

Основы медицинской статистики

Кафедра общественного здоровья и
здравоохранения

доцент, к.м.н.

Чебыкин Дмитрий Викторович

Статистика

– общественная наука, изучающая количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной

Основная задача статистики

-выявление тенденций и закономерностей в социальных и экономических процессах

Особенности статистики

1. Массовость исследуемых явлений;
2. Изучение общественных явлений в условиях конкретного места и времени;
3. Изучение динамики общественных явлений.

Медицинская статистика

- отраслевая наука, использующаяся для изучения тенденций и закономерностей в общественном здоровье, здравоохранении, общей гигиене и научных медицинских исследованиях.

Разделы медицинской статистики

1. Санитарная статистика

- Статистика здоровья (показатели заболеваемости, рождаемости, смертности, инвалидизации, физического развития);
- Статистика здравоохранения (показатели работы ЛПУ, финансирования учреждений здравоохранения и т.д.);

2. Статистические методы в лабораторных, биологических и других научных исследованиях.

Основные направления применения статистических методов в медицине и здравоохранении:

- Сбор данных и обобщение результатов исследования с помощью относительных и средних величин (показатели здоровья и здравоохранения);
- Сравнение и определение достоверности различия двух и более групп результатов:
 - несопряженные выборки (сравнение групп результатов случай-контроль);
 - сопряженные выборки (сравнение групп результатов у одних и тех же людей до воздействия и после);
- Изучение взаимосвязи между факторами (корреляционный и другие виды анализа);
- Анализ динамики процессов.

Основные категории (терминология) статистики

Статистическая совокупность – это совокупность, состоящая из относительно однородных элементов (единиц наблюдения), взятых вместе в известных границах времени и пространства.

Особенности статистической совокупности:

- Множество единиц совокупности;
- Наличие единого свойства, объединяющего элементы в совокупность;
- Разнообразие признаков каждого отдельного элемента совокупности.

Основные категории (терминология) статистики

Виды статистических совокупностей:

- **Генеральная совокупность** - состоит из всех единиц наблюдения, которые могут быть к ней отнесены в соответствии с целью исследования.
- **Выборочная совокупность** – часть генеральной совокупности, отобранная специальным методом и предназначенная для характеристики генеральной совокупности.

Объект наблюдения – это статистическая совокупность о которой собираются сведения

Единица совокупности - составная часть, первичный элемент статистической совокупности, наделенный общими свойствами и имеющий индивидуальные отличительные признаки.

Признак – характерная особенность, отличительное свойство единицы наблюдения:

- количественные (возраст, рост, вес);
- качественные (пол, профессия);
- альтернативные (принимающие противоположное значение).

Этапы статистического исследования

1. Подготовительный этап;
2. Этап сбора материала
(статистическое наблюдение);
3. Сводка и группировка материала;
4. Анализ материала с помощью
обобщающих показателей;
5. *Выводы, предложения, внедрение в
практику.*

I этап - Подготовительный

- программно-методологические вопросы: (цель, задачи, выбор объекта и единицы наблюдения, определение способов сбора материала, подбор инструментария);
- организационные вопросы: (время и место проведения исследования, ресурсы, исполнители, руководитель).

Основные понятия подготовительного этапа статистического наблюдения:

- Цель исследования – конкретный результат, который мы хотим достигнуть;
- Научная гипотеза – предполагаемый результат (предшествует формулировке цели);
- Задачи исследования – это те действия, которые позволяют нам достигнуть цели, т.е. необходимого результата исследования;
- Регистрационная форма – учетный документ для сбора статистического материала (официальная форма, самостоятельно составленная учетная форма).

Основные понятия подготовительного этапа статистического наблюдения:

- План исследования – систематизирует решение организационных вопросов (место и сроки статистического наблюдения, кадровые и материальные ресурсы и т.д.)
В некоторых источниках, план исследования – это подробное описание последовательности выполнения действий.
- Программа исследования - подразумевает выбор объекта и единицы наблюдения, способов сбора материала, определение необходимого числа наблюдений, выбор учетных признаков.

