8 ОТОБРАЖЕНИЕ ВЗАИМОСВЯЗИ МЕЖДУ ПАРАМЕТРАМИ С ПРИМЕНЕНИЕМ МНОЖЕСТВЕННОГО РЕГРЕССИОННОГО АНАЛИЗА

Множественный регрессионный анализ – это метод математической статистики, который позволяет найти наиболее точное и достоверное отображение (модель, аппроксимацию, уравнение регрессии) стохастической зависимости между откликом Y и несколькими факторами X_1, X_2, … , X_j, … , X_m. Для решения данной задачи необходимо:

1. Определить вид уравнения регрессии.
2. Оценить значимость коэффициентов регрессии.
3. Оценить допустимость отображения исследуемой зависимости выбранным уравнением регрессии.
4. Исследовать остатки (отклонения действительных значений отклика от предсказываемых по уравнению регрессии).

8.1. Определение вида уравнения множественной регрессии

Учитывая возможные отклонения, модель связи отклика с некоторым комплексом факторов $\vec{X} = \{X_1, \ldots, X_j, \ldots, X_m\}$ должна быть представлена в виде двух составляющих:

$$y = \varphi\left(\vec{X}\right) + \varepsilon, \tag{8.1}$$

где $\varphi\left(\vec{X}\right)$ - систематическая (объясненная) составляющая. Она обусловлена существованием зависимости между откликом и комплексом факторов;

ε - случайная составляющая. Она обусловлена разнообразными возмущениями и вызывает отклонения y от значений, соответствующих реальной зависимости.

Для построения множественной регрессионной модели (иначе – множественного уравнения регрессии или просто множественной регрессии) необходимо решить следующие задачи:

Задача определения вида уравнения множественной регрессии состоит в нахождении систематической составляющей $\varphi\left(\vec{X}\right)$. Однако, поскольку используются выборки ограниченного объема ($n \lll \infty$), могут быть найдены лишь оценки истинных параметров.

Пусть, например, действительная зависимость отклика от комплекса факторов является линейной:

$$y = \varphi(\vec{X}) = \beta_0 + \sum_{j=1}^{m} \beta_j X_j . \qquad (8.2)$$

Оценкой (моделью, отображением, аппроксимацией) этой связи также может быть линейное выражение:

$$\widehat{y} = \widehat{\varphi}(\vec{X}) = b_0 + \sum_{j=1}^{m} b_j X_j . \qquad (8.3)$$

В выражении (8.3), которое и есть уравнение регрессии, коэффициенты регрессии b_0 и b_j ($j = 1,2,\ldots,m$) представляют собой оценки коэффициентов истинной зависимости ($b_0 \approx \beta_0$ и $b_j \approx \beta_j$).

Для подбора уравнения $\widehat{y} = \widehat{\varphi}(\vec{X})$, которое наилучшим образом отображает стохастическую связь между откликом и рассматриваемыми факторами, используют метод наименьших квадратов (МНК). Согласно МНК наилучшей оценкой исследуемой зависимости является та, которая дает наименьшую сумму квадратов отклонений наблюдаемых значений отклика y_i от рассчитанных по уравнению регрессии $\widehat{y}_i$ при тех же значениях факторов $\{x_{1i},\ldots,x_{ji},\ldots,x_{mi}\}$. Это условие выражается следующим образом:

$$\sum_{i=1}^{n} {e_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \min . \qquad (8.4)$$

Исходя из условия (8.4) задача определения коэффициентов уравнения регрессии сводится практически к определению минимума функции нескольких переменных и решена математической статистикой для линейного уравнения. Значения коэффициентов регрессии в (8.3) вычисляются решением системы из n линейных уравнений с m неизвестными (здесь n - число наблюдений).

В *MS Excel* коэффициенты линейной аппроксимации могут быть определены с использованием статистической функции ЛИНЕЙН(). Синтаксис функции и вопросы, связанные с ее использованием приведены в приложении 11. Для комплексного решения задачи множественного регрессионного анализа в *MS Excel* имеется инструмента «Регрессия».

8.2. Оценивание качества множественной аппроксимации

Для оценивания качества множественной аппроксимации необходимо определить ее статистическую надежность, проверить значимость коэффициентов регрессии и проверить выполнение допущений относительно остатков. Оценивание статистической надежности уравнения множественной регрессии выполняется так же, как и парной аппроксимации (см. подраздел 7). Поэтому здесь рассмотрим только две последние задачи.

