

О ПРИМЕНЕНИИ СЕКВЕНЦИАЛЬНЫХ КРИТЕРИЕВ ПРИ ОПРЕДЕЛЕНИИ ИСКОПАЕМЫХ БРАХИОПОД

М. РУБЕЛЬ

При определении ископаемых брахиопод исследователь нередко имеет дело с малочисленными выборками или же многочисленная выборка сравнивается им с известным видом, охарактеризованным несовершенно. Это затрудняет идентификацию данной выборки на основании статистических приемов.

В настоящей статье рассматриваются вопросы определения ископаемых брахиопод при помощи секвенциальных критериев, применение которых позволяет уменьшить необходимое для видовой идентификации количество экземпляров в принципе в два раза по сравнению с другими статистическими приемами.

Материалом для примера такого определения служат экземпляры *Linoporella*, вымытые из 50 кг пробы глинистой породы биогерма Хиллисте западной Эстонии (тамсалуский горизонт, лландовери). Автор признателен Л. Выханду (Тартуский госуниверситет) за ценные замечания и оказанную помощь в написании данной статьи.

В диагностике видов рода *Linoporella* Schuchert et Cooper 1931 (Linoporellidae, Enteletacea) наиболее существенным признаком является очертание раковины, ее форма. Для сравнения отдельных видов по названным диагностическим признакам целесообразно их выразить при помощи отношений следующих стандартных размеров (рис. 1): x_1 — длина брюшной створки, x_2 — длина спинной створки, x_3 — ширина арея, x_4 — толщина раковины, x_5 — наибольшая ширина раковины.

Все отношения приводятся по x_5 , т. е. x_i/x_5 , где $i = 1, 2, 3, 4$.

Отметим, что абсолютные размеры раковины данного вида могут принимать значения в большом интервале. Правильное представление для определения получается только на основании наибольших раковин, являющихся предполагаемо взрослыми, тогда как вычисление средних и дисперсии какого-либо признака по всем экземплярам выборки может дать искаженное представление, так как в различных выборках количество молодых особей может быть неодинаковым. Вообще, абсолютными размерами можно пользоваться только в тех случаях, когда мы уверены в статистической однородности сравниваемых выборок по данным признакам.

Пользуясь при определении видов отношениями абсолютных размеров, в большинстве случаев мы не имеем дело с постоянными показателями, так как и отношения зависят от относительного возраста (абсолютных размеров) раковины (рис. 2).

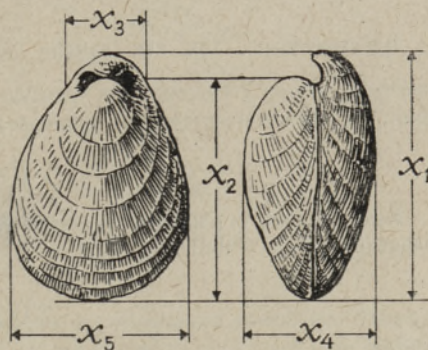


Рис. 1. Обозначение размеров раковины *Linoporella*.

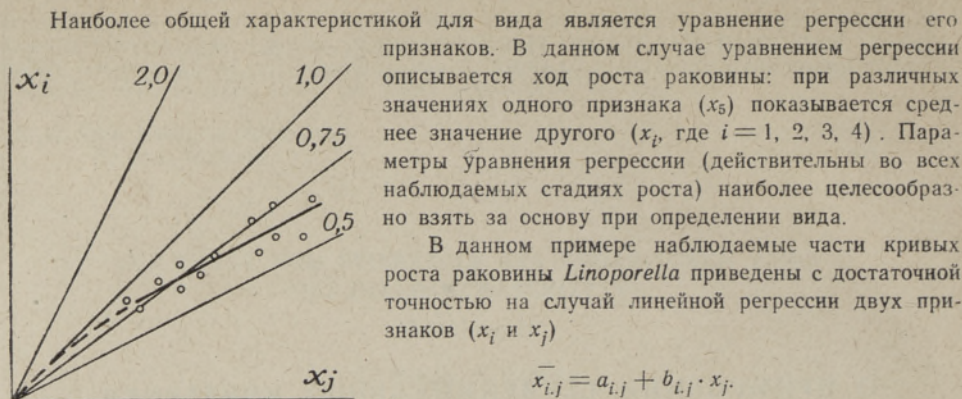


Рис. 2. Отношения (кружки) и линия регрессии (в наблюдаемой части — прямолинейное — непрерывная линия).

