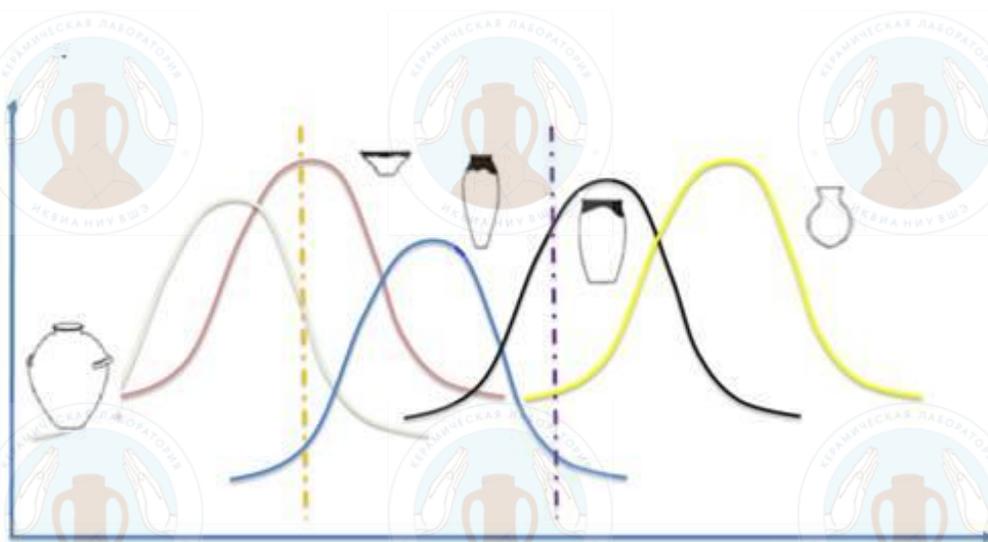




Керамическая лаборатория ИКВИА НИУ ВШЭ

Лекции, семинары, практикумы Статистико-комбинаторные методы в археологии



лекция А.Н.Васильева
(ЦАВАРХ ИКВИА НИУ ВШЭ)



4 марта 2024

Watermarkly



КУРС РАЗРАБОТАН В РАМКАХ РЕАЛИЗАЦИИ ГРАНТ-ПРОЕКТА НАУЧНОГО ФОНДА НИУ ВШЭ № 24-00-048

«МЕЖДИСЦИПЛИНАРНОЕ ИССЛЕДОВАНИЕ ТОРГОВО-ЭКОНОМИЧЕСКИХ СВЯЗЕЙ АНТИЧНЫХ ГОРОДОВ СЕВЕРНОГО ПРИЧЕРНОМОРЬЯ И КОЛХИДЫ В III ВВ. ДО.Н.Э.-II В.Н.Э., НА ПРИМЕРЕ АМФОР КОЛХИДСКОГО КРУГА.»



СОДЕРЖАНИЕ:

- ❖ **История возникновения статистико-комбинаторных методов в археологии**
- ❖ **Основные понятия статистики в археологии**
- ❖ **Примеры применения статистики в исследованиях**
- ❖ **Ограничения метода статистики**



Статистика is...

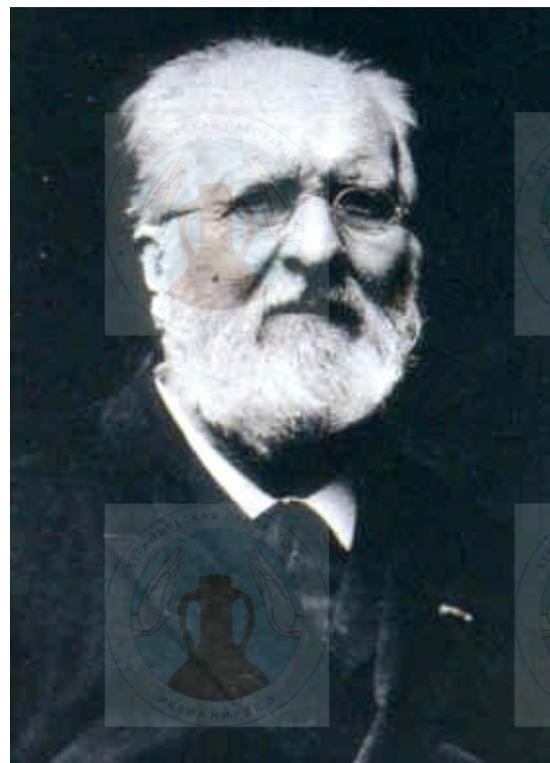
Статистика – наука о работе с большими объёмами данных

- Впервые понятие ввёл Готфрид Ахенвалль в сер. XVIII в. (в контексте работы со сведениями о налогах и сборах для государственных служащих)
- В археологии статистика появилась в 1860-х гг. в работах, связанных с созданием *типологий*
- В истории статистики в археологии можно выделить важный хронологический рубеж – появление доступных для «гражданской» науки ЭВМ – 1960-е гг.



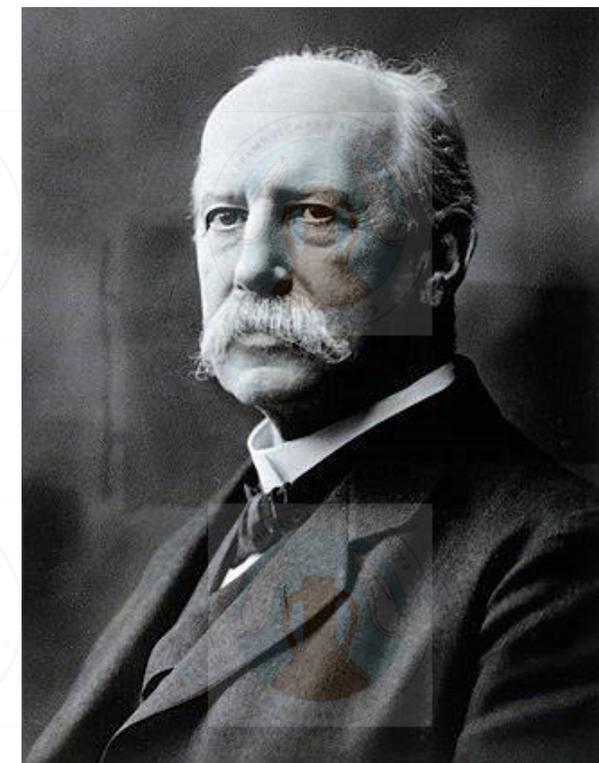
➤ В археологии статистика впервые появилась в создании концепции типологического ряда

➤ Применялся как количественный, так и комбинированный – количественный и качественный (объединяющий предметы в группы по тому или иному признаку) подход



Габриель де Мортилье
(1821-1898)

- Создатель классификации палеолитических индустрий

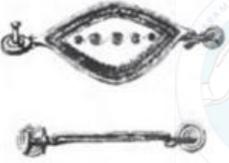
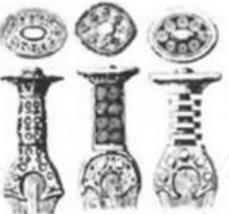
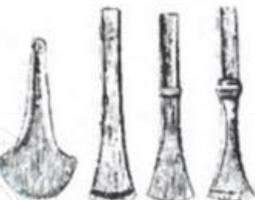
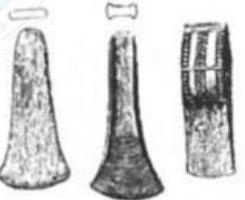


Оскар Монтелиус (1843-1921)

- Автор типологического метода в археологии (на материалах эпохи бронзы РЖВ Северной Европы)

➤ В археологии статистика впервые появилась в создании концепции типологического ряда

