



10. Численность населения по возрастным группам на начало 2016 г. по Саратовской области // Официальный сайт Территориального органа Федеральной службы государственной статистики по Саратовской области. – Режим доступа: <http://srtv.gks.ru>.

11. Bylina S., Korostelev V. Forecasting the Number of Rural Population of Able-Bodied Ages: Methodology and Program // The Village in Transition: Papers' abstracts of the XIV International scientific conference "Vlasina Encounters 2008" / ed. by D. Stevanovic. – Beograd: Mladost biro, 2008. – P. 99.

Блинова Татьяна Викторовна, д-р экон. наук, главный научный сотрудник, Федеральное государственное бюджетное

учреждение науки Институт аграрных проблем РАН. Россия.

Былина Светлана Геннадиевна, канд. экон. наук, старший научный сотрудник, Федеральное государственное бюджетное учреждение науки Институт аграрных проблем РАН. Россия.

Коростелев Вячеслав Геннадьевич, канд. экон. наук, заместитель директора по научной работе, Федеральное государственное бюджетное учреждение науки Институт аграрных проблем РАН. Россия.

410012, г. Саратов, ул. Московская, 94.

Тел.: (8452) 26-25-38.

Ключевые слова: сельское население; рождаемость; смертность; миграция; сценарное прогнозирование; демографическое развитие; риски; Саратовская область.

THE DEMOGRAPHIC PROSPECTS OF SARATOV'S COUNTRYSIDE: ANALIS AND PROJECTIONS

Blinova Tatiana Viktorovna, Doctor of Economic Sciences, Professor, Main Researcher, Institute of Agrarian Problems of the Russian Academy of Sciences, Russia.

Bylina Svetlana Gennadiyevna, Candidate of Economic Sciences, Senior Researcher, Institute of Agrarian Problems of the Russian Academy of Sciences, Russia.

Korostelev Viacheslav Gennadiyevich, Candidate of Economic Sciences, Deputy Director on Science of the Institute of Agrarian Problems of the Russian Academy of Sciences, Russia.

Keywords: rural population; natality; mortality; migration; scenario forecasting; demographic development; risks; Saratov Oblast.

The paper discusses the results of forecasting the dynamics of the number and age structure of the rural population of Saratov Oblast. It presents medium-term forecast scenarios and a comparative analysis of them. The paper also contains an assessment of strategic demographic risks and constraints for socio-demographic development of rural areas of Saratov Oblast. The results of the forecasting show that in the medium term, the trends of reducing the number of the rural population of the Oblast in general, as well as that of the working-age population, will remain. The conclusion is made about the "fading effect" of the active demographic policy and the need to implement additional measures instrumental in "retaining" the rural population of Saratov Oblast.

УДК 338.432

СТАТИСТИЧЕСКАЯ ОЦЕНКА НАДЕЖНОСТИ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОЙ МОДЕЛИ ПОКАЗАТЕЛЕЙ ПРОИЗВОДСТВА СЕЛЬСКОХОЗЯЙСТВЕННОГО ПРЕДПРИЯТИЯ

КАНЕВСКАЯ Ирина Юрьевна, Саратовский государственный аграрный университет имени Н.И. Вавилова

Метод корреляционно-регрессионного анализа позволяет решить целый ряд необходимых эконометрических задач. Проведена проверка качества регрессионной модели зависимости объемов валового производства от различных факторов. Установлено, что обязательное включение таких элементов оценки качества, как тесты на наличие мультиколлинеарности независимых переменных, автокорреляции и гетероскедастичности остатков, позволяют значительно повысить их качество, а использование инструментария MS Excel даст возможность существенно снизить трудоемкость данных операций.

В арсенале средств современных экономических исследований довольно часто применяют метод корреляционно-регрессионного анализа, использование которого позволяет решить целый ряд необходимых эконометрических задач: определить степень влияния каждого фактора на результативный индикатор и отсеять наиболее значимые факторы; построить корреляционно-регрессионную модель исследуемого экономического процесса; определить прогнозные значения результатов управленческого воздействия на изменения факторов модели.

Современные средства автоматизации, такие как программный продукт MS Excel и входящий в него пакет «Анализ данных», позволяют при наличии полноты исходных данных с минимальными временными затратами построить корреляционную модель и представить ее в готовом виде.