8.2.1. Проверка значимости коэффициентов регрессии

Коэффициенты регрессии b_j являются случайными величинами с математическими ожиданиями β_j и дисперсиями, которым соответствуют стандартные отклонения S_{bj}. Значение b_j признается статистически значимым, если выполняется условие:

$$t_{bj} = \frac{|b_j|}{S_{bj}} > t[\alpha; n-k], \tag{8.5}$$

где t_{bj} и $t[\alpha; n-k]$ - расчетное и табличное числа Стьюдента.

Если условие (8.5) не выполнено, то следует признать, что $b_j = 0$ и влияние фактора X_j на отклик несущественное. В таком случае рекомендуется повторить регрессионный анализ без учета фактора X_j.

8.2.2. Анализ остатков

Остатками принято называть отклонения действительных значений отклика от рассчитанных по уравнению регрессии (для i-го наблюдения остаток $e_i = y_i - \hat{y}$). Отклонения обусловлены наличием в (8.1) случайной составляющей ε, относительно которой делают следующие предположения:

1. Это нормально распределенная случайная переменная.
2. Математическое ожидание случайной составляющей равно нулю - $M(\varepsilon) = 0$. Считают, что данная гипотеза выполняется, если среднее выборочное остатков можно считать равным нулю:

$$\bar{e} = \frac{1}{n}\sum_{i=1}^{n} e_i \approx 0. \tag{8.6}$$

3. Дисперсия случайной составляющей постоянна - $D(\varepsilon) = Const$. Гипотеза может быть проверена, например, построением графиков остатков в зависимости от каждого фактора. Если на всех таких графиках остатки примерно равномерно рассеяны в пределах области, параллельной оси X_j, то гипотезу $D(\varepsilon) = Const$ считают справедливой.
4. В различных наблюдениях значения ε не зависят друг от друга. Для проверки гипотезы о независимости отклонений в различных наблюдениях оценивают автокорреляцию остатков с применением критерия Дарбина-Уотсона (критерия DW):

$$DW = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}. \tag{8.7}$$

Строгое условие отсутствия автокорреляции DW =2. Однако, с учетом особенностей распределения критерия Дарбина-Уотсона, ориентировочно можно считать, что автокорреляция остатков отсутствует при $1{,}2 \le DW \le 2{,}8$. В противном случае следует признать, что гипотеза о независимости остатков в рассматриваемом случае не верна.

Если анализ остатков обнаруживает несоответствия указанным гипотезам, то уравнение регрессии, относительно которого данные остатки получены, следует считать неудовлетворительным, т. к. правомерность применения МНК и указанных выше оценок и для множественного регрессионного анализа может быть поставлена под сомнение. В таком случае рекомендуют рассмотреть уравнение иного вида (например, нелинейное вместо линейного), включить неучтенные ранее факторы, выделить в области варьирования факторов различные подобласти.

8.3. Пример множественного регрессионного анализа в *MSExcel* с применением инструмента «Регрессия»