➤ Применялся как количественный, так и комбинированный – количественный и качественный (объединяющий предметы в группы по тому или иному признаку) подход

	FIBULAE	SWORDS AND DAGGERS	AXES
Period III			
Period II			
Period I			

■ Типологические ряды О. Монтелиуса были основаны не только на внешних морфологических признаках и стратиграфии, но и на количественных (число тех или иных предметов) или качественных (предметов с тем или иным признаком) характеристиках

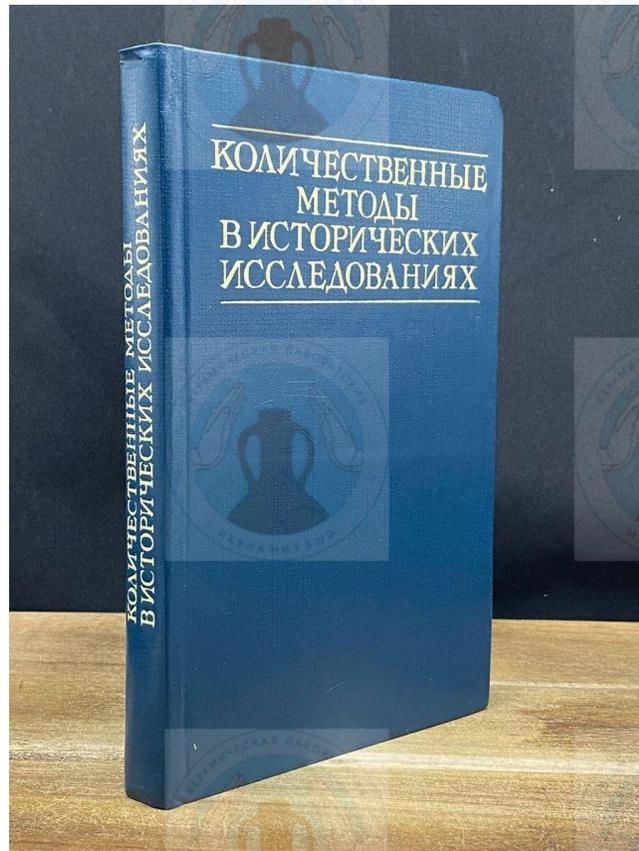
Информационная революция и археология

- С 1960-х гг., с распространением ЭВМ и появлением первых компьютеров, возможности статистического анализа данных в гуманитарных науках и в археологии сильно возросли
- 1960-е -1970-е подъём интереса к статистико-комбинаторным методам в истории и археологии

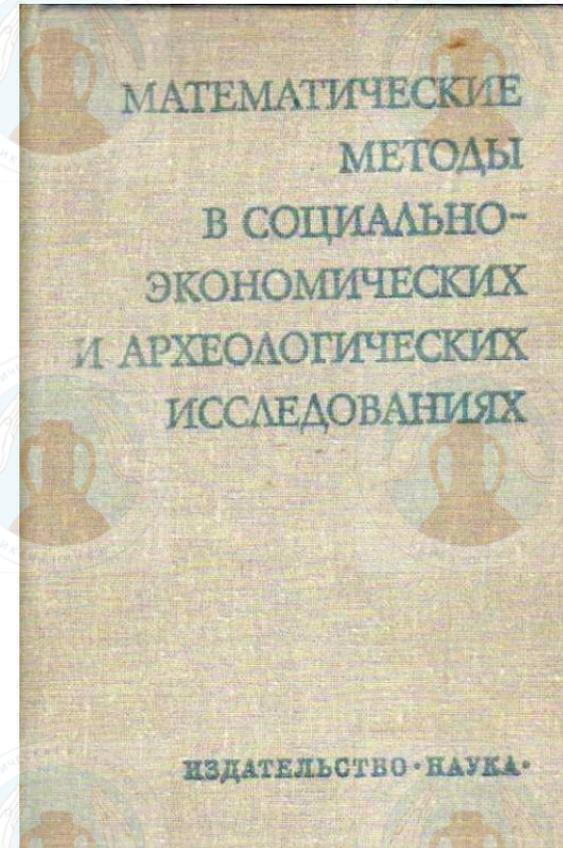


Советская ЭВМ Урал-1, 1959 г.

Информационная революция и археология



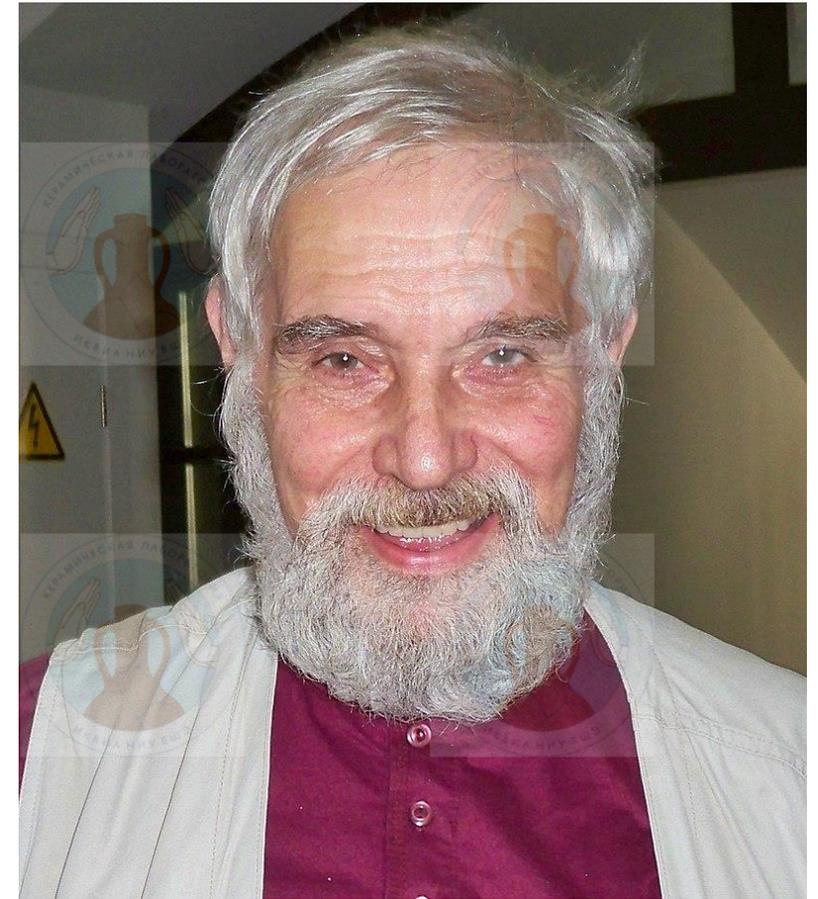
«Количественные методы в исторических исследованиях»
М.: 1984