NB! 2/3 успеха исследования зависит от правильной его подготовки – формулирования целей, задач, составления программы, определения объекта и единицы исследования, способов сбора материала.

Требования, предъявляемые к статистическому наблюдению:

- наблюдаемые явления должны иметь научную или практическую ценность;
- полнота данных;
- достоверность собранных данных;
 - компетентность работника, участвующего в статистическом наблюдении,
 - совершенство инструментария (бланков, инструкций),
 - проверка (контроль) качества собираемых фактов;
- обоснованность отбора той части совокупности, о которой собираются данные;
- сопоставимость данных;
- своевременность.

ФОРМЫ СБОРА СТАТИСТИЧЕСКОГО МАТЕРИАЛА:

Специально организованное статистическое наблюдение –это сбор информации, организуемый с какой-либо целью, на определенную дату с целью получения сведений, о которых не предоставляется отчетность.

Статистическая отчетность –это форма наблюдения, при которой стат.данные поступают в органы статистики от предприятий и организаций в виде обязательных отчетов

Требования к отчетности:

- обязательность;
- строгие сроки подачи;
- ответственность подающего за достоверность;
- фиксированная программа отчета.

Классификация отчетности:

- по принадлежности (общегосударственная, ведомственная);
- по длительности периода отчета (помесячная, поквартальная, годовая);
- по содержанию (типовая и специализированная).

Способы (виды) сбора статистического материала:

1. По временному критерию:

- Единовременное;
- Текущее;
- Интервальное.

Способы (виды) сбора статистического материала:

2. По способу регистрации данных:

- Метод непосредственного наблюдения;
- Метод выкопировки данных;
- Метод опроса (устный, письменный).

Способы (виды) сбора статистического материала:

3. По полноте охвата единиц совокупности:

- Сплошное

- Несплошное

- монографическое;

- основного массива;

- выборочное.

Выборочное наблюдение

- это вид несплошного наблюдения, при котором производится отбор единиц наблюдения из генеральной совокупности.

Положительные стороны:

- ✓ Может быть единственно доступным методом во время исследования (при порче или уничтожении единиц наблюдения);
- ✓ Экономичность;
- ✓ Сжатые сроки, быстрое получение результата.

Выборочное наблюдение

- это вид несплошного наблюдения, при котором производится отбор единиц наблюдения из генеральной совокупности.

Отрицательные стороны:

- Неизбежность ошибки в исследовании, связанная с тем, что берутся не все единицы наблюдения;
- Для редких событий можно не накопить достаточного количества единиц наблюдения;
- При социологических исследованиях могут вызывать чувство дискриминации у населения.

Способы формирования выборки:

- **Случайный;**
- **Механический;**
- **Типический (типологический);**
- **Серийный;**
- **Многоступенчатый;**
- **Когортный;**
- **Направленного отбора;**
- **Метод копи-пара.**

Ошибки статистического наблюдения:

1. Ошибки регистрации

- Случайные ошибки;
- Систематические;
- Преднамеренные;
- Непреднамеренные.

Ошибки статистического исследования:

Систематическая ошибка, обусловленная неправильно составленным регистрационным документом в программе наблюдения:

Распределение студентов педиатрического факультета по возрастным интервалам:

16-18

18-20

20-22

22-24

24-26

Ошибки статистического исследования:

Правильное распределение студентов педиатрического факультета по возрастным интервалам:

16-18	- 123 студента
19-21	- 215 студентов
22-24	- 180 студентов
24-26	- 230 студентов
26-28	- 150 студентов

Интервал группировки $i = 3$

(в каждый интервал входит по три возраста)

2. Методические ошибки

□ Ошибки репрезентативности

- ошибка качественной репрезентативности – нарушение случайности отбора;

- ошибка количественной репрезентативности – недостаточность числа наблюдений;

□ Использование средних величин в неоднородных группах и в группах с выбросами

Ошибки статистического исследования:

Расчет средних в неоднородной группе:

Возраст больных серозным менингитом
(данные инфекционной больницы №1)

2,2,3,3,4,4,4,4,5,5,5,6,6,7,8,8,9,10,10,52,53,53
,54,55,58,59,63,63,64

Средняя величина = 23,4

Расчет средних в группе с выбросами:

2,2,3,3,4,4,4,4,5,5,5,6,6,62

Средняя 8,1

III ЭТАП СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ

Статистическая сводка –это обработка материалов статистического наблюдения для полной и объективной характеристики всей совокупности.