В качестве примера рассмотрим задачу построения зависимости предела текучести металла, прокатанного на широкополосном стане горячей прокатки (ШСГП) от температур конца прокатки (tкп) и смотки (tсм). Исходные данные заносятся на рабочий лист с клавиатуры (на рис. 8.1 и 8.2 они расположены в ячейках А1:С29).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ИСХОДНЫЕ ДАННЫЕ				ВЫВОД ИТОГОВ			УРАВНЕНИЕ РЕГРЕССИИ						
2	Sт, МПа	tкп,С	tсм,С					Sт = 1152,965 - 0,498*tкп -0,695*tсм						
3	364,0	788	560		*Регрессионная статистика*			ЗНАЧИМОСТЬ КОЭФФИЦИЕНТОВ						
4	341,2	788	584		Множественный R	0,937		t[0,05;n-k]	b0	Значим				
5	338,4	788	611		R-квадрат	0,879		2,0595	b1	Значим				
6	284,1	788	639		Нормированный R-квадрат	0,869			b2	Значим		АВТОКОРРЕЛЯЦИЯ ОСТАТКОВ		
7	330,9	788	657		Стандартная ошибка	18,56		НАДЕЖНОСТЬ АППРОКСИМАЦИИ						
8	281,9	788	715		Наблюдения	28		F[0,05;k-1;n-k]	Аппроксимация			DW	2,3644	
9	265,9	788	728					3,3852	надежная			Отсутствует		
10	336,4	834	560		Дисперсионный анализ									
11	320,5	834	584			*df*	*SS*	*MS*	*F*	*Значимость F*				
12	336,8	834	611		Регрессия	2	62268,036	31134,018	90,4187	3,6E-12				
13	299,4	834	639		Остаток	25	8608,291	344,332						
14	269,3	834	657		Итого	27	70876,327							
15	240,3	834	715											
16	206,3	834	728			*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 90,0%*	*Верхние 90,0%*	
17	334,2	875	560		Y-пересечение	1152,695	73,484	15,6863	1,9E-14	1001,352	1304,039	1027,174	1278,217	
18	290,3	875	584		tкп,С	-0,498	0,073	-6,7941	4,03E-07	-0,649	-0,347	-0,623	-0,373	
19	318,5	875	611		tсм,С	-0,695	0,060	-11,6051	1,47E-11	-0,818	-0,572	-0,797	-0,593	
20	283,2	875	639											
21	281,0	875	657											
22	233,1	875	715											
23	200,0	875	728		ВЫВОД ОСТАТКА									
24	304,4	917	560											
25	307,5	917	584		*Наблюдение*	*Предсказанное Sт, МПа*	*Остатки*	$(e_i - e_{i-1})^2$						
26	268,8	917	611		1	371,3	-7,3	53,6						
27	225,4	917	639		2	354,6	-13,4	37,48						
28	270,3	917	657		3	335,9	2,5	254,80						

Рис. 8.1. Фрагмент рабочего листа с результатами множественного регрессионного анализа инструментом «РЕГРЕССИЯ»

8.3.1. Аппроксимация с применением инструмента «Регрессия»

При выполнении работы настройки инструмента «Регрессия» должны соответствовать указанным в приложении 10.

В этом случае основные результаты (ячейки E1:M19 на рис. 8.1) будут дополнены таблицей остатков (ячейки E23:G23 на рис. 8.2).

	A	B	C	D	E	F	G	H
22	233,1	875	715					
23	200,0	875	728		ВЫВОД ОСТАТКА			
24	304,4	917	560					
25	307,5	917	584		*Наблюдение*	*Предсказанное Sm, МПа*	*Остатки*	$(e_i - e_{i-1})^2$
26	268,8	917	611		1	371,3	-7,3	53,6
27	225,4	917	639		2	354,6	-13,4	37,48
28	270,3	917	657		3	335,9	2,5	254,80
29	181,3	917	715		4	316,4	-32,3	1214,00
30	175,3	917	728		5	303,9	27,0	3517,48
31					6	263,6	18,3	75,61
32					7	254,6	11,3	48,53
33					8	348,4	-12,0	545,15
34					9	331,8	-11,3	0,60
35					10	313,0	23,8	1229,38
36					11	293,5	5,9	321,94
37					12	281,0	-11,7	309,47
38					13	240,7	-0,4	127,80
39					14	231,7	-25,4	623,31
40					15	328,0	6,2	996,21
41					16	311,3	-21,0	741,05
42					17	292,6	25,9	2205,48
43					18	273,1	10,1	250,99
44					19	260,6	20,4	106,26
45					20	220,3	12,8	57,69
46					21	211,3	-11,3	579,18
47					22	307,1	-2,7	73,28
48					23	290,4	17,1	391,16
49					24	271,7	-2,9	397,50
50					25	252,2	-26,8	573,25
51					26	239,7	30,6	3295,72
52					27	199,4	-18,1	2371,23
53					28	190,4	-15,1	9,20
54								

Рис. 8.2. Таблица остатков, полученная с использованием инструмента «РЕГРЕССИЯ»

Таблицу остатков необходимо дополнить столбцом, во всех строках которого, начиная со второй, вычислить значения $(e_i - e_{i-1})^2$. Эти данные будут в дальнейшем использованы для расчета критерия DW. Например, в ячейке H27:

=(G27-G26)^2.

На основании результатов работы инструмента записать уравнение регрессии в содержательной форме, оценить значимость коэффициентов регрессии, надежность аппроксимации и автокорреляцию остатков.

Уравнение регрессии в содержательной форме (ячейки H1:J2) записывается с клавиатуры.

Оценивание значимости коэффициентов регрессии (ячейки H3:J6). В ячейке H5 определяется табличное число Стьюдента. Для рассматриваемого примера (число коэффициентов регрессии k =3):

=СТЬЮДРАСПОБР(0,05;F8-3).

