Для освоения темы студенту необходимо:

1. Прочитать учебный материал
2. Общественное здоровье и здравоохранение : учебник / В.А. Медик, В.К. Юрьев. – М.: Гэотар-Медиа, 2020.
3. Методические материалы по теме занятия

**СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ СОЦИАЛЬНО-ГИГИЕНИЧЕСКИХ ИССЛЕДОВАНИЙ: КОРРЕЛЯЦИЯ, СТАНДАРТИЗАЦИЯ, АНАЛИЗ ДИНАМИКИ ЯВЛЕНИЯ**

**Общая цель учебного элемента:** обучение студентов статистическим методам обработки результатов медико-биологических исследований, методам оценки взаимодействия факторов (корреляция, стандартизация).

*В результате изучения темы студент будет:*

- способенприменять в дальнейшей научно-практической работе врача знания и умения по обработке результатов медико-биологических исследований, методам оценки взаимодействия факторов (корреляция, стандартизация);

- способен и готов применять методики медико-статистического анализа информации о показателях здоровья взрослого и детского населения на уровне различных подразделений медицинских организаций в целях разработки научно-обоснованных мер по улучшению и сохранению здоровья населения.

**Мотивация** В начале занятия преподаватель обращает внимание студентов, что актуальность изучения данной темы определена тем, что деятельность врача в условиях развития высоких технологий предусмат­ривает проведение научных исследований, которые не возможны без знаний методов обработки данных, позволяющих установить взаимосвязь между различными признаками, дать объективную характеристику состояния показателей в разных по структуре статистических совокупностях.

**Вопросы для обсуждения по теме занятия**

1. Дать определение понятия «стандартизация»

2. Перечислить этапы прямого метода стандартизации

3. Дать определение понятия «корреляции»

4. Привести примеры прямой и обратной корреляционной связи

5. Перечислить этапы расчета коэффициента корреляции по методу рангов

**Тестовые задания для самоконтроля**

*1. Метод корреляции используется для:*

1) определения числа необходимых наблюдений

более точных расчетов в выборочных исследованиях

2) определения взаимосвязи между признаками

3) для точных расчетов в генеральной совокупности

4) для сравнения двух неоднородных совокупностей

*2. При корреляционной связи:*

1) каждому значению одного признака соответствует строго определенное значение другого признака

2) каждому значению одного признака соответствует несколько значений другого признака

3) значения признаков не зависят друг от друга

4) все вышеперечисленное

*3. При функциональной связи:*

1) каждому значению одного признака соответствует строго определенное значение другого признака

2) каждому значению одного признака соответствует несколько значений другого признака

3) значения признаков не зависят друг от друга

4) все вышеперечисленное

*4. Метод стандартизации используется для:*

1) определения числа необходимых наблюдений

2) более точных расчетов в выборочных исследованиях

3) определения взаимосвязи между признаками

4) для точных расчетов в генеральной совокупности

5) для сравнения двух неоднородных совокупностей

*5. Для устранения различий в составе сравниваемых совокупностей используется:*

1) метод стандартизации

2) расчет относительных величин

*6.Стандартизованные показатели применяются:*

1) для характеристики первичного материала

2) для анализа полученных данных

3) для сравнения между собой

*7. Метод стандартизации применяется:*

1) для определения характера и силы связи между двумя явлениями (признаками)

2) при сравнении интенсивных показателей для устранения влияния на их величину определенных факторов, обуславливающих качественную неоднородность совокупностей

3) для определения достоверности различия двух сравниваемых показателей.

*8. К методам стандартизации относят:*

1) прямой

2) косвенный

3) обратный

***БЛОК-СХЕМА ТЕМЫ ЗАНЯТИЯ***

***СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ МЕДИКО-БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ***

**МЕТОД СТАНДАРТИЗАЦИИ**

**КОРРЕЛЯЦИОННЫЙ АНАЛИЗ**

*КОЭФФИЦИЕНТ РАНГОВОЙ КОРРЕЛЯЦИИ, ПОРЯДОК РАСЧЕТА*

*РАНЖИРОВАНИЕ ПОКАЗАТЕЛЕЙ*

*КОРРЕЛЯЦИОННАЯ СВЯЗЬ, ВИДЫ, ХАРАКТЕР,*

*СИЛА СВЯЗИ*

*ПРЯМОЙ МЕТОД*

*КОСВЕННЫЙ МЕТОД*

*ОБРАТНЫЙ МЕТОД*

 ***ПОШАГОВАЯ СИСТЕМА РАСЧЕТА КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ ПО МЕТОДУ РАНГОВ:***

1**. Обозначить рангами (порядковыми номерами) места значений показателей x и y в рядах.**

**2.Найти разность между рангами (d) по каждому значению показателя.**

**3.Возвести разность между рангами в квадрат (d2).**

**4. Найти по формуле коэффициент корреляции. Дать оценку характера и силы связи.**

**5. Определить ошибку и достоверность коэффициента корреляции.**

**ОСНОВНЫЕ ПОЛОЖЕНИЯ ТЕМЫ**

**МЕТОДИКА ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ**

**ПО МЕТОДУ РАНГОВ.**

 Различают два вида связи между признаками: функциональную и корреляционную.

