

# Статистический анализ в медицине

**Жилина Наталья Михайловна**

# Для чего нужна статистика?

- Для *строгого доказательства эффективности* методов диагностики и лечения (какому проценту больных помогает лечение и в какой степени);
- Для *оценки эффективности работы врача* (сокращение пребывания больного в стационаре (не в ущерб пациенту) позволяет экономить значительные средства);
- Для *поиска новых методов диагностики и лечения, выбора наилучшего метода из существующих.*

# Этапы исследования:

- *Планирование и организация исследования;*
- *Проведение наблюдения (собственно исследование);*
- *Обработка и анализ данных, выводы, оформление результатов исследования.*

## Простейшие ошибки:

- **Отсутствие контрольной группы;**
- **Использование неслучайных выборок;**
- **Пренебрежение статистической проверкой гипотез.**

# Вероятностный характер

- В результате применения статистических методов мы получаем не истину в последней инстанции, а оценку вероятности того или иного предположения.
- Каждый статистический метод основан на собственной математической модели, и результаты его правильны настолько, насколько эта модель соответствует действительности.

# ВХОД→ПРОЦЕСС→ВЫХОД

Врач Иванов лечил n пациентов с исходами «выздоровление», «улучшение состояния», «без изменений», «ухудшение», «летальный исход».

Тогда описание *модели* имеет вид:

## На «входе»:

- экспериментатор (врач);
- единица наблюдения (пациент);
- ресурсы (можно рассмотреть всевозможные сочетания).

На «процессе» - лечение n больных (операции, процедуры и т.д.).

На «выходе» - различные исходы: «выздоровление», «улучшение состояния», «без изменений», «ухудшение», «летальный исход».

# Корректность модели:

Если у врача X больше благоприятных исходов, чем у врача Y, означает ли это, что врач Y – «хуже»? Мы не можем так утверждать, не выяснив **начальные условия**.

Что может повлиять на **исход лечения** (рассмотреть различные сочетания):

- **Состояние больного:** диагноз, тяжесть состояния больного, возраст, сопутствующие заболевания, запущенность случая.
- **Профессионализм врача:** уровень подготовки, опыт работы, контакт с пациентом и т.д.
- **Ресурсы:** наличие необходимых средств, материальная база – необходимая аппаратура, лекарства и т.д.

# Гипотезы:

- **Гипотеза** – предположение о сущности данного факта (или ряда фактов).
- Гипотеза, принятая исследователем называется **рабочей гипотезой**. Противоположная ей – **альтернативная гипотеза**.
- **Нулевая гипотеза** предполагает, что не существует значимого различия (например, между контрольной и экспериментальной группой). В статистике принято нулевую гипотезу считать **рабочей**, а ей противоположную – **альтернативной**.
- Гипотеза, имеющая большую вероятность (больше 95 или 99%) **статистически значима**.

## Примеры гипотез (Стругацкие. «Стажеры»):

- В древней пещере «первояпонцев» обнаружено множество небольших окаменевших следов босых ног, а в центре пещеры один след огромной рифленой подошвы ботинка. Дать объяснения возможных ситуаций. (Придумать ряд гипотез).

## Единицы и признаки наблюдения

- **Единицы наблюдения** – отдельные случаи изучаемого явления. Например, при исследовании заболеваемости раком, единицей наблюдения является **больной раком**.
- **Признаки наблюдения** – общие и важные характеристики, соответствующие цели исследования. Например: возраст больного, вес, давление, пульс, диагноз, осложнение, срок реабилитации и т.д.

# Типы признаков:

- **Количественные признаки** измеряются числовыми значениями (например, возраст, рост, вес, давление).
- **Порядковые признаки** – могут быть измерены в шкалах (например, школьные оценки, степень тяжести заболевания – легкая (1), средняя (2), тяжелая (3) и т.д.).
- **Качественные признаки** – характеризуют некоторое состояние объекта, но не могут быть измерены количественно (например, пол, профессия, диагноз).

## База данных (БД):

- БД – формализованная таблица, состоящая из единиц наблюдения с их признаками;
- **Важно:** в **одной** клетке таблицы – **одно** число (или запись)!!!
- Признаки наблюдения должны быть классифицированы

# Фрагмент БД в MS EXCEL:

Excel - запрос 200 человек

Файл Вид Вставка Формат Сервис Данные Окно Справка

Введите вопрос

Arial Cyr 10 Ж К Ч

РАЙОН

А	В	С	Д	Е	Ф	Г	Н
ФИО	Дата освидетельств	Дата рождения	Пол	Социальное положение	Диагноз основной ВТЭК	Диагноз сопутс.ВТЭК	Группа инва
А ТК	12.06.2004	03/31/1956	женский	служащий	E01.0	111.9	3
КЕВИЧ ТВ	11/22/2004	10.02.1956	женский	домохозяйка с полисом	E01.0	125.2	2
В ВЕ	10/28/2004	06/21/1957	мужской	неработающ. с полисом	E01.1		3
А НН	04/15/2004	12/24/1953	женский	служащий	E01.1	167.2	2
НА ОС	11.01.2004	08.02.1983	женский	домохозяйка с полисом	E03.1	G93.4	2
ОВА ДФ	03/25/2004	04.05.1981	женский	рабочий	E03.1		2
НЕВА СВ	05/24/2004	09/20/1966	женский	служащий	E03.9		3
ЦОВА НВ	09.06.2004	04.12.1950	женский	рабочий	E06.3	125.2	2
ЦОВА НМ	11.01.2004	04/15/1941	женский	пенсинер	E06.3		2
НВ	11/22/2004	11/20/1986	мужской	неработающ. с полисом	E10.2		3
В КС	03.09.2004	12.10.1980	мужской	служащий	E10.3		3
А	08/18/2004	01/18/1964	мужской	рабочий	E10.3		3
РМ	09.06.2004	02/19/1935	женский	пенсинер	E10.4		2
ВА	08/24/2004	04.02.1966	мужской	неработающ. с полисом	E10.5		2
ОВА ОВ	02.02.2004	02.11.1984	женский	студент ВУЗа	E10.5		3
СКИЙ МН	09/20/2004	12/15/1985	мужской	учащийся ГПТУ, техникума	E10.7		3
ОВА КС	10/20/2004	03/20/1936	женский	пенсинер	E10.7	125.2	2
ЛГ	06/15/2004	08/20/1940	женский	домохозяйка с полисом	E10.8		2
А ПА	04.08.2004	06/22/1976	женский	домохозяйка с полисом	E10.9		2

