

УТВЕРЖДЕНО
Проректор по учебной работе
А. А. Воронов
15 января 2026 г.

ПРОГРАММА

по дисциплине: Математическая статистика
по направлению подготовки: 09.03.01 «Информатика и вычислительная техника»
физтех-школа: ПИИ ФАЛТ
кафедра: высшей математики
курс: 2
семестр: 4

лекции — 30 часов
практические (семинарские)
занятия — 30 часов
лабораторные занятия — нет

Диф. зачёт — 4 семестр

ВСЕГО АУДИТОРНЫХ ЧАСОВ — 60

Самостоятельная работа:
теор. курс — 75 часов

Программу составил

к. ф.-м. н., доцент С. Д. Животов

Программа принята на заседании кафедры
высшей математики 23 октября 2025 г.

Заведующий кафедрой
д. ф.-м. н., профессор

Г. Е. Иванов

1. Основные распределения: нормальное, χ^2 , Стьюдента, Фишера. Квантили.
2. Лемма Фишера. Теорема Фишера
3. Выборка и генеральная совокупность (вероятностная модель). Обработка выборки – размах, медиана, мода, коэффициент асимметрии, вариационный ряд. Законы распределения элементов вариационного ряда. Bootstrap.
4. Требования к оценкам. Достаточное условие состоятельности. Единственность оптимальной оценки.
5. Неравенство Крамера–Рао и следствия из него. Многомерный аналог неравенства. Критерий Бхаттачария и его многомерный аналог. Эффективность оценки.
6. Функция распределения и гистограмма.
7. Частота появления события как несмещенная, состоятельная и эффективная оценка вероятности. Оценка начальных моментов. Оценки математического ожидания и дисперсии.
8. Методы получения оценок: метод моментов, метод максимального правдоподобия. Асимптотические свойства оценок, полученных методом моментов. Асимптотические свойства оценок, полученных методом максимального правдоподобия.
9. Доверительные интервалы и методы их построения (точный, асимптотический, численный). Асимптотические доверительные интервалы для функции распределения и моментов. Асимптотический доверительный интервал для вероятности появления события.
10. Точные доверительные интервалы для параметров нормального распределения.
11. Проверка статистических гипотез. Ошибки первого и второго рода. Мощность критерия.
12. Критерий χ^2 для проверки гипотезы о законе распределения.
13. Критерий χ^2 для проверки гипотез однородности и независимости.
14. Критерий Колмогорова для проверки гипотезы о законе распределения, критерий Смирнова для проверки гипотезы однородности, критерий проверки гипотезы случайности.
15. Проверка параметрических гипотез. Теорема Неймана–Пирсона для проверки простой гипотезы против простой альтернативы.
16. Критерий отношения правдоподобия для проверки сложных гипотез. Асимптотический метод проверки сложных гипотез. Проверка гипотез о вероятности появления события.
17. Проверка гипотез о параметрах нормального распределения. Проверка гипотез о равенстве средних и равенстве дисперсий.

18. Регрессионный анализ. Несмещенность, состоятельность и эффективность оценки коэффициентов регрессии. Оценка остаточной дисперсии. Закон распределения коэффициентов регрессии. Закон распределения остаточной суммы квадратов.
19. Статистический анализ коэффициентов регрессии: проверка значимости коэффициентов, проверка простейших линейных гипотез. Проверка значимости всей регрессии, сравнение двух уравнений регрессии. Коэффициент детерминации.
20. Проверка адекватности модели. Кросс-проверка. Проверка основных предположений регрессионного анализа. Доверительный интервал для прогноза. Методы регуляризации.

