ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ МЕДИЦИНСКИЙ УНИВЕРСИТЕТ»
МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Кафедра медицинской физики и информатики

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ
к практическим занятиям

Дисциплина МЕДИКО-БИОЛОГИЧЕСКАЯ СТАТИСТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЗДРАВООХРАНЕНИИ

Направление подготовки 34.04.01 Управление сестринской деятельностью

Квалификация Магистр

Курс 2

Семестр 4

Уфа

Рецензенты:

Заведующий кафедрой управления сестринской деятельностью ФГАОУ ВО «Российский университет дружбы народов Имени Патриса Лумумбы», д.м.н., профессор, И.В. Радыш.

Работодатель:

Президент Региональной общественной организации «Профессиональной ассоциации специалистов с высшим сестринским, средним медицинским и фармацевтическим образованием Республики Башкортостан» Э.Ю. Ахметшина.

Автор: доцент Войтик В.В.

Утверждена на заседании № 8 кафедры медицинской физики и информатики, от «16» апреля 2024 г.

1. Тема занятия № 1 и её актуальность. Статистическое распределение. Характеристики статистического распределения. Характеристики положения и вариации. Точечная и интервальная оценки параметров генеральной совокупности нормального распределения по ее выборке. Коэффициент Стьюдента.

Актуальность. Статистические ряды распределения являются базисным методом для любого статистического анализа. Понимание данного метода и навыки его использования необходимы для проведения статистических исследований. Проанализировать выборочный метод как один из аспектов исследования и научиться его планированию и проведению

2. Учебные цели:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен **знать:**

- метод наименьших квадратов;
- формулы для нахождения коэффициентов линейного и квадратического тренда;

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**

- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:

1. Понятие генеральной совокупности; методы исследования генеральной совокупности;
2. Выборочная совокупность; репрезентативность и объем выборки;
3. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
4. Построение графиков вариационных рядов;
5. Понятие об ошибках репрезентативности;
6. Доверительный интервал и доверительная вероятность;
7. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка. 8. Уравнение регрессии

9. Коэффициенты корреляции

10.Нулевая и конкурентная гипотезы

4. Вид занятия: практическое занятие.

5. Продолжительность занятия: 4 часа.

6. Оснащение:

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

1). Статистическое распределение - это:

1. Совокупность вариант; 2. Относительная частота вариант; 3. Совокупность вариант и соответствующих им частот; 4. Совокупность данных и закон их распределения; 5. Совокупность данных.

2) Построить эмпирическую функцию и ее график по данным

❍

$$F^{¿}(x)=\begin{cases} 0 \quad при \quad x \le 0; \\ 0{,}2 \quad при \quad 0 < x \le 2; \\ 0{,}4 \quad при \quad 2 < x \le 3; \\ 0{,}7 \quad при \quad 3 < x \le 4; \\ 0{,}8 \quad при \quad 4 < x \le 10; \\ 1 \quad при \quad x > 10; \end{cases}$$

3) Дана выборка значений случайной величины X объема 20:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12

18, 17, 15, 13, 17, 14, 14, 13, 14, 16

Требуется: - построить дискретный вариационный ряд;

- найти размах варьирования R, моду M_0, медиану M_e;

- построить полигон частостей.

4) Результаты измерений отклонений от нормы диаметров

50 подшипников дали численные значения (в мкм), приведенные в табл. 4.

Таблица 4.

-1,760	-0,291	-0,110	-0,450	0,512

-0,158	1,701	0,634	0,720	0,490
1,531	-0,433	1,409	1,740	-0,266
-0,058	0,248	-0,095	-1,488	-0,361
0,415	-1,382	0,129	-0,361	-0,087
-0,329	0,086	0,130	-0,244	-0,882
0,318	-1,087	0,899	1,028	-1,304
0,349	-0,293	0,105	-0,056	0,757
-0,059	-0,539	-0,078	0,229	0,194
0,123	0,318	0,367	-0,992	0,529

Для данной выборки: - построить интервальный вариационный ряд;

- построить гистограмму и полигон частостей.

5). Величина нормированного отклонения (распределение Стьюдента) зависит от а) Уровня значимости; б) Выборочного среднего; в)Ошибки выборочного среднего;

г) Объема выборки; д) От уровня значимости и от объѐма выборки.

6). Для интервала оценки необходимо знать:

1. X, t, S; 2. X, t, S, n; 3. σX, S, , n; 4. X, t, т; 5.t, S, m

7). Характеристики положения случайной величины:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

8). Характеристики рассеяния:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

9). Сумма всех относительных частот равна

1. единице; 2. объему выборки; 3. 100; 4. 10; 5. 0.

Типовые задачи.

1) Дана функция распределения непрерывной случайной величины X

2) Дана функция распределения непрерывной случайной величины X
Найти плотность распределения f (x).

5) Непрерывная случайная величина X задана плотностью распределения f (x) =3sin3x в интервале (0; π /3); вне этого интервала f (x) = 0 . Найти вероятность того, что X примет значение, принадлежащее интервалу (π/6, π/4).

6) Случайная величина X задана плотностью распределения f(x)=2x в интервале [0,1]. Вне этого интервала f(x)=0. Найти математическое ожидание, и среднее квадратичное отклонение величины X.

7) Случайная величина X задана плотностью распределения f(x)=x/2 в интервале [0,2]; вне этого интервала. f(x) = 0. Найти M(X) и D(X).

8) Случайная величина X задана плотностью распределения f(x) = C(x^2+2x) в интервале [0,1]; вне этого интервала. f(x)=0. Найти: а) параметр C; б) математическое ожидание величины X.

9) Математическое ожидание нормально распределенной случайной величины X равно μ = 3 и среднее квадратическое отклонение σ=2. Написать плотность вероятности X.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Корреляционную зависимость Y от X можно описать с помощью уравнения вида: M(Y)x = f(x), где M(Y)x – условное математическое ожидание величины Y, соответствующее данному значению x; f(x) – некоторая функция.

Если функция f(x) линейная, то уравнение регрессии можно записать в виде: M(Y)x = Ax + B, где A и B – параметры.

Корреляционная таблица содержит всю информацию, полученную в результате выборочных наблюдений.

yj xi	x1	x2	···	xk	ny
y1	n11	n21	···	nk1	ny1
y2	n12	n22	···	nk2	ny2
…	…	…	…	…	…
yl	n1l	n2l	···	nkl	nyl
nx	nx1	nx2	···	nxk	N

С помощью корреляционной таблицы для каждого значения xi можно записать соответствующее распределение величины Y.

В случаях, когда существует линейная зависимость между величинами X и Y, для описания корреляционной зависимости вводятся выборочные уравнения линейной = ρ + ρ регрессии: $Y^{x}{}_{yx}x\ b$, где yx –

выборочный коэффициент регрессии, имеющий смысл выборочной оценки коэффициента А (см. формулу

4.17), условное среднее $\bar{Y}_x$ является оценкой условного математического ожидания M(Y)x, а параметр b – оценкой B.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx} \frac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy} \frac{\sigma_y}{\sigma_x}$.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5 - 15\cdot 2{,}35}{\sqrt{50\cdot 0{,}7025}} = \frac{40{,}5 - 35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

Проверка статистических гипотез

1. Проверка значимости выборочного коэффициента линейной корреляции. После вычисления выборочного коэффициента линейной корреляции r всегда возникает необходимость проверки гипотезы о наличии существенности линейной корреляционной зависимости между изучаемыми величинами, или гипотезы о значимости выборочного коэффициента линейной корреляции. Для этого, предполагая, что величины X и Y распределены по нормальному закону, вычисляют экспериментальное значение критерия: $t_{эксп} = r\frac{\sqrt{n-2}}{\sqrt{1-r^2}}$, где n – объем выборки (N).

Далее по таблице критических значений распределения Стьюдента (таблица 4 Приложения) при заданном уровне значимости p (с заданной вероятностью γ = 1 – p) и числе степеней свободы f = n – 2 находят критическое значение критерия: tкр = t(f).

Если |tэксп| >tкр, то выборочный коэффициент значим.

