

**И. Г. Ширококов**

## **ВЕРОЯТНОСТНЫЙ ПОДХОД К ОЦЕНКЕ ГИПОТЕЗ И ПРОБЛЕМА МНОЖЕСТВЕННЫХ СРАВНЕНИЙ В КРАНИОЛОГИИ**

**АННОТАЦИЯ.** Проведение попарных сравнений выборок по группе признаков сопряжено с неизбежным риском получения ложноположительных результатов. Однако необходимость применения поправки на множественные сравнения большинством антропологов игнорируется. Нередко исследователи имеют дело с небольшими выборками, и введение поправки настолько снижает вероятность выявления реально существующих различий, что лишает процедуру тестирования смысла. Один из путей решения проблемы заключается в проверке конкретных статистических гипотез о причинах различий между выборками или их неоднородности. Задача сводится к оценке вероятности совпадения списков признаков с некоторыми значениями, реально наблюдаемыми в выборке, и значениями, ожидаемыми при условии справедливости различных гипотез об истоках формирования исследуемого населения. Наблюдаемые значения признаков определяются их внутригрупповой изменчивостью относительно средних величин. Ожидаемые значения признаков — их способностью дифференцировать краниологические выборки, характеристики которых предположительно близки к характеристикам предковых популяций. Для оценки вероятности совпадения списков может быть использовано биномиальное распределение. Анализ краниологической серии из Ижевска (конец XVIII — начало XX в.) свидетельствует о работоспособности предлагаемого подхода. Характеристика выборки в наилучшей степени согласуется с данными письменных источников. В рассматриваемое время состав жителей города формировался преимущественно за счет русского населения, связанного своим происхождением с Волго-Уральским регионом, при участии групп удмуртов и казанских татар.

**КЛЮЧЕВЫЕ СЛОВА:** физическая антропология, краниология, множественные сравнения, внутригрупповая изменчивость, межгрупповая изменчивость

УДК 572.08

DOI 10.31250/2618-8619-2018-1-87-96

ШИРОКОКОВ ИВАН ГРИГОРЬЕВИЧ — к.и.н., н.с. отдела антропологии, Музей антропологии и этнографии им. Петра Великого (Кунсткамера) РАН (Россия, Санкт-Петербург)

E-mail: ivansmith@bk.ru

При проведении большого числа попарных сравнений (средних значений, дисперсий, коэффициентов корреляций) неизбежно растет вероятность совершения ошибки первого рода — решения отклонить нулевую гипотезу в случае, когда она верна. В общем случае нулевая гипотеза предполагает, что реальные различия между характеристиками сравниваемых групп не существуют. Для снижения риска ложноположительного решения об отклонении нулевой гипотезы следует пользоваться специальными поправками (поправка Бонферрони, метод Холма, Бенджамини-Хохберга и др.), подробные описания которых доступны в сети, в том числе в руководствах по статистике для биологов (см. напр.: Шитиков, Розенберг 2013: 62–64). Между тем необходимость использования поправки на множественные сравнения игнорируется в абсолютном большинстве краниологических исследований. Сомнительно, что такая позиция антропологов является сознательной, однако отчасти она может быть оправдана. Поскольку исследователи, как правило, работают с небольшими выборками, применение поправки настолько повышает вероятность ошибки второго рода (принятия нулевой гипотезы в случае, когда она неверна),<sup>1</sup> что лишает смысла саму процедуру тестирования.

В то же время следует признать, что в антропологии не имеет самостоятельного смысла и формальная оценка статистической значимости различий между средними значениями или дисперсиями признаков в сравниваемых выборках. Чаще всего такие сопоставления проводятся между выборочной дисперсией признаков и верхним пределом ее средней величины, характеризующей однородные группы (в соответствии с данными таблиц краниометрических констант из: Алексеев, Дебеч 1964). Однако всякая предполагаемая неоднородность выборки требует объяснения и выдвижения гипотез о ее причинах. Эти гипотезы могут опираться на исторические, археологические, лингвистические и другие сведения, касающиеся проблемы формирования исследуемого населения. Невозможность выдвижения гипотез о причинах неоднородности выборки превращает процедуру проведения статистического анализа в пустую формальность, поскольку сами понятия однородности и неоднородности в краниологии являются условными и контекстозависимыми.

