

ОСОБЕННОСТИ ПРИМЕНЕНИЯ ОПИСАТЕЛЬНОЙ И АНАЛИТИЧЕСКОЙ СТАТИСТИКИ В СОЦИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

FEATURES OF THE USE OF DESCRIPTIVE AND ANALYTICAL STATISTICS IN SOCIOLOGICAL RESEARCH

V. Zamyatin
E. Molchanova

Summary. The purpose of the article is to identify the features of the analysis of sociological distributions of indices, linearly expressed based on data from large random samples. The relationship between the general drug addiction risk indicator and the normal type of distribution is considered. A data matrix for 75 indicators was obtained by testing 2109 ISTU students using the ASSOL IS. It is shown that, contrary to the opinion of Russian authoritative scientists, methods of descriptive and analytical statistics based on the normal distribution law can be applied to indices calculated from groups of factors consisting of ordinal and binary indicators. Using the 3σ criterion, respondents with an increased risk of drug addiction were identified. Estimates of pair correlation according to Pearson, according to Spearman and using the phi coefficient (for binary scales) were compared. If there are related ranks in the data sample, the calculation of the Spearman coefficient should be carried out taking into account corrections. All three types of correlation coefficients gave similar results. It has been established that when analyzing large samples there are no clear criteria for assessing the strength of the connection. It has been shown that identifying strong pairwise correlations between indicators makes it possible to reduce the size of the questionnaire without loss of reliability.

Keywords: sociology, Pearson correlation coefficient, Spearman correlation coefficient, phi correlation coefficient, analytical statistics, descriptive statistics, standard normal distribution, random quantity, indicator, index, scale.

Замятин Владимир Михайлович

Аспирант, Федеральное Государственное бюджетное образовательное учреждение высшего образования «Иркутский Государственный Университет Путей Сообщения»
zamvov01@yandex.ru

Молчанова Елена Ивановна

Профессор, доктор технических наук, Федеральное Государственное бюджетное образовательное учреждение высшего образования «Иркутский Государственный Университет Путей Сообщения»
moleli59@gmail.com

Аннотация. Цель статьи — выявить особенности анализа социологических распределений индексов, линейно выраженных по данным больших случайных выборок. Рассмотрено соотношение показателя общего риска наркотизации с нормальным типом распределения. Матрица данных по 75 индикаторам получена при тестировании 2109 студентов ИРГУПС с помощью ИС «АССОЛЬ». Показано, что вопреки мнению российских авторитетных ученых, методы описательной и аналитической статистики, основанные на нормальном законе распределения, можно применять к индексам, рассчитанным по группам факторов, состоящих из порядковых и бинарных индикаторов. С помощью 3σ -критерия выявлены респонденты с повышенным риском наркотизации. Сопоставлены оценки парной корреляции по Пирсону, по Спирмену и с помощью фи-коэффициента (для бинарных шкал). При наличии связанных рангов в выборке данных расчет коэффициента Спирмена следует выполнять с учетом поправок. Все три типа коэффициентов корреляции дали сходный результат. Установлено, что при анализе больших выборок нет четких критериев для оценки силы связи. Показано, что выявление сильных парных корреляций между индикаторами позволяет уменьшить размер анкеты без потери надежности.

Ключевые слова: социология, коэффициент корреляции Пирсона, коэффициент корреляции Спирмена, фи-коэффициент корреляции, аналитическая статистика, описательная статистика, нормальный закон распределения, случайная величина, индикатор, индекс, шкала.

Часто при обработке данных социологических и психологических исследований расчет коэффициентов корреляции выполняется и обсуждается без учета типа измерительных шкал и форм распределения факторов, например, [1], а коэффициент — альфа Кронбаха оценивают без указания типа показателя корреляции [2, С. 651–657]. Тогда как от результатов применения описательной статистики, полученных на первом этапе статистического анализа данных, должен зависеть дальнейший подход к исследованию статистических гипотез.

Автор работы [3, С.36–37] считает, что идея соотношения социологических распределений (шкал, выборок, коэффициентов корреляции) с нормальным распределением часто используется, но не обоснована в социологии. Приведем несколько цитат из этой статьи: «В частности, не всегда уточняется распределение, по каким шкалам может подчиняться закону Гаусса. Распределения по атрибутивным шкалам не могут рассматриваться на предмет соответствия нормальному распределению, так как это независимые шкалы и у них не может быть среднего значения и, соответственно, отклонений». «Порядковые шкалы также не имеют корректного среднего

значения, поэтому также не могут соотноситься с нормальным распределением». Вместе с тем, атрибутивные шкалы при обработке, как правило, сводятся к дихотомическим. А дихотомические шкалы, по мнению того же автора «зачастую являются свернутыми порядковыми, сводимыми к абсолютным». Но Ильясов Ф.Н. утверждает, что и «среди распределений по абсолютным шкалам не наблюдается типов распределений, соответствующих нормальному». Он цитирует Крыштановского А.О. [4, С.110], Орлова А.И. [5, С.64–66][6] и делает вывод, что «идея нормального распределения в социологии не обоснована. Ряд методов измерения и анализа данных в социологии зиждется на гипотезе нормального распределения. Поскольку сама гипотеза нормальности не находит обоснования и подтверждения, то и методы, на ней основанные, могут быть недостаточно достоверными».

