

Применение многофакторной динамической модели прогноза в исследовании здоровья населения

© 2009 П.А. Смелов

© 2009 М.В. Карманов

доктор экономических наук, профессор
Московский государственный университет экономики,
статистики и информатики (МЭСИ)

Рассматривается методика построения многофакторной динамической модели прогноза. Представлен алгоритм расчета модели на примере прогнозирования средней ожидаемой продолжительности жизни населения России, как одного из интегральных показателей здоровья.

Ключевые слова: прогнозирование, многофакторные модели, динамические модели, здоровье, средняя ожидаемая продолжительность жизни.

При изучении закономерностей формирования уровня ожидаемой продолжительности жизни населения следует учитывать сложное переплетение множества факторов, выявление связей и взаимозависимостей между которыми поможет воздействовать на рост ожидаемой продолжительности жизни. В настоящее время достижения науки позволяют не только установить сам факт наличия таких связей, но и оценить их количественно с помощью многофакторных регрессионных моделей прогноза.

Многофакторные регрессионные модели прогноза ожидаемой продолжительности жизни населения могут строиться как по пространственной информации, так и по рядам динамики. Пространственная информация может отражать влияния предшествующих периодов времени и воздействовать на формирование информации, которая будет характеризовать ожидаемую продолжительность жизни населения в будущем. Но этот динамический характер пространственной информации, используемой для построения многофакторных моделей прогноза, не дает возможности выявить изменение влияний факторных признаков на продолжительность жизни во времени и не позволяет учесть запаздывание влияний этих факторов.

По нашему мнению, лишь совместное использование пространственной информации позволяет построить многофакторную динамическую модель ожидаемой продолжительности жизни населения, пригодную для практического использования.

Многомерная регрессионная модель прогноза ожидаемой продолжительности жизни населения, построенная по временным рядам, будет считаться динамической, если она позволяет учесть полное изменение состояния исследуемого процесса. По определению Г.В. Розанова¹, динамическая модель должна учитывать:

¹ Френкель А.А. Прогнозирование производительности труда: методы и модели. М., 1989. С. 214.

- общие закономерности изменения явления во времени;
- запаздывание влияний факторов-аргументов;
- закономерности изменения во времени структуры влияний факторов-аргументов.

Динамический характер многофакторной модели прогноза ожидаемой продолжительности жизни, удовлетворяющей перечисленным выше требованиям, позволит учесть временной вид информации, по которой строится модель, выявить существование динамики влияния факторов в целом, а также характер влияния различных факторов на результирующий показатель.

Существование динамики влияния факторов выражается в изменении времени коэффициентов регрессии по причине их нестабильности при разработке многофакторных регрессионных моделей по данным, взятым за определенный период, и сравнении между собой полученных результатов. При рассмотрении динамики коэффициентов регрессии следует учитывать, что чаще всего коэффициенты регрессии имеют ярко выраженную тенденцию, причем наибольшим изменениям подвержен свободный член. Попытка анализа причин, из-за которых возникает тенденция изменения во времени коэффициентов регрессии, была предпринята Л. М. Шнайманом², однако провести четкое разграничение, за счет чего появляется эта тенденция во временных рядах коэффициентов регрессии, на наш взгляд, вряд ли вообще возможно.

Основным достоинством многофакторных динамических моделей прогноза ожидаемой продолжительности жизни населения является возможность получения итогового уровня исследуемого показателя в будущем в зависимости от изменения во времени факторов, оказывающих влияние на результирующий показатель.

² Френкель А.А. Указ. соч. С. 215.

Прогнозные значения факторных признаков, влияющих на ожидаемой продолжительности жизни, можно получить, используя любые методы прогнозирования. В этом случае модель прогноза будет иметь следующий вид:

$$\hat{y}_t = a_0 + a_1x_{1t} + a_2x_{2t} + \dots + a_nx_{nt},$$

где \hat{y}_t - прогнозируемый показатель;

x_{it} - прогнозируемые значения факторных признаков;

a_{it} - прогнозируемые оценки параметров моделей регрессии;

t - период прогнозирования.

В данной модели учитываются результаты, полученные при анализе временных рядов факторов, оказывающих влияние на ожидаемую продолжительность жизни населения, коэффициенты регрессии и динамика изменения коэффициентов регрессии во времени. Отметим, что тенденции во временных рядах коэффициентов регрессии складываются в результате изменения во времени как структур влияния факторов на ожидаемую продолжительность жизни, так и значений самих факторов.

Для построения адекватных динамических моделей следует отобрать наиболее значимые факторные признаки, которые оказывают решающее влияние на прогнозируемый показатель. Задача отбора факторных признаков может быть решена на основе эвристических или многомерных статистических методов анализа.

Исходной информацией для построения многофакторной динамической модели ожидаемой продолжительности жизни населения Российской Федерации послужили данные по 80 регионам Российской Федерации за период 1990 - 2007 гг. Прогнозирование ожидаемой продолжительности жизни населения по пространственно-временной информации позволит выявить закономерность

формирования уровня исследуемого показателя в зависимости от факторов, оказывающих влияние как за счет времени, так и за счет изменения структуры связей показателей.

