

сферах его деятельности. Эти преобразования должны учитывать современные технические возможности предоставления информационных услуг и обеспечивать распространение информации в обществе, формировать глобальное информационное пространство. Важнейшей частью этих процессов является компьютеризация общества.

Использование в образовании электронных и информационно-образовательных систем облегчает поиск необходимой информации в стремительно возрастающем ее потоке. С помощью информационных технологий эффективно решается проблема доступности и наглядности обучения [2]. Электронные информационные технологии широко используются в обучающих целях, способствуют усвоению и отработке учебного материала при подготовке к семинарским и практическим занятиям поиску необходимого учебного материала для докладов, рефератов, презентаций. Широко применяются электронные информационные технологии при проведении студентами научно-исследовательской работы, подготовке курсовых и дипломных работ, проектов.

Эффективность использования информационных технологий зависит и от организации самостоятельной работы. В сторону, которой изменено соотношение времени, отводимого на аудиторную и самостоятельную работу.

В информационном обществе информация является главнейшим и определяющим фактором его развития. Современные технологии помогают человеку познавать новое и соответствовать времени и требованиям.

#### ЛИТЕРАТУРА

- 1 Абасов, С.Э. Современные информационно-коммуникационные технологии в образовании / С.Э. Абасов. – 2011.
- 2 Кажаява, Е.А. Внедрение ИКТ-технологий в образовательный процесс / Е.А. Кажаява. – 2013.

УДК 378.016:[51:004.42

### **ИСПОЛЬЗОВАНИЕ КОМПЬЮТЕРНОЙ ПРОГРАММЫ ДЛЯ ПЕРЕБОРА МОДЕЛЕЙ РЕГРЕССИИ В РАМКАХ УЧЕБНОГО КУРСА «МАТЕМАТИЧЕСКАЯ СТАТИСТИКА В МЕДИЦИНЕ»**

**А.В. Копыцкий, В.Н. Хильманович**

УО «Гродненский государственный медицинский университет» (Республика Беларусь, 230009, г. Гродно, ул. М. Горького, 80; e-mail: fizika@grsmu.by)

Аннотация. В статье рассматривается использование компьютерной программы для перебора моделей регрессии в образовательном процессе в рамках факультативного курса «Математическая статистика в медицине» для профессорско-преподавательского состава Гродненского государственного медицинского университета.

Ключевые слова: программное обеспечение, математическая статистика, регрессионные модели, медицинское образование.

## USING OF COMPUTER PROGRAM PURPOSED FOR ENUMERATION OF REGRESSION MODELS WITHIN THE FRAMEWORK OF TRAINING COURSE “STATISTICS IN MEDICINE”

**A.V. Kapytski, V.N. Khilmanovich**

El «Grodno State Medical University» (Belarus, Grodno, 230009, 80 M. Gorkogo st.; e-mail: fizika@grsmu.by)

Summary. Application of computer program, purposed for enumeration of regression models, in educational process of facultative course “Statistics in medicine” for faculty stuff of Grodno state medical university is considered considered.

Key words: software, mathematical statistics, regression models, medical education.

Математическая статистика – это научная дисциплина, предметом изучения которой является разработка методов регистрации, описания и анализа статистических экспериментальных данных, полученных в результате наблюдений массовых случайных явлений. Учебная дисциплина «Математическая статистика в медицине» занимает важное место в системе подготовки врача по всем специальностям. В нашем университете эта дисциплина преподаётся с 2012 года как компонент УВО, с 2017 года и по сегодняшний день данная дисциплина в расширенном варианте преподаётся факультативно для профессорско-преподавательского состава университета, аспирантов и соискателей учёной степени; с осени 2019 года на базе нашей кафедры открываются курсы повышения квалификации по дисциплине «Математическая статистика в медицинских исследованиях».