Однако применение данного метода в статистическом анализе экономических процессов требует всесторонней проверки качества полученной в Excel регрессионной модели на предмет оценки ее надежности и способности с минимальными искажениями прогнозировать эко-



номические процессы. Данные меры позволяют выявить, значимо ли выбранное уравнение регрессии, подходят ли отобранные независимые переменные или необходимо исключить некоторые из них, а также подойдет ли в целом данный набор независимых переменных; не завышен ли уровень точности модели, насколько сильно отличается дисперсия случайной ошибки [2].

В качестве базы для исследования обозначенной проблематики была построена корреляционно-регрессионная модель зависимости объемов валового производства от ряда факторов.

Объектом исследования послужило сельскохозяйственное предприятие ЗАО «ПЗ «Мелиоратор» Марковского района Саратовской области.

Рассматривались следующие влияющие факторы:

X_1 – среднемесячный размер заработной платы одного работника;

X_2 – фондовооруженность;

X_3 – производительность труда (определяемая объемами производства валовой продукции в денежном выражении в расчете на одного работника);

X_4 – фондоотдача;

X_5 – количество занятых работников;

X_6 – объем используемых основных фондов.

Входные данные представлены в табл. 1, содержащей результирующие и факторные по-

казатели, выстроенные в соответствии с системой, предложенной автором за период с 2008 по 2015 г.

Определение зависимостей между изучаемыми показателями и отбор наиболее значимых независимых переменных являются важнейшей задачей следующего этапа оценки надежности выстраиваемой модели. Проведена оценка тесноты и направления связи между переменными величинами посредством построения матрицы парной корреляции (табл. 2). Расчетные значения получены с помощью инструмента MS Excel «Корреляция» (пакет «Анализ данных»).

Значение коэффициентов парной корреляции лежит в интервале от (-1) до $(+1)$. Положительное значение показателя свидетельствует о прямой связи между переменными, отрицательное – об обратной (когда с ростом одной переменной другая уменьшается). Качественная оценка тесноты связи между величинами может быть произведена на основе шкалы Чеддока: связь признается достаточно сильной (высокой), если коэффициент корреляции по абсолютной величине (без учета положительного или отрицательного знака) превышает 0,7, и слабой, если он меньше 0,3.

В результате расчета парных коэффициентов корреляции было выявлено, что факторы «Фондоотдача X_4 » и «Количество занятых работников X_5 » необходимо исключить из регрессионной

Таблица 1

Входные данные для корреляционно-регрессионного анализа [4]

Отчетный период, год	Объем валового производства Y, тыс. руб.	Факторные показатели					
		размер з/п X_1 , тыс. руб.	фондовооруженность X_2 , тыс. руб./чел.	производительность труда X_3 , тыс. руб./чел.	фондоотдача X_4	количество занятых работников X_5 , чел.	объем используемых основных фондов X_6 , млн руб.
2008	205 088	8,2	0,36	681,3	1,34	301	152,6
2009	219 972	10,1	0,43	766,5	1,31	287	168,4
2010	248 389	10,5	0,26	868,5	1,45	286	170,8
2011	292 693	12,4	1,00	1076,1	1,39	272	210,7
2012	300 618	13,2	4,33	1109,3	0,95	271	315,1
2013	332 552	13,7	2,48	1283,9	0,91	259	366,6
2014	363 543	15,5	4,31	1477,8	0,93	246	392,2
2015	403 938	17,3	3,46	1779,5	0,92	227	437,1

Таблица 2

Матрица парной корреляции исследуемых показателей

	Y	X_1	X_2	X_3	X_4	X_5	X_6
Y	1						
X_1	0,99	1					
X_2	0,80	0,82	1				
X_3	0,99	0,99	0,77	1			
X_4	-0,80	-0,80	-0,917	-0,78	1		
X_5	-0,99	-0,99	-0,76	-0,99	0,77	1	
X_6	0,96	0,95	0,88	0,95	-0,93	-0,94	1

Примечание: табл. 2 и все последующие таблицы и рисунок рассчитаны автором с помощью программы MS Excel.