Вместе с тем, авторы статьи [7, С.148] не соглашались с этим выводом «Статистика — не в природе вещей, а в методе исследователя (хорошо или плохо описывает построенную исследователем систему)». На взгляд авторов, «позиция вышеназванных российских ученых свидетельствует лишь о малом количестве построенных исследователями моделей социальных систем, где нормальное распределение успешно использовалось». Они исследовали вопрос: почему нормальное распределение является повсеместно встречающимся. Приводят ссылки на источники, доказывающие этот тезис математически и источники, рассматривающие его с позиции философии статистики. На основании последних источников приводится заключение, что «закон нормального распределения наилучшим образом моделирует общий способ устройства социальных и физических процессов» [7, С.147].

Автором статьи [8, С.4-16] теоретически доказано, что индексы, которые линейно выражаются по данным из случайных выборок через частоты ответов респондентов, должны иметь приблизительно нормальный закон распределения. Тогда к ним можно применять методы математической статистики. «Мнение о том, что временной ряд некорректно соотносить с гауссовым распределением [3, С. 37], несомненно верно в случае, когда соответствующий индекс X не является квазистационарным. Однако очевидно, что на коротком промежутке времени $[t_1; t_n]$, когда индекс X является квазистационарным, числа $x(t_1), x(t_2), \dots, x(t_n)$ можно считать выборкой значений случайной величины. Последняя может быть (или не быть) нормальной».

Целью наших исследований являлось изучить особенности применения описательной и аналитической статистики на примере результатов тестирования студентов ИРГУПС в количестве $N=2109$ человек с помощью ИС социологического и психологического тестирования

«АССОЛЬ» [9] по анкете «Исходная оценка наркотизации» ($Q=75$ вопросов), модифицированной и адаптированной Центром развития человеческого капитала ИРГУПС, с целью выявления социально-психологических факторов риска и защиты от немедицинского употребления наркотиков.

Оценка применимости методов описательной статистики

Одним из важнейших этапов описательной статистики является построение гистограмм, результаты которого дают представление о наблюдаемой функции плотности распределения вероятностей [10].

Рассмотрим генеральную совокупность полученных значений показателя общего риска наркотизации, распределение которой неизвестно. Проверим гипотезу о том, что генеральная совокупность распределена по нормальному закону. Выборочные данные группируются в дискретный вариационный ряд (см. рисунок 1) с вариантами x_i и соответствующими частотами n_j . Здесь же представлен ряд теоретических частот nt_j , рассчитанных по формуле: $n_{tj} = \Delta \cdot m \cdot f(x_i)$, где

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \bar{x}_e}{\sigma}\right)^2} \text{ — функция Гауса, } m \text{ — число}$$

интервалов, Δ — ширина интервалов. Ширина каждого интервала равна 1.

Сначала оценили принадлежность вариационного ряда показателя общего уровня риска наркотизации к нормальному закону с помощью мер центральной тенденции и вариации. Автор [11] перечисляет 3 аналитических критерия для нормального распределения:

1) У нормального распределения оценка математического ожидания по выборочной средней \bar{x}_e совпадает с модой M_0 и медианой me . В нашем случае рассчитанные выборочные показатели весьма близки друг к другу: $\bar{x}_e = 143$, $me = 141$, $M_0 = 131$.

2) Практически все значения нормальной случайной величины должны находиться в интервале $a - 3\sigma$, $a + 3\sigma$. Найдём этот интервал для нашей выборки. Математическое ожидание «а» оценивается выборочной средней $\bar{x}_e = 143$, а стандартное отклонение «сигма» — выборочным стандартным отклонением $\sigma_e = 20.84$. Таким образом, наш эмпирический интервал: (80.5; 205.6) — и в него попадают практически все значения, кроме 9 (0.4 % от выборки). Именно эти точки и являются искомыми тревожными значениями показателя общего уровня риска наркотизации в нашем исследовании.