**Функциональная связь** – изменение какого-либо одного явления вызывает строго определенные изменения другого (ph, химическая реакция).

**Корреляционная связь** – каждому определенному значению одного признака соответствует несколько значений другого взаимосвязанного с ним признака. Выявляется при массовом сопоставлении изучаемых явлений.

**По своему характеру может быть**:

-прямой,

-обратной.

**По силе:**

-полной,

-сильной,

-средней,

-слабой,

-отсутствовать

 Характер и силу связи определяют с помощью коэффициента корреляции.

Методика расчета коэффициента корреляции по методу рангов основана на использовании формулы:

$β\_{xy}$ = 1 -$\frac{6∑d^{2}}{n(n^{2}-1)}$ ,где

X и y - признаки , между которыми определяется связь;

6 – постоянный коэффициент;

d – разность рангов ;

n – число наблюдений.

**Порядок расчета:**

1. Обозначить рангами (порядковыми номерами) места показателей x и y в рядах.

2. Найти разность между рангами (d) по каждому значению показателя.

3. Возвести разность между рангами в квадрат (d2).

4. Найти по формуле коэффициент корреляции.

5. Дать оценку характера и силы связи (таблица).

6. Определить ошибку коэффициента корреляции по формуле:

$m\_{p}$$m\_{p}$ = ±$\sqrt{\frac{1-p^{2}y}{n-2}}$

7. Определяем достоверность коэффициента корреляции по формуле:

 t = $ \frac{ραy}{mρ}$$\frac{ρ}{m}$

**При выполнении ранжирования** показателей руководствуются следующими правилами:

1. Обозначение показателей рангами начинают с меньшего (или с большего) в обоих рядах (x и y).

2. Если отдельные значения показателя ряда встречаются несколько раз, то ранги проставляются следующим образом:

а) находим средний порядковый номер (как среднее арифметическое порядковых номеров, занимаемых одинаковыми значениями показателей ряда);

б) проставляем каждому одинаковому значению ранг, соответствующий его среднему порядковому номеру.

3. При оценке характера и силы связи необходимо ориентироваться на данные таблицы.

|  |  |  |
| --- | --- | --- |
| Характер и сила связи | Прямая (+) | Обратная (-) |
| полная | +1 | -1 |
| сильная | от +1 до +0.7 | от -1 до -0.7 |
| средняя | от +0.7 до +0.3 | от -0.7 до -0.3 |
| слабая | от +0.3 до + 0 | от -0.3 до 0 |
| отсутствует | 0 | 0 |

 При оценке достоверности коэффициента корреляции критерий должен быть равен или больше 3, что соответствует вероятности безошибочного прогноза

р ≥ 99%.

***Учебный пример:***Определить коэффициент корреляции между ростом и весом студентов и оценить его достоверность. Исходные данные представлены в таблице:

|  |  |  |
| --- | --- | --- |
| Порядковый номер студента | Рост, стоя, см (х) | Вес, кг (у) |
| 1 | 155 | 56 |
| 2 | 165 | 57 |
| 3 | 167 | 58 |
| 4 | 162 | 60 |
| 5 | 171 | 63 |
| 6 | 174 | 65 |
| 7 | 167 | 67 |
| 8 | 176 | 72 |
| 9 | 170 | 79 |
| 10 | 180 | 82 |

Решение. 1. Строим расчетную таблицу.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Порядковый номер студента | Рост стоя, см. (х) | Вес, кг. (у) | Рангих у | Разность рангов d | Квадрат разности рангов d 2 |
| 1 | 155 | 56 | 1 1 | 0 | 0 |
| 2 | 165 | 57 | 3 2 | 1 | 1 |
| 3 | 167 | 58 | 4,5 3 | 1,5 | 2,25 |
| 4 | 162 | 60 | 2 4 | 2 | 4 |
| 5 | 171 | 63 | 7 5 | 2 | 4 |
| 6 | 174 | 65 | 8 6 | 2 | 4 |
| 7 | 167 | 67 | 4,5 7 | 2,5 | 6,25 |
| 8 | 176 | 72 | 9 8 | 1 | 1 |
| 9 | 170 | 79 | 6 9 | 3 | 9 |
| 10 | 180 | 82 | 10 10 | 0 | 0 |

