

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«ДАГЕСТАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет математики и компьютерных наук
Кафедра прикладной математики

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по дисциплине
«Технологии обработки статистических данных»

Кафедра прикладной математики
факультета математики и компьютерных наук

Образовательная программа бакалавриата
01.03.05 – Статистика

Направленность (профиль) программы
Анализ больших данных

Форма обучения
Очная

Статус дисциплины:
входит в часть ОПОП, формируемую участниками образовательных отношений;
дисциплина по выбору


Махачкала, 2023

Фонд оценочных средств по дисциплине «Технологии обработки статистических данных» составлена в 2023 году в соответствии с требованиями ФГОС ВО бакалавриата по направлению подготовки 01.03.05 - статистика от 14.08.2020 г. № 1032

Разработчики:

кафедра прикладной математики, Лугуева А.С. к.ф.-м. н., доцент;

Фонд оценочных средств по дисциплине «Технологии обработки статистических данных» одобрен:
на заседании кафедры Прикладной математики от «20» января 2023 г.,
протокол № 5

Зав. кафедрой  Кадиев Р.И.

на заседании Методической комиссии факультета МикН от
«25» января 2023г., протокол №4.

Председатель  Ризаев М.К.

Фонд оценочных средств «Технологии обработки статистических данных»
согласован с учебно-методическим управлением

«20» февраля 2023 г 

**1. ПАСПОРТ
ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ
по дисциплине**

«Технологии обработки статистических данных»

1.1. Основные сведения о дисциплине

Общая трудоемкость дисциплины составляет 3 зачетные единицы (108 академических часов).

Вид работы	Трудоемкость, академических часов		
	б семестр		всего
Общая трудоёмкость	108		108
Контактная работа:	32		32
Лекции (Л)	16		16
Практические занятия (ПЗ)	-		-
Лабораторные занятия (ЛЗ)	16		16
Консультации			
Промежуточная аттестация (зачет, экзамен)	экзамен		экзамен
Самостоятельная работа			
1. работа с лекционным материалом, с учебной литературой	8		8
2. опережающая самостоятельная работа (изучение нового материала до его изложения на занятиях)	8		8
3. выполнение домашних заданий	8		8
4. подготовка к лабораторным работам, к практическим занятиям	8		8
5. подготовка к коллоквиуму			
6. подготовка к контрольным работам	8		8
7. подготовка к экзамену	36		36

1.2. Требования к результатам обучения по дисциплине, формы их контроля и виды оценочных средств

*ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ
по дисциплине «Технологии обработки статистических данных»*

№ п/п	Контролируемые модули, разделы (темы) дисциплины	Код контролируемой компетенции (или её части)	Оценочные средства		Способ контроля
			наименование	№№ заданий	
1	Модуль 1. Одномерные статистические модели при обработке и анализе информации	ОПК-3 УК-1	Вопросы для собеседования	1-10	устно
		ОПК-3 УК-1	Контрольные работы	1-3	письменно

		ОПК-3 УК-1	Лабораторные работы		письменно
2	Модуль 2. Статистические исследования зависимостей	ОПК-3 УК-1	Вопросы для собеседования	11-20	устно
		ОПК-3 УК-1	Контрольные работы	2	письменно
		ОПК-3 УК-1	Лабораторные работы	4-6	письменно
	Модуль 3. Экзамен	ОПК-3 УК-1	Вопросы для подготовки к экзамену		Устно- письменно

