МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

ЭКОНОМЕТРИКА (ПРОДВИНУТЫЙ УРОВЕНЬ). СБОРНИК ЗАДАНИЙ И КЕЙСОВ ДЛЯ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

Рекомендовано редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» в качестве методических указаний для студентов Самарского университета, обучающихся по основной образовательной программе высшего образования по направлению подготовки 38.04.01 Экономика

Составители: *Е.А. Блинова,*
О.А. Кузнецова

Самара
Издательство Самарского университета
2020

УДК 330.4(075)
ББК 65в6я7

Составители: ***Е.А. Блинова, О.А. Кузнецова***

Рецензент: д-р экон. наук, проф. Д. Ю. Иванов

Эконометрика (продвинутый уровень). Сборник заданий и кейсов для практических занятий: методические указания / *Е.А. Блинова, О.А. Кузнецова*; Министерство науки и высшего образования Российской Федерации, Самарский университет. – Самара: Издательство Самарского университета, 2020. – 1 CD-ROM (1,2 Мб). – Загл. с титул. экрана. – Текст: электронный.

В сборнике заданий и кейсов приведены задачи для анализа на практических занятиях по курсу эконометрика.

Предназначены для студентов магистратуры Самарского университета, обучающихся по направлению подготовки 38.04.01 Экономика.

Подготовлены на кафедре «Математических методов в экономике».

УДК 330.4(075)
ББК 65в6я7

Минимальные системные требования:
PC, процессор Pentium, 160 МГц;
Microsoft Windows XP и выше; мышь; дисковод CD-ROM;
Adobe Acrobat Reader.

Редактор Л.Р. Д м и т р и е н к о
Компьютерная верстка Л.Р. Д м и т р и е н к о

Подписано для тиражирования 17.09.2020.
Объем издания 1,2 Мб.
Количество носителей 1 диск.
Тираж 10 дисков.

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)
443086 Самара, Московское шоссе, 34.

Изд-во Самарского университета.
443086 Самара, Московское шоссе, 34.

ПРАКТИКА 1
ПОСТРОЕНИЕ ЭКОНОМЕТРИЧЕСКОЙ МОДЕЛИ

Задание 1

Составить спецификацию модели макроэкономики.

Экономический объект – закрытая национальная экономика. Ее состояние в текущем периоде t описывается переменными: Yt – уровень ВВП, Ct – потребление, It – инвестиции, Gt – государственные расходы.

Требуется, применяя четыре принципа спецификации эконометрических моделей, составить спецификацию макромодели, при составлении которой учесть следующие экономические утверждения: а) текущее потребление объясняется автономным потреблением и уровнем ВВП в предыдущем периоде, возрастая вместе с ним, но с меньшей скоростью; б) величина инвестиций прямо пропорциональна приросту ВВП за предшествующий период; в) государственные расходы возрастают с постоянным темпом роста g; г) текущее значение ВВП есть сумма текущих уровней потребления, инвестиций и государственных расходов.

Задание 2

Построена прогнозная модель множественной регрессии. Разобрать решение задачи и найти ошибки.

В качестве статистических данных для анализа выбраны основные годовые показатели деятельности компании за 12 лет –

с 2005 г. по 2017 г., взятые из бухгалтерской финансовой отчетности ПАО «Газпром» [1].

Введем следующие обозначения:

x_1 – средний обменный курс рубля к доллару за период;

x_2 – списочная численность работников Группы, тыс. человек;

x_3 – объем НИОКР в денежном выражении, выполненных по заказу Группы Газпром (без НДС), млрд руб.;

x_4 – выручка от продажи газа (за вычетом НДС, акциза и таможенных пошлин), млн руб.;

x_5 – выручка от реализации нефти и газового конденсата (за вычетом НДС, акциза и таможенных пошлин), млн руб.;

x_6 – выручка от реализации продуктов нефтегазопереработки (за вычетом НДС, акциза и таможенных пошлин), млн руб.;

x_7 – среднегодовая цена нефти Brent (Dated);

x_8 – среднегодовая цена нефти Urals (среднее CIF MED/ RDAM);

x_9 – ввод в эксплуатацию новых магистральных газопроводов и отводов, км.

x_{10} – добыча природного и попутного газа, млрд $м^{3;}$

x_{11} – добыча газового конденсата, млн т.;

x_{12} – добыча нефти, млн т.;

x_{13} – цена за акцию на закрытие торгов на ПАО Московская Биржа на конец года, руб.

В качестве результирующего фактора выбрана выручка предприятия (y).

Таблица 1. **Корреляционная матрица**

	Y	*x1*	*x2*	*x3*	*x4*	*x5*	*x6*	*x7*	*x9*	*x10*	*x11*	*x12*
Y	1,00											
x1	0,68	1,00										
x2	0,66	0,42	1,00									
x3	0,88	0,48	0,50	1,00								
x4	0,99	0,74	0,68	0,84	1,00							
x5	0,68	0,72	0,64	0,46	0,73	1,00						
x6	0,96	0,75	0,68	0,80	0,96	0,70	1,00					
x7	0,22	-0,54	0,20	0,28	0,16	-0,13	0,12	1,00				
x9	-0,08	-0,44	-0,19	0,00	-0,10	-0,22	-0,16	0,59	1,00			
x10	-0,81	-0,81	-0,47	-0,80	-0,82	-0,53	-0,81	0,21	0,29	1,00		
x11	0,82	0,84	0,52	0,54	0,84	0,65	0,92	-0,13	-0,32	-0,72	1,00	
x12	0,83	0,63	0,92	0,65	0,85	0,75	0,86	0,11	-0,17	-0,69	0,73	1,00

В результате получено следующее уравнение регрессии:

$$Y = b_0+b_1*x_1+b_2*x_2+b_3*x_3+b_4*x_4+b_5*x_5+b_6*x_6+b_{10}*x_{10}.$$

Задание 3

Построена прогнозная модель множественной регрессии. Разобрать решение задачи и найти ошибки.

Для построения прогнозной модели множественной регрессии использованы статистические данные с 2006 по 2019 гг. (всего 14 наблюдений). В качестве параметров выбраны:

y – объем продаж новых автомобилей;

x_1 – курс доллара;

x_2 – цены на бензин;

x_3 – доходы населения;

x_4 – ставка кредитования;

x_5 – цены на отечественные автомобили;

x_6 – цены на иномарки;

x_7 – стоимость обслуживания отечественных автомобилей;

x_8 – стоимость обслуживания иномарок.