Для сравнения отдельных экземпляров с параметрами уравнения регрессии известной выборки (вида) выражаем один из параметров, $b_{i,j}$, при зафиксированном значении другого, $a_{i,j}$,

$$\bar{x}_{i,j} = a_{i,j} + b_{i,j} \cdot x_j.$$

$$b_{i,j} = \frac{\bar{x}_{i,j} - a_{i,j}}{x_j}$$

и заменяем $\bar{x}_{i,j}$, обозначая новую переменную $b_{(i,j)}$

$$\frac{x_i - a_{i,j}}{x_j} = b_{(i,j)}. \quad (*)$$

Покажем, что переменная $b_{(i,j)}$ имеет среднее $b_{i,j}$ или

$$\Sigma (b_{(i,j)} - b_{i,j}) = 0.$$

Действительно, так как замена переменной $\frac{x_i - a_{i,j}}{x_j} = b_{(i,j)}$ значит, что мы сдвинули общую регрессионную прямую вниз, чтобы она прошла через начало координат, то уравнение регрессии приобретает вид

$$x'_i = b_{i,j} x_j.$$

Коэффициент $b_{i,j}$ найдем в случае применения метода наименьших квадратов из уравнения (Weber, 1961, стр. 241)

$$\Sigma (x'_i - b_{i,j} x_j) x_j = 0.$$

При предположении $x_j \neq 0$ найдем сразу, что

$$\Sigma \left(\frac{x'_i}{x_j} - b_{i,j} \right) = 0$$

или, принимая во внимание отношение (*),

$$\Sigma (b_{(i,j)} - b_{i,j}) = 0.$$

Переменное $b_{(i,j)}$ имеет дисперсию

$$s_0^2 = \frac{\Sigma (b_{(i,j)} - b_{i,j})^2}{n - 2}.$$

Доверительные пределы для некоторого будущего $(n+1)$ -го $b_{(i,j)}$, совместимого с измерениями данной выборки, найдем из уравнения (по Налимову, 1960, стр. 172)

$$b_{i,j} - b_{(i,j)(n+1)} = t \cdot s_0 \sqrt{\frac{n+1}{n}},$$

причем для нахождения значения t из соответствующих таблиц (например, Налимов, 1960, табл. 3 в приложении) возьмем уровень значимости 0,05 при $f = n - 1$.

На основании величин s_0^2 , $b_{i,j} \pm t \cdot s_0 \sqrt{\frac{n+1}{n}}$ определяем и двухсторонние секвенциальные критерии (см. Weber, 1961, стр. 362; Налимов, 1960, стр. 182) для сравнения выборок графическим путем.

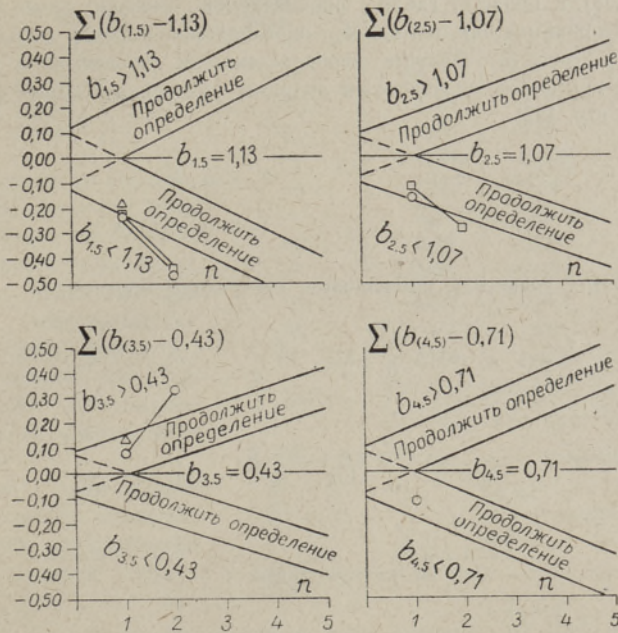


Рис. 3. Секвенциальные критерии для *Linoporella cf. punctata* по $b_{1.5}$, $b_{2.5}$, $b_{3.5}$, $b_{4.5}$ со сравнением с двумя видами: \circ — *Linoporella punctostriata*, Δ — *L. punctata* с о-ва Готланд, \square — *L. punctata* из Англии.