1.3. Показатели и критерии определения уровня сформированности компетенций

№ п/п	Код компет енции	Уровни сформированности компетенции			
		Недостаточный	Удовлетворительный (достаточный)	Базовый	Повышенный
		Отсутствие признаков удовлетворительного уровня	Знать: Уметь: Владеть:	Знать: Уметь: Владеть:	Знать: Уметь: Владеть:
1	ОПК-3	Не знает на достаточном уровне общую методику статистического исследования и способы количественной формализации объекта наблюдений Не умеет на достаточном уровне применять математический и эконометрический инструментарий для анализа количественных данных, в том числе с применением информационных систем и технологий Не владеет на достаточном уровне навыками выбора и применения	Знает на достаточном уровне общую методику статистического исследования и способы количественной формализации объекта наблюдений Умеет на достаточном уровне применять математический и эконометрический инструментарий для анализа количественных данных, в том числе с применением информационных систем и технологий Владеет на достаточном уровне навыками выбора и применения	Знает на хорошем уровне общую методику статистического исследования и способы количественной формализации объекта наблюдений Умеет на хорошем уровне применять математический и эконометрический инструментарий для анализа количественных данных, в том числе с применением информационных систем и технологий Владеет на	Знает в совершенстве общую методику статистического исследования и способы количественной формализации объекта наблюдений Умеет в совершенстве применять математический и эконометрический инструментарий для анализа количественных данных, в том числе с применением информационных систем и технологий Владеет в совершенстве навыками выбора и применения

		инструментальных средств для обработки количественных данных, навыками интерпретации результатов и формулирования выводов и рекомендаций для подготовки аналитических материалов.	инструментальных средств для обработки количественных данных, навыками интерпретации результатов и формулирования выводов и рекомендаций для подготовки аналитических материалов.	хорошем уровне навыками выбора и применения инструментальных средств для обработки количественных данных, навыками интерпретации результатов и формулирования выводов и рекомендаций для подготовки аналитических материалов.	инструментальных средств для обработки количественных данных, навыками интерпретации результатов и формулирования выводов и рекомендаций для подготовки аналитических материалов.
2	УК-1	Не знает на достаточном уровне структуру задач в области математики, теоретической механики и физики, а также базовые составляющие таких задач. Не умеет на достаточном уровне анализировать постановку данной математической задачи, необходимость и (или) достаточность информации для ее решения. Не владеет на достаточном уровне навыками сбора, отбора и обобщения научной информации в области математических дисциплин	Знает на достаточном уровне структуру задач в области математики, теоретической механики и физики, а также базовые составляющие таких задач. Умеет на достаточном уровне анализировать постановку данной математической задачи, необходимость и (или) достаточность информации для ее решения. Владеет на достаточном уровне навыками сбора, отбора и обобщения научной информации в области математических дисциплин	Знает на хорошем уровне структуру задач в области математики, теоретической механики и физики, а также базовые составляющие таких задач. Умеет на хорошем уровне анализировать постановку данной математической задачи, необходимость и (или) достаточность информации для ее решения. Владеет на хорошем уровне навыками сбора, отбора и обобщения научной информации в области математических дисциплин	Знает в совершенстве структуру задач в области математики, теоретической механики и физики, а также базовые составляющие таких задач. Умеет в совершенстве анализировать постановку данной математической задачи, необходимость и (или) достаточность информации для ее решения. Владеет в совершенстве навыками сбора, отбора и обобщения научной информации в области математических дисциплин

2. КОНТРОЛЬНЫЕ ЗАДАНИЯ И ИНЫЕ МАТЕРИАЛЫ ОЦЕНКИ знаний, умений, навыков и (или) опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения дисциплины «Технологии обработки статистических данных»

Контрольные работы Контрольная работа 1

1. Даны две независимые выборки объема 11 и 14, извлеченные из нормальных совокупностей X , Y . Известны также исправленные дисперсии, равные соответственно 0,75 и 0,4. Необходимо проверить нулевую гипотезу о равенстве генеральных дисперсий при уровне значимости $\gamma=0,05$. Нулевая гипотеза записывается следующим образом:

$$H_0 : D(X) = D(Y).$$

В качестве конкурирующей гипотезы рассмотрим следующую: Для

$$H_1 : D(X) > D(Y).$$

решения задачи

1.1 Вычислите отношение большей исправленной дисперсии к меньшей

1.2. По таблице для уровня значимости 0,05 и числам степеней свободы равным 10 ($11 - 1 = 10$) и 13 ($14 - 1 = 13$) соответственно найдите критическую точку

1.3. Сделайте вывод о проверке гипотезы

а) есть основания отвергнуть гипотезу о равенстве генеральных дисперсий

б) нет оснований отвергнуть гипотезу о равенстве генеральных дисперсий

2. Произведено 5 независимых наблюдений над случайной величиной $X \sim N(a, 20)$. Результаты наблюдений таковы: $x_1 = -25$, $x_2 = 34$, $x_3 = -20$, $x_4 = 10$, $x_5 = 21$. Построить для выборочного среднего 95%-й доверительный интервал.

