

Оценка и управление состоянием здоровья на основе моделей Г. Раша

Аннотация

Рассматриваются вопросы оптимизации пространства информативных признаков и синтеза правил прогнозирования, диагностики и управления состоянием здоровья человека с использованием теории измерения латентных переменных на основе статистической модели Г. Раша.

Введение

Значительное число задач оценки и управления состоянием здоровья организма человека (прогнозирование, ранняя диагностика, выбор схем профилактики и лечения при дефиците информации и т. д.) относится к классу плохо формализуемых задач, поскольку в условиях существующих медико-технических ограничений на количество, качество и время сбора необходимой информации исследуемые классы состояний сильно пересекаются и не имеют четких границ [1]-[4].

Такой тип задач удобно рассматривать с позиций теории измерения латентных переменных (IRT) в том смысле, что результаты принимаемых решений (прогноз, диагноз, эффективность лечения и другие медицинские выводы) не поддаются прямым измерениям, в то время как исходные данные для их определения можно считать индикаторными (измеряемыми) переменными.

В теории IRT устанавливается связь между двумя множествами значений латентных переменных. Первое множество – значения латентных переменных, характеризующих уровень качества объектов наблюдения θ_i , где i – номер объекта и $i = 1, 2, \dots, n$. Второе множество – значения латентных переменных, определяющих значимость j -го индикатора $\beta_j, j = 1, 2, \dots, m$ [5].

В предлагаемой статье исследуется возможность применения теории измерения латентных переменных для решения задач прогнозирования и ранней диагностики заболеваний, что представляет уверенность в принятии решений как латентную переменную, определяемую по ряду получаемых в процессе опросов, осмотров, лабораторных и инструментальных исследований информативных признаков (индикаторные переменные).

Возможно применение этой теории и для оценки эффективности профилактических и лечебных мероприятий с использованием исследуемой эффективности как латентной переменной, а составляющих лечебно-профилактических мероприятий – как индикаторных переменных.

Метод

В теории измерения латентных переменных, основанной на модели Г. Раша, обоснованной математической зависимостью, является логистическая модель, которая для медицинских приложений имеет вид [5]

$$P_{ij} = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}}, \quad (1)$$

где P_{ij} – вероятность достижения медицинской цели (прогноз, диагноз, эффективное лечение) у пациента с номером i при использовании индикаторных переменных S_{ij} (информативные признаки, составляющие схем лечения и т. д.); θ_i – эффективность достигаемой цели у пациента i ; β_j – информативность (эффективность) использования j -го признака (воздействия) при решении поставленной задачи.

Все латентные переменные измеряются в логитах, которые формируются при переходе от количественных шкал исходных информативных признаков X_i к интегральным шкалам K_i , вычисляемым по формуле

$$K_i = \begin{cases} \text{int} \left(\frac{X_i - X_{i,\min}}{R_i} \right) & \text{при } X_i < X_{i,\max}; \\ n-1 & \text{при } X_i = X_{i,\max}, \end{cases} \quad (2)$$

где X_i – текущее натуральное значение i -го индикатора; $X_{i,\min}$ – наименьшее значение i -го индикатора; $X_{i,\max}$ – наибольшее значение i -го индикатора; n – основание качественной шкалы ($n = 2, 3, \dots$), выбираемое в соответствии с рекомендациями [5]; $\text{int}(\dots)$ – целая часть числа.

Исследование роли индикаторных переменных в измерении латентных переменных удобно проводить с помощью стандартного пакета диалоговых прикладных программ RUMM 2020 (Rasch Unidimensional Measurement Models) [5]. Используя значения индикаторных переменных, переведенных в логиты, пакет RUMM 2020 строит теоретические кривые модели Г. Раша, по которым судят о соответствии индикаторных переменных этой модели и формируют пространство информативных признаков и (или) рациональных схем лечебно-оздоровительных мероприятий.

Рассмотрим более подробно механизм формирования пространства информативных признаков с использованием диалогового пакета RUMM 2020. В ходе реализации этого пакета формируется семейство графиков теоретических кривых Г. Раша (характеристических кривых), ось абсцисс которого является шкала логитов индикаторных переменных, а осью ординат – шкала логитов латентной переменной.

Относительно теоретической кривой по обучающей выборке формируются три примерно равные группы с низким, средним и высоким уровнем исследуемых состояний человека или эффективности лечебно-профилактических мероприятий, для которых определяются координаты их средних значений. Считается, что если индикатор «хорошо» соответствует общему набору индикаторов, то точки, определяющие «слабые», «средние» и «сильные» уровни, располагаются близко.