# **Основные типы статистических задач:**

- 1. Как сжато описать данные?**
- 2. Статистическая оценка значимости различий признаков в группах, проверка гипотез.**

<b>ПРИЗНАК</b>	<b>ИССЛЕДОВАНИЕ</b>				
	<b>Две группы</b>	<b>Более двух групп</b>	<b>Группа до и после лечения</b>	<b>Одна группа несколько видов лечения</b>	<b>Связь признаков</b>
<b>Количественный (нормальное распределение*)</b>	<b>Критерий Стьюдента</b>	<b>Дисперсионный анализ</b>	<b>Парный критерий Стьюдента</b>	<b>Дисперсионный анализ повторных измерений</b>	<b>Линейная регрессия, корреляция, или метод Блэнда-Алтмана</b>
<b>Качественный</b>	<b>Критерий <math>\chi^2</math> Z-критерий</b>	<b>Критерий <math>\chi^2</math></b>	<b>Критерий Мак-Нимара</b>	<b>Критерий Кокрена</b>	<b>Коэффициент сопряженности</b>
<b>Порядковый</b>	<b>Критерий Манна Уитни</b>	<b>Критерий Крускала Уоллиса</b>	<b>Критерий Уилкоксона</b>	<b>Критерий Фридмана</b>	<b>Коэффициент ранговой корреляции Спирмена</b>

# Принцип действия критериев:

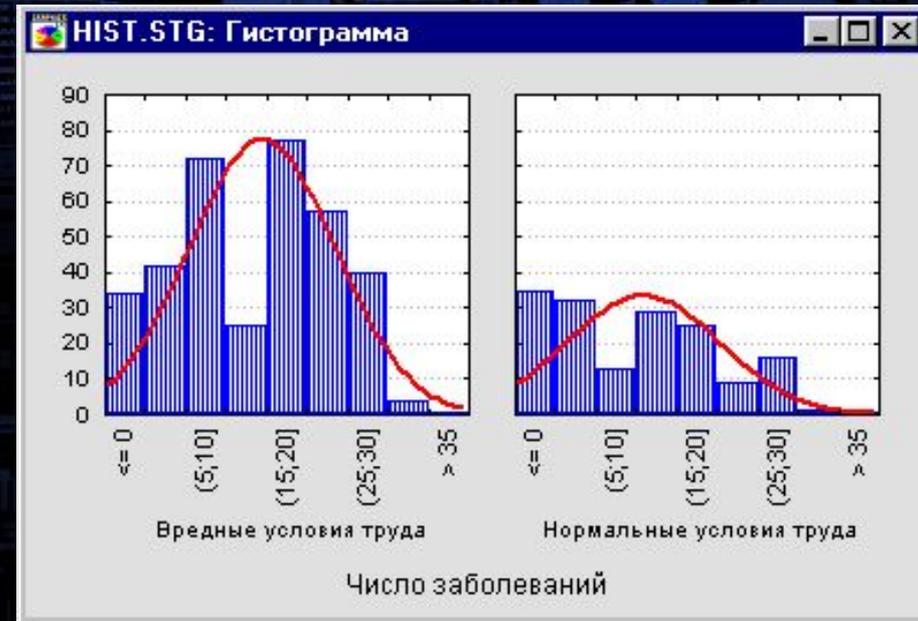
- Сравниваются нужные признаки в соответствующем виде эксперимента.
- Проверяется нулевая гипотеза. Находится фактическая вероятность ошибки **отклонить верную нулевую гипотезу ( $P$ )**. Говоря упрощенно,  $P$  – это вероятность справедливости нулевой гипотезы.
- Максимальную приемлемую вероятность отвергнуть нулевую гипотезу называют уровнем значимости и обозначают  $\alpha$ . Обычно в медико-биологических исследованиях принимают  $\alpha = 0.05$ .
- Если  $P < 0,05$  – нулевая гипотеза отвергается, следовательно найдено статистически значимое различие в сравниваемых группах.

# Как выбрать метод?

- Если Вы имеете дело с порядковыми и качественными признаками, то подходят только непараметрические методы.
- Если признак числовой, стоит подумать, нормально ли его распределение.
- Если данных мало (или Вы не хотите думать о типе распределения) - воспользуйтесь непараметрическими методами.

# Основные цели первичного (разведочного) анализа

Определение характера  
распределений  
переменных, визуальный  
анализ зависимостей и  
идентификация  
возможных выбросов.

