Тестирование автокорреляции

Понятие автокорреляции

Модель называется автокоррелированной, если не выполняется третья предпосылка теоремы Гаусса-Маркова: $Cov(u_i, u_j) \neq 0$ при $i \neq j$.

Автокорреляция чаще всего появляется в моделях временных рядов и моделировании циклических процессов.

Причина – неправильный выбор спецификации модели.

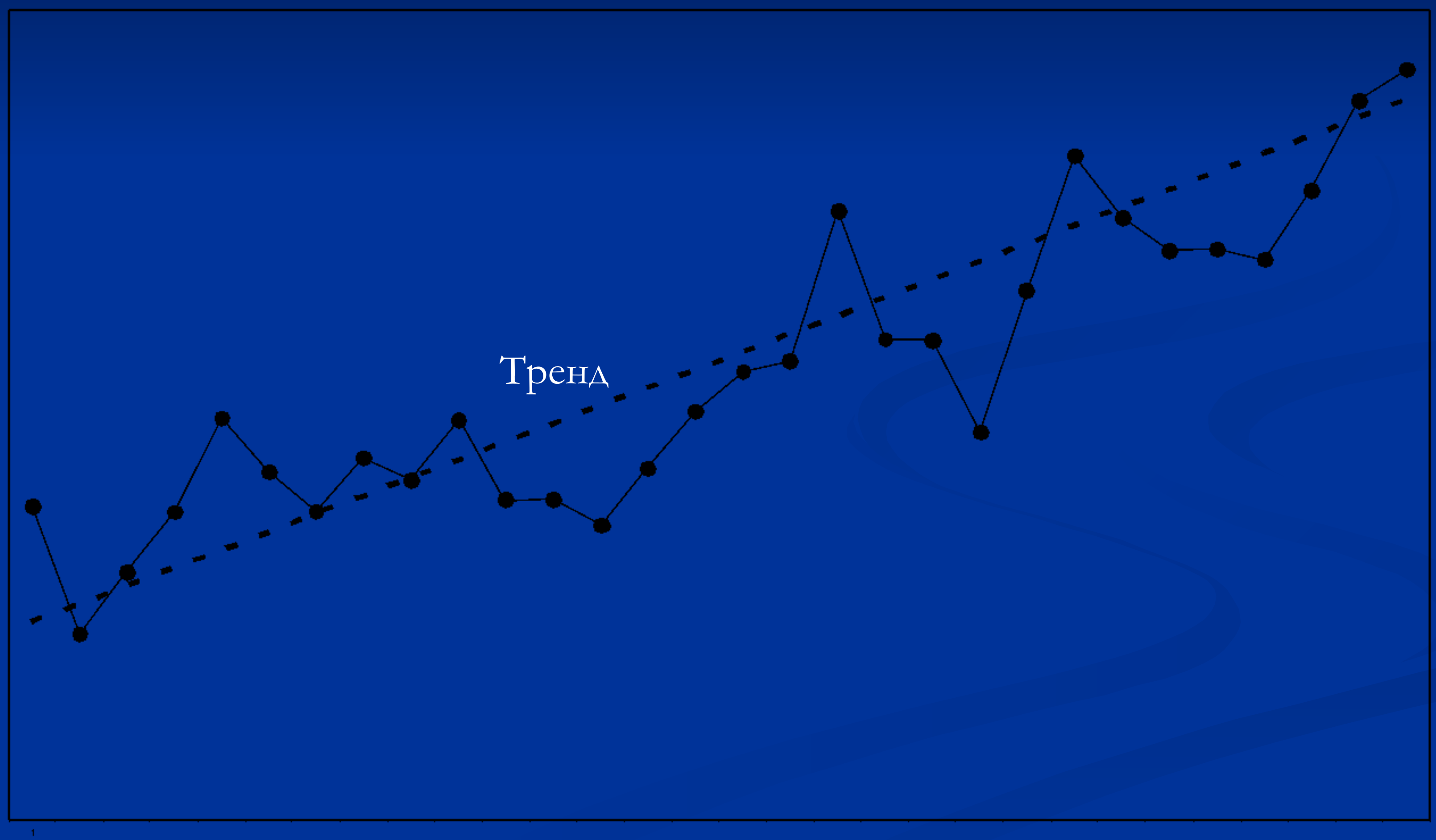
Последствия автокорреляции.