При этом главным условием для проверки выдвигаемых гипотез является возможность выведения из них некоторых следствий, которые, во-первых, отражаются в антропологической характеристике серии, а во-вторых, могут быть сформулированы на языке статистики.

Для статистической оценки неоднородности выборки является крайне важной возможность выделения той группы краниометрических показателей, которые дифференцируют ее предполагаемые предковые группы и, следовательно, по крайней мере, отчасти характеризуются повышенной внутригрупповой изменчивостью. В таком случае задача может быть сведена к проведению сравнительной оценки вероятности совпадения списка признаков с некими реально наблюдаемыми свойствами (например, повышенной внутригрупповой изменчивостью) и списка признаков со свойствами, ожидаемыми при условии справедливости различных гипотез об истоках формирования исследуемого населения (способностью дифференцировать предковые популяции).

В рамках такого подхода каждому краниометрическому признаку присваивается одно из двух значений: «0» или «1». В анализируемой краниологической выборке признаку присваивается значение «0», если его выборочная дисперсия не превышает верхний предел средних величин дисперсии, установленных Алексеевым и Дебечем (1964), и значение «1», если дисперсия признака его превышает. Полученный набор и будет набором признаков с реально наблюдаемыми значениями.

<sup>1</sup> Строго говоря, исследователь приходит к заключению не о том, что нулевая гипотеза истинна, а о том, что для отклонения нулевой гипотезы имеющихся данных недостаточно.

Для определения ожидаемых значений того же набора признаков требуется проведение серии межгрупповых сопоставлений. При этом значение признака определяется его способностью разграничивать предполагаемые предковые группы, изменчивость характеристик которых могла повлиять на характеристику исследуемой выборки. Признаку присваивается значение «0», если он не дифференцирует предковые группы, и значение «1», если его межгрупповая изменчивость имеет существенную величину. Так, например, если для проведения межгрупповых сопоставлений используется канонический дискриминантный анализ, искомое ожидаемое значение признака может определяться величиной коэффициента корреляции между его исходной величиной и первым каноническим вектором: статистически значимым коэффициентом соответствует значение «1», незначимым — значение «0».

Способы определения наборов бинарных признаков с некими реально наблюдаемыми в выборке и ожидаемыми в рамках конкретных гипотез значениями могут быть разными. При их вычислении вовсе необязательно исходить из искусственно установленного порога статистической значимости. При принятии решения об отклонении нулевой гипотезы или решения о том, что для ее отклонения данных недостаточно, исследователи, как правило, отталкиваются от величины  $p$ -значения. Эта величина выражает вероятность получить наблюдаемые и еще более отклоняющиеся значения признака при условии, что нулевая гипотеза верна. Мы часто забываем, что  $p$ -значение не является вероятностью нулевой гипотезы. Различия в его величине, установленные для разных признаков, не позволяют определять, для какого из них вероятность справедливости нулевой гипотезы ниже. Между тем получение статистически незначимых различий ( $p < 0.05$ ) не дает нам оснований полагать, что выборки действительно не различаются между собой (Goodman 2008). Значение  $p$  вообще не позволяет нам оценивать вероятности справедливости нулевой и альтернативных гипотез. «Всякий раз, принимая решение отклонить или принять нулевую гипотезу, мы совершаем акт интеллектуальной смелости. И этот акт является внестатистическим» (Хромов-Борисов 2016: 45–46). По этой причине для определения значений бинарных признаков могут быть использованы несколько способов, в том числе как учитывающие результаты применения  $F$ -критерия, так и не привязанные к значению  $p$  как таковому.

