

Учреждение образования
«Белорусский государственный университет
культуры и искусств»

УТВЕРЖДАЮ

Ректор БГУКИ

_____ Ю. П. Бондарь

« ____ » _____ 2015 г.

Регистрационный № УД-____/баз.

**МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ
БИБЛИОТЕЧНО-ИНФОРМАЦИОННОЙ
ДЕЯТЕЛЬНОСТИ**

*Учебная программа учреждения высшего образования
по учебной дисциплине для специальности
1-23 01 11 Библиотечно-информационная деятельность
(по направлениям)*

Минск
БГУКИ
2015

СОСТАВИТЕЛЬ

В. В. Нешитой, профессор кафедры информационных ресурсов учреждения образования «Белорусский государственный университет культуры и искусств», доктор технических наук, профессор

РЕЦЕНЗЕНТЫ:

П. В. Гляков, заведующий кафедрой информационных технологий в культуре учреждения образования «Белорусский государственный университет культуры и искусств», кандидат физико-математических наук, доцент;

А. Г. Буравкин, ведущий научный сотрудник Объединенного института проблем информатики Национальной академии наук Беларуси, кандидат технических наук, доцент

РЕКОМЕНДОВАНО К УТВЕРЖДЕНИЮ:

кафедрой информационных ресурсов учреждения образования «Белорусский государственный университет культуры и искусств» (протокол № 9 от 25.04.2014 г.);

президиумом научно-методического совета учреждения образования «Белорусский государственный университет культуры и искусств» (протокол № 5 от 18.06.2014 г.)

Ответственный за редакцию: В. Б. Кудласевич

Ответственный за выпуск: В. В. Нешитой

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебная программа «Математико-статистические методы библиотечно-информационной деятельности» разработана для студентов факультета информационно-документных коммуникаций БГУКИ в соответствии с требованиями образовательного стандарта по специальности 1-23 01 11 Библиотечно-информационная деятельность (по направлениям).

Актуальность изучения дисциплины обусловлена необходимостью использования статистических методов анализа в библиотечно-информационной деятельности для решения различных практических задач, в том числе: выявление статистической структуры библиотечного фонда; выработка критериев степени использования фонда; определение оптимального объема фонда и его информационной полноты комплектования.

Изучение дисциплины «Математико-статистические методы библиотечно-информационной деятельности» требует от студентов знаний, полученных ими при изучении таких дисциплин, как «Основы информационных технологий», «Автоматизация обработки текстовой информации» и некоторых других. В то же время знания и умения, полученные при изучении дисциплины, будут востребованы при изучении других дисциплин, например, «Библиотечные фонды», а также при выполнении курсовых и дипломных работ.

Цель изучения дисциплины – сформировать у студентов знания, умения и навыки в области статистического моделирования библиотечного фонда на базе четырехпараметрических обобщенных распределений, вычисления универсальных законов старения и рассеяния публикаций.

Задачи:

- изучение элементов теории обобщенных распределений;
- приобретение практических навыков по вычислению на базе статистических данных теоретических непрерывных законов распределения, в том числе ранговых, и оценок параметров;
- приобретение практических навыков по вычислению дискретных распределений.

В результате изучения дисциплины «Математико-статистические методы библиотечно-информационной деятельности» студенты должны *знать*:

- методы вычисления непрерывных и дискретных законов распределения по статистическим данным;
- известные и новые методы оценивания параметров;
- математико-статистические модели, которые могут быть использованы в библиотечно-информационной деятельности, в том числе универсальные законы старения и рассеяния публикаций;

уметь:

- использовать математико-статистические методы для построения моделей библиотечных процессов;
- вычислять законы распределения по статистическим данным;
- вычислять оценки параметров универсальных законов рассеяния и старения публикаций;
- извлекать информацию из библиотечной статистики;

владеть:

- методикой вычисления законов распределения, в том числе ранговых, и оценок параметров по статистическим данным;
- методикой оценивания оптимального объема библиотечного фонда по теоретическому ранговому распределению;
- методикой извлечения информации из библиотечной статистики.