модели, так как они имеют отрицательную корреляцию с зависимым показателем.

Вместе с тем следует отметить высокую связь между следующими факторами: размером заработной платы, производительностью труда, объемом используемых основных фондов и объемом валового производства (коэффициенты корреляции имели значения от 0,96 до 0,99). Также стоит выделить высокую связь между факторами: размером заработной платы, производительностью труда и объемом используемых основных фондов (коэффициенты корреляции составили от 0,95 до 0,99); фондовооруженностью и объемом используемых фондов (коэффициент корреляции 0,88); производительностью труда и объемом используемых фондов (коэффициент корреляции 0,95). Данные результаты свидетельствуют о вероятном возникновении явления мультиколлинеарности.

Природа мультиколлинеарности нагляднее всего проявляется, когда между объясняющими переменными существует строгая линейная связь. Это строгая мультиколлинеарность, когда невозможно разделить вклад каждой переменной в объяснение поведения результативного показателя. Мультиколлинеарность между факторами приводит к тому, что дальнейшие вычисления становятся невозможными, а содержательная интерпретация параметров модели крайне затруднительна.

Основные последствия мультиколлинеарности:

1) понижается точность оценки параметров регрессии, что проявляется в следующих аспектах: ошибки некоторых оценок становятся очень большими; выборочные дисперсии сильно возрастают;

2) коэффициенты некоторых введенных в регрессию переменных оказываются незначимыми, но в силу экономических соображений именно эти переменные должны оказывать заметное влияние на объясняемую переменную;

3) оценки коэффициентов становятся очень чувствительными к выборочным наблюдениям (небольшое увеличение объема выборки приводит к очень сильным сдвигам в значениях оценок).

Для более точного выявления данного явления необходимо проверить качество модели. Один из наиболее часто используемых вариантов проверки заключается в следующем. Для полученного уравнения регрессии определяют F -статистику (критерий Фишера) – характеристику точности уравнения регрессии, представляющую собой отношение той части дисперсии зависимой переменной, которая объяснена уравнением регрессии к необъясненной (остаточной) части дисперсии.

Для вычисления данного критерия необходимо рассчитать регрессионное уравнение на основе 4 из 6 изучаемых факторов:

X_1 – среднемесячный размер заработной платы одного работника;

X_2 – фондовооруженность;

X_3 – производительность труда (определяемая объемами производства валовой продукции в денежном выражении в расчете на одного работника);

X_4 – объем используемых основных фондов.

Инструмент «регрессия», входящий в пакет «анализ данных», позволил определить параметры искомого уравнения, имеющего вид:

$$Y = 94506,2 + 1255,5X_1 + 2057X_2 + 151,2X_3 + 41,7X_4.$$

В окне «вывод итогов» в разделе «дисперсионный анализ» регрессионного уравнения получаем фактическое значение F -критерия $F_{\text{факт}}$, равное 41,36. Критическое значение данного критерия $F_{\text{крит}}$ определяется с помощью функции «fраспобр». Для его расчета необходимо указать числа степеней свободы Df_1 и Df_2 , вычисляемых по формулам:

$$Df_1 = N;$$

$$Df_2 = M - N,$$

где N – количество независимых переменных; M – количество наблюдений.

В окне данной функции также необходимо указать значение уровня значимости (как правило, в большинстве исследований указывают стандартное значение 0,05). Расчетное значение $F_{\text{крит}}$ составило 6,388. Поскольку $F_{\text{факт}} > F_{\text{крит}}$, то регрессионная модель считается значимой.

Однако критерий Фишера свидетельствует лишь об общей жизнеспособности модели. Следовательно, для более детальной оценки ее качественных сторон необходимо прежде всего оценить, наблюдается ли явление автокорреляции остатков. Для этого обычно используют две проверочные методики: расчет критерия Дарбина – Уотсона и методику Бреуша – Годфри.

Критерий Дарбина – Уотсона рассчитывают по формуле:

$$DW = \frac{\sum (E_i - E_{i-1})^2}{\sum E_i^2},$$

где E_i – значение остатков в текущем периоде; E_{i-1} – значение остатков в предшествующем периоде.

Промежуточные данные отобразим в табл. 3.

