

Лабораторная работа

Корреляционный анализ Часть 3

Вычисление коэффициентов корреляции

Часть 3.1 Теория

Джозеф Вольпе приводит упорядоченный перечень из наиболее часто встречающихся у современного человека «бесполезных», по его обозначению, страхов, которые не несут сигнального значения и лишь мешают полноценно жить и действовать. В отечественном исследовании, проведенном М. Э. Раховой 32 испытуемых должны были по 10-балльной шкале оценить, насколько актуальным для них является тот или иной вид страха из перечня Вольпе. Обследованная выборка состояла из студентов Гидрометеорологического и Педагогического университетов Санкт-Петербурга: 15 юношей и 17 девушек в возрасте от 17 до 28 лет, средний возраст 23 года.

Данные, полученные по 10-балльной шкале, были усреднены по 32 испытуемым, и средние проранжированы. В таблице 1 представлены ранговые показатели, полученные Дж. Вольпе и М. Э. Раховой. Совпадают ли ранговые последовательности 20 видов страхов?

Решение

Сформулируем гипотезы.

H_0 : Корреляция между упорядоченными перечнями страхов в американской и отечественной выборках не отличается от нуля.

H_1 : Корреляция между упорядоченными перечнями страхов в американской и отечественной выборках статистически значимо отличается от нуля.

Все расчеты выполняем, используя таблицу 1

Таблица 1

№	Вид страха	Ранг в американской выборке	Ранг в российской выборке	d	d ²
1	Страх публичного выступления	1	7		
2	Страх полета	2	12		
3	Страх совершить ошибку	3	10		
4	Страх неудачи	4	6		
5	Страх неодобрения	5	9		
6	Страх отвержения	6	2		
7	Страх злых людей	7	5		
8	Страх одиночества	8	1		
9	Страх крови	9	16		
10	Страх открытых ран	10	13		
11	Страх дантиста	11	3		

12	Страх уколов	12	19		
13	Страх прохождения тестов	13	20		
14	Страх полиции	14	17		
15	Страх высоты	15	4		
16	Страх собак	16	11		
17	Страх пауков	17	18		
18	Страх искалеченных людей	18	8		
19	Страх больниц	18	15		
20	Страх темноты	20	14		
Суммы					

Задание 1.

Изучить теоретический материал (представлен в данной лабораторной работе).

Задание 2. Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти эмпирическое значение r_s коэффициента ранговой корреляции Спирмена.

$$r_s = 1 - \frac{6 \sum(d^2)}{n(n^2 - 1)}$$

где d – это разница рангов (например: $1 - 7 = -6$)

Задание 3. По данным таблицы 2 (Критические значения выборочного коэффициента корреляции рангов) определить критическое значение $r_{s\text{кр}}$ при $n = 20$ при уровне значимости 0.05 и 0.01.

Таблица 2

Критические значения выборочного коэффициента корреляции рангов (по В.Ю. Урбаху, 1964)

Связь достоверна, если $r_{s\text{ эмп}} \geq r_{s\text{ табл}}$, 0,05, и тем более достоверна, если $r_{s\text{ эмп}} \geq r_{s\text{ табл}}$, 0,01

n	p		n	p		n	p	
	0,05	0,01		0,05	0,01		0,05	0,01
5	0,94	—	17	0,48	0,62	29	0,37	0,48
6	0,85	—	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,40	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40

Задание 4. Сделать вывод о том, какая гипотеза принимается (H_0 или H_1).

Замечание:

Если корреляция между упорядоченными перечнями страхов в американской и отечественной выборках не достигает уровня статистической значимости, т.е. значимо не отличается от нуля, то принимается гипотеза H_0 . В противном случае принимается гипотеза H_1 .

Задание 5. Результаты выполненной работы разместите в Теме № 6 ЭУК Moodle.

Часть 3.2

Теория

Корреляция между индивидуальным и среднегрупповым профилями.

Задание 1.

Изучить теоретический материал (представлен в данной лабораторной работе).

Выборке петербуржцев в возрасте от 20 до 78 лет (31 мужчины и 46 женщин), уравниваемой по возрасту таким образом, что в возрасте старше 55 лет составляли в ней 50%, предлагалось ответить на вопрос: «Какой уровень развития каждого из перечисленных ниже качеств необходим для депутата Городского собрания Санкт-Петербурга?» Оценка производилась по 10-балльной шкале. Параллельно с этим обследовалась выборка из депутатов и кандидатов в депутаты в Городское собрание Санкт-Петербурга ($n = 14$). Индивидуальная диагностика политических деятелей и претендентов производилась с помощью Оксфордской системы экспресс-видеодиагностики по тому же набору личных качеств, который предъявлялся выборке избирателей.

В таблице 3 представлены средние значения, полученные для каждого из качеств в выборке избирателей («эталонный ряд») и индивидуальные значения одного из депутатов Городского собрания.

Необходимо определить, насколько индивидуальный профиль депутата К-ва коррелирует с эталонным профилем.

Таблица 3.