Список литературы

1. *Ивченко Г. И. Медведев Ю. И.* Введение в математическую статистику. — Москва : ЛКИ, 2010, 2014, 2015.
2. *Лагутин М. Б.* Наглядная математическая статистика. — 9-е изд. — Москва : Лаборатория знаний, 2023.
3. *Чернова Н. И.* Математическая статистика. Учебное пособие. — Новосибирск, 2014. https://tvims.nsu.ru/chernova/ms/ms_nsu14.pdf
4. *Боровков А. А.* Математическая статистика. — 3-е изд. — Москва : Физматлит, 2007.
5. *Кобзарь А. И.* Прикладная математическая статистика: для инженеров и научных работников. — 2-е изд., испр. — Москва : Физматлит, 2012.

ПЕРВОЕ ЗАДАНИЕ

(срок сдачи 13–18 апреля)

1. Случайная величина распределена равномерно на отрезке $[0, \theta]$. По выборке объема n найдены оценки параметра θ : $\tilde{\theta}_1 = 2\bar{x}$, $\tilde{\theta}_2 = x_{\min}$,

$$\tilde{\theta}_3 = x_{\max}, \quad \tilde{\theta}_5 = \left(x_1 + \frac{\sum_{k=2}^n x_k}{(n-1)} \right).$$

- а) Проверить оценки на несмещенность и состоятельность. Исправить эти оценки, если необходимо.
 - б) Какая из исправленных оценок более эффективна?
2. Случайная величина имеет экспоненциальный закон распределения

$$p(x) = \begin{cases} e^{-x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Сгенерируйте выборку объема $n = 25$.

- a) Определить по выборке моду, медиану, размах, оценку коэффициента асимметрии.
- b) Построить эмпирическую функцию распределения, гистограмму и boxplot.
- c) Сравнить оценку плотности распределения среднего арифметического элементов выборки, полученную с помощью ЦПТ, с бутстраповской оценкой этой плотности.
- d) Найти бутстраповскую оценку плотности распределения коэффициента асимметрии и оценить вероятность того, что коэффициент асимметрии будет меньше 1.
- e) Сравнить плотность распределения медианы выборки с бутстраповской оценкой этой плотности.

3. Случайная величина имеет экспоненциальное распределение

$$p(x) = \begin{cases} e^{-x/\theta}/\theta, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad \theta > 0. \text{ По выборке объема } n = 3 \text{ найдены}$$

оценки параметра $\theta : \tilde{\theta}_1 = \bar{x}, \tilde{\theta}_3 = x_{(2)}$ (второй член вариационного ряда).

- a) Проверить оценки на несмещенность. Исправить эти оценки, если необходимо.
- b) Какая из исправленных оценок более эффективна?
- c) Исследовать эти оценки на эффективность с помощью неравенства Крамера–Рао.

4. Случайная величина с равной вероятностью принимает любое значение на интервале $(-1, 1)$, кроме 0, и принимает значения 0 и 2 с одинаковой вероятностью.

- a) По выборке объема n найти оценки параметров случайной величины методом моментов и методом максимального правдоподобия.
- b) Проверить оценки на несмещенность и состоятельность.
- c) Исследовать эти оценки на эффективность с помощью неравенства Крамера–Рао.

5. Случайная величина распределена равномерно на отрезке $[\theta, 2\theta]$.

- a) По выборке объема n найти оценки параметра θ методом моментов и методом максимального правдоподобия.
- b) Проверить оценки на несмещенность и состоятельность. Исправить эти оценки, если необходимо.
- c) Сравнить эффективность исправленных оценок.
- d) Построить точный доверительный интервал для параметра θ .

- e) Построить асимптотический доверительный интервал для параметра θ .
- f) Сгенерировать выборку объема $n = 100$ для некоторого значения параметра θ . Вычислите указанные выше доверительные интервалы для доверительной вероятности 0.95.
- g) Численно постройте бутстраповский доверительный интервал.
- h) Сравнить все интервалы.