Расчеты показали, что значение критерия Дарбина – Уотсона составило 2,53, что свидетельствует об отсутствии явления автокорреляции в модели (согласно типовым табличным данным).

Проверим наличие автокорреляции с помощью теста Бреуша – Годфри. Необходимо рас-



Промежуточные данные для измерения критерия Дарбина – Уотсона

E_i	E_{i-1}	E_i^2	$(E_i - E_{i-1})^2$
1714,9	-11025,6	–	–
9055,7	1714,9	2940 839,304	162 319 131
-245,5	9055,7	82 006 203,37	53 887 933,5
6317,4	-245,5	60 255,52616	86 512 278
896,4	6317,4	39 909 905,16	43 071 639,3
-6713,37	896,4	803 464,03	29 387 968,8
–	–	45 069 342,8	57 908 021,1

считать значение коэффициента корреляции между значениями остатков E_i и E_{i-1} и проверить его значимость с помощью T -критерия (критерия Стьюдента).

Значение наблюдаемого T -критерия вычислим по формуле:

$$T = r\sqrt{n-2} / \sqrt{1-r^2},$$

где r – коэффициент парной корреляции между значениями остатков (E_i и E_{i-1}); N – число наблюдений.

Критическое значение T находим с помощью встроенной функции «студент обр».

Таким образом, $T_{\text{факт}} = -0,25$ (при $r = -0,1$), $T_{\text{крит}} = 0,067$. Следовательно, расчетные данные свидетельствуют о том, что между остатками отсутствует явление автокорреляции, что указывает на относительную надежность построенной модели.

Рассмотрим подробнее последствия обнаружения автокорреляции в остатках модели:

оценки параметров, оставаясь линейными и несмещенными, перестают быть эффективными; дисперсии оценок являются смещенными. Зачастую дисперсии, вычисляемые по стандартным формулам, являются заниженными, что приводит к увеличению t -статистик. Это может привести к признанию статистически значимыми объясняющие переменные, которые в действительности таковыми могут и не являться [2].

В силу вышеизложенного выводы по t - и F -статистикам, определяющим значимость коэффициентов регрессии и коэффициента детерминации, возможно, будут неверными. Вследствие этого ухудшаются прогнозные качества модели.

Основной причиной наличия случайного члена в модели являются несовершенные знания о причинах и взаимосвязях, определяющих то или иное значение зависимой переменной. Поэтому свойства случайных отклонений, в том числе и автокорреляция, в первую очередь зависят от выбора формулы зависимости и состава объясняющих переменных. Так как автокорреляция чаще всего вызывается неправильной спецификацией модели, то для ее устранения необходимо прежде всего попытаться скорректировать саму модель.

Возможно, автокорреляция вызвана отсутствием в модели некоторой важной объясняющей переменной. Необходимо попытаться определить данный фактор и учесть его в уравнении регрессии. Также можно попробовать изменить формулу зависимости (например, линейную на лог-линейную, линейную на гиперболическую и т. д.).

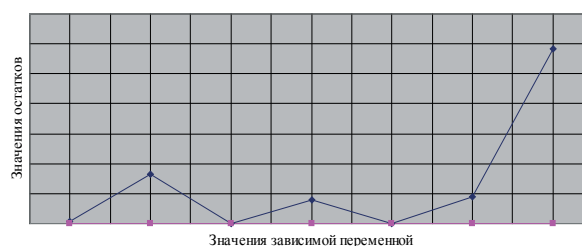
В нашем примере обе тестовые методики – критерий Дарбина – Уотсона и тест Бреуша – Годфри не обнаружили явления автокорреляции, что свидетельствует о надежности модели.

Для оценки качества регрессионной модели также необходимо выявить, имеется ли в ней зависимость дисперсий ошибок от факторных признаков (явление гетероскедастичности) или нет (гомоскедастичность). Обнаружение явления гетероскедастичности свидетельствует о снижении качества модели и искажении оценок коэффициентов регрессии за счет сокращения значения их стандартных ошибок [1].

Одной из распространенных методик, реализуемых с помощью инструментария MS Excel по проверке модели на гетероскедастичность, является графический метод, отражающий зависимость величины Y и квадратов остатков E_i^2 (см. рисунок).

