

УДК 636.2.082.26:575.174.4

## ОСОБЕННОСТИ КОРРЕЛЯЦИОННОЙ СВЯЗИ МЕЖДУ УДОЕМ И ЖИРНОМОЛОЧНОСТЬЮ В СТАДАХ КРУПНОГО РОГАТОГО СКОТА

Г. П. АНТИПОВ

(Кафедра генетики и разведения с.-х. животных)

Предложена формула для интерпретации коэффициента корреляции как отношения разности между количеством общих факторов одинакового и противоположного действия к сумме всех факторов. Показана возможность ее использования для моделирования и содержательного описания структуры совместного распределения признаков удоя и жирномолочности в стадах крупного рогатого скота разного направления продуктивности, а также прогнозирования направления изменений корреляционной связи между признаками в процессе селекции.

В теории селекции важное место занимает корреляционный анализ связей как между разными признаками у особей той или иной популяции, так и между одними и теми же признаками у родственников. В первом случае коэффициент корреляции используется в качестве меры относительной изменчивости признаков, во втором — меры сходства между родственниками, как генотип, обуславливающий это сходство и одновременно проявляющийся в нем.

Корреляционный анализ применяется в практике селекции животных для оценки наследуемости и повторяемости признаков, связей между ними, препотентности производителей, прогноза коррелированного эффекта селекции и для

выбора ее методов. При этом существенное значение приобретают не только содержательная интерпретация получаемых оценок коррелятивной связи, но и прогнозирование последней при реализации тех или иных селекционных решений.

Однако до самого последнего времени именно прогноз изменений коррелятивных связей в процессе селекции весьма затруднителен из-за отсутствия в статистической трактовке коэффициента корреляции какого-либо принципа или указания на такую возможность.

В большинстве руководств по статистике, биометрии и популяционной генетике коэффициент корреляции (или его квадрат) трактуется как доля общей диспер-

сии того или иного признака ( $y$ ), обусловленная разнообразием другого признака ( $x$ ) при линейной форме интерпретации зависимости между ними:

$$r^2 = \frac{\sigma_{y/x(\text{линейная})}^2}{\sigma_y^2}.$$

Квадрат же корреляционного отношения, оценивающий степень связи между признаками при нелинейной их зависимости, выглядит как

$$\eta_{y/x}^2 = \frac{\sigma_{y/x(\text{нелинейная})}^2}{\sigma_y^2}.$$

Отсюда вытекают два принципиально важных следствия:

$$r^2 \leq \eta^2,$$

и если  $\eta_{y/x}^2 = \eta_{x/y}^2 = \eta^2$ , то связь линейная и, следовательно,  $\eta^2 = r^2$ .

Эти соотношения использованы Н. А. Плохинским [4] для проверки линейности связи. При достоверной разности ( $\eta^2 - r^2$ ) связь можно считать линейной с заданным уровнем надежности, при достоверной — нелинейной. Использование линейной модели в данном случае не запрещается, но ошибка будет тем больше, чем значительнее отклонение от линейности.

Разумеется, если взять какой-либо признак в качестве независимой переменной ( $x$ ) и оценить его влияние на разнообразие признака  $y$ , можно получить весьма существенную с практической точки зрения информацию, однако причины связи и формирование структуры совместного распределения признаков остаются за пределами возможностей анализа.

Существует еще одна трактовка коэффициента корреляции, согласно которой он «выражает отношение числа факторов, общих для

обеих случайных величин, ко всему числу факторов. Например, найдя, что коэффициент корреляции между объемным весом и модулем упругости при сжатии вдоль волокон древесины резонансовой ели равен приблизительно 0,9, мы можем допустить, что среди десяти некоторых факторов имеется девять, одинаковым образом влияющих на изменение объемного веса и модуля упругости» [2, с. 138].

Эта трактовка используется и в некоторых частных задачах популяционной генетики. Так, Ч. Ли приводит вывод формулы коэффициента корреляции для частного случая в виде:

$$r = \frac{C}{\sqrt{n_1 \cdot n_2}},$$

где  $n_1$  — сумма всех факторов, определяющих разнообразие 1-го признака;  $n_2$  — то же для 2-го признака;  $C$  — количество общих факторов для 1-го и 2-го признаков. Отметим далее, что если на оба признака действует одинаковое число факторов, т. е.  $n_1 = n_2 = n$ ,

$$r = \frac{C}{n}, \text{ Ч. Ли заключает: «В общем}$$

коэффициент корреляции представляет собой долю одинаково важных, но некоррелирующих факторов, которые проявляются в обоих выбранных признаках» [1, с. 274].

