

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
"Казанский (Приволжский) федеральный университет"  
Институт вычислительной математики и информационных технологий



УТВЕРЖДАЮ  
Проректор по образовательной деятельности КФУ  
Проф. Д.А. Гафурский

» \_\_\_\_\_ 20\_\_ г.

*подписано электронно-цифровой подписью*

## Программа дисциплины

Математическая лингвистика Б1.В.ДВ.02.01

Направление подготовки: 09.04.02 - Информационные системы и технологии

Профиль подготовки: Компьютерная лингвистика

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2019

**Автор(ы):** Галимянов А.Ф.

**Рецензент(ы):** Гафаров Ф.М.

### **СОГЛАСОВАНО:**

Заведующий(ая) кафедрой: Галимянов А. Ф.

Протокол заседания кафедры No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 20\_\_ г.

Учебно-методическая комиссия Института вычислительной математики и информационных технологий:

Протокол заседания УМК No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 20\_\_ г.

## Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы
2. Место дисциплины в структуре основной профессиональной образовательной программы высшего образования
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
  - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
  - 4.2. Содержание дисциплины
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
  - 6.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы и форм контроля их освоения
  - 6.2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания
  - 6.3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы
  - 6.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций
7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)
  - 7.1. Основная литература
  - 7.2. Дополнительная литература
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

Программу дисциплины разработал(а)(и) доцент, к.н. (доцент) Галимянов А.Ф. (Кафедра информационных систем, отделение фундаментальной информатики и информационных технологий), Anis.Galimjanoff@kpfu.ru

### 1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Выпускник, освоивший дисциплину, должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-1	Создание педагогических условий для развития группы обучающихся по программам ВО, организационно-методическое, преподавание для программ бакалавриата, ориентированным на соответствующий уровень квалификации
ПК-4	Выполнение работ и управление работами по созданию, модификации и сопровождению ИС
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий
УК-2	Способен управлять проектом на всех этапах его жизненного цикла
УК-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

Выпускник, освоивший дисциплину:

Должен знать:

1. Основы математического анализа и алгебры, применяющийся в задачах лингвистики
2. Основы алгебры логики
3. Основы теории вероятностей и математической статистики
4. Способы применения математики для исследования лингвистических явлений

Должен уметь:

1. Анализировать лингвистические явления с точки зрения математики
2. Анализировать вероятностные распределения
3. Строить таблицы истинности

Должен владеть:

1. Навыками решения задач по теории вероятностей и математической статистике
2. Навыками решения задач по математической логике
3. Навыками решения задач по математическому анализу и линейной алгебре

Должен демонстрировать способность и готовность:

применять математические методы для решения лингвистических задач

### 2. Место дисциплины в структуре основной профессиональной образовательной программы высшего образования

Данная учебная дисциплина включена в раздел "Б1.В.ДВ.02.01 Дисциплины (модули)" основной профессиональной образовательной программы 09.04.02 "Информационные системы и технологии (Компьютерная лингвистика)" и относится к дисциплинам по выбору.

Осваивается на 1 курсе в 2 семестре.

### 3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 6 зачетных(ые) единиц(ы) на 216 часа(ов).

Контактная работа - 36 часа(ов), в том числе лекции - 0 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 36 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 126 часа(ов).

Контроль (зачёт / экзамен) - 54 часа(ов).

Форма промежуточного контроля дисциплины: экзамен во 2 семестре.

#### 4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

##### 4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Лингвистические явления и их исследование с точки зрения математики	2	0	0	12	42
2.	Тема 2. Динамика лингвистических процессов и ее описание с помощью приемов дифференциального исчисления	2	0	0	12	42
3.	Тема 3. Вероятностно-информационные оценки нормы языка и статистическое построение текста	2	0	0	12	42
	Итого		0	0	36	126

##### 4.2 Содержание дисциплины

###### Тема 1. Лингвистические явления и их исследование с точки зрения математики

Множество лингвистических объектов. Операции над множествами. Применение диаграмм Эйлера-Венна при решении практических задач. Бинарные отношения. Прямое произведение множеств. Свойства отношений. Лингвистическое явление как математическая величина. Функции. Диахронический скачок и его моделирование с помощью элементарных функций.

