



Направление подготовки
31.05.02 Педиатрия (врач педиатр)

Учебный План утвержден решениями Ученого совета НГМУ
Протокол №3 от 17.04.2018 г.:

Учебная дисциплина
Б1.Б.12 МАТЕМАТИКА



КАФЕДРА
МАТЕМАТИКИ
НГМУ



ЛЕКЦИОННОЕ ЗАНЯТИЕ

Тема:

*Основные понятия математической статистики.
Статистические оценки параметров распределения.*



**КАФЕДРА
МАТЕМАТИКИ**
НИГУ



МАТЕМАТИКА

Рабочая программа дисциплины

(лекционные занятия)

Раздел 1. Теория вероятностей

8 час

1	Тема-1.1 Введение в теорию вероятностей. Логические операции над множествами. Элементы комбинаторики. Вероятность события - определения, основные свойства и формулы вычисления.	2 час
2	Тема-1.2 Алгебра событий. Теоремы сложения и умножения вероятностей. Формула полной вероятности и формула Байеса. Формула Бернулли и формула Пуассона.	2 час
3	Тема-1.3 Случайные величины и их числовые характеристики. Основные законы распределения дискретных случайных величин.	2 час
4	Тема-1.4 Непрерывные случайные величины и их числовые характеристики. Основные законы распределения непрерывных случайных величин. Правило трех сигм.	2 час

Раздел 2. Математическая статистика

4 час

5	Тема-2.1 Основные понятия математической статистики. Статистические оценки параметров распределения.	2 час
6	Тема-2.2 Основные понятия теории статистических гипотез. Проверка статистических гипотез. Основы теории корреляции.	2 час



Тема - 2.1

План лекционного занятия

(лекционное занятие)

1	Основные понятия математической статистики. <ul style="list-style-type: none">• Генеральная и выборочная совокупности• Репрезентативная и представительная выборки• Способы и виды отбора• Вариативный ряд, размах варьирования• Относительная частота• Статистическое распределение выборки• Полигон и гистограмма выборки	30 мин
2	Числовые характеристики выборки.	30 мин
3	Оценка параметров генеральной совокупности по ее выборке.	30 мин



1. Основные понятия математической статистики

Математическая статистика – это раздел математики, изучающий способы сбора статистической информации и методы ее обработки.

В математической статистике выделяют два основных направления исследований:

- 1. Оценка параметров генеральной совокупности.*
- 2. Проверка статистических гипотез.*



1. Основные понятия математической статистики

• Генеральная и выборочная совокупности

Генеральная совокупность (N) – это множество всех изучаемых объектов.

Выборочная совокупность, выборка (n) – это часть генеральной совокупности, выбранная некоторым (случайным) образом.

Объемом совокупности (генеральной или выборочной) называют число объектов этой совокупности.



1. Основные понятия математической статистики

- Генеральная и выборочная совокупности

Пример: из десяти тысяч студентов отобрано для обследования 100 человек.

Объем генеральной совокупности $N = 10\ 000$

Объем выборки $n = 100$



1. Основные понятия математической статистики

• Репрезентативность случайной выборки

Выборка должна быть **репрезентативной**, то есть давать правильное представление о пропорциях генеральной совокупности.

Выборка будет **репрезентативной**, если её осуществить случайно, так как все объекты имеют одинаковую вероятность попасть в выборку.



1. Основные понятия математической статистики

• Представительная выборка

В силу закона больших чисел можно утверждать, что **выборка будет представительной**, если:

- её *объем* достаточно велик, а её *значения* независимы;
- выборка сделана *случайно* и все её объекты имеют одинаковую вероятность попасть в выборку.

Примечание: Если объем совокупности достаточно велик, а выборка составляет лишь незначительную часть этой совокупности, то различие между повторными и безповторными выборками **практически исчезает**.



