

*Министерство науки и высшего образования
Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение
высшего образования «Северо-Осетинский
государственный университет
имени Коста Левановича Хетагурова»*

**РАБОЧАЯ ПРОГРАММА
ДИСЦИПЛИНЫ
«Теория вероятностей и
математическая статистика»**

Направление подготовки 39.03.01
Социология

Профиль: "Социология"

Форма обучения – очная

Владикавказ, 2022

Программа составлена в соответствии с Федеральным государственным образовательным стандартом по направлению подготовки 39.03.01. Социология, утвержденным приказом Министерства образования и науки Российской Федерации от 12 ноября 2015 г. № 1328, учебным планом подготовки бакалавров по направлению 39.03.01. Социология, утвержденным ученым советом ФГБОУ ВО «Северо-Осетинский государственный университет имени Коста Левановича Хетагурова» 31.05. 2022г., протокол № 13.

Одобрена советом исторического факультета

(протокол № 9 от «08» апреля 2022 г.)

Председатель Совета факультета
З.Т. Плиева



Цели освоения дисциплины

Целями освоения дисциплины «Теория вероятностей и математическая статистика» являются:

- изучение и практическое освоение теории вероятностей, математической статистики и базовых методов статистического анализа данных в социальных науках,
- изучение и практическое освоение компьютерной программы, применяемой для математико-статистического анализа данных (пакета SPSS),
- приобретение понимания специфики работы с количественными данными в социальных науках, понимания типов задач, которые могут быть решены с помощью математико-статистических методов.

Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать основные понятия теории вероятностей, математической статистики, методы статистического анализа данных в пределах программы курса,
- Уметь ставить и понимать социологические задачи, которые могут быть решены с помощью методов теории вероятностей и математической статистики; понимать специфику данных, используемых в математико-статистическом анализе.
- Иметь навыки самостоятельного статистического анализа данных на компьютере в программе SPSS.

В результате освоения дисциплины студент осваивает следующие компетенции:

<i>Компетенция</i>	<i>КОД по ОС НИУ ВШЭ</i>	<i>Дескрипторы</i>	<i>Формы и методы обучения, способствующие формированию и развитию компетенции</i>
Способен работать с информацией: находить, оценивать и использовать информацию из различных источников, необходимую для решения научных и профессиональных задач (в том числе на основе системного	УК-5	Демонстрирует знание основных информационных ресурсов (источников) в профессиональной области (РБ)	Лекции, семинары
		Знает способы структурирования информации (РБ)	Семинары
		Находит информацию по заданным критериям (СД)	Семинары, выполнение самостоятельных работ
		Сохраняет результаты поиска информации (СД)	Семинары, выполнение самостоятельных работ

подхода)		Пользуется различными источниками информации (СД)	Семинары, выполнение самостоятельных работ
		Анализирует полученную информацию по заданным критериям (СД)	Лекции, семинары, выполнение самостоятельных работ
Способен использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа и моделирования, теоретического и экспериментального исследования при решении профессиональных	ПК-5	Демонстрирует умение решать поставленные профессиональные задачи с помощью статистических методов анализа (РБ)	Лекции, семинары, выполнение самостоятельных работ
		Понимает основной принцип причинноследственного анализа (РБ)	Лекции, семинары, выполнение самостоятельных работ
		Формулирует требующую решения исследовательскую задачу на основе имеющихся данных (РБ)	Лекции, семинары, выполнение самостоятельных работ
		Понимает особенности методов анализа, применяемых в научных исследованиях (РБ)	Лекции, семинары, выполнение самостоятельных работ
		Применяет методы статистического анализа данных для исследования общественных	Семинары, выполнение самостоятельных работ
задач		закономерностей (СД)	
		Интерпретирует результаты статистического анализа данных (формулирует к задаче статистический и содержательный выводы) (СД)	Лекции, семинары, выполнение самостоятельных работ
		Правильно использует статистических метод в качестве инструмента решения исследовательской задачи (СД)	Лекции, семинары, выполнение самостоятельных работ
		Демонстрирует знание основных методов статистического анализа данных (РБ)	Лекции, семинары

Способен обрабатывать и анализировать данные для подготовки аналитических решений, экспертных заключений и рекомендаций	ПК-10	Знает способы кодирования, формализации информации для количественного анализа данных (РБ)	Семинары
		Соблюдает правила академической этики при использовании информации и подготовке аналитических решений (МЦ)	Семинары, выполнение самостоятельных работ
		Признает необходимость изучения новой информации, данных в профессиональной области и подготовке аналитических решений (МЦ)	Семинары, выполнение самостоятельных работ

Место дисциплины в структуре образовательной программы

Настоящая дисциплина является обязательной дисциплиной профессионального цикла бакалаврской программы направления 39.03.01 Социология.

Изучение данной дисциплины базируется на следующих дисциплинах: «Алгебра и анализ».

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

«Анализ социологических данных», «Методология и методы социологического исследования», «Анализ данных в социальных науках: методы факторизации и кластеризации», Научно-исследовательский семинар, и др.

Тематический план учебной дисциплины

Название темы	Всего часов	Аудиторные занятия		
		Лекции	Семинары, практические занятия	Самостоятельная работа
Модуль 1. Теория вероятностей.	144	36	36	72

1	Предмет теории вероятностей. Случайный эксперимент. Выбор из конечной совокупности. Пространство элементарных исходов (событий). Классическое и статистическое определение понятия вероятности.	8	2	2	4
2	События и операции над ними. Правила сложения и умножения вероятностей. Условная вероятность. Зависимость/независимость событий	8	2	2	4
3	Формула полной вероятности. Формула Байеса.	8	2	2	4
4	Случайные величины – дискретные и непрерывные. Характеристики случайной величины – математическое ожидание, дисперсия.	8	2	2	4
5	Испытания Бернулли. Биномиальное распределение.	8	2	2	4
6	Функция и плотность распределения. Квантили распределения. Равномерное распределение, его числовые характеристики и квантили.	8	2	2	4
7	Нормальное распределение: ситуации возникновения, функция и плотность распределения. Произвольное и стандартное нормальное распределение. Работа с таблицами нормального распределения. Квантили распределения.	8	2	2	4
8	Теорема Муавра – Лапласа. Центральные предельные теоремы.	8	2	2	4
9	Закон больших чисел.	8	2	2	4

Название темы	Всего часов	Аудиторные занятия		
		Лекции	Семинары, практические занятия (к/класс)	Самостоятельная работа

