

Лекция 2

# Компьютерный анализ медицинских данных

*К.п.н., преподаватель  
кафедры ФММИ  
Арзуманян Н.Г.*

# Виды медико-биологических данных:

- Электромагнитные сигналы;
- Звуковые сигналы;
- Механические сигналы.

Информация о:

- Росте;
- Массе тела;
- Составе крови и др. биологических жидкостей;
- Жалобы больного;
- Лихорадка;
- Желтуха и т.д.

Сигналы могут быть  
зарегистрированы приборами  
либо поступать в мозг  
человека.

**Зарегистрированные  
сигналы в информатике  
называются данными.**

# Медико-биологические данные

## 1. Количественные данные (параметры):

- Рост пациента;
- Концентрация в крови ферментов;
- Заболеваемость;
- Количество вич-инфицированных и т.д.

## 2. Качественные данные (признаки):

- Цвет кожных покровов;
- Наличие болей;
- Качество жизни человека;

# Медико-биологические данные

## 3. Статические картины органов человека или всего тела:

- Патологические изменения на рентгенограмме грудной клетки;
- Изображение головного мозга на компьютерной томограмме и т.д.

## 4. Динамические картины органов человека:

- Регистрация движущихся органов: (сердце, легкие);
- Изучение быстроменяющихся картин прохождения по организму контрастных веществ.

# Медико-биологические данные

## 5. Динамические данные физиологических функций:

- Электрокардиограмма;
- Электроэнцефаллограмма;
- Кривые,  
зарегистрированные при  
прохождении  
радиоактивного вещества  
по организму и т.д.

# Оценка медико-биологических данных

## Признак

- Это характеристика, которая может иметь только 2 значения: наличие или отсутствие.
- Примеры: наличие или отсутствие болей, лихорадки, покраснения кожных покровов.

## Параметр

- Это величина, характеризующая свойство процесса, явления или системы в абсолютных или относительных величинах.
- Примеры: показатели температуры тела и давление, концентрации в крови отдельных веществ и т.д.

# Шкалы измерения

## Шкала наименований

- Группировка объектов и их производных в ряд непересекающихся классов.
- Например: симптомы и синдромы, цвет кожных покровов и т.д.

## Шкала порядка

- Это упорядоченная шкала наименований, на которой отражена, в основном, тенденция процесса.
- Например, концентрация гормонов, степень желтушности и т.д.



# Шкалы измерения

## Интервальная шкала

- Это шкала с наличием единицы измерения.
- Примеры: температура термометра.

## Шкала отношений

- Это интервальная шкала с нулевой точкой, т.е. имеющей такую точку, в которой данный параметр отсутствует.
- Пример: ростомер, весы, линейка.

# Этапы операций с медико-биологическими данными

## 1. Сбор и первичная обработка данных.

- это накопление результатов исследований в том объеме, который задан условиями поставленной задачи или необходимостью принять адекватное решение.

## 2. Оценка эффективности измерения данных

- это определение степени точности и величины погрешности зарегистрированных сигналов и полученных данных.

# Этапы операций с медико-биологическими данными

## 3. Сохранение данных

- это регистрация данных в виде твердых копий или на магнитных носителях.

## 4. Формализация и стандартизация данных

- это сведение всех полученных данных к единой форме, которая должна соответствовать требованиям компьютерной обработки и обеспечивать сопоставимость всех данных между собою

# Этапы операций с медико-биологическими данными

## 5. Фильтрация и очищение данных

- это отсеивание лишних сигналов, обусловленных неточностью работы регистрирующих приборов, некорректно собранной информацией о состоянии изучаемого явления.

## 6. Кодировка данных

- это унификация формы представления данных на бумажных или магнитных носителях.

# Этапы операций с медико-биологическими данными

## 7. Сортировка данных

- это упорядочение данных по заданному признаку или совокупности их характеристик

## 8. Преобразование данных

- это изменение формы данных по заданному алгоритму или между различными типами носителей.

# Этапы операций с медико-биологическими данными

## 9. Сжатие и архивация данных

- уплотнение данных на носителях и организация их хранения, нередко связана с изменением их формы

## 10. Защита данных

- приведение данных по специальному алгоритму к форме, которая недоступна для несанкционированного их использования

# Этапы операций с медико-биологическими данными

## 11. Транспортировка данных

- это передача данных на расстояния с помощью механических или телекоммуникационных каналов связи.

# Критерии оценки эффективности методов измерения медико- биологических данных

## Точность измерений

- это соответствие результатов измерения истинному значению определяемой величины.