###### Тема 2. Динамика лингвистических процессов и ее описание с помощью приемов дифференциального исчисления

Моделирование информационного построения речи. Глоттохронология, информационная схема текста и их моделирование с помощью аппарата бесконечно малых величин и пределов. Понятие бесконечно малой величины. Свойства бесконечно малых функций. Методы вычисления пределов. Глоттохронология. Использование дифференциала для приближенных вычислений в лингвистических задачах.

###### Тема 3. Вероятностно-информационные оценки нормы языка и статистическое построение текста

Лингвистические задачи, приводящие к понятию интеграла. Интегралы и их применение к лингвистическим задачам. Вероятностно-информационные оценки нормы языка и статистическое построение текста. Информационные измерения в тексте. Случайная лингвистическая величина, ее характеристики и функция распределения. Случайные процессы. Методы организации статистического наблюдения над текстом.

#### 5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства образования и науки Российской Федерации от 5 апреля 2017 года №301).

Письмо Министерства образования Российской Федерации №14-55-996ин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений".

Положение от 29 декабря 2018 г. № 0.1.1.67-08/328 "О порядке проведения текущего контроля успеваемости и промежуточной аттестации обучающихся федерального государственного автономного образовательного учреждения высшего образования "Казанский (Приволжский) федеральный университет".

Положение № 0.1.1.67-06/24/15 от 14 декабря 2015 г. "О формировании фонда оценочных средств для проведения текущей, промежуточной и итоговой аттестации обучающихся федерального государственного автономного образовательного учреждения высшего образования "Казанский (Приволжский) федеральный университет"".

Положение № 0.1.1.56-06/54/11 от 26 октября 2011 г. "Об электронных образовательных ресурсах федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"".

Регламент № 0.1.1.67-06/66/16 от 30 марта 2016 г. "Разработки, регистрации, подготовки к использованию в учебном процессе и удаления электронных образовательных ресурсов в системе электронного обучения федерального государственного автономного образовательного учреждения высшего образования "Казанский (Приволжский) федеральный университет"".

Регламент № 0.1.1.67-06/11/16 от 25 января 2016 г. "О балльно-рейтинговой системе оценки знаний обучающихся в федеральном государственном автономном образовательном учреждении высшего образования "Казанский (Приволжский) федеральный университет"".

Регламент № 0.1.1.67-06/91/13 от 21 июня 2013 г. "О порядке разработки и выпуска учебных изданий в федеральном государственном автономном образовательном учреждении высшего профессионального образования "Казанский (Приволжский) федеральный университет"".

## 6. Фонд оценочных средств по дисциплине (модулю)

### 6.1 Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы и форм контроля их освоения

Этап	Форма контроля	Оцениваемые компетенции	Темы (разделы) дисциплины
<b>Семестр 2</b>			
	<b>Текущий контроль</b>		
1	Лабораторные работы	УК-6 , УК-2 , УК-1 , ПК-4 , ПК-1	1. Лингвистические явления и их исследование с точки зрения математики 2. Динамика лингвистических процессов и ее описание с помощью приемов дифференциального исчисления 3. Вероятностно-информационные оценки нормы языка и статистическое построение текста
2	Письменная работа	УК-6 , УК-2 , УК-1 , ПК-4 , ПК-1	1. Лингвистические явления и их исследование с точки зрения математики 2. Динамика лингвистических процессов и ее описание с помощью приемов дифференциального исчисления
3	Письменная работа	УК-6 , УК-2 , УК-1 , ПК-4 , ПК-1	3. Вероятностно-информационные оценки нормы языка и статистическое построение текста
	<b>Экзамен</b>	ПК-1, ПК-4, УК-1, УК-2, УК-6	

