

# Справочные материалы

Множественная регрессия,  
тестирование предпосылок

# Пример результатов регрессионного анализа

	A	B	C	D	E	F	G	H	I	J
1	ВЫВОД ИТОГОВ									
2										
3	<i>Регрессионная статистика</i>									
4	Множественный R	0,879224018								
5	R-квадрат	0,773034874								
6	Нормированный R-квadra	0,740611285								
7	Стандартная ошибка	0,735114383								
8	Наблюдения	25								
9										
10	Дисперсионный анализ									
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>				
12	Регрессия	3	38,65174372	12,88391	23,8417427	5,82374E-07				
13	Остаток	21	11,34825628	0,540393						
14	Итого	24	50							
15										
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>			
17	Y-пересечение	-0,684816194	1,113384882	-0,61508	0,54511279	3,0002268	1,630594			
18	Использованный капитал, млрд. долл. США	0,0195467	0,01114282	1,754197	0,09397987	-0,003626062	0			
19	Численность служащих, тыс. чел.	0,006025897	0,001461831	4,122136	0,00048519	0,002985852	0			
20	Рыночная капитализация компании, млрд. долл. США	0,03648291	0,034520894	1,056836	0,30259326	-0,035307214	0,108273			
21										

Число степеней свободы 1 для F критерия

Число степеней свободы 2 для F-критерия  
Число степеней свободы для критерия Стьюдента (t-статистики)

Остаточная  
сумма  
квадратов

# Поиск табличных значений критериев F и t средствами Excel

Аргументы функции

FRASПОБР

Вероятность	0,05	= 0,05
Степени_свободы1	3	= 3
Степени_свободы2	21	= 21

= 3,072467001

Возвращает обратное значение для F-распределения вероятностей: если  $p = \text{FRASП}(x, \dots)$ , то  $\text{FRASПОБР}(p, \dots) = x$ .

**Вероятность** вероятность, связанная с F-интегральным распределением, число в диапазоне от 0 до 1 включительно.

Значение: 3,072467001

[Справка по этой функции](#)

OK Отмена

Аргументы функции

СТЬЮДРАСПОБР

Вероятность	0,05	= 0,05
Степени_свободы	21	= 21

= 2,079613837

Возвращает обратное распределение Стьюдента.

**Степени\_свободы** положительное целое число степеней свободы, характеризующее распределение.

Значение: 2,079613837

[Справка по этой функции](#)

OK Отмена

# Алгоритм применения теста Голдфелда-Квандта

Государственные расходы на образование в  
различных странах

Страна	Рас- ходы	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

Применение ф-ии «ЛИНЕЙН»

