

УДК 378:519

ПРОЦЕССУАЛЬНЫЙ АСПЕКТ ПРОФЕССИОНАЛЬНОЙ НАПРАВЛЕННОСТИ МАТЕМАТИЧЕСКОГО ОБРАЗОВАНИЯ СТУДЕНТОВ-ФИЛОЛОГОВ И ЕГО РЕАЛИЗАЦИЯ ПРИ ИЗУЧЕНИИ СТОХАСТИЧЕСКИХ МЕТОДОВ МАТЕМАТИЧЕСКОЙ ЛИНГВИСТИКИ

Кокорина И. В.

Северный (Арктический) федеральный университет имени М. В. Ломоносова, Архангельск, Россия (163000, г. Архангельск, наб. Северной Двины, д. 17), e-mail: webmaster@narfu.ru

В статье описан результат работы по созданию комплекса методических средств, систематическое применение которых способствует формированию у студентов филологических специальностей опыта использования методов теории вероятностей и математической статистики в профессиональной деятельности. Теоретическую основу предлагаемого в статье подхода составляет концепция контекстного обучения, развиваемая в трудах А. А. Вербицкого. Предлагаемый автором статьи комплекс методических средств основан на идее обеспечения поэтапного перехода от обучения решению учебных стохастических задач к профессионально-ориентированным, а от них и к профессиональным лингвистическим стохастическим задачам. Представлен курс профессионально ориентированных лабораторных работ для студентов-филологов, приводится решение профессиональной задачи лингвистического анализа с помощью факторного дисперсионного анализа на примере одной из таких работ. Освещается вопрос обучения использованию программных продуктов специального назначения, в частности, пакета анализа данных SPSS Statistics, для обработки и анализа статистических данных.

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

REMEDIAL ASPECT OF A PROFESSIONAL ORIENTATION OF MATHEMATICAL FORMATION OF STUDENTS-PHILOLOGISTS AND ITS REALIZATION AT STUDYING STOCHASTIC METHODS OF MATHEMATICAL LINGUISTICS

Kokorina I. V.

Northern (Arctic) federal university named after M. V. Lomonosov, Arkhangelsk, Russia (163002, Arkhangelsk, Severnaya Dvina Emb. 17), e-mail: webmaster@narfu.ru

The article describes the result of the work on creation of a complex of teaching students of philology, the systematic application of which contributes to the formation of the experience of using methods of the theory of probability and mathematical statistics in the professional activity. The approach proposed in the article is based on the concept of contextual education developed in the works of A. A. Verbitsky. The complex of teaching offered by the author of the article is based on a gradual transition training to solve from stochastic problems to professionally-oriented ones, and from the latter to the professional linguistic stochastic problems. The course of professionally oriented laboratory works for students of philology is presented; the solution to professional tasks of linguistic analysis with the help of factor analysis of variance is given. The coverage of the issue of teaching to use software products of special purpose, in particular Statistical Package for the Social Sciences for the purpose of processing and analyzing statistical data is presented.

Key words: mathematics education in institutions of higher education; professional orientation; probability theory and mathematical statistics; methods of linguistic analysis; informational support for education.

В современный период развития Высшего профессионального образования большое значение придаётся профессиональной направленности обучения. По мнению Г. И. Худяковой, «принцип профессиональной направленности в вузе разрешает противоречие между теоретическим характером изучаемых дисциплин и практическим умением применять эти знания в профессиональной деятельности» [5]. Особенно это значимо для студентов гуманитарных специальностей, так как профессиональная

направленность является «основой мотивации и гуманизации обучения студентов будущей профессии» [5].

С целью усиления профессиональной направленности обучения математике для студентов филологических специальностей профилей подготовки «Иностранный язык», «Русский язык», «Литература», реализуемых в рамках общего направления подготовки 050100.62 «Педагогическое образование» стандартами нового поколения, вместо привычной дисциплины «математика», «математика и информатика» введена дисциплина «Основы математической обработки информации».