Для оценки гипотез о причинах повышенной изменчивости конкретных признаков в исследуемой выборке мы можем воспользоваться оценкой вероятности ( $P$ ) того, что значения дискретных признаков с вероятностью  $p$  совпадут не менее  $k$  раз из  $n$  возможных (где  $n$  ограничено набором признаков) на внутри- и межгрупповом уровнях сопоставления. Искомая вероятность может быть рассчитана по формуле для биномиального распределения (функция доступна в Microsoft Excel, а пояснения — в сети). Величина  $p$  нам неизвестна, однако условно мы можем принять вероятность совпадения реального и ожидаемого значений признаков равной вероятности их несовпадения, т. е. есть  $p = 0.5$ .

Для корректной оценки вероятностей требуется соблюдение нескольких условий: 1) при определении наблюдаемых и ожидаемых значений рассматривается один и тот же набор признаков; 2) априорные вероятности  $p$  должны быть независимыми; 3) способ определения значений бинарных признаков опирается на один критерий (необязательно общий для набора признаков с ожидаемыми и наблюдаемыми величинами); 4) проводится вероятностная оценка справедливости сразу нескольких альтернативных гипотез, в рамках каждой из которых осуществляется строгий отбор выборок, претендующих на репрезентативность по отношению к популяциям, гипотетически рассматриваемым в качестве предковых групп.

Следует отметить, что даже в случаях, когда историческая интерпретация краниологических материалов не вызывает сомнений, не следует ожидать полного соответствия наблюдаемой и ожидаемой комбинаций значений бинарных признаков. Во-первых, избежать ошибок первого и второго рода при оценке нулевой гипотезы принципиально невозможно. Во-вторых, возможности анализа всегда ограничены материалом. Характеристики привлекаемых к анализу сравнительных серий скорее всего будут несколько отличаться от характеристик предполагаемых предковых групп, а численность анализируемой выборки может оказаться недостаточной для корректного проведения статистического анализа.

Протестировать работоспособность предлагаемого метода логично на примере представительной краниологической выборки, варианты интерпретации характеристики которой могут быть ограничены небольшим числом гипотез, опирающихся на материалы исторических источников. В качестве такого примера рассмотрим выборку мужских черепов из раскопок Троицкого и Михайловского кладбищ города Ижевска. Выборка была получена в ходе археологических работ, проводившихся в течение 2000–2010-х гг. (Карпушкина и др. 2018; Макаров и др. 2014). Датировка выборки укладывается в период конца XVIII — начала XX в.

Известно, что первоначально состав ижевцев формировался за счет различных групп русского населения Волго-Уральского региона, в основном выходцев с уральских заводов, жителей Казанской губернии, Пермского и Вятского наместничеств. С начала XIX в. этнический состав населения Ижевского завода заметно расширился за счет привлечения удмуртов (преимущественно), татар, марийцев, чувашей, а также представителей других народов. В середине XIX в. удмурты и татары составляли около 4% населения Ижевска. К началу XXI в. относительная численность первых возросла до 15%, численность вторых — до 9% (Пислегин, Чураков 2017).

Очевидно, что напрямую проецировать эти данные на краниологический материал невозможно: информация об этнической идентичности ижевцев еще не позволяет судить о соотношении в их составе генетических компонентов разного происхождения. Однако важно, что эти данные позволяют очертить тот круг краниологических серий и тот географический ареал, которые тесно связаны с историей формирования местного населения.

Рассмотрим этот вопрос в рамках трех гипотез. Краниологическая характеристика ижевской выборки в первую очередь отражает ожидаемые результаты: А) смешения различных групп русского населения Волго-Уральского региона; В) смешения групп удмуртов, казанских татар и русских Волго-Уральского региона; С) смешения групп удмуртов, казанских татар и русских различных областей европейской части России и Урала.

