

ВЕЛИЧИНА КОРРЕЛЯЦИИ И СИЛА СВЯЗИ

ВЕЛИЧИНА КОРРЕЛЯЦИИ И СИЛА СВЯЗИ

План:

- **1. ЗНАЧЕНИЕ КОРРЕЛЯЦИИ**
- **2. КАКОЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ВЫБРАТЬ**
- **3. ОБРАБОТКА НА КОМПЬЮТЕРЕ**

ЗНАЧЕНИЕ КОРРЕЛЯЦИИ

Коэффициенты корреляции были специально разработаны для численного определения силы и направления связи между двумя свойствами, измеренными в числовых шкалах (метрических или ранговых).

Максимальной силе связи соответствуют значения корреляции $+1$ (прямо пропорциональная связь) и -1 (обратно пропорциональная связь), отсутствию связи соответствует корреляция, равная нулю.

В ряде случаев разные коэффициенты корреляции имеют различную эффективность, а иногда все они оказываются нечувствительными к связям.

ЗНАЧЕНИЕ КОРРЕЛЯЦИИ

Выбросы и отклонения распределений от нормальности

Выбросы — это экстремально большие или малые значения признака. В наиболее существенной степени выбросы влияют на корреляцию r -Пирсона

Влияние «третьей» переменной

Иногда корреляция между двумя переменными обусловлена не связью между соответствующими свойствами, а влиянием некоторой общей причины совместной изменчивости этих переменных, которая зачастую выпадает из поля зрения исследователя.

Нелинейные связи

Еще одним источником низкой эффективности корреляций являются возможный нелинейный характер связи между переменными. Важен визуальный анализ связи с помощью таких графиков во всех случаях применения корреляций.

КАКОЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ВЫБРАТЬ

При изучении связей между переменными наиболее предпочтительным является случай применения *r-Пирсона* непосредственно к исходным данным.

Применяя *r-Пирсона*, необходимо убедиться, что:

- обе переменные не имеют выраженной асимметрии;
- отсутствуют выбросы;
- связь между переменными прямолинейная.

Можно применить ранговые коэффициенты корреляции: *r-Спирмена* или *t-Кендалл*:

- обе переменные представлены в количественной шкале (метрической или ранговой);
- связь между переменными является монотонной (не меняет свои знак с изменением величины одной из переменных).

КАКОЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ВЫБРАТЬ

Применение ранговых коэффициентов корреляции при расчете «вручную» требует предварительного ранжирования переменных. Если при этом встречаются одинаковые значения признаков (связи в рангах), применяется формула r -Пирсона для предварительно ранжированных переменных (в случае с r -Спирмена) либо вводятся поправки на связанные ранги (в случае с τ -Кендалла).

Если значение частной корреляции двух переменных по абсолютной величине заметно меньше, чем их парная корреляция, то парная корреляция обусловлена влиянием третьей переменной.

Применяя коэффициенты корреляции, особое внимание следует уделять *графикам двумерного рассеивания*.

ОБРАБОТКА НА КОМПЬЮТЕРЕ

Графики двумерного рассеивания.

Выбираем **Graphs... > Scatter... > Simple**. Нажимаем **Define**.
Получаем график рассеивания назначенных переменных.

Вычисление парных корреляций.

Выбираем **Analyze > Correlate > Bivariate...**

В открывшемся окне диалога переносим интересующие переменные из левой части в правую

Вычисление частной корреляции.

Выбираем **Analyze > Correlate > Partial...**

В открывшемся окне диалога переносим интересующие переменные из левой части в правое верхнее окно (**Variables:**). Затем при помощи нижней кнопки из левой части в правое нижнее окно (**Controlling for:**) переносим переменную, значения которой хотим фиксировать.