а) $(-13,5; 21,5)$,

б) $(-10,5; 10,5)$

3. В таблице приведены сгруппированные данные измерений роста у 50 случайно отобранных студентов.

Рост студентов, см	162–166	166–170	170–174	174–178	178–182	182–186
Число студентов	3	7	15	13	11	1

Построить доверительный интервал для среднего роста студентов с надёжностью 0,9

а) $P(172,86 < M(X) < 175,14) = 0,9$,

б) другой ответ

4. С целью определения средней суммы вкладов в сберегательной кассе, имеющей 7200 вкладчиков, произведено выборочное обследование (бесповторный отбор) 111 вкладчиков, которое дало следующие результаты:

Сумма вклада, тыс. руб.	10 – 30	30 – 50	50 – 70	70 – 90	90 – 110	110 – 130
Число вкладов	1	3	10	30	60	7

Пользуясь этими данными, найдите доверительные границы для генерального среднего, которые можно было бы гарантировать с вероятностью 0,95.

а) $(86.5274, 93.2926)$,

б) другой

Контрольная работа 2

1. Определите численность выборки при обследовании остатков на расчетных счетах у клиентов коммерческого банка, чтобы с вероятностью 0,683 ошибка репрезентативности не превышала 5 тыс. руб., если $\sigma_{ген} = 120$ тыс. руб.

а) 576,

б) другой ответ

2. По случайной выборке измерений роста 20 студентов первого курса университета вычислена несмещенная оценка генеральной дисперсии $s^2 = 0.002$. Найдите 95%-ный доверительный интервал для среднего квадратического отклонения роста всех студентов первого курса университета, если распределение роста нормально.

а) $(0.0342, 0.0657)$,

б) другой ответ

3. Для определения процента вкладов, не превышающих 100000 руб., произведена повторная выборка 900 лицевых счетов. Среди них оказалось 30% вкладов не более 100000 каждый. С какой доверительной вероятностью можно утверждать, что процент таких вкладов в данной кассе будет отличаться от найденного не более чем на 2%?

а) 0.8098,

б) другой ответ

4. За последние 5 лет годовой рост цены актива A составлял в среднем 20% со средним квадратическим отклонением (исправленным) 5%. Построить доверительный интервал с вероятностью 95% для цены актива в конце следующего года, если и начале года она равна 100 ден. ед.

а) цена актива в следующем году составит от 105 до 135 ден. ед., б) другой ответ

Критерии оценки:

- оценка «отлично» выставляется студенту, если верно и правильно выполнено 90%-100% заданий;
- оценка «хорошо» выставляется студенту, если верно и правильно выполнено 70%-80% заданий;
- оценка «удовлетворительно» выставляется студенту, если верно и правильно решено 50%-60% заданий, возможны некоторые исправления при решении;
- оценка «неудовлетворительно» выставляется студенту, если верно выполнено менее 50% заданий;

Вопросы для коллоквиумов, собеседования

Модуль 1. Одномерные статистические модели при обработке и анализе информации

1. Основные требования к выборочной совокупности
2. Инструментарий информационных технологий в статистических исследованиях.
3. Одномерные статистические модели. Условия применения. Генеральная и выборочные совокупности. Основные требования к выборочной совокупности (массовость, однородность, случайность, независимость).
4. Статистическая сводка. Статистическая группировка. Система группировок. Статистические таблицы.
5. Составление выборок, подготовка данных для статистической обработки.
6. Абсолютные и относительные величины.
7. Графический метод. Понятия и правила построения.
8. Расчет статистических параметров распределения анализируемых параметров. Построение и анализ гистограмм, проверка статистических гипотез распределения с использованием инструментов Microsoft Excel.
9. Сущность и задачи средних величин.
10. Расчет ошибки выборочной средней с использованием электронных таблиц.