В ячейках J4:J6, с использованием функции ЕСЛИ() программируется вывод о значимости коэффициентов регрессии. Например, для ячейки J4:

=ЕСЛИ(ABS(H17)>$H$5;”Значим”;”Не значим”) .

Внимание! Если коэффициент регрессии при факторе Xj оказался не значимым, необходимо повторить регрессионный анализ, не включая во входной интервал X столбец со значениями Xj. Поскольку входной интервал X должен состоять из смежных столбцов, может оказаться необходимым перегруппировать столбцы со значениями факторов.

Оценивание надежности аппроксимации в примере на рис. 8.1 выполняется в ячейках H7:J9. В ячейке H9, с помощью статистической функции FРАСПОБР(), определяется табличное значение числа Фишера:

=FРАСПОБР(0,05;2;F8-3) .

Слово «Аппроксимация» в ячейки I8:J8 введено с клавиатуры. Собственно вывод («надежная» или «не надежная») формируется в ячейке I9 с применением функции ЕСЛИ() путем сравнения табличного (из ячейки H9) и рассчитанного (из ячейки I12) чисел Фишера:

=ЕСЛИ(H9>I12;”надежная”;”не надежная”) .

Оценивание автокорреляции остатков выполнено в ячейках L7:M9. Значение критерия Дарбина-Уотсона в ячейке M8 вычисляется по формуле, которая для рассматриваемого примера программируется следующим образом:

=СУММ(H27:H53)/G13 .

Вывод в ячейке L9 формируется с применением функции ЕСЛИ(), которая проверяет условие $1{,}2 \le DW \le 2{,}8$:

ЕСЛИ(M8<1,2;”Существует”;ЕСЛИ(M8>2,8;”Существует”;”Отсутствует”)).

Пример множественного регрессионного анализа в программе Statistica приведен в приложении 12.

8.3.2. Анализ результатов множественного регрессионного анализа

Анализируя результаты множественного регрессионного анализа необходимо ответить на следующие вопросы:

1. Связь между какими величинами анализировалась?
2. Значимы ли коэффициенты регрессии?

3. Как выглядит уравнение множественной регрессии, полученное в результате выполнения работы?

4. Можно ли считать полученное уравнение множественной регрессии статистически надежной аппроксимацией анализируемой зависимости?

Применительно к рассмотренному примеру можно сказать следующее.Анализировалась связь между пределом текучести металла $\sigma_т$, температурой конца прокатки $t_{кп}$ и смотки $t_{см}$ на ШСГП.

С доверительной вероятностью р = 95% коэффициенты регрессии $b(t_{кп})$=-0,498 и $b(t_{см})$=-0,695 являются статистически значимыми, т. к. соответствующие числа Стьюдента $|t(t_{кп})|$=6,7941 и $|t(t_{см})|$=11,6051 больше табличного t[0,05;25]=2,0595.

Для рассмотренных условий множественная линейная аппроксимация зависимости предела текучести металла от температуры конца прокатки и смотки на ШСГП имеет вид:

$$\sigma_{\grave{O}} = 1152{,}695 - 0{,}498 t_{\hat{e}\ddot{i}} - 0{,}695 t_{\tilde{n}\grave{i}} \ .$$

С доверительной вероятностью 95% полученное уравнение регрессии можно считать статистически надежной аппроксимацией исследуемой зависимости, т. к. рассчитанное число Фишера $F_р$=90,4187 больше табличного F[0,05;2;25]=3,3352.

8.4. Пример нелинейного множественного регрессионного анализа в MS Excel

В качестве примера рассмотрим задачу построения аппроксимации зависимости напряжения текучести при горячей пластической деформации от термомеханических параметров процесса. Массив исходных данных записан на рабочем листе (рис. 8.3) в ячейках B3:E32. Напряжение текучести σ_s является откликом (Y), а термомеханические параметры степень ε, скорость u и температура t деформации – факторами X_1, X_2 и X_3 соответственно.