На рисунке 2 представлена диаграмма типа Диаграмма размаха («ящик с усами»), изображающая медиану

вариационные ряды частот общий балл

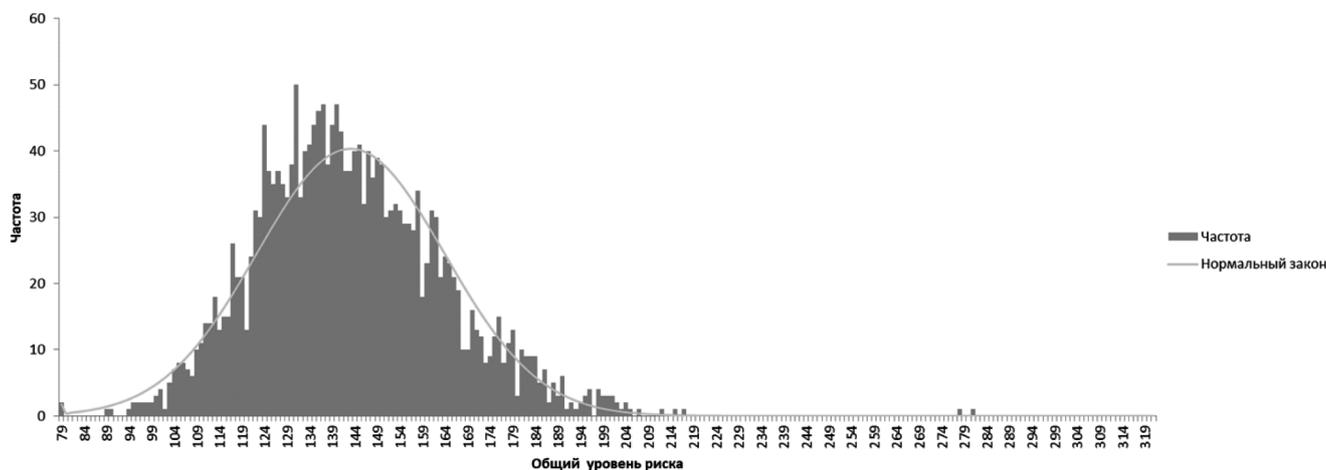


Рис. 1. Дискретный вариационный ряд показателя общего уровня риска наркотизации — частота n_j и теоретическая частота — n_{ij} для нормального закона

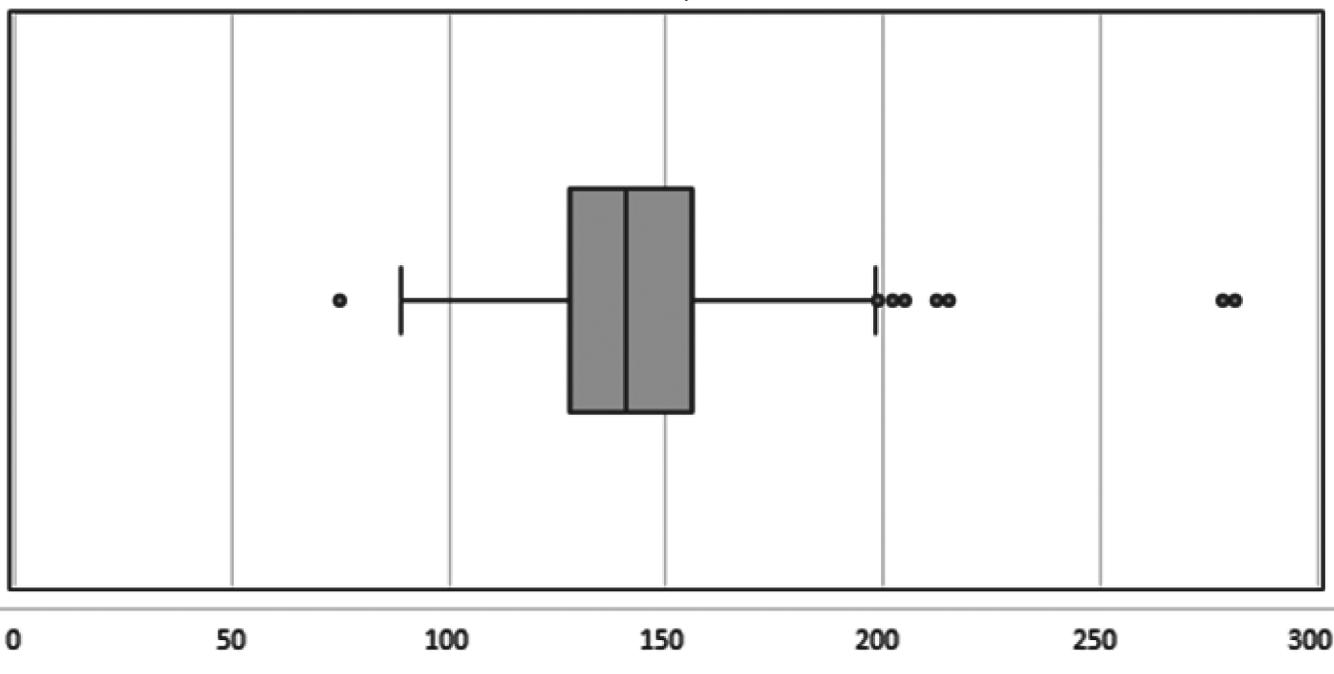


Рис. 2. Диаграмма размаха

показателя общего уровня риска наркотизации с 95 % доверительным интервалом, интерквартильный размах и выбросы.