### 6.2 Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Форма контроля	Критерии оценивания				Этап
	Отлично	Хорошо	Удовл.	Неуд.	
<b>Семестр 2</b>					
<b>Текущий контроль</b>					
Лабораторные работы	Оборудование и методы использованы правильно. Проявлена превосходная теоретическая подготовка. Необходимые навыки и умения полностью освоены. Результат лабораторной работы полностью соответствует её целям.	Оборудование и методы использованы в основном правильно. Проявлена хорошая теоретическая подготовка. Необходимые навыки и умения в основном освоены. Результат лабораторной работы соответствует её целям.	Оборудование и методы частично использованы правильно. Проявлена удовлетворительная теоретическая подготовка. Необходимые навыки и умения частично освоены. Результат лабораторной работы частично соответствует её целям.	Оборудование и методы использованы неправильно. Проявлена неудовлетворительная теоретическая подготовка. Необходимые навыки и умения не освоены. Результат лабораторной работы не соответствует её целям.	1

Форма контроля	Критерии оценивания				Этап
	Отлично	Хорошо	Удовл.	Неуд.	
Письменная работа	Правильно выполнены все задания. Продемонстрирован высокий уровень владения материалом. Проявлены превосходные способности применять знания и умения к выполнению конкретных заданий.	Правильно выполнена большая часть заданий.	Задания выполнены более чем наполовину.	Задания выполнены менее чем наполовину.	2 3
		Присутствуют незначительные ошибки. Продемонстрирован хороший уровень владения материалом. Проявлены средние способности применять знания и умения к выполнению конкретных заданий.	Присутствуют серьезные ошибки. Продемонстрирован удовлетворительный уровень владения материалом. Проявлены низкие способности применять знания и умения к выполнению конкретных заданий.	Присутствуют серьезные ошибки. Продемонстрирован неудовлетворительный уровень владения материалом. Проявлены недостаточные способности применять знания и умения к выполнению конкретных заданий.	
<b>Экзамен</b>	Обучающийся обнаружил всестороннее, систематическое и глубокое знание учебно-программного материала, умение свободно выполнять задания, предусмотренные программой, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой дисциплины, усвоил взаимосвязь основных понятий дисциплины в их значении для приобретаемой профессии, проявил творческие способности в понимании, изложении и использовании учебно-программного материала.	Обучающийся обнаружил полное знание учебно-программного материала, успешно выполнил предусмотренные программой задания, усвоил основную литературу, рекомендованную программой дисциплины, показал систематический характер знаний по дисциплине и способен к их самостоятельному пополнению и обновлению в ходе дальнейшей учебной работы и профессиональной деятельности.	Обучающийся обнаружил знание основного учебно-программного материала в объеме, необходимом для дальнейшей учебы и предстоящей работы по профессии, справился с выполнением заданий, предусмотренных программой, знаком с основной литературой, рекомендованной программой дисциплины, допустил погрешности в ответе на экзамене и при выполнении экзаменационных заданий, но обладает необходимыми знаниями для их устранения под руководством преподавателя.	Обучающийся обнаружил значительные пробелы в знаниях основного учебно-программного материала, допустил принципиальные ошибки в выполнении предусмотренных программой заданий и не способен продолжить обучение или приступить по окончании университета к профессиональной деятельности без дополнительных занятий по соответствующей дисциплине.	

**6.3 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы**

**Семестр 2**

**Текущий контроль**

**1. Лабораторные работы**

Темы 1, 2, 3

Примеры задач:

- 1.1. Найти параметр, по которому русские слова забота, открываться, вечерний, очи можно объединить в одно множество.
- 1.2. Найти параметр, по которому английские слова woman, mouse, phenomenon можно объединить в одно множество.
- 1.3. Найти параметр, по которому латынь, японский и литовский языки можно объединить в одно множество.
- 1.4. Найти три параметра, по которым французское слово liberte ?свобода? и русские слова кантата, поклажа, полоска можно объединить в одно множество.
- 1.5. Определить взаимное расположение множества изменяемых слов (А), множества существительных (В), множества числительных (С) и множества прилагательных (D) в русском языке.