В результате вычислений получена прогнозная модель.

$$y = -12{,}9x_1 + 25{,}06x_2 + 162{,}6x_4 + 0{,}004x_5 + 4{,}1x_6.$$

Задание 4

По статистическим данным определить вид модели и найти коэффициенты парной регрессии.

Таблица 2. **Статистика распределения расходов на потребление продуктов питания и средней заработной платы**

№	y	x
1	33,66	120,00
2	28,56	112,30
3	20,40	107,25
4	27,54	107,25
5	48,96	127,25
6	21,42	112,20
7	22,44	113,85
8	43,86	122,10
9	47,94	122,10
10	72,42	132,10
11	79,56	127,05
12	55,08	127,05
13	72,42	128,70
14	77,52	132,00
15	81,60	133,65
16	56,10	128,60
17	77,52	140,25
18	51,00	130,25
19	66,30	130,25
20	64,26	131,90
21	66,30	136,90

Задание 5

Годовые доходности акций компаний *А и В,* принадлежащих одной отрасли, приводятся в табл. 3. Построить модель, позволяющую оценить значения годовых доходностей акций компании *А* по значениям годовых доходностей акций компании *В*.

Таблица 3. **Статистические данные по компаниям**

t – номер года	**Доходность компании *А***	**Доходность компании *В***
1	-2,54	-5.31
2	26,50	16,84
3	4,44	0,07
4	17,12	10,03
5	10,19	4,98
б	13,88	7,52
7	4,55	0,23
8	10,28	5,53
9	11,76	5,94
10	11,89	6.09
11	5.14	0.93
12	7,70	3,22
13	7,17	2,08
14	7.57	2,81
15	17,46	10,73

Решение. Предположим, что доходности акций компании *А* и компании *В* связаны линейной стохастической зависимостью, тогда можно записать схему Гаусса-Маркова:

$$Y_t = a + bX_t + \varepsilon_t, \qquad t = 1,...,15,$$

где Y_t – доходность акций компании *A* за год *t*;

X_t —доходность акции компании *B* за год *t*.

Нам понадобятся значения средних по выборке:

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t = \frac{1}{15}\sum_{t=1}^{15} X_t = \frac{71,6900}{15} = 4,7793,$$

$$\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t = \frac{1}{15}\sum_{t=1}^{15} Y_t = \frac{153,1100}{15} = 10,2073.$$

Результаты необходимых предварительных вычислений (центрированные значения переменных $x_t = X_t - \bar{X}$, $y_t = Y_t - \bar{Y}$, вычисленные произведения $x_t y_t$ и квадраты центрированных значений переменных x_t^2, y_t^2) внести в табл. 4.

Таблица 4. **Вспомогательная таблица**

t	**x_t**	**y_t**	**$x_t y_t$**	x_t^2	y_t^2
1	**2**	**3**	**4**	**5**	**6**
1					
2					
3					
4					
5					
6					
7					

Окончание табл. 4

8					
9					
10					
11					
12					
13					
14					
15					
		$\sum$			

Оценки параметров равны:

$$\hat{b} = \frac{\sum x_t y_t}{\sum x_t^2} = \frac{508,400}{392,701} = 1,2946,$$

$$\hat{a} = \overline{Y} - \hat{b}\overline{X} = 10,2073 - 1,2946 \cdot 4,7793 = 4,0200.$$

Таким образом, уравнение регрессии с оцененными параметрами имеет вид:

$$\hat{Y}_t = \hat{a} + \hat{b}X_t = 4,0200 + 1,2946X_t.$$

Задание 6. Используя данные задания 5, определить границы доверительных интервалов для параметров парной регрессионной модели.

Решение. Для определения границ доверительных интервалов необходимо вычислить оценки дисперсий: остатков регрессии

s^2 и коэффициентов модели: $s_{\hat{a}}^2, s_{\hat{b}}^2$. Предварительные вычисления оформить, для наглядности, в виде следующей таблицы:

Таблица 5. **Предварительные вычисления**

t	**Y_t**	**X_t**	$\widehat{Y}_t$	**e_t**	e_t^2	$(\widehat{Y} - \overline{Y})^2$
1	**2**	**3**	**4**	**5**	**6**	**7**
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
				$\sum$		

Оценка доходности компании *A* (столбец 4 табл. 5) выполнена в соответствии с моделью, настроенной по выборочным данным в задании 5:

$$\widehat{Y}_t = \widehat{a} + \widehat{b}X_t = 4,0200 + 1,2946X_t .$$

Суммируя данные столбца 6, получим: $\sum_{t=1}^{15} e_t^2 = 2,306.$

Оценку дисперсии возмущений регрессии вычислим по формуле:

$$s^2 = \widehat{\sigma}^2 = \frac{1}{n-2}\sum_{t=1}^{15} e_t^2 = \frac{2,306}{13} = 0,177,$$

оценки дисперсий коэффициентов:

$$s_{\widehat{a}}^2 = s^2 \frac{\sum X_t^2}{n\sum x_t^2} = 0,177\frac{735,333}{15\cdot 392,702} = 0,022,$$

$$s_{\widehat{b}}^2 = \frac{s^2}{\sum x_t^2} = \frac{0,177}{392,702} = 0,0005.$$

Табличное значение t-статистики для степени свободы n – 2 = 13 и уровня значимости α = 5% равно $t_{кр}$ = 2,160, поэтому границы доверительных интервалов параметров модели принимают значения:

$$\widehat{a} - t_{кр}\cdot s_{\widehat{a}} = 4,020 - 2,160\cdot\sqrt{0,022} = 3,699;$$

$$\widehat{a} + t_{кр}\cdot s_{\widehat{a}} = 4,020 + 2,160\cdot\sqrt{0,022} = 4,341,$$

т. е.

$$3,699 < a < 4,341,$$

$$\widehat{b} - t_{кр}\cdot s_{\widehat{b}} = 1,295 - 2,160\cdot\sqrt{0,0005} = 1,246;$$

$$\widehat{b} + t_{кр}\cdot s_{\widehat{b}} = 1,295 + 2,160\cdot\sqrt{0,0005} = 1,343,$$

т.е. 1,246<b<1,343.