Группировка –это процесс разбиения совокупности на однородные группы по определенным существенным для них признакам.

Виды статистических таблиц:

1. Простые;
2. Сложные:
 - групповые;
 - комбинационные.

Простые таблицы:

№ детского сада	Количество детей
№3	112
№7	143
№12	132
№21	118

Образование персонала	Количество сотрудников
Среднее	27
Средне-специальное	45
Высшее	56
Ученая степень, звание	22

Групповые таблицы:

№ детского сада	Количество детей	Количество сотрудников
№3	112	14
№7	143	21
№12	132	16

№ детского сада	пол		возраст		
	мальчики	девочки	0-2	2-4	4-6
№3	62	50	32	45	35
№7	83	60	43	51	49
№12	42	90	42	40	50

КОМБИНАЦИОННАЯ ТАБЛИЦА

№ д/сад	МАЛЬЧИКИ			ДЕВОЧКИ		
	0-2	2-4	4-6	0-2	2-4	4-6
№3	-	-	-	-	-	-
№7	-	-	-	-	-	-
№12	-	-	-	-	-	-

Вариационный ряд – это ряд, который строится из числовых значений:

5, 6, 8, 9, 11, 12, 7, 11, 16, 12, 8, 9, 8, 10

Элементы вариационного ряда:

- **Варианта (V);**
- **Частота (p);**
- **Общее количество значений (n).**

КЛАССИФИКАЦИЯ ВАРИАЦИОННЫХ РЯДОВ:

- **Простой:**
5, 6, 7, 8, 8, 8, 9, 9, 10, 11, 11, 12, 12, 16
Сгруппированный;
5-7 - 3
8-10 - 6
11-13- 4
14-16- 1.
- **Ранжированный:**
5, 6, 7, 8, 8, 8, 9, 9, 10, 11, 11, 12, 12, 16;
Неранжированный:
5, 6, 8, 9, 11, 12, 7, 11, 16, 12, 8, 9, 8, 10;
- **Непрерывный (состоящий из дробных чисел):**
36,6; 36,9; 37,5; 38,1; 39,0 –t тела 5-ти пациентов;
Прерывный (состоящий из целых чисел):
122, 134, 132, 154, 99, 121 –количество детей в 6-ти д/с

Показатели вариации

Амплитуда –определяется как разность между крайними вариантами

$$A_m = V_{\max} - V_{\min};$$

Лимит –это соотношение крайних вариант

$$L_{\text{im}} = V_{\max} / V_{\min};$$

Среднеквадратическое отклонение (σ) –характеризует структуру вариационного ряда, а также типичность средней величины;

Средняя ошибка средней арифметической (m)

Коэффициент вариации –позволяет делать вывод об однородности совокупности

$$C_v = \sigma / M * 100\%$$

0 - 10% -слабое разнообразие признака в совокупности

11 – 25% -среднее разнообразие признаков в совокупности

>25% -высокое (сильное) разнообразие признаков в совокупности

Виды статистических величин

Абсолютные величины

Производные:

- относительные;
- средние.

Абсолютные величины –отражают размеры явлений и процессов и получаютя в результате стат.наблюдения:

- Индивидуальные;
- Суммарные.

Относительные величины –отражают соотношение двух или нескольких абсолютных величин

1. Экстенсивный показатель

– это показатель структуры распределения, отражает в процентах удельный вес части явления в целом.

Экстенсивный
показатель = $\frac{\text{часть явления}}{\text{целое явление}} \times 100\%$

% мужского
населения в
популяции = $\frac{\text{число мужчин}}{\text{общая численность населения}} \times 100\%$

Таким же образом рассчитывается структура населения по возрасту, структура по причинам смерти, структура заболеваемости, лейкоцитарная формула и др.

Относительные величины

2. Интенсивный показатель

–это показатель частоты (распространенности) явления в среде продуцирующей данное явление.

