

Оценка эффективности практического применения модели непрерывного образования по прикладной статистике в медицинских вузах с применением программных решений на языке «R»

А. В. Копыцкий,

магистр естественных наук,

В. Н. Хильманович,

кандидат педагогических наук, доцент;

Гродненский государственный медицинский
университет

Прикладная статистика сегодня становится одним из базисных инструментов современной медицины. Это связано в первую очередь с широким внедрением в медицинскую практику положений доказательной медицины, исходя из которых специалист с высшим медицинским образованием должен правильно выбирать для диагностики, лечения и профилактики заболеваний методы, эффективность которых была доказана результатами специально организованных клинических испытаний или исследований.

Пандемия коронавирусной инфекции актуализировала и популяризировала среди академического сообщества, практикующих врачей, студентов медицинских вузов, а также пациентов интерес к научным медицинским исследованиям и интерпретации их результатов. Именно доказательная медицина располагает необходимой методологией для проведения научных медицинских исследований, их верификации и обобщения. Неотъемлемой частью доказательной медицины является прикладная статистика, дающая математический инструментарий, требуемый для правильного планирования эксперимента, сбора информации, ее последующего хранения, обработки и анализа. Поэтому важным является внедрение в учебный процесс студентов медицинских вузов дисциплин, связанных с прикладной статистикой.

Начиная с 2014 г. в учебных планах трех из четырех медицинских вузов Республики Беларусь появляются как компоненты УВО для большинства специальностей дисциплины, связанные с изучением основ математической статистики, а с 2021/2022 учебного года дисциплина «Биомедицинская статистика» приобретает статус государственного компонента в образовательных стандартах следующих специальностей: «Лечебное дело», «Педиатрическое дело», «Медико-психологическое дело», «Медико-профилактическое дело», «Сестринское дело». В свете вышеизложенного актуальной является такая организация процесса обучения прикладной статистике, которая будет учитывать специфику медицинского вуза и одновременно специфику статистических методов, используемых в медицинских и биологических научных исследованиях.

С момента введения в учебные планы студентов лечебного, педиатрического и медико-диагностического факультетов УО «Гродненский государственный медицинский университет» дисциплины «Основы статистики» были начаты поиски способов такой организации учебного процесса, при которой эффективность преподавания статистики была бы высокой.

В 2017/2018 учебном году кафедра медицинской и биологической физики университета разработала учебную программу для дисциплины «Основы статистики для профессорско-преподавательского состава» (позже – «Математическая статистика в медицине для профессорско-преподавательского состава»), а в 2019/2020 учебном году – для дисциплины курсов повышения квалификации «Математическая статистика в медицинских исследованиях». Кроме того, с 2019/2020 учебного года кафедра осуществляет обучение магистрантов по дисциплинам «Биоинформационный анализ биологических и медицинских данных» и «Медицинская статистика», где значительную часть объема учебного материала занимает прикладная статистика.

Таким образом, поиск подходящей модели организации обучения статистике является важным в рамках непрерывного образования по прикладной статистике в медицинском вузе. Модель должна иметь следующие отличительные особенности:

- соответствовать уровню математической подготовки обучающихся;
- быть практико-ориентированной;
- иметь связующие элементы для перехода от одной ступени образования к другой;
- быть гибкой (в контексте изменчивости образовательных стандартов, типовых учебных планов и учебных программ).

Нами была разработана и внедрена в учебный процесс прикладной статистики педагогическая модель непрерывного образования для медицинских вузов Республики Беларусь, обладающая перечисленными выше качествами [1]. Отличительной особенностью этой модели является то, что в ней используются оригинальные программные решения на языке программирования «R» [2].

Программные решения на языке «R» как неотъемлемая часть модели непрерывного образования в области прикладной статистики в медицинском вузе

Первое решение, которое является отличительной особенностью разработанной модели, – генератор шаблонных заданий по прикладной статистике. Данная программа используется на всех ступенях образовательного процесса в связке со средой дистанционного обучения (СДО) «Moodle» для создания большого количества однотипных заданий, различающихся числовыми данными. Отметим, что создание автоматизированных генераторов тестов на сегодняшний день является актуальной задачей при преподавании различных дисциплин [3–5], не имеющей универсального решения. Наша программа – генератор

тестовых заданий позволяет за короткое время создать необходимые тестовые задачи, объединить их в один файл со специфической разметкой (например, «gift» или «docx»). Далее полученный файл импортируется в СДО или используется непосредственно как раздаточный материал на письменных самостоятельных или контрольных работах. Программа является гибкой и позволяет путем изменения уже существующих шаблонов строить новые шаблоны и далее на их основе генерировать уже новые оригинальные задачи.