2. Проверка статистической гипотезы о равенстве средних значений двух произвольно распределенных случайных величин по их оценкам. Большие независимые выборки ($n \geq 30$)

Пусть рассматриваются две случайные величины x и y и по результатам выборочных измерений этих величин требуется проверить так называемую нулевую гипотезу о равенстве их средних значений. Сначала рассчитывается экспериментальное значение критерия:

$$Z_{эксп} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}},$$

где $\bar{x}$, $\bar{y}$ – выборочные средние величин x и y, nx, ny – объемы выборок величин x и y, S_x^2, S_y^2

$$\Phi(Z_{кр}) = \frac{1-p}{}$$

– оценки дисперсий величин x и y. Затем находится критическое значение критерия:

, где Φ(Z) – функция Лапласа (таблица 3 Приложения), p – заданный уровень значимости. 2

После этого сравниваются полученные значения, и если |Zэксп| <Zкр, то при заданном уровне p делают

вывод о равенстве средних значений $\bar{x}, \bar{y}$.

3. Проверка статистической гипотезы о равенстве средних значений двух нормально распределенных случайных величин. Малые независимые выборки (n< 30)

Если известно, что величины X и Y – нормально распределенные и их генеральные дисперсии равны (или оценки дисперсий различаются незначимо), то для проверки гипотезы используется следующий метод. Сначала рассчитывается экспериментальное значение критерия:

$$t_{эксп} = \frac{\bar{x}-\bar{y}}{\sqrt{(n_x-1)S_x^2+(n_y-1)S_y^2}} \cdot \sqrt{\frac{n_x n_y (n_x+n_y-2)}{n_x+n_y}}$$

Затем находят критическое значение критерия: $t_{kp} = t(p; f)$, где t(p; f) – находится по таблице распределения Стьюдента (таблица 5 Приложения), где p – уровень значимости, f = nx + ny – 2.

После этого полученные значения сравниваются, и если |tэксп| <tкр, то гипотеза о равенстве средних значений при заданном p является согласующейся с результатами наблюдений.

4. Проверка статистической гипотезы о равенстве дисперсий двух нормально распределенных величин по их оценкам

Пусть по результатам наблюдений двух нормально распределенных случайных величин x и y (объемы выборок nx и ny) найдены оценки их дисперсий Sx2 и Sy2. Для проверки значимости различия между ними выдвигается нулевая гипотеза о равенстве дисперсий (x2 = y2). Задается уровень значимости p, с которым проводится проверка. Вычисляется экспериментальное значение критерия:

$$F_{эксп.} = \frac{S_б^2}{S_м^2}$$, т.е. отношение большей из оценок дисперсий к меньшей. Затем по

таблице критических значений распределения Фишера-Снедекора (таблица 6

Приложения) при заданном уровне значимости p находят критическое значение критерия:

$$F_{kp} = F\left(\frac{p}{2}; f_1; f_2\right)$$, где f1 = nб – 1, f2 = nм – 1 (nб, nм – объемы выборок с большей и меньшей

дисперсией соответственно).

Если Fэксп<Fкр, то нулевая гипотеза считается согласующейся с результатами наблюдений.

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

7.4. Самостоятельная работа обучающихся под контролем преподавателя.

Типовые задачи.

Задача 1. Частота пульса по данным медицинского осмотра 17 девочек-первоклассниц: 76 76 70 66 68 70 72 74 76 78 70 82 68 74 70 70 70. Найти по выборочным данным точечные оценки параметров генеральной совокупности и оценить истинное значение генерального среднего с доверительной вероятностью 0.95.

Задача 2. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных $NaNO^3$ (Y) при соответствующих температурах (X) раствора:

частей

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие

$NaNO^3$ и

какова уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO^3$, если растворилось 84 условных частей. температура раствора

7.5. Контроль конечного уровня усвоения темы:

Подготовка к выполнению практических приѐмов по теме занятия.

Материалы для контроля усвоения темы: набор тестовых заданий.

Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.

Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.

8. Литература (см. приложение) (

1.Тема занятия № 2 и еѐ актуальность. **Проверка гипотез о равенстве генеральных средних и дисперсий.Проверка гипотезы о нормальном распределении Непараметрические критерии.**

Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.

2. Учебные цели:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен **знать:**

- метод наименьших квадратов;
- формулы для нахождения коэффициентов линейного и квадратического тренда;

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**

- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:

1. Понятие генеральной совокупности; методы исследования генеральной совокупности;
2. Выборочная совокупность; репрезентативность и объем выборки;
3. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
4. Построение графиков вариационных рядов;
5. Понятие об ошибках репрезентативности;
6. Доверительный интервал и доверительная вероятность;
7. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка.
8. Уравнение регрессии
9. Коэффициенты корреляции

4. Вид занятия: практическое занятие.

5. Продолжительность занятия:4 часа.

6. Оснащение:

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

1). Статистическое распределение - это:

1. Совокупность вариант; 2. Относительная частота вариант; 3. Совокупность вариант и соответствующих им частот; 4. Совокупность данных и закон их распределения; 5. Совокупность данных.

2). Распределение может быть:

1. Одномодальным; 2. Двумодальным; 3. Многомодальным; 4. Среди ответов 1-3 нет верного;

2) Построить эмпирическую функцию и ее график по данным

❍

$$F^{¿}(x)=\begin{cases} 0 \quad при \quad x \le 0; \\ 0{,}2 \quad при \quad 0 < x \le 2; \\ 0{,}4 \quad при \quad 2 < x \le 3; \\ 0{,}7 \quad при \quad 3 < x \le 4; \\ 0{,}8 \quad при \quad 4 < x \le 10; \\ 1 \quad при \quad x > 10; \end{cases}$$

3) Дана выборка значений случайной величины X объема 20:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12

18, 17, 15, 13, 17, 14, 14, 13, 14, 16

Требуется: - построить дискретный вариационный ряд;

- найти размах варьирования R, моду M_0, медиану M_e;

- построить полигон частостей.

5). Величина нормированного отклонения (распределение Стьюдента) зависит от а) Уровня значимости; б) Выборочного среднего; в)Ошибки выборочного среднего;

г) Объема выборки; д) От уровня значимости и от объѐма выборки.

6). Для интервала оценки необходимо знать:

1. X, t, S; 2. X, t, S, n; 3. X, S, , n; 4. X, t, т; 5.t, S, m

7). Характеристики положения случайной величины:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

8). Характеристики рассеяния:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

9). Сумма всех относительных частот равна

1. единице; 2. объему выборки; 3. 100; 4. 10; 5. 0.

Типовые задачи.

1) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases} 0, \text{при } x \le 0 \\ \sin x, \text{при } 0 < x \le \pi/2 \\ 1, \text{при } x > \pi/2 \end{cases} \qquad F(X)=\begin{cases} 0, \text{при } x \le 0 \\ \sin 2x, \text{при } 0 < x \le \pi/4 \\ 1, \text{при } x > \pi/4 \end{cases}$$

Найти плотность распределения f (x).

2) Дана функция распределения непрерывной случайной величины X

Найти плотность распределения f (x).

3) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases} 0, \text{при } x \le 0 \\ ax^2, \text{при } 0 < x \le 2 \\ 1, \text{при } x > 2 \end{cases} \qquad f(X)=\begin{cases} 0, \text{при } x \le 0 \\ \cos x, \text{при } 0 < x \le \pi/2 \\ 0, \text{при } x > \pi/2 \end{cases}$$

Найти коэффициент a и плотность вероятности попадания случайной величины X в интервал (0; 1). 4) Задана плотность распределения непрерывной случайной величины X:

Найти функцию распределения F(x).

5) Непрерывная случайная величина X задана плотностью распределения

f (x) =3sin3x в интервале (0; π /3); вне этого интервала f (x) = 0 . Найти вероятность того, что X примет значение, принадлежащее интервалу (π/6, π/4).

6) Случайная величина X задана плотностью распределения f(x)=2x в интервале [0,1]. Вне этого интервала f(x)=0. Найти математическое ожидание, и среднее квадратичное отклонение величины X.

7) Случайная величина X задана плотностью распределения f(x)=x/2 в интервале [0,2]; вне этого интервала. f(x) = 0. Найти M(X) и D(X).

