При обработке результатов исследования может возникнуть вопрос о том, как связаны между собой различные критерии (признаки, факторы) социальной стратификации. Например, верно ли, что чем выше уровень располагаемого дохода домохозяйства, тем выше средний уровень образованности участников данного домохозяйства; чем выше страна стоит в рейтинге стран по индексу развития человеческого потенциала, тем ниже она стоит в рейтинге стран по индексу валового счастья? Подобная взаимосвязь называется ***корреляционной связью*** или ***корреляцией***. Если такая связь велика, говорят, что признаки тесно (или сильно) коррелируют, в противном случае - они слабо коррелируют. Мерой зависимости (теснотой связи) между признаками является коэффициент корреляции, а его вычисление - корреляционным анализом.

При расчетах коэффициент корреляции получается с положительным или отрицательным знаком. Знак "+" указывает на прямую (положительную) корреляцию, знак "-" на обратную (отрицательную).

*Прямая* корреляция отражает однотипность в изменении признаков: с увеличением значений первого признака увеличивается значение и другого, или с уменьшением первого уменьшается второй. Например, повышение месячного располагаемого дохода физического лица приводит увеличению объема потребления, а уменьшение средних значений располагаемого дохода социальных групп приводит к уменьшению их политической лояльности.

*Обратная* корреляция указывает на увеличение первого признака при уменьшении второго, или уменьшение первого признака при увеличении второго. Например, уменьшение индивидуальных показателей IQ приводит к повышению индивидуальной же агрессивности.

Как видим, знак коэффициента корреляции отражает только направленность зависимости между показателями.

Качественную меру связи оценивают по абсолютному значению коэффициента (от 0 до 1). Тесноту взаимосвязи принято считать по нескольким уровням. Так, если коэффициент корреляции равен [0,99;0,7], то это сильная статистическая взаимосвязь; [0,5;0,69] - средняя; [0,2;0,49] - слабая; [0,09;0,19] - очень слабая. При коэффициенте корреляции, равном нулю, корреляция отсутствует (данные факторы между собой нейтральны).

Корреляция называется *линейной*, когда направление связи между изучаемыми признаками графически и аналитически выражается прямой линией. Анализ линейной корреляции осуществляется с помощью вычисления коэффициентов корреляций. Для измерения *нелинейной*, т.е. криволинейной, зависимости используется показатель, называемый корреляционным отношением.

Выбор коэффициентов корреляции также обусловлен шкалой измерений результатов. В нашем случае – рейтинги стран мира по разным основаниям – шкала является очевидно *порядковой*. Так что будем применять ***ранговый коэффициент корреляции Спирмена.*** При ранговой корреляции сравнивают не сами значения измерений, а только порядок (ранги), поэтому вычисление рангового коэффициента возможно только тогда, когда результаты измерений получены на основе **шкалы не ниже порядковой**. Ранговый коэффициент **r s**не рекомендуется применять, если связанных пар меньше 5 и больше 20.

Итак, как рекомендуется выполнять последний расчет в ПР №1.

1. Записать значения двух рядов измерения.
   1. Т.е. если вам по варианту полагается вычислить коэффициент корреляции, скажем, между страновым индексом обеспеченности ИКТ (информационно-коммуникативными технологиями) и страновым индексом душевного здоровья, то вы выбираете из статистического приложения две соответствующих сводки.
   2. Затем, формируете выборку стран, в которой должны быть страны упомянутые в обеих сводках. Выборка объемом 12 экземпляров должна включать 3 страны из Европы, 3 страны из Азии, 3 страны из Африки, 2 страны из Америки (Северной + Южной), и 1 страна из Австралии и Океании. Внутри каждой континентальной локации страны выбираются случайным образом.
   3. Начинаем составлять таблицу, расставляя страны в порядке убывания значения любого из двух, произвольно выбранного индекса:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Страна** | **Ряды измерений индексов** | | **Ранговые числа** | | **Разность рангов** | |
| **Индексикт** | **Индекспсихов** | **Xi** | **Yi** | **d=xi-yi** | **d2** |
| *Браззавиль* | 0, 98 | 0,78 | 1 | 4,5 | -3,5 | 12,25 |
| *Зубровка* | 0,95 | 0,34 | 2 | 6 | -4 | 16 |
| *Дримлэнд* | 0,80 | 0,99 | 3 | 1 | 2 | 4 |
| *Ла-лалэнд* | 0,79 | 0,78 | 4 | 4,5 | -0,5 | 0,25 |
| *Внутренняя Монголия* | 0,74 | 0,95 | 5 | 2 | 3 | 9 |
| *………* | 0,65 | 0,81 | 6 | 3 | 3 | 9 |
| *……* | 0, 61 | 0,33 | 7 | 7 | 0 | 0 |
| *………..* | 0,59 | 0,28 | 8 | 9 | -1 | 1 |
| *………..* | 0,58 | 0,10 | 9 | 12 | -3 | 9 |
| *………..* | 0.40 | 0,15 | 10 | 11 | -1 | 1 |
| *…….* | 0,39 | 0,21 | 11 | 10 | 1 | 1 |
| *…………* | 0,28 | 0,31 | 12 | 8 | 4 | 16 |
|  |  |  | **78** | **78** |  | **78,5** |