Интенсивный показатель = $\frac{\text{явление, продуцируемое средой}}{\text{среда, продуцирующая явление}} \times 1000$

Показатель заболеваемости в Новосиб-ке = $\frac{\text{Число зарегистрированных заболеваний}}{\text{Общее численность населения}} \times 1000$

Таким же образом рассчитывается смертность, рождаемость, инвалидизация населения.

Относительные величины

3. Показатель соотношения

–характеризует численное соотношение двух не связанных между собой совокупностей.

$$\text{Показатель соотношения} = \frac{\text{явление, непродуцируемое средой}}{\text{среда, непродуцирующая явление}} \times 1000$$

$$\text{Показатель обеспеченности населения стац. койками} = \frac{\text{Число стационарных коек}}{\text{Общее численность населения}} \times 1000$$

4. Показатель наглядности

– используется с целью сравнения рядов различных величин. Указывают на сколько процентов произошло увеличение или уменьшение сравниваемых величин.

Средние величины – отражают общую меру исследуемого признака в совокупности

Мода (M_o) –соответствует величине признака, который чаще всего встречается в совокупности

5, 6, 7, 8, 8, 8, 9, 9, 10, 11, 11, 12, 12, 16

Медиана (M_e)–величина признака, занимающая срединное положение в вариационном ряду

5, 6, 7, 8, 8, 8, 9, 9, 10, 11, 11, 12, 12

($n=13$)

Средняя арифметическая (M)–вычисляется при учете всех единиц наблюдения и является характеристикой для всего вариационного ряда

Виды средних арифметических:

1. Простая средняя арифметическая

Вычисляется в вариационном ряду, в котором каждая варианта встречается одинаковое число раз;

2. Взвешенная средняя арифметическая

Вычисляется из вариационного ряда, в котором отдельные варианты встречаются различное число раз;

3. Ср. арифметическая, вычисляемая по способу моментов -рассчитывается в сгруппированном интервальном ряду.

Расчет простой средней

Факультеты	% курящих
Лечебный	24%
Педиатр	32%
Стоматологический	45%
Экономики в здравоохранении	80%
Средняя	45,25

Расчет взвешенной средней

Факультеты	% курящих	кол-во студентов	абс. число курящих
Лечебный	24	2000	480
Педиатрический	32	1000	320
Стоматологический	45	500	225
Экономики в здравооо-нии	80	100	80
		3600	1105
Процент курящих		30,7	