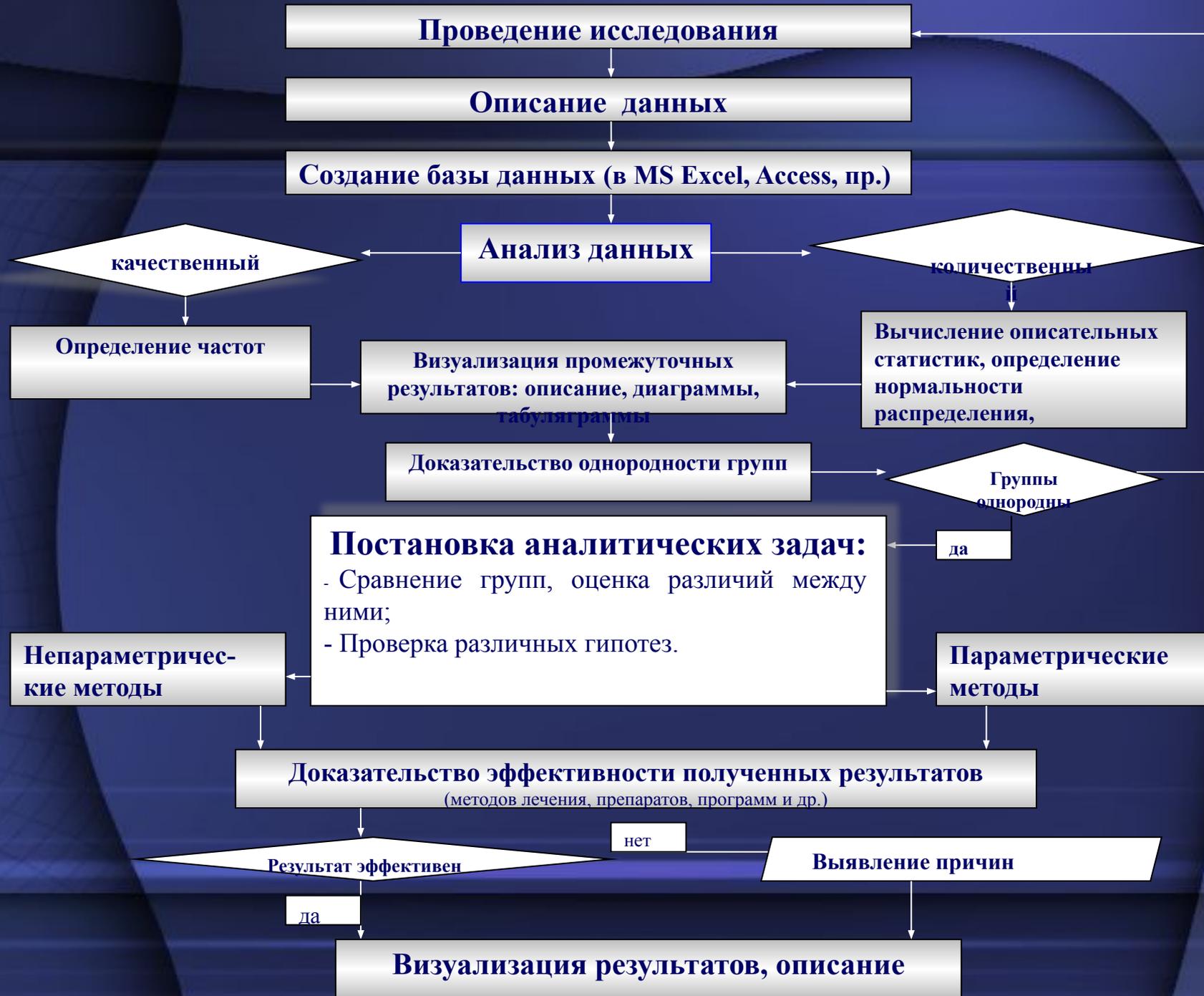


## Нормальное

Можно применять  
стандартные методы:  
t-критерии и  
дисперсионный анализ.

## Отличное от нормального (или малая выборка)

Необходимо использовать  
непараметрические критерии.



## Примеры статистических задач:

**Задача 1.** В группе из 20 человек, вакцинированных от гриппа, заболело 4 человека, а в группе из 20 не вакцинированных заболело 12. Определить, эффективна ли вакцинация?

# Таблица сопряженности:

Группы	Заболело		Не заболело	
	факт	ожд.	факт	ожд.
Не вакцинированные	факт	12	факт	8
	ожд.	8	ожд.	12
Вакцинированные	факт	4	факт	16
	ожд.	8	ожд.	12

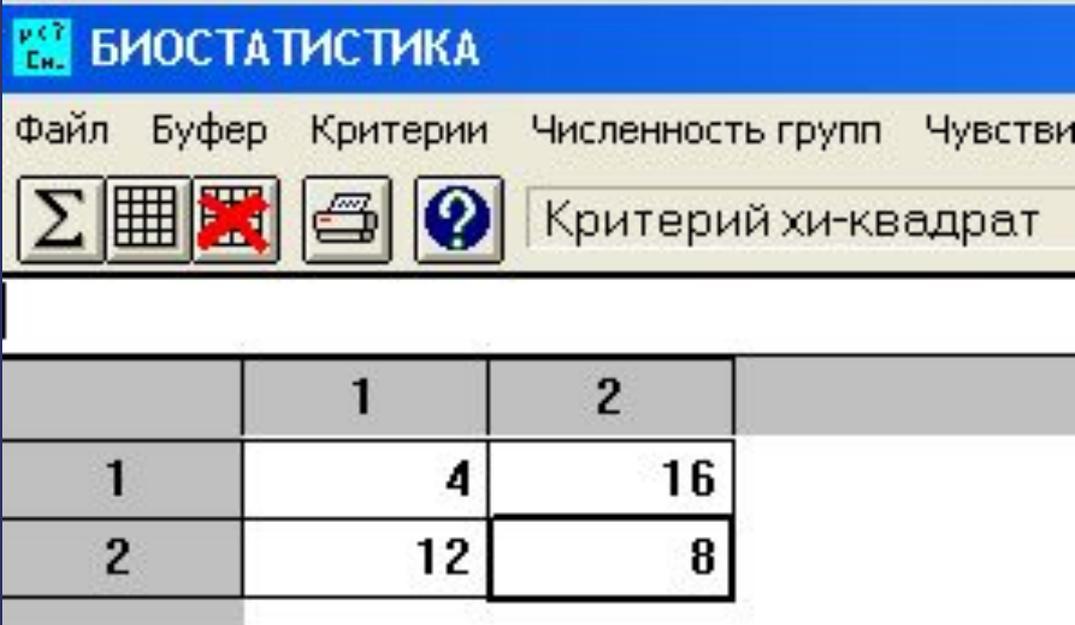
**Ожидаемые значения – при условии справедливости нулевой гипотезы**

# Решение задачи 1 в пакете Биостатистика:

Признак качественный («заболел» с возможными значениями «да» или «нет»), две независимых группы. Следовательно, необходимо воспользоваться критерием  $\chi^2$ .