3) Коэффициенты асимметрии и эксцесса нормального распределения равны нулю. В нашем случае эти характеристики не равны нулю, но и невелики: $A_3 = 0.6$, $E_k = 1.6$. Однако, если коэффициент асимметрии > 0.5 , то она считается существенной. После исключения из выборки 9 грубых выбросов коэффициент асимметрии распределения уменьшился до 0.19, и она стала незначима, эксцесс поменял знак -1.6 .

Следуя рекомендациям литературных источников, заключаем, что перечисленные выше предпосылки ещё

не означают, что соответствующая гипотеза будет принята. Далее применили графический метод Квантиль-Квантиль (Q-Q plot). Согласно источнику [12] Этот метод «позволяет:

- оценить степень отклонения данных от теоретического распределения;
- графически проиллюстрировать такие параметры как расположение данных (медиану), масштаб (дисперсию) и скошенность (наклон функции плотности распределения);
- сравнить две выборки между собой;
- делать выводы, не основываясь на таких спорных показателях как $p.value$. Фактически, $p.value$ в случае Q-Q Plot будет оценивать человеческий мозг на основе визуального изучения».

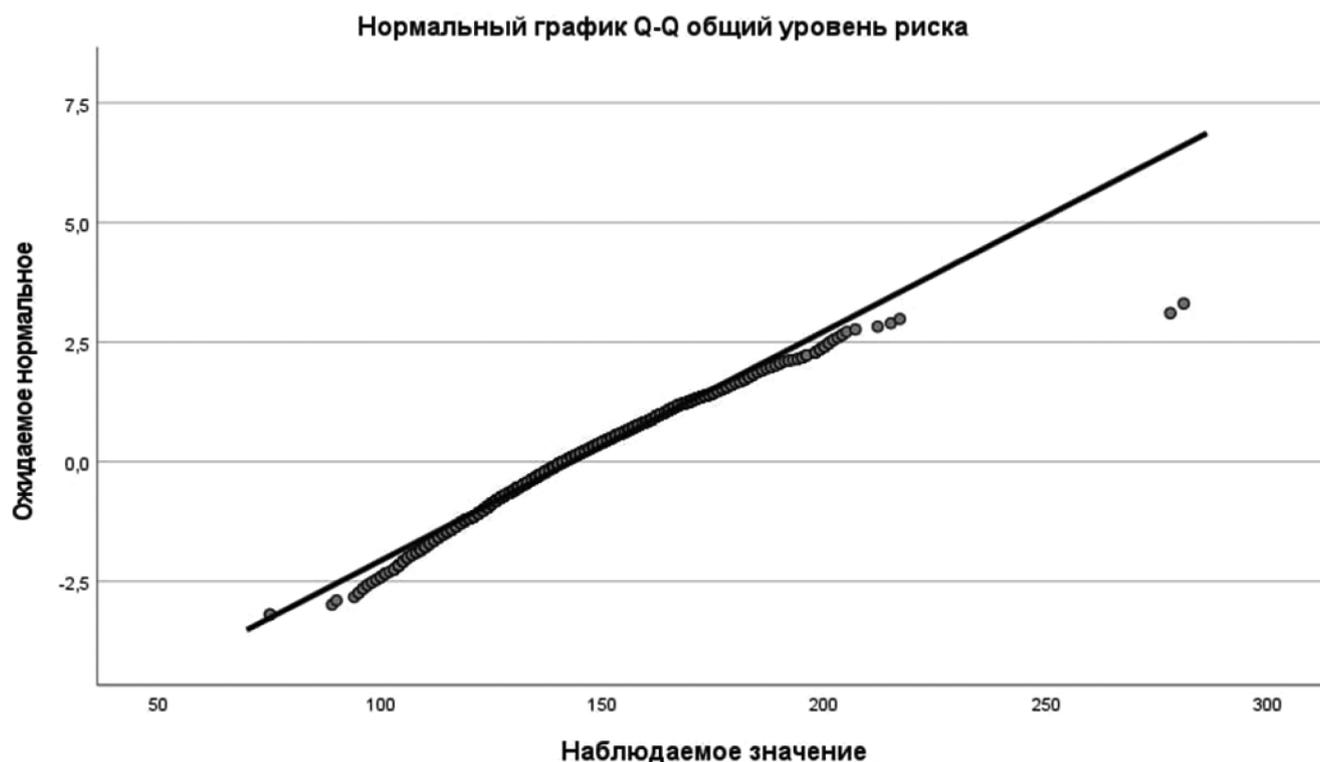


Рис. 3. Диаграмма Квантиль-Квантиль (Q-Q plot) для показателя общего уровня риска наркотизации

На рисунке 3 представлен график типа Q-Q plot, полученный средствами статистического пакета IBM SPSS Statistics. Как видно частота n_j и теоретическая частота nt_j хорошо согласуются, а выбросы соответствуют точкам, выделенным на Диаграмме размаха.

Оценка применимости методов аналитической статистики

Результаты полученные, методами описательной статистики позволяют, нам и при решении задач аналитической статистики применять критерии, основанные на нормальном законе распределения случайных величин.

