

Министерство образования Республики Беларусь  
Учреждение образования  
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет компьютерного проектирования

Кафедра проектирования информационно-компьютерных систем

**ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ  
МНОЖЕСТВЕННОЙ РЕГРЕССИИ И ПРОВЕРКА ЗНАЧИМОСТИ В  
РЕЖИМЕ РЕГРЕССИИ**

Студент гр. 315401

Е.М. Косарева

Проверил

Г.А. Пискун

Минск 2023

## Цель работы

Используя режим Регрессия, вычислить вектор коэффициентов уравнения регрессии.

## Теоретические сведения

Табличный процессор *MS Excel* содержит модуль **Анализ данных**. Этот модуль позволяет выполнить статистический анализ выборочных данных (построение гистограмм, вычисление числовых характеристик и т. д.). Режим работы **Регрессия** этого модуля осуществляет вычисление коэффициентов линейной множественной регрессии с  $k$  переменными, построение доверительных интервалов и проверку значимости уравнения регрессии.

1. *Входной интервал  $Y$*  – вводится диапазон адресов ячеек, содержащих значения  $y_i$  (ячейки должны составлять один столбец).
2. *Входной интервал  $X$*  – вводится диапазон адресов ячеек, содержащих значения независимых переменных. Значения каждой переменной представляются одним столбцом. Количество переменных – не более 16.
3. *Метки* – включается, если первая строка во входном диапазоне содержит заголовок. В этом случае автоматически будут созданы стандартные названия.
4. *Уровень надежности* – при включении этого параметра задается надежность  $\gamma$  при построении доверительных интервалов.
5. *Константа-ноль* – при включении этого параметра коэффициент  $b_0 = 0$ .
6. *Выходной интервал* – при включении активизируется поле, в которое необходимо ввести адрес левой верхней ячейки выходного диапазона, который содержит ячейки с результатами вычислений режима **Регрессия**.
7. *Новый рабочий лист* – при включении этого параметра открывается новый лист, в который, начиная с ячейки **A1**, вставляются результаты работы режима **Регрессия**.
8. *Новая рабочая книга* – при включении этого параметра открывается новая книга, на первом листе которой, начиная с ячейки **A1**, вставляются результаты работы режима **Регрессия**.
9. *Остатки* – при включении вычисляется столбец, содержащий невязки  $y_i - \hat{y}_i, i = 1, \dots, n$ .
10. *Стандартизованные остатки* – при включении вычисляется столбец, содержащий стандартизованные остатки.
11. *График остатков* – при включении выводятся точечные графики невязки  $y_i - \hat{y}_i, i = 1, \dots, n$  в зависимости от значений переменных  $x_j, j = 1, \dots, k$ . Количество графиков равно числу  $k$  переменных  $x_j$ .
12. *График подбора* – при включении выводятся точечные графики предсказанных по построенной регрессии значений  $\hat{y}_i$  от значений переменных  $x_j, j = 1, \dots, k$ . Количество графиков равно числу  $k$  переменных  $x_j$ .

## Реализация решения задачи

Регрессионная статистика представлена на рисунке 1:

7	Регрессионная статистика					
8	R	0,3667	R-квадрат	0,1345	Скорректированный R-квадрат	-0,1128
9	MSE	5,9843	S	2,4463	MAPE	57,3239
10	Дарбин—Уотсон (DW)	1,9857	Log-правдоподобие	-21,3517		
11	Инф. критерий Акаике (AIC)	4,8703	AICc	5,0418		
12	Критерий Шварца (BIC)	4,9611	Критерий Хеннана-Куинна (HQC)	4,7708		
13	PRESS	162,9872	PRESS RMSE	4,0372	Предсказанный R-квадрат	-2,3675
14						
15						

Рисунок 1 – Регрессионная статистика

На рисунке 2 представлен дисперсионный анализ:

15						
16	$Y = 2.9732 + 0.0558 * X1 + 0.1249 * X2$					
17						
18	Дисперсионный анализ					
19		d.f.	SS	MS	F	p
20	Регрессия	2,	6,5098	3,2549	0,5439	0,6032
21	Остаток	7,	41,8902	5,9843		
22	Итого	9,	48,4000			
23						

Рисунок 2 – Дисперсионный анализ

Коэффициенты представлены на рисунке 3:

13						
14		Коэффициенты	Станд. ошибка	LCL	UCL	t-статистика
15	Константа	2,9732	2,0223	-1,8087	7,7552	1,4702
16	X1	0,0558	0,2027	-0,4235	0,5351	0,2752
17	X2	0,1249	0,1203	-0,1595	0,4092	1,0382
18	T (5%)	2,3646				
19	LCL - Нижняя граница 95% доверительного интервала					
20	UCL - Верхняя граница 95% доверительного интервала					
21						

Рисунок 3 – Коэффициенты

Остатки представлены на рисунке 4:

34	Остатки										
35	Наблюдение	Y	Предсказанное Y	Остаток	Стандартизованные	Стьюдентизированные	Удаленные t	Разбалансировка	Cook's D	DFIT	PRESS
36	1	4,0000	6,3176	-2,3176	-0,9474	-1,9601	-2,7019	0,7664	4,2014	-4,8937	-9,9207
37	2	2,0000	3,7808	-1,7808	-0,7280	-0,8304	-0,8097	0,2315	0,0693	-0,4444	-2,3173
38	3	5,0000	4,1394	0,8606	0,3518	0,3962	0,3710	0,2117	0,0141	0,1923	1,0918
39	4	6,0000	3,5709	2,4291	0,9930	1,1479	1,1795	0,2517	0,1477	0,6841	3,2462
40	5	2,0000	3,6825	-1,6825	-0,6878	-0,7605	-0,7351	0,1821	0,0429	-0,3469	-2,0571
41	6	2,0000	4,0571	-2,0571	-0,8409	-0,8978	-0,8837	0,1228	0,0376	-0,3307	-2,3451
42	7	6,0000	4,8089	1,1911	0,4869	0,7150	0,6876	0,5363	0,1971	0,7394	2,5686
43	8	8,0000	4,6388	3,3612	1,3740	1,5582	1,7850	0,2224	0,2315	0,9547	4,3227
44	9	2,0000	3,8631	-1,8631	-0,7616	-0,8235	-0,8023	0,1447	0,0382	-0,3300	-2,1783
45	10	7,0000	5,1409	1,8591	0,7600	0,9287	0,9182	0,3304	0,1418	0,6449	2,7762
46	Минимум	2,0000	3,5709	-2,3176	-0,9474	-1,9601	-2,7019	0,1228	0,0141	-4,8937	-9,9207
47	Максимум	8,0000	6,3176	3,3612	1,3740	1,5582	1,7850	0,7664	4,2014	0,9547	4,3227
48	Среднее	4,4000	4,4000	0,0000	0,0000	-0,0526	-0,0991	0,3000	0,5122	-0,3130	-0,4813
49											

Рисунок 4 – Остатки

График остатков представлен на рисунке 5:

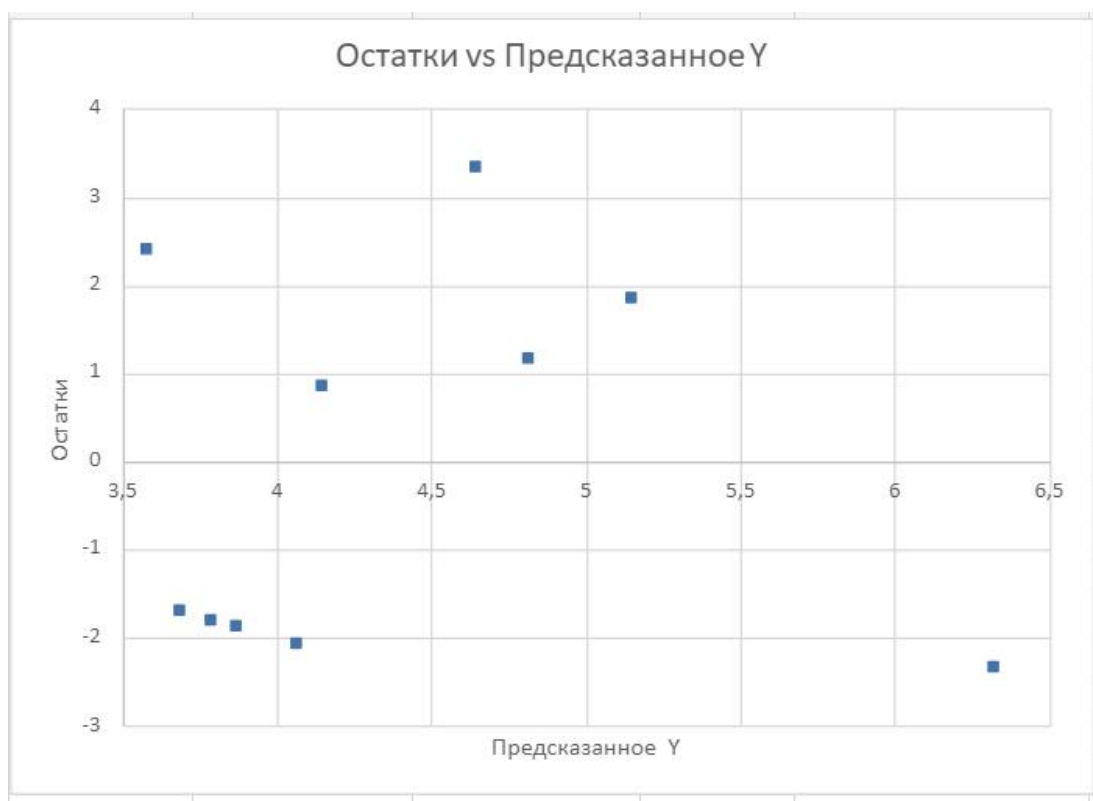


Рисунок 5 – График остатков

## Выводы

В результате применения режима Регрессия был получен комплексный анализ данных.