**Важно:** в условии задачи дана численность всей группы и число заболевших.

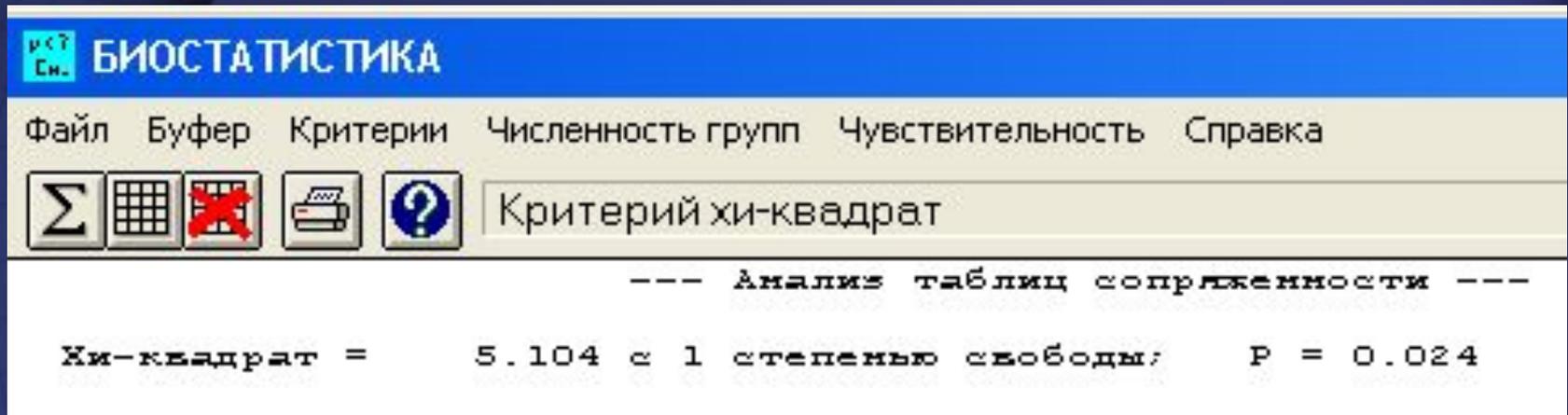
Во входную таблицу для пакета «Биостатистика» необходимо по группам ввести число *заболевших* и *не заболевших* (как разность между общей численностью группы и числом заболевших).



The screenshot shows the 'БИОСТАТИСТИКА' software interface. The title bar reads 'БИОСТАТИСТИКА'. The menu bar includes 'Файл', 'Буфер', 'Критерии', 'Численность групп', and 'Чувствительность'. Below the menu bar are several icons: a summation symbol ( $\Sigma$ ), a grid, a grid with a red 'X', a printer, and a question mark. A dropdown menu is open, showing 'Критерий хи-квадрат'. Below the menu bar is a 2x2 contingency table with the following data:

	1	2	
1	4	16	
2	12	8	

# Результат и интерпретация:



- Результат вычислений представлен на рисунке.
- **Интерпретация результата.** Нулевая гипотеза предполагает, что между числом заболевших в группах нет статистически значимого различия. В пакете «Биостатистика» вычислена ее вероятность  $P = 0,024$ .  $P < 0,05$ , следовательно, нулевая гипотеза отвергается. То есть, найдено статистически значимое различие по признаку «заболел гриппом» между группой вакцинированных и не вакцинированных. Значит **вакцинация эффективна.**

## Задача 2.

Существует ли связь между проницаемостью сосудов сетчатки (X) и электрической активностью сетчатки (Y)?

$X = \{19.5, 15.0, 13.5, 23.3, 6.3, 2.5, 13.0, 1.8, 1.6, 6.5, 1.8\};$

$Y = \{0.0, 38.5, 59.0, 97.4, 119.2, 129.5, 191.7, 248.7, 318.0, 438.5\}.$

## Решение задачи 2:

- Для того чтобы не думать, нормально ли распределение, решим задачу с помощью непараметрического метода, а именно критерия ранговой корреляции Спирмена.
- Выбираем нужный критерий в выпадающем меню критериев. Заполняем входную таблицу. Таблица входных данных представлена на рисунке



The screenshot shows the BIOSTATISTIKA software interface. The title bar reads "БИОСТАТИСТИКА". The menu bar includes "Файл", "Буфер", "Критерии", "Численность групп", "Чувствительность", and "Справка". The "Критерии" menu is open, showing "Коэффициент ранговой корреляции Спирмена". Below the menu is a table with 10 rows and 3 columns. The columns are labeled "X" and "Y". The data points are as follows:

	X	Y
1	19.5	0
2	15	38.5
3	13.5	59
4	23.3	97.4
5	6.3	119.2
6	2.5	129.5
7	13	191.7
8	1.8	248.7
9	6.5	318
10	1.8	438.5

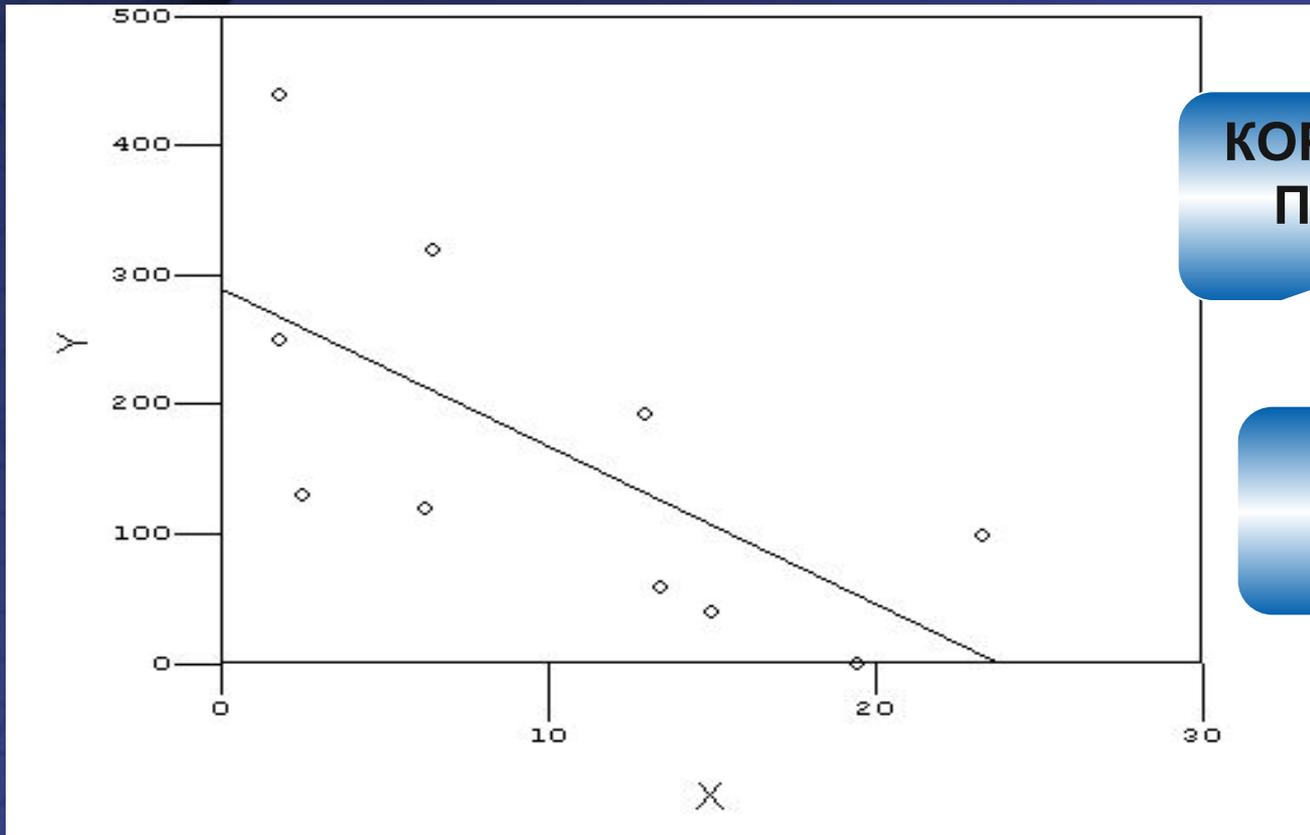
# Результат и интерпретация:

The screenshot shows the BIOSTATISTIKA software interface. The title bar reads "БИОСТАТИСТИКА". The menu bar includes "Файл", "Буфер", "Критерии", "Численность групп", "Чувствительность", and "Справка". The toolbar contains icons for summation, grid, deletion, printing, and help. The main window title is "Коэффициент ранговой корреляции Спирмена". The output text is as follows:

```
--- Коэффициент ранговой корреляции Спирмена ---  
  
n = 10  
Коэффициент ранговой корреляции Спирмена = -0.779  
P = 0.010
```

- **Интерпретация результата.** Нулевая гипотеза, предполагающая, что связь между признаками статистически незначима, имеет вероятность  $P = 0.01$ .
- Таким образом, нулевая гипотеза отвергается и **выявлена сильная обратная связь** между проницаемостью сосудов сетчатки и электрической активностью сетчатки.
- Величину связи характеризует коэффициент ранговой корреляции Спирмена ( $r_s$ ),  $r_s = -0.78$ . Знак «минус» указывает на обратный тип связи, то есть **чем выше значение одного признака, тем меньше значение другого.**

# Корреляция и линейная регрессия. Метод наименьших квадратов:



КОРРЕЛЯЦИОННОЕ  
ПОЛЕ (ОБЛАКО)

ЛИНИЯ  
РЕГРЕССИИ

Линия, сумма квадратов расстояний до которой из точек корреляционного поля минимальна называется линией регрессии.

Метод можно использовать для прогноза.

# Контрольные вопросы

**1.** Цель исследования определяется на этапе:

- планирования и организации исследования
- проведения наблюдения
- обработки и анализа данных
- оформления результатов

## 2. Объект наблюдения, это:

- отдельный случай изучаемого явления
- явление, подлежащее исследованию
- качественный признак исследования
- метод анализа данных

### **3.** Единица наблюдения, это:

- количественный признак наблюдения;
- явление, подлежащее исследованию;
- отдельный случай изучаемого явления;
- качественный признак наблюдения.

## 4. Сплошное исследование:

- изучает все единицы, входящие в объект наблюдения;
- не имеет цели;
- изучает микроорганизмы;
- изучает часть единиц объекта наблюдения.

## 5. Выборочное исследование:

- выбирает вид наблюдения;
- изучает все единицы, входящие в объект наблюдения;
- изучает часть единиц объекта наблюдения;
- имеет несвоевременно определенную цель.

**6.** Число единиц наблюдения должно **БЫТЬ:**

- очень большим;
- очень маленьким;
- оптимальным (не слишком малым, но и не неоправданно большим);
- равно 10.

## 7. Точность результата:

- приближение, с которым можно говорить о подлинности результата;
- порядковый признак наблюдения;
- ошибка, получаемая из-за несплошного характера исследования;
- метод исследования.

## 8. Научная гипотеза:

- рассказ об исследовании;
- метод анализа данных;
- предположение о сущности факта или ряда фактов;
- совпадает с целью исследования.

## 9. Что такое вероятность события $A$ ?:

- частота события  $A$  при достаточно большом числе экспериментов;
- метод изучения;
- наблюдение за событием  $A$ ;
- модель эксперимента.

## **10.** Статистические результаты, как правило:

- точные;
- вероятностные;
- невозможно описать;
- не встречаются в жизни.

# 11. Нулевая гипотеза:

- принимается в статистике в качестве рабочей гипотезы;
- принимается в качестве альтернативной гипотезы;
- имеет большую вероятность;
- имеет малую вероятность.

## 12. Критический уровень значимости:

- максимально приемлемая вероятность отвергнуть справедливую нулевую гипотезу (обычно в мед. исследованиях 0,05);
- вероятность альтернативной гипотезы;
- достижимая цель;
- вероятностное событие.

## **13.** Если условия экспериментов неоднородны:

- нельзя сравнивать их исходы;
- эксперименты надо объединять;
- исследования не следует проводить;
- это не влияет на результат сравнения ИСХОДОВ.

## 14. Количественный признак:

- выражается и измеряется числовыми значениями;
- измеряется в шкалах (ранжируется);
- не может быть измерен количественно;
- встречается только в медицине.

## 15. Качественный признак:

- выражается и измеряется числовыми значениями;
- измеряется в шкалах (ранжируется);
- не может быть измерен количественно;
- встречается редко.

## **16. Порядковый признак:**

- выражается числовыми значениями;
- измеряется в шкалах (ранжируется);
- не может быть измерен количественно;
- встречается только в жизни.

## 17. База данных:

- таблица, содержащая *единицы наблюдения* и характеризующие их *признаки*;
- полигон проведения эксперимента;
- оборудование для исследования;
- не нуждается в описании.

## 18. Статистика может:

- улучшить выборку;
- дать статистическое оценивание результатов исследования;
- Исправить ошибки в измерениях;
- Оценить неизвестные признаки.

## **19. Статистика не может:**

- Исправить ошибки в измерениях;
- провести статистическое оценивание;
- выполнить Проверку гипотез;
- провести статистическое моделирование.

## **20.** Возможные проблемы статистической обработки:

- сравнение групп;
- некорректное использование статистических методов;
- вычисление описательных статистик;
- расчет частот качественных признаков.