1. Основные понятия математической статистики

• Способы отбора

Способы отбора

Отбор не требующий разбиения генеральной совокупности на части

Отбор требующий разбиения генеральной совокупности на части

Простой случайный бесповторный отбор

Простой случайный повторный отбор

Механический отбор

Типический отбор

Серийный отбор



1. Основные понятия математической статистики

• Виды отбора

Простой случайный отбор

- Потеря
- Таблицы случайных чисел
- Генерация случайных чисел на ПК

Механический отбор

Генеральную совокупность "механически" делят на столько групп, сколько объектов должно войти в выборку, а из каждой группы отбирают один объект.
Механический отбор может не обеспечить репрезентативности выборки.

Типический отбор

Объекты отбираются не из всей генеральной совокупности, а из каждой ее "типической" части.

Используется тогда, когда обследуемый признак заметно колеблется в различных типических частях генеральной совокупности *например, если продукция изготавливается на нескольких машинах, среди которых есть более и менее изношенные, то типический отбор целесообразен.*

Серийный отбор

Объекты отбирают из генеральной совокупности не по одному, а сериями, которые подвергаются сплошному обследованию.

Используются тогда, когда обследуемый признак мало колеблется в различных сериях.

На практике часто применяется комбинированный отбор, при котором сочетаются указанные выше способы. Например, иногда разбивают генеральную совокупность на серии одинакового объема, затем простым случайным отбором выбирают несколько серий и, наконец, из каждой серии простым случайным отбором извлекают отдельные объекты.



1. Основные понятия математической статистики

- Вариативный ряд, размах варьирования

Вариантой x_j называют каждый элемент выборки.

Вариативный ряд – это последовательность вариантов, записанных *в возрастающем порядке*.

Частота встречаемости n_j – это частота (число наблюдений) варианты.

Размах варьирования – это величина $R = x_{\max} - x_{\min}$



1. Основные понятия математической статистики

• Относительная частота

Относительная частота – это отношение частоты к объему выборки

$$W_i = \frac{n_i}{n}$$



1. Основные понятия математической статистики

- Статистическое распределение выборки

Статистическое распределение выборки – это перечень вариантов и соответствующих им частот (или относительных частот)

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k



1. Основные понятия математической статистики

- Статистическое распределение выборки

Пример:

Задано распределение частот выборки объема 20.

Написать распределение относительных частот.

x_i	2	5	9
n_j	3	10	7

Решение:

$$W_1 = \frac{3}{20} = 0,15 \quad W_2 = \frac{10}{20} = 0,5 \quad W_3 = \frac{7}{20} = 0,35$$