Модули 1-2. Математическая статистика.				
10 Введение в статистический анализ данных. Выборка и генеральная совокупность. Вероятностный отбор. Простая случайная выборка (SRS).	8	2	2	4
11 Данные и переменные. Описательная статистика (меры среднего и меры разброса для разных типов шкал) и статистические графики.	8	2	2	4
12 Точечное и интервальное оценивание. Доверительные интервалы для среднего и доли. Размер выборки.	8	2	2	4
13 Статистические гипотезы и их проверка. Основные понятия и алгоритмы. Статистическая значимость, p-value. Ошибки статистического вывода (1-го и 2-го рода).	8	2	2	4
14 Совместное распределение переменных. Таблицы сопряженности. Критерий Хи-квадрат. Анализ стандартизированных остатков	8	2	2	4
15 Парные коэффициенты корреляции (линейный, ранговые). Проверка статистической значимости коэффициентов корреляции.	8	2	2	4
16 Проверка гипотезы о математических ожиданиях (средних). t-критерий Стьюдента (одновыборочный). Гипотеза о равенстве доли определенному числу.	8	2	2	4
17 Гипотеза о равенстве дисперсий в двух группах. t-критерий Стьюдента для двух независимых выборок. t-критерий Стьюдента для двух связанных выборок. Гипотеза о равенстве долей в двух группах.	8	2	2	4
18 Дисперсионный анализ (ANOVA)	8	2	2	4

Формы контроля знаний студентов

	Форма контроля	2 курс		Параметры
		1 модуль	2 модуль	
Текущий контроль	Контрольная работа*	1	1	<p>Каждая контрольная состоит из набора микропроверочных работ.</p> <p>КР №1 (состоит из двух микропроверочных работ)</p> <p>КР №2 (состоит из двух микропроверочных работ)</p> <p>Каждая микропроверочная работа выполняется в течение 40-60 минут в аудитории (учебном классе).</p>
Итоговый контроль	Экзамен	1	1	Выполняется в течение ≤ 160 минут (письменно) в аудитории.

* Контрольные работы разбиваются преподавателем на части (содержательные задачи), которые выполняются поэтапно в течение курса в виде микроконтрольных работ.

Критерии оценки знаний, навыков

Оценки по всем формам текущего и итогового контроля выставляются по 10-ти балльной шкале.

Оценки за контрольные работы и экзаменационные работы выставляются, исходя из следующих критериев:

1. правильность решения задачи,
2. полнота решения задачи, аргументация решения,
3. корректность оформления статистического вывода,
4. наличие и корректность содержательной интерпретации статистического вывода.

Содержание дисциплины

Тема 1. Предмет теории вероятностей. Случайный эксперимент. Выбор из конечной совокупности. Пространство элементарных исходов (элементарных событий).

Классическое и статистическое определение понятия вероятности. Случайность как предмет изучения. Предмет теории вероятностей. Случайный эксперимент. Пространство элементарных исходов (событий). Различные подходы к изучению вероятности. Объективизм и субъективизм. Классическое и статистическое определение понятия вероятности. Относительная частота и вероятность.

Литература основная

1. Макаров А.А., Пашкевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.9-24.
2. Мостеллер Ф., Рурке Р., Томас Дж. Вероятность. М.: МЦНМО, 2015. С. 15-70.
3. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.29-52
4. Bluman A. Elementary statistics. McGraw-Hill, 2008. 2nd ed. Pp.153-163

Литература дополнительная

1. Мизес Р. Вероятность и статистика. М.: КомКнига, 2007
2. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007.

Тема 2. События и операции над ними.

Правила сложения и умножения вероятностей. Условная вероятность. Зависимость/независимость событий. Совместимые и несовместимые события.

Литература основная

1. Макаров А.А., Пашкевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.25-36.
2. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.60-63, 68-77
3. Мостеллер Ф., Рурке Р., Томас Дж. Вероятность. М.: МЦНМО, 2015. С. 77-138, 139-173.
4. Bluman A. Elementary statistics. McGraw-Hill, 2008. Pp.163-174

Тема 3. Формула полной вероятности и формула Байеса.

Полная группа событий. Формула полной вероятности. Априорные и апостериорные вероятности. Формула Байеса.

Литература основная

1. Макаров А.А., Пашкевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.37-52.
2. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.77-79
3. Мостеллер Ф., Рурке Р., Томас Дж. Вероятность. М.: МЦНМО, 2015. С. 173-184.
4. Bluman A. Elementary statistics. McGraw-Hill, 2008. Pp.174-182

Тема 4. Случайные величины - дискретные и непрерывные. Характеристики случайной величины – математическое ожидание, дисперсия.

Дискретные и непрерывные СВ. Основные социологические шкалы. Математические операции со случайными величинами. Характеристики случайной величины – математическое ожидание и дисперсия.

Литература основная

1. Макаров А.А., Пашкевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.61-74.
2. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.20-22

3. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И.* Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.84-112
4. *Мостеллер Ф., Рурке Р., Томас Дж.* Вероятность. М.: МЦНМО, 2015. С. 185-227.
5. *Bluman A.* Elementary statistics. McGraw-Hill, 2008. Pp. 69-90, 199-206,

Тема 5. Испытания Бернулли. Биномиальное распределение.

Вероятность в дискретных пространствах. Серия испытаний Бернулли.

Литература основная

1. *Макаров А.А., Пашкевич А.В.* Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.53-60.
2. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И.* Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.63-68
3. *Мостеллер Ф., Рурке Р., Томас Дж.* Вероятность. М.: МЦНМО, 2015. С. 244-276.
4. *Bluman A.* Elementary statistics. McGraw-Hill, 2008. Pp.207-214

Литература дополнительная

1. *Сигел Э.* Практическая бизнес-статистика. М.: Издательский дом "Вильямс", 2004

Тема 6. Функция и плотность распределения случайной величины.

Функция и плотность распределения. Квантили распределения. Равномерное распределение, его числовые характеристики и квантили.

Литература основная

1. *Макаров А.А., Пашкевич А.В.* Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.91-108.
2. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И.* Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.185-195
3. *Bluman A.* Elementary statistics. McGraw-Hill, 2008. Pp.194-199, 227-242

Литература дополнительная

1. *Хили Дж.* Статистика. Социологические и маркетинговые исследования. СПб.: Питер, 2005

Тема 7. Нормальное распределение: произвольное и стандартное.

Нормальное распределение: ситуации возникновения, функция и плотность распределения. Произвольное и стандартное нормальное распределение. Работа с таблицами нормального распределения. Квантили распределения.

Литература основная

1. *Макаров А.А., Пашкевич А.В.* Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.91-108.
2. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.51-53
3. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И.* Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.173-185,

4. *Bluman A. Elementary statistics. McGraw-Hill, 2008. Pp.242-251*

Литература дополнительная

1. *Хили Дж. Статистика. Социологические и маркетинговые исследования. СПб.: Питер, 2005*

Тема 8. Теорема Муавра - Лапласа. Центральные предельные теоремы. Теорема Муавра - Лапласа. Центральные предельные теоремы.

Литература основная

1. *Макаров А.А., Паишевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016). С.109-120.*
2. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.126-136, 148-156*
3. *Bluman A. Elementary statistics. McGraw-Hill, 2008. Pp.207-214, 214-220*

Литература дополнительная

1. *Сигел Э. Практическая бизнес-статистика. М.: Издательский дом «Вильямс», 2004*

Тема 9. Закон больших чисел.