Применение описанного решения дает ряд полезных преимуществ по сравнению с использованием традиционных тестов с ограниченным спектром задач:

1. Число получаемых тестовых заданий настолько велико, что у студентов нет возможности за разумные сроки собрать базу правильных ответов. Кроме этого, у преподавателей есть возможность регулярно (по окончании семестра или учебного года) обновлять все задания. Таким образом, студенты вынуждены самостоятельно решать предоставленные им задания, что повышает объективность полученных результатов.

2. Снижается нагрузка на преподавателей, которые ведут практические занятия по дисциплинам, связанным с прикладной статистикой, в контексте сообщения, закрепления и проверки навыков решения практических задач. Разнообразие заданий, решаемых самостоятельно и многократно обучающимися, наличие системы размещения и автоматической их проверки позволяют преподавателю сконцентрировать свое внимание на том, как лучше подать материал и проработать его со студентами. Рутинная часть организации контроля знаний и скрупулезной проверки решенных заданий может быть исключена из учебной деятельности преподавателя.

3. Простота и гибкость программы-генератора позволяют оперативно адаптировать ее под нужды учебных программ, если в последних происходят изменения или необходимы изменения в сложности заданий. Данное программное решение можно менять для получения заданий и по другим дисциплинам. Так, наше решение было быстро переписано для генерации шаблонных заданий по темам дисциплины «Медицинская и биологическая физика» для обеспечения учебного процесса в 2019/2020 учебном году в условиях пандемии COVID-19 [6]. На базе данной программы было построено решение для быстрого наполнения тестами и практическими заданиями тем дисциплин, появившихся в 2021/2022 учебном году в учебных планах специальностей «Педиатрия» (дисциплина «Биомедицинская статистика») и «Медико-диагностическое дело» (факультативная дисциплина «Математическая статистика в медицинских исследованиях»). Указанное решение используется также для генерации заданий,

необходимых для закрепления навыков использования программы Statistica 10 для обработки медико-биологической информации слушателями проекта «Школа молодого ученого» (на базе УО «Гродненский государственный медицинский университет»).

4. Данное решение позволяет также создавать задания-тренажеры для обучающихся на второй ступени и курсах повышения квалификации.

Таким образом, программа – генератор тестовых заданий используется на всех ступенях образования как неотъемлемая часть педагогической модели непрерывного образовательного процесса в области прикладной статистики в медицинском вузе.

Еще одно программное решение, применяемое на второй ступени образования и курсах повышения квалификации, – программа для построения и фильтрации множеств регрессионных моделей, которые могут быть построены на выборках ограниченного объема. Данная программа используется в двух контекстах: в образовательном контексте – при прохождении обучающимися второй ступени тем, связанных с моделями множественной регрессии, в научно-образовательном контексте – при обработке результатов диссертационных исследований в случае необходимости построения регрессионных моделей на выборках ограниченного объема.

Отметим, что на сегодняшний день нет универсального подхода [7], позволяющего строить такие модели на выборках ограниченного объема: при числе переменных, сопоставимых с числом измерений, использовать методы пошагового включения или исключения не представляется возможным. Остается построение всех моделей с одной независимой переменной с одной и той же переменной-откликом. В дальнейшем независимые переменные, регрессионные коэффициенты которых были статистически значимыми, используются как возможные регрессоры, к которым уже применяются методы пошагового отбора. Такой подход имеет ряд недостатков: плохо работает на таблицах данных, содержащих пропущенные значения; не работает тогда, когда после его использования число регрессоров все еще сопоставимо с числом наблюдений; в нем не решается проблема мультиколлинеарности регрессоров. Альтернативой данному подходу является прямой перебор всех возможных моделей, построенных для всех возможных сочетаний предикторов, с последующей фильтрацией этих моделей. У такого экстенсивного метода есть ряд обоснований: число переменных-регрессоров в итоговых моделях, как правило, невелико (в медицине и биологии оно редко превышает 8–10); модели обычно содержат только те предикторы, коэффициенты которых статистически значимы; число возможных моделей может быть существенно ограничено путем

исключения моделей, содержащих статистически значимо связанные предикторы (так одновременно решается проблема интеркорреляции); благодаря цифровым технологиям время сбора медико-биологической информации обычно существенно больше времени ее обработки, поэтому несколько часов перебора и фильтрации моделей не могут сравниться с несколькими месяцами или годами сбора информации.