- 1.6. Определить взаимное расположение множества слов мужского рода (A), множества слов в именительном падеже (B) и множества слов во множественном числе (C) в русском языке.
- 1.7. Определить взаимное расположение множества германских языков (A), множества романских языков (B), множества индоевропейских языков (C) и множества языков Европы (D).
- 1.8. Определить взаимное расположение множества алфавитных письменностей (A), множества консонантных письменностей (B), множества фонетических письменностей (C), множества арабского письма (D) и множества грузинского письма (E).
- 1.9. Определить дополнение повелительного наклонения до множества наклонений русского языка.
- 1.10. Определить дополнение множества именительного, родительного и винительного падежей до множества падежей немецкого языка.
- 1.11. Определить дополнение множества имперфекта и перфекта до множества прошедших времён старославянского языка.
- 1.12. Определить мощность множества букв слова successfully.
- 1.13. Определить максимально возможную мощность палиндрома, составленного из 9 букв.
- 1.14. Определить мощность дополнения множества слоговых знаков до множества графем в японском письме кана.
- 1.15. Определить мощность двузначных числительных французского языка, содержащих элемент quatre ?4?.
- 1.16. Придумать русский глагол, равномогущий по числу букв с его немецким переводом.
- 1.17. Английское слово mine является омонимом. Составить декартово произведение этого слова и его переводов на русский язык.
- 1.18. Придумать описание чтения английской буквы g через декартово произведение этой буквы и содержащих её слов, начинающихся на букву e.
- 1.19. Составить декартово произведение множества глагольных сказуемых (видеть, слышать) и множества прямых дополнений (друга, крик, рисунок), отбросив не подходящие по смыслу пары.
- 1.20. В упражнении даны два столбца, содержащих по 6 слов. Известно, что сопоставление слов из разных столбцов однозначно. Найти разность между мощностями декартового произведения столбцов и числа правильных вариантов.
- 1.21. Найти пересечение японских иероглифов A ?стол?, Ш ?мост?, Ш ?груша? как графем и попытаться определить значение их общего элемента (?ключа?).
- 1.22. Дано множество глаголов во времени Present Simple (A) и множество изменяемых слов (B) в английском языке.
  - 2.1. Привести примеры атомарных и составных термов ЛП ?письменность?.
  - 2.2. Привести примеры атомарных термов ЛП ?транспорт? для разных исторических эпох и регионов.
  - 2.3. Придумать ЛП, атомарными термами которой являются зелёный и жёлтый.
  - 2.4. Привести примеры атомарных и составных термов ЛП ?эпитет?.
  - 2.5. Придумать ЛП с областью определения (1, 2, 3).
  - 2.6. Придумать несколько ЛП со схожей или совпадающей областью определения.
  - 2.7. Пусть A - множество слов китайского языка, B - множество слов монгольского языка, C - множество слов вьетнамского языка, а x означает тон
- 2.8. Пусть x означает артикль, изменяющийся по роду, числу и падежу. Придумать такое языковое множество A, что  $MA(x) = 1$ .
- 2.9. Найти характеристические функции слов лъзя, лобзанье, эстакада на множестве лексики современного русского языка.
- 2.10. Найти характеристическую функцию фонемы ч на множестве согласных французского языка.
- 2.11. Найти характеристическую функцию японской каны на множестве силлабических письменностей.
- 2.12. Найти степень истинности нечётких высказываний ?Лёша, вероятно, успеет на матч?, ?Серёжа вряд ли успеет на матч?, ?Вадим, возможно, тоже успеет на матч? их чёткому аналогу.
- 2.13. Найти степень истинности нечётких высказываний ?Она, уж конечно, запомнила это правило?, ?Было бы странно, если б она запомнила это правило? их чёткому аналогу.
- 2.14. Найти степень истинности сокращённых вариантов кон., к., КР ЛП ?контрольная работа? во фразе ?Завтра - ...?.
- 2.15. Придумать индифферентное высказывание.
- 2.16. У слова мало значений. Найти функцию принадлежности числа значений слова. Проверить нормальность и унимодальность множества.
- 2.17. В восьмисловной японской фразе почти не встречается заимствованных слов. Найти функцию принадлежности числа заимствованных слов фразы. Построить дополнение множества. Найти моды полученных множеств.
- 2.18. Придумать нечёткое множество, носителем которого является интервал [6, 10].
- 2.19. Придумать нечёткое множество, точкой перехода которого является значение  $x = 1$ .
- 2.20. В десятистраничном тексте пассивные причастия встречаются редко, а причастные обороты - очень редко. Найти функции принадлежности числа страниц, на которых встречаются пассивные причастия и причастные обороты.

