

КОКОРИНА Ирина Владимировна, соискатель кафедры методики преподавания математики Института математики и компьютерных наук Северного (Арктического) федерального университета имени М.В. Ломоносова, старший преподаватель кафедры естественнонаучных дисциплин и информационных технологий Гуманитарного института филиала Северного (Арктического) федерального университета имени М.В. Ломоносова в г. Северодвинске. Автор трех научных публикаций

ОСОБЕННОСТИ ОБУЧЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ СТУДЕНТОВ-ФИЛОЛОГОВ

В статье рассматриваются вопросы методики обучения теории вероятностей и математической статистике студентов-филологов в рамках дисциплины «основы математической обработки информации». Основу предлагаемой методики составляет идея усиления профессиональной направленности за счет выделения в качестве целевых элементов содержания вероятностно-статистических методов лингвистического анализа и использования ИКТ.

Ключевые слова: обучение математике в вузе, профессиональная направленность, теория вероятностей и математическая статистика, методы лингвистического анализа, информационная поддержка обучения.

В мае 2010 года Министерством образования и науки были утверждены ФГОС ВПО третьего поколения. Их идейную основу составляет компетентностный подход, призванный решить задачу усиления профессиональной направленности всех составляющих программы ВПО.

Реализация этой идеи применительно к математическим дисциплинам, входящим в программу подготовки бакалавров по направлению 050100.62 – педагогическое образование, филологических профилей: иностранный

язык, русский язык и литература, выражается в замене традиционной дисциплины «математика и информатика» на дисциплину «основы математической обработки информации», одной из целей которой, по мнению М.Н. Дмитриевой, является развитие навыков использования «математических методов и основ математического моделирования»¹.

Само название этой новой дисциплины нацеливает разработчиков учебных программ на отбор и достаточно полное представление

в содержании дисциплины тех математических методов обработки информации, которые наиболее значимы для специалистов данных профилей.

Проведенный нами анализ трудов отечественных и зарубежных лингвистов и математиков (В.В. Кроммер, Дж.Х. Гринберг, М.В. Панов, Б.Н. Головин, Р.Г. Пиотровский, А.В. Гладкий, Н.Ф. Алефиренко, З.И. Тарланов, В.А. Звегинцев и др.) показал, что наиболее широкое применение в современных лингвистических исследованиях находят методы, опирающиеся на положения теории вероятностей и математической статистики. Основной целью этих методов Н.Ф. Алефиренко называет «раскрытие закономерностей функционирования единиц языка и речи, а также установление закономерностей построения текстов»².

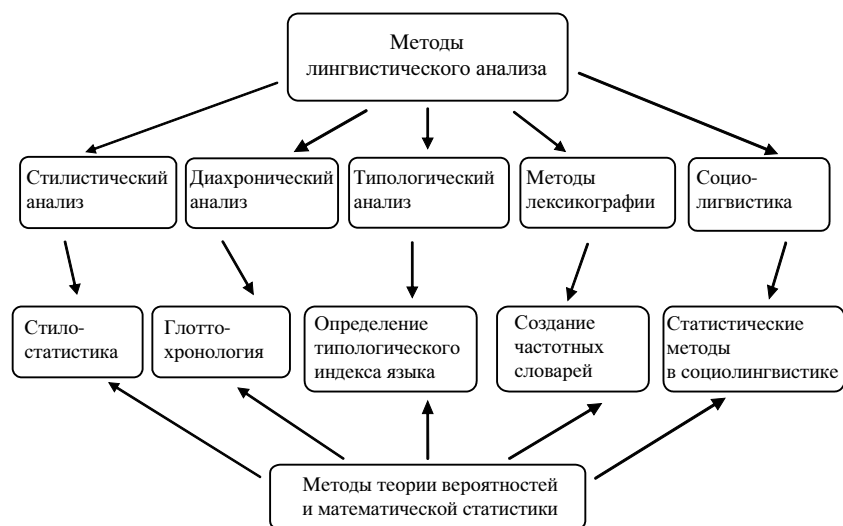
Наглядное представление о месте методов теории вероятностей и математической статистики в системе методов лингвистического анализа может дать *рисунок*.

Принцип профессиональной значимости требует переосмысления традиционного содержания с позиции профессиональной значимости его элементов, а также дополнения его новыми вопросами, составляющими теорети-

ческую основу наиболее популярных методов лингвистического анализа. Полученная таким образом номенклатура вопросов составляет целевой компонент содержания дисциплины, т. е. является основной целью ее изучения. Последний принцип – логической целостности требует дополнения целевого содержания дисциплины вспомогательными элементами, теми элементами содержания, которые хоть и не являются профессионально значимыми, но необходимы для понимания целевых элементов.

Новая дисциплина должна, по нашему мнению, не только сохранить прежнюю связь с информатикой (установленную дисциплиной «Математики и информатика»), но даже усилить ее. Необходимость такого усиления определяется тем, что применение математических методов для решения задач лингвистического анализа сегодня не обходится без использования компьютерных средств.