Расчет простой средней

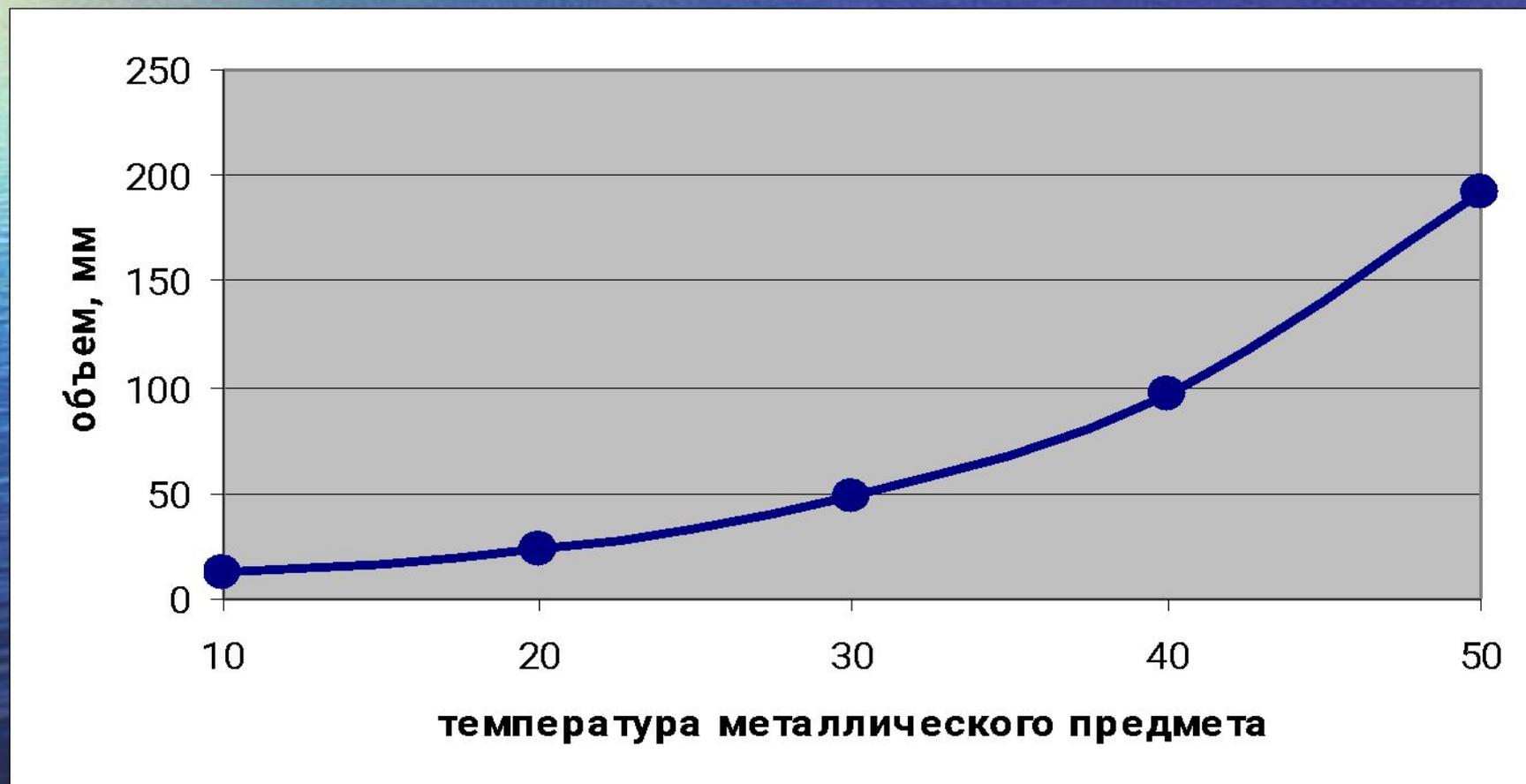
ЛПУ района	Ср. длительность госпитализации
Частная клиника	7,8
Участковая больница №1	11,5
Участковая больница №2	15
ЦРБ	16,8
Участковая больница №3	20,5
Средняя величина	14,32

Расчет взвешенной средней

ЛПУ района	Ср. длительность госпитализации	Количество пациентов	Количество койко-дней
Частная клиника	7,8	10	78
Участковая больница №1	11,5	68	782
Участковая больница №2	15	79	1185
ЦРБ	16,5	226	3729
Участковая больница №3	20,5	76	1558
Итого		459	7332
Средняя величина	16,0		

Виды взаимосвязей между явлениями

1. Функциональная связь -каждому значению одного признака соответствует строго определенное значение другого признака (обычно проявляется в физике, химии);



Виды взаимосвязей между явлениями

2. Корреляционная связь –каждому значению одного признака соответствует несколько значений другого признака (характерна для социально-гигиенических и медико-биологических процессов).

Пример:

- одному росту может соответствовать разный вес пациентов;
- при одинаковом уровне вакцинации в соседних районах (процент привитых)–уровень заболеваемости может быть разным.

Корреляционная связь рассчитывается по специальным формулам Пирсона (метод квадратов) и Спирмена (метод рангов).

Классификация корреляционной связи

- I. По направлению:
- прямая (+) –при увеличении одного признака увеличивается другой или при уменьшении одного признака другой также уменьшается, т.е. присутствует однонаправленность изменения величины признака.
(пример: с увеличением температуры тела –ЧСС также возрастает).
 - обратная (-) –при увеличении одного признака другой уменьшается или при уменьшении одного признака другой –увеличивается.
(пример: с увеличением возраста –снижается *visus* –острота зрения; с уменьшением вакцинации –заболеваемость увеличивается).