«Количественные методы в социально-экономических исследованиях» М.: 1981

Дега Витальевич Деопик

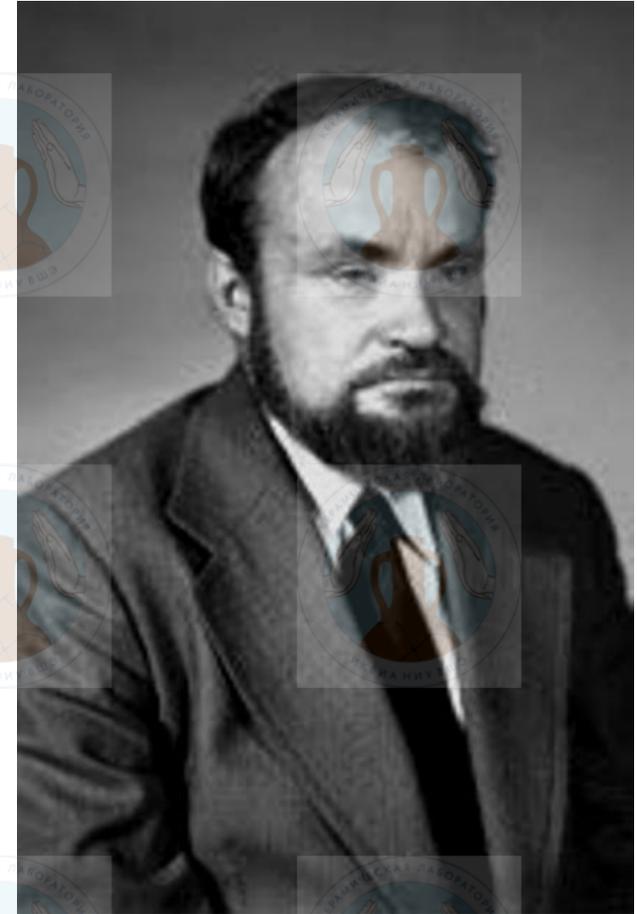
- пионер в разработке статистико-комбинаторных методов в поселенческой археологии в СССР (материалы античных городов Причерноморья, позже – Вьетнама Юго-Восточной Азии)
- Ввёл в научный оборот для археологов *теорию вероятностей*



Дега Витальевич Деопик (род. 1932)

Герман Алексеевич Фёдоров-Давыдов

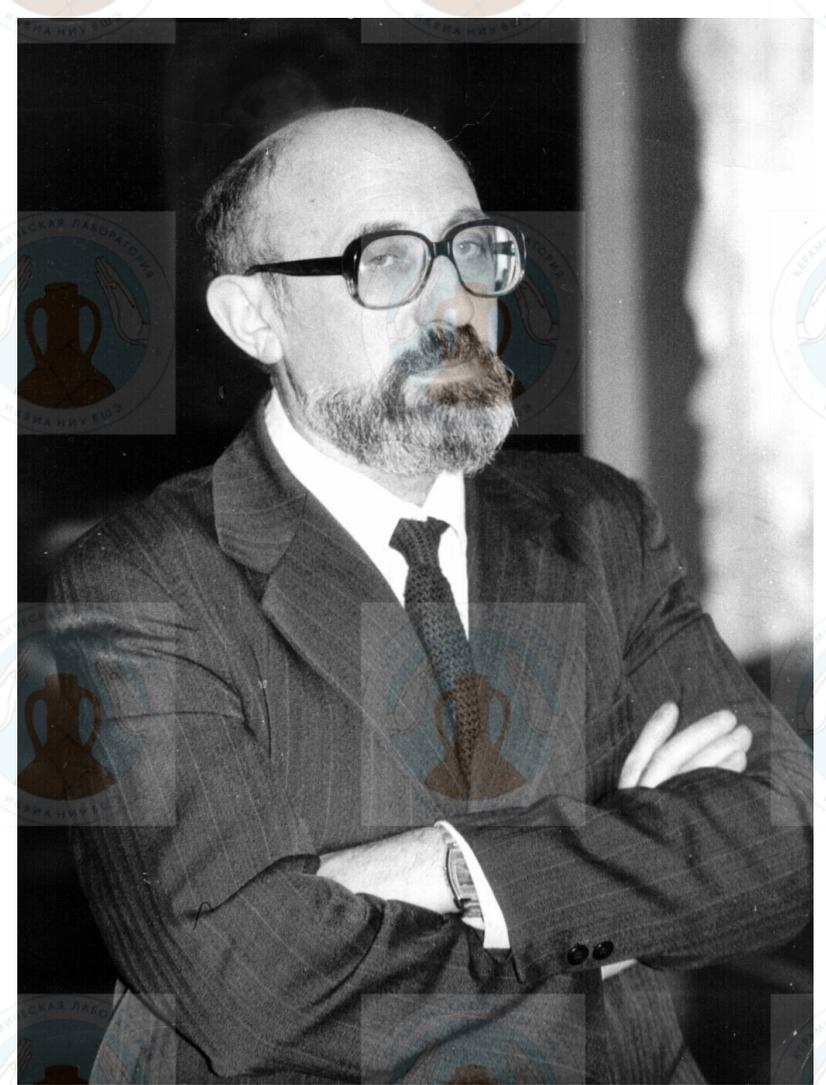
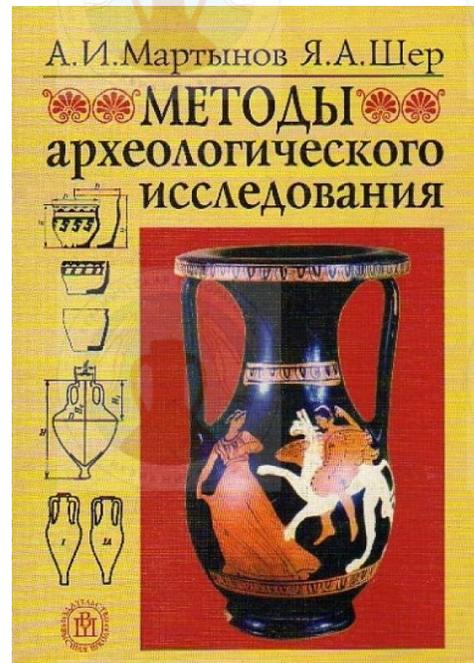
- Специалист по археологии Поволжья и Средней Азии эпохи Средневековья
- *Применял теорию вероятностей для создания типологий*



Герман Алексеевич Фёдоров-Давыдов
(1931-2000)

Яков Абрамович Шер

- *Использовал теорию вероятностей при составлении типологий и датировок петроглифов*
- *Автор учебных пособий по статистическим методам в археологии*



Яков Абрамович Шер
(1931-2019)

Теория вероятностей и археология

- Теория вероятностей – раздел математики, изучающий случайные события, случайные величины свойства и операции над ними
- *Особенность применения теории в археологии – далеко не всегда у нас есть понятная система мер, к которой можно апеллировать – её мы определяем сами для каждого конкретного исследования*

-> Не всегда просто сопоставлять количественные археологические исследования между собой, даже если они посвящены одной группе памятников/артефактов

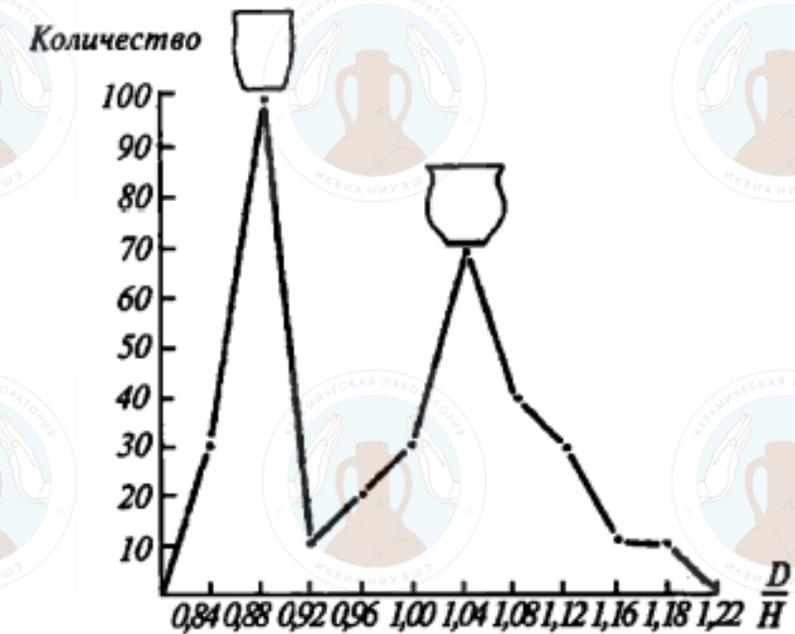
Основные понятия

- **Среднее арифметическое** – сумма всех значений множества чисел, делёная на их количество
- **Среднее медианное** – число, которое находится в середине ряда чисел, построенных в порядке возрастания от самого большого к самому малому.
- **Мода** – одно или несколько значений во множестве, встречающееся наиболее часто (мода = типичность)
- **Корреляция** – взаимовстречаемость, статистическая взаимосвязь двух и более величин. В археологии различаем корреляцию **качественных** и **количественных** признаков

Основные понятия

- **Частота (относительная частота) встречаемости** – доля объектов с тем или иным признаком от общей их совокупности
- **Нормальное распределение** – непрерывное распределение вероятностей с пиком в центре и симметричными боковыми сторонами

Зачем нам это всё надо?