Поскольку отечественными краниологами межгрупповые сопоставления проводятся преимущественно по 14 признакам (1, 8, 17, 9, 45, 48, 55, 54, 51, 52, 77, zm, SS:SC, 75(1), для простоты расчетов проведем все анализы в рамках той же программы измерений. Результаты внутригруппового анализа показывают, что повышенной изменчивостью у ижевцев обладает только один признак — назомалярный угол ( $F = 1.76$ ,  $p = 0.0030$ ,  $df_1 = 39$ ,  $df_2 = \infty$ ). Таким образом, при переводе в бинарную систему значения 13 признаков могут быть приравнены «0» и только значение назомалярного угла — «1». Если исходить из требований поправки Бонферрони, полученное  $p$ -значение является достаточным для того, чтобы считать отклонение дисперсии от верхнего предела средних величин статистически значимым ( $\alpha = 0.05/14 = 0.0036$ ), однако для решения поставленной задачи это не так важно.

Результаты межгруппового анализа, проведенного путем расчета расстояний Махаланобиса, говорят о том, что наиболее близкими к ижевцам являются сборные выборки Вятской и Казанской

губерний, а также Пысгора, Гольян, Сарапула и Симбирска. В общем списке близких серий удмурты занимают 23-е, а казанские татары — 25-е место, уступая позиции выборкам русских европейской части России, однако опережая представителей других этнических групп. В целом эти результаты можно было бы трактовать как согласующиеся с любой из трех гипотез. Между тем, поскольку результаты сопоставления определяются набором сравниваемых признаков, очевидно, что даже в первой части списка могут присутствовать серии, сблизившиеся с ижевской выборкой по случайным причинам, не связанным с их реальным генетическим родством. Для установления причин предполагаемой внутригрупповой неоднородности ижевской выборки включение в анализ сравнительных групп, отобранных по принципу хронологического или территориального единства, нельзя считать достаточным условием.

При помощи канонического дискриминантного анализа определим, какие именно признаки обладают наибольшей дифференцирующей способностью для разграничения предполагаемых предковых групп в рамках каждой из трех гипотез.

Для выявления дифференцирующих признаков в рамках гипотезы А были привлечены серии населения Вятской губернии (Алексеев 1969; неопубликованные данные автора), Казани (Алексеев 1969; Газимзянов, Макарова 2012), Свияжска (Макарова 2011), Пысгора (Макарова 2016), Гольян, Сарапула (неопубликованные данные автора) и русских Урала (неопубликованные данные Е. О. Святовой<sup>2</sup>). Для проверки гипотезы В привлекались те же серии, а также серии татар (Алексеев 1971) и удмуртов (Широбоков 2014). Для проверки гипотезы С привлекались те же серии, что и в предыдущем случае, а также выборки русских Новгородской (Пежемский 2000; Евтеев 2015; Евтеев, Олейников 2015), Ярославской (Гончарова 2011), Тверской (Харламова 2012; Васильев, Боруцкая 2013), Псковской (Пежемский 2013), Московской (Дубов, Дубова 2000; Евтеев 2011), Костромской (Васильев, Боруцкая 2004), Курской (Безбородых, Долженко 2014), Смоленской, Пензенской, Витебской, Орловской (Алексеев 1969) губерний.

Признаки, связанные статистически значимой корреляцией ( $p < 0.05$ ) с первым каноническим вектором, условно считались дифференцирующими, а их значения равными «1». Значения остальных признаков приравнивались «0». Общие данные сведены в таблице 1. В каждом из трех случаев была рассчитана вероятность того, что ожидаемые и наблюдаемые значения признаков совпадут не менее  $k$  раз из  $n$ , где  $n = 14$ . Если считать, что исходная вероятность совпадения  $p$  для каждого из признаков равна 0.5,<sup>3</sup> то искомые вероятности составят:  $P_A = 0.09$ ;  $P_B = 0.03$ ;  $P_C = 0.60$ . Задача аналогична расчету вероятности выпадения орла не менее чем в  $k$  случаях при бросании монетки  $n$  раз и  $p = 0.5$ . Так, например, поскольку в рамках гипотезы А наблюдаемые и ожидаемые значения бинарных признаков совпадают в 10 из 14 случаев, то по формуле Бернулли:

$$P_A = \left( \frac{14!}{10!(14-10)!} + \frac{14!}{11!(14-11)!} + \frac{14!}{12!(14-12)!} + \frac{14!}{13!(14-13)!} \right) * \left( \frac{1}{2} \right)^{14} + \left( \frac{1}{2} \right)^{14} = 0.0898.$$

<sup>2</sup> Автор выражает искреннюю признательность Е. О. Святовой за возможность использовать при проведении сравнительных анализов неопубликованные данные.