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<sub>кр</sub>=3.0

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

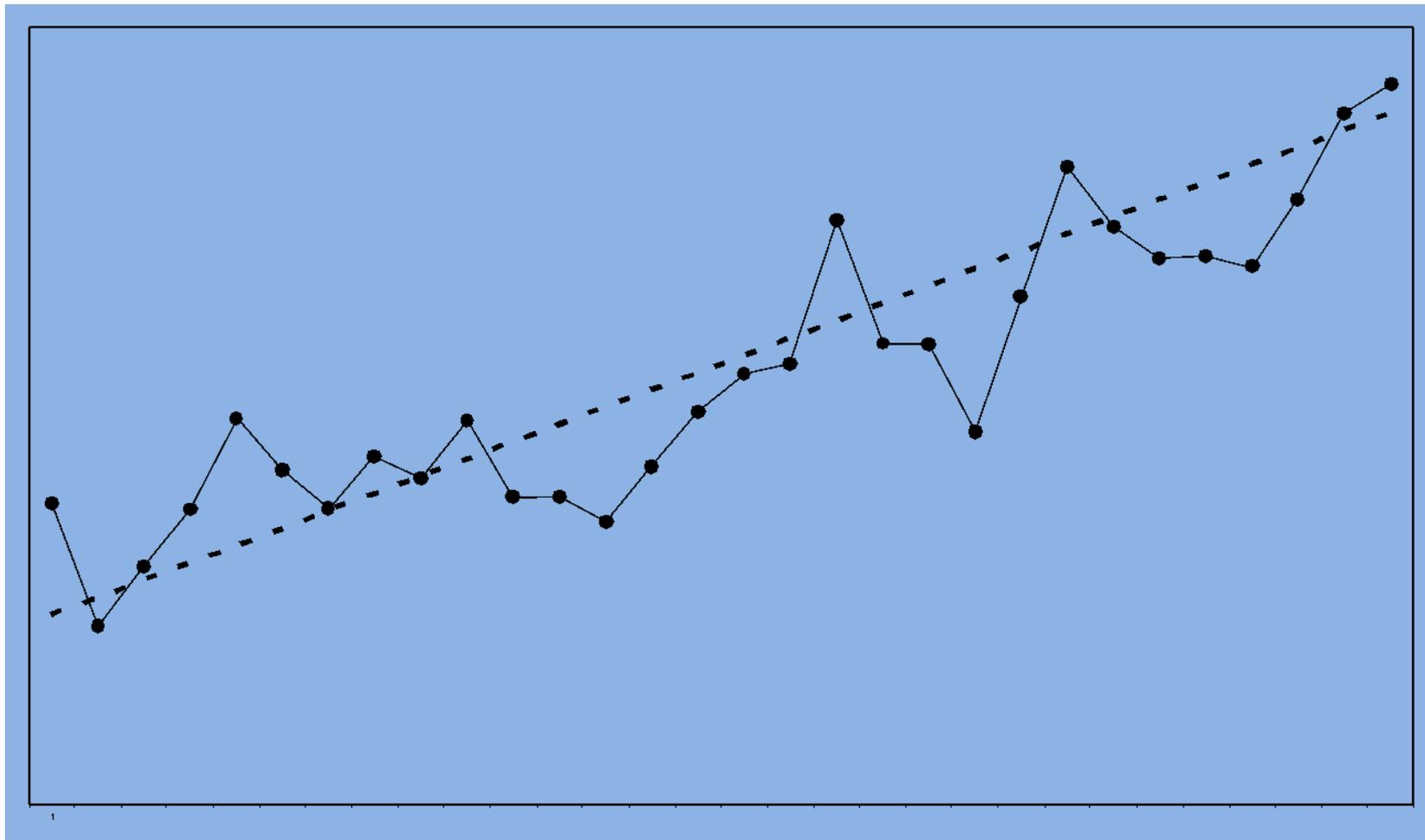
Модель гетероскедастична

Модель: (по всей выборке)