Задание 7. Выполним проверку значимости регрессора для сквозного задания 5. Оцененная регрессионная модель в стандартной форме записывается следующим образом:

$$Y_t = \underset{(S_a)}{\widehat{a}} + \underset{(S_b)}{\widehat{b}} \cdot X_t + \underset{(S)}{\widehat{\varepsilon}_t},$$

с учетом полученных оценок

$$Y_t = \underset{(0{,}148)}{4{,}02} + \underset{(0{,}022)}{1{,}295} \cdot X_t + \underset{(0{,}421)}{\widehat{\varepsilon}_t}.$$

Для проверки значимости регрессора вычислим *t*-статистику и сравним ее значение с критическим $t_{\text{кр}}$, например, для уровня значимости α= 5 % и числа степеней свободы v = 13:

$$|t_b| = \left|\frac{\widehat{b}}{s_{\widehat{b}}}\right| = \left|\frac{1{,}295}{0{,}0212}\right| \approx 60{,}93 > t_{кр} = 2{,}16;$$

поскольку значение значительно больше критического, нулевая гипотеза $H_0{:}b = 0$ отвергается при уровне значимости α = 5 %. Коэффициент признается значимым, а, следовательно, значим и регрессор.

Задание 8. Выполнить оценку доходности компании *A* для момента t = 3, при котором доходность компании *B* принимает значение X_3 = 0,07. Построить доверительные интервалы для оценок среднего и индивидуального значений доходности.

Решение. Используя уравнение регрессии с оцененными параметрами (для сквозного примера), получим оценку ожидаемого значения доходности компании *A* для момента t = 3:

$$\widehat{Y}_t = \widehat{a} + \widehat{b}X_t = 4,0200 + 1,2946X_t = 4,0200 + 1,2946 \cdot 0,07 = 4,111.$$

Для определения доверительного интервала вычислим оценку дисперсии значения зависимой переменной:

$$s_{\hat{Y}}^2 = s^2\left(\frac{1}{n} + \frac{\left(x_p\right)^2}{\sum x_t^2}\right) = 0,177\left(\frac{1}{15} + \frac{\left(0,07 - 4,779\right)^2}{392,702}\right) = 0,022,$$

где $x_p = (0,070 - 4,779) = -4,709$ – центрированное по выборке значение регрессора. Таким образом, границы доверительного интервала среднего значения доходности равны

$$4,111 - 2,160 \cdot \sqrt{0,022} = 3,791;$$

$$4,111 + 2,160 \cdot \sqrt{0,022} = 4,431.$$

Для определения границ доверительного интервала для отдельных индивидуальных значений зависимой переменной Y_t необходимо вычислить оценку дисперсии ошибки e_p, которая на интервале оценивания определяется по формуле:

$$s_p^2 = s^2 - s_{\hat{Y}}^2 = s^2\left(1 - \frac{1}{n} - \frac{\left(x_p\right)^2}{\sum x_t^2}\right) = 0,177\left(1 - \frac{1}{15} - \frac{\left(-4,709\right)^2}{392,702}\right) = 0,155.$$

Таким образом, границы для доверительного интервала индивидуальных значений Y_t равны

$$4,111 - 2,160 \cdot \sqrt{0,155} = 3,259;$$
$$4,111 + 2,160 \cdot \sqrt{0,155} = 4,962.$$

Задание 9. Известный английский эконометрист С. Лизер ставил задачу по построению модели, которая давала бы возможность объяснить величину сбережений S_t домашних хозяйств Соединенного Королевства в текущем году t текущим уровнем Y_t их располагаемого дохода. Для решения данной задачи С. Лизер собрал статистическую информацию об объекте-оригинале в виде конкретных годовых значений величин S_t и Y_t за отрезок времени с 1947 по 1962 гг. Их значения, выраженные в миллиардах фунтов стерлингов (в ценах базисного года), приведены в табл. 6. Оценить по данным табл. 6 спецификацию вида

$$S_t = b \cdot Y_t^2 + w_t, \quad b > 0,$$

и проверить адекватности модели.

Таблица 6. **Исходные данные для построения модели**

Год	S_t	Y_t	Y_t^2
1946	0,36	8.8	77,44
1947	0,21	9,4	88,36
1948	0,08	10,0	100,00
1949	0,20	10,6	112,36
I9S0	0,10	11,0	121,00
1951	0,12	11,9	141,61
1952	0,41	12,7	161,29
1953	0,50	13,5	182,25
1954	0,43	14,3	204,49
1955	0,59	15,5	240,25
1956	0,90	16,7	278,89
1957	0,95	17,7	313,29
1958	0,82	18,6	345,96

Продолжение табл. 6

1959	1,04	19,7	388,09
1960	1,53	21,1	445,21
1961	1,94	22,8	519,84
1962	1,75	23,9	571,21
1963	1,99	25,2	635,04

Решение. Поделить выборочные данные на две части: обучающую и контролирующую. В обучающую выборку включить наблюдения за период с 1946 по 1962 гг., в контролирующую – данные за 1963 г. Результаты оценивания следующие:

$$S_t = \underset{(0,0002)}{0,003} \cdot Y_t^2 + \underset{(0,190)}{\hat{\varepsilon}_t} .$$

Вычислить прогноз величины сбережений на 1963 г.:

$$\hat{S}_t = \hat{b} \cdot Y_t^2 = 0,003 \cdot 635,04 = 1,905.,$$

Для определения границ доверительного интервала для *индивидуального значения* зависимой переменной $S_{p=1963}$ необходимо вычислить оценку СКО:

$$s_p = s\left(1 + \frac{1}{n} + \frac{\left(x_p\right)^2}{\sum x_t^2}\right)^{0,5} = 0,19\left(1 + \frac{1}{16} - \frac{\left(635,04 - 263,381\right)^2}{354936}\right) = 0,229.$$

Границы доверительного интервала прогноза значения эндогенной переменной из контролирующей выборки, при $t_{кр}$ (0.05; 16) = 2,12, равны:

$$1,903 – 2,12 \cdot 0,228 = 1,418;$$

$$1{,}903 + 2{,}12 \cdot 0{,}228 = 2{,}389.$$

Истинное значение эндогенной переменной S_{1963} =1,99 накрывается данным интервалом, поэтому оцененная модель достаточно хорошо согласуется со статистическими данными (модель адекватна результатам наблюдений).

Задание 10. В качестве примера вычислим коэффициент детерминации R^2 и F-статистику для регрессии, оцененной в сквозном примере. Значение *TSS:*

$$TSS = \sum y_t^2 = \sum (Y_t - \bar{Y})^2 = 660{,}480.$$

В последней строке табл. 5 подсчитаны значения *ESS* = 2,306 и *RSS* = 658,178, расположенные в шестом и седьмом столбцах, соответственно. Таким образом, имеем:

$$R^2 = 1 - \frac{2{,}306}{660{,}481} = \frac{658{,}176}{660{,}481} = 0{,}996.$$

Для вычисления значения F-статистики воспользуемся формулой:

$$F = \frac{R^2}{\left(1 - R^2\right) / \left(n - 2\right)} = \frac{0{,}996}{(1 - 0{,}996) / 13} = 3701{,}286.$$

При пятипроцентном уровне значимости α = 5 %, степенях свободы $к_1$= 1 и $к_2$= n – 2 = 13 критическое значение критерия $F_{кр}$ = 4,67, т.е. $F_{выч} > F_{кр}$. Это свидетельствует о том, что значение коэффициента детерминации, близкое к единице, получено не случайно и между переменными *Y* и *X* действительно существует линейная стохастическая зависимость.