Наше решение возвращает электронную таблицу, содержащую характеристики всех возможных моделей, которые можно построить из заданного количества возможных регрессоров. Данная таблица используется как дидактический материал при рассмотрении тем, связанных с моделями множественной регрессии. По ней видно, как меняются характеристики качества подгонки регрессионных моделей при различных сочетаниях предикторов, что позволяет лучше понять роль этих характеристик и сам процесс классического пошагового включения/исключения регрессоров. Использование данной таблицы на занятиях с кейс-методом обучения позволяет оценить оптимальность выбора регрессионной модели исследователем в конкретном случае, согласованность логической и математической интерпретаций регрессионной модели. Эта же электронная таблица может быть использована для выбора оптимальной статистической модели связи между переменными при обработке реальных медико-биологических данных, полученных на выборках ограниченного объема. На основании описанной модели могут быть разработаны важные для медицины документы: методические рекомендации, патенты на изобретения, инструкции по применению препаратов или методов лечения, диагностики, прогноза. В биологических исследованиях статистическое моделирование может применяться для изучения закономерностей расселения популяции по ареалу обитания, временной динамики численности популяции, порогов действия различных сочетаний стимулов и факторов на живые организмы и т. д. Описанные научные аспекты результатов работы нашей оригинальной программы-переборщика оказываются важными для диссертационного исследования или научной работы обучающихся на второй ступени образования или курсах повышения квалификации. Подробно преимущества программы-переборщика моделей множественной регрессии описаны в нашей работе [8].

Таким образом, можно говорить о том, что наши программные решения позволяют технологизировать и цифровизировать многие аспекты образовательного процесса, акцентировать внимание на практикоориентированности, снизить нагрузку на преподавателей, корректно применять статистические методы в научных и диссертационных исследованиях.

Эффективность применения педагогической модели непрерывного образования по прикладной статистике в медицинском вузе

Программа-генератор была внедрена в учебный процесс по прикладной статистике в УО «Гродненский государственный медицинский университет» с 2015/2016 учебного года в рамках дисциплины «Основы статистики» и успешно используется по настоящее время. Потребность во внедрении связана с рядом особенностей преподавания данной дисциплины в медицинском вузе. Во-первых, уровень математической подготовки студентов-медиков не позволяет использовать в обучении элементы математического анализа (пределы, производные и интегралы), во-вторых, для сообщения и закрепления практических навыков существует необходимость привлечения дополнительных методов и цифровых технологий в образовательный процесс.

Анализ учебных программ показал, что в содержание учебного материала должны быть включены только те вопросы и методы, которые являются актуальными при статистической обработке результатов медико-биологических исследований и не требуют при их рассмотрении глубокой предварительной математической подготовки обучающихся в области высшей математики. Более сложные вопросы (наподобие множественной регрессии, статистического моделирования и пр.) целесообразно рассматривать на второй ступени образования, где увеличено количество аудиторных часов. Так, в содержательную часть учебной программы были включены такие методы прикладной статистики, которые чаще всего используются в медицинских и биологических исследованиях, встречаются в научных студенческих работах. Теоретические аспекты происхождения этих методов, доказательства и теоремы исключены. Акцент сделан на корректном применении этих методов (последовательность шагов, ограничения) и интерпретации их результатов.

Таким образом, был реализован инструктивный подход, что, по нашему мнению, является необходимым и обоснованным на обеих ступенях образования. Так как данный подход предполагает формирование и закрепление навыков путем многократного повторения шагов (инструкций), то актуальность программы-генератора состоит в следующем: она позволяет генерировать большое количество шаблонных тестовых практико-ориентированных заданий с постоянно изменяющимися числовыми данными, что исключает возможность заучивания правильных ответов и стимулирует обучающихся к самостоятельному решению. Это приводит к закреплению требуемого практического навыка как части компетенций будущего врача.