$$Y = -2.32 + 0.067X \quad (10.4)$$

# Автокорреляция анализ графика остатков

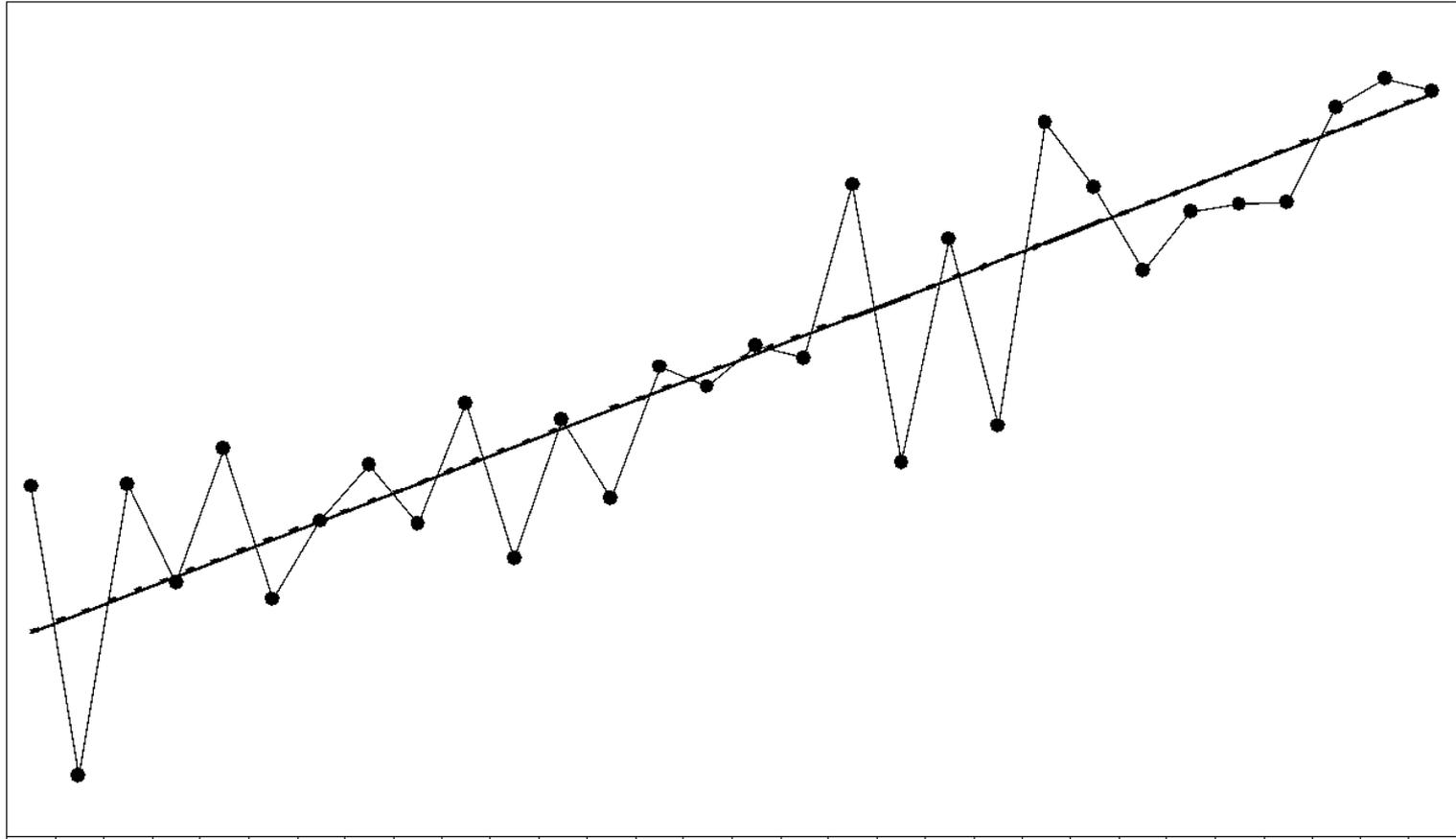
Диаграмма рассеяния с положительной автокорреляцией



Признак – чередование зон с повышенными и заниженными значениями по отношению к тренду

# Автокорреляция анализ графика остатков

Пример отрицательной автокорреляции случайных возмущений



Признак – наблюдения действуют друг на друга по принципу «маятника»

# Тест Дарбина-Уотсона

## 1. Предпосылки теста

Случайные возмущения распределены по нормальному закону

Имеет место авторегрессия первого порядка:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

$$M(\varepsilon_t) = 0 \quad \sigma^2(\varepsilon_t) = \sigma_0$$

2. Статистика для проверки гипотезы:

$$DW = \frac{\sum (u_i - u_{i-1})^2}{\sum (u_i^2)}$$

# Тест Дарбина-Уотсона

Особенности статистики DW:

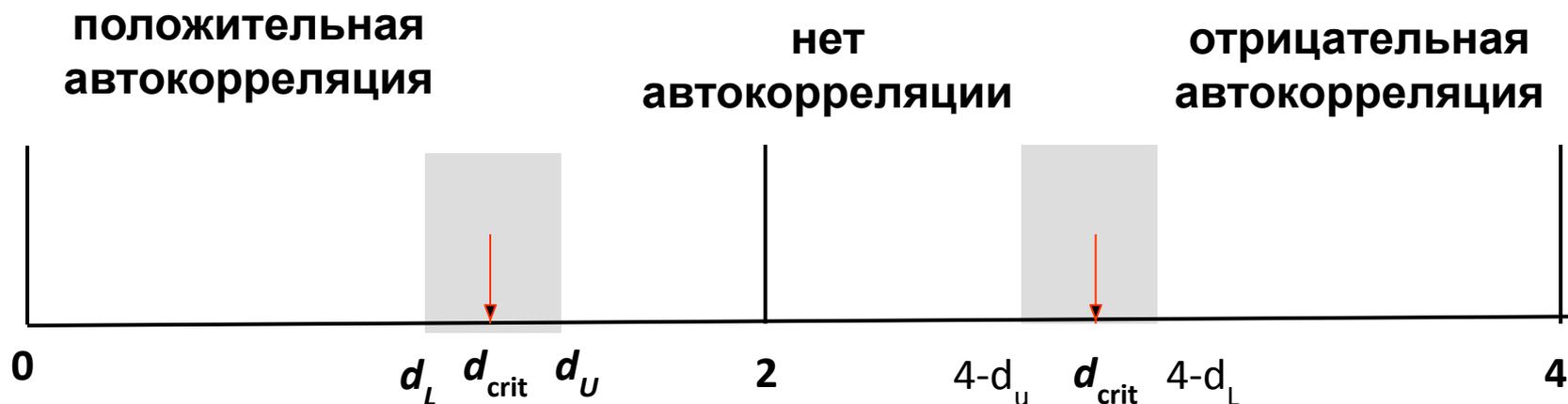
Для статистики DW не возможно найти критическое значение, т.к. оно зависит не только от  $R_{\text{дов}}$  и степеней свободы  $k$  и  $n$ , но и от абсолютных значений регрессоров

Возможно определить границы интервала  $D_L$  и  $D_U$  внутри которого критическое значение  $DW_{\text{кр}}$  находится:

$$D_L \leq DW_{\text{кр}} \leq D_U$$

Значения  $D_U$  и  $D_L$  находятся по таблицам

# Тест Дарбина-Уотсона



Нет автокорреляции  $DW \rightarrow 2$

Положительная автокорреляция  $DW \rightarrow 0$

Отрицательная автокорреляция  $DW \rightarrow 4$

Интервалы  $(D_L, D_U)$  и  $(4-D_L, 4-D_U)$  зоны неопределенности

# Значения $d_L$ , $d_U$ для выборки объёмом $n$ и числа факторов $k$

$n$	$k^1 = 1$		$k^1 = 2$		$k^1 = 3$		$k^1 = 4$		$k^1 = 5$	
	$d_L$	$d_U$								
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

# Постановка задачи

Анализируются продажи некоторого товара в нескольких магазинах. Предполагается, что объем продаж зависит от цены товара и расходов на рекламу.

## Задани

- Построить уравнение зависимости объема продаж от определяющих факторов, выбрав в качестве формы спецификации линейную. Сделать вывод о качестве модели.
- Оценить остатки модели. Сделать вывод о соответствии предпосылкам теоремы Гаусса-Маркова.

## План работы

- Используя метод включения или исключения проверить существенность факторов (изменение остаточной суммы квадратов и коэффициента детерминации)
- Проверить линейную независимость факторов (Инструмент корреляция надстройки Анализа данных)
- Оценить параметры уравнения регрессии (инструмент Регрессия надстройки Анализ данных)
- Построить график зависимости остатков от номера наблюдения, сделать предположение о гомоскедастичности и наличии/отсутствии автокорреляции остатков
- Проверить гомоскедастичность остатков при помощи теста Голдфельда-Квандта, проверить автокорреляцию остатков с помощью теста Дарбина-Уотсона.