Пути решения задачи усиления профессиональной направленности подсказывают трактовки понятия «профессиональная направленность обучения» Г. И. Худяковой, М. И. Зайкина, С. В. Щербатых. Они выделяют **содержательный и процессуальный аспекты** решения данной задачи.

Содержательный аспект проявляется в насыщении курса профессионально значимым содержанием – сведениями о теоретических основах решения профессиональных задач. В настоящее время ему уделяется наиболее пристальное внимание. Проведенный нами анализ трудов отечественных и зарубежных лингвистов и математиков, таких как И. А. Бодуэн де Куртенэ, М. Н. Петерсон, В. В. Виноградов, Р. Г. Пиотровский, А. В. Гладкий, В. А. Звегинцев, М. В. Панов, Н. Ф. Алефиренко, Б. Н. Головин и других, показал, что решение проблемы усиления профессиональной направленности математического образования студентов-филологов видится в увеличении доли теории вероятностей и математической статистики в структуре курса, а также включении в его содержание стохастических методов математической лингвистики: стилостатистики, глоттохронологии, статистических методов в лексикографии и других.

Менее разработанным, но не менее важным является **процессуальный аспект** усиления профессиональной направленности. Он связан с созданием комплекса методических средств, систематическое применение которых способствует формированию у студентов опыта использования научных знаний в профессиональной деятельности. Мы предлагаем формировать этот опыт путем организации поэтапного перехода от обучения решению учебных стохастических задач к профессионально-ориентированным, а от них и к профессиональным лингвистическим стохастическим задачам. Для понимания различий между этими видами задач необходимо дать характеристику различий видов деятельности, которую они определяют.

Таблица 1

Характеристика различий видов деятельности

| <i>Виды</i> | <i>Виды</i> | <i>Характеристика</i> | <i>Место</i> |
|-------------|-------------|-----------------------|--------------|
|-------------|-------------|-----------------------|--------------|

| <i>деятельности обучающихся в системе ВПО (А. Вербицкий)</i> | <i>лингвистических стохастических задач</i> | <i>видов лингвистических стохастических задач</i> | <i>в структуре курса</i> |
|--|---|--|---|
| Учебная деятельность академического типа | Учебные стохастические задачи | Задачи на применение теоретических положений ТВ и МС на типичных моделях лингвистических опытов со случайным исходом | Лекционные, практические занятия |
| Квазипрофессиональная | Профессионально-ориентированные стохастические задачи | Сюжетно построенные лингвистические стохастические задачи, сюжет которых вводит в профессиональную ситуацию постановки таких задач | Практические занятия, подготовительный этап лабораторных работ |
| Учебно-профессиональная | Профессиональные стохастические задачи | Задачи, которые реально возникли или могли бы возникнуть в процессе лингвистического исследования и решение которых требует обращения к стохастическим методам | Лабораторные занятия, научно-исследовательская практика, курсовые работы, выпускные квалификационные работы |

Для содержательной иллюстрации различий задач, представленных в таблице, рассмотрим конкретные примеры.

Пример 1 (учебная лингвистическая стохастическая задача)

Лингвистами установлено, что «статистическая вероятность появления глагола в произведениях Л. Н. Толстого равна 0,21, в произведениях А.И. Куприна – 0,15» [2]. Из текста каждого автора случайным образом выбирают по одному слову. Составить закон распределения случайной величины X – «Количество глаголов в выборке» и определить среднее квадратическое отклонение $\sigma(X)$ случайной величины X . Для данной случайной величины найти функцию распределения $F(x)$ и построить её график.

Приведенный пример показывает, что сюжетная составляющая такой задачи выполняет лишь мотивационную функцию, показывая студентам применимость изучаемой теории в лингвистике. Описание ситуации применения теории и требований здесь представлено уже с использованием всей необходимой стохастической терминологии, поэтому успешность решения задачи зависит лишь от наличия знаний соответствующей математической теории. Такие задачи могут быть использованы на лекциях для демонстрации типичных ситуаций применения введенной теории или на практических занятиях для формирования навыков ее применения.