В ходе реализации пакета RUMM 2020 рассчитываются: степень соответствия индикаторных переменных модели измерения латентной переменной $\chi^2_{\text{критич. Prob}}$ – χ^2 квадрат – $\chi^2_{\text{критич. Prob}}$; местоположения индикаторной переменной, измеряемой в логитах, – $Location$; погрешность измерения местоположения индикаторной переменной, измеряемой в логитах, SE ; величина, характеризующая суммарное отклонение значений данного индикатора от ожидаемых значений на основе модели – $FitResid$.

Если для каждой индикаторной переменной $\chi^2_{\text{критич. Prob}} \geq 0,05$, то принимается решение о том, что все выбранные исходные индикаторы совместимы между собой и определяют одну и ту же латентную переменную.

Индикаторные переменные, для которых $\chi^2_{\text{критич. Prob}} < 0,05$, корректируются или исключаются из списка информативных признаков.

Для откорректированного набора индикаторов проводится новый анализ степени соответствия индикаторных переменных модели измерения латентной переменной по величине $\chi^2_{\text{критич. Prob}}$.

Перечисленные процедуры повторяются до тех пор, пока для всех индикаторных переменных не будет выполняться условие $\chi^2_{\text{критич. Prob}} > 0,05$.

На заключительном этапе выбора списка информативных признаков (схем лечебно-оздоровительных мероприятий) реализуется контрольная проверка с использованием графиков соответствия между уровнем интегрального показателя латентной переменной и значимостью индикаторных переменных (рис. 1).

На рис. 1 ось абсцисс является шкала исследуемой латентной переменной L . По оси ординат верхней части рисунка откладывается число объектов (людей из обучающей выборки), попадающих в полученные интервалы латентной переменной. То есть фактически рис. 1 представляет собой гистограмму распределения объектов наблюдения по шкале L . По оси ординат нижней части рисунка откладывается число информативных переменных, приходящихся на выделенные интервалы латентной переменной.

Средняя значимость информативных признаков (индикаторов) не должна отличаться от среднего уровня интегрального показателя более чем на 0,5 логит [4]. Если это условие выполняется, то можно сделать вывод о том, что система информативных признаков (индикаторов) соответствует измеряемой латентной переменной и набор используемых информативных признаков (индикаторов) является эффективным для целей измерения обобщенной латентной переменной.

Используя теорию измерения латентных переменных, основанную на модели Г. Раша, можно не только получить число, обозначающее информативность того или иного признака, но и дать развернутое описание ряда его качественных свойств: способности дифференцировать шкалы принимаемых решений (уверенность в прогнозе, диагнозе); способности признака дифференцировать исследуемую латентную переменную в противоположность ожидаемой ситуации; способности определять признаки со сверхвысокой и сверхнизкой дискриминирующей способностью, признаки с «неупорядоченной» классификацией и т. д. [5].

Такая гибкость анализа системы информативных признаков на основе модели Г. Раша позволяет рекомендовать ее как один из основных механизмов проведения разведочного ана-

лиза при решении общих задач синтеза моделей вычисления латентной переменной для прогностических и диагностических решающих правил.

В пакете RUMM 2020 предусмотрена возможность определения функциональной связи между латентной переменной L , определяемой в логитах, и выбранным набором индикаторных переменных S_j :

$$L = f_L(S_j), \quad (3)$$

где f_L – вид функциональной зависимости L от S_j .

На рис. 1 латентная переменная L представлена осью абсцисс.

Часто при синтезе решающих правил, в условиях нечеткости и неопределенности, используется аппарат нечеткой логики принятия решений с определением функций принадлежности к исследуемым классам состояний и (или) уверенности в принимаемых решениях [1]-[4], [6]-[8].

Для согласованного применения нечеткой логики принятия решений целесообразно шкалу L в логитах использовать как базовую переменную для построения функций принадлежности $\mu_{\omega_j}(L)$ объекта исследований к классам состояний ω_j .

Для перехода от L к $\mu_{\omega_j}(L)$ удобно использовать такую обобщенную характеристику, определяемую пакетом RUMM 2020, как гистограмма распределения объектов исследования на шкале латентной переменной L (рис. 1). При этом функции $\mu_{\omega_j}(L)$ удобно строить, используя рекомендации работ [3], [9].

Таким образом, модель Г. Раша может быть использована для нечеткой оценки состояния объекта исследования с расчетом уверенности в классификации в соответствии с выражением

$$UP_i = \mu_{\omega_j}(L). \quad (4)$$

В качестве примера в табл. 1 приведен фрагмент экспериментальных данных, полученных для решения задачи прогнозирования появления язвы желудка по факторам риска (индикаторным переменным – ИП), приведенным в работах [4], [9].