Пример применения секвенциальных критериев по отдельным признакам приведен на рис. 3. Для нанесения данных измерений экземпляров из определяемой выборки на график преобразуем соответствующие размеры по формуле

$$b_{(i,j)} = \frac{x_i - a_{i,j}}{x_j},$$

где $b_{(i,j)}$ — оцененный коэффициент регрессии для сравнения (определения), $a_{i,j}$ — параметр по известной популяции, x_i и x_j — соответствующие абсолютные размеры по определяемым экземплярам.

Сравнение проводится по отдельным значениям $b_{(i,j)}$, причем после измерения и сравнения каждого определяемого экземпляра (из определяемой выборки) с известной выборкой (принимающей здесь характеристикой популяции какого-либо вида) по секвенциальным критериям делается одно из трех следующих заключений:

(1) проверяемая гипотеза о принадлежности данной выборки к известной популяции по данному признаку принимается,

(2) эта гипотеза отвергается или

(3) определение надо продолжить.

По оси абсцисс соответствующего графика (рис. 3) отложено число последовательных экземпляров (n), а по оси ординат накопленные суммы результатов последовательных преобразований по формуле

$$\frac{x_i - a_{i,j}}{x_j} - b_{i,j}$$

На графике двумя парами линий ограничена область неопределенности (продолжить определение) и области, где $b_{i,j}$ определяемой выборки равно известной $b_{i,j}$ (идентификация принимается) или где $b_{i,j}$ определяемой выборки меньше или больше от известной $b_{i,j}$ (рис. 3). Область неопределенности, зависящей в большой мере от изменчивости признаков, определяется линиями, уравнения которых следующие (по Налимову, 1960, стр. 182):

$$\begin{aligned} T_0 &= a_0 + b \cdot n & T'_0 &= a'_0 + b' \cdot n \\ T_1 &= a_1 + b \cdot n & T'_1 &= a'_1 + b' \cdot n \end{aligned}$$

Параметры этих уравнений вычисляются следующим образом:

$$a_0 = 2,3 \frac{s_0^2}{\delta} \cdot \log \frac{\beta}{1 - \frac{\alpha}{2}} \quad a'_0 = -a_0$$

$$a_1 = 2,3 \frac{s_0^2}{\delta} \cdot \log \frac{1 - \beta}{\frac{\alpha}{2}} \quad a'_1 = -a_1$$

$$b = 1/2 \cdot \delta \quad b' = -1/2 \cdot \delta,$$

причем $\delta = t_{(0,05)} \cdot s_0 \sqrt{\frac{n+1}{n}}$.

Определяемые экземпляры должны представлять собой выборку из одного вида, из одного обнажения и одного слоя. Соблюдение этих условий обеспечит наибольшую информацию для оценки определения вида. Иными словами, на основании предложенного здесь приема сравниваются выборки из естественных популяций (см. Imbrie, 1956). При этом, особенно для получения настоящего заключения о данных экземплярах (выборке) относительно известной популяции (вида), остается возможность соединить несколько выборок, например малочисленные выборки из близких обнажений одной и той же пачки, если между такими выборками по отдельным признакам (их коэффициентам регрессии) нет статистически существенных различий.

Исходя из идентификации выборок на уровне популяции, конечная оценка на уровне вида решается на основании трех вышеназванных заключений следующим образом. По заключению (1) можно говорить об идентичности до известной популяции (по сравниваемым признакам), т. е. несомненная идентичность на уровне вида. По заключению (2) различие заслуживает доверия только на уровне популяции, причем при двухсторонних критериях показывается и относительный знак (больше или меньше сравниваемого) коэффициента регрессии. Оценка на уровне вида при заключении (2) должна быть проведена из общих представлений на объем вида. По заключению (3), если даже все имеющиеся экземпляры исчерпаны, выборка остается неопределенной до новых дополнительных выборок с того же места.

Приведенные секвенциальные критерии по известной выборке действительны в наблюдаемом промежутке роста, т. е. они действительны для сравнения тех экземпляров, размеры которых не меньше или не больше наименьших или соответственно наибольших известных экземпляров.