Закон больших чисел.

Литература основная

1. *Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.64-68*
2. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009. С.206-221, 226-233*
3. *Bluman A. Elementary statistics. McGraw-Hill, 2008. Pp.251-263*

Тема 10. Введение в статистический анализ данных. Выборка и генеральная совокупность. Вероятностный отбор. Простая случайная выборка (SRS).

Выборка и генеральная совокупность. Вероятностный отбор. Простая случайная выборка (SRS). Для решения каких задач в социальных науках используется статистический анализ данных. Примеры исследований, построенных на статистическом анализе данных. Источники данных для статистического анализа: опрос, перепись, агрегированная статистика. Основные методы анализа. Обзор содержания курса. Основы работы в SPSS. Архитектура пакета. Синтаксис SPSS.

Литература основная

1. *Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.19-70.*
2. *Толстова Ю.Н. Методология математического анализа данных // Толстова Ю.Н. Социология и математика. М.: Научный мир, 2003. С.80-94.*
3. *Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.10-13*
4. *Паишевич А.В. Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.8-28.*

5. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 15-44, 296-314.
6. *Моосмюллер Г., Ребик Н.Н.* Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007. – С. 21-41.
7. *Чуриков А.В.* Случайные и неслучайные выборки в социологических исследованиях // Социальная реальность. 2007. №4. С. 89-109.
8. *Чуриков А.В.* Основы формирования выборки: лекции для студентов направления 040100.62 (Социология) – М.: ГУ-ВШЭ, 2005, с. 8-38.
9. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.1-30

Литература дополнительная

1. *Айвазян С. А., Мхитарян В. С.* Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. *Бююль А., Цёфель П.* SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.
3. *Девятко И.Ф.* Методы социологического исследования. / И.Ф. Девятко. – 4-е изд. – М.: КДУ, 2006.
4. *Наследов А.* SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
5. *Толстова Ю.Н.* Анализ социологических данных. М.: Научный мир, 2000
6. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
7. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 11. Данные и переменные. Описательная статистика (меры среднего и меры разброса для разных типов шкал) и статистические графики.

Шкалы измерений (интервальная, порядковая, категориальная). Средние величины, медиана, мода. Дисперсия. Распределения и их характеристики. Статистическая визуализация. График плотности распределения. Диаграммы рассеяния. Гистограммы. Столбиковые диаграммы. Линейные диаграммы. Коробчатые диаграммы. Круговые диаграммы.

Литература основная

1. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.19-70.
2. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007, С.10-36.
3. *Пашкевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.29-76.
4. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 44-58.
5. *Моосмюллер Г., Ребик Н.Н.* Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007. – С.42-60.
6. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.31-72.

Литература дополнительная

1. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. Бьюль А., Цёфель П. SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.
3. Девятко И.Ф. Методы социологического исследования. / И.Ф. Девятко. – 4-е изд. – М.: КДУ, 2006.
4. Наследов А. SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
5. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000
6. Бослаф С. Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
7. Bluman A. Elementary statistics. McGraw-Hill, 2008.

Тема 12. Точечное и интервальное оценивание. Доверительные интервалы для среднего и доли. Размер выборки.

Два вида оценивания: точечное и интервальное. Понятие стандартной ошибки. Доверительные интервалы для среднего арифметического и для доли. Уровень доверительной вероятности, его содержательный смысл. Связь точности исследования и объема выборки исследования.

Литература основная

1. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.71-100.
2. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.37-39
3. Паишевич А.В. Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.124-176.
7. Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б. Математическая статистика для социологов: задачник: учебное пособие для вузов / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010, С.45-65.
4. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С.125-146.
5. Моосмюллер Г., Ребик Н.Н. Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007. – С.42-60.
6. Чуриков А.В. Основы формирования выборки: лекции для студентов направления 040100.62 (Социология) – М.: ГУ-ВШЭ, 2005, С. 8-38.
7. Agresti, A. and Finlay, B. Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.73-142.

Литература дополнительная

1. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. Бьюль А., Цёфель П. SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.
3. Девятко И.Ф. Методы социологического исследования. / И.Ф. Девятко. – 4-е изд. – М.: КДУ, 2006.
4. Наследов А. SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
5. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000

6. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
7. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 13. Статистические гипотезы и их проверка. Основные понятия и алгоритмы.

Статистическая значимость.

Введение в проверку статистических гипотез. Базовые понятия: статистическая гипотеза, нулевая гипотеза, альтернативная гипотеза, направленные и ненаправленные альтернативные гипотезы, статистика критерия, фиксированный уровень значимости, минимальный уровень значимости (p -value), критическая область, ошибки 1 и 2 рода, статистический вывод и мн.др. Примеры.

Литература основная

1. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.101-107, 120-122, 130-140.
2. *Пашкевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.177-197.
3. *Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б.* Математическая статистика для социологов: задачник: учебное пособие для вузов / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010, С.66-68, 106-108.
4. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С.82-96.
5. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.143-147.

Литература дополнительная

1. *Айвазян С. А., Мхитарян В. С.* Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
3. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 14. Совместное распределение переменных. Таблицы сопряженности. Критерий Хиквадрат. Анализ стандартизированных остатков.

Понятие о двумерной частотной таблице и способах отражения в ее виде независимости рассматриваемых признаков. Ожидаемые частоты и наблюдаемые частоты.

Логика проверки гипотезы об отсутствии связи между двумя номинальными (или рассматриваемыми как номинальные) признаками на основе критерия Хи-квадрат. Основная модель. Интерпретация остатков. Примеры.

Литература основная

1. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS: учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.57-67, 81.
2. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С.240-253, 263-267.

3. *Пашкевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.197-215.
4. *Моосмюллер Г., Ребик Н.Н.* Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007. – С.74-80.
5. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.221-255.

Литература дополнительная

1. *Айвазян С. А., Мхитарян В. С.* Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. *Бююль А., Цёфель П.* SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.
3. *Наследов А.* SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
4. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
5. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009
6. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 15. Парные коэффициенты корреляции

Коэффициента линейной корреляции Пирсона (для переменных, измеренных на уровне шкал высокого порядка), коэффициенты ранговой корреляции Спирмена и Кендалла (для переменных, измеренных на порядковом уровне). Проверка статистических гипотез о равенстве коэффициентов 0. Соотнесение с ограничениями социологических шкал.

Литература основная

1. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.126-127, 128-129.
2. *Пашкевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.77-107, 264-287.
3. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS: учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.67-81, 115-146, 166-182.
4. *Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б.* Математическая статистика для социологов: задачник: учебное пособие для вузов / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010, С.34-44, 102-104.
5. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С.253-263, 267-269, 208-219, 230-239.
6. *Моосмюллер Г., Ребик Н.Н.* Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007. – С.104-127.
7. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.255-300.

Литература дополнительная

1. *Айвазян С. А., Мхитарян В. С.* Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. *Бююль А., Цёфель П.* SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.