Однако и при такой интерпретации коэффициента корреляции возможности для содержательного анализа ситуаций остаются весьма ограниченными и совершенно не затрагивают вопроса о прогнозе изменения силы, направления или формы связи в процессе селекции. Так, установив путем статистической обработки эмпирических материалов факт изменения корреляций под действием селекционных

мероприятий, исследователи ограничиваются простой констатацией принципиальной возможности подобных изменений, поскольку имеющиеся интерпретации корреляции не позволяют найти способы, которые давали бы возможность прогнозировать направление ее изменений или их пределы.

Поэтому в зоотехнических селекционных исследованиях при обнаружении отрицательной зависимости между селекционируемыми признаками, существенно затрудняющей скорость и эффективность отбора на одновременное увеличение данных признаков, в качестве выхода из сложившейся ситуации ставится задача «преодоления» отрицательной связи. Причем под «преодолением» подразумевается перестройка связи, замена ее на положительную. Такая теоретическая рекомендация вполне естественно вытекает из рассмотренных общетеоретических представлений о формировании коррелятивных связей.

По крайней мере, в доступной литературе не удалось встретить более обоснованное положение, как: успех селекции на одновременное увеличение двух признаков, связанных отрицательной коррелятивной зависимостью, может сопровождаться не преодолением, а усилением отрицательной связи между ними.

В связи с изложенным целью нашей работы являются анализ особенностей формирования коррелятивных связей между признаками и разработка методов содержательной характеристики структуры их совместного распределения на примере важнейших селекционных признаков крупного рогатого скота — уоя и жирномолочности — в стадах разного уровня и направления продуктив-

ности и в различных селекционных ситуациях.

Анализ особенностей формирования коррелятивных связей и их изменений под действием некоторых форм отбора и систем спаривания проводили путем построения теоретических моделей и последующей их проверки на соответствие фактическим распределениям признаков в реальных стадах. Соответствие фактических распределений теоретически ожидаемым устанавливали методом минимизации значения критерия  $\chi^2$  [4]. Использовали фактические распределения (корреляционные решетки) уоя и жирномолочности по данным сводных бонитировочных ведомостей за ряд лет и других материалов племенного учета в стадах крупного рогатого скота симментальской породы учхоза ТСХА «Муммовское» Саратовской области, черно-пестрой породы учхоза «Михайловское» и совхоза «Истринский» Московской области.

Модель формирования коррелятивных связей между признаками в результате действия факторов, влияющих на разнообразие признаков, составлена с учетом следующих положений:

- 1) факторы независимы;
- 2) каждый фактор представлен двумя градациями — плюс и минус (генетические основания такого подразделения — диплоидность организмов и менделевские законы сегрегации и агрегации генов);
- 3) вероятность плюс-градации  $j$ -того фактора —  $P_j$ , минус-градации —  $q_j$ ;
- 4) вклад каждой градации соответствующего фактора в фенотипическое значение признака для плюс-градации —  $a_j$ , для минус-градации — 0;
- 5) вклады аддитивны.

$$\text{Тогда } \bar{x} = \sum_{j=1}^n a_j P_j; \quad \sigma_x^2 = \\ = \sum_{j=1}^n a_j^2 P_j q_j, \quad j=1, 2, \dots, n, \quad \text{где } n —$$

количество факторов, действующих на разнообразие признака.

Известно, что при равенстве всех вкладов плюс-градаций факторов двум и минус-градаций — нулю (что эквивалентно различию между вкладами в размере двух единиц измерения, чаще всего описываемому после некоторых преобразований выражением вклада плюс-градации в виде  $+1$  и минус-градации в виде  $-1$ ) и равенстве вероятностей  $P_j = 0,5$  среднее значение признака ( $\bar{x}$ ) в этом частном случае биномиального распределения становится равным вариансу ( $\sigma_x^2$ ), а обе величины — числу действующих факторов ( $n$ ), т. е.  $\bar{x} = \sigma_x^2 = n$ .

Обычно считают, что одна часть факторов является общими, одинаково действующими на оба признака, другая — специфическими, действующими на разнообразие только одного признака, и исходя из этого определяют коэффициент корреляции как отношение количества общих факторов к сумме всех факторов (при равенстве количества факторов, действующих на оба признака).