Зависимость напряжения текучести от термомеханических параметров часто отображают степенной зависимостью:

$$\sigma_s = b_0 \varepsilon^{b_1} u^{b_2} \left(t/1000\right)^{b_3} . \qquad (8.6)$$

Для построения нелинейных аппроксимаций поступают следующим образом. Сначала уравнение переводят в линейную форму. Линеаризацию зависимости (8.6) осуществляют логарифмированием:

$$ln(\sigma_s) = ln(b_0) + b_1 ln(\varepsilon) + b_2 ln(u) + b_3 ln(t/1000) . \qquad (8.7)$$

Таким образом, получают новые значения переменных $y' = ln(y)$ и $x'_j = ln(x_j)$, с использованием которых строят линейную аппроксимацию

	A	B	C	D	E	G	H	I	J	L	M	N	O	Q	R	S
1	№	σs, МП	ε, %	u, 1/с	t,°C	ПРЕОБРАЗОВАНИЕ				t	u	ε		ОСТАТКИ		
2	п/п	(Y)	(X1)	(X2)	(X3)	ln(σs)	ln(ε)	ln(u)	ln(t/1000)	РЕЗУЛЬТАТЫ ЛИНЕЙН()				y^	e	E
3	1	6,2	5	10	1000	1,8	1,61	2,30	0,000	-3,082	0,498	0,264	0,269	6,30	0,098	-2,97
4	2	7,6	10	10	1000	2,0	2,30	2,30	0,000	0,158	0,017	0,017	0,071	7,56	0,003	-1,61
5	3	8,3	14	10	1000	2,1	2,64	2,30	0,000	0,983	0,045	#Н/Д	#Н/Д	8,27	-0,051	-0,85
6	4	9,0	18	10	1000	2,2	2,89	2,30	0,000	490,052	26	#Н/Д	#Н/Д	8,83	-0,164	-0,17
7	5	9,5	22	10	1000	2,2	3,09	2,30	0,000	3,036	0,054	#Н/Д	#Н/Д	9,31	-0,137	0,28
8	6	9,8	26	10	1000	2,3	3,26	2,30	0,000	РЕГРЕССИОННАЯ СТАТИСТИКА				9,73	-0,095	0,66
9	7	10,2	30	10	1000	2,3	3,40	2,30	0,000	**R²**	**Se**	**n-k**		10,11	-0,098	1,04
10	8	10,4	34	10	1000	2,3	3,53	2,30	0,000	0,983	0,045	26	0,054	10,45	0,015	1,26
11	9	10,6	38	10	1000	2,4	3,64	2,30	0,000	∑E²	∑e²	Fp		10,76	0,175	1,41
12	10	10,7	40	10	1000	2,4	3,69	2,30	0,000	3,0362	0,054	490,052	3,036	10,91	0,246	1,49
13	1	5,5	10	5	1000	1,7	2,30	1,61	0,000	ТАБЛИЧНЫЕ ЗНАЧЕНИЯ				5,36	-0,104	-3,71
14	2	7,6	10	10	1000	2,0	2,30	2,30	0,000	**P,%**	**t[0,05;n-k]**	**F[0,05;k-1;n-k]**		7,56	0,003	-1,61
15	3	10,4	10	14	1000	2,3	2,30	2,64	0,000	95	2,0555	3,3690		8,94	-1,454	1,22
16	4	11,2	10	18	1000	2,4	2,30	2,89	0,000	ОЦЕНИВАНИЕ КОЭФФИЦИЕНТОВ				10,13	-1,102	2,06
17	5	11,8	10	22	1000	2,5	2,30	3,09	0,000	b_3	b_2	b_1	b_0	11,20	-0,562	2,59
18	6	12,4	10	26	1000	2,5	2,30	3,26	0,000	-3,082	0,498	0,264	0,269	12,17	-0,221	3,22
19	7	13,2	10	30	1000	2,6	2,30	3,40	0,000	S_{b3}	S_{b2}	S_{b1}	S_{b0}	13,07	-0,163	4,06
20	8	13,8	10	34	1000	2,6	2,30	3,53	0,000	0,158	0,017	0,017	0,071	13,91	0,152	4,58
21	9	14,0	10	38	1000	2,6	2,30	3,64	0,000	t_{b3}	t_{b2}	t_{b1}	t_{b0}	14,70	0,734	4,79
22	10	14,2	10	40	1000	2,7	2,30	3,69	0,000	-19,443	28,475	15,105	3,810	15,08	0,904	5,00
23	1	9,8	10	10	900	2,3	2,30	2,30	-0,105	ЗНАЧИМ	ЗНАЧИМ	ЗНАЧИМ	ЗНАЧИМ	10,46	0,635	0,66
24	2	9,1	10	10	930	2,2	2,30	2,30	-0,073	ОЦЕНКА АППРОКСИМАЦИИ				9,46	0,386	-0,10
25	3	8,5	10	10	960	2,1	2,30	2,30	-0,041	АППРОКСИМАЦИЯ НАДЕЖНАЯ				8,58	0,109	-0,70
26	4	7,6	10	10	990	2,0	2,30	2,30	-0,010					7,80	0,240	-1,61
27	5	7,2	10	10	1020	2,0	2,30	2,30	0,020	ДЕЙСТВИТЕЛЬНОЕ УРАВНЕНИЕ				7,11	-0,067	-1,99
28	6	6,4	10	10	1050	1,9	2,30	2,30	0,049	b_3	b_2	b_1	b_0	6,51	0,081	-2,74
29	7	5,9	10	10	1080	1,8	2,30	2,30	0,077	-3,082	0,498	0,264	1,309	5,97	0,069	-3,27
30	8	5,4	10	10	1110	1,7	2,30	2,30	0,104	1,309 e^0,264 u^0,498 (t/1000)^-3,082				5,48	0,115	-3,80
31	9	4,9	10	10	1140	1,6	2,30	2,30	0,131					5,05	0,136	-4,26
32	10	4,2	10	10	1200	1,4	2,30	2,30	0,182	Se^2	S_E^2	R^2	**Fp**	4,31	0,078	-4,94
33										0,205	7,83	0,974	38,135			
34																