## **21.** Возможная статистическая ошибка:

- вычисление описательных статистик количественных признаков;
- расчет частот качественных признаков;
- использование неслучайных выборок;
- правильная интерпретация данных.

## **22.** Основные описательные статистики количественного признака, это:

- среднее, стандартное отклонение, ошибка среднего, процентиля (нижний квантиль, медиана, верхний квантиль);
- нулевая гипотеза;
- цель исследования;
- тип признака.

## 23. Распределение признака близко к нормальному, если:

- выборка представительна;
- среднее признака близко к медиане (различаются не более, чем на 20%) и в интервал "среднее плюс-минус ст.откл. попадает до 70% значений признака;
- описательные статистики не вычисляются;
- распределение равномерно.

**24.** Параметрические методы применяют ТОЛЬКО для анализа:

- качественных признаков;
- порядковых признаков;
- нерепрезентативных выборок;
- количественных признаков с нормальным распределением.

## **25.** Непараметрические методы применяют для анализа:

- объема выборки;
- качественных, порядковых признаков и количественных, если распределение не является нормальным;
- разброса признака вокруг среднего;
- нерепрезентативных выборок.

## **26. Рандомизированное исследование, это:**

- исследование со случайно отобранной контрольной группой;
- ретроспективное исследование;
- проспективное исследование;
- только основная группа наблюдения.

## **27.** Группы независимы, если:

- единицы наблюдения основной и контрольной групп различны;
- исследуются одни и те же больные до и после лечения;
- исследуется одна группа больных с несколькими видами лечения;
- вид исследования не определен.

**28.** При сравнении нескольких независимых групп с нормальным распределением признака нужно применять:

- критерий Стьюдента;
- дисперсионный анализ;
- корреляцию качественных признаков;
- z-критерий.

## **29.** Поправка Бонферрони используется:

- в критерии Хи-квадрат;
- в корреляционном анализе;
- при применении критерия Стьюдента, если число групп больше 2-х;
- при расчете чувствительности.

**30.** Можно ли применять парный критерий Стьюдента для независимых выборок?

- да;
- нет;
- если выборка ретроспективна;
- если выборка репрезентативна.

**31.** Если сравнивают две независимые группы с качественным дихотомическим признаком:

- применяют критерий Хи-квадрат с поправкой Йейтса;
- применяют критерий Фридмана;
- вычисляют описательные статистики;
- определяют коэффициент корреляции Пирсона.

## **32.** Чувствительность критерия проверяется:

- если различия в группах выявлены;
- если найденные различия статистически незначимы;
- если группы очень велики по объему;
- если исследователю грустно.

**33.** Если различий не выявлено при чувствительности **90%**, значит:

- на самом деле различия есть;
- надо применять другие методы;
- надо правильно определить цель исследования;
- на самом деле различия в группах не значимы.

## 34. Линейная регрессия

применяется:

- для вычисления прогнозных значений количественных признаков с нормальным распределением;
- нахождения различия в двух группах связанных выборок;
- нахождения различия в трех группах независимых выборок;
- для вычисления частот качественных признаков.

## **35. Экспорт данных:**

- перевод данных из одной компьютерной программы в другую;
- выезд за границу;
- ввоз товаров из-за рубежа;
- получение данных по электронной почте.

## 36. Что означает репрезентативность выборки:

- идет сравнение "коров с курицами»;
- однородность выборки;
- выборка отражает свойства основной совокупности (данные извлечены случайным образом);
- исследование проспективно.

**37.** В статистическом анализе наиболее важно:

- уметь корректно поставить задачу исследования, выбрать нужный метод;
- освоить современный инструмент компьютерного анализа;
- знать формулы расчета;
- завуалировать недостатки данных.

**38.** Связь между степенью тяжести послеоперационного осложнения и временем восстановительного периода в группе оперированных можно найти с помощью:

- критерия Стьюдента;
- линейной корреляции Пирсона;
- дисперсионного анализа;
- корреляции Спирмена.

## **39.** От чего зависит выбор статистического критерия:

- от типа признака и вида исследования;
- от возможности исследователя;
- от знаний руководителя;
- от настроения и времени года.

## 40. Метод многокритериального выбора "Анализ иерархий" можно применить:

- для определения цели исследования;
- для вычисления описательных статистик;
- для выбора оптимального решения из нескольких возможных альтернатив;
- для расчета частот качественных признаков.

**41.** Какая из программ наиболее мощная для проведения статистического анализа:

- MS Excel;
- Биостатистика;
- Instat;
- SPSS

**42.** Нормально ли распределение:  
**1,1,1,1,1,1,1,1,1,1,5,10,10:**

- да;
- нет;
- не знаю;
- сложный вопрос.

**43.** Чему равен критический уровень значимости ( $p$ ) в классическом медицинском исследовании, если для трех групп корректно применен критерий Стьюдента с поправкой Бонферрони:

- $p=0,05$ ;
- $p=0,01$ ;
- $p=0,05:3=0,017$ ;
- $p=0,1$

**44.** Какой метод статистического анализа следует применить для сокращения числа переменных (редукции данных)?

- факторный анализ;
- расчет описательных статистик;
- z-критерий;
- критерий Фридмана.

## 45. Что значит «выявлены статистически значимые различия признака в группах сравнения»?

- уровень значимости различия  $p < 0,05$ ;
- значения признаков визуально близки;
- значения признаков не сравнимы;
- сравниваются только две группы.

**46. Можно ли применить критерий Стьюдента к сравнению признака «рост» с признаком «вес»?**

- можно;
- проверив нормальность распределения одного из признаков;
- при нормальном распределении обоих признаков;
- нельзя.

## 47. Формула критерия Стьюдента

- работает на разности рангов признаков;
- основана на разности средних значений признака в группах сравнения;
- включает в себя коэффициент корреляции;
- рассчитывается для таблицы сопряженности.

## 48. Таблица сопряженности, это

- число возможных сочетаний числа выборок и значений признака;
- возможные значения уровня значимости различия;
- формализованное представление нулевой гипотезы;
- описание графического представления.

49. Что такое «ожидаемое значение признака» в таблице сопряженности?

- фактическое значение признака;
- критическая величина уровня значимости;
- значение признака при условии выполнения нулевой гипотезы;
- значение признака при отрицании нулевой гипотезы.

