УДК 519.257

Валитов Д.И.
ФГБОУ ВО Башкирский государственный университет, Уфа,
e-mail: damirvalitov@yandex.ru

РАЗРАБОТКА МЕХАНИЗМА РАСЧЕТА СТРАХОВЫХ СТАВОК В МОДЕЛИ ЗАБОЛЕВАНИЯ САХАРНЫМ ДИАБЕТОМ

Ключевые слова: статистические данные, страхование, актуарные расчеты, марковская модель, интенсивности перехода, система уравнений Колмогорова-Чепмэна, сахарный диабет, граф состояний.

В основе статьи лежат актуарные расчеты на базе статистических данных. Актуарные расчёты – это расчёты страховых ставок на базе математической статистики. За статистическую базу взяты данные заболеваемости сахарным диабетом. На основе реальных данных, предоставленных в Медицинском Информационно-Аналитическом Центре при Министерстве Здравоохранения Республики Башкортостан, получены необходимые статистические совокупности, построена математическая модель, поставлены 2 задачи (обратная и прямая), их решение получено с использованием вероятностно-статистического анализа, марковских процессов, теории графов, интервального анализа, принципа двойственности в линейном программировании. При решении поставленной задачи использовались научные исследования ученых, занимающихся исследованием математической статистики, актуарной математикой, математического моделирования при страховании. Написана программа для решения обратных задач определения интенсивностей переходов в системе многих состояний процесса. Используя результаты исследования и на основе вероятностно-статистического моделирования полученных результатов можно предсказывать, управлять и оценивать риски, специалистам по медицинскому страхованию -рассчитывать тарифы. Это дает определенный импульс в анализе рисков при расчете границ величин страховых взносов в задачах профилактики и лечения сахарным диабетом.

Введение

В последнее время начали широко применяться актуарные исследования в анализе рисков. Существует большое количество моделей, и одной из них является стохастическая модель Уилки. Такие модели описывают случайные процессы, в которых подразумевается, что случайность тех или иных явлений выражается в терминах вероятности. Немаловажную роль в их построении также играет и статистическое моделирование. В основе регулирования страховых отношений между страховщиком и страхователем лежат методы, опирающиеся на математические и статистические закономерности. Они и называются актуарными расчетами.

Большинство популярных приложений используют принцип, использующий марковские процессы при моделировании условий поведения системы. Это задачи с начальными статистическими данными, например, задачи системам массового обслуживания: почта, мастерские, заправочные станции, магазины и т.д. Так для представления состояния страхователя используется тот же подход.

В результате математического анализа получаем 2 противоположные задачи: прямая, где необходимо найти вероятности нахождения индивида в том или ином состоянии при известных параметрах модели и обратная, где необходимо найти параметры модели, используя известные характеристики из эксперимента.

Обратные задачи для марковских процессов начали развиваться сравнительно недавно в отличии от методов решения прямых задач. Так, рассматривая модель медицинского страхования, определим, что интенсивности переходов в нашей случае – это количество заболевающих, которым требуется медицинская помощь, количество выздоравливающих и количество умерших в конкретную единицу времени. Решение обратной задачи это поиск интервала в заданном пространстве. Интервал характеризуется уровнем варьирования величины интенсивности перехода, оставляющей неизменным некоторые характеристики качества процесса [1].

Цель работы: определение метода решения обратных задач процесса заболеваемости сахарным диабетом на основе статистической базы в математических модели случайного процесса.

Задачи исследования:

– статистическое исследование базы данных;

– получение математической модели медицинского страхования;

– написание программы на языке программирования C++ для решения обратных задач

– приближенное решение для математической модели

– на основе вероятностно-статистического моделирования полученных результатов предсказание и оценка рисков, расчет тарифов при страховании;

Материал и методы исследования

Случайный процесс, протекающий в какой-либо системе, именуется марковским, если он удовлетворяет следующему свойству: в любое время на вероятность наступления любого состояния модели в будущем влияет только ее состояние в настоящем. С учетом непрерывности времени, протекающего в географических системах. В нашем случае состояния дискретны и время непрерывно. Марковское моделирование дает возможность исследователям видеть изменения заболевания в определенный период времени. В марковской схеме начальные параметры – это интенсивности перехода из одного состояния в другое состояние, которые находятся на основе статистической информации. Исходными данными являются состояния людей сначала в одном временном интервале, а затем в другом временном интервале (например, первый год, а затем второй год). Человек, обладающий некоторыми свойствами, является основным элементом анализируемой системы. В простейшем случае он может находиться лишь в двух состояниях: 1 – здоров, 2 – умер (рис. 1).