Согласно нашему представлению, которое отличается от данного подхода (точнее развивает его), общие факторы могут действовать не только в одинаковом направлении, но и в противоположном. При одинаковом действии плюс-градация фактора (с вероятностью  $P$ ) увеличивает значение обоих признаков, а минус-градация — соответственно уменьшает (или не изменяет) их значение. При противоположном дей-

ствии плюс-градация фактора увеличивает значение одного признака и уменьшает (или не меняет) значение другого, а минус-градация уменьшает значение одного из них и увеличивает значение другого. Иными словами, конкретная градация какого-либо фактора для одного признака будет плюс-градацией, а для другого — минус-градацией. Тогда коэффициент корреляции между признаками  $x$  и  $y$  может быть описан следующей формулой:

$$r_{x/y} = \frac{n_i - n_o}{\sqrt{n_x n_y}}, \quad (1)$$

где  $n_i$  — количество общих факторов одинакового действия на признаки  $x$  и  $y$ ;  $n_o$  — количество общих факторов противоположного действия на  $x$  и  $y$ ;  $n_x$ ,  $n_y$  — количество всех факторов, формирующих разнообразие признаков, — соответственно  $x$  и  $y$ .

Введем следующие обозначения:  $n_c$  — количество общих факторов, действующих на разнообразие признаков  $x$  и  $y$ ;

$$n_c = n_i + n_o, \quad (2)$$

$n_{sx}$  — количество специфических (добавочных) факторов, действующих только на разнообразие признака  $x$  и не действующих на признак  $y$ ;  $n_{sy}$  — количество специфических (добавочных) факторов, действующих только на  $y$ . Тогда

$$n_x = n_i + n_o + n_{sx} = n_c + n_{sx}, \quad (3)$$

$$n_y = n_i + n_o + n_{sy} = n_c + n_{sy}. \quad (4)$$

Если  $n_{sx} = n_{sy} = n_s$ , то

$$n_x = n_y = (n_c + n_s) = n, \quad (5)$$

и коэффициент корреляции между признаками  $x$  и  $y$  можно выразить в виде

$$r_{xy} = \frac{n_i - n_o}{n}. \quad (6)$$

Если принять, что  $n^{(G)}$  и  $n^{(π)}$  — количество соответственно генетических (генов, их блоков или других наследуемых единиц) и паратипических факторов, действующих на разнообразие признаков, тогда коэффициент генетической корреляции между признаками можно выразить

$$r_G = \frac{n_i^{(G)} - n_o^{(G)}}{\sqrt{n_x^G n_y^G}}, \quad (7)$$

паратипической —

$$r_\pi = \frac{n_i^{(\pi)} - n_o^{(\pi)}}{\sqrt{n_x^{(\pi)} n_y^{(\pi)}}}, \quad (8)$$

фенотипической —

$$r_\varphi = \frac{(n_i^{(G)} - n_o^{(G)}) + (n_i^{(\pi)} - n_o^{(\pi)})}{\sqrt{n_x n_y}}, \quad (9)$$

где  $n_x = n_i^G + n_o^G + n_i^\pi + n_o^\pi + n_{sx}^G + n_{sx}^\pi$ ,  $n_y = n_i^G + n_o^G + n_i^\pi + n_o^\pi + n_{sy}^G + n_{sy}^\pi$ .

Полученные выражения (1) и (6) позволяют охарактеризовать внутреннюю содержательную структуру корреляции, вскрыть особенности формирования коррелятивной зависимости между признаками и дать некоторые прогнозы относительно возможных изменений корреляций в процессе селекции.

Из формул (1) и (6) следует, что селекция на одновременное увеличение признаков  $x$  и  $y$  возможна прежде всего при переводе в плюс-состояние (т. е. в гомозиготное) факторов, одинаково влияющих на эти признаки. Но перевод факторов в одну из градаций (в нашем случае в плюс-градацию) есть исключение источника разнообразия. Таким образом,

уменьшение количества одинаково действующих (генетических) факторов или исключение их приведет либо к уменьшению положительной корреляции между признаками, либо к проявлению факторов, действующих противоположно, т. е. к усилению отрицательной связи между признаками.

С этой точки зрения находят свое объяснение факты усиления отрицательной (фенотипической и генетической) корреляции при одновременном повышении удоя и жирномолочности в отдельных стадах и породах крупного рогатого скота. Так, по нашим и литературным данным [3], при повышении удоев в стаде ярославского скота племхоза «Горшиха» до 4,5—5 тыс. кг на корову и содержания жира до 4,5—4,6% значительно возросла отрицательная корреляция между этими признаками — до  $-0,4...-0,5$ . Столь же высокие значения отрицательной корреляции между удоем и жирномолочностью характерны для высокопродуктивных стад гернзейской и джерсейской пород.

