

КЫРГЫЗСКО-РОССИЙСКИЙ СЛАВЯНСКИЙ УНИВЕРСИТЕТ

**КЫРГЫЗСКАЯ ГОСУДАРСТВЕННАЯ
МЕДИЦИНСКАЯ АКАДЕМИЯ**

**КАФЕДРА СОЦИАЛЬНОЙ МЕДИЦИНЫ
И ОРГАНИЗАЦИИ ЗДРАВООХРАНЕНИЯ
им. проф. А.А.АЙДАРАЛИЕВА**

**К.У.Акынбеков, Н.Е.Чернова, К.Д.Абдуллин,
А.А.Орозалиева**

БИОСТАТИСТИКА

(УЧЕБНОЕ ПОСОБИЕ)

**Издательство Кыргызско-Российского
Славянского университета
Бишкек · 2001**

*Акынбеков К.У., Чернова Н.Е.,
Абдуллин К.Д., Орозалиева А.А.*

БИОСТАТИСТИКА. Учебное пособие /Под ред. К.У.Акынбекова. Кыргызско-Российский Славянский университет. – Бишкек, 2001. – 109 с.

Даны основы биostatистики, организация статистического исследования и ее этапы, представлены методы вычисления основных статистических показателей. По каждой теме предложены контрольные вопросы и задачи для самостоятельной работы. Приведены рисунки, диаграммы, таблицы и примеры.

Для студентов, ординаторов, аспирантов, научных работников, организаторов здравоохранения и практических врачей.

Рецензенты:

доктор медицинских наук, проф. Б.С.Мамбеталиев;
доктор медицинских наук М.М.Каратаев

© КРСУ, 2001 г.

СОДЕРЖАНИЕ

Предисловие	4
1. ВВЕДЕНИЕ В БИОСТАТИСТИКУ	5
2. ОРГАНИЗАЦИЯ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ.....	11
3. ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ.....	20
4. ВАРИАЦИОННЫЙ РЯД, СРЕДНИЕ ВЕЛИЧИНЫ, СРЕДНЕЕ КВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ И КОЭФФИЦИЕНТ ВАРИАЦИИ	33
5. ОЦЕНКА ДОСТОВЕРНОСТИ СРЕДНИХ, ОТНОСИТЕЛЬНЫХ ВЕЛИЧИН И РАЗНОСТИ РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЙ	45
6. ОЦЕНКА ДОСТОВЕРНОСТИ РАЗЛИЧИЯ СРАВНИВАЕМЫХ ГРУПП ПО КРИТЕРИЮ СООТВЕТСТВИЯ χ^2 (ХИ-КВАДРАТ) ...	54
7. ВЫБОРОЧНЫЙ МЕТОД В СТАТИСТИЧЕСКОМ ИССЛЕДОВАНИИ	59
8. СТАНДАРТИЗАЦИЯ ПОКАЗАТЕЛЕЙ	66
9. ИЗМЕРЕНИЕ СВЯЗИ МЕЖДУ ЯВЛЕНИЯМИ ИЛИ ПРИЗНАКАМИ. КОРРЕЛЯЦИЯ	72
10. КОЭФФИЦИЕНТ РЕГРЕССИИ	82
11. ДИНАМИЧЕСКИЕ РЯДЫ И ИХ АНАЛИЗ	87
12. ГРАФИЧЕСКИЕ ИЗОБРАЖЕНИЯ	95
Приложения	104
Литература	109

ПРЕДИСЛОВИЕ

Современный этап развития здравоохранения внес значительные изменения в организацию лечебно-профилактических учреждений. Накапливаются новые данные: о состоянии здоровья населения и факторах, определяющих его; о сети деятельности и кадрах лечебных учреждений; о результатах лечебно-оздоровительных мероприятий, которые используют при поиске путей улучшения здоровья населения и дальнейшего совершенствования здравоохранения. В связи с этим возникла необходимость в широком применении разнообразных статистических, в том числе биостатистических методов обработки этих данных. Биостатистика – это статистическая информация, а также методы, в которых используются специальные рекомендации, для изучения вопросов здравоохранения и социальных проблем.

Большие возможности для дальнейшего перспективного применения статистики в медицине и здравоохранении создают вычислительная техника, а также новые математические методы. Разнообразные электронно-вычислительные машины облегчают расчеты и обработку потока массовой информации, а также решение логических задач и осуществление математического моделирования различных явлений и процессов в медицине и здравоохранении. Следовательно, каждый врач должен хорошо знать теоретические основы статистики, уметь правильно использовать биостатистические методы и оценивать информацию, накопленную в различных областях его деятельности.

Тема №1

ВВЕДЕНИЕ В БИОСТАТИСТИКУ

Цель темы:

Студент должен *знать*:

- Содержание биostatистики (медстатистики).
- Статистическую совокупность, ее виды.

Студент должен *уметь*:

- Определять объем, объект и виды статистической совокупности, единицу наблюдения и учетные признаки.

План изучения темы

Разбор темы по контрольным вопросам

1. Определения биostatистики, статистики и медицинской статистики.
2. Задачи медицинской статистики.
3. Разделы медицинской статистики.
4. Показатели состояния здоровья населения.
5. Показатели статистики здравоохранения.
6. Понятие о статистической совокупности.
7. Единицы наблюдения – понятие.
8. Учетные признаки – понятие.
9. Виды учетных признаков.
10. Виды статистической совокупности и их сущность.
11. Репрезентативность – понятие, требования.
12. Свойства статистической совокупности.
13. Типы распределения признаков в статистической совокупности.
14. Виды статистических величин.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Биostatистика – статистическая информация, а также методы, в которых используются специальные рекомендации для изучения вопросов здравоохранения и социальных проблем. Биostatистика охва-

тывает, перекрывает и в некоторой степени является аналогом витальной статистики (рождаемость, смертность) и в целом демографии.

Основным методом биostatистики является статистика.

Статистика – самостоятельная общественная наука. Она изучает количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной.

Статистика изучает массовые явления, их общие свойства и закономерности, которые нельзя выявить при единичных наблюдениях.

Медицинская статистика (санитарная статистика) – отрасль статистики, изучающая количественные закономерности состояния и динамики здоровья населения и системы здравоохранения, а также разрабатывающая методы статистического анализа клинических и лабораторных данных.

Задачи медицинской статистики:

1. Выявление особенностей в состоянии здоровья населения и факторов, определяющих его.
2. Изучение данных о сети, деятельности и кадрах в органах и учреждениях здравоохранения.
3. Статистическая обработка материала в экспериментальных, клинических, гигиенических и лабораторных исследованиях.

Разделы медицинской статистики:

1. Теоретические и методические основы.
2. Статистика здоровья населения.
3. Статистика здравоохранения.
Статистика здоровья населения изучает состояние здоровья, которое состоит из 4 групп показателей:
 1. Физического развития.
 2. Медико-демографических показателей.
 3. Заболеваемости и травматизма.
 4. Инвалидности.Статистика здравоохранения изучает деятельность органов и учреждений здравоохранения по следующим основным показателям:
 1. Объему выполненной работы.
 2. Количественной и качественной характеристики работников.
 3. Качественных показателей работы.
 4. Нормам и нормативам здравоохранения.
 5. Организационным формам работы.
 6. Финансовой деятельности.

Изучение тех или иных явлений с применением статистических методов требует от врача прежде всего умелого подхода к выбору объекта исследования (так называемой статистической совокупности), единицы наблюдения и ее учитываемых признаков.

Статистическая совокупность – это группа однородных единиц наблюдений, взятых во времени и пространстве. Например, группа больных в данное время в данной больнице.

Статистическая совокупность состоит из отдельных единиц наблюдений.

Единицы наблюдения – первичные элементы данной совокупности, сходные между собой. Например, больные данной больницы. Численность единиц наблюдения в совокупности определяет объем исследования и обозначается буквой «n».

Учетные признаки – различия между первичными элементами. Например, пол, возраст, профессия, национальность и др.

Учетные признаки по характеру бывают:

1. Количественные, выражаемые числом. Например, возраст, вес, среднесуточная температура тела, артериальное давление и т.д.
2. Качественные или атрибутивные, выражаемые словесно и имеющие определенный характер, содержание. Например, пол, профессия, национальность, диагноз.

По своей роли в составе совокупности учетные признаки делятся на:

1. Факторные, которые влияют на изменение другого признака. Например, рост ребенка: с возрастом увеличивается вес ребенка.
2. Результативные, которые изменяются в зависимости от изменения факторного признака. Например, вес ребенка: его увеличение зависит от увеличения роста ребенка.

Различают два вида статистической совокупности:

- Генеральную.
- Выборочную.

Генеральная совокупность состоит из всех единиц наблюдения, которые могут быть к ней отнесены в соответствии с целью исследования. Например, беременность и роды, осложненные тяжелой формой акушерского травматизма – разрывом матки. Если бы удалось собрать информацию о всех женщинах с указанной патологией на определенной территории за определенный период времени, то эта информация была бы о группе женщин генеральной совокупности. Генеральная совокупность формируется только при сплошном исследовании, при котором необходимо установить абсолютные размеры явления, такие как: общая числен-

ность населения данной территории, численность врачей, больничных коек и т.д. Организация сплошного исследования сложна и громоздка, требует больших материальных затрат и времени. Однако в ряде случаев невозможно, а иногда и не целесообразно подвергать изучению (анализу) все единицы наблюдения, составляющие генеральную совокупность, тогда исследование ограничивается определенной частью единиц, т.е. **выборочной совокупностью**.

Выборочная совокупность – часть генеральной совокупности, отобранная специальным методом и предназначенная для характеристики генеральной совокупности. Выборочная совокупность должна быть репрезентативной, т.е. в отобранной части должны быть представлены все элементы и в том соотношении, как и в генеральной совокупности. Для обеспечения репрезентативности выборочной совокупности к ней предъявляют два основных требования:

- Выборочная совокупность должна обладать характерными чертами генеральной совокупности, т.е. по составу быть максимально похожей на нее.
- Выборочная совокупность должна быть достаточной по объему (по числу наблюдений) (см. тему №7).

Свойства статистической совокупности:

1. Характер распределения явления: частота, доля.
2. Средний уровень явления: мода, медиана, средняя арифметическая.
3. Разнообразие единиц наблюдения: лимит, амплитуда, среднее квадратическое отклонение, коэффициент вариации.
4. Взаимосвязь между признаками: коэффициент корреляции, коэффициент регрессии.
5. Репрезентативность признаков: средняя ошибка, коэффициент соответствия.

Типы распределения признаков в статистической совокупности:

1. Альтернативный, т.е. имеющие противоположное значение (да, нет). Например, исход лечения: умер, выжил (не умер).
2. Нормальный или симметричный, когда наибольшее число случаев приходится на середину ряда. Например, наибольшее число больных в стационаре (в днях) приходится на середину недели.
3. Асимметричный, когда наибольшее число случаев сдвигается в сторону: правосторонняя, левосторонняя; или ряд имеет две вершины – двугорбый или бимодальный.

Для оценки изучаемых явлений, составляющих статистическую совокупность, используют статистические величины:

1. Абсолютные числа.
2. Относительные величины.
3. Средние величины.

Самостоятельная работа студентов

Пример

В институте обучается 8000 студентов, в том числе на 5 курсе 1500 человек, из них курящих 500 студентов. Были обследованы студенты 5 курса на предмет заболеваний органов пищеварения в зависимости от курения: из числа курящих у 20 студентов выявлена язвенная болезнь желудка; у 45 – хронические гастриты; из числа некурящих соответственно – 35 и 55.

Определить: объем совокупности, вид совокупности, единицу наблюдения, признаки наблюдения.

Образец выполнения:

Объем совокупности – 8000, т.е. все студенты института.

Вид совокупности – выборочная, т.к. обследована часть студентов института, а именно 5 курса в количестве 1500 человек.

Единица наблюдения – студент, как первичный элемент данной совокупности.

Признаки наблюдения – курящий, некурящий, язвенная болезнь, хронический гастрит. Они являются различиями между первичными элементами.

Задача 1

Врач Т. изучил распространенность заболеваний органов дыхания у шахтеров 2-ой железорудной шахты в области N и исследовал жалобы на кашель.

Результаты исследования шахтеров, работающих в забое, показали, что из 80 шахтеров в возрасте до 30 лет 12 человек предъявили жалобы на кашель с мокротой, а из 300 шахтеров в возрасте старше 30 лет жаловались на кашель с мокротой 58 человек.

В группе 150 шахтеров этой шахты в возрасте до 30 лет, не работающих в забое, предъявили жалобы на кашель 12 человек, а из 450 рабочих в возрасте старше 30 лет жаловались на кашель 45 человек.

Определить: объем совокупности, вид совокупности, единицу наблюдения, признаки наблюдения.

Задача 2

Врач-кардиолог изучал отдаленные последствия специального экспериментального метода лечения 200 больных ишемической болезнью (из них 100 больных до 50 лет и 100 – старше 50 лет).

Результаты исследования показали, что через год после первого инфаркта наступил рецидив у 20 больных в возрасте до 50 лет и у 30 больных в возрасте старше 50 лет.

Определить: объем совокупности, вид совокупности, единицу наблюдения, признаки наблюдения.

Задача 3

Группа терапевтов изучала последствия гриппа у рыбаков. Результаты исследования показали, что из 150 предварительно иммунизированных больных, перенесших грипп, вызванный вирусом A₂, у 50 были выявлены изменения ЭКГ.

Из 400 больных без предварительной иммунизации, перенесших грипп A₂, у 200 были выявлены изменения ЭКГ.

Определить: объем совокупности, вид совокупности, единицу наблюдения, признаки наблюдения.

Тема №2

ОРГАНИЗАЦИЯ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ

Цель темы:

Студент должен *знать*:

- Правила и порядок проведения статистического исследования.
- Особенности проведения социально-медицинских и клинико-статистических исследований.
- Основные виды ошибок при анализе материала.

Студент должен *уметь*:

- Составлять программу статистического исследования.
- Определять необходимый объем наблюдений, проводить разработку, сводку и анализ материала.

План изучения темы

Разбор темы по контрольным вопросам

1. Этапы статистического исследования.
2. Элементы первого этапа статистического исследования и их содержание.
3. Элементы второго этапа статистического исследования и их содержание.
4. Организационные элементы плана исследования.
5. Сущность характеристики объекта исследования.
6. Способы формирования статистической совокупности.
7. Разделы программы исследования.
8. Содержание программы сбора материала.
9. Содержание программы разработки материала.
10. Содержание программы анализа материала.
11. Группировка материала и требования к нему.
12. Статистические таблицы, их содержание и отличия.
13. Стадии четвертого этапа статистического исследования.
14. Условия анализа материала.
15. Ошибки при анализе материала.
16. Особенности статистического исследования в социально-медицинских и клинико-статистических работах.

Самостоятельная работа студентов

- Составление плана и программы статистических исследований на заданную тему.

Содержание темы

В практической деятельности врача часто возникает необходимость проведения статистического исследования с целью получения данных, необходимых для разработки предложений по улучшению медицинской помощи. Так, например, для того, чтобы наметить пути повышения эффективности использования коечного фонда, необходимо изучить сроки пребывания больных в стационаре в зависимости от пола, возраста, заболевания, сроков госпитализации и других факторов.

Указанные и множество других вопросов требуют знания этапов статистического исследования и их содержания, умения формулировать цель и задачи исследования, умения правильно определять единицу наблюдения, составлять макеты статистических таблиц, опре-

делять необходимое число наблюдений, проводить шифровку, группировку, сводку и анализ данных.

Статистические исследования проводятся в 5 этапов:

1. Формулировка цели и задач исследования.
2. Организационный.
3. Сбор материала.
4. Обработка материала.
5. Анализ, выводы и предложения.

1 этап. Формулировка цели и задач исследования. Цель – это конкретное конечное состояние или желаемый результат, которого стремятся достичь. Например, улучшение состояния здоровья населения. Улучшение медицинской помощи населению. Повышение качества профилактики заболевания и т.п. Цель отвечает на вопрос: – Зачем проводится данное исследование? Цель должна быть сформулирована четко и недвусмысленно.

Для раскрытия цели необходимо определить задачи, которые в краткой и конкретной форме освещают все стороны поставленной цели.

Задача – это предписанная работа. Например, для улучшения состояния здоровья необходимо изучить:

- 1) показатели здоровья населения;
- 2) отрицательные факторы, влияющие на состояние здоровья;
- 3) положительные факторы, влияющие на здоровье населения.

Число задач для исследований намечают несколько: от 3 до 6. Задача отвечает на вопрос: – Что делать?

2 этап. Организационный. Состоит из 2 частей:

1. Составление плана исследования (где, сколько, как, когда и кто).
2. Составление программы исследования (что исследовать и в каком направлении).

1 часть. В плане отражаются организационные элементы работы:

1. Характеристика объекта исследования, т.е. совокупность явлений, предметов, о которых должны быть собраны статистические сведения. Например, население города, больные лица, кадры и др. Объект исследования (совокупность) должен быть определен в пространстве (территория), во времени (период), в объеме (число) наблюдений (n), по признакам, характеризующим подлежащее (пол, возраст, национальность и т.п.).

2. Способы формирования статистической совокупности. Они бывают по полноте охвата наблюдений – сплошные и выборочные; по времени – текущие и единовременные; по виду – непосредственные (выкопировка сведений), анамнестические (анкетный и опросный).

При выборочной совокупности предусматривают способы отбора единицы наблюдения. Они бывают:

- **Случай-контроль.** Отличием этого исследования является то, что субъекты исследования отбираются по наличию или отсутствию у них изучаемого заболевания. Структура исследования подразумевает наличие двух групп наблюдения, при этом в опытную группу (случай) включают лиц, у которых выявлено данное заболевание (или иной изучаемый исход), а в контрольную (группа сравнения, референс-группа) – лиц, у которых данное заболевание отсутствует (контроль). Опытную и контрольную группы затем делят на подгруппы «экспонированных» и «неэкспонированных» к изучаемому фактору риска и проводят оценку распределения изучаемого фактора риска по подгруппам.
- **Механический** – проводится по какому-либо признаку. Например, первая буква фамилии, номер истории болезни.
- **Типологический** – позволяет производить выбор единиц наблюдения из типичных групп всей генеральной совокупности. Например, сначала все единицы наблюдения группируются по возрасту, затем из каждой возрастной группы производят случайный или механический отбор.
- **Серийный (гнездовой)** – предусматривает выбор из генеральной совокупности не отдельных единиц, а выбор серий (гнезд), отражающих свойства генеральной совокупности. А внутри серий (гнезд) наблюдение производится сплошным методом.
- **Парносопряженный** – каждой единице наблюдения подбирают пару в контрольной группе. В контрольной группе в каждой единице наблюдения все признаки наблюдения должны быть сходными с опытной группой, за исключением исследуемых признаков.
- **Направленный** – исследуются лица одного возраста и пола. Например, женщины 20-24 лет для изучения деторождаемости.
- **Когортный** – подразумевает формирование опытной группы (когорты), в которую включают лиц, экспонированных к изучаемому фактору риска, и контрольной группы неэкспонированных, при этом ни один из наблюдаемых к началу исследования не имеет рассматриваемого заболевания (иного исхода).

3. Срок работы.
4. Исполнители.
5. Финансовые вопросы.
6. Проведение инструкций, семинаров для исполнителей.

2 часть. В программе отражаются методические элементы работы.

1. Программа сбора материала: определение единицы наблюдения, перечня учетных признаков, источников получения материала, составление карты (анкеты, бланка) для заполнения сведений, составление инструкций по заполнению карт, шифровальный код, методы обработки материала (ручная, машинная).
2. Программа разработки материала: необходимо наметить группировку материала. Под группировкой понимается распределение совокупности единиц наблюдения в однородные группы по одному или нескольким признакам. Например, по возрастным, образовательным группам. Группировка является творческим процессом, для этого необходимо знать существо материала. Группировка должна быть научно обоснованной. Неправильная группировка материала может привести к ошибочным выводам. Один и тот же материал при разной группировке может привести к различным выводам. Затем приступают к составлению макетов статистических таблиц для последующего их заполнения. Таблицы должны иметь четкие и краткие заглавия. В них различают подлежащее (о чем говорится) и сказуемое (объясняет подлежащее). Подлежащее помещают слева (строка), сказуемое – сверху (графы).

Различают три вида статистических таблиц:

1. Простая статистическая таблица имеет одно статистическое подлежащее, позволяет провести сводку материала по одному признаку (табл. 1).

Таблица 1
Распределение больных острой пневмонией, лечившихся в стационаре N-ой больницы, по исходам заболеваний (в абс. числах)

Исход заболевания	Число больных
Выздоровление	
Улучшение	
Без изменений	
Ухудшение	
Летальный исход	
Всего:	

2. Групповая статистическая таблица имеет одно статистическое подлежащее и несколько не связанных между собой статистических сказуемых (табл. 2).

Таблица 2

Распределение больных острой пневмонией, лечившихся в стационаре N-ой больницы, по исходам заболевания, возрасту и полу (в абс. числах)

Исход заболевания	Возраст в годах			Пол		Итого
	До 30	31-59	60 и ст.	Муж.	Жен.	
Выздоровление						
Улучшение						
Без изменений						
Ухудшение						
Летальный исход						
Всего						

3. Комбинационная статистическая таблица имеет одно или несколько взаимосвязанных статистических подлежащих и несколько взаимосвязанных структур (табл. 3).

Таблица 3

Распределение больных острой пневмонией, лечившихся в стационаре N-ой больницы, по исходам заболевания, возрасту и полу (в абс. числах)

Исход заболевания	Возраст в годах						Итого
	До 30		31-59		60 и старше		
	М.	Ж.	М.	Ж.	М.	Ж.	
Выздоровление							
Улучшение							
Без изменений							
Ухудшение							
Летальный исход							
Всего							

Таблицы не должны быть громоздкими, не более 3-4 признаков.

3 этап. Сбор материала. Сбор материала проводят по плану и программе для получения необходимых сведений.

4 этап. Обработка материала.

1. Контроль качества собранного материала.
2. Шифровка материала (если материал большой).
3. Группировка материала.
4. Сводка материала (заполнение таблиц).
5. Вычисление статистических показателей.
6. Графическое изображение материала.

5 этап. Анализ материала.

Условиями для его проведения являются:

- Всестороннее знание исследователем сущности изучаемого явления.
- Владение методикой статистического исследования, в том числе методикой статистического анализа.
- Правильное выполнение предыдущих этапов статистического исследования.