Модуль 2. Статистические исследования зависимостей

11. Показатели вариации.
12. Расчет показателей вариации инструментарием Microsoft Excel.
13. Ряды динамики и их анализ
14. Ряды динамики. Изучение характеристик рядов динамики средствами Microsoft Excel. Понятие корреляционной зависимости. Характеристика корреляционной связи по тесноте и форме.
15. Изучение корреляционных зависимостей табличным, графическим и аналитическими методами.
16. Статистическое изучение взаимосвязей социальных явлений.
17. Построение линейного уравнения регрессии для эмпирических данных в Microsoft Excel.
18. Статистическое измерение связи. Виды и формы связей.
19. Основы выборочного метода. Основные виды выборки, способы отбора.
20. Статистическое изучение взаимосвязей в социальных явлениях. Расчет критерия Хи-квадрат в Microsoft Excel.
21. Статистическая обработка малых выборок.

Критерии оценки:

- оценка «отлично» выставляется студенту, если изложение полученных знаний в устной форме полное, в системе, в соответствии с требованиями учебной программы; допускаются единичные несущественные ошибки, самостоятельно исправляемые учащимися;
- оценка «хорошо» выставляется студенту, если изложение полученных знаний в устной форме полное, в системе, в соответствии с требованиями учебной программы; допускаются, отдельные несущественные ошибки, исправляемые учащимися после указания преподавателя на них;
- оценка «удовлетворительно» выставляется студенту, если изложение полученных знаний неполное, однако это не препятствует усвоению последующего программного материала; допускаются отдельные существенные ошибки, исправляемые с помощью преподавателя;
- оценка «неудовлетворительно» выставляется студенту, если изложение учебного материала неполное, бессистемное, что препятствует усвоению последующей учебной информации; существенные ошибки, не исправляемые даже с помощью преподавателя;

Комплект тестовых заданий для контроля

А. Примерный перечень заданий для подготовки к тестированию

1. Половина всех значений в исследуемом ряду будет меньше медианы, а другая половина – больше ее
 - а) да
 - б) нет
2. Мода – наиболее часто встречающееся значение
 - а) да
 - б) нет
3. Стандартное отклонение определяется следующим образом: вычисляется отклонение между каждым значением и средним, затем отклонения возводятся в квадрат и суммируются, сумма квадратов отклонений делится на счет минус один. В результате получается дисперсия выборки. Квадратный корень из дисперсии является стандартным отклонением.
 - а) да
 - б) нет
4. Эксцесс является показателем симметричности распределений.
 - а) да,
 - б) нет
5. Среднее зависит от экстремальных значений данных. Экстремальные значения в положительной части оси увеличивают среднее. Экстремальные значения в отрицательном направлении уменьшают среднее, и среднее становится меньше медианы
 - а) да
 - б) нет
6. Коэффициент корреляции – это общая характеристика двумерных данных, отражающая существующую между ними

а) нелинейную зависимость, б)

линейную зависимость

7. Простая линейная регрессия используется для определения линейного уравнения, описывающего

а) линейную зависимость между двумя переменными, б)

нелинейную зависимость между двумя переменными

8. Уровень надежности равняется половине длины 95% доверительного интервала для среднего

а) да,

б) нет

9. Статистической гипотезой называется предположение о свойстве генеральной совокупности, которое можно проверить, опираясь на данные выборки

а) да,

б) нет

10. Для проверки статистических гипотез используются параметрические и непараметрические методы. Параметрические методы служат для проверки гипотез о неизвестных параметрах генеральной

совокупности, когда закон распределения случайной величины известен.

Непараметрические методы применяются в тех случаях, когда закон распределения случайной величины неизвестен, или когда условия применения параметрических методов не выполняются.

а) да,

б) нет

Б. Примерный перечень заданий для решения стандартных задач

Дана выборка объемом $n=21$:

21, 23, 24, 20, 21, 23, 24, 20, 19, 25, 26, 27, 21, 22, 25, 26, 27, 28, 21, 21, 22

Необходимо определить числовые характеристики выборки.