Пример 2 (профессионально-ориентированная лингвистическая стохастическая задача)

«Для расчёта памяти автомата, распознающего устную речь, и построения алгоритма его работы приходится вычислять вероятность совпадения хотя бы одной из словоформ обрабатываемого текста с соответствующей лексемой, заданной в словаре автомата. Выбрано два одинаковых по объёму отрывка текста, из каждого отрывка случайным образом выбирается слово. Нужно определить вероятность того, что хотя бы одно из двух выбранных слов будет местоимением *он*» [4].

Этой задаче присущи черты как учебной, так и профессиональной задачи. В сюжетной составляющей на языке учебной информации описываются условия и содержание реального производственного процесса. Успех решения будет зависеть не только от знания теоретических положений теории вероятностей, но и от выбора вероятностной модели, которая будет использована. Задачи такого вида могут быть предложены студентам для решения на практических занятиях, а также в качестве подготовительных, на лабораторных работах.

Пример 3 (профессиональная лингвистическая задача)

Установить специфичность описанного в труде Б. Н. Головина «Язык и статистика» [2] факта существования отрицательной корреляционной зависимости между количеством существительных и количеством местоимений в русских литературных текстах для а) художественных текстов; б) для публицистических текстов. Задача данного вида уже возникала в реальном профессионально-лингвистическом исследовании. Подтверждая результаты этого исследования, студенты фактически осуществляют профессиональную деятельность. Для успешности этой деятельности необходимо использовать знания, полученные при решении учебных и профессионально-ориентированных задач. По словам А. А. Вербицкого, студенты «остаются в позиции обучающихся и в то же время по целям, содержанию, формам, процессу и требованиям к получаемым результатам в позиции специалистов» [5]. Профессиональные лингвистические задачи обычно решаются студентами на лабораторных работах, а также при выполнении курсовых и дипломных работ.

Так как процесс решения профессиональных стохастических задач является длительным и многогранным, то для проведения учебных занятий (лабораторных работ) мы решили использовать так называемую «сквозную» задачу лингвистического анализа, т.е. задачу статистического анализа одной базы данных лингвистического эксперимента, в решении которой студент продвигается на протяжении целой серии учебных занятий. Такую базу данных студенты формируют самостоятельно на этапе подготовки к лабораторным работам. Для этого им предлагается, например, выполнить следующее *домашнее задание*:

«Выбрать по 20 фрагментов в 100 словоупотреблениях из двух произведений различных периодов творчества любимого писателя: 1 – ранний период, 2 – поздний период. Подсчитать количество глагольных форм в каждом фрагменте. Занести полученные данные в электронную таблицу SPSS (форма такой таблицы может быть предложена студентам в готовом виде)».

Характер продвижения в решении сквозной задачи при проведении серии лабораторных работ представлен таблицей 2.

Таблица 2

Лабораторные работы для студентов-филологов со «сквозной задачей»

| <i>Тема лабораторной работы</i> | <i>Профессиональная подзадача</i> |
|---|---|
| Первичная обработка лингвистической информации. Описательные статистики выборки | По лингвистической выборке построить статистический ряд, получить геометрическую интерпретацию полученного статистического распределения. Определить числовые характеристики выборки. Проверить результаты в пакете анализа данных SPSS |
| Проверка гипотезы о нормальности распределения генеральной лингвистической совокупности | Сформировать статистические гипотезы о виде распределения генеральной лингвистической совокупности и проверить их с помощью различных критериев согласия используя программу SPSS |
| Оценка параметров нормально распределённой лингвистической случайной величины по выборке. Определение достаточного объёма выборки в лингвистических исследованиях | Построить доверительные интервалы для параметров нормально распределённой лингвистической величины, найти ошибки измерения. Определить достаточный объём выборки по заданной возможной ошибке. Произвести проверку в SPSS |
| Проверка гипотезы о статистической значимости различия средних двух независимых лингвистических совокупностей | Сформулировать гипотезы о параметрах распределения лингвистической информации. Проверить гипотезы, используя параметрические критерии. Проверить результаты в программе SPSS |
| Корреляционный анализ. Проверка гипотезы о статистической зависимости двух лингвистических совокупностей | Оценить величину корреляционной зависимости лингвистических величин с помощью коэффициента корреляции Пирсона. Построить простую линейную регрессию по экспериментальным данным, используя программу SPSS |
| Однофакторный дисперсионный анализ. Исследование влияния фактора на лингвистическую СВ | Проверить влияние независимой СВ (номинативный признак) на зависимую лингвистическую СВ, измеренную в метрической шкале. Применить при проверке однофакторный дисперсионный анализ. Использовать программу SPSS |