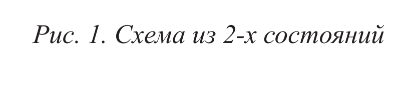

Рис. 1. Схема из 2-х состояний

В данном случае индивид может перейти только из состояния «жив» в состояние «мертв». При разных схемах страхования страховые пособия выплачиваются по-разному. Например, аннуитетные выплаты идут пока клиент находится в состоянии 1, а после перехода в состояние 2 – прекращаются. Либо производятся при переходе в состояние 2 и т.д. Все эти схемы просты и легко вычисляются.

Непростая ситуация складывается когда состояний системы больше 2-х. На рисунке 2 представлена схема с тремя возможными вариантами: «здоров», «болен», «мертв». Другими словами – описана ситуация, когда индивид застрахован по полису наступления потери трудоспособности. В таком случае взносы будут выплачиваться при нахождении страхователя в первом состоянии, а выплаты будут осуществляться при нахождении во втором состоянии.

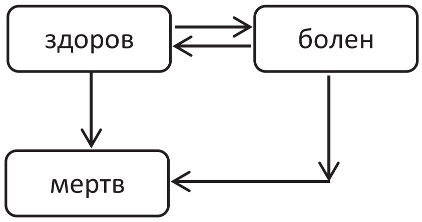

Рис. 2. Схема из 3-х состояний

В данной схеме актуарные расчеты усложняются за счет того, что застрахованный индивид может переходить из состояния в состояние по несколько раз в обоих направлениях. По данной причине в расчете как правило закладывают невозможным переход из «болен» в «здоров» [2].

В общем случае марковкий процесс имеет следующий вид:

Обозначим через X(t) состояние индивида в возрасте t (t≥0). Определим стохастический процесс, как {X(t), t≥0}. Предположим, что имеется конечное число состояний, пронумерованных 1,2,...n, т.е. процесс имеет пространство состояний {1,2,…,n}. Тогда {X(t), t≥0} – Марковский процесс, если для любых s, t≥0 и i, j, x(u)∈{1,2,…n},

$$\Pr\{X(s+t)=j\,|\,X(s)=i, X(u)=x(u), 0\le u<s\}=$$

$$=\Pr\{X(s+t)=j\,|\,X(s)=i\}$$

Описанный процесс зависит только от «настоящего» (в момент времени s) и не зависит от «прошлого» (до момента времени s).

Функция вероятности перехода будет иметь вид:

$$p_{ij}(s, s+t) \equiv \Pr\{X(s+t) = j \mid X(s) = i\},$$

$$i, j \in \{1, 2, \ldots n\},$$

и положим, что $\sum_{j=1}^{n} p_{ij}(s, s+t) = 1$ для любого $t \geq 0$.

Предположим также существование пределов

$$\mu_{ij}(t) = \lim_{h \to 0} \frac{p_{ij}(t, t+h)}{h},$$

$i, j \in \{1, 2 \ldots n\}$, $i \neq j$.

При $i \neq j$ μ_{ij} – это интенсивность перехода их состояния i в состояние j. Легко видеть, что при $s, t, u \geq 0$

Уравнение Колмогорова-Чепмэна:

$$p_{ij}(s, s+t+u) =$$

$$= \sum_{l=1}^{n} p_{il}(s, s+t) p_{lj}(s+t, s+t+u), \quad (1)$$

$i, j \in \{1, 2, \ldots n\}$.