Пример влияния критериев показателей

2 могильника одной и той же культурной общности

Сравниваем комплекс инвентаря из погребения – считаем количество каждого вида предмета, выводим средние значения, подмечаем корреляции

В могильнике А в погребении – в среднем 7 наконечников; те погребения, где $12 >$ – имеют сопроводительные остатки конской сбруи

В могильнике Б – в среднем, по 3 наконечника и погребения с конской сбруей есть только с числом предметов $17 >$

...при этом в некрополе А считалась медианная, и был разброс от 2 до 14, а в некрополе Б – арифметическая, и разброс от 2 до 17

При том, что оба некрополя были похожи, разница в критерии (подсчёте средних показателей) привела к тому, что в могильнике Б была гораздо сильнее выражена количественная дифференциация инвентаря

Если бы при этом в некрополях были бы погребения, не содержащие в себе стрел вообще и мы бы учли их в общем подсчёте, они также могли сильно исказить статистику – хотя их логичнее считать отдельно (иной культурный/ социальный/ гендерный маркер) – если проигнорировать контекст, мы рискуем получить *статистические аномалии*



Статистические аномалии

- любые числовые отклонения от средних значений в множестве
- могут иметь «естественное» происхождение: выделяющийся своими характеристиками среди прочих аналогичных ему типологически предмет/участок и т. д.
- чаще связаны с ошибками в выборке (разные культурные признаки, попавшие в один статистический ряд)
- могут также объясняться обстоятельствами сбора материала (разграбленное погребение, откуда пропал весь металл, оказалось включено в выборку вместе с непо потревоженными)



Пример аномалии на графике



Выборка

- ключевое понятие археологической статистики: совокупность элементов множества, которые охватываются исследованием
- от корректности сформулированных условий напрямую будет зависеть результат, который мы получим в исследовании
- археолог должен учитывать не только особенности материалов памятника, с которым он работает, но и особенности выполнения работ и все прочие факторы, возможно влиявшие на итоговый состав предметов

Особенности выборки на городском материале Танаиса (по Д. В. Деопику)

Распределение материала в 43 000 фрагментов керамики по выборкам, каждая из которых – фрагменты из одного штыка одного квадрата => учёт *территориального и хронологического* признака в распределении, по мысли автора, позволяет экстраполировать данные на всю западную часть городища (*репрезентативность выборки*)

Определение *степени случайности выборок*, так как, будучи неслучайными и, следовательно, подчиненными каким-то закономерностям отбора, выборки могли бы показать *закономерности отбора*, а не общие закономерности распределения интересующих нас видов керамики во множестве (совокупности) фрагментов, которое оценивается при помощи выборок

– поскольку территория городища раскапывалась сплошными участками, сохраняющими на своём протяжении всю совокупность слоёв греческого и римского времени, автор полагает, что данные репрезентативны и достаточно случайны (т. е. не будет аномалий), чтобы можно было говорить об обнаруженных корреляциях как о **биномиальном распределении**

Биномиальное распределение

– распределение количества «успехов» в последовательности из n независимых случайных экспериментов, таких, что вероятность «успеха» в каждом из них постоянна и равна p

Иначе говоря, если мы проведём серию шурфовок в случайных местах городища, в каждом слое мы с той же вероятностью (P), что и в контрольной группе (нашей выборке) найдём интересующий нас тип керамики – чем больше таких шурфовок (n), тем выше эта вероятность

Искажения «эффекта малых величин»

Число профильных частей сосуда (донце, ручки и венчик), остающихся от каждого конкретного экземпляра, сравнительно невелико – если мы попытаемся построить **кривую распределения** по слоям только на их основе, мы рискуем не заметить искажения, вызванные **эффектом малых величин** – случайно улетевший в отвал венчик, перемещённая мышью ручка амфоры могут сильно снизить их итоговый процент распространения в слое

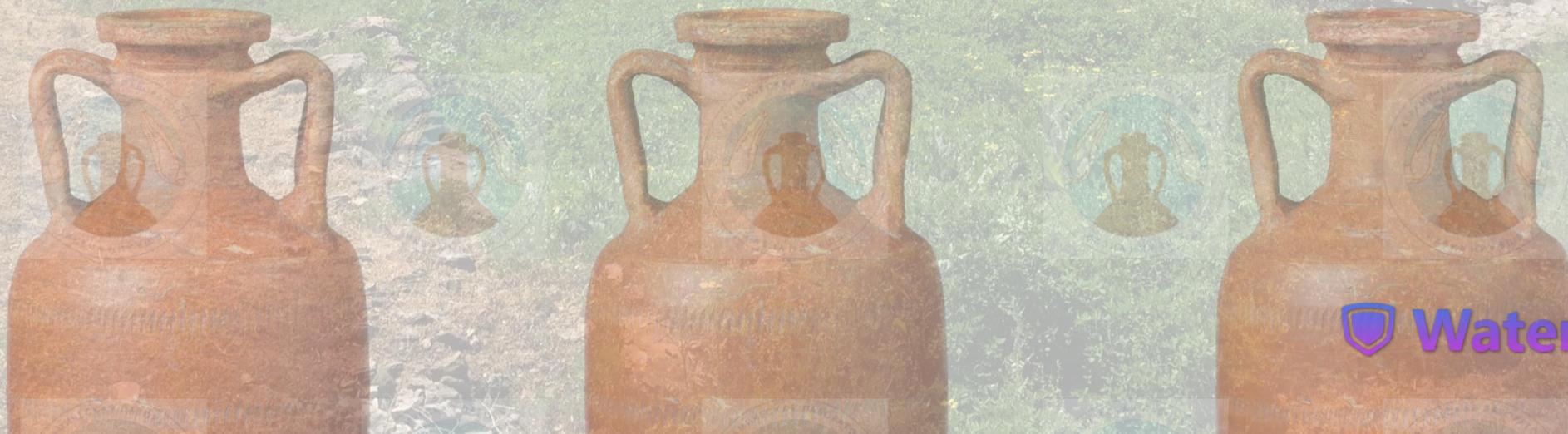
Как этого избежать?