Вопрос о необходимом количестве экземпляров выборки для вычисления секвенциальных критериев оцениваем в данном случае на основании достоверности среднего. Вычисление вероятности (P) о нерепрезентативности среднего ($b_{i,j}$) на основании распределения Стюдента

Таблица 1

Статистики изучаемой выборки
Linoporella cf. punctata

Пары признаков	x_1/x_5	x_2/x_5	x_3/x_5	x_4/x_5
Статистики				
n	10	11	11	10
$b_{i,j}$	1,13	1,07	0,43	0,71
$a_{i,j}$	-0,70	-1,05	1,76	-0,39
s_0	0,081	0,059	0,056	0,069
δ	0,20	0,14	0,13	0,17
v_b	7,17%	5,51%	13,02%	9,72%
m_b	0,02	0,02	0,02	0,03
P	<0,001	<0,001	<0,001	<0,001
a_0	-0,098	-0,074	-0,071	-0,084
a_1	0,120	0,091	0,087	0,103
b	0,098	0,070	0,065	0,085
ω	5,1—13,7 5,3—12,1	4,5—12,7 5,3—12,1	4,1—7,3 5,3—10,3	3,2—8,9 5,3—12,1

$$t = \frac{b_{i,j}}{m_b}$$

$$m_b = \sqrt{\frac{\sum (b_{(i,j)} - b_{i,j})^2}{n(n-1)}}$$

показывает, что такая гипотеза имеет в данном случае $P < 0,001$ (табл. 1). Поэтому мы считаем данную выборку *Linoporella cf. punctata* еще пригодной для вычисления секвенциальных критериев. Также оценка $b_{i,j}$ как коэффициента регрессии утверждает его достоверное значение или по крайней мере $b_{i,j} \neq 0$.

Примером определения проведем сравнение многочисленной выборки по *Linoporella*

из биогерма Хиллисте (табл. 1) с двумя видами из трех местонахождений (табл. 2).

Примечание. v_b — коэффициент вариации $b_{(i,j)}$;
 ω — размах варьирования в мм.

Исходные данные по двум видам *Linoporella*
для сравнения в мм

Таблица 2

Признаки	x_1	x_2	x_5	x_3	x_4	Примечания
Вид						
1. <i>Linoporella punctostriata</i>	14,5 22,5 —	— — 13,1	16,3 25,0 15,3	— — 9,6	— — 8,6	Hall, 1852, табл. 52, фиг. 5а, d, e, f; Северная Америка.
2. <i>L. punctata</i>	7,5 22,0	~7,5 ~22,0	9,0 25,0	— —	— —	Whittard and Barker, 1950, стр. 580; Англия.
3. <i>L. punctata</i>	10,6	—	12,0	8,6	—	Schuchert and Cooper, 1932, табл. 18, фиг. 33; о-в Готланд.

В данном случае размеры раковин видов *Linoporella punctostriata* и *L. punctata* из биогерма Хиллисте, поэтому сравнение дает предполагаемые результаты.

Процедура при сравнении следующая. Секвенциальные критерии даны по признакам определяемой выборки (табл. 1, рис. 3). По x_1/x_5 известно, что $a_{1,5} = -0,7$ и

$b_{1,5} = 1,13$. Для *Linoporella punctostriata* соответствующие $b_{(1,5)}$ получим: для первого экземпляра

$$b_{(1,5)} = \frac{x_1 - (-0,7)}{x_5} = \frac{14,5 + 0,7}{16,3} = 0,93$$

и для второго

$$b_{(1,5)} = \frac{22,5 + 0,7}{25,0} = 0,93,$$

т. е. случайно тот же результат, что и для первого. После вычитания из каждого $b_{(1,5)}$ величину 1,13, т. е.

$$b_{(1,5)} - b_{1,5},$$

получим для обоих экземпляров одинаковые величины $-0,23$. Для нанесения этих данных на график суммируем их последовательно

$$-0,23 \text{ и } -0,23 + (-0,23) = -0,46.$$

По этим накопленным суммам ($\sum (b_{(1,5)} - b_{1,5})$) ломанная выходит уже после второго измерения из области неопределенности, и можно принять гипотезу отличия выборки из биогерма Хиллисте от сравниваемых экземпляров *Linoporella punctostriata* по данному признаку (рис. 3). Таким же образом (см. рис. 3) имеет *Linoporella punctata* из Англии меньшее значение $b_{1,5}$, чем 1,13. Данные по *Linoporella punctata* с о-ва Готланд являются малочисленными для заключения идентности их с выборкой *Linoporella* из биогерма Хиллисте по секвенциальным критериям.