- оценки коэффициентов теряют эффективность;

- стандартные ошибки коэффициентов занижены.

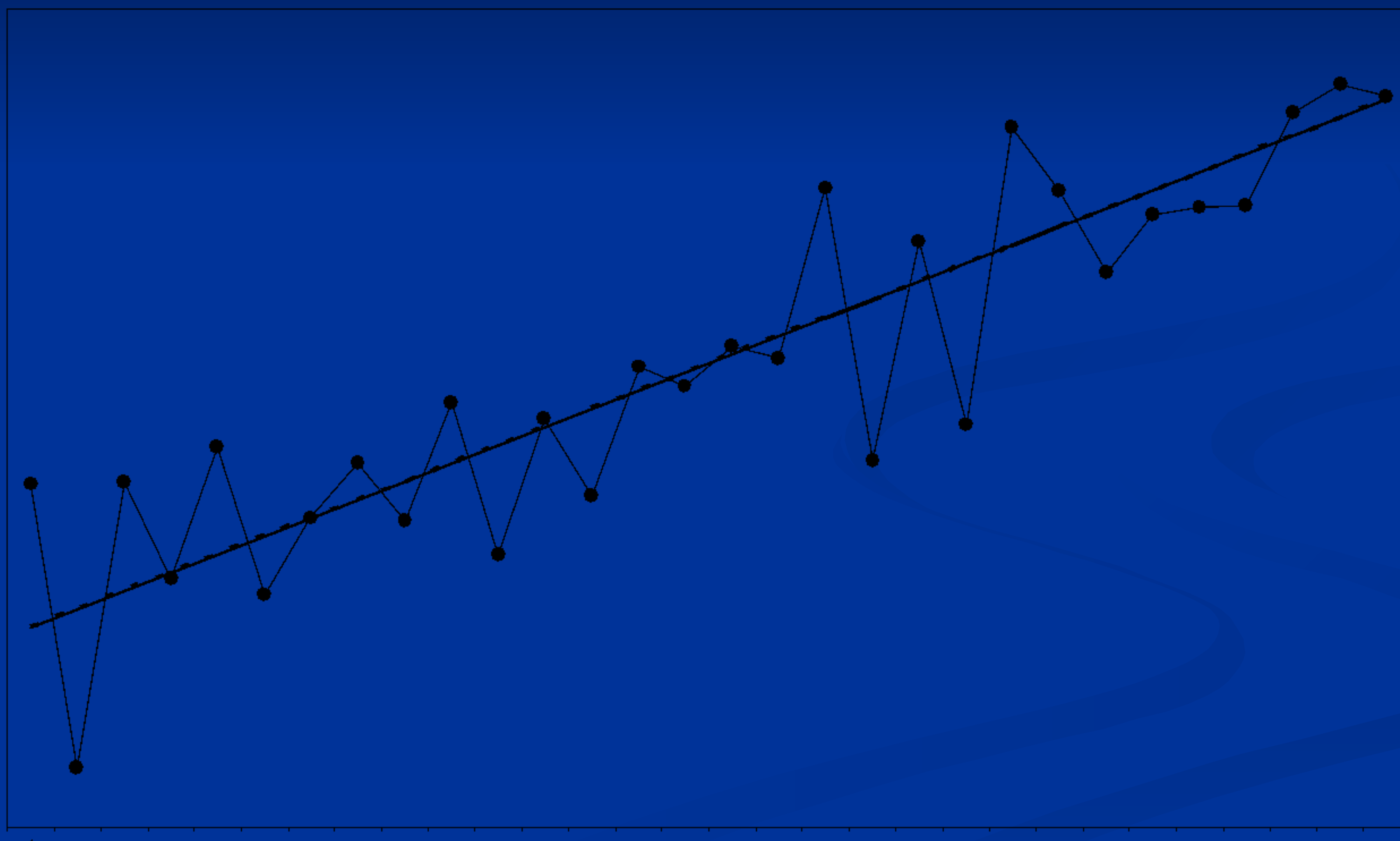
Понятие автокорреляции

Диаграмма рассеяния с положительной автокорреляцией.



Понятие автокорреляции

Пример отрицательной автокорреляции случайных возмущений.