<sup>3</sup> На самом деле нельзя исключать того, что это не совсем так, поскольку на внутригрупповом уровне у признаков различается число наблюдений, а также средние коэффициенты вариации, но для простоты расчетов мы пренебрежем этими различиями.

## Краниометрическая характеристика ижевской выборки и комбинации признаков, предположительно дифференцирующие ее предковые группы

Номера признаков по Мартину и др.	Мужчины			Признаки, дифференцирующие предковые группы ижевцев в рамках трех разных гипотез		
	n	X	sd	A	B	C
1.	54	179.3	6.1	0	1	0
8.	51	142.3	5.6	0	0	0
17.	47	135.1	5.2	0	0	0
9.	55	96.7	4.8	0	0	1
45.	38	132.9	6.4	1	0	1
48.	33	71.2	3.4	0	0	1
55.	42	51.7	2.5	0	0	1
54.	42	25.4	1.6	0	0	0
51.	45	43.3	1.9	0	0	0
52.	46	32.3	1.9	0	1	0
77.	40	140.1	6.1	0	1	1
Hzm.	35	126.2	5.8	0	0	1
SS:SC	40	45.6	10.5	1	0	1
75 (1).	31	5.8	1.9	1	1	1

См. пояснения в тексте.

Следовательно, поскольку в случае справедливости гипотезы В (при условии репрезентативности исследуемых групп) ожидаемая вероятность совпадения значений признаков наиболее низка, гипотезу В формально следует предпочесть гипотезе А, а вероятность справедливости гипотезы А считать более высокой, чем у гипотезы С. Таким образом, наиболее правдоподобной является гипотеза о том, что антропологический состав ижевцев сформировался в ходе смешения групп удмуртов, казанских татар и русского населения Волго-Уральского региона. Вычисленные расстояния Махаланобиса показывают, что вклад последнего был наиболее значимым. Эти результаты ничего не говорят о возможном присутствии в выборке людей, происхождение которых связано с какими-либо группами населения, проживающими за пределами Волго-Уральского региона. Они лишь свидетельствуют о том, что доля последнего в выборке недостаточно велика для того, чтобы оказывать влияние на ее краниометрическую характеристику. Если быть еще более точным, можно говорить, что вклад таких групп не может быть выявлен без привлечения дополнительных источников. Очевидно, что чем более подробной исторической и археологической информацией мы обладаем о составе исследуемых выборок, тем более детально сформулированные гипотезы об их происхождении мы можем проверять. Так, например, воспользовавшись возможностью хронологического разделения ижевской выборки на раннюю и позднюю подгруппы, мы обнаружим, что ранняя подгруппа ижевцев (1760–1810 гг.), происходящая из раскопок Михайловского кладбища, значительно сильнее отличается от любой из удмуртских выборок, чем поздняя подгруппа из Троицкого кладбища (1810–1928 гг.). Соответствующие средние расстояния Махаланобиса составляют 13.2 и 3.4. Такие результаты хорошо согласуются с историческими сведениями об отсутствии удмуртов среди первых жителей Ижевского завода.

Насколько надежными являются полученные результаты? Приходится признать, что доля изменчивости краниометрических признаков в ижевской выборке, обусловленная включением в ее состав морфологического комплекса, типичного для удмуртов, не очень велика. Так, например, случайным образом разделив выборку на две равные подгруппы, мы обнаружим, что дисперсия признаков ни в одной из них не превышает верхний предел средних значений по Алексееву, Дебелцу (1964). Следовательно, при других обстоятельствах (например, в случае худшей сохранности костей) наблюдаемые значения всех бинарных признаков можно было признать равными «0», а выборку однородной. В рамках предложенного подхода это означает, что если бы ижевская выборка имела меньшую численность, гипотеза А получила бы больше статистических оснований, чем гипотеза В, поскольку в таком случае вероятности совпадения дискретных значений признаков составили бы  $P_A = 0.03$  и  $P_B = 0.08$ .

