

сравнение построенных оценок методом кросс-валидации и реальных значений ряда с помощью выбранных метрик.

Исследование параметра «радиус влияния» модели семивариограммы показало, что с увеличением его значения растет точность прогноза.

К. Г. Дюндя

(ГрГУ им. Я. Купалы, Гродно)

ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ КЛАССИФИКАЦИИ МНОГОМЕРНЫХ НАБЛЮДЕНИЙ

Кластерный анализ – процедура, предназначенная для группировки наблюдений или переменных в значимые кластеры на основе сходства между ними. Кластер – это группа объектов, внутригрупповое сходство между которыми максимально, а межгрупповое сходство сведено к минимуму по какому-то объективному критерию [1]. Большинство методов кластерного анализа производит объединение наблюдений в кластеры на основе матрицы расстояний.

В ходе проделанной работы были проанализированы данные по семи странам, отражающие значения таких признаков, как число врачей на 10 тыс. населения, смертность на 100 тыс. населения, ВВП, расходы на медицину. При выполнении работы использовались методы кластерного анализа: иерархический агломеративный метод и метод k -средних, реализованные в пакете STATISTICA. В результате проведенного исследования оптимальной была признана классификация, выделяющая два кластера, представленная в таблице 1.

Таблица 1 – Результаты классификации

Кластер 1	Кластер 2
Россия	Австралия
Беларусь	Австрия
Азербайджан	Бельгия
Армения	

Результаты дисперсионного анализа подтверждают, что классификация является обоснованной по всем признакам, за исключением одного. Наблюдается несущественная разница между кластерами лишь по числу врачей на 10 тыс. населения. Страны первого кластера

имеют существенно большую смертность, но существенно меньшие расходы на медицину и ВВП.

Литература

1 Берестнева, О. Г. Прикладная математическая статистика / О. Г. Берестнева, О. В. Марухина, Г. Е. Шевелёв. – Томск : ТПУ, 2012. – 188 с.

М. Ф. Жихарко

(ГрГУ им. Я. Купалы, Гродно)

ПРИМЕНЕНИЕ СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ В СТРАХОВАНИИ

Потоки исков, поступающих в страховую компанию (СК) в случайные моменты времени, случайные длительности интервалов времени, необходимых для их обработки, предопределили необходимость использования методов теории сетей массового обслуживания (СеМО) для разработки математических моделей процессов обработки исков.

Рассмотрим СК, имеющую K полисодержателей и страхующую n типов рисков: авто, имущество, здоровье, ответственность и другие. Каждый застрахованный может находиться в одном из следующих состояний: S_0 – нет необходимости предъявлять иск, S_{n+1} – иск требует оплаты одним из m_{n+1} кассиров, S_i – иск требует обслуживания одним из m_i оценщиков исков i -го типа, $i = \overline{1, n}$. В качестве математической модели используем замкнутую марковскую СеМО с узлами S_i типа M/m_i , $i = \overline{1, n+1}$, нулевой узел – источник заявок. Каждое обращение клиента с иском в СК соответствует поступлению заявки в СеМО. Пребывание заявки в определенной СМО и ее маршрутизация между СМО соответствуют статусу иска клиента в СК и процессу его маршрутизации между оценщиками и кассирами.

Пусть $k_i(t)$ обозначает количество заявок в состоянии S_i в момент времени t , $i = \overline{1, n+1}$. Очевидно, что $\sum_{i=0}^{n+1} k_i(t) = K$. Состояние исследуемой сетевой модели в момент t определяется случайным процессом $k(t) = (k_1(t), k_2(t), \dots, k_{n+1}(t))$.