Изложение материала по дисциплине «Математико-статистические методы библиотечно-информационной деятельности» ориентировано на современные методы статистического анализа данных, основанные на теории обобщенных распределений, которая включает три системы непрерывных четырехпараметрических распределений, систему дискретных распределений, методы вычисления теоретического закона распределения, в том числе рангового, и оценок параметров по статистическому распределению. В случае непрерывных четырехпараметрических распределений для вычисления теоретического закона распределения и оценок его параметров используются два новых метода: универсальный метод моментов и общий устойчивый метод. В некоторых частных случаях могут быть использованы классические методы оценивания параметров: метод наименьших квадратов Гаусса, метод моментов Пирсона, метод наибольшего правдоподобия Фишера. В курсе особое внимание уделяется ранговым распределениям. Они находят широкое применение в библиотековедении, ин-

форматике, математической лингвистике, социологии и других областях знания.

При проведении практических занятий рассматриваются примеры вычисления непрерывных законов распределения и оценок параметров по универсальному методу моментов, общему устойчивому методу, а также классическому методу моментов и методу наименьших квадратов.

Для вычисления дискретных распределений используются два метода: метод моментов и упрощенный метод, основанный на кривой роста новых событий.

В соответствии с учебным планом на изучение дисциплины «Математико-статистические методы библиотечно-информационной деятельности» всего предусмотрено 52 часа, из которых 34 часа – аудиторные занятия (22 часа – лекции и 12 часов – практические).

Рекомендуемая форма контроля знаний студентов – зачет.

ПРИМЕРНЫЙ ТЕМАТИЧЕСКИЙ ПЛАН

Темы	Количество аудиторных часов		
	всего	лекции	практ.
<i>Введение</i>	2	2	
<i>Тема 1.</i> Функции одной переменной. Производная функции	4	2	2
<i>Тема 2.</i> Некоторые понятия теории вероятностей и математической статистики	2	2	
<i>Тема 3.</i> Построение кривых роста и непрерывных распределений новых событий	2	2	
<i>Тема 4.</i> Методы построения обобщенных непрерывных распределений	2	2	
<i>Тема 5.</i> Классические методы оценивания параметров непрерывных распределений	4	2	2
<i>Тема 6.</i> Универсальный метод моментов вычисления закона распределения и оценок параметров	4	2	2
<i>Тема 7.</i> Общий устойчивый метод вычисления закона распределения и оценок параметров	4	2	2
<i>Тема 8.</i> Ранговые распределения в библиотечно-информационной деятельности	4	2	2
<i>Тема 9.</i> Построение системы дискретных распределений	2	2	
<i>Тема 10.</i> Оценивание параметров дискретных распределений. Расчет и прогнозирование частотного спектра	4	2	2
Итого...	34	22	12

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Введение

Предмет и содержание дисциплины. Цели, задачи и особенности дисциплины. Требования к базовым знаниям и навыкам студентов. Анализ литературных источников по теме курса. Связь с другими дисциплинами. Основные виды учебных занятий, обеспеченность учебно-методической литературой, формы контроля знаний студентов.

Тема 1. Функции одной переменной. Производная функции

Функция. Способы задания функций. Основные элементарные функции и их графики. Гамма-функция. Понятие эмпирической функции.

Задачи, приводящие к понятию производной. Определение и смысл производной. Касательная. Вычисление производных. Производные элементарных функций. Правила дифференцирования.

Тема 2. Некоторые понятия теории вероятностей и математической статистики

Предмет теории вероятностей. Случайное событие. Испытание. Относительная частота и вероятность. Основные теоремы теории вероятностей. Непрерывные и дискретные случайные величины и их законы распределения. Числовые характеристики случайных величин.

Предмет математической статистики. Генеральная и выборочная совокупности. Простой статистический ряд. Вариационный ряд. Интервальный ряд распределения. Гистограмма, полигон, кривая распределения. Плотность распределения. Функция распределения.

Тема 3. Построение кривых роста и непрерывных распределений новых событий

Понятия математического ожидания случайной функции, нового события и кривой роста новых событий. Связь кривой роста с законами распределения вероятностей разных и новых событий. Установление статистической структуры выборки по кривой роста новых событий. Восстановление кривой роста новых событий по статистической структуре выборки. Форму-

лы В. Калинина. Построение системы кривых роста и непрерывных распределений новых событий. Классификация распределений и кривых роста.

Тема 4. Методы построения обобщенных непрерывных распределений

Построение непрерывных распределений по кривым роста новых событий. Метод К. Пирсона. Метод обобщения. Распределение функций случайных аргументов. Классификация распределений. Форма кривых распределения. Характерные точки кривых распределения.

Тема 5. Классические методы оценивания параметров непрерывных распределений

Метод наименьших квадратов. Метод наибольшего правдоподобия. Классический метод моментов. Критерии для установления типа выравнивающей кривой по методу моментов. Использование номограммы.