При построении прогноза рассматривались три фактора, оказывающих существенное влияние на уровень ожидаемой продолжительности жизни населения: коэффициент младенческой смертности, коэффициент смертности населения в трудоспособном возрасте, коэффициент заболеваемости населения, отобранные в результате проведенного регрессионного анализа.

На первом этапе прогнозирования по временным рядам рассматриваемых показателей для каждого года в период с 1990 по 2007 г. были построены матрицы парных коэффициентов корреляции в целях выявления мультиколлинеарности между анализируемыми показателями и изменений направления связи между ними. Анализ матриц парных коэффициентов корреляции между ожидаемой продолжительностью жизни населения в регионах и выбранными факторами (коэффициент младенческой смертности, коэффициент смертности населения в трудоспособном возрасте, коэффициент заболеваемости населения) подтвердил отсутствие мультиколлинеарности. Направление связи между исследуемыми показателями в наблюдаемые периоды времени также не менялось.

Далее для каждого года были построены уравнения регрессии, выражающие зависимость ожидаемой продолжительности жизни населения (\hat{y}_t) от коэффициента младенческой смертности (x_1), коэффициента смертности населения в трудоспособном возрасте (x_2) и коэффициента заболеваемости населения (x_3) по всем регионам России за период с 1990 по 2007 г. (табл. 1).

Таблица 1. Многофакторные регрессионные модели ожидаемой продолжительности жизни населения Российской Федерации в 1990 - 2007 гг.

Период	Уравнение регрессии	R ²	F-критерий
1990	$\hat{y}_1 = 74,923 - 0,327 * x_1 - 0,023 * x_2 - 0,001 * x_3$	0,899	16,18
1991	$\hat{y}_1 = 74,439 - 0,317 * x_1 - 0,042 * x_2 - 0,001 * x_3$	0,816	11,23
1992	$\hat{y}_1 = 74,354 - 0,343 * x_1 - 0,076 * x_2 - 0,001 * x_3$	0,894	10,14
1993	$\hat{y}_1 = 72,336 - 0,260 * x_1 - 0,111 * x_2 - 0,001 * x_3$	0,875	5,17
1994	$\hat{y}_1 = 76,948 - 0,438 * x_1 - 0,264 * x_2 - 0,002 * x_3$	0,885	9,72
1995	$\hat{y}_1 = 73,144 - 0,310 * x_1 - 0,326 * x_2 - 0,002 * x_3$	0,824	7,12
1996	$\hat{y}_1 = 73,123 - 0,267 * x_1 - 0,167 * x_2 - 0,001 * x_3$	0,882	9,70
1997	$\hat{y}_1 = 74,807 - 0,363 * x_1 - 0,117 * x_2 - 0,001 * x_3$	0,847	13,12
1998	$\hat{y}_1 = 73,116 - 0,235 * x_1 - 0,141 * x_2 - 0,001 * x_3$	0,835	7,59
1999	$\hat{y}_1 = 74,452 - 0,226 * x_1 - 0,330 * x_2 - 0,001 * x_3$	0,897	10,43
2000	$\hat{y}_1 = 73,288 - 0,202 * x_1 - 0,396 * x_2 - 0,001 * x_3$	0,866	8,94
2001	$\hat{y}_1 = 76,637 - 0,319 * x_1 - 0,462 * x_2 - 0,001 * x_3$	0,891	15,85
2002	$\hat{y}_1 = 77,997 - 0,334 * x_1 - 0,532 * x_2 - 0,002 * x_3$	0,831	18,68
2003	$\hat{y}_1 = 77,935 - 0,326 * x_1 - 0,554 * x_2 - 0,002 * x_3$	0,874	22,23
2004	$\hat{y}_1 = 78,594 - 0,333 * x_1 - 0,603 * x_2 - 0,001 * x_3$	0,848	20,05
2005	$\hat{y}_1 = 78,067 - 0,306 * x_1 - 0,608 * x_2 - 0,001 * x_3$	0,856	20,64
2006	$\hat{y}_1 = 74,846 - 0,139 * x_1 - 0,542 * x_2 - 0,001 * x_3$	0,844	12,91
2007	$\hat{y}_1 = 76,956 - 0,354 * x_1 - 0,616 * x_2 - 0,001 * x_3$	0,818	17,69

Таблица 2. Модели прогноза для коэффициентов регрессии средней ожидаемой продолжительности жизни населения РФ в 1990–2007 гг.

Коэффициент регрессии	Уравнение тренда
a_0	$\bar{y}_{a_0} = 74,8 - 0,292 * t + 0,0298 * t^2$
a_1	$\bar{y}_{a_1} = -0,342 + 0,0208 * \ln(t)$
a_2	$\bar{y}_{a_2} = 0,0291 - 0,0376 * t$
a_3	$\bar{y}_{a_3} = -0,00128 + 0,000287/t$

Таблица 3. Прогнозные значения коэффициентов регрессии РФ на 2008 - 2010 гг.