Взять массовый материал, который бы нивелировал эффект малых величин – т. е., *стенки сосудов*

Искажения «эффекта малых величин»

Малая информативность единицы учета – обломка стенки сосуда – перекрывается большей достоверностью (в силу массовости материала) полученных цифр и, наоборот, большая информативность профильных частей сосудов частично обесценивается малочисленностью самих предметов

– таким образом, при расчёте процента встречаемости того или иного типа керамики в слоях учитывались не отдельные профильные и непрофильные части, а был проведён независимый расчёт по каждой из них, после чего был сделан сводный график со средним арифметическим % встречаемости



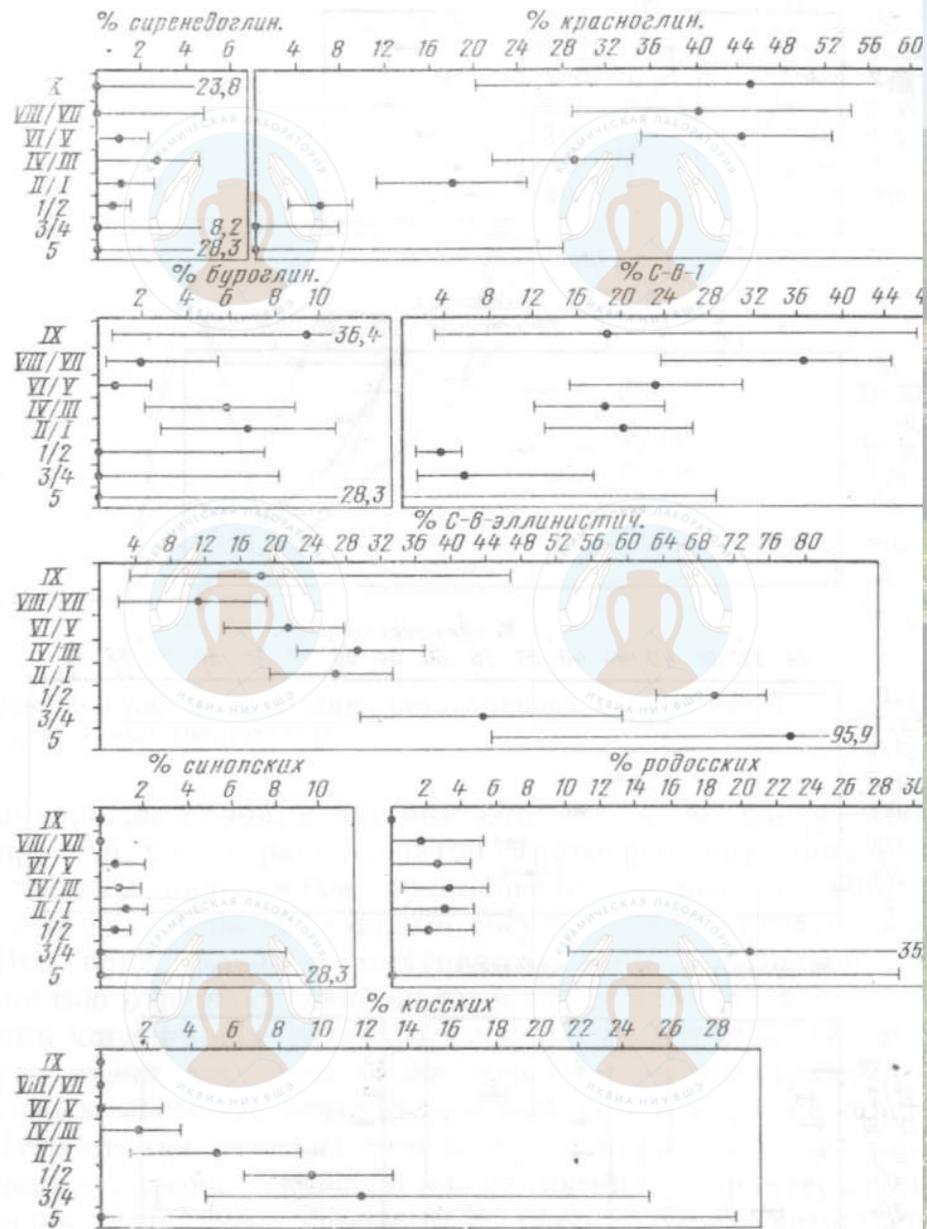


Рис. 3. Веники амфор

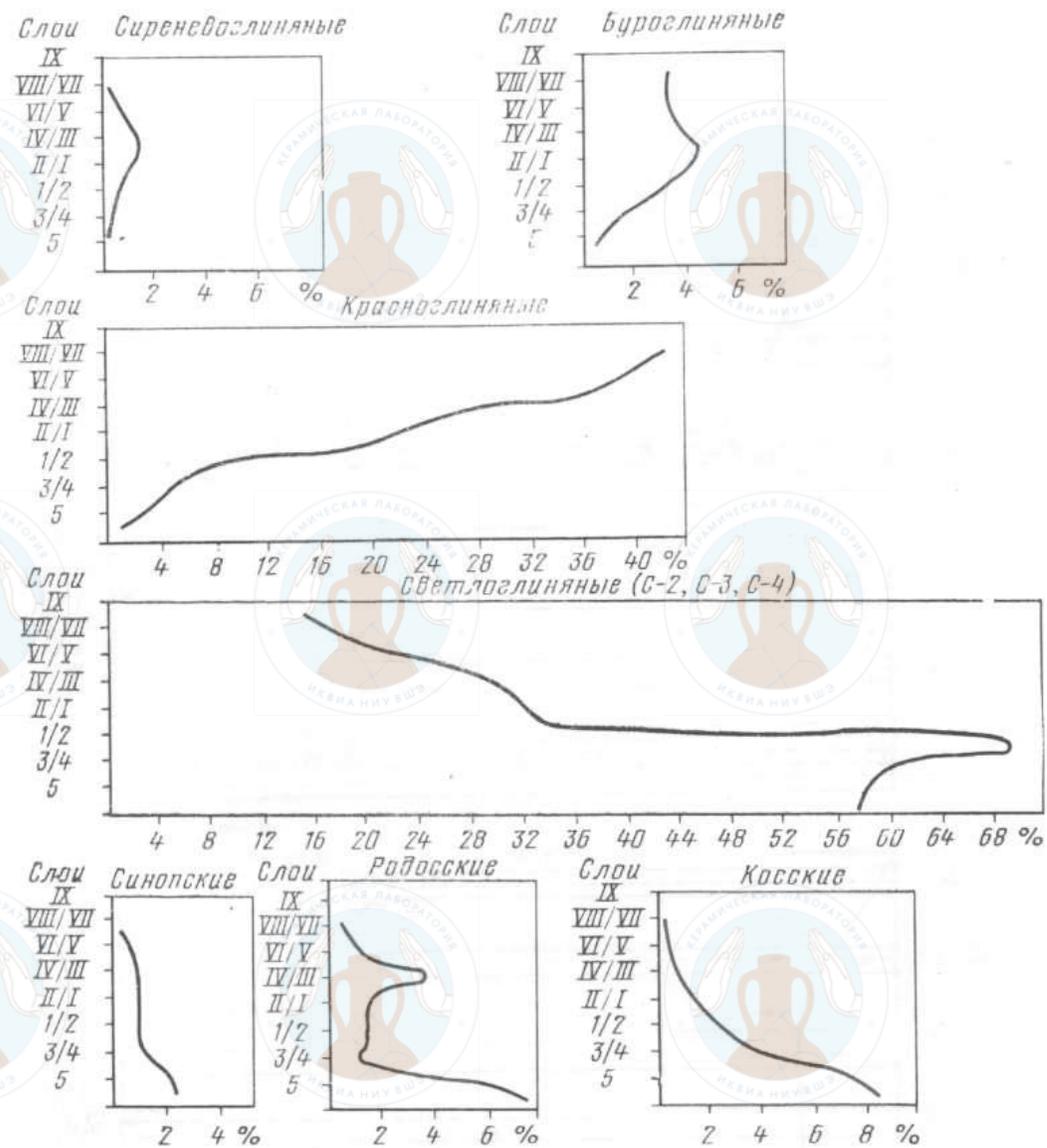


Рис. 5. Итоговые кривые эволюции доли УРФ в слоях

Watermark



Группировка материала

Поскольку исследование проводилось на многослойном памятнике, затрагивающем длительное время бытования сосудов различного происхождения и формы, Д. В. Деопик ввёл особое понятие – *устойчивую разновидность формы*, включающую в себя как качественные признаки – форму профильных частей (*коды*), так и количественные – толщину стенок, предположительный объём тары и т. д.

