



Прикладная статистика и анализ данных

Съезд XI



Анализ причинности

Парадокс Симпсона

	Лекарство	Плацебо
Мужчины	81 из 87 = 93%	234 из 270 = 87%
Женщины	192 из 263 = 73%	55 из 80 = 69%
Все вместе	273 из 350 = 78%	289 из 350 = 83%

у 234-х из них
есть эффект

270 мужчин
принимали
плацебо

Выводы:

- На женщин лекарство влияет положительно.
- На мужчин лекарство влияет положительно.
- В целом по всем лекарство влияет отрицательно.

Следствие:

- Если знаем пол пациента, то нужно назначить лекарство.
- Если пол не знаем, то не нужно.



Парадокс Симпсона

	Лекарство	Плацебо
Мужчины	81 из 87 = 93%	234 из 270 = 87%
Женщины	192 из 263 = 73%	55 из 80 = 69%
Все вместе	273 из 350 = 78%	289 из 350 = 83%

у 234-х из них
есть эффект

270 мужчин
принимали
плацебо

Разбираемся в причинно-следственной связи:

- Женщины в целом реже выздоравливают чем мужчины.
- Случайный человек, принимающий лекарство, скорее всего женщина.
- Случайный человек, не принимающий лекарство, скорее всего мужчина.

В этом и заключается
причина эффекта
в сгруппир. данных.

Итог: нужно рассматривать сгруппированные данные, т.е. лекарство полезно.

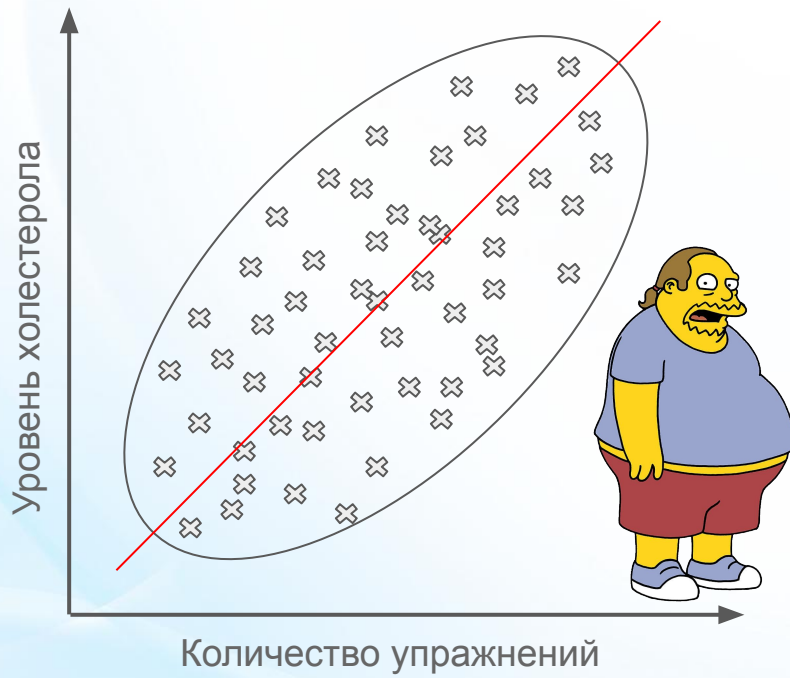


Что же делать на практике?

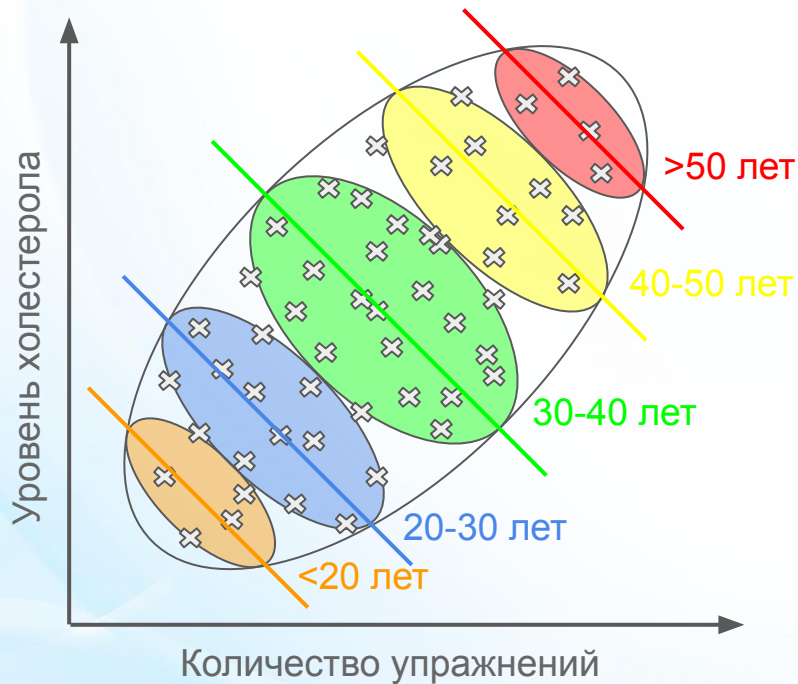
- Рассматривать выводы по группам.
- Требовать тестовую группу одного размера.
- Использовать взвешенное усреднение.

Статистика может доказать
что угодно, даже правду.

Другой пример



Другой пример



В каждой группе:
чем больше упражнений,
тем меньше холестерина.

В целом по всем:
чем больше упражнений,
тем больше холестерина.

Причинно-следственные связи:

- молодые делают меньше упражнений;
- но и в целом имеют меньший уровень холестерина.

**Итог: нужно рассматривать сгруппированные данные,
т.е. упражнения полезны для уменьшения холестерина.**

Всегда ли так?

Пациенты принимают лекарство.
Спустя час измеряется кровяное давление

	Плацебо	Лекарство
Низкое давление	81 из 87 = 93%	234 из 270 = 87%
Высокое давление	192 из 263 = 73%	55 из 80 = 69%
Все вместе	273 из 350 = 78%	289 из 350 = 83%

270 мужчин
принимали
лекарство

у 234-х из них
есть эффект

Причинно-следственные связи:

- Тестовая и контрольная группы равны.
- У принимающих лекарство оказалось в основном низкое давление.
- У принимающих плацебо — в основном высокое.

Вывод: лекарство влияет на понижение давления.

**Итог: нельзя рассматривать сгруппированные данные;
лекарство эффективно.**