

*И.А. Бушмин, А.Б. Панюкова***Прогнозирование показателей занятости населения в Алтайском крае на основе интервального представления временных рядов****I.A. Bushmin, A.B. Panyukova***Forecasting Population Employment Indicators in the Altai Territory on the Basis of Interval Representation of Time Series**

В статье рассмотрено прогнозирование показателей занятости. За основу взято интервальное представление временных рядов. Также в результате факторного анализа определены основные социально-экономические параметры, влияющие на занятость населения в Алтайском крае. Базой для решения поставленной задачи служит система математических моделей.

Ключевые слова: моделирование, безработица, занятость, полином, горизонт прогнозирования.

Введение. Прогнозирование показателей занятости – одна из важных задач экономического анализа рыночной ситуации, используется для принятия обоснованных стратегических решений как органами государственной власти, так и крупными бизнес-структурами.

Обоснованным требованием к выполнению прогноза выступает полнота охвата показателей, с помощью которых реализуется экономическая ситуация, и точность полученных прогнозов на год, два и более лет.

В данной статье рассматривается прогноз основных показателей занятости населения регионального рынка труда. Полнота показателей прогноза, характеризующих занятость населения на уровне субъекта федерации, в качестве которого выступает Алтайский край, зависит от их иерархической системы.

Система показателей включает количество безработных, численность населения, состояние рынка, уровень заработной платы, сферу подготовки.

Рассмотрим проблему точности прогноза. Известно, что точность обеспечивается обоснованным выбором уравнений прогноза и факторов, а также методикой обработки данных [1, 2].

Большинство прогнозных показателей осуществляется путем экстраполяции временных рядов. Указанный подход является естественным, поскольку данные временных рядов несут определенную информацию о развитии процессов в настоящем и на период прогноза. Однако прогнозные возможности

This article considers employment indicators forecasting. Interval representation of time series is assumed as a basic. Also as a result of the factorial analysis, the basic social and economic parameters influencing on the population employment in the Altai Territory were determined. The system of mathematical models serves as a base for the decision of this task.

Key words: modelling, unemployment, employment, polynomial, forecasting horizon.

рассматриваемых совокупностей временных рядов не оценены. В литературе по математической статистике и теории случайных процессов методики таких оценок отсутствуют либо носят сложный характер и требуют дополнительной информации (о распределении вероятностных процессов), которые в реальных условиях трудно получить.

Указанная задача решается с использованием нестатистических подходов к анализу данных и применительно к прогнозу временных рядов изложена в работе [3].

В данной статье решается задача прогнозирования показателей службы занятости населения с использованием этого интервального подхода.

В качестве исходных данных рассмотрены временные ряды численности безработных по Алтайскому краю за 1999–2010 гг., а результаты предложены к использованию в Управлении Алтайского края по труду и занятости населения.

Описание методики анализа временных рядов. Традиционно обработка дискретных временных рядов осуществляется методом наименьших квадратов с использованием авторегрессионных моделей, представлением закономерностей от времени или их сочетанием.

В работе [3] используется только способ полиномиальных закономерностей полинома степени n по времени t .

$$y(t) = \sum_{i=0}^n A_i t^i. \quad (1)$$

В данном случае построение уравнений прогноза (1) сводится к оценке коэффициентов полинома (1) методом наименьших квадратов (МНК) и выбором степени полинома, который определяется за прогнозируемым процессом $\{y(t)\}, t = \overline{1, T}$.

* Данная работа выполнена при финансовой поддержке РФФИ (проект №10-01-98005 р_сибирь_a) и ведомственно-аналитической программы «Развитие научного потенциала высшей школы 2009–2010 гг.» (проект №2.2.2.4/4278).

Численность безработных на начало периода

t	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
y(t)	24962	25660	22880	28587	40124	44654	52936	74221	76259	61379	42077	56480

Пример такого временного ряда представлен в таблице 1.