Задание 11. Бюджетное обследование пяти случайно выбранных семей дало следующие результаты (в тыс. руб.):

Таблица 7. **Данные бюджетного обследования**

Семья	**Накопления, *S***	**Доход, *Y***	**Имущество, *W***
1	3.0	40	60
2	6,0	55	36
3	5.0	45	36
4	3,5	30	15
5	1.5	30	90

Оценить регрессионную модель зависимости накоплений от дохода и имущества семьи.

Решение. В данной модели два регрессора: доход и имущество семьи, поэтому спецификация имеет вид:

$$S_t = \beta_1 + \beta_2 \cdot Y_t + \beta_3 \cdot W_t + \varepsilon_t, \quad t = 1,...,5.$$

β_j **,**j = 1,..,3 – параметры модели; ε_t – случайное возмущение, удовлетворяющее условиям Гаусса-Маркова.

МНК-оценку вектора $\beta = (\beta_1, \beta_2, \beta_3)^T$ параметров модели вычислим по формуле:

$$\widehat{\beta} = (X^T X)^{-1} X^T Y,$$

где

Y = (3,0;6,0;5,0;3,5;1,5)T

– вектор-столбец выборочных данных эндогенной переменной модели;

$$\underset{15,3}{X} = \begin{bmatrix} 1 & 40 & 60 \\ 1 & 55 & 36 \\ 1 & 45 & 36 \\ 1 & 30 & 15 \\ 1 & 30 & 90 \end{bmatrix}$$

– матрица регрессоров X. Матрица $X^T X$ и правая часть системы нормальных уравнений $X^T Y$ определяются непосредственным умножением, или по формулам (по предварительно вычисленным суммам):

$$\underset{3,5}{X^T}\,\underset{5,3}{X} = \begin{bmatrix} n & \sum X_{t2} & \sum X_{t3} \\ \sum X_{t2} & \sum X^2_{t2} & \sum X_{t2} X_{t3} \\ \sum X_{t3} & \sum X_{t2} X_{t3} & \sum X^2_{t3} \end{bmatrix} = \begin{bmatrix} 5 & 200 & 237 \\ 200 & 8450 & 9150 \\ 237 & 9150 & 14517 \end{bmatrix},$$

$$\underset{3,5}{X^T}\,\underset{5,1}{Y} = \begin{bmatrix} \sum Y_t \\ \sum Y_t X_{t2} \\ \sum Y_t X_{t3} \end{bmatrix} = \begin{bmatrix} 19,0 \\ 825,0 \\ 763,5 \end{bmatrix}.$$

Элементы обратной матрицы равны:

$$(X^T X)^{-1} = \begin{bmatrix} 5,6916 & -0,1074 & -0,0252 \\ -0,1074 & 0,0024 & 0,0002 \\ -0,0252 & 0,0002 & 0,0003 \end{bmatrix}.$$

Таким образом, решением системы является вектор

$$\hat{\beta} = (X^T\ X)^{-1} X^T Y = \begin{bmatrix} 0.279 \\ 0.123 \\ -0.029 \end{bmatrix},$$

и, следовательно, оцененная модель определяется выражением

$$Y_t = 0,279 + 0,123 X_{t2} - 0,029 X_{t3} + e_t.$$

Задание 12. По данным табл. 7 оценим СКО возмущений модели. Воспользуемся оцененной моделью и рассчитаем оценки значений эндогенной переменной для всех семей (t = 1,...,5). Результаты разместим в третьем столбце табл. 8.

В четвертом столбце данной таблицы приведены значения вектора остатков, далее, в пятом столбце, рассчитаны квадраты остатков. В последней строке пятого столбца приведено значение суммы квадратов остатков. Теперь можно воспользоваться формулой:

$$s^2 = \tilde{\sigma}^2 = \frac{\sum e_t^2}{n-k} = \frac{e^T e}{n-k} = \frac{0,281}{5-3} = 0,141,$$

и, таким образом, оценка СКО дисперсии возмущений по выборочным данным равна

$$s = \tilde{\sigma} = \sqrt{0,141} = 0,375.$$

Таблица 8. **Вспомогательные вычисления**

Семья	**Накопления, S**	**Оценка, S**	**Остатки,** $e_t = S_t - \hat{S}_t$	$e^2{}_t$
1	**2**	**3**	**4**	**5**
1	3	3,429		

Продолжение табл. 8

2	б	5.978		
3	5	4,750		
4	3,5	3,524		
5	1.5	1,318		
			$\sum$	

Задание 13. Применяя стандартную процедуру, построим доверительный интервал, который с доверительной вероятностью 0,95 накроет истинные значения параметров β_j **,**j = 1,..,3, модели. Критическое (табличное) значение t-статистики равно t(0,95;2) = 4,303. Вычислим оценки СКО оценок $\hat{\beta}_j$ **,**j = 1,..,3 параметров модели через диагональные элементы матрицы $\hat{C}_{\hat{\beta}\hat{\beta}}$:

$$\hat{C}_{\hat{\beta}\hat{\beta}} = s^2 \cdot (X^T X)^{-1} = 0,141 \begin{bmatrix} 5,6916 & -0,1074 & -0,0252 \\ -0,1074 & 0,0024 & 0,0002 \\ -0,0252 & 0,0002 & 0,0003 \end{bmatrix},$$

таким образом,

$$s_{\hat{\beta}_1} = \sqrt{0,141 \cdot 5,6916} = 0,896,$$
$$s_{\hat{\beta}_2} = \sqrt{0,141 \cdot 0,0024} = 0,018,$$
$$s_{\hat{\beta}_1} = \sqrt{0,141 \cdot 0,0003} = 0,007.$$

По формуле вычислим границы доверительных интервалов для параметров модели:

• для параметра β_1**:**

$$\widehat{\beta}_1 - t_\alpha \cdot s_{\widehat{\beta}_1} = 0,279 - 4,303 \cdot 0,896 \approx -3,576,$$
$$\widehat{\beta}_1 + t_\alpha \cdot s_{\widehat{\beta}_1} = 0,279 + 4,303 \cdot 0,896 \approx 4,134,$$