Исходя из того, что индикаторы, соответствующие баллам на вопросы теста, были представлены в порядковой и бинарных шкалах, сопоставили результаты оценки коэффициентов корреляции Пирсона, Спирмена и Фи-коэффициента. Для того, чтобы контролировать алгоритм расчета, первоначально вычисления выполняли с помощью приложения MS Excel, а не специализированных статистических пакетов. В частности, при оценке коэффициентов Спирмена необходимо учитывать число совпадающих (связанных) рангов, и расчетная формула должна содержать поправки. Согласно [13, С.35].

При наличии связанных рангов выражение r^s приобретает вид:

$$r^s = \frac{\frac{1}{6}(N^3 - N) - \sum_{i=1}^N (x^i - y^i)^2 - T - U}{\sqrt{\left[\frac{1}{6}(N^3 - N) - 2T \right] \left[\frac{1}{6}(N^3 - N) - 2U \right]}} \quad (1)$$

где

$$T = \frac{1}{12} \sum_{i=1}^k t_i (t_i^2 - 1), \quad U = \frac{1}{12} \sum_{i=1}^l u_i (u_i^2 - 1) \quad (2)$$

k и l — число случаев появления связанных рангов; t_i и u_i — число совпадающих рангов в каждом из упорядочений; N — число респондентов; r^s — коэффициент корреляции Спирмена.

Сформировалось мнение, что интерпретация матрицы коэффициентов корреляции Пирсона имеет смысл только при метрических шкалах, а большинство социальных признаков измеряются категориальными шкалами [14]. Авторы на основании литературных источников обсуждают несколько способов легитимизовать расчет коэффициентов корреляции Пирсона, для порядковых индикаторов, в том числе принять допущение о том, что порядковые индикаторы можно считать интервальными, или дихотомизировать их. В наших исследованиях использовано первое допущение. При этом измерения проводили по четырехбалльной шкале Лайкерта. Наибольшую проблему при оценке коэффициентов корреляции Пирсона должны представлять индикаторы, выраженные в бинарной шкале. В нашем исследовании это были восемь следующих вопросов: 60) Меняли ли Вы

Таблица 3.

Сопоставление коэффициентов корреляции для бинарных индикаторов

№ Вопрос №	Коэффициент	60)	43)	24)	47)	48)	49)	63)	57
60)	Фи	1.000	0.043	-0.016	0.053	0.035	0.063	0.075	0.089
	Пирсон	1.000	0.043	-0.016	0.053	0.035	0.063	0.075	0.089
	Спирмен	1.000	0.046	-0.016	0.054	0.037	0.065	0.076	0.092
43)	Фи	0.043	1.000	-0.092	0.092	0.092	0.051	0.075	0.059
	Пирсон	0.043	1.000	-0.092	0.092	0.092	0.051	0.075	0.059
	Спирмен	0.046	1.000	-0.087	0.096	0.098	0.055	0.082	0.070
24)	Фи	-0.016	-0.092	1.000	-0.068	-0.084	-0.055	-0.032	-0.059
	Пирсон	-0.016	-0.092	1.000	-0.068	-0.084	-0.055	-0.032	-0.059
	Спирмен	-0.016	-0.087	1.000	-0.067	-0.082	-0.054	-0.030	-0.055
47)	Фи	0.053	0.092	-0.068	1.000	0.524	0.498	0.220	0.052
	Пирсон	0.053	0.092	-0.068	1.000	0.524	0.498	0.220	0.052
	Спирмен	0.054	0.096	-0.067	1.000	0.525	0.499	0.222	0.056
48)	Фи	0.035	0.092	-0.084	0.524	1.000	0.445	0.205	0.075
	Пирсон	0.035	0.092	-0.084	0.524	1.000	0.445	0.205	0.075
	Спирмен	0.037	0.098	-0.082	0.525	1.000	0.446	0.208	0.081
49)	Фи	0.063	0.051	-0.055	0.498	0.445	1.000	0.195	0.025
	Пирсон	0.063	0.051	-0.055	0.498	0.445	1.000	0.195	0.025
	Спирмен	0.065	0.055	-0.054	0.499	0.446	1.000	0.197	0.029
63)	Фи	0.075	0.075	-0.032	0.220	0.205	0.195	1.000	0.119
	Пирсон	0.075	0.075	-0.032	0.220	0.205	0.195	1.000	0.119
	Спирмен	0.076	0.082	-0.030	0.222	0.208	0.197	1.000	0.124
57)	Фи	0.089	0.059	-0.059	0.052	0.075	0.025	0.119	1.000
	Пирсон	0.089	0.059	-0.059	0.052	0.075	0.025	0.119	1.000
	Спирмен	0.092	0.070	-0.055	0.056	0.081	0.029	0.124	1.000

местожительство за последний год?; 43) Пробовали ли Вы когда-нибудь наркотики?; 24) Являетесь ли Вы участником какого-либо неформального общественного движения или общественной организации?; 47) Знаете ли Вы людей, которые в прошлом употребляли марихуану или другие наркотики?; 48) Знаете ли Вы людей, которые в прошлом продавали или имели дело с наркотиками?; 49) Знаете ли Вы людей, которые в прошлом делали незаконные вещи, которые могли привести к осложнениям в отношениях с полицией?; 63) Были ли когда-нибудь у кого-либо в Вашей семье проблемы с алкоголем или наркотиками?; 57) Меняли ли Вы ВУЗ в прошлом году?