1. Среднее

а) 23.14,

б) 20.34

2. Медиан

а

а) 22,

б) 23

3. Мода

а) 21,

б) 23

4. Размах выборки

а) 9,

б) 10

5. Стандартное отклонение а)

2.65,

б) 2.34

6. Стандартная ошибка а)

0.58,

б) 0.71

7. Уровень надежности
а) 1.21,
б) 0.8
8. Эксцесс
а) -1.11,
б) -2.1
9. Дисперсия выборки
а) 7.45,
б) 7.03
10. Ассиметричность
а) 0.31,
б) 0.21

Критерии оценки:

- оценка «отлично» выставляется студенту, если верно и правильно выполнено 90%-100% заданий;
- оценка «хорошо» выставляется студенту, если верно и правильно выполнено 70%-80% заданий;
- оценка «удовлетворительно» выставляется студенту, если верно и правильно решено 50%-60% заданий, возможны некоторые исправления при решении;
- оценка «неудовлетворительно» выставляется студенту, если верно выполнено менее 50% заданий;

Темы лабораторных работ

Модуль 1. Одномерные статистические модели при обработке и анализе информации

Лабораторная работа 1. Одномерные статистические модели. Условия применения. Генеральная и выборочные совокупности.

Инструментарий информационных технологий в статистических исследованиях. Одномерные статистические модели. Условия применения. Генеральная и выборочные совокупности. Основные требования к выборочной совокупности (массовость, однородность, случайность, независимость). Статистическая сводка. Статистическая группировка. Система группировок. Статистические таблицы.

Лабораторная работа 2. Составление выборок, подготовка данных для статистической обработки.

Составление выборок, подготовка данных для статистической обработки. Абсолютные и относительные величины. Графический метод. Понятия и правила построения.

Лабораторная работа 3. Расчет статистических параметров распределения анализируемых параметров

Расчет статистических параметров распределения анализируемых параметров. Построение и анализ гистограмм, проверка статистических гипотез распределения с использованием инструментов Microsoft Excel. Сущность и задачи средних величин. Расчет ошибки выборочной средней с использованием электронных таблиц.

Модуль 2. Статистические исследования зависимостей

Лабораторная работа 4. Расчет показателей вариации инструментарием Microsoft Excel.

Показатели вариации. Расчет показателей вариации инструментарием Microsoft Excel. Ряды динамики и их анализ. Ряды динамики. Изучение характеристик рядов динамики средствами Microsoft Excel. Понятие корреляционной зависимости. Характеристика корреляционной связи по тесноте и форме.

Лабораторная работа 5. Изучение корреляционных зависимостей табличным, графическим и аналитическими методами.

Изучение корреляционных зависимостей табличным, графическим и аналитическими методами. Статистическое изучение взаимосвязей социальных явлений. Построение линейного уравнения регрессии для эмпирических данных в Microsoft Excel.

Лабораторная работа 6. Статистическое измерение связи. Виды и формы связей.

Статистическое измерение связи. Виды и формы связей. Основы выборочного метода. Основные виды выборки, способы отбора. Статистическое изучение взаимосвязей в социальных явлениях. Расчет критерия Хи-квадрат в Microsoft Excel. Статистическая обработка малых выборок.

Критерии оценки:

- оценка «отлично» выставляется студенту, если выполнены все задания лабораторной работы, составлен отчет по работе, согласно требованиям;
- оценка «хорошо» выставляется студенту, если выполнены почти все задания, за исключением отдельных пунктов, лабораторной работы, составлен отчет по работе, согласно требованиям;

- оценка «удовлетворительно» выставляется студенту, если выполнены больше половины заданий лабораторной работы, составлен отчет по работе, согласно требованиям;
- оценка «неудовлетворительно» выставляется студенту, если выполнены меньше половины заданий лабораторной работы и не составлен отчет по работе, согласно требованиям;

Темы эссе (рефератов, докладов, сообщений)

1. Статистическая обработка опытных данных
2. Статистическая обработка экспериментальных данных
3. Статистическая отчетность
4. Статистическая отчетность о деятельности транспортных предприятий
5. Статистическая отчетность по основным фондам и порядок её оформления
6. Статистическая оценка вероятности наступления неблагоприятных событий
7. Статистическая оценка влияния факторов и прогнозирование показателей уровня жизни населения России
8. Статистическая оценка динамики тарифов на электроэнергию
9. Статистическая оценка качества и конкурентоспособности продукции
10. Статистическая оценка рынка труда. Критическая зона рынка труда

Реферат оценивается следующим образом:

- соответствие содержания теме- 4 балла;
- глубина проработки материала, 3 балла;
- грамотность и полнота использования источников, 1 балл;
- соответствие оформления реферата требованиям, 2 балла;
- доклад, 5 баллов;
- умение вести дискуссию и ответы на вопросы, 5 баллов.