• для параметра β_2:

$$\widehat{\beta}_2 - t_\alpha \cdot s_{\widehat{\beta}_2} = 0,123 - 4,303 \cdot 0,018 \approx 0,045,$$
$$\widehat{\beta}_2 + t_\alpha \cdot s_{\widehat{\beta}_2} = 0,123 + 4,303 \cdot 0,018 \approx 0,201,$$

• для параметра β_3:

$$\widehat{\beta}_3 - t_\alpha \cdot s_{\widehat{\beta}_3} = -0,029 - 4,303 \cdot 0,0003 \approx -0,030,$$
$$\widehat{\beta} + t_\alpha \cdot s_{\widehat{\beta}_3} = -0,029 + 4,303 \cdot 0,0003 \approx -0,028.$$

Задание 14. Для проверки статистической значимости каждого коэффициента множественной регрессии из задания 11 воспользуемся t-статистикой:

$$t = \frac{\widehat{\beta}_i}{s_{\widehat{\beta}_i}} \approx t(n-k),$$
$$t_1 = \frac{\widehat{\beta}_1}{s_{\widehat{\beta}_1}} = \frac{0,279}{0,896} \approx 0,31,$$
$$t_2 = \frac{\widehat{\beta}_2}{s_{\widehat{\beta}_2}} = \frac{0,123}{0,0183} \approx 6,69,$$
$$t_3 = \frac{\widehat{\beta}_3}{s_{\widehat{\beta}_3}} = \frac{-0,029}{0,0068} \approx -4,33.$$

Сравним вычисленные значения с критическим, выбранным из таблиц t-распределения по числу степеней свободы n-k = 5 -3 = 2 и уровню значимости α=0,05:

$$t_{кр}(0,95;2) = 4,303,$$

$|t_1| = 0,31 < t_{кр}$ – оценка параметра β_1, статистически незначима;

$|t_2| = 6,69 > t_{кр}$ – оценка параметра β_2, статистически значима;

$|t_3| = 4,33 > t_{кр}$ – оценка параметра β_3, статистически значима.

Задание 15. Проверим качество оцененной модели задания 11. Воспользуемся коэффициентом детерминации, значение которого рассчитаем по формуле:

$$R^2 = 1 - \frac{ESS}{TSS} = 1 - \frac{e^T e}{y^T y},$$

где $e^T e = \sum_{t=1}^{5} e_t^2 = 0,281.$

Вычисление *TSS* оформим в виде табл. 3.3.

Таблица 9. **Вычисление *TSS***

S_t	$S_t - \bar{S}$	$(S_t - \bar{S})^2$
3	-0,8	0,64
6	2,2	4,84
5	1,2	1,44
3,5	-0,3	0,09
1,5	-2,3	5,29
$\bar{S} = 3,8$	$\sum$	**12,30**

TSS = 12,30. Таким образом:

$$R^2 = 1 - \frac{0,281}{12,3} = 1 - 0,023 = 0,977,$$

что указывает на хорошее качество регрессионной модели (регрессоры хорошо объясняют значения эндогенной переменной). Проверим статистическую значимость коэффициента детерминации. Рассчитаем F-статистику по формуле:

$$F = \frac{RSS/(k-1)}{ESS/(n-k)} = \frac{(TSS - ESS)/(k-1)}{ESS/(n-k)} = \frac{12,019/2}{0,281/2} = 42,722.$$

Критическое значение для доверительной вероятности 0,95 равно

$$F_{0,95}(2,2) = 19,0.$$

Таким образом, $F_{выч} > F_{кр}$, и коэффициент детерминации признается статистически значимым. Скорректируем значение коэффициента детерминации с учетом числа регрессоров:

$$R^2 = 1 - \frac{e^T e/(n-k)}{y^T y/(n-1)} = 1 - \frac{0,281/2}{12,3/4} = 0,954,$$

что несколько меньше, чем обычный коэффициент детерминации.

Задание 16. Имеются данные о динамике товарооборота и доходов населения России за 1997-1999 гг. Оценить регрессионную зависимость товарооборота от доходов населения в рамках модели с распределенными лагами. Максимальная величина лага, включаемая в модель, равна трем.

Таблица 10. **Исходные данные к задаче**

Месяц	**Товарооборот *Y*, % к предыдущему месяцу**	**Доходы *X* населения, % к предыдущему месяцу**
январь	91,5	79,5
февраль	92,8	100,3
март	104,3	102.9
апрель	101,5	106.6
май	97,9	92,5
июнь	98,7	110,1
июль	100.8	96,6
август	103,7	97,1
сентябрь	104,6	98,5
октябрь	100,3	105,7
ноябрь	101,5	97,4
декабрь	116	129,9
январь	82,3	63,9
февраль	91,6	104.3
март	103,4	101,7
апрель	100,3	105,5
май	99,2	91,3
июнь	99	102,6
июль	102,3	102,6
август	106,8	96,6
сентябрь	96,7	81,5
октябрь	92,7	107,8
ноябрь	100,4	69,7
декабрь	108,1	122,8

Окончание табл. 10

январь	80	63.9
февраль	96,9	107,4
март	106	103,7
апрель	97,6	108,1
май	100.2	93.9
июнь	100,7	104,1
июль	100	97,2
август	106,5	104,6
сентябрь	100,5	98,6
октябрь	102,1	104,5
ноябрь	100,5	99,9
декабрь	116	136,9

Решение. Спецификация модели включает конечное число лагов (к = 3):

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \varepsilon_t, \quad t = 4, ..., 36.$$

Воспользуемся методом замены переменных ${X^*}_{it} = X_{t-i}, \quad i = 0,1,2,3, \quad t = 4,...,36$, и вычислим значения четырех регрессоров:

$$
\begin{gathered}
{X^*}_{0t} = X_t, \quad {X^*}_{1t} = X_{t-1}, \quad {X^*}_{2t} = X_{t-2}, \quad {X^*}_{3t} = X_{t-3}, t = 4,...,36. \\
Y_t = \alpha + \beta_0 {X^*}_{0t} + \beta_1 {X^*}_{1t} + \beta_2 {X^*}_{2t} + \beta_3 {X^*}_{3t} + \varepsilon_t, \quad t = 4,...,36.
\end{gathered}
$$

Схема Гаусса-Маркова:

$$Y_4 = \alpha + \beta_0 X_4 + \beta_1 X_3 + \beta_2 X_2 + \beta_3 X_1 + \varepsilon_4,$$
$$Y_5 = \alpha + \beta_0 X_5 + \beta_1 X_4 + \beta_2 X_3 + \beta_3 X_2 + \varepsilon_5,$$
$$Y_6 = \alpha + \beta_0 X_6 + \beta_1 X_5 + \beta_2 X_4 + \beta_3 X_3 + \varepsilon_6,$$
$$\dots$$
$$Y_{36} = \alpha + \beta_0 X_{36} + \beta_1 X_{35} + \beta_2 X_{34} + \beta_3 X_{33} + \varepsilon_{36}.$$