- 2.21. Среди шести информантов не менее половины являются билингвами. Некоторые информанты живут в горных сёлах. Найти функцию принадлежности числа информантов, являющихся билингвами или проживающих в горных сёлах.
- 2.22. Некоторые письменности можно освоить за несколько дней с помощью пособия, а некоторые - почти за неделю самостоятельно. Найти функцию принадлежности числа дней, необходимых для изучения письменности не с помощью пособия, а самостоятельно.
- 2.23. Дана фраза: ?Коля думает только о себе?. Расставить индексы референции.
- 2.24. Дана фраза: ?Я увидел Настю с Ксюшей, она была в хорошем настроении?. Расставить индексы референции.
- 2.25. Дана фраза: ?Сергей сослался на него в статье, опубликованной в его журнале?. Расставить индексы референции.
- 2.26. Дана фраза: ?Он сказал, что сосед его крепко спит?. Расставить индексы референции.

## 2. Письменная работа

Темы 1, 2

Примеры задач:

- 17.1. Дана фраза: По мнению лингвистов, баскский язык является одним из сложнейших языков Европы. Определить её качественные характеристики для нефилолога.
- 17.2. Дана фраза: Как написано в одной книге, в арабском языке корень состоит в основном из согласных. Определить её качественные характеристики для филолога.
- 17.3. Дана фраза: Говорят, что стихи Ахматовой было сложно переводить на другие языки. Определить её качественные характеристики для человека, далёкого от поэзии.
- 17.4. Дана фраза: И тогда я увидел ... как бы это лучше выразить... увидел странный... не знаю, что бы это значило... Определить её качественные характеристики.
- 17.5. Определить прагматическую информацию общения для театрала: Ожидаются выступления известных представителей МДТ и ДДТ.
- 17.6. Определить прагматическую информацию сообщения для филолога: Готовятся переиздания книг Фриш, Смирнова, Гумилёва и Солнцева.
- 17.7. Придумать фразу, не содержащую никакой прагматической информации.
- 17.8. Найти энтропию фразы: В сентябре (октябре, ноябре, декабре) на первую (вторую, третью, четвёртую) пару опоздали (пришли вовремя) пять (шесть, семь, восемь) студентов.
- 17.9. Найти энтропию фразы: Завтра он улетает с женой (невестой) на неделю (месяц) в Рим (Париж, Прагу, Вену, Брюссель, Лондон, Амстердам, Копенгаген).
- 17.10. Найти энтропию фразы, написанной с сокращениями: ?Вых. д. - п.?
- 17.11. Найти энтропию фразы, написанной с сокращениями: ?Поед. 2 ост. на тр. 16, пересяд. на тр. 29?.
- 17.12. В сборнике 10 пьес, которые автор советует прочесть в прямом или обратном (от 10 к 1) порядке. Найти энтропию числа способов прочтения пьес, предложенных автором.
- 17.13. Из четырёх слов можно составить двухсловные синтагмы, в которых порядок слов имеет значение, а слова могут повторяться. Найти энтропию числа способов образования синтагм.
- 17.14. Найти синтаксическую информацию подчёркнутого словосочетания во фразе Мероприятие переносится с завтрашнего дня на будний день (кроме среды) следующей недели.
- 17.15. Найти синтаксическую информацию подчёркнутого словосочетания во фразе В рамках доклада конференции обсуждается глагольная морфология одного скандинавского (но не шведского) языка.
- 17.16. Найти синтаксическую информацию подчёркнутого словосочетания во фразе Грани шести кубиков с цифрами, больше 2, была разного цвета.
- 17.17. Определить число букв с нулевой синтаксической информацией слова кенгуру.
- 17.18. Найти синтаксическую информацию пятой буквы английского слова earlier.

