

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Северный (Арктический) федеральный университет
имени М.В. Ломоносова»

И.В. Кокорина

Основы математической обработки информации в филологии

*Комбинаторика, теория вероятностей
и математическая статистика*

Учебно-методическое пособие

Архангельск



ИД САФУ
2014

УДК 519.2:81
ББК 22.171+22.172+81.1
К53

*Рекомендовано к изданию редакционно-издательским советом
Северного (Арктического) федерального университета
имени М.В. Ломоносова*

Рецензенты:

доктор педагогических наук, профессор **М.В. Шабанова**,
доктор филологических наук, профессор **Т.В. Симашко**,
кандидат технических наук, доцент **Н.И. Черников**

Кокорина, И.В.

К53 Основы математической обработки информации в филологии:
комбинаторика, теория вероятностей и математическая статистика:
учеб.-метод. пособие / И.В. Кокорина; Сев. (Арктич.) федер. ун-т
им. М.В. Ломоносова. – Архангельск: ИД САФУ, 2014. – 115 с.
ISBN 978-5-261-00928-3

Пособие содержит теоретические основы комбинаторики, теории вероятностей и математической статистики, описываются методы этих наук применительно к филологии и языкознанию. Даны практические задания, лабораторные работы, рекомендуемая литература, программа курса.

Издание адресовано студентам, обучающимся по направлению подготовки бакалавриата 050100.62 Педагогическое образование (профиль «Филология»).

УДК 519.2:81
ББК 22.171+22.172+81.1

ISBN 978-5-261-00928-3

© Кокорина И.В., 2014
© Северный (Арктический)
федеральный университет
им. М.В. Ломоносова, 2014

Оглавление

Часть 1

Основные понятия комбинаторики, теории вероятностей и математической статистики

§ 1. Математические методы в филологии.....	6
§ 2. Комбинаторика	8
2.1. Правила комбинаторики.....	8
2.2. Основные понятия комбинаторики	9
§ 3. Теория вероятностей и ее основные понятия.....	12
3.1. Начальные понятия теории вероятностей.....	13
3.2. Определения вероятности событий.....	15
§ 4. Основные теоремы теории вероятностей случайных событий	18
4.1. Действия над событиями.....	19
4.2. Вероятность суммы событий	20
4.3. Вероятность произведения событий	21
4.4. Формула полной вероятности. Формула Байеса.....	23
4.5. Теорема Бернулли	25
4.6. Предельные теоремы в схеме Бернулли.....	26
§ 5. Случайная величина	28
5.1. Начальные понятия	28
5.2. Функция распределения $F(x)$	30
5.3. Функция плотности вероятности $f(x)$ непрерывной случай- ной величины	32
5.4. Числовые характеристики случайных величин.....	32
5.5. Виды законов распределения случайных величин.....	34
5.6. Вероятность попадания непрерывной случайной величины в заданный промежуток.....	39
§ 6. Двумерная дискретная случайная величина.....	40
6.1. Начальные понятия.....	40
6.2. Операции над независимыми случайными величинами.....	40
6.3. Числовые характеристики системы двух случайных величин	41
§ 7. Предельные теоремы теории вероятностей	43
7.1. Закон больших чисел	43
7.2. Центральная предельная теорема	46
§ 8. Основы математической статистики.....	47
8.1. Предмет математической статистики. Генеральная и выбо- рочная совокупности.....	48
8.2. Статистическое распределение выборки и его графическое изображение	50
8.3. Числовые характеристики статистического распределения.....	54
8.4. Числовые характеристики вариационного ряда	55

§ 9. Элементы теории статистических оценок и проверки гипотез	57
9.1. Оценка параметров генеральной совокупности по выборке	57
9.2. Точечные оценки математического ожидания, дисперсии и вероятности	58
9.3. Интервальное оценивание параметров	59
9.4. Доверительные интервалы для параметров нормального распределения	59
9.5. Интервальная оценка параметров генеральной лингвистической совокупности	60
9.6. Число степеней свободы	63
9.7. Определение минимально достаточного объема выборки в лингвистических исследованиях	63
§ 10. Проверка статистических гипотез. Метод гипотез	64
10.1. Статистические гипотезы	64
10.2. Статистический критерий	65
10.3. Принцип проверки статистических гипотез	65
10.4. Ошибки при проверке гипотез	66
10.5. Проверка лингвистических гипотез с помощью параметрических критериев	66
10.6. Проверка лингвистических гипотез с помощью непараметрических критериев	68
10.7. Однофакторный дисперсионный анализ	71

Часть 2 Практикум

Практические занятия

<i>Практическое занятие № 1. Тема 1. Элементы комбинаторики. Тема 2. Начальные понятия теории вероятностей</i>	73
<i>Практическое занятие № 2. Основные теоремы теории вероятностей</i>	76
<i>Практическое занятие № 3. Случайные величины</i>	78
<i>Практическое занятие № 4. Элементы математической статистики</i>	80

Задания для самостоятельной работы

<i>Задание № 1. Тема 1. Элементы комбинаторики. Тема 2. Начальные понятия теории вероятностей</i>	82
<i>Задание № 2. Основные теоремы теории вероятностей</i>	83
<i>Задание № 3. Случайные величины</i>	84
<i>Задание № 4. Элементы математической статистики</i>	85

Лабораторные работы

<i>Лабораторная работа № 1. Первичная обработка лингвистической информации</i>	86
<i>Лабораторная работа № 2. Проверка гипотезы о нормальности распределения глагольных форм в литературных текстах.....</i>	88

<i>Лабораторная работа № 3. Оценка параметров нормально распределенной лингвистической случайной величины по выборке. Определение необходимого объема выборки в лингвистических исследованиях.....</i>	91
<i>Лабораторная работа № 4. Проверка гипотезы о статистической значимости различия средних частот употребления глаголов у двух авторов</i>	92
<i>Лабораторная работа № 5. Парный корреляционный анализ. Построение модели линейной регрессии лингвистической информации.....</i>	95
<i>Лабораторная работа № 6. Однофакторный дисперсионный анализ. Влияние стиля речи на частоту употребления глагольных форм</i>	97

Часть 3

Учебно-методические материалы

Рабочая программа модуля «Комбинаторика, теория вероятностей и математическая статистика» (для филологов).....	99
Примерные вопросы к зачету	102
Рекомендуемая литература	103
Приложение. Таблицы математической статистики	
1. Значения интегральной функции Лапласа $\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$	105
2. Критические значения χ^2 (распределение Пирсона).....	107
3. Критические значения критерия t (распределение Стьюдента)	109
4. Критические значения F -Фишера (для проверки направленных альтернатив).....	111
5. Критические значения F -Фишера (для проверки ненаправленных альтернатив).....	112
Библиографический список	113