3. *Наследов А.* SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
4. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
5. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 16. Проверка гипотезы о математических ожиданиях (средних). Т-тест (одновыборочный). Гипотеза о равенстве доли определенному числу.

Алгоритм (правила) проверки гипотез. Статистический критерий проверки гипотез о равенстве математического ожидания и доли определенному числу (константе). Статистический критерий проверки гипотезы о равенстве двух математических ожиданий для случая парных (связных) выборок. Содержательный характер исследовательских задач. Примеры.

Литература основная

1. *Паишевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.226-228.
2. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.94-96
3. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С.147-158.
4. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009. Pp.147-159.

Литература дополнительная

1. *Айвазян С. А., Мхитарян В. С.* Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
3. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Тема 17. Гипотеза о равенстве дисперсий в двух группах. Т-тест для двух независимых выборок. Т-тест для двух связанных выборок. Гипотеза о равенстве долей в двух группах. Алгоритм (правила) проверки гипотез. Статистический критерий проверки гипотезы о равенстве двух математических ожиданий для случая независимых выборок. Статистические критерии проверки гипотез о равенстве дисперсий и долей в двух группах. Содержательный характер исследовательских задач. Примеры.

Литература основная

1. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007. С.113-119.
2. *Паишевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014. С.215-263.
3. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007. – С.82-94, 97-99.

4. Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б. *Математическая статистика для социологов: задачник: учебное пособие для вузов* / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010, С.79-105
5. Тюрин Ю.Н., Макаров А.А. *Анализ данных на компьютере: учебное пособие*. – М.: ИД «ФОРУМ», 2008. – С.158-169.
6. Моосмюллер Г., Ребик Н.Н. *Маркетинговые исследования с SPSS: Учебное пособие*. – М.: ИНФРА-М, 2007. – С.81-90.
7. Agresti, A. and Finlay, B. *Statistical Methods for the Social Sciences*. 4th ed. Pearson Prentice Hall, 2009. Pp.183-187, 191-197.

Литература дополнительная

1. Айвазян С. А., Мхитарян В. С. *Теория вероятностей и прикладная статистика*, ч.1. М.: Юнити, 2001.
2. Бююль А., Цёфель П. *SPSS: искусство обработки информации*. СПб.: ДиаСофтЮП, 2005.
3. Наследов А. *SPSS 19: профессиональный статистический анализ данных*. – СПб.: Питер, 2011.
4. Бослаф С. *Статистика для всех*. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
5. Bluman A. *Elementary statistics*. McGraw-Hill, 2008.

Тема 18. Дисперсионный анализ.

Дисперсионный анализ как проверка статистической гипотезы о равенстве средних (формулировка гипотезы, используемый критерий, логика проверки). Однофакторный дисперсионный анализ. Основная модель. Класс решаемых задач. Понимание причинно-следственной связи при использовании однофакторного дисперсионного анализа (априорная модель такой связи). Примеры.

Литература основная

1. Крыштановский А.О. *Анализ социологических данных с помощью пакета SPSS : учебное пособие для вузов*. – М.: Изд. дом ГУ ВШЭ, 2007. – С.99-114.
2. Толстова Ю.Н. *Математико-статистические модели в социологии*. М.: ИД ГУ-ВШЭ, 2007. С.144-190.
3. Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б. *Математическая статистика для социологов: задачник: учебное пособие для вузов* / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010, С.113-124, 126-135.
4. Тюрин Ю.Н., Макаров А.А. *Анализ данных на компьютере: учебное пособие*. – М.: ИД «ФОРУМ», 2008. – С.170-193.
5. Моосмюллер Г., Ребик Н.Н. *Маркетинговые исследования с SPSS: Учебное пособие*. – М.: ИНФРА-М, 2007. – С.90-103.
6. Agresti, A. and Finlay, B. *Statistical Methods for the Social Sciences*. 4th ed. Pearson Prentice Hall, 2009. Pp.369-376.

Литература дополнительная

1. Айвазян С. А., Мхитарян В. С. *Теория вероятностей и прикладная статистика*, ч.1. М.: Юнити, 2001.
2. Бююль А., Цёфель П. *SPSS: искусство обработки информации*. СПб.: ДиаСофтЮП, 2005.

3. *Наследов А.* SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
4. *Бослаф С.* Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
5. *Bluman A.* Elementary statistics. McGraw-Hill, 2008.

Образовательные технологии

Занятия проводятся в форме лекций и семинаров.

Оценочные средства для текущего контроля и аттестации студента

Тематика заданий текущего контроля

В 1 и 2 модулях студентам предлагаются две контрольные работы по задачам теории вероятностей и математической статистики, которые выполняются в классе.

В течение всего учебного года студентам предлагаются дополнительные задания, выполняемые по желанию (добровольно) в виде «Тренировочных задач» для самоподготовки (самопроверки) дома к каждой изученной теме курса. Эти задания требуют решения (расчетов) «вручную» или на компьютере (в программе SPSS). Тренировочные задачи обычно включают в себя самостоятельное решение и содержательную интерпретацию результатов анализа.

Порядок формирования оценок по дисциплине

Аттестация знаний осуществляется в 10-балльной шкале.

Итоговая оценка за курс вычисляется следующим образом:

$$O_{\text{итоговая}} = O_{\text{накопленная}} * 0,6 + O_{\text{экзамен}} * 0,4.$$

Перед вычислением итоговой оценки: **накопленная оценка** (отражающая успешность академической работы в течение 1-2 модулей) и **оценка за экзамен** (отражающая успешность выполнения письменной работы в конце 2-го модуля) округляются до целого. Округление производится в сторону ближайшего целого (арифметический способ).

Формами текущего контроля по данному курсу являются: 1 контрольная работа в первом модуле (она состоит из двух частей, распределенных во времени, – **микроконтрольная А** и **микроконтрольная Б**), 1 контрольная работа во втором модуле (она состоит из двух частей, распределенных во времени, – **микроконтрольная А** и **микроконтрольная Б**). Обе контрольные работы оцениваются в 10-балльной целочисленной шкале (с грейдами: 10, 9, 8, 7, ..., 0 баллов).

Формами итогового контроля выступают: письменный экзамен (который проводится в *конце 1 модуля*) по материалам теории вероятностей и письменный экзамен (который проводится в *конце 2 модуля*) по материалам всего курса целиком с преимущественным акцентом на математическую статистику.

$$O_{\text{накопленная}} = O_{\text{Теория вероятностей}} * 0,45 + O_{\text{Математическая статистика}} * 0,55$$

Оценка за **первую часть дисциплины** («Теория вероятностей», **1 модуль**) формируется следующим образом:

$O_{\text{Теория_вероятностей}} = 0,6 * O_{\text{ТВ_1}} + 0,4 * O_{\text{накопленная_в_1ом_модуле}}$, где

- $O_{\text{ТВ_1}}$ – оценка за письменную экзаменационную работу в 1-ом модуле (с весом 60%)
- $O_{\text{накопленная_в_1ом_модуле}}$ – накопленная оценка в 1-ом модуле, учитывающая оценку за контрольную работу в 1-ом модуле (с весом 50%) и активную академическую работу на семинарских занятиях (решение задач у доски), выполнение домашних заданий для самоподготовки и тренировки к каждому семинару по изученному материалу в 1-ом модуле (тоже с весом 50%).