Классификация корреляционной связи

II. По силе (знак не учитывается):

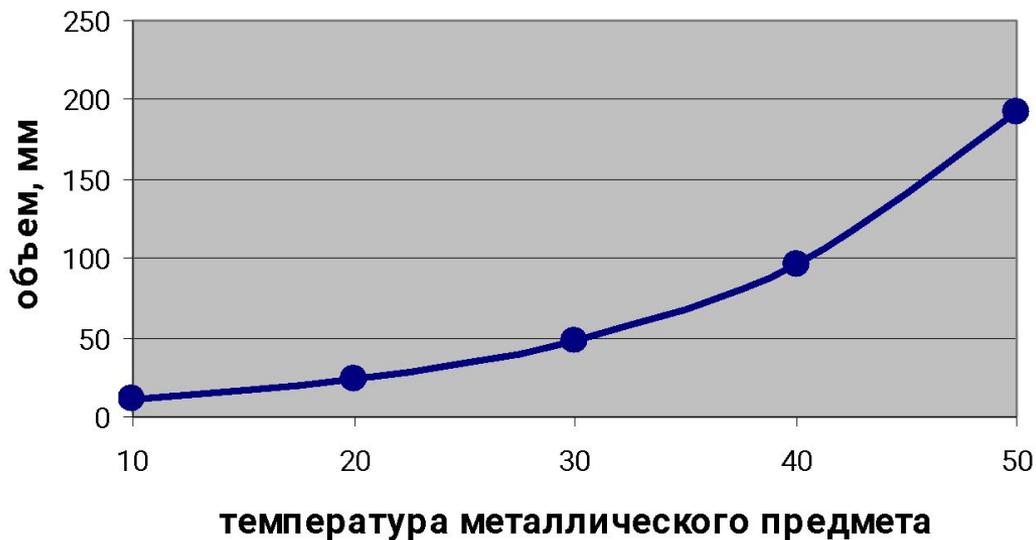
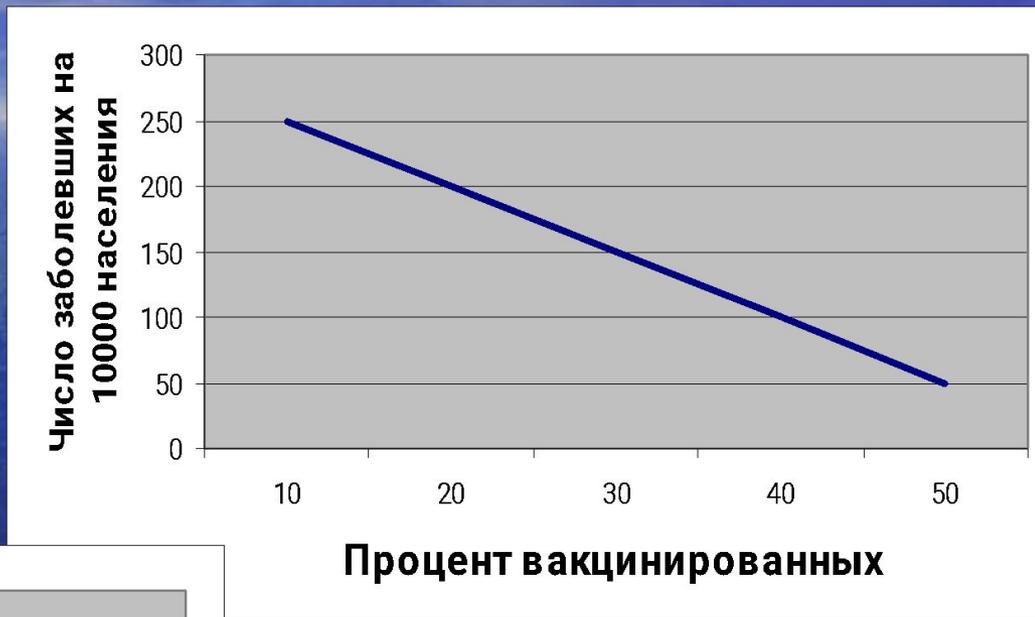
- Сильная от 0,7 до 1
- Средняя от 0,3 до 0,7
- Слабая от 0 до 0,3
- 0 –отсутствие связи
- 1 –связь полная

Классификация корреляционной связи

III.