Проведенное нами моделирование отбора на увеличение значений двух признаков (тандемного, по независимым уровням и индексу) показало, что при наличии факторов, противоположно действующих на признаки, в процессе селекции на одновременное увеличение последних происходит отбор в пользу гетерозигот. Это приводит к установлению в популяции стационарного сбалансированного состояния с соответствующими частотами генов по локусам с противоположным действием на признаки и объясняет довольно значительную гетерозиготность высокопродуктивных племенных стад крупного рогатого скота по ряду хозяйственно цен-

ных признаков, в частности по удою и жирномолочности.

Формулы также позволяют моделировать совместные (двумерные, а в общем случае —  $n$ -мерные) распределения признаков и оценивать изменения в структуре их совместного разнообразия под действием селекционных мероприятий в конкретных стадах. Нами разработан соответствующий алгоритм. Следует подчеркнуть, что задача описания совместных распределений (построение корреляционных решеток) удою и жирномолочности гораздо легче решается для коров-первотелок и труднее — для взрослых животных с разными лактациями вследствие значительного увеличения числа факторов, оказывающих система-

Таблица 1

Фактическое и теоретическое распределение коров-первотелок черно-пестрой породы по удою и жирномолочности (1986 г., учхоз «Михайловское»,  $n_i=0$ ,  $n_o=1$ ,  $n_{sx}=3$ ,  $n_{sy}=2$ ,  $P_o=P_{sx}=0,5$ ,  $P_{sy}=0,4$ )

Содержание жира (y), %	Удой за лактацию (x), кг					Частоты (y)
	2000	3000	4000	5000	6000	
4,4	4	13	15	3	—	35
	4	12	13	4		33
4,0	9	33	56	30	1	129
	12	41	49	25	4	131
3,6	8	45	69	53	9	184
	9	40	64	46	12	171
3,2		14	19	21	5	59
		9	27	27	9	72
Частоты (x)	21	105	159	107	15	N=407
	25	102	153	102	25	

Примечание. Здесь и в табл. 2 числитель — фактические частоты, знаменатель — теоретически ожидаемые.  $\chi^2_{факт. (x)} = 5,3$ ;  $\chi^2_{табл.} = 9,5$ ; 13,3; 18,5 ( $\nu=4$ );  $\chi^2_{факт. (y)} = 3,7$ ;  $\chi^2_{табл.} = 7,8$ ; 11,3; 16,3 ( $\nu=3$ );  $\chi^2_{факт. (y/x)} = 20,2$ ;  $\chi^2_{табл.} = 21,0$ ; 26,2; 31,9 ( $\nu=12$ ).

Таблица 2

Фактическое и теоретическое распределение коров-первотелок симментальной породы по удою и жирномолочности (1988 г., учхоз «Муммовское»,  $n_i=1$ ,  $n_o=2$ ,  $n_{sx}=0$ ,  $n_{sy}=2$ ,  $P_i=0,5$ )

Содержание жира (y), %	Удой за лактацию (x), кг				Частоты (y)
	1500	2500	3500	4500	
4,4	—	2	—	—	2
		3			3
4,2	0	11	6	—	17
	3	7	6		16
4,0	1	10	14	1	26
	6	10	13	3	32
3,8	4	13	10	6	33
	3	13	10	6	32
3,6		7	11	4	22
		7	7	3	17
3,4		2			2
		3			3
Частоты (x)	5	43	43	11	N=103
	12	40	39	13	

Примечание.  $\chi^2_{факт. (x)} = 5,2$ ;  $\chi^2_{табл.} = 7,8$ ; 11,3; 16,3 ( $\nu=3$ );  $\chi^2_{факт. (y)} = 4,4$ ;  $\chi^2_{табл.} = 11,1$ ; 15,1; 20,5 ( $\nu=5$ );  $\chi^2_{факт. (y/x)} = 14,3$ ;  $\chi^2_{табл.} = 25,0$ ; 30,6; 37,7 ( $\nu=15$ ).

тическое и случайное («шумовое») влияние и тем самым искажающих значения признаков. Это приводит к тому, что в бонитировочных ведомостях таблицы зависимости между удою и жирномолочностью для коров по последней законченной лактации в большинстве случаев представляют собой композицию нескольких распределений.

Примеры описания структуры разнообразия и особенностей формирования коррелятивных связей между удою и жирномолочностью у коров-первотелок черно-пестрой породы учхоза «Михайловское» и симментальской породы учхоза «Муммовское» представлены в табл. 1 и 2.