Основные виды ошибок при анализе:

- Методические: недостаточное число наблюдений, неправильное определение единицы наблюдения, неправильная группировка.
- Неправильная оценка статистических величин: суждение о динамике явления на основе экстенсивных показателей, оценка темпа роста без учета исходного уровня; использование средних в неоднородных группах, недостаточная или неправильная статистическая обработка материала, чрезвычайное увлечение «математизацией».
- Логические: сравнение данных без учета их качественной характеристики, смешение причины и следствия, недоучет связей явления, слабое знание сущности явления.

Особенности статистического исследования

В социально-медицинских исследованиях	В клинко-статистических исследованиях
1. Изучаются, в основном, социально-медицинские вопросы	Изучаются, в основном, клинические вопросы
2. В основном, изучается влияние социально-медицинских факторов на здоровье населения	В основном, изучается влияние клинических вопросов (диагностика, лечение и др) на состояние здоровья больного
3. Как правило, не имеется контрольной группы наблюдения	Необходима контрольная группа наблюдения
4. Требуется большое количество единиц наблюдения	Достаточно малое количество единиц наблюдения (n<30)
5. Материал собирается путем выкопировок, анамнеза (анкетирование, опрос)	Материал собирается путем непосредственного наблюдения за больными и выкопировок
6. Единицей наблюдения является здоровый и больной человек	Единицей наблюдения является больной

Самостоятельная работа студентов

Образец

Задание на составление первых двух этапов статистического исследования на тему «Изучение распространения гастрита среди студентов КГМА»

1 этап

Цель исследования

Задачи исследования

2 этап

План исследования:

- Объект исследования или статистическая совокупность;
- Способы формирования статистической совокупности;
- Способ отбора материала.

Программа исследования:

- Программа сбора материала:
 1. Единица наблюдения
 2. Учетные признаки
 3. Составление карты
 1. ФИО
 2. Возраст
 3. Пол
 4. Национальность
 5. Курс обучения
 6. Вид обучения: бюджетный, контрактный (нужное подчеркнуть)
 7. С кем проживает: в общежитии, в квартире с родителями, в квартире без родителей (нужное подчеркнуть)
 8. Диагноз

Рекомендации по снижению заболеваемости гастритом среди студентов.

- Изучить распространение гастрита.
- Выявить основные причины заболевания гастритом.

Студенты академии

По охвату – выборочная
По виду – выкопировка сведений

Серийный (обследовать одну группу с каждого курса сплошным методом)

Случай болезни гастритом

Возраст, пол, национальность, курс обучения, вид обучения, с кем проживает, диагноз

Карта обследования

- Программа разработки материала:

Группировка материала

1. Возраст: до 20 лет, 20 лет и старше
2. Пол: мужской/женский
3. Национальность: европейской нац., коренной нац., другие, иностранцы
4. Курс обучения: 1-3, 4-6
5. Диагноз: острый гастрит, хронический гастрит

2. Составление макета таблицы

Задание 1

Цель исследования – изучить распространение курения среди студентов КГМА, способствующие ему причины, и наметить мероприятия по его уменьшению.

В соответствии с целью:

1. Сформировать основные задачи исследования.
2. Определить единицу и признаки наблюдения.
3. Составить:
 - Карту обследования.
 - Группировку материала.
 - Макеты таблиц.

Задание 2

Цель исследования – изучить распространение язвенной болезни среди студентов КГМА, влияющие на него факторы и наметить пути по снижению заболеваемости.

В соответствии с целью:

1. Сформировать основные задачи исследования.
2. Определить единицу и признаки наблюдения.
3. Составить:
 - Карту обследования.
 - Группировку материала.
 - Макеты таблиц.

Задание 3

Цель исследования – изучить уровень информации студентов института искусств о венерических болезнях и источники получения этой информации. Наметить мероприятия по улучшению знаний студентов об этих болезнях.

В соответствии с целью:

1. Сформировать основные задачи исследования.
2. Определить единицу и признаки наблюдения.
3. Составить:
 - Карту обследования.
 - Группировку материала.
 - Макеты таблиц.

Тема №3

ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

Цель темы:

Студент должен *знать*:

- Виды относительных величин и их сущность.
- Различие между относительными величинами.

Студент должен *уметь*:

- Вычислять относительные величины.

План изучения темы

Разбор темы по контрольным вопросам

1. Определение относительных величин.
2. Виды относительных величин.
3. Интенсивные показатели, их применение, техника вычисления.
4. Экстенсивные показатели, их применение, техника вычисления.
5. Отличие интенсивных и экстенсивных показателей.
6. Коэффициенты соотношения, их применение, техника вычисления, отличие от интенсивных показателей.
7. Относительные числа наглядности, их применение, техника вычисления.
8. Общие и специальные показатели относительных величин (интенсивности, соотношения).

Самостоятельная работа студентов

- Решение задач.

Содержание темы

В медицине и здравоохранении для характеристики изучаемой совокупности используют относительные величины. Они необходимы для сравнения и сопоставления одной совокупности с другой.

Однако это не означает, что абсолютные величины вообще не применяются при анализе. Абсолютные величины сами по себе несут важную информацию о размере того или иного явления: количестве больных, родившихся, числе коек в стационаре и т.д.

Но чаще всего абсолютные величины без преобразования их в относительные показатели имеют ограниченное познавательное значение. Например, в городе А было зарегистрировано в 2000 г. 1500 случаев вирусного гепатита, а в городе Б – 900 случаев этого заболевания. По этим абсолютным числам нельзя сделать вывод о том, что населения города А болеет вирусным гепатитом чаще, чем население города Б. Чтобы сделать правильный вывод, необходимо учесть различия в численности населения этих двух городов, что достигается путем преобразования абсолютных величин в относительные величины.

Таким образом, *относительные величины* – это результат сравнения (отношения) двух величин. Различают следующие виды относительных величин:

1. Относительные величины интенсивности (частоты, силы, распространения).
2. Относительные величины экстенсивности (структуры, удельного веса).
3. Относительные величины соотношения.
4. Относительные величины наглядности.

1. Интенсивные показатели характеризуют частоту (силу) распространения явления в среде, в которой оно происходит непосредственно и органически связанное с ней, как бы порождается этой средой. Например: роды, в т.ч. осложненные кровотечением, где роды – среда, кровотечение в родах – явление, вызванное (порожденное) родами. Из этого следует, что для вычисления интенсивных показателей необходимо иметь среду и явление, происходящее в ней, т.е. две самостоятельных, качественно различных статистических совокупности.

Среда, в статистическом смысле этого слова, представляет собой статистическую совокупность, в которой происходят изучаемые явления (процессы).

Средой могут быть население или его различные контингенты. Например: контингент больных, к которому определяются показатели исхода заболевания (выздоровления, инвалидности, летальности и т.д.).

Интенсивные показатели вычисляются к единому условному знаменателю (основанию) – к единице с нулями (100, 1000, 10000, 100000), в зависимости от частоты изучаемого явления. Если за основание принимается 100 – показатель выражается в процентах (%), 1000 – в промиллях (‰), 10000 – процедимиллях (‱) и т.д. Чем реже явление происходит в своей среде, тем больше следует брать основание, чтобы избежать получения показателя в малых дробных величинах. Эта рекомендация вытекает из неудобства использования коэффициентов с большим количеством десятичных знаков. При изменении основания величина коэффициента изменяется в соответствующее число раз (10, 100 и более). Следовательно, каждый интенсивный показатель, вычисленный к любому основанию, может быть пересчитан к другому основанию путем переноса запятой вправо или влево на определенное число десятичных знаков. При вычислении интенсивных показателей рождаемости, общей заболеваемости, смертности за основание обычно принимают 1000 человек населения; при вычислении размеров заболеваемости, смертности от отдельных редко встречающихся заболеваний за основание принимают 10000, 100000 и т.д.

Исходной средой в медицинских исследованиях может быть население или отдельные его контингенты (беременные женщины, роженицы, новорожденные дети и др.), по отношению к которой вычисляются интенсивные показатели рождаемости, заболеваемости, материнской и перинатальной смертности и др.

Например: требуется установить, как часто в общей численности беременных встречаются женщины с резус-отрицательной принадлежностью крови (общая численность беременных – основная, исходная среда, беременные с резус-отрицательной кровью – явление). Так, в женской консультации г. N состоялось на диспансерном учете 1210 беременных женщин, из них 58 с резус-отрицательной принадлежностью крови. Определить интенсивность распространения беременных с Rh-кровью.

Техника вычисления:

1. Составление пропорции.

На 1210 беременных – 58 с Rh-кровью

На 1000 беременных – X

2. Вычисление интенсивного показателя.

$$X = \frac{58 \times 1000}{1210} = 47,9\%$$

Если за основание взять 100 – интенсивный коэффициент будет равен 4,79%, т.е. запятая переносится влево на один десятичный знак.

Если интенсивный показатель вычисляется за период времени меньше года, для сравнения его надо привести к годовому уровню, т.е. определить чему бы он равнялся, если бы частота изучаемого явления в течение всего года была такой, как в период наблюдения. Для этого числитель умножается на 12 (число месяцев в году), а знаменатель на число месяцев периода наблюдения.

Например: в г. N с населением 180000, за 5 месяцев родилось 1750 детей. Коэффициент рождаемости за 5 месяцев составил:

$$P = \frac{1750 \times 1000 \times 12}{180000 \times 5} = 23,3\%$$
 – коэффициент рождаемости за 5 месяцев.

При пользовании интенсивными коэффициентами всегда следует указывать к какому основанию они вычислены (на 100, 1000, 10000 и т.д.)

2. Экстенсивные показатели (экстенсивные отношения). Вторыми по частоте применения, после интенсивных коэффициентов, в статистической обработке материалов исследования являются экстенсивные показатели или относительные числа распределения или, иначе, экстенсивные отношения.

Экстенсивные показатели характеризуют распределение, или расчленение целой совокупности (среды) на ее составные части, или соотношения частей к целому, принятому за 100, и выражаются в процентах (%), или за 1000 и выражаются в промиллях (‰). Экстенсивные показатели (отношения) освещают внутренний состав, или структуру одного явления, относительное чередование его составных элементов. Например: в 2000 г. в г. N зарегистрировано 4878 больных, из них 2120 терапевтического профиля, 650 – хирургического, 1090 – гинекологического и 1018 – других профилей (табл. 1).

Техника вычисления:

1. Составление пропорции.

Целое явление 4878 – 100%

Часть явления 2120 – X

2. Вычисление экстенсивного показателя.

$$X = \frac{2120 \times 100}{4878} = 43,5\%$$

Таблица 1

Структура зарегистрированных больных в городе N в 2000 г.

Наименование профилей	Абсолютные числа	Процент к итогу (экстенсивные показатели)
Терапевтический	2120	43,5%
Хирургический	650	13,3%
Гинекологический	1090	22,3%
Другие профили	1018	20,9%
Итого:	4878	100,0

Если целое число (4878) принять за 1000, получим промилльные отношения (%_о). И в том и в другом случае полученные коэффициенты показывают распределение всего числа больных на его составные части по профилю заболевания, но не указывают на частоту распространения заболеваний.

При вычислении экстенсивных показателей берется только одна статистическая совокупность, определяется ее внутренняя структура путем распределения на ее составные части.

Характерной особенностью экстенсивных показателей является их взаимосвязанность, вызывающая определенный автоматизм сдвигов его составных частей, т.к. сумма их всегда равна 100 (1000). Экстенсивные показатели дают представление об удельном весе, о величине доли того или иного явления по отношению к другой доле только в данной среде (совокупности).

Например: из 50 заболевших новорожденных (общая исходная среда) 30 заболели пневмонией, что составило от общего числа 60% [(30·100):50], 10 – ОРЗ или 20%; 10 – прочими заболеваниями или 20%. Общая сумма всех составных частей должна составлять 100% (60,0%+20,0%+20,0%=100%). Тем не менее, экстенсивные показатели можно сравнивать. Например: до 1979 г. в структуре материнской смертности Кыргызской Республики первое место занимали акушерские кровотечения, второе – экстрагенитальные заболевания, третье – послеродовые септические заболевания и т.д. В последующие годы в структуре материнской смертности первое место заняли гестозы беременных, второе – акушерские кровотечения, третье – послеродовые септические заболевания. Эти данные свидетельствуют о том, что в силу каких-то причин в структуре материнской смертности по сравнению с

1979 г. произошли сдвиги ее составных частей: увеличилась доля умерших от гестозов, уменьшилась – от акушерских кровотечений, прежнее место заняли послеродовые септические заболевания. Эти сведения имеют определенное значение при анализе изучаемого явления. Однако они не указывают на частоту распространения изучаемого явления, на его динамику во времени. Поэтому экстенсивными показателями нельзя пользоваться для сравнения изучаемого явления (материнской смертности) в динамике или между отдельными регионами, нельзя определить их достоверность путем вычисления того или иного критерия достоверности (t, F и др.). Эту функцию могут выполнить только интенсивные показатели (частоты).

Экстенсивными показателями можно пользоваться только для определения внутренней структуры явления (смертности, заболеваемости), для выявления роли ее отдельных частей, относительного порядка их чередования.

В ряде медицинских исследований нередко можно встретить необоснованные выводы (заключения) в связи с недостаточной осведомленностью в основах статистики. Например: утверждение об увеличении заболеваемости раком женских половых органов в молодом возрасте исследователь делает только на основе относительного увеличения экстенсивного коэффициента смертности от указанной патологии в молодом возрасте по сравнению с другими повозрастными коэффициентами. Однако увеличение экстенсивного показателя смертности от рака гениталий в молодом возрасте могло произойти и без действительного изменения смертности в этом возрасте, а лишь отражать собой уменьшение смертности от других причин. Поэтому утверждения об «омоложении» рака гениталий, сделанные только на сравнении экстенсивных показателей не верны, не убедительны особенно в научных исследованиях. Подобные сравнения вполне правомерны, если на их основе не делаются обобщающие, необоснованные заключения о частоте явления.

При анализе относительных величин наиболее частые ошибки допускаются при использовании интенсивных и экстенсивных показателей.

В тех случаях, когда статистическая природа коэффициентов не вполне ясна, необходимо помнить следующее:

- При интенсивных коэффициентах мы всегда имеем дело с двумя строго ограниченными совокупностями, коллективами, с двумя самостоятельными, качественно различными совокупностями, одна

из которых характеризует среду, вторая – явление, распространенное в ней.

- При экстенсивных показателях исследователь имеет дело с одной статистической совокупностью – одним коллективом (только с больными, только с умершими). Поэтому, как бы детально не дифференцировался ее внутренний состав, понятия о частоте получить нельзя, отсутствует среда – основной фон. Это следует твердо помнить, т.к. наиболее частой ошибкой неопытных в знании статистики исследователей является то, что они делают выводы о частоте на основании данных распределения.

3. Коэффициенты соотношения. Коэффициенты соотношения характеризуют численное соотношение двух не связанных между собой разнородных совокупностей – «замкнутых величин». Например: обеспеченность населения родильными койками, где население – одна самостоятельная «замкнутая» совокупность, родильные койки – вторая самостоятельная «замкнутая» совокупность. Обе указанные совокупности не связаны друг с другом, одна совокупность не порождает другую. Аналогичными примерами являются обеспеченность населения машинами «скорая медицинская помощь», выполнение плана строительства медицинских учреждений, показатели применения физиотерапевтического лечения, лечебной физкультуры, переливания крови и кровезаменителей и другие лечебно-профилактические методы.

Техника вычисления интенсивных показателей и показателей соотношения одинакова. Основанием для них являются 100, 1000, 10000 и т.д. и выражаются в процентах (%), промиллях (‰), продецимиллях (‱). Однако при вычислении интенсивного показателя в числителе ставится величина, отражающая явление (число родившихся, заболевших, умерших) в знаменателе – величина, отражающая среду (население), в которой совершаются явления (процессы), т.е. обе эти величины органически связаны между собой, среда порождает явление. Поэтому, число, стоящее в числителе входит в состав числа, стоящего в знаменателе.

При вычислении показателя соотношения в числителе и в знаменателе для сопоставления ставятся две совершенно разнородные по своему значению величины. Поэтому величина в числителе не входит в состав величины знаменателя. Коэффициенты соотношения показывают частоту, но не вскрывают внутренних связей.

Интенсивные показатели преимущественно отражают состояние здоровья населения, а показатели соотношения – различные виды и стороны деятельности медицинских учреждений.

Подтверждением этого служат следующие примеры:

1. Уровень госпитализации больных. В г. N 120000 населения. В амбулаторно-поликлинических учреждениях зарегистрировано впервые установленных в данном году 180000 заболеваний, из них госпитализировано 24400 больных. При отношении последних двух величин к населению, как органически связанными с ним, получим интенсивные показатели.

Техника вычисления:

1. Составление пропорции.
2. Вычисление интенсивных показателей.

- *Общая заболеваемость:*

$$\frac{120000 - 180000}{1000 - x} \quad x = \frac{180000 \times 1000}{120000} = 1500\%_0.$$

- *Госпитализированная заболеваемость:*

$$\frac{120000 - 24400}{1000 - x} \quad x = \frac{24400 \times 1000}{120000} = 203\%_0.$$

При отношении двух самостоятельных величин амбулаторной и больничной помощи получим показатель соотношения, характеризующий уровень госпитализации.

1. Составление пропорции.
2. Вычисление показателей соотношения.

- *Коэффициент соотношения:*

$$\frac{180000 - 24400}{1000 - x} \quad x = \frac{24400 \times 100}{180000} = 13,6\%$$

Вторым, не менее убедительным примером, являются рождаемость и частота аборт на 1000 населения – интенсивные показатели, а отношение числа абортов к числу родов – коэффициент соотношения и т.д. В этом и заключается смысл логических различий коэффициентов интенсивности и коэффициентов соотношения.

4. Относительные числа (показатели) наглядности характеризуют отношение каждой из сравниваемых величин к исходному уровню, принятому за условную величину, обычно за 100. При статистической обработке материалов исследования или анализе

деятельности учреждений здравоохранения наряду с перечисленными относительными коэффициентами нередко применяются относительные числа (показатели) наглядности. Коэффициенты наглядности не имеют качественно нового содержания, представляют собой способ преобразования самостоятельных абсолютных, относительных и средних величин в легко обозримую, наглядную форму в динамике. С этой целью составляется динамический ряд сравниваемых величин (абсолютных, относительных или средних). Одна из величин динамического ряда, наиболее яркая, приравнивается к какому-либо круглому числу (100,1000,10000 и т.д.) и принимается за исходную – базисную, все остальные числа ряда пересчитываются по отношению к базисной величине. Содержание динамического ряда при этом не меняется.

Например: рост численности врачей акушеров-гинекологов в Кыргызской Республике за период 1940-1980 гг. может быть представлен в следующем виде (табл. 2):

Таблица 2

Рост численности врачей акушеров-гинекологов в Кыргызской Республике с 1940 по 2000 гг. в динамике

Годы	Численность врачей акушеров-гинекологов	Показатель наглядности (в %) по отношению к 1940 г.
1940	36	100
1950	126	350
1960	255	708
1970	385	1069
1980	606	1683
1990	1015	2819
2000	1206	3350

Примечание: при построении динамического ряда для большей наглядности необходимо соблюдать равные временные интервалы между рядами (5 лет, 10 лет и т.д.).

В приведенном примере за исходную (базисную) величину принята численность врачей акушеров-гинекологов 1940 г. – 36. Эта величина приравнена к 100 и по отношению к ней определен рост численности врачей в последующие годы в процентах (%).

Техника вычисления:

1. Составление пропорции. 2. Вычисление показателей наглядности.

$$1940 \text{ г.: } 36 - 100\%$$

$$1950 \text{ г.: } 126 - x\%$$

$$x = \frac{126 \times 100}{36} = 350,0\%$$

т.е., за 10 лет численность врачей акушеров-гинекологов увеличилась на 350%, т.е. в 3,5 раза.

$$1940 \text{ г.: } 36 - 100\%$$

$$1960 \text{ г.: } 255 - x\%$$

$$x = \frac{255 \times 100}{36} = 708,3\% \text{ и т.д.}$$

За базисную величину можно принять любую, наиболее яркую величину динамического ряда: начальную, среднюю, конечную ряда.

Для примера приведем следующие данные: если за базисную величину будет принята первая, как в нашем примере, то ряд относительных величин наглядности будет нарастающим (от 100 до 3350%). Если за базисную величину будет взята одна из величин, стоящих в середине ряда (второй вариант), то вычисленные относительные показатели наглядности, расположенные выше базисной, будут меньше ее, а ниже – больше. В третьем варианте, если за базисную будет взята конечная величина динамического ряда, то все относительные величины, расположенные выше базисной, будут меньше ее.

За базисную величину можно взять не только любую величину динамического ряда, но и отсутствующую в нем. Например: за исходную величину может быть принята численность врачей акушеров-гинекологов 1913 г. (если имеются данные) или любую другую предыдущего года.

Эти данные свидетельствуют, насколько широко применяются показатели наглядности в научных исследованиях и практическом здравоохранении.

Интенсивные коэффициенты и некоторые коэффициенты соотношения могут быть общими и специальными.

Общие показатели характеризуют явление в целом по отношению ко всей среде. Например: общая заболеваемость, смертность, рождаемость вычисляется из числа всех заболевших, умерших, родившихся ко всему населению, общая летальность – отношение числа умерших к числу заболевших и т.д. В зарубежной статистике (английской) общие показатели считаются грубыми («crude»), т.к. они показывают только самую грубую ориентировку, сравнение динамики во времени и пространстве.

Для углубленного, более точного дифференцированного статистического анализа необходимо пользоваться специальными показателями, при которых можно выявить более тесные связи между изучаемым явлением и средой, в которой оно возникло. Специальные коэффициенты характеризуются разделением среды на ее составные части, применением более узкого основания, или расчленением явления.

Например: в г. N всего 5600 детей до 14 лет, из них умерло в течение года 72 ребенка, показатель общей детской смертности равен 12,8‰ [$P=(72 \times 1000):5600$]. Из общей численности детей до одного года жизни – 585, умерло 22 ребенка, смертность детей до одного года равна [$P=(22 \times 1000):585=37,6$]. 37,6‰ – это специальный показатель смертности детей до одного года жизни. Аналогично можно определить повозрастную смертность всего населения, смертность любой социальной группы населения, смертность от отдельных нозологических форм. При этом в числителе ставится число умерших (заболевших), т.е. изучаемого явления определенной возрастной группы (паритета беременности, родов и др.), в знаменателе – среда той же возрастной группы (социальной группы, по паритету родов и др.).

Например: в районе А 5900 детей до 14 лет, из них 180 детей до одного года. Из числа детей до одного года жизни в течение года первично заболело 37; заболеваемость детей до одного года жизни: $P=37 \times 1000:180=205,5$ ‰ – специальный показатель повозрастной заболеваемости.