Оценка за **вторую часть дисциплины** («Математическая статистика», **2 модуль**) формируется следующим образом:

$O_{\text{Математическая_статистика}} = 0,8 * O_{\text{МС_1}} + 0,2 * O_{\text{МС_2}}$, где

- $O_{\text{МС_1}}$ – оценка за контрольную работу во 2-ом модуле (с весом 80%).
- $O_{\text{МС_2}}$ – оценка за активную академическую работу на семинарских занятиях (решение задач у доски), выполнение домашних заданий для самоподготовки и тренировки к каждому семинару по изученному материалу в 1-ом и 2-ом модулях (с весом 20%).

При округлении накопленной и итоговой оценок за курс «Теория вероятностей и математическая статистика» производится в соответствии с арифметическим правилом округления. Так, например, оценка 3,5 балла округляется до 4.

Учебно-методическое и информационное обеспечение дисциплины

Базовые учебники

1. *Макаров А.А., Пашкевич А.В.* Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: МЦНМО, 2015 (2016).
2. *Толстова Ю.Н.* Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ, 2007.
3. *Тюрин Ю.Н., Макаров А.А., Симонова Г.И.* Теория вероятностей: учебник для экономических и гуманитарных специальностей. М.: МЦНМО, 2009.
4. *Крыштановский А.О.* Анализ социологических данных с помощью пакета SPSS: учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2007.
5. *Пашкевич А.В.* Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высш. образования/ Науч. ред.: А. А. Макаров. М. : Издательский центр «Академия», 2014.
6. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. А также: *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере: учебное пособие. – М.: МЦНМО, 2016.
7. *Agresti, A. and Finlay, B.* Statistical Methods for the Social Sciences. 4th ed. Pearson Prentice Hall, 2009.

Основная литература

1. Мостеллер Ф., Рурке Р., Томас Дж. Вероятность. М.: МЦНМО, 2015.
2. Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б. Математическая статистика для социологов: задачник: учебное пособие для вузов / отв. ред. Ю.Н. Толстова. – М.: Изд.дом ГУ-ВШЭ, 2010
3. Моосмюллер Г., Ребик Н.Н. Маркетинговые исследования с SPSS: Учебное пособие. – М.: ИНФРА-М, 2007.

Дополнительная литература

1. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001.
2. Бьюль А., Цёфель П. SPSS: искусство обработки информации. СПб.: ДиаСофтЮП, 2005.
3. Девятко И.Ф. Методы социологического исследования. / И.Ф. Девятко. – 4-е изд. – М.: КДУ, 2006.
4. Наследов А. SPSS 19: профессиональный статистический анализ данных. – СПб.: Питер, 2011.
5. Сигел Э. Практическая бизнес-статистика. М.: Издательский дом «Вильямс», 2004
6. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000
7. Хили Дж. Статистика. Социологические и маркетинговые исследования. СПб.: Питер, 2005
8. Чуриков А.В. Основы формирования выборки: лекции для студентов направления 040100.62 (Социология) – М.: ГУ-ВШЭ, 2005
9. Бослаф С. Статистика для всех. / Пер. с англ. П.А. Волкова, И.М. Флямер, М.В. Либерман, А.А. Галицына. – М.: ДМК Пресс, 2015.
10. Bluman A. Elementary statistics. McGraw-Hill, 2008.

Образцы письменных заданий

Большое количество примеров и упражнений (математических задач и разъяснений к ним) и образцов контрольной и экзаменационной работ приведены в учебниках:

Макаров А.А., Пашкевич А.В. Задачник по теории вероятностей для студентов социальногуманитарных специальностей. – М.: Московский центр непрерывного математического образования (МЦНМО), 2016.

Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. – М.: МЦНМО, 2009. Упражнения – в конце параграфов.

Толстова Ю. Н. Математико-статистические модели в социологии (математическая статистика для социологов). М.: Издательский дом ГУ-ВШЭ, 2007. Упражнения – в конце параграфов.

Толстова Ю.Н., Куликова А.А., Рыжова А.В., Юдин Г.Б. Математическая статистика для социологов: задачник / Отв. ред.: Ю.Н. Толстова. М.: Издательский дом ГУ-ВШЭ, 2010.

Пашкевич А.В. Теория вероятностей и математическая статистика для социологов и менеджеров: учебник для студентов учреждений высшего образования / Науч. ред.: А.А. Макаров. М.: Издательский центр «Академия», 2014. Задачи для самостоятельного решения приведены в завершении каждой главы учебника на следующих страницах: С.69-76, 98-107, 122-123, 167-176, 276-287. Образцы экзаменационных заданий: С.288-304.

Примеры комплектов экзаменационных заданий в 2017-2018 учебном году в курсе «Теория вероятностей и математическая статистика», Социология:

Контрольная работа №1 (теория вероятностей): примеры задач

Задача 1. Известно, что в женской косметичке царит «идеальный порядок» – увы, женщине подчас трудно найти в содержимом нужную вещь, и она долго копошится. В сумочке некоторой женщины находятся 17 предметов губной помады – 4 с оттенком «вишня», 4 – с оттенком «коралл», 4 – с оттенком «роза», остальные – с оттенком «малина». Какова вероятность того, что при случайном выборе четырех предметов из косметички ими окажутся помады одного оттенка?

Задача 2. Вероятность того, что шоколадное мороженое окажется в наличии к концу рабочего дня в первом киоске на железнодорожной станции, равна 0,09. Вероятность наличия этого же сорта мороженого к вечеру рабочего дня во втором киоске равна 0,17. Продажи мороженого в этих киосках осуществляются независимо друг от друга. Какова вероятность, что к концу рабочего дня шоколадное мороженое останется в наличии *только в одном из двух* киосков?

Задача 3. В рамках проекта «Университет для города», было проведено исследование о целевой аудитории публичных лекций университета. В опросе участвовали 450 человек, среди них 260 студентов, 100 – вузовские преподаватели, остальные – “обычные” прохожие. По данным опроса, 45% студентов, 20% преподавателей и 10% “обычных” прохожих планируют посещать лекции в московских музеях, которые запланированы на новый сезон. Исследователь для глубинного интервью случайно выбрал респондента, положительно ответившего на данный вопрос. Какова вероятность того, что случайно выбранный респондент – *не “обычный”* прохожий?

Задача 4. По результатам некоторого социологического исследования, проведенного среди офисных служащих одной корпорации, стало известно, что 60% работников не успевают утром завтракать дома, а уезжают на работу голодными. Найдите вероятность того, что среди 6-ти случайно выбранных респондентов (работников этой корпорации) *менее трех* успевают утром завтракать дома?

