

К.А. Колобова

Институт археологии и этнографии СО РАН
Новосибирск, Россия
E-mail: kolobovak@yandex.ru

Применение z-оценки для визуализации и корреляции характеристик археологических комплексов

Стандартизация – это переработка данных с целью приведения их к одному формату или диапазону. Например, в случае сравнения двух археологических коллекций, количество артефактов в которых отличается на порядок. Обычно для таких сравнений используют арифметический подсчет процентов. Однако на процентные показатели оказывает влияние множество факторов. В статье приводится пример определения конкретного типа сколов (пластинок с изогнутым латеральным профилем), получаемых с нуклеусов определенной технологической направленности (кареноидных). С целью сравнения двух различных категорий артефактов из нескольких выборок подсчет процентов не подходит, поскольку в анализе учитываются только две категории артефактов: пластинки с изогнутым профилем и кареноидные нуклеусы. При этом для подсчета процентов используются все категории в выборках сколов (пластинки с изогнутым, закрученным и прямым профилями) или нуклеусов (призматические, торцовые, кареноидные, плоскостные). Соответственно, на процентную долю одного типа артефактов оказывает влияние количество артефактов других типов. На процентную долю пластинок с изогнутым профилем оказывает влияние количество в выборке пластинок с прямыми и закрученными профилями. Таким образом, данные значительно изменяются, что может оказать влияние на результаты подсчетов и корреляций. В этой ситуации наиболее рациональным является использование z-оценки, или стандартизации данных, которая не влияет на структуру выборки, оставляет все выбросы, однако приводит данные к одному диапазону. Метод z-оценки является перспективным в силу простоты своего применения, а также широкого круга решаемых проблем в археологии, среди которых сравнение разных комплексов и категорий артефактов, особенно включая заметно различающиеся по своему количественному наполнению.

Ключевые слова: археология, математическая статистика, сравнение индустрий и артефактов, z-стандартизация.

K.A.Kolobova

Institute of Archaeology and Ethnography SB RAS
Novosibirsk, Russia
E-mail: kolobovak@yandex.ru

Application of Z-score to Visualization and Correlation of Archaeological Complexes

Standardization is reprocessing of data aimed at bringing it to the same format or range, for example, while comparing two archaeological collections with the number of artifacts differing by an order of magnitude. Arithmetic calculation of percentages is commonly applied to such comparisons. However, percentages are influenced by many factors. This article provides an example of identifying a specific type of spalls (bladelets with curved lateral profile) produced from specific (carinated) cores. Calculation of percentages was not suitable for comparing two different categories of artifacts from several samples because only two categories of artifacts – curved bladelets and carinated cores – were considered in the analysis. For calculating percentages, all categories of spalls (bladelets with curved, twisted, and straight profiles) or cores (prismatic, narrow-faced, carinated, and flat-faced) are used. Accordingly, the percentage of one type of artifacts would be influenced by the number of artifacts of other types. Thus, the percentage of bladelets with curved profiles would be influenced by the number of bladelets with straight and twisted profiles in the sample thereby significantly changing the data and affecting results of counts and correlations. In this situation, the most rational solution would be using z-score or data standardization, which does not affect the sampling structure, leaves all outliers, but brings the data to the same range. The z-score method is promising due to simplicity of its application as well as wide range of problems in archaeology it may solve, such as comparison of different complexes and categories of artifacts, especially including those which markedly differ in their quantitative content.

Keywords: archaeology, mathematical statistics, comparison of industries and artifacts, z-score.

При проведении анализа и описания каменных индустрий исследователям необходимо визуализировать результаты для читателей научных трудов. Различного рода графики вполне пригодны для этой цели, часто в качестве значений переменных выбираются количественные, либо процентные данные. Однако в настоящее время доступны простые статистические инструменты, которые упрощают работу с разноразмерными выборками.

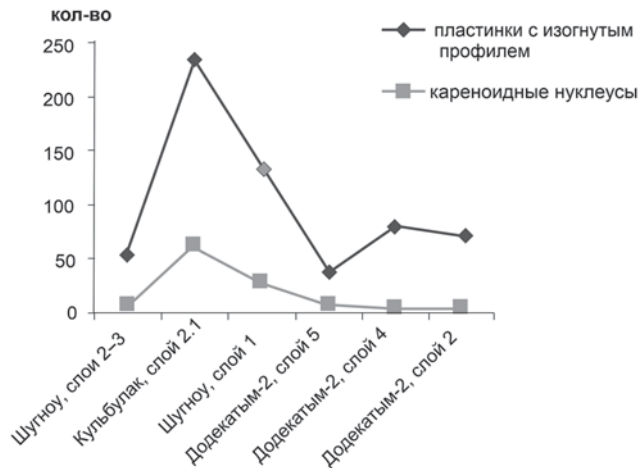
В статье на примере верхнепалеолитических нуклеусов из комплексов западной части Центральной Азии и реализуемых с них сколов демонстрируется применение z-стандартизации, результаты которой сравниваются с количественными данными и с процентными показателями. Кроме того, применяется коэффициент корреляции в совокупности с коэффициентом детерминации с целью определения конкретных сколов, реализуемых с технологически разных ядрищ [Hammer, Harper, Ryan, 2001].