По сути, выделение УРФ – создание альтернативной классификации амфорных сосудов, за счёт сочетания количественных и качественных признаков позволяющее, во-первых, включать в классификацию непрофильные части, во-вторых, дающее возможность за счёт изменения количественных показателей *кодов* проследить динамику эволюции того или иного типа амфор.

- При этом за счёт привязок к обычной типологии амфор и соотнесения с абсолютной хронологией слоёв памятника появилась возможность выявить чисто эллинистические, чисто римские и «общие» УРФ (по времени бытования).

% смещений

- В слое в зависимости от его перемешанности имеется определенное количество более раннего и более позднего материала (следствие перекопов и неизбежных дефектов раскопок в многослойных памятниках)
- Появление раннего материала в позднем слое более вероятно, чем обратное явление
- Выделив среди УРФ чисто эллинистические и чисто римские, можно выделить % положительного смещения (снизу вверх) и отрицательного смещения (сверху вниз) – всё это для чистоты гипотезы следует проделывать отдельно для каждого слоя и с каждым типом фрагмента керамики
 - в западной части Танаиса % (общий, арифметическая средняя) отрицательного смещения составил 1,2 %, положительного 10 % (= «коэффициент переотложенности»)



Дополнительные выводы: «цикл импорта»

Кривые распространённости УРФ по слоям представляют собой не время бытования того или иного типа «в вакууме», а связаны с распространением и бытованием УРФ в месте производства, его торговом обращении и бытовом использовании (пока он не попадёт в слой) – *«циклы производства»* и *«циклы бытования»*

Но!

Если учесть характер памятника (торговый эмпорий) кривая искажается изменениями степени интенсивности торговых контактов между центром производства и изучаемым поселением; для импортера эти изменения образуют **«цикл импорта»** - единственный достоверно нами фиксируемый в слое

правда, и он тоже искажен зависимостью от изменения численности других типов амфор (последняя зависит от шпроты и интенсивности торговых связей в целом и искажает общую картину для каждого конкретного типа)



Доверительный интервал

интервал, в пределах которого с заданной вероятностью лежат выборочные оценки статистических характеристик выборки.

$$CI = \bar{x} \pm z \frac{\sigma}{\sqrt{n}}$$

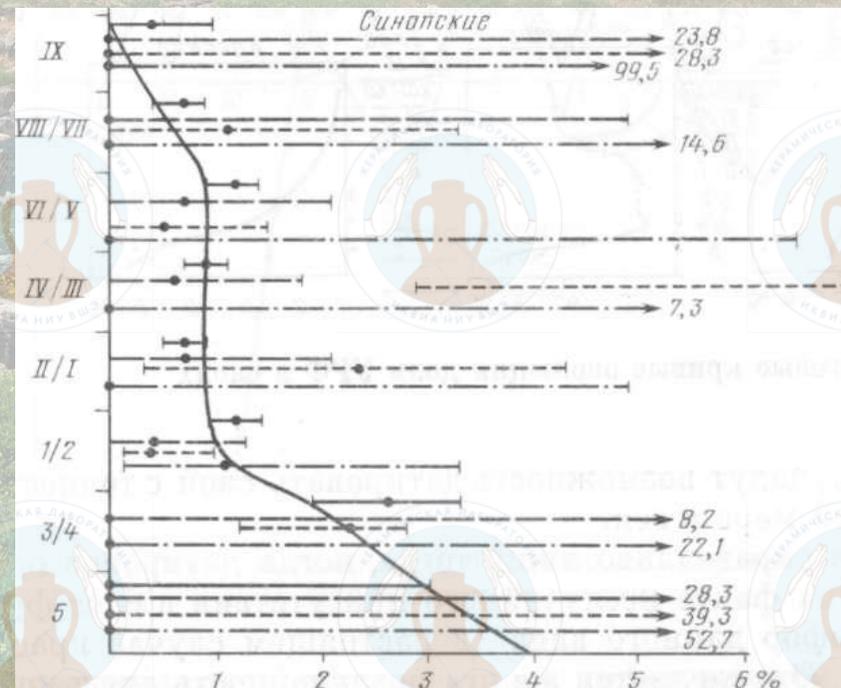
CI = доверительный интервал

σ = стандартное отклонение выборки

\bar{x} = выборочное среднее

n = размер выборки

z = значение уровня достоверности



Наиболее часто используемые интервалы – 95% и 99%; Обычно обозначается буквой Σ (σ)

«Условно-целые амфоры»

Если провести серию экспериментов: изготовить некоторое количество амфор, соответствующих древним образцам по физическим параметрам, а потом разбить – и посчитать среднее медианное распределения числа осколков, можно пересчитывать всю массу керамического материала в «условно-целые единицы» амфор.

Практические результаты

- Были получены кривые распределения УРФ по слоям, на основании чего можно сделать выводы о динамике торговли Танаиса с другими полисами греческого мира
- За счёт установления групп чисто эллинистических и чисто римских УРФ была доказана стратиграфическая граница между временем существования жилых построек и мусорным заполнением руин тарой римских типов (ок. кон. I в. до н. э. – нач. I н. э.)
- За счёт уточнения хронологии УРФ появилась возможность уже датировать аналогичные слои в варварских памятниках на периферии Танаиса, содержащих в себе амфорную тару



Практические результаты

- Были получены кривые распределения УРФ по слоям, на основании чего можно сделать выводы о динамике торговли Танаиса с другими полисами греческого мира
- За счёт установления групп чисто эллинистических и чисто римских УРФ была доказана стратиграфическая граница между временем существования жилых построек и мусорным заполнением руин тарой римских типов (ок. кон. I в. до н. э. – нач. I н. э.)
- За счёт уточнения хронологии УРФ появилась возможность уже датировать аналогичные слои в варварских памятниках на периферии Танаиса, содержащих в себе амфорную тару



Г. Е. Фёдоров-Давыдов и типология средневековых бусин

Цель: составить типологию бус, установив чёткие критерии перво- и второстепенных признаков

Параметров: количество отверстий, материалы, техника, форма, пропорции, цвет, прозрачность, наибольший размер

– не все признаки сочетаемы между собой (например, прозрачность и материал - серебро)



Энтропия

Энтропия — величина, которая измеряет степень неопределенности или неопределенности случайной величины

$$H(I) = - \sum_{i=1}^{m(I)} p(I)_i \lg p(I)_i$$

$p(I)_i$ - частота (вероятность) бус с i -м значением признака I ; $m(I)$ - число значений признака I