Для доказательства эффективности модели с программой-генератором рассмотрим динамику успеваемости студентов всех специальностей, изучающих прикладную статистику на первой ступени образования в УО «Гродненский государственный медицинский университет».

В таблице приведены описательные статистики (в виде $Me(Q1;Q3)$, где Me – медиана, $Q1$, $Q3$ – первый и третий квартили соответственно) среднего балла по дисциплинам, связанным с прикладной статистикой: «Основы статистики» (специальности 1-79 01 01, 1-79 01 04, 1-79 01 02) и «Математическая статистика в медицине» (специальность 1-79 01 05). Описательные статистики демонстрируют динамику возрастания среднего балла в период 2014/2015–2020/2021 учебных годов, а рост показателя для всех специальностей подтверждается критерием тенденций Джокхир – Терпстры ($p < 0,05$ для всех специальностей).

Приведем для большей наглядности графики динамики медиан среднего балла в описываемом временном интервале. Как видно из таблицы и рисунка, в 2015/2016 и 2016/2017 учебных годах происходит резкое увеличение медианных значений среднего балла для всех специальностей, после чего скорость нарастания показателя уменьшается (в 2017/2018 учеб-

Таблица

Описательные статистики среднего балла по прикладной статистике в период 2014/2015–2020/2021 учебных годов

Специальность	Учебный год							КДТ
	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020	2020/2021	
ЛД	4,1 (3,4;5,2)	6,7 (6;7,3)	7,3 (6,5;8)	7,7 (6,5;8,4)	7,7 (6,7;8,2)	7,3 (6,6;8)	7,6 (6,6;8,1)	$z = 21,24,$ $p = 0$
МДД	4,7 (3,7;5,4)	7 (6,3;7,5)	7,1 (6,4;7,8)	7,9 (7,6;8,3)	8 (7,3;8,4)	7,9 (7,4;8,4)	7,8 (7,5;8,4)	$z = 12,42,$ $p = 0$
ПД	4,1 (3,4;4,8)	6,2 (5,8;6,8)	7,4 (7;8,1)	7,2 (6,5;7,9)	7,4 (6,7;8,1)	7,9 (7,1;8,3)	7,8 (7,4;8,3)	$z = 13,74,$ $p = 0$
МПД	4,7 (3,9;5,7)	6 (5;6,5)	6,5 (6;7,2)	7,1 (6,5;7,9)	6,9 (5,9;7,6)	7,7 (7,2;8,1)	7,3 (6,6;7,9)	$z = 10,84,$ $p = 0$

Примечание: КДТ – статистика критерия Джокхир – Терпстры, ЛД – 1-79 01 01 «Лечебное дело», МДД – 1-79 01 04 «Медико-диагностическое дело», ПД – 1-79 01 02 «Педиатрическое дело», МПД – 1-79 01 05 «Медико-психологическое дело».