Результаты оценивания спецификации в стандартном виде:

$$Y_t = \underset{(21,3)}{46,8} + \underset{(0,1)}{0,4}\, X^*_{0t} + \underset{(0,086)}{0,1}\; X^*_{1t} - \underset{(0,09)}{0,042}\, X^*_{2t} + \underset{(0,074)}{0,031}\, X^*_{3t} + \underset{(4.687)}{\widehat{\varepsilon}_t}\;;$$
$$R^2 = 0,623, \quad DW = 2,28.$$

Табличное значение t-статистики для доверительной вероятности 0,95 равно:

$$t\kappa p(28) = 2,1.$$

Вычисленные значения t-статистик для оценок параметров модели равны:

$$t_\alpha = 2,20, \quad t_{\beta_0} = 6,53, \quad t_{\beta_1} = 1,37, \quad t_{\beta_2} = -0,47, \quad t_{\beta_3} = 0,42,$$

и, следовательно, значимыми являются только константа и оценка параметра $\widehat{\beta}_0$ при текущем значении регрессора.

Задание 17. Проверим идентифицируемость уравнений модели «равновесного рынка»:

$$Y_t^d = a_0 + a_1 \cdot P_t + a_2 \cdot X_t + \upsilon_t^d,$$
$$Y_t^s = b_0 + b_1 \cdot P_t + b_2 \cdot P_{t-1} + \upsilon_t^s,$$
$$Y_t^d = Y_t^s,$$

где величины спроса Y_t^d, предложения Y_t^s, цены P_t, дохода X_t и случайных возмущений υ_t^d, υ_t^s относятся к текущему моменту *t*.

$Y_t = (Y_t^d, Y_t^s, P_t)^T$ – вектор-столбец эндогенных переменных,

$X_t = (1,\ X_t,\ P_{t-1})^T$ – расширенный вектор-столбец предопределенных переменных.

Проверим *необходимое условие идентификации* каждого уравнения. Для этого выпишем следующие характеристики системы:

- число предопределенных переменных в системе $\kappa = 3$ (1, X_t – доход, P_{t-1} – лаговое значение цены);
- число эндогенных переменных в системе $m = 3$ (P_t – цена, Y_t^d – величина спроса, Y_t^s – величина предложения).

Характеристики уравнений:

1) число предопределенных переменных: $p_1 = 2(1, X_t)$; $p_2 = 2(1, P_{t-1})$;

2) число эндогенных переменных: $q_t = q_2 = 2(P_t, Y_t)$.

Проверка на идентификацию.

1. Первое уравнение точно идентифицируемо по критерию порядкового условия, так как неравенство $k - p_1 \ge q_1 - 1$ приводит к точному равенству:

$$3-2 = 2-1 = 1.$$

2. Второе уравнение точно идентифицируемо:

$$k - p_2 \ge q_2 - 1; \quad 3 - 2 = 2 - 1 = 1.$$

Задание 18.

Исследуется зависимость вида:

$$\begin{cases} y_1 = f_1(y_2, x_1), \\ y_2 = f_2(y_1, x_2), \end{cases}$$

где y_1 – годовое потребление свинины на душу населения (в фунтах);

y_2 – оптовая цена за фунт (в долл.);

x_1 – доход на душу населения (в долл.);

x_2 – расходы по обработке мяса (в % к цене).

Выборочные данные за пять лет представлены в табл. 11.

Таблица 11. **Данные о годовом потреблении населения**

Год	**Y_1**	**Y_2**	**X_1**	**X_2**	**y_1**	**y_2**	**x_1**	**x_2**	$\widehat{y}_2$	$\widehat{y}_1$
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1990	60	5,0	1300	60	-3	0,6	-200	3	0,277	-2,013
1991	62	4,0	1300	56	-1	-0,4	-200	-1	-0.171	-0,954
1992	65	4,2	1500	56	2	-0,2	0	-1	-0,112	0,265
1993	62	5,0	1600	63	-1	0,6	100	6	0,702	-0,980
1994	66	3.8	1800	50	3	-0,6	300	-7	-0,696	3,681
среднее	63	4,4	1500	57	0	0	0	0		

Используя КМНК, оценить структурные параметры СОУ:

$$\begin{aligned} &y_{1t} = a_{12} \cdot y_{2t} + b_{11} x_{1t} + v_{1t}, \\ &y_{2t} = a_{21} \cdot y_{1t} + b_{22} x_{2t} + v_{2t}, \\ &t = 1, \ldots, n. \end{aligned}$$

Решение. Оба уравнения точно идентифицируемы, поэтому для оценки структурных параметров можно применить КМНК. Составим приведенную форму модели:

$$y_{1t} = m_{11} \cdot x_{1t} + m_{12} x_{2t} + u_{1t},$$
$$y_{2t} = m_{21} \cdot x_{1t} + m_{22} x_{2t} + u_{2t}.$$

Используя выборочные данные 6, 8 и 9 столбцов табл. 11, найдем МНК-оценки первого уравнения системы:

$$y_{1t} = \underset{(0,003)}{0,006} \cdot x_{1t} - \underset{(0,13)}{0,265} \cdot x_{2t} + \underset{(1,22)}{e_t};$$
$$R^2 = 0,81, \quad F = 6,6.$$

Стандартная форма второго уравнения, оцененного по данным 7, 8 и 9 столбцов:

$$y_{2t} = \underset{(0,0006)}{0,0003} \cdot x_{1t} + \underset{(0,028)}{0,112} \cdot x_{2t} + \underset{(0,25)}{e_t};$$
$$R^2 = 0,85, \quad F = 8,9.$$

КМНК имеет серьезный недостаток. Если уравнение сверхидентифицируемо, то число уравнений превышает число неизвестных, и один и тот же структурный параметр допускает разные выражения через коэффициенты приведенной формы. В данном случае применяется двухшаговый метод наименьших квадратов (ДМНК).

Задание 19. Оценить ДМНК структурные параметры модели (задание 18):

$$\begin{cases} y_{1t} = a_{12} \cdot y_{2t} + b_{11} x_{1t} + v_{1t}, \\ y_{2t} = a_{21} \cdot y_{1t} + b_{22} x_{2t} + v_{2t}, \end{cases}$$
$$t = 1, \ldots, n.$$

Решение. Необходимое условие идентификации выполняется для обоих уравнений модели, поэтому оценки ДМНК должны совпадать с оценками КМНК. Проверим это свойство.