 ∑ d 2= 32

2. Делаем расчеты по формулам:

$β\_{xy}$ = 1 - $\frac{6∑d^{2}}{n(n^{2}-1)}$ = 1 - $\frac{6×32}{10(10^{2}-1)}$ = 1 - $\frac{192}{990}$ = 1 - 0.2 = 0.8

m =$\sqrt{\frac{1-p^{2}}{n-2}}$ =$\sqrt{\frac{1-0.64}{10-2}}$ = $\sqrt{\frac{0.36}{8}}$ = $\sqrt{0.45}$ = ±0.2

t = $\frac{ρ}{m}$ = $\frac{0.8}{0.2}$ = 4

**Ответ.** Наблюдается прямая сильная связь между ростом и весом студентов. Коэффициент корреляции достоверен (t > 3).

**МЕТОД СТАНДАРТИЗАЦИИ**

При изучении общественного здоровья и здравоохранения в научных или практических целях исследователю нередко приходится доказывать влияние факторных признаков на результативные при сравнении двух или более совокупностей. При сравнении двух неоднородных совокупностей по какому-либо признаку (составу) применяются методы стандартизации.

При сопоставлении показателей рождаемости, смертности, заболеваемости, инвалидности и т.п. следует учесть, что названные показатели необходимо вычислять на совокупностях, в каждой из которых все включенные возрастно-половые контингенты должны находиться в определенном пропорциональном отношении, что в реальных условиях не всегда получается.

**Стандартизация** – метод расчета условных показателей, позволяющий исключить количественную диспропорцию элементов (единиц наблюдения) сопоставляемых совокупностей.

Существует три метода расчета стандартизованных показателей:

**-прямой,**

**-косвенный**

**-обратный.**

**Прямой метод стандартизации** применяется при сравнении двух неоднородных совокупностей по какому-либо признаку (составу) (например, по возрасту, полу, профессиям и т.д.). Для вычисления стандартизованных показателей прямым методом исследователь должен иметь данные о составе населения и состав изучаемого явления.

**Сущность метода:** условно принимают какой-либо состав населения за стандарт и считают его одинаковым в сравниваемых совокупностях. Затем, учитывая действительные размеры явления по групповым показателям, вычисляют общие стандартизованные коэффициенты.

*Например:*

*-состав населения двух районов по возрасту и состав умерших по этим двум районам по возрасту*

*- состав больных по одинаковой тяжести заболевания в двух сравниваемых больницах*

**Стандартизованные показатели** **-** это условные, гипотетические величины, они не отражают истинных размеров явлений. Они свидетельствуют о том, каковы были бы значения сравниваемых интенсивных показателей, если бы были исключены различия в составах совокупностей.

**Этапы расчета стандартизованных показателей:**

**I этап.** Вычисление групповых показателей:

- общих - по совокупностям в целом;

- частных - по признаку различия (полу, возрасту, стажу работы и т.д.).

**II этап.** Определение стандарта. Стандартом в прямом методе называют состав населения, условно принимаемый одинаковым в сравниваемых совокупностях. За стандарт может быть принят:

1) средний состав или состав обеих групп, вместе взятых *(например: все выписавшиеся больные в двух больницах или все умершие в двух больницах;*

2) состав третьего объекта, известного по другим материалам или по предыдущим исследованиям *(например: при сравнении летальности в конкретной больнице по двум отделениям скорой помощи за стандарт может быть выбран состав больных любой другой больницы скорой помощи).*

**III этап.** Вычисление «ожидаемого» числа больных (или умерших) по стандарту.

**IV этап.** Вычисление стандартизованных показателей для сравниваемых совокупностей. Сопоставление соотношений стандартизованных и интенсивных показателей. Выводы.