Структурные возможности функций можно получить, если представить, что процесс {y(t)} подчинен обыкновенному дифференциальному уравнению с ненулевой правой частью F(t). Разностная аппроксимация для n = 1 имеет следующий вид:

$$y(t) = a_0 y(t-1) + F(t). \quad (2)$$

Авторегрессия получается из модели (2) когда F ≡ 0, т.е. когда отсутствуют внешние воздействия.

Модель (1) получается, когда a₀ = 0.

В этом случае закономерность изменений временного ряда определяют только внешние причины. Таким образом, при построении уравнения прогноза обоснованию подлежат структура уравнения (2), n, T, период наблюдения за прогнозируемым процессом, способ определения коэффициентов используемого уравнения прогноза и показателей точности ε в зависимости от периода прогноза.

Рассмотрим на примере уравнения (1) оценку коэффициентов МНК [4].

Описание метода наименьших квадратов. Задан равномерный временной ряд {y(t)}, t = 1, T, требуется подобрать полином степени N

$$P_N(t) = \sum_{i=0}^N a_i t^i,$$

в наилучшей степени соответствующий заданному ряду. Решение такой задачи сводится к нахождению коэффициентов полинома из условия минимума суммы квадратов их невязок.

Критерием отбора полиномов считается точность приближения к заданной статистике. В качестве показателя точности S(N) примем среднеквадратичную величину расхождения значений полинома P_N и временного ряда в узловых точках [3]:

$$S(N) = \sqrt{\frac{1}{N} (Q_N)_{\min}}.$$

С увеличением степени полинома (при N, близких к (T - 1)) этот показатель резко убывает и в случае точного воспроизведения (при N = T - 1) полинома в узлах становится равным нулю, т.е. S(T - 1) = 0.

Варианты приближений представлены на рисунке 1. Таким образом, с точки зрения точности приближения, казалось бы, выгодно выбирать полиномы наиболее высоких степеней, включая полиномы абсолютного точного приближения. Однако поведение полиномов высоких степеней в большинстве случаев не соответствует общим свойствам моделируемого процесса и не позволяет надеяться на составление надежного прогноза. Поэтому подходящий

полином предлагается найти эксперту, исходя из его собственных представлений о моделируемом процессе.

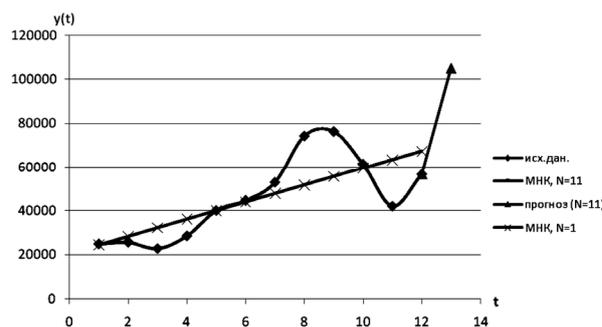


Рис 1. Приближения 1 и 11 степеней

Интервальное представление полиномиальных регрессий. Теперь рассмотрим интервальное представление формулы прогноза.

Интервальное представление базируется на подходах, описанных в работах [1, 3, 4].

Полином N-й степени в интервальном представлении можно записать следующим образом:

$$[P_N(t)] = \sum_{i=0}^N [a_i] t^i; [a_i] = [a_i^{(1)}; a_i^{(2)}].$$

Здесь точность приближения будем определять следующим образом:

$$\varepsilon^* = \min_{(a_0, \dots, a_n) \in R^{n+1}} \max_{t \in [t^1, t^2]} [F(t) - P_N(t)].$$

Примеры расчетов для моделей представлены на рисунках 2 и 3.

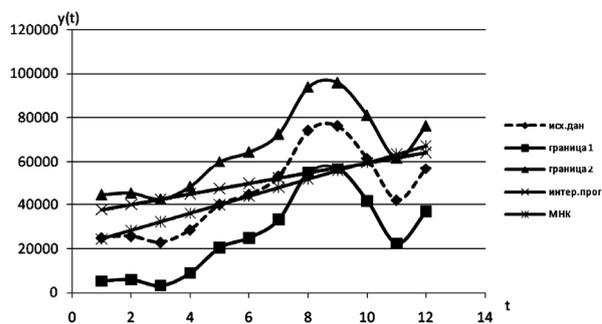


Рис. 2. Полиномы первой степени

Параметры интервал-полинома: МНК оценки:
 a₀ = 35352,6; a₀ = 20538,92;
 a₁ = 2399,62; a₁ = 3895,79;
 ε = 19671,44; ε = 11847,05.