- Оценивание параметров *первого уравнения* системы.

В соответствии с алгоритмом ДМНК на *первом шаге* оценим по выборочным данным табл. 11 (столбцы 7, 8, 9) регрессию:

y2=m21x1+m22x2+u2,

в обозначениях: $Y^{(1)} = y_2, X = (x_1, x_2), M_1 = (m_{21}, m_{22})^T$.

Оцененная спецификация имеет вил:

$$y_{2t} = \underset{(0,0006)}{0,0003} \cdot x_{1t} + \underset{(0,028)}{0,112} \cdot x_{2t} + \underset{(0,25)}{e_t};$$
$$R^2 = 0,85, \quad F = 8,9.$$

Оценки значений вектора

$$\hat{y}_{2t} = \underset{(0,0006)}{0,0003} \cdot x_{1t} + \underset{(0,028)}{0,112} \cdot x_{2t},$$
$$t = 1, \ldots, 5,$$

приведены в табл. 11 в столбце 10.

На *втором шаге* двухшагового метода по данным 6, 10 и 8 столбцов табл. 11 оценим регрессию вида:

$$y_{1t} = a_{12} \cdot \hat{y}_{2t} + b_{11} x_{1t} + \varepsilon_{1t},$$
$$t = 1, \ldots, 5,$$

в обозначениях: $Y_1 = y_1, \hat{Y}^{(1)} = \hat{y}_2, X^{(1)} = x_1, A^{(1)} = a_{12}, B^{(1)} = b_{11}$.

Результат оценивания:

$$y_{1t} = \underset{(1,247)}{-2,363} \cdot \hat{y}_{2t} + \underset{(0,003)}{0,0067} \cdot x_{1t} + \underset{(1,218)}{e_t};$$

$$R^2 = 0,81, \quad F = 6,6.$$

Для того чтобы перейти к оцененной форме первого структурного уравнения системы, необходимо вычислить оценки СКО возмущений и оценок параметров. Состоятельные оценки СКО возмущений S = 1,218 и СКО оценок параметров $S_{a_{12}} = 1,247, S_{b_{11}} = 0,003$ вычислены в ЛИНЕЙН через остатки, в которых используются значения оценок эндогенной переменной $\hat{y}_{2t}$:

$$e_t = y_{1t} - \left(\hat{a}_{12} \cdot \hat{y}_{2t} + \hat{b}_{11} \cdot \hat{x}_{1t}\right),$$

а должны быть использованы наблюдения y_{2t}, т. е.

$$e_t = y_{1t} - \left(\hat{a}_{12} \cdot y_{2t} + \hat{b}_{11} \cdot x_{1t}\right).$$

Предварительные вычисления представим в виде табл. 12.

Таблица 12. **Вспомогательная таблица к заданию 19**

y_{2t}	$\hat{y}_{1t}$	e_t	e^2_t
-3	-2.775	-0,225	0,050
-1	-0,413	-0.587	0,345
2	0,473	1,527	2,333
-1	-0.739	-0,261	0.068
3	3,454	-0,454	0,206
		$\sum$	**3,003**

Определим несмещенную оценку СКО возмущения по формуле:

$$\hat{\sigma} = \sqrt{\sum_{t=1}^{n} e_t^2 / (n-k)} = \sqrt{\frac{3,003}{5-2}} = 1,001.$$

Оцененное первое уравнение модели имеет вид

$$y_{1t} = \underset{(1,025)}{-2,363} \cdot y_{2t} + \underset{(0,002)}{0,0067} \cdot x_{1t} + \underset{(1,001)}{e_t}.$$

Описать полученный результат

- Оценки параметров *второго уравнения* системы.

В соответствии с алгоритмом ДМНК на *первом шаге* оценим по выборочным данным табл. 11 (столбцы 6, 8, 9) регрессию:

y1=m11x1+m12x2+u1,

в обозначениях: $Y^{(2)} = y_1, X = (x_1, x_2), M_1 = (m_{11}, m_{12})^T$.

После вычисления оценок СКО возмущения и оценок параметров, результат оценивания второго уравнения модели принимает вид:

$$\hat{y}_{2t} = \underset{(0,123)}{0,048} \cdot y_{1t} + \underset{(0,056)}{0,125} \cdot x_{2t} + \underset{(0,286)}{e_t}.$$

Таким образом, оценки структурных параметров ДМНК следующие:

$$a_{11} = 1, \quad \hat{a}_{12} = -2,363, \hat{b}_{11} = 0,0067, b_{12} = 0,$$

$$\hat{a}_{21} = 0,048, \; a_{22} = 1, \quad \hat{b}_{21} = 0, \quad \hat{b}_{22} = 0,125.$$

Задание 20.

Имеется выборка из 540 респондентов.

Изучаем зависимость окончания школы для индивидов в зависимости от ряда факторов.

Определим бинарную зависимую переменную:

GRAD = 1, если индивид закончил школу;

GRAD = 0, в противном случае.

Для этих респондентов фиксировались также:

- совокупные результаты тестирования познавательных способностей: ASVABC (по 100-балльной шкале);
- число лет обучения матери респондента SF;
- число лет обучения отца SM;
- пол респондента MALE (фиктивная переменная, 1 – для мужчин, 0 – для женщин).

В таблице приведены результаты оценивания логит-модели. Использовался ММП, несколько итераций. Также в нижней части таблицы рассчитаны предельные эффекты.

Таблица 13. **Результаты оценивания логит-модели**

Переменная	Коэф.	Станд.ош.	z	P>|z|	Доверит.инт.	
ASVABC	0,1329	0,024	5,42	0,00	0,085	0,18
SM	-0,023	0,087	-0,23	0,79	-0,19	0,15
SF	0,0122	0,072	0,15	0,87	-0,13	0,15
MALE	0,128	0,398	0,32	0,75	0,65	0,9
_cons	-3,25	1,066	-3,00	0,001	-5,34	-1,15

Логит-оценивание: зависимая переменная *:GRAD*

Переменная	Среднее	*b*	Среднее x *b*	*f(Z)*	*bf(Z)*
ASVABC	51,25	0,13	6,6625	0,0273	0,0037

SM	11,2	-0,02	-0,224	0,0273	-0,0007
SF	11,5	0,01	0,115	0,0273	0,003
MALE	0,6	0,12	0,072	0,0273	0,0036
Постоянный член	1,0	-3,25	-3,25	0,0273	
Итого			3,5143		

Некоторые пояснения к построенной модели:

$$Z = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + \beta_5 MALE.$$

Средние значения по каждой переменной приведены в нижней части таблицы. С их помощью рассчитывается среднее значение Z по приведенной выше формуле и имеющимся оценкам параметров, оцененных ММП. Затем считается

$$f(\overline{Z}) = \frac{e^{-\overline{Z}}}{\left(1+e^{-\overline{Z}}\right)^2} = \frac{e^{-3,5143}}{\left(1+e^{-3,5143}\right)} = 0,0281$$

.