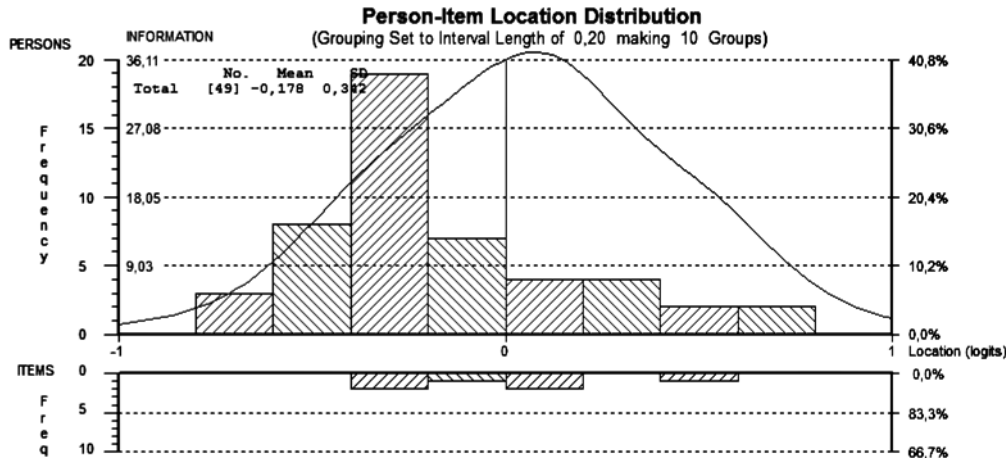


Рис. 1. Соответствие между уровнем интегрального показателя латентной переменной (вверху) и местоположением информативных признаков – индикаторов (внизу)

Таблица 1

Фрагмент таблицы экспериментальных данных, подготовленной для обработки пакетом RUMM 2020 по задаче оценки уверенности в возникновении у обследуемого язвы желудка

Номер обследуемого	ИП											
	GI	AI	Tk	Br	E21	E36	V21	V43	VB24	YP	YU	
1	0	0	0	0	5	10	5	15	10	0,3	0,1	
2	5	10	0	1	0	3	2	5	5	0,1	0,2	
...	
99	25	30	25	2	80	70	95	100	110	0,9	0,3	
100	30	28	32	3	95	85	100	95	80	0,8	0,4	

В этой *таблице* принято такое же обозначение индикаторных переменных, что и в работах [4], [9]: *Gl* – прием лекарственных средств; *Al* – прием алкоголя; *Tk* – табакокурение; *Br* – болезни желудка у родственников; отклонение сопротивления биологически активных точек, связанных с заболеванием желудка *E21*, *E36*, *V21*, *V43*, *VB24*, от своих номинальных значений; уровень психоэмоционального напряжения *УР* и уровень утомления *УУ*.

Результат исследования

Предложенный механизм использования модели Г. Раша в медицинских приложениях был опробован на решении задачи синтеза решающего правила для прогнозирования возникновения гастрита и язвы желудка с латентной переменной «Уверенность в прогнозе возникновения заболеваний желудка». Индикаторные переменные (факторы риска) те же, что и в *табл. 1*. После перехода к шкале K_i [формула (2)] данные модифицированной ТЭД были подвергнуты обработке диалоговым пакетом RUMM 2020. В результате обработки были получены следующие результаты.

Расчетное критическое значение уровня значимости, характеризующее соответствие данного набора индикаторов модели Раша, $\chi^2_{\text{критич. Prob}} = 0,954865$, что больше 0,05. Индекс сепарабельности (*Separation Index*) равен 0,98552. Коэффициент альфа Кронбаха (*Cronbach Alpha*) равен 0,9906. Мощность (оценка) системы индикаторов (*Power of Test-of-Fit*) оценива-

ется как превосходная (*excellent*). Полученные данные свидетельствуют о том, что все индикаторные переменные удовлетворяют модели Г. Раша.

Гистограмма распределения объектов обучающей выборки по шкале латентной переменной (уверенность в распределении объектов по следующим группам: низкий, средний и высокий риск заболеваний желудка) представлена на *рис. 2*.