Проводя снова и снова сопоставление анализа изменчивости признаков для случайным образом извлекаемых из выборки подгрупп черепов, мы можем рассчитать процент случаев, в которых преимущество получает та или иная гипотеза. По результатам извлечения 10 случайных подгрупп гипотеза А получила преимущество в 80% случаев, гипотеза В — в 20%, гипотеза С — в 0%. При этом только в двух случаях повышенной внутригрупповой изменчивостью характеризовался назомаллярный угол, в двух случаях — ширина лица, в трех случаях — оба признака одновременно, и еще в трех случаях — ни один из признаков. Сопоставление всех полученных результатов позволяет прийти к заключению о том, что антропологический состав ижевцев в конце XVIII — начале XX в. формировался преимущественно за счет русского населения Волго-Уральского региона при участии удмуртского и татарского населения. Приведенный выше пример показывает, что если тестируемая гипотеза предполагает незначительную роль какого-то компонента в формировании исследуемого населения, для его выявления может быть использован один из множества методов, позволяющих извлекать из имеющейся выборки множество случайных подгрупп (псевдовыборок) и получать выборочные оценки параметров (Шитиков, Розенбург 2013). Таким образом, для проверки каждой из гипотез может быть не только проведен расчет вероятности совпадения ожидаемой и наблюдаемой в конкретной выборке комбинации бинарных признаков, но и получен диапазон вероятностей, в рамках которых гипотезу можно было бы считать справедливой с учетом случайного фактора анализируемой выборки. Казалось бы, логично было бы не извлекать из выборки подгруппы вручную, а воспользоваться одной из статистических программ с встроенной процедурой бутстрепа, позволяющей выстраивать распределения на основе извлечений нескольких тысяч псевдовыборок. Однако для предложенного способа тестирования гипотез, объясняющих предполагаемую неоднородность выборки, существенной является оценка не только дисперсии каждого признака в отдельности, но и возможной связи между изменением дисперсии одного признака и дисперсии другого.

Подход к определению ожидаемой в рамках той или иной гипотезы комбинации признаков является одним из наиболее важных этапов исследования. К сожалению, канонический дискриминантный анализ не может рассматриваться в качестве универсального метода — возможно, даже и в рассмотренном примере его использование нельзя считать бесспорным. В случаях, когда информация о гипотетических предковых группах ограничена небольшим числом выборок, ожидаемые значения бинарных признаков могут быть определены при помощи *t*-критерия или дисперсионного анализа (ANOVA). Признаку присваивается значение «1», если после проверки равенства его средних величин в предковых группах был получен отрицательный результат, и значение «0» в противоположном случае. Правда, такой анализ требует соблюдения некоторых условий, в частности нормальности распределения исходных признаков и равенства выборочных дисперсий.

Возможны также ситуации, при которых предполагаемые предковые популяции не представлены антропологическими материалами непосредственно и об их характеристиках мы можем судить только по материалам более поздних эпох. В таком случае для установления ожидаемых значений бинарных признаков, вероятно, может быть использован  $U$ -критерий Манна-Уитни. При этом оценка различий проводится между двумя наборами средневывборочных значений признаков, каждый из которых включает локальные варианты группы-потомков одной из двух искомым популяций, характеристики которых нам неизвестны, но различия между которыми предположительно и послужили основной причиной неоднородности исследуемой группы. Применение непараметрического критерия в данном случае опирается на предположение о том, что признаки, величина которых систематически различается в локальных вариантах, разграничивали и предковые популяции.