Для сохранения масштаба индикаторов при вычислении агрегированных показателей, их кодировали как 1 или 4. Как видно из таблицы 3, для бинарных переменных, значения всех трех рассматриваемых коэффициентов корреляции близки.

Вероятно, при таком способе кодирования бинарных индикаторов среднее арифметическое может служить оценкой математического ожидания несмотря на то, что значения не дихотомизированы до уровня $0 \div 1$. Для подтверждения этой гипотезы для восьми бинарных индикаторов рассчитали новые оценки среднего арифметического, для случая дихотомизированной переменной, интерпретируя его, как долю объектов, имеющих значение 1 среди всех изученных объектов [16]. При этом требование симметричности распределения превращается в требование его равномерности. А непрерывность дихотомических переменных проявляется в способности характеризовать совокупность изучаемых объектов любой точкой, отражающей (якобы) среднее арифметическое на отрезке $[0;1]$ [14].

На рисунке 4 сопоставили результаты традиционной оценки среднего арифметического и по способу, пред-

ложенному в работе [15]. Чтобы масштаб сравниваемых величин был сопоставим, традиционную оценку представили в долях от 4. Как видно традиционные оценки чуть выше, но тренд значений совпадает.

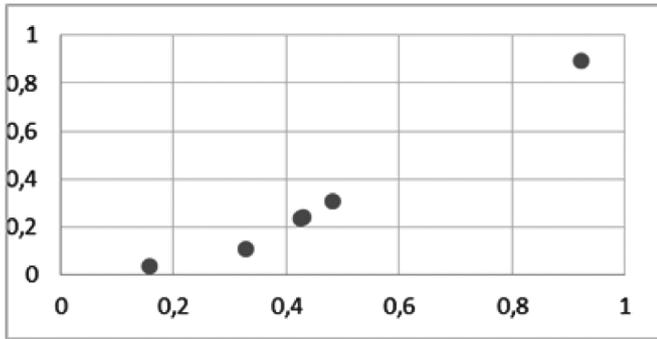


Рис. 4. Диаграмма рассеяния для средних арифметических восьми бинарных переменных, рассчитанных по значениям соответствующего индикатора (ось X) и по способу, предложенному в работе [15] (ось Y)

Для порядковых индикаторов сопоставили только коэффициенты Пирсона и Спирмена. На рисунке 5 приведены точечные диаграммы рассеяния значений этих коэффициентов для показателя общего уровня наркотизации и первого вопроса. Как видно значения коэффициентов практически совпадают.

В работе [16] говорится о большом влиянии выбросов на коэффициент корреляции. Рекомендуется исключить их из выборки и делается оговорка, что это возможно только на основании соображений предметной области, а не на применении статистических критериев. В нашем случае последняя оговорка очень важна, так как выбросы справа на Диаграмме размаха соответствуют результату достижения цели исследования, т.е. выделения респондентов с высоким уровнем риска наркотизации. Выбросы не удаляли. Значения индикаторов и общего показателя для них не являются погрешностями эксперимента, а как раз наоборот, проявлением закономерностей в данных опроса.

Значимость коэффициентов корреляции сначала оценивали по шкале Чеддока. На рисунке 6а представлена тепловая карта коэффициентов корреляции Спирмена. Красным цветом обозначены значения, превышающие 0.3, желтым -0.5. При разработке анкеты вопросы были сгруппированы по пяти социометрическим факторам: близкое окружение, индивидуальный, макросоциальная среда, образовательная среда, семейный. И показатель общего уровня риска наркотизации является суммой баллов, полученных респондентом по индикаторам каждой группы. Здесь приведены не только коэффициенты парной корреляции между индикаторами и показателем общего уровня риска, но и корреляции с агрегированными показателями по выделенным пяти группам факторов. Именно они соответствуют желтому прямоугольнику в нижнем углу карты 6а.

Однако в работе Орлов А.И. [16] утверждает, что критическое значение значимости коэффициента корреляции зависит от объема выборки и стремится к нулю при больших объемах. Вместе с тем, достаточно сложно найти таблицы с рассчитанными критическими значениями при больших N. Например, в работе [17] для числа степеней свободы $f = 120$ приводится значение корреляции Пирсона $r_{кр} = 0.15$ и следующее $f = \infty$ $r_{кр} = 0.07$. В работе [18] наибольшее значение $f = 1000$, а $r_{кр} = 0.06$, а для корреляции Спирмена значения f заканчиваются на 40. На рисунке 6б приведена тепловая карта коэффициентов Спирмена, где желтым цветом выделены значения, превышающие 0.06. Как видно, наибольшая часть коэффициентов оказывается значимой.