Анализ гистограммы, представленной на *рис. 2*, показывает, что данные группируются в 3 блока, которые с учетом данных разведочного анализа и экспертных заключений легко интерпретируются как: 1) обследуемые не болеют исследуемым классом заболеваний (левый блок данных); 2) обследуемые с индикаторными переменными, составляющими «зону неопределенности» (средний блок данных); 3) обследуемые, которые в процессе обучения получили исследуемые заболевания желудка (правый блок данных). С учетом специфики решаемой задачи и руководствуясь рекомендациями работ [3], [5], эксперты, используя метод Делфи, на координатах *рис. 2* построили график функции принадлежности $\mu_{\omega_{ПЖ}}$ к классу «Высокий риск заболеваний желудка» $\omega_{ПЖ}$ с базовой переменной $L = UFF_{ПЖ}$ (сплошная квадратичная линия).

В соответствии с рекомендациями [1]-[3] уверенность в правильном прогнозе по заболеванию желудка может быть определена по формуле

$$UP_{Ж} = \mu_{\omega_{ПЖ}}(UFF_{ПЖ}). \quad (5)$$

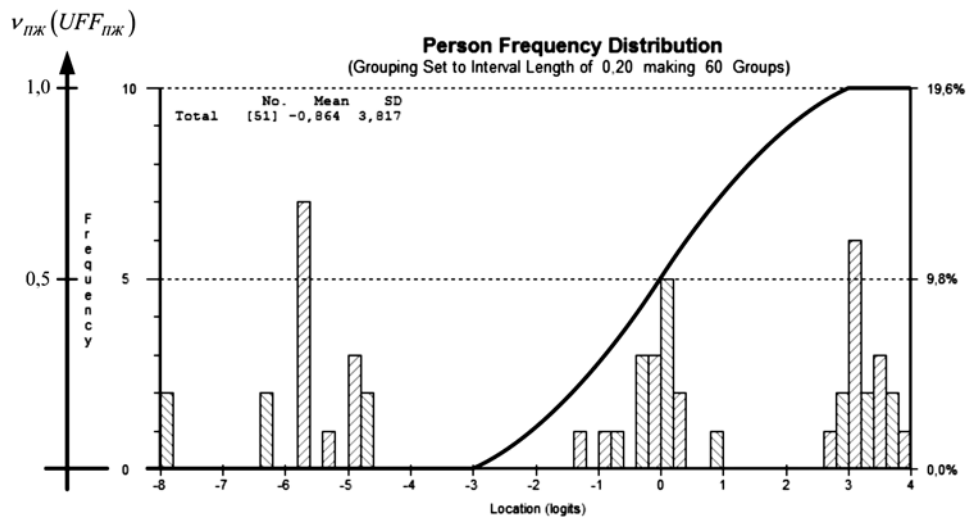


Рис. 2. Распределение интегрального показателя «Прогноз возникновения заболеваний желудка» в логитах

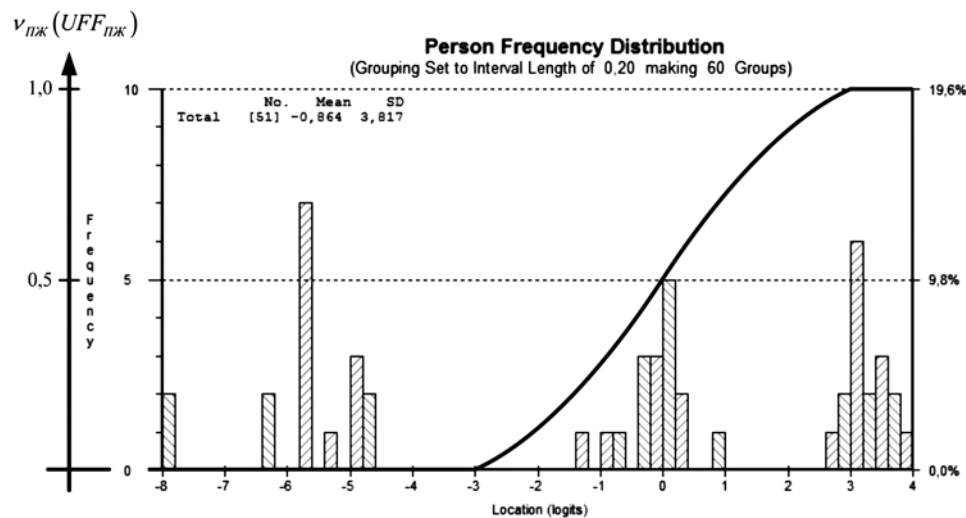


Рис. 3. График функции принадлежности $\mu_{\omega_{ПЖ}}$ к классу «Высокий риск заболеваний желудка» $\omega_{ПЖ}$ с базовой переменной $L = UFF_{ПЖ}$ (сплошная квадратичная линия)

Результат математического моделирования показал, что уверенность в принимаемых решениях по модели (5) составляет 0,88.