Может быть предложено множество способов оценки краниометрической характеристики группы. Вероятно, для проверки гипотез о причинах неоднородности исследуемой выборки могут быть также использованы методы, не связанные ни с вычислением вероятностей совпадения наблюдаемых и ожидаемых значений наборов бинарных признаков, ни с оценкой  $p$ -значений. В качестве одного из альтернативных подходов может использоваться вычисление и сравнение величин ранговых коэффициентов Спирмена между разностями наблюдаемых и ожидаемых величин дисперсий и различиями в величине средних значений соответствующих признаков предполагаемых предковых групп. Однако в любом случае, сама необходимость проведения в рамках краниологических исследований статистических процедур, позволяющих оценивать соответствие наблюдаемой характеристики картине, ожидаемой в рамках различных гипотез об истоках происхождения исследуемого населения, сегодня не должна вызывать сомнений. Только в этом случае исследователь может определить, действительно ли полученные им краниологические данные подтверждают справедливость конкретной гипотезы или же они одновременно согласуются с несколькими альтернативными гипотезами и, следовательно, не должны быть использованы для избирательной поддержки одной из них.

#### СПИСОК ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

Алексеев В. П. Происхождение народов Восточной Европы (краниологическое исследование). М.: Наука, 1969. 324 с.

Алексеев В. П. Очерк происхождения тюркских народов Восточной Европы в свете данных краниологии // Вопросы этногенеза тюркоязычных народов Среднего Поволжья. Казань, 1971. С. 232–271.

Алексеев В. П., Дебец Г. Ф. Краниометрия. Методика антропологических исследований. М.: Наука, 1964. 128 с.

Безбородых В. И., Долженко Ю. В. Палеоантропология городских жителей Белгорода XVII–XIX вв. // Физическая антропология: методики, базы данных, научные результаты. СПб.: МАЭ РАН, 2014. С. 104–120.

Васильев С. В., Боруцкая С. Б. Палеоантропология населения Костромской области XVII в.: по материалам могильника Исупово // Расы и народы. М., 2004. Вып. 30. С. 249–267.

Васильев С. В., Боруцкая С. Б. Палеоантропологический анализ погребений из Кашинского Кремля (Воскресенский 1) // Вестник антропологии. 2013. Вып. 3 (25). С. 107–120.

Газимзянов И. Р., Макарова Е. М. К антропологической экспертизе черепов из храма «Во имя святой великомученицы Параскевы Пятницы» // Актуальные вопросы археологии Поволжья. К 65-летию студенческого научного археологического кружка Казанского университета. Казань: Изд-во «ЯЗ», 2012. С. 166–179.



Гончарова Н. Н. Формирование антропологического разнообразия средневековых городов: Ярославль, Дмитров, Коломна // Вестник антропологии. 2011. Вып. 19. С. 202–216.

Дубов А. И., Дубова Н. А. Антропологическая характеристика четырех краниологических серий с территории Москвы // Народы России: от прошлого к настоящему. Антропология. М.: Старый сад, 2000. Ч. 2. С. 130–150.

Евтеев А. А. Краниологические материалы из некрополя с. Козино: предварительные результаты анализа // Актуальные вопросы антропологии. Минск: Беларуская навука, 2011. Вып. 6. С. 474–485.

Евтеев А. А. Палеоантропологические материалы из раскопок Пантелеймонова монастыря // Новгород и Новгородская земля. История и археология: материалы XXIX конф., посвящ. 150-летию Новгородского музея-заповедника. СПб.: Первый ИПХ, 2015. Вып. 29. С. 313–326.

Евтеев А. А., Олейников О. М. Археологические и палеоантропологические исследования на Даныславле улице в Великом Новгороде // Российская археология. 2015. № 1. С. 176–192.

Карпушкина О. А., Бернц В. А., Вязников А. М., Митряков А. Е., Перовошиков С. Е., Пастушенко И. Ю., Сабиров Т. Р., Широбокова Н. Ф., Шишкина Н. Ф., Широбоков И. Г., Емельянова А. Ю., Хайруллина О. Ф. Спасательные раскопки на Троицком (Нагорном) кладбище Ижевска // Археологические открытия в Удмуртии — 2017. Ижевск: Сарапульское полиграфическое предприятие, 2018. С. 12–13.