x_i	2	5	9
W_j	0,15	0,5	0,35

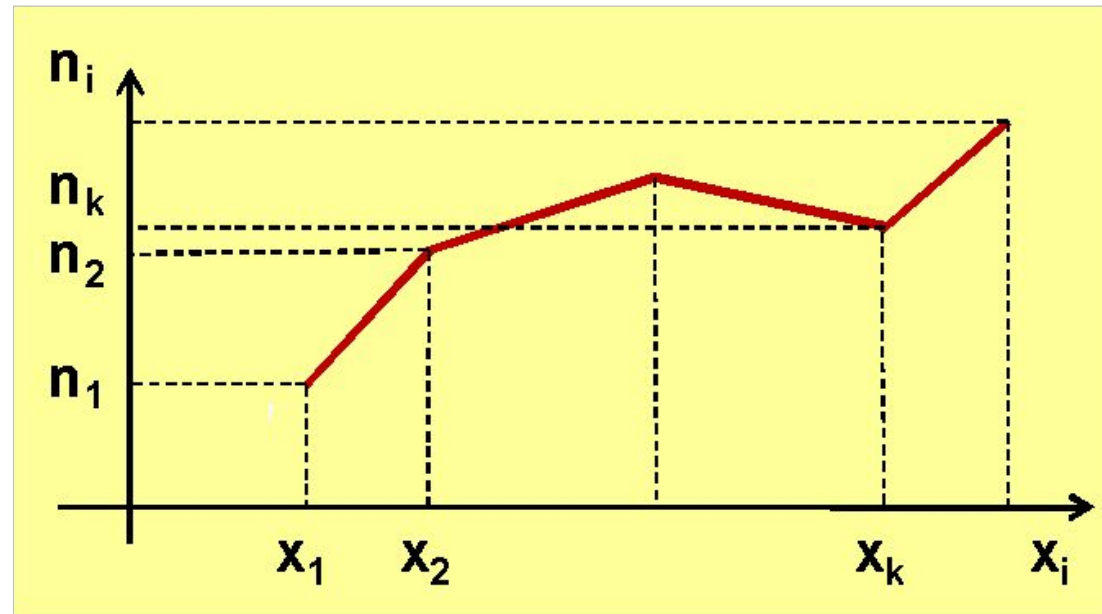
Контроль: $0,15 + 0,5 + 0,35 = 1$



1. Основные понятия математической статистики

• Полигон и гистограмма выборки

Полигон частот выборки – это ломаная линия, соединяющая точки $(x_1; n_1), (x_2; n_2), \dots, (x_k; n_k)$



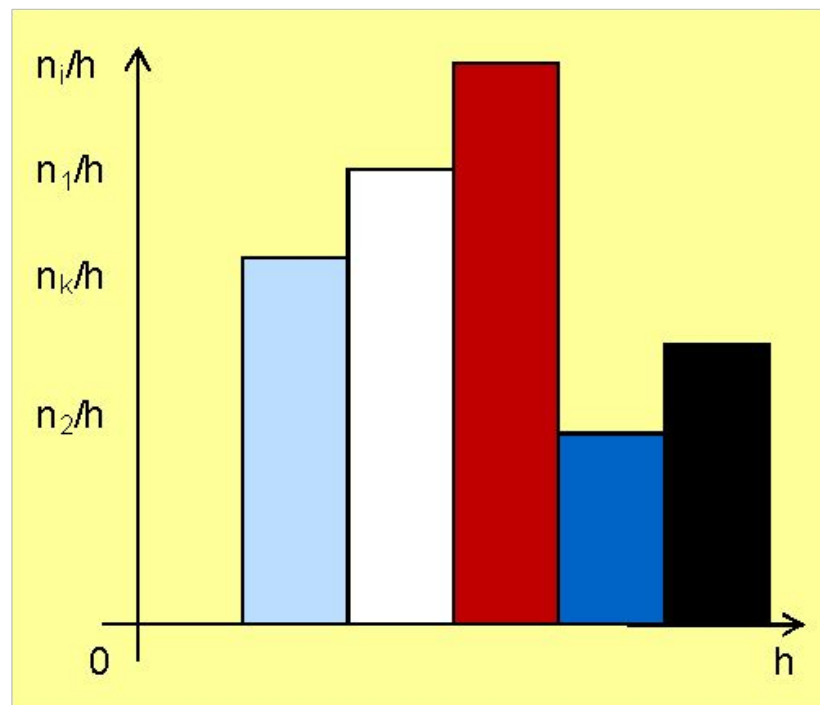
Полигон относительных частот выборки – это ломаная линия, соединяющая точки $(x_1; W_1), (x_2; W_2), \dots, (x_k; W_k)$



1. Основные понятия математической статистики

• Полигон и гистограмма выборки

Гистограмма частот – это ступенчатая фигура, состоящая из прямоугольников, основаниями которых являются частичные интервалы длины h , а высоты равны отношению n_i к h (плотность частоты n_i/h).





Тема - 2.1

План лекционного занятия

(лекционное занятие)

1	Основные понятия математической статистики.	30 мин
2	Числовые характеристики выборки. <ul style="list-style-type: none">• Характеристики положения, модальность• Модальность интервальных рядов• Медиана• Выборочная средняя• Выборочная дисперсия• Среднее квадратическое отклонение• Коэффициент вариации	30 мин
3	Оценка параметров генеральной совокупности по ее выборке.	30 мин



2. Числовые характеристики выборки

ХАРАКТЕРИСТИКИ ВЫБОРКИ

Характеристики
положения

Характеристики
рассеяния
вариант вокруг
своего среднего



2. Числовые характеристики выборки

• Характеристики положения, модальность

Мода (M_0) – это такое значение варианты, что предшествующее и следующее за ним значения имеют меньшие частоты встречаемости.

Для одномодальных распределений мода – это наиболее часто встречающаяся варианта в данной совокупности.