Следует отметить: 1) при вычислении специальных показателей расчленяется или числитель, или знаменатель, или то и другое; 2) один и тот же показатель в одном исследовании может быть общим, в другом – специальным. Например: при изучении общей заболеваемости – заболеваемость детей до 14 лет будет специальным показателем, а при изучении детской – этот показатель будет общим; 3) общий показатель для данного исследования должен быть один (заболеваемость, смертность и др.), в то время как специальных показателей по изучаемому явлению может быть много (повозрастная заболеваемость населения, повозрастная детская заболеваемость и т.д.).

Самостоятельная работа студентов

Задача 1

Число умерших детей от инфекционных заболеваний в возрасте до 7 лет

Нозологические формы	Коклюш	Корь	Грипп	Другие	Всего
Число умерших	87	78	56	448	669

Численность детей – 203000. Вычислить структуру и частоту смертности детей.

Задача 2

Глазные заболевания среди населения некоторых районов Кыргызстана

Локализация заболеваний	Число заболеваний
Век и соединительных оболочек	6553
Роговой оболочки	1806
Сосудистого тракта	18
Хрусталика	47
Мышечного аппарата	21
Слезного аппарата	24
Прочих	39
Итого:	8508

Обследовано 14048 человек. Вычислить структуру и частоту заболеваемости.

Задача 3

Число обращений населения Ат-Башинского района в амбулаторно-поликлинические учреждения

Классы болезней	Число умерших
Болезни органов дыхания	20266
Болезни органов пищеварения	24800
Болезни нервной системы и органов чувств	8554
Болезни органов кровообращения	2709
Новообразования	506
Всего	56835

Численность населения 27200 человек. Вычислить структуру и частоту обращений.

Задача 4

Число впервые вышедших на инвалидность рабочих и служащих в г. N

Классы болезней	Число инвалидов
Системы кровообращения	732
Органов дыхания	221
Органов пищеварения	218
Новообразования	475
Костно-мышечной системы и соединительной ткани	340
Всего	1986

Численность рабочих и служащих – 779600. Вычислить структуру и частоту выхода на инвалидность

Задача 5

Число первичных посещений рабочих завода стоматологической поликлиники и санированных

Год	Число рабочих	Число посещений	Число санированных
1971	5500	2557	1211
1972	5000	2108	1122
1973	7000	4117	1786
Всего	17500	8782	4119

Вычислить структуру и частоту обращаемости и санированных.

Задача 6

Число заболеваний населения г. N отдельными инфекционными заболеваниями

Нозологическая форма	Бруцеллез	Корь	Инфекционный гепатит	Всего
Число заболеваний	270	4400	10200	14870

Население города – 250000 человек. Вычислить структуру и частоту заболеваемости.

Задача 7

Численность населения и врачей по зонам Кыргызстана в 1996 году

Зоны республики	Численность населения (тыс.)	Число врачей
г. Бишкек	595,3	2714
Чуйская область	762,5	1831
Иссык-Кульская область	429,9	1176
Нарынская область	264,7	675
Таласская область	207,9	514
Ошская область	1479,7	2950
Джалал-Абадская область	856,9	1751
По республике	4596,9	11611

Вычислить показатели соотношения.

Задача 8

Показатели рождаемости в Кыргызстане за 1992-1996 гг.

Год	1992	1993	1994	1995	1996
Рождаемость на 1000 населения	28,6	26,1	24,6	26,0	23,6

Вычислить показатели наглядности.

Задача 9

Численность населения и средних медицинских работников в Кыргызстане в 1996 г.

Зоны республики	Численность населения (в тыс.)	Число средних медработников
г. Бишкек	595,3	3479
Чуйская область	762,5	5259
Иссык-Кульская область	429,9	3550
Нарынская область	264,7	2663
Таласская область	207,9	1923
Ошская область	1479,7	12020
Джалал-Абадская область	856,9	7845
По республике	4596,9	36739

Вычислить показатели соотношения.

Задача 10

Показатели естественного прироста в Кыргызстане за 1992-1996 гг.

Год	1992	1993	1994	1995	1996
Показатель естественного прироста	21,4	18,4	16,3	17,8	16,0

Вычислить показатели наглядности.

Тема № 4

ВАРИАЦИОННЫЙ РЯД, СРЕДНИЕ ВЕЛИЧИНЫ, СРЕДНЕЕ КВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ И КОЭФФИЦИЕНТ ВАРИАЦИИ

Цель темы:

Студент должен *знать*:

- Вариационный ряд, его характеристику и виды.
- Основные положения теории средних величин, их практическое применение в работе врача.
- Значение среднего квадратического отклонения и коэффициента вариации.

Студент должен *уметь*:

- Составлять простой и сгруппированный вариационный ряды.
- Вычислять средние величины, среднее квадратическое отклонение и коэффициент вариации.

План изучения темы:

Разбор темы по контрольным вопросам

1. Понятие вариационного ряда.
2. Характеристика вариационного ряда.
3. Разница между простым и сгруппированным вариационным рядом.
4. Определение средней величины.
5. Виды средних величин.
6. Понятие моды (M_o) и медианы (M_e).
7. Способ вычисления простой средней арифметической.
8. Различие при вычислении центральной варианты интервального и непрерывного ряда.
9. Способ вычисления средней арифметической взвешенной.
10. Характеристика среднего квадратического отклонения σ (сигма).
11. Способы вычисления σ (сигмы).
12. Значение коэффициента вариации (C_v), критерий оценки, способ вычисления.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Вариационный ряд является одним из типов распределения признака в статистической совокупности и служит для вычисления средних величин.

Вариационный ряд – это ряд числовых измерений определенного признака, отличающихся друг от друга по своей величине, расположенных в определенном порядке.

Характеристиками вариационного ряда являются:

- Вариант (V) – числовое значение изучаемого признака.
- Частота (P) – число, указывающее сколько раз встречается данная варианта в вариационном ряду.
- Общее число наблюдений ($n=\Sigma P$), где Σ – знак суммы.

Вариационный ряд может быть простым, где каждая варианта обозначается отдельно, или сгруппированным, где варианты объединяются в группы с указанием частоты встречаемости всех вариантов, входящих в данную группу.

Простой вариационный ряд составляется обычно при малом числе наблюдений ($n<30$), а сгруппированный – при большом числе наблюдений ($n>30$).

После составления вариационного ряда определяется средний уровень изучаемого признака – вычисляется средняя величина, т.е. обобщающая характеристика статистической совокупности по определенному меняющемуся количественному признаку.

Средний уровень измеряют с помощью критериев, которые носят название средних величин. Под **средней величиной** понимают число, выражающее общую меру исследуемого признака в совокупности.

Общепотребительными являются три вида средних величин:

- Мода (M_o),
- Медиана (M_e).
- Средняя арифметическая (M).

Для определения любой средней величины необходимо использовать результаты индивидуальных измерений, записав их в виде вариационного ряда (табл. 1).

Таблица 1

Результаты измерения массы тела у юношей в возрасте 18 лет
(вариационный ряд)

Масса тела, кг (V)	59	60	61	62	63	64	Всего (n)
Число лиц (P)	1	4	6	9	3	2	25

Мода (M_o) – соответствует величине признака, которая чаще других встречается в данной совокупности. Иначе говоря, за моду принимают варианту, которой соответствует наибольшее количество частот (P) вариационного ряда. В нашем примере $M_o=62$ кг, т.к. эта масса тела наблюдается у 9 из 25 юношей.

Медиана (M_e) – величина признака, занимающая срединное положение в вариационном ряду. Она делит ряд на две равные части по числу наблюдений. Для определения медианы надо найти середину ряда.

Имеется несколько методов вычисления средней арифметической величины:

1. Средняя арифметическая простая вычисляется в том случае, если частоты всех вариантов равны единице, т.е. каждая варианта в своем вариационном ряду встречается один раз или же все варианты

встречаются с одинаковой частотой в совокупности, где $n < 30$. Вычисление средней арифметической простой производится по формуле

$$M = \frac{\sum V}{n}, \text{ где } \Sigma - \text{сумма, } V - \text{варианта, } n - \text{число наблюдений.}$$

Например: определить средний вес в кг 8 женщин, родивших в возрасте 18 лет:

$$M = \frac{\sum V}{n} = \frac{49,2 + 49,8 + 55,0 + 51,0 + 52,6 + 52,8 + 54,3 + 54,9}{8} = \frac{419,6}{8} = 52,5 \text{ кг}$$

Средний вес 8 женщин получен путем сложения всех значений вариант и делением суммы (419,6) на число наблюдений (8).

2. Средняя арифметическая взвешенная вычисляется тогда, когда все или несколько вариант встречаются в своем вариационном ряду несколько раз, где $n > 30$. Вычисление средней арифметической взвешенной проводится по формуле: $M = \frac{\sum VP}{n}$, где M – средняя арифметическая взвешенная, Σ – сумма произведений вариант на их частоту (VP), n – число наблюдений.

Например: определение средней арифметической в дискретном (прерывном) вариационном ряду (табл. 2).

Таблица 2

Распределение женщин по числу перенесенных в году гинекологических заболеваний

Число заболеваний V	Число женщин P	Произведения VP
1	2	3
0	12	0
1	44	44
2	86	172
3	119	357
4	112	448
5	76	380
6	49	294
Всего	n=500	$\Sigma VP=1698$

Методика вычисления (обычный метод):

- Определение суммы частот (ΣP) или числа наблюдений «n» путем последовательного сложения ($\Sigma P=500$).
- Определение произведений вариант на частоту (VP) – путем умножения каждой варианты на ее частоту (графа 3).

- Определение суммы произведений вариант на частоту ($\Sigma VP=1698$).
- Определение средней арифметической взвешенной дискретного ряда

$$M = \frac{\sum VP}{n} = \frac{1698}{500} = 3,4 - \text{заболевания в году на одну женщину (в среднем).}$$

Как отмечалось выше, средняя величина одним числом характеризует целую совокупность. Однако согласно нашему примеру получена средняя величина с десятичным знаком (3,4), но заболеваний, выраженных дробными долями быть не может. Следовательно, средняя ряда может не совпасть по величине ни с одним из индивидуальных значений признака и принимать любое значение в пределах между наименьшим и наибольшим значением ряда.

3. Средняя арифметическая сгруппированного (интервального) ряда. Этапы вычисления средней арифметической сгруппированного (интервального) ряда представлены в табл. 3.

Таблица 3

Определение средней заработной платы рабочих швейной фабрики

Заработная плата (в рублях)	Число рабочих	Середина интервала (центральная варианта $V_{ц}$)	VP
До 140	25	130	3250
140-160	50	150	7500
160-180	66	170	11220
180-200	99	190	18810
200-220	108	210	22680
220-240	32	230	7360
240 и больше	20	250	5000
Всего	n=400		$\Sigma VP=75820$

- Определение середины ряда (центральной варианты). Расчет в данном примере и аналогичных ему осложняется тем, что на основе интервального ряда нельзя определить заработок каждого отдельного рабочего. Известно только, что 25 рабочих первой группы ряда получили заработную плату до 140 руб. каждый, из 50 рабочих следующей группы – от 140 до 160 руб. и т.д. Поэтому для определения общей суммы заработка рабочих каждой группы приходится делать определенные допущения. Обычно считают, что внутри каждого варианта заработка рабочих распределены равномерно и середина интервала является средним значением варианты (центральная варианта) для данной группы рабочих. Тогда общий заработок каждой группы ра-

бочих определяется как произведение середины интервала на частоту, т.е. по обычному методу.

Средины интервалов легко найти для всех групп, кроме первой и последней, путем сложения минимальной и максимальной вариант каждой группы и делением их суммы пополам, т.е. они равны полусумме максимальной и минимальной вариант.

Например, середина интервала (центральная варианта) второй группы равна: $(140+160):2 = 150$ и т.д.

Однако у первой группы нет нижней границы (максимальной варианты интервала), а у последней – нет верхней границы интервала. В случаях, когда интервалы не определены, произвольно устанавливают их границы, чаще всего формально, считая, что первый интервал равен следующему за ним, т.е. $120 - 140$, середина интервала равна $(120+140):2=130$, для последней группы – $(240+260):2 = 250$. Так определены все середины интервалов $V_{ц}$ (графа 3, табл. 3).

- Определение произведений центральных вариант на частоту (графа 4).
- Определение суммы произведений вариант на частоту ($\Sigma VP=75820$; табл.3).
- Определение средней арифметической интервального ряда заработка рабочих: $x=189,6$ руб.

$$M = \frac{\Sigma VP}{n} = \frac{75820}{400} = 189,6 \text{ руб.}$$

Среднее квадратическое отклонение σ

Среднее квадратическое отклонение σ (сигма) характеризует степень рассеяния вариант вокруг средней арифметической и является вторым параметром вариационного ряда, т.е. величиной характеризующей вариационный ряд. Основное свойство среднего квадратического отклонения: чем меньше его числовое значение, тем компактнее расположены варианты вокруг своей средней ряда, и наоборот.

Для примера метода вычисления среднего квадратического отклонения приведем следующие данные (табл. 4).

Таблица 4

Вычисление средней арифметической роста женщин в 25-35 лет и среднего квадратического отклонения σ в населенном пункте N

Рост женщин в 25-35 лет в см (V)	Центральная варианта ($V_{ц}$)	Число женщин (P)	VP	Отклонения $V_{ц}$ от средней ($d=V_{ц}-x$)	Квадраты отклонений d^2	Pd^2
1	2	3	4	5	6	7
155,0-156,9	156,0	13	2028	-6	36	468
157,0-158,9	158,0	15	2370	-4	16	240
159,0-160,9	160,0	39	6240	-2	4	156
161,0-162,9	162,0	50	8100	0	0	0
163,0-164,9	164,0	40	6560	+2	4	160
165,0-166,9	166,0	16	2656	+4	16	256
167,0-168,9	168,0	3	504	+6	36	108
Всего		$\Sigma P=176$	$\Sigma VP=28458$			$\Sigma Pd^2=1388$

$$M = \frac{\Sigma VP}{n} = \frac{28458}{176} = 161,7 \approx 162; \quad \sigma = \sqrt{\frac{\Sigma Pd^2}{n}} = \sqrt{\frac{1388}{176}} = \sqrt{7,9} = 2,8.$$

Из этого следует, что среднее квадратическое отклонение – σ равняется квадратному корню из суммы произведений частот вариационного ряда на квадрат отклонений вариант от средней арифметической ряда (ΣPd^2), деленной на число наблюдений (n).

Этапы вычисления среднего квадратического отклонения сгруппированного ряда:

1. Составить вариационный ряд из представленной совокупности.
2. Определить центральную варианту ($V_{ц}$), равную полусумме начальных значений двух смежных интервалов.
3. Определить сумму частот вариант ряда (ΣP), или число наблюдений (n).
4. Определить произведения центральных вариант на частоту ($V_{ц}P$) и их сумму (ΣVP).
5. Вычислить среднюю арифметическую ряда: $M = \frac{\Sigma VP}{n}$.
6. Вычислить отклонения всех центральных вариант (d) от общей средней, равной 162 ($d=V_{ц}-x$).
7. Возвести в квадрат все отклонения (d^2).

8. Определить произведения частот на квадрат отклонений (Pd^2) и сумму произведений частот на квадрат отклонений ($\sum Pd^2$).

9. Определить среднее квадратическое отклонение $\sigma = \sqrt{\frac{\sum Pd^2}{n}}$; если n равен или меньше 30, то $\sigma = \sqrt{\frac{\sum Pd^2}{n-1}}$.

Вычисление среднего квадратического отклонения по амплитуде

В ряде случаев отсутствуют исходные данные для вычисления σ – среднего квадратического отклонения. В подобной ситуации может быть использован приближенный способ вычисления σ по амплитуде вариационного ряда, т.е. разности между минимальной и максимальной вариантами. Среднее квадратическое отклонение, вычисленное по амплитуде ряда, несколько отличается от такового, вычисленного обычным путем. При этом разница тем больше, чем больше число наблюдений, использованных для составления вариационного ряда. Поэтому, определение σ по амплитуде целесообразно использовать при небольшом числе наблюдений и при ориентировочных расчетах. Вычисление σ по амплитуде ряда производится по следующей формуле:

$$\sigma = \frac{\text{ampl}}{K} = \frac{V_{\max} - V_{\min}}{K},$$

где ampl – амплитуда ряда; K – коэффициент, соответствующий числу наблюдений (n). Коэффициент K определяется по специальной таблице (приложение 1), рассчитанной С.И.Ермолевой.

Методика использования таблицы K (см. приложение 1):

- в первой вертикальной графе числа n означают десятки, а в первой горизонтальной строке – единицы наблюдений. Например: для числа наблюдений 57 ($n=57$) – величина K будет находиться на пересечении по горизонтали на уровне 50 и вертикали – на уровне 7, и будет составлять 4,60, для $n=19$, $K=3,69$ и т.д.
- если $n=100$ и более, K определяется по двум нижним горизонтальным строкам.

Например: центральная наибольшая варианта равна 180, минимальная – 120, $n=200$:

$$\sigma = \frac{180-120}{K} = \frac{60}{5,49} = 10,93, \text{ где } K \text{ при } n=200=5,49.$$

Если имеются сведения о величине крайних вариантов (минимальной и максимальной), то можно пользоваться указанной методикой определения σ , т.к. разность между σ , вычисленной по этой очень простой методике и обычным методом, не очень велика.

Значение среднего квадратического отклонения

Средняя арифметическая является обобщающей характеристикой признака статистической совокупности, т.е. средняя арифметическая одной величиной характеризует весь вариационный ряд. Однако следует отметить, что чем больше варианты отличаются друг от друга по своей величине и чем больше они варьируют, тем менее точно средняя арифметическая характеризует свой вариационный ряд. Чем больше амплитуда, тем больше среднее квадратическое отклонение. Из этого следует, что две средние арифметические, имеющие одинаковое числовое выражение, но вычисленные из вариационных рядов с различной по величине амплитудой, не в одинаковой степени характеризуют свои вариационные ряды. Средняя, которая имеет меньшее среднее квадратическое отклонение, полученная из вариационного ряда, в котором варианты меньшей величиной отличаются друг от друга, будет своим размером больше приближаться к величине значительного большинства единиц ряда. В противоположность этому, средняя, имеющая большую σ , полученная из ряда с большей амплитудой, по своим размерам будет отличаться от реальных размеров значительного количества чисел ряда.

Например: при анализе заболеваемости раком гениталий в двух населенных пунктах установлены сроки выявляемости: в первом населенном пункте в среднем они выявлялись через 9 месяцев с момента заболевания со средним квадратическим отклонением 3 месяца. Это значит, что около 70% больных женщин взяты на учет в сроке 9 ± 3 месяца, т.е. в сроке от 6 до 12 месяцев с момента заболевания.

Во втором населенном пункте около 70% больных выявлено и взято на учет в среднем сроке с момента заболевания через $9 \pm 1,5$, т.е. от 7,5 до 10,5 месяцев, т.е. до 90% во втором пункте взяты на учет в пределах от 7,5 до 10,5 месяцев, в то время как в первом населенном пункте 70% в пределах от 6 до 12 месяцев с момента заболевания. Эти данные указывают, что при одинаковой средней величине равной 9, но при различном среднем квадратическом отклонении ($\pm 3,0$ и $\pm 1,5$) можно

сделать вывод, что во втором населенном пункте профилактическая работа по раннему выявлению больных с раком гениталий организована лучше, т.к. в пределах от 7,5 до 10,5 месяца с момента заболевания на учет взято около 90%, в то время как в первом пункте в пределах от 6 до 12 месяцев взято на учет 70% больных.

Таким образом, среднее квадратическое отклонение служит для измерения изменчивости (вариабельности) вариационного ряда.

Коэффициент вариации (C_V)

Среднее квадратическое отклонение является именованной величиной. Поэтому сравнивать изменчивость (вариабельность) двух средних величин, выраженных в различных единицах измерения (средний возраст беременных, заболевших эклампсией в годах, средняя продолжительность лечения в днях) по абсолютным размерам среднего квадратического отклонения невозможно. В этих случаях используется относительная величина – коэффициент вариации, который обозначается буквой латинского алфавита C_V и выражается в процентах (%).

Коэффициент вариации (C_V) – это отношение среднего квадратического отклонения к средней арифметической, умноженное на

100. Вычисляется по формуле: $C_V = \frac{\sigma}{M} \times 100$,

при $C_V < 10\%$ – слабое разнообразие признака,
 $C_V = 10-20\%$ – среднее разнообразие признака,
 $C_V > 20\%$ – сильное разнообразие признака.

Пример: средняя арифметическая равна 9, а среднее квадратическое отклонение – в одном случае 1,5, в другом – 3,0.

$$C_{V1} = \frac{1,5}{9} \times 100 = 16,6\% ; C_{V2} = \frac{3}{9} \times 100 = 33,3\% .$$

Меньший размер коэффициента вариации $C_{V1}=16,6\%$ по сравнению с $C_{V2}=33,3\%$ указывает на более высокий уровень организации профилактической работы органов и учреждений здравоохранения, направленный на выявление больных раком гениталий.

Самостоятельная работа студентов

Задача 1

Результаты измерения роста у группы мальчиков

Рост (см)	116	117	118	119	120	121	122	123	124	125	126	127	128	129	Всего
Число лиц	2	1	8	4	13	15	10	10	6	8	2	4	1	1	85

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 2

Результаты измерения роста у группы девочек

Рост (см)	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	Всего
Число лиц	2	2	7	8	4	10	9	20	16	10	15	4	7	3	1	118

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 3

Результаты изучения количества сигарет, выкуриваемых в день у группы мужчин

Кол-во сигарет	5	10	15	20	25	30	35	40	45	50	Всего
Число лиц	5	10	12	40	60	30	15	10	2	1	185

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 4

Результаты измерения максимального АД у студентов перед экзаменами

Максимальное АД (мм рт.ст.)	100-104	105-109	110-114	115-119	120-124	125-129	130-134	135-139	140-144	145-149	Всего
Число студентов	2	2	4	5	4	6	9	5	2	1	40

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 5

Результаты измерения частоты пульса у студентов перед экзаменами

Частота пульса (число ударов в минуту)	55-64	65-74	75-84	85-94	95-104	105-114	115-124	125-134	Всего
Число студентов	2	3	10	5	22	6	4	3	55

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 6

Число пораженных кариесом зубов у подростков

Вариант кариозности (по числу зубов)	1	2	3	4	5	6	7	Всего
Число случаев кариеса	2	3	5	10	5	3	2	30

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 7

Число случаев кариеса у девочек

Вариант кариозности (по числу зубов)	1	2	3	4	5	6	7	8	9	10	Всего
Число случаев кариеса	2	1	3	9	5	4	4	2	3	1	34

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 8

Средний рост 14-летних девочек

Рост в см	131-135	136-140	141-145	146-150	151-155	156-160	161-165	166-170	171-180	Всего
Число девочек	2	6	20	30	85	35	15	6	1	200

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Задача 9

Результаты исследования сыворотки на холестерин (в мг %) у мужчин (46-50 лет) во время гипертонического криза: 210, 215, 230, 231, 232, 236, 238, 240, 245.