Теперь считаем предельные эффекты. Для примера рассчитаем предельный эффект для ASVABC:

$$\frac{\partial p}{\partial ASVABC} = f(\overline{Z}) \cdot \beta_{ASVABC} = 0,0281 \cdot 0,1329 = 0,0037$$

Полученное значение говорит о том, что при увеличении результата теста индивида на 1 балл, вероятность окончить школу возрастет на 0,37% пункта. Из таблицы также видно, что от уровня образования родителей эффект незначителен.

Анализ кредитоспособности заемщиков на основе построения эконометрической логит-модели с использованием бинарной зависимой переменной путем адаптации методики анализа пяти основных финансовых показателей предприятия, характеризующих ликвидность, финансовую устойчивость, деловую активность, обслуживание долга и рентабельность [6, с. 215; 7]:

- D – коэффициент финансовой зависимости характеризует зависимость фирмы от внешних займов. Рассчитывается как отношение собственного капитала к заемным средствам. Оптимальное значение для показателя – 0,5.
- D_2 – коэффициент рентабельности, который показывает долю прибыли в каждом заработанном рубле. Обычно рассчитывается как отношение чистой прибыли (прибыли после налогообложения) за определенный период к выраженному в денежных средствах объему продаж за тот же период.
- D_3 – коэффициент финансовой устойчивости показывает, какая часть актива финансируется за счет устойчивых источников, рассчитывается как отношение суммы собственного капитала и долгосрочных кредитов и займов к валюте баланса. Рекомендуемое же значение – не менее 0,75.
- D_4 – коэффициент оборачиваемости оборотных активов показывает активность использования и скорость обращения оборотных активов. Рассчитывается как отношение выручки к среднегодовой стоимости оборотных активов (дебиторская задолженность, денежные средства, запасы и расходы будущих периодов, краткосрочные финансовые вложения).
- D_5 – коэффициент покрытия активов измеряет способность организации погасить свои долги за счет имеющихся активов. Коэффициент покрытия активов = ((Активы –

Нематериальные активы) – (Краткосрочные обязательства – Краткосрочные кредит и займы)) / Обязательства. В промышленности нормальным считается коэффициент покрытия активов не менее 2, в обслуживающих компаниях – 1,5.

Модель бинарного выбора предполагает, что зависимая переменная модели принимает только два альтернативных значения: 1 и 0. Поэтому такую модель удобно применять при прогнозировании решения заимодателя о выдаче кредита: бинарная зависимая переменная принимает значение единицы в случае положительного решения и значение нуля в случае отказа. Тем самым вектор $Y = (y_1, y_2, y_n)$ исходных данных будет содержать только бинарные признаки 0 и 1

В качестве функции $F(*)$ используются известные функции распределения вероятностей. Выбор функции F(*) определяет тип бинарной модели. Если используют функцию логистического распределения, то модель бинарного выбора называют логит-моделью (logit model)

$$P(y_t = 1) = p_t = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_t)}} = \frac{1}{1 + e^{-z}}$$

$$P(y_t = 0) = 1 - p_t = \frac{1}{1 + e^{z_t}}$$

$$\frac{p_t}{1 - p_t} = \frac{1 + e^{z}}{1 + e^{-z}} = e^{z_t}$$

$$ln \frac{p_t}{1 - p_t} = z_t = \beta_0 + \beta_1 x_t$$

Для построения бинарной логит-модели кредитоспособности была сформирована обучающая выборка из пятидесяти предприятий промышленности и оптовой торговли на основе сайта раскрытия корпоративной информации «Интерфакс»:

1) $D_1 = 0{,}534$; $D_2 = 0{,}068$; $D_3 = 0{,}762$; $D_4 = 3{,}234$; $D_5 = 2{,}106$;

2) $D_1 = 0{,}670$; $D_2 = 0{,}098$; $D_3 = 0{,}773$; $D_4 = 1{,}480$; $D_5 = 1{,}890$;

3) $D_1 = 0{,}375$; $D_2 = 0{,}102$; $D_3 = 0{,}810$; $D_4 = 6{,}653$; $D_5 = 1{,}172$.

В ходе исследования методом максимального правдоподобия получены оценки логит-модели.

$$Y = -5{,}648 - 20{,}361D_1 + 0{,}458D_4 + 9{,}831D_5 + \varepsilon;$$

$$Y = -5{,}648 - 20{,}361D_1 + 0{,}458D_4 + 9{,}831D_5 + \varepsilon;$$

Оценить целесообразность предоставления кредита претендентам.

СПИСОК ЛИТЕРАТУРЫ

1. Велицкая, С.В. Корреляционно-регрессионный анализ влияния показателей на выручку предприятия (на примере ПАО «Газпром») / С.В. Велицкая // Вектор экономики. – 2019. – № 6.
2. Бабешко, Л.О. Основы эконометрического моделирования: учебное пособие / Л.О. Бабешко. – Изд. 3-е, стереотипное. – М.: Комкнига, 2007. – 432 с.
3. Айвазян С.А. Эконометрика-2. Продвинутый курс с приложениями в финансах: учебник / С.А. Айвазян. – М.: Магистр, 2015. – 944 с.
4. Бородич, С.А. Эконометрика. Практикум / С.А. Бородич. – М: ИНФРА-М, 2014. – 329 с.
5. Герасимов, А.Н. Эконометрика. Теория и практика / А.Н. Герасимов, А.В. Гладилин. М.: Изд-во КноРус, 2011.
6. Гладилин, А.В. Эконометрика / А.В. Гладилин, А.Н. Герасимов, Е.И. Громов. – М: Феникс, 2011. – 304 с.
7. Костромин, А.В. Эконометрика / А.В. Костромин. – М.: Изд-во КноРус, 2015. – 232 с.
8. Новиков, А.И. Эконометрика / А.И. Новиков. – М: ИНФРА-М, 2014. –- 272 с.
9. Соколов, Г.А. Эконометрика. Теоретические основы: учеб. пособие / Г.А. Соколов. – М: ИНФРА-М, 2012. – 216 с.
10. Эконометрика / И.И. Елисеева [и др.] ; под редакцией И.И. Елисеевой. – М.: Изд-во Юрайт, 2019. – 449 с.