8) Случайная величина X задана плотностью распределения $f(x) = C(x^2+2x)$ в интервале [0,1]; вне этого интервала. $f(x)=0$. Найти: а) параметр C; б) математическое ожидание величины X.

9) Математическое ожидание нормально распределенной случайной величины X равно $\mu = 3$ и среднее квадратическое отклонение $\sigma=2$. Написать плотность вероятности X.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \dfrac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx}\dfrac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy}\dfrac{\sigma_y}{\sigma_x}$.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5 - 15\cdot 2{,}35}{\sqrt{50\cdot 0{,}7025}} = \frac{40{,}5 - 35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

7.4. Самостоятельная работа обучающихся под контролем преподавателя.

Задача. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей.

7.5. Контроль конечного уровня усвоения темы:

Подготовка к выполнению практических приѐмов по теме занятия.
Материалы для контроля усвоения темы: набор тестовых заданий.
Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.
8. Литература (см. приложение) (см. приложение):

1.Тема занятия № 3 и её актуальность. Основные понятия корреляционного анализа. Коэффициент корреляции визуализация данных. Выборочное уравнение линейной регрессии.

Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.

2. **Учебные цели**:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен **знать:**
- метод наименьших квадратов;
- формулы для нахождения коэффициентов линейного и квадратического тренда;

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**
- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы:

Вопросы для самоподготовки:

4. Вид занятия: практическое занятие. **5. Продолжительность занятия**: 4 часа

6. Оснащение:

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов. Типовые задачи.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Корреляционную зависимость Y от X можно описать с помощью уравнения вида: M(Y)x = f(x), где M(Y)x – условное математическое ожидание величины Y, соответствующее данному значению x; f(x) – некоторая функция.

Если функция f(x) линейная, то уравнение регрессии можно записать в виде: M(Y)x = Ax + B, где A и B – параметры. Корреляционная таблица содержит всю информацию, полученную в результате выборочных наблюдений.

yj xi	x1	x2	· · ·	xk	ny
y1	n11	n21	· · ·	nk1	ny1
y2	n12	n22	···	nk2	ny2
· · ·	· · ·	· · ·	· · ·	· · ·	· · ·
yl	n1l	n2l	···	nkl	nyl
nx	nx1	nx2	· · ·	nxk	N

С помощью корреляционной таблицы для каждого значения xi можно записать соответствующее распределение величины Y.

В случаях, когда существует линейная зависимость между величинами X и Y, для описания корреляционной зависимости вводятся выборочные уравнения линейной регрессии: $\bar{Y}_x = \rho_{yx} x + b$, где ρyx – выборочный коэффициент регрессии, имеющий смысл выборочной оценки коэффициента A (см. формулу 4.17), условное среднее $\bar{Y}_x$ является оценкой условного математического ожидания M(Y)x, а параметр b – оценкой B.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \dfrac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx}\dfrac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy}\dfrac{\sigma_y}{\sigma_x}$.

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5 - 15\cdot 2{,}35}{\sqrt{50\cdot 0{,}7025}} = \frac{40{,}5 - 35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

7.4. Самостоятельная работа обучающихся под контролем преподавателя.
Задача. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей.

7.5. Контроль конечного уровня усвоения темы:
Подготовка к выполнению практических приѐмов по теме занятия.
Материалы для контроля усвоения темы: набор тестовых заданий.
Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.
8. Литература (см. приложение)

1.Тема занятия № 4 и её актуальность Сущность дисперсионного анализа. Однофакторный дисперсионный анализ вероятностных законов распределения, которым подчиняются данные; выявление различий между группами; определение взаимосвязей между переменными; предварительный выбор методов анализа.
Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.
2. Учебные цели:
- закрепить основные понятия математической статистики; научиться проверять первичные статистические гипотезы.

Для формирования профессиональных компетенций студент должен **знать:**
- метод наименьших квадратов;
- формулы для нахождения коэффициентов линейного и квадратического тренда;

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**
- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.
3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:
1. Понятие генеральной совокупности; методы исследования генеральной совокупности;
2. Выборочная совокупность; репрезентативность и объем выборки;
3. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
4. Построение графиков вариационных рядов;
5. Понятие об ошибках репрезентативности;
6. Доверительный интервал и доверительная вероятность;
7. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка. 8. Уравнение регрессии

9. Коэффициенты корреляции

10.Нулевая и конкурентная гипотезы

4. Вид занятия: практическое занятие.
5. Продолжительность занятия: 4 часа
6. Оснащение:
- 6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.
- 6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

1. Проверка статистической гипотезы о равенстве дисперсий двух нормально распределенных величин по их оценкам

Пусть по результатам наблюдений двух нормально распределенных случайных величин x и y (объемы выборок n_x и n_y) найдены оценки их дисперсий S_x^2 и S_y^2. Для проверки значимости различия между ними выдвигается нулевая гипотеза о равенстве дисперсий ($\sigma_x^2 = \sigma_y^2$). Задается уровень значимости p, с которым проводится проверка. Вычисляется экспериментальное значение критерия:

$F_{эксп} = \frac{S_б^2}{S_м^2}$, т.е. отношение большей из оценок дисперсий к меньшей. Затем по таблице критических значений распределения Фишера-Снедекора (таблица 6 Приложения) при заданном уровне значимости p находят критическое значение критерия:

$F_{kp} = F\left(\frac{p}{2};\ f_1;\ f_2\right)$, где $f_1 = n_б - 1$, $f_2 = n_м - 1$ ($n_б$, $n_м$ – объемы выборок с большей и меньшей дисперсией соответственно).

Если $F_{эксп} < F_{кр}$, то нулевая гипотеза считается согласующейся с результатами наблюдений.

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

Гипотеза о равенстве среднего значения числу

Пример 1. Продавец утверждает, что средний вес пачки чая составляет 100 г. Из партии извлечена выборка и взвешена. Вес каждой пачки - 98, 104, 97, 97, 101, 100, 99, 101, 99, 98. Не противоречит ли это утверждению продавца? Доверительная вероятность 99%. Вес пачек чая распределен нормально.

$$= \frac{\sum x_i n_i}{n} = \frac{994}{10} = 99,4$$

Исправленная выборочная дисперсия $S^2 = \frac{\sum \left(x_i - \overline{x}\right)^2 n_i}{n-1} = \frac{42,4}{9} = 4,71$. Выборочное исправленное среднеквадратичное отклонение: S ≈ 2,171. Расчеты в таблице:

x_i	n_i	$x_i n_i$	$(x_i - \overline{x})^2 n_i$
97	2	194	11,52
98	2	196	3,92
99	2	198	0,32
100	1	100	0,36
101	2	202	5,12
104	1	104	21,16
Сумма	10	994	42,4

Введем нулевую гипотезу H_0 : a=100 при конкурирующей гипотезе H_1: a ≠100 . Вычисляем наблюдаемое значение критерия

$T_{наблюд.} = \frac{x - \mu}{S}\sqrt{n} = \frac{99,4 - 100}{2,171}\sqrt{10} = -0,87$. По таблице критических точек распределения Стьюдента

Решение. Вычислим показатели выборки. ~~х~~ найдем критическую точку по уровню значимости α = 0,01 и числу степеней свободы k = 9 , откуда $t_{кр} \approx 3,25$. Так как $|T_{наблюд.}| = 0,87 < 3,25 = t_{кр.}$, то нулевую гипотезу о равенстве среднего веса 100 г можно принять.

Гипотеза о равенстве дисперсии числу Пример 3. По результатам n=7n=7 независимых измерений найдено, что x¯¯¯=82,48x¯=82,48 мм, а S=0,08 мм. Допустив, что ошибки измерения имеют нормальное распределение проверить на уровнезначимости α=0,05α=0,05 гипотезу H0:σ2=0,01H0:σ2=0,01 мм22. против конкурирующей гипотезы H0:σ2=0,005H0:σ2=0,005 мм22. В ответе записать разность между фактическим и табличным значениями выборочной характеристики.

Вычисляем наблюдаемое значение критерия:

$$\chi^2_{набл} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{6 \cdot 0,08^2}{0,01} = 3,84 .$$

Вычисляем критическое значение $\chi^2_{кр}(1-\alpha; n-1) = \chi^2_{кр}(0,95; 6) = 12,583$.