Задача 5. По оценкам экспертов, ожидаемая прибыль компании (X) в предстоящем году составит:

X, усл. ед.	– 100	0	+200	+300
P (вероятность)	0,15	0,20		0,35

Найдите математическое ожидание, дисперсию и стандартное отклонение случайной величины X. Единицы измерения найденных параметров подписывайте.

Задача 6. X и Y независимы. Математическое ожидание случайной величины X равно $+7$, а стандартное отклонение равно 6 . Математическое ожидание случайной величины Y равно -5 , а стандартное отклонение равно 9 . Найдите математическое ожидание и дисперсию случайной величины $L = -10X - 3Y - 8$.

Экзаменационная работа (теория вероятностей): примеры задач

Задача 1. Социологический опрос содержит 3 вопроса, на каждый из которых есть 4 варианта ответа. Сколькими различными способами можно ответить на все вопросы, предложенные респондентам в этом исследовании?

Задача 2. Закончите фразу: «Случайные события A и B в теории вероятностей называются независимыми, если ...».

Задача 3. Чему равна вероятность того, что случайная величина, имеющая стандартное нормальное распределение, примет значение большее 5?

Задача 4. На родительском собрании присутствует 16 человек, среди которых 11 имеют высшее образование. Какова вероятность того, что при случайном выборе шести человек в родительских комитет, в него войдет ровно три человека без высшего образования и три человека с высшим образованием?

Задача 5. Судоходная компания в течение лета организует круизы по реке. Очевидно, наибольшая коммерческая эффективность будет достигнута, если все каюты корабля будут полностью заняты отдыхающими. Представитель туристического агентства, сотрудничающий с компанией, предсказывает, что если экономическая среда останется стабильной, то вероятность того, что корабль будет заполнен в течение сезона, равна $0,93$, а если ухудшится, то с вероятностью $0,85$. По оценкам экономистов, вероятность того, что в течение сезона экономическая ситуация сохранится стабильной, равна $0,74$. Найдите вероятность того, что билеты на круизы будут проданы не полностью.

Задача 6. По многолетним наблюдениям геологической разведки, около 20% участков земли, как правило, содержат газ. Какова вероятность, что при бурении 9 скважин, месторождение газа не будет обнаружено более чем в четырех из них?

Задача 7. Легковых автомобилей у автозаправочной станции (АЗС) проезжает вчетверо больше, чем грузовых. Вероятность того, что проезжающий автомобиль подъедет на заправку, составляет для грузового автомобиля $0,05$, для легкового – $0,15$. Только что от бензоколонки отъехал заправленный автомобиль. Какова вероятность, что это был грузовик?

Задача 8. Рассмотрите две случайные величины: X – температура воздуха в градусах Цельсия и Y – температура воздуха в градусах Фаренгейта. Чтобы перевести температуру из шкалы Фаренгейта в шкалу Цельсия, нужно вычесть из числа градусов Фаренгейта 32 и результат разделить на $1,8$. В некотором помещении математическое ожидание температуры воздуха равно 21°C . Найдите математическое ожидание этой температуры в градусах Фаренгейта.

Задача 9. У фирмы, продающей дорогую корпусную мебель, есть два магазина в разных частях города. Вероятность возникновения претензии (рекламации) от покупателей в первом из них равна $0,05$. Вероятность возникновения претензии (рекламации) от покупателей во втором

магазине равна 0,14. Претензии, поступающие от клиентов, происходят в этих магазинах независимо друг от друга. По длительному опыту и наблюдениям известно, что в течение месяца в каждом магазине может случиться не более одной претензии. Рассмотрите случайную величину X – «общее количество поступивших претензий в течение месяца в двух магазинах вместе». Найдите: 1) закон распределения случайной величины X , б) математическое ожидание, дисперсию и стандартное отклонение этой случайной величины. (Единицы измерения при расчетах параметров подписывайте).

Задача 10. X и Y независимы. Математическое ожидание случайной величины X равно -7 , а стандартное отклонение равно 4. Математическое ожидание случайной величины Y равно -9 , а стандартное отклонение равно 5. Найдите математическое ожидание и дисперсию случайной величины $L = 6X - 3Y - 10$.

Задача 11. Случайная величина X нормально распределена со средним 15 и дисперсией 16. Найдите квантиль уровня 0,36.

Задача 12. В среднем, количество появляющихся новых учебников по математике в крупном книжном магазине города составляет 18 за год с дисперсией 16. Найдите вероятность того, что в течение года в книжном магазине появится от 9 до 21 новинок учебной литературы по математике (если считать, что число новых учебников по математике – случайная величина, которая описывается моделью нормального распределения с указанными параметрами). Схематично отразите искомую вероятность на графике нормального распределения и на графике стандартного нормального распределения. Рисунки должны быть аккуратными.

Задача 13. Можно сказать, что массы животных имеют нормальное распределение. Пусть коровы определенной молочной породы имеют средний вес, равный 470 кг. Чему равно стандартное отклонение веса коров этой породы, если известно, что 35% животных этой породы имеют вес больше 490 кг?

Задача 14. В лыжной гонке на 50 км участвует 1000 человек. В среднем лишь 80% участников выдерживают испытание до конца, а остальные сходят с дистанции. Оцените вероятность того, что в этой гонке к финишу придет не менее 780 человек. (Вычисления производите с точностью до второго знака после запятой).

Задача 15. При работе с документацией секретарь допускает опечатки. Число опечаток – случайная величина, распределенная по закону Пуассона, со средним значением 1 опечатка на одной странице. Какова вероятность того, что при просмотре 3 страниц текста в них обнаружится не менее четырех опечаток?

Задача 16. Совместный закон распределения случайных величин X и Y задан таблицей (для каждой пары (т.е. комбинации, сочетания) значений X и Y приводится своя вероятность). Значения X – указаны в строках, значения Y – в столбцах. Найдите: $\text{Cov}(X, Y)$ и $\text{Corr}(X, Y)$. Интерпретируйте полученное значение корреляции.

$X \backslash Y$	-1	0	2
-1	0.10	?	0.10
-3	0.20	0.30	0.10

Контрольная работа №2 (математическая статистика): примеры задач

Задача 1. Четырнадцать секретарей соревновались в финале конкурса по «слепой» печати на клавиатуре. Было зафиксировано, сколько опечаток (в расчете на 100 знаков) сделал каждый участник конкурса. Дана выборка, X : 15, 12, 29, 3, 27, 21, 13, 25, 4, 3, 32, 19, 7, 19 (опечаток). По методу Дж. Тьюки найдите медиану, квартили (нижнюю и верхнюю), квартильный размах. Определите, есть ли в выборке нетипичные (нехарактерные) значения наблюдений – «подозрительные» и/или «выбросы». Обоснуйте расчетами. Интерпретируйте значение верхней квартили.

Задача 2. Имеется следующая информация (выборка) относительно оценок, полученных школьниками на выпускных экзаменах ЕГЭ (100-балльная система) по двум предметам: физике и химии. Можно ли считать два этих академических показателя монотонно коррелированными? Рассчитайте коэффициент ранговой корреляции Спирмена. Развернуто интерпретируйте полученный ответ.