Для проверки достоверности полученных результатов повторили расчеты коэффициентов корреляции с помощью аналитической low-code платформы Logiplot [19]. Значения матрицы коэффициентов корреляции Пирсона совпали с рассчитанными нами в Excel. Что касается матрицы коэффициентов корреляции Спирмена, то платформа Logiplot позволяет выполнять расчет только для случая отсутствия связанных рангов в порядковых

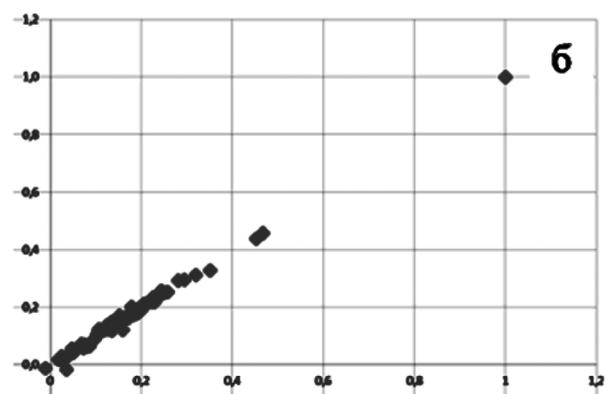
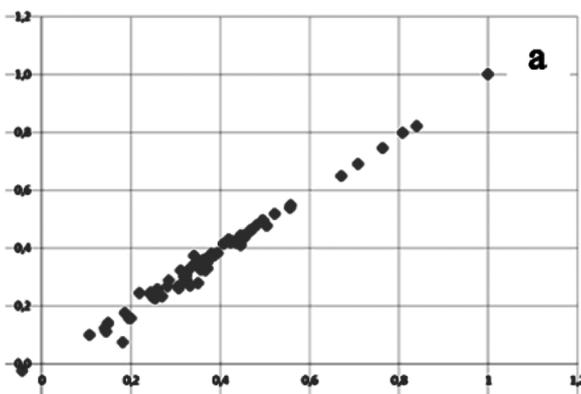


Рис. 5. Диаграммы рассеяния коэффициентов Пирсона и Спирмена для показателя общего уровня наркотизации (а) и первого вопроса (б)

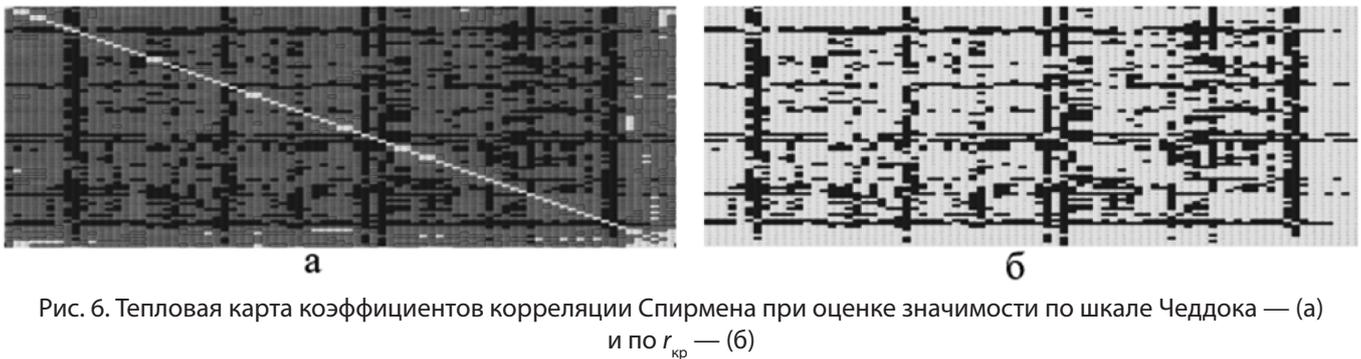


Рис. 6. Тепловая карта коэффициентов корреляции Спирмена при оценке значимости по шкале Чеддока — (а) и по $r_{кр}$ — (б)

индикаторах. Если такие совпадающие значения рангов наблюдаются, то выполняется расчет коэффициентов Пирсона. В наших данных, имеет место очень большое число связанных рангов, поэтому при запросе на вычисление коэффициентов корреляции Спирмена мы получили опять матрицу коэффициентов Пирсона. Но это вполне справедливо, так как наши результаты расчета коэффициентов корреляции Спирмена по формуле с поправками тоже очень близки к матрице коэффициентов Пирсона. Пренебрежение поправками сильно завышает значения коэффициентов корреляции Спирмена.