Так как $\chi^2_{кр} = 12,583 > 3,84 = \chi^2_{набл}$, нулевую гипотезу следует отвергнуть.

Разность между фактическим и табличным значениями $3,84 - 12,583 = -8,743$.

Ответ: -8,743

Пример 4. Компания не осуществляет инвестиционных вложений в ценные бумаги с дисперсией годовой доходности более чем 0,04. Выборка из 52 наблюдений по активу А показала, что выборочная дисперсия ее доходности равна 0,045.Выяснить, допустимы ли для данной компании инвестиционные вложения в актив А на уровне значимости: а) 0,05; б) 0,01.

Решение. Нулевая гипотеза задачи $H_0 : \sigma^2 = \sigma_0^2 = 0,04$. Конкурирующая гипотеза: $H_1 : \sigma^2 > 0,04$.

Вычисляем наблюдаемое значение критерия:

$$\chi^2_{набл} = \frac{(n-1)s_x^2}{\sigma_0^2} = \frac{(52-1) \cdot 0,045}{0,04} = 57,375 .$$

Конкурирующая гипотеза имеет вид $H_1 : \sigma^2 > 0,04$, поэтому критическая область правосторонняя.

А) По таблице вычисляем критическое значение $\chi^2_{кр}(\alpha; n-1) = \chi^2_{кр}(0,05; 51) = 68,669$.

Так как $\chi^2_{набл} = 57,375 < 68,669 = \chi^2_{кр}$, можно принять нулевую гипотезу на данном уровне значимости 0,05.

Б) По таблице вычисляем критическое значение $\chi^2_{кр}(\alpha; n-1) = \chi^2_{кр}(0,01; 51) = 77,386$.
Так как $\chi^2_{набл} = 57,375 < 77,386 = \chi^2_{кр}$, можно принять нулевую гипотезу на данном уровне значимости 0,01.

Таким образом и на уровне значимости 0,05, и на уровне значимости 0,01, для компании инвестиционные вложения в актив А допустимы.

Гипотеза о равенстве вероятности числу

Пример 5. Фирма рассылает рекламные каталоги возможным заказчикам. Как показал опыт, вероятность того, что организация получившая каталог, закажет рекламируемое изделие, равна 0,08. Фирма разослала 1000 каталогов новой, улучшенной, формы и получила 100 заказов. На уровне значимости 0,05 выяснить, можно ли считать, что новая форма рекламы существенно лучше прежней.

Введем нулевую гипотезу $H_0: \ p = p_0 = 0,08$ при конкурирующей гипотезе $H_1: p > 0,08$.
Вычислим наблюдаемое значение критерия по формуле:

$$U_{\text{íàáë}} = \frac{(m/n - p_0)\sqrt{n}}{\sqrt{p_0(1-p_0)}} = \frac{(100/1000 - 0,08)\sqrt{1000}}{\sqrt{0,08 \cdot 0,92}} \approx 2,33.$$

Найдем критическую точку правосторонней критической области из равенства:

$$\Phi(u_{\text{êð}}) = \frac{1-2\alpha}{2} = \frac{1-0,1}{2} = 0,45, \text{ откуда } u_{кр} = 1,645.$$

Так как $U_{набл} > u_{кр}$, то следует отвергнуть нулевую гипотезу. Новая форма рекламы значимо эффективнее прежней.

Пример 6. Обычно применяемое лекарство снимает послеоперационные боли у 80% пациентов. Новое лекарство, применяемое для тех же целей, помогло 90 пациентам из первых 100 оперированных. Можно ли на уровне значимости а = 0,05 считать, что новое лекарство лучше? А на уровне а = 0,01?

Здесь нулевая гипотеза состоит в том, что новое лекарство действует как и прежнее (излечивает боли в 80% случаев):

$$H_0: \ p = 0,8.$$

Альтернативная гипотеза - H_1 является левосторонней — действие лекарства более сильное:

$$H_1: \ p > 0,8.$$

Уровень значимости и объем выборки даны в условии: $\alpha_1 = 0,05$; $\alpha_2 = 0,01$, n=100.
Если при проверке фиксируется число пациентов K, которым лекарство помогло, то эта величина может принимать значения из множества

$$\Omega = \{0, 1, ..., 100\}.$$

В случае истинности гипотезы H_0 величина K является биномиальной случайной величиной $Bi(100; \ 0,8)$. Следовательно,

$$K \approx N\left(100 \cdot 0,8; \ \sqrt{100 \cdot 0,8 \cdot 0,2}\right) \approx N(80; 4)$$

Пусть $\alpha_1 = 0,05$. Найдем границу критической области по формуле:

$$Z_{1-2\alpha} = Z_{0.9} = 1,645,$$

$$x_{\text{ïðàâ}} = 80 + 4 \cdot 1,645 = 86,58$$

Таким образом, критическая область может быть записана в виде

$$S = \{87, ..., 100\}.$$

Область принятия нулевой гипотезы соответственно такова:

$$\{0, 1, ..., 86\}.$$

Так как новое лекарство помогло 90 пациентам из первых 100 оперированных, то следует отвергнуть нулевую гипотезу.

Пусть $\alpha_2 = 0,01$. Найдем границу критической области по формуле:

$$Z_{1-2\alpha} = Z_{0.98} = 2,33,$$

$$x_{\text{ïðàâ}} = 80 + 4 \cdot 2,33 = 89,32$$

Таким образом, критическая область может быть записана в виде

$$S = \{90, ..., 100\}.$$

Область принятия нулевой гипотезы соответственно такова:

$$\{0, 1, ..., 89\}.$$

Гипотеза о равенстве средних

Пример 7. Ожидается, что добавление специальных веществ уменьшит жесткость воды. По оценке жесткости воды до после добавления специальных веществ по 40-ка и 50-ти пробам соответственно получим средние значения жесткости (в стандартных единицах), равные 4,0 и 0,8. Дисперсия измерений в обоих случаях предполагается равно 0,25. Подтверждают ли эти результаты ожидаемый эффект? Принять α=0,05α=0,05. Контролируемая величина имеет нормальное распределение.

Введем нулевую гипотезу $H_0 : M(X) = M(Y)$, Где X, Y - жесткость воды до и после добавления реагента. Альтернативная гипотеза $H_1 : M(X) \neq M(Y)$.

Вычислим наблюдаемое значение критерия:

$$Z_{набл} = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}} = \frac{4,0 - 0,8}{\sqrt{\frac{0,25}{40} + \frac{0,25}{50}}} \approx 30,17 .$$

По таблице функции Лапласа найдем критическую точку из условия

$$\Phi\left(z_{кр}\right) = \frac{1-\alpha}{2} = \frac{1-0,05}{2} = 0,475, \text{ откуда } z_{кр} = 1,96 .$$

Так как $\left|Z_{набл}\right| = 30,17 > 1,96 = z_{кр}$, то нулевую гипотезу $H_0 : M(X) = M(Y)$ следует отвергнуть.

Влияние реагента существенно, результаты подтверждают ожидаемый эффект

Пример 8. Производительность каждого из агрегатов А и В составила (в кг вещества за час работы)

Номер замера	1	2	3	4	5
Агрегат А	14,1	13,1	14,7	13,7	14,0
Агрегат В	14,0	14,5	13,7	12,7	14,1

Можно ли считать производительность агрегатов А и В одинаковой в предложении, что обе выборки получены из нормально распределенных генеральных совокупностей, при уровне значимости а = 0,1?

Обозначим первую выборку (агрегат А) X, вторую выборку (агрегат В) Y. Найдем выборочные числовые характеристики выборок.

						Сумма
x_i	14,1	13,1	14,7	13,7	14	**69,6**
$(x_i-\bar{x})^2$	0,0324	0,6724	0,6084	0,0484	0,0064	**1,368**

Выборочное среднее $\bar{x}=\frac{1}{n}\sum x_i=\frac{69,6}{5}=13,92$.

Выборочная исправленная дисперсия $s_x^2=\frac{1}{n-1}\sum(x_i-\bar{x})^2=\frac{1,368}{4}=0,342$.