Большой интерес для специалиста, исследующего сложные явления, представляют связи между изучаемыми переменными, а также совокупное влияние переменных («предикторов») на некоторую целевую переменную («отклик»). Это влияние изучается во множественном регрессионном анализе или его обобщённом варианте – обобщённой линейной модели [1].

Обобщённая линейная модель с бинарным откликом с логит- или пробит-функцией связи может использоваться для классификации. Однако зачастую из-за малого объёма выборки исследователь не может использовать такие хорошо известные методы построения регрессионных моделей, как пошаговое включение и исключение предикторов, что не позволяет ему определить их наилучшее сочетание для предсказания значений данной переменной-отклика. Исследователю в таком случае остаётся выполнять только ручной перебор возможных сочетаний предикторов, но при большом числе предикторов счёт их возможных комбинаций идёт на тысячи, десятки и сотни тысяч и даже на миллионы. Данный ручной перебор может быть заменён автоматизированным, выполняемым с помощью компьютера.

Одной из тем научной работы нашей кафедры является разработка программного обеспечения для автоматизированного перебора всех возможных регрессионных моделей для некоторой переменной-отклика при  $m$  возможных предикторах. Нами уже разработан пакет расширения языка программирования «R», позволяющий строить множества моделей бинарной регрессии с логит- и пробит-функциями связи. Сама программа находится в свободном доступе, например в [2]. Данный пакет используется для анализа данных пилотных исследований, в которых как раз собирается большое

количество данных на малых выборках. Результаты работы данного «переборщика» позволяют исследователям наметить пути дальнейшего углубления исследования.

Хотя данная компьютерная программа предназначена для научных целей, она нашла также применение в образовательном процессе. Во-первых, она используется для демонстрации возможностей языка программирования «R», предназначенного для проведения статистического анализа данных. Во-вторых, с помощью программы мы тщательно рассматриваем методологические аспекты регрессионного моделирования. Так как программа возвращает чаще всего большое количество возможных моделей регрессии, то мы рассматриваем количественные меры качества подгонки моделей и то, как они меняются при переходе от модели к модели. Определяем, какими показателями можно пользоваться для выбора оптимальной модели. Мы акцентируем внимание слушателей курса на том, что нельзя слепо выбирать модель с наилучшими чисто математическими показателями, ведь одно из главных достоинств регрессионных моделей – их интерпретируемость. Поэтому мы рекомендуем выбирать для дальнейшего анализа не просто наилучшую с точки зрения математики модель, но наиболее просто и логично интерпретируемую модель, пусть и с худшими математическими характеристиками. В-третьих, так как программа строит все возможные модели регрессии, то вполне вероятно появление среди них моделей, где обнаруживается чисто случайная связь между данной переменной-откликом и некоторой линейной комбинацией предикторов. При этом никакой внятной интерпретации с точки зрения медицины или биологии данная модель не имеет. Это позволяет нам в очередной раз акцентировать внимание на том, что обнаруженные на выборочных данных связи между переменными могут иметь чисто случайный характер, поэтому требуется крайне осторожно описывать результаты корреляционного и регрессионного анализа, помня о том, что обнаруженные связи могут не иметь места в генеральной совокупности. Отсюда мы приходим к выводу о том, что найденные в результате одного исследования связи должны быть, по возможности, перепроверены в ходе более масштабного исследования, включающего большее количество испытуемых, или в исследованиях других авторов, что осуществляется в ходе, например, метаанализа источников. Таким образом, мы подчёркиваем важность научных работ, посвящённых метаанализу, так с их помощью можно уверенно говорить об эффективности, например, новой терапии или препарата.

#### ЛИТЕРАТУРА

1. Мاستицкий, С. Э. Статистический анализ и визуализация данных с помощью R. Цветное зрение / С.Э. Мاستицкий, В.К. Шитиков. – Москва: ДМК-Пресс, 2015. – 496с.
2. R Core Team. R: A Language and Environment for Statistical Computing / R Core Team. – Vienna, Austria: R Foundation for Statistical Computing, 2107.