Вычислить среднюю величину, среднее квадратическое отклонение и коэффициент вариации.

Тема № 5

ОЦЕНКА ДОСТОВЕРНОСТИ СРЕДНИХ, ОТНОСИТЕЛЬНЫХ ВЕЛИЧИН И РАЗНОСТИ РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

Цель темы:

Студент должен *знать*:

- Сущность понятий «достоверность средних и относительных величин», «критерий достоверности» результатов статистических исследований, их практическое применение.
- Методы расчета средних ошибок средних арифметических и относительных величин – m (ошибка репрезентативности), доверительных границ этих величин.
- Методику расчета достоверности разности средних и относительных величин (по критерию t).

Студент должен *уметь*:

- Владеть методикой расчета достоверности разности средних и относительных величин.
- Оценивать достоверность результатов исследования.

План изучения темы:

Разбор темы по контрольным вопросам

1. Определение понятия «достоверность».
2. Критерии достоверности – ошибки средних величин и показателей, формулы вычисления их при $n < 30$ и при $n > 30$.
3. Понятие о вероятности безошибочного прогноза (P) и критерии достоверности (t).
4. Доверительные границы средних и относительных величин и их определение при $n < 30$ и $n > 30$.

5. Достоверность разности между двумя средними величинами (или показателями) и ее определение.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Под достоверностью статистических показателей следует понимать степень их соответствия отображаемой ими действительности. Достоверными результатами считаются те, которые не искажают и правильно отображают объективную реальность.

Оценить достоверность результатов исследования значит перенести вероятность безошибочного прогноза, с которой результаты исследования, полученные на основании изучения выборочной совокупности, на генеральную совокупность.

Оценка достоверности результатов исследования предусматривает определение:

- 1) ошибок репрезентативности (средних ошибок средних арифметических и относительных величин) – m .
- 2) доверительных границ средних (или относительных) величин.

1. Определение средней ошибки средней величины (ошибка репрезентативности) – m .

Ошибка репрезентативности (m) является важнейшей статистической величиной, необходимой для оценки достоверности результатов исследования. Эта ошибка возникает в тех случаях, когда требуется по части охарактеризовать явление в целом.

Каждая средняя величина – M (средняя длительность лечения, средний рост, средняя масса тела, средний уровень белка в крови и др.), а также каждая относительная величина – P (уровень летальности, заболеваемости и др.) должны быть представлены со своей средней ошибкой – m . Так, средняя арифметическая величина выборочной совокупности (M) имеет ошибку репрезентативности, которая называется средней ошибкой средней арифметической (m_M) и определяется по формуле:

$$m_M = \frac{\sigma}{\sqrt{n}}$$

Как видно из этой формулы, величина средней ошибки средней арифметической прямо пропорциональна степени разнообразия признака и обратно пропорциональна корню квадратному из числа наблюдений. Следовательно, уменьшение величины этой ошибки при

определении степени разнообразия (σ) возможно путем увеличения числа наблюдений. На этом принципе основан метод определения достаточного числа наблюдений для выборочного исследования.

Относительные величины (P), полученные при выборочном исследовании, также имеют свою ошибку репрезентативности, которая называется средней ошибкой относительной величины и обозначается m_P .

Для определения средней ошибки относительной величины (P) используется следующая формула:

$$m_P = \sqrt{\frac{P \times q}{n}}$$

где P – относительная величина. Если показатель выражен в процентах, то $q=100-P$, если P – в промиллях, то $q=1000-P$, если P – в продецимиллях, то $q=10000-P$ и т.д.;

n – число наблюдений. При числе наблюдений менее 30 в знаменателе следует взять $n-1$.

Таблица 1

Примеры определения средних ошибок средних и относительных величин

$m_M = \frac{\sigma}{\sqrt{n}}$	$m_P = \sqrt{\frac{P \times q}{n}}$
<p>У 49 больных гипертиреозом исследован уровень пепсина $n=49$ $M=1,0$ г, % $\sigma=\pm 0,35$ г, % $m_M=?$</p> $m_M = \frac{0,39}{\sqrt{49}} = \pm 0,05$	<p>Исследовано 110 больных с абсцессом легкого, из них у 44 обнаружены дистрофические изменения пародонта $n=110$ $P = \frac{44 \times 100}{110} = 40\%$ лиц с дистрофическими изменениями пародонта $q=100-40=60\%$ лиц без дистрофических изменений пародонта $m_P=?$</p> $m_P = \sqrt{\frac{40 \times 60}{110}} = \pm 4,7\%$

Каждая средняя арифметическая или относительная величина, полученная на выборочной совокупности, должна быть представлена со своей средней ошибкой. Это дает возможность рассчитать доверительные границы средних и относительных величин, а также определить

достоверность разности (t) сравниваемых показателей (результатов исследования).

2. Определение доверительных границ M и P. Определяя для средней арифметической (или относительной) величины два крайних значения: минимально и максимально возможное, находят пределы, в которых может быть искомая величина генерального параметра. Эти пределы называют доверительными границами.

Доверительные границы – границы средних (или относительных) величин, выход за пределы которых вследствие случайных колебаний имеет незначительную вероятность.

Доверительные границы средней арифметической в генеральной совокупности определяют по формуле:

$$M_{ген} = M_{выб} \pm t m_M.$$

Доверительные границы относительной величины в генеральной совокупности определяют по следующей формуле:

$$P_{ген} = P_{выб} \pm t m_P,$$

где $M_{ген}$ и $P_{ген}$ – значения средней и относительной величин, полученных для генеральной совокупности;

$M_{выб}$ и $P_{выб}$ – значения средней и относительной величин, полученных для выборочной совокупности;

m_M и m_P – ошибки репрезентативности выборочных величин;

t – доверительный критерий (критерий точности, который устанавливают при планировании исследования);

t_m – доверительный интервал;

$t_m = \Delta$, где Δ – предельная ошибка показателя, полученного при выборочном исследовании.

Размеры предельной ошибки (Δ) зависят от коэффициента t, который избирает сам исследователь, исходя из необходимости получить результат с определенной степенью точности.

Понятие «вероятность безошибочного прогноза» (p) – это вероятность, с которой можно утверждать, что в генеральной совокупности M будет находиться в пределах $M \pm t m_M$ (или P – в пределах $P \pm t m_P$).

Если $n < 30$: при $p = 95\%$ } критерий t находится
при $p = 99\%$ } по таблице Стюдента (приложение 2)

Если $n > 30$: при $p = 95\%$ } $t = 2$
при $p = 99\%$ } $t = 3$

Для медицинских исследований степень вероятности безошибочного прогноза (p) должна быть не менее 95% (приложение 2).

Пример.

Определить доверительные границы среднего уровня пепсина у больных гипертиреозом с 95% вероятностью безошибочного прогноза (p=95%) (см. табл. 1).

Техника вычисления:

- Условие задачи: $n = 49$ $M_{выб} = 1 \text{ г, \%}$
 $m_M = \pm 0,05 \text{ г, \%}$ $p = 95\%$ (следовательно при $n=49, t=2$).
- Определить: доверительные границы средней величины в генеральной совокупности.
- Формула $M_{ген} = M_{выб} \pm t m_M$.
- Решение: $M_{ген} = 1 \text{ г, \%} \pm 2 \times 0,05 \text{ г, \%}$
 $M_{ген} \begin{cases} \text{не более } 1 \text{ г, \%} + 0,1 \text{ г, \%} = 1,1 \text{ г\%} \\ \text{не менее } 1 \text{ г, \%} - 0,1 \text{ г, \%} = 0,9 \text{ г\%} \end{cases}$
- Вывод: Установлено с вероятностью безошибочного прогноза (p=95%), что средний уровень пепсина в генеральной совокупности у больных с гипертиреозом не превышает 1,1 г, % и не ниже 0,9 г, %.

Пример.

Определить доверительные границы показателя частоты дистрофии пародонта у больных с абсцессом легкого с вероятностью безошибочного прогноза p=95% (см. табл. 1).

- Условие задачи: $n = 110$
 $P_{выб} = 40\%$
 $m_P = \pm 4,7 \%$
 $p = 95\%$ (следовательно при $n=110, t=2$).
- Определить: доверительные границы относительного показателя в генеральной совокупности.
- Формула: $P_{ген} = P_{выб} \pm t m_P$.
- Решение: $P_{ген} = 40\% \pm 2 \times 4,7\%$
 $P_{ген} \begin{cases} \text{не более } 40\% + 9,4 = 49,4\% \\ \text{не менее } 40\% - 9,4 = 30,6\% \end{cases}$
- Вывод: Установлено с 95% вероятностью безошибочного прогноза (p=95%), что дистрофические изменения пародонта в генеральной совокупности наблюдаются у больных с абсцессом легкого не чаще, чем в 49,4%, и не реже, чем в 30,6% случаев.

Определение достоверности разности средних (или относительных) величин (по критерию t)

Во всех случаях при сопоставлении двух сравниваемых величин возникает необходимость не только определить их разность, но и оценить ее достоверность. Достоверность разности величин, полученных при выборочных исследованиях, означает, что вывод об их различии может быть перенесен на соответствующие генеральные совокупности.

Достоверность выборочной разности измеряется доверительным критерием (критерием точности t), который рассчитывается по специальным формулам для средних и относительных величин.

Формула оценки достоверности разности сравниваемых средних величин такова:

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

и для относительных величин:

$$t = \frac{P_1 - P_2}{\sqrt{m_1^2 + m_2^2}},$$

где M_1, M_2, P_1, P_2 – параметры, полученные при выборочных исследованиях;

m_1 и m_2 – их средние ошибки;

t – критерий точности.

Разность достоверна при $t \geq 2$, что соответствует вероятности безошибочного прогноза, равной 95% и более ($p \geq 95,0\%$, приложение 2).

Пример.

Рост 16-летних школьников

	M	$\pm m$
Школа 1	165	1,02
Школа 2	162	0,59

$$t = \frac{165 - 162}{\sqrt{1,02^2 + 0,59^2}} = \frac{3}{1,2} = 2,5$$

Поскольку $t > 2$, можно с вероятностью безошибочного прогноза ($p > 95\%$ или $< 0,05$) утверждать, что рост школьников двух школ существенно (достоверно) различается.

Самостоятельная работа студентов

Задача 1

Заболееваемость гриппом между привитыми и непривитыми детьми

	Количество детей	Количество больных
Привитые	1000	100
Непривитые	400	52

Определить, существенны ли различия показателей результатов лечения между привитыми и непривитыми детьми.

Задача 2

Результаты исследования молока на зараженность бруцеллезом

Годы	Исследовано проб молока	Число положительных реакций на бруцеллез
1995	1786	109
1996	7248	510

Определить, существенны ли различия показателей положительных реакций на бруцеллез в сравниваемых годах.

Задача 3

Результаты дегельминтизации при лечении в амбулаторных и стационарных условиях

Место лечения	Число детей	Дегельминтизировано
В амбулатории	360	308
В стационаре	50	43

Определить, существенны ли различия показателей результатов лечения в амбулаторных условиях и в стационаре.

Задача 4

Восстановление трудоспособности у инвалидов от травматизма

Год	Число инвалидов	Трудоспособность восстановлена
1995	145	13
1996	157	12

Определить, существенны ли различия показателей восстановления трудоспособности в сравниваемых годах.

Задача 5

Распространенность кариеса зубов среди детей различного возраста

Возраст	Количество обследованных	Выявлено детей с кариесом
5 лет	434	304
6 лет	389	300

Определить, существенны ли различия показателей кариеса зубов в сравниваемых возрастах.

Задача 6

Выполнение плана периодических медицинских осмотров

	Подлежало осмотрам	Всего осмотрено
Рабочие совхозов	75,7	69,3
Колхозники	114,2	73,7

Определить, существенны ли различия показателей осматриваемых лиц.

Задача 7

Распространение атеросклероза по полу среди служащих по материалам прозектуры

Пол	Всего наблюдений	Умерли от атеросклероза
Мужчины	231	58
Женщины	93	16

Определить, существенны ли различия показателей умерших от атеросклероза по полу.

Задача 8

Выход на инвалидность от болезней органов кровообращения

Год	1995	1996
Численность рабочих и служащих	16600	18400
Число вышедших на инвалидность	23	29

Определить, существенны ли различия показателей выхода на инвалидность в сравниваемых годах.

Задача 9

Успеваемость неработающих студентов и совмещающих учебу с работой

Студенты	Средний балл М	$\pm m$
Неработающие	4,10	$\pm 0,09$
Совмещающие учебу с работой	3,86	$\pm 0,05$

Определить достоверность различий в сравниваемых группах и сделать выводы.

Задача 10

Средние показатели физического развития рабочих-подростков и учащихся ПТУ

Показатель	Возраст в годах	Учащиеся ПТУ		Рабочие-подростки	
		М	$\pm m$		
Рост (в см)	16	168,62	$\pm 0,50$	172,88	$\pm 1,02$
	17	169,49	$\pm 0,68$	173,50	$\pm 0,49$

Определить достоверность различий роста в сравниваемых группах и сделать выводы.

Задача 11

Успеваемость мужчин и женщин – студентов

Студенты	Средний балл	$\pm m$
Женщины	4.03	± 0.04
Мужчины	3.86	± 0.04

Определить достоверность различий в сравниваемых группах и сделать выводы.

Задача 12

Результаты изучения токсического воздействия сернистого кадмия на развитие плода (эксперимент на белых мышах)

Группы	Количество мышей	Вес плаценты животных (в гр.)	$\pm m$
Контрольная	14	0,63	$\pm 0,07$
Опытная	14	0,53	$\pm 0,02$

Определить достоверность различий в сравниваемых группах и сделать выводы.

Задача 13

Затраты времени на подсчет лейкоцитов у лаборантов с различным стажем работы

Стаж работы (в годах)	Затрата времени (в минутах) $M \pm m$
До 5 лет	$4,7 \pm 0,03$
5-9	$4,2 \pm 0,03$
10-14	$3,3 \pm 0,02$

Определить достоверность различий в сравниваемых группах и сделать соответствующие выводы.

Задача 14

Рост семилетних школьников

Возраст	Мальчики		Девочки	
	М	$\pm m$	М	$\pm m$
7 лет	120,9	0,57	119,6	0,65

Определить достоверность различий в сравниваемых группах и сделать выводы.

Задача 15

Рост тринадцатилетних школьников

Возраст	Мальчики		Девочки	
	М	$\pm m$	М	$\pm m$
13 лет	139,2	0,80	140,1	0,82

Определить достоверность различий в сравниваемых группах и сделать выводы.

Задача 16

Результаты обследования на выявление рака легкого в группах мужчин

Группа мужчин	Число лиц	Из них имели рак легкого
Курящие (не менее 1 пачки в день) в течение 10 лет	1800	20
Некурящие	1200	3

Определить, влияет ли курение на возникновение рака легкого. Сделать выводы.

Задача 17

Успеваемость выпускников в зависимости от стажа работы до поступления в ВУЗ

Стаж работы	Число выпускников	Из них имеющих средний балл «удовлетворительно»
До 4 лет	1500	50
От 6 лет и более	100	60

Определить, влияет ли стаж работы до поступления в институт на успеваемость студентов. Сделать выводы.

Тема № 6

ОЦЕНКА ДОСТОВЕРНОСТИ РАЗЛИЧИЯ СРАВНИВАЕМЫХ ГРУПП ПО КРИТЕРИЮ СООТВЕТСТВИЯ χ^2 (ХИ-КВАДРАТ)

Цель темы:

Студент должен *знать*:

- Отличие критерия соответствия (Хи-квадрат) от критерия достоверности t .
- Применение Хи-квадрат в статистических исследованиях.

Студент должен *уметь*:

- Владеть методикой вычисления Хи-квадрат.

План изучения темы:

Разбор темы по контрольным вопросам

1. Применение критерия соответствия χ^2 в статистических исследованиях.
2. Отличие χ^2 от критерия достоверности (t).
3. Понятие о «нулевой гипотезе».
4. Формула вычисления χ^2 .
5. Этапы вычисления χ^2 .

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Критерий χ^2 (в отличие от критерия t) применяется в тех случаях, когда нет необходимости знать величину (среднюю или относительный показатель) и требуется оценить достоверность различия не только двух, но и большего числа групп, выраженных в абсолютных величинах.

Так, критерий χ^2 может быть использован для ответа на следующие вопросы: существенно ли отличаются друг от друга группы вакцинированных и невакцинированных по распределению их на больных и здоровых (т.е. эффективна ли вакцина); существенно ли отличаются группы населения с разным среднедушевым доходом по распределению их на больных и здоровых (т.е. влияет ли материальное обеспечение на уровень заболеваемости) и т.д.

Критерий χ^2 (хи-квадрат) определяется по формуле:

$$\chi^2 = \frac{\sum(\varphi - \varphi_1)^2}{\varphi_1},$$

где φ – фактические (эмпирические) данные,
 φ_1 – «ожидаемые» (теоретические) данные, вычисленные на основании нулевой гипотезы;
 Σ – знак суммы.

«Нулевая гипотеза» – это предположение о том, что в сравниваемых группах отсутствует различие в распределении частот. Например, допускают одинаковое распределение больных и здоровых в группах вакцинированных и невакцинированных.

Определение критерия соответствия χ^2 основано на расчете разницы между фактическими и «ожидаемыми» данными. Чем больше эта разность ($\varphi - \varphi_1$), тем с большей вероятностью можно утверждать, что существуют различия в распределении сравниваемых выборочных совокупностей и, наоборот, чем меньше разность ($\varphi - \varphi_1$), тем меньше шансов на то, что сравниваемые выборочные совокупности различны между собой.

Последовательность расчета критерия χ^2 (5 этапов) представим на примере.

Требуется оценить при помощи критерия χ^2 , различаются ли по срокам постановки диагноза (менее 15 дней, 15 дней и более с момента обращения) группы больных ревматизмом, обратившихся за помощью в

поликлинику, где имеется кардиоревматологический кабинет, и в поликлинику, где он отсутствует.

Этапы расчета представлены в табл. 1.

Таблица 1

Распределение больных ревматизмом по срокам установления диагноза в поликлиниках с разной системой организации специализированной помощи

Кардиоревматологический кабинет	Число больных	I этап		II этап		III этап		IV этап		V этап	
		Фактические числа		«Ожидаемые» числа		$\varphi - \varphi_1$		$(\varphi - \varphi_1)^2$		$\frac{(\varphi - \varphi_1)^2}{\varphi_1}$	
		<15	15 и более	<15	15 и более	<15	15 и более	<15	15 и более	<15	15 и более
Есть	73	54	19	47,4	25,6	+6,6	-6,6	43,56	43,56	0,9	1,7
Нет	21	7	14	13,6	7,4	-6,6	+6,6	43,56	43,56	3,2	5,9
Итого	94	61	33								

Описание последовательности расчета критерия χ^2 :

I этап – распределение фактических данных (φ) по всем группам, суммирование итогов $61+33=94$.

II этап – определение ожидаемых величин (φ_1) на основе «нулевой гипотезы». Согласно нулевой гипотезе, допускают, что наличие или отсутствие в поликлинике кардиоревматологических кабинетов не влияет на сроки постановки диагноза у больных ревматизмом. В этом случае распределение двух групп больных, обслуживаемых с участием специалистов кардиоревматологического кабинета и без него, по срокам постановки диагноза (менее 15 дней и 15 дней и более с момента обращения в поликлинику) должно быть одинаковым и соответствовать итоговому фактическому распределению всех наблюдаемых больных, т.е. 61 и 33. При таком условии в первой группе (кардиоревматологический кабинет есть) «ожидаемое» число больных со сроком установления диагноза менее 15 дней определяется из следующей пропорции:

$$\frac{94 - 61}{73 - x} = \frac{73 \times 61}{94} = 47,4.$$

«Ожидаемое» число больных со сроком установления диагноза 15 дней и более получается путем вычисления $73-47,4=25,6$. Подобным же способом рассчитывают «ожидаемые» числа больных второй группы. Со сроком меньше 15 дней $= \frac{21 \times 61}{94} = 13,6$; со сроком 15 и более дней =

$21-13,6=7,4$. Полученные «ожидаемые» числа по всем группам заносят в таблицу.

III этап – определяют разность между фактическими и «ожидаемыми» числами ($\varphi - \varphi_1$). Первая группа больных ($\varphi - \varphi_1$)= $54-47,4=+6,6$; $19-25,6=-6,6$. Вторая группа больных ($\varphi - \varphi_1$)= $7-13,6=-6,6$; $14-7,4=+6,6$. В числовом отношении разность между фактическими и «ожидаемыми» числами ($\varphi - \varphi_1$) одинакова, что позволяет проверить правильность расчетов.

IV этап – определяют квадрат разностей $(\varphi - \varphi_1)^2$ по всем группам.

V этап – квадрат разности делят на «ожидаемое» число $\frac{(\varphi - \varphi_1)^2}{\varphi_1}$ во всех группах и результаты заносят в таблицу.

Например, $\frac{43,56}{47,4} = 0,9$, $\frac{43,56}{25,6} = 1,7$ и т.д.

VI этап – критерий соответствия χ^2 определяется путем суммирования предыдущих результатов $\frac{(\varphi - \varphi_1)^2}{\varphi_1}$ по всем группам:

$$\chi^2 = \sum \frac{(\varphi - \varphi_1)^2}{\varphi_1} = 0,9 + 1,7 + 3,2 + 5,9 = 11,7.$$

Для оценки критерия χ^2 учитывают число граф (R) и число строк (S) распределения фактических чисел (без итоговых) и на основании этих данных вычисляют так называемое число степеней свободы $n^1 = (R-1) \times (S-1)$. В нашем примере $R=2$, $S=2$, при этом $n^1 = (2-1) \times (2-1) = 1$. Полученную величину $\chi^2=11,7$, при $n^1=1$ оценивают по специальной таблице (табл. 2).

Для того, чтобы опровергнуть «нулевую гипотезу», вычисленный критерий соответствия (χ^2) должен быть равен или больше табличного (критического) значения χ^2 при уровне вероятности «нулевой» гипотезы $p_1=5\%$.