Наличие в матрице коэффициентов корреляции индикаторов с большой силой парной корреляции говорит о возможности сокращения числа вопросов в анкете без потери ее согласованности. Из результатов теста исключили 5 вопросов из скоррелированных пар. Так же был выявлен и исключен вопрос, незначимо коррелирующий с другими индикаторами и показателем общего риска наркотизации. Оценили коэффициенты «альфа» Кронбаха по результатам теста до 0.895 и после 0.889 исключения, отобранных шести вопросов. Как видно, согласованность теста не уменьшилась, а сокращение числа вопросов уменьшает время его прохождения.

Заключение

Выполненные исследования подтвердили мнение авторов [20] о том, что гипотеза нормальности и методы, на ней основанные, могут быть достоверно применены при обработке результатов социологических тестов, несмотря на мнение других авторитетных российских ученых. Так, например, применения 3-сигма критерия позволяет выявить респондентов с повышенным риском наркотизации. Наши результаты, однако, говорят о том, что социологические данные при проверке гипотез согласия и корреляции необходимо обрабатывать несколькими способами и сопоставлять полученные результаты. Подобный вывод сделан Орловым А.И. [16], на основе теории устойчивости. Это требует высокой квалификации аналитика, а простое применение статистических пакетов может привести ошибочным выводам. В частности, при оценке корреляции Спирмена надо учитывать связанные ранги, а при выявлении грубых выбросов, опираться на содержательную часть исследования. Анализ коэффициентов парной корреляции между индикаторами позволяет уменьшить объем анкеты без потери согласованности. Полученные нами рекомендации и алгоритмы визуализации будут реализованы в ИС «АССОЛЬ» в качестве подсистемы поддержки принятия решения инженером-социологом, разработанной в ИргУПС.

ЛИТЕРАТУРА

1. Crossman, Ashley. «Correlation Analysis in Research.» ThoughtCo, Apr. 5, 2023, thoughtco.com/what-is-correlation-analysis-3026696.
2. Муслон С.А. Вычисление α Кронбаха при тестировании и опросах по исследованию качества жизни на занятиях по статистике в медицинских вузах — Москва: ООО «Издательство АЛЕФ», 2023.
3. Ильясов Ф.Н. Типы шкал и анализ распределений в социологии // Мониторинг общественного мнения: экономические и социальные перемены. 2014. №4. С. 24–40.
4. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. М.: ГУ ВШЭ, 2006.
5. Орлов А.И. Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория. 1991. № 7.
6. Орлов А.И. Прикладная статистика: учебник. М.: Экзамен, 2006.
7. В. Комарова, Ю.М. Вассерман, Н. Селиванова-Федорова. Уровень развития человеческого потенциала: глобальная дифференция // Вестник Пермского национального исследовательского политехнического университета. Социально-экономические науки 2019. № 4. С.145–160.
8. А.Д. Зотьева. О социологических индексах с нормальным законом распределения//Мониторинг общественного мнения: экономические и социальные перемены. 2020. №4.
9. Информационная система социологического и психологического тестирования «АССОЛЬ» URL: <https://www.irgups.ru/eis/aist/> (дата обращения: 01.03.2024).
10. Петрушин В.Н., Ульянов М.В., Чертихина И.А., Никульче Е.В. Бикритериальный метод построения и оценки качества гистограмм // информационные технологии и вычислительные системы. 2012. №4.

11. Емелин А. Теория вероятностей. Базовые термины и понятия URL: http://mathprofi.ru/teorija_verojatnostei.html (дата обращения: 01.03.2024).
12. Силкин В. Q-Q Plots. URL: <https://habr.com/ru/articles/578754/> (дата обращения: 01.03.2024).
13. Низаметдинов Ш.У., Румянцев В.П. Анализ данных: учебное пособие. М. НИЯУ МИФИ, 2012. 288 с.
14. Зангиева И.К., Ротмистров А.Н. Сравнительный анализ способов проведения факторного анализа на порядковых переменных // Мониторинг общественного мнения: Экономические и социальные перемены. 2018. № 3. С. 29–46
15. Толстова Ю.Н. Анализ социологических данных // Методология, дескриптивная статистика, изучение связей между номинальными признаками — М.: Научный мир, 2000. 352с.
16. Орлов А.И. Ошибки при использовании коэффициентов корреляции и детерминации // Заводская лаборатория. Диагностика материалов. 2018. №3.
17. Татарчук Д.П. Математические основы психологии: учебное пособие. Орский гуманитарно-технологический институт. 2015.
18. Fisher R.A., Frank Y. Statistical Tables for Biological, Agricultural and Medical Research, 6th ed., published by Longman Group, Ltd., London (previously published by Oliver and Boyd, Edinburgh), Table VII.
19. Loginom Коэффициент ранговой корреляции Спирмена URL: <https://wiki.loginom.ru/articles/rank-correlation-spearman.html>. (дата обращения: 01.03.2024).
20. Зотьева И.К., Комарова И.В., Вассерман Ю.М., Селиванова-Федорова Н. Уровень развития человеческого потенциала // Вестник ПНИПУ. Социально-экономические науки. 2019. №4.

© Замятин Владимир Михайлович (zamvov01@yandex.ru); Молчанова Елена Ивановна (moleli59@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»