## 3. Письменная работа

Тема 3

Примеры задач:

- 20.1. Первая фраза абзаца содержит 4 прилагательных, вторая и третья - 3 прилагательных и четвёртая - 2 прилагательных. Найти выборочное среднее числа прилагательных в абзаце.
- 20.2. Текст состоит из двадцати фраз. Четверть фраз этого текста содержит четыре слова, записанного катаканой, ещё четверть - два таких слова, остальные фразы - одно такое слово. Найти выборочное среднее числа слов, записанных катаканой.
- 20.3. Самое короткое слово в тексте стоит из одной буквы, а самое длинное слово содержит редуцированный пятибуквенный корень и трёхбуквенный аффикс. Найти размах выборки длин слов текста.
- 20.4. В каждой строфе поэмы встречается как минимум один эпитет. Если число эпитетов, содержащихся в строфах поэмы, выстроить в вариационный ряд и разбить его на три равных интервала, то длина и число интервалов совпадут. Найти максимально возможное число эпитетов, встречающихся в поэме.
- 20.5. При проверке диктанта было записано число ошибок, допущенных учениками: 4, 11, 9, 16, 0, 18, 7, 20, 13, 6, 8, 11, 15, 7, 9, 13, 8, 17, 14, 8. Составить вариационный интервальный ряд, взяв четыре равных интервала. Вычислить относительные частоты, моду и модальный интервал. Построить гистограмму распределения частот.

20.6. При анализе статей по лингвистике было зафиксировано число лингвистических терминов: 5, 8, 6, 9, 4, 11, 2, 10, 3, 6, 2, 14, 7, 4, 2. Составить вариационный интервальный ряд, взяв четыре равных интервала. Вычислить относительные частоты, среднее выборочное и среднее интервальное значения. Построить гистограмму распределения частот.

20.7. Некоторые из пятнадцати поэтов ратовали за чистоту родного языка, а некоторые специально употребляли заимствования. Дана выборка числа заимствованных слов в их самых известных стихах: 9, 14, 3, 1, 4, 0, 19, 7, 3, 17, 24, 2, 3, 21, 10. Вычислить среднее выборочное значение и среднее интервальное значение при разбиении выборки на три, четыре и шесть интервалов.

20.8. В десяти книгах по лингвистике было подсчитано число опечаток: 6, 2, 3, 12, 5, 7, 0, 3, 8, 4. На какое максимальное число интервалов можно разбить данные выборки так, чтобы первый интервал оказался модальным?

20.9. Сотрудник отвечает по телефону: Директор поехал в филиал, равномерно меняя порядок слов. Вычислить относительную частоту вариантов произнесения этой фразы.

20.10. Арабское слово  $\hat{}$  как существительное может читаться как malak ?ангел?, malik ?король?, mulk ?царство; имущество? и milk ?имущество?. Построить статистический ряд чтений слова полагая, что их частоты соотносятся как 1:10:3:2, и найти моду.