						Сумма
y_i	14	14,5	13,7	12,7	14,1	**69**
$(y_i-\bar{y})^2$	0,04	0,49	0,01	1,21	0,09	**1,84**

Выборочное среднее $\bar{y}=\frac{1}{m}\sum y_i=\frac{69}{5}=13,8$.

Выборочная исправлена дисперсия $s_y^2=\frac{1}{m-1}\sum(y_i-\bar{y})^2=\frac{1,84}{4}=0,46$.

Исправленные дисперсии различны, поэтому проверим гипотезу о равенстве дисперсий, используя критерий Фишера-Снедекора. Найдем отношение большей дисперсии к меньшей:

$F_{набл}=\frac{0,46}{0,342}\approx 1,345$.

В качестве конкурирующей гипотезы примем гипотезу $H_1: D(X)\neq D(Y)$. По таблице при уровне значимости $\alpha/2=0,1/2=0,05$ и числам степеней свободы $k_1=m-1=4$ и $k_2=n-1=4$ найдем критическую точку $F_{кр}(0,05,4,4)=6,39$. Так как $F_{набл}=1,345<6,39=F_{кр}$, нет оснований отвергнуть нулевую гипотезу о равенстве генеральных дисперсий.

Итак, предположение о равенстве генеральных дисперсий выполняется, поэтому сравним средние.

Введем нулевую гипотезу: $H_0: M(X)=M(Y)$. Вычислим наблюдаемое значение критерия по формуле:

$$T_{набл.}=\frac{\bar{x}-\bar{y}}{\sqrt{(n-1)s_x^2+(m-1)s_y^2}}\cdot\sqrt{\frac{n\cdot m\cdot(n+m-2)}{n+m}}.$$

Получаем:

$$T_{набл.}=\frac{13,92-13,8}{\sqrt{4\cdot 0,342+4\cdot 0,46}}\cdot\sqrt{\frac{5\cdot 5\cdot 8}{10}}\approx 0,3.$$

Находим критическую точку (двусторонняя область) из таблицы Стьюдента при уровне значимости $\alpha=0,1$ и числу степеней свободы $k=m+n-2=8$ $t_{дв.кр}=1,86$. Так как наблюдаемое значение критерия 0,3 меньше критического, нет оснований отвергнуть нулевую гипотезу, то есть можно сказать, что производительность агрегатов А и В одинакова.

Гипотеза о равенстве дисперсий

Пример 9. До наладки станка была проверена точность изготовления 10 втулок и найдено значение оценки дисперсии диаметра s21=9,6s12=9,6 мкм22. После наладки подверглись контролю еще 15 втулок и получено новое значение оценки дисперсии s22=5,7s22=5,7 мкм22. Можно ли считать, что в результате наладки станка точность изготовления деталей увеличилась? Принять α=0,05α=0,05.

Проверим нулевую гипотезу $H_0: D(X)=D(Y)$ о равенстве генеральных дисперсий, при конкурирующей гипотезе $H_1: D(X)>D(Y)$.

Используем критерий Фишера-Снедекора. Найдем отношение большей дисперсии к меньшей:

$F_{набл}=\frac{9,6}{5,7}=1,68$.

По условию конкурирующая гипотеза имеет вид $H_1: D(X)>D(Y)$, потому критическая область – правосторонняя.

По таблице при уровне значимости $\alpha=0,05$ и числам степеней свободы $k_1=n_1-1=9$ (большая дисперсия) и $k_2=n_2-1=14$ (меньшая дисперсия) найдем критическую точку $F_{кр}(0,05;9;14)=2,64$. Так как $F_{набл}=1,68<2,64=F_{кр}$, нет оснований отвергнуть нулевую гипотезу о равенстве генеральных дисперсий, дисперсии различаются незначимо.

Гипотеза о равенстве вероятностей

Пример 11. Из 200 задач первого раздела курса математики, предложенных для решения, абитуриенты решили 130, а из 300 задач второго раздела абитуриенты решили 120. Можно ли при α=0,01 утверждать, что первый раздел школьного курса абитуриенты усвоили лучше, чем второй.

РЕШЕНИЕ. Пусть p_1 - процент абитуриентов, решающих задачи первого раздела, p_2 - процент абитуриентов, решающих задачи второго раздела. Введем нулевую гипотезу $H_0 : p_1 = p_2$ при конкурирующей гипотезе $H_1 : p_1 > p_2$. Вычислим наблюдаемое значение критерия по формуле:

$$U_{набл.} = \frac{\frac{m_1}{n_1} - \frac{m_2}{n_2}}{\sqrt{\frac{m_1+m_2}{n_1+n_2}\left(1-\frac{m_1+m_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$, где $m_1 = 130,\ n_1 = 200,\ m_2 = 120,\ n_2 = 300$

Подставляем

$$U_{набл.} = \frac{\frac{130}{200} - \frac{120}{300}}{\sqrt{\frac{130+120}{200+300}\left(1-\frac{130+120}{200+300}\right)\left(\frac{1}{200}+\frac{1}{300}\right)}} \approx 5,48$$

Найдем критическую точку $U_{кр.}$ из условия $\Phi(U_{\partial\partial.}) = \frac{1-2\alpha}{2} = \frac{1-0,02}{2} = 0,49$, $U_{\partial\partial.} = 2,33$.

Так как $|U_{набл.}| = 5,48 > 2,33 = U_{кр.}$, нулевую гипотезу следует отвергнуть на данном уровне значимости, можно считать, что первый раздел усвоен лучше.

7.4. Самостоятельная работа обучающихся под контролем преподавателя. Типовые задачи.

Задача 1. При уровне значимости α=0,1α=0,1 проверить гипотезу о равенстве дисперсий двух нормально распределенных случайных величин X и Y на основе выборочных данных (табл. 4) при альтернативной гипотезе H1:σ2x≠σ2yH1:σx2≠σy2.

Задача 2. Выборочная проверка надежности материнских плат 2-х производителей дала следующие результаты: в течения месяца после продажи в 15 из 200 материнских плат производителя А обнаружены дефекты, тогда как среди 400 материнских плат производителя В 8% оказались дефектами. Существенны ли различия в надежности материнских плат производителей А и В? Уровень значимости принять равным 0,01.

Задача 3. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей. Решить задачи:

№ 16-17 (Высшая математика: учебник для студентов высш. пед. учеб.заведений Баврин, И. И. - 5-е изд., стер. - М. : Академия, 2005. - 611 с., с. 602).

Основы высшей математики: учебник. Лобоцкая, Н. Л. - 2-е изд., перераб. и доп., стереотипное издание.

Перепечатка с издания 1978 г. - М.: Альянс, 2015. - 479 с. № 1-17 (с. 31).

7.5. Контроль конечного уровня усвоения темы:

Подготовка к выполнению практических приѐмов по теме занятия.

Материалы для контроля усвоения темы: набор тестовых заданий.

1. Статистическая гипотеза – это:

а) любое предположение, используемое в статистическом исследовании;

б) предположение, которое можно проверить с использованием имеющейся статистической информации;

в) научное предположение, выдвигаемое для объяснения какого-либо явления и требующее проверки на опыте.

2. Критерий – это:
а) отличительный признак, принимаемый за норму, мерило;

б) то, что удостоверяет объективную истинность познания;

в) набор правил, принимаемых для проверки статистической гипотезы.

3. Мощность критерия представляет собой:
а) объекты, вводимые в процесс производства;

б) способность критерия четко различать нулевую и альтернативную статистические гипотезы;

в) величина, которой определяется количество энергии, развиваемой двигателем.

4. Ошибка первого ряда – это:
а) принятие статистической гипотезы, когда она ошибочна;

б) отклонение статистической гипотезы, когда она правильна;

в) ошибка при установлении истинного значения признака;

г) ошибка при исчислении статистического показателя.

5. Ошибка второго ряда – это:
а) принятие статистической гипотезы, когда она ошибочна;

б) отклонение статистической гипотезы, когда она правильна;

в) ошибка при установлении истинного значения признака;

г) ошибка при исчислении статистического показателя.