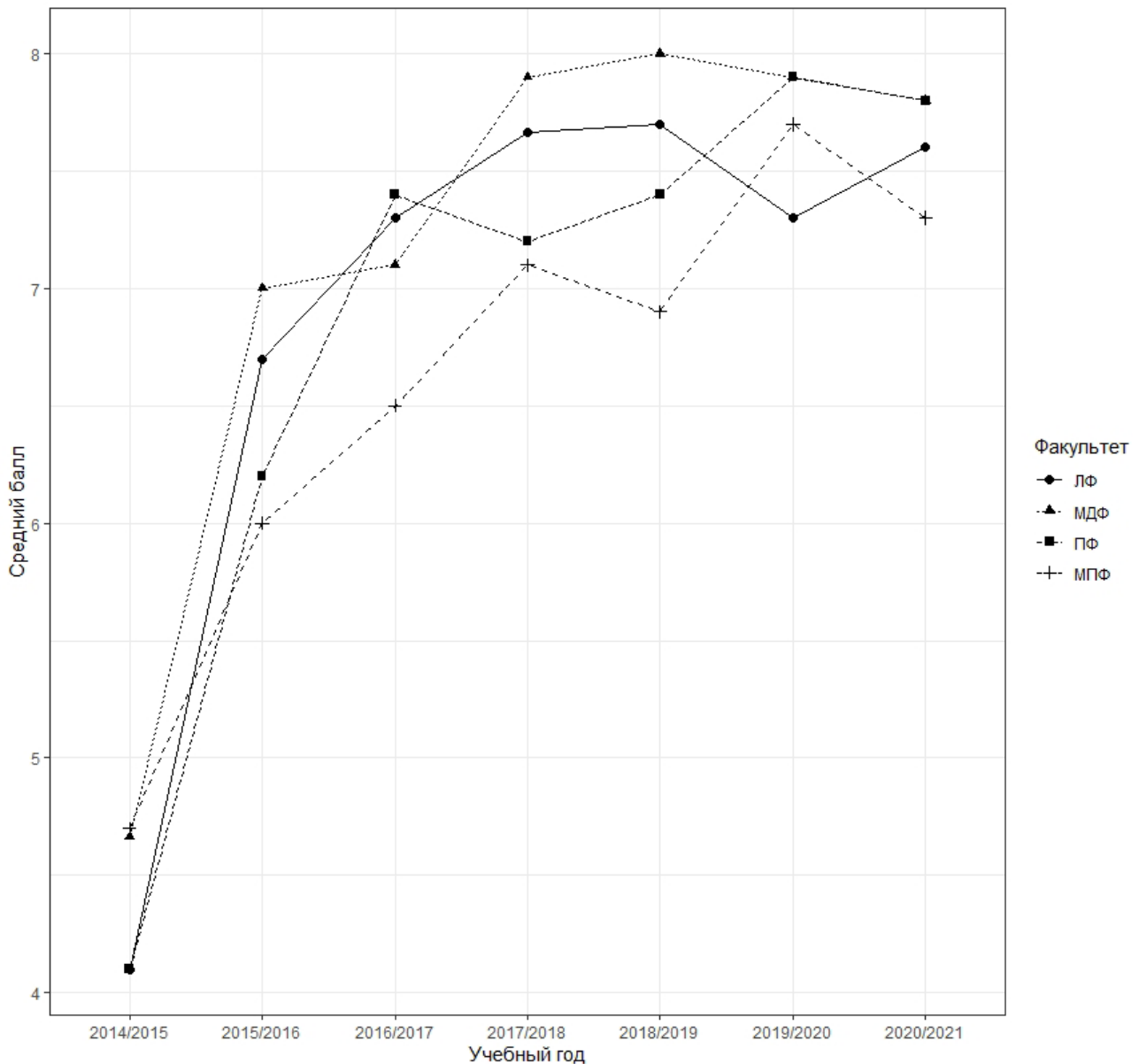


Рис. Графики динамики медиан среднего балла по прикладной статистике для специальностей 1-79 01 01, 1-79 01 04, 1-79 01 02, 1-79 01 05 в 2014–2021 учебных годах (см. обозначения в примечании таблицы)

ном году) и наступает своеобразное плато с небольшими флуктуациями среднего балла. Такое поведение можно объяснить именно внедрением в образовательный процесс нашего практико-ориентированного подхода с использованием генератора тестовых заданий на всех практических занятиях (9 занятий) дисциплины «Основы статистики». Внедрение программы-генератора в учебный процесс дисциплины «Математическая статистика в медицине» занимало более длительный временной период в силу возросшего числа аудиторных занятий, поэтому динамика роста среднего балла здесь несколько отличается.

Таким образом, модель непрерывного образования по прикладной статистике для медицинских ву-

зов с применением программы-генератора на языке «R» является эффективной на первой ступени образования.

На второй ступени образования эффективность модели обучения связана с результатами закрепления практических навыков применения специализированных компьютерных программ для статистического анализа. Здесь программа-генератор, реализованная на языке «R», используется в качестве тренажера. Численность обучающихся на второй ступени невелика: в период 2014–2021 гг. в нашем вузе обучение по прикладной статистике прошли три магистранта. Они успешно овладели практическими навыками и получили высокий результат итогового тестирования

в рамках зачетного или экзаменационного занятия по статистике (средний балл обучающихся составил 8,5).

Курсы повышения квалификации и переподготовки кадров в области прикладной статистики прошли 30 слушателей в период с 2019 по 2021 г. и 33 слушателя курса «Математическая статистика в медицине» для профессорско-преподавательского состава нашего вуза, организованного на базе кафедры медицинской и биологической физики в период с 2017 по 2019 г. На этой ступени актуальным является также наше программное решение для перебора и фильтрации регрессионных моделей. За описываемый период им воспользовались 12 обучающихся, с 2 из них получены патенты на полезные изобретения, 5 обучающихся использовали результаты работы программы-переборщика при написании диссертационных исследований.

