Казахский национальный университет имени аль-Фараби
Факультет медицины и здравоохранения
Кафедра политики и организации здравоохранения

Тема: «Понятие о корреляционной зависимости. Коэффициент корреляции Пирсона для количественных данных, имеющих нормальное распределение (параметрический). Коэффициент корреляции Спирмена для порядковых данных или количественных данных, не удовлетворяющих условиям нормальности (непараметрический). Интерпретация силы (величины) и направления (знака)»

Лектор: к.м.н., доцент Калмаханов С.Б.

Кафедра политики и организации здравоохранения

Понятие корреляционной зависимости

Корреляция – это статистическая взаимосвязь между двумя количественными признаками. Она показывает, насколько изменение одного показателя связано с изменением другого, но не доказывает причинно-следственную связь.

Примеры в здравоохранении:

- связь между уровнем дохода населения и частотой обращений за платными медицинскими услугами;
- зависимость между индексом массы тела (ИМТ) и уровнем артериального давления;
- взаимосвязь между удовлетворенностью пациентов и эффективностью менеджмента клиники.

Типы корреляции

Тип зависимости	Характеристика	Пример
Положительная (прямая)	При увеличении одного признака увеличивается и другой	Чем выше доход, тем выше удовлетворённость медицинским обслуживанием
Отрицательная (обратная)	При увеличении одного признака другой уменьшается	Чем больше стаж, тем меньше количество ошибок при ведении документации
Отсутствие корреляции	Изменение одного признака не связано с изменением другого	Рост пациента и уровень холестерина

Коэффициент корреляции Пирсона (r)

- **Когда используется:**
  - оба признака – количественные (интервальные или шкала отношений);
  - данные имеют нормальное распределение;
  - связь между переменными линейная.
- **Формула (для понимания принципа):**

$$r = \frac{cov(X,Y)}{s_X s_Y}$$

где

cov(X, Y) – ковариация между переменными X и Y,

s_X, *s_Y* – стандартные отклонения X и Y.

Значение *r* всегда лежит в диапазоне от **–1 до +1.**

Интерпретация коэффициента корреляции Пирсона (r)

Значение r	Сила связи	Интерпретация	Пример из менеджмента здравоохранения
0.00-0.19	Очень слабая	Линейная связь практически отсутствует	Между возрастом руководителя и средним временем ожидания пациентов нет заметной связи
0.20-0.39	Слабая	Незначительная линейная зависимость	Слабая связь между числом коек и уровнем удовлетворенности пациентов
0.40-0.59	Средняя	Умеренная линейная зависимость	Умеренная зависимость между расходами на обучение персонала и производительностью труда
0.60-0.79	Сильная	Явная линейная связь	Сильная положительная зависимость между числом врачей и количеством обслуженных пациентов
0.80-1.00	Очень сильная	Почти прямая (линейная) зависимость	Очень сильная связь между количеством койко-дней и общими затратами на лечение

Коэффициент корреляции Спирмена (ρ)

- **Когда используется:**

- данные порядковые (например, ранги удовлетворенности);
- или количественные данные, не распределенные нормально;
- или когда связь нелинейная, но монотонная (т.е. при увеличении одного признака другой либо растет, либо уменьшается).

- **Принцип расчёта:**

Вместо исходных значений используются **ранги** (места по порядку).
Формула:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

где
d_i – разность между рангами каждой пары наблюдений,
n – количество пар наблюдений.

Значения ρ также лежат от **–1 до +1**.

Интерпретация знака и силы корреляции Спирмена (ρ)

Значение ρ	Сила связи	Интерпретация	Пример из менеджмента здравоохранения
0.00-0.19	Очень слабая	Монотонная зависимость практически отсутствует	Ранг удовлетворенности пациентов почти не зависит от уровня оснащённости клиники
0.20-0.39	Слабая	Слабая монотонная зависимость	Незначительная связь между уровнем стресса персонала и оценками командного взаимодействия
0.40-0.59	Средняя	Умеренная монотонная зависимость	Умеренная зависимость между рангами качества управления и уровнем удержания сотрудников
0.60-0.79	Сильная	Значимая монотонная зависимость	Сильная отрицательная зависимость: чем выше ранг удовлетворенности пациентов, тем ниже частота жалоб
0.80-1.00	Очень сильная	Почти идеальная монотонная зависимость	Очень сильная положительная связь между рангом эффективности отделений и рейтингом клиники

Сравнение

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ПИРСОНА (R)

Преимущества: Более точный результат, если условия нормальности выполнены.

Недостатки: Высокая чувствительность к выбросам в данных, которые могут исказить результат.

Применение: Изучение пропорциональной изменчивости двух переменных, например, рост и вес.

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ СПИРМЕНА (P)

Преимущества: Более универсален и менее строг, чем Пирсон; устойчив к выбросам.