Тема 6. Универсальный метод моментов вычисления закона распределения и оценок параметров

Критерии для вычисления типа выравнивающей кривой. Номограмма (типы кривых в координатах B_1, B_2). Вычисление оценок параметров. Расширение трех систем непрерывных распределений.

Тема 7. Общий устойчивый метод вычисления закона распределения и оценок параметров

Понятие устойчивости. Критерии для вычисления типа выравнивающей кривой по устойчивому методу. Номограмма (типы кривых в координатах B, H). Вычисление оценок параметров.

Тема 8. Ранговые распределения в библиотечно-информационной деятельности

Построение эмпирической кривой рангового распределения. Форма представления ранговых распределений. Критерий однородности ранговых распределений. Выделение неоднородной части. Законы распределения документальных информационных потоков (ранговые распределения журналов). Уни-

версальные законы старения и рассеяния публикаций. Оптимальный объем библиотечного фонда.

Тема 9. Построение системы дискретных распределений

Построение системы дискретных распределений по системе непрерывных распределений. Построение системы дискретных распределений по кривым роста новых событий на основе формулы В.Калинина. Классификация дискретных распределений.

Тема 10. Оценивание параметров дискретных распределений.

Расчет и прогнозирование частотного спектра

Порядок установления типа выравнивающего распределения и нахождения оценок параметров. Критерий степени неравномерности появления событий. Ранжирование слов по степени семантической нагрузки. Прогнозирование кривых роста новых лексических единиц и частотного спектра. Расчет достоверной части частотного словаря на выборке заданного объема.

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Литература

Основная

1. *Вентцель, Е. С.* Теория вероятностей / Е. С. Вентцель. – М. : Физматгиз, 1969. – 576 с.
2. *Герасимович, А. И.* Математическая статистика / А. И. Герасимович, Я. И. Матвеева. – Минск : Выш. шк., 1978. – 200 с.
3. *Гмурман, В. Е.* Теория вероятностей и математическая статистика: учеб. пособие для втузов / В. Е. Гмурман. – 5-е изд., перераб. и доп. – М. : Высш. шк., 1977. – 479 с.
4. *Калинин, В. М.* Некоторые статистические закономерности математической лингвистики / В. М. Калинин // Проблемы кибернетики. – М. : Наука, 1964. – Вып. 11. – С. 246–255.
5. *Нешиной, В. В.* Методы статистического анализа на базе обобщенных распределений : учеб.-метод. пособие / В. В. Нешиной. – Минск : Веды, 2001. – 168 с.
6. *Нешиной, В. В.* Элементы теории обобщенных распределений : монография / В. В. Нешиной. – Минск : РИВШ, 2009. – 204 с.
7. *Нешиной, В. В.* Математико-статистические методы анализа в библиотечно-информационной деятельности : учеб.-метод. пособие / В. В. Нешиной. – Минск : БГУКИ, 2009. – 203 с.
8. *Нешиной, В. В.* Методы статанализа в библиотечной деятельности: вычисление непрерывных распределений : учеб.-метод. пособие / В. В. Нешиной. – Минск : Беларус. гос. ун-т культуры и искусств, 2010. – 61 с.
9. *Нешиной, В. В.* Методы статанализа в библиотечно-информационной деятельности: вычисление дискретных распределений и кривых роста / В. В. Нешиной. – Минск : РИВШ, 2012. – 134 с.

Дополнительная

1. *Нешиной, В. В.* Форма представления ранговых распределений / В. В. Нешиной // Ученые записки Тартуского гос. ун-та, 1987. – Вып. 774. – С. 123–134.
2. *Нешиной, В. В.* Математические модели роста словаря и информационных потоков / В. В. Нешиной // Ученые записки Тартуского гос. ун-та, 1989. – Вып. 872. – С. 83–102.

3. *Нешиной, В. В.* Универсальные законы рассеяния и старения публикаций / В. В. Нешиной // Весн. Беларус. дзярж. ун-та культуры і мастацтваў. – 2007. – № 2. – С. 128–132.

4. *Нешиной, В. В.* Статистическое моделирование библиотечного фонда / В. В. Нешиной // Науч. и техн. б-ки. – 2009. – № 12. – С. 36–46.

5. *Нешиной, В. В.* Законы Ципфа, Бредфорда и универсальные модели / В. В. Нешиной // НТИ. – Сер. 2. – 2010. – № 1. – С. 26–33.