Энтропия считается отдельно для каждого признака, после чего возможно оценить уровень связей между каждыми отдельно взятыми признаками (здесь – I и II):

$$U(I/II) = U(II/I) = H(I) + H(II) - H(I, II).$$



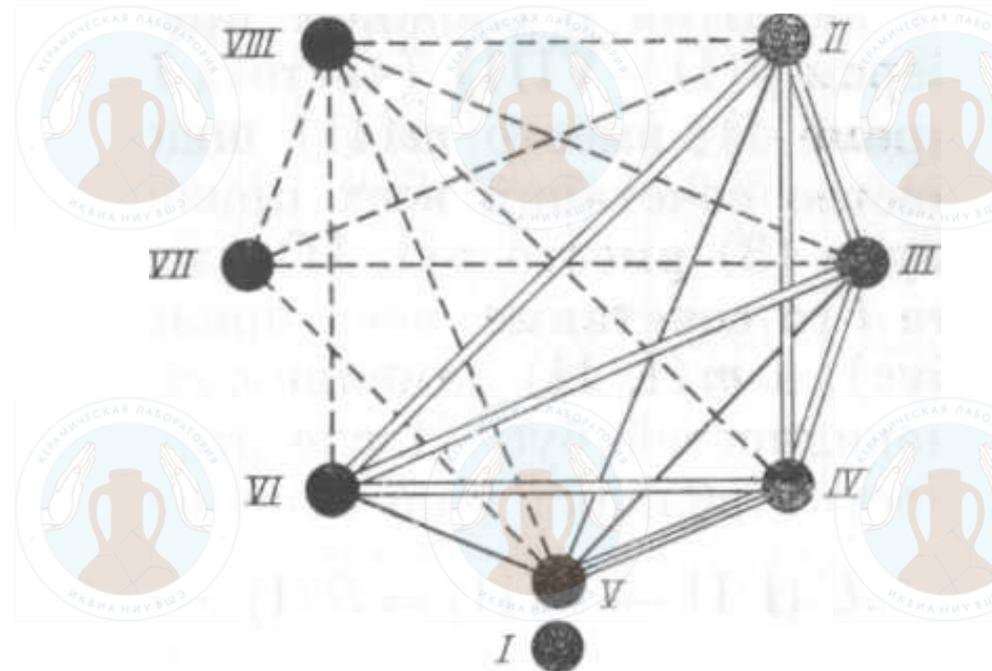
Структура признаков

На основе полученных значений энтропии и связей между признаками становится возможным построить граф:

Самый неинформативный признак – I

есть корреляция между признаками II, III, IV, V, VI

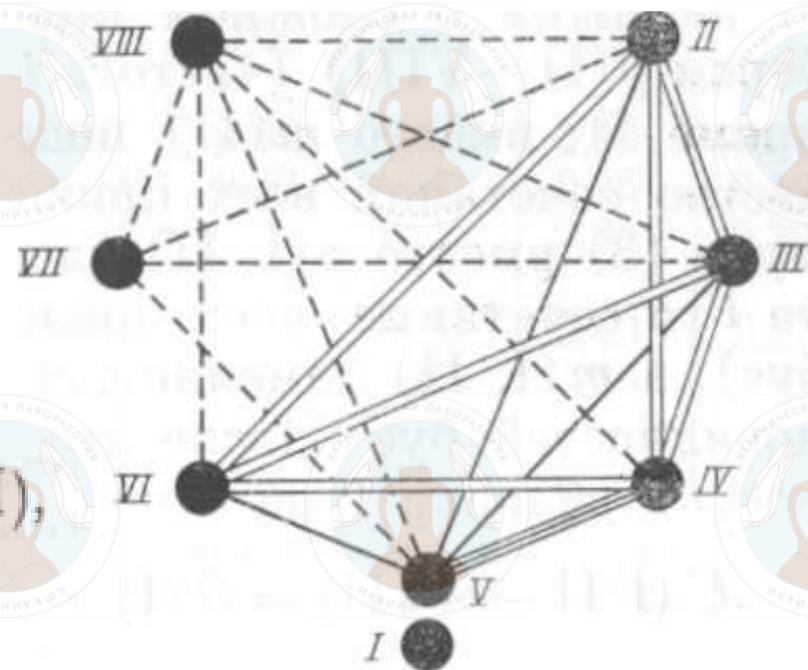
Признаки IV и V образуют особый кластер



Коэффициент информативности

С учётом полученной сети связей (графа) становится возможным посчитать коэффициент информативности для каждого признака по отношению к остальным (здесь – для признака I)

$$U(I/II - VIII) = \mathcal{E}(I) + \mathcal{E}(II - VIII) - \mathcal{E}(I - VIII),$$



После этого становится возможным определить меру зависимости его от других признаков (= нормированную информативность признака) (здесь – для признака I)

$$Q(I) = \frac{U(I/II - VIII)}{\mathcal{E}(I - VIII)}$$

U – информативность, Э – энтропия, Q – информативность (зависимость) признака

Энтропия

Степень информативности

Признаки

Высокая (0,27 – 0,64)

Техника, форма, цвет

Малая информативность (0,01-0,22)

Материал, прозрачность, размер

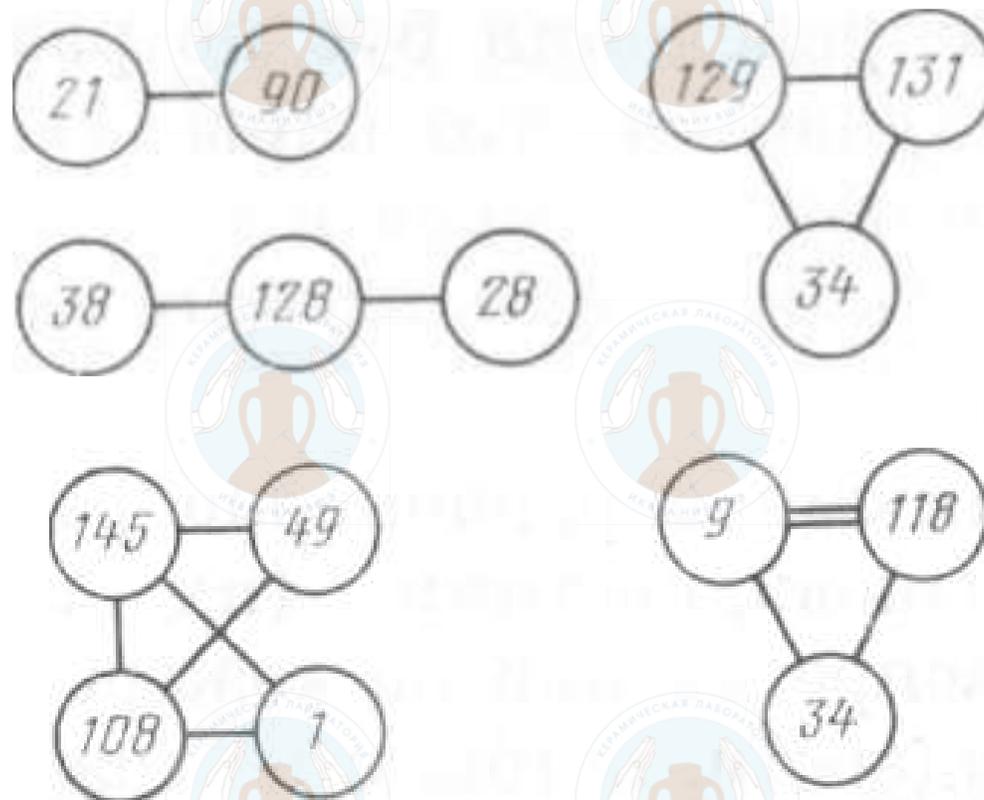


Распределение бус по ожерельям

Принцип подбора бусин в ожерелья, найденные в погребениях, возможно исследовать, воспользовавшись **коэффициентом сходства:**

$$K = \frac{m}{n} \sum \sqrt{P_i Q_i}$$

m — число бус в одном ожерелье; n — число бус в другом; информативности каждого Σ -признака
 P — доля бус i -й разновидности в одном ожерелье; q — доля бус i -й разновидности в другом ожерелье