Максимальное количество баллов: 20.

Критерии оценки:

- оценка «отлично» выставляется студенту, если набрал 19-20 баллов;
- оценка «хорошо» выставляется студенту, если набрал 15-18 баллов;
- оценка «удовлетворительно» выставляется студенту, если набрал 10-14 баллов;
- оценка «неудовлетворительно» выставляется студенту, если набрал менее 10 баллов;

Вопросы к экзамену

1. Ошибки измерения. Классификация ошибок.

2. Распределение случайных ошибок измерения. Вероятностная модель.
3. Основные понятия теории вероятностей
4. Классическое определение вероятности
5. Основные свойства вероятностей
6. Случайные величины. Законы распределения.
7. Дискретные случайные величины. Ряд распределения.
8. Числовые характеристики дискретных случайных величин.
9. Математическое ожидание и его свойства.
10. Дисперсия дискретной случайной величины.
11. Функция распределения.
12. Непрерывные случайные величины.
13. Плотность распределения вероятностей.
14. Числовые характеристики непрерывных случайных величин.
15. Математическое ожидание, дисперсия.
16. Моменты случайных величин. Мода и медиана распределения.
17. Асимметрия. Эксцесс.
18. Квантили распределения.
19. Распределения, используемые в анализе данных (нормальное, равномерное, показательное, хи-квадрат, Стьюдента, Фишера).
20. Генеральная совокупность и выборка.
21. Выборочные характеристики: вариационный ряд, эмпирическая функция распределения.
22. Выборочные моменты: среднее, выборочная дисперсия, квантили, выборочная корреляция.
23. Наглядные методы описательной статистики: группировка, гистограмма и полигон частот.
24. Понятие статистической оценки
25. Свойства оценок: состоятельность несмещенность, эффективность.
26. Интервальные оценки: понятие доверительного интервала.
27. Доверительный интервал для среднего.
28. Доверительный интервал для дисперсии.
29. Доверительные интервалы для среднего и дисперсии.
30. Статистические гипотезы. Статистические критерии.
31. Основной принцип проверки статистических гипотез.
32. Построение критических областей.
33. Проверка гипотезы о виде функции распределения (критерии согласия).
34. Критерий согласия хи- квадрат Пирсона.
35. Проверка гипотезы о математическом ожидании нормальногораспределения.
36. Проверка гипотез однородности и симметрии распределения.
37. Критерии однородности хи-квадрат. Ранговые критерии.

38. Непараметрические критерии проверки однородности.
39. Линейный регрессионный анализ. Метод наименьших квадратов.
40. Алгоритмы формирования последовательности случайных чисел.
41. Моделирование непрерывных распределений.
42. Моделирование дискретных случайных величин и случайных событий

Критерии оценки:

«отлично» - теоретическое содержание курса освоено полностью, без пробелов, умения сформированы, все предусмотренные программой учебные задания выполнены, качество их выполнения оценено высоко;

«хорошо» - теоретическое содержание курса освоено полностью, без пробелов, некоторые умения сформированы недостаточно, все предусмотренные программой учебные задания выполнены, некоторые виды заданий выполнены с ошибками;

«удовлетворительно» - теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые умения работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий содержат ошибки;

«неудовлетворительно» - теоретическое содержание курса не освоено, необходимые умения не сформированы, выполненные учебные задания содержат грубые ошибки.

Рекомендуемые границы оценок:

«отлично» - не менее 86% правильных ответов,

«хорошо» - 66-85% правильных ответов,

«удовлетворительно» - 51-65% правильных ответов,

«неудовлетворительно» - менее 50% правильных ответов.