6. Случайная величина имеет распределение Парето:

$$p(x) = \begin{cases} \frac{\theta - 1}{x^\theta}, & x \geq 1 \\ 0, & x < 1 \end{cases}, \quad \theta > 1.$$

- a) По выборке объема n найти оценку параметра θ методом максимального правдоподобия.
- b) Построить доверительный интервал для медианы.
- c) Построить асимптотический доверительный интервал для параметра θ .
- d) Сгенерировать выборку объема $n = 100$ для некоторого значения параметра θ . Вычислите указанные выше доверительные интервалы для доверительной вероятности 0.95.
- t) Численно постройте бутстраповский доверительный интервал двумя способами, используя параметрический бутстрап и непараметрический бутстрап.
- f) Сравнить все интервалы.

7. Ниже приводятся ставшие классическими данные Борткевича о числе лиц, убитых ударом копыта в 10 прусских армейских корпусах за 20 лет (1875—1894).

| | | | | | |
|--|-----|----|----|---|---|
| i – число смертей в одном корпусе за год | 0 | 1 | 2 | 3 | 4 |
| число случаев, когда произошло i смертей | 109 | 65 | 22 | 3 | 1 |

Проверить гипотезу о том, что число смертей в одном корпусе за год подчиняется распределению Пуассона.

- 8. Произведено измерение размеров деталей в двух партиях по 100 деталей в каждой партии. В первой партии оказалось 25 деталей с заниженным размером, 50 деталей с точным размером, 25 деталей с завышенным размером, а во второй партии аналогичные числа оказались равны 52, 41, 7 соответственно. Проверить гипотезу о независимости номера партии деталей и размера детали.
- 9. Участники олимпиады по математике разбиты на два потока по 300 человек в каждом. Итоги олимпиады оказались следующими: в первом

потоке оценки 2, 3, 4, 5 получили соответственно 33, 43, 80 и 144 человека, соответствующие данные для второго потока 39, 35, 72 и 154. Можно ли считать оба потока однородными?

10. При снятии показаний измерительного прибора десятые доли деления шкалы прибора оцениваются «на глаз» наблюдателем. Количества цифр 0, 1, 2, ..., 9, записанных наблюдателем в качестве десятых долей при 100 независимых измерениях, равны 5, 8, 6, 12, 14, 18, 11, 6, 13, 7 соответственно.

- a) Проверить гипотезу о согласии данных с законом равномерного распределения с помощью критерия χ^2 и с помощью критерия Колмогорова. Сравнить результаты.
- b) Проверить гипотезу о согласии данных с законом нормального распределения с помощью критерия χ^2 и с помощью критерия Колмогорова. Сравнить результаты.

11. Основная гипотеза H_0 состоит в том, что случайная величина имеет распределение с плотностью $p_0(x) = \begin{cases} 1, & x \in (0, 1) \\ 0, & x \notin (0, 1) \end{cases}$. Альтернатива H_1

состоит в том, что случайная величина имеет распределение с плотно-

стью $p_1(x) = \begin{cases} \frac{e^{1-x}}{e-1}, & x \in (0, 1) \\ 0, & x \notin (0, 1) \end{cases}$.

- a) Построить наиболее мощный критерий проверки этих гипотез по выборке объема $n = 1$ с уровнем значимости α , найти ошибки первого и второго рода, мощность критерия.
- b) Построить наиболее мощный критерий проверки этих гипотез по выборке объема $n = 2$ с уровнем значимости α , найти ошибки первого и второго рода, мощность критерия.
- c) Построить наиболее мощный асимптотический критерий проверки этих гипотез по выборке объема n с уровнем значимости α , найти ошибки первого и второго рода, мощность критерия, исследовать состоятельность этого критерия.
- d) Построить оптимальный критерий по выборке объема n с критической областью $x_{\min} < c$ и уровнем значимости α , найти ошибки первого и второго рода, мощность критерия, исследовать состоятельность этого критерия.