Период упреждения	Прогноз	Нижняя граница прогноза	Верхняя граница прогноза
Коэффициент a_0			
2008	77,993	75,645	80,34
2009	78,863	75,973	81,753
2010	79,793	76,291	83,295
Коэффициент a_1			
2008	-0,281	-0,325	-0,237
2009	-0,280	-0,325	-0,234
2010	-0,279	-0,325	-0,232
Коэффициент a_2			
2008	-0,686	-0,756	-0,615
2009	-0,723	-0,800	-0,647
2010	0,761	-0,843	-0,679
Коэффициент a_3			
2008	-0,001	-0,0015	-0,0005
2009	-0,001	-0,0015	-0,0005
2010	-0,001	-0,0015	-0,0005

Анализируя построенные статические (пространственные) модели в динамике, видно, что значения переменных и величины коэффициентов регрессии изменялись во всех уравнениях связи.

Статистическая проверка моделей показала, что все уравнения регрессии адекватны, так как расчетные значения критерия Фишера - Снедекора (F-критерия) по всем построенным моделям намного превышают табличное 2,60 для 5%-ного уровня значимости. Коэффициенты множественной детерминации (R^2) показывают, что отобранные факторы объясняют от 81,6 до 89,9% вариации результативного признака.

Значения средних ошибок аппроксимации незначительны и не превышают 13%, что свидетельствует о том, что построенные модели регрессии в целом достаточно хорошо отражают зависимость ожидаемой продолжительности жизни населения от коэффициента младенческой смертности, коэффициента смертности населения в трудоспособном возрасте и коэффициента заболеваемости населения и отражают закономерности развития продолжительности жизни за 1990-2007 гг.

Анализ динамики структуры влияния факторных признаков на уровень ожидаемой продолжительности жизни населения с учетом особенностей законов изменения каждого рассмат-

риваемого фактора во времени позволил выявить заметную тенденцию увеличения влияния общего коэффициента смертности населения на итоговый уровень прогнозируемого показателя. С учетом сказанного можно предположить, что в планируемом периоде не намечается резких изменений в структуре показателей, поэтому коэффициенты регрессии могут быть использованы для прогноза ожидаемой продолжительности жизни мужчин.

На следующем этапе построения прогнозной модели были рассчитаны трендовые модели коэффициентов регрессии, отображающих динамику влияния на среднюю ожидаемую продолжительность жизни, по временным рядам коэффициентов регрессии (табл. 2).

Построенные динамические модели коэффициентов регрессии были оценены статистическими характеристиками. Приведенные значения статистических показателей точности и адекватности полученных моделей прогноза коэффициентов регрессии наилучшим образом аппроксимируют реально существующие тенденции и закономерности изменения рассмотренных показателей.

Прогнозные значения коэффициентов регрессии динамической модели зависимости средней ожидаемой продолжительности жизни населения от коэффициента младенческой смертно-

Таблица 4. Прогнозные значения ожидаемой продолжительности жизни населения Российской Федерации в 2008–2010 гг.

Период упреждения	Прогноз	Нижняя граница	Верхняя граница
2008	68,48	66,58	71,22
2009	69,53	67,47	72,31
2010	70,65	68,91	73,04

ти, коэффициента смертности населения в трудоспособном возрасте, коэффициента заболеваемости населения представлены в табл. 3.

На основе моделей прогноза коэффициентов регрессии была построена динамическая обобщенная прогнозная модель ожидаемой продолжительности жизни населения, которая имеет следующий вид:

$$\hat{y} = 74,8 - 0,292 * t + (-0,342 + 0,0208 * \ln(t)) * x_1 + 0,0298 * t^2 (0,0291 - 0,0376 * t) * x_2 + (-0,00128 + 0,000287 / t) * x_3$$

Данная модель учитывает, с одной стороны, изменение структуры влияния факторов на уровень ожидаемой продолжительности жизни населения Российской Федерации, а с другой стороны, закономерности изменения и влияния результативного и факторных признаков за исследуемый период времени.

Подставив в обобщенную многофакторную динамическую модель средней ожидаемой продолжительности жизни населения соответствующие значения t -факторов, мы получили про-

гнозные значения результативного показателя на три года (табл. 4).

Из таблицы видно, что тенденция к увеличению ожидаемой продолжительности жизни населения России будет наблюдаться на протяжении всего периода упреждения, т.е. до 2010 г., и максимальный прогнозируемый уровень составит, соответственно, 70,65 года.

Таким образом, на основании полученной динамической многофакторной регрессионной модели ожидаемой продолжительности жизни населения Российской Федерации за период с 1990 по 2007 г. можно сделать выводы о том, что она дает достаточно точный прогноз. В целом, полученная прогнозная модель служит практически ценным инструментом в процессе аргументации и обоснования самых разнообразных программ развития всех без исключения сфер и направлений жизнедеятельности, потому что средняя ожидаемая продолжительность жизни населения является барометром здоровья населения, отношение к ней предопределяет зрелость общества и институтов государственной власти.

Поступила в редакцию 09.01.2009 г.