Типы автокорреляции

Рассматриваем модель парной регрессии.

$$y_t = a_0 + a_1 x_t + u_t$$

Авторегрессия 1-го порядка : AR(1)

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

Авторегрессия 5-го порядка : AR(5)

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \rho_4 u_{t-4} + \rho_5 u_{t-5} + \varepsilon_t$$

Авторкорреляция скользящих средних 3-го порядка:

$$u_t = \lambda_0 \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \lambda_2 \varepsilon_{t-2} + \lambda_3 \varepsilon_{t-3}$$

Тест Дарбина-Уотсона

1. Предпосылки теста.

Случайные возмущения распределены по нормальному закону.

Имеет место авторегрессия первого порядка:

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

$M(\varepsilon_t)=0$; $\sigma 2(\varepsilon_t)=Const$

2. Статистика для проверки гипотезы:

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

Тест Дарбина-Уотсона

3. Свойства статистики DW.

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

$$=2\left(1-\frac{\sum(u_i-\bar{u})(u_{i-1}-\bar{u})}{\sum u_i^2}\right)=2\left(\frac{Cov(u_i,u_{i-1})}{\boldsymbol{\sigma}(u_i)\boldsymbol{\sigma}(u_i)}\right)\approx 2(1-r)$$

где: **r**- коэффициент корреляции между случайными возмущениями.

Из этого выражения следует:

DW изменятся в пределах (0 – 4).

При этом если r = 1, DW=0- положительная корреляция;

если r = 0, DW=2-; отсутствие корреляции;

если r=-1, DW=4- отрицательная корреляция.

Тест Дарбина-Уотсона

Для статистики DW не возможно найти критическое значение, т.к. оно зависит не только от $P_{дов}$ и степеней свободы k и n-1, но и от абсолютных значений регрессоров.

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

$$\mathbf{D_L \le DW_{кр} \le D_u}$$

Значения D_u и D_L находятся по таблицам.

Тест Дарбина-Уотсона

Нет автокорреляции $DW \to 2$

Положительная автокорреляция $DW \to 0$

Отрицательная автокорреляция $DW \to 4$

Интервалы (D_L, D_u) и (4-D_L, 4-D_u) зоны неопределенности.

Тестирование автокорреляции

Государственные расходы на образование в различных странах

Расхо-ды (Y)	ВВП (X)	ỹ	U=Y-ỹ	U_i-U_{i-1}	$(U_i-U_{i-1})^2$	Расхо-ды (Y)	ВВП (X)	ỹ	U=Y-ỹ	U_i-U_{i-1}	$(U_i-U_{i-1})^2$
0,34	5,67	-1,9	2,28			5,31	101,65	4,48	0,83		
0,22	10,13	-1,6	1,86	-0,42	0,18	6,4	115,97	5,44	0,96	-0,13	0,02
0,32	11,34	-1,6	1,88	0,02	0,00	7,15	119,49	5,67	1,48	-0,51	0,26
1,23	18,88	-1,1	2,29	0,41	0,16	11,22	124,15	5,98	5,24	-3,76	14,12
1,81	20,94	-0,9	2,73	0,44	0,20	8,66	140,98	7,11	1,55	3,69	13,58
1,02	22,16	-0,8	1,86	-0,87	0,76	5,56	153,85	7,97	-2,41	3,96	15,69
1,27	23,83	-0,7	2,00	0,14	0,02	13,41	169,38	9,01	4,40	-6,81	46,39
1,07	24,67	-0,7	1,74	-0,26	0,07	5,46	186,33	10,14	-4,68	9,08	82,52
0,67	27,56	-0,5	1,15	-0,59	0,35	4,79	211,78	11,85	-7,06	2,37	5,63
1,25	27,57	-0,5	1,73	0,58	0,34	8,92	249,72	14,38	-5,46	-1,59	2,53
0,75	40,15	0,4	0,38	-1,34	1,80	18,9	261,41	15,17	3,73	-9,20	84,60
2,8	51,62	1,1	1,67	1,28	1,65	15,95	395,52	24,14	-8,19	11,92	142,10
4,9	57,71	1,5	3,36	1,69	2,87	29,9	534,97	33,46	-3,56	-4,62	21,36
3,5	63,03	1,9	1,60	-1,76	3,08	33,59	655,29	41,51	-7,92	4,36	18,99
4,45	66,32	2,1	2,33	0,73	0,53	38,62	815	52,20	-13,58	5,65	31,96
1,6	66,97	2,2	-0,56	-2,89	8,37	61,61	1040,5	67,28	-5,67	-7,91	62,56
4,26	76,88	2,8	1,44	2,00	3,99	181,3	2586,4	170,69	10,61	-16,28	265,03