6. Уровень значимости – это:
а) вероятность, с которой гарантируется надежность результата исчисления того или иного показателя;

б) величина количественного показателя или степень проявления качественного показателя;

в) вероятность, соответствующая отклонению верной гипотезы.

7. Критическая область значений – это:
а) максимальные значения признака;

б) минимальные значения признака;

в) область, попадание значения статистического критерия вкоторою, приводит к отклонению испытываемой статистической гипотезы.

Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.

8. **Литература (см. приложение) (см. приложение)**:

1.Тема занятия № 5 и ее актуальность Типы временных рядов. Простейшие показатели временных рядов и методы их оценки. Тренд временного ряда. Способы задания тренда. Выравнивание временного ряда.

Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.

2. Учебные цели:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен **знать:**

-тренд временного ряда.

- способы задания тренда.

- выравнивание временного ряда.

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**

- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:

10. Понятие генеральной совокупности; методы исследования генеральной совокупности;
11. Выборочная совокупность; репрезентативность и объем выборки;
12. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
13. Построение графиков вариационных рядов;
14. Понятие об ошибках репрезентативности;
15. Доверительный интервал и доверительная вероятность;
16. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка.
17. Уравнение регрессии
18. Коэффициенты корреляции

4. Вид занятия: практическое занятие.

5. Продолжительность занятия:4 часа.

6. Оснащение:

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

1). Статистическое распределение - это:

1. Совокупность вариант; 2. Относительная частота вариант; 3. Совокупность вариант и соответствующих им частот; 4. Совокупность данных и закон их распределения; 5. Совокупность данных.

2). Распределение может быть:

1. Одномодальным; 2. Двумодальным; 3. Многомодальным; 4. Среди ответов 1-3 нет верного;

2) Построить эмпирическую функцию и ее график по данным

❍

$$F^{¿}(x)=\begin{cases} 0 \quad при \quad x \le 0; \\ 0{,}2 \quad при \quad 0 < x \le 2; \\ 0{,}4 \quad при \quad 2 < x \le 3; \\ 0{,}7 \quad при \quad 3 < x \le 4; \\ 0{,}8 \quad при \quad 4 < x \le 10; \\ 1 \quad при \quad x > 10; \end{cases}$$

3) Дана выборка значений случайной величины X объема 20:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12

18, 17, 15, 13, 17, 14, 14, 13, 14, 16

Требуется: - построить дискретный вариационный ряд;

- найти размах варьирования R, моду M_0, медиану M_e;

- построить полигон частостей.

5). Величина нормированного отклонения (распределение Стьюдента) зависит от а) Уровня значимости; б) Выборочного среднего; в)Ошибки выборочного среднего;

г) Объема выборки; д) От уровня значимости и от объѐма выборки.

6). Для интервала оценки необходимо знать:

1. X, t, S; 2. X, t, S, n; 3. X, S, , n; 4. X, t, т; 5.t, S, m

7). Характеристики положения случайной величины:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

8). Характеристики рассеяния:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

9). Сумма всех относительных частот равна

1. единице; 2. объему выборки; 3. 100; 4. 10; 5. 0.

Типовые задачи.

1) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases}0, \text{при } x \le 0\\ \sin x, \text{при } 0 < x \le \pi/2\\ 1, \text{при } x > \pi/2\end{cases} \qquad F(X)=\begin{cases}0, \text{при } x \le 0\\ \sin 2x, \text{при } 0 < x \le \pi/4\\ 1, \text{при } x > \pi/4\end{cases}$$

Найти плотность распределения f (x).

4) Дана функция распределения непрерывной случайной величины X

Найти плотность распределения f (x).

5) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases}0, \text{при } x \le 0\\ ax^2, \text{при } 0 < x \le 2\\ 1, \text{при } x > 2\end{cases} \qquad f(X)=\begin{cases}0, \text{при } x \le 0\\ \cos x, \text{при } 0 < x \le \pi/2\\ 0, \text{при } x > \pi/2\end{cases}$$

Найти коэффициент a и плотность вероятности попадания случайной величины X в интервал (0; 1). 4) Задана плотность распределения непрерывной случайной величины X:

Найти функцию распределения F(x).

5) Непрерывная случайная величина X задана плотностью распределения

f (x) =3sin3x в интервале (0; π /3); вне этого интервала f (x) = 0 . Найти вероятность того, что X примет значение, принадлежащее интервалу (π/6, π/4).

10) Случайная величина X задана плотностью распределения f(x)=2x в интервале [0,1]. Вне этого интервала f(x)=0. Найти математическое ожидание, и среднее квадратичное отклонение величины X.

11) Случайная величина X задана плотностью распределения f(x)=x/2 в интервале [0,2]; вне этого интервала. f(x) = 0. Найти M(X) и D(X).

12) Случайная величина X задана плотностью распределения $f(x) = C(x^2+2x)$ в интервале [0,1]; вне этого интервала. f(x)=0. Найти: а) параметр C; б) математическое ожидание величины X.

13) Математическое ожидание нормально распределенной случайной величины X равно μ = 3 и среднее квадратическое отклонение σ=2. Написать плотность вероятности X.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \dfrac{\overline{XY} - \overline{X}\,\overline{Y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx}\dfrac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy}\dfrac{\sigma_y}{\sigma_x}$.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5 - 15\cdot 2{,}35}{\sqrt{50}\cdot 0{,}7025} = \frac{40{,}5 - 35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

7.4. Самостоятельная работа обучающихся под контролем преподавателя.

Задача. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей.

7.5. Контроль конечного уровня усвоения темы:
Подготовка к выполнению практических приѐмов по теме занятия.
Материалы для контроля усвоения темы: набор тестовых заданий.
Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.
8. Литература (см. приложение) (см. приложение):

1.Тема занятия № 6 и ее актуальность Прямые измерения. Погрешности прямых измерений Методы оценки случайных погрешностей косвенных измерений.

Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.

2. Учебные цели:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен знать:

- прямые измерения
- погрешности прямых измерений.
- методы оценки случайных погрешностей косвенных измерений

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**

- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:

19. Понятие генеральной совокупности; методы исследования генеральной совокупности;
20. Выборочная совокупность; репрезентативность и объем выборки;
21. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
22. Построение графиков вариационных рядов;
23. Понятие об ошибках репрезентативности;
24. Доверительный интервал и доверительная вероятность;
25. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка.
26. Уравнение регрессии
27. Коэффициенты корреляции

4. **Вид занятия**: практическое занятие.

5. **Продолжительность занятия**:6 часа.

6. **Оснащение:**

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

1). Статистическое распределение - это:

1. Совокупность вариант; 2. Относительная частота вариант; 3. Совокупность вариант и соответствующих им частот; 4. Совокупность данных и закон их распределения; 5. Совокупность данных.

2). Распределение может быть:

1. Одномодальным; 2. Двумодальным; 3. Многомодальным; 4. Среди ответов 1-3 нет верного;

2) Построить эмпирическую функцию и ее график по данным

❍

$$F^{¿}(x)=\begin{cases} 0 \quad при \quad x \le 0; \\ 0{,}2 \quad при \quad 0 < x \le 2; \\ 0{,}4 \quad при \quad 2 < x \le 3; \\ 0{,}7 \quad при \quad 3 < x \le 4; \\ 0{,}8 \quad при \quad 4 < x \le 10; \\ 1 \quad при \quad x > 10; \end{cases}$$

3) Дана выборка значений случайной величины X объема 20:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12

18, 17, 15, 13, 17, 14, 14, 13, 14, 16

Требуется: - построить дискретный вариационный ряд;

- найти размах варьирования R, моду M_0, медиану M_e;

- построить полигон частостей.

5). Величина нормированного отклонения (распределение Стьюдента) зависит от а) Уровня значимости; б) Выборочного среднего; в)Ошибки выборочного среднего;

г) Объема выборки; д) От уровня значимости и от объѐма выборки.

6). Для интервала оценки необходимо знать:

1. X, t, S; 2. X, t, S, n; 3. X, S, , n; 4. X, t, т; 5.t, S, m

7). Характеристики положения случайной величины:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

8). Характеристики рассеяния:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

9). Сумма всех относительных частот равна

1. единице; 2. объему выборки; 3. 100; 4. 10; 5. 0.