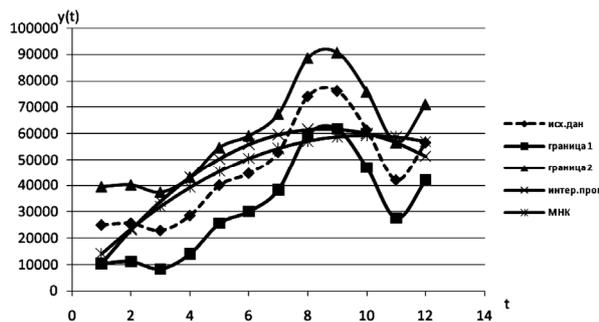


Рис. 3. Полиномы второй степени

Параметры интервал-полинома: $a_0 = -4056,25$;
 $a_1 = 15382,09$;
 $a_2 = -896,996$;
 $\epsilon = 14533,16$;
 МНК оценки: $a_0 = 3490,11$;
 $a_1 = 11202,4$;
 $a_2 = -562,05$;
 $\epsilon = 10257,57$.

Форма и параметры интервал-полиномов заметно отличаются от полиномов, построенных по методу наименьших квадратов. При этом оказалось, что показатель точности интервал-полинома превышает значение среднеквадратичного отклонения в среднем в 1,5 раза. Однако сравнение показателей точности, полученных различными методами, некорректно, во-первых, из-за отсутствия объективных правил определения ширины доверительного интервала (в среднеквадратичном приближении), а во-вторых, из-за различного вида самих полиномов приближения.

Определим горизонт прогнозирования. Горизонтом прогнозирования относительно текущего момента T понимается промежуток времени ΔT и, соответственно, точка на временной оси $(T+\Delta T)$, за пределами которой прогноз перестает удовлетворять требуемым свойствам (надежности, точности и т.п.). Оценки горизонта ΔT во многом определяются состоятельностью прогноза и возможностью метода прогнозирования.

Максимальная степень интервального полинома при заданной статистике T не должна превышать

$$N_{\max} \leq \left(\frac{T}{2} - 2\right).$$

Определим состоятельность и обоснованность прогноза. Значение горизонта прогнозирования представим в виде $K = [K_1, K_2]$, где K_1 и K_2 – нижняя и верхняя оценки горизонта соответственно, при этом $K_1(N)$ ограничивает сверху область состоятельных прогнозов, а $K_2(N)$ определяет снизу область заведомо ненаблюдаемых изменений. Полоса между K_1 и K_2 составляет зону рискованных прогнозов и решений.

K_1 можно вычислить как минимальное расстояние между точками альтернанса, K_2 – максимальное. Альтернанс (чередование) – это совокупность номеров точек ряда, находящихся на границах интервал-полинома.

Определяя горизонты прогнозирования для заданной статистики, определили, что $K_1 = 2, K_2 = 5$. Зона рискованных прогнозов – от 2 до 5. Прогноз более чем на пять шагов будет несостоятельным.

Если стоит задача прогнозирования более чем на один шаг, то ограничения на объем статистики становятся жесткими: $N \leq (T - 5) / 3$.

Учитывая вышесказанные ограничения и выполняя необходимые условия состоятельности прогноза по заданной статистике, определяем наилучшую степень полинома: $N = 3$.

Результаты расчетов представлены на рисунке 4.