Тестирование автокорреляции

Модель: $Y=-2.32 + 0.669X + U$

$(0.9) \quad (0.002)$

$ESS=\Sigma U_i^2=710.34$

$\Sigma(U_i-U_{i-1})^2 = 832.4$

$DW = 832.4/710.3=1.17$

Границы интервала – $d_L=1.35$; $d_u=1.49$

$DW< d_L$

Вывод: модель автокоррелирована

Тестирование автокорреляции

Относительные расходы на образование в различных странах

Y/ВВП	1/ВВП	Y*пр	U	U_i-U_{i-1}	$(U_i-U_{i-1})^2$	Y/ВВП	1/ВВП	Y*пр	U	U_i-U_{i-1}	$(U_i-U_{i-1})^2$
0,060	0,1764	0,042	0,0183			0,052	0,0098	0,05256	-0,00032	-0,00338	0,00001
0,022	0,0987	0,047	-0,0250	-0,0433	0,00188	0,055	0,0086	0,05264	0,00255	0,00287	0,00001
0,028	0,0882	0,047	-0,0192	0,0058	0,00003	0,060	0,0084	0,05265	0,00718	0,00463	0,00002
0,065	0,0530	0,050	0,0154	0,0346	0,00120	0,090	0,0081	0,05268	0,03770	0,03052	0,00093
0,086	0,0478	0,050	0,0364	0,0209	0,00044	0,061	0,0071	0,05274	0,00869	-0,02901	0,00084
0,046	0,0451	0,050	-0,0042	-0,0406	0,00165	0,036	0,0065	0,05278	-0,01664	-0,02533	0,00064
0,053	0,0420	0,050	0,0028	0,0071	0,00005	0,079	0,0059	0,05282	0,02635	0,04299	0,00185
0,043	0,0405	0,051	-0,0072	-0,0100	0,00010	0,029	0,0054	0,05285	-0,02355	-0,04990	0,00249
0,024	0,0363	0,051	-0,0265	-0,0193	0,00037	0,023	0,0047	0,05289	-0,03028	-0,00673	0,00005
0,045	0,0363	0,051	-0,0055	0,0210	0,00044	0,036	0,0040	0,05294	-0,01722	0,01306	0,00017
0,019	0,0249	0,052	-0,0329	-0,0274	0,00075	0,072	0,0038	0,05295	0,01935	0,03657	0,00134
0,054	0,0194	0,052	0,0023	0,0352	0,00124	0,040	0,0025	0,05304	-0,01271	-0,03206	0,00103
0,085	0,0173	0,052	0,0328	0,0305	0,00093	0,056	0,0019	0,05308	0,00281	0,01552	0,00024
0,056	0,0159	0,052	0,0034	-0,0295	0,00087	0,051	0,0015	0,05310	-0,00184	-0,00465	0,00002
0,067	0,0151	0,052	0,0149	0,0115	0,00013	0,047	0,0012	0,05312	-0,00574	-0,00389	0,00002
0,024	0,0149	0,052	-0,0283	-0,0432	0,00187	0,059	0,0010	0,05314	0,00607	0,01181	0,00014
0,055	0,0130	0,052	0,0031	0,0314	0,00099	0,070	0,0004	0,05318	0,01692	0,01084	0,00012

Тестирование автокорреляции

Модель: $0.0530 - 0.66X + U$

$(0.004) \quad (0.1)$

$ESS = \Sigma U_i^2 = 0.012$

$\Sigma(U_i - U_{i-1})^2 = 0.0229$

$DW = 0.0229/0.012 = 1.79$

Границы интервала – $d_L = 1.35$; $d_u = 1.49$

$d_L < DW < d_u$

Вывод: модель неавтокоррелирована

Метод исправления автокорреляции

Рассматривается случай авторегрессии первого порядка:

$$Y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + U_t$$

$$U_t = \rho U_{t-1} + \varepsilon_t$$

При этом:

$$M(\varepsilon_t) = 0 \qquad \sigma^2(\varepsilon_t) = \sigma^2_t \qquad |\rho| < 1$$