График свидетельствует о наличии гетероскедастичности, так как имеет место быть непостоянство дисперсии остатков. В противном случае, при гомоскедастичности модели, область расположения точек, отражающих значения остатков, была бы на графике параллельна оси абсцисс.

В качестве расчетной методики, позволяющей обнаружить явление гетероскедастичности, воспользуемся тестом Парка. Данный тест предпо-



Графическая проверка наличия гетероскедастичности исследуемой регрессионной модели



лагает построение вспомогательного регрессионного уравнения, в котором в качестве зависимой переменной будет выступать значение натурального логарифма квадрата остатков, а факторных показателей – значения натурального логарифма переменных. Цель проведения теста – выявить статистическую значимость коэффициентов уравнения согласно критерию Стьюдента. Если коэффициенты окажутся статистически значимыми, то явление гетероскедастичности присутствует.

В табл. 4 отображены исходные данные для проведения теста Парка.

Воспользуемся функцией MS Excel «Регрессия» для получения вспомогательного уравнения и расчета соответствующих коэффициентов. В результате получаем следующие выходные данные (табл. 5).

Выходные данные свидетельствуют о незначимости коэффициентов вспомогательного регрессионного уравнения, поскольку значения плотности вероятности распределения Стьюдента (P -значение) приближены к единице, следовательно, данный тест определил отсутствие явления гетероскедастичности в представленной модели.

Таким образом, проверка качества регрессионной модели зависимости объемов валового производства от различных факторов, позволила выявить следующие результаты:

согласно критерию Фишера, модель является статистически значимой;

проведенные тест на критерий Дарбина-Уотсона и тест Бреуша – Годфри не выявили явления автокорреляции остатков модели;

графический анализ отразил наличие гетероскедастичности модели, однако для более детальной проверки проведен тест Парка, который влияние данного явления на модель.

В целом, помимо использования традиционного арсенала проверочных средств статистической значимости корреляционно-регрессионных моделей (критерии Фишера и Стьюдента), обязательное включение таких элементов оценки качества, как тесты на наличие мультиколлинеарности независимых переменных, автокорреляции и гетероскедастичности остатков, позволят значительно повысить их качество, а использование инструментария MS Excel – существенно снизить трудоемкость данных операций.

Представленный подход к изучению и исследованию данного вопроса позволяет прогнозировать – предсказывать состояния какого-либо объекта, процесса или явления в будущем. Фактически любое моделирование дает прогноз, т.е. отвечает на вопрос: «Что будет, если...?». Если, например, увеличить на 20 % значение каждого фактора отчетного периода (чем больше величина, тем более заметна чувствительность влияния каждого фактора), то можно получить следующие результаты: рост заработной платы X_1 на данную величину приведет к росту объема валового производства с 403,9 млн до

Таблица 4

Исходные данные для проведения Теста Парка

$\ln E_i^2$	$\ln X_1$	$\ln X_2$	$\ln X_3$	$\ln X_4$
18,62	2,10	-1,02	6,52	5,03
14,89	2,31	-0,84	6,64	5,13
18,22	2,35	-1,35	6,77	5,14
11,01	2,52	0,00	6,98	5,35
17,50	2,58	1,47	7,01	5,75
13,60	2,62	0,91	7,16	5,90
17,62	2,74	1,46	7,30	5,97
19,49	2,85	1,24	7,48	6,08

Таблица 5

Выходные данные для проведения теста Парка

	t -статистика	P -значение
У-пересечение	-0,325 273 552	0,766 338 975
Переменная X_1	-0,390 387 382	0,722 319 394
Переменная X_2	-0,375 698 086	0,732 133 997
Переменная X_3	0,119 646 561	0,912 325 789
Переменная X_4	0,522 659 996	0,637 354 355

415,0 млн руб.; рост фондовооруженности X_2 повлечет увеличение зависимой переменной с 403,9 млн до 412,0 млн руб.; рост производительности труда X_3 на 20 % повлияет на увеличение объема валового производства с 403,9 млн до 464,4 млн руб.; увеличение состава основных средств X_4 на 20 % приведет к расширению объема валового производства с 403,9 млн до 414,3 млн руб. Посредством полученной модели можно определить степень влияния каждого фактора на объем валового производства ЗАО «ПЗ «Мелиоратор». Следовательно, согласно представленной модели, результирующий показатель наиболее чувствителен к фактору производительности труда. Чтобы достичь обозначенной цели – добиться роста объема валового производства на 15 % в плановом периоде – руководству предприятия рекомендуется направить усилия на поиск необходимых резервов и путей увеличения производительности труда до величины 20 %:

сократить до минимума время необоснованных простоев техники и оборудования;

увеличить автоматизацию трудоемких производственных процессов;

повысить интенсивность труда работников основного производства, а также их квалификационный уровень;

провести дальнейшие усовершенствования уровня химизации и мелиорации.

Тем не менее, с помощью построенной модели можно получить лишь сугубо приближенные данные прогнозных значений объема валового производства ЗАО «ПЗ «Мелиоратор», поскольку корреляционный анализ выявил мультиколлинеарность влияющих переменных. Следовательно, для построения более надежной и качественной эконометрической модели, с авторской точки зрения, в данном случае необходимо принять во внимание более широкий спектр влияющих факторов и расширить количественный состав независимых переменных, либо исключить из нее менее существенные объясняющие переменные.

Математические модели имеют двоякую особенность [3]. Во-первых, при описании реальных процессов и систем они действи-

тельность упрощают – идеализируют. Сама жизнь гораздо сложнее, чем любая модель, с помощью которой мы пытаемся реальность проанализировать и прогнозировать ее будущее. Исследователи это понимают и не строят иллюзий относительно точности своего анализа и степени достоверности своего прогноза: все выводы и предсказания теории содержат некоторую погрешность, которую мы показали в своих расчетах. Отдавая себе в этом отчет и не претендуя на абсолютную точность, при построении модели мы учитывали главные факторы и отбрасывали второстепенные, несущественные, это понимали еще классики экономической науки Леон Мати Эспри Вальрас, Уильям Стэнли Джевонс и др. Построить теорию легче, чем научиться использовать ее в реальных ситуациях. Нужен период адаптации теории к практике.

СПИСОК ЛИТЕРАТУРЫ

1. Андреев В.И., Исаева Т.А. Использование методических подходов А. Лафера в анализе влияния налогового бремени на динамику макроэкономических показателей // Аграрный научный журнал. – 2016. – № 2. – С. 66–70.
2. Воротников И.Л., Розанов А.В., Котова М.В. Мониторинг динамики и прогнозирование уровня потребительских цен на продукты питания в Саратовской области // Аграрный научный журнал. – 2015. – № 7. – С. 71–75.
3. Красс М.С., Чупрынов Б.П. Математика в экономике. Математические методы и модели: учебник. – М.: Финансы и статистика, 2007. – 544 с.
4. Территориальный орган Федеральной службы государственной статистики по Саратовской области. – Режим доступа: http://srtv.gks.ru/wps/wcm/connect/rosstat_ts/srtv/ru/reporting/.

Каневская Ирина Юрьевна, канд. с.-х. наук, доцент кафедры «Математика и математическое моделирование», Саратовский государственный аграрный университет имени Н.И. Вавилова. Россия.

410012, г. Саратов, Театральная пл., 1.
Тел.: (8452) 23-32-92.

Ключевые слова: корреляционно-регрессионная модель; показатели производства; оценка надежности.

STATISTICAL EVALUATION OF THE RELIABILITY OF THE CORRELATION AND REGRESSION MODEL OF PRODUCTION FIGURES OF AGRICULTURAL ENTERPRISE

Kanevskaya Irina Yurievna, Candidate of Agricultural Sciences, Associate Professor of the chair "Mathematics and Mathematical Modeling", Saratov State Agrarian University named after N. I. Vavilov. Russia.

Keywords: correlation and regression model; production figures; assessment of reliability.

Method of correlation and regression analysis allows solving a range of econometric problems. The checking

the quality of regression model of the dependence of the gross production from a variety of factors is performed. It is established that the mandatory inclusion of such elements of quality assessment, as the tests for the presence of multicollinearity of the independent variables, autocorrelation and heteroscedasticity of the residuals significantly improve their quality and the use of the tools of MS Excel will allow to significantly reduce the complexity of these operations.