Для получения вероятностей перехода решаем систему дифференциальных уравнений. А полученные вероятности переходов связаны с интенсивностями переходов в прямой и обратной задачах [3].

$$\frac{\partial}{\partial t} p_{ij}(s, s+t) = \sum_{l=1}^{n} p_{il}(s, s+t) \mu_{lj}(s+t) \quad (2)$$

$$\frac{\partial}{\partial t} p_{ij}(s, s+t) = -\sum_{l=1}^{n} \mu_{li}(s) p_{lj}(s, s+t) \quad (3)$$

Граничные условия для системы: $p_{ij}(s,s) = \delta_{ij}$, где – $\delta_{ij} = \begin{cases} 1, & \text{если } i = j, \\ 0 & \text{иначе.} \end{cases}$

Если в марковском процессе при $\mu_{ij}(t) = \mu_{ij}$ для любого t выводятся точные выражения для функции вероятностей переходов, то этот марковский процесс называется однородным относительно времени. При постоянных интенсивностях перехода время в каждом состоянии имеет экспоненциальное распределение и функция $p_{ij}(s,s+t)$ одинакова для всех s, и ее можно обозначить просто p_{ij}.

Модель при известных интенсивностях с любым числом возможных состояний системы (в случае прямой задачи) ведет к решению системы дифференциальных уравнений Колмогорова. А модель с неизвестными интенсивностями (в случае обратной задачи) ведет к оцениванию интенсивностей по имеющимся статистическим исходным данным.

При моделировании схемы марковского процесса необходимо ввести числовую оценку показателя качества, находящуюся внутри определенного интервала. Другими словами – интенсивности переходов между состояниями принадлежат определенному промежутку, у которого границы заданы условиями самого процесса.

Таким образом проверка на попадание величины интенсивности в сам интервал есть решение обратной задачи. Возникает задача регулирования качества, решаемая путем анализа экспериментальных данных. Критерием контроля качества будет являться возможность описания входных данных в рамках модели.

$p_{ij}(t)$ известны нам из статистической базы. Зададим условие, при котором модель будет описывать измерения в пределах их точности:

$$\left| p_i^{cm} - p_i^{p} \right| \leq \varepsilon_i, \quad (4)$$

где p_i^{cm} – табличные данные по вероятности; p_i^{p} – вероятности, рассчитанные по системе (1).

Выделим для всех из констант m_{ij} интервал неопределенности как некоторый диапазон вариации Dij, внутри которого сохраняется совместность системы (4).

$$D_{ij} = [\, \min m_{ij}, \max m_{ij} \,] \quad (5)$$

Эту задачу поиска интервалов (5) при удовлетворении неравенству (4) ставил еще в 1962г. основоположник линейного программирования Л.В. Канторович [4].

Результаты исследования

Исследования проводились при анализе процесса, в котором модель многих состояний, описывала состояния застрахованного лица с заболеванием сахарного диабета. Статистические данные по заболеванию сахарным диабетом получены в Медицинском Информационно-Аналитическом Центре при Министерстве Здравоохранения Республики Башкортостан.

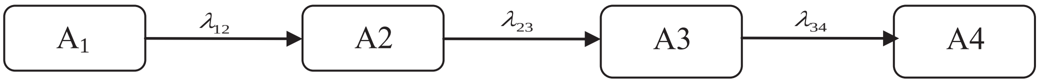

Рис. 3. Схема состояний

Рассматриваемая модель состоит из четырех состояний: «здоров», «болен сахарным диабетом инсулинонезависимый тип», «болен сахарным диабетом инсулинозависимый тип», «мертв», где λ_{ij} – интенсивности переходов между состояниями.

По представленной на рисунке 3 схеме составим систему дифференциальных уравнений Колмогорова для поиска вероятности нахождения индивида в каждом из состояний:

$$\begin{cases} \dfrac{dp_1(t)}{dt} = -\lambda_{12} p_1(t) \\ \dfrac{dp_2(t)}{dt} = \lambda_{12} p_1(t) - \lambda_{23} p_2(t) \\ \dfrac{dp_3(t)}{dt} = \lambda_{23} p_2(t) - \lambda_{34} p_3(t) \\ \dfrac{dp_4(t)}{dt} = \lambda_{34} p_3(t) \end{cases} \quad (6)$$

$p_i(t)\ (i = \overline{1,4})$ – вероятность состояния A_i.