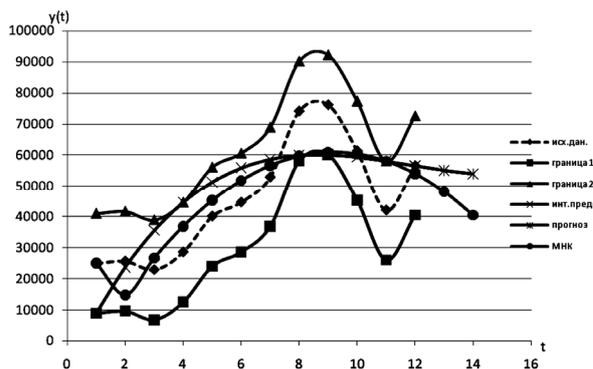


Рис. 4. Прогнозирование с помощью интервал-полинома на два шага

Параметры интервал-полинома: $a_0 = -9294,92$;
 $a_1 = 19953,08$;
 $a_2 = -1821,47$;
 $a_3 = -51,36$;
 $\epsilon = 16073,95$;
 МНК оценки: $a_0 = -14431,114$;
 $a_1 = 16388,3849$;
 $a_2 = -890,272$;
 $a_3 = 23895,1935$;
 $\epsilon = 9146,01$.

Таким образом, интервальный подход является предпочтительнее. Он позволяет обосновать выбор степени полинома, обосновать горизонт прогнозирования. Полученные прогнозы могут быть использованы для принятия управленческих решений.

Социально-экономические факторы, влияющие на рынок труда. В соответствии со сценарием экономического развития края можно говорить о прогнозе численности безработных с учетом факторов, оказывающих на это существенное влияние.

В результате факторного анализа определены основные социально-экономические параметры, влияющие на занятость населения в Алтайском крае. Введем основные исследуемые показатели рынка труда:

- x_1 – темп роста (снижения) численности занятого населения;
- x_2 – индекс физического объема промышленного производства;
- x_3 – индекс потребительских цен, в процентах к декабрю предыдущего года;
- x_4 – инвестиции в основной капитал, в процентах к предыдущему году (в сопоставимых ценах);
- x_5 – сальдо миграции населения;
- x_6 – темп роста (снижения) численности постоянного населения;
- x_7 – темп роста (снижения) общей численности безработных;

x_8 – темп роста (снижения) численности безработных, зарегистрированных в государственной службе занятости населения;

x_9 – темп роста (снижения) среднедушевых доходов населения в месяц (в сопоставимых ценах);

x_{10} – темп роста (снижения) потребности в рабочих местах, заявленной предприятиями в государственную службу занятости населения.

Анализ данных проводился с 1999 по 2008 г. с использованием программной системы Deductor [5] и прикладной программы SPSS.

Исходные данные представлены в таблице 2.

Для анализа рассчитана матрица корреляции темп роста (снижения) численности безработных, зарегистрированных в государственной службе занятости населения с остальными показателями (рис. 5).

Наблюдается тесная связь темпа роста (снижения) численности занятого населения и индекса потребительских цен (на нее указывает коэффициент парной корреляции $r_{13} = 0,633$). Между темпом роста (снижения) численности занятого населения и индексом физического объема промышленного производства (коэффициент парной корреляции $r_{12} = 0,6$). Тесная связь между темпом роста (снижения) чис-

ленности безработных граждан, зарегистрированных в службе занятости, и темпом роста (снижения) потребности в рабочих местах ($r_{810} = -0,776$ – обратная сильная связь) обусловлена реальными экономическими процессами, поскольку непосредственно от количества вакантных рабочих мест зависит численность безработных граждан. Также наблюдается тесная связь между динамикой численности безработных граждан, зарегистрированных в службе занятости, и миграцией населения ($r_{85} = 0,655$). Частный коэффициент эластичности между факторным признаком – темпом роста (снижения) потребности в рабочих местах – и результативным признаком – темпом роста (снижения) численности безработных, зарегистрированных в государственной службе занятости населения, равен $-0,932$, это значит, что при увеличении количества вакансий на 1% численность безработных граждан, зарегистрированных в службе занятости, уменьшится на 0,932%. Частный коэффициент эластичности между факторным признаком – сальдо миграции населения и результативным признаком – темпом роста (снижения) численности безработных, зарегистрированных в государственной службе занятости населения,