Типовые задачи.

1) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases} 0, \text{при } x \le 0 \\ \sin x, \text{при } 0 < x \le \pi/2 \\ 1, \text{при } x > \pi/2 \end{cases} \qquad F(X)=\begin{cases} 0, \text{при } x \le 0 \\ \sin 2x, \text{при } 0 < x \le \pi/4 \\ 1, \text{при } x > \pi/4 \end{cases}$$

Найти плотность распределения f (x).

6) Дана функция распределения непрерывной случайной величины X
Найти плотность распределения f (x).

7) Дана функция распределения непрерывной случайной величины X

0, при x 0 F X ax, при 0 ^{2} x 2 1, при x

$$f(X)=\begin{cases} 0, \text{при } x \le 0 \\ \cos x, \text{при } 0 < x \le \pi/2 \\ 0, \text{при } x > \pi/2 \end{cases}$$

2 f X

Найти коэффициент a и плотность вероятности попадания случайной величины X в интервал (0; 1). 4) Задана плотность распределения непрерывной случайной величины X:

Найти функцию распределения F(x).

5) Непрерывная случайная величина X задана плотностью распределения

f (x) =3sin3x в интервале (0; π /3); вне этого интервала f (x) = 0 . Найти вероятность того, что X примет значение, принадлежащее интервалу (π/6, π/4).

14) Случайная величина X задана плотностью распределения f(x)=2x в интервале [0,1]. Вне этого интервала f(x)=0. Найти математическое ожидание, и среднее квадратичное отклонение величины X.

15) Случайная величина X задана плотностью распределения f(x)=x/2 в интервале [0,2]; вне этого интервала. f(x) = 0. Найти M(X) и D(X).

16) Случайная величина X задана плотностью распределения $f(x) = C(x^2+2x)$ в интервале [0,1]; вне этого интервала. $f(x)=0$. Найти: а) параметр C; б) математическое ожидание величины X.

17) Математическое ожидание нормально распределенной случайной величины X равно $\mu = 3$ и среднее квадратическое отклонение $\sigma=2$. Написать плотность вероятности X.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \dfrac{\overline{xy} - \overline{x}\,\overline{y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx}\dfrac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy}\dfrac{\sigma_y}{\sigma_x}$.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5 - 15\cdot 2{,}35}{\sqrt{50\cdot 0{,}7025}} = \frac{40{,}5 - 35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

7.4. Самостоятельная работа обучающихся под контролем преподавателя.

Задача. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей.

7.5. Контроль конечного уровня усвоения темы:

Подготовка к выполнению практических приѐмов по теме занятия.
Материалы для контроля усвоения темы: набор тестовых заданий.
Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.
8. Литература (см. приложение) (см. приложение):

1.Тема занятия № 7 и ее актуальность Анализ зависимостей (корреляции, ассоциации). Коэффициент корреляции Пирсона. Коэффициент корреляции Спирмена. Методы регрессионного анализа. Множественная линейная регрессия. Множественная нелинейная регрессия. Бинарная логистическая регрессия.

Актуальность изучения статистики обусловлена тем, что статистические представления являются важнейшей составляющей интеллектуального багажа современного медика. Они нужны и для повседневной жизни в современном цивилизованном обществе, и для продолжения образования практически во всех сферах человеческой деятельности, особенно, в медицине.

2. Учебные цели:

- закрепить основные понятия математической статистики; научиться выполнять первичный статистический анализ выборочных данных.

Для формирования профессиональных компетенций студент должен знать:

- прямые измерения
- погрешности прямых измерений.
- методы оценки случайных погрешностей косвенных измерений

Для формирования профессиональных компетенций обучающийся должен **владеть и уметь:**

- применять метод наименьших квадратов для нахождения коэффициентов линейного и квадратического тренда;
- составлять уравнения линейного и квадратического тренда;
- применять методы сглаживания временного ряда;
- составлять линейную и квадратическую модель прогноза

и овладеть следующими **компетенциями:** способность и готовность использовать в познавательной и профессиональной деятельности базовые знания в области математики и естественных наук, применять методы математического анализа и моделирования, теоретического и экспериментального исследования.

3. Материалы для самоподготовки к освоению данной темы: Вопросы для самоподготовки:

28. Понятие генеральной совокупности; методы исследования генеральной совокупности;
29. Выборочная совокупность; репрезентативность и объем выборки;
30. Группировка выборочных данных: построение ранжированного вариационного ряда распределения; составление равноинтервального ряда распределения;
31. Построение графиков вариационных рядов;
32. Понятие об ошибках репрезентативности;
33. Доверительный интервал и доверительная вероятность;
34. Оценка параметров генеральной совокупности по выборочным данным: точечная и интервальная оценка.
35. Уравнение регрессии
36. Коэффициенты корреляции

4. **Вид занятия**: практическое занятие.

5. **Продолжительность занятия**:6 часа.

6. **Оснащение:**

6.1. Дидактический материал: кино и видеофильмы, плакаты, интерактивная доска, компьютерная программа, таблицы, графики.

6.2. ТСО: Мультимедийный проектор, ноутбук, компьютеры.

7. Содержание занятия:

Задания для самоконтроля: решение студентами индивидуальных наборов тестовых заданий по теме:

7.1. Контроль исходного уровня знаний обучающихся с применением тестов.

1). Статистическое распределение - это:

1. Совокупность вариант; 2. Относительная частота вариант; 3. Совокупность вариант и соответствующих им частот; 4. Совокупность данных и закон их распределения; 5. Совокупность данных.

2). Распределение может быть:

1. Одномодальным; 2. Двумодальным; 3. Многомодальным; 4. Среди ответов 1-3 нет верного;

2) Построить эмпирическую функцию и ее график по данным

❍

$$F^{¿}(x)=\begin{cases} 0 \quad при \quad x \le 0; \\ 0{,}2 \quad при \quad 0 < x \le 2; \\ 0{,}4 \quad при \quad 2 < x \le 3; \\ 0{,}7 \quad при \quad 3 < x \le 4; \\ 0{,}8 \quad при \quad 4 < x \le 10; \\ 1 \quad при \quad x > 10; \end{cases}$$

3) Дана выборка значений случайной величины X объема 20:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12

18, 17, 15, 13, 17, 14, 14, 13, 14, 16

Требуется: - построить дискретный вариационный ряд;

- найти размах варьирования R, моду M_0, медиану M_e;

- построить полигон частостей.

5). Величина нормированного отклонения (распределение Стьюдента) зависит от а) Уровня значимости; б) Выборочного среднего; в)Ошибки выборочного среднего;

г) Объема выборки; д) От уровня значимости и от объѐма выборки.

6). Для интервала оценки необходимо знать:

1. X, t, S; 2. X, t, S, n; 3. X, S, , n; 4. X, t, т; 5.t, S, m

7). Характеристики положения случайной величины:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

8). Характеристики рассеяния:

1. Дисперсия; 2. Мода; 3. Среднее квадратическое отклонение; 4. Выборочное среднее;

5. Медиана.

9). Сумма всех относительных частот равна

1. единице; 2. объему выборки; 3. 100; 4. 10; 5. 0.

Типовые задачи.

1) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases}0,\text{при } x\le 0\\ \sin x,\text{при } 0<x\le \pi/2\\ 1,\text{при } x>\pi/2\end{cases}\qquad F(X)=\begin{cases}0,\text{при } x\le 0\\ \sin 2x,\text{при } 0<x\le \pi/4\\ 1,\text{при } x>\pi/4\end{cases}$$

Найти плотность распределения f (x).

8) Дана функция распределения непрерывной случайной величины X

Найти плотность распределения f (x).

9) Дана функция распределения непрерывной случайной величины X

$$F(X)=\begin{cases}0,\text{ при } x\le 0\\ ax^2,\text{ при } 0<x\le 2\\ 1,\text{ при } x>2\end{cases}\qquad f(X)=\begin{cases}0,\text{при } x\le 0\\ \cos x,\text{при } 0<x\le \pi/2\\ 0,\text{при } x>\pi/2\end{cases}$$

Найти коэффициент a и плотность вероятности попадания случайной величины X в интервал (0; 1). 4) Задана плотность распределения непрерывной случайной величины X:

Найти функцию распределения F(x).