## Правильность измерений

- это качество измерения характеризует величину систематических погрешностей



# Критерии оценки эффективности методов измерения медико- биологических данных

## Воспроизводимость измерений

- этот критерий показывает, как близки между собою будут результаты измерений, выполненных в различных условиях.

## Сходимость измерений

- данное качество измерения характеризует величину случайных ошибок. Чем они меньше, тем лучше сходимость измерения.

Показатели измерения	Нет ошибок	Рандомизированные ошибки	Систематические ошибки
«Охота за мишенью»			
Точность (достоверность)	Очень хорошая	Удовлетворительная	Плохая
Сходимость	Хорошая	Плохая	Очень хорошая

Рис.1.1. «Охота за мишенью». Связь между достоверностью (точностью) и сходимостью результатов, рандомизированными и систематическими ошибками исследований

# Математическая статистика

- это наука, изучающая методы обработки результатов наблюдений массовых случайных явлений, обладающих статистической устойчивостью, закономерностью с целью выявления этой закономерности по исследованию части этого массива данных.

# Основные задачи математической статистики:

- 1) задача нахождения закона распределения случайной величины по наблюдаемым данным;
- 2) задача нахождения параметров распределения;
- 3) проверка согласованности теории с данными опыта;
- 4) задача установления и исследования различного рода зависимостей на основании экспериментальных данных.

# Генеральная совокупность и выборка

## Генеральная совокупность

- Совокупность всех исследуемых объектов

## Выборка

- совокупность случайно отобранных объектов из генеральной совокупности.

Характеристики:

- Объем;
- Репрезентативность;
- Размах выборки.

# Статистическое распределение (вариационный ряд)

Статистическое распределение – это совокупность вариантов и соответствующих им частот.

$x_i$
$n_i$

-варианта

- частота встречаемости

Пример: Рост 175 см встретился 5 раз;  
рост 168 см – 7 раз; 180 см – 8 раз.

Вариационный ряд - это та же самая выборка, но расположенная в порядке возрастания элементов.

Пример: 168 см – 7 раз; 175 см – 5 раз;  
180 см – 8 раз.

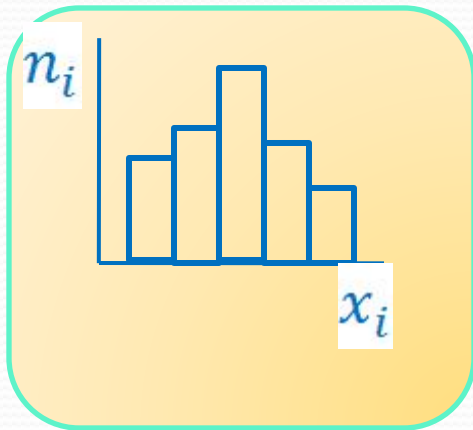
# Пример

- 10 студентов получили следующие оценки на коллоквиуме по физике: 3, 3, 3, 5, 4, 2, 4, 5, 3, 4.
- Составим вариационный ряд:

Оценка (варианта)	Частота	Относительная частота
2	1	0,1
3	4	0,4
4	3	0,3
5	2	0,2

# Гистограмма

Гистограмма – это *ступенчатая* фигура, состоящая из смежных *прямоугольников*, построенных на одной прямой, основания которых одинаковы и равны ширине класса, а высоты равны относительной частоте.



вариационный размах

$$i = \frac{x_{\max} - x_{\min}}{1 + 3,32 \lg n}$$

Формула  
Стерджеса

Ширина класса



# Точечная оценка

- это выборочная характеристика, используемая в качестве **приближенного** значения неизвестной генеральной характеристики.
- Определяется одним числом (точкой на числовой оси).
- Выборка должна быть **большого** объема.
- Дает лишь некоторое приближенное значение параметра.

## Основные величины:

- Генеральное среднее;
- Генеральная дисперсия;
- Исправленная дисперсия;
- Математическое ожидание.

# Интервальная оценка

- это числовой интервал, содержащий неизвестный параметр генеральной совокупности с заданной вероятностью.
- Определяется двумя числами – границами интервала.
- Более точная, надежная и информативная, так как дает информацию о степени близости к соответствующему теоретическому параметру.
- Используется, если выборка **малого** объема.

## Основные величины:

- Генеральное среднее;
- Генеральная дисперсия;
- Исправленная дисперсия;
- Математическое ожидание.



**Благодарю за внимание!**