Изначально предполагаем, что индивид находится в состоянии «здоров», т.е. задаем следующие начальные условия:

$$p_1(0) = 1,\ p_2(0) = 0,\ p_3(0) = 0,\ p_4(0) = 0, \quad (7)$$

В любой момент времени выполняется нормировочное условие, т.е. индивид находится в каком-то из 4-х состояний:

$$p_1(t) + p_2(t) + p_3(t) + p_4(t) = 1 \quad (8)$$

Решаем прямую и обратную задачи. Соответственно ищем решение уравнений Колмогорова при известных интенсивностях переходов , а при неизвестных интенсивностях переходов – по статистическим данным оцениваем интенсивности переходов.

Имея средние значения интенсивностей, которые были найдены на базе статистических данных, решена обратная задача методом Рунге-Кутта 4 порядка, реализованного в программе, написанной на языке программирования C++. Вычисления были реализованы на суперкомпьютере математического факультета БашГУ.

Для вероятности p_1 из описанной схемы выражаем интенсивность перехода

$$\lambda_{12} = -\frac{\ln\left(\dfrac{p_1}{0.998458}\right)}{t} .$$

Другие интенсивности выражаем таким же образом. Анализируя средние значения показателей интенсивностей выводим расчетные значения вероятностей нахождения в состояниях A_1 – «здоров», A_2 – «болен сахарным диабетом инсулинонезависимый тип», A_3 – «болен сахарным диабетом инсулинозависимый тип», A_4 – «мертв».

На основании сопоставимости расчетных и экспериментальных данных, показанных на графиках (рис. 4), можно заключить, что удовлетворяющая системе (6) модель соответствует реальным данным и может быть применена в практических расчетах по определению, проверке, прогнозу диапазона необходимых денежных средств, выделяемых для борьбы с сахарным диабетом [5].

Решая систему неравенств (4), определим области неопределенности по интенсивностям λ_{ij}, взяв p_i^{cm} – из табличных данных по вероятностям заболевания сахарным диабетом, а p_i^{p} – как вероятности, полученных из системы (6). Полученный интервал изменения интенсивностей необходим для интервала изменения средств.[6]

В настоящее время лечение сахарного диабета в подавляющем большинстве случаев является симптоматическим и направлено на устранение имеющихся симптомов без устранения причины заболевания, так как эффективного лечения диабета ещё не разработано. От сахарного диабета как такого никто еще не умирал, больные умирают от осложнений, которые вызывает сахарный диабет. Стоимость лечения сахарного диабета может варьироваться в достаточно большом интервале в зависимости от осложнений.