Еще одной особенностью решения задачи профессиональной направленности обучения математике в процессуальном аспекте является обучение использованию программных продуктов специального назначения для обработки и анализа статистических данных. Для обработки лингвистической информации на компьютере мы выбрали пакет анализа данных SPSS Statistics. Данная программа проста в употреблении, универсальна, не требует от студентов знания математических формул, владения техникой вычислений, поэтому наиболее удобна для студентов-гуманитариев.

Для демонстрации специфики постановки профессиональных стохастических лингвистических задач на лабораторных работах рассмотрим задачу на проверку влияния периода творчества автора на частоту употребления глагольных форм. Эта задача решается методом однофакторного дисперсионного анализа. Так как фактор имеет только две градации, то вместо дисперсионного анализа можно использовать критерий t-Стьюдента для сравнения средних двух независимых выборок. В случае если выборки имеют различные объёмы и при проверке равенства дисперсий будет получен отрицательный результат, то может быть использован наиболее чувствительный аналог критерия Стьюдента для независимых выборок – критерий U-Манна – Уитни.

Пример 4. Лабораторная работа на тему «Исследование влияния периода творчества автора на частоту употребления глагольных форм»

Цель работы: научиться применять однофакторный дисперсионный анализ для проверки влияния независимой СВ (номинативный признак – фактор) на зависимую лингвистическую СВ, измеренную в метрической шкале (результативный признак).

Постановка задачи. В лингвистике часто требуется проверить влияние на нормально распределённую лингвистическую случайную величину определённого фактора, имеющего несколько уровней (градаций), например, исследовать влияние на распределение частоты употребления в художественных текстах, относящихся к различным периодам творчества автора одних и тех же словоформ.

Используя программу SPSS и составленную при выполнении подготовительного домашнего задания базу данных об употреблении глагольных форм в текстах любимого писателя в разные периоды его творчества, проверить на уровне значимости $p=0,05$ гипотезу о зависимости числа употребления (средней частоты) глагольных форм от периода написания произведения. Для решения задачи используйте результат лабораторной работы № 2 (доказано на уровне значимости $p=0,05$, что случайная величина Y – «число употребления глагольных форм», имеет нормальное распределение). Предварительно докажете возможность использования для решения задачи дисперсионного однофакторного анализа (проверьте гипотезу об однородности дисперсий).

План решения задачи вручную

1) Сформировать нулевую H_0 и альтернативную H_1 гипотезы однофакторного дисперсионного анализа.

2) Найти общее среднее \bar{x} и групповые средние значения: $\bar{x}_1; \bar{x}_2$.

3) Вычислить суммы квадратов отклонений значений выборки от среднего:

$$Q_{\text{общее}} = \sum_{i=1}^n (x_i - \bar{x})^2; \quad Q_{\text{межгр.}} = \sum_{j=1}^k n \cdot (\bar{x}_j - \bar{x})^2; \quad Q_{\text{вн.гр.}} = Q_{\text{общее}} - Q_{\text{межгр.}}$$

4) Определить число степеней свободы: для общей суммы квадратов $df_{\text{общ}} = n - 1$; для межгрупповой – $df_{\text{межгр.}} = k - 1$; для внутригрупповой – $df_{\text{вн.гр.}} = n - k$, где n – число всех значений выборки (для нашей задачи $n=15$), k – число групп ($m=2$).