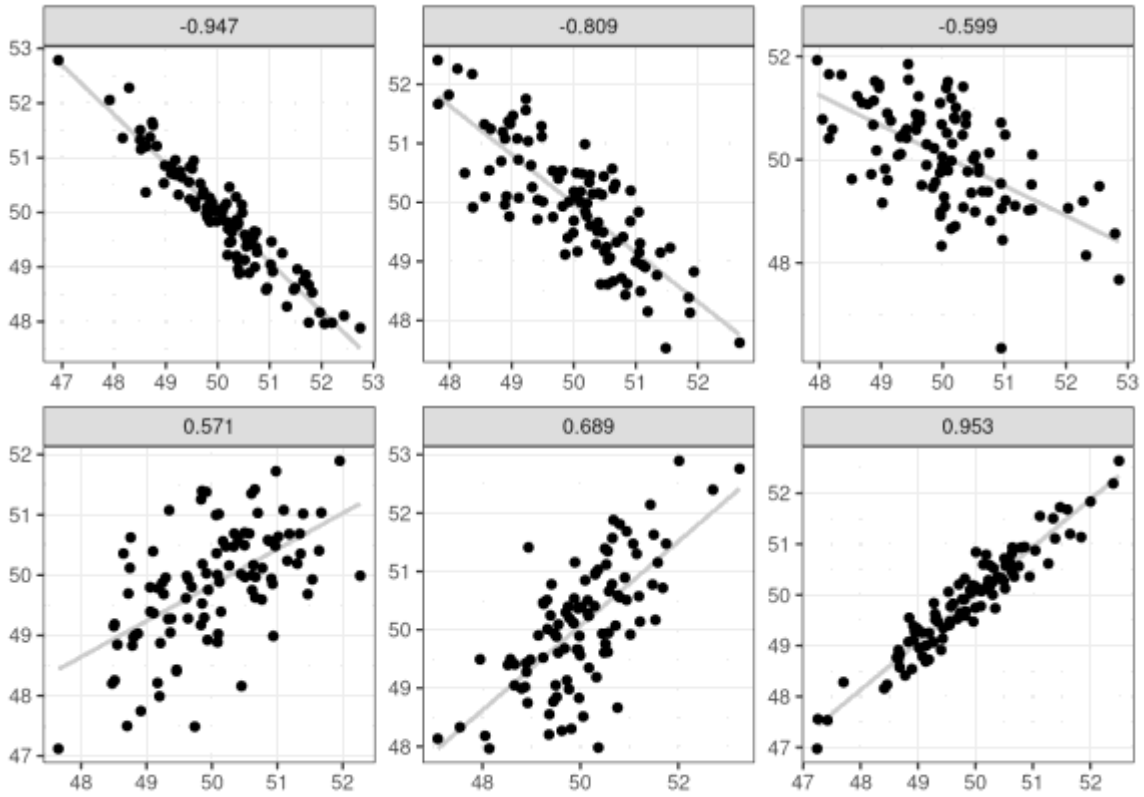
Недостатки: Менее точный, если данные соответствуют всем условиям для Пирсона, но все же часто используется из-за своей универсальности.

Применение: Оценка связи между двумя рейтингами, например, оценка двух разных экспертов для одной и той же группы людей.

Направление корреляции (знак)

Знак	Тип связи	Интерпретация	Пример
«+»	Прямая	При увеличении одного признака увеличивается другой	Больше врачей – больше пациентов
«–»	Обратная	При увеличении одного признака другой уменьшается	Выше нагрузка – выше удовлетворенность
0	Отсутствует	Изменения признаков не связаны	Число коек не связано с качеством менеджмента

Распределения с разными значениями корреляции



Как видно из этого графика, чем ближе модуль корреляции к 1, тем более компактно расположены точки друг к другу, чем ближе к 0, тем более рассеяны значения.

Важно: высокая корреляция не означает, что одно явление вызывает другое!

Например, рост числа пациентов и рост зарплаты персонала могут коррелировать, но причина – ***общий рост финансирования системы здравоохранения.***

Пример интерпретации в контексте исследования

Задача: определить связь между уровнем удовлетворенности и количеством жалоб.

r = –0.68 → ***сильная отрицательная корреляция***

Чем выше удовлетворенность, тем меньше жалоб.

Задача: определить связь между числом врачей и количеством пациентов, обслуженных в день.

r = +0.74 → ***сильная положительная корреляция***

Рост числа обслуженных пациентов связан увеличением количества врачей.

Пошаговый подход для анализа в практике менеджера здравоохранения

1. **Определить тип данных** (количественные или порядковые).
2. **Проверить нормальность распределения** (например, тестом Шапиро-Уилка).
3. **Выбрать подходящий метод:**
   - Нормальные данные → коэффициент Пирсона (r).
   - Ненормальные/порядковые данные → коэффициент Спирмена (ρ).
4. **Рассчитать коэффициент** (в Excel, SPSS).
5. **Интерпретировать** величину и знак.
6. **Сделать управленческие выводы**, например:
   - выявить факторы, влияющие на качество медицинских услуг;
   - определить взаимосвязь между удовлетворенностью пациентов и нагрузкой на персонал;
   - обосновать приоритеты при распределении ресурсов.

Заключение

Корреляционный анализ – мощный инструмент для управленческих решений в здравоохранении.

Он помогает определять взаимосвязи, оценивать тенденции и обосновывать стратегические решения на основе данных.

Главное – правильно выбрать метод, интерпретировать результаты и помнить: **корреляция ≠ причинность**.

Ситуационное задание 1

По данным 20 поликлиник региона выявлено: r = +0.67 между финансированием и охватом профилактическими осмотрами.

- Как можно интерпретировать этот результат?
- Можно ли утверждать, что рост финансирования вызывает повышение охвата?
- Какие иные факторы могут влиять на показатель?

Ситуационное задание 2

В клинике выявлена отрицательная корреляция ($\rho = -0.56$) между уровнем стресса медперсонала и удовлетворенностью пациентов.

- Какие управленческие решения можно предложить на основе этого результата?
- Как можно подтвердить, что именно стресс влияет на качество обслуживания?

Спасибо за внимание!