Из данных табл. 1 и 2 видно, что фактические распределения признаков как отдельно по удою и жирномолочности, так и совместно с очень высокой надежностью (во всех случаях  $\chi^2_{\text{факт.}} < \chi^2$ ) соответствуют теоретически ожидаемым. Существенный интерес в связи с этим представляет сравнительная оценка изменений в структуре связей между признаками в процессе селекции при чистопородном разведении и скрещивании (табл. 3). В стаде учхоза «Михайловское» за 5-летие средние значения удоя и жирномолочности существенно не изменились, но при этом снизилось количество факторов разнообразия такого признака, как содержание жира, что можно объяснить фактическим исчерпанием источников разнообразия общих одинаково действующих на удою и жирномолочность факторов ( $n_i=0$ ). В то же время при скрещивании черно-пестрых коров с голштино-фризскими бы-

ками заметно изменяется структура связи. Хотя общее количество факторов не изменилось, у помесей 1-го поколения значительно увеличился удою и прежде всего за счет гомозиготности производителей по блоку генов, положительно влияющих на удою и жирномолочность ( $n_i=0$ ), а также за счет повышения генетического разнообразия по специфически действующим факторам.

Предлагаемый метод позволяет оценивать и структурные изменения в потомстве отдельных производителей. Особого внимания заслуживает тот факт, что во всех случаях как при чистопородном разведении, так и при скрещивании наиболее стабильным оказывается количество общих факторов разнообразия с противоположным действием на удою и жирномолочность ( $n_0$ ). Следовательно, отбор на повышение (или удержание на высоком уровне)

Таблица 3

Изменение структуры коррелятивной связи между удою (числитель — кг) и жирномолочностью (знаменатель — %) у коров черно-пестрой породы в процессе селекции

Показатель	M	n	$n_1$	$n_0$	$n_2$	r (удой — содержание жира)
<i>Учхоз «Михайловское»</i>						
Чистопородные:						
1981 г. (N=329)	$\frac{3953}{3,74}$	$\frac{4}{6}$	0	1	$\frac{3}{5}$	-0,20
1986 г. (N=408)	$\frac{3949}{3,72}$	$\frac{4}{3}$	0	1	$\frac{3}{2}$	-0,28
<i>Совхоз «Истринский» и др.</i>						
Чистопородные (N=320)	$\frac{4500}{3,80}$	$\frac{5}{7}$	1	2	$\frac{2}{4}$	-0,29
Помеси F <sub>1</sub> от голштино-фризских быков (N=349)	$\frac{5200}{3,70}$	$\frac{5}{7}$	0	2	$\frac{3}{5}$	-0,33
Дочери быка Грея (3/4 по голштино-фризской породе, N=112)	$\frac{3700}{3,50}$	$\frac{4}{7}$	2	2	$\frac{0}{3}$	0,0

двух признаков одновременно приводит к исчерпанию факторов разнообразия, одинаково действующих на оба признака, и сохранению факторов разнообразия противоположного действия, т. е. сохранению и поддержанию соответствующего уровня гетерозиготности, поскольку выключение из разнообразия этих факторов (перевод в гомозиготное состояние) отрицательно сказалось бы на одном из признаков.

### Выводы

1. Предложена формула для интерпретации коэффициента корреляции как отношения разности между количеством общих факторов одинакового и противоположного действия к сумме всех факторов, формирующих разнообразие признаков.

2. Выявлена возможность использования формулы для моделирования и описания особенностей совместного распределения важных хозяйственно полезных признаков у крупного рогатого скота —

удоя и содержания жира в молоке.

3. Показана возможность сравнительной оценки изменений структуры связи между удоем и жирномолочностью в процессе селекции.

4. При наличии генетических факторов (генов или блоков генов) с противоположным действием, определяющих разнообразие признаков, отбор на повышение значений двух признаков одновременно является отбором в пользу гетерозигот. Успеху селекции в этом случае будет сопутствовать усиление отрицательной коррелятивной связи между селекционируемыми признаками.

### ЛИТЕРАТУРА

1. Ли Ч. Введение в популяционную генетику.— М.: Мир, 1978.— 2. Митропольский А. К. Техника статистических вычислений.— М.: Наука, 1971.— 3. Никоро Э. С., Стакан Г. А., Харитонова Э. Н. и др. Теоретические основы селекции животных.— М.: Колос, 1968.— 4. Плохинский Н. А. Алгоритмы биометрии (2-е изд.).— М.: Изд-во МГУ, 1980.

Статья поступила 28 января 1993 г.

### SUMMARY

The new formula for correlation coefficient interpretation is proposed.

The ratio of difference between the number of general factors with similar and opposite effect to the sum of all influencing factors together is corresponded with correlation coefficient. This principle is suitable for modeling and substantial description of correlative distribution of milk yield and fat in different dairy cattle herds and for prognosing correlation changes between milk productive traits in process of selection.