Как видно из таблицы 2, при числе степеней свободы $n^1=1$ и уровне вероятности подтверждения «нулевой гипотезы» $p_1=5\%$ критическое значение $\chi^2=3,8$; для уровня $p_2=1\%$ значение $\chi^2=6,6$, а для $p_3=0,1\%$ значение $\chi^2=10,8$. Вычисленная же нами величина $\chi^2=11,7$ больше критического значения χ^2 , при котором уровень вероятности подтверждения «нулевой гипотезы» будет равен 0,1%. Это дает основание опровергнуть

Таблица 2

Таблица оценки значений критерия соответствия χ^2

Число степеней свободы (n^1)	Уровень вероятности подтверждения «нулевой гипотезы» (p)		
	$p_1=5\%$	$p_2=1\%$	$p_3=0,1\%$
1	3,8	6,6	10,8
2	5,9	9,2	13,8
3	7,8	11,3	16,3
4	9,5	13,3	18,5
5	11,1	15,1	20,5
и т.д.	и т.д.		

«нулевую гипотезу» и признать существенными различия в распределении по срокам постановки диагноза двух сравниваемых групп больных ревматизмом (обследованных специалистом, работающим в кардиоревматологическом кабинете, и при отсутствии такого специалиста). Эти данные позволяют утверждать, что организация специализированных кардиоревматологических кабинетов в поликлиниках позволяет снизить сроки обследования больных ревматизмом.

Самостоятельная работа студентов

Задача 1

Количество детей в семьях в зависимости от уровня образования матери (на 100 семей)

Образование матери	Всего членов семьи	В том числе детей
Начальное	470	246
Среднее	466	229
Среднее специальное, неполное высшее, высшее	430	188

Определить, влияет ли уровень образования матери на количество детей в семье. Сделать выводы.

Задача 2

Число электросварщиков с поражениями бронхов и легких в зависимости от стажа работы

Стаж работы (в годах)	Поражение бронхов и легких		Итого
	Выявлено	Не выявлено	
До 5 лет	25	159	184
5-9	35	91	126
10-14	39	45	84
15 и больше	57	22	79
Всего	156	317	473

Определить, влияет ли стаж работы на заболеваемость бронхов и легких. Сделать выводы.

Тема № 7

ВЫБОРОЧНЫЙ МЕТОД В СТАТИСТИЧЕСКОМ ИССЛЕДОВАНИИ

Цель темы:

Студент должен *знать*:

- Основные положения методики определения объема выборки при средних (M) и относительных (P) величинах.

Студент должен *уметь*:

- Вычислять необходимый объем выборочного наблюдения.

План изучения темы

Разбор темы по контрольным вопросам

1. Формула расчета объема выборки для средних величин при неизвестной численности генеральной совокупности.
2. Формула расчета объема выборки для средних величин при известной численности генеральной совокупности.
3. Формула расчета объема выборки для относительных величин при неизвестной численности генеральной совокупности.
4. Формула расчета объема выборки для относительных величин при известной численности генеральной совокупности.
5. Понятие о вероятности безошибочного прогноза (P), способ определения.
6. Понятие о максимальной ошибке (Δ) и ее предельная величина при $P=95\%$.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Часть генеральной совокупности, отобранная по определенной методике для обследования и изучения, называется выборочной совокупностью. Характеристиками выборочной совокупности являются относительные показатели (P) и средние величины (M). Они в какой-то степени отличаются от аналогичных количественных характеристик генеральной совокупности (P_0 , M_0). Эта разность между количественными характеристиками генеральной и выборочной совокупностей на-

зывается ошибкой репрезентативности (m), которая свойственна и определяется только к относительным показателям и средним величинам. Ошибка репрезентативности при выборочной совокупности возникает потому, что изучается не вся генеральная совокупность, а только часть, которая не вполне точно ее воспроизводит (представляет).

Для определения степени точности выборочного наблюдения, необходимо оценить величину ошибки (m), которая может возникнуть в процессе выборки. Величина или размер ошибки позволяет установить степень точности выборочного наблюдения, обосновать меры, направленные на уменьшение отклонения выборочных характеристик от соответствующих характеристик генеральной совокупности.

Теория выборочного метода возникла как теория случайной выборки, при которой создаются условия для действия закона больших чисел (А.М. Мерков, Л.Е. Поляков, 1974), он указывает, что при достаточно большом числе наблюдений более ярко выступают те черты, которые наиболее свойственны всем единицам данной совокупности. Действие закона больших чисел проявляется в стремлении ошибки репрезентативности к нулю при бесконечном увеличении числа наблюдений выборочной совокупности. Однако не величина численности выборки определяет степень точности, а величина ошибки репрезентативности.

Репрезентативность обеспечивается предварительным расчетом объема выборки и специальным методом отбора единиц наблюдения.

Для иллюстрации методики расчета объема (числа) наблюдений (n) с целью получения достоверных средних величин или относительных показателей при планировании выборочного исследования приведем следующие примеры (числа взяты произвольно).

1. Определить объем наблюдений (n) для установления относительно достоверного среднего числа дыхания роженицы во время схватки, используя исходные данные, полученные из литературы или при предварительном (пробном) исследовании (M , σ , Δ , pq).

Допустим, что получены следующие данные при аналогичных исследованиях: $M=34$; $\sigma=3,2$; $\Delta_x=1$. При условии, если вероятность безошибочного прогноза $P=95\%$, то $t=2$ (критерий достоверности). Подставляем исходные данные в формулу:

$$n = \frac{t^2 \sigma^2}{\Delta^2} = \frac{2^2 \cdot 3,2^2}{1^2} = \frac{4 \cdot 10}{1} = \frac{40}{1} = 40.$$

Вывод: для определения среднего дыхания у рожениц во время родовой схватки достаточно провести измерение у сорока рожениц.

2. Если конечный результат должен быть выражен относительными показателями, то объем выборки определяется по формуле:

$$n = \frac{t^2 pq}{\Delta^2},$$

где n – требуемое число наблюдений (объем выборки);

t – доверительный коэффициент, который при вероятности безошибочного прогноза (P), равного 95% составляет 2;

P – относительный показатель в % (%_o), который берется из опубликованной литературы, или получен при предварительном исследовании, или берется с расчетом, чтобы произведение было максимальным, что возможно при $P=50\%$, т.е. $p=q$ или $pq=50 \cdot 50=2500$; Δ – доверительный интервал ($P \pm$), т.е. предельно допустимая ошибка – « $tm\%$ ».

Например: определить частоту развития позднего токсикоза среди беременных группы «риска» по возникновению этой патологии. С этой целью необходимо заранее рассчитать достаточную численность беременных, угрожаемых по возникновению позднего токсикоза и наблюдать их до конца (до исхода) беременности. Если в литературе нет исходных данных, закономерности этой патологии не изучены и не представляется возможным провести пробные исследования, то исследователь допускает (предполагает), что среди беременных из группы «риска» поздний токсикоз возникает в 50%, в то же время среди этого контингента беременных у 50% не возникает поздний токсикоз (q). Далее исследователь задает предельную ошибку, равную 8. При вероятности безошибочного прогноза $P=95\%$, критерий вероятности будет равен 2. Подставляя все вышеуказанные исходные данные в формулу, получим:

$$n = \frac{t^2 \times pq}{\Delta^2} = \frac{4 \times 50 \times 50}{8^2} = \frac{10000}{64} = 156$$

Вывод: для определения частоты возникновения позднего токсикоза среди беременных группы «риска» по возникновению этой патологии необходимо взять под постоянное наблюдение 156 беременных группы «риска» и наблюдать их до исхода беременности при одновременном профилактическом лечении.

Приведенные примеры относятся к клинко-статистическим исследованиям, при которых группы исследования в большинстве своем не многочисленны (40, 156 и т.д.). Следует отметить, что

Самостоятельная работа студентов

Задача 1

Изучить длительность пребывания в хирургическом отделении стационара больных острым холециститом для того, чтобы наметить пути повышения эффективности использования коечного фонда.

При расчете необходимого числа наблюдений вероятность безошибочного прогноза p примите равной 95%; $\sigma=1,9$ дня, $\Delta=0,5$ дня.

Задача 2

Изучить исходы заболевания у больных, лечившихся в стационаре по поводу язвенной болезни желудка, чтобы наметить пути улучшения результатов лечения.

При расчете необходимого числа наблюдений для определения среди больных числа нуждающихся в оперативном лечении (в %) $\Delta=5\%$, вероятность безошибочного прогноза $p=95\%$; учтите, что по данным ранее проводившихся аналогичных исследований терапевтическое лечение не дало результатов в 25% случаев.

Задача 3

Изучить организацию диспансеризации больных ревматизмом, чтобы наметить пути повышения ее эффективности.

При расчете необходимого числа наблюдений для определения доли (в %) больных, снятых с учета в связи с излечением, $\Delta=5\%$, вероятность безошибочного прогноза $p=95\%$; учтите, что по данным ранее проведенных аналогичных исследований доля снятых с учета в связи с выздоровлением составляла от 60 до 85%.

Задача 4

Изучить заболеваемость с временной утратой трудоспособности в связи с гипертонической болезнью у больных, состоящих на диспансерном учете по поводу этого заболевания, чтобы наметить пути снижения временной нетрудоспособности у этой группы больных.

При расчете необходимого числа наблюдений вероятность безошибочного прогноза p примите равной 95%, $\sigma=1,5$ случая, предельная ошибка $\Delta=0,2$ случая.

Задача 5

Изучить организацию диспансерного наблюдения за больными бронхиальной астмой, чтобы наметить пути его улучшения.

вышеуказанные формулы $n = \frac{t^2 \sigma^2}{\Delta^2}$, $n = \frac{t^2 pq}{\Delta^2}$ предназначены для определения объема выборки при неизвестной численности (объеме) генеральной совокупности, в то время как при известном объеме генеральной совокупности для вычисления объема выборки применяются следующие формулы:

$$1) n = \frac{t^2 \sigma^2 N}{N \Delta^2 + \sigma^2 t^2}$$

при конечном результате, выраженном средними величинами, где N – объем генеральной совокупности

$$2) n = \frac{t^2 pq N}{N \Delta^2 + pqt^2}$$

при конечном результате, выраженном относительными показателями, где N – объем (численность) генеральной совокупности (Е.Н. Шиган, 1987). Если при клинико-статистических исследованиях группы наблюдения, как правило, небольшие, то при социально-медицинских исследованиях они значительно большего объема.

3. Например: в г. N с числом женщин 80000 в возрасте 15 лет и старше, зарегистрировано в течение года 20% доброкачественных новообразований гениталий. С целью установления причинно-следственных закономерностей возникновения указанной патологии провести выборочное социально-медицинское исследование. Одним из важных разделов для выполнения этого исследования является определение объема выборки при условии, что генеральная совокупность (N) равна 80000 человек, $P=20,0\%$, $q=1000,0-20,0=980,0\%$, предельная ошибка $\Delta=2\%$, при вероятности безошибочного прогноза $P=95\%$ и критерия достоверности $t=2$. Подставляя исходные данные в формулу, получим:

$$n = \frac{t^2 pq N}{N \Delta^2 + pqt^2} = \frac{4 \times 20,0 \times 980,0 \times 80000}{80000 \times 4 + 20,0 \times 980,0 \times 4} = 15742,9 \approx 16000,0$$

Вывод: чтобы установить с необходимой точностью закономерность частоты возникновения доброкачественных новообразований гениталий необходимо обследовать 16000 женщин. Эти данные свидетельствуют, что в большинстве социально-медицинских исследований для выявления закономерностей того или иного явления необходим значительный объем выборки.

При расчете необходимого числа наблюдений для выявления процента больных, нерегулярно наблюдающихся врачом, $\Delta = 5\%$, вероятность безошибочного прогноза $p = 95\%$; учтите, что в ранее проводившихся исследованиях этот показатель составлял от 10% до 32%.

Задача 6

Изучить результаты стационарного лечения детей, больных лимфогранулематозом, чтобы наметить пути повышения эффективности лечения.

При расчете необходимого числа наблюдений для определения процента выписавшихся из стационара с улучшением $\Delta = 5\%$, вероятность безошибочного прогноза $p = 95\%$; учтите, что по данным ранее проводившихся исследований этот показатель составлял от 60% до 80%.

Задача 7

Изучить сроки стационарного лечения детей, больных ревматизмом, чтобы наметить пути повышения эффективности использования коечного фонда при лечении этой группы больных.

При расчете необходимого числа наблюдений для определения средней длительности пребывания в стационаре госпитализированных в неактивной фазе ревматизма, вероятность безошибочного прогноза p примите равной 95%, $\sigma = 3.5$ дня, $\Delta = 0.5$ дня.

Задача 8

Изучить длительность стационарного лечения детей, оперированных по поводу паховой грыжи, чтобы наметить пути повышения эффективности использования коечного фонда при лечении этой группы больных.

При расчете необходимого числа наблюдений для определения средней длительности предоперационного периода у детей, оперированных по поводу паховой грыжи, вероятность безошибочного прогноза p примите равной 95%, $\sigma = 3.8$ дня, $\Delta = 0.5$ дня.

Задача 9

Изучить эффективность диспансеризации детей, больных хронической пневмонией, чтобы наметить пути повышения эффективности диспансерного наблюдения.

При расчете необходимого числа наблюдений для определения процента снятых с диспансерного учета в связи с излечением $\Delta = 5\%$, вероятность безошибочного прогноза $p = 95\%$; учтите, что в аналогичных исследованиях этот показатель составлял от 20% до 60%.

Задача 10

Изучить эффективность диспансеризации школьников, страдающих язвенной болезнью двенадцатиперстной кишки, чтобы наметить пути повышения эффективности диспансерного наблюдения.

При расчете необходимого числа наблюдений для определения средней длительности заболевания вероятность безошибочного прогноза p примите равной 95%, $\sigma = 2.5$ дня, $\Delta = 0.25$ года.

Задача 11

Изучить причины нарушения сроков проведения прививок у детей первых трех лет жизни, чтобы наметить пути совершенствования организации прививок детской поликлиники.

При расчете необходимого числа наблюдений для определения процента отводов от прививок $\Delta = 5\%$, вероятность безошибочного прогноза p составляет 95%; учтите, что по данным ранее проводившихся исследований доля детей, имевших отводы от прививок, составляла от 10% до 25%.

Задача 12

Изучить причины несвоевременного проведения вакцинации и ревакцинации детей против дифтерии, коклюша, столбняка, чтобы наметить пути улучшения работы прививочного кабинета детской поликлиники.

При расчете необходимого числа наблюдений для определения частоты отводов от прививок $\Delta = 3\%$, вероятность безошибочного прогноза p составляет 95%; учтите, что в ранее проводившихся исследованиях этот показатель составлял от 6% до 15%.

Тема № 8

СТАНДАРТИЗАЦИЯ ПОКАЗАТЕЛЕЙ

Цель темы:

Студент должен *знать*:

- Основные положения метода стандартизации статистических показателей.
- Этапы прямого метода стандартизации.

Студент должен *уметь*:

- Вычислять стандартизованные показатели.

План изучения темы

Разбор темы по контрольным вопросам

1. Понятие о стандартизованных показателях.
2. Методы стандартизации.
3. Суть и применение прямого метода стандартизации.
4. Этапы прямого метода стандартизации.
5. Способы выбора (или расчета) стандарта.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Практическому врачу или научному работнику нередко приходится сравнивать статистические показатели, вычисленные в качественно неоднородных по составу группах. И в этих случаях применяется метод стандартизации показателей, который устраняет (элиминирует) различия в изучаемых группах. **Стандартизация** – метод расчета условных (стандартизованных) показателей, заменяющих общие интенсивные величины в тех случаях, когда их сравнение затруднено из-за несопоставимости состава групп. Полученные стандартизованные показатели при сопоставлении их с обычными интенсивными показателями позволяют сделать вывод, связаны ли различия в интенсивных показателях с неоднородностью составов сравниваемых групп. Стандартизованные показатели являются условными и не отражают истинных размеров изучаемых явлений.

Например, сравнивая показатели рождаемости на двух территориях, следует иметь в виду, что на одной территории в структуре населения могут преобладать мужчины, на другой – женщины; или на одной

территории преобладают лица молодых возрастов, на другой – преимущественно люди средних и старших возрастов. Естественно, что с указанными обстоятельствами могут быть связаны уровни показателей рождаемости.

Аналогичным образом следует относиться и к показателям общей смертности в группах населения неоднородных, например, по возрастному составу, т.к. уровень смертности в разных возрастных группах неодинаков, и это обстоятельство может влиять на уровень общего показателя. Возможное влияние неоднородности составов следует иметь в виду при сопоставлении показателей заболеваемости, травматизма, летальности и многих других, если они заведомо вычислены в неоднородных группах (по полу, возрасту, составу больных по отделениям, профессиям, стажу работы и т.д.).

Для того чтобы при сравнении интенсивных показателей, вычисленных по разнородным совокупностям, увидеть влияние на эти показатели непосредственно самих составов по тому или иному признаку, нужно исключить (элиминировать) эту разнородность. Это достигается применением метода стандартизации – прямого, косвенного или обратного. Суть наиболее часто применяемого прямого метода состоит в вычислении показателей, которые имели бы место, если бы состав совокупностей был одинаков (по возрасту, полу или другому признаку). За стандарт можно принять данные одной из сравниваемых групп, суммируя или полусумму обеих групп.

Например, на основе приведенных в табл. 1 данных требуется:

1. Вычислить стандартизованные показатели, используя прямой метод стандартизации.
2. Сравнить:
 - Общие показатели, рассчитанные обычным способом.
 - Частные показатели в отдельных группах.
 - Стандартизованные и обычные показатели.
3. Сделать вывод, вытекающий из сопоставления обычных и стандартизованных показателей в сравниваемых группах.

Таблица 1

Распределение больных и умерших по отделениям больниц А и Б (условные данные)

Отделение	Больница А		Больница Б	
	Число прошедших больных	Из них умерло	Число прошедших больных	Из них умерло
Терапевтическое	600	30	200	12
Хирургическое	300	6	700	21
Инфекционное	100	4	100	5
Всего	1000	40	1000	38
Летальность в %	4,0		3,8	

За стандарт принять полусумму больных по каждому отделению больниц А и Б.

Техника вычисления:

Расчет стандартизованных показателей летальности прямым методом.

I этап метода стандартизации: расчет интенсивных показателей (в данном случае летальности) в двух сравниваемых совокупностях. Летальность рассчитывается в процентах (табл. 2).

Таблица 2

Показатели летальности по отделениям и по больницам А и Б в целом (в % к числу больных) (I этап метода стандартизации)

Отделение	Больница А	Больница Б
Терапевтическое	5,0	6,0
Хирургическое	2,0	3,0
Инфекционное	4,0	5,0
По больнице в целом	4,0	3,8

Следовательно, если из 600 больных терапевтического отделения умерло 30, то показатель летальности $(30 \times 100) : 600 = 5\%$ (и так по всем отделениям больниц А и Б). По больницам А и Б в целом летальность составляет 4,0 и 3,8% соответственно (табл. 2).

II этап. Определение стандарта. Поскольку условием задачи предусмотрено за стандарт принять полусумму больных по каждому отделению больниц А и Б, то проводили следующие вычисления (табл. 3).

Таблица 3

Расчет стандарта (II этап метода стандартизации)

Отделение	Стандарт (число больных)
Терапевтическое	$600 \text{ (больница А)} + 200 \text{ (больница Б)} : 2 = 400$
Хирургическое	$(300 + 700) : 2 = 500$
Инфекционное	$(100 + 100) : 2 = 100$
Всего	$(1000 + 1000) : 2 = 1000$ или $400 + 500 + 100 = 1000$

III этап. Расчет ожидаемых величин (в данном случае числа умерших) в каждой группе стандарта.

Если из 600 больных терапевтического отделения больницы А умерло 30, то сколько умерло бы, если число прошедших больных составляло бы 400 (стандарт)? Составляем пропорцию:

$$\frac{600 - 30}{400 - x} = \frac{400 \times 30}{600} = 20.$$

Далее: если из 200 больных терапевтического отделения больницы Б умерло 12, то сколько умерло бы, если число прошедших больных составляло бы 400 (т.е. число больных в каждом из терапевтических отделений было бы одинаковым, «стандартным»). Из аналогичной пропорции получаем ожидаемую величину – 24. Полученные аналогичным образом данные по всем отделениям и по больницам А и Б в целом представлены в табл. 4.

Таблица 4

Расчет ожидаемых величин (числа умерших) в каждой группе (по отделениям и больницам А и Б в целом) стандарта (III этап метода стандартизации)

Отделение	Ожидаемое число умерших	
	Больница А	Больница Б
Терапевтическое	$600 - 300$ $400 - x$ $x = 20$	$200 - 12$ $400 - x$ $x = 24$
Хирургическое	$300 - 6$ $500 - x$ $x = 10$	$700 - 21$ $500 - x$ $x = 15$
Инфекционное	$100 - 4$ $100 - x$ $x = 4$	$100 - 5$ $100 - x$ $x = 5$
Всего	34	44

IV этап. Расчет стандартизованных показателей. При условии, что в каждой больнице число прошедших больных составляло 1000 (стандарт, табл. 3), рассуждаем следующим образом. Из 1000 прошедших больных в больнице А ожидаемое число умерших составляет 34, следовательно, показатель летальности вычисляется на основе пропорции: $1000 - 34$ $1000 - 44$
 $100 - x$ $x = 3,4$, а в больнице Б: $100 - x$ $x = 4,4$.

Это и есть стандартизованные показатели, т.е. показатели, вычисленные при условии, что состав больных в каждой из больниц одинаковый (стандартный).

Анализ летальности в больницах А и Б выявил следующее:

- показатель летальности по больнице А в целом выше, чем по больнице Б (4,0% > 3,8%);
- и показатели летальности по всем отделениям, напротив, выше в больнице Б (в терапевтических отделениях – 6,0 и 5,0%, в хирургических – 3,0 и 2,0%, в инфекционных отделениях – 5,0 и 4,0%).

Более высокий показатель летальности в больнице А объясняется различиями в составах больных и преобладанием в ней больных терапевтического профиля, имеющих самую высокую летальность, а более низкий показатель летальности в больнице Б обусловлен преобладанием в ней больных хирургического профиля, имеющих самую низкую летальность.

Стандартизованный показатель летальности выше в больнице Б. Таким образом, если составы больных в больницах А и Б были бы одинаковыми, то летальность была бы выше в больнице Б.

Самостоятельная работа студентов

Задача 1

Распределение случаев инфаркта миокарда в секционном материале больниц

Возраст	Мужчины		Женщины	
	Вскрытый	Инфаркт миокарда	Вскрытый	Инфаркт миокарда
20-29	260	-	256	-
30-39	381	3	325	12
40-49	440	24	345	12
50-59	658	95	450	33
60-69	578	104	431	70
70-79	311	69	246	49
80 и старше	92	13	77	10
Итого	2720	308	2130	186

Вычислить простые и стандартизованные показатели распространенности инфаркта миокарда по полу, сделать выводы. За стандарт взять полусумму возрастного состава обоих полов.