Таблица 2

Динамика основных социально-экономических показателей рынка труда Алтайского края

ГОД	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1999	99,7	113,3	135,4	105,4	152,7	100	84,7	111,2	137,5	54,5
2000	99,6	100,1	122,6	108,7	134,2	-26,8	91,0	64,8	139,4	357,6
2001	99,6	103,2	119,4	112,2	128,1	100,6	77,2	89,2	137,0	128,4
2002	99,2	103,8	113,1	110,3	117,7	73,4	71,6	124,9	137,3	84,7
2003	99,3	93,5	112,8	118,7	126,3	90,3	192,0	140,4	122,6	98,7
2004	99,3	106,6	112,9	99,6	135,5	81,4	76,5	111,3	111,1	112,4
2005	99,3	99,1	113,6	120	94,9	79,0	99,3	118,5	131,0	79,6
2006	99,1	101,6	107,7	122,6	105,4	88,5	99,2	140,2	136,6	86,3
2007	99,2	101,6	112,1	126,1	108,9	71,5	73,2	102,7	118,9	131,6
2008	99,4	98,5	114,5	106,7	103,7	61,1	129,8	80,5	131,1	167,9



Рис. 5. Корреляционная матрица

Матрица главных компонент

Показатель	Компонента			
	1	2	3	4
Год	-0,839	0,033	0,049	-0,334
Рост постоянного населения	0,845	-0,325	-0,048	0,218
Рост занятого населения	0,563	0,260	-0,760	-0,022
Индекс потребительских цен	0,880	-0,086	-0,182	0,268
Инвестиции в основной капитал	-0,676	0,232	0,242	0,402
Индекс промышленного производства	0,931	0,004	-0,072	-0,141
Сальдо миграции	-0,005	0,925	-0,061	-0,103
Численность общих безработных	-0,021	0,115	0,967	-0,086
Зарегистрированных безработных	-0,247	0,859	0,235	-0,008
Доходы населения	0,207	-0,158	-0,092	0,905
Потребности в работниках	0,083	-0,967	0,115	0,062

равен -0,434, это означает, что для уменьшения уровня безработных на 1% необходимо уменьшить миграцию населения на 2,326%.

Факторный анализ проводился с использованием прикладной программы SPSS по методу отбора главных компонент, метод вращения выберем варимакс. Результаты представлены в таблице 3. Как видно из таблицы, выделились четыре фактора. Опишем их:

Первый фактор – экономика, в него вошли следующие показатели:

- 1) год;
- 2) рост постоянного населения;
- 3) индекс потребительских цен;
- 4) индекс промышленного производства.

Второй фактор – регистрируемый рынок труда:

- 1) сальдо миграции;

- 2) численность зарегистрированных безработных;
- 3) потребности в работниках.

Третий фактор – демографические процессы:

- 1) рост занятого населения;
- 2) численность общих безработных.

Четвертый фактор – доходы населения:

- 1) инвестиции в основной капитал;
- 2) доходы населения.

Заключение. Рассмотренный подход, применимо к прогнозированию рынка труда, позволяет комплексно отслеживать влияние основных социально-экономических факторов на ситуацию на рынке труда.

С помощью предложенного инструментария можно разработать варианты прогноза основных параметров рынка труда, занятости населения в зависимости от макроэкономического сценария развития экономики.

Библиографический список

1. Подружко А.А., Подружко А.С. Интервальное представление полиномиальных регрессий. – М., 2003.
2. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. Приближение функций, дифференциальные и интегральные уравнения. – М., 1984.
3. Шокин Ю.И. Интервальный анализ. – Новосибирск, 1981.
4. Максимов А.В., Оскорбин Н.М. Многопользовательские информационные системы: основы теории и методы исследования: монография. – Барнаул, 2005.
5. BaseGroup Labs: Технология анализа данных [Электронный ресурс]. – Режим доступа: <http://www.basegroup.ru>