Пример: Для нижеприведенного распределения $M_0 = 18$

x_i	16	17	18	20
n_i	5	1	20	6



2. Числовые характеристики выборки

• Модальность интервальных рядов

Мода интервальных рядов вычисляется по формуле:

$$M_0 = x_{\text{НИЖ}} + i \cdot \frac{n_2 - n_1}{2n_2 - n_1 + n_3}$$

Где:

$x_{\text{НИЖ}}$ - нижняя граница модального класса, т. е. класса с наибольшей частотой встречаемости;

n_2 - частота модального класса;

n_1 - частота класса, предшествующего модальному;

n_3 - частота класса, следующего за модальным;

i - ширина классового интервала.



2. Числовые характеристики выборки

• Модальность интервальных рядов

Пример: Найти моду распределения роста 1000 взрослых мужчин.

Рост	143-146	146-149	149-152	152-155	155-158	158-161	161-164	164-167	167-170	170-173	173-176	176-179	179-182	182-185	185-188
n	1	2	8	26	65	120	181	201	170	120	64	28	10	3	1

Решение:

$$M_0 = 3 \cdot \frac{201 - 181}{2 \cdot 201 - 181 + 170} = 164,21$$



2. Числовые характеристики выборки

• Медиана

Медиана (M_E) – это значение признака, относительно которого ряд распределения делится на 2 равные по объему части.

Пример: Дано распределение

12	14	16	18	20	22	24	26	28
----	----	----	----	----	----	----	----	----

Медианой данного распределения является центральная варианта, то есть $M_E=20$, так как по обе стороны от нее стоит по 4 варианты.

Для ряда с четным числом членов медианой будет полусумма его центральных членов, например:

6	8	10	12	14	16	18	20	22	24
---	---	----	----	----	----	----	----	----	----

$$M_E = (14 + 16) : 2 = 15$$



2. Числовые характеристики выборки

• Выборочная средняя

Выборочная средняя – это среднее арифметическое значение вариантов статистического ряда

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i n_i$$



2. Числовые характеристики выборки

- Выборочная дисперсия

Выборочная дисперсия – это среднее арифметическое квадратов отклонения вариантов от их среднего значения.

$$D_B = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot n_i$$

$$D_B = \frac{1}{n} \sum_{i=1}^k x_i^2 \cdot n_i - (\bar{x}_B)^2$$



2. Числовые характеристики выборки

• Среднее квадратическое отклонение

Среднее квадратическое отклонение – это квадратный корень из выборочной дисперсии

$$\sigma_B = \sqrt{D_B}$$



2. Числовые характеристики выборки

• Коэффициент вариации

Коэффициент вариации – это отношение среднего квадратического отклонения к средней величине признака, выраженное в процентах.

$$CV = \frac{\sigma_B}{\bar{x}_B} \cdot 100 \%$$

Коэффициент вариации – это мера относительной изменчивости случайной величины.

То есть можно сравнивать разнородные величины, например, частоту сердечных сокращений, артериальное давление и температуру в единых единицах – процентах.



Тема - 2.1

План лекционного занятия

(лекционное занятие)

1	Основные понятия математической статистики.	30 мин
2	Числовые характеристики выборки.	30 мин
3	Оценка параметров генеральной совокупности по ее выборке. <ul style="list-style-type: none">• Способы статистического оценивания• Точечная оценка генеральной совокупности• Несмещенная и смещенная оценка• Стандартное отклонение, стандартная ошибка среднего• Интервальная оценка, доверительный интервал и вероятность• Доверительный интервал выборки малого объема нормального распределения	30 мин



3. Оценка параметров генеральной совокупности

Смысл статистических методов заключается в том, чтобы по выборке ограниченного объема, т. е. по некоторой части генеральной совокупности, высказать обоснованное суждение, о ее свойствах в целом.

Числовые значения, характеризующие генеральную совокупность, называются параметрами.

Одна из задач математической статистики – определение параметров большого массива по исследованию его части.



3. Оценка параметров генеральной совокупности

• Способы статистического оценивания

Способы статистического оценивания

Точечная оценка

Оценка, которая дается для некоторой определенной точки.

Интервальная оценка

По данным выборки оценивается интервал, в котором лежит истинное значение с заданной вероятностью.