Рис. 8.3. Рабочий лист для построения и оценивания степенной аппроксимации

$y' = b'_0 + \sum b'_j x'_j$. В заключение необходимо перейти от значений коэффициентов b'_0 и b'_j к коэффициентам b_0 и b_j исходного уравнения. Применительно к уравнениям, подобным зависимости (8.6), обратные преобразования выполняются следующим образом: $b_j = b'_j$ и $b_0 = exp\left(b'_0\right)$.

Линеаризующие преобразования выполнены в ячейках G3:J32. Например, в G3, H3, I3 и J3 запрограммировано:

G3:=ln(B3), H3: =ln(C3), I3: =ln(D3), J3: = ln(E3/1000).

Для построения линейной аппроксимация использована функция ЛИНЕЙН(). С учетом особенностей ее применения (приложение 11) сначала выделили блок ячеек L3:O3. Затем в ячейку L3 запрограммировали

=ЛИНЕЙН(G3:G32;H3:J32;;ИСТИНА).

после чего нажали комбинацию клавиш <Ctr>+<Shift>+<Inter>. Результаты также выводятся в ячейки L3:O3. Символы #Н/Д, которые в MS Excel обычно указывают на ошибку при обращении к функции, в данном случае содержательного значения не имеют. Смысл остальных результатов иллюстрируется рис. 8.4.

В ячейках L9:N12 воспроизведены характеристики статистической надежности линейной аппроксимации. Их обозначения введены с клавиатуры, а значения принимаются ссылкой на соответствующие результаты функции ЛИНЕЙН():
L10: =L5, M10: =M5, N10: =M6, L12: =L7, M12: =M7, N12: =L6.

	K	L	M	N	O	P
1						
2		РЕЗУЛЬТАТЫ ЛИНЕЙН()				
3		b_3	b_2	b_1	b_0	
4		S_{b3}	S_{b2}	S_{b1}	S_{b0}	
5		R^2	Se	#Н/Д	#Н/Д	
6		Fp	n-k	#Н/Д	#Н/Д	
7		$\sum E^2$	$\sum e^2$	#Н/Д	#Н/Д	
8						

Рис. 8.4. – Смысл результатов работы функции ЛИНЕЙН()

В ячейках L15:N15 записаны значения доверительной вероятности (принято p=95%), также табличные значения чисел Стьюдента и Фишера

M15: =СТЬЮДРАСПОБР(1-L15/100;N10);

N15: =FРАСПОБР(1-L15/100;3;N10).

Далее с их помощью оцениваются значимость коэффициентов линейной аппроксимации (в ячейках L17:O23) и ее статистическая надежность (в ячейках L25:O25).

Для оценки значимости коэффициентов регрессии сначала в ячейки L18:O118 воспроизводятся их значения

L18: =L3, M18: =M3, N18: =N3 и O18: =O3,

а в ячейках L20:O20 – значения соответствующих стандартных отклонений:

L20: =L4, M20: =M4, N20: =N4 и O20: =O4.

