

С.В. Вараксин, Н.В. Вараксина, Н.П. Гончарова (Барнаул)

ПРИМЕНЕНИЕ МЕТОДА НЕЧЕТКОЙ ЛИНЕЙНОЙ РЕГРЕССИИ К МОДЕЛИРОВАНИЮ ДЕМОГРАФИЧЕСКИХ ПРОЦЕССОВ¹

Научные исследования, привлекающие технику нечетких множеств и нечеткой логики, ведут начало от основополагающей работы Л.Заде [1]. Отличие нечеткого множества от обычного в том, что характеристическая функция принадлежности $\mu(x)$ нечеткого множества может принимать произвольные значения от 0 до 1, соответственно степени достоверности принадлежности элемента этому множеству. Нечеткое число - это нечеткое множество с выпуклой унимодальной (т.е. имеющей один максимум) функцией принадлежности, значение в точке максимума которой равно 1. Одними из наиболее употребительных нечетких чисел являются треугольные числа (a,b,c) , ненулевые части графика функции принадлежности которых образованы двумя наклонными отрезками. Нахождение значений функций принадлежности нечетких множеств на основе обычных, четких данных называют фазификацией, а нахождение обычного, нечеткого значения для нечеткого множества – дефазификацией. Нечетким временным рядом называют набор соответствующих определенным моментам времени нечетких чисел.

1. Построение модели нечеткой линейной регрессии демографического поведения.

В стандартной вероятностной регрессии возникают проблемы при исследовании коротких временных рядов, с неясными законом распределения и зависимостью между входной и выходной переменными. При построении модели нечеткой регрессии не требуются подобные предположения. Первой в разработке моделей нечеткой линейной регрессии стала работа Х. Танака [2], базовая модель нечеткой линейной регрессии в которой имеет вид $x=kt+b$, где k и b – некоторые треугольные числа. В

* Статья выполнена в рамках гранта РФФИ №№16-06-00350 «Прогностический потенциал теории нечетких временных рядов в построении модели демографического поведения населения» 2016-2018 г.

данной работе использован частный случай модели Танака, равномерной нечеткой регрессии, в которой k является четким числом, а b – нечетким числом вида $(b-\Delta, b, b+\Delta)$. Исходный временной ряд $\{x(t)\}$ является обычным числовым рядом, а его значения считаются дефазификациями соответствующих нечетких чисел A_i . Согласно работе Пономарева И.В., Славского В.В [3], параметры подобной нечеткой линейной регрессии находятся из условия наибольшего правдоподобия как параметры наиболее узкой полосы между двумя параллельными линиями, содержащей все точки временного ряда $\{x(t)\}$. Эта задача сводится к задаче линейного программирования, которую можно решить симплекс-методом в два этапа с введением искусственных переменных, т.к. исходная задача не имеет подходящего исходного опорного решения.

2. Моделирование динамики численности населения России.

Численность населения России с 1991 по 2016 годы задана последовательностью (в тыс. жителей) 148543, 148704, 148673, 148366, 148306, 147976, 147502, 147105, 146693, 145925, 146304, 145649, 144964, 144168, 143474, 142754, 142220, 141980, 141900, 142962, 142914, 143103, 143395, 143700, 146267, 146546. Применение симплекс-метода применительно ко всему периоду дает коэффициенты $k=-90$, $b=330330$, $\Delta=2650$. В ряду данных четко видна точка минимума в 2009 году, поэтому применим данный алгоритм к периодам 1991-2009 и 2009-2016 г., получим соответственно $k=-400$, $b=936220$, $\Delta=860$ и $k=70$, $b= -118660$, $\Delta=1002$.

3. Математическая модель коэффициента рождаемости.

Исходные данные с 1991 по 2015 годы: 12,1, 10,7, 9,4, 9,6, 9,3, 8,9, 8,6, 8,8, 8,3, 8,7, 9, 9,7, 10,2, 10,4, 10,2, 10,4, 11,3, 12,1, 12,4, 12,5, 12,6, 13,3, 13,2, 13,3, 13,3. Применение симплекс-метода применительно ко всему периоду периоду дает коэффициенты $k= 0.057$, $b= -104$, $\Delta=2.13$. Также найдем линии регрессии для интервалов 1991-1999 и 1999-2015 гг.: $k = -0.47$, $b = 950$, $\Delta = 0.88$ и $k=0.32$, $b = -635$, $\Delta = 0.53$ соответственно.

4. Математическая модель коэффициента смертности в России.

Найдем параметры линейной регрессии коэффициента смертности в России с 1991 по 2015 годы. Исходные данные: 11,4, 12,2, 14,5, 15,7, 15, 14,2, 13,8, 13,6, 14,7, 15,4, 15,6, 16,2, 16,4, 16, 16,1, 15,2, 14,6, 14,6, 14,2, 14,2, 13,5, 13,3, 13,1, 13,1, 13,1 на тысячу жителей. Применение симплекс-метода применительно ко всему периоду дает коэффициенты $k = 0.071$, $b = -128$, $\Delta = 2.1$, а отдельно с 1991 по 2003 и с 2003 по 2015 гг.: $k = 0.31$, $b = -613$, $\Delta = 1.7$ и $k = -0.30$, $b = -617$, $\Delta = 0.45$ соответственно.

Подобные исследования показывают, что метод нечеткой линейной регрессии вполне применим для моделирования демографических процессов. Сочетание математического моделирования и социологического анализа позволит разрабатывать более точные и многоаспектные прогнозные модели демографического поведения населения.

Библиографический список

1. Zadeh L. A., Fuzzy sets // Information and Computation, vol. 8, pp. 338–353, 1965.
2. Tanaka H., Uejima S., Asai K., Linear Regression Analysis with Fuzzy Model // IEEE Transactions on Systems, Man and Cybernetics. 1982. Vol. 12. No. 6. P. 903–907
3. Пономарев И.В., Славский В.В. Нечеткая модель линейной регрессии // Доклады Академии наук. - 2009. - Т. 428, №5. - С. 598-600.