Макаров Л. Д., Останина Т. И., Перевозчикова С. А., Пастушенко И. Ю., Черных Е. М., Перовошиков С. Е., Бернц В. А., Князева Л. Ф. История города по археологическим данным // Ижевск: история, повседневность, образы. Ижевск: Пешта, 2014. С. 7–25.

Макарова Е. М. Первые поселенцы острова-города Свияжск. К вопросу об антропологическом составе населения // Историко-культурное наследие и современная этнология: материалы конф. молодых ученых. М.: ИЭА РАН, 2011. С. 21–32.

Макарова Е. М. Антропологические материалы из раскопок православного некрополя Пыскорского Спасо-Преображенского монастыря XVI в. Предварительные результаты исследования // XV Бадеровские чтения по археологии Урала и Поволжья: материалы всерос. науч.-практ. конф. Пермь, 2016. С. 208–211.

Пежемский Д. В. Новые материалы по краниологии позднесредневековых новгородцев // Народы России. Антропология. М.: Старый сад, 2000. Ч. 2. С. 95–129.

Пежемский Д. В. Новые краниологические материалы по позднесредневековому населению Пскова // Вестник антропологии. 2013. Вып. 3 (25). С. 121–126.

Пислегин Н. В., Чураков В. С. Первые ижевцы и воткинцы. Население Ижевского и Воткинского заводов по данным III ревизии. Ижевск: Институт компьютерных исследований, 2015. 192 с.

Харламова Н. В. Тверское население XVI–XX веков по данным краниологии // Вестник антропологии. 2012. Вып. 21. С. 49–58.

Хромов-Борисов Н. Н. Гармонизация статистических доказательств и предсказаний в биомедицине // Мы продолжаем традиции российской статистики. Новосибирск: НГУЭУ, 2016. Т. 4. Теоретические поиски и предложения. С. 40–53.

Широбоков И. Г. Материалы к антропологии пермских народов XVII–XIX вв. (краниологические данные) // Ежегодник финно-угорских исследований. 2014. Вып. 3. С. 80–98.

Шитиков В. К., Розенберг Г. С. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R. Тольятти: Кассандра, 2013. 314 с.

Goodman S. A Dirty Dozen: Twelve p-value Misconceptions // Seminar in Hematology. 2008. Vol. 45. P. 135–140.

## PROBABILISTIC APPROACH TO EVALUATING HYPOTHESES AND MULTIPLE COMPARISONS PROBLEM IN CRANIOLOGY

**ABSTRACT.** Pairwise comparisons of samples are associated with the inevitable risk of obtaining false-positive results. However, most anthropologists ignore the importance of the correction for multiple comparisons. Often, researchers deal with small samples, and the multiple testing correction reduces the probability of identifying real differences enough to make the procedure meaningless. One way to solve this problem is to test specific statistical hypotheses about the reasons for the differences between the samples or their heterogeneity. The overall objective is to estimate the likelihood that the list of trait meanings actually observed in the sample will be consistent with the list of trait meanings expected on the assumption that specific hypothesis about the origins of the population is true. The observed meanings of traits are determined by their intra-group variability in comparison with the average values, typical for a homogenous groups. The expected meanings of traits determined by their ability to differentiate cranial samples from ancestral populations or populations close to them. The binomial distribution can be used to estimate the probability that meanings of traits will coincide in certain cases. The analysis of the craniological series from Izhevsk (late 18<sup>th</sup> — early 20<sup>th</sup> century) demonstrates the productiveness of the approach. The sample characteristics best correspond to the data of written sources. The population of the town was formed mainly at the expense of the Russian people from the Volga-Ural Region, with the participation of Udmurts and Kazan Tatars.

**KEYWORDS:** physical anthropology, craniology, multiple comparisons, intra-group variability, inter-group variability

IVAN G. SHIROBOKOV — Candidate of Historical Sciences, Peter the Great Museum of Anthropology and Ethnography (Kunstkamera) of the Russian Academy of Sciences (Russia, Saint Petersburg)  
E-mail: ivansmith@bk.ru