Задача 2

Распределение рабочих по полу и случаев заболевания с временной утратой трудоспособности на двух заводах

Пол	Число рабочих		Число заболеваний с временной утратой трудоспособности	
	Завод №1	Завод №2	Завод №1	Завод №2
Мужчины	6000	12000	500	960
Женщины	10000	4000	500	200
Всего:	16000	16000	1000	1160

Вычислить простые и стандартизованные показатели заболеваемости, сделать выводы. За стандарт взять сумму чисел рабочих обоих предприятий по каждой группе. Сделать выводы.

Задача 3

Летальность детей в двух больницах

Возраст	Больница №1		Больница №2	
	Число больных	Число умерших	Число больных	Число умерших
До 1 года	1500	90	500	40
1-4	500	10	500	15
5-7	500	5	1500	22
Всего	2500	105	2500	77

Вычислить простые и стандартизованные показатели летальности в детских больницах и сделать выводы. За стандарт взять полусумму больных в 2-х больницах.

Задача 4

Распределение рабочих в цехах и числа заболевших по профессиям

Профессия	Цех А		Цех Б	
	Число рабочих	Число больных	Число рабочих	Число больных
Прессовщики	255	41	124	22
Вулканизаторы	153	11	215	19
Вальцовщики	111	5	364	23
Всего	519	57	703	64

Вычислить простые и стандартизованные показатели заболеваемости в двух цехах. Сделать выводы. За стандарт взять состав рабочих по профессиям в обоих цехах.

Задача 5

Распределение случаев инфаркта миокарда в секционном материале больниц

Возраст	Городское население		Сельское население	
	Вскрытый	Инфаркт миокарда	Вскрытый	Инфаркт миокарда
20-29	270	-	246	1
30-39	393	1	313	2
40-49	493	27	292	9
50-59	774	109	334	19
60-69	744	147	265	27
70-79	451	107	106	11
80 и старше	145	23	24	-
Итого	3270	414	1580	68

Вычислить простые и стандартизованные показатели распространенности инфаркта миокарда среди городского и сельского населения и сделать выводы. За стандарт взять полусумму возрастного состава городского и сельского населения.

Задача 6

Распределение больных с непроходимостью кишечника и числа умерших от этого заболевания в больницах А и Б по срокам поступления

в стационар от начала заболевания

Срок поступления в стационар от начала заболевания	Больница А		Больница Б	
	Число больных	Число умерших	Число больных	Число умерших
До 6 часов	250	42	170	20
6-24 часов	273	49	215	37
Свыше 24 часов	201	30	415	116
Всего	724	121	800	173

Вычислить простые и стандартизованные показатели летальности в обеих больницах. Сделать выводы. За стандарт взять состав больных обеих больниц по срокам поступления в стационар от начала заболевания.

Задача 7

Распределение городского и сельского населения и числа умерших по возрастам

Возраст	В городах		В сельской местности	
	Численность населения	Умерло	Численность населения	Умерло
До 15 лет	9795	152	22651	369
15-49	20269	82	29386	94
50 и старше	3716	85	9181	187
Всего:	33780	319	61220	650

Вычислить простые и стандартизованные показатели смертности городского и сельского населения. Сделать выводы. За стандарт взять возрастной состав городского населения.

Тема № 9

ИЗМЕРЕНИЕ СВЯЗИ МЕЖДУ ЯВЛЕНИЯМИ ИЛИ ПРИЗНАКАМИ. КОРРЕЛЯЦИЯ

Цель темы:

Студент должен *знать*:

- Сущность корреляционной связи между признаками.
- Оценку достоверности коэффициента корреляции.

Студент должен *уметь*:

- Владеть методикой вычисления коэффициента корреляции по методу рангов и оценки характера, силы и достоверности связи.

План изучения темы

Разбор темы по контрольным вопросам

1. Виды связи между явлениями и признаками и их характеристика.
2. Определение понятия «корреляционная связь».
3. Коэффициент корреляции и способы его вычисления.
4. Определение характера и силы связи по коэффициенту корреляции.
5. Формула вычисления коэффициента корреляции по способу квадратов (Пирсона).
6. Формула вычисления коэффициента ранговой корреляции (Спирмана).
7. Формулы вычисления достоверности коэффициента корреляции (Пирсона и Спирмана).
8. Коэффициент детерминации, применение и техника вычисления.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Все явления, процессы, происходящие в природе и обществе, находятся во взаимосвязи. Различают две формы связи между явлениями (признаками): *функциональную* и *корреляционную*. Функциональная связь характеризует строгую зависимость явлений (признаков), при которой изменение одного явления влечет за собой строго определенное изменение другого. Функциональная связь свойственна для химико-физиологических и природных процессов (явлений). Наглядными примерами функциональной связи являются: расширение объема тела по мере повышения температуры нагревания; увеличение долготы дня при строго определенном одновременном уменьшении долготы ночи и др.

При корреляционной связи изменение какого-либо признака связано не исключительно с изменением другого. Корреляционная связь проявляется в большинстве своем при массовом сопоставлении изучаемых явлений. В каждом отдельном случае она может не

проявляться. Такой вид связи характерен для социально-медицинских процессов, клинической медицины и биологии.

По своему характеру корреляционная связь может быть прямой (положительной со знаком +) и обратной (отрицательной со знаком -); по силе взаимодействия – сильной, средней, слабой. Кроме того, связь может отсутствовать или быть полной (табл. 1).

Таблица 1

Схема оценки характера и силы корреляционной связи по коэффициенту корреляции

Сила связи	Характер связи	
	Прямая	Обратная
Полная	+1	-1
Отсутствует	0	0
Сильная	от+1 до+07	от -1 до -07
Средняя	от +07 до +03	от -07 до 03
Слабая	от+03 до 0	от-03 до 0

При прямой связи с увеличением значений одного признака возрастает среднее значение другого признака. Например, с повышением температуры тела увеличивается частота пульса у большинства инфекционных больных; с увеличением роста ребенка увеличивается масса его тела. При обратной связи: с увеличением одного признака убывает среднее значение другого признака. Например, чем ниже температура воздуха в осенний период, тем выше заболеваемость детей острым бронхитом.

Для измерения связи имеется несколько способов вычисления коэффициента корреляции (r_{xy}). Выбор этих способов в каждом конкретном случае определяется целью, характером и объемом исследования, наличием или отсутствием вычислительной техники. Наиболее часто применяемым методом получения коэффициента корреляции является *способ квадратов (Пирсона)* и используется в случаях, где $n > 30$.

Формула коэффициента корреляции:

$$r_{xy} = \frac{\sum dx \times dy}{\sqrt{\sum d_x^2 \times \sum d_y^2}},$$

где: r_{xy} – коэффициент корреляции;

x и y – признаки, между которыми определяется связь;

dx и dy – отклонения каждой варианты от средней величины, вычисленной в ряду признака (x) и в ряду признака (y);

Σ – знак суммы.

Для того, чтобы убедиться в том, что коэффициент корреляции, вычисленный по данным выборочного исследования, будет соответствовать размеру связи в генеральной совокупности, необходимо определить среднюю ошибку коэффициента корреляции (m_r) и критерий t .

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n - 2}}, \quad t = \frac{r_{xy}}{m_r},$$

где m_r – средняя ошибка коэффициента корреляции;

$n-2$ – число парных вариантов.

t оценивается по таблице критерия t (приложение 2).

Для примера определим наличие хронических гинекологических заболеваний в зависимости от возраста (табл. 2).

Подставляя результаты расчетов в формулу, получим следующие данные:

$$r_{xy} = \frac{\sum dx \times dy}{\sqrt{\sum d^2 x \times \sum d^2 y}} = \frac{11610}{\sqrt{440 \times 310791}} = \frac{11610}{11694} = +0,9.$$

Полученный коэффициент корреляции свидетельствует, что между возрастом женщин и наличием у них хронических гинекологических заболеваний имеется прямая (со знаком +), сильная (+0,9) связь – с увеличением возраста женщин увеличивается частота хронических гинекологических заболеваний. Чем ближе размер коэффициента корреляции к +1 или к -1, тем больше, теснее измеряемая им прямая или обратная связь.

Таблица 2

Корреляция между возрастом женщин и выявленных у них хронических гинекологических заболеваний при профилактическом осмотре (на 1000 осмотренных)

№ обследования	Возраст женщин (x)	Кол-во хронич. гинекол. заб-ний (y)	dx	dy	d ² x	d ² y	dxdy
1	2	3	4	5	6	7	8
1	20	180	-10	-254	100	64516	2540
2	22	230	-8	-204	64	41616	1632
3	24	275	-6	-159	36	25281	954
4	26	340	-4	-94	16	8336	376
5	28	390	-2	-44,5	4	1936	88
6	30	420	0	-14	0	196	0
7	32	470	+2	+36	4	1296	72
8	34	490	+4	+56	16	3136	224
9	36	585	+6	+151	36	22801	906

1	2	3	4	5	6	7	8
10	38	660	+8	+226	64	51076	1808
11	40	735	+10	+301	100	90601	3010
n=11	$\Sigma x=330$	$\Sigma y=4775$	$\Sigma dx=0$	$\Sigma dy=0$	$\Sigma d^2x=440$	$\Sigma d^2y=310791$	$\Sigma dx dy=11610$
	$M_x=30,0$	$M_y=434,0$					

Степень вероятности коэффициента корреляции (r_{xy}) по способу квадратов определяется по таблице (приложение 4).

Техника вычисления коэффициента корреляции – r_{xy} :

1. Определить по каким признакам, в каком количественном размере (от этого будет зависеть выбор способа вычисления r_{xy}) необходимо провести корреляцию, определить число коррелируемых пар – n (в нашем примере графа первая).

2. Составить коррелируемые ряды: «x» и «y» (в нашем примере: x – возраст женщин; y – количество хронических гинекологических заболеваний на 1000 женщин этого возраста – графы 2,3).

3. Определить суммы (Σ) коррелируемых рядов «x» и «y» (в нашем примере $\Sigma x=330$, $\Sigma y=4775$).

4. Вычислить средние арифметические (M_x и M_y) каждого коррелируемого ряда путем деления суммы рядов на число коррелируемых пар ($M_x = \frac{\Sigma x}{n}$; $M_y = \frac{\Sigma y}{n}$). В нашем примере ($M_x=30$, $M_y=434$).

5. Определить разность (отклонения – dx и dy) между соответствующими значениями «x» и «y» и средними арифметическими этих рядов (M_x и M_y) – графы 4 и 5. В нашем примере отклонение для первой коррелируемой пары: $dx = x - M_x = 20 - 30 = -10$; $dy = y - M_y = 180 - 434 = -254$ и т.д.

6. При правильном составлении коррелируемых рядов суммы dx и dy равны 0.

7. Поочередно возвести в квадрат отклонения dx и dy и определить их суммы (в нашем примере $\Sigma dx^2 = 440$, $\Sigma dy^2 = 310791$).

8. Перемножая попарно dx и dy получим их произведения – dx·dy (графа 8).

9. Определить сумму произведений $\Sigma dx \cdot dy$. В нашем примере она равна 11610.

10. Вычислить коэффициент корреляции по способу квадратов по вышеуказанной формуле путем подставления числовых значений.

Из вышеизложенного следует, что коэффициент корреляции, вычисленный по способу квадратов, равен сумме попарных произведений отклонений от средних арифметических коррелируемых рядов, деленной на корень квадратный из произведения сумм квадратов этих отклонений.

Из приведенной схемы видно, что этот метод вычисления парного коэффициента корреляции довольно простой, особенно при малых числах наблюдения. При больших числах наблюдения трудности заключаются только в математических расчетах.

Вторым методом получения коэффициента корреляции является **метод рангов (Спирмена)**. Он применяется обычно в случаях, где $n < 30$ и если достаточно иметь лишь ориентировочные данные для оценки характера и силы связи. Коэффициент корреляции рангов обозначается греческой буквой ρ (ро), аналогичен коэффициенту парной корреляции (r) по значению, но с несколько меньшей точностью. Формула коэффициента корреляции:

$$\rho_{xy} = 1 - \frac{6 \times \Sigma d^2}{n(n^2 - 1)},$$

где x и y – признаки, между которыми определяется связь;

6 – постоянный коэффициент;

d – разность рангов;

n – число наблюдений.

Средняя ошибка (m_p) ранговой корреляции и критерия t определяются по формуле:

$$m_p = \sqrt{\frac{1 - \rho_{xy}^2}{n - 2}} \quad t = \frac{\rho_{xy}}{m_p}.$$

Для иллюстрации техники вычисления ρ приведем следующий пример.

Существует предположение, что с увеличением смертности от рака молочной железы, уменьшается смертность от рака матки. Требуется установить – существует ли зависимость (связь) между этими явлениями? На этот вопрос можно ответить путем вычисления ρ .

Техника вычисления ρ делится на 4 этапа, представленные в табл. 3.

I этап – составление корреляционных рядов смертности от рака молочной железы – «x» и смертности от рака матки – «y» в нескольких регионах республики (2 и 3 графы табл. 3).

Таблица 3

**Зависимость между показателями смертности
от рака молочной железы и рака матки**

Регионы республики	I этап		II этап		III этап	IV этап
	Смертность от рака молочной железы на 100000 женщин (x)	Смертность от рака матки на 100000 женщин (y)	Порядковые номера (ранги) по уровню показателей смертности от рака			
			молочной железы	матки	Разность рангов (d)	Квадрат разности рангов (d ²)
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>
Регион А	31,2	17,5	1	4	3	9
Регион Б	25,9	15,8	2	5	3	9
Регион В	23,2	18,4	3	2	1	1
Регион Г	8,2	17,7	4	3	1	1
Регион Е	5,9	21,7	5	1	4	16
		-	-		-	Σd ² =36

II этап – определение порядкового номера (ранга) региона в зависимости от уровня показателя смертности от рака, т.е. проводится ранжирование (графа 4 и 5). Обычно в графе «x» данные ряда проставляются в возрастающем порядке (как в нашем примере – от меньшего к большему) или в порядке уменьшения (от большего к меньшему); в графе «y» ранги определяются соответственно величине их показателей.

III этап – определение разности рангов – d (графа 6).

IV этап – возведение разности рангов в квадрат – d² и определение их суммы – Σd².

Проведя все этапы предварительных действий и расчетов (4 этапа), определяем коэффициент ранговой корреляции.

$$\rho = 1 - \frac{6 \times 36}{5(5^2 - 1)} = 1 - \frac{216}{5 \times 24} = 1 - 1,8 = -0,8.$$

Обратная величина коэффициента ранговой корреляции указывает на то, что с увеличением смертности от рака грудной железы уменьшается смертность от рака матки. Однако, прежде чем делать подобный вывод о степени связи между изучаемыми явлениями, необходимо оценить существенность (достоверность различий) полученного коэффициента (ρ). Степень вероятности коэффициента корреляции (ρ) находят по таблице Спирмана (приложение 5).

Если в коррелируемом ряду встречаются два и более одинаковых по величине числа, порядковый номер каждого из них следует обозначать средней из суммы очередных порядковых номеров.

Так, если бы в нашем примере в регионе Б, занимающим второе место по уровню показателя смертности от рака молочной железы и в регионе А – занимающим первое место, имели бы одинаковые уровни показателей, следовало бы обозначить каждый из них порядковым номером 1,5 (средняя из 1 и 2 равна (1+2):2=1,5).

Коэффициент детерминации. Применение математико-статистических методов обработки материалов научных исследований значительно увеличилось в связи с развитием социальной медицины и организации здравоохранения. С целью выявления изменения взаимосвязи и взаимодействия различных факторов в большинстве социально-медицинских работ используют относительные показатели и средние величины, измерения связи, корреляцию, факторный, дисперсионный и регрессионный анализы. Однако несмотря на многообразие статистических методов, основным ведущим методом (инструментом) выявления связи между изучаемым явлением и воздействующими факторами является парный коэффициент корреляции (Пирсона).

Почти во всех социально-медицинских и клинических исследованиях возникает необходимость установить наличие связи между изучаемым явлением и различными факторами, воздействующими на это явление. В то же время, оценить сложность воздействия, учесть все факторы, оказывающие влияние на явление-следствие невозможно, поэтому выбирают важнейшие из них и изучают их действие путем парной корреляции, параметры которой относительно четко показывают, как меняется зависимая переменная величина (изучаемое явление) при изменении независимой (действующего фактора).

Например, Г.И.Вихристюк, при изучении кратности заболеваний детей первого года жизни в зависимости от некоторых социальных и биологических факторов, доказала связь между уровнем показателя кратности заболеваний детей первого года жизни и материнским уходом, жилищно-бытовыми условиями, материальной обеспеченностью семьи путем вычисления парного коэффициента корреляции (r).

В корреляционном анализе для определения степени влияния рассматриваемого фактора и степени влияния других факторов, наряду с коэффициентом парной корреляции используются его производные коэффициенты. Одним из них является коэффициент детерминации, который показывает какая доля изменений в изучаемом явлении-следствии вызвана изменениями в исследуемом факторе-причине.

Коэффициент детерминации определяется путем возведения в квадрат коэффициента корреляции, умножением на 100 и выражается в процентах:

$$R=r^2 \times 100,$$

где R – коэффициент детерминации,
r – коэффициент корреляции.

Для примера обратимся к конкретным данным (табл. 4).

Таблица 4

Степень влияния изучаемых факторов на состояние здоровья детей 1-го года жизни. Коэффициент детерминации

Влияющий фактор	Коэффициент парной корреляции	Коэффициент детерминации $R=r^2 \times 100$	Коэффициент неопределенности в % (слияние др. факторов)
Материнский уход	0,36	$0,36^2 \times 100 = 13\%$	87
Обеспеченность семьи жилой площадью	0,51	$0,51^2 \times 100 = 26\%$	74
Комплексная оценка жилища	0,64	$0,64^2 \times 100 = 41\%$	59
Средний доход на душу	0,57	$0,57^2 \times 100 = 32\%$	68

Из таблицы видно, что материнский уход лишь в 13% является причиной изменений заболеваемости детей, остальные 87% изменений в заболеваемости детей вызваны другими факторами. Аналогично можно оценить влияние обеспеченности семьи жилой площадью – на этот фактор приходится 26% изменений в кратности заболеваемости детей, а 74% – на другие факторы и т.д. Коэффициент детерминации не дает возможности точно определить другие влияющие факторы в данной паре связи, но он определяет степень влияния конкретно рассматриваемого фактора. Коэффициент детерминации не дает возможности рассматривать все влияющие факторы в едином сочетании.

Самостоятельная работа студентов

Задача 1

Уровень холестерина по возрастным группам у обследованного мужского сельского населения Чуйского района

Возраст	20-29	30-39	40-49	50-59	60-69	70 и старше
Холестерин в мг %	185,1	195,0	238,0	217,5	204,4	233,7

Определить направление и силу связи между этими явлениями путем вычисления ранговой корреляции. Оценить достоверность и сделать выводы.

Задача 2

Зависимость между длительностью охлаждения организма (2 часа ежедневно) и уровнем молочной кислоты в крови (мг %)

Дни охлаждения	1	2	3	4	5	6	7	8	9	10
Молочная кислота в мг %	7,0	7,0	7,2	7,1	8,5	8,9	8,7	9,0	9,5	9,3

Определить направление и силу связи между двумя показателями путем вычисления коэффициента корреляции. Оценить достоверность и сделать выводы.

Задача 3

Суточная потребность в белках у 8-летних девочек

Вес девочек в кг	20	22	23	25	26	27	28
Суточная потребность белка в гр.	62	66	62	75	75	78	82

Определить коэффициент корреляции между весом девочек и суточной потребностью у них в белках. Оценить достоверность и сделать выводы.

Задача 4

Распространенность пневмокониоза среди шахтеров

Возраст	Выявлено на 100 рабочих
До 20 лет	0,0
21-30	0,3
31-40	1,3
41-50	8,8
50 и старше	5,0

Определить направление и силу связи между этими явлениями путем вычисления коэффициента ранговой корреляции. Оценить достоверность и сделать выводы.

Задача 5

Поглощение радиоактивного йода щитовидной железой крыс в разные сроки их пребывания в высокогорье

Дни пребывания	10	15	20	25	30	35
Радиоактивность в %	12,0	15,0	15,0	17,0	20,0	15,0

Определить коэффициент корреляции между днем пребывания в высокогорье и поглощением радиоактивного йода щитовидными железами. Оценить достоверность и сделать выводы.

Задача 6

Результаты изучения брачной плодовитости женщин в связи с доходом семьи

Группы по среднему месячному душевому доходу	Коэффициент брачной плодовитости (на 1000)
До 30	158
31-50	110
51-70	86
71-100	65
Свыше 100	35

Определить направление и силу связи между этими явлениями путем вычисления ранговой корреляции. Оценить достоверность и сделать выводы.

Тема № 10

КОЭФФИЦИЕНТ РЕГРЕССИИ

Цель темы:

Студент должен *знать*:

- Применение метода регрессионного анализа в практике врача.
- Принципы построения шкалы регрессии.

Студент должен *уметь*:

- Владеть методикой вычисления коэффициента регрессии
- Строить график регрессии.

План изучения темы

Разбор темы по контрольным вопросам

1. Понятие о регрессии.
2. Формула вычисления коэффициента регрессии, привести пример.
3. Формула вычисления «уравнения линейной регрессии» и его применение.
4. Формула вычисления сигмы регрессии.
5. Принцип построения графика или шкалы регрессии.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Рассмотренный нами коэффициент корреляции указывает лишь на направление и силу связи между двумя переменными величинами, но не дает возможности судить о том, как количественно меняются величины одного по мере изменения величин другого признака. Ответ на этот вопрос дает применение метода регрессии.

Регрессия – функция, позволяющая по величине одного коррелируемого (связанного) признака определить средние величины другого признака.

С помощью регрессии ставится задача выяснить, как количественно меняется одна величина при изменении другой величины на единицу. Для определения размера этого изменения применяется специальный коэффициент – коэффициент регрессии.

Коэффициент регрессии $R_{y/x}$ – абсолютная величина, на которую в среднем изменяется признак при изменении другого признака на единицу.

Формула коэффициента регрессии:

$$R_{y/x} = r_{xy} \times \frac{\sigma_y}{\sigma_x},$$

где $R_{y/x}$ – коэффициент регрессии;

r_{xy} – коэффициент корреляции;

σ_x и σ_y – средние квадратические отклонения для ряда «x» и «y».

Проследим вычисление коэффициента регрессии на примере.

Пример.