20.11. Найти относительные частоты повторяющихся букв и моду во фразе Вдали отчётливо зазвонил колокол.

20.12. Найти букву, ранг которой равен 3, во фразе Он не хотел, но вышло именно так.

20.13. Найти префикс, обладающий наибольшей частотой в строфе из песни В. Высоцкого:

Смешно! Не правда ли, смешно! Смешно!

А он шутил - недошутил,

Недораспробовал вино И даже недопригубил.

20.14. Составить полный частотный словарь японских слов, встречающихся в стихотворении Ильи Пушкина (в японском тексте слова пишутся слитно, для отделения слов в строке используется полужирное начертание):

### Экзамен

Вопросы к экзамену:

1. Множество лингвистических объектов
2. Операции над множествами  $\mathcal{P}$
3. Применение диаграмм Эйлера-Венна при решении практических задач
4. Бинарные отношения. Прямое произведение множеств
5. Свойства отношений
6. Лингвистическое явление как математическая величина
7. Функции
8. Диахронический скачок и его моделирование с помощью элементарных функций
9. Моделирование информационного построения речи
10. Глоттохронология, информационная схема текста и их моделирование с помощью аппарата бесконечно малых величин и пределов
11. Понятие предела функции в точке.
12. Определение предела функции в точке.
13. Свойства пределов.
14. Понятие бесконечно малой величины. Свойства бесконечно малых функций.
15. Методы вычисления пределов.
16. Глоттохронология
17. Динамика лингвистических процессов и ее описание с помощью приемов дифференциального исчисления
18. Использование дифференциала для приближенных вычислений в лингвистических задачах.
19. Лингвистические задачи, приводящие к понятию интеграла
- 6.1. Интегралы и их применение к лингвистическим задачам 52
20. Вероятностно-информационные оценки нормы языка и статистическое построение текста
21. Информационные измерения в тексте
22. Случайная лингвистическая величина, ее характеристики и функция распределения
23. Случайные процессы
24. Статистика
25. Методы организации статистического наблюдения над текстом.

### 6.4 Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

В КФУ действует балльно-рейтинговая система оценки знаний обучающихся. Суммарно по дисциплине (модулю) можно получить максимум 100 баллов за семестр, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов.

Для зачёта:

56 баллов и более - "зачтено".

55 баллов и менее - "не зачтено".

Для экзамена:

86 баллов и более - "отлично".

71-85 баллов - "хорошо".

56-70 баллов - "удовлетворительно".

55 баллов и менее - "неудовлетворительно".

Форма контроля	Процедура оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций	Этап	Количество баллов
<b>Семестр 2</b>			
<b>Текущий контроль</b>			
Лабораторные работы	В аудитории, оснащённой соответствующим оборудованием, обучающиеся проводят учебные эксперименты и тренируются в применении практико-ориентированных технологий. Оцениваются знание материала и умение применить его на практике, умения и навыки по работе с оборудованием в соответствующей предметной области.	1	30
Письменная работа	Обучающиеся получают задание по освещению определённых теоретических вопросов или решению задач. Работа выполняется письменно и сдаётся преподавателю. Оцениваются владение материалом по теме работы, аналитические способности, владение методами, умения и навыки, необходимые для выполнения заданий.	2	10
		3	10
<b>Экзамен</b>	Экзамен нацелен на комплексную проверку освоения дисциплины. Экзамен проводится в устной или письменной форме по билетам, в которых содержатся вопросы (задания) по всем темам курса. Обучающемуся даётся время на подготовку. Оценивается владение материалом, его системное освоение, способность применять нужные знания, навыки и умения при анализе проблемных ситуаций и решении практических заданий.		50