По всем признакам (рис. 3) *Linoporella punctostriata* отличается от нашей выборки по $b_{1,5}$ и $b_{3,5}$. Данные по *Linoporella punctata* из Англии показывают различия при $b_{1,5}$ и $b_{2,5}$. Так как очертание раковины представителей рода *Linoporella* является наиболее существенным диагностическим признаком, возможно в дальнейшем придется выделить на основании материала из биогерма Хиллисте новый вид. По имеющимся данным определена названная выборка, однако, как *Linoporella* cf. *punctata*.

ЛИТЕРАТУРА

- В. В. Налимов, 1960. Применение математической статистики при анализе вещества. Гос. изд-во физ.-матем. лит., М.
- J. Hall, 1852. Palaeontology of New York, vol. II. Albany.
- J. Imbrie, 1956. Biometrical methods in the study of invertebrate fossils. Amer. Mus. Nat. Hist., Bull., vol. 108, No. 2.
- Ch. Schuchert and G. A. Cooper, 1932. Brachiopod genera of the suborders Orthoidea and Pentameroidea. Peabody Mus. Nat. Hist., Mem., vol. 4, pt. 1.
- E. Weber, 1961. Grundriss der biologischen Statistik für Naturwissenschaftler, Landwirte und Mediziner. VEB Gustav Fischer Verlag, Jena.
- W. F. Whittard and G. H. Barker, 1950. The upper valentian brachiopod fauna of Shropshire, Ann. Mag. Nat. Hist. XII ser., vol. 3, No. 31.

Институт геологии
Академии наук Эстонской ССР

Поступила в редакцию
8. V 1963

JÄRJENDANALÜÜSI KASUTAMISEST FOSSILSETE BRAHHIOPOODIDE MÄÄRAMISEL

M. Rubel

Resüme

Artiklis kirjeldatakse Hilliste biohermist (tamsalu lade, ländouveri, Lääne-Eesti) kogutud brahhiopoodi *Linoporella* cf. *punctata* väljavõtte alusel liigilise kuuluvuse määramist kahepoolse järjendanalüüsi graafilise võtte abil. Identifitseerimine teiste sama perekonna liikidega toimub populatsioonilisel tasemel üksikute tunnuste regressioonikoefitsientide järgi. Üksikeksemplari identifitseerimine statistiliselt iseloomustatud väljavõtte regressioonikoefitsiendiga $b_{i,j}$ toimub muutuja $b_{(i,j)}$ põhjal, kusjuures

$$b_{(i,j)} = \frac{x_i - a_{i,j}}{x_j},$$

milles x_i ja x_j on määratava eksemplari tunnused ning $a_{i,j}$ statistiliselt iseloomustatud väljavõtte samade tunnuste regressioonivõrrandi vabaliige. Näidatakse, et

$$\Sigma (b_{(i,j)} - b_{i,j}) = 0.$$

Eesti NSV Teaduste Akadeemia
Geoloogia Instituut

Saabus toimetusse
8. V 1963

ON THE APPLICATION OF SEQUENTIAL ANALYSIS IN IDENTIFYING FOSSIL BRACHIOPODS

M. Rubel

Summary

The author presents a method of applying two-sided sequential analysis by graphic mode for identifying a sample of brachiopod *Linoporella* cf. *punctata* collected from Hilliste bioherm (Tamsalu stage, Llandoverian, West-Estonia). Identification with other species of the same genus is carried out on populational level by regression coefficients of certain properties. The identification of a single individual of the sample characterized by regression coefficient $b_{i,j}$ is effected on the basis of the variable $b_{(i,j)}$, where

$$b_{(i,j)} = \frac{x_i - a_{i,j}}{x_j},$$

in which x_i and x_j are the properties of the individual to be identified and $a_{i,j}$ is the first member of the regression equation of the same properties of the statistically characterized sample. It is shown that

$$\Sigma (b_{(i,j)} - b_{i,j}) = 0.$$

Academy of Sciences of the Estonian S.S.R.,
Institute of Geology

Received
May 8th, 1963