Определить массу тела по росту у 9-летних девочек. Обозначим через «y» их массу и через «x» – их рост. Известно, что сигма роста девочек этого возраста $\sigma_x = \pm 5,8$ см, сигма массы тела $\sigma_y = \pm 4,2$ кг, коэффициент корреляции между ростом и массой тела равен $r_{xy} = +0,6$. Коэффициент регрессии по росту равен:

$$R_{y/x} = +0,6 \times \frac{4,2}{5,8} = +0,43 \text{ кг/см.}$$

Вывод: при увеличении среднего роста 9-летних девочек на 1 см средняя масса их увеличивается на 0,43 кг.

С помощью коэффициента регрессии без специальных измерений можно определить величину одного из признаков (например, массы тела), зная значение другого (роста). Для этих целей служит **уравнение линейной регрессии**:

$$y = M_y + R_{y/x}(x - M_x),$$

где y – искомая величина массы тела,

x – известная величина роста;

$R_{y/x}$ – коэффициент регрессии массы тела по росту;

M_y – среднее значение массы тела, характерное для данного возраста;

M_x – среднее значение роста.

Пример.

Известны среднее значение массы тела девочек девятилетнего возраста $M_y=30,3$ кг и среднее значение роста $M_x=135,5$ см. Какова будет средняя масса тела у девятилетних девочек, имеющих рост 132 см? Подставим в уравнение регрессии известные нам данные и получим: $y=30,3+0,43 \times (132-135,5)=28,8$ кг.

Вывод: росту девятилетних девочек, равному 132 см, соответствует средняя масса тела 28,8 кг.

В практике исследования физического развития детей и подростков широко распространен метод оценки показателей роста, массы, окружности груди по шкале регрессии. Индивидуальные значения отдельных признаков очень разнообразны: так, у людей с одинаковым ростом показатели массы и окружности грудной клетки могут колебаться в самых широких пределах. Мера разнообразия индивидуальных размеров признаков характеризует сигма регрессии ($\sigma_{R_{y/x}}$), которая вычисляется по следующей формуле:

$$\sigma_{R_{y/x}} = \sigma_y \sqrt{1 - r_{xy}^2},$$

где σ_y – среднее квадратическое отклонение измеряемого признака (например, массы);

r_{xy} – коэффициент корреляции.

Чем меньше значение сигмы регрессии, тем более узкими будут пределы колебаний индивидуальных признаков относительно средней. В нашем примере:

$$\left[\sigma_y = \pm 4,2 \text{ кг}, \quad r_{xy} = +0,6, \quad \sigma_{R_{y/x}} = 4,2 \sqrt{1 - 0,6^2} = \pm 3,36 \text{ кг} \right].$$

Зная коэффициент регрессии, используя уравнение регрессии и сигму регрессии, можно построить шкалу регрессии.

Пример расчета шкалы регрессии.

Измерен рост (x) и вес (y) мальчиков 6 лет. $M_x=118,8$ см, $\sigma_x=\pm 5,8$ см, $M_y=24,1$ кг, $\sigma_y=\pm 2,6$ кг, $r_{xy}=\pm 0,7$.

1. Коэффициент регрессии массы тела по росту:

$$R_{y/x} = \pm 0,7 \times \frac{2,6}{5,8} = 0,3 \text{ кг/см.}$$

2. По росту находим массу тела с помощью уравнения регрессии

($y = R_{y/x} \times x + a$), где $a = M_y - R_{y/x} \times M_x$, $a = 24,1 - 0,3 \times 118,8 = -11,54$.

Нахождение массы тела по росту:

Средний рост, см	Средняя масса тела, кг
$x_1 = 100$	$y_1 = 0,3 \times 100 - 11,54 = 18,46$
$x_2 = 110$	$y_2 = 0,3 \times 110 - 11,54 = 21,46$
$x_3 = 120$	$y_3 = 0,3 \times 120 - 11,54 = 24,46$

3. Для определения пределов колебаний индивидуальной массы тела по среднему росту используют сигму регрессии:

$$\sigma_{R_{y/x}} = \sigma_y \sqrt{1 - r_{xy}^2} = 2,6 \times \sqrt{1 - 0,7^2} = \pm 1,3 \text{ кг.}$$

Таблица 1

Рост, x	Масса, $y \pm \sigma_{R_{y/x}}$, кг	Пределы колебаний
100	$18,46 \pm 1,3$ кг	$19,76 \pm 17,16$
110	$21,46 \pm 1,3$ кг	$22,76 \pm 20,16$
120	$24,46 \pm 1,3$ кг	$25,76 \pm 23,16$

На основании данных таблицы 1, можно построить график. Линия должна ограничиваться найденными точками (рис. 1).

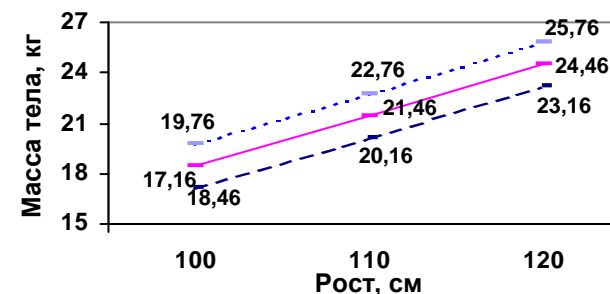


Рис. 1. Шкала регрессии массы тела по росту шестилетних мальчиков.

По этому графику можно определить как средние значения массы тела по росту, так и индивидуальные колебания массы тела по росту каждого ребенка не только по полученным точкам, но и в промежутках между ними. На этом принципе расчетов строятся стандарты физического развития, которые позволяют сопоставить рост каждого ребенка с массой его тела, окружностью груди, ростом сидя и т.д.

Применяя регрессию в исследовании, мы можем по величине какого-либо признака судить о средних размерах другого, взаимосвязанного с ним признака, при этом не производить каждый раз новых опытов и новых измерений. Например, по среднему росту можно судить о возможном уровне массы, по уровню максимального давления крови – о минимальном уровне, по количеству заболеваний – о среднем количестве бактерийносителей и т.д. На этом принципе строят номограмму – специальный график – сетку, на которой откладывают значения признака на одной линии, а на другой – соответствующие значения другого признака. Метод регрессионного анализа имеет большое значение в науке и практике.

Самостоятельная работа студентов

Задача 1

Результаты измерения уровней систолического и диастолического давления у 9 здоровых юношей 18 лет

Уровни систолического давления	110	110	115	120	120	115	125	130	120
Уровни диастолического давления	60	65	65	70	70	70	75	80	85

Вычислить коэффициент регрессии уровня систолического давления по диастолическому. Определить ожидаемый уровень систолического давления при диастолическом, равном 85, 90 и 85 мм ртутного столба. Начертить линию регрессии.

Задача 2

Результаты определения суточной потребности в белках у 8-летних девочек

Вес (кг)	20	22	23	25	26	27	28
Суточная потребность белка (в г)	62	66	62	75	75	78	82

Вычислить коэффициент регрессии суточной потребности в белках по весу девочек. Определить суточную потребность белка у девочек весом 21, 24 и 29 кг. Начертить линию регрессии.

Задача 3

Показатели заболеваемости

Возраст	10	15	20	25	30	35	40
На 1000 населения	0,15	0,71	3,56	20,42	40,11	64,20	69,51

Вычислить коэффициент регрессии заболеваемости по возрасту. Определить ожидаемую заболеваемость в возрасте 25 и 35 лет. Начертить линию регрессии.

Задача 4

Заболеваемость болезнями женских половых органов

Возраст	15	20	25	30	35	40
На 1000 женщин	0,87	3,75	18,33	25,04	18,71	9,62

Вычислить коэффициент регрессии заболеваемости по возрасту женщин. Определить ожидаемую заболеваемость в возрасте 20, 30 и 40 лет. Начертить линию регрессии.

Тема № 11

ДИНАМИЧЕСКИЕ РЯДЫ И ИХ АНАЛИЗ

Цель темы:

Студент должен *знать*:

- Роль динамических рядов в анализе научных исследований.
- Способы выравнивания динамических рядов.

Студент должен *уметь*:

- Вычислять и анализировать показатели динамических рядов.

План изучения темы

Разбор темы по контрольным вопросам

1. Определение динамического ряда.
2. Виды (типы) динамических рядов.
3. Определение моментного динамического ряда, его применение.
4. Определение интервального динамического ряда, его применение.
5. Способы выравнивания динамического ряда, их характеристика.
6. Анализ динамических рядов.
7. Показатели, характеризующие динамический ряд.
8. Абсолютный прирост, техника вычисления.
9. Темп прироста, техника вычисления.

10. Темп роста, техника вычисления.
 11. Значение 1% прироста, техника вычисления.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

Для того, чтобы анализировать динамику того или иного процесса, необходимо уметь сопоставлять динамические ряды разных типов, уметь их выравнять и анализировать.

Динамическим рядом называется совокупность однородных статистических величин, отражающих изменение какого-либо явления за определенный промежуток времени.

Числа динамического ряда принято называть уровнями ряда. Уровни ряда могут быть представлены абсолютными, относительными или средними величинами.

Виды (типы) динамических рядов

Ряды могут быть простыми (состоять из абсолютных величин) и сложными (состоять из относительных или средних величин). Простой динамический ряд может быть двух типов: моментный и интервальный.

1. **Моментный** – это такой динамический ряд, который состоит из величин, характеризующих размеры явления на определенные даты – моменты. Уровни моментного ряда не подлежат дроблению. Например, число врачей акушеров-гинекологов в Кыргызской Республике на конец соответствующего года (табл. 1).

Таблица 1

Число акушеров-гинекологов в Кыргызской Республике в 1992-2000 гг.

Год	1992	1993	1994	1995	1996	1997	1998	1999	2000
Число акушеров-гинекологов в Кыргызской Республике на конец года	883	884	853	845	848	803	1200	1264	1200

2. **Интервальный** ряд – ряд чисел, характеризующих какие-либо итоги за определенный интервал времени (сутки, неделя, декада, месяц, год).

Например, в акушерских стационарах Кыргызской Республики проведено родов за определенные годы, т.е. в таком динамическом ряду

каждое, отдельно взятое число указывает на численность родов, проведенных в акушерских стационарах в течение всего (целого) года (табл. 2).

Таблица 2

Число родов в акушерских стационарах Кыргызской Республики в 1992-2000 гг.

Год	1992	1993	1994	1995	1996	1997	1998	1999	2000
Число родов, проведенных в акушерских стационарах республики	104107	109116	115092	119056	122206	126637	102372	100102	96506

Интервальный ряд в отличие от моментного можно разделить на более дробные периоды, а также можно укрупнить интервалы.

Так, в нашем примере (табл. 2), вместо общего числа родов, проведенных в акушерских стационарах за год, можно было бы привести данные о поквартальном, помесечном числе родов. Из абсолютных величин интервального ряда можно получить средние значения за определенный дробный период времени. Например, разделив годовые величины родов на 12 (число месяцев в году), получим среднемесячное число родов для 1992 г. – $104107:12 = 8675$; для 1993 г. – $109116:12 = 9093$ и т.д. Эти же числа можно было бы разделить на 366 дней и получить среднедневное число родов для 1992 г. – $104107:366 = 284$, для 1993 г. – $109116:366 = 298$ и т.д.

Интервальные ряды могут состоять не только из чисел родившихся, но и из чисел умерших, заболевших и др., т.е. представляют данные, которые накапливаются за те или иные промежутки времени.

Выбор величины периода для интервального ряда (год, месяц, неделя, день, час и т.д.) в известной мере определяется степенью изменчивости явления (смертность, заболеваемость, рождаемость и т.д.). Чем медленнее изменяется явление во времени, тем крупнее могут быть периоды наблюдения.

Простые ряды (как моментные, так и интервальные) являются исходными для построения сложных рядов. Сложные же ряды состоят из средних величин (средняя длительность лечения, среднегодовое число коек и пр. за несколько лет) или из относительных величин (заболеваемость, смертность, рождаемость и пр. за несколько лет) (табл. 3).

Таблица 3

**Перинатальная смертность в республике
(на 1000 родившихся живыми и мертвыми)**

Год	1992	1993	1994	1995	1996	1997	1998	1999	2000
Перинатальная смертность	6,0	6,3	5,9	5,7	6,6	7,8	11,4	12,3	12,6

Выравнивание уровней динамических рядов. Динамический ряд не всегда состоит из уровней, последовательно изменяющихся в сторону снижения или увеличения. Нередко некоторые уровни в динамическом ряду представляют значительные колебания, что затрудняет возможность проследить основную закономерность, свойственную явлению в наблюдаемый период.

В этих случаях для выявления общей динамической тенденции рекомендуется произвести **выравнивание ряда**. Существует несколько способов выравнивания динамического ряда:

- Укрупнение интервала.
- Сглаживание ряда при помощи групповой средней.
- Сглаживание ряда при помощи скользящей средней и др.

Однако выравнивание уровней динамических рядов необходимо осуществлять только после глубокого и всестороннего анализа причин, обусловивших колебания этих уровней. Механическое выравнивание может искусственно сгладить уровни и завуалировать причинно-следственные связи.

1. **Укрупнение интервала** производят путем суммирования данных за ряд смежных периодов (табл. 4).

Как видно из табл. 4, месячные числа заболеваний ангиной то увеличиваются, то уменьшаются. После укрупнения интервалов по кварталам года можно увидеть определенную закономерность, наибольшее число заболеваний приходится на летне-осенний период.

Таблица 4

Сезонные колебания случаев ангины в городе N за 2000 г.

Месяц	I	II	III	IV	V	VI
Число заболеваний по месяцам по кварталам	129	193	133	387	230	288
	455			905		
VII	VIII	IX	X	XI	XII	Итого за 2000 г.
530	370	380	231	137	260	3268
1280			628			3268

2. **Вычисление групповой средней.** Для каждого укрупненного периода: суммируют смежные уровни соседних периодов, а затем полученную сумму делят на число слагаемых (табл. 5).

Таблица 5

**Динамика процента расхождений клинических
и патологоанатомических диагнозов
по данным областной больницы города N за 1993-2000 гг.**

Годы	1993	1994	1995	1996	1997	1998	1999	2000
Процент расхождения диагнозов	11,0	9,8	8,0	9,2	8,2	8,6	8,5	7,9
Групповая средняя	10,4		8,6		8,4		8,2	

Для уровней динамического ряда, представленных в табл. 5, характерны волнообразные колебания. Выравнивание ряда путем вычисления групповой средней позволило получить данные, иллюстрирующие довольно четкую тенденцию к постепенному снижению процента расхождений диагнозов в областной больнице.

3. **Вычисление скользящей средней** позволяет каждый уровень заменить на среднюю величину из данного уровня и двух соседних с ним (табл. 6).

Таблица 6

Методика расчета скользящей средней

Годы	1993	1994	1995	1996	1997	1998	1999	2000
Процент расхождения диагнозов	11,0	9,8	8,0	9,2	8,2	8,6	8,5	7,9
Скользящая средняя	–	9,6	9,0	8,7	8,6	8,4	8,3	–

Пример расчета: для 1994 г. $(11,0+9,8+8,0):3=9,6$, для 1995 г. $(9,8+8,0+9,2):3=9,0$ и т.д.

Ряд, выравненный при помощи скользящей средней, представляет последовательную тенденцию снижения процента расхождения диагнозов. Таким образом, скользящая средняя является простейшим способом выравнивания рядов. Этот метод дает возможность сгладить, устранить резкие колебания динамического ряда.

Анализ динамических рядов. Для изучения динамического ряда лучше всего сначала изобразить его графически. Графические изображения уровней динамического ряда позволяют в наглядной форме обнаружить последовательность изменения изучаемого явления.

Уровни показателей частоты расхождения клинических и патологоанатомических диагнозов в областной больнице ежегодно колебались (рис. 1). Ряд, выравненный с помощью скользящей средней, иллюстрирует общую четкую тенденцию: постепенное уменьшение показателей частоты расхождений диагнозов.

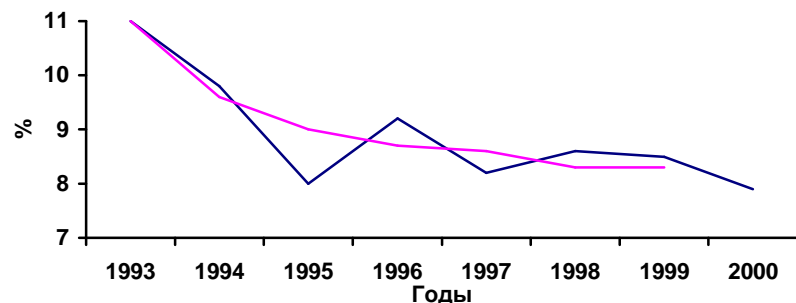


Рис. 1. Динамика частоты расхождений клинических и патологоанатомических диагнозов (в %) по данным областной больницы за 1993-2000 гг.

Для анализа динамических рядов используют следующие показатели:

- Абсолютный прирост (или убыль).
- Темп прироста (убыли).
- Темп роста.
- Абсолютное значение одного процента прироста (убыли) – табл. 7.

Таблица 7

Временная нетрудоспособность по поводу болезней периферической нервной системы рабочих завода за 1997-2000 гг. (число дней на 100 среднегодовых рабочих)

Показатели	1997 г.	1998 г.	1999 г.	2000 г.	Итого за 4 года
Число дней на 100 рабочих	39,8	44,6	55,5	59,7	-
Абсолютный прирост	-	+4,8	+10,9	+4,2	+19,9
Темп прироста	-	+12,1	+24,4	+7,5	+50,0
Темп роста	-	112,1	124,4	107,5	150,0

Техника вычисления показателей:

1. Абсолютный прирост – разность уровней данного года и предыдущего. Если динамический ряд имеет тенденцию к увеличению – прирост будет иметь положительное значение, и пишется со знаком плюс; если динамический ряд убывающий, прирост будет иметь отрицатель-

ное значение, и пишется со знаком минус. Например, для 1998 г.: $44,6 - 39,8 = +4,8$ и т.д.

2. Темп прироста – процентное отношение абсолютного прироста % к предыдущему уровню. Например, для 1998 г.:

$$x = \frac{4,8 \times 100}{39,8} = +12,1\%$$

3. Темп роста – процентное отношение последующего уровня к предыдущему уровню и указывает на сколько процентов увеличился или уменьшился последующий уровень. Например, для 1998 г.:

$$x = \frac{44,6 \times 100}{39,8} = 112,1\%$$

4. Если величина уровней различна, то, чтобы получить истинные данные об интенсивности роста (уменьшения) исследуемого явления, необходимо кроме темпа прироста вычислить значение 1%, т.е. какая величина прироста приходится на 1%. Для примера определим рост больничных коек в городах и сельской местности N-ой области (табл. 8).

Таблица 8

Интенсивность роста больничных коек в городах и сельской местности N-ой области в 1990-2000 гг.

Регионы	Рост коек		Абсолютный прирост	Темп прироста	Значение 1% (роста)
	1990	2000			
Города	850	1150	300	35,3 %	8,5
Село	15	245	95	63,3 %	1,5

Из таблицы видно, что темп прироста в селе выше (63,3%), чем в городе (35,3%). Однако если рассчитать прирост на 1%, то получим противоположные данные.

Техника вычисления:

- составление пропорции:

для города – на 35,3% – 300 коек, $x = \frac{300 \cdot 1,0}{35,3} = 8,5\%$

на 1% – x

для села – на 63,3% – 95 коек, $x = \frac{95 \cdot 1,0}{63,3} = 1,5\%$

на 1% – x,

т.е. интенсивность роста в городе почти в 6 раз выше, чем в селе.

Вывод: в тех случаях, когда исходные и конечные уровни сравниваемых явлений различны, нельзя судить о росте (уменьшении) явле-

ния только по темпу прироста, т.к. можно сделать ошибочный вывод. В подобной ситуации, наряду с темпом прироста, необходимо определить прирост на 1%.

Используя статистический метод для характеристики динамических рядов, следует всегда исходить из необходимости предварительного качественного анализа сущности изучаемого явления. Без этого не может быть осмыслена статистика динамических рядов.

Самостоятельная работа студентов

Задача 1

Численность населения (в тыс.)

Год	1991	1992	1993	1994	1995	1996
Численность населения	2532	2615	2696	2777	2859	2933

Вычислить и проанализировать показатели динамического ряда.

Задача 2

Численность многодетных матерей в г. N, получающих ежемесячные пособия (тыс.)

Год	1993	1994	1995	1996
Численность матерей, получающих пособие	35,0	74,3	115,0	132,0

Вычислить и проанализировать показатели динамического ряда.

Задача 3

Число больничных учреждений

Год	1992	1993	1994	1995	1996
Число учреждений	112	138	261	284	273

Вычислить и проанализировать показатели динамического ряда.

Задача 4

Обеспеченность населения средними медицинскими работниками (на 10000 человек населения)

Год	1992	1993	1994	1995	1996
Число средних медицинских работников	16,1	27,0	48,8	57,9	72,1

Вычислить и проанализировать показатели динамического ряда.

Задача 5

Число зарегистрированных больных с дизентерией

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	Всего
Число больных	20	23	30	31	29	42	68	100	93	71	35	32	574

Измерить сезонное колебание методом обычных средних.

Задача 6

Рождаемость среди городского населения (на 1000 человек населения)

Год	1991	1992	1993	1994	1995	1996
Рождаемость	33,1	31,2	29,3	31,3	24,9	24,6

Вычислить и проанализировать показатели динамического ряда.

Задача 7

Смертность населения от болезней системы кровообращения (на 1000 человек населения)

Год	1990	1991	1992	1993	1994	1995	1996
Смертность от болезней системы кровообращения	15,2	17,2	20,9	21,0	22,6	24,0	27,0

Вычислить и проанализировать показатели динамического ряда.

Задача 8

Число зарегистрированных случаев ревматизма

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	Всего
Число случаев	95	102	112	80	68	62	63	66	76	90	95	108	1017

Измерить сезонное колебание методом отношения среднедневного значения за каждый месяц к среднедневному их значению за год.

Задача 9

Распределение скоропостижно умерших от острого нарушения мозгового кровообращения при гипертонической болезни и атеросклерозе

Времена года	Зима	Весна	Лето	Осень	Всего
Количество случаев	595	504	344	546	1989

Измерить сезонное колебание методом обычных средних.

ТЕМА № 12

«ГРАФИЧЕСКИЕ ИЗОБРАЖЕНИЯ»

Цель темы:

Студент должен *знать*:

- Цель применения графических изображений в статистическом исследовании и общие правила их построения.

Студент должен *уметь*:

- Подбирать способ и строить графические изображения.