*Пример:*При сравнении показателей летальности в 2-х больницах более высоким оказался показатель в больнице №1 по сравнению с больницей №2. Вычислить стандартизованные показатели летальности по возрасту. Оценить достоверность.

|  |  |  |
| --- | --- | --- |
| Возраст больных | Больница №1 | Больница №2 |
| Число выбывших больных | Число умерших больных | Число выбывших больных | Число умерших больных |
| 0-19 лет | 1000 | 66 | 500 | 40 |
| 20-29 лет | 700 | 21 | 300 | 12 |
| 30-39 лет | 500 | 10 | 1000 | 20 |
| 40 лет и старше | 300 | 3 | 700 | 8 |
| Всего | 2500 | **100** | 2500 | **80** |

 Решение: ***I этап. Вычисление повозрастных показателей летальности (в %)***

|  |  |  |
| --- | --- | --- |
| Возраст больных | В больнице 1 | В больнице 2 |
| 0-19 лет | = 6,6 | = 8,0 |
| 20-29 лет | = 3,0 | = 4,0 |
| 30-39 лет | = 2,0 | = 2,0 |
| 40 лет и старше | = 1,0 | = 1,0 |
| Общий показатель летальности | = 4,0 | = 3,2 |

Сравнивая общие показатели летальности можно сделать заключение, что летальность выше в первой больнице (4,0%) по сравнению со второй больницей (3,2%).

***II этап. Определение стандарта.*** В данном случае за стандарт принимается средний состав больных по возрасту в двух больницах.

|  |  |  |  |
| --- | --- | --- | --- |
| Возраст в годах | Число больных | В двух больницах | Распределение больных в стандарте (в %) |
| Больница №1 | Больница №2 |
| 0-19 лет | 1000 | 500 | 1500 | 1500х100% = 30 **5000** |
| 20-29 лет | 700 | 300 | 1000 | 1000х100% = 20 5000 |
| 30-39 | 500 | 1000 | 1500 | 1500х100% = 30 5000 |
| 40 лет и старше | 300 | 700 | 1000 | 1000х100% = 20 5000 |
| всего | 2500 | 2500 | **5000** | 100% |

Таким образом, условно принимаем, что состав больных по возрасту в двух больницах одинаков и соответствует распределению, принятому за стандарт.

***III этап. Вычисление «ожидаемого» числа умерших по стандарту.***

Каждая больница имеет свой фактический показатель летальности больных в зависимости от возраста. Рассмотрим, какое число умерших могло бы быть в стандартном числе больных.

Рассуждение ведется следующим образом: задается вопрос, какое число умерших было бы среди 30 больных в возрасте 0-19 лет с учетом того, что в 1 больнице летальность у пациентов этой группы составила 6,6 на 100 больных, а во второй 8,0 на 100 больных.

|  |  |  |  |
| --- | --- | --- | --- |
| Возраст | Показатели летальности | Распределение больных в стандарте | Вычисление ожидаемого числа умерших по стандарту |
| Больница №1 | Больница №2 | Больница №1 | Больница №2 |
| 0-19 лет | 6,6 | 8,0 | 30,0 | **6,6** х **30** = 1,9 100  | **8,0** х **30** = 2,4 100 |
| 20-29 лет | 3,0 | 4,0 | 20,0 | 3,0 х 20 = 0,6 100  | 4,0 х 20 = 0,8 100 |
| 30-39 лет | 2,0 | 2,0 | 30,0 | 2,0 х 30 = 0,6 100  | 2,0 х 30 = 0,6 100 |
| 40 лет и старше | 1,0 | 1,0 | 20,0 | 1,0 х 20 = 0,2 100  | 1,0 х 20 = 0,2 100 |
| Всего | 12,6 | 14,0 | 100 | **3,3** | **4,0** |

Ответ по данному этапу: на 30 больных в возрасте 0-19 лет в 1 больнице возможна летальность 1,9, во второй 2,4. Аналогично рассчитываем по всем возрастам.

***IV этап. Вычисление стандартизованных показателей для сравниваемых совокупностей.***

Суммирование «ожидаемых» чисел умерших по возрасту в одной и другой больнице. Это и есть стандартизованные показатели (Р1 и Р2).

В больнице №1 Р1 =1,9+0,6+0,6+0,2 = **3,3%** (на 100 рабочих стандартов).

В больнице №2 Р2 =2,4+0,8+0,6+0,2 = **4,0%** (на 100 рабочих стандартов).

*Определение достоверности разности стандартизованных показателей.*

**Критерий достоверности Стьюдента**

$t=\frac{P\_{1}-P\_{2}}{\sqrt{m\_{1}^{2}+m\_{2}^{2}}}\geq 2$

**Ошибка:** $m=\pm \sqrt{\frac{Р×q}{n}}$**,** где q=100-Р, т.е. $m=\pm \sqrt{\frac{Р×(100-Р)}{n}}$     n1=2500; n2=2500

**Вывод:** При одинаковом составе больных по возрасту в обеих больницах показатели летальности несколько выше в больнице №2. Разница в показателях летальности больниц №1 и №2 статистически недостоверна (t=1,4)

**ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ**

**I. ЗАДАНИЕ ПО РАЗДЕЛУ «МЕТОДИКА ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ»:**

***Для студентов лечебного факультета*:**

Определить взаимосвязь между признаками по коэффициенту корреляции рангов и сделать выводы

|  |  |  |
| --- | --- | --- |
| № п\п | Возраст | Число больных сердечной недостаточностью |
| 1 | До 20 лет | 12 |
| 2 | 21-30 | 24 |
| 3 | 31-40 | 30 |
| 4 | 41-50 | 23 |
| 5 | 51 и старше | 32 |

***Для студентов педиатрического факультета:***

Определить взаимосвязь между признаками по коэффициенту корреляции рангов и сделать выводы

|  |  |  |
| --- | --- | --- |
| № п\п | Средний возраст первородящих женщин | Средний вес ребенка при рождении, в гр. |
| 1 | 18,2 | 2800 |
| 2 | 20,1 | 3700 |
| 3 | 24,3 | 3600 |
| 4 | 33,5 | 3820 |
| 5 | 41,0 | 3700 |

***Для студентов стоматологического факультета:***

Определить взаимосвязь между признаками по коэффициенту корреляции рангов и сделать выводы

|  |  |  |
| --- | --- | --- |
| № п\п | Среднее число обращений пациентов к стоматологу в год | Число запущенных случаев кариеса  |
| 1 | 0,5 | 32 |
| 2 | 1,0 | 14 |
| 3 | 2,0 | 4 |
| 4 | 1,5 | 6 |
| 5 | 2,5 | 1 |

**II. ЗАДАНИЕ ПО РАЗДЕЛУ «СТАНДАРТИЗАЦИЯ»**

***Для студентов лечебного факультета:*** Определить: отличаются ли показатели плодовитости женщин в двух районах (А. и Б.), если известно, что возрастной состав женщин, проживающих в названных районах, различен. Вычислить стандартизованные показатели. Оценить достоверность разности их.



***Для студентов педиатрического факультета:*** Определить**:** отличаются ли показатели летальности детей в двух больницах в зависимости от возраста детей (в абс. числах). Вычислить стандартизованные показатели. Оценить достоверность разности их.

|  |  |  |
| --- | --- | --- |
| Возраст | Больница №1 | Больница №2 |
| Число больных | Число умерших | Число больных | Число умерших |
| 0-1 | 1000 | 66 | 500 | 40 |
| 2-3 | 700 | 21 | 300 | 12 |
| 4-7 | 500 | 10 | 1000 | 30 |
| Старше 7 | 300 | 3 | 700 | 8 |
| Всего | 2500 | 100 | 2500 | 90 |

***Для студентов стоматологического факультета:***

В таблице представлено распределение школьников, страдающих кариесом зубов, по классам школ А и Б города N. (в абс. числах) Вычислить стандартизованные показатели. Оценить достоверность разности их.

|  |  |  |
| --- | --- | --- |
| Класс | Школа А | Школа Б |
| Число детей | Из них с кариесом | Число детей | Из них с кариесом |
| 1-3 | 400 | 180 | 200 | 140 |
| 4-9 | 400 | 300 | 600 | 500 |
| 10-11 | 200 | 100 | 200 | 100 |
| Всего | 1000 | 580 | 1000 | 740 |

**Тестирование по итогам занятия.**

*1.Величина стандартизованных показателей истинному размеру явления:*

1) соответствует

2) не соответствует

*2. Основное условие применения метода стандартизации:*

1) достаточное число наблюдений

2) наличие данных о структуре среды и явления

3) резкие отличия состава изучаемых совокупностей

*3. Стандартизованные показатели применяются:*

1) для характеристики первичного материала

2) для анализа полученных данных

3) для сравнения между собой

*4. При корреляционной связи:*

1) изменение какого-либо одного явления вызывает строго определенные изменения другого.

2) изменение какого-либо одного явления иногда вызывает строго определенные изменения другого

3) изменение какого-либо одного явления вызывает обратные изменения другого

4) каждому определенному значению одного признака соответствует несколько значений другого взаимосвязанного с ним признака

*5. Поучить представление о силе и направлении связи между признаками можно с помощью:*

1) таблиц;

2) графического изображения зависимости;

3) коэффициента корреляции.

*6. Направление корреляционной зависимости может быть представлено с помощью:*

1) таблиц;

2) диаграмм рассеяния;

3) коэффициента корреляции.

*7. Только наличие и направление связи между коррелируемыми признаками можно представить с помощью:*

1) коэффициента корреляции;

2) диаграмм рассеяния;

3) таблиц.