5) Непрерывная случайная величина X задана плотностью распределения

f (x) =3sin3x в интервале (0; π /3); вне этого интервала f (x) = 0 . Найти вероятность того, что X примет значение, принадлежащее интервалу (π/6, π/4).

18) Случайная величина X задана плотностью распределения f(x)=2x в интервале [0,1]. Вне этого интервала f(x)=0. Найти математическое ожидание, и среднее квадратичное отклонение величины X.

19) Случайная величина X задана плотностью распределения f(x)=x/2 в интервале [0,2]; вне этого интервала. f(x) = 0. Найти M(X) и D(X).

20) Случайная величина X задана плотностью распределения $f(x) = C(x^2+2x)$ в интервале [0,1]; вне этого интервала. f(x)=0. Найти: а) параметр С; б) математическое ожидание величины X.

21) Математическое ожидание нормально распределенной случайной величины X равно μ = 3 и среднее квадратическое отклонение σ=2. Написать плотность вероятности X.

7.2. Разбор с преподавателем узловых вопросов, необходимых для освоения темы данного занятия.

Линейная регрессия, корреляция

Статистическая зависимость величины Y от величины X - это такая зависимость, при которой каждому значению величины X из множества ее возможных значений соответствует некоторое множество возможных значений величины Y, характеризуемое определенным законом распределения. Частным случаем статистической зависимости является корреляционная зависимость между величинами, когда изменение одной из величин влечет изменение математического ожидания другой.

Для полного описания корреляционной связи недостаточно найти форму корреляционной зависимости между величинами, необходимо еще определить силу этой зависимости по величине коэффициентов регрессии. Для количественной характеристики силы (тесноты) связи вводится понятие выборочного коэффициента линейной корреляции, определяемого отношением: $r = \dfrac{\overline{XY} - \overline{X}\,\overline{Y}}{\sigma_x \sigma_y}$, или $r = \rho_{yx}\dfrac{\sigma_x}{\sigma_y}$, или $r = \rho_{xy}\dfrac{\sigma_y}{\sigma_x}$.

Пример 4.5. Пользуясь данными примера 4.4, найти выборочный коэффициент корреляции между массой таблетки и скоростью ее растворения.

Решение: воспользовавшись формулой (4.22), найдем коэффициент r:

$$r = \frac{40{,}5-15\cdot 2{,}35}{\sqrt{50\cdot 0{,}7025}} = \frac{40{,}5-35{,}25}{\sqrt{35{,}125}} = \frac{5{,}25}{5{,}93} = 0{,}885$$

7.3. Демонстрация преподавателем методики практических приемов по данной теме.

7.4. Самостоятельная работа обучающихся под контролем преподавателя.

Задача. В книге "Основы химии" Д.И. Менделеева приводятся данные о растворимости азотнокислого натрия в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ (Y) при соответствующих температурах (X) раствора:

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Постройте корреляционное поле.

Предполагая, что зависимость между X и Y близка к линейной, найдите выборочный коэффициент парной корреляции и оцените достоверность выборочного значения коэффициента парной корреляции. Найдите уравнения линейной регрессии Y на X и X на Y. Постройте линии регрессии. Используя соответствующие уравнения регрессии, найдите, при какой температуре раствориться 100 условных частей $NaNO_3$ и какова температура раствора $NaNO_3$, если растворилось 84 условных частей.

7.5. Контроль конечного уровня усвоения темы:
Подготовка к выполнению практических приѐмов по теме занятия.
Материалы для контроля усвоения темы: набор тестовых заданий.
Место проведения самоподготовки: читальный зал, учебная комната для самостоятельной работы обучающихся, компьютерный класс.
Учебно-исследовательская работа обучающихся по данной теме (проводится в учебное время): работа с основной и дополнительной литературой.

8. Литература (см. приложение) (см. приложение):

ПРИЛОЖЕНИЕ

Перечень основной и дополнительной литературы, необходимой для освоения учебной дисциплины

	Основная литература	
	Обмачевская, С. Н. Медицинская информатика. Курс лекций : учебное пособие для вузов / С. Н. Обмачевская. — 3-е изд., стер. — Санкт-Петербург : Лань, 2021. — 184 с. — ISBN 978-5-8114-7053-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/154391	Неограниченный доступ
	Зарубина, Т. В. Медицинская информатика : учебник / Зарубина Т. В. [и др.] - Москва : ГЭОТАР-Медиа, 2018. - 512 с. - ISBN 978-5-9704-4573-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970445730.html	Неограниченный доступ
	Омельченко, В. П. Информатика, медицинская информатика, статистика : учебник / В. П. Омельченко, А. А. Демидова. - Москва : ГЭОТАР-Медиа, 2021. - 608 с. - ISBN 978-5-9704-5921-8. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970459218.html	Неограниченный доступ
	Царик, Г. Н. Информатика и медицинская статистика / под ред. Г. Н. Царик - Москва : ГЭОТАР-Медиа, 2017. - 304 с. - ISBN 978-5-9704-4243-2. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970442432.html	Неограниченный доступ
	Дополнительная литература	
	Диденко Г. А. Теоретические основы медицинской информатики / Г. А. Диденко, А. А. Мукашева, О. А. Степанова. - Челябинск : ЮУГМУ, 2017. - 175 с. - Текст : электронный // ЭБС "Букап" : [сайт]. - URL : https://www.books-up.ru/ru/book/teoreticheskie-osnovy-medicinskoj-informatiki-15045004/	Неограниченный доступ
	Медицинская информатика : учебное пособие / Н. В. Маркина, Г. А. Диденко, А. А. Мукашева и др. - Челябинск : ЮУГМУ, 2017. - 145 с. - Текст : электронный // ЭБС "Букап" : [сайт]. - URL : https://www.books-up.ru/ru/book/medicinskaya-informatika-11851468/	Неограниченный доступ
	Медицинская информатика: параметрические и непараметрические методы статистики на компьютере / Н. В. Маркина, Э. И. Беленкова, Г. А. Диденко и др. - Челябинск : ТЕТА, 2022. - 138 с. - Текст : электронный // ЭБС "Букап" : [сайт]. - URL : https://www.books-up.ru/ru/book/medicinskaya-informatika-parametricheskie-i-neparametricheskie-metody-statistiki-na-kompyutere-15440733/	Неограниченный доступ
	Семенова О. Л. Медицинская информатика: в 2 ч. Часть 1 : учебное пособие / О. Л. Семенова, Н. Ю. Часовских, А. Ю. Гречишникова. - Томск : Издательство СибГМУ, 2021. - 79 с. - Текст : электронный // ЭБС "Букап" : [сайт]. - URL : https://www.books-up.ru/ru/book/medicinskaya-informatika-v-2-ch-chast-1-12564392/	Неограниченный доступ
	Статистические методы в медицине и здравоохранении [Электронный ресурс] : учеб. пособие / ФГБОУ ВО «Баш. гос. мед. ун-т» МЗ РФ ; сост. Н. Х. Шарафутдинова [и др.]. - Электрон. текстовые дан. - Уфа, 2018. - Текст: электронный // БД «Электронная учебная библиотека» .- **URL:**	Неограниченный доступ

	http://library.bashgmu.ru/elibdoc/elib719.pdf	
	Таллер В. А. Медицинская информатика / В. А. Таллер. - Витебск : ВГМУ, 2019. - 225 с. - ISBN 9789854669809. - Текст : электронный // ЭБС "Букап" : [сайт]. - URL : https://www.books-up.ru/ru/book/medicinskaya-informatika-12137206/	Неограниченный доступ
	ЭБС "Букап"	https://www.books-up.ru/
	Электронно-библиотечная система «Консультант студента» для ВПО	www.studmedlib.ru
	База данных «Электронная учебная библиотека»	http://library.bashgmu.ru

Перечень ресурсов информационно-телекоммуникационной сети «интернет», необходимых для освоения учебной дисциплины

1. https://www.medicinform.net/ (Медицинская информационная сеть)
2. https://www.studentlibrary.ru/ (Консультант студента)