По линейности

- Прямолinéйная
- Криволинейная



Ряд динамики –это ряд последовательно расположенных в хронологическом порядке показателей, которые отражают изменение явления во времени.

1. По времени:

- моментный; - интервальный;

Временной критерий (год)	Уровень ряда (число жителей)
01.01.2005	143 тыс.
01.01.2006	142,5 тыс.
01.01.2007	141,2 тыс.
01.01.2008	142,5 тыс.

Временной критерий (годы)	Уровень ряда (тонны продукции)
1980-1985	48 тыс.
1985-1990	54 тыс.
1990-1995	81,2 тыс.
1995-2000	84,5 тыс.

2. По виду величин:

- абсолютные; - относительные; - средние

Временной критерий (год)	Число жителей в тыс.
01.01.05	143
01.01.06	142,5
01.01.07	141,2
01.01.08	142,5

Временной критерий (годы)	Уровень рожд-ти на 1000 населения
`86	16,4
`87	16,1
`88	15,7
`89	15,0

Временной критерий (годы)	Средняя заработная плата
2005	8,9 тыс.
2006	10,4 тыс.
2007	11,2 тыс.
2008	12,6 тыс.

3. По полноте охвата во времени:

- полный; - неполный;

Временной критерий (год)	Уровень ряда (число жителей)
2005	143 тыс.
2006	142,5 тыс.
2007	141,2 тыс.
2008	142,5 тыс.

Временной критерий (годы)	Уровень ряда (тонны продукции)
2004	48 тыс.
2006	54 тыс.
2007	81,2 тыс.
2009	84,5 тыс.

Показатели анализа рядов динамики

1. Абсолютный прирост –это
разница последующего и
предыдущего уровней;

Для 95-96гг. $10 - 5 = 5$

Для 95-99гг. $4 - 5 = -1$

Годы	Тонны продукции
1995	5
1996	10
1997	15
1998	7
1999	4

Показатели анализа рядов динамики

2. Темп роста –это процентное соотношение последующего и предыдущего уровней;

Для 95-96гг. $(10 / 5) * 100\% = 200\%$

Для 95-99гг. $(4 / 5) * 100\% = 80\%$

Годы	Тонны продукции
1995	5
1996	10
1997	15
1998	7
1999	4

Показатели анализа рядов динамики

3. Темп прироста –это процентное соотношение абсолютного прироста и предыдущего уровня;

Для 95-96гг. $(5 / 5) * 100\% = 100\%$

Для 95-99гг. $(-1 / 5) * 100\% = -20\%$

Годы	Тонны продукции
1995	5
1996	10
1997	15
1998	7
1999	4

Показатели анализа рядов динамики

4. Значение 1% прироста – это соотношение абсолютного прироста и темпа прироста.

Для 95-96гг. $5 / 100\% = 0,05$ тонны

Для 95-99гг. $-1 / -20\% = 0,05$

Годы	Тонны продукции
1995	5
1996	10
1997	15
1998	7
1999	4

Метод стандартизации

Отделения	Больница А		Больница В	
	Всего больных	Число умерших	Всего больных	Число умерших
Хирургическое	1500	180	500	80
Терапевтическое	500	30	500	40
Инфекционное	500	20	1500	90
Всего	2500	230	2500	210

Показатели летальности в больнице А и В соответственно:

$$\text{Показатель госпитальной летальности} = \frac{\text{Общее число умерших больных}}{\text{Общее число пролеченных больных}} \times 100$$

Больница А - 9,2 на 100 человек

Больница В - 8,4 на 100 человек

Метод стандартизации (общая таблица)

Отделе- ния	Больница А			Больница В			Стан- дарт	Ожидаемое число умерших больных	
	Всего больн ых	Число умер ших	Лета льно сть	Всего боль ных	Числ о умер ших	Лета льно сть		А	В
Хирургия	1500	180		500	80				
Терапия	500	30		500	40				
Инфекция	500	20		1500	90				
Всего	2500	230	9,2	2500	210	8,4			

Возможные причины разницы в показателях летальности:

- Более низкая квалификация персонала;
- Недооснащенность современными технологическими средствами;
- Высокий процент тяжелобольных!!!;
- Больше количество хирургических больных в стационаре!!!