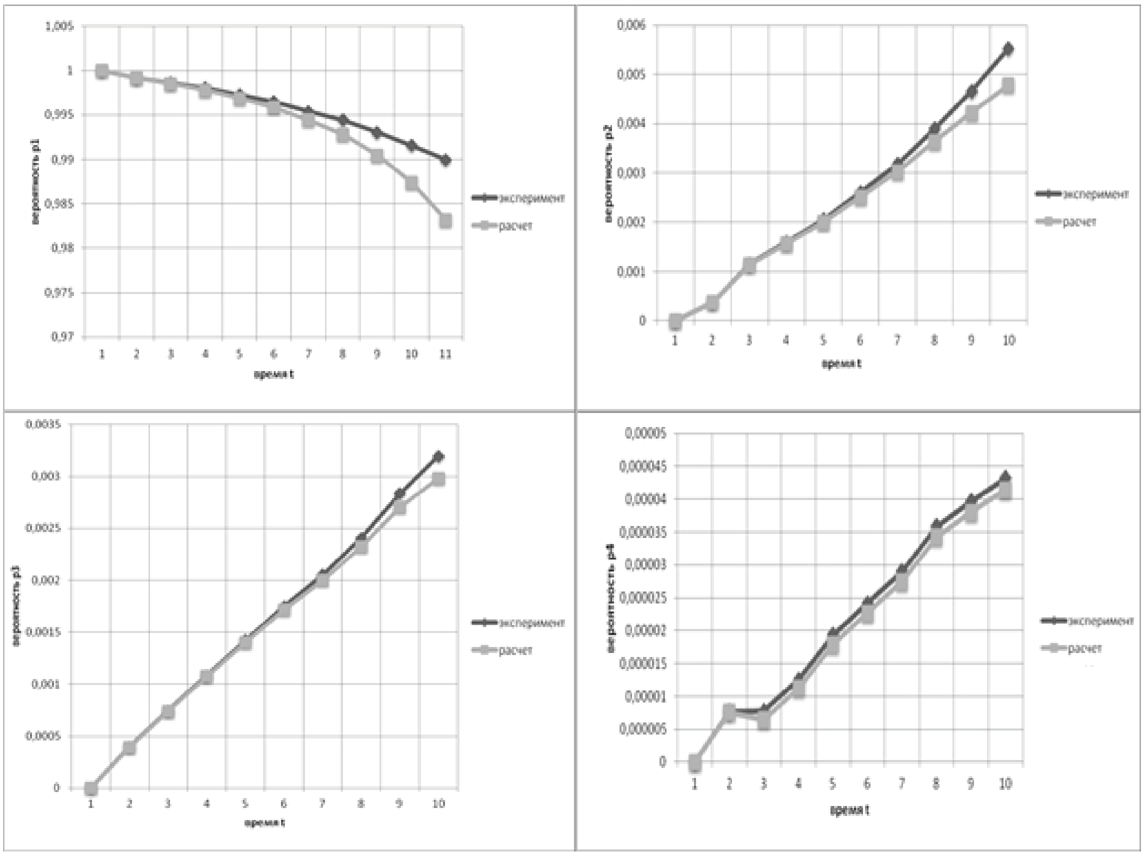

Рис. 4. Графики сравнения расчета и эксперимента

На примере 2013 года, Распространенность сахарного диабета в Российской Федерации по данным ГосРегистра 3 779 423 человек. По данным IDF на лечение 1 больного тратится 22750 рублей в год. Интервал изменения этой величины [22 000 руб.; 156 000 руб.]

На основе математических ожиданий численности количества больных определены величины расходов на медицинскую помощь больным сахарным диабетом [8,3*109; 5,89*1011]. Из интервала изменения интенсивности перехода был построен интервал изменения средств, необходимых для лечения больных сахарным диабетом [5,3*1010; 8,2*1010]. Найденный интервал входит в интервал изменения средств, выделяемых для борьбы с сахарным диабетом. Полученный интервал требуемых средств на медицинское страхование сопоставим с реальным. Таким образом данный метод определения интервала изменения требуемых средств на медицинское страхование позволяет эффективно решать задачи, возникающие в страховании.

Заключение

Разработанная математическая модель и комплекс компьютерных программ использованы при анализе реальных статистических данных по заболеванию сахарным диабетом, полученных в Медицинском Информационно-Аналитическом Центре при Министерстве Здравоохранения Республики Башкортостан. На основе вероятностно-статистических моделирования полученные результаты можно использовать для предсказания и управления. Проделанный анализ дает страховщикам в области медицины реальный механизм расчета страховых ставок.

Библиографический список

1. Валитов Д.И., Спивак С.И., Райманова Г.К. Актуарные расчеты при медицинском страховании процесса заболевания сахарным диабетом // Управление риском. – 2016. – №2. – С. 26–34.
2. Фалин Г.И., Фалин А.И. Введение в математику финансов и инвестиций для актуариев. – М: МАКС Пресс, 2017. – 327 с.
3. Спивак С.И., Райманова Г.К., Абдюшева С.Р. Обратные задачи для марковских моделей медицинского страхования // Страховое дело. – 2008. – №9. – С. 36–42.
4. Канторович Л.В. О некоторых новых подходах к вычислительным методам // Сибирский математический журнал. – 1962. – Т. 3. – № 5. – С. 701–709.
5. Спивак С.И., Райманова Г.К., Валитов Д.И. Актуарные расчеты при медицинском страховании процесса заболевания сахарным диабетом // Известия Уфимского научного центра РАН. – 2017. – № 1. – С. 17–22.
6. Спивак С.И., Райманова Г.К. Математическая модель процесса заболевания туберкулезом // Системы управления и информационные технологии. – 2009. – №36. – С. 293–297.