## 50. Что такое «наблюдаемое значение признака» в таблице сопряженности?

- фактическое значение признака;
- критическая величина уровня значимости;
- значение признака при условии выполнения нулевой гипотезы;
- значение признака при отрицании нулевой гипотезы.

## 51. Критерий хи-квадрат нельзя применять, если

- таблица сопряженности имеет размерность больше, чем  $2 \times 2$ ;
- сравниваются признаки в независимых группах;
- ожидаемое число в ячейке таблицы сопряженности меньше 5;
- сравниваются качественные признаки.

## 52. Точный критерий Фишера применяется

- если находится связь признаков;
- если вы хотите построить график;
- ожидаемое число в ячейке таблицы сопряженности меньше 5;
- для множественного сравнения.

## **53. Признак называется дихотомическим, если он**

- принимает одно из двух возможных значений;
- количественный;
- входит в основную группу;
- входит в группу сравнения.

## 54. Если сравнивают качественный дихотомический признак в двух независимых группах:

- применяют критерий хи-квадрат с поправкой Йейтса;
- применяют критерий Фридмана;
- вычисляют описательные статистики;
- определяют коэффициент корреляции Пирсона.

## 55. Коэффициент корреляции

- всегда положительный;
- всегда отрицательный;
- по модулю больше 1;
- определяет степень связи признаков.

## 56. Корреляционная связь признаков прямая, если

- признаки независимы;
- признаки порядковые;
- признаки качественные;
- с увеличением (уменьшением) значений одного признака, увеличиваются (уменьшаются) значения другого.

## 57. Корреляционная связь признаков обратная, если

- с уменьшением значений одного признака, увеличиваются значения другого;
- признаки независимы;
- признаки порядковые;
- признаки качественные.

**58.** Корреляция сильная, если коэффициент корреляции

- больше единицы;
- меньше -1;
- по модулю больше или равен 0,7;
- близок к нулю.

## 59. Корреляция слабая, если коэффициент корреляции

- по модулю больше или равен 0,7;
- близок к нулю, (по модулю меньше 0,3);
- больше единицы;
- меньше -1.

# 60. Линейная корреляция

## Пирсона применяется

- для определения связи двух рядов количественных признаков с нормальным распределением;
- для определения статистической значимости различия качественных признаков;
- для определения парных связей порядковых признаков;
- для множественного сравнения.

# 61. Ранговая корреляция Спирмена используется

- для определения парной связи количественных признаков с нормальным распределением;
- для определения статистической значимости различия качественных признаков;
- для определения парных связей порядковых признаков;
- для множественного сравнения.

## 62. Критерий Мак-Нимара применяется:

- для сравнения повторных измерений качественных признаков;
- для сравнения качественных признаков в независимых группах;
- для определения связи качественных признаков;
- для определения значимости различия независимых порядковых признаков.

63. Какие из приведенных данных о послеоперационном больном являются полными, а не цензурированными:

- послеоперационный больной погиб от несчастного случая;
- отказался от наблюдения;
- расторг договор страхования;
- выздоровел.

## 64. Что не требуется в анализе выживаемости?

- чтобы все данные были полными, а не цензурируемыми;
- для всех исследуемых известно время начала наблюдения;
- для всех исследуемых известно время окончания наблюдения и исход - полный случай или цензурируемый (выбывание);
- выбор наблюдаемых – случаен.

## 65. Что является невозможным исследованием при анализе **выживаемости**:

- изучение продолжительности жизни (исход – смерть);
- изучение срока лечения определенного заболевания (исход – ремиссия);
- Длительность лечения бесплодия или эффективность контрацепции (исход – беременность);
- исследование зависимости уровня сахара от веса пациента.

## 66. Биомедицинская статистика

- не является необходимой областью знания для врача;
- это инструмент для анализа экспериментальных данных и клинических наблюдений;
- слишком сложна для восприятия медицинского работника;
- не нужна в публикационных медицинских материалах.

# 67. «Температура больного» – относится к следующему типу переменной

- количественная;
- качественная;
- порядковая;
- дискретная.

## 68. Что такое «распределение признака»?

- область возможных значений признака;
- абсолютная или относительная частота встречаемости конкретных значений признака;
- тип признака;
- стандартное отклонение.

## 69. Медиана количественного признака

- делит распределение пополам (половина значений признака меньше медианы, половина больше);
- это сумма значений признака, деленная на объем выборки;
- значение признака с наибольшей частотой встречаемости;
- это некорректное утверждение.

## 70. Распределение не является нормальным, если:

- значение среднего обязательно близко к медиане (различие не более 20%);
- в интервал  $\pm\sigma$  попадает 68% значений признака;
- в интервал  $\pm 2\sigma$  попадает 95% значений признака;
- оно ассимметрично.

# 71. Медиана качественного признака

- делит распределение пополам (половина значений признака меньше медианы, половина больше);
- это сумма значений признака, деленная на объем выборки;
- значение признака с наибольшей частотой встречаемости;
- это некорректное утверждение.

## 72. Для чего нужны описательные статистики?

- для краткого описания большого массива количественных данных;
- для сжатого описания качественных признаков;
- для расчета коэффициента ранговой корреляции;
- для определения частоты признака.

## 73. Выборка является репрезентативной, если

- она отражает свойства основной совокупности, то есть данные случайно (равновероятно) извлечены из совокупности;
- если содержит только качественные признаки;
- если относится к социально-гигиеническим исследованиям;
- если состоит только из клинических наблюдений.

## 74. Что характеризует стандартное (среднеквадратичное) отклонение?

- уровень значимости различия признаков;
- разброс значений количественного признака с нормальным распределением от среднего арифметического (ширину нормального распределения);
- отклонение значения признака от медианы;
- степень связи двух признаков.

## 75. Какие характеристики хорошо описывают ассиметричное распределение?

- среднее и стандартное (среднеквадратичное) отклонение;
- среднее и стандартная ошибка среднего;
- коэффициенты корреляции;
- квантили (персентили).

## 76. Как лучше сохранить результаты эксперимента (наблюдений):

- в памяти;
- в рукописном описании;
- в виде базы данных на компьютере;
- в таблице на бумаге.