Тогда:

$$\sigma^2(U_t) = \rho^2 \sigma^2(U_{t-1}) + \sigma^2_t + 2Cov(\rho, U_{t-1})$$

$Cov(\rho, U_{t-1}) = 0$, т.к. $\rho = Const$

Следовательно

$$\sigma^2(U_t) = \rho^2 \sigma^2(U_{t-1}) + \sigma^2_t \qquad (10.1)$$

Метод исправления автокорреляции

Т.к. U0 отсутствует, полагаем, что $\sigma^2(U_1) = \sigma^2(U_0)$
Тогда из (10.1) следует:

$$\sigma_{u_0}^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad \textbf{(10.2)}$$

Множитель $(1-\rho^2)$ обеспечивает стационарность $\sigma^2(U_t)$, т.е. постоянство $\sigma^2(Ut)$

Выражение (10.2) – начальное условие для $\sigma^2(U_0)$

Из выражения (10.1) с учетом (10.2) вытекает:

$$\sigma_{u_2}^2 = \rho^2 \sigma_{u_1}^2 + \sigma_\varepsilon^2 = \rho^2 \frac{\sigma_\varepsilon^2}{1-\rho^2} + \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

Метод исправления автокорреляции

Для произвольного наблюдения t в силу рекурентности (10.1) имеем:

$$\sigma_{u_t}^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad \textbf{(10.3)}$$

Вывод: введение корректирующего множителя **(1-ρ^2)** обеспечивает постоянство σ^2(U) во всех наблюдениях и, следовательно, отсутствие автокорреляции между случайными возмущениями.

Метод устранения автокорреляции

Рассмотрим два последовательных уравнения наблюдения

$$y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + U_t \qquad \textbf{(10.4)}$$

$$y_{t-1} = a_0 + a_1 x_{1t-1} + a_2 x_{2t-1} + U_{t-1} \qquad \textbf{(10.5)}$$

Умножим уравнение (10.5) на ρ и вычтем из (10.4)

$$y_t - \rho y_{t-1} = a_0(1-\rho) + a_1(x_{1t} - \rho x_{2t-1}) + a_2(x_{2t} - \rho x_{2t-1}) + U_t - \rho U_{t-1}$$

Учитывая, что $\mathbf{U_t - \rho U_{t-1} = \varepsilon_t}$ и делая замену переменных

$$y_t^* = y_t - \rho y_{t-1}; \quad b_0 = a_0(1-\rho); \quad z_{1t} = (x_{1t} - \rho x_{2t-1}); \quad z_{2t} = (x_{2t} - \rho x_{2t-1})$$

получим систему уравнений, в которых дисперсия случайных возмущений постоянна.

$$y_t^* = b_0 + a_1 z_{1t} + a_2 z_{2t} + \boldsymbol{\varepsilon}_t \qquad \textbf{(10.6)}$$

Метод устранения автокорреляции

Параметры уравнения (10.6) можно оценить с помощью МНК. Если значение ρ известно, то решение окончено.

Замечание. Уравнения (10.6) имеют смысл при t=2, т.к. при t=1 оно не может быть получено.

Для включения первого уравнения наблюдений в систему (10.6) его умножают на $(1-\rho)^{1/2}$.

Этот множитель (поправка Прайса-Уинстона) обеспечивает уменьшение влияния первого уравнения на все остальные при ρ близких к единице.

Тогда окончательно система уравнений наблюдений принимает вид:

$$\sqrt{(1-\boldsymbol{\rho})}\,y_1=\sqrt{(1-\boldsymbol{\rho})}\,a_0+\sqrt{(1-\boldsymbol{\rho})}\,a_1x_1+\sqrt{(1-\boldsymbol{\rho})}\,a_2x_2+\sqrt{(1-\boldsymbol{\rho})}\,U_1$$

$$y_t^*=a_0(1-\boldsymbol{\rho})+a_1z_{1t}+a_2z_{2t}+\boldsymbol{\varepsilon}_t \quad (t=2,3,\ldots,n)$$