Выпускник	Ваня	Таня	Коля	Оля	Ром	Тома	Саша	Даша	Вася	Ася
Физика, балл ЕГЭ	56	46	47	93	81	51	65	51	74	65
Химия балл ЕГЭ	64	41	52	87	71	57	67	54	70	73

Задача 3. Менеджер университетского буфета оценивает, имеется ли взаимосвязь в заказах двух блюд меню – кофе и пирожного. Выборка наблюдений состоит из 7 дней октября. Постройте диаграмму рассеяния и рассчитайте коэффициент линейной корреляции. Развернуто интерпретируйте ответ. Можно ли считать два этих показателя линейно коррелированными?

День	21 окт.	22 окт.	23 окт.	24 окт.	25 окт.	26 окт.	27 окт.
Кофе, число заказов	25	40	31	48	50	27	17
Пирожное, число заказов	30	25	20	25	40	31	18

Задача 4. Используя данные предыдущей задачи (см. №3), рассмотрите величину «Число заказов пирожного». Укажите коэффициент вариации. Содержательно (и развернуто) интерпретируйте полученное значение коэффициента вариации. Укажите стандартную ошибку и 90%-ый доверительный интервал. Содержательно (и развернуто) интерпретируйте полученный интервал.

Задача 5. В сентябре 2017 г. в ходе всероссийского исследования случайной выборке населения РФ был задан вопрос «Как часто в течение недели Вы смотрите телевизор (не важно, где именно и какой канал)?». Постройте доверительный интервал для доли жителей России, которые смотрят телевизор «Реже, чем раз в неделю или не смотрят вовсе». Уровень доверия примите равным 0,99. Исходные данные:

Категория ответа	Количество респондентов в выборке
Практически каждый день	870
3-5 дней в неделю	135
1-2 дня в неделю	210
Реже, чем раз в неделю	90
Не смотрю телевизор	195

Задача 6. Пусть в некотором исследовании стремились понять, есть ли взаимосвязь между цветом машины, которую инспектор ГИБДД останавливает для проверки документов, и тем, кто управляет автомобилем – мужчина или женщина. В результате выборочного исследования построена таблица перекрестной классификации. При $\alpha = 0.05$ проверьте соответствующую гипотезу исследования. О чем говорят значения стандартизированных остатков? Какой вывод можно сделать? В ячейках исходной таблицы указано количество наблюдений:

Автомобилист (кто за рулем)	Инспектор останавливает автомобиль для проверки документов:		
	Автомобиль красного цвета	Автомобиль белого цвета	Автомобиль черного цвета
Мужчина	22	24	42
Женщина	30	26	16

Задача 7. Среди офисных служащих некоторой корпорации было проведено исследование о том, придерживаются ли они здорового образа жизни – питание, сон, пребывание на свежем воздухе. Так, в случайной выборке мужчин объемом 54 человек оказалось, что средний километраж прогулок на свежем воздухе в течение недели составляет 12 км, а в случайной выборке женщин объемом 50 человек этот показатель оказался равен 11.1 км. Можно ли при имеющихся данных считать различия в средних показателях статистически значимыми (достоверными), если известно, что оценка стандартного отклонения в выборке мужчин составила 3.1, а в выборке женщин – 1.8? Примите уровень доверительной вероятности 0.95. Конкурирующую статистическую гипотезу сформулируйте как *ненаправленную*.

Задача 8. При проверке гипотезы о равенстве среднего определенному числу (константе) против *направленной* конкурирующей гипотезы в случайной выборке из 10 наблюдений критерий Стьюдента для имеющихся данных принял значение 2.99. Значим ли полученный результат на 5%-м уровне? Значим ли полученный результат на 1%-м уровне? Укажите ответ и дайте обоснование (почему).

Экзаменационная работа (математическая статистика): примеры задач

Задача 1. Планируется опрос жителей в крупных городах России с целью узнать об их планах на предстоящие зимние праздники – какая часть населения никуда не поедет, а останется праздновать Новый год дома, в России. Какой объем выборки потребуется исследователям, чтобы обеспечить выполнение следующих требований: уровень надежности выводов будущего исследования должен составить 90%, а доля жителей, которые останутся на Новый год дома, в России, должна быть оценена с точностью $\pm 5\%$? По предварительным оценкам экспертов, стандартное отклонение составляет 40%.

Задача 2. Верно ли, что среднее значение переменной всегда больше медианы? а)

Неверно.

б) Верно, но только для нормально распределенных переменных.

в) Верно для всех переменных.

г) Верно для категориальных переменных (измеренных в номинальных или порядковых шкалах).

Задача 3. Значение коэффициента вариации, равное 0.15 говорит:

а) О гарантированном наличии в выборке статистических «выбросов».

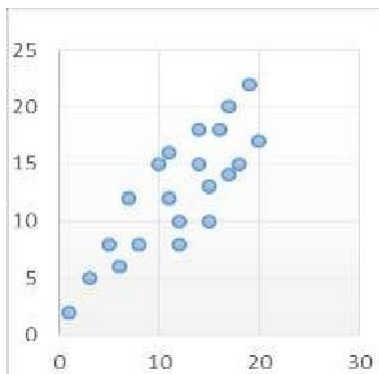
б) О низкой однородности выборки.

в) О высокой однородности выборки.

г) О том, что лучше использовать среднее, а не медиану, для понимания центральной тенденции.

д) Недостаточно информации для ответа на поставленный вопрос **Задача**

4. Рассмотрите следующую диаграмму рассеяния:



Что из перечисленного является корректной оценкой коэффициента корреляции между X и Y?

- а) -0.80 , б) -0.20 , в) -0.50 , г) 0 , д) 0.20 , е) 0.50 , ж) 0.80

Задача 5. Для некоторого набора данных в выборке имеются статистические выбросы. Какое сочетание статистических параметров, указывающих на центральную тенденцию и вариацию предпочтительно выбрать в этой ситуации?

- а) Медиана и дисперсия, б) Среднее и дисперсия, в) Медиана и размах вариации
г) Среднее и размах вариации, д) Медиана и квартильный размах

Задача 6. Продолжительность жизни людей в некоторой стране имеет стандартное отклонение, равное 15 лет. Чему равно среднее значение продолжительности жизни, если известно, что 30% населения этой страны живут дольше 74 лет? Ответьте на поставленный вопрос, исходя из предположения, что продолжительность жизни людей, живущих в этой стране, – случайная величина, имеющая нормальное распределение с указанными параметрами.