*Рядом со значениями индекса X для каждого испытуемого проставить значения показателей индекса Y.*

*По каждому признаку проставить ранговые числа, т.е. место которое занимает данная страна по данному индексу в выборке из 12 стран. При этом, когда попадаются одинаковые значения, в этом случае общим для обоих значений будет среднеарифметический ранг* *(4 и 5 колонки таблицы). Но лучше избегать такого совпадения значений у разных стран на стадии формирования выборки.*

*Вычислить разность рангов* ***d=xi-yi****с сохранением соответствующего знака - 6 колонка.*

*Возвести разность рангов в квадрат* ***d2*** *7 колонка.*

1. Вычислить сумму квадратов разности рангов ∑d2

∑d2 = 78,5

1. Рассчитываем коэффициент ранговой корреляции **Rs** по формуле Спирмена:

Rs = , где n – это количество пар значений, ну т.е. количество стран в выборке в нашем случае

Rs = 0,7255

1. Формулируем вывод.

На основании полученного результата выявляем связь между изучаемыми признаками стратификации:

Если коэффициент имеет положительный знак (+), то связь положительная, и, наоборот, при отрицательном знаке (-) - связь отрицательная.

По абсолютному значению коэффициента (от 0 до 1) оцениваем количественную меру связи:

- если Rs = 0 - корреляция отсутствует (данные признаки стратификации между собой нейтральны);

- если Rs принадлежит [0,09;0,19] - статистическая взаимосвязь очень слабая;

- если Rs принадлежит [0,2;0,49] - статистическая взаимосвязь слабая;

- если Rs принадлежит [0,5;0,69] - статистическая взаимосвязь средняя;

- если Rs принадлежит [0,99;0,7] - статистическая взаимосвязь сильная.

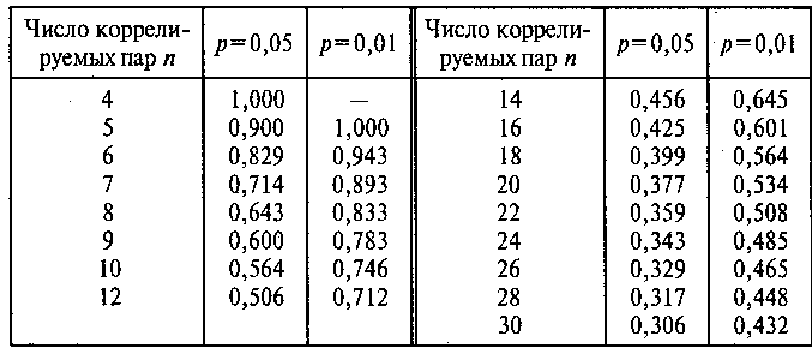
Т.о., на основании расчетного Rs делается вывод о том, что между исследуемыми признаками стратификации существует слабая (средняя, сильная) положительная (отрицательная) связь. В нашем случае, связь между страновым индексом обеспеченности ИКТ (информационно-коммуникативными технологиями) и страновым индексом душевного здоровья является сильной положительной корреляцией, если растет один индекс, будет расти и другой и темп прироста будет примерно одинаковым (чем больше социальных взаимодействий в стране опосредованы информационно-коммуникативными технологиями, тем больше в ней психически здоровых людей (или тем меньше в ней психов). Мы установили эмпирическую закономерность, а теоретическая социология попытается ее объяснить.

5. Проводим проверка достоверности выявленной связи (корреляции) осуществляется сравнением Rs и Rs крит (таблица). Что такое ***p***? Это уровень значимости статистического теста. Обычно он выбирается из априорных соображений (т.е. вне связи с темой конкретного исследования). Выбор часто сводится к назначению одного из популярных вариантов p = 0,005, 0,05, 0,01, 0,1. Нет никаких специальных причин для выбора именно этих значений, поэтому мы выбираем значение p = 0,05

На основании того, что Rs > Rs крит, наличие обнаруженной связи считается *достоверным* при p = 0,05.

На основании того, что Rs < Rs крит, наличие обнаруженной связи считается *недостоверным* при p = 0,05.

Таблица критических значений Rs крит



Итак, в нашем примере количество коррелируемых пар – 12, следовательно, 8 строка / 2 столбец Rs крит =0,506. Rs (0,7255) > Rs крит (0,506). Следовательно, обнаруженная ***корреляция достоверна.***

Все. Конец.