Затем в ячейках L22:O22 вычисляются расчетные числа Стьюдента. Например

L22: =L18/L20.

Выводы относительно значимости коэффициентов регрессии программируются в ячейках L23:O23 с применением функции ЕСЛИ(). Например

L23: =ЕСЛИ(ABS(L22)>$M$15;"ЗНАЧИМ";"НЕ ЗНАЧИМ").

Такая же функция применяется для вывода относительно надежности линейной аппроксимации в ячейке L25

=ЕСЛИ(N12>N15;"НАДЕЖНАЯ";"НЕ НАДЕЖНАЯ").

Действительные коэффициенты исходной зависимости (8.6) определяются в ячейках L29:O29

L29: =L18, M29: =M18, N29: =N18, O29: =EXP(O18).

Для оценивания статистической надежности действительного уравнения сначала вычисляются оценки значений отклика (в ячейках Q3:Q32), квадраты остаточных (ячейки R3:R32) и объясненных (ячейки S3:S32) отклонений. Например

Q3: =$O$29*C3^$N$29*D3^$M$29*(E3/1000)^$L$29;
R3: =(Q3-B3)^2;
Q3: =(B3-СРЗНАЧ($B$3:$B$32))^2.

Далее, в ячейках L31:O31 вычисляются остаточная и объясненная дисперсии, показатель достоверности аппроксимации и соответствующее расчетное число Фишера

L31: =СУММ(R3:R32)/((СЧЁТ(R3:R32)-СЧЁТ(L18:O18)));
M31: =СУММ(S3:S32)/(СЧЁТ(L18:O18)-1);
N31: =1-L31/M31;
O31: =M31/L31.

Вывод относительно надежности действительного уравнения запрограммирован в ячейке L32

=ЕСЛИ(O31>N15;"НАДЕЖНОЕ";"НЕ НАДЕЖНОЕ").

Полученные результаты позволяют сделать следующие выводы.

1. Для условий выполненного эксперимента зависимость напряжения текучести от термомеханических параметров процесса горячей пластической деформации имеет вид

$$\sigma_s = 1{,}309\varepsilon^{0{,}264}u^{0{,}498}\left(t/1000\right)^{-3{,}082}.$$

2. С доверительной вероятностью 95% данная аппроксимация является статистически надежным отображением влияния термомеханических параметров на напряжение текучести, так как рассчитанное число Фишера F_p =330,501 больше табличного F_{95} =2,9752.

Для иллюстрации качества аппроксимации можно построить график соответствия расчетных и экспериментальных значений (рис. 8.5).

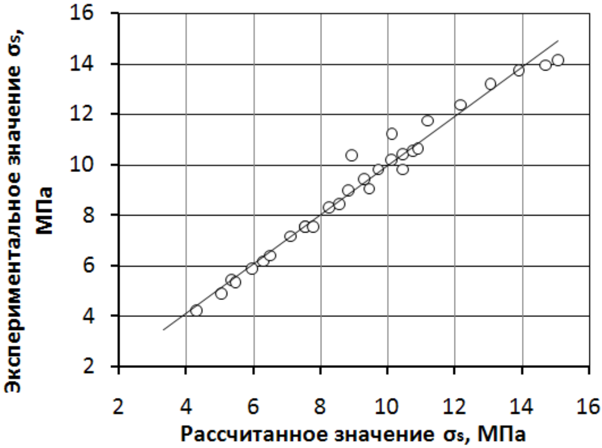

Рис. 8.5. – График соответствия расчетных и экспериментальных значений напряжения текучести

Для построения графика соответствия необходимо применить диаграмму типа «Точечная». При выборе данных значения для оси абсцисс задаются ссылкой на ячейки Q3:Q32, а значения для оси ординат - ссылкой на ячейки B3:B32. Далее необходимо добавить линейный тренд, т. к. на подобные графики наносят именно прямую линию. Чем ближе угол наклона линии к 45 градусам и чем меньше рассеяние точек относительно ее, тем точнее полученное уравнение отображает исследуемую зависимость.

8.5. Контрольные вопросы

1. Поясните сущность и укажите этапы множественного регрессионного анализа.
2. Укажите допущения множественного регрессионного анализа.
3. Запишите модель множественного регрессионного анализа.
4. Что представляет собой уравнение множественной регрессии?
5. Как определить качество уравнения множественной регрессии?
6. Как график соответствия расчетных и экспериментальных значений отклика характеризует качество аппроксимации?