Задача 7. Диаграмма рассеяния в основном используется для визуального (наглядного) представления связи между:

- а) Двумя категориальными (измеренными в номинальных шкалах) переменными.
б) Двумя количественными переменными.
в) Количественной и категориальной (в номинальной шкале) переменными.
г) Двумя порядковыми переменными.
д) Двумя переменными, независимо от типа шкал, в которых они измерены

Задача 8. Дана выборка X: 5, 7, 5, 3, 5, 2, 4, 0, 9, 0. Найдите медиану, нижнюю и верхнюю квартили. Зафиксируйте диапазон характерных значений. Есть ли в этой выборке нетипичные (нехарактерные) значения наблюдений?

Задача 9. За месяц до президентских выборов в стране был проведен телефонный социологический опрос общественного мнения (случайная выборка телефонных номеров –

стационарных и мобильных). В ходе исследования было установлено, что 360 человек в возрасте 18+ из 400 опрошенных уже решили, за кого будут голосовать. Насколько мы можем быть уверены в том, что к этому моменту времени истинная процентная доля избирателей, определившихся с политическим выбором, находится в интервале между 87,5% и 92,5%?

Задача 10. Компания предложила двум экспертам оценить по 15-балльной шкале эстетическое восприятие нового журнала для автомобилистов, выполненного в 10-ти разных обложках. Более высокие оценки соответствуют более приятному эстетическому восприятию журнала. В какой мере оценки (мнения) двух экспертов согласуются между собой? Вычислите коэффициент ранговой корреляции Спирмена и дайте его содержательную интерпретацию.

Обложка журнала	Эксперт 1, оценки:	Эксперт 2, оценки:
А	8	6
Б	5	9
В	15	13
Г	6	7
Д	7	5
Е	4	4
Ж	10	8
З	9	10
И	2	1
К	13	13

Задача 11. Что такое p-value? Дайте развернутый содержательный ответ.

Задача 12. Какова вероятность ошибки 2-го рода в том случае, когда проверка нуль-гипотезы осуществляется при 1% уровне значимости ($\alpha = 0,01$)

а) Недостаточно информации для ответа на вопрос

б) 0.001, в) 0.01, г) 0.05, д) 0.10, е) 0.95, ж) 0.99

Задача 13. Критерий хи-квадрат может принимать следующие значения:

а) Любое положительное число или ноль. б) Любое число.

в) Любое число от -1 до 1. г) Любое число от 0 до 1. д) Ничего из перечисленного

Задача 14. В одном кулинарном журнале для домашних хозяек была приведена информация о том, что среди опрошенных 200 женщин (выборка случайная) 40% сказали, что на завтрак для семьи всегда готовят кашу. Приняв уровень надежности статистического вывода равным 90%, укажите границы доверительного интервала для истинной процентной доли домашних хозяек, которые на завтрак для семьи всегда готовят кашу.

Задача 15. Служба занятости города М* решила выяснить, является ли уверенное знание английского языка важным условием для получения вакансии «менеджер-администратор». Руководитель службы занятости организовал случайный отбор 450 вакансий работодателя с пометкой «менеджер-администратор» и нашел, что для 198 требуется уверенное знание

английского языка. На основании полученных данных можем ли мы поддержать гипотезу о том, что половина рабочих мест на этом рынке труда предусматривает уверенное знание английского языка? Уровень доверия примите равным 0,99. Конкурирующую гипотезу сформулируйте *направленной*, исходя из полученных данных исследования.

Задача 16. В ходе некоторого исследования (выборка случайная), посвященного изучению бюджета времени выпускников университета, было зафиксировано: сколько времени в неделю они затрачивали на спорт в студенческие годы и в первый год работы после окончания вуза. Можно ли говорить о том, что бюджет времени на занятия спортом в студенческие годы в среднем выше, чем после окончания университета? Проверку соответствующей статистической гипотезы осуществите на уровне доверительной вероятности 0,90. Конкурирующая гипотеза *направленная* (исходя из данных). Данные исследования таковы:

Респондент	Время на спорт (количество часов в неделю)	
	В студенческие годы	В первый год работы после окончания вуза
И.	2	1
Ж.	0	0
П.	5	3
Л.	3	4
А.	2	0
О.	7	5
Ю.	1	0

Задача 17. Объем двух зависимых (связных) выборок равен 20 наблюдений. Каким будет аналитическое заключение при 10%-м уровне значимости (α), если при проверке гипотезы H_0 : «средние равны» против альтернативы H_1 : $\mu_1 < \mu_2$ (второе среднее достоверно больше первого) было получено P-value (2-стороннее), равное 0,12? Сформулируйте ответ, исходя из того, являются ли наблюдаемые различия в тестируемых выборках статистически достоверными или нет, и почему.

Задача 18. В ходе исследования читательских предпочтений горожан относительно трех жанров – исторического романа, детектива и научной фантастики – были зафиксированы ответы мужчин и женщин. Рассмотрите таблицу сопряженности. При $\alpha = 0,05$ выясните, можно ли утверждать, что рассмотренные признаки связаны. (В ячейках указано количество респондентов).

	Исторический роман	Детектив	Научная фантастика
Мужчины	42	18	30
Женщины	30	24	24

Задача 19. Автоматическая фильтрация почтового «мусора» (спам) основана на механизме распознавания писем как легитимных и нежелательных, но в зависимости от настроек системы она не всегда способна к нормальной классификации почты. Статистическая нуль-гипотеза, сформулированная в отношении всех писем, направленных получателю, такова: «*Письмо является легитимным*». Каковы в таком случае последствия ошибок 1-го и 2-го рода в контексте сформулированной нуль-гипотезы?

- а) Ошибка 1-го рода: *Легитимное e-mail сообщение доставлено своему получателю*; Ошибка 2-го рода: *Нежелательное e-mail сообщение заблокировано*.
- б) Ошибка 1-го рода: *Легитимное e-mail сообщение классифицировано как спам*; Ошибка 2-го рода: *Нежелательное e-mail сообщение классифицировано как легитимное*.
- в) Ошибка 1-го рода: *Нежелательное e-mail сообщение классифицировано как легитимное*; Ошибка 2-го рода: *Нежелательное e-mail сообщение заблокировано*.
- г) Ошибка 1-го рода: *Нежелательное e-mail сообщение классифицировано как легитимное*; Ошибка 2-го рода: *Легитимное e-mail сообщение классифицировано как спам*.
- д) Ошибка 1-го рода: *Нежелательное e-mail сообщение заблокировано*; Ошибка 2-го рода: *Легитимное e-mail сообщение классифицировано как спам*.

Задача 20. Тренер хочет узнать, есть ли разница в прибавке веса у лыжников-биатлонистов, следующих специальной назначенной диете. В ходе обследования спортсмены по жребии (т.е. случайным образом) были разделены на три группы: каждая группа следовала определенной диете 6 календарных недель. Спустя время прибавка в весе (в кг) указана в таблице. На уровне значимости 0,05 определите, может ли исследователь утверждать, что наблюдается статистически достоверная разница в результате в зависимости от применённой диеты.

Трек 1: Диета «А»	Трек 2: Диета «В»	Трек 3: Диета «С»
3, 6, 7, 4	7, 10, 10, 8, 7, 6	7, 3, 2, 4