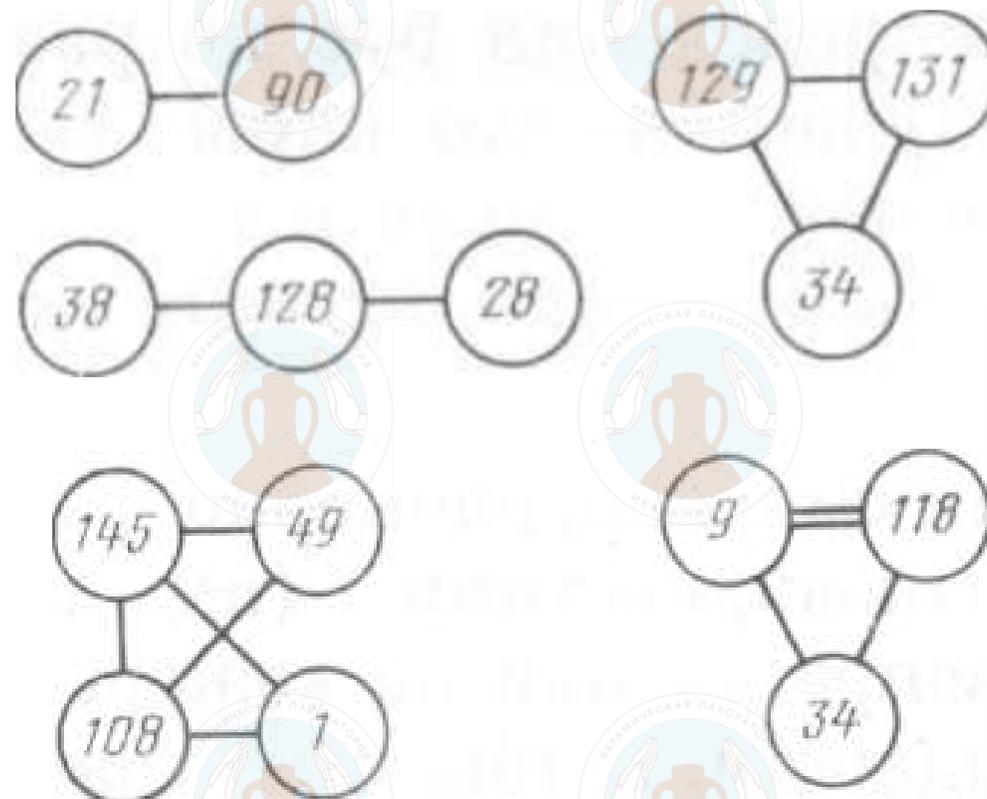


Распределение бус по ожерельям

Используя значения коэффициента сходства, возможно построить сеть связей между бусами в ожерельях для каждого из 8 исследуемых могильников

Большинство массовых разновидностей бус распределено по погребениям так, что все бусы одного вида присутствуют только в одном ожерелье (погребении). Следовательно, все ожерелья не могут группироваться по сходству своего состава бус с какими бы то ни было другими ожерельями

Наиболее частный признак бусин, по которым они сгруппированы в ожерелья – цвет



Социальные и политические реконструкции в истории и археологии

Table VIII. Ideological position of sovereign

	Mythical charter	'Divine' status	Genealogies	Relationship to supernatural forces	Acquisition of sacral status	Ritual upon ruler's death	Human sacrifices	Ritual duties	Taboos
Angkor	X	S	X	M	I	G	0	R	X
Ankole	X	S	—	M	0	A	0	R	0
Axum	X	S	X	M	0	G	0	R	X
Aztecs	X	S	X	M	I	G	X	R	X
China	X	S	X	M	I	G	X	H	X
Egypt	X	S	X	0	I	G	X	H	X
France	X	S	X	M	I	G	—	R	X
Hawaii	X	S	X	M	C	A	X	R	X
Iberia	0	0	0	0	0	0	0	R	0
Inca	X	S	X	M	I	G	X	R	X
Jimma	—	B	—	—	—	0	—	R	X
Kachari	0	S	0	0	0	0	X	R	0
Kuba	X	S	X	M	I	A	X	H	X
Maurya	X	S	X	M	I	0	—	0	0
Mongolia	X	S	X	0	0	0	0	0	0
Norway	0	S	X	M	0	G	—	R	0
Scythia	X	S	X	0	I	G	X	R	0
Tahiti	X	S	X	M	C	A	X	R	X
Volta	X	S	X	M	I	A	—	0	0
Yoruba	X	S	X	M	I	G	X	H	X
Zande	X	0	X	—	I	G	0	R	0

Abbreviations: S = sacral ruler; B = non-divine human being; M = middleman; I = only through ritual inauguration; C = sacral already before inauguration; G = mourning only; A = ritual anarchy; R = performs rites; H = ruler is high priest as well.

- Возможны, но следует подходить осторожно
- Высока доля субъективной оценки исследователя как при определении критериев сравнения, так и при измерении «величин»

- Велик риск «эффекта накопления погрешностей»

О роли субъективного

- Там, где начинается историческая интерпретация материала, статистика бессильна
- Необходимо не злоупотреблять и всегда учитывать долю не поддающегося типологизации материала – «серого пятна»

В 1960-е гг., во время т. н. «второй дискуссии о варяжском вопросе» Д. А. Авдусин представил в качестве доказательства незначительности доли скандинавов в Гнёздово % скандинавских погребений *от общего числа* – 3,7%

В ответ на это Л. С. Клейн представил % скандинавских

Некоторые выводы

«Один показатель – не показатель, один результат – не результат»

– Л. С. Клейн, «Время в археологии»

- Помним о несопоставимости результатов, полученных разным методическим путём
- Ищем и стараемся понять причину возникающих статистических аномалий – это позволит избежать ошибок
- Помним об ошибках, вызванных «эффектом малых величин» и «эффектом накопления погрешностей» (корректируется использованием доверительного интервала)



Работы по статистической обработке массового материала городища Джанкент. Аул Джанкент, лето 2023 г.

Watermarkly

**ВСЕ ИЗОБРАЖЕНИЯ, ИСПОЛЬЗОВАННЫЕ В
ПРЕЗЕНТАЦИИ, НАХОДЯТСЯ В ОТКРЫТОМ ДОСТУПЕ.**

**ФОТОГРАФИЯ РАБОТ В ЭКСПЕДИЦИИ «ДЖАНКЕНТ»
СДЕЛАНА ИКРАМОВОЙ А.**

**НА ФОТО: ВАСИЛЬЕВ А.Н, ГЕРАСИМОВА В.В., ИВАНОВ
К.А.**

**ТАБЛИЦЫ И СХЕМЫ ВЗЯТЫ ИЗ РАБОТ
Г. Е. ФЁДОРОВА-ДАВЫДОВА, Д.В.ДЕОПИКА, Я.Б.ШЕРА
И О. МОНТЕЛИУСА. ПУБЛИКАЦИИ НАХОДЯТСЯ В
ОТКРЫТОМ ДОСТУПЕ.**

СПАСИБО ЗА ВНИМАНИЕ



СПИСОК ЛИТЕРАТУРЫ

1. Деопик Д. В. Керамический комплекс и культурный слой // Математические методы в социально-экономических и археологических исследованиях. – М.: «Наука», 1981. – С. 222-266.
2. Фёдоров-Давыдов Г. Е. Археологическая типология и процесс типобразования (На примере средневековых бус) // Математические методы в социально-экономических и археологических исследованиях. – М.: «Наука», 1981. – С. 267-317.
3. Квирквелия О. Р. Краткий обзор советской литературы по вопросам применения статистико-математических методов исследования в археологии // Математические методы в социально-экономических и археологических исследованиях. – М.: «Наука», 1981. – С. 318-334.
4. Мартынов А. И., Шер Я. А. Методы археологического исследования. / 2-е изд, испр. и доп. – М.: Высш. шк., 2002. – 240 с.