Усредненные эталонные оценки избирателей ($n = 77$) и индивидуальные показатели депутата К-ва по 18 личностным качествам экспресс-видеодиагностики

Таблица 3

Наименования качества		Усредненные эталонные оценки избирателей	Индивидуальные показатели депутата К-ва
1	Общий уровень культуры	8.64	15
2	Обучаемость	7.89	7
3	Логика	8.38	12
4	Способность к творчеству нового	6.97	5
5	Самокритичность	8.28	14
6	Ответственность	9.56	18
7	Самостоятельность	8.12	13
8	Энергия, активность	8.41	17
9	Целеустремленность	8.00	19
10	Выдержка, самообладание	8.71	9
11	Стойкость	7.74	16

12	Личностная зрелость	8.10	11
13	Порядочность	9.02	12
14	Гуманизм	7.89	10
15	Умение общаться с людьми	8.74	8
16	Терпимость к чужому мнению	7.84	6
17	Гибкость поведения	7.67	4
18	Способность производить благоприятное впечатление	7.23	8

Из таблицы 3 оценки избирателей и индивидуальные показатели депутата варьируют в разных диапазонах. Оценки избирателей были получены по 10-балльной шкале, а индивидуальные показатели по экспресс-диагностике измеряются по 20-балльной шкале. Ранжирование позволяет перевести обе шкалы измерения в единую шкалу, где единицей измерения будет 1 ранг, а максимальное значение составит 18 рангов.

Ранжирование необходимо произвести отдельно по каждому ряду значений. В данном случае целесообразно начислять большему значению меньший ранг, чтобы сразу можно было увидеть, на каком месте по значимости (для избирателей) или по выраженности (у депутата) находится то или иное качество.

Результаты ранжирования представлены в таблице 4. Качества перечислены в последовательности, отражающей эталонный профиль.

Таблица 4

Наименования качества		Ряд 1: ранг качества в эталонном профиле	Ряд 2: ранг качества в индивидуальном профиле	d	d ²
1	Ответственность	1	2		
2	Порядочность	2	8.5		
3	Умение общаться с людьми	3	13.5		
4	Выдержка, самообладание	4	12		
5	Общий уровень культуры	5	5		
6	Энергия, активность	6	3		
7	Логика	7	8.5		
8	Самокритичность	8	6		
9	Самостоятельность	9	7		
10	Личностная зрелость	10	10		
11	Целеустремленность	11	1		
12	Обучаемость	12.5	15		
13	Гуманизм	12.5	11		
14	Терпимость к чужому мнению	14	16		
15	Стойкость	15	4		
16	Гибкость поведения	16	18		
17	Способность производить благоприятное впечатление	17	13.5		
18	Способность к творчеству нового	18	17		

Суммы				
-------	--	--	--	--

Сформулируем гипотезы.

H₀ : Корреляция между индивидуальным профилем депутата К-ва и эталонным профилем, построенным по оценкам избирателей, не отличается от нуля.

H₁ : Корреляция между индивидуальным профилем депутата К-ва и эталонным профилем, построенным по оценкам избирателей, значительно отличается от нуля.

Решение.

Поскольку в обоих сопоставляемых ранговых рядах присутствуют группы одинаковых рангов, перед вычислением коэффициента ранговой корреляции необходимо внести поправки на одинаковые ранги **T_a** и **T_b** :

Задание 2.

Все расчеты выполняются, используя таблицу 4.

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена (Примеры 1 и 2 из материалов лекции).

$$T_a = \sum (a^3 - a)/12$$

$$T_b = \sum (b^3 - b)/12$$

где **a** – объем каждой группы одинаковых рангов в ранговом ряду А,
b - объем каждой группы одинаковых рангов в ранговом ряду В.

В данном случае, в ряду А (эталонный профиль) присутствует одна группа одинаковых рангов качества «обучаемость» и «гуманизм» имеют один и тот же ранг 12.5; следовательно, **a = 2**.

$$T_a = \frac{(2^3 - 2)}{12} = 0.5.$$

В ряду В (индивидуальный профиль) присутствуют две группы одинаковых рангов, при этом **b₁=2** и **b₂=2**.

$$T_b = [(2^3 - 2) + (2^3 - 2)]/12 = 1.00$$

Для подсчета эмпирического значения **r_s** используем формулу

$$r_s = 1 - 6 \frac{\sum d^2 + T_a + T_b}{n(n^2 - 1)}$$

В данном случае :

$$r_s = 1 - \frac{6 * 487.5 + 0.50 + 1.00}{18 * (18^2 - 1)} = 1 - \frac{2926.5}{5814} = 0.4967$$

Если бы поправка на одинаковые ранги не вносилась, то величина r_s была бы лишь на (на 0.0002) выше:

$$r_s = 1 - \frac{6 * 487.5}{18 * (18^2 - 1)} = 1 - \frac{2925}{5814} = 0.4969$$

при больших количествах одинаковых рангов изменения r_s могут оказаться гораздо более существенными. Наличие одинаковых рангов означает меньшую степень дифференцированности упорядоченных переменных и, следовательно, меньшую возможность оценить степень связи между ними.

Задание 3. По таблице 2 определите критические значения r_s при $n = 18$ при уровне значимости 0.05 и 0.01.

Сделайте вывод о том, какая гипотеза отвергается, а какая принимается.

Общий вывод: из таблицы 4 видно, что депутат К-в имеет более низкий ранг по шкалам Умения общаться с людьми и более высокие ранги по шкалам Целеустремленности и Стойкости, чем это предписывается избирательским эталоном. Этими расхождениями, главным образом, и объясняется некоторое снижение полученного r_s .

Задание 4. Результаты выполненной работы разместите в Теме № 6 ЭУК Moodle.