Программа новой дисциплины должна предусматривать три типа занятий: лекционные, практические и лабораторные. Лекционные занятия имеют целью ознакомление студентов с математическими методами и их теоретическими основами. Практические занятия направлены на включение студентов в деятельность,



Методы теории вероятностей и математической статистики в лингвистическом анализе

**ЦЕЛЕВОЕ СОДЕРЖАНИЕ РАЗДЕЛА «ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА» ДИСЦИПЛИНЫ
«ОСНОВЫ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ ИНФОРМАЦИИ»**

Элементы	Области лингвистических приложений элементов
<i>1. Традиционно рассматриваемые вопросы</i>	
Определения вероятности события	Составление частотных словарей
Вероятность суммы и произведения событий	Вычисление вероятности цепочек языковых элементов
Формула полной вероятности. Теорема Байеса	Определение логической вероятности. Измерение семантической информации текста
Формула Бернулли. Предельные теоремы в схеме Бернулли	Вероятностное моделирование порождения текста и составляющих его единиц
Случайные величины. Основные законы распределения СВ	Случайная лингвистическая величина. Моделирование образования языковых единиц текста
Нормальное распределение НСВ	Логнормальное распределение вероятностей длин текстовых словоупотреблений
Двумерная случайная величина	Корреляция случайных лингвистических величин
Основные понятия математической статистики. Выборки и их характеристики	Первичная статистическая обработка текста
<i>2. Дополнительные вопросы</i>	
Элементы теории оценок	Статистическая модель текста. Вероятностные характеристики норм языка
Проверка гипотез	Исследование вероятности свойств языка и статистики текста с помощью метода гипотез
Корреляционный анализ	Проверка гипотез о связи двух лингвистических выборок
Дисперсионный анализ	Сравнение более двух лингвистических выборок по признаку, измеренному в метрической шкале
Факторный анализ	Анализ зависимости множества признаков лингвистических переменных

позволяющую осмыслить суть представленных понятий, положений и методов, а лабораторные занятия – на овладение компьютерными средствами решения профессиональных задач с использованием изученных методов математики.

Приведем пример задачи лингвистического исследования³, которая может быть решена «вручную» на практическом занятии и с помощью компьютерных программ «Excel», «Statistica» или «SPSS» на лабораторной работе.

Задача. Ученые, которые занимаются вопросами характеристики авторских стилей, используют в качестве параметра стиля среднюю длину словоупотребления. Можно ли среднюю длину словоупотребления использовать также в качестве параметра для различения языков? Дай-

те обоснованный ответ на этот вопрос, опираясь на данные о средних длинах словоформ в языках мира, которые представлены в *таблице*⁴.

Решение.

1. Если вариационный ряд средних длин словоформ близок к нормальному распределению, то средние длины словоформ равномерно группируются вокруг средней, задаваемой возможностями оперативной памяти человека. Отклонение от этой средней в каждом конкретном языке будет рассматриваться как результат случайных воздействий.

2. Применим критерии асимметрии и эксцесса. Эти критерии определяют допустимую степень отклонения эмпирических значений асимметрии и эксцесса от нулевых значений, соответствующих нормальному распределению.

Для решения приведенной задачи с использованием программы SPSS необходимо выбрать «Анализ», «Описательные статистики» и задать параметры «Асимметрия» и «Экссесс». В таблице результатов содержатся значения асимметрии и эксцесса и соответствующие им стандартные ошибки. «Распределение соответствует нормальному виду, если для соответствующей переменной абсолютные значения асимметрии и эксцесса не превышают свои стандартные ошибки»⁵. В нашем примере это справедливо

Ответ. Средняя длина словоформ во всех языках мира группируется около среднего значения, определенного возможностями опе-

ративной памяти человека, а не типическими особенностями каждого языка. Следовательно, **средняя длина словоформ не может считаться параметром для различения языков мира.**

Приведенный пример показывает, что компьютерное решение профессиональной задачи концентрирует внимание студентов не на технике математических вычислений, а на условиях выбора математического метода и интерпретации результатов его применения.

Курс «Основы математических методов обработки информации» построенный подобным образом, будет не только понятен, но интересен и полезен студентам-филологам.

Примечания

¹ Дмитриева М.Н. Методическая система интенсификации обучения математике студентов гуманитарных специальностей вузов // Вестн. Помор. ун-та. Сер.: Гуманит. и соц. науки. № 6. 2009.

² Алефиренко Н.Ф. Современные проблемы науки о языке: учеб. пособие. М., 2005. С. 383.

³ Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А. Математическая лингвистика: учеб. пособие для пед. ин-тов. М., 1997. С. 262.

⁴ Там же. С. 259.

⁵ Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных: учеб. пособие. СПб., 2004. С. 60.

Kokorina Irina Vladimirovna

Northern (Arctic) Federal University named after M.V. Lomonosov,
Institute of Mathematics and Computer Sciences,
Humanitarian Institute in Severodvinsk

PECULIARITIES OF TEACHING PHILOLOGY STUDENTS THE PROBABILITY THEORY AND MATHEMATICAL STATISTICS

The article deals with the issues of teaching philology students the theory of probability and mathematical statistics at the course of lectures on the «Basics of Mathematical Information Processing». The suggested method is based on the idea of strengthening professional orientation by taking probabilistic and statistical methods of linguistic analysis as key elements of the content as well as applying information and communication technologies.

Key words: *teaching mathematics in university, professional orientation, the theory of probability and mathematical statistics, methods of linguistic analysis, information support of teaching.*

Контактная информация:
e-mail: kokorina_ira@mail.ru

Рецензент – Шабанова М.В., доктор педагогических наук, профессор кафедры методики преподавания математики Института математики и компьютерных наук Северного (Арктического) федерального университета имени М.В. Ломоносова