В качестве демонстрационных археологических материалов нами использовались нуклеусы и сколы из комплексов кульбулакской культуры: Кульбулак, Узбекистан, слой 2.1; Шугноу, Таджикистан, слои 2–3, 1; Додекатым-2, Узбекистан, слои 5–2 [Kolobova et al., 2021].

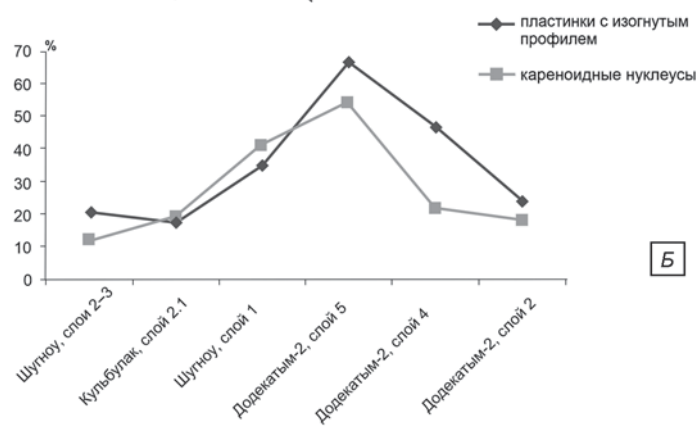
Для кульбулакской верхнепалеолитической культуры в западной части Средней Азии (39–23 кал. тыс. л.н.) характерно мелкопластинчатое производство в рамках кареноидной технологии с призматических и торцовых нуклеусов. В орудийных наборах значительное количество орудий с притупляющей ретушью и геометрических микролитов в форме неравносторонних треугольников. В процессе развития культуры наблюдается синхронное уменьшение количества кареноидных нуклеусов (увеличение количества призматических) и увеличение количества негеометрических и геометрических микролитов, изготовленных на пластинках с прямым профилем. Появление этого технокомплекса в регионе связано с возросшей, по сравнению с предыдущими периодами, мобильностью населения [Ibid.].

Всего было проанализировано 112 кареноидных нуклеусов, 130 призматических нуклеусов для пластинок и 2 551 пластинка с различными типами латеральных профилей (прямые, закрученные и изогнутые).

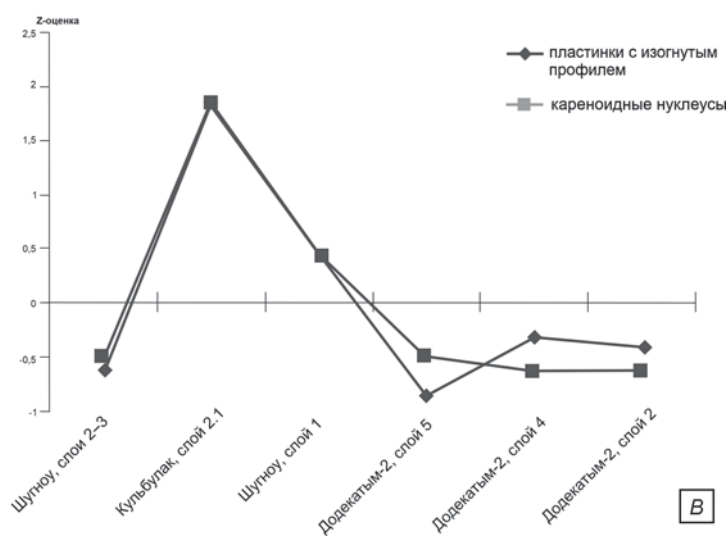
Основной метод, используемый в этой работе, – статистическая z-стандартизация или z-оценка (z-score). Она представляет собой преобразование зна-



А



Б



В

Визуализация количества двух категорий артефактов в нескольких хронологически последовательных комплексах кульбулакской культуры.

А – количественные данные; Б – процентные данные; В – стандартизированные данные.

чений в выборках с разными диапазонами значений. В результате ее применения выборки приводятся к одному диапазону. Преимущество этого вида преобразования данных по сравнению с нормализацией данных (data normalization) заключается в сохранении структуры выборки, в том числе и экстремальных значений (выбросов) и отсутствие ограничительного диапазона

значений. В математической статистике z-оценка используется для подготовки данных с различными диапазонами перед тем, как они используются в мультивариантных анализах [Колобова и др., 2021]. Однако и само по себе это преобразование может быть очень полезным в археологических исследованиях.