Таким образом, модель Г. Раша может быть использована для решения задач прогнозирования исследуемого класса заболеваний.

Заключение

Теория измерения латентных переменных с помощью модели Г. Раша в России и за рубежом хорошо зарекомендовала себя в педагогике, психологии, экономике и ряде других приложений. Ее возможности в медицине изучены достаточно слабо.

Проведенные нами исследования при решении задач прогнозирования и диагностики ранней стадии заболеваний желудка, прогнозирования обострения и оценки стадий заболеваний генитальным герпесом, выборе эффективных схем профилактики и лечения генитального герпеса показали перспективность использования моделей Г. Раша в медицинских приложениях при построении соответствующих систем поддержки принятия решений.

Список литературы:

1. *Корневский Н.А.* Использование нечеткой логики принятия решений для медицинских экспертных систем // Медицинская техника. 2015. № 1. С. 33-35.
2. *Корневский Н.А., Руцкой Р.В., Долженков С.Д.* Метод прогнозирования и диагностики состояния здоровья на основе коллективов нечетких решающих правил // Системный анализ и управление в биомедицинских системах. 2013. Т. 12. № 4. С. 905-909.
3. *Корневский Н.А., Филлист С.А., Устинов А.Г., Рябкова Е.Б.* Геометрический подход к синтезу нечетких решающих правил для решения задач прогнозирования и медицинской диагностики // Биомедицинская радиоэлектроника. 2012. № 4. С. 20-25.
4. *Al-Kasasbeh R., Korenevskiy N., Ionescou F., Alshamasin M.* Prediction and prenosological diaghostics of gastrointestinal tract diseases basec on enercteristic of acupuncture points and fuzzy logic / In: Bioinformatica biomedical technology: Proceedings of the 3th International Conference on Bioinformatics and Biomedical Technology (ICBVT 2011). Sanya, China, 2011. PP. 307-311.
5. *Маслак А.А.* Измерение латентных переменных в социально-экономических системах. Монография. – М., 2006. 333 с.

6. *Корневский Н.А., Крупчатников Р.А., Горбатенко С.А.* Синтез нечетких сетевых моделей, обучаемых по структуре данных для медицинских экспертных систем // Медицинская техника. 2008. № 2. С. 18-24.
7. *Корневский Н.А., Дегтярев С.В., Серегин С.П., Новиков А.В.* Интерактивный метод классификации в задачах медицинской диагностики // Медицинская техника. 2013. № 4. С. 1-3.
8. *Корневский Н.А.* Метод синтеза гетерогенных нечетких правил для анализа и управления состоянием биотехнических систем // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. 2013. № 2. С. 99-103.
9. *Корневский Н.А., Крупчатников Р.А., Аль-Касасбех Р.Т.* Теоретические основы биофизики акупунктуры с приложениями в медицине, психологии и экологии на основе нечетких сетевых моделей. – Старый Оскол: ТНТ, 2013. 528 с.

*Николай Алексеевич Корневский,
д-р техн. наук, профессор,
зав. кафедрой биомедицинской инженерии,
ФГБОУ ВПО «Юго-Западный
государственный университет»,
г. Курск,*

*Александр Николаевич Шуткин,
канд. физ.-мат. наук,
зам. начальника института по учебной работе,
ФГБОУ ВПО «Воронежский институт
государственной противопожарной
службы МЧС России»,
г. Воронеж,*

*Елена Алексеевна Бойцова,
канд. техн. наук, доцент,
зав. кафедрой высшей математики,
ФГБОУ ВПО «Юго-Западный
государственный университет»,
г. Курск,*

*Валентина Викторовна Дмитриева,
канд. техн. наук, ст. преподаватель,
кафедра электрофизических установок,
Национальный исследовательский
ядерный университет «МИФИ»,
г. Москва,
e-mail: kstu-bmi@yandex.ru*

ВНИМАНИЮ ПОДПИСЧИКОВ, РУКОВОДИТЕЛЕЙ СЛУЖБ ИНФОРМАЦИИ И БИБЛИОТЕК!

ПРЕДЛАГАЕМ ПОДПИСАТЬСЯ НА ЖУРНАЛ «МЕДИЦИНСКАЯ ТЕХНИКА»

НА 2016 ГОД.

Индекс по каталогу «Роспечать» – 72940.

В редакции можно оформить и оплатить льготную подписку с любого месяца.

Стоимость подписки : 1200 руб. – за один номер,

3600 руб. – на первое полугодие 2016 года (3 номера), 7200 руб. – на 2016 год (6 номеров).

Наши тел.: (495) 695-10-70, 695-10-71.