## 7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

### 7.1 Основная литература:

- Чернявская, В.Е. Лингвистика текста. Лингвистика дискурса : учеб. пособие / В.Е. Чернявская. - 4 -е изд., стер. - Москва : Флинта, 2016. - 203 с. - ISBN 978-5-9765-1454-6. - Режим доступа: <http://znanium.com/catalog/product/1034619>
- Информатика и лингвистика : учеб. пособие / Т.М. Волосатова, Н.В. Чичварин. - М. : ИНФРА-М, 2018. ? 196 с.? (Высшее образование: Бакалавриат). - Режим доступа: <http://znanium.com/catalog/product/938009>

### 7.2. Дополнительная литература:

- Грудева, Е.В. Корпусная лингвистика : учеб. пособие / Е.В. Грудева. - 3-е изд., стер. - Москва : ФЛИНТА, 2017. - 165 с. - ISBN 978-5-9765-1497-3. - Режим доступа: <http://znanium.com/catalog/product/1032488>
- Щипицина, Л. Ю. Информационные технологии в лингвистике [Электронный ресурс] : учеб. пособие / Л. Ю. Щипицина. ? М. : ФЛИНТА, 2013. ? 128 с. - ISBN 978-5-9765-1431-7 - Режим доступа: <http://znanium.com/catalog/product/462989>

## 8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

- МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА? - <https://texts.news/lingvistika-perevod/matematiceskaya-lingvistika-79219.html>
- Математическая лингвистика - [http://informaticslib.ru/news/c0016\\_1.shtml](http://informaticslib.ru/news/c0016_1.shtml)
- Математическая лингвистика. Практикум - ЭБС IPRbooks - <http://www.iprbookshop.ru/67815.html>

## 9. Методические указания для обучающихся по освоению дисциплины (модуля)

Вид работ	Методические рекомендации
лабораторные работы	Лабораторные работы выполняются по темам, определенным учебным планом. Легенды для конкретной работы предлагаются преподавателем. Каждая лабораторная работа завершается отчетом. В отчете должны быть четко определены постановка задачи, используемый инструментарий, пути решения задачи, подробный ход решения задачи, выводы. Приветствуется обсуждение и возможные альтернативные варианты решения. Инструментарий зависит от имеющегося программного обеспечения.

Вид работ	Методические рекомендации
самостоятельная работа	Самостоятельная работа включает в себя работу с лекционным материалом, подготовку к лабораторным работам и выполнение лабораторных работ вне аудитории, если это предлагается преподавателем, подготовку отчета, а также изучение нового материала по сети. Изучение нового материала по теме должно обязательно сопровождаться ознакомлением с новейшими достижениями, так как данная сфера относится к быстро развивающимся областям. Поэтому приветствуется включение в отчеты по лабораторным работам а также вопросы во время лекций по новейшим достижениям по изучаемой теме, это может поощряться преподавателем дополнительными баллами.
письменная работа	Письменная работа выполняются по темам, определенным учебным планом. Легенды для конкретной работы предлагаются преподавателем. В работе должны быть четко определены постановка задачи, используемый инструментарий, пути решения задачи, подробный ход решения задачи, выводы. Приветствуется обсуждение и возможные альтернативные варианты решения. Работа должна быть законченным исследованием по той или иной микротеме.
экзамен	Экзамен проводится в форме тестирования. Все вопросы и весь материал имеется в виртуальной аудитории. Время тестирования варьируется так, чтобы на ответ на один вопрос отводился от одного до трех минут. Обычно тест открывается на сутки, количество попыток регламентируется преподавателем. Окончательная оценка ставится как арифметическое среднее оценки всех попыток, но может изменяться преподавателем.

#### **10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)**

Освоение дисциплины "Математическая лингвистика" предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows Professional 7 Russian

Пакет офисного программного обеспечения Microsoft Office 2010 Professional Plus Russian

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен обучающимся. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.

#### **11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)**

Освоение дисциплины "Математическая лингвистика" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

## **12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья**

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 09.04.02 "Информационные системы и технологии" и магистерской программе Компьютерная лингвистика .