Для вычисления z-оценки каждого наблюдения в выборке:

$$X_{\text{станд.}} = \frac{X_i - \mu}{\sigma}, \text{ где}$$

$X_{\text{станд.}}$ – стандартизированный элемент выборки,
 X_i – наблюдение в выборке,
 μ – среднее арифметическое,
 σ – стандартное отклонение по выборке.

Стандартизировать выборки можно в Excel либо в любой статистической программе, например бесплатном программном продукте PAST-4 [Hammer, Harper, Ryan, 2001].

В качестве примера стандартизации данных будет проиллюстрировано соотношение кареноидных нуклеусов и пластинок с изогнутым профилем в комплексах кульбулакской культуры. В научной литературе обычно кареноидные нуклеусы упоминаются в связи с пластинками с закрученным профилем, однако в кульбулакских комплексах была зафиксирована в первую очередь связь с изогнутыми пластинками.

На первом графике представлены количественные данные по соотношению изогнутых пластинок и кареноидных нуклеусов в комплексах, распределенных согласно хронологической последовательности (см. рисунок, А). Как мы видим, количество пластинок с изогнутым профилем значительно превышает количество кареноидных нуклеусов. Глядя на график, мы можем предположить, что количество артефактов этих категорий может коррелироваться.

На втором графике те же данные представлены в виде процентных соотношений (см. рисунок, Б). Как мы видим, максимальные значения графиков значительно сместились, что не соответствует первоначальным значениям. На это обстоятельство оказало влияние значительное количество сколов и нуклеусов других категорий в исследуемых комплексах.

На третьем графике представлена z-оценка все тех же данных (см. рисунок, В). Как следует из графика, данные были приведены к одному диапазону, что позволяет легко проводить их сравнение и дальнейшую корреляцию.

Наборы данных, полученных в результате z-стандартизации, легко применимы к различным тестам математической статистики. Так, в первую очередь, логично применение различных коэффициентов корреляции. Например, коэффициент корреляции Пирсона, примененный к данным на последнем графике, имеет значение $r = 0,96$ ($p = 0,0014$), что демонстрирует очень сильную связь двух переменных. Коэффициент детерминации $r^2 = 0,94$, это означает, что

94 % вариабельности выборок общие. Такие высокие коэффициенты корреляции и детерминации показывают статистически значимую зависимость между переменными, из которой можно сделать вывод, что в комплексах кульбулакской культуры пластинки с изогнутым латеральным профилем в первую очередь производились с кареноидных нуклеусов на протяжении нескольких тысячелетий.

Впоследствии стандартизированные данные могут применяться в мультивариантных анализах, например анализе главных компонент (ГКА) [Kolobova et al., 2019; Колобова и др., 2021].

Продемонстрированный метод визуализации археологических данных является крайне перспективным в силу своей простоты и доступности. С его помощью можно решать следующие задачи для исследования археологических комплексов любых эпох:

1. Корреляция двух или нескольких археологических комплексов между собой. В случае решения такой задачи предлагаемый метод позволит на более высоком методическом уровне проводить сравнения между культурными индустриями одного памятника или среди памятников одного региона, чем это было принято ранее [Рыбин, Колобова, 2004].

2. Сравнение двух и более категорий артефактов из одного или нескольких комплексов. Например, определение конкретных сколов или орудий, изготовленных на них, реализующихся с нуклеусов определенного типа [Хаценович, Рыбин, 2015].

3. Сравнение между комплексами одного или нескольких памятников, количество артефактов в которых значительно отличается [Shunkov et al., 2019; Павленок и др., 2019].

Благодарности

Исследование проведено в рамках проекта НИР ИАЭТ СО РАН № FWZG-2022-0009 «Цифровизация процессов изучения древнейшей и древней истории Евразии».

Список литературы

Колобова К.А., Зоткина Л.В., Маркин С.В., Васильев С.К., Чистяков П.В., Бочарова Е.Н., Харевич А.В. Комплексное изучение персонального украшения из резца сурка в раннеголоценовом комплексе пещеры Каминная (Российский Алтай) // *Stratum plus. Археология и культурная антропология.* – 2021. – № 1. – С. 319–335.

Павленок Г.Д., Анойкин А.А., Бочарова Е.Н., Кулик Н.А., Ульянов В.А. Слой 5.1 верхнепалеолитической стоянки Ушбулак: археология, петрография, планиграфия // *Теория и практика археологических исследований.* – 2019. – № 4. – С. 154–163.

