(\widetilde{y}_i - \overline{\widetilde{y}}\right)^2$. И разделим обе часть тождества (4.15) на велечину TSS в итоге придём к формуле (4.16).

$$R^2 = 1 - \frac{ESS}{TSS} \tag{4.16}$$

Рассматривая (4.16) мы констатируем, что R^2 – это доля эндогенной переменной модели, которая объясняется предопределёнными перменными R^2.
Скоррективонная коэффицент детерминации рассчитывается по следующему правилу:

$$\overline{R}^2 = 1 - \frac{\frac{ESS}{n-(k+1)}}{\frac{TSS}{n-1}}$$

$$\overline{R}^2 \leqslant R^2$$

Числитель и позволяет отбирать в модель объясняющие перменные. Если при включении в модель новой объясняющей переменной велечина $\overline{R}^2$ возрастает, то включение этой переменной в модель полезно, если убывает, то бесмысленно.

Замечание. В числителе вычитаемого $\dfrac{ESS}{n-(k+1)}$ размещается оценка дисперсии случайного возмущения. В знаменателе находится дисперсия эндогенной переменной. При добавлении в модель новой объясняющей переменной меняется только числитель (знаменатель остаётся всегда неизменным) и поэтому увеличение $\overline{R}^2$ равносильно снижению дисперсии случайного возмущения и значит уменьшению экзогенных переменных. Добавим, что всегда имеет место следующее неравенство, при чём $\overline{R}^2$ может быть меньше 0.

39. **Связь коэффициента детерминации с коэффициентом корреляции эндогенной переменной и её оценки.**

Коэффициент детерминации равен квадрату модуля коэффициента корреляции прогноза $\tilde{y}$ и y. Он показывает, какая доля дисперсии результативного признака объясняется влиянием независимых переменных.

Доказательство:

$$Cor(y,\tilde{y}) = \frac{Cov(y,\tilde{y})}{\sqrt{Var(y)}\sqrt{Var(\tilde{y})}} = \frac{Cov(\tilde{y}+u,\tilde{y})}{\sqrt{Var(y)}\sqrt{Var(\tilde{y})}} = \frac{Cov(\tilde{y},\tilde{y})+Cov(u,\tilde{y})}{\sqrt{Var(y)}\sqrt{Var(\tilde{y})}} =$$

$$= \frac{Var(\tilde{y})+Cov(u,\tilde{y})}{\sqrt{Var(y)}\sqrt{Var(\tilde{y})}} = \frac{Var(\tilde{y})}{\sqrt{Var(y)}\sqrt{Var(\tilde{y})}} = \frac{\sqrt{Var(\tilde{y})}}{\sqrt{Var(y)}} = \sqrt{\frac{Var(\tilde{y})}{Var(y)}} = \sqrt{R^2}$$

***При условии, что** (для парной регрессии)*

$\tilde{y} = a_0 + a_1 x;\ \ u = y - a_0 - a_1 x;\ b = \frac{Cov(x,y)}{Var(x)}$ *и с использованием ковариационных правил, можно доказать, что*

$$Cov(u,\tilde{y}) = Cov(y - a_0 - a_1 x, a_0 + a_1 x) = Cov(y - a_1 x, a_1 x) = Cov(y, a_1 x) - Cov(a_1 x, a_1 x) =$$

$$= a_1 Cov(y,x) - {a_1}^2 Cov(x,x) = a_1\big(Cov(y,x) - a_1 Cov(x,x)\big) = a_1(a_1 Var(x) - a_1 Var(x)) = 0$$

(для множественной регрессии аналогично).

40. **F-тест качества спецификации эконометрической модели.**

F – тест - процедура проверки гипотезы о неудовлетворительной спецификации эконометрической модели:

$$H_0 : a_1 = a_2 = \ \dots \ = a_k = 0$$

То есть гипотеза о том, что ни одна объясняющая переменная не несёт в себе информацию об эндогенной переменной y. Альтернативой для H_0 служит гипотеза:

$$H_1 = \overline{H}_0$$

Означающая, что хотя бы один из коэффициентов отличны от нуля.