План изучения темы

Разбор темы по контрольным вопросам

1. Графические изображения – понятия, применения.
2. Статистические величины, используемые для построения графических изображений.
3. Основные виды графических изображений.
4. Виды диаграмм.
5. Линейные диаграммы, применение.
6. Столбиковые диаграммы, применение.
7. Радиальные диаграммы, применение.
8. Картограмма – понятие.
9. Картодиаграмма – понятие.
10. Общие правила к построению графических изображений.

Самостоятельная работа студентов

- Решение задач.

Содержание темы

При сводке материалов статистического исследования составляются таблицы различного вида и объема. В таблицы вносятся абсолютные и относительные величины. По данным таблиц проводится научный анализ материалов исследования.

Для большей наглядности, лучшего усвоения результатов исследования вместо статистических таблиц применяются графические изображения в виде диаграмм, картограмм, картодиаграмм. Графические изображения в ряде случаев помогают не только лучшему усвоению статистического материала, но и облегчают его научный анализ. Они могут быть построены по абсолютным и относительным величинам.

Различают следующие виды графических изображений:

1. Диаграммы: линейные, столбиковые, внутрискладовые, секторные, радиальные (на системе полярных координат), фигурные или объемные.

а) **Линейная диаграмма** (рис. 1) применяется для изображения явления во времени. Строится на прямоугольной системе координат. При этом, на оси абсцисс (горизонтальной линии) откладываются слева направо равные отрезки числовых значений одного ряда величин (временные интервалы в днях, месяцах, годах), на оси ординат (верти-

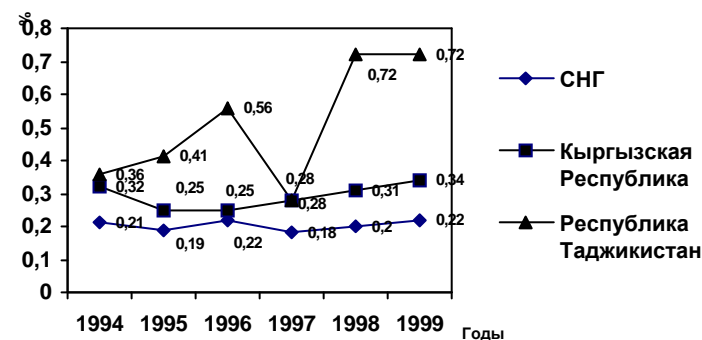


Рис. 1. Динамика коэффициента материнской смертности в странах СНГ, Кыргызской Республике и Республике Таджикистан за 1994-1999 гг. (на 1000 родов).

кальной линии) снизу вверх – значения другого ряда (изучаемого явления). Точка пересечения оси ординат и оси абсцисс дает нулевую точку пересечения обеих шкал.

Из точек, отложенных на оси абсцисс, снизу вверх проводятся линии, параллельные оси ординат, высота которых соответствует величине изучаемого явления на определенный временной промежуток, отложенных на оси ординат. Конечные точки параллельных линий соединяются ломаной линией слева направо, начиная от первой точки отложенной на оси ординат. Ярким примером линейной диаграммы является температурный лист, который служит повседневным ориентиром для врача о состоянии больного. Прежде чем чертить диаграмму, надо определить масштаб, т.е. расстояние на котором будут откладываться точки на оси ординат и оси абсцисс (см, граммах и т.д.). В нашем примере на оси ординат отложение точек ведется через 2 см, абсцисс – через 1,5 см. Произвольно определяются условные обозначения (цвет, штриховка и др.).

В примере №2 будет показано графическое изображение показателя материнской смертности за период 1958–1963 гг. (на 1000 родов).

б) **Столбиковые диаграммы** (рис. 2) обычно используются для изображения динамики или статике явления. Могут быть вертикальные и горизонтальные, строятся по тому же принципу, что и линейные. Для большей наглядности столбиковые диаграммы используются для

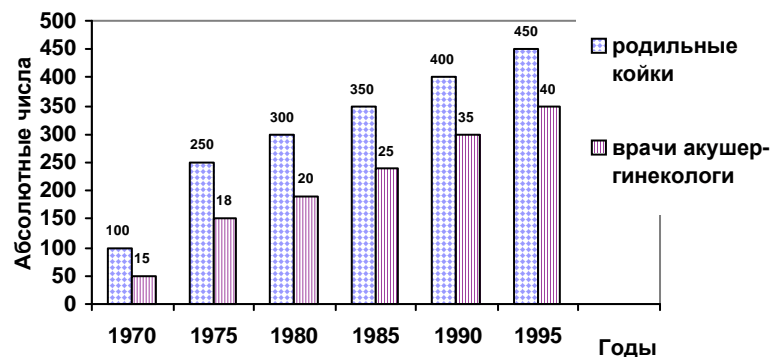


Рис. 2. Динамика роста родильных коек и врачей акушеров-гинекологов (1970-1995) в N области.

изображения одного или двух явлений. Например: рост численности акушерских коек – это диаграмма для одного явления. Но эту диаграмму можно скомбинировать с другим явлением – рост численности врачей акушеров-гинекологов.

Аналогичные столбиковые диаграммы можно чертить в горизонтальном положении.

в) Часто в практике применяются *секторные диаграммы* (рис. 3). Этот вид диаграмм применяется для изображения структуры явления за определенный промежуток времени или в динамике, или для сопоставления аналогичных явлений. Например, структуры заболеваемости, или структуры причин смерти населения.

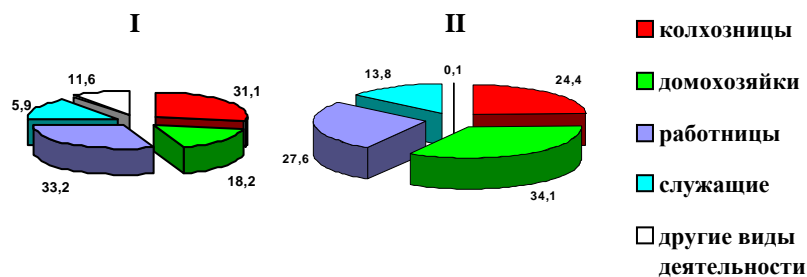


Рис. 3. Социально-общественная структура родивших (I) и умерших (II) женщин беременных, рожениц и родильниц в N области (1985-1998).

Секторные диаграммы представляют собой круг, отдельные секторы которого соответствуют отдельным долям изучаемого явления (удельному весу), они применяются для изображения распределения того или иного явления, т.е. для графического изображения совокупности экстенсивных коэффициентов.

Из диаграммы видно, что среди родивших женщин наибольший удельный вес составляют работницы – 33,2%, а среди умерших – домохозяйки – 34,1%.

Эти диаграммы можно построить в динамике за какое-то число лет только для родивших и только для умерших женщин. Однако при сравнении диаграмм в динамике нельзя делать выводы о распространенности явления, т.к. экстенсивный показатель в динамике дает представление об изменении долей: одна или две доли в целой совокупности уменьшаются или увеличиваются, что влечет за собой изменение остальных долей. Иногда пытаются интенсивные показатели в динамике, или при сравнении нескольких показателей, отражающих частоту аналогичных явлений изобразить секторной диаграммой. Это не верно. Секторная диаграмма предназначена для отражения только структуры явления, а не его интенсивности (частоты).

Структуру можно изобразить столбиковой диаграммой (рис. 4). Для иллюстрации приведем тот же пример.

В отличие от столбиковой диаграммы, отражающей изменение частоты явления в динамике, где каждый столбик отражает изменение всей совокупности в динамике, в столбиковых диаграммах структуры в одном столбике отражается вся структура одного явления (100%). Если изменяется при этом одна или две доли, то это влечет за собой изменение всех долей одной и той же совокупности. Расцветка долей в столбиковой диаграмме произвольная.

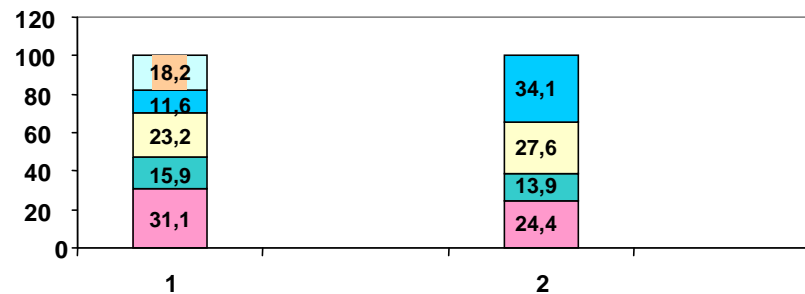


Рис. 4. Социально-общественная структура родивших и умерших беременных, рожениц и родильниц.

г) **Радиальные диаграммы** (рис. 5) – диаграммы, построенные на системе полярных координат. Эти диаграммы предназначены для изображения сезонных колебаний (подекадного, помесячного, поквартального и т.д.) заболеваемости, смертности, рождаемости и т.д.

При построении такой диаграммы круг делят на такое число секторов, на сколько частей делится год. Если год делится на 12 месяцев – круг делится на 12 секторов, на 4 квартала – на 4 сектора и т.д.

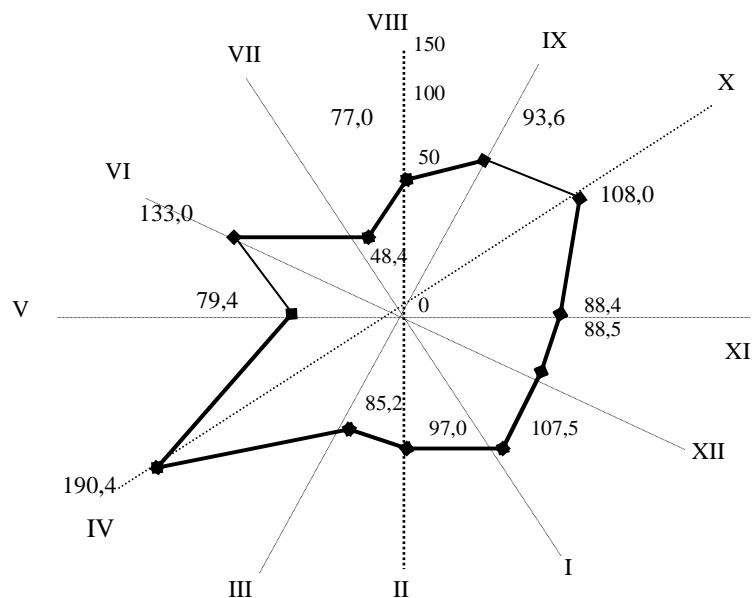


Рис. 5. Сезонные колебания заболеваемости эклампсией беременных, рожениц и родильниц (помесечная) за 1985-1998 гг.

Данные диаграммы указывают, что при измерении сезонных колебаний заболеваемости эклампсией (помесечно), установлено, что первая волна подъема этого показателя начинается в январе (индекс сезонного колебания равен 107,5), вторая волна и самая высокая в апреле (индекс=190,4), третья – в июне (индекс=133,0), четвертая в октябре (индекс=107,9). Самая низкая заболеваемость в июле, августе и мае (индексы=48,4, 77,0, 79,4).

2. **Картограммы** – это диаграммы, в которых изображено распределение какого-либо явления по территории. Например,

необходимо распределить районы области по рождаемости. Для этого определяются коэффициенты рождаемости каждого района, входящего в состав области. Готовится территориальная карта области, определяется раскраска (штриховка) соответственно коэффициента рождаемости и покрываются определенной раскраской (штриховкой) отдельные районы:

Например, если в районе N, входящего в состав изучаемой области коэффициент рождаемости равен 17,2‰, то на карте, на всю его территорию наносится соответствующая штриховка (окраска).

Картодиаграммы – это диаграммы, которые также рисуются на соответствующих территориальных картах (районных, областных, республиканских) для изображения распространения изучаемого явления на данной территории путем столбиковых или секторных диаграмм, а также различных фигур, которые вписываются в отдельные территориальные единицы при строгом соблюдении правил построения диаграмм (масштаб, форма, штриховка).

Общие правила построения графических изображений:

1. Каждая диаграмма, к какому бы типу она не относилась, должна иметь четкое, краткое, ясное описание, поясняющее изображение.
2. Шкалы диаграмм должны быть снабжены указателями размеров.
3. Числа показателей рекомендуется надписывать на самой диаграмме или в прилагаемой таблице.
4. Должны быть объяснены все условные обозначения.

Самостоятельная работа студентов

Задача 1

Структура затрат времени на прием одного больного врачами-терапевтами в одной из поликлиник

Элементы работы	Затрата времени в минутах (в % к итогу)
Подготовка и ознакомление с документацией	10.6
Опрос больного	15.1
Осмотр и обследование больного	25.9
Другие элементы работы	48.4
Всего	100.0

Составить графическое изображение, выбрав подходящий его вид.

Задача 2

Контингент больных с ревматическими заболеваниями (на 1000 человек населения)

Возраст	5-9	10-14	15-19	20-29	30-39	40-49	50-59	60 и >
Число больных	6,3	23,5	32,5	26,4	26,1	30,5	37,6	23,6

Составить графическое изображение, выбрав подходящий его вид.

Задача 3

Структура причин смерти населения по основным осложнениям и сопутствующим заболеваниям

Причины смерти	Основные заболевания	Осложнения	Сопутствующие заболевания
Несчастные случаи, отравления и травмы	34,6	1,4	10,9
Болезни системы кровообращения	31,4	68,7	32,7
Новообразования	13,3	1,6	16,5
Болезни органов дыхания	7,6	10,9	14,8
Болезни органов пищеварения	5,6	7,9	10,9
Другие классы болезней	7,5	9,5	14,2
Всего	100,0	100,0	100,0

Составить графическое изображение, выбрав подходящий его вид.

Задача 4

Гнойничковые заболевания пальцев правой кисти в некоторых отраслях промышленности

Пальцы	1 палец	2 палец	3 палец	4 палец	5 палец	Всего
В % к итогу	35,9	28,2	24,9	8,0	3,0	100,0

Составить графическое изображение, выбрав подходящий его вид.

Задача 5

Причины аборта

Причины	Безусловно устранимые	Условно устранимые	Неясные	Неустраняемые	Итого
В % к итогу	33,9	18,3	29,1	18,7	100,0

Составить графическое изображение, выбрав подходящий его вид.

Задача 6

Детская смертность в зависимости от месяца рождения (к среднегодовой принятой за 1000)

Месяцы	В городских поселениях	В сельских местностях
Январь	1201	1006
Февраль	892	737
Март	482	1099
Апрель	639	962
Май	590	910
Июнь	310	1087
Июль	1027	1233
Август	1437	936
Сентябрь	1178	970
Октябрь	1349	936
Ноябрь	1462	1149
Декабрь	1636	1094
За год	1000	1000

Составить графическое изображение, выбрав подходящий его вид.

Приложение 1

Расчет коэффициента К по размаху

n	0	1	2	3	4	5	6	7	8	9
0			1,13	1,69	2,06	2,33	2,53	2,70	2,85	2,97
10	3,08	3,17	3,26	3,34	3,41	3,47	3,53	3,59	3,64	3,69
20	3,73	3,78	3,82	3,86	3,90	3,93	3,96	4,00	4,03	4,06
30	4,09	4,11	4,14	4,16	4,19	4,21	4,24	4,26	4,28	4,30
40	4,32	4,34	4,36	4,38	4,40	4,42	4,43	4,45	4,47	4,48
50	4,50	4,51	4,53	4,54	4,56	4,57	4,59	4,60	4,61	4,63
60	4,64	4,65	4,66	4,68	4,69	4,70	4,71	4,72	4,73	4,74
70	4,75	4,77	4,78	4,79	4,80	4,81	4,82	4,83	4,83	4,84
80	4,85	4,86	4,87	4,88	4,89	4,90	4,91	4,91	4,92	4,93
90	4,94	4,95	4,96	4,96	4,97	4,98	4,99	4,99	5,00	5,01
n	100	200	300	400	500	600	700	800	900	1000
K	5,02	5,49	5,76	5,94	6,07	6,18	6,28	6,42	6,42	6,48

Приложение 2

Таблица значений критерия t (Стьюдента) и критерия вероятности безошибочного прогноза – p

Вероятность ошибки (P) Число степеней свободы n-1	0,05=95%	0,01=99%	0,001=99,9%
	1	12,7	63,6
2	3,1	5,8	12,9
3	2,7	4,6	8,6
4	2,5	4,0	6,8
5	2,4	3,7	5,9
6	2,3	3,5	5,4
7	2,3	3,3	5,1
8	2,2	3,2	4,7
9	2,2	3,1	4,6
10	2,2	3,1	4,4
11	2,2	3,0	4,3
12	2,1	3,0	4,2
13	2,1	2,9	4,1
14	2,1	2,9	4,0
15	2,1	2,9	4,0
16	2,1	2,9	4,0
17	2,1	2,8	3,9
18	2,1	2,8	3,9
19	2,0	2,8	3,8
20	2,0	2,8	3,8
21	2,0	2,8	3,8
22	2,0	2,8	3,7
23	2,0	2,8	3,7
24	2,0	2,7	3,7
25	2,0	2,7	3,7
26	2,0	2,7	3,7
27	2,0	2,7	3,6
28	2,0	2,7	3,6
29	2,0	2,7	3,6
30	2,0	2,7	3,6
∞	1,9	2,5	3,2

Приложение 3

Значение критерия соответствия χ^2
(хи-квадрат)

n-1	0,05	0,01	0,001
1	3,841	6,635	10,88
2	5,991	9,210	13,82
3	7,815	11,34	16,27
4	9,488	13,28	18,47
5	11,07	15,09	20,15
6	12,59	16,81	22,46
7	14,07	18,48	24,32
8	15,51	20,09	26,13
9	16,92	21,67	27,88
10	18,31	23,21	29,59
11	19,58	24,72	31,26
12	21,03	26,22	32,91
13	22,36	27,69	34,53
14	23,68	29,14	36,12
15	25,00	30,58	37,70
16	26,30	32,00	39,25
17	27,59	33,41	40,79
18	28,87	34,81	42,31
19	30,14	36,19	43,82
20	31,41	37,57	45,31
21	32,67	38,93	46,80
22	33,92	40,29	48,27
23	35,17	41,64	49,73
24	36,42	42,98	51,18
25	37,65	44,31	52,62
26	38,89	45,64	54,05
27	40,11	46,96	55,48
28	41,34	48,28	56,89
29	42,56	49,59	58,30
30	43,77	50,89	59,70

Приложение 4

Значение коэффициента корреляции(r_{xy}) по способу квадратов

Число степеней свободы n-2	Уровень вероятности P (в процентах)		
	95,0	99,0	99,9
1	0,99692	0,9988	0,999988
2	0,9500	0,9800	0,9990
3	0,878	0,9587	0,9911
4	0,811	0,9172	0,9741
5	0,754	0,875	0,9509
6	0,707	0,834	0,9249
7	0,666	0,798	0,898
8	0,632	0,765	0,872
9	0,602	0,735	0,847
10	0,676	0,708	0,823
11	0,553	0,684	0,801
12	0,532	0,661	0,780
13	0,514	0,641	0,760
14	0,497	0,623	0,742
15	0,482	0,606	0,725
16	0,468	0,590	0,708
17	0,456	0,575	0,693
18	0,444	0,561	0,679
19	0,433	0,549	0,665
20	0,423	0,537	0,652
25	0,381	0,487	0,597
30	0,349	0,449	0,554
35	0,325	0,418	0,519
40	0,304	0,393	0,490
45	0,288	0,372	0,465
50	0,273	0,354	0,443
60	0,250	0,325	0,408
70	0,232	0,302	0,408
80	0,217	0,283	0,357
90	0,205	0,267	0,338
100	0,195	0,254	0,321

Приложение 5

Значение коэффициента корреляции рангов (R) Спирмана (по В.Ю.Урбаху)

P n	P		P n	P	
	0,05	0,01		0,05	0,01
5	0,94	-	23	0,42	0,53
6	0,85	-	24	0,41	0,52
7	0,78	0,94	25	0,40	0,51
8	0,72	0,88	26	0,39	0,50
9	0,68	0,83	27	0,38	0,49
10	0,64	0,79	28	0,38	0,48
11	0,61	0,76	29	0,37	0,48
12	0,58	0,73	30	0,36	0,47
13	0,56	0,70	31	0,36	0,46
14	0,54	0,68	32	0,36	0,45
15	0,52	0,66	33	0,34	0,45
16	0,50	0,64	34	0,34	0,44
17	0,48	0,62	35	0,33	0,43
18	0,47	0,60	36	0,33	0,43
19	0,46	0,58	37	0,33	0,42
20	0,45	0,57	38	0,32	0,41
21	0,44	0,56	39	0,32	0,41
22	0,43	0,54	40	0,31	0,40

Коэффициент корреляции незначим при $p < 0,05$

Литература

1. Социальная гигиена и организация здравоохранения /Под ред. А.Ф.Серенко – 2-е изд. – М.: Медицина, 1984. – 640 с.
2. Руководство к практическим занятиям по социальной гигиене и организации здравоохранения /Под ред. Ю.П.Лисицына – 2-е изд. – М.: Медицина, 1984. – 400 с.
3. Социальная медицина и организация здравоохранения: Руководство для студентов, клинических ординаторов и аспирантов. – В 2 томах. – Т.1 /В.А.Миняев, Н.И.Вишняков и др. – СПб., 1998. – 219 с.
4. Практикум по медицинской статистике /Под ред. К.У.Акынбекова. – Бишкек, 1999. – 129 с.
5. Медицинская статистика /Под ред. К.Д.Абдуллина. – Бишкек, 1999. – 133 с.
6. Основы инфекционного контроля. Практическое руководство /Под ред. Е.А.Бурганской. – М., 1997. – 450 с.

К.У.Акынбеков, Н.Е.Чернова, К.Д.Абдуллин, А.А.Орозалиева

БИОСТАТИСТИКА

Редактор И.С.Волоскова
Технический редактор Э.К.Гаврина
Корректор О.А.Матвеева
Компьютерная верстка Е.Г.Шевёлкина

Подписано к печати 11.12.2001. Формат 60x84 ¹/₁₆.
Офсетная печать. Объем 7,0 п.л.
Тираж 300 экз. Заказ 243.

Издательство Кыргызско-Российского Славянского университета
720000, Бишкек, Киевская, 44

Отпечатано в типографии КРСУ
720000, Бишкек, Шопокова, 68

КЫРГЫЗСКО-РОССИЙСКИЙ СЛАВЯНСКИЙ УНИВЕРСИТЕТ

**КЫРГЫЗСКАЯ ГОСУДАРСТВЕННАЯ
МЕДИЦИНСКАЯ АКАДЕМИЯ**

**К.У.АКЫНБЕКОВ,
Н.Е.ЧЕРНОВА,
К.Д.АБДУЛЛИН,
А.А.ОРОЗАЛИЕВА**

БИОСТАТИСТИКА

Бишкек · 2001