Рыбин Е.П., Колобова К.А. Структура каменных индустрий и функциональные особенности палеолитических памятников Горного Алтая // *Археология, этнография и антропология Евразии.* – 2004. – № 4. – С. 20–34.

Хаценович А.М., Рыбин Е.П. Геометрические изделия в позднем верхнем палеолите Монголии // Проблемы археологии, этнографии, антропологии Сибири и сопредельных территорий. – Новосибирск: Изд-во ИАЭТ СО РАН, 2015. – Т. XXI. – С. 161–165.

Hammer Øyvind, Harper D.A.T., Ryan. P.D. PAST: Paleontological Statistics Software Package for Education and Data Analysis // *Palaeontologia Electronica*. – 2001. – Vol. 4, N 1. – URL: https://palaeo-electronica.org/2001_1/past/past.pdf (дата обращения 25.09.2023).

Kolobova K.A., Kharevich V.M., Kharevich A.V., Fedorchenko A.Y., Bocharova E.N., Krivoschapkin A.I., Olsen J.W., Kurbanov R., Flas D. Archaeological and Experimental Studies of Splintered Pieces in the Central Asian Upper Paleolithic // *Archaeological and Anthropological Sciences*. – 2021. – Т. 13, N 2. – P. 28.

Kolobova K.A., Shalagina A.V., Vasiliev S.V., Markin S.V., Krivoschapkin A.I., Chabai V.P., Krajcarz M.T., Krajcarz M., Rendu W. Exploitation of the Natural Environment by Neanderthals from Chagyrskaya Cave (Altai) // *Quartar*. – 2019. – Vol. 66. – P. 7–31.

Shunkov M.V., Anoikin A.A., Pavlenok G.D., Kharevich V.M., Shalagina A.V., Zotkina L.V., Taimagambetov Zh.K. Nouveau site Paléolithique supérieur ancien au nord de l'Asie Centrale (New Initial Upper Paleolithic site in northern Central Asia) // *L'Anthropologie*. – 2019. – Vol. 123. – P. 438–451.

References

Hammer Øyvind, Harper D.A.T., Ryan. P.D. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 2001. Vol. 4, N 1. URL: https://palaeo-electronica.org/2001_1/past/past.pdf (Accessed: 25.09.2023).

Khatsenovich A.M., Rybin E.P. The Geometric Tools in the Late Upper Paleolithic of Mongolia. In *Problems of*

Archaeology, Ethnography, Anthropology of Siberia and Neighboring Territories. Novosibirsk: IAET SB RAS Publ., 2015. Vol. 21. P. 161–165. (In Russ.).

Kolobova K.A., Kharevich V.M., Kharevich A.V., Fedorchenko A.Y., Bocharova E.N., Krivoschapkin A.I., Olsen J.W., Kurbanov R., Flas D. Archaeological and Experimental Studies of Splintered Pieces in the Central Asian Upper Paleolithic. *Archaeological and Anthropological Sciences*, 2021. Vol. 13, N 2. P. 28.

Kolobova K.A., Shalagina A.V., Vasiliev S.V., Markin S.V., Krivoschapkin A.I., Chabai V.P., Krajcarz M.T., Krajcarz M., Rendu W. Exploitation of the Natural Environment by Neanderthals from Chagyrskaya Cave (Altai). *Quartar*, 2019. Vol. 66. P. 7–31.

Kolobova K.A., Zotkina L.V., Markin S.V., Vasiliev S.K., Chistyakov P.V., Bocharova E.N., Kharevich A.V. Complex Study of a Personal Ornament Made on a Marmot Incisor from the Early Holocene Complex of Kaminnaya Cave (Russian Altai). *Stratum plus. Arkheologiya i kul'turnaya antropologiya*, 2021. N 1. P. 319–335. (In Russ.).

Pavlenok G.D., Anoikin A.A., Bocharova E.N., Kulik N.A., Ulianov V.A. Artificial Construction in the Upper Palaeolithic Layer 5.1 of Ushbulak site: planigraphic Aspect. *Theory and Practice of archaeological Research*, 2019. N 4. P. 154–163. (In Russ.).

Rybin E.P., Kolobova K.A. Structure of stone industries and functional features of paleolithic sites in the Altai Mountains. *Archaeology, Ethnology and Anthropology of Eurasia*, 2004. N 4. P. 20–34. (In Russ.).

Shunkov M.V., Anoikin A.A., Pavlenok G.D., Kharevich V.M., Shalagina A.V., Zotkina L.V., Taimagambetov Zh.K. Nouveau site Paléolithique supérieur ancien au nord de l'Asie Centrale (New Initial Upper Paleolithic site in northern Central Asia). *L'Anthropologie*. 2019. Vol. 123. P. 438–451.

Колобова К.А. <https://orcid.org/0000-0002-5757-3251>