5) Найти средние квадраты отклонений: $\bar{Q}_{\text{межгр.}} = \frac{Q_{\text{межгр.}}}{df_{\text{межгр.}}}$; $\bar{Q}_{\text{вн.гр.}} = \frac{Q_{\text{вн.гр.}}}{df_{\text{вн.гр.}}}$.

6) Вычислить эмпирическое F -отношение: $F_9 = \frac{\bar{Q}_{\text{межгр.}}}{\bar{Q}_{\text{вн.гр.}}}$.

7) По таблице критических значений F -распределения для направленных альтернатив (распределение Фишера – Снедекора) найти критическое значение $F_{кр}$ с найденным числом степеней свободы числителя $df_{\text{межгр.}}$ и числом степеней свободы знаменателя $df_{\text{вн.гр.}}$.

8) Сделать вывод относительно принятия или опровержения гипотез H_0 и H_1 : если $F_9 \leq F_{кр}$, то гипотеза H_0 не отвергается, т. е. влияние фактора на результативный признак не доказано, если $F_9 > F_{кр}$, то гипотеза H_0 отвергается с вероятностью ошибки p (уровень значимости указан в таблице), т. е. фактор существенно влияет на результативный признак Y .

Для решения задачи в программе SPSS необходимо:

– открыть программу SPSS, в главном окне перейти на вкладку «переменные», ввести следующие переменные: N (номер), числовая переменная, шкала порядковая;

Y (число глаголов), числовая переменная, шкала количественная; X (период творчества), числовая переменная, шкала номинальная. Установить метки значений переменной X: 1 – «ранний период», 2 – «поздний период», перейти в окно «данные» и ввести исходные значения переменных;

– выбрать «Анализ»→«Сравнение средних»→«Однофакторный дисперсионный анализ»;

– перенести в окно «зависимые переменные» переменную Y, а в окно «фактор» переменную X (период творчества);

– открыть вкладку «параметры» и отметить «описательные», «проверка однородности дисперсии» и «график средних» → «продолжить» →ОК.

Получаем: 1) описательные статистики; 2) критерий однородности дисперсии по статистике Ливиня (если уровень значимости ($Z_{нч}$) больше 0,05, то отличие дисперсий статистически не значимо; 3) дисперсионный анализ (если уровень значимости не превышает 0,05, то отличия средних значений в группах статистически значимо); 4) график зависимости среднего значения количества глаголов от периода творчества.

В результате решения задачи студенты получают возможность сделать следующий качественный вывод: частота употребления глаголов меняется во время творческой деятельности автора, что говорит о изменении динамичности, действенности повествования автора.

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.

Практическая апробация представленной методики формирования у студентов-филологов опыта использования стохастических методов в лингвистических исследованиях показывает, что такой подход формирует положительную мотивацию студентов к изучению математики, снижает их тревожность, связанную с успешностью овладения содержанием курса, снимает формализм результатов обучения студентов-филологов математике в вузе.

Список литературы

1. Вербицкий А. А. Компетентностный подход и теория контекстного обучения. – М.: ИЦ ПКПС, 2004. – С. 91.
2. Головин Б. Н. Язык и статистика. – М.: Просвещение, 1971. – С. 14, 161.
3. Наследов А. Д. Математические методы психологического исследования. Анализ и интерпретация данных: Учебное пособие. – СПб.: Речь, 2004. – С. 63.
4. Пиотровский Р. Г., Бектаев К. Б., Пиотровская А. А. Математическая лингвистика: Учебное пособие для пед. ин-тов. – М.: Высш. школа, 1997. – С.125.
5. Худякова Г. И. Системообразующая роль принципа профессиональной направленности в обучении математике // Ярославский вестник. – 2009. – № 4. – С.117–118.

Рецензенты:

Шабанова Мария Валерьевна, доктор педагогических наук, профессор, зав. кафедрой методики преподавания математики Северного (Арктического) федерального университета имени М. В. Ломоносова, г. Архангельск.

Симашко Татьяна Васильевна, доктор филологических наук, профессор, зав. кафедрой языкознания Гуманитарного института Северодвинского филиала Северного (Арктического) федерального университета имени М. В. Ломоносова, г. Архангельск.