Порядок F – теста

Шаг 1. Модель оценивается методом наименьших квадратов и рассчитывается статистика F критерия гипотезы H_0:

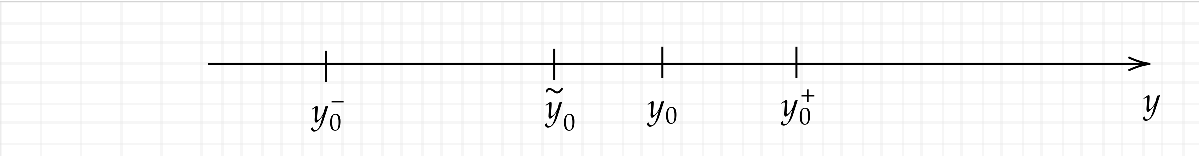
$$F = \frac{R^2 / k}{\left(1 - R^2\right) / (n - (k+1))} \qquad \text{(енто дробь)}$$

Если эта гипотеза верна, то случайная переменная F имеет закон распределения Фишера с кол-ми степеней свободы $k,\ n-k+1$. Если велина F превосходит квантиль распределения Фишера уровня $1-\alpha$, где $\alpha = 0.01\ - 0.05$, то гипотеза H_0 отвергается. Эта квантиль обозначена $F_{\text{крит}}$.

Вывод: F – тест позволяет объективно объяснить качество перменных модели.

41. **Процедура интервального прогнозирования значений эндогенной переменной по оценённой линейной эконометрической модели с гомоскедастичным неавтокоррелированным случайным возмущением.**

оптимальный прогноз $\widetilde{y}_0$.



Помимо точечного прогноза в финансово-экономической сфере прогноз искомой величины y_0 часто строится в виде интервала с левой границей y_0^- и правой границей y_0^+. Этот интервал накрывает неизвестное значение y_0 с заданной доверительной вероятностью. В основании лежит следующая теорема:
Пусть в моделе выполнены все предпосылки теоремы Гаусса-Маркова и случайные возмущения имеют нормальный закон распределения
Тогда следующая дробь:

$$t = \frac{\widetilde{y}_0 - y_0}{\widetilde{\sigma}\left(\triangle \widetilde{y}_0\right)} \sim t(m), \text{ где } m = m - k + 1$$

имеет закон распределения Стьюдента с количеством степеней свододы $n - k + 1$.
Из теоремы выше вытекает равенство:

$$P\left(\left|\frac{\widetilde{y}_0 - y_0}{S\widetilde{y}_0}\right| \leqslant t_{\text{крит}}\right) = 1 - \alpha$$

$t_{\text{крит}}$ имеет значение двухсторонней квантили с кол-ом степеней свободы $n - k + 1$.
Освобождаясь от модуля мы перепишем формулу в следующем виде:

$$P\left(\widetilde{y}_0 - t_{\text{крит}} \cdot \widetilde{\sigma}\left(\triangle \widetilde{y}_0\right) \leqslant y_0 \leqslant \widetilde{y}_0 + t_{\text{крит}} \cdot \widetilde{\sigma}\left(\triangle \widetilde{y}_0\right)\right) = 1 - \alpha$$

$\left[y_0^-,\ y_0^+\right]$ – доверительный интервал

42. **Процедура проверки адекватности оценённой линейной эконометрической модели.**

1. Результаты наблюдений объекта следует разделить на два класса. В первый класс (обучающую выборку) включить основной объем результатов наблюдений 95% выборки. Оставшиеся результаты наблюдений (например, пара (x_0, y_0)) составят контролирующую выборку.

2. По обучающей выборке $(\vec{y}, X)$ оценить модель:

$$\begin{cases} y_t = \quad \tilde{a}_0 \; + \tilde{a}_1 x_t \; + \; u_t, \\ \qquad\quad (S\tilde{a}_0) \quad (S\tilde{a}_1) \quad (\tilde{\sigma}_u) \\ R^2 = \cdots \end{cases} \quad \mathbf{(2)}$$

3. Задаться доверительной вероятностью $(1-\alpha)$ и по значениям регрессоров, входящих в контролирующую выборку, построить доверительные интервалы $[Y_0^-, Y_0^+]$ для соответствующих этим регрессорам значений эндогенной переменной модели.

$$y_0^- = \tilde{y}_0 - t_{\text{крит}} S_{\tilde{y}_0}, \quad y_0^+ = \tilde{y}_0 + t_{\text{крит}} S_{\tilde{y}_0}$$

$$\tilde{y}_0 = \tilde{a}_0 + \tilde{a}_1 x_0$$

$$S_{\tilde{y}_0} = \tilde{\sigma}_u \sqrt{1 + q_0}$$

$$q_0 = \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$t_{\text{крит}}$ – двусторонний $(1-\alpha)$–квантиль распределения Стьюдента с количеством степеней свободы $v_2 = n - (k+1)$, где $(k+1) = 2$ – количество оцениваемых коэффициентов модели **(1)**.

Проверить, попадают ли значения эндогенной переменной из контролирующей выборки в соответствующие доверительные интервалы (в интервал $[y_0^-, y_0^+]$). Если да, то признать оцененную модель адекватной; если нет, то оцененная модель не может быть признана адекватной и подлежит доработке.