## 77. Чем не является дисперсия?

- средним квадратом отклонения от среднего арифметического количественного признака;
- характеристикой разброса признака;
- стандартным (среднеквадратичным) отклонением во второй степени;
- частотой признака.

**78. Для чего в критерии хи-квадрат (при таблице сопряженности  $2 \times 2$ ) применяется поправка на непрерывность (поправка Йейтса)?**

- чтобы задать вопрос соискателю на защите;
- для компенсации излишнего «оптимизма» (несколько ужесточает критерий);
- для усложнения вычислений;
- для популяризации ученого.

## 79. Нужно ли вычислять описательные статистики для качественных признаков?

- обязательно, чтобы сжато описать выборку;
- не обязательно;
- после проверки статистических гипотез;
- нет, следует вычислить их частоты.

## 80. Как выбрать статистический критерий для решения конкретной задачи?

- по типу признака и виду исследования;
- по объему выборки;
- по предпочтениям исследователя;
- по значениям описательных статистик.

## 81. Чем близки различные статистические критерии?

- имеют одинаковые формулы;
- созданы одним и тем же ученым;
- имеют сходный принцип действия: формулирование нулевой гипотезы, нахождение уровня значимости различия, сравнение его с критическим, вывод;
- одинаковыми видами исследования и типами признаков.

**82. Что выполняется раньше: проверка нормальности распределения количественного признака или критериальный анализ сравнения признаков в группах?**

- критериальный анализ;
- проверка нормальности распределения;
- не имеет значения;
- проверку на нормальность распределения можно не проводить.

## 83. Какой из критериев используется для проверки нормальности распределения?

- Колмогорова-Смирнова;
- линейной корреляции Пирсона;
- дисперсионный анализ;
- Мак-Нимара.

## 84. Нулевая гипотеза

- предполагает, что различия статистически значимы;
- стандартизует один из признаков, присваивая ему нулевое значение;
- это отсутствие всяческих предположений;
- предполагает, что различия значений признака в сравниваемых группах статистически незначимы.

## 85. Различия признака в сравниваемых группах статистически значимы, если

- нулевая гипотеза отвергнута ошибочно;
- вероятность ошибки отвергнуть справедливую нулевую гипотезу меньше 5% ( $p < 0,05$ );
- если значения признаков в группах сравнения достаточно близки;
- если нулевая гипотеза принимается.

86. Если корректно примененный критерий не нашел статистически значимых различий ( $p > 0,05$ ):

- значит различий на самом деле нет;
- нужно проверить чувствительность критерия;
- критерий не заслуживает доверия;
- статистика правды не скажет.

## 87. Если чувствительность критерия низкая:

- нужно попытаться увеличить объем выборки, проанализировать выбросы;
- ничего не нужно делать;
- применить другие критерии;
- сказать, что различия статистически не значимы.

**88. Если коэффициент корреляции близок к 1, но  $p > 0,05$ , это означает:**

- связь между признаками сильная, но результат не является закономерным;
- связь не является сильной;
- различия между признаками статистически значимы;
- различия незначимы.

## 89. Что такое «нижний квартиль»?

- это медиана;
- это 25-й перцентиль: значение признака, которое делит распределение на 25% и 75%, четверть значений меньше нижнего квартиля, три четверти – больше;
- это десятый перцентиль;
- это 75-й перцентиль.

## 90. Что такое «верхний квартиль»?

- это 25-й перцентиль;
- это медиана;
- это 75-й перцентиль: значение признака, которое делит распределение на 75% и 25%, т.е. три четверти значений меньше верхнего квартиля, четверть – больше;
- это десятый перцентиль.

## 91. Однофакторный дисперсионный анализ (ANOVA)

- определяет значимость различия количественного признака с нормальным распределением в нескольких несвязных группах;
- ищет связи порядковых признаков;
- сравнивает качественные признаки в связанных группах;
- сравнивает качественные признаки в независимых группах.

**92. С помощью какого критерия можно сравнить вес исследуемых в четырех возрастных группах?**

- парный критерий Стьюдента;
- критерий Мак-Нимара;
- критерий Крускала-Уоллиса;
- дисперсионный анализ.

## 93. Что важнее в статистическом анализе?

- корректно применять статистические критерии;
- знать формулы вычисления;
- иметь самую новую версию статистического пакета;
- знать иностранные языки.

## 94. Что не требуется при описании материалов исследования?

- описание гипотезы, подлежащей проверке;
- описание данных и способа их получения из основной совокупности (включая метод рандомизации);
- перечень статистических методов оценки гипотезы;
- семейное положение исследователя.

## 95. Что не является уровнем значимости различия ( $p$ )?

- ошибка первого рода;
- вероятность ошибки отвергнуть справедливую нулевую гипотезу;
- ошибка второго рода;
- вероятность найти различия там, где их на самом деле нет.

## 96. Что такое ошибка первого рода?

- вероятность найти различия там, где их на самом деле нет;
- вероятность не найти различий там, где они есть;
- некорректное применение статистического критерия;
- отсутствие научной гипотезы.

## 97. Что такое ошибка второго рода?

- вероятность найти различия там, где их на самом деле нет;
- вероятность не найти различий там, где они есть;
- некорректное применение статистического критерия;
- отсутствие научной гипотезы.

## 98. Что такое двойное слепое исследование?

- О том, какой препарат принимается, не знают ни пациенты, ни врач;
- исследование с двумя случайными группами сравнения;
- двух этапное исследование инвалидов по зрению;
- исследование двух групп с нечетко поставленной целью.

## 99. Проспективное исследование

- исследование нерандомизированное;
- исследователь пользуется имеющимися данными;
- исследование слепое;
- исследователь сам проводит наблюдение за пациентами, сбор и анализ данных.

# 100. Продольные исследования

- это одномоментные срезы для оценки распространенности определенного заболевания, факторов риска и т.д.;
- выполняются несколькими исследователями;
- продолжаются определенный период времени, когда принимается лечение и проводится наблюдение за пациентами;
- включают несколько целей.

## **101.** Применение статистических методов в медицине требует:

- заучивания формул;
- умения отыскать табличное значение;
- понимания сути задачи, возможностей и ограничений статметодов, корректной интерпретации результатов;
- использования конкретной компьютерной программы.



*СПАСИБО ЗА  
ВНИМАНИЕ!*

*m. (384-3)-796-770  
jilina@ivcgzo.nkz.ru*