I этап стандартизации – расчет относительных величин
(расчет показателей летальности по отделениям)

Отделение	Больница А			Больница В		
	Всего больных	Число умерших	Летальность на 100 больных	Всего больных	Число умерших	Летальность на 100 больных
Хирургич	1500	180	12	500	80	16
Терапевт ич	500	30	6	500	40	8
Инфекцио нное	500	20	4	1500	90	6
Всего	2500	230		2500	210	

II этап стандартизации –определение стандарта (за стандарт берется полусумма больных по отделению)

Отделе- ния	Больница А			Больница В			Стан- дарт	Ожидаемое число умерших больных	
	Всего больн ых	Число умер ших	Лета льно сть	Всего боль ных	Числ о умер ших	Лета льно сть		А	В
Хирургия	1500	180	12	500	80	16	1000		
Терапия	500	30	6	500	40	8	500		
Инфекция	500	20	4	1500	90	6	1000		
Всего	2500	230		2500	210		2500		

III этап стандартизации – расчет ожидаемых величин

Отделе- ния	Больница А			Больница В			Станд арт	Ожидаемое число умерших больных	
	Всего больн ых	Число умер ших	Лета льно сть на 100	Всего боль ных	Числ о умер ших	Лета льно сть на 100		А	В
Хирургия	1500	180	12	500	80	16	1000	120	160
Терапия	500	30	6	500	40	8	500	30	40
Инфекция	500	20	4	1500	90	6	1000	40	60
Всего	2500	230		2500	210		2500	190	260

IV Этап

Расчет стандартизированных показателей

- Для больницы А:
2500 больных -100%
190 предполагаемо умерших – x %;
- Для больницы В:
2500 больных -100%
260 предполагаемо умерших – x %;

Стандартизированные показатели летальности:

Для больницы А: 7,6

Для больницы В: 10,4

Таким образом, если бы состав больных по отделениям в количественном соотношении в больнице В был бы такой же как в больнице А, то показатель летальности в больнице В был бы существенно выше.

Проведя стандартизацию мы исключили влияние фактора разности объемов оказываемой помощи по различным отделениям.

Оценка достоверности результатов исследования:

- Параметрические критерии (t-критерий Стьюдента) –применяются в больших выборках с правильным распределением признака.
- Непараметрические критерии (критерий знаков, критерий Вилкоксона, критерий Х-квадрат) –применяются для оценки достоверности исследования в малых выборках с ассиметричным распределением признака.

Критерий Стьюдента (t)

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

M_1 и M_2 – средние величины;

m_1 и m_2 – ошибки средних величин

(показывают вариабельность признака в совокупности).

Критерий Стьюдента (t)

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

Вес 10-ти детей до отправление в пионерский лагерь:

33, 34, 38, 33, 34, 37, 35, 36, 39, 37

$M_1 = \mathbf{35,6}$ $m_1 = 0,7$

После пионерского лагеря средний вес этих же детей составил $M_2 = \mathbf{37,3}$

Равномерное увеличение массы тела:

37, 35, 39, 35, 36, 37, 38, 38, 39, 39 $m_2 = 0,5$

Неравномерное увеличение

33, 35, 39, 33, 34, 37, 35, 38, 46, 47 $m_2 = 1,6$

Критерий Стьюдента (t)

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}} = \frac{35,6 - 37,3}{\sqrt{0,7^2 + 0,5^2}} = 2$$

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}} = \frac{35,6 - 37,3}{\sqrt{0,7^2 + 1,6^2}} = 0,9$$

КРИТЕРИЙ СТЬЮДЕНТА

- Если вычисленное значение t окажется меньше 2, то различие между средними признается случайным, статистически не значимым;
- При $t > 2$ это различие можно считать значимым с вероятностью 95%;
- При $t > 2.6$ – значимым с вероятностью $> 99\%$;
- При $t > 3.3$ – с вероятностью более 99,9%.



СПАСИБО ЗА ВНИМАНИЕ!