12. У игрока, наблюдавшего за игрой в четырехгранные кости, создалось впечатление, что четверка выпадает в 8 случаях из 24, тройка в 4, а остальные две грани выпадают равновероятно. Получив приглашение

принять участие в игре, игрок попросил разрешения предварительно проверить свою гипотезу на двух производимых подряд бросаниях кости. Единственная рассматриваемая им альтернатива состоит в том, что игральная кость сделана «честно». Найти наиболее мощный критерий с уровнем значимости 0,2. Найти мощность критерия.

13. Исследование длины и ширины 139 черепов, найденных в Верхнем Египте и относимых к расе, жившей за 8000 лет до нашей эры, показало, что стандартное отклонение длины и ширины черепа 5,722 и 4,612 мм соответственно. Те же величины, выведенные на основании обследования 1000 европейцев, оказались равными 6,161 и 5,055 мм. Предполагая, что законы распределения длины и ширины черепа нормальные, выяснить, можно ли считать расхождение стандартов случайным. Построить график мощности критерия.
14. Пусть x_n и y_m — независимые случайные выборки из нормального распределения с параметрами $a, \sigma_x^2 = 2$ и $b, \sigma_y^2 = 1$ соответственно. Используя реализации случайных выборок: $x = \{-1.11, -6.10, 2.42\}$, $y = \{-2.29, -2.91\}$, проверить гипотезу о равенстве средних против альтернатив $a > b$. Построить график мощности критерия.

ВТОРОЕ ЗАДАНИЕ

(срок сдачи 11–16 мая)

1. Случайный вектор $(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \eta)$ имеет компоненты, распределенные по следующему закону: $\xi_k \sim R(-1, 1)$, $\eta \sim N(2 + 3x_1 - 2x_2 + x_3 + x_4 - x_5, 1.5^2)$, где x_k это значения, которые принимает случайная величина ξ_k . Сгенерировать выборку объема $n = 50$.
 - a) Проверить переменные ξ_k на мультиколлинеарность.
 - b) Определить уравнение линейной регрессии $\eta = \beta_0 + \sum_{k=1}^5 \beta_k \xi_k$ и проверить значимость коэффициентов.
 - c) Определить коэффициент детерминации и проверить его значимость.
 - d) Найти значение в точке $x_k = 0$ и построить 95% доверительный интервал.
 - e) Проверить предположение о независимости ошибок измерения.
 - f) Проверить предположение о нормальности распределения ошибок.
 - g) Исследовать регрессию на выбросы.
 - h) Провести кросс-проверку регрессии.
 - i) Проверить адекватность регрессии, сделав 5 повторных измерений в одной точке.

- j) Удалить переменную, соответствующую наименее значимому коэффициенту и повторить пункты *b* и *c*, сравнить уравнения регрессии.
- к) Сравнить уравнения регрессии бутстрапом.

2. В таблице приведены данные о содержании иммуноглобулина *Ig A* в сыворотке крови у больных пяти возрастных групп:

| № | Содержание <i>Ig A</i> (%) |
|---|--|
| 1 | 83, 85 |
| 2 | 84, 85, 85, 86, 86, 87 |
| 3 | 86, 87, 87, 87, 88, 88, 88, 88, 89, 90 |
| 4 | 89, 90, 90, 91 |
| 5 | 90, 92 |

- а) Определить влияние возраста на содержание иммуноглобулина в крови с помощью регрессионного анализа.
- б) Провести попарное сравнение средних в рамках регрессионной модели, с учетом множественности проверяемых гипотез.

3. По трем значениям случайного вектора $(\xi, \eta, \omega) : (1, 0, 1), (0, 1, 5)$ и $(1, 1, 2)$:

- а) найти уравнение линейной регрессии $\omega = a\xi + b\eta$;
- б) найти уравнение *ridge* регрессии $\omega = a\xi + b\eta$, построить график *CVSS* и определить значение параметра регуляризации;
- с) найти уравнение *lasso* регрессии $\omega = a\xi + b\eta$, построить график *CVSS* и определить значение параметра регуляризации.