6. *Нешиной, В. В.* Методы вычисления границ ядра и зон рассеяния публикаций / В. В. Нешиной // НТИ. – Сер. 2. – 2013. – № 11. – С. 1–11.

РЕПОЗИТОРИЙ БГУКИ

Тематика и содержание практических занятий

Тема 1. Функции одной переменной. Производная функции

Рассмотреть основные элементарные функции. Повторить определение и смысл производной, таблицу производных и правила дифференцирования.

Вычислить производные функций, предложенных преподавателем.

Тема 5. Классические методы оценивания параметров непрерывных распределений

Для выравнивания статистического распределения использовать обобщенные распределения группы А.

Оценить параметры выравнивающего распределения группы А по методу наименьших квадратов.

Установить тип выравнивающего распределения группы Б по показателям U , L .

Найти оценки параметров по классическому методу моментов.

Тема 6. Универсальный метод моментов вычисления закона распределения и оценок параметров

По статистическому распределению вычислить среднее значение случайной величины, центральные моменты (2–4-го) порядков.

Вычислить показатели асимметрии B_1 и островершинности B_2 . Вычислить критерий B . Нешиного L_1 .

С помощью номограммы установить тип выравнивающего распределения.

Вычислить оценки параметров по универсальному методу моментов.

Вычислить значения плотности вероятностей и сравнить с эмпирическими плотностями.

Тема 7. Общий устойчивый метод вычисления закона распределения и оценок параметров

По статистическому распределению вычислить критерии B , H .

С помощью номограммы установить тип выравнивающего распределения.

Вычислить оценки параметров по общему устойчивому методу. Вычислить значения плотности вероятностей.

*Тема 8. Ранговые распределения
в библиотечно-информационной деятельности*

Рассмотреть несколько примеров по вычислению ранговых выравнивающих распределений. Объяснить, почему найденное распределение может служить универсальным законом рассеяния публикаций.

Вычислить координаты трех характерных точек по формулам, которые следуют из обобщенных четырехпараметрических распределений. Сравнить объемы ядра и оптимального фонда.

*Тема 10. Оценивание параметров дискретных
распределений. Расчет и прогнозирование
частотного спектра*

По статистическим распределениям слов в подвыборках одинакового объема установить тип выравнивающих дискретных распределений для разных слов, найти оценки параметров и ранжировать слова по степени семантической нагрузки.

Методические указания по организации и выполнению самостоятельной работы студентов

Студент должен всегда иметь в виду, что любое образование – это прежде всего самообразование.

Для успешного освоения программы по дисциплине «Математико-статистические методы библиотечно-информационной деятельности» студенту необходимо посещать все лекции и практические занятия, на которых принимать активное участие в решении задач, предложенных преподавателем.

После прохождения каждой темы студент должен освоить теоретический материал по конспекту и литературным источникам, приведенным в списке литературы. Можно также использовать их электронные варианты, размещенные в репозитории библиотеки БГУКИ.

При подготовке к зачету следует повторить курс лекций и рассмотреть все решенные на практических занятиях примеры и задачи. Кроме того необходимо самостоятельно решить примеры, приведенные в учебно-методическом комплексе, используя вычислительную технику. При возникновении вопросов обращаться за консультацией к преподавателю.

Выполнение этих указаний гарантирует успешную сдачу зачета.

Перечень используемых средств диагностики результатов учебной деятельности

Для оценки результатов учебной деятельности используются средства диагностики: контроль посещаемости занятий; участие в практических занятиях; собеседование; опрос; тестирование; проверка правильности самостоятельно решенных задач; зачет.

Учебное издание

**МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ
БИБЛИОТЕЧНО-ИНФОРМАЦИОННОЙ ДЕЯТЕЛЬНОСТИ**

*Учебная программа учреждения высшего образования
по учебной дисциплине для специальности
1-23 01 11 Библиотечно-информационная деятельность
(по направлениям)*

Корректор В. Б. Кудласевич
Технический редактор Л. Н. Мельник

Подписано в печать 2015. Формат 60x84 ¹/₁₆.
Бумага писчая № 2. Ризография.
Усл. печ. л. 0,87. Уч.-изд. л. 0,44. Тираж экз. Заказ .

Издатель и полиграфическое исполнение:
УО «Белорусский государственный университет культуры и искусств».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1/177 от 12.02.2014.
ЛП № 02330/456 от 23.01.2014.
Ул. Рабкоровская, 17, 220007, г. Минск.