Как следует из вышеизложенного, разработанная нами модель непрерывного образования по прикладной статистике в медицинском вузе является эффективной на основании проверки с помощью методов математической статистики. Отличительными особенностями педагогической модели являются:

1. Практикоориентированность, обусловленная отсутствием у студентов предварительных знаний по высшей математике и малым количеством аудиторных часов.

2. Инструктивный подход при подаче учебного материала.

3. Наличие двух программных решений на языке программирования «R». Первое решение позволяет в связке с СДО технологизировать и цифровизировать процесс обучения прикладной статистике, повысить успеваемость студентов, а также упростить его методическое сопровождение. Второе решение позволяет применять его результаты для научных исследований, связанных с множественной регрессией и стратегиями отбора предикторов в статистические модели.

Список использованных источников

1. Копыцкий, А. В. Трехступенчатая модель непрерывного образования по прикладной статистике в медицинских вузах с применением программных решений на языке «R» / А. В. Копыцкий, В. Н. Хильманович // Педагогическая наука и образование. – 2022. – № 1(38). – С. 56–64.

2. Копыцкий, А. В. Программная среда «R» как составная часть методики преподавания математической статистики в медицинском вузе / А. В. Копыцкий, В. Н. Хильманович // Информатизация и методика электронного обучения: цифровые технологии в образовании: материалы V Междунар. науч. конф. / Сибир. федерал. ун-т. – Красноярск, 2021. – С. 245–249.

3. Посов, И. А. Обзор генераторов и методов генерации учебных заданий / И. А. Посов // Образовательные технологии и общество. – 2014. – Т. 17, № 4. – С. 593–609.

4. Романенко, В. В. Обзор технологий автоматизированного создания и публикации тестовых заданий в системах дистанционного обучения / В. В. Романенко, И. О. Аксененко // Информационные технологии в науке, управлении, социальной сфере и медицине: сб. науч. тр. IV Междунар. науч. конф.: в 2 ч. / под ред. О. Г. Берестневой, А. А. Мицеля, Т. А. Гладковой; Том. политехн. ун-т. – Томск: Изд-во Том. политехн. ун-та, 2017. – Ч. 1. – С. 317–321.

5. Швецов, А. Н. Опыт применения метода автоматической генерации тестовых заданий / А. Н. Швецов, А. П. Сергушичева // Образовательные технологии и общество. – 2017. – Т. 20, № 4. – С. 318–333.

6. Копыцкий, А. В. Использование языка программирования «R» для генерации тестовых заданий по дисциплине «Медицинская и биологическая физика» / А. В. Копыцкий, В. Н. Хильманович, С. И. Клинецвич // Информатизация образования и методика электронного обучения: цифровые технологии в образовании: материалы IV Междунар. науч. конф., Красноярск, 6–9 окт. 2020 г. – Красноярск, 2020. – С. 191–195.

7. Мاستицкий, С. Э. Статистический анализ и визуализация данных с помощью R / С. Э. Мастицкий, В. К. Шитиков. – М.: ДМК Пресс, 2015. – 496 с.

8. Копыцкий, А. В. Информационно-вычислительная технология с применением языка «R» в рамках второй ступени высшего образования в медицинских вузах / А. В. Копыцкий // Высшая школа. – 2021. – № 3. – С. 18–22.

Аннотация

В работе описана модель непрерывного образования по прикладной статистике в медицинских вузах с применением программных решений на языке программирования «R». Рассмотрены предпосылки создания модели, требования, которым удовлетворяет оригинальная модель, ее особенности, роль применяемых в ней программных решений. Показан статистически значимый рост ($p < 0,05$) среднего балла обучающихся прикладной статистике на первой ступени и востребованность программы-переборщика на второй ступени в УО «Гродненский государственный медицинский университет» в период с 2014 по 2021 г.

Abstract

The paper describes a model of continuous education in applied statistics in medical universities using software solutions in the R programming language. The prerequisites for creating the model, the requirements that the original model satisfies, its features, the role of software solutions used in it are considered. A statistically significant increase ($p < 0,05$) in the average score of students in applied statistics at the first stage and the demand for a program-enumerator at the second one at the Grodno State Medical University in the period from 2014 to 2021 are shown.