3. Оценка параметров генеральной совокупности

- Точечная оценка генеральной совокупности

Точечная оценка – это оценка, которая определяется одним числом. Это число определяется по выборке. Это функция результатов выборки.

Качество точечного оценивания

Состоятельная оценка

Если при увеличении объема выборки **выборочная характеристика** стремится к соответствующей характеристике генеральной совокупности

Эффективная оценка

Если она имеет **наименьшую дисперсию** выборочного распределения по сравнению с другими аналогичными оценками

Несмещенная оценка

Если ее **математическое ожидание** равно оцениваемому параметру при любом объеме выборки.



3. Оценка параметров генеральной совокупности

- Несмещенная и смещенная оценка

Несмещенной оценкой генеральной средней (математического ожидания) служит выборочная средняя

$$\bar{X}_{ГЕН} \cong \bar{X}_B$$

Выборочная дисперсия не обладает свойством несмещенности. **Это смещенная оценка генеральной дисперсии.**

На практике используют исправленную выборочную дисперсию, которая является несмещенной оценкой дисперсии генеральной совокупности:

$$S^2 = \frac{n \cdot D_B}{n - 1}$$

$$D_{ГЕН} \cong S^2$$



3. Оценка параметров генеральной совокупности

- Стандартное отклонение, стандартная ошибка среднего

Кроме того, в расчетах используют S - *исправленное среднее квадратическое отклонение*, называемое **стандартным отклонением**.

$$\sigma_{ГЕН} \cong S$$

и ошибку выборочной средней (**стандартную ошибку среднего**)

$$m_{\bar{x}} = \frac{S}{\sqrt{n}}$$



3. Оценка параметров генеральной совокупности

- Интервальная оценка, доверительный интервал и вероятность

Интервальная оценка – это числовой интервал, содержащий неизвестный параметр генеральной совокупности.

Доверительный интервал – это интервал, в котором с той или иной заранее заданной вероятностью находится неизвестный параметр генеральной совокупности.

Доверительная вероятность p - это такая вероятность, что событие вероятности **$1 - p$** можно считать невозможным.

$\alpha = 1 - p$ – это уровень значимости.

Обычно, в качестве *доверительных вероятностей* используют вероятности близкие к 1: $P \approx 0,95$; $P \approx 0,99$; $P \approx 0,999$. Тогда событие, что интервал накроет характеристику, будет практически достоверным.



3. Оценка параметров генеральной совокупности

- Доверительный интервал выборки малого объема

Доверительный интервал для выборки малого объема
нормально распределенного количественного признака X
может иметь вид:

$$\bar{x}_B - m_{\bar{x}} t < \bar{x}_Г < \bar{x}_B + m_{\bar{x}} t$$

где t - нормированный показатель распределения
Стьюдента, с $(n-1)$ степенями свободы, который определяется
вероятностью попадания генерального параметра в данный
интервал.



3. Оценка параметров генеральной совокупности

Пример: Найти точечные оценки генеральной совокупности по данному статистическому распределению выборки.

x_i	3-5	5-7	7-9	9-11
n_i	5	10	20	15

Построить гистограмму частот. С надежностью 95% указать доверительный интервал для оценки генерального среднего.

Решение: Дан интервальный ряд распределения. Составим дискретный ряд, находя середины интервалов.

x_i	4	6	8	10
n_i	5	10	20	15



3. Оценка параметров генеральной совокупности

Продолжения решения. Сделаем следующие вычисления.

$$\bar{x}_B = \frac{4 \cdot 5 + 6 \cdot 10 + 8 \cdot 20 + 10 \cdot 15}{50} = 7,8 \quad \Rightarrow \quad \bar{x}_G \cong 7,8$$

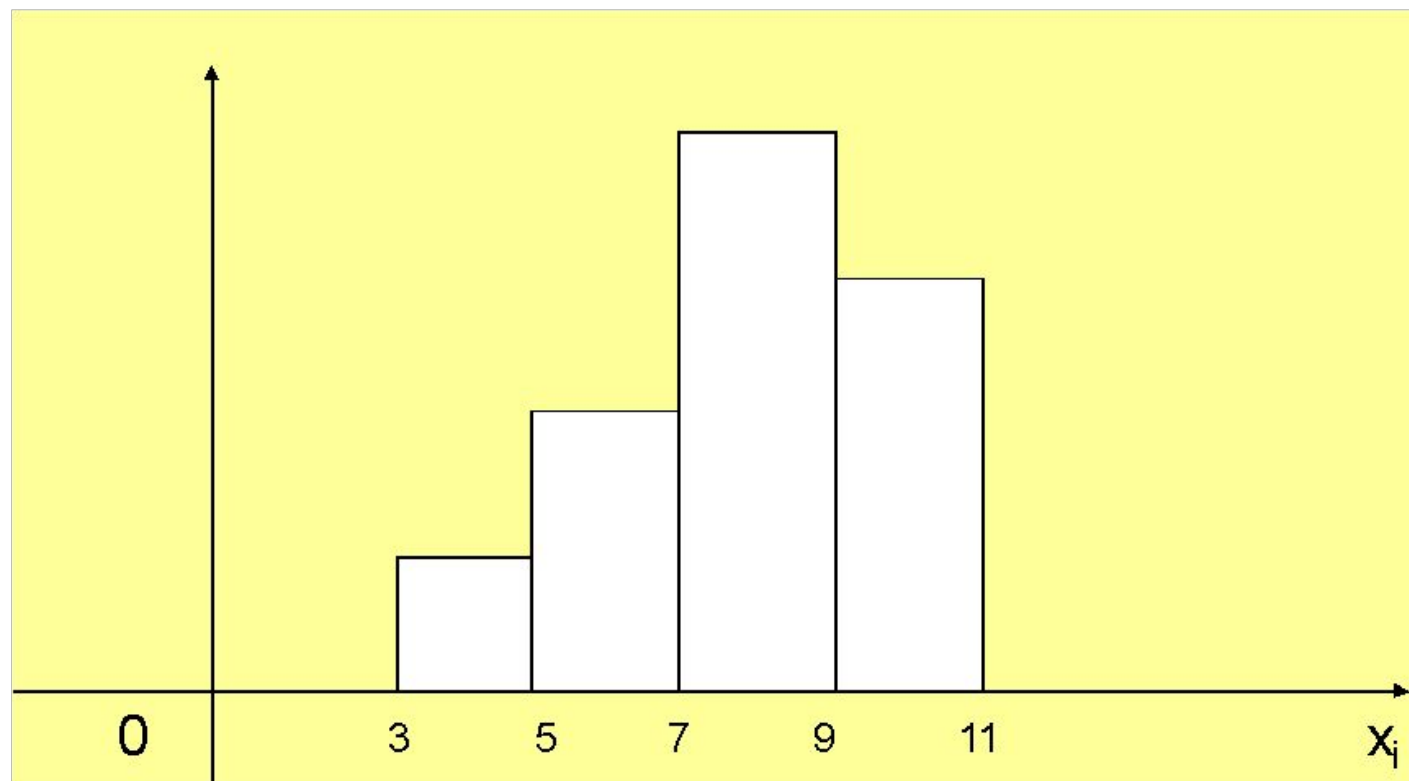
$$D_B = \frac{4^2 \cdot 5 + 6^2 \cdot 10 + 8^2 \cdot 20 + 10^2 \cdot 15}{50} - (7,8)^2 = 3,56$$

$$S^2 = \frac{50 \cdot 3,56}{50 - 1} = 3,63 \quad \Rightarrow \quad D_G \cong 3,63$$

$\frac{n_i}{h}$ 

3. Оценка параметров генеральной совокупности

Продолжения решения. Построим гистограмму частот.





3. Оценка параметров генеральной совокупности

Продолжения решения. Вычислим ошибку средней.

$$m_{\bar{x}} = \frac{1,906}{\sqrt{50}} = 0,27$$

Найдем значение нормированного показателя распределения Стьюдента .

$$t = t(\gamma; n) = t(0,95; 50) = 2,009$$

$$7,26 < \bar{x}_T < 8,34$$



СПАСИБО ЗА ВНИМАНИЕ