

# **АВТОКОРРЕЛЯЦИЯ**

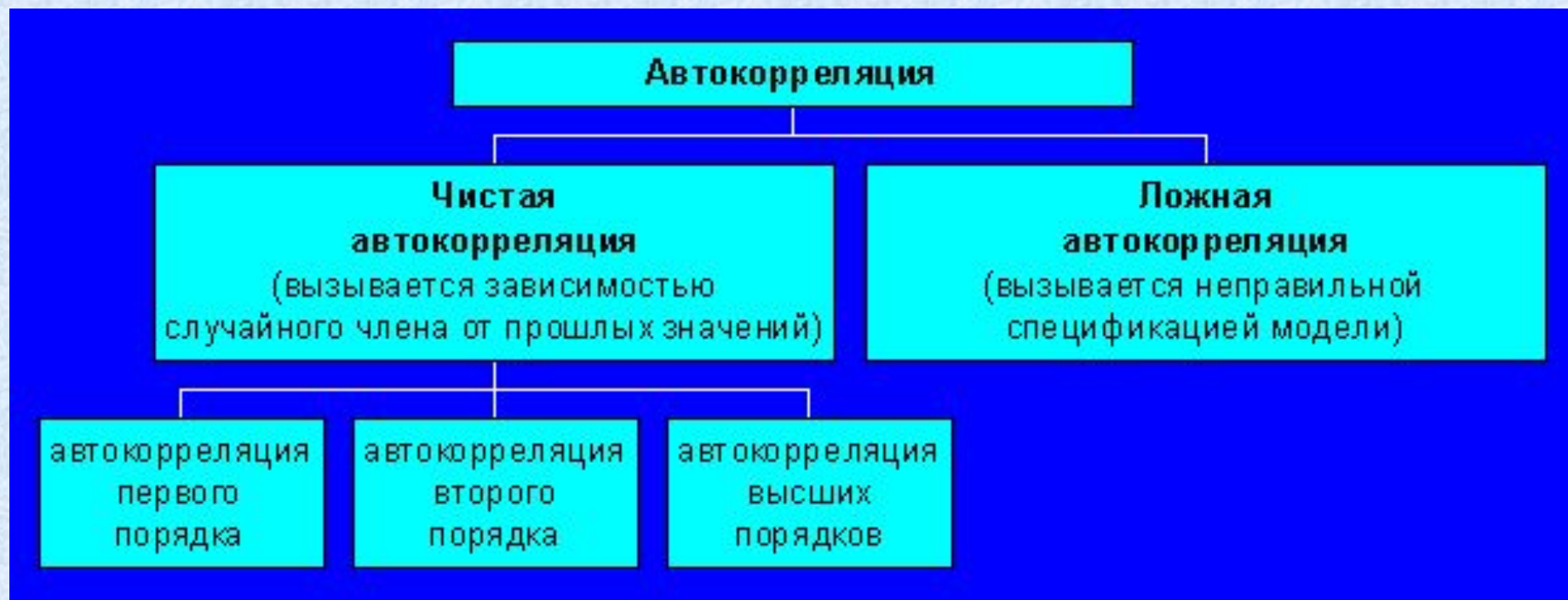
# Определение автокорреляции

**Автокорреляция (последовательная корреляция)** – это корреляция между наблюдаемыми показателями во времени (временные ряды) или в пространстве (перекрестные данные).

**Автокорреляция остатков** характеризуется тем, что не выполняется предпосылка  $3^0$  использования МНК:

$$3^0. \sigma_{\varepsilon_i \varepsilon_j} = \text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} 0, & i \neq j \\ \sigma^2, & i = j \end{cases}$$

# Виды автокорреляции



# Причины чистой автокорреляции

## 1. Инерция.

Трансформация, изменение многих экономических показателей обладает инерционностью.

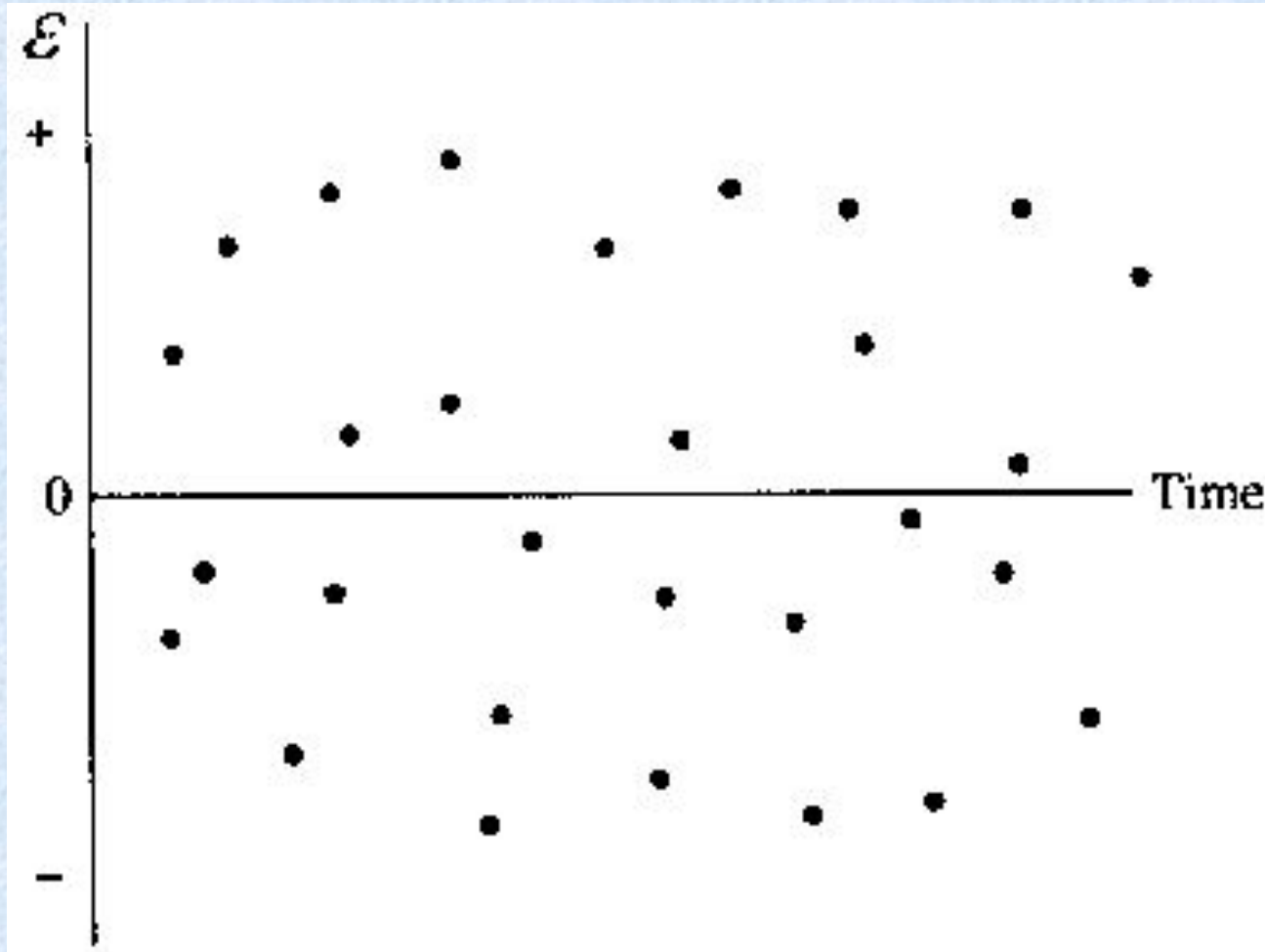
## 2. Эффект паутины.

Многие экономические показатели реагируют на изменение экономических условий с запаздыванием (временным лагом)

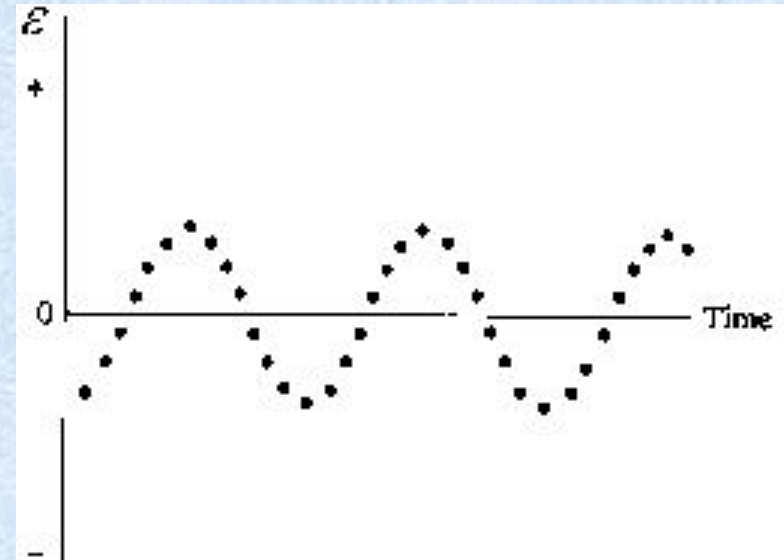
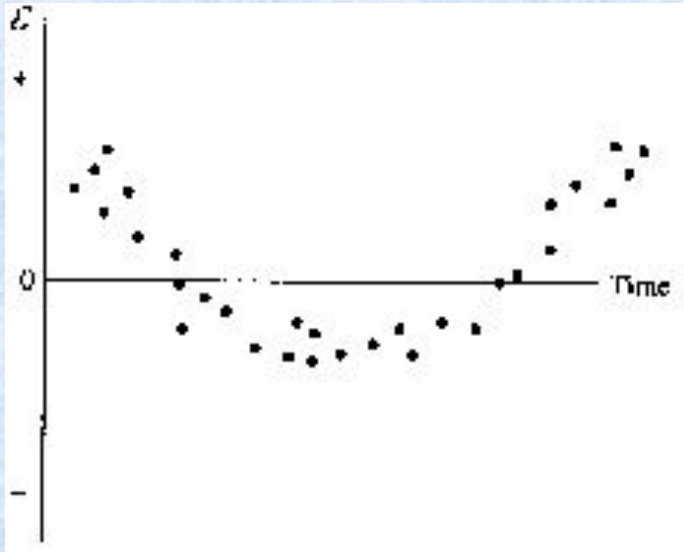
## 3. Сглаживание данных.

Усреднение данных по некоторому продолжительному интервалу времени.

# Классический случайный член $\varepsilon$ (автокорреляция отсутствует)



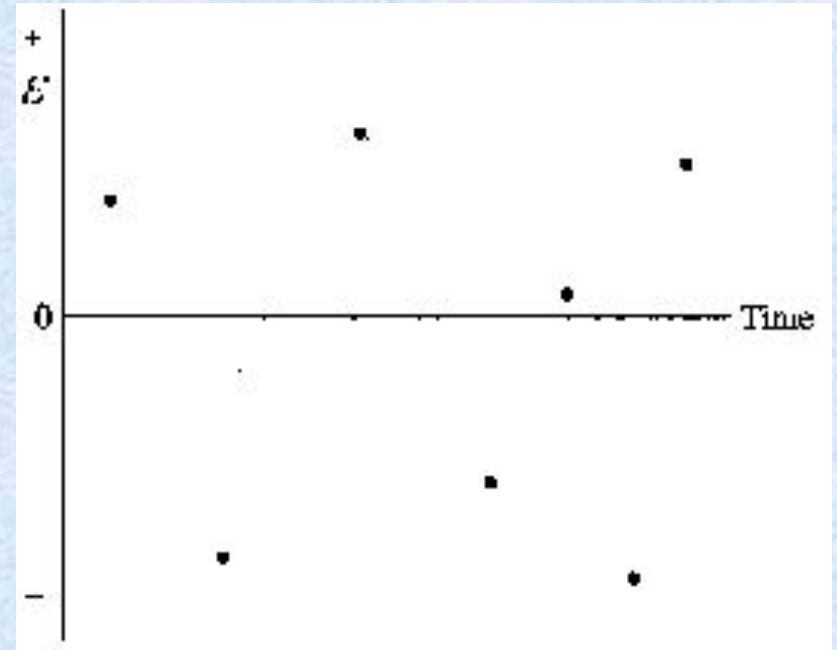
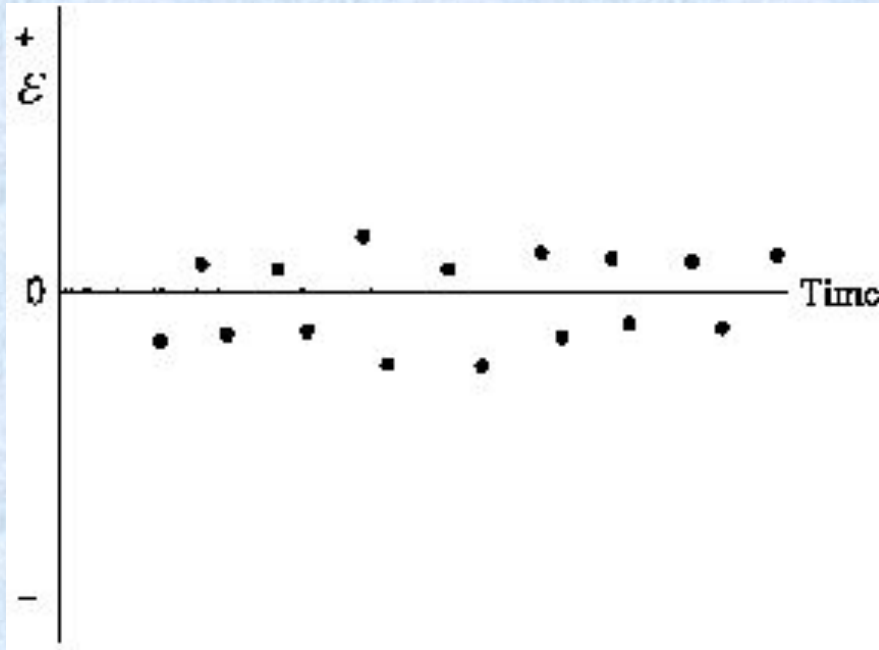
# Положительная автокорреляция



$$\varepsilon_t = \rho \varepsilon_{t-1} + \eta_t \quad \rho > 0$$

Положительная автокорреляция – наиболее важный для экономики случай

# Отрицательная автокорреляция



$$\varepsilon_t = \rho \varepsilon_{t-1} + \eta_t$$

$$\rho < 0$$

# Ложная автокорреляция

(автокорреляция, вызванная ошибочной спецификацией)

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \varepsilon_t$$

$$Y_t = \beta_0 + \beta_1 X_{t1} + \varepsilon_t^*$$

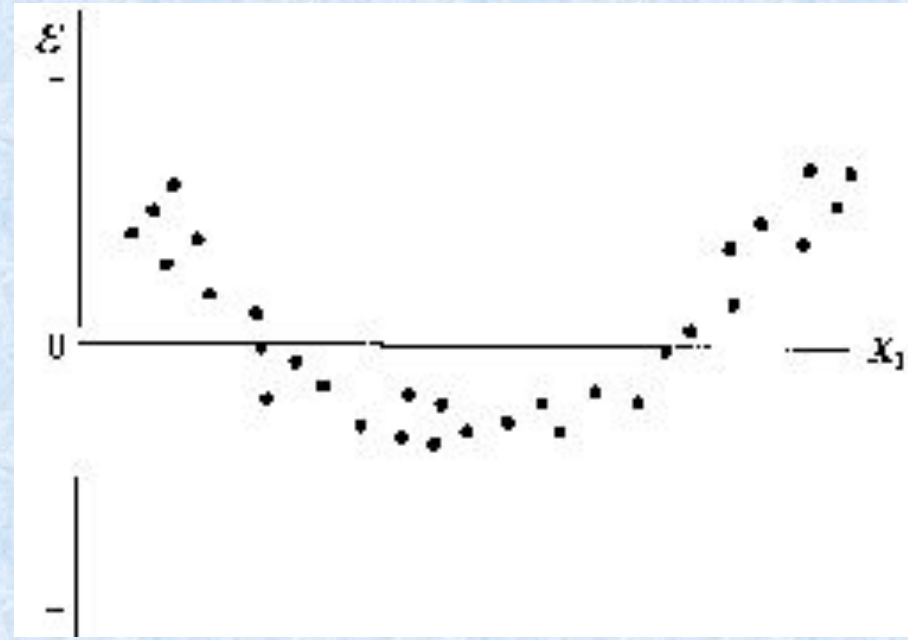
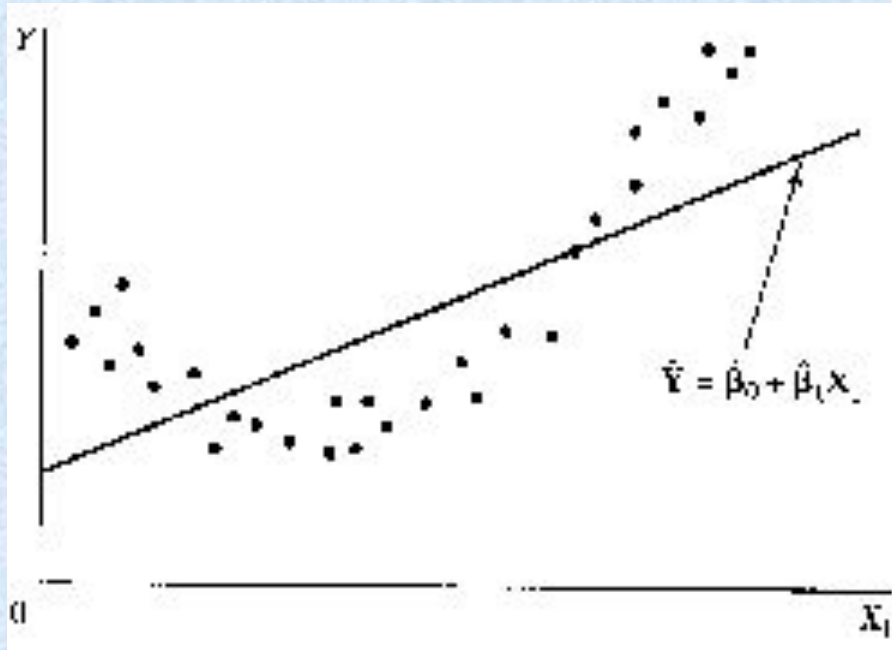
$$\varepsilon_t^* = f(\beta_2 X_{t2} + \varepsilon_t)$$

$X_2$  – сама является автокоррелированной переменной,  
Значение  $\varepsilon$  мало по сравнению с величиной  $\beta_2 \overline{X_2}$



# Ложная автокорреляция как результат неправильного выбора функциональной формы

$$\ln Y_t = \beta_0 + \beta_1 \ln X_{t1} + \varepsilon_t \quad Y_t = \alpha_0 + \alpha_1 X_{t1} + \varepsilon_t^*$$



# Последствия автокорреляции

1. Истинная автокорреляция не приводит к смещению оценок регрессии, но оценки перестают быть эффективными.
2. Автокорреляция (особенно положительная) часто приводит к уменьшению стандартных ошибок коэффициентов, что влечет за собой увеличение  $t$ -статистик.
3. Оценка дисперсии остатков  $S_e^2$  является смещенной оценкой истинного значения  $\sigma_e^2$ , во многих случаях занижая его.
4. В силу вышесказанного выводы по оценке качества коэффициентов и модели в целом, возможно, будут неверными. Это приводит к ухудшению прогнозных качеств модели.

# Обнаружение автокорреляции

1. Графический метод.
2. Метод рядов.
3. Специальные тесты.

# Критерий восходящих и нисходящих серий

Проверяемая гипотеза:

$H_0$ : автокорреляция отсутствует

Последовательность проведения критерия

1. Вычислить остатки
2. Вычислить разницу между соседними остатками,  
 $\Delta_t = e_{t+1} - e_t$
3. Приписать каждой разнице у знак (+/-)
4. Построить ряд знаков

При отсутствии автокорреляции ряд должен носить случайный характер

5. Подсчитать общее количество серий  
(последовательностей постоянного знака) -  $v(n)$
6. Подсчитать длину самой длинной серии -  $t(n)$
7. Сравнить полученные значения с критическими

# Критерий восходящих и нисходящих серий

Проверяемая гипотеза:

$H_0$ : автокорреляция отсутствует

Приблизительный критерий проверки гипотезы на уровне значимости  $\sim 2,5\% \div 5,0\%$  :

При истинности гипотезы должна выполняться система неравенств:

$$\left\{ \begin{array}{l} v(n) > \left[ \frac{1}{3} \cdot (2n - 1) - 1.96 \sqrt{\frac{16n - 29}{90}} \right] \\ \tau(n) < \tau_0(n), \end{array} \right.$$

$n$	$n \leq 26$	$26 < n \leq 153$	$153 < n < 1170$
$\tau_0(n)$	5	6	7

# Обнаружение автокорреляции. Тест Дарбина-Уотсона

Критерий Дарбина-Уотсона предназначен для обнаружения автокорреляции **первого порядка**.

Он основан на анализе остатков уравнения регрессии.

# Тест Дарбина-Уотсона. Ограничения

## Ограничения:

1. Тест не предназначен для обнаружения других видов автокорреляции (более чем первого) и не обнаруживает ее.
2. В модели должен присутствовать свободный член.
3. Данные должны иметь одинаковую периодичность (не должно быть пропусков в наблюдениях).
4. Тест не применим к авторегрессионным моделям, содержащих в качестве объясняющей переменной зависимую переменную с единичным лагом:

$$Y_t = \beta_0 + \sum_{j=1}^m \beta_1 X_{tj} + \gamma Y_{t-1} + \varepsilon_t$$

# Статистика Дарбина-Уотсона

Статистика Дарбина-Уотсона имеет вид:

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$$

$T$  – число наблюдений (обычно временных периодов)

$e_t$  – остатки уравнения регрессии



# Границы для статистики Дарбина-Уотсона

Можно показать, что:  $DW \approx 2(1 - r_{e_t e_{t-1}})$

Отсюда следует:  $0 \leq DW \leq 4$

При положительной корреляции:  $r_{e_t e_{t-1}} \approx 1 \Rightarrow DW \approx 0$

При отрицательной корреляции:  $r_{e_t e_{t-1}} \approx -1 \Rightarrow DW \approx 4$

При отсутствии корреляции:  $r_{e_t e_{t-1}} \approx 0 \Rightarrow DW \approx 2$

# Критические точки распределения Дарбина-Уотсона

Для более точного определения, какое значение  $DW$  свидетельствует об отсутствии автокорреляции, а какое – о ее наличии, построена таблица критических точек распределения Дарбина-Уотсона.

По этой таблице для заданного уровня значимости  $\alpha$ , числа наблюдений  $n$  и количества объясняющих переменных  $m$  определяются два значения:

$d_l$  – нижняя граница,  $d_u$  – верхняя граница

# Критические точки распределения Дарбина-Уотсона

d-статистика Дарбина-Уотсона:  $d_1$  и  $d_u$ , уровень значимости в %

N	K=1		K=2		K=3		K=4		K=5	
	$d_1$	$d_u$	$d_1$	$d_u$	$d_1$	$d_u$	$d_1$	$d_u$	$d_1$	$d_u$
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
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
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78

# Расположение критических точек распределения Дарбина-Уотсона



При положительной корреляции:  $d \rightarrow 0$

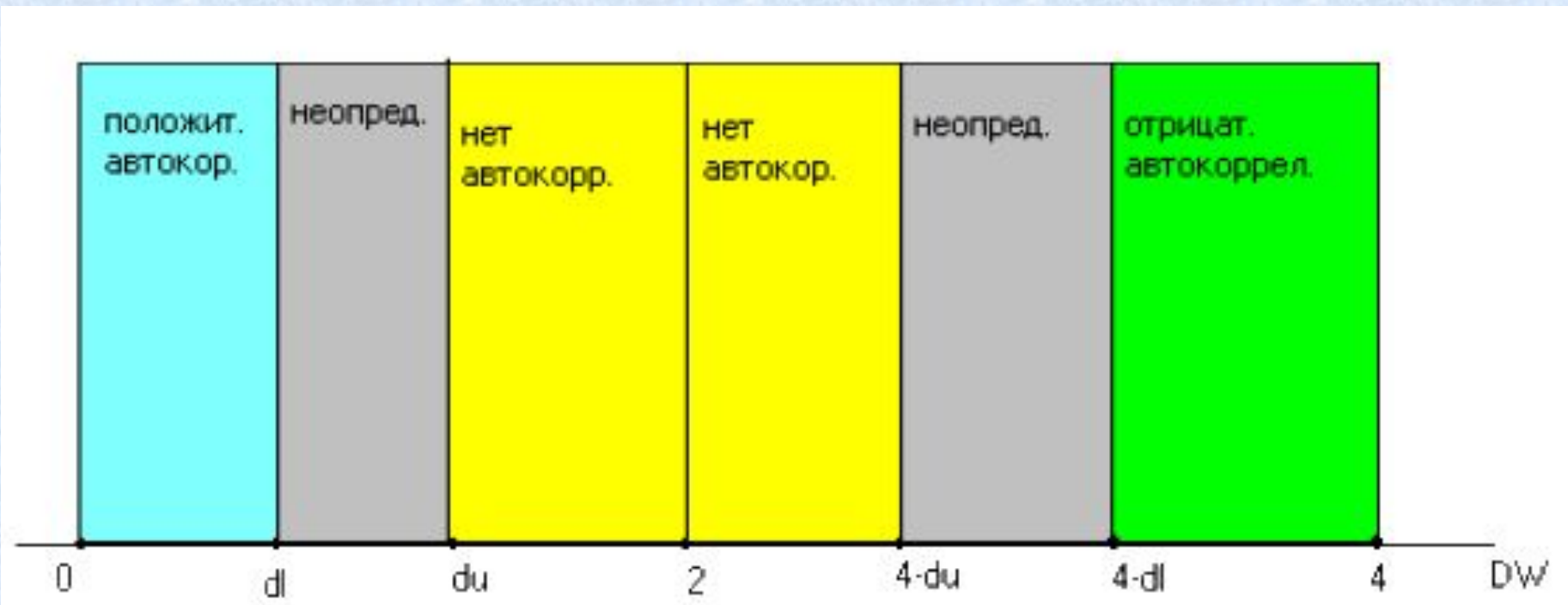
При отрицательной корреляции:  $d \rightarrow 4$

При отсутствии корреляции:  $d \rightarrow 2$

# Практическое использование теста Дарбина-Уотсона

<u>Величина статистики DW</u>	<u>Результат</u>
$4 - d_1 < DW < 4$	В модели с некоторой вероятностью существует отрицательная автокорреляция первого порядка
$4 - d_u < DW < 4 - d_1$	Результат неопределенный.
$2 < DW < 4 - d_u$	В модели с некоторой вероятностью нет автокорреляции первого порядка
$d_u < DW < 2$	В модели с некоторой вероятностью нет автокорреляции первого порядка
$d_1 < DW < d_u$	Результат неопределенный.
$0 < DW < d_1$	В модели с некоторой вероятностью существует положительная автокорреляция первого порядка

# Интерпретация результата теста Дарбина-Уотсона при некотором уровне значимости



# Устранение автокорреляции первого порядка

(при известном коэффициенте автокорреляции)

Пусть имеем:  $Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$      $\varepsilon_t = \rho \varepsilon_{t-1} + \eta_t$   
( $\rho$  – известно)

Процедура устранения автокорреляции остатков:

$$Y_t^* = Y_t - \rho Y_{t-1}, \quad X_t^* = X_t - \rho X_{t-1}, \quad \beta_0^* = \beta_0(1 - \rho)$$

Отсюда:  $Y_t^* = \beta_0^* + \beta_1 X_t^* + \nu_t$

Проблема потери первого наблюдения преодолевается с помощью поправки Прайса-Винстена:

$$Y_1^* = \sqrt{1 - \rho^2} Y_1, \quad X_1^* = \sqrt{1 - \rho^2} X_1$$

# Устранение автокорреляции первого порядка. Обобщения

Рассмотренное авторегрессионное преобразование может быть обобщено на:

- 1) Произвольное число объясняющих переменных
- 2) Преобразования более высоких порядков  $AR(2)$ ,  $AR(3)$  и т.д.:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \eta_t \quad \varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \rho_3 \varepsilon_{t-3} + \eta_t$$

Однако на практике значения коэффициента автокорреляции  $\rho$  обычно неизвестны и его необходимо оценить. Существует несколько методов оценивания.



# Способы оценивания коэффициента автокорреляции $\rho$

1. На основе статистики Дарбина-Уотсона.
2. Процедура Кохрейна-Оркатта.
3. Процедура Хилдрета-Лу.
4. Процедура Дарбина
5. Метод первых разностей.

# Определение коэффициента $\rho$ на основе статистики Дарбина-Уотсона

$$\left\{ \begin{array}{l} DW \approx 2(1 - r_{e_t e_{t-1}}) \\ \hat{\rho} = r_{e_t e_{t-1}} \end{array} \right. \Rightarrow \hat{\rho} \approx 1 - \frac{DW}{2}$$

Этот метод дает удовлетворительные результаты при большом числе наблюдений.

# Итеративная процедура Кохрейна-Оркатта (на примере парной регрессии)

1. Определение уравнения регрессии и вектора остатков:

$$(*) \quad \hat{Y}_t = b_0 + b_1 X_t \Rightarrow b_0, b_1, e_i, i = \overline{1, n}$$

2. В качестве приближенного значения  $\rho$  берется его МНК-оценка:

$$e_t = \rho^* e_{t-1} + v_t$$

3. Для найденного  $\rho^*$  оцениваются коэффициенты  $\alpha_0$   $\alpha_1$ :

$$(Y_t - \rho^* Y_{t-1}) = \alpha_0 (1 - \rho^*) + \alpha_1 (X_t - \rho^* X_{t-1}) + (e_t - \rho^* e_{t-1}) + \eta_t$$

4. Подставляем  $b_0 = \alpha_0 (1 - \rho^*)$ ,  $b_1 = \alpha_1$  в (\*) и вычисляем  $e_i, i = \overline{1, n}$   
Возвращаемся к этапу 2.

Критерий остановки: разность между текущей и предыдущей оценками  $\rho^*$  стала меньше заданной точности.

# Итеративная процедура Хилдрета-Лу (поиск по сетке)

1. Определение уравнения регрессии и вектора остатков:

$$\hat{Y}_t = b_0 + b_1 X_t \quad \Rightarrow \quad b_0, b_1, e_i, i = \overline{1, n}$$

2. Оцениваем регрессию

$(Y_t - \rho^* Y_{t-1}) = \alpha_0(1 - \rho) + \alpha_1(X_t - \rho X_{t-1}) + (e_t - \rho e_{t-1}) + \eta_t$   
для каждого возможного значения  $\rho \in [-1, 1]$  с некоторым достаточно малым шагом, например 0,001; 0,01 и т.д.

3. Величина  $\rho^*$ , обеспечивающая минимум стандартной ошибки регрессии принимается в качестве оценки автокорреляции остатков.

# Итеративные процедуры оценивания коэффициента $\rho$ .

## Выводы

1. Сходимость процедур достаточно хорошая.
2. Метод Кохрейна-Оркатта может «попасть» в локальный (а не глобальный) минимум.
3. Время работы процедуры Хилдрета-Лу значительно сокращается при наличии априорной информации об области возможных значений  $\rho$ .

# Процедура Дарбина

(на примере парной регрессии)

Пусть имеет место автокорреляция остатков:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \quad \Rightarrow \quad Y_t = \beta_0 + \beta_1 X_t + \rho \varepsilon_{t-1} + \eta_t$$
$$\varepsilon_t = \rho \varepsilon_{t-1} + \eta_t$$

$$Y_{t-1} = \beta_0 + \beta_1 X_{t-1} + \varepsilon_{t-1} \quad \Rightarrow \quad \rho Y_{t-1} = \rho \beta_0 + \rho \beta_1 X_{t-1} + \rho \varepsilon_{t-1}$$

$$Y_t - \rho Y_{t-1} = \beta_0(1 - \rho) + \beta_1(X_t - \rho X_{t-1}) + \eta_t$$

$$Y_t = \beta_0(1 - \rho) + \beta_1 X_t - \rho \beta_1 X_{t-1} + \rho Y_{t-1} + \eta_t$$

# Процедура Дарбина

(на примере парной регрессии)

Процедура Дарбина представляет собой традиционный МНК с нелинейными ограничениями типа равенств:

$$Y_t = \beta_0^* + \beta_1^* X_t + \beta_2^* X_{t-1} + \rho Y_{t-1} + \eta_t$$

$$\beta_0^* = \beta_0(1 - \rho); \quad \beta_2^* = -\rho\beta_1^*$$

## Способы решения:

1. Решать задачу нелинейного программирования.
2. Двухшаговый МНК Дарбина (полученный коэффициент автокорреляции используется в поправке Прайса-Винстена).
3. Итеративная процедура расчета.

# Итеративная процедура метода Дарбина

1. Считается регрессия и находятся остатки.
2. По остаткам находят оценку коэффициента автокорреляции остатков.
3. Оценка коэффициента автокорреляции используется для пересчета данных и цикл повторяется.

Процесс останавливается, как только обеспечивается достаточная точность (результаты перестают существенно улучшаться).



# Обобщенный метод наименьших квадратов. Замечания

Не следует применять обобщенный МНК автоматически

1. Значимый коэффициент  $DW$  может указывать просто на ошибочную спецификацию.
2. Последствия автокорреляции остатков иногда бывают незначительными.
3. Качество оценок может снизиться из-за уменьшения числа степеней свободы (нужно оценивать